

PUBLIC VERSION

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

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

v.

GOPRO, INC.,
Patent Owner

Case IPR2024-01164
U.S. Patent No. 10,958,840

**DECLARATION OF HASSAN FOROOSH PH.D.
IN SUPPORT OF PATENT OWNER'S RESPONSE**

PUBLIC VERSION

1. I have been retained on behalf of GoPro, Inc. (“Patent Owner”) in connection with the above-captioned *inter partes* review (IPR). I have been retained to provide my opinions in support of Patent Owner’s Response. I am being compensated for my time at my standard rate of \$650 per hour. I have no interest in the outcome of this proceeding.

2. In preparing this declaration, I have reviewed and am familiar with the Petition for IPR2024-01164, U.S. Patent No. 10,958,840 (“the ’840 patent”) and its file history, and all other materials cited and discussed in the Petition (including the declaration of Petitioner’s expert Dr. Aggelos Katsaggelos), and cited and discussed in this Declaration.

3. The statements made herein are based upon my own knowledge and opinion. This Declaration represents only the opinions I have formed to date. I may consider additional documents as they become available or other documents that are necessary to form my opinions. I reserve the right to revise, supplement, or amend my opinions based on new information and on my continuing analysis.

I. QUALIFICATIONS

4. I have been retained to serve as a technical expert in this matter based upon my educational and research experience, and specifically my work in the fields of computer vision, image processing, and optics. My curriculum vitae

PUBLIC VERSION

(“CV”) submitted as Exhibit 2002, which contains details about my education, experience, publications, and other qualifications to render an expert opinion in this matter.

5. I am a Professor of Computer Science at the University of Central Florida. My primary research interests include computer vision, image/video processing, and machine learning, including deep learning and large language models.

6. I received my Ph.D. in Computer Science from INRIA (French National Institute for Research on Informatics and Control), University of Nice Sophia Antipolis in 1996. I also earned a Master of Science in Computer Science from the University of Nice Sophia Antipolis in 1993, a Master of Science in Electrical Engineering from Kingston University, U.K., in 1990, and a Bachelor of Science in Electrical Engineering from Kingston University, U.K., in 1989.

7. Prior to research and teaching, I worked as an R&D Engineer at Innovative Engineering Solutions Ltd. in London from 1990 to 1992. Between 1993 and 1996, I served as a Research Assistant in the Scene Analysis and Symbolic Image Processing Group of INRIA, where I performed research in areas such as 3D super-resolution, sub-pixel image registration, and image-based depth and shape recovery and analysis. From 1997 to 2000, I served as an Assistant

PUBLIC VERSION

Research Scientist at the University of Maryland, College Park, where I investigated topics such as image registration (2D and 3D), super-resolution of video data for unmanned aerial vehicles, super-resolution of Synthetic Aperture Radar (SAR) and Foliage Penetrating SAR (FOPEN-SAR), and Automatic Target Detection, Tracking, and Recognition (ATD/R). During this time, I authored one of the most cited works on sub-pixel registration and alignment based on phase correlation: (Hassan Foroosh (Shekarforoush), Josiane Zerubia and Marc Berthod, Extension of Phase Correlation to Sub-pixel Registration, IEEE Transactions on Image Processing, vol. 11, Issue 3, pp. 188-200, 2002). Between 2000 to 2002, I served as Senior Research Scientist at the University of California, Berkeley where I also served as the Technical Coordinator and Principal Research Scientist of a Multi-University Research Initiative (MURI) funded by the Office of Naval Research. As part of the project, I conducted research on multi-sensor acquisition of 3D data for autonomous 3D modeling and visualization of urban environments.

8. I joined the Department of Computer Science at the University of Central Florida (UCF) as an Assistant Professor in 2002. Since then, I have served as an Associate Professor, Professor, and CAE Link Professor of Computer Science in the same department. I also served in various other capacities within the department, such as the Graduate Program Coordinator, chair of various leadership

PUBLIC VERSION

committees, and most recently in the Faculty Senate Strategic Planning Council for UCF.

9. I have taught a number of courses directed to computer vision and 3D modeling, including 3D Computer Vision, Image Processing, Special Topics in 3D Computer Vision, and Visual Simulation. Since joining UCF, I have also advised and chaired the dissertation of over 30 doctoral students, 28 of whom have already successfully defended their doctoral dissertation, and secured academic positions or industry jobs with major high-tech corporations. As one example, I served as the advisor for a dissertation directed to “*Improving Matching and Classification Through Deep Learning of Structure and Varying Illumination*”. This dissertation involved also working with GoPro camera systems for mobile aerial platforms.

10. I have been the principal investigator or co-principal investigator of numerous research projects funded by the National Science Foundation (NSF), NASA, DoD, and industry partners, with total funding exceeding \$14M since joining UCF, and ongoing funded projects for an additional \$3M. As one related example, I recently served as the principal investigator in two projects funded by the National Science Foundation (NSF) on 3D modeling and camera systems, and the role of 3D information in machine understanding of digitized 3D world.

PUBLIC VERSION

11. I have also conducted extensive research in partnership with industry in the field of computer vision, in particular 3D computer vision and camera systems, including body-worn cameras and cameras on agile mobile platforms, or cameras operating in various different modalities (EO, FLIR, polarized, etc.). Below, I highlight some of the projects that were funded by industry and state-sponsored technology transfer initiatives that specifically involved using GoPro camera systems or similar platforms. For example, from 2012 to 2018, I received funding from Jackson Technologies related to Airborne Multisensor Visual Data Processing, which studied automated methods of reconstructing 3D panoramic models around a flying aircraft that used the GoPro camera system and software tools extensively; From 2013 to 2018, I received related funding from the Florida High Tech Corridor (“FHTC”) for Rapid Automated 3D Terrain Modeling from Airborne Data, which similarly studied automated methods of reconstructing 3D models of a terrain from a flying aircraft, again using extensively the GoPro cameras and software tools; From 2008 to 2009, I received funding from Electronic Arts – Tiburon related to Modeling and Tracking Human Activities in Sport Events. The goal of this project was to develop algorithms for tracking and recognition of human activities in sport events, using body-worn cameras similar

PUBLIC VERSION

to GoPro systems, as well as surrounding cameras, network of cell phone cameras, and other sensors.

12. I have authored or co-authored over 200 articles in peer reviewed journals or conference proceedings, relating to the areas of Computer Vision (CVPR, ICCV, ECCV, etc.), Machine Learning (ICLR, AAAI, ACL, EMNLP, etc.), and Image Processing and Pattern Recognition (ICIP, ICPR). Some related example publications include: Yang Zhang, Zixiang Zhou, Philip David, Xiangyu Yue, Zerong Xi, Boqing Gong, Hassan Foroosh, PolarNet: An Improved Grid Representation for Online LiDAR Point Clouds Semantic Segmentation, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020; Chuan Sun, Imran Junejo, Marshall Tappen, Hassan Foroosh, *Exploring Sparseness and Self-Similarity for Action Recognition*, IEEE Transactions on Image Processing, vol. 24, issue 8, pp. 2488-2501, 2015; Chen Shu, Luming Liang, Wenzhang Liang, Hassan Foroosh, *3D Pose Tracking with Multi-Templates Warping and SIFT Correspondences*, IEEE Trans. on Circuits and Systems for Video Technology, 2015; Nazim Ashraf, Chuan Sun, and Hassan Foroosh, *View-Invariant Action Recognition Using Projective Depth*, Journal of Computer Vision and Image Understanding (CVIU), vol. 123, pp. 41-52, 2014; Yuping Shen and Hassan Foroosh, View-invariant Action Recognition from Point

PUBLIC VERSION

Triplets, IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 31, no. 10, pp. 1898-1905, 2009; Nazim Ashraf, Yuping Shen, Xiaochun Cao, and Hassan Foroosh, *View-Invariant Action recognition using Weighted Fundamental Ratios*, Journal of Computer Vision and Image Understanding (CVIU), vol. 117, pp. 587-602, 2013.

13. Based on my work, I have also applied for and received several patents in the field of image processing, including U.S. Pat. No. 8,755,569 entitled Methods for Recognizing Pose and Action of Articulated Objects with Collection of Planes and Motion, which was issued in 2014.

14. I have served as the editor and/or committee member on numerous professional organizations. For instance, I have been on the editorial board of the IEEE Transactions on Image Processing (Associate Editor, 2003-2008, 2011-2016); IEEE Conference on Computer Vision and Pattern Recognition (CVPR) serving as area chairs, session chairs and Technical Committee member over many years (2004-2012, 2015-2016, 2019-2023); Organizing Committee and sponsorship chair of ACM Multimedia Conference in 2014; and frequent reviewer on various IEEE transactions, including the IEEE Transactions on Pattern Analysis and Machine Intelligence.

PUBLIC VERSION

15. For my work, I have been honored with numerous awards, including the *Piero Zamperoni Award*, International Association for Pattern Recognition (IAPR) (2004). I have also received several Best Paper awards, including “*The Best Scientific Paper Award*” in the IAPR International Conference on Pattern Recognition in 2008, “*The Best Paper Award*” in the IEEE International Conference on Image Processing (ICIP) in 2018, and the *Distinguished Researcher*, UCF College of Engineering and Computer Science (2004-2005, 2005-2006).

16. In sum, I have studied, taught, practiced, and conducted research in the areas of computer vision, image processing, machine learning (deep learning and large language models), and optics.

II. LEGAL UNDERSTANDING**A. The Person of Ordinary Skill in the Art (“POSITA”)**

17. I understand that a person of ordinary skill in the relevant art at the time of the invention (also referred to herein as “ordinary artisan”) is presumed to be aware of all pertinent art, thinks along conventional wisdom in the art, and is a person of ordinary creativity—not an automaton.

18. I have been asked to consider the level of ordinary skill in the field that someone would have had at the time the claimed invention was made. In deciding the level of ordinary skill, I considered the following:

PUBLIC VERSION

- the levels of education and experience of persons working in the field;
- the types of problems encountered in the field; and
- the sophistication of the technology.

19. I have reviewed the declaration of Dr. Katsaggelos, including his opinions regarding the Person of Ordinary Skill in the Art, which is a person with “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” or alternatively “four years of experience working on image or video processing, or digital cameras.” Ex. 1003, ¶¶ 31-33.

20. My opinions set forth in this Declaration apply Dr. Katsaggelos’s definition of a POSITA.

21. I am well-qualified to determine the level of ordinary skill in the art. I was a person of at least ordinary skill in the art at the time of the September 2018 provisional application priority date of the ’840 patent under Dr. Katsaggelos’s definition of a POSITA. If I do not explicitly state that my statements below are based on this timeframe, all of my statements are to be understood as an ordinary artisan would have understood something as of the priority date of the ’840 patent.

PUBLIC VERSION**B. Legal Principles**

22. I am not a lawyer and will not provide any legal opinions. Though I am not a lawyer, I have been advised that certain legal standards are to be applied by technical experts in forming opinions regarding the meaning and validity of patent claims.

Anticipation

23. I understand that to obtain a patent, a claimed invention must have, as of the priority date, been novel. I understand that to prove that a prior art reference anticipates a patent, the four corners of the prior art reference must describe element of the claimed invention, either expressly or inherently. I also understand that the prior art reference must describe the elements arranged as in the claim and that it is improper to combine disparate embodiments to establish anticipation unless a POSITA when reading the reference could at once envisage such a combination.

Obviousness

24. I understand that to obtain a patent, a claimed invention must also have, as of the priority date, been nonobvious in view of prior art in the field. I understand that an invention is obvious when the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a

PUBLIC VERSION

whole would have been obvious at the time the invention was made to a person having ordinary skill in the art.

25. I understand that to prove that prior art, or a combination of prior art, renders a patent obvious, it is necessary to: (1) identify the particular references that singly, or in combination, make the patent obvious; (2) specifically identify which elements of the patent claim appear in each of the asserted references; and (3) explain how the prior art references could have been combined to create the inventions claimed in the asserted claim, including how there would have been a reasonable expectation of success.

26. I understand that a patent composed of several elements is not proved obvious merely by demonstrating that each of its elements was, independently, known in the prior art, and that obviousness cannot be based on the hindsight combination of components selectively culled from the prior art to fit the parameters of the patented invention.

27. I also understand that a reference may be said to teach away when a person of ordinary skill, upon reading the reference, would be discouraged from following the path set out in the reference, or would be led in a direction divergent from the path that was taken by the applicant. Even if a reference is not found to teach away, I understand its statements regarding preferences are relevant to a

PUBLIC VERSION

finding regarding whether a skilled artisan would be motivated to combine that reference with another reference.

My Understanding of Claim Construction Law

28. I have been informed that patent claims are construed from the viewpoint of a person of an ordinary artisan and that patent claims generally should be understood consistent with their ordinary and customary meaning at the time of the invention. A review of the patent claim language, the patent specification, and its prosecution history are also necessary to determine the proper meaning and scope of the term at issue.

29. I have further been informed that in the specification and prosecution history the patentee may define a claim term in a way that differs from the ordinary and customary meaning. I understand that during prosecution of the patent before the U.S. Patent and Trademark Office, the Applicant may make representations or provide definitions of terms that may affect the scope of the patent claims. In particular, the Applicant may, during the course of prosecution, limit the scope of the claims to overcome prior art and/or disavow claim coverage by making clear and unambiguous statements to that effect.

30. I have been informed that an ordinary artisan may, among other things, consider dictionaries, publications, other patents, and treatises (*i.e.*,

PUBLIC VERSION

“extrinsic evidence”). I understand that extrinsic evidence may generally not be relied on if it contradicts the meaning of claim language provided by the intrinsic evidence, such as express definitions given for terms in the specification.

Secondary Considerations of Non-Obviousness

31. I understand that one of the factual inquiries when determining the obviousness of a patent is to consider objective evidence related to the issue of obviousness, sometimes called “secondary considerations of non-obviousness” or “objective indicia of non-obviousness.” I understand that such evidence can often be probative of the issues of obviousness.

32. I understand that examples of objective evidence include (1) commercial success, (2) copying, (3) industry praise, (4) skepticism, (5) long-felt but unsolved need, and (6) failure of others.

33. I understand that there must be a nexus between the patented invention and the evidence of secondary considerations. I further understand that there is a presumption of such a nexus if the claims of the patent are coextensive with a product or products that are the subject of the secondary considerations evidence. I further understand that in order for a claim to be coextensive with a product, one must look at the claims at the whole and not just the new claimed

PUBLIC VERSION

feature. In addition, the claimed invention should not be merely a subcomponent of the products sold but should encompass the entire apparatus.

III. CLAIM CONSTRUCTION

34. I understand that: in an *inter partes* review, claim terms receive their plain and ordinary meaning in light of the specification of the patent in which they appear. Petitioner states that no constructions are necessary. See Pet. at 5-6. However, a dispute has arisen regarding the meaning of the “low-light high-pass parameter.”

35. Based on the preliminary arguments, the Board understood a “low-light high-pass parameter” as “a parameter that controls the amount or strength of smoothing used in motion blur compensation.” (I.D. at 8-9). While the Board correctly notes that a low-light high-pass parameter is a parameter that controls the amount or strength of smoothing used in motion blur compensation, I understand that it is also important that the claim terms “low-light” and “high-pass” be accorded meaning.

36. In my opinion, Patent Owner’s proposed claim construction set for the below is consistent with the meaning of those terms in view of the specification and the understanding of a POSITA at the time of the invention:

PUBLIC VERSION

Claims term	Patent Owner's Proposed Construction
"low-light high-pass parameter"	A parameter that controls the amount or strength of smoothing used in motion blur compensation by allowing more high frequencies to pass through based on low-light conditions.

37. In my opinion, consistent with the plain claim language, GoPro's "low-light high-pass" parameter is a targeted parameter that selectively addresses blur based on "low-light" conditions while otherwise maintaining smoothing strength. GoPro's parameter also controls smoothing in a specific manner, *i.e.* as a "high-pass" parameter which allows more high frequencies to pass as opposed to simply attenuating the smoothing strength across the entire range of frequencies.

38. In my opinion, GoPro's proposed construction is supported by how the "low-light high-pass" parameter is described in the intrinsic evidence. The specification explains how the "low-light high-pass" parameter addresses an issue that occurs specifically in low-light conditions due to longer exposure times. Ex. 1001 at 18:1-19:59; *see also* at 19:18-21 ("[L]onger exposure time may be used with low light conditions and shorter exposure time may be used with high light conditions."). As the patent explains, camera motion during the exposure period—termed "intra-frame" motion—can become "encrusted" in the video frame, leading to undesirable blur artifacts. *Id.* at 18:1-24; 21:66-22:7. Strong smoothing in this

PUBLIC VERSION

situation can cause the motion blur to become more evident, creating what the patent calls “motion blur pulsations” and “poor visual impression.” *Id.* at 22:2-5.

39. To address this particular cause of motion blur, the specification describes when and how the “low-light high-pass” parameter is utilized. For example, the amount of smoothing/stabilization may vary based on the frequency of motion as well as the exposure time. *Id.* at 18:43-19:26; Figure 7. As a result, “[c]ertain portion[s] of high frequencies in the input may be kept based on the image capture configuration, such as exposure start time and exposure duration time . . .” *Id.* at 17:61-18:1. Thus, the output motion can be controlled to “include high frequencies that mask motion blur within the visual content.” *Id.* at 22:16-21. By controlling smoothing in this targeted manner, the claimed “low-light high-pass” parameter addresses motion blur arising from long exposure times, *id.* at 22:8-21, while still maintaining smoothing strength for other types of motion subject only to the “margin constraint,” *id.* at 21:64-66.

40. GoPro’s proposed construction is also consistent with the extrinsic evidence. There is no dispute that “low-light” conditions, as indicated by longer exposure times, is a possible cause of blur. In his declaration, Petitioner’s expert Dr. Katsaggelos states: “A POSITA at the time of the ’840 patent filing would understand that blurs can be caused by low-light conditions because a shutter is

PUBLIC VERSION

open for a longer time in video capture under low-light conditions, causing movement to add more blur to a frame.” Ex. 1003 at ¶ 111; ¶ 237 (same quote); and ¶ 123 (“as explained above, blur may be caused by longer device shutter integration time under low light conditions.”).

41. “Low-light,” however is just one of many potential causes of blur, as Dr. Katsaggelos admitted at his deposition. Ex. 2020 at 103:22-104:8 (testifying that motion blur can be caused “due to many different reasons”). Given that the term specifically recites a “*low-light*” high pass parameter, GoPro’s proposed construction appropriately distinguishes the claimed parameter from techniques for mitigating blur due to other factors.

42. Patent Owner’s proposed construction is also consistent with the well-understood meaning of “high-pass.” As Dr. Katsaggelos confirmed at his deposition, a “high-pass” system is one that allows high frequencies to pass through the smoothing process while still attenuating the low frequencies. *Id.* at 94:6-15 (Q. “[W]ould a high pass system be one that allows most of the information of the signal at high frequencies to go through while attenuating the information at the low frequency of the signal? . . . THE WITNESS: That would be consistent with the first definition. It’s a matter of what is attenuated and what is allows to go through with some smaller distortion, yeah.”). Further, when teaching

PUBLIC VERSION

his students, Dr. Katsaggelos has described “high-pass” systems as those that “allow the high frequencies to go through the system unchanged.” Ex. 2022 (Coursera screenshot).

43. For all these reasons, in my opinion, the term “low-light high-pass parameter” should be construed to mean: “a parameter that controls the amount or strength of smoothing used in motion blur compensation by allowing more high frequencies to pass through based on low-light conditions.”

I. OVERVIEW OF THE '840 PATENT

44. The '840 patent, titled “Systems and Methods for Stabilizing Videos,” discloses an advanced system designed to stabilize video content in diverse circumstances and conditions. The system measures position information characterizing the rotational positions of the image capture device during the capture duration. This allows the system to generate a “trajectory of the image capture device during the capture duration,” which “reflect[s] the rotational positions of the image capture device at different moments within the capture duration.” Ex. 1001, 1:18-29; 30:18-23 (claim 1); 32:28-33 (claim 16).

45. Among the '840 patent’s key innovations is the introduction of tunable parameters that allow the system to adaptively modulate the stabilization process based on environmental conditions and/or the characteristics of the

PUBLIC VERSION

captured motion. Ex. 1001, 21:17-25. These tunable parameters include the “low-light high-pass parameter” (*id.*, 21:59-22:21), “weight-balance parameter” (Ex. 1001, 17:32-60, 21:17-58), and “stickiness parameter” (*id.*, 22:22-41), which enable fine tuning of the stabilization process under diverse circumstances and conditions. *Id.*, 30:35-38 (claim 1), 32:44-48 (claim 16).

Low-light high pass parameter

46. The low-light high-pass parameter addresses the challenge of stabilizing videos captured in low-light environments by controlling the extent to which the system compensates for motion blur that occurs during frame exposure. Ex. 1001, 21:59-22:21.

47. A central concern addressed by the patent is how to stabilize video in “low-light” conditions, where longer exposure times increase the likelihood of motion blur. The specification describes how camera motion during the exposure period—termed “intra-frame” motion—can become “encrusted” in the video frame, leading to undesirable blur artifacts. *Id.* at 18:1–24; 21:66-22:7. In other words, strong smoothing in this situation creates what the patent calls a “poor visual impression.” *Id.* at 22:2–5. The goal of the invention, therefore, is to stabilize in a way where “intra-frame motion is taken into account so that the

PUBLIC VERSION

smoothed trajectory is smooth for inter-frame phases while corresponding to the original motion for intra-frame phases.” *Id.* at 18:18-20.

48. To mitigate this problem, the specification describes modifying the smoothing strength based on exposure time and other factors such as motion of the image capture device. *Id.* at 18:25–46. For example, in some implementations, a certain portion of “***high frequencies in the input may be kept based on the image capture configuration, such as exposure start time and exposure duration time***” to compensate for “***motion blur and/or lowlight image capture conditions.***” *Id.* at 17:61–18:9.

49. Figure 7 and its accompanying discussion further elaborate on this by illustrating how smoothing values (S1, S2, S3, S4) can be modulated depending on exposure time and frequency of motion within the frame. *Id.* at 18:43-19:59. In particular, the specification explains that “***longer exposure time may be used with low light conditions,***” and that “***higher exposure time may result in greater amount of motion blur.***” *Id.* at 19:18-21; 19:25–26. Additionally, intra-frame motion “higher than or the same as [a] high motion” threshold is considered high-frequency motion to further tune strength of smoothing. *Id.* at 18:64-19:2.

50. Within this context of adaptive smoothing in low-light conditions, the specification introduces the claimed “low-light high-pass parameter” as a

PUBLIC VERSION

parameter that “controls the amount/strength of smoothing performed to generate the smoothed trajectory.” *Id.* at 21:59-61. The specification reiterates the distinction between “inter-frame motion,” which is smoothed subject only to a margin constraint, and “intra-frame motion” which is addressed by a higher value of the low-light high-pass parameter to let the smoothed trajectory follow the observed trajectory more closely. *Id.* at 21:61-22:21. To do so, the parameter reduces the strength of smoothing in a “high-pass” manner—*i.e.* a manner that maintains more high frequency motion while continuing to attenuate low frequency motion. *Id.* at 22:8–25. (“[T]he output motion may include high frequencies that mask motion blur within the visual content.”).

51. Thus, as described in the specification, the low-light high-pass parameter is targeted to a specific issue—stabilizing video captured in low-light conditions, where longer exposure durations create intra-frame blur. It does so by modifying the strength of smoothing in a “high-pass” manner that seeks to preserves high frequency motion, thereby reducing the appearance of blur in this specific circumstance, while attempting to maintain the stabilization strength for low frequency motion in other circumstances.

52. The result is a more natural and visually pleasing video output for GoPro’s specific use cases. As used in the claims, the “low-light high-pass

PUBLIC VERSION

parameter” is therefore best understood as “a parameter that controls the amount or strength of smoothing used in motion blur compensation by allowing more high frequencies to pass through based on low-light conditions.”

Weight-balance parameter

53. The weight-balance parameter provides a mechanism for stabilizing video by determining how much emphasis to place on minimizing rotational velocity versus rotational acceleration. Ex. 1001, 17:32-60, 21:17-58. This allows the system to adapt to different types of motion and produce smoother, more natural video. *Id.*

Rotational Motion of the Camera Housing

54. Rotational motion is typically described using yaw, pitch, and roll angles (“Euler angles”), which represent rotation around three perpendicular axes:

- **Yaw**— Rotation around the **vertical axis** (turning left or right).
- **Pitch**— Rotation around the **lateral axis** (tilting up or down).
- **Roll**— Rotation around the **longitudinal axis** (tilting side to side).

55. While the claims do not explicitly reference **yaw**, **pitch**, or **roll**, minimizing rotational velocity and acceleration of a camera housing—a three-dimensional object in space—necessarily involves at least one of these rotational angles.

PUBLIC VERSION

Minimizing a Combination of Rotational Velocity and Acceleration of the Camera Housing

56. The '840 patent improves video stability by computing a capture trajectory that smooths camera motion through the minimization of a combination of rotational velocity and rotational acceleration of the camera housing. This sort of optimization inherently involves non-linear and trigonometric relationships to account for rotational motion accurately.

57. In the preferred embodiment, a cost function applies weighting factors to different rotational motion derivatives to balance stabilization performance. Specifically, Equation 54 (*id.* at 11:2-22) minimizes both the first and second time derivatives of a function expressed in terms of rotational angle (θ)—These derivatives correspond to rotational velocity and rotational acceleration of the housing:

$$\min \left(\left\| \frac{d}{dt} \theta \right\|^2 + \left\| \frac{d^2}{dt^2} \theta \right\|^2 + \dots \right)$$

where θ represents the yaw, pitch, or roll angle of the camera housing in the 3D space. The two terms of rotational velocity and rotational acceleration in the above equation describe the 3D rotational motion of the camera housing through a set of non-linear relationships (e.g. Euler's equations for angular velocity) that are

PUBLIC VERSION

studied in the kinetics of moving 3D rigid bodies, including in sensor-based tracking of objects. In fact, the non-linearity of such relationships, even when involving only the first-order time-derivatives of the Euler angles (i.e. only the rotational velocity) results in the governing relationships in the form of a set of simultaneous first order non-linear differential equations, that are in general solved only numerically.

58. Minimizing these time derivatives is one example of optimizing rotational velocity and acceleration in accordance with the '840 patent claims. By minimizing rotational velocity and acceleration of the camera housing, the '840 patent accounts for the real-world motion of a camera rotating in the 3D space, ensuring smoother video output.

59. The specification explains that “the smoothed trajectory may be determined based on minimization of a rotational velocity . . . and a rotational acceleration of the image capture device . . . while respecting a set of constraints.” *Id.* at 17:32–37. For example, a “smooth path” of yaw, pitch, or roll angle may be generated by “finding a path of the image capture device . . . that minimizes a combination of a time derivative, a second time derivative, and/or other time derivative(s)” of these angles. *Id.* at 17:37–53.

PUBLIC VERSION

60. The specific combination that is minimized is set by the weight-balance parameter, which “may range between 0 and 1, with 0 corresponding to minimization of angular velocity and 1 corresponding to minimization of angular acceleration.” *Id.* at 21:36–39. Intermediate values allow the system to minimize both types of motion, “with minimization of different types of motion weighed according to the value.” *Id.* at 21:39-42. The patent explains the practical effect of different values:

- Minimizing **only angular velocity** (value = 0) may lead to trajectories with sharp, piecewise-linear corners that, as the patent explains, “prefer[] to keep a steady position when possible” and “minimize apparent motion between video frames”—but the resulting motion “may appear unnatural.” *Id.* at 21:44–50.
- Minimizing **only angular acceleration** (value = 1) may reduce corner points and result in smoother acceleration, but it “may produce more motion than [s]ole minimization of angular velocity.” *Id.* at 21:50–54.
- Minimizing **a combination** (value between 0 and 1) produces a trajectory that blends these two effects, with weights determined by the parameter value. *Id.* at 17:54–58.

61. In this way, the weight-balance parameter gives the system the ability to fine-tune stabilization based on the type of motion encountered, enabling more natural and visually appealing video.

Stickiness parameter

62. The stickiness parameter “control[s] the extent by which preceding portion(s) of the (observed) trajectory impacts determination of the smoothed

PUBLIC VERSION

trajectory.” *Id.* at 22:22–25. The parameter may range between 0 and 1—where a value of 1 may result in a “lazy” stabilizer that “maximizes the impact of the preceding motion/position,” and a value of 0 may result in the smoothed trajectory being generated without taking the preceding motion into account. *Id.* at 22:25–33. Values between 0 and 1 provide a flexible weighting of the preceding motion’s influence, allowing for a dynamic balance between responsiveness and stability. The parameter may also “impact bootstrapping”—that is, the way the system initially estimates the smoothed trajectory based on prior motion. *Id.* at 22:36–42.

63. Together, these parameters reflect a sophisticated, configurable approach to digital video stabilization—one capable of adapting to varied motion types and lighting conditions.

Temporal horizon enabling approximation of intentional motion

64. In addition, unlike conventional approaches that distinguish between stationary and moving cameras to infer the presence of jitter, the ’840 patent’s “*temporal horizon enabling approximation of intentional motion*” is directed to leveraging a duration of future motion to estimate user intent at a given moment, regardless of whether the camera is stationary or in motion.

65. The ’840 patent explains that “intentional motion may include a forward motion to follow an object of interest, a panning motion to shoot a wide

PUBLIC VERSION

view of a scenery, and/or other motion,” while “unintentional motion may result in noise . . . caused by shaking hands, vibrations, bumps.” Ex. 1001 at 9:8-16. To distinguish between these types of motion, the ’840 patent teaches that “intentional motion may be characterized by lower frequency motion than unintentional motion,” and that “a sufficiently long temporal horizon of experienced motion may need to be used” to identify it. *Id.* at 9:17-23.

66. The specification explains that “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.” *Id.* at 15:1-6.

67. Figure 5C provides a useful illustration of how the temporal horizon enables approximation of whether a given movement is intentional or not—regardless of the state of the camera (e.g., stationary or in motion).

PUBLIC VERSION

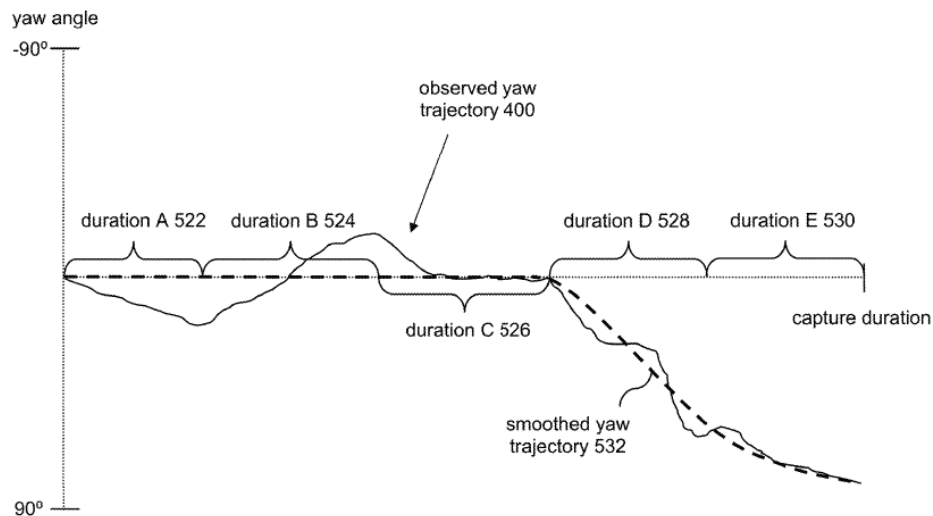


FIG. 5C

68. In processing durations A, B, and C for example, the system's ability to look-ahead across a sufficiently long temporal horizon of the observed trajectory (400) allows the system to infer that measured motion within these durations was, in fact, unintentional and process the smoothed trajectory (532) accordingly. In processing duration D, by contrast, the temporal horizon allows the system to understand that the detected motion persists into the future and is therefore intentional. Thus, the system can immediately begin processing a trajectory at the beginning of duration D that smoothly tracks the path of intended motion. The description in the specification is instructive:

Based on the subsequent portion(s) (look ahead) of the observed yaw trajectory 400, the smoothed trajectory component 104 may determine that the staggered rotation of the image capture device to the right during durations 528, 530 included unintended motions (non-

PUBLIC VERSION

continuous rotation to the right), and may determine the smoothed yaw trajectory 532 to include continuous rotation to the right during durations 528, 530.

69. Thus, the “temporal horizon *enabling approximation of intentional motion*” requires inferring intentional versus unintentional motion in a specific manner. The claim requires more than simply comparing detected motion in a current frame to a motion threshold, or even comparing the difference in motion between two successive frames to a motion threshold. The claim requires looking ahead over a temporal horizon that is sufficiently long to enable an approximation of intentional motion in producing the smoothed trajectory.

70. In my opinion, these innovations are integral to GoPro’s HyperSmooth technology, which has played a significant role in the commercial success and industry recognition of GoPro’s cameras.

II. OVERVIEW OF THE CITED ART

71. The petition challenges the ’840 patent’s claims 1 to 21 on seven grounds. Ground 1 asserts that the combination of Bell and Shi renders the independent claims 1 and 16, along with certain dependent claims, obvious. Grounds 2 through 7 address the remaining dependent claims by asserting the combination of Bell and Shi, along with one or two other references. Ground 5

PUBLIC VERSION

asserts that the combination of Bell, Shi, and Kwatra renders dependent claim 13 obvious. Below are short summaries of the primary references.

B. Ground 1 Cited Art**1. *Bell (Ex. 1006)***

72. Bell discloses a video stabilization system that uses gyroscopes to measure device motion and applies a virtual camera model to reduce shake by smoothing the trajectory. Ex. 1006, 1:15-29; 6:9-18; 8:51-56.

73. However, Bell does not disclose key features of the '840 patent—such as the low-light high-pass parameter, which reduces smoothing in low-light conditions; the weight-balance parameter, which adjusts smoothing by minimizing a combination of rotational velocity and acceleration; or the stickiness parameter, which flexibly controls the influence of prior motion. Ex. 1001, 17:40-53; 21:17-58; 22:22-41.

74. Notably, the Bell patent is closely related to a contemporaneous publication, the Bell Paper (Ex. 1004)—and the Petition notes that “Bell and Bell Paper have the same inventors or authors and are directed to similar subjects,” and that “[a] POSITA consulting Bell would have been motivated to turn to Bell’s related publication, Bell Paper . . .” Petition at 68. The Bell Paper expressly teaches away from using acceleration data in stabilization, stating that such data is “difficult and of limited value.” Ex. 1004 at 4.

PUBLIC VERSION

75. These distinctions make Bell materially different from the '840 patent, particularly in its capacity to handle a wide variety of recording situations.

2. Shi (Ex. 1007)

76. Shi discloses a video stabilization method for handheld devices such as smartphones, using a hybrid approach that combines optical and electronic image stabilization (OIS/EIS). Ex. 1007 at 1:5-15; 1:36-37. For example, it discloses that “[b]y using OIS lens shift positions, the [EIS] stabilization processing can correct for distortion caused by the OIS systems . . .” *Id.* at 1:40-42.

77. Shi’s method applies a first transformation during capture using OIS and device position data. *Id.* at Abstract; 10:4-12; 12:30-34. After capture, a second transformation smooths the trajectory based on a filtered device position referred to as $V_0(T)$, which may be further adjusted to $V_1(T)$ and other virtual camera positions. *Id.* at 12:44-13:49. Notably, the smoothing process in the second transformation begins with a Gaussian “low-pass” filter that **removes** high frequencies—regardless of lighting conditions. *Id.* at 12:54-57; 16:66-17:2.

78. Only *after* this removal does Shi perform a separate blur estimation step as part of an optional third transformation ($V_2(T)$). *Id.* at 17:53-56. More specifically, Shi estimates a motion blur kernel to determine the amount of blur incurred in the captured frame. *Id.* at 18:1-26. It does so by analyzing motion

PUBLIC VERSION

associated with a frame based only on the start and end positions of the camera for that frame—comparing projection matrices at the start and end of the exposure period and examining the difference between the start and end—to estimate how much blur occurred in the frame. *Id.* at 3:19-24.

79. The resulting measure, called a blur kernel, is then used to calculate a coefficient that governs the degree to which the smoothed trajectory is blended with the original trajectory. *Id.* at 18:27-40. Shi teaches the use of spherical linear interpolation (Slerp)¹, weighted by that coefficient, to reintroduce components of the original motion, regardless of frequency, and mix them with the smoothed trajectory to obtain a third trajectory that mitigates the effect of the detected blur.

¹ Slerp is a function that can interpolate between two vectors along the geodesic path on the unit hypersphere. Since it is designed to perform interpolation along a path on a hypersphere, it is a suitable method of interpolating vectors that are defined on a sphere or a hypersphere (e.g. interpolating Euler angles or quaternions, etc.). One key point about Slerp is that it is defined in terms of Dirichlet function. Therefore, one can readily verify that its Fourier spectrum is essentially a low-pass filter that attenuates high frequencies, which is an expected behavior of an interpolating function.

PUBLIC VERSION

Id. at 18:1-6; 18:30-33 (“interpolating between $V_1(T)$ and $R(T)$ according to the amount of image blur detected”).

80. Shi’s approach is triggered by the amount of estimated blur, not by “low-light” conditions. Indeed, Shi does not even mention low-light environments or lighting conditions at all. It is also not a “high-pass” system that allows high-frequency motion to pass unaffected, but rather one that interpolates between the original and smoothed trajectories.

3. Kwatra (Ex. 1012)

81. Kwatra discloses a video stabilization method that estimates camera motion by tracking image features across frames and fitting ***linear motion models*** to the detected feature motion. Ex. 1005, 13:23-14:52. Kwatra then solves for a smoothed camera path comprised of linear segments using a cost function that minimizes at least one derivative along each segment. *Id.*, 1:48-57, 2:6-14, 6:1-28, 6:40-42, 13:23-48.

82. As shown below, Kwatra’s optimization framework is based entirely around linear motion models and explicitly avoids the complexities involved in minimizing rotational dynamics.

PUBLIC VERSION

Kwatra's Camera Path Model Uses Linear Motion Fitting

83. Kwatra's video stabilization methodology begins by using "feature tracking" of the image data to determine an "original camera path" ("C(t)"). The camera path is modeled by fitting a linear motion model describing the motion of tracked features from one frame to the next. To determine C(t), Kwatra fits a single linear model to all the feature points tracked between two successive frames, regardless of their depth. The concatenation of a sequence of linear motion models for the tracked features over time (i.e. across successive frame-pairs) provides a model for the original camera path, C(t). Ex. 1005, 5:41-53, 12:11-37, 19:39-58, claim 6. Therefore C(t) is a concatenation of a sequence of linear 2D transformations that describe how a set of 2D image feature points are roughly mapped between successive frames in least-squares sense.

84. Kwatra then conceptualizes a "smooth" or "optimal" camera path ("P(t)"), wherein P(t) is a function of the original camera path C(t) multiplied by some linear transformation B(t), *id.* at 13:49-56:

$$P(t)=C(t)*B(t)$$

where B(t) is a linear 2D transformation, such as similarity, affine, or a homography. Crucially, Kwatra teaches that the homogenous representations of its choice of linear transformation models are in the form of 3x3 matrices with

PUBLIC VERSION

specific number of degrees of freedom (d.o.f.), which are well-known to correspond to 2D transformations (not 3D). In particular, Kwatra clearly states that its similarity transformation has 4 d.o.f., its affine transformation has 6 d.o.f., and its homography has 8 d.o.f. A POSITA would have known that these numbers of degrees of freedom correspond to 2D transformations (not 3D). In the 3D space, a similarity transformation has 7 d.o.f., an affine transformation has 12 degrees of freedom, and a homography has 15 d.o.f. Kwatra never discloses any transformation with the numbers of d.o.f. corresponding to the 3D cases, because it is exclusively working in 2D image domain, i.e. linear 2D transformations applied to 2D feature points extracted from images. Kwatra never considers 3D motions of the camera housing as disclosed by the '840 patent.

85. Kwatra teaches solving for the $B(t)$ transformation to be applied to each video frame to obtain the smoothed camera path. *Id.* at 13:57-62.

86. For example, Kwatra teaches that the “smooth” camera path can be “partitioned” into segments, and $B(t)$ can be solved for each segment to either result in a “static camera” path, a “constant velocity” path,” or a “path of constant acceleration.” *Id.* at 13:23-49. A “stable” camera path may be determined by minimizing a cost function of the first, second, and third-order derivatives of the smoothed camera path. *Id.* at 13:62-14:13.

PUBLIC VERSION

87. Kwatra states that it solves for the unknown elements of $B(t)$ transformation that minimize its cost function, i.e. the unknown elements of transformation matrices across successive frames, and not directly for specific degrees of freedom of such transformations. Kwatra teaches away from smoothing parameters associated with specific degrees of freedom (such as scale or the image plane rotation angle), since minimizing the derivatives of such parameters would involve non-linear relationships that Kwatra purposefully avoids. Kwatra's method is based on linear programming, a well-known method that can be used to optimize a cost function based only on linear constraints (linear equalities or inequalities).

Kwatra's Cost Function Minimizes Only Derivatives of a Linear Path Representation

88. Kwatra's optimization may involve minimizing the cost function (Equation 7) comprised of first, second, and third derivatives of the smoothed camera path $P(t)$ as follows:

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

where a , b , c are weighting factors. *Id.* at 13:62-14:12. "Equation (7) can be minimized by setting each derivative equal to zero and solved, for example. Thus, the derivatives can be defined at different points along the camera path (e.g., per

PUBLIC VERSION

frame), and a derivative can be determined at every point (every frame).” *Id.* at 14:8-12.

Kwatra Intentionally Avoided Non-Linear Minimization

89. Kwatra’s stabilization method is based on linear motion models and explicitly avoids non-linear formulations due to computational constraints.

Kwatra’s inventors expressly chose a linear framework and rejected non-linear (e.g., rotational) optimization due to computational complexity, stating that such an approach was “not fast enough for real-time performance.” Instead, they deliberately constrained their approach to linear programming, as confirmed in Kwatra’s provisional application (incorporated by reference in Kwatra) (Ex. 2019):

- “The optimal algorithm could estimate all parameters for all frames jointly, similar to Bundle Adjustment. ***However, the concatenation of parameters is not linear anymore, requiring a more complex non-linear optimization method which is not fast enough for real-time performance. We therefore opt for a simpler linear update scheme, here described . . .***” (*Id.* at 3).
- “***Since imposing smoothness constraints directly on s and θ would be non-linear, we instead constrain just a and b*** but use the estimated camera path to ensure that s and θ do not deviate too much from that.” (*Id.* at 5).
- “**Linear programming:** This results in a constrained L1 minimization problem which ***can be solved using linear programming.***” (*Id.* at 4).
- “***We fit several 2D linear motion models (translation, similarity, affine)*** to the tracked features describing the motion of the camera between two frames.” (*Id.* at 1).

PUBLIC VERSION

- “Another interpretation is that rotational motion can be well *approximated by a series of parabolic [2D] motion segments.*” (*Id.* at 12).
- “Note that the objective *now becomes linear in the number of slack variables and this optimization can be solved using linear programming.*” (*Id.* at 4).
- “This formulation can be extended to other *linear transforms* such as similarity and affine transforms (for $P(t)$) easily, since *all terms above are linear in $P(t)$ or parameters of $P(t)$.*” (*Id.* at 5).
- “*This, however is expressed in terms of inverse of A , which is non-linear (specifically quadratic) in the parameters of A . To preserve our linear framework, we make the assumption* that the rotation and scale components of A are small enough that if we would transform the frame window corners, they would not move by more than the distance to the nearest crop window corner.” (*Id.* at 7).

90. By deliberately avoiding non-linear computations, Kwatra’s method is fundamentally incapable of minimizing rotational velocity and acceleration of a camera housing. In fact, as I explained above, Kwatra does not even minimize the derivatives of the image plane 2D rotation, let alone the derivatives of the camera housing’s 3D rotation, because even image plane rotation (1 d.o.f of rotation) would lead to non-linear relations that cannot be solved using linear programming.

Kwatra’s References to Rotation

91. Kwatra’s references to rotation appear in contexts unrelated to its cost function and optimization framework, *i.e.* nowhere teaching the minimization of the angular velocity and angular acceleration of the camera housing. Kwatra’s

PUBLIC VERSION

references to rotation are restricted to stating that its 2D linear transformation models have one degree of freedom associated with image plane rotation, and rather than optimizing this rotation Kwatra merely bounds image plane rotation to assert a quality control (Ex. 1005, 11:5-11, 11:63-66, 16:31-35):

- **Rotation as an Image Degree of Freedom, Not Camera Housing Rotation:** Kwatra defines rotation of an image as a degree of freedom within its linear transformation model (*id.* at 11:5-11) but does not describe the rotation of a camera housing, i.e. no mention of the three degrees of freedom associated with the rotation of camera housing.
- **Rotation Flags Unreliable Feature Matching, Not an Optimization Criterion:** Kwatra detects “too much rotation” as a quality-control measure to flag unreliable frame-pair linear motion models, rather than incorporating it into its optimization function by directly minimizing its derivatives. Ex. 1005, 11:63-66.
- **Bounding Constraints on Rotation:** Kwatra applies bounding constraints on scale and rotation (*id.* at 16:31-35) to limit excessive image transformations. However, these bounding constraints are applied to prevent excessive motion changes in the optimized linear path model, which is not the same as minimizing the rotational velocity or acceleration of the camera housing.

III. GROUND 1: PETITIONER FAILS TO ESTABLISH THAT BELL AND SHI RENDER CLAIMS 1, 14 AND 16 OBVIOUS

A. Claims 1, 14, and 16: Bell and Shi Do Not Disclose Or Render Obvious “Determine a Smoothed Trajectory of the Housing Based On . . . A Low-Light High Pass Parameter . . .”

92. Claims 1 and 16 require the stabilization system to “determine a smoothed trajectory of the housing based on a look-ahead of the trajectory and one

PUBLIC VERSION

or more of a weight-balance parameter, *a low-light high pass parameter*, and/or a stickiness parameter.”² (Emphasis added).

93. Claim 14 further requires that “*the low-light high-pass parameter controls [the] amount or strength of smoothing performed in smoothed trajectory generation and an extent of motion blur compensation applied in determination of the smoothed trajectory.*” (Emphasis added). In my opinion, the Petition fails to establish that the prior art renders this limitation obvious based on the “low-light high-pass parameter.”

94. As discussed above, the claimed “low-light high-pass parameter” means “a parameter that controls the amount or strength of smoothing used in motion blur compensation by allowing more high frequencies to pass through based on low-light conditions.”

95. The Petition relies on Shi’s disclosure of a “coefficient” based on a “motion blur kernel” as allegedly disclosing a “low-light high-pass parameter.” But Shi addresses blur in a fundamentally different way *after* determination of the

² Petitioner has not asserted that claim 1 is obvious based on any prior art allegedly teaching the other two recited parameters.

PUBLIC VERSION

smoothed trajectory, without regard to light conditions, and in a manner inconsistent with the claimed high-pass system.

96. Specifically, Shi's process begins with a stand-alone smoothing process in which "[t]he positions $R(t)$ of the recording device 102 . . . can be filtered to smooth motion over the range, for example, using a Gaussian filter." Ex. 1007 at 12:54-57. The result of this process is an "initial virtual camera position, $V_0(T)$." *Id.* at 12:57-59, 16:41-17:4.

97. Only after this smoothing process is complete does Shi take steps to remove other "undesirable movement in the device positions or other factors that result in undesired motion." *Id.* at 12:59-62. For example, Shi teaches the possibility of generating a "second virtual camera position, $V_1(T)$ " by interpolating between $V_0(T)$ and a camera position for the previous frame based on an amount of motion occurring over future frames. *Id.* at 13:4-15, 17:5-52. Since interpolation is essentially a low-pass filtering process³, $V_1(T)$ is a virtual camera position that is obtained by further low-pass filtering of adjacent frames.

³ An interpolating function is a smooth function that provides an estimate of new intermediate data points from a set of sampled data points, such that the interpolating function goes through every sampled data point. This is in contrast to a fitting

PUBLIC VERSION

98. Shi then teaches the possibility of generating a “third virtual camera position, $V_2(T)$ ” by interpolating $V_1(T)$ with the real device position $R(T)$ based on the amount of camera motion blur present in the frame 211. *Id.* at 13:16-23, 17:53-18:40. Since this blending of $V_1(T)$ and $R(T)$ is done through an interpolation process with Slerp, it is not and cannot be a high-pass filtering process, but is rather a low-pass filtering, i.e. a weighted sum of $R(t)$ and $V_1(T)$, with the weights determined as a function of estimated motion blur. This weighted sum may be viewed as a convolution with a two-tap low-pass FIR filter that attenuates high

function (such as least squares) that is not required to go through the sampled data points, although both are smoothing estimators. Interpolating functions vary in terms of accuracy and smoothness, with the latter property typically expressed in terms of the number of continuous derivatives of the function, or monotonicity of the function in sampling intervals. For instance, polynomials of different orders or splines of different orders are among the popular interpolating functions used in various different applications of signal/image processing. The smoothness constraints in practice translate to the Fourier spectrum of the function attenuating high frequencies, or better stated, makes them behave like low-pass filters, with the cut-off frequency depending on the level of smoothness.

PUBLIC VERSION

frequencies based on the estimated motion blur strength. As a result, Shi's system performs an interpolation between the smoothed and original trajectories using a low-pass filter (Slerp) that blends high- and low- frequency motions from the original camera trajectory.

99. As part of determining the third virtual camera position $V_2(T)$, Shi estimates a "motion blur kernel" to determine the amount of blur incurred in the video frame. *Id.* at 18:7-9. Shi does so by comparing projection matrices at the start and end of the exposure period and analyzing the motion of a reference offset point to estimate how much blur occurred in the frame. *Id.* at 18:10-26.

100. The resulting estimated motion blur kernel is then used to control the degree to which original trajectory motion is *reintroduced* to the previously smoothed trajectory. Specifically, the motion blur kernel is used to determine a coefficient that is used as weighting factor in a spherical linear interpolation ("SLERP") between the original trajectory ($R(t)$) and second trajectory ($V_1(T)$) to obtain a third trajectory ($V_2(T)$). *Id.* at 18:27-40 (teaching spherical linear interpolation between $V_1(T)$ and $R(T)$ according to the amount of image blur detected). SLERP is a low-pass filter that interpolates (with constant angular velocity) along the shortest path (a.k.a. geodesic path) on the unit hypersphere between $V_1(T)$ and $R(T)$ represented for instance in quaternions or Euler angles.

PUBLIC VERSION

This blending of the second trajectory with the original camera trajectory, by low-pass filtering through SLERP to handle motion blur, goes against the teachings of the '840 patent of “high-pass” filtering in “low-light” conditions.

101. Therefore, not only Shi is not teaching $V_2(T)$ is obtained through a “high-pass” process, but also does not even use “low-light” as an indication to compensate for motion blur. Instead, it relies on comparing projection matrices at the start and end of the exposure period and analyzing the motion of a reference offset point to estimate how much blur occurred in the frame. This is fundamentally different from the teaching of the '840 patent that relies on “low-light” as an indication to perform “high-pass” filtering in order to compensate for motion blur.

102. Petitioner contends that the weighting coefficient, based on the motion blur kernel, satisfies the “low-light high-pass” parameter of the claims. For multiple reasons, however, Shi’s disclosures fail to render obvious the claimed requirement that the system “determine a smoothed trajectory based on a look-ahead of the trajectory and . . . *a low-light high pass parameter . . .*”

1. *Shi Does Not “Determine A Smoothed Trajectory” Based On The Purported Low-Light High Pass Parameter.*

103. First, Shi does not “determine a smoothed trajectory” based on the purported “low-light high-pass parameter,” as the plain claim language requires. In

PUBLIC VERSION

Shi, smoothing is accomplished by determination of the first virtual camera position ($V_0(T)$) based on low-pass filtering the positions $R(t)$ of the recording device. Ex. 1007 at 12:54-57 (“The positions $R(t)$ of the recording device 102 over this time period can be filtered *to smooth motion* over the range, for example, using a Gaussian filter.” (emphasis added)). This smoothing process does not involve any consideration of motion blur, nor is it “based on” the purported low-light high-pass parameter. Rather, Shi’s blur correction is only performed later, as a separate step.

104. Shi’s blur correction is also not performed as part of “determin[ing] a smoothed trajectory,” but rather as part of a new process involving reintroduction of real trajectory data to the already smoothed trajectory through another low-pass filter (SLERP) used as an interpolator. This is fundamentally different from what the claim language requires.

2. *Shi’s Parameter is Calculated Independently Of Low Light Conditions.*

105. Second, Shi’s purported low-light high-pass parameter (the weighting coefficient based on the estimated motion blur kernel) is calculated based strictly on the amount of motion incurred in a given frame. It is *not* based on low-light conditions. Indeed, Shi does not even mention low-light conditions. And Petitioner’s expert Dr. Katsaggelos confirmed that Shi’s motion blur kernel is

PUBLIC VERSION

calculated independent of the cause of the blur and is not tied specifically to low-light conditions. *See* Ex. 2020 at 103:22-104:8 (THE WITNESS: “So it’s just understood here that this is a method that is independent of what the cause of the motion blur is.”); *id.* at 107:11-108:3 (Q. “This example is not tied specifically to the nature of the lighting conditions; isn’t that right, Doctor? . . . THE WITNESS: This example gives us the algorithm of estimating this motion parameter, motion blur parameter. It can be applied to low-lighting conditions, normal lighting conditions. ***It’s not tied to any of these situations.*** As long as blur, due to whatever it is, is one way to estimate the magnitude of this blur.”) (Emphasis added).

106. To illustrate why Shi’s coefficient based on the motion blur kernel is not a low-light parameter consider the following: it is entirely possible that Shi’s system could calculate a relatively high value for its weighting coefficient for fast moving objects even in well-lit conditions. By the same token, Shi may well calculate a relatively low value for its weighting coefficient when capturing slow moving or stationary objects even in low-lit conditions. This demonstrates that Shi’s coefficient is not based on low-light conditions and thus is ***not*** a low-light high-pass parameter targeted to the specific issue of blur addressed by the ’840 patent. To further highlight these fundamental differences consider a camera jittering during exposure period. It is quite conceivable that the projection matrices

PUBLIC VERSION

may end up being the same at the start and the end of the exposure, resulting in a false negative scenario affecting Shi's estimated weighting coefficient, despite the fact that jitter might have caused motion blur in low light condition during exposure period. Indeed, regardless of lighting conditions, motion blur may be even caused due to differences in depth for points in the scene (including Shi's own reference offset point used for analysis), which can create both false positive or false negative analysis of motion blur for Shi, and result in a weighting coefficient that is very different from the "low-light high-pass" parameter. These simple examples demonstrate that Shi's weighting coefficient determined by comparing camera projection matrices at the two ends of the frame exposure period and analyzing the motion of a reference offset point is not the same as the "low-light high-pass" parameter.

3. *Shi's Parameter Does Not Allow More High Frequency Motion To Pass Through And Thus Is Not A "High Pass" Parameter*

107. Third, Shi's blur mitigation system does not function as a "high-pass" system—*i.e.*, one allowing more high-frequency motion to pass while continuing to attenuate low frequency motion. Rather, Shi's system performs a spherical linear interpolation (*i.e.* a low-pass filtering using SLERP) between the smoothed and original trajectories, reintroducing both high- and low- frequency motion from the original camera trajectory. Because the linear blending of smoothed and

PUBLIC VERSION

original trajectories does not function as a “high-pass” system, this further confirms that the weighting coefficient is not a low-light high-pass system as claimed.

4. *Shi’s Approach To Blur Management Results In Practical Differences From The Claimed Invention.*

108. In sum, Shi’s approach to blur management operates after determination of the smoothed trajectory to reintroduce original camera motion, regardless of the frequency of motion and without regard to lighting conditions. This is vastly different from the claimed approach of determining the smoothed trajectory based on a parameter that controls the amount or strength of smoothing used in motion blur compensation by allowing more high frequencies to pass through based on low-light conditions.

109. The differences between Shi and the claimed technique have real world implications. For example, GoPro’s invention is directed to cameras designed for fast moving sporting activities like cycling or running where camera shake is a major issue. In these circumstances, it makes sense to maintain the strength of smoothing, even at the expense of some motion blur—especially in fast-paced, high-motion scenes. Thus, GoPro’s low-light high-pass parameter is tailored to a specific type of motion blur—*i.e.*, one that occurs in a “low-light” scenario characterized by longer exposure times—and does so in a targeted way—

PUBLIC VERSION

by allowing high frequencies to pass while attempting to maintain attenuation of lower frequency motion. For other scenarios, however, GoPro's invention prioritizes smoothing of camera shake over concerns about blur.

110. Shi's approach, on the other hand, is directed to more causal video recording devices such as cameras on cellular telephones or tablet computers. Ex. 1007 at 1:5-7. For these types of devices, it may make sense to prioritize blur correction, even if it comes at the expense of maintaining smoothing strength. That is why Shi never mentions low-light conditions, and why it does not implement a high-pass system for correcting blur. Accordingly, in my opinion, Shi's "coefficient based on the motion blur kernel" does not teach nor suggest the claimed "low-light high-pass parameter."

5. *A POSITA Would Not Have Been Motivated To Incorporate Shi's Parameter Into Bell With A Reasonable Expectation of Success.*

111. Finally, even if Shi did disclose the claimed low-light high pass parameter, in my opinion, Petitioner fails to adequately explain why a person of ordinary skill in the art would have been motivated to incorporate Shi's blur-based coefficient into Bell's approach to stabilization with a reasonable expectation of success.

PUBLIC VERSION

112. Bell’s stabilization approach is fundamentally based on defining a virtual camera frame within the larger image frame, and then “dampening” or “accelerating” the movement of that virtual camera frame based on its location relative to an “inner region” of a padding space. Ex. 1006, claim 1; *see also id.* at 9:22-25 (“If the hypothetical crop polygon 340 falls outside the inner region, then the virtual camera needs to move faster or slower to match the physical camera and avoid letting the crop polygon go out of frame.”). For example, to calculate the virtual position of the camera, Bell discloses a “decay” factor when the prior position is within the inner region, and a “mixing weight” when the prior position is outside of the inner region.

113. Thus, a POSITA would have understood that Bell’s system is based on a delicate balance of slowing down or speeding up the rate of change of the virtual camera to provide acceptable stabilization while keeping the virtual camera crop window within the boundaries of the image frame. Petitioner provides no explanation for how Bell’s system can accommodate Shi’s coefficient based on a motion blur kernel without compromising the balance that Bell establishes between dampening and accelerating the virtual camera motion. Neither Petitioner nor its expert considered or addressed this interaction and thus in my opinion have failed to establish that their purported combination renders the claim limitation obvious.

PUBLIC VERSION

B. Claims 1 and 16: Bell and Shi Do Not Disclose Or Render Obvious the Claimed “Temporal Horizon Enabling Approximation of Intentional Motion”

114. Claims 1 and 16 require: “the determination of the smoothed trajectory based on the look-ahead of the trajectory 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.”

115. The ’840 patent teaches that because intentional motion “may be characterized by lower frequency motion than unintentional motion,” intentional motion can be identified by considering a “sufficiently long temporal horizon of experienced motion.” Ex. 1001 at 9:17-23. Thus, in implementing the claimed “look ahead,” the ’840 patent teaches a **tradeoff** between **processing delay** and **accuracy** in estimating whether a given motion is an intentional motion. *Id.* at 9:25-28 (“Longer intervals of the temporal horizon may enable better approximation of the intentional motion while causing longer delays (longer wait to determine smoothed trajectory).”). Bell and Shi make no such tradeoff; they apply short buffering and fixed heuristics to avoid abrupt virtual camera changes—not to approximate user intention.

PUBLIC VERSION

1. *Bell Fails To Teach Or Suggest A Temporal Horizon For Enabling Approximation Of Intentional Motion.*

116. Petitioner contends that Bell discloses a “temporal horizon enabling approximation of intentional motion” because it “define[s] a new virtual camera that stays static when the motion is so small that the intention is probably that the camera should remain static.” Pet. at 39 (citing Ex. 1006, 8:52-55). Petitioner also cites Bell’s statement 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.” *Id.* (citing Ex. 1006 at 10:10-13). From this, Petitioner concludes that Bell uses motion experienced during future frames to “approximate intentional motion.” *Id.* at 39-40.

117. Petitioner’s argument conflates two distinct aspects of Bell’s system and mischaracterizes both. Understood in context, Bell does not disclose a “temporal horizon enabling approximation of intentional motion.”

118. First, Bell’s primary method of determining whether and by how much to move the virtual camera does not involve any look-ahead, temporal horizon, or approximation of intentional motion. Bell’s stabilization technique uses a “crop polygon” to track the motion of a “virtual camera” between frames and adjust the center of the view accordingly. For any given frame, the orientation

PUBLIC VERSION

of the virtual camera is initially determined based only on *past information*—*i.e.*, a multiplication of the virtual camera orientation from the prior frame by the virtual camera velocity in that prior frame. Ex. 1006 at 9:1-5 (describing “Relation 4”). This technique does not involve any consideration of a temporal horizon or any approximation of intentional motion.

119. To maintain a stable and bounded virtual camera path, Bell also conceptualizes a “padding space” with an “inner region” and an “outer region.” *Id.* at 9:16-20. If, for any given frame, the virtual camera is located within the inner region, the calculation provided by Relation 4 is adopted, and Bell applies a “decay factor” to slow down the rate of change of the virtual camera. *Id.* at 9:20-36. This has the effect of dampening the rate of change of the virtual camera. Alternatively, if part of the virtual camera falls outside the inner region of the padding space, a “mixing weight” is applied to speed up the angular change of the virtual camera. *Id.* at 9:43-10:7. Critically, this process of applying dampening and acceleration factors is based entirely on the newly calculated virtual camera orientation relative to a fixed padding space—again, it makes no use of any look-ahead, a temporal horizon, or any approximation of intentional motion.

120. Second, Bell notes that while the previously described “algorithm works well,” it “occasionally has to make quick changes in velocity when the crop

PUBLIC VERSION

polygon suddenly hits the edge.” Ex. 1006 at 10:8-10. Only in this context—*i.e.*, to avoid abrupt corrections—does Bell propose “a small buffer of future frames, e.g., five or less,” *id.* at 8:66-67, and only for the purpose of ensuring that “sharp changes can be preemptively avoided,” *id.* at 10:10-16.

121. Bell’s look-ahead is thus limited to projecting the crop polygon 2 to 5 frames into the future for the purpose of determining whether the crop polygon will lie outside the inner region at future time. *Id.* at 10:17-25. If the crop polygon is projected to lie outside the inner region, Bell teaches to apply a further “mixing factor” to further increase the velocity of the virtual camera and provide it with a “preemptive nudge in the right direction.” *Id.* at 10:25-37. Bell’s look-ahead system is used to adjust the velocity in calculating the next frame in order to make sure that the crop polygon does not hit the frame boundary, regardless of whether the camera motion is intentional or not. This shows a fundamental difference between Bell’s look ahead system and the claimed temporal horizon that is by design for enabling approximation of intentional motion. Significantly, Bell does ***not*** update the previously calculated virtual position of the camera based on the lookahead as would be expected from a system seeking to approximate intentional motion—Bell merely changes the velocity that is used in calculation of the next frame. *Id.* at 10:34-37.

PUBLIC VERSION

122. Thus, while Bell does teach a very short “look ahead,” it does so only for the limited purpose of predicting whether the virtual camera will approach the frame boundary at a future time so that a further correction value may be preemptively applied to avoid “abrupt” corrections in the future. This does not teach the use of the temporal horizon for purposes of approximating whether motion at a current time is intentional. Bell’s limited look-ahead, is merely a safeguard to prevent the crop polygon from drifting too close to the boundary, resulting in a sudden corrective movement at a future time.

123. Indeed, according to Bell’s teachings, the dampening and acceleration factors are applied without regard to whether actual camera motion at the current time was intentional or not. Rather, Bell applies these correction factors based strictly on whether current and future frames fall outside the inner region of the padding space.

124. Accordingly, in my opinion, Bell does not disclose or suggest the claimed “temporal horizon enabling approximation of intentional motion.”

2. *Shi Likewise Fails To Teach Or Suggest A Temporal Horizon For Enabling Approximation Of Intentional Motion.*

125. As a fallback argument, Petitioner contends that Shi discloses the temporal horizon limitation, because Shi’s “system can reduce the impact of large or sudden movements using look-ahead frames” and “evaluate whether consistent

PUBLIC VERSION

movement, such as panning, is occurring and adjust the video frames to reflect the consistent movement.” Pet. at 40 (citing Ex. 1007, 5:17-24). Neither of these assertions withstand scrutiny.

126. First, Shi’s use of look-ahead frames to “reduce the impact of large or sudden movements” does not teach the use of a temporal horizon to enable an approximation of intentional motion as claimed. In fact, Shi’s approach contradicts the claimed approach, because it involves intentionally introducing motion to frames where the system determines that no motion actually exists. For example, Shi explains that “for a current frame under analysis, the position of the recording device may not have moved significantly since the previous frame.” Ex. 1007 at 10:46-49. “*Nevertheless*, Shi teaches, “if the engine 170 determines that significant motion occurs in future frames, the second transformation 172 can be generated to *shift or otherwise alter the current frame to introduce apparent movement in the video*, so that large future movement can be spread out as a series of gradual changes rather than an abrupt change.” *Id.* at 10:49-54. This is the opposite of the using a temporal horizon to enable approximation of intentional motion—Shi teaches a different approach of intentionally *altering* motion across frames to avoid abrupt to sudden changes.

PUBLIC VERSION

127. Second, Petitioner’s reliance Shi’s disclosure of “evaluat[ing] whether consistent movement, such as panning, is occurring,” Pet. 40 (citing Ex. 1006 at 5:22-24) is similarly misplaced. Shi teaches that this technique involves calculation of a fourth virtual camera position ($V_3(T)$) that “is generated by filtering the device patterns to conform to a predetermined, consistent motion pattern.” Ex. 1006 at 13:35-40. Shi itself admits that this technique “impose[s] a *predetermined* shape or pattern,” *id.* at 13:35, as distinguished from prior filtering steps that “smooth[] but generally follow[] the changes in estimated actual device positions over time,” *id.* at 13:32-34. This approach of “imposing” a predetermined shape or pattern is not the same as using the temporal horizon to “enable approximation of intentional motion” of the image capture device. In fact, a POSITA recognizes that resort to such a “heuristic”-based approach indicates an inability of the system to implement the claimed temporal horizon to enable approximation of intentional motion while determining a smoothed trajectory as claimed.

IV. GROUND 2-4 AND 6-7 FAIL ON THE MERITS

128. Petitioner asserts that the combination of Bell and Shi in combination with various other references render the remaining dependent claims of the ’840 patent obvious. However, except for claim 13 (addressed below), these combinations fail for the same reason as Ground 1.

PUBLIC VERSION

V. GROUND 5: PETITIONER FAILS TO ESTABLISH THAT BELL, SHI, AND KWATRA RENDER CLAIM 13 OBVIOUS

129. Dependent claim 13 further requires that “*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.*” (Emphasis added). Petitioner fails to establish that this claim is obvious.

B. Claim 13: Kwatra Does Not Disclose Or Render Obvious The “Weight-Balance Parameter” For Controlling Rotational Velocity and Rotational Acceleration

130. As in IPR2024-01186, Petitioner relies on Kwatra’s disclosures about optimizing certain *linear motion models* in Kwatra (Ex. 1012). These disclosures, however, fail to teach or suggest a “weight balance” parameter for controlling the extent to which rotational velocity and rotational acceleration are minimized. Kwatra, in fact, intentionally avoided the type of non-linear calculations that minimizing *rotational* velocity and acceleration would entail.

1. Kwatra’s Linear Framework Does Not Provide For Minimization Of Rotational Velocity And Rotational Acceleration Of The Housing As Required

131. Due to concerns with computational constraints, Kwatra’s motion estimation and smoothing approach is fundamentally based on linear motion models. Kwatra expressly rejects the sort of non-linear computations that would be

PUBLIC VERSION

required to optimize rotational dynamics. Thus, Kwatra's technique is incapable of controlling or balancing the minimization of rotational velocity and rotational acceleration.

132. Kwatra discloses a video stabilization method that estimates camera motion by tracking image features across frames and fitting *linear motion models* to the detected feature motion. Ex. 1012, 13:23-14:52. The motion trajectory is then smoothed using an optimization framework that minimizes derivatives of the estimated motion path based on a linear cost function. *Id.*, 1:48-57, 2:6-14, 6:1-28, 6:40-42, 13:23-48.

133. Kwatra's approach begins with an analysis of the image information to determine an "original camera path" ($C(t)$). First, "trackable features between two frames may be extracted by extracting a number of features in a first video frame (e.g., based on x and y location) and tracking the extracted features in a next video frame." Ex. 1005 at 7:58-62. Once features are extracted and tracked, "two-dimensional (2D) linear motion models" are "fit to the trackable feature-matches to describe motion of the camera between video frames. *Id.* at 9:51-45. "An estimate of the original camera path ($C(t)$) can be determined by fitting linear motion models to the tracked feature pair matches resulting in a linear transform for each feature pair." *Id.* at 9:57-60. The original camera path $C(t)$ is a simple

PUBLIC VERSION

concatenation of the linear 2D motion models for feature-pair matches between consecutive frames. *Id.* at 12:11-21; *see also* claim 6. Therefore, in order to determine $C(t)$, Kwatra fits a sequence of linear 2D motion models (linear 2D transformations) to every pair of successive frames. However, Kwatra fits a single linear transformation to all the feature points tracked between two frames regardless of their depth.

134. Kwatra then conceptualizes a “smooth” or “optimal” camera path ($P(t)$), wherein $P(t)$ is defined as:

$$P(t)=C(t)*B(t)$$

where $C(t)$ represents the estimated original camera path and $B(t)$ is a linear transformation matrix applied to each frame to generate a stabilized sequence. Ex. 1005, 13:52-56.

135. Petitioner contends that the cost function (Equation 7) discloses that the first, second, and third derivatives of the smoothed camera path $P(t)$ are minimized:

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

Ex. 1005, 13:62-14:12.

PUBLIC VERSION

136. However, Kwatra is minimizing this cost function in terms of the elements in the smoothing transformation B . Minimizing the derivatives of the smoothed camera path as discussed in Kwatra does **not** result in minimization of a combination of **rotational** velocity and acceleration. As the discussion from columns 14 through 15 make clear, the derivatives in column 7 are computed strictly as a function of the *linear* transforms between a series of consecutive frames. See, e.g., Ex. 1005 at 14:54-15:56. Equation 12, for example, makes this clear. Equation 12 expresses the linear frame-pair transform that represents motion from a first position (x_1) to a second position (x_2) based on linear relationship between four degree of freedom parameters (a_t , b_t , c_t , and d_t) and the translation in the x and y dimensions (dx_t and dy_t).

137. As the Board recognized in its Institution Decision, Kwatra does teach that the 2-dimensional linear motion model may be represented as a “similarity” or “affine” transformation with degrees of freedom for scale and rotation. Ex. 1005 at 11:3-21. Kwatra, however, expressly rejects an approach that minimizes rotation or scale in its smoothing/optimization step.

138. As an initial matter, it is important to note that the rotation degree of freedom parameter discussed in Kwatra refers to the rotation of a trackable feature

PUBLIC VERSION

in an image, *not* rotation of the camera housing.⁴ As discussed below, rotation of a trackable feature and rotation of the camera housing are not the same. Even more fundamentally, however, despite teaching motion models that may include a rotation or scale parameter, Kwatra expressly rejects an approach that would attempt to minimize these changes to rotation or scale as part of its smoothing/optimization step.

139. Specifically, for similarity or affine transformations, Kwatra teaches to impose smoothness constraints only on the degree of freedom parameters “a” and “b,” rather than directly on the scale (“s”) and the image rotation angle (“ θ ”) of trackable features. *Id.* at 17:5-20. Kwatra acknowledges that this approach of

⁴ Kwatra’s teaching is strictly about the 2D rotation of the image plane (i.e. the image plane rotating either clockwise or anti-clockwise, corresponding to the one d.o.f. of rotation that Kwatra refers to) – not the rotations of the camera housing that has 3 d.o.f. associated with its rotation in the 3D space. Even disregarding the fact that Kwatra’s reference to rotation is the image plane rotation, Kwatra expressly rejects an approach that minimizes this rotation in its smoothing/optimization step, due to the fact that minimization of rotational gradients for smoothing would lead to non-linearities that Kwatra considered undesirable and rejected.

PUBLIC VERSION

constraining smoothness on only a and b results in a path where “rotation and scale may not remain smooth.” *Id.* at 17:11-18. To address this issue, Kwatra teaches instead to place bounding constraints (*i.e.*, numerical limits) on changes to scale and rotation in the smoothed camera path “to ensure s and θ do not deviate too much.” Ex. 1005 at 17:19-20.

140. Thus, Kwatra itself recognizes that its linear optimization model does not result in a smoothed θ —*i.e.*, it does not minimize any derivatives of θ . As Kwatra teaches, this simplification is necessary because “imposing smoothness constraints on s and θ may be non-linear,” *id.* at 17:18-19, which would subvert Kwatra’s consistent teaching for linear models and using linear programming.

141. It is also worth noting that Kwatra does not even teach bounding θ itself but rather discloses bounding the sine or cosine of the image rotation angle, which is not the same as bounding the image angle θ itself. Ex. 1005 at 17:37-38. But more importantly, placing bounds on the extent the scale s and the $\cos(\theta)$ or $\sin(\theta)$ are allowed to change during estimation of the camera path is not the same as controlling a minimization of rotational velocity and rotational acceleration of the camera housing.

142. Kwatra’s provisional application (Ex. 2019), which is incorporated into Kwatra by reference, Ex. 1005 at 1:8-11, provides further explanation and

PUBLIC VERSION

support for the foregoing understanding of Kwatra. The provisional application notes how the “optimal algorithm could estimate all parameters for all frames jointly,” however this “concatenation of parameters is not linear anymore, requiring a more complex, non-linear optimization method *which is not fast enough for real-time performance.*” Ex. 2019 at 6 (emphasis added). Kwatra’s inventors noted that they “therefore opt for a simpler linear update scheme” *Id.* “To preserve our linear framework,” Kwatra assumes that “*the rotation and scale components of A are small enough* that if we would transform the frame window corners, they would not move by more than the distance to the nearest crop corner.” *Id.* at 10 (emphasis added).

143. Thus, by deliberately avoiding non-linear computations, Kwatra does not and cannot control any minimization of rotational velocity and acceleration of a camera housing path. Kwatra thus does not teach or suggest any weight-balance parameter.

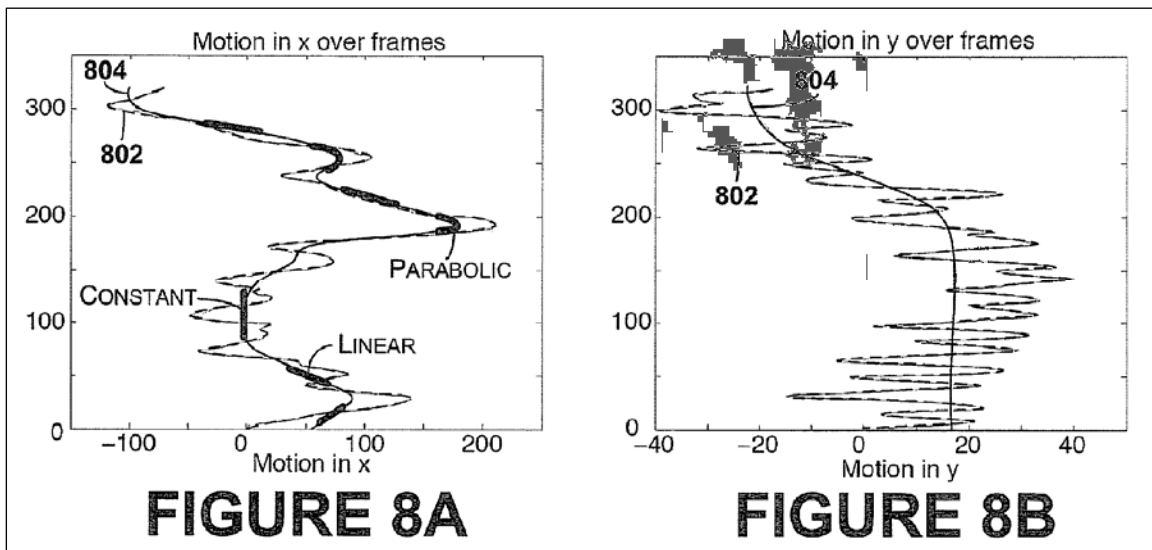
144. In forming his invalidity opinions, Dr. Katsaggelos did not review Kwatra’s provisional application which was incorporated by reference in Kwatra. *See* Ex. 2020, Depo. Tr. at 19:25-20:5 (“Q. So in forming your invalidity opinions, you did not review the provisional application that the inventors in Kwatra filed; is that correct? . . . THE WITNESS: This is a document that --

PUBLIC VERSION

application that I did not analyze.”). Thus, Dr. Katsaggelos was not aware of and did not consider the inventors statements that they avoided more complex nonlinear optimization methods, as “not fast enough for real-time performance.” *See id.* at 21:1-14 (“Q. Do you understand what’s being described there? . . . THE WITNESS: As I mentioned, I haven’t analyzed this document. So I cannot offer an opinion just on the fly by just looking at a sentence or two.”).

145. Further demonstrating that Kwatra fails to render the weight-balance parameter limitation of claim 13 obvious are Kwatra Figures 8A and 8B and their associated description. Figures 8A and 8B depict the “original camera path motion 802 versus an updated or smooth camera path motion 804.” Ex. 1005 at 19:4-6. Kwatra states that Figures 8A and 8B are depicting the “optimal L1 path” comprised of segments with minimized derivatives of the smooth camera path $P(t)$ consistent with Equation 7. Yet, Figures 8A and 8B track camera motion (both original and smooth) exclusively as a function of position in the X and Y dimensions, respectively.

PUBLIC VERSION



146. Figures 8A and 8B make clear that Kwatra's aim is not to control the minimization of any combination of rotational velocity and rotational acceleration. Indeed, Petitioner's own expert at trial, Mr. Goodin, conceded that a POSITA would understand Kwatra's smoothed motion model "P," as represented in Figure 8, to be translational, not rotational. Ex. 2021, Tr. (Goodin) 1368:6-1369:7. And, of course, there is no teaching or suggestion in Kwatra that results in any smoothed camera position as function of a rotational position or angle. This directly contradicts Petitioner's assertion that Kwatra teaches minimizing rotational velocity and acceleration of the camera housing. Indeed, Kwatra's express rejection of non-linear optimization precludes it from teaching or suggesting the claimed weight-balance parameter of claim 13 of the '840 patent.

PUBLIC VERSION

2. *Kwatra's Two-Dimensional Approach To Motion Estimation And Optimization Is Inconsistent With Controlling A Minimization Of Rotational Velocity And Rotational Acceleration Of The Housing As Claimed*

147. Moreover, Dr. Katsaggelos conceded at deposition that modeling rotational velocity of a camera housing requires a three-dimensional model. Ex. 2020 at 79:2-22 (“Q. Can you have angular velocity considered in a two dimensional model? . . . THE WITNESS: The two dimensional models we talk about cannot – cannot cover that, cannot cover the camera rotation in the 3D space, cannot cover the ang[ular] velocity of the camera. I need a 3D model to be able to capture this quantity.”). Kwatra’s motion estimation and optimization techniques, however, are consistently described in the context of *two-dimensional* linear models. Ex. 1005 at 9:41-43 (“two-dimensional (2D) linear motion models (e.g., translation, similarity, affine) can be fit to the trackable feature-matches . . .”), at 19:39-41 (“As described above using the example methods shown in FIGS. 3-4, a 2D original camera path $C(t)$ may be first estimated, and a new camera path $P(t)$ can be determined.”).

148. While Dr. Katsaggelos opined that Kwatra suggested modeling and optimizing motion in three dimensions, he failed to identify anything in the reference that actually discloses such an approach. For example, despite using the word “dimensional,” Kwatra’s mention of “estimates from lower to higher

PUBLIC VERSION

dimensional motion models (e.g., similarity --> homography)” at column 11, lines 58-61 is understood by a POSITA as merely referring to an estimate of a 2D model with more degrees of freedom (*i.e.*, a homography transform) from another 2D model with fewer degrees of freedom (*i.e.*, a similarity). Again, the actual motion modeling and optimization disclosures of Kwatra relied upon by Petitioner are exclusively described in the context of two-dimensional models. Nowhere does Kwatra teach or disclose a method for modeling or optimizing three-dimensional motion.

149. Thus, because the motion modeling and optimization that Petitioner relies upon is inherently a two-dimension model, and because Petitioner’s own expert agrees that the controlling the minimization of rotational velocity and rotational acceleration inherently requires a three-dimensional model, Kwatra fails to disclose or render obvious the claim limitation. Petitioner’s theory fails for this independent reason.

150. Dr. Katsaggelos misunderstands and misinterprets the terms “lower” and “higher dimensional motion models” at column 11, lines 58-61 of Kwatra. Here Kwatra clearly refers to the dimensionality in terms of the degrees of freedom (*i.e.* dimensionality of the parameter space of its linear model). This is clear because of the context that Kwatra is discussing the degrees of freedom of its

PUBLIC VERSION

transformation models and immediately after that statement qualifies it by stating “(e.g., similarity --> homography)”. Crucially, Kwatra teaches that its homogeneous linear transformation models are in the form of 3x3 matrices with specific number of degrees of freedom (d.o.f.), which are well-known to correspond to 2D transformations (not 3D). In particular, Kwatra clearly states that its similarity transformation has 4 d.o.f. (2 for translation, 1 for scale, 1 for rotation), its affine transformation has 6 d.o.f., and its homography has 8 d.o.f. A POSITA would have known that these numbers of degrees of freedom correspond to 2D transformations (not 3D). In the 3D space, a similarity transformation has 7 d.o.f. (3 for rotation, 3 for translation, and 1 for scale), a 3D affine transformation has 12 degrees of freedom, and a 3D homography has 15 d.o.f. Kwatra never discloses any transformation with the numbers of d.o.f. corresponding to 3D cases, because it is exclusively working in the 2D image domain, i.e. linear 2D transformations applied to 2D feature points extracted from images. Therefore, Kwatra never considers 3D motions of camera housing as disclosed by the '894 patent.

PUBLIC VERSION

3. *Kwatra Solves for a Linear Transform $B(t)$ of Each Video Frame, Not for a Trajectory that Minimizes Rotational Velocity and Acceleration of a Camera Housing*

151. The Petition also fails to disclose the weight-balance parameter, because it conflates minimizing derivatives of what Kwatra refers to as the smooth camera path $P(t)$ with controlling minimization of rotational velocity and acceleration of a *camera housing* as required.

152. Independent claim 1 makes clear that claimed system “determine[s] a smoothed trajectory *of the housing*” that results in “smoother changes in the rotational positions of the housing than the trajectory.” Thus, in this context, a POSITA understands that the claimed “weight-balance” parameter of claim 13 must control the “rotational velocity and rotational acceleration” of the housing, as opposed to specific tracked features within the captured image.

153. Kwatra fails to disclose this aspect of the claims. As discussed above, Kwatra’s motion model is based on extracting and tracking feature pairs across successive frames. Thus, Kwatra’s “camera path” does *not* represent the motion of the physical camera housing in space, but rather “describe[s] motion of pixels between frames—such as feature moved 10 pixels to the right, equivalent to movement of the camera to the left by 10 pixels.” Ex. 1005 at 9:49-54.

PUBLIC VERSION

154. The number of pixels that a particular feature moves across a frame is *not* equivalent to the amount of rotation of the camera housing in three-dimensional space. Rather, the relationship between movement of a feature and motion of the camera depends critically upon the distance of that trackable feature from the camera. For example, a background feature and a foreground feature at different depths will move by a different number of pixels for a given displacement of the camera housing in physical space. Petitioner's expert admitted this fact at his deposition, Ex. 2020 at 52:13-21, but did not account for it in his analysis. Mathematically, pixel motions in the frames of a moving camera are described as the sum of two terms. One that depends on the intrinsic parameters of the camera and the relative rotation of the camera between frames (this term is known as the infinite homography), and the second term that also depends on the intrinsic parameters of the camera and is proportional to the relative translation of the camera between frames but inversely scaled by the depth of the 3D point from the camera.

155. Thus "minimizing" derivatives of Kwatra's "camera path," at most, minimizes the apparent motion of pixels for some composite set of features in the image, without taking into account depth variations for feature points. That is because Kwatra is **fitting a motion model** to a set of tracked image feature points

PUBLIC VERSION

regardless of their depths using least squares. Therefore, Kwatra is fitting the parameters of its model (e.g. 2D similarity, affine, or homography) to a set of tracked feature points regardless of their depth, which results in a model that is on average representing the motion of the tracked 2D feature points. In other words, Kwatra's simplified 2D motion models (e.g. 2D similarity, affine, or homography) are underfitting through a linear regression as they ignore depth by limiting the least squares fitting to 2D models. Kwatra itself recognizes and acknowledges that least squares fitting of 2D motion model leads to an averaging regression of the 2D model. Some examples include: "For mathematical translation, the least square fit can be **an average of translations** for each feature pair match." Ex1005, 9:54-56. "As described above using the example methods shown in FIGS. 3-4, the original camera path motion is approximately an inverse of **an average feature path transformation** (e.g., as a camera moves to the left, the image pixel content or features move to the right)." Ex1005, 19:59-63. Although Kwatra suggests using RANSAC to improve its 2D linear fitting or replace L2 norm (least squares) with L1 norm (median type estimator) for minimization, it acknowledges that these are only reducing the effect of outliers in the tracked feature points by either removing them (RANSAC) or making the minimization robust to them (L1 instead of L2). None of these options account for depth

PUBLIC VERSION

differences in the tracked feature points, which makes 2D motion of pixels (along the two image axes) distributed around some average or median value. Kwatra simply takes that average or median to represent all the points in the distribution. This is a fundamental limitation of Kwatra as a 2D method, which by design cannot account for 3D information (i.e. depth) that creates such distributions of observed motion of pixels in the 2D image domain.

156. Petitioner fails to demonstrate how linear average regression of 2D motion models through least square fitting or even L1 minimization would result in the minimization of a combination of rotational velocity and acceleration *of the camera housing* as claimed, which is a 3D optimization.

157. The failure of Kwatra to teach or disclose an approach that controls minimization of the camera housing speaks to the fundamental differences between Kwatra and the '840 patented approach. Kwatra is directed to optimizing a 2-dimensional camera path based on linear motion models extracted from trackable features. The '840 patent, by contrast, is directed to optimizing the trajectory of a video so that it appears as if it were captured by a camera whose housing underwent a minimal combination of rotational velocity and rotational acceleration *of the camera* during capture. Kwatra fails to render the claim obvious for this independent reason.

PUBLIC VERSION**A. A POSITA Would Not Have Been Motivated to Combine Bell, Shi, and Kwatra**

158. In my opinion, the Petition fails to account for the fundamental different image stabilization techniques employed by Bell and Shi on the one hand, and Kwatra on the other hand. The embodiments that Petitioner relies upon from Bell and Shi teach video image stabilization strategies, implemented in mobile devices that make use of an on-board position sensor to estimate the 3-dimensional motion of the mobile device itself. Kwatra, on the other hand, teaches a post-processing, server-based approach that estimates motion through feature tracking within the image data. Ex. 1005 at 3:47 (“In one example, a video may be stabilized by performing post-processing techniques.”), at 5:21-26 (“The video hosting service 200 comprises the camera path translation system 202, a video server 210, an ingest server 212, and a video database 216.”); *see also* FIG. 2.

159. While Petitioner’s expert contends that Kwatra contemplates implementing its techniques in a mobile device, this assertion is unsupported and contradicted by Kwatra itself. For example, while Dr. Katsaggelos cites Figure 13 and its associated discussion, Ex. 1002 at ¶ 117, that discussion merely provides that generic “computing device 1300” can comprise a server or a mobile device and “may be implemented as a client device, a server, a system, a combination thereof, or as a portion of components *described in FIGS. 1-2.*” Ex. 1005 at

PUBLIC VERSION

26:11-21 (emphasis added). Figures 1 and 2 provide an architecture in which the “camera path translation system” (202) is implemented exclusively in the “video hosting service” (200) and only communicates with mobile “client” devices (204) through a “network” (206). Thus, while Kwatra’s disclosure of generic “computing device 1300” is broad, Kwatra teaches that this computing device is for implementing components of the specific architecture of Figure 2 in which camera path estimation and modeling are performed exclusively in a server-based system.

160. This understanding of Kwatra is consistent with the relevant state of the art. A POSITA would have recognized that solving for a complex, multi-derivative trajectory optimization problem was a computationally intensive process that was not considered suitable for a resource-constrained mobile device. Moreover, performing this sort of trajectory optimization on a real-time stream of three-dimensional gyroscope data (*i.e.*, like that collected by systems disclosed in Bell and Shi) would have been a complex, challenging endeavor that is not contemplated by any of the prior art of record. A POSITA would have recognized that Kwatra’s server-based, post-processing optimization of two-dimensional linear motion models was not compatible with Bell/Shi’s mobile-device based stabilization of three-dimensional rotational data. Nor is there any reason to believe

PUBLIC VERSION

that a POSITA would have had a reasonable expectation of success in attempting to perform this combination.

161. Petitioner's expert admitted that the image-based approach of motion estimation taught by Kwatra was a different approach from the gyroscope based motion estimation approach taught by Shi and Bell. Ex. 2020 at 14:12-19; *see also* Ex. 1004 at 1:29-34 (teaching that stabilizing a video by tracking salient features requires "a priori knowledge of the image's content to, for example, identify and track a person or other salient object in the scene.").

162. Moreover, Bell explains why a POSITA implementing a gyroscope-based system would not be motivated to look to the image-based systems like Kwatra:

Using the on-board multi-axis motion sensing system of a mobile phone or other device for motion tracking has a number of advantages. ***Filtering and analyzing the motion sensing system measurements in real time requires much less processing power in comparison to trying to estimate camera rotations from the image data after the fact.*** Motion sensing system estimates can be calculated with a relative few multiplications and additions, while ***image-based methods require processing on thousands or even millions of picture elements.*** Additionally, image-based motion estimation methods can fail when features are sparse or when large objects move in the foreground. A motion sensing system, by contrast, always reports the device motion regardless of how much and in which direction(s) the objects in the scene move.

PUBLIC VERSION

Ex. 1006 at 6:45-59. This passage underscores the clear differences between the two alternative approaches to motion sensing and smoothing and demonstrates why a POSITA would not consider the teachings from one to be easily substituted into the other.

163. Petitioner’s asserted motivation to combine is conclusory and unsupported by substantial evidence, merely suggesting—without detailed analysis or explanation—that combining Bell and Shi’s 3D sensor-based approaches with Kwatra’s 2D image-based linear optimization would yield better stabilization results.

164. Given the clear computational and functional disparities between the references, in my opinion, there is no persuasive motivation or reasonable expectation of success for a POSITA to combine Bell and Shi with Kwatra. Therefore, in my opinion, the Petition fails to demonstrate obviousness and claims 1-21 should be upheld.

VI. SECONDARY CONSIDERATIONS CONFIRM THE NON-OBVIOUSNESS OF THE ’840 PATENT

165. In my opinion, there is strong, objective evidence that the ’840 patent is not obvious. Patent Owner GoPro, Inc. (“GoPro”) successfully manufactured and sold action cameras that include the HyperSmooth functionality, which embodies the ’840 patent.

PUBLIC VERSION**A. HyperSmooth Practices the '840 Patent**

166. In my opinion, GoPro action cameras with HyperSmooth practice at least claims 1, 13 and 14 of the '840 patent. My analysis below focuses on the GoPro HERO7 Black, [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED] In my opinion, Petitioner is wrong as to claim limitation 1[g(iii)]: the HERO7 Black practices this limitation and every other limitation of claims 1, 13, and 14 of the '840 patent.

1[pre]: An image capture system that stabilizes videos, the image capture system comprising:

167. In my opinion, the HERO7 Black practices this limitation. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

168. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

PUBLIC VERSION

[REDACTED]

[REDACTED]

[REDACTED]

1[a]: a housing of an image capture device:

169. In my opinion, the HERO7 Black practices this limitation. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

1[b]: an optical element carried by the housing and configured to guide light within a field of view to an image sensor:

170. In my opinion, the HERO7 Black practices this limitation. [REDACTED]

[REDACTED]

[REDACTED]

171. [REDACTED]

[REDACTED].

1[c]: 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, 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:

PUBLIC VERSION

172. In my opinion, the HERO7 Black practices this limitation. [REDACTED]

[REDACTED]

[REDACTED]

173. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

1[d]: 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, the position output signal conveying position information that characterizes rotational positions of the housing at different moments within the capture duration:

174. In my opinion, the HERO7 Black practices this limitation. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

175. [REDACTED]

[REDACTED]

PUBLIC VERSION

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

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

176. In my opinion, the HERO7 Black practices this limitation. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

177. [REDACTED]

[REDACTED]

[REDACTED].

1[f]: 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, 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:

PUBLIC VERSION

178. In my opinion, the HERO7 Black practices this limitation. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

179. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

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:

180. In my opinion, the HERO7 Black practices this limitation. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

PUBLIC VERSION

181. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

182. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

183. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

PUBLIC VERSION

184. [REDACTED]

[REDACTED].

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 smoothed trajectory having smoother changes in the rotational positions of the housing than the trajectory:

185. In my opinion, the HERO7 Black practices this limitation. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

1[g(iii)]: wherein the determination of the smoothed trajectory based on the look-ahead of the trajectory 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

PUBLIC VERSION

capture device while causing a delay in the determination of the smoothed trajectory:

187. In my opinion, the HERO7 Black practices this limitation. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

188.

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

PUBLIC VERSION

190.

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

l[h]: 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 viewing window defining one or more extents of the visual content to be included within a punchout of the visual content:

191. In my opinion, the HERO7 Black practices this limitation. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

192.

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

PUBLIC VERSION

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

1[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, 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:

193. In my opinion, the HERO7 Black practices this limitation. [REDACTED]

[REDACTED]

[REDACTED]

194. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

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:

195. In my opinion, the HERO7 Black practices this limitation. [REDACTED]

[REDACTED]

PUBLIC VERSION

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

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:

197. In my opinion, the HERO7 Black practices this limitation. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

PUBLIC VERSION

198. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

B. Nexus

199. Here, secondary considerations—commercial success, industry praise, long-felt need, failure of others, unexpected results, and copying—are directly attributable to the unique stabilization approach claimed in the '840 patent. GoPro's HERO7 Black's HyperSmooth system leverages a weight-balance parameter to adjust the emphasis between rotational velocity and acceleration, a low-light high-pass parameter to preserve high-frequency motion in low-light conditions, and a temporal horizon to approximate intentional motion. This multifaceted optimization represented a substantial departure from prior approaches like Shi's blur-detection-based blending or Kwatra's linear image-

PUBLIC VERSION

driven smoothing, enabling GoPro to achieve stabilization praised by consumers and industry experts alike.

200. HERO7 Black embodies the inventions claimed in the '840 patent, including claims 1, 13, and 14. While Petitioner disputes that HERO7 Black meets limitation 1[g(iii)]—requiring a temporal horizon approximating intentional motion and introducing a processing delay—that argument mischaracterizes the HERO7's stabilization.

201. I have confirmed that HERO7 Black includes a temporal horizon as understood by a POSITA. Specifically, the HERO7 Black [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

202. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

PUBLIC VERSION

203. The HERO7 Black was the first GoPro product to commercialize HyperSmooth and embody the claimed inventions of the '840 patent. *See* Ex. 2024 (Lema Decl.), ¶¶ 6, 10. Its launch marked a paradigm shift for GoPro and the industry, driven by HyperSmooth's capabilities. *See id.* at ¶ 10. The resulting commercial success and industry recognition—including HERO7 Black being called one of GoPro's most important releases—were directly tied to the stabilization enabled by the '840 patent. *See id.* at ¶ 11.

204. Accordingly, the HERO7 Black's commercial success, industry praise, and copying by competitors are specifically tied to the patented features. The record thus establishes a clear nexus between the '840 patent and the secondary considerations supporting non-obviousness.

C. Commercial Success

205. I understand that GoPro's HyperSmooth technology, which practices the '840 patent, has driven significant commercial success. According to GoPro's head of product, Mr. Pablo Lema:

The release of HyperSmooth was a “shock to the market.” HERO7 Black became the fastest-selling production launch in GoPro's history at the time. Our sales jumped 170% compared to HERO6 in the same post-launch window, driven almost entirely by the introduction of HyperSmooth. From a feature standpoint, the HERO6 camera is nearly indistinguishable from the HERO7 Black with the exception of HyperSmooth.

PUBLIC VERSION

Ex. 2024 at ¶ 11.

206. Mr. Lema also cited internal survey data confirming that HyperSmooth and immersive lens options were among “the most important and valued features driving customer demand across our HERO customer base,” and that GoPro’s “product planning decisions—including resource allocation and feature prioritization—relied heavily on this data.” *Id.* at ¶19.

207. Specifically, in a [REDACTED] (Ex. 2040, also Ex. A to Lema Dec., CX-590C), “Advanced video stabilization” ranked as the top feature for the Core HEROes segment and among the top three for the market segments representing GoPro’s historical customer base. *Id.* at ¶¶ 14-19.

208. In a separate [REDACTED] (Ex. 2041, also Ex. B to Lema Dec., CX-649C), HERO11 Black buyers selected “Advanced video stabilization (HyperSmooth 5.0 with AutoBoost)” as the #1 driver of purchase—above all other features. *Id.* at ¶ 20.

209. These results, according to Mr. Lema, “reflect the role that GoPro’s patented stabilization and lens features play in driving consumer demand and purchasing behavior.” *Id.* at ¶ 21. He concluded that “the patented technologies in the ’894, ’840, ’832, and ’413 patents were directly responsible for GoPro’s

PUBLIC VERSION

sustained commercial success with HERO cameras, especially from the HERO7 onward.” *Id.*

210. GoPro’s dominant position in the U.S. action camera market—holding more share than all competitors combined—further illustrates the commercial strength of its patented technologies. Partnerships with brands like Red Bull, NHL, and Adobe underscore this market confidence. *See* Ex. 2005 (NHL Senior VP of Programming and Production said, “[O]nly GoPro has the technology to help us and our media partners showcase the beauty and intensity of hockey in new and deeper ways.”); Ex. 2008 (GoPro “named the exclusive action camera of Red Bull” mountain bike competition; GoPro athletes will capture their runs wearing HERO11 featuring “GoPro’s Emmy Award-winning HyperSmooth video stabilization”); Ex. 2009 (Vice President of Adobe Stock said, “Adding GoPro’s unique footage to Adobe Stock gives our community access to a library of immersive ‘wow’ moments”).

D. Industry Praise and Recognition

211. I understand that GoPro’s HyperSmooth technology has received widespread industry recognition, including a 2019 CES Innovation Award and a 2022 Technology and Engineering Emmy. *See* Ex. 2010 (““In Camera Sensor and Software Stabilization – Awarded to GoPro and Socionext” . . . “The technology

PUBLIC VERSION

has created the ability for camera operators and producers to obtain stabilized action footage when a subject is rapidly moving easily and economically in unstable situation.”); Ex. 2011; Ex. 2012 (“This Emmy is in recognition of the technology powering GoPro’s industry-leading HyperSmooth video stabilization”).

212. Mr. Lema confirms, “GoPro won the 73rd Technology and Engineering Emmy award specifically for its in-camera sensor and software image stabilization, introduced with HERO7.” Ex. 2024 at ¶¶ 13, 22. He also notes “a vast amount of industry praise for HyperSmooth and HERO7” in the form of customer reviews, publications, and tech blogs. *Id.* at ¶ 23. These technologies enabled “a step-change” in user experience that drove demand, growth and recognition for GoPro’s products. *Id.* at ¶ 24.

213. Third-party reviews echo this view. Analysts at JP Morgan called GoPro’s “gimbal like hypersmooth stabilization (developed in-house . . . a significant upgrade” and “key selling point” for the HERO7. Ex. 2013.; Tech reviewers wrote: “GoPro’s HyperSmooth image stabilization is the best we’ve seen on any action camera.” Ex. 2014 at 5.

214. Technical reviews of the subsequently released HERO8 Black note how HyperSmooth had already “eliminated the need for using a motorized

PUBLIC VERSION

gimbal,” and how the electronic image stabilization of HyperSmooth 2.0 was “unfathomably good.” Ex. 2042 (CX-1388) at 3.

E. Long-Felt but Unmet Need

215. Before HyperSmooth, users of action cameras struggled to capture stable footage during high-vibration activities like mountain biking, hiking, or surfing. Mechanical gimbals were often used but were bulky, fragile, and poorly suited to these dynamic conditions. The ’840 patent’s digital stabilization filled this long-felt need.

216. According to GoPro’s Mr. Lema, HyperSmooth was a “shock to the market.” Ex. 2024 at ¶ 10. Public reaction was “one of disbelief”—some users assumed GoPro had faked footage using mechanical stabilizers. But once they tried the product, “the response was overwhelmingly positive.” *Id.* at ¶ 12.

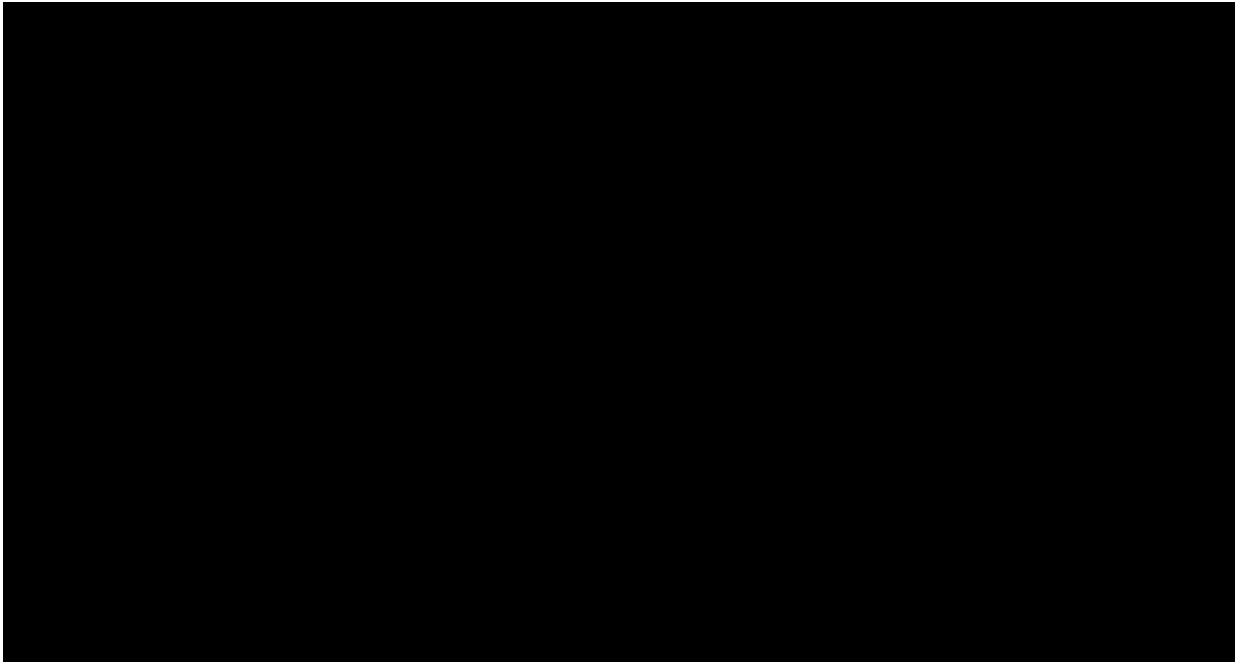
217. These outcomes reflect how HyperSmooth filled a long-felt need—further demonstrating the non-obviousness of the claimed inventions.

F. Failure of Others

218. GoPro was not the only camera maker seeking to perfect digital image stabilization in 2018. Some of the world’s greatest technology leaders, including Sony, Google, Samsung, and Apple, were all intently pursuing their own solutions. Many of these approaches involved hybrid electronic image stabilization (“EIS”) and optical image stabilization (“OIS”) solutions—as disclosed in Shi, a Google

PUBLIC VERSION

patent, Ex. 1006 at 1:36-44 and Zhou, an Apple patent, Ex. 1004 at 1:34-45—
which were widely believed to be necessary given the computational requirements
of pure EIS solutions and the limitations of mobile devices. As reflected in

**G. Unexpected Results**

220. GoPro’s system consistently outperformed traditional stabilization
techniques, which often failed under rapidly changing motion. HyperSmooth
maintained stabilization where others did not. Ex. 2014 (“GoPro’s HyperSmooth

PUBLIC VERSION

image stabilization is the best we’ve seen on any action camera”); Ex. 2024 at ¶ 11 (HyperSmooth was a “shock to the market”).

221. These unexpected results further reinforce the non-obviousness of the ’840 patent.

H. Copying by Competitors

222. I understand that in this case, Petitioner Insta360, a direct competitor, has been accused of copying GoPro’s HyperSmooth technology, underscoring the value and innovation of the ’840 patent. Ex. 2006 (Complaint).

223. I understand that beginning in 2021, Insta360 began releasing its own version of “in-camera” electronic image stabilization technology that relied upon look-ahead smoothing, optimization of the smooth camera trajectory path, and specialized tunable parameters to control the strength of smoothing. Ex. 2048 (Insta360 Speeds Up Workflow with New 360 Reframing Tool, Better Stabilization & Quick Reader). [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

- [REDACTED]

PUBLIC VERSION

- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]

224. In my opinion, these secondary considerations—commercial success, industry praise, long-felt need, unexpected results, and copying— confirm that the '840 patent's inventions were not obvious. HyperSmooth revolutionized the action camera market, drove demand, earned recognition, and prompted imitation. In my opinion, the claims should be upheld.

IV. CONCLUSION

225. In signing this declaration, I recognize that the declaration will be filed as evidence in a contested case before the Patent Trial and Appeal Board of the United States Patent and Trademark Office. I also recognize that I may be subject to

PUBLIC VERSION

cross-examination in the case and that cross-examination will take place within the United States. If cross-examination is required of me, I will appear for cross-examination within the United States during the time allotted for cross-examination.

226. I hereby declare that all statements made herein of my own knowledge are true and that all statements made on the information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Executed: April 19, 2025

/s/ Hassan Foroosh

Hassan Foroosh, Ph.D.