

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

ARASHI VISION (U.S.) LLC (D/B/A INSTA360)
Petitioner,

v.

GOPRO, INC.
Patent Owner.

Patent No. 10,958,840 B2

Inter Partes Review No. IPR2024-01164

**PETITION FOR *INTER PARTES* REVIEW
UNDER 35 U.S.C. §§ 311-319 AND 37 C.F.R. § 42.100 *et seq.***

TABLE OF CONTENTS

	Page
I. INTRODUCTION	1
II. THE '840 PATENT	2
A. Summary	2
B. Effective Filing Date	4
C. Prosecution History of the '840 Patent	5
D. Level of Ordinary Skill in the Art	5
III. CLAIM CONSTRUCTION	5
IV. PRECISE REASONS FOR RELIEF REQUESTED	6
A. Ground 1: Bell and Shi Render Obvious Claims 1-3, 5, 8-11, 14, 16, 18, and 19	8
1. Summary of Bell	8
2. Summary of Shi	9
3. Claim 1	10
4. Claim 2: “The image capture system of claim 1, wherein the position information further characterizes translational positions of the housing at the different moments within the capture duration, the trajectory further reflects the translational positions of the housing at the different moments within the capture duration, and the smoothed trajectory has smoother changes in the translational positions of the housing than the trajectory.”	48
5. Claim 3	50
6. Claim 5: “The image capture system of claim 1, wherein a longer interval of the temporal horizon enables better approximation of the intentional motion of the image capture device while causing a longer delay in the determination of the smoothed trajectory.”	52

TABLE OF CONTENTS
(continued)

	Page
7. Claim 8: “The image capture system of claim 1, wherein the determination of the smoothed trajectory based on the look-ahead of the trajectory includes weighing past positions of the housing less than future positions of the housing.”	54
8. Claim 9: “The image capture system of claim 1, wherein the determination of the smoothed trajectory based on the look-ahead of the trajectory includes a prior portion of the trajectory having less impact on the determination of the smoothed trajectory than a future portion of the trajectory.”	55
9. Claim 10: “The image capture system of claim 1, wherein the determination of the smoothed trajectory based on the look-ahead of the trajectory includes using a first amount of a prior portion of the trajectory and a second amount of a future portion of the trajectory, the first amount being smaller than the second amount.”	55
10. Claim 11: “The image capture system of claim 1, wherein the visual content are stored in a buffer during smoothed trajectory generation such that the visual content corresponding to the portion of the smoothed trajectory are stored in the buffer during the determination of the portion of the smoothed trajectory using the look-ahead of the trajectory, and the visual content stored in the buffer are used after the determination of the portion of the capture trajectory to generate the stabilized visual content.”	56
11. Claim 14: “The image capture system of claim 1, wherein the low-light high-pass parameter controls amount or strength of smoothing performed in smoothed trajectory generation and an extent of motion blur compensation applied in determination of the smoothed trajectory.”	59

TABLE OF CONTENTS
(continued)

	Page
12. Claim 16	60
13. Claim 18	61
14. Claim 19	61
B. Ground 2: Bell, Shi, and Karpenko Render Obvious Claims 4 and 20	62
1. Summary of Karpenko	62
2. Claim 4: “The image capture system of claim 3, wherein the smoothed trajectory is determined further based on iterations of smoothing including a fine scale approach to smooth on a small scale and a coarse approach to smooth on a large scale, the iterations of smoothing resulting in convergence of the smoothed trajectory.”	63
3. Claim 20	65
C. Ground 3: Bell, Shi, and Bell Paper Render Obvious Claims 6, 7, and 21	65
1. Summary of Bell Paper	66
2. Claim 6: “The image capture system of claim 1, wherein the smoothed trajectory is determined further based on minimization of a value, a score, or a metric that expresses an amount of high frequencies while respecting a crop constraint.”	66
3. Claim 7: “The image capture system of claim 6, wherein the smoothed trajectory is determined further based on iterative minimization, the iterative minimization including iterations of steps that alter latest estimate, the steps including an update step that updates a current estimate to decrease a value of an objective function and a projection step that applies one or more constraints, the current estimate being re-projected based on the one or more constraints not being satisfied.”	68
4. Claim 21	71

TABLE OF CONTENTS (continued)

	Page
D. Ground 4: Bell, Shi, and Sokeila Render Obvious Claim 12.....	71
1. Summary of Sokeila.....	71
2. Claim 12: “The image capture system of claim 11, wherein the buffer includes a circular buffer, and the visual content are stored in the circular buffer as video frames to be stabilized.”	72
E. Ground 5: Bell, Shi, and Kwatra Render Obvious Claim 13.....	73
1. Summary of Kwatra	74
2. Claim 13: “The image capture system of claim 1, wherein the weight-balance parameter controls types of motion that are minimized in smoothed trajectory generation, the types of motion including rotational velocity and rotational acceleration.”	74
F. Ground 6: Bell, Shi, and Corey Render Obvious Claim 15	76
1. Summary of Corey	76
2. Claim 15: “The image capture system of claim 1, wherein the stickiness parameter controls an extent by which a preceding portion of the trajectory impacts determination of the smoothed trajectory.”	76
G. Ground 7: Bell, Shi, Kwatra, and Corey Render Obvious Claim 17	79
1. Claim 17	79
V. THE BOARD SHOULD NOT DENY INSTITUTION UNDER <i>FINTIV</i>	80
VI. THE BOARD SHOULD NOT EXERCISE ITS DISCRETION TO DENY INSTITUTION UNDER § 325(D).....	80
VII. NOTICES AND STATEMENTS.....	80

TABLE OF AUTHORITIES

	Page(s)
Cases	
<i>In re Translogic Tech., Inc.</i> , 504 F.3d 1249 (Fed. Cir. 2007)	6
Statutes	
35 U.S.C. §§ 311-319	1
35 U.S.C. § 311(b)	6
37 C.F.R. § 42.8(b)(1).....	80
37 C.F.R. § 42.8(b)(2).....	80
37 C.F.R. § 42.8(b)(3).....	81
37 C.F.R. § 42.8(b)(4).....	81
37 C.F.R. § 42.15(a).....	82
37 C.F.R. § 42.104(a).....	82
37 C.F.R. § 42.104(b)	6
37 C.F.R. § 42.104(b)(4).....	8

Exhibit List

Exhibit Description	Exhibit #
U.S. Patent No. 10,958,840 B2 to Douady <i>et al.</i>	1001
File History of U.S. Patent No. 10,958,840 B2	1002
Expert Declaration of Aggelos Katsaggelos	1003
Steven Bell, Alejandro Troccoli, and Kari Pulli, “A Non Linear Filter for Gyroscope-Based Video Stabilization,” <i>ECCV (4)</i> , Vol. 8692 of <i>Lecture Notes in Computer Science</i> , pp. 294-308, Springer (2014).	1004
Scanned copy of Steven Bell, Alejandro Troccoli, and Kari Pulli, “A Non-Linear Filter for Gyroscope-Based Video Stabilization,” <i>ECCV (4)</i> , Vol. 8692 of <i>Lecture Notes in Computer Science</i> , pp. 294-308, Springer (2014).	1005
U.S. Patent No. 10,027,893 to Bell <i>et al.</i>	1006
U.S. Patent No. 10,462,370 to Shi <i>et al.</i>	1007
U.S. Patent No. 8,994,838 to Sokeila <i>et al.</i>	1008
U.S. Patent No. 9,071,756 to Karpenko	1009
Declaration of Sylvia D. Hall-Ellis	1010
U.S. Patent Application Publication No. US 2013/0044228 A1 to Corey <i>et al.</i>	1011
U.S. Patent No. 8,531,535 to Kwatra <i>et al.</i>	1012
Complaint in Certain Cameras, Camera Systems, and Accessories Used Therewith, Inv. No. 337-TA-1400, March 2024	1013

IPR2024-01164

Patent 10,958,840 B2

Exhibit Description	Exhibit #
'840 Patent Infringement Claim Chart in Certain Cameras, Camera Systems, and Accessories Used Therewith, Inv. No. 337-TA-1400, March 2024	1014
U.S. Patent No. 9,674,438 to Zhou <i>et al.</i>	1015

Petitioner Arashi Vision (U.S.) LLC (d/b/a Insta360) (“Petitioner” or “Insta360”) respectfully petitions for *inter partes* review of claims 1-21 of U.S. Patent No. 10,958,840 B2 (“the ’840 patent” (EX1001)) in accordance with 35 U.S.C. §§ 311-319 and 37 C.F.R. § 42.100 *et seq.*

I. INTRODUCTION

The ’840 patent relates to stabilizing video content captured by an image capture device. (EX1001, Abstract.) The ’840 patent does so by generating a smoothed trajectory, which has smoother changes (*e.g.*, in rotational or translational positions) than the actual trajectory of device. (*Id.*, 2:56-58.) The smoothed trajectory is generated by using a look-ahead, or later portion, of the trajectory of the device. (*Id.*, 8:43-46.)

Stabilized visual content is then generated by relying on the smoothed trajectory. (EX1001, Abstract.) Specifically, the smoothed trajectory is used to determine the placement of a viewing window, which is smaller than the captured image frame, within the frame. (*Id.*, 1:40-44.) Visual content within the viewing window is then cropped and used to generate the stabilized video. (*Id.*, 25:4-9.)

Determining a smoothed trajectory by using a look-ahead of the trajectory was not new at the time of the ’840 patent in 2018. The prior art commonly looked into the future, *i.e.*, considered the motion data corresponding to future video frames, when determining the smoothed trajectory.

For example, Bell discloses “us[ing] data from a look-ahead buffer to calculate a smoother path.” (EX1006, 10:15-16.) Specifically, Bell discloses determining virtual positions of the device using “the physical camera parameters from . . . a small buffer of future frames.” (EX1006, 8:65-67.)

Similarly, generating stabilized video content by using a window within the captured frame was well-known in 2018. Bell discloses a “crop polygon [that] is smaller than the input size.” (EX1006, 9:11-15; Fig. 4.) The “region within the polygon is projectively warped to create the output video.” (EX1006, 7:3-5.)

Petitioner submits that the challenged claims are unpatentable for the reasons set forth in this Petition, as supported by Dr. Katsaggelos. (EX1003.)

II. THE '840 PATENT

A. Summary

The independent claims of the '840 patent are directed to an image capture device and a method that stabilizes videos. (EX1001, Claims 1, 16.)

As shown in Figure 5C, a captured video is stabilized by determining a smoothed trajectory corresponding to the original trajectory of an image capture device. (EX1001, Abstract, 1:18-47.) The smoothed trajectory is determined based on a look ahead of the trajectory, *i.e.*, subsequent portions of the trajectory. (EX1001, 8:41-46.)

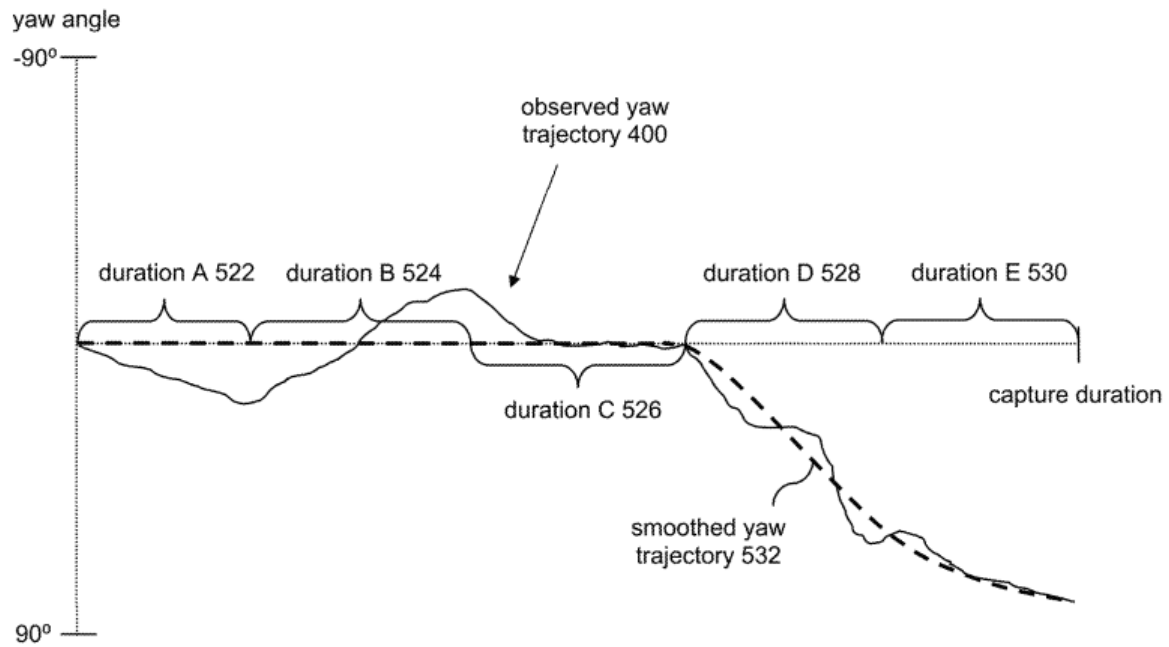


FIG. 5C

The generation of the stabilized video is based on the *viewing window*, which is smaller than the *field of view* of the captured frame¹ (*Id.*)

¹ All emphasis and annotations, such as color, highlighting, etc., are added, unless otherwise noted.

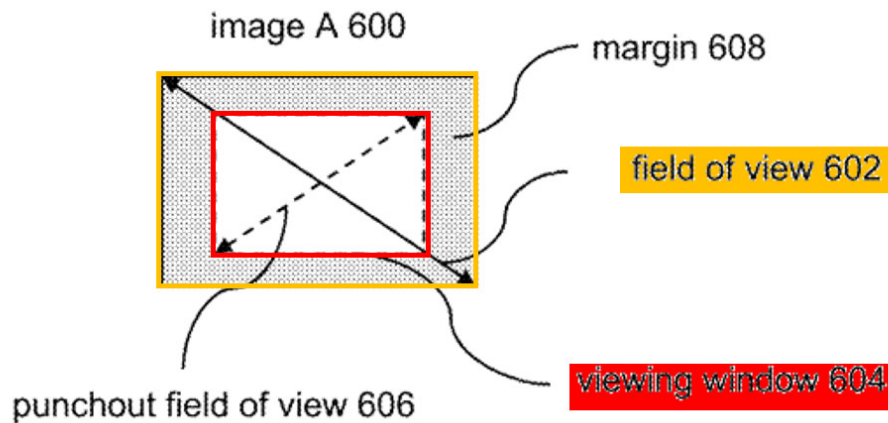
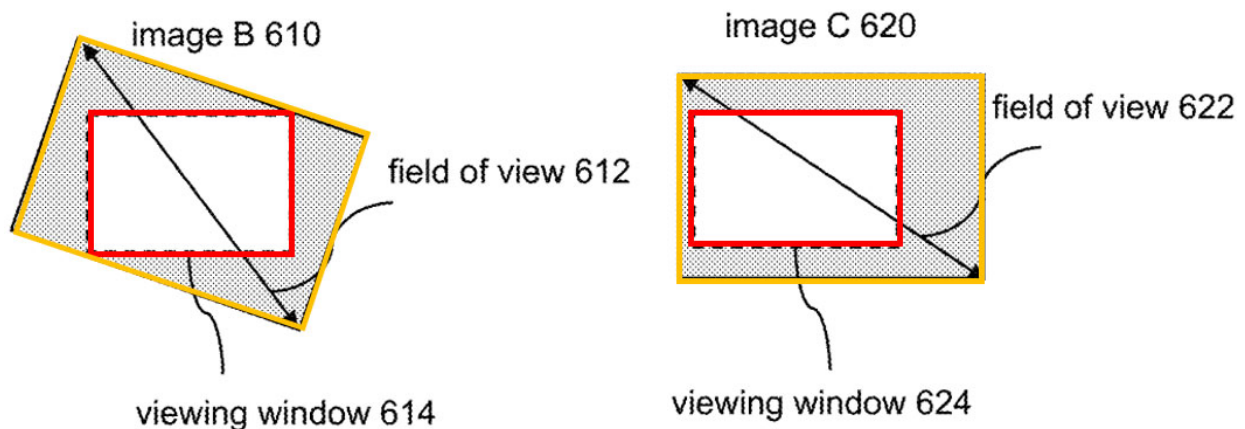


FIG. 6A

Placement of the **viewing window** with respect to the **field of view** is determined based on the smoothed trajectory. (*Id.*) For example, based on the smoothed trajectory, the viewing window may be “rotated” or “laterally moved” with respect to the field of view, as shown in Figure 6B below. (*Id.*, 20:37-43.)



B. Effective Filing Date

The '840 patent was filed on December 13, 2019, issued on March 23, 2021,

and claims priority to a provisional application filed on September 19, 2018. For purposes of this Petition only, Petitioner relies on September 19, 2018 as the earliest effective filing date. (EX1003, ¶9.)

C. Prosecution History of the '840 Patent

The Examiner issued one lone prior-art-based Office Action. (EX1002, 91-101.) The Examiner issued the Notice of Allowance after Applicant (1) amended original claim 5 to include all limitations in original claim 1, (2) cancelled original claim 1 and amended all claims dependent from it such that they depend from allowable claim 5, and (3) amended original claim 17 such that it parallels claim 5. (EX1002, 32-38, 53-66.)

D. Level of Ordinary Skill in the Art

A person of ordinary skill in the art (POSITA) in the field of the '840 patent would have a bachelor's degree in engineering or computer science, or equivalent field, and two years of experience working on image or video processing, or digital cameras. Alternatively, a POSITA would have four years of experience working on image or video processing, or digital cameras. An advanced graduate degree in the disciplines cited above may make up for the work experience indicated above. (EX1003, ¶32.)

III. CLAIM CONSTRUCTION

For purposes of this Petition only, Petitioner does not believe any terms

require a formal construction.² Each of the terms carries its “ordinary and customary meaning,” which is “the meaning that the term would have to a person of ordinary skill in the art in question” at the time of the invention.³ *In re Translogic Tech., Inc.*, 504 F.3d 1249, 1257 (Fed. Cir. 2007) (citations omitted); EX1003, ¶35.)

IV. PRECISE REASONS FOR RELIEF REQUESTED

Pursuant to 37 C.F.R. § 42.104(b), Petitioner requests cancellation of the challenged claims based on the following references (EX1003, ¶¶36-43):

Prior Art Reference	Abbreviation (Exhibit No.)
U.S. Patent No. 10,027,893	“Bell” (EX1006)
U.S. Patent No. 10,462,370	“Shi” (EX1007)

² Petitioner interprets the challenged claims based on the breadth of GoPro’s infringement allegations in the related International Trade Commission (ITC) proceeding. (EX1013; EX1014.) Petitioner reserves the right to argue in the ITC proceeding that GoPro’s construction of the terms are incorrect and/or not supported by the the ’840 patent specification.

³ An IPR petition may only rely on prior art patents or printed publications. 35 U.S.C. § 311(b). Petitioner reserves the right to raise other invalidity issues, such as invalidity under § 101 or § 112, in the related ITC proceeding.

Prior Art Reference	Abbreviation (Exhibit No.)
U.S. Patent No. 8,994,838	“Sokeila” (EX1008)
U.S. Patent No. 9,071,756	“Karpenko” (EX1009)
Steven Bell, Alejandro Troccoli, and Kari Pulli, “A Non-Linear Filter for Gyroscope-Based Video Stabilization,” <i>ECCV (4)</i> , Vol. 8692 of <i>Lecture Notes in Computer Science</i> , pp. 294-308, Springer (2014).	“Bell Paper” (EX1004)
U.S. Patent Application Publication No. US 2013/0044228 A1	“Corey” (EX1011)
U.S. Patent No. 8,531,535	“Kwatra” (EX1012)

The statutory grounds for the challenge of each claim are set forth below.

All statutory citations are AIA.

Ground	35 U.S.C. §	Claims	Prior Art Reference
1	103	1-3, 5, 8-11, 14, 16, 18, 19	Bell, Shi
2	103	4, 20	Bell, Shi, Karpenko
3	103	6, 7, 21	Bell, Shi, Bell Paper
4	103	12	Bell, Shi, Sokeila
5	103	13	Bell, Shi, Kwatra
6	103	15	Bell, Shi, Corey
7	103	17	Bell, Shi, Kwatra, Corey

Below, Petitioner discusses why the challenged claims are unpatentable

under the statutory grounds raised, including how and where the prior art satisfies each element of each challenged claim, as required by 37 C.F.R. § 42.104(b)(4).

Petitioner’s showing establishes that there is a reasonable likelihood that it will prevail on each ground of unpatentability to each challenged claim.

A. Ground 1: Bell and Shi Render Obvious Claims 1-3, 5, 8-11, 14, 16, 18, and 19

1. Summary of Bell

Bell was filed on May 10, 2016 and issued on July 17, 2018. It qualifies as prior art to the ’840 patent at least under §§ 102(a)(1) and 102(a)(2). The Examiner never found, and Applicant never cited, Bell during prosecution.

Bell discloses “systems and methods for real-time video stabilization for mobile devices based on on-board motion sensing.” (EX1006, 11:50-53.) Bell’s inventions “analyz[e] the motion sensing system measurements,” which include “angular velocity” and “device orientation.” (EX1006, 6:45-50, 7:43-48.) Those measurements describe the motion of the device. (EX1006, 7:40-48.)

For Bell, “[t]he goal of the smoothing method is to define a new virtual camera that stays static when the motion is so small that the intention is probably that the camera should remain static, and otherwise follows the original camera with smooth changes in angular velocity.” (EX1006, 8:52-56.)

Bell moves a “*crop polygon*” within the captured image, and “the region

within the *polygon* is projectively warped to create the output video.” (EX1006, 7:1-5.) The placement of the *crop polygon* is based on the virtual camera positions. (EX1006, 9:1-9.)

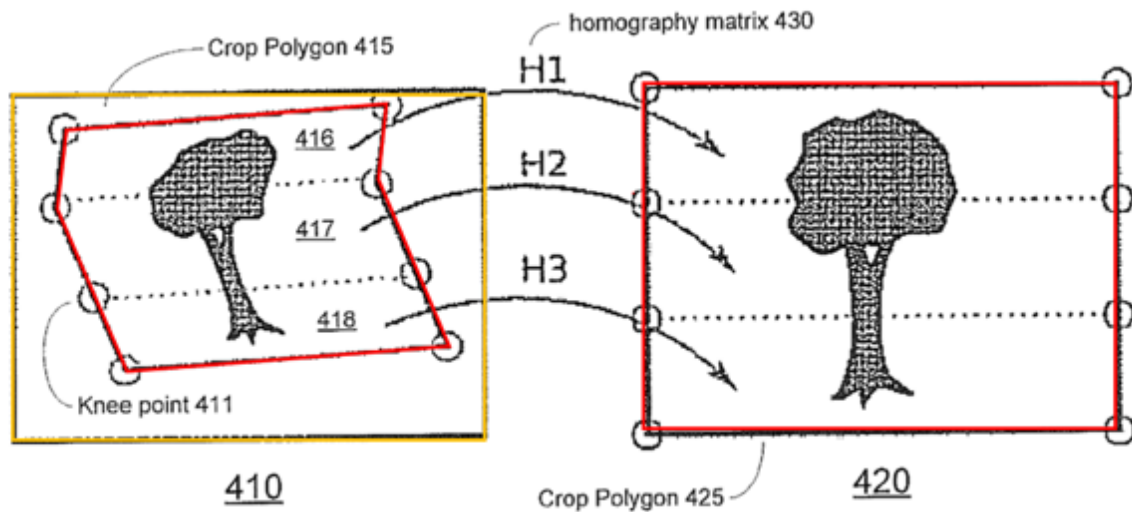


Fig. 4

2. Summary of Shi

Shi was filed on October 3, 2017 and issued on October 29, 2019. It qualifies as prior art to the '840 patent at least under § 102(a)(2). The Examiner never found, and Applicant never cited, Shi during prosecution.

Shi relates to video stabilization. (EX1007, Abstract.) Shi discloses determining the virtual camera position, which “can represent an adjusted or hypothetical pose of the recording device . . . that would stabilize the video being recorded.” (EX1007, 10:62-66.)

Shi further discloses adjusting the virtual camera position “based on

movement detected before and after a frame, based on an amount of blur in the frame, [and] based on a likelihood that panning is occurring.” (EX1007, 11:18-24.)

3. Claim 1

a. Preamble: “An image capture system that stabilizes videos, the image capture system comprising:”

To the extent the preamble is limiting, Bell discloses the preamble of claim 1.

Bell relates to “systems and methods for real-time *video stabilization* for mobile devices based on on-board motion sensing.” (EX1006, 11:50-52.) Bell also discloses that the “mobile devices” are image capture systems as they are for “*recording casual video.*” (EX1006, 1:15-16.)

b. 1[a]: “a housing of an image capture device;”

Figure 1 of Bell “illustrates an exemplary block diagram of an exemplary” mobile device, or “electronic system 100[.]” (EX1006, 3:54-57; 4:5.)

A POSITA would have understood that a mobile device has a *housing* that encloses the components shown in Figure 1. (EX1006, 3:58-65; EX1003, ¶54.)

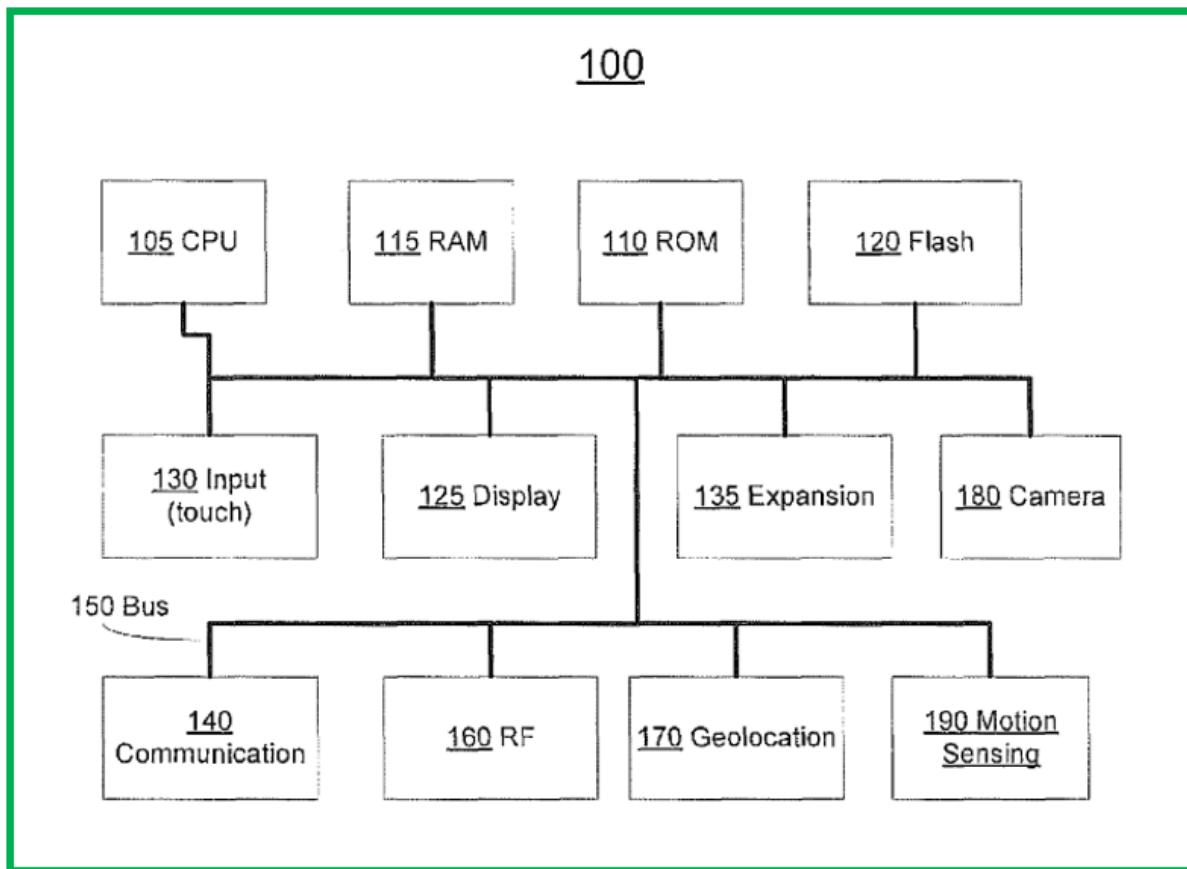


Fig. 1

- c. **1[b]: “an optical element carried by the housing and configured to guide light within a field of view to an image sensor;”**

Optical element. Per the '840 patent, “the optical element” may include “one or more of lens, mirror, prism, and/or other optical elements.” (EX1001, 6:32-34.)

Bell discloses that electronic system 100, the “image capture device,” includes “camera 180,” which includes “fixed and/or auto-focus optics, wafer-scale lenses, liquid lenses.” (EX1006, 6:1-5.) Bell therefore discloses optical elements.

(EX1003, ¶¶55-58.)

As those optical elements are included in electronic system 100, a POSITA would have understood that the optical elements are carried by the *housing* of electronic system 100. (EX1003, ¶58.)

Image sensor. Bell discloses that the camera in electronic system 100 includes “back-side illuminated CMOS image sensor(s).” (EX1006, 6:5-8.)

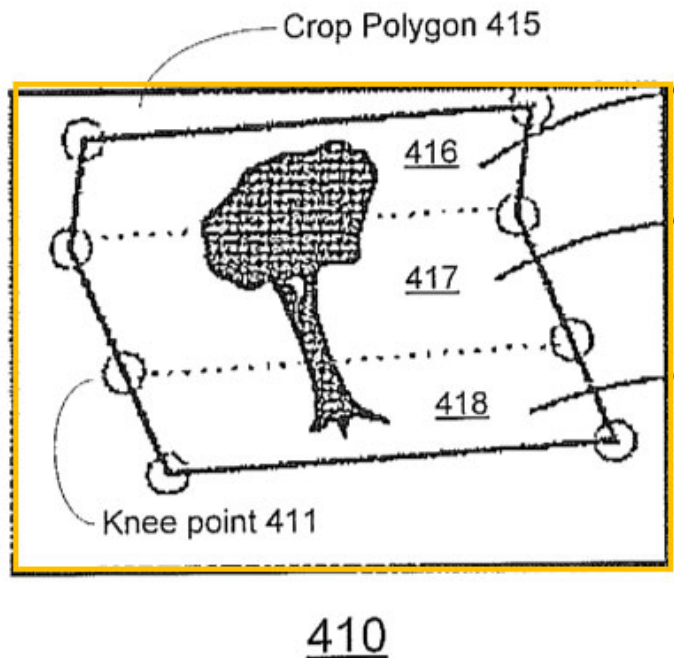
The optical elements, including “fixed and/or auto-focus optics, wafer-scale lenses, [and] liquid lenses,” are configured to guide light within a field of view captured by them to the image sensor. (EX1003, ¶60.)

- d. 1[c][i]: “the image sensor carried by the housing and configured to generate a visual output signal conveying visual information based on light that becomes incident thereon during a capture duration,”

Image sensor. A POSITA would have understood that the image sensor is carried by the *housing* of electronic system 100. (EX1003, ¶61.)

An image sensor is configured to generate visual output signals based on light that strikes upon the image sensor. (EX1003, ¶62.)

Visual information. The images in the video captured by electronic system 100 are the “visual information.” (EX1003, ¶63.) For example, Bell discloses that *image 410* in Figure 4 “illustrates a frame as captured by an image sensor.” (EX1006, 10:38-42.)



The *image 410* constitutes “visual information.” (EX1003, ¶¶65.) A POSITA would have understood that *image 410* is conveyed by the visual output signals generated by the image sensor. (EX1003, ¶¶63-66.)

Bell also discloses “a series of *frame images*” captured by electronic system 100, as Figure 2 shows. (EX1006, 7:9-11.) Those *frame images* also constitute “visual information” and are conveyed by the visual output signals generated by the image sensor. (EX1003, ¶¶63-66.)

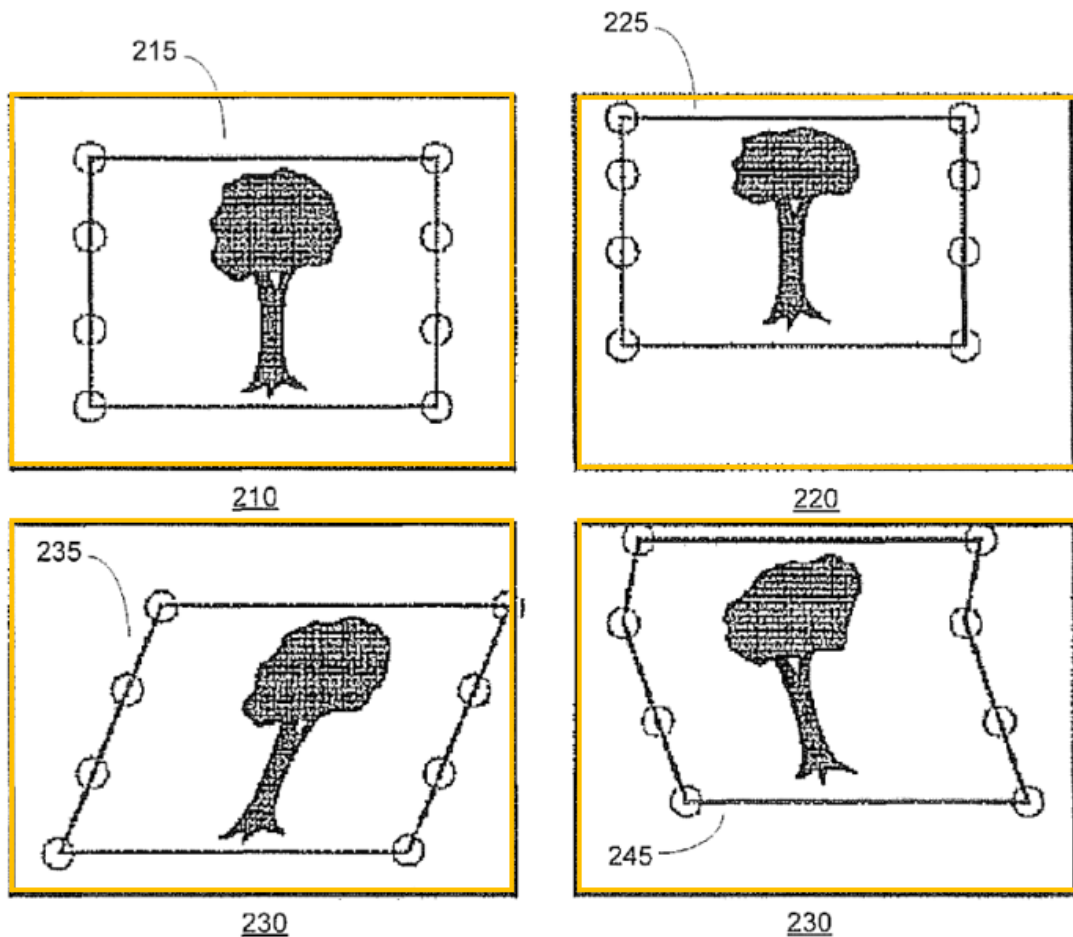


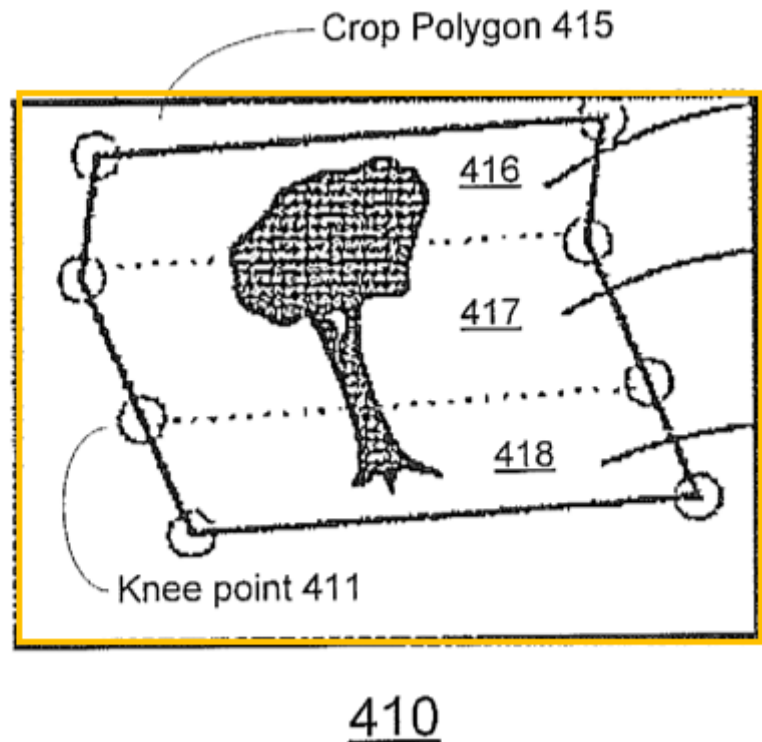
Fig. 2

Capture duration. Electronic system 100 captures and records videos. (EX1006, 1:15-16.) As Bell discloses that electronic system 100 performs “real-time video stabilization” and “correct[s] the video before it is sent to the hardware video encoder” (EX1006, 1:33-43, 6:27-30), a POSITA would have understood that a video is captured by electronic system 100. (EX1003, ¶¶67-68.) The duration of the captured video is the “capture duration.”

Therefore, Bell discloses that *images, such as image 410* in Figure 4, are conveyed by visual output signals generated by the image sensor, based on light that becomes incident on the image sensor, during the capture of the video by electronic system 100. (EX1003, ¶69.)

e. 1[c][ii]: “the visual information defining visual content having the field of view,”

The *images* captured by electronic system 100 are the “visual information.” (*Supra* § IV.A.3.d.) The captured images also are called *input frames*. (EX1006, 10:51-53, 10:64-67.) As the images contain “scene content,” or visual content (EX1006, Abstract), the captured images define the visual content. (EX1003, ¶70.) For example, *image 410* contains visual content that includes a tree. (EX1003, ¶70.)



The *entire field within image 410* is the “field of view.” (EX1003, ¶71.)

- f. 1[c][iii]: “the field of view being larger than a size of a viewing window for generating stabilized visual content;”

As the claim recites “the field of view being larger than a size of a viewing window,” the claim requires a “viewing window” that is smaller than the field of view. (EX1003, ¶72.)

Bell discloses a “viewing window”: a “*crop polygon* [that] is smaller than the *input size*, which leaves a small amount of ‘padding’ between the polygon borders and the input frame edges.” (EX1006, 9:12-15.)

As Figure 4 shows, the *field of view of image 410* is large than the size of

crop polygon 415. (EX1003, ¶74.)

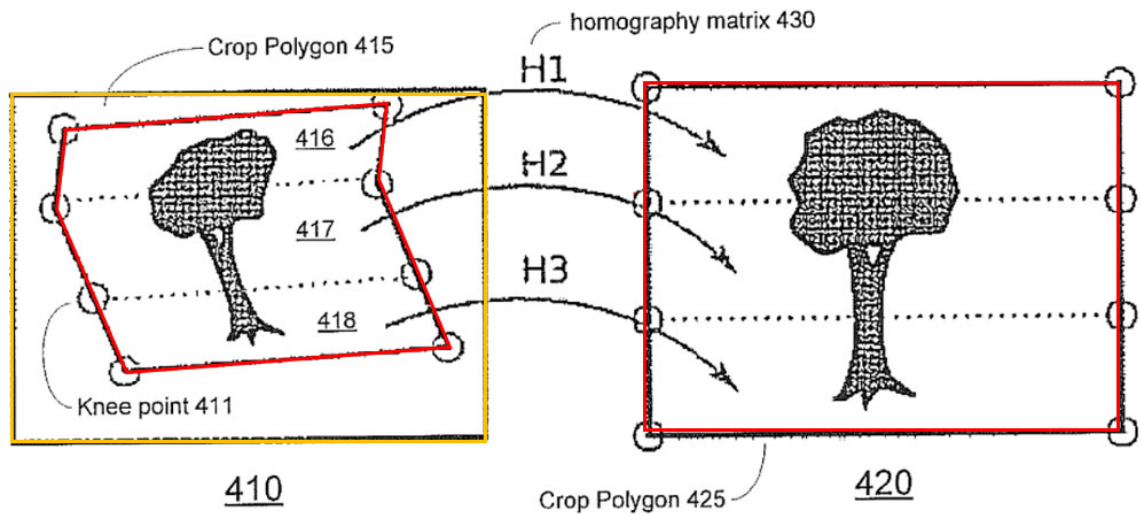


Fig. 4

Bell also discloses that *crop polygon* is “for generating stabilized visual content.” (EX1003, ¶75.) According to Bell, image 420, shown to the right of image 410, “illustrates an exemplary corrected image made from the image 410 video capture information.” (EX1006, 10:48-50.) In other words, image 420 is stabilized image, or stabilized visual content. (EX1003, ¶76.) Stabilized image 420, or “the output,” is “synthesized by projectively warping *the crop polygon* from the video input to the virtual camera.” (EX1006, 10:51-53.)

- g. **1[d][i]: “a position sensor carried by the housing and configured to generate a position output signal based on positions of the housing during the capture duration,”**

Position sensor carried by the housing. According to the '840 patent, “the position sensor may include one or more of a gyroscope, an accelerometer, and/or

an inertial measurement unit.” (EX1001, 2:20-22.)

Bell discloses that “[e]lectronic system 100 further comprises an *on-board motion sensing system 190*.” (EX1006, 6:9-10.) “*On-board motion sensing system 190* may comprise any suitable linear and/or angular motion sensing technology, including, for example, accelerometer(s), gyroscope(s), compass(es), barometer(s) and the like.” (EX1006, 6:10-14.) *Motion sensing system 190* therefore includes the “position sensor.”

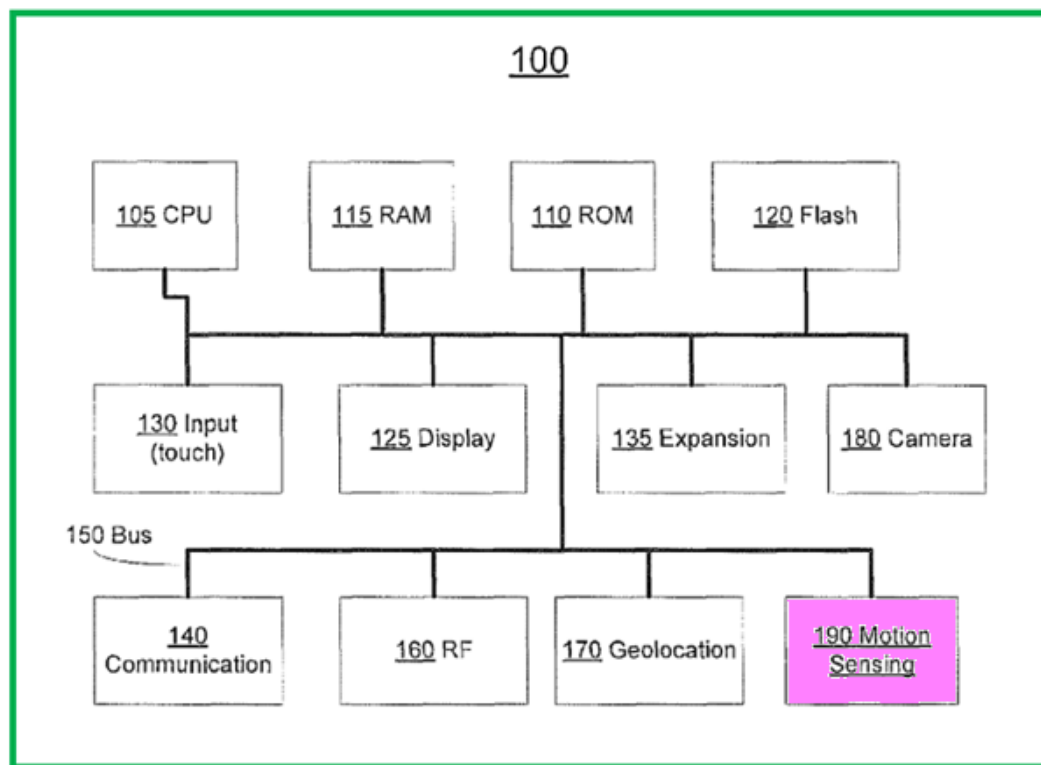


Fig. 1

A POSITA would have understood that *motion sensing system 190* is carried by the *housing* of the device. (EX1003, ¶80.)

Positions of the housing during the capture duration. The time duration of the captured video in Bell is the “capture duration.” (*Supra* § IV.A.3.d.) Bell discloses that the “motion sensing system provides a series of angular velocity measurements with timestamps,” which describe “the device orientation at any time t.” (EX1006, 7:42-45.) The device’s orientation describes the position of the device (electronic system 100). (EX1003, ¶81.)

Because the device includes the housing, the position (*i.e.*, orientation) of the device reflects the position information of the housing. (EX1003, ¶82.) As electronic system 100 also includes a camera (EX1006, 6:1-2), the position of the device also reflects the position of the included camera. (EX1003, ¶82.) Bell therefore discloses measuring the positions of the housing during the duration of the captured video. (EX1003, ¶82.)

Position output signal. The position information is conveyed by position output signals generated by *motion sensing system 190*. (EX1003, ¶83.)

- h. 1[d][ii]: “the position output signal conveying position information that characterizes rotational positions of the housing at different moments within the capture duration; and”**

Bell discloses position information conveyed by position output signals generated by motion sensing system 190. (*Supra* § IV.A.3.g.)

Bell further discloses that the position information characterizes rotational

positions of the housing. (EX1003, ¶85.) Bell discloses “model[ing] the camera motion as a *rotation* in a global coordinate frame.” (EX1006, 7:40-42.) Bell does that by using “[t]he motion sensing system [to] provide[] a series of angular velocity measurements with timestamps, which are interpolated and integrated to produce a function that describes the device *orientation* at any time t , expressed as *the quaternion $q_g(t)$* .” (EX1006, 7:42-46.) A POSITA would have understood that the device orientation characterizes the rotational positions of electronic system 100, and therefore, its housing. (EX1003, ¶86.) Additionally, quaternion is a mathematical notation for rotational positions. (EX1003, ¶85.)

Bell discloses that the measured device orientations correspond to different moments within the duration of the captured video. (EX1003, ¶¶87-88.) Bell discloses that the “motion sensing system measurements [occur] in real time,” “as part of [the] video capture.” (EX1006, 6:47-50, 6:24-26.) Bell also discloses that the device orientation can be expressed as quaternion $q_g(t)$. (EX1006, 7:42-46.) t refers to a moment within the duration of the captured video. (EX1003, ¶88.)

i. 1[e]: “one or more physical processors configured by machine-readable instructions to:”

Bell discloses that the electronic system includes “a *processor* coupled to the image sensor and to the motion sensing system.” (EX1006, 1:63-67, claim 6.) Bell also discloses that “[e]lectronic system 100 includes . . . *a central processor*

105,” which “may comprise a single processor or multiple processors, e.g., a multi-core processor, or multiple separate processors.” (EX1006, 3:66-4:5.)

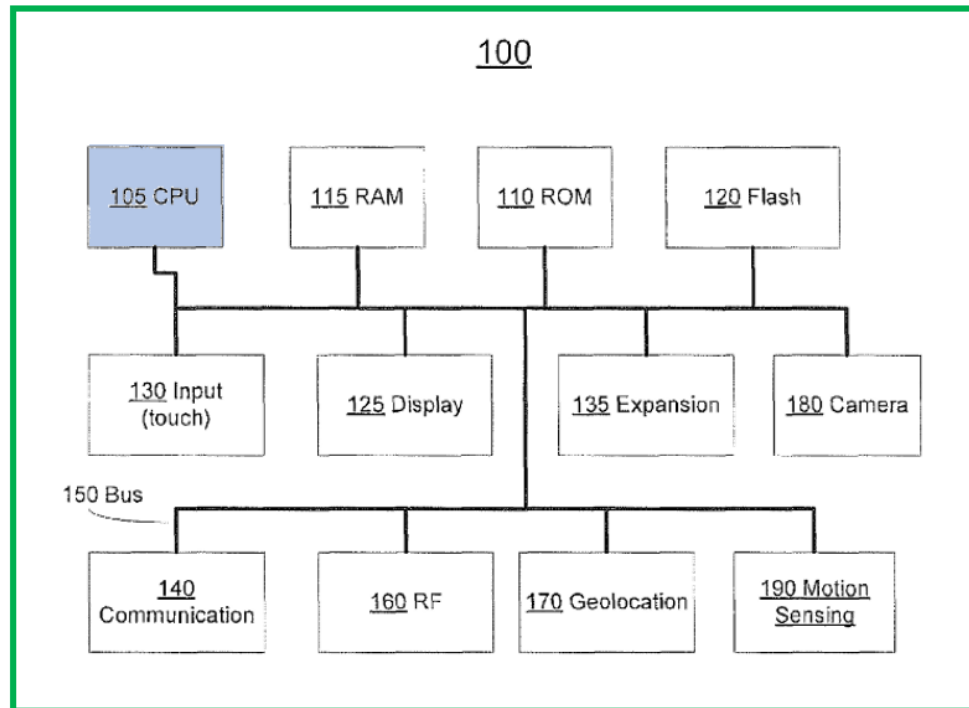


Fig. 1

Bell further discloses “a non-transitory computer readable medium having instructions stored thereon that . . . cause [the] electronic system to perform operations.” (EX1006, 13:21-24.) Per Bell, the non-transitory computer readable medium can be “a volatile memory 115,” “a non-volatile memory 110,” or “a changeable, non-volatile memory 120.” (EX1006, 4:5-12.) All of them “stor[e] information and instructions for the central processor.” (EX1006, 4:12-13.) Bell therefore discloses processors configured by machine-readable instructions to perform operations. (EX1003, ¶90.)

- j. **1[f][i]: “determine a trajectory of the housing during the capture duration based on the position information, the trajectory reflecting the rotational positions of the housing at the different moments within the capture duration,”**

Bell discloses this limitation under GoPro’s infringement theory. In the related ITC action, GoPro alleges that the accused products “form[] a moment-by-moment listing of the camera position based on the six-axis gyroscope generated position output signals to create *an observed trajectory* of the housing during the capture duration.” (EX1014, at 13.)

Determine a trajectory of the housing during the capture duration based on the position information. Bell discloses “a function that describes the device orientation at any time t , expressed as the quaternion $q_g(t)$.” (EX1006, 7:44-46.) Bell therefore discloses a moment-by-moment listing of the device’s position information, and therefore determines a trajectory of the device and its housing based on the position information. (EX1003, ¶93.)

The trajectory reflecting the rotational positions of the housing at the different moments within the capture duration. The time duration of the captured video is the “capture duration” (*supra* § IV.A.3.d); the device orientations measured by motion sensing system 190 are the “position information,” which characterizes the rotational positions of the device at different moments within the duration of the captured video (*supra* § IV.A.3.h). Bell’s trajectory (moment-by-

moment listing of position information) therefore reflects the rotational positions of the housing at the different moments within the duration of the captured video. (EX1003, ¶94.)

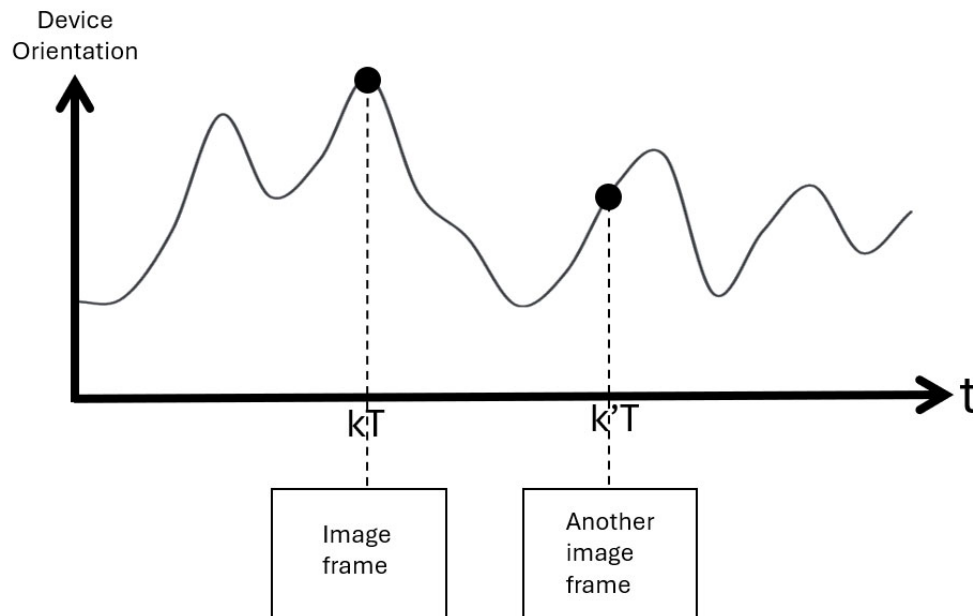
To the extent that the trajectory requires a continuous path, it would have been obvious to connect the device orientations (measured as quaternion $q_g(t)$) across the entire duration of the captured video to form the path. (EX1003, ¶95.)

- k. 1[f][ii]: “the trajectory including a first portion corresponding to a first moment within the capture duration and a second portion corresponding to a second moment subsequent to the first moment within the capture duration;”**

Bell discloses that within the duration of the captured video, image frames are captured. (EX1006, 6:28-30, 11:18-26.) Bell also discloses “access[ing] a first image frame from a camera at a first time” and “access[ing] a second image frame from the camera at a subsequent time.” (EX1006, 2:1-3.) The images frames therefore correspond to different moments within the duration of the captured video. (EX1003, ¶96.)

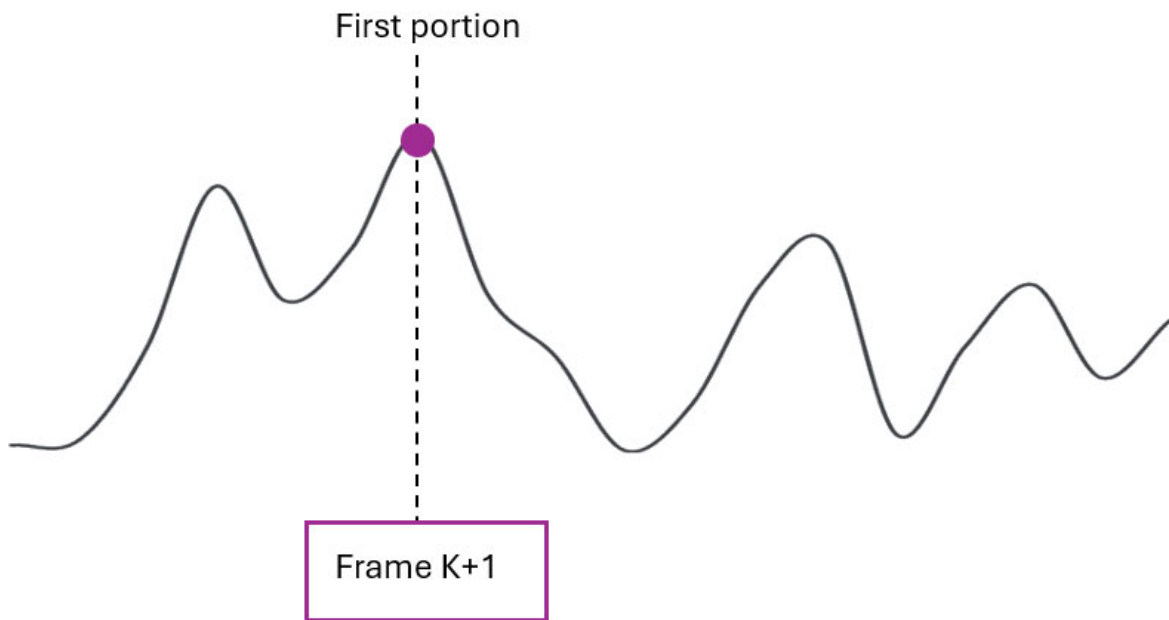
Moreover, as discussed for element 1[d][ii], the measured device orientations correspond to different moments within the duration of the captured video. (*Supra* § IV.A.3.h.) Therefore, a POSITA would have understood that there is a corresponding relationship among the image frames of the captured video, the measured device orientations, and the moments within the duration of

the captured video. (EX1003, ¶97.) As the below illustration shows, the image frames correspond to the device orientations, as both correspond to the moments within the duration of the captured video. (EX1003, ¶¶97-98.)



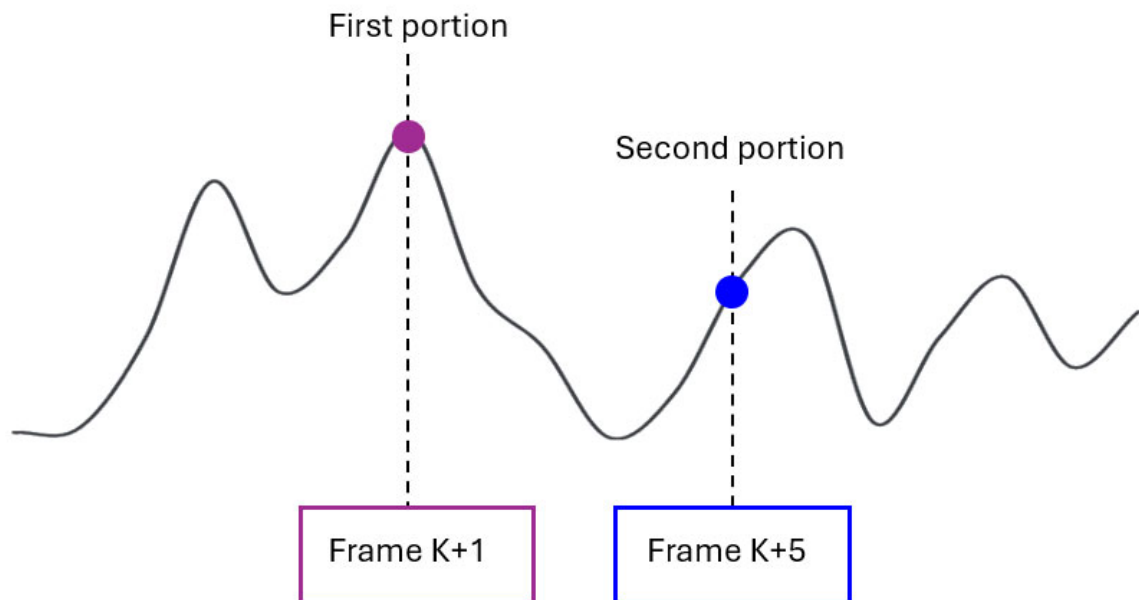
A first portion corresponding to a first moment. Bell discloses that the “physical . . . camera orientation[] at frame k ” is represented with the quaternion $p(k)$. (EX1006, 8:57-58.) A POSITA would have understood that physical camera orientation at the subsequent frame, frame $k + 1$, is represented by quaternion $p(k + 1)$. (EX1003, ¶100.) Bell therefore discloses position of electronic system 100 at a moment within the duration of the captured video corresponding to captured frame $k + 1$. (EX1003, ¶100.) That moment can be a first moment. (EX1003, ¶100.)

As the trajectory is formed by electronic system 100's orientations, or position information, across the duration of the captured video (*supra* § IV.A.3.j), the trajectory includes a first portion corresponding to position of electronic system 100 at a first moment, corresponding to *frame k + 1*, as the below illustration shows. (EX1003, ¶101.)



A *second portion* corresponding to *a second moment* subsequent to the first moment. Bell also discloses considering “future frames, e.g., five or less” from the current frame *k*. (EX1006, 8:63-67.) A future frame such as *frame k + 5* corresponds to a second moment after the first moment corresponding to *frame k + 1*. (EX1003, ¶102.) As the “motion sensing system” measures “the device

orientation at any time t ,” a POSITA would have understood that the orientation of electronic system 100 at *frame $k + 5$* (i.e., $p(k + 5)$) also is measured. (EX1003, ¶103.) Therefore, the trajectory includes a second portion corresponding to position of electronic system 100 at a second moment, corresponding to *frame $k + 5$* , as the below illustration shows. (EX1003, ¶103.)



- I. 1[g][i]: “determine a smoothed trajectory of the housing based on a look-ahead of the trajectory and one or more of a weight-balance parameter, a low-light high-pass parameter, and/or a stickiness parameter,”

Determine a smoothed trajectory based on a look-ahead of the trajectory.

Bell discloses “us[ing] data from *a look-ahead buffer* to calculate a *smoother*

path.” (EX1006, 10:10-16.) “The goal of the smoothing method is to define a new virtual camera” and calculate the movement of the virtual camera. (EX1006, 8:52-56, claim 1.) The movement of the virtual camera is the smoothed trajectory. (EX1003, ¶105.)

Bell further discloses that its “smoothing algorithm” uses “physical camera parameters from . . . a small buffer of *future frames*, e.g., five or less.” (EX1006, 8:63-67.) (*Infra* § IV.A.3.k.) As discussed for element 1[f][i], physical camera parameters include device orientations, which form the trajectory. (*Supra* § IV.A.3.j.)

Bell therefore discloses determining the movement of the virtual device (*i.e.*, smoothed trajectory) based on future position information of the device (*i.e.*, a look-ahead of the trajectory). (EX1003, ¶¶105-107.)

Determine a smoothed trajectory based on a low-light high-pass parameter.⁴

According to the ’840 patent, “smoothing the trajectory may result in the motion blur becoming more evident.” (EX1001, 18:37-42.) Therefore, a “low-light high-pass parameter” is used to control the balance between strength of smoothing and masking motion blur. (EX1001, 22:8-21.)

Specifically, a “low-light high-pass parameter may range between 0 and 1,

⁴ Claim 14 also is disclosed by Bell. (*Infra* § IV.A.11.)

with 0 corresponding to no motion blur compensation and 1 corresponding to full motion blur compensation.” (EX1001, 22:8-10.)

Therefore, a value of 1 for the parameter means there is no smoothing performed, which results in the resulting trajectory “includ[ing] high frequencies that mask motion blur within the visual content.” (EX1001, 22:14-18.) When no smoothing is performed, “the smoothed trajectory [is the] same as the (observed) trajectory during the exposure periods.” (EX1001, 22:5-7.) Conversely, a value of 0 (no compensation) “may result in the observed motion being fully attenuated” and the motion blur is left unaddressed. (EX1001, 22:10-14.)

a. Shi discloses determining a smoothed trajectory based on a low-light high-pass parameter.

Shi discloses that “[i]f a significant amount of motion blur is present in the frame . . . and the processed video is made very stable, the motion blur may become very noticeable and unnatural.” (EX1007, 17:59-63.) A POSITA would have understood that blurs can be caused by low light conditions. (EX1003, ¶111.) “To reduce a viewer’s perception of the blur,” Shi performs less smoothing: “the change in the virtual camera position from one frame to the next may be allowed to at least partially track the real camera position.” (EX1007, 17:63-66.)

Shi discloses using a low-light high-pass parameter—the “coefficient based on the motion blur kernel.” (EX1003, ¶¶110-120; EX1007, 18:30-33.)

Shi discloses using the coefficient to “adjust[] the second virtual camera position, $V_1(T)$, toward the real device position $R(T)$ to an extent determined based on the *amount of blur* identified.” (EX1007, 18:27-30.) Per Shi, determining the second virtual camera position is part of determining a smoothed trajectory. (EX1003, ¶116.) Therefore, Shi discloses that the coefficient controls (1) amount or strength of smoothing performed in smoothed trajectory generation and (2) an extent of motion blur compensation applied in determination of the smoothed trajectory. (EX1003, ¶117.)

More specifically, Shi discloses using the “coefficient to weight a spherical interpolation between positions $V_1(T)$ and $R(T)$.” (EX1007, 18:30-33.) A POSITA would have understood that less smoothing is involved if the second virtual camera position, $V_1(T)$, is adjusted toward the real device position $R(T)$. (EX1003, ¶118.) Conversely, more smoothing is involved if $V_1(T)$ is not adjusted toward $R(T)$, *i.e.*, there is no interpolation with $R(T)$. (EX1003, ¶119.) As the coefficient controls the “spherical interpolation between positions $V_1(T)$ and $R(T)$ ” (EX1007, 18:30-33), it controls the amount of smoothing performed. (EX1003, ¶120.)

Shi also discloses that “interpolation with $R(T)$ is only performed if the *amount of blur* exceeds a minimum amount represented by the threshold.” (EX1007, 18:34-37.) Shi therefore also discloses that the coefficient controls an

extent of motion blur compensation applied. (EX1003, ¶120.)

b. Combination of Shi and Bell.

It would have been obvious to combine Bell and Shi. Both Bell and Shi are directed to the same goal of stabilizing videos captured by mobile, or handheld, devices. (EX1006, Abstract; EX1007, Abstract.) Bell discloses that one issue with videos captured by mobile devices is that those videos “are plagued by camera shake.” (EX1006, 1:17-21.) Shi shares the same concern. According to Shi, “videos recorded using handheld recording devices suffer from unintentional movement that reduces video quality. One of the major causes of movement is hand shake.” (EX1007, 1:7-12.)

A POSITA looking at Bell’s video stabilization technique for mobile devices would have understood that the videos captured by mobile devices have undesired blurs. (EX1003, ¶123.) For example, Shi explains that the “undesired motions [from mobile devices] can cause *blur* . . . in recorded video footage.” (EX1007, 1:13-15.) The POSITA would have been motivated to turn to a reference that taught how to mitigate the blurs. (EX1003, ¶123.) It would have been obvious to use the known motion blur compensation technique in Shi to improve Bell’s similar video stabilization technique for mobile devices. (EX1003, ¶123.)

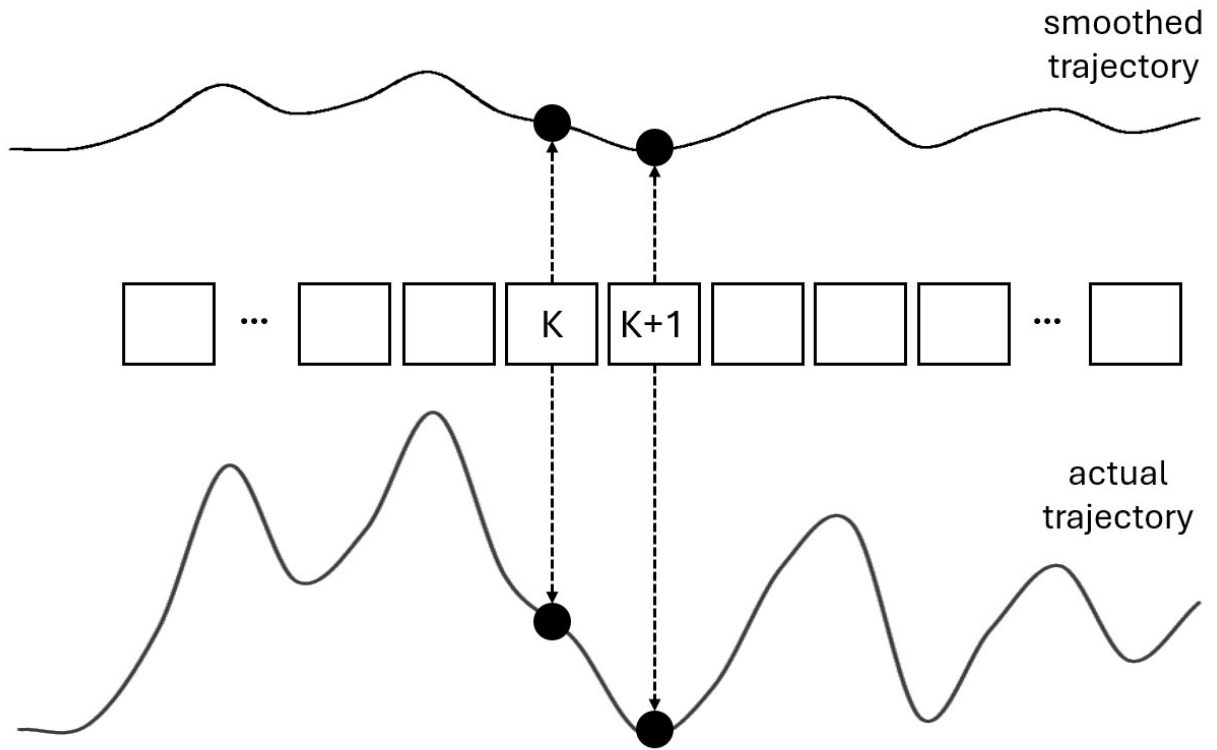
The Bell-Shi combination improves Bell’s video stabilization technique for mobile devices by addressing motion blur and would have had a reasonable

expectation of success. (EX1003, ¶124.)

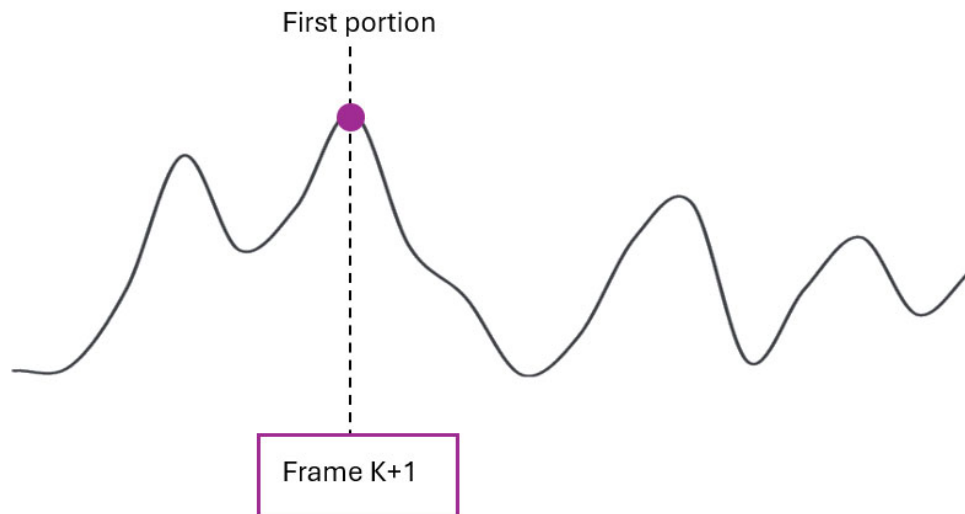
Thus, the Bell-Shi combination discloses determining a smoothed trajectory of the image capture device based on a low-light high-pass parameter, which is “one or more of a weight-balance parameter, a low-light high-pass parameter, and/or a stickiness parameter.” (EX1003, ¶125.)

- m. **1[g][ii]: “the look-ahead of the trajectory including use of a subsequent portion of the trajectory to determine a preceding portion of the smoothed trajectory such that a portion of the smoothed trajectory corresponding to the first portion of the trajectory is determined based on the second portion of the trajectory,”**

The trajectory corresponds to the smoothed trajectory. Bell discloses that the physical and virtual camera orientations have a one-to-one corresponding relationship. (EX1003, ¶126.) For example, Bell discloses that “physical and virtual camera orientations at frame k are represented with the quaternions $p(k)$ and $v(k)$.” (EX1006, 8:57-58.) As the below illustration shows, a POSITA would have understood that a portion of the trajectory corresponding to frame k , $p(k)$, has a corresponding portion of the smoothed trajectory ($v(k)$), which also corresponds to frame k . (EX1003, ¶126.) Likewise, a portion of the trajectory corresponding to frame $k + 1$ corresponds to a portion of the smoothed trajectory that corresponds to frame $k + 1$. (EX1003, ¶127.)



A portion of the smoothed trajectory corresponding to the first portion of the trajectory. As discussed for element 1[f][ii], physical camera orientation at *frame $k + 1$* corresponds to a first portion of the trajectory. (*Supra* § IV.A.3.k.)

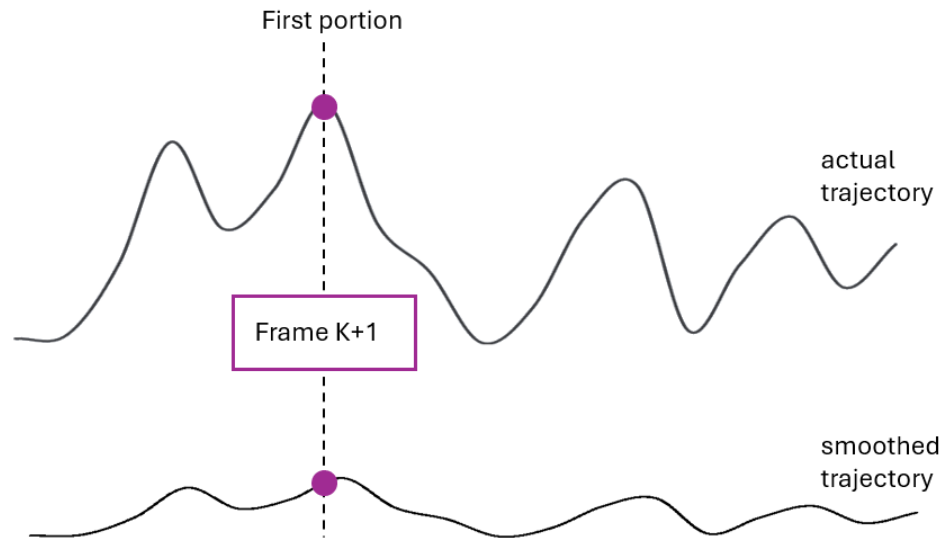


Bell further discloses a portion of the smoothed trajectory corresponding to *frame k + 1*. (EX1003, ¶¶127-130.)

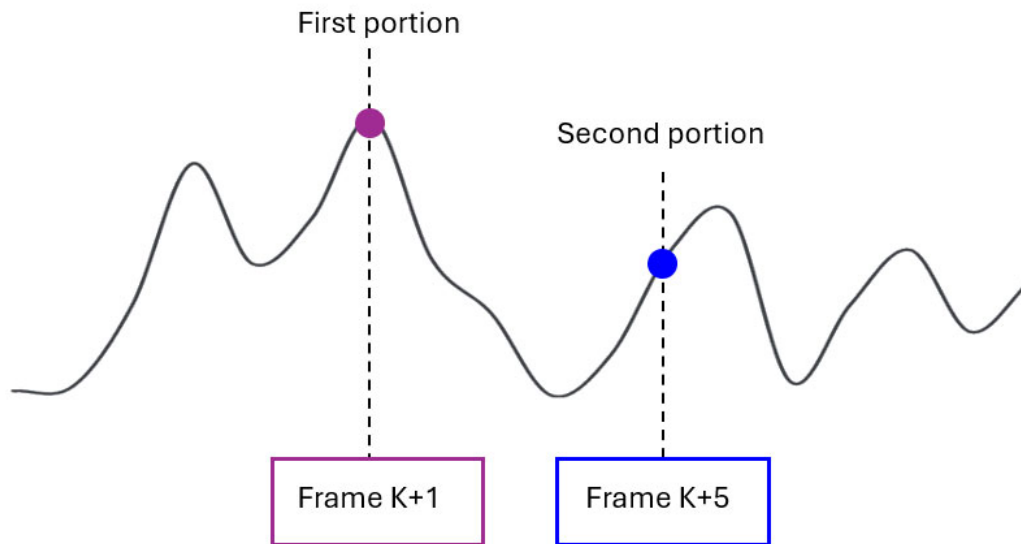
Bell discloses that its “smoothing algorithm determines $v(k)$ and $v_{\Delta}(k)$.” (EX1006, 8:61-67.) $v(k)$ represents “virtual camera orientation[] at frame k .” (EX1006, 8:57-58.) $v_{\Delta}(k)$ represents “virtual angular velocit[y],” which is the “discrete angular change[] from frame k to *frame k+1*.” (EX1006, 8:58-60.) With $v(k)$ and $v_{\Delta}(k)$, the virtual camera orientation at the next frame, *frame k + 1*, can be determined. (EX1003, ¶129.) That is because the calculation of $v(k)$ is as follows: $v(k)=v(k-1) \cdot v_{\Delta}(k-1)$. (EX1006, 9:30-36.) Substituting k with $k + 1$ results in: $v(k+1) = v(k) \cdot v_{\Delta}(k)$. (EX1003, ¶129.)

As there is a one-to-one corresponding relationship between the trajectory and the smoothed trajectory, $v(k+1)$ (i.e., virtual camera orientation at *frame k + 1*)

corresponds to a portion of the smoothed trajectory corresponding to the first portion of the trajectory. (EX1003, ¶130.)



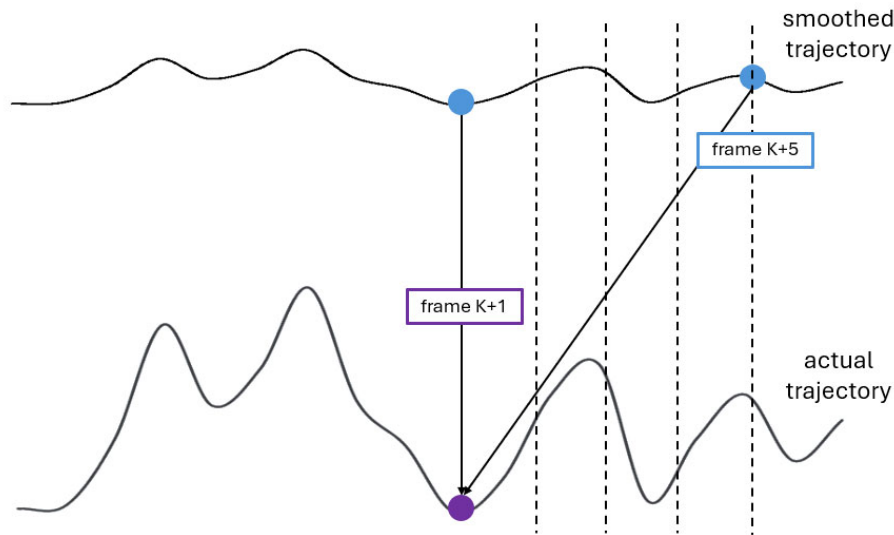
A portion of the smoothed trajectory . . . is determined based on the second portion of the trajectory. As discussed in the above paragraph, $v(k+1)$ is the claimed “portion of the smoothed trajectory.” As discussed in element 1[f][ii], $p(k+5)$, which is physical camera orientation at $frame\ k+5$, corresponds to a second portion of the trajectory. (*Supra* § IV.A.3.k.)



Bell discloses “project[ing] the orientation forward in time and compar[ing] it to the *orientation at the ‘future’ time.*” (EX1006, 10:17-19.) “Exemplary values for α ,” “the number of frames to look ahead,” “may be 2 to 5 frames.” (EX1006, 10:17-23.) Bell further discloses that its “smoothing algorithm determines $v(k)$ and $v_{\Delta}(k)$ ” using “*physical camera parameters* from . . . a small buffer of future frames, e.g., five or less.” (EX1006, 8:61-67.) Thus, Bell discloses using physical camera orientation at future frames, including *frame $k + 5$* (*i.e.*, *$p(k + 5)$*), when looking ahead. (EX1003, ¶132.)

Specifically, Bell discloses that after “look[ing] ahead,” “ $v_{\Delta}(k)$ may be updated.” (EX1006, 10:5-27.) In other words, “the virtual camera velocity that will be used for the next frame [*i.e.*, *frame $k + 1$*] is updated.” (EX1006, 10:34-

37.) And as explained before, $v(k+1) = v(k) \cdot v_{\Delta}(k)$. Therefore, looking ahead at future frames, including *frame k + 5*, determines the virtual camera position at *frame k + 1*. (EX1003, ¶134.) Bell thus discloses determining $v(k+1)$ based on $p(k+5)$. (EX1003, ¶¶132-134.)



The look-ahead of the trajectory including use of a subsequent portion of the trajectory to determine a preceding portion of the smoothed trajectory. As Bell discloses determining $v(k+1)$ based on $p(k+5)$, *i.e.*, using a portion of the trajectory corresponding to *frame k + 5* to determine a portion of the smoothed trajectory corresponding to *frame k + 1*, Bell therefore discloses “use of a subsequent portion of the trajectory to determine a preceding portion of the smoothed trajectory.” (EX1003, ¶136.) And as discussed for element 1[g][i], Bell discloses “us[ing] data from *a look-ahead buffer* to calculate a smoother path.”

(*Supra* § IV.A.3.1.)

- n. **1[g][iii]: “the smoothed trajectory having smoother changes in the rotational positions of the housing than the trajectory,”**

Bell’s device orientations measured by motion sensing system 190 characterize the rotational positions of the housing, and they form the trajectory.

(*Supra* §§ IV.A.3.h and IV.A.3.j.)

Bell discloses that “[t]he goal of the smoothing method is to define a new virtual camera that . . . follows the original camera with *smooth changes in angular velocity*.” (EX1006, 8:52-56.) A POSITA would have understood that angular velocity characterizes rotational positions. (EX1003, ¶138.)

As discussed before, the virtual camera positions correspond to the smoothed trajectory, and the actual camera positions correspond to the trajectory.

(*Supra* § IV.A.3.1.) Therefore, Bell teaches this limitation. (EX1003, ¶139.)

- o. **1[g][iv]: “wherein the determination of the smoothed trajectory based on the look-ahead of the trajectory”**

Bell discloses this limitation. (*Supra* § IV.A.3.1 and IV.A.3.lm.)

- p. **1[g][v]: “includes use of a temporal horizon of motion experienced by the image capture device, the temporal horizon enabling approximation of intentional motion of the image capture device while causing a delay in the determination of the smoothed trajectory”;**

Use of a temporal horizon of motion experienced by the image capture

device, the temporal horizon enabling approximation of intentional motion of the image capture device.

According to the '840 patent, “[i]ntentional motion may refer to the motion that was intended by the user of the image capture device,” such as “a forward motion to follow an object of interest, a panning motion to shoot a wide view of a scenery, and/or other motion.” (EX1001, 9:3-5, 9:8-10.) On the other hand, “[u]nintentional motion may be characterized by higher frequency motion (e.g., jitter, shake) than intentional motion.” (EX1001, 9:18-20.)

The '840 patent provides no definition of “temporal horizon.” (EX1003, ¶143.) The '840 patent, however, states that “[t]he positions/motions of the image capture device in the ‘future’ may be analyzed (look ahead) to determine whether particular position(s)/motion(s) of the image capture device at a moment is an *intended motion* or an unintended motion (e.g., shaking due to vibration, rotation due to bumping/mishandling the image capture device).” (EX1001, 15:1-6.) The '840 patent also states that “when determining the smoothed trajectory for a moment (e.g., corresponding to the 1000th video frame) within the capture duration, the position(s)/motion(s) of the image capture device for a duration subsequent to the moment (e.g., corresponding to the next 30 video frames, corresponding to next second of capture) may be analyzed to determine whether the position/motion of the image capture device at the moment was *intended* or

not.” (EX1001, 15:6-15.) The ’840 patent additionally states that “[u]sing the subsequent portion(s) of the (observed) trajectory may enable the smoothed trajectory component 104 to determine a smoothed trajectory that preserves a user’s *intended motion* for the image capture device.” (EX1001, 14:52-58.)

Bell discloses this limitation. Bell discloses that its goal is to “define a new virtual camera that stays static when the motion is so small that the *intention* is probably that the camera should remain static.” (EX1006, 8:52-55.) Bell further discloses that it does so by using the device orientations corresponding to successive future frames. (EX1003, ¶¶145-147.) Bell discloses that “[i]f frames can be buffered within the camera pipeline for a short time before being processed, then a larger time window of motion sensing system data can be examined, and sharp changes can be preemptively avoided.” (EX1006, 10:10-13.) As “motion sensing system data” is utilized, a POSITA would have understood that Bell’s stabilization technique uses motion experienced by the device during future frames. (EX1003, ¶¶145-147.)

A POSITA would have further understood that the motion experienced by the device during future frames is used to approximate intentional motion, as Bell discloses that “sharp changes can be preemptively avoided” by using the motion data in the “look-ahead buffer.” (EX1003, ¶147; EX1006, 10:10-16.) This allows the “virtual camera [to] stay[] static when the motion is so small that the *intention*

is probably that the camera should remain static.” (EX1006, 8:52-55.) The motion experienced by the device during successive future frames, which is used to approximate intentional motion, is the “temporal horizon of motion experienced by the image capture device.” (EX1003, ¶148.)

Shi also discloses this limitation. Shi discloses that the stabilization of a current frame uses “future frames during processing” and that the future frames correspond to the motion of the device. (EX1007, 12:25-29, 2:62-65, 5:56-64, 10:43-59; EX1003, ¶¶149-157.) Additionally, Shi discloses that consistent movement, such as panning, is the “intentional motion.” (EX1007, 1:24-27, 13:35-40.) Shi discloses that its “system can reduce the impact of large or sudden movements using look-ahead frames” and “evaluate whether consistent movement, such as panning, is occurring and adjust the video frames to reflect the consistent movement.” (EX1007, 5:17-24.) Shi therefore also discloses using motion experienced by the device in successive future frames to evaluate and preserve intentional motion, such as panning. (EX1003, ¶¶149-157.)

Causing a delay in the determination of the smoothed trajectory. Bell discloses using “data from a look-ahead buffer to calculate a smoother path” and preemptively avoid sharp changes and that “frames [are] buffered within the camera pipeline for a short time before being processed.” (EX1006, 10:10-16.) A POSITA would have understood that as the determination of the portion of the

smoothed trajectory corresponding to the current frame must wait for future frames to be captured and then buffered, there is a delay. (EX1003, ¶¶158-159.)

Shi also discloses this limitation. (EX1003, ¶¶161-164.) For example, Shi discloses that “analysis of future frames can be performed by implementing a small delay between frame capture and . . . processing to allow an appropriate number of later-captured frames to become available for use in stabilizing earlier-captured frames.” (EX1007, 2:1-5, 12:26-29, 5:17-19; *see also id.*)

Combination. As discussed above, Bell already discloses this limitation. Additionally, a POSITA would have been motivated to combine the known technique disclosed in Shi with Bell. (EX1003, ¶¶165-170.)

It would also have been obvious to use the known intentional motion approximation technique in Shi to improve Bell’s similar video stabilization technique for mobile devices. (EX1003, ¶169.)

The Bell-Shi combination would have had a reasonable expectation of success as both Bell and Shi use the look-ahead frames to determine the smoothed trajectory and perform video stabilization. (EX1003, ¶170.)

- q. **1[h][i]: “determine placement of the viewing window for the visual content with respect to the field of view of the visual content based on the smoothed trajectory of the housing,”**

The content within the captured image, or input frame, is the “visual

content.” (*Supra* § IV.A.3.e.) The entire field within the *captured image* is its field of view. (*Supra* § IV.A.3.e.) The *crop polygon* within *captured image* is the “viewing window.” (*Supra* § IV.A.3.f.)

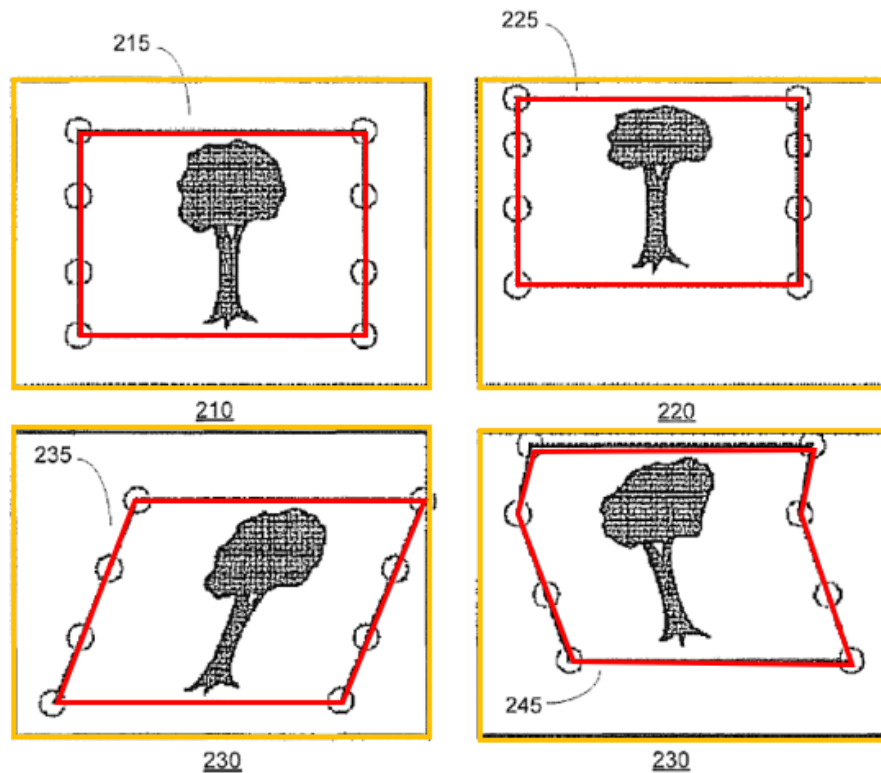


Fig. 2

Bell additionally discloses that there is “a small amount of ‘padding’ between the *polygon borders* and the *input frame edges*.” (EX1006, 9:12-15.) In other words, padding = *captured image* - *crop polygon*. (EX1003, ¶172.)

Bell discloses determining placement of the *crop polygon* with respect to the *captured frame* based on the smoothed trajectory. (EX1003, ¶173.) Bell discloses first determining an “orientation for the virtual camera,” which corresponds to the

smoothed trajectory. (EX1006, 9:1-6; *supra* § IV.A.3.1.)

Bell additionally discloses that the “[virtual] camera orientation” is “used to determine the coordinates of the corners of the resulting *crop polygon*.” (EX1006, 9:6-9.) Determining the coordinates of the corners of the resulting *crop polygon* means determining the placement of the *crop polygon* with respect to the field of view of *captured frame*. (EX1003, ¶¶174-178.) This is because Bell discloses that the “*crop* [*polygon* is] within the *input frame*” and also discloses determining if “any corner of the *crop polygon* would fall outside the *input frame*.” (EX1006, 6:66-7:1,10:3-7.)

As the determination of the placement of the *crop polygon* with respect to the field of view of *captured frame* is based on the orientation of the virtual camera, which corresponds to a portion of the smoothed trajectory, Bell discloses determining placement of the viewing window with respect to the field of view of the visual content, based on the smoothed trajectory of the housing. (EX1003, ¶¶171-178.)

- r. **1[h][ii]: “the viewing window defining one or more extents of the visual content to be included within a punchout of the visual content; and”**

According to the ’840 patent, a “punchout of visual content may refer to using one or more portions of the visual content for presentation, such as a cut-out of the image or a cropped portion of the image.” (EX1001, 12:52-55.) The ’840

patent additionally states that “[t]he size of the viewing window may refer to the field of view of the punchout for generating a video based on visual content.”

(EX1001, 20:10-12.) In other words, the claimed “punchout of the visual content” defines the portion of the displayed video frames and is of the same size of the viewing window. (EX1003, ¶179.)

As discussed for element 1[c][iii], the *crop polygon* is “for generating stabilized visual content.” (*Supra* § IV.A.3.f.) Bell additionally discloses that “the region within the polygon is projectively warped to create the output video.” (EX1006, 7:4-5.) A POSITA would have understood that the region within the *crop polygon* is the “punchout of the visual content.” (EX1003, ¶181.) Therefore, the *crop polygon* defines the extents of the *captured frame* included within the region within the *crop polygon*. (EX1003, ¶181.)

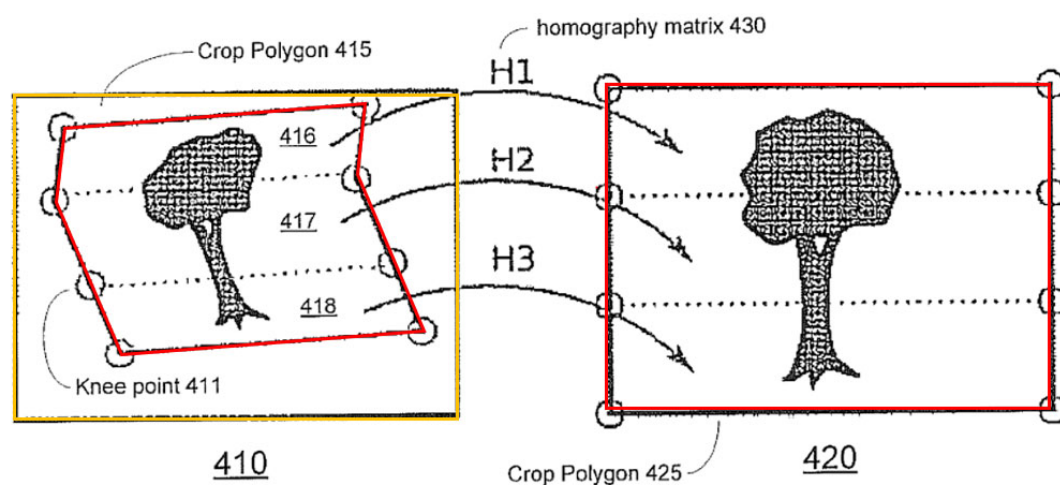


Fig. 4

- s. **1[i][i]: “generate the stabilized visual content of a video based on the viewing window, the stabilized visual content including the punchout of the one or more extents of the visual content within the viewing window,”**

Generate the stabilized visual content of a video based on the viewing window. The **crop polygon** is “for generating stabilized visual content.” (*Supra* § IV.A.3.f.)

The stabilized visual content including the punchout of the one or more extents of the visual content within the viewing window. The region within the **crop polygon** is the “punchout of the visual content.” (*Supra* § IV.A.3.r.) Bell also discloses that the stabilized image, or “the output,” is “synthesized by projectively warping the **crop polygon** from the video input to the virtual camera.” (EX1006, 10:51-53.) Therefore, the stabilized image includes the region within the **crop polygon**, *i.e.*, the punchout, which includes the portion of the **captured frame** that is within the **crop polygon**. (EX1003, ¶¶182-185.)

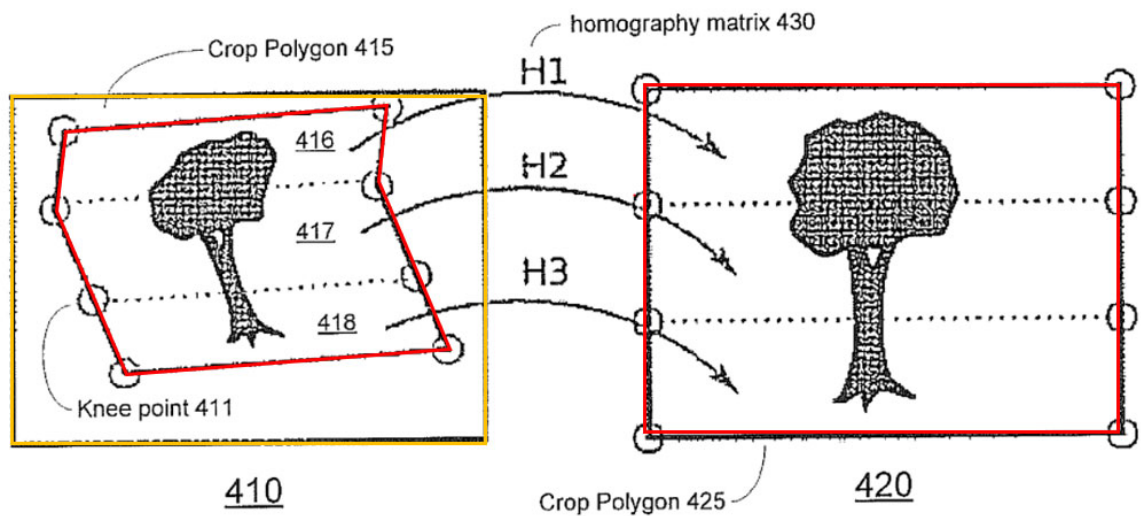


Fig. 4

- t. 1[i][ii]: “wherein inclusion in the stabilized visual content of the one or more extents of the visual content within the viewing window effectuates stabilization of the visual content via selective cropping.”

Inclusion in the stabilized visual content of the one or more extents of the visual content within the viewing window. As discussed above, Bell discloses that the stabilized video includes the images within the *crop polygon*. (*Supra* § IV.A.3.s.)

Effectuate stabilization of the visual content via selective cropping. Bell discloses that “the crop polygon . . . remove[s] motion distortions.” (EX1006, Abstract.) The crop polygon serves the same purpose as the conventional crop rectangle, which “tracks the scene content across a time-sequence of video frames as the camera shakes around [T]he content within the crop rectangle is mostly

stable, producing a smooth output.” (EX1006, 6:63-7:1.) Bell therefore discloses that the stabilization of the video is performed by using the *crop polygon*.

(EX1003, ¶186.)

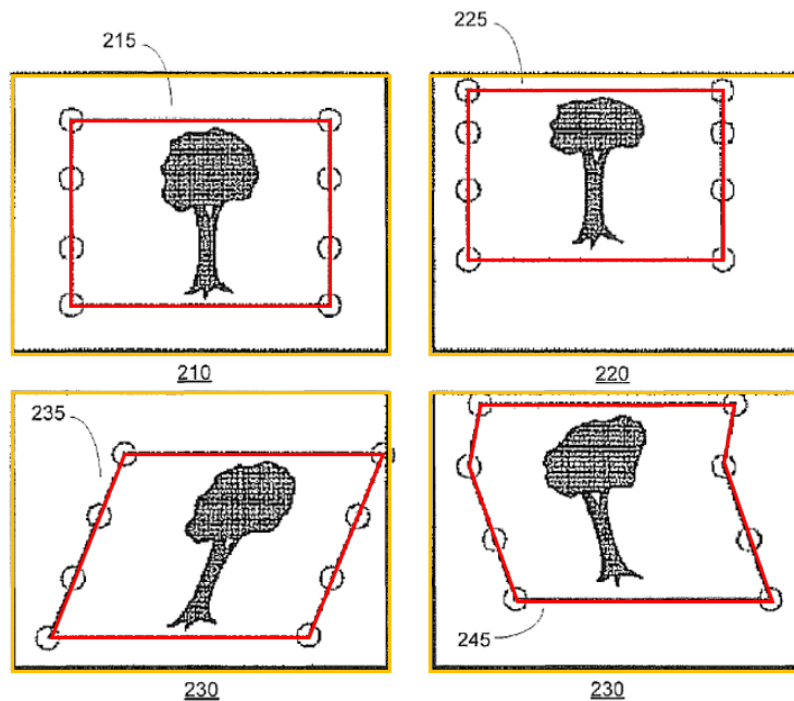


Fig. 2

Bell further discloses selective cropping. (EX1003, ¶187.) Specifically, Bell discloses that the “*crop polygon* is smaller than the *input size*,” with the difference being “a small amount of ‘padding’ between the *polygon borders* and the *input frame edges*.” (EX1006, 9:12-15.) Therefore, a POSITA would have understood that the padding is selectively removed from the *input frame* to “create the output video.” (EX1006, 7:1-5; EX1003, ¶188.) As Figure 4 below shows, only the image within the *crop polygon* is included in the stabilized image; the

padding is removed and not part of the output image. (EX1003, ¶188.)

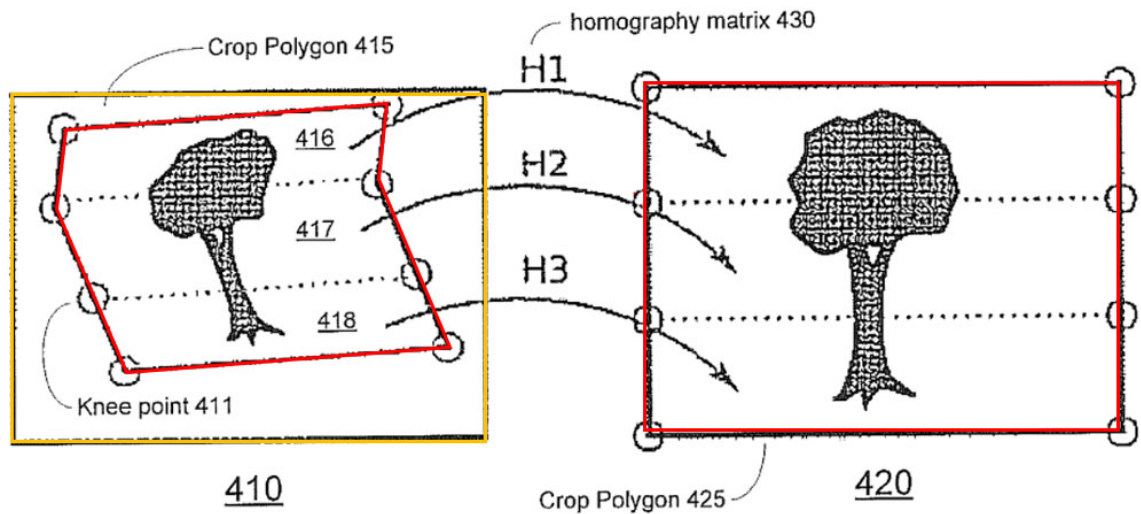


Fig. 4

4. **Claim 2: “The image capture system of claim 1, wherein the position information further characterizes translational positions of the housing at the different moments within the capture duration, the trajectory further reflects the translational positions of the housing at the different moments within the capture duration, and the smoothed trajectory has smoother changes in the translational positions of the housing than the trajectory.”**

The orientation of electronic system 100, measured by the motion sensing system 190, is the “position information.” (*Supra* § IV.A.3.h.)

Position information further characterizes translational positions. Bell discloses that “motion sensing system 190 may comprise any suitable *linear* and/or angular motion sensing technology.” (EX1006, 6:10-14.) Additionally, “motion sensing system 190 may sense and report motion in multiple axes, including for example, *X*, *Y*, and/or *Z* axes.” (EX1006, 6:14-16.) Bell further discloses that

“camera movement [can be] in more than one *linear axis* and/or one or more angular rotations.” (EX1006, 7:31-34.) A POSITA would have understood that the orientation information characterizes translational positions of electronic system 100 and its housing. (EX1003, ¶190.)

As discussed for element 1[d][ii], Bell discloses that the measured position information, *i.e.*, device orientations, corresponds to different moments within the duration of the captured video. (*Supra* § IV.A.3.h.) Bell therefore discloses that the position information further characterizes translational positions of the housing at the different moments within the capture duration. (EX1003, ¶191.)

The trajectory further reflects the translational positions. As the trajectory is formed by electronic system 100’s position information, across the duration of the captured video (*supra* § IV.A.3.j), the trajectory further reflects the translational positions of the housing at the different moments within the capture duration. (EX1003, ¶192.)

The smoothed trajectory has smoother changes in the translational positions. Bell discloses generating a smoothed trajectory with smoother changes in motion. (EX1003, ¶193.) Bell discloses “us[ing] data from a look-ahead buffer to calculate a *smoother* path” and preemptive avoid “sharp changes.” (EX1006, 10:14-16.)

A POSITA would have understood that the smoother changes in motion include motion in the translational positions — as the motion sensing system

measures motion in translational positions and the camera movement can be translational. (EX1003, ¶193.)

5. Claim 3

- a. 3[i]: “The image capture system of claim 1, wherein the smoothed trajectory is determined further based on a trajectory retrieval algorithm,”**

Bell discloses a “smoothing method [which] define[s] a new virtual camera.” (EX1006, 8:52-56.) The movement of the virtual camera is the smoothed trajectory. (*Supra* § IV.A.3.1.) Specifically, “[f]or each new frame k , [Bell’s] smoothing algorithm determines $v(k)$ and $v_{\Delta}(k) \dots$ ” (EX1006, 8:63-67.) $v(k)$ and $v_{\Delta}(k)$ represent the virtual movement of the device and determine the smoothed trajectory. (*Supra* § IV.A.3.m.)

Bell’s smoothing algorithm therefore is the “trajectory retrieval algorithm.” (EX1003, ¶¶194-195.)

- b. 3[ii]: “input to the trajectory retrieval algorithm including the trajectory of the housing and a bootstrap trajectory, the bootstrap trajectory including a current or an initial guess of the smoothed trajectory.”**

Bootstrap trajectory. Bell discloses that “[a]s a first step, $\dots$ the present invention hypothesize[s] a new orientation for the virtual camera $\dots$ ” (EX1006, 9:1-3.) The hypothetical camera orientation $\hat{v}(k)$ is determined based on the virtual camera orientation and velocity of the last frame, $k-1$. (EX1003, ¶196.) Bell

discloses that $\hat{v}(k) = v(k-1) \cdot v_{\Delta}(k-1)$. (EX1006, 9:5.)

The hypothetical camera orientation at frame k , $\hat{v}(k)$, is a current or an initial guess of the device orientation at frame k . (EX1003, ¶197.) As discussed for element 1[g][ii], the virtual device orientation at frame k corresponds to a portion of the smoothed trajectory. (*Supra* § IV.A.3.m.) Therefore, $\hat{v}(k)$ is an initial, estimated portion of the smoothed trajectory. (EX1003, ¶197.) $\hat{v}(k)$ and the already known virtual orientations at past frames (0 to $k-1$) form a current guess of the smoothed trajectory. (EX1003, ¶198.)

The estimated smoothed trajectory is used as an input to Bell's smoothing algorithm. (EX1003, ¶199.) Bell discloses using $\hat{v}(k)$ to determine the resulting crop polygon, and further decide whether to update the estimated virtual orientation at frame k . (EX1006, 9:1-9, 9:30-37.)

Trajectory of the housing. Bell also discloses that as part of the smoothing algorithm, "if any part of the hypothetical crop polygon lies outside the inner region, the angular change is updated to bring it closer to the physical camera's rate of change" (EX1006, 9:43-45.) In order to do that, the algorithm "calculates the frame to frame motion of the physical camera" (EX1006, 9:54-56.) Specifically, the physical positions of the device at frames k and $k-1$ are used to calculate the motion of the physical camera. (EX1003, ¶¶200-204.) Also, the positions of the device at different moments constitute the trajectory of the

housing. (*Supra* § IV.A.3.j.) Therefore, the trajectory of the housing is an input to the trajectory retrieval algorithm. (EX1003, ¶¶200-204.)

6. Claim 5: “The image capture system of claim 1, wherein a longer interval of the temporal horizon enables better approximation of the intentional motion of the image capture device while causing a longer delay in the determination of the smoothed trajectory.”

Bell and Shi disclose an image capture system that approximates intended motion using the motion experienced by the system during future frames. (*Supra* § IV.A.3.p.)

Bell discloses that if “a *larger time window* of motion sensing system data can be examined, and sharp changes can be preemptively avoided.” (EX1006, 10:10-13.) Therefore, Bell discloses that if motion experienced by the device during more future frames is considered, then intentional motion can be better approximated, and sharp changes can be avoided. (EX1003, ¶207.)

Shi discloses the use of “look-ahead frames to increase the accuracy of [image stabilization] processing.” (EX1007, 5:17-19.) Shi also discloses “determin[ing] a probability, S, that consistent motion (e.g., panning) of the recording device is occurring over a time range.” (EX1007, 18:53-55.) Panning is one example of “intentional motion.” (*Supra* § IV.A.3.p.) As “panning may occur over a second or several seconds,” it is beneficial to consider “*a relatively long history of device positions* determined over the course of the video capture.”

(EX1007, 18:60-66.) Therefore, Shi discloses that more device positions, *i.e.*, motion corresponding to more future frames, enable better approximation of the intentional motion. (EX1003, ¶¶208-209.)

Additionally, it would have been obvious to try to better approximate intentional motion by analyzing motion experienced during more future frames. (EX1003, ¶¶210-211.) Statistically speaking, the more frames that are considered, the more accurate the determination of the probability S is. (*Id.*) A POSITA therefore would have been motivated to consider more frames to determine whether panning is considering. (*Id.*)

It would also been obvious to try utilizing a longer interval of the temporal horizon of the motion experienced by the device, and a POSITA would have a reasonable expectation of success. (EX1003, ¶212.) To approximate intentional motion such as panning, a POSITA is faced with three, finite options: using motion experienced during more future frames, fewer future frames, or the same amount of future frames. (EX1003, ¶212.)

As discussed for element 1[g][v], the determination of the smoothed trajectory is delayed “by the time required to capture” the future frames. (*Supra* § IV.A.3.p.) A POSITA would have understood that considering more frames would cause a longer delay. (EX1003, ¶213.)

7. **Claim 8: “The image capture system of claim 1, wherein the determination of the smoothed trajectory based on the look-ahead of the trajectory includes weighing past positions of the housing less than future positions of the housing.”**

Bell discloses determining a smoothed trajectory based on a look-ahead of the trajectory. (*Supra* § IV.A.3.1.)

Bell discloses that that its “smoothing algorithm determines $v(k)$ and $v_{\Delta}(k)$ using . . . the *physical camera parameters from the last frame*, the current frame, and, optionally, *a small buffer of future frames, e.g., five or less.*” (EX1006, 8:63-67.) $v(k)$ represents the virtual position of the device at frame k . (*Supra* § IV.A.3.m.) Additionally, knowing $v(k)$ and $v_{\Delta}(k)$ allows the determination of $v(k+1)$, the virtual position of the device at frame $k + 1$. (*Supra* § IV.A.3.m.) Therefore, Bell discloses that in determining the virtual positions of the device, *i.e.*, the smoothed trajectory, Bells considers just one past frame but five or fewer future frames. (EX1003, ¶216.)

If two or more than two future frames are considered, then more future frames are considered than past frames. (EX1003, ¶217.) Bell indeed discloses that “[e]xemplary values” for the “number of frames to look ahead,” “may be 2 to 5 frames.” (EX1006, 10:17-25.) If device orientations corresponding to fewer past frames compared to future frames are used for determining the smoothed trajectory, that means past device orientations are weighed less compared to future

device orientations. (EX1003, ¶¶217-218.)

Moreover, it would have been obvious to try to weigh past positions of the device (and its housing) less than future positions. (EX1003, ¶219.) In determining the stabilized trajectory, a POSITA is faced with three finite options: weigh past and future positions equally, weigh past positions less than future positions, and weigh past positions more than future positions. (EX1003, ¶219.) Given Bell’s disclosure of the benefit of “looking-ahead,” a POSITA would be motivated to try weighing past positions less than future positions with reasonable expectation of success. (EX1003, ¶220.)

8. **Claim 9: “The image capture system of claim 1, wherein the determination of the smoothed trajectory based on the look-ahead of the trajectory includes a prior portion of the trajectory having less impact on the determination of the smoothed trajectory than a future portion of the trajectory.”**

As discussed for claim 8, Bell discloses or makes obvious weighing the position information of past frames less than that of future frames when determining the smoothed trajectory. (*Supra* § IV.A.7.) A POSITA would have understood that means past frames have less impact on the determination of the smoothed trajectory than future frames. (EX1003, ¶221.)

9. **Claim 10: “The image capture system of claim 1, wherein the determination of the smoothed trajectory based on the look-ahead of the trajectory includes using a first amount of a prior portion of the trajectory and a second amount of a**

future portion of the trajectory, the first amount being smaller than the second amount.”

As discussed for claim 8, Bell discloses using one past frame but five or less future frames, when determining the smoothed trajectory. (*Supra* § IV.A.7.) Bell therefore teaches this claim. (EX1003, ¶222.)

Moreover, it would have been obvious to try to use fewer past frames than future frames when determining the smoothed trajectory. (EX1003, ¶223.) In determining the smoothed trajectory, a POSITA is faced with three finite options: considering (1) an equal amount of past and future frames, (2) fewer past frames than future frames, and (3) more past frames than future frames. (EX1003, ¶223.) Given Bell’s emphasis on the unique advantages of future frames, a POSITA would be motivated to try considering fewer past frames than future frames when determining the smoothed trajectory. (EX1003, ¶224.)

- 10. Claim 11: “The image capture system of claim 1, wherein the visual content are stored in a buffer during smoothed trajectory generation such that the visual content corresponding to the portion of the smoothed trajectory are stored in the buffer during the determination of the portion of the smoothed trajectory using the look-ahead of the trajectory, and the visual content stored in the buffer are used after the determination of the portion of the capture trajectory to generate the stabilized visual content.”**

Visual content are stored in a buffer during smoothed trajectory generation.⁵

The content within the images captured by electronic system 100 is the “visual content.” (*Supra* § IV.A.3.e.)

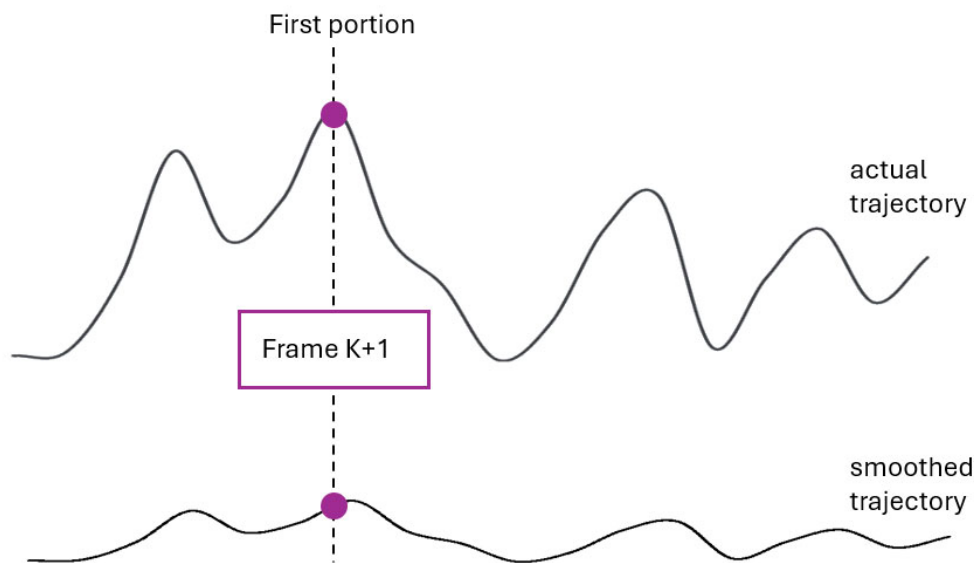
Bell discloses determining the virtual camera positions (*i.e.*, smoothed trajectory) and using them to determine the placement of the crop polygon with respect to the captured frame. (*Supra* § IV.A.3.q.) Afterwards, “the region within the polygon is projectively warped to create the output video.” (EX1006, 7:1-5.) A POSITA would have understood that the captured image must have been stored by the device during the generation of the smoothed trajectory; otherwise, it would be impossible to later create the output video, which includes a portion of the captured frame. (EX1003, ¶¶225-228.)

It would have been obvious to store the captured images in a buffer. (EX1003, ¶229.) Bell discloses storing image frames in a buffer: “frames can be buffered within the camera pipeline for a short time before being processed” (EX1006, 10:10-13.)

Visual content corresponding to the portion of the smoothed trajectory are stored in the buffer during the determination of the portion of the smoothed

⁵ This limitation is not grammatically correct. It should be “visual content[s] are stored” or “visual content [is] stored.”

trajectory using the look-ahead of the trajectory.⁶ “The portion of the smoothed trajectory” refers to “a portion of the smoothed trajectory corresponding to the first portion of the trajectory” in claim 1. (EX1003, ¶230.) The claimed “the portion of the smoothed trajectory” therefore is a point on the smoothed trajectory that corresponds to frame $k + 1$, *i.e.*, virtual camera position corresponding to frame $k + 1$, $v(k + 1)$. (*Supra* § IV.A.3.m.) The claimed “visual content corresponding to the portion of the smoothed trajectory” is frame $k + 1$. (*Id.*)



Bell additionally discloses determining the virtual camera position corresponding to frame $k + 1$, $v(k + 1)$, using a look-ahead of the trajectory. (*Supra* § IV.A.3.1.)

⁶ This limitation is not grammatically correct.

As Bell discloses that virtual camera position is used to determine the placement of the crop polygon within the captured image, a POSITA would have understood that the captured image (*i.e.*, frame $k + 1$) is stored (*e.g.*, in a buffer) during the determination of $v(k + 1)$. (EX1003, ¶232.) Otherwise, the device would not be able to use the captured image after the virtual camera position ($v(k + 1)$) is decided. (EX1003, ¶232.)

The visual content stored in the buffer are used after the determination of the portion of the capture trajectory⁷ to generate the stabilized visual content. Bell discloses that “[a]fter determining a new orientation of the virtual camera, the output may be synthesized by projectively warping the crop polygon from the video input to the virtual camera.” (EX1006, 10:51-53.) Therefore, Bell discloses that after $v(k + 1)$, which corresponds to the portion of the smoothed trajectory, is determined, the captured image for frame $k + 1$ stored in the buffer is used to generate the stabilized video. (EX1003, ¶¶233-235.)

11. Claim 14: “The image capture system of claim 1, wherein the low-light high-pass parameter controls amount or

⁷ “The capture trajectory” has no antecedent basis. Claim 1, from which claim 11 depends, does not recite a “capture trajectory.” For purposes of this proceeding only, Petitioner interprets “the capture trajectory” to refer to the smoothed trajectory.

strength of smoothing performed in smoothed trajectory generation and an extent of motion blur compensation applied in determination of the smoothed trajectory.”

As shown in claim element 1[g][i], Bell and Shi disclose a low-light high-pass parameter that meets this claim. (*Supra* § IV.A.3.1.)

12. Claim 16

- a. Preamble: “A method for stabilizing videos, the method performed by an image capture system including one or more processors, an optical element, an image sensor, and a position sensor, the optical element carried by an image capture device and configured to guide light within a field of view to the image sensor, the method comprising:”**

To the extent the preamble is limiting (which Petitioner does not concede), Bell discloses the preamble of claim 16.

As shown for the preamble of claim 1, Bell discloses “systems and methods for real-time *video stabilization* for mobile devices based on on-board motion sensing.” (*Supra* § IV.A.3.a.) The “mobile devices” are image capture systems. (*Id.*) Bell further discloses or renders obvious the components included in the “image capture system,” as shown in elements 1[b]-1[e]. (*Supra* §§ IV.A.3.c - IV.A.3.i.)

- b. 16[a]: “generating, by the image sensor, a visual output signal conveying visual information based on light that becomes incident thereon during a capture duration, the visual information defining visual content having the field of view, the field of view being**

larger than a size of a viewing window for generating stabilized visual content”;

Element 16[a] recites limitations similar to 1[c][i], 1[c][ii], and 1[c][iii], except that it requires that the image sensor generate a visual output signal.

(EX1003, ¶240.) Image sensors generate a visual output signal. (EX1003, ¶241.)

- c. 16[b]: “generating, by the position sensor, a position output signal based on positions of the image capture device during the capture duration, the position output signal conveying position information that characterizes rotational positions of the image capture device at different moments within the capture duration”;**

Element 16[b] recites limitations similar to 1[d][i] and 1[d][ii], except that it requires that the position sensor generate a position output signal. (EX1003, ¶242.)

Motion sensors generate a position output signal. (EX1003, ¶243.)

- d. 16[c] – 16[f]**

Elements 16[c]-16[f] mirror elements 1[f]-1[i], respectively, except that they recite trajectories and positions of the image capture device, instead of the housing.

(EX1003, ¶¶244-245.)

13. Claim 18

Claim 18 mirrors claim 3, except that it requires the trajectory of the image capture device, instead of the housing. (EX1003, ¶246.)

14. Claim 19

- a. **Preamble: “The method of claim 16, wherein the determination of the smoothed trajectory based on the look-ahead of the trajectory includes:”**

To the extent the preamble is limiting (which Petitioner does not concede), Bell and Shi disclose the preamble of claim 19. (*Supra* § IV.A.12.)

- b. **19[a]: “weighing past positions of the image capture device less than future positions of the image capture device;”**

Element 19[a] mirrors a subset of claim 8, except that it recites positions of the image capture device, instead of the housing. (EX1003, ¶248.)

- c. **19[b]: “a prior portion of the trajectory having less impact on the determination of the smoothed trajectory than a future portion of the trajectory; or”**

Element 19[b] mirrors a subset of claim 9. (EX1003, ¶249.)

- d. **19[c]: “using a first amount of the prior portion of the trajectory and a second amount of the future portion of the trajectory, the first amount being smaller than the second amount.”**

Element 19[c] mirrors a subset of claim 10. (EX1003, ¶250.)

B. Ground 2: Bell, Shi, and Karpenko Render Obvious Claims 4 and 20

1. Summary of Karpenko

Karpenko was filed on December 9, 2013 and issued on June 30, 2015, qualifying as prior art to the '840 patent at least under §§ 102(a)(1) and 102(a)(2). The Examiner cited Karpenko's related publication but never applied it during prosecution.

Karpenko discloses “[s]ystems and methods for digital video stabilization via constraint-based rotation smoothing” (EX1009, Abstract.) Specifically, Karpenko discloses an algorithm that performs a “sufficiently high” number of “iterations” that “result in smooth rotations.” (EX1009, 13:19-35.)

Karpenko additionally discloses a way to “improve the efficiency” of the algorithm by “running it in coarse to fine mode.” (EX1009, 16:19-21, 16:59-60.) This allows the algorithm “to be run for significantly fewer iterations at each step.” (EX1009, 17:1-2.)

2. **Claim 4: “The image capture system of claim 3, wherein the smoothed trajectory is determined further based on iterations of smoothing including a fine scale approach to smooth on a small scale and a coarse approach to smooth on a large scale, the iterations of smoothing resulting in convergence of the smoothed trajectory.”**

Iterations of smoothing. Karpenko discloses that a constraint-based rotation smoothing algorithm, algorithm 1, that uses “sufficiently high” “number of *iterations*” to “result in smooth rotations.” (EX1009, 13:32-33.) Karpenko discloses that after the iterations of smoothing, the resulting ϕ_i , “new smoothed camera orientations ϕ_i ,” are used to “produce a stabilized output video.” (EX1009, 10:17-25, 13:14-34; EX1003, ¶¶256-260.)

A fine scale approach and a coarse approach. Karpenko discloses that that the efficiency of the smoothing algorithm can be improved by “running it in *coarse*

to fine mode.” (EX1009, 16:57-60.) Specifically, Karpenko discloses that algorithm 1 may be “initialized by supplying and solving for every 16th orientation ϕ_i .” (EX1009, 16:30-61.) Karpenko further discloses that the algorithm “may be run again, but now optimizing every 8th orientation for smoothness and constraints.” (EX1009, 16:64-67.) According to Karpenko, “[r]unning coarse to fine may permit algorithm 1 to be run for significantly fewer iterations at each step. The overall camera motion may be smoothed out and then the orientations may be refined to meet constraints at increasingly smaller frame intervals while still remaining smooth.” (EX1009, 17:1-5.)

Convergence of the smoothed trajectory. Karpenko discloses that the coarse to fine mode can “improv[e] the rate of *convergence* of algorithm 1.” (EX1009, 16:57-60.) As a result, “a smoothed orientation has been computed for all frames.” (EX1009, 17:9-11.) A POSITA would have understood that Karpenko’s algorithm 1’s iterations of the coarse to fine smoothing result in convergence of the smoothed trajectory. (EX1003, ¶265.)

Combination. It would have been obvious to combine Karpenko with the Bell-Shi combination. (EX1003, ¶266.) Both Bell and Shi emphasize the importance of computing efficiency. (EX1003, ¶266.) For example, Bell discloses that its method provides advantages over existing prior art, including “requir[ing] much less processing power” and eliminating unnecessary steps. (EX1006, 6:47-

50) Additionally, as Bell’s inventions relate to “[r]eal-time video stabilization for mobile devices,” whose computing power is limited, a POSITA would have understood the importance of efficiency, so that “stabilization [can] be done on the device without inducing any lag.” (EX1006, 6:42-44; EX1003, ¶267.) Similarly, Shi discloses that its video stabilization techniques “can be performed in a manner that is computationally feasible and power efficient on a battery-powered mobile electronic device, such as a mobile phone.” (EX1007, 5:32-35.)

Therefore, a POSITA would have been motivated to turn to a prior art reference that discloses the details of improving the efficiency of video stabilization. (EX1003, ¶268.) Given that Karpenko’s coarse to fine smoothing approach improves the efficiency of calculating the smoothed trajectory by allowing “significantly fewer iterations at each step” to be run, a POSITA therefore would have been motivated to adopt Karpenko’s efficient approach to smoothing the trajectory. (EX1003, ¶269.) The combination would have a reasonable expectation of success because the iterative smoothing disclosed by Karpenko is constraint-based like Bell. (EX1003, ¶270.)

3. Claim 20

Claim 20 mirrors claim 4. (EX1003, ¶271.)

C. Ground 3: Bell, Shi, and Bell Paper Render Obvious Claims 6, 7, and 21

1. Summary of Bell Paper

Bell Paper was published in *ECCV (4)*, Volume 8692 of *Lecture Notes in Computer Science* in 2014 and made publicly available in 2015, qualifying as prior art to the '840 patent at least under § 102(a)(1). (EX1004, EX1005, EX1010.) The Examiner never found, and Applicant never cited, Bell Paper during prosecution.

Bell Paper discloses “a method for video stabilization . . . for videos captured on mobile devices.” (EX1004, 294.) Bell Paper discloses “us[ing] the gyroscope to track the camera motion” and “model[ing] the camera motion as a rotation in a global coordinate frame.” (EX1004, 297.) Bell Paper discloses that the stabilized video produced with its method is based on a smoothed camera path, which “is nearly free of rotational motions” (EX1004, 304.)

2. Claim 6: “The image capture system of claim 1, wherein the smoothed trajectory is determined further based on minimization of a value, a score, or a metric that expresses an amount of high frequencies while respecting a crop constraint.”

A value . . . that expresses an amount of high frequencies. The '840 patent describes that “the smoothed trajectory . . . may be determined based on minimization of a value, a score, and/or a metric that expresses the amount of high frequencies”; the “value/score/metric may be determined based on an objective function that consists of groups of terms where one term expresses the L2-norm of interframe *angular velocity* and the other term expresses the *angular*

acceleration.” (EX1001, 9:31-37.) Therefore, according to the ’840 patent, angular velocity and angular acceleration can be examples of a value, score, or metric that expresses an amount of high frequencies. (EX1003, ¶275.)

Bell discloses that “[m]otion may be described, for example, by ***angular velocity*** [and] ***acceleration***” (EX1006, 6:16-18.)

Bell paper also discloses “us[ing] the data from an on-board gyroscope to track the camera’s ***angular velocity***” and measuring “***angular acceleration.***” (EX1004, Abstract, 294, 304.)

Minimization of the value. Bell Paper discloses that “[w]hen the camera is nearly still, [its] virtual camera is fixed, removing all jitter.” (EX1004, 296.) Removing all jitter when the camera is still means that the rotational velocity is minimized. (EX1003, ¶278.)

Bell Paper further discloses that its method “removes most of the ***angular acceleration***, producing an output video which is pleasing to view.” (EX1004, 304.) Therefore, Bell and Bell Paper disclose determining a smoothed trajectory based on minimizing angular velocity and angular acceleration. (EX1003, ¶¶278-279.)

Respect a crop constraint. According to the ’840 patent, the “crop constraint may require the viewing window for punchout of the visual content to remain within the field of view of the visual content.” (EX1001, 9:45-47.) Bell and Bell

Paper disclose respecting a crop constraint. (EX1003, ¶¶280-281.) Bell discloses using a “crop polygon [that] is smaller than the input size” and “avoid[ing] letting the crop polygon go out of frame.” (EX1006, 9:10-25.) Similarly, Bell Paper discloses that its method “keep[s] the crop window inside the input frame” and ensures that “the crop window never leaves the input frame.” (EX1004, 296.)

Combination. It would have been obvious to combine Bell, Shi, and Bell Paper. Bell and Bell Paper have the same inventors or authors and are directed to similar subjects. (EX1003, ¶282.) A POSITA consulting Bell would have been motivated to turn to Bell’s related publication, Bell Paper, and would have reasonably expected that combining Bell and Bell Paper would be successful. (EX1003, ¶282.) Furthermore, as Bell discloses that the “goal of the smoothing method is to define a new virtual camera that stays static” or “follow[] the original camera with smooth changes in angular velocity,” a POSITA would have been motivated to combine it with its related publication Bell Paper to minimize angular velocity and acceleration, so that the resulting smoothed trajectory is smoother. (EX1006, 8:52-56; EX1003, ¶283.)

3. **Claim 7: “The image capture system of claim 6, wherein the smoothed trajectory is determined further based on iterative minimization, the iterative minimization including iterations of steps that alter latest estimate, the steps including an update step that updates a current estimate to decrease a value of an objective function and a projection step that applies one or more constraints, the current**

estimate being re-projected based on the one or more constraints not being satisfied.”

As shown for Claim 6, Bell and Bell Paper disclose determining a smoothed trajectory based on minimizing angular velocity. (*Supra* § IV.C.2.) Per Bell, virtual angular velocity represents “the discrete angular changes from” frame to frame. (EX1006, 8:58-61.) “Since the frame rate is constant, time is implicit in this representation of the velocity.” (EX1006, 8:61-63.) Thus, for example, virtual angular velocity from frame k to frame $k+1$ can be calculated as follows: $v_{\Delta}(k) = v(k+1) - v(k)$. (EX1003, ¶286.) Therefore, virtual angular velocity is the value of an objective function. (EX1003, ¶¶289, 292, 296.)

Bell further discloses iterations of steps to minimize the virtual angular velocity, as shown below. (EX1003, ¶¶284-306.)

Update step that updates a current estimate to decrease a value of an objective function. Bell discloses that “[a]s a first step, . . . the present invention hypothesize[s] a new orientation for the virtual camera” (EX1006, 9:1-3.) For example, “hypothetical camera orientation $\hat{v}(k)$ ” at frame k “is . . . determine[d]” (EX1006, 9:6-8.) Bell therefore discloses determining a current estimate of the trajectory of the virtual camera, *i.e.*, smoothed trajectory. (*Supra* § IV.A.5.b.)

As Bell discloses that it aims “to allow the motion decay toward zero” and to

“biase[] the virtual camera toward remaining still when possible,” (EX1006, 9:26-29), a POSITA would have understood that the current estimate of the smoothed trajectory decreases virtual angular velocity, which is the value of an objective function. (EX1003, ¶¶290-298.)

Projection step that applies one or more constraints, the current estimate being re-projected based on the one or more constraints not being satisfied. As discussed for claim 6, the crop constraint includes the requirement that the crop polygon stay within the borders of the captured image (input frame). (*Supra* § IV.C.2.)

Bell discloses calculating the location of the crop polygon based on the estimated smoothed trajectory. (EX1003, ¶¶299-301.) For example, Bell discloses “determine[ing] the coordinates of the corners of the resulting crop polygon” using the “hypothetical camera orientation.” (EX1006, 9:6-10.) If “any corner of the crop polygon would fall outside the input frame,” Bell “caus[es] the virtual camera to keep up with the physical camera.” (EX1006, 10:3-7.) Specifically, Bell discloses updating $v_{\Delta}(k)$, which represents the “discrete angular change[] from frame k to frame $k+1$.” (*Supra* § IV.A.3.m; EX1003, ¶302.) Because the movement of the virtual camera is changed to “keep up with the physical camera,” Bell discloses re-projecting the estimated smoothed trajectory. (EX1003, ¶303.)

Iterative steps of update and projection. Bell discloses that “[f]or each new

frame,” its inventions perform the same smoothing steps. (EX1006, 8:61-67.) A POSITA would have understood that Bell calculates a hypothetical camera orientation for each frame. (EX1003, ¶304.) Therefore, Bell discloses iteratively updating the estimated smoothed trajectory. (EX1003, ¶304.)

A POSITA would also have understood that for each frame, Bell checks if “any corner of the crop polygon would fall outside the input frame” and updates the estimated motion of the virtual camera if that is the case. (EX1003, ¶305.)

Additionally, it would have been obvious to a POSITA to perform iterative minimization steps so that a different minimum value of the objective function is achieved with more desirable properties. (EX1003, ¶¶287-298.) Accordingly, each calculation of a different minimum value of the objective function will result in a projection step to see if the resulting crop polygon stays within the input frame, which is the iterative steps of projection. (EX1003, ¶¶287-298.)

Bell therefore discloses iterative steps of estimating the smoothed trajectory, including steps of updating the estimate and projecting to see if the resulting crop polygon stays within the input frame. (EX1003, ¶306.)

4. Claim 21

Claim 21 mirrors claim 7. (EX1003, ¶307.)

D. Ground 4: Bell, Shi, and Sokeila Render Obvious Claim 12

1. Summary of Sokeila

Sokeila was filed on April 16, 2013 and issued on March 31, 2015, qualifying as prior art to the '840 patent at least under §§ 102(a)(1) and 102(a)(2). The Examiner never found, and Applicant never cited, Sokeila during prosecution.

Sokeila discloses “methods, apparatuses, and software related products” for “video stabilization using motion adaptive cropping.” (EX1008, 2:41-44.) The cropping is “based at least on a future prediction of motion in a frame” to “make the cropping factor change smoothly over time.” (EX1008, 2:48-56.)

2. Claim 12: “The image capture system of claim 11, wherein the buffer includes a circular buffer, and the visual content are stored in the circular buffer as video frames to be stabilized.”

As discussed for claim 11, Bell discloses, or renders obvious, storing captured images in a buffer during the video stabilization operation. (*Supra* § IV.A.10.) A POSITA would have known that there are multiple widely known types of buffer, including linear buffer and ring buffer (also known as circular buffer). (EX1003, ¶310.) Given there are finite types of buffers, it would have been obvious to try to use a known option, such as a ring buffer, with a reasonable expectation of success. (EX1003, ¶310.) Additionally, it would have been obvious to apply a known type of buffer, a ring buffer, to a known image capture device to yield predictable results. (EX1003, ¶311.)

Further, as Bell does not discuss what type of buffer to use, a POSITA

would have been motivated to turn to a prior art reference that discloses the specific type of buffer to use for storing video frames. (EX1003, ¶311.)

Sokeila, like Bell, relates to video stabilization using cropping and “future prediction.” (EX1008, 2:50.) Sokeila discloses a “*ring buffer 14* [] configured to hold a certain number of frames and may be continuously or periodically or repeatedly updated with new frames.” (EX1008, 4:20-22.) It would have been obvious to combine Bell’s video stabilization techniques with the known ring buffer disclosed in Sokeila. (EX1003, ¶314.)

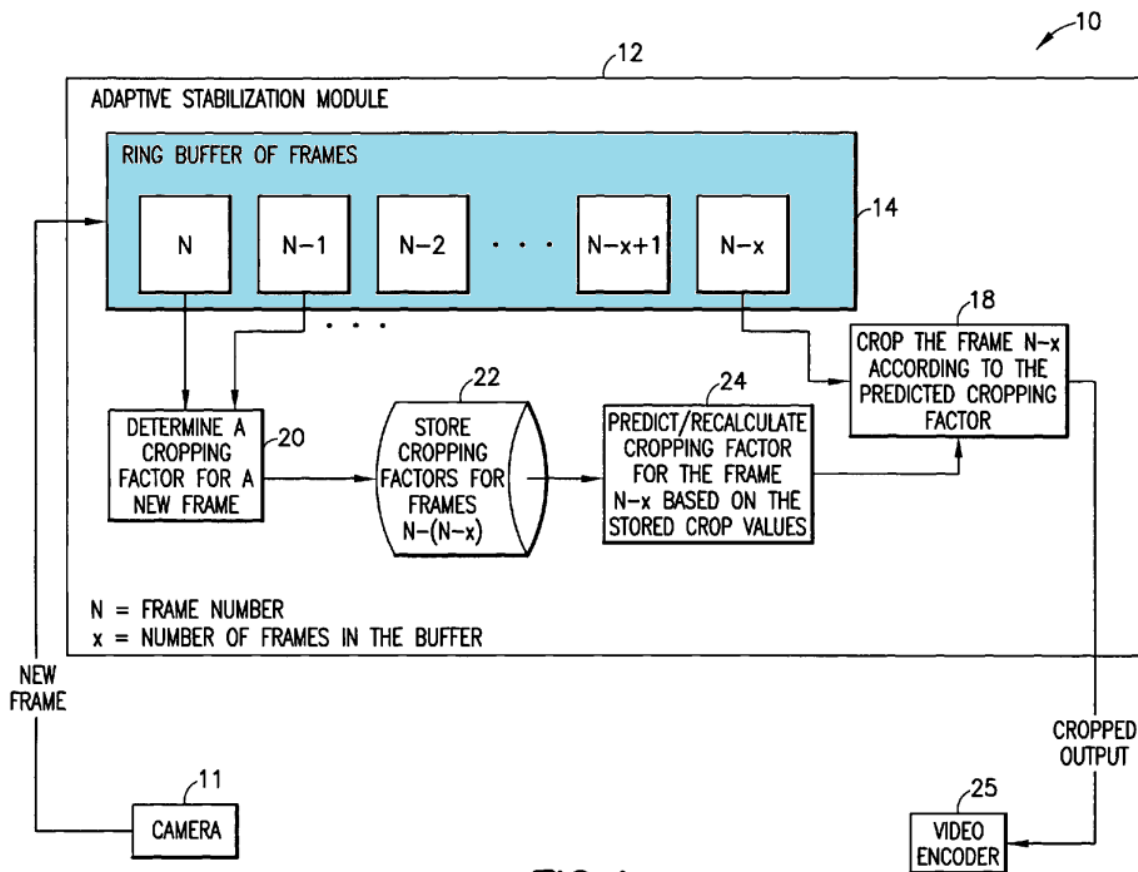


FIG. 1

E. Ground 5: Bell, Shi, and Kwatra Render Obvious Claim 13

1. Summary of Kwatra

Kwatra was filed on February 8, 2011 and issued on September 10, 2013, qualifying as prior art to the '840 patent at least under §§ 102(a)(1) and 102(a)(2). The Examiner cited but never applied Karpenko during prosecution.

Kwatra discloses “[m]ethods and systems for processing a video for stabilization” (EX1012, Abstract.) Kwatra discloses that its techniques “attempt to render the recorded video as the video would have been recorded from a smooth or stable camera path.” (*Id.*, 1:37-40.) Kwatra further discloses “determining a modified motion camera path of the original motion by considering velocity and acceleration. (*Id.*, 1:52-57.)

2. Claim 13: “The image capture system of claim 1, wherein the weight-balance parameter controls types of motion that are minimized in smoothed trajectory generation, the types of motion including rotational velocity and rotational acceleration.”

Kwatra discloses this claim under GoPro’s infringement theory. GoPro alleges that the accused products include the “weight-balance parameter” because “angular velocity and angular acceleration are considered as part of the process for smoothing videos.” (EX1014, at 14.)

The types of motion. Kwatra discloses “determining a modified motion camera path of the original motion path of the camera including for each time *t* the substantially constant path, the substantially constant velocity, or the substantially

constant acceleration of the original motion path of the camera.” (EX1012, 1:52-57.) A POSITA would have understood that the camera motion include rotational movement, and the velocity and acceleration include rotational velocity and rotational acceleration. (EX1003, ¶¶317-319.) This is because Kwatra discloses that “[i]mage stabilization techniques may compensate for . . . *angular* movement[] of a camera” (EX1012, 1:17-19.) Additionally, “constraints can be imposed on the optimal camera path” such that “a range of change in zoom and *rotation*” can be limited. (EX1012, 16:13-32.)

The weight-balance parameter. Kwatra discloses that a “smooth camera path can be estimated using minimization of derivatives of the original camera path” (EX1012, 5:64-66.) For example, “a constant path [may] represent[] a static camera,” denoted as: $\frac{dP}{dt} = 0$; “a path of constant velocity [may] represent[] a panning or a dolly shot,” denoted as: $\frac{d^2P}{dt^2} = 0$; and “a path of constant acceleration [may] represent[] ease-in and out transitions between static and panning cameras,” denoted as: $\frac{d^3P}{dt^3} = 0$. (EX1012, 13:23-48.)

$\frac{dP}{dt}$ represents velocity while $\frac{d^2P}{dt^2}$ represents acceleration. (EX1003, ¶322.)

Constant path means zero velocity, and constant velocity means zero acceleration. (*Id.*)

Kwatra further discloses “determine[ing] a stable camera path” by

minimizing the below “cost function”:

$$a \left| \frac{dP}{dt} \right| + b \left| \frac{d^2P}{dt^2} \right| + c \left| \frac{d^3P}{dt^3} \right|$$

Weights a, b, and c determine how much the stabilization techniques control the path, velocity, and acceleration being constant, respectively. (EX1012, 13:66-14:37.)

Therefore, weights a and b are parameters that control the minimization of rotational velocity and rotational acceleration. (EX1003, ¶¶323-324.) A POSITA would have understood that one parameter can represent both weights a and b, as Kwatra discloses that in some scenarios, “a=b.” (EX1012, 14:31-32; EX1003, ¶324.) That parameter is the claimed “weight-balance parameter.” (EX1003, ¶325.)

F. Ground 6: Bell, Shi, and Corey Render Obvious Claim 15

1. Summary of Corey

Corey was filed on August 15, 2011 and published on February 21, 2013. It qualifies as prior art to the '840 patent at least under § 102(a)(2). The Examiner never found, and Applicant never cited, Corey during prosecution.

2. Claim 15: “The image capture system of claim 1, wherein the stickiness parameter controls an extent by which a preceding portion of the trajectory impacts determination of the smoothed trajectory.”

Corey discloses a stickiness parameter: M, the number of previously

captured frames used to estimate motion. (EX1003, ¶¶326-334.)

Corey discloses that “the motion of a frame may be characterized with respect to a specified number of ‘neighbor’ frames” (EX1011, [0021].)

Corey additionally discloses that “the motion of a current frame (F_c) captured at time t_d may be characterized by M number of previously captured frames . . . and N number of subsequently captured frames” (*Id.*) For example, M can be 3 frames (“captured at prior times t_a , t_b , and t_c ”) and N can be 3 frames (“captured at later times t_e , t_f , and t_g ”). (*Id.*) Corey further discloses that “the choice of 3 frames before and 3 frames after the current frame is a design choice and may vary from implementation to implementation” (*Id.*)

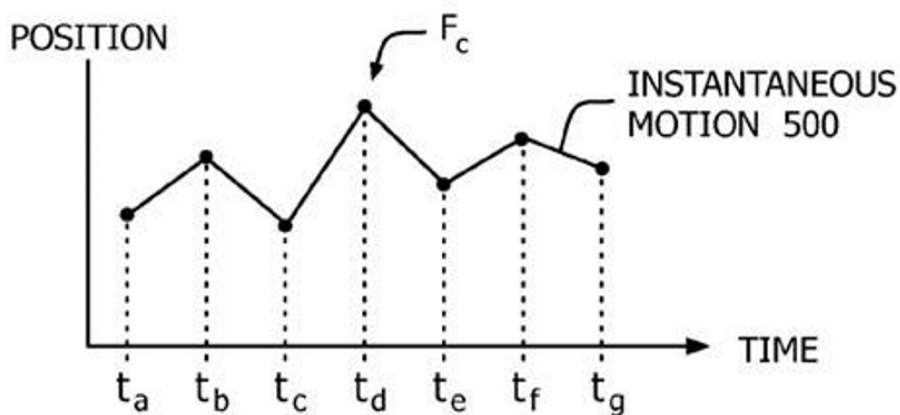


FIG. 5A

Corey discloses that after “the motion of a frame” is characterized using $M + N$ number of frames, the unwanted motion may be estimated. (EX1011, [0022].) Unwanted motion, such as “motion introduced by shaky hands or walking,” can

therefore be removed. (EX1011, [0002, 0004].) After “an estimate of the unwanted motion for a frame has been determined,” “a perspective transformation” is generated and “applied to modify or compensate for the frame’s estimated unwanted motion” (EX1011, [0023].) “The result is stabilized video sequence.” (*Id.*)

The trajectory after the transformation to compensate for the unwanted motion is the smoothed trajectory. (EX1011, [0022]; EX1003, ¶331.)

Furthermore, as M determines how many past frames are considered for the determination of the smoothed trajectory, it controls the extent by which the past frames (*i.e.*, “a preceding portion of the trajectory”) impacts determination of the smoothed trajectory. (EX1003, ¶332.) For example, if M is 0, then no past frames are considered and therefore past frames have no impact on the determination of the smoothed trajectory. (EX1003, ¶333.) Conversely, if M is equal to the number of all past frames, then past frames have a substantial impact on the determination of the smoothed trajectory. (EX1003, ¶333.)

It would have been obvious to combine Bell, Shi, and Corey. (EX1003, ¶¶335-338.) Like Bell and Shi, Corey is directed to videos captured by personal electronic devices, which cause unwanted motion of the camera, including motion introduced by shaky hands or walking. (EX1011, [0002]; *supra* § IV.A.3.1.) It would have been obvious to use the known motion estimation technique in Corey

to improve Bell and Shi's similar video stabilization techniques for mobile devices. (EX1003, ¶¶335-338.) The combination would have had a reasonable expectation of success, especially when Bell and Shi already consider the motion of past frames. (EX1006, 8:64-67; EX1007, 12:14-22; EX1003, ¶¶337-338.)

G. Ground 7: Bell, Shi, Kwatra, and Corey Render Obvious Claim 17

1. Claim 17

a. Preamble: "The method of claim 16, wherein:"

To the extent the preamble is limiting (which Petitioner does not concede), Bell and Shi disclose the preamble of claim 17. (*Supra* § IV.A.12.)

b. 17[a]: "the weight-balance parameter controls types of motion that are minimized in smoothed trajectory generation, the types of motion including rotational velocity and rotational acceleration";

Element 17[a] mirrors claim 13. (EX1003, ¶340.)

c. 17[b]: "the low-light high-pass parameter controls amount or strength of smoothing performed in smoothed trajectory generation and an extent of motion blur compensation applied in determination of the smoothed trajectory; and"

Element 17[b] mirrors claim 14. (EX1003, ¶341.)

d. 17[c]: "the stickiness parameter controls an extent by which a preceding portion of the trajectory impacts determination of the smoothed trajectory."

Element 17[c] mirrors claim 15. (EX1003, ¶342.)

V. THE BOARD SHOULD NOT DENY INSTITUTION UNDER *FINTIV*

“The plain language of the *Fintiv* factors is directed to district court litigation and does not apply to parallel U.S. International Trade Commission (ITC) proceedings.” (USPTO Interim Procedure for Discretionary Denials, 2-3 (June 21, 2022).)

VI. THE BOARD SHOULD NOT EXERCISE ITS DISCRETION TO DENY INSTITUTION UNDER § 325(D)

Five of the seven references presented in this Petition—Bell, Shi, Sokeila, Bell Paper, and Corey—were not cited during prosecution of the ’840 patent. And while Karpenko’s related publication and Kwatra were cited during prosecution, the Examiner never substantively discussed or applied them against the claims.

VII. NOTICES AND STATEMENTS

Pursuant to 37 C.F.R. § 42.8(b)(1), Arashi Vision Inc. (d/b/a Insta360) is the real party-in-interest for Petitioner.

Pursuant to 37 C.F.R. § 42.8(b)(2), Petitioner identifies the following related matters:

- *In the Matter of Certain Cameras, Camera Systems, and Accessories Used Therewith*, Inv. No. 337-TA-1400 (USITC).

- *GoPro, Inc. v. Arashi Vision Inc. et al.*, No. 8:24-cv-00681-HDE-JCV
(C.D. Cal.).⁸

Pursuant to 37 C.F.R. § 42.8(b)(3), Petitioner identifies the following counsel (and a power of attorney accompanies this Petition).

Lead Counsel	First Backup Counsel
Alex S. Yap Registration No.: 60,609 MORRISON & FOERSTER LLP 707 Wilshire Blvd., Suite 6000 Los Angeles, CA 90017 Email: AYap@mofo.com Tel: 213-892-5200 Fax: 213-892-5454	Mehran Arjomand Registration No.: 48,231 MORRISON & FOERSTER LLP 707 Wilshire Blvd., Suite 6000 Los Angeles, CA 90017 Email: MArjomand@mofo.com Tel: 213-892-5200 Fax: 213-892-5454
Second Backup Counsel	
W. Stella Mao <i>pro hac vice</i> to be filed MORRISON & FOERSTER LLP 755 Page Mill Road Palo Alto, CA 94304-1018 Email: SMao@mofo.com Tel: (650) 813-5600 Fax: (650) 494-0792	

Pursuant to 37 C.F.R. § 42.8(b)(4), service information for lead and back-up counsel is provided above. Petitioner consents to electronic service by email to

⁸ IPR2024-00984 and IPR2024-01052 also are related to the ITC and C.D. Cal. litigations.

IPR2024-01164
Patent 10,958,840 B2

MoFo_Insta360-IPRs@mofo.com.

Pursuant to 37 C.F.R. § 42.104(a), Petitioner certifies that the '840 patent is available for *inter partes* review and that Petitioner is not barred or estopped from requesting an *inter partes* review challenging the patent claims on the grounds identified in this Petition.

The USPTO is authorized to charge any required fees, including the fee as set forth in 37 C.F.R. § 42.15(a) and any excess claim fees, to Deposit Account No. **03-1952** referencing Docket No. **33316-00000.02.**

Dated: July 12, 2024

Respectfully submitted,

/s/ Alex S. Yap

Alex S. Yap

Reg. No.: 60,609

MORRISON & FOERSTER LLP

707 Wilshire Blvd., Suite 6000

Los Angeles, CA 90017

Email: AYap@mofo.com

Tel: 213-892-5200

Counsel for Petitioner

Certification of Word Count (37 C.F.R. § 42.24)

I hereby certify that this Petition for *Inter Partes* Review has 13,949 words (as counted by the “Word Count” feature of the Microsoft Word™ word-processing system), exclusive of “a table of contents, a table of authorities, mandatory notices under § 42.8, a certificate of service or word count, or appendix of exhibits or claim listing.”

Dated: July 12, 2024

/s/ Alex S. Yap
Alex S. Yap

IPR2024-01164
Patent 10,958,840 B2

Certificate of Service (37 C.F.R. § 42.6(e)(4) and 42.205(a))

I hereby certify that the attached Petition for *Inter Partes* Review and supporting materials were served as of the below date by UPS, which is a means at least as fast and reliable as U.S. Express Mail, on the Patent Owner at the correspondence address indicated for U.S. Patent No. 10,958,840:

GoPro Issuance
3025 Clearview Way
San Mateo, CA 94402

Dated: July 12, 2024

/s/ Alex S. Yap
Alex S. Yap