

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 IPR2024-01186
U.S. Patent No. 10,574,894

**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 IPR2024-01186, U.S. Patent No. 10,574,894 (“the ’894 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.

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.

LEGAL UNDERSTANDING**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. 1002, ¶¶36-37.

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 May 2018 provisional application priority date of the ’894 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 ’894 patent.

PUBLIC VERSION**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.

CLAIM CONSTRUCTION

34. In my opinion, all claim terms should be given their ordinary and customary meaning in light of the specification.

I. OVERVIEW OF THE '894 PATENT

35. The '894 patent, titled “Systems and Methods for Stabilizing Videos,” discloses a video stabilization system that determines a capture trajectory to smooth the observed trajectory of the camera by minimizing a combination of rotational velocity and rotational acceleration of the camera housing. Ex. 1001, 10:61-11:15.

Rotational Motion of the Camera Housing

36. 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).

37. While the claims do not explicitly reference **yaw**, **pitch**, or **roll**, minimizing rotational velocity and acceleration of a camera housing—a three-

PUBLIC VERSION

dimensional object in space—necessarily involves at least one of these rotational angles.¹

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

38. The '894 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.

39. 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:

¹ Furthermore, claim 5 in the '894 patent limits the position sensor in independent claim 1 to be “a gyroscope, an accelerometer, and/or an inertial measurement unit” (IMU). Both gyroscope and IMU are sensors that are used to provide yaw, pitch, roll angular measurements for a moving body in the 3D space.

PUBLIC VERSION

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

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 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.

40. Minimizing these time derivatives is one example of optimizing rotational velocity and acceleration in accordance with the '894 patent claims. By minimizing rotational velocity and acceleration of the camera housing, the '894 patent accounts for the real-world motion of a camera rotating in the 3D space, ensuring smoother video output.

PUBLIC VERSION**Sensor-Based Position Determination**

41. The '894 patent's approach to stabilization is further distinguished by its reliance on onboard motion sensors to determine camera position, independent of image data. Claim 5 explicitly requires that position information be determined "*independent of the image information*," meaning the system utilizes gyroscopes, accelerometers, and/or inertial measurement units (IMUs) to measure the camera's movement directly.²

42. This sensor-based measurement method contrasts with image-based stabilization approaches, which estimate camera motion by tracking image features across frames. *Cf.* Ex. 1004 ("Zhou") at 1:29-34 ("One way to stabilize a video is to track a salient feature in the image and use this as an anchor point to cancel out all perturbations relative to it. This approach requires a priori knowledge of the image's content to, for example, identify and track a person or other salient object in the scene."). By relying on sensor-based motion estimation rather than feature tracking from image data, the '894 patent ensures that stabilized videos appear as if

² By contrast, image-based stabilization methods estimate position by tracking visual features across frames, meaning their accuracy depends on image content rather than direct motion sensing.

PUBLIC VERSION

they were captured by a camera whose rotational velocity and rotational acceleration were minimized during the capture duration.

II. OVERVIEW OF THE CITED ART

43. The Petition challenges claims 1-20 of the '894 patent on two grounds. Ground 1 asserts that the combination of Zhou and Kwatra renders the claims obvious, while Ground 2 asserts that the combination of Zhou, Kwatra, and Shi renders the claims obvious. Below are summaries of the cited references.

B. Ground 1 Cited Art**1. Zhou (Ex. 1004)**

44. Zhou discloses a system for stabilizing video by smoothing camera motion using a “stabilization metric” derived from motion data obtained from gyroscopes or accelerometers. Ex. 1004, Abstract, 3:47-4:19. Zhou’s approach is designed for handheld cameras in low-light conditions and estimates motion blur in each frame to improve video quality. Ex. 1004, 1:5-25.

45. However, as Petitioner concedes, Zhou does not disclose or suggest minimizing a combination of rotational velocity and acceleration of the camera housing, as required by the '894 patent.

2. Kwatra (Ex. 1005)

46. Kwatra discloses a video stabilization method that estimates camera motion by tracking image features across frames and fitting *linear motion models*

PUBLIC VERSION

to the tracked feature points using a least squares method or L1 minimization. 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.

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

Kwatra's Camera Path Model Uses Linear Motion Fitting

48. 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.

PUBLIC VERSION

49. 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 $B(t)$:

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

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 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 '894 patent.

PUBLIC VERSION

50. Kwatra then 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.

51. 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.

52. 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).

PUBLIC VERSION

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

53. Kwatra's optimization may involve minimizing the cost function (Equation 7) comprised of the 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 frame), and a derivative can be determined at every point (every frame)." *Id.* at 14:8-12.

Kwatra Intentionally Avoided Non-Linear Minimization

54. 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

PUBLIC VERSION

Kwatra's provisional application (incorporated by reference in Kwatra) (Ex. 2019)

(emphasis added below):

- “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).
- “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

PUBLIC VERSION

the frame window corners, they would not move by more than the distance to the nearest crop window corner.” (*Id.* at 7).

55. 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

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

PUBLIC VERSION

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.

C. Ground 2 Cited Art

1. Shi (Ex. 1006)

57. Shi focuses on trajectory estimation and correction to stabilize videos by adjusting camera position and rotation based on motion data. The system applies: (1) a first transformation using optical image stabilization (OIS) data, and (2) then calculates a second transformation based on additional positional data collected after the capture of the frame. Ex. 1006, 3:52-4:11.

58. The '894 patent advances beyond Shi by focusing on rotational stabilization and requiring a minimization of a combination of rotational velocity and rotational acceleration of the camera housing. Therefore, it teaches determining an optimal capture trajectory that yields rotational stabilization of the camera housing, which Shi does not disclose.

PUBLIC VERSION

III. GROUND 1: PETITIONER FAILS TO ESTABLISH THAT ZHOU AND KWATRA RENDER CLAIMS 1-20 OBVIOUS**A. Claims 1 and 11: The Petition Fails to Show a Capture Trajectory “That Minimizes a Combination of Rotational Velocity and Acceleration of the [Camera] Housing”**

59. Claims 1 and 11 require that the “*capture trajectory is determined to include a path that minimizes a combination of rotational velocity and rotational acceleration of the housing.*”

60. Petitioner does not assert that Zhou itself discloses the determination of such a trajectory, instead relying on a combination with Kwatra. However, Kwatra similarly fails to disclose this limitation and, in fact, intentionally avoided the type of non-linear motion models that a minimization of *rotational* velocity and acceleration would entail.

1. *Kwatra’s Linear Framework Does Not Result In A Path That Minimizes Rotational Velocity And Rotational Acceleration of the Housing As Claimed*

61. Due to concerns regarding 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 required to optimize rotational dynamics. Thus, Kwatra’s technique is incapable of resulting in a path that minimizes a combination of rotational velocity and acceleration as required by the claim.

PUBLIC VERSION

62. 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. 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.

63. 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 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

PUBLIC VERSION

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.

64. 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 $B(t)$ are linear transformation matrices applied to each frame to generate a stabilized sequence.

Ex. 1005, 13:52-56.

65. 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:

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

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

66. 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

PUBLIC VERSION

combination of *rotational* velocity and acceleration as claimed. 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).

67. 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 one degree of freedom for scale and one degree of freedom for rotation. Ex. 1005 at 11:3-21. Kwatra, however, expressly rejects an approach that minimizes rotation or scale in its smoothing/optimization step.

68. 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 in an image, *not* rotation of the camera housing.³ As discussed below, rotation of a

³ 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

PUBLIC VERSION

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 rotation or scale as part of its smoothing/optimization step.

69. Specifically, for similarity or affine transformations, Kwatra teaches to impose smoothness constraints only on the elements *a* and *b* of its linear transformation matrix rather than directly on the degrees of freedom associated with image scale (“*s*”) and image rotation angle “*θ*” of the trackable 2D image features. *Id.* at 17:5-20. Kwatra acknowledges that this approach of constraining smoothness only on the elements *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

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

and rotation in the smoothed camera path “to ensure s and θ do not deviate too much.” Ex. 1005 at 17:19-20.

70. Thus, even setting aside the fact that the angle “ θ ” in Kwatra is referring to the image rotation (not camera housing rotations), 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.

71. 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 cosine or sine of the 2D image rotation θ are allowed to change during estimation of the smooth camera path is not the same as computing a smooth 3D camera path that minimizes a combination of rotational velocity and acceleration of the camera housing over the smoothed path.

72. 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). Thus, by deliberately avoiding non-linear computations, Kwatra does not and cannot minimize rotational velocity and acceleration of a camera housing path.

73. In forming his invalidity opinions, Petitioner’s expert 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 -- 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

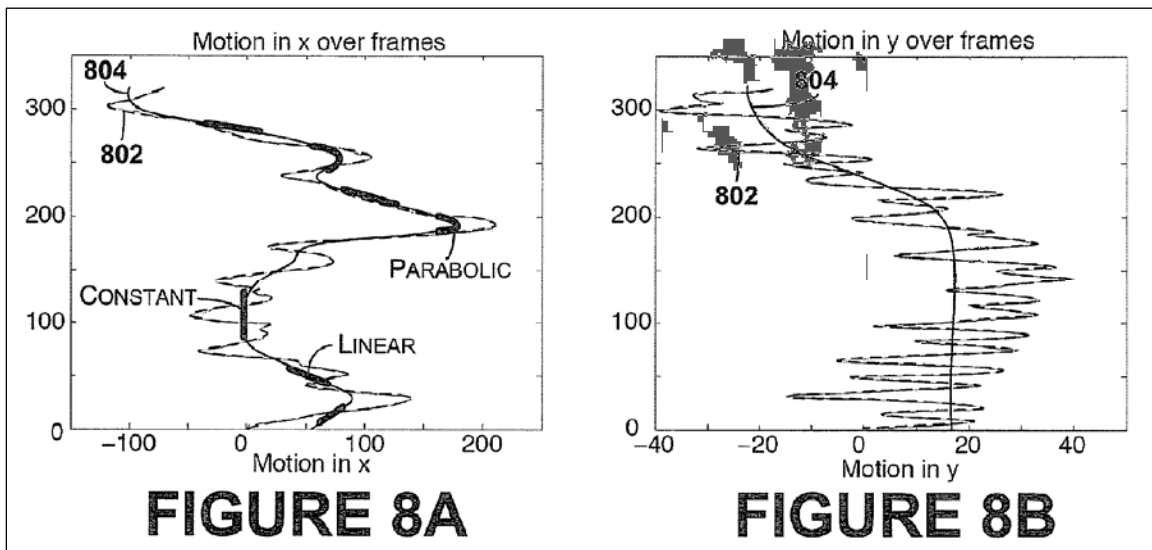
PUBLIC VERSION

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.”).

74. Further demonstrating that Kwatra fails to render the ’894 minimization requirements 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



75. Figures 8A and 8B make clear that Kwatra's aim is not to minimize 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, Trial Tr. (Goodin) at 1368:6-1369:7. And, of course, there is no teaching or suggestion in Kwatra that results in any smoothed camera position as a 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 minimization approach of the '894 patent.

PUBLIC VERSION

2. *Kwatra's Two-Dimensional Approach To Motion Estimation And Optimization Is Inconsistent With Determining A Path That Minimizes Rotational Velocity And Rotational Acceleration of the Housing As Claimed*

76. 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.”).

77. 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 mentioning of “estimates from lower to higher

PUBLIC VERSION

dimensional motion models (e.g., similarity --> homography)” at column 11, lines 58-61, are understood by a POSITA as merely referring to an estimate of a 2D model with more degrees of freedom (*i.e.*, a homography transform) instead of another 2D model with fewer degrees of freedom (*e.g.*, a similarity or an affine transformation). Again, the actual motion modeling and optimization disclosures of Kwatra that are 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.

78. 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 claimed minimization of rotational velocity and rotational acceleration of the camera housing inherently requires a three-dimensional model, Kwatra fails to disclose or render obvious the claim limitation. Petitioner’s theory fails for this independent reason.

79. 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 Capture Trajectory that Minimizes Rotational Velocity and Acceleration of a Camera Housing, as Claimed*

80. The Petition also fails to disclose the claimed limitation because it conflates minimizing derivatives of what Kwatra refers to as the smooth camera path $P(t)$ with minimizing rotational velocity and acceleration of a *camera housing* as claimed.

81. 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. Rather, it "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.

82. 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. As a camera housing moves, the extent of motion of a particular feature point across frames (measured in pixels) depends not only on the actual motion of the camera, but also on the depth or the distance of the 3D point from

PUBLIC VERSION

the camera that is imaged as that trackable feature point in the image. For example, a background feature and a foreground feature at different depths will move by a different number of pixels for a given displacement (translation and rotation) of the camera housing in the physical 3D 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.

83. 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 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

PUBLIC VERSION

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 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

PUBLIC VERSION

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.

84. 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.

85. The failure of Kwatra to teach or disclose an approach that minimizes motion of the camera housing speaks to the fundamental differences between Kwatra and the '894 patented approach. Kwatra is directed to optimizing a 2-dimensional camera path based on linear motion models extracted from trackable features. The '894 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 acceleration *of the camera* during capture. Kwatra fails to render the claim obvious for this independent reason.

PUBLIC VERSION**B. Claim 5: The Petition Fails to Show “the position information is determined independent of the image information”**

86. Claims 5 requires that the “the position sensor includes a gyroscope, an accelerometer, and/or an inertial measurement unit, *and the position information is determined independent of the image information.*”

87. As discussed above, Petitioner relies exclusively on Kwatra for the requirement that the capture trajectory minimizes the rotational velocity and acceleration of the camera housing. Kwatra’s approach to motion estimation and minimization, however, contradicts the additional limitations of dependent claim 5. Again, Kwatra extracts motion exclusively from tracked image features. Ex. 1005, 5:43-45. Kwatra’s “determination of position information” is entirely dependent on, not independent of, image information as claimed.

88. Petitioner purports to rely on Zhou’s use of a gyroscope to compensate for Kwatra’s failure to satisfy this limitation. But Petitioner cannot simply cherry-pick the disclosures from the prior art in this manner without a reasoned basis for doing so. The minimization disclosures that Petitioner relies upon from Kwatra are inextricably intertwined with how Kwatra estimates the camera path using two-dimensional linear motion models. Petitioner, for example, offers no explanation of how Kwatra’s optimization techniques (which are fundamentally based on its 2D linear motion modeling) could be implemented in a

PUBLIC VERSION

device like Zhou with a completely different estimated motion model derived from gyroscope data. One cannot simply replace P with θ in Kwatra's optimization equation (7) and assume that the rest of Kwatra's teachings would hold. Kwatra's solution to optimize (7) is linear and cannot handle such naïve substitution, and Zhou and Shi do not disclose any method to optimize the alleged modification of equation (7) in view of Zhou/Shi.

C. A POSITA Would Not Have Been Motivated to Combine Zhou (and/or Shi) and Kwatra

89. The Petition fails to account for the fundamental different image stabilization techniques employed by Zhou and Shi on the one hand, and Kwatra on the other hand. The embodiments that Petitioner relies upon from Zhou 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.

PUBLIC VERSION

90. 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 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.

91. 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.

PUBLIC VERSION

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 Zhou 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 Zhou/Shi's mobile-device based stabilization of three-dimensional rotational data. Nor is there any reason to believe that a POSITA would have had a reasonable expectation of success in attempting to perform this combination.

92. 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 Zhou. 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."). And, while not at issue in this Petition, U.S. Patent 10,027,893 ("Bell") (Ex. 2023) which is relied on by Petitioner in parallel cases for related technology, explains why a POSITA implementing a gyroscope-based system like Zhou or Shi would not be motivated to look to the image-based systems like Kwatra:

PUBLIC VERSION

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.

Ex. 2023 at 6:45-59 (emphasis added). 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.

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

94. Given the clear computational and functional disparities between the references, there is no persuasive motivation or reasonable expectation of success

PUBLIC VERSION

for a POSITA to combine Zhou and Shi with Kwatra. Therefore, in my opinion, the Petition fails to demonstrate obviousness and claims 1-20 should be upheld.

IV. GROUND 2 FAILS ON THE MERITS

95. Petitioner asserts that the combination of Zhou, Kwatra, and Shi renders claims 2, 3, 7, 12, 13, and 17 of the '894 patent obvious. However, this combination fails for the same reason as Ground 1, as Shi does not (and is not even alleged to) cure the deficiencies of Zhou and Kwatra with respect to the requirements of independent claims 1 and 11 or dependent claim 5.

V. SECONDARY CONSIDERATIONS CONFIRM THE NON-OBVIOUSNESS OF THE '894 PATENT

96. In my opinion, there is strong, objective evidence that the '894 patent is not obvious. Patent Owner GoPro, Inc. ("GoPro") successfully manufactured and sold action cameras that include the HyperSmooth functionality, which embodies the '894 patent.

A. HyperSmooth Practices the '894 Patent

97. In my opinion, GoPro action cameras with HyperSmooth practice at least claims 1 and 5 of the '894 patent. My analysis below focuses on the GoPro HERO7 Black, [REDACTED]

[REDACTED]

PUBLIC VERSION

[REDACTED]

[REDACTED]

1[pre]: A system that stabilizes videos, the system comprising:

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

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

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

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

[REDACTED]

[REDACTED]

[REDACTED]

PUBLIC VERSION

1[b]: An optical element carried by the housing and configured to guide light within an optical field of view to an image sensor, the optical field of view being greater than a capture field of view for generating video content:

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

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

1[c]: The image sensor carried by the housing and configured to generate an image output signal based on light that becomes incident thereon during a capture duration, the image output signal conveying image information that defines images with the optical field of view:

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

[REDACTED]

[REDACTED]

104. [REDACTED]

[REDACTED]

[REDACTED]

PUBLIC VERSION

[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 the positions of the housing at different moments within the capture duration, the positions of the housing including rotational positions of the housing:

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

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

106. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

PUBLIC VERSION

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

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

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

108. [REDACTED]

[REDACTED]

[REDACTED]

1[f]: Determine an observed trajectory of the housing during the capture duration based on the position information, the observed trajectory reflecting the positions of the housing at the different moments within the capture duration, the observed 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:

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

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

PUBLIC VERSION

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

1[g(i)]: Determine a capture trajectory of the housing based on a look ahead of the observed trajectory, the capture trajectory reflecting actual and/or virtual positions of the housing from which orientations of the capture field of view with respect to the optical field of view of the images are determined:

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

[REDACTED]

[REDACTED]

[REDACTED]

112. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

1[g(ii)]: The look ahead of the observed trajectory including use of a subsequent portion of the observed trajectory to determine a

PUBLIC VERSION

preceding portion of the capture trajectory such that a portion of the capture trajectory corresponding to the first portion of the observed trajectory, which corresponds to the first moment within the capture duration, is determined based on the second portion of the observed trajectory, which corresponds to the second moment within the capture duration, the capture trajectory having smoother changes than the observed trajectory:

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

[REDACTED]

[REDACTED]

[REDACTED]

114. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

1[g(iii)]: Wherein the capture trajectory is determined to include a path that minimizes a combination of rotational velocity and rotational acceleration of the housing, the combination of the rotational velocity and the rotational acceleration of the housing weighing the rotational velocity and the rotational acceleration equally or differently:

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

[REDACTED]

[REDACTED]

PUBLIC VERSION

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

l[h]: Determine the orientations of the capture field of view for the images with respect to the optical field of view of the images based on the capture trajectory of the housing:

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

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

l[i]: Generate the video content based on a punch-out of visual content of the images within the capture field of view, wherein use of

PUBLIC VERSION

the visual content of the images within the capture field of view to generate the video content provides stabilization that creates a more stable view than use of entire visual content of the images:

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

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

5: The system of claim 1, wherein the position sensor includes a gyroscope, an accelerometer, and/or an inertial measurement unit, and the position information is determined independent of the image information:

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

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

PUBLIC VERSION

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

B. Nexus

123. A number of secondary considerations—commercial success, industry praise, long-felt need, failure of others, unexpected results, and copying by competitors—are specifically attributable to the unique stabilization approach claimed in the '894 patent, notably the minimization of a combination of rotational velocity and rotational acceleration feature. This rotational optimization technique was a substantial departure from prior approaches (like Kwatra's linear image-dependent framework), enabling GoPro's HyperSmooth technology to achieve the unprecedented stabilization praised by consumers and industry experts alike.

124. I understand that it is undisputed that HyperSmooth, as implemented in GoPro's HERO7 Black camera, practices claims 1 and 5 of the '894 patent. Petitioner did not dispute and its expert at the ITC trial provided no opinion regarding the '849 patent's coverage of GoPro's domestic industry product the HERO7 Black camera. *See generally* Ex. 2021, Trial Tr. (Goodin) at 1134:9-

PUBLIC VERSION

1135:6. I have confirmed that HyperSmooth embodies each and every limitation of claims 1 and 5 of the '894 patent. *See above* ¶¶ 97-122.

125. The transformative stabilization experience provided by HyperSmooth is based on the novel features claimed in the '894 patent, including: the minimization of a combination of rotational velocity and rotational acceleration (Claim 1) and the use of gyroscope and accelerometer data to obtain position information independent of image information (Claim 5).

126. The HERO7 Black was the first GoPro product to commercialize HyperSmooth and thus the first to embody the claimed invention of the '894 patent. Ex. 2024 (Lema Decl.), ¶¶ 6, 10. The launch of the HERO7 Black marked a paradigm shift for both GoPro and the action camera industry, driven by the capabilities of HyperSmooth. *See id.* at ¶ 10. The widespread commercial success and industry recognition that followed—including HERO7 Black being heralded as one of GoPro's most important releases—were directly tied to the stabilization performance enabled by the claimed inventions of the '894 patent. *See id.* at ¶ 11.

127. I have also reviewed [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

PUBLIC VERSION

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

128. The HERO7 Black’s commercial success, industry praise, and Petitioner’s reverse engineering and copying is specifically tied to the claimed features of Claims 1 and 5. *See* Ex. 2024 (Lema Decl.), at ¶¶ 11, 21. Accordingly, the record establishes a clear nexus between Claims 1 and 5 of the ’894 patent and the secondary considerations attributable to these claimed features in the HERO7 Black.

129. In the case of GoPro’s ’894 patent, the following secondary considerations strongly support non-obviousness:

PUBLIC VERSION**C. Commercial Success**

130. I understand that GoPro’s HyperSmooth technology, which practices the claimed inventions of the ’894 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.

Ex. 2024 at ¶ 11.

131. Further, according to Mr. Lema, internal survey data confirmed 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.

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

PUBLIC VERSION

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

134. According to Mr. Lema, these survey results “reflect the role that GoPro’s patented stabilization and lens features play in driving consumer demand and purchasing behavior,” and in his view, “the patented technologies in the ’894, ’840, ’832, and ’413 patents were directly responsible for GoPro’s sustained commercial success with HERO cameras, especially from the HERO7 onward.” *Id.* at ¶ 21.

135. GoPro’s dominant position in the U.S. action camera market, holding a larger share than all competitors combined, further illustrates the commercial strength of its patented technologies. GoPro’s partnerships with leading brands like Red Bull, NHL, and Adobe reflect the substantial demand for and market confidence in GoPro’s products, driven in large part by its stabilization technologies. *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;

PUBLIC VERSION

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 that will take their audience to new places and deliver an additional level of depth and perspective to their creations”).

D. Industry Praise and Recognition

136. I understand that GoPro’s HyperSmooth technology has garnered widespread industry recognition. It earned a CES Innovation Award in 2019 and a Technology and Engineering Emmy in 2022, further confirming the value and non-obviousness of GoPro’s patented stabilization methods. *See* Ex. 2010 (“73rd Annual Technology & Engineering Emmys” . . . “In Camera Sensor and Software Stabilization – Awarded to GoPro and Socionext” . . . “The technology 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 (“GoPro Showcases Award-Winning HERO7 Cameras at CES 2019”); Ex. 2012 (“GoPro Wins Second Emmy® Award for Groundbreaking Digital Imaging Technology”; “This Emmy is in recognition of the technology powering GoPro’s industry-leading HyperSmooth video stabilization”).

PUBLIC VERSION

137. As Mr. Lema explains in his declaration, “GoPro won the 73rd Technology and Engineering Emmy award specifically for its in-camera sensor and software image stabilization, introduced with HERO7.” Ex. 2024, Lema Dec. at ¶¶ 13, 22. GoPro also received “a vast amount of industry praise for HyperSmooth and HERO7 generally in the form of customer reviews, industry publications, tech blogs over the years.” *Id.* at ¶ 23. According to Lema, these technologies enabled “a step-change” in user experience, which in turn drove consumer demand, sales growth and “recognition for GoPro’s products.” *Id.* at ¶ 24.

138. Third-party reviews consistently highlight HyperSmooth as a groundbreaking feature that sets GoPro apart from competitors, with praise for its ability to deliver stable, smooth video footage in challenging conditions. Ex. 2013 (analysts at JP Morgan stating GoPro’s “gimbal like hypersmooth stabilization (developed in-house software and hardware)” was “a significant upgrade” when introduced and was the “key selling point” in the HERO7 camera); Ex. 2014 at 5 (“GoPro’s HyperSmooth image stabilization is the best we’ve seen on any action camera”).

139. 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

140. Before the introduction of HyperSmooth, users of action cameras faced persistent challenges in capturing stable footage during high-vibration activities like mountain biking, hiking, or surfing. Mechanical gimbals, while used for stabilization, were bulky, fragile, and unsuitable for dynamic environments of such applications. This long-standing issue, now addressed by the digital stabilization system claimed in the ’894 patent, underscores the invention's non-obvious nature.

141. According to GoPro’s head of product, Mr. Lema, the release of HyperSmooth was a “shock to the market.” Ex. 2024 at ¶ 10. The initial public reaction was “one of disbelief at the quality of the stabilization technology—some assumed we had faked the footage using mechanical stabilizers. But once users got the product in hand, their response was overwhelmingly positive.” *Id.* at ¶ 12.

142. The success of these advanced methods filled a long-felt gap in the market, further demonstrating the non-obviousness of the claimed inventions.

F. Failure of Others

143. GoPro was not the only camera maker seeking to perfect digital image stabilization in 2018. Some of the world’s greatest technology leaders, including

PUBLIC VERSION

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 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

[REDACTED]

[REDACTED]:

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED] The failure

PUBLIC VERSION

of others in the industry to achieve HyperSmooth-level results, particularly in a software-only solution, is strong evidence of non-obviousness.

G. Unexpected Results

144. GoPro’s system consistently outperformed traditional stabilization techniques, which often failed to maintain stabilization under rapidly changing motion characteristics, showcasing the unexpected success of the HyperSmooth system over prior technologies. Ex. 2014 (“GoPro’s HyperSmooth image stabilization is the best we’ve seen on any action camera”); Ex. 2024 at ¶ 11 (introduction of HyperSmooth was a “shock to the market”).

145. These unexpected results further reinforce the non-obviousness of the ’894 patent.

H. Copying by Competitors

146. 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).

147. 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

PUBLIC VERSION

Stabilization & Quick Reader). Additionally, [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

- 1 [REDACTED]
- 2 [REDACTED]
- 3 [REDACTED]
- 4 [REDACTED]
- 5 [REDACTED]
- 6 [REDACTED]
- 7 [REDACTED]
- 8 [REDACTED]
- 9 [REDACTED]

148. 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

PUBLIC VERSION

camera market, drove demand, earned recognition, and prompted imitation. In my opinion, the claims should be upheld.

CONCLUSION

149. 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 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.

150. 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 18, 2025

/s/ Hassan Foroosh
Hassan Foroosh, Ph.D.