

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

INTELLIGENT PROTECTION MANAGEMENT CORP.,
Petitioner,

v.

CISCO TECHNOLOGY, INC.,
Patent Owner.

Case IPR2025-01588
U.S. Patent No. 8,830,293

**PETITION FOR *INTER PARTES* REVIEW
OF U.S. PATENT NO. 8,830,293**

TABLE OF CONTENTS

I. STATEMENT OF PRECISE RELIEF REQUESTED1

II. The '293 Patent.....3

 A. Overview3

III. Level of Ordinary Skill.....10

IV. Claim Construction.....11

V. Claims 1–20 are Unpatentable Over the Prior Art12

 A. Ground 1: Tyso Renders Obvious Claims 1-2012

 1. Independent Claim 1: [1.pre] “A method comprising”.....19

 2. Claim 2: “The method of claim 1, wherein the subject image of video frames of the first video stream comprises images of multiple videoconference participants and the subject image of video frames of the second video stream comprises images of multiple videoconference participants.”37

 3. Claim 3: [3.pre] “The method of claim 1, wherein combining comprises:38

 4. Claim 4: “4. The method of claim 3, wherein the second direction is opposite to the first direction.”.....42

 5. Claim 5: [5.pre] “5. The method of claim 1, wherein combining further comprises:”.....43

 6. Claim 6: “6. The method of claim 1, wherein the first and second video streams are produced at a same video conferencing site.”45

 7. Claim 7: “7. The method of claim 1, wherein the first and second video streams are produced at different video conferencing sites.”46

8.	Claim 8: “8. The method of claim 1, wherein combining is performed at a site other than where the first and second video streams are produced.”	47
9.	Claim 9: “9. The method of claim 1, wherein the subject images of video frames of the first and second video streams are images of video conference participants.”	48
10.	Claim 10: [10.pre] 10. An apparatus comprising:.....	49
11.	Claim 11: [11.pre, 11.a] “11. The apparatus of claim 10, wherein the at least one data processor is configured to combine the subject images of video frames by:”	51
12.	Claim 12: [12.pre] “12. A system comprising the apparatus of claim 10, and further comprising:”	52
13.	Claim 13: [13.pre] “13. Logic encoded in one or more tangible non-transitory storage media for execution and when executed operable to:”	56
14.	Claim 14: “14. The logic of claim 13, wherein the logic that receives is configured to receive video frames of the first video stream comprising images of multiple videoconference participants in the subject image and video frames of the second video stream comprising images of multiple videoconference participants in the subject image.....	58
15.	Claim 15: [15.pre] “15. The logic of claim 13, wherein the logic that combines is configured to:”	58
16.	Claim 16: “16. The logic of claim 15, wherein the second direction is opposite to the first direction.”	59
17.	Claim 17: [17.pre] “17. The logic of claim 13, wherein the logic that combines is configured to:”	60
18.	Claim 18: “18. The logic of claim 13, wherein the first and second video streams are produced at the same video conferencing site.”	61

19.	Claim 19: “19. The logic of claim 13, wherein the first and second video streams are produced at different video conferencing sites.”	61
20.	Claim 20: “20. The logic of claim 13, wherein the logic that combines is configured to be performed at a site other than where the first and second video streams are produced.”	61
B.	Ground 2: Tysso in view of GIMP Renders Obvious Claims 1-20	62
1.	Independent Claim 1: [1.pre]	64
2.	Claim 2	65
3.	Claim 3: [3.pre]	65
4.	Claims 4-9	66
5.	Claim 10: [10.pre], [10.a], [10.b], and [10.c].....	66
6.	Claim 11	67
7.	Claim 12	68
8.	Claim 13: [13.pre]	68
9.	Claim 14	69
10.	Claim 15: [15.pre]	69
11.	Claim 16-20.....	70
VI.	THE BOARD SHOULD INSTITUTE REVIEW	71
VII.	MANDATORY NOTICES	71
A.	Related Matters.....	71
B.	Real Party-in-Interest	71
C.	Lead and Back-Up Counsel and Service Information	72

VIII. GROUNDS FOR STANDING.....72

IX. CONCLUSION.....72

TABLE OF AUTHORITIES

Cases

Nidec Motor Corp. v. Zhongshan Broad Ocean Motor Co.,
868 F.3d 1013 (Fed. Cir. 2017)11

Phillips v. AWH Corp.,
415 F.3d 1303 (Fed. Cir. 2005) (en banc)11

Statutes

35 U.S.C. § 3111

35 U.S.C. § 325(d)70

Rules

37 C.F.R. §42.100(b)11

PETITIONER’S EXHIBIT LIST

Exh. No.	Description
1001	U.S. Patent No. 8,830,293 (“the ’293 patent”)
1002	File History of U.S. Patent No. 8,830,293
1003	Expert Declaration of Omid Kia, Ph. D.
1004	U.S. Patent No. 8,345,082 (“Tysso”)
1005	U.S. Patent Application No. 61/103,588 (“Tysso Provisional”)
1006	User Manual for GNU Image Manipulation Program (dated May 1, 2007) (“GIMP”)
1007	Printout of https://web.archive.org/web/20070521020845/http://docs.gimp.org/download.html , captured by Internet Archive on May 21, 2007
1008	Screenshot of https://web.archive.org/web/20070512113240/http://docs.gimp.org/en.pdf , captured by Internet Archive on May 12, 2007
1009	Printout of https://archive.org/legal/affidavit

CLAIM LISTING

1.pre	1. A method comprising:
1.a	receiving at least first and second real-time video streams, each of the first and second real-time video streams comprising video frames containing a picture comprising a subject image and a background image;
1.b	combining the subject images of corresponding video frames of the first and second video streams into a combined frame of a combined video stream such that the subject image of the first video stream is positioned in an anterior portion of the combined frame and the subject image of the second video stream is positioned in a posterior portion of the combined frame, wherein the combining comprises:
1.c	scaling the video frames of the first video stream and repositioning in a first direction the resulting pictures in the video frames of the first video stream to produce a first sequence of scaled video frames;

1.d	removing the background image in the first sequence of scaled video frames to produce first background separated video frames to be displayed in the anterior portion of the combined frames of the combined video stream; and
1.e	superimposing the first background separated video frames onto corresponding ones of the video frames of the second video stream to produce combined video frames of the combined video stream; and
1.f	supplying the combined video stream to a video display for displaying the combined video stream.
2	2. The method of claim 1, wherein the subject image of video frames of the first video stream comprises images of multiple videoconference participants and the subject image of video frames of the second video stream comprises images of multiple videoconference participants.
3.pre	3. The method of claim 1, wherein combining comprises:

3.a	scaling the video frames of the second video stream and repositioning in a second direction the resulting pictures in the video frames of the second video stream to produce a second sequence of scaled video frames;
3.b	extending the background image in the second sequence of scaled video frames to produce background extended video frames to be displayed in the posterior portion of the combined frames of the combined video stream; and
3.c	superimposing the first background separated video frames onto corresponding ones of background extended video frames to produce combined video frames of the combined video stream.
4	4. The method of claim 3, wherein the second direction is opposite to the first direction.
5.pre	5. The method of claim 1, wherein combining further comprises:
5.a	removing the background image in the video frames of the first and second video streams to produce first and second background separated video frames;
5.b	generating supplemental background image video frames that comprise a supplemental background image; and

5.c	superimposing corresponding ones of the first background separated video frames and the second background separated video frames onto corresponding ones of the supplemental background video frames to produce the combined video frames of the combined video stream.
6	6. The method of claim 1, wherein the first and second video streams are produced at a same video conferencing site.
7	7. The method of claim 1, wherein the first and second video streams are produced at different video conferencing sites.
8	8. The method of claim 1, wherein combining is performed at a site other than where the first and second video streams are produced.
9	9. The method of claim 1, wherein the subject images of video frames of the first and second video streams are images of video conference participants.
10.pre	10. An apparatus comprising:
10.a	a first memory configured to store data for at least first and second real-time video streams, each of the first and second real-time video streams comprising video frames containing a picture comprising a subject image and a background image; and

10.b	at least one data processor configured to:
10.c	combine the subject images of corresponding video frames of the first and second video streams into a combined frame of a first combined video stream such that the subject image of the first video stream is positioned in an anterior portion of the combined frame and the subject image of the second video stream is positioned in a posterior portion of the combined frame, wherein the at least one data processor is configured to combine the subject images of video frames by:
10.d	scaling the video frames of the first video stream and repositioning in a first direction the resulting pictures in the video frames of the first video stream to produce a first sequence of scaled video frames;
10.e	removing the background image in the first sequence of scaled video frames to produce first background separated video frames to be displayed in the anterior portion of the combined frames of the combined video stream; and

10.f	superimposing the first background separated video frames onto corresponding ones of the video frames of the second video stream to produce combined video frames of the combined video stream.
11.pre	11. The apparatus of claim 10,
11.a	wherein the at least one data processor is configured to combine the subject images of video frames by:
11.b	scaling the video frames of the second video stream and reposition in a second direction the resulting pictures in the video frames of the second video stream to produce a second sequence of scaled video frames;
11.c	extending the background image in the second sequence of scaled video frames to produce background extended video frames to be displayed in the posterior portion of the combined frames of the combined video stream; and
11.d	superimposing the first background separated video frames onto corresponding ones of background extended video frames to produce combined video frames of the combined video stream.

12.pre	12. A system comprising the apparatus of claim 10, and further comprising:
12.a	a first video conferencing endpoint configured to communicate with the apparatus, the first endpoint including:
12.b	a first video camera configured to generate the first real-time video stream, wherein the subject image of the first real-time video stream comprises an image of a first video conferencing participant at the first endpoint;
12.c	a second video camera configured to generate the second real-time video stream, wherein the subject image of the second real-time video stream comprises an image of a second video conferencing participant at the first endpoint; and
12.d	a first control unit configured to forward the first and second real-time video streams to the apparatus; and
12.e	a second video conferencing endpoint configured to communicate with the apparatus, the second endpoint including:
12.f	a second control unit configured to receive the combined video stream of the apparatus; and

12.g	a display configured to receive and render the combined video stream from the control unit.
13.pre	13. Logic encoded in one or more tangible non-transitory storage media for execution and when executed operable to:
13.a	receive at least first and second real-time video streams, each of the first and second real-time video streams comprising video frames containing a picture comprising a subject image and a background image;
13.b	combine the subject images of corresponding video frames of the first and second video streams into a combined frame of a combined video stream such that the subject image of the first video stream is positioned in an anterior portion of the combined frame and the subject image of the second video stream is positioned in a posterior portion of the combined frame, wherein the subject images are combined by:
13.c	scaling the video frames of the first video stream and repositioning in a first direction the resulting pictures in the video frames of the first video stream to produce a first sequence of scaled video frames;

13.d	removing the background image in the first sequence of scaled video frames to produce first background separated video frames to be displayed in the anterior portion of the combined frames of the combined video stream; and
13.e	superimposing the first background separated video frames onto corresponding ones of the video frames of the second video stream to produce combined video frames of the combined video stream; and
13.f	supply the combined video stream to a video display for displaying the combined video stream.
14	14. The logic of claim 13, wherein the logic that receives is configured to receive video frames of the first video stream comprising images of multiple videoconference participants in the subject image and video frames of the second video stream comprising images of multiple videoconference participants in the subject image.
15.pre	15. The logic of claim 13, wherein the logic that combines is configured to:

15.a	scale the video frames of the second video stream and reposition in a second direction the resulting pictures in the video frames of the second video stream to produce a second sequence of scaled video frames;
15.b	extend the background image in the second sequence of scaled video frames to produce background extended video frames to be displayed in the posterior portion of the combined frames of the combined video stream; and
15.c	superimpose the first background separated video frames onto corresponding ones of background extended video frames to produce combined video frames of the combined video stream.
16	16. The logic of claim 15, wherein the second direction is opposite to the first direction.
17.pre	17. The logic of claim 13, wherein the logic that combines is configured to:
17.a	remove the background image in the video frames of the first and second video streams to produce first and second background separated video frames;

17.b	generate supplemental background image video frames that comprise a supplemental background image; and
17.c	superimpose corresponding ones of the first background separated video frames, and the second background separated video frames onto corresponding ones of the supplemental background video frames to produce the combined video frames of the combined video stream.
18	18. The logic of claim 13, wherein the first and second video streams are produced at the same video conferencing site.
19	19. The logic of claim 13, wherein the first and second video streams are produced at different video conferencing sites.
20	20. The logic of claim 13, wherein the logic that combines is configured to be performed at a site other than where the first and second video streams are produced.

Intelligent Protection Management Corp. (“Petitioner” or “IPM”) petitions for *Inter Partes* Review (“IPR”) of claims 1-20 (“the Challenged Claims”) of U.S. Patent No. 8,830,293 (“the ’293 patent”). The Grounds presented in this Petition are supported by the Declaration of Dr. Omid Kia, submitted along with this Petition as Ex1003.

I. STATEMENT OF PRECISE RELIEF REQUESTED

Petitioner requests review under 35 U.S.C. § 311 of claims 1-20 of the ’293 patent and their cancellation in view of the following:

Prior Art References	
Ex1004	U.S. Patent No. 8,345,082 (“Tyso”), issued Jan. 1, 2013, filed Oct. 8, 2009
Ex1005	U.S. Patent Application No. 61/103,588 (“Tyso Provisional”), filed Oct. 8, 2008
Ex1006	User Manual for GNU Image Manipulation Program (dated May 1, 2007) (hereinafter, this User Manual will be referred to as “GIMP”)

The ’293 patent was filed May 26, 2009. Tyso is entitled to an effective filing date of October 8, 2008, by virtue of a claim to priority to U.S. Provisional Patent Application 61/103,588 (Ex1005) (“Tyso Provisional” or “Tyso Prov.”), filed Oct. 8, 2008, and is available as a prior art reference under pre-AIA 35 U.S.C. § 102(e). As explained below, there is no meaningful difference between Tyso and the Tyso Provisional. The Tyso Provisional contains the same or substantially the same disclosure as Tyso and supports all claims of Tyso.

Ex1003, ¶43. In the detailed discussion of the Grounds below, Petitioner cites to both Tyso and the corresponding portions of the Tyso Provisional to establish that Tyso and all the relied-upon disclosures therein are available as of the provisional application's filing date. While Petitioner cites to both Tyso and the corresponding portions of the Tyso Provisional, the asserted Ground 1 of Unpatentability is based on Tyso having the priority date of Tyso's Provisional, and Ground 2 is based on Tyso (having the priority date of Tyso's Provisional) in combination with "GIMP."

GIMP, which is a User Manual for GNU Image Manipulation Program, was publicly available as of at least May 21, 2007, and is prior art under at least 35 U.S.C. § 102(a) and (b). GIMP has a "Written By" date of May 1, 2007 (see Ex1006, 3), and was indexed on a documentation page for the GIMP image processing project at docs.gimp.org/download.html at least as early as May 21, 2007. Ex1008 (archive.org page showing availability of 626 page, 19 MB English version)

<https://web.archive.org/web/20070521020845/http://docs.gimp.org/download.html>¹;

¹ The Internet Archive assigns a URL on its site to the archived files in the format [http://web.archive.org/web/\[Year in yyyy\]\[Month in mm\]\[Day in dd\]\[Time code in hh:mm:ss\]/\[Archived URL\]](http://web.archive.org/web/[Year in yyyy][Month in mm][Day in dd][Time code in hh:mm:ss]/[Archived URL]) aka an "extended URL". Thus, the extended URL

see also Ex1009 (excerpt of archive.org page showing availability of Ex1007 on May 12, 2007 at docs.gimp.org/en.pdf). Ex1003, ¶146.

Ground of Unpatentability	
1	Ground 1: Claims 1-20 are obvious under 35 U.S.C. § 103 in view of Tyssso
2	Ground 2: Claims 1-20 are obvious under 35 U.S.C. § 103 over Tyssso in view of GIMP

II. THE '293 PATENT

A. Overview

The '293 patent is entitled, "Video Superposition for Continuous Presence," and describes a system for creating video conferencing displays that appear to place multiple participants in the same room, ('293 patent,) in contrast to other arrangements where participant images are reduced and stacked in rows/columns. '293 patent, 2:23-29, 1:20-28.

Per the abstract of the '293 patent, the system receives two or more real-time video streams, each with a video frame having a subject image and a background

<http://web.archive.org/web/19970126045828/http://www.archive.org/> would be the URL for the record of the Internet Archive home page HTML file (<http://www.archive.org/>) archived on January 26, 1997 at 4:58 a.m. and 28 seconds (1997/01/26 at 04:58:28). *See* Ex1009.

image. The system combines the video frames of the video streams into combined video frames to produce a combined real-time video stream. The system positions a subject image of the first video stream in an anterior (front) portion of the combined frame, and positions a subject image of the second video stream in a posterior (back) portion of the combined frame. The '293 patent can scale the superimposed subject images of the combined video frame to “near life-size,” and arranges the subject images so they appear to be sitting in the same room, one in front of another, like people seated in stadium seating. '293 patent, 5:20-35.

The '293 patent discloses that it utilizes unused background space in a video frame to maintain near “life-size” participant images while achieving continuous presence. '293 patent, 5:65-6:3.

FIG. 4 of the '293 patent, reproduced below, illustrates a method 720 of combining the subject images of video frames into a combined video frame ('293 patent, 6:53-67):

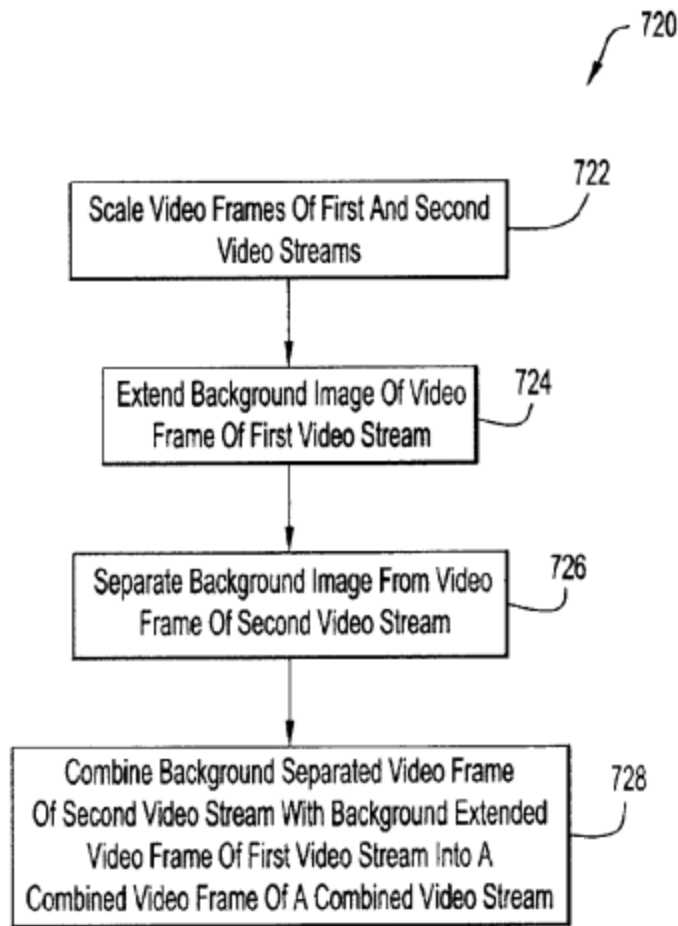


FIG.4

Video frames of a first video stream (FIG. 5A, '293 patent, 7:5-8) and a second video stream (FIG. 5B, '293 patent, 7:8-9) captured at substantially the same time are scaled in step 722.

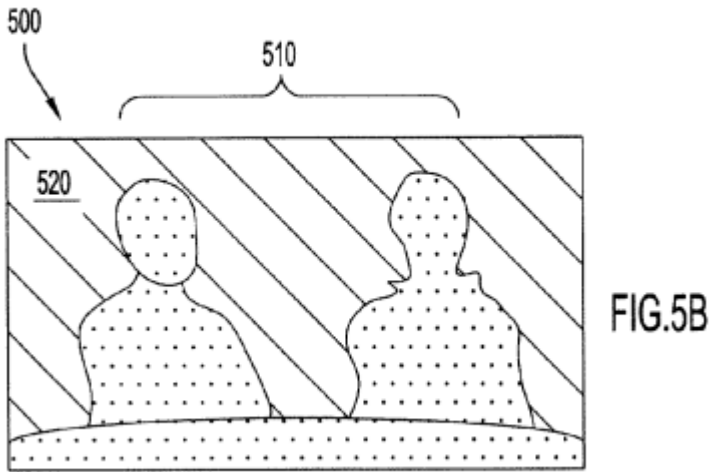
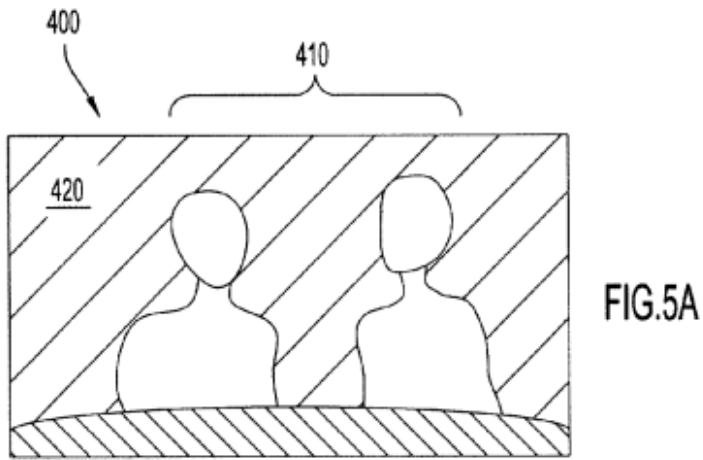


FIG. 5C and 5D illustrate the scaling function 722 applied to video frames of the first and second real-time video streams:

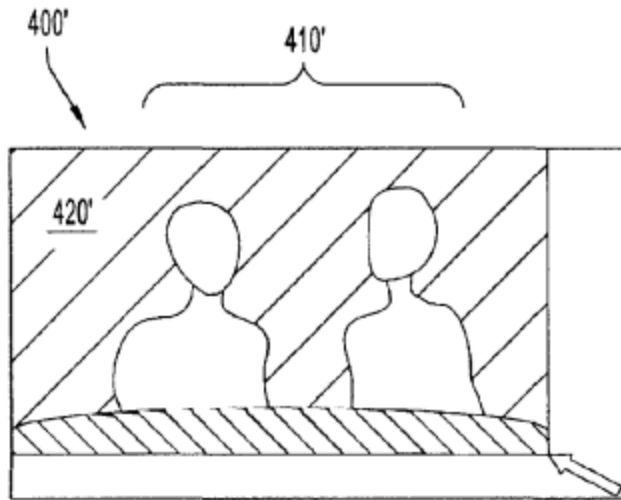


FIG.5C

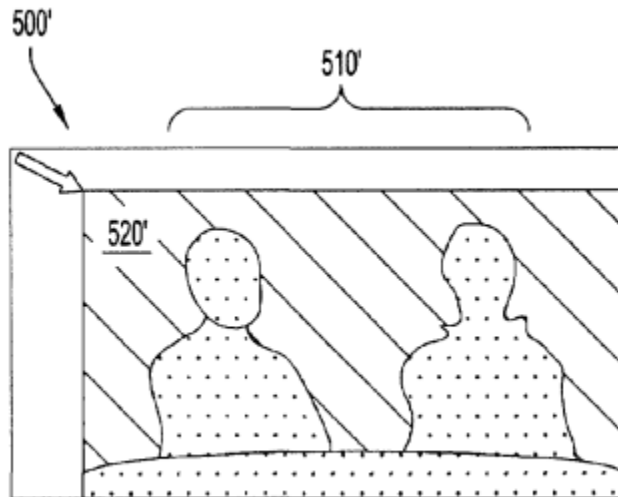
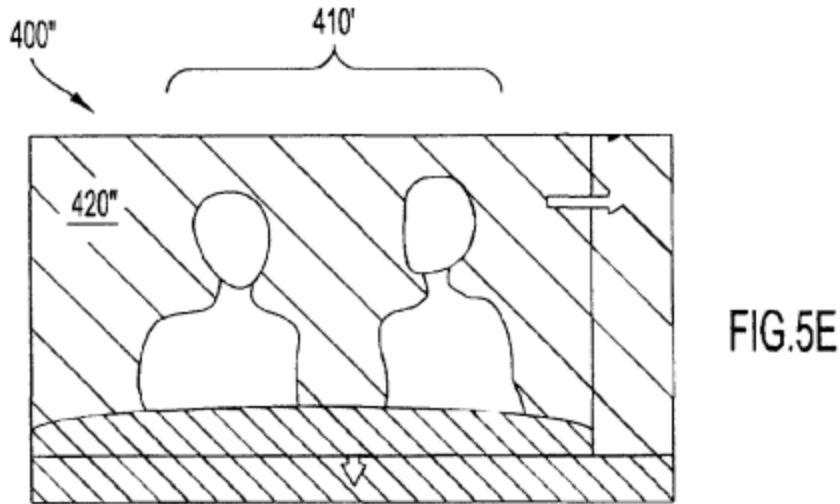


FIG.5D

Scaled video frames include a scaled subject image (410' and 510') and scaled background images (420' and 520'). '293 patent, 7:13-19.

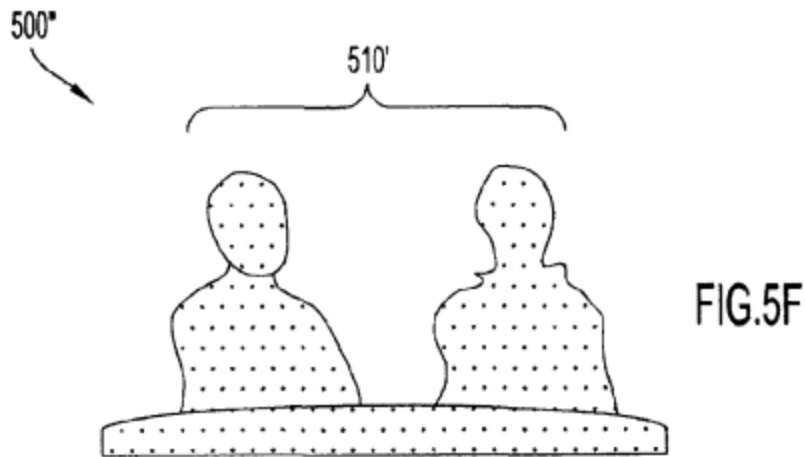
In step 724, the background image of the video frame of the first video stream is extended. FIG. 5E illustrates the background extension function 724 applied to the scaled background image 420' of the scaled video frame 400' of FIG. 5C ('293 patent, 7:20-23):



The '293 patent discloses that “[e]xtending of the background image ... may be performed via a video enhancing technique (e.g., inpainting).” '293 patent 7:27-29.

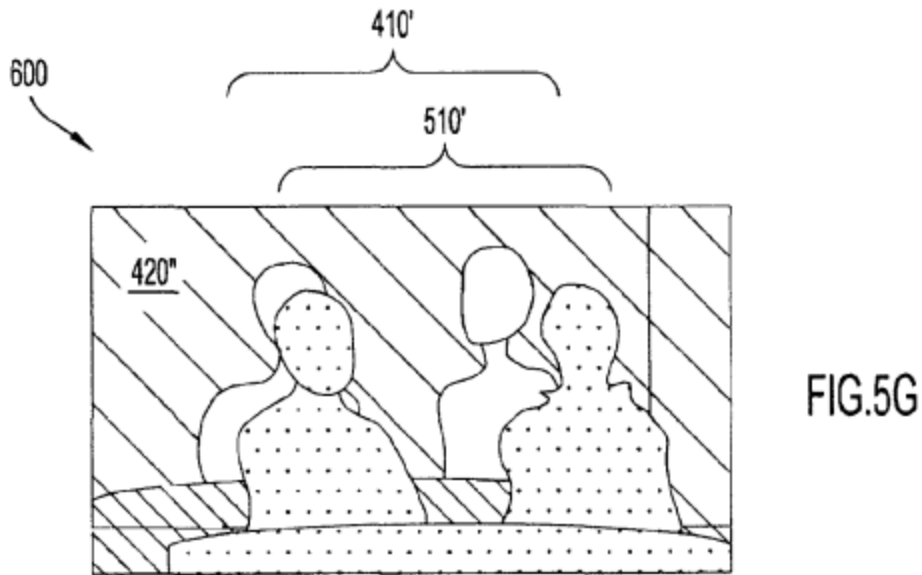
The '293 patent further discloses additional details of “inpainting” technique in various arrangements such as extending the background image in horizontal or vertical directions. '293 patent 7:30-52. In another embodiment, a supplemental background image can be used instead of a background image from one of the video streams. '293 patent, FIGS. 6, 7A-7H, '293 patent, 8:54-9:40.

Turning back to FIG. 4, in step 726, the background image of second video stream is separated from a video frame of the second video stream. FIG. 5F illustrates a background separated video frame 500'' of the second real-time video stream that results from the background separation function 726 ('293 patent, 7:53-55):



The background image 520 is separated from the scaled subject image 510' of the scaled video frame 500' of the second video stream, resulting in a background separated video frame 500" of the second real-time video stream. '293 patent, 7:55-59.

In step 728, the background separated video frame of the second video stream and the background extended video frame of the first video stream are combined into a combined video frame of a real-time combined video stream. FIG. 5G illustrates a combined video frame that results from the combining function 728, where the background extended video frame 400" of the first video stream shown in FIG. 5E is combined with the background separated video frame 500" of the second video stream shown in FIG. 5F into a combined video frame 600 ('293 patent, 7:60-65):



The combined video frame 600 includes the scaled subject image 510' of the second video stream in an anterior portion, and the scaled subject image 410' and the extended background image 420'' of the video frame of the first video stream in a posterior portion. '293 patent, 7:66-8:7.

III. LEVEL OF ORDINARY SKILL

A person of ordinary skill in the art (“POSITA”) relevant to the '293 patent would be familiar with the technical field relating to video conferencing. Ex1003, ¶26. A POSITA at the time of the invention would have had at least a computer science, computer engineering, or electrical engineering degree, or a related field, and two years of experience in image or video processing. Work experience can substitute for formal education and additional formal education can substitute for work experience. Ex1003, ¶26.

IV. CLAIM CONSTRUCTION

The Board construes claims under *Phillips v. AWH Corp.*, 415 F.3d 1303 (Fed. Cir. 2005) (en banc). 37 C.F.R. §42.100(b) (2018). Under this standard, terms receive their plain and ordinary meaning as understood by one of ordinary skill in the art, consistent with the disclosure and prosecution history. *Phillips*, 415 F.3d at 1314–19. Claims should only be construed to the extent necessary to resolve a controversy. *Nidec Motor Corp. v. Zhongshan Broad Ocean Motor Co.*, 868 F.3d 1013, 1017 (Fed. Cir. 2017). Except as noted below regarding the order of claimed steps, no claim terms need to be construed by the Board at this time, and all should be given their ordinary meanings to the extent they are not indefinite.

Elements [1.c], [3.a], [10.d], [11.b], [13.c], and [15.a] in claims 1, 3, 10, 11, 13, and 15 each requires scaling the video frames of a video stream and repositioning “the resulting pictures.” *See, e.g.*, claim 1 (“scaling the video frames of the first video stream and repositioning in a first direction the resulting pictures in the video frames of the first video stream **to produce a first sequence of scaled video frames**”) (emphasis added). Elements [1.d], [3.b], [10.e], [11.c], [13.d], and [15.b] in claims 1, 3, 10, 11, 13, and 15 then recite removing (or extending) the background image in the scaled video frames to produce first background separated video frames. *See, e.g.*, claim 1 (“removing the background image in **the**

first sequence of scaled video frames to produce first background separated video frames to be displayed in the anterior portion of the combined frames of the combined video stream”) (emphasis added).

To give meaning to the antecedence used in the emphasized claim language above, a POSITA would interpret the claims to have the order set out in the claims so that the “scaling” of received frames would occur before “removing” or “extending” the background of a frame. Ex1003, ¶42.

V. CLAIMS 1–20 ARE UNPATENTABLE OVER THE PRIOR ART

A. Ground 1: Tyso Renders Obvious Claims 1-20

Tyso is entitled, “System and Associated Methodology for Multi-Layered Site Video Conferencing,” and just like the ’293 patent, describes a video conferencing that displays multiple participants from different locations on a single monitor using a dynamic layering approach. Tyso, 1:15-19; Tyso Prov., 1:2-4. Also like the ’293 patent, Tyso specifically addresses problems with then conventional video conferencing systems where participants are displayed in windows “as opposed to being in the same room.” Tyso, 1:30-40; Tyso Prov., 1:27-32. Tyso also addresses a situation where participants are scaled differently, thus causing people at greater-populated sites to appear smaller in the combined video. *Id.*; see also Ex1003, ¶45.

The dynamic layering approach of Tysso separates participant images from their corresponding backgrounds and merges them to make all participants appear the same size on the monitor, as if they were all located in the same room. See Tysso, Abstract, 2:63-66; Tysso Prov., 2:31-33. FIG. 1 of Tysso is a diagram illustrating a process of combining images of participants situated at different locations to display the participants of a video conference as if located in one room (Tysso, FIG. 1, 2:39-43; Tysso Prov., FIG. 1, 3:1-2):

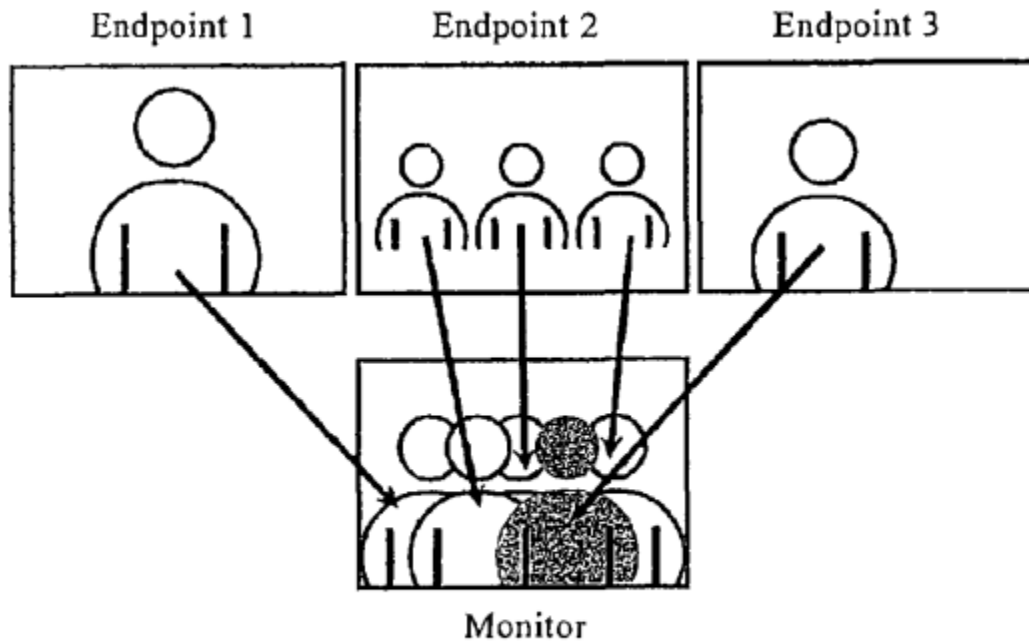


Fig. 1

In FIG. 1, endpoints 1 and 3 include only one participant and are sized differently than endpoint 2, which includes three participants. The arrows in FIG. 1 indicate that the participant images from endpoints 1-3 are “scaled” to have a same size and are merged by overlapping and aligning them for display on Monitor. Tysso, 3:22-

30; Tyso Prov., 5:6-14. Participant image data is extracted from the received video data in order to separate the participant image data from background image data included in the video data. Tyso, 5:11-17; Tyso Prov., 4:36-5:2. Thus, the processed participant images are merged “wherein the participant images are overlapped.” Tyso, 5:17-20; Tyso Prov., 5:3-4 (“participant images are also overlapped and aligned horizontally so that participants appear to be located in a same room when displayed on monitor 4.”); Ex1003, ¶¶45-46.

As shown in the chart below and explained by Dr. Kia, the claims of Tyso are supported by the Tyso Provisional (Ex1003, ¶47):

Tyso Claim	Examples of Support in Tyso Provisional
1. A method to cause a video conferencing system to display a plurality of participants, comprising:	2:10-25; 7:1-19.
receiving, at a multipoint control unit, video data including sound and image data from a plurality of endpoints connected to the multipoint control unit, the video data corresponding to participants located at each of the plurality of endpoints, at least one of the endpoints capturing video data corresponding to more than one participant;	2:13-16; 7:7-8; Fig. 1 (endpoint 2).
extracting, at the multipoint control unit, participant image data from the sound and image data included in the video data;	6:37-38.

Tysso Claim	Examples of Support in Tysso Provisional
processing, at the multipoint control unit, the extracted participant image data to have a same appearance;	3:32-35; 4:11-14.
combining, at the multipoint control unit, the processed participant image data to generate a merged image data of participants from other endpoints, the merged image data excluding surrounding background image data corresponding to each of the participants, images of the participants overlapping in the merged image data;	Fig. 1; 3:15-17; 3:32-35; 5:3-5.
transmitting the merged image data from the multipoint control unit to a respective endpoint; and	Fig. 1; 6:26-30.
displaying, at a monitor of the respective endpoint, the merged image data received from the multipoint control unit, the monitor displaying participants from other endpoints as being located in a same room.	Fig. 1; 6:26-30.
2. The method according to claim 1, further comprising: identifying, at the multipoint control unit, an endpoint with a highest sound level as an active endpoint.	7:21-24.
3. The method according to claim 1, wherein the extracted image data corresponds to a body image of the participants without corresponding background imagery.	7:32-36.

Tysso Claim	Examples of Support in Tysso Provisional
<p>4. The method according to claim 1, wherein the participant image data is extracted from the video data using a synthetic blue or green screen and an algorithm to analyze background imagery and image depth information.</p>	<p>8:6-9.</p>
<p>5. The method according to claim 1, wherein the participants are displayed aligned horizontally on the monitor, the horizontal alignment and overlapping dynamically varying according to changes in a number of participants in the video conference.</p>	<p>8:11-14.</p>
<p>6. The method according to claim 1, wherein the processing of the extracted participant image data includes determining mean values for image size, color balance, contrast and intensity based on the extracted participant image data, and based on the mean values determined scaling each participant image to a same size, adjusting each participant image to a same color balance and adjusting each participant image to a same contrast and intensity.</p>	<p>8:16-22.</p>
<p>7. The method according to claim 2, wherein a participant image from the active endpoint is highlighted on the monitor.</p>	<p>8:24-25.</p>

Tysso Claim	Examples of Support in Tysso Provisional
8. The method according to claim 7, wherein the participant image from the active endpoint is highlighted with borders.	8:27-29.
9. The method according to claim 2, wherein participant images from endpoints other than the active endpoint are softened.	8:31-33.
10. The method according to claim 2, wherein a participant image corresponding to the active endpoint is displayed in front of other participant images on the monitor.	8:36-37.
11. The method according to claim 2, wherein a participant image corresponding to the active endpoint is faded out according to a predetermined delay, and a participant image corresponding to a next active endpoint is faded in according to the predetermined delay.	9:1-4.
12. A video conferencing system for displaying a plurality of participants as being located in a single room, comprising: [...].	Claim 12 is a system claim reciting components configured to perform operations substantially similar to steps claimed in claim 1. Thus, examples of support for claim 12 are identical to those cited for claim 1.

Tysso Claim	Examples of Support in Tysso Provisional
<p>13. The system according to claim 12, wherein the plurality of endpoints are configured to extract the participant image data before transmitting the video data to the multipoint control unit.</p>	<p>4:6-10.</p>
<p>14. The system according to claim 12, wherein the MCU separates the extracted participant image data from background image data.</p>	<p>See claim 1; 4:2-5; 6:37-38.</p>
<p>15. The system according to claim 14, wherein the image data is extracted by applying a blue or green screen behind each participant and analyzing corresponding depth information.</p>	<p>See claim 4.</p>
<p>16. A multipoint control unit device for a video conferencing system comprising: [...].</p>	<p>Claim 16 is a device claim reciting components configured to perform operations substantially similar to steps claimed in claim 1. Thus, examples of support for claim 16 are identical to those cited for claim 1 above.</p>
<p>17. The multipoint control unit device according to claim 16, wherein the extracting unit extracts participant body image data from surrounding background image data, the extracted participant body image data being the extracted participant image data.</p>	<p>See claims 1, 13, 14.</p>

Tysso Claim	Examples of Support in Tysso Provisional
18. An endpoint device in a video conferencing system, comprising: [...].	Claim 18 is a device claim reciting components configured to perform operations substantially similar to steps claimed in claim 1. Thus, examples of support for claim 18 are identical to those cited for claim 1 above.
19. The endpoint device according to claim 18, wherein the extracting unit extracts participant body image data from surrounding background image data, the extracted participant body image data being the extracted participant image data.	See claim 3.
20. The endpoint device according to claim 18, wherein a blue or green screen is placed behind the participant, and the extracting unit extracts the participant image data using a predetermined algorithm to analyze image depth information.	See claim 4.
21. A non-transitory computer-readable medium storing computer-readable instructions thereon, the computer-readable instructions when executed by a computer cause the computer to perform a method comprising: [...].	Claim 21 recites operations substantially similar to steps claimed in claim 1. Thus, examples of support for claim 21 are identical to those cited for claim 1 above.

1. Independent Claim 1: [1.pre] “A method comprising”

Tysso discloses a method. Tysso, 7:4-31; Tysso Prov., 2:10-25; 7:1-19;

Ex1003, ¶48.

- a) **[1.a] “receiving at least first and second real-time video streams, each of the first and second real-time video streams comprising video frames containing a picture comprising a subject image and a background image;”**

Tysso discloses receiving real-time video streams with video frames containing a picture comprising a subject image and a background image. As explained in the Tysso Provisional:

“a method for presenting a plurality of participants in a video conference on one monitor in such a way that *participants located at at least three endpoints* seems to be located in the same room, and where the person currently talking is highlighted on the monitor. The method is characterized by the following steps performed in an MCU connected to said endpoints through a network: - *receiving video signals comprising sound- and image data of the participants at each endpoint;*”

Tysso Prov., 2:10-16 (emphasis added); *see also id.*, 3:21-22; Tysso, 2:1-7. Tysso discloses processing “received video data” and “extracts participant image data from the video data received” (Tysso Prov. 3:21-30, Tysso, 5:8-11) indicating that Tysso has video streams with video frames having subject images and background images. Also, Tysso discloses sending image frames with both participant images and corresponding background to Tysso’s multipoint control unit (“MCU”) (Tysso Prov., 4:2-5; 6:37-38; Tysso, 5:23-32) or only participant images (Tysso Prov., 4:6-10; Tysso, 5:39-46); Ex1003, ¶49.

- b) [1.b] “combining the subject images of corresponding video frames of the first and second video streams into a combined frame of a combined video stream such that the subject image of the first video stream is positioned in an anterior portion of the combined frame and the subject image of the second video stream is positioned in a posterior portion of the combined frame, wherein the combining comprises:”

A POSITA would understand Tysson to disclose combining the subject images of corresponding video frames of the first and second video streams into a combined frame of a combined video stream as claimed. Ex1003, ¶¶50-64.

As shown in Fig. 1 of the Tysson Provisional and Tysson, Tysson’s MCU uses a dynamic layering technique to combine participants (the subject images) of corresponding video frames of video streams into a combined frame of a combined video stream in an overlapping manner. Because the dynamic layering technique allows the participants to be overlapped in the combined frame, a participant in one video stream can be positioned in “an anterior portion” of the combined frame and a participant from another video stream can be positioned in a “posterior portion” of the combined frame.

Specifically, the Tysson Provisional discloses, “An object of the present invention is to provide a method and a device for presenting several participants located at different endpoints on a single monitor by *displaying a dynamic layered multi-site.*” Tysson Prov., 1:2-4, 2:7-9 (emphasis added); Tysson, 1:15-20. Tysson achieves this by “merging video images from each end point in such a way that the

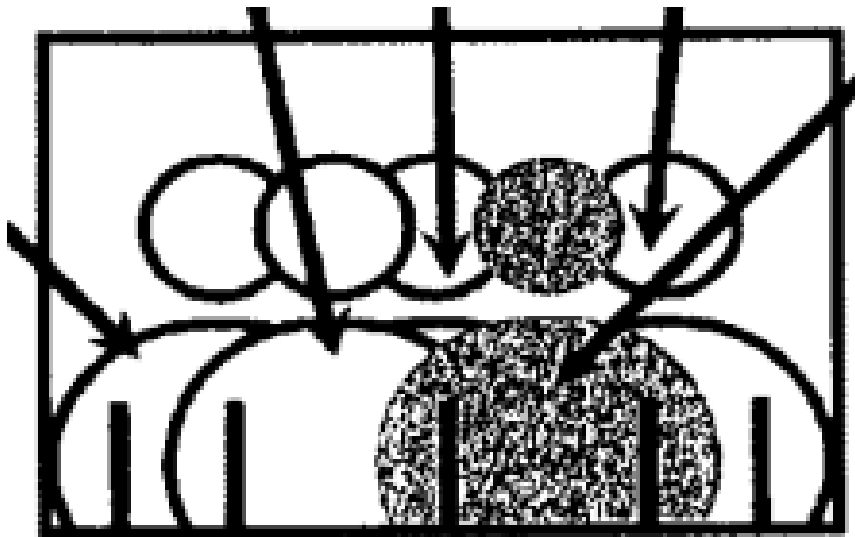
participants appear to be sitting in the same room.” Tyso Prov., 2:31-33; Tyso, 2:63-67. Furthermore, “[t]he resulting images of the participants will then be presented aligned horizontally and ***overlapped on a monitor***. This process is performed dynamically according to the number of participants in the video conference.” Tyso Prov., 5:3-5 (emphasis added); Tyso, 3:27-30, 7:42-46.

As explained by Dr. Kia (Ex1003, ¶53), it was well within the knowledge and skill of a POSITA (and taught by Tyso) that presentation of multiple objects from different sources can benefit from a visual equalizing process where objects are scaled to be of similar size but also can be equalized in contrast, color balance and intensity as well to be visually pleasing. Tyso Prov. 4:36-5:14; Tyso 3:25-33. Tyso recognized the relationship between scale and presentation of large number of objects in a limited space. When visualizing multiple objects, it can be advantageous to scale the objects to fit into a limited space. At a certain level of scaling, where objects are made smaller by scaling them down, the visual acuity of the represented objects diminishes, and the objects may not be recognizable by the viewer. An option is to overlap the objects where the exposed portions provide detailed information to the viewer while some of the objects are covered by others. This approach is addressed in the Tyso Provisional at 5:6-14 and in Tyso at 3:28-33. Tyso discloses “[t]he single monitor shows all participants aligned horizontally, scaled to the same sized and possibly overlapped, giving the

impression that they are located in the same conference room.” Tyso Prov. 5:11-14; Tyso 3:28-30.

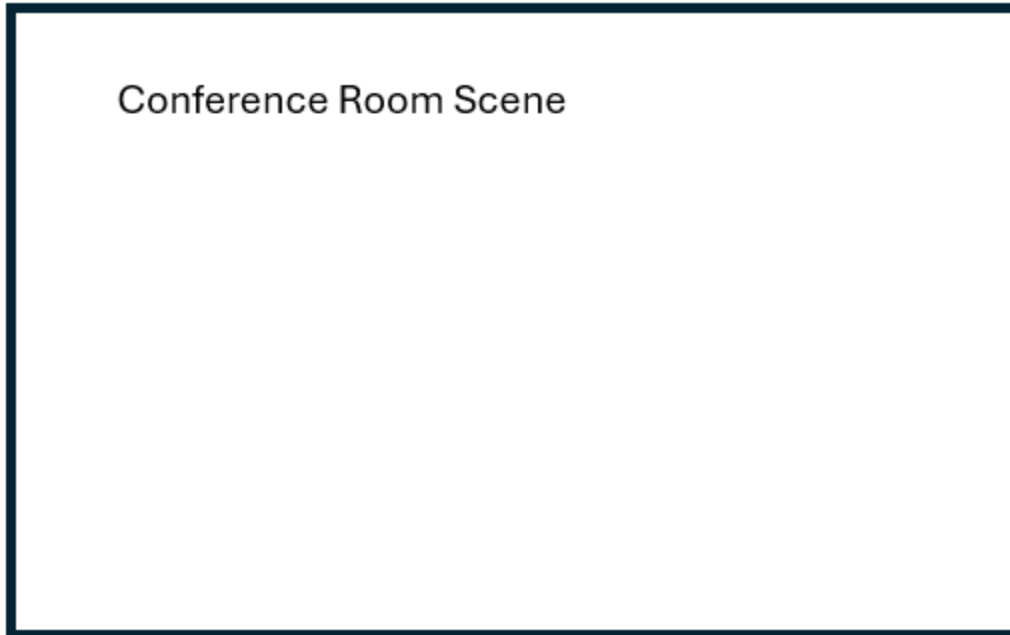
Accordingly, Dr. Kia provides an example (Ex1003, ¶¶54-63) of how a POSITA would understand Tyso’s process of creating the combined frame shown in Tyso’s FIG. 1 with respect to a layer of Tyso’s dynamic layered multi-site for one endpoint. This example is provided for Endpoint 1, but it is understood per Dr. Kia that a POSITA would have understood that similar processing would occur for Endpoints 2 and 3. Ex1003, ¶54.

First, a POSITA would recognize that Tyso creates a combined frame for display. To achieve this, Tyso places layers for the participants in each Endpoint into a frame forming a base for the final combined image, as shown below:



As explained in Tyso, that base frame could have a lower layer, such as a conference room scene. Tyso Prov. 5:11-14 (“The single monitor shows all

participants aligned horizontally, scaled to the same sized and possibly overlapped, giving the impression that they are located in the same conference room.”); Tyso 3:28-30.



For Endpoint 1, Tysso's MCU would receive a frame from Endpoint 1 with participant image and a background, e.g., a green screen background:



Tysso discloses removing the existing background from Endpoint 1, which could be achieved by changing the background pixels in Endpoint 1 with alpha

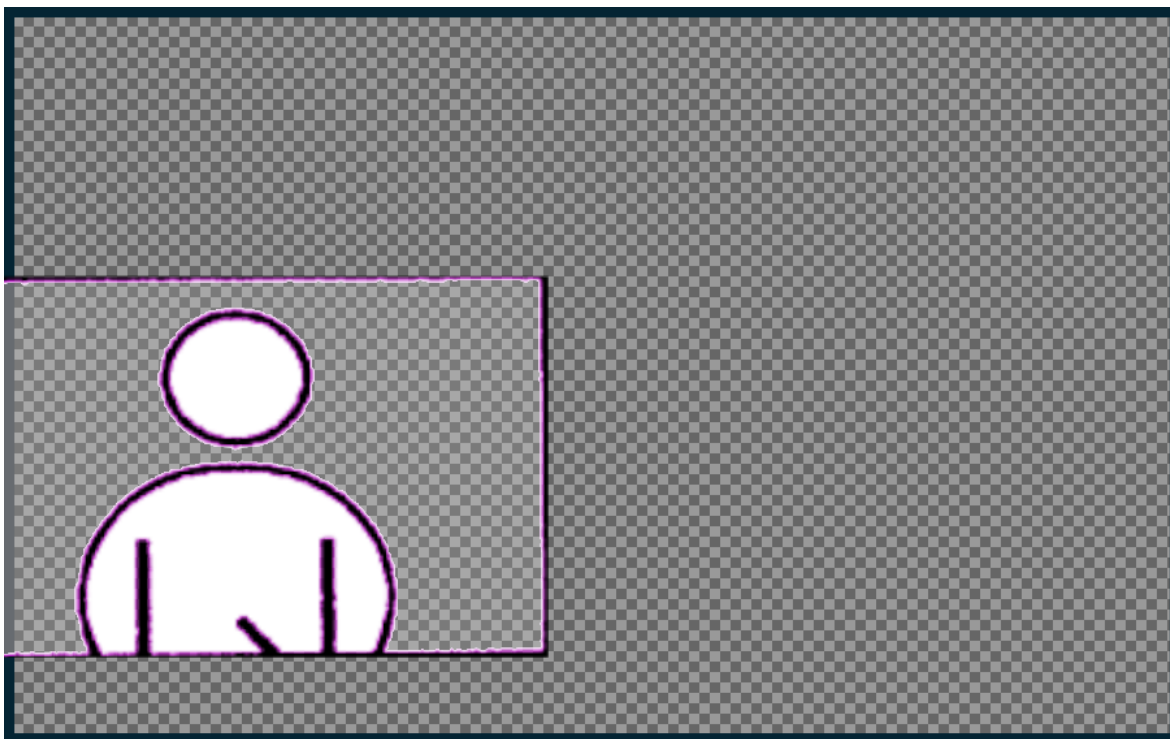
channel pixels, to create transparent portions around the image:



Tysso then discloses scaling Endpoint 1 before placing it in the background layer, which could be achieved by changing the vertical and horizontal dimensions of the frame:

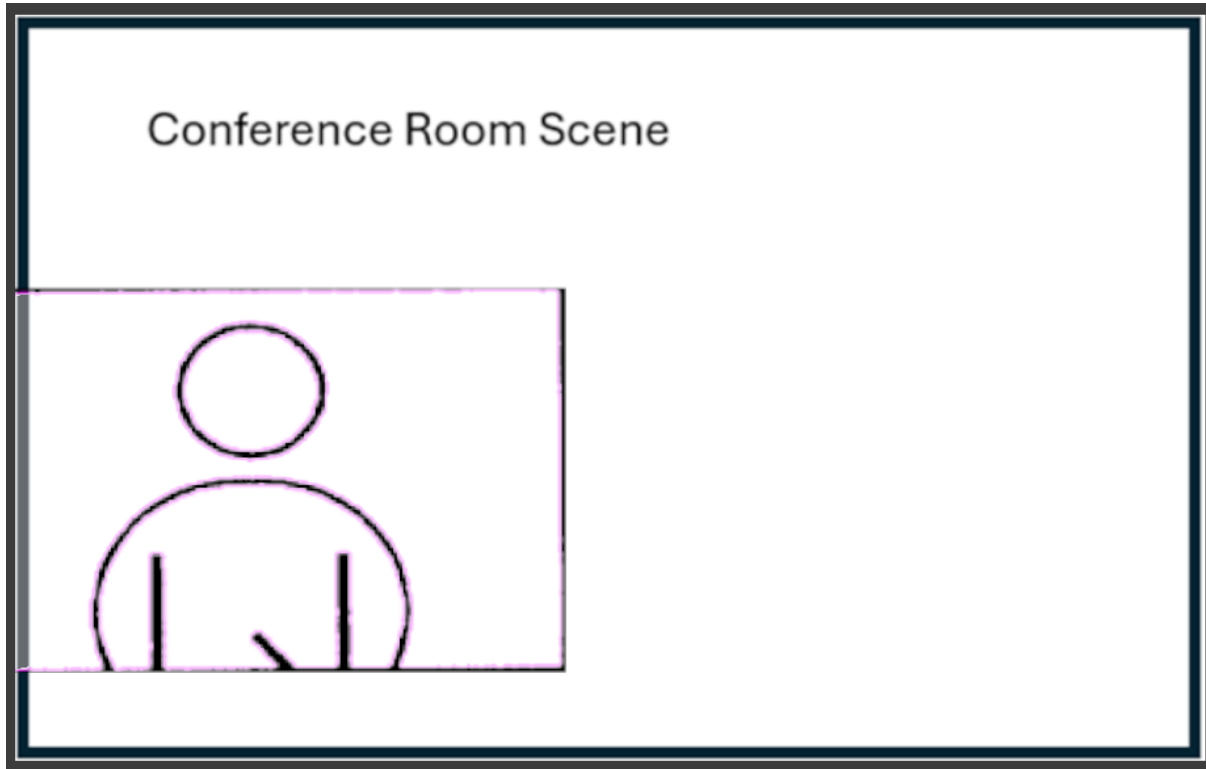


Due to the use of an alpha channel, the pixels surrounding the resized received frame in the layer for the received frame would also be made transparent to “extend” the background of the frame to match the size of the base frame, if necessary, for the combined image when repositioning the participant in the frame (the purple and black frame around the image is provided for illustration and would not be present as each background pixel would be set as transparent).

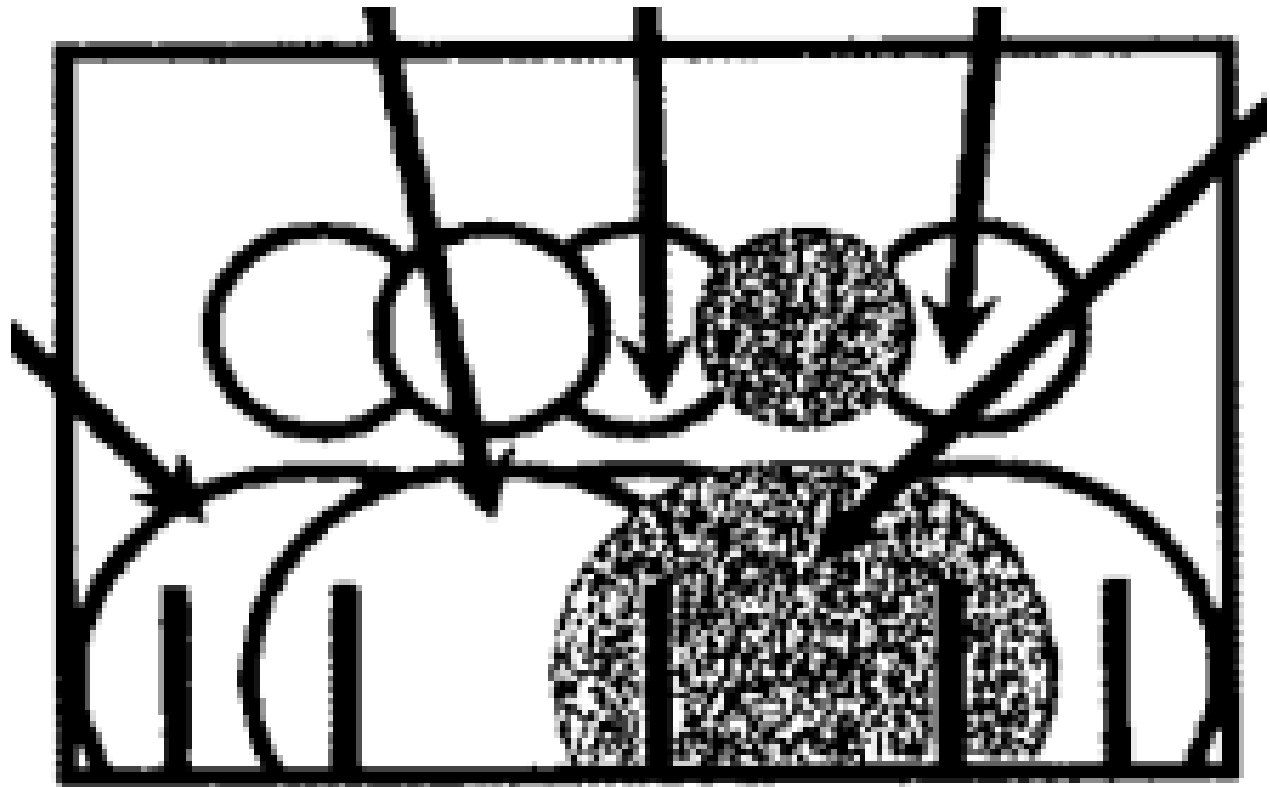


With the use of transparency in the alpha channel of the received layer for Endpoint 1, placing the layer over the base layer for the combined image would result in that participant image being shown in the conference room scene (with the

outline of the resized and repositioned received frame shown for clarity):



The layering process would proceed to build the combined frame until all of the Endpoints and participants are processed:



While this exemplary process described by Dr. Kia (*see* Ex1003, ¶¶54-63) would remove the existing background (e.g., green screen) of the layer for the rearmost Endpoint, a POSITA would understand that the background of the rearmost Endpoint layer could alternatively be retained or re-added in a similar manner (with a suitable filler of transparent pixels or generic background pixels to account for a reduction in size of received frames).

Further, a POSITA would understand that the challenged claims do not contain a requirement concerning any degree of separation between the recited

“anterior” and “posterior” portions. Ex1003, ¶¶64. Nevertheless, as Tyso shows the visual effect of layering to align all participants horizontally, a POSITA would have found it obvious, once images are provided in layers, to arrange those layers with any number of intermediate layers by scaling and shifting the participants to provide additional video effects to provide an illusion of 3D depth in the combined frame. *Id.*

- c) **[1.c] “scaling the video frames of the first video stream and repositioning in a first direction the resulting pictures in the video frames of the first video stream to produce a first sequence of scaled video frames;”**

Tyso discloses, or at least suggests to a POSITA, scaling video frames of first video stream and repositioning in a first direction the resulting pictures in the video frames of the first video stream to produce a first sequence of scaled video frames. *See* Ex1003, ¶¶65-73.

Tyso’s Provisional discloses that “[w]hen the images of the participants in the video conference are extracted from the background image, they will be subjected to further processing in the MCU by processing the image data of the participants in such a way that the images of each participant located at the endpoints are *scaled to the same size*, and adjusted to have similar colour balance, contrast and intensity by comparing the image data received from each endpoint

and finding a mean value for said parameters.” Tysso Prov., 4:36-5:2 (emphasis added); Tysso, 6:3-9.

Figure 1 of Tysso shows the pictures in video streams being repositioned in a first direction in at least two ways (resulting in different “first directions”).

Ex1003, ¶67. With regard to Endpoint 3, a picture of a person is scaled and moved in a direction relatively from the left of the frame to the right. In this way, the person in Endpoint 3 is moved in a “first direction” relatively to the right. Also, in Endpoint 3, the picture of the person is placed in the front-most layer (anterior) of the frame to overlap other pictures. In this way, the first direction that the picture of the person moves is relatively forward (to the anterior), *see* Tysso Prov., 5:6-14; *see also* Tysso, 3:22-33:

Figure 1 illustrates visually the process performed for presenting participants located at different endpoint on a single monitor. As seen from the figure, the participants at endpoints 1 through 3 have different sizes, and different number of participants at the endpoints. When they are presented merged on one monitor, the image data from each end point has been processed and scaled such that all participants appear to have the same size on the monitor. The single monitor shows all participants aligned horizontally, scaled to the same sized and possibly overlapped, giving the impression that they are located in the same conference room.

The other pictures move in other directions, as will be discussed further with regard to claim 3 below. For example, with regard to Endpoint 1, a picture of a person is scaled and moved from the center of the frame to the left of the frame, and placed in a back layer (posterior) of the frame, thus allowing the picture of the person to be overlapped by other pictures. With regard to Endpoint 1, the person's picture is moved in a second direction relatively to the left or in a second direction relatively backwards. *See* Tyso Prov., 5:6-14; *see also* Tyso, 3:22-33; Ex1003, ¶68.

As explained above for element [1.b], Tyso's description of removing a background of a frame and scaling an image would disclose, or at least suggest, scaling a received video frame. To the extent that the claim is interpreted to require a specified order of steps, where received frames are scaled and then background processing occurs to remove the received background, such operation would have been obvious. *See* Ex1003, ¶¶69-73.

As Dr. Kia explains, in Tyso, it is not critical where and when a background is removed from an image, and Tyso allows the background to be removed prior to transmission to the MCU or at the MCU, as explained above. Tyso Prov., 4:2-10; 6:37-38; Tyso, 5:23-32, 5:39-46. Specifically, Tyso's endpoints can provide full frames to the MCU, as noted above, or Tyso's endpoints can provide frames with backgrounds removed. Similarly, a POSITA would find no criticality in

removing a background before or after a scaling operation, as the steps of removing a background (that is determining background pixels and setting them to alpha pixels) would merely be an engineering design choice. Ex1003, ¶¶70-71.

The decision of when and where to remove a background presents engineering design choices that would have been known to a POSITA. *Id.* For example, the choice to transmit frames with backgrounds to the MCU comes at the expense of bandwidth to transmit the frames, which a POSITA could weigh against whether to provide processing power at an endpoint to remove backgrounds before sending them to the MCU. Similarly, the choice to remove the background of a full-resolution frame would allow a higher quality of background removal, which a POSITA could weigh against the less processing power needed to remove the background of a scaled down frame. Thus, a POSITA would not consider such choices concerning the order of processing to remove background from a frame to create a non-obvious distinction given these well-known processing tradeoffs. *Id.*

Accordingly, a POSITA would have found it to be an obvious alternative for scaling the pictures in the frames by either (1) extracting a background from the frame and then scaling the remaining picture containing the subject to the desired size (as expressly taught in Tyso) or (2) scaling of a frame and then extracting the background, as either equivalent alternative would result in a picture that has been scaled to an appropriate size. Ex1003, ¶72. And Tyso expressly discloses that the

manipulation of the frames to “process and extract image data of participants” can involve multiple sub-steps. Tyso Prov., 3:32-4:5; *see also* Tyso, 5:33-38.

Thus, a POSITA would have known that contents of a frame could be scaled by scaling a frame to decrease (or increase) the size of the contents of the frame to a certain size or by extracting the contents of the frame and scaling the extracted content. Thus, there being a limited number of known alternatives to scale the contents of a video frame containing a subject and a background, namely scaling either the entire video frame or scaling a frame with only the subject, either choice would have been obvious.

d) [1.d] “removing the background image in the first sequence of scaled video frames to produce first background separated video frames to be displayed in the anterior portion of the combined frames of the combined video stream; and”

Tyso discloses, or at least suggests to a POSITA, a step of removing the background image of scaled video frames to produce background separated video frames to be displayed in the anterior portion of the combined frames of the combined video stream. Ex1003, ¶¶74-76. As noted above for element [1.c], Tyso discloses removing the background image in the first sequence of scaled video frames to produce first background separated video frames to be displayed in the anterior portion of the combined frames of the combined video stream.

For example, Endpoint 3’s video frame is shown with its background image removed and scaled to place the video frame for Endpoint 3 in the front layer of the video frame of the merged video frame. As explained in the Tyso Provisional, “[t]he second step is to process and extract image data of participants from their backgrounds. . . . This second step can be performed with several sub-steps including extracting the participants from the image data. For doing this, the MCU processes the received image data by extracting each participant from the background surroundings in the image in such a way that only the bodies of the participants are included in the extracted image data.” Tyso Prov., 3:32-4:5; *see also* Tyso, 5:33-38.

Further, under the equivalent obvious alternatives discussed for element [1.c] above, a POSITA would have found it to be an obvious alternative for scaling the pictures in the frames by either (1) extracting a background from the frame and then scaling the remaining picture containing the subject to the desired size (as expressly taught in Tyso) or (2) scaling of an entire frame and then extracting the background, as either equivalent alternative would result in a picture that has been scaled to an appropriate size. Ex1003, ¶76. There are a limited number of options for scaling the pictures in the frames, and either option would have been obvious to a POSITA. Thus, regardless whether claim 1 of the ’293 patent requires an order of steps of scaling prior to background processing (that is, an order where

removing the background image occurs after scaling), either process would have been obvious in view of Tyso.

- e) [1.e] “superimposing the first background separated video frames onto corresponding ones of the video frames of the second video stream to produce combined video frames of the combined video stream; and”

Tyso discloses, or at least suggests to a POSITA, the step of superimposing first background separated video frames onto corresponding ones of the video frames of second video stream to produce combined video frames of the combined video stream. Ex1003, ¶77. As shown in Fig. 1 of Tyso, Tyso superimposes background separated video frames onto other video frames from other endpoints to produce combined video frames of the combined video stream.

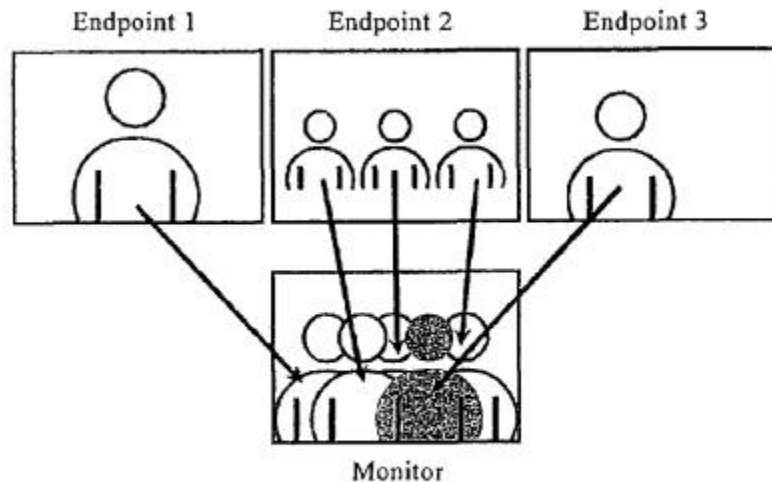


Fig. 1

Tysso, Fig. 1; Tysso, Prov., Fig. 1. Again, as explained in the Tysso Provisional, “participants at endpoints 1 through 3 have different sizes, and different number of participants at the endpoints. When they are presented merged on one monitor, the image data from each end point has been processed and scaled such that all participants appear to have the same size on the monitor. The single monitor shows all participants aligned horizontally, scaled to the same sized and possibly overlapped, giving the impression that they are located in the same conference room.” Tysso Prov., 5:6-14; Tysso, 3:22-33. In Tysso’s Fig. 1, the frames of Endpoint 3 are placed in the front layer of the combined frame in a manner that overlaps pictures in an intermediate layer from Endpoint 2 and the back layer of Endpoint 1.

f) [1.f] “supplying the combined video stream to a video display for displaying the combined video stream.”

Tysso discloses, or at least suggests to a POSITA, supplying the combined video stream to a video display for displaying the combined video stream.

Ex1003, ¶78. Tysso displays the combined video stream on a monitor, as shown in Fig. 1 of Tysso. Tysso Prov., Fig. 2; 1:2-4; 2:7-9; 5:3-17; Tysso, 3:11-14, 3:28-30, 7:42-46.

2. Claim 2: “The method of claim 1, wherein the subject image of video frames of the first video stream comprises images of multiple videoconference participants and the subject image of video frames of the second video stream comprises images of multiple videoconference participants.”

Tysso discloses, or at least suggests to a POSITA, that the subject image of video frames of the first video stream comprises images of multiple videoconference participants and the subject image of video frames of the second video stream comprises images of multiple videoconference participants. Ex1003, ¶79. Tysso discloses that each endpoint may have multiple participants. Tysso Prov., 5:3-9; Tysso, 3:11-14, 3:28-30, 7:42-46.

3. Claim 3: [3.pre] “The method of claim 1, wherein combining comprises:

Tysso discloses, or at least suggests to a POSITA, the claimed method of claim 1 for the reasons stated above and the claimed combining of claim 3 as discussed below. Ex1003, ¶80.

a) [3.a] “scaling the video frames of the second video stream and repositioning in a second direction the resulting pictures in the video frames of the second video stream to produce a second sequence of scaled video frames;”

Tysso discloses, or at least suggests to a POSITA, scaling the video frames of the second video stream and repositioning in a second direction the resulting pictures in the video frames of the second video stream to produce a second sequence of scaled video frames. Ex1003, ¶¶81-82. Tysso scales video frames from each endpoint, including the “second video stream.” For example and as explained above with regard to elements [1.b]-[1.d], Fig. 1 of Tysso and the Tysso Provisional show a reduction in the scale of the image from Endpoint 1.

A POSITA would understand that Tyssó’s reduction in scale would effectively reduce the ratio of the participant image relative to the original frame size, as explained above with regard to element [1.b]. Ex1003, ¶82. Further, and as also explained above with regard to elements [1.b] and [1.c], Tyssó then repositions the picture in the video frames for Endpoint 1 to the back and to the left of the combined frame. Tyssó Prov., 5:6-14; Tyssó, 3:22-33; Ex1003, ¶82. Thus, with regard to Endpoint 1, the person’s picture is moved in a second direction relatively to the left or in a second direction relatively backwards (towards the posterior).

- b) [3.b] “extending the background image in the second sequence of scaled video frames to produce background extended video frames to be displayed in the posterior portion of the combined frames of the combined video stream; and”**

Tyssó at least suggests to a POSITA and therefore renders obvious extending the background image in the second sequence of scaled video frames to produce background extended video frames to be displayed in the posterior portion of the combined frames of the combined video stream. Ex1003, ¶¶83-86. As Dr. Kia explains, Tyssó suggests to a POSITA extending the background image of, e.g., the rearmost Endpoint layer, if its size has been reduced. Ex1003, ¶83. The preferred embodiment of Tyssó can remove all the backgrounds from the source video streams. But a background would still be needed and found desirable by a

POSITA, and a POSITA could select from a number of predictable and finite alternatives, such as using a generic background or using a live background from one of the images, as described above with regard to element [1.b]. Also, the Tyso Provisional explains that “[t]he surroundings can be extracted from an image by applying a blue- or green screen behind the participant(s) that are filmed at the endpoint together with depth information.” Tyso Prov., 9:29-21; Tyso, 5:47-51. A POSITA would understand the “green screen” process to suggest use of supplemental background image that would then be added as a layer to final video image to supplement video frame with a desired background image, such as to show that a participant is in a studio, particular landscape, etc.

As noted above, Tyso already scales the participant image for Endpoint 1. This scaling could be achieved by scaling the entire video frame or the picture in the video frame. Under either approach, the reduction in size of the participant image (or frame) would result in the image size (or frame size) being reduced relative to the original size of the image and frame (or frame) for Endpoint 1.

A POSITA would know that extending a background image would be desirable, and possibly required, because the reduction in size of the image would leave a gap relative to original-size image (if the image were replaced in the frame) or original-size frame (comparing the reduced-size frame to the original-size frame) or the combined frame to be displayed (when it is larger than the size of a

reduced frame, which would be typical when the original-size frame and combined frame are the same size). Ex1003, ¶¶84-85. A POSITA would understand that there are a limited number of ways to address this gap in the reduction in size of an image or frame that would effectively “extend” the background image of the frame. In particular, additional image processing to fill the gap caused by a reduction in image or frame size would increase the size of the background to replace the pixels lost by reduction in scale. Also, as explained regarding element [1.b] above, simply adding additional transparent pixels to allow a frame to be transparent around a participant image would “extend” the background of the frame to allow transparency in all areas of the frame other than the participant image, which would be desirable to allow a predetermined background, such as Tyso’s conference room setting, to be displayed. Similarly, instead of transparent pixels, filler pixels (such as those corresponding to Tyso’s conference room setting) could be added to the rear-most participant layer of Tyso’s combined video image to eliminate gaps in the combined video image. Ex1003, ¶85.

Thus, a POSITA would have understood Tyso to disclose, or at least suggest to a POSITA, using an extended background produced from a reduction in scale of the picture as the background for all of the endpoints, or an extended green screen, to provide a consistent experience across endpoints. To do so would be

consistent with Tyso's aim to give "the impression that they are located in the same conference room." Tyso Prov., 5:13-14; Tyso, 2:59-63; Ex1003, ¶86.

- c) **[3.c] "superimposing the first background separated video frames onto corresponding ones of background extended video frames to produce combined video frames of the combined video stream."**

Tyso discloses, or at least suggests to a POSITA, superimposing the first background separated video frames onto corresponding ones of background extended video frames to produce combined video frames of the combined video streams. Ex1003, ¶87. Tyso discloses, or at least suggests to a POSITA, this feature for the reasons stated above with regard to element [1.e].

4. Claim 4: "4. The method of claim 3, wherein the second direction is opposite to the first direction."

Tyso discloses, or at least suggests to a POSITA, the method of claim 3 as discussed above. Tyso discloses, or at least renders obvious, that the second direction is opposite to the first direction. Ex1003, ¶88. For example, and as explained with regard to elements [1.c] and [3.a], the pictures of Endpoint 1 are moved backwards and to the left so that the claimed "second direction" could be satisfied by either the backward movement or the movement to the left and the pictures of Endpoint 3 are moved forward and to the right so that the claimed "first direction" could be satisfied by either the forward movement (opposite to the

backward movement of Endpoint 1) or the movement to the right (opposite to the movement to the left of Endpoint 1). Tyso Prov., 5:6-14; Tyso, 3:22-33.

As explained with regard to element [1.b], this manipulation of participant images in Tyso recognizes the relationship between scale and presentation of large number of objects in a limited space. Tyso's movement of the images relative to the video frame allows more users to be shown in a combined video frame in a clear manner and with minimal obstruction by only overlapping the participant images as necessary. Ex1003, ¶89.

5. Claim 5: [5.pre] “5. The method of claim 1, wherein combining further comprises:”

Tyso discloses, or at least suggests to a POSITA, the method of claim 1 for the reasons discussed above, and also discloses, or at least suggests to a POSITA, the claimed combining as discussed below. Ex1003, ¶90.

a) [5.a] “removing the background image in the video frames of the first and second video streams to produce first and second background separated video frames;”

Tyso discloses, or at least suggests to a POSITA, removing the background image in the video frames of video streams to produce first and second background separated video frames. Ex1003, ¶91. As explained in the Tyso Provisional, “the MCU processes the received image data by extracting each participant from the background surroundings in the image in such a way that only the bodies of the

participants are included in the extracted image data.” Tyso Prov., 4:1-5; Tyso, 5:33-38.

b) [5.b] “generating supplemental background image video frames that comprise a supplemental background image; and”

Tyso renders obvious generating a supplemental background image video frames with supplemental background images. Ex1003, ¶¶92-93. As noted above with regard to element [3.b], the preferred embodiment of Tyso can remove all the backgrounds from the source video streams. But a background would still be needed and found desirable by a POSITA, and a POSITA could select from a limited number of known alternatives, such as using a generic background of a conference room or using a live background from one of the images. *See, e.g.*, Tyso Prov., 9:29-21 (discussing green screen processing); Tyso, 5:47-51.

It would have been obvious to use a generic conference room image as the rear-most layer of the combined video frame as the background for all of the endpoints to provide a consistent experience across endpoints. To do so would be consistent with Tyso’s aim to give “the impression that they are located in the same conference room.” Tyso Prov., 5:11-14; Tyso, 2:59-63; Ex1003, ¶93.

c) [5.c] “superimposing corresponding ones of the first background separated video frames and the second background separated video frames onto corresponding ones of the supplemental background video frames to produce the combined video frames of the combined video stream”

Tysso discloses, or at least suggests to a POSITA, superimposing corresponding ones of the first background separated video frames and the second background separated video frames onto corresponding ones of the supplemental background video frames to produce the combined video frames of the combined video stream. Ex1003, ¶94. As noted above and with regard to elements [1.e] and [5.b], Tysso discloses, or at least suggests to a POSITA, superimposing the first background separated video frames (i.e., frames of having only picture data) onto a generic conference room background so as to give the impression that all participants “are located in the same conference room.” Tysso Prov., 5:11-14; Tysso, 2:59-63.

6. Claim 6: “6. The method of claim 1, wherein the first and second video streams are produced at a same video conferencing site.”

Tysso discloses, or at least suggests to a POSITA, the method of claim 1 for the reasons discussed above, and Tysso discloses, or at least renders obvious, the same first and second video streams are produced at a same video conferencing site. Ex1003, ¶¶95-97.

The '293 patent does not define “site.” It does, however, refer to videoconferencing as allowing “two or more locations to interact via simultaneous two-way video and audio transmissions.” '293 patent, 1:12-14. Further, the '293 patent states, the components of its video distribution system “may be distributed

over a wide area network (e.g., WAN, internet, etc.) or may be distributed over a local area network (LAN) within *a same facility (e.g., building, campus, etc.)*.”

’293 patent, 3:30-34 (emphasis added). Thus, a POSITA would have understood that a “site” refers to a physical facility, such as an office building. Ex1003, ¶96.

Also, it would be obvious to a POSITA that the video streams in Tyso can and would be allowed to originate from the same building, such as to allow two or more office workers to conduct a meeting in a “virtual” conference room. Ex1003, ¶97. A POSITA would understand that the problems associated with multi-site teleconferencing and utilization of “IP and ISDN based teleconferencing” (Tyso Prov. 3:9-10; Tyso, 3:38-39), such as bandwidth utilization and latency, would not be present or at least not as prevalent in local areas networks such as a single conferencing site. It is well understood that the solutions to the problems associated with bandwidth utilization and latency are easier to solve in a local area than in a wide area, indicating that the POSITA would be well informed in solving a solution for teleconferencing in a single site if a solution was already there for a multi-site configuration. No technical impediment would preclude Tyso’s system from working in a local environment using a local area network (LAN). Ex1003, ¶97.

7. Claim 7: “7. The method of claim 1, wherein the first and second video streams are produced at different video conferencing sites.”

Tysso discloses, or at least suggests to a POSITA, the method of claim 1 for the reasons discussed above, and Tysso discloses, or at least suggests to a POSITA, producing video streams at different video conferencing “sites” (*see* discussion of “site” in the analysis of claim 6). Ex1003, ¶98. The Tysso Provisional discloses, “Figure 1 illustrates a plurality of participants located at different locations merged and presented on a single monitor.” Tysso Prov., 3:1-2; Tysso, 3:11-14. Furthermore, Tysso Provisional discloses utilization of “IP and ISDN-based videoconferencing” (Tysso Prov. 3:9-10) and “Internet” (Tysso Prov. Fig. 2) which are common technologies to interconnect video conferencing devices at different sites, as was well known to a POSITA.

8. Claim 8: “8. The method of claim 1, wherein combining is performed at a site other than where the first and second video streams are produced.”

Tysso discloses, or at least suggests to a POSITA, the method of claim 1 for the reasons discussed above, and Tysso discloses, or at least suggests to a POSITA, combining video streams at a “site” (*see* discussion of “site” in the analysis of claim 6) other than where the first and second video streams are produced. Ex1003, ¶99. As shown in Fig. 2 of the Tysso provisional, the MCU in Tysso may be at a location remote from the endpoints in Tysso:

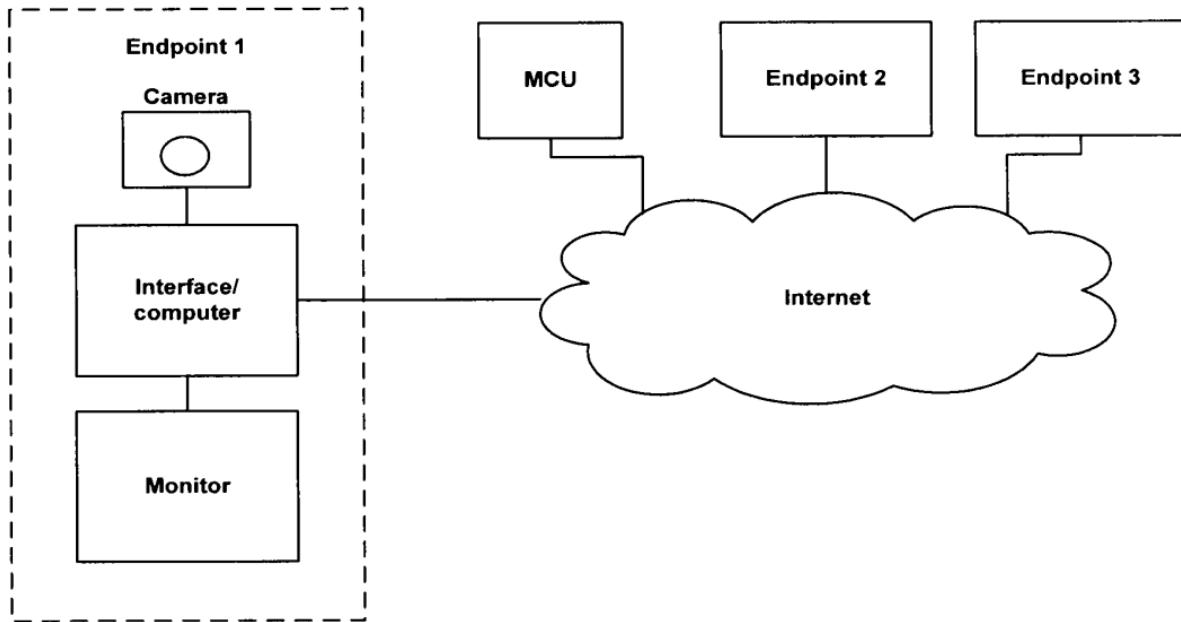


Fig. 2

Tysso Prov., Fig. 2; 6:18-30 (“Each endpoint sends image and sound data of the participants in a video conference to the MCU, and the MCU processes the data before sending the resulting processed video signals to respective endpoints.”); *see also* Tysso, Fig. 2, 3:52-62.

9. Claim 9: “9. The method of claim 1, wherein the subject images of video frames of the first and second video streams are images of video conference participants.”

Tysso discloses, or at least suggests to a POSITA, the method of claim 1 for the reasons discussed above, and Tysso discloses that the subject images of the video frames are images of video conference participants. Ex1003, ¶¶100-01. The Tysso Provisional discloses, “The last step of the inventive method is to present processed video signals on said single monitor at respective endpoints showing video images of the participants at other endpoints than the endpoint receiving the

video signals, and *where all participants in the video conference*, appear removed from their surroundings and seems to be located in the same room.” Tyso Prov., 6:13-17 (emphasis added); Tyso, 5:33-38, 6:28-37.

Further confirming the video conferencing nature of Tyso, Tyso discloses various techniques to focus on certain participants. See Tyso Prov., 6:2-9 (“One way of showing that a certain participant currently is talking is to place that participant, at endpoint marked as active, in front of the others. Another way of highlighting the participant is by making borders around the participant. Yet another way is to soften the other participants that are not currently talking, i.e. at endpoint not marked as active”); Tyso, 6:12-24.

10. Claim 10: [10.pre] 10. An apparatus comprising:

Tyso discloses an apparatus. Tyso, 7:1-31; Tyso Prov., Fig. 2; 6:18-20; 2:10-25; 7:1-19; 9:7-25; Ex1003, ¶102.

- a) **[10.a] “a first memory configured to store data for at least first and second real-time video streams, each of the first and second real-time video streams comprising video frames containing a picture comprising a subject image and a background image; and”**

Tyso discloses, or at least suggests to a POSITA, the claimed first memory configured to store data for at least first and second real-time video streams, each of the first and second real-time video streams comprising video frames containing a picture comprising a subject image and a background image for the reasons

discussed above with regard to element [1.a]. In addition, Tyso discloses, or at least suggests to a POSITA, a “first memory configured to store data” because Tyso’s system includes a computer-implemented processes to interface with Tyso’s MCU. Tyso Fig. 2 at 22, 24 (“Interface/computer”, “MCU”); 3:1-10, 3:63-65; Tyso Prov., Fig. 2 (“Interface/computer”, “MCU”); 3:5-11; 6:21-23; Ex1003, ¶103.

b) [10.b] “at least one data processor configured to:”

Tyso discloses, or at least suggests to a POSITA, the claimed data processor with its MCU for the reasons discussed above with regard to element [1.b]. Tyso Prov., 3:18-20 (“In order to do so, pictures from the different end points are processed by performing a method comprising different steps performed in an MCU connected to said endpoints.”); Tyso, 3:53-55 (“the MCU 22 processes the data before sending the resulting processed data to respective endpoints 1-3”); Ex1003, ¶104.

c) [10.c] “combine the subject images of corresponding video frames of the first and second video streams into a combined frame of a first combined video stream such that the subject image of the first video stream is positioned in an anterior portion of the combined frame and the subject image of the second video stream is positioned in a posterior portion of the combined frame, wherein the at least one data processor is configured to combine the subject images of video frames by:”

Tysso discloses or at least suggests to a POSITA the recitations of element [10.c] for the reasons discussed above with regard to element [1.b]. Ex1003, ¶105.

- d) **[10.d] “scaling the video frames of the first video stream and repositioning in a first direction the resulting pictures in the video frames of the first video stream to produce a first sequence of scaled video frames;**

Tysso discloses or at least suggests to a POSITA the recitations of element [10.d] for the reasons discussed above with regard to element [1.c]. Ex1003, ¶106.

- e) **[10.e] “removing the background image in the first sequence of scaled video frames to produce first background separated video frames to be displayed in the anterior portion of the combined frames of the combined video stream; and**

Tysso discloses or at least suggests to a POSITA the recitations of element [10.e] for the reasons discussed above with regard to element [1.d]. Ex1003, ¶107.

- f) **[10.f] “superimposing the first background separated video frames onto corresponding ones of the video frames of the second video stream to produce combined video frames of the combined video stream.”**

Tysso discloses or at least suggests to a POSITA the recitations of element [10.f] for the reasons discussed above with regard to element [1.e]. Ex1003, ¶108.

11. Claim 11: [11.pre, 11.a] “11. The apparatus of claim 10, wherein the at least one data processor is configured to combine the subject images of video frames by:”

Tysso discloses, or at least suggests to a POSITA, the apparatus of claim 10 for the reasons discussed above, and Tysso discloses, or at least suggests to a

POSITA, the requirements of the apparatus with a data processor configured to combine subject images of video frames as noted above with regard to element [3.pre]. Ex1003, ¶109.

- a) **[11.b] “scaling the video frames of the second video stream and reposition in a second direction the resulting pictures in the video frames of the second video stream to produce a second sequence of scaled video frames;”**

Tysso discloses or at least suggests to a POSITA the recitations of element [11.b] for the reasons discussed above with regard to element [3.a]. Ex1003, ¶110.

- b) **[11.c] “extending the background image in the second sequence of scaled video frames to produce background extended video frames to be displayed in the posterior portion of the combined frames of the combined video stream; and”**

Tysso discloses or at least suggests to a POSITA the recitations of element [11.c] for the reasons discussed above with regard to element [3.b]. Ex1003, ¶111.

- c) **[11.d] “superimposing the first background separated video frames onto corresponding ones of background extended video frames to produce combined video frames of the combined video stream.”**

Tysso discloses or at least suggests to a POSITA the recitations of element [11.d] for the reasons discussed above with regard to element [3.c]. Ex1003, ¶112.

12. Claim 12: [12.pre] “12. A system comprising the apparatus of claim 10, and further comprising:”

Tysso discloses, or at least suggests to a POSITA, the apparatus of claim 10 for the reasons discussed above, and Tysso discloses a system comprising the

apparatus for the reasons stated above with regard to element [10.pre]. Ex1003, ¶113.

- a) **[12.a] “a first video conferencing endpoint configured to communicate with the apparatus, the first endpoint including:”**

Tysso discloses or at least suggests to a POSITA the recitations of element [12.a]. The Tysso Provisional discloses multiple endpoints. *See* Tysso Prov., Fig. 2; Ex1003, ¶114.

- b) **[12.b] “a first video camera configured to generate the first real-time video stream, wherein the subject image of the first real-time video stream comprises an image of a first video conferencing participant at the first endpoint;”**

Tysso discloses, or at least suggests to a POSITA, a system with video cameras configured to generate real-time video streams having subject images being images of video conferencing participants at endpoints. Tysso Prov., Fig. 2, 6:18-30 (“Each endpoint comprises at least a camera with means for recording images and sound from the endpoint, a computer, and a monitor, all connected together through interfaces. The monitor is the means for displaying the other participants in the video conference. The computer is further connected to said MCU that is managing an ongoing conference, and is carrying out the inventive method.”); Tysso, Fig. 2, 3:63-65; Tysso Prov., 6:26-28 (“Each endpoint sends image and sound data of the participants in a video conference to the MCU, and

the MCU processes the data before sending the resulting processed video signals to respective endpoints.”); Tyso, 3:52-62, 6:28-37; Ex1003, ¶115.

- c) **[12.c] “a second video camera configured to generate the second real-time video stream, wherein the subject image of the second real-time video stream comprises an image of a second video conferencing participant at the first endpoint; and”**

Tyso suggests to a POSITA the recitations of element [12.c] for the reasons discussed above with regard to element [12.b]. Ex1003, ¶116.

The level of ordinary skill of a POSITA would include capturing a scene with multiple cameras. Ex1003, ¶117. Thus, a POSITA would have found it obvious to provide more than one camera at an endpoint so as to improve the functionality of the system by capturing participants that could not fit within the field of view of a camera or by capturing participants located in different rooms of a same building and then forwarding all captured streams to the MCU. *Id.*

- d) **[12.d] “a first control unit configured to forward the first and second real-time video streams to the apparatus; and”**

Tyso discloses, or at least suggests to a POSITA, the recitations of element [12.d] for the reasons discussed above with regard to element [12.c]. Ex1003, ¶118.

When a POSITA would have provided more than one camera at an endpoint so as improve the functionality of the system by capturing participants that could

not fit within the field of view of a camera or by capturing participants located in different rooms of a same building and then forwarding all captured streams to the MCU, the POSITA would have found it obvious to forward those video signals to Tysso's MCU with a control unit, as Tysso expressly discloses a control unit for a preferred embodiment having a camera. Tysso Prov., Fig. 1

(“Interface/computer”); Tysso, Fig. 1 (“Interface/computer 24”); Ex1003, ¶119.

The level of skill of a POSITA would include knowledge that computers at the time would often interface with distinct peripherals. Thus, a POSITA would have found that allowing a single computer to interface with multiple cameras to have been obvious in order to avoid the use of duplicate computing resources, such redundant processor, memory, power supply, and other components for interfacing with each camera. Ex1003, ¶120.

- e) **[12.e] “a second video conferencing endpoint configured to communicate with the apparatus, the second endpoint including:”**

Tysso discloses or at least suggests to a POSITA the recitations of element [12.e] for the reasons discussed above with regard to element [12.a]. Ex1003, ¶121.

- f) **[12.f] “a second control unit configured to receive the combined video stream of the apparatus; and”**

Tysso discloses, or at least suggests to a POSITA, the recitations of element [12.f] for the reasons discussed above with regard to elements [1.f] and [12.d] and

as Tyso discloses multiple endpoints for receiving the combined video stream from the MCU. Tyso Prov., Fig. 2; Tyso, Fig. 2; Ex1003, ¶122.

g) [12.g] “a display configured to receive and render the combined video stream from the control unit.”

Tyso discloses, or at least suggests to a POSITA, the recitations of element [12.g] for the reasons stated above with regard to elements [1.f] and [12.f]. Tyso Prov., 6:21-24, Fig. 2; Tyso, 3:63-65, Fig. 2; Ex1003, ¶123.

13. Claim 13: [13.pre] “13. Logic encoded in one or more tangible non-transitory storage media for execution and when executed operable to:”

To the extent that the preamble is limiting, Tyso discloses, or at least suggests to a POSITA, logic encoded in one or more tangible non-transitory storage media for execution and when executed operable to an apparatus in a system because Tyso discloses a system operating on computer equipment. Tyso Prov., Fig. 2; Tyso, Fig. 2; Ex1003, ¶124.

a) [13.a] “receive at least first and second real-time video streams, each of the first and second real-time video streams comprising video frames containing a picture comprising a subject image and a background image;”

Tyso discloses or at least suggests to a POSITA the recitations of element [13.a] for the reasons stated above with regard to element [1.a]. Ex1003, ¶125.

b) [13.b] “combine the subject images of corresponding video frames of the first and second video streams into a combined frame of a combined video stream such that the subject image of the first video stream is

positioned in an anterior portion of the combined frame and the subject image of the second video stream is positioned in a posterior portion of the combined frame, wherein the subject images are combined by:”

Tysso discloses or at least suggests to a POSITA the recitations of element [13.b] for the reasons stated above with regard to element [1.b]. Ex1003, ¶126.

- c) **[13.c] “scaling the video frames of the first video stream and repositioning in a first direction the resulting pictures in the video frames of the first video stream to produce a first sequence of scaled video frames;”**

Tysso discloses or at least suggests to a POSITA the recitations of element [13.c] for the reasons stated above with regard to element [1.c]. Ex1003, ¶127.

- d) **[13.d] “removing the background image in the first sequence of scaled video frames to produce first background separated video frames to be displayed in the anterior portion of the combined frames of the combined video stream; and”**

Tysso discloses or at least suggests to a POSITA the recitations of element [13.d] for the reasons stated above with regard to element [1.d]. Ex1003, ¶128.

- e) **[13.e] “superimposing the first background separated video frames onto corresponding ones of the video frames of the second video stream to produce combined video frames of the combined video stream; and”**

Tysso discloses or at least suggests to a POSITA the recitations of element [13.e] for the reasons stated above with regard to element [1.e]. Ex1003, ¶129.

- f) [13.f] “supply[ing] the combined video stream to a video display for displaying the combined video stream.”

Tysso discloses or at least suggests to a POSITA the recitations of element [13.f] for the reasons stated above with regard to element [1.f]. Ex1003, ¶130.

14. Claim 14: “14. The logic of claim 13, wherein the logic that receives is configured to receive video frames of the first video stream comprising images of multiple videoconference participants in the subject image and video frames of the second video stream comprising images of multiple videoconference participants in the subject image.

Tysso discloses, or at least suggests to a POSITA, logic of claim 13 for the reasons discussed above, and Tysso discloses, or at least renders obvious, the recitations of claim 14 for the reasons stated above with regard to claim 2.

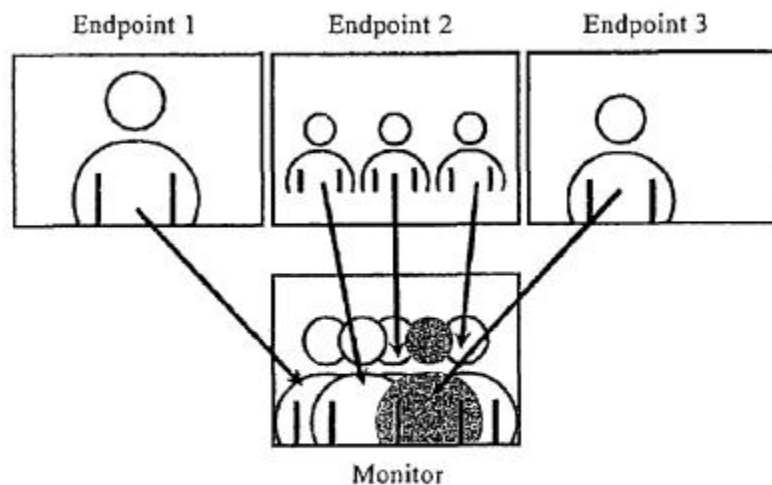


Fig. 1

Tysso, Fig. 1; Tysso, Prov., Fig. 1; Ex1003, ¶131.

15. Claim 15: [15.pre] “15. The logic of claim 13, wherein the

logic that combines is configured to:”

Tysso discloses, or at least suggests to a POSITA, the logic of claim 13 for the reasons discussed above, and Tysso discloses, or at least suggests to a POSITA, the claimed logic that combines as discussed below. Ex1003, ¶132.

- a) **[15.a] “scale the video frames of the second video stream and reposition in a second direction the resulting pictures in the video frames of the second video stream to produce a second sequence of scaled video frames;”**

Tysso discloses or at least suggests to a POSITA the recitations of element [15.a] for the reasons stated above with regard to element [3.a]. Ex1003, ¶133.

- b) **[15.b] “extend the background image in the second sequence of scaled video frames to produce background extended video frames to be displayed in the posterior portion of the combined frames of the combined video stream; and”**

Tysso discloses or at least suggests to a POSITA the recitations of element [15.b] for the reasons stated above with regard to element [3.b]. Ex1003, ¶134.

- c) **[15.c] “superimpose the first background separated video frames onto corresponding ones of background extended video frames to produce combined video frames of the combined video stream.”**

Tysso discloses or at least suggests to a POSITA the recitations of element [15.c] for the reasons stated above with regard to element [3.c]. Ex1003, ¶135.

16. Claim 16: “16. The logic of claim 15, wherein the second direction is opposite to the first direction.”

Tysso discloses, or at least suggests to a POSITA, the logic of claim 15 for the reasons discussed above, and Tysso discloses, or at least suggests to a POSITA, the recitations of claim 16 for the reasons stated above with regard to claim 4.

Ex1003, ¶136.

17. Claim 17: [17.pre] “17. The logic of claim 13, wherein the logic that combines is configured to:”

Tysso discloses, or at least suggests to a POSITA, the logic of claim 13 for the reasons discussed above, and Tysso discloses, or at least suggests to a POSITA, the claimed logic that combines as discussed below. Ex1003, ¶137.

- a) **[17.a] “remove the background image in the video frames of the first and second video streams to produce first and second background separated video frames;”**

Tysso discloses or at least suggests to a POSITA the recitations of element [17.a] for the reasons stated above with regard to elements [1.d] and [5.a]. Ex1003, ¶138.

- b) **[17.b] “generate supplemental background image video frames that comprise a supplemental background image; and”**

Tysso discloses or at least suggests to a POSITA the recitations of element [17.b] for the reasons stated above with regard to element [5.b]. Ex1003, ¶139.

- c) **[17.c] “superimpose corresponding ones of the first background separated video frames, and the second background separated video frames onto corresponding ones of the supplemental background**

video frames to produce the combined video frames of the combined video stream.”

Tysso discloses or at least suggests to a POSITA the recitations of element [17.c] for the reasons stated above with regard to element [5.c]. Ex1003, ¶140.

18. Claim 18: “18. The logic of claim 13, wherein the first and second video streams are produced at the same video conferencing site.”

Tysso discloses, or at least suggests to a POSITA, the logic of claim 13 for at least the reasons set forth above, and Tysso discloses or at least suggests to a POSITA the recitations of claim 18 for the reasons stated above with regard to claim 6. Ex1003, ¶136.

19. Claim 19: “19. The logic of claim 13, wherein the first and second video streams are produced at different video conferencing sites.”

Tysso discloses, or at least suggests to a POSITA, the logic of claim 13 for at least the reasons set forth above, and Tysso discloses or at least suggests to a POSITA the recitations of claim 19 for the reasons stated above with regard to claim 7. Ex1003, ¶137.

20. Claim 20: “20. The logic of claim 13, wherein the logic that combines is configured to be performed at a site other than where the first and second video streams are produced.”

Tysso discloses, or at least suggests to a POSITA, the logic of claim 13 for at least the reasons set forth above, and Tysso discloses or at least suggests to a

POSITA the recitations of claim 20 for the reasons stated above with regard to claim 8. Ex1003, ¶138.

B. Ground 2: Tyso in view of GIMP Renders Obvious Claims 1-20

As noted above, Tyso in view of GIMP demonstrates the obviousness of claims 1-20 more particularly. The User Manual (“GIMP”) for GNU Image Manipulation Program, demonstrates that a POSITA would regard scaling image frames and pictures in image frames to be alternate ways to manipulate and scale images, and contemplates that the use of either technique is within the level of skill of a POSITA. GIMP is analogous art to the ’293 patent and to Tyso as it deals with techniques to process images, and video can be regarded as a collection of images. Ex1003, ¶¶144-46.

In GIMP, a “canvas” corresponds to the visible area of an image. GIMP, 346. An image in GIMP consists of “layers” stacked on each other, and GIMP allows the layers to be individually manipulated. *See, e.g.*, GIMP, 82-100. For example, an image on one layer can overlap an image found on another layer, so that only the upper layer covers the layers below it, which only appear when the upper layer has some transparency in areas surrounding, e.g., the picture in the upper layer. GIMP, 86, 422 (“The Color to Alpha filter makes transparent all pixels with a selected color”). Ex1003, ¶148.

GIMP allows pictures, layers, and canvases to be “scaled,” which enlarges or reduces the size of the picture, layers, or canvases. GIMP, 189-91 (“The Scale Tool is used to scale layers, selections, or paths”); GIMP, 349 (“The Scale Image command enlarges or reduces the physical size of the image by changing the number of pixels it contains. It changes the size of the contents of the image and resizes the canvas accordingly. ... If you only want to scale a particular layer, use the Scale Layer command.”); GIMP, 387-88 (“The Scale Layer command resizes the layer and its contents.”). Ex1003, ¶149.

When reducing the size of an opaque upper layer in GIMP, portions under the layer can be revealed, creating an “extended” background for the upper layer. Ex1003, ¶150. For example, as Dr. Kia explains, a POSITA would understand with regard to Tyso that the processing of a video frame to be presented could include, as a lowest layer, a transparent layer. As disclosed in Tyso, a received video frame could be processed to select a picture in that image according to the principles of GIMP and scale the picture, and apply it as a layer over the transparent layer. *Id.* Alternatively, a received frame with a green screen, as contemplated by GIMP, could be layered over a lowest green screen fill layer. When the received frame layer is scaled over the lower fill layer, the scaling of the received layer would reveal the underlying green screen fill, to keep the green screen extended to the size of the frame. Ex1003, ¶151.

A POSITA would find these alternate ways to manipulate and scale images to be obvious alternatives. Ex1003, ¶152. In fact, GIMP expressly suggests alternative image processing options dependent of the preferences of a POSITA. GIMP, 349 (“If you only want to scale a particular layer, use the Scale Layer command”); GIMP, 189 (“The Scale Tool is used to scale layers, selections or paths (the Object).”); GIMP, 387 (“The Layer to Image Size command resizes the layer boundaries to match the image boundaries, without moving the contents of the layer with respect to the image”).

Once scaled, the background of a received image could be removed to prevent it from obscuring lower layers, as disclosed in Tyso and GIMP by selecting by color and deleting it. Ex1003, ¶153.

Thus, a POSITA would have been motivated to combine the imaging techniques of GIMP with Tyso as discussed, and claims 1-20 are obvious in view of that combination. Ex1003, ¶154.

1. Independent Claim 1: [1.pre]

Tyso discloses, or at least suggests to a POSITA, element [1.pre] for the reasons stated above in Ground 1 with regard to element [1.pre]. Ex1003, ¶155.

a) Elements [1.a]-[1.b]

Tysso discloses, or at least suggests to a POSITA, elements [1.a] and [1.b] for the reasons stated above in Ground 1 with regard to elements [1.a] and [1.b]. Ex1003, ¶¶156-57.

b) [1.c]

Tysso discloses, or at least suggests to a POSITA, element [1.c] for the reasons stated above in Ground 1 with regard to element [1.c]. In addition, as noted above, Tysso in combination with GIMP renders obvious scaling entire video frames as layers. Ex1003, ¶¶158-59.

c) [1.d]-[1.f]

Tysso discloses, or at least suggests to a POSITA, elements [1.d]-[1.f] for the reasons stated above in Ground 1 with regard to elements [1.d]-[1.f]. Ex1003, ¶¶160-62.

2. Claim 2

Tysso discloses, or at least suggests to a POSITA, claim 2 for the reasons stated above in Ground 1 with regard to claim 2. Ex1003, ¶163.

3. Claim 3: [3.pre]

Tysso discloses, or at least suggests to a POSITA, element [3.pre] for the reasons stated above in Ground 1 with regard to element [3.pre]. Ex1003, ¶164.

a) [3.a]

Tysso discloses, or at least suggests to a POSITA, element [3.a] for the reasons stated above in Ground 1 with regard to element [3.a]. In addition, as

noted above, Tysso in combination with GIMP renders obvious scaling entire video frames as layers. Ex1003, ¶¶165-66.

b) [3.b] “extending the background image in the second sequence of scaled video frames to produce background extended video frames to be displayed in the posterior portion of the combined frames of the combined video stream; and”

Tysso discloses, or at least suggests to a POSITA, element [3.b] for the reasons stated above in Ground 1 with regard to element [3.b]. In addition, as noted above, Tysso in combination with GIMP renders obvious extending a background image for a rear layer of a video frame. For example, when a background is transparent, GIMP’s “Layer to Image Size” command (GIMP, 387) would resize the layer of a reduced-sized scaled video frame to extend its background to match the frame size of the output video frame so as not to move “the contents of the layer with respect to the image.” Ex1003, ¶¶167-68.

c) [3.c]

Tysso discloses, or at least suggests to a POSITA, element [3.c] for the reasons stated above in Ground 1 with regard to element [3.c]. Ex1003, ¶169.

4. Claims 4-9

Tysso discloses, or at least suggests to a POSITA, claims 4-9 for the reasons stated above in Ground 1 with regard to claims 4-9. Ex1003, ¶¶170-75.

5. Claim 10: [10.pre], [10.a], [10.b], and [10.c]

Tysso discloses, or at least suggests to a POSITA, elements [10.pre], [10.a], [10.b], and [10.c] for the reasons stated above in Ground 1 with regard to elements [10.pre], [10.a], [10.b], and [10.c]. Ex1003, ¶¶176-79.

a) [10.d]

Tysso discloses, or at least suggests to a POSITA, element [10.d] for the reasons stated above in Ground 1 with regard to element [10.d]. In addition, as noted above, Tysso in combination with GIMP renders obvious scaling entire video frames as layers. Ex1003, ¶¶181-82.

b) [10.e]-[10.f]

Tysso discloses, or at least suggests to a POSITA, elements [10.e]-[10.f] for the reasons stated above in Ground 1 with regard to elements [10.e]-[10.f]. Ex1003, ¶¶182-83.

6. Claim 11

Tysso discloses, or at least suggests to a POSITA, elements [11.pre] and [11.a] for the reasons stated above in Ground 1 with regard to elements [11.pre] and [11.a]. Ex1003, ¶184.

a) [11.b]

Tysso discloses, or at least suggests to a POSITA, element [11.b] for the reasons stated above in Ground 1 with regard to element [11.b]. In addition, as noted above, Tysso in combination with GIMP renders obvious scaling entire video frames as layers. Ex1003, ¶¶185-86.

b) [11.c]

Tysso discloses, or at least suggests to a POSITA, element [11.c] for the reasons stated above in Ground 1 with regard to element [11.c]. In addition, as noted above, Tysso in combination with GIMP renders obvious extending a background image for a rear layer of a video frame. Ex1003, ¶¶187-88.

c) [11.d]

Tysso discloses, or at least suggests to a POSITA, element [11.d] for the reasons stated above in Ground 1 with regard to element [11.d]. Ex1003, ¶189.

7. Claim 12

Tysso discloses, or at least suggests to a POSITA, claim 12 for the reasons stated above in Ground 1 with regard to claim 12. Ex1003, ¶190.

8. Claim 13: [13.pre]

Tysso discloses, or at least suggests to a POSITA, element [13.pre] for the reasons stated above in Ground 1 with regard to element [13.pre]. Ex1003, ¶191.

a) [13.a]-[13.b]

Tysso discloses, or at least suggests to a POSITA, elements [13.a]-[13.b] for the reasons stated above in Ground 1 with regard to elements [13.a]-[13.b].

Ex1003, ¶¶192-93.

b) [13.c] “scaling the video frames of the first video stream and repositioning in a first direction the resulting pictures in the video frames of the first video stream to produce a first sequence of scaled video frames;”

Tysso discloses, or at least suggests to a POSITA, element [13.c] for the reasons stated above in Ground 1 with regard to element [13.c]. In addition, as noted above, Tysso in combination with GIMP renders obvious scaling entire video frames as layers. Ex1003, ¶¶194-95.

c) [13.d]-[13.f]

Tysso discloses, or at least suggests to a POSITA, elements [13.d]-[13.f] for the reasons stated above in Ground 1 with regard to elements [13.d]-[13.f]. Ex1003, ¶¶196-98.

9. Claim 14

Tysso discloses, or at least suggests to a POSITA, claim 14 for the reasons stated above in Ground 1 with regard to claim 14. Ex1003, ¶199.

10. Claim 15: [15.pre]

Tysso discloses, or at least suggests to a POSITA, element [15.pre] for the reasons stated above in Ground 1 with regard to element [15.pre]. Ex1003, ¶200.

a) [15.a] “scale the video frames of the second video stream and reposition in a second direction the resulting pictures in the video frames of the second video stream to produce a second sequence of scaled video frames;”

Tysso discloses, or at least suggests to a POSITA, element [15.a] for the reasons stated above in Ground 1 with regard to element [15.a]. In addition, as noted above, Tysso in combination with GIMP renders obvious scaling entire video frames as layers. Ex1003, ¶¶201-02.

b) [15.b]

Tysso discloses, or at least suggests to a POSITA, element [15.b] for the reasons stated above in Ground 1 with regard to element [15.b]. In addition, as noted above, Tysso in combination with GIMP renders obvious extending a background image for a rear layer of a video frame. Ex1003, ¶¶203-04.

c) [15.c]

Tysso discloses, or at least suggests to a POSITA, element [15.c] for the reasons stated above in Ground 1 with regard to element [15.c]. Ex1003, ¶205.

11. Claim 16-20

Tysso discloses, or at least suggests to a POSITA, claims 16-20 for the reasons stated above in Ground 1 with regard to claims 16-20. Ex1003, ¶¶206-10.

VI. THE BOARD SHOULD INSTITUTE REVIEW

Petitioner believes that discretionary denial is unwarranted. If Patent Owner requests discretionary denial, Petitioner will address discretionary denial factors—including as they relate to 35 U.S.C. § 325(d) and the prior art cited herein—in that briefing as instructed by the Director.

VII. MANDATORY NOTICES

A. Related Matters

The '293 patent was asserted in the following litigation:

Cisco Systems, Inc. et al v. Intelligent Protection Management Corp., Civ.

No. 1-25-cv-00271 (D. Del. March 7, 2025).

B. Real Party-in-Interest

The Petitioner and real party-in-interest is Intelligent Protection Management Corp.

The Patent Owner is Cisco Technology, Inc. In the Cisco litigation, Patent Owner asserts that “Cisco, with its subsidiaries, is the current owner and assignee of over 20,000 patents and over 700 WebEx-related patents. Cisco’s portfolio includes the '293 Patent and '708 Patent” and defines “Cisco” as “Cisco Systems, Inc. (‘Cisco Systems’) and Cisco Technology, Inc. (‘Cisco Technology’) (collectively, ‘Cisco’).”

C. Lead and Back-Up Counsel and Service Information

Lead Counsel	Back-up Counsel
Wayne M. Helge (Reg. No. 56,905) whelge@bdiplaw.com Bunsow De Mory LLP 277 S. Washington St., Suite 210, #1088 Alexandria, VA 22314 Telephone: 571-208-0186 Facsimile: 415-426-4744	James T. Wilson (Reg. No. 41,439) jwilson@bdiplaw.com Bunsow De Mory LLP 277 S. Washington St., Suite 210, #1088 Alexandria, VA 22314 Telephone: 571-222-1022 Facsimile: 415-426-4744

Petitioner consents to electronic service at the email addresses listed above (whelge@bdiplaw.com, jwilson@bdiplaw.com) with additional cc: to erobinson@bdiplaw.com.

VIII. GROUNDS FOR STANDING

Petitioner certifies the '293 patent is available for *inter partes review* and that Petitioner is not barred or estopped.

IX. CONCLUSION

For the reasons presented above, the Petition for institution of *inter partes review* should be granted and the challenged claims cancelled.

Dated: October 6, 2025

/Wayne M. Helge/
Wayne M. Helge, Reg. No. 56,905
James T. Wilson, Reg. No. 41,439
BUNSOW DE MORY LLP
277 S. Washington St., Suite 210 #1088
Alexandria, VA 22314
T: (571) 208-0186
Email: whelge@bdiplaw.com

Email : jwilson@bdiplaw.com

Lead Counsel for Intelligent Protection
Management Corp.

CERTIFICATE OF WORD COUNT

The undersigned certifies that the foregoing PETITION FOR *INTER PARTES* REVIEW OF U.S. PATENT NO. 8,830,293 complies with the type-volume limitation in 37 C.F.R. § 42.24(a)(1)(i). According to the word-processing system's word count, the brief contains 13,506 words, excluding the parts of the brief exempted by 37 C.F.R. § 42.24(a).

By: /Wayne M. Helge/
Wayne M. Helge (Reg. No. 56,905)
Lead Counsel for Intelligent Protection
Management Corp.

CERTIFICATE OF SERVICE

The undersigned certifies that the foregoing Petition for *Inter Partes* Review of U.S. Patent No. 8,830,293, the associated Power of Attorney, and Exhibits EX1001 through EX1009 were served on this date by Priority Mail Express® (or by means at least as fast and reliable as Priority Mail Express®) to the correspondence address of record indicated in the Patent Office's Patent Center system for U.S. Patent No. 8,830,293:

EDELL, SHAPIRO & FINNAN, LLC
9801 WASHINGTONIAN BLVD.
SUITE 750
GAITHERSBURG, MD 20878

Dated: October 6, 2025

/Wayne M. Helge/
Wayne M. Helge (Reg. No. 56,905)
BUNSOW DE MORY LLP
277 S. Washington St., Suite 210, #1088
Alexandria, VA 22314
Email: whelge@bdiplaw.com
Telephone: 571-208-0186

Lead Counsel for Petitioner Intelligent
Protection Management Corp.