

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 IPR2025-00017
U.S. Patent No. 10,529,052

**DECLARATION OF HASSAN FOROOSH PH.D.
IN SUPPORT OF PATENT OWNER'S PRELIMINARY 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 Preliminary 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 IPR2025-00017, U.S. Patent No. 10,529,052 (“the ’052 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, 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 Research

PUBLIC VERSION

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

PUBLIC VERSION

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 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

PUBLIC VERSION

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.

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*

PUBLIC VERSION

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:

- 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.

PUBLIC VERSION

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 Electrical Engineering, Computer Science, Physics, Optics, or an equivalent field, and one to two years of academic or industry experience working on image or video processing, or comparable academic or industry experience.” Ex. 1002, ¶¶17-18.

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 20, 2015, the priority date of the ’052 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 ’052 patent.

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.

PUBLIC VERSION1. 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 each 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.

2. 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 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

PUBLIC VERSION

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

3. *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

PUBLIC VERSION

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.*, “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.

4. *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

PUBLIC VERSION

“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 as a whole and not just the new claimed 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. In my opinion, all claim terms should be given their ordinary and customary meaning in light of the specification.

I. OVERVIEW OF THE '052 PATENT

35. The '052 Patent describes a system and method for simulating a virtual lens effect by selecting sub-frames from different positions of areas of the input images and processing them to achieve desired lens distortions, centered in

PUBLIC VERSION

the reduced fields of view of the sub-frames and consistent with the input lens distortion. The system accepts “input images . . . including fields of view of a scene captured through a lens of a camera, the input images depicting the scene with an input lens distortion centered in the fields of view.” Ex. 1001, 13:28–31. Sub-frames are then selected, representing “portions of the input images,” with “reduced fields of view of the scene smaller than the fields of view of the input images.” *Id.*, 13:32–35. These sub-frames include “different lens distortion effects as a function of the input lens distortion,” which vary based on “different positions of areas of the input images . . . and a size of the areas of the input images.” *Id.*, 13:35–40.

36. The system further processes the sub-frames to remap the input lens distortion to a “desired lens distortion centered in the reduced fields of view,” ensuring that the output images maintain coherent distortion effects. *Id.*, 13:42–54. This allows the output images to “simulate the same effect that would have been achieved by a camera operator manually re-orienting and/or physically moving the camera.” *Id.*, 3:24–28.

37. Applications of the '052 Patent include virtual zooming, object tracking, and panning effects. For instance, sub-frames may be selected from different regions of an input image to track a moving subject or to create zoom

PUBLIC VERSION

effects. *Id.*, 1:32–37; 3:16–28. By processing the sub-frames to maintain consistent lens characteristics, the invention enables seamless and realistic transitions between frames in the output video or image sequences.

38. In summary, the '052 Patent provides a transformative solution for producing high-quality images and videos by selecting and processing sub-frames from input images, ensuring that output images exhibit coherent and realistic lens distortion effects across selected sub-frames and with the input lens distortion.

II. OVERVIEW OF CERTAIN CITED ART

39. The petition challenges the '052 Patent's claims 1 to 20 on three grounds. Ground 1 asserts that Okubo renders the claims obvious. Ground 2 asserts that the combination of Okubo and Sokeila also renders the claims obvious. Ground 3 asserts that Eder renders the claims obvious. Below are short summaries of these references.

A. Ground 1 Cited Art**1. Okubo (Ex. 1004)**

40. Okubo is directed to an “image pick-up apparatus and image restoration method.” Ex. 1004, Title. It discloses two distinct modes: a “moving image” or “video mode” and a “still image” mode, controlled by a mode switch. *Id.*, Fig. 7, [0093] (“it is judged whether a mode is a still or moving image mode (S106), and the process shifts to S120 of FIG. 8 in the moving image mode in

PUBLIC VERSION

which the mode key 7 is positioned on the M-side”). In the still image mode, processing continues through blocks 122 and 123 of the circuit. *Id.*, [0095] (“After picking up the still image, the resultant image is processed by the image processing circuit 117,” and discussing use of circuits 122 and 123). In the moving image mode, processing follows a separate through blocks 131 and 132. *Id.*, [0096] (“a main flowchart for the moving image will be described with reference to FIG. 8,” and describing use of blocks 131 and 132).

41. Okubo’s “two-stage lens distortion correction process,” described in connection with Fig. 12, applies exclusively to still images. *Id.*, [0013], [0055]. The process includes a single step, step S301, to “pick up image.” *Id.*, Fig. 12, [0113]. No intermediate step is disclosed in which a larger image is recorded, and a smaller sub-frame is later selected.

42. In still image mode, the apparatus “picks up” only the portion of the sensor’s range necessary for its further processing. For example, Okubo states that “the picked-up image in this *image pick-up range* [i.e., the cutout range] is subjected to the electronic vibration correcting, and *recorded*.” *Id.*, [0088] (emphasis added). This “image cutout” is described as the recorded image. *See id.*, [0090] (“range to be picked up and recorded”), [0091] (same). Okubo confirms that “picked up” refers to the image captured directly by the sensor: “A light beam

PUBLIC VERSION

passed through the lenses **12, 13, 14** is formed into an image on the image pick-up device **114**, and photoelectrically converted by each pixel of the image pick-up device. Accordingly, the ***image is picked up.***” *Id.*, [0072] (emphasis added).

43. Finally, Okubo’s image cutouts are fixed in size and remain centered on the CCD sensor. For example, in the moving image mode, the cutout corresponds to 70% of the sensor’s diagonal length, while in the still image mode, it corresponds to 95%. *Id.*, [0088]-[0089], Fig. 6B.

B. Ground 2 Cited Art

1. Sokeila (Ex. 1005)

44. Sokeila discloses a system for “motion adaptive cropping for video stabilization,” which dynamically adjusts the cropping factor of video frames to stabilize images. Ex. 1005, Title, Abstract. The system determines “an amount of motion in a second frame of video with reference to a first frame of video” and uses this information to “[a]daptively predict[] a cropping factor for the second frame.” *Id.*, Abstract. The cropping factor specifies the portion of the frame to exclude, with the goal of “providing a stable video image with a maximized field-of-view and a maximized level of quality.” *Id.*, 5:25-30. Adjustments are based on motion analysis, allowing for a “smooth change for the cropping factor over time (to avoid bad user experience).” *Id.*, 5:25-34.

PUBLIC VERSION

45. Sokeila focuses exclusively on global camera motion-based cropping adjustments for stabilization and does not involve or disclose any processing related to lens distortion correction or remapping. The system's adaptive cropping prioritizes preserving important content, noting that "if there is a face (or another object of interest) detected in the received image frames, it may be desired to avoid cropping the object altogether." *Id.*, 3:32-35. These features aim to maintain content visibility while minimizing reductions in the field of view. *Id.*, 5:25-34. Additionally, the system supports stabilization "either during the capture or after the capture as an editor feature." *Id.*, 3:65-67.

C. Ground 3 Cited Art**1. Eder (Ex. 1009)**

46. Eder discloses a "rear view mirror simulation" system for vehicles that replaces or supplements traditional mirrors with a camera-based system. Ex. 1009, Title. The system includes an optical sensor that captures "wide-angle video images," which are processed by an "image rectification device [that] modifies the continuous images and transmits the improved images by video data to the display device 20, in order to help the driver." *Id.*, 3:14-16, 3:67-4:2. This rectification addresses "visual distortion, which is caused by a wide-angle lens used in the camera," and uses a mathematical model to adjust pixel positions. *Id.*, 4:15-25.

PUBLIC VERSION

47. The rectified images are further processed to simulate traditional mirror appearances, with Eder explaining that the system “distorts the image of the wide-angle camera and applies post-distortion to this image, in order to give this image the same view as that of the desired mirror glass.” *Id.*, 4:4-7. This includes simulating mirrors with “an aspheric curve” or “a convex additional mirror,” depending on the specific mirror type being replicated. *Id.*, 4:66-5:6.

48. Eder describes that its system processes these rectified images to display them on a screen, providing the driver with a view that is “adjusted depending on the application of the vehicle.” *Id.*, 5:23-26. The system uses fixed parameters for processing, including a “defined number of pixels” for specific mirror effects. *Id.*, 5:6-8, 5:33-35. Eder also discloses asymmetric and radially asymmetric displays, such as those depicted in Figure 2A, for simulating rear-view mirrors with customized shapes. *See id.*, Fig. 2A.

III. GROUND 1: PETITIONER FAILS TO ESTABLISH THAT OKUBO RENDERS CLAIMS 1-20 OBVIOUS

A. Okubo Does Not Disclose “Selecting Sub-Frames Representing Portions Of The Input Images, The Sub-Frames Including Reduced Fields Of View Of The Scene Smaller Than The Fields Of View Of The Input Images”

49. Claims 1 and 11 require “accessing input images, the input images including fields of view of a scene captured through a lens of a camera,” and then “selecting sub-frames representing portions of *the* input images, the sub-frames

PUBLIC VERSION

including reduced fields of view of the scene smaller than the fields of view of the input images.” Ex. 1001, 13:28-35, 14:18-25. The claims therefore plainly require a first image, the “input image,” and then selecting sub-frames that represents a smaller portion of that image.

50. In my opinion, Okubo does not disclose selecting sub-frames from any image—rather, it “picks up” only that portion of the sensor’s range necessary for its further processing. The Petition relies on Okubo’s disclosure of an “image cutout” as representing the sub-frame. Pet., 9, 20. However, Okubo discloses that this “image cutout” does not represent a portion of a larger input image, but rather is the size of the image that is “picked up” and recorded in the first instance. That is, Okubo does not use the entirety of the sensor, but only a portion that it needs, and that portion of the sensor that is used is the image cutout, which is the input that is recorded. For example, Okubo discloses that “the picked-up image in this *image pick-up range* [i.e., the cutout range] is subjected to the electronic vibration correcting, and *recorded*.” Ex. 1004, [0088] (emphasis added). Subsequent paragraphs also refer to the “image cutout” as the image that is recorded in the first instance. *See id.* [0090] (“range to be picked up and recorded”), [0091] (same).

51. Prior disclosures in Okubo confirm that Okubo uses the term “picked up” to refer to the image that is taken in the first instance. For example, paragraph

PUBLIC VERSION

[0072] states that a “light beam passed through the lenses **12, 13, 14** is formed into an image on the image pick-up device **114**, and photoelectrically converted by each pixel of the image pick-up device. Accordingly, the *image is picked up.*” *Id.*

(emphasis added)

52. The flow-charts relied upon in the Petition further confirm that there is no two step process in which an image that represents 100% of the CCD is recorded, and then a sub-frame is taken from that image. The Petition relies on the “two stage lens correction process” of Fig. 12. Pet, 10, 13, 15, 20, 27-29. That process shows only a single step, step S301, to “pick up image.” Ex. 1004, Fig. 12. The corresponding discussion contains no disclosure of any alleged selection of sub-frames from a larger image. Ex. 1004, [0113]. There is no disclosure or discussion of picking up a 100% size image and then selecting a portion of that as the image cutout. This description confirms the discussion above that, rather than recording a full-size image and selecting a crop thereof, Okubo simply records only that portion of the sensor it requires.

53. The Petitioner euphemistically argues that the image cutout is a sub-frame representing part of some “CCD image.” Pet 9, 20. But the CCD is the *sensor*, not an image itself. Ex. 1004, [0072]. The ’052 patent recognizes the

PUBLIC VERSION

distinction between a sensor, and an image recorded by the sensor. Ex. 1001, 7:23-26 (“image sensor 414 . . . captures images”).

B. Okubo Does Not Disclose “Selecting Sub-Frames Representing Portions of the Input Images” with “Different Lens Distortion Effects” Based on “Different Positions” or “Sizes” of the Areas

54. Claims 1 and 11 also require:

selecting sub-frames representing portions of the input images, the sub-frames including reduced fields of view of the scene smaller than the fields of view of the input images, *the sub-frames including different lens distortion effects as a function of the input lens distortion present in the input images, different positions of areas of the input images included within the reduced fields of view, and a size of the areas of the input images included within the reduced fields of view.*

Ex. 1001, 13:32-41; 14:22-32 (emphasis added).

55. The Petition, at most, appears to argue that “the sub-frames include[e] different lens distortion effects as a function of the input lens distortion present in the input images.” In my opinion, it fails to address at all the requirements that the sub-frames be a function of “different positions” and a function of “size” of the sub-frames included in the reduced fields of view.

56. The ’052 patent discloses that “output images may include zooming and/or panning effects in which a reduced field of view image may be extracted which may be of *varying size and location* in different images.” Ex. 1001, 2:46-

PUBLIC VERSION

49. That is the effect of this claim limitation, which states that the distortion in the sub-frames is a function of, not only the input lens distortion in the input image, but also of the size and position of those sub-frames. The claim language directed to “different positions of areas of the input images included within the reduced fields of view, and a size of the areas of the input images included within the reduced fields of view” would have no patentable weight otherwise.

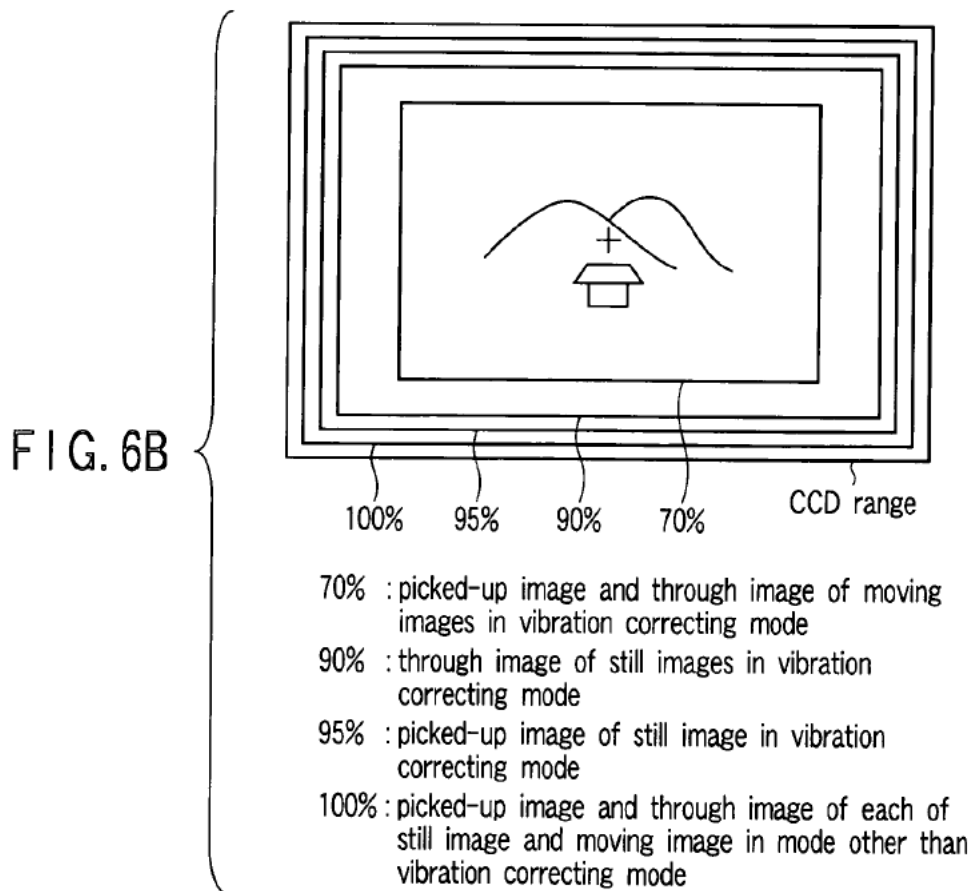
57. In my opinion, the Petition’s conclusory analysis only confirms that Okubo does not disclose this limitation. It states that a “POSITA would further understand the sub-frames to include different lens distortion effects as a function of different positions of areas of the input images included within the reduced fields of view, and a size of the areas of the input images included within the reduced fields of view” because “as described above for FIGs. 11D and 11E, the lens distortion effects in an input image—and thus a sub-frame of that image—are expressed as a function of the position and/or size (i.e., image height d) of an area within the input image.” Pet., 26. But what the Petition is describing there is not different positions and sizes of the “areas of the input images included within the reduced fields of view,” i.e., of the location and size of the sub-frames within the input. The Petition there is merely restating its position that there is varying distortion in the input image.

PUBLIC VERSION

58. In fact, in my opinion, this argument in the Petition confirms that Okubo fails to disclose this limitation. It states that the distortion effects are a function of position and size of “an area” within the input image. *Id.* But that is not what the claim requires—the claim requires that the distortion effects are a function of the position and the size of the “areas of the input images included within the reduced fields of view,” i.e. the area captured by the sub-frame. The Petition fails to even allege that this limitation is met.

59. In my opinion, the Petition’s failure in this regard is no accident. The alleged “sub-frames” that the Petition points to, the image cutouts, are both fixed in size in any given mode on which the Petition relies (*see* Pet., 9, relying on the 70% cutout range associated with the “moving image” mode), and remain centered on the CCD sensor, as shown in Fig. 6B:

PUBLIC VERSION



60. In other portions of the Petition, Petitioner in fact agrees that the alleged sub-frame size in Okubo is fixed. Pet., 55 (Okubo “has fixed cropping factors that are based on metadata”).

IV. GROUND 2: PETITIONER FAILS TO ESTABLISH THAT OKUBO IN VIEW OF SOKEILA RENDERS CLAIMS 1-20 OBVIOUS

A. Sokeila Does Not Cure Okubo’s Deficiencies

61. Petitioner relies on Sokeila, a reference that Petitioner correctly notes is related to “stabilizing video” (Pet., 54), for its disclosure of adaptive cropping of

PUBLIC VERSION

video on a frame-by-frame basis. *See* Ex. 1005 at Abstract (“Determining an amount of motion in a second frame of video with reference to a first frame of video. Adaptively predicting a cropping factor for the second frame in response to the determined amount of motion.”).

62. But even if Okubo were modified to apply Sokeila’s adaptive cropping of video frames, in my opinion, Sokeila cannot remedy the deficiencies of Okubo identified with respect to Ground 1. That is because the “two-stage lens distortion correction process” Petitioner relies on (and indeed requires) to demonstrate created output images with “desired distortion,” is limited to *still* images—Okubo does not use that process for “moving images,” i.e., video. Thus, although Okubo modified in view of Sokeila may allegedly modify the size and position of sub-frames from frame to frame of a video, the alleged “desired lens distortion” would never be applied to such sub-frames.

63. Okubo discloses two distinct modes, a “moving image” or “video mode,” and a “still image” mode, that is controlled by a mode switch. Ex. 1004, Fig. 7, [0093] (“it is judged whether a mode is a still or moving image mode (S106), and the process shifts to S120 of FIG. 8 in the moving image mode in which the mode key 7 is positioned on the M-side”). In the still image mode, processing continues to, e.g., blocks 122 and 123 of the circuit. Ex. 1004, [0095]

PUBLIC VERSION

(“After picking up the still image, the resultant image is processed by the image processing circuit 117,” and discussing use of circuits 122 and 123).

64. In the moving image mode, processing takes a different path and continues to different circuits, i.e. circuits 131 and 132. Ex. 1004, [0096] (“a main flowchart for the moving image will be described with reference to FIG. 8,” and describing use of blocks 131 and 132).

65. The Petition relies exclusively on the “two-stage lens distortion correction process,” described in connection with Fig. 12, for the “desired lens distortion.” Pet., 27-28. However, this process *only applies to still images, not video*. That is evinced by the fact that Okubo discloses the alleged desired distortion being applied in circuit 123 which, as mentioned previously, is only in the still image mode. Ex. 1004, [0013]. The Petition’s analysis of other limitations confirms that the Petition is relying on the still image path for its mapping. For example, in connection with claim 2, the Petition relies on Okubo’s disclosure of “correction of the vibration in a still image.” Pet., 35, Ex. 1004, [0055].

66. In my opinion, Petitioner’s proposed modification of Okubo in view of Sokeila therefore does nothing to address the deficiencies of Okubo, in that

PUBLIC VERSION

Sokeila would not modify the still image mode of Okubo that Petitioner relies on for its “two-stage” distortion correction process.

III. GROUND 3: PETITIONER FAILS TO ESTABLISH THAT EDER RENDERS CLAIMS 1-20 OBVIOUS

A. Eder Does Not Disclose “Selecting Sub-Frames Representing Portions of the Input Images,” Where “The Input Images Including Fields Of View Of A Scene Captured Through A Lens Of A Camera, The Input Images Depicting The Scene With An Input Lens Distortion Centered In The Fields Of View”

67. In my opinion, Eder does not disclose selecting sub-frames representing portions of the input images, where the input images include fields of view captured through a lens of the camera.

68. The Petition plainly relies on “area A” in Fig. 8 as the alleged “sub-frame,” and “B” as the input image. Pet. 75-76 (“as annotated below with respect to Fig. 8, area A represents a subset, and thus a portion, of the total pixels B”).

This is demonstrated through Petitioner’s annotation of Fig. 8:

PUBLIC VERSION

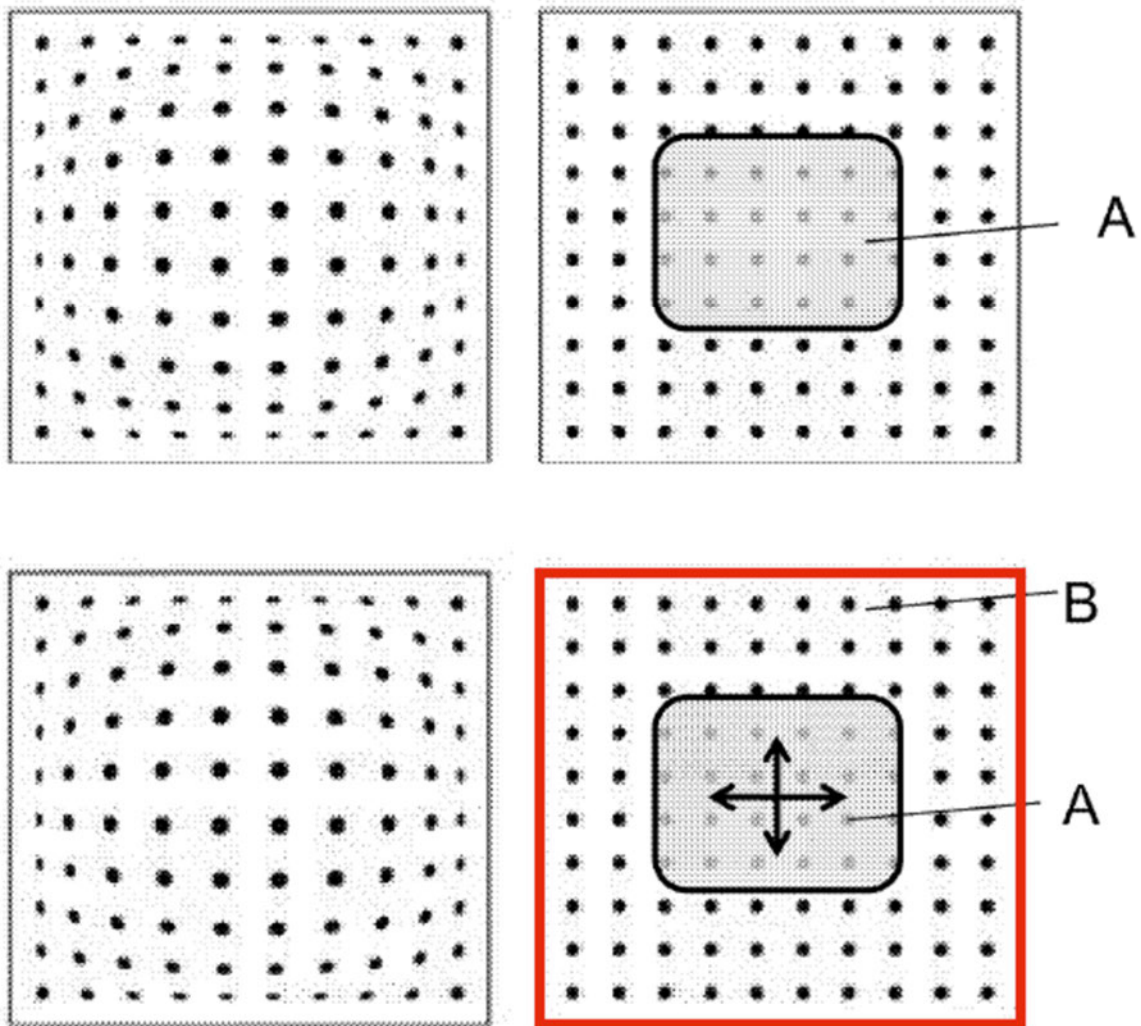


Fig. 8

Pet., 76 (annotating Ex. 1009, Fig. 8).

69. However, the images on the right (including area B that Petitioner identified as the input image) is *not* an “input image[] including fields of view of a scene captured through a lens of a camera.” The right images in Fig. 8 are not the

PUBLIC VERSION

images that were captured through the lens of the camera, but are rather “rectified” images. The images on the *left* of Fig. 8 represent images captured through the lens of the camera. But the Petition does not allege that those images have sub-frames selected from them. The Petition in fact acknowledges that the images on the left are the input images, as it pointed to the *left* images as those showing the “input lens distortion” that is “centered in the field of view.” Pet., 74. The images on the right are not distorted, and thus plainly not the claimed “input images” having input lens distortion.

B. Eder Does Not Disclose “Selecting Sub-Frames Representing Portions of the Input Images” with “Different Lens Distortion Effects” Based on “Different Positions” or “Sizes” of the Areas

70. Claims 1 and 11 of the '052 Patent also require: “selecting sub-frames representing portions of the input images, the sub-frames including reduced fields of view of the scene smaller than the fields of view of the input images, the sub-frames including different lens distortion effects as a function of the input lens distortion present in the input images, different positions of areas of the input images included within the reduced fields of view, and a size of the areas of the input images included within the reduced fields of view.” Ex. 1001, 13:32-41; 14:22-32.

PUBLIC VERSION

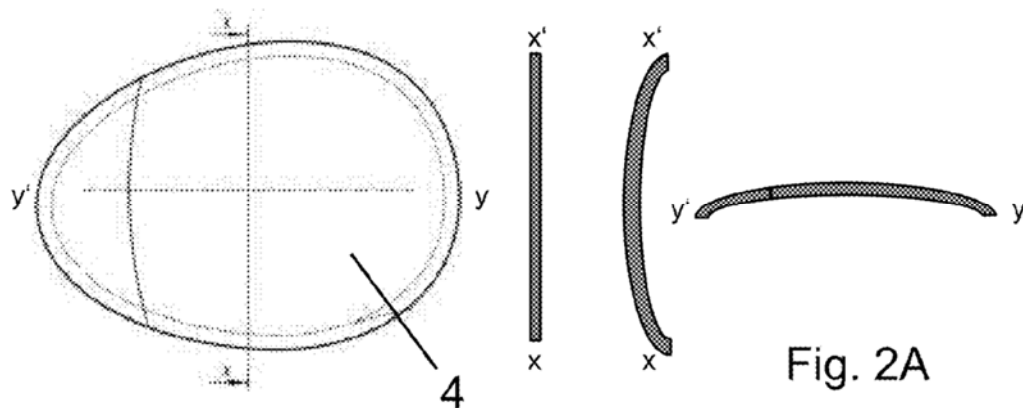
71. Eder discloses that its alleged sub-frames are always fixed in size. Area A in Fig. 8 relates to what Eder calls the “plane mirror. Ex. 1009 at 5:8-9 (“it is area A which shows plane surfaces in the middle of the optical sensor”). Eder plainly discloses that the size of area A is fixed. For example, it states that “a defined number of pixels is chosen for the display of the plane mirror surface.” Ex. 1009 at 5:7-8. It similarly states again that the “number of pixels, and therefore the size of the section, is not changed.” *Id.* at 5:33-35.

C. Eder Does Not Disclose “The Output Images Include The Sub-Frames Remapped From The Input Lens Distortion Centered In The Fields Of View Of The Input Images To The Desired Lens Distortion Centered In The Reduced Fields Of View”

72. Claims 1 and 11 require that the output images include the sub-frames remapped to the desired lens distortion *centered in the reduced field of view*. The Petition never once alleges that the desired lens distortion is centered in the reduced fields of view. *See* Pet., 84 *et seq.*

73. There is good reason why that would be the case. Eder discloses an asymmetric, and radially asymmetric display for use as a simulated rear-view mirror. *See* Ex. 1009, Fig. 2A:

PUBLIC VERSION



74. Although Eder also discloses a mirror with an ovoid shape (Fig. 2B), in my opinion, Eder is entirely lacking in any disclosure about how the alleged desired distortion is applied to these shapes.

75. Petitioner's expert similarly fails to even allege that Eder discloses that the output images include the alleged desired distortion *centered* in the reduced fields of view. Ex. 1002 at 95-98.

V. **SECONDARY CONSIDERATIONS CONFIRM THE NON-OBVIOUSNESS OF THE '052 PATENT**

76. In my opinion, there is strong, objective evidence that the '052 Patent is not obvious. Patent Owner GoPro, Inc. ("GoPro") successfully manufactured and sold wearable action cameras that include GoPro's novel virtual lens technology, which I understand GoPro asserts is an embodiment of the '052 Patent. Ex. 2006 (Complaint), ¶21.

PUBLIC VERSION

77. I have reviewed GoPro documents describing its virtual lens technology. In my opinion, there is a nexus between the virtual lens technology claimed by the '052 Patent and the secondary considerations of non-obviousness.

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]. Ex. 2015 at 2-3.

78. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]. *Id.* at 4, 6. [REDACTED]

[REDACTED]

[REDACTED]. *Id.* at 8, 10, 13. [REDACTED]

[REDACTED]

[REDACTED]. *Id.* at 14-16, 19-21.

PUBLIC VERSION

79. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]. *Id.* at 23-24. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]. *Id.* at 27-28.

80. I have seen evidence of commercial success, specifically a press release indicating GoPro has sold over 50 million action cameras since 2009, with its HERO series cameras, featuring virtual lens technology, at the forefront of this success. *See* Ex. 2003 (“GoPro Celebrates Surpassing 50 Million Cameras Sold”).

[REDACTED]

[REDACTED]

[REDACTED]. *See* Ex. 2004 ([REDACTED]).

81. 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

PUBLIC VERSION

Red Bull, NHL, and Adobe reflect the substantial demand for and market confidence in GoPro's products, driven in large part by its patented 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; GoPro athletes will capture their runs wearing HERO11); 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").

82. This success is closely linked to specific features of GoPro's patented technologies, such as GoPro's virtual lens technology described in the '052 Patent.

83. I have seen evidence of long-felt, but unmet need. The '052 Patent addresses a long-standing problem in image and video processing: how to remap lens distortions in cropped images or video while maintaining consistent distortion characteristics. Before the '052 Patent, cropping or zooming into wide-angle images often produced inconsistencies in distortion effects, resulting in a disjointed viewing experience. Existing methods failed to preserve the appearance that the

PUBLIC VERSION

cropped or zoomed frame had been captured independently by a virtual lens, making it difficult to simulate seamless camera movements or object tracking.

84. The '052 Patent solves this problem by introducing a method of remapping lens distortions across sub-frames, ensuring that cropped or zoomed sub-frames maintain distortion characteristics consistent with the original images. This innovation enables virtual lens simulation, creating the impression that a camera has physically moved to reframe a scene, even when it has not. The claimed method eliminates the undesirable frame-to-frame variation in distortion effects, enhancing the usability of cropping and zooming functions for wide-angle or spherical cameras. It further enables the selection of sub-frames from distinct regions of the input image, ensuring seamless transitions in scenarios like object tracking or dynamic zooming.

85. I have seen evidence of unexpected results. For example, the '052 Patent yields unexpected results by enabling realistic manipulation of sub-frames while preserving consistent lens distortion effects across the output images and with the original images. Prior to the '052 Patent, methods for processing cropped or zoomed frames often led to visible inconsistencies in lens distortion, particularly when sub-frames were selected from different regions of the input images. This lack of consistency hindered the ability to produce visually cohesive outputs and

PUBLIC VERSION

disrupted the viewer's experience during tasks such as object tracking or virtual zooming.

86. The '052 Patent addresses these challenges by enabling the selection of sub-frames with varying lens distortion effects—based on their size and position within the input image—and ensuring that output images maintain a cohesive and natural appearance. This capability was unexpected in the industry, as prior art systems were incapable of seamlessly combining sub-frames from different regions while preserving a consistent distortion effect. The ability to simulate smooth panning, zooming, and object tracking across such diverse sub-frames represents a breakthrough in wide-angle imaging technology.

87. By achieving these results, the '052 Patent provides a transformative advancement over the prior art and demonstrates unexpected benefits that further support the non-obviousness of the claimed invention.

88. I have seen evidence of possible copying, as I understand in an ITC investigation, Insta360, a direct competitor, is accused of copying GoPro's virtual lens technology described in the '052 Patent.

IV. CONCLUSION

89. 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

PUBLIC VERSION

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.

90. 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: January 30, 2025

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