UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

~~APPLE INC.~~
GOOGLE LLC
Petitioner

v.

GESTURE TECHNOLOGY PARTNERS LLC
Patent Owner

Case No. ~~IPR2021-00921~~IPR2022-00362
U.S. Patent No. 8,878,949

PETITION FOR *INTER PARTES* REVIEW OF U.S. PATENT NO. 8,878,949

TABLE OF CONTENTS

I. INTRODUCTION

Petitioner ~~Apple Inc.~~Google LLC ("Petitioner") requests an *Inter Partes* Review ("IPR") of claims 1–18 (the "Challenged Claims") of U.S. Patent No. 8,878,949 ("the '949 Patent"). This petition is substantively the same as IPR2021-00921 (which is instituted), and is being filed concurrently with a motion for joinder with respect to that proceeding.

II. SUMMARY OF THE '949 PATENT

A. The '949 Patent's Alleged Invention

Generally directed to digital imaging, the '949 Patent seeks to automate the process of taking a picture by analyzing the scene and capturing an image when "certain poses of objects, sequences of poses, motions of objects, or any other states or relationships of objects are represented." *'949 Patent* (Ex. 1001), 1:50-2:8. The patent describes a number of different scenarios that, when detected, cause the camera to capture an image. Some examples include detecting (1) a "[s]ubject in a certain pose," (2) a "[s]ubject in a sequence of poses," (3) a "[p]ortion of [s]ubject in a sequence of poses (e.g., gestures)," (4) a "[s]ubject or portion(s) in a specific location or orientation," (5) a "[s]ubject in position relative to another object or person" such as a "bride and groom kissing in a wedding," and (6) "a subject undertak[ing] a particular signal comprising a position or gesture" such as "raising one's right hand." *Id*. at 5:30-49. Only gestures are claimed, however. Each of the

Challenged Claims requires detecting or determining a "gesture has been performed." *Id*. at Independent Claims 1, 8, 13.

The '949 Patent contemplates multiple image sensors to accomplish its goal. For example, a "central camera . . . is for picture taking and has high resolution and color accuracy," while "lower resolution" cameras "with little or no accurate color capability . . . are used to simply see object positions." *Id*. at 5:1-6. Although the term is not used outside the claims, all Challenged Claims refer to the gesture-capturing sensor as an "electro-optical sensor." *Id*. at Independent Claims 1, 8, 13.

B. The '949 Patent's Prosecution

The Application that resulted in the '949 Patent was filed on August 7, 2013. The Application claims priority to provisional patent application No. 60/133,671, filed May 11, 1999. *Id*. at (22), (60). For purposes of this petition and without waiving its right to challenge priority in this or any other proceeding, Petitioner adopts May 11, 1999 as the invention date for the Challenged Claims.

A first office action rejected all initially presented claims as anticipated or obvious over U.S. Patent No. 6,359,647 to Sengupta et al. ("*Sengupta*"). *'949 File History* (Ex. 1002), 136-144. The examiner noted that *Sengupta* teaches an electro-optical sensor separate from a digital camera, which triggers an image capture when it detects movement within the sensor's field of view. *Id*. at 140-141.

In response, the Applicant characterized *Sengupta* as a system comprising multiple security cameras that transitions to an appropriate camera when an object moves from one camera's field of view to another's. *Id*. at 167-168. Focusing on structural distinctions, the Applicant argued that Sengupta did not teach "a device housing including a forward facing portion having an electro-optical sensor and a digital camera" as required by Claim 1 and its dependents. *Id*. at 168. The Applicant drew a functional distinction with respect to the claims that ultimately issued as independent Claims 8 and 13 (and their dependents), arguing *Sengupta* does not "identify a particular gesture apart from a plurality of gestures, where the particular gesture corresponds to an image capture command." *Id*. at 169-170.

A second office action rejected the Applicant's alleged distinctions, finding the structural point was "not clearly defined in claim 1" and "the term 'gesture' [] not clearly defined in the claim[s]" to support the purported distinction regarding independent Claims 8 and 13. *Id*. at 186. Following an examiner interview on August 7, 2014 (*Id*. at 199), the Applicant further amended the claims to distinguish the claimed invention from *Sengupta*. *Id*. at 210-217. The Applicant noted, "[w]ith respect to [the] amended independent claims . . . , Sengupta does not disclose, teach or suggest: a) a device housing including a forward facing portion that encompasses an electro-optical sensor and a digital camera; or b) a processor to determine a

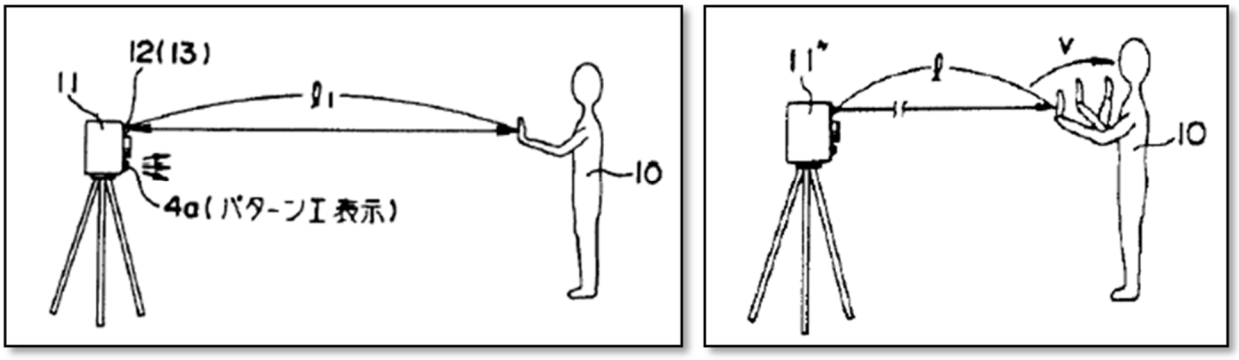
gesture corresponds to an image capture command, which causes the camera to store or capture an image." *Id*. at 215. Regarding the claimed "gesture," the Applicant distinguished *Sengupta's* generic movement tracking from the claimed invention, which must "identif[y] a particular gesture" that triggers the device to capture an image. *Id*. at 216.

In response, the Examiner held the pending claims allowable, and the application was deemed allowable on September 18, 2014 after multiple double patenting issues were resolved. *Id*. at 254-260.

C. Overview of the Proposed Grounds

Neither the camera structure—multiple forward-facing image sensors—nor the concept of detecting a specific gesture that causes a camera to capture an image were new. As set forth in detail below, U.S. Patent 6,144,366 to Numazaki et al. ("*Numazaki*") teaches portable devices equipped with multiple forward-facing image sensors, detects gestures performed by users, connects those gestures with specific commands, and captures images with a separate sensor from those used to detect gestures. Although *Numazaki* does not expressly contemplate detecting a gesture that causes an image capture command to be processed, this was also known in the art. Namely, JPH4-73631 to Osamu Nonaka ("*Nonaka*") (a Japanese Unexamined Patent Application Publication filed by Olympus Optical Co., Ltd.)

teaches a process by which "the photographer gives a release instruction" for the camera to capture an image by performing a "predetermined motion towards the camera." *Nonaka* (Ex. 1005), 3:34-36. The camera "detects this motion by the subject [] and exposure is carried out." *Id*. at 37-38. *Nonaka* contemplates multiple predetermined gestures such as holding one's hand out toward the camera (as depicted in Fig. 3 below left) or moving one's hand toward the camera (as depicted in Fig. 7 below right).



Id. at Figs. 3, 7. For the reasons discussed in the proposed grounds below, a PHOSITA would have been motivated to implement such image capture gesture functionality in *Numazaki's* gesture-recognizing, multi-camera devices.

D. A Person Having Ordinary Skill in the Art

A person having ordinary skill in the art ("PHOSITA") at the time of the '949 Patent would have had at least a bachelor's degree in electrical engineering or equivalent with at least one year of experience in the field of human computer

interaction. Additional education or experience might substitute for the above requirements. *Bederson Dec.* (Ex. 1003), ¶¶ 29-31.

III. REQUIREMENTS FOR IPR UNDER 37 C.F.R. § 42.104

A. Standing Under 37 C.F.R. § 42.104(A)

Petitioner certifies that the '949 Patent is available for IPR and that Petitioner is not barred or estopped from requesting an IPR challenging the claims of the '949 Patent. Specifically, (1) Petitioner is not the owner of the '949 Patent, (2) Petitioner has not filed a civil action challenging the validity of any claim of the '949 Patent, and (3) ~~this Petition is filed less than one year after the~~ Petitioner ~~was~~has not been served with a complaint more than one year ago alleging infringement of the '949 ~~Patent~~patent.

B. Challenge Under 37 C.F.R. § 42.104(B) and Relief Requested

In view of the prior art and evidence presented, claims 1–18 of the '949 Patent are unpatentable and should be cancelled. 37 C.F.R. § 42.104(b)(1). Further, based on the prior art references identified below, IPR of the Challenged Claims should be granted. 37 C.F.R. § 42.104(b)(2).

Proposed Ground of Unpatentability	Exhibits
Ground 1: Claims 1-18 are obvious under pre-AIA 35 U.S.C. § 103 over U.S. Patent 6,144,366 to Numazaki, et al. ("*Numazaki*") in view of JPH4-73631 to Osamu Nonaka ("*Nonaka*")	Ex. 1004, Ex. 1005

Ground 2: Claims 6, 11, and 12 are obvious under pre-AIA 35 U.S.C. § 103 over *Numazaki* in view of *Nonaka* and in further view of U.S. Patent No. 5,666,157 to David G. Aviv ("*Aviv*")	Ex. 1004, Ex. 1005, Ex. 1006

Section IV identifies where each element of the Challenged Claims is found in the prior art. 37 C.F.R. § 42.104(b)(4). The exhibit numbers of the evidence relied upon to support the challenges are provided above and the relevance of the evidence to the challenges raised is provided in Section IV. 37 C.F.R. § 42.104(b)(5). **Exhibits 1001-~~1017~~1018** are also attached.

C. Claim Construction Under 37 C.F.R. § 42.104(B)(3)

In this proceeding, claims are interpreted under the same standard applied by Article III courts (i.e., the *Phillips* standard). *See* 37 C.F.R § 42.100(b); *see also* 83 Fed. Reg. 197 (Oct. 11, 2018); *Phillips v. AWH Corp.*, 415 F.3d 1303, 1312 (Fed. Cir. 2005) (*en banc*). Under this standard, words in a claim are given their plain meaning which is the meaning understood by a person of ordinary skill in the art in view of the patent and file history. *Phillips*, 415 F.3d 1303, 1212–13. With the single exception discussed below in this section, Petitioner proposes that no terms require express construction to resolve the proposed grounds presented herein.

a. *Capture and store an image*

All three independent claims require capturing and/or storing "an image" in response to detecting an image capture gesture. *See '949 Patent* (Ex. 1001), Claim

1 ("the image capture command causes the digital camera to store an image to memory"), Claim 8 ("capturing an image to the digital camera in response to . . . the image capture command"), Claim 13 ("correlate the gesture detected . . . with an image capture function and subsequently capture an image using the digital camera"). For a number of reasons, these limitations should be construed broadly enough to encompass capturing/storing video or still images. First, the '949 Patent repeatedly teaches that its invention applies both to still image and video capture. *See, e.g., id*. at 2:19-30 (describing efficiencies that can be realized in "movie making" by implementing the invention, noting it "allows more cost effective film production by giving the director the ability to expose the camera to the presence of masses of data, but only saving or taking that data which is useful), 8:10-30 (describing "digital or other camera 320" that "record[s] the picture" when the subjects are posed in a predetermined manner, noting "camera 320 may be a video camera and recorder").

Second, the '949 Patent teaches that captured video comprises a sequence of images. *Id*. at 8:10-30 (describing an embodiment in which "camera 320 may be a video camera and recorder which streams in hundreds or even thousands of frames of image data[,]" noting the subjects may then select individual prints from the captured image data). The Federal Circuit "has repeatedly emphasized that an

indefinite article 'a' or 'an' in patent parlance carries the meaning of 'one or more' in open-ended claims containing the transitional phrase 'comprising.'" *Baldwin Graphic Sys., Inc. v. Siebert, Inc.*, 512 F.3d 1338, 1342 (Fed. Cir. 2008) (quoting *KCJ Corp. v. Kinetic Concepts, Inc.,* 223 F.3d 1351, 1356 (Fed. Cir. 2000)). Indeed, this principle "is best described as a rule, rather than merely as a presumption or even a convention" and "[t]he exceptions to this rule are extremely limited." *Id*. (noting the patentee must evince a clear intent to limit "a" or "an" to "one"). Here, all independent claims are open-ended containing the transitional phrase "comprising," and each requires capturing/storing "an image." Applying the Federal Circuit's well-established "rule," the claims should be construed to capture capturing/storing *one or more* images to memory in response to detecting the image capture gesture. Because the '949 Patent teaches that captured video is a sequence of still image frames, the claims should be construed to encompass capturing/storing still images or a video sequence of images.

Finally, as discussed above, the prior art-based rejections during prosecution all focused on U.S. Patent No. 6,359,647 to Sengupta, et al. ("*Sengupta*") (Ex. 1014). *Sengupta* discloses "a system for controlling multiple video cameras" that supports "an automated camera handoff for selecting and directing cameras" by "tracking a figure." The applicant conceded that "the

security system in Sengupta tracks movement of an object 'from one camera's field of view to another camera's field of view,'" but disputed that Sengupta taught detecting a gesture that corresponds to an image capture command. '949 File History (Ex. 1002), 216 (arguing that *Sengupta* simply "'reports the location' of the moving objects"). Critically, however, the applicant did not attempt to distinguish Sengupta on the basis that it captures video, rather than still images. Accordingly, the prosecution history supports a conclusion that the applicant intended the Challenged Claims to encompass both capturing/storing still images and video.

IV. THE CHALLENGED CLAIMS ARE UNPATENTABLE

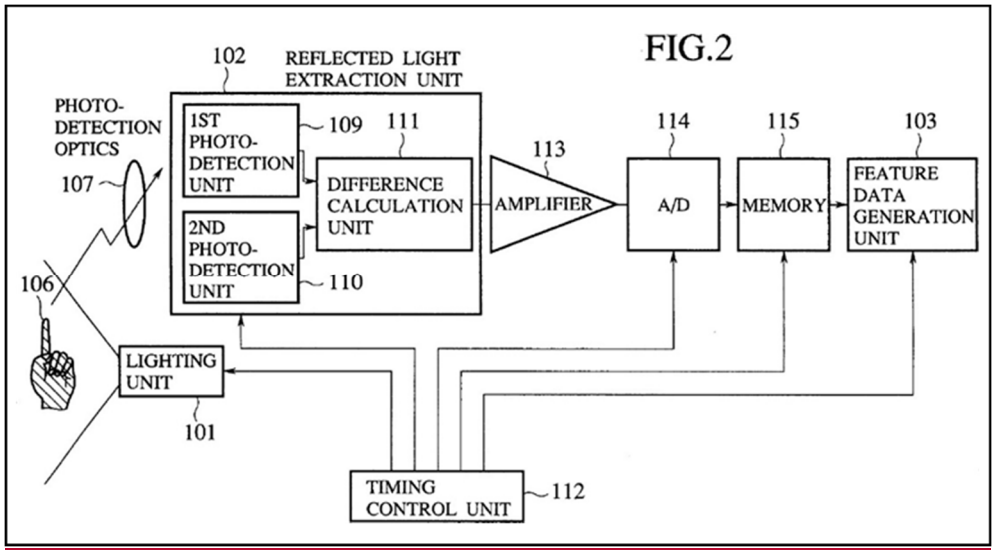
A. Ground 1: Claims 1-18 are obvious under pre-AIA 35 U.S.C. § 103 over *Numazaki* in view of *Nonaka*

Overview of Numazaki

U.S. Patent No. 6,144,366 to Numazaki et al. (*"Numazaki"*) (Ex. 1004) was filed on October 17, 1997, issued on November 7, 2000, and is prior art to the '949 Patent under at least 35 U.S.C. § 102(e) (pre-AIA). *Numazaki* was not cited or considered during prosecution of the '949 Patent. *'949 Patent* (Ex. 1001).

Numazaki is generally directed to a method for detecting a gesture or the movement of a user's hand. *Numazaki* (Ex. 1004), Abstract, 4:9-40. *Numazaki* purports to have improved upon prior methods by using a controlled light source to

illuminate the target object (e.g., the user's hand), a first camera unit (referred to by *Numazaki* as a "photo-detection unit"),[1] and a second camera unit. *Id*. at 11:9-23. This arrangement is illustrated in Fig. 2 below:



Id. at Fig. 2. A timing control unit is used to turn lighting unit 101 on (i.e., illuminating the target object) when the first camera unit is active and off when the second camera unit is active. *Id*. at 11:20-32. The result of this light control is the

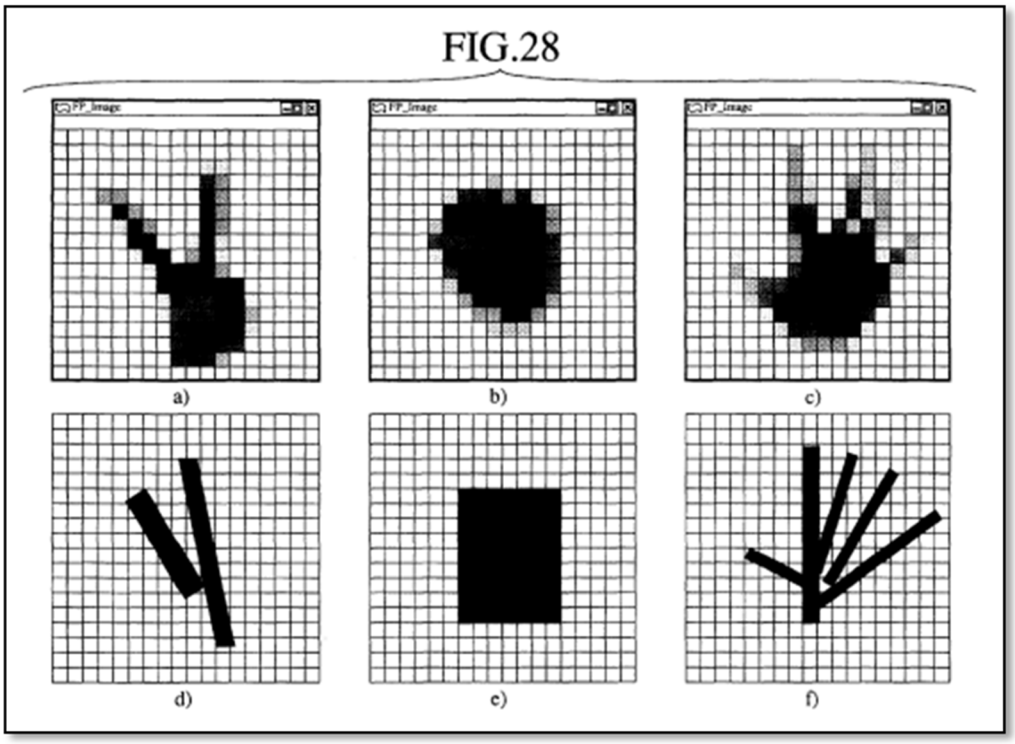
[1] A PHOSITA would have considered *Numazaki's* photo-detection units to be camera units. *Bederson Dec.* (Ex. 1003), ¶ 35 (explaining that *Numazaki* describes using CMOS or CCD sensor units at 15:24-16:19, which were two of the more common optical sensors used in camera units at the time).

first camera unit captures an image of the target object illuminated by both natural light and the lighting unit 101 and the second camera unit captures an image of the target object illuminated by only natural light. *Id*. at 11:33-39. The difference between the two images—obtained by difference calculation unit 111—represents the "reflected light from the object resulting from the light emitted by the lighting unit 101." *Id*. at 11:43-51. This information is then used by feature data generation unit 103 to determine gestures, pointing, etc. of the target object that may be converted into commands executed by a computer. *Id*. at 10:57-66.

Throughout its lengthy disclosure, *Numazaki* describes various features and functionality based on this two-camera structure. Although referred to as "embodiments," as discussed in detail below, a PHOSITA would have understood that many are complementary and would have been motivated to combine certain of these features and functionalities in a single device.

Numazaki's third embodiment is "another exemplary case of the feature data generation unit of the first embodiment, which realizes a gesture camera for recognizing the hand action easily and its application as a pointing device in the three-dimensional space." *Id*. at 29:4-8. In this embodiment, data reflecting pre-registered gestures or hand positions is stored in "shape memory unit 332." *Id*. at 29:9-38. This stored data is compared with the output from the two-sensor "reflected

light extraction unit" (i.e., component 102 from first embodiment) to identify when the user has performed a pre-registered gesture and instructs the device to implement a "command corresponding to that [gesture]." *Id*. at 29:22-30:5. In addition to "hand [g]esture recognition as the means for inputting the command into the three-dimensional space and the like into the computer," the third embodiment contemplates additional gesture-based instructions such as "instructing the power ON/OFF of the TV, the lighting equipment, etc." *Id*. at 31:3-10. *Numazaki* cautions that gestures can be erroneously detected when the user did not intend to command the system and proposes a sequence of gestures to prevent such "erroneous commanding." *Id*. at 31:11-44, Fig. 29 (describing a sequence of "two fingers, fist, five fingers" that causes the system to "power on"). Such a sequence (frames a-c) and the rectangular data extracted from the images (frames d-f) are illustrated in Fig. 28 below:

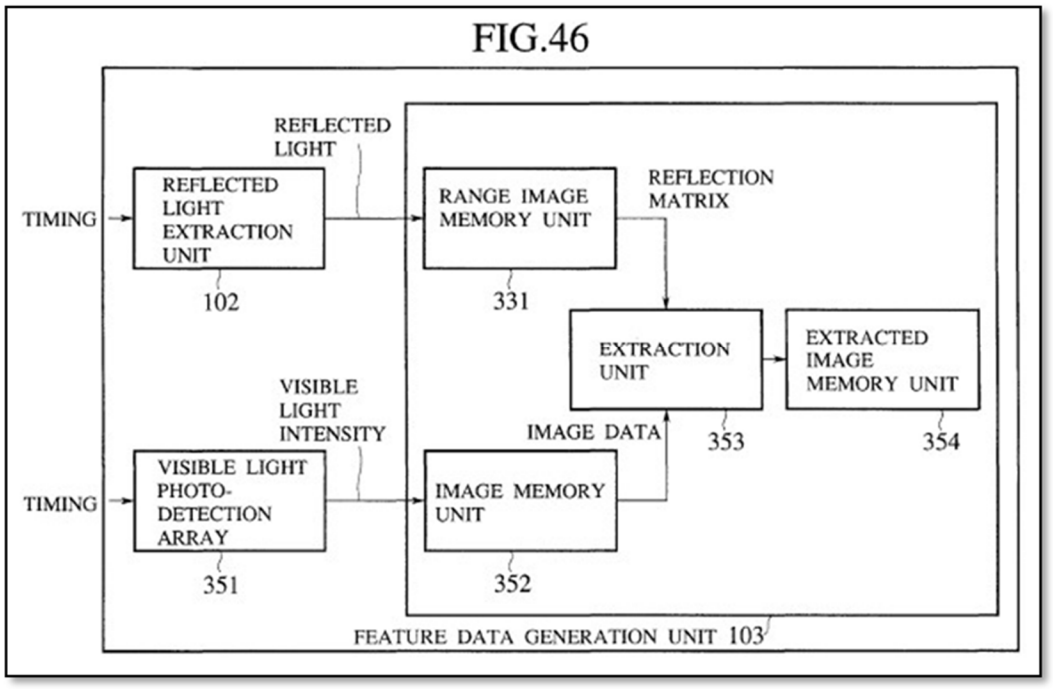


Id. at Fig. 28.

Numazaki's fifth embodiment is directed to video capture and transmission for applications such as videoconferencing. Specifically, the fifth embodiment recognizes the difficulties involved with videoconference applications using a portable device, including communication costs and power consumption. *Id*. at 39:6-16. To combat these difficulties, *Numazaki* teaches a "TV telephone" that extracts and transmits only the faces of the participating users, removing extraneous background information that would otherwise consume bandwidth and battery power. *Id*. at 39:12-20. To accomplish this improvement, *Numazaki* uses the same two-camera "reflected light extraction unit 102" (i.e., the same component 102 from

first embodiment) in conjunction with a separate "visible light photo-detection array 351 which is generally used as a CCD camera for taking video images." *Id.* at 39:21-49. This structure is depicted in Fig. 46 below:

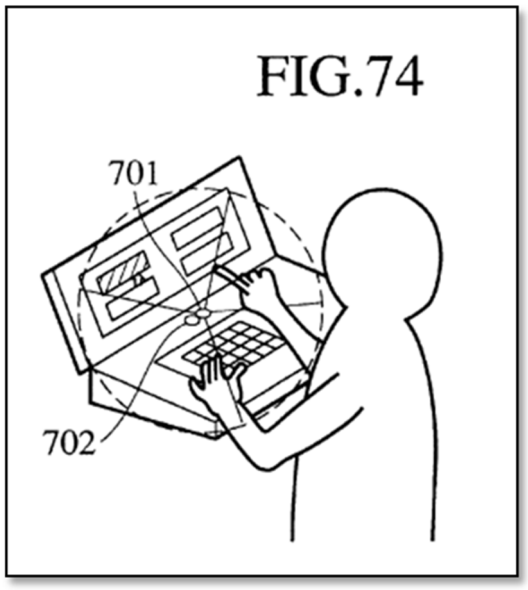
FIG.46



Id. at Fig. 46. Using this arrangement, the fifth embodiment processes the output of two-sensor structure 102 to identify an outline of the subject and subtracts everything outside this outline from the image captured by the visible light sensor 351. *Id.* at 39:24-60. In doing so, the fifth embodiment arrives at image information that contains only the subject without any background image information and uses much less data as a result of removing the extraneous image information. *Id.* at 40:32-35

("By recording the extracted image, it is possible to reduce the recording capacity considerably (in an order of one tenth to one hundredth)[.]").

Numazaki's eighth embodiment describes various portable computers such as laptops and handheld devices on which the earlier-described hardware and functionalities may be implemented. *Id*. at 50:19-24 For example, Fig. 74 depicts a laptop computer with a lighting unit 701 and camera unit 702:



Id. at Fig. 74, 50:25-37 (describing the same). Similar implementations in which a user can control a cursor by moving a finger are described with reference to Fig. 78 (a PDA device) and Fig. 79 (a wristwatch device). *Numazaki* expressly teaches that these portable devices implement the information input generation apparatus described in the preceding embodiments. *Id*. at 50:19-24 ("This eighth embodiment

is directed to a system configuration incorporating the information input generation apparatus of the present invention as described in the above embodiments."). For example, *Numazaki's* eighth embodiment devices include the controlled light and two-camera configuration described in its first embodiment. As described with respect to the first embodiment above, *Numazaki's* information input generation apparatus uses controlled lighting and two camera sensors to isolate an image of an object (such as a user's hand) by subtracting the image taken when the light is off from the image taken when the light is on, thereby obtaining an image that focuses on the illuminated object. *Numazaki* teaches that this precise image difference calculation is included in the eighth embodiment. *Id*. at 53:22-36. Accordingly, a PHOSITA would have understood that *Numazaki's* eighth embodiment portable devices incorporate the controlled lighting and two-camera sensor structure described with respect to the first embodiment. *Bederson Dec*. (Ex. 1003), ¶ 44. As discussed below in the motivations to combine *Numazaki* and *Nonaka* section, a PHOSITA would also have been motivated to implement in the eighth embodiment laptop the third embodiment's feature that maps specific gestures to specific commands and the fifth embodiment's video capture feature.

Because *Numazaki*, like the '949 Patent, discloses a camera system that may be controlled by human gesture input, *Numazaki* is in the same field of endeavor as

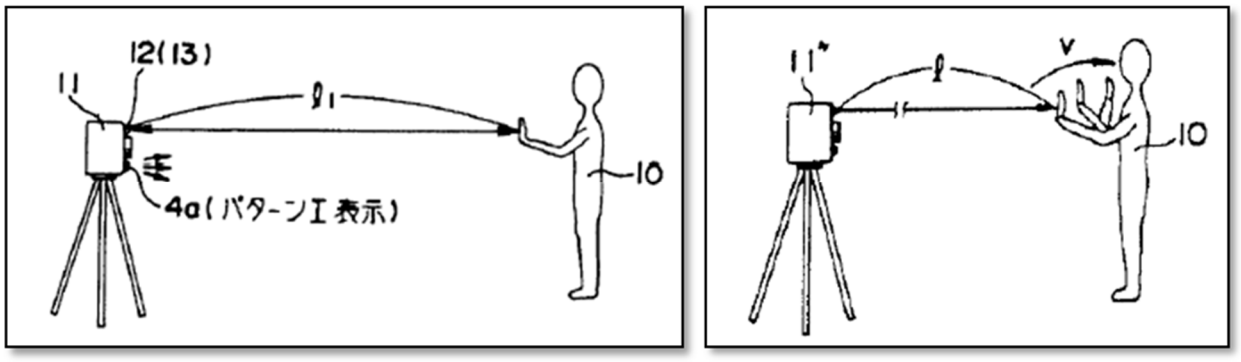
the ’949 Patent. *Compare Numazaki* (Ex. 1004), 31:3-10 (noting gesture recognition is used as “the means for inputting the command . . . into the computer” and noting gestures can be used to turn on/off equipment such as a TV or lighting equipment), 50:25-48 (describing a portable laptop device equipped with a light source and camera such that a user can control certain functionalities using hand gestures) *with ’949 Patent* (Ex 1001), 5:24-49 (describing the process of identifying a subject has performed a predetermined gesture and controlling the camera device to capture an image in response). *Numazaki* is therefore analogous art to the ’949 Patent. *Bederson Dec.* (Ex. 1003), ¶¶ 33-45.

Overview of Nonaka

Japanese Unexamined Patent Application Publication JPH4-73631 to Osamu Nonaka (“*Nonaka*”) (Ex. 1005) published on March 9, 1992 and is prior art to the ’949 Patent under at least 35 U.S.C. § 102(b) (pre-AIA). *Nonaka* was not cited or considered during prosecution of the ’949 Patent. *’949 Patent* (Ex. 1001).

Nonaka teaches a “remote release device-equipped camera” that allows a user to signal a desire for the camera to take a picture by “mak[ing] a predetermined motion.” *Nonaka* (Ex. 1005), 15:11-14. *Nonaka* contemplates multiple predetermined gestures such as holding one’s hand out toward the camera (as

depicted in Fig. 3 below left) or moving one's hand toward the camera (as depicted in Fig. 7 below right):



Id. at Figs. 3, 7, 3:34-4:4 (describing "the photographer giv[ing] a release instruction by means of a predetermined motion towards the camera . . . such as moving his or her hand forward"), 6:11-22 (describing the photographer "giv[ing] a release instruction . . . [by] mov[ing] his or her hand, etc. at a certain speed").

Because *Nonaka*, like the '949 Patent, discloses a camera system that may be controlled by human gesture input, *Nonaka* is in the same field of endeavor as the '949 Patent. *Compare Nonaka* (Ex. 1005), 14:32-36 ("[D]uring the remote release operation of the camera according to the present exemplary embodiment, if the photographer makes a motion such as moving his or her hand . . . when he or she intends to carry out the release operation, the photographer is thus able to communicate that intent to the camera.") *with '949 Patent* (Ex 1001), 5:24-49

(describing the process of identifying a subject has performed a predetermined gesture and controlling the camera device to capture an image in response). *Nonaka* is therefore analogous art to the '949 Patent. *Bederson Dec.* (Ex. 1003), ¶¶ 46-47.

Motivation to Combine Numazaki and Nonaka

As discussed above, *Numazaki's* first embodiment teaches a two-camera "reflected light extraction unit" through which an object such as a user's hand can be detected. *Numazaki's* third embodiment takes the output from the two-sensor "reflected light extraction unit," compares it with stored data reflecting pre-registered gestures or hand positions in "shape memory unit 332" to detect when the user has performed a pre-registered gesture, and instructs the device to implement a "command corresponding to that [gesture]." *Numazaki* (Ex. 1004), 29:19-30:5. Such gesture commands, for example, can instruct the system to turn devices on or off. *Id*. at 31:3-10. In its fifth embodiment, *Numazaki* adds an additional "visible light photo-detection array 351 . . . for taking video images." *Id*. at 39:21-49. Using this arrangement, the fifth embodiment processes the output of two-sensor "reflected light extraction unit" 102 to identify an outline of the subject and subtracts everything outside this outline from the image captured by the visible light sensor 351, thereby decreasing the size of the ultimate video file. *Id*. at 39:24-60, 4:32-35. Finally, *Numazaki's* eighth embodiment teaches portable devices that implement the

information input generation apparatus described in the preceding embodiments. *Id*. at 50:19-24. Although *Numazaki* does not expressly describe combining all these features into a single portable device such that a user could perform a gesture command (pursuant to its third embodiment) that causes video capture to initiate (pursuant to its fifth embodiment), a PHOSITA would have been motivated to implement *Numazaki's* portable device in this manner pursuant to *Nonaka's* image capture command gesture teachings.

First, implementing *Numazaki's* features as proposed would have improved *Numazaki's* portable devices in the same way *Nonaka's* gesture-based image capture functionality benefits its camera device. *Bederson Dec*. (Ex. 1003), ¶¶ 48-49 (explaining that *Numazaki's* native functionality of associating hand gestures with commands would have been a natural fit as a means to initiate video capture); *KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 417 (2007) (obvious to use known techniques to improve similar devices in the same way). *Nonaka* explains that users desired the ability to remotely trigger image capture, but that then-existing options were limited to self-timer mechanisms and expensive wireless remote controls—both of which were undesirable. *Nonaka* (Ex. 1006), 2:6-25. According to *Nonaka*, its gesture-based image capture solution "achiev[es] a higher degree of freedom, good portability, and cost benefits." *Id*. at 2:26-29. A PHOSITA would have recognized

that these same benefits would be realized in *Numazaki's* laptop. Namely, implementing *Numazaki's* laptop as proposed would provide a greater degree of freedom than alternative means of trigger image capture such as timers or manual initiation, ensuring the user is able to get in position and prepared before the video capture begins. *Bederson Dec*. (Ex. 1003), ¶ 49. Using gesture-based image capture initiation also maintains *Numazaki's* portability and reduces costs by employing *Numazaki's* native gesture recognition functionality, rather than adding wireless or wired remote control hardware. *Id*.

Second, a PHOSITA would have been motivated to utilize distinct sensor units to perform the gesture recognition and video capture. Specifically, the lower resolution "reflected light extraction unit" 102 is well suited to monitor the scene and detect any gestures performed by the user, while the higher resolution "visible light photo-detection array 351" is better suited for taking full quality images of the scene (when such video is requested by the user). *Bederson Dec*. (Ex. 1003), ¶ 50. Indeed, as Dr. Bederson explains, it was widely recognized that lower resolution monitoring enables a system to quickly respond to changes in the environment and is preferable to higher resolution monitoring:

> A key requirement of many image processing systems is that they operate quickly enough to respond to changes in the visual environment in a timely fashion. This is especially true for systems that need to

> respond to events as they happen in the physical world. Because the essential nature of image processing is that the pixels in the image are examined, the more pixels there are, the heavier the computational requirements are, and the slower the system can respond. A tension is that if an image does not hold enough detail about a scene, the image processing algorithm will not have enough information to correctly analyze that scene. Thus, it is not surprising that managing the resolution of images for processing is an essential part of image processing systems. People have for decades found ways to work with lower resolution images—sometimes starting with lower resolution images at first, and then working with higher resolution images later for fine-tuned processing.

Id. at ¶ 50 (further discussing a 1982 textbook that describes image resolution algorithms that are used to process images from coarse-to-fine resolution as required by the application, which it describes as "vastly more efficient than [] single-level, high resolution-only" algorithms). As another specific example, Dr. Bederson describes his own work with image processing that utilized hybrid resolution imaging to "balance the tension between number of pixels and acuity." *Id*.

Third, a PHOSITA would have anticipated success in implementing *Numazaki* in this manner given that *Numazaki* already includes the technical hardware and programming necessary to detect gestures, associate gestures with commands, and capture video, and expressly contemplates incorporating these early-

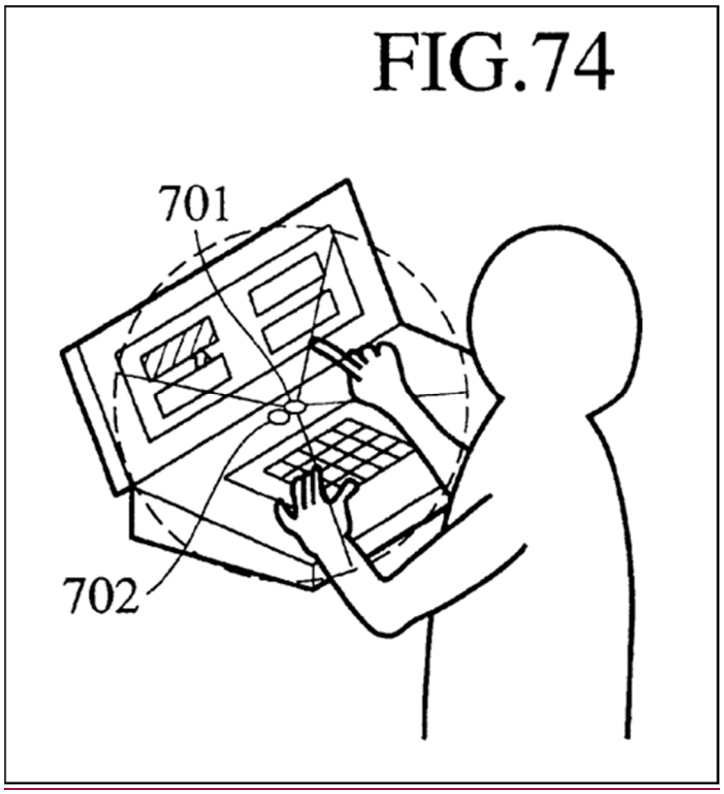
described embodiments in the eighth embodiment portable devices. *Bederson Dec.* (Ex. 1003), ¶ 51 (discussing the third embodiment's gesture recognition and gesture-command mapping, the fifth embodiment's video capture, and the eighth embodiment's portable devices); *see also KSR*, 550 U.S. at 416 (obvious to combine prior art elements according to known methods to yield predictable results). Indeed, as discussed at length above, the third embodiment uses *Numazaki's* two-sensor reflected light extraction unit 102 to recognize hand-based gestures and then maps these gestures to commands such as turning a device on or off. *Bederson Dec.* (Ex. 1003), ¶¶ 36, 39-43. The fifth embodiment uses the same two-sensor reflected light extraction unit 102 and adds a third "visible light photo-detection array" that serves as a video camera. *Id*. at 39-43. The eighth embodiment's portable device, such as Fig. 74's laptop, already includes "the information input generation apparatus of the present invention as described in the above embodiments," which a PHOSITA would have recognized includes at least the two-sensor reflected light extraction unit 102 and gesture recognition functionality. *Id*. at ¶ 51 (discussing *Numazaki* at 50:19-24, which teaches the earlier described embodiments are incorporated into the portable devices and 53:22-36, which expressly teaches the two-sensor differencing process for gesture recognition). A PHOSITA would have understood that adding a third image sensor to the portable laptop in support of the fifth embodiment's video

capture functionality would have been straightforward and well within the skill set of a PHOSITA. Indeed, *Numazaki* even notes that the multiple sensors could be "mixed[] within the same photo-detection array, and share a photo-detection lens so as to reduce the size and the weight." *Numazaki* (Ex. 1004), 39:42-49; *Bederson Dec*. (Ex. 1003), ¶ 51 (discussing the same and noting that videoconferencing technology had been incorporated into even smaller devices such as cell phones at the time).

i. Claim 1

[1(P)] A portable device comprising:

To the extent the preamble is limiting, *Numazaki's* eighth embodiment teaches a computer implemented method for controlling functions on a portable laptop device through gestures or pointing. *Numazaki* (Ex. 1004), 50:38-43. This arrangement is illustrated in Fig. 74 below:



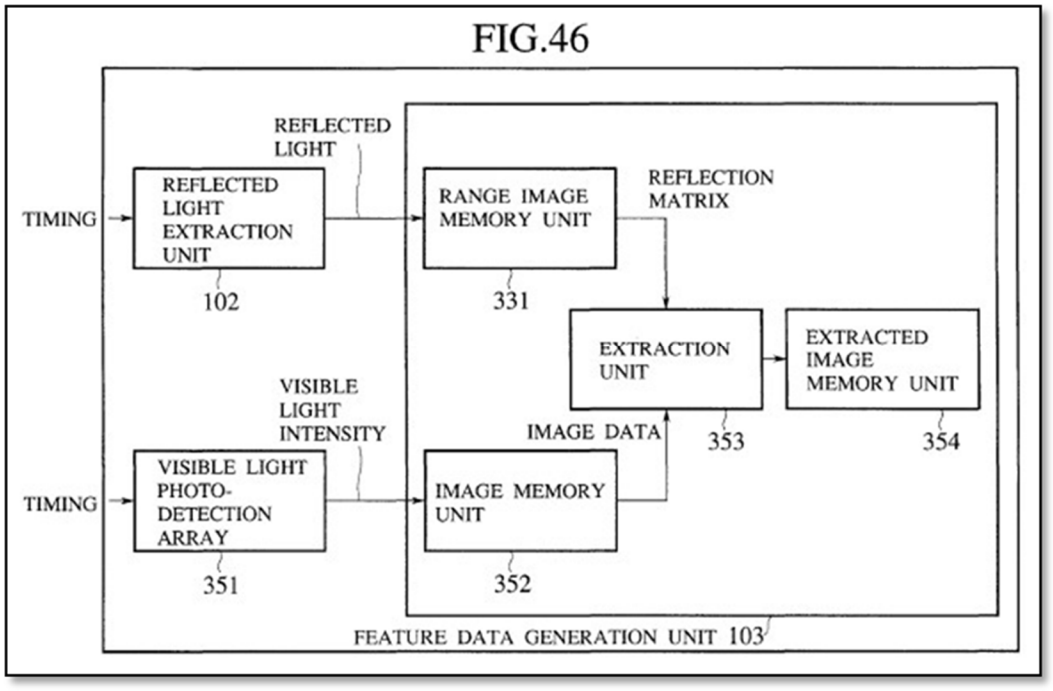
Id. at Fig. 74.

[1(a)] a device housing including a forward facing portion, the forward facing portion of the device housing encompassing an electro-optical sensor having a field of view and including a digital camera separate from the electro-optical sensor; and

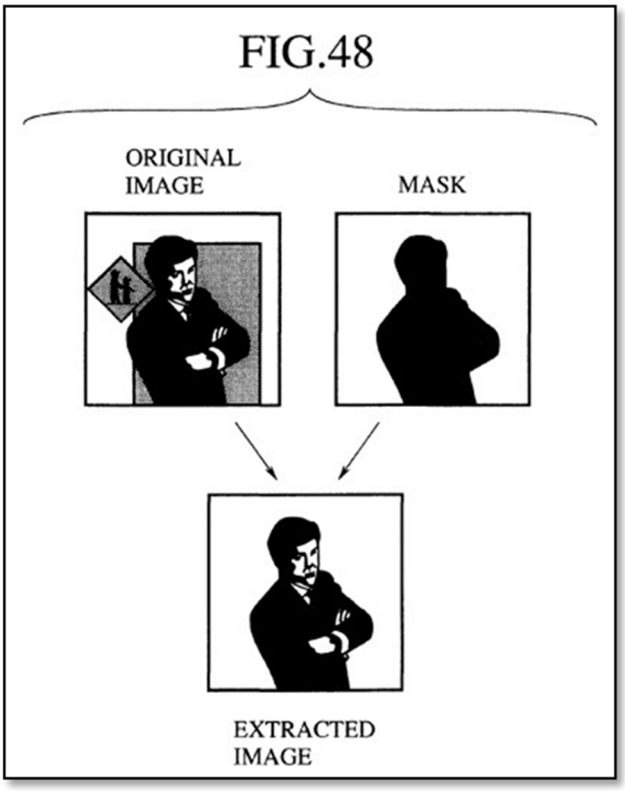
As set forth above, a PHOSITA would have been motivated to implement the fifth embodiment's videoconference functionality in the eighth embodiment's laptop device. *Numazaki's* fifth embodiment teaches "TV telephone" functionality that extracts and transmits only the faces of the participating users, removing extraneous background information that would otherwise consume bandwidth and battery power. *Id*. at 39:12-20. To accomplish this improvement, *Numazaki* uses the same two-camera "reflected light extraction unit 102" (i.e., the same component 102 from

first embodiment)[2] in conjunction with a separate "visible light photo-detection array 351 which is generally used as a CCD camera for taking video images." *Id*. at 39:21-49. This structure is depicted in Fig. 46 below:

[2] A PHOSITA would have understood the "reflected light extraction unit 102" in the fifth embodiment is the same two-sensor structure as "reflected light extraction unit 102" from the first embodiment. *Bederson Dec*. (Ex. 1003), ¶¶ 39-43 (noting (1) both use the same numerical designations and identical language, (2) both refer to the output of component 102 as a "range image," and (3) the fifth embodiment expressly notes it is "another exemplary case of the feature data generation unit in the first embodiment").



Id. at Fig. 46. Using this arrangement, the fifth embodiment processes the output of two-sensor structure 102 to identify an outline of the subject and subtracts everything outside this outline from the image captured by the visible light sensor 351. *Id*. at 39:24-60. Given this functionality in which the output of unit 102 is used to define which portions the video captured by camera 351 are retained, a PHOSITA would have understood that both are forward facing and have overlapping fields of view. *Bederson Dec*. (Ex. 1003), ¶ 52 (explaining that the two outputs must have overlapping fields of view for one to define what is extracted from the other). This process is illustrated in Fig. 48 below:



Numazaki (Ex. 1004), Fig. 48, 39:50-60 (explaining that the "original image" is the output of camera 351 and the mask is the output of sensors 102).

Finally, the "reflected light extraction unit 102" is an "electro-optical sensor" unit and that "visible light photo-detection array 351" is a digital camera as claimed. Starting with the former, although the '949 Patent does not define "electro-optical sensor," dependent claim 7 specifies that the sensor is either a "CCD detector" or "CMOS detector." *'949 Patent* (Ex. 1001), 15:50-52. *Numazaki* expressly teaches that the "reflected light extraction unit 102 . . . has the photo-detection section formed by CMOS sensors." *Numazaki* (Ex. 1004), 12:56-57; *see also id*. at 15:23-27 (noting it is also possible to implement the section using "CCD image sensors").

Turning to the "digital camera," the '949 Patent notes that "representative . . . patents on [d]igital cameras" describe CCD matrix arrays. *'949 Patent* (Ex. 1001), 1:15-18. *Numazaki* teaches that its "visible light photo-detection array 351" is a "CCD camera for taking video images."

[1(b)] a processing unit within the device housing and operatively coupled to an output of the electro-optical sensor, wherein the processing unit is adapted to: determine a gesture has been performed in the electro-optical sensor field of view based on the electro-optical sensor output, and

As set forth above, a PHOSITA would have been motivated to implement *Numazaki's* third embodiment gesture recognition in the eighth embodiment's laptop device. In this embodiment, data reflecting pre-registered gestures or hand positions is stored in "shape memory unit 332." *Numazaki* (Ex. 1004), 29:19-55. This stored data is compared with the output from the two-sensor "reflected light extraction unit"[3] (i.e., component 102 from first embodiment) to identify when the user has performed a pre-registered gesture and instructs the device to implement a "command corresponding to that [gesture]." *Id*. at 29:22-30:5.

A PHOSITA would have understood that *Numazaki's* third embodiment gesture detection process would be implemented by "a processing unit" within

[3] A PHOSITA would have understood that the output of unit 102 necessarily reflects the unit's "field of view." *Bederson Dec*. (Ex. 1003), ¶ 53.

Numazaki's laptop device and adapted (via software) to detect a user's gesture (or sequence of gestures). *Id*. at ¶¶ 53-54 (explaining that laptop computers are equipped with general purpose processors that implement a wide variety of software-based routines and that a PHOSITA would have understood the described gesture detection is a software-based process). To the extent Patent Owner argues *Numazaki's* gesture-detection is not implemented through software, a PHOSITA would have been motivated to implement this gesture detection process in software. *Id*. (explaining that it would have been well within the abilities of a PHOSITA to use the laptop processor to implement *Numazaki's* gesture detection in software and that a PHOSITA would have recognized a software implementation is well suited to the described functionality).

[1(c)] control the digital camera in response to the gesture performed in the electro-optical sensor field of view, wherein the gesture corresponds to an image capture command, and wherein the image capture command causes the digital camera to store an image to memory.

As set forth with respect to the preceding limitation, *Numazaki's* third embodiment detects when a user has performed a pre-registered gesture and instructs the device to implement a "command corresponding to that [gesture]." *Numazaki* (Ex. 1004), 29:22-30:5. As discussed in the Motivation to Combine section above, a PHOSITA would have been motivated to implement this gesture recognition as a means of allowing the user to initiate (or turn on) the fifth embodiment's

videoconferencing functionality. Specifically, pursuant to *Nonaka's* teachings, the user experience would be improved by allowing users to position themselves in place before the video camera and initiate video capture through a gesture, rather than a physical input or timer mechanism. All arguments and evidence from the Motivation to Combine section are hereby incorporated by reference.

Numazaki's fifth embodiment uses a "visible light photo-detection array 351 . . . as a CCD camera for **taking video images**, and an image memory unit 352 for **storing the video images**." *Numazaki* (Ex. 1004), 39:32-35 (emphasis added). Using this arrangement, the fifth embodiment processes the output of two-sensor structure 102 to identify an outline of the subject and subtracts everything outside this outline from the image captured by the digital camera 351. *Id*. at 39:24-60. In doing so, the fifth embodiment arrives at image information that contains only the subject without any background image information and uses much less data as a result of removing the extraneous image information. *Id*. at 40:32-35 ("By recording the extracted image, it is possible to reduce the recording capacity considerably (in an order of one tenth to one hundredth)[.]"). *Numazaki* teaches that the "extracted image is stored in the extracted image memory unit 354." *Id*. at 39:59-60.

Accordingly, in the proposed implementation of *Numazaki* pursuant to the teachings of *Nonaka*, *Numazaki's* laptop causes digital camera 351 to capture and

store images in response to detecting a predetermined "image capture gesture" in the sensor unit 102's field of view. As explained in the Motivation to Combine section, a PHOSITA would have been specifically motivated to utilize the lower resolution "reflected light extraction unit" 102 to monitor the scene and detect any gestures performed by the user and the higher resolution "visible light photo-detection array 351" to take full quality images of the scene (when such video is requested by the user). *Bederson Dec*. (Ex. 1003), ¶ 50 (explaining that it was widely recognized in the art that lower resolution monitoring enables a system to quickly respond to changes in the environment and is preferable to higher resolution monitoring).

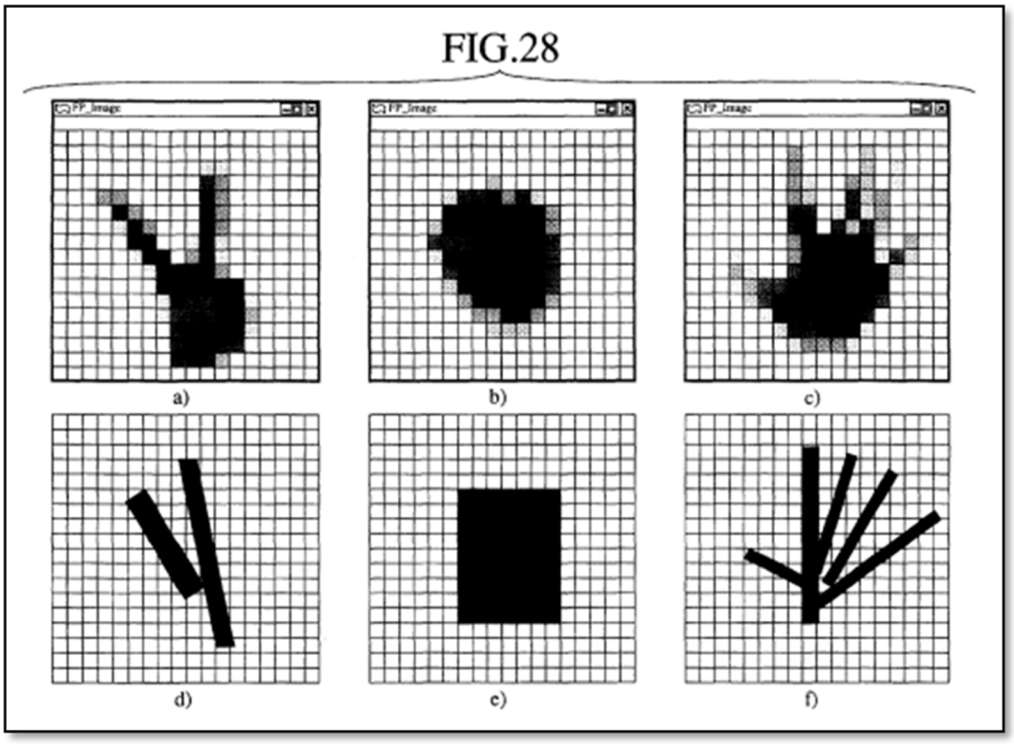
As with the preceding limitation, a PHOSITA would have understood that *Numazaki's* gesture detection and image capture processes would be implemented by "a processing unit" within *Numazaki's* laptop device and adapted (via software) to perform these functions. *Bederson Dec.* (Ex. 1003), ¶ 55 (explaining that laptop computers are equipped with general purpose processors that implement a wide variety of software-based routines and that a PHOSITA would have understood the described image capture and storage is a software-directed process). To the extent Patent Owner argues that *Numazaki's* image capture process is not governed by software, a PHOSITA would have been motivated to implement this imaging process in software. *Id*. (explaining that it would have been well within the abilities

of a PHOSITA to use the laptop processor to implement *Numazaki's* gesture detection and image capture and storage in software and that a PHOSITA would have recognized a software implementation is well suited to the described functionality).

ii. Claim 2

The portable device of claim 1 wherein the determined gesture includes a hand motion.

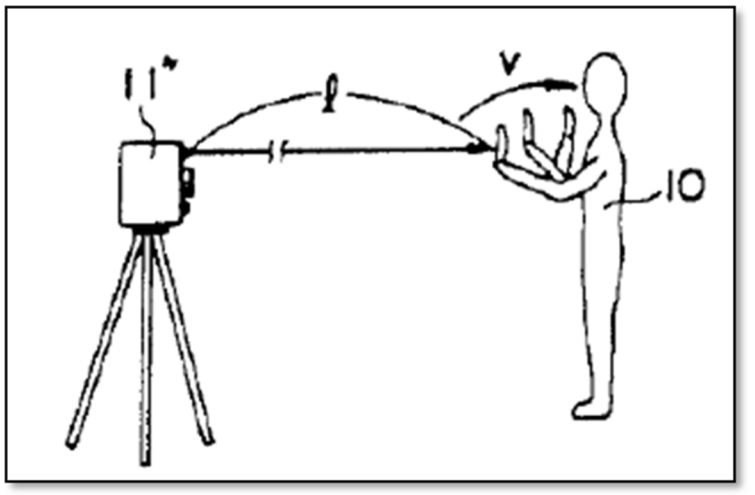
As discussed above with reference to limitation [1(b)], *Numazaki* teaches detecting predetermined gestures and converting them into commands. However, *Numazaki* cautions that gestures can be erroneously detected when the user did not intend to command the system and proposes a sequence of gestures to prevent such "erroneous commanding." *Numazaki* (Ex. 1004), 31:11-44, Fig. 29 (describing a sequence of "two fingers, fist, five fingers" that causes the system to "power on"). Such a sequence (frames a-c) and the rectangular data extracted from the images (frames d-f) are illustrated in Fig. 28 below:



Id. at Fig. 28. *Nonaka* is similarly concerned with ensuring a user intends to have performed an image capture gesture, and proposes its own solution for avoiding erroneous gesture detection. *Nonaka* (Ex. 1005), 3:34-5:26 (describing a process by which the user must perform "the motion for instructing release . . . twice in a row, [so] there is no uncertainty about the remote release operation"). Given both references' concerns over avoiding erroneous gesture detection, a PHOSITA would have been motivated to implement *Numazaki's* multiple-gesture approach to triggering image capture. *Bederson Dec*. (Ex. 1003), ¶ 56 (explaining that *Numazaki's* "two fingers, fist, five fingers" is sufficiently unique that it would be effective at avoiding erroneous gesture detection). A PHOSITA would have further

understood that sequencing between these hand signals necessarily involves hand motion. *Id*.

Alternatively, it would have been obvious to implement *Nonaka's* velocity-based image capture gesture in *Numazaki's* laptop device. *Nonaka* contemplates multiple predetermined gestures, including moving one's hand toward the camera as depicted in Fig. 7 below:



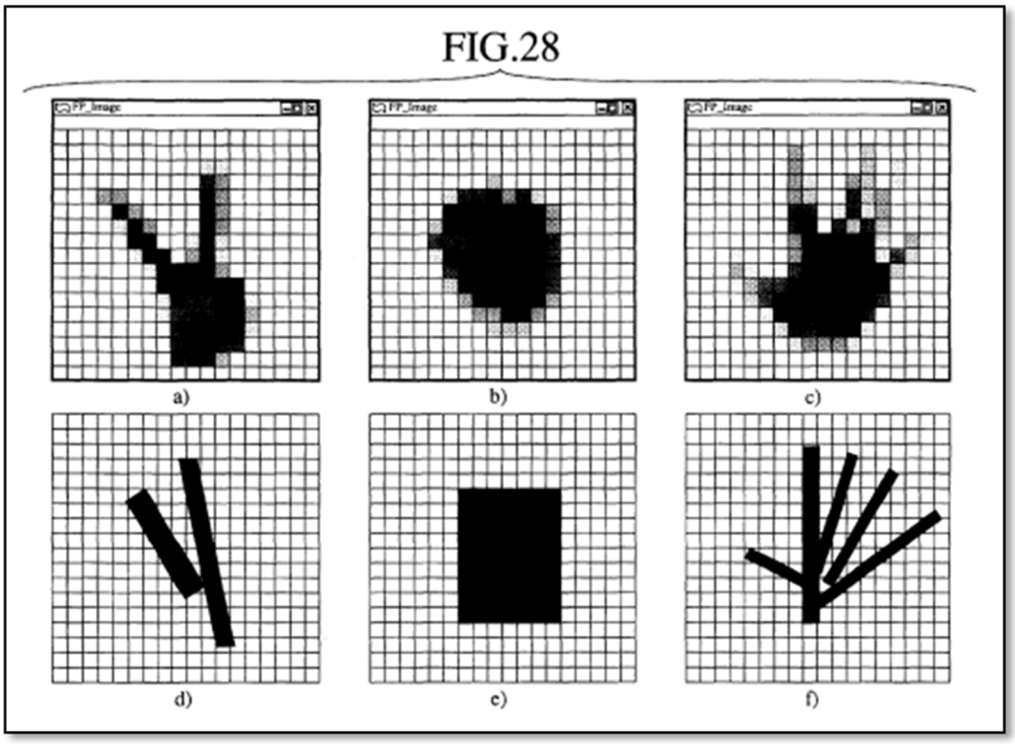
Nonaka (Ex. 1005), Fig. 7, 6:11-22 (describing "the photographer giv[ing] a release instruction . . . [by] mov[ing] his or her hand, etc. at a certain speed"). *Numazaki's* third embodiment expressly contemplates detecting gestures comprised of a user's hand (or finger) toward the device. Namely, *Numazaki* teaches that frames (a) and (b) in Fig. 27 represent the "same hand action," but are distinguished in that (b) represents a hand closer to the device than (a). *Numazaki* (Ex. 1004), 30:9-38. By

measuring the "transition time" between these two states, the system can detect a hand "mov[ing] forward into the depth direction" and convert this into a command. *Id*. A PHOSITA would have expected success in implementing such a motion-based gesture in *Numazaki's* laptop embodiment as a trigger to initiate video capture as discussed above. *Bederson Dec*. (Ex. 1003), ¶ 57.

iii. Claim 3

The portable device of claim 1 wherein the determined gesture includes a pose.

As discussed above with reference to Claim 2, a PHOSITA would have been motivated to implement *Numazaki's* process that detects a predefined sequence of gestures such as a hand motion that sequences through "two fingers, fist, five fingers" as illustrated in Fig. 28 below:



Numazaki (Ex. 1004), Fig. 28. The '949 Patent distinguishes between a single pose and a sequence of poses and notes that a sequence of poses constitutes a "gesture." *'949 Patent* (Ex. 1001), 5:34-36. Based on this disclosure, a PHOSITA would have considered *Numazaki's* "two fingers, fist, five fingers" transition to be a sequence of poses. *Bederson Dec*. (Ex. 1003), ¶ 58. Accordingly, *Numazaki's* gesture "includes a pose" as claimed.

iv. Claim 4

The portable device of claim 1 wherein the electro-optical sensor is fixed in relation to the digital camera.

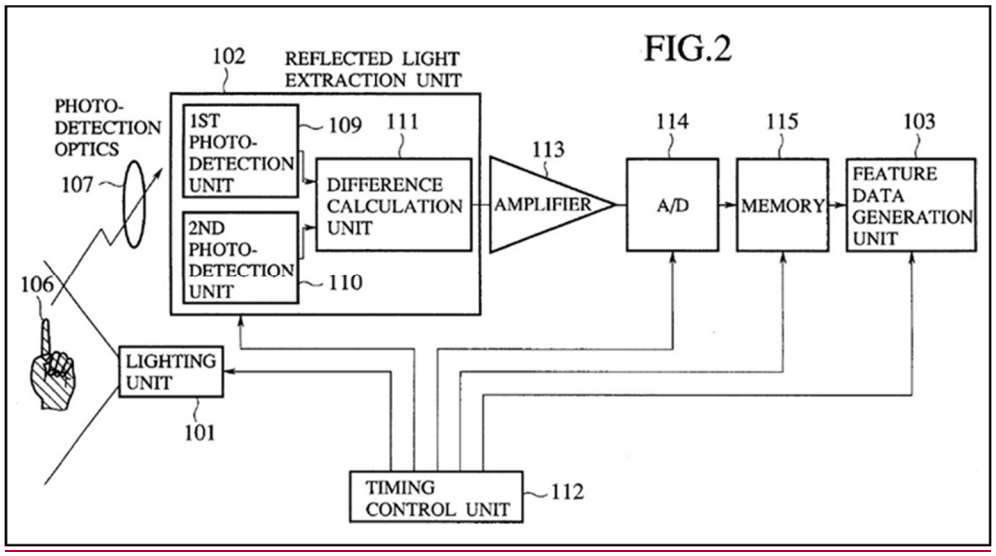
As discussed above with reference to limitation [1(a)], *Numazaki's* fifth embodiment positions an electro-optical sensor (i.e., "reflected light extraction unit

102) and digital camera (i.e., visible light photo-detection array 351) side-by-side such that they have overlapping fields of view. Indeed, *Numazaki* expressly teaches that "visible light photo-detection array 351 and the reflected light extraction unit 102 are arranged in parallel." *Numazaki* (Ex. 1004), 39:4-44.

v. Claim 5

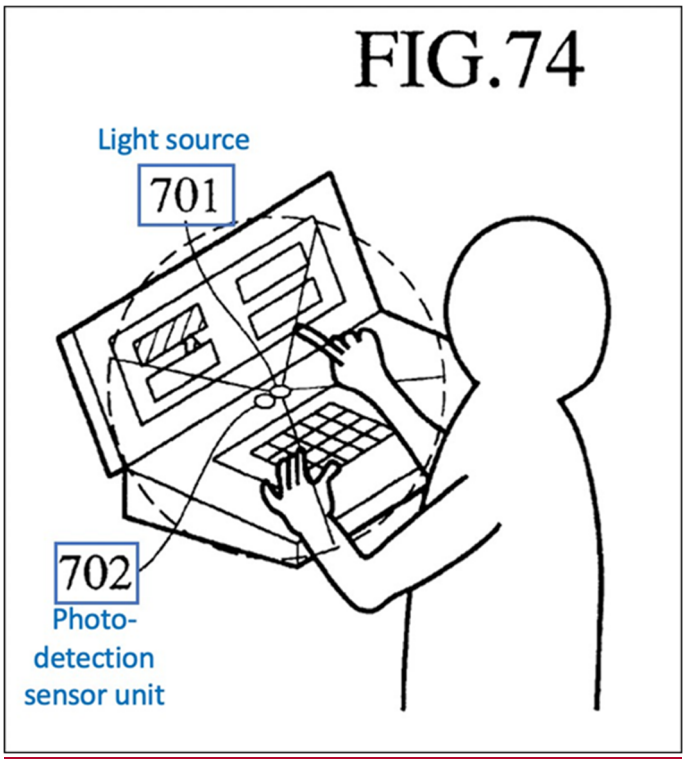
The portable device of claim 1 further including a forward facing light source.

As discussed above with reference to limitation [1(a)], *Numazaki's* fifth embodiment electro-optical sensor (i.e., "reflected light extraction unit 102) is the same two-sensor structure as "reflected light extraction unit 102" from the first embodiment. The premise of the first embodiment is that a controlled light source is used to illuminate the target object (e.g., the user's hand) and two image sensors are used to obtain an image of the object. *Numazaki* (Ex. 1004), 11:9-23. This arrangement is illustrated in Fig. 2 below:



Id. at Fig. 2. A timing control unit is used to turn lighting unit 101 on (i.e., illuminating the target object) when the first camera unit is active and off when the second camera unit is active. *Id*. at 11:20-32. The result of this light control is the first camera unit captures an image of the target object illuminated by both natural light and the lighting unit 101 and the second camera unit captures an image of the target object illuminated by only natural light. *Id*. at 11:33-39. The difference between the two images—obtained by difference calculation unit 111—represents the "reflected light from the object resulting from the light emitted by the lighting unit 101." *Id*. at 11:43-51. This information is then used by feature data generation unit 103 to determine gestures, pointing, etc. of the target object that may be converted into commands executed by a computer. *Id*. at 10:57-66. *Numazaki* further clarifies that both the light source and image sensor structures are implemented in its eighth embodiment laptop device. *Id*. at 50:29-37. (noting lighting unit 701 and a

photo-detection sensor unit 702" are implemented in the laptop). This arrangement is illustrated in Fig. 74 below:



Id. at Figure 74 (annotated to indicate light source 701 and photo-detection sensor unit 702).

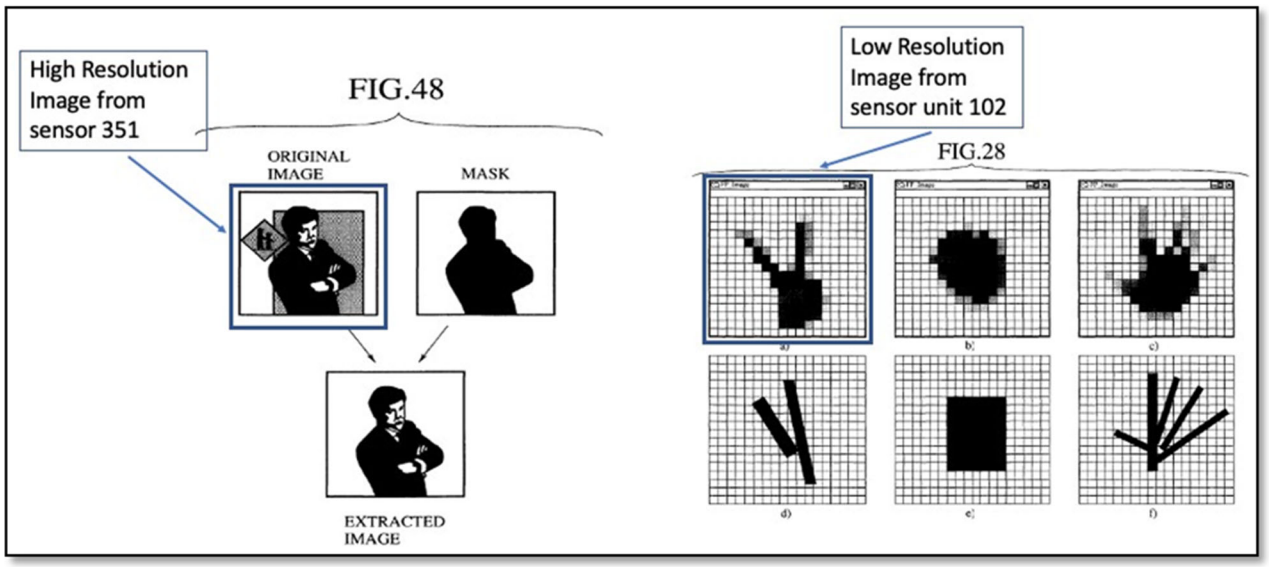
iv. Claim 6

The portable device of claim 1 wherein the electro-optical sensor defines a resolution less than a resolution defined by the digital camera.

Numazaki teaches that its electro-optical sensor (i.e., reflected light extraction unit) "can be manufactured at lower cost . . if it is possible to obtain" the desired information about the user's hand position/gesture "from the rough reflected light

image." *Numazaki* (Ex. 1004), 17:51-56. *Numazaki* further explains that "it is possible to obtain[] the pointing position [of a finger] by the resolution above the pixels." *Id*. at 18:15-21. A PHOSITA would have understood the principle described here to mean that a lower cost electro-optical sensor should be used when lower resolution imaging is sufficient for the given application. *Bederson Dec*. (Ex. 1003), ¶ 59 (explaining that choosing less expensive imaging components where higher resolution isn't required by the application is consistent with standard product design and component sourcing principles).

Numazaki's gesture recognition using its reflected light extraction unit 102 requires significantly less resolution than the visible light photo-detection array 351—the electro-optical sensor and digital camera, respectively, in *Numazaki's* fifth embodiment. Even a cursory assessment of these sensor outputs demonstrates that *Numazaki's* video camera is designed with much higher resolution than its gesture-detecting reflected light extraction unit 102:



Numazaki (Ex. 1004), Fig. 48 (annotated to indicate output of digital camera 351), 39:50-60 (describing the same), Fig. 28 (annotated to indicate output of reflected light extraction unit 102), 29:22-53 (noting the gesture recognition is performed using the output of reflected light extraction unit), 31:11-26 (describing Fig. 28). A PHOSITA would have understood that, based on the outputs of these units and *Numazaki's* express teaching that low resolution gesture detection sensors should be used to save costs, *Numazaki's* reflected light extraction unit 102 would have less resolution than the visible light photo-detection array 351. *Bederson Dec*. (Ex. 1003), ¶ 60 (further citing U.S. Patent No. 5,666,157 to David G. Aviv ("*Aviv*") (Ex. 1006) as evidence that it was known in the art to utilize a lower resolution image

sensor for scene monitoring and a higher resolution camera to provide a detailed video recording when a decision is made to capture the scene).[4]

vii. Claim 7

The portable device of claim 1 wherein the electro-optical sensor includes at least one of a CCD detector and a CMOS detector.

Numazaki expressly teaches that the "reflected light extraction unit 102 . . . has the photo-detection section formed by CMOS sensors." *Numazaki* (Ex. 1004), 12:56-57; *see also id*. at 15:23-27 (noting it is also possible to implement the section using "CCD image sensors").

viii. Claim 8

[8(P)] A computer implemented method comprising:

As set forth above with respect to Claim 1, the proposed combination teaches a computer implemented method of detecting an image capture gesture and initiating video capture in response.

[8(a)] providing a portable device including a forward facing portion encompassing a digital camera and an electro-optical sensor, the electro-optical sensor having an output and defining a field of view;

See limitations [1(P)] and [1(a)], *supra.*

[4] Proposed ground 2 below presents an alternative mapping that relies on *Aviv* as a formal component of the obviousness rejection.

[8(b)] determining, using a processing unit, a gesture has been performed in the electro-optical sensor field of view based on the electro-optical sensor output, wherein the determined gesture corresponds to an image capture command; and

See limitations [1(b)] and [1(c)], *supra.*

[8(c)] capturing an image to the digital camera in response to the determined gesture corresponding to the image capture command.

See limitations [1(c)], *supra.*

ix. Claim 9

The method according to claim 8 wherein the determined gesture includes a hand motion.

See Claim 2, *supra*.

x. Claim 10

The method according to claim 8 wherein the determined gesture includes a pose.

See Claim 3, *supra*.

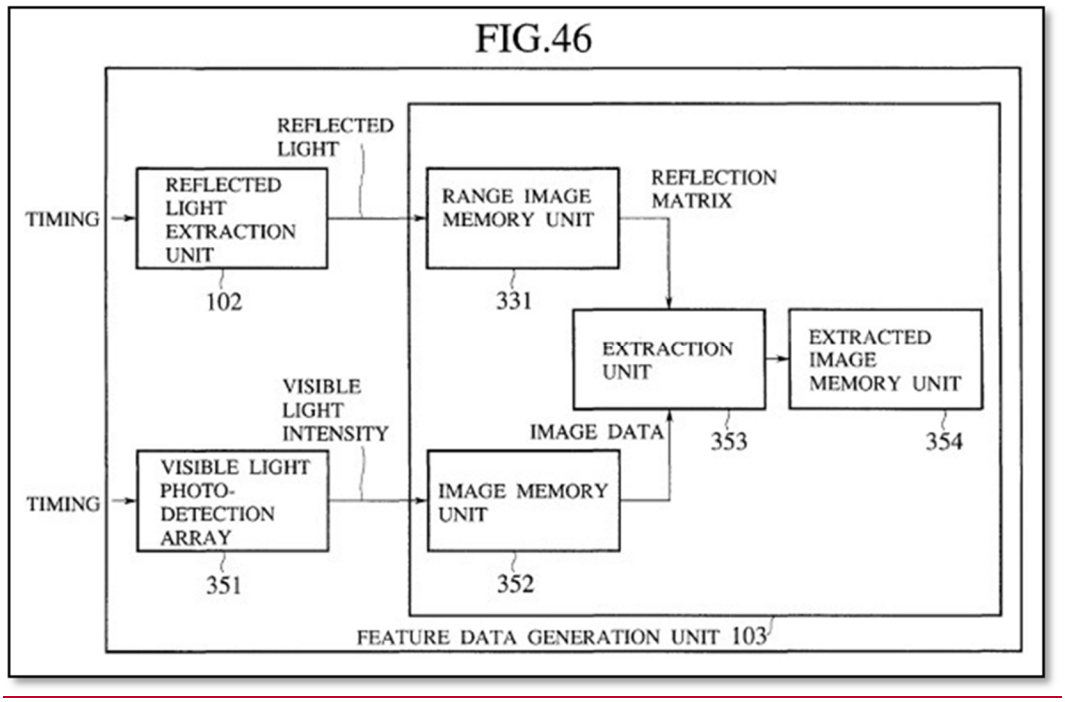
xi. Claim 11

The method according to claim 8 wherein the electro-optical sensor includes first and second sensors in fixed relation relative to the digital camera.

As discussed above with reference to limitation [1(a)], *Numazaki's* fifth embodiment teaches "TV telephone" functionality that extracts and transmits only the faces of the participating users, removing extraneous background information that would otherwise consume bandwidth and battery power. *Id*. at 39:12-20. To accomplish this improvement, *Numazaki* uses the same two-camera "reflected light

extraction unit 102" (i.e., the same component 102 from first embodiment)[5] in conjunction with a separate "visible light photo-detection array 351 which is generally used as a CCD camera for taking video images." *Id*. at 39:21-49. This structure is depicted in Fig. 46 below:

[5] A PHOSITA would have understood the "reflected light extraction unit 102" in the fifth embodiment is the same two-sensor structure as "reflected light extraction unit 102" from the first embodiment. *Bederson Dec*. (Ex. 1003), ¶¶ 39-43 (noting (1) both use the same numerical designations and identical language, (2) both refer to the output of component 102 as a "range image," and (3) the fifth embodiment expressly notes it is "another exemplary case of the feature data generation unit in the first embodiment").



Id. at Fig. 46. Using this arrangement, the fifth embodiment processes the output of two-sensor structure 102 to identify an outline of the subject and subtracts everything outside this outline from the image captured by the visible light sensor 351. *Id*. at 39:24-60.

Numazaki's third embodiment also uses the two-sensor reflected light extraction unit 102 for its gesture detection functionality. *Numazaki* explains that its "third embodiment is directed to another exemplary case of the feature data generation unit of the first embodiment, which realizes a gesture camera for recognizing the hand action easily[.]" *Id*. at 29:4-8. It also expressly notes that a "range image [is] extracted by the reflected light extraction unit." *Id.* at 29:11-13. As Dr. Bederson discussed with reference to the fifth embodiment, *Numazaki* refers

to the two-sensor gesture detection sensor in the first embodiment by this exact name and to its output as a "range image." For these same reasons, a PHOSITA would understand that the two-sensor reflected light extraction unit 102 from *Numazaki's* first embodiment is utilized (1) for gesture detection in the third embodiment and (2) to identify a subject's outline for the fifth embodiment's video image extraction functionality. *Bederson Dec*. (Ex. 1003), ¶¶ 39-43.

Finally, as addressed above with reference to Claim 4, *Numazaki's* electro-optical sensor (i.e., the two-sensor reflected light extraction unit 102) is in fixed relation relative to the digital camera (i.e., visible light photo-detection array 351).

xii. Claim 12

The method according to claim 8 wherein the electro-optical sensor defines a resolution less than a resolution defined by the digital camera.

See Claim 6, *supra*.

xiii. Claim 13

[13(P)] An image capture device comprising:

See limitations [1(P)] and [1(a)], *supra*.

[13(a)] a device housing including a forward facing portion, the forwarding facing portion encompassing a digital camera adapted to capture an image and having a field of view and encompassing a sensor adapted to detect a gesture in the digital camera field of view; and

See limitations [1(a-c)], *supra*.

[13(b)] a processing unit operatively coupled to the sensor and to the digital camera, wherein the processing unit is adapted to: detect a gesture has been performed in the electro-optical sensor field of view based on an output of the electro-optical sensor, and

See limitations [1(b)], *supra*.

[13(c)] correlate the gesture detected by the sensor with an image capture function and subsequently capture an image using the digital camera, wherein the detected gesture is identified by the processing unit apart from a plurality of gestures.

See limitations [1(c)], *supra*.

Further, the image capture gesture in the proposed combination is identified "apart from a plurality of gestures" as claimed. *Numazaki* teaches that multiple gestures can be stored in "shape memory unit" such that, "when [a] matching shape is found, a command corresponding to that shape is outputted. *Numazaki* (Ex. 1004), 29:54-30:3; see also id. at 31:3-10 (noting that commands can include "means for instructing the power ON/OFF of the TV, the lighting equipment, etc."). In the proposed combination, *Numazaki* is implemented to detect a gesture that corresponds to an image capture command pursuant to the teachings of *Nonaka*. Because *Numazaki* can distinguish between different gestures, each of which corresponds to a unique command, when the proposed combination detects this image capture command gesture, it is identified "apart from a plurality of [other] gestures."

xiv. Claim 14

The image capture device of claim 13 wherein the detected gesture includes a hand motion.

See Claim 2, *supra*.

xv. Claim 15

The image capture device of claim 13 wherein the detected gesture includes a pose.

See Claim 3, *supra*.

xvi. Claim 16

The image capture device of claim 13 further including a forward facing light source.

See Claim 5, *supra*.

xvii. Claim 17

The image capture device of claim 13 wherein the sensor defines a resolution less than a resolution defined by the digital camera.

See Claim 6, *supra*.

xviii. Claim 18

The image capture device of claim 13 wherein the sensor is fixed in relation to the digital camera.

See Claim 4, *supra*.

B. Ground 2: Claims 6, 12, and 17 are obvious under pre-AIA 35 U.S.C. § 103 over *Numazaki* in view of *Nonaka* and in further view of *Aviv*

Overview of Aviv

U.S. Patent 5,666,157 to David G. Aviv ("*Aviv*") (Ex. 1006) issued on September 9, 1997, and is prior art to the '949 Patent under at least 35 U.S.C. § 102(b) (pre-AIA). *Aviv* was not cited or considered during prosecution of the '949 Patent. *'949 Patent* (Ex. 1001).

Aviv teaches a camera-based surveillance system for identifying "known characteristics of movements which are indicative of individuals having a criminal intent." *Aviv* (Ex. 1007), Abstract. Reference gestures indicative of criminal intent ("signatures of known criminal acts") are stored in memory and used to compare newly captured gestures to identify subjects engaged in criminal activity:

> **[A] collection of differencing signals for each moved object** of subsequent sampled frames of images (called a "track") **allows a determination of the type, speed and direction** (vector) **of each motion involved**, processing which will extract acceleration, i.e., note of change of velocity: and change in acceleration with respect to time (called "jerkiness"), **and correlating this with stored signatures of known physical criminal acts**. **For example,** subsequent differencing signals may reveal that **an individual's arm is moving to a high position, such as the upper limit of that arm's motion, i.e., above his head) at a fast speed**. This particular movement **could be perceived**, as described below, **as a hostile movement with a possible criminal activity** requiring the expert analysis of security personnel.

Id. at 6:13-26 (emphasis added), 7:5-10 ("[P]rocessed signal is sent to comparison means 14 which compares selected f[r]ames of the video signals from the picture input means 10 with 'signature' video signals stored in memory 16. The signature signals are representative of various positions and movements of the body p[a]rts of an individual having various levels of criminal intent.").

Aviv teaches that "the picture input means 10, may be any conventional electronic picture pickup device operational within the infrared or visual spectrum (or both) including a vidicon and a CCD/TV camera of **moderate resolution**." *Id*. at 4:19-23 (emphasis added). If the system detects a subject performing a gesture that matches a stored signature indicative of criminal intent, "an output 'alert' signal is sent from the comparison means 14 to a controller 18. The controller 18 controls the operation of a secondary, **high resolution** picture input means (video camera)." *Id*. at 7:14-18 (emphasis added). "The purpose of the secondary camera 20 is to provide a **detailed video signal** of the individual having assumed criminal intent." *Id*. at 7:26-28 (emphasis added).

Because *Aviv*, like the '949 Patent, discloses a camera system that may be controlled by human gesture input, *Aviv* is in the same field of endeavor as the '949 Patent. *Compare Aviv* (Ex. 1006), 6:13-26, 7:5-10 (describing gesture recognition of a subject indicative of criminal intent) *with '949 Patent* (Ex 1001), 5:24-49

(describing the process of identifying a subject has performed a predetermined gesture and controlling the camera device to capture an image in response). *Aviv* is therefore analogous art to the '949 Patent. *Bederson Dec.* (Ex. 1003), ¶¶ 61-63.

Motivation to Modify Numazaki in view of Nonaka and in further view of Aviv

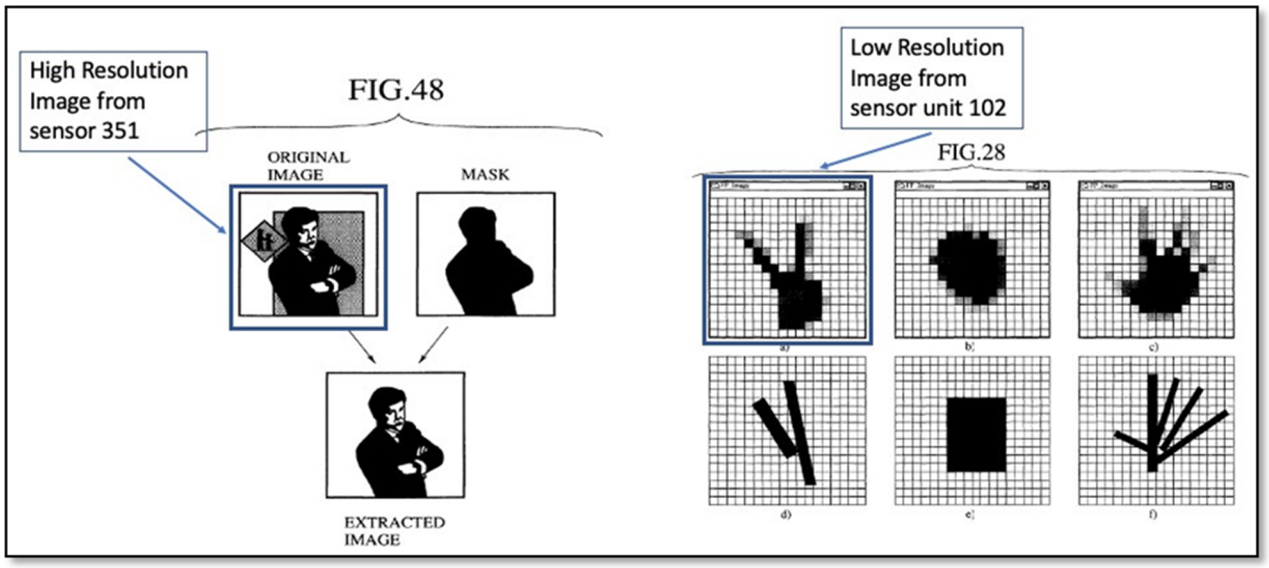
As set forth in Ground 1, a PHOSITA would have been motivated to implement *Numazaki's* eighth embodiment laptop device with its third embodiment gesture recognition and gesture-command mapping and with its fifth embodiment's video capture functionality such that a user can perform a gesture to indicate that the device should begin capturing video, i.e., an image capture command gesture pursuant to the teachings of *Nonaka*. All arguments and evidence from Ground 1 are hereby incorporated by reference.

As further set forth in the Ground 1 mapping of Claim 6, a PHOSITA would have understood that, based on (1) the outputs of *Numazaki's* reflected light extraction unit 102 and visible light photo-detection array 351 and (2) *Numazaki's* express teaching that low resolution gesture detection sensors should be used to save costs, *Numazaki's* reflected light extraction unit 102 would have less resolution than the visible light photo-detection array 351. In the alternative, a PHOSITA would have been motivated to implement *Numazaki's* laptop device in this way pursuant to *Aviv's* teachings.

Aviv summarizes its invention as providing "a surveillance system having at least **one primary video camera** for translating real images of a zone into electronic video signals **at a first level of resolution** . . . [and] means for activating **at least one secondary sensor** and associated recording device **having a second higher level of resolution**, said activating means being in response to determining a predetermined level of criminal activity." *Aviv* (Ex. 1006), 3:17-31. *Numazaki* similarly teaches that its scene monitoring and gesture detecting electro-optical sensor (i.e., reflected light extraction unit) "can be manufactured at lower cost . . if it is possible to obtain" the desired information about the user's hand position/gesture "from the rough reflected light image." *Numazaki* (Ex. 1004), 17:51-56. *Numazaki* further explains that "it is possible to obtain[] the pointing position [of a finger] by the resolution above the pixels." *Id*. at 18:15-21. A PHOSITA would have understood the principle described here to mean that a lower cost electro-optical sensor should be used when lower resolution imaging is sufficient for the given application. *Bederson Dec*. (Ex. 1003), ¶¶ 64-65 (explaining that choosing less expensive imaging components where higher resolution isn't required by the application is consistent with standard product design and component sourcing principles). To accommodate this design goal, a PHOSITA would have been motivated to implement *Numazaki's* gesture detecting electro-

optical sensor (i.e., reflected light extraction unit) at a lower resolution pursuant to the teachings of *Aviv*. *Id*.

Further supporting the conclusion that a PHOSITA would have been motivated to implement *Numazaki's* gesture recognizing reflected light extraction unit 102 with lower resolution than the visible light photo-detection array 351, *Numazaki's* illustrations of these sensor outputs demonstrate unit 351 produces a higher resolution image than unit 102:



Numazaki (Ex. 1004), Fig. 48 (annotated to indicate output of digital camera 351), 39:50-60 (describing the same), Fig. 28 (annotated to indicate output of reflected light extraction unit 102), 29:22-53 (noting the gesture recognition is performed using the output of reflected light extraction unit), 31:11-26 (describing Fig. 28). A

PHOSITA would have understood that, based on the outputs of these units and *Numazaki's* express teaching that low resolution gesture detection sensors should be used to save costs, *Numazaki's* reflected light extraction unit 102 should be implemented with lower resolution than the visible light photo-detection array 351, pursuant to the teachings of *Aviv*. *Bederson Dec*. (Ex. 1003), ¶ 66.

xix. Claim 6

The portable device of claim 1 wherein the electro-optical sensor defines a resolution less than a resolution defined by the digital camera.

As set forth above, a PHOSITA would have been motivated to implement *Numazaki's* gesture recognizing reflected light extraction unit 102 with lower resolution than the visible light photo-detection array 351.

xx. Claim 12

The method according to claim 8 wherein the electro-optical sensor defines a resolution less than a resolution defined by the digital camera.

See Claim 6, *supra*.

xxi. Claim 17

The image capture device of claim 13 wherein the sensor defines a resolution less than a resolution defined by the digital camera.

See Claim 6, *supra*.

V. DISCRETIONARY CONSIDERATIONS

A. The *Fintiv* Factors Favor Institution

Under the "*Fintiv* factors," the Board has asserted that it may consider parallel litigation, including an early trial date, in determining whether to institute under 35 U.S.C. § 314(a). *NHK Spring Co., Ltd., v. Intri-Plex Technologies, Inc.*, IPR2018-00752, Paper 8 at 19–20 (PTAB Sept. 12, 2018) (precedential); *Apple Inc. v. Fintiv, Inc.*, IPR2020-00019, Paper 11 at 2–3, 6 (PTAB Mar. 20, 2020) (enumerating six factors that "relate to whether efficiency, fairness, and the merits support the exercise of authority to deny institution in view of an earlier trial date in [a] parallel proceeding") (precedential) [6]. Here, the *Fintiv* factors weigh strongly in favor of institution, because Petitioner merely seeks to join an already instituted IPR. *See AT&T Servs., Inc., et al. v.* ~~). Those factors favor institution here~~*Broadband iTV, Inc.*, IPR2021-00556, Paper 14 at 9 (P.T.A.B. Feb. 19, 2021) (granting institution and a motion for joinder of a petition that was "identical," noting that "the issue of whether to exercise our discretion to deny institution of a trial based on the instant grounds of unpatentability was addressed and decided in the [instituted IPR]").

[6] Petitioner reserves the right to challenge the *Fintiv* framework because: (1) it exceeded the Director's authority; (2) it is arbitrary and capricious; and (3) it was adopted without notice-and-comment rulemaking.

Nevertheless, if the Board proceeds with analyzing the *Fintiv* factors, they favor institution here, as discussed in detail below.

~~**a. The *Fintiv* factors strongly favor institution**~~

~~1.~~1. ***Whether the court granted a stay or evidence exists that one may be granted if a proceeding is instituted.***

~~The litigation~~Google is not a named party in any pending proceeding.

Judge Albright in the District Court for the Western District of Texas ~~has just begun. Neither party has requested a stay yet, though Petitioner intends to do so if institution is~~ granted Apple Inc.'s motion to stay pending the final written decision of the instituted underlying *inter partes* review of the '949 patent (IPR2021-00921). *Gesture Technology Partners, LLC v.* ~~. Petitioner has acted expeditiously in filing this Petition since the~~*Apple Inc.*, No. 6:21-cv-00121 (W.D. Tex. Dec. 29, 2021).

Motions to stay are pending in two other pending district court cases involving the '949 Patent. *Gesture Technology Partners, LLC v. LG Electronics, et al.*, No. 2:21-cv-19234, Dkt. 39 (D.N.J. Nov. 24, 2021) and *Gesture Tech. Partners, LLC v. Huawei Device Co., Ltd. et al.*, No. ~~litigation began.~~ 2:21-cv-00040, Dkt. 157 (E.D.

Tex. Dec. 16, 2021) (the motion is filed by the Samsung defendants in 2:21-cv-00041 which is consolidated with 2:21-cv-00040).[7]

Judge Albright has granted a motion to dismiss the case against Lenovo (United States) Inc. and Motorola Mobility LLC for improper venue. *Gesture Technology Partners, LLC v. Lenovo Group Ltd. et al.*, No. 6:21-cv-00122, Dkt. 43 (W.D. Tex. Dec. 29, 2021). The remaining defendant in that case—Lenovo Group Ltd.—was not served. *Id.* at 1 n.1.

Accordingly, this factor ~~should favor~~favors institution. At worst, this factor is neutral because the Board "will not attempt to predict" how the district court will proceed. *Sand Revolution II, LLC v. Continental Intermodal Group-Trucking LLC*, IPR2019-01393, Paper 24 at 7 (PTAB June 16, 2020) (informative).

~~2.~~2. ***Proximity of the court's trial date to the Board's projected statutory deadline for a final written decision***.

~~Although no trial date~~ As mentioned above, the litigation against Apple has been ~~set, the projected trial date in the Texas Litigation should be scheduled in early~~

[7] In *Gesture Tech. Partners, LLC v. Huawei Device Co., Ltd. et al.*, the action against Huawei was dismissed with prejudice.

~~2023 or later.[8] The trial therefore is expected to occur *after* the expected~~ stayed pending the final written decision on the IPRs, so there is no trial date. *Gesture Technology Partners, LLC v. Apple Inc.*, No. ~~(FWD). "This factor looks at~~6:21-cv-00121 (W.D. Tex. Dec. 29, 2021).

As mentioned above, the ~~proximity~~litigation against two of the ~~trial date to the date of [FWD] to assess~~defendants in the case against Lenovo has been dismissed, and the third defendant was not served. *See Gesture Technology Partners, LLC v. Lenovo Group Ltd. et al.*, No. ~~weight to be accorded a trial date set earlier than~~6:21-cv-00122, Dkt. 43 (W.D. Tex. Dec. 29, 2021).

In the ~~expected [FWD] date." IPR2020-00944, Paper~~consolidated case numbers 2:21-cv-00040 and 2:21-cv-00041, the current trial date is March 7, 2022, but that date is likely to change, as the claims against Huawei were dismissed with prejudice and the Samsung defendants have filed a motion to stay pending the outcome of the instituted IPRs against the '949 Patent, and the day is triple booked with trials schedule to begin on the same day. *Gesture Tech. Partners, LLC v. Huawei Device Co., Ltd., et al.*, No. 2:21-cv-00040, Dkt. 168 (E.D. Tex. Dec. 22,

~~[8] A recent Order Governing Proceedings in patent cases issued by Judge Alan D. Albright proposes a *Markman* hearing 23 weeks after the Case Management Conference ("CMC") and trial 52 weeks after *Markman*. Ex. 1007 at 8, 10. Here, no CMC has been scheduled, and Petitioner does not anticipate a CMC in the coming months.~~

2021); *Gesture Tech. Partners, LLC v. Huawei Device Co., Ltd. et al.*, No. 2:21-cv-00040, Dkt. 157 (E.D. Tex. Dec. 16, 2021) (the motion is filed by the Samsung defendants in 2:21-cv-00041 which is consolidated with 2:21-cv-00040).

The case against LG was only recently transferred to the district of New Jersey and does not yet even have a scheduling order. The LG defendants have moved to stay the case pending IPR. *Gesture Technology Partners, LLC v. LG Electronics, et al.*, No. 2:21-cv-19234, Dkt. 39 (D.N.J. Nov. 24, 2021). The plaintiff requested an extension of the deadline for filing opposition and reply briefs concerning the motion to stay pending IPR. *Gesture Technology Partners, LLC v. LG Electronics, et al.*, No. 2:21-cv-19234, Dkt. 44 (D.N.J. Dec. 14, 2021). The court granted the request, except that any opposition shall be filed by December 31st, and any reply shall be filed by January 7, 2022. *Gesture Technology Partners, LLC v. LG Electronics, et al.*, No. 2:21-cv-19234, Dkt. 45 (D.N.J. Dec. 15, 2021). The court has ordered a scheduling conference on January 20, ~~61~~2022. *Gesture Technology Partners, LLC v. LG Electronics, et al.*, No. 2:21-cv-19234, Dkt. 37 (D.N.J. Oct. 27, 2021).

Because of the distant and uncertain trial dates in the pending litigations, this factor should favor institution. As recognized by the Board, where "there is at least some persuasive evidence that delays are possible," trial dates upward of six months

before the FWD are insufficient to deny institution. *Id*. Here, the law and facts support the same conclusion.

As the Federal Circuit explained in *In re Apple*, "a court's general ability to set a fast-paced schedule is not particularly relevant," especially where "the forum ~~[i.e., W.D. Tex.]~~ itself has not historically resolved cases so quickly." *In re Apple Inc*., No. 20-135, slip op. at 16 (Fed. Cir. Nov. 9, 2020) (Ex. 1008); *see* IPR2020-01280, Paper 17, 13-16. Indeed, the Federal Circuit found error in reliance on the case's scheduled trial date. *In re Apple Inc*., No. 20-135, slip op. at 15. Similarly, adhering to the Federal Circuit's guidance, it would be error for the Board to rely upon the current schedule. ~~And, if the Board instead projects the trial date using the WDTX's average time to trial—which itself leaves the Board to error-prone speculation as to the timing of this particular trial—that speculated trial date is (at best) concurrent with the timing of the final written decision. *Id.*~~

~~Notably, here, concerns over speculation transcend the legal error noted by the Federal Circuit, as averages are inherently unreliable in resolving trial dates. This is evident from statistics concerning strikingly frequent trial slippage in WDTX. "70% of [WDTX] trial dates initially relied upon by the PTAB to deny petitions have slid," as of July 2020. District Court Trial Dates Tend to Slip After PTAB Discretionary Denials (Ex. 1009). Such delays even impacted the seminal *NHK* and*Fintiv* cases, where, after the Board denied institution, associated trial dates were delayed to after the expected FWD dates by the courts—the same WDTX court in *Fintiv* as is handling the Texas Litigation. *See* IPR2018-01680, Paper 22 at 17, n. 6(PTAB Apr. 3, 2019) ("In the district court case running parallel to *NHK Spring*, the court ultimately moved the trial date back six months, illustrating the uncertainty associated with litigation schedules."); Order~~

~~Resetting *Fintiv* Jury Trial (Ex. 1010)(resetting *Fintiv* trial to October 4, 2021, nearly five months after the FWD would have been due in the associated IPR). In contrast, despite the pandemic, the Board has adhered to the one-year statutory deadline for FWDs prescribed by 35 U.S.C. §316(a)(11).~~

~~3.~~3. ***Investment in the parallel proceeding by the court and the parties.***

~~The~~ Petitioner is not a named party to any parallel litigation.

In the case against Samsung (2:21-cv-00040 (E.D. Tex.) (lead case) and 2:21-cv-00041 (E.D. Tex.) (member case)), the parties exchanged contentions and received a claim construction order.

In the case against LG (2:21-cv-19234 (D.N.J.)), the parties in that case have invested little ~~in the parallel proceeding. No~~. Before transfer, the case effectively never began. "Although the parties served preliminary infringement ~~or~~and invalidity contentions ~~have been exchanged~~, no claim construction positions have been taken, and no non-venue discovery has been conducted. ~~This factor weighs in favor of institution.~~And the District of New Jersey has yet to set any case schedule or proceed with the litigation in any way." *LG Elecs., Inc., et al. v. Gesture Tech. Partners, LLC*, IPR2022-00091, Paper 1 at 65 (Nov. 5, 2021).

~~4.~~ In the litigation against Apple (6:21-cv-00121 (W.D. Tex.)), Patent Owner submitted an opening claim construction brief and Apple submitted a reply claim construction brief, but the case has been stayed as noted above.

In the litigation against Lenovo, (6:21-cv-00122 (W.D. Tex.)), Patent Owner had submitted its opening claim construction brief, and the hearing was scheduled for February 4, 2022. *Gesture Technology Partners, LLC v. Lenovo Group Ltd. et al.,* No. 6:21-cv-00122, Dkt. 30 (W.D. Tex. Aug. 6, 2021). But, as mentioned above, the litigation against two of the defendants in the case against Lenovo has been dismissed, and the third defendant was not served. *See Gesture Technology Partners, LLC v. Lenovo Group Ltd. et al.,* No. 6:21-cv-00122, Dkt. 43 (W.D. Tex. Dec. 29, 2021).

Because parties in the copending litigations involving the '949 Patent have not invested significantly in those proceedings, and because a great majority of the costs to be expended in litigating the validity of this patent remain yet future, this factor favors institution.

4. ***Overlap between issues raised in the petition and in the parallel proceeding***.

~~No~~There is no evidence that the same invalidity ~~contentions~~grounds sought here are at issue in the other pending litigations involving the '949 patent or even that all of the claims challenged here have been ~~served in the district court litigation, sothere is presently no overlap.~~challenged in any of the pending litigations. This factor ~~weights~~weighs in favor of institution.

5. ***Whether the petitioner and the defendant in the parallel proceeding are the same party***.

~~The parties are the same in the district court litigation. However, members of the Board have noted that *Fintiv* addresses only the scenario in which the petitioner is unrelated to a defendant in a parallel proceeding, finding this should weigh against denying institution, but that *Fintiv* "says nothing about situations in which the petitioner is the same as, or is related to, the district court defendant." *Cisco Sys., Inc*~~Google has not been named as a defendant in any parallel proceeding. Because no parallel litigation between the parties exists, this factor weighs in favor of institution.

6.~~. *v. Ramot at Tel Aviv Univ. Ltd.*, IPR2020-00122, Paper 15, at *10 (PTAB May15, 2020) (APJ Crumbley, dissenting) (noting that disfavoring a "defendant in the district court" is "contrary to the goal of providing district court litigants an alternative venue to resolve questions of patentability").~~

6. ***Other circumstances that impact the Board's exercise of discretion, including the merits***.

As set forth above, the strength of the proposed grounds weighs strongly in favor of institution.

Additionally, the *Fintiv* factors require the Board to take "a holistic view of whether efficiency and integrity of the system are best served by denying or instituting review." *Fintiv*, Paper 11 at 6. Denying institution here is both inefficient and would raise significant concerns about the patent system's integrity. Were the Board to deny institution, years of district court litigation would be required just to

obtain a finding of invalidity. Such a scenario runs directly counter to *Fintiv's* goals of preserving efficiency and the system's integrity.

In sum, denying institution here creates vast inefficiencies, forcing the parties to litigate for potentially years in the district court and harms system integrity by rejecting the PTAB as the "lead agency" in assessing patentability. The Board should decline to exercise its discretion for these additional reasons.

~~**b. The *Fintiv* Framework Should Be Overturned**~~

~~Apart from Petitioner's showing that the *Fintiv* factors favor institution, the *Fintiv* framework should be overturned because it is both legally invalid and unwisepolicy. Specifically, the framework (1) exceeds the Director's authority, (2) is arbitrary and capricious, (3) and was adopted without notice-and-comment rulemaking.~~

~~1. The *Fintiv* framework exceeds the Director's authority~~

~~Under the America Invents Act ("AIA"), the Director has no authority to deny IPR petitions based on parallel infringement litigation if the IPR petition was timelyunder § 315(b). Section 314(a) permits the Director to determine whether "there is areasonable likelihood that the petitioner would prevail with respect to at least 1 of the claims challenged." That phrasing is not an invitation to create grounds for denying institution that conflict with other aspects of the statute. Accordingly, evenassuming § 314(a) gives the Director a degree of discretion to deny institution, he cannot exercise that discretion in a manner that violates or exceeds his authority under the AIA. And the AIA's text, structure, and purpose show that Congress withheld the authority to deny institution based on parallel infringement litigation.~~

~~For example, § 315(b) makes clear that IPR is permissible when parallel litigation is pending if the petition is timely. Because statutorily defined time limitations inherently "take[] account of delay," other "case-specific circumstances" like the *Fintiv* factors "cannot be invoked to preclude adjudication of a claim … brought within the [statutory] window." *Petrella v. Metro-Goldwyn-Mayer, Inc.*, 572 U.S. 663, 667, 677-680, 685 (2014).~~

The AIA's structure confirms that the Director lacks authority to adopt and apply the *Fintiv* framework. Various provisions specify how the Director may or must handle parallel proceedings, showing that Congress carefully considered how to promote both efficiency and the purposes of IPR in the face of parallel proceedings and intended for IPR to be available even when parallel infringement litigation is pending if the IPR petition is timely. *See, e.g.*, §§ 315(a), (d), 325(d). But none of those provisions grants the Director discretion to reject IPR because there is a parallel lawsuit. Thus, Congress "knew how to draft the kind of statutory language that [the Director] seeks to read into" the AIA. *State Farm Fire & Cas. Co. v. United States ex rel. Rigsby*, 137 S. Ct. 436, 443-444 (2016). "[H]ad Congress intended to" grant the Director the discretion he has asserted in the *NHK-Fintiv* framework, Congress "would have said so." *Id.*

Finally, by treating overlap as a reason to deny institution, the *Fintiv* framework contravenes a central purpose of IPR, namely, providing a more efficientadditional pathway for resolving the same issues that the challenger could otherwisehave brought only in litigation. *See* H.R. Rep. No. 98, at 48, 112th Cong., 1st Sess. (2011); *Thryv, Inc. v. Click-To-Call Techs., LP*, 140 S. Ct. 1367, 1374-1375 (2020).

2. The *Fintiv* framework is arbitrary and capricious

The *Fintiv* framework is arbitrary and capricious, and therefore unlawful, in several respects. First, the framework requires the Board to make decisions based on speculation about the course of parallel litigation, producing irrational outcomes and disparities between similarly situated IPR petitioners. *See, e.g.*, *Horsehead Resource Dev. Co. v. Browner*, 16 F.3d 1246, 1269 (D.C. Cir. 1994) ("agency actions based upon speculation are arbitrary and capricious"). Second, the *NHK-Fintiv* factors areso vague and malleable that they lead the Board to "treat similar situations differently without reasoned explanation"—a hallmark of arbitrary and capricious agencyaction. *Port of Seattle v. FERC*, 499 F.3d 1016, 1034 (9th Cir. 2007). And third, the*NHK-Fintiv* framework is not rationally connected to its ostensible purpose of promoting efficiency. On the contrary, the rule undermines that goal by, for example, pressuring infringement defendants to file premature IPR petitions and allowing infringement plaintiffs to block IPR entirely through forum shopping.

3. The *Fintiv* framework was impermissibly adopted withoutnotice-and-comment rulemaking

The *Fintiv* framework is also invalid because it is a rule that was impermissibly adopted without notice-and-comment rulemaking. Through the Director's

~~designationof *NHK* and *Fintiv* as precedential, those decisions' framework became "binding" onthe Board "in subsequent matters involving similar facts or issues," Patent Trial andAppeal Board, Standard Operating Procedure 2 (Rev. 10) ("SOP-2"), at 11 (Sept. 20, 2018) – that is, they became a "rule" as defined in the Administrative Procedure Act("APA"), *see* 5 U.S.C. § 551(4) (defining "rule"). But that designation process did notentail public notice and an opportunity or public comment, *see* SOP-2 at 1-4, 8-12; *Aqua Prods., Inc. v. Matal*, 872 F.3d 1290, 1331-1332 (Fed. Cir. 2017) (Moore, J., concurring), contrary to the requirements of both the APA and the AIA. *See Kisor v.Wilkie*, 139 S. Ct. 2400, 2420 (2019); §§ 2(b)(2), 316(a).~~

VI. CONCLUSION

For the ~~forgoing~~foregoing reasons, Petitioner respectfully requests *inter partes* review of claims 1–18 of the '949 Patent.

Respectfully submitted,

By: ~~/s/ Adam P. Seitz~~
~~Adam P. Seitz~~Erika H. Arner/
Erika H. Arner, Reg. No. ~~52,206~~
~~Paul R. Hart, Reg. No. 59,646~~57,540
COUNSEL FOR PETITIONER

VII. MANDATORY NOTICES UNDER 37 C.F.R. § 42.8(A)(1)

A. Real Party-In-Interest

Petitioner is the real party-in-interest. 37 C.F.R. § 42.8(b)(1~~).~~) [9].

B. Related Matters

Pursuant to 37 C.F.R. § 42.8(b)(2), Petitioner is aware of the following matters involving the '949 Patent: *Gesture Technology Partners, LLC v. LG Electronics, Inc. et al.*, No. 6:21-cv-00123 (W.D. Tex.) (transferred to District of New Jersey as case number 2:21-cv-19234); *Gesture Technology Partners, LLC v. Apple Inc.*, No. 6:21-cv-00121 (W.D. Tex.); *Gesture Technology Partners, LLC v. Lenovo Group Ltd. et al.*, No. 6:21-cv-00122 (W.D. Tex.); *Gesture Technology Partners, LLC v. ~~LG Electronics, Inc. et al.~~*, ~~No. 6:21-cv-00123 (W.D. Tex.); *Gesture Technology Partners, LLC v.~~ *Huawei Device Co., Ltd. et al.*, No. 2-:21-cv-00040 (E.D. Tex.); *Gesture Technology Partners, LLC v. Samsung Electronics Co., Ltd. et al.*, 2-:21-cv-00041 (E.D. Tex~~.)~~.); and *Gesture Technology Partners, LLC v. LG Electronics, et al.*, No. 2:21-cv-19234 (D.N.J.). IPR2021-00921, filed on June 2, 2021,

[9] Google LLC is a subsidiary of XXVI Holdings Inc., a subsidiary of Alphabet Inc. XXVI Holdings Inc. and Alphabet Inc are not real parties in interest to this proceeding.

challenged claims 1-18 of the '949 Patent. IPR2022-00092, filed on November 5, 2021, with a contemporaneous motion for joinder with IPR2021-00921, challenged claims 1-18 of the '949 Patent.

C. Lead and Back-Up Counsel

Petitioner provides the following designation and service information for lead and back-up counsel. 37 C.F.R. § 42.8(b)(3) and (b)(4).

Lead Counsel	**Back-Up Counsel**
~~**Adam P. Seitz**~~Erika H. Arner (Reg. No. ~~**52,206**~~57,540) ~~**Postal and Hand-Delivery Address:**~~ ~~**Erise IP, P.A.**~~ ~~**7015 College Blvd., Ste. 700 Overland Park, Kansas 66211Telephone: (913) 777-5600**~~ ~~**Fax: (913) 777-5601**~~erika.arner@finnegan.com Finnegan, Henderson, Farabow, Garrett, & Dunner LLP 1875 Explorer Street, Suite 800 Reston, VA 20190-6023 Tel: 571-203-2700 Fax: 202-408-4400	~~**Paul R. Hart**~~Daniel C. Cooley (Reg. No. 59,~~**646**~~639) ~~**Postal and Hand-Delivery Address:**~~ ~~**Erise IP, P.A.**~~ ~~**5299 DTC Blvd., Ste. 1340**~~ ~~**Greenwood Village, Colorado 80111**~~ ~~**Telephone: (913) 777-5600**~~ daniel.cooley@finnegan.com Finnegan, Henderson, Farabow, Garrett, & Dunner LLP 1875 Explorer Street, Suite 800 Reston, VA 20190-6023 Tel: 571-203-2700 Fax: ~~**(913) 777-5601**~~202-408-4400

	Mingji Jin (Reg. No. 69,674) mingji.jin@finnegan.com Finnegan, Henderson, Farabow, Garrett, & Dunner LLP 901 New York Avenue, NW Washington, DC 20001-4413 Tel: 202-408-4000 Fax: 202-408-4400

APPENDIX OF EXHIBITS

Exhibit 1001	U.S. Patent No. 8,878,949 ("'949 Patent")
Exhibit 1002	Prosecution History for the '949 Patent ("'949 File History")
Exhibit 1003	Bederson Declaration ("Bederson Dec.")
Exhibit 1004	U.S. Patent 6,144,366 to Numazaki et al. ("Numazaki")
Exhibit 1005	JPH4-73631 to Osamu Nonaka ("Nonaka")
Exhibit 1006	U.S. Patent 5,666,157 to David G. Aviv ("Aviv")
Exhibit 1007	U.S. District Court, W.D. Texas Time to Jury Trial
Exhibit 1008	*In re Apple Inc*., No 20-135 (Fed. Cir. Nov. 9, 2020)
Exhibit 1009	District Court Trial Dates Tend to Slip After PTAB Discretionary Denials
Exhibit 1010	Order Resetting *Fintiv* Jury Trial
Exhibit 1011	"Miniature Space-Variant Active Vision System: Cortex-I," Ph.D. Thesis, NYU, 1992
Exhibit 1012	In Conference Companion on Human Factors in Computing Systems (CHI '95), I. Katz, R. Mack, and L. Marks (Eds.). ACM, New York, NY, USA, 210-211. DOI=http://dx.doi.org/10.1145/223355.223526
Exhibit 1013	"CCD and CMOS Imaging Array Technologies," by Stuart Taylor ("*Taylor*") (Ex. 1014)
Exhibit 1014	U.S. Patent No. 6,359,647 to Sengupta, et al. ("*Sengupta*")
Exhibit 1015	Dana Harry Ballard and Christopher M. Brown. 1982. "Computer Vision" (1st. ed.). Prentice Hall Professional Technical Reference
Exhibit 1016	Wallace, R.; Ong, P.; Bederson, B.; Schwartz, E. 1993. "Space Variant Image Processing", New York University TR-1993-633
Exhibit 1017	"First Mobile Videophone Introduced," May 18, 1999
Exhibit 1018	Comparison between the current Petition and petition in IPR2021-00921

CERTIFICATION OF WORD COUNT

The undersigned certifies pursuant to 37 C.F.R. §42.24 that the foregoing Petition for *Inter Partes* Review, excluding any table of contents, mandatory notices under 37 C.F.R. §42.8, certificates of service or word count, or appendix of exhibits, contains 11,~~959~~676 words according to the word-processing program used to prepare this document (Microsoft Word).

Dated: ~~June 2, 2021~~January 10, 2022

BY: /~~s/ Adam P. Seitz~~
~~Adam P. Seitz~~Erika H. Arner/
Erika H. Arner, Reg. No. ~~52,206~~57,540

COUNSEL FOR PETITIONER

CERTIFICATE OF SERVICE ON PATENT OWNER UNDER 37 C.F.R. § 42.105

Pursuant to 37 C.F.R. §§ 42.6(e) and 42.105, the undersigned certifies that on ~~June 2, 2021~~ January 10, 2022, a ~~complete and entire~~ copy of ~~this~~the foregoing **Petition for *Inter Partes* Review** ~~including exhibits was provided via Federal Express to the~~ **of U.S. Patent** ~~Owner by serving~~**No. 8,878,949, the associated power of attorney,** and **Exhibits 1001-1018** were served by FedEx Priority Overnight on the correspondence address of record ~~for~~indicated in the ~~'949~~ Patent ~~as listed on~~ Office's public PAIR system for U.S. Patent No. 8,878,949:

Vito A. Ciaravino
Warner Norcross + Judd LLP
Intellectual Property Group
150 Ottawa Ave. NW, Suite 1500
Grand Rapids, MI 49503

Further, a courtesy copy of ~~this Petition for *Inter Partes* Review~~the foregoing was also sent via email to Patent Owner's litigation counsel:

Fred I. Williams (fwilliams@wsltrial.com)
Michael Simons (msimons@wsltrial.com)
Jonathan L. Hardt (jhardt@wsltrial.com)
Chad Ennis (cennis@wsltrial.com)
Todd E. Landis (tlandis@wsltrial.com)
John Wittenzellner (johnw@wsltrial.com)

~~BY: /s/ Adam P. Seitz~~
~~Adam P. Seitz, Reg.~~BY: /Lisa C. Hines/

Lisa C. Hines
Senior Litigation Legal Assistant

FINNEGAN, HENDERSON, FARABOW, GARRETT, & DUNNER, LLP
~~No. 52,206~~

~~COUNSEL FOR PETITIONER~~