

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent of: Hideki Takehara, et al.
U.S. Patent No.: 9,986,303 Attorney Docket No.: 50095-0243IP1
Issue Date: May 29, 2018
Appl. Serial No.: 14/657,359
Filing Date: March 13, 2015
Title: VIDEO IMAGE CODING DATA TRANSMITTER, VIDEO
IMAGE CODING DATA TRANSMISSION METHOD, VIDEO
IMAGE CODING DATA RECEIVER, AND VIDEO IMAGE
CODING DATA TRANSMISSION AND RECEPTION SYSTEM

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. 9,986,303
PURSUANT TO 35 U.S.C. §§ 311–319, 37 C.F.R. § 42

TABLE OF CONTENTS

I.	REQUIREMENTS FOR IPR	1
A.	Grounds for Standing.....	1
B.	Challenge and Relief Requested.....	1
C.	Claim Construction	2
D.	Level of Ordinary Skill in the Art.....	4
II.	THE '303 PATENT.....	5
A.	Brief Description.....	5
B.	Examination History	7
III.	THE CHALLENGED CLAIMS ARE UNPATENTABLE.....	8
A.	GROUND 1A – DEMIRCIN RENDERS CLAIMS 1-2 OBVIOUS.....	8
1.	Overview of Demircin.....	8
2.	Analysis	16
B.	GROUND 1B – DEMIRCIN IN VIEW OF KIMOTO RENDERS CLAIMS 1-2 OBVIOUS	46
C.	GROUND 2 – KIMOTO IN VIEW OF DEMIRCIN RENDERS CLAIM 1 OBVIOUS.....	51
1.	Overview of Kimoto.....	51
2.	Combination of Kimoto and Demircin	52
3.	Reasons to Combine Kimoto and Demircin.....	53
4.	Analysis	59
IV.	PTAB DISCRETION SHOULD NOT PRECLUDE INSTITUTION.....	74
V.	CONCLUSION AND FEES	74
VI.	MANDATORY NOTICES UNDER 37 C.F.R § 42.8(a)(1).....	74
A.	Real Party-In-Interest Under 37 C.F.R. § 42.8(b)(1).....	74
B.	Related Matters Under 37 C.F.R. § 42.8(b)(2).....	74
C.	Lead And Back-Up Counsel Under 37 C.F.R. § 42.8(b)(3).....	75
D.	Service Information	76

LIST OF EXHIBITS

APPLE-1001	U.S. Patent No. 9,986,303 to Takehara, et al. (“the ’303 Patent”)
APPLE-1002	Excerpts from the Prosecution History of the ’303 Patent (“the ’303 Patent File History”)
APPLE-1003	Expert Declaration of Dr. Sheila Hemami, Ph.D.
APPLE-1004	Complaint, <i>Advanced Coding Techs. LLC v. Apple Inc.</i> , 2:24-cv-00572-JRG, E.D. Tex., filed July 22, 2024
APPLE-1005	U.S. Patent Publication No. 2008/0181302 (“Demircin”)
APPLE-1006	U.S. Patent Publication No. 2007/0153916 (“Demircin916”)
APPLE-1007	U.S. Patent Publication No. 2011/0211631 (“Kimoto”)
APPLE-1008	Excerpts from the Prosecution History of Kimoto (“Kimoto File History”)
APPLE-1009	ITU-T Recommendation H.264 (11/07)
APPLE-1010	U.S. Patent Publication No. 2004/0196972 (“Zhu”)
APPLE-1011	U.S. Patent No. 8,335,252 (“Chen”)
APPLE-1012	U.S. Patent No. 8,175,154 (“Luthra”)
APPLE-1013	U.S. Patent No. 9,094,641 (“Ansley”)
APPLE-1014	Wu Jiyan, Shang Yanlei, Cheng Bo, Wu Budan, and Chen Junliang; “ <i>Loss Tolerant Bandwidth Aggregation for Multihomed Video Streaming over Heterogeneous Wireless Networks</i> ”; <i>Wireless Personal Communications: An International Journal</i> . 75. 1265-1282. 10.1007/s11277-013-1422-3; Published online October 5, 2013 (“Wu”)
APPLE-1015	ISO/IEC 11172-2: 1993 (E) (MPEG1)
APPLE-1016	ISO/IEC 13818-2: 2000 (E) (MPEG2)

APPLE-1017	Infringement Contentions, <i>Advanced Coding Technologies LLC v. Samsung Electronics Co., Ltd. et al.</i> , 2:22-cv-00499-JRG, E.D. Tex., filed May 16, 2023
APPLE-1018	Schwarz et al., <i>Overview of the Scalable Video Coding Extension of the H.264/AVC Standard</i> , IEEE Transactions on Circuits and Systems for Video Technology, v.17, n.9 (September 2007)
APPLE-1019	ITU-T Recommendation H.262 (02/2000)
APPLE-1020	ITU-T Recommendation H.263 (02/1998)
APPLE-1021	Reader, Intraframe and Interframe Adaptive Transform Coding, Efficient Transmission of Pictorial Information, SPIE v.66, 108 (1975)
APPLE-1022	Reader, MPEG4: Coding for Content, Interactivity, and Universal Accessibility, Optical Engineering 35(1), 104 (January 1996)
APPLE-1023	Reader, MPEG Patents, in MPEG Video Compression Standard, Chapter 16, at 357-362 (Mitchell et al ed.) (1996)
APPLE-1024	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-1025	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-1026	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-1027	ISO/IEC 14496-2: 2004 (MPEG4)
APPLE-1028	ISO/IEC 13818-2: 1996 (MPEG2)
APPLE-1029	[RESERVED]

APPLE-1030	[RESERVED]
APPLE-1031	U.S. Patent No. 6,292,512 to Radha et al.
APPLE-1032	U.S. Patent No. 6,697,426 to Van Der Schaar et al.
APPLE-1033	U.S. Patent No. 5,886,736 to Chen
APPLE-1034	Apostolopoulos et al., <i>Video Streaming: Concepts, Algorithms, and Systems</i> , HP Laboratories (Sept. 18, 2002)
APPLE-1035	U.S. Serial No. 11/479,077 to Demircin, Application as Filed for Demircin916 (“Demircin077”)
APPLE-1036	Petition (Paper 2), IPR2024-00327, <i>Samsung Electronics. Co., Ltd. v. Advanced Coding Techs., LLC</i> (Jan. 5, 2024)
APPLE-1037	Patent Owner’s Preliminary Response (Paper 6), IPR2024-00327, <i>Samsung Electronics. Co., Ltd. v. Advanced Coding Techs., LLC</i> (April 22, 2024)
APPLE-1038	Petitioner’s Pre-Institution Reply to Patent Owner’s Preliminary Response (Paper 9), IPR2024-00327, <i>Samsung Electronics. Co., Ltd. v. Advanced Coding Techs., LLC</i> (June 24, 2024)
APPLE-1039	Patent Owner’s Pre-Institution Sur-Reply to Petitioner’s Reply to Patent Owner’s Preliminary Response (Paper 10), IPR2024-00327, <i>Samsung Electronics. Co., Ltd. v. Advanced Coding Techs., LLC</i> (July 1, 2024)
APPLE-1040	Decision Instituting <i>Inter Partes</i> Review (Paper 11), IPR2024-00327, <i>Samsung Electronics. Co., Ltd. v. Advanced Coding Techs., LLC</i> (July 18, 2024)
APPLE-1041	Joint Claim Construction and Prehearing Statement, Dkt. 87, <i>Advanced Coding Techs. LLC v. Samsung Electronics. Co. Ltd.</i> , Case No. 2:22-cv-00499-JRG (E.D. Tex.), filed March 5, 2024
APPLE-1042	U.S. Patent Publication No. 2006/0095944 (“Demircin944”)

- APPLE-1043 Mokhtarian, K, et al., Authentication of Scalable Video Stream with Low Communication Overhead, IEEE Transactions on Multimedia, 12:7, 730 (November 2010)
- APPLE-1044 US Patent Publication No. 2009/0060035 (“He2009”)
- APPLE-1045 P. Amon et al., File format for scalable video coding, IEEE Transactions on Circuits and Systems for Video Technology, Vol 17, No. 9, September 2007, 639 (“Amon”)
- APPLE-1046 Tian et al., Sub-sequence video coding for improved temporal scalability, 2005 IEEE International Symposium on Circuits and Systems (ISCAS) Date of Conference: 23-26 May 2005 (“Tian”)
- APPLE-1047 Iain Richardson, The H.264 Advanced Video Compression Standard, Second Edition, Wiley, 2010 (“Richardson”)
- APPLE-1048 Wien, et al., Real-Time System for Adaptive Video Streaming Based on SVC, IEEE Transactions on Circuits and Systems for Video Technology, VOL. 17, NO. 9, SEPTEMBER 2007, 1227 (“Wien”)
- APPLE-1049 Wiegand et al., Introduction to the Special Issue on Scalable Video Coding—Standardization and Beyond, IEEE Transactions on Circuits and Systems for Video Technology, VOL. 17, NO. 9, SEPTEMBER 2007, 1009 (“Wiegand”)
- APPLE-1050 Tektronix, A Guide to MPEG Fundamentals and Protocol Analysis – Updated with QoS/QoE Diagnostics and Troubleshooting, Primer, *retrieved from* <https://download.tek.com/document/25W-11418-10.pdf>
- APPLE-1051 Wang, Scalable Video Coding, Slides for Video Processing & Communications, Polytechnic Institute of NYU (2011)
- APPLE-1052 Steward Memorandum, Interim Processes for PTAB Workload Management (March 26, 2025)

LISTING OF CHALLENGED CLAIM ELEMENTS

Claim 1	
[1.P]	A video image coding data receiver comprising:
[1.1]	a processor; and
[1.2]	a memory unit having instructions stored which, when executed by the processor, cause the processor to perform operations comprising:
[1.3]	receiving basic video image coding data;
[1.4]	decoding the received basic video image coding data so as to reproduce a video image;
[1.5.1]	receiving supplementary video image coding data including a supplementary hierarchical picture
[1.5.2]	whose coding order and display order are earlier by a factor of a group of pictures including an intra coded picture and a plurality of inter prediction coded pictures than those of a basic hierarchical picture included in the basic video image coding data,
[1.5.3]	a basic hierarchy and a supplementary hierarchy being set in units of the group of pictures;
[1.6]	acquiring basic video image coding data received before supplementary video image coding data that has been received at the moment; and
[1.7]	reconstructing video image coding data from the basic video image coding data and the supplementary video image coding data.
Claim 2	
[2.P]	A video image coding data transmission and reception system comprising:
[2.1]	a video image coding data transmitter that codes and transmits a video image; and
[2.2]	a video image coding data receiver that receives and reconstructs video image coding data,
[2.3.1]	wherein the video image coding data transmitter includes: a first processor; and

[2.3.2]	a first memory unit having instructions stored which, when executed by the first processor, cause the first processor to perform operations comprising:
[2.4]	acquiring the transmission rate of a network;
[2.5]	setting a transmission structure including a basic hierarchy and a supplementary hierarchy, the basic hierarchy and the supplementary hierarchy being set in units of a group of pictures including an intra coded picture and a plurality of inter prediction coded pictures;
[2.6.1]	transmitting basic video image coding data of the basic hierarchy,
[2.6.2]	wherein the first memory unit is further configured to store supplementary video image coding data of the supplementary hierarchy;
[2.7]	transmitting the supplementary video image coding data stored in the memory unit; and
[2.8]	controlling the transmission of the supplementary video image coding data according to the transmission rate, and
[2.9]	wherein the video image coding data receiver includes:
[2.10]	a second processor; and
[2.11]	a second memory unit having instructions stored which, when executed by the second processor, cause the second processor to perform operations comprising:
[2.12]	receiving the basic video image coding data;
[2.13]	decoding the received basic video image coding data so as to reproduce a video image;
[2.14]	receiving the supplementary video image coding data including a supplementary hierarchical picture having coding order and display order that are earlier by a factor of the group of pictures than those of a basic hierarchical picture included in the basic video image coding data;
[2.15]	acquiring basic video image coding data received before supplementary video image coding data that is currently received; and
[2.16]	reconstructing video image coding data from the basic video image coding data and the supplementary video image coding data.

Apple Inc. (“Petitioner” or “Apple”) petitions for *Inter Partes* Review (“IPR”) of claims 1-2 (“the Challenged Claims”) of U.S. Patent No. 9,986,303 (“the ’303 Patent”).

I. REQUIREMENTS FOR IPR

A. Grounds for Standing

Apple certifies that the ’303 Patent is available for IPR. This Petition is being filed within one year of service of a complaint against Apple in *Advanced Coding Techs. LLC v. Apple Inc.*, 2:24-cv-00572-JRG (E.D. Tex.). *See* APPLE-1004. Apple is not barred or estopped from requesting review of the Challenged Claims on the grounds identified below.

B. Challenge and Relief Requested

Apple respectfully requests institution of IPR and cancellation of the Challenged Claims based on the following grounds¹:

Ground	Claim(s)	35 U.S.C. §103
1A	1-2	Demircin
1B	1-2	Demircin in view of Kimoto

¹ The Board **previously instituted** IPR on the ’303 Patent in IPR2024-00327 based on the **same prior art** combinations and related obviousness theories presented in Grounds 1A and 2 of this Petition. *See* APPLE-1036, -1040.

Ground	Claim(s)	35 U.S.C. §103
2	1	Kimoto in view of Demircin

Grounds 1A-2 are further supported by the expert testimony of Dr. Sheila Hemami. *See generally*, APPLE-1003. Dr. Hemami provides an overview of the state of the art and explains in detail why a POSITA would have found the Challenged Claims obvious before the '303 Patent. *E.g., id.*, ¶¶241, 300-304.

The '303 Patent claims priority to a Japanese patent application filed on March 31, 2014. APPLE-1001, Face. Apple does not concede that the Challenged Claims are entitled to the benefit of the Japanese application's filing date, but Apple nonetheless treats March 31, 2014 as the Critical Date for purposes of the analysis in this Petition. APPLE-1003, ¶130. The prior art references that form the basis of Grounds 1A-2 were each filed and published before the Critical Date:

Reference	Effective Filing Date	Publication Date	AIA Prior Art Status
Demircin	Jan. 25, 2007	July 31, 2008	§102(a)(1)-(2)
Kimoto	Aug. 1, 2007	Sept. 1, 2011	§102(a)(1)-(2)

C. Claim Construction

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. Apple submits that no formal construction are presently necessary for purposes of

demonstrating obviousness of the Challenged Claims based on Grounds 1A-2.

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’”) (quoting *Vivid Techs., Inc. v. Am. Sci. & Eng’g, Inc.*, 200 F.3d 795, 803 (Fed.Cir.1999)). Apple reserves the right to respond to any constructions or other arguments implicating claim scope raised by the Patent Owner or the Board. Apple is not conceding that the Challenged Claims satisfy all statutory requirements, nor is Apple waiving any arguments concerning claim scope or grounds that can only be raised in district court. For this Petition, Petitioner applies prior art in a manner consistent with Patent Owner’s allegations of infringement before the district court.

Apple notes that that the parties in a prior litigation involving the ’303 Patent—i.e., a litigation in which Apple was *not* a party—stipulated to constructions of the following claim terms:

Claim Term	Claims	Stipulated Construction
“a factor of a group of pictures”	1-2	“one or more groups of pictures”
“reconstructing video image coding data from the basic video image coding data and the supplementary video image coding data”	1-2	“rearranging basic video image coding data with supplementary video image coding data”

APPLE-1041, 2.

While Apple does not take a position on the stipulated constructions from the prior litigation at this time, Grounds 1A-2 of this Petition nonetheless confirm that the Challenged Claims would have been obvious even if those constructions applied.² *Wellman*, 642 F.3d at 1361; APPLE-1003, ¶¶238-240.

D. Level of Ordinary Skill in the Art

A person of ordinary skill in the art (“POSITA”) relating to the subject matter of the ’303 Patent as of the Critical Date would have had at least a bachelor’s degree in electrical engineering, computer engineering, or computer science, with 2-3 years of experience in digital video processing and familiarity with standards. APPLE-1003, ¶¶137-139. Additional graduate education could substitute for professional experience, and *vice versa*. *Id.*

² Patent Owner previously conceded in IPR2024-00327 that the stipulated construction of “a factor of a group of pictures” had “no bearing” on the Demircin and Kimoto-based obviousness grounds that were similarly advanced in that proceeding. APPLE-1037, 13. In its institution decision in IPR2024-00327, the Board likewise found a reasonable likelihood that the Demircin and Kimoto-based grounds rendered obvious the “reconstructing” limitations of claims 1 and 2 even under the stipulated construction. APPLE-1040, 9-10, 16-17.

II. THE '303 PATENT

A. Brief Description

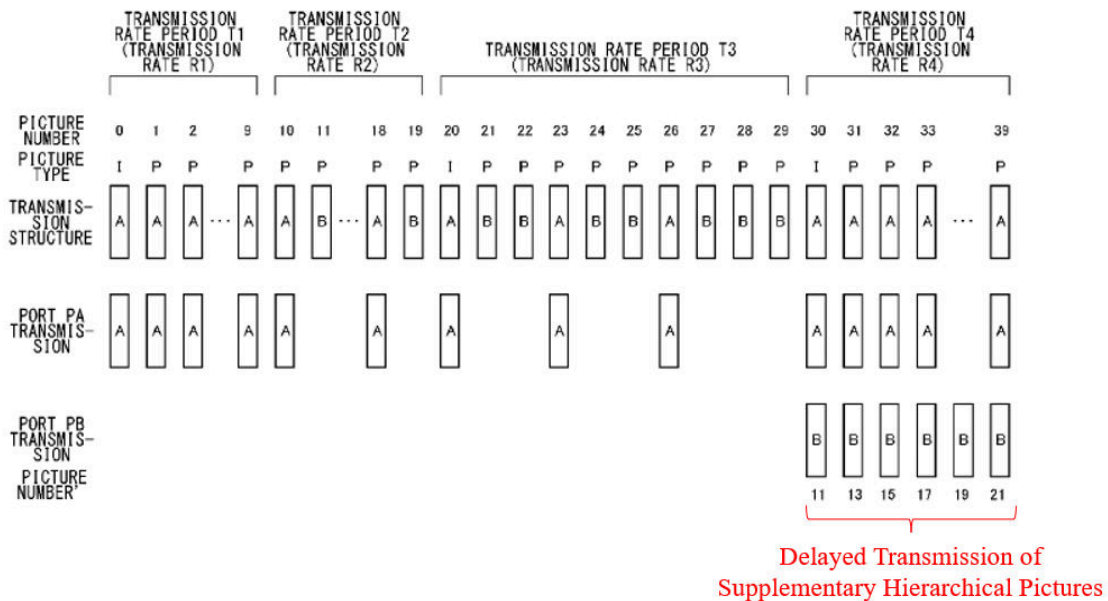
The '303 Patent describes “video image coding data transmission technology” for hierarchically coded video. APPLE-1001, 1:18-19; APPLE-1003, ¶¶140-181. Digital video can be hierarchically encoded using a “basic” structure comprised of “basic hierarchical pictures” and a “supplementary” structure comprised of “supplementary hierarchical pictures.” APPLE-1001, 3:58-67, 9:27-36; *see also id.*, 4:34-42, 4:59-65, 5:42-44, 8:39-42, 9:57-67, 11:22-36. “Basic” video coding data generally refers to video data that can be decoded without reference to “supplementary” data, thereby allowing video to be played at a basic quality level from only basic coding data. *Id.*, 9:57-67. “Supplementary” coding data can be used to supplement “basic” video coding data and thereby enhance the decoded video stream. *Id.*, 4:50-5:44, 8:39-42, FIGS. 1-4.

“Supplementary” coding data generally cannot be decoded without reference to basic data. *Id.* For example, supplementary video coding data can be used to increase the frame rate of a basic video stream, decrease coding errors in the basic stream, or add color to an otherwise monochromatic stream coded in the basic hierarchy. *Id.*, 8:39-42, 14:28-15:27, 15:35-48; APPLE-1003, ¶¶140-181.

The '303 Patent further explains how the division of coded video data into basic and supplementary hierarchies can be beneficial to facilitate transmission of

coded video over a channel whose bandwidth or transmission rate changes over time. APPLE-1003, ¶¶140-181. Basic data can be transmitted in *real-time* during periods of lower channel bandwidth while transmission of supplementary data is *delayed* until channel bandwidth increases sufficiently. APPLE-1001, 11:38-12:25, FIG. 5. The receiver can thus decode and play a video stream from the basic video image coding data in real-time at the basic quality level. *Id.* The receiver can later play the video at a higher quality once the supplementary data is received by reconstructing pictures from both the basic and supplementary hierarchies. *Id.* 11:38-12:25; *see also id.*, 8:60-10:36, 12:27-36, 4:40-42, FIGS. 4-5.

FIG. 5



APPLE-1001, FIG. 5 (annotated).

B. Examination History

During original examination of the '303 Patent, the Examiner initially rejected claim 8³ as obvious based on Radha (APPLE-1031) and Van Der Schaar (APPLE-1032). APPLE-1002, 74-75. The applicant responded to the rejection by emphasizing that claim 8 required a “supplementary hierarchical picture whose coding order *and* display order are earlier than those of a basic hierarchical picture included in the basic video image coding data.” APPLE-1002, 84. The applicant argued that, by contrast, the cited teachings from Van Der Schaar were allegedly limited to “re-arranging the transmission order of the enhancement frames ‘to be the same as the *display order* of the enhancement frames’—but not the coding order. *Id.*

The Examiner subsequently rejected claim 8 as obvious based on Radha and Chen (APPLE-1033). APPLE-1002, 127-128. In response, the applicant amended claim 8 to require that the coding order and display order are earlier “by a factor of a group of pictures including an intra coded picture and a plurality of inter prediction coded pictures” than those of a basic hierarchical picture included in the basic video image coding data, “the basic hierarchy and a supplemental hierarchy being set in units of the groups of pictures.” *Id.*, 158. Following this amendment,

³ Application claim 8 later issued as claim 1 of the '303 Patent.

the Examiner allowed claim 8. *Id.*, 178.

Significantly, however, the file history contains no indication that the Examiner ever considered the prior art relied upon in this Petition (i.e., Demircin or Kimoto) in connection with examination of the '303 Patent. If the art and Grounds asserted herein had been considered during examination, the '303 Patent never should have issued.

III. THE CHALLENGED CLAIMS ARE UNPATENTABLE

A. GROUND 1A – DEMIRCIN RENDERS CLAIMS 1-2 OBVIOUS

1. Overview of Demircin

Demircin, like the '303 Patent, describes techniques for transmitting coded “audiovisual (AV) data from a source device (e.g., server) to one or more destination devices (e.g., clients) over a packet network.” Demircin, [0023]; APPLE-1003, ¶¶195-230. Demircin explains that “high-quality AV data streams require a very high bandwidth continuously,” but “time-varying channel conditions” can make the available bandwidth of a communication channel “unpredictable” and can lead to degradation in video quality when the bandwidth decreases. Demircin, [0024].

Demircin proposes to address these challenges in the context of a Scalable Video Coding (SVC) system. Demircin, [0027]. SVC was well-known by the Critical Date and had been defined in standards since at least MPEG-2. APPLE-

1016, 10, 17, 109 (describing temporal scalability); APPLE-1003, ¶¶10-129, 195-

230. Demircin provides a basic overview of SVC and explains how “SVC enables fine-grained adaptation of a bit stream in spatial, temporal and SNR dimensions.”

Demircin, [0026]. “SVC enables coding video in a scalable manner, resulting in a scalable bit stream, for example by coding the video data in several layers.” *Id.*

“[T]he base layer provides a minimum level of video quality,” while the

“enhancement layers provide additional quality in an incremental manner.”

Demircin, [0035]. “Decoding of the first enhancement layer requires availability of the base layer.” *Id.* “Decoding of each subsequent enhancement layer requires

availability of all previous enhancement layers, as well as the base layer.” *Id.*

FIG. 1 below “illustrates an exemplary prediction structure, GOP structure and layering structure of a sample SVC bit stream” that includes a base layer 10 and enhancement layers 11, 12. *Id.*, [0036].

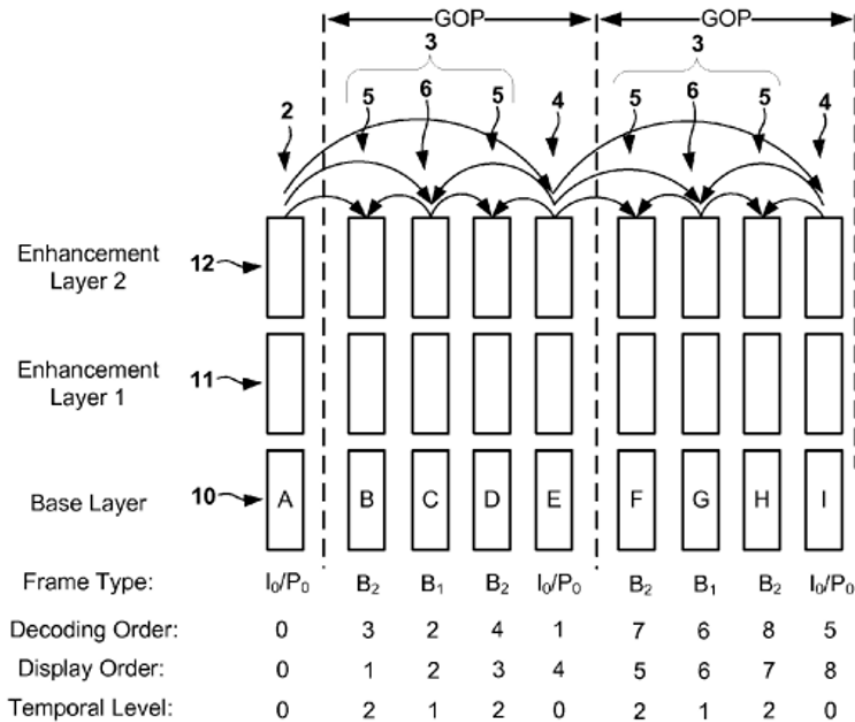


FIG. 1

Demircin, FIG. 1; APPLE-1003, ¶¶195-230.

SVC defines different forms of scalability that can be used individually or in combination. APPLE-1003, ¶¶10-159, 195-230. These include temporal scalability for scaling the frame rate, spatial scalability for scaling the resolution of video frames, and signal-to-noise ratio (SNR) scalability for scaling the visual quality of video frames. Demircin, [0026]. The meaning of “base” and “enhancement” in video depends on the type of scalability at issue. For example, in spatial and SNR scalability, each frame is coded partly in the base layer and partly in one or more enhancement layers. APPLE-1003, ¶¶195-230. This is shown in FIG. 1 where the coding data for each of frames A-I includes a base layer

component 10 and two enhancement layer components 11, 12 that can be combined with the base layer components to enhance the quality of the base layer. *Id.*; Demircin, FIG. 1.

On the other hand, temporal scalability “utilizes the concept of a temporal level, which indicates the level of a video frame in a hierarchical temporal prediction structure.” *Id.*, [0035]-[0036] (“Temporal scalability is achieved using an IBP prediction structure with hierarchical B-frames.”). APPLE-1003, ¶¶195-230. The temporally scalable structure shown in FIG. 1 consists of (i) I-frames at the lowest temporal level (e.g., level 0) that are intra-predicted without reference to other pictures and (ii) B-frames at higher temporal levels (e.g., levels 1 or 2) that are inter-predicted with reference to other frames at a lower temporal level. Demircin, [0036], FIG. 1. In the example shown in Figure 1, “the GOP size is equal to 4 frames.” *Id.* The frames at lower temporal level(s) (e.g., the I-frames at level 0) constitute a “base” layer while the frames at higher temporal level(s) constitute “enhancement” layers in the context of temporal scalability (e.g., B-frames at levels 1 and 2). Demircin, [0036] (“The key pictures are [] intra-coded (I-frame 2 Key pictures may be considered a base layer for temporal scalability and may be assigned temporal level 0.”); APPLE-1003, ¶¶195-230. The following simplified version of Demircin’s FIG. 1 is modified to focus on a video stream that uses only temporal scalability (without spatial or SNR scalability):

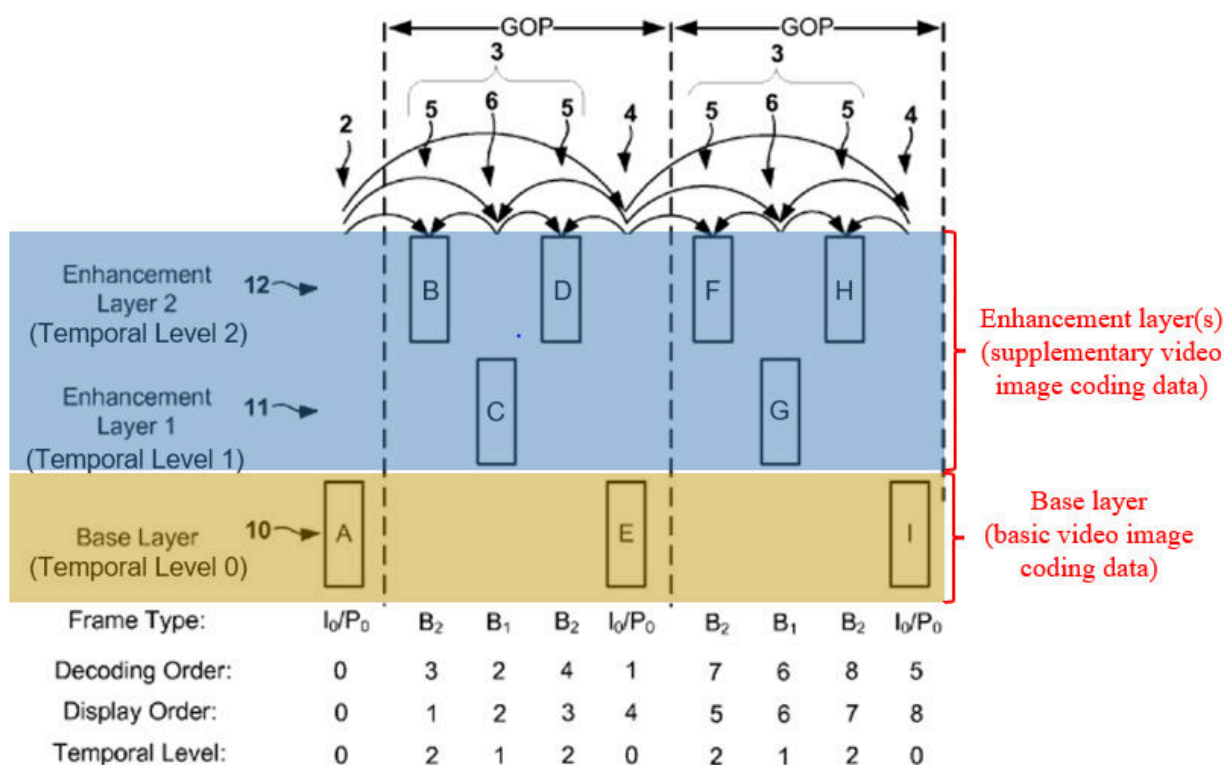


FIG. 1

Demircin, FIG. 1 (modified and annotated⁴); APPLE-1003, ¶¶195-230.

⁴ The prediction structure shown in FIG. 1 is merely exemplary, and as Dr. Hemami explains, a POSITA would have known that many different prediction structures were known and commonly employed in the art before the Critical Date. APPLE-1003, ¶¶53, 195-230, n.4. For example, FIG. 1 depicts a prediction structure with hierarchical B-frames, but a POSITA would have known (and Demircin expressly acknowledges) that P-frames could be used in place of B-

The temporal levels are indicated in the last line of labels below the frame types, and relate to numbers of prediction dependencies. In the example in FIG. 1, frames in display order 0, 4, and 8 are explained to be key pictures, and either I or P frames, which are either intra-coded (I-frames) or predicted from other key pictures (P-frames), and are explained to be a “base layer for temporal scalability.” These frames are assigned a temporal level of 0. Other frames are assigned temporal levels in increasing order based upon the prediction structure, and as such

frames at higher temporal levels. Demircin, [0036], [0039], [0043], [0095] (“a P-frame 164[] ... may be assigned a temporal level value of 1”); Demircin916, FIGS. 4-5 (depicting additional prediction structures and GOP sizes), APPLE-1003, ¶¶195-230, n.4. Demircin’s FIG. 1 also depicts “open” GOPs in which pictures are predicted with reference to decoded picture in a prior GOP. But a POSITA would have appreciated that Demircin’s techniques are equally suitable for “closed” GOPs that do not reference pictures from other GOPs. APPLE-1003, ¶¶195-230, n.4; APPLE-1050, 21, 46, 104. Closed GOPs were standard in MPEG-based video coders and it would have been obvious to a POSITA to use closed GOPs to create splice points in a video sequence, e.g., for scene changes and to provide random access points at which a user could choose to begin playing the video. APPLE-1003, ¶¶195-230, n.4; *cf.* APPLE-1001, 12:37-50.

the frames in each non-zero temporal level n comprise enhancement layer n .

Demircin, [0036] [0095].

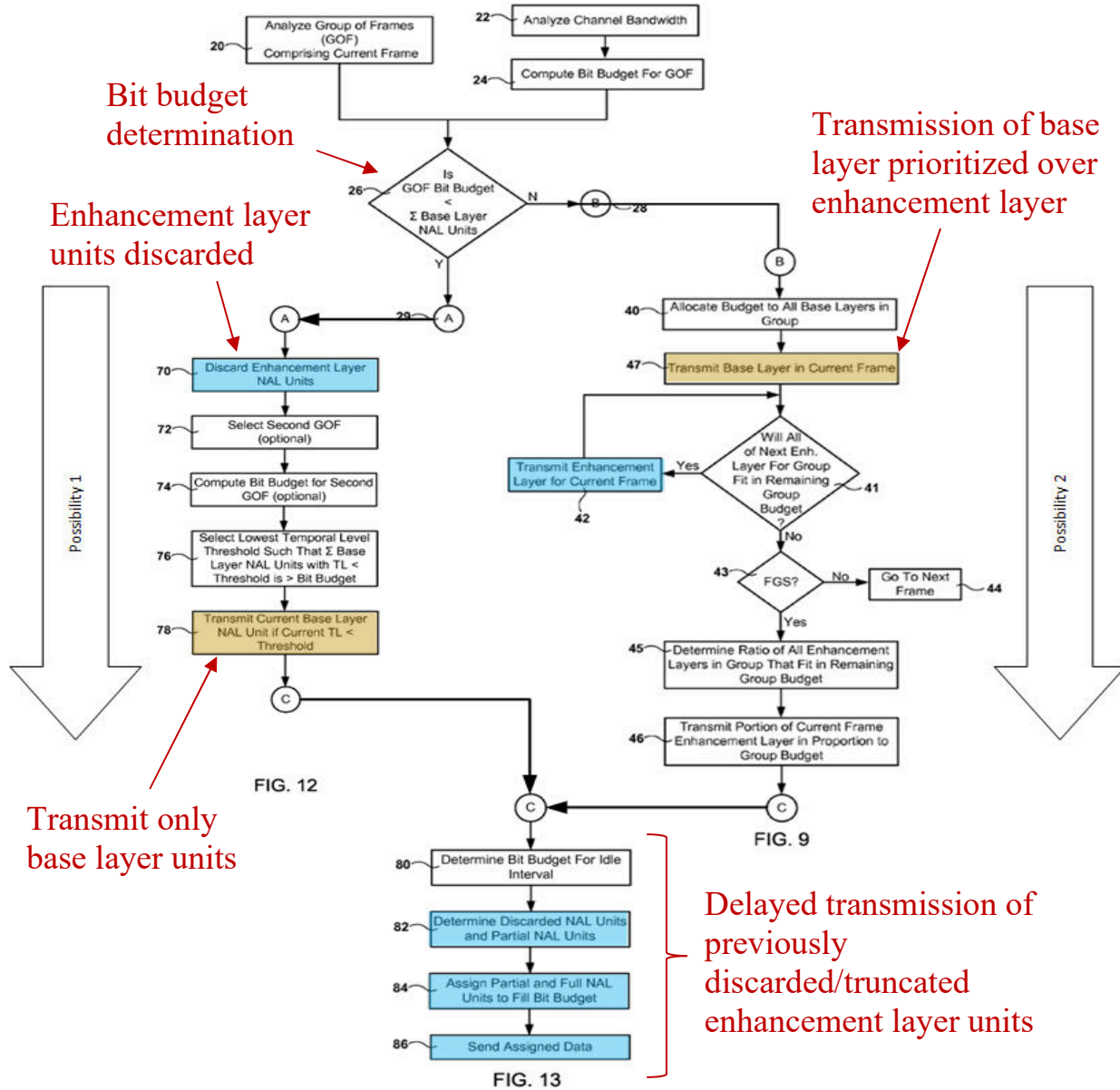
The temporally scalable structure shown in FIG. 1 consists of (i) I-frames at the lowest temporal level (e.g., level 0) that are intra-predicted without reference to other pictures and (ii) B-frames at higher temporal levels (e.g., levels 1 or 2) that are inter-predicted with reference to other frames at a lower temporal level.

Demircin, [0036], FIG. 1. In the example shown in Figure 1, “the GOP size is equal to 4 frames.” *Id.* The frames at lower temporal level(s) (e.g., the I-frames at level 0) constitute a “base” layer while the frames at higher temporal level(s) constitute “enhancement” layers in the context of temporal scalability (e.g., B-frames at levels 1 and 2). Demircin, [0036] (“The key pictures are [] intra-coded (I-frame 2 Key pictures may be considered a base layer for temporal scalability and may be assigned temporal level 0.”). The collection of temporal levels could be $\{(0), (1\ 2)\}$ or $\{(0\ 1), (2)\}$,⁵ when using Demircin’s length-4 GOP depicted in FIG. 1. APPLE-1003, ¶¶195-230, 302.

Demircin further describes techniques for “dynamically adapt[ing] a given SVC bit stream for high-quality video streaming, by intelligently selecting which parts of the bit stream should be transmitted, which parts should be discarded

⁵ The notation of the temporal levels is as follows: $\{(\text{lower}), (\text{upper})\}$. APPLE-1003, ¶302.

entirely, and which parts should be transmitted partially (truncated), in order to meet time-varying rate and/or delay constraints.” Demircin, [0027]. In this way, Demircin adapts the bit rate to varying channel conditions. *Id.* For example, Demircin’s system sets a “bit budget” for each group of frames (GOF) to be transmitted to a receiver based on detected channel conditions. *Id.*, [0061]-[0062]. Demircin then selectively transmits all or a portion of the coded data for a GOF that fits within the allocated bit budget, prioritizing base layer data over enhancement layer data. *Id.*, [0129]-[0135]. If the bit budget cannot accommodate transmission of an enhancement layer frame, the enhancement layer frame can be temporarily discarded and belatedly transmitted during an “idle interval.” *Id.*, [0104]-[0122]. The flowcharts in Figures 7, 9, 12, and 13 illustrate these concepts:



Demircin, FIGS. 7, 9, 12, 13 (annotated); APPLE-1003, ¶¶195-230.

2. Analysis

(a) Claim 1

[1.P]

To the extent the preamble is limiting, Demircin renders obvious [1.P].

APPLE-1003, ¶¶195-230, 241-262.⁶ For example, Demircin discloses “systems for rate-adaptive transmission of video.” Demircin, [0001]. Demircin discloses a destination device (*a video image coding data receiver*) that receives video image coding data from a source / server (*a video image coding data transmitter*) over a packet network. Demircin, [0023] (“transmission of audiovisual (AV) data from a source device (e.g., server) to one or more destination devices (e.g., clients) over a packet network”), [0024] (“a receiver”), [0119] (“client station”); APPLE-1003, ¶¶241-262. The transmitter encodes and transmits video image coding data and the receiver decodes and reconstructs video image coding data. *Infra*, [1.3], [1.5.1], [1.7].

⁶ Dr. Hemami reviewed the Petition and confirmed that the analysis herein accurately reflect a POSITA’s view that Claims 1 and 2 would have been obvious based on Demircin and Claim 1 would have been obvious based on Kimoto in view of Demircin. APPLE-1003, ¶¶241, 284. To provide a more focused discussion in the declaration, Dr. Hemami does not repeat all of the analysis in the Petition but offers additional observations with respect to various claim limitations and obviousness contentions.

Demircin also incorporates by reference U.S. Patent Application Serial No. 11/479,077 (“Demircin077” (APPLE-1035), published as “Demircin916” (APPLE-1006)). Demircin, [0063]. Through Demircin916, Demircin discloses additional details regarding video receivers. For example, with reference to FIG. 1, Demircin916 describes “a plurality of client receiver units 230, 240, 250,” each of which constitutes a video image coding data receiver as recited in [1.P]. See APPLE-1006, [0052], FIG. 1; *see also id.*, [0054]-[0064], [0237], FIG. 30 (depicting “receiver”). Demircin916 also teaches a transmitter “sends a packet burst to the receiver.” APPLE-1006, [0236].

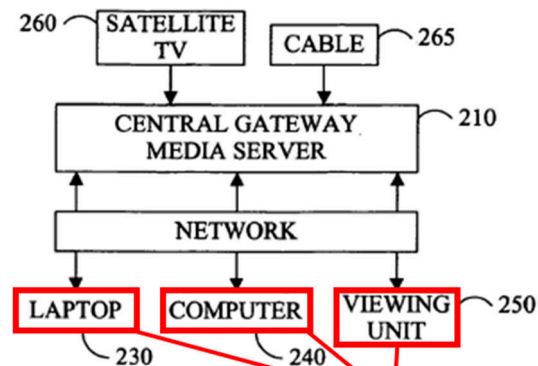


FIG. 1 Video image coding
data receiver

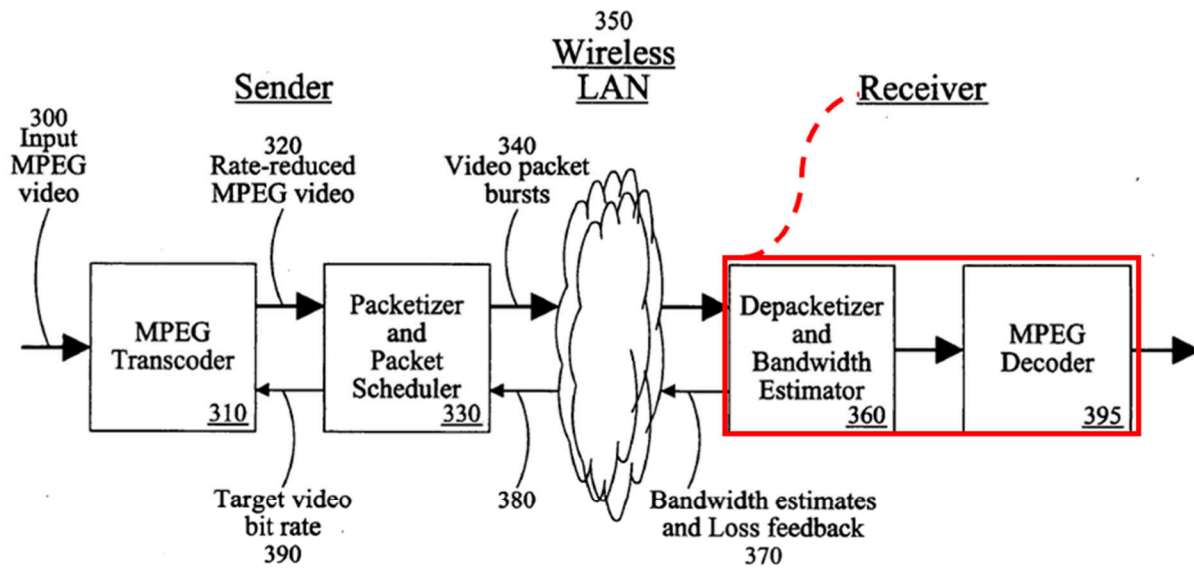


FIG. 30

APPLE-1006, FIGS. 1, 30 (annotated); APPLE-1003, ¶¶241-262.

See also, APPLE-1035, [0052], [0055], [0194]-[0195].

[1.1]

Demircin renders obvious [1.1]. APPLE-1003, ¶¶195-230, 241-262. For example, Demircin discloses “system[s] for adapting a video transmission rate” that require “computational resources” for processing coded video data. Demircin, Abstract, [0025]-[0026]. A POSITA would have understood, and it would have been obvious, that Demircin’s “computational resources” would have included a processor for decoding and performing other processing on coded video data at the receiver. Demircin, [0019]; APPLE-1003, ¶¶241-262.

Demircin further discloses through Demircin916 that “[d]ifferent devices interconnected to the network have different resources and different usage paradigms” including “*different microprocessors.*” APPLE-1006, [0061] (emphasis added); APPLE-1035, [0055]; APPLE-1003, ¶¶241-262. In this context, a POSITA would have understood, and it would have been obvious, to implement Demircin’s video image coding data receiver with a processor that would have allowed the receiver to carry out the disclosed operations for processing coded video data. APPLE-1003, ¶¶241-262.

[1.2]

Demircin renders obvious [1.2]. APPLE-1003, ¶¶195-230, 241-262. For example, Demircin teaches that “[e]lements of embodiments of the present invention may be embodied in hardware, firmware and/or software,” and a POSITA would have known that such a memory unit as recited in [1.2] was conventionally included in “hardware, firmware and/or software” implementations of video image coding data receivers. Demircin, [0019]; APPLE-1003, ¶¶247-249.

Demircin further discloses through Demircin916 that “[d]ifferent devices interconnected to the network have different resources and different usage paradigms” including “*different memory requirements.*” APPLE-1006, [0061] (emphasis added); APPLE-1035, [0055]; APPLE-1003, ¶¶241-262. Thus, a POSITA would have understood, and it would have been obvious, for Demircin’s

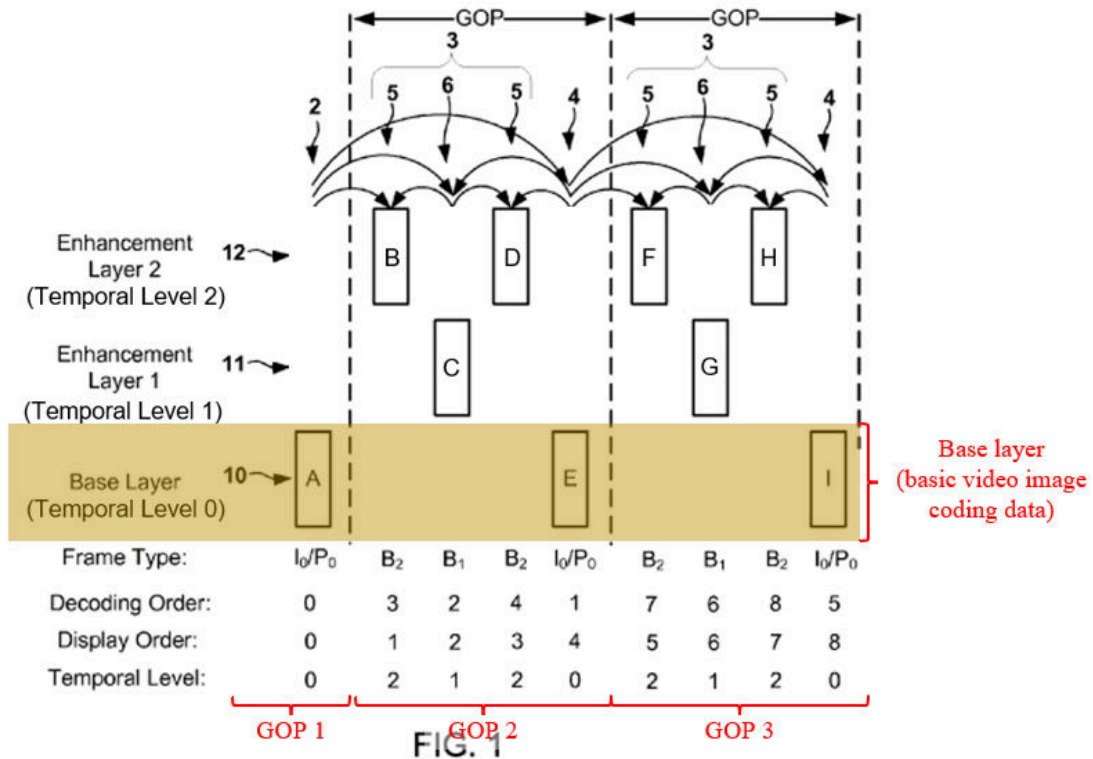
video image coding data receiver to have included a memory unit having instructions as recited in [1.2] to store instructions for carrying out the disclosed processing functions of the receiver. APPLE-1003, ¶¶241-262.

[1.3]

Demircin renders obvious [1.3]. APPLE-1003, ¶¶195-230, 241-262. For example, Demircin discloses that the transmitter transmits coded video data including base layer coding data and enhancement layer coding data to a receiver. Demircin, [0095]-[0103], [0129]-[0132], FIGS. 6-9 & 12-13; APPLE-1003, ¶¶241-262. The transmission system withholds transmission of enhancement layer coding data when the “bit budget” is insufficient to transmit the enhancement layer coding data along with the base layer coding data. *Id.* Demircin’s disclosure of transmitting base layer coding data (i.e., with or without enhancement layer coding data) implies and renders obvious that a receiver would receive the base layer coding data. APPLE-1003, ¶¶241-262.

Demircin describes “Scalable Video Coding (SVC),” which “is an extension of the H.264/MPEG AVC video codec.” Demircin, [0035], [0022] (incorporating by reference the SVC extension to H.264/MPEG-4). “A base layer provides a minimum level of video quality, and may provide compatibility with H.264/AVC.” *Id.*; *see also id.*, [0038]. As described above (*supra*, §III.A.1), one form of scalable coding disclosed in Demircin is temporal scalability, and in the context of

temporal scalability, frames at a lowest temporal level that are coded without reference to other pictures (e.g., I-frames) would have been understood to form a “base” layer while frames at higher temporal levels that are coded with reference to other pictures at lower temporal levels would have been understood as “enhancement” layers. Demircin, [0035]-[0036] (“The key pictures are [] intra-coded (I-frame 2 Key pictures may be considered a base layer for temporal scalability and may be assigned temporal level 0.”); APPLE-1003, ¶¶241-262. Like the basic hierarchical pictures described in the ’303 Patent, Demircin’s temporal base layer frames can be decoded independent of temporal enhancement layer frames such that video can be played at a lower frame rate from the base layer alone. *Id.*; *cf.*, APPLE-1001, 9:57-67. Demircin’s receiver that receives coding data for a temporal base layer (e.g., I-frames at temporal level 0) thus discloses and renders obvious receiving “*basic video image coding data*” as recited in claim 1. APPLE-1003, ¶¶241-262.



Demircin, FIG. 1 (modified and annotated); *see also id.*, FIG. 6; APPLE-1003, ¶¶241-262.

[1.4]

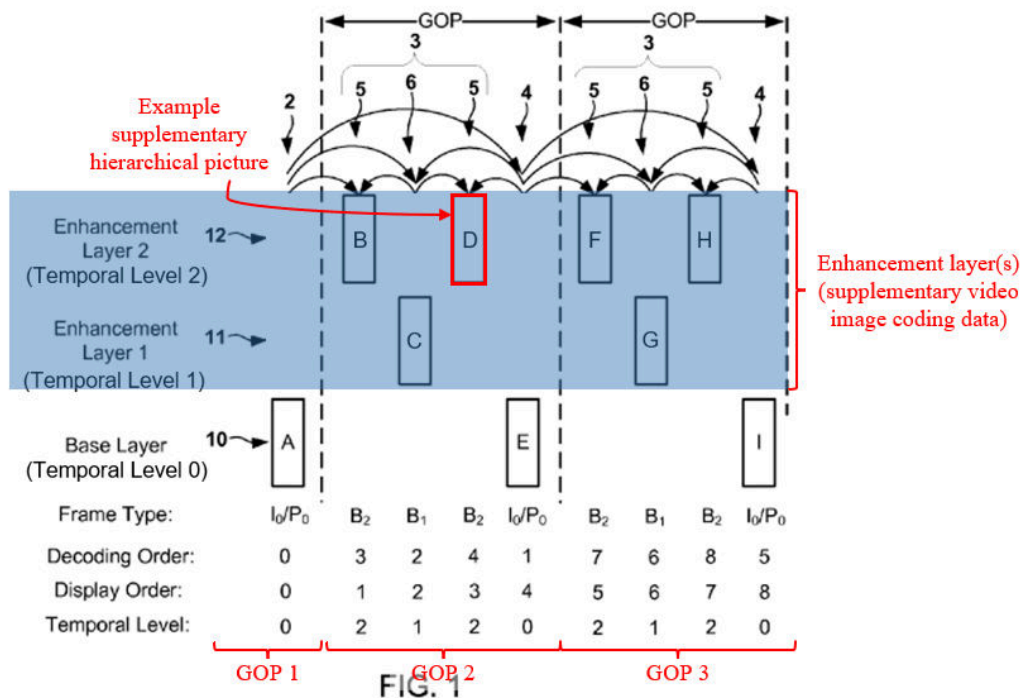
Demircin renders obvious [1.4]. APPLE-1003, ¶¶195-230, 241-262; *supra* [1.3]. For example, Demircin discloses that video data can be coded in a “layered structure” in which the “base layer provides a minimum level of video quality.” Demircin, [0035], FIG. 1. Depending on current network conditions, Demircin’s transmitter will transmit, and the receiver will receive, only coded video frames at the lowest temporal level (*basic video image coding data*) to the exclusion of coded video frames at higher temporal levels. Demircin, [0061] (“[T]he SVC stream bit rate may be adjusted in real-time utilizing ... temporal scalability

features. ... [T]he temporal resolution of the video may be reduced up to a minimum level where only key frames are transmitted, by dropping certain non-key frames.”), [0095]-[0103], [0129]-[0132], FIGS. 6-9 & 12; APPLE-1003, ¶¶241-262. In such instances, it would have been understood and obvious to a POSITA that the receiver would decode video frames at the lowest temporal level since coded frames at the higher temporal levels are not available at the receiver. Demircin, [0035]-[0036]; APPLE-1003, ¶¶241-262. Additionally, in the context of temporal scalability, the received basic video image coding data would be decoded so as to reproduce video images even when the receiver has received coded data for frames at a higher temporal level since playing video at a higher frame rate involves decoding frames both at the lower and higher temporal levels. Demircin, [0035]-[0036], [0061], [0095]-[0103], FIG. 12; *see also id.*, [0131], FIGS. 6-9, 13; APPLE-1003, ¶¶241-262.

[1.5.1]

Demircin renders obvious [1.5.1]. APPLE-1003, ¶¶195-230, 241-262. For example, Demircin’s transmitter transmits, and the receiver thus receives, coded data for video frames at higher temporal levels (e.g., P- or B-frames at temporal levels 1, 2), which corresponds to “*supplementary video image coding data*” including a “*supplementary hierarchical picture*” (e.g., a picture for a P- or B-frame) as recited in claim 1. Demircin, [0022], [0035]-[0036]; *cf.*, APPLE-1001,

9:57-67, 11:22-24 (“a simpler structure can be employed where basic hierarchical pictures and supplementary hierarchical pictures are set to be I pictures and P pictures, respectively”); *supra*, [1.3]-[1.4]; APPLE-1003, ¶¶241-262.

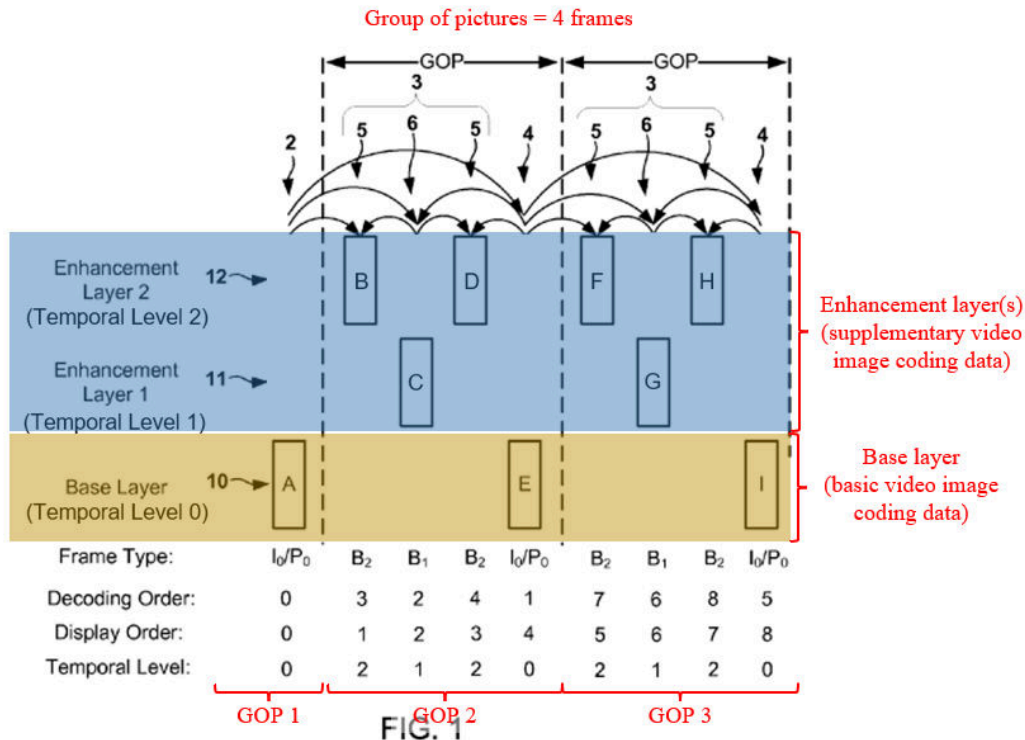


Demircin, FIG. 1 (modified and annotated); *see also id.*, FIG. 6; APPLE-1003, ¶¶241-262.

[1.5.2]

Demircin renders obvious [1.5.2]. APPLE-1003, ¶¶195-230, 241-262. For example, in FIG. 1, Demircin illustrates a GOP structure that “uses I-frames 2, which are intra-predicted from within the frame; B-frames 3, which may be predicted from one or more other frames,” and optionally “P-frames 4 which are predicted from another I-frame.” Demircin, [0036]; *see also id.*, [0039], [0043],

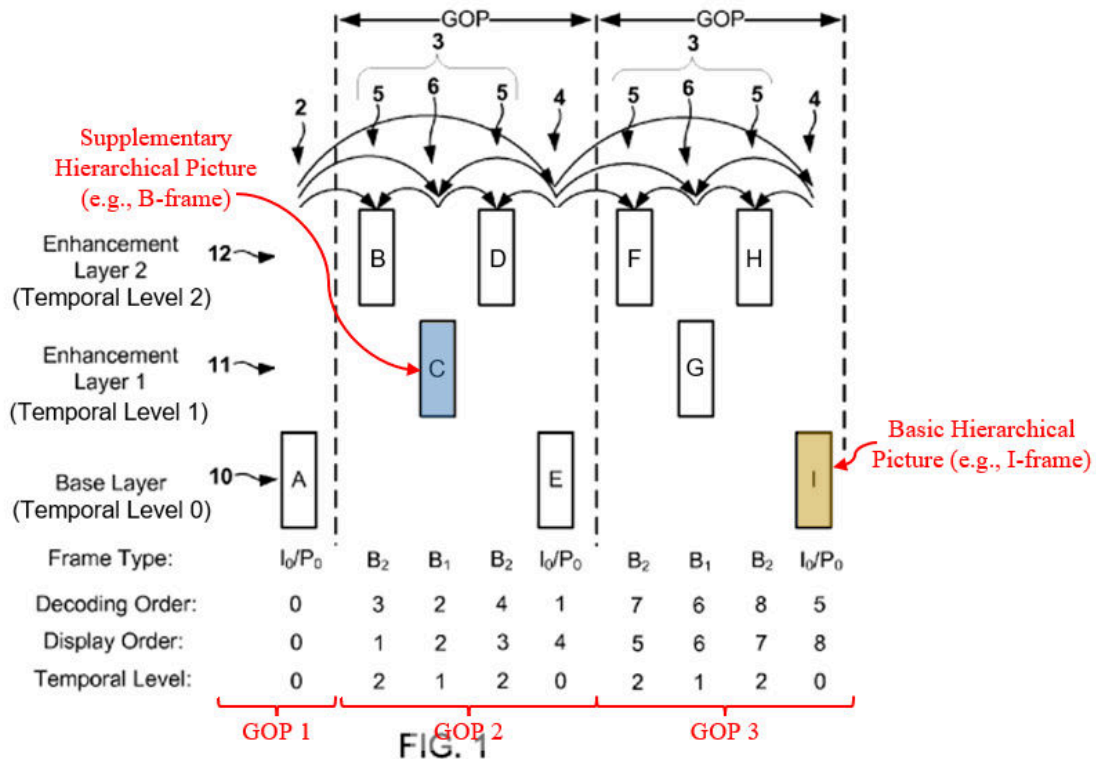
[0095]. The I-frame within Demircin's GOP is an "*intra coded picture*" and the series of B-frames (and/or P-frames) in the GOP constitutes "*a plurality of inter prediction coded pictures*" as recited in [1.5.2]. *Id.*, [0036]; *cf.* APPLE-1001, 9:19-23, 11:22-37; APPLE-1003, ¶¶241-262.



Demircin, FIG. 1 (modified and annotated); *see also id.*, FIG. 6; APPLE-1003, ¶¶241-262.

Demircin also discloses and renders obvious receiving a supplementary hierarchical picture whose coding order and display order are earlier by a factor of a group of pictures including an intra-coded picture and a plurality of inter-prediction coded pictures than those of a basic hierarchical picture included in the basic video image coding data. APPLE-1003, ¶¶241-262; APPLE-1041, 2 (ACT

construing “a factor of a group of pictures” as “*one* or more groups of pictures”);
supra, §I.C. Consider the following version of Demircin’s FIG. 1, which is
adapted to emphasize aspects of Demircin’s teachings related to temporal
scalability:



Demircin, FIG. 1 (modified and annotated); *see also id.*, FIG. 6; APPLE-1003,
¶¶241-262.

As shown above, picture C in GOP 2 is coded as a B-frame and corresponds
to a “*supplementary hierarchical picture*” (SHP) as claimed, while picture I in
GOP 3 is coded as an I-frame and corresponds to a “*basic hierarchical picture*” as
claimed. Demircin, [0035]-[0036], FIG. 1; APPLE-1003, ¶¶241-262. The B-
frame for picture C in this example has a display order that is earlier by a factor of

a group of pictures than the I-frame for picture I, which is evident from FIG. 1's display order indices of 2 and 8 for pictures C and I, respectively. APPLE-1003, ¶¶241-262. The GOP size in FIG. 1 is just four frames including an intra-coded picture and a plurality of inter-predicted pictures, and the display order of picture C is therefore earlier by a factor of a group of pictures than the display order of picture I. *Id.*

Likewise, the B-frame for picture C in this example has a coding order that is earlier by a factor of a group of pictures than the I-frame for picture I, which is evident from FIG. 1's decoding order indices of 2 and 5 for pictures C and I, respectively. APPLE-1003, ¶¶241-262. Demircin's decoding order indicates the order in which frames are decoded at the receiver and corresponds to the order in which frames are encoded at the source. *Id.* The GOP size in FIG. 1 is just four frames including an intra-coded picture and a plurality of inter-predicted pictures, and the decoding order of picture C is therefore earlier by a factor of a group of pictures than the decoding order of picture I. APPLE-1003, ¶¶241-262.

Demircin's receiver receives supplementary hierarchical pictures like the B-frame for picture C having the claimed coding order and display order characteristics, and thus satisfies [1.5.1]-[1.5.2]. APPLE-1003, ¶¶241-262. Demircin renders these limitations obvious when the basic hierarchical picture is received before the supplementary hierarchical picture due to a delay in

transmitting the supplementary hierarchical picture, for example. *Id.* Indeed, a POSITA would have understood and found obvious based on Demircin that a temporal enhancement frame (e.g., SHP) from a prior GOP, such as the blue-shaded B-frame for picture C, would be received after a temporal base frame (e.g., BHP) from a subsequent GOP, such as the gold-shaded I-frame for picture I, when the transmission of the temporal enhancement frame is delayed due to bandwidth or bit budget constraints at the first opportunity for transmission of that frame. APPLE-1003, ¶¶241-262.

For example, Demircin teaches that frames are grouped into groups of frames (GOFs) for purposes of computing “bit budgets” and setting transmission parameters. Demircin, [0029] (“[A] group of frames comprising a current frame and several subsequent frames is considered, a bit budget for that group of frames is calculated, and then bits are allocated ... according to several methods. ... The use of a moving window comprising a group of multiple frames reduces video quality variations over time due to the variations in the bit sizes of individual frames.”); *see also id.*, [0099], [0128]. Demircin’s system then determines which frames to transmit within a GOF based on the frames’ respective temporal levels, where “a lower temporal level indicates greater importance.” *Id.*, [0146]. Frames are selected for transmission “in order of most important temporal level until the bit budget [for the GOF] is exhausted,” which “may be done by selecting a

temporal level threshold or by other methods.” *Id.*, [0147]; *see also id.*, [0061], [0144]-[0147], [0097] (“[T]he system may decide whether the current frame is transmitted or dropped based on the importance of the frame in terms of its temporal level $\tau(i)$.”), [0095]-[0103], FIGS. 6, 12.⁷ Frames in a GOF whose transmission cannot be accommodated within the bit budget for that GOF (e.g., frames at one or more higher temporal levels) are dropped/discarded and are not immediately transmitted with other frames in that GOF (e.g., not transmitted with other frames at one or more lower temporal levels). *Id.*; APPLE-1003, ¶¶241-262.

⁷ Demircin highlights an example embodiment in paragraphs [0061], [0095]-[0103], and [0144]-[0147], and in FIGS. 1 and 12, that involves both SNR and temporal scalability. For this reason, Demircin first discards the SNR-based enhancement layer NAL units for all frames in a GOF before using temporal scalability to reduce the frame rate selecting frames according to temporal level. APPLE-1003, ¶¶241-262. As explained above, however, Demircin also describes the use of temporal scalability even in the absence of SNR scalability. Demircin, [0035]-[0036]; *supra*, §III.A.1; APPLE-1003, ¶¶212, 255, 258. These embodiments would not require first discarding SNR-based enhancement layer NAL units before prioritizing frames for transmission based on temporal level. *Id.*

Demircin further describes techniques for “Delayed NAL Unit Transmission.” Demircin, [0104]; *see also id.*, [0104]-[0122], [0123]-[0128], FIGS. 13-14. With these techniques, “when the system detects there may be additional channel bandwidth available, transmission of NAL units that were previously discarded or transmitted only partially may be considered and performed” and “[e]ligible NAL units may be reconsidered for transmission in delayed manner.” *Id.*, [0034] (“This method results in high video quality because it increases channel bandwidth utilization.”). The system “may detect idle channel intervals (instances when the channel is not fully utilized), and subsequently transmit previously dropped or truncated NAL units whose decoding deadlines have not yet expired.” *Id.*, [0105]; *see also id.*, [0106]-[0108], [0148] (“untransmitted units may be assigned to the supplementary bit budget” and “[t]he assigned units maybe then be transmitted 86 during this interval”), [0149] (“After transmission of these frames, it may be determined that the transmission medium has available bandwidth[] ... Full units and partial units that were previously omitted from transmission may be considered 91 for transmission.”). Demircin’s scheme for “Delayed NAL Unit Transmission” allows frames that were dropped/discarded during an initial transmission window for those frames’

respective GOF(s) to be later transmitted on a delayed basis.⁸ APPLE-1003, ¶¶241-262.

Consistent with Demircin’s description of “Delayed NAL unit Transmission” and delayed transmission techniques, the blue-shaded B-frame for picture C (an exemplary SHP) identified in the adapted version of FIG. 1 above would be transmitted by the transmitter and received by the receiver *after* the gold-shaded I-frame for picture I or even after subsequent I-frames (exemplary BHPs), when the B-frame for picture C is initially discarded due to insufficient bit budget, but is later transmitted on a delayed basis during a suitable idle interval detected after transmission/receipt of the BHP (e.g., the gold-shaded I-frame for picture I), which is transmitted without delay. APPLE-1003, ¶¶241-262. Such delay in transmission by the transmitter, and hence receipt by the receiver, of the B-frame for picture I would occur, for example, when the bit budget in the GOF for picture I is insufficient to accommodate transmission of any frames beyond the higher-priority base layer frames encoded as I-frames. Demircin, [0061] (“temporal resolution of the video may be reduced up to a minimum level where only key frames are transmitted, by dropping certain non-key frames”), [0097] (“the system

⁸ “NAL units” correspond to an access unit of a coded frame such as a B- or P-frame in the context of temporal scalability. APPLE-1003, ¶¶241-262.

may decide whether the current frame is transmitted or dropped based on the importance of the frame in terms of its temporal level $\tau(i)$ ”), [0095]-[0104], [0144]-[0147], FIG. 12; APPLE-1003, ¶¶241-262.

The amount of delay incurred by a discarded but later-transmitted temporal enhancement frame will depend on various factors such as how soon idle interval(s) occur after the initial transmission opportunity for the discarded unit, the bit budget available in the idle interval(s), and the priority of the discarded frame relative to other discarded frames that are eligible for delayed transmission. APPLE-1003, ¶¶241-262. However, based on Demircin’s teachings, it would have been understood and obvious to a POSITA that Demircin’s system would allow for the discarded B-frame for picture C to be transmitted in a delayed fashion during an idle interval *after* transmission of the I-frame for picture I or the I-frame of the next GOP in various conditions. *Id.*

To this point, Demircin explains that discarded frames can be transmitted on a delayed basis so long as their “decoding deadlines have not yet expired.” Demircin, [0105]. A “decoding deadline” refers to the time that all of the bits representing a frame must be available such that the decoder can fetch them and decode the frame. APPLE-1003, ¶¶241-262; *see also id.*, APPLE-1034, 9-10, 15-16; Demircin, [0035]-[0036], [0099], [0104], [0145], FIG. 12; Demircin916, [0252], [0255], [0261], FIG. 39 (decoder buffer 528). As Dr. Hemami explains,

the amount of time between the transmission or receipt of coded data for a frame and its decoding deadline depends to a large extent on the length of time that coded data is buffered at the receiver. APPLE-1003, ¶¶35-37, 45-46, 77, 259, 241-262.

The buffer size or duration, in turn, can vary depending on the needs and objectives of a particular system design or application, and a POSITA would have known how to select a suitable buffer size/duration according to the needs and objectives of the system at hand. *Id.* Longer buffers can absorb greater variability in channel conditions as well as in the size of coded video frames (resulting from the nature of hybrid video coding data, as well as from the use of scalability), and therefore provide a more robust system. *Id.* Longer buffering times also reduce the risk of coded data for frames not being received by their decoding deadlines (i.e., reduce the risk of freezes) and provide greater opportunity for retransmission of lost or corrupted packets and for delayed transmission of discarded enhancement layer frame data. APPLE-1003, ¶¶241-262; APPLE-1034, 9-10, 15-16. On the other hand, longer buffering times increase latency before video playback can begin or resume and require more memory at the receiver. APPLE-1003, ¶¶241-262. The size/duration of the coded picture buffer amounts to no more than an obvious design choice and is a common design parameter (e.g., a result-effective variable) that a POSITA would have found obvious to routinely optimize according to the needs/objectives of a particular design. APPLE-1003, ¶¶241-262.

Nonetheless, a POSITA would have known that receiver buffering durations of ~5-15 seconds at the receiver were common before the Critical Date, and that buffering durations in this range would have been suitable and obvious to apply in many ordinary implementations of a system like Demircin's. *Id.*; APPLE-1034, 9-10, 15-16. With so much time afforded by the buffering delay before a frame is decoded, Demircin's transmitter would have had ample opportunity to delay transmission of the B-frame for picture C (an exemplary SHP) past transmission of the I-frame for picture I, or even past transmission of subsequent I-frames from subsequent GOPs (additional exemplary BHPs), before the decoding deadline for picture C expires. APPLE-1003, ¶¶241-262. Indeed, even if the receiver buffered coded data for just 5 seconds corresponding to 30 frames-per second (FPS) video, the buffer would collect up to ~150 frames before decoding—i.e., far more than the few frames separating pictures C and I in consecutive GOPs, and far more than the number of frames separating picture C from the I-frames in even several subsequent GOPs. APPLE-1003, ¶¶241-262; Demircin916, [0312]. A POSITA would have recognized that 5 seconds (or even less) of buffering time would sufficiently allow for transmission of I-frames for *multiple* GOPs before transmitting the B- or P-frames from an earliest one of those GOPs since the buffering duration is long relative to the size of the GOPs disclosed in Demircin. APPLE-1003, ¶¶241-262. A GOP of size 4 like that shown in FIG. 1, for example,

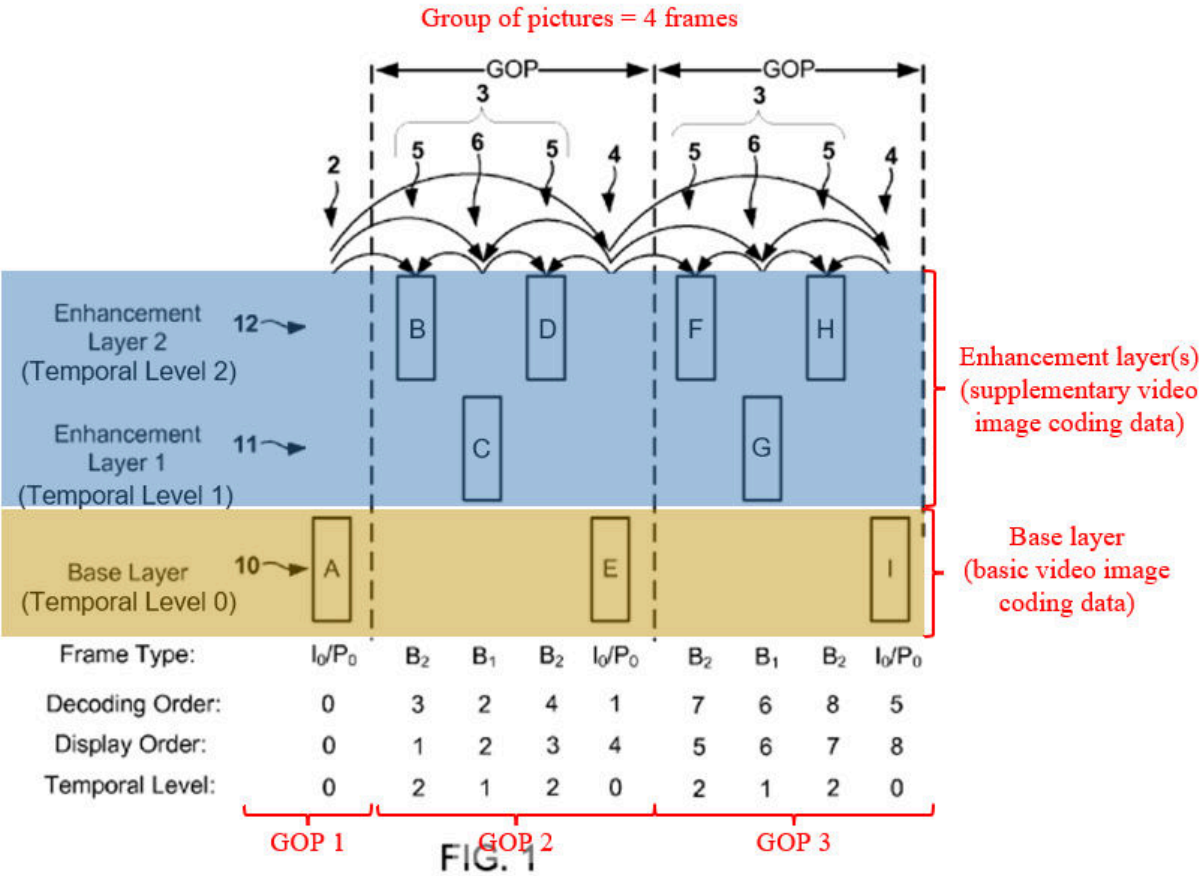
represents just 133 ms of video at 30 FPS, or 67 ms of video at 60 FPS, meaning that coded data for multiple GOPs could be buffered even with much short buffering times as well (e.g., <300 ms) or even with longer GOPs. *Id.*

Moreover, and to the extent claim 1 does not require receipt of the BHP before the SHP (e.g., no requirement for delayed transmission of the SHP), this scenario is likewise rendered obvious by Demircin. APPLE-1003, ¶¶241-262. For example, the B-frame for picture C (e.g., SHP), which is assigned a temporal level of 1, would be transmitted and received with its current GOF without delay when the bit budget for the GOF is sufficiently large to accommodate transmission of frames with temporal level 1. APPLE-1003, ¶¶241-262. The I-frame for picture I (e.g., BHP), which is assigned a temporal level of 0, would then be transmitted and received after picture C when picture I is included in a GOF that follows the GOF of picture C. *Id.*; *see also* Demircin, FIGS. 7-9, 12. The same result would be similarly attained for any other temporal enhancement layer frame (SHP) transmitted and received without delay in a GOF preceding a subsequent GOF for which a temporal base layer frame (BHP) is then transmitted and received. APPLE-1003, ¶¶241-262.

[1.5.3]

Demircin renders obvious [1.5.3]. APPLE-1003, ¶¶195-230, 241-262. For example, as shown in FIG. 1, the base layer and the enhancement layers are set in

units of GOPs. Demircin, FIG. 1. The “GOP size is equal to 4 frames” and the depicted structure “uses I-frames 2, which are intra-predicted from within the frame; B-frames 3, which may be predicted from one or more other frames,” and optionally “P-frames 4 which are predicted from another I-frame.” *Id.*; see also Demircin, [0039], [0043], [0095]; APPLE-1006, [0084] (disclosing option for 9 frame GOP), FIG. 4; APPLE-1035, [0069]; *supra*, Footnote 5.



Demircin, FIG. 1 (modified and annotated); APPLE-1003, ¶¶241-262.

For temporal scalability, I-frames are key pictures and may be considered a base layer with temporal level 0. Demircin, [0036]; APPLE-1003, ¶¶241-262.

“The remaining pictures of a GOP are hierarchically predicted as illustrated by arrows in FIG.1. These non-key pictures form successive temporal levels for temporal scalability, each layer incrementing the temporal level by 1.” Demircin, [0036]; APPLE-1003, ¶¶241-262.

[1.6]-[1.7]

Demircin renders obvious [1.6]-[1.7]. APPLE-1003, ¶¶195-230, 241-262. For context, the '303 Patent teaches that “the coding data reconstruction unit 202 *stores the basic video image coding data provided from the receiving unit 200 in a memory . . .*” and “the coding data reconstruction unit 202 is simply required to be able to *acquire basic video image coding data received before supplementary video image coding data that has been received at the moment.*” APPLE-1001, 7:34-43 (emphasis added). Thus, the '303 Patent contemplates acquiring basic video image coding data by retrieving it from memory. APPLE-1003, ¶¶241-262.

Demircin provides “acquiring” operations much like those disclosed in the '303 Patent's specification. APPLE-1003, ¶¶241-262. For example, as discussed in further detail above (*supra*, [1.5.2]), and based on Demircin's teachings to prioritize immediate transmission of key frames (e.g., I-frames at temporal level 0) (*basic video image coding data*) and delay transmission of non-key frames at

higher temporal levels (*supplementary video image coding data*) when the channel bandwidth dips, it would have been obvious to delay the transmission of frames in a temporal enhancement layer (e.g., the B-frame for picture C in FIG. 1) by a factor of one or more GOPs after transmission of the key frame for that GOP (e.g., the I-frame for picture E in FIG. 1). Demircin, [0061], [0095]-[0103], [0126], [0144]-[0147], FIGS. 6, 12; *see also id.*, [0104]-[0122], [0127], [0148]-[0150], FIGS. 13-14; *supra*, [1.5.2]; APPLE-1003, ¶¶241-262. Also as discussed above (*supra*, [1.5.2]), Demircin discloses and it would have been obvious for the receiver to store coded picture data including key frames (e.g., I-frames at temporal level 0) (*basic video image coding data*) received earlier in time in a buffer before those frames are retrieved for decoding and playback. *Supra*, [1.5.2]; Demircin, [0104]; Demircin916, [0298], FIG. 39 (decoder buffer 528); APPLE-1034, 9-10, 15-16; APPLE-1003, ¶¶241-262. Thus, similar to the '303 Patent, Demircin's coded key frame(s) stored in the decoder buffer (*basic video image coding data*) would be acquired from the decoder buffer at least when they are passed to the decoder, where the acquired base layer frame(s) were received before frames at higher temporal levels (*supplementary video image coding data*) received on a delayed basis at the moment. *Id.* If the B-frame(s) for the frame(s) at higher temporal level(s) are received just in time for the start of decoding of their respective GOP(s) in a corresponding GOF, the video can then be decoded at a

higher frame rate according to Demircin's use of temporal scalability. Demircin, [0035]-[0036], [0095]-[0102], [0144]-[0147], FIGS. 1, 6, 12; APPLE-1003, ¶¶241-262.

Further, because frames in a GOP are decoded according to a prescribed decoding order where the decoded I-frame is used as a reference picture for prediction and decoding of other frames in the GOP, and further because the frames in one GOP are decoded before any frames in a subsequent GOP in a prediction structure like that illustrated in Demircin's FIG. 1, it would have been understood and obvious to a POSITA that the receiver would reconstruct video image coding data by re-arranging delayed temporal enhancement frames (e.g., B- or P-frames) that have been received later in time with temporal base frames (e.g., I-frames / key frames) that have been received earlier in time according to their prescribed decoding order. Demircin, [0036], FIG. 1 (depicting decoding order of frames); *cf.* APPLE-1001, 5:19-25, 8:1-4, 7:51-55, 12:1-5, FIG. 3 (S1104); APPLE-1041, 2; *supra*, §I.C; APPLE-1003, ¶¶241-262. The reconstruction would occur, for example, by passing the earlier-received I-frame in a GOP (*basic video image coding data*) to the decoder before the later-received B- or P-frame(s) in the GOP (*supplementary video image coding data*) according to the decoding order prescribed in Demircin (*see e.g.*, FIG. 1), which would itself entail reconstructing (e.g., rearranging) the earlier received I-frame and the later received B- or P-

frame(s), especially since the decoding order of the frames is different from the order in which the frame was received. APPLE-1003, ¶¶241-262. A POSITA would have appreciated that this re-ordering could be implemented in any of a variety of ways at the receiver, including storing the coded frame data in decoding order in the buffer or retrieving the coded frame data to be passed to the decoder in the correct order from one or more memories in the receiver. Demircin916, [0261] (“the receiver buffers the incoming data units from the wireless network before passing the data units to the decoder”); Demircin, [0026]-[0028], [0035]-[0036]; APPLE-1051, [0082] (“extract the DUs for decoding in a decoding order”); APPLE-1003, ¶¶241-262; *see also* APPLE-1040, 16-17 (“We are currently persuaded that Demircin teaches or reasonably suggests that such delayed NAL units could be received out of order and would be rearranged, as claimed.”).

(b) Claim 2

[2.P]-[2.2]

Supra, [1.P]; APPLE-1003, ¶¶195-230, 241-262.

[2.3.1]

It would have been obvious to a POSITA for Demircin’s source/server (*video image data transmitter*) to include a processor for substantially the same reasons as Demircin’s receiver’s processor would have been obvious as discussed

above in [1.1], *supra*. Demircin, Abstract, [0019], [0025]-[0026]; APPLE-1006, [0061]; APPLE-1035, [0055]; APPLE-1003, ¶¶195-230, 241-262.

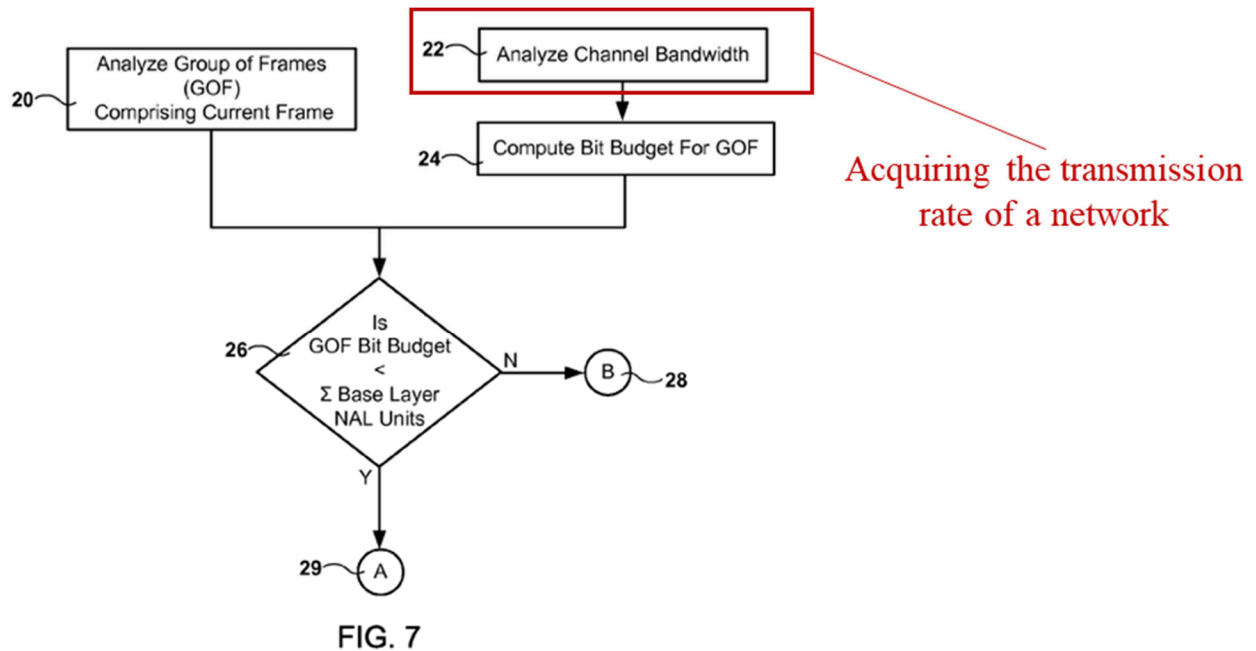
[2.3.2]

It would have been obvious to a POSITA for Demircin's source/server (*video image data transmitter*) to include a first memory unit having instructions executable by a first processor for substantially the same reasons as it would have been obvious for Demircin's receiver's to include a memory unit having executable instructions as discussed above in [1.2], *supra*. Demircin, Abstract, [0019], [0025]-[0026]; APPLE-1006, [0061]; APPLE-1035, [0055]; APPLE-1003, ¶¶195-230, 241-262.

[2.4]

Demircin teaches and renders obvious [2.4]. APPLE-1003, ¶¶241-262. Demircin discloses "methods and systems for adapting a video transmission rate." Demircin, Abstract. Demircin discloses that the source/server "determine[s] the bandwidth of the communication medium over which a signal may be sent" and "[b]ased on the available bandwidth, a GOF bit budget may be computed 24." Demircin, [0129]. "The bit budget may represent an estimate of the amount of data that may be transmitted across the communication medium at the time of the estimate." *Id.*; *see also id.*, [0144]-[0145], [0027]-[0029], [0040] ("bit budget may

be based on available communication bandwidth”), [0148]; APPLE-1003, ¶¶195-230, 241-262.



Demircin, FIG. 7 (annotated); APPLE-1003, ¶¶195-230, 241-262.

[2.5]

Supra, [1.5.1]-[1.5.3]; APPLE-1003, ¶¶195-230, 241-262.

[2.6.1]

Supra, [1.3]; APPLE-1003, ¶¶195-230, 241-262.

[2.6.2]

It would have been obvious to a POSITA for the first memory unit of Demircin’s source/server to store coded enhancement data (*supplementary video image coding data*) such as B-frames encoded at a non-zero temporal level or other

enhancement layer NAL units. APPLE-1003, ¶¶195-230, 241-262. Demircin explicitly teaches through Demircin916, which is incorporated by reference, that “[t]he sender (encoding) part of the system typically includes a buffer at the output of the encoder.” Demircin916, [0252]; *see also id.*, [0255] (“The system may include an encoder buffer 520 after the encoder that receives data from the encoder”). Storing coded enhancement data until an appropriate transmission time would have been obvious to prevent buffer overflows at the receiver and to ensure that any enhancement layer data that was initially discarded due to insufficient bit budget would be available for delayed transmission during a subsequent idle interval. Demircin, [0061] (transmission buffer/queue), [0034] (“Eligible NAL units may be reconsidered for transmission in delayed manner.”), [0104]-[0106], FIGS. 7, 12-14; APPLE-1003, ¶¶241-262. It would have been obvious that the discarded pictures would be stored so as to be available for delayed transmission, and a POSITA would have known and found obvious that the storage of the discarded pictures, like much other digital content, would be in a memory unit. Demircin, [0061]-[0062]; APPLE-1003, ¶¶241-262.

[2.7]

Supra, [1.5.1]-[1.5.3]; APPLE-1003, ¶¶195-230, 241-262.

[2.8]

Demircin's system determines whether to transmit, and thereby controls the transmission of, enhancement layer data for frames in a GOF, according to the transmission rate of a channel as indicated by a calculated bit budget. Demircin, [0129] (“[C]ommunication bandwidth is also analyzed 22 to determine the bandwidth of the communication medium over which a signal may be sent. Based on the available bandwidth, a GOF bit budget may be computed 24. The bit budget may represent an estimate of the amount of data that may be transmitted across the communication medium at the time of the estimate.”), FIGS. 7, 12-14; *supra*, [2.4]; APPLE-1003, ¶¶195-230, 241-262.

Demircin further describes techniques for “Delayed NAL Unit Transmission,” which involves controlled transmission of supplementary video image coding data. *Supra*, [1.P], [1.5.1]-[1.5.3]; Demircin, [0034], [0104]-[0122], [0123]-[0128], FIGS. 7-9, 12-14. With these techniques, “when the system detects there may be additional channel bandwidth available, transmission of NAL units that were previously discarded or transmitted only partially may be considered and performed” and “[e]ligible NAL units may be reconsidered for transmission in delayed manner.” *Id.*, [0034]; APPLE-1003, ¶¶241-262.

[2.9]

Supra, [1.P]; APPLE-1003, ¶¶195-230, 241-262.

[2.10]

Supra, [1.1]; APPLE-1003, ¶¶195-230, 241-262.

[2.11]

Supra, [1.2]; APPLE-1003, ¶¶195-230, 241-262.

[2.12]

Supra, [1.3]; APPLE-1003, ¶¶195-230, 241-262.

[2.13]

Supra, [1.4]; APPLE-1003, ¶¶195-230, 241-262.

[2.14]

Supra, [1.5.1]-[1.5.3]; APPLE-1003, ¶¶195-230, 241-262.

[2.15]

Supra, [1.6]; APPLE-1003, ¶¶195-230, 241-262.

[2.16]

Supra, [1.7]; APPLE-1003, ¶¶195-230, 241-262.

**B. GROUND 1B – DEMIRCIN IN VIEW OF KIMOTO RENDERS
CLAIMS 1-2 OBVIOUS**

Beyond the relevance of Demircin’s teachings on its own (Ground 1A), a
POSITA also would have found the features recited in [1.6]-[1.7] and [2.15]-[2.16]

obvious based on the teachings of Demircin as modified in view of Kimoto.⁹

APPLE-1003, ¶¶195-237, 263-266. This Petition addresses Kimoto in further detail below with respect to Ground 2, but it notably discloses a receiving unit 20 that receives “lower layer” coding data for a video long before receiving any “upper layer” data for the video. Kimoto, [0041]-[0043]; *infra*, §III.C.1, [1.6]-[1.7]. Like Demircin and the ’303 Patent, Kimoto explains that its video is “hierarchically coded” and “is scalable in one or a plurality of directions including time (frame rate).” Kimoto, [0037]-[0038]. In the context of temporal scalability, a POSITA thus would have understood that Kimoto’s lower layer data would comprise key frames (e.g., I-frames at temporal level 0) while the upper layer data would comprise non-key frames (e.g., B- or P-frames at higher temporal levels), in some examples. APPLE-1003, ¶¶195-237, 263-266, 302; Demircin, [0035]-[0036]. Kimoto further explains that, in preparation for decoding, the lower layer data received earlier in time is acquired from a lower layer data storing unit, and is provided to a hierarchical data synthesizing unit to be synthesized with upper layer data received at the moment, thereby forming high quality data. Kimoto, [0069]-

⁹ Demircin-Kimoto render obvious all other elements of claims 1-2 for the reasons described above with respect to those elements in Ground 1A. APPLE-1003, ¶¶263-266.

[0070], [0080], [0083], [0109], Abstract, FIGS. 1-2, 6-7.

A POSITA would have found it obvious to apply Kimoto's teachings in this regard to Demircin such that Demircin's receiver would likewise acquire temporal base layer frame(s) received earlier in time from memory, and would provide the temporal base layer frames(s) to a hierarchical data synthesizing unit to be synthesized with temporal enhancement layer frame(s) received at the moment (e.g., frame(s) whose transmission was delayed per Demircin's teachings), thereby forming high quality data suitable for decoding at a high frame rate as described in Kimoto. APPLE-1003, ¶¶195-237, 263-266. A POSITA would have appreciated and found obvious that synthesizing temporal base/lower layer data (*basic video image coding data*) with temporal enhancement/upper layer data (*supplementary video image coding data*) as taught in Kimoto would involve reconstructing the video image coding data by re-arranging the base/lower layer frame data and the enhancement/upper layer frame data in the correct decoding as opposed to the order in which the frames were received, for reasons discussed in detail above—and especially since Kimoto explains that the result of the synthesizing operation is to produce high-quality data that is suitable for decoding. Kimoto, [0070] (“reproducing unit 201 then decodes the high quality data 1006”); *supra*, Ground 1A, [1.6]-[1.7]; APPLE-1003, ¶¶195-237, 263-266.

A POSITA would have sought to integrate Kimoto's teachings in Demircin

in this regard to better ensure Demircin's temporal base and enhancement layer frames are properly ordered for decoding, despite the delayed transmission of enhancement layer frames. APPLE-1003, ¶¶195-237, 263-266. This would serve as a reliable way of ensuring that frames are not inadvertently sent to the decoder before their respective reference picture(s) have been decoded, and would further ensure that the receiver does not inadvertently bypass decoding delayed enhancement layer frames that were not received contemporaneously with the base layer frame in a GOP; *see also 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). A POSITA would have reasonably expected success implementing the combination, especially since Demircin and Kimoto describe similar hierarchical coding techniques and re-arranging coded frames was a known technique in the art. APPLE-1003, ¶¶195-237, 263-266.

Additionally, while Demircin describes a real-time streaming system in which the delayed transmission of enhancement layer data is limited by real-time decoding deadlines at the receiver, Kimoto describes another option in which all lower layer data for a video is transmitted to and received by the receiver before the upper layer data. Kimoto, [0041]-[0043]. A POSITA would have applied Kimoto's teachings in this regard to Demircin such that Demircin's receiver would

receive all temporal base layer frames (e.g., I-frames) for a video before temporal enhancement layer frames (e.g., B- or P-frames). APPLE-1003, ¶¶195-237, 263-266. In the combination, and to the extent required by the claims, the resulting system would thus delay transmission and receipt, by at least a factor of a group of pictures, of supplementary video image coding data including a supplementary hierarchical picture (e.g., B- or P-frames) relative to basic video image coding data including a basic hierarchical picture (e.g., I-frames). APPLE-1003, ¶¶195-237, 263-266; *supra*, [1.5.1]-[1.5.3], [1.6]-[1.7]. A POSITA would have sought to temporally separate transmission of basic and enhancement layer data in the combination as taught in Kimoto to achieve predictable benefits, including allowing users to view video at a lower quality even before any enhancement layer data is received. APPLE-1003, ¶¶195-237, 263-266. This would have been desirable in various circumstances, such as where the transmission channel is initially bandwidth limited, where the user wishes to view the video first on a smaller display, or where the user may first wish to preview the video at a lower quality without consuming additional bandwidth. Kimoto, [0042], [0091]-[0095]; APPLE-1003, ¶¶195-237, 263-266. Kimoto's approach would also reduce complexity of Demircin's transmitter by avoiding the need to actively manage transmission of enhancement layer frames while transmitting base layer frames in the first instance. APPLE-1003, ¶¶195-237, 263-266. A POSITA would have

reasonably expected success implementing the combination since it merely involves techniques that were well known before the Critical Date for transmitting base and enhancement layer data at different times. *Id.*

C. GROUND 2 – KIMOTO IN VIEW OF DEMIRCIN RENDERS CLAIM 1 OBVIOUS

1. Overview of Kimoto

Kimoto discloses a “moving image data distribution system for distributing reproduction content data that is hierarchically structured so that lower layer data is synthesized with upper layer data to form high quality data.” Kimoto, Abstract, [0019]-[0024], [0063], [0069]-[0070]; APPLE-1003, ¶¶231-237, 267-308. Kimoto perceived that “the amount of code to be generated varies greatly from one frame to another” when using coding methods defined in, e.g., MPEG. Kimoto, [0003]. Likewise, “[i]n a viewing environment of limited bandwidth, there can occur fluctuations in the reception time of the frame data.” *Id.*

Considering these issues, Kimoto’s solution achieved “quick data initialization even when the reproduction contents have a high bit rate” by “transmitting only the upper layer data when low quality data has already been viewed on a receiving terminal that contains the lower layer data.” *Id.*, [0025], [0041]-[0043]. “When reproducing the contents,” Kimoto describes that a “hierarchical data synthesizing unit” synthesizes the received upper layer data and

the corresponding part of the stored lower layer data “to generate high quality data.” *Id.*, [0069].

Kimoto explains that its techniques are broadly “applicable” to “international standards of the hierarchical coding methods” including “MPEG-2 scalable, MPEG-4 scalable, MPEG-4 FGS, H.264 / MPEG-4 AVC scalable video coding, and Motion JPEG 2000,” as well as “other hierarchical coding methods.” Kimoto, [0038]. Kimoto techniques use hierarchical coding and temporal scalability. Kimoto, [0037].

2. Combination of Kimoto and Demircin

Kimoto describes receiving, storing, and displaying the lower layer data earlier in time; then, later in time, receiving upper layer data, and synthesizing the upper layer data with corresponding stored lower layer data to generate high quality data. Kimoto, [0069]-[0070]. Accordingly, Kimoto teaches substantially the same features as claim 1 of the '303 Patent. APPLE-1003, ¶¶231-237, 267-308. Kimoto does not, however, expressly disclose setting the hierarchically coded lower and upper data as GOPs, which have an intra coded picture and a plurality of (i.e., at least two) inter prediction coded pictures. APPLE-1003, ¶¶195-237, 267-

308.¹⁰ Doing so, however, would have been obvious in Kimoto, as described in more detail below. *Id.*

In a predictable combination based on Kimoto and Demircin, the hierarchically coded lower and upper data of Kimoto would be set in GOPs that include an intra coded picture and a plurality of inter prediction coded pictures, like the GOP structure taught by Demircin's Figure 1. APPLE-1003, ¶¶195-237, 267-308.

3. Reasons to Combine Kimoto and Demircin

A POSITA would have been motivated to modify Kimoto based on the teachings of Demircin such that Kimoto's lower and upper layer data would be organized in GOPs consisting of an intra coded picture (*e.g.*, an I-frame) and a plurality of inter-prediction coded pictures (*e.g.*, hierarchical B-frames), *e.g.*, consistent with the structure depicted in FIG. 1 of Demircin. APPLE-1003, ¶¶195-237, 267-308. This would have been obvious and beneficial to establish accurate and efficient synthesis of newly received upper layer data with stored lower layer

¹⁰ Dr. Hemami notes that Kimoto states that the upper layer data is transmitted (and hence received) at a point in time after the lower layer data has been received, decoded, and reproduced. *E.g.*, Kimoto, [0042]; APPLE-1003, ¶¶195-237, 267-308.

data, which supports Kimoto's delayed playback of high quality contents. APPLE-1003, ¶¶38-56, 195-237, 267-308. Additionally, a POSITA would have recognized the advantages of using a GOP structure, as taught by Demircin, to support random playback capabilities and account for varying channel bandwidths. *Id.*

Accurate and Efficient Synthesis to Support Delayed Playback

At the outset, a POSITA would have turned to Demircin as analogous art because both Kimoto and Demircin were concerned with prioritizing the transmission of base layer (or lower layer) data over the enhancement layer (or upper layer) data, and both involved the synthesis of lower and upper layer data later in time. *E.g.*, Kimoto, [0041]-[0042], [0067]; Demircin, [0135]. And both Kimoto and Demircin describe synthesizing high-quality data at the receiver by reconstructing lower layer data with upper layer data. *E.g.*, Kimoto, [0069]; Demircin, [0035]; APPLE-1003, ¶¶195-237, 267-308.

Kimoto describes the use case in which a viewer “views low quality contents for some period of time before viewing the contents in high quality again.” Kimoto, [0095] (emphasis added); *see also id.*, [0042]. The delay between the initial (low quality) viewing and the start of the (high quality) re-viewing is either fixed, *id.*, [0072], or at the request of the user. *Id.*, [0073]; APPLE-1003, ¶¶195-237, 267-308. A POSITA also would have been motivated to turn to Demircin for

structure of the lower and upper layer data in the form of GOPs because Demircin describes a mechanism to assign data coded using that GOP structure to lower and upper layers. *E.g.*, Demircin, [0035]-[0036]; APPLE-1003, ¶¶195-237, 267-308. Such an assignment facilitates the matching of receiving upper layer data with the “corresponding part” of the stored lower layer data during synthesis, thereby increasing efficiency and accuracy of the synthesis. Kimoto, [0070]; APPLE-1003, ¶¶195-237, 267-308. Additionally, such assignments also facilitate random playback. APPLE-1003, ¶¶195-237, 267-308. Thus, a POSITA would have recognized the tangible benefits to Kimoto of applying Demircin’s GOP structure. *Id.*; Kimoto, [0070].

Variable Bandwidth

Kimoto describes a first embodiment involves sending the lower layer (low-quality video) to a mobile terminal, and a POSITA would have recognized such transmissions may have required sending lower layer data over a link that has varying bandwidth. Kimoto, [0015], [0042]; APPLE-1003, ¶¶195-237, 267-308. A POSITA would have understood that, in some cases, the mobile unit may have supported a lower bandwidth than the receiving unit; and in that scenario, Kimoto would have benefited from the reduced-rate representation of the video, as taught by Demircin. Demircin, [0035]-[0036], [0135]-[0138]; APPLE-1003, ¶¶195-237, 267-308.

Because Kimoto teaches that the low quality contents still need to be reproduced on the mobile terminal, a POSITA would be motivated to look for ways to ensure that the perceived quality of video (e.g., frame rate) at the mobile terminal was as high as possible to improve the user's experience. APPLE-1003, ¶¶195-237, 267-308. A POSITA would have understood that use of Demircin's approach for bit rate adaptation using an appropriate GOP prediction structure would have resulted in the highest-frame-rate-possible video being transmitted to the mobile terminal and would improve the user experience. Demircin, [0035]-[0036], [0135]-[0138]; APPLE-1003, ¶¶20-56, 112-122, 195-237, 267-308.

Moreover, Demircin teaches flexibility in the GOP structure to adapt to varying bandwidths that are often experienced by mobile terminals. Demircin, [0036] (length 4 GOPs), FIGS. 6 (length 8 and 12 GOFs); APPLE-1003, ¶¶10-129, 195-237, 267-308. A POSITA would have recognized the benefits of Demircin's adaptability for providing low quality contents for viewing on Kimoto's mobile terminal. APPLE-1003, ¶¶195-237, 267-308. Likewise, a POSITA would have looked to Demircin's teachings, incorporated in Demircin, to estimate the available channel bandwidth to select the best GOP structure to support transmission of the lower layer data. APPLE-1003, ¶¶10-159, 195-237, 267-308. A POSITA would use Demircin's teachings to map the scalable video provided by a hybrid video coder with an appropriate GOP prediction structure into Kimoto's

lower layer and upper layer. APPLE-1003, ¶¶195-237, 267-308. Accordingly, a POSITA would have recognized the obvious benefits of using Demircin's approach for bit rate adaptation as resulting in the highest-frame-rate-possible video being transmitted to the mobile terminal. APPLE-1003, ¶¶195-237, 267-308.

The application of known techniques for organizing a video stream into GOPs, as taught in Demircin, also would have improved the similar type of scalable video coding system of Kimoto in a similar way and would have yielded merely predictable results. APPLE-1003, ¶¶195-237, 267-308. Implementing GOPs and frame structures like those in Demircin in other scalable video contexts such as that described in Kimoto would have been well within the norms of standard practice at the time. *Id.*; Kimoto, [0038]; Demircin, [0020], [0026].

Reasonable Expectation of Success

Setting the lower and upper layer data of Kimoto in GOPs, like that of Demircin, would have been met with a reasonable likelihood of success, especially since the use of GOPs with intra coded pictures and a plurality of inter prediction coded pictures was a well-known video coding technique at the time. APPLE-1003, ¶¶10-159, 195-237, 267-308; MPEG1, FIGS. 1, 15. GOP prediction structures were well known and used as early as 1993, and were understood to provide temporal scalability. APPLE-1003, ¶¶10-159, 195-237, 267-308; MPEG1,

FIGS. 1, 15. A POSITA would have been able to structure Kimoto's lower and upper layer data in units of GOPs, like that described by Demircin, because doing so was a common and standard practice at that time. *E.g.*, APPLE-1003, ¶¶10-159, 195-237, 267-308; MPEG1, FIGS. 1, 15.

Kimoto already explains that its "content data" is "hierarchically coded" and is "scalable in one or a plurality of directions including time (frame rate)."

Kimoto, [0037]. Kimoto is broadly applicable to hybrid video coding. *Id.*, [0038]. Likewise, Demircin involves hierarchical scalable video coding, such as that prescribed by MPEG and other standards bodies. *E.g.*, Demircin [0035] (citing H.264/MPEG AVC/SVC); APPLE-1003, ¶¶195-237, 267-308.

Also, Kimoto already contemplates structuring pictures as GOPs, as shown in Kimoto, FIG. 4. Kimoto describes that its hierarchical structure includes a non-reference frame and a plurality of corresponding reference frames. *E.g.*, Kimoto [0056]. Thus, the underlying frame structure for intra- and inter-prediction in Kimoto already includes the same number of frames necessary to satisfy the claim limitation of the '303 Patent. *See* '303 Patent, 9:33-36 (describing reference frames); APPLE-1003, ¶¶195-237, 267-308. The GOP of Demircin already included one intra-coded picture and a plurality of inter-prediction coded pictures, which was a commonplace GOP structure at the time as discussed at length above. *E.g.*, Demircin, FIG. 1, [0036].

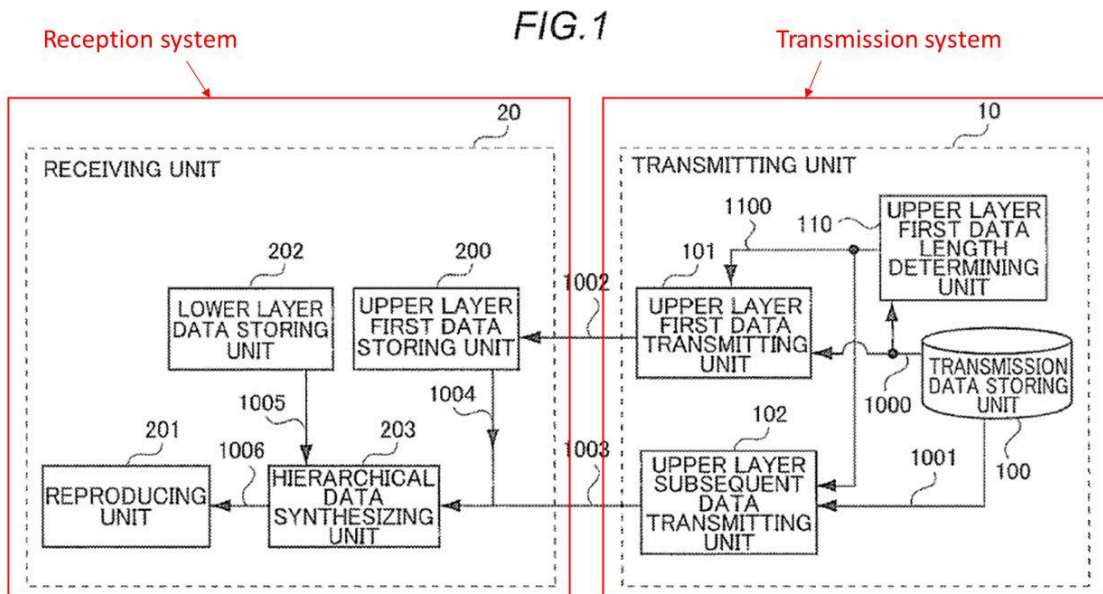
Additionally, Demircin's description of structuring data in units of GOPs, as well as its description of adaptability, is thorough and comprehensive. APPLE-1003, ¶¶195-237, 267-308. Thus, a POSITA reading Demircin would not have had to undertake undue experimentation to structure the lower and upper layer data of Kimoto in a GOP structure that allowed both 1) sufficiently high quality viewing of lower layer data, 2) adaptive lower (and upper) layer data structuring, and 3) efficient and accurate synthesis of received upper layer data with the corresponding part of the stored lower layer data to produce high quality contents. APPLE-1003, ¶¶195-237, 267-308. The GOP prediction structure taught in Demircin naturally supported scalability and hence the mapping of frames to lower and upper layers for each GOP. *See, e.g., supra* §§ II.D, V.A.5, & VI.A.

4. Analysis

(a) Claim 1

[1.P]

To the extent the preamble is limiting, the Kimoto-Demircin combination renders it obvious. APPLE-1003, ¶¶195-237, 267-308. For example, Kimoto FIG. 1 illustrates a “moving image distribution system” that includes “a receiving unit 20” (*video image coding data receiver*). Kimoto [0035]. Kimoto FIG. 1 illustrates the receiving unit that can perform image decoding:



Kimoto, FIG. 1 (annotated).

Kimoto explains that receiving unit 20 is “capable of reproducing contents.”

Kimoto, [0040].

[1.1]

The Kimoto-Demircin combination renders obvious [1.1]. APPLE-1003, ¶¶195-237, 267-308. For example, according to Kimoto, “the moving image data distribution system may be implemented by hardware, software, or a combination of these.” Kimoto, [0111]. A POSITA would have known and found obvious that a system that was implemented in hardware and that executed software included a processor. APPLE-1003, ¶¶267-308.

Kimoto also teaches that “a computer” can be made to “function as a receiving apparatus of a moving image data distribution system” Kimoto,

[0024]. A POSITA would have known that computers included processors, and thus would have recognized and found obvious that Kimoto's receiving unit would have included a processor. APPLE-1003, ¶¶267-308. Indeed, Kimoto teaches that the receiving unit in each embodiment "performs the same *processing*." Kimoto, [0108] (emphasis added). A POSITA would have known and found it obvious for the receiving unit to have performed "processing" with a processor. APPLE-1003, ¶¶267-308.

To the extent Kimoto does not expressly disclose the recited processor, a POSITA nonetheless would have found it obvious to have implemented a processor in the receiving unit to carry out the various functions of the receiving unit as described in Kimoto and provided in the Kimoto-Demircin combination, including receiving, storing, retrieving, decoding, synthesizing, and reproducing moving images, as was commonly practiced at the time (as also evidenced by Demircin itself). APPLE-1003, ¶¶267-308; *supra* §A.2-[1.1]. A POSITA would have had a reasonable expectation of success implementing a processor in the receiving unit, especially in view of the common use of processors for carrying out processing functions in receiving units and other types of computers as of the Critical Date. APPLE-1003, ¶¶267-308.

[1.2]

The Kimoto-Demircin combination renders obvious [1.2]. APPLE-1003, ¶¶195-237, 267-308. For example, according to Kimoto, “the moving image data distribution system may be implemented by hardware, software, or a combination of these.” Kimoto, [0111]. A POSITA would have known and found obvious that a system that was implemented in hardware and that executed software would have included a memory to store the software instructions. APPLE-1003, ¶¶267-308.

Kimoto also teaches that a “computer” can be made to “function as a receiving apparatus of a moving image data distribution system” Kimoto, [0024]. A POSITA would have known that computers commonly included memory to store software instructions, and thus would have recognized and found obvious that Kimoto’s receiving unit would have included such a memory. APPLE-1003, ¶¶267-308.

To the extent Kimoto does not teach the recited memory unit, a POSITA would have found it obvious to have included a memory in the receiving unit to store instructions for executing the various functions of the receiving unit as described in Kimoto and provided in the Kimoto-Demircin combination. APPLE-1003, ¶¶267-308. A POSITA would have had a reasonable expectation of success implementing a memory storing instructions in the receiving unit, especially in

view of the common use of memory units for storing instructions in receiving units and other types of computers as of the Critical Date. *Id.*; *supra* §A.2-[1.2].

[1.3]

The Kimoto-Demircin combination renders obvious [1.3]. APPLE-1003, ¶¶195-237, 267-308. For example, Kimoto describes “lower layer data” (***basic video image coding data***) from which moving image contents can be viewed at a low quality. Kimoto [0025]. Kimoto explains that the “lower layer data on the reproduction contents is previously stored in the lower layer data storing unit 202 by the point in time when the contents are reproduced.” Kimoto [0041]. In situations where a user has already viewed low quality contents, “it is assumed that the low quality content data, or lower layer data, has already been transferred from a server to the receiving unit 20.” Kimoto [0043]-[0044], [0025]. Thus, in Kimoto, at some time before receiving the upper layer data, the receiving unit receives and stores the lower layer data. *Id.*; APPLE-1003, ¶¶267-308. Like basic video image coding data in the ’303 Patent, Kimoto’s lower layer data can be decoded independently and reference to the upper layer or supplementary video image coding data. Kimoto, [0041]-[0042]; APPLE-1001, 9:57-67; APPLE-1003, ¶¶267-308.

The lower layer data of Kimoto is hierarchically coded according to industry standards such as H.264/MPEG-4 advanced video coding (AVC) scalable video

coding (SVC). Kimoto, [0019]-[0024]; Demircin, [0020]-[0021], [0026], [0035]-[0037]. According to both Kimoto and Demircin, hierarchical coding of the video coding data (including lower layer data) can take different forms, including temporal scalability. Kimoto, [0037]; Demircin, [0026]; APPLE-1003, ¶¶267-308.

Based on the exemplary context illustrated by Demircin, in the combination, the hierarchical structuring of lower layer data in Kimoto includes basic hierarchical pictures organized in units of GOPs. *Supra*, III.A.1-2; APPLE-1003, ¶¶267-308.

As described above (*supra*, §§III.A.1 & III.A.2.[1.3]), Kimoto and Demircin both describe temporal scalability. Kimoto, [0037]; Demircin, [0035]-[0036]. In temporal scalability, frames at a lowest temporal level that are coded without reference to other pictures (e.g., I-frames) would have been understood to form a “base” layer (or lower layer), while frames at higher temporal levels that are coded with reference to other pictures at lower temporal levels would have been understood as “enhancement” layers (or upper layers). Demircin, [0035]-[0036]; APPLE-1003, ¶¶267-308.

Like the basic hierarchical pictures described in the ’303 Patent, base layer frames in temporally scalable video are decoded independent of enhancement layer frames such that video can be played at a lower frame rate from the base layer alone. *Id.*; *cf.*, APPLE-1001, 9:57-67. Kimoto’s receiver that receives lower layer

data such as temporal base layer frames (e.g., I-frames at temporal level 0) thus discloses and renders obvious receiving “**basic video image coding data**” as recited in claim 1. APPLE-1003, ¶¶267-308.

[1.4]

The Kimoto-Demircin combination renders obvious [1.4]. APPLE-1003, ¶¶195-237, 267-308.

As one example, Kimoto discloses “a situation where a user *who has viewed low quality contents* on a mobile terminal such as a cellular phone and a PDA (Personal Data Assistant) views the same contents in high quality on the receiving unit 20 which is a part of the embodiment of the present invention.” Kimoto, [0042] (emphasis added). In such cases, “it is assumed that the low quality content data, or lower layer data, has *already been transferred* from a server to the receiving unit 20.” *Id.*, [0043]; *see also id.*, [0025] (“reduce the first data transmission time by transmitting only the upper layer data when low quality data has already been viewed on a receiving terminal that contains the lower layer data”). Viewing low quality contents involves decoding lower layer data (**basic video image coding data**), e.g., by decoding temporal base layer frames at a lower frame rate. APPLE-1003, ¶¶267-308. It also would have been obvious to decode the lower layer data at the receiving unit 20 to provide greater flexibility for a user to view low quality contents on a device other than a mobile terminal without

requiring the user to wait for receipt of upper layer data. *Id.*

As another example, Kimoto's receiver decode the lower layer data (*basic video image coding data*) as part of the decoding of high quality data in which the lower layer data has been synthesized with upper layer data. Kimoto, [0069]-[0071]; APPLE-1003, ¶¶267-308. In temporal scalability, lower layer data is decoded without reference to upper layer data even after the lower layer data and upper layer data has been reconstructed. This is because temporal base layer frames (e.g., I-frames) are decoded independently of temporal enhancement layer frames (e.g., B- or P-frames), the latter of which must be decoded with reference to pictures at lower temporal levels. Thus, even when decoding high quality data at a high frame rate, a subset of the decoded frames are frames from the lower layer data. APPLE-1003, ¶¶267-308.

Additionally, Kimoto describes "a case where low quality contents and high quality contents are switched and viewed on an identical receiving terminal."

Kimoto, [0091]; *see also*, [0092]-[0095] ("the viewer views low quality contents for a certain period of time"). Viewing the low quality contents in this case involves decoding the lower layer data to reproduce a video image. APPLE-1003, ¶¶267-308.

[1.5.1]

The Kimoto-Demircin combination renders obvious [1.5.1]. APPLE-1003,

¶¶195-237, 267-308. Kimoto teaches receiving upper layer data (*supplementary video image coding data*). *Id.*; *supra* [1.3]-[1.4]. In addition, Kimoto teaches that the upper layer data is hierarchically coded. Kimoto, [0019]-[0024].

In Kimoto, the upper layer data can be split into components to account for reproduction delays. Kimoto, [0047]-[0049]. Kimoto refers to these components as upper layer first data and upper layer subsequent data. Kimoto, [0047], [0068]. The upper layer data meets the claim limitation whether it includes upper layer first data, upper layer subsequent data, or both because Kimoto describes synthesizing the upper layer data with corresponding lower layer data for reconstruction. Kimoto, [0069]; APPLE-1003, ¶¶267-308.

After the receiving unit receives the lower layer data (*basic video image coding data*), Kimoto, [0042]-[0043], the receiver “receives the upper layer first data 1002” and “stores the upper layer first data 1002.” Kimoto, [0067]. The hierarchical data synthesizing unit “synthesizes the upper layer first data” with the “corresponding part” of the stored lower layer data to generate high quality data. *Id.*, [0069]. Likewise, the hierarchical data synthesizing unit “subsequently synthesizes the upper layer subsequent data” with “the corresponding part” of the lower layer data to generate high quality data. *Id.*, [0070]. Since the upper layer first data and the upper layer subsequent data are each synthesized with the lower layer data, either or both read onto the claimed supplementary video image coding

data. *Id.*, APPLE-1003, ¶¶267-308. Thus, Kimoto discloses and the combination renders obvious receiving supplementary image coding data (i.e., the upper layer data). APPLE-1003, ¶¶267-308.

Kimoto further discloses, and the combination renders obvious, that the upper layer data (whether upper layer first data, upper layer subsequent data, or both) includes a supplementary hierarchical picture. APPLE-1003, ¶¶267-308.

For example, the upper layer data is hierarchically coded using an international standard of hierarchical coding, such as H.264/MPEG-4 AVC SVC. Kimoto, [0019]-[0024], [0038]; Demircin, [0020]-[0021], [0026], [0035]-[0037].

According to both Kimoto and Demircin, hierarchical coding of the video coding data (including upper layer data) can include temporal scalability. Demircin, [0026]; Kimoto, [0037]; APPLE-1003, ¶¶267-308.

As described above (*supra*, §§III.A.1 & III.A.2.[1.3]), Kimoto and Demircin each disclose temporal scalability. Kimoto, [0037]; Demircin, [0035]-[0036]. In the context of temporal scalability, I-frames would have been understood to form a “base” layer, while frames at higher temporal levels that are coded with reference to other pictures at lower temporal levels would have been understood as “enhancement” layers, in some examples. Demircin, [0035]-[0036]; APPLE-1003, ¶¶267-308. The temporal upper layer data include inter-predicted pictures such as B- or P-frames that constitute a “*supplementary hierarchical picture*,” as recited

in claim 1. APPLE-1003, ¶¶267-308.

[1.5.2]

The Kimoto-Demircin combination renders obvious [1.5.2]. APPLE-1003, ¶¶195-237, 267-308. First, Kimoto teaches that upper layer data (*supplementary video image coding data*) is not received by the receiving unit until the lower layer data (*basic video image coding data*) has already been received, stored, decoded, and viewed. Kimoto, [0041]-[0043], [0025]; *supra* B.4-[1.3]. The reception of lower layer data sufficient to decode and view a series of pictures in the moving image content at a low quality, before the reception of any upper layer data for those pictures, means that the later received upper layer data includes at least one supplementary hierarchical picture whose coding order and display order is earlier than a basic hierarchical picture in the earlier received and stored lower layer data. APPLE-1003, ¶¶267-308.

Further, in the combination, Kimoto's upper- and lower-layer data are organized in groups of pictures (GOPs) using a frame structure like that in Demircin's FIG. 1. *Supra*, [1.3]-[1.4], [1.5.1]; *infra* [1.5.3]. With a GOP size of just four frames, for example, it would have been obvious to a POSITA for a received supplementary hierarchical picture in the upper layer to be earlier by a factor of a GOP including an I-frame and a plurality of B-frames than those of a basic hierarchical picture included in the lower-layer data. APPLE-1003, ¶¶267-

308. Indeed, many GOPs, and even substantially all GOPs in the coded video, would separate transmission of temporal base layer frames and temporal enhancement layer frames due to the bifurcated transmission of lower and upper layer data as taught in Kimoto APPLE-1003, ¶¶267-308.

Therefore, in the Kimoto-Demircin combination, it would have been obvious to a POSITA that there would be many upper layer pictures having a coding order and display order earlier by a factor of a GOP than a lower layer picture. APPLE-1003, ¶¶267-308.

[1.5.3]

The Kimoto-Demircin combination renders obvious [1.5.3]. APPLE-1003, ¶¶195-237, 267-308. In the combination, the lower layer data and the upper layer data of Kimoto are set in units of GOPs, such as the GOP described by Demircin. Demircin, [0036]-[0037], FIG. 1; *supra*, §III.C.2-3 APPLE-1003, ¶¶267-308. The upper layer data and the lower layer data in the combination are hierarchically coded using an international standard of hierarchical coding, such as H.264/MPEG-4 AVC SVC. Kimoto, [0019]-[0024], [0037]-[0038]; *see also* Demircin, [0020]-[0021], [0026], [0035]-[0037], FIG.1; *supra* [1.3]; APPLE-1003, ¶¶267-308. According to MPEG, a GOP is “the smallest coding *unit* that can be independently decoded.” APPLE-1015, 54 (emphasis added). In the combination, the lower and upper layer data of Kimoto are likewise set in units of GOPs that can

be decoded independently of pictures in other GOPs. APPLE-1003, ¶¶267-308.

[1.6]

The Kimoto-Demircin combination renders obvious [1.6]. APPLE-1003, ¶¶195-237, 267-308.

First, Kimoto teaches that “it is assumed that the low quality content data, or lower layer data, has already been transferred from a server to the receiving unit 20.” Kimoto, [0025], [0044]. Thus, the lower layer data (*basic video image coding data*) is already stored at the receiver before upper layer data (*supplementary video image coding data*) is received. Kimoto, [0025], [0044]. This lower layer data is basic video image coding data, which includes a basic hierarchical picture. *Supra*, [1.3].

Then, Kimoto teaches that the receiver “receives the upper layer first data 1002 from the transmitting unit 10,” and “stores the upper layer first data 1002.” Kimoto, [0067]. Next, “[w]hen reproducing the contents, the hierarchical data synthesizing unit 203 synthesizes the upper layer first data 1004 stored in above S03 and the corresponding part 1005 of the lower layer data to generate high quality data 1006 (S04).” Kimoto, [0067]-[0070]. Because the lower layer data is stored in the lower layer data storing unit 202, it would have been obvious that the receiving unit acquires the lower layer data from lower layer data storing unit 202 in order for the hierarchical data synthesizing unit 203 to carry out operations for

synthesizing the upper layer data with the lower layer data. Kimoto, [0067]-[0070], FIGS. 1, 6-7; APPLE-1003, ¶¶267-308. This process of acquiring earlier-received lower layer data (*basic video image coding data*) from memory is congruent with the understanding of the term “acquiring” from the ’303 Patent specification, which discloses that, “after supplementary hierarchical pictures are received, video image coding data to be reconstructed from basic hierarchical pictures stored in the memory and the received supplementary hierarchical pictures.” APPLE-1001, 12:6-9.

Kimoto also teaches that reproduction of high-quality images is performed in real time upon receipt of upper layer data after a short buffering interval. Kimoto, [0048], [0053], [0060]-[0064], FIGS. 3-5. “The receiving unit 20 becomes capable of reproducing contents when upper layer first data is stored in the upper layer first data storing unit 200” and “[t]he lower layer data on the reproduction contents is previously stored in the lower layer data storing unit 202 by the point in time when the contents are reproduced.” *Id.*, [0040]-[0041], [0069]-[0071].

[1.7]

The Kimoto-Demircin combination renders obvious [1.7]. APPLE-1003, ¶¶195-237, 267-308. Kimoto teaches that the receiving unit reconstructs video image coding data by synthesizing the lower layer data (*basic video image coding*

data) with the upper layer data (*supplementary video image coding data*) to generate high quality data for reproduction. Kimoto, [0019], [0043], [0069]-[0071], FIG. 2 (S05-S06). Note that the '303 Patent also describes “synthesizing basic video image coding data and supplementary image coding data” using a “coding data reconstruction unit.” APPLE-1001, 12:53-55 (emphasis added).

A POSITA also would have understood Kimoto’s synthesizing operation to include rearranging basic video coding data with supplementary video image coding data. APPLE-1003, ¶¶267-308; *supra*, § I.C. For example, Kimoto’s lower layer data (*basic video image coding data*) is received (and is stored locally) earlier in time than upper layer data (*supplementary video image coding data*). Kimoto [0025], [0044], [0067]-[0070]. The lower layer data can be used by itself to reproduce a video with low quality contents, just like the claimed basic video image coding data. Kimoto, [0041]-[0042]. When the upper layer data is received, the hierarchical data synthesizing unit synthesizes the upper layer first data with a corresponding part of the lower layer data stored in memory. Kimoto, [0069]-[0070]. Because of the significant difference in arrival times of the upper layer data compared to the lower layer data, the synthesizing of the upper layer data with a corresponding part of the lower layer data naturally involves a rearranging of the upper and lower layer data to realign the data and properly order the upper layer SHPs with the lower layer BHPs. *Id.*; APPLE-1003, ¶¶267-308.

IV. PTAB DISCRETION SHOULD NOT PRECLUDE INSTITUTION

Apple believes that discretionary denial is unwarranted, and yet, Apple intends to utilize the bifurcated briefing process contemplated by the March 26, 2025, Stewart Memorandum, APPLE-1052, to rebut contentions if offered by Patent Owner to the contrary. See <https://www.uspto.gov/about-us/events/learn-aboutnew-interim-processes-relating-institution-aia-proceedings> (April 24, 2025); <https://uspto.cosocloud.com/p9xpz6s4jn75> (Video).

V. CONCLUSION AND FEES

The Challenged Claims are unpatentable. Apple authorizes charge of fees to Deposit Account 06-1050.

VI. 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)

Apple Inc. is the sole petitioner and real party-in-interest.

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

The '303 Patent is the subject of the following pending civil actions:

Advanced Coding Technologies LLC v. Apple Inc., 2:24-cv-00572 (E.D. Tex.), filed July 22, 2024 (see APPLE-1004) and *Advanced Coding Technologies LLC v. Google LLC*, 2:24-cv-00353 (E.D. Tex.), filed May 10, 2024.

The '303 Patent was also previously asserted in *Advanced Coding Technologies LLC v. Samsung Electronics Co. Ltd.*, 2:22-cv-00499 (E.D. Tex.) and *Advanced Coding Technologies LLC v. LG Electronics, Inc.*, 2:22-cv-00501 (E.D.

Tex.). Samsung previously filed an IPR on the '303 Patent in IPR2024-00327.

Apple understands that these prior district court actions and IPR involving the '303 Patent have all been terminated or dismissed due to settlement between the parties in those proceedings. Apple was not a party to those proceedings.

Apple is not aware of any other disclaimers, reexamination certificates, or IPR petitions addressing the '303 Patent.

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

Apple 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: IPR50095-0243IP1@fr.com	Jeremy J. Monaldo, Reg. No. 58,680 Nicholas Stephens, Reg. No. 74,320 Rishi Gupta, Reg. No. 64,768 60 South Sixth Street, Suite 3200 Minneapolis, MN 55402 Tel: 202-783-5070 Fax: 877-769-7945 IPR50095-0243IP1@fr.com

D. Service Information

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

Apple consents to electronic service by email at IPR50095-0243IP1@fr.com.

Respectfully submitted,

Dated: May 9, 2025

/W. Karl Renner/

W. Karl Renner, Reg. No. 41,265
Jeremy J. Monaldo, Reg. No. 58,680
Nicholas Stephens, Reg. No. 74,320
Rishi Gupta, Reg. No. 64,768
Fish & Richardson P.C.
60 South Sixth Street, Suite 3200
Minneapolis, MN 55402
T: 202-783-5070
F: 877-769-7945

(Control No. IPR2025-00983)

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 13,639 words, which is less than the 14,000 allowed under 37 CFR §42.24.

Dated: May 9, 2025

/W. Karl Renner/

W. Karl Renner, Reg. No. 41,265
Jeremy J. Monaldo, Reg. No. 58,680
Nicholas Stephens, Re. No. 74,320
Rishi Gupta, Reg. No. 64,768
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 May 10, 2025, a complete and entire copy of this Petition for *Inter Partes* Review, Power of Attorney, and all supporting exhibits were provided via USPS Priority Mail Express, to the Patent Owner, by serving the correspondence address of record as follows:

VENABLE LLP
P.O. BOX 34385
WASHINGTON, DC 20043-9998

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