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.: 39843-0171IP1
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 3

II. THE '303 PATENT 3
- **A.** Brief Description 3
- **B.** Prosecution History 7

III. THE CHALLENGED CLAIMS ARE UNPATENTABLE 9
- **A.** GROUND 1 – DEMIRCIN RENDERS CLAIMS 1-2 OBVIOUS 9
  - 1. Overview of Demircin 9
  - 2. Analysis 13
- **B.** 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 52
  - 4. Analysis 57

IV. PTAB DISCRETION SHOULD NOT PRECLUDE INSTITUTION 74
- **A.** 35 U.S.C. §325(d) – *Advanced Bionics* 74
- **B.** 35 U.S.C. §314(a) – *Fintiv* 76

V. CONCLUSION AND FEES 76

VI. MANDATORY NOTICES UNDER 37 C.F.R § 42.8(a)(1) 77
- **A.** Real Party-In-Interest Under 37 C.F.R. § 42.8(b)(1) 77
- **B.** Related Matters Under 37 C.F.R. § 42.8(b)(2) 77
- **C.** Lead And Back-Up Counsel Under 37 C.F.R. § 42.8(b)(3) 77
- **D.** Service Information 78

EXHIBITS

SAMSUNG-1001	U.S. Patent No. 9,986,303 to Takehara, et al. ("the '303 Patent")
SAMSUNG-1002	Excerpts from the Prosecution History of the '303 Patent ("the '303 Patent File History")
SAMSUNG-1003	Declaration and Curriculum Vitae of Dr. Clifford Reader.
SAMSUNG-1004	Complaint, *Advanced Coding Technologies LLC v. Samsung Electronics Co., Ltd. et al.*, 2:22-cv-00499-JRG, E.D. Tex., filed December 30, 2022
SAMSUNG-1005	U.S. Patent Publication No. 2008/0181302 ("Demircin")
SAMSUNG-1006	U.S. Patent Publication No. 2007/0153916 ("Demircin916")
SAMSUNG-1007	U.S. Patent Publication No. 2011/0211631 ("Kimoto")
SAMSUNG-1008	File History for Kimoto ("Kimoto File History")
SAMSUNG-1009	ITU-T Recommendation H.264 (11/07)
SAMSUNG-1010	U.S. Patent Publication No. 2004/0196972 ("Zhu")
SAMSUNG-1011	U.S. Patent No. 8,335,252 ("Chen")
SAMSUNG-1012	U.S. Patent No. 8,175,154 ("Luthra")
SAMSUNG-1013	U.S. Patent No. 9,094,641 ("Ansley")
SAMSUNG-1014	Wu Jiyan, Shang Yanlei, Cheng Bo, Wu Budan, and Junliang Chen; "*Loss Tolerant Bandwidth Aggregation for Multihomed Video Streaming over Heterogeneous Wireless Networks";* Wireless Personal Communications: An International Journal. 75. 1265-1282. 10.1007/s11277-013-1422-3; Published online October 5, 2013 ("Wu")
SAMSUNG-1015	ISO/IEC 11172-2: 1993 (E)
SAMSUNG-1016	ISO/IEC 13818-2: 2000 (E)

SAMSUNG-1017 Infringement Contentions, *Advanced Coding Technologies LLC v. Samsung Electronics Co., Ltd. et al.*, 2:22-cv-00499-JRG, E.D. Tex., filed December 30, 2022

SAMSUNG-1018 Schwarz et al., *Overview of the Scalable Video Coding Extension of the H.264/AVC Standard*, IEEE Transactions on Circuits and Systems for Video Technology, v.17, n.9 (September 2007)

SAMSUNG-1019 ITU-T Recommendation H.262 (02/2000)

SAMSUNG-1020 ITU-T Recommendation H.263 (02/98)

SAMSUNG-1021 Reader, Intraframe and Interframe Adaptive Transform Coding, EFFICIENT TRANSMISSION OF PICTORIAL INFORMATION, SPIE v.66, 108 (1975)

SAMSUNG-1022 Reader, MPEG4: Coding for Content, Interactivity, and Universal Accessibility, OPTICAL ENGINEERING 35(1), 104 (January 1996

SAMSUNG-1023 Reader, MPEG Patents, in MPEG VIDEO COMPRESSION STANDARD, Chapter 16, at 357-362 (Mitchell et al ed.) (1996)

SAMSUNG-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)

SAMSUNG-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)

SAMSUNG-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)

SAMSUNG-1027 ISO/IEC 14496-2: 2004 (MPEG4)

SAMSUNG-1028 ISO/IEC 13818-2: 1996 (MPEG2)

SAMSUNG-1029	Samsung Stipulation Letter Regarding IPR grounds in District Court Litigation
SAMSUNG-1030	Memorandum, Interim Procedure for Discretionary Denials in AIA Post-Grant Proceedings with Parallel District Court Litigation (USPTO June 21, 2022) ("Director's Guidance")
SAMSUNG-1031	U.S. Patent No. 6,292,512 to Radha et al.
SAMSUNG-1032	U.S. Patent No. 6,697,426 to Van Der Schaar et al.
SAMSUNG-1033	U.S. Patent No. 5,886,736 to Chen
SAMSUNG-1034	Apostolopoulos et al., *Video Streaming: Concepts, Algorithms, and Systems*, HP Laboratories (Sept. 18, 2002)
SAMSUNG-1035	U.S. Serial No. 11/479,077 to Demircin, Application As Filed for Demircin916 ("Demircin077")

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]	[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,
[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.

Samsung Electronics Co., Ltd. ("Petitioner" or "Samsung") 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

Petitioner certifies that the '303 Patent is available for IPR. This Petition is being filed within one year of service of a complaint against Samsung. Samsung is not barred or estopped from requesting review of the Challenged Claims on the grounds identified below.

B. Challenge and Relief Requested

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

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

Grounds 1-2 are further supported by the expert testimony of Dr. Clifford Reader, Ph.D. *See, generally,* SAMSUNG-1003. Dr. Reader provides a detailed overview of the state of the art before the '303 Patent in paragraphs ¶¶30-100.

The '303 Patent claims priority to a Japanese patent application filed on

March 31, 2014 ("the Critical Date"). SAMSUNG-1001, Face. The prior art references forming the basis of Grounds 1-2 pre-date the Critical Date. Petitioner does not concede that the Challenged Claims are entitled to the benefit of Critical Date.

Reference	Effective Filing Date	Publication Date	AIA Prior Art
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. Petitioner submits that no formal construction are presently necessary for purposes of demonstrating obviousness of the Challenged Claims based on Grounds 1-2. *Wellman, Inc. v. Eastman Chem. Co.*, 642 F.3d 1355, 1361 (Fed. Cir. 2011).

Petitioner reserves the right to respond to any constructions offered by Patent Owner or adopted by the Board. Petitioner is not conceding that the Challenged Claims satisfy all statutory requirements, nor is Petitioner waiving any arguments concerning indefiniteness or claim scope that can only be raised in district court or otherwise outside the context of an IPR. *Mylan Pharm. Inc. v. Horizon Pharma USA, Inc.*, IPR2018-00272, Paper 35, 6-8 (PTAB 2019). The Board has regularly compared indefinite claims to the prior art for purposes of considering unpatentability. *See Chicago Mercantile Exchange, Inc. v. 5th Market, Inc.*,

CBM2013-00027, Paper 33, 3 (PTAB 2014); *SAP America, Inc. v. Lakshmi Arunachalam*, CBM2013-00013, Paper 61, 29 (PTAB 2014). For this Petition, Petitioner applies prior art in a manner consistent with disclosure of the claimed features in the '303 Patent itself and Patent Owner's allegations of infringement before the district court.

D. Level of Ordinary Skill in the Art

A person of ordinary skill in the art ("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 computer science, electrical engineering, computer engineering, or a related field, with 2-3 years of experience in digital video processing, or a related field. SAMSUNG-1003, §IV. Additional graduate education could substitute for professional experience, and *vice versa*. *Id*.

II. THE '303 PATENT

A. Brief Description

The '303 Patent describes "video image coding data transmission technology" for hierarchically coded video streams. SAMSUNG-1001, 1:18-19; SAMSUNG-1003, §V. In the'303 Patent, hierarchical video coding refers to a coding structure in which pictures are coded either as "basic hierarchical pictures" in a "basic hierarchy" or as "supplemental hierarchical pictures" in a "supplemental hierarchy." SAMSUNG-1001, 3:63-67. Pictures in the basic hierarchy are decoded without reference to pictures in the supplemental hierarchy, while pictures in the

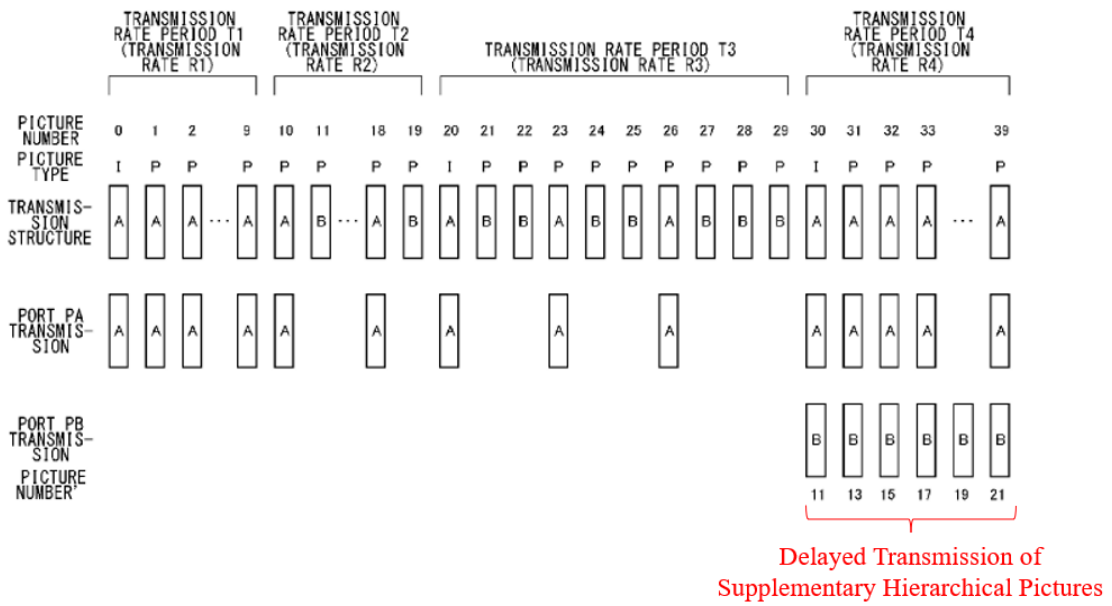
supplemental hierarchy are decoded with reference to other pictures in the supplemental hierarchy and/or pictures in the basic hierarchy. SAMSUNG-1001, 9:57-67 ("limiting basic hierarchical pictures to refer to only basic hierarchical pictures allows for decoding in the presence of only basic video image coding data"), 3:63-67. Supplementary hierarchical pictures generally have greater compression and are smaller in size than basic hierarchical pictures. *Id.*, 9:64-67 ("increases the coding efficiency"); SAMSUNG-1003, ¶233.

When the bandwidth of the transmission channel is insufficient to support transmission of all pictures in the stream as a basic hierarchical picture, the coder increases the proportion of pictures in the stream coded as supplementary hierarchical pictures and decreases the proportion of pictures coded as basic hierarchical pictures. SAMSUNG-1001, 9:37-48, 10:1-15. The basic hierarchical pictures are initially transmitted while the bandwidth is low, but transmission of the supplementary hierarchical pictures deferred until the bandwidth increases sufficiently. *Id.* Later, the supplementary hierarchical pictures are sent. *Id*., 4:34-42. The video stream can thus be reproduced initially at a lower quality or frame rate by transmitting and decoding only pictures in the basic hierarchy and can be reproduced later at a higher quality or frame rate by transmitting the supplementary hierarchical pictures in a delayed fashion so that the receiver can combine the supplementary hierarchical pictures with the earlier received basic hierarchical pictures.

Id., 11:22-36. For example, supplementary hierarchical pictures can be used to increase the frame rate or add color to basic hierarchical pictures in a video stream. SAMSUNG-1001, 12:6-25, 15:35-48, 10:25-37, 11:38-12:25.

Figure 5 below depicts an example transmission structure for pictures in a video stream. The "basic hierarchical pictures" (labeled "A") "are transmitted in real time from the port PA." SAMSUNG-1001, 11:55-56; *see also id.*, 9:27-36. During transmission rate periods T2 and T3, the channel bandwidth has decreased and the respective transmission rates R2 and R3 are unable to support transmission of all basic hierarchical pictures. *Id.*, 11:38-12:25. Accordingly, the coder changes the transmission structure of the video stream to code a subset of pictures in the stream during periods T2 and T3 as supplementary hierarchical pictures. *Id.* Only basic hierarchical pictures are transmitted during periods T2 and T3 to conserve channel bandwidth, while still allowing a receiver to decode and play the video in real-time at a lower quality. *Id.*

FIG.5



SAMSUNG-1001, FIG. 5 (annotated).

Later, during transmission rate period T4, "[w]hen there is a margin in the transmission rate (in a case where the transmission rate is R4), the supplementary hierarchical pictures" (labeled "B") whose transmissions were deferred during periods T2 and T3 "are transmitted from the port PB in the order stored without exceeding the transmission rate." SAMSUNG-1001, 11:56-60, FIG. 5. The transmitter "belatedly transmits supplementary video image coding data of the past when there is a margin in the transmission rate." SAMSUNG-1001, 4:40-42. Hence, "supplementary hierarchical pictures having coding order and display order that are earlier than those of basic hierarchical pictures transmitted in a given transmission interval are transmitted." *Id.*, 11:64-67. If the receiver stored the earlier

received (and displayed) basic hierarchical pictures, the receiver can then use the stored basic and received supplementary hierarchical pictures together to reconstruct and replay the video stream from periods T2 and T3 at a higher quality (e.g., at a higher frame rate or with additional color depth) at a later time.

The '303 Patent further discloses that "basic video image coding data (basic hierarchical GOP [group of pictures]) and supplementary video image coding data (supplementary hierarchical GOP) are set in units of GOPs" rather than being "set in units of pictures." SAMSUNG-1001, 12:27-36. As described by Dr. Reader, the use of GOPs for structuring hierarchical video data was a well known technique as early as 1993, having been described in MPEG standards, as well as in the patent literature. SAMSUNG-1003, ¶54-64, 126-134.

B. Prosecution History

The application (U.S. Serial No. 14/657,359) resulting in the '303 Patent was filed on March 13, 2015, by JVC Kenwood. The '303 Patent claims foreign priority to JP2014-071975, filed March 31, 2014. The '303 Patent issued on May 29, 2018. SAMSUNG-1001, Face; SAMSUNG-1003, ¶¶136-144.

The Examiner initially rejected claim 8[1] as obvious using Radha (U.S. 6,292,512) and Van Der Schaar (U.S. 6,697,426). SAMSUNG-1002, 74-75. The

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

Examiner contended that Van Der Schaar discloses an enhancement (supplemental) picture E2 having a coding order and display order earlier than a basic picture P3. *Id*. In response, the Applicant noted 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." SAMSUNG-1002, 84. The Applicant replied that Van Der Schaar was 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 using Radha and Chen (U.S. 5,886,736). SAMSUNG-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, the Examiner allowed claim 8. *Id.*, 178.

Significantly, the file history contains no indication that the Examiner ever considered any of the prior art cited in this Petition in connection with examination of the application that issued as 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 1 – 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][2]; SAMSUNG-1003, ¶¶146-151. Demircin explains that "high-quality AV data streams require a very high bandwidth continuously," but "time-varying channel conditions" result in unpredictable bandwidth" causing video quality degradation. Demircin, [0024].

Demircin addresses these challenges using Scalable Video Coding (SVC) techniques. Demircin, [0027]. "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.*, [0026]. Figure 1 below "illustrates an exemplary prediction structure, GOP structure and layering structure of a sample SVC bit stream" that includes a

[2] Demircin incorporates by reference SAMSUNG-1035 in its entirety. Demircin, [0063].

base layer 10 and enhancement layers 11, 12. *Id.*, [0036]. "[T]he base layer provides a minimum level of video quality," while the "enhancement layers provide additional quality in an incremental manner." *Id.*, [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.* The enhancement layers can be used to increase the signal-to-noise ratio (SNR) of coded video content, for example, although Demircin explains that its techniques also apply to other forms of scalability including temporal scalability (frame rate) and spatial scalability (picture resolution). *Id.*

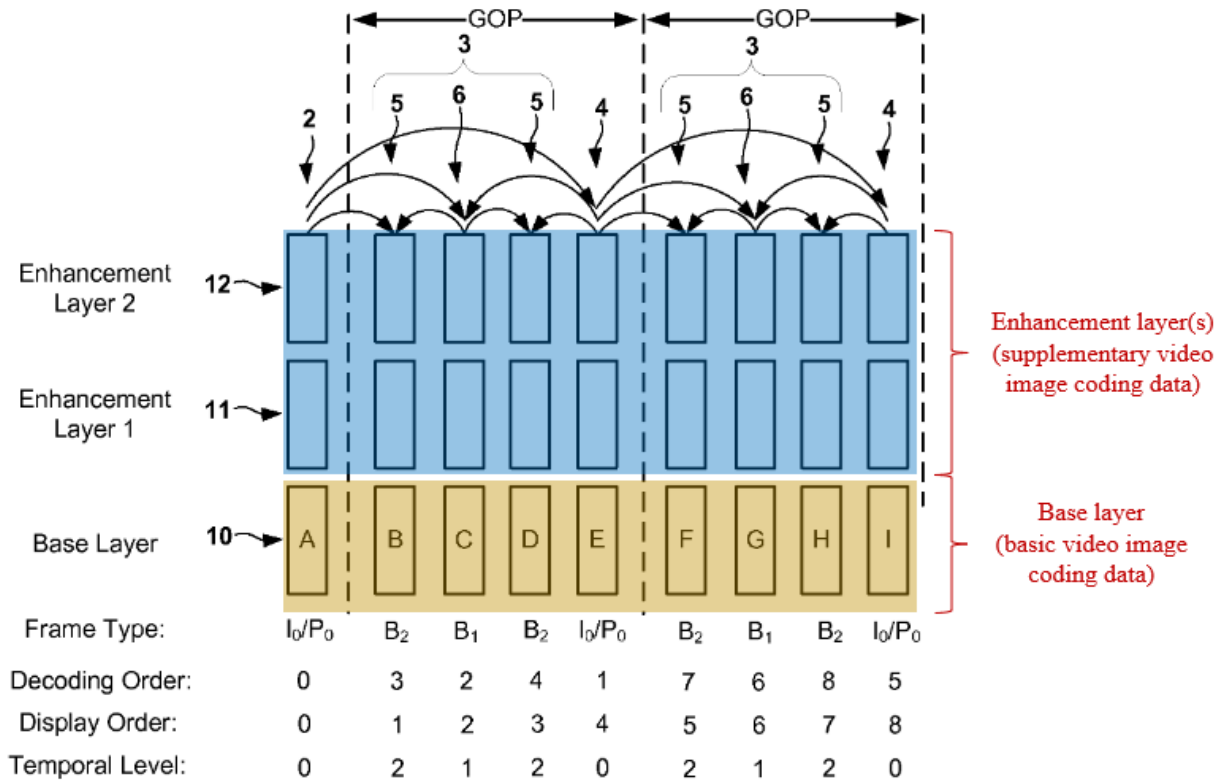
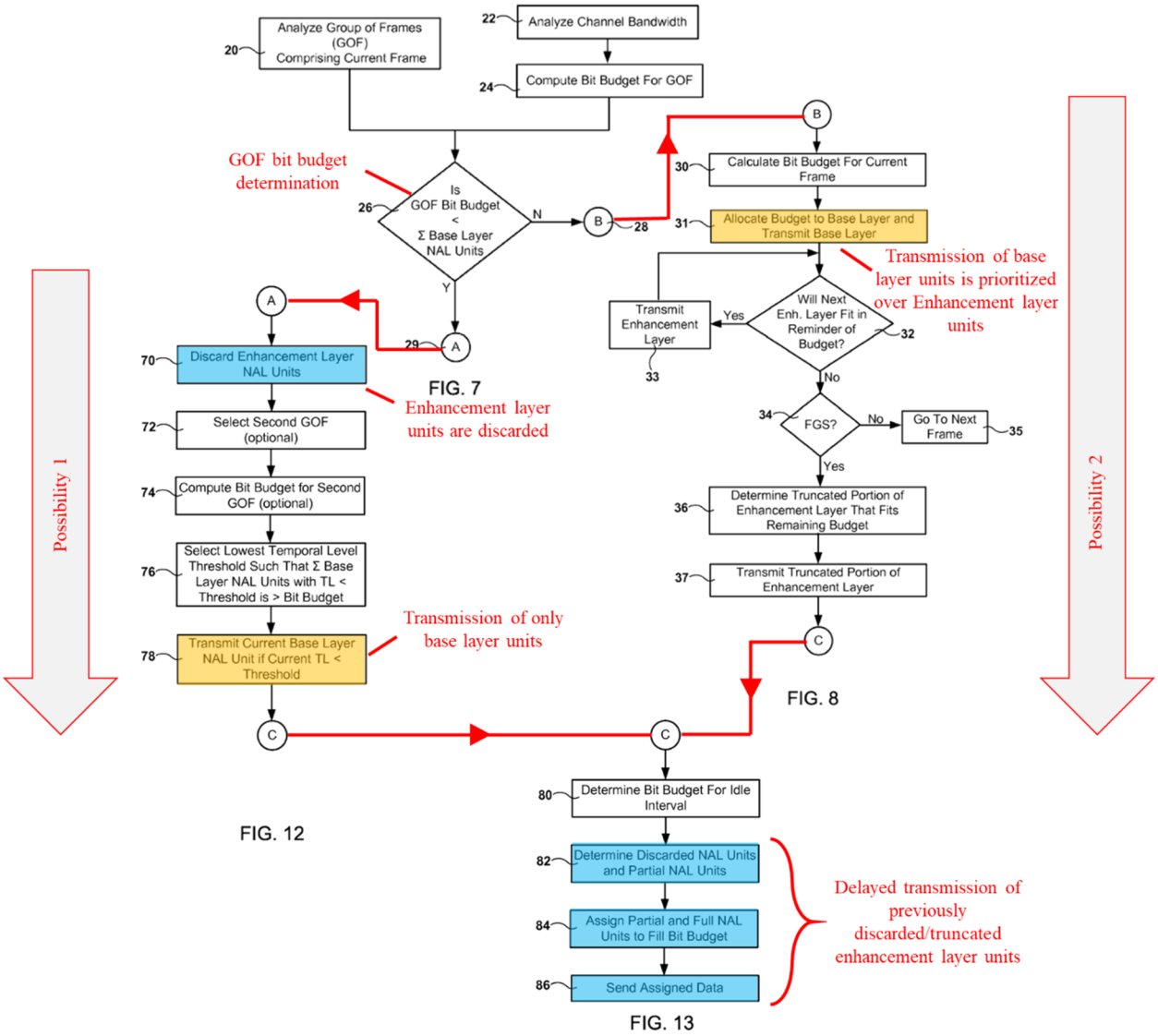


FIG. 1

Demircin, FIG. 1 (annotated); SAMSUNG-1003, ¶148.

Demircin's describes "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 entirely, and which parts should be transmitted partially (truncated), in order to meet time-varying rate and/or delay constraints." Demircin, [0027]. For example, Demircin's transmission 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, 8, 12, and 13 illustrate these concepts:



Demircin, FIGS. 7, 8, 12, 13 (annotated); SAMSUNG-1003, ¶151.

2. Analysis

(a) Claim 1

[1.P]

To the extent the preamble is limiting, Demircin discloses and renders obvious a video image coding data receiver in a video image coding transmission and reception system. SAMSUNG-1003, ¶¶162-164. 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"); SAMSUNG-1003, ¶¶162-164. 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].

Demircin also incorporates by reference Demircin077. Demircin, [0063]; SAMSUNG-1035. Through Demircin077, Demircin discloses additional details regarding video receivers. For example, with reference to FIG. 1, Demircin077 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* SAMSUNG-1035, [0052], FIG. 1; *see also id.*, [0055], [0195], FIG. 30 (depicting "receiver").

Demircin077 also teaches a transmitter "sends a packet burst to the receiver."

SAMSUNG-1035, [0194].

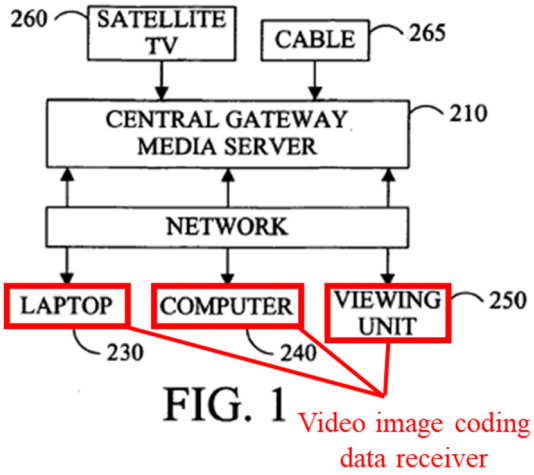


FIG. 1

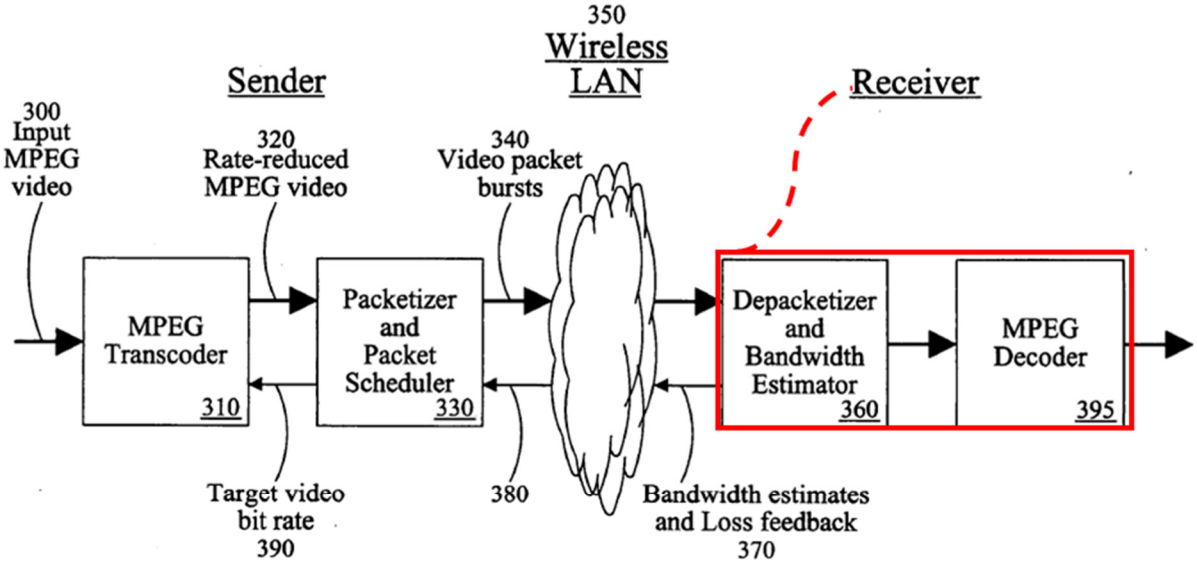


FIG. 30

SAMSUNG-1035, FIGS. 1, 30 (annotated); SAMSUNG-1003, ¶¶162-164.

[1.1]

Demircin discloses and renders obvious a "processor" in the video image coding data receiver as recited in [1.1]. SAMSUNG-1003, ¶¶165-166. 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] ("Elements of embodiments of the present invention may be embodied in hardware, firmware and/or software. While exemplary embodiments revealed herein may only describe one of these forms, it is to be understood that one skilled in the art would be able to effectuate these elements in any of these forms while resting within the scope of the present invention."); SAMSUNG-1003, ¶¶165-166.

Demircin further discloses through Demircin077 that "[d]ifferent devices interconnected to the network have different resources and different usage paradigms" including "***different microprocessors***." SAMSUNG-1035, [0055] (emphasis added); SAMSUNG-1003, ¶¶165-166. 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. SAMSUNG-1003, ¶¶165-166.

[1.2]

Demircin discloses and renders obvious a memory unit having instructions stored which, when executed by the processor, cause the processor to perform operations. SAMSUNG-1003, ¶¶167-168. 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 commonly and ordinarily included in "hardware, firmware and/or software" implementations of video image coding data receivers. Demircin, [0019]; SAMSUNG-1003, ¶¶167-168.

Demircin further discloses through Demircin077 that "[d]ifferent devices interconnected to the network have different resources and different usage paradigms" including "***different memory requirements***." SAMSUNG-1035, [0055] (emphasis added); SAMSUNG-1003, ¶¶167-168. 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. SAMSUNG-1003, ¶¶167-168.

[1.3]

Demircin discloses and renders obvious [1.3]. SAMSUNG-1003, ¶¶169-173. For example, Demircin discloses that the receiver receives network abstraction layer (NAL) units for base layer coding data (***basic video image coding data***), which were transmitted by a source/server (***video image coding data transmitter***). Demircin, [0131], SAMSUNG-1003, ¶169. Demircin discloses that the transmitter transmits base layer NAL units to the receiver. Demircin, [0129]-[0132], FIGS. 7-8 & 12. Demircin discards enhancement layer NAL units, when the "bit budget" for individual frames or the bit budget for a GOF is not less than required to transmit both the enhancement layer NAL units and the base layer NAL units. *Id*. The transmission of the base layer NAL units (i.e., with or without enhancement layer NAL units) implies the reception of the base layer NAL units at the receiver. SAMSUNG-1003, ¶169.

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.* "The SVC bitstream comprises NAL (Network Abstraction Layer) units similar to H.264/AVC." *Id.*, [0038]. "The base layer data of a video frame may be encapsulated into a NAL unit[] …." *Id.*

As shown in FIG. 1, Demircin teaches "a layered structure for each frame" that includes "a base layer 10 and two enhancement/refinement layers 11, 12." Demircin, [0036]-[0037]. The coded data in base layer 10 for one or more frames (pictures) corresponds to "***basic video image coding data***" as recited in [1.3]. SAMSUNG-1003, ¶171. Like the basic hierarchical pictures described in the '303 Patent, which are coded without reference to any pictures in the supplementary hierarchy, the pictures in Demircin's base layer 10 are coded without reference to any pictures in enhancement layers 11, 12. Demircin, [0035]-[0036]; SAMSUNG-1001, 9:57-67; SAMSUNG-1003, ¶171.

For temporal scalability, the pictures at temporal level 0 in Demircin's base layer 10 are I-pictures that can be coded without reference even to other pictures at higher temporal levels in base layer 10. Demircin, [0035]-[0036]; SAMSUNG-1001, 11:22-36; SAMSUNG-1003, ¶172. The base layer 10 data that is transmitted from the transmitter is received by the receiver. Demircin, [0024], [0061] ("received by the client"), [0062], [0063] ("deliver frames to the client"), [0119].

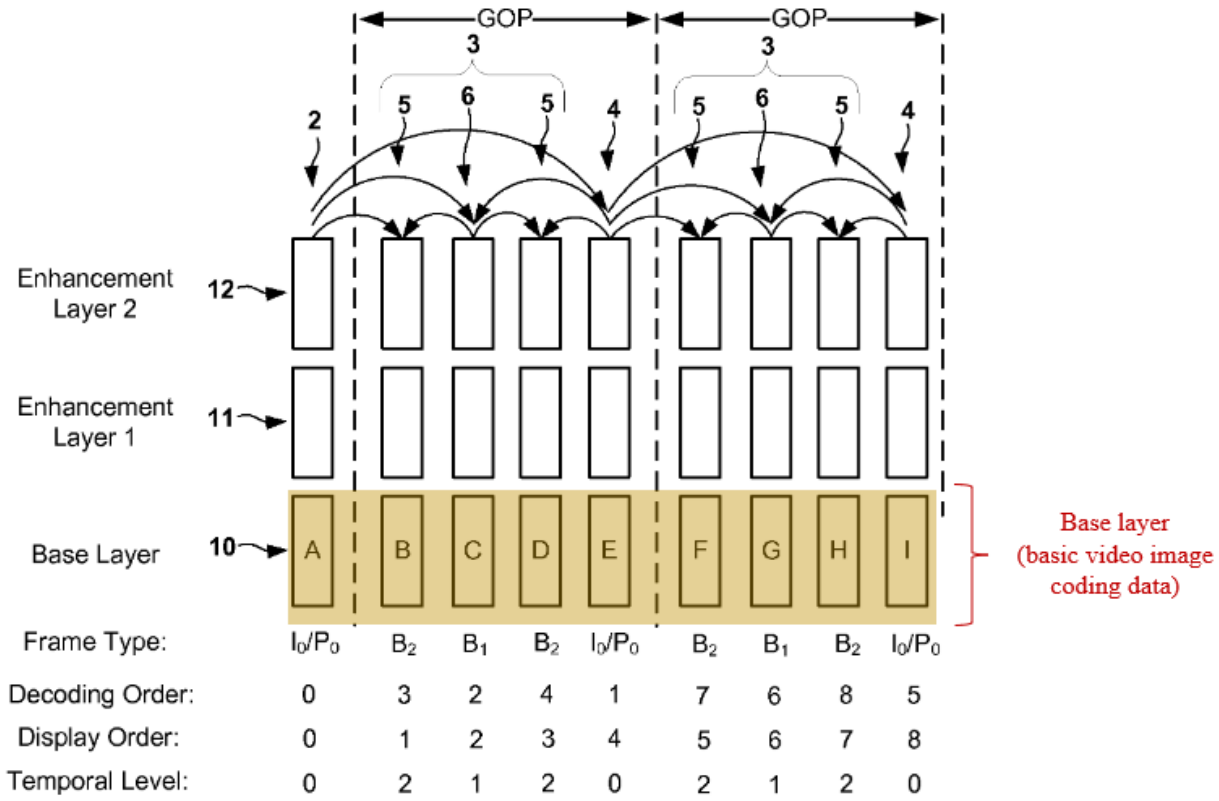


FIG. 1

Demircin, FIG. 1 (annotated); SAMSUNG-1003, ¶¶169-173.

[1.4]

Demircin discloses and renders obvious [1.4]. SAMSUNG-1003, ¶¶174-178; *supra* [1.3]. 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 the base layer NAL units for one or more video frames (***basic video image coding data***) to the exclusion of corresponding enhancement layer NAL units. Demircin, [0129]-[0132], FIGS. 7-8 & 12; SAMSUNG-1003, ¶174. In such instances, a POSITA would

have understood and it would have been obvious that the receiver would have reproduced the frames with a "minimum level of video quality" by decoding only the base layer NAL units for those frames without the corresponding enhancement layer NAL units. Demircin, [0035]; SAMSUNG-1003, ¶174.

Demircin discloses that the transmitter transmits base layer NAL units to the receiver. Demircin, [0129]-[0132], FIGS. 7-8 & 12. Demircin discards enhancement layer NAL units, when the "bit budget" for individual frames or the bit budget for a GOF is insufficient to allow for transmission of both the enhancement layer NAL units and the base layer NAL units. *Id*. The transmission of the base layer NAL units (i.e., with or without enhancement layer NAL units) implies the reception of the base layer NAL units at the receiver. SAMSUNG-1003, ¶175. As Demircin explains, the base layer data may be transmitted; and enhancement layer data may or may not be transmitted, depending on bit budget. Demircin, [0131]-[0132] (emphasis added); *also id.*, [0144] ("When all base layers do not fit in the GOF bit budget, all enhancement layer units may be discarded or omitted from transmission at the current time 70."). If the decoding deadline has expired, dropped enhancement layer NAL units are not transmitted. *See* Demircin, [0104]-[0105]. Therefore, the receiver would receive and decode only the base layer NAL units. *Id.*; SAMSUNG-1003, ¶175.

A POSITA would have understood, and it would have been obvious, that the receiver (decoding device) only would have decoded base layer data for a frame, (i.e., without combining such base layer data with enhancement layer data) if the receiver has only received base layer NAL units for the frame by the frame's decoding deadline to support streaming video. Demircin, [0063], [0105], FIGS. 1-8, 12; SAMSUNG-1003, ¶¶175-177. Indeed, this is evident from Demircin's explanation that "dropped or truncated NAL units" for enhancement layer frames will only be transmitted if their frames' "decoding deadlines have not yet expired." Demircin, [0105]. Demircin thus discloses and renders obvious the scenario where coded base layer data (***basic video image coding data***) is decoded so as to reproduce a video image, even without use of an enhancement layer data. SAMSUNG-1003, ¶¶175-177.

To the extent the breadth of [1.4] allows for decoding of received basic video image coding data together with supplementary video image coding data so as to reproduce a video image, Demircin further discloses and renders obvious such features as well. SAMSUNG-1003, ¶178. For instance, Demircin teaches that "[o]ne or more enhancement layers provide additional quality in an incremental manner, also referred to as progressive refinement," "[d]ecoding of the first enhancement layer requires availability of the base layer," and "[d]ecoding of each subsequent enhancement layer requires availability of all previous enhancement

layers, as well as the base layer." Demircin, [0035]; SAMSUNG-1003, ¶178. When sufficient bit budget exists for both the base layer and enhancement layer NAL units for a frame (***video image***) to be transmitted by the transmitter and received by the receiver before the frame's decoding deadline expires, the receiver decodes the base layer and enhancement layer NAL units to reproduce the frame at a higher quality than if only the base layer NAL units were decoded. Demircin, [0131] ("If the budget remainder is at least as large as the next most important enhancement layer, that enhancement layer may be allocated to the budget thereby creating an updated budget remainder and the enhancement layer may be transmitted 33. This process may be repeated until all enhancement layers for the current frame have been transmitted or until the next most important enhancement layer is larger than the updated budget reminder."), FIGS. 7-8; SAMSUNG-1003, ¶178.

[1.5.1]

Demircin discloses and renders obvious [1.5.1]. SAMSUNG-1003, ¶¶179-182; *supra* [1.3]-[1.4]. For example, Demircin discloses that the receiver receives network abstract layer (NAL) units for enhancement layer coding data (***supplementary video image coding data***), including enhancement layer NAL unit(s) for an individual frame or picture (***supplementary hierarchical picture***). Demircin, FIGS. 7-8 & 12-13 and accompanying text; *supra* [1.3]-[1.4].

Demircin discloses SVC. Demircin, [0035]. In SVC, "[o]ne or more enhancement layers provide additional quality in an incremental manner, also referred to as progressive refinement." *Id*. "Decoding of the first enhancement layer requires availability of the base layer" and "[d]ecoding of each subsequent enhancement layer requires availability of all previous enhancement layers, as well as the base layer." *Id.* "The SVC bitstream comprises NAL (Network Abstraction Layer) units similar to H.264/AVC." *Id.*, [0038]. "[T]he data for each enhancement layer of a video frame may be encapsulated into a NAL unit." *Id.*

In FIG. 1 below, Demircin illustrates "a layered structure for each frame" that includes "a base layer 10 and two enhancement/refinement layers 11, 12." Demircin, [0036]-[0037]; SAMSUNG-1003, ¶181. The coded data in either or both of enhancement layers 11, 12 for one or more frames (pictures) corresponds to the "***supplementary video image coding data***" in [1.5.1]; and for each frame, an enhancement layer picture ("***supplementary hierarchical picture***") is defined at each enhancement layer includes enhancement layer NAL unit(s) for a "***supplementary hierarchical picture***" for a given frame in the video image sequence. SAMSUNG-1001, 6:24-31. Like the supplementary hierarchical pictures described in the '303 Patent, which can be coded with reference to pictures in both the basic and supplementary hierarchies, the pictures in Demircin's enhancement layers 11, 12 can also be coded with reference to both pictures in both the base

layer 10 and other pictures in the enhancement layers 11, 12. Demircin, [0035]-[0036]; SAMSUNG-1001, 9:57-67; SAMSUNG-1003, ¶181.

In the special case of temporal scalability, the pictures at temporal levels 1 and 2 in Demircin's both base layer 10 and enhancement layers 11, 12 are bi-directionally inter-predicted with reference to other pictures at lower temporal levels. Demircin, [0035]-[0036]; SAMSUNG-1001, 11:22-36; SAMSUNG-1003, ¶182. In this case, the pictures at temporal levels 1 and 2 in base layer 10 can alternatively provide "***supplementary video image coding data including a supplementary hierarchical picture***" as recited in claim 1 and only the pictures at temporal level 0 in base layer 10 are characterized as "***basic video image coding data***" including a basic hierarchical picture. Demircin, [0035]-[0036]; SAMSUNG-1003, ¶182.

Data from enhancement layers 11, 12 that is transmitted from the transmitter is received by the receiver. Demircin, [0024], [0061] ("received by the client"), [0062], [0063] ("deliver frames to the client"), [0119]. Demircin's transmitter "transmits enhancement layer NAL units of the current frame provided that the enhancement NAL units of other frames at the same layer fit into the bit budget for the group," for example. Demircin, [0030]; SAMSUNG-1003, ¶182.

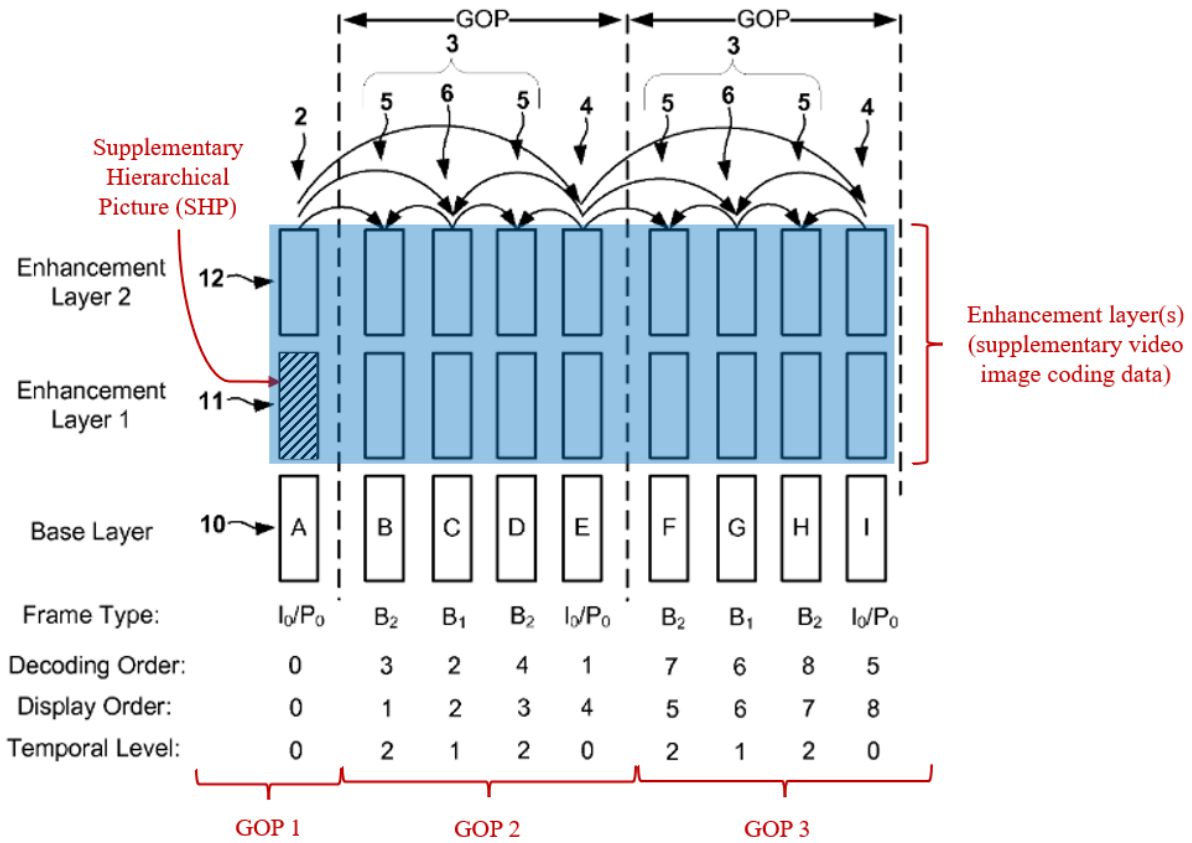


FIG. 1

Demircin, FIG. 1 (annotated); SAMSUNG-1003, ¶¶179-182.

[1.5.2]

Demircin discloses and renders obvious the features recited in [1.5.2]. SAMSUNG-1003, ¶¶197-202.

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 three B-frames in the GOP constitutes "***a plurality of inter prediction coded pictures***" as recited in [1.5.2]. *Id.*, [0036] ("A hierarchical B-frame 5 may be predicted from other B-frames as well as I-frames and P-frames while a standard B-frame 6 may only be predicted from I-frames and P-frames."). Demircin's description of I-, P, and B- frames is consistent with the description of intra-coded and inter-predicted coded pictures in the specification of the '303 Patent itself. SAMSUNG-1001, 9:19-23 ("Picture types include an I picture for performing intra coding, a P picture for performing inter prediction coding, and a B picture for performing inter prediction coding in a bidirectional manner."); SAMSUNG-1003, ¶184.

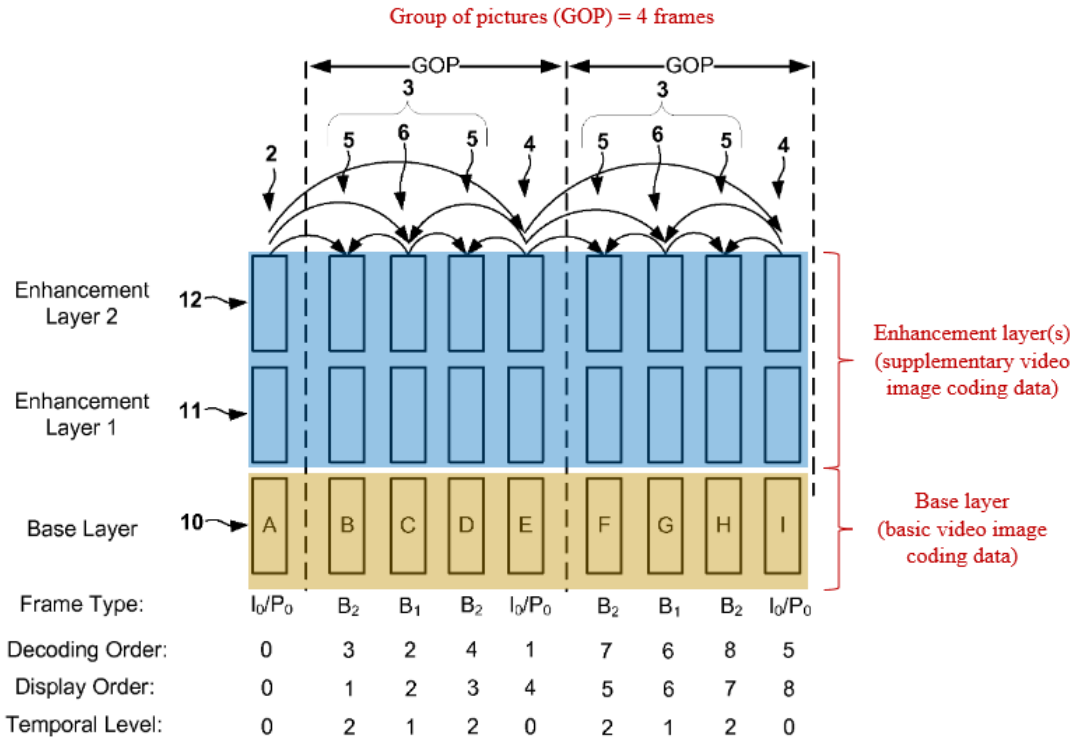


FIG. 1

Demircin, FIG. 1 (annotated); SAMSUNG-1003, ¶185.

Demircin also discloses and renders obvious 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. SAMSUNG-1003, ¶¶186. Consider the following illustration from FIG. 1 of Demircin, for example:

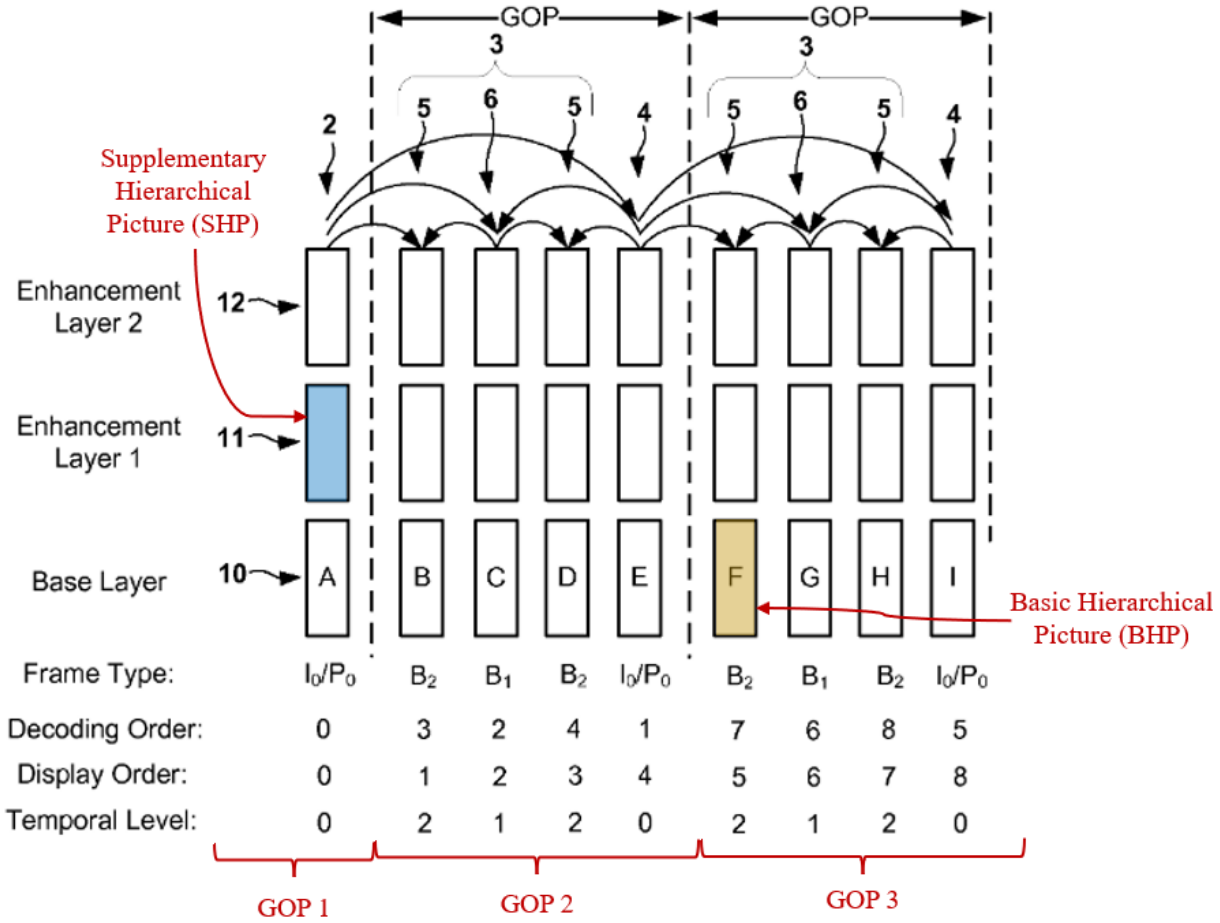


FIG. 1

Demircin, FIG. 1 (annotated); SAMSUNG-1003, 186.

As shown above, Demircin's prediction, GOP, and layering structure includes an enhancement layer picture for frame 0 in display order (***supplementary hierarchical picture*** ("SHP") indicated by blue shading and a base layer picture for frame 5 in display order (***basic hierarchical picture*** ("BHP")) indicated by gold shading. The blue-shaded SHP in this example has a display order that is earlier by a factor of a group of pictures than the gold BHP, which is evident from the display order annotations in FIG. 1. SAMSUNG-1003, ¶¶186. Specifically, the

blue-shaded SHP in the enhancement layer 11 has a display order index of "0" while the gold-shaded BHP in the base layer 10 has a display order index of "5." *Id.* The GOP size in FIG. 1 is just four frames, and thus the display order of the blue-shaded SHP is earlier by a factor of a group of pictures including an intra-coded picture and a plurality of inter-predicted pictures than the display order of the gold-shaded BHP. *Id.*

Likewise, the blue-shaded SHP in this example has a coding order that is earlier by a factor of a group of pictures than the gold-shaded BHP, which is evident from the decoding order annotations in FIG. 1. SAMSUNG-1003, ¶¶186-187. Demircin's decoding order indicates the order in which frames are decoded at the receiver and is both a type of "coding order" itself, and is the same as the order in which frames are encoded at the source and corresponds to the order in which frames are encoded as well. *Id*. Here, the blue-shaded SHP has a decoding order index of "0" while the gold-shaded BHP has a decoding order index of "7." *Id*. The GOP size in FIG. 1 is just four frames, and thus the decoding order of the blue-shaded SHP is earlier by a factor of a group of pictures including an intra-coded picture and a plurality of inter-predicted pictures than the decoding order of the gold-shaded BHP. *Id.*

Although the specification of the '303 Patent describes certain embodiments in which the transmission of a supplementary hierarchical picture for a frame is delayed by a period of time (e.g., multiple transmission rate periods) relative to transmission of a basic hierarchical picture (SAMSUNG-1001, 11:38-12:25 and FIG. 5), claim 1 recites no language that requires the "supplementary hierarchical picture" (SHP) in [1.5.2] to be received after the "basic hierarchical picture" (BHP) in [1.5.2]. SAMSUNG-1003, ¶188. Nonetheless, Demircin discloses and renders obvious [1.5.2] and the features of claim 1 as a whole regardless of whether the receiver must receive the SHP after the BHP. *Id.* This follows because Demircin discloses both scenarios in which (1) an enhancement layer NAL unit for a frame (SHP) is transmitted, and hence received, without delay in the ordinary course of transmission/reception for the GOP to which the frame belongs, as well as scenarios in which (2) the enhancement layer NAL unit for a frame (SHP) is transmitted, and hence received, on a delayed basis by well over a factor of a GOP. *Id*.

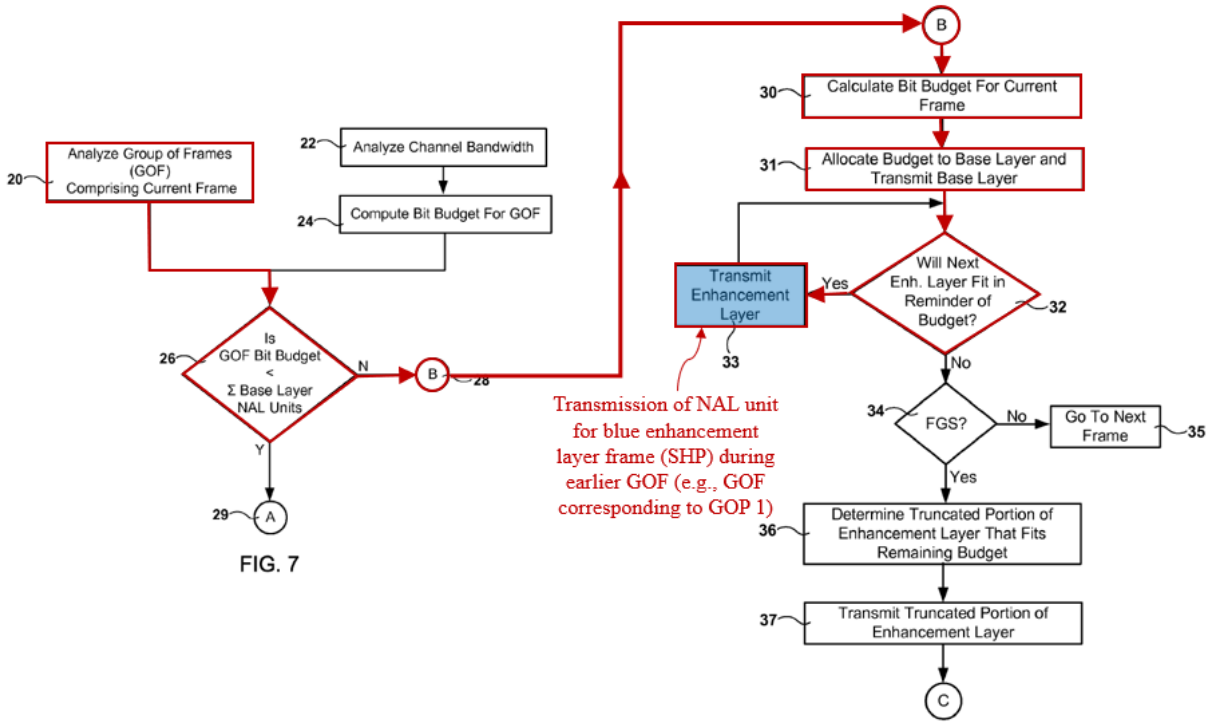
Demircin's First Example Scenario: Non-Delayed Transmission of Enhancement Layer NAL Unit(s)

Considering the ***first example scenario*** in Demircin in which an enhancement layer NAL unit for a frame (SHP) is transmitted, and hence received, without delay in the ordinary course of transmission/reception for the GOP to which the frame belongs, a POSITA would have understood and it would have been obvious

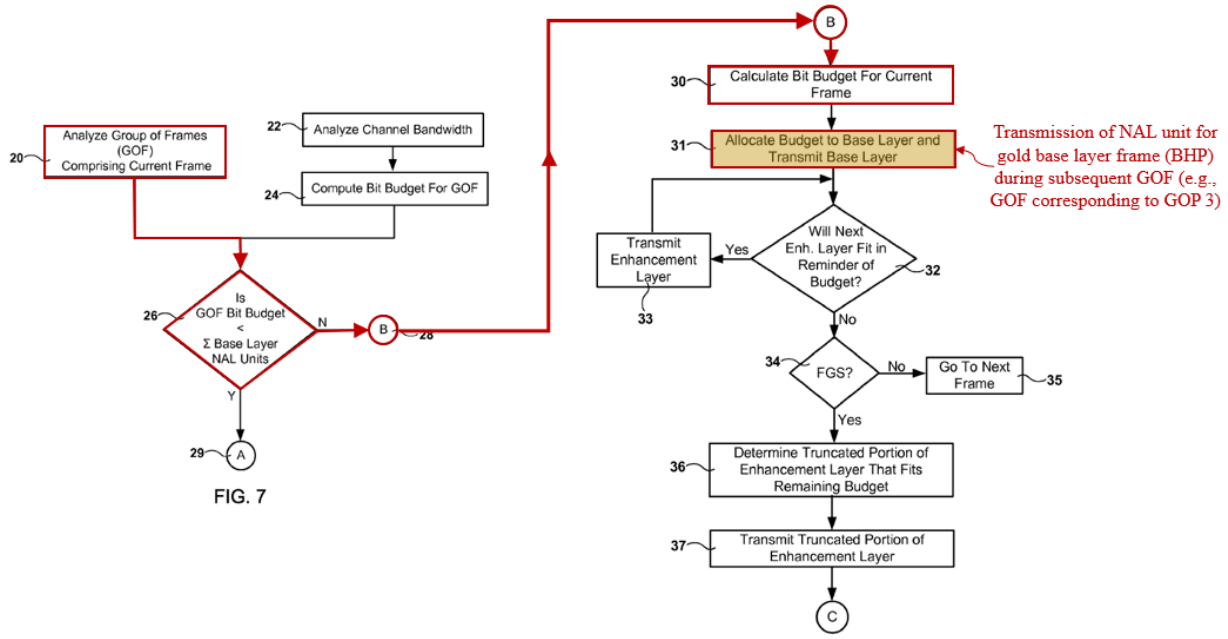
that the receiver would have received the blue-shaded SHP identified in the annotated version of FIG. 1 (above) ***before*** the gold-shaded BHP. SAMSUNG-1003, ¶¶189-190. For example, Demircin teaches that frames are grouped into groups of frames (GOFs) for purposes of computing "bit budgets" and setting 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 to the Network Abstraction Layer (NAL) units of the current frame 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].

Following the path through steps 20, 26, 28, 30, 31, 32, and 33 in FIGS. 7-8, Demircin discloses how an enhancement layer picture (SHP) for a GOF (such as that indicated by the blue shading in the annotated version of FIG. 1) can be transmitted without delay during an initial time interval allotted for transmission, provided that the bit budget for that frame's GOF is not less than the total size of the base layer NAL units in the GOF and provided that the bit budget allocated to that particular frame is of sufficient size to accommodate both the base layer and enhancement layer NAL units for that frame. Demircin, [0128]-[0135], FIGS. 7-8. Demircin's process later repeats for the base layer picture (BHP) like that indicated

by the gold shading in the annotated version of FIG. 1 above (and its GOF), and the gold base layer picture (BHP) would be subsequently transmitted following the path through steps 20, 26, 28, 30 and 31 in FIGS. 7-8, for example. *Id*.



Demircin, FIGS. 7 and 8 (annotated); SAMSUNG-1003, ¶190.



Demircin, FIGS. 7 and 8 (annotated); SAMSUNG-1003, ¶190.

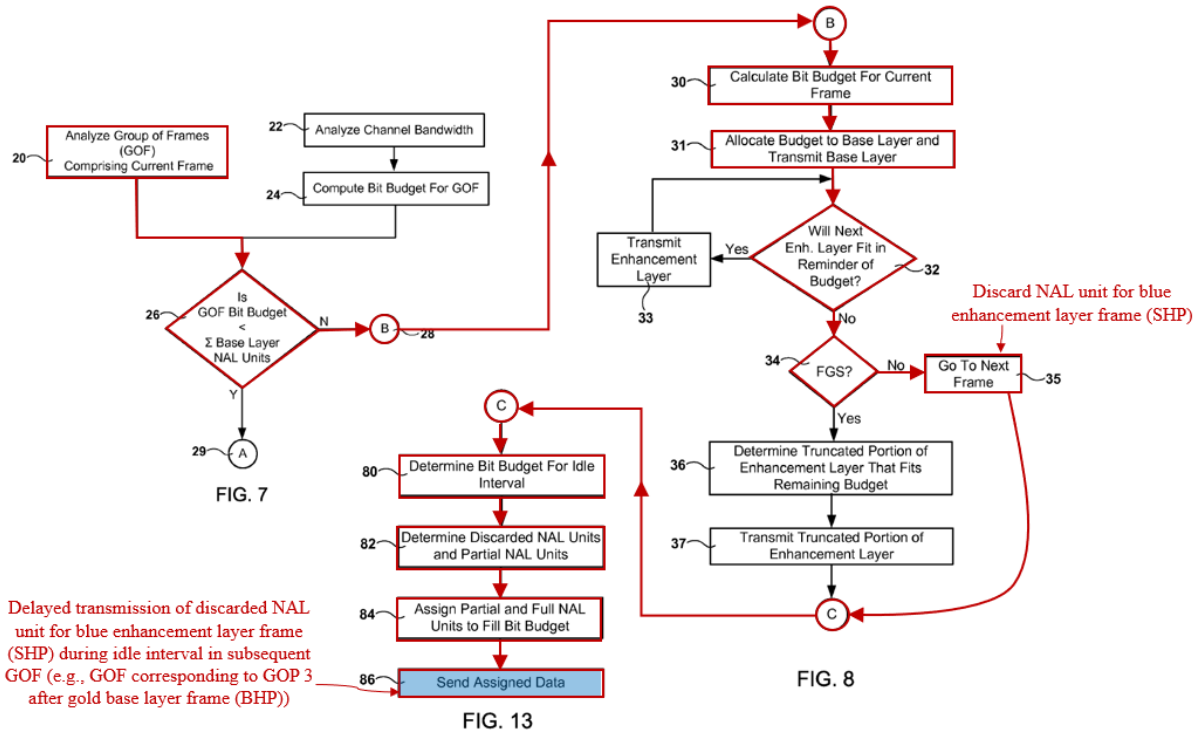
Demircin's Second Example Scenario: Delayed Transmission of Enhancement Layer NAL Unit(s)

Considering Demircin's ***second example scenario*** in which an NAL unit for an enhancement layer picture (SHP) is transmitted, and hence received, on a delayed basis after the ordinary time for transmission/reception of the GOP to which the enhancement layer picture belongs, a POSITA would have understood and it would have been obvious for the receiver to receive the blue-shaded SHP identified in the annotated version of FIG. 1 (above) ***after*** the gold-shaded BHP—particularly when transmission of the blue-shaded SHP is delayed by at least five frame transmission intervals. SAMSUNG-1003, ¶¶191-192.

Relevant to this scenario, Demircin describes techniques for "Delayed NAL Unit Transmission." Demircin, [0104]; [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."). In particular, 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]-[0149].

Thus, consistent with Demircin's description of "Delayed NAL Unit Transmission," the blue-shaded SHP identified in the annotated version of FIG. 1 (above) would be transmitted by the transmitter and received by the receiver ***after*** the gold-shaded BHP when the blue-shaded SHP is initially discarded due to insufficient bit budget, but is later transmitted/received in a delayed fashion during a suitable idle interval detected after transmission/receipt of the gold-shaded BHP (which is transmitted without delay). SAMSUNG-1003, ¶¶193194. Such delay in transmission by the transmitter, and hence receipt by the receiver, of the blue-shaded SHP would occur, for example, following the path through steps 20, 26, 28,

30, 31, 32, 34, and 35 of Demircin's FIGS. 7-8 for the blue-shaded SHP—which results in the blue enhancement layer frame initially being discarded/not transmitted during its ordinary transmission interval—and then following the path through steps 80, 82, 84, and 86 in FIG. 13 for delayed transmission of the previously discarded blue-shaded SHP during a detected idle interval. Demircin, [0128]-[0134], [0148].



Demircin, FIGS. 7-8, 13 (annotated); SAMSUNG-1003, ¶¶194-195.

The specific delay incurred by a discarded but later-transmitted enhancement layer NAL unit will depend on a number of 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 NAL unit

relative to other discarded or truncated NAL units that are eligible for delayed transmission. SAMSUNG-1003, ¶196. However, it would have been obvious to a POSITA that Demircin's system would allow for the discarded blue-shaded SHP to be transmitted in a delayed fashion during an idle interval *after* transmission of the gold-shaded BHP. *Id.* In particular, Demircin explains that discarded NAL units 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 when a frame that has been received by a decoding device is scheduled to be retrieved from a coded frame buffer and decoded. SAMSUNG-1003, ¶196; SAMSUNG-1034, 9-10, 15-16. If both the base and enhancement layer NAL units for a frame are available in the coded frame buffer by the decoding deadline for that frame, then the decoder may decode the enhancement layer picture with reference to the base layer picture (and potentially other pictures). *Id.* But if only the base layer NAL unit for a frame is available in the buffer by the frame's decoding deadline, then the decoder can only use the base layer NAL unit to decode that frame at a lower quality, for example. *Id.*

A POSITA would have known that decoding deadlines, as well as "play-back" or "display" deadlines, in streaming video systems depend in large part on how long the receiver buffers coded frame data before the frames are decoded and

displayed. SAMSUNG-1003, ¶197. Long buffering times typically lead to performance improvements and reduce the risk of coded frame data being unavailable by the time that a frame is to be decoded and displayed, albeit at the expense of delayed playback and larger buffer size requirements. *Id.* Notably, buffering durations of 5-15 seconds were common before the Critical Date. *Id.*; SAMSUNG-1034, 9-10, 15-16. With so much time available before a frame is decoded, Demircin's transmitter would have had ample ability to delay transmission of the NAL unit for a discarded enhancement layer picture (e.g., the blue SHP) past transmission of the NAL unit for a base layer picture (e.g., the gold BHP) whose coding order and display order are later by a factor of a group of pictures than the coding order and display order of the enhancement layer picture (e.g., the blue SHP) before the decoding deadline for the discarded enhancement layer picture (e.g., the blue SHP) expires. SAMSUNG-1003, ¶197. Thus, even when a frame's enhancement layer picture is transmitted well after that frame's base layer picture (e.g., delayed by ~5-15 seconds), the receiver of a streaming video program having a typical frame buffer can still receive the enhancement layer picture in sufficient time to decode it with reference to the earlier-received and buffered base layer picture before the decoding deadline expires. *Id.*

Demircin does not otherwise restrict use of its techniques for delayed NAL unit transmissions, despite noting that other embodiments where "NAL units may

have only one transmission opportunity at their first scheduled transmission … perform well when the initial buffering duration is relatively long, for example longer than 300 ms[] … ." Demircin, [0104]. This disclosure does not preclude use of delayed NAL transmissions when the initial buffering duration is relatively long, and in any event, a POSITA would have been motivated to use delayed NAL unit transmissions in many cases when the initial buffering duration is relatively long to provide additional opportunities to transmit enhancement layer pictures that could not be initially transmitted if decoding deadline has not expired (such that the receiver could receive and decode the enhancement layer picture to render higher quality video). SAMSUNG-1003, ¶198.

Still, Dermicin meets and renders obvious the requirements of claim 1 in the scenario where the buffering duration is shorter (e.g., less than 300 ms). SAMSUNG-1003, ¶199. In this case, if the enhancement layer NAL units for a frame are received in the buffer within 300 ms of arrival of the base layer picture for the frame, then the enhancement layer picture can be decoded for that frame. *Id.* It follows that, for transmission of frames in display order, the blue-shaded SHP would need to be delayed by five frame transmission intervals to be transmitted, and hence received, after the gold-shaded BHP. *Id.* At a frame rate for decoding and presentation of 30 frames-per-second (FPS), this would implicate a delay of

just ~167 ms—i.e., far below the 300 ms buffering duration and well before the decoding deadline for frame 0 in display and decoding order. *Id.* At a frame rate for decoding and presentation of 60 FPS, this would implicate a delay of just ~83.3 ms—i.e., again, far below the 300 ms buffering duration and well before the decoding deadline for frame 0. *Id.* Certainly, if a common buffering duration such as 5-15 seconds applied, the delays of ~167 or ~83.3 ms would be far less than the buffering duration and the decoding deadlines would not have expired. *Id.*

Similarly, to the extent frames are transmitted in decoding order, the blue-shaded SHP would need to be delayed by seven frame transmission intervals to be transmitted, and hence received, after the gold-shaded BHP. SAMSUNG-1003, ¶200. At 30 frames-per-second (FPS), this would require a delay of ~233.3 ms—i.e., still well below the 300 ms buffering duration and decoding deadline for frame 0. *Id.* At 60 FPS, this would require a delay of just ~116.7 ms—i.e., again, far below the 300 ms buffering duration and well before the decoding deadline for frame 0 in display and decoding order. *Id.*

While the GOP size depicted in FIG. 1 accommodates 4 frames, Demircin broadly contemplates other GOP sizes and structures as well. For example, Demircin further discloses through Demircin077 an arbitrarily sized GOP of length N_{GOP} (9 frames): "The length of a GOP in number of pictures is denoted by

N_{GOP}… in FIG. 4 where N_{GOP}=9 for illustration purposes." SAMSUNG-1035, [0069], FIG. 4; SAMSUNG-1003, ¶201.

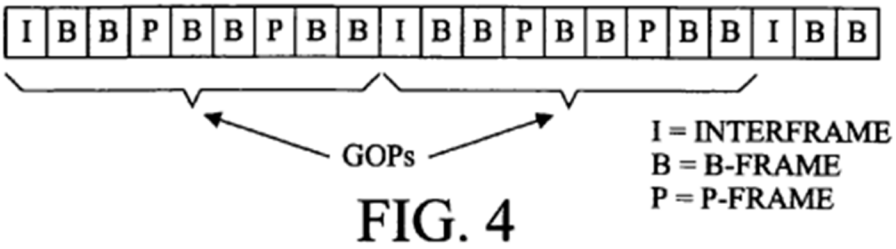


FIG. 4

Demircin077, FIG. 4.

[1.5.3]

Demircin discloses and renders obvious [1.5.3]. SAMSUNG-1003, ¶¶202-203. For example, in FIG. 1, the base layer and the enhancement layers are set in units of GOPs. *See also* SAMSUNG-1015, 54 (a GOP is "the smallest coding unit that can be independently decoded within a sequence"). 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].

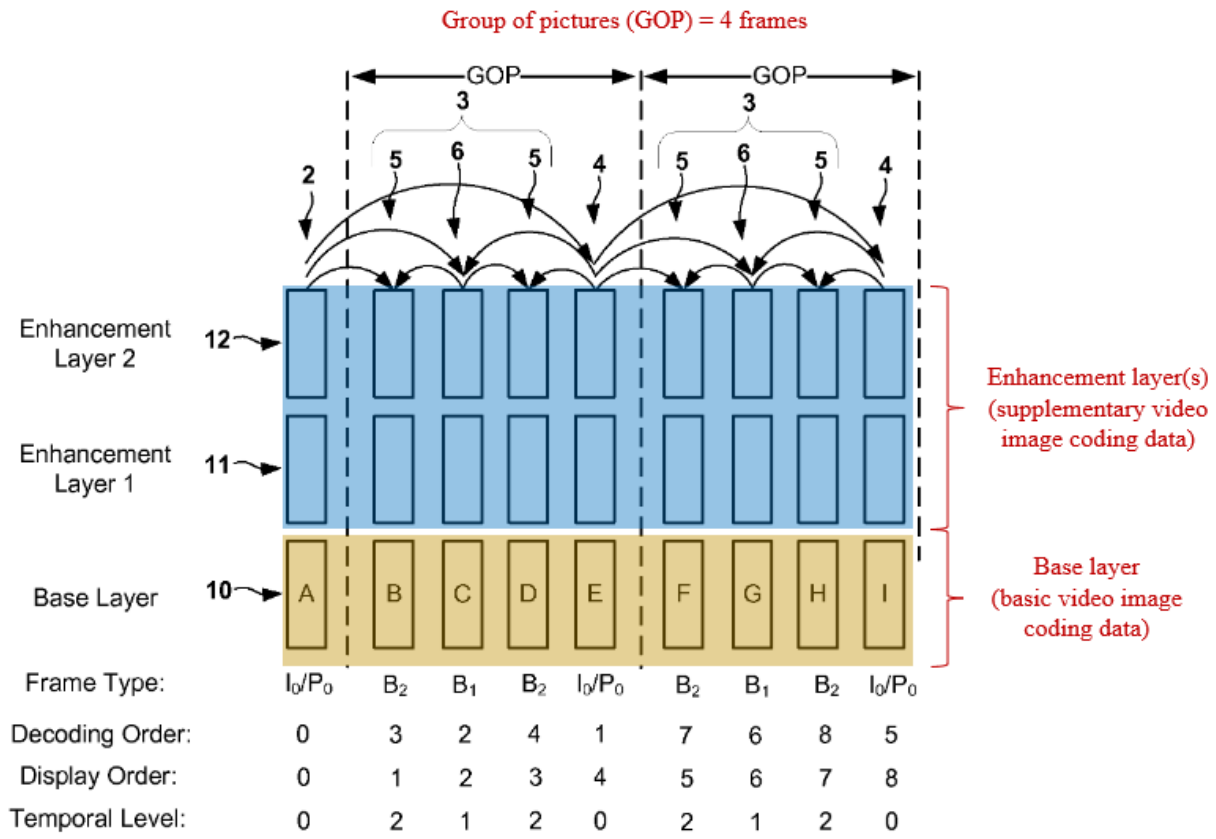


FIG. 1

Demircin, FIG. 1 (annotated).

"Coarse-grained SNR scalability is achieved using a layered structure for each frame, with a base layer 10 and two enhancement/refinement layers 11, 12 in this example." Demircin, [0037].

For temporal scalability, key pictures may be considered a base layer with temporal level 0. Demircin, [0036]. "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." *Id*.

In each case (SNR and temporal scalability), the basic and supplementary hierarchies are set in units of the group of pictures. Demircin, FIG. 1 (four frames across each of layers 10, 11, 12).

[1.6]

Demircin discloses and renders obvious [1.6]. SAMSUNG-1003, ¶¶204-208. 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***." SAMSUNG-1001, 7:34-43 (emphasis added). Thus, the '303 Patent contemplates acquiring basic video image coding data by retrieving it from memory. SAMSUNG-1003, ¶204.

Demircin discloses and renders obvious similar "acquiring" actions to those disclosed in the '303 Patent's specification. SAMSUNG-1003, 204. Demircin teaches receiving base layer NAL units (***basic video image coding data***) at an earlier time, storing the base layer NAL units in memory (e.g., a buffer), later receiving enhancement layer NAL units (***supplementary video image coding data***) corresponding to the base layer NAL units stored in memory, acquiring the base layer NAL units from memory (e.g., the buffer), and reconstructing video image coding

data using the stored and acquired base layer NAL units and the newly received enhancement layer NAL units. Demircin, [0131]-[0132], [0148]; SAMSUNG-1003, ¶204. The frames can be decoded and rendered to yield higher-quality pictures than if the enhancement layer NAL units had not been received. SAMSUNG-1003, ¶205.

For example, Demircin describes "Delayed NAL Unit Transmission" techniques. Demircin, [0104]; [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."). In particular, 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]-[0149]; SAMSUNG-1003, ¶206.

Demircin allocates bit budgets to groups of frames (GOFs) and individual frames within GOFs, and prioritizes transmission of base layer NAL units before enhancement layer NAL units. Demircin, FIG. 8 (steps 30-33) & [0131]-[1032]; *id.*, FIG. 7 (step 26) & [0130], FIG.12. If the GOF bit budget is less than the total size of all base layer NAL units in a GOF, Demircin discards all enhancement layer NAL

units in the GOF and they are eligible for delayed transmission in idle intervals, even as all or some of the base layer NAL units are transmitted earlier in scheduled intervals. *Id.*, [0144] ("When all base layers do not fit in the GOF bit budget, all enhancement layer units may be discarded or omitted from transmission at the current time 70."), [0129]-[0130], FIGS. 7, 12, 13. If the GOF bit budget can accommodate all base layer NAL units in a GOF but the bit budget for a particular frame is only sufficient to accommodate the frame's base layer NAL unit, then that frame's enhancement layer NAL unit(s) are discarded and provided for delayed transmission in idle intervals. *Id.*, [0132]-[0134]. In each of these examples, Demircin teaches that base layer NAL units for one or more frames (***basic video image coding data***) are transmitted before enhancement layer NAL units for the one or more frames (***supplementary video image coding data***). NAL units transmitted earlier in time will ordinarily and commonly be received at the receiver earlier in time than NAL units transmitted later, and in this context, it would have been obvious for Demircin's explicitly disclosed delay in transmitting enhancement layer NAL units to have resulted in a situation where the receiver receives the base layer NAL units for one or more frames before the corresponding enhancement layer NAL units for those frames. SAMSUNG-1003, ¶207.

Demircin further discloses that the receiver stores received NAL units for frames that are to be later decoded in a "playout buffer." Demircin, [0065] ("playout

buffer at the client"), [0096] ("playout buffer underflows"), [0104]. Demircin077 explains that "Buffering of audio/video data is necessary and commonplace in audio/video transmission systems" and "[t]he receiving (decoding) part of the system typically includes a buffer at the input of the decoder." SAMSUNG-1035, [0252], [0255] & FIG. 39 (decoder buffer 528). Demircin077 further teaches that "the receiver buffers the incoming data units from the wireless network before passing the data units to the decoder." *Id.*, [0261]. The receiver thus acquires data unit(s) for a frame from the buffer when the frame is to be reconstructed (e.g., including decoding operations), and it would have been obvious in such cases for Demircin's receiver to have acquired from the buffer base layer NAL units for one or more frames (***basic video image coding data***) that were received before the delayed enhancement layer NAL units for the one or more frames (***supplementary video image coding data***) received at the moment—including, for example, when the delayed enhancement layer NAL units are received just in time "at the moment" the one or more frames are to be decoded. SAMSUNG-1003, ¶208; *see also supra*, [1.5.1]-[1.5.2].

[1.7]

Demircin discloses and renders obvious [1.7]. SAMSUNG-1003, ¶209. As described above with respect to [1.6], it would have been obvious for Demircin's receiver to have acquired from the buffer base layer NAL unit(s) for one or more

frames (***basic video image coding data***) that were received before the delayed enhancement layer NAL units for those frames (***supplementary video image coding data***) received at the moment. *Supra*, [1.6]. Demircin also teaches Scalable Video Coding (SVC). Demircin, [0026], [0035]-[0036]. In SVC, "[a] base layer provides a minimum level of video quality, and may provide compatibility with H.264/AVC." *Id.*, [0035]. "One or more enhancement layers provide additional quality in an incremental manner, also referred to as progressive refinement." *Id.* "Decoding of the first enhancement layer requires availability of the base layer" and "[d]ecoding of each subsequent enhancement layer requires availability of all previous enhancement layers, as well as the base layer." *Id.* Thus, Demircin discloses and/or it would have been obvious that, when the receiver receives the base layer and enhancement layer NAL units for a frame before the decoding deadline, the receiver decodes the base layer and enhancement layer data, and reconstructs higher-quality video image coding data by combining use of the base layer and enhancement layer NAL units consistent with SVC techniques. *Id.*; SAMSUNG-1003, ¶208.

(b) Claim 2

[2.P]-[2.2]

Supra, [1.P]; SAMSUNG-1003, ¶¶162-164, 210.

[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]; SAMSUNG-1035, [0055]; SAMSUNG-1003, ¶211.

[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]; SAMSUNG-1035, [0055]; SAMSUNG-1003, ¶¶212.

[2.4]

Demircin teaches and renders obvious [2.4]. SAMSUNG-1003, ¶213. 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.*; SAMSUNG-1003, ¶212.

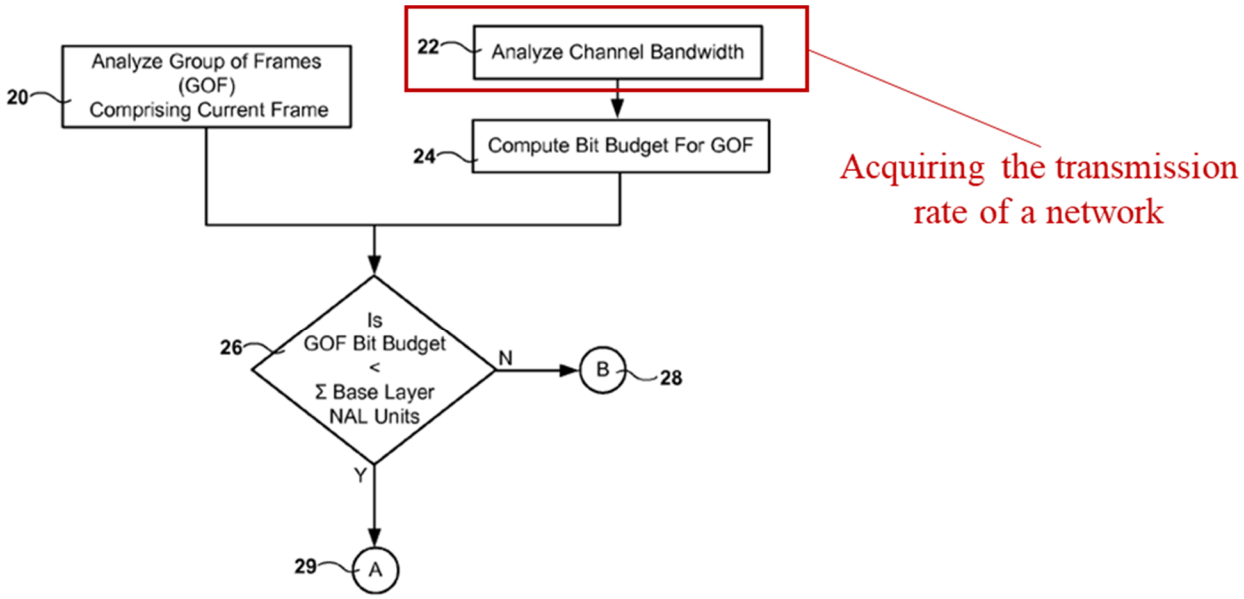


FIG. 7

Demircin, FIG. 7 (annotated).

[2.5]

Supra, [1.5.1]-[1.5.3]; SAMSUNG-1003, ¶214.

[2.6.1]

Supra, [1.3]; SAMSUNG-1003, ¶¶215.

[2.6.2]

It would have been obvious to a POSITA that the source/server of Demircin would have a memory for storing the NAL units for enhancement layer pictures (***supplementary video image coding data***), particularly in view of Demircin teaching

that enhancement layer pictures can be transmitted on a delayed basis in idle intervals. Demircin, [0061] (transmission buffer/queue), [0104]-[0106], FIGS. 12-13; SAMSUNG-1003, ¶215. 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]; SAMSUNG-1003, ¶216.

[2.7]

Supra, [1.5.1]-[1.5.3]; SAMSUNG-1003, ¶217.

[2.8]

Demircin's system determines whether to transmit (thus controls the transmission of) enhancement layer data for frames in a GOF, based on a bit budget for the GOF calculated based on the acquired bandwidth (i.e., transmission rate) of the network. Demircin, [0129], FIGS. 7-9; *supra*, [2.4]; SAMSUNG-1003, ¶¶218-219.

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, [0104]-[0122], [0123]-[0128], FIGS. 7-8, 12-14. With these techniques, "when the system detects there may be additional channel bandwidth available, transmission of NAL units that were previously dis-

carded or transmitted only partially may be considered and performed" and "[e]ligible NAL units may be reconsidered for transmission in delayed manner." *Id.*, [0034].

[2.9]

Supra, [1.P]; SAMSUNG-1003, ¶220.

[2.10]

Supra, [1.1]; SAMSUNG-1003, ¶221.

[2.11]

Supra, [1.2]; SAMSUNG-1003, ¶222.

[2.12]

Supra, [1.3]; SAMSUNG-1003, ¶223.

[2.13]

Supra, [1.4]; SAMSUNG-1003, ¶224.

[2.14]

Supra, [1.5.1]-[1.5.3]; SAMSUNG-1003, ¶225.

[2.15]

Supra, [1.6]; SAMSUNG-1003, ¶226.

[2.16]

Supra, [1.7]; SAMSUNG-1003, ¶227.

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

1. Overview of Kimoto

Kimoto describes 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]; SAMSUNG-1003, 152-157. Kimoto acknowledges 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].

Considering these issues, Kimoto's system achieves "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.*, [1018] & [0025]. "When reproducing the contents," Kimoto describes synthesizing the received upper layer data and the corresponding part of the stored lower layer data "to generate high quality data." Kimoto, [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 teaches scalability and hierarchical coding. Kimoto, [0037] (describing time, space, and image quality scalability).

2. Combination of Kimoto and Demircin

Kimoto describes synthesizing received hierarchically coded upper layer data with stored hierarchically coded lower layer data to reconstruct a high-quality video image. Kimoto, [0037]-[0045], [0069]-[0070]. Accordingly, Kimoto teaches substantially the same features as claim 1 of the '303 Patent. SAMSUNG-1003, ¶¶114-125, 152-157, 228-285. Kimoto does not expressly disclose setting the hierarchically coded lower and upper data as GOPs. SAMSUNG-1003, ¶229. Doing so, however, would have been an obvious modification to Kimoto's system, as described in more detail below. *Id*., 228-285.

In the combination, the hierarchically coded lower and upper data of Kimoto are set in GOPs, like the GOP structure taught by Demircin's Figure 1. SAMSUNG-1003, 229. A POSITA would have found it obvious to modify the lower and upper layer data of Kimoto to be set in units of GOP, in a manner similar to Demircin, to achieve the predictable results of accurate and efficient synthesis of newly received upper layer data with stored lower layer data, and to support delayed playback described in Kimoto. *Id*.

3. Reasons to Combine Kimoto and Demircin

As explained by Dr. Reader, there are several reasons a POSITA would have

been motivated to modify Kimoto at the time so that Kimoto's lower and upper layer data would be organized in units of GOPs consisting of an intra-coded picture (e.g., an I-frame) and a plurality of inter-prediction coded pictures (e.g., B-frames), consistent with the structure depicted in FIG. 1 of Demircin, for example. SAMSUNG-1003, ¶¶229-243. As discussed at length by Dr. Reader, setting hierarchical pictures into GOPs was well known and used as early as 1993, having been described in MPEG standards as well as in various patents. *Id*. And Kimoto makes it clear that it is applicable to hierarchical coding methods including MPEG and H.264, which included GOPs. Kimoto, [0038]; SAMSUNG-1003, ¶¶54-64, 230. A POSITA would have found it obvious to set the frames of Kimoto into units of GOP because doing so would have been standard practice at the time, and for other reasons described herein. SAMSUNG-1003, ¶230.

As mentioned, GOP structure including one intra-coded picture (one I-frame) and a plurality of inter-prediction coded pictures (B-frames) were commonplace before the Critical Date. SAMSUNG-1003, ¶¶54-64, 126-134, 230-243. Demircin teaches such a GOP structure. Demircin, FIG. 1. A POSITA would have been motivated to turn to Demircin specifically because both Kimoto and Demircin concern prioritizing the transmission of base layer (or lower layer) data over the enhancement layer (or upper layer) data, and the synthesis of lower and upper layer data later in time. Kimoto, [0040]-[0041], [0067]; Demircin, [0129]-

[0132]; SAMSUNG-1003, ¶¶230-243. And both Kimoto and Demircin describe reconstructing high-quality data at the receiver by synthesizing upper layer data with corresponding lower layer data. Kimoto, [0069]; Demircin, [0035], [0131]. Demircin provides structural information in the form of GOPs that Kimoto can use to accomplish the same goal of prioritizing the transmission of lower layer data, while later transmitting upper layer data, and synthesizing lower and upper layer data to reconstruct a high-quality video. SAMSUNG-1003, ¶¶230-243. Applying Demircin's teachings of GOPs would have benefited Kimoto to achieve Kimoto's desired goals. *Id*.

Kimoto is silent as to the details of how the synthesis of lower and upper layer data works, and a POSITA would have been motivated to seek a solution, like that taught by Demircin, which includes setting the lower and upper layer data into GOPs. Kimoto, [0069]; SAMSUNG-1003, ¶¶230-243. First, using GOPs allows the stored lower layer data to be aligned quickly and efficiently with corresponding upper layer data for synthesis and reproduction. SAMSUNG-1003, ¶¶230-243. In addition, using GOPs aids Kimoto in decoding a video stream by allowing it to recognize sets of pictures that can be decoded without reference to any other pictures in the stream. *Id*. Therefore, a POSITA seeking to implement the prioritization scheme of Kimoto would have been motivated to seek out and use the GOP teachings of Demircin. *Id*. Moreover, Demircin's GOP applied to

Kimoto would include one intra-coded picture and a plurality of inter-prediction coded pictures, as claimed in the '303 Patent. Demircin, FIG. 1, [0036], SAMSUNG-1003, ¶¶54-64, 126-134, 230-243. Thus, the combination of Kimoto with Demircin's GOP renders the '303 Patent obvious and invalid. SAMSUNG-1003, ¶¶228-285.

In addition, organizing pictures in the lower and upper layer data of Kimoto into GOPs of size four, like in Demircin FIG. 1, would have allowed a POSITA to optimize one or more characteristics of the video stream (although smaller or larger GOP sizes would also have been suitable in the combination). SAMSUNG-1003, ¶¶230-243. Kimoto considers delay times and transmission bandwidths to determine the length of the upper layer first data and upper layer subsequent for transmission. Kimoto, claim 42. The selection of GOP size and IBP pattern would have amounted to an obvious design choice for Kimoto to achieve lower and upper layer data transmission, particularly in view of Kimoto's desire to compensate for reproduction delays caused, in part, by transmission bit rate limitations. Kimoto, [0047]-[0048], cl.42; SAMSUNG-1003, ¶¶230-243. Shorter GOPs can preserve more detail in high-motion data, for example, while longer GOPs improve the amount of compression that can be achieved since intra-coded frames occur less frequently. SAMSUNG-1003, ¶¶230-243; *see also* SAMSUNG-1015, 15-16 (describing flexible GOP length can be selected based on compression rate, desired

video quality, bandwidth for transmission, decoding time and resources, etc.). Thus, a POSITA would have known to size the GOP for Kimoto with transmission bandwidth in mind. SAMSUNG-1003, ¶¶230-243.

As discussed by Dr. Reader, GOPs were used by MPEG for random access, fast-forward playback, or fast and normal reverse playback. SAMSUNG-1015, 15-16; SAMSUNG-1003, ¶¶54, 230-243. Recall that Kimoto is about reconstructing a video using stored lower layer data and received upper layer data at a later time, after a first viewing of lower layer data. Kimoto, [0040]-[0041], [0069]. Thus, a POSITA would have recognized that using GOPs would have also aided the delayed viewing of the reconstructed video image data, e.g., by facilitating random access, fast-forward playback, or fast and normal reverse playback. SAMSUNG-1015, 16; SAMSUNG-1003, ¶¶54-64, 230-243. GOPs were designed to facilitate random access, and a POSITA trying to reconstruct a video for subsequent (and non-real-time) viewing would have recognized the advantages of GOP for implementing random access in delayed video playback. SAMSUNG-1003, ¶¶230-243. Thus, a POSITA would have been motivated to use GOPs, like that in Demircin, to support Kimoto's playback of reconstructed video data. *Id*.

A POSITA would have had a reasonable expectation of success setting the lower and upper layer data of Kimoto in GOPs like that of Demircin considering

that the use of GOPs was a widely-used video coding technique at the time. SAMSUNG-1003, ¶¶54-64, 126-134, 230-243. First, Kimoto already explains that its "content data" is "hierarchically coded" and "scalable," and is broadly applicable to MPEG and other standards. *Id*., [0037]-[0038]. Likewise, Demircin involves hierarchical scalable video coding and is applicable to MPEG standards. Demircin [0035] (citing H.264/MPEG AVC/SVC). As Dr. Reader explains, GOPs were described in MPEG standards as early as 1993 and were widely used for scalable vide coding before the Critical Date. SAMSUNG-1003, ¶¶54-64, 126-134, 230-243. Thus, it would have been well within the skills of a POSITA to structure the hierarchically coded lower and upper layer data in units of GOPs because doing so was widely accepted, standard practice at the time.

For at least these reasons, and for those further articulated by Dr. Reader, a POSITA would have found it obvious to modify Kimoto with the GOP of Demircin.

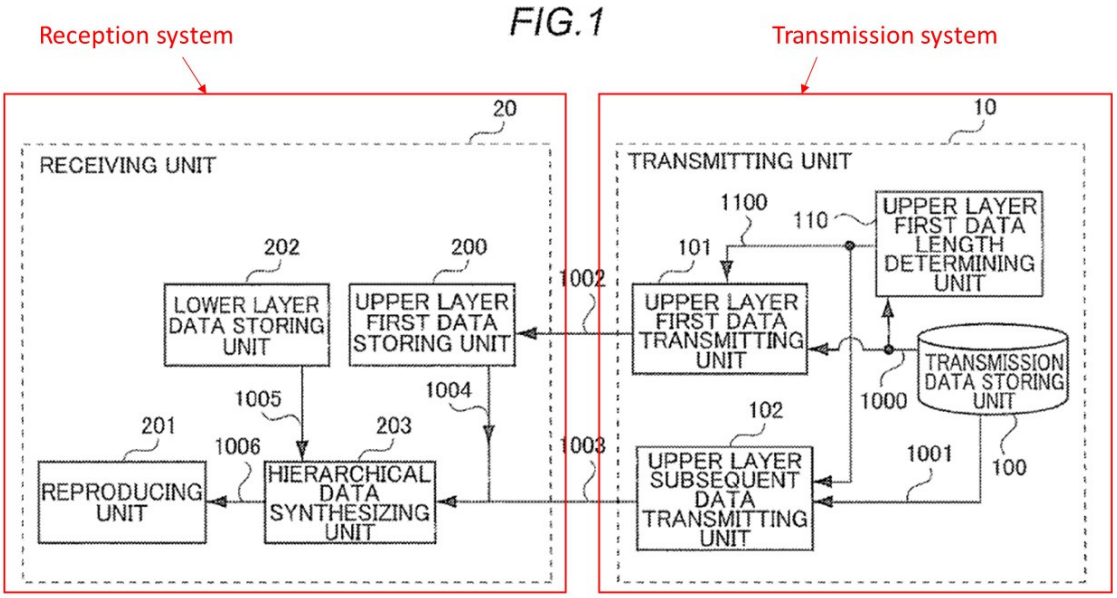
4. Analysis

(a) Claim 1

[1.P]

To the extent the preamble is limiting, the Kimoto-Demircin combination renders it obvious. SAMSUNG-1003, ¶¶245-249. 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 [1.1] obvious in view of Kimoto's teaching of a processor. SAMSUNG-1003, ¶¶250-252. 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. SAMSUNG-1003, ¶250.

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. SAMSUNG-1003, ¶251. 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. SAMSUNG-1003, ¶251.

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). SAMSUNG-1003, ¶250-252; *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 of the ’303 Patent. SAMSUNG-1003, ¶¶250-252.

[1.2]

The Kimoto-Demircin combination renders [1.2] obvious. SAMSUNG-1003, ¶¶253-256. 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. SAMSUNG-1003, ¶253.

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. SAMSUNG-1003, ¶¶254-255.

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 de-

scribed in Kimoto and provided in the Kimoto-Demircin combination. SAMSUNG-1003, ¶¶168, 256. 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 of the '303 Patent. *Id.*; *supra* §A.2-[1.2].

[1.3]

The Kimoto-Demircin combination renders [1.3] obvious. SAMSUNG-1003, ¶¶257-260. 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, stores, and even decodes the lower layer data. *Id.*; SAMSUNG-1003, ¶257. Lower layer data of Kimoto reads onto basic video image coding data because the lower layer

data of Kimoto can be decoded without the upper layer data. Kimoto, [0041]-[0042]; SAMSUNG-1003, ¶257.

The lower layer data of Kimoto is hierarchically coded according to international standards, including 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 and signal-to-noise ratio (SNR) scalability (also referred to as image quality scalability). Demircin, [0026]; Kimoto, [0037]; SAMSUNG-1003, ¶258.

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; SAMSUNG-1003, ¶¶172-173, 202-203, 259. As Demircin further explains, coarse-grained SNR scalability "is achieved using a layered structure for each frame, with a base layer 10 and two enhancement/refinement layers"—although "[m]ore or fewer enhancement layers may be used." Demircin, [0037]. The following drawing (Illustration 1) based on Demircin FIG. 1 illustrates an example of SNR scalability for a known frame structure in the Kimoto-Demircin combination, which includes (i) a base

layer containing lower layer or basic video image coding data and (ii) one or more enhancement layers containing upper layer or supplementary video coding data:

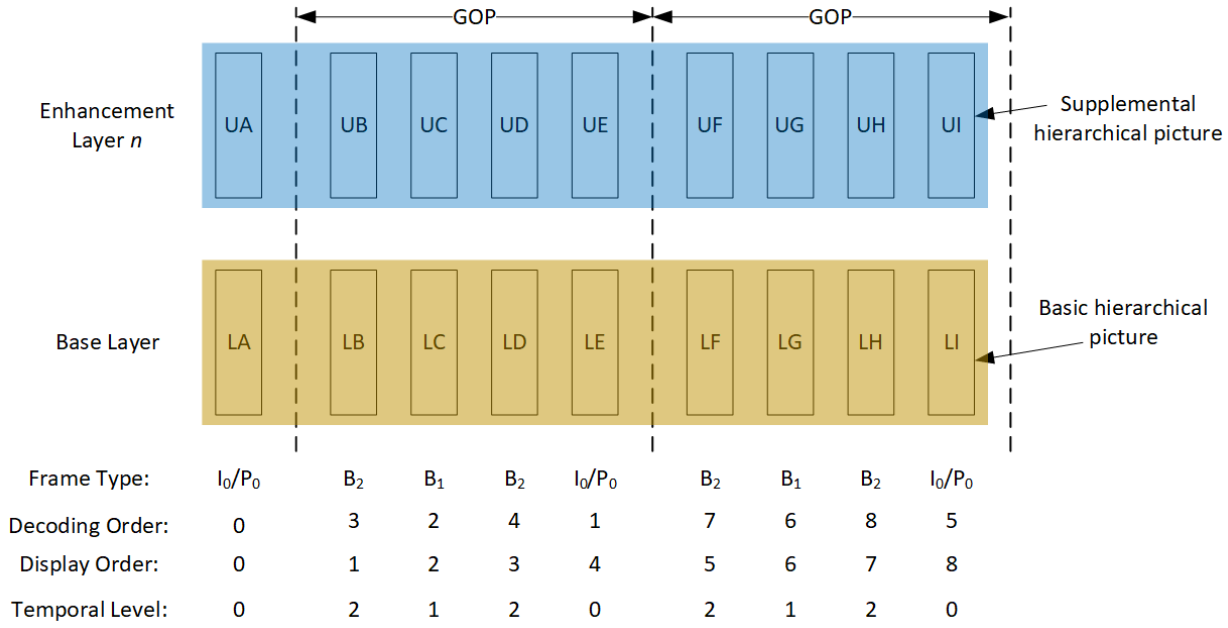


Illustration 1 of Kimoto-Demircin Combination

SAMSUNG-1003, ¶259 (based on Demircin, FIG. 1).

As shown in Illustration 1, the base (lower) layer includes an intra-predicted I-frame and a plurality of B-frames inter-predicted from one or more other frames. Demircin, [0036]; *see also id.*, FIG. 1. The I-frames and B-frames of the base layer are BHPs (gold). SAMSUNG-1003, ¶260. Demircin's disclosure of a base layer, which corresponds to the lower layer of Kimoto in the combination, is organized into groups of pictures (GOPs), where the "GOP size is equal to 4 frames." Demircin, [0036]. Thus, in the combination, the lower layer data of Kimoto is set in units of GOP, as taught by Demircin, where each GOP includes an intra coded

picture and a plurality of inter prediction coded pictures. SAMSUNG-1003, ¶260.

[1.4]

The Kimoto-Demircin combination renders [1.4] obvious. SAMSUNG-1003, ¶¶261-263. For 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").

A POSITA would have understood and found obvious that the receiving unit decodes the received lower layer data since coded video data is ordinarily, if not always, decoded for it to be viewed. SAMSUNG-1003, ¶¶261-263. Thus, in Kimoto, the lower layer data is decoded and viewed as low quality contents at one point in time; and later, after the upper layer data is received, the lower layer data is acquired from the lower layer data storing unit for reproduction as a higher quality content with the upper layer data. *Id.*; Kimoto, [0043], [0069].

[1.5.1]

The Kimoto-Demircin combination renders [1.5.1] obvious. SAMSUNG-1003, ¶¶264-270. 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]; SAMSUNG-1003, ¶¶264-265.

After the receiving unit receives, decodes, and reproduces 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.*, SAMSUNG-1003, ¶¶264-270. Thus, Kimoto discloses and the combination renders obvious receiving supplementary image coding data (i.e., the upper layer data). SAMSUNG-1003, ¶¶264-270.

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. SAMSUNG-1003, ¶¶267-268. 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 take different forms, including temporal scalability and signal-to-noise ratio (SNR) scalability (also referred to as image quality scalability). Demircin, [0026]; Kimoto, [0037]; SAMSUNG-1003, ¶¶267-268.

As Demircin explains, coarse-grained SNR scalability "is achieved using a layered structure for each frame, with a base layer 10 and two enhancement/refinement layers"—although "[m]ore or fewer enhancement layers may be used." Demircin, [0037]. The following drawing based on Demircin's FIG. 1 illustrates an example of SNR scalability for a known frame structure in the Kimoto-

Demircin combination, which includes (i) a base layer containing lower layer or basic video image coding data and (ii) one or more enhancement layers containing upper layer or supplementary video coding data:

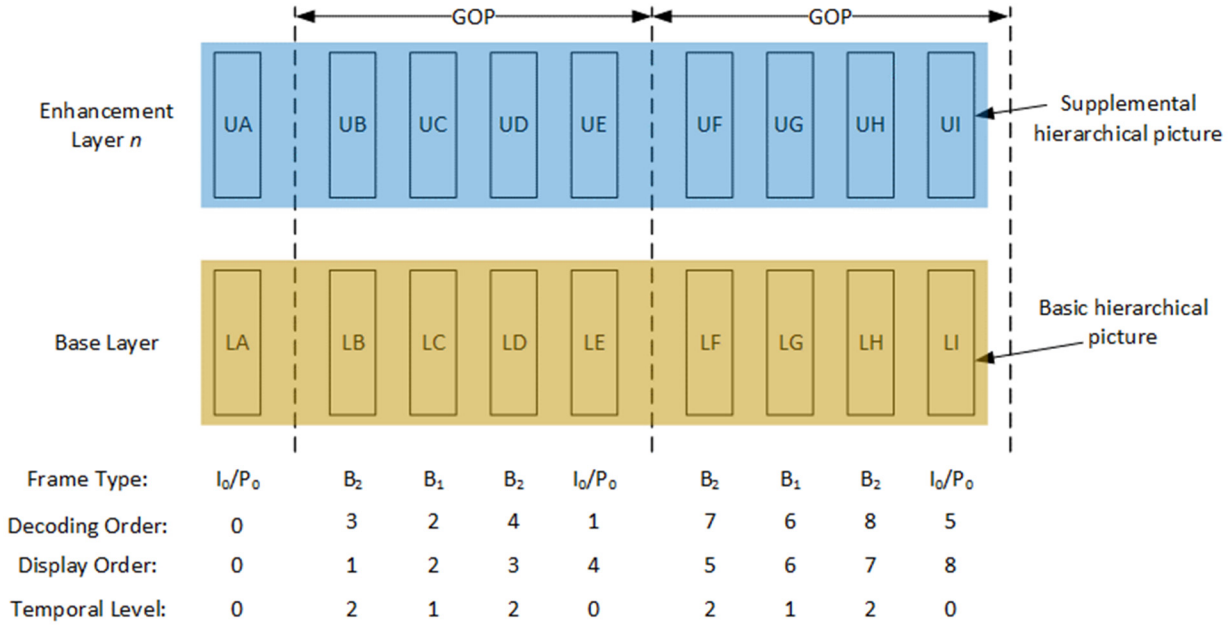


Illustration 1 of Kimoto-Demircin Combination.

SAMSUNG-1003, ¶269; *supra*, [1.3]-[1.4].

Illustration 1 shows one enhancement layer *n*, though Demircin teaches that more than one enhancement layer can be used in coarse-grain SNR scalability. Demircin, [0037]. The orange highlighted boxes are basic hierarchical pictures within the base layer and the blue highlighted boxes are supplementary hierarchical pictures within the enhancement layer. The illustrated enhancement layer corresponds to Kimoto's upper layer data and the claimed supplementary video image coding data. SAMSUNG-1003, ¶270. When performing decoding with SNR

scalability, decoding the "first enhancement layer requires availability of the base layer," and decoding "each subsequent enhancement layer requires availability of all previous enhancement layers, as well as the base layer." Demircin, [0035].

[1.5.2]

The Kimoto-Demircin combination renders [1.5.2] obvious. SAMSUNG-1003, ¶¶271-277. 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, prior to the reception of any upper layer data for those pictures, demonstrates 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. SAMSUNG-1003, ¶271.

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. SAMSUNG-1003, ¶272. Illustration 2 demonstrates this point for SNR scalability:

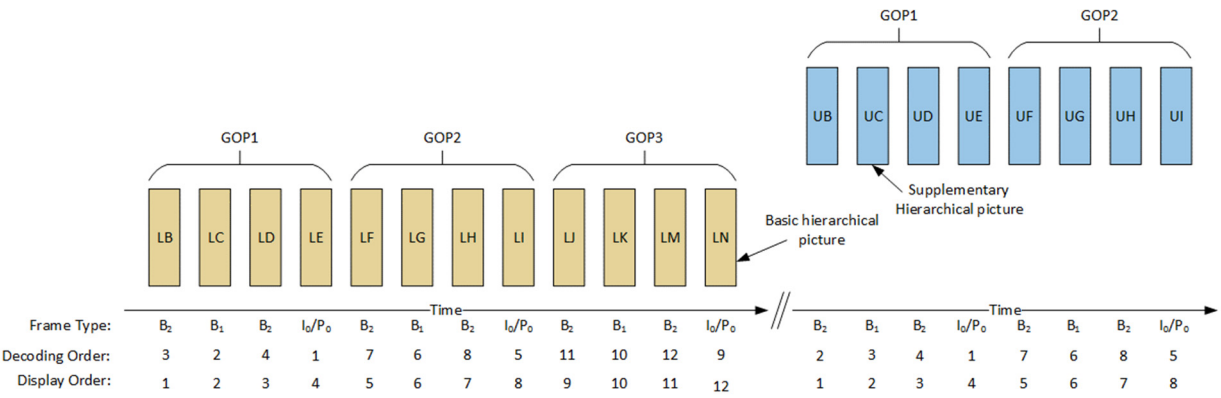


Illustration 2 of Kimoto-Demircin Combination

SAMSUNG-1003, ¶272.

In Illustration 2, lower layer data including basic hierarchical pictures (LB-N) for frames 1-12 are received and stored at an earlier time. Kimoto, [0041]-[0043]. In the combination, based on Demircin's frame structure, the lower layer data includes GOP1 with intra-coded picture LE and inter-prediction coded pictures LD, LC, and LB; GOP2 with intra-coded picture LI and inter-prediction

coded pictures LH, LG, and LF; GOP3 with intra-coded picture LN and inter-prediction coded pictures LM, LK, and LJ.[3] Later in time (as shown by the broken arrow), based on Kimoto's teachings, the upper layer data for frames corresponding to those in the lower-layer data is received (e.g., frames 1-8). Kimoto, [0067]. In the combination, based on Demircin's frame structure, the upper layer data includes GOP1 with intra-coded picture UE and inter-prediction coded pictures UD, UC, and UB; GOP2 with intra-coded picture UI and inter-prediction coded pictures UH, UG, and UF. Though not shown, a POSITA would have recognized that the upper layer data can also have included GOP3. SAMSUNG-1003, ¶¶272-273. As taught by Kimoto, however, the upper layer data can be "upper layer first data," which in this example includes GOP1 and GOP2. See Kimoto [0067]. Later, subsequent upper layer data is received in Kimoto, which can include GOP3. *See id.*, [0068].

As shown by Illustration 2, at least one (but really, multiple) upper layer pictures have a coding order and display order earlier by a factor of a GOP than at least one lower layer picture. SAMSUNG-1003, ¶274. For example, UC in GOP1

[3] The number of pictures in the lower layer data depends on the size of the video being reproduced, and the number shown in the figure is purely for illustration purposes. SAMSUNG-1003, ¶¶272-273.

is a supplemental hierarchical pictures that has a coding order of 3 and a display order of 2 while LN in GOP3 has a coding order of 9 and a display order of 12. GOP1 is separated by GOP3 by GOP2; more particularly, UC in GOP1 is separated from LN in GOP3 by the entirety of GOP2. *Id*. Thus, UC has a coding order and a display order earlier than LN by a factor of a GOP. SAMSUNG-1003, ¶¶271-274.

In the example, a GOP is equal to one I-frame and three B-frames, for a total of 4 frames. Demircin, [0037], FIG. 1. An upper layer picture (UC) having a coding order of 3 and a display order of 2 has a coding order and a display order more than one GOP (and four frames) from a lower layer picture (LN) having a coding order of 9 and a display order of 12. SAMSUNG-1003, ¶275.

All of the lower layer data shown in Illustration 2 is received and stored before any of the upper layer data is received. Kimoto, [0041]-[0043]. Therefore, it would have been obvious 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. SAMSUNG-1003, ¶271-277.

[1.5.3]

The Kimoto-Demircin combination renders [1.5.3] obvious. SAMSUNG-1003, ¶¶278-279. 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; SAMSUNG-1003, ¶¶172-173, 202-203, 278. 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]; SAMSUNG-1003, ¶278. According to MPEG, a GOP is "the smallest coding ***unit*** that can be independently decoded." SAMSUNG-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. SAMSUNG-1003, ¶¶202-203, 278.

[1.6]

The Kimoto-Demircin combination renders [1.6] obvious. SAMSUNG-1003, ¶¶280-284.

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]; SAMSUNG-1003, ¶283. 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." SAMSUNG-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 [1.7] obvious. SAMSUNG-1003, ¶285. Kimoto teaches that the receiving unit reconstructs (e.g., synthesizes) 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 "***synthesizes*** basic video image coding data and supplementary image coding data" using a "coding data reconstruction unit." SAMSUNG-1001, 12:53-55 (emphasis added).

IV. PTAB DISCRETION SHOULD NOT PRECLUDE INSTITUTION

A. 35 U.S.C. §325(d) – *Advanced Bionics*

This Petition presents new prior art and arguments for which the file history contains no indication the Office ever previously considered in connection with the '303 Patent. Indeed, none of the prior art that forms the basis of Grounds 1 or 2 of

this Petition—i.e., neither Kimoto nor Demircin—is identified on the face of the ’303 Patent as having been cited by either the applicant or the Examiner during original examination, and certainly neither reference was applied in a rejection or identified in the Examiner’s reasons for allowance.

Because the ’303 Patent’s original examination did not involve negotiation between the applicant and the Examiner regarding the teachings of the prior art applied in the instant Petition, this Petition’s art and arguments cannot be said to be the same as, or substantially similar to, the art or arguments previously presented to the Office in connection with the ’303 Patent. Petitioner notes that the Demircin reference (U.S. Pub. 2006/0095944) cited by the Examiner is different than the reference relied on in Grounds 1-2 of this Petition. The Examiner’s Demircin reference is materially different from the Demircin (SAMSUNG-1005) analyzed in this Petition (e.g., delayed transmission of NAL units in SVC).

Under these circumstances, Petitioner submits that exercise of the Board’s discretion under § 325(d) is unwarranted. To the extent file history contains any prior art teachings or arguments cumulative to those presented in this Petition, the Examiner erred in not recognizing the significance of those teachings or arguments in relation to the unpatentability of claim 1. *Advanced Bionics LLC v. MED-EL Elektromedizinische Gerate GmbH*, IPR2019-01469, Paper 6 (PTAB Feb. 13, 2020) (precedential) (citing *Becton, Dickinson and Co. v. B. Braun Melsungen AG*,

IPR2017-01586 Pap. 8 (PTAB Dec. 15 2017) (precedential)).

B. 35 U.S.C. §314(a) – *Fintiv*

In *Apple Inc. v. Fintiv, Inc.*, the Board enumerated six factors that it will assess in determining whether to exercise discretion to deny institution under §314(a) based on co-pending litigation. IPR2020-00019, Paper 11 at 2-3 (PTAB "precedential" Mar. 20, 2020) ("*Fintiv*"). In June 2022, the USPTO issued updated discretionary denial guidance in view of the *Fintiv* factors and expressly noted that ***the PTAB will not discretionarily deny institution if Petitioner presents a Sotera stipulation***. SAMSUNG-1030, 3, 7-8. Here, Petitioner has eliminated any risk of duplicated effort by voluntarily presenting a *Sotera* stipulation to Patent Owner. SAMSUNG-1029; *Sotera Wireless, Inc. v. Masimo Corp.*, IPR2020-01019, Paper 12 (PTAB Dec. 1, 2020). For at least this reason, and considering the compelling merits of Grounds 1-2, the Board should not discretionarily deny this Petition.

V. CONCLUSION AND FEES

The Challenged Claims are unpatentable. Petitioner 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)

Samsung Electronics Co., Ltd. is the petitioner and real party-in-interest. Samsung Electronics America, Inc. is an additional 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 civil actions: *Advanced Coding Technologies LLC v. Samsung Electronics Co., Ltd. et al.*, 2:22-cv-00499-JRG, E.D. Tex., filed December 30, 2022 (SAMSUNG-1004) and *Advanced Coding Technologies LLC v. LG Electronics, Inc. et al.*, 2:22-cv-00501-JRG, E.D. Tex., filed December 30, 2022.

Petitioner 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)

Petitioner provides the following designation of counsel.

Lead Counsel	Backup counsel
W. Karl Renner, Reg. No. 41,265 Fish & Richardson P.C. 60 South Sixth Street, Suite 3200 Minneapolis, MN 55402 Tel: 202-783-5070 Fax: 877-769-7945 Email:IPR39843-0171IP1@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 PTABInbound@fr.com

D. Service Information

Please address all correspondence and service to the address listed above. Petitioner consents to electronic service by email at IPR39843-0171IP1@fr.com (referencing No. 39843-0171IP1 and cc'ing PTABInbound@fr.com, axf-ptab@fr.com, jjm@fr.com, nstephens@fr.com, and rgupta@fr.com).

Respectfully submitted,

Dated 01/05/2024

/Nicholas W. Stephens/

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. IPR2024-00327)

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 13,855 words, which is less than the 14,000 allowed under 37 CFR §42.24.

Dated 01/05/2024

/Nicholas W. Stephens/

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 January 5, 2024, 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

/Michael Stanwyck/
Michael Stanwyck
Fish & Richardson P.C.
60 South Sixth Street, Suite 3200
Minneapolis, MN 55402
(202) 626-7790