

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

META PLATFORMS, INC.,
Petitioner,

v.

MULLEN INDUSTRIES LLC,
Patent Owner.

Case No. IPR2025-00744
U.S. Patent No. 12,019,791

PATENT OWNER'S PRELIMINARY RESPONSE

TABLE OF CONTENTS

	Page
I. Introduction.....	1
II. The Challenged Patent.....	4
III. The Proper Claim Construction.....	7
IV. The Petition.....	7
A. Fager.....	8
B. Martins.....	10
V. Board Should Deny Institution Because The Petition Does Not Establish The Required Reasonable Likelihood Of Success	10
A. Ground 1: Fager In View Of Martins.....	10
1. Ground 1 Fails To Disclose “A Head-Mounted Display Comprising A Camera” (Claim 21)	10
2. Ground 1 Fails Because The Petition Does Not Articulate A Sufficient Motivation To Combine Fager And Martins To Achieve “Utilizing Said Images To Detect A Surface Of A Physical Object Included In Said Images” (Claim 21).....	18
3. Ground 1 Fails To Disclose “A Hand-Held Controller For Providing A First Three-Dimensional Control Signal To Move Said Virtual Object And A Second Three- Dimensional Control Signal To Change The Movement Of Said Virtual Object From The Movement Provided From Said First Three-Dimensional Control Signal, Wherein Said First Three-Dimensional Control Signal Moves Said Virtual Object Three-Dimensionally Over At Least Three Axes Of Movement In A First Manner And Said Second Three- Dimensional Control Signal Moves Said Virtual Object Three-Dimensionally Over At Least Three Axes Of Movement In A Second Manner” (Claim 21)	32
a. The Virtual Racket As A Virtual Object Fails Because The Controller Does Not Move The Virtual Racket Because They Are One And The Same.....	33
b. The Virtual Ball As Virtual Object Also Fails, As The Alleged Second Control Signal Does Not Change The	

	Ball's Movement From That Provided By The First Control Signal.....	35
VI.	Conclusion	36

TABLE OF AUTHORITIES

	Page(s)
Statutes	
35 U.S.C. § 314(a)	2

Mullen Industries LLC (“Patent Owner” or “Mullen”) submits this Preliminary Response to the Petition for *Inter Partes* Review (“Petition”) of U.S. Patent No. 12,019,791 (“the ’791 Patent”) filed by Meta Platforms, Inc. (“Petitioner”).

I. Introduction

The ’791 patent improves upon prior augmented reality (“AR”) systems by providing advanced methods to detect and interact with physical objects and surfaces within a user’s environment, and by enabling sophisticated three-dimensional control of virtual objects through a handheld controller. Unlike prior art systems that required external sources of object information, the ’791 patent utilizes a head-mounted camera and detects surfaces of objects using images obtained from a camera and incorporates a handheld controller that can move the virtual objects during use.

There are two grounds presented. Ground 1 is the only ground that challenges the sole independent claim challenged in the Petition, claim 21, and all challenged dependent claims depend from claim 21. Ground 2 merely adds a reference to attempt to address a dependent claim, but does not modify the claim 21 analysis and thus Ground 2 fails for the same reasons as Ground 1. Because Ground 1 fails to disclose or suggest all claim requirements of independent claim 21, each of the two grounds fail.

The Board should deny institution under 35 U.S.C. § 314(a) for each of the following reasons.

First, Ground 1, the sole ground relied on to challenge independent claim 21, fails to disclose “a head-mounted display comprising a camera,” as required by the claim. The Petition relies on base reference Fager and its “obtaining means,” which the Petition incorrectly characterizes as a camera. In fact, the obtaining means is not a camera but rather is explicitly described in Fager as part of the display, and in particular it is the portion of the display that receives picture information for display to the user. The obtaining means obtains such picture information from the “control- and calculating unit,” which generates the visual for the user. It simply is not a camera. There is no modification proposed and thus Ground 1 fails.

Second, Ground 1 also fails because there is no motivation to incorporate Martins’s very limited image-based plane detection system, which is designed for extracting simple 2D planes from analyzing a video taken while walking around the room, into Fager’s system architecture that uses prestored 3D models from a computer database. In particular, Fager must model a wide variety of objects, from lamps-posts and trees to bushes and hills and simply adding the ability to model planes only adds expense, complexity, and wasted user time, because Fager would need to retain its existing capabilities *in addition* to the plane-modeling capabilities of Martins in order to model all the objects it needs to. Martins would thus only be

a pure addition to Fager and save nothing. It would not even have a benefit as to modeling planes, because Fager already models planes, not to mention the drawbacks of having two modeling systems, rather than the one that Fager already has.

Third, Ground 1 fails to disclose a handheld controller that provides two distinct control signals from the same controller that move a virtual object in two different manners, whereby the second signal changes the object's movement as created by the first signal. The Petition points to a virtual tennis example in which a virtual racket and virtual ball are both alleged to be the virtual object. The virtual racket is not a virtual object that is moved by the controller. Instead, it is merely a representation of the controller itself, whereas the virtual object must be interacted with by the controller. The virtual ball fails, as the alleged second signal, which is the second hit of the racket on the ball by the same user as the first hit, does not change the motion from that created by the first control signal. Rather, there is a second user who hits the ball between the first and second hits by the first user. Thus, the second hit by the first user, the alleged second signal, does not "change the movement of said virtual object from the movement provided from said first three-dimensional control signal," as required in claim 21.

Ground 2 does not change the independent claim 21 analysis, and so it fails for the same reasons Ground 1 fails.

For each of these independent reasons, the Petition should be denied.

II. The Challenged Patent

Augmented Reality (AR) is the integration of digital information into a user's environment by overlaying content onto the real-world using devices like headsets to display virtual images or information within the user's physical environment.

The '791 patent describes a head-mounted system that enables users to detect real-world surfaces using a camera integrated into a head-mounted display and a hand-held controller to interact with virtual objects within a three-dimensional environment using distinct control signals.

This integration of real-world video with virtual imagery enhances usability by allowing users to visualize and interact with virtual objects overlaid on real surfaces. For example, the patent explains that “video game indicia may be interlaced back into image data that was captured (e.g., video) and this interlaced image data may be displayed to a user (e.g., via a head-mounted display) such that the user can see an image of his/her environment and the interlaced video game indicia.” Ex. 1001, 1:32–37.

Image data is captured by a camera on the head-mounted display that “continually capture[s] images (e.g., a video camera) of a user's environment.” *Id.*, 1:53–55. The system can then render virtual elements that appear to interact with those surfaces, where “virtual game indicia may interact with objects as data is

collected, processed, and utilized to affect video game indicia in real-time.” *Id.*, 1:57–59. For instance, “if a user adds a physical object to the table like a large book, the detection system (e.g., video camera or radar and associated processing) may detect the surfaces of the book.” *Id.*, 2:3–6. Moreover, virtual objects may traverse those surfaces, for example “a virtual object may be provided that is able to drive on the top surface of a table, the side surfaces of a table, and the bottom surface of a table.” *Id.*, 7:27–29. The patent further describes flexibility in play environments: “a single game may have levels that utilize different types of physical surfaces (e.g., tabletop, wall, stairwell, floor, sink, bed, couch, TV screen).” *Id.*, 3:24–27.

The system also supports detailed user motion-based control of virtual objects. For example, the patent explains that a user may first “be perceived as holding a virtual object such as a basketball” and then “gripping the controller, shoot the virtual basketball.” *Id.*, 14:44–49. The patent similarly describes a “virtual object” such as a “virtual jar” which a user may swipe their hand through so the jar may “stick to a user’s hand as the user grips the controller.” *Id.*, 15:5–11.

The ’791 patent provides improvements over prior art by enabling real-time video capture and integration of virtual content into the user’s view of the physical environment. The patent enables real-time detection of physical surfaces and objects via a camera on a head-mounted display, allowing virtual game elements to respond dynamically to the user’s actual environment. *See Id.*, 1:53–59, 2:3–6. The patent

further enables the ability to overlay virtual objects on diverse physical surfaces such as tables, walls, or couches, and to support complex user interactions including detailed user motion-based control enhancing usability and immersion. *See Id.*, 3:24–27, 7:27–29, 14:44–49, 15:5–11. By enabling support for a broad range of physical play areas, the '791 patent advances gaming experiences beyond static or pre-modeled virtual environments. *See Id.*, 2:6–8, 3:24–27, 7:27–29. These innovations collectively provide more realistic and engaging gameplay, improving user control and interaction with virtual elements in a flexible and adaptable real-world context.

These innovations are captured in the claims as demonstrated by claim 21, which states:

21. A system comprising:

a head-mounted display comprising a camera for obtaining images in view of said head-mounted display and utilizing said images to detect a surface of a physical object included in said images, wherein said head-mounted display is operable to display a virtual object associated with an application; and

a hand-held controller for providing *a first three-dimensional control signal to move said virtual object and a second three-dimensional control signal to change*

the movement of said virtual object from the movement provided from said first three-dimensional control signal, wherein said first three-dimensional control signal moves said virtual object three-dimensionally over at least three axes of movement in a first manner and said second three-dimensional control signal moves said virtual object three-dimensionally over at least three axes of movement in a second manner.

Id., claim 21.

III. The Proper Claim Construction

Patent Owner submits that no express claim construction is necessary for the purposes of this Preliminary Response. The challenged claims should be given their plain and ordinary meaning, consistent with the specification and as understood by a person of ordinary skill in the art. Even under this standard, however, the Petition fails to demonstrate that the prior art discloses or renders obvious all of the limitations of the challenged claims.

IV. The Petition

The Petition asserts two grounds:

Ground	Claims	Basis for Challenge under §103
1	21, 23, 37, 41, 42	Fager in view of Martins
2	31	Ground 1 Prior Art + Maguire

The sole challenged independent claim is 21 and is addressed in this preliminary response, as Ground 2 does not modify the claim 21 analysis and thus fails for the same reasons as Ground 1.

A. Fager

Fager describes a device and method for detecting and modeling physical objects to influence virtual gameplay using a headset and a hand-held controller. Fager describes a transducer which can determine “its position and/or orientation relative to the real environment” by receiving “incident optical signals” from signals sources (i.e., light sources) from the environments. Ex. 1003, ¶¶ [0031], [0033]. This transducer 5 can be included in a “headset” as well as a “component” which is “intended to constitute a tool for performing the actual playing act” and is “designed as a handle of an ordinary tennis racket.” *Id.*, ¶¶ [0077], [0079]. Each transducer may “determine its position and/or orientation relative to the real environment” and sends this information to a “control- and calculating unit.” *Id.*, ¶ [0079].

Fager describes a “control- and calculating unit” that is “arranged to receive . . . model information from an external computer 18.” *Id.*, ¶ [0067]. For example, the system receives “model information from an external computer 18”, such as a “suspended lamp,” and this information is sent to the control- and calculating unit 17 which uses the information to generate the model of the suspended lamp. *Id.*, ¶¶ [0067], [0080].

Fager also describes determining the position and and/or orientation of the user's eyes and of the transferring unit for visual information relative to the environment, so "these positions and orientations may be inserted into a three-dimensional model of the environment which model is stored in computer 18." *Id.*, ¶ [0052]. The control- and calculating unit 17 uses this information along with the model from computer 18 to calculate what "visual information" should be overlaid for the user, such as the presence and behavior of virtual game elements like a "fictitious ball 29, a fictitious tennis net 24 and fictitious tennis lines 25." *Id.*, ¶¶ [0052], [0080]. The visual overlays are projected directly onto the user's retina. *Id.*, ¶ [0081] ("projecting a picture directly on the retina" when discussing the tennis controller's racket), ¶ [0092] ("projection of the picture on the retina of the eye" when discussing use with an "impermeable shield" to block external light); *see also id.*, ¶¶ [0027], [0069], [0072], [0078], [0094].

Fager describes its interactions as multiplayer, where two users participate in a shared virtual environment for a game of tennis. *See id.*, ¶¶ [0086]–[0089]. In one configuration, both users are physically present in the same space and "the both devices according to the invention share one control-and calculating unit in common." *Id.*, ¶ [0086]. Each device transmits "position- and orientation information . . . to the control- and calculating unit in common for calculating the

path of the fictitious ball, and the respective fictitious racket of both players, and the net and lines.” *Id.*, ¶ [0086].

B. Martins

Martins describes a system that surveys a room by having the user with a camera walk “around the room, capturing frames of two-dimensional (2D) video data.” Ex. 1004, ¶ [0013]. The process then recognizes “planes and reference points, such as corners, in the real-world environment” and “performs standard positions (‘pose’) estimation to determine the location of camera 12.” *Id.*, ¶¶ [0014]–[0015]. The process then “recognizes planes by identifying clusters of three or more points (e.g., pixels) of the 2D video that behave similarly during motion” of the video taken by the camera. *Id.*, ¶ [0017]. The system further allows a user to manually “populate this virtual 3D environment with virtual objects” by “select[ing] a virtual 3D object from a database” that is pre-loaded with models of objects such as “furniture or the like.” *Id.*, ¶¶ [0019]–[0020].

V. Board Should Deny Institution Because The Petition Does Not Establish The Required Reasonable Likelihood Of Success

A. Ground 1: Fager In View Of Martins

1. Ground 1 Fails To Disclose “A Head-Mounted Display Comprising A Camera” (Claim 21)

Ground 1 exclusively relies on Fager for disclosure of the claimed head-mounted camera. Pet., 13. The Petition points to the “obtaining means for obtaining

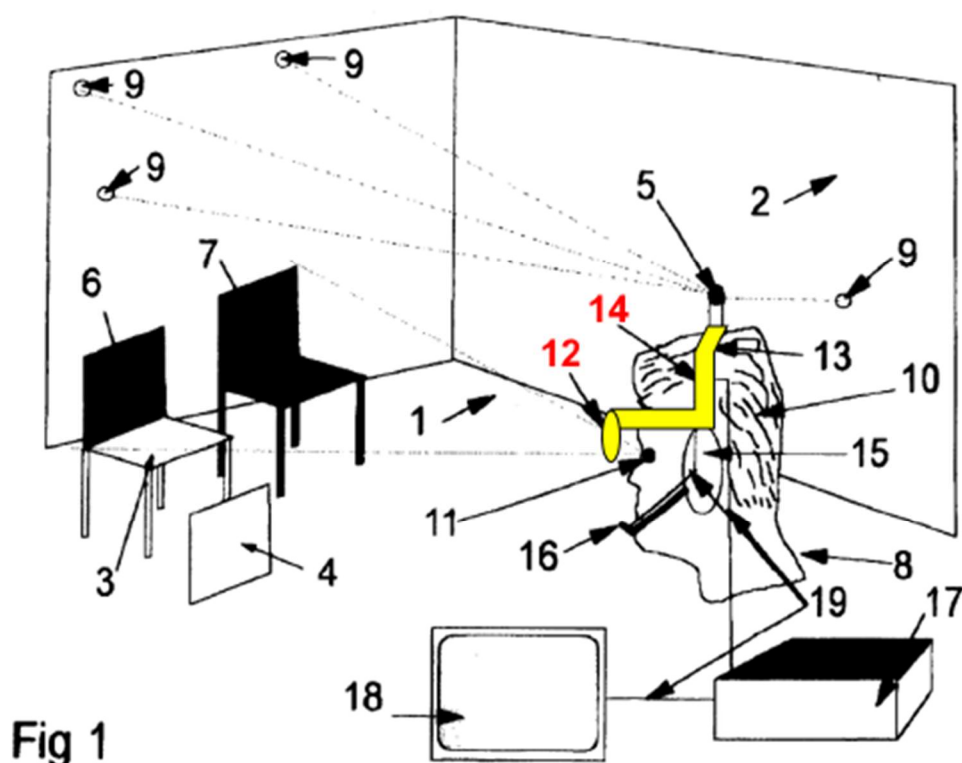
picture information” described in Fager as the claimed camera. *Id.*, 15. But the obtaining means is not a camera. It is simply a part of a display that *receives* (“obtains”) images from somewhere else for display to a user. The odd phrasing, which is an issue throughout Fager, may come from the fact that Fager was originally filed as a Swedish application. Ex. 1003, cover.

As the Petition cites and argues, the obtaining means is *part of* the “transferring means” 12. Pet., 14 (citing Ex. 1003, ¶ [0094], stating “means 12 of the transferring unit for transferring visual information comprises an obtaining means for obtaining picture information.”). As the Petition further argues, transferring means 12 is what displays images to the user, and in particular is “a transparent LCD-display.” *Id.*, 11–13 (citing Ex. 1003, ¶ [0084]). As explained more fully below, the obtaining means is the part of the LCD display or other type of display that receives the “picture information” for display to the user. It is not a camera.

Fager states that “*a unit 12 for transferring* stimuli is further included, which comprises two stimulation means, one for each eye, each consisting of a means for *picture generation directly to the retina*” and refers to “*means 12 of the transferring unit* for transferring visual information in such a form and in such a way that the means *for creating pictures on the retina*.” Ex. 1003, ¶¶ [0069],

[0072]. The Petition agrees that the transferring means 12 is a display. Pet., 11–13.

The Petition annotates unit 12 as below, which is clearly a display, and not a camera:

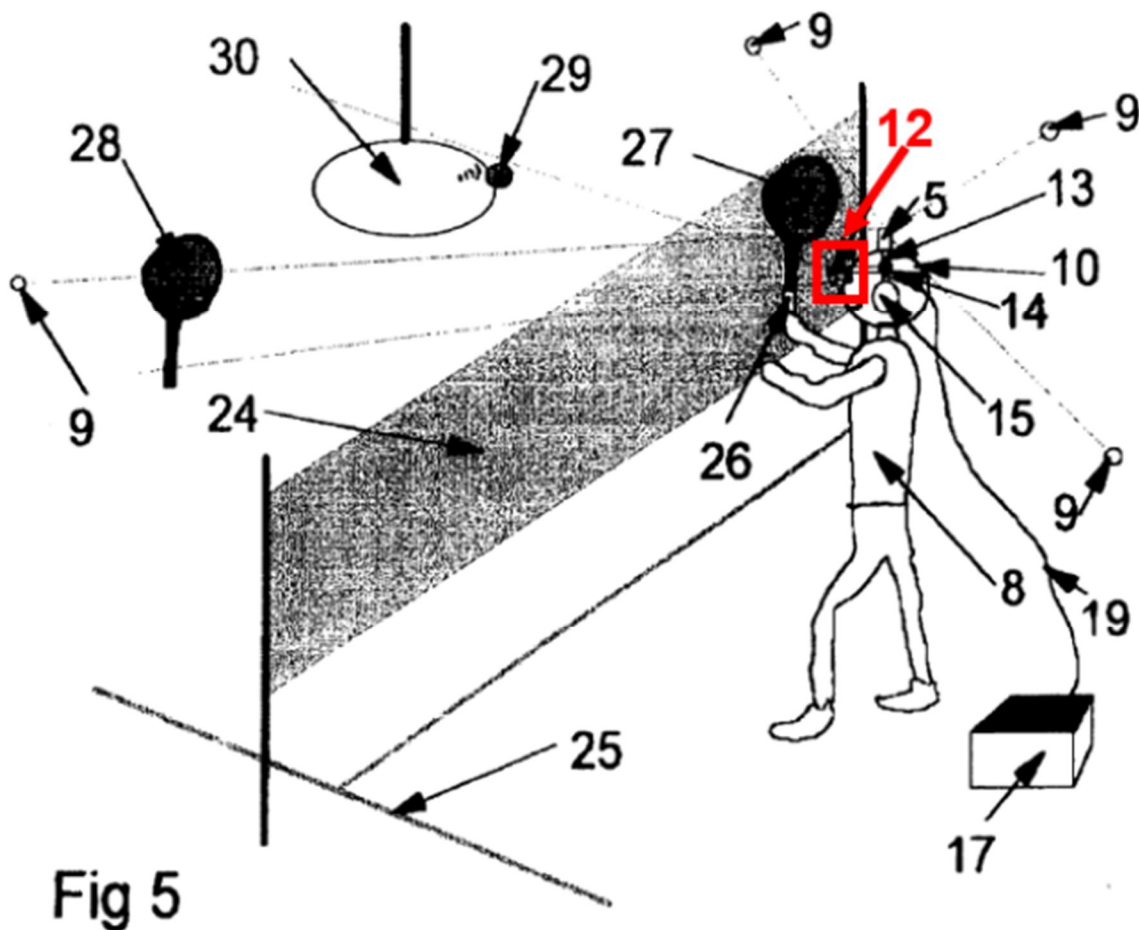


Id., 12 (annotated in original); Ex. 1003, Fig. 1.

The Petition relies on a quotation that states that “the means 12 of the transferring unit for transferring visual information **comprises** [1] an obtaining means for obtaining picture information, [2] a means which blocks the light of the environment, and a [3] transferring means for visual information for directly projection of pictures on the retina” for alleged disclosure of a camera, arguing that the “obtaining means” is a camera. Pet., 14–15 (citing Ex. 1003, ¶ [0094]). But that quotation is not describing a camera, which would be an entirely different

component from a display. Rather, it is describing the components of the means 12 itself. There is a “transferring means” for “directly projection of pictures on the retina,” which works in conjunction with “a means which blocks the light of the environment” so that the user can see the image better, and there is “an obtaining means for obtaining picture information,” so that the transferring means has something to display to the user.

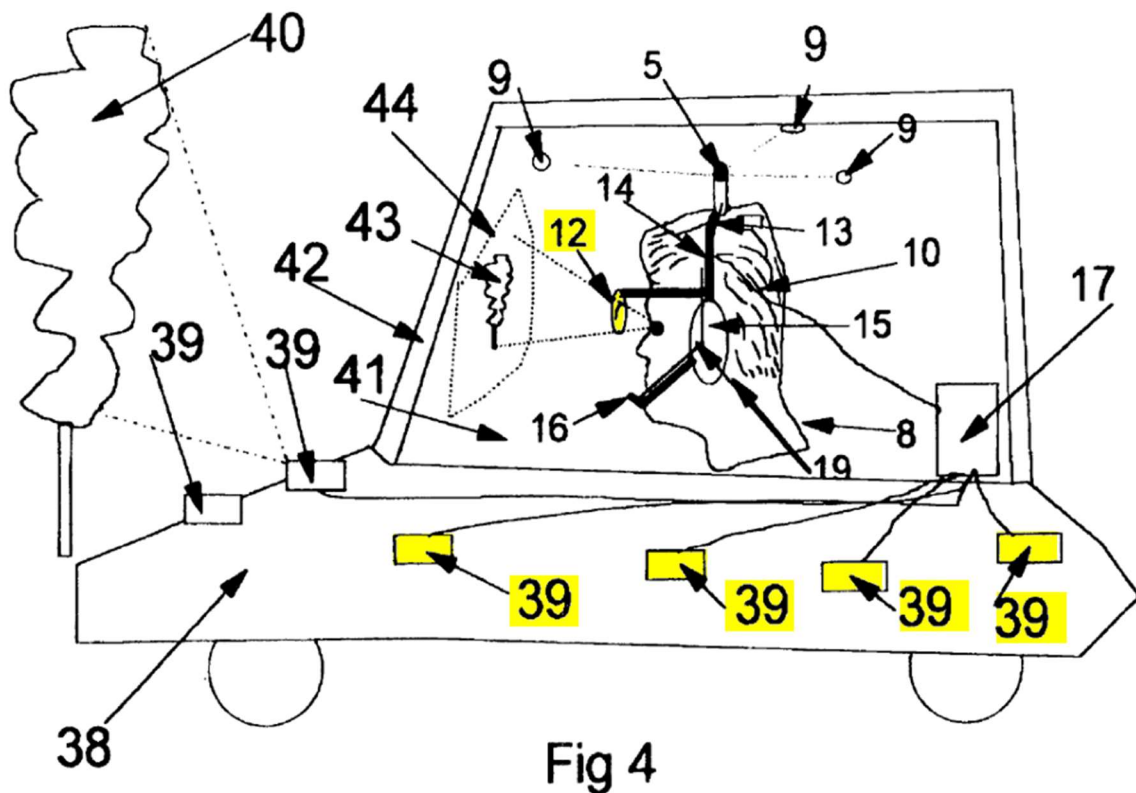
As annotated by the Petition below (and is true in all discussion, including Figure 1 above, and Figure 5 regarding the relied-upon tennis example), that obtaining means receives its “picture information” from “the control- and calculating unit 17,” which as explained more fully in the next section, runs the simulation based on “model information from an external computer 18” and creates what the user sees, and sends that picture information via “signal path[] 19” to the obtaining means and the transferring means (the transparent LCD panel) for display to the user. Ex. 1003, ¶ [0052]:



Pet., 14 (annotated in original); Ex. 1003, Fig. 5. As explained above in Section IV.A and below in Section V.A.2, unit 17 receives the picture information from “computer 18,” shown in Figure 1. As explained above, the display is a “transparent LCD,” so that the user can see the real world and the generated picture images as well. There is no basis to read the quotation above, or any part of Fager, as disclosing the addition of a *camera* to the display, *i.e.*, the transferring means 12.

The Petition also references Figure 4 below and the “*external* obtaining means 39,” which Fager states may “include IR-cameras.” Pet., 16; Ex. 1003, ¶ [0074].

This, too, makes it clear that the obtaining means in head-mounted transferring means **12** is *not* a camera—external obtaining means **39** is the camera. Figure 4 shows that external obtaining means **39** is not head worn, and is instead part of a vehicle, and it feeds in information to control- and calculating unit **17**, which, just as in the tennis example above, then feeds into the display (transferring means) **12**, which contains the obtaining means (*i.e.*, not a camera, but which can obtain information from an external camera):



Ex. 1003, Fig. 4 (annotated). Thus, even the portion the Petition cites to makes clear that transferring means **12** is just a display, and it receives its “picture information” from control- and calculating unit **17** in all embodiments. This mirrors the fact that

control- and calculating unit 17 models and outputs what the user sees in all embodiments, as described in more detail in the next section.

In addition, transferring means 12 and the control- and calculating unit 17 are not described as being any different in Figures 5 regarding the tennis embodiment or Figure 1 from what it is in Figure 4. Indeed, it is the same transferring means, as seen by the identical depiction of it in all the figures. *Id.*, ¶¶ [0069], [0077], Figs. 1, 4, 5. Thus, the “obtaining means” that the Petition points to that is part of transferring means 12 is not a camera of a head-mounted display and does not render obvious a camera of a head-mounted display.

Further, the only disclosure of information being transmitted from the user’s apparatus to anywhere has nothing to do with a video feed or any picture information or a camera whatsoever; it is merely that control-and calculating unit 17 “receive[s] position- and orientation information from the transducer 5 via, in this case, a cable 19,” as shown in Figure 1. *Id.*, ¶ [0067]; *see also id.*, ¶ [0051].

above, demonstrate that Fager uses a “video see-through head-mounted display,” as the Petition argues. *Id.*, 16. It is simply a means to block light to provide a clearer image and is part of the display, as would be common to create a better image. Nor is it evidence that there is an item providing such video specifically mounted on the head, as opposed to elsewhere in the system, as is done in Figure 4 above with the external cameras 39.

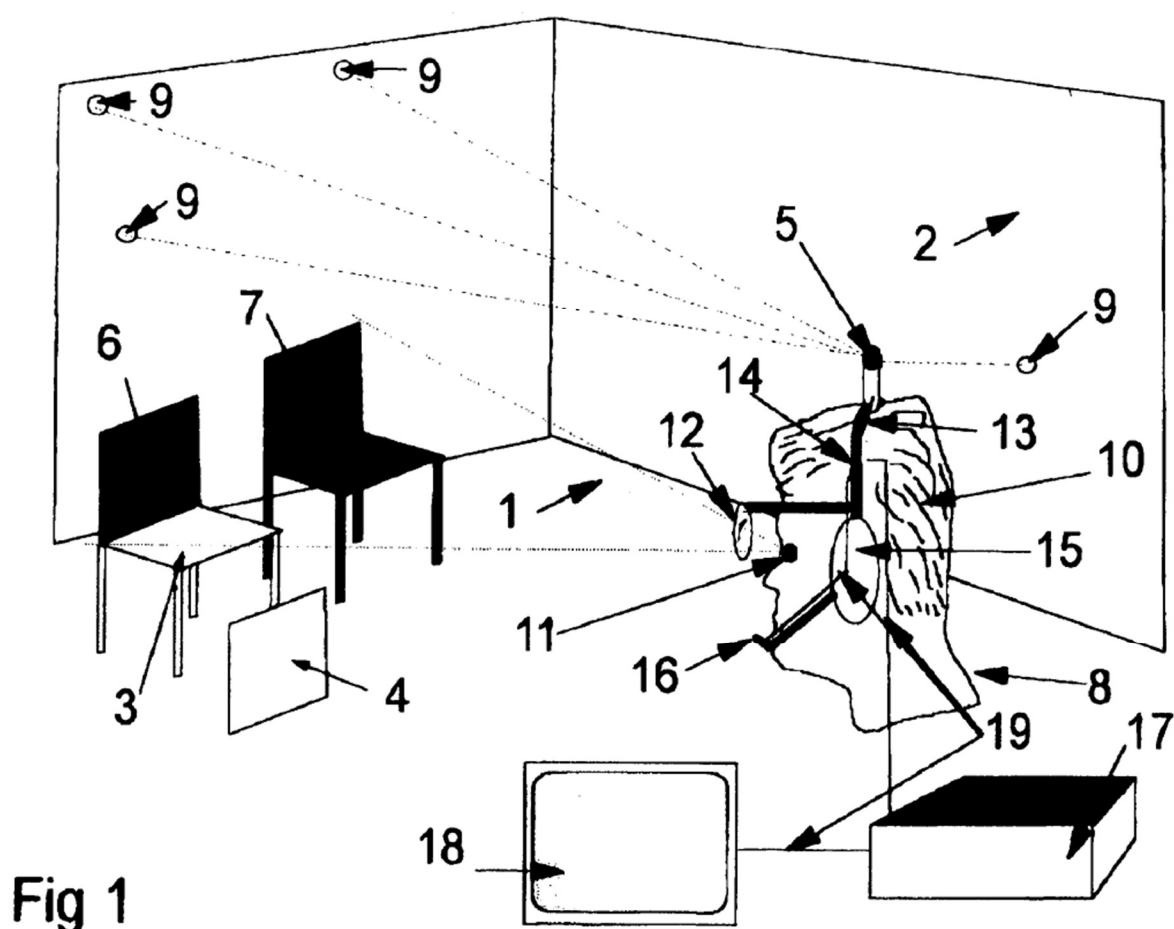
In sum, Fager does not disclose or suggest a head-mounted display comprising a camera, and thus Ground 1 fails and the Petition should be rejected.

2. Ground 1 Fails Because The Petition Does Not Articulate A Sufficient Motivation To Combine Fager And Martins To Achieve “Utilizing Said Images To Detect A Surface Of A Physical Object Included In Said Images” (Claim 21)

The Petition argues that a POSITA would have been motivated to combine Fager and Martins and in doing so the combination would disclose “utilizing said images to detect a surface of a physical object included in said images.” Pet., 18–24. That is incorrect.

The Petition argues that “Fager generally describes detecting and modeling physical objects in the real world” but concedes it does not disclose “utilizing said images” (*i.e.*, images obtained from a head-mounted camera) “to detect a surface of a physical object included in said images.” *Id.*, 18, 20.

Fager does not disclose a head-mounted camera, as explained above, but even if it does, the Petition is correct that Fager does not disclose using images (much less images from the head-mounted camera) to detect anything about an object. As explained more fully below regarding motivation to combine, Fager uses “a three-dimensional model of the environment which model is stored in a computer 18” and then fed into “control- and calculating unit 17,” which then feeds to the transferring means 12 (the alleged display) to display to the user, as depicted below:



Ex. 1003, Fig. 1; *see also id.*, ¶ [0052], Figs. 4, 5. It does not use a head-mounted camera nor does it analyze images to detect any objects; it uses pre-stored 3-D models.

The Petition relies on combination with Martins to supply the missing functionality. Pet., 20–21. But there is no motivation to combine Martins with Fager, which is substantially different from Martins and has no use for Martins’s video-based plane-detection mechanism, as it is insufficient for Fager’s needs; Martins would only add complexity while still requiring all of Fager’s functionality to remain, because Martins can exclusively only detect and model basic planes based on images, whereas Fager must model a wide variety of shapes and objects.

Martins cannot model general objects, rather, it can only model basic planar surfaces. Martins provides a process to “generate a 3D model” of a room. Ex. 1004, ¶ [0013]. The process relied upon in the Petition is that “process 16 extracts features of the room (the real-world environment, in this example) from the frames of 2D video data.” *Id.*, ¶ [0014]; Pet., 20-21. These “features” are things such as “planes and reference points, such as corners.” Ex. 1004, ¶ [0014]. Such “planes and corners” are shown in Figure 3:

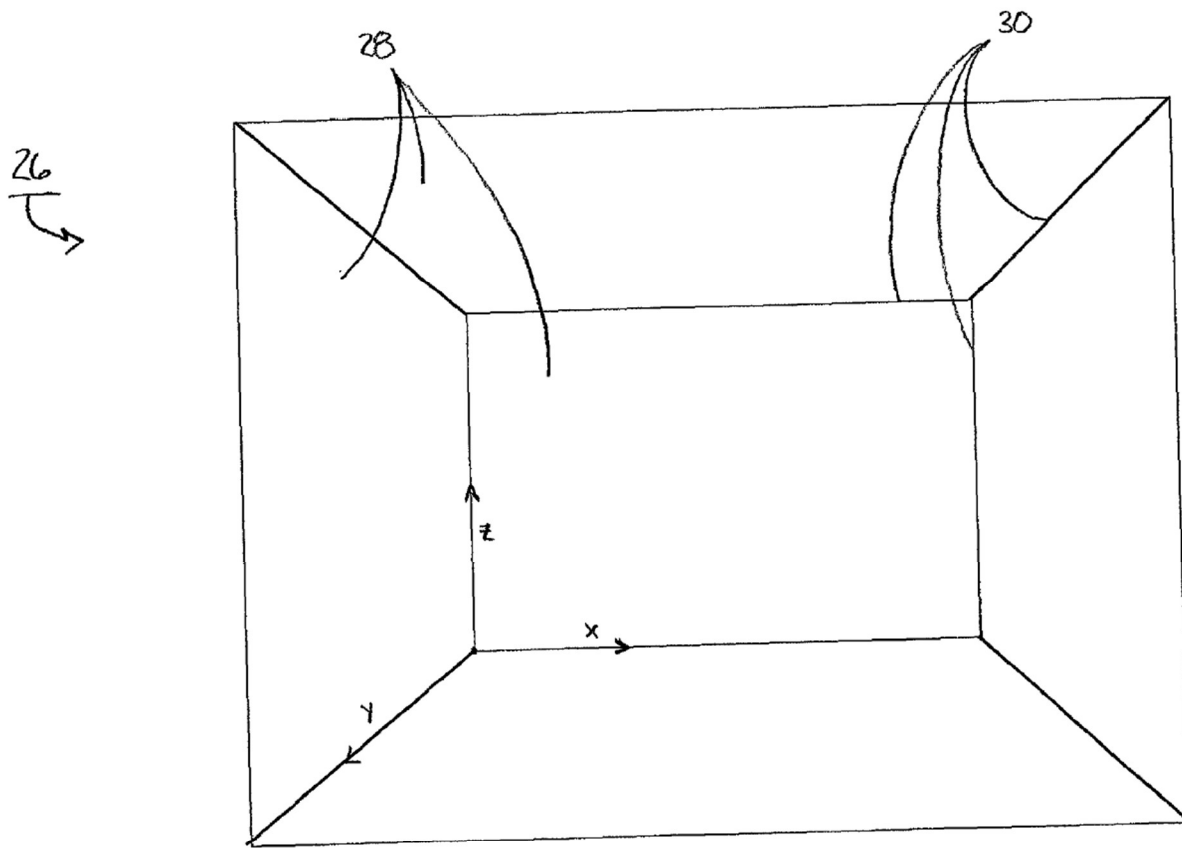


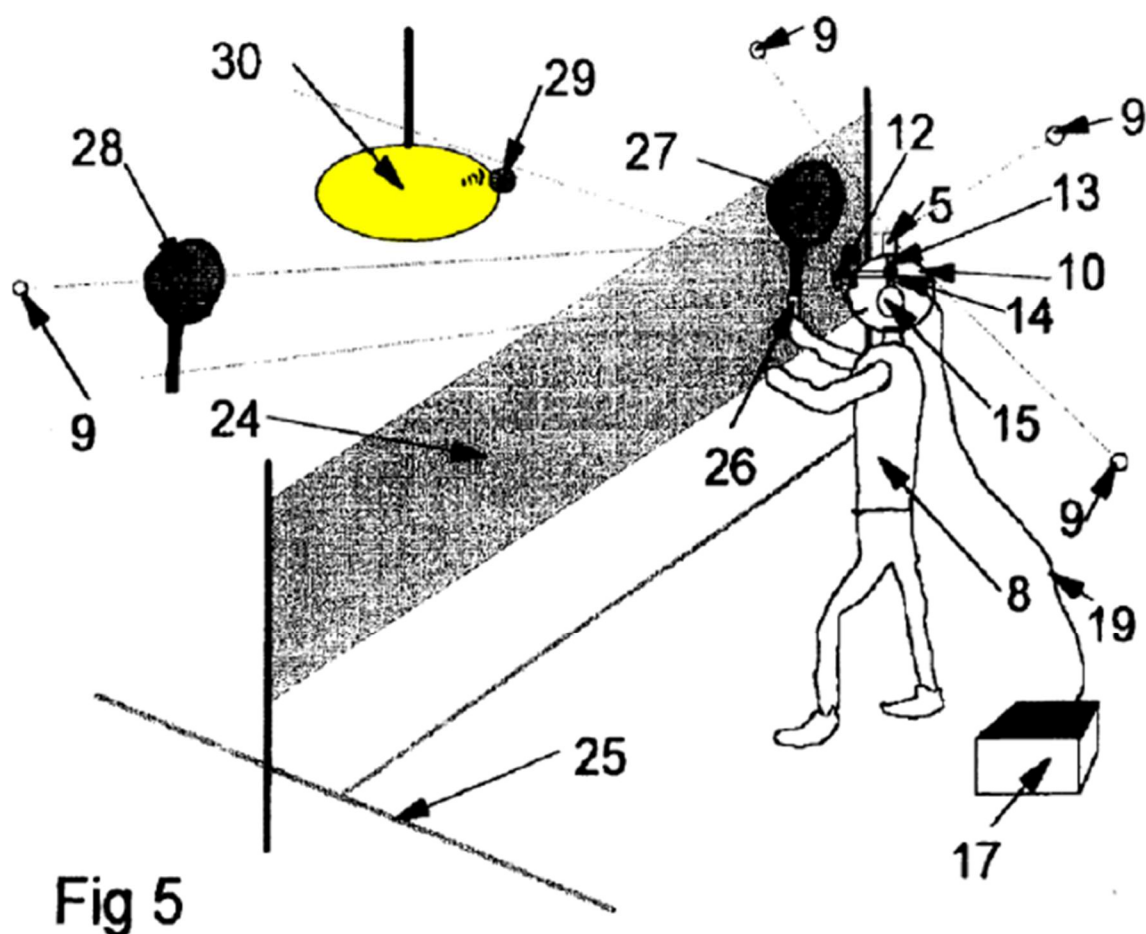
FIG. 3

Id., Fig. 3, ¶ [0014]. Planes may include all sides of the room, i.e., the walls, floor, or ceiling. *Id.*, ¶ [0017]. These features, and the knowledge of the “camera position” are what are used to create “a virtual 3D version of the room.” *Id.*, ¶ [0016]. Martins explains that “Process 16 recognizes planes by identifying clusters of three or more points (e.g., pixels) of the 2D video that behave similarly during motion of camera

12.” *Id.*, ¶ [0017]. That is, the system uses 3D analysis to find where structural elements of a plane, *i.e.*, the corners, are.

But Fager already has a process. Fager uses pre-existing 3-D models imported from an external computer for modeling everything it needs, including general objects, which Martins cannot do at all, and the surrounding environment of walls and floors. There is no reason why a POSITA would undergo the massive redesign required to change Fager from using predefined 3-D models for all objects and environmental components (walls, floors, etc.) in its simulation to Martins’s procedure that relies on a user walking around an area with a camera to record video and then performing image-processing analysis to determine planar surfaces, especially when it would only work for some small *subset* of the modeling that Fager needs to do.

The Petition largely points to Fager’s modeling of “suspended lamp 30,” which is “modelled [*sic*] in the control- and calculating unit 17” (as the Petition agrees):



Pet., 19; Ex. 1003, Fig. 5 (annotated), ¶ [0080]. As the Petition concedes, Fager does not utilize images obtained from a camera to model suspended lamp 30, or any object. Pet., 20 (“Fager does not appear to disclose that the images obtained using the camera of the head-mounted display are utilized ‘to detect a surface of a physical object included in said images,’ but this would have been obvious in view of Martins.”). Instead, the information for the lamp model comes from a database and is input into “control- and calculating unit 17,” which models the lamp. Ex. 1003, ¶ [0080]. Fager states that “[c]ontrol- and calculating unit 17” is “arranged to

transducer 5 via, in this case, a cable 19.” *Id.*, ¶ [0067]. There is no video feed or other picture information coming from the user.

The Petition argues that Martins contains disclosure for modeling objects like suspended lamp 30 in Fager and so a POSITA would use Martins’s object detection methods, relying on analyzing images from a camera, in Fager. Pet., 20–24. But Martins cannot model generic items such as a lamp based on photo or video—it is far more limited and can only recognize an ordinary plane bounded by its corners using cameras. Martins would not accomplish what Fager needs and what the Petition proposes.

Martins can only model simple planar surfaces that are bounded by corners using images, and does so by detecting the corners. The Petition points to paragraph 17 of Martins:

Process 16 **recognizes (34) planes** in the real-world environment **that are bounded by the corner**. Process 16 recognizes planes by identifying clusters of three or more points (e.g., pixels) of the 2D video that behave similarly during motion of camera 12. For example, as camera 12 moves toward a cluster of pixels, the pixels may appear to “grow”, i.e., they may appear larger because they become closer to the camera. Pixels on the same plane may appear to “grow” by about the same amount.

Examples of planes that may be identified include a floor, ceiling, and walls of a room.

Pet., 21; Ex. 1004, ¶ [0017].

The Petition cites no evidence, and there is none in Martins, of a procedure that can detect general objects, such as a lamp, based on images. *See* Ex. 1004, ¶¶ [0012]–[0019]. In fact, for shapes other than basic planes, Martins explicitly depends on a preconstructed “database” where the user themselves must manually “select” and “place” anything that is not a corner- bounded plane. *Id.*, Abstract (“retrieves a virtual object from a database comprised of pre-stored virtual objects”), ¶ [0020] (“In more detail, the user selects a virtual 3D object from a database, along with a location in the virtual 3D environment where the selected virtual 3D object is to be displayed. . . . the virtual 3D objects may be virtual 3D representations of furniture or the like. Process 16 retrieves (44) the selected virtual 3D object from the database and positions (46) it at the appropriate location.”). That is just like the approach Fager already takes: it receives “model information from an external computer 18.” Thus, the stated rationale that in Fager one could model objects such as the lamp using Martins’s method fails, as Martins cannot use images to detect such objects. Pet., 23-24.

The Petition also argues that a POSITA would modify Fager with Martins’s image analysis functionality to detect the floor, walls, and ceiling. *Id.*, 21–22. But

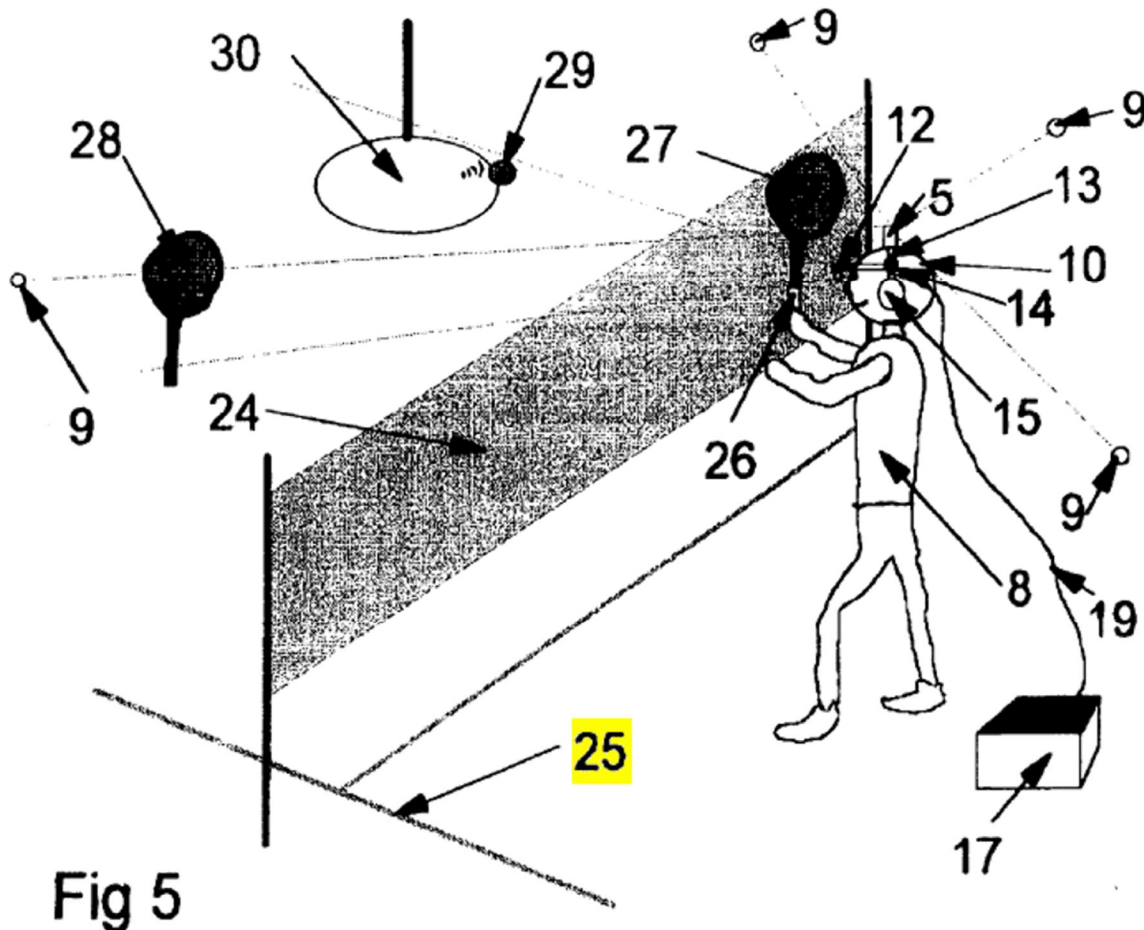
as with the lamp and all other objects, the floor, walls, and ceiling, *i.e.*, the “virtual environment,” are already modeled by “control- and calculating unit 17” and are imported to that unit from a model already existing in “computer 18”:

In this connection, it is interesting to determine the position and/or orientation of the eyes 11 and of the means of the transferring unit for transferring visual information 12, relative to the real environment 1 so these positions and orientations may be inserted into *a three-dimensional model of the environment which model is stored in a computer 18*. In this connection, the control- and calculating unit 17 included in the stimulation generating member, *from the model taken from the computer 18*, the position and orientation of the person and models input to the control- and calculating unit by programmes which may generate different visual objects, in the current example chairs and parts thereof, may calculate the information which is to be communicated to the means for transferring visual impressions 12 of the transferring unit, so that generated pictures of the fictitious phenomena may be incorporated with great accuracy in the real picture information of the room 1, the milieu being realised for the person.

Ex. 1003, ¶ [0052]. As to the floor, for example, Fager states:

Furthermore, in the control- and calculating unit 17 there is stored information about the fictitious ball 29, a fictitious tennis net 24 and *fictitious tennis lines* 25, calculation models for how a tennis ball physically acts and other rules for how the game tennis is performed in the real world.

Id., ¶ [0080]. The lines 25 denote the floor in Figure 5:



Id., Fig. 5 (annotated). The Petition’s own citation confirms this as well: “Correspondingly, for example the fictitious sound sources that are present are modelled in the *control- and calculating unit 17*. For example, sounds from how it would sound when a ball hits a racket under given conditions or when a ball hits the floor and bounces.” Pet., 20 (citing Ex. 1003, ¶ [0082]). If there were any ambiguity, Fager confirms in multiple other places that the floor, for example, is modeled in the control- and calculating unit 17, stating that “the choice of colour of the fictitious *lines 25 on the fictitious tennis court* may be changed by that control signals are transmitted to the control- and calculating unit 17” (Ex. 1003, ¶ [0093]), and that “[b]y an easy exchange of models in the control- and calculating unit, the *ground* may be altered to games, for example golf” (*Id.*, ¶ [0095]). More generally, it is clear that all aspects of the environment are modeled in the control- and calculating unit 17. *Id.*, ¶ [0067] (“In the control- and calculation unit calculations are made based on the model information obtained, and the position- and orientation information.”), ¶ [0081] (“Information about where and how the pictures of the fictitious net 24, the fictitious ball 29, the fictitious lines and the fictitious racket 28 of the simulated opponent are to be generated, is then transmitted from the control- and calculating unit 17 to the transferring unit”). This is indeed confirmed throughout the specification. *See id.*, ¶¶ [0067]–[0097] (regarding “control- and calculating unit”).

Why then would a POSITA viewing Fager undergo the large-scale redesign to incorporate image-processing-based object detection? Such a change is not a trivial matter. And, critically, Martins cannot model all the objects Fager needs to model: it can only model basic plane shapes. There is no reason provided in the Petition, nor can there be a plausible reason, why a POSITA would change Fager's database-based system into an image-analysis-based system for just a *small subset* of the objects that it needs to model. Fager would still need to use its database-driven approach for the vast majority of the objects it models. The Petition states that Fager's systems is "intended to be applied primarily 'indoors' and 'semi-indoors'," where semi-indoors are places "located outdoors in the vicinity of constructions or naturally occurring objects" such as "lamp-posts, bridges, trees, rock formations, stones, bushes, long valleys and hills, puddles, shorelines, variations in vegetation, etc., are intended." Pet., 22-23; *see also* Ex. 1003, ¶¶ [0011], [0080]. None of that can be modeled with the images taken in Martins's system.

Nor is the approach in Martins simple or easy to integrate with Fager. Martins's approach requires moving cameras around an area to determine planes based on analyzing the "relative motion" of the corners: "The user walks camera 12 around the room, capturing frames of two-dimensional (2D) video data from different positions in the room. Process 16 uses these frames to generate a 3D model

of the room. . . . It is noted that the corners themselves do not move within the room, but rather they move relative to camera 12 exclusively due to the motion of camera 12. It is this relative motion that is being tracked. . . . Process 16 recognizes planes by identifying clusters of three or more points (e.g., pixels) of the 2D video that behave similarly during motion of camera 12.” Ex. 1004, ¶¶ [0013]–[0017]. This requires an unnecessarily complicated and distinct step requiring involved user interaction in a system that is supposed to provide a real-time tennis game or similar.

It simply makes no sense to incorporate such a burdensome, complicated process into a system that already has a solution that works, especially where adding such features would not be able to fully replace the corresponding features already in place because the modeling it provides is limited to such simple shapes as planes, thus requiring both modeling systems to exist and to synchronize with the game system. It merely adds hardware, expense, time, and complication, for absolutely zero value. The Petition’s argument boils down to its statement that “images captured by that camera *could be used* in accordance with Martins’ teachings,” (Pet., 24) but ignores the significant downsides and provides no countervailing reason why it *would be used*.

Thus, Ground 1 fails because the Petition does not articulate a sufficient rationale, supported by evidence and technical reasoning, for why a POSITA would have been motivated to combine the fundamentally different systems of Fager and

Martins to achieve the claimed functionality of “utilizing images to detect a surface of a physical object included in said images.” The proposed combination is technically incompatible, would require significant redesign, and would not provide a reasonable expectation of success in achieving that functionality. Accordingly, Ground 1 fails and the Petition should be denied.

3. Ground 1 Fails To Disclose “A Hand-Held Controller For Providing A First Three-Dimensional Control Signal To Move Said Virtual Object And A Second Three-Dimensional Control Signal To Change The Movement Of Said Virtual Object From The Movement Provided From Said First Three-Dimensional Control Signal, Wherein Said First Three-Dimensional Control Signal Moves Said Virtual Object Three-Dimensionally Over At Least Three Axes Of Movement In A First Manner And Said Second Three-Dimensional Control Signal Moves Said Virtual Object Three-Dimensionally Over At Least Three Axes Of Movement In A Second Manner” (Claim 21)

The Petition relies exclusively on Fager without modification for this element.

Pet., 27–32. The claim language requires:

1. A handheld controller for providing a first and second three-dimensional control signal
2. The first three-dimensional control signal moves a virtual object
3. The second three-dimensional signal changes the movement of the virtual object from the movement provided by the first control signal

4. The first and second signals move the virtual object three-dimensionally in a first and a second manner

The Petition relies on Fager’s virtual tennis example. *Id.* The Petition relies on “component 26” in Fager as the handheld controller, and has two theories as to what the virtual object is: the virtual tennis racket and the virtual tennis ball. *Id.*, 28. Both theories fail.

a. The Virtual Racket As A Virtual Object Fails Because The Controller Does Not Move The Virtual Racket Because They Are One And The Same

The Petition’s first theory is that the virtual tennis racket is the claimed virtual object. *Id.* But that argument fails. The virtual racket simply mirrors the physical motion of the controller. *See* Ex. 1003, ¶¶ [0081]. Under this theory, every movement of the controller would inherently result in the same movement of the virtual object, from the time the tennis example is initiated until it is turned off.

The claim requires “***a hand-held controller*** for providing a first three-dimensional control signal ***to move said virtual object*** and a second three-dimensional control signal to change the movement of said virtual object.” Ex. 1001, claim 21. This language requires the controller to interact with a distinct virtual object, rather than the virtual object being a representation of the controller itself. The controller does not move the virtual object; it simply is the virtual object, under the Petition’s read. There would be no separation between input and effect. That

interpretation collapses the controller and the virtual object into the same thing, which contradicts the claim language that treats them as separate components with a defined relationship: the controller sends signals, and the virtual object responds. Thus, whether the rationale is that the virtual racket is not a virtual object, or that the controller does not move the alleged virtual object, the result is the same. The virtual racket fails to meet the claim language.

The structure of the claim requires a user holding a controller that operates on (“moves”) a virtual object. For example, the patent describes that “a user may be perceived as holding *a virtual object such as a virtual basketball*” which a user may then shoot. Ex. 1001, 14:42–61. The specification also provides other examples, such as a “*virtual object*” being a “*virtual jar*” which may stick to the user’s hand and may be released or drunk. *Id.*, 15:7–13. These examples make clear that the virtual object is distinct from the controller and reacts to it, not that the virtual object is the controller.

Because the virtual racket in Fager is not a virtual object moved by the controller but merely a mirrored image of the controller itself, it does not satisfy the claim’s requirement.

b. The Virtual Ball As Virtual Object Also Fails, As The Alleged Second Control Signal Does Not Change The Ball's Movement From That Provided By The First Control Signal

The Petition's apparent second theory, woven in with the first regarding the virtual racket, is that the virtual tennis ball is the virtual object. Pet., 30. But the claims require that two control signals *from the same hand-held controller* (and thus the same user) move the same virtual object in two ways. The claim recites "a first three-dimensional control signal to move said virtual object and a second three-dimensional control signal *to change the movement of said virtual object from the movement provided from said first three-dimensional control signal.*"

In the virtual tennis example, however, even accepting the Petition's allegations as true, the admitted player-alternating nature of the tennis example prevents this limitation from being met. The Petition's theory is that the first control signal is from a first swing of the controller by the first user and moves the virtual ball in a first movement; then, the second user, using a different controller, hits the virtual ball back; and subsequently, a second swing from the first user creates a second control signal, which moves the virtual ball in a second manner. Pet., 29–31.

However, in that example, the alleged second control signal from the second swing by the first user's controller only changes the direction of the virtual object

from that provided by the second user's controller, *not* by the first control signal from the first user's controller as claimed. That is, it does not describe two control signals from the same controller moving the virtual object in two different movements, a first and then a second that changes it from the first. That is, the alleged second control signal does not "change the movement of said virtual object from the movement provided from said first three-dimensional control signal." It changes the movement provided from a different impetus, *i.e.*, the one provided by a different user and a different controller, not from the movement provided by the same first user's first control signal, as required. Ex. 1003, ¶¶ [0086]–[0087].

The virtual ball example thus fails as well, and thus Ground 1 fails to disclose claim limitation 21[b] and the Petition should be rejected for this additional reason.

VI. Conclusion

For the above-described reasons, Patent Owner respectfully requests that the Board deny institution.

Dated: July 18, 2025

Respectfully submitted,

/Peter F. Snell/

Peter F. Snell (Reg. No. 52,235)
Brad M. Scheller (Reg. No. 61,022)
MINTZ, LEVIN, COHN, FERRIS,
GLOVSKY AND POPEO, P.C.
919 Third Avenue
New York, New York 10022
Telephone: 212-935-3000
Facsimile: 212-983-3115
E-mails: PFSnell@mintz.com;
BMScheller@mintz.com

Reza Dokhanchy (Reg. No. 62,795)
MINTZ, LEVIN, COHN, FERRIS,
GLOVSKY AND POPEO, P.C.
3580 Carmel Mountain Rd Suite 300,
San Diego, CA 92130
Telephone: 858-314-1596
Facsimile: 858-314-1501
E-mail: RDokhanchy@mintz.com

CERTIFICATE OF WORD COUNT

The undersigned certifies that the portions of the above-captioned PATENT OWNER'S PRELIMINARY RESPONSE specified in 37 C.F.R. § 42.24 have 6,855 words, in compliance with the 14,000 word limit set forth in 37 C.F.R. § 42.24(b)(1). This word count was prepared by the word-processing program used to prepare this paper (Microsoft Word).

Dated: July 18, 2025

/Peter F. Snell/

Peter F. Snell (Reg. No. 52,235)

CERTIFICATE OF SERVICE

I certify that a copy of this Patent Owner's Preliminary Response is being served by electronic mail on the following counsel of record:

Lead Counsel

Heidi L. Keefe
(Reg. No. 40,673)
COOLEY LLP
1299 Pennsylvania Ave. NW,
Suite 700
Washington, DC 20004
Telephone: (650) 843-5001
Facsimile: (650) 849-7400
hkeefe@cooley.com

Backup Counsel

Phillip E. Morton
(Reg. No. 57,835)
pmorton@cooley.com
Andrew C. Mace
(Reg. No. 63,342)
amace@cooley.com
Mark R. Weinstein (Admission *pro hac vice* to be requested)
mweinstein@cooley.com
Lowell D. Mead (Admission *pro hac vice* to be requested)
lmead@cooley.com
Patrick Lauppe (Admission *pro hac vice* to be requested)
plauppe@cooley.com
COOLEY LLP
ATTN: PATENT GROUP
1299 Pennsylvania Ave. NW,
Suite 700
Washington, DC 20004
Telephone: (650) 843-5001
Facsimile: (650) 849-7400

Dated: July 18, 2025

/Peter F. Snell/

Peter F. Snell (Reg. No. 52,235)