

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent of: Satoru Sakazume
U.S. Patent No.: 9,042,448 Attorney Docket No. 50095-0244IP1
Issue Date: May 26, 2015
Appl. Serial No.: 12/995,039
Filing Date: May 28, 2009
Title: 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, AND
MOVING PICTURE REENCODING PROGRAM

Mail Stop Patent Board

Patent Trial and Appeal Board
U.S. Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450

**PETITION FOR *INTER PARTES* REVIEW OF UNITED STATES PATENT
NO. 9,042,448 PURSUANT TO 35 U.S.C. §§ 311–319, 37 C.F.R. § 42**

TABLE OF CONTENTS

I.	REQUIREMENTS FOR IPR UNDER 37 C.F.R. § 42.104.....	1
A.	Grounds for Standing Under 37 C.F.R. § 42.104(a).....	1
B.	Challenge Under 37 C.F.R. § 42.104(b) and Relief Requested	1
C.	Claim Construction under 37 C.F.R. §§ 42.104(b)(3).....	3
D.	Level of Ordinary Skill in the Art.....	4
II.	SUMMARY OF THE '448 PATENT	4
A.	Technical Background	4
B.	Brief Description.....	6
C.	Examination History	11
III.	THE CHALLENGED CLAIMS ARE UNPATENTABLE.....	13
A.	[GROUND 1A] – Phek in view of Segall and He Renders Claims 1-3 Obvious	13
1.	Phek Overview	13
2.	Segall Overview	14
3.	He Overview.....	17
4.	Combination of Phek, Segall, and He	19
5.	Analysis	44
B.	[GROUND 1B] – Phek in view of Segall, He, and Takahashi Renders Claims 1-3 Obvious	69
1.	Takahashi Overview	69
2.	Combination of Phek, Segall, He, and Takahashi.....	72
3.	Analysis	75
C.	[GROUND 1C] – Phek in view of Segall and Takahashi Renders Claims 1-3 Obvious.....	80
D.	[GROUND 1D] – Phek in view of Takahashi Renders Claim 3 Obvious	81
IV.	DISCRETIONARY DENIAL IS NOT WARRANTED.....	82
V.	PAYMENT OF FEES – 37 C.F.R. § 42.103.....	82
VI.	CONCLUSION.....	82
VII.	MANDATORY NOTICES UNDER 37 C.F.R § 42.8(a)(1).....	82
A.	Real Party-In-Interest Under 37 C.F.R. § 42.8(b)(1).....	82
B.	Related Matters Under 37 C.F.R. § 42.8(b)(2).....	82
C.	Lead And Back-Up Counsel Under 37 C.F.R. § 42.8(b)(3).....	83
D.	Service Information	83

LIST OF EXHIBITS

APPLE-1001	U.S. Patent No. 9,042,448 to Sakazume (“the ’448 Patent”)
APPLE-1002	Excerpts from the Prosecution History of the ’448 Patent (“the ’448 Patent File History”)
APPLE-1003	Declaration and Curriculum Vitae of David Lett
APPLE-1004	Complaint, <i>Advanced Coding Technologies LLC v. Apple Inc.</i> , 2-24-cv-00572-JRG, E.D. Tex., filed Jul. 22, 2024
APPLE-1005	English Translation of Japanese Patent Publication No. 2007316161 A with Translation Certificate (“Phek”)
APPLE-1006	Segall et al., <i>Spatial Scalability Within the H.264/AVC Scalable Video Coding Extension</i> ; IEEE Transactions on Circuits & Systems for Video Tech., Vol. 17, No. 9, September 2007 (“Segall”)
APPLE-1007	Martins et al., <i>A Unified Approach to Restoration, Deinterlacing and Resolution Enhancement in Decoding MPEG-2 Video</i> , IEEE Transactions on Circuits & Systems for Video Tech., Vol. 12, No. 9, September 2002 (“Martins”)
APPLE-1008	U.S. Patent Publication No. 2008/0137753 (“He”)
APPLE-1009	U.S. Patent No. 5,886,736 (“Chen”)
APPLE-1010	Schwarz et al., <i>Overview of the Scalable Video Coding Extension of the H.264/AVC Standard</i> , IEEE Transactions on Circuits & Systems for Video Tech., Vol. 17, No. 9, September 2007 (“Schwarz”)
APPLE-1011	U.S. Patent Publication No. 2009/0154567 (“Lei”)
APPLE-1012	Park et al., <i>Super-Resolution Image Reconstruction: A Technical Overview</i> , IEEE Signal Processing Magazine, pp. 21-36, May 2003 (“Park”)

APPLE-1013	English Machine Translation of Japanese Patent Publication No. 2008053848 (“Hatanaka”)
APPLE-1014	U.S. Patent Publication No. 2010/0165077 (“Yin”)
APPLE-1015	U.S. Patent Publication No. 2009/0257664 (“Kao”)
APPLE-1016	U.S. Patent Publication No. 2009/0046995 (“K anumuri”)
APPLE-1017	U.S. Pat. 6,470,051 (“Campisano”)
APPLE-1018	U.S. Patent Publication No. 2006/0013306 (“Kim”)
APPLE-1019	Lu et al., Mechanisms of MPEG Stream Synchronization, ACM SIGCOMM Computer Communication Review, Vol. 24, Issue 1, pp. 57-67 (January 1994) (“Lu”)
APPLE-1020	U.S. Patent Publication No. 2003/0021345 (“Brusewitz”)
APPLE-1021	U.S. Patent Publication No. 2003/0021347 (“Lan”)
APPLE-1022	EE Times, <i>How Video Compression Works</i> (Aug. 6, 2007), available at https://www.eetimes.com/how-video-compression-works (retrieved Dec. 26, 2023) (“EETimes”)
APPLE-1023	Bier, <i>Introduction to Video Compression</i> , Berkeley Design Tech., Inc. (October 2005), available at https://www.bdti.com/MyBDTI/pubs/20051024_GSPx05_Video_Intro.pdf (retrieved Dec. 30, 2023) (“Bier”)
APPLE-1024	U.S. Patent No. 7,379,496 (“Holcomb”)
APPLE-1025	U.S. Patent Publication No. 2008/0043832 (“Barkley”)
APPLE-1026	U.S. Patent No. 6,173,013 (“Suzuki”)
APPLE-1027	Memorandum, Interim Procedure for Discretionary Denials in AIA Post-Grant Proceedings with Parallel District Court Litigation (USPTO June 21, 2022) (“Director’s Guidance”)
APPLE-1028	[RESERVED]
APPLE-1029	Declaration of June Ann Munford

APPLE-1030	Ely, <i>MPEG Video Coding: A Simple Introduction</i> , EBU Technical Review (Winter 1995) (“Ely”)
APPLE-1031	Japanese Patent Publication No. 2007316161 A with English Machine Translation
APPLE-1032	ITU-T Recommendation H.264 (11/07)
APPLE-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)
APPLE-1034	ISO/IEC 11172-2: 1993 (E) (MPEG1 Video)
APPLE-1035	ISO/IEC 13818-2: 2000 (E) (MPEG2/H.262 Video)
APPLE-1036	ITU-T Recommendation H.263 (02/98)
APPLE-1037	Reader, Intraframe and Interframe Adaptive Transform Coding, Efficient Transmission of Pictorial Information, SPIE v.66, 108 (1975)
APPLE-1038	Reader, MPEG4: Coding for Content, Interactivity, and Universal Accessibility, Optical Engineering 35(1), 104 (January 1996)
APPLE-1039	Reader, MPEG Patents, in MPEG Video Compression Standard, Chapter 16, at 357-362 (Mitchell et al ed.) (1996)
APPLE-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)
APPLE-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)
APPLE-1042	ISO/IEC 14496-2: 2004 (MPEG4)
APPLE-1043	ISO/IEC 13818-2: 1996 (MPEG2)

APPLE-1044	Petition (Paper 2), IPR2024-00372, <i>Samsung Electronics. Co., Ltd. v. Advanced Coding Techs., LLC</i> (Jan. 8, 2024)
APPLE-1045	Decision Instituting <i>Inter Partes</i> Review (Paper 9), IPR2024-00372, <i>Samsung Electronics. Co., Ltd. v. Advanced Coding Techs., LLC</i> (July 18, 2024)
APPLE-1046	U.S. Patent Publication No. 2001/0043751 (“Takahashi”)
APPLE-1047	Peter J. Burt et al., <i>The Laplacian Pyramid as a Compact Image Code</i> , IEEE TRANSACTIONS ON COMMUNICATIONS, VOL. COM-31, NO. 4, APRIL 1983 (“Burt”)
APPLE-1048	Michael Elad et al., <i>Restoration of a Single Superresolution Image from Several Blurred, Noisy, and Undersampled Measured Images</i> , IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 6, NO. 12, DECEMBER 1997 (“Elad”)
APPLE-1049	Michal Irani et al., <i>Super resolution from image sequences</i> , [1990] Proceedings. 10th International Conference on Pattern Recognition, Atlantic City, NJ, USA, 1990, pp. 115-120 vol.2, doi: 10.1109/ICPR.1990.119340 (“Irani”)
APPLE-1050	U.S. Patent Publication No. 2007/0160133 (“Bao”)
APPLE-1051	U.S. Patent Publication No. 2004/0170330 (“Fogg”)
APPLE-1052	U.S. Patent Publication No. 2006/0126952 (“Suzuki”)
APPLE-1053	U.S. Patent Publication No. 2004/0022318 (“Garrido318”)
APPLE-1054	U.S. Patent Publication No. 2005/0019000 (“Lim”)
APPLE-1055	U.S. Patent Publication No. 2004/0156561 (“YuChuan”)
APPLE-1056	U.S. Patent Publication No. 2008/0018506 (“Raveendran”)
APPLE-1057	Hunt B R, <i>Super-Resolution of Images- Algorithms, Principles, Performance</i> , Intl. Jnl. of Imaging Sys. and Technol., Vol. 6, pp. 297-304, Wiley, 1995, Oct. 1995 (“Hunt_Oct_1995”)

APPLE-1058	Kramer P, O. Hadar, Benois-Pineau J, Domenger J-P, <i>Use of Motion Information in Super-Resolution Mosaicing</i> , ICIP, Oct. 2006 (“Kramer”)
APPLE-1059	Patti A J, Sezan M I, Tekalp M, <i>Superresolution video reconstruction with arbitrary sampling lattices and nonzero aperture time</i> , IEEE Trans. on Im. Proc., Vol. 6, No. 8, pp. 10 (“Patti”)
APPLE-1060	Erickson K J, Schultz R R, <i>MPEG-1 super-resolution decoding for the analysis of video still images</i> , 4th IEEE Southwest Symp. on Im. Anal. and Interp., Apr. 2000 (“Erickson”)
APPLE-1061	Tsai R Y, Huang T S, <i>Multiframe image restoration and registration</i> , Advances in Computer Vision and Image Processing, Vol. 1, T. S. Huang, ed., Jai Press, pp. 317-319, 1984 (“Tsai”)
APPLE-1062	WIPO Patent Publication No. WO2008/057308A2 (“Kao308”)
APPLE-1063	U.S. Patent No. 5,606,369 (“Keesman”)
APPLE-1064	U.S. Patent Publication No. 2007/0201560 (“Segall560”)
APPLE-1065	U.S. Patent Publication No. 2004/0022518 (“Taga”)
APPLE-1066	U.S. Patent No. 6,510,177 (“DeBonet”)
APPLE-1067	U.S. Patent No. 6,023,535 (“Aoki”)
APPLE-1068	U.S. Patent Publication No. 2007/0160153 (“Sullivan”)
APPLE-1069	U.S. Patent Publication No. 2006/0222079 (“Park079”)
APPLE-1070	Exhibit SAMSUNG-1006 as submitted in IPR2024-00327
APPLE-1071	Exhibit SAMSUNG-1007 as submitted in IPR2024-00327
APPLE-1072	Exhibit SAMSUNG-1010 as submitted in IPR2024-00327
APPLE-1073	JP2009182776A and English Machine Translation (“Kuroda”)
APPLE-1074	U.S. Patent No. 8,023,561 (“Garrido561”)

APPLE-1075	U.S. Patent No. 6,580,754 (“Wan”)
APPLE-1076	U.S. Patent No. 5,408,270 (“Lim270”)
APPLE-1077	EP 1811786 (“Wang”)
APPLE-1078	Declaration of Dr. Clifford Reader from IPR2024-00372

LISTING OF CLAIMS

Claim 1	
[1.P]	A moving picture encoding system that makes an encoding of a sequence of moving pictures with a resolution higher than a standard resolution using moving pictures contents which include a sequence of moving pictures with the standard resolution and do not include a sequence of moving pictures with a resolution higher than the standard resolution, the moving picture encoding system comprising:
[1.1]	a first encoder configured to work on 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 sequence of moving pictures with the standard resolution to implement a process for a first super-resolution enlargement to create a set of super-resolution enlarged pictures with a resolution higher than the standard resolution;
[1.3]	a second super-resolution enlarger configured to acquire the set of decoded pictures from the first encoder to implement thereon a process for a second super-resolution enlargement to create a set of super-resolution enlarged decoded pictures with a resolution higher than the standard resolution;
[1.4]	a third resolution converter configured to acquire the set of decoded pictures from the first encoder to implement thereon a process for a third resolution conversion to create a set of resolution converted enlarged decoded pictures with a resolution higher than the standard resolution; and
[1.5]	a third encoder configured to have the set of super-resolution enlarged pictures from the first super-resolution enlarger as a set of encoding target pictures, employing the set of super-resolution enlarged decoded pictures from the second super-resolution

	enlarger and the set of resolution converted enlarged decoded pictures from the third resolution converter as sets of reference pictures, to implement thereon a third combination of processes for a prediction and an encoding to create a third sequence of encoded bits,
[1.6]	wherein a spatial resolution of the set of super-resolution enlarged pictures, that of the set of super-resolution enlarged decoded pictures, and that of the set of resolution converted enlarged decoded pictures are made equal; and
[1.7]	wherein the third encoder controls a selection of a set of reference pictures and creates a set of data on the selection of the set of reference pictures to identify a selected set of reference pictures during the process for prediction of the third combination.
Claim 2	
[2.P]	A moving picture encoding system that makes an encoding of a sequence of moving pictures with a resolution higher than a standard resolution using moving pictures contents which include a sequence of moving pictures with the standard resolution and do not include a sequence of moving pictures with a resolution higher than the standard resolution, the moving picture encoding system comprising:
[2.1]	a first encoder configured to work on 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;
[2.2]	a first super-resolution enlarger configured to work on the sequence of moving pictures with the standard resolution to implement a process for a first super-resolution enlargement to create a set of super-resolution enlarged pictures with a resolution higher than the standard resolution;
[2.3]	a third resolution converter configured to acquire the set of decoded pictures from the first encoder to implement thereon a

	process for a third resolution conversion to create a set of resolution converted enlarged decoded pictures with a resolution higher than the standard resolution; and
[2.4]	a third encoder configured to have the set of super-resolution enlarged pictures from the first super-resolution enlarger as a set of encoding target pictures, employing the set of resolution converted enlarged decoded pictures from the third resolution converter as a set of reference pictures, to implement thereon a third combination of processes for a prediction and an encoding to create a third sequence of encoded bits,
[2.5]	wherein a spatial resolution of the set of super-resolution enlarged pictures, and that of the set of resolution converted enlarged decoded pictures are made equal; and
[2.6]	wherein the third encoder controls a selection of a set of reference pictures and creates a set of data on the selection of the set of reference pictures to identify a selected set of reference pictures during the process for prediction of the third combination.
Claim 3	
[3.P]	A moving picture encoding system that makes an encoding of a sequence of moving pictures with a resolution higher than a standard resolution using moving pictures contents which include a sequence of moving pictures with the standard resolution and do not include a sequence of moving pictures with a resolution higher than the standard resolution, the moving picture encoding system comprising:
[3.1]	a first encoder configured to work on 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;
[3.2]	a first super-resolution enlarger configured to work on the sequence of moving pictures with the standard resolution to implement a process for a first super-resolution enlargement to

	create a set of super-resolution enlarged pictures with a resolution higher than the standard resolution;
[3.3]	a second super-resolution enlarger configured to acquire the set of decoded pictures from the first encoder to implement thereon a process for a second super-resolution enlargement to create a set of super-resolution enlarged decoded pictures with a resolution higher than the standard resolution; and
[3.4]	a third encoder configured to have the set of super-resolution enlarged pictures from the first super-resolution enlarger as a set of encoding target pictures, employing the set of super-resolution enlarged decoded pictures from the second super-resolution enlarger as a set of reference pictures, to implement thereon a third combination of processes for a prediction and an encoding to create a third sequence of encoded bits,
[3.5]	wherein a spatial resolution of the set of super-resolution enlarged pictures, and that of the set of super-resolution enlarged decoded pictures are made equal; and
[3.6]	wherein the third encoder controls a selection of a set of reference pictures and creates a set of data on the selection of the set of reference pictures to identify a selected set of reference pictures during the process for prediction of the third combination.

Apple Inc. (“Apple” or “Petitioner”) petitions for *Inter Partes* Review (“IPR”) of claims 1-3 (“the Challenged Claims”) of U.S. Patent No. 9,042,448 (“the ’448 patent”).

I. REQUIREMENTS FOR IPR UNDER 37 C.F.R. § 42.104

A. Grounds for Standing Under 37 C.F.R. § 42.104(a)

Petitioner certifies that the ’448 Patent is available for IPR. This Petition is being filed within one year of service of a complaint against Apple. Apple is not barred or estopped from requesting review of the Challenged Claims on the grounds identified below.

B. Challenge Under 37 C.F.R. § 42.104(b) and Relief Requested

Apple requests institution of IPR and cancellation of the Challenged Claims based on the following grounds:

Ground	Claims	35 U.S.C. § 103
1A	1-3	Phek in view of Segall and He
1B	1-3	Phek in view of Segall, He, and Takahashi
1C	1-3	Phek in view of Segall and Takahashi
1D	3	Phek in view of Takahashi

Grounds 1A-1D are further supported by the expert testimony of Mr. David B. Lett (APPLE-1003) and additional corroborating evidence cited throughout the

Petition.

The '448 Patent claims earliest priority to Japanese Application No. 2008-142433, filed on May 30, 2008. APPLE-1001, Cover. Petitioner does not concede that the Challenged Claims are entitled to the benefit of the filing date of any earlier-filed application in the alleged priority chain, but this Petition nonetheless treats May 30, 2008 as the Critical Date of the Challenged Claims for purposes of this IPR. Each reference that forms the basis of Grounds 1A-1D qualified as prior art even as early as May 30, 2008:

Reference	U.S. Filing Date	Publication Date	Pre-AIA Prior Art
Phek	---	12/06/2007	§102(a)
Segall		09/2007	§102(a)
He	12/08/2006	06/12/2008	§102(e)
Takahashi	3/16/1998	11/22/2001	§102(a)-(b)

Petitioner also submits evidence demonstrating that Segall was published by IEEE well before the Critical Date of the '448 Patent. *See* IPR2014-00527, Paper 41, 11 (PTAB May 18, 2015) (“IEEE is a well-known, reputable compiler and publisher of scientific and technical publications.”). Ms. June Munford, an expert librarian, reviewed MARC records and a physical copy of an IEEE journal in which Segall was published, and Ms. Munford explains how this evidence confirms that Segall was available to the interested public before the Critical Date.

APPLE-1029, ¶¶6-9, 14-18; *generally id.*, ¶¶1-19; *see also* APPLE-1070, -1071, -1072 (corresponding exhibits reviewed by Ms. Munford).

C. Claim Construction under 37 C.F.R. §§ 42.104(b)(3)

All claim terms should be construed according to the *Phillips* standard.

Phillips v. AWH Corp., 415 F.3d 1303 (Fed. Cir. 2005); 37 C.F.R. §42.100.

Petitioner submits that no formal constructions are presently necessary for purposes of demonstrating obviousness of the Challenged Claims based on Grounds 1A-1D. *Wellman, Inc. v. Eastman Chem. Co.*, 642 F.3d 1355, 1361 (Fed. Cir. 2011) (“claim terms need only be construed to the extent necessary to resolve the controversy”) (internal quotations omitted).

Petitioner reserves the right to respond to any constructions offered by Patent Owner or adopted by the Board. Petitioner is not conceding that the Challenged Claims satisfy all statutory requirements, nor is Petitioner waiving any arguments concerning indefiniteness or claim scope that can only be raised in district court or otherwise outside the context of an IPR. IPR2018-00272, Paper 35, 6-8 (PTAB 2019). The Board has regularly compared indefinite claims to the prior art for purposes of considering unpatentability. *See, e.g.*, CBM2013-00027, Paper 33, 3 (PTAB 2014); CBM2013-00013, Paper 61, 29 (PTAB 2014). For this Petition, Petitioner applies prior art in a manner consistent with disclosure of the claimed features in the ’448 Patent itself and Patent Owner’s allegations of

infringement before the district court.

D. Level of Ordinary Skill in the Art

A person of ordinary skill in the art relating to the subject matter of the '448 Patent as of the Critical Date ("POSITA") would have had a bachelor's degree in computer science, electrical engineering, computer engineering, or a related field, with 2-3 years of experience in digital video processing, or a related field.

APPLE-1003, ¶¶37-39. Additional graduate education could substitute for professional experience, and *vice versa*. *Id.*

II. SUMMARY OF THE '448 PATENT

A. Technical Background

The '448 Patent relates to video coding technology, including the use of "super-resolution" techniques to enhance pictures in a video coding system.

The '448 Patent discloses a predictable arrangement of conventional video coding features that were well known to a POSITA before the Critical Date.

APPLE-1003, ¶¶25-27. These features include:

(i) **motion-compensated coding** in which pictures (or regions of pictures) in a video stream are compressed by using a previously decoded picture (or a region of the previously decoded picture)—referred to as a "reference picture"—to predict other picture(s) (or region(s) of the other picture(s)) in the video stream;

(ii) **loop filtering** in which the previously decoded pictures—i.e., “reference pictures”—are enhanced before being used to predict other picture(s);

(iii) **super-resolution processing** in which multiple lower resolution pictures of the same scene are registered and used to form a super-resolution picture of the scene that has higher information content (e.g., more detail) than any of the individual lower-resolution input pictures and a higher spatial resolution (e.g., higher pixel count¹) than the input pictures; and

(iv) **layered, hierarchical coding** in which two or more related video streams are encoded together to provide a base layer that can be decoded independent of the other layers and one or more enhancement layers that are decoded in dependence on the base layer.²

¹ The '448 Patent sometimes uses the term “resolution” to refer to the amount of detail (e.g., information content) in a picture and at other times uses the term “resolution” to refer to the size (e.g., number of pixels) in the picture. APPLE-1003, ¶¶25-27. The meaning is context dependent. *Id.*

² Common forms of layered, hierarchical coding include multi-view coding (e.g., for stereoscopic video) and scalable coding for transmission of multiple streams

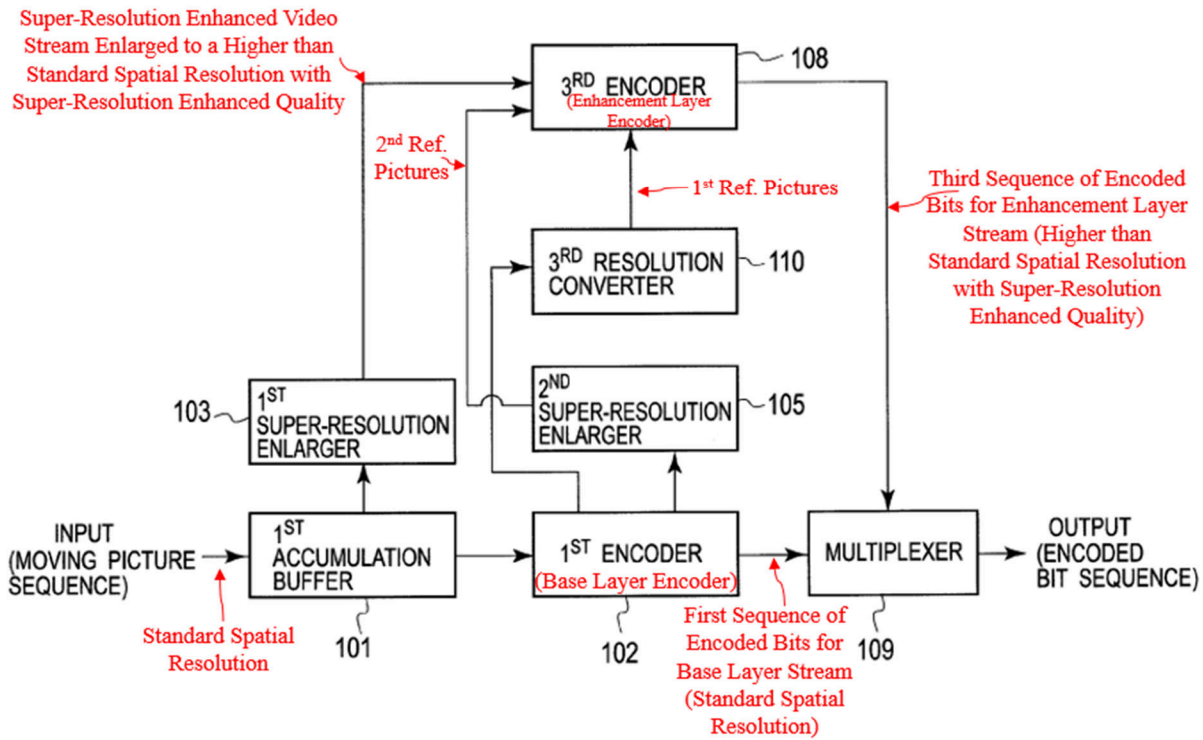
B. Brief Description

The '448 Patent describes layered or hierarchical video coding systems that apply conventional super-resolution techniques to enhance both (i) an original video stream that is provided as input to a video encoding/compression process and (ii) reference pictures used in a process for predicting pictures in the super-resolution enhanced video stream. APPLE-1003, ¶28. FIG. 13's depiction of an encoding apparatus is instructive:

differentiated by signal-to-noise ratio, frame rate, and/or spatial resolution.

APPLE-1003, ¶26; *generally id.*, 25-27.

FIG. 13



APPLE-1001, FIG. 13 (annotated)

As shown above, the encoding apparatus passes pictures from the input video sequence through an accumulation buffer 101 to each of two encoding paths. APPLE-1001, 33:60-34:4. In the first path, the original, base version of the input video having a standard spatial resolution is encoded by the first encoder 102 to produce a first coded bitstream for the base version of the video at the standard spatial resolution. *Id.*, 37:22-29. The first encoder 102 is also operable to decode the base version of the encoded input video at the standard resolution. *Id.*, 37:30-31. The set of decoded pictures with the standard resolution are provided from the first encoder 102 to the second super-resolution enlarger 105 and the third

resolution converter 110, each of which operates on the set of decoded pictures with the standard resolution to generate a respective set of reference pictures with a resolution higher than the standard resolution. *Id.*, 37:36-52. APPLE-1003, ¶29.

In the second path, a set of encoding target pictures are created by processing the pictures from the input video through a first super-resolution enlarger 103. APPLE-1001, 34:46-61. The first super-resolution enlarger 103 applies the same or similar super-resolution processing techniques as the second super-resolution enlarger 105 to enhance the spatial resolution and information content (e.g., quality) of the pictures from the original input video sequence. *Id.*, 34:13-18. For each input video picture that is to be subjected to super-resolution processing, the first accumulation buffer 101 accumulates subsequences of pictures from the neighborhood of the input video picture that is to be super-resolution enhanced, in which the pictures in the subsequences are registered with sub-pixel accuracy to provide a corresponding super-resolution image. *Id.*, 33:57-40:64; *see also id.*, 4:27-42, 33:57-40:64; APPLE-1003, ¶30. These super-resolution enlarged input pictures are provided to the third encoder 108 as encoding target pictures. *Id.*

As noted above, the encoding apparatus processes the set of decoded pictures with the standard resolution from the first encoder 102 in at least two ways. The second super-resolution enlarger 105 processes the set of decoded pictures with the standard resolution to create super-resolution enlarged decoded

pictures with a resolution higher than the standard resolution. APPLE-1001, 33:57-40:64. The third resolution converter 110 processes the same set of decoded pictures with the standard resolution using conventional resolution conversion (rather than super-resolution techniques) to create resolution-converted enlarged decoded pictures with a resolution matching that of the super-resolution enlarged pictures.³ *Id.* APPLE-1003, ¶31.

³ In the '448 patent, resolution conversion can include increasing or decreasing the size (and hence the spatial resolution) of a picture depending on whether either an enlarged or shrunken version of the picture will be produced for further processing. *See* APPLE-1001, 14:59–15:5 (describing the resolution conversion ratio); 26:28–27:9 (defining the resolution conversion ratio and explaining that a ratio less than 1 indicates downscaling, whereas a ratio greater than 1 indicates upscaling); 36:45–64. For example, the third resolution converter 110 performs upscaling to increase the resolution of an input picture since the third encoder 108 processing pictures at a resolution higher than the standard resolution. *Id.*, 34:26–45 (describing third resolution converter 110); 33:57–34:25, 36:53–64, 37:46–56 (explaining that the third resolution converter generates output pictures having a higher resolution than the input standard-resolution pictures); FIG. 13. By contrast, the first resolution

The third encoder 108 receives three inputs: (1) super-resolution enlarged pictures from the first super-resolution enlarger 103 as encoding target pictures, (2) resolution-converted enlarged decoded pictures from the third resolution converter 110 as first reference pictures, and (3) super-resolution enlarged decoded pictures from the second super-resolution enlarger 105 as second reference pictures.

APPLE-1001, 34:46-61. The third encoder 108 implements a combination of processes for prediction and encoding to create a third sequence of encoded bits representing an encoded/compressed version of the super-resolution enlarged pictures from the first super-resolution enlarger 103. *Id.* Further, the third encoder

converter 104 and second resolution converter 106 perform downscaling to decrease the size of pictures produced for further processing. *Id.*, 14:17–26 (describing first resolution converter 104); 14:49–15:5 (describing second resolution converter 106); 26:37–47 (indicating that the first and second resolution converters decimate higher-resolution pictures to produce standard-resolution pictures); 25:1–26:36; FIGS. 1, 6. The '448 Patent discloses example resolution converters that resize an image according to a conversion ratio by applying up-sampling, filtering, and/or decimation processes on a given image without referencing additional pixels from other images of a scene, as the disclosed super-resolution enlargers are configured to operate. *Id.*, 25:1-27:17, FIG. 6.

108 makes a selection of reference picture for each picture or for each set of a prescribed number of pictures. *Id.*, 35:6-14 . The third encoder 108 may be configured to create a set of data on the selection of reference pictures to identify a first reference picture or a second reference picture whichever is used. *Id.* Finally, the multiplexer 109 combines the first sequence of encoded bits from the first encoder 102 and the third sequence of encoded bits from the third encoder 108 into a single output signal. *Id.*, 36:13-44. APPLE-1003, ¶32.

The configuration of the encoding apparatus depicted in FIG. 13 permits encoding of higher-resolution content while still maintaining the base layer encoding for compatibility with standard resolution decoders. APPLE-1001, 33:57-41:22, FIGs. 13 and 15. By providing both super-resolution enlarged reference pictures and conventionally upsampled reference pictures to the third encoder, the system allows for flexible reference picture selection to optimize coding efficiency at the higher resolution. *Id.* APPLE-1003, ¶¶33-34.

C. Examination History

The '448 Patent issued on May 26, 2015 from U.S. Patent Application No. 12/995,039 (“the '039 application”), which was filed on November 29, 2010 with originally 26 claims. *See* APPLE-1002, 15-25 (NOA), 271-288 (original claims).

The Examiner mailed only a single office action during prosecution. In the non-final action of May 15, 2014, the Examiner rejected the pending claims as

obvious over Wan (APPLE-1075) in view of Phek (APPLE-1005) and Holcomb (APPLE-1024). *See* APPLE-1002, 90-107 (Non-final rejection), 197-203 (OA Election/restriction). The Examiner asserted that Wan's scalable video encoding system, when modified by Phek's super-resolution processing and Holcomb's multi-resolution techniques, would render application claims 7-9 (corresponding to issued claims 1-3) obvious. In response, the applicant amended the claims by adding further detail to the claim preambles and adding the limitations recited in elements [1.6]-[1.7] of the issued claims. On January 22, 2015, following the applicant's response, the Examiner issued a notice of allowance. *See* APPLE-1002, 56-75 (amended claims in applicant response), 76-82 (applicant argument to the OA), 15-25 (NOA). The Examiner's reasons for allowance reflect an erroneous belief that the prior art did not disclose or render obvious the predictable collection of conventional features recited by each of claims 1-3. *Id.*

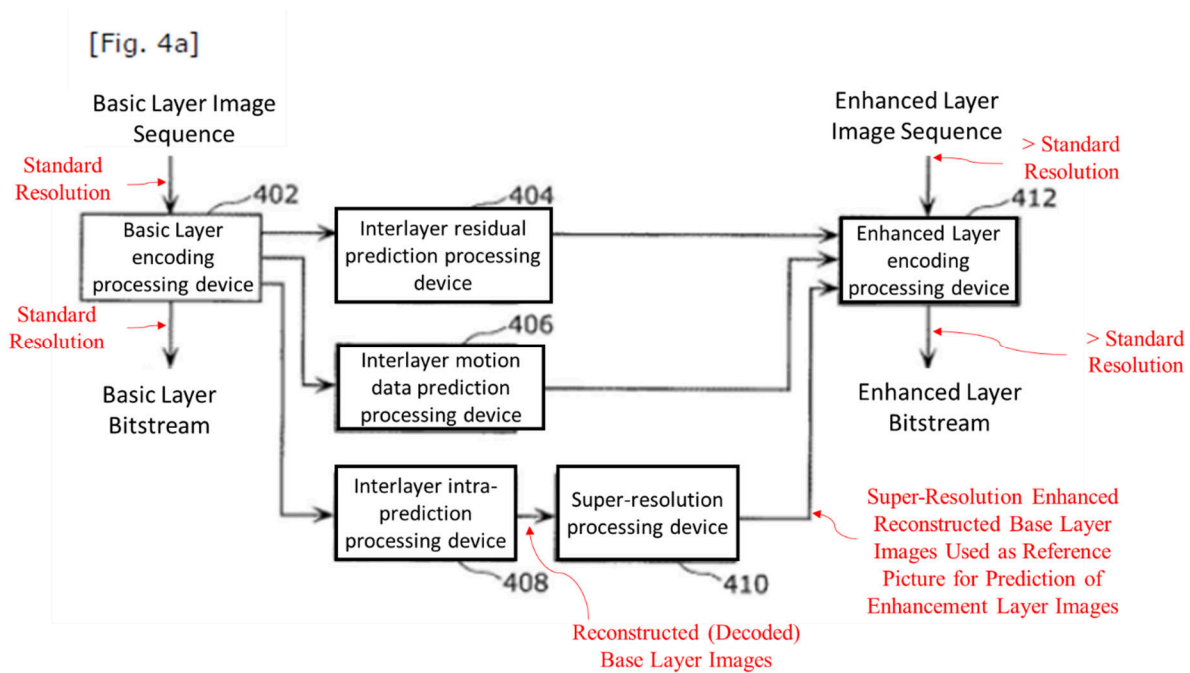
There is no indication in the file history, however, that the Examiner ever considered Phek's teachings in combination with Segall, He, and Takahashi as set forth in Grounds 1A-1D of this Petition. If these combinations had been considered, the application never should have been allowed. APPLE-1003, ¶¶35-36.

III. THE CHALLENGED CLAIMS ARE UNPATENTABLE

A. [GROUND 1A] – Phek in view of Segall and He Renders Claims 1-3 Obvious

1. Phek Overview

Phek discloses a video coding system that uses “[i]nterlayer prediction ... for hierarchical encoding in order to improve encoding efficiency.” APPLE-1005, [0014]. “One example of hierarchical encoding that uses interlayer prediction is ... scalable video encoding.” *Id.* Phek’s interlayer coding system includes a basic layer and an enhanced layer, 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.* Moreover, “super-resolution processing ... is used to improve these additional reference images” from the basic layer. *Id.*; *see also id.*, [0011]-[0020], FIGS. 1a-1b, 3a-3b, 4a-4b, 5. APPLE-1003, ¶41.

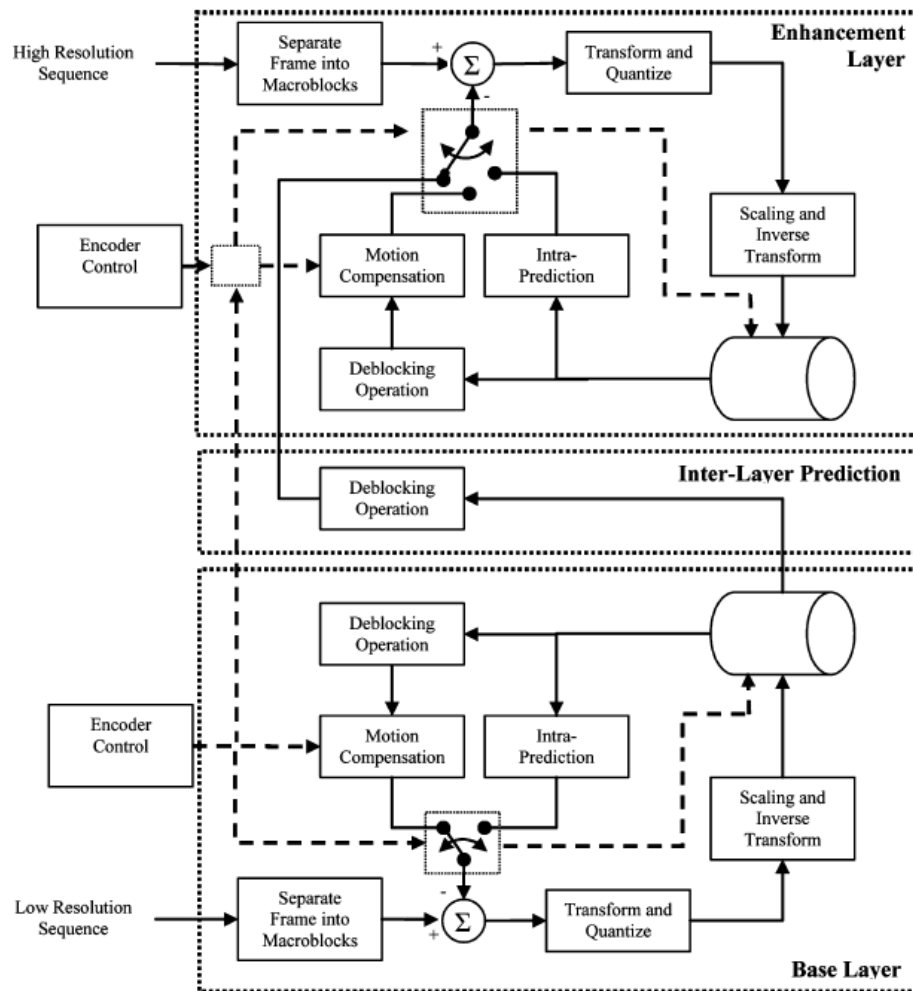


APPLE-1005, FIG. 4a (modified and annotated)⁴; APPLE-1003, ¶41.

2. Segall Overview

⁴ To aid in the presentation of Phek’s figures, Petitioner has added labels to the boxes in Phek’s figures using text of the labels that are external to the boxes in the certified English translation. *See* APPLE-1005. Petitioner also notes that the English translations of Phek (APPLE-1005 and -1031) use the terms “base” and “basic” layers (or streams) interchangeably, and further use the terms “extended,” “enhanced,” and “enhancement” layers (or streams) interchangeably. This Petition cites and quotes disclosure from the certified translation that uses different ones of these terms, but this Petition typically refers simply to “base” and “enhancement” layers (or streams) consistent with common usage of these terms in the field.

Segall is an IEEE paper titled “Spatial Scalability Within the H.264/AVC Scalable Video Coding Extension.” APPLE-1006, §I. Segall surveys aspects of 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*.” *Id.* (emphasis in original). Segall describes spatial scalability with reference to a layered coding system that, like Phek’s system, employs inter-layer prediction as illustrated in FIG. 1: APPLE-1003, ¶42.



APPLE-1006, FIG. 1

According to Segall, “[t]he ‘base layer’ encoder takes a lower resolution video sequence as input and encodes it with the H.264/AVC video coding standard while conforming to a legacy profile.” APPLE-1006, §II. “The enhancement layer encoder takes a higher resolution sequence as input” and encodes it “with ordinary H.264/AVC technologies.” *Id.* “Moreover, inter-layer prediction can be used to provide additional coding choices” and “reconstructed intensities provide a prediction for the enhancement layer.” *Id.* Segall’s system “up-samples the

decoded samples of the lower resolution layer to form a prediction” and “[t]hen, when indicated by the encoder, it receives residual difference information in the enhancement bit stream and adds this residual difference data to the prediction.”

Id., §III.C; *see also id.*, (“up-sampled to the higher resolution grid”), §III.A; APPLE-1003, ¶43.

3. He Overview

He discloses techniques for selective in-loop filtering in a scalable video coding system. APPLE-1003, ¶44. He specifically focuses on selectively enhancing reference pictures via in-loop deblock filtering to control deblock filtering parameters at each layer of a scalable coding hierarchy. APPLE-1008, [0002]-[0006], [0022], [0028], [0032]-[0037], FIGS. 1-4. APPLE-1003, ¶44.

According to He, “[t]he deblocking filter enhances the decoded picture quality and increases the encoding efficiency by removing the block effects in the boundary of each 4x4, 8x8, and/or 16/16 block,” although “[t]he deblocking filter[] ... also consumes a significant amount of power and processing cycles.” APPLE-1008, [0006]. “Disabling the entire deblocking filter [] provides a significant reduction of computational complexity and power consumption, especially in the decoder of a video processing system.” *Id.*, [0028]. He thus proposes techniques for examining “the 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 of the video information” and “also based on power savings considerations.” *Id.*, [0028], [0037]. APPLE-1003, ¶45.

FIG. 2 (reproduced below) depicts a video decoder 200 configured to implement selective deblock filtering. APPLE-1008, [0032], FIG. 2. The decoder 200 receives a coded bitstream (BTS) and performs prediction and reconstruction operations to decode at least a portion of a picture, thereby generating “unfiltered reconstructed frame information URF’.” *Id.*, [0032]. URF’ is provided to a deblocking filter 211, controlled by a deblock control circuit 221. *Id.*, [0033]-[0034]. Deblock control circuit 221 issues a “decoder filter control DFC signal” that selectively enables or disables the deblocking filter 211. *Id.*, [0034]. When filter 211 is disabled, deblock filtering operations are bypassed and “URF’ is passed as the ‘filtered’ reconstructed information RF’ to a reference frame buffer 213 for use as a reference picture for motion-compensated prediction processes. *Id.* When filter 211 is enabled, URF’ is filtered to generate RF’ for use as a reference picture in the motion compensated prediction processes. *Id.* APPLE-1003, ¶46.

He further explains that “the video encoder 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.” APPLE-1008, [0034]. APPLE-1003, ¶47.

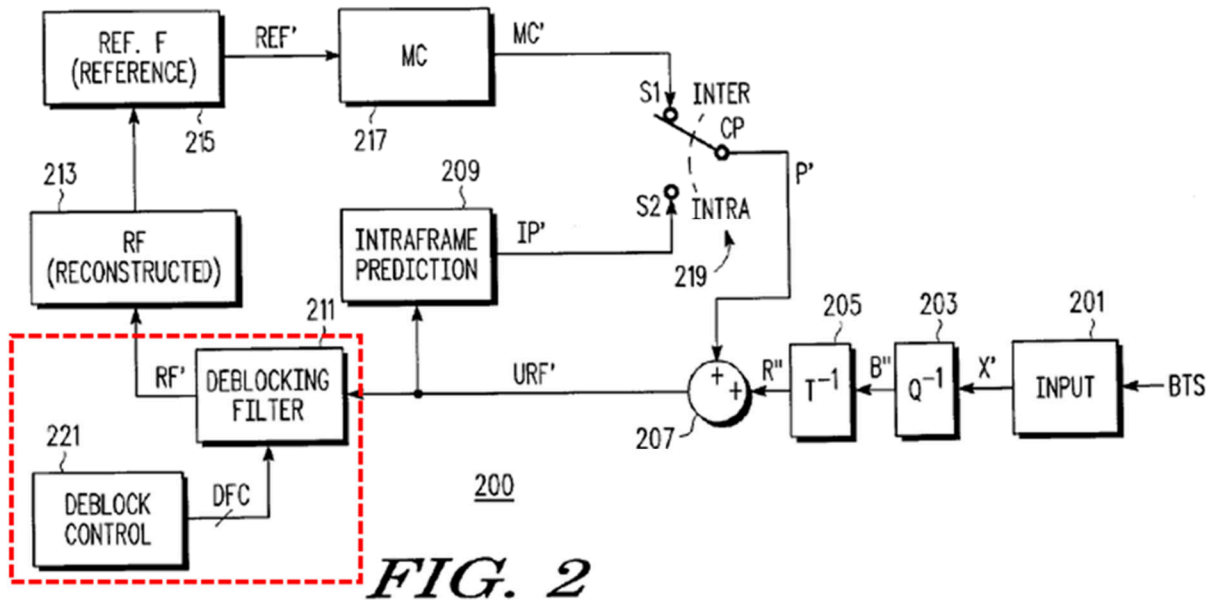


FIG. 2
In-Loop Deblocking Filter 211 and
Control Circuit 221

APPLE-1008, FIG. 2 (annotated); APPLE-1046, [0131], [0253], FIGs. 6, 12, 15.

4. Combination of Phek, Segall, and He

Like the '448 Patent, Phek discloses a layered video coding system that processes decoded base layer pictures to create reference pictures for the prediction of pictures in an enhancement layer. *Supra*, §II.A-B. While Phek's disclosure focuses on a super-resolution algorithm used to enhance decoded base layer pictures, Phek's description of the layered coding system itself (which uses the proposed super-resolution algorithm) is less detailed. Indeed, Phek acknowledges that its encoder/decoder descriptions are "simplified" and "do not have to be

performed as shown.” APPLE-1005, [0013]. Phek’s system is thus ripe for improvement. APPLE-1003, ¶48.

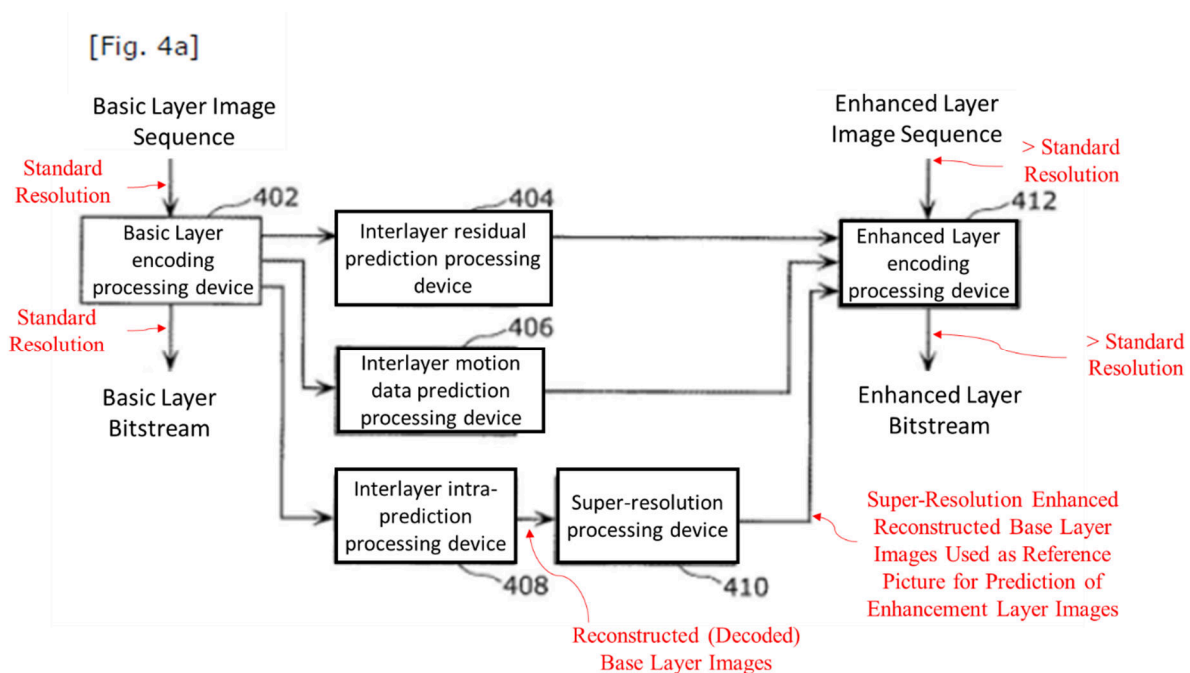
As described below, a POSITA reviewing Segall and He would have found it obvious to improve Phek’s layered coding system in at least the following ways.⁵ APPLE-1003, ¶48-71.

(a) Obvious based on Phek to form an enhancement layer video input by super-resolution processing a base layer video input

Phek teaches that “Fig. 4a and Fig. 4b are two block diagrams showing use of a super-resolution processing device for intra-layer prediction in, respectively, an encoding device (encoder) and decoding device (decoder).” APPLE-1005, [0014]. According to Phek, “[i]nterlayer prediction is used for hierarchical encoding in order to improve encoding efficiency” and “[o]ne example of hierarchical encoding that uses interlayer prediction is ... known as scalable video encoding.” *Id.* As shown in FIG. 4A, “[a] basic layer image sequence is encoded using a basic layer encoding processing device 402” and “[a]n enhanced layer

⁵ This overview of the Phek-Segall-He combination is referenced and incorporated in the analysis of each claim element in §III.A.5 below. The overviews of the combinations in each ground of this Petition are likewise incorporated in the respective analyses of the claim elements in each ground.

image sequence is encoded using an enhanced layer encoding processing device 412,” where “[t]he enhanced layer encoding processing device 412 uses 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.* APPLE-1003, ¶50.



APPLE-1005, FIG. 4A.

To the extent Phek does not expressly describe how the enhanced layer image sequence encoded by enhanced layer encoding processing device 412 is formed, it would have been obvious for the enhanced layer image sequence to be formed by applying Phek’s super-resolution processing to the base layer image sequence. APPLE-1003, ¶51. Multiple reasons would have prompted a POSITA

to apply Phek's super-resolution processing to the base layer image sequence to generate the enhanced layer image sequence. APPLE-1003, ¶51.

First, Phek's super-resolution processing represents a known and beneficial option for *enhancing* a video stream. It would have been obvious and beneficial to generate an enhanced layer image sequence to achieve Phek's express goal of obtaining such an *enhanced* layer image sequence for encoding. APPLE-1003, ¶52. Indeed, Phek specifically teaches that its "super-resolution processing method ... improves the resolution and quality of image sequences" while "reduc[ing] grid artifact[s]" in the enhanced images. APPLE-1005, (57); *see also id.*, [0002], [0003], [0011], [0013]. Considering the quality- and resolution-based enhancements offered by super-resolution processing—which are expressly touted in Phek—super-resolution processing would have been obvious to use in forming an enhanced layer image sequence that could be encoded by the enhanced layer encoding processing device 412. APPLE-1003, ¶52.

Second, a POSITA would have recognized that applying super-resolution processing to the base layer image sequence to produce an enhanced layer image sequence for encoding is consistent with Phek's multiple teachings for applying super-resolution processing to video inputs that are to be encoded/compressed. APPLE-1003, ¶53. To this point, Phek explains that "[w]hen super-resolution processing is used in recording devices ..., *the resolution and quality of images*

subject to transcoding and storage *can be improved.*” APPLE-1005, [0002]

(emphasis added). Phek also discloses with respect to FIG. 2 that super-resolution processing can be used in a “pre-processing device” that obtains an uncompressed image sequence D254 and that thereafter applies super-resolution processing to produce a super-resolution image sequence D266 that can be displayed, stored, or *compressed* (which a POSITA would have understood to include encoding).

APPLE-1005, [0010] (“[c]ompression may also be used”), FIG. 2; APPLE-1003,

¶53. Thus, even if Phek does not specifically discuss details of the enhanced layer image sequence in FIG. 4A, Phek does expressly disclose how a source image sequence D254 analogous to the base layer image sequence of FIG. 4A can be enhanced by super-resolution processing and subsequently encoded. *Id.* In this context, a POSITA naturally would have understood that Phek’s own enhancement technique (i.e., super-resolution processing) would apply to obtain an enhanced layer image for encoding by the layered encoding system of FIG. 4A. At minimum, it would have been obvious to do so based on the context of Phek’s disclosure to achieve known advantages of super-resolution enhancements as described in this Petition and in Phek. *Boston Scientific Scimed, Inc. v. Cordis Corp.*, 554 F.3d 982, 991 (Fed. Cir. 2009) (“Combining two embodiments disclosed adjacent to each other in a prior art patent does not require a leap of inventiveness”); APPLE-1003, ¶53.

Third, it would have been obvious to form the enhanced layer image sequence by applying super-resolution processing to the base layer image sequence in order to produce an enhanced layer image sequence that is related to the base layer image sequence consistent with conventional scalable coding techniques.

APPLE-1003, ¶54. A POSITA would have appreciated that the relationship between the resulting base and enhanced layer image sequences would have improved coding efficiency because the layered coding system removes redundancies between the base layer image sequence and the enhanced layer image sequence during coding. *Id.*; APPLE-1005, [0014] (“Interlayer prediction is used for hierarchical encoding in order to improve encoding efficiency. ... The enhanced layer encoding processing device 412 uses 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.”). Super-resolution processing also has the effect of both increasing the quality (e.g., information content, signal-to-noise ratio (SNR)) and size/spatial resolution of an input image sequence, which are two characteristics of base and enhancement layer image sequences that are commonly differentiated between layers of a scalable video coding system like that disclosed in Phek. APPLE-1005, (57), [0002]-[0004], [0011], [0013]; *see also* APPLE-1010, 1103-1104; APPLE-1003, ¶54. A POSITA thus would have seen super-resolution processing as a

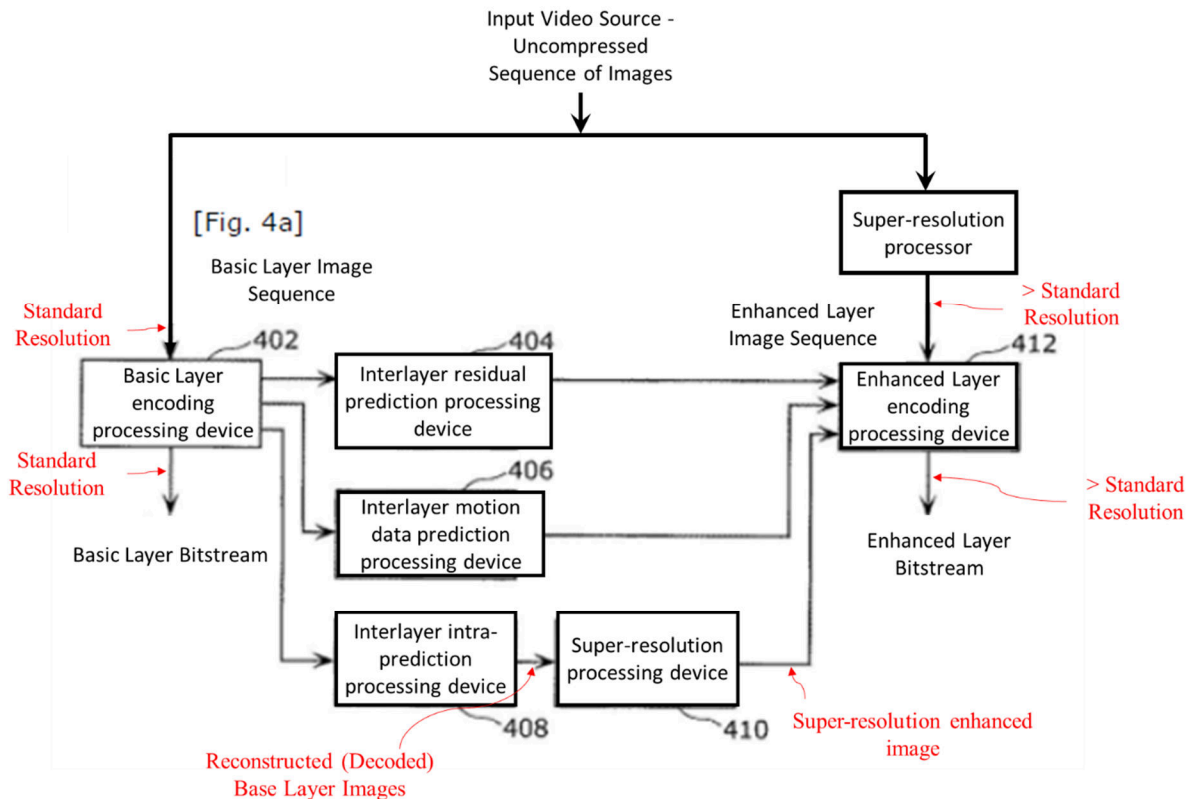
beneficial technique for generating an enhanced layer image sequence that could be suitably encoded by a layered/scalable coding system as taught in Phek.

APPLE-1003, ¶54.

Fourth, a POSITA would have sought to form the enhanced layer image sequence through super-resolution processing to overcome limitations that may have suppressed the quality and size of the original input video (base layer image sequence). APPLE-1003, ¶55. For example, the original input video may have been captured using a camera that was only capable of recording video at standard definition (e.g., 480i or 480p), but a user may have a need or desire to code and transmit the video for viewing in high definition (e.g., 1080i or 1080p) to increase viewer satisfaction, and to better utilize the higher-resolution display capabilities of the receiver. APPLE-1003, ¶55. In this case, super-resolution processing would have been obvious and beneficial to enhance the quality and resolution of the original standard definition video, so that the video could be later coded, decoded, and viewed at a higher definition. *Id.* Moreover, by performing super-resolution at the encoding system before compression and storage or transmission of the video, the viewer's device would beneficially be able to view the video at a higher quality and resolution simply by decoding the enhancement layer stream, without need to further apply super-resolution post-processing to a decoded version of the base layer image sequence at the receiver/viewing device. *Id.*; *see also* APPLE-

1007, 804-805 (describing super-resolution post-processing performed at the receiver/viewing device).

Fifth, a POSITA would have found it obvious to generate Phek's enhanced layer image sequence by applying super-resolution processing to the base layer image sequence because doing so merely entails the predictable application of known technique (e.g., super-resolution processing) to a known system (e.g., Phek's layered coding system) to achieve predictable results. *KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 417 (2007) (a supposed invention that "simply arranges old elements with each performing the same function it had been known to perform" is obvious). "It's enough... to show that there was a known problem... in the art, that [another known teaching] ... helped address that issue, and that combining the teachings of ... wasn't beyond the skill of an ordinary artisan. Nothing more is required to show a motivation to combine under *KSR*." *Intel Corp. v. PACT XPP Schweiz AG*, 61 F.4th 1373, 1380-81, (Fed. Cir. 2023). Here, Phek expressly describes how to perform super-resolution processing and the implementation of such techniques was well within the capability of a POSITA. APPLE-1003, ¶56. A POSITA would have reasonably expected success using super-resolution to form the enhanced layer image sequence in Phek's system, especially since such super-resolution techniques were well known and expressly taught by Phek before the Critical Date. *Id.*



First Intermediate Illustration of Proposed Combination (Phek)

- (b) **Obvious based on Segall to form enhancement layer reference pictures by applying conventional resolution conversion to upscale decoded base layer pictures**

Phek's interlayer predictive encoding device of FIG. 4A "use[s] ... a super-resolution processing device for intra-layer prediction." APPLE-1005, [0014], FIG. 4A. "The 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 "[a] super-resolution processing device 410 is used to improve these additional reference

images.” *Id.* As discussed, super-resolution processing has the effect of both increasing the spatial resolution of an image and improving quality of the image by adding detail based on information about a scene obtained from nearby frames in a video image sequence. *Id.*, [0002] (“improve the resolution and quality of the inputted image sequences to be displayed”), [0004], [0016]; APPLE-1012, 22 (“The basic premise for increasing the spatial resolution in SR [super-resolution] techniques is the availability of multiple LR [low-resolution] images captured from the same scene In SR, typically, the LR images represent different ‘looks’ at the same scene. That is, LR images are subsamples (aliased) as well as shifted with subpixel precision. . . . In this case, the new information contained in each LR image can be exploited to obtain an HR [high-resolution] image.”). By applying super-resolution processing to reconstructed images of the base layer, Phek’s encoding device increases the spatial resolution of decoded base layer images so that they can be used as a reference for prediction of higher-resolution enhancement layer images. APPLE-1003, ¶57. *See also*, APPLE-1024, 6:8-43, 6:64-7:21, FIG. 2 (referring to resolution converters 210 and 230); APPLE-1026, 25:27-37, FIG. 1 (referring to resolution converter 205); APPLE-1050, [0053]-[0058], [0098], FIGs. 3, 14 (illustrating upsampling process); APPLE-1051, FIGs. 1, 21, [0164]; APPLE-1052, [0110], [0060]-[0079], FIGs. 3-5, 9; APPLE-1053, [0034], [0123]-[0140], [0152]-[0162], FIGs. 2d, 6b, 5a-d; APPLE-1076, 5:63-6:56,

8:14-57, FIGs. 3A-B. Coding higher-resolution images at an enhancement layer was commonly performed in systems like Phek's, particularly in the context of spatial scalability. APPLE-1003, ¶57; *see also* APPLE-1006, 1 ("The SVC [scalable video coding] extension ... provides a mechanism for reusing an encoded lower resolution version of an image sequence for the coding of a corresponding higher resolution sequence.")

Although Phek discloses the option of super-resolution processing as a way of increasing the spatial resolution of base layer images used for interlayer prediction, a POSITA would have understood that other techniques could also be used to increase the spatial resolution of an image, including to form higher-resolution reference pictures for interlayer prediction. APPLE-1003, ¶58; APPLE-1073, [0052] ("The super-resolution processor 120 uses multiple frames of the input image to generate a high-resolution image. ... [T]he processing content of the super-resolution processing unit does not necessarily have to be super-resolution processing as described in the prior art section, and the processing content is not particularly limited as long as it is processing that improves the resolution of the input image."). In general, reference pictures for image prediction can be prepared using any suitable technique so long as the reference picture sufficiently approximates the target image that is to be predicted/decoded. APPLE-1003, ¶58. *See also*, APPLE-1017, 6:6-23, 7:39-44; APPLE-1050, [0079]; APPLE-1015,

[0010], [0014], [0018]-[0019], FIG. 1 (referring to reference picture buffer 180); APPLE-1062, FIG 1; APPLE-1005, [0014]. Better compression is typically achieved when the reference picture more closely approximates the target image. APPLE-1003, ¶58. However, the desire for high-quality reference pictures would have been balanced by a POSITA against the increased computational expense needed to generate such pictures—i.e., computational expense that can increase coding time, increase power consumption, and increase the cost of hardware needed to support the additional processing complexity attendant to the use of higher-quality reference pictures. APPLE-1008, [0005]-[0006], [0009], [0028]-[0029], [0037], [0049]; APPLE-1003, ¶58.

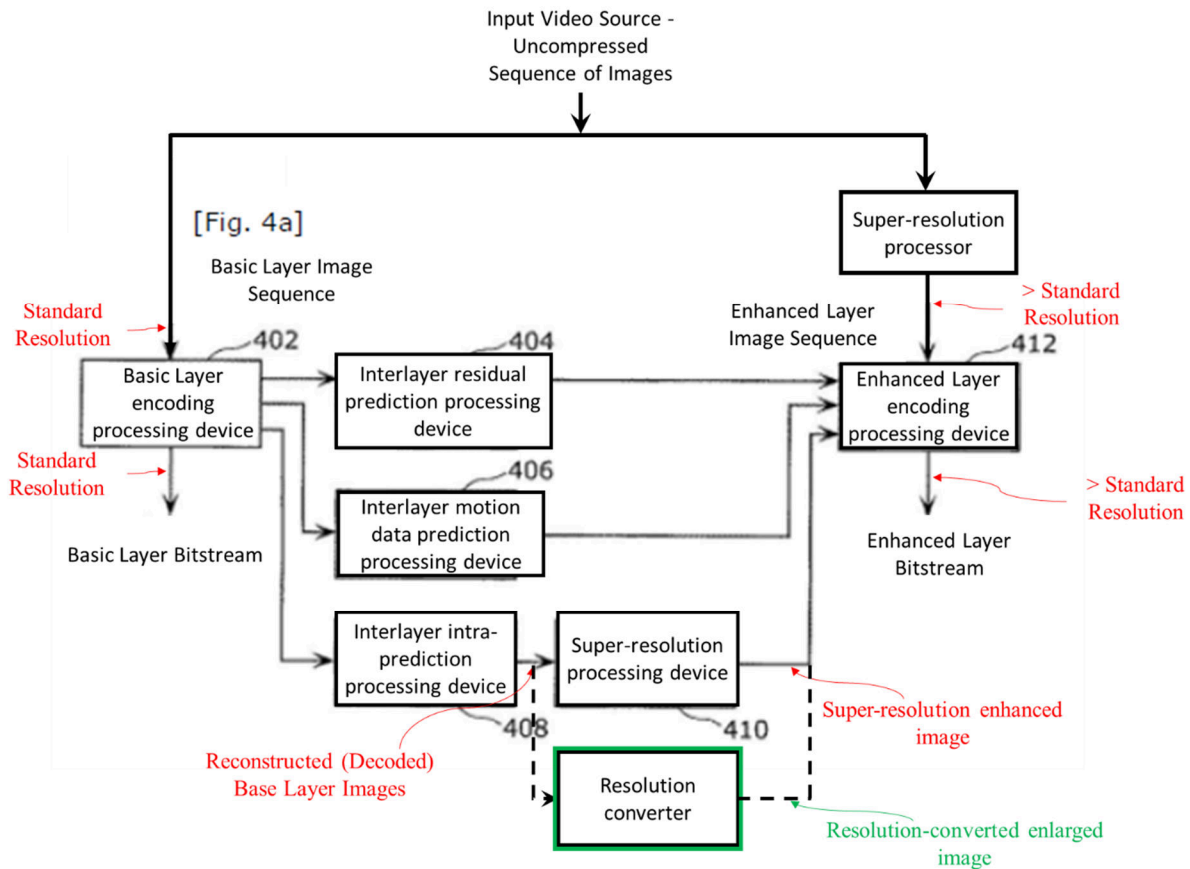
With these considerations in mind, and recognizing that super-resolution processing is a computationally expensive technique for increasing the spatial resolution of an image, a POSITA would have looked for additional or alternative methods beyond Phek's super-resolution processing that could achieve the result of increasing the spatial resolution of decoded base layer images used as reference pictures for interlayer prediction (even if the quality/detail of the base layer images is not improved). APPLE-1003, ¶59; *see also* APPLE-1055, [0014] (explaining with respect to super-resolution processing that “the runtime of image reconstruction by the iterative method ... depends on the performance of computer system”); APPLE-1074, 4:60-67 (“Inter-frame methods such as super-resolution

attempt to fuse sub-pixel (or “sub-sample”) detail that has been scattered over several pictures by aliasing and other diffusion methods, and can in fact restore definition above the Nyquist limit implied by the standard definition sampling lattice. However such schemes are computationally expensive, non-linear, and do not always yield consistent quality gains frame-to-frame.”).

Segall discloses one such conventional technique for increasing the spatial resolution of an image without the higher computational demands associated with super-resolution processing. APPLE-1003, ¶60. For example, Segall discloses a resolution conversion technique that involves “up-sampl[ing] the decoded samples of the lower resolution layer to form a prediction.” APPLE-1006, 6. “The up-sampling operation for the luma component consists of applying a separable four-tap poly-phase interpolation filter” while “[t]he chroma component of the signal is also up-sampled but with a different (simpler) interpolating kernel.” *Id.*; *generally id.*, §§III.A (describing detailed techniques for up-sampling and calculation of corresponding spatial positions between base and enhancement layer images), III.C (describing details of up-sampling decoded samples of the lower layer), FIG. 1. Segall’s teachings are illustrative of a class of well-known resolution conversion techniques that function to increase image resolution by interpolating additional pixels in an image from other pixels within the same image, but without the need for a series of images as required by super-resolution processing. *See, e.g.*,

APPLE-1012, 22-23 (“[I]mage interpolation ... has been used to increase the size of a single image. ... [S]ingle image interpolation cannot recover the high-frequency components lost or degraded during the LR [low-resolution] sampling process. For this reason, image interpolation methods are not considered as SR [super-resolution] techniques.”); APPLE-1074, 4:53-60 (“Prior art interpolation methods can interpolate a standard definition video signal to, for example, a high definition display, but do not add or restore content beyond the limitations of the standard-definition sampling lattice. Prior art methods include, from simplest to most complex: sample replication (“zero order hold”), bi-linear interpolation, poly-phase filters, spline fitting, POCS (Projection on Convex Sets), and Bayesian estimation.”). Indeed, resolution conversion based on conventional interpolation techniques like those disclosed in Segall were commonly used by the Critical Date of the ’448 patent to adjust the size of reference pictures in scalable coding systems. *See, e.g.*, APPLE-1024, 6:8-43 (disclosing resolution converters (210 and 230) as standard components in multi-resolution encoding systems that receive input frames and output converted frames with parameters for use by other encoder modules), FIG. 2; APPLE-1026, 25:27-37 (“converts the resolution of the base layer image signal to the resolution corresponding to the enhancement image signal by means of filtering operation.”), FIG. 1 (depicting resolution converter 205); APPLE-1003, ¶60.

A POSITA would have been motivated based on Segall to implement a conventional interpolation-based resolution converter in Phek's system as an alternative to super-resolution processing for increasing the resolution of decoded base layer pictures used as reference pictures in the enhancement layer. APPLE-1003, ¶61. The use of a conventional interpolation-based resolution converter in Phek's system would have been obvious and beneficial to increase the size of the decoded base layer pictures as needed for prediction in the enhancement layer, and to do so in a less computationally demanding manner than would be required by super-resolution processing. *Id.* Conventional interpolation-based resolution conversion would balance processing complexity against image quality, and would predictably require less time/processing cycles, less power consumption, and less costly hardware (e.g., less memory and processing capability) than super-resolution processing. APPLE-1003, ¶61.



Second Intermediate Illustration of Proposed Combination (Phek-Segall)

A POSITA would have recognized that the addition of a conventional resolution converter represents a straightforward implementation of Segall's spatial scalability architecture within Phek's hierarchical encoding framework. APPLE-1005, [0014]; APPLE-1006, §II. Segall explicitly teaches the "calculation of corresponding positions in different resolution layers" and "methods for inter-layer prediction of various data," providing the technical foundation for implementing conventional resolution conversion alongside super-resolution processing. APPLE-1006, §II. The resulting combination maintains the system's ability to "enhance the quality of interlayer prediction" while introducing a resolution

conversion pathway that offers reduced computational complexity for appropriate images. APPLE-1005, [0014]; APPLE-1003, ¶61.

The use of a conventional resolution converter also addresses practical implementation considerations for system designers. APPLE-1005, [0014]. Segall recognizes that “one potential approach” to scalable encoding “would be to perform multiloop decoding,” but notes that a “lower complexity approach is adopted” in scalable coding designs. APPLE-1006, §II. The conventional resolution converter implements a lower-complexity approach, providing a computationally efficient pathway that complements the more sophisticated super-resolution processing. APPLE-1055, [0030]. This configuration enables system designers to make targeted trade-offs to optimize between processing complexity and image quality, enhancing the system’s adaptability to diverse application requirements. APPLE-1005, [0014]; APPLE-1006, §II. APPLE-1003, ¶63.

Finally, a POSITA would have found it obvious to apply Segall’s suggestion for interpolation-based resolution conversion to Phek’s system as the combination entails nothing more than the use of a known technique as taught by Segall to a known system as taught by Phek to achieve merely predictable results. *KSR Int’l Co. v. Teleflex Inc.*, 550 U.S. 398, 417 (2007) (a supposed invention that “simply arranges old elements with each performing the same function it had been known to perform” is obvious). “It’s enough... to show that there was a known problem...

in the art, that [another known teaching] ... helped address that issue, and that combining the teachings of ... wasn't beyond the skill of an ordinary artisan. Nothing more is required to show a motivation to combine under *KSR*.” *Intel Corp. v. PACT XPP Schweiz AG*, 61 F.4th 1373, 1380-81, (Fed. Cir. 2023). Here, Segall expressly teaches how to use conventional interpolation-based resolution conversion techniques in a scalable coding system like Phek's and the implementation of such techniques was well within the capability of a POSITA. APPLE-1003, ¶64. A POSITA would have reasonably expected success applying these techniques to Phek's system, especially since such techniques were well known and commonly implemented before the Critical Date. *Id.*

(c) Obvious based on He to implement reference picture selection functionality

As discussed, it would have been obvious based on Segall to increase the spatial resolution of decoded base layer pictures in Phek's system using conventional interpolation-based techniques as an alternative to super-resolution processing. *Supra*, §III.4.(b); APPLE-1003, ¶65. The choice of whether to apply interpolation-based resolution conversion or super-resolution would be informed by known and predictable factors such as the compression efficiency achieved by each technique and the relative computational demands of the different techniques. APPLE-1003, ¶65. For example, super-resolution processing typically yields a

higher-quality reference picture that more closely predicts the decoding target in the enhancement layer (especially since the enhancement layer input stream is also super-resolution enlarged in the Phek-Segall-He combination (*see supra*, §III.4.(a)), and super-resolution processing can therefore increase compression efficiency for many images. However, despite potential gains in compression efficiency, super-resolution processing can also require more processing power, memory, and longer processing times than less complex interpolation-based upscaling methods like those represented by Segall, which do not require processing a series of images to resolve a higher resolution image. APPLE-1003, ¶65; *see also* APPLE-1012, 22-23; APPLE-1074, 4:53-67; APPLE-1024, 6:8-43, FIG. 2; APPLE-1026, 25:27-37, FIG. 1; APPLE-1006, 1-6. A POSITA would have recognized that, in some cases, the improvements in compression efficiency and/or image quality achieved by enlarging reference pictures using super-resolution processing rather than conventional single-image interpolation methods do not justify the increased processing demands associated with super-resolution enlargement. APPLE-1003, ¶65. In other cases, the benefits attendant to enlarging reference pictures by super-resolution would be insignificant and would not justify the requisite increase in processing demands since conventional single-image interpolation methods would yield comparable compression efficiency and image quality. *Id.*; *see also* APPLE-1073, [0052]. Accordingly, a POSITA would have

looked for options to realize the benefits of both approaches (i.e., super-resolution enlargement and single-image interpolation-based resolution conversion) and thereby optimize compression efficiency, image quality, and memory/processor utilization during the coding process. APPLE-1003, ¶65.

He teaches one such option that would have been obvious to apply in combination with Phek and Segall to optimize the selection of reference pictures in Phek's enhancement layer coding process. APPLE-1003, ¶66. As discussed, He discloses techniques for selectively enhancing reference pictures via in-loop filtering in a scalable video coding system. *Supra*, §III.A.3. For example, He describes selective in-loop deblock filtering to control deblock filtering parameters at each layer of a scalable coding hierarchy. APPLE-1008, [0002]-[0006], [0022], [0028], [0032]-[0037], FIGS. 1-4. APPLE-1003, ¶66. According to He, "[t]he deblocking filter enhances the decoded picture quality and increases the encoding efficiency by removing the block effects in the boundary of each 4x4, 8x8, and/or 16/16 block," although "[t]he deblocking filter[] ... also consumes a significant amount of power and processing cycles." APPLE-1008, [0006]. "Disabling the entire deblocking filter [] provides a significant reduction of computational complexity and power consumption, especially in the decoder of a video processing system." *Id.*, [0028]. He proposes techniques for examining "the 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 of the video information” and “also based on power savings considerations.” *Id.*, [0028], [0037]. Based on results of this examination, He uses a switching signal to selectively enable or disable the reference picture enhancement deblocking filter. APPLE-1008, [0034], FIG. 2; APPLE-1003, ¶66.

He thus describes a switching capability that allows a video coder to selectively activate or de-activate a reference picture enhancement process. While He addresses de-block filtering as an example reference picture enhancement process, in the context of the Phek-Segall combination, a POSITA would have understood that He’s teachings are similarly applicable to super-resolution enlargement as a reference picture enhancement process/filter. *Supra*, §III.A.4.(b); APPLE-1003, ¶67. That is, as compared to conventional resolution conversion techniques as taught by Segall, Phek’s super-resolution enlargement further enhances reference pictures by adding detail to a depicted scene indicated by other images. *Supra*, §III.A.4.(b); APPLE-1005, [0002] (explaining that super-resolution processing “improve[s] the resolution *and quality* of the inputted image sequences”); APPLE-1012, 2-3 (“single image interpolation cannot recover the high-frequency components” and “[f]or this reason, image interpolation methods are not considered as SR [super-resolution] techniques”). A POSITA reviewing

He would have found it obvious and straightforward to similarly implement in the Phek-Segall system a switching capability that would allow selective use of the enhancements offered by super-resolution processing when forming reference pictures for interlayer prediction. APPLE-1003, ¶67.

For example, based on He's teachings, when content characteristics of the video indicate that the balance between coding efficiency and processing demands (e.g., power savings considerations) would be optimized by enhancing a reference picture, the system would apply super-resolution processing to the decoded base layer image to form the reference picture. APPLE-1008, [0028], [0034], [0037]; APPLE-1003, ¶67. But when content characteristics of the video indicate that the balance between coding efficiency and processing demands (e.g., power savings considerations) would be optimized by not similarly enhancing a reference picture, the system would apply conventional single-image interpolation-based upscaling as suggested in Segall to the decoded base layer image to form the reference picture. *Id.*

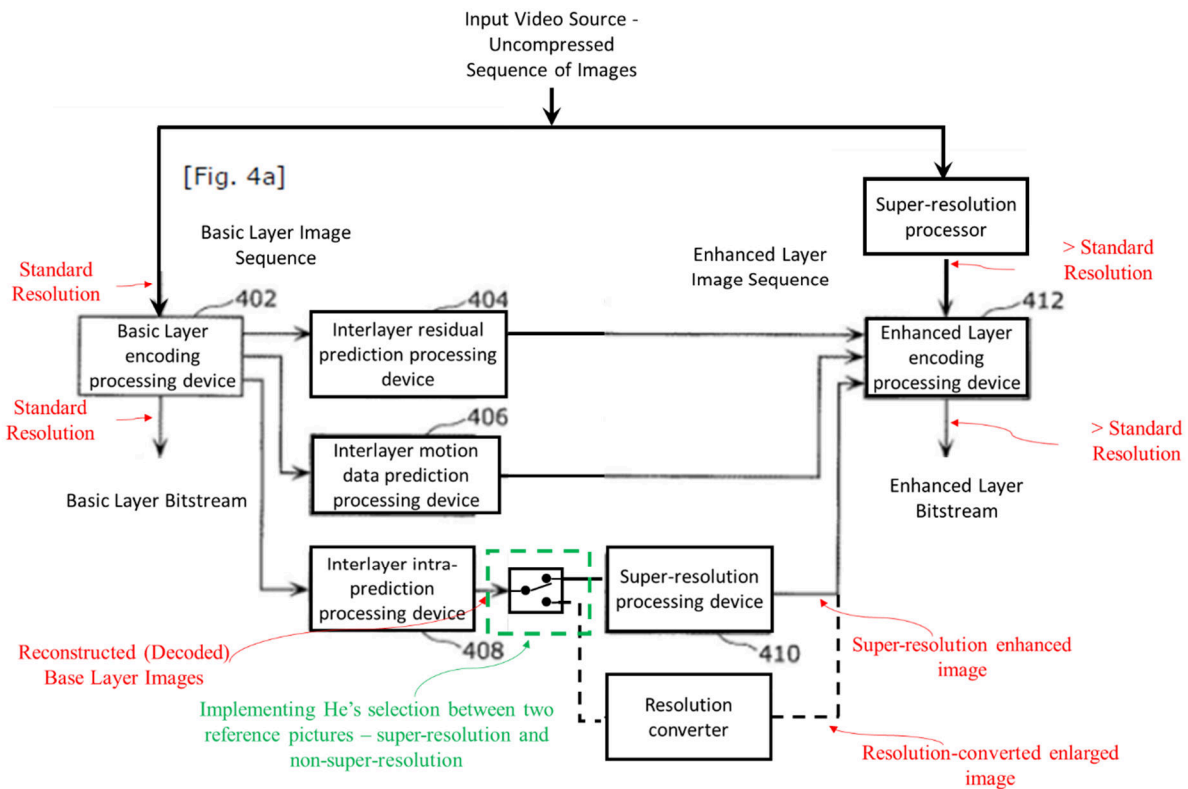


Illustration of Proposed Combination (Phek-Segall-He)

A POSITA would have been motivated to implement reference picture selection in the Phek-Segall-He combination to maintain the benefits of improved coding efficiency and picture quality offered by super-resolution processing while reducing the processing demands and power consumption of the system by using conventional resolution conversion when super-resolution processing is less beneficial. APPLE-1003, ¶68. A POSITA also would have sought to adopt such reference picture selection techniques because reference picture selection was a common and well-understood technique in the field before the Critical Date of the '448 Patent. *Id.* Even Phek and Segall acknowledge different forms of reference

picture selection can be used to optimize the performance of a video coder. *See, e.g.,* APPLE-1005, [0014] (explaining that “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”); APPLE-1006, 3 (“Considerable flexibility is ... provided in regard to the selection of the reference layer”), 2 (“[I]nter-layer prediction can be used to provide **additional** coding choices.”); APPLE-1046, [0137]-[0140], [0176], [0251], FIG. 6 (describing reference picture selectors).

As depicted above, the system in the resulting combination would predictably allow the enhanced layer encoding processing device (412) to utilize reference pictures from either a super-resolution processing device like Phek’s or a resolution converter like Segall’s, optimizing the efficiency-quality tradeoff in hierarchical video coding systems. APPLE-1005, [0014]; APPLE-1006, §II. APPLE-1003, ¶69. The implementation of He’s selection functionality directly addresses computational complexities that arise from Phek’s super-resolution processing. APPLE-1003, ¶69; *see also* APPLE-1055, [0014]. The selective application of super-resolution processing only when warranted by content characteristics provides a significant optimization of system resources. APPLE-1008, [0006]. He teaches that analogous enhancements like deblock filtering “consume a significant amount of power and processing cycles” and that

selectively disabling such enhancements “provides a significant reduction of power and processing complexity, especially in the decoder.” APPLE-1008, [0006].

APPLE-1003, ¶69. The selection functionality enables the system to route only appropriate content through the more intensive super-resolution path. APPLE-1008, [0069], [0006]; APPLE-1015, [0068]; APPLE-1011, [0022]-[0023], FIGS. 2-5); *supra*, §II.A.

A POSITA further would have recognized that the significant computational burden of super-resolution processing makes selective application highly beneficial for system efficiency. APPLE-1003, ¶70; *see also* APPLE-1055, [0014]. Second, the varying effectiveness of super-resolution enhancement across different content types creates a clear need for content-adaptive processing. APPLE-1003, ¶70; *see also* APPLE-1055, [0070]. Third, the need to maintain high output quality while managing computational resources demands a flexible approach that can dynamically select the appropriate processing path. *See* APPLE-1008, [0069]. APPLE-1003, ¶70. Implementation of this selection functionality follows established techniques in video coding. APPLE-1008, [0069]. He teaches that selective filtering allows systems to maintain a “desired tradeoff between video quality and complexity reduction.” APPLE-1008, [0069].

A POSITA also would have found it obvious to apply He’s reference picture selection functionality to Phek’s system as the combination entails nothing more

than the use of a known technique as taught by He to a predictable system as taught by Phek and Segall to achieve merely predictable results. *KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 417 (2007) (a supposed invention that “simply arranges old elements with each performing the same function it had been known to perform” is obvious). “It’s enough... to show that there was a known problem... in the art, that [another known teaching] ... helped address that issue, and that combining the teachings of ... wasn’t beyond the skill of an ordinary artisan. Nothing more is required to show a motivation to combine under *KSR*.” *Intel Corp. v. PACT XPP Schweiz AG*, 61 F.4th 1373, 1380-81, (Fed. Cir. 2023). Here, He expressly teaches how to use switching signals to selectively enable or disable reference picture enhancements in a scalable coding system like Phek’s and the implementation of such techniques was well within the capability of a POSITA. APPLE-1003, ¶71. A POSITA would have reasonably expected success applying these techniques to Phek’s system, especially since such techniques were well known and commonly implemented before the Critical Date. *Id.*

5. Analysis

(a) Claim 1

[1.P]

To the extent the preamble is limiting, the Phek-Segall-He combination renders obvious [1.P]. APPLE-1003, ¶¶72-78. For example, in FIG. 4a, Phek

illustrates a hierarchical video coding system that encodes moving picture sequences at a base layer and an enhancement layer. APPLE-1005, [0002], [0014], FIG. 4a. Phek teaches super-resolution processing as “a technique for generating high-resolution images from a plurality of low-resolution images.” APPLE-1005, [0002].

Phek’s hierarchical video coding system includes a basic layer encoder and an enhanced layer encoder, with “interlayer intra-prediction processing” using reconstructed standard-resolution images from the base layer as reference pictures for encoding the enhanced layer. APPLE-1005, [0014], FIG. 4A. The super-resolution processing device in Phek “improves these additional reference images” upscaled from the basic layer. APPLE-1005, [0014]. This upscaling is used to enhance reference frames for “motion detection, motion compensation prediction, and display.” APPLE-1005, [0013].

In the Phek-Segall-He combination as described above (*supra*, §III.A.4), the encoding system makes an encoding of a sequence of moving pictures with a resolution higher than a standard resolution using moving pictures contents which include a sequence of moving pictures with the standard resolution and do not include a sequence of moving pictures with a resolution higher than the standard resolution. For example, as described in detail in the overview of the combination (*supra*, §III.A.4(a)), it would have been obvious based on Phek for the encoding

system to produce a super-resolution enlarged video stream from an original uncompressed base layer video stream whose images are formatted at a standard resolution. *Supra*, §III.A.4(a); The resulting enhanced layer video stream thus includes moving pictures whose resolution is higher than the standard resolution as a result of super-resolution processing, and the enhanced layer video stream is encoded by the enhanced layer encoding processing device 412. APPLE-1005, [0002], [0011], [0013]-[0015], FIG. 1a (depicting super-resolution processing device), 3a (depicting super-resolution processing device in an encoder), 4a (depicting super-resolution processing device in an encoder using interlayer prediction). APPLE-1005, [0024].

A POSITA also would have understood and found obvious that the original input video sequence, which is used to form the enhanced layer video stream, and which is encoded by basic layer encoding processing device 402, would include only moving pictures formatted at the standard resolution, and would include no moving pictures formatted at a higher resolution, especially since each layer of a scalable coding systems like Phek's conventionally encodes pictures only at a single spatial resolution different from the spatial resolutions of pictures encoded in different layers. APPLE-1005, [0014], FIG. 4A; *see also* APPLE-1006, 2-3 ("The lowest resolution video data in a spatially scalable system is sometimes referred to as the *base layer* (especially when it is decodable by an ordinary non-

scalable single-layer decoder), and the higher resolution video data is often referred to as the *enhancement layer*.”), 3 (“two layers—a lower resolution base layer and a higher resolution spatial scalability enhancement layer”). The base layer video stream would thus include pictures only at a standard resolution for coding at the standard resolution and the enhanced layer video stream would thus include pictures only at a resolution higher than the standard resolution for coding at the higher resolution. This would have been understood, obvious, and beneficial because input video streams commonly encode pictures at a single spatial resolution and because it would simplify the system by processing all pictures from the input video at the same spatial resolution.

Phek’s decoding path reinforces this understanding, applying super-resolution to reconstructed basic-layer content to produce enhanced-layer outputs, confirming no high-resolution input is needed. APPLE-1005, [0015], FIG. 4b. Additionally, Phek’s super-resolution method performs “motion compensation on observed low-resolution images to convert them into a common high-resolution grid,” demonstrating the sufficiency of standard-resolution input for high-resolution encoding. APPLE-1005, [0003].

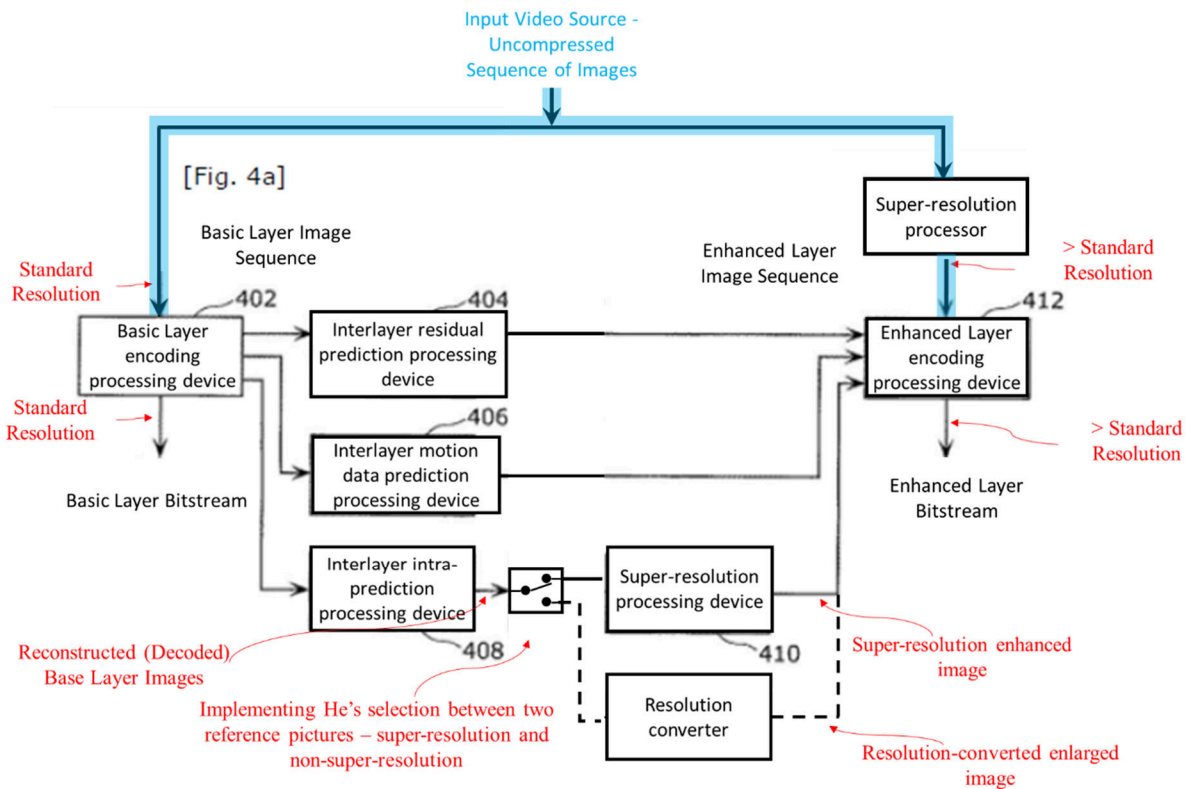


Illustration of Proposed Combination (Phek-Segall-He)

[1.1]

Phek-Segall-He renders obvious [1.1]. APPLE-1003, ¶¶72-73. For example, Phek's hierarchical encoding system includes a basic layer encoding processing device 402 that, alone or together with interlayer intra-prediction processing device 408, provides a "first encoder" that works on a basic layer image sequence ("a sequence of moving pictures with standard resolution") to generate encoded bits and decoded pictures at the same resolution. APPLE-1005, [0014]. The "basic layer encoding processing device 402" operates on basic layer image sequence having low (standard) resolution images, performing encoding processes

to create a basic layer bitstream. APPLE-1005, [0014] (“A basic layer image sequence is encoded using a basic layer encoding processing device 402.”), [0002]. Phek states that “the 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.” APPLE-1005, [0014].

The coding process implemented by Phek’s basic layer encoder includes decoding steps to generate decoded or reconstructed images. APPLE-1005, [0014]. Phek explains that the encoder utilizes “reconstructed images” as a reference for interlayer prediction, confirming that the basic layer encoder implements both encoding and decoding processes to produce the images used as references for interlayer prediction. APPLE-1005, [0014]. This reconstruction occurs as part of the encoder’s standard operation, ensuring compatibility with inter-layer prediction mechanisms. APPLE-1005, [0014]. Phek describes the operation of encoding devices like the basic layer encoding processing device 402 with respect to FIG. 3A and explains that the encoding device implements a combination of processes for an encoding and a decoding to create a sequence of encoded bits and a set of decoded pictures of the same resolution. APPLE-1005, [0011] (“In the encoder, the image sequence is intra-coded or inter-coded” and “an encoded video bit stream is produced. The target frame is reconstructed using a

local decoder ..., and the reconstructed target frame is stored as a reference frame for use by the motion detection processing device 302.”), FIG. 3A.

A POSITA would have understood that Phek’s basic layer encoder performs these dual functions to facilitate hierarchical encoding. The generation of decoded pictures at the input resolution is disclosed in Phek’s framework, which uses reconstructed images for reference in the enhanced layer encoding process.

APPLE-1005, [0014].

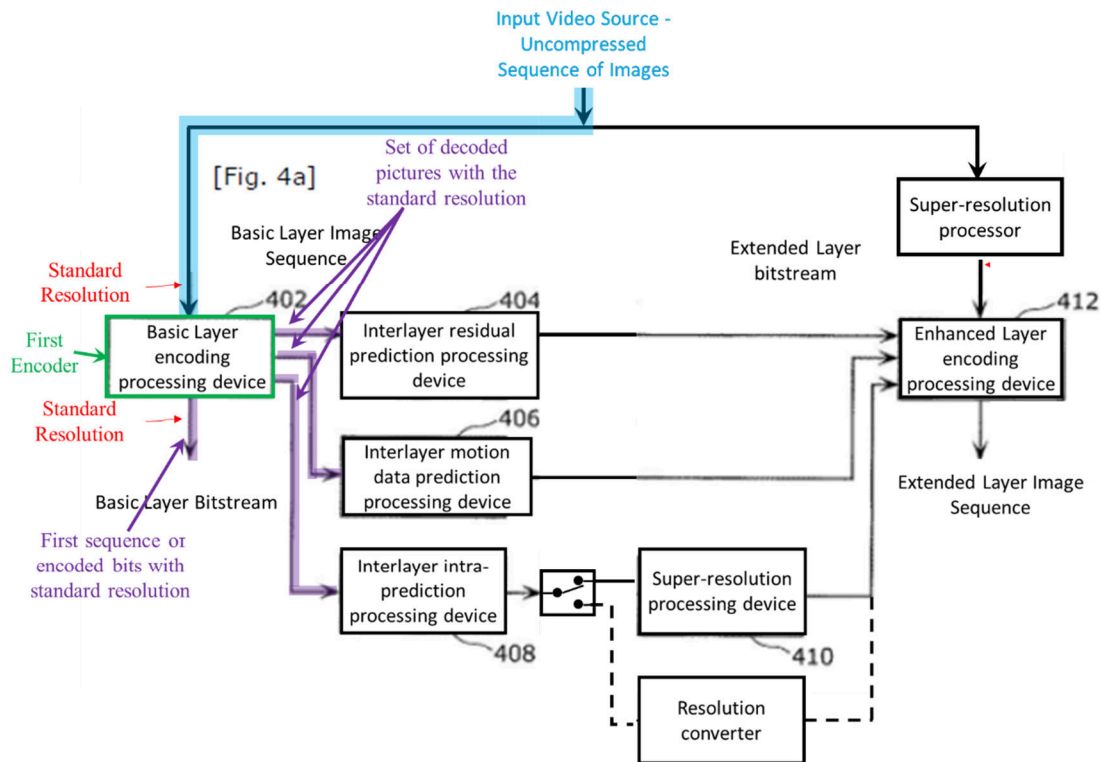


Illustration of Proposed Combination (Phek-Segall-He)

[1.2]

Phek-Segall-He renders obvious [1.2]. APPLE-1003, ¶¶72, 74. For example, Phek explains that “[s]uper-resolution processing is a technique for generating high-resolution images from a plurality of low-resolution images.” APPLE-1005, [0002]. “Super-resolution processing methods include methods that involve image registration. In these methods, motion compensation is performed on observed low-resolution images to convert them into a common high-resolution grid known as a registered image. An initial estimated high-resolution frame is obtained using a spatial interpolation method such as bicubic interpolation. This estimated frame is improved by repeatedly subjecting it to blurring, frame difference, point spreading inverse function, regularization, and refinement processing. A residual image is generated during the operation of finding frame differences. The registered image is thus subtracted from the blurred high-resolution estimated frame.” *Id.*, [0003]; *see also id.*, [0004]-[0009]. Phek further describes an improved super-resolution processing method with respect to FIGS. 1, 5, and 6. *Id.*, [0016]-[0020]. The effect of applying super-resolution processing to an image is to enhance the quality and spatial resolution of the image by incorporating into the image additional information about a scene depicted in other images in the moving pictorial sequence. *Id.*, [0011], FIGS. 1, 5, 6a-6b.

Phek's hierarchical encoding system includes a super-resolution processing device 410 that operates on decoded pictures from the basic layer encoder to generate super-resolution enlarged decoded pictures with a resolution higher than the standard resolution. APPLE-1005, [0014], FIG. 4A. While Phek focuses on improving reference images for inter-layer prediction by super-resolution processing, Phek explains that super-resolution processing can also be applied for a variety of purposes including "to improve the resolution and quality of [] inputted image sequences to be displayed" and of "images subject to transcoding and storage." APPLE-1005, [0002]; *see also id.*, [0010]-[0011], [0014], FIGS. 2, 4A-B.

Phek's hierarchical encoding system includes a basic layer encoding processing device 402 configured to encode a basic layer image sequence and an enhanced layer encoding processing device 412 configured to encode an enhanced layer image sequence. APPLE-1005, [0014], FIG. 4A. Based on Phek's specific focus on enhancing images through super-resolution processing, and for the reasons discussed in further detail above (*supra* §III.A.4.(a)), it would have been obvious to a POSITA for the extended layer image sequence to be formed by applying super-resolution processing to the basic layer image sequence, thereby enhancing the quality and spatial resolution of the extended layer image sequence. APPLE-1005, [0014], FIG. 4A; *supra*, §III.A.4.(a). The encoding system would

thus employ a super-resolution processing device, e.g., device 208, configured to work on a sequence of moving pictures from the basic layer image sequence with the standard resolution to implement a process for a first super-resolution enlargement to create a set of super-resolution enlarged pictures with a resolution higher than the standard resolution as the extended layer image sequence. APPLE-1005, [0002]-[0003], [0016]-[0020], [0014], FIGS. 1-2, 4A, 5, 6A-6B; *supra*, §III.A.4.(a).

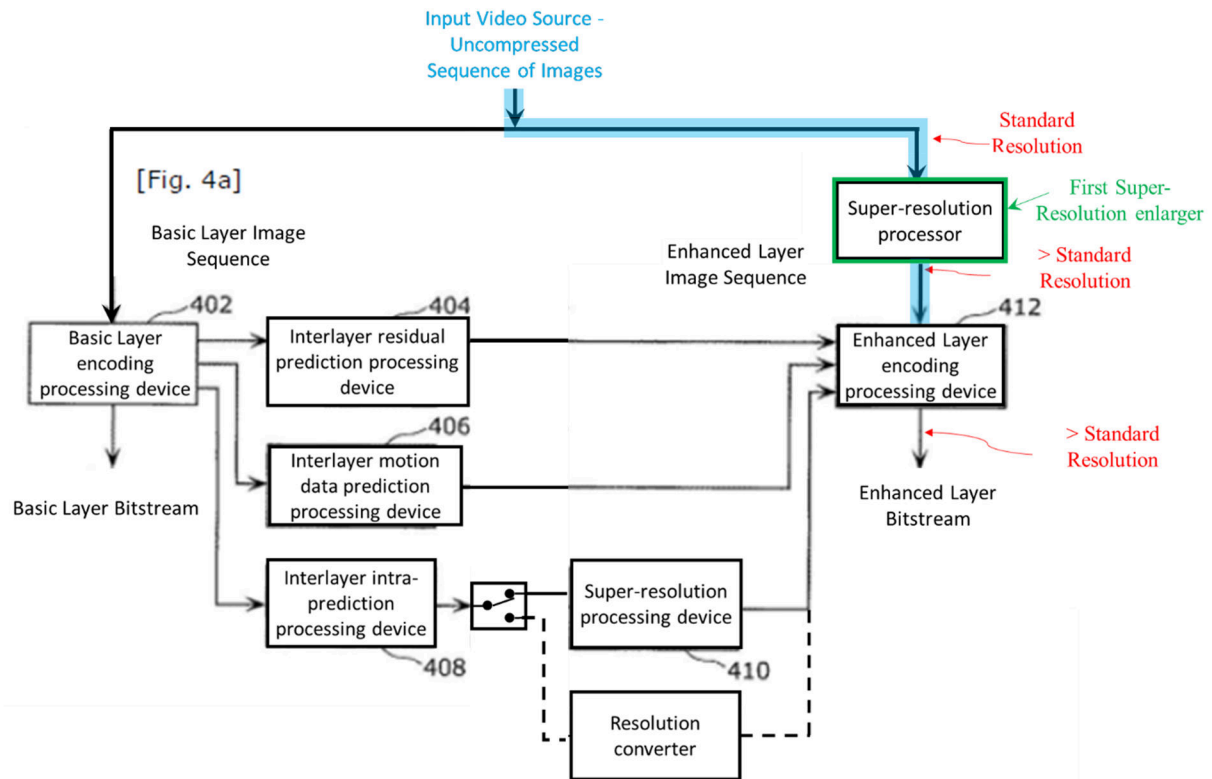


Illustration of Proposed Combination (Phek-Segall-He)

[1.3]

Phek-Segall-He renders obvious [1.3]. APPLE-1003, ¶¶72, 75. As discussed above, Phek discloses that “[s]uper-resolution processing is a technique for generating high-resolution images from a plurality of low-resolution images.” APPLE-1005, [0002]. “Super-resolution processing methods include methods that involve image registration. In these methods, motion compensation is performed on observed low-resolution images to convert them into a common high-resolution grid known as a registered image. An initial estimated high-resolution frame is obtained using a spatial interpolation method such as bicubic interpolation. This estimated frame is improved by repeatedly subjecting it to blurring, frame difference, point spreading inverse function, regularization, and refinement processing. A residual image is generated during the operation of finding frame differences. The registered image is thus subtracted from the blurred high-resolution estimated frame.” *Id.*, [0003]; *see also id.*, [0004]-[0009]. Phek further describes an improved super-resolution processing method with respect to FIGS. 1, 5, and 6. *Id.*, [0016]-[0020]. The effect of applying super-resolution processing to an image is to enhance the quality and spatial resolution of the image by incorporating into the image additional information about a scene depicted in other images in the moving pictorial sequence. *Id.*, [0011], FIGS. 1, 5, 6a-6b.

Phek’s hierarchical encoding system includes a super-resolution processing device 410, i.e., a second super-resolution enlarger, that operates on decoded

pictures from the basic layer encoder to generate super-resolution enlarged decoded pictures with a resolution higher than the standard resolution as reference pictures for the enhancement layer encoding device 412. APPLE-1005, [0014], FIG. 4A. For example, Phek explains that “[t]he 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 “[a] super-resolution processing device 410 is used to improve these additional reference images.” *Id.*, [0014]; *see also id.*, [0011] (“The present invention uses a super-resolution processing device 320 to improve the quality and resolution of reference frames used by the motion detection processing device 302 and motion compensation processing device 318.”), FIGS. 3A-3B, 4B.

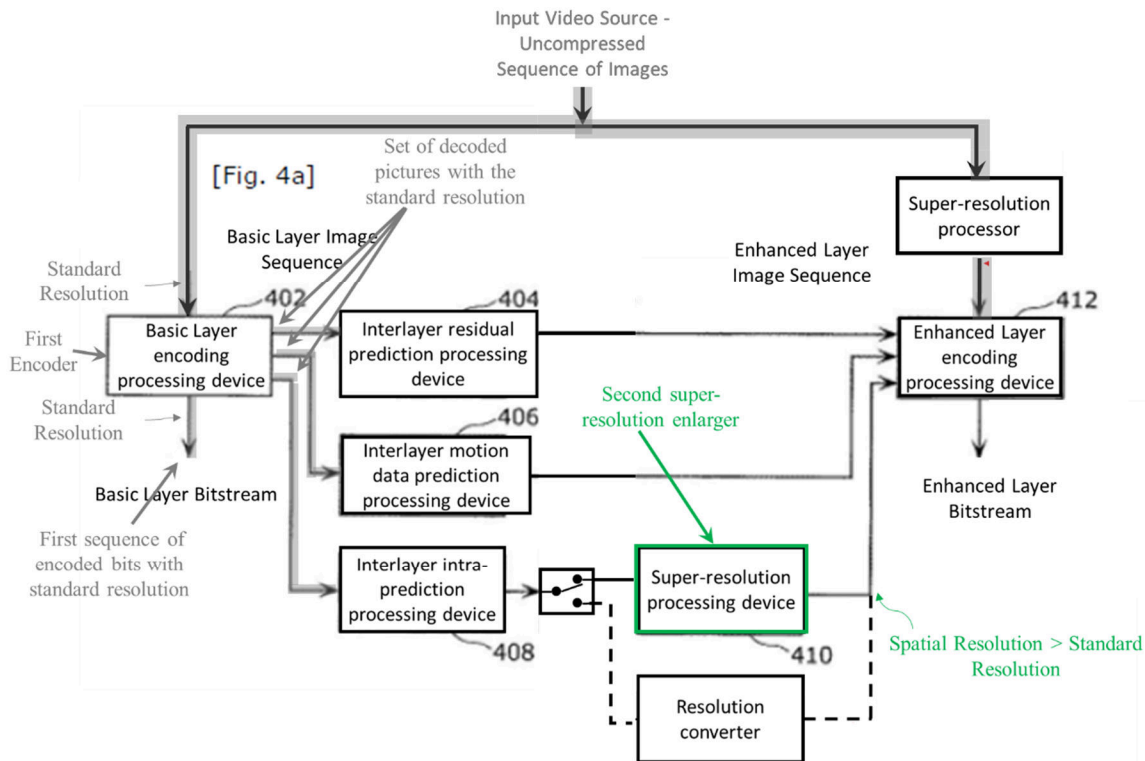


Illustration of Proposed Combination (Phek-Segall-He)

[1.4]

Phek-Segall-He renders obvious [1.4]. APPLE-1003, ¶¶72, 75. Phek discloses a super-resolution processing device 410 that enhances decoded basic layer images through advanced algorithms, although the system does not restrict or limit resolution conversion to super-resolution methods alone. APPLE-1005, [0014]. As discussed above (supra §III.A.4.(b)), a POSITA would have recognized that conventional resolution conversion techniques, such as interpolation-based up-sampling as taught in Segall provides a predictable

alternative to super-resolution processing for increasing the spatial resolution of decoded pictures.

Segall teaches a spatial scalability framework that uses conventional resolution conversion for generating reference pictures for inter-layer prediction. APPLE-1006, §II. Segall explains that “inter-layer prediction of various data” is obtained by up-sampling base layer images, stating that “for the case of intra-picture coded blocks in the base layer, reconstructed intensities provide a prediction for the enhancement layer.” APPLE-1006, §II. This process involves up-sampling decoded base layer images to match the enhancement layer resolution through standard interpolation methods, as Segall’s system operates on “multiple display resolutions within a single compressed bit stream.” APPLE-1006, §I. For example, Segall discloses that “[t]he up-sampling operation for the luma component consists of applying a separable four-tap poly-phase interpolation filter” while “[t]he chroma component of the signal is also up-sampled but with a different (simpler) interpolating kernel.” APPLE-1006, §III.C; *see also id.*, §§III.A (describing detailed techniques for up-sampling and calculation of corresponding spatial positions between base and enhancement layer images), III.C (describing details of up-sampling decoded samples of the lower layer), Table 1 (“Interpolation Filter for Luma Up-Sampling”), Table 2 (“INTERPOLATION FILTER FOR CHROMA UP-SAMPLING”), FIG. 1. Further, as discussed, a POSITA would

have understood Segall's teachings to exemplify a class of well-known resolution conversion techniques that increase image resolution by interpolating additional pixels in an image from other pixels within the same image, without the need for working on a series of images as required by super-resolution enlargement. *See, e.g.,* APPLE-1012, 22-23; APPLE-1074, 4:53-60; APPLE-1024, 6:8-43, FIG. 2; APPLE-1026, 25:27-37, FIG. 1.

For the reasons detailed above, a POSITA would have found it obvious to apply Segall's resolution conversion methods to Phek's system thereby providing another pathway for enlarging reference pictures for use by the enhancement layer coding device 412. APPLE-1005, [0014]; APPLE-1006, §II; *supra*, §III.4.(b)-(c). For example, in the combination, Phek's interlayer intra-prediction processing device 408 routes decoded basic layer images to enhancement layer components according to a switching signal based on He, and a POSITA would have logically implemented Segall's upsampling in this path to produce "resolution converted enlarged decoded pictures" through conventional interpolation. APPLE-1006, §II. Through this configuration, a third resolution converter is provided that acquires the set of decoded pictures from the first encoder (e.g., basic layer encoding processing device 402) to implement thereon a process for a resolution conversion to create a set of resolution converted enlarged decoded pictures with a resolution higher than the standard resolution. Segall's resolution conversion allows decoded

base layer images to be upsampled to the enhancement layer resolution through
“methods for inter-layer prediction of various data.” APPLE-1006, §§II-III.

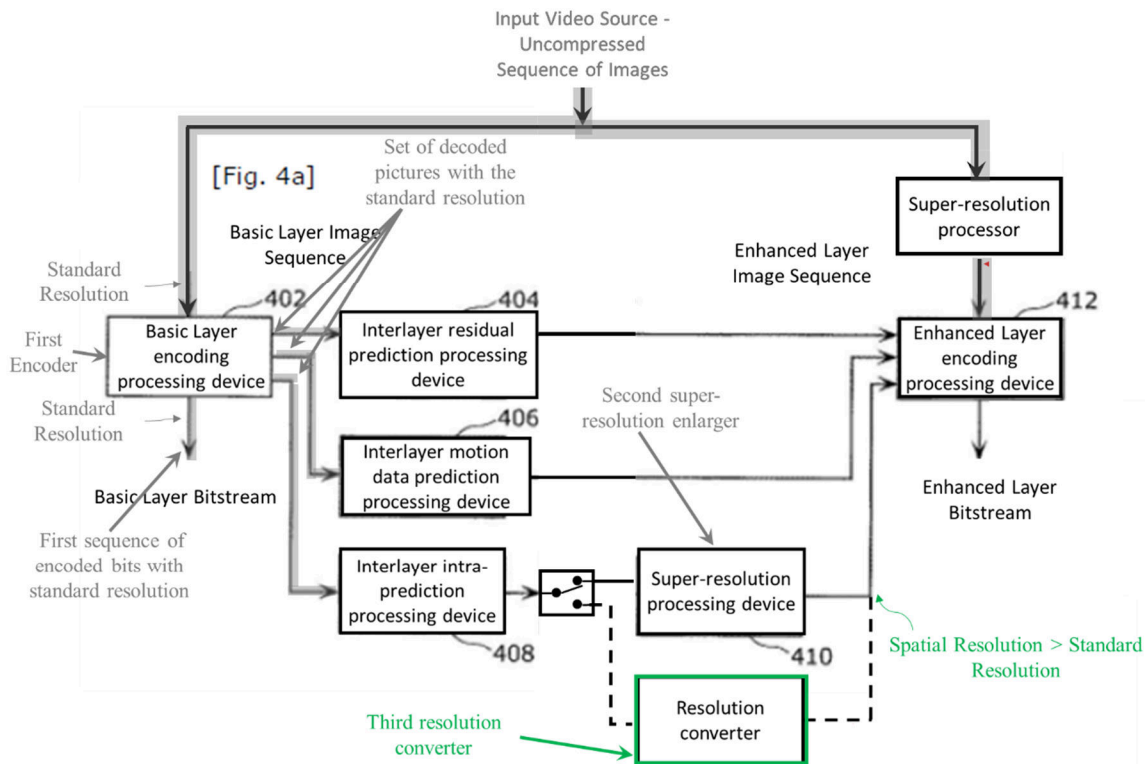


Illustration of Proposed Combination (Phek-Segall-He)

[1.5]

Phek-Segall-He renders obvious [1.5]. APPLE-1003, ¶¶72, 76-78. Phek’s hierarchical encoding system includes an enhanced layer encoding processing device 412 that corresponds to the claimed “third encoder.” *Id.* In Phek’s system, enhanced layer encoding processing device 412 obtains an enhanced layer image sequence as a set of encoding target pictures and uses a super-resolution enlarged version of a decoded base layer picture from super-resolution processing device

410 as a reference picture for predicting a target picture in the enhancement layer. APPLE-1005, [0014] (“An enhanced layer image sequence is encoded using an enhanced layer encoding processing device 412. The enhanced layer encoding processing device 412 uses 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. ... The 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. A super-resolution processing device 410 is used to improve these additional reference images.”), FIG. 4A; *see also id.*, [0011], FIG. 3A. The enhanced layer encoding processing device 412 implements a combination of processes for a prediction and an encoding to create an enhanced layer bitstream (e.g., a third sequence of encoded bits). APPLE-1005, [0011], [0014], FIG. 4A.

As discussed in detail above, it would have been obvious to generate the enhanced layer image sequence that forms the set of encoding target pictures for the enhanced layer encoding processing device 412 by super-resolution enlarging the base layer image sequence. *Supra*, §III.A.4.(a). In this case, the enhanced layer encoding processing device 412 (third encoder) would be configured to have

the set of super-resolution enlarged pictures from the first super-resolution enlarger as the set of encoding target pictures.

As further discussed in detail above, it would have been obvious based on the teachings of Segall and He for the enhanced layer encoding processing device 412 (third encoder) to use a set of super-resolution enlarged decoded pictures from the super-resolution processing device 410 (second super-resolution enlarger) and the set of resolution converted enlarged decoded pictures from the Phek-based super-resolution converter as alternate sets of reference pictures depending on the content characteristics of the decoded base layer picture, processing demands, and/or other information that affect a switching signal for choosing between the different types of reference pictures based on He. *Supra*, §III.A.4.(b)-(c).

For example, Segall's spatial scalability framework provides the second set of reference pictures through conventional upsampling by interpolating luma and chroma information from a single image. APPLE-1006, §II. Segall explicitly describes "inter-layer prediction of various data" using upsampled base layer images, stating that "reconstructed intensities provide a prediction for the enhancement layer" when upsampled to match the enhancement layer resolution. APPLE-1006, §II. A POSITA would have recognized that this conventional resolution conversion represents a predictable alternative to super-resolution processing for generating reference pictures. APPLE-1006, §II. Further, based on

He's teachings, the enhanced layer encoding processing device 412 selects among these reference picture sets using Phek's inter-layer prediction mechanisms.

APPLE-1008, [0002]-[0006], [0022], [0028], [0032]-[0037], FIGS. 1-4; APPLE-1005, [0014]. In the combination, Phek's enhanced layer encoder "utilizes interlayer prediction to reduce redundancy between the enhanced layer and the basic layer," employing both super-resolution and conventionally upsampled reference pictures as part of this process. APPLE-1005, [0014]. Segall reinforces this by teaching that "inter-layer prediction can be used to provide additional coding choices," including the use of multiple reference types. APPLE-1006, §II; *supra*, [1.3]-[1.4].

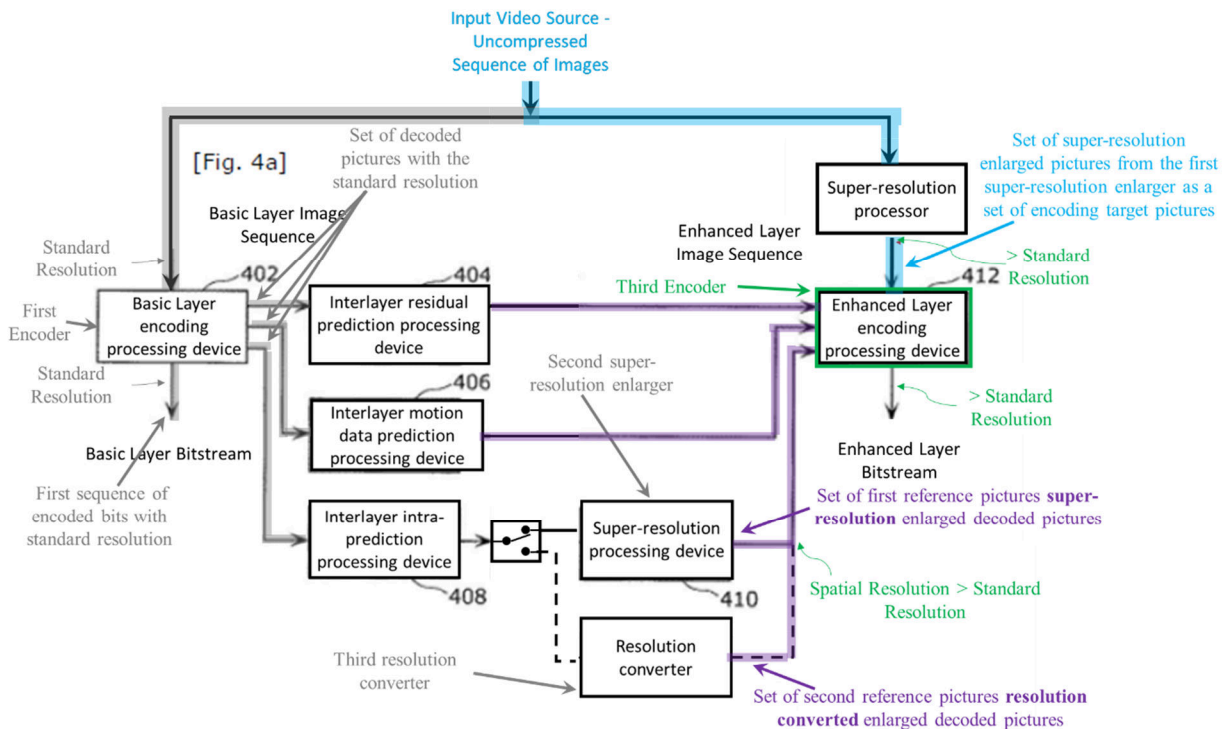


Illustration of Proposed Combination (Phek-Segall-He)

[1.6]

Phek-Segall-He renders obvious [1.6]. APPLE-1003, ¶¶72-78. A POSITA would have understood that [1.6] merely reflects the conventional spatial scalable coding practice of upscaling decoded lower-layer images to the *same* resolution of upper layer images for when the decoded lower-layer image is to be used as a reference picture for interlayer prediction. *See also* APPLE-1006, §II (Segall describing baseline profile in which “the width and height of the scaled regions of lower resolution and higher resolution pictures must have the same scaling ratio”); APPLE-1026, 25:27-37 (“the resolution converter 205 converts the resolution of the base layer image signal to the resolution corresponding to the enhancement image signal”); APPLE-1046, [0128] (“a resolution conversion circuit 3 for subjecting regenerated data ... to resolution conversion in a way that its resolution becomes equal to that of the high-resolution image signal”); APPLE-1024, 9:30-35 (“the encoder sets the resolution for the reference frame (630) before setting the resolution for the current frame (620)”), 14:40-46 (“the one or more spatial resolution scaling factors are constrained to be identical to the one or more spatial resolution scaling factors for the reference”), FIG. 6; APPLE-1077, (describing an inter-layer “transform circuit 13” that “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.”).

Consistent with this conventional practice, it would have been obvious and a POSITA would have naturally implemented the encoding system in the Phek-Segall-He combination such that the reference pictures and encoding target pictures processed by the enhancement layer encoding processing device 412 (e.g., third encoder) would have the same spatial resolution. A POSITA would have set the resolutions to be equal as recited in [1.6] to simplify and facilitate the prediction of enhancement layer pictures from the set of super-resolution enlarged decoded pictures and the resolution converted decoded pictures that are used as references for the prediction. Indeed, Segall explicitly teaches that “corresponding positions in adjacent levels of [a] pyramid hierarchy” in a scalable video coding system should be calculated so that “sample locations in a lower resolution layer” can be identified with respect to “correspond[ing] [] sample locations in the enhancement layer.” APPLE-1006, §III.A. By enlarging reference pictures to the same resolution as the images they are being used to predict in the enhancement layer, each sample location in the enhancement layer picture would have a corresponding sample location in the upscaled base layer picture used as reference, an appropriate interlayer intra- or inter-prediction coding technique can be performed. A POSITA would have reasonably expected success setting the resolution of the recited sets of pictures to be equal in the Phek-Segall-He combination since it had been common practice for long before the Critical Date of

the '448 Patent to set the resolution of the encoding target and its reference picture(s) to be the same.

[1.7]

Phek-Segall-He renders obvious [1.7]. APPLE-1003, ¶¶72-78. At the outset, for instance, Phek explains that the enhanced layer coding device 412 controls reference picture selection by “utiliz[ing] interlayer prediction to reduce redundancy between the enhanced layer and the basic layer.” APPLE-1005, [0014]. Phek further teaches that “the 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,” which indicates that Phek’s system makes a selection among multiple reference pictures for encoding enhancement layer images. APPLE-1005, [0014] (emphasis added).

As discussed (§§III.A.3, III.A.4.(c)), He’s teachings are also relevant, and the application of He’s teachings in combination with Phek and Segall render obvious [1.7]. He describes techniques for selectively enabling/disabling reference picture filters in a scalable coding system. APPLE-1008, [0028], [0034], [0037], [0069]; *supra*, III.A.3, FIGS. 1-4. He also describes control flags and other signaling that indicates whether enhanced (filtered) reference pictures are to be used for the prediction of a current frame. *Id.* A POSITA would have recognized that He’s signaling directly applies to reference picture selection, enabling the third

encoder to create a set of data on the selection of the set of reference pictures through analogous signaling. APPLE-1008, [0034] (“To ensure equivalency of data between the encoder 100 and the decoder 200, the video encoder asserts control bits in the bitstream BTS received by the video decoder 200 to control the activation state of the [] filter ...”), [0047]-[0048], [0052], [0060]-[0061], [0069], FIGS. 5, 7-8. In the combination, and based on He’s teachings, the enhanced layer processing device 412, would control a selection of a set of reference pictures (e.g., the super-resolution enlarged or resolution converted reference pictures) and would create a set of data to identify the selected set of reference pictures for prediction of enhancement layer pictures of the third combination as claimed. *Id.* Based on He, the set of data can indicate whether the enhancement filter is to be applied and thus whether super-resolution enlargement should be applied to enhance the quality of the reference picture or whether conventional resolution conversion based on interpolation of information in a single image should be used to upscale the reference picture without further enhancement. *Supra*, §III.A.4.(c). For example, in some implementations, a “reference_selection_flag” in the header of a coded picture signal can indicate whether the prediction uses super-resolution enlarged decoded pictures (flag=0) or resolution-converted enlarged decoded pictures (flag=1). This flag “identifies the selected set of reference pictures” during prediction. APPLE-1008, [0008], [0055], [0063], [0068]-[0069]. This would have

been obvious and beneficial to ensure the decoder and any downstream coding devices make the same selection of reference pictures during the downstream coding process as was originally used to encode the video, thereby ensuring that the downstream coding process accurately codes/decodes the video consistent with the original encoding.

Further, Segall's spatial scalability architecture reinforces this obviousness of [1.7] by enforcing alignment of reference picture resolutions across layers. APPLE-1006, §II. Segall teaches that "inter-layer prediction can be used to provide additional coding choices," including selecting between upsampled base-layer references and enhanced-layer predictions. APPLE-1006, §II. Based on Segall's teachings, the system "[calculates] corresponding positions in different resolution layers" and "maintain flexible cropping, scaling, and alignment relationships," involving explicit identification of selected reference pictures during prediction. APPLE-1006, §§IIb (Inter-Layer Spatial Relationships and Profile Constraints), III (Coding Tools).

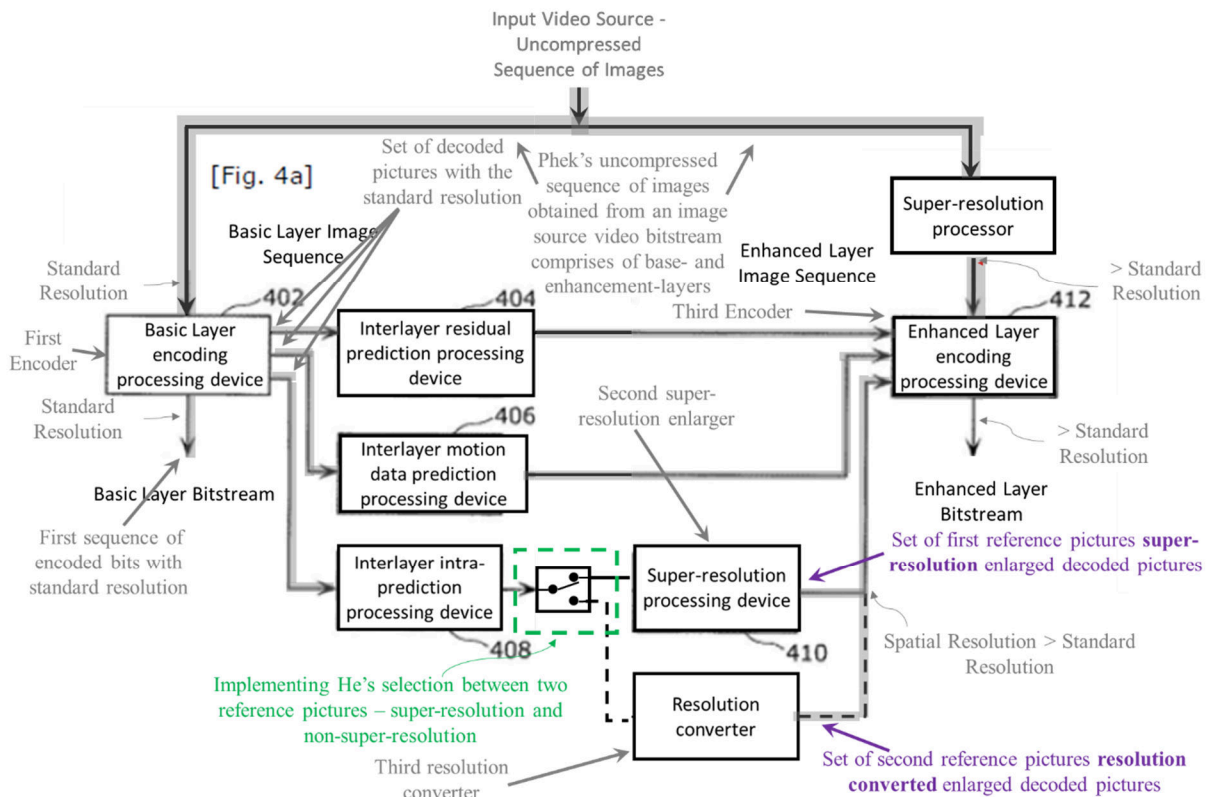


Illustration of Proposed Combination (Phek-Segall-He)

(b) Claim 2

[2.P]

Supra §III.A.5.[1.P]. APPLE-1003, ¶¶72-78.

[2.1]

Supra §III.A.5.[1.1]. APPLE-1003, ¶¶72-78.

[2.2]

Supra §III.A.5.[1.2]. APPLE-1003, ¶¶72-78.

[2.3]

Supra §III.A.5.[1.4]. APPLE-1003, ¶¶72-78.

[2.4]

Supra §III.A.5.[1.5]. APPLE-1003, ¶¶72-78.

[2.5]

Supra §III.A.5.[1.6]. APPLE-1003, ¶¶72-78.

[2.6]

Supra §III.A.5.[1.7]. APPLE-1003, ¶¶72-78.

(c) Claim 3

[3.P]

Supra §III.A.5.[1.P]. APPLE-1003, ¶¶72-78.

[3.1]

Supra §III.A.5.[1.1]. APPLE-1003, ¶¶72-78.

[3.2]

Supra §III.A.5.[1.2]. APPLE-1003, ¶¶72-78.

[3.3]

Supra §III.A.5.[1.3]. APPLE-1003, ¶¶72-78.

[3.4]

Supra §III.A.5.[1.5]. APPLE-1003, ¶¶72-78.

[3.5]

Supra §III.A.5.[1.6]. APPLE-1003, ¶¶72-78.

[3.6]

Supra §III.A.5.[1.7]. APPLE-1003, ¶¶72-78.

**B. [GROUND 1B] – Phek in view of Segall, He, and Takahashi
Renders Claims 1-3 Obvious**

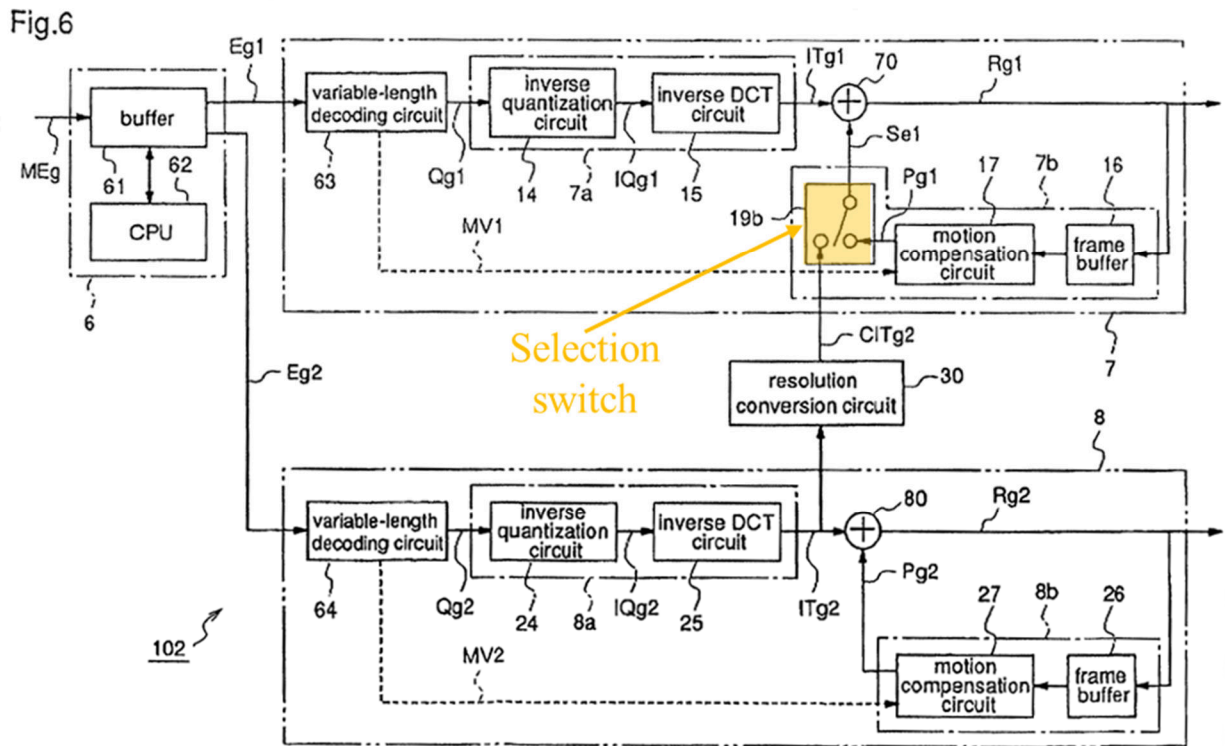
1. Takahashi Overview

Like the '448 Patent, Takahashi discloses a hierarchical image coding system configured to code digital image signals across multiple resolutions. APPLE-1046, [0001], Abstract. Through hierarchical coding, Takahashi enables simultaneous transmission of high-definition and standard-definition coded video data within a single bitstream, thereby supporting resolution-scalable transmission and decoding. APPLE-1046, [0004]. APPLE-1003, ¶79. Although Takahashi describes the switching mechanism in the context of decoding, a POSITA would have understood that the same switching principles are equally applicable to the encoding process. APPLE-1003, ¶79. Takahashi's teachings regarding selection between reference picture sets based on resolution and coding efficiency are not limited to the decoding architecture and would have been readily recognized by a POSITA as symmetrical and transferable to the encoder-side operations to maintain consistency and optimize predictive coding performance. APPLE-1003, ¶79.

Takahashi's coding device includes a selection switch 19b within a first decoding unit 7 that selects between different reference picture signals based on criteria bearing on compression and computational efficiency, such as the coding mode of the frame. APPLE-1046, [0131], [0176], Fig. 6. For example, in some embodiments, for I-frames, the switch selects resolution-converted data from a second decoding unit 8 as a reference picture signal, whereas for P- or B-frames,

the switch selects high-resolution reference picture data from decoding unit 7.

APPLE-1046, [0137]-[0140]. As illustrated in FIG. 6, selection switch 19b selects either regenerated resolution-converted data (CITg2) or predicted data (Pg1), depending on the coding mode. APPLE-1046, [0176], [0251]. APPLE-1003, ¶80.



Takahashi, FIG. 6 (annotated)

Takahashi explains that the pixel (or information) densities of high- and low-resolution reference pictures differ significantly, thereby warranting selection of different reference pictures to promote efficient processing. APPLE-1046, [0139]-[0140]. The implementation of reference picture selection across multiple embodiments – through components such as switches 19a, 19b, 29a, and 29b, as

illustrated in FIGs. 1, 6, 12, and 15, respectively – demonstrates that reference picture selection was a foundational concept in Takahashi’s hierarchical coding system. APPLE-1046, [0131]-[0137], [0176], [0229]-[0235], [0253], [0262]; APPLE-1003, ¶81.

2. Combination of Phek, Segall, He, and Takahashi

As detailed above (*supra* §§III.A.1, III.A.4.(a)-(c)), Phek describes a hierarchical video coding system with super-resolution processing, while Segall and He describe similar scalable coding techniques that use interlayer prediction to reduce coding redundancies between a base layer and one or more enhancement layers. *Supra* III.A.1-4, III.A.5.[1.1]-5.[1.5]. APPLE-1003, ¶82. Takahashi describes analogous techniques and offers a complementary solution designed to reduce “the amount of operational processing” while maintaining efficient coding through “simple circuit configuration.” APPLE-1046, [0001], [0039], [0042]-[0045]. APPLE-1003, ¶82. For example, Takahashi discloses an enhancement layer coding unit that selects between resolution-converted data from a lower layer or predicted data from the enhancement layer for use as a reference signal for the prediction of a frame in the enhancement layer based on a coding mode (e.g., whether the frame is an intra-coded I-frame or an inter-predicted B- or P-frame). APPLE-1046, [0131], [0139]-[0142], [0176], [0253], [0262], FIG. 6; *see also id.*, [0131]-[0137], [0176], [0189], [0229]-[0235], [0253], [0262]. A POSITA would

have found it obvious to implement the hierarchical coder of the Phek-Segall-He combination further in accordance with Takahashi's teaching to select between resolution-converted reference data or enhancement layer reference data based on coding mode. APPLE-1003, ¶82. Multiple reasons would have prompted a POSITA to implement a Phek-Segall-He-Takahashi combination in this manner. *Id.*

First, for example, Takahashi's reference picture selection framework provides improved efficiency through simplified architecture. By implementing the selective reference signal approach, "the configuration ... becomes simpler" because resolution conversion or super-resolution enlargement of decoded base layer pictures will be unnecessary where the coding mode indicates that an enhancement layer picture can be used as a reference. APPLE-1046, [0042], [0141], [0219], [0282]-[0283]; APPLE-1003, ¶83.

Second, the implementation reduces computational complexity by providing a clear mechanism to choose between different reference signals, eliminating unnecessary processing operations during the decoding process. APPLE-1046, [0190]-[0192]. *See also*, APPLE-1046, [0001], [0039]-[0045], [0055], [0063]-[0066], [0192], [0204], [0213]-[0219], [0266]; APPLE-1003, ¶84.

Third, Takahashi's reference picture selection capability improves memory bandwidth management—a critical consideration in video coding systems. By

streamlining the reference selection process, Takahashi's approach reduces unnecessary memory accesses, which would be particularly valuable in the context of the Phek-Segall combination. APPLE-1046, [0191]-[0192], [0213]-[0219]; APPLE-1003, ¶85.

Fourth, Takahashi's reference picture selection capability improves processing and compression efficiency. For example, Takahashi explains that "since an image signal corresponding to the I frame is not subjected to temporal prediction, for the I frame, predictive coding in which the low-resolution regenerated data corresponding to the target processing signal is used as predicted data is more efficient than the processing in which the target processing signal of the I frame is coded as it is. However, for other frames, even when the image signal having a different resolution is not used as predicted data, the compression efficiency is not reduced much compared with the case that all image signals are subjected to predictive coding." APPLE-1046, [0140]; *see also id.* [0138]-[0142]; APPLE-1003, ¶86.

Fifth, a POSITA would have recognized that implementing Takahashi's selection switch simplifies hardware requirements while maintaining decompression efficiency. Takahashi expressly teaches the "reduc[tion] of [the] amount of operation and simplifying the configuration of the hierarchical decoding apparatus." APPLE-1046, [0047]. APPLE-1003, ¶87.

Sixth, a POSITA would have found it obvious to apply Takahashi's suggestion for selecting reference pictures based on coding mode because doing so merely entails the predictable use of a known technique as taught by Takahashi to a similar scalable coding system as taught by Phek to achieve predictable results. *KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 417 (2007) (a supposed invention that "simply arranges old elements with each performing the same function it had been known to perform" is obvious). "It's enough... to show that there was a known problem... in the art, that [another known teaching] ... helped address that issue, and that combining the teachings of ... wasn't beyond the skill of an ordinary artisan. Nothing more is required to show a motivation to combine under *KSR*." *Intel Corp. v. PACT XPP Schweiz AG*, 61 F.4th 1373, 1380-81, (Fed. Cir. 2023). Here, Takahashi expressly describes how to perform reference picture selection in a scalable coding system based on coding mode and the implementation of such techniques was well within the capability of a POSITA. APPLE-1003, ¶88. A POSITA would have reasonably expected success, especially since reference picture selection techniques like those described in Takahashi were commonly implemented before the Critical Date. *Id.*

3. Analysis

(a) Claim 1

[1.P]-[1.6]

Phek-Segall-He-Takahashi renders obvious [1.P]-[1.6] based on the teachings of Phek, Segall, and Takahashi for the reasons described above in the analysis of the corresponding claim elements in Ground 1A. *Supra* §§III.A.5.[1.P]-[1.6]. APPLE-1003, ¶¶72, 89.

[1.7]

Phek-Segall-He-Takahashi renders obvious [1.7]. APPLE-1003, ¶¶72, 89. As discussed in further detail above (*supra*, §III.B.1-2), Takahashi discloses an enhancement layer coding unit that selects between resolution-converted data from a lower layer or predicted data from the enhancement layer for use as a reference signal for the prediction of a frame in the enhancement layer based on a coding mode (e.g., whether the frame is an intra-coded I-frame or an inter-predicted B- or P-frame). APPLE-1046, [0131], [0139]-[0142], [0176], [0253], [0262], FIG. 6; *see also id.*, [0131]-[0137], [0176], [0189], [0229]-[0235], [0253], [0262]. A POSITA would have found it obvious to implement a Phek-Segall-He-Takahashi combination with an enhancement layer coding device that likewise makes a selection of a set of reference pictures by selecting between resolution-converted reference data from the base layer or other reference data from the enhancement

layer based on the coding mode of a frame (e.g., whether the frame is intra- or inter-predicted).⁶ *Supra*, §III.B.2.

Takahashi and He also both describes the known option of creating a set of data on the selection of the set of reference pictures to identify a selected set of reference pictures during the encoding and decoding processes . *See*, e.g., APPLE-1046, [0265]-[0267] (“control signal”; “identification flag”), [0176]; APPLE-1008, [0034] (“control bits in the bitstream BTS received by the video decoder 200 to control the activation state”), [0008], [0055], [0063], [0068]-[0069] (“reference_selection_flag”). In the combination, it likewise would have been obvious to create such selection data on the selection of the set of reference pictures to ensure encoding is implemented deterministically while optimizing coding and computational efficiency. This would allow the decoder and any downstream coding devices to make the same selection of reference pictures

⁶ Petitioner notes that the reference picture selection techniques made in the Phek-Segall-He-Takahashi combination based on coding mode according to Takahashi’s teachings provide an independent basis for meeting the corresponding features recited in [1.7]. Thus, [1.7] is rendered obvious based on the integration of Takahashi’s teachings, even apart from He’s teachings of selective reference picture enhancement.

Supra §III.B.3.[1.2]. APPLE-1003, ¶89.

[2.3]

Supra §III.B.3.[1.4]. APPLE-1003, ¶89.

[2.4]

Supra §III.B.3.[1.5]. APPLE-1003, ¶89.

[2.5]

Supra §III.B.3.[1.6]. APPLE-1003, ¶89.

[2.6]

Supra §III.B.3.[1.7]. APPLE-1003, ¶89.

(c) Claim 3

[3.P]

Supra §III.B.3.[1.P]. APPLE-1003, ¶89.

[3.1]

Supra §III.B.3.[1.1]. APPLE-1003, ¶89.

[3.2]

Supra §III.B.3.[1.2]. APPLE-1003, ¶89.

[3.3]

Supra §III.B.3.[1.3]. APPLE-1003, ¶89.

[3.4]

Supra §III.B.3.[1.5]. APPLE-1003, ¶89.

[3.5]

Supra §III.B.3.[1.6]. APPLE-1003, ¶89.

[3.6]

Supra §III.B.3.[1.7]. APPLE-1003, ¶89.

**C. [GROUND 1C] – Phek in view of Segall and Takahashi
Renders Claims 1-3 Obvious**

Grounds 1A and 1B rely on He’s teachings directed to selective application of reference picture enhancements in a scalable coding system. APPLE-1003, ¶¶72, 89-90. In the context of the combinations addressed above, He’s teachings as applied to the Phek-Segall system would have rendered obvious the features recited in [1.7], [2.6], and [3.6]. However, as further described in Ground 1B, the application of Takahashi’s teachings related to reference picture selection based on coding mode to the Phek-Segall system provides an additional and independent basis for rendering obvious [1.7], [2.6], and [3.6]. Accordingly, even if He’s teachings were not considered, a Phek-Segall-Takahashi combination would have rendered obvious all claims 1-3 of the ’448 Patent and their respective limitations for the reasons addressed above in Grounds 1A-1B without reference to He’s teachings. Additionally, since neither claims 2 nor 3 require *both* a second super-resolution enlarger *and* a third resolution converter (claim 2 recites only the latter and claim 3 the former), a system with only *one* of these devices in the Phek-Segall-Takahashi combination would suffice to address claims 2 and 3, and He’s selection capability would be inapplicable when only one of the second super-resolution enlarger and the third resolution converter is implemented.

D. [GROUND 1D] – Phek in view of Takahashi Renders Claim 3 Obvious

Grounds 1A and 1B rely on He's teachings directed to selective application of reference picture enhancements in a scalable coding system. APPLE-1003, ¶91. In the context of the combinations addressed above, He's teachings as applied to the Phek-Segall system would have rendered obvious the features recited in [3.6]. However, as further described in Ground 1B, the application of Takahashi's teachings related to reference picture selection based on coding mode to Phek's system provides an additional and independent basis for rendering obvious [3.6]. Additionally, since claim 3 does not require *both* a second super-resolution enlarger *and* a third resolution converter, and instead requires only a second super-resolution enlarger (regardless of whether the system further includes a third resolution converter), a system with only a second super-resolution enlarger suffices to meet claim 3, and neither Segall's nor He's teachings are necessary to provide a third resolution converter or a selection capability between the output of a second super-resolution enlarger and a third resolution converter. Accordingly, even if He's and Segall's teachings were not considered, a Phek-Takahashi combination alone would have rendered obvious claim 3 and each of its limitations for the reasons addressed above in Grounds 1A and 1B with reference to the teachings of Phek and Takahashi alone.

IV. DISCRETIONARY DENIAL IS NOT WARRANTED

Petitioner believes that discretionary denial is unwarranted, and yet, Petitioner intends to utilize the bifurcated briefing process contemplated by the March 26, 2025, Stewart Memorandum to rebut contentions if offered by Patent Owner to the contrary. See <https://www.uspto.gov/about-us/events/learnaboutnew-interim-processes-relating-institution-aia-proceedings> (April 24, 2025); <https://uspto.cosocloud.com/p9xpz6s4jn75> (Video).

V. PAYMENT OF FEES – 37 C.F.R. § 42.103

Petitioner authorizes the Patent and Trademark Office to charge Deposit Account No. 06-1050.

VI. CONCLUSION

The Challenged Claims are unpatentable. Petitioner submits that trial should be instituted and the Challenged Claims found unpatentable based on Grounds 1A-1D.

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

A. Real Party-In-Interest Under 37 C.F.R. § 42.8(b)(1)

Petitioner, Apple Inc. is the real party-in-interest.

B. Related Matters Under 37 C.F.R. § 42.8(b)(2)

Petitioner is not aware of any disclaimers, reexamination certificates or petitions for inter partes review for the '448 patent. The '448 patent is the subject

of civil actions including: *Advanced Coding Technologies LLC v. Apple Inc.*, 2-24-cv-00572 (EDTX), filed July 22, 2024, and *Advanced Coding Technologies LLC v. Google LLC*, 2-24-cv-00353 (EDTX), filed May 10, 2024.

Apple has also filed an IPR petition on U.S. Patent No. 10,218,995, which is a divisional of the '448 Patent, in IPR2025-00984. Google filed an IPR petition on the '995 Patent in IPR2025-01000.

C. Lead And Back-Up Counsel Under 37 C.F.R. § 42.8(b)(3)

Petitioner provides the following designation of counsel.

Lead Counsel	Backup counsel
W. Karl Renner, Reg. No. 41,265 Fish & Richardson P.C. 60 South Sixth Street Suite 3200 Minneapolis, MN 55402 Tel: 202-783-5070 Fax: 877-769-7945 Email: IPR50095-0244IP1@fr.com	Jeremy J. Monaldo, Reg. No. 58,680 Nicholas Stephens, Reg. No. 74,320 Fish & Richardson P.C. 60 South Sixth Street, Suite 3200 Minneapolis, MN 55402 Tel: 202-783-5070 Fax: 877-769-7945 Email: IPR50095-0244IP1@fr.com

D. Service Information

Please address all correspondence and service to the address listed above.

Petitioner consents to electronic service by email at IPR50095-0244IP1@fr.com.

Respectfully submitted,

Dated 06/02/2025

/Nicholas W. Stephens/

W. Karl Renner, Reg. No. 41,265

Jeremy J. Monaldo, Reg. No. 58,680

Nicholas Stephens, Reg. No. 74,320

Fish & Richardson P.C.

60 South Sixth Street, Suite 3200

Minneapolis, MN 55402

T: 202-783-5070

F: 877-769-7945

(Control No. IPR2025-01070)

Attorneys for Petitioner

CERTIFICATION UNDER 37 CFR § 42.24

Under the provisions of 37 CFR § 42.24(d), the undersigned hereby certifies that the word count for the foregoing Petition for *Inter partes* Review totals 13,946 words, which is less than the 14,000 allowed under 37 CFR § 42.24.

Dated 06/02/2025_____

/Nicholas W. Stephens/_____

W. Karl Renner, Reg. No. 41,265
Jeremy J. Monaldo, Reg. No. 58,680
Nicholas Stephens, Reg. No. 74,320
Fish & Richardson P.C.
60 South Sixth Street, Suite 3200
Minneapolis, MN 55402
T: 202-783-5070
F: 877-769-7945

Attorneys for Petitioner

CERTIFICATE OF SERVICE

Pursuant to 37 CFR §§ 42.6(e)(4)(i) *et seq.* and 42.105(b), the undersigned certifies that on June 2, 2025, a complete and entire copy of this Petition for *Inter partes* Review, Power of Attorney, and all supporting exhibits were provided via Federal Express, to the Patent Owner by serving the correspondence address of record as follows:

JOSHUA GOLDBERG
Nath, Goldberg & Meyer
112 S. West Street
Alexandria, VA 22314
UNITED STATES
(703) 548-6284

/Crena Pacheco/

Crena Pacheco
Fish & Richardson P.C.
60 South Sixth Street, Suite 3200
Minneapolis, MN 55402
pacheco@fr.com