

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

AMAZON.COM, INC., and
AMAZON WEB SERVICES, INC.,

Petitioner,

v.

DIVX, LLC,

Patent Owner.

Case: IPR2025-01223

U.S. Patent No. 11,611,785

DECLARATION OF HARI KALVA

TABLE OF CONTENTS

	Page
I. INTRODUCTION	14
II. BACKGROUND AND QUALIFICATIONS	14
III. COMPENSATION AND RELATIONSHIP WITH PARTIES.....	19
IV. MATERIALS CONSIDERED	19
V. BASIS OF OPINION	20
A. Legal Standard for Anticipation.....	20
B. Legal Standard for Obviousness	21
C. Claim Construction.....	23
VI. RELATED LITIGATION AND RELATED PATENTS	23
VII. OVERVIEW OF THE 785 PATENT.....	24
A. Priority Date	24
B. Technical Background.....	24
1. Introduction to Adaptive Bitrate Streaming.....	24
2. How Adaptive Streaming Works	27
3. Selecting the Bitrate and Resolution for Adaptive Bitrate Streaming	34
C. Patent Description	40
D. Level of Ordinary Skill in the Art.....	45
VIII. PRIOR ART TO THE CHALLENGED PATENT.....	47
A. Summary of the Prior Art.....	47
1. Ozer	47

TABLE OF CONTENTS (Continued)

	Page
2. Liao	53
3. Gu	56
4. Ronca	57
B. General Motivations to Combine Prior Art References	58
1. Ozer in view of Liao	58
2. Ozer in view of Liao and Gu	65
3. Ronca in view of Liao	67
4. Ronca in view of Liao and Gu	69
IX. INVALIDITY	71
A. Ground 1: Claims 1-4, 7-14, and 17-19 of 785 Patent are Rendered Obvious by Ozer in view of Liao	71
1. Claim 1	71
(1) [1.pre] “A method for encoding source content as a plurality of alternative video streams and adaptively streaming the plurality of alternative video streams, each having a bitrate, the method comprising:”	71
(2) [1.a] “identifying source video content;”	72
(3) [1.b] “identifying a plurality of resolutions for encoding of video content;”	73
(4) [1.c(i)] “for each particular resolution of the plurality of resolutions: encoding at least a portion of the source video content multiple times using the particular resolution and multiple different target bitrates for the particular resolution; and”	75

TABLE OF CONTENTS
(Continued)

	Page
(5) [1.c(ii)] “for each particular resolution of the plurality of resolutions: . . . evaluating quality for each of the multiple encodings”	78
(6) [1.d] “selecting a plurality of resolution and target bitrate combinations for a plurality of alternative streams based on the evaluated quality of each of the multiple encodings;”	81
(7) [1.e] “uploading encodings of a plurality of sections of the source video content at each of the selected plurality of resolution and target bitrate combinations for the plurality of alternative streams to a set of servers that form part of a content distribution system;”	83
(8) [1.f] “generating entries in a top level index describing at least one of the plurality of alternative streams;”	85
(9) [1.g] “providing the generated top level index from the content distribution system to a one or more playback devices capable of performing adaptive streaming using the encodings of the plurality of sections; and”	87
(10) [1.h] “responding to requests from the one or more playback devices, where the requests include requests for specific encodings from the encodings of the plurality of sections.”	88
2. Claim 2: “The method of claim 1 further comprising assembling the encodings of the plurality of sections for a given alternative stream into a video container file.”	90

TABLE OF CONTENTS
(Continued)

	Page
3. Claim 3: “The method of claim 2 further comprising generating entries in a container index within each container file indicating sizes of portions of the given alternative stream within the video container file.”	91
4. Claim 4: “The method of claim 1, wherein adaptively streaming the encodings comprises receiving at least one request for an encoded section of one of the plurality of alternative streams from the one or more playback devices based on the generated top level index.”	94
5. Claim 7: “The method of claim 1, wherein both a width and a height of each resolution of the plurality of resolutions have an integer number of pixels.”	96
6. Claim 8: “The method of claim 1, wherein each resolution of the plurality of resolutions has a same display aspect ratio and at least two resolutions of the plurality of resolutions have different sample aspect ratios.”	99
7. Claim 9: “The method of claim 1, wherein the target bitrate for each resolution is determined using a formula.”	103
8. Claim 10: “The method of claim 1, wherein the target bitrate of a given resolution is determined using experimentation.....	104
9. Claim 11	105

TABLE OF CONTENTS
(Continued)

	Page
(1) [11.pre] “A non-transitory machine readable medium containing processor instructions for encoding source content as a plurality of alternative video streams and adaptively streaming the plurality of alternative video streams, each having a bitrate, where execution of the instructions by a processor causes the processor to perform a process that comprises:”	105
(2) [11.a] “identifying source video content;”	106
(3) [11.b] “identifying a plurality of resolutions for encoding of video content;”	106
(4) [11.c(i)] “for each particular resolution of the plurality of resolutions: encoding at least a portion of the source video content multiple times using the particular resolution and multiple different target bitrates for the particular resolution; and”	106
(5) [11.c(ii)] “for each particular resolution of the plurality of resolutions: . . . evaluating quality for each of the multiple encodings;”	107
(6) [11.d] “selecting a plurality of resolution and target bitrate combinations for a plurality of alternative streams based on the evaluated quality of each of the multiple encodings;”	107

TABLE OF CONTENTS
(Continued)

	Page
(7) [11.e] “uploading encodings of a plurality of sections of the source video content at each of the selected plurality of resolution and target bitrate combinations for the plurality of alternative streams to a set of servers that form part of a content distribution system;”	108
(8) [11.f] “generating entries in a top level index describing at least one of the plurality of alternative streams;”	108
(9) [11.g] “providing the generated top level index from the content distribution system to one or more playback devices capable of performing adaptive streaming using the encodings of the plurality of sections; and”	109
(10) [11.h] “responding to requests from the one or more playback devices, where the requests include requests for specific encodings from the encodings of the plurality of sections.”	109
10. Claim 12: “The non-transitory machine readable medium of claim 11, wherein the process further comprises assembling the encodings of the plurality of sections for a given alternative stream into a video container file.”	110
11. Claim 13: “The non-transitory machine readable medium of claim 12, wherein the process further comprises generating entries in a container index within each container file indicating sizes of portions of the given alternative stream within the video container file.”	110

TABLE OF CONTENTS
(Continued)

	Page
12. Claim 14: “The non-transitory machine readable medium of claim 11, wherein adaptively streaming the encodings comprises receiving at least one request for an encoded section of one of the plurality of alternative streams from the one or more playback devices based on the generated top level index.”	111
13. Claim 17: “The non-transitory machine readable medium of claim 11, wherein both a width and a height of each resolution of the plurality of resolutions have an integer number of pixels.”	111
14. Claim 18: “The non-transitory machine readable medium of claim 11, wherein each resolution of the plurality of resolutions has a same display aspect ratio and at least two resolutions of the plurality of resolutions have different sample aspect ratios.”	112
15. Claim 19: “The non-transitory machine readable medium of claim 11, wherein the target bitrate for each resolution is determined using a formula.”	112
B. Ground 2: Claims 6 and 16 of 785 Patent are Rendered Obvious by Ronca in view of Liao and Gu.....	113
1. Claim 6: “The method of claim 1, wherein identifying a plurality of resolutions comprises identifying resolutions with target bitrates within a predetermined percentage of a bitrate associated with an alternative stream of the plurality of alternative streams.”	113
2. Claim 16: “The non-transitory machine readable medium of claim 11, wherein identifying a plurality of resolutions comprises identifying resolutions with target bitrates within a predetermined percentage of a bitrate associated with an alternative stream of the plurality of alternative streams.”	114
C. Ground 3: Claims 1-4, 7, 9-14, 17, and 19 of 785 Patent are Rendered Obvious by Ronca in view of Liao and Gu	114

TABLE OF CONTENTS
(Continued)

	Page
1. Claim 1	114
(1) [1.pre] “A method for encoding source content as a plurality of alternative video streams and adaptively streaming the plurality of alternative video streams, each having a bitrate, the method comprising:”	114
(2) [1.a] “identifying source video content;”	115
(3) [1.b] “identifying a plurality of resolutions for encoding of video content;”	116
(4) [1.c(i)] “for each particular resolution of the plurality of resolutions: encoding at least a portion of the source video content multiple times using the particular resolution and multiple different target bitrates for the particular resolution; and”	117
(5) [1.c(ii)] “for each particular resolution of the plurality of resolutions . . . evaluating quality for each of the multiple encodings;”	119
(6) [1.d)] “selecting a plurality of resolution and target bitrate combinations for a plurality of alternative streams based on the evaluated quality of each of the multiple encodings;”	121
(7) [1.e] “uploading encodings of a plurality of sections of the source video content at each of the selected plurality of resolution and target bitrate combinations for the plurality of alternative streams to a set of servers that form part of a content distribution system;”	123

TABLE OF CONTENTS
(Continued)

	Page
(8) [1.f] “generating entries in a top level index describing at least one of the plurality of alternative streams;”	124
(9) [1.g] “providing the generated top level index from the content distribution system to a one or more playback devices capable of performing adaptive streaming using the encodings of the plurality of sections; and”	125
(10) [1.h] “responding to requests from the one or more playback devices, where the requests include requests for specific encodings from the encodings of the plurality of sections.”	126
2. Claim 2: “The method of claim 1 further comprising assembling the encodings of the plurality of sections for a given alternative stream into a video container file.”	127
3. Claim 3: “The method of claim 2 further comprising generating entries in a container index within each container file indicating sizes of portions of the given alternative stream within the video container file.” /.....	128
4. Claim 4: “The method of claim 1, wherein adaptively streaming the encodings comprises receiving at least one request for an encoded section of one of the plurality of alternative streams from the one or more playback devices based on the generated top level index.”	130
5. Claim 7: “The method of claim 1, wherein both a width and a height of each resolution of the plurality of resolutions have an integer number of pixels.”.....	132
6. Claim 9: “The method of claim 1, wherein the target bitrate for each resolution is determined using a formula.”	133

TABLE OF CONTENTS
(Continued)

	Page
7. Claim 10: “The method of claim 1, wherein the target bitrate of a given resolution is determined using experimentation.....	135
8. Claim 11	135
(1) [11.pre] “A non-transitory machine readable medium containing processor instructions for encoding source content as a plurality of alternative video streams and adaptively streaming the plurality of alternative video streams, each having a bitrate, where execution of the instructions by a processor causes the processor to perform a process that comprises:”	135
(2) [11.a] “identifying source video content;”	136
(3) [11.b] “identifying a plurality of resolutions for encoding of video content;”	136
(4) [11.c(i)] “for each particular resolution of the plurality of resolutions: encoding at least a portion of the source video content multiple times using the particular resolution and multiple different target bitrates for the particular resolution; and”	137
(5) [11.c(ii)] “for each particular resolution of the plurality of resolutions: . . . evaluating quality for each of the multiple encodings;”	137
(6) [11.d] “selecting a plurality of resolution and target bitrate combinations for a plurality of alternative streams based on the evaluated quality of each of the multiple encodings;”	137

TABLE OF CONTENTS
(Continued)

	Page
(7) [11.e] “uploading encodings of a plurality of sections of the source video content at each of the selected plurality of resolution and target bitrate combinations for the plurality of alternative streams to a set of servers that form part of a content distribution system;”	138
(8) [11.f] “generating entries in a top level index describing at least one of the plurality of alternative streams;”	138
(9) [11.g] “providing the generated top level index from the content distribution system to one or more playback devices capable of performing adaptive streaming using the encodings of the plurality of sections; and”	139
(10) [11.h] “responding to requests from the one or more playback devices, where the requests include requests for specific encodings from the encodings of the plurality of sections.”	139
9. Claim 12: “The non-transitory machine readable medium of claim 11, wherein the process further comprises assembling the encodings of the plurality of sections for a given alternative stream into a video container file.”	140
10. Claim 13: “The non-transitory machine readable medium of claim 12, wherein the process further comprises generating entries in a container index within each container file indicating sizes of portions of the given alternative stream within the video container file.”	140

TABLE OF CONTENTS
(Continued)

	Page
11. Claim 14: “The non-transitory machine readable medium of claim 11, wherein adaptively streaming the encodings comprises receiving at least one request for an encoded section of one of the plurality of alternative streams from the one or more playback devices based on the generated top level index.”	141
12. Claim 17: “The non-transitory machine readable medium of claim 11, wherein both a width and a height of each resolution of the plurality of resolutions have an integer number of pixels.”	141
13. Claim 19: “The non-transitory machine readable medium of claim 11, wherein the target bitrate for each resolution is determined using a formula.”	142
D. Ground 4: Claims 6 and 16 of 785 Patent are Rendered Obvious by Ronca in view of Liao and Gu.....	142
1. Claim 6: “The method of claim 1, wherein identifying a plurality of resolutions comprises identifying resolutions with target bitrates within a predetermined percentage of a bitrate associated with an alternative stream of the plurality of alternative streams.”	142
2. Claim 16: “The non-transitory machine readable medium of claim 11, wherein identifying a plurality of resolutions comprises identifying resolutions with target bitrates within a predetermined percentage of a bitrate associated with an alternative stream of the plurality of alternative streams.”	144
X. CONCLUSION.....	144

I, Dr. Hari Kalva, declare as follows:

I. INTRODUCTION

1. I have been asked by, Amazon.com, Inc. and Amazon Web Services, Inc. (“Petitioner”), to provide my expert opinions in support of the petition for Inter Partes review of U.S. Patent No. 11,611,785 (the “785 patent” or “challenged patent”) challenging the patentability of claims 1-4, 6-14, 16-19 (the “challenged claims”).

2. It is my opinion that the challenged claims are disclosed and rendered obvious by the prior art. I set forth below my detailed opinions.

II. BACKGROUND AND QUALIFICATIONS

3. I am Professor and Department Chair in the Department of Electrical Engineering and Computer Science at Florida Atlantic University. I regularly teach courses on Engineering Design, Video Communication, and Cloud Computing. I am also Director of Florida Atlantic University’s Multimedia Lab, in which capacity I conduct research in the broad area of visual computing that includes video compression, communication, and analysis. Part of the research relates to development and optimization of video compression standards.

4. I received my Doctorate in Electrical Engineering from Columbia University in 2000. Prior to that, I received a Master of Philosophy in Electrical Engineering from Columbia University in 1999, a Master of Science in Computer

Engineering from Florida Atlantic University in 1994, and a Bachelor of Technology in Electronics and Communications Engineering from the N.B.K.R. Institute of Science and Technology at Sri Venkateswara University in Tirupati, Andhra Pradesh, India.

5. I have taught in the Department of Electrical Engineering and Computer Science at Florida Atlantic University since 2003. Before that, I was a Consultant in the Video Processing Group for Mitsubishi Electric Research Labs from 2001 to 2003; Co-Founder and Vice President of Engineering at Flavor Software, a New York based startup developing MPEG-4 standard based multimedia solutions, from 2000 to 2001; a Research Staff Associate in the Image and Advanced Television Laboratory at Columbia University from 1995 to 1996; and a Teaching Assistant in the Department of Electrical Engineering and Computer Science at Florida Atlantic University from 1993 to 1994.

6. I have extensive experience in multimedia data compression technologies. In addition to my education and nearly three decades of professional experience in data compression technologies, at both academic and commercial levels, I have been granted numerous patents for inventions relating to such technologies; I have authored over 100 peer reviewed publications regarding such technologies; and I have contributed to several international video compression standards, including contributions of standard essential patents.

7. I am a 2025 Inductee to the Florida Inventors Hall of Fame and recognized for “Significant contributions to international standards for video compression, communications, and analytics that have enabled video streaming technologies, fundamentally transforming education, entertainment, and the workplace.” I am a Fellow of the National Academy of Inventors (NAI), an honor bestowed upon individuals whom the NAI’s Selection Committee deems to “have demonstrated a highly prolific spirit of innovation in creating or facilitating outstanding inventions that have made a tangible impact on the quality of life, economic development and the welfare of society.” I am also a senior member of the Institute of Electrical and Electronics Engineers (IEEE); a member of the IEEE’s Consumer Electronics, Circuits and Systems, Communications, and Signal Processing Societies, including the Multimedia Technical Committee of the IEEE’s Communications Society; Chair of the IEEE’s Palm Beach Section; and a member of the Association for Computer Machinery.

8. I have been granted 70 U.S. patents, covering inventions including systems and methods for encoding multimedia content; systems and methods for processing object-based audiovisual architecture; systems and methods for video and multimedia rendering, composition, and user-interactivity; methods for encrypting and compressing video, object-based audiovisual terminal and bitstream structures. I have also been granted over a dozen foreign patents covering similar

types of inventions. In addition, I have sold patent applications to Mitsubishi Electric Company covering multi-level bi-directional prediction and methods for encoding and decoding video.

9. In particular, I am an inventor of patented technology essential to the implementation of three video compression standards: H.264, also known as MPEG-AVC (for “Advanced Video Coding”) or MPEG-4 Part 10; H.265, also known as HEVC (for “High Efficiency Video Coding”); and H.266, also known as VVC (for “Versatile Video Coding”) or MPEG-I Part 3.

10. I was a key member of the team that initiated the MPEG-4 file format standards activity and wrote the software for the proposed file format. Based on the ideas and experiences with implementing the initial MPEG-4 file format, MPEG-4 video, and MPEG-4 Systems technologies, I developed requirements for the MPEG-4 file format that is used today. I have experience writing software for video compression and streaming systems including MPEG-1, MPEG-2, H.263, MPEG-4 part 2, and MPEG-4 part 10.

11. I have made over three dozen contributions to the development of international video compression standards from the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC). Among other video compression standards, I made a variety of contributions to the MPEG-4 group of international standards for compressing digital audio and visual

data, multimedia systems, and file storage formats. A list of my contributions to international standards is provided in my curriculum vitae (EX1004).

12. I served as U.S. delegate to the ISO standards committees during the development of the MPEG-4, MPEG-7, VVC, and VCM international video compression standards. Over the past three decades, I have also served on the technical program committees of a variety of conferences of the International Conference on Consumer Electronics, the International Conference on Multimedia and Expo, the International Workshop on Multimedia Signal Processing, the IEEE's International Symposium on Multimedia, and other organizations.

13. I have published numerous peer reviewed articles, chapters, and books pertaining to video compression technology over the past three decades. Many of these publications have focused on various international video compression standards, including MPEG-2, H.264, and H.265, also known as HEVC or MPEG-H Part 2. I have also authored many publications on video compression technologies such as intra-frame coding, a technique that allows smaller files sizes and lower bit rates with little to no loss in quality. A list of my peer reviewed publications, including all publications authored in the previous ten years, is provided in my curriculum vitae (EX1004).

14. I have previously been retained as an expert witness in cases involving patents before the Patent Trial and Appeal Board of the U.S. Patent and Trademark

Office to assess the relevance of prior art to patents related to HEVC video decoding. A list of cases in which I have been retained as an expert witness, including all cases at which I testified as an expert at trial or by deposition during the previous four years, is provided in my curriculum vitae (EX1004).

III. COMPENSATION AND RELATIONSHIP WITH PARTIES

15. I am being compensated for my time at a rate of \$650 per hour. This compensation is not contingent upon my performance, the conclusions I reach in my analysis, the outcome of this matter, or any issues involved in or related to this matter.

16. I have no financial interest in Petitioner or any related parties.

IV. MATERIALS CONSIDERED

17. I have reviewed and considered, in preparing this declaration, the 785 patent (EX1001), its prosecution history (EX1002), and the patents and publications noted in this declaration.

18. I am also aware of, and considered in forming my opinions presented in this declaration, other information available to a person of ordinary skill in the art (“POSITA”) at the time of the effective filing date of the challenged patent, such as

other patents, reference materials, articles, and standards cited in this declaration, all of which reflect the general knowledge of a POSITA.¹

19. I understand that Petitioner and Patent Owner are involved in a parallel litigation involving seven patents, including the challenged patent.

20. Should new information be disclosed or discovered during this proceeding, I may supplement and further elaborate on the opinions I express in this declaration as necessary. Nothing in this declaration should be interpreted as an opinion that no grounds of invalidity exist other than the ones in the petition. I do not express such an opinion and have not reached such an opinion.

V. BASIS OF OPINION

A. Legal Standard for Anticipation

21. I understand that a claimed invention is unpatentable as “anticipated” by a single prior art reference if it discloses each and every element of the claim. I also understand that, even if a claim is not explicitly disclosed in the prior art reference, the prior art reference may nevertheless anticipate the claim if the subject matter of the claim is inherently present from the perspective of a POSITA. Moreover, I understand that even if the prior art reference does not disclose the

¹ I understand that the challenged patent claims to have an effective filing date of August 30, 2011, as I discuss in more detail below.

limitations of a claim in the same arrangement or combination of the limitations of the claim, it may nevertheless anticipate if a POSITA would envisage the claimed arrangement or combination.

B. Legal Standard for Obviousness

22. I understand that a claimed invention is unpatentable as “obvious” if the differences between the invention and the prior art are such that the subject matter of the claim as a whole would have been obvious to a POSITA.

23. It is my understanding that “obviousness” is a question of law based on underlying factual issues including (1) the scope and content of the prior art, (2) the differences between the prior art and the asserted claims, (3) the level of ordinary skill in the pertinent art, and (4) the existence of secondary considerations such as commercial success, long-felt but unresolved needs, and failure of others.

24. I understand that a single reference can render a patent claim obvious if any differences between that reference and the claims would have been obvious to a POSITA. Alternatively, the teachings of two or more references may be combined in the same way as disclosed in the claims, if such a combination would have been obvious to a POSITA. In determining whether a combination based on either a single reference or multiple references would have been obvious, I understand that it is appropriate to consider the following factors:

- Whether the teachings of the prior art references disclose known concepts combined in familiar ways, and when combined, would yield predictable results;
- Whether a POSITA could implement a predictable variation, and would see the benefit of doing so;
- Whether a POSITA would have recognized a reason to combine known elements in the manner described in the claim;
- Whether there is some teaching or suggestion in the prior art to make the modification or combination of elements claimed in the patent; and
- Whether the innovation applies a known technique that had been used to improve a similar device or method in a similar way.

25. In an obviousness evaluation based on a combination of multiple prior art references, I understand that the prior art references themselves may provide a suggestion, motivation, or reason to combine, but other times the nexus linking two or more prior art references is simple common sense. I further understand that obviousness analysis recognizes that market demand, rather than scientific literature, often drives innovation, and that a motivation to combine references may be supplied by the direction of the marketplace.

C. Claim Construction

26. I understand that, in *Inter Partes* Review (“IPR”) proceedings, a claim of a patent shall be construed using the same claim construction standard that would be used to construe the claim in a civil action filed in a U.S. District Court (which I understand is called the “*Phillips*” claim construction standard), including construing the claim in accordance with the plain-and-ordinary meaning of such claim as understood by one of ordinary skill in the art and the prosecution history pertaining to the patent.

27. I applied plain-and-ordinary meaning to all terms of the challenged claims for my analysis in this proceeding.

28. I am not an attorney, and my opinions are limited to what I believe a POSITA would have understood the meaning of certain claim terms to be, based on the patent specification and prosecution history in light of the POSITA’s knowledge.

VI. RELATED LITIGATION AND RELATED PATENTS

29. I understand that the 785 patent is the subject of pending litigation involving Petitioner and Patent Owner: *DivX, LLC v. Amazon.com, Inc., and Amazon.com Web Services, Inc.*, No. 1:24-cv-02061 (Eastern District of Virginia, November 15, 2024) (the “parallel litigation”). I understand that Patent Owner sued Petitioner for alleged infringement of the 785 patent.

VII. OVERVIEW OF THE 785 PATENT

A. Priority Date

30. The 785 patent is a continuation of application No. 16/789,303, filed on February 12, 2020, now U.S. Pat. No. 10,931,982, which is a continuation of application No. 15/922,198, filed on March 15, 2018, now U.S. Pat. No. 10,645,429, which is a continuation of application No. 13/600,046, filed on August 30, 2012, now U.S. Pat. No. 9,955,195. The 785 patent claims priority to August 30, 2011 through provisional application No. 61/529,201. EX1001, 1.

B. Technical Background

1. Introduction to Adaptive Bitrate Streaming

31. Video streaming allows a client device to play video content over the internet, without the need to download the entire file beforehand. EX1009, 1; EX1010, 2:32-37; EX1011, 5; EX1012, 5-7; EX1005, 18. Prior to streaming, a client device would download an entire video file to a device before playback began. EX1010, 2:20-31; *see also* EX1012, 5-7. But this process could take a long time, was highly dependent on network bandwidth, and was inefficient. EX1010, 2:20-31; *see also* EX1012, 5-7. For example, a client device interested in only a portion of the video file still had to download the entire video file. EX1010, 2:20-31; *see also* EX1012, 5-7. Now, with streaming, sequential portions of a video file can be delivered in a continuous stream of data, allowing for immediate playback at a client

device (e.g., mobile phone, tablet, computer, television, etc.) once the first portion is received. EX1010, 2:32-37; *see also* EX1012, 5-7; EX1005, 18.

32. The video data is delivered to a client device at a specific transmission rate. EX1005, 18; *see also* EX1013, 873. Specifically, to avoid any disruptions to the stream, video streaming requires transmitting video streams at rates comfortably below the available capacity of the network (e.g., the network's bandwidth) and the available resources of the client device (e.g., the device's buffer capacity), but not so low that video is being transmitted more slowly than it is being played. EX1005, 18; EX1013, 873. This is done in part to ensure that client devices does not experience playback disruptions due to a mismatch between their available network bandwidth and/or available resources and the transmission rates of the received video streams. EX1005, 18; EX1013, 873. Because both network bandwidth and the available resources of the client device tend to fluctuate over time, the transmission rate of the stream needs to be adjusted for this variability. EX1009, 1-2; EX1012, 5-6, 7-9; EX1005, 19. In general, the transmission rate limits the amount of video data that can be transmitted at a given period of time. Ex. EX1012, 5-6; EX1005, 19.

33. The amount of data in an encoded video is described in terms of the average data per second of video (bitrate). EX1012, 5-6; EX1005, 16, 18. This is known as the encoded video's bitrate. EX1012, 5-6; EX1005, 16, 18. Transmitting

a video encoded at a higher bitrate requires a higher transmission rate, while transmitting a video encoded at a lower bitrate requires a lower transmission rate. *See* EX1012, 5-6 (“The server sends the data packets to the client at a real-time rate only—that is, the bit rate at which the media is encoded. For example, a video encoded at 500 kilobits per second (kbps) is streamed to clients at approximately 500 kbps.”); *see also* EX1005, 19. The need to adapt to the varying network and client resources led to the introduction of adaptive bitrate streaming. EX1009, 1-2; *see also* EX1012, 7-9; EX1014, 3-4; EX1005, 19, 115.

34. Adaptive streaming allows a client device to dynamically adjust between streams encoded at a higher and a lower bit rate, thereby resulting in higher or lower transmission rates. EX1009, 1 *see also* EX1012, 7-9. Specifically, if network capacity changes, a client device can request a lower bit rate or higher bitrate version of the video content or stream on the fly, ensuring that the client device can continuously stream the video content. EX1009, 1-2; *see also* EX1015, ¶¶155-57; EX1016, ¶45-51; EX1012, 7-9; EX1005, 19, 115.

35. The client device dynamically chooses the version of the video content (the alternate stream) based on information such as buffer conditions, client bandwidth, CPU use, and playback window size. EX1009, 1-2, Fig. 2; *see also* EX1016, ¶¶48-51; EX1015, ¶¶156-57; EX1005, 116-17. Ozer provides more detail regarding how the client device knows when to change streams.

How does the system know when to change streams? This depends upon multiple factors, which can include:

Buffer conditions. Each player has a playback buffer-say 5 seconds of video data that ensures that if the connection stops for intermittent periods, playback will continue. An adaptive streaming technology can monitor the buffer, switching to higher- bandwidth streams when the buffer stays full for a certain duration (indicating a robust connection) or drop to a lower-bandwidth streams when the buffer drops under certain levels (say 2 seconds).

Client bandwidth. For technologies that monitor effective client bandwidth, the system can adjust the stream to match the effective bandwidth.

CPU use. Systems can monitor dropped frames, which are an indication of the CPU power available for video decode. For example, even if the buffer was full and bandwidth sufficient for a higher-quality stream, the system wouldn't switch if the player was currently dropping frames, since a higher-bandwidth, higher-quality stream requires more CPU to decode, and would cause more dropped frames.

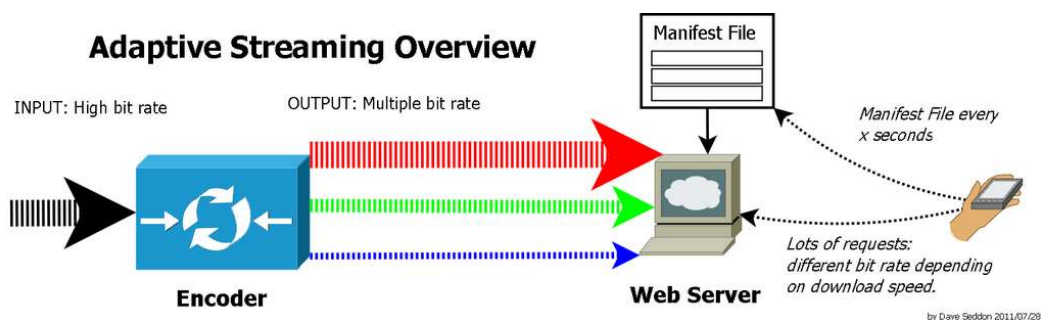
Playback window size. Some systems monitor playback size to determine whether to switch to a higher-quality stream. For example, if the video is being played back in a 640x360 window-even if buffer, bandwidth and CPU status indicate the ability to retrieve and playback a higher-bandwidth stream-the system wouldn't switch to a 720p stream because the viewer wouldn't notice the quality difference within the 640x360 viewing window, essentially wasting the additional bandwidth. If the viewer switched to full screen, however, the system would then switch to the higher-quality stream.

EX1005, 116-17.

2. How Adaptive Streaming Works

36. To allow for switching between different streams, a single source video needs to be encoded into multiple encodings with different bitrate and resolutions. EX1016, ¶¶45-46, Fig. 1; *see also* EX1015, ¶¶156-57; EX1014, 3-4; EX1005, 115-18. These multiple encodings are used to facilitate the streaming across a broad range of network bandwidths, screen sizes, and aspect ratios. EX1016, ¶¶45-46, Fig. 1; *see also* EX1015, ¶¶156-57; EX1014, 3-4; EX1005, 115-18. To provide for this functionality, as shown below, adaptive bitrate streaming systems include at least a

few basic components: an encoder, one or more servers (e.g., a content distribution network), and a client. EX1017, 1 (available at https://en.wikipedia.org/wiki/Adaptive_bitrate_streaming); *see also* EX1016, Fig. 1.; *see also* EX1005, 115-118, Fig. 7-1. In a more complex adaptive bitrate streaming system architecture, the server can take the form of a content distribution network or system (CDN). The CDN may include many servers that store video encodings and their associated index and manifest files. The CDN can then transmit the content and files to clients upon request to facilitate adaptive streaming. EX1005, 115-18, 137-37, Fig. 7-1, Fig. 7-9 (depicting a content delivery network consisting of a network of connected servers); EX1011, 9, 12, Fig. 1-1; EX1008, ¶¶28-31.

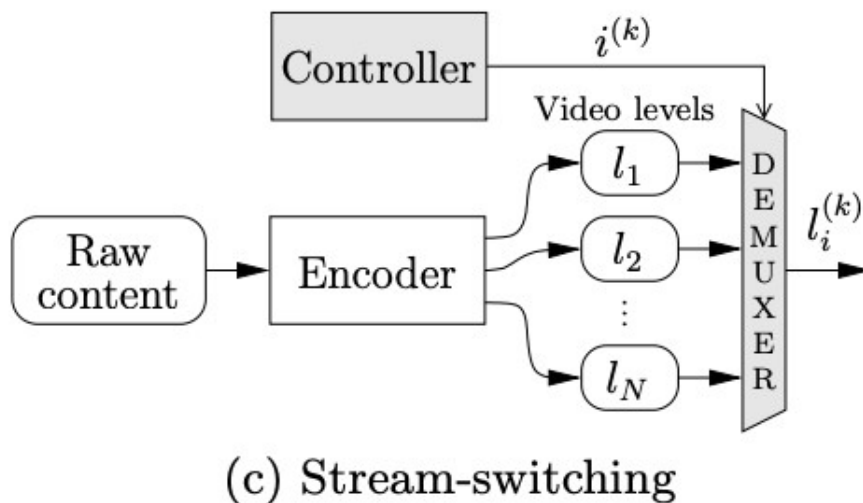


EX1017, 1; *see also* EX1016, Fig. 1.

37. As shown above, an encoder receives a source video with a high bitrate (INPUT). EX1017, 1; *see also* EX1016, Fig. 1. The encoder then encodes the source video into multiple versions (OUTPUT) each with a different bitrate. EX1017, 1; *see also* EX1016, Fig. 1. In the above figure, the red, green, and blue arrows

represent multiple encodings at different bit rates, where the red arrow has the highest bitrate and the blue arrow has the lowest bitrate. EX1017, 1; *see also* EX1016, Fig. 1.

38. Typically, as shown below in Figure 1(c), the server encodes the INPUT video content at increasing bitrates and resolutions resulting in some predetermined number of versions (“N”), or video levels (e.g., $l_1, l_2, \dots, l_N$) of the content. EX1009, 1, Fig. 1(c), Table 1; *see also* EX1015, ¶¶155-57; EX1016, ¶¶45-46; EX1012, 7-9; EX1018, 20; EX1005, 115-18.



EX1009, Fig. 1(c).

39. The table below shows an example of the resulting encodings, including five versions of the source video content, or five video levels, at five different bit rates (300, 700, 1500, 2500, 3500 (kbps)) clustered around three different resolutions (320x180, 640x360, 1280x720 (width x height)).

Video level	Bitrate (kbps)	Resolution (width×height)
l_0	300	320x180
l_1	700	640x360
l_2	1500	640x360
l_3	2500	1280x720
l_4	3500	1280x720

Table 1: Set of available video levels $\mathcal{L}$

EX1009, Table 1.

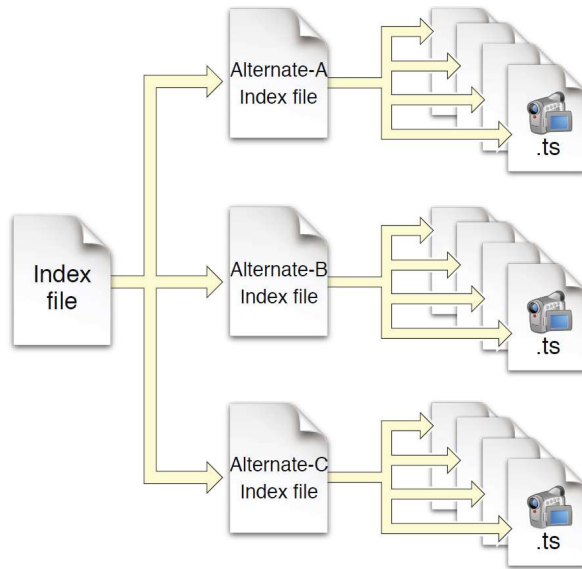
40. The multiple versions of the video encodings (of the original video INPUT video content) are then transmitted to a web server or a network of servers, known as a CDN, for storage. EX1017, 1; *see also* EX1016, ¶47, Fig. 1; EX1015, ¶157; EX1005, 115-18. Each version of the video encoding is stored in short 2-to-10 second segments (or chunks). EX1012, 7-9; *see also* EX1015, ¶¶158, 168; EX1005, 115-18. So, for example, given three versions of the video encoding, there would exist three encoded video chunks of 2-to-10 seconds for each portion of the original video content (each encoded at different resolutions and bitrates as shown in the above table). EX1012, 7-9; *see also* EX1015, ¶¶158, 168; EX1016, ¶ 46.

41. In addition to storing the video encodings, the web server or CDN also stores a file including information describing each of the different video encodings, often referred to as a “Manifest File” (as shown in the Figure above) or “index file.” EX1017, 1; EX1016, ¶47; *see also* EX1015, ¶158; EX1011, 16-17; EX1005, 115-

18. It is also often called an index file because it contains an indexing of all the multiple encoded versions of a given source video file that are stored on the web server or the CDN. EX1011, 16-17; EX1005, 117-18.

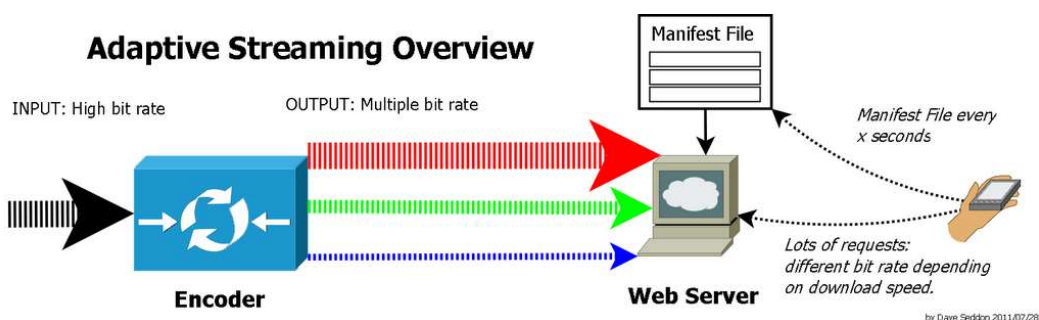
42. An index file may take on different forms and include different kinds of information. Generally, the index file includes information identifying each of the video encodings (e.g., a title, a synopsis, a content identifier (ID)), information describing the encoding characteristics of the video encodings (resolution, bit rate, size), and they often provide an address to locate each of the video encodings (e.g., a content uniform resource locator (URL)) to facilitate requests for them. EX1016, ¶¶47, 84-120, 157, 167, Fig. 13A; EX1015, ¶¶158, 168, 175; EX1011, 6-7. One example of an index file, used by HTTP-based systems like Apple's HTTP Live Streaming, is shown below. EX1011, Fig. 2-2. Each index file identifies the alternative versions of the video source file (Alternate-A, Alternate-B, and Alternate-C) encoded at a given bitrate and resolution. EX1011, 11-12, 16-18, Fig. 2-1; EX1005, 117-18, 127-31. The index file also identifies the 2-10 second chunks of video (".ts" transport stream) for each alternate version of the video. EX1011, 16-18, 30, Fig. 2-1; EX1005, 117-18, 120.

Figure 2-1 Stream alternates



EX1011, 16, Fig. 2-1.

43. As shown in the below image, when a client device wants to initiate playback of a video it first obtains the “Manifest File” or index file from the web server or CDN. EX1017, 1; *see also* EX1016, ¶47; EX1015, ¶158; EX1005, 115-18.



EX1017, 1. As noted above, the “Manifest File” or index file includes information describing all the multiple versions of the video encodings of the video, such as the given bitrate and resolution of each video encoding and where to locate the chunks

making up each video encoding. EX1016, ¶¶47, 155, Fig. 13A; *see also* EX1015, ¶¶ 156, 158; EX1005, 115-18. The client device receives this index and uses it to request the version of the video, encoded at a given bitrate and resolution, appropriate for its network capacity and resources. EX1016, ¶¶47-50, 155; *see also* EX1015, ¶¶156, 158. For example, if the network capacity and/or available resources of the client device decreases, a client device will look at the index, identify a lower bitrate stream in the index, and send a request to the web server or CDN for that lower bitrate video. EX1016, ¶¶48-50, Fig. 1; *see also* EX1015, ¶¶155-56, 158; EX1005, 115-18; EX1012, 7-9; EX1011, 16-17.

44. When the client device requests a chunk of the desired video encoding, it requests a chunk at a time, allowing the device to re-evaluate which video encoding to request based on the streaming environment between requests for chunks. EX1015, ¶¶156, 158; *see also* EX1005, 115-18. A client may determine its streaming environment (based on for example the bandwidth of its network) using any suitable means of measuring or estimating a streaming environment. EX1016, ¶51; EX1005, 115-18. Ha provides more details below.

The client 130 may estimate a streaming environment by using any method of estimating a streaming environment based on, for example, at least one of the bandwidth of the network 140 and the hardware resources that may be used by at least one of the server 120 and the client 130. For example, the client 130 may estimate the streaming environment based on a time stamp and a bit error rate (BER) of received media data. The streaming environment may be determined to be deteriorated when media data is received slower than a reproduction speed by checking time stamps of the received media data. Alternatively, the streaming environment may be determined to be deteriorated when BERs of the received media

data are increased.

EX1016, ¶51.

3. Selecting the Bitrate and Resolution for Adaptive Bitrate Streaming

45. Streaming protocols are ubiquitous in the field of video streaming. *See generally* EX1019. Streaming protocols define the mechanism by which video data is streamed across the Internet to client devices. EX1019, 5 (“Streaming protocols define the standards and methods for transmitting multimedia content over the internet from one device to another.”) One of the first streaming protocols (Real Time Messaging) was introduced in 1996. EX1019, 5. The timeline also Apple’s HTTP Live Streaming (“HLS”) was introduced in 2009. EX1022; *see also* EX1005, 118-21.

46. Providers of streaming protocols, like Apple, the provider of the HLS protocol, often provide tables (*see* Table 2-1 below) that recommend encoder settings (e.g., combinations of resolution and bitrate) for a content provider’s multiple alternate streams. *See, e.g.*, EX1011, Table 2-1; *see also* EX1005, 126-27, 137-38, 139-41. Table 2-1 below is an example of the recommended encoder settings provided by Apple. EX1011, Table 2-1. Each of the rows below represents one alternate stream and includes a recommended bitrate and resolution. *Id.* It should be noted that “Total bit rate” reflects the audio bit rate plus the “Video bit rate.” *Id.* As can be seen below, several of the streams (red box below) share the

same resolution (480 x 224) but have a different video bitrate (110, 200, 400 kbps).

Id.

Table 2-1 Encoder settings for iPhone, iPod touch, iPad, and Apple TV, 16:9 aspect ratio

Connection	Dimensions	Total bit rate	Video bit rate	Keyframes
Cellular	480 x 320	64 kpbs	audio only	none
Cellular	480 x 224	150 kpbs	110 kbps	30
Cellular	480 x 224	240 kpbs	200 kbps	45
Cellular	480 x 224	440 kpbs	400 kbps	90
WiFi	640 x 360	640 kpbs	600 kbps	90

EX1011, Table 2-1.

47. Encoder settings factor into the video bitrate. A lower resolution video can be represented by less data and thus a lower bitrate. EX1015, ¶¶155-57; *see also* EX1006, 7:62-8:9. A higher resolution video can be represented by more data and thus a higher bitrate. EX1015, ¶¶155-57; EX1005, 17. Resolution and level of compression also impact the video bitrate. EX1005, 17. Higher levels of compression can results in lower bitrates and compression, while lower levels of compression can results in higher bitrates and compression. *Id.*; *see generally* EX1014; EX1011, 16-18; EX1005, 127-131.

48. Customizing encoder settings involves selecting the appropriate number of alternative streams and selecting the bitrate and resolution for each of the alternate streams to maximize perceived quality across a broad range of content, network bandwidths, screen sizes, and aspect ratios. EX1014, 3-4; *see also* EX1005, 127-131; EX1015, ¶155; EX1009, 2. Content providers select a number of alternative streams and their bitrates to ensure coverage for a broad range of streams, assuring smooth switching between them. EX1014, 3-4. Too many alternative streams packed closely together might result in too much stream switching, and too few alternative streams may result in too little stream switching. EX1005, 127-28, 130-31. Either state will have a negative impact on viewing experience. EX1005, 127-28, 130-31. For content at a certain bit rate, there is an optimal resolution that maximizes the perceived quality of each of the streams (maximizes the tradeoff between lower resolutions that provide a slightly blurred image with fewer artifacts versus higher resolutions that provide a sharper image but possibly with more artifacts), and, thus, the goal for each combination of bitrate and resolution for each stream is to optimize the perceived quality of each stream. EX1014, 3-4; *see also* EX1015, ¶155; EX1005, 127-31.

49. POSITAs use experimentation to determine the resolution and bitrates of each of the streams. Specifically, experimentation used to maximize the quality of the different combinations of resolution and bitrates. The experimentation takes

the form of repeatedly encoding a video source file using multiple combinations of bitrates and resolutions and then comparing those encodings using various quality metrics.² EX1014, 3-8; EX1021, 258; EX1005, 127-131. The quality of each of the resulting encodings can be evaluated using structural similarity index measure (“SSIM”). EX1014, 5-8; *see also* EX1006, 8:37-9:16. SSIM is a “quality and fidelity metric that measures the similarity between the original image and the same image after processing.” EX1014, 5-8; *see also* EX1006, 8:37-9:16.

50. SSIM was first introduced in a paper published in IEEE Transactions on Image Processing in April of 2004. EX1035; *see also* EX1034. SSIM subsequently found strong adoption in the image processing community and in the television and social media industries. EX1034. The paper from which it stemmed has subsequently been cited over 50,000 times, and it was recognized with the IEEE Signal Processing Society Best Paper Award for 2009. *Id.* Thus, its use as a quality heuristic was well known in the industry prior to 2011.

51. Using SSIM to compare the encodings, the content provider could determine the optimal resolution for each bit rate range by comparing the SSIM of

² The optimal bitrate for any content at a specific resolution can also be determined through experimentation. EX1014, 10-11.

the encodings or by ensuring that a given encoding achieves a minimum SSIM. EX1014, 7-8, Fig. 2; *see also* EX1006, 8:37-9:16, 10-11.

52. The graph below demonstrates bitrate selection based on SSIM. The graph shows the same video encoded with four different resolutions (480 x 720 (“720p 2/3”)³, 640 x 480, 480 x 320, and 320 x 240). EX1014, 6-7, Figs. 1, 2. Each of the four videos were then repeatedly reencoded at different bitrates, and each of these encodings were evaluated using SSIM. EX1014, 7-8, Figs. 1, 2. As shown below, the SSIM of the 320 x 240 resolution encodings (QVGA (light blue)) is highest amongst the alternative encodings at 600 kbps, while the SSIM index of the 480 x 320 resolution encodings (HVGA (green)) is the highest amongst the alternative encodings at 1000 kbps. EX1014, 8, Fig. 2. Thus, based on this quality evaluation, a content provider would select a first alternate stream having a resolution of 320 x 240 and a bitrate of 600 kbps, a second alternate stream having

³ An image with a resolution of 720p 2/3 has 720 vertical pixels and aspect ratio of 2/3. The aspect ratio is the ratio the width to the height of an image, so here, the ratio of width to height is 2 to 3. Thus, here, the resolution in terms of width to height is 480 x 720 because 720 vertical pixels serves as the height and 720 is 150% (3/2) of 480.

a resolution of 480 x 320 and a bitrate of 1125 kbps, a third stream having a resolution of 640 x 480 and a bitrate of 2000 kbps, and so on.

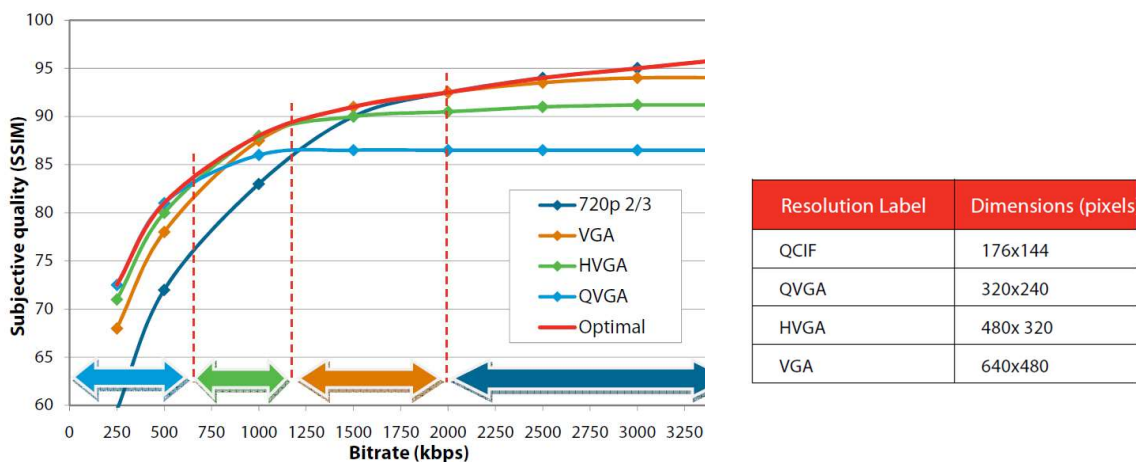


Figure 2. Optimal resolution for the sequence "Public Television" at various bit rates

EX1014, 7-8, Fig. 2 (annotated).

53. The selection of appropriate resolution and bitrate combinations for encoder settings based on quality is baked into widely used streaming protocols. For example, Dynamic Adaptive Streaming over HTTP (MPEG-DASH), with its original draft released in January of 2011, includes a “qualityRanking” attribute. For example, the index could include a “qualityRanking” attribute that directly “[p]rovides a quality ranking of the [alternative video stream] relative to other [alternative video streams].” EX1018, 20. Also, for example, the index could include a “qualitydependency” attribute that describes how to weigh quality factors like resolution, bandwidth, and framerate. EX1016, ¶¶ 170-178, Fig. 8A. These attributes allow a client device to make an informed decision regarding which

alternative stream to request at any given time based on “quality.” EX1016, ¶ 170.

Ha provides more detail below with respect to the “qualitydependency” attribute.

In adaptive streaming, the priority among the plurality of elements for determining the quality may function as an important selection standard when the client 130 selects media data appropriate for a streaming environment and requests the server 120 for the media data from among a plurality of media data. Thus, the media presentation descriptions of FIGS. 14A through 14D according to the present embodiment define the information about the priority among the plurality of elements for determining the quality, by using a “qualitydependency” attribute, so that the client 130 may select optimal media data appropriate for the streaming environment.

EX1016, ¶ 170.

C. Patent Description

54. U.S. Patent No. 11,611,785 B2, titled “Systems and Methods for Encoding and Streaming Video Encoded Using a Plurality of Maximum Bitrate Levels,” (the “785 patent,” EX1001), lists Kourosh Soroushian as an inventor, and was filed on February 22, 2021 and issued on March 21, 2023.

55. The 785 patent is directed to well-known, conventional techniques for encoding videos for adaptive bitrate streaming. EX1001, 1:22-27, 1:39-42. The Background of the 785 patent defines adaptive bitrate streaming or adaptive streaming as follows: “Adaptive bitrate streaming or adaptive streaming involves detecting the present streaming conditions (e.g. the user’s network bandwidth and CPU capacity) in real time and adjusting the quality of the streamed media accordingly.” EX1001, 1:39-42. As further explained in the Background section, adaptive bitrate streaming on a playback device requires encoding a source video

into several alternative streams. *Id.*, 1:43-47. Further, each of these alternate “stream[s] is typically stored in one or more container files.” *Id.* And during streaming, a “top level index” pointing to these alternative streams is used to “adjust[] the quality of the streamed media” based on “present streaming conditions (e.g., user’s network bandwidth and CPU capacity).” *Id.*, 1:34-47.

56. Additionally, during prosecution the applicant highlighted that the purported novelty of the claims was “evaluating quality of each of the multiple encodings” and “selecting a plurality of resolution and target bitrate combinations . . . based on the evaluated quality.” EX1002, 116. As explained above, POSITAs knew about and used quality evaluation to select alternative streams encoded at different resolutions and bitrates long before the priority date of the 785 patent.

57. Figure 1 of the 785 patent, shown below, and the accompanying disclosure provide an example of a common, well-known adaptive bitrate streaming system at a high-level.

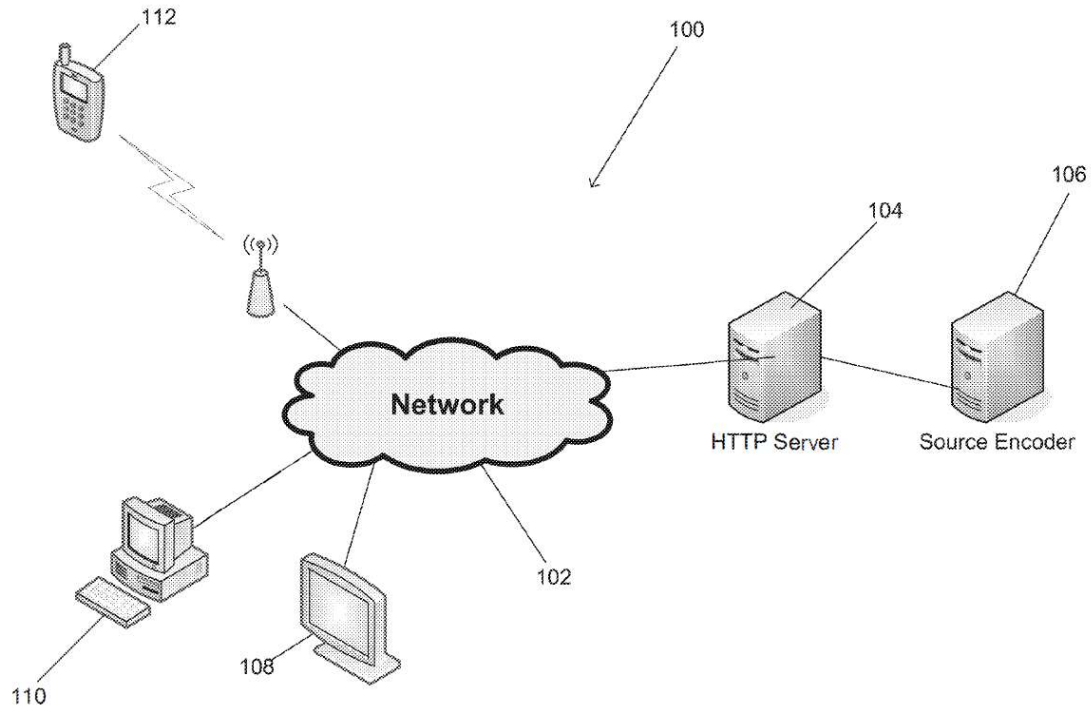


FIG. 1

EX1001, Fig. 1.

58. The adaptive streaming system includes a source encoder 106, an HTTP server 104, and several playback devices 108, 110, and 112, all connected via a network 102. *Id.*, 6:20-7:25, Fig. 1. The source encoder 106 is “configured to encode source video as a number of alternative video streams[,]” which “are streams that encode the same media content in different ways.” *Id.*, 6:21-25; *see also id.*, 7:27-31. For example, “alternative streams [may] encode media content (such as but not limited to video) at different resolution and sample aspect ratio combinations and different maximum bitrates.” *Id.*, 6:25-28. The source encoder 106 “encodes each alternative stream into a separate container file and generates a top level index

file describing each of the streams contained within the container files.” *Id.*, 6:38-42. Both “[t]he top level index file and the container files are [then] uploaded to an HTTP server 104.” *Id.*, 6:42-44.

59. Figure 1 of the 785 patent, shown above, also includes playback devices such as personal computers 110, CE players 108, and mobile phones 112. EX1001, 6:53-55, Fig. 1. But playback devices are not limited to these specific devices. EX1001, 6:55-59. The 785 patent explains “playback devices can include consumer electronics devices such as DVD players, Blu-ray players, televisions, set top boxes, video game consoles, tablets, and other devices that are capable of connecting to a server via HTTP and playing back encoded media.” *Id.* Generally, in the context of an adaptive streaming system, the playback devices operate as follows:

In the illustrated embodiment, a variety of playback devices use HTTP or another appropriate stateless protocol to request portions of a top level index file and the container files via a network 102 such as the Internet. Prior to a playback device performing adaptive bitrate streaming using portions of media from alternative streams contained within the container files, a bandwidth probe can be performed by the playback device to determine available bandwidth. Once the bandwidth probe has been completed, the playback device can utilize data within the top level index including (but not limited to) the maximum bitrate of each of the available streams to determine the initial streams from which to commence requesting portions of encoded media as part of an adaptive streaming process.

Once playback of content from the initial set of streams commences, the playback device utilizes the top level index to perform adaptive bitrate streaming of the content in response to changes in streaming

conditions. In many adaptive bitrate streaming systems, the playback device can progress through a series of operational phases in which the playback device responds differently in each phase to changes in the streaming conditions.

EX1001, 6:59-7:14.

60. In sum, the playback devices 108, 110, and 112 use HTTP or another appropriate stateless protocol to request portions of the top level index file, and the container files from the HTTP Server 104 via the network 102. *Id.* The playback devices 108, 110, and 112 decide which portion of the top level index file and corresponding container files to request from HTTP server 104 based on streaming conditions—e.g., available bandwidth. *Id.* And as streaming conditions change, the playback devices will utilize the top level index to perform adaptive streaming. *Id.* To provide an example, the playback devices request portions of a lower quality container file when available bandwidth decreases. *Id.* All of the adaptive streaming functionality described above, including that of the source encoder, the HTTP server, the playback devices, and the network, were well-known in the field of encoding long before the priority of date of the 785 patent. *See supra* § VII.B.

61. As was noted above, in an adaptive streaming system, the source encoder is “configured to encode source video as a number of alternative video streams.” EX1001, 6:21-25; *see also id.*, 7:27-31. To do so the coding parameters of alternative streams must be set. The 785 patent claims a method of generating

said encoder settings, but as will be shown below, the 785 patent’s claimed method simply describes well known steps taken in a well-known order to generate encoder settings. *See generally* EX1001; *see also id.*, Claim 1.

62. According to the 785 patent, for a video content at a given resolution and frame rate, video quality tends to reach a point of diminishing returns once the bitrate climbs past a certain bitrate threshold. EX1001, 4:28-32.

63. The 785 patent refers to the above bitrate threshold— i.e., the bitrate beyond which video quality does not improve appreciably as it is increased for a given resolution and frame rate— as the optimal target maximum bitrate. *See* EX1001, 4:29-40; *see also id.*, 8:17-20 (“Stated another way, the optimal target maximum bitrate is the lowest bitrate at which the highest quality video can be achieved at a given resolution and frame rate.”) According to the 785 patent, the optimal target maximum bit rate for a specific combination of resolution and frame rate can be determined by a formula, or it can be determined through testing for quality. EX1001, 4:34-40. Either way, the 785 patent explains choosing a specific resolution and a target bitrate combination based on the quality. EX1001, 4:34-37.

D. Level of Ordinary Skill in the Art

64. I understand that the 785 patent should be interpreted from the perspective of a person having “ordinary skill in the art” as of the effective filing date of the patent. For the purposes of this declaration, I assume that the effective

filing date of the patent application leading to the 785 patent is no earlier than August 30, 2011, and I understand that Patent Owner has not claimed an earlier priority date in the parallel litigation. I am familiar with the technological field at issue and the state of the art as of that date.

65. I understand that the education level of those in the technological field, sophistication of the technology, types of problems encountered, prior art solutions to those problems, and the speed at which innovations are made in the field may factor into establishing what the appropriate level of ordinary skill should be.

66. A POSITA would have a bachelor's or master's degree in electrical engineering, computer engineering, or a related field, with specialized training or coursework in digital video coding and processing and related technologies. The individual would have at least four years of practical experience in digital video coding and processing, including software and/or hardware architectures used to perform digital video coding and processing. A person could also have qualified, as a POSITA, with more formal education and less technical experience, or vice versa.

67. I had at least the requisite level of skill with respect to the subject matter of the challenged patent at the time of the effective filing date as demonstrated by my qualifications listed above and my CV (EX1004).

VIII. PRIOR ART TO THE CHALLENGED PATENT

A. Summary of the Prior Art

1. Ozer

68. “Video Compression for Flash, Apple Devices and HTML5” is a printed publication written by Jan Ozer and was published on May 2, 2011 (“Ozer”). EX1005, 1.

69. Ozer introduces adaptive streaming, stating that “[a]daptive streaming technologies make multiple video streams available to [] viewers and dynamically switch streams to adapt to changing connections speeds, CPU use statistics and other heuristics.” EX1005, 115. To allow for this adaptive streaming, Ozer teaches encoding a source video to produce multiple alternative streams. *Id.*, 116, Fig. 7-1.

70. This is shown in Figure 7-1 where an encoder (shown in blue below) encodes the source video (shown in red below) into multiple alternative streams (shown in green). *Id.* Specifically, Figure 7-1 of Ozer depicts the common components of all adaptive streaming systems: the encoder (shown in blue below), server and or CDN, and the client devices (shown in purple).

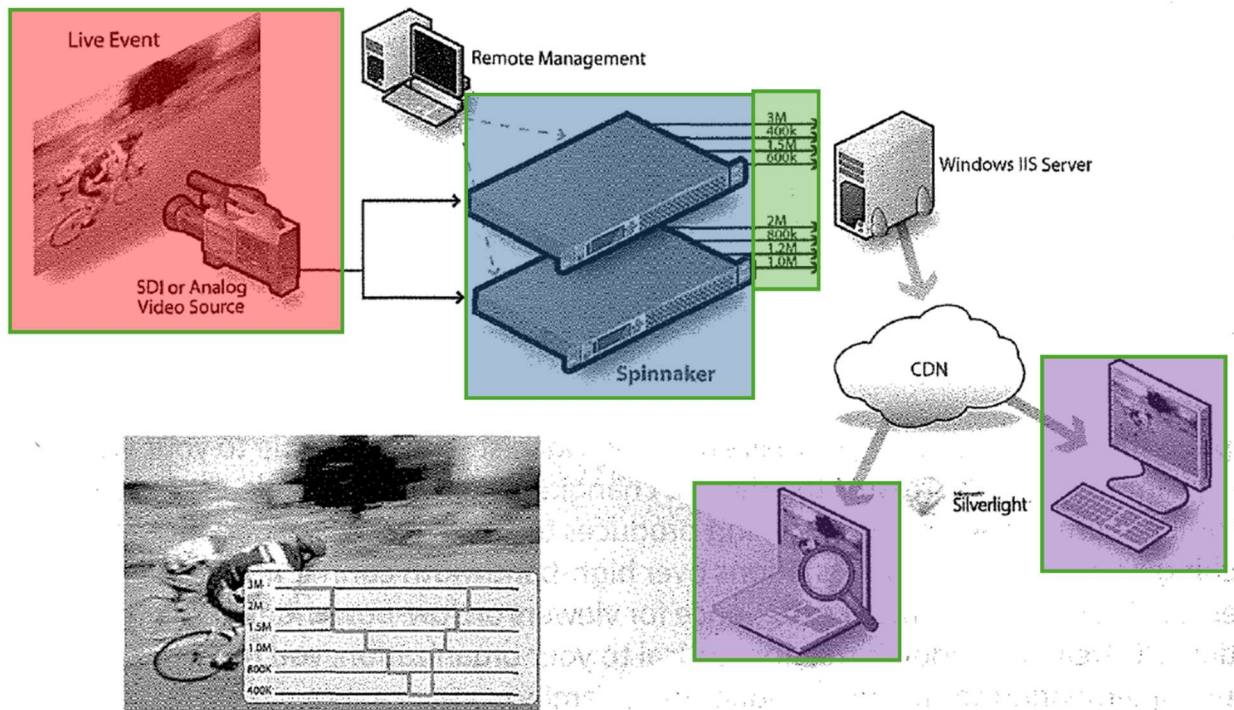


Figure 7-1. Using Inlet Technologies' Spinnaker encoder to produce multiple streams for a Microsoft Smooth Streaming event. Usage courtesy Inlet Technologies.

Id., 115-116, 118, Fig. 7-1 (annotated).

71. Each of the alternative streams is encoded using a pre-specified resolution and a video bitrate. *Id.*, 16, 127-129. The table below provides guidance as to the number of alternative streams, each having a recommended bitrate (V), audio bitrate, and total bitrate (AV), all with an eye towards switching smoothness and stream quality. *Id.*

16x9				
Stream #	Picture Size	V	A	AV
1	256x144	150	64	214
2	256x144	250	64	314
3	512x288	450	64	514
4	512x288	600	64	664
5	512x288	800	64	864
6	512x288	1200	64	1264
7	768x432	1400	64	1464
8	1280x720	1700	64	1764
9	1280x720	2500	64	2564
10	1280x720	3500	64	3564

Figure 7-4. Adobe's Levkov recommends clustering your streams around certain frame sizes.

Id., 129, Fig. 7-4.

72. When encoding a stream, Ozer suggests “choos[ing] the desired resolution, then output[ing] at the data rate necessary to produce acceptable quality at that resolution.” *Id.*, 17. As shown in the above table, for example, a picture resolution of 1280x720 can be encoded using three different bitrates 1700, 2500, and 3500. *Id.*, Fig. 7-4. Indeed, because streams of identical resolution switch most smoothly, Ozer “recommends clustering streams at specific window sizes as shown in Figure 7-4,” shown above. As can be seen in Figure 7-4 above, Ozer recommended clustering its encoder settings around four resolutions: 256x144 (2 streams), 512x288 (4 streams), 768x432 (1 stream), and 1280x720 (3 streams). *Id.*, Fig. 7-4.

73. Ozer provides examples of real-world organizations adopting this resolution clustering method, including Adobe and MTV Networks. *Id.*, 129-130, Table 7-3 and Fig. 7-5 (both depicted below).

Preset	Resolution	Video bitrate	Profile	Key Frame	Audio bitrate	Channels
Phone and Table, 3G, 16x9	512x288	300 kbps	Baseline	72	48 kbps	Stereo
Phone and Table, Low, 16x9	512x288	450 kbps	Baseline	72	48 kbps	Stereo
Phone and Table, WiFi, 16x9	512x288	650 kbps	Baseline	72	48 kbps	Stereo
PC & TV, SD, Med, 16x9	768x432	1,140 kbps	Main	72	64 kbps	Stereo
PC & TV, SD, High, 16x9	768x432	1,140 kbps	Main	72	64 kbps	Stereo
PC & TV, HD, Low, 16x9	1280x720	2,440 kbps	High	72	64 kbps	Stereo
PC & TV, HD, High, 16x9	1280x720	3,440 kbps	High	72	64 kbps	Stereo

Table 7-3. The configurations of adaptive streaming presets released by Adobe in late 2010.

Scenario	Format	Frame Size	Total Bitrate	Audio Bitrate	bits/pixel *frame @ 30 fps	bits/pixel *frame @ 24 fps
Mobile & constrained (low)	baseline, mono, 10 fps	448x252	150	48	0.09	0.09
Mobile & constrained (high)	baseline, mono	448x252	450	48	0.12	0.15
Sidebar placements	main profile, stereo	384x216	400	96	0.12	0.15
Small In-page	main profile, stereo	512x288	750	96	0.15	0.18
Medium In-page	main profile, stereo	640x360	1200	96	0.16	0.20
Large In-page	main profile, stereo	768x432	1700	96	0.16	0.20
Full size In-page	main profile, stereo	960x540	2200	96	0.14	0.17
HD 720p (full screen)	high profile, stereo	1280x720	3500	96	0.12	0.15

Figure 7-5. MTV's encoding schema targets video resolutions to specific window sizes. Image courtesy MTV.

74. Ozer also offers specific recommendations regarding selection of bit rates, referred to as video data rates. *Id.*, 128-130. Ozer explains that “encoding too few streams with a huge gap in bit rates would not provide the best quality or the most optimal stream for a particular bandwidth environment” but “[k]eeping too

many bit rates close to one another could result in too many stream switches.” *Id.*, 130. Ozer also explains that “[t]he most successful tests [] conducted include a number of streams with fairly close-together bit rates.” *Id.*, 130-31. The quality of each of these encoded streams can then be evaluated and a bitrate selected. *Id.*, 154, 156.

75. The resulting streams are split into 2 to 10-second chunks. *Id.*, 118, 136, Fig. 7-9. An index file (also known as a manifest file) identifies and describes the chunks of the streams and their locations. *Id.*, 118, 134-41. The chunks are provided to the playback devices upon request. *Id.* Figure 7-2 of Ozer depicts a manifest file as used by HTTP-based systems like Apple’s HTTP Live Streaming. *Id.*, 118; *see also id.*, 136-138 (describing Apple’s HTTP Live Streaming). In these systems, the manifest files are stored on an a CDN and, as noted above, are provided to playback devices upon request. *Id.*, 118; *see also id.*, 136-138 (describing Apple’s HTTP Live Streaming). As shown below, the manifest file identifies the available files (Alternate-Low, Alternate-Mid, and Alternate-hi) and their locations on the CDN and refers to specific chunks of each stream (e.g., Low_01.ts, Low_02.ts, and so on). *Id.*, 118, Fig. 7-2.

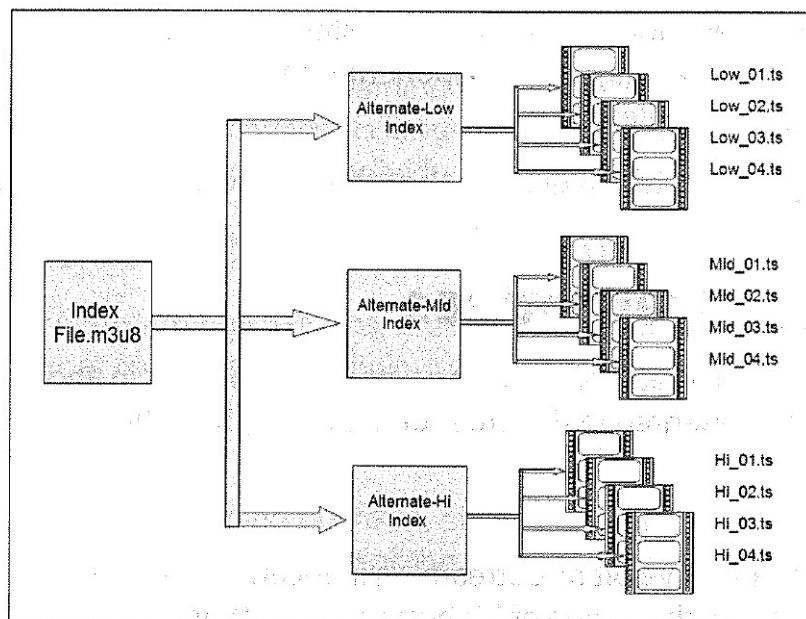


Figure 7-2. The manifest file identifies the different files and their locations. From an article that I wrote for Millimeter Magazine, and usage is courtesy Millimeter.

Id., 117, Fig. 7-2.

76. Once the playback devices obtain the manifest file from the one or more servers, the playback devices use the index file (manifest file) to identify the desired chunks of streams and request them from the server or CDN. *See id.*, 117-18, 136-38, Fig. 7-2, Fig. 7-9. Specifically, Ozer explains that “[t]he player might start out playing the lowest quality stream (Low_01.ts), but if it determines [sic] that it can handle a higher quality stream, it would check the manifest file for the next chunk of a higher-quality stream (say, Mid_02.ts) and retrieve and play that chunk.” *Id.*, 118.

77. The server or CDN receives these requests for chunks of the alternative streams from the playback devices, and provides the requested chunks of the alternative streams to the playback devices. *Id.*, 117-18, 134-41.

2. Liao

78. United States Patent No. 9,832,540 B2, titled “Method and Apparatus for Hybrid Transcoding of a Media Program” (“Liao”, EX1006), to Yupeng Liao, Kang Kang, and Zhibing Wang, and was filed for on December 15, 2010. EX1006, 1.

79. Liao describes “a method and apparatus for encoding a source media program.” *Id.*, 1:47-49. The invention of Liao seeks to encode source media programs in such a way as to maximize the use of available bandwidth of the communications link, and to produce video streams at a constant quality. *Id.*, 1:38-43, 7:63-9:16, 9:45-57.

80. Liao describes preparing alternative streams of video content for use in an adaptive streaming system. *Id.*, 7:63-8:26. Specifically, Liao explains that “[s]ince customers’ Internet service varies widely in bandwidth, different versions of the media program can be generated, each with a different resolution, typically 480p, 360p, 288p or 240p.” *Id.*, 7:67-8:3. This allows “[l]ower resolution versions (e.g., 240p) [to be] transmitted when the bandwidth of the communications link is lower and higher resolution versions (e.g. 480p) [to be] transmitted when

communications bandwidth permits.” *Id.*, 8:3-6. Liao explains that “[t]his functionality is typically implemented by the media program player 304 which selects the appropriate version based upon estimated bandwidth.” *Id.*, 8:6-9.

81. When preparing these alternative streams, each with a different resolution, Liao explains that the convention is to encode each of the alternative streams at the same bit rate, regardless of content. *Id.*, 8:10-13. “For example, 360p transcoded versions of the media program series ‘FAMILY GUY’ and the media program series ‘PRISON BREAK’ may both be transcoded using the same bit rate.” *Id.*, 8:14-17. But according to Liao, “since these media programs have different characteristics (one has a great deal more movement and action), the media programs should be encoded with different bit rates to account for these different characteristics, while still achieving the same image quality.” *Id.*, 8:17-22.

82. To determine the appropriate bitrate, Liao employs a measure known as structural similarity index (SSIM) to evaluate the quality of the encodings. *Id.*, 8:38-9:16, Fig. 6. SSIM “is a measure of the similarity between two images.” *Id.*, 8:41-42. It “is computed on windows of an image (typically applied only over luminescence information), wherein a measure of similarity between two windows x and y of common size $N \times N$ is defined as:”

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

wherein:

μ_x =the average of x;

μ_y =the average of y;

σ_x^2 =the variance of x;

σ_y^2 =the variance of y;

σ_{xy} =the covariance of x and y;

$c_1=(k_1L)^2$, $c_2=(k_2L)^2$ two variables to stabilize the division with weak denominator;

L=dynamic range of the pixel values

$$\left(\text{typically } 2^{\frac{\#bits}{pixel}} - 1 \right);$$

$k_1=0.01$; and

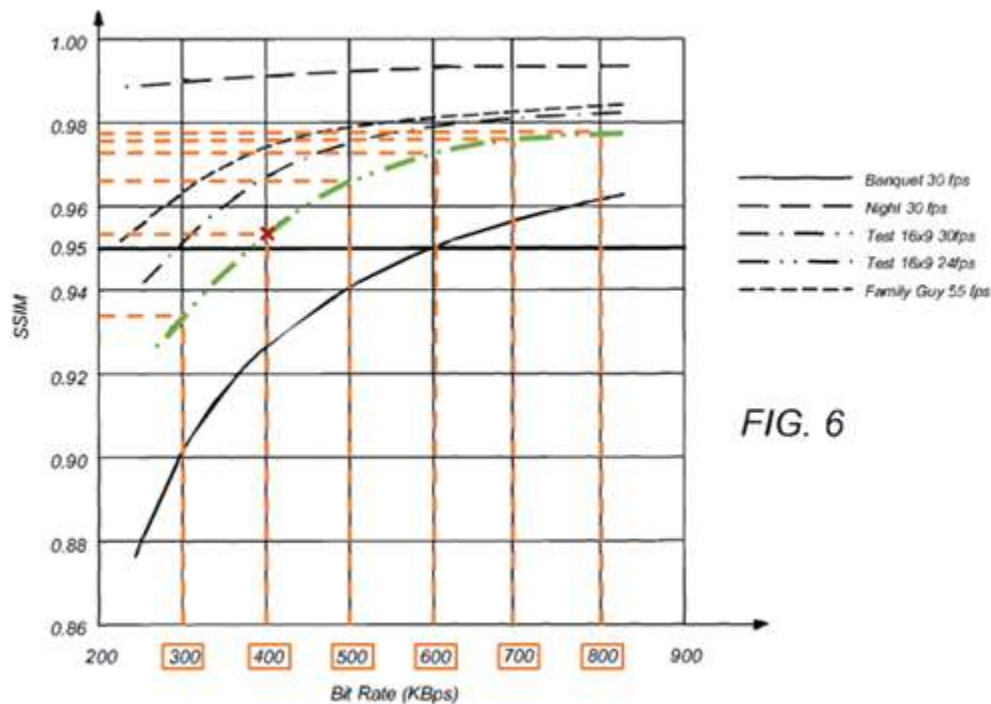
$k_2=0.03$.

Id., 8:42-67.

83. Figure 6 of Liao below, evaluates the quality of and compares the video encodings of 5 different media programs using SSIM (a video quality heuristic). *Id.*, 8:38-9:16, Fig. 6. Specifically, each of the 5 media programs are encoded at several different bit rates, and each of the encodings are then evaluated using SSIM so that the appropriate coding bitrate for the media program can be determined. *Id.*, 8:17-22, 9:7-16, Fig. 6.

84. For example, Figure 6 depicts a TEST media program (shown below in green) encoded multiple times at different bitrates ranging from 300 to 800 kbps and an SSIM is calculated for each of the different bitrates (ranging from 0.93 to

0.98). *Id.* Liao then teaches selecting the encoding having the lowest bitrate that achieves a minimum SSIM of 0.95. *Id.* As shown in the below figure, a bitrate of 400 kbps achieves this minimum SSIM.



Id., Fig. 6 (annotated).

3. Gu

85. U.S. Patent No. 8,396,114 to Chuang Gu *et al.*, titled “Multiple Bit Rate Video Encoding Using Variable Bit Rate and Dynamic Resolution for Adaptive Video Streaming,” was filed for on January 29, 2009, and published on July 29, 2010 (“Gu”). EX1007, 1.

86. Gu discloses a video encoding system for streaming videos having multiple bitrates. *Id.*, 2:18-35. Gu explains that the target bitrates for alternative

streams should “decrease progressively in a proportional, logarithmic or other decreasing manner.” *Id.*, 7:58-61.

87. To do so, Gu recommends that the target bitrate of each alternate stream is “decreased proportionally” by “ $\frac{1}{4}$ ” (25%) from the highest bitrate alternate stream. *Id.*, 7:61-64. In practice, if four alternate streams were to be encoded, where the highest alternative bitrate stream has a bitrate of 1000 kbps, the remaining three streams would have bitrates of 750 kbps, 500 kbps, and 250 kbps, respectively. Gu explains that having bitrates that proportionally decrease by “ $\frac{1}{4}$ ” or 25% result in “swift video switching.” *Id.*, 7:64-67.

4. Ronca

88. U.S. Patent Pub. No. 2011/0246616 A1, invented by David Ronca et al., is entitled “Dynamic Virtual Chunking of Streaming Media Content.” (“Ronca”) EX1008, 1. Ronca was filed for on April 2, 2010, and published on October 6, 2011. *Id.*

89. Ronca describes an adaptive bitrate streaming method that encodes, stores, and distributes video content. *Id.*, Abstract, ¶¶ 24, 38-39, 42. An “encoding tool” is used to “generate multiple video (and potentially audio) encodings of the same media presentation, each encoded at different bitrates.” *Id.*, ¶¶ 38-39. This creates a “number of different alternate streams containing the same material

encoded at a variety of data rates.” *Id.*, ¶ 24. These alternate stream are each uploaded to a content delivery network where they are stored. *Id.*, ¶¶ 28-36.

90. A header file with an index is then generated that “defin[es] what chunks, or segments of multiplexed audiovisual content, are present within that stream,” and provides their locations. *Id.*, ¶¶ 22, 41-42, 60-61. Because of the potential number of chunks that would be necessary to store larger alternative streams, Ronca provides an example in which these streams are physically stored as one piece in memory. But, to accommodate streaming schemes such as HTTP Live Streaming, the header file with an index still virtually breaks up each of the alternative streams into chunks. *Id.*, ¶¶ 25-26. During playback a client device uses this index to selectively request specific chunks of specific alternative streams. *Id.*, ¶¶ 22, 60-62, Fig. 8.

B. General Motivations to Combine Prior Art References

1. Ozer in view of Liao

91. A POSITA would have been motivated to combine Ozer with Liao. Specifically, a POSITA would combine Ozer’s architecture, including streaming, encoding, and indexing, with Liao’s quality evaluation. *See* EX1005, 115-144; EX1006, 7:63-8:26.

92. Liao and Ozer disclose creating alternate versions of source video content to be used in adaptive streaming. EX1005, 115-16; EX1006, 7:63-8:9.

While encoding the source content to create these versions, Ozer explains that determining the number of alternative streams to generate, and maintaining quality across streams is important. EX1005,127. Specifically, selecting appropriate encoding bitrates for each stream is crucial. *Id.*, 15-17, 126-127. “[C]hoosing a[n encoding bitrate] is largely a cost/quality trade-off.” *Id.*, 15. The higher the bitrate, generally, the higher the quality of the encoded stream, and the more network bandwidth/resources that are needed to transmit the stream. *Id.* Thus, the “cost” of increasing encoding bitrate is transmission time and effort, while the benefit is that the “quality” of each video encoding improves. *Id.*, 15-17, 126-127.

93. Liao teaches how to optimize bitrate selection so that high quality streams are available that will not overburden transmission resources. Specifically, Liao teaches using the SSIM heuristic to compare the quality of encodings. EX1006, 8:38-9:16. For example, as shown in the annotated version of Figure 6 below, Liao encodes a TEST media program (annotated in green) at different target bit rates (*e.g.*, 300-800 kbps), and evaluates the quality of each encoding using SSIM. *Id.*, 9:13-16, FIG. 6 (annotated). Because the quality level of each encoding is determined, Liao teaches selecting the lowest bitrate corresponding to some minimum quality threshold. *Id.*, 9:13-16, Fig. 6 (annotated). For example, if an SSIM of 0.95 is desired, then a 400 kbps bitrate stream may be selected, and the adaptive streaming system need not waste additional transmission resources on 500, 600, 700, or 800

kbps streams. *Id.*, 9:13-16, Fig. 6 (annotated). In other words, the 400 kbps bitrate stream represents the best tradeoff of quality for ease of transmission because it is the lowest bitrate stream to achieve the desired quality level; the lower bitrate 300 kbps stream does not achieve the minimum SSIM requirement (quality level), while the higher bitrate encodings (500, 600, 700, or 800 kbps) are harder to transmit and require more resources to transmit than the 400 kbps stream. *Id.*, 8:38-9:16; EX1005, 15-17, 126-127. Although bitrate encodings (500, 600, 700, or 800 kbps) result in a higher SSIM than the 400 kbps, the 400 kbps meets the minimum SSIM requirement, and any additional resources devoted to it would result in minimal quality gain compared to the resources that need to be exerted. EX1006, 8:38-9:16.

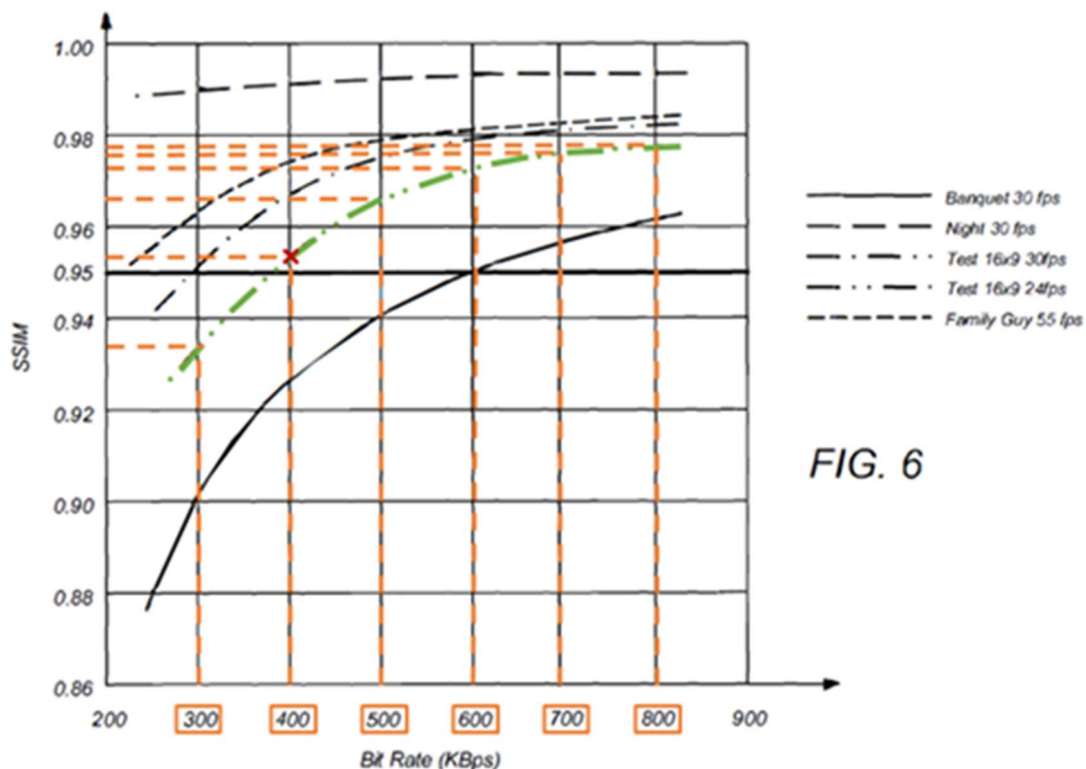


FIG. 6

Id., Fig. 6 (annotated).

94. Thus, Ozer in view of Liao will select the optimal bitrate streams among all encodings, with each encoding having its own bitrate and resolution combination. Liao's evaluation process may be repeated for each pre-specified encoding resolution to determine the optimal bitrate stream for a specific resolution, which is the lowest bitrate stream that meets a pre-determined SSIM threshold. *Id.*, 8:38-9:16; EX1005, 127-132. Because, a user may specify the target SSIM, a plurality of target SSIMs may also be specified if streams of different quality levels (e.g. a low quality, medium quality, and high quality stream) are desired for adaptive streaming. EX1006, 8:38-9:16; EX1005, FIG. 7-2, 127-132, 117-18, 136-37; EX1014, 3-4, 7-8, 10.

95. A POSITA combining the teachings of Ozer and Liao would have had a reasonable expectation of success. Such a combination simply requires using the quality evaluation and selection of Liao (EX1006, 8:38-9:16) to determine which combinations of resolution and target bit rate to determine the settings for the adaptive streaming encoder taught in Ozer. EX1005, 115-144, Table 7-3, Figs. 7-4, 7-5; EX1014, 3-4, 7-8.

96. First, source video is encoded at a plurality of resolution and target bitrate combinations, to generate encodings as demonstrated by Ozer, as shown below. EX1005, 128-130, Table 7-3, Figs. 7-4, 7-5; *see also* EX1014, 3-4, 7-8, 10.

16x9				
Stream #	Picture Size	V	A	AV
1	256x144	150	64	214
2	256x144	250	64	314
3	512x288	450	64	514
4	512x288	600	64	664
5	512x288	800	64	864
6	512x288	1200	64	1264
7	768x432	1400	64	1464
8	1280x720	1700	64	1764
9	1280x720	2500	64	2564
10	1280x720	3500	64	3564

Figure 7-4. Adobe's Levkov recommends clustering your streams around certain frame sizes.

Preset	Resolution	Video bitrate	Profile	Key Frame	Audio bitrate	Channels
Phone and Table, 3G, 16x9	512x288	300 kbps	Baseline	72	48 kbps	Stereo
Phone and Table, Low, 16x9	512x288	450 kbps	Baseline	72	48 kbps	Stereo
Phone and Table, WiFi, 16x9	512x288	650 kbps	Baseline	72	48 kbps	Stereo
PC & TV, SD, Med, 16x9	768x432	1,140 kbps	Main	72	64 kbps	Stereo
PC & TV, SD, High, 16x9	768x432	1,140 kbps	Main	72	64 kbps	Stereo
PC & TV, HD, Low, 16x9	1280x720	2,440 kbps	High	72	64 kbps	Stereo
PC & TV, HD, High, 16x9	1280x720	3,440 kbps	High	72	64 kbps	Stereo

Table 7-3. The configurations of adaptive streaming presets released by Adobe in late 2010.

Scenario	Format	Frame Size	Total Bitrate	Audio Bitrate	bits/pixel *frame @ 30 fps	bits/pixel *frame @ 24 fps
Mobile & constrained (low)	baseline, mono, 10 fps	448x252	150	48	0.09	0.09
Mobile & constrained (high)	baseline, mono	448x252	450	48	0.12	0.15
Sidebar placements	main profile, stereo	384x216	400	96	0.12	0.15
Small in-page	main profile, stereo	512x288	750	96	0.15	0.18
Medium in-page	main profile, stereo	640x360	1200	96	0.16	0.20
Large in-page	main profile, stereo	768x432	1700	96	0.16	0.20
Full size in-page	main profile, stereo	960x540	2200	96	0.14	0.17
HD 720p (full screen)	high profile, stereo	1280x720	3500	96	0.12	0.15

Figure 7-5. MTV's encoding schema targets video resolutions to specific window sizes. Image courtesy MTV.

EX1005, Table 7-3, Figs. 7-4, 7-5.

97. After the encodings are generated, the SSIM of each is calculated and evaluated using the below SSIM formula. EX1006, 8:42-9:16; EX1014, 3-4, 7-8.

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

wherein:

μ_x =the average of x;

μ_y =the average of y;

σ_x^2 =the variance of x;

σ_y^2 =the variance of y;

σ_{xy} =the covariance of x and y;

$c_1=(k_1L)^2$, $c_2=(k_2L)^2$ two variables to stabilize the division with weak denominator;

L=dynamic range of the pixel values

$$\left(\text{typically } 2^{\frac{\#bits}{pixel}} - 1 \right);$$

$k_1=0.01$; and

$k_2=0.03$.

98. These quality evaluations are then used to compare the encodings and select one or more encodings having a specific bitrate-resolution combination. EX1006, 8:42-9:16; EX1014, 3-4, 7-8. For example, for each encoding resolution, SSIM may be used to identify the lowest bitrate encoding that achieves a pre-specified SSIM as discussed above. EX1006, 8:38-9:16; EX1005, 127-132. If multiple quality levels for each resolution are desired, then the process may be repeated for additional SSIM thresholds. EX1006, 8:38-9:16; EX1005, Fig. 7-2, 127-132, 117-18, 136-37.

99. After the combinations of resolution and target bit rates are selected using SSIM, the encoded selected combinations may be uploaded to the content distribution network of Ozer and streamed using Ozer's adaptive bitrate system, and/or the encoder settings of the adaptive bitrate system may be set according to the quality evaluations. EX1005, 136, Fig. 7-9.

100. The combination of Ozer and Liao requires no change or modification in any of the underlying technology disclosed by either system. It just adds the extra evaluation and selections step taught by Liao (EX1006, 8:38-9:16) to Ozer's adaptive streaming system and encoder (EX1005, 115-144, Table 7-3, Figs. 7-4, 7-5). Adding this extra selection step during encoding poses no technical complication. Indeed, Ozer teaches that encoders are customizable for specific streaming protocols or implementations. EX1005, 118. Adding Liao's quality

evaluation process (EX1006, 8:38-9:16) to Ozer would ensure that the combined system selects optimal resolution-bitrate combinations that achieve a pre-specified quality level but do not unnecessarily expend transmission. Thus, a POSITA would therefore know how to combine the encoding, indexing, and streaming of Ozer with the quality evaluation of Liao with minimal design impact. *See, e.g.*, EX1014, 3-4, 7-8, 10.

2. Ozer in view of Liao and Gu

101. As was noted above, a POSITA would have combined Ozer's architecture, including streaming, indexing, and encoding, with the quality evaluation of Liao, to select bitrate-resolution combinations of appropriate quality. This ensures not wasting transmission resources by transmitting unnecessarily high bitrate encodings. *Supra* § VIII.B.1. A POSITA would also have been motivated to include Gu's teachings of choosing initial combinations of resolution and a plurality of target bitrates to the combination of Ozer and Liao.

102. Ozer discloses that the optimal number of streams and target encoding bitrates of these streams must be identified to assure a good viewing experience. EX1005, 127-28, 130-31. "[K]eeping too many bitrates close to one another could result in too many stream switches, even with smaller bandwidth fluctuations." *Id.*, 130. But, "encoding too few streams with a huge gap in bitrates would not provide the best . . . stream for a particular bandwidth environment" because the abrupt

changes in quality rates during stream switching would lead to distracting and very noticeable video quality changes for the viewer. *Id.*

103. Gu teaches a way to identify a reasonable gap between encoding bitrates in different streams. Specifically, Gu recommends decreasing “progressively in a proportional, logarithmic or other decreasing manner,” the target encoding bitrates for a plurality of alternative streams. EX1007, 7:58-61. For example, the target bitrate of each alternate stream may iteratively be “decreased proportionally” by 25% relative to the highest bitrate (e.g. encoding the first stream at 75% of the bitrate of the highest bitrate stream, the second at 50% of the bitrate of the highest encoded stream, etc.). *Id.*, 7:61-64. Gu explains that having bitrates that proportionally decrease by, for example, “ $\frac{1}{4}$ ” or 25%, results in “swift video switching” that improves the viewing experience. *Id.*, 7:64-67. Proportionally decreasing the bitrate gap between alternate encodings will ensure that there is a reasonable gap between alternate encodings. *Id.*, 7:64-67; EX1005, 127-28, 130-31. A POSITA would have used the teachings of Liao to evaluate and maintain the quality of those streams by selecting an encoding bitrate that achieves a pre-specified quality level without unnecessarily expending transmission resources. EX1006, 8:38-9:16; EX1014, 3-4, 7-8; *see generally* EX1005, FIG. 7-2, 127-132, 117-18, 136-37. A POSITA combining the teachings of Ozer, Liao, and Gu would have had

a reasonable expectation of success in implementing these well-known encoding techniques.

3. Ronca in view of Liao

104. A POSITA would have been motivated to combine Ronca with Liao. Specifically, a POSITA would have sought to combine the architecture of Ronca, including streaming, indexing, and encoding, with the encoding and quality evaluation of Liao. *See* EX1008, ¶¶ 24, 38-39, 42; EX1006, 7:63-8:26.

105. Both Ronca and Liao teach encoding multiple different versions of a media program or a source video to be used in adaptive streaming. EX1008, ¶ 38; EX1006, 7:63-8:9. Liao teaches a way to ensure consistent video quality across the different streams. EX1006, 8:38-9:16. Specifically, Liao teaches evaluating the quality (SSIM) of each of the different streams at different target bit rates, and then selecting a resolution and bitrate combination and meets a minimum desired SSIM. *Id.*, 9:13-16, Fig. 6; *see supra* § VIII.A.2. A POSITA would have understood that setting a minimum quality level for each of the alternative streams, as taught in Liao, would improve the viewing experience when switching between the alternative streams, as taught in Ronca. EX1014, 3-4, 7-8, 10. It would also allow the ideal resolution-bitrate combination to be selected as discussed above for Ozer, by evaluating the lowest bitrate necessary to achieve a pre-specified quality level so that transmission resources are not unnecessarily wasted on streams of a higher bitrate.

EX1006, 8:38-9:16; EX1014, 3-4, 7-8, 10. For example, if a minimum SSIM of 0.95 is desired, and there exists a 300 kbps stream that does not meet the minimum SSIM of .95 and there exists 400 500, 600, 700, or 800 kbps streams that meet the minimum SSIM of .95, then a 400 kbps bitrate stream may be selected, and the adaptive streaming system need not waste additional transmission resources on 500, 600, 700, or 800 kbps streams.. *Id.*, 9:13-16, Fig. 6. In other words, the 400 kbps bitrate stream represents the best tradeoff of quality for ease of transmission because it is the lowest bitrate stream to achieve the desired quality level; the lower bitrate 300 kbps stream does not achieve the minimum SSIM requirement (quality level), while the higher bitrate encodings (500, 600, 700, or 800 kbps) are harder to transmit and require more resources to transmit than the 400 kbps stream. EX1006, 8:38-9:16; EX1005, 15-17, 126-127.

106. A POSITA combining the teachings of Ronca and Liao would have had a reasonable expectation of success because it would simply require using the quality evaluation of Liao to select from the multiple video encodings generated by the adaptive streaming encoder of Ronca. *See, e.g.*, EX1014, 3-4, 7-8, 10. Indeed, Ronca's encoder is adaptable "for a broad variety of streaming media protocols." EX1008, ¶ 27. Ronca also already discloses an encoder used to generate multiple video encodings or steams of the same media file having different bitrates. *Id.*, ¶ 38. A POSITA would simply incorporate Liao's quality evaluation (SSIM) to compare

these encodings at different bitrates and resolutions to select an optimal set of alternative streams that would ensure quality. EX1006, 8:38-9:16; *see also, e.g.*, EX1014, 3-4, 7-8.

4. Ronca in view of Liao and Gu

107. As was noted above, a POSITA would have combined Ronca's architecture, including streaming, indexing, and encoding, with the encoding and quality evaluation of Liao, to maintain quality across the different encodings used in adaptive streaming. *Supra* §VIII.B.3. A POSITA would also have been motivated to include Gu's teachings of choosing initial encoding resolution-bitrate combinations.

108. Both Ronca and Liao teach encoding alternate versions of source content to be used in adaptive streaming. EX1008, ¶38; EX1006, 7:63-8:9. Ronca "generat[ing] multiple video...encodings of the same media presentation, each encoded at different bit rates" using an encoder. EX1008, ¶¶24, 38. Liao teaches that in addition to encoding at "different bit rates," each of these encodings should have "a different resolution, typically 480p, 360p, 288p, 240p" to account for variations in "customers' Internet service" during adaptive streaming. EX1006, 7:67-8:6, 8:17-22, 8:38-40. Gu describes a way to determine a reasonable spacing between these different bitrates.

109. Specifically, Gu describes varying the target bitrates for alternative streams by iteratively “decreas[ing] progressively in a proportional, logarithmic or other decreasing manner” the target bitrate of each successive stream. EX1007, 7:58-61. In one example Gu recommends that the target bit rate of each encoding may be “decreased proportionally” by 25% (or $\frac{1}{4}$) from the highest bitrate streams (e.g. encoding a first stream at the highest bitrate, the second at 75% of the first stream’s bitrate, and a third stream at 50% of the first stream’s bitrate). *Id.*, 7:61-64. Gu explains that having bitrates that proportionally decrease by, for example, “ $\frac{1}{4}$ ” or 25% results in “swift video switching.” *Id.*, 7:64-67. As discussed above for Ground 2, a POSITA would have looked to Gu’s teaching of decreasing the target bit rates for alternate streams by a predetermined percentage such as 25% to determine the proper spacing between the streams and target bitrates. *Supra* §VIII.B.2. A POSITA would have looked to Gu’s teaching of 25% decrease in bitrate as an example of reasonable spacing between the streams and target bitrates taught by Ronca and Liao. A POSITA combining the teachings of Ronca, Liao, and Gu would have had a reasonable expectation of success to implement these well-known encoding techniques.

IX. INVALIDITY

A. Ground 1: Claims 1-4, 7-14, and 17-19 of 785 Patent are Rendered Obvious by Ozer in view of Liao

1. Claim 1

- (1) [1.pre] “A method for encoding source content as a plurality of alternative video streams and adaptively streaming the plurality of alternative video streams, each having a bitrate, the method comprising:”

110. To the extent the preamble of claim 1 is limiting, the combination of Ozer in view of Liao and general knowledge of a POSITA, as described below, renders claim 1 obvious.

111. Ozer discloses that “[a]daptive streaming technologies make multiple video streams available for [for] viewers.” EX1005, 115-16. Specifically, adaptive streaming allows viewers’ playback devices to “dynamically switch” between alternatively encoded “streams to adapt to changing connection speeds, CPU use statistics and other heuristics.” *Id.*

112. The first step of “[a]daptive streaming” is encoding source video (annotated in red below) using an encoder (annotated in blue below) into a plurality of alternative streams (annotated in green below) each having a pre-specified bitrate (*e.g.*, 3 M (“mbps”), 400 kbps, 1.5 mbps, 600 kbps). *Id.*, 116, Fig. 7-1 (annotated). Thus, the combination of Ozer and Liao discloses the preamble.

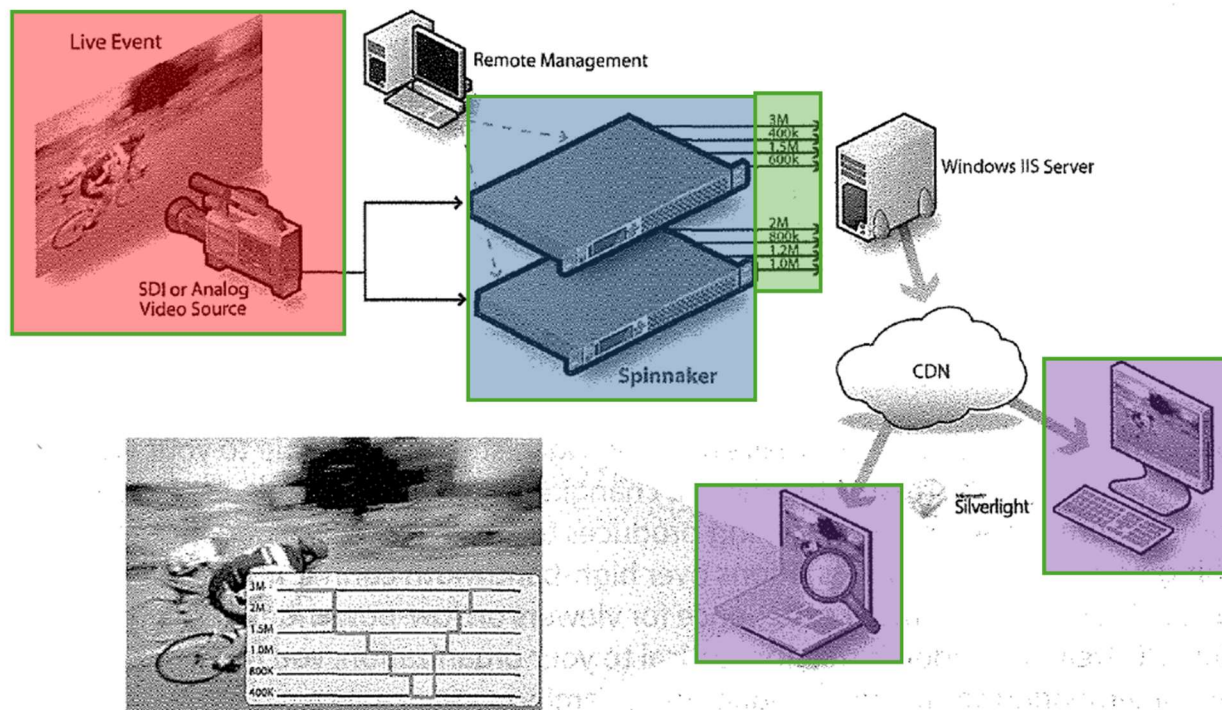


Figure 7-1. Using Inlet Technologies' Spinnaker encoder to produce multiple streams for a Microsoft Smooth Streaming event. Usage courtesy Inlet Technologies.

Id., 116, Fig. 7-1 (annotated).

(2) [1.a] “identifying source video content;”

113. The combination of Ozer in view of Liao and general knowledge of a POSITA, as described below, renders limitation [1.a] obvious.

114. Figure 7-1 of Ozer discloses “source video content” (annotated in red) in the form of a “Live Event” recorded using an “SDI [serial digital interface] or Analog Video Source.” EX1005, 115-116, Figure 7-1. This “source video content” is fed to the encoder (annotated in blue) where it is encoded into a plurality of alternate streams (annotated in green). *Id.* While Figure 7-1 is directed to “source

video content” in the form of a “Live Event,” Ozer discloses that “source video content” may also include “on-demand video files.” *Id.*, 116.

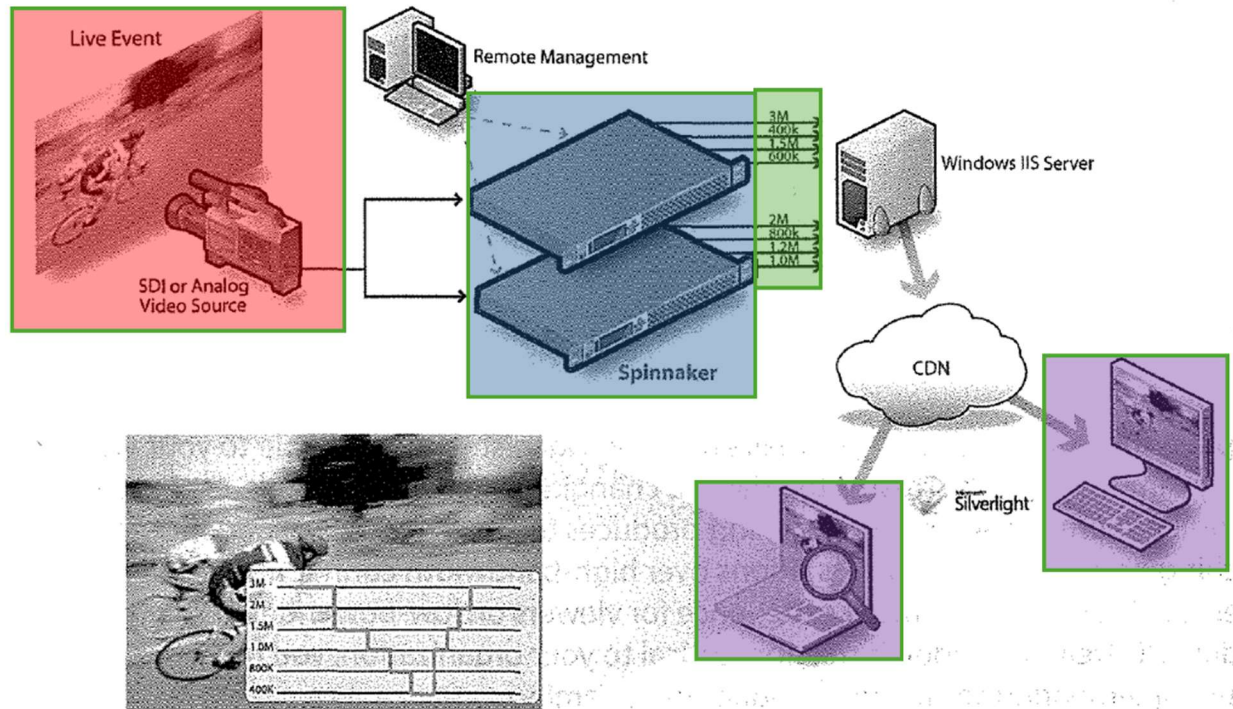


Figure 7-1. Using Inlet Technologies' Spinnaker encoder to produce multiple streams for a Microsoft Smooth Streaming event. Usage courtesy Inlet Technologies.

Id., 116, Fig. 7-1 (annotated).

- (3) [1.b] “identifying a plurality of resolutions for encoding of video content;”

115. The combination of Ozer in view of Liao and general knowledge of a POSITA, as described below, renders limitation [1.b] obvious.

116. Specifically, Ozer teaches that “source video content” may be encoded using multiple video resolutions. EX1005, 128-130. Adaptive streaming inherently includes encoding “source video content” at different resolutions so that a client

device may dynamically adjust between streams of different resolutions as necessary based on network conditions and device resources. *Id.*, 128 (disclosing that there “will inevitably be some switching between streams with different frame sizes” (resolutions) in adaptive bitrate streaming).

117. While adaptive streaming requires encoding video at a plurality of different resolutions, Ozer advises clustering encoding parameters around a smaller subset of resolutions to ensure smooth transitions during streaming. *Id.*, 128. For example, video may be encoded at four resolutions (picture sizes): 256x144, 512x288, 768x432, and 1280x720. *Id.* at 128-29, Fig. 7-4.

16x9				
Stream #	Picture Size	V	A	AV
1	256x144	150	64	214
2	256x144	250	64	314
3	512x288	450	64	514
4	512x288	600	64	664
5	512x288	800	64	864
6	512x288	1200	64	1264
7	768x432	1400	64	1464
8	1280x720	1700	64	1764
9	1280x720	2500	64	2564
10	1280x720	3500	64	3564

Figure 7-4. Adobe's Levkov recommends clustering your streams around certain frame sizes.

Id., Fig. 7-4 (annotated).

118. Alternatively, video may be encoded at three resolutions: 512x288, 768x432, and 1280x720. *Id.* at 128-29, Table 7-3.

Preset	Resolution	Video bitrate	Profile	Key Frame	Audio bitrate	Channels
Phone and Table, 3G, 16x9	512x288	300 kbps	Baseline	72	48 kbps	Stereo
Phone and Table, Low, 16x9	512x288	450 kbps	Baseline	72	48 kbps	Stereo
Phone and Table, WiFi, 16x9	512x288	650 kbps	Baseline	72	48 kbps	Stereo
PC & TV, SD, Med, 16x9	768x432	1,140 kbps	Main	72	64 kbps	Stereo
PC & TV, SD, High, 16x9	768x432	1,140 kbps	Main	72	64 kbps	Stereo
PC & TV, HD, Low, 16x9	1280x720	2,440 kbps	High	72	64 kbps	Stereo
PC & TV, HD, High, 16x9	1280x720	3,440 kbps	High	72	64 kbps	Stereo

Table 7-3. The configurations of adaptive streaming presets released by Adobe in late 2010.

Id., Table 7-3 (annotated).

119. Accordingly, Ozer and Liao disclose “identifying a plurality of resolutions for encoding of video content.”

- (4) [1.c(i)] “for each particular resolution of the plurality of resolutions: encoding at least a portion of the source video content multiple times using the particular resolution and multiple different target bitrates for the particular resolution; and”

120. The combination of Ozer in view of Liao and general knowledge of a POSITA, as described below, renders limitation [1.c(i)] obvious.

121. Ozer discloses “identifying a plurality of resolutions for encoding of video content.” *Supra* § IX.A.1(3). Each of these resolutions may be encoded into a plurality of alternative encodings with a plurality of different “targeted bit[]rates.” EX1005, 128, Fig. 7-4; *see also id.*, Table 7-3.

122. For example, Fig. 7-4 of Ozer discloses source video encoded into four resolutions/Picture Size (256x144, 512x288, 768x432, and 1280x720), each of which is encoded at anywhere from one to four different bitrates resulting in ten

alternative streams of the source content. *Id.*, 128-29, Fig. 7-4 (annotated).⁴ The streams encoded with a resolution of 256x144 (streams 1 and 2) are encoded at “targeted bit[]rates” of 150 kbps and 250 kbps. *Id.*, 128-29, Fig. 7-4 (annotated). The streams encoded with a resolution of 512x288 (streams 3-6) are encoded at “targeted bit[]rates”: 450 kbps, 600 kbps, 800 kbps, and 1200 kbps. *Id.*, 128-29, Fig. 7-4 (annotated). And the streams encoded with a resolution of 1280x720 (streams 8-10) are encoded at “targeted bit[]rates” of 1700 kbps, 2500 kbps, and 3500 kbps. *Id.*, 128-29, Fig. 7-4 (annotated).

16x9				
Stream #	Picture Size	V	A	AV
1	256x144	150	64	214
2	256x144	250	64	314
3	512x288	450	64	514
4	512x288	600	64	664
5	512x288	800	64	864
6	512x288	1200	64	1264
7	768x432	1400	64	1464
8	1280x720	1700	64	1764
9	1280x720	2500	64	2564
10	1280x720	3500	64	3564

Figure 7-4. Adobe's Levkov recommends clustering your streams around certain frame sizes.

⁴ A POSITA would understand that the “A” column in Fig. 7-4 represents audio bitrate while the “V” column represents video bitrate.

Id., Fig. 7-4 (annotated).

123. Thus, it would have been obvious to a POSITA that Ozer discloses this limitation by showing that for *each* particular encoding resolution, video may be encoded using multiple target bitrates. EX1014, 3-4, 7-8. As discussed above, selecting the target bitrate of each stream from a plurality of bitrates helps to achieve the proper balance between costs and quality. EX1005 at 15-16. Undergoing this same target bitrate selection process for each stream will enhance these benefits and help to achieve cost savings across all streams while ensuring that each of the streams maintains a minimum quality. *See, e.g.*, EX1014, 3-4, 7-8.

124. Liao also discloses this limitation. Liao discloses “[s]ince customers’ Internet service varies widely in bandwidth, different versions of the media program can be generated, each with a different resolution, typically 480p, 360p, 288p or 240p.” EX1006, 7:67-8:6. Each of these alternate resolution versions “should be encoded with different bit rates.” *Id.*, 8:17-22, 8:38-40. For example, in Figure 6, Liao discloses a representation of a TEST source video (annotated below in green) that is encoded at multiple target bitrates: 300 kbps, 400 kbps, 500 kbps, 600 kbps, 700 kbps, and 800 kbps. *Id.*, 8:27-37, Fig. 6 (annotated). Given the above, a POSITA would understand that the TEST source video should be encoded at different resolutions, each with multiple target bitrates. *Id.*, 7:67-8:6, 8:17-22, 8:27-40, Fig. 6.

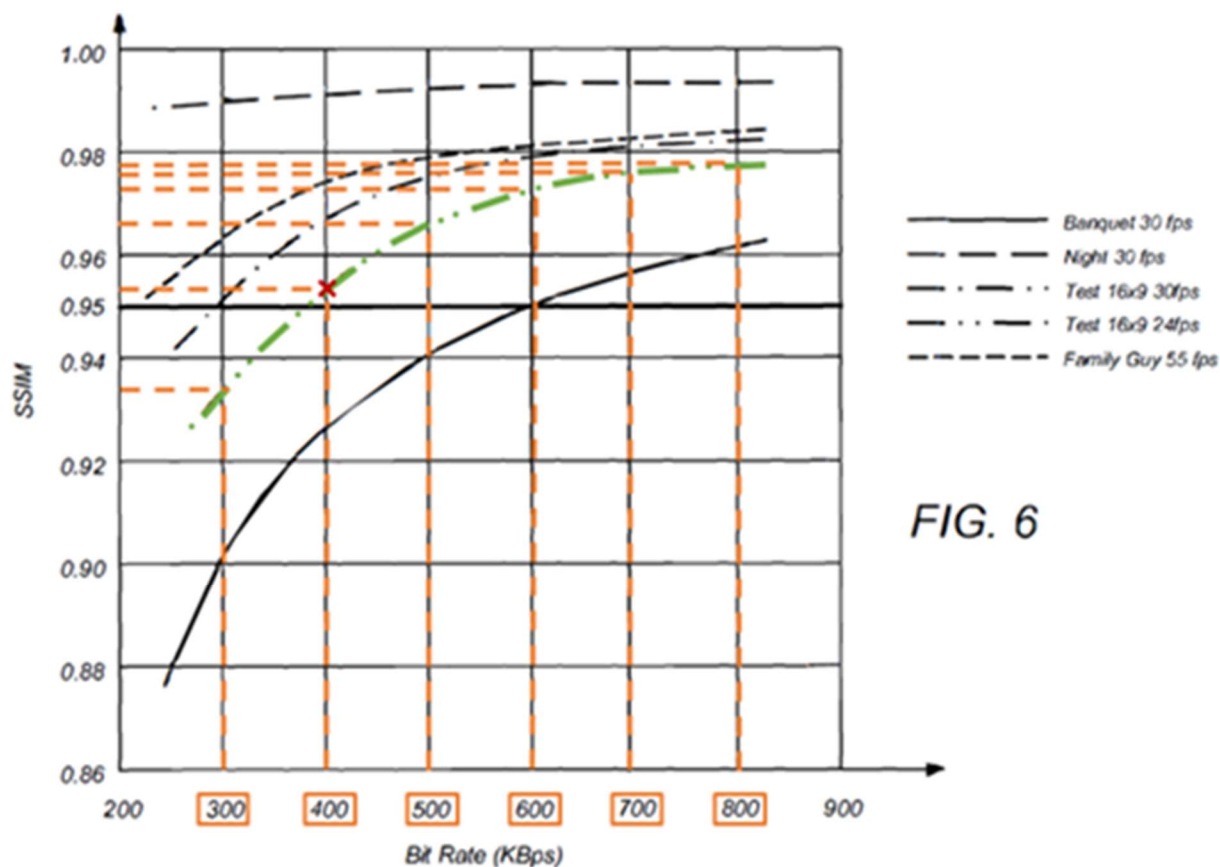


FIG. 6

Id., Fig. 6.

125. Thus, both Ozer and Liao disclose this limitation.

(5) [1.c(ii)] “for each particular resolution of the plurality of resolutions: . . . evaluating quality for each of the multiple encodings”

126. The combination of Ozer in view of Liao and general knowledge of a POSITA, as described below, renders limitation [1.c(ii)] obvious.

127. Ozer and Liao both disclose encoding a source video at different resolutions, each encoded using multiple target bit rates. *Supra* § IX.A.1(4). EX1005, 115-16.

128. Ozer highlights the importance of maintaining quality level for each stream to prevent the unnecessary expenditure of transmission resources. *Id.*, 127-128. Liao teaches using SSIM to determine video quality level by evaluating how similar an encoded stream is to a reference video. EX1006, 8:40-67.

129. SSIM may be used to determine the “relationship between video quality and bitrate for different transcoded media programs.” *Id.*, 8:38-40. Liao discloses “[s]ince customers’ Internet service varies widely in bandwidth, different versions of the media program can be generated, each with a different resolution, typically 480p, 360p, 288p or 240p.” *Id.*, 7:67-8:6. Figure 6 (below) of Liao then teaches that each version of a source video, such as TEST source video, is encoded at multiple target bitrates: 300 kbps, 400 kbps, 500 kbps, 600 kbps, 700 kbps, and 800 kbps. *Id.*, 8:27-37, Fig. 6 (annotated).

130. TEST (green below) encoded at a resolution and a 16x9 aspect ratio with a target bitrate of 400 kbps results in an SSIM between 0.95 and 0.96 (blue dashed line below). *Id.*, 8:1-2, Fig. 6 (annotated). Encoding TEST at a resolution and a 16x9 aspect ratio with a target bitrate of 500 kbps results in an SSIM between 0.96 and 0.97 (purple dashed line below). *Id.* And encoding TEST at a resolution and a 16x9 aspect ratio with a target bitrate of 600 kbps has an SSIM between 0.97 and 0.98 (yellow dashed line below). *Id.*, Fig. 6 (annotated).

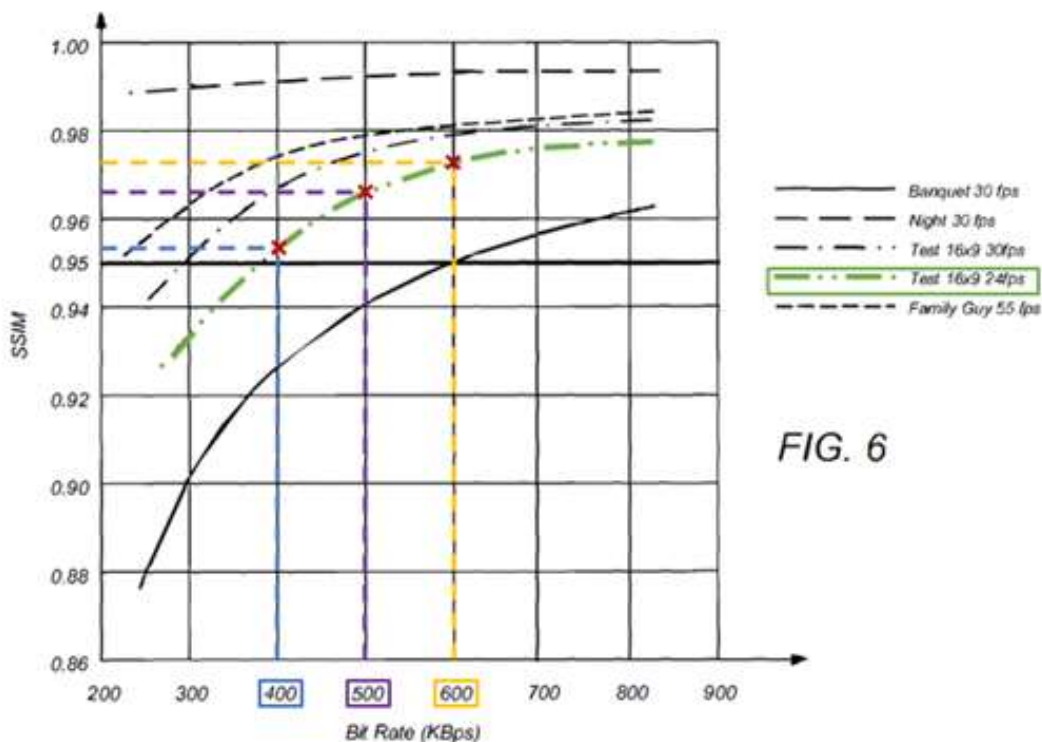


FIG. 6

Id., Fig. 6 (annotated).

131. It would have been obvious to a POSITA to use this quality evaluation for each of the “different versions of the media program . . . each [encoded] with a different resolution, typically 480p, 360p, 288p or 240p.” *Id.*, 7:67-8:6; EX1014, 3-4, 7-8. As explained above, applying the quality evaluation of Liao for a stream makes it possible to identify an optimal bitrate that achieves a pre-specified quality level (e.g., a minimum of 0.95 SSIM) without unnecessarily expending transmission resources on streams with a bitrate that is needlessly high. EX1006, 8:38-9:16; *see generally* EX1005, FIG. 7-2, 127-132, 117-18, 136-37. A POSITA would understand that providing this quality evaluation on each of the “different versions of the media program . . . each [encoded] with a different resolution” would enhance

these benefits by achieving greater resource savings. EX1006, 8:38-9:16; *see generally* EX1005, FIG. 7-2, 127-132, 117-18, 136-37; EX1014, 3-4, 7-8.

132. Thus, the combination of Ozer in view of Liao discloses this limitation.

- (6) [1.d] “selecting a plurality of resolution and target bitrate combinations for a plurality of alternative streams based on the evaluated quality of each of the multiple encodings;”

133. The combination of Ozer in view of Liao and general knowledge of a POSITA, as described below, renders limitation [1.d] obvious.

134. Ozer and Liao teach “a plurality of alternative streams” that are each encoded at a particular target bitrate and resolution, and “evaluating [the] quality for each of the multiple encodings.” *Supra* §§ IX.A.1(4), -(5). The result of the quality evaluation of Liao is “selecting a plurality of resolution and target bitrate combinations” for the “plurality of alternative streams” “based on the evaluated quality of each of the multiple encodings.”

135. Liao teaches selecting combinations of resolution and bitrate by selecting encodings having a desired a minimum target SSIM (quality evaluation metric) of 0.95. EX1006, 9:7-16. For example, Liao discloses encoding a TEST source video at multiple target bitrates: 300 kbps, 400 kbps, 500 kbps, 600 kbps, 700 kbps, and 800 kbps. *Id.*, 8:1-3, 8:27-37, Fig. 6. Liao teaches selecting the specific combinations of resolution and bitrate from these encodings by selecting the lowest bitrate encoding having a minimum SSIM (a minimum quality) of at least 0.95,

which in this case will result in the selection of a 400 kbps stream. *Id.*, 9:7-16, Fig. 6. A stream encoded at 300 kbps has an SSIM that is lower than the minimum of 0.95. *Id.* Meanwhile, streams encoded at 500 kbps, 600 kbps, 700 kbps, and 800 kbps all have SSIMs that are higher than 0.95, but require more resources to transmit with little quality return. *Id.*, Fig. 6. Accordingly, a POSITA would have known to encode a stream at 400 kbps to achieve the optimum cost/quality trade-off. *Id.*, 9:7-16.

136. As discussed above, a POSITA would have known to select a target bitrate based on SSIM's quality evaluation for each of the "different versions of a media program each [having] a different resolution, typically 480p, 360p, 288p or 240p" is obvious. *Id.* 8:1-9, Fig. 6; EX1014, 3-4, 7-8. Performing such an evaluation ensures that the lowest bitrate stream of a desired quality level is selected for each resolution so that transmission resources are not wasted on higher bitrate streams. *Id.*, 8:38-9:16; *see generally* EX1005, Fig. 7-2, 127-132, 117-18, 136-37; EX1014, 3-4, 7-8. A POSITA would further understand that multiple selections using SSIM may also be made for each individual resolution if multiple quality streams are desired (e.g. a high quality, medium quality, and low quality stream) to facilitate stream switching during adaptive bitrate streaming. EX1006, 8:38-9:16; *see generally* EX1005, Fig. 7-2, 127-132, 117-18, 136-37; EX1014, 3-4, 7-8.

137. Accordingly, the combination of Ozer and Liao teaches selecting a plurality of resolution (one of 480p, 360p, 288p or 240p) and bitrate (over 300 kbps and over 400 kbps) combinations for a plurality of alternative streams based on the evaluated quality (SSIM greater than 0.95) of each of the multiple encodings.

- (7) [1.e] “uploading encodings of a plurality of sections of the source video content at each of the selected plurality of resolution and target bitrate combinations for the plurality of alternative streams to a set of servers that form part of a content distribution system;”

138. The combination of Ozer in view of Liao and general knowledge of a POSITA, as described below, renders limitation [1.e] obvious.

139. As noted above, Ozer and Liao teach “select[ing a] plurality of resolution and target bitrate combinations for the plurality of alternative streams.” *Supra* § IX.A.1(6). Liao discloses “encod[ing] the media program to a specific target video quality” based on the selection. EX1006, 9:17-22.

140. Ozer discloses uploading the “plurality of alternative streams” created during encoding to a content distribution network/system (CDN). EX1005, 115-16, 128. This is shown explicitly in Figure 7-1 of Ozer, where the source video content (annotated in red) is encoded into a number of different streams (annotated in green) and uploaded to a content distribution network (“CDN”) (annotated in yellow). *Id.*, 115-16. The source video content is split into a plurality of chunks (sections), each representing 2 to 10 seconds of video. *Id.*, 116-17, 136.

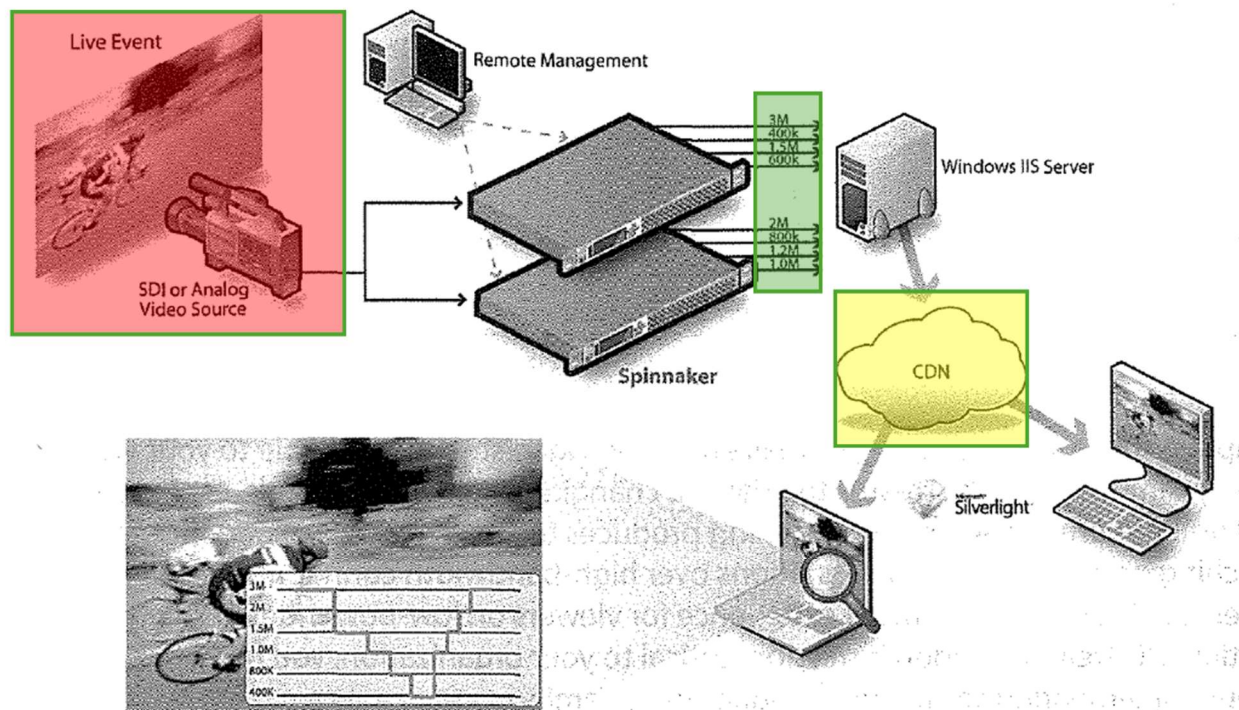


Figure 7-1. Using Inlet Technologies' Spinnaker encoder to produce multiple streams for a Microsoft Smooth Streaming event. Usage courtesy Inlet Technologies.

Id., Fig. 7-1 (annotated).

141. A CDN is formed using a set of servers. *Id.*, 136-37. For example, in Figure 7-9, Ozer illustrates using the Akamai HTTP Network as a CDN, and the Akamai HTTP Network is composed of a plurality of servers (green shaded circles), as shown in Figure 7-9 below. *Id.*, 136-37. A POSITA would have understood at the time of invention that a CDN is a content distribution system.

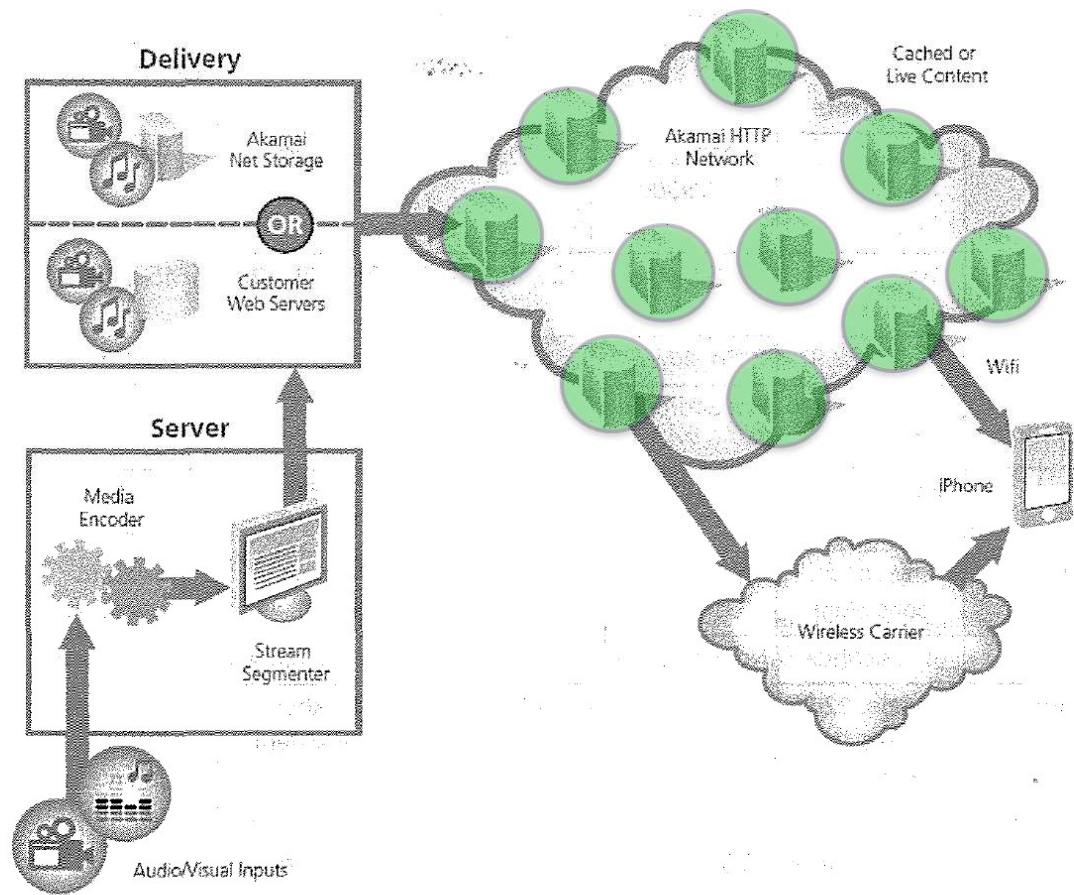


Figure 7-9. Overview of Apple's HTTP Live Streaming with Akamai as the CDN. Image courtesy Akamai.

142. Accordingly, Ozer in view of Liao discloses “uploading encodings of a plurality of sections” (chunks) at “each of the selected plurality of resolution and target bitrate combinations for the plurality of alternative streams to a set of servers that form part of a content distribution system” (i.e., CDN).

- (8) [1.f] “generating entries in a top level index describing at least one of the plurality of alternative streams;”

143. The combination of Ozer in view of Liao and general knowledge of a POSITA, as described below, renders limitation [1.f] obvious.

144. Ozer and Liao teach uploading plurality of alternative streams, each having a different combination of bitrate and resolution, to a CDN. *Supra* § IX.A.1(7). To keep track of the video encodings uploaded to the CDN, Ozer teaches generating entries in a top level index that “identifies the available files and their locations on the [CDN] and refers to specific chunks of each stream.” EX1005, 117-18, 136-37. This top level index is stored in and distributed from the CDN. *Id.*, 136. Ozer discloses, as an example top level index, a file titled “Index File.m3u8.” *Id.*, 117-18. This file is the top-level index file typically associated with Apple’s HTTP Live Streaming. *Id.*, 117-18. Figure 7-2 of Ozer depicts an “Index File.m3u8,” and describes three alternative streams: a “Low” level stream, a “Mid” level stream, and a “Hi[gh]” level stream (the descriptors of each stream constituting entries in the top level index). *Id.*, Fig. 7-2, 117-18, 136-37. Ozer explains that “[i]n an on-demand presentation, the manifest file is static, while in a live event, the manifest file is continually updated with the identity and location of all streams and chunks of streams.” *Id.*, 118. The manifest file is the “top-level index” (Index File.m3u8).

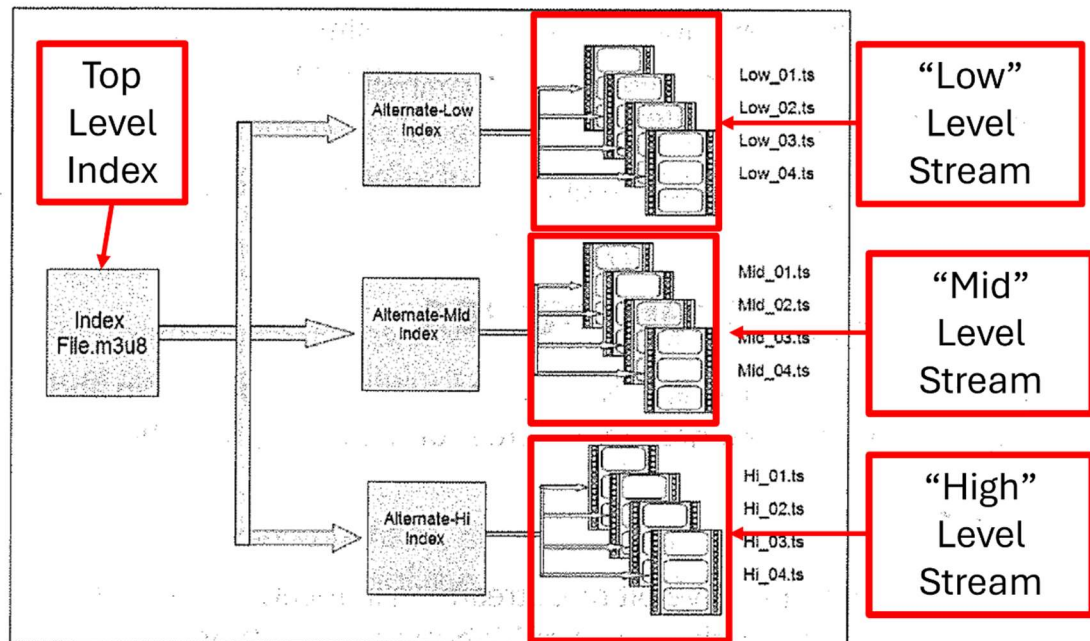


Figure 7-2. The manifest file identifies the different files and their locations. From an article that I wrote for Millimeter Magazine, and usage is courtesy Millimeter.

Id., Fig. 7-2 (annotated).

145. Accordingly, Ozer in view of Liao disclose “generating entries in a top level index describing at least one of the plurality of alternative streams.”

- (9) [1.g] “providing the generated top level index from the content distribution system to a one or more playback devices capable of performing adaptive streaming using the encodings of the plurality of sections; and”

146. The combination of Ozer in view of Liao and general knowledge of a POSITA, as described below, renders limitation [1.g] obvious.

147. As was noted above, Ozer and Liao teach generating a “top level index” that describes the plurality of alternative streams and “uploading encodings of a

plurality of sections of the source video content” to a content distribution network/system. *Supra* §§ IX.A.1(7), -(8).

148. “Adaptive streaming technologies make multiple video streams available to your viewers and dynamically switch streams to adapt to changing connection speeds, CPU use statistics and other heuristics.” EX1005, 115. For a viewer’s device to switch between these streams they must “check the manifest file for the next chunk.” *Id.*, 118. In other words, the viewer’s device must use the top level index to adaptively stream chunks (plurality of sections) of video content. *Id.*, 117-18, 136. The top level index is provided to the viewer’s device, which the device later uses to perform adaptive streaming. *See Id.*, 117-18, 136. For example, the viewer’s device “might start out playing the lowest-quality stream (Low_01 .ts), but if it determines [sic] that it can handle a higher-quality stream, it would check the manifest file for the next chunk of a higher-quality stream (say, Mid_02.ts) and retrieve and play that chunk.” *Id.*, 118.

149. Accordingly, Ozer in view of Liao discloses “providing the generated top level index from the content distribution system to a one or more playback devices capable of performing adaptive streaming using the encodings of the plurality of sections.”

- (10) [1.h] “responding to requests from the one or more playback devices, where the requests include

requests for specific encodings from the encodings of the plurality of sections.”

150. The combination of Ozer in view of Liao and general knowledge of a POSITA, as described below, renders limitation [1.h] obvious.

151. As noted above, Ozer in view Liao discloses “providing the generated top level index” to the viewer to “perform[]adaptive streaming using the encodings of the plurality of sections.” *Supra* § IX.A.1(9). In response “to changing connection speeds, CPU use statistics and other heuristics” (e.g., changes in buffer conditions, client bandwidth, CPU use, and playback window size) a viewer’s device (playback device) will request chunks of specific encodings (the alternative streams) having an encoding resolution and target bitrate. EX1005, 115, 117-18.

152. Ozer provides several examples. For instance, Ozer discloses that a playback device “might start out playing the lowest quality stream (Low_01.ts), but if it determines [sic] that it can handle a higher-quality stream, it would check the manifest file for the next chunk of a higher-quality stream (say, Mid_02.ts) and retrieve and play that chunk.” *Id.*, 118. “[R]etrieval” of the requested “next chunk” is a response to the playback device’s request for a specific encoding. *Id.*

153. Accordingly, Ozer in view of Liao teaches “responding to requests from the one or more playback devices, where the requests include requests for specific encodings from the encodings of the plurality of sections.”

2. Claim 2: “The method of claim 1 further comprising assembling the encodings of the plurality of sections for a given alternative stream into a video container file.”

154. In addition to teaching all the limitations of claims 1 and 11, Ozer in view of Liao and general knowledge of a POSITA, discloses claims 2 and 12.

155. Ozer in view of Liao discloses encoding a number of sections (chunks) of alternative encodings or streams. *Supra* § IX.A.1(4); EX1005, 118 (“All HTTP-based adaptive technologies split their [container] files into 2 to 10-seconds chunks, which are the video chunks on the right in []Figure 7-2 (Hi_01.ts, Hi_02.ts, etc.)”). Each of these chunks (plurality of sections) is assembled into video container files. EX1005, 117-118, 120.

156. Specifically, a POSITA would understand that Figure 7-2 of Ozer shows container files for three alternate streams: a “Hi[gh]” stream (blue shaded box below), “Mid” stream (red shaded box below), and “Low” stream (yellow shaded box below). *Id.*, Fig. 7-2 (annotated), 120.

157. The container file for the “Hi[gh]” stream includes four assembled chunks: Hi_01.ts, Hi_02.ts, Hi_03.ts, and Hi_04.ts. EX1005, 117, Fig. 7-2. Each of the container files for “Low” and “Mid” streams include four assembled chunks. *Id.*

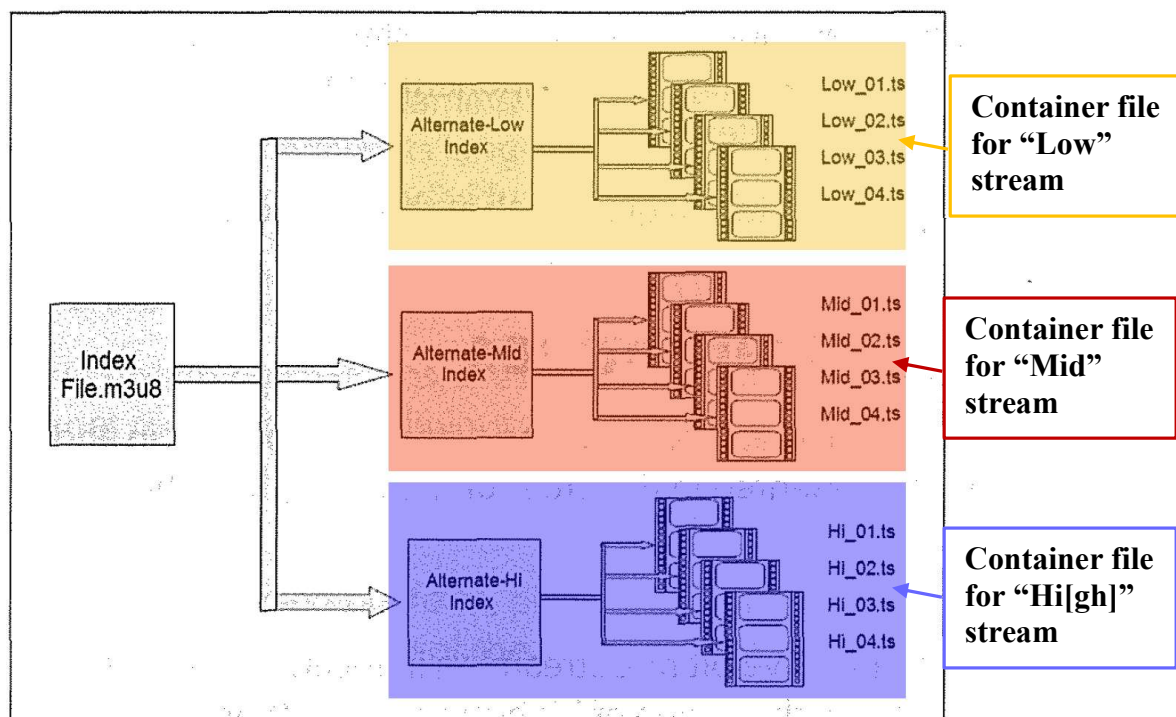


Figure 7-2. The manifest file identifies the different files and their locations. From an article that I wrote for Millimeter Magazine, and usage is courtesy Millimeter.

Id., 117, Fig. 7-2 (annotated).

158. Accordingly, Ozer in view of Liao teaches “assembling the encodings of the plurality of sections for a given alternative stream into a video container file.”

3. Claim 3: “The method of claim 2 further comprising generating entries in a container index within each container file indicating sizes of portions of the given alternative stream within the video container file.”

159. In addition to teaching all the limitations of claims 1 and 2 and general knowledge of a POSITA, Ozer in view of Liao, as described below, renders claim 3 obvious.

160. As explained above, Ozer discloses assembling a plurality of sections (chunks) for a given alternative stream into one or more video container files. *Supra* § IX.A.2. Ozer also teaches that a container index within these container files describes the chunks of an alternative stream. EX1005, 117-18, Fig. 7-2. The size of each chunk is one of the attributes described by this container index. *Id.*, 117-18, Fig. 7-2. As shown below in Figure 7-2, the container files for a “Hi[gh]” stream (blue shaded box), “Mid” stream (red shaded box), and “Low” stream (yellow shaded box), each include a container index, respectively the Alternate-Hi Index, Alternate-Mid Index, and Alternate-Low Index. *Id.*, 117-18, Fig. 7-2 (annotated). These container indexes identify and point to the plurality of segments (chunks) making up their respective alternative streams. *Id.*, 118.

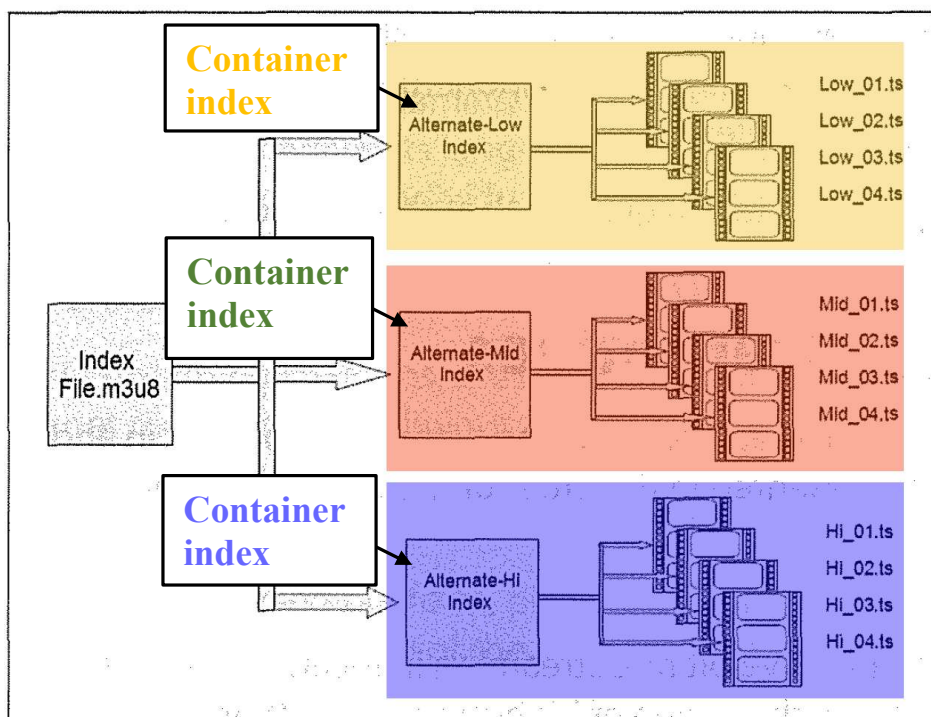


Figure 7-2. The manifest file identifies the different files and their locations. From an article that I wrote for Millimeter Magazine, and usage is courtesy Millimeter.

Id., Fig. 7-2 (annotated)

161. Ozer also teaches that the container index indicates sizes of portions of the given alternative stream within the video container file. Specifically, in the container index, the size of the chunk would be represented in the form of time entries indicating the duration of each chunk. In the plain text file version of the above container index, there exists a tag named “EXT-X-TARGETDURATION.” EX1022, 2, 6-7. This TARGETDURATION tag indicates the maximum duration in seconds of any chunk in the container file. *Id.*, 6-7. Each portion of the alternative stream (one or more chunks) noted in the container index is less than or equal to the target duration tag value. *Id.* The “EXT-X-TARGETDURATION” tag appears at

least once in the container index plain text file. *Id.* The duration of each individual segment is specified using the EXTINF tag, which appears immediately before each chunk URL. *Id.*

162. A POSITA would understand that the noted size of each chunk within the container index is used by the playback device to successfully locate and request a given chunk from each available stream. EX1005, 117-18, Fig. 7-2. This information could help to inform the viewer's device regarding how many segments (chunks) can be loaded into the viewer's playback buffer, or setting the length of the playback buffer. *Id.*, 116. Using the size of the segments (chunks) in this manner was well-known in the art. *See, e.g.*, EX1015, ¶¶175-76.

163. Accordingly, Ozer in view of Liao teaches “[t]he method of claim 2 further comprising generating entries in a container index within each container file indicating sizes of portions of the given alternative stream within the video container file.”

- 4. Claim 4: “The method of claim 1, wherein adaptively streaming the encodings comprises receiving at least one request for an encoded section of one of the plurality of alternative streams from the one or more playback devices based on the generated top level index.”**

164. In addition to teaching all the limitations of claim 1 and general knowledge of a POSITA, Ozer in view of Liao, as described below, renders claim 4 obvious.

165. As noted above, Ozer in view of Liao discloses “providing the generated top level index” to the viewer’s device, “performing adaptive streaming using the encodings of the plurality of sections,” and “responding to request from the one or more playback devices.” *Supra* §§ IX.A.1(9), -(10). Ozer teaches that the top level index is used to make this request. EX1005, 117. Specifically, because the top level index (or manifest file) “identifies the available files and their locations on the HTTP server and refers to specific chunks of each stream,” Ozer discloses that the playback device “might start out playing the lowest quality stream (Low_01.ts), but if it determines [sic] that it can handle a higher-quality stream, it would ***check the manifest file*** for the next chunk of a higher-quality stream (say, Mid_02.ts) and retrieve and play that chunk.” *Id.*, 118 (emphasis added). This “retriev[al]” of the requested “next chunk” based on “check[ing] the manifest file” requires that the CDN receives a request and provides an encoded section (chunk) of the plurality of alternative streams in response to that request. *Id.*

166. Accordingly, Ozer in view of Liao teaches “[t]he method of claim 1, wherein adaptively streaming the encodings comprises receiving at least one request for an encoded section of one of the plurality of alternative streams from the one or more playback devices based on the generated top level index.”

5. Claim 7: “The method of claim 1, wherein both a width and a height of each resolution of the plurality of resolutions have an integer number of pixels.”

167. In addition to teaching all the limitations of claim 1, Ozer in view of Liao and general knowledge of a POSITA, as described below, renders claim 7 obvious.

168. As noted above, Ozer in view of Liao discloses “identifying a plurality of resolutions for encoding of video content.” *Supra* § IX.A.1(3). Ozer discloses that a “[r]esolution is the height and width of the video in pixels.” EX1005, 17. For example, as shown below in Figure 1-2, a “320x240 video has 76,800 pixels in each frame, while a 640x480 video [] has 307,200 pixels, or four times as many.” *Id.*, 17, Fig. 1-2. 320 represents the number of pixels in the width of each resolution, while 240 represents the number of pixels in the height of each resolution. *Id.*

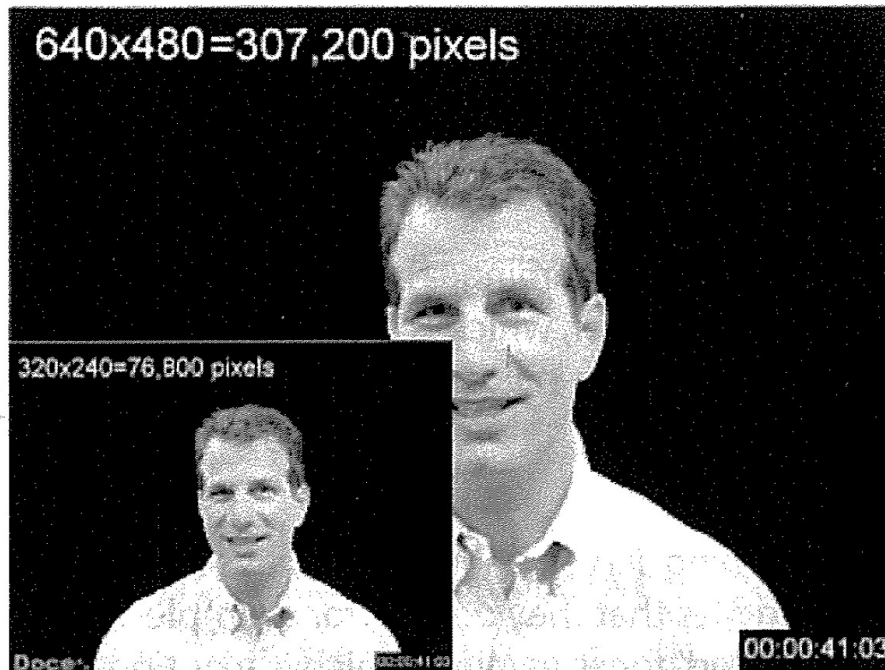


Figure 1-2. Scaling the 640x480 file down to 320x240 reduces the number of pixels by 75%, making the smaller file easier to compress.

169. Another example is shown below in Figure 7-4 of Ozer where the identified resolutions (“Picture Size[s]”) are listed using a width and height (256x144, 512x288, 768x432, and 1280x720). *Id.*, Fig. 7-4. Each of 256, 512, 768, and 1280 represents the number of pixels in the width of each resolution, while each of 144, 288, 432, and 720 represents the number of pixels in the height of each resolution. *Id.*

16x9				
Stream #	Picture Size	V	A	AV
1	256x144	150	64	214
2	256x144	250	64	314
3	512x288	450	64	514
4	512x288	600	64	664
5	512x288	800	64	864
6	512x288	1200	64	1264
7	768x432	1400	64	1464
8	1280x720	1700	64	1764
9	1280x720	2500	64	2564
10	1280x720	3500	64	3564

Figure 7-4. Adobe's Levkov recommends clustering your streams around certain frame sizes.

Id., Fig. 7-4.

170. Yet another example is shown below in Table 7-3 of Ozer where the identified resolutions are listed using a width and height (512x288, 768x432, and 1280x720) each having a corresponding number of pixels (147,456, 331,776, 921,600). *Id.*, Table 7-3.

Preset	Resolution	Video bitrate	Profile	Key Frame	Audio bitrate	Channels
Phone and Table, 3G, 16x9	512x288	300 kbps	Baseline	72	48 kbps	Stereo
Phone and Table, Low, 16x9	512x288	450 kbps	Baseline	72	48 kbps	Stereo
Phone and Table, WiFi, 16x9	512x288	650 kbps	Baseline	72	48 kbps	Stereo
PC & TV, SD, Med, 16x9	768x432	1,140 kbps	Main	72	64 kbps	Stereo
PC & TV, SD, High, 16x9	768x432	1,140 kbps	Main	72	64 kbps	Stereo
PC & TV, HD, Low, 16x9	1280x720	2,440 kbps	High	72	64 kbps	Stereo
PC & TV, HD, High, 16x9	1280x720	3,440 kbps	High	72	64 kbps	Stereo

Table 7-3. The configurations of adaptive streaming presets released by Adobe in late 2010.

Id., Table 7-3.

171. Accordingly, Ozer in view of Liao teaches that “both a width and a height of each resolution of the plurality of resolutions have an integer number of pixels.”

6. Claim 8: “The method of claim 1, wherein each resolution of the plurality of resolutions has a same display aspect ratio and at least two resolutions of the plurality of resolutions have different sample aspect ratios.”

172. In addition to teaching all the limitations of claim 1, the combination of Ozer in view of Liao and general knowledge of a POSITA, as described below, renders claim 8 obvious. Each video is encoded and stored at a specific resolution, such as 1280x720, identifying the width and height of each video frame in pixels. *See supra* § IX.A.1(3); *see also* EX1036. Thus, each encoding has a width-to-height ratio (*e.g.*, 16:9 and 4:3). POSITAs were also aware of two other width and height relationships used to display an encoded video: display aspect ratio and sample aspect ratio. *See, e.g.*, EX1001, 8:28-32, 10:8-11. Display aspect ratio is the proportion between the width and height of a device’s display. *See, e.g., id.*, 8:28-32; *see also* EX1036. Sample aspect ratio is the proportion between the width and the height of each individual pixel that makes up a video frame. *See, e.g., id.*, 10:8-11; *see also* EX1036. Discrepancies between a video’s resolution and these ratios, if not addressed, may result in distortion during playback. EX1005, 24-26; *see also* EX1036. A POSITA would have understood that creating video encodings with different combinations of these three characteristics would help to avoid potential

distortions. *See id.* For example, a POSITA would have known to encode a video with one display aspect ratio using different resolutions with different sample aspect ratios. *See id.*

173. Ozer teaches that “each resolution of the plurality of resolutions has a same display aspect ratio.” *Id.*, Table 7-3 (annotated); *see also id.*, Table 7-4 (same). This is shown at least in Table 7-3, where each resolution has a display aspect ratio of 16x9 (shaded yellow box below).

Preset	Resolution	Video bitrate	Profile	Key Frame	Audio bitrate	Channels
Phone and Table, 3G, 16x9	512x288	300 kbps	Baseline	72	48 kbps	Stereo
Phone and Table, Low, 16x9	512x288	450 kbps	Baseline	72	48 kbps	Stereo
Phone and Table, WiFi, 16x9	512x288	650 kbps	Baseline	72	48 kbps	Stereo
PC & TV, SD, Med, 16x9	768x432	1,140 kbps	Main	72	64 kbps	Stereo
PC & TV, SD, High, 16x9	768x432	1,140 kbps	Main	72	64 kbps	Stereo
PC & TV, HD, Low, 16x9	1280x720	2,440 kbps	High	72	64 kbps	Stereo
PC & TV, HD, High, 16x9	1280x720	3,440 kbps	High	72	64 kbps	Stereo

Table 7-3. The configurations of adaptive streaming presets released by Adobe in late 2010.

Id., Table 7-3.

174. Ozer also teaches that “at least two resolutions of the plurality of resolutions have different sample aspect ratios” to account for different types of viewer devices. As Ozer explains, video “designed to be viewed on a television set has a fundamentally different aspect ratio when viewed on a computer.” *Id.*, 24. The “frame is horizontally squeezed by 10% when displayed on a tv set,” resulting in a distorted image, as shown in Figure 1-10 below. *Id.*, 24, Fig. 1-10. Ozer recommends having encodings with different sample aspect ratios to account for

different types of client devices to avoid this distortion. *Id.*, 25, 129, Table 5-13. This is because different types of client devices may have different pixel dimensions and resolutions that may not match the dimensions of the encoded video. *See id.*, 24-26. To scale the video for these different displays and avoid distortion, Ozer recommends having encodings with different sample aspect ratios to accommodate different types of client devices (e.g., tvs and computers) with the same display aspect ratios. *Id.*, 25, 129, Table 5-13.



Figure 1-10. SD video displays differently on computers and TV sets.

Id., Fig. 1-10.

175. Ozer provides another reason for using different sample aspect ratios and the same display aspect ratio: the different width-to-height ratios of stored encodings. Table 5-13 of Ozer discloses alternate encodings at multiple resolutions having two different width-to-height ratios of 16:9 and 4:3. *Id.*, Table 5-13. To account for these differences in width-to-height ratios, the sample aspect ratio is adjusted to avoid distortion. *Id.* For example, displaying the 16:9 encodings in Table 7-3 on a device having a display aspect ratio of 16x9, requires a sample aspect ratio

of 1:1. *Id.*, 24-25. Displaying the 4:3 encodings in Table 7-3 on the same device, requires a sample aspect ratio of 4:3 to avoid distortions. *Id.* As such, a POSITA would have understood that using a different sample aspect ratio would allow playback of alternate encodings with different width-to-height ratios on devices with the same display aspect ratio without distortion.

30 fps		Low Motion		High Motion	
Width	Height	Data Rate (kbps)	Bits per Pixel	Data Rate (kbps)	Bits per Pixel
16:9					
320	180	244,376	0.14	366,564	0.21
480	270	448,948	0.12	673,421	0.17
640	360	691,200	0.10	1,036,800	0.15
853	480	1,063,860	0.09	1,595,790	0.13
1280	720	1,955,009	0.07	2,932,513	0.11
1920	1080	3,591,581	0.06	5,387,371	0.09
4:3					
320	240	325,835	0.14	488,752	0.21
400	300	455,368	0.13	683,052	0.19
480	360	598,597	0.12	897,895	0.17
640	480	921,600	0.10	1,382,400	0.15

Table 5-13. Recommended data rates and bits-per-pixel values for video produced at 30 fps.

Id., Table 5-13 (annotated).

176. Accordingly, Ozer and Liao teach that “each resolution of the plurality of resolutions has a same display aspect ratio and at least two resolutions of the plurality of resolutions have different sample aspect ratios.”

7. Claim 9: “The method of claim 1, wherein the target bitrate for each resolution is determined using a formula.”

177. In addition to teaching all the limitations of claim 1, Ozer in view of Liao and general knowledge of a POSITA, as described below, renders claim 9 obvious.

178. Ozer discloses a “Power of .75 rule,” which is a formula used to determine the target bitrate for each resolution. EX1005, 94; *see also generally* EX1033. As noted in Ozer: “[a]ccording to the Power of .75 rule, for lower-resolution videos, the bits-per pixel value should increase, while it should decrease for higher-resolution videos.” EX1005, 94; *see also generally* EX1033. For example, based on a “target bitrate” of 691.2 kbps (shaded in yellow below) at a resolution of 640x360, a POSITA would know to calculate the target bitrate for a resolution of 1280x720 using the following formula:

$$\text{target bitrate} = \frac{(1280 \times 720)^{.75}}{(640 \times 360)} \times 691.2 \text{ kbps}$$

EX1005, 94-95; *see also* EX1033 (describing the Power of .75 rule (formula) used to create Table 5-13 of Ozer).

179. The resulting target bitrate is 1,955.009 kbps, as shown in Table 5-13 of Ozer, below (shaded in green below). EX1005, 95. This formula is used to

determine the bitrate at each of the resolutions shown in Table 5-13. *Id.*, 94-95; *see also* EX1033.

30 fps		Low Motion		High Motion	
Width	Height	Data Rate (kbps)	Bits per Pixel	Data Rate (kbps)	Bits per Pixel
16:9					
320	180	244,376	0.14	366,564	0.21
480	270	448,948	0.12	673,421	0.17
640	360	691,200	0.10	1,036,800	0.15
853	480	1,063,860	0.09	1,595,790	0.13
1280	720	1,955,009	0.07	2,932,513	0.11
1920	1080	3,591,581	0.06	5,387,371	0.09
4:3					
320	240	325,835	0.14	488,752	0.21
400	300	455,368	0.13	683,052	0.19
480	360	598,597	0.12	897,895	0.17
640	480	921,600	0.10	1,382,400	0.15

Table 5-13. Recommended data rates and bits-per-pixel values for video produced at 30 fps.

EX1005, Table 5-13 (annotated).

180. Accordingly, Ozer in view of Liao teach that “the target bitrate for each resolution is determined using a formula.”

8. Claim 10: “The method of claim 1, wherein the target bitrate of a given resolution is determined using experimentation.”

181. In addition to teaching all the limitations of claim 1, Ozer in view of Liao and general knowledge of a POSITA, as described below, renders claim 10 obvious.

182. As discussed above, Ozer in view of Liao teaches iteratively encoding source video content, evaluating the quality of these encodings using SSIM, then

selecting resolution-target bitrate combinations for the plurality of alternative streams based on the SSIM (evaluated quality of each of the encodings). *Supra* IX.A.1. A POSITA would have understood this iterative process, involving evaluation and selection, is a form of experimentation. Thus, Ozer in view of Liao teaches this claim.

9. Claim 11

- (1) [11.pre] “A non-transitory machine readable medium containing processor instructions for encoding source content as a plurality of alternative video streams and adaptively streaming the plurality of alternative video streams, each having a bitrate, where execution of the instructions by a processor causes the processor to perform a process that comprises:”

183. To the extent the preamble of claim 11 is considered limiting, Ozer in view of Liao and general knowledge of a POSITA, discloses it. Ozer and Liao teach a processor for encoding source content and adaptively stream the plurality of streams. For example, Ozer discloses multi-thread encoding using an exemplary “3.33 GHz Xeon processor[.]” EX1005, 173-176. Further, Liao discloses encoders with a “special purpose processor.” EX1006, 8:27-37. A POSITA would have also understood that the servers described in Ozer’s Figure 7-9 as part of the Akamai CDN also include processors. *See, e.g.*, EX1005, Fig 7-9. The preamble of claim 1 is otherwise substantively identical to the preamble of claim 11. Therefore, I incorporate by reference my analysis of the preamble of claim 1, as my analysis of

the preamble of claim 1 also applies to the preamble of claim 11. The combination of Ozer in view of Liao renders the preamble of claim 11 obvious for the same reasons as discussed regarding the preamble of claim 1. *See supra* § IX.A.1(1).

(2) [11.a] “identifying source video content;”

184. Ozer in view of Liao and general knowledge of a POSITA discloses limitation [11.a]. Limitation [11.a] is substantively identical to limitation [1.a]. Therefore, I incorporate by reference my analysis of limitation [1.a], as my analysis of limitation [1.a] also applies to limitation [11.a]. The combination of Ozer in view of Liao renders limitation [11.a] obvious for the same reasons as discussed regarding limitation [1.a]. *See supra* § IX.A.1(2).

(3) [11.b] “identifying a plurality of resolutions for encoding of video content;”

185. Ozer in view of Liao and general knowledge of a POSITA discloses limitation [11.b]. Limitation [11.b] is substantively identical to limitation [1.b]. Therefore, I incorporate by reference my analysis of limitation [1.b], as my analysis of limitation [1.b] also applies to limitation [11.b]. The combination of Ozer in view of Liao renders limitation [11.b] obvious for the same reasons as discussed regarding limitation [1.b]. *See supra* § IX.A.1(3).

(4) [11.c(i)] “for each particular resolution of the plurality of resolutions: encoding at least a portion of the source video content multiple times using

the particular resolution and multiple different target bitrates for the particular resolution; and”

186. Ozer in view of Liao and general knowledge of a POSITA discloses limitation [11.c(i)]. Limitation [11.c(i)] is substantively identical to limitation [1.c(i)]. Therefore, I incorporate by reference my analysis of limitation [1.c(i)], as my analysis of limitation [1.c(i)] also applies to limitation [11.c(i)]. The combination of Ozer in view of Liao renders limitation [1c(i)] obvious for the same reasons as discussed regarding limitation [11.c(i)]. *See supra* § IX.A.1(4).

- (5) [11.c(ii)] “for each particular resolution of the plurality of resolutions: . . . evaluating quality for each of the multiple encodings;”

187. Ozer in view of Liao and general knowledge of a POSITA discloses limitation [11.c(ii)]. Limitation [11.c(ii)] is substantively identical to limitation [1.c(ii)]. Therefore, I incorporate by reference my analysis of limitation [1.c(ii)], as my analysis of limitation [1.c(ii)] also applies to limitation [11.c(ii)]. The combination of Ozer in view of Liao renders limitation [11.c(ii)] obvious for the same reasons as discussed regarding limitation [11.c(ii)]. *See supra* § IX.A.1(5).

- (6) [11.d] “selecting a plurality of resolution and target bitrate combinations for a plurality of alternative streams based on the evaluated quality of each of the multiple encodings;”

188. Ozer in view of Liao and general knowledge of a POSITA discloses limitation [11.d]. Limitation [11.d] is substantively identical to limitation [1.d]. Therefore, I incorporate by reference my analysis of limitation [1.d], as my analysis

of limitation [1.d] also applies to limitation [11.d]. The combination of Ozer in view of Liao renders limitation [11.d] obvious for the same reasons as discussed regarding limitation [1.d]. *See supra* § IX.A.1(6).

- (7) [11.e] “uploading encodings of a plurality of sections of the source video content at each of the selected plurality of resolution and target bitrate combinations for the plurality of alternative streams to a set of servers that form part of a content distribution system;”

189. Ozer in view of Liao and general knowledge of a POSITA discloses limitation [11.e]. Limitation [11.e] is substantively identical to limitation [1.e]. Therefore, I incorporate by reference my analysis of limitation [1.e], as my analysis of limitation [1.e] also applies to limitation [11.e]. The combination of Ozer in view of Liao renders limitation [11.e] obvious for the same reasons as discussed regarding limitation [1.a]. *See supra* § IX.A.1(7).

- (8) [11.f] “generating entries in a top level index describing at least one of the plurality of alternative streams;”

190. Ozer in view of Liao and general knowledge of a POSITA discloses limitation [11.f]. Limitation [11.f] is substantively identical to limitation [1.f]. Therefore, I incorporate by reference my analysis of limitation [1.f], as my analysis of limitation [1.f] also applies to limitation [11.f]. The combination of Ozer in view of Liao renders limitation [11.f] obvious for the same reasons as discussed regarding limitation [1.f]. *See supra* § IX.A.1(8).

- (9) [11.g] “providing the generated top level index from the content distribution system to one or more playback devices capable of performing adaptive streaming using the encodings of the plurality of sections; and”

191. Ozer in view of Liao and general knowledge of a POSITA discloses limitation [11.g]. Limitation [11.g] is substantively identical to limitation [1.g]. Therefore, I incorporate by reference my analysis of limitation [1.g], as my analysis of limitation [1.g] also applies to limitation [11.g]. The combination of Ozer in view of Liao renders limitation [11.g] obvious for the same reasons as discussed regarding limitation [1.g]. *See supra* § IX.A.1(9).

- (10) [11.h] “responding to requests from the one or more playback devices, where the requests include requests for specific encodings from the encodings of the plurality of sections.”

192. Ozer in view of Liao and general knowledge of a POSITA discloses limitation [11.h]. Limitation [11.h] is substantively identical to limitation [1.h]. Therefore, I incorporate by reference my analysis of limitation [1.h], as my analysis of limitation [1.h] also applies to limitation [11.h]. The combination of Ozer in view of Liao renders limitation [11.h] obvious for the same reasons as discussed regarding limitation [1.h]. *See supra* § IX.A.1(10).

10. Claim 12: “The non-transitory machine readable medium of claim 11, wherein the process further comprises assembling the encodings of the plurality of sections for a given alternative stream into a video container file.”

193. In addition to disclosing each of the limitations of Claim 11, Ozer in view of Liao and general knowledge of a POSITA discloses claim 12. Claim 12 is substantively identical to Claim 2. Therefore, I incorporate by reference my analysis of Claim 2, as my analysis of Claim 2 also applies to Claim 12. The combination of Ozer in view of Liao renders Claim 12 obvious for the same reasons as discussed regarding Claim 2. *See supra* § IX.A.2.

11. Claim 13: “The non-transitory machine readable medium of claim 12, wherein the process further comprises generating entries in a container index within each container file indicating sizes of portions of the given alternative stream within the video container file.”

194. In addition to disclosing each of the limitations of Claims 11 and 12, Ozer in view of Liao and general knowledge of a POSITA discloses claim 13. Claim 13 is substantively identical to Claim 3. Therefore, I incorporate by reference my analysis of Claim 3, as my analysis of Claim 3 also applies to Claim 13. The combination of Ozer in view of Liao renders Claim 13 obvious for the same reasons as discussed regarding Claim 3. *See supra* § IX.A.3.

- 12. Claim 14: “The non-transitory machine readable medium of claim 11, wherein adaptively streaming the encodings comprises receiving at least one request for an encoded section of one of the plurality of alternative streams from the one or more playback devices based on the generated top level index.”**

195. In addition to disclosing each of the limitations of Claim 11, Ozer in view of Liao and general knowledge of a POSITA discloses claim 14. Claim 14 is substantively identical to Claim 4. Therefore, I incorporate by reference my analysis of Claim 4, as my analysis of Claim 4 also applies to Claim 14. The combination of Ozer in view of Liao renders Claim 14 obvious for the same reasons as discussed regarding Claim 4. *See supra* § IX.A.5.

- 13. Claim 17: “The non-transitory machine readable medium of claim 11, wherein both a width and a height of each resolution of the plurality of resolutions have an integer number of pixels.”**

196. In addition to disclosing each of the limitations of Claim 11, Ozer in view of Liao and general knowledge of a POSITA discloses claim 17. Claim 17 is substantively identical to Claim 7. Therefore, I incorporate by reference my analysis of Claim 7, as my analysis of Claim 7 also applies to Claim 17. The combination of Ozer in view of Liao renders Claim 17 obvious for the same reasons as discussed regarding Claim 7. *See supra* § IX.A.5.

14. Claim 18: “The non-transitory machine readable medium of claim 11, wherein each resolution of the plurality of resolutions has a same display aspect ratio and at least two resolutions of the plurality of resolutions have different sample aspect ratios.”

197. In addition to disclosing each of the limitations of Claim 11, Ozer in view of Liao and general knowledge of a POSITA discloses claim 18. Claim 18 is substantively identical to Claim 8. Therefore, I incorporate by reference my analysis of Claim 8, as my analysis of Claim 8 also applies to Claim 18. The combination of Ozer in view of Liao renders Claim 18 obvious for the same reasons as discussed regarding Claim 8. *See supra* § IX.A.6.

15. Claim 19: “The non-transitory machine readable medium of claim 11, wherein the target bitrate for each resolution is determined using a formula.”

198. In addition to disclosing each of the limitations of Claim 11, Ozer in view of Liao and general knowledge of a POSITA discloses claim 19. Claim 19 is substantively identical to Claim 9. Therefore, I incorporate by reference my analysis of Claim 9, as my analysis of Claim 9 also applies to Claim 19. The combination of Ozer in view of Liao renders Claim 19 obvious for the same reasons as discussed regarding Claim 9. *See supra* § IX.A.7.

B. Ground 2: Claims 6 and 16 of 785 Patent are Rendered Obvious by Ronca in view of Liao and Gu

- 1. Claim 6: “The method of claim 1, wherein identifying a plurality of resolutions comprises identifying resolutions with target bitrates within a predetermined percentage of a bitrate associated with an alternative stream of the plurality of alternative streams.”**

199. As noted above, Ozer in view of Liao teach claim 1. The combination of Ozer, Liao, and Gu and general knowledge of a POSITA teaches Claim 6.

200. Ozer in view of Liao teaches identifying a plurality of resolutions, each having a plurality of target bitrates, encoding source video content at each resolution and their respective target bit rates, and selecting a plurality of alternative streams based on the evaluated quality of the resulting encodings. *See supra* §§ IX.A.1, -11.

201. Ozer describes a need to find reasonable “gap[s]” between the target bitrates of corresponding streams, each having a particular resolution. EX1005, 130-131 (“Keeping too many bit rates close to one another could result in too many stream switches,” even with smaller bandwidth fluctuations during a streaming session, while “encoding too few streams with a huge gap in bit rates would not provide the best quality or the most optimal stream for a particular bandwidth environment.”) Gu explains how to select appropriate target bitrates that have a reasonable bitrate gap between alternate streams. Specifically, Gu teaches “decreas[ing] progressively in a proportional, logarithmic or other decreasing manner” the target bit rates of alternative streams. EX1007, 7:58-61. One specific

example of this decrease that Gu provides is decreasing the target bitrate of each alternative stream by a predetermined proportion of 25% (or ¼) from the preceding stream (e.g. a first stream with the highest bitrate, a second stream with a bitrate 75% of the first bitrate, a third stream with a bitrate 50% of the first bitrate). *Id.*, 7:54-67.

202. Accordingly, a POSITA would have understood that the combination of Ozer, Liao, and Gu teaches “identifying a plurality of resolutions comprises identifying resolutions with target bitrates within a predetermined percentage of a bitrate associated with an alternative stream of the plurality of alternative streams.”

2. Claim 16: “The non-transitory machine readable medium of claim 11, wherein identifying a plurality of resolutions comprises identifying resolutions with target bitrates within a predetermined percentage of a bitrate associated with an alternative stream of the plurality of alternative streams.”

203. The combination of Ozer, Liao, and Gu and general knowledge of a POSITA teaches Claim 16.

204. Claim 16 is substantively identical to Claim 6. Therefore, I incorporate by reference my analysis of Claim 6, as my analysis of Claim 6 also applies to Claim 16. The combination of Ozer in view of Liao renders Claim 16 obvious for the same reasons as discussed regarding Claim 6. *See supra* § IX.B.1.

C. Ground 3: Claims 1-4, 7, 9-14, 17, and 19 of 785 Patent are Rendered Obvious by Ronca in view of Liao and Gu

1. Claim 1

- (1) [1.pre] “A method for encoding source content as a plurality of alternative video streams and

adaptively streaming the plurality of alternative video streams, each having a bitrate, the method comprising:”

205. To the extent the preamble is limiting, Ronca in view of Liao and general knowledge of a POSITA teaches the preamble.

206. Ronca discloses “generate[s] multiple video . . . encodings” or video streams of the “same media presentation, each encoded at different bit rates” using an encoder. EX1008, ¶ 38. This allows a client device to adaptively select between these “different alternate streams containing the same material encoded at a variety of data rates, allowing the streaming session to adapt to an available data rate” of the device’s network. *Id.*, ¶ 24; *see also id.*, Fig. 9.

207. Accordingly, Ronca in view of Liao discloses “A method for encoding source content as a plurality of alternative video streams and adaptively streaming the plurality of alternative video streams, each having a bitrate.”

(2) [1.a] “identifying source video content;”

208. Ronca in view of Liao and general knowledge of a POSITA teaches claim limitation [1.a].

209. Ronca explains that “[d]igital content distribution systems typically include a content server, a content player, and a communications network connecting the content server to the content player,” where “[t]he content server is configured to store files (or ‘streams’) available for download from the content server to the

content player.” EX1008, ¶ 4. Ronca discloses that each of these streams originates from source video content, such as “a movie, a television program, a sporting event, a staged or live event captured by recorded video, etc.” *Id.* This source video content is formatted as an “elementary video stream file” later processed by an encoder. *Id.*, ¶¶ 4, 38.

210. Accordingly, Ronca in view of Liao discloses “identifying source video content.”

(3) [1.b] “identifying a plurality of resolutions for encoding of video content;”

211. Ronca in view of Liao and general knowledge of a POSITA teaches claim limitation [1.b].

212. Both Ronca and Liao describe preparing multiple encodings of the same source video content for use in adaptive bitrate streaming. EX1008, ¶¶ 24, 38; EX1006, 7:67-8:6. Specifically, Ronca discloses “generat[ing] multiple video . . . encodings of the same media presentation, each encoded at different bit rates,” using an encoder. EX1008 ¶¶ 24, 38. This allows a client device to adaptively select between these “different alternate streams containing the same material encoded at a variety of data rates, allowing the streaming session to adapt to an available data rate” of the device’s network. *Id.*, ¶ 24; *see also id.*, Fig. 9.

213. Liao teaches that in addition to encoding video at “different bit rates,” source video should also be encoded with “different resolution[s], typically 480p,

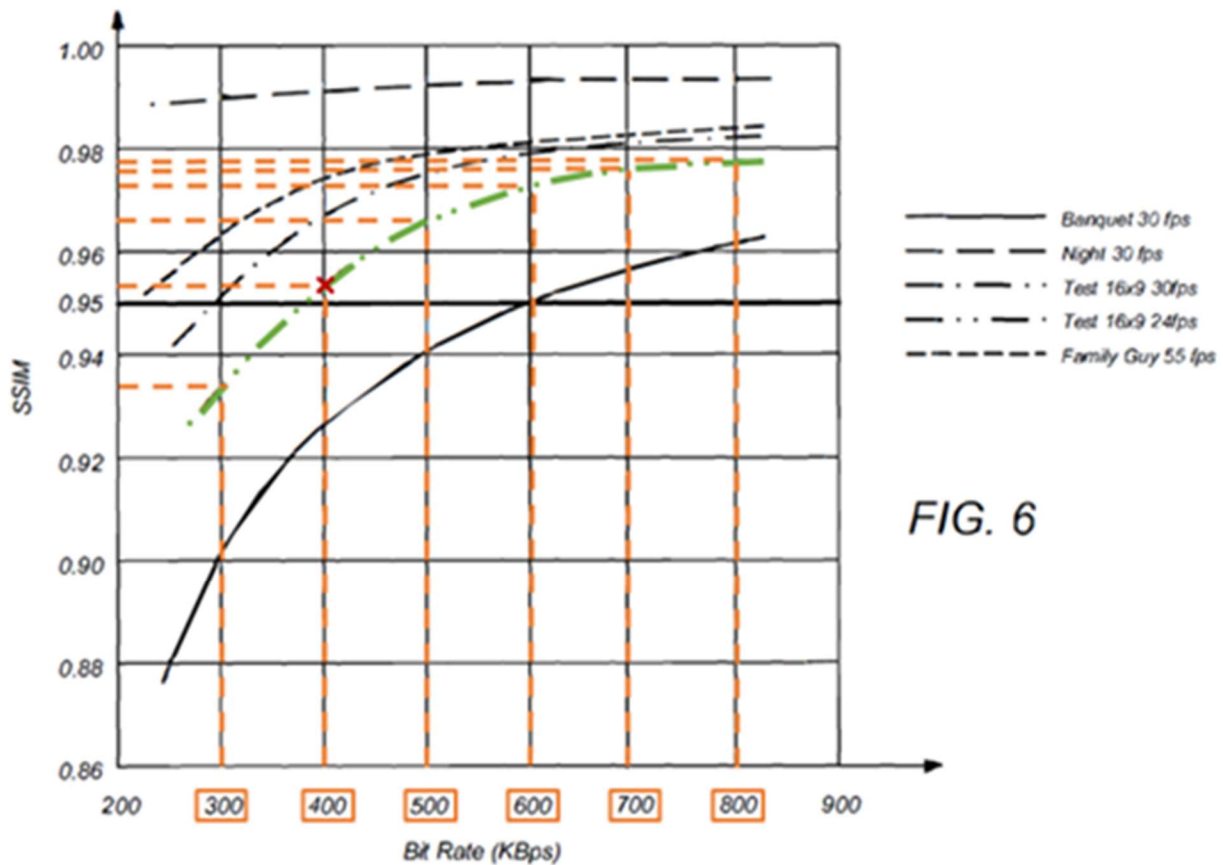
360p, 288p, 240p” to account for variations in “customers’ Internet service” during adaptive streaming. EX1006, 7:67-8:6, 8:17-22, 8:38-40. Liao explains that “[l]ower resolution versions (e.g. 240p) are transmitted when the bandwidth of the communications link is lower and higher resolution versions (e.g. 480p) are transmitted when communications bandwidth permits.” *Id.*, 8:3-6.

214. Accordingly, Ronca in view of Liao discloses “identifying a plurality of resolutions for encoding of video content.”

- (4) [1.c(i)] “for each particular resolution of the plurality of resolutions: encoding at least a portion of the source video content multiple times using the particular resolution and multiple different target bitrates for the particular resolution; and”

215. Ronca in view of Liao and general knowledge of a POSITA renders claim limitation [1.c(i)] obvious. The combination of Ronca and Liao teaches identifying a “plurality of resolutions” for the encoding of source content. *Supra* § IX.C.1(3). Liao further teaches encoding each of these particular resolutions at multiple different target bitrates. EX1006, Fig. 6. As noted in Liao, “[s]ince customers’ Internet service varies widely in bandwidth, different versions of the media program can be generated, each with a different resolution, typically 480p, 360p, 288p or 240p.” *Id.*, 7:67-8:6. Each of these alternate resolution versions “should be encoded with different bit rates.” *Id.*, 8:17-22, 8:38-40. For example, Liao discloses a representation of a TEST source video (annotated below in green)

that is encoded at multiple target bitrates: 300 kbps, 400 kbps, 500 kbps, 600 kbps, 700 kbps, and 800 kbps. *Id.*, 8:27-37, Fig. 6 (annotated).



Id., Fig. 6 (annotated).

216. Accordingly, Ronca in view of Liao discloses “for each particular resolution of the plurality of resolutions: encoding at least a portion of the source video content multiple times using the particular resolution and multiple different target bitrates for the particular resolution.”

- (5) [1.c(ii)] “for each particular resolution of the plurality of resolutions . . . evaluating quality for each of the multiple encodings;”

217. Ronca in view of Liao and general knowledge of a POSITA teaches claim limitation [1.c(ii)]. Ronca and Liao teach encoding a source video at different resolutions using multiple target bit rates for each resolution. *Supra* § IX.C.1(4). A POSITA would have looked to Liao’s quality evaluation to select the optimal resolution-bitrate combinations. *Supra* § VIII.B.3. Specifically, Liao teaches a way to evaluate the quality of each encoding using SSIM, “which is a measure of the similarity between two images.” EX1006, 8:40-42. As explained above, SSIM may be used to determine the “relationship between video quality and bitrate for different transcoded media programs.” *Id.*, 8:38-40. For example, Liao discloses a TEST source video encoded at multiple target bitrates: 300 kbps, 400 kbps, 500 kbps, 600 kbps, 700 kbps, and 800 kbps. *Id.*, 8:27-37, Fig. 6 (annotated).

218. TEST (green below) encoded at a resolution having a 16x9 aspect ratio with a target bitrate of 400 kbps results in an SSIM between 0.95 and 0.96 (blue dashed line below). *Id.*, 8:1-2, Fig. 6 (annotated). Encoding TEST at a resolution having a 16x9 aspect ratio with a target bitrate of 500 kbps results in an SSIM between 0.96 and 0.97 (purple dashed line below). *Id.* In addition, encoding TEST at a resolution having a 16x9 aspect ratio with a target bitrate of 600 kbps has an SSIM between 0.97 and 0.98 (yellow dashed line below). *Id.*, Fig. 6 (annotated).

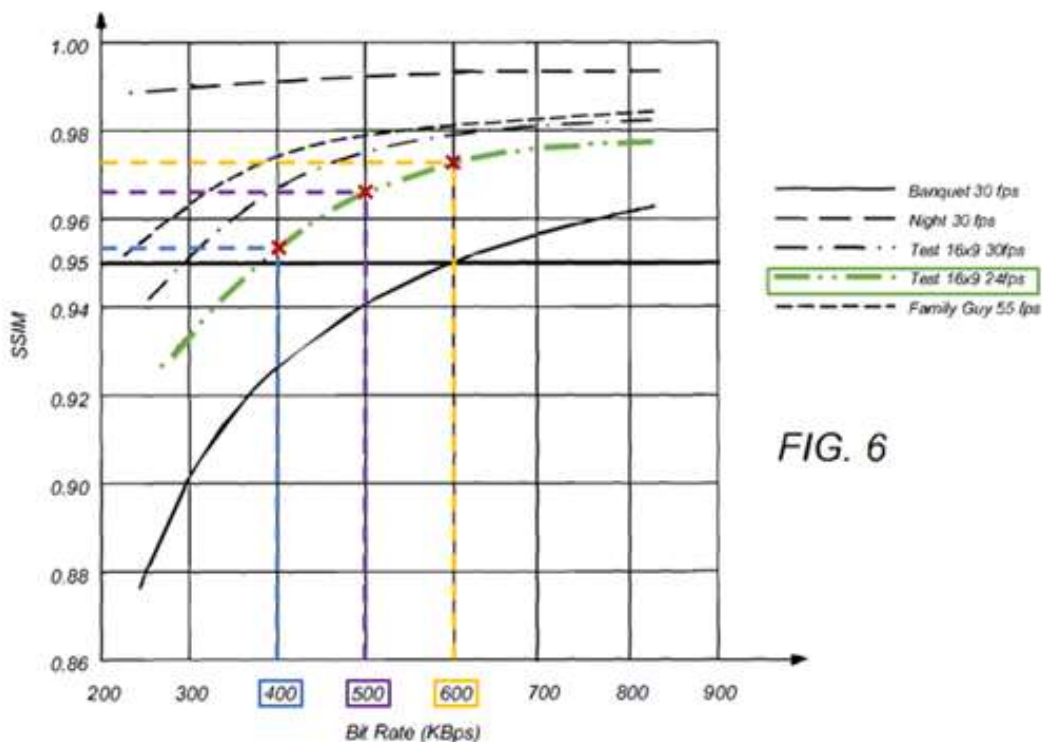


FIG. 6

Id., Fig. 6 (annotated).

219. It would have been obvious to a POSITA to use this quality evaluation for encodings of each encoding resolution. *Id.*, 8:38-9:16; *see also* EX1005, FIG. 7-2, 127-132, 117-18, 136-37; EX1014, 3-4, 7-8, 10. Adaptive bitrate streaming generally involves “different versions of the media program . . . each [encoded] with a different resolution, typically 480p, 360p, 288p or 240p.” EX1006, 7:67-8:6. As explained above, applying the quality evaluation of Liao for a stream makes it possible to identify an optimal bitrate that achieves a pre-specified quality level without expending transmission resources on streams with a bitrate that is needlessly high. *Id.*, 8:38-9:16; *see also* EX1005, Fig. 7-2, 127-132, 117-18, 136-37; EX1014, 3-4, 7-8, 10. A POSITA would understand that performing this quality evaluation

on each of the “different versions of the media program . . . each [encoded] with a different resolution” serves to enhance these resource savings. EX1006, 8:38-9:16; *see also* EX1005, Fig. 7-2, 127-132, 117-18, 136-37; EX1014, 3-4, 7-8, 10.

220. Accordingly, the combination of Ronca and Liao teaches “evaluating quality for each of the multiple encodings” using the SSIM heuristic for each resolution of the plurality of resolutions (each encoded at multiple target bit rates).

- (6) [1.d)] “selecting a plurality of resolution and target bitrate combinations for a plurality of alternative streams based on the evaluated quality of each of the multiple encodings;”

221. Ronca in view of Liao and general knowledge of a POSITA teaches claim limitation [1.d]. Ronca and Liao teach encoding source video content at a number of resolutions and for each of these resolutions, encoding the source video content using multiple target bitrates, and “evaluating [the] quality for each of the multiple encodings.” *Supra* §§ IX.C.1(4), -(5). The quality evaluation of Liao results in the **selection** of a “plurality of resolution and target bitrate combinations” based on SSIM.

222. Liao teaches selecting combinations of resolution and bitrate by selecting encodings having a minimum desired SSIM (quality evaluation metric). EX1006, 9:7-16. For example, Liao discloses encoding a TEST source video at multiple target bitrates: 300 kbps, 400 kbps, 500 kbps, 600 kbps, 700 kbps, and 800 kbps. *Id.*, 8:1-3, 8:27-37, Fig. 6. Liao teaches selecting specific combinations of

resolution and bitrate from these encodings by selecting the lowest bitrate encoding having a minimum SSIM (a minimum quality) of at least 0.95, so that adaptive streaming system need not waste additional transmission resources on higher bitrate encodings, which in this case will result in the selection of a 400 kbps stream. EX1006, 9:7-16, Fig. 6. In other words, the 400 kbps bitrate stream represents the best tradeoff of quality for ease of transmission because it is the lowest bitrate stream to achieve the desired quality level; the lower bitrate 300 kbps stream does not achieve the minimum SSIM requirement (quality level), while the higher bitrate encodings (500, 600, 700, or 800 kbps) are harder to transmit and require more resources to transmit than the 400 kbps stream. EX1006, 8:38-9:16; EX1005, 15-17, 126-127.

223. As discussed above, a POSITA would have understood that selecting a target bitrate based on SSIM's quality evaluation would be performed for each of the "different versions of a media program each [having] a different resolution, typically 480p, 360p, 288p or 240p." *Id.* 8:1-9, Fig. 6; *see also* EX1014, 3-4, 7-8. Performing such an evaluation ensures that the lowest bitrate stream of a desired quality level is selected for each resolution so that transmission resources are not wasted on higher bitrate streams. *Id.*, 8:38-9:16; *see also* EX1005, Fig. 7-2, 127-132, 117-18, 136-37; EX1014, 3-4, 7-8. A POSITA would further understand that multiple selections using SSIM may be made for each individual resolution for

multiple streams (e.g., a high quality, medium quality, and low quality stream) to facilitate stream switching during adaptive bitrate streaming. EX1006, 8:38-9:16; *see also* EX1005, Fig. 7-2, 127-132, 117-18, 136-37; EX1014, 3-4, 7-8.

224. Accordingly, the combination of Ronca and Liao teaches selecting a plurality of resolution (one of 480p, 360p, 288p or 240p) and bitrate (over 300 kbps and over 400 kbps) combinations for a plurality of alternative streams based on the evaluated quality (e.g., having a minimum SSIM, such as 0.95).

- (7) [1.e] “uploading encodings of a plurality of sections of the source video content at each of the selected plurality of resolution and target bitrate combinations for the plurality of alternative streams to a set of servers that form part of a content distribution system;”

225. Ronca in view of Liao and general knowledge of a POSITA teaches claim limitation [1.e]. Ronca and Liao teach “select[ing a] plurality of resolution and target bitrate combinations for the plurality of alternative streams.” *Supra* § IX.C.1(6). Ronca teaches that these multiple video encodings, encoded with a plurality of resolution-bitrate combinations and comprised of a plurality of sections or chunks, are uploaded or transmitted through data communications network 108 to a content delivery network (CDN) 104. EX1008, ¶¶ 28-38, Figs. 1-2. The CDN 104 includes “one or more” servers (e.g. 102, 106) and is part of a larger content distribution system 100. *Id.*, ¶¶ 28, 29, Fig. 1.

226. Accordingly, the combination of Ronca and Liao teach “uploading encodings of a plurality of sections of the source video content at each of the selected plurality of resolution and target bitrate combinations for the plurality of alternative streams to a set of servers that form part of a content distribution system.”

- (8) [1.f] “generating entries in a top level index describing at least one of the plurality of alternative streams;”

227. Ronca in view of Liao and general knowledge of a POSITA teaches claim limitation [1.f]. Ronca and Liao each teach uploading or transmitting a plurality of alternative streams—each representing a different combination of resolution and bit rate—to a CDN. *Supra* § IX.C.1(7). To appropriately track the uploaded encodings in the CDN, Ronca describes generating a header with an index (top level index) that describes “the audiovisual data in the corresponding stream[s] as a plurality of chunks.” EX1008 ¶¶ 7, 39, 60, 54, 55, cl. 1, Fig. 8. Specifically, the header “includes an index object” comprised of “index entries” that “may correspond to one of the chunks” making up the stream. *Id.*, ¶44. In other words, the header, generated by a multiplexer during encoding, “describe[s] the layout of a given AV stream” so that the CDN may properly catalog the stream. *Id.*, ¶¶ 7, 39, 43-44, 60.

228. Accordingly, Ronca in view of Liao discloses “generating entries in a top level index describing at least one of the plurality of alternative streams.”

- (9) [1.g] “providing the generated top level index from the content distribution system to a one or more playback devices capable of performing adaptive streaming using the encodings of the plurality of sections; and”

229. Ronca in view of Liao and general knowledge of a POSITA teaches claim limitation [1.g].

230. Ronca and Liao teach generating a “top level index” that describes at least one of the plurality of alternative streams, and “uploading encodings of a plurality of sections of the source video content” to a content distribution network/system. *Supra* §§ IX.C.1.(7), -(8).

231. As Ronca explains, a client device is provided with a header that includes the “top level index” from the CDN. EX1008 ¶¶ 7, 24, 41, 51, 54, 60, Fig. 8, cl. 1. For example, Ronca explains that the client device “may be configured to obtain the header files” or download the header files containing the “top level index” from one or more servers. *Id.*, ¶¶ 51, 54. The client device uses this header when selecting between “different alternate stream” during adaptive bitrate streaming. *Id.*, ¶¶ 7, 24, 41, 51, 54, 60, Fig. 8, cl. 1.

232. Accordingly, Ronca in view of Liao discloses “providing the generated top level index from the content distribution system to a one or more playback devices capable of performing adaptive streaming using the encodings of the plurality of sections.”

- (10) [1.h] “responding to requests from the one or more playback devices, where the requests include requests for specific encodings from the encodings of the plurality of sections.”

233. Ronca in view of Liao and general knowledge of a POSITA teaches claim limitation [1.h]. As noted above, Ronca and Liao teach “providing the generated top level index” to the viewer to “perform[]adaptive streaming using the encodings of the plurality of sections.” *Supra* § IX.C.1(9).

234. Ronca explains that streaming begins when “the client device receives a user selection of a media title.” EX1008, ¶ 60. “In response, the client device requests a list of URLs [from the CDN] identifying one or more [] streams available from the server.” *Id.*, ¶¶ 60-62, Figs. 8-9. Each of these streams are available for request as chunks (sections of the particular stream/encoding). *Id.*, ¶¶ 60-64, Figs. 8-9. Then, once the client device is authenticated, it “may download a header” with the index (top level index) “from the content server.” *Id.*, ¶¶ 60-61. Using the top level index, the client device begins downloading the chunks (segments) of the desired alternative stream. *Id.*, ¶¶ 62-64.

235. Accordingly, the combination of Ronca and Liao teaches “responding to requests from the one or more playback devices, where the requests include requests for specific encodings from the encodings of the plurality of sections.”

2. Claim 2: “The method of claim 1 further comprising assembling the encodings of the plurality of sections for a given alternative stream into a video container file.”

236. Ronca in view of Liao and general knowledge of a POSITA discloses all the limitations of claims 1 and 11 as discussed above. Ronca in view of Liao also teaches claims 2 and 12. As was noted, Ronca discloses “uploading encodings of a plurality of sections of the source video content . . . for the plurality of alternative streams.” *Supra* § IX.C.1(7).

237. Prior to uploading, Ronca teaches assembling the plurality of sections of the multiple encodings into a video container file. Specifically, Ronca discloses an encoding tool “configured to format” “video streams 265_[1-N]” (composed of a plurality of sections) into an “MPEG-2 [sic] compliant storage transport stream container” (video container file). EX1008, ¶ 39. “[V]ideo streams 265_[1-N]” represent “cop[ies] of the same general media file, encoded at a different bit rate.” As such, Ronca discloses “assembling the encodings . . . into a video container file.” *Id.*, ¶¶ 36, 39; Ex. After being packaged into an MPEG-2 video container file, the alternative streams are uploaded to the CDN. *Supra* § IX.C.1(7). Accordingly, Ronca in view of Liao discloses this claim.

3. Claim 3: “The method of claim 2 further comprising generating entries in a container index within each container file indicating sizes of portions of the given alternative stream within the video container file.” /

238. Ronca in view of Liao and general knowledge of a POSITA discloses all the limitations of claims 1 and 11 as discussed above. Ronca in view of Liao also teaches claims 3 and 13.

239. Ronca in view of Liao discloses “generating entries in a top level index describing at least one of the plurality of alternative streams” that encoded at a different bit rate.” *Supra* § IX.C.1(8); EX1008, ¶55. Ronca in view of Liao also discloses “assembling the encodings of the plurality of sections for a given alternative stream into a video container file.” *Supra* § IX.C.2.

240. Ronca provides an example of a sub-playlist (container index) with uniform resource locators (URLs) identifying at least three different encodings or alternate streams located within a video container file (orange shaded boxes). EX1008, Table I (annotated below), ¶¶54-55.

TABLE I

M3U8 Multiplexed AV Streams Playlist Example

```
#EXTM3U
#EXT-X-MEDIA-SEQUENCE:1
#EXT-X-TARGETDURATION:6
#EXT-X-KEY:METHOD=AES-128,URI="keys/crypt0.key"
#EXTINF:6,
http://ehub.netflix.com/high.tsrange/00001-1000000
#EXTINF:6,
http://ehub.netflix.com/high.tsrange/1000001-2000000
#EXTINF:6,
http://ehub.netflix.com/high.tsrange/2000001-3000000
...
```

241. Ronca also teaches that sub-playlist (container index) indicates sizes of portions of the given alternative stream within the video container file. The sub-playlist, or container index, points to a specific portion of a specific encoding or stream, in this case the “high” bitrate stream, one of “three [alternative] streams (‘low’ ‘mid’ ‘high’ stream).” *Id.* Each of the URL also specifies a specific chunk or byte range (00001-1000000, 1000001-2000000, and 2000001-3000000 (highlighted in blue boxes above)) within the “high” bitrate stream. EX1008, Table I, ¶¶ 54-55. A POSITA would have understood that the byte range of the portion of the “high” bitrate stream indicates the size of portion of the given alternative stream.

242. Also, as was noted above, the size of the chunk could be represented in the form of time entries indicating the duration of each chunk. *Supra* § IX.A.3. This

is demonstrated above in Table I, with the target duration tag, which indicates the target duration in seconds of the portions of the alternative streams. EX1008, Table 1 (annotated); EX1022, 2. Each portion of the alternative stream included in the container index is less than or equal to the target duration tag value. EX1022, 2. This tag appears at least once in the container index plain text file. *Id.*

243. A POSITA would have understood the duration of the portions of the alternative stream indicates the size of portion of the given alternative stream.

244. Accordingly, Ronca in view of Liao discloses “generating entries in a container index within each container file indicating sizes of portions of the given alternative stream within the video container file.”

4. Claim 4: “The method of claim 1, wherein adaptively streaming the encodings comprises receiving at least one request for an encoded section of one of the plurality of alternative streams from the one or more playback devices based on the generated top level index.”

245. Ronca in view of Liao and general knowledge of a POSITA discloses all the limitations of claims 1 and 11 as discussed above. Ronca in view of Liao also teaches claims 4 and 14.

246. Ronca in view of Liao discloses “providing the generated top level index” to the viewer’s device, “performing adaptive streaming using the encodings of the plurality of sections,” and “responding to request from the one or more playback devices.” *Supra* §§ IX.C.1(9), -(10).

247. Playback devices “performing adaptive streaming” request a specific encoding from the CDN using the “generated top level index.” Specifically, Ronca discloses that the top-level index may generate a playlist describing the location of each chunk within an alternative stream as a sequence of URLs, “where each URL refers to the chunks . . . as though they were present on the server as a collection of discrete files.” EX1008, ¶¶45, 61, Fig. 8.

TABLE I

M3U8 Multiplexed AV Streams Playlist Example	
#EXTM3U	
#EXT-X-MEDIA-SEQUENCE:1	
#EXT-X-TARGETDURATION:6	
#EXT-X-KEY:METHOD=AES-128,URI=“keys/crypt0.key”	
#EXTINF:6,	
http://ehub.netflix.com/high.ts	range/00001-1000000
#EXTINF:6,	
http://ehub.netflix.com/high.ts	range/1000001-2000000
#EXTINF:6,	
http://ehub.netflix.com/high.ts	range/2000001-3000000
...	

Id., Table I (annotated in pink to highlight URLs specifying the location of a chunk).

248. The client device references the index to request chunks (sections) associated with a specific alternate stream and stored in a specific location within the CDN. *Id.*, ¶¶57-58, 61, Fig. 8. The CDN in turn, “sends, in response to request [sic] from the client device, [the requested] chunk files”. *Id.*, ¶¶57-58, 61, Fig. 8.

249. Accordingly, Ronca in view of Liao discloses this claim.

5. Claim 7: “The method of claim 1, wherein both a width and a height of each resolution of the plurality of resolutions have an integer number of pixels.”

250. Ronca in view of Liao and general knowledge of a POSITA discloses all the limitations of claims 1 and 11 as discussed above. Ronca in view of Liao also teaches claims 7 and 17.

251. “[D]ifferent versions of the media program can be generated, each with a different resolution, typically 480p, 360p, 288p or 240p.” EX1006, 7:48-8:9. A POSITA would have understood that “480p, 360p, 288p or 240p” are resolutions with “both a width and a height” representing “an integer number of pixels.” *See, e.g.,* EX1005, 17. For example, the resolution “480p,” corresponds to 854 pixels (width) x 480 pixels (height). *See, e.g., id.*, 17. 854 pixels x 480 pixels corresponds to an integer 16x9 aspect ratio, which Fig. 6 of Liao also discloses for several exemplary streams.

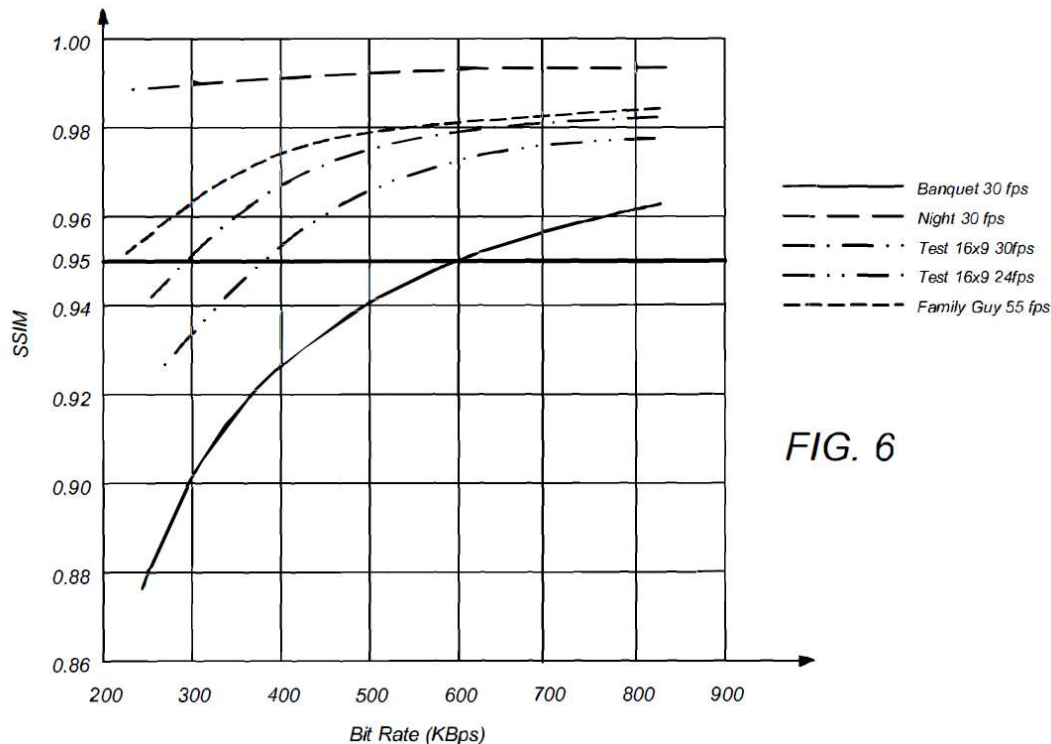


FIG. 6

Id., Fig. 6.

252. Thus, Ronca in view of Liao discloses that “both a width and a height of each resolution of the plurality of resolutions have an integer number of pixels.”

6. Claim 9: “The method of claim 1, wherein the target bitrate for each resolution is determined using a formula.”

253. Ronca in view of Liao and general knowledge of a POSITA discloses all the limitations of claims 1 and 11 as discussed above. Ronca in view of Liao also teaches claims 9 and 19.

254. Liao discloses using SSIM, “which is a measure of the similarity between two images,” to evaluate the quality of the multiple encodings, and selecting the target bit rate for each resolution based on the evaluated quality. *Supra* §§

IX.C.1(5), -(6); EX1006, 8:37-67. Liao explains that “SSIM is computed on windows of an image (typically applied only over luminescence information), wherein a measure of similarity between two windows x and y of common size NxN is defined” by the formula below. EX1006, 8:37-67.

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

wherein:

μ_x =the average of x;

μ_y =the average of y;

σ_x^2 =the variance of x;

σ_y^2 =the variance of y;

σ_{xy} =the covariance of x and y;

$c_1=(k_1L)^2$, $c_2=(k_2L)^2$ two variables to stabilize the division with weak denominator;

L=dynamic range of the pixel values

$$\left(\text{typically } 2^{\frac{\#bits}{pixel}} - 1 \right);$$

$k_1=0.01$; and

$k_2=0.03$.

Id., 8:38-67.

255. Accordingly, Ronca in view of Liao teach that “the target bitrate for each resolution is determined using a formula.”

7. Claim 10: “The method of claim 1, wherein the target bitrate of a given resolution is determined using experimentation.”

256. Ronca in view of Liao and general knowledge of a POSITA discloses all the limitations of claim 1 as discussed above. Ronca in view of Liao also teaches claim 10.

257. As discussed above, Ronca in view of Liao teaches iteratively encoding source video content, evaluating the quality of these encodings using SSIM, then selecting resolution-target bitrate combinations for the plurality of alternative streams based on the SSIM (evaluated quality of each of the encodings). *Supra* § IX.C.1. A POSITA would have understood such an iterative process to be a form of experimentation. Thus, Ronca in view of Liao teaches this claim.

8. Claim 11

- (1) [11.pre] “A non-transitory machine readable medium containing processor instructions for encoding source content as a plurality of alternative video streams and adaptively streaming the plurality of alternative video streams, each having a bitrate, where execution of the instructions by a processor causes the processor to perform a process that comprises:”

258. To the extent the preamble of claim 11 is considered limiting, Ronca in view of Liao and general knowledge of a POSITA discloses it. Ronca and Liao teach that an encoder and content distribution system, which includes the CDN, has “a central processing unit” (processor). EX1008, ¶34, Fig. 1-2; *see also* EX1006,

8:27-37 (disclosing encoders with a “special purpose processor”). The preamble of claim 1 is otherwise substantively identical to the preamble of claim 11. Therefore, I incorporate by reference my analysis of the preamble of claim 1, as my analysis of the preamble of claim 1 also applies to the preamble of claim 11. The combination of Ronca in view of Liao renders the preamble of claim 11 obvious for the same reasons as discussed regarding the preamble of claim 1. *See supra* § IX.C.1(1).

(2) [11.a] “identifying source video content;”

259. Ronca in view of Liao and general knowledge of a POSITA discloses limitation [11.a]. Limitation [11.a] is substantively identical to limitation [1.a]. Therefore, I incorporate by reference my analysis of limitation [1.a], as my analysis of limitation [1.a] also applies to limitation [11.a]. The combination of Ronca in view of Liao renders limitation [11.a] obvious for the same reasons as discussed regarding limitation [1.a]. *See supra* § IX.C.1(2).

(3) [11.b] “identifying a plurality of resolutions for encoding of video content;”

260. Ronca in view of Liao and general knowledge of a POSITA discloses limitation [11.b]. Limitation [11.b] is substantively identical to limitation [1.b]. Therefore, I incorporate by reference my analysis of limitation [1.b], as my analysis of limitation [1.b] also applies to limitation [11.b]. The combination of Ronca in view of Liao renders limitation [11.b] obvious for the same reasons as discussed regarding limitation [1.b]. *See supra* § IX.C.1(3).

- (4) [11.c(i)] “for each particular resolution of the plurality of resolutions: encoding at least a portion of the source video content multiple times using the particular resolution and multiple different target bitrates for the particular resolution; and”

261. Ronca in view of Liao and general knowledge of a POSITA discloses limitation [11.c(i)]. Limitation [11.c(i)] is substantively identical to limitation [1.c(i)]. Therefore, I incorporate by reference my analysis of limitation [1.c(i)], as my analysis of limitation [1.c(i)] also applies to limitation [11.c(i)]. The combination of Ronca in view of Liao renders limitation [1c(i)] obvious for the same reasons as discussed regarding limitation [11.c(i)]. *See supra* § IX.C.1(4).

- (5) [11.c(ii)] “for each particular resolution of the plurality of resolutions: . . . evaluating quality for each of the multiple encodings;”

262. Ronca in view of Liao and general knowledge of a POSITA discloses limitation [11.c(ii)]. Limitation [11.c(ii)] is substantively identical to limitation [1.c(ii)]. Therefore, I incorporate by reference my analysis of limitation [1.c(ii)], as my analysis of limitation [1.c(ii)] also applies to limitation [11.c(ii)]. The combination of Ronca in view of Liao renders limitation [11.c(ii)] obvious for the same reasons as discussed regarding limitation [11.c(ii)]. *See supra* § IX.C.1(5).

- (6) [11.d] “selecting a plurality of resolution and target bitrate combinations for a plurality of alternative

streams based on the evaluated quality of each of the multiple encodings;”

263. Ronca in view of Liao and general knowledge of a POSITA discloses limitation [11.d]. Limitation [11.d] is substantively identical to limitation [1.d]. Therefore, I incorporate by reference my analysis of limitation [1.d], as my analysis of limitation [1.d] also applies to limitation [11.d]. The combination of Ronca in view of Liao renders limitation [11.d] obvious for the same reasons as discussed regarding limitation [1.d]. *See supra* § IX.C.1(6).

- (7) [11.e] “uploading encodings of a plurality of sections of the source video content at each of the selected plurality of resolution and target bitrate combinations for the plurality of alternative streams to a set of servers that form part of a content distribution system;”

264. Ronca in view of Liao and general knowledge of a POSITA discloses limitation [11.a]. Limitation [11.a] is substantively identical to limitation [1.a]. Therefore, I incorporate by reference my analysis of limitation [1.a], as my analysis of limitation [1.a] also applies to limitation [11.a]. The combination of Ronca in view of Liao renders limitation [11.a] obvious for the same reasons as discussed regarding limitation [1.a]. *See supra* § IX.C.1(7).

- (8) [11.f] “generating entries in a top level index describing at least one of the plurality of alternative streams;”

265. Ronca in view of Liao and general knowledge of a POSITA discloses limitation [11.f]. Limitation [11.f] is substantively identical to limitation [1.f].

Therefore, I incorporate by reference my analysis of limitation [1.f], as my analysis of limitation [1.f] also applies to limitation [11.f]. The combination of Ronca in view of Liao renders limitation [11.f] obvious for the same reasons as discussed regarding limitation [1.f]. *See supra* § IX.C.1(8).

- (9) [11.g] “providing the generated top level index from the content distribution system to one or more playback devices capable of performing adaptive streaming using the encodings of the plurality of sections; and”

266. Ronca in view of Liao and general knowledge of a POSITA discloses limitation [11.g]. Limitation [11.g] is substantively identical to limitation [1.g]. Therefore, I incorporate by reference my analysis of limitation [1.g], as my analysis of limitation [1.g] also applies to limitation [11.g]. The combination of Ronca in view of Liao renders limitation [11.g] obvious for the same reasons as discussed regarding limitation [1.g]. *See supra* § IX.C.1(9).

- (10) [11.h] “responding to requests from the one or more playback devices, where the requests include requests for specific encodings from the encodings of the plurality of sections.”

267. Ronca in view of Liao and general knowledge of a POSITA discloses limitation [11.h]. Limitation [11.h] is substantively identical to limitation [1.h]. Therefore, I incorporate by reference my analysis of limitation [1.h], as my analysis of limitation [1.h] also applies to limitation [11.h]. The combination of Ronca in

view of Liao renders limitation [11.h] obvious for the same reasons as discussed regarding limitation [1.h]. *See supra* § IX.C.1(10).

9. Claim 12: “The non-transitory machine readable medium of claim 11, wherein the process further comprises assembling the encodings of the plurality of sections for a given alternative stream into a video container file.”

268. In addition to disclosing each of the limitations of Claim 11, Ronca in view of Liao and general knowledge of a POSITA discloses claim 12. Claim 12 is substantively identical to Claim 2. Therefore, I incorporate by reference my analysis of Claim 2, as my analysis of Claim 2 also applies to Claim 12. The combination of Ronca in view of Liao renders Claim 12 obvious for the same reasons as discussed regarding Claim 2. *See supra* § IX.C.2.

10. Claim 13: “The non-transitory machine readable medium of claim 12, wherein the process further comprises generating entries in a container index within each container file indicating sizes of portions of the given alternative stream within the video container file.”

269. In addition to disclosing each of the limitations of Claims 11 and 12, Ronca in view of Liao and general knowledge of a POSITA discloses claim 13. Claim 13 is substantively identical to Claim 3. Therefore, I incorporate by reference my analysis of Claim 3, as my analysis of Claim 3 also applies to Claim 13. The combination of Ronca in view of Liao renders Claim 13 obvious for the same reasons as discussed regarding Claim 3. *See supra* § IX.C.3.

- 11. Claim 14: “The non-transitory machine readable medium of claim 11, wherein adaptively streaming the encodings comprises receiving at least one request for an encoded section of one of the plurality of alternative streams from the one or more playback devices based on the generated top level index.”**

270. In addition to disclosing each of the limitations of Claim 11, Ronca in view of Liao and general knowledge of a POSITA discloses claim 14. Claim 14 is substantively identical to Claim 4. Therefore, I incorporate by reference my analysis of Claim 4, as my analysis of Claim 4 also applies to Claim 14. The combination of Ronca in view of Liao renders Claim 14 obvious for the same reasons as discussed regarding Claim 4. *See supra* § IX.C.4.

- 12. Claim 17: “The non-transitory machine readable medium of claim 11, wherein both a width and a height of each resolution of the plurality of resolutions have an integer number of pixels.”**

271. In addition to disclosing each of the limitations of Claim 11, Ronca in view of Liao and general knowledge of a POSITA discloses claim 17. Claim 17 is substantively identical to Claim 7. Therefore, I incorporate by reference my analysis of Claim 7, as my analysis of Claim 7 also applies to Claim 17. The combination of Ronca in view of Liao renders Claim 17 obvious for the same reasons as discussed regarding Claim 7. *See supra* § IX.C.5.

13. Claim 19: “The non-transitory machine readable medium of claim 11, wherein the target bitrate for each resolution is determined using a formula.”

272. In addition to disclosing each of the limitations of Claim 11, Ronca in view of Liao and general knowledge of a POSITA discloses claim 19. Claim 19 is substantively identical to Claim 9. Therefore, I incorporate by reference my analysis of Claim 9, as my analysis of Claim 9 also applies to Claim 19. The combination of Ronca in view of Liao renders Claim 19 obvious for the same reasons as discussed regarding Claim 9. *See supra* § IX.C.6.

D. Ground 4: Claims 6 and 16 of 785 Patent are Rendered Obvious by Ronca in view of Liao and Gu

1. Claim 6: “The method of claim 1, wherein identifying a plurality of resolutions comprises identifying resolutions with target bitrates within a predetermined percentage of a bitrate associated with an alternative stream of the plurality of alternative streams.”

273. As noted above, Ronca in view of Liao teach claim 1. The combination of Ronca, Liao, and Gu and general knowledge of a POSITA teaches Claim 6.

274. Ronca in view of Liao teaches identifying a plurality of resolutions, each having a plurality of target bitrates, encoding source video content at each resolution and their respective target bit rates, and selecting a plurality of alternative streams based on the evaluated quality of the resulting encodings. *See supra* § IX.C.1.

275. Both Ronca and Liao each teach encoding multiple versions of source content to be used in adaptive streaming. EX1008, ¶38; EX1006, 7:63-8:9. Ronca teaches using an encoder to “generat[e] multiple video . . . encodings of the same media presentation, each encoded at different bit rates”. *Id.*, ¶¶24, 38. Liao further teaches that in addition to “different bit rates,” each of these encodings should have “a different resolution, typically 480p, 360p, 288p, 240p” to account for variations in “customers’ Internet service” during adaptive streaming. EX1006, 7:67-8:6, 8:17-22, 8:38-40. Gu describes a way to determine a reasonable spacing between these different bitrates.

276. Specifically, Gu teaches “decreas[ing] progressively in a proportional, logarithmic or other decreasing manner” the target bit rates of alternative streams. EX1007, 7:58-61. One specific example is decreasing the target bitrate of each alternative stream by a predetermined proportion of 25% (or $\frac{1}{4}$) from the preceding stream (*e.g.*, a first stream with the highest bitrate, a second stream with a bitrate 75% of the first bitrate, a third stream with a bitrate 50% of the first bitrate). *Id.*, 7:54-67.

277. Accordingly, a POSITA would have understood that the combination of Ronca, Liao, and Gu teaches “identifying a plurality of resolutions comprises identifying resolutions with target bitrates within a predetermined percentage of a bitrate associated with an alternative stream of the plurality of alternative streams.”

2. **Claim 16: “The non-transitory machine readable medium of claim 11, wherein identifying a plurality of resolutions comprises identifying resolutions with target bitrates within a predetermined percentage of a bitrate associated with an alternative stream of the plurality of alternative streams.”**

278. The combination of Ronca, Liao, and Gu and general knowledge of a POSITA teaches Claim 16.

279. Claim 16 is substantively identical to Claim 6. Therefore, I incorporate by reference my analysis of Claim 6, as my analysis of Claim 6 also applies to Claim 16. The combination of Ronca in view of Liao renders Claim 16 obvious for the same reasons as discussed regarding Claim 6. *See supra* § IX.D.1.

X. CONCLUSION

280. For the reasons I have presented above, it is my opinion that the references serving as the bases for the grounds of rejection in the Petition render claims 1-4, 6-14, and 16-19 of the challenged patent unpatentable.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Dated: 6/30/2025



Dr. Hari Kalva