UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

JUMIO CORPORATION,
Petitioner

v.

FACETEC, INC.,
Patent Owner.

IPR2025-00109

U.S. Patent No. 11,874,910

PETITION FOR *INTER PARTES* REVIEW OF U.S. PATENT NO. 11,874,910

Mail Stop PATENT BOARD
Patent Trial and Appeal Board
U.S. Patent & Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450

TABLE OF CONTENTS

TABLE OF CONTENTS
(continued)

TABLE OF CONTENTS
(continued)

Page

TABLE OF CONTENTS
(continued)

TABLE OF CONTENTS
(continued)

Page

TABLE OF CONTENTS
(continued)

Page

EXHIBIT LIST

Exhibit	Reference
1001	U.S. Patent No. 11,874,910 ("'910 Patent")
1002	File History of the '910 Patent ("'910 File History")
1003	Declaration of Dr. Chris Daft
1004	Dr. Chris Daft Curriculum Vitae
1005	U.S. Patent 8,437,513 ("Derakhshani")
1006	U.S. Patent Application Publication 2011/0299741 ("Zhang")
1007	U.S. Patent Application Publication 2002/0113884 ("Tanii")
1008	U.S. Patent Application Publication 2014/0028823 ("Tahk")
1009	U.S. Patent Application Publication 2004/0239799 ("Suzuki")
1010	U.S. Patent Application Publication 2010/0014720 ("Hoyos")
1011	U.S. Patent 9,077,891
1012	U.S. Patent 8,965,064
1013	Zhengyou Zhang et al., *A Robust Technique for Matching Two Uncalibrated Images Through the Recovery of the Unknown Epipolar Geometry*, Institut National De Recherche En Informatique Et En Automatique (May 1994)
1014	U.S. Patent Application Publication 2010/0158319
1015	U.S. Patent Application Publication 2007/0127787
1016	U.S. Patent Application Publication 2004/0036574
1017	Eugene Hect & Alfred Zajac, *Optics*, Addison-Wesley Publishing Co. (1974)
1018	Anil K. Jain et al., *Introduction to Biometrics*, Springer (2011)
1019	B. Honlinger & H.H. Nasse, *Distortion*, Zeiss (Oct. 2009)
1020	Hue and Hatchet, *What is the Focal Plane?*, https://hueandhatchet.com/what-is-the-focal-plane/
1021	*Introduction to Astronomy*, https://physics.weber.edu/palen/phsx1040/lectures/ldistcomp.html

Exhibit	Reference
1022	Daniel Baker, *Face distortion is not due to lens distortion*, (May 5, 2012), https://bakerdh.wordpress.com/2012/05/05/face-distortion-is-not-due-to-lens-distortion/
1023	Richard Hartley & Andrew Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge Univ. Press (2nd ed. 2011)
1024	Daniel Moreno & Gabriel Taubin, *Simple, Accurate, and Robust Projector-Camera Calibration*, Brown Univ., mesh.brown.edu/calibration/
1025	Serge Belongie, *CSE 252B: Computer Vision II, Lecture 4: Planar Scenes and Homography* (Apr. 7, 2004)
1026	Merriam-Webster's Collegiate Dictionary (11th ed. 2014)
1027	Laurenz Wiskott, *Face Recognition by Elastic Bunch Graph Matching* (1999)
1028	Axis Communications, *CCD and CMOS sensor technology: Technical White Paper* (2010)
1029	Jeff Meyer, *What is Depth of Field? How aperture, focal length and focus control sharpness*, Digital Camera World (July 17, 2013), https://prism.org.gg/wp-content/uploads/2016/07/Depth-of-Field.pdf
1030	Brian Klug, *Understanding Camera Optics & Smartphone Camera Trends, A Presentation by Brian Klug*, Anandtech (Feb. 22, 2013), https://www.anandtech.com/Show/Index/6777?cPage=2&all=False&sort=0&page=2&slug=understanding-camera-optics-smartphone-camera-trends
1031	*Portrait mode now available on iPhone 7 Plus with iOS 10.1*, Apple (Oct. 24, 2016)
1032	*The Timeline of Evolution of the Camera from the 1600s to 21st Century*, Capture.com (May 5, 2023), https://www.capture.com/blogs/insights/evolution-of-the-camera
1033	John Biggs, *A Nokia Camera Phone That's More Like a Camera With a Phone Attached*, The New York Times (Nov. 29, 2007), ttps://www.nytimes.com/2007/11/29/technology/personaltech/29phone.html

Exhibit	Reference
1034	U.S. Patent 7,412,081
1035	U.S. Patent 8,457,367
1036	U.S. Patent 8,260,008
1037	Zac Hall, *Over a decade of selfies, starting with iPhone 4*, 9to5Mac (July 1, 2021), https://9to5mac.com/2021/07/01/iphone-4-selfie-celebration/
1038	Duane C. Brown, *Decentering Distortion of Lenses*, Photogrammertic Engineering (1966)
1039	Reserved
1040	Reserved
1041	Minute Entry, *FaceTec, Inc. v. Jumio Corp.*, C.A. No. 24-cv-03623, ECF No. 38 (N.D. Cal. Sept. 18, 2024)
1042	Reserved
1043	Reserved
1044	Reserved
1045	Reserved
1046	Complaint, *FaceTec, Inc. v. Jumio Corp.*, C.A. No. 24-cv-03623, ECF No. 1 (N.D. Cal. June 14, 2024)
1047	Reserved
1048	Reserved

LISTING OF CHALLENGED CLAIMS

Reference	Claim Limitations
Claim 1	
1[pre]	A computing device for verifying three-dimensionality of a user via a user's camera equipped computing device, the computing device comprising:
1[a]	a processor configured to execute machine executable code;
1[b]	a screen configured to provide a user interface to the user;
1[c]	a camera configured to capture images;
1[d]	one or more memories configured to store machine readable instructions that are stored on the memory of the authentication server which when executed by the processor, cause the computing device to:
1[d1]	capturing at least one first image of the user taken with the camera of the computing device at a first location which is a first distance from the user;
1[d2]	processing the at least one first image or a portion thereof to create first data;
1[d3]	capturing at least one second image of the user taken with the camera of the computing device is at a second distance from the user, the second distance being different than the first distance, the capturing at least one second image of the user occurring after movement of the camera or the user to establish the camera at the second distance from the user;
1[d4]	processing the at least one second image or a portion thereof to create second data;
1[d5]	comparing the first data to the second data to determine whether expected differences exist between the first data and the second data which indicates three-dimensionality of the user;
1[d6]	verifying the images of the user exhibit three dimensional traits when the expected differences exist between the first data and the second data as a result of capturing the at least one first image and the at least one second image at different distances from the user.
Claim 2	
2[a]	The system according to claim 1, further comprising: interpolating the first data and the second data to obtain estimated intermediate data;

2[b]	capturing at least one third image of the user taken with the camera of the computing device at a third distance from the user, the third distance being between the first distance and the second distances;
2[c]	processing the at least one third image or a portion thereof to obtain third data; and
2[d]	comparing the estimated intermediate data with the third data to determine whether the third data matches the estimated intermediate data.
Claim 3	
3	The system according to claim 1, further comprising verifying the presence of one or more features on a side of a user's head in the at least one first image, and verifying the absence or reduced visibility of the one or more features on the side of the user's head in the at least one second image due to image capture at different distances from the user's head, wherein the first distance is larger than the second distance.
Claim 4	
4	The system according to claim 1, wherein the machine readable instructions is configured to display one or more prompts on the screen of the computing device to guide the user to capture the at least one first image at the first distance and the at least on second image at the second distance.[1]
Claim 5	
5[a]	The system according to claim 1, further comprising comparing the first data, second data, or both to enrollment data derived from an enrollment image, the enrollment image captured and stored prior to an authenticating; and
5[b]	only authenticating the user when the first data, the second data, or both match the enrollment data within a predetermined threshold.

[1] Petitioner treats this claim as intending to recite: "machine readable instructions [**are**] configured to display one or more prompts" and "to guide the user to capture the at least one first image at the first distance and the at least on**[e]** second image."

Claim 6	
6	The system according to claim 1, wherein the computing device is a hand-held device, and the user holds the device at the first and second distance to capture the at least one first image and the at least one second image.
Claim 7	
7	The system according to claim 1, wherein the first data and the second data comprise biometric data.
Claim 8	
8	The system according to claim 1, wherein the first data and the second data comprise a mapping of facial features.
Claim 9	
9	The method according to claim 1, wherein the first image and the second image is of the user's face and the user's head and facial features are held steady and without movement during capture of the first image and the second image.
Claim 10	
10[pre]	A method for evaluating three-dimensionality of a user, the method comprising:
10[a]	capturing at least one first image of the user taken with a camera at a first location which is a first distance from the user;
10[b]	processing the at least one first image or a portion thereof to create first data;
10[c]	moving the camera from the first location to a second location, the second location being a second distance from the user, or the user moving to change the distance between the user and the camera from the first distance to the second distance;
10[d]	capturing at least one second image of the user taken with the camera when the camera is the second distance from the user, the second distance being different than the first distance;
10[e]	processing the at least one second image or a portion thereof to create second data;
10[f]	comparing the first data to the second data to determine whether expected differences exist between the first data and the second data which indicate three-dimensionality of the user;
10[g]	verifying the images of the user exhibit three-dimensional traits when the first data and the second data have expected differences resulting from the at least one first image being captured with the

	camera at a different distance from the user than when the at least one second image is captured.
Claim 11	
11[a]	The method according to claim 10, further comprising: interpolating the first data and the second data to obtain estimated intermediate data;
11[b]	capturing at least one third image of the user taken with the camera at a third distance from the user, the third distance being between the first distance and the second distances;
11[c]	processing the at least one third image or a portion thereof to obtain third data; and
11[d]	comparing the estimated intermediate data with the third data to determine whether the third data matches the estimated intermediate data.
Claim 12	
12	The method according to claim 10, further comprising verifying the presence of ears of the user in the at least one first image, and verifying the absence or reduced visibility of the ears in the at least one second image, wherein the first distance is larger than the second distance.
Claim 13	
13	The method according to claim 10, further comprising one or more prompts on a screen to guide the user to capture the at least one first image at the first distance and the at least on second image at the second distance.
Claim 14	
14	The method according to claim 13, wherein the one or more prompts are an oval shape guide on the screen within which an image of a face of the user is aligned to capture the at least one first image and the at least one second image.
Claim 15	
15	The method according to claim 10, wherein the camera is part of a computing device is a hand-held device, and the user holds the computing device at the first distance when capturing at least one first image and at the second distances when capturing the at least one second image.
Claim 16	
16	The method according to claim 10, wherein the first data and the second data comprise biometric data.

Claim 17	
17	The method according to claim 10, wherein the first data and the second data comprise a map of facial features.
Claim 18	
18	The method according to claim 10, further comprising illuminate a screen of a computing device while capturing the at least one first image and/or the at least one second image, and processing the at least one first image and/or the at least one second image to detect a reflection of the illumination from a face of the user.
Claim 19	
19	The method according to claim 10, wherein a face of the user is held steady when capturing the at least one first image and the at least one second image and the camera moves from the first location to the second location.
Claim 20	
20	The method according to claim 10, wherein the first data and the second data are maintained on a computing device.
Claim 21	
21	The method of claim 10 wherein the camera is part of is one of a smartphone, tablet, laptop, or desktop computer.
Claim 22	
22[pre]	A method, performed by a user using a user's computer device, for verifying three-dimensionality of the user, the method comprising:
22[a]	capturing a first image of the user's head with a camera at a first distance from the user, the camera associated with the user's computing device;
22[b]	changing a distance between the user and the camera to a second distance by the user moving the camera, or the user moving relative to the camera, or both;
22[c]	capturing a second image of the user's head with the camera when the camera is at the second distance from the user, the second distance being different than the first distance;
22[d]	comparing one or more aspects of the user's head from the first image to one or more aspects of the user's head from the second image to determine whether expected differences, between the first image and the second image, exist which indicates three-dimensionality of the user, such that the expected differences between the first image and the second image result from the first

	image being captured when the camera is at a different distance from the user than when the second image is captured; and
22[e]	responsive to the comparing determining that expected differences between the first image and the second image exist, providing notice to the user, a third party, or both that the three-dimensionality of the user is verified.
Claim 23	
23	The method of claim 22 wherein the one or more aspects of the user's head from the first image is first data resulting from processing the first image and the one or more aspects of the user's head from the second image is second data resulting from processing the second image.
Claim 24	
24	The method of claim 22 wherein the user's head is the user's face.

I. INTRODUCTION

Unique biometric features such as fingerprints, eyes, and facial features are routinely used to authenticate a user's identity. But as long as biometric authentication has existed in computer systems, spoofers have sought to trick these systems into thinking the real user is present.

One way to spoof a biometric-authentication system is to present a picture of the biometric feature instead of the real thing. This attack has been mitigated in different ways for different biometric features. For fingerprints, sensors may check whether the object placed on the scanner conducts electrical current like a real finger. And facial-authentication systems have long used well-known camera-optics principles to verify the presence of a three-dimensional face rather than a two-dimensional picture.

U.S. Patent No. 11,874,910 ("the '910 Patent," Ex-1001) presents another facial-authentication system that seeks to distinguish real faces from pictures of a face. The only purported novelty, however, is the specific camera-optics principle used. But that principle was already known, and applying it for facial authentication was not new or non-obvious. For this reason, Jumio Corporation ("Petitioner") requests *inter partes* review ("IPR") of claims 1-24 ("Challenged Claims") of the '910 Patent.

II. GROUNDS FOR STANDING

Petitioner certifies the '910 Patent is available for IPR, and Petitioner is not barred or estopped from requesting IPR on the grounds herein. Petitioner files this Petition within one year of service of PO's district court complaint (*see* Ex-1046).

III. STATEMENT OF PRECISE RELIEF REQUESTED FOR EACH CLAIM CHALLENGED

Petitioner respectfully requests review and cancellation under 35 U.S.C. §311 of the Challenged Claims in view of the following prior art and opinions of Dr. Chris Daft ("*Daft*," Ex-1003):

Prior Art[2]
Derakhshani, U.S. Patent No. 8,437,513, filed August 10, 2012; issued May 7, 2013; prior art under 35 U.S.C. §§102(a)(1), (a)(2) (Ex-1005)
Zhang, U.S. Patent Application Publication No. 2011/0299741, filed June 8, 2010; published December 8, 2011; prior art under 35 U.S.C. §§102(a)(1), (a)(2) (Ex-1006)
Tanii, U.S. Patent Application Publication No. 2002/0113884, filed February 15, 2002; published August 22, 2002; prior art under 35 U.S.C. §§102(a)(1), (a)(2) (Ex-1007)
Tahk, U.S. Patent Application Publication No. 2014/0028823, filed May 22, 2013; published January 30, 2014; prior art under 35 U.S.C. §102(a)(2) (Ex-1008)
Hoyos, U.S. Patent Application Publication No. 2010/0014720, filed October 2, 2007; published January 21, 2010; prior art under 35 U.S.C. §§102(a)(1), (a)(2) (Ex-1010)

Ground (35 U.S.C. §103)	**Claims**	**References**
1A	1-13, 15-24	Derakhshani, Tanii
1B	14	Derakhshani, Tanii, Tahk
2A	1-3, 5-12, 15-17, 19-24	Zhang, Tanii
2B	4, 13-14	Zhang, Tanii, Tahk
2C	18	Zhang, Tanii, Hoyos

Daft, ¶¶98, 133.

[2] Applying post-AIA §102.

IV. ’910 PATENT OVERVIEW

A. Specification

The ’910 Patent is titled, “Facial Recognition Authentication System Including Path Parameters,” was filed October 22, 2021, and claims priority to several earlier applications dating back to August 28, 2014. Ex-1001; *Daft*, ¶¶87, 90.[3] For purposes of this petition only, Petitioner assumes the ’910 patent is entitled to claim priority back to this date.

The ’910 Patent describes facial-authentication systems and methods to distinguish a three-dimensional face from a two-dimensional picture of a face by looking for well-known optical effects—such as a “fish-eye” effect—that naturally occur in images to different degrees due to the three-dimensionality of a face and the distance between the face and the camera. Ex-1001, 1:45-48, 28:53-29:20; *Daft*, ¶88. The patent proposes capturing at least *two* images of the face: one “close” and one “far.” Ex-1001, 29:52-55. If the face is three-dimensional, the “close” image should exhibit the known effects (fish-eye distortion), but the “far” image should have much less. *Id.*, 29:59-67. If both images lack such “expected” (fish-eye) effects, however, it likely indicates a spoofing attempt using a two-dimensional picture. *Id.*; *Daft*, ¶89.

---

[3] Dr. Daft provides an overview of biometric authentication and camera optics at ¶¶37-53 (citing Ex-1017–19, Ex-1038).

B. Prosecution History

The ’910 Patent only received an obvious-type double patenting rejection during prosecution. Ex-1002, 1175-1181. After the Applicant filed Terminal Disclaimers (along with a handful of claim amendments), (*id*., 1242-44, 1252-59), the patent issued (*id*., 1265-67). *Daft*, ¶¶90-91.

V. LEVEL OF ORDINARY SKILL

A person of ordinary skill in the art (“POSITA”) in August 2014 would have had a Bachelor’s degree in electrical engineering, computer engineering, computer science, physics, or a related field, and two years’ work experience related to biometrics, facial authentication, computer vision, and/or optics, such that they would have had significant academic and/or work experience in both software development and optics. *Daft*, ¶¶129-31. Formal education can substitute for work experience and relevant work experience could substitute for formal education. *Id.*

Dr. Daft was at least a POSITA in the field as of August 2014 and is qualified to offer opinions here. *Id.*, ¶¶1-53, 132.

VI. CLAIM CONSTRUCTION

Petitioner is unaware of any prior claim construction determination related to the ’910 Patent. 37 C.F.R. §42.100(b). Nor does Petitioner believe any formal claim constructions are necessary to resolve the issues presented in this petition. *Wellman,* *Inc. v. Eastman Chem. Co*., 642 F.3d 1355, 1361 (Fed. Cir. 2011) (cleaned up). *Daft*, ¶¶92-93, 96-97.

Although there are several indefiniteness issues, none preclude the Board from evaluating patentability here. *Intel Corp. v. Qualcomm Inc.*, 21 F.4th 801, 813 (Fed. Cir. 2021) ("[I]ndefiniteness...precludes a patentability determination only when the indefiniteness renders it logically impossible for the Board to reach…a decision"). For instance, although the claims require "expected differences" between two images, which may be indefinite, the prior art can nonetheless be evaluated to determine whether it teaches types of distortion like those ("perspective," "fish-eye," and/or "barrel") identified in the '910 Patent that purportedly produce the "expected differences." Ex-1001, 28:57-29:4. *Daft*, ¶94.

Moreover, claim 1 requires "[a] computing device" for "verifying three-dimensionality of a user via a user's *camera equipped computing device*" that has a processor, screen, camera, and memory. Yet the memory is "configured to store machine readable instructions *that are stored on* ***the memory of the authentication server***,"[4] which is then executed by "the processor, caus[ing] the computing device" to execute steps. By associating "the memory" with an authentication server rather than the "computing device," it is unclear whether claim 1 is directed to a ***server***

[4] All emphasis added unless otherwise noted.

computing device that must have a processor, screen, and camera; or a networked, client-computing device that merely receives instructions from an authentication server. However, the Board need not resolve this ambiguity if the prior art teaches either interpretation (which it does). *Daft*, ¶95.

Finally, claim 9 refers to the "***method*** according to claim 1," but claim 1 recites "[a] ***computing device***." Ex-1001, claims 1, 9. For purposes of this Petition, however, Petitioner assumes there is no conflict between these two claims (e.g., that claim 9 is directed to a "system"). *Daft*, ¶95.

VII. DETAILED EXPLANATION OF GROUNDS

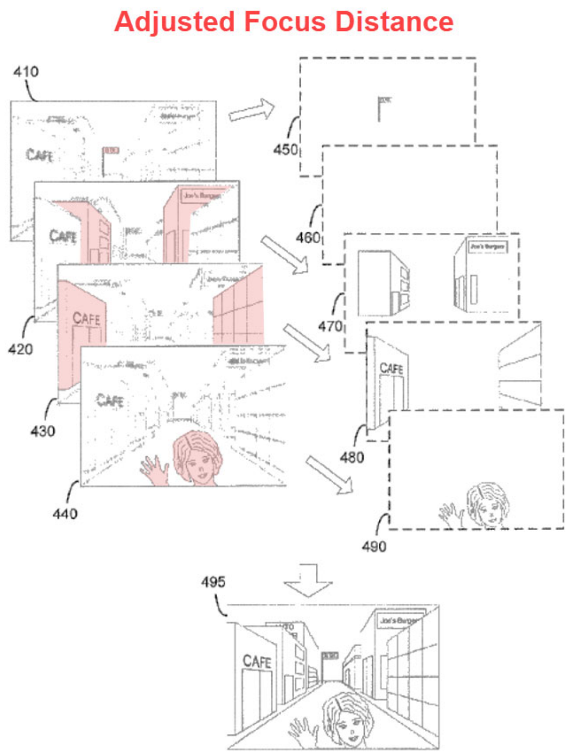
A. Ground 1A: Obviousness in view of Derakhshani and Tanii (Claims 1-13, 15-24)

1. Derakhshani (Ex-1005)

Derakhshani discloses a "biometric authentication" process using a device (e.g., computer or smartphone) that captures images with a camera. *Derakhshani*, 1:11-25, 5:22-27, 6:3-5, 9:10-22, 18:1-3; *Daft*, ¶¶99-100.

Although a user's identity is authenticated by evaluating the user's eyes, the authentication process also verifies whether the user's face is three-dimensional by calculating a "spatial metric" for the entire face. *Derakhshani,* Abstract (ocular authentication), 16:44-18:4 (three-dimensional verification). To verify three-dimensionality, Derakhshani exploits one or more known optics principles, such as: (1) adjusting the *focus* distance of the camera to determine whether different facial features exhibit different amounts of blur across multiple images, indicating the face has depth; and/or (2) evaluating whether "parallax" exists between two images captured from different camera positions, also indicating the face has depth. *Id.*, 16:44-17:11, 17:45-18:4; *Daft*, ¶¶101-02.

For the focus-distance approach, Derakhshani explains: "[a] landmark's representation in a particular image has a degree of focus that depends on how far the object corresponding to the landmark is from an in-focus point in the field of view of the sensor. Degree of focus is a measure of the extent…the image of the landmark is blurred by optical effects." *Derakhshani*, 16:54-59. Other prior art shows this principle in action, depicting several images (410, 420, 430, 440) captured at different focus distances, altering which objects are blurred and which are clear based on their distance from the camera:



See, e.g., Ex-1011, Fig. 4B (annotated); *Daft*, ¶¶103-04; Ex-1020, 4. For parallax, Derakhshani explains: "[a] plurality of images taken from different perspectives on the subject may result in landmarks within the images appearing to move by different amounts because of differences in their distance from the sensor." *Derakhshani*, 17:49-52. Other prior art likewise shows this principle in action, depicting how the position of a user's glasses relative to the eyes shift as the user moves toward the camera:

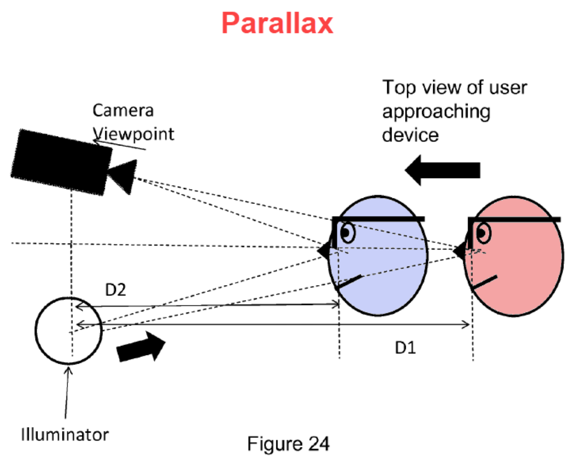
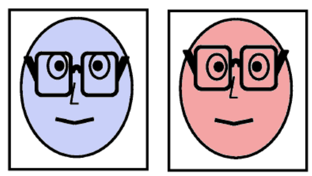
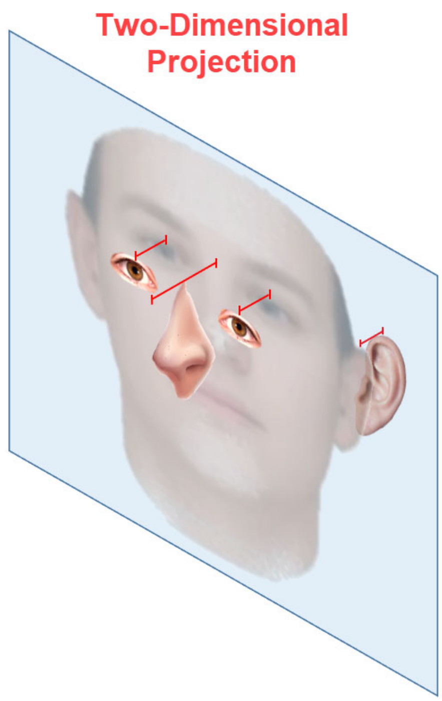


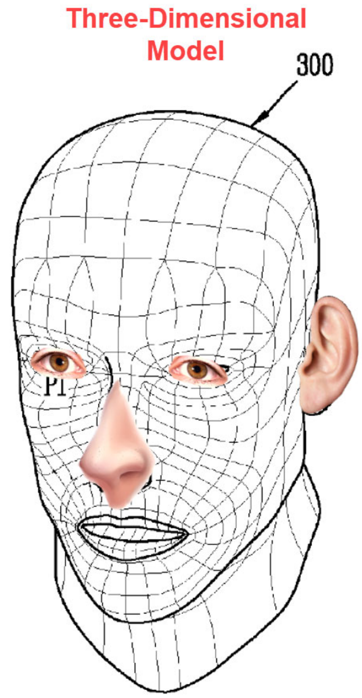
Figure 24



Ex-1012, Figs. 20, (bottom), 24 (top) (annotated); *Daft*, ¶¶105-06; Ex-1021, 2.

Derakhshani also discloses the distance information derived from these approaches can be mapped to either a projected two-dimensional plane (left example) or three-dimensional model (right example) to determine whether the distance information matches expectations. *Derakhshani*, 17:12-44.





Daft, ¶¶107-08.

2. Tanii (Ex-1007)

Tanii recognizes a well-known issue with certain camera systems: capturing a three-dimensional object (e.g., a face) at close range—especially when using a wide-angle lens common in mobile devices—can produce distortions in the resulting image based in part on the distance between the object and the camera. *Tanii*, [0005], [0007], [0009]; *Daft*, ¶¶109-10. Specifically, when an object (e.g., face) is close to the camera, it produces an "unnatural image…in which the perspective is exaggerated."

Fig.4A

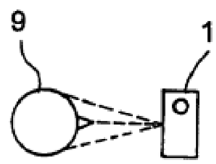
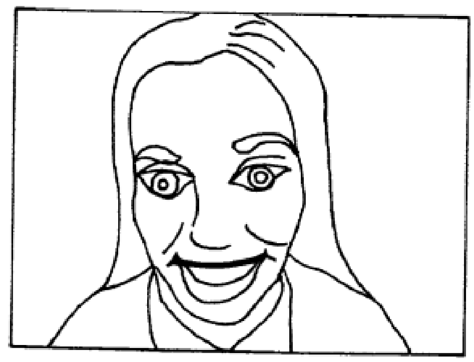


Fig.4B



Tanii, [0047], Figs. 4A-B. But when an object is further from the camera, "a natural image can be obtained."

Fig.3A

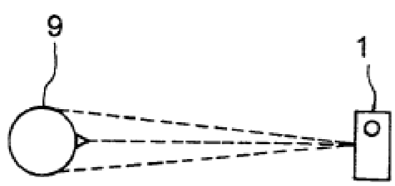
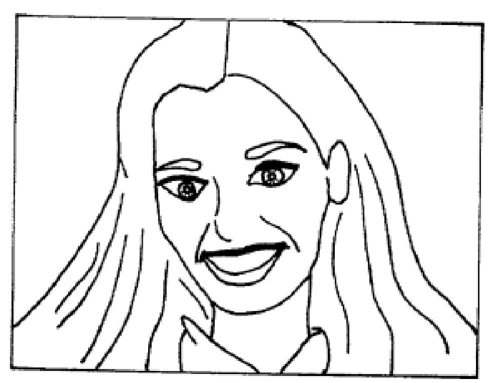


Fig.3B



Id., [0047], Figs. 3A-B; *Daft*, ¶111.

According to Tanii, this distortion arises because the face "has an essentially convex configuration that protrudes toward the [camera]," which causes the peripheral areas of the user to appear smaller relative to the center. *Tanii*, [0048]. Tanii then provides a procedure to correct this distortion by enlarging the image's peripheral areas relative to the center to produce an undistorted image. *Id*., [0056]; *Daft*, ¶¶111-13.

3. Motivation to Combine

A POSITA would have been motivated to combine Derakhshani and Tanii because both concern identifying and accounting for the three-dimensionality of a face when capturing an image. *Daft*, ¶877. They differ, however, in what principles are used to account for that three-dimensionality: Derakhshani uses changes in focus distance and/or parallax effect (§VII.A.1-(Derakhshani)), whereas Tanii looks for distortions in a face induced by the camera's lens (§VII.A.2-(Tanii); *Daft*, ¶877). A POSITA would have appreciated, however, that Tanii merely recognizes an obvious alternative to evaluating the depth of a face—consistent with Derakhshani's existing two approaches—and a POSITA would have been motivated to incorporate Tanii's evaluation into Derakhshani. *Id.*

Specifically, a POSITA would have recognized Derakhshani's focus-distance approach and Tanii's evaluation of distance-induced distortions are both attributable to optical effects caused by (among other factors) distances between the camera and the object(s) being captured. *Derakhshani*, 16:57-60 ("Degree of focus is a measure of the extent…the image of the landmark is blurred by optical effects…(e.g., due to *diffraction* and convolution with the aperture shape."); *Tanii*, [0048] (noting the "unnatural image" is caused by the angles of the face relative to the angle of the camera lens); *Daft*, ¶878.

Derakhshani and Tanii differ, however, in the type of diffraction that occurs. *Daft*, ¶879. Specifically, Derakhshani leverages the blurring of objects at distances *other than* the camera's focal plane—often called a "bokeh" effect—which makes those objects appear unfocused. *Derakhshani*, 16:54-59; *Daft*, ¶879. By adjusting the focus distance and evaluating when objects (or features of an object) in an image are clear versus when they are blurry, distance information can be derived. *Derakhshani*, 16:51-63; *Daft*, ¶879.

Tanii, on the other hand, leverages geometric distortion arising from the interaction between the shapes of objects being imaged and the camera's lens. *Daft*, ¶880. As Tanii explains, the convex shape of a three-dimensional face near the lens exacerbates this type of distortion. *Tanii*, [0048]; *Daft*, ¶880. Thus, particularly when a camera incorporates a wide-angle lens, images of a face close to the camera—in which the face occupies most of the image—will exhibit significant distortion. *Tanii*, [0047]; *Daft*, ¶880. But when the face is further from the camera and occupies less of the image, less distortion will be apparent. *Tanii*, [0047]; *Daft*, ¶880. A POSITA would have appreciated, however, that when evaluating multiple images taken at either different *focus* distances (Derakhshani) or *actual* distances (Tanii), these different types of diffraction provide information about an object's depth. *Daft*,

¶881. In other words, a POSITA would have understood Tanii merely teaches an obvious alternative to Derakhshani's existing two approaches to evaluate whether a face being captured is three-dimensional. *Id.*

A POSITA would have also had particular motivation to substitute Derakhshani's existing approaches with Tanii's distance-induced distortion analysis. *Id.*, ¶882. Specifically, a POSITA would have understood that implementing Derakhshani's focus-distance approach requires a camera with the ability to set a focal plane at a distance with extreme sensitivity to selectively blur nearby objects at different distances. *See Derakhshani*, 16:48-51; *Daft*, ¶882 (citing Ex-1029). A POSITA would have also understood the cameras often found in mobile devices do not have this ability; mobile devices typically incorporate wide-angle lenses to capture a wide field of view with a large depth of field because of their small size. *Tanii*, [0007]; *Daft*, ¶883 (citing Ex-1017, Ex-1030). In other words, a POSITA would have known there is not enough room in mobile devices to incorporate large image sensors and optics to fine-tune the focus distance to induce blurring of out-of-plane objects. *Daft*, ¶883; Ex-1031.

Accordingly, a POSITA would have been motivated to modify Derakhshani—especially when implementing biometric authentication in a mobile device as Derakhshani already envisions (*Derakhshani*, 5:23-26)—to capture at least two images at different *actual* distances and evaluate whether they exhibit different degrees of distance-induced (barrel/fish-eye) distortion, as Tanii suggests. *Daft*, ¶884. A POSITA would have found such a modification obvious because both techniques merely involve the application of well-known optics principles relating camera design and object distance from the camera, and Tanii already taught a mechanism to identify such distance-induced distortions and thus indicate when the face being captured has depth. *See, e.g.*, *Tanii*, [0056]; *Daft*, ¶884. This modification would have been nothing more than the use of a known technique (adjusting *actual* distance to evaluate depth based on distance-induced distortions) to improve similar devices (devices that adjust *focus* distance to evaluate depth based on lens-induced blurring), or a simple substitution of elements. *KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 417 (2007).

Finally, although Derakhshani discloses a separate process to verify the three-dimensionality of a face using parallax, a POSITA would have understood that evaluating for distance-induced distortion consistent with Tanii would be easier for users on a mobile device. *Daft*, ¶885. Specifically, mobile devices (such as phones or laptops) typically capture images of users at arms-length distances. *Id.* And facial features do not have *significant* differences in their depth (on the order of a few centimeters, as opposed to meters between the face and a background). *Id.* Thus, to evaluate for parallax in a face at hand-held distances with suitable accuracy, a user may need to move the device around their head (or vice versa)—out of their direct line of sight—to create substantial differences in perspective and thus more parallax to accurately verify the three-dimensionality of the face. *Id.* Evaluating instead for distance-induced distortions when the camera is held at different distances consistent with Tanii—and all within the user's direct line of sight—would therefore be easier to verify that a user's face is, in fact, three dimensional. *Id.* That said, a POSITA would have appreciated that evaluating for distance-induced distortion consistent with Tanii could be *supplemented* by also evaluating for any parallax effects that also arise from capturing images of the face from different distances. *Id.* In other words, a POSITA would have understood the techniques of Derakhshani and Tanii as complimentary and would have been motivated to use both or either, depending upon the application. *Id.*

4. Independent Claim 1

a. 1[pre]

If limiting, Derakhshani discloses or suggests 1[pre]. *Daft*, ¶¶886-87.

Derakhshani discloses "biometric authentication" systems and methods using a camera-equipped computing device. *Derakhshani*, 1:11-25, 5:22-27, 6:3-5, 9:10-22, 18:1-3. As part of the authentication process, Derakhshani discloses verifying the user's face is three-dimensional by capturing multiple images of a user's face to calculate a "spatial metric" representing the face's three-dimensionality. *Id.,* 1:11-25, 3:14-15, 16:44-18:4; *Daft*, ¶887.

b. 1[a]

Derakhshani discloses or suggests 1[a]. *Daft*, ¶¶888-89.

Derakhshani discloses the invention can be implemented in computing devices such as a "smart phone, a tablet computer, a television, a laptop computer, or a personal computer" (*Derakhshani*, 5:22-27), which incorporate a processor configured to execute machine-readable code (*see, e.g.*, *id*., Fig. 9, 2:4-12, 2:31-38, 7:15-20, 22:12-44, 23:26-37, 24:49-25:8).

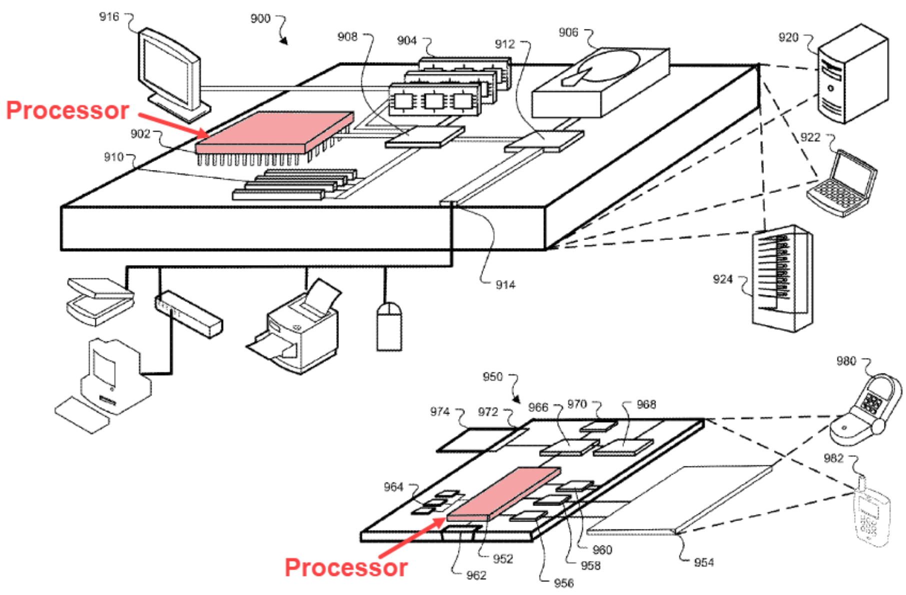


Fig. 9

Id., Fig. 9 (annotated). Derakhshani also discloses the computing device may be a server that also comprises a processor. *Id.*, 7:38-50, 8:48-9:4, 9:27-31, 10:16-19 ("the server system 514 is a data processing apparatus that includes one or more processors."), 23:14-44; *Daft*, ¶889.

c. 1[b]

Derakhshani discloses or suggests 1[b]. *Daft*, ¶¶890-91.

Derakhshani's computing device incorporates a display to provide an interface to the user. *See, e.g.*, *Derakhshani.* 6:8-11, 9:22-24, 14:35-37, 22:33-38, 23:48-52.

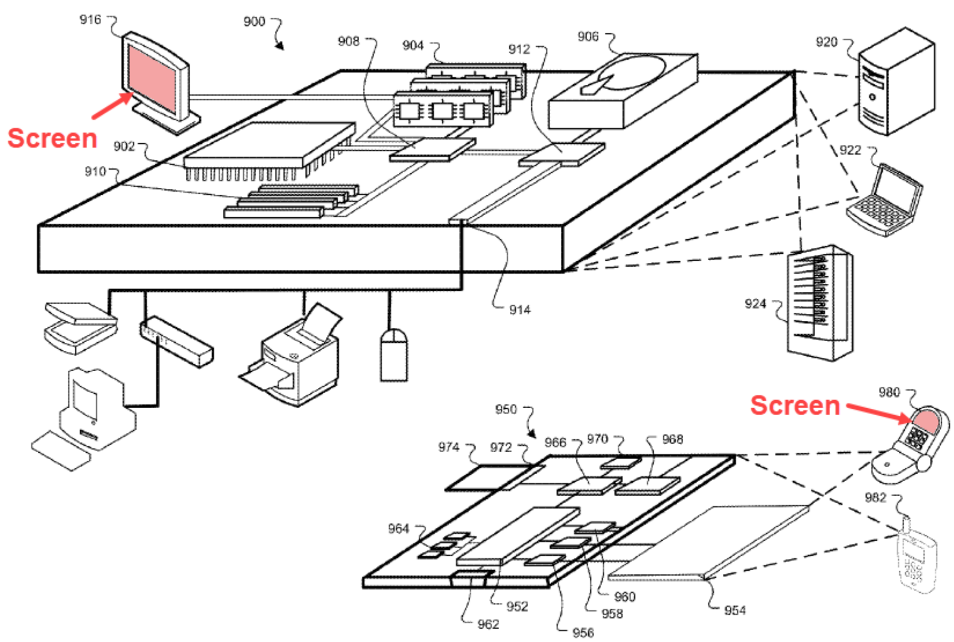


Fig. 9

Id., Fig. 9 (annotated). Derakhshani also discloses the computing device may be a server that also comprises a screen. *Id.*, 23:14-44, Fig. 9; *Daft*, ¶891.

d. 1[c]

Derakhshani discloses or suggests 1[c]. *Daft*, ¶¶892-93.

Derakhshani's computing device incorporates a camera configured to capture images. *See, e.g.*, *Derakhshani*, 5:23-27, 6:3-10; *Daft*, ¶893.

e. 1[d]

Derakhshani discloses or suggests 1[d] under both potential interpretations. §VI; *Daft*, ¶¶894-99.

Derakhshani's computing devices (e.g., a personal computer or phone) have memory that stores machine-readable instructions that are executed by the processor. *Derakhshani*, 2:4-12, 2:31-38, 22:26-44, 24:49-25:8.

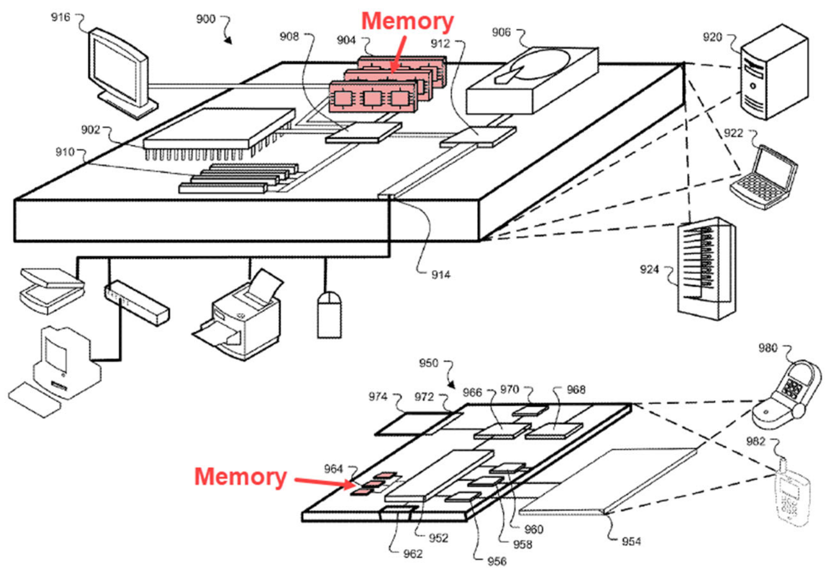
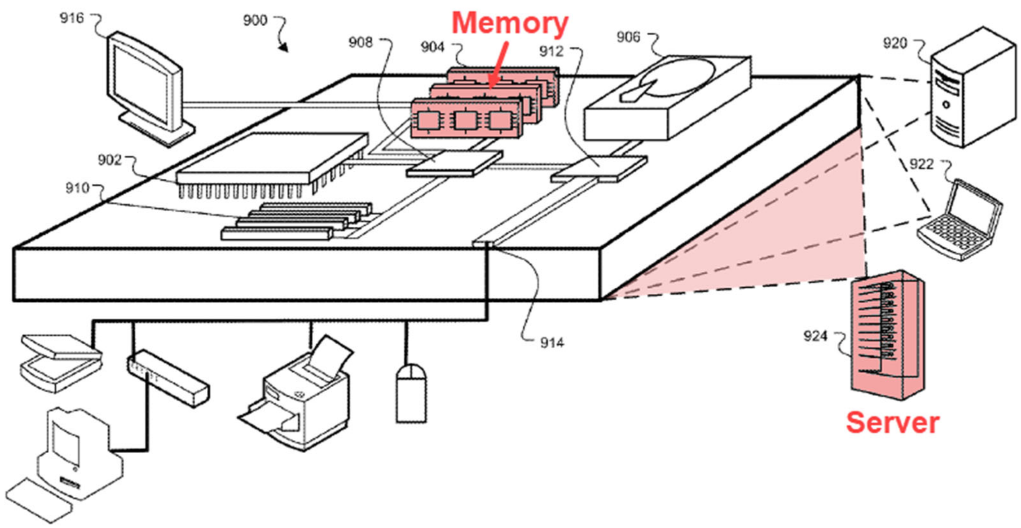


Fig. 9

Id., Fig. 9 (annotated); *Daft*, ¶896.

Derakhshani also discloses different embodiments of networked authentication systems, including: (1) a user-facing computing device that interacts with a "secure transaction service 523 hosted, e.g., by [a remote] server system" with "an authentication module 525 that coordinates authentication of users from the secured server's side of the interaction" (*Derakhshani*, 8:29-39); and (2) a user-facing computing device that hosts a local application that interacts with a remote authentication server (*id.*, 9:10-34). And Derakhshani discloses more generally that "authentication functions may be distributed between the client and the server side processes in a manner suited [to] a particular application." *Id.*, 9:27-58, 10:1-24; *Daft*, ¶896. Accordingly, a POSITA would have understood Derakhshani to teach different system configurations, including one in which a remote authentication server drives all aspects of the authentication process for a user-facing device such that the server directs the user-facing device to carry out certain aspects of Derakhshani's procedure (e.g., user image capture) using local memory that stores machine readable instructions originating from the memory of the authentication server. *Id.*, ¶897.

Derakhshani alternatively discloses the computing device itself can be a server containing the components depicted in Figure 9 (such as the display 916 and processor 902). *Derakhshani*, 22:12-18, 23:14-25.



Id., Fig. 9 (annotated); *Daft*, ¶898. Although Derakhshani does not show a camera as part of the server system, a POSITA would have understood Derakhshani envisions the server itself needing biometric protection (because servers are widely known to store sensitive information) such that it would have all the components necessary (including a camera, 1[c], per VII.A.4.d) for biometric authentication to access the server. *Daft*, ¶¶898-99.

f. 1[d1]

Derakhshani discloses or suggests 1[d1]. *Daft*, ¶¶900-02.

Derakhshani discloses, as part of the verification process, "two or more images of a subject" are captured using the camera of the computing device. *Derakhshani*, Fig. 8A, 1:44-46, 16:44-17:11, 17:45-18:4. A POSITA would have understood Derakhshani captures an image at a first location at a first distance because some distance must exist between the user and camera to capture a picture of the user's face. *Derakhshani*, 16:44-17:11; *Daft*, ¶¶901-02.

g. 1[d2]

Derakhshani discloses or suggests 1[d2]. *Daft*, ¶¶903-05.

Derakhshani discloses, as part of the verification process, "a landmark (e.g., an iris, an eye corner, a nose, an ear, or a background object) may be identified and located in the plurality of images." *Derakhshani*, 16:44-54 (focus distance approach), 17:32-36, 17:45-64 (parallax approach). A POSITA would have understood Derakhshani's identification of facial landmarks constitutes creating data (specifically, biometric data) because the identification involves using a computer (which operates on data) to characterize the unique physical characteristics of an individual, which would include the positions of "landmarks" such as a user's eyes, nose, ears, and other features on the user's head. *Daft*, ¶¶904-05.

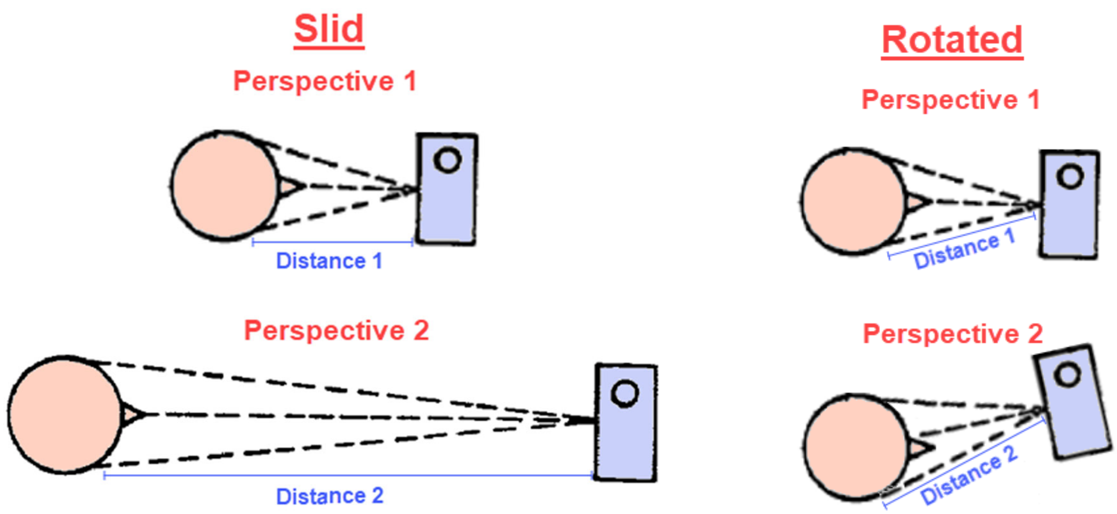
h. 1[d3]

Derakhshani, alone or combined with Tanii, teaches 1[d3]. *Daft*, ¶¶906-15.

Derakhshani discloses capturing "two or more images of a subject" using the camera. *Derakhshani*, 1:44-46, 16:44-17:11, 17:45-18:4; §VII.A.4.f-(1[d1]); *Daft*, ¶¶906-08.

For Derakhshani's focus-distance approach, a POSITA would have understood the distance between the user and the camera would need to change if the camera has a fixed focus distance (like in many mobile devices). *Id.,* ¶909. In other words, a POSITA would have understood and found it obvious that when using Derakhshani's focus-distance approach with a fixed-focus-distance camera, Derakhshani captures multiple images from multiple distances. §VII.A.3-(motivation); *Daft*, ¶910.

A POSITA would have also understood, however, that Derakhshani's parallax approach captures multiple images from multiple distances, because Derakhshani discloses that "[a] plurality of images [are] taken from different perspectives on the subject," such as: (1) when "a single camera [is] rotated *or slid* slightly"; (2) "a user is prompted to move" between image captures; or (3) the sensor moves naturally, such as "where the sensor is a camera in a hand-held user device (e.g., a smartphone or tablet) [that] may naturally move relative to the users face due to involuntary haptic motion." *Derakhshani*, 17:45-18:4. Thus, Derakhshani envisions capturing a second image after sliding or rotating the camera, which a POSITA would have understood changes the distance from landmarks of the face, if not the entire face, as demonstrated below (showing an exemplary top-down view of a face and camera):



Daft, ¶¶911-13. A POSITA would have therefore understood Derakhshani to disclose or suggest taking a second image of a user from a second distance that is different than the first to evaluate for parallax. *Id.,* ¶914.

But even if Derakhshani does not expressly disclose taking two images at different distances, doing so would have been obvious in view of other prior art. A POSITA would have understood, for instance, that distance-induced distortions indicate three-dimensionality of the object being captured, as Tanii teaches. *Tanii*, [0048]; *Daft*, ¶915; §VII.A.3-(motivation). Accordingly, a POSITA would have been motivated to modify Derakhshani in view of Tanii to expressly capture a second image at a second distance, and look for different degrees of distance-induced distortions to verify the face is three-dimensional. *Daft*, ¶915; §VII.A.3-(motivation).

Because Derakhshani, alone or combined with Tanii, teaches capturing a plurality of images at different distances between the user and the camera, a POSITA would have understood that the camera or user must move relative to the other between image captures. *Daft*, ¶915.

i. 1[d4]

Derakhshani discloses or suggests 1[d4]. *Daft*, ¶¶916-17.

Derakhshani discloses processing the captured images (including a second image) to identify biometric data "landmarks" in the face ("second [biometric] data") as part of the three-dimensional verification process. *Derakhshani*, 17:45-52; §VII.A.4.g-(1[d2]); *Daft*, ¶917.

j. 1[d5]

Derakhshani, alone or combined with Tanii, teaches 1[d5]. *Daft*, ¶¶918-23.

Derakhshani discloses that, for either the focus-distance or parallax approach, biometric features are compared between each of the images to match them. *Derakhshani*, 16:66-17:2 ("comparing the degree of focus for a landmark in images with different focus distances."); 17:45-64 (evaluating relative displacement of identified landmarks between images); *Daft*, ¶919. Moreover, a POSITA would have appreciated that, when modifying Derakhshani to evaluate distance-induced distortions consistent with Tanii, biometric data would be compared between images to determine whether they exhibit distance-induced distortion. *Id.*

Accordingly, a POSITA would have understood that each of the approaches to evaluate an object's depth taught by Derakhshani (focus distance or parallax), alone or combined with Tanii (distance-induced distortion), would compare biometric data points between multiple images to match the biometric data between each of the images (e.g., matching the ears, eyes, and nose in one image to those same features in another) to evaluate differences between them once matched. *Id.*

Derakhshani also discloses comparing the first biometric data to the second biometric data to determine whether expected differences between the two exist. *Derakhshani*, 16:66-17:2 (focus distance), 17:55-59 (parallax). Specifically, Derakhshani's focus-distance approach looks for expected differences in the blurriness/clarity of facial landmarks by changing *actual* distance (for fixed-focus cameras), and the parallax approach looks for different relative displacements of landmarks by also changing *actual* distance. *See id.*, 16:66-17:2, 17:55-59; §VII.A.4.h-(1[d3]); *Daft*, ¶¶920-22.

Moreover, when modifying Derakhshani in view of Tanii, a POSITA would have understood the biometric data from the first and second images would be compared (as Derakhshani already discloses for the focus-distance and parallax approaches) to determine whether the images exhibit expected differences in the amount of distance-induced distortion present in the image (as described by Tanii). §§VII.A.4.h-(1[d3]), VII.A.3-(motivation); *Daft*, ¶923. A POSITA would have understood that evaluating for different degrees of distance-induced distortion relies on a well-known optical effect of camera systems—particularly wide-angle camera systems common in mobile devices—to provide an alternative or supplemental, user-friendly verification of three-dimensionality of a face. §VII.A.3-(motivation); *Daft*, ¶923.

k. 1[d6]

Derakhshani, alone or combined with Tanii, teaches 1[d6]. *Daft*, ¶¶924-27.

Derakhshani discloses verifying a face is three-dimensional when expected differences exist between the biometric landmarks (e.g., biometric data) using either the focus-distance or parallax approach attributable to changes in distance between images. §§VII.A.4.h-(1[d3]), VII.A.4.j-(1[d5]); *Daft*, ¶925.

Moreover, a POSITA would have understood that, when modifying Derakhshani in view of Tanii, a face would be determined to be three-dimensional when the two images display expected differences in the biometric data due to different degrees of distance-induced distortion attributable to the change in distance between images. §§VII.A.3-(motivation), VII.A.4.h-(1[d3]), VII.A.4.j-(1[d5]); *Daft*, ¶926.

5. Claims 2 and 11

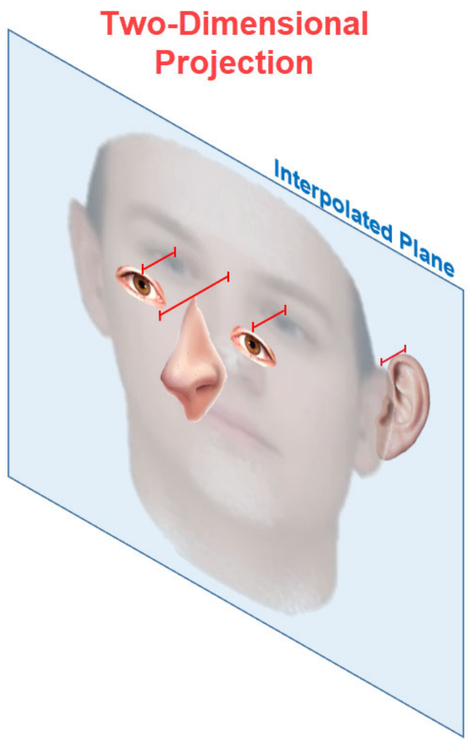
a. 2[a] and 11[a]

Derakhshani, alone or combined with Tanii, teaches 2[a] and 11[a].[5] *Daft*, ¶¶928-35, 980.

Derakhshani discloses verifying the face's three-dimensionality by fitting "the locations of multiple landmarks…to the closest two dimensional plane and the average distance of the landmarks from this fit plane can be determined as the spatial metric." *Derakhshani*, 17:12-26; *Daft*, ¶929.

A POSITA would have understood this two-dimensional plane would be generated by interpolating the position of facial landmarks onto a single intermediate plane (constituting "estimated intermediate biometric data," as claimed), and then the distance of those landmarks from the intermediate plane would be estimated to determine whether the face is three-dimensional. *Daft,* ¶930 (citing Ex-1026).

[5] Substantially identical claims are analyzed together within each ground throughout this petition.



A POSITA would have further understood that a similar interpolated model could be determined from Derakhshani's parallax process. *Id.*, ¶¶931-33 (citing Ex-1015, Ex-1018, Ex-1027).

Moreover, a POSITA would have been motivated to interpolate intermediate biometric data when modifying Derakhshani in view of Tanii. Specifically, a POSITA would have understood that a relationship exists between the degree of distance-induced distortion and distance: distortion increases as distance decreases, and vice versa. *Tanii*, [0048]; *Daft*, ¶934; Ex-1022. A POSITA would have therefore been motivated to implement Derakhshani's interpolated modeling approach in combination with Tanii, but instead of a single plane or three-dimensional model, interpolate biometric data simulating different degrees of distance-induced distortion that would be expected at intermediate distances. *Daft*, ¶935. A POSITA would have appreciated that building such a model would enable comparing additional, intermediate images to the model to further validate the three-dimensionality of the face based on a range of different distances. *Id.*

b. 2[b] and 11[b]

Derakhshani, alone or combined with Tanii, teaches 2[b] and 11[b]. *Daft*, ¶¶936-40, 981.

Derakhshani discloses verifying the three-dimensionality of the face by capturing "a plurality" of images. *Derakhshani*, 16:44-46 (focus-distance embodiment), 17:45-47 (parallax embodiment). A POSITA would have generally understood that capturing more images would produce more accurate verifications because there would be more samples to evaluate. *Daft*, ¶¶937-38. Thus, a POSITA would have been motivated to capture *more* than two images at different distances. *Id.* A POSITA also would have appreciated that, in any set with more than two images captured at different distances (§VII.A.4.h-(1[d3])); capturing images at different distances), there would be one image at a maximum distance, one at a minimum distance, and the rest at intermediate distances (*Daft*, ¶939).

Moreover, when modifying Derakhshani in view of Tanii to interpolate intermediate biometric data, a POSITA would have been motivated to capture a third image at an intermediate distance to compare it to the interpolated biometric data for verification. *Daft*, ¶940; §§VII.A.4.h-(1[d3]), VII.A.5.a-(2[a]).

c. 2[c] and 11[c]

Derakhshani, alone or combined with Tanii, teaches 2[c] and 11[c]. *Daft*, ¶¶941-43, 982.

Derakhshani discloses processing the images to identify feature landmarks in each (§§VII.A.4.g-(1[d2]), VII.A.4.i-(1[d4])), which a POSITA would have understood would be performed on all captured images (including a third image). *Daft,* ¶942.

Moreover, when modifying Derakhshani in view of Tanii, a POSITA would have considered it obvious to process the third, intermediate image to obtain biometric data for comparison to the interpolated, intermediate biometric data for verification. §VII.A.5.b-(2[b]); *Daft*, ¶943.

d. 2[d] and 11[d]

Derakhshani, alone or combined with Tanii, teaches 2[d] and 11[d]. *Daft*, ¶¶944-46, 983.

Derakhshani discloses comparing multiple images to the interpolated two-dimensional projection or three-dimensional model. *Derakhshani*, 17:12-26; §VII.A.5.a-(2[a]). A POSITA would have understood, therefore, that once a two-dimensional projection is generated consistent with Derakhshani, if a third image is captured, that too would be compared to the projection to estimate the distance of any facial landmarks in that image from the projection. §VII.A.5.a-(2[a]); *Daft*, ¶945.

Moreover, when modifying Derakhshani in view of Tanii, a POSITA would have found it obvious to acquire a third image and extract biometric data from the third image to compare it to the interpolated positions of the biometric data based on the first and second images to determine if there is a match between the two. §VII.A.5.b-(2[b]); *Daft*, ¶946.

6. Claims 3 and 12

Derakhshani, alone or combined with Tanii, teaches claims 3 and 12. *Daft*, ¶¶947-52, 984.

Derakhshani discloses "a landmark…an ear…may be identified and located," (*Derakhshani*, 16:51-54; *see also id.,* 17:14-19), which a POSITA would have appreciated is a "feature[] on a side of a user's head," as claimed. Ex-1001, 29:34-50 (describing a user's ear as a facial feature). For Derakhshani's focus-distance approach, a POSITA would have appreciated that the ear would have reduced visibility (i.e., blurry) when it does not lie in the focal plane, and would be present (i.e., clear) when it lies in the focal plane, and that distances in which the ear would be clear would be greater than those with reduced visibility (e.g., when the focal plane is aligned behind the ears). *Daft*, ¶¶948-49. For Derakhshani's parallax approach, a POSITA would have appreciated some perspectives would capture both ears, and other perspectives would only capture one (when one is obstructed by the head), which would indicate the user's face is three-dimensional. §VII.A.4.h-(1[d3]) (depicting a camera rotation that would obfuscate one ear); *Daft*, ¶950. Changing the camera perspective would not have the same effect if imaging a two-dimensional picture of a face with ears. *Daft*, ¶950.

Moreover, when modifying Derakhshani in view of Tanii, Tanii teaches the absence or presence of an ear is a natural result of changing the distance between the user and camera. Specifically, given sufficient distance between the face and camera, the ears are captured because there is enough distance for light rays from the ears to reach the camera's sensor. *See Tanii*, [0047]-[0048], Figs. 3A-B; *Daft*, ¶951.

Fig.3A

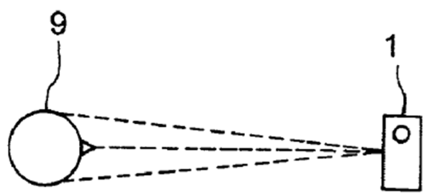
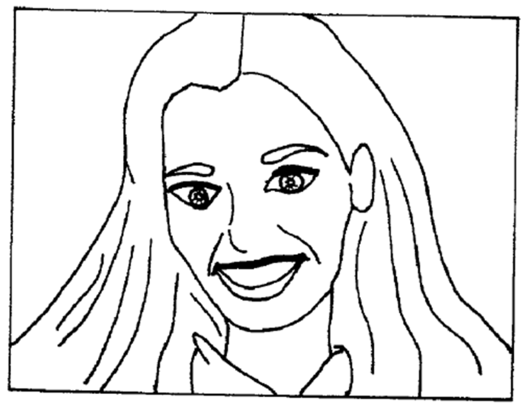


Fig.3B



But when a face is too close, the ears will not be captured because light rays from the ears cannot reach the camera's sensor. *See Tanii*, [0047]-[0048], Figs. 4A-B; *Daft*, ¶951.

Fig.4A

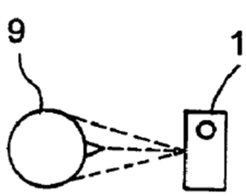
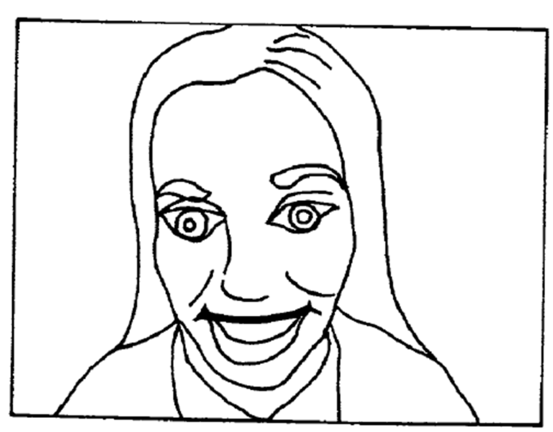


Fig.4B



This was a well-known consequence of distance-induced distortion when capturing images of a face. *Daft*, ¶951.

Accordingly, a POSITA would have appreciated (based on Tanii) that the presence of features on the side of the user's head (e.g., ears) in one image and absence in another image at a closer distance would be indicative of a three-dimensional face, and would have found it obvious to modify Derakhshani to verify the presence and absence of these features between images. *Id.*, ¶952.

7. Claims 4 and 13

Derakhshani, alone or combined with Tanii, teaches claims 4 and 13. *Daft*, ¶¶953-55, 985.

Derakhshani discloses displaying prompts to guide the user when capturing images of the user's face for authentication (*Derakhshani*, 5:23-32, 6:8-16, 9:22-26), including at more than one distance (*id.*, 17:64-66; §VII.A.4.h-(1[d3])). *Daft*, ¶954.

When modifying Derakhshani to look for distance-induced distortions by capturing images at different distances, consistent with Tanii (§VII.A.4.h-(1[d3])), a POSITA would have found it obvious to also provide prompts to a user to ensure the images are captured at the correct distances, as Derakhshani already discloses providing prompts to correctly orient the user relative to the camera. *Daft*, ¶955.

8. Claim 5

a. 5[a]

Derakhshani discloses or suggests 5[a]'s additional limitation. *Daft*, ¶¶956-57.

Derakhshani discloses capturing and analyzing multiple images of a user and comparing the user's features to a previously stored "reference record" to authenticate the user. *Derakhshani*, 4:19-24; 7:20-34; 8:60-64; 9:31-34. To create the enrollment reference record, the system captures one or more initial reference images of the user during an enrollment or registration process, extracts features from the reference images, and stores the extracted features as an enrollment reference record. *Id*., 7:19-34; 9:31-34; 13:62-14:9. During the authentication process, Derakhshani compares the extracted features from the captured images to the user's enrollment reference record to determine a match score. *Id*. 9:59-67, 13:62-14:9; 17:32-36; *Daft*, ¶957.

b. 5[b]

Derakhshani discloses or suggests 5[b]'s additional limitation. *Daft*, ¶¶958-59.

During the authentication process, Derakhshani calculates a match score based on the comparison of features extracted from the first and second image to the corresponding features in an enrollment reference record. *Derakhshani*, 13:62-66.

The match score is then compared to a predetermined threshold. *Id*. 14:25-28. When the match score is within that predetermined threshold, it is determined that the user who is depicted in the captured image is the same individual associated with the enrollment reference record, and the user is authenticated. *Id*. 14:25-45; *Daft*, ¶959.

9. Claims 6 and 15

Derakhshani, alone or combined with Tanii, teaches claims 6 and 15. *Daft*, ¶¶960-62, 986.

Derakhshani's biometric-authentication process can be implemented on hand-held computing devices, such as "a laptop computer, a handheld computer…, a tablet computing device, a personal digital assistant (PDA), a cellular telephone…, a camera, a smart phone," and more. *See, e.g.*, *Derakhshani*, 8:11-28, 18:1-4. Moreover, Derakhshani recognizes that, to verify three-dimensionality of the face, "a single camera may be rotated or slide slightly," or even a hand-held device "may naturally move relative to the users face due to involuntary haptic motion" that may sufficiently capture a parallax effect. *Id*., 17:59-18:4. Moreover, Tanii notes that distance-induced distortions often occur in mobile devices that have incorporated wide-angle lenses, and the amount of distortion depends upon the distance between the user and the camera. *Tanii*, [0007], [0047]-[0048], Figs. 3A-B, 4A-B; *Daft*, ¶961.

When implementing a three-dimensional verification process on a handheld mobile computing device consistent with Derakhshani, alone or combined with

Tanii (§§VII.A.4.h-(1[d3]), VII.A.7-(cl.4)), a POSITA would have understood that the user holds the computing device at a first distance for the first image and a second distance for the second image (e.g., by extending and retracting the user's arm). *Daft*, ¶962. That is a convenient and obvious way of changing the distance between a hand-held device and the user's face, and Derakhshani envisions evaluating depth based on movement of the user's arm while holding the device. §VII.A.1-(Derakhshani); *Daft*, ¶962.

10. Claims 7 and 16

Derakhshani, alone or combined with Tanii, teaches claims 7 and 16 for the reasons discussed in §§VII.A.4.g-(1[d2]), VII.A.4.i-(1[d4]). *Daft*, ¶963-64, 987.

11. Claims 8 and 17

Derakhshani discloses or suggests the additional limitation of claims 8 and 17. *Daft*, ¶¶965-66, 988.

Derakhshani discloses processing the captured images to identify and locate facial biometric "landmarks" (e.g., an iris, an eye corner, a nose, a mouth, an ear) in a three-dimensional verification process. *Derakhshani*, 16:44-54. A POSITA would have understood that the identification of facial landmarks would include their locations relative to one another, thus constituting a mapping of facial features. *Daft*, ¶966. Moreover, a POSITA would have understood that computer-based facial recognition conventionally involves mapping facial features to data structures. *Id*.

12. Claim 9

Derakhshani, alone or combined with Tanii, teaches claim 9. *Daft*, ¶¶967-70.

Derakhshani and Tanii both teach or suggest moving the camera to capture images at two different distances. §VII.A.4.h-(1[d3]). A POSITA would have understood that, when capturing images from different distances, either the user's face would remain stationary (e.g., steady) and the camera would move, or the user's face would move and the camera would remain stationary. *Daft*, ¶¶968-69. However, a POSITA would have appreciated that holding the user's face and head steady and moving the camera closer and further away would be more user friendly with a handheld device, for example, than forcing the user to move their head closer and further from the camera while holding the camera steady. *Id.*, ¶¶969-70.

13. Independent Claim 10

a. 10[pre]

If limiting, Derakhshani discloses or suggests 10[pre]. *Daft*, ¶971-72.

Derakhshani discloses a method for using a camera-equipped computing device for "biometric authentication." *See Derakhshani*, 1:11-25, 5:22-27, 6:3-5, 9:10-22, 18:1-3. As part of the authentication process, Derakhshani discloses verifying the user's face is three-dimensional—as opposed to a two-dimensional picture—by capturing multiple images of a user's face at different focus distances or from different perspectives to calculate a "spatial metric" that represents the three-dimensionality of the face. *Id.,* 1:11-25, 3:14-15, 16:44-18:4; §VII.A.4.a-(1[pre]); *Daft*, ¶972.

b. 10[a]

Derakhshani discloses or suggests 10[a] for the reasons discussed in §VII.A.4.f-(1[d1]). *Daft*, ¶973.

c. 10[b]

Derakhshani discloses or suggests 10[b] for the reasons discussed in §VII.A.4.g-(1[d2]). *Daft*, ¶974.

d. 10[c]

Derakhshani, alone or combined with Tanii, teaches 10[c] for the reasons discussed in §VII.A.4.h-(1[d3]). *Daft*, ¶975.

e. 10[d]

Derakhshani, alone or combined with Tanii, teaches 10[d] for the reasons discussed in §VII.A.4.h-(1[d3]). *Daft*, ¶976.

f. 10[e]

Derakhshani discloses or suggests 10[e] for the reasons discussed in §VII.A.4.i-(1[d4]). *Daft*, ¶977.

g. 10[f]

Derakhshani, alone or combined with Tanii, teaches 10[f] for the reasons discussed in §VII.A.4.j-(1[d5]). *Daft*, ¶978.

h. 10[g]

Derakhshani, alone or combined with Tanii, teaches 10[g] for the reasons discussed in §VII.A.4.k-(1[d6]). *Daft*, ¶979.

14. Claim 18

Derakhshani discloses or suggests claim 18's additional limitation. *Daft*, ¶¶989-90.

Derakhshani discloses calculating a "reflectance metric" that measures changes in surface glare on the eye due to light—such as the illumination of the screen—to further verify the "liveness" of the user. *Derakhshani*, 18:8-19 ("The reflectance metric may be a measure of changes in glare or specular reflection patches… [from] a dynamic light source (e.g.,… LCD screen….)"). A POSITA would have understood that the "LCD screen" is the screen of the computing device. *Daft*, ¶990. Derakhshani also discloses processing images to detect a reflection of the illumination from a face of the user. *Derakhshani*, 18:8-19; *Daft*, ¶990.

A POSITA would have therefore understood that illuminating the screen while capturing the first and second images, and measuring the reflectance from the illumination in the captured images, serves as another liveness measure to ensure the imaged face is from a real user, not a spoofer. *Daft*, ¶990.

15. Claim 19

Derakhshani, alone or combined with Tanii, teaches holding the user's face steady and moving the camera from a first to a second location. *See* §VII.A.12-(cl.9); *Daft*, ¶991.

16. Claim 20

Derakhshani discloses or suggests claim 20's additional limitation. *Daft*, ¶¶992-93.

Derakhshani discloses the biometric-authentication process can be performed locally on the device, on a server, or split between the two. *See, e.g.*, *Derakhshani*, 9:27-58, 10:1-24. In an implementation in which Derakhshani's three-dimensional verification—part of the biometric-authentication process—is performed locally on a device, a POSITA would have understood that the biometric data would be maintained on that device. *Daft*, ¶993.

17. Claim 21

Derakhshani, alone or combined with Tanii, teaches claim 21's additional limitation. *Daft*, ¶¶994-95.

Derakhshani discloses that images "may be captured with a sensor (e.g., a camera) that is integrated into a computing device such as, for example, a smart phone, a tablet computer, a television, a laptop computer, or a personal computer." *Derakhshani*, 5:22-27. Derakhshani further discloses that the computing device may be a "desktop computer [or] a laptop computer." *Id.*, 8:11-28, *see also, e.g.*, 9:18-22, 18:1-4; *Daft*, ¶995.

18. Independent Claim 22

a. 22[pre]

If limiting, Derakhshani discloses or suggests 22[pre] for the reasons discussed in §VII.A.4.a-(1[pre]). *Daft*, ¶996.

b. 22[a]

Derakhshani discloses or suggests 22[a] for the reasons discussed in §VII.A.4.f-(1[d1]). *Daft*, ¶997.

c. 22[b]

Derakhshani, alone or combined with Tanii, teaches 22[b] for the reasons discussed in §VII.A.4.h-(1[d3]). *Daft*, ¶998.

d. 22[c]

Derakhshani, alone or combined with Tanii, teaches 22[c] for the reasons discussed in §VII.A.4.h-(1[d3]). *Daft*, ¶999.

e. 22[d]

Derakhshani, alone or combined with Tanii, teaches 22[d] for the reasons discussed in §§VII.A.4.j-(1[d5]), VII.A.4.k1-(1[d6]). *Daft*, ¶1000.

f. 22[e]

Derakhshani discloses or suggests 22[e]. *Daft*, ¶¶1001-03.

Derakhshani discloses that, when rejecting an authentication attempt, the system provides notice to the user or a third party, and that the authentication is one of at least liveliness and/or three-dimensionality. *Derakhshani*, 8:67-9:4, 11:17-26. Separately, Derakhshani discloses providing notice that the user's identify has been authenticated. *Id.*, 14:45-58; *Daft*, ¶1002.

To the extent PO argues that Derakhshani does not expressly disclose notifying the user when their face has been verified as three-dimensional, providing such notice would have been obvious to a POSITA. *Daft*, ¶1003. For instance, if the biometric-authentication system comprised separate liveness evaluations performed in series, as Derakhshani discloses (such as behavioral, spatial, and reflectance) (*Derakhshani*, 19:46-48; *see also generally id*. 15:26-19:48), a POSITA would have been motivated to provide notice after each separate evaluation to indicate the next had begun, particularly since Derakhshani already envisions providing notifications during the authentication process. *Daft*, ¶1003. Implementing such notifications would have involved routine software coding that would have been within the skill of a POSITA. *Id*.

19. Claim 23

Derakhshani discloses or suggests claim 23's additional limitation for the reasons discussed in §§VII.A.4.g-(1[d2]), VII.A.4.i-(1[d4]). *Daft*, ¶1004.

20. Claim 24

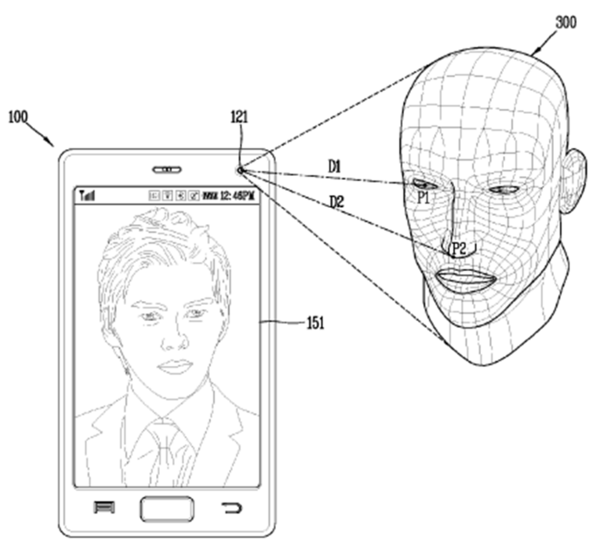
Derakhshani discloses or suggests claim 24's additional limitation for the reasons discussed in §§VII.A.4.g-(1[d2]), VII.A.12-(cl.9). *Daft*, ¶1005.

B. Ground 1B: Obviousness in view of Derakhshani, Tanii, and Tahk (Claim 14)

1. Tahk (Ex-1008)

Tahk discloses a facial-recognition procedure for a mobile terminal (e.g., a cellphone) that captures at least two images of a user's face at different distances and uses the "stereoscopic shape" of the user's face from those two images to distinguish between a live, three-dimensional face and two-dimensional pictures of a face. *Tahk*, Abstract, [0023], [0117], [0122], [0130]-[0131].

FIG. 5



Tahk, Fig. 5; *Daft*, ¶¶122-23. To capture the one or more images, Tahk not only presents a live image of the user, but also provides users written prompts and oval overlays to ensure the user's face is appropriately distanced from the camera.

FIG. 8A

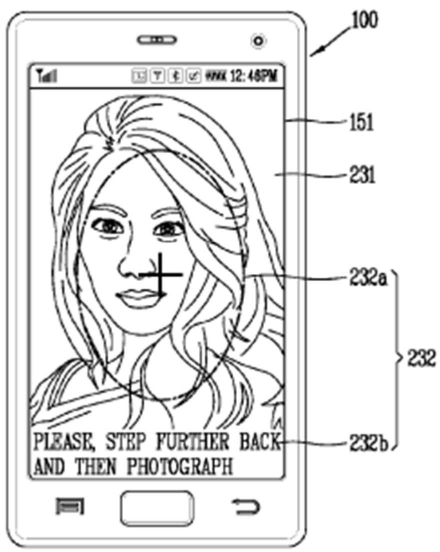
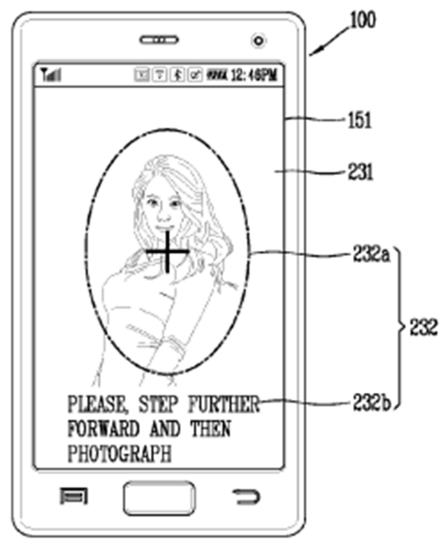


FIG. 8B



See, e.g., *Tahk*, Figs. 8A-B, [0118], [0129], [0135], [0139], [0143], [0144]; *Daft*, ¶124.

2. Motivation to Combine

A POSITA would have been motivated to combine Tahk with Derakhshani (with or without Tanii) because Derakhshani envisions prompting a user to capture a self-portrait. *See Derakhshani*, 5:23-32, 6:8-16, 9:22-26, 17:64-66. Tahk merely provides more-explicit, user-friendly ways of ensuring that a face presented for authentication is properly framed, which would be useful for the authentication procedure in Derakhshani and Tanii that requires multiple images of the face to be taken at different distances. *Daft*, ¶1006. A POSITA would have understood that providing a user with real-time feedback regarding how the face is framed for imminent image capture would allow the user to adjust position relative to the camera, and on-screen prompts such as written instructions and ovals to frame the face would provide users clear instructions to ensure an image is captured that is useable for authentication. *Id.*, ¶¶1007.

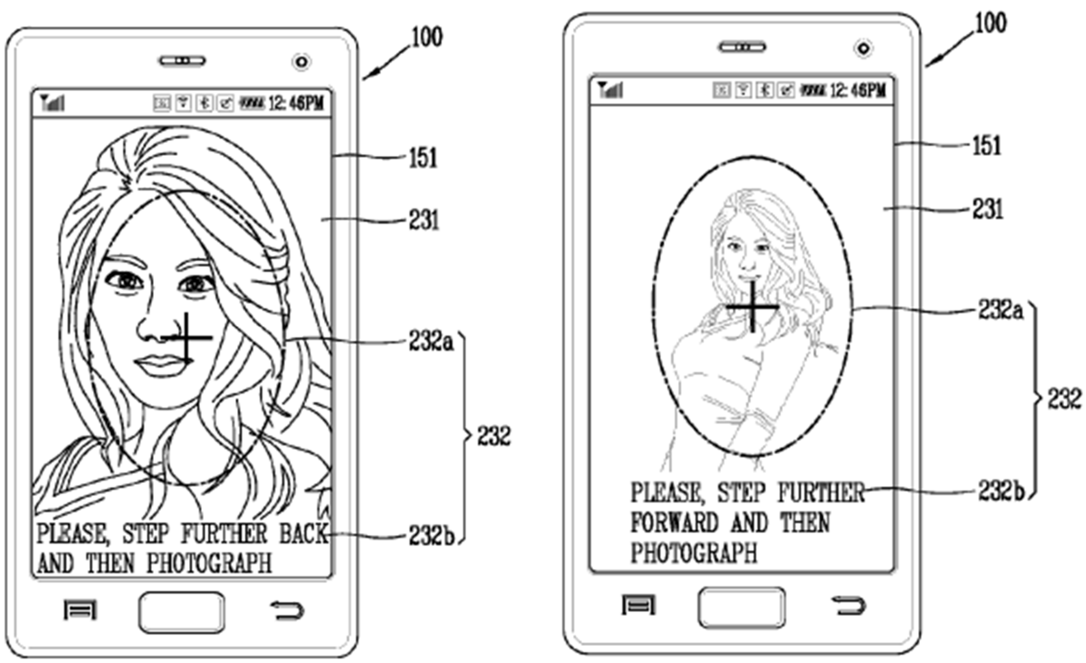
Moreover, providing real-time image feedback, written instructions, and oval shapes to frame a face during authentication were all well-known and conventional techniques at the time of the invention. *See, e.g.*, Ex-1009, 7:16-8:7, Figs. 6B-7C (live image previews and oval prompts); Ex-1010, 5:31-32 ("The computing device may present prompts that instruct the user to perform one or more liveness gestures"), 6:3-4 (same); Ex-1032. A POSITA would have therefore been motivated to implement these types of feedback to ensure that images are captured properly for facial authentication. *See Samsung Elecs. V. Lynk Labs, Inc.*, IPR2022-00100, Paper 30, 23-24, 34-35 (June 7, 2023) (conventional teachings supported by background references obvious); *Daft*, ¶1008 (citing Ex-1034, Ex-1035).

3. Claim 14

Derakhshani, combined with Tanii and/or Tahk, teaches claim 14's additional limitation. *Daft*, ¶¶1009-11.

Derakhshani, alone or combined with Tanii, teaches prompting a user to properly frame an image at different distances to capture images for biometric authentication. §VII.A.7-(cl.13); *Daft,* ¶1010.

Although neither Derakhshani nor Tanii expressly teaches using oval-shaped prompts to guide a user during the facial-authentication process, Tahk teaches using written and oval prompts to ensure images of the face are captured at the correct distances. *See, e.g.*, *Tahk*, Figs. 8A-B ("Please step further back" and "Please step further forward," and presenting an oval to frame the face at the correct distance):



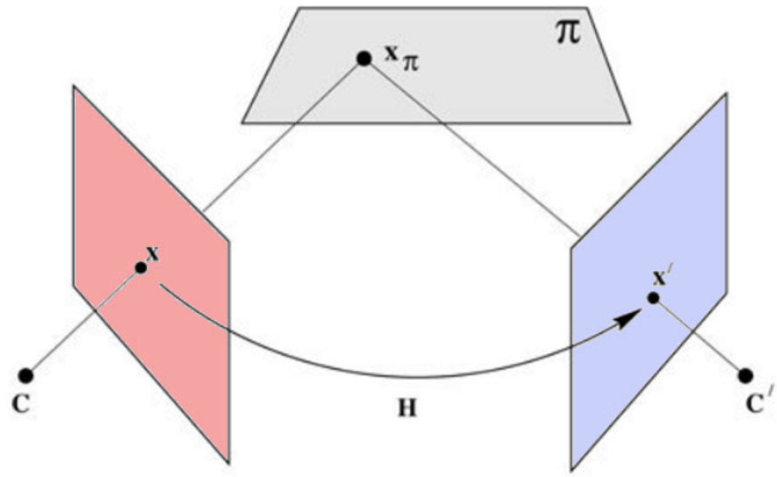
A POSITA would have been motivated to modify Derakhshani (alone or combined with Tanii), to provide such oval-shaped prompts because they are a natural shape to appropriately size and frame a face at different distances. §§VII.B.1-(Tahk), VII.B.2-(motivation); *Daft*, ¶¶1010-11.

C. Ground 2A: Obviousness of Zhang and Tanii (Claims 1-3, 5-12, 15-17, 19-24)

1. Zhang (Ex-1006)

Zhang discloses distinguishing "live faces" from two-dimensional pictures of a face during authentication. *See, e.g.*, *Zhang*, Title, Abstract, [0012], [0016]. To do so, Zhang's "image capture component" captures a series of images of a face. *Id.*, [0016]. The images are passed to a "live face detection module" that processes them to determine whether the face is live or not. *Zhang*, [0017]; *Daft*, ¶¶114-16.

One way Zhang distinguishes three-dimensional faces from two-dimensional pictures is through a "homography based technique" that utilizes a well-known relationship that "two views of a flat (planar) surface are related based on a homography matrix." *Zhang*, [0024]. For background, a paper Zhang references in the specification (*id.*, [0027] (also written by Zhang)) briefly explains the theory behind similar transformations. Ex-1013, 6-7. When two images of the same scene are captured from different perspectives—e.g., when multiple cameras capture the same object from different positions, or a single camera captures the same object from different positions—a mathematical relationship exists between different points in the two images. *Id.*; *Daft*, ¶¶117-18.



Id.; Ex-1023, 40-41; Ex-1025. A homography matrix defines the relationship between these two perspectives, and enables, *inter alia*, one image (e.g., red) to be mathematically transformed to match the perspective of another image (e.g., blue) of the same scene (e.g., grey). *Daft*, ¶119; Ex-1024.

To accomplish this transformation, Zhang discloses that facial features are extracted and matched between the first and second images to serve as reference points relating the two perspectives. *Zhang*, [0027]-[0028] ("This matching of the feature points across the first and second images refers to identifying the locations of the same feature points in each of the two images."); *Daft*, ¶119. These matched feature points are inputs to generate the "homography matrix." *Zhang*, [0029].

After the homography matrix is calculated, it is applied to every pixel in the first image to generate a "warped image" intended to match the perspective of the second image. *Id.*, [0030]. The "warped" (transformed) image is compared to the second image. *Id.*, [0031]-[0032]. If the differences between the "warped" (transformed) image and second image meet a threshold, the image is determined to be a live, three-dimensional face. *Id.*, [0034]. Otherwise, the image is determined to be a two-dimensional "imposter." *Id.*; *Daft*, ¶120.

Zhang takes advantage of an assumption in homography that the object captured from different perspectives exists on a two-dimensional plane. *Id.*, ¶121. Because a picture of a face exists on a two-dimensional plane (the paper it is printed on), performing a homography transformation on the picture of a face should produce a near-identical image no matter into which perspective it is transformed. *Id*. A live, three-dimensional face, however, does not exist on a single plane (e.g., the nose is closer to the camera than the ears) and therefore produces distortions (e.g., the ears would distort relative to the nose) when an image of a live face is transformed from one perspective to another. *Id*.

2. Motivation to Combine

A POSITA would have been motivated to combine Zhang and Tanii because both concern identifying and accounting for the three-dimensionality of a face when capturing an image. *Daft*, ¶1012. They differ, however, in what information is available to identify a face as three-dimensional: Zhang looks for distortions in a homography transform, whereas Tanii looks for distortions caused by the interaction between the user's face and the camera's lens. *Id.* A POSITA would have appreciated, however, that Tanii's teachings have broad application, including to Zhang. *Id.*

As Tanii recognizes, distance-induced distortions occur because of the interactions between the shape of the camera lens and shape of the face, and the extent of distortion depends upon the distance between the face and camera. §VII.A.2-(Tanii); *Tanii*, [0048]; *Daft*, ¶1013. Accordingly, a POSITA would have understood from Tanii that, by taking two images from two different distances, different degrees of distance-induced distortion of the face indicate whether a face is three-dimensional. *Daft*, ¶1013.

A POSITA would have therefore appreciated from Tanii that, even without modification, images captured by Zhang may exhibit distance-induced distortions, particularly when a wide-angle camera common in computers and mobile devices is used. *Id.*, ¶1014. However, a POSITA would have also appreciated that any distance-induced distortions would *enhance* Zhang's process because a homography transformation would not correct for distance-induced distortions. *Daft*, ¶1014. For instance, applying a homography transformation to Tanii's Figure 4B (as Zhang's "first image") to compare to Figure 3B (as Zhang's "second image"), the transformation would not account for distortion-related differences, such as the missing ears in Figure 4B and the differences in relative distance between facial features. *Id.*, ¶1015.

Fig.4B

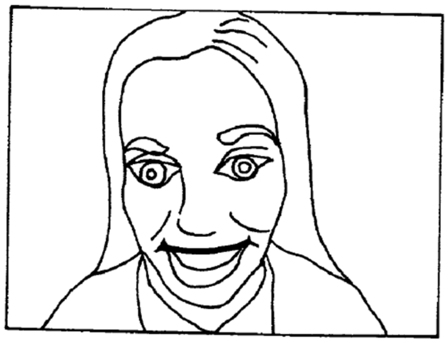
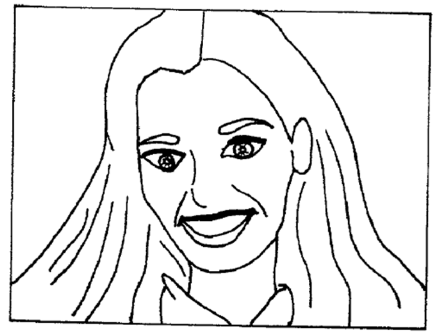
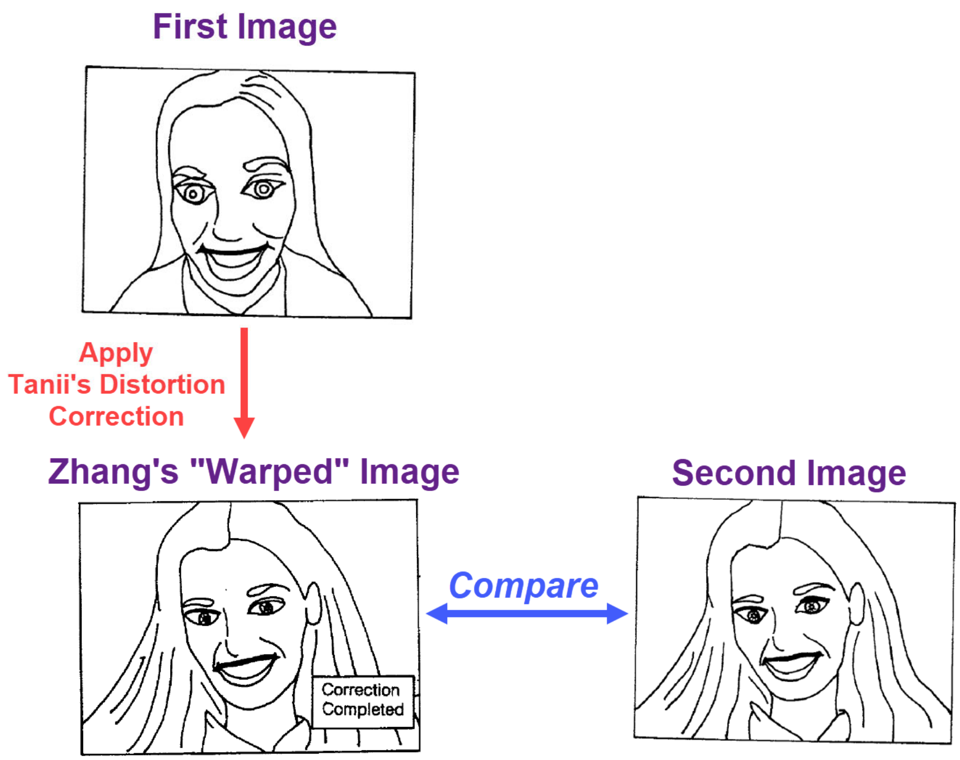


Fig.3B



Accordingly, a POSITA would have been motivated to use Zhang's existing process, but take advantage of the distance-induced distortions identified by Tanii by using a (wide-angle) camera that induces these distortions at different distances, and capturing two images of the face at different distances, because Zhang's procedure would identify the two images as different due to the differences in distance-induced distortion and verify three-dimensionality. *Id.*, ¶¶1015-16.

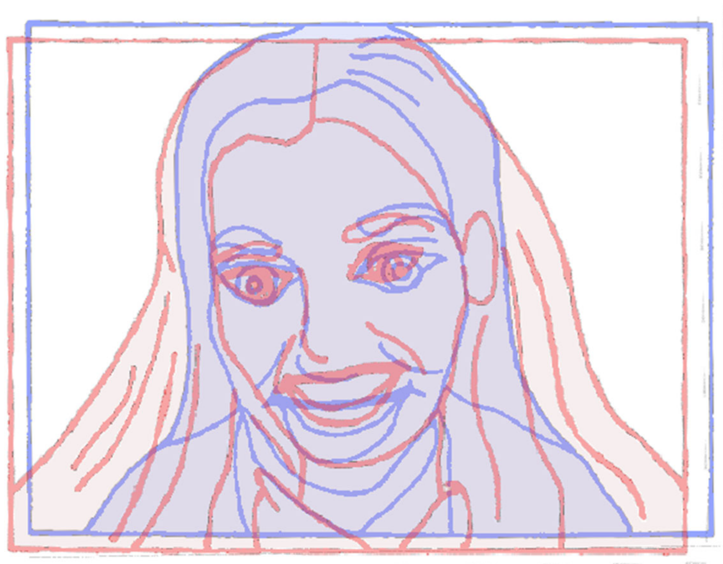
But a POSITA would have also been motivated to modify Zhang's process in view of Tanii in either of two additional ways. First, a POSITA would have been motivated to perform Zhang's three-dimensional verification process, but rather than "warp" one image using a homography transform to compare it to the second image and evaluate the two for differences, a POSITA would "warp" (i.e., correct) the "close" image using Tanii's distortion-correction procedure to compare it to the "far" image and evaluate the two for differences.



Tanii, Figs. 3B, 4B, 9; *Daft*, ¶¶1017-18.

A POSITA would have understood that if the "warped" (corrected) image and second image are sufficiently similar, it indicates a three-dimensional face because Tanii corrects for distortions attributable the three-dimensionality of the user's face. *Daft*, ¶1019. A POSITA would have appreciated this modification would reduce spoofing of Zhang's system (which only evaluates whether a face is non-planar, which can be spoofed), and provides a simpler mathematical operation that may require less computational resources. *Id.*, ¶¶1020-21 (citing Ex-1014).

Second, a POSITA would have alternatively appreciated the combination of Zhang and Tanii could be further simplified by eliminating the need to mathematically transform one image. *Daft*, ¶1020. Instead, when evaluating two images for different degrees of distance-induced distortion, the two images (one "close", one "far") could be compared directly to determine whether they exhibit expected differences in degree of distance-induced distortions. *Daft*, ¶¶1019-20.

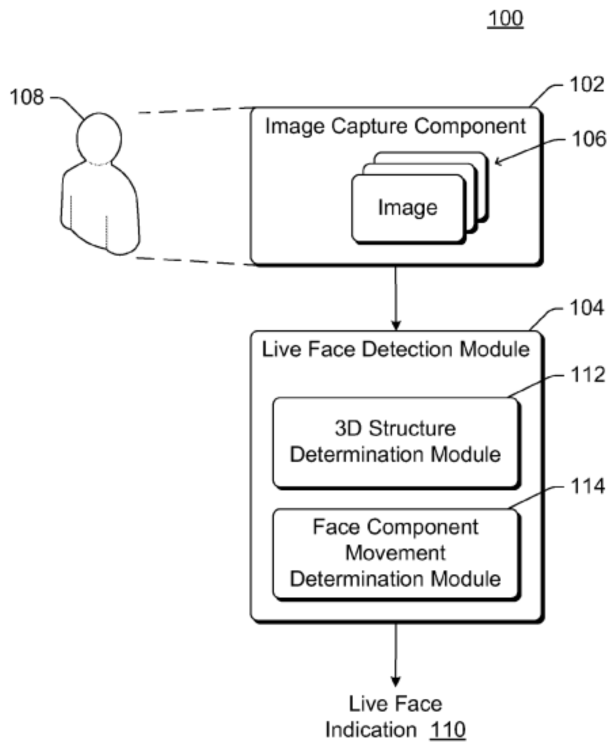


3. Independent Claim 1

a. 1[pre]

If limiting, Zhang discloses or suggests 1[pre]. *Daft*, ¶¶1023-25.

Zhang discloses a system "to determine whether a face in multiple images is a 3D structure or a flat surface" (*Zhang*, [0026], Figs. 1-3; *see also, e.g., id.*, Abstract, [0003], [0013]) to "authenticate a user for particular access" (*id.*, [0012]). To do so, Zhang captures and analyzes multiple images of a user's face using image capture component 102 implemented in a computing device (e.g., "a desktop computer, a laptop or notebook computer…[or] a cellular or other wireless phone”). *Zhang*, [0012]-[0013], [0016]; *Daft*, ¶1024. A POSITA would have understood that the “image capture component 102” is a camera as conventionally used in computing devices to capture images. *Zhang*, [0016] (explaining image capture component 102 uses, e.g., “CCDs” and “CMOS” sensors); *Daft*, ¶1025 (citing Ex-1009, Ex-1028).

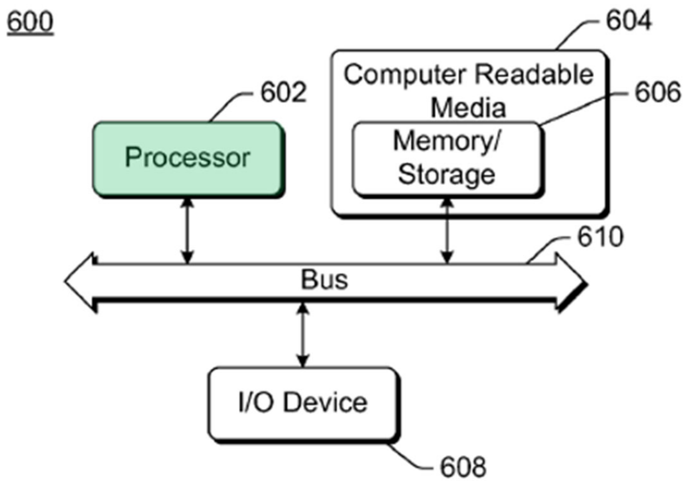


Id., Fig. 1.

b. 1[a]

Zhang discloses or suggests 1[a]. *Daft*, ¶¶1026-27.

Zhang discloses a computing device that contains a processor and computer-readable media (e.g., memory) storing software instructions. *Zhang*, Fig. 6 (annotated), [0063]-[0067]; *Daft*, ¶1027.



c. 1[b]

Zhang discloses or suggests 1[b]. *Daft*, ¶¶1028-29.

Zhang discloses a computing device that contains a display that allows a user to interact with the device and presents information to the user. *Zhang*, [0067]; *Daft*, ¶1029.

d. 1[c]

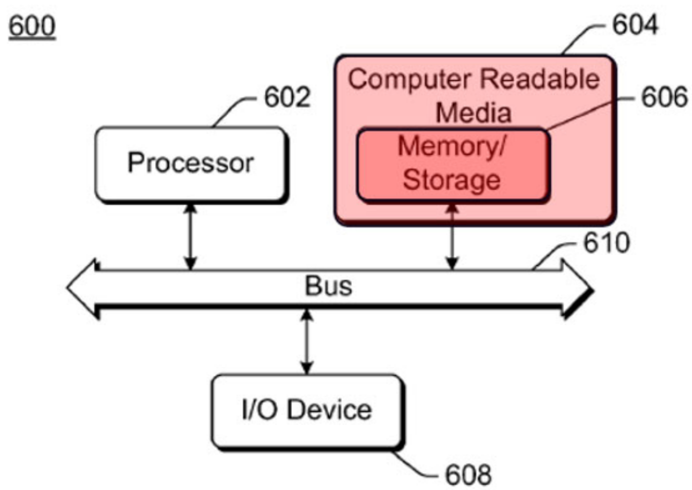
Zhang discloses or suggests 1[c]. *Daft*, ¶¶1030-31.

Zhang discloses a computing device that contains an image capture component, which a POSITA would have understood to be a camera. §VII.C.3.a-(1[pre]); *Daft*, ¶1031.

e. 1[d]

Zhang discloses or suggests 1[d] under both potential interpretations. §VI; *Daft*, ¶¶1032-36.

Zhang discloses a computing device that contains computer-readable media (e.g., memory) storing software instructions. *Zhang*, Fig. 6, [0063]-[0067].



Zhang also discloses that the image capture component and live face detection module (104) can be separate computing devices that communicate and send data (including biometric facial feature data) over a variety of networks, such as the Internet, a local area network (LAN), an intranet, etc. *Zhang*, [0014]; *Daft*, ¶1034.

Although Zhang does not expressly mention that the data is sent to a "server," a POSITA would have found it obvious that Zhang's separate computing device is a server because servers were well-known and conventional network infrastructure that could handle remote processing. *Daft*, ¶1035; Ex-1016, Abstract, [0040]-[0043];

Ex-1012, Fig. 1A, 5:24-50. A POSITA would have further understood that, when using a server as the separate computing device running live face detection module (104), the server would store machine-readable instructions to carry out Zhang's disclosed process and would send those instructions to the separate image-capture device, which, in turn, would store those instructions in its own local memory to execute facial recognition. *Daft*, ¶1035. Alternatively, a POSITA would have found it obvious that a server would be provided with a facial-authentication system (with a camera, 1[c]) due to the sensitivity of information that can be stored on servers. *Id.*, ¶1036.

f. 1[d1]

Zhang discloses or suggests 1[d1]. *Daft*, ¶¶1037-1039.

Zhang discloses capturing a first image of a user as part of the authentication method. *Zhang*, [0016], [0021]. A POSITA would have understood that Zhang's process captures an image of the user's face at a first location which is a first distance from the image capture component 102. *Zhang*, [0016]; *Daft*, ¶¶1038-39.

g. 1[d2]

Zhang discloses or suggests 1[d2]. *Daft*, ¶¶1040-1042.

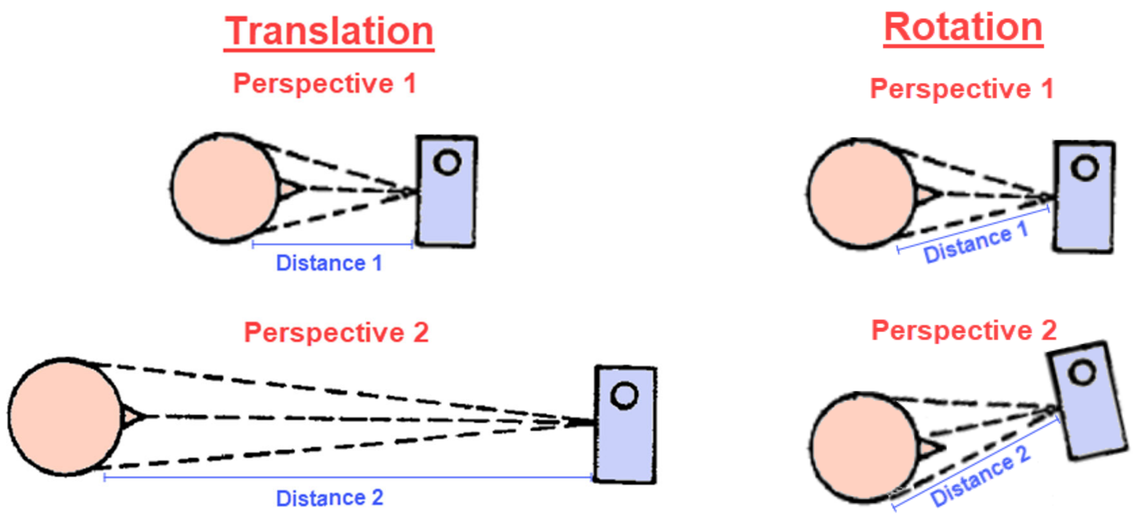
Zhang discloses processing the first image to extract "feature points" of the face from the image. *Zhang*, [0027] ("[O]ne or more feature points are extracted from two images… A variety of different feature points can be extracted, such as a corner of an eye, a corner of a mouth, a tip of a nose, and so forth."), [0026] (disclosing "software, firmware, hardwire, or combin[ed]" implementations). A POSITA would have understood these extracted feature points to be data (and specifically biometric data) because the identification involves using a computer operating on data to characterize an individual's unique physical characteristics. *Daft*, ¶¶1041-42.

h. 1[d3]

Zhang, alone or combined with Tanii, teaches 1[d3]. *Daft*, ¶¶1043-48.

Zhang discloses capturing a second image of a user as part of the authentication method. *Zhang*, [0016] ("Image capture component 102 captures multiple images"). *Daft*, ¶1045.

Although Zhang does not expressly disclose that the second image is captured at a second distance different from the first distance of the first image, this would have been obvious to a POSITA. *Daft*, ¶1046; §VII.C.1-(Zhang). Specifically, Zhang discloses a "3D structure determination module 112" that uses a "homography" technique to distinguish between a real face and a picture of a face by, *inter alia*, transforming a first image to the perspective of a second image and comparing the two. *Zhang*, [0024], [0026]-[0035]; *Daft*, ¶1046; §VII.C.1-(Zhang). A POSITA therefore would have understood that, to perform a homography transform, the "multiple images" Zhang captures would be from different perspectives of the face. *Daft*, ¶1046. And a POSITA would have further understood that, to take multiple images of different perspectives of the face, the position of the face relative to the camera would have to change—either by rotating/translating the position of the user or camera to the side, or changing the distance between the user and camera, as depicted below (showing a top-down view of a face and camera):



Not only would a POSITA have understood that providing images at different distances allows for a greater understanding of depth between objects in the scene, as explained in the Zhang paper (*Daft*, ¶1046; Ex-1013, 22-25), but also that taking pictures at different distances with a camera lens that induces distortions would *enhance* the performance of Zhang's homography-transformation procedure. *Daft*, ¶1046; §VII.C.2-(motivation). Thus, Zhang at least renders this limitation obvious.

Moreover, taking two images at different distances would have been obvious in view of other prior art. A POSITA would have understood that distortions caused by camera lenses can indicate depth in the object being captured, as exemplified by Tanii. *Daft*, ¶1047; §VII.C.2-(motivation). Accordingly, to the extent Zhang does not disclose or suggest this limitation, a POSITA would have been motivated to modify Zhang in view of Tanii to capture a second image at a second distance, while taking advantage of these distance-induced distortions to distinguish between three-dimensional faces and two-dimensional pictures. *Daft*, ¶1047; §VII.C.2-(motivation).

Because Zhang (alone or combined with Tanii) teaches capturing a plurality of images at different distances between the user and the camera, a POSITA would have understood that either the camera or user must physically move in between image captures. *Daft*, ¶1048. And where the camera is attached to a mobile computing device (e.g., a phone), it would have been obvious that the user could easily move the camera relative to the user's head. *Id.*

i. 1[d4]

Zhang discloses or suggests 1[d4]. *Daft*, ¶¶1049-50.

Zhang discloses processing the second image to obtain second feature-point biometric data from the image. *Zhang*, [0026]-[0027]; §VII.C.3.g-(1[d2]); *Daft*, ¶1050.

j. 1[d5]

Zhang, alone or combined with Tanii, teaches 1[d5]. *Daft*, ¶¶1051-55.

Zhang discloses that "[t]he feature points extracted…are matched across the first and second images (act 304)," and those feature points constitute data. *Zhang*, [0028]; §VII.C.3.i-(1[d4]). Furthermore, Zhang discloses that the matching process may "determine[] whether the first and second images include the same face," including "during the matching of feature points in 304, if all (or at least a threshold number) of the feature points cannot be matched then it is determined that the first and second images are of different faces." *Zhang*, [0038]; *Daft*, ¶1052.

After determining the homography matrix between the first and second images, a "warped" version of the first image is created and then compared to the second image to determine whether differences exist. *Zhang*, [0025], [0031]. As part of the comparison, "any of a variety of conventional face detection algorithms or face recognition algorithms can be used to detect the face within each image, and the selected locations are the locations that are part of a face within at least one of the warped and second images." *Zhang*, [0032]. A POSITA would have understood Zhang discloses comparing a first biometric data (e.g., the facial-feature locations in the first warped image) and second biometric data (e.g., the facial-feature locations in the second image) to determine whether differences between the two exist, in which it would be expected that a live face would produce sufficient differences between the two images due to relative movement of the image capture component 102 (camera). *Daft*, ¶¶1053-54.

However, a POSITA would have also been aware (as exemplified by Tanii) that differences between two images—one with distance-induced distortions and one without—can also be used to distinguish between three-dimensional face and a two-dimensional picture. §VII.C.2-(motivation); *Daft*, ¶1055. And a POSITA would have been motivated to utilize these expected distortions as either a supplemental or alternative verification of three-dimensionality of a face. §VII.C.2-(motivation); *Daft*, ¶1055. In doing so, a POSITA would have understood verification using distance-induced distortion matches the positions of features across first and second images—as Zhang already discloses. *Id.* But rather than use that comparison to calculate a homography matrix, the comparison would evaluate whether one image exhibits the expected distance-induced distortion when the user's face is captured at a close distance to the camera, and the other image does not exhibit similar distance-induced distortion when captured further from the camera. *Id*. Thus, a POSITA would have appreciated that, when modifying Zhang to evaluate differences caused by distance-induced distortions, the presence of such distortions (due to the change in distance of the camera) indicates a three-dimensional face. *Zhang*, [0025], [0034]; *Daft*, ¶1055.

k. 1[d6]

Zhang, alone or combined with Tanii, teaches 1[d6]. *Daft*, ¶¶1056-58.

Zhang discloses captured images are determined to be of a three-dimensional face when sufficient differences in the image exist (*Zhang*, [0031]), including when first data (the position of facial features after the first image is "warped" (undergoes homography transformation)) does not match the second data (*Zhang*, [0032]-[0034]). These differences would be expected due to a change in perspective (rotation and/or distance) of the camera between the two images. §§VII.C.3.h-(1d[3]), VII.C.3.j-(1[d5]); *Daft*, ¶1057.

Moreover, to the extent Zhang does not disclose this limitation, Zhang combined with Tanii does. §§VII.C.3.h-(1d[3]), VII.C.3.j-(1[d5]); *Daft*, ¶1058. Specifically, a POSITA would have appreciated that, when modifying Zhang to evaluate differences caused by distance-induced distortions, a three-dimensional face would be indicated when one of the two sets of biometric data exhibits expected distance-induced distortions due to the change in distance of the camera. §§VII.C.3.h-(1d[3]), VII.C.3.j-(1[d5]); *Daft*, ¶1058.

4. Claims 2 and 11

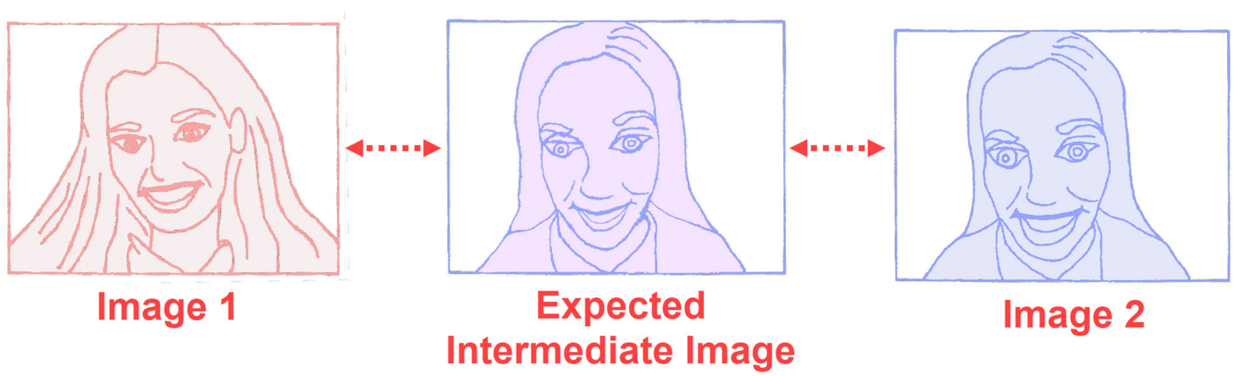
a. 2[a] and 11[a]

Zhang, alone or combined with Tanii, teaches 2[a] and 11[a]. *Daft*, ¶¶1059-63, 1102.

Zhang discloses that, as part of the authentication process, the two images being compared may be "non-adjacent," meaning additional images exist between the two images. *Zhang*, [0035]-[0036]; *Daft*, ¶1060. In such instances, Zhang discloses performing some of the processes—such as "feature point extraction and feature point matching"—using the intermediate images to "facilitate the feature matching process when matching features across two images with one or more intervening images." *Zhang*, [0036]; *Daft*, ¶1060. Moreover, Zhang discloses that the homography-transformation process can be applied to multiple pairs of images. *Zhang*, [0037]; *Daft*, ¶1060.

When a set of intermediate images exist *between* the first and second images, a POSITA would have been motivated to generate interpolated predictions of what those intermediate images should look like, and compare the intermediate images to those predictions because building predictive, interpolative models for biometric comparison was well known and further prevented spoofers from making artificial changes between image captures. *Daft*, ¶1061; Ex-1036, 8:19-27.

Similarly, Zhang combined with Tanii would further render this limitation obvious. Specifically, a POSITA would have found it obvious to modify Zhang to evaluate images for distance-induced distortions. *Id.*, ¶1062. Tanii further discloses that distance-induced distortions increase as distance between the face and camera decreases. *See Tanii*, [0048]. Thus, a relationship exists between the extent of distance-induced distortion and distance, such that, when two (e.g., non-adjacent) images are captured of a face—with different degrees of distance-induced distortion—there is some intermediate set of distances between the face and camera in which distance-induced distortion is reduced. *Daft*, ¶1062. A POSITA reading Zhang—which discloses processing and evaluating these intermediate images—in view of Tanii therefore would have been motivated to interpolate intermediate biometric data with an intermediate, interpolated amount of distance-induced distortion based on the two non-adjacent images captured to create an array of potential distance-induced distortions that would indicate depth in a three-dimensional face (e.g. intermediate positions of the facial features representing a transition from the distance-distorted blue lines to normal red-lined images below). *Id.*, ¶1063.



b. 2[b] and 11[b]

Zhang, alone or combined with Tanii, teaches 2[b] and 11[b]. *Daft*, ¶¶1064-66, 1103.

Zhang discloses capturing a series of intermediate images between two non-adjacent images. *Zhang*, [0035]-[0037]; §VII.C.4.a-(2[a]). A POSITA would have understood these intermediate images are at different positions (e.g., rotation or translation) between the first and second images. *Daft*, ¶1065; §VII.C.4.a-(2[a]).

Moreover, when modifying Zhang in view of Tanii to interpolate intermediate biometric data attributable to distance-induced distortions, a POSITA would have been further motivated to capture a third image at a distance that correlates to one of the interpolated data sets for further authentication of three-dimensional depth of the face in the captured images. *Daft*, ¶1066; §VII.C.4.a-(2[a]).

c. 2[c] and 11[c]

Zhang, alone or combined with Tanii, teaches 2[c] and 11[c]. *Daft*, ¶¶1067-69, 1104.

Zhang discloses processing a third (intermediate) image to obtain third data. *Zhang*, [0036] ("the feature point extraction and feature point matching in acts 302 and 304 can be generated for each adjacent pair of images in the sequence, which can facilitate the feature matching process when matching features across two images with one or more intervening images."); *Daft*, ¶1068.

Moreover, when modifying Zhang in view of Tanii, a POSITA would have found it obvious to acquire a third image and extract biometric data from the third image to compare it to the interpolated positions of the biometric data based on the first and second images. §§VII.C.4.a-VII.C.4.b (2[a]-2[b]); *Daft*, ¶1069.

d. 2[d] and 11[d]

Zhang, alone or combined with Tanii, teaches 2[d] and 11[d]. *Daft*, ¶¶1070-72, 1105.

Zhang discloses tracking and comparing biometric features between non-adjacent and intermediate images. *Zhang*, [0036]-[0037]; §VII.C.4.a-(2[a]); *Daft*, ¶1071. A POSITA would have understood that, when predicting (i.e., interpolating) intermediate images, the biometric features from the *actual* intermediate images would be compared to the predicted features to determine whether there is a match. *Daft*, ¶1072.

Moreover, when modifying Zhang in view of Tanii, a POSITA would have found it obvious to acquire a third image and extract data from the third image to compare it to the interpolated positions of the data based on the first and second images to determine if there is a match. §VII.C.3.j-(1[d5]); *Daft*, ¶1073.

5. Claims 3 and 12

Zhang combined with Tanii teaches claim 3 and 12. *Daft*, ¶¶1074-76, 1106.

Zhang does not expressly disclose a process for verifying the presence and absence of a feature on the side of a user's head (which a POSITA would have known includes an ear). *Id.*, ¶1076. But as Tanii shows, the absence or presence of an ear is a natural result of the distance between the user and camera. *Id*. Specifically, given sufficient distance between the face and camera, the ears are captured because there is enough distance for light rays from the ears to reach the camera's sensor. *See Tanii*, [0047]-[0048], Figs. 3A-B; *Daft*, ¶1076.

Fig.3A

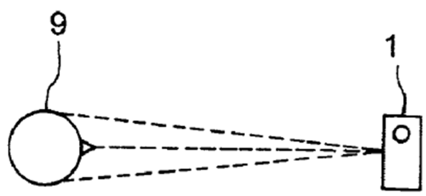
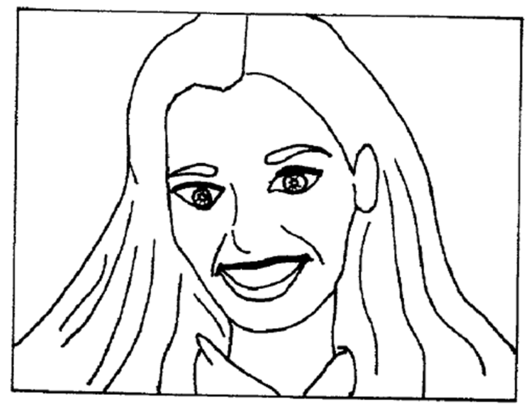


Fig.3B



But when a face is too close, the ears will not be captured because light rays from the ears cannot reach strike the camera's sensor. *See Tanii*, [0047]-[0048], Fig. 4B; *Daft*, ¶1076.

Fig.4A

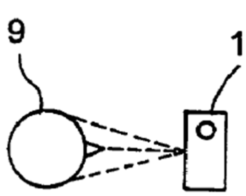
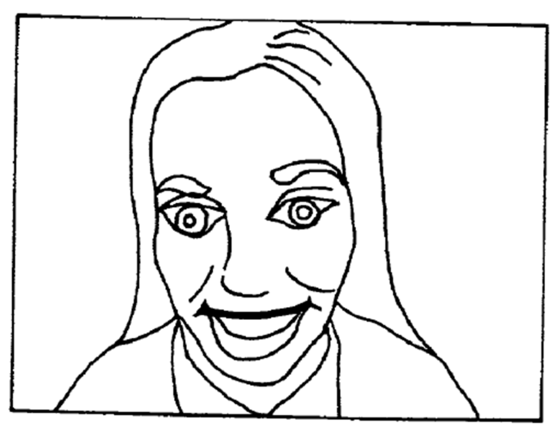


Fig.4B



This was a well-known consequence of distance-induced distortion when capturing images of a face. *See Daft*, ¶1076.

Accordingly, a POSITA would have appreciated (based on Tanii) that the presence of features on the side of a user's head (e.g., ears) in one image and absence of the user's ears in another image at a closer distance would further indicate a three-dimensional face, and a POSITA would have found it obvious to modify Zhang to verify this. *Daft*, ¶1076.

6. Claim 5

a. 5[a]

Zhang discloses or suggests claim 5[a]. *Daft*, ¶¶1077-79.

Zhang discloses that a user can be authenticated "by comparing one or more of the images 106 to previously captured images of the user 108." *Zhang*, [0017]. A POSITA would have understood that Zhang describes conventional biometric authentication that compares recently acquired biometric data to biometric data previously acquired during an enrollment process. *Daft*, ¶¶1078-79.

b. 5[b]

Zhang discloses or suggests claim 5[b]. *Daft*, ¶¶1080-82.

Zhang discloses a conventional facial authentication procedure by comparing biometric data to previously acquired biometric data during enrollment. §VII.C.6.a-(5[a]); *Daft*, ¶1081. Zhang also discloses images can be evaluated to determine whether they capture the same user's face by matching feature points across both images, consistent with conventional practices. *Zhang*, [0038]; *Daft*, ¶1081.

A POSITA would have understood that, although Zhang discloses matching feature points across images using a three-dimensional verification procedure, these same processes are conventionally used when authenticating a user as well. *Daft*, ¶¶1081-82. Moreover, a POSITA would have understood that, because facial authentication must be able to account for different conditions (e.g., perspective of the face in the image, lighting, facial accessories, facial hair, etc.) the use of thresholds was a common practice to provide a *minimum* similarity between faces to be considered a match. *Id.*, ¶1081. Thus, a POSITA would have been motivated to utilize thresholds when matching faces, consistent with both convention and Zhang's disclosure. *Id.*, ¶1082.

7. Claims 6 and 15

Zhang combined with Tanii teaches claim 6. *Daft*, ¶¶1083-85, 1107.

Zhang's facial-authentication process can be implemented on a variety of hand-held computing devices, such as a cellular phone, a digital camera, or a video camera. *Zhang*, [0013]. Moreover, Tanii notes that distance-induced distortions often occur in mobile devices that have wide-angle lenses, and the amount of distortion is dictated by the distance between the user and the camera. *Tanii*, [0007], [0047]-[0048], Figs. 3A-B, 4A-B; *Daft*, ¶1084.

When implementing a facial-authentication process on a handheld mobile device consistent with Zhang (alone or combined with Tanii, §VII.C.3.h-(1[d3])), a POSITA would have further understood the user holds the computing device at a first distance for the first image, and a second distance for the second image (e.g., by extending and retracting the user's arm) because that is a convenient and obvious way of changing the distance between the device and the user's face. *Daft*, ¶1085.

8. Claims 7 and 16

Zhang discloses or suggests claim 7 and 16's additional limitation. *Daft*, ¶¶1086-87, 1108. §§VII.C.3.g-(1[d2]), VII.C.3.i-(1[d4]).

Zhang discloses processing the first image to extract "feature points" from the image. *Zhang*, [0027] ("[O]ne or more feature points are extracted from two images… A variety of different feature points can be extracted, such as a corner of an eye, a corner of a mouth, a tip of a nose, and so forth."). A POSITA would have understood these extracted feature points to be biometric data that identifies unique physical characteristics (e.g., characteristics of a face). *Daft*, ¶¶1086-87.

9. Claims 8 and 17

Zhang discloses or suggests claim 8 and 17's additional limitation. *Daft*, ¶¶1088-89, 1109.

Zhang discloses processing the multiple images to extract "feature points" from the image that correspond to characteristics of a user's face. *Zhang*, [0027], ("[O]ne or more feature points are extracted from two images… A variety of different feature points can be extracted, such as a corner of an eye, a corner of a mouth, a tip of a nose, and so forth."). A POSITA would have understood these extracted feature points to comprise a mapping of the user's facial features. *Daft*, ¶1089. Moreover, a POSITA would have understood that computer-based facial recognition conventionally involves mapping facial features to data structures. *Daft*, ¶1089.

10. Claim 9

Zhang, alone or combined with Tanii, teaches claim 9. *Daft*, ¶¶1090-92.

Zhang and Tanii both teach or suggest moving the camera to capture images at two different distances. §VII.C.3.h-(1[d3]). A POSITA would have understood that, when moving the camera to capture images from different distances, either the user's face would remain stationary (e.g., steady) and the camera would be moved, or the user's face would be moved and the camera would remain stationary. *Daft*, ¶¶1091-92. However, a POSITA would have appreciated that holding the user's face steady and moving the camera would be more user friendly with, for example, a handheld computing device than forcing the user to move closer and further from the camera while holding the device steady. *Id*.

11. Independent Claim 10

a. 10[pre]

If limiting, Zhang discloses or suggests 10[pre]. *Daft*, ¶¶1093-94.

Zhang discloses a method "to determine whether a face in multiple images is a 3D structure or a flat surface," *Zhang*, [0026], Figs 2-3; *see also, e.g.*, *id.*, Abstract, [0003], to "authenticate a user for particular access," *id.*, [0012]; *Daft*, ¶1094.

b. 10[a]

Zhang discloses or suggests 10[a] for the reasons discussed in §VII.C.3.f-(1[d1]). *Daft*, ¶1095.

c. 10[b]

Zhang discloses or suggests 10[b] for the reasons discussed in §VII.C.3.g-(1[d2]). *Daft*, ¶1096.

d. 10[c]

Zhang, alone or combined with Tanii, teaches 10[c] for the reasons discussed in §VII.C.3.h-(1[d3]). *Daft*, ¶1097.

e. 10[d]

Zhang, alone or combined with Tanii, teaches 10[d] for the reasons discussed in §VII.C.3.h-(1[d3]). *Daft*, ¶1098.

f. 10[e]

Zhang discloses or suggests 10[e] for the reasons discussed in §VII.C.3.i-(1[d4]). *Daft*, ¶1099.

g. 10[f]

Zhang, alone or combined with Tanii, teaches 10[f] for the reasons discussed in §VII.C.3.j-(1[d5]). *Daft*, ¶1100.

h. 10[g]

Zhang, alone or combined with Tanii, teaches 10[g] for the reasons discussed in §VII.C.3.k-(1[d6]. *Daft*, ¶1101.

12. Claim 19

Zhang, alone or in combination with Tanii, teaches holding the user's face steady and moving the camera from a first to a second location. §VII.C.10-(cl.9); *Daft*, ¶1110.

13. Claim 20

Zhang discloses or suggests claim 20's additional limitation. *Daft*, ¶¶1111-12.

Zhang discloses that image capture component (102) and live face detection module (104) can both be implemented on the same computing device. *Zhang*, Fig. 6, [0014]; *Daft*, ¶1112. In that implementation, it would have been obvious to a POSITA to maintain the biometric data in the device memory to perform the live face detection locally. *Daft*, ¶1112.

14. Claim 21

Zhang, alone or combined with Tanii, teaches claim 21. *Daft*, ¶¶1113-14.

Zhang discloses an image capture component (102) and a live face detection module (104) that can both be implemented on the same computing device. *Zhang*, Fig. 6, [0014]; *Daft*, ¶1114.

15. Independent Claim 22

a. 22[pre]

If limiting, Zhang discloses or suggests 22[pre]. §VII.C.3.a-(1[pre]); *Daft*, ¶1115.

b. 22[a]

Zhang discloses or suggests 22[a] for the reasons discussed in §VII.C.3.f-(1[d1]). *Daft*, ¶1116.

c. 22[b]

Zhang, alone or combined with Tanii, teaches 22[b] for the reasons discussed in §VII.C.3.h-(1[d3]). *Daft*, ¶1117.

d. 22[c]

Zhang, alone or combined with Tanii, teaches 22[c] for the reasons discussed in §VII.C.3.h-(1[d3]). *Daft*, ¶1118.

e. 22[d]

Zhang, alone or combined with Tanii, teaches 22[d] for the reasons discussed in §§VII.C.3.j-(1[d5], VII.C.3.k-(1[d6]). *Daft*, ¶1119.

f. 22[e]

Zhang discloses or suggests 22[e]. *Daft*, ¶¶1120-21.

After Zhang's live face detection module analyzes two images, the module "outputs an indication 110 of whether images 106 include a live face or a picture of a face." Zhang, [0017]. The indication may be, e.g., "yes" and "authenticate" (i.e., user is verified) or "no" and "do not authenticate" (i.e., the user is not verified). *Id*.; *Daft*, ¶1121.

16. Claim 23

Zhang, alone or combined with Tanii, teaches 22[b] for the reasons discussed in §§VII.C.3.g-(1[d2]), VII.C.3.i-(1[d4]). *Daft*, ¶1122.

17. Claim 24

Zhang, alone or combined with Tanii, teaches claim 24's additional limitation. §§VII.C.3.g-(1[d2]), VII.C.3.i-(1[d4]), VII.C.10-(cl.9); *Daft*, ¶1123.

D. Ground 2B: Obviousness in view of Zhang, Tanii, and Tahk (Claims 4, 13-14)

1. Motivation to Combine

A POSITA would have been motivated to combine Tahk with Zhang, alone or combined with Tanii, because Tahk provides user-friendly ways of ensuring a face presented for authentication is framed properly for image capture, which a POSITA would have recognized as useful for the authentication procedure taught by the Zhang-Tanii combination that requires multiple images of the face. *Daft*, ¶1124. A POSITA would have understood that providing a live preview would allow the user to adjust position relative to the camera, and on-screen prompts such as written instructions and ovals to frame the face would provide express guidance to ensure an image is captured that is useable for authentication. *Id.*, ¶1125.

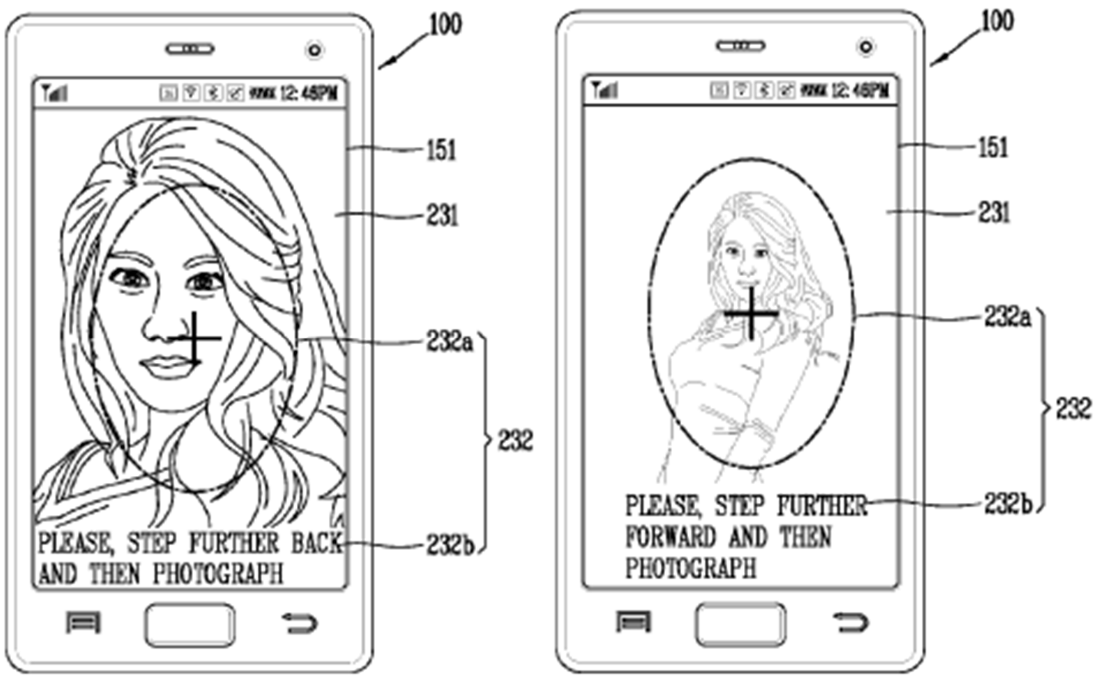
Moreover, providing real-time image feedback, written instructions, and oval shapes to frame a face during an authentication process were well-known and conventional techniques at the time of the ’910 Patent. *See, e.g.*, Ex-1009, 7:16-8:7, Figs. 6B-7C (live image preview and oval prompts); Ex-1010, 5:31-32, 6:3-4; Ex-1032. A POSITA would have therefore been motivated to implement these types of feedback to ensure images are captured properly for facial authentication. *See Samsung*, IPR2022-00100, Paper 30, 23-24, 34-35 (conventional teachings supported by background references obvious); *Daft*, ¶1126 (citing Ex-1034, Ex-1035).

2. Claims 4 and 13

Zhang combined with Tanii and Tahk teaches claims 4 and 13. *Daft*, ¶¶1127-30.

Zhang discloses taking a series of images sufficient to calculate a homography matrix. *See, e.g.*, *Zhang*, [0026], Figs. 1, 3 and a POSITA would have been motivated to take the series of images at different distances, (§§VII.D.1-(motivation), VII.C.3.h -(1[d3])). However, Zhang does not expressly disclose providing prompts to guide a user through different camera positions that would enhance calculations of the homography matrix. *Daft*, ¶1128.

Tahk teaches that using one or more prompts on a screen ensures images of the face are captured at the correct distances. *See, e.g.*, Tahk, Figs. 8A-B ("Please step further back" and "Please step further forward," and presenting an oval to frame the face at the correct distance):



Daft, ¶ 1129.

A POSITA would have been motivated by Tahk to modify Zhang (alone or combined with Tanii) to prompt a user to alter the distance of the camera to either capture sufficiently different images for a homography transformation (Zhang) or to capture an image with distance-induced distortion (Tanii) to distinguish a live, three-dimensional face from a two-dimensional photo. *Id.*; *see also* §VII.D.1-(motivation). *Daft*, ¶1129.

3. Claim 14

Zhang combined with Tanii and Tahk teaches claim 14's additional limitation. *Daft*, ¶¶1131-32.

Although neither Zhang nor Tanii expressly teach prompting a user during the facial-authentication process, Tahk does, including using oval prompts to frame a user's face. §VII.D.2-(cl.4). A POSITA would have been motivated to modify Zhang (alone or combined with Tanii) to provide such oval-shaped prompts because they are a natural shape to appropriately size and frame a face at different distances. §VII.D.1-(motivation); *Daft*, ¶1132.

E. Ground 2C: Obviousness in view of Zhang, Tanii, and Hoyos (Claim 18)

1. Hoyos (Ex-1010)

Hoyos discloses another facial-authentication system for a computer device involving a "liveness" determination of the user's face. *Hoyos*, Abstract. To perform this check, Hoyos displays images on the device screen while capturing first and/or second images of the user and evaluates the captured images for reflections of the displayed images. *Hoyos*, [0018]-[0019]; [0033]-[0035]; *Daft*, ¶¶127-28.

2. Motivation to Combine

Zhang discloses implementing a process to verify the three-dimensionality of a user's face by capturing images of a user with a mobile computing device such as a phone or laptop. *See, e.g.*, *Zhang*, [0013]. For mobile devices, a POSITA would have understood that mobile devices are often provided with user-facing cameras. *Daft*, ¶1133.

Although Zhang discloses using homography transformation to distinguish real, three-dimensional faces from pictures of a face, (§VII.C.1-(Zhang)) a POSITA would recognized that Zhang's process could be spoofed by presenting a non-planar picture of a face (e.g., by bending the picture in a way to trick the system), because the homography transformation looks for *planar* pictures. *Daft*, ¶1134. Thus, a POSITA would have been motivated to look for secondary methods to ensure the user's face is truly a real, three-dimensional face, as is common for biometric authentication. *Id.* To that end, a POSITA would have been aware of the use of reflectance of light off a face as a liveness check, as exemplified by Hoyos. *Id.*, ¶1135. A POSITA would have therefore incorporated a secondary liveness check based on reflectance to guard against spoofing of Zhang's homography-transformation process. *Id.*

A POSITA would have had a reasonable expectation of success in making this modification because Zhang's homography transformation and Hoyos's reflectance measure operate on two distinct, modular principles that can work together; Zhang requires two images from two different perspectives, and Hoyos requires reflecting different light patterns during image capture. *Id.*, ¶1136.

3. Claim 18

Zhang combined with Tanii and/or Hoyos teach claim 18's additional limitation. *Daft*, ¶¶1137-38.

Zhang does not expressly disclose illuminating the device's screen when capturing images to detect a reflection of the displayed image off the user's face. Hoyos, however, teaches that measuring reflectance of displayed images is a well-known method to verify the liveness of the user. *Hoyos*, [0018]-[0019]; [0033]-[0035]. A POSITA would have been motivated to modify Zhang to incorporate Hoyos's reflectance determination to provide further verification that the user is presenting a real, three-dimensional face instead of a picture of a face that can be altered to bypass Zhang's homography transformation process. §VII.E.2-(motivation); *Daft*, ¶1138.

VIII. DISCRETIONARY DENIAL IS INAPPROPRIATE

A. *Fintiv*

Petitioner expects that PO will not raise *Fintiv* here, but even if it does, the Board should not deny institution on that basis.

In June 2024, PO filed suit against Petitioner (*FaceTec, Inc. v. Jumio Corp.*, 3:24-cv-3623 (N.D. Cal.)) asserting infringement of four patents, including the '910 Patent. *See* Ex-1046. The Court set a trial date of August 10, 2026, which would be after any expected final written decision here (May 2026). Ex-1041. But

Petitioner also intends to request a stay of the district court litigation to allow the IPR to be decided.

Accordingly, discretionary denial under *Fintiv* is unwarranted.

B. 325(d)

The Board should not exercise its discretion under §325(d) to deny institution. *See Becton, Dickinson & Co. v. B. Braun Melsungen AG*, IPR2017-01586, Paper 8, 17–18 (Dec. 15, 2017) (precedential); *Advanced Bionics, LLC v. MED-EL Elektromedizinische Geräte GmbH*, IPR2019-01469, Paper 6, 9–11 (Feb. 13, 2020) (precedential).

Considering the first part of the *Advanced Bionics* framework, including *Becton, Dickinson* factors (a), (b), and (d), a related Derakhshani publication (2014/0198959) was cited—but never applied—during prosecution. Indeed, *no* art was applied during prosecution. That is not a basis to invoke discretionary denial. *See GMG Prods. LLC v. Traeger Pellet Grills LLC*, PGR2019-00024, Paper 17, 27 (July 17, 2019) (finding listing on an IDS alone factor "does not favor denying institution" because "the Examiner did not provide any detailed assessment of [the prior art, and], instead, only indicat[ed] the references had been considered"). Regardless, the Examiner never considered Zhang, or any prior-art combination involving Tanii. Accordingly, the same or substantially the same art or arguments were not previously presented to or considered by the Office; there is no need to go to *Advanced Bionics* part two. But even if the Board were to reach step two, the Examiner materially erred by failing to apply Derakhshani's teachings, alone or combined with other art, related to verifying the three-dimensionality of a face presented for authentication.

IX. Conclusion

Inter Partes Review of the challenged claims is respectfully requested.

X. Mandatory Notices

A. Real Party in Interest

Petitioner identifies itself as real parties in interest.

B. Related Matters

To the best of Petitioner's knowledge, the '910 Patent has been involved in the following district court litigation:

FaceTec, Inc. v. Jumio Corporation, Case No. 5:24-cv-3623 (N.D. Cal.)

C. Notice of Counsel and Service Information

Lead Counsel:	**Backup Counsel:**
Jessica C. Kaiser, Reg. No. 58,937 kaiser-ptab@perkinscoie.com PERKINS COIE LLP 1900 Sixteenth St., Suite 1400 Denver, CO 80202-5255 (303) 291-2300 (telephone) (303) 291-2400 (fax)	Jon R. Carter, Reg. No. 75,145 Carter-ptab@perkinscoie.com PERKINS COIE LLP 1155 Avenue of the Americas, 22nd Fl. New York, NY 10036 Tel: (212) 261-6826 Fax: (212) 977-1649

Lead Counsel:	**Backup Counsel:**
	David Pekarek Krohn, Reg. No. 69,443 pekarekkrohn-ptab@perkinscoie.com PERKINS COIE LLP 33 East Main Street, Suite 201 Madison, WI 53703-3095 Tel: (608) 663-7460 Fax: (608) 663-7499
	Grant Kinsel Pro Hac Vice to be submitted kinsel-ptab@perkinscoie.com PERKINS COIE LLP 1201 Third Avenue, Suite 4900 Seattle, WA 98101-3099 Tel: (206) 359-8000
	John C. Moulder, Reg. No. 70,490 moulder-ptab@perkinscoie.com PERKINS COIE LLP 700 Thirteenth Street, N.W. Suite 800 Washington, DC 20005-3960 Telephone: (202) 654-3337
	Tyler Knox, Reg. No. 79,948 knox-ptab@perkinscoie.com PERKINS COIE LLP 1900 Sixteenth Street, Suite 1400 Denver, CO 80202 Tel: (303) 291-2300
	Matthew A. Lembo, Reg. No. 75,633 lembo-ptab@perkinscoie.com PERKINS COIE LLP 1155 Avenue of the Americas, 22nd Fl. New York, NY 10036 Tel: (212) 261-6826 Fax: (212) 977-1649

Respectfully submitted,

/ *Jessica Kaiser* /

Jessica Kaiser
Reg. No. 58,937
Attorney for Petitioner

PERKINS COIE LLP
1900 Sixteenth Street, Suite 1400
Denver, CO 80202-5255

Date: November 7, 2024

CERTIFICATE OF WORD COUNT UNDER 37 CFR §42.24(D)

Pursuant to 37 C.F.R. §42.24(a), Petitioner hereby certifies that portions of the above-captioned Petition for *Inter Partes* Review of U.S. Patent 11,874,910, in accordance with and reliance on the word count provided by the word-processing system used to prepare this Petition, that the number of words in this paper is 13,831. Pursuant to 37 C.F.R. §42.24(a), this word count is in compliance and excludes the table of contents, table of authorities, mandatory notices under §42.8, certificate of service, certificate of word count, appendix of exhibits, and any claim listing. This word count was prepared using Microsoft Word.

Respectfully submitted,

/ *Jessica Kaiser* /

Jessica Kaiser
Reg. No. 58,937
Attorney for Petitioner

Date: November 7, 2024
PERKINS COIE LLP
1900 Sixteenth Street, Suite 1400
Denver, CO 80202-5255

CERTIFICATE OF SERVICE

The undersigned hereby certifies that true copies of the Petition for *Inter Partes* Review of U.S. Patent No. 11,874,910 and supporting materials (Exhibits and Power of Attorney) were served via overnight delivery on the Patent Owner at the correspondence address of record as listed on PAIR:

WEIDE & MILLER, LTD.
10655 PARK RUN DRIVE
SUITE 100
LAS VEGAS, NV 89144

A courtesy copy was also sent via electronic mail to Patent Owner's litigation counsel listed below:

Nathaniel L. Dilger (CA Bar No. 196203)
Email: ndilger@onellp.com
Peter R. Afrasiabi (CA Bar No. 193336)
Email: pafrasiabi@onellp.com
Taylor C. Foss (CA Bar No. 253486)
tfoss@onellp.com
William J. O'Brien (CA Bar No. 99526)
Email: wobrien@onellp.com

Respectfully submitted,

/ *Jessica Kaiser* /

Jessica Kaiser
Reg. No. 58,937
Attorney for Petitioner

Date: November 7, 2024