

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

DECLARATION OF Dr. Kevin C. Almeroth UNDER 37 C.F.R. § 1.68
IN SUPPORT OF PETITION FOR *INTER PARTES* REVIEW OF
U.S. PATENT NO. 10,135,952

TABLE OF CONTENTS

I.	Introduction.....	3
II.	Qualifications	5
III.	Understanding of Patent Law.....	12
IV.	Person of Ordinary Skill in the Art	15
V.	Background and Summary of the '952 Patent	15
	A. Background of the Field Relevant to the '952 Patent	15
	B. Summary of the '952 Patent.....	19
	C. The '952 Patent Claims	20
	D. The '952 Patent Specification	22
	E. The '952 Patent Prosecution History	27
VI.	Claim Construction	30
VII.	Overview of the Prior Art.....	30
	A. <i>Irie</i>	30
	B. <i>Van Deventer</i>	32
VIII.	Claims 1-6, 8-10, and 15-21 Are Unpatentable in View of the Prior Art	33
	A. Ground I: Claims 7 and 11-14 are Obvious Over <i>Irie</i> in view of <i>Van Deventer</i>	34
IX.	Secondary Considerations of Non-Obviousness	56
X.	Conclusion	57

I, Dr. Kevin Almeroth, do hereby declare as follows:

I. INTRODUCTION

1. I have been retained on behalf of Avigilon USA Corporation and Avigilon Corporation (collectively, “Avigilon”) for the above-captioned Petition for *Inter Partes* Review (“IPR”) of U.S. Patent No. 10,135,952 (“the ’952 patent”). I am being compensated for my time in connection with this IPR at my standard consulting rate of \$700 per hour. My compensation is not affected by the outcome of this matter.

2. I have been asked to provide my opinions regarding whether claims 7 and 11-14 (the “Challenged Claims”) of the ’952 patent are invalid because they would have been obvious to a person having ordinary skill in the art at the time of the alleged invention.

3. 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).

4. The face of the '952 patent names Franck Denoual, Hervé Le Floch, Frédéric Maze, Jean Le Feuvre, and Cyril Concolato as the inventors and identifies Canon Kabushiki Kaisha as the assignee of the patent. Ex. 1100 at Cover.

5. In preparing this Declaration, I have reviewed the '952 patent and its file history, numerous prior art references, and technical references from the time of the alleged invention.

6. In forming the opinions expressed in this Declaration, I relied upon my education and experience in the relevant field of the art, and have considered the viewpoint of a person having ordinary skill in the relevant art, as of April 16, 2013, the effective priority date of the '952 patent. My analysis does not differ if the '952 patent has an effective priority date of October 12, 2012. My opinions are based, at least in part, on the following:

Reference	Date of Public Availability
<i>Irie</i> - Japanese published unexamined patent application No. 2009-212821 to Yuji Irie, <i>et al.</i> , (" <i>Irie</i> ")	<i>Irie</i> was published on September 17, 2009, and is attached as Ex. 1103 to the IPR.
<i>Van Deventer</i> – U.S. Patent No. 9,860,572 to Mattijs Oskar van Deventer, <i>et al.</i> (" <i>Van Deventer</i> ")	<i>Van Deventer</i> was filed as PCT/EP2012/060801 on June 7, 2012, claiming foreign application priority to EP 11169068 filed on June 8, 2011. It is attached as Ex. 1104 to the IPR.

II. QUALIFICATIONS

7. My curriculum vitae (“CV”) is attached hereto as Exhibit A and provides an accurate identification of my background and experience.

8. I am currently a Professor in the Department of Computer Science at the University of California, Santa Barbara (UCSB). At UCSB, I also hold faculty appointments and am a founding member of the Computer Engineering (CE) Program, Media Arts and Technology (MAT) Program, and the Technology Management Program (TMP). I also served as the Associate Director of the Center for Information Technology and Society (CITS) from 1999 to 2012. I have been a faculty member at UCSB since July 1997.

9. I hold three degrees from the Georgia Institute of Technology: (1) a Bachelor of Science degree in Information and Computer Science (with minors in Economics, Technical Communication, and American Literature) earned in June 1992; (2) a Master of Science degree in Computer Science (with specialization in Networking and Systems) earned in June 1994; and (3) a Doctor of Philosophy (Ph.D.) degree in Computer Science (Dissertation Title: Networking and System Support for the Efficient, Scalable Delivery of Services in Interactive Multimedia System, minor in Telecommunications Public Policy) earned in June 1997. During my education, I have taken a wide variety of courses as demonstrated by my minor.

10. One of the major concentrations of my research has been the delivery of multimedia content and data between computing devices, including various network architectures. In my research, I have studied large-scale content delivery systems, and the use of servers located in a variety of geographic locations to provide scalable delivery to hundreds or thousands of users simultaneously. I have also studied smaller-scale content delivery systems in which content is exchanged between individual computers and portable devices. My work has emphasized the exchange of content more efficiently across computer networks, including the scalable delivery of content to many users, mobile computing, satellite networking, delivering content to mobile devices, and network support for data delivery in wireless networks.

11. An important component of my research has been investigating the challenges of communicating multimedia content, including video, between computers and across networks including the Internet. Although the early Internet was used mostly for text-based, non-real time applications, the interest in sharing multimedia content, such as video, quickly developed. Multimedia-based applications ranged from downloading content to a device to streaming multimedia content to be instantly used. One of the challenges was that multimedia content is typically larger than text-only content, but there are also opportunities to use different delivery techniques since multimedia content is more resilient to errors. I

have worked on a variety of research problems and used a number of systems that were developed to deliver multimedia content to users. One content-delivery method I have researched is the one-to-many communication facility called “multicast,” first deployed as the Multicast Backbone, a virtual overlay network supporting one-to-many communication. Multicast is one technique that can be used on the Internet to provide streaming media support for complex applications like video-on-demand, distance learning, distributed collaboration, distributed games, and large-scale wireless communication. The delivery of media through multicast often involves using Internet infrastructure, devices and protocols, including protocols for routing and TCP/IP.

12. As a parallel research theme, starting in 1997, I began researching issues related to wireless devices and sensors. In particular, I was interested in showing how to provide greater communication capability to “lightweight devices,” *i.e.*, small form-factor, resource-constrained (e.g., CPU, memory, networking, and power) devices. Starting in 1998, I published several papers on my work to develop a flexible, lightweight, battery-aware network protocol stack. The lightweight protocols we envisioned were similar in nature to protocols like Universal Plug and Play (UPnP) and Digital Living Network Alliance (DLNA).

13. Protecting networks, including their operation and content, has been an underlying theme of my research almost since the beginning. Starting in 2000, I

have also been involved in several projects that specifically address security, network protection, and firewalls. After significant background work, a team on which I was a member successfully submitted a \$4.3M grant proposal to the Army Research Office (ARO) at the Department of Defense to propose and develop a high-speed intrusion detection system. Once the grant was awarded, we spent several years developing and meeting the milestones of the project. I have also used firewalls in developing techniques for the classroom to ensure that students are not distracted by online content.

14. Between 2002 and 2004, I authored several papers directed to my research on distance learning systems and infrastructures that is particularly relevant to the technology of the '952 patent. I developed a digital classroom at UCSB to enable distance learning, realtime collaboration, and for lectures and presentations to be archived and reviewed at a later time. One goal of this research was to ensure distance learning students felt included in the lectures they attended remotely. To this end, the digital classroom I helped to develop utilized multiple cameras distributed across different locations in a room that, together, captured all of the classroom activity. All the cameras were connected to an electronic network so their video feeds could be combined and shared with remote students.

15. Employing such a system for distance learning required experimentation with a variety of different video encoders to select one that is most

suitable for livestreaming a lecture. Over the duration of the research, the cameras we used required periodic software updates to enable new features or render them compatible with new technologies. In particular, controlling and maintaining numerous, heterogeneous video capturing devices introduced a non-trivial bit of complexity that required developing a specialized digital classroom infrastructure. My work developing and programming this video and camera-based distance learning system spanned several years and is described in my refereed conference papers listed in my CV.

16. As an additional important component of my research program, I have been involved in the development of academic research into available technology in the market place. One aspect of this work is my involvement in the Internet Engineering Task Force (IETF). The IETF is a large and open international community of network designers, operators, vendors, and researchers concerned with the evolution of the Internet architecture and the smooth operation of the Internet. I have been involved in various IETF groups including many content delivery-related working groups like the Audio Video Transport (AVT) group, the MBone Deployment (MBONED) group, Source Specific Multicast (SSM) group, the Inter- Domain Multicast Routing (IDMR) group, the Reliable Multicast Transport (RMT) group, the Protocol Independent Multicast (PIM) group, etc. I have also served as a member of the Multicast Directorate (MADDOGS), which

oversaw the standardization of all things related to multicast in the IETF. Finally, I was the Chair of the Internet2 Multicast Working Group for seven years.

17. My involvement in the research community extends to leadership positions for several academic journals and conferences. I am the co-chair of the Steering Committee for the ACM Network and System Support for Digital Audio and Video (NOSSDAV) workshop and on the Steering Committees for the International Conference on Network Protocols (ICNP), ACM Sigcomm Workshop on Challenged Networks (CHANTS), and IEEE Global Internet (GI) Symposium. I have served or am serving on the Editorial Boards of IEEE/ACM Transactions on Networking, IEEE Transactions on Mobile Computing, IEEE Network, ACM Computers in Entertainment, AACE Journal of Interactive Learning Research (JILR), and ACM Computer Communications Review. I have co-chaired a number of conferences and workshops including the IEEE International Conference on Network Protocols (ICNP), IEEE Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON), International Conference on Communication Systems and Networks (COMSNETS), IFIP/IEEE International Conference on Management of Multimedia Networks and Services (MMNS), the International Workshop On Wireless Network Measurement (WiNMee), ACM Sigcomm Workshop on Challenged Networks (CHANTS), the Network Group

Communication (NGC) workshop, and the Global Internet Symposium, and I have served on the program committees for numerous conferences.

18. Furthermore, in the courses I teach at UCSB, a significant portion of my curriculum covers aspects of the Internet and network communication including the physical and data link layers of the Open System Interconnect (OSI) protocol stack, and standardized protocols for communicating across a variety of physical media such as cable systems, telephone lines, wireless, and high-speed Local Area Networks (LANs). The courses I have taught also cover most major topics in Internet communication, including data communication, multimedia encoding, and mobile application design. My research and courses cover a range of physical infrastructures for delivering content over networks, including cable, Integrated Services Digital Network (ISDN), Ethernet, Asynchronous Transfer Mode (ATM), fiber, and Digital Subscriber Line (DSL). For a complete list of courses I have taught, see my CV, attached as Exhibit A.

19. In addition, I co-founded a technology company called Santa Barbara Labs that was working under a sub-contract from the U.S. Air Force to develop very accurate emulation systems for the military's next generation internetwork. Santa Barbara Labs' focus was in developing an emulation platform to test the performance characteristics of the network architecture in the variety of environments in which it was expected to operate, and, in particular, for network

services including IPv6, multicast, Quality of Service (QoS), satellite-based communication, and security. Applications for this emulation program included communication of a variety of multimedia-based services, including video conferencing and video-on-demand.

20. In addition to having co-founded a technology company myself, I have worked for, consulted with, and collaborated with companies for nearly 30 years. These companies range from well-established companies to start-ups and include IBM, Hitachi Telecom, Turner Broadcasting System (TBS), Bell South, Digital Fountain, RealNetworks, Intel Research, Cisco Systems, and Lockheed Martin.

21. I am a Member of the Association of Computing Machinery (ACM) and a Fellow of the Institute of Electrical and Electronics Engineers (IEEE).

22. As set forth in my CV, I have authored nearly 200 refereed papers on topics including mesh networks, wireless network monitoring and management, information dissemination in mobile networks, multicast applications support, multicast security, and tools for multicast monitoring and management.

23. Additional details about my employment history, fields of expertise, and publications are further included in my CV.

III. UNDERSTANDING OF PATENT LAW

24. I have been advised by counsel of the following legal standards, which I have applied to my analysis.

25. I understand that prior art to the '952 patent includes patents and printed publications in the relevant art that may predate the priority date of the inventions recited in the '952 patent. For the purposes of my analysis, I have assumed that the priority date for the '952 patent is the earliest filing date claimed on the face of the patent, which is October 12, 2012, the date of the Great Britain application 1218370.3.

26. I understand that a claim is invalid if it is anticipated or obvious. Anticipation of a claim requires that every element of a claim be disclosed expressly or inherently in a single prior art reference, arranged in the prior art reference as arranged in the claim. I further understand that for an element of a claim to be disclosed inherently, the prior art must necessarily include the claim element, or the claimed element must inevitably result from the steps disclosed in the prior art.

27. Obviousness of a claim requires that the claim be obvious from the perspective of a person having ordinary skill in the relevant art at the time the alleged invention was made. I understand that a claim may be obvious from a combination of two or more prior art references.

28. For my obviousness analysis, I have been instructed to consider the following: (a) the scope and content of the prior art; (b) the differences between the prior art and the subject matter of the claimed invention; (c) the level of ordinary skill in the art at the time of the claimed invention; (d) the real-world factors

constituting objective indicia of non-obviousness such as long-felt need, failure of others to solve the problems addressed by each invention, initial skepticism, unexpected results, recognition in the industry, copying, and commercial success. I have been informed that when determining invalidity for obviousness by combining prior art references, it is relevant whether a person having ordinary skill in the art (also known as a “POSITA”) would have been motivated, taught, or suggested to combine the relevant teachings of the prior art and form the claimed invention. I have been informed that a POSITA is presumed to have had knowledge of analogous prior art in the field of invention. I have been informed that the motivation, teaching, or suggestion to combine prior art may come from the prior art itself, the knowledge of a POSITA, or from the nature of the problem to be solved. I also understand that it is not permissible to use hindsight in assessing whether the invention is invalid for obviousness. I understand that one should not look at the invention knowing what persons of ordinary skill in the art know today. Rather, you must place yourself in the shoes of a POSITA in the field of the technology of the patent at the time of the filing of the patent application.

29. I further understand that in an *inter partes* review a petitioner carries the burden of proof of showing unpatentability by a preponderance of the evidence.

IV. PERSON OF ORDINARY SKILL IN THE ART

30. I have been informed that I should consider the following factors to define the level of ordinary skill in the art: (i) the types of problems encountered in the art; (ii) the prior art solutions to those problems; (iii) the rapidity with which innovations are made; (iv) the sophistication of the technology; and (v) the educational level of active workers in the field.

31. In my opinion, POSITA at the time of the '952 patent would have (i) 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, or (ii) additional education that might compensate for a deficiency in experience, and vice-versa.

32. It is also my opinion that a person of ordinary skill in the art would be knowledgeable and familiar with the video encoding or processing concepts and techniques recited in the claims of the '952 patent, as they were quiet common and well known at the time of the '952 patent.

33. As of this time, I was at least a person of ordinary skill in the art, and for all my opinions in this declaration, I have applied the knowledge of a person having ordinary skill in the art at the time of the '952 patent.

V. BACKGROUND AND SUMMARY OF THE '952 PATENT

A. Background of the Field Relevant to the '952 Patent

1. History of Streaming Standards

34. Video content delivery via the Internet began in the late nineteen-nineties. In the early days, due to limitations of Internet speeds and old-fashioned CPUs, all video content had to be downloaded completely before viewing. One of the primary ways in which online video was primarily delivered through the real-time messaging protocol (“RTMP”), which is a Flash-based standard. But RTMP required plugins for video playback, which was neither efficient nor reliable. Thus, fewer and fewer devices supported this aging protocol nowadays.

35. As a result of the decline of RTMP and Flash-based content delivery, the media streaming industry developed HTTP-based content delivery protocols by the mid-2000s. In 2007, HTTP-based adaptive streaming was introduced, which adjusted the quality of a video stream according to a user’s bandwidth and CPU capacity. Vendor-specific live-streaming protocols soon followed like Microsoft’s Smooth Streaming in 2008, Apple’s HTTP Live Streaming (HLS) in 2009, and Adobe’s HTTP Dynamic Streaming in 2010. These proprietary, adaptive streaming technologies allow for video streaming by breaking up video data into a sequence of small HTTP-based file segments that contain a short interval of playback content. This way video streaming occurs over the Internet so that the client device does not have to download the entire video file before playing it.

2. Industry Need for an Independent Standard

36. HTTP-based adaptive streaming was quickly adopted but competing protocols created market fragmentation causing a barrier to further growth of the online video ecosystem. In April 2009, the telecommunications group 3rd General Partnership Project (3GPP) began efforts to standardize adaptive streaming protocols to potentially replace existing proprietary technologies like Apple's HLS, Microsoft's Smooth Streaming, and Adobe's HDS. More than 50 companies, in coordination with other industry organizations, were involved in developing MPEG's Dynamic Adaptive Streaming over HTTP ("MPEG-DASH")—a vendor-independent, international protocol. MPEG-DASH is an open-sourced standard that became the alternative, unifying streaming protocol that the content streaming industry needed.

37. The MPEG-DASH specification (ISO/IEC 23009-1) was ratified in 2011 and first published in early 2012. The first major trial utilizing MPEG-DASH occurred in 2012 for the London Olympics, when media consumers experienced the Olympic Games broadcast on their personal devices via MPEG-DASH. This effort demonstrated the benefits of the standardization for adaptive streaming.

38. Because several major media companies, including Adobe, Samsung, Microsoft, Dolby, Netflix, and Cisco, took part in the standardization efforts, MPEG-DASH helped eliminate technical issues in delivery and compression. In effect, it combined all of the existing technologies and standards into one, making

streaming support seamless on all devices. As a result of MPEG-DASH, content publishers are now able to generate a single set of files for encoding and streaming so that consumers can play the content they want to watch on any supporting devices.

39. While MPEG-DASH is audio/video codec agnostic, there are profiles in the specification that indicate how media is to be segmented on the server for ISO Base Media File Format (BMFF) and MPEG2 Transport Stream container formats. Additionally, both live and on-demand media types have been given special consideration.

3. Benefits of MPEG-DASH

40. MPEG-DASH offers several benefits over its Flash-based and vendor-specific predecessors. For example, it reduces the startup delays and buffering time during videos by continually adapting to the consumer's available bandwidth. This allows a mobile user to receive a full HD stream using high-speed Internet access and continue watching the same video, but at a lower quality on a 4G network. MPEG-DASH also utilizes existing and cost-effective HTTP-based content delivery networks, proxies, and caches. This means that, unlike older technologies, MPEG-DASH streams are delivered through standard web servers, which provide maximum speed with minimum expense and setup costs for content creators and publishers. These benefits are actualized, at least in part, by the way the MPEG-DASH standard

requires media files to be segmented and encoded at different bitrates or spatial resolutions.

41. The '952 patent claims nothing more than well-known practices of splitting up video into segments 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 that I will discuss in details below.

42. Each of these 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.

B. Summary of the '952 Patent

43. The '952 patent was filed on October 13, 2017 and is entitled “Method and Corresponding Device for Streaming Video Data.” Ex. 1100 at Cover. The patent states that it is a “[c]ontinuation” of an application filed October 11, 2013. *Id.*

44. 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. 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 the alleged procedure.

45. 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,” 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.*

C. The ’952 Patent Claims

46. 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.

47. The four independent claims 1, 3, 16, and 19 are similar in scope. Each claims dividing a video into n partitions in order to create n independent videos. *See, e.g.,* Ex. 1100 at claim 1. 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.*

48. Dependent claims 2, 4, 17, and 20 claim spatial information that includes both a width and height of the regions. Dependent claims 5 and 6 include partition number information. Dependent claim 7 claims 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 8 claims spatial size information that represents a spatial size of a video corresponding to the spatial regions. Dependent claim 9 claims that each frame is divided into rectangles. 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. Dependent claim 15 claims a notification that “includes a parameter relating to at least one of an image resolution, a frame rate and a bandwidth of the video.” *Id.* at

claim 15. Dependent claim 18 claims a user interface to determine the Region Of Interest. Dependent claim 21 claims a camera unit for capturing the video data.

49. Representative claim 3 is reproduced below:

3. A transmitting method for transmitting video data to a client, the method comprising:

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,

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, and

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,

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.

D. The '952 Patent Specification

50. 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.

51. 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 1:43-44. 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:

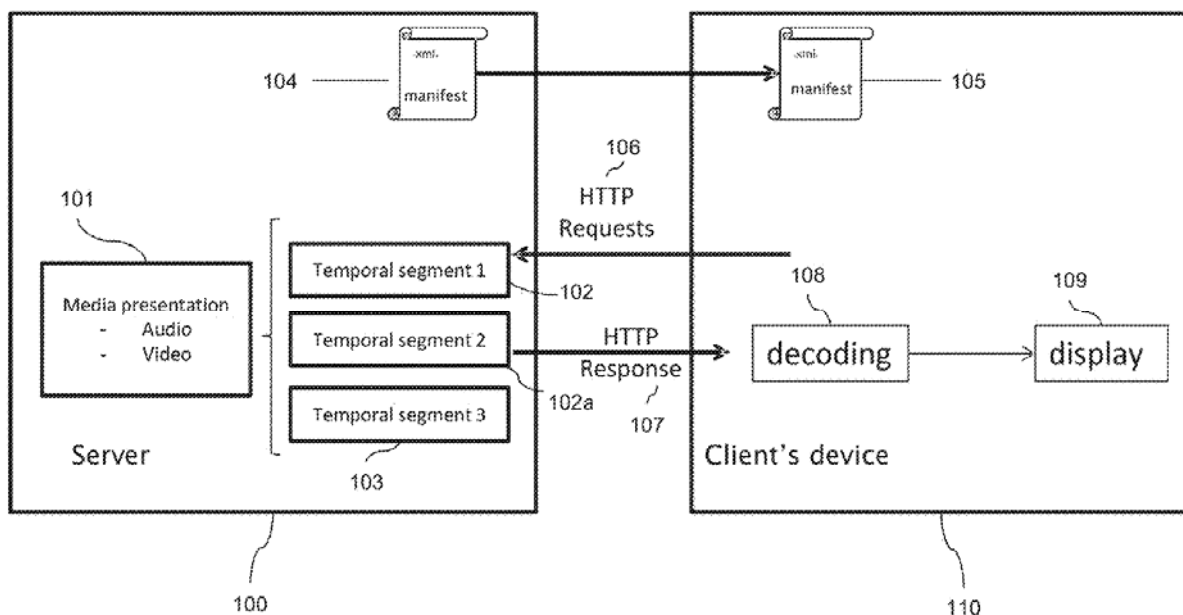


Figure 1a

Ex. 1100 at Fig. 1a.

52. 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 2:5-9. 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.*

53. 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.

54. The patent then acknowledges that other systems have accomplished the same solution: in an article by *Mavlankar et al.*, 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.

55. 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:

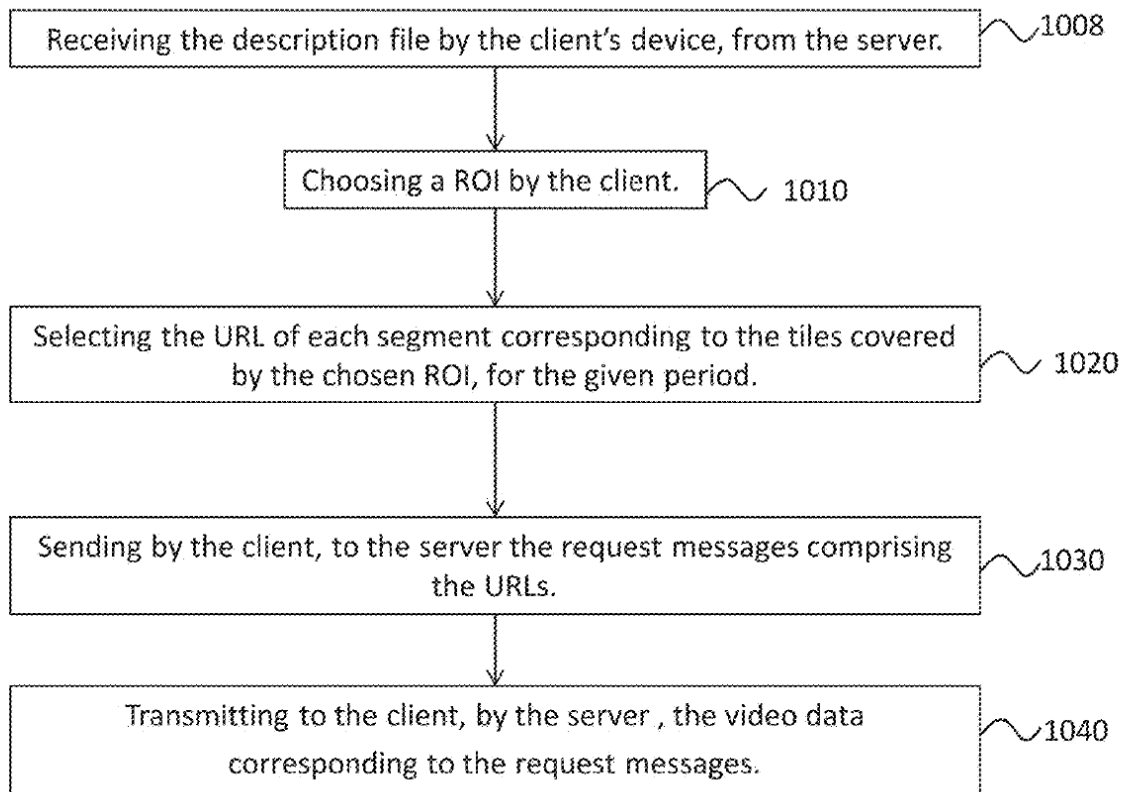


Figure 1b

Id. at Fig. 1b.

56. 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 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.

E. The ’952 Patent Prosecution History

57. As part of my analysis, I read and considered the ’952 patent and related prosecution history before the Patent Office. The following overview is not meant to describe my full understanding of the ’952 patent and prosecution history, but rather to highlight the general aspects of the ’952 patent and prosecution history.

58. As I discussed above 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.*

59. According to the prosecution history, 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 and this is depicted in the figure below.

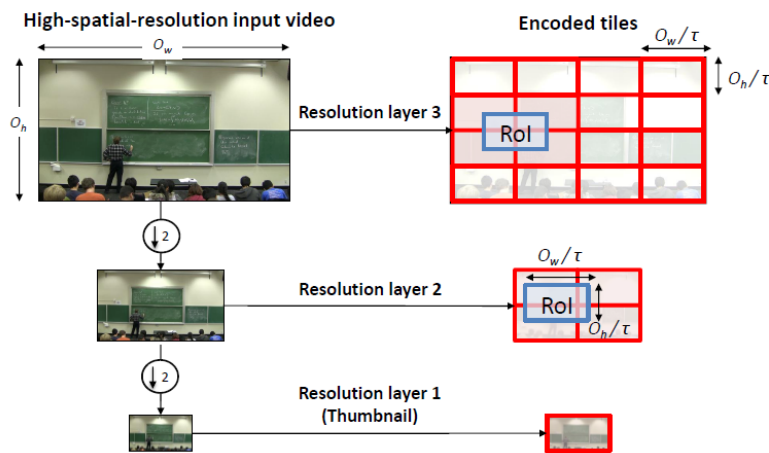


Fig. 4. Proposed coding scheme shown for the case of three resolution layers. The original high-spatial-resolution video is dyadically downsampled to create multiple resolution layers. Each resolution layer is further subdivided into tiles. Each tile is independently coded using an H.264/AVC video encoder.

Ex. 1105 at Fig. 4; Ex. 1102 at 238.

60. The Mavlankar Publication further disclosed transmitted the requested high-resolution video tile to a client. Ex. 1105 at Fig. 4; Ex. 1102 at 238. I understand that 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.

61. After the rejection, the Patent Owner did not contest that the Mavlankar Publication in view of Koyanagi rendered obvious the original claims. Rather, 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. After an interview, the Examiner issued a notice of allowance along with an Examiner’s Amendment. *Id.* at 324-25. 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.

62. 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.” Ex. 1102 at 330. But, as I would explain below, *Irie* and *Van Deventer*—which were not before the Examiner during the original prosecution of the ’952 patent—expressly teach this functionality.

VI. CLAIM CONSTRUCTION

63. I understand that for the purposes of this *inter partes* review only, Avigilon proposed that the challenged claims be given their plain and ordinary meaning under the “*Phillips* standard,” which instructs claim terms to be construed in accordance with the ordinary and customary meaning of such claims as understood by a POSITA and the prosecution history pertaining to the patent. I am not aware of any particular interpretations of the claims that would lead them to fall outside of the scope of what is disclosed in the grounds, but I reserve the right to respond to any proffered claim constructions by the patent owner.

VII. OVERVIEW OF THE PRIOR ART

64. Before providing a detailed analysis of how the prior art invalidates the Challenged Claims, I provide a brief summary of the asserted prior art references.

A. *Irie*

65. *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.

66. *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.

67. *Irie* also 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.

68. Similar to 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—which is 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.

69. As I mentioned earlier, *Irie* was not in front of the Examiner during the 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 in view of *Irie*.

B. Van Deventer

70. *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:

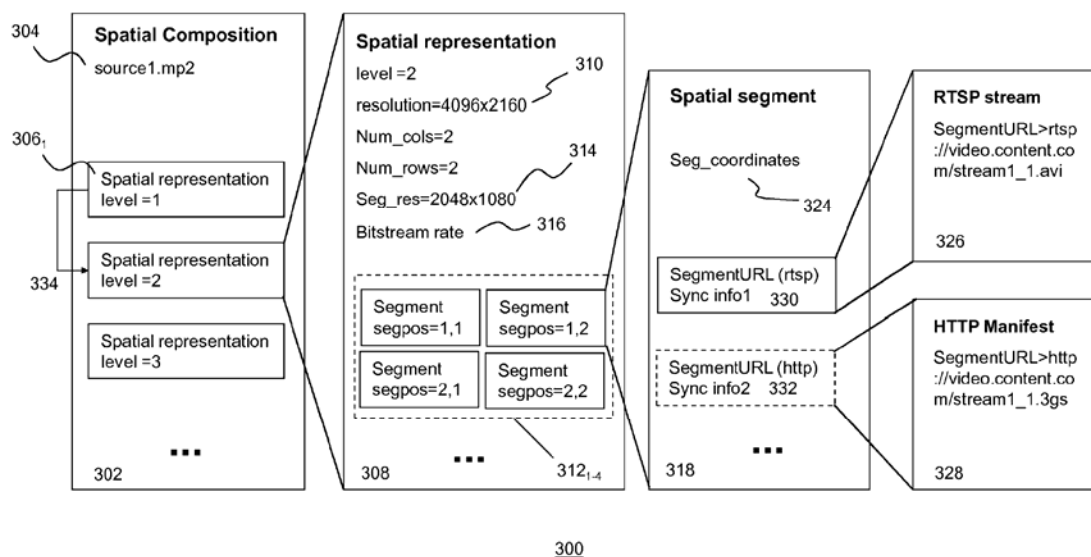


Figure 3

Id. at Fig. 3.

71. 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 3:67-4:8. *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.

VIII. CLAIMS 1-6, 8-10, AND 15-21 ARE UNPATENTABLE IN VIEW OF THE PRIOR ART

72. I have been asked to provide an opinion as to whether the Challenged Claims of the '952 patent are invalid in view of the prior art identified in the proposed grounds. The discussion below provides a detailed invalidity analysis of how the prior art references identified in Section VII invalidate the Challenged Claims of the '952 patent.

73. As part of my obviousness analysis, I have considered the scope and content of the prior art, and whether any differences between the alleged invention and the prior art are such that the subject matter, as a whole, would have been obvious to a person having ordinary skill in the art at the time of the alleged invention. I have also considered the level of ordinary skill in the pertinent art. I have also considered the secondary considerations of non-obviousness.

74. I describe in detail below the scope and content of the prior art, as well as any differences between the alleged invention and the prior art, on an element-by-element basis for the Challenged Claims of the '952 patent.

75. It is my opinion that claims 7 and 11-14 are obvious in view of the combination of *Irie* and *Van Deventer*.

A. Ground I: Claims 7 and 11-14 are Obvious Over *Irie* in view of *Van Deventer*

1. Motivation to Combine *Irie* with *Van Deventer*

76. A POSITA would have been motivated to combine *Irie* and *Van Deventer* and would have had a reasonable expectation of success in doing so. *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 ¶ 56 (citing Ex. 1103). From a technical perspective, *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* specifically discloses the use of its invention in content-delivery networks, which is the same subject matter described in *Van Deventer*. *Id.* ¶ 1; Ex. 1104 at 2:33-40. *Van Deventer* explains that its system provides content delivery “without requiring changes to conventional [content-delivery network] infrastructure.” Ex. 1104 at 2:36-38. The *Van Deventer* system discloses a method for partitioning and streaming video data, including a Region of Interest, focusing on the “spatial manifest file” transmitted to the client. This is accomplished using “spatially-segmented content delivery” that improves the transmission of high-resolution video content. See Ex. 1104 at 1:15-23. *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.

77. The techniques described in *Irie* have broad applicability, as *Irie* itself notes. Ex. 1103 ¶¶ 1-3 (discussing IP multicast delivery services and VoD multicast delivery services). *Irie* provides detailed information to enable users to enlarge regions of interest from videos from various types of content delivery services without reducing the resolution. *Id.* ¶ 9. Although *Van Deventer* also describes users enlarging regions of interests in videos, 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. 1104 at 2:33-40. The advantage of *Van Deventer*’s system is that it can be used with a wide variety of exist content-delivery networks, such as the ones described in *Irie*, and a POSITA would have understood that *Irie*’s video-enlarging techniques could easily be adapted to use the spatial manifest file described by *Van Deventer*. See *id.* at 2:32-40, 2:63-67. For example, the spatial manifest file described by *Van Deventer* could be used to carry the information discussed in the notifications disclosed in *Irie*. A POSITA would thus have had a reasonable expectation of success in combining *Irie* and *Van Deventer*, as the two systems are complementary and compatible. See *id.*

78. 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

79. For the reasons explained below, claims 7 and 11-14 are obvious in view of *Irie* and *Van Deventer*. These claims depend on independent claim 3, thus I start by setting forth why claim 3 is rendered obvious by *Irie* in view of *Van Deventer*.

80. Claim 3 begins with the following preamble: “A transmitting method for transmitting 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, it is my opinion that *Irie* discloses the preamble of claim 3.

81. The first element of claim 3 reads as follows: “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.

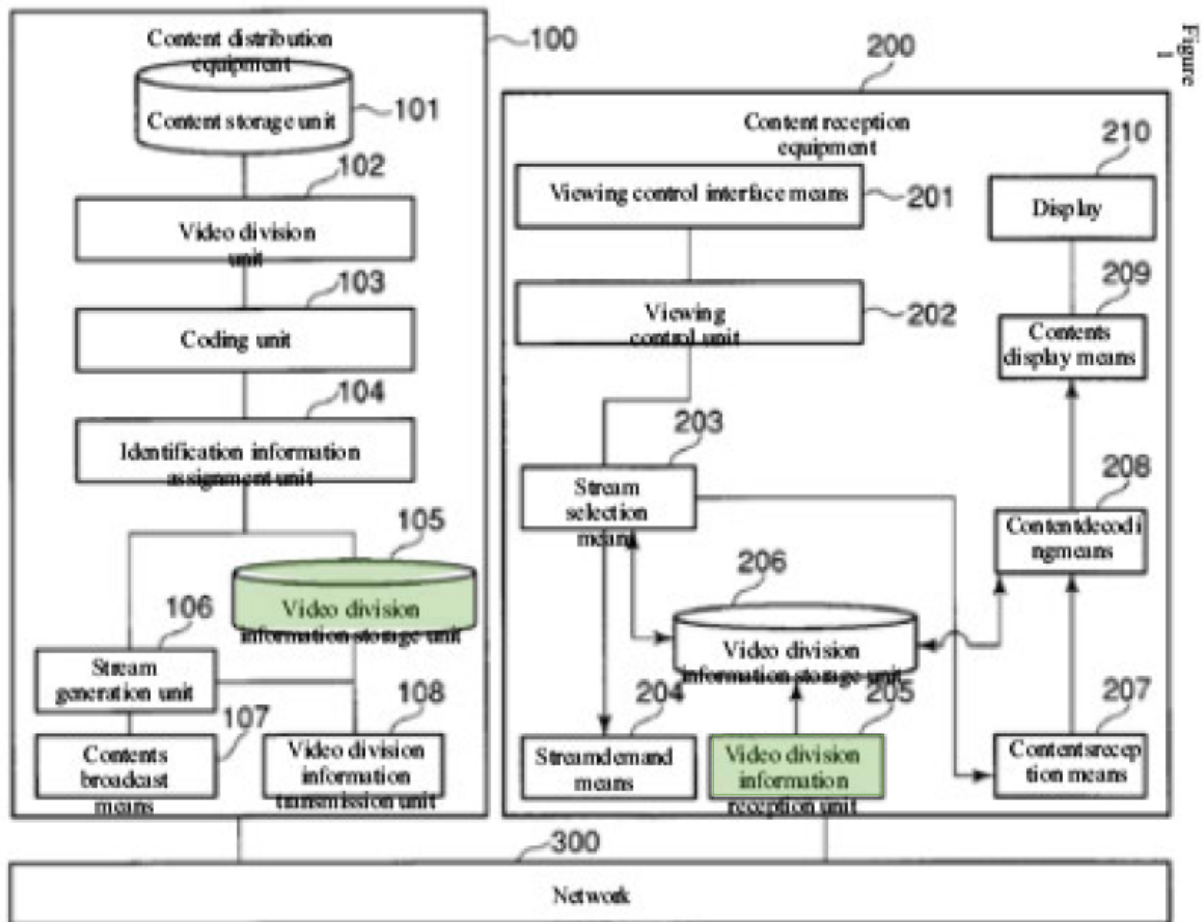
82. *Irie* discloses transmitting to a client a notification that includes the spatial position information represented by a pixel position and spatial size information. For example, *Irie* explains that it “assigns identification data to the multiple blocks to identify each of them” and “sends video 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.

83. *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.

84. 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:



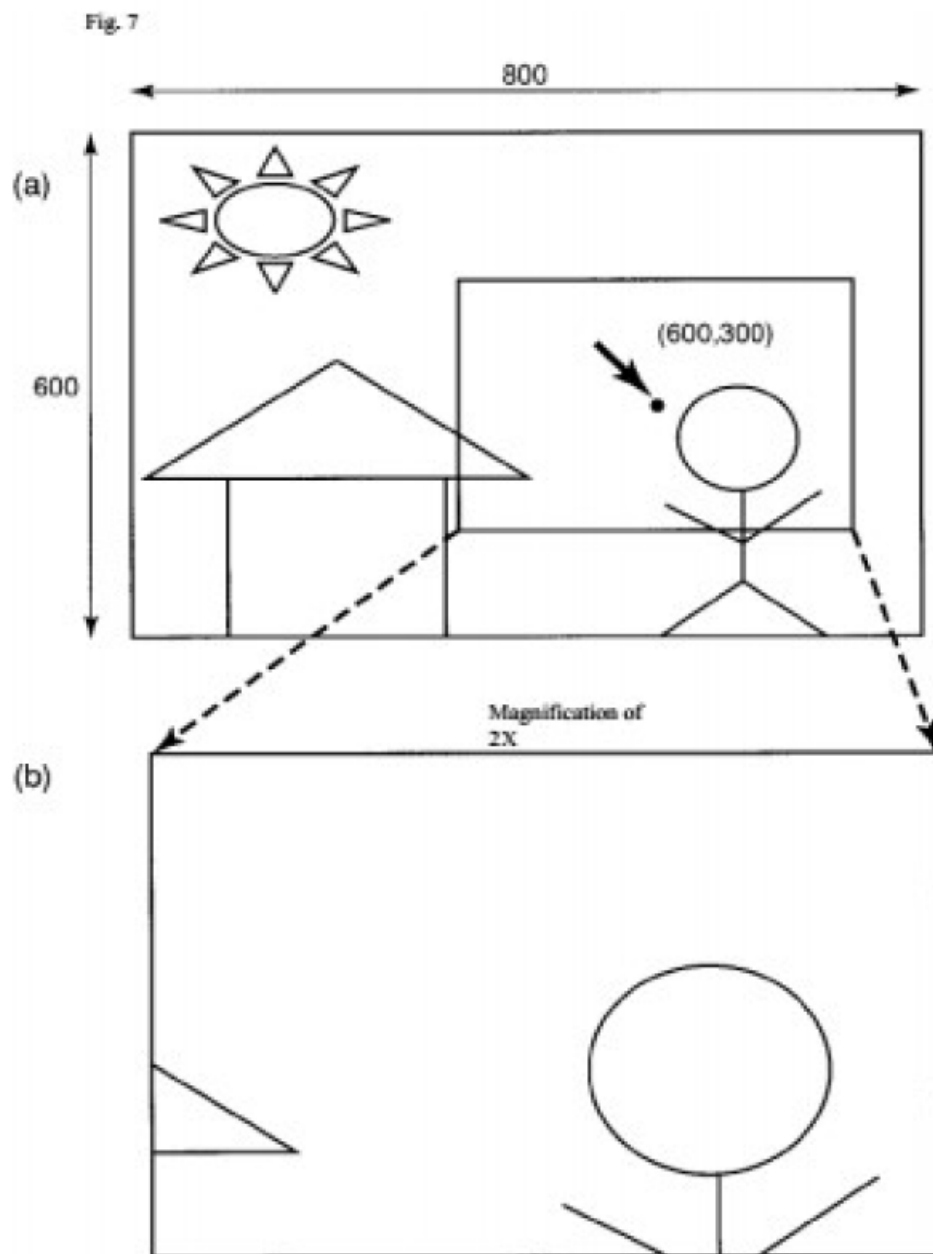
Id. at Fig. 1 (annotated), ¶¶ 18, 19.

85. 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, it is my opinion that *Irie* discloses the first element of claim 3.

86. The second element of claim 3 reads as follows: “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.

87. *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.

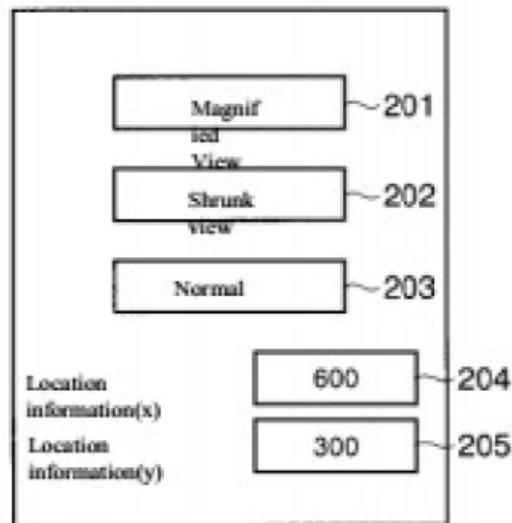
88. *Irie* illustrates this functionality in Figure 7, where a specific region of interest is selected in Figure 7(a) and magnified using a higher-resolution image in Figure 7(b):



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

89. *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

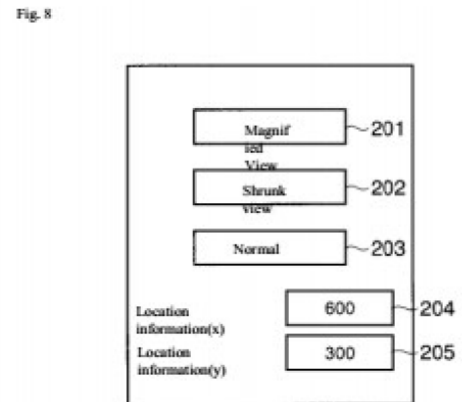
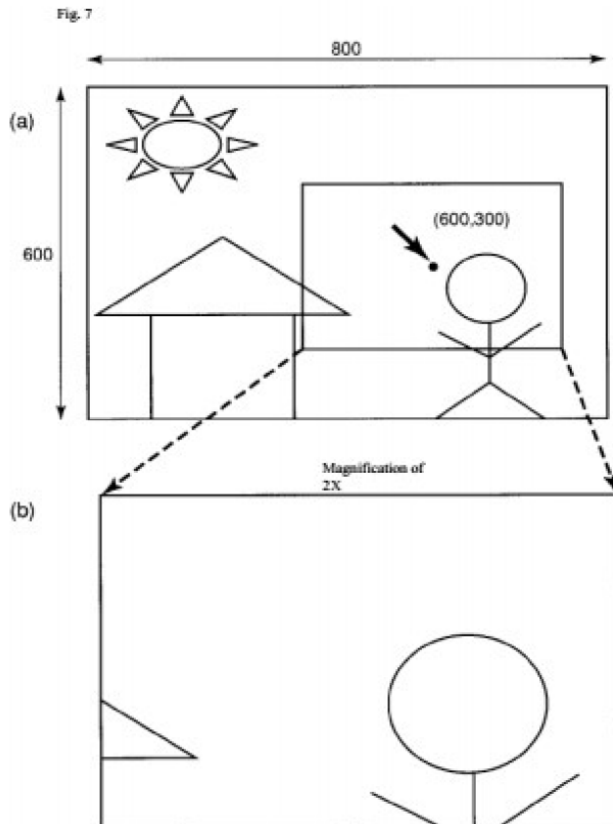


Id. at Fig. 8.

90. Therefore, it is my opinion that *Irie* discloses the second element of claim 3.

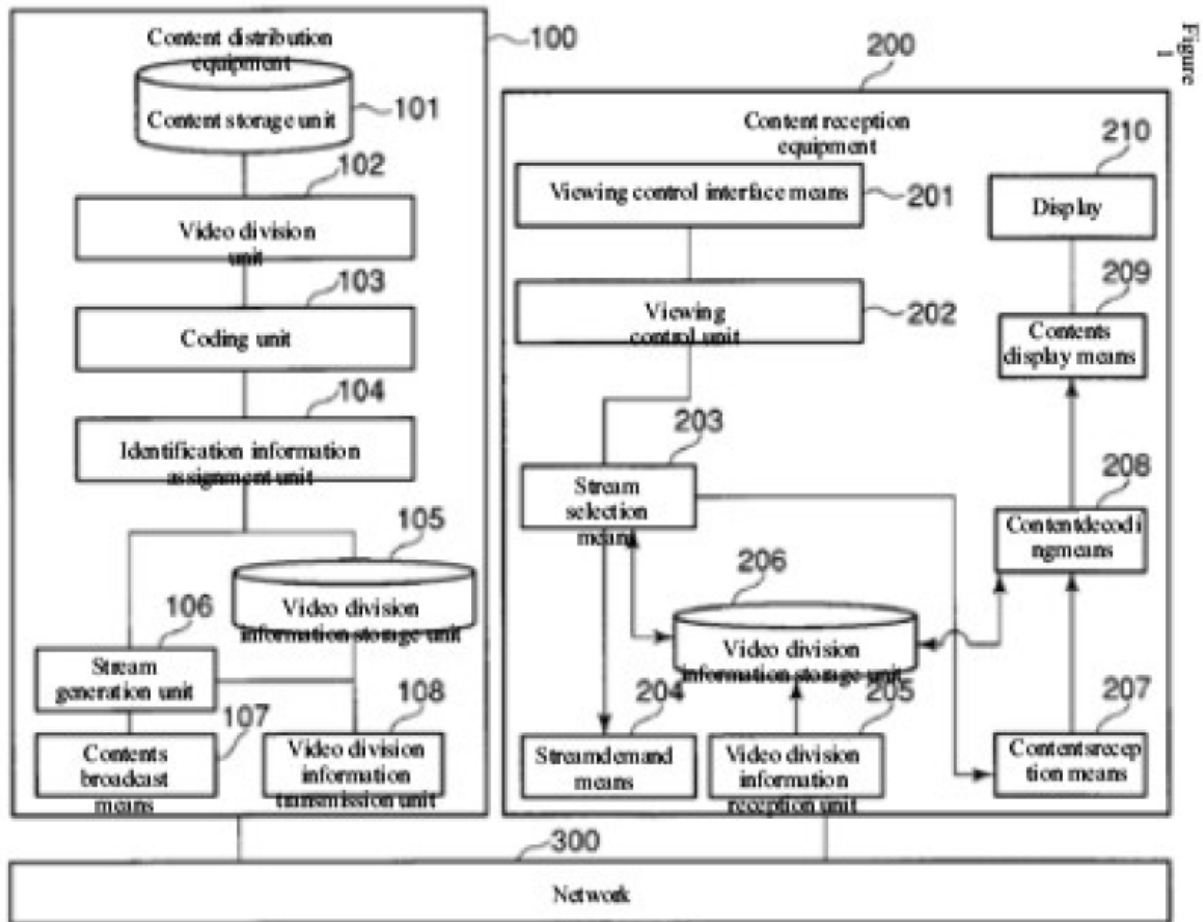
91. The third element of claim 3 reads as follows: “transmitting, to the client, the video data corresponding to the one or more spatial regions which are identified according to the information received from the client.” Ex. 1100, claim 3.

92. *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:



Ex. 1103 at Figs. 7, 8.

93. *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

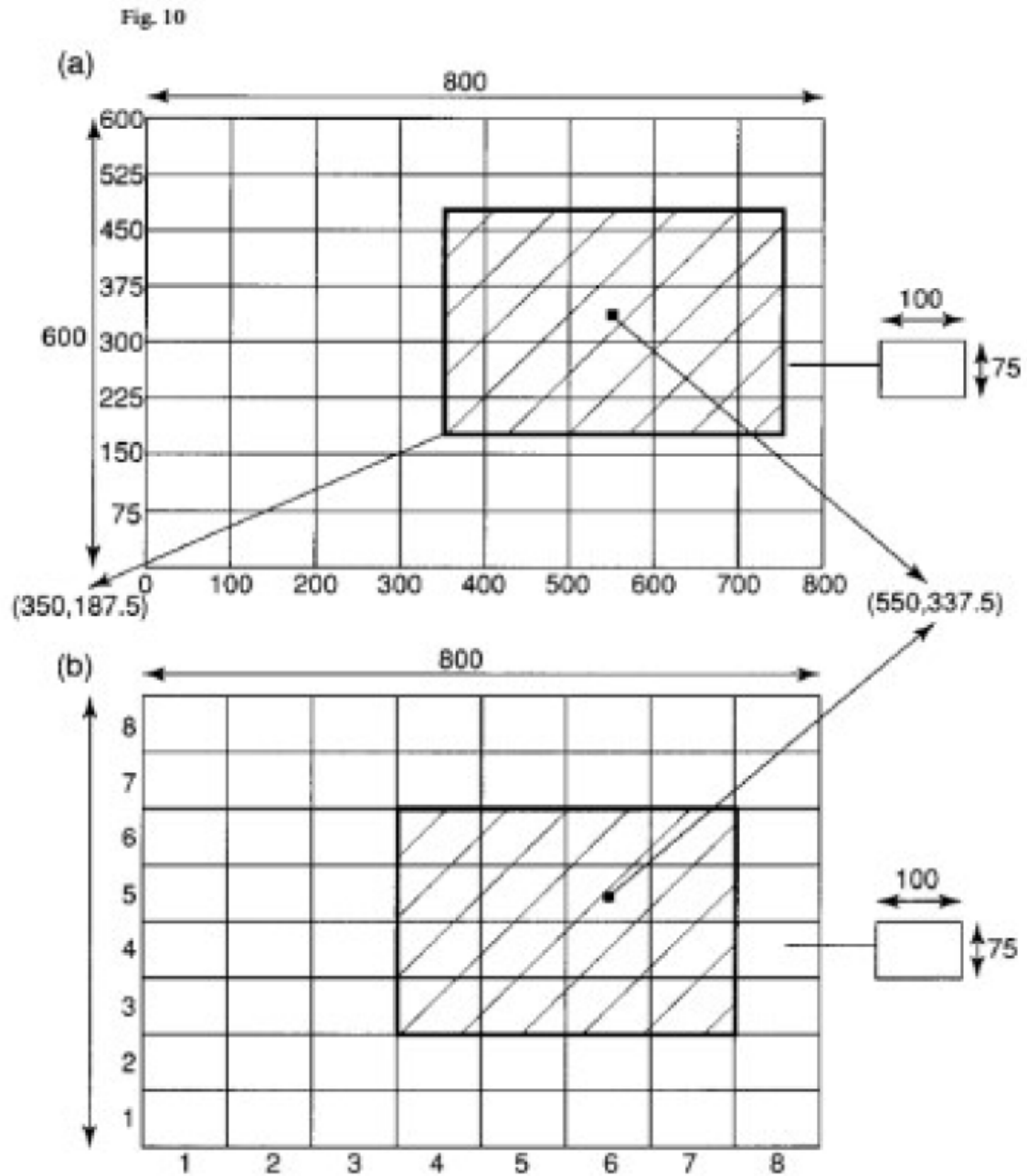
94. *Irie* in paragraph 19 describes that the content receiving device 200 contains the viewing control interface part 201, a viewing control part 202, a stream selection part 203, a stream request part 204, a video division information receiving part 205, a content receiving part 207, a content decoding 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.

95. 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.

96. Therefore, it is my opinion that *Irie* discloses the third element of claim 3.

97. The fourth element of claim 3 reads as follows: “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.

98. *Irie* discloses this claim limitation. As noted above for claim element the first element of claim 3, 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:



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

99. Therefore, it is my opinion that *Irie* discloses the fourth element of claim 3.

3. Claim 7

100. 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.”

101. 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 size information transmitted to the client in *Irie* can be of different sizes given that there are “no restriction on 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.

102. Nonetheless, *Irie* in view of *Van Deventer* also discloses this claim. *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):

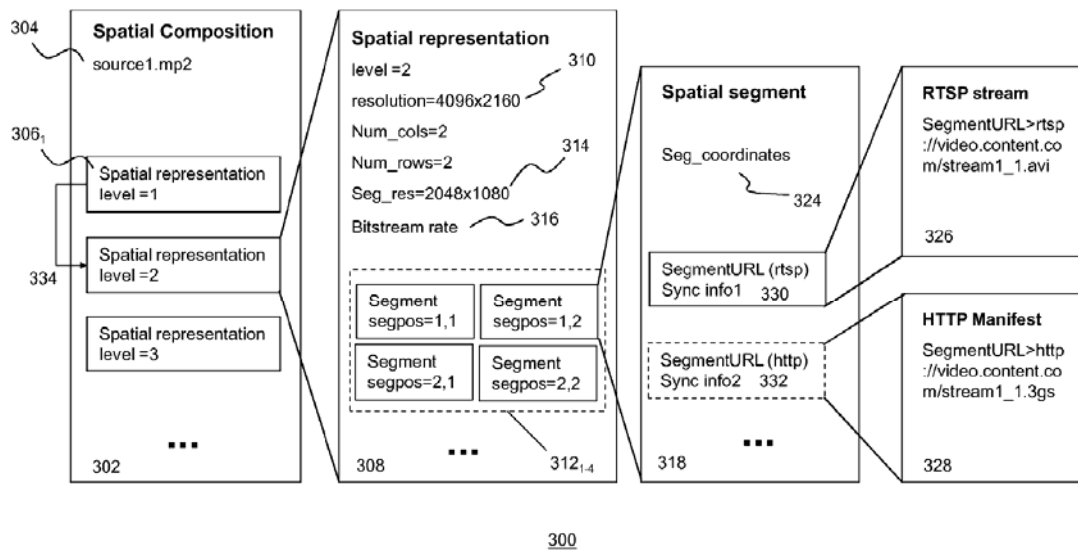


Figure 3

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

103. Therefore, for the reasons discussed above, it is my opinion that a POSITA would have been motivated to use the information disclosed in *Van Deventer*'s notifications in combination with *Irie*, thus disclosing claim 7 of the '952 patent.

4. Claim 11

104. 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.” Ex. 1100 at claim 11. *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.

105. A POSITA would thus have known that information from the client like the multicast address would include metadata that includes “one or more URLs

associated with the one or more spatial regions” as a way to identify the spatial region the client is interested in. Ex. 1100, claim 11.

106. 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 8:7-12. 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, it is my opinion that a POSITA would have been motivated to use the information disclosed in *Van Deventer*’s notifications in combination with *Irie*.

5. Claim 12

107. 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 ¶ 14. 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.

108. 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.

109. Accordingly, it is my opinion that 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. 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*.

6. Claim 13

110. Claim 13, which depends on only 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.” Ex. 1100 at claim 13. *Irie* in view of *Van Deventer* discloses this claim.

111. *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.

112. The content delivery device 100 shown in Figure 1 of *Irie* contains a “video division information storage part 105” and a “video division information receiving part 205.” *Id.* ¶¶ 18,19. 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.

113. *Van Deventer* adds on to this and 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.

114. For the reasons discussed above, it is my opinion that a POSITA would have been motivated to use the information disclosed in *Van Deventer*’s notifications in combination with *Irie*.

7. Claim 14

115. 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.” Ex. 1100 at claim 14. *Irie* in view of *Van Deventer* discloses this claim.

116. As explained in above in Section VIII.A.13, *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, each instance of spatial size information and spatial position information in *Irie* corresponds to its respective time period of the video frame.

117. Furthermore, *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, it is my opinion that a POSITA would have been motivated to use the information disclosed in *Van Deventer*'s notifications in combination with *Irie*.

IX. SECONDARY CONSIDERATIONS OF NON-OBVIOUSNESS

118. Should Patent Owner assert that secondary considerations support a finding of non-obviousness, I reserve the right to submit a Declaration addressing those new assertions.

X. CONCLUSION

119. I hereby declare under penalty of perjury under the laws of the United States of America that the foregoing is true and correct, and that all statements made of my own knowledge are true and that all statements made on information and belief are believed to be true. I understand that willful false statements are punishable by fine or imprisonment or both. *See* 18 U.S.C. § 1001.

Date: April 24, 2020

Respectfully submitted,



Dr. Kevin C. Almeroth