

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

Avigilon USA Corporation and Avigilon Corporation,
Petitioners,

v.

Canon Inc.,
Patent Owner.

Case No. IPR2020-TBD
Patent No. 10,135,952

PETITION FOR *INTER PARTES* REVIEW

TABLE OF CONTENTS

I.	INTRODUCTION	1
II.	MANDATORY NOTICES	3
	A. Real Party-In-Interest	3
	B. Related Matters.....	3
	C. Counsel And Service Information.....	3
III.	PAYMENT OF FEES UNDER 37 C.F.R. § 42.103.....	4
IV.	CERTIFICATION OF GROUNDS FOR STANDING	5
V.	OVERVIEW OF CHALLENGE AND RELIEF REQUESTED.....	5
	A. Prior Art Patents And Printed Publications.....	5
	B. Relief Requested.....	6
VI.	THE '952 PATENT.....	7
	A. Claims.....	7
	B. Summary Of The Specification.....	8
	C. Summary Of The Prosecution History	12
	D. Person Of Ordinary Skill In The Art.....	14
VII.	CLAIM CONSTRUCTION	15
VIII.	OVERVIEW OF THE PRIOR ART	15
	A. Overview Of <i>Irie</i>	15
	B. Overview of <i>Van Deventer</i>	16
IX.	SPECIFIC GROUND FOR PETITION	18
	A. Ground I: Claims 7 and 11-14 Are Rendered Obvious By <i>Irie</i> In View Of <i>Van Deventer</i>	18

X.	CONCLUSION.....	40
----	-----------------	----

TABLE OF AUTHORITIES

	Page(s)
Cases	
<i>Canon Inc. v. Avigilon USA Corp. et al.</i> , Case No. 19-cv-10931	3
<i>Phillips v. AWH Corp.</i> , 415 F.3d 1303 (Fed. Cir. 2005)	15
Statutes	
35 U.S.C. § 102	5, 6
35 U.S.C. § 103	2, 6, 7, 13
Rules	
Rule 42.104(a)	5
Other Authorities	
37 C.F.R. § 42.10(b)	4
37 C.F.R. § 42.15(a)(1)	4
37 C.F.R. § 42.103	4

LIST OF EXHIBITS

Exhibit	Description
1100	U.S. Patent No. 10,135,952
1101	Declaration of Dr. Kevin Almeroth
1102	U.S. Patent No. 10,135,952 File History
1103	Japanese Patent Application Publication No. 2009-212821 (with Translation and Certificate) (“ <i>Irie</i> ”)
1104	U.S. Patent No. 9,860,572 (“ <i>Van Deventer</i> ”)
1105	<i>Mavlankar et al., An Interactive Region-of-Interest Video Streaming System for Online Lecture Viewing</i> (2010) (“ <i>Mavlankar Publication</i> ”)
1106	U.S. Patent Application Publication No. 2003/0142872A1 (“ <i>Koyanagi</i> ”)
1107	Inquiry of History Information - Notice of Reasons for Refusal (JP Patent Application 2014-514071)

I. INTRODUCTION

U.S. Patent No. 10,135,952 (Ex. 1100, “the ’952 patent”) purports to invent a technique for transmitting video data over IP networks by dividing video data into multiple partitions. The claims of the ’952 patent reflect such a procedure. For example, the challenged claims claim transmitting a notification that includes spatial information to a client, a receipt of a message from a client identifying a “Region Of Interest,” (“ROI”) and transmitting the corresponding identified videos to the client. Ex. 1100 at claim 1. The spatial information in the notification represents a spatial position of each of the n partial videos represented by a pixel position, where the pixel position “is represented by a value according to a number of pixels between a reference pixel position on the frame and a pixel position in a top-left corner of each of the n partial videos in an x-axis direction and a y-axis direction.” *Id.*

But the claimed invention is unpatentable. At bottom, it claims nothing more than well-known practices of splitting up video into tiles and transmitting information regarding the video taught in numerous prior art patents and publications and, in particular, in the *Irie* Japanese patent application and the *Van Deventer* patent. Each of those prior art references teaches splitting video into tiles and sending it to a client for viewing. *Irie* teaches the transmission of video data from a source to a client with “multiple blocks” based on video resolution, from which the client chooses a Region Of Interest for higher resolution video blocks.

Van Deventer describes methods and systems for processing “spatially segmented content originating from a content delivery network” through the transmission of a spatial manifest data structure or file. *See* Ex. 1104 at Abstract.

The prosecution history of the '952 patent confirms that transmitting video using tiles was well known at the time of the invention. Indeed, when the original claims directed to transmitting a notification to a client including spatial information for “n partial videos” and having a client select partial videos according to a Region Of Interest were rejected by the examiner under 35 U.S.C. § 103 over *Mavlankar et al.*, *An Interactive Region-of-Interest Video Streaming System for Online Lecture Viewing* (2010) (Ex. 1105, “*Mavlankar Publication*”), in view of U.S. Patent Application No. 2003/0142872A1 (Ex. 1106, “*Koyanagi*”), Patent Owner did not argue that the claims were non-obvious as written. Instead, Patent Owner amended the claim to add a requirement that the spatial information include “a pixel position”—spatial information using pixels—and argued that such information was novel. The examiner then further amended the claims to specify that the pixel position was represented by a value relative to a reference pixel position and a pixel position in the top-left corner of the videos. The Patent Owner’s argument in combination with the examiner’s amendment convinced the examiner to overcome the prior art of record, but was incorrect. *Irie* and *Van Deventer*, which were not before the Patent Office during prosecution, disclose a tiled video-delivery system

that transmits notification from a server containing the positions of the tiles and receives a request from a client to provide a specific high-resolution portion of the video.

II. MANDATORY NOTICES

A. Real Party-In-Interest

Avigilon USA Corporation, Avigilon Corporation, and Motorola Solutions, Inc. are the real parties-in-interest.

B. Related Matters

The '952 patent is involved in litigation *Canon Inc. v. Avigilon USA Corp. et al.*, Case No. 19-cv-10931, filed on April 23, 2019 in the United States District Court for the District of Massachusetts. Petitioners also filed a separate petition for IPR for the '952 patent, IPR2020-00656, on February 25, 2020 where claims 1-6, 8-10, and 15-21 were challenged. These cases may affect, or be affected by, decisions in this proceeding.

C. Counsel And Service Information

Lead Counsel	Back-up Counsel
Reza Dokhanchy, (Reg. No. 62,795) KIRKLAND & ELLIS LLP 555 California Street San Francisco, CA 94104 Telephone: 415-439-1400 Facsimile: 415-439-1500 reza.dokhanchy@kirkland.com Avigilon_Axis@kirkland.com	Adam R. Alper, (to seek <i>pro hac vice admission</i>) KIRKLAND & ELLIS LLP 555 California Street San Francisco, CA 94104 Telephone: 415-439-1400 Facsimile: 415-439-1500 adam.alper@kirkland.com

Lead Counsel	Back-up Counsel
	<p>Michael W. De Vries, (to seek <i>pro hac admission</i>) KIRKLAND & ELLIS LLP 555 South Flower Street Los Angeles, CA 90071 Telephone: 213-680-8400 Facsimile: 213-680-8500 michael.devries@kirkland.com</p> <p>Akshay S. Deoras, (to seek <i>pro hac vice admission</i>) KIRKLAND & ELLIS LLP 555 California Street San Francisco, CA 94104 Telephone: 415-439-1400 Facsimile: 415-439-1500 akshay.deoras@kirkland.com</p>

Avigilon concurrently submits a Power of Attorney, 37 C.F.R. § 42.10(b), and consents to electronic service directed to the following email address: Avigilon_Axis@kirkland.com.

III. PAYMENT OF FEES UNDER 37 C.F.R. § 42.103

The undersigned authorizes the Office to charge the fee set forth in 37 C.F.R. § 42.15(a)(1) for this Petition to Deposit Account No. 506092. Review of 5 claims is requested, and thus no excess claim fees are required. The undersigned further

authorize payment for any additional fees that may be due in connection with this Petition to be charged to the above-referenced deposit account.

IV. CERTIFICATION OF GROUNDS FOR STANDING

Petitioners certify pursuant to Rule 42.104(a) that the '952 patent is available for *inter partes* review and that Petitioners are not barred or estopped from requesting an *inter partes* review challenging the patent claims on the grounds identified in this Petition.

V. OVERVIEW OF CHALLENGE AND RELIEF REQUESTED

Petitioners challenge claims 7 and 11-14 ("the challenged claims") of the '952 patent and requests that each challenged claim be canceled.

A. Prior Art Patents And Printed Publications

Petitioners rely upon the patents and publications listed in the List of Exhibits, and in particular on the following prior art patents:

- **Irie** - Japanese published unexamined patent application No. 2009-212821 to Yuji Irie, *et al.*, filed March 4, 2008, and published on September 17, 2009. *Irie* (Ex. 1006) is prior art under pre-AIA 35 U.S.C. § 102(a) if the priority date for the '952 patent is October 12, 2012 and it is a prior art under AIA 35 U.S.C. § 102(a)(1) if the priority date for the '952 patent is April 16, 2013.
- **Van Deventer** – U.S. Patent No. 9,860,572 to Van Deventer, *et al.*, filed as PCT/EP2012/060801 on June 7, 2012, and published as WO2012/168365 on

December 13, 2012, with a claim of foreign application priority to EP 11169068 filed on June 8, 2011. *Van Deventer* (Ex. 1104) issued on January 2, 2018, is prior art under pre-AIA 35 U.S.C. § 102(e) if the priority date for the '952 patent is October 12, 2012 because the PCT was filed after November 29, 2000 and published in the English language. *Van Deventer* is a prior art under AIA 35 U.S.C. § 102(a)(2) if the priority date for the '952 patent is April 16, 2013.

As explained in more detail below, the '952 patent claims foreign priority to two Great Britain patent applications GB 1218370.3, filed October 12, 2012, and GB 1306897.8, filed April 16, 2013. Although Petitioners do not concede the validity of any of these foreign priority claims, the prior art references relied upon by Petitioners predate the '952 patent regardless of whether the priority claims apply or not. Accordingly, Petitioners do not address here the issue of foreign priority, but do reserve the right to challenge foreign priority at a later date or in other proceedings.

B. Relief Requested

Petitioners request that the Patent Trial and Appeal Board cancel the challenged claims because they are unpatentable under 35 U.S.C § 103. The specific ground for the Petition is set forth in the table below, and is supported by the declaration of Dr. Kevin Almeroth (Ex. 1101).

Ground	Claims	Proposed Statutory Rejection
I	7 and 11-14	Obvious under § 103 by <i>Irie</i> in view of <i>Van Deventer</i>

VI. THE '952 PATENT

The '952 patent was issued on November 20, 2018, from U.S. Application No. 15/783,999 (“the '999 Application”) filed on October 13, 2017. The '999 application claims priority to U.S. Application No. 14/434,741, filed as PCT Application No. PCT/EP2013/071370 on October 11, 2013, and claims foreign priority to GB Application No. 1218370.3 (filed October 12, 2012) and GB Application No. 1306897.8 (filed April 16, 2013). The named inventors are Franck Denoual, Hervé Le Floch, Frédéric Maze, Jean Le Feuvre, and Cyril Concolato. Canon Kabushiki Kaisha is the original assignee. The '952 patent purports to address the problem of displaying high-quality videos received over the Internet on mobile devices. The allegedly inventive solution is to split the video into “tiles,” which a user can then request by sending a message to the server.

A. Claims

The '952 patent has 21 claims, including 4 independent claims (claims 1, 3, 16, and 19), and 17 dependent claims. Claims 7 and 11-14 are challenged in this Petition.

The four independent claims 1, 3, 16, and 19 are similar in scope. Each claim recites dividing a video into a number (“n”) of partitions in order to create n

independent videos. *See, e.g.*, '952 patent at claim 3. Information regarding the position of the videos is transmitted to a client, who requests one of the videos by selecting a particular "Region Of Interest" within the video. *Id.* The server then sends the corresponding videos to the client. *Id.*

Dependent claim 7 claims a spatial size that is different from the spatial size of another spatial region of the frame of the video, and the spatial size information representing both of the spatial size of one spatial region and another spatial region. Dependent claim 10 claims that each frame is compressed independently. Dependent claims 11 and 12 require HTTP addresses / URLs associated with the information regarding the spatial regions received from the client. Dependent claim 13 requires the spatial position of one time period to be different from a second time period, and the first notifications for the first and second time periods be transmitted to the client. Dependent claim 14 further requires the notifications to include (i) spatial position information and (ii) spatial size information related to each of the time periods.

B. Summary Of The Specification

The specification of the '952 patent discusses techniques for transmitting information from a server to a client to inform the client of the positions of the divisions of transmitted video streams. The '952 patent begins by describing two well-known prior art concepts. First, the patent describes a prior-art standard for

media streaming called “Dynamic Adaptive Streaming over HTTP,” or “DASH.” Ex. 1100 at 1:36-40. The prior-art DASH standard, as the patent explains, creates an association between the contents of a media presentation and HTTP addresses. *Id.* at 1:40-42. This association information is stored within a “manifest file,” also called a “Media Presentation Description” (“MPD”) file. *Id.* at 1:43-46. The manifest file is sent to the client, which can use the MPD file to request specific media content to download and play. *Id.* at 1:47-55. This functionality is illustrated in Figure 1a, which shows a client requesting specific segments of a video identified by a manifest file:

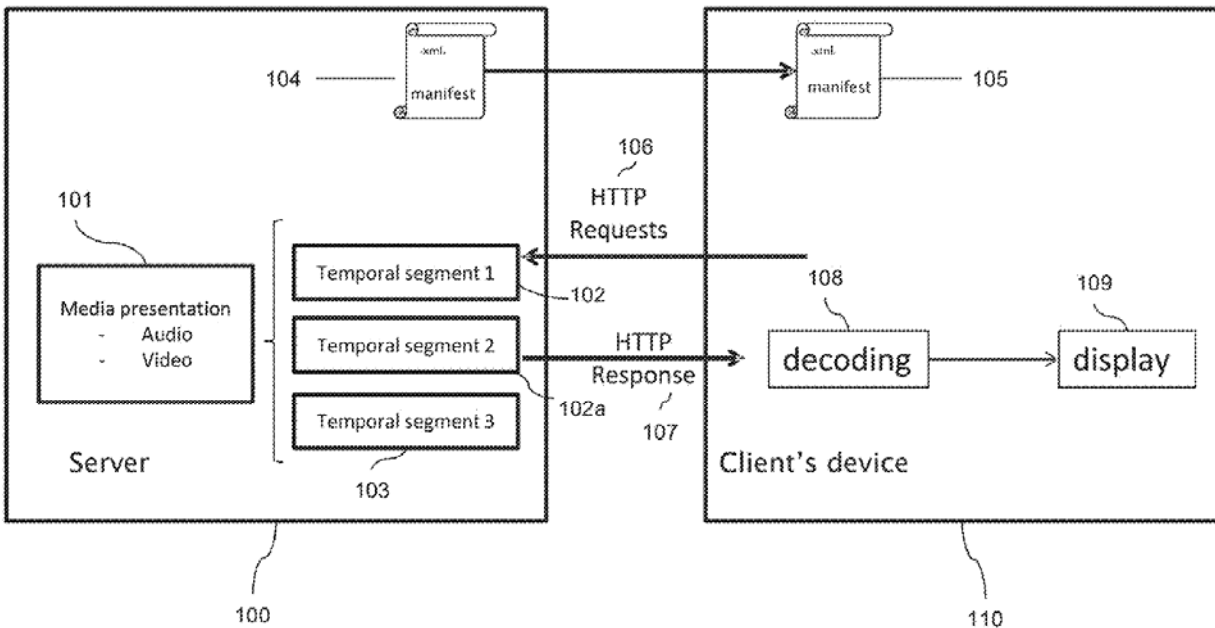


Figure 1a

Ex. 1100 at Fig. 1a.

Second, the patent describes prior-art solutions to display high-resolution content on mobile devices. One such prior-art solution involved the use of “tiles” to split up a video into sub-parts. *Id.* at 2:5-8. As the patent explains “[i]f the user of a mobile application wants to display or focus on sub-parts of the video, only the tiles corresponding to the sub-part are transmitted.” *Id.*

The patent evidently wants to combine these two known prior-art solutions, but contends that this combined scheme suffers from a problem, however—namely, that “[t]here is no way in the [DASH] manifest to signal that each track is a sub-part of the same video.” *Id.* at 2:15-20. Although a client could request specific tiled video sub-parts using the prior-art DASH standard, the patent claims that the process would require multiple HTTP requests. *Id.* at 2:20-29.

The patent then acknowledges that other systems have accomplished the same solution: in the *Mavlankar Publication*, a “specific manifest” is proposed that “describ[es] tiles for a scalable video” by providing “an identifier and a piece of location information for each tile.” *Id.* at 2:30-38. Clients request a specific tile by using an HTTP request. *Id.* at 2:38-41. This system is criticized as allegedly “requir[ing] processing at server-side [*sic*] to retrieve,” which can allegedly only be done “by a very specific server.” *Id.* at 2:41-45.

The ’952 patent goes on to describe a purported solution to this problem. Like the prior art, the patent proposes dividing the video into tiles, or sub-tracks. *Id.* at

2:51-55. A description file is sent to a client including information about how each of the tiles is organized, along with URLs to allow the client to access those files. *Id.* at 2:57-60. After the client receives the description file, the client can choose a URL associated with a particular “Region Of Interest” and request the corresponding video tile or sub-track from the server. *Id.* at 2:61-3:3. This process is illustrated in the flow diagram in Figure 1b:

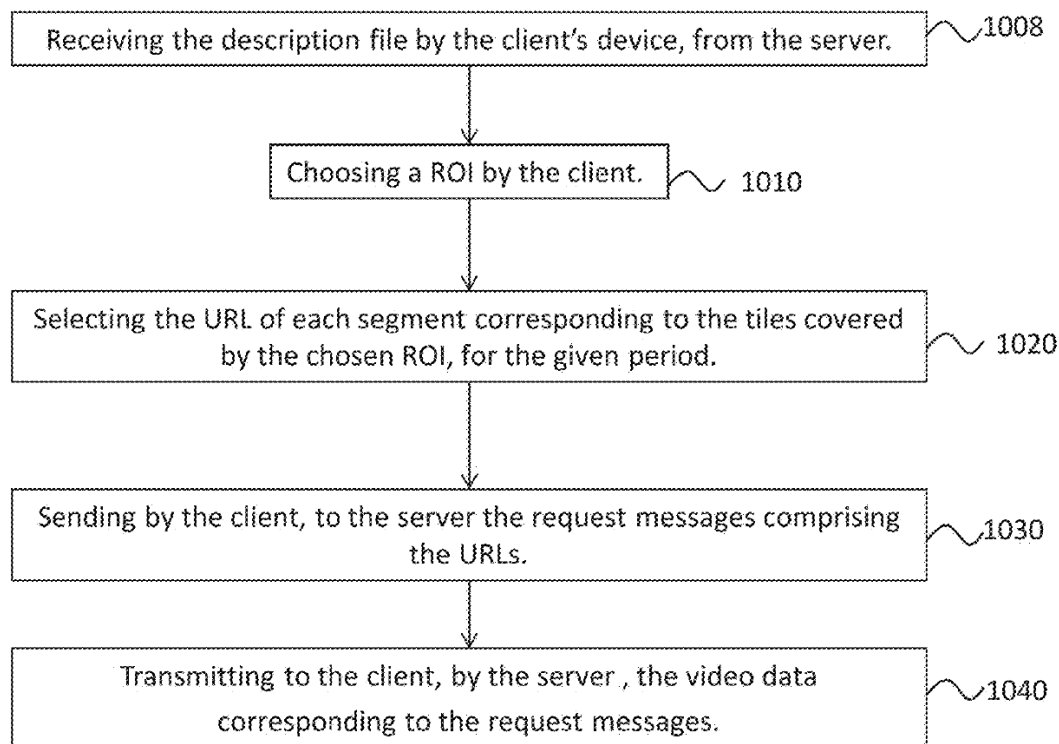


Figure 1b

Id. at Fig. 1b.

The claims of the '952 patent are focused on a particular technique for transmitting the information in the description file—namely, by identifying

individual tiles in the description file by using a “pixel position” on a video frame, where the pixel position “is represented by a value according to a number of pixels between a reference pixel position on the frame and a pixel position in a top-left corner of each of the n partial videos in an x-axis direction and y-axis direction.” *Id.* at claim 1. While the specification of the ’952 patent does describe indicating the position of a tile using its position in pixels (*see, e.g., id.* at 4:34-46), the term “reference pixel” does not appear in the specification of the ’952 patent. Instead, that term was added during prosecution as described in Section VI.C below.

C. Summary Of The Prosecution History

As stated in the introductory part of this section, Section VI, the ’952 patent issued from the ’999 Application claims priority to a foreign application filed in Great Britain on October 12, 2012. The original claims of the ’999 application claimed “[a] method for streaming data corresponding to a given video, each frame of the video being divided into n partitions . . . in order to create n independent partial videos” that comprises transmitting a notification to the client with “spatial information about a spatial organization of the n partial videos.” Ex. 1102 at 4. The method then required “receiving from the client which received the notification “messages for requesting one or more partial videos according to a Region Of Interest selected based on an operation by a user of the client.” *Id.* Finally, the

server transmitted to the client video data corresponding to the one or more partial videos in response to the messages sent by the client. *Id.*

The Examiner rejected the original claims under 35 U.S.C. § 103 as being obvious over the *Mavlankar Publication* in view of *Koyanagi*. As the Examiner explained, the *Mavlankar Publication* described dividing a video up into tiles so that the client could select a Region Of Interest for higher-resolution video:

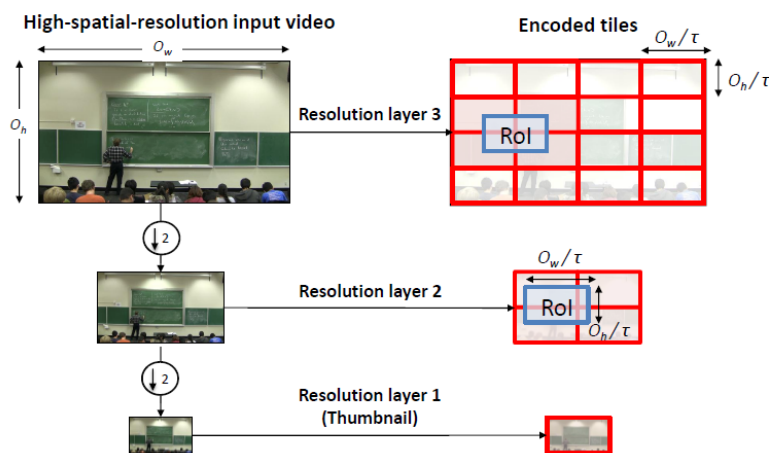


Fig. 4. Proposed coding scheme shown for the case of three resolution layers. The original high-spatial-resolution video is dyadically downsampled to create multiple resolution layers. Each resolution layer is further subdivided into tiles. Each tile is independently coded using an H.264/AVC video encoder.

Ex. 1105 at Fig. 4; Ex. 1102 at 238. The *Mavlankar Publication* further disclosed transmitted the requested high-resolution video tile to a client. Ex. 1105 at Fig. 4; Ex. 1102 at 238. The Examiner determined that the *Mavlankar Publication* did not explicitly disclose transmitting a notification to a client including spatial information

regarding the tiles, but found such a notification in analogous art of *Koyanagi*. Ex. 1102 at 239.

In response, Patent Owner did not contest that the *Mavlankar Publication* in view of *Koyanagi* rendered obvious the original claims. Instead, Patent Owner amended the claims to recite that the individual tiles were identified in the manifest file by a pixel position on a frame of the video. *Id.* at 303. Following an interview, the Examiner issued a notice of allowance including an Examiner’s Amendment. *Id.* at 325-31. In the Examiner’s Amendment, the Examiner specified that “the pixel position of each of the n partial videos is represented by a value according to a number of pixels between a reference pixel position on the frame and a pixel position in a top-left corner of each of the n partial videos in an x-axis direction and a y-axis direction.” *Id.* at 326. Thus, patentability of the claims of the ’952 patent hinged on the specific representation of the pixel position in the notification by a “reference pixel position” and a “pixel position in the top-left corner of the video.” But, as explained below, *Irie*—which was not before the Examiner during the original prosecution of the ’952 patent—expressly teaches this functionality.

D. Person Of Ordinary Skill In The Art

A person having ordinary skill in the art at the time of the ’952 patent would have been a person having a Bachelor’s degree in Electrical Engineering or Computer Science degree with at least two years of academic or industry experience

in video encoding or processing. Ex. 1101 ¶ 31. Additional education might compensate for a deficiency in experience, and vice-versa.

VII. CLAIM CONSTRUCTION

For the purposes of this *inter partes* review only, the challenged claims should be given their plain and ordinary meaning in light of the specification. *Phillips v. AWH Corp.*, 415 F.3d 1303 (Fed. Cir. 2005).

VIII. OVERVIEW OF THE PRIOR ART

A. Overview Of *Irie*

Irie describes a system for transmitting video between a source and a client that allows the client to “easily enlarge and display the video images distributed” by the source. Ex. 1103 at Summary. First, *Irie* teaches the transmission of video data from a source to a client with a lower resolution capability. *Id.* ¶ 10. The source “divides a frame of the video into multiple blocks” and “assigns identification data to the multiple blocks to identify each of them.” *Id.* This “video division information include[s] identification data of each block and information on the position of each block in the frame.” *Id.* The source sends the “video division information” to “the content reception apparatus.” *Id.* ¶¶ 49, 51. The client chooses a Region Of Interest, *id.* ¶ 17, and requests that video data. *See id.* ¶¶ 52, 53, 56, 62, 64, 66, 67. Then the server transmits the requested video to the client. *See id.* ¶¶ 56, 67, 68, 69. Second, *Irie* teaches the same system from the perspective of the

client device. *See id.* ¶ 12. *Irie* explains the “content receiving device can easily enlarge a part of a video image using video images multicasted from a content delivery device.” *Id.* ¶ 13.

Like the ’952 patent, *Irie* addresses the problem of transmitting high-resolution video data to a client device that either cannot handle the resolution or does not want to use the data required for a full high-resolution data stream. *Irie* describes the exact same “solution” that the ’952 patent contends is novel—namely, the use of identification information for the tiles specifying the pixel position of each of the tiles in a frame of the video. *See, e.g., id.* ¶ 54.

Irie was not in front of the Examiner during prosecution of the ’952 patent. Patent Owner’s assertion during prosecution that the claims were patentable because they claimed a notification including a pixel position (and the Examiner’s Amendment regarding the pixel position), is incorrect.

B. Overview of *Van Deventer*

Van Deventer describes methods and systems for processing “spatially segmented content originating from a content delivery network.” *See* Ex. 1104 at Abstract. One embodiment “comprises a client in a media processing device receiving a spatial manifest information comprising one or more spatial representations of a source stream, each spatial representation identifying one or more spatial segment streams, location information for locating one or more delivery

nodes in said content delivery network, and optionally, position information for stitching spatial segment frames in said segment streams into a video frame for display.” *Id.* *Van Deventer* provides significant detail as to the structure of its spatial manifest file, including by reference to examples showing different spatial representations and segments:

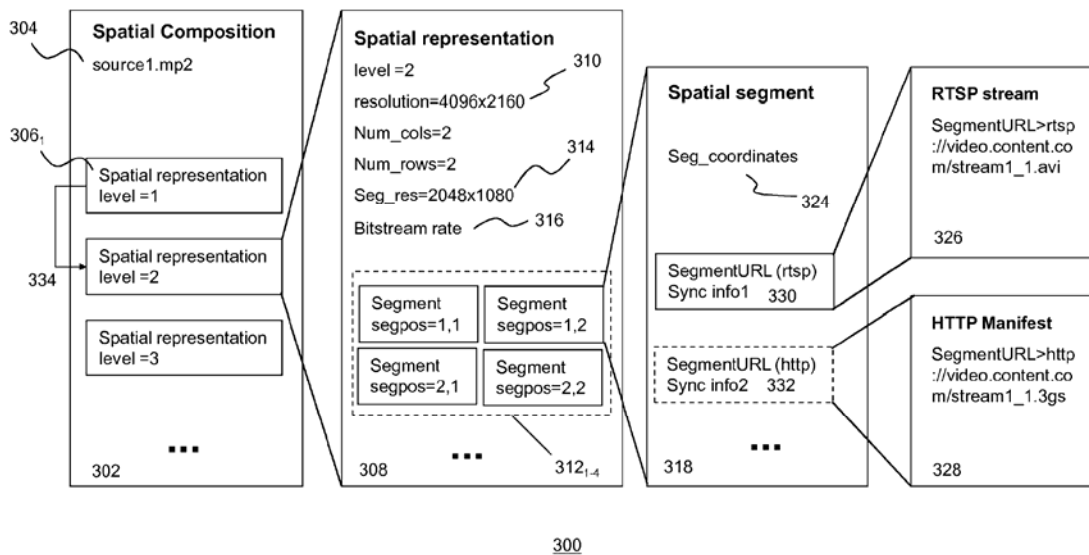


Figure 3

Id. at Fig. 3.

Like the '952 patent, *Van Deventer* discloses a method of “receiving user input associated with a region of interest, said region of interest defining a selected area of said displayed spatially-segmented content; on the basis of positioning information in said spatial segment data structure and said region of interest determining a second spatial representation in accordance with one or more media processing criteria, preferably said processing criteria including at least one of

bandwidth, screen resolution and/or encoding quality.” *Id.* at 3:67-4:9. *Van Deventer* further discloses “a user selecting a particular area 512 (the region of interest or ROI) of the display 514 as schematically depicted in Fig. 5B. On the basis of the size and coordinates of the selected ROI, the control engine may select an appropriate spatial representation from the SMF.” *Id.* at 12:63-67.

Van Deventer was cited on the face of the ’952 patent, but was not substantively discussed during prosecution.

IX. SPECIFIC GROUND FOR PETITION

A. Ground I: Claims 7 and 11-14 Are Rendered Obvious By Irie In View Of Van Deventer

For the reasons explained below, claims 7 and 11-14 are obvious in view of *Irie* and *Van Deventer*. The ground presented in this Petition is not redundant of the grounds presented in IPR2020-00656. Although IPR2020-00656 concerned the same ’952 patent, the only claims challenged here—claims 7 and 11-14—were not challenged in IPR2020-00656. Moreover, IPR2020-00656 did not include the *Van Deventer* reference at issue here.

1. Motivation To Combine *Irie* With *Van Deventer*

A POSITA would have been motivated to combine *Irie* and *Van Deventer* and would have had a reasonable expectation of success in doing so. Ex. 1101 ¶¶ 133-138. As an initial matter, *Van Deventer* cites *Irie* on the face of the patent, and thus a POSITA would have known of and been motivated to combine *Irie* with *Van*

Deventer. See Ex. 1104 at 2 (citing JP 2009-212821). Moreover, *Irie* was cited in the prosecution of the Japanese counterpart application (JP Pat. Appln. No. 2014-415071) (Ex. 1107 at 2).

Irie and *Van Deventer* describe complementary methods for streaming media data. *Irie* is directed to a multicast system that transmits video between a source and a client that allows a user to select a particular region to be enlarged. For example, *Irie* discloses “a content delivery device which delivers a video image of which resolution is higher than that of a content receiving device, divides a frame of the video image into multiple blocks, assigns identification data to the multiple blocks to identify each of them, sends video division information including identification data of each block and information on the position of the block in the frame to the content receiving device, and sends a multicast stream of the video of each block with the identification data.” Ex. 1103 ¶ 10. *Irie* explains that its system applies to “IP multicast video delivery systems.” *Id.* ¶ 1. *Van Deventer* likewise relates to spatially-segmented content delivery and “a content delivery network configured to distribute spatially segmented content.” Ex. 1004 at 1:15-19. Moreover, *Van Deventer* discloses a method for partitioning and streaming video data, including a Region of Interest, focusing on the “spatial manifest file” transmitted to the client. For example, *Van Deventer* discloses “spatially-segmented content delivery” to improve the transmission of “high-resolution video content.” *Id.* at 2:33-40. *Van*

Deventer also discloses a method of “receiving user input associated with a region of interest, said region of interest defining a selected area of said displayed spatially-segmented content; on the basis of positioning information in said spatial segment data structure and said region of interest determining a second spatial representation in accordance with one or more media processing criteria, preferably said processing criteria including at least one of bandwidth, screen resolution and/or encoding quality.” *Id.* at 3:67-4:8.

While *Irie* is primarily concerned with ensuring that users can enlarge videos from content delivery services without reducing the resolution, *Van Deventer* is focused on improving the ability for content-delivery networks to provide information about the different resolutions of the video partitions to users. *See* Ex. 1103 at ¶ 5; Ex. 1104 at 2:33-40. As Dr. Almeroth explains, a POSITA would recognize that *Irie* is not limited to the specific notification method disclosed but could be used with other types of notifications, such as the one described by *Van Deventer*. Ex. 1101 ¶¶ 76-78. The spatial manifest file described by *Van Deventer* could be used to carry the information discussed in the notifications disclosed in *Irie*. *See id.* A POSITA would also have appreciated that using the detailed spatial manifest file provided in *Van Deventer* to send the video-display information from *Irie* to accommodate different types of content-delivery networks, as *Van Deventer* explicitly explains. *See* Ex. 1104 at 2:32-40, 2:63-67. Accordingly, a POSITA

would have had a reasonable expectation of success in combining *Irie* and *Van Deventer*, as the two systems are complementary and compatible. *Id.*

The alleged invention of the '952 patent would have been obvious to one having ordinary skill in the art in view of the combined teachings of *Irie* and *Van Deventer*. These references disclose a method for streaming data by “receiving from the client which received the notification by the server, one or more messages for requesting one or more partial videos according to a Region Of Interest selected based on an operation by a user of the client.” (Ex. 1100, claim 1). For example, a person of ordinary skill in the art would have been motivated to apply the techniques described in the *Van Deventer* to the system disclosed in *Irie*. The common subject matter and similar approaches to similar problems conceived and developed in a similar timeframe would have motivated one having ordinary skill in the art to combine the teachings of *Irie* and *Van Deventer*.

2. Claim 3

Claims 7 and 11-14 of the '952 Patent all depend on claim 3. Accordingly, Petitioners first set forth why claim 3 is invalid over *Irie* in combination with *Van Deventer*.

a. Preamble

Claim 3 of the '952 patent starts with the following preamble: “A transmitting method for video data to a client, the method comprising.” Ex. 1100, claim 3. *Irie*

discloses a method for transmitting video data to a client. *See, e.g.*, Ex. 1103 ¶ 10. Therefore, *Irie* discloses the preamble of claim 1.

b. Element 3(a)

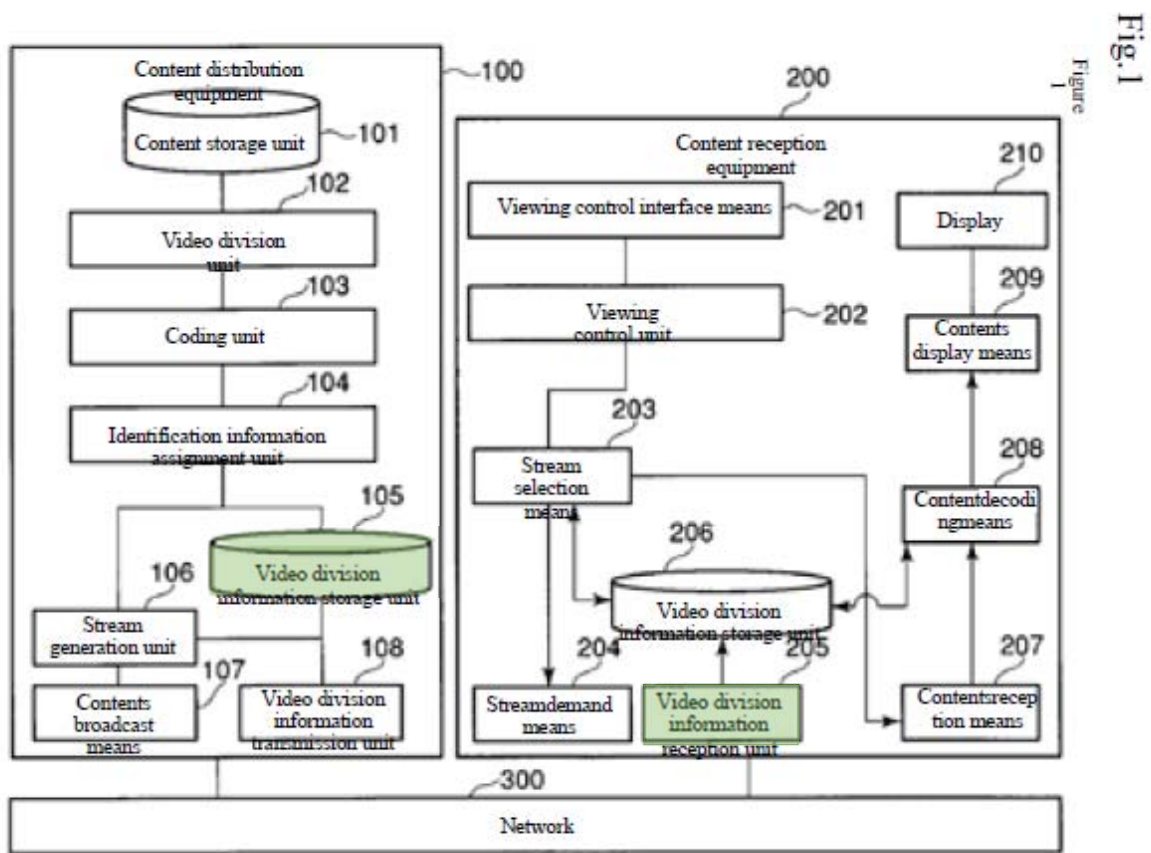
The first limitation of claim 3 reads: “transmitting, to the client, a notification including (i) spatial position information representing, by a pixel position on a frame of a video, a spatial position of each of n spatial regions of the video, n being an integer, and (ii) spatial size information representing a spatial size of each of the n spatial regions of the video.” Ex. 1100, claim 3. *Irie* discloses transmitting to a client a notification that includes the spatial position information represented by a pixel position and spatial size information. For example, *Irie* explains that it “assigns identification data to the multiple blocks to identify each of them” and “sends video division information including identification data of each block and information on the position of the block in the frame to the content receiving device.” Ex. 1103 ¶ 10.

Irie also explains that a frame of the video image can be split into blocks in multiple different ways:

Figure 2 shows division of a frame of the video image into multiple video blocks by the video dividing part 102 of the content delivery device 100 (Step S1). There is no restriction on sizes of video blocks. A frame can be divided into macro blocks, which is the minimum unit of encoding, and 1 video block can be 1 macro block or multiple macro blocks (a group of macro blocks) (for example, 160 x 160 pixels). A frame can also be divided into pixels, and 1 image block can be 1 pixel (1 x 1 pixel) or a group of multiple pixels.

Id. ¶ 21.

The content delivery device 100 shown in Figure 1 of *Irie* below contains a “video division information storage 105” and a “video division information receiving part 205,” highlighted in green below:



Id. at Fig. 1 (annotated), ¶ 19.

This information includes, among other things, “[p]osition information of each video block (x, y coordinates).” *Id.* ¶¶ 36-37. This position information can

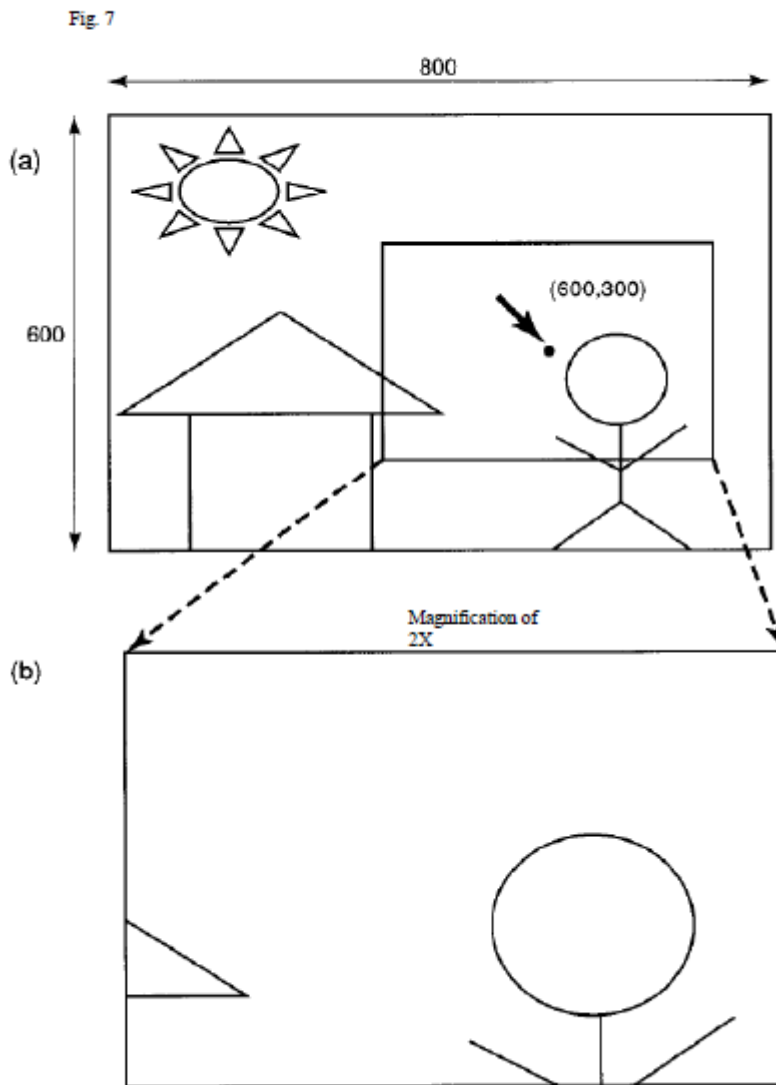
be represented in terms of pixels. *See, e.g., id.* ¶ 54. Therefore, *Irie* discloses the first element of claim 3.

c. Element 3(b)

The second limitation of claim 3 reads: “receiving, from the client which received the notification, information relating to identification of one or more spatial regions from among the n spatial regions defined in the notification, the one or more spatial regions corresponding to video data to be transmitted to the client.” Ex. 1100, claim 3. *Irie* discloses receiving from the client which received the notification, information relating to the identification of one or more spatial regions according to a Region Of Interest selected based on an operation by a user of the client. Video division information is transmitted to the content receiving device, or client, and stored at the client device. Ex. 1103 ¶ 51. When a user at the client selects a desired position on the display for interest, the position information is detected for viewing a different resolution image of that region of interest. *See, e.g., id.* ¶¶ 53-57.

Irie illustrates this functionality in Figure 7, where a specific region of interest is selected in Figure 7(a) and magnified using a higher-resolution image in Figure 7(b):

Fig. 7

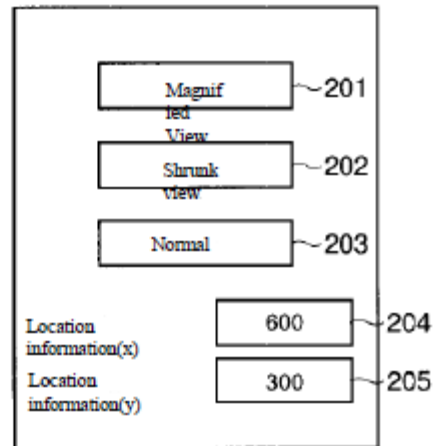


Id. at Fig. 7; *see also id.* ¶¶ 72-75.

Irie also shows a dialog box that provides position information for the region of interest and options for the user to view other resolution versions of that particular region of interest:

Fig. 8

Fig. 8



Id. at Fig. 8.

Therefore, *Irie* discloses the second element of claim 3.

d. Element 3(c)

The third limitation of claim 3 reads: “transmitting, to the client, the video data corresponding to the one or more spatial regions which are identified according to the information received from the client.” Ex. 1100, claim 3. *Irie* discloses transmitting to the client video data corresponding to the one or more spatial regions, in response to the one or more messages from the client. For example, Figures 7 and 8 of *Irie* show the client user interface for selecting a particular region of interest:

Fig. 7

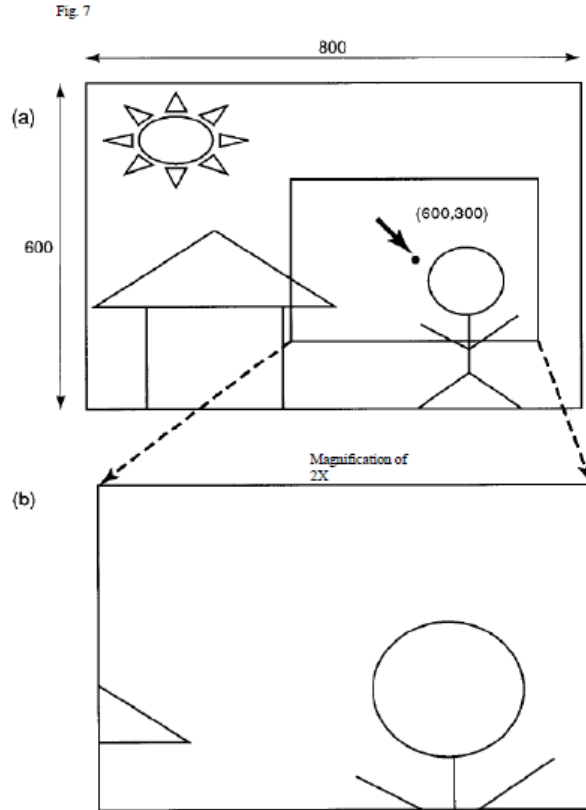
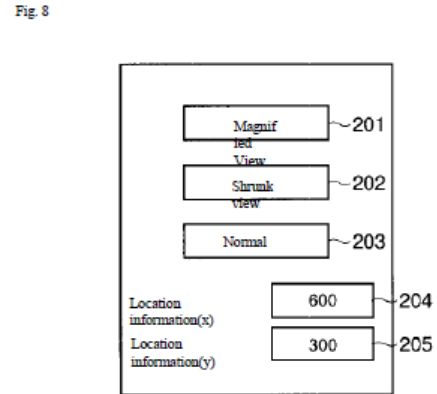
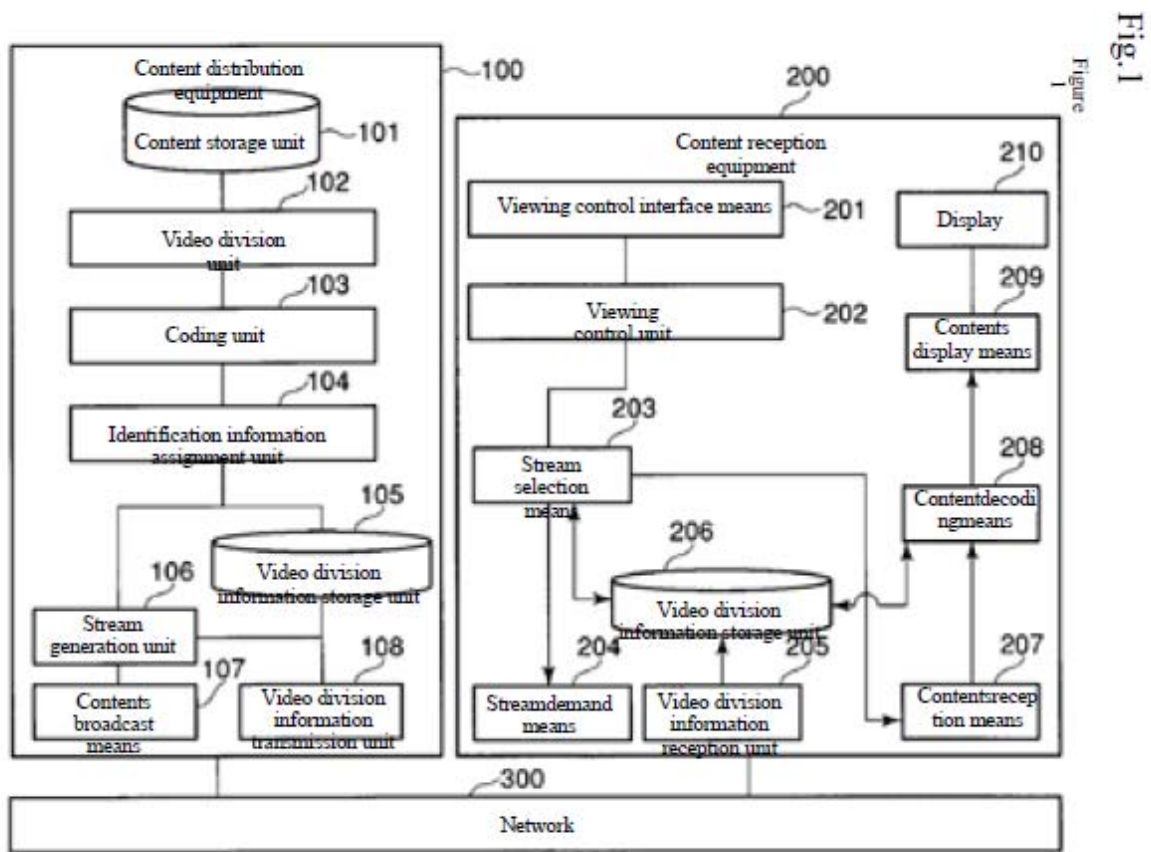


Fig. 8



Ex. 1103 at Figs. 7, 8.

Irie also explains that the client selects a particular region of interest—here, identified as the position (600, 300)—and selects a particular action to take, such as magnified view or shrunk view. *Id.* at Fig. 8, ¶¶ 53-55. This information is used by the client—the content receiving device—to request a particular stream and resolution:



Id. at Fig. 1, ¶ 19

Irie in paragraph 19 describes that the content receiving device 200 contains the viewing control interface part 201, a viewing control part 202, a stream selection part 203, a stream request part 204, a video division information receiving part 205, a content receiving part 207, a content decoding part 208, a content displaying part 210, and a display 210. It also describes that a request is transmitted to the content distribution device to retrieve the selected stream. *Id.* ¶ 67.

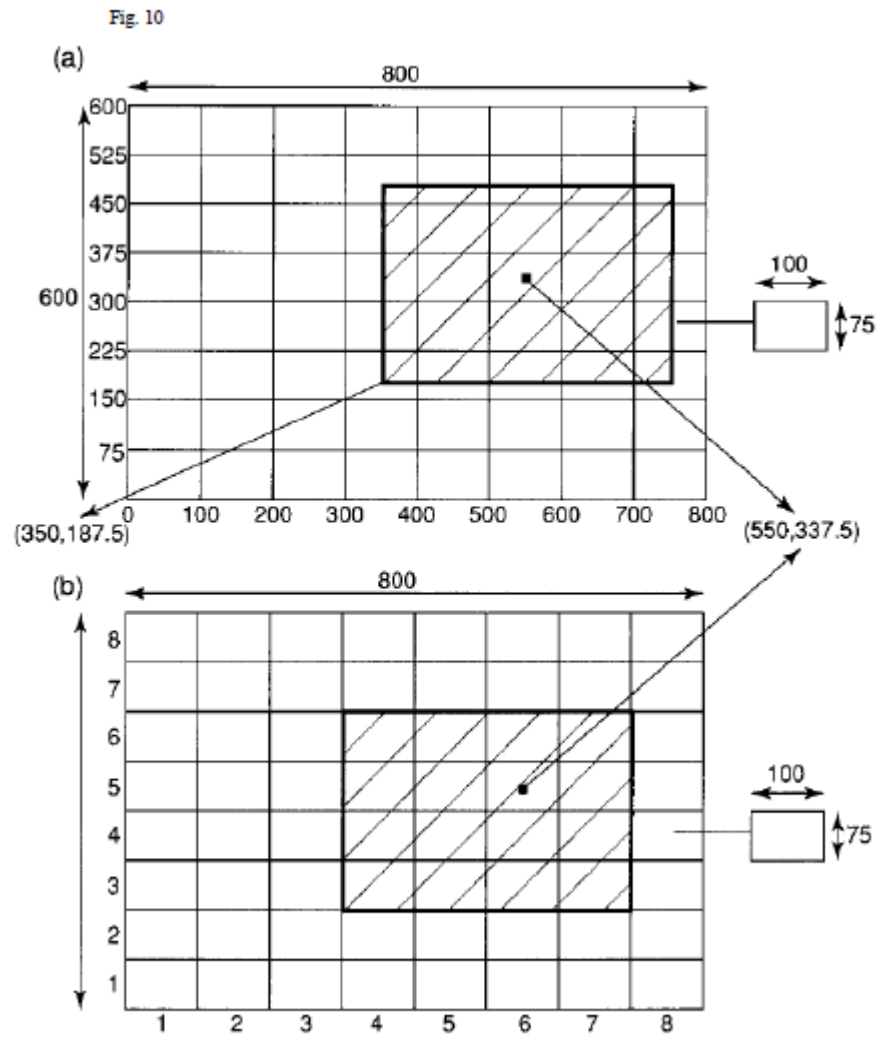
Furthermore, Figure 7 shows that the requested stream is transmitted by the server to the client. *Id.* at Fig. 7; *see also id.* ¶¶ 67-70, 80, 97, 98 (“At step 6, content distribution unit 107 transmits the multicast stream generated at step 5.”), 99.

Therefore, *Irie* discloses the third element of claim 3.

e. Element 3(d)

The fourth limitation of claim 3 reads: “wherein the pixel position of each of the n spatial regions is represented by a value according to a number of pixels between a reference pixel position and a pixel position in a top-left corner of each of the n spatial regions in an x-axis direction and a v-axis direction.” Ex. 1100, claim 3. *Irie* discloses this claim limitation. As noted above for claim element 3(a) *supra*, the video division information sent to the client includes, among other things, “[p]osition information of each video block (x, y coordinates),” Ex. 1103 ¶¶ 36-37, that can be represented in terms of pixels. *See, e.g., id.* ¶¶ 36-37, 54. This is illustrated in Figure 10 of *Irie*, which shows individual blocks identified by a pixel position represented by a value according to a number of pixels between a reference pixel position on the frame and a pixel position in a top-left corner of each of the n partial videos in an x-axis direction and a y-axis direction:

Fig. 10



Id. at Fig. 10; *see also id.* ¶¶ 76-77.

Therefore, *Irie* discloses the fourth element of claim 3.

3. Claim 7

Claim 7, which depends only on claim 3, requires “a spatial size of at least one spatial region among the *n* spatial region of the frame of the video is different

from a spatial size of another spatial region among the n spatial region of the frame of the video, and the spatial size information represents a both of the spatial size of the at least one spatial region and the spatial size of said another spatial region.” As explained above, *Irie* discloses all the elements of claim 3. Claim 7 does not add anything patentably distinct to claim 3, as a POSITA would understand that the spatial segments sent in *Irie* can be of different sizes. *See* Ex. 1101 at ¶¶ 100-103. Nothing in *Irie* mandates that they must be the same size.

Nonetheless, *Irie* in view of *Van Deventer* discloses this claim. *Irie* explains that there are “no restriction on the sizes of the video blocks. A frame can be divided into macro blocks, which is the minimum unit of encoding, and 1 video block can be 1 macro block or multiple macro blocks (a group of macro blocks) (for example, 160 x 160 pixels). A frame can also be divided into pixels, and 1 image block can be 1 pixel (1 x 1 pixel) or a group of multiple pixels.” Ex. 1103 ¶ 21; *see also id.* at Fig. 2. Accordingly, *Irie* is compatible with and can be used with a system in which spatial regions are of different sizes. *Van Deventer* specifically explains that “the spatial division of image frames into a segment group is not limited to those schematically depicted in Fig. 2A. An image frame may be spatially divided into a segment group defining a matrix of spatial segment frames of equal dimension. Alternatively, ***in other embodiment[s], spatial segment frames in a segment group may have different segment frame dimensions, e.g. small segment frames***

associated with the center of the original image frame and larger sized spatial segment frames at the edges of an image frame.” Ex. 1104 at 7:58-67 (emphasis added). Moreover, *Van Deventer* explains that its “spatial manifest data structure (file)” —which is transmitted to a client so that the client can identify the various spatial representations of a video stream (*see id.* at Abstract)—may comprise “positioning information” for each of the spatial frames into which the video has been divided. *See id.* at 3:1-23. *Van Deventer* explains how the spatial manifest file includes this positioning information for each of the spatial segments (which may be small or large regions):

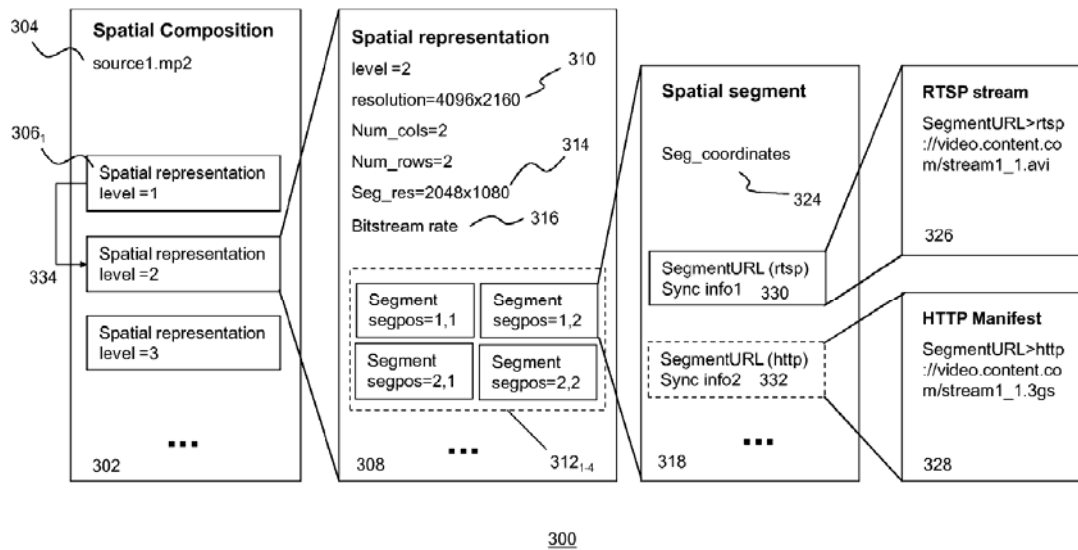


Figure 3

Id. at Fig. 3; *see also id.* at 10:57-67.

Therefore, for the reasons discussed above, a POSITA would have been motivated to use the information disclosed in *Van Deventer*'s notifications in combination with *Irie* to disclose the requirements of claim 7.

4. Claim 11

Claim 11, which depends only on claim 3, requires “the information received from the client includes one or more URLs associated with the one or more spatial regions.” *Irie* in view of *Van Deventer* discloses this claim. *Irie* explains that the “video division information including identification data assigned to respective blocks obtained by dividing a frame of a video image” is sent from the client. Ex. 1103 ¶ 12. In certain cases, “a multicast address and a port number are used as address information.” *Id.* ¶ 46; *see also id.* at Fig. 5. “When generating a multicast stream of each hierarchical video block, the address information in the video division information corresponding to the hierarchical video block is acquired from video division information storage unit 105, and the acquired multicast address and port number are used as the destination address to generate a multicast stream.” *Id.* at ¶ 47; *see also id.* at ¶¶ 48-49, 67, 82, 99, 107, and 118. A POSITA would thus have known that information from the client like the multicast address would include “one or more URLs associated with the one or more spatial regions.” Ex. 1100, claim 11; Ex. 1101 ¶¶ 104-105.

Furthermore, *Van Deventer* discloses receiving “one or more URLs associated with the one or more spatial regions.” Ex. 1100, claim 11. For example, it explains that “[t]he information establishing the spatial and/or temporal relation between spatial segment streams in each spatial representation and the location information, e.g. URLs, for retrieving the spatial segment streams may be hierarchically ordered and stored as a spatial manifest file (SMF).” Ex. 1104 at 8:7-11. It further explains that “[i]n order to ingest spatially segmented content, a CDNCF 212 may receive a SMF from the content processing entity wherein the SMF comprises location information, e.g. URLs, where the spatially segmented content associated with the content processing entity is stored. Based on the location information in the SMF, the CDNCF may then initiate transport of spatial segment streams to the ingestion node, which may temporarily store or buffer the ingested content in an ingestion cache 214.” *Id.* at 8:23-31; *see also e.g., id.* at 6:54-60, 8:52-58, 8:65-0:2, 9:5-11, 10:19-26, 12:51-60, Fig. 2C, 3, and 4. Figures 3, 4, 6A, and 6B further disclose the use of RTSP-based URLs to provide streams to the client. For the reasons discussed above, a POSITA would have been motivated to use the information disclosed in *Van Deventer*’s notifications in combination with *Irie* to disclose the requirements of claim 11.

5. Claim 12

Claim 12, which depends only on claim 3, requires “the information received from the client includes one or more HTTP addresses associated with the one or more spatial regions.” *Irie* in view of *Van Deventer* discloses this claim. *Irie* explains that video-on-demand multicast delivery systems uses IP communication protocols to transfer content by “RTP (Real-Time Transport Protocol) or HTTP (Hyper Text Transfer Protocol).” Ex. 1103 ¶¶ 4-5. Furthermore, *Irie* explains that the “video division information including identification data assigned to respective blocks obtained by dividing a frame of a video image” is sent from the client. Ex. 1103 ¶ 12. In certain cases, “a multicast address and a port number are used as address information.” *Id.* ¶ 46; *see also id.* at Fig. 5. “When generating a multicast stream of each hierarchical video block, the address information in the video division information corresponding to the hierarchical video block is acquired from video division information storage unit 105, and the acquired multicast address and port number are used as the destination address to generate a multicast stream.” *Id.* at ¶ 47; *see also id.* at ¶¶ 48-49, 67, 82, 99, 107, and 118.

Similarly, *Van Deventer* discloses “[s]patial segment information may further include one or more spatial segment stream locators 326,328 (e.g. one or more URLs) for locating delivery nodes, which are configured to transmit the particular spatial segment stream to the client. The spatial segment information may also comprise protocol information, indicating which protocol (e.g. RTSP or HTTP) is

used for controlling the transmission of the streams to the client, and synchronization information 330,332 for a synchronization module in the client to synchronize different segment streams received by a client.” Ex. 1104 at 11:28-37; *see also e.g., id.* at 6:65-7:7, 9:36-46, 9:61-10:4, 12:51-60, 14:60-15:3, 18:48-55, 19:13-21, Fig. 3 and 10A. *Van Deventer* also explains that “various streaming protocols may be used to deliver the content to the client. Such protocols may include HTTP and RTP type streaming protocols. In a preferred embodiment an adaptive streaming protocol such as HTTP adaptive streaming (HAS), DVB adaptive streaming, DTG adaptive streaming, MPEG DASH, ATIS adaptive streaming, IETF HTTP Live streaming and related protocols may be used.” *Id.* at 6:67-7:7.

Accordingly, a POSITA would have known that when transferring content by HTTP, the content metadata would include information related to “one or more HTTP addresses associated with the one or more spatial regions.” Ex. 1100, claim 12; Ex. 1101 ¶¶ 107-109. For the reasons discussed above, a POSITA would have been motivated to use the information disclosed in *Van Deventer*’s notifications in combination with *Irie* to disclose the requirements of claim 12.

6. Claim 13

Claim 13, which depends on claim 3, discloses that “in a case where a spatial position of the n spatial regions for a first time period and a spatial position of the n spatial regions for a second time period are different from each other, a first

notification for the first time period and a second notification for the second time period are transmitted to the client.” *Irie* in view of *Van Deventer* discloses this claim. *Irie* explains that it “assigns identification data to the multiple blocks to identify each of them” and “sends video division information including identification data of each block and information on the position of the block in the frame to the content receiving device.” Ex. 1103¶ 10. These notification in *Irie* correspond each time period of the video frame transmitted to the content receiving device.

The content delivery device 100 shown in Figure 1 of *Irie* contains a “video division information storage 105” and a “video division information receiving part 205.” *Id.* ¶ 19. This information includes, among other things, “[p]osition information of each video block (x, y coordinates).” *Id.* ¶¶ 36-37. This position information can be represented in terms of pixels. *See, e.g., id.* ¶ 54.

Van Deventer explains that multiple spatial representations can be sent in spatial manifest files to define different video streams that are transmitted to the client at different times. *See* Ex. 1104 at 8:11-13, 12:31-13:9, 13:15-18. For example, *Van Deventer* explains that spatial segments in video streams will change over time, and must be identified accordingly: “[s]ynchronization information may comprise timing information in the headers of the packets in a spatial segment stream. Such timing information may include time stamps, e.g. RTP time stamps,

indicating the relative position of an image frame in a stream. Timing information may also relate to RTP time stamp information exchanged between an RTSP server and a client. In further embodiment, timing information may relate to time segments, e.g. DASH time segments specified in a HTTP DASH manifest file and time stamped frames (e.g. MPEG) or sequentially numbered frames (e.g. AVI) contained in such DASH time segment.” *Id.* at 17:59-18:2; *see also e.g., id.* at 19:62-20:13, Fig. 10B. Moreover, *Van Deventer* also describes a first spatial manifest file and an “updated” spatial manifest file that can be sent to a client from the network after the network device receives updated information regarding the spatial segments. *See id.* at 10:19-26. Thus, as in *Irie*, each instance of spatial information sent to the client in *Van Deventer* corresponds to its respective time period.

For the reasons discussed above, a POSITA would have been motivated to use the information disclosed in *Van Deventer*’s notifications in combination with *Irie* to disclose the requirements of claim 13.

7. Claim 14

Claim 14, which depends on claim 13, discloses “the first notification includes (i) first spatial position information representing a spatial position of each of the n spatial regions for the first time period and (ii) first spatial size information representing a spatial size of each of the n spatial regions for the first time period, and the second notification includes (iii) second spatial position information

representing a spatial position of each of the *n* spatial regions for the second time period and (iv) second spatial size information representing a spatial size of each of the *n* spatial regions for the second time period.” *Irie* in view of *Van Deventer* discloses this claim. As explained in Section IX.A.6 *supra*, *Irie* in view of *Van Deventer* discloses transmitting multiple notifications to the client for different time periods. *Irie* and *Van Deventer* discloses transmitting to the client multiple notifications that include spatial location information and spatial size information. For example, *Irie* describes “[a] content delivery device which delivers a video image of which resolution is higher than that of a content receiving device divides a frame of the video image into multiple blocks, assigns identification data to the multiple blocks to identify each of them, sends video division information including identification data of each block and information on the position of the block in the frame to the content receiving device, and sends a multicast stream of the video of each block with the identification data.” Ex. 1103 ¶ 10. Furthermore, the notifications sent to the client includes “video block size [and] position information of each video block (x, y coordinates).” *See id.* ¶¶ 36-37; *see also* Ex. 1103 at Fig. 4, 5, 6, and 9. As noted above in Section IX.A.6, each instance of spatial size information and spatial position information in *Irie* corresponds to its respective time period of the video frame.

Similarly, as explained above in Section IX.A.6, *Van Deventer* also discloses transmitting spatial position information and spatial size information to the client for multiple time periods. Therefore, for the reasons discussed above, a POSITA would have been motivated to use the information disclosed in *Van Deventer*'s notifications in combination with *Irie* to disclose the requirements of claim 14.

X. CONCLUSION

As set forth in the foregoing and in the supporting declaration of Dr. Almeroth (Ex. 1101), the challenged claims of the '952 patent are rendered obvious by the prior art references cited in this Petition. Accordingly, Petitioners request institution of an *inter partes* review to cancel those claims.

Respectfully submitted,

/s/ Reza Dokhanchy

Reza Dokhanchy (Reg. No. 62,795)
Adam R. Alper (to seek *pro hac vice*)
Akshay S. Deoras (to seek *pro hac vice*)
KIRKLAND & ELLIS LLP
555 California Street
San Francisco, California 94104
Telephone: (415) 439-1400
Facsimile: (415) 439-1500
reza.dokhanchy@kirkland.com
aalper@kirkland.com
adeoras@kirkland.com

Michael W. De Vries (to seek *pro hac vice*)

KIRKLAND & ELLIS LLP
555 South Flower Street
Los Angeles, CA 90071
Telephone: 213-680-8400
Facsimile: 213-680-8500
michael.devries@kirkland.com

**Attorneys For Petitioners Avigilon
Corporation and Avigilon USA
Corporation**

CERTIFICATE OF COMPLIANCE

This Petition complies with the type-volume limitations as mandated in 37 C.F.R. § 42.24, totaling 7,667 words. Counsel has relied upon the word count feature provided by Microsoft Word.

/s/ Reza Dokhanchy
Reza Dokhanchy

CERTIFICATE OF SERVICE

The undersigned hereby certifies that a copy of the foregoing Petition for *Inter Partes* Review of U.S. Patent No. 10,135,952, along with all exhibits and other supporting documents, was served on April 24, 2020 to the assignee for the patent and counsel:

Canon U.S.A. Inc. Intellectual Property Division Via Fed-Ex
15975 Alton Parkway
Irvine CA 92618-3731

Joseph A. Calvaruso (Reg. No. 28,287) Via Email
Richard F. Martinelli (Reg. No. 52,003)
Orrick, Herrington & Sutcliffe LLP
51 West 52nd Street
New York, NY 10019-6142
Telephone: (212) 506-5140
Fax: (212) 506-5151
JVCPTABDocket@orrick.com
RFMPTABDocket@orrick.com
Canon-Avigilon_OHS@orrick.com

/s/ Reza Dokhanchy
Reza Dokhanchy