

**UNITED STATES PATENT AND TRADEMARK OFFICE BEFORE THE
PATENT TRIAL AND APPEAL BOARD**

Inter IKEA Systems B.V.,

Petitioner

v.

Pointwise Ventures, LLC

Patent Owner

Case No.: 2025-00798

Patent No. 8,471,812

Declaration of Dr. Oded Gottesman

Regarding U.S. Patent No. 8,471,812

TABLE OF CONTENTS

	Page
I. Introduction & Qualifications.....	1
II. Summary of My Qualifications.....	3
A. Education, Employment, and Research	3
B. Publications	6
C. Prior Testimony.....	7
D. Materials Considered.....	7
III. Legal Standards Applied in This Report.....	7
A. Anticipation	9
B. Obviousness	10
IV. Technical Background	13
A. The '812 Patent's Background.....	13
B. Cell Phones with Digital Cameras	15
C. Cameras and Reticles	19
D. Object Detection.....	25
V. The '812 Patent.....	29
A. The '812 Specification	29
B. The '812 Patent Claims	35
C. The '812 Patent's Prosecution History.....	36
VI. The Relevant Technical Field and the POSITA	37
VII. Claim Construction	38
VIII. Overview of the Prior Art.....	39
A. Darrell.....	39
B. Boncyk.....	44
C. Fukuda	47
IX. The Asserted Claims are Invalid	48
A. Anticipation by Darrell	49
B. Obviousness over Boncyk and Darrell.....	71

Declaration of Dr. Oded Gottesman Regarding U.S. Patent No. 8,471,812;
IPR No. 2025-00798

C. Obviousness over Boncyk, Darrell, and the Level of Ordinary Skill in the Art.....	93
D. Obviousness over Boncyk, Darrell, and Fukuda	94
E. Secondary Considerations.....	98

I. Introduction & Qualifications

1. My name is Dr. Oded Gottesman. I have been retained as an expert witness by counsel for Inter IKEA Systems B.V. (“Inter IKEA”) in the above captioned proceeding. I have been asked to provide certain opinions regarding U.S. Patent No. 8,471,812 (“the ’812 patent”).

2. It is my opinion that claims 1, 3, and 10-12 of the ’812 patent are anticipated by U.S. Application Patent Publication No. 2005/0162523 A1 (“Darrell”).

3. It is my opinion that claims 1-12 of the ’812 patent are obvious over U.S. Patent Application Publication No. US 2004/0208372 A1 (“Boncyk) in view of Darrell.

4. It is my opinion that claims 5-9 of the ’812 patent are obvious over Boncyk, in view of Darrell, and in view of a person of ordinary level of skill in the art’s knowledge.

5. It is my opinion that claims 1-12 of the ’812 patent are obvious over Boncyk, in view of Darrell, and in view of U.S. Patent Application Publication No. 2004/0125220 (“Fukuda”).

6. This report contains a statement of the opinions that I have formed regarding invalidity to date and provides the basis and reasons for my opinions. In connection with those opinions, I have provided my opinion as to the level of skill

of a person of ordinary skill in the art (“POSITA”), and the appropriate meaning of one of the claim terms to a POSITA. I may offer additional opinions based on further review of materials in this proceeding, including opinions and/or testimony of other experts, including those of Patent Owner.

7. I am being compensated for my time in connection with this matter at my standard hourly rate of \$700/hour. My compensation is not affected by my opinions or the outcome of this proceeding. I do not have any financial interest in the outcome of this proceeding.

8. I reserve the right to amend and/or supplement my opinions to respond to any declarations, reports, or testimony from any of Patent Owner’s expert(s) or witnesses; any briefs or pleadings that Patent Owner may file; and any other matters relating to invalidity of the ’812 claims, the level of skill of a POSITA, and/or the interpretation of claim terms that are proposed by Petitioner or the Patent Trial and Appeal Board (“the Board”). I also reserve the right to prepare and rely on demonstrative exhibits concerning the opinions and their bases as an aid for trial or to assist the Board in connection with this proceeding.

9. I reserve the right to make demonstratives (including product demonstrations, product usage, and videos thereof), charts, graphs, or other similar visual aids for trial based upon the opinions expressed in this report, the data

contained in this report, the exhibits or other things cited in this report and/or attached as exhibits to this report.

II. Summary of My Qualifications

A. Education, Employment, and Research

10. A copy of my curriculum vitae (CV) is Exhibit 1014. As set forth more fully in my CV, over the course of my 30-year career in communication system hardware and software development, I have studied and worked in many aspects of communication systems, including wired, wireless, cellular and GPS communication, image and video processing, including pattern recognition and detection, networks including Internet, and database design and usage.

11. I am currently the President and CEO of Compandent, Inc., a company I founded in February 2001. My work at Compandent includes developing and implementing digital signal processing hardware and software used by various types of communication networks. The hardware and software developed by Compandent is used in a wide variety of communication networks, including cellular networks, satellite networks, telephone networks, packetized networks, and military networks. I have partnered with SignalCom (a Microsoft company), MIT Lincoln Labs, the U.S. Department of Defense (DoD, NSA, DARPA), and North Atlantic Treaty Organization (NATO) to develop and implement a variety of communication system

hardware and software solutions, some of which were adopted as standards by the U.S. DoD and NATO.

12. I have many years of experience with mobile devices, wireless communications, mobile phones with digital camera technology, and processing digital image data. For example, I worked with chipsets for a cellular phone including the cellular network interface which involved wireless communications. I also worked on developing streaming-media systems that included wireless communications. I have also worked with peripheral devices associated with computers, such as computer mice, which involve wireless communications. I am also familiar with laser technology from my advanced degrees in electrical engineering.

13. In the late 1980s, I developed a database application for business management using the MAGIC application generator. Over the years I have developed and implemented many interfaces and protocols for device communication, voice communication, etc., including IP based device communication for VoIP (Voice Over Internet). I monitored IP based interfaces using tools such as Wireshark. I have developed and used many software packages that use data structures, objects, and databases. I have also developed software packages that included APIs and used others' packages using their APIs. I have developed web sites and mobile device applications that use server-client services,

using different APIs, *e.g.*, Google's Android APIs. I have also served as an expert in cases related to client-server systems and databases for e-commerce.

14. My post-secondary education began in 1988 when I graduated from Ben-Gurion University in Be'er Sheva, Israel, where I obtained a Bachelor of Science in Electrical and Computer Engineering, *cum laude*. My B.Sc. study included specializing in digital signal processing, such as audio, image and video processing, pattern recognition, computers, and communications systems. In 1992, I received a Master of Science in Electrical and Computer Engineering from Drexel University in Philadelphia, Pennsylvania. I also received a Ph.D. from the Electrical and Computer Engineering Department of the University of California, Santa Barbara ("UCSB") in 2000. My Ph.D. research included a dissertation entitled "*Analysis-by-Synthesis Waveform Interpolative Coding of Speech at Low Bit Rates.*" After receiving my Ph.D., I remained at UCSB for about a year performing post-doctorate research related to Voice-over-Internet Protocol (VoIP) and other issues related to packetized communication networks. My M.Sc. and Ph.D. studies included advanced studies and research in digital signal processing, such as audio, image and video processing and enhancement, computers, networks, and communications systems.

15. My study and work over the years also included many subjects such as laser radiation, image and video processing, pattern recognition including object

recognition and optical character recognition (OCR). I worked at Optisk Laboratorium in Denmark, a company that manufactured optical filters for lasers and measuring instruments.

16. I have many years of experience in photography and videography. In the early days, I operated film cameras which included operating indicators and sensors through the viewfinder, and I even developed films and photos in my home laboratory. For decades I've been using single-lens reflex (SLR) cameras, both analog and digital SLR (DSLR), in both automatic and manual mode where needed to point the camera to certain object or region and operate it accordingly.

B. Publications

17. I have authored or co-authored numerous publications, which are set forth fully in my CV. I am an author or co-author of the following exemplary publications relating to communication systems and networking:

- Oded Gottesman "Redundant compression techniques for transmitting data over degraded communication links and/or storing data on media subject to degradation," US Patent No. 6,614,370, issued Sep. 2, 2003.
- O. Gottesman and A. Gersho, "Enhanced Waveform Interpolative Coding at Low Bit-Rate," in IEEE Transactions on Speech and Audio Processing, vol. 9, no. 8, Nov. 2001, pp. 786-98.

18. I have been recognized as an expert in patent cases involving mobile devices, cellular phones, wireless location, wireless communications, and digital signal processing. Based on my education and experience, I consider myself to be an expert in the fields of cameras, processing camera images for object detection, laser pointing devices, computer mice, and wireless communication, and including the technologies described and claimed in the '812 patent.

C. Prior Testimony

19. I have previously served as a testifying expert witness in patent cases as set forth more fully in my CV (Ex-1014).

D. Materials Considered

20. I have considered information from various sources in forming my opinions. I have also reviewed the '812 patent, its file history, as well as prior art. All of the materials I have considered are referenced in this report.

III. Legal Standards Applied in This Report

21. I am not an attorney and will not offer opinions on the law. I have, however, an understanding of several principles concerning invalidity that I have used in arriving at my stated conclusions in this report.

22. First, I understand that for an invention claimed in a patent to be found patentable, it must be, among other things, new and not obvious from what was known before the invention was made.

23. I understand that the information that is used to evaluate whether an invention is new and not obvious is generally referred to as “prior art” and generally includes patents and printed publications (*e.g.*, books, journal publications, articles on websites, product manuals, *etc.*). I understand that whether a document constitutes prior art is a legal conclusion.

24. I understand that, in this proceeding, the Petitioner has the burden of showing a reasonable likelihood that at least one of the Challenged Claims of the ’812 patent is invalid in order to institute the inter partes review and has the burden of persuasion by a preponderance of the evidence in a final decision in an inter partes review. I understand that preponderance of the evidence means more likely than not.

25. As I discuss further in the claim construction section below, I understand that the ’812 patent claims must be interpreted in accordance with certain standards of claim construction, and my understanding of these standards is set forth in greater detail below. After being construed, the construed claims are compared to the prior art.

26. I understand that there are two ways in which prior art may render a patent claim unpatentable. First, the prior art can be shown to “anticipate” the claim. Second, a claim is invalid if it is obvious over the prior art.

27. In determining the appropriate level of ordinary skill, my understanding is that the following factors can be considered: the inventor's educational background and experience, the types of problems encountered in the art, prior art solutions to those problems, the rapidity with which innovations are made in the art, the sophistication of the technology, the educational background and experience of those working in the art, and the like.

A. Anticipation

28. I understand that the following standards govern the determination of whether a patent claim is "anticipated" by the prior art. I have applied these standards in my evaluation of whether the claims of the patent are anticipated by the prior art.

29. I understand that, for a patent claim to be "anticipated" by the prior art, each and every claim limitation must be found, expressly or inherently, in a single prior art reference as recited in the claim. I understand that claim limitations that are not expressly described in a prior art reference are still deemed to be present if they are "inherent" to the thing or process being described in the prior art.

30. I understand that if a reference incorporates other documents by reference, the incorporating reference and the incorporated reference(s) should be treated as a single prior art reference for purposes of analyzing anticipation. I understand that it is acceptable to consider evidence other than the information in a

particular prior art document to determine if a feature is necessarily present in or inherently described by that reference.

B. Obviousness

31. I understand that a claimed invention is not patentable if it would have been obvious to a person of ordinary skill in the field of the invention at the time the invention was made.

32. I understand that a patent may not be obtained even though the invention is not identically disclosed or anticipated by the prior art, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. I also understand that patentability shall not be negated by the manner in which the invention was made.

33. I understand that the following standards govern the determination of whether a claim in a patent is obvious. I have applied these standards in my evaluation of whether the asserted claims of the patent-at-issue here would have been considered obvious as of the relevant priority date. I have been asked to assume that the priority date for the claims at issue here is September 23, 2005.

34. I understand that to find a claim in a patent obvious, one must consider certain factual questions regarding the claimed invention and the prior art.

Specifically, I understand that the question of whether an alleged invention is obvious requires consideration of four factors:

- The scope and content of the prior art;
- The differences between the prior art and the claims at issue;
- The knowledge of a person of ordinary skill in the pertinent art; and
- Whatever objective factors indicating obviousness or non-obviousness may be present in any particular case.

35. In addition, I understand that the obviousness inquiry should not be done based on hindsight but must be done using the perspective of a person of ordinary skill in the relevant art as of the priority date of the patent claim.

36. I understand that the objective factors indicating obviousness or non-obviousness may include: commercial success of products covered by the patent claims; a long-felt need for the invention; failed attempts by others to make the invention; copying of the invention by others in the field; unexpected results achieved by the invention; praise of the invention by those in the field; the taking of licenses under the patent by others; expressions of surprise by experts and those skilled in the art at the making of the invention; and the patentee proceeded contrary to the accepted wisdom of the prior art. I also understand that any of this evidence must be specifically connected to the invention rather than being associated with the prior art or with marketing or other efforts to promote an invention.

37. I understand that the obviousness analysis can consider the predictability or unpredictability of changes to the prior art and the amount of experimentation needed to make variations to the prior art. I also understand that, in order to make an obviousness determination, sometimes it will be necessary to look to interrelated teachings of multiple patents; the effects of demands known to the design community or present in the marketplace; and the background knowledge possessed by a person having ordinary skill in the art. I understand that all these issues may be considered to determine whether there was an apparent reason to combine the known elements in the fashion claimed by the patent at issue.

38. I understand that the obviousness analysis cannot be confined by a formalistic conception of the words “teaching, suggestion, and motivation.” It is my understanding that any motivation that would have been known to a POSITA, including common sense and the nature of the problem to be solved, may be considered.

39. I understand that an invention that might be considered an obvious variation or modification of the prior art may be considered non-obvious if one or more prior art references discourages or leads away from the line of inquiry disclosed in the reference(s). A reference does not “teach away” from an invention simply because the reference suggests that another embodiment of the invention is better or preferred. My understanding of the doctrine of teaching away requires a clear

indication that the combination should not be attempted (*e.g.*, because it would not work or due to explicit statements saying the combination should not be made).

IV. Technical Background

A. The '812 Patent's Background

40. The '812 patent is entitled "Pointing and Identification Device" (Ex-1001, Cover Page). The '812 patent describes its "Field of the Invention" as "directed to a pointing device and more particularly to a pointing device which can allow a user to point to a spot and determine that spot's absolute location, either on a TV screen, a computer screen, or in the real world" (Ex-1001, 1:5-8).

41. The '812 patent notes that a computer mouse is well-known (Ex-1001, 1:11-15).

42. The '812 patent describes that a computer mouse detects relative motion and through a central processing unit imparts the mouse's relative motion to a display screen (Ex-1001, 1:15-24).

43. The '812 patent describes that "low-cost, low power lasers" have pointing applications (Ex-1001, 1:31-32). The '812 patent states that laser pointing applications include "a presentation controller with cursor control and a laser pointer, laser tag for game playing, and military applications such as guns with laser dots and laser-guided bombs" (Ex-1001, 1:32-36).

44. The '812 patent provides examples of pointing lasers that were commercially available such as the “laser pointing” device by Omegascope® and a LaserMouse from Penta Performance of Albuquerque, New Mexico (Ex-1001, 1:37-55).

45. The '812 patent also describes commercial products such as the Nintendo “Revolution” pointing device for a video game, the PaperClick® product by NeoMedia which reads and decodes UPC/EAN or other bar codes to link users to the internet, and the PaperClick® Mobile Go-Window™, which provides a horizontal bar on the screen of a wireless device (Ex-1001, 1:56-2:4).

46. The '812 patent also recognizes that the prior technologies “permitted users to use cameras in their mobile phones to take pictures of offline media such as magazines, posters, product packaging, and logos” (Ex-1001, 2:21-24).

47. The '812 patent also recognized that prior technologies included object detection or “visual matching technology” to identify pictures and connect the image to relevant information (Ex-1001, 1:2:21-25).

48. The '812 patent recognizes that “reticles” were commonly used (Ex-1001, 2:26-28). For example, the '812 patent describes that cameras have been combined with reticles for such purposes as gun scopes and periscopes (Ex-1001, 2:26-28).

49. According to the '812 patent, “there is still no solution in the art to provide a pointer for pointing directly at, clicking-on, and identifying a distant absolute location and for doing so on a TV screen, on a computer screen, or in the real world” (Ex-1001, 2:29-32).

B. Cell Phones with Digital Cameras

50. The '812 patent did not invent cell phones with a digital camera and wireless transmission capability for transmitting digital images.

51. U.S. Patent No. 6,651,053 B1 (Ex-1017) referenced in the '812 patent describes mobile cellular phones that had internet capabilities such as “Internet Portable Digital Assistants (PDA’s),” including the Palm VII™ wireless PDA (Ex-1017, 1:49-64). An example of the Palm III-V™ PDA is shown below (https://en.wikipedia.org/wiki/Palm_VII#/media/File:Palm7-dock.jpg). But the Plam VII™ did not have a camera.



52. As the '812 patent admits mobile cell phones were well-known to a POSITA (Ex-1001, 2:21:25). The paper entitled “IDEixis-Searching the Web with Mobile Images for Location-Based Information” from Tollmar (2004) also recognized that “camera-equipped mobile devices were becoming commonplace” (Ex-1009, p. 2).

53. The '812 patent admits that e-commerce has been enabled on a camera cell phone, such as the NOKIA 3650 model (Ex-1001, 1:62-67). Further, Tollmar references using the NOKIA 3650 phone and taking “advantage of its built in camera” (Ex-1009, p. 5). Tollmar explains that the NOKIA 3650 phone has a “built-in camera” having a resolution of “640x480” and having “support for Multimedia

Messaging Service (MMS), using C++ on Symbian OS” (Ex-1009, p. 4). The NOKIA 3650 phone is shown below (https://www.gsmarena.com/nokia_3650-345.php) (Ex-1015; *see also* Ex. 1016).

NOKIA 3650 Mobile Phone



54. The NOKIA 3650 specifications included having a single VGA camera and could take video (https://www.gsmarena.com/nokia_3650-345.php) (Ex-1015; *see also* Ex-1016).

55. One of the first cell phones having a digital camera was released in 2000 by Samsung and was known as the Samsung SCH-V200. The Samsung SCH-V200, had a digital camera capable of taking 20 photos at 350,000 pixel resolution, which is 0.35 megapixels (<https://www.digitaltrends.com/mobile/camera-phone-history>)

(Ex-1011). But the SCH-V200 could only transmit the digital images by wired transmission.

56. Another early cell phone having a digital camera was the Sharp J-phone which was released in 2000 (Ex-1011). The Sharp J-Phone could take pictures at 110,000-pixel resolution or 0.11-megapixels (Ex-1011).

Sharp J-Phone (Ex-1011)



57. The Sanyo SCP-5300 cell phone having a camera was released in 2002 and had a 0.3 megapixel capability (Ex-1011).

58. In July 2004, the Audiovox PM8920 was available. The Audiovox PM8920 had a digital camera and could send the digital images wirelessly (<https://www.digitaltrends.com/mobile/camera-phone-history/>) (Ex-1011).

Audiovox PM8920



© 2004 CNET Networks, Inc.

C. Cameras and Reticles

59. The '812 patent recognizes that “reticles” were commonly used (Ex-1001, 2:26-28). The '812 patent did not invent “reticles.” For example, the '812 patent describes that cameras have been combined with reticles for such purposes as gun scopes and periscopes (Ex-1001, 2:26-28).

60. I agree that reticles were well-known to a POSITA in 2004-2005.

61. Reticles were used in digital cameras. For example, the article “Understanding the Role of Image Recognition in Mobile Tour Guides” (“Davies”) describes the HP Photosmart Mobile Camera (Ex-1010, p. 193). Davies describes using “a PDA (HP iPAQ 5550) with built-in 802.11 networking,” and “an attached mobile camera (HP Photosmart Mobile Camera)” (Ex-1010, p. 193). Davies describes that the “HP Photosmart Mobile Camera” has a 320×240 resolution and

an expansion-pack with a 4.66 GB PC card hard-drive to store the application content, meaning the camera was digital (Ex-1010, p. 193).

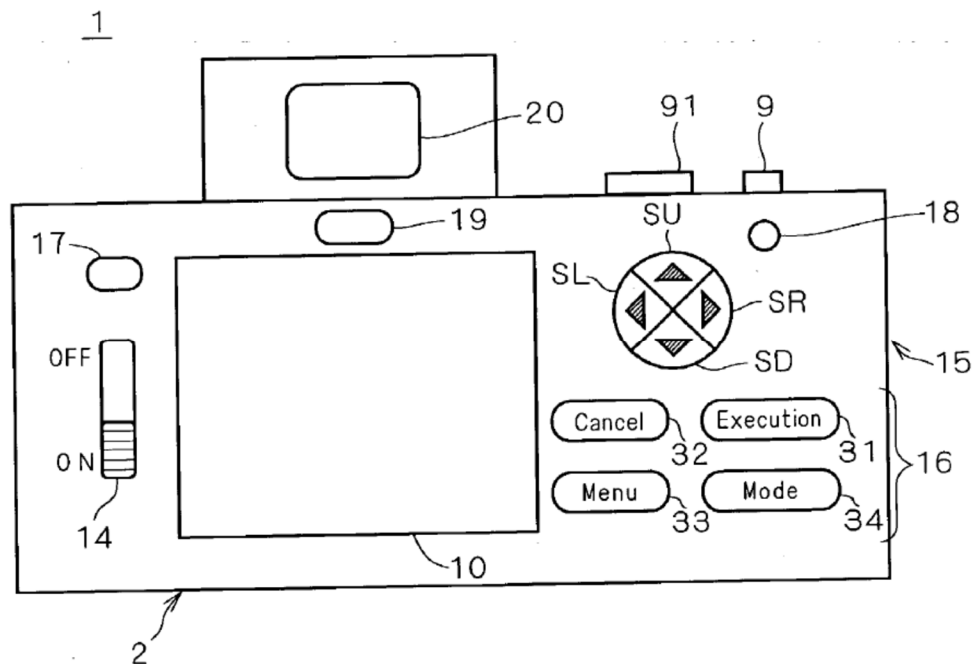
62. The HP Photosmart Mobile Camera, referenced in Davies, is shown in a video uploaded to YouTube (<https://www.youtube.com/watch?v=EfoPRZURqHk>). As shown from the image below, the HP Photosmart Mobile Camera had a reticle that is shown by the four green corners in the camera's field of view:

HP Photosmart Mobile Camera



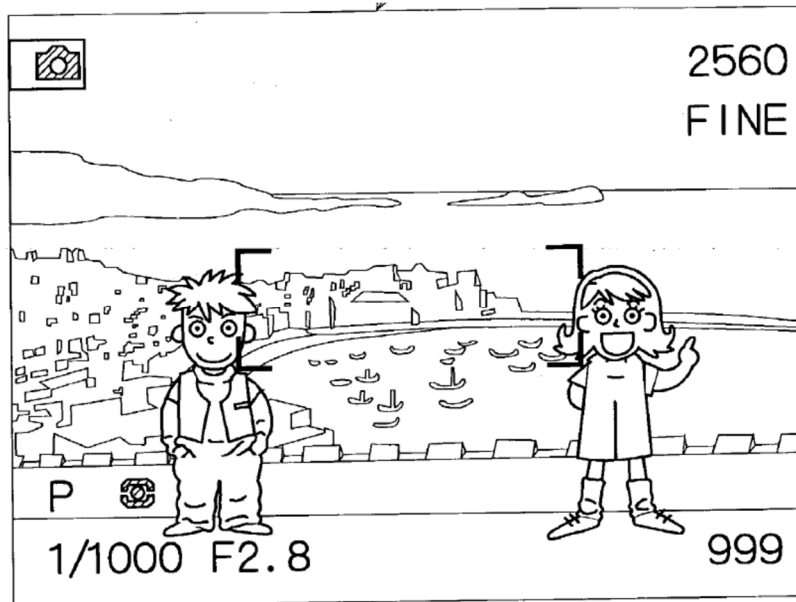
63. A patent describing a reticle in a digital camera is U.S. Patent Application Publication No. 2004/0125220 (“Fukuda”), which published on July 1, 2004 (Ex-1006). Fukuda discloses a digital camera in Figure 1 shown below (Ex-1006, ¶¶ [0020]-[0021]).

Fukuda Figure 1



64. Fukuda discloses an “optical viewfinder” or “reticle” in Figure 7 denoted by four corners (Ex-1006, ¶ [0067])

Fukuda Figure 7



65. Fukuda describes a “digital camera” that “has a display such as an LCD” and an “electronic viewfinder” or “EVF” (Ex-1006, Abstract; ¶ [0028]). Fukuda explains that the digital camera may include a “CCD (Charge Coupled Device)” that “functions as an image capturing part for capturing a viewfinder image (preview image) of the subject before photographing” (Ex-1006, ¶ [0024]).

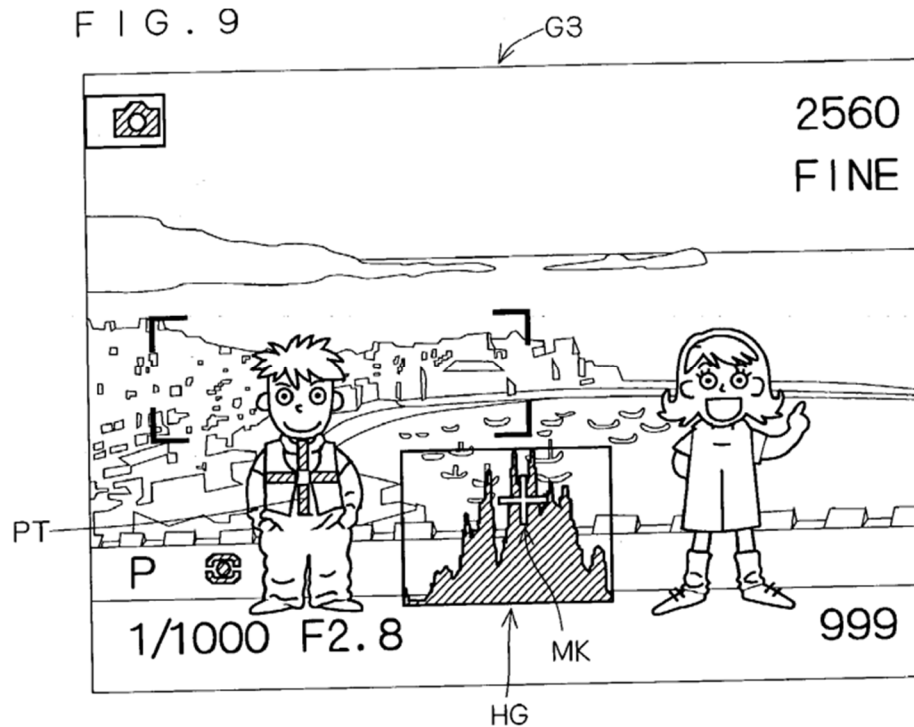
66. Fukuda teaches that a “viewfinder image,” which is captured by the CCD and is shown in FIG. 7 (reproduced below) “is displayed as a live view on the LCD 10,” but that when a user’s eye approaches a “sensor 19” of the digital camera, “the LCD 10 is turned off, the EVF 20 is turned on, and the viewfinder image G1 is displayed on the EVF 20” (Ex-1006, Abstract; ¶¶ [0024], [0061], FIG. 7). Therefore,

Fukuda teaches that the viewfinder image G1 can be presented in either the LCD or the EVF before a picture is taken using the digital camera.

67. In Figure 7, Fukuda illustrates a reticle denoted by four corners in the viewfinder image. Fukuda discloses that a user can move the reticle to adjust the camera's point of focus. Fukuda describes that "a pointer" may be displayed in the viewfinder display and that the user may "move the pointer PT up, down, right and left in the screen" to set a designation position in the screen (Ex-1006, ¶ [0070]).

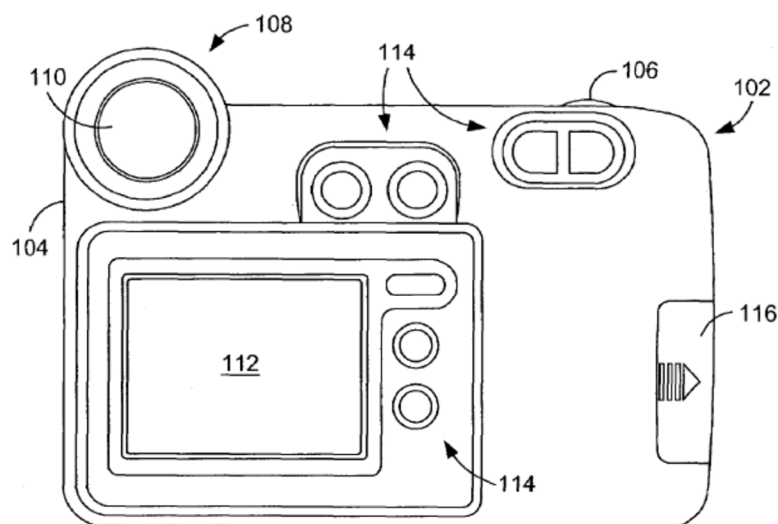
68. Fukuda illustrates in FIG. 9 (reproduced below) that "a focus point may be associated with the pointer PT" (Ex-1006, ¶ [0071], FIG. 9). Fukuda describes that the "in-focus area" is defined by "a rectangular area which is set around the pointer PT" and the user can move the rectangular area "interlockingly with the pointer" to allow for "focusing control" (Ex-1006, ¶ [0071], FIG. 9). Therefore, Fukuda discloses that a user may move a pointer around a viewfinder display in order to move or adjust a reticle, defined by the "rectangular area," to adjust the camera's point of focus.

Fukuda Figure 9



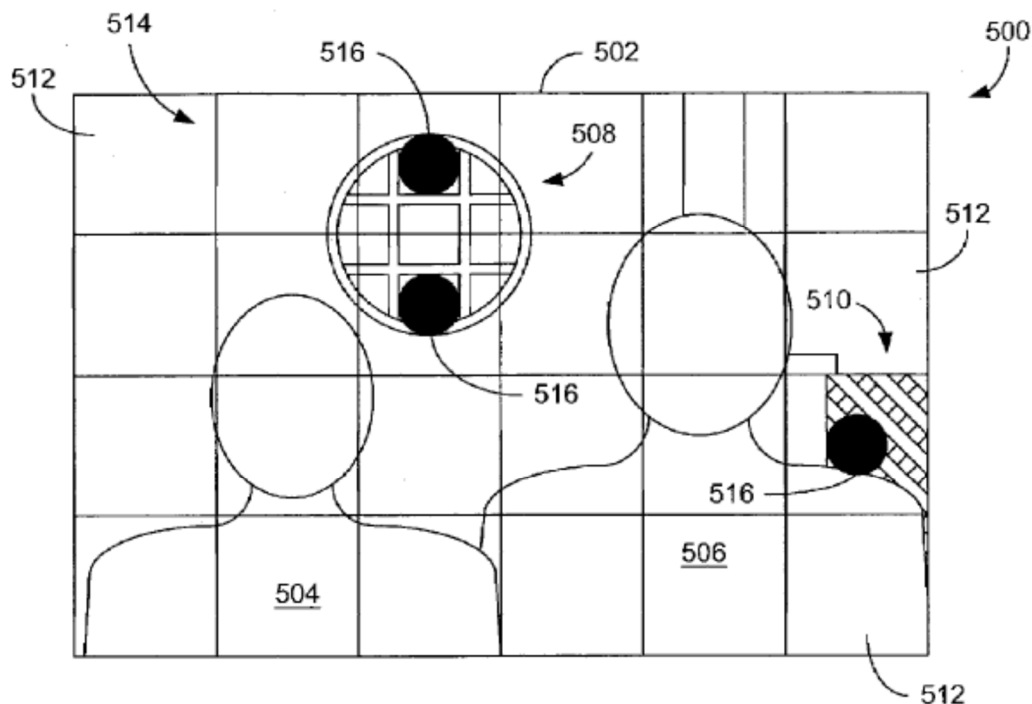
69. U.S. Patent No. 7,248,301 ("Voss") also discloses a digital camera having a reticle (Ex-1007). Voss discloses a digital camera (Ex-1007, 2:55-3:44).

Voss Figure 1



70. Voss discloses “graphical indicia” or reticles “for display to the user that identify the portions of the image having the greatest focus” (Ex-1007, 5:31-35). Voss describes that the “indicia can take several different forms” (Ex-1007, 5:35). Voss provides reticle examples of an array grid 514 or dots 516 as shown in Figure 5C.

Voss Figure 5C



D. Object Detection

71. A POSITA was familiar with many object detection techniques for detecting objects in a digital image in 2004. Object detection is a task or application within the broader fields of computer vision and image processing that can use pattern recognition techniques to not only recognize but also to precisely locate and

classify objects within images and videos. For example, U.S. Patent Application No. 10/492,243 (“Boncyk”) published as U.S. Pub. No. 2004/0208372 on October 21, 2004. Boncyk describes many types of object detection (Ex-1005).

72. Boncyk discloses “image type determination 26” (Ex-1005, ¶ [0047]). Boncyk discloses shape comparison, grayscale comparison, wavelet comparison, color cube comparison, and calculating a combined match score that are used to detect objects in a digital image (Ex-1005, ¶ [0070]). These are examples of computer vision technologies that were used for object detection prior to September 2005.

73. The ’812 patent did not invent object detection. Rather, the ’812 patent describe that a system can provide a “list of likely pointed-to objects” but does not provide details of any object detection (Ex-1001, 6:30-33). The ’812 patent describes capturing an image of a television screen, determining the frame of the television show, and determining objects from knowing the frame number (Ex-1001, 6:14-21; 10:57-11:54).

74. Computer vision is a field that enables computers to “see” and interpret images and videos, much like humans do, by using algorithms and machine learning techniques to identify objects, patterns, and scenes. Computer vision aims to replicate human vision by equipping computers with the ability to analyze visual

data, such as images and videos, and derive meaningful information from them.

Computer vision started in the 1960s and evolved significantly since the 1970s .

75. Pattern recognition, especially in the context of identifying objects in images, leverages Digital Signal Processing (DSP) techniques to extract meaningful features and classify them, enabling tasks like object detection and image segmentation. Pattern recognition in image processing involves identifying and classifying visual patterns, objects, or events within digital images and videos. It uses DSP techniques for image enhancement, edge detection, segmentation, relevant parameter extraction, and transformations to domains that are more relevant and efficient to the pattern recognition's objective. Object detection, which is essentially a branch in pattern recognition, is aimed for identifying and locating specific objects within an image or video.

Early Approaches for object detection (1960s-2000s) included Template Matching. Template Matching involved comparing new images to predefined templates or shapes to detect objects.

76. Other object detection techniques are known as feature-based methods. Feature-based methods identified unique features within an object (*e.g.*, edges or corners) that were invariant to scale and rotation.

77. In the late 1990s and early 2000s, Haar Cascade emerged as a popular method, particularly for face detection, using a sliding window approach to scan for

specific features (see also Paul Viola and Michael J. Jones. Robust Real-Time Face Detection. *International Journal of Computer Vision*, 57(2):137–154, 2004; and Rainer Lienhart and Jochen Maydt. An Extended Set of Haar-like Features for Rapid Object Detection. In *Proc. Int. Conf. Image Processing*. 2002, volume 1, pages 1-900. IEEE, 2002).

78. The object detection algorithm known as the Viola-Jones Algorithm was developed and formulated in 1993 by Paul Viola and Michael Jones, and used statistical classification known as Adaptive Boosting (AdaBoost) learning, Haar-like features, and a cascade of classifiers to detect objects, particularly faces, in real-time (Viola, P.; Jones, M., Rapid Object Detection Using a Boosted Cascade of Simple Features (2001). *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. CVPR 2001. Vol. 1. IEEE Comput. Soc. doi:10.1109/cvpr.2001.990517. ISBN 0-7695-1272-0. S2CID 2715202; Viola, Paul and Jones, Michael J., 57(2) *International Journal of Computer Vision*, Robust Real-Time Face Detection pp. 137–154 (May 2004).

79. Another known object-detection technique was known as the Histogram of Oriented Gradients (HOG), which analyzed light and shadow changes in image sections, while Support Vector Machines (SVMs) classified these features to identify objects (see US Patent No. 4,5676,10A).

80. Another object detection technique was known as SIFT (Scale-Invariant Feature Transform). SIFT identified key points (corners and edges) in images and generated local descriptors, allowing for object recognition across different scales and rotations (Lowe, David G., Proceedings of the International Conference on Computer Vision Object recognition from local scale-invariant features., pp. 1150-1157 (1999). Vol. 2).

V. The '812 Patent

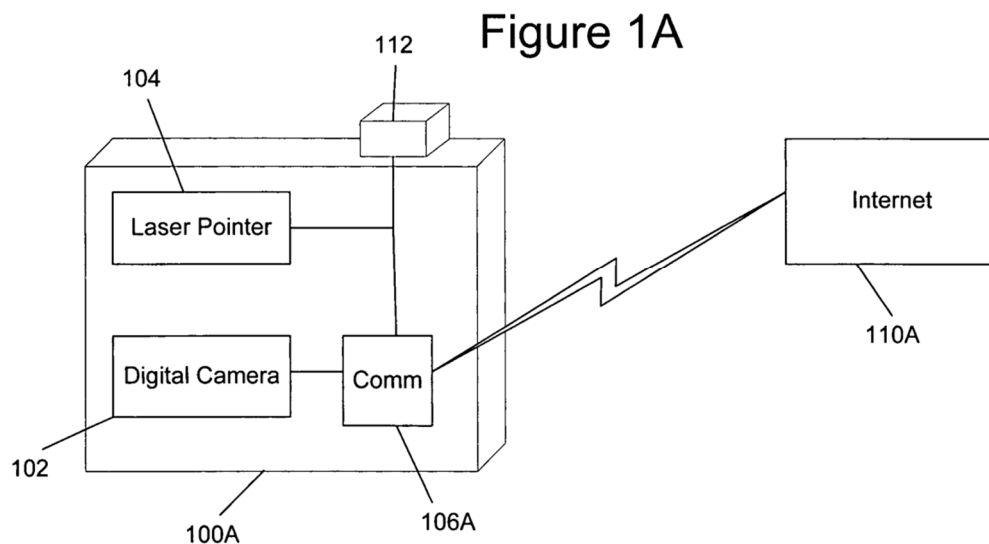
A. The '812 Specification

81. The '812 patent claims the benefit of the filing date of United States patent application serial no. 11/233,043, filed on September 23, 2005 (Ex-1001, Cover Page (21)). I applied September 23, 2005 as the effective filing date.

82. The '812 patent describes that a “mouse” used with a computer is “well known,” but has some disadvantages (Ex-1001, 1:120-30). Some alleged drawbacks are that a computer mouse detects “relative” motion, not “absolute motion,” because the mouse does not know its own location (Ex-1001,1:15-24). Another alleged drawback is that the mouse cannot point to a “real-world” object, such as a computer or TV screen (Ex-1001, 1:24-30). The '812 patent states that the problem to be solved was that there was “no solution in the art to provide a pointer for pointing directly at, clicking-on, and identifying a distant absolute location and for doing so on a TV screen, on a computer screen, or in the real world” (Ex-1001, 2:29-32).

83. The invention as described in Figure 1A includes a pointing and identification device (“PID”) 100A (Ex-1001, 5:8-10). The PID 100A “includes a digital camera” (*e.g.*, digital video camera) 102 for taking real-time images, a laser pointer 104, and a wireless communication component 106A (Ex-1001, 5:10-18). The PID has “at least one mouse button 112” (Ex-1001, 5:22-23).

’812 Patent Fig. 2



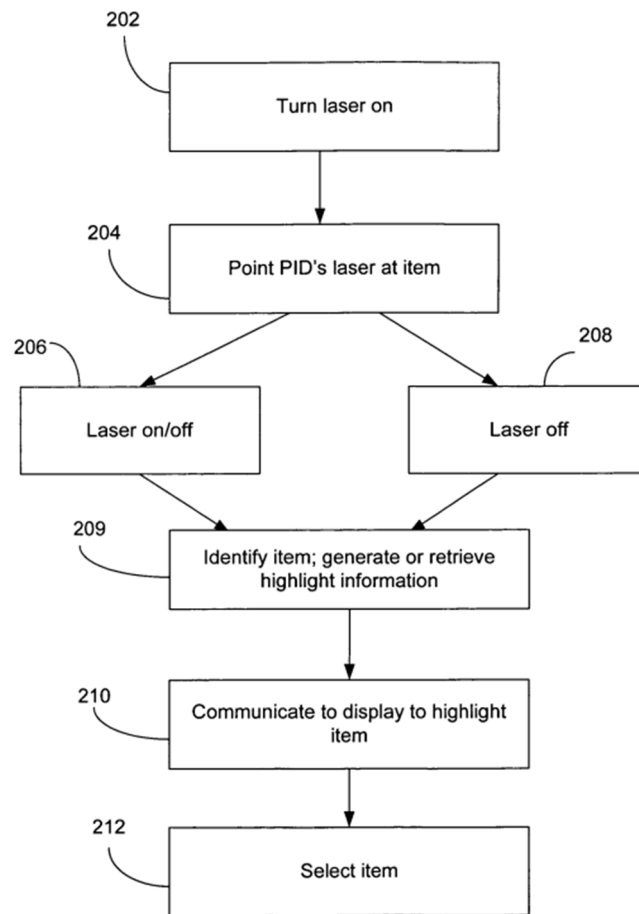
84. The PID 100A is for pointing to either (i) “a Frame on a TV screen and a location on that Frame;” or (ii) “a view of” a “real world and a location in that” real world view (Ex-1001, 5:10-13). The digital camera 102 is for “taking real-time images” (Ex-1001, 5:13-14). The laser pointer 104 is for “projecting a dot onto the object to be pointed to,” and the wireless communication component 106A is for “communicating with a computer or Internet 110A” (Ex-1001, 5:16-18). The wireless communication can be “any suitable standard for digital cellular telephony”

and the wireless communication component 106A “can implement a complete cellular telephone” (Ex-1001, 5:24-30). The ’812 patent does not describe a new laser, camera, or communication device. Rather, the ’812 patent describes known components (*e.g.*, a laser pointer, a digital camera, and a communication device) in one PID 100A.

85. As the ’812 patent describes, the PID 100A can point at a television or a real-world object and take a picture of a television frame or a real-world object (Ex-1001, 5:8-23). The PID 100A can interface with the internet 110A or a remote computer 110B (Ex-1001, 5:10-23). The remote system can run software for recognizing a television frame captured by the PID 100A and determine the frame pointed to by knowing the television frames associated with the frame (Ex-1001, 5:17-20). The remote system software can also recognize an object in a view of the real world and/or determine the location pointed to on or near the recognized object (Ex-1001, 5:20-23).

86. A method of using the PID 100A is described with reference to Figure 2 (Ex-1001, 5:61-63).

'812 Patent Fig. 2

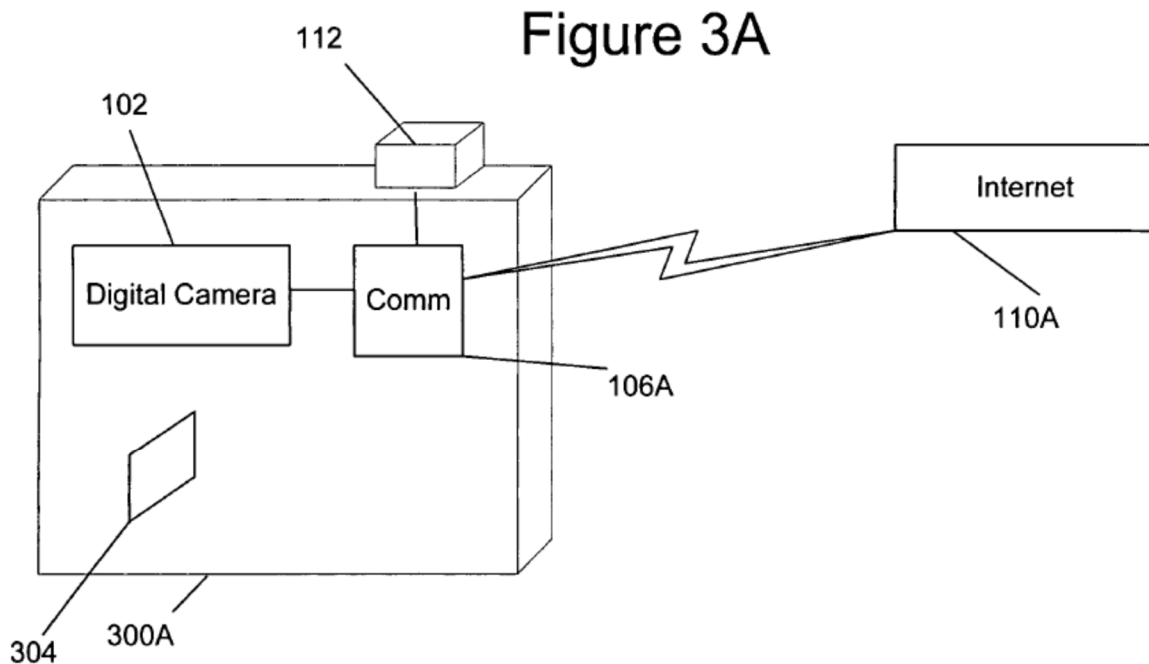


87. In step 202, the laser is turned on by for example holding down on the button for the laser (Ex-1001, 5:64-67). The laser is pointed at an item in step 204 (Ex-1001, 6:1-3). An image is taken in either step 206 or 208. At step 209, an item is identified in the image and information for highlighting the item is retrieved (Ex-1001, 6:22-24). At step 210, the PID communicates with the display to highlight the item, and at step 212, the user can select the item (Ex-1001, 6:22-30).

88. The '812 patent provides a second embodiment shown in Figure 3A. This embodiment uses a camera with a reticle-aimed camera mouse, as shown in

Figure 3A below (Ex-1001, 9:1-3). The reticle is denoted as 304 in Figure 3A (Ex-1001, 9:4-6).

'812 Patent Fig. 3A

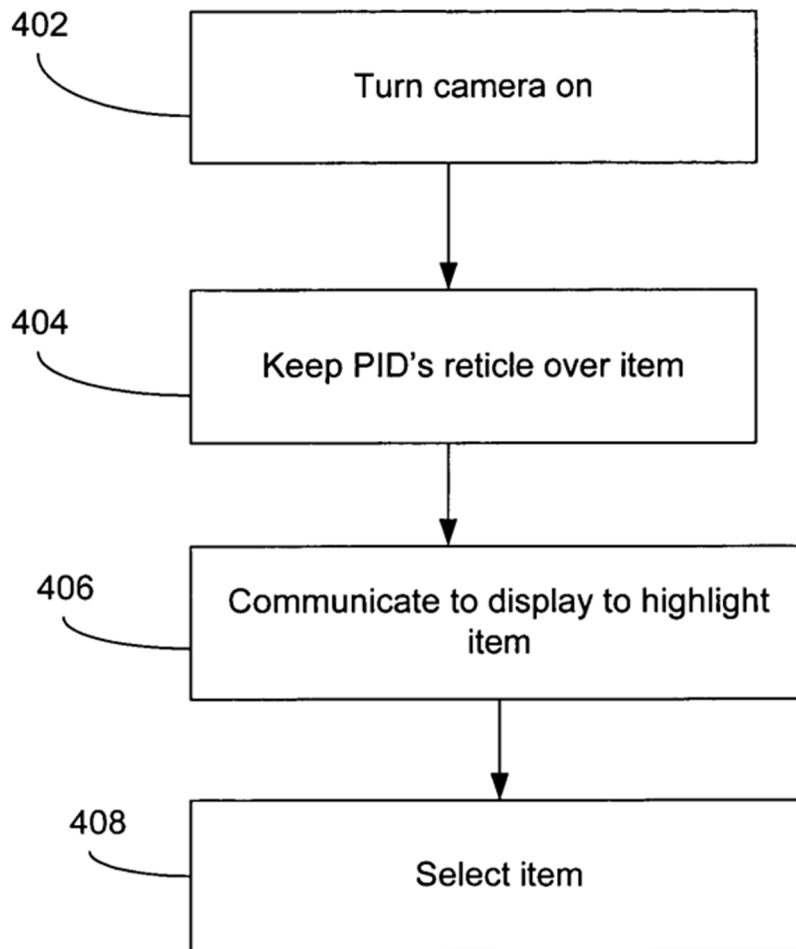


89. The '812 patent describes that a reticle “refers to any means that the user sees in a viewfinder to designate a spot in the view or a subarea of the view” (Ex-1001, 9:12-15). As the '812 patent describes, the reticle may be fixed or mobile with respect to the camera image (Ex-1001, 9:20-26). The '812 patent explains that a “fixed reticle (FiRe) is useful when the intention is to move about the pointing of the camera” and is “typically in the center of the view” (Ex-1001, 9:20-23). The '812 patent also discusses that a “mobile reticle (MoRe) is most useful when it is

most convenient to keep the camera fixed” (Ex-1001, 9:23-25). According to the ’812 patent, the reticle may be switchably fixed or mobile (Ex-1001, 9:25-26).

90. The described method of using the second embodiment is set forth in Figure 4. The method is fairly simple. The camera is turned on at 402. The user keeps the reticle over an item (404) (Ex-1001, 9:11-12). At 406, the pointing and identification device optionally communicates to the display to highlight an item (Ex-1001, 9:16-17). The user clicks to select an item at 408 (Ex-1001, 9:17-19).

’812 Patent Fig. 4



B. The '812 Patent Claims

91. While the specification describes problems with computer mice in terms of relative and absolute mice location and pointing directly at, clicking-on, and identifying a distant absolute location or object, none of that is set forth in the claim. Claim 1 is a method claim involving four steps.

92. Claim 1's first step is "providing" a PID comprising "at least one actuation means," "a digital camera," and a "communication device" (Ex-1001, 49:9-21). The second step is "communicating" a digital image to a "different location" (Ex-1001, 49:22-23). The third step is "automatically identifying" a list of likely pointed-to-objects from the digital image and occurs at "the different location" (Ex-1001, 49:24-26). The fourth step is "returning" the list to the user (Ex-1001, 49:27-28).

93. Claim 1's elements are referred to using the identifiers below.

Claim Identifier	Claim Element
Preamble	A method for identifying an object, the method comprising:
Providing a PID	(a) providing a pointing and identification device for pointing at the object, the pointing and identification device comprising:
Actuation Means	at least one actuation means for actuation by the user when the user points the pointing and identification device at the object;

Digital Camera	digital camera for forming a digital image of the object or of a portion of the object when the user points the pointing and identification device at the object and actuates the at least one actuation means; and
Communication Device	a communication device for communicating the digital image to a different location when the user actuates the at least one actuation means;
Communicating	(b) communicating the digital image to the different location;
Identifying	c) automatically identifying a list of likely pointed-to objects from the digital image at the different location to return the list of likely pointed-to objects; and
Returning	(d) returning the list of likely pointed-to objects to the user to select one of the likely pointed-to objects;
Object	wherein the object is at least one of a spot on a displayed image on a display, a subarea of a space on the displayed image on the display, one of a plurality of objects in the displayed image on the display, an object in space, or near an object in space, a subarea of a surface in space, or one of a plurality of objects in space.

C. The '812 Patent's Prosecution History

94. The '812 patent was filed as U.S. App. No. 11/233,043 on May 22, 2015 (Ex-1001 at Cover Page). The as-filed claims were narrowed during prosecution in order to achieve allowance (Ex-1002).

95. The USPTO rejected '043 application claims 100-105 as anticipated by Boncyk or as obvious over Boncyk in view of other art (Ex-1002, pp.44-45, 59-61, 82-90, 106-112) in four rejections. After an interview, the applicant amended application claim 100 ('812 patent claim 1) to recite the automatically identifying a “list of likely pointed-to objects” and “returning the list” (Ex-1002, p. 36). Following this amendment, the USPTO allowed the '812 patent claims (Ex-1002, p. 15).

VI. The Relevant Technical Field and the POSITA

96. A POSITA in the field of the '812 patent at the time of the effective filing date would have had: (1) a bachelor's degree in computing science, electrical engineering, or equivalent degree, and (2) at least 2 years of experience with cameras, processing camera images for object detection, laser pointing devices, computer mice, and/or wireless communication. More education can compensate for less experience, and more experience can be compensated for less education.

97. I believe such a person also would have been familiar with technologies such as mobile devices, wireless communications, laser pointers, mice, digital cameras, and object detection.

98. As explained above, I understand that the relevant timeframe is September 23, 2005.

99. As shown in my CV, I have extensive experience in the relevant technical field. Based on my experience, I have an understanding of the relevant field in the relevant timeframe.

100. I consider myself to be a person of at least ordinary skill in the art for the '812 patent. As of 2005, I had a Bachelors, Masters, and Ph.D degrees in electrical and computer engineering and at years of experience in digital signal processing that includes cell phone technology, wireless communication technology, location technology, computer mice, and object detection in digital images.

VII. Claim Construction

101. I understand that, in this proceeding, the claims are construed according to their ordinary and customary meaning, in view of the specification and prosecution history, as understood by a POSITA as of the effective filing date of the underlying patent application.

102. I have been instructed that the claims of the '812 patent are to be given their ordinary and customary meaning as understood by one of ordinary skill in the art at the time of the invention in light of the specification and the prosecution history pertaining to the patent. I have followed these principles in my analysis.

103. I have been asked to apply the following construction of the claimed “actuation means.” The function of the claimed “actuation means” is: “for actuation by the user when the user points the pointing and identification device at the object,”

“forming a digital image of the object or of a portion of the object when the user points the pointing and identification device at the object and actuates the at least one actuation means;” and “communicating the digital image to a different location when the user actuates the at least one actuation means.” The corresponding structure for the claimed “actuation means” is a user interface such as a button(s) or a keyboard.

VIII. Overview of the Prior Art

A. Darrell

104. Darrell was filed on January 22, 2004, and published as U.S. Patent Pub. No. US 2005/0162523 on July 28, 2005 (Ex-1004 at Cover Page).

105. Darrell describes that the “present invention provides ... a novel point-by-photography paradigm” that is possible using, inter alia, “a new generation of camera-phones and content based image retrieval (CBIR) techniques applied to mobile imagery and the world wide web” (Ex-1004, ¶ [0020]).

106. Darrell describes a mobile device that “includes a camera to capture an image and a wireless handheld device, coupled to the camera and to a wireless network, to communicate the image with existing databases to find similar images” (Ex-1004 at Abstract). The Darrell mobile device “further includes a processor, coupled to the device, to process found database records related to similar images and a display to view found database records that include web pages including

images” (Ex-1004 at Abstract). Darrell touts that, with “such an arrangement, users can specify a location of interest by simply pointing a camera-equipped cellular phone at the object of interest and by searching an image database or relevant web resources, users can quickly identify good matches from several close ones to find an object of interest” (Ex-1004 at Abstract).

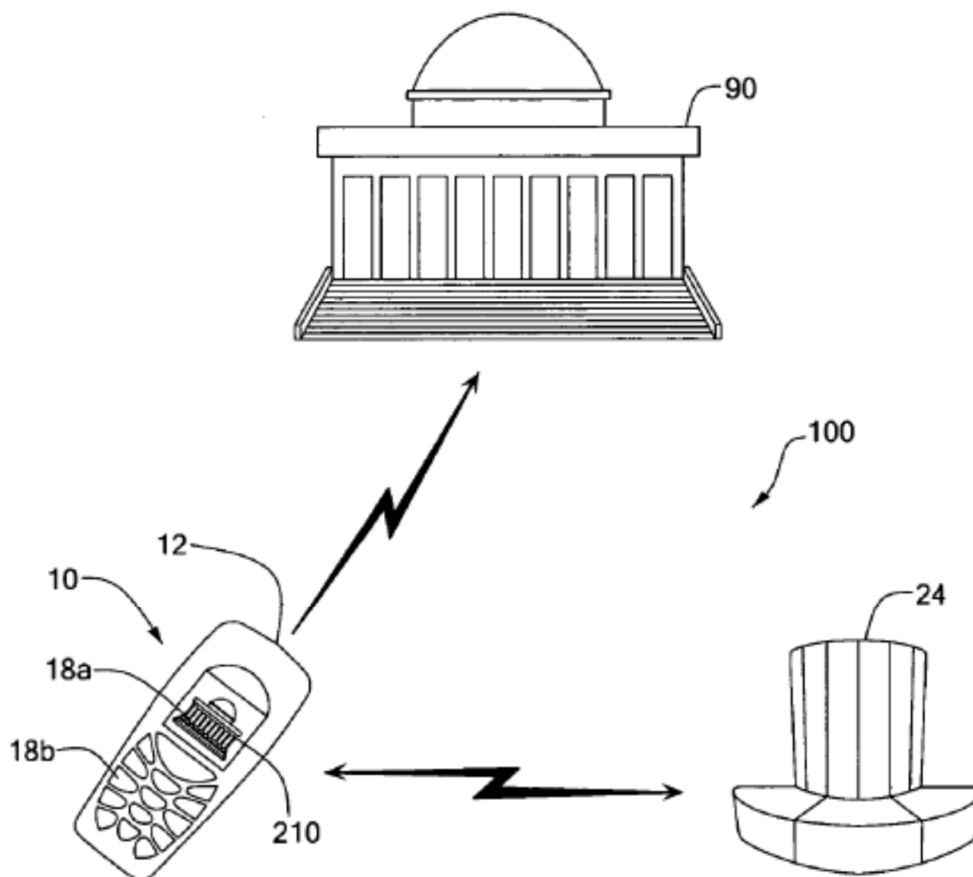
107. Darrell teaches that, “[u]sing the hand held device 10 with a camera 12, an image-based query can be formed simply by pointing with the camera 12 and snapping a photo” (Ex-1004 at 5:8-10).

108. Darrell also describes that, “[i]n operation, a mobile user can point the camera 12 to the view of interest, take photos, and send them wireless as queries (via multimedia SMS, a.k.a. MMS) to the web database 25” (Ex-1004 at 5:22-25). Darrell explains that the “handheld device 10 further includes a wireless communication device 14 coupled to the camera and to a wireless network 16 to communicate the image 210 with a computer 24 having a database 25 with computer files 26 to find similar images and a user interface 18, having here a display 18a and a keyboard 18b, coupled to the wireless communication device 14, to communicate to an user [*sic*] any results of found similar images” (Ex-1004, 4:6-16).

109. In Fig. 1, Darrell describes a handheld device 10 or mobile deixis device 10 having a “camera 12 to capture an image 210 of an object 90.” Darrell

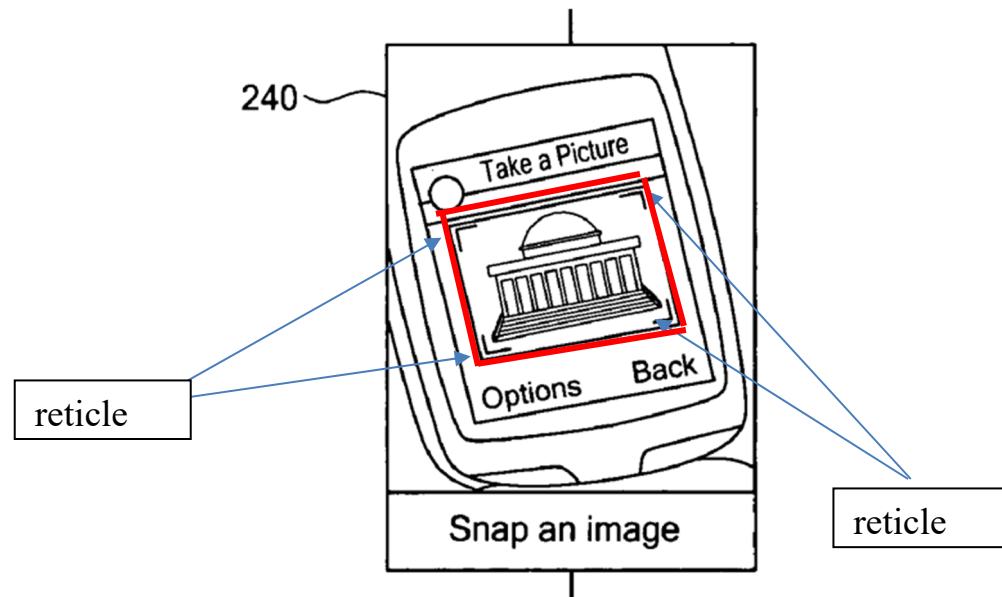
also refers to the handheld device as a “cellular phone” (Ex-1004, ¶ [0006], Abstract).

Darrell, Fig. 1



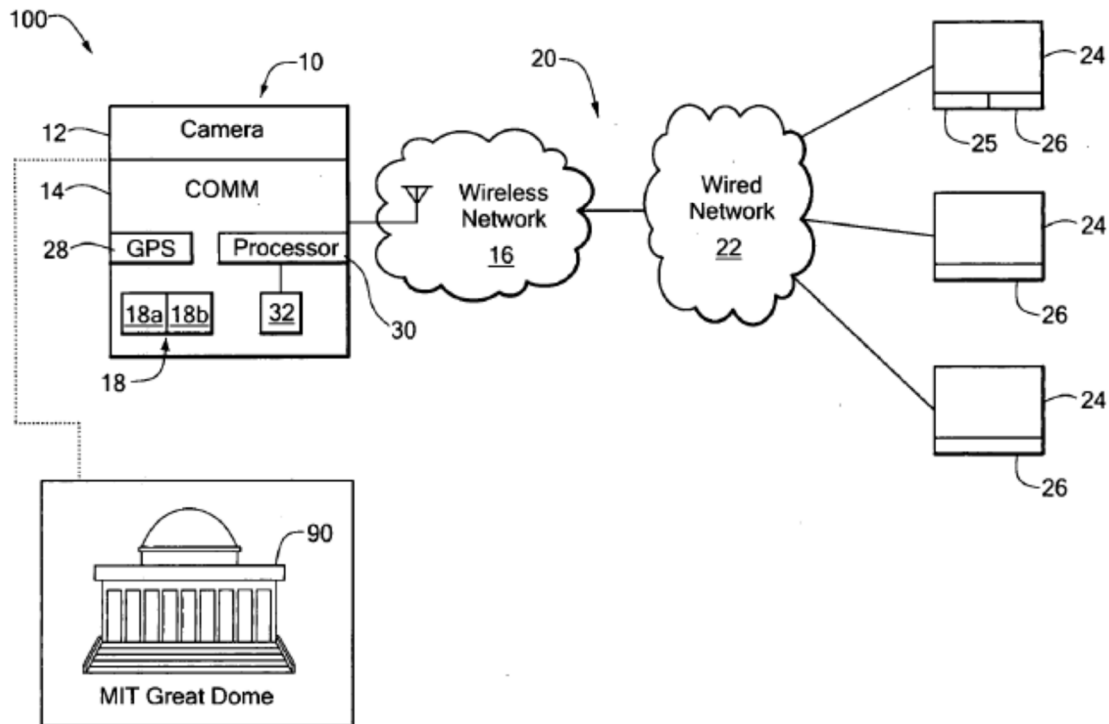
110. Darrell discloses that the handheld device 10 has a reticle as shown by the four outlining corners of the camera window in the “snap an image” portion of Fig. 3. As shown, Darrell’s reticle is inside of the camera viewfinder window (annotated in red) and surrounds the object – the building – that is in the camera viewfinder window.

Darrell, Fig. 3 Cropped, Annotated



111. Darrell describes (Fig. 1A) that the handheld device 10 includes a “wireless communication device 14 coupled to the camera and to a wireless network 16 to communicate the image 210” to a “computer 24 having a database 25 with computer files 26” (Ex-1004, ¶ [0021]). The database 25 includes features of the object in the image and a uniform resource locator (URL) (Ex-1004, ¶ [0022]).

Darrell, Fig. 1A



112. Darrell describes that “A mobile user can take photos of a place of interest” and that the “photos are sent to the image database 25 via a wireless link” (Ex-1004, ¶ [0021]). Darrell teaches that a “search engine looks for a set of images most similar to the query image” (*e.g.*, the photo sent from cell phone 10) (Ex-1004, ¶ [0025]). Darrell’s search “result” includes “a list of (candidate image, source URL) pairs” (Ex-1004, ¶ [0025]). Darrell describes that “a list of possible matching images” can be provided in response to the image captured by the camera (Ex-1004, claim 26).

113. Darrell teaches that the “mobile device 10 displays the result by arranging candidate images into a thumbnail mosaic 220” (Ex-1004, ¶ [0025]). Darrell further describes that the “user, as the final judge, can easily identify what sub-images are ‘really relevant’” (Ex-1004, ¶ [0025]). Darrell also describes that when “a thumbnail is selected, the source URL is retrieved and the content from that URL is shown on the mobile device 10” (Ex-1004, ¶ [0025]).

114. Darrell provides the example of a mobile user taking a picture of a building (*e.g.*, Killian Court), wirelessly communicating the image to server 24, the server searching for images of the building, and returning the most relevant web pages to the user (*e.g.*, URLs) (Ex-1004, ¶ [0027]).

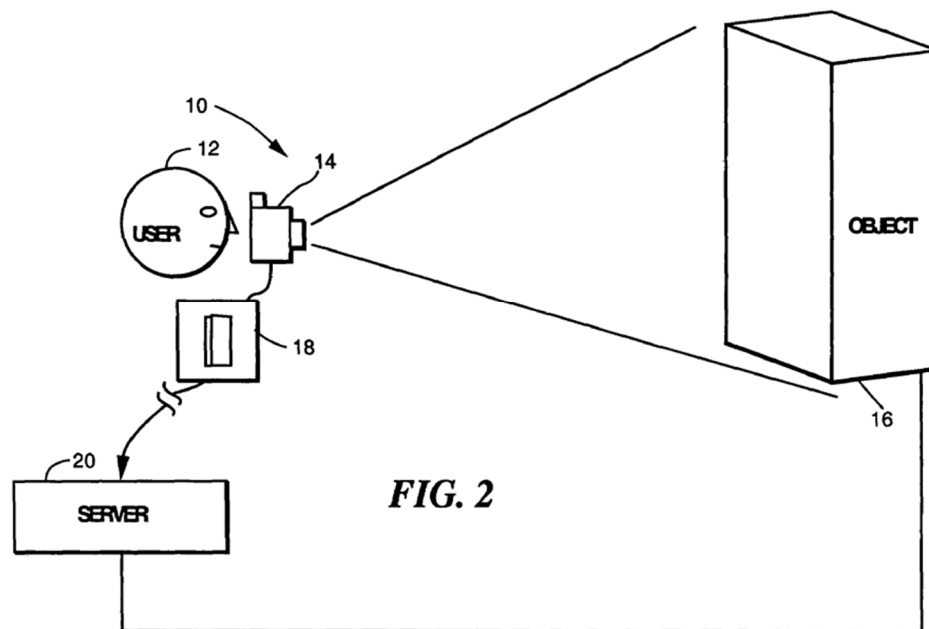
B. Boncyk

115. Boncyk was filed as U.S. Patent Application No. 10/492,243 on April 9, 2004, and published as U.S. Pub. No. 2004/0208372 on October 21, 2004 (Ex-1005 at Cover Page).

116. Boncyk describes, in its Abstract and Field of the Invention, “a method by which information and communication pertinent to an object is provided based on imagery of the object” where a “digital image of the object is captured and the object is recognized from a plurality of objects in a database” (Ex-1005 at Abstract; 1:7-10).

117. Boncyk discloses (Fig. 2) a system that includes “a computer, mobile telephone, personal digital assistant, or other similar device 14 equipped with an image sensor (such as a CCD or CMOS digital camera)” that “captures a digital image 18 of the scene at which it is pointed” (Ex-1005, ¶ [0045]). Boncyk also describes a “server 20” that may be “physically separate from the device 14” Ex-1005, ¶ [0046]). Boncyk describes that “user acquired images are transmitted from the device 14 to the Image Processor/Server 20 using a conventional digital network or wireless network means” (Ex-1005, ¶ [0046]).

Boncyk Figure 2

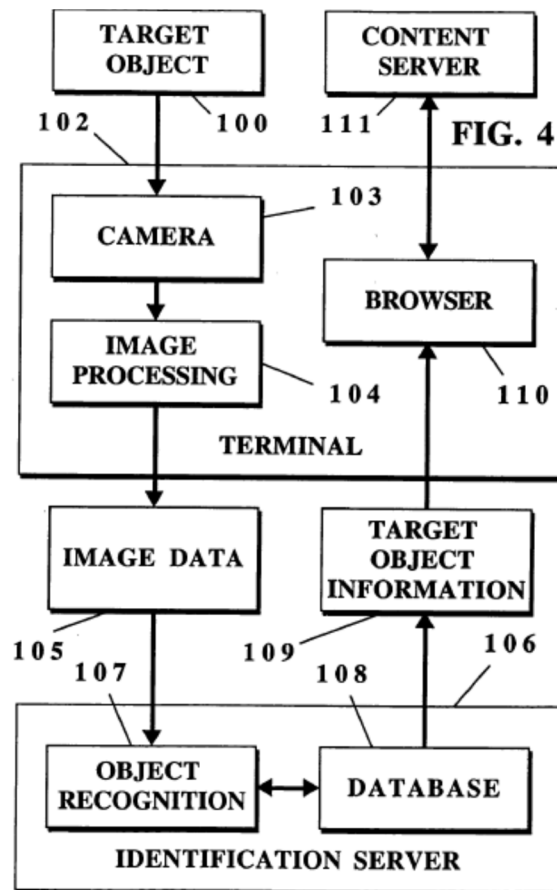


118. Boncyk also teaches that the server 20 may perform multiple “database comparison processes,” such as “Shape Comparison, Grayscale Comparison,

Wavelet Comparison, and Color Cube Comparison” to identify matches for the object that is in the digital image. (Ex-1005, ¶ [0070]).

119. Boncyk also discloses a “terminal 102” that “is a computing device that has an ‘image’ capture device such as digital camera 103, a video camera, or any other device that [c]an convert a physical object into a digital representation of the object” (Ex-1005, ¶ [0128], Fig. 4).

Boncyk Figure 4



120. As Boncyk explains, “[a]fter the camera 103 captures the digital imagery of the target object 100, image preprocessing 104 software converts the digital imagery into image data 105 for transmission to and analysis by an identification server 106” (Ex-1005, ¶ [0129]). Identification server 106’s overall function is “to determine and provide the target object information 109 corresponding to target object 100, based on image 105” (Ex-1005, ¶ [0143]). Boncyk describes an identification server 106 having object recognition 107 that “detects and recognizes images of the target object 100 or portions thereof” (Ex-1005, ¶ [00148]). The identification server 106 sends target object information 109 to terminal 102 (Ex-1005, ¶¶ [0147], [0171]). The terminal can use the target object information 109 to identify the target object (Ex-1005, ¶ [0171]).

121. Boncyk discloses that the user can “purchase” the “target object 100 via on-line transaction, with the purchase amount billed to an account linked to the terminal 102, to the individual user, to a bank account, or to a credit card” (Ex-1005, ¶ [0265]).

C. Fukuda

122. Fukuda was filed as U.S. Patent App. No. 10/421,791 on April 24, 2003, and published as U.S. Pub. No. 2004/0125220 on July 1, 2004 (Ex-1006 at Cover).

123. Fukuda describes a “digital camera” that “has a display such as an LCD” and an “electronic viewfinder” or “EVF” (Ex-1006, Abstract; ¶ [0028]). Fukuda explains that the digital camera may include a “CCD (Charge Coupled Device)” that “functions as an image capturing part for capturing a viewfinder image (preview image) of the subject before photographing” (Ex-1006, ¶ [0024]).

124. Fukuda teaches that a “viewfinder image,” which is captured by the CCD “is displayed as a live view on the LCD 10,” but that when a user’s eye approaches a “sensor 19” of the digital camera, “the LCD 10 is turned off, the EVF 20 is turned on, and the viewfinder image G1 is displayed on the EVF 20” (Ex-1006, Abstract; ¶¶ [0024], [0061], FIG. 7).

125. Fukuda describes that “a pointer” may be displayed in the viewfinder display and that the user may “move the pointer PT up, down, right and left in the screen” to set a designation position in the screen (Ex-1006, ¶ [0070]). Fukuda describes that “a focus point may be associated with the pointer PT” and that “the in-focus area” is defined by “a rectangular area which is set around the pointer PT” and that the rectangular area can move “interlockingly with the pointer” to allow for “focusing control” to be performed by the user (Ex-1006, ¶ [0071], FIG. 9).

IX. The Asserted Claims are Invalid

126. In my opinion, claims 1, 3, and 10-12 are anticipated by Darrell. It is also my opinion that claims 1-12 are obvious over Boncyk in view of Darrell. It is

further my opinion that dependent claims 5-9 are obvious over Boncyk in view of Darrell and the level of ordinary skill in the art. It is also my opinion that claims 1-12 are obvious over Boncyk, in view of Darrell, and further in view of Fukuda.

A. Anticipation by Darrell

1. Claim 1

127. Darrell describes each limitation in claim 1 and anticipates claim 1.

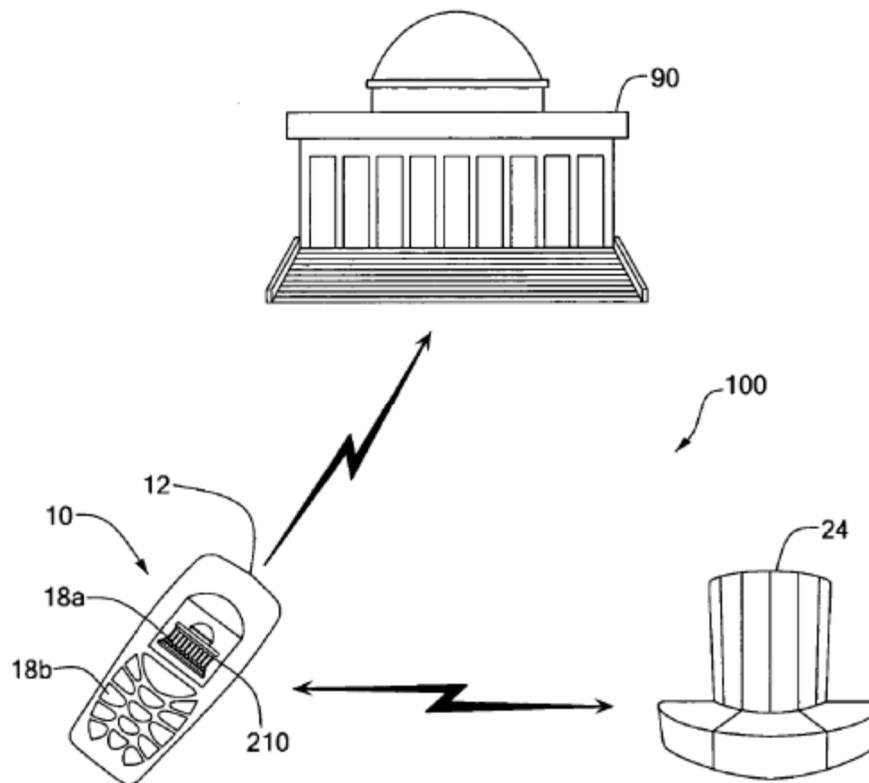
a. Preamble

128. The preamble is a “method for identifying an object.” Darrell discloses the “preamble” (Ex-1004, Abstract, Field of the Invention, ¶ [0020]). For example, Darrell describes that a “mobile device” or “hand held device 10” includes “a camera to capture an image” (Ex-1004, ¶ [0021], Abstract). Darrell teaches that the mobile device 10 is configured such that “users can specify a location of interest by simply pointing a camera-equipped cellular phone at the object of interest and by searching an image database or relevant web resources, users can quickly identify good matches from several close ones to find an object of interest” (Ex-1004, Abstract, ¶ [0024]). Darrell’s Field of the Invention section states that this “invention relates generally to location and object awareness systems and more particularly to systems and techniques to identify a location or object in a person's field of view” (Ex-1004, ¶ [0001]). Therefore, Darrell teaches methods “for identifying an object.”

b. Providing a PID

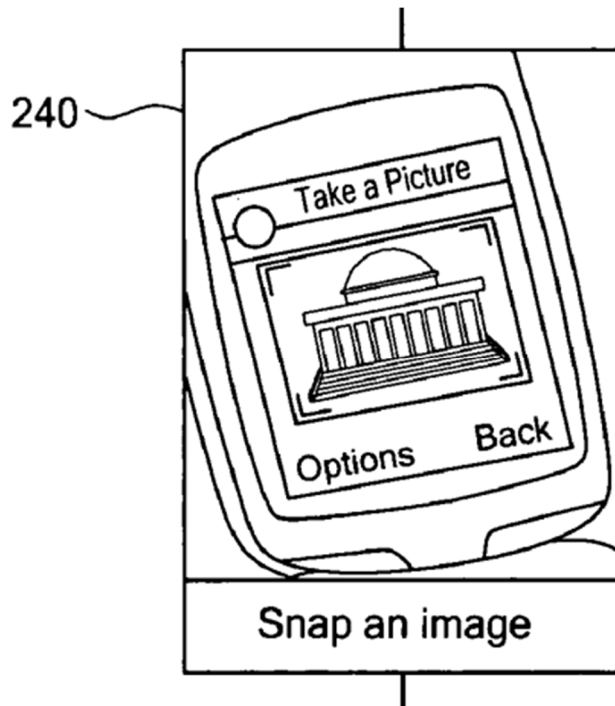
129. The “Providing a PID” limitation states “providing a pointing and identification device for pointing at an object.” Darrell discloses the “Providing a PID” limitation. Darrell discloses providing a “mobile device” or a “handheld device 10” that includes “a camera to capture an image” (Ex-1004, ¶ [0021]). Darrell’s mobile or handheld device constitutes the claimed “a pointing and identification device for pointing at an object.” In Darrell, the “mobile device” or “hand held device 10” (Fig. 1) operates as a “pointing and identification device” as recited by claim 1 of the ’812 patent because the “mobile device” or “hand held device 10” includes a camera that allows a user to specify and identify an object of interest “by simply pointing” the camera-equipped mobile phone or handheld device at the object of interest (Ex-1004, Abstract, ¶¶ [0006], [0023], [0024]). Darrell explains that the “present invention provides ... a novel point-by-photography paradigm” (Ex-1004, ¶ [0006]). Darrell has “a new pointing interface” that combines “a new generation of camera-phones and content based image retrieval (CBIR) techniques applied to mobile imagery and the world wide web” (Ex-1004, ¶ [0006]).

Darrell, Fig. 1



130. To the extent that Pointwise argues that the claimed “pointing and identification device” should be construed to include a reticle or laser, Darrell also shows that the display of the handheld device 10 includes a reticle in Figure 3. For example, the portion of Figure 3 of Darrell reproduced below illustrates, *inter alia*, a reticle as four corners in the camera viewfinder window of the handheld device 10. Darrell’s corner reticles identify or designate an area in the handheld camera’s field of view. The ‘812 patent describes a reticle as the “means that the user sees in a viewfinder to designate a spot in the view or a subarea of the view” (Ex-1001, 9:12-15). Darrell’s four corners constitute a reticle as defined by the ’812 patent.

A partial reproduction of FIG. 3 of Darrell (Ex-1004 at FIG. 3)



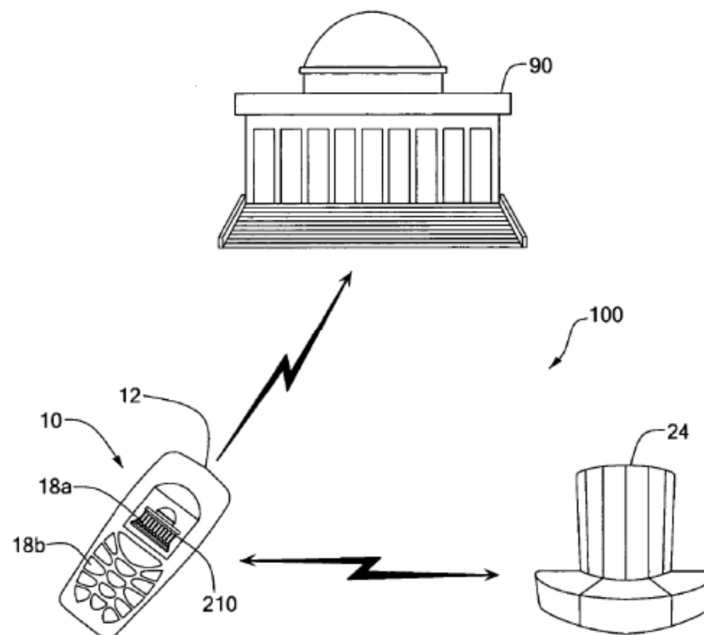
c. Actuation Means

131. Claim 1 of the '812 patent claims that the PID comprises “at least one actuation means for actuation by the user when the user points the pointing and identification device at the object” (Ex-1001, 49:12-14). I was asked to apply the construction of “actuation means” where the claimed function for the “actuation means” is “for actuation by the user when the user points the pointing and identification device at the object,” “forming a digital image of the object or of a portion of the object when the user points the pointing and identification device at the object and actuates the at least one actuation means;” and “communicating the digital image to a different location when the user actuates the at least one actuation

means.” I was asked to apply the construction of “actuation means” where the corresponding structure for the claimed “actuation means” constitutes a user interface such as a button(s) or a keyboard.

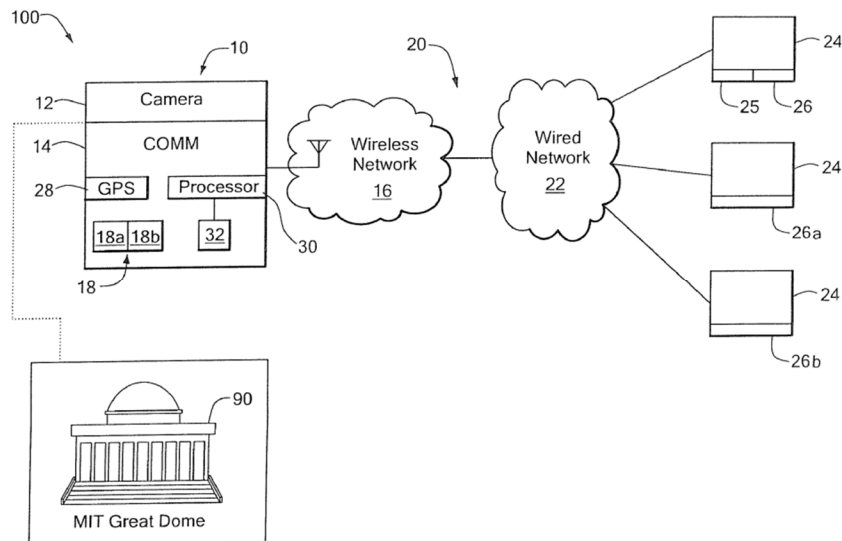
132. Darrell discloses the claimed “at least one actuation means” as cell phone buttons, a user interface, and/or a keyboard. For example, Darrell discloses in Figure 1 that the “mobile device 10” or PID has buttons, a “user interface,” and a “keyboard 18b” for “actuation by the user when the user points the pointing and identification device at the object,” which each constitute the claimed “at least one actuation means” (Ex-1004, ¶ [0021]). Darrell shows buttons for operating the cell phone camera in Figure 1.

Darrell Fig. 1



133. Darrell also describes that the handheld device 10 includes “a user interface 18, having here a display 18a and a keyboard 18b, coupled to the wireless communication device 14, to communicate to an user [*sic*] any results of found similar images” (Ex-1004, ¶ [0021]). Darrell further shows the device 10 or claimed PID has a user interface or “keyboard 18b” in Figure 1A for actuation by the user to take a picture and wirelessly communicate a digital image.

Darrell FIG. 1A



134. Darrell discloses that cell phone buttons, a user interface, and/or a keyboard perform the claimed functions. For example, Darrell discloses that: (i) the handheld device 10 includes a user interface, such as the “keyboard 18b” or the “small hand held computer” “to facilitate user interaction;” (ii) “[u]sing the hand held device 10 with a camera 12, an image-based query can be formed simply by pointing with the camera 12 and snapping a photo;” (iii) “a user causes the handheld

device 10” “to capture an image to send as a query;” and (iv) “wireless capability permits communication with a remote database” (Ex-1004, ¶ [0021], [0024], [0028], [0031], [0033], and [0042]).

135. In view of Darrell’s disclosure, a POSITA understands that Darrell discloses a cell phone having buttons and/or a keyboard for interfacing with a user, so that the user can control the cell phone’s operations. Further, Darrell describes that a user can take a picture with the cell phone and wirelessly communicate the digital image with Darrell’s cell phone. A POSITA understands that Darrell discloses that the cell phone having buttons and/or a keyboard are used by the user to conduct these functions.

136. To the extent that Pointwise contends that the claimed “actuation means” should not be construed under 35 U.S.C. § 112(f) and takes its plain meaning, Darrell discloses the claimed “actuation means” under a plain meaning construction for the reasons set forth above as Darrell discloses buttons, a user interface, and a keyboard.

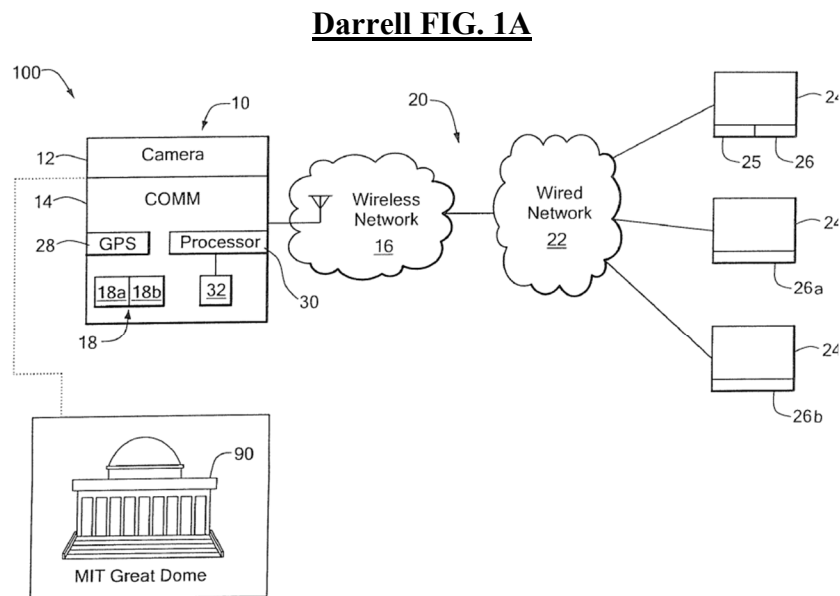
d. Digital Camera

137. The “Digital Camera” limitation states “a digital camera for forming a digital image of the object or a portion of the object when the user points the pointing and identification device at the object and actuates the at least one actuation means” (Ex-1001, 49:15-18).

138. Darrell describes that the camera can be “a Nokia 3650 phone taking advantage of its built-in camera (640×480 resolution) and the support for Multimedia Messaging Service (MMS)” that uses “C++ on Symbian OS to implement the required programming steps” (Ex-1004, ¶ [0021]). As I explained above, the built-in camera of the Nokia 3650 phone is a digital camera.

139. A POSITA further understands that Darrell describes the claimed “Digital Camera” because Darrell’s “mobile device” or “hand held device 10” is used to capture, snap, or take an image (Ex-1004, Abstract; ¶¶ [0021], [0024]), forms “an image-based query . . . simply by pointing with the camera 12 and snapping a photo” (Ex-1004, ¶ [0024]), and wirelessly communicates, or sends to, database 25 of a computer 24 using “a second program” stored on the storage medium of the handheld device to find similar images (Ex-1004, Abstract; ¶¶ [0021], [0026], and [0028]; FIG. 2; and FIG. 3). Accordingly, a POSITA would understand Darrell’s description to mean Darrell’s camera is a digital camera because Darrell is describing how to operate a digital camera, save the captured image digitally, and transmit wirelessly, which means that the images are digital. A POSITA understands that Darrell’s cell phone has a “digital camera” because that was the state of the art by September 2005. Analog cell phone cameras were not used as the state of the art in September 2005, as described above.

140. A POSITA further understands that Darrell describes the claimed “Digital Camera” because Darrell discloses a digital communication system in Figure 1A for communicating the image, meaning that the image must be digital. As illustrated in Figure 1A of Darrell, the camera 12 is part of the handheld device 10 that includes a processor 30, a communication device 14, and a storage medium 32. Darrell describes that the handheld device 10 includes “a first program stored on the storage medium 32 being operative to interact with the processor 30 to capture the image from the camera 12” (Ex-1004, ¶ [0026]). If the “first program” interacts with “the processor 30” to capture the image from the camera, then the camera and the image must be digital.



e. Communicating Device and Communicating

141. The “Communication Device and Communicating” limitation provides “a communication device for communicating the digital image to a different location when the user actuates the at least one actuation means” (Ex-1001, 49:12-21).

142. Darrell discloses this limitation because Darrell discloses (Fig 1A) that the “handheld device 10 further includes a wireless communication device 14 coupled to the camera and to a wireless network 16 to communicate the image 210 with a computer 24 having a database 25 with computer files 26 to find similar images and a user interface 18 ...” (Ex-1004, ¶ [0021]; *see also* step 114 of FIG. 4A, step 204 of FIG. 5A, step 254 of FIG. 5B, step 274 of FIG. 5C; ¶¶ [0021], [0028], [0031], [0033], and [0042]).

143. Darrell further discloses the “Communication Device and Communicating” Limitation throughout the specification. Darrell describes that, in “operation, a mobile user can point the camera 12 to the view of interest, take photos, and send them wireless as queries (via multimedia SMS, a.k.a. MMS) to the web database 25” (Ex-1004, ¶ [0025]). Further, Darrell teaches that a “mobile user can take photos of a place of interest,” and that the “photos are sent to the image database 25 via a wireless link” (Ex-1004, ¶ [0025]).

144. Darrell also describes that the handheld device 10 includes “a user interface 18” and a keyboard 18b (*e.g.*, “actuation means”) “coupled to the wireless communication device 14, to communicate to an user [*sic*] any results of found similar images” (Ex-1004, ¶ [0021]).

145. Darrell shows in Fig. 1A that the digital image is communicated from the handheld device 10 to a “different location,” where the computer 24 and database 25 are located. The mobile device 10 is at a “different location” than the computer 24 and database 25. Darrell describes the “database 25” as being a “remote database 25,” so the “different location limitations are satisfied (Ex-1004, ¶ [0024]).

Darrell FIG. 1A

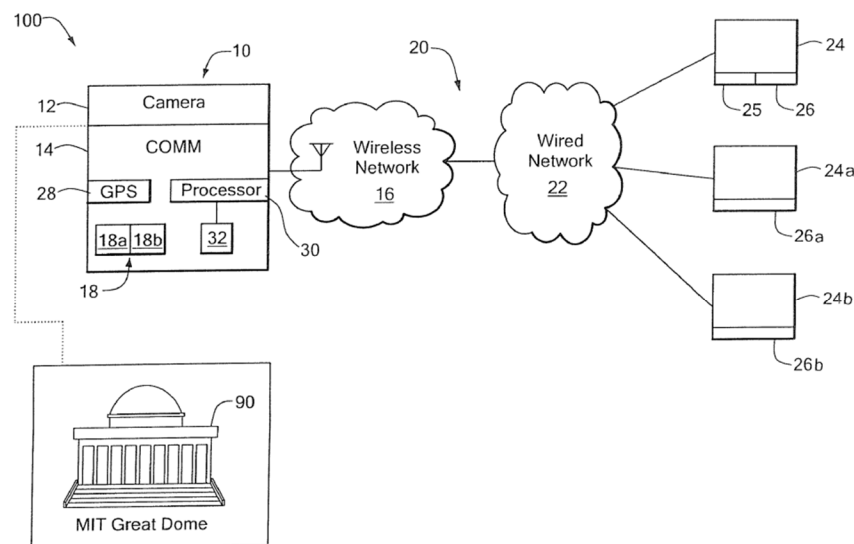


FIG. 1A

146. Darrell also describes that, in “operation, a mobile user can point the camera 12 to the view of interest, take photos, and send them wireless as queries (via

multimedia SMS, a.k.a. MMS) to the web database 25” (Ex-1004 at 5:21-25). Darrell teaches that a “mobile user can take photos of a place of interest,” and that the “photos are sent to the image database 25 via a wireless link” (Ex-1004 at 5:31-33). Darrell describes that the handheld device 10 includes a user interface, such as a keyboard 18b or “small hand held computer ... to facilitate user interaction” (Ex-1004 at 4:6-28).

147. Based on Darrell’s disclosure, Darrell discloses that the handheld device includes “a wireless communication device 14” that is for communicating “the image 210” to a different location (e.g., the computer 24 that has the database 25), and the handheld device includes a user interface, such as a keyboard 18b or “small hand held computer ... to facilitate user interaction.”

148. Darrell shows in Fig. 1A that the digital image is communicated to a different location to the computer 25 and database 25 through the networks. Darrell describes the “database 25” as being a “remote database 25,” so the “different location limitations are satisfied (Ex-1004, ¶ [0024]).

149. For further evidence that the Communication Device limitation and the Communicating limitation of claim 1 of the ‘812 patent are taught by Darrell, see step 114 of FIG. 4A, step 204 of FIG. 5A, step 254 of FIG. 5B, step 274 of FIG. 5C (Ex-1004 at 6:33-39; 7:33-39; 8:11-15; 10:64-11:3).

f. Identifying

150. The “Identifying” limitation provides “automatically identifying a list of likely pointed-to objects from the digital image at the different location to return the list of likely pointed-to objects” (Ex-1004, 49:24-26). Darrell teaches this limitation with a computer 24 that can return the claimed list.

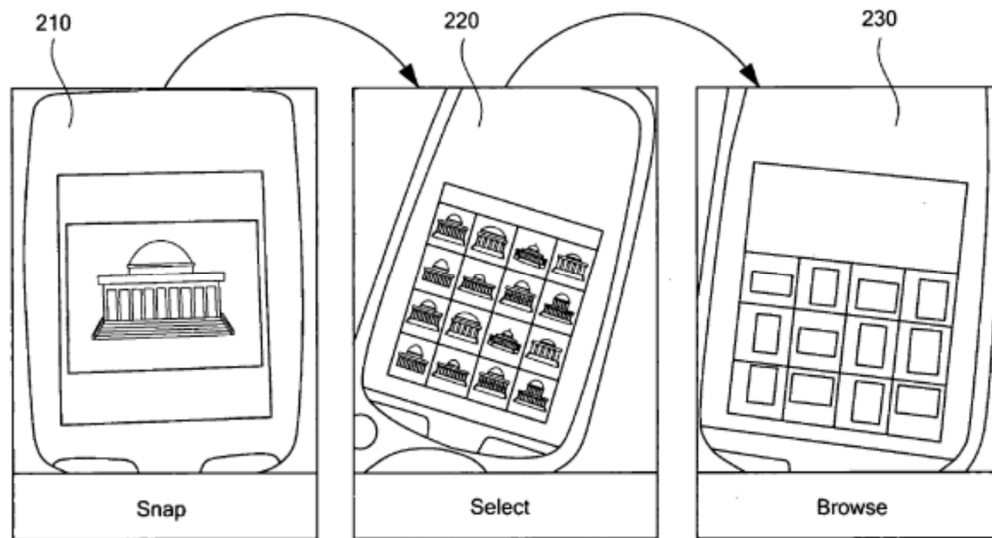
151. Darrell teaches “identifying a list of likely pointed-to objects from the digital image” because Darrell describes that the wireless communication device of the handheld device can “communicate the image 210 with a computer 24 having a database 25 with computer files 26 to find similar images . . . to communicate to an user [*sic*] any results of found similar images” (Ex-1004, ¶ [0021]). Darrell teaches that the computer 24 or server includes a search engine will provide a result that “consist of a list of (candidate image, source URL) pairs,” so that the user can “identify what sub-images” are relevant (Ex-1004, ¶ [0025]).

152. Darrell also describes that, it “should now be appreciated that a method comprises the steps of: (i) providing a database of images, each image having an associated URL that includes said image and a description of the image; (ii) comparing an image of an unknown location with images from the database of images and providing a list of possible matching images; and (iii) reviewing the images in the list of possible matching images until the correct location is identified” (Ex-1004, claim 26).

153. Darrell describes that the “identifying” is “automatically” performed as Darrell teaches that, using “the handheld device 10 with a camera 12, an image-based query can be formed simply by pointing with the camera 12 and snapping a photo” (Ex-1004, ¶ [0024]). Darrell also describes that, in “operation, a mobile user can point the camera 12 to the view of interest, take photos, and send them wireless as queries (via multimedia SMS, a.k.a. MMS) to the web database 25” (Ex-1004, ¶ [0025]). Darrell explains that the “identifying” is “automatically” performed in response without any further interaction (Ex-1004, ¶¶ [0025]-[0027]).

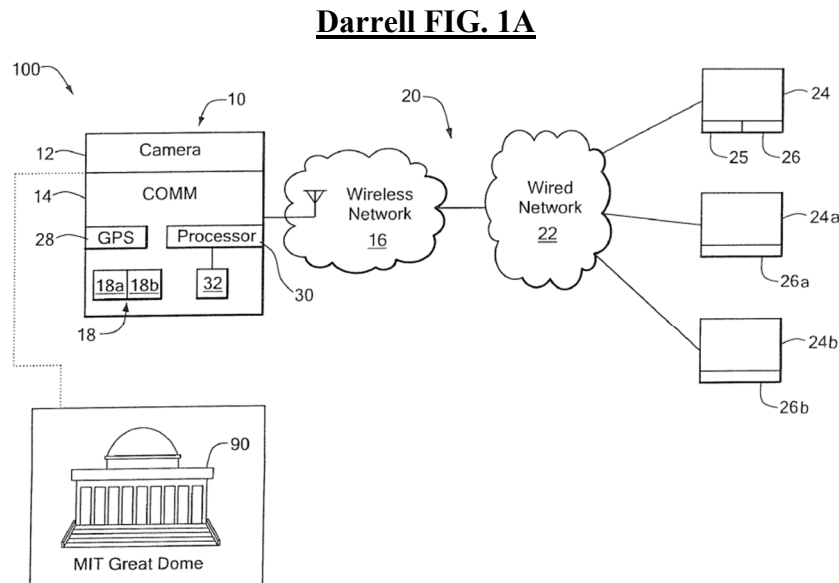
154. Darrell teaches “to return the list of likely pointed-to objects.” As Darrell explains, a result “will consist of a list of (candidate image, source URL) pairs,” so that the user can “identify what sub-images” are relevant (Ex-1004, ¶¶ [0025], [0028]). Figure 2 of Darrell shows the return of thumbnails or a list of likely pointed-to-objects (Ex-1004, ¶ [0028]).

Darrell Figure 2



155. Darrell teaches “automatically identifying a list of likely pointed-to objects from the digital image” (*e.g.*, Darrell’s “a list of (candidate image, source URL) pairs” and “list of possible matching images”) “at the different location to return the list of likely pointed-to objects” (Ex-1004, ¶ [0025], claim 26; *see also* ¶ [0028]). As Darrell explains, a result “will consist of a list of (candidate image, source URL) pairs,” so that the user can “identify what sub-images” are relevant (Ex-1004, ¶¶ [0025], [0028]).

156. Darrell teaches the “at a different location” because the computer or server that includes the search engine and image database 25 in Darrell is at a “different location” from the handheld device 10 as shown in Fig. 1A. Also, Darrell describes the “database 25” as being a “remote database 25,” so the “different location limitations are satisfied (Ex-1004, ¶ [0024]).



157. Darrell provides an example that discloses this limitation. In the example, a mobile user is uncertain whether a building is “Killian Court” (Ex-1004, ¶ [0027]). The user takes a picture of the building with the mobile and sends the image to the server 24 (Ex-1004, ¶ [0027]). The server automatically returns “the most relevant web pages” without any intervention (Ex-1004, ¶ [0027]).

158. Darrell also discloses the “Identifying” limitation throughout Darrell (Ex-1004, step 116 of FIG. 4A, step 206 of FIG. 5A, step 256 of FIG. 5B, and FIG. 6; ¶¶ [0021], [0028], [0031], [0033], and [0042]).

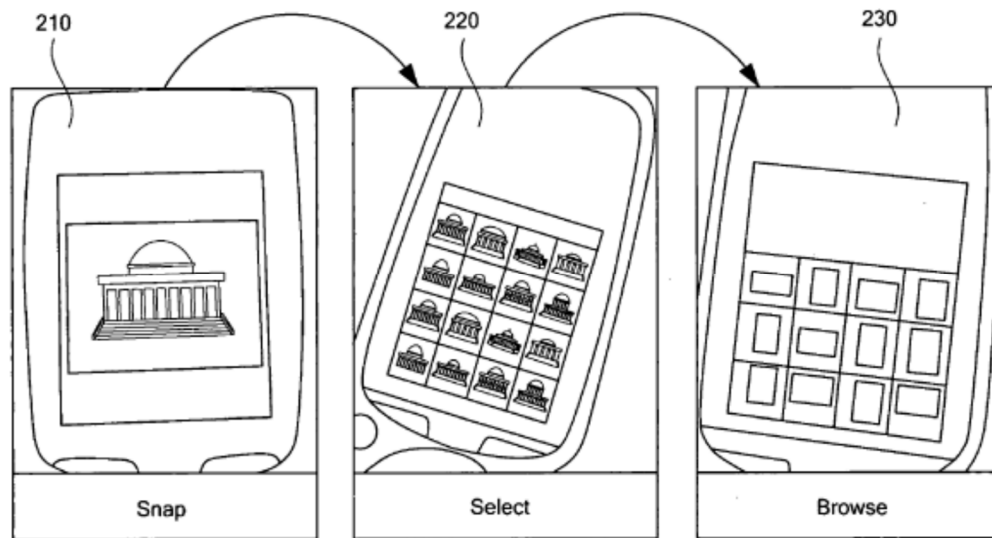
g. Returning

159. The “Returning” limitation states “returning the list of likely pointed-to objects to the user to select one of the likely pointed-to objects” (Ex-1001, 49:27-28). Darrell describes that the wireless communication device of the handheld

device can “communicate the image 210 with a computer 24 having a database 25 with computer files 26 to find similar images . . . to communicate to an user [*sic*] any results of found similar images” (Ex-1004, ¶ [0021]). Specifically, Darrell teaches that the computer or server that includes a search engine, which is at a different location, will provide a result that “consist of a list of (candidate image, source URL) pairs,” so that the user can “identify what sub-images” are relevant (Ex-1004, ¶ [0025], *see also* claim 26).

160. Darrell describes that “a list of possible matching images” can be provided by the search engine back to the handheld device in response to the image captured by the camera, so that the user can identify, or select, what “sub-images”, or subset of images from the provided list, are most relevant (Ex-1004, ¶ [0025]). Darrell discloses returning a “list of likely pointed to objects” in Figure 2 where a series of thumbnails or list were provided to the user (Ex-1004, ¶ [0028]).

Darrell Figure 2



h. Object

161. Darrell describes that the “a location awareness system 100 includes a handheld device 10 (sometimes also referred to as mobile deixis device 10) having a camera 12 to capture an image 210 of an object 90” (Ex-1004, 4:6-9). Darrell teaches throughout the disclosure that the “mobile device” or “hand held device 10” includes “a camera to capture an image” (Ex-1004 at Abstract; and 5:22-25). Darrell also discloses that “users can specify a location of interest by simply pointing a camera-equipped cellular phone at the object of interest and by searching an image database or relevant web resources, users can quickly identify good matches from several close ones to find an object of interest” (Ex-1004 at Abstract; and 5:22-25).

162. Darrell's Field of the Invention section states that this "invention relates generally to location and object awareness systems and more particularly to systems and techniques to identify a location or object in a person's field of view" (Ex-1004 at 1:5-9). An object "in a person's field of view" is an "object in space." Therefore, Darrell teaches that the object 90 is at least "an object in space." As such, Darrell discloses the "Object" limitation of claim 1 of the '812 patent.

2. Darrell Anticipates Claim 3

163. Claim 3 recites the "method of claim 1, further comprising receiving a selection by the user of said one of the likely pointed-to objects as the object and an instruction from the user to obtain information on the object" (Ex-1001, 50:8-12).

164. Darrell discloses "receiving a selection by the user of said one of the likely pointed-to objects as the object" as Darrell teaches that the "mobile device 10 displays the result by arranging candidate images into a thumbnail mosaic 220," the "user, as the final judge, can easily identify what sub-images are 'really relevant'" by selecting an image, and when "a thumbnail is selected, the source URL is retrieved and the content from that URL is shown on the mobile device 10" (Ex-1004, ¶ [0025]). Similarly, in FIG. 5C, Darrell discloses that "a user can then select one of the images and the content from the associated URL is then displayed" (Ex-1004, FIG. 5C).

165. Darrell also teaches that the handheld device receives “an instruction from the user to obtain information on the object” with a URL. For example, Darrell teaches when “a thumbnail is selected, the source URL is retrieved and the content from that URL is shown on the mobile device 10” (Ex-1004, ¶ [0025]). Moreover, Darrell describes that the handheld device can receive an instruction in the form of “keywords” from the user to search for “additional web pages that might share conceptual similarity with the query image (Ex-1004, ¶ [0039]; Fig. 6). Darrell also describes that “keywords” can be “linked to each similar image and the server database is capable of initiating a further search using the keywords to find additional similar images” (Ex-1004, ¶ [0007]). Darrell describes that “automatically extracted keywords are displayed side-by-side with the thumbnail image,” and that if “further searching is required, as shown by decision block 258, the process continues with process step 262 where selecting a keyword initiates a keyword-based search on Google to find more information” (Ex-1004, ¶ [0031]). Darrell describes that “keywords can be found in the surrounding text and used directly as a location context cue, or used for further interactive browsing to find relevant information resources” (Ex-1004, ¶ [0034]).

3. Darrell Anticipates Claim 10

166. Claim 10 depends from claim 1 and states “wherein the object is an object in space, or near an object in space” (Ex-1001, 50:28-29). Darrell teaches

that the object is an object in space (*see supra* “Object” limitation). Darrell discloses taking a picture of an “object” in the form of a building in Figure 2 that is an “object in space.” Darrell discloses taking a picture of “Killian Court” which is an “object in space” (Ex-1004, ¶ [0027]). Darrell also discloses Darrell teaches taking a picture that includes “people in front of a building” and people and building are “objects in space” (Ex-1004, ¶ [0037]). As previously noted, Darrell’s Field of the Invention section states that “This invention relates generally to location and object awareness systems and more particularly to systems and techniques to identify a location or object in a person's field of view” (Ex-1004, ¶ [0001]). An object “in a person’s field of view” is an “object in space.” As such, Darrell anticipates claim 10 of the ’812 patent.

4. Darrell Anticipates Claim 11

167. Claim 11 depends from claim 1 and states “wherein the object is a subarea of a surface in space” (Ex-1001, 50:30-31). Darrell teaches that the object is a subarea of a surface in space as shown in Figure 2 where the building is a “subarea” of a “surface” (people and building) in “space” (*e.g.*, field of view) (*see supra* “Object” Limitation). Darrell also describes that one image detection technique could be based on the “energy spectrum” of the image, and that “different placements of people in front of a building should not affect its the [*sic*] image representation too dramatically” (Ex-1004, ¶ [0037]). Darrell also teaches that the

“system allows users to stay where they are and point at a remote place in sight simply by taking photographs” (Ex-1004, ¶ [0023]). Darrell teaches taking a picture that includes “people in front of a building” (Ex-1004, ¶ [0037]). The building and each person are an “object” that “is a subarea of a surface in space.” Since there may be people in front of the building that block at least a portion of the surfaces of the building and/or the user may be remote from the object being identified, Darrell teaches that the object being identified is a subarea of a surface in space.

5. Darrell Anticipates Claim 12

168. Claim 11 depends from claim 1 and states “wherein the object is one of a plurality of objects in space.” (Ex-1001, 50:30-31). Darrell teaches that the object is a subarea of a surface in space as shown in Figure 2 (*see supra* “Object” Limitation). Darrell also describes that one image detection technique could be based on the “energy spectrum” of the image, and that “different placements of people in front of a building should not affect its the [*sic*] image representation too dramatically.” (Ex-1004, ¶ [0037]). Therefore, Darrell teaches taking a picture that includes an “object” (*e.g.*, building) that is “one of a plurality of objects in space” (*e.g.*, people and building). Since there may be people in front of the building that block at least a portion of the surfaces of the building, Darrell teaches that the object is one of a plurality of objects in space.

B. Obviousness over Boncyk and Darrell

169. The USPTO allowed the '812 patent claims over Boncyk primarily based on the argument that Boncyk returned a detected object rather than the claimed “list” of likely objects. Darrell discloses the missing limitation. Claims 1-12 are obvious over Boncyk in view of Darrell.

1. Claim 1 is Obvious over Boncyk in view of Darrell

170. Boncyk in view of Darrell renders claim 1 obvious as described below.

a. Preamble

171. Boncyk discloses a “method for identifying an object” throughout the disclosure, for example, at least in its Abstract and Field of the Invention, where Boncyk discloses “a method by which information and communication pertinent to an object is provided based on imagery of the object” where a “digital image of the object is captured and the object is recognized from a plurality of objects in a database” (Ex-1005, Abstract, see also ¶ [0026]).

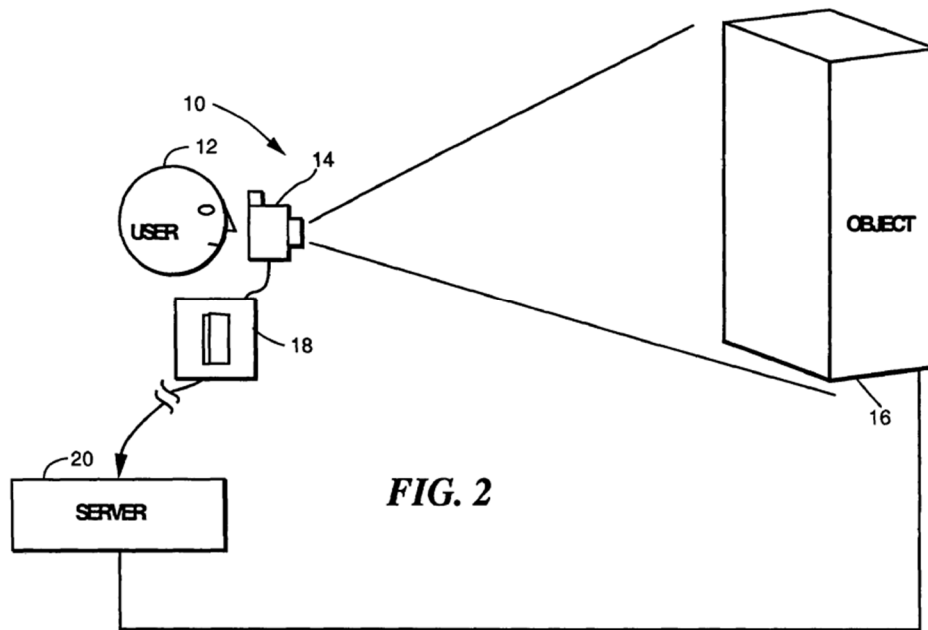
172. Darrell discloses the Preamble.

b. Providing a PID

173. Boncyk discloses an “image capture device 14,” “terminal 102, and “terminal 202” (Ex-1005, ¶¶ [0045], [0128], [0269]; Figs. 2, 4, 5) each of which are the claimed “pointing and identification device for pointing at the object.” For example, Boncyk discloses a system that includes “a computer, mobile telephone,

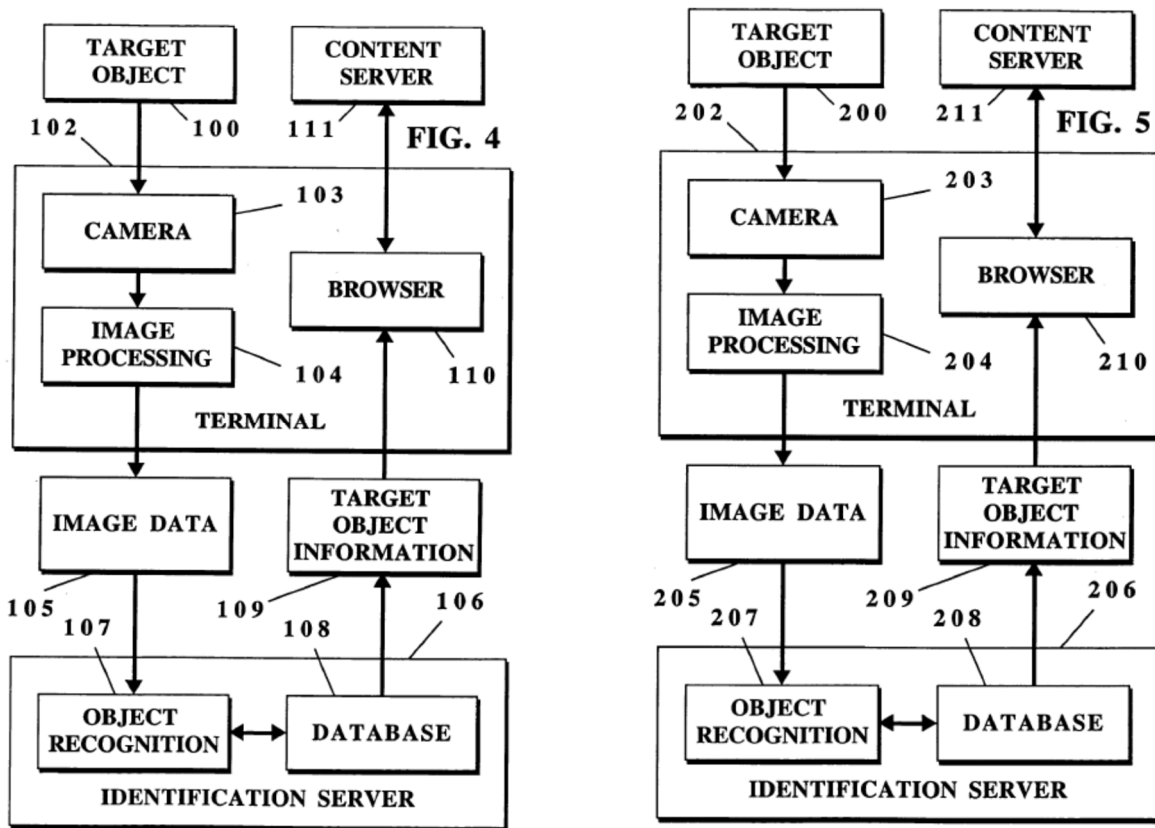
personal digital assistant, or other similar device 14 equipped with an image sensor (such as a CCD or CMOS digital camera)” that “captures a digital image 18 of the scene at which it is pointed” (Ex-1005, ¶ [0045]).

Boncyk Figure 2



174. Boncyk discloses that each of “terminal 102” (which has an “image capture device such as digital camera 103” (Ex-1005, ¶ [0128]; Fig. 4)) and “terminal 202” can be a “cellular telephone, PDA, or such mobile device equipped with computational capability, a digital camera, and wireless connection (Ex-1005, ¶ [0269]).

Boncyk Figures 4-5



175. Boncyk describes that a user 12 “aligns the sensor of the image capture device 14 with the object 16 of interest,” and that the “device 14 captures a digital image 18 of the scene at which it is pointed” (Ex-1005, ¶ [0045]; Fig. 5).

176. Boncyk also discloses that the “user acquired images are transmitted from the device 14 to the Image Processor/Server 20” and that a “Best Match 38” is determined from image “Database Matching 36” (Ex-1005, ¶¶ [0046], [0049]).

177. Darrell discloses the claimed “Providing a PID.”

c. Actuation Means

178. Boncyk discloses the claimed “actuation means” for actuation by the user when the user points the pointing and identification device at the object.

179. The claimed function for the “actuation means” is “for actuation by the user when the user points the pointing and identification device at the object,” “forming a digital image of the object or of a portion of the object when the user points the pointing and identification device at the object and actuates the at least one actuation means;” and “communicating the digital image to a different location when the user actuates the at least one actuation means.” The corresponding structure for the claimed “actuation means” constitutes a user interface such as a button(s) or a keyboard.

180. Boncyk discloses that the “image capture device 14,” “terminal 102, and “terminal 202” perform the recited functions of “actuation by the user,” and “forming a digital image.” For example, Boncyk discloses that the terminal 102 has a “digital camera 103” (Fig. 4), terminal 202 has a “digital camera” (Fig 5), and image capture device 14 “captures a digital image” (Fig. 2) (e.g., has a digital camera) (Ex-1005, ¶¶ [0045], [0128], [0269]).

181. Boncyk further discloses the “communicating” function. Boncyk discloses that the that the “image capture device 14” transmits “acquired images” to server 20 using a “wireless network means” (Ex-1005, ¶ [0046]; Fig. 2). Likewise,

Boncyk describes that the “terminal 102” sends image data 105 to the identification server 106 via networks such as a “cellular telephone network” and “wireless internet” (Ex-1005, ¶¶ [0140], [0141]; Fig. 4). Similarly, Boncyk describes that the “terminal 202” transfers “image data 205 to the identification server 206” and can do so via networks such as a cellular telephone network” (Ex-1005, ¶¶ [0269], [0270]; Fig. 5).

182. Boncyk discloses that the corresponding structure for the claimed “actuation means” constitutes a user interface such as a button(s) or a keyboard. For example, Boncyk also discloses that “‘point and click’ refers to the operation of the user capturing imagery of the target object 200 and initiating the transfer of the image data 205 to the identification server 206” (Ex-1005, ¶ [0270]). As “buttons” are clicked, Boncyk discloses buttons as corresponding structure. Similarly, Boncyk discloses that the image capture device 14 has “any suitable means such as a button on the device 14 or sensor” for image capture (Ex-1005, ¶ [0045]). Moreover, Boncyk discloses that the “image capture device 14,” “terminal 102, and “terminal 202” all have a “digital camera” and can be “mobile devices” which all have buttons and/or keyboards for actuation, forming a digital image, and communicating the digital image (Ex-1005, ¶¶ [0045], [0128], [0269]; Figs. 2, 4, 5).

183. Darrell discloses the claimed “Actuation Means.”

d. Digital Camera

184. Boncyk discloses the claimed “Digital Camera” limitation. Boncyk describes, in its Abstract and Field of the Invention, “a method by which information and communication pertinent to an object is provided based on imagery of the object” where a “digital image of the object is captured and the object is recognized from a plurality of objects in a database” (Ex-1005, Abstract, ¶ [0001]).

185. Boncyk discloses a system that includes “a computer, mobile telephone, personal digital assistant, or other similar device 14 equipped with an image sensor (such as a CCD or CMOS digital camera)” that “captures a digital image 18 of the scene at which it is pointed” (Ex-1005, ¶ [0045]).

186. Boncyk also explains that “terminal 102” that “is a computing device that has an ‘image’ capture device such as digital camera 103, a video camera, or any other device that [c]an convert a physical object into a digital representation of the object” (Ex-1005, ¶ [0128]). Boncyk likewise discloses that terminal 202 has a “digital camera” (Fig 5) for “capturing imagery” (Ex-1005, ¶ [0269]).

187. Boncyk also discloses that the “actuation means” constitutes a user interface such as a button(s) or a keyboard for forming the digital image.

188. Darrell also discloses the Digital Camera limitation.

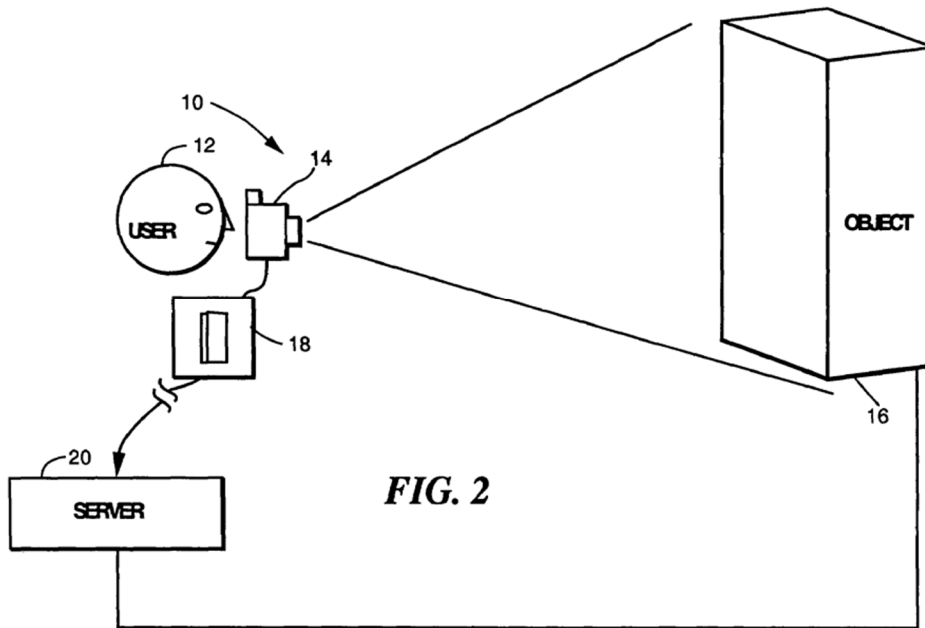
e. Communication Device and Communicating

189. Boncyk discloses “a communication device for communicating the digital image to a different location when the user actuates the at least one actuation means.” Darrell also discloses this limitation as describe above.

190. Boncyk discloses that each of the “image capture device 14,” “terminal 102, and “terminal 202” have the claimed “communication device.” For example, Boncyk (Fig. 2) describes that the system may include “a computer, mobile telephone, personal digital assistant, or other similar device 14 equipped with an image sensor (such as a CCD or CMOS digital camera)” that “captures a digital image 18 of the scene at which it is pointed” (Ex-1005, ¶ [0045]).

191. Boncyk also describes that “user acquired images are transmitted from the device 14 to the Image Processor/Server 20 using a conventional digital network or wireless network means,” and notes that the “server 20” may be “physically separate from the device 14” (Ex-1005, ¶ [0046], see also ¶ [0030]).

Boncyk Figure 2



192. Similarly, terminal 102 has the claimed “communication device” because terminal 102 can transmit image data 105 wirelessly to an identification server 106 that is “usually separate from the terminal 102” (Ex-1005, ¶ [0141]). Likewise, terminal 202 has the claimed “communication device” because terminal 202 can be a cellular telephone, PDA, mobile device, with a “wireless connection” for communicating with server 211 (Ex-1005, ¶ [0269]).

193. Boncyk discloses the “different location” limitation. The Boncyk device 14 may be a mobile telephone. Boncyk describes that the images are “transmitted from the device 14” to a “physically separate” “server 20” which is shown to be at a different location in Figure 2 (Ex-1005, ¶ [0046], see also ¶ [0030]).

194. As described above, Boncyk discloses the claimed “actuation means” for communicating as claimed.

195. Darrell discloses the Communication Device and Communicating limitations.

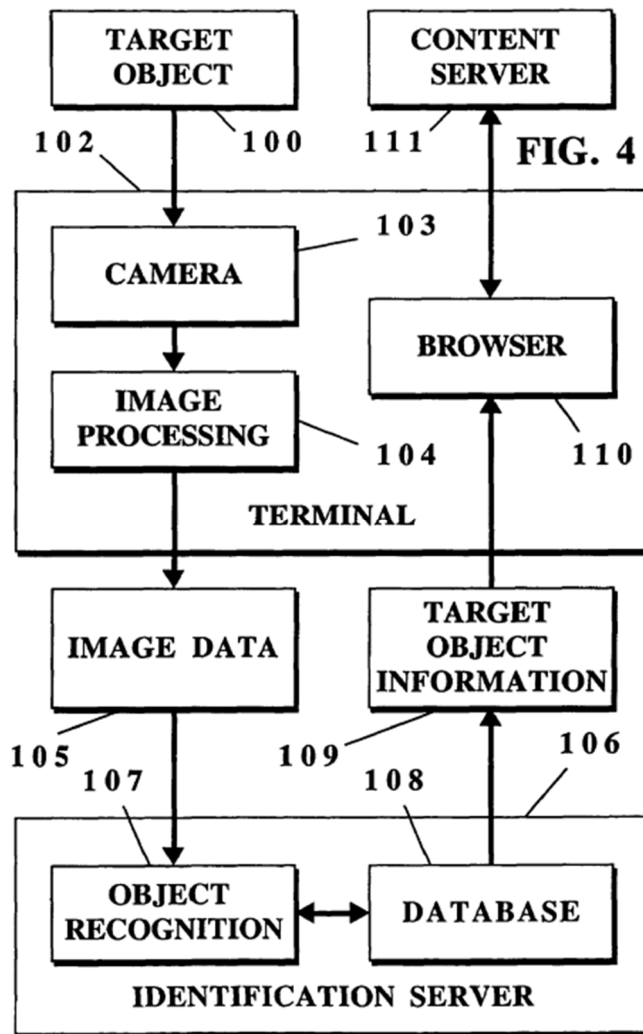
f. Identifying

196. Boncyk discloses the “Identifying” limitation with the exception of expressly disclosing “to return the list of likely pointed-to-objects.” However, as described above, Darrell discloses “to return the list of likely pointed-to-objects.” This limitation is obvious over Boncyk in view of Darrell.

197. Boncyk disclose the first part of the “Identifying” limitation: “automatically identifying a list of likely pointed-to objects from the digital image at the different location” (Ex-1001, 49:24-25). Boncyk discloses that after “the camera 103 captures the digital imagery of the target object 100, image preprocessing 104 software converts the digital imagery into image data 105 for transmission to and analysis by an identification server 106,” and that the “[i]mage data 105 is data extracted or converted from the original imagery of the target object 100 and has information content appropriate for identification of the target object 100 by the object recognition 107, which may be software or hardware” (Ex-1005, ¶ [0128]). Boncyk discloses that the “overall function of the identification server 106” is to determine “target object information 109” that corresponds to “the target

object 100” (Ex-1005, ¶ [0143]). This is illustrated in at least Figures 4 (reproduced below) and Figure 5.

Boncyk Figure 4



198. Boncyk teaches multiple algorithms at the identification server (e.g., shape comparison, grayscale comparison, wavelet comparison, and color cube comparison) that run in parallel that generate “matches” or “a list of likely pointed-to objects from the digital image” (Ex-1005, ¶¶ [0070], [0118], [0119]). Boncyk

discloses that the “matches” are weighed and ranked and those above a threshold are assessed in a “Combined Match Score evaluation” (Ex-1005, ¶¶ [0118]-[0121]). Similarly, Boncyk discloses there may be “many such objects that have similar appearance” as the target object (Ex-1005, ¶ [0171]).

199. Boncyk discloses returning “target object information” to the image capture device 14, terminal 102, or terminal 202 (Ex-1005, ¶ [0237]). The “target object information” 109 is the information “corresponding to the target object” such as “information address,” “identity” and “position” (Ex-1005, ¶¶ [0238]-[0244]).

200. Boncyk also teaches the “automatically” limitation as Boncyk discloses that “the object recognition technique” is “[a]utonomous (no human assistance required in the recognition process)” (Ex-1005, ¶ [0156]).

201. Boncyk discloses the “at a different location” limitation in Figs 2-5. Boncyk discloses that the “server 20” may be “physically separate from the device 14” (Ex-1005, ¶ [0046], see also ¶ [0030]). Further, Boncyk describes that the identification server 106 can “exist on [a] computing platform separate from the terminal 102” (Ex-1005, ¶ [0141]).

202. Although Boncyk determines a list of potential matches to the target object, Boncyk does not expressly state that the returned target object information 109 can include a list or more than one “target object information” corresponding to the target object. However, Darrell teaches that identifying the claimed list is “to

return the list of likely pointed-to-objects.” During prosecution of the ’812 patent before the USPTO, it was argued that Boncyk does not teach that “a list of likely pointed-to objects” is identified or returned (Ex-1002, p. 36, 39-40). However, the Examiner did not cite the Darrell reference.

203. It would have been obvious to a POSITA to combine Boncyk and Darrell to form the claimed invention. Boncyk and Darrell disclose all of the claim limitations. As noted above, Darrell teaches that a search engine will identify and return a result that will “consist(s) of a list of (candidate image, source URL) pairs,” and also describes that “a list of possible matching images” can be identified and provided to the user for review in response to the image captured by the camera (Ex-1004, ¶ [0025], claim 26). Thus, there is no difference between the claim and the prior art.

204. There are no known secondary considerations of nonobviousness.

205. A POSITA would have been motivated to combine Boncyk with Darrell prior to September 2005 to form the claimed invention. Boncyk generates a likely list of matches and can have best matches that are above a threshold (Ex-1005, ¶¶ [0070], [0118]-[0121]). Darrell teaches the benefit of sending a likely list to the user instead of only one match because the user can examine the list to determine the best match. As Darrell explains, the user can be the “final judge” and can “easily identify what sub-images” are “really relevant” (Ex-1004, ¶ [0025]). Thus, Darrell’s

likely list avoids errors because with the Boncyk approach the best software determined match may not be correct, and if Boncyk determines an incorrect match, the user will either be using incorrect information or recognize that the match is incorrect and will be without the correct information because Boncyk provides no other options.

206. For example, Boncyk discloses that the “Calculate Combined Match Score” algorithm may result in “statistically equivalent” matches when resolving close rankings, but provides no method of resolving the equivalency (Ex-1005, ¶ [0118]-[0122]). Further, Boncyk discloses that the “target object information 109” “can be used by the terminal 102 or content server 111 to identify the specific target object 100, out of many such objects that have similar appearance and differ only in the identifying marks” (Ex-1005, ¶ [0171]). Therefore, Boncyk appreciates that the object will need to be identified “out of many such objects that have similar appearance” and the potential for errors and close calls (Ex-1005, ¶ [0171]).

207. Moreover, Boncyk and Darrell are in the same technical fields (image recognition and retrieval). Combining the references leads to a predictable result — sending a list of target object information, instead of one target object information. There is no technical difficulty in sending more than one target object information.

g. Returning

208. As described, Boncyk discloses returning target object information 109 to the user but does not disclose the claimed list of likely pointed-to objects that is returned to the user so the user can select one of the likely pointed to objects. Darrell discloses this limitation. The “Returning” limitation is obvious over Boncyk in view of Darrell as explained in the “Identifying” limitation analysis.

h. Object

209. Boncyk discloses the “Object” limitation. For example, Boncyk describes the “target object information 109 can be used by the terminal 102 or content server 111 to identify the specific target object 100,” from the many such objects that have “similar appearance and differ only in the identifying marks” (Ex-1005, ¶ [0171]). Boncyk also describes that the object recognition is capable of “recognizing both 2-dimensional (i.e., flat) and 3-dimensional objects” and of “discriminating the target object 100 from any foreground or background objects or image information, i.e., be robust with respect to changes in background.” (Ex-1005, ¶¶ [0152]-[0154]). Therefore, Boncyk discloses that the object could be “an object in space” as provided by the Object limitation of claim 1 of the ‘812 patent.

210. Darrell teaches the Object limitation.

2. Claim 2 is Obvious over Boncyk and Darrell

211. Boncyk discloses a method of (i) receiving a designation of the object as designated by the user; (ii) recording a frame image comprising the designated object; (iii) determining a context type of the frame image from the frame image; and (iv) identifying the object, as recited by claim 2 of the '812 patent.

212. Boncyk discloses that the server 20 receives a designation of the object in the form of image data. For example, Boncyk discloses that the “identification server 106” and the “object recognition 107” receive “the image data 105” from the terminal 102 (Ex-1005, ¶ [0140]). Alternatively, Boncyk discloses “receiving a designation of the object as designated by the user” as Boncyk discloses that the object received by the server 20 may have “recognizable symbols, such as bar codes, matrix codes, or alphanumeric codes” (Ex-1005, ¶ [0047]).

213. Boncyk states that the imagery that includes the target object 100 “can be a single image, a series of images, or a continuous video stream” (Ex-1005, ¶ [0128]). Therefore, Boncyk “records” or saves a frame image that includes the target object 100 to analyze the frame image for target information.

214. Boncyk discloses “determining a context type of the frame image from the frame image.” For example, Boncyk discloses an algorithm that can “find an object or portion thereof in an input image, given a database of multiple objects with multiple views (from different angles) of each object” (Ex-1005, ¶ [0200]). Boncyk

teaches that this could be done using a “Local Image Descriptor (LID)” process that finds the “best matches for multiple input image LIDs” (Ex-1005, ¶¶ [0201]-[0218]). Each of these views of a multiple input image is a different context of the image, and these views are used to find the “best matches” for the target object 100.

3. Claim 3 is Obvious over Boncyk and Darrell

215. Claim 3 recites “receiving a selection by the user of said one of the likely pointed-to objects as the object and an instruction from the user to obtain information on the object” (Ex-1001, 50:8-12). Boncyk discloses that “the target object information 109 may include, but not be limited to, one or more of the following items of information pertaining to the target object 100: Information address (e.g., Internet URL), Identity (e.g., object name, number, classification, etc.); Position; Orientation; Size; Color; Status; Information decoded from and/or referenced by symbols (e.g. information coded in a barcode or a URL referenced by such a barcode); and other data (e.g. alphanumerical text),” such that “the identification server determines the identity and/or various attributes of the target object 100 from the image data 105” (Ex-1005, ¶¶ [0237]-[0246]).

216. Boncyk describes that, after the target object information 109 is returned to the terminal 102, the “user can perform ongoing interactions with the content server 111” (Ex-1005, ¶ [0263]). For example, Boncyk teaches that the user can “Listen to streaming audio samples” or “Purchase the target object 100 via on-

line transaction.” (Ex-1005, ¶¶ [0234]-[0265]). These “ongoing interactions” with the content server 111 include both “selection by the user” and “an instruction from the user to obtain information on the object.”

217. Darrell discloses claim 3.

4. Claim 4 is Obvious over Boncyk and Darrell

218. Claim 4 recites “further comprising receiving a selection by the user of said one of the likely pointed-to objects as the object and an instruction from the user to purchase the object” (Ex-1001, 50:12-15). Boncyk describes that, after the target objection information 109 is returned to the terminal 102, the “user can perform ongoing interactions with the content server 111” (Ex-1005, ¶ [0263]). For example, Boncyk discloses that the user can “[p]urchase the target object 100 via on-line transaction, with the purchase amount billed to an account linked to the terminal 102, to the individual user, to a bank account, or to a credit card” (Ex-1005, ¶ [0265]). Therefore, Boncyk teaches “ongoing interactions” including the user “selecting” the object in order to provide an instruction to “purchase” the “the target object 100 via on-line transaction.”

5. Claims 5-9 are Obvious over Boncyk and Darrell

219. Claim 5 states “wherein the object is at least one of the spot on the displayed image on the display, the sub-area of the space on the displayed image on

the display, and one of the plurality of objects in the displayed image on the display” (Ex-1001, 50:16-19).

220. Claim 6 depends from claim 5 and states “wherein the object is the sub-area of the space on the displayed image on the display” (Ex-1001, 50:20-21).

221. Claim 7 depends from claim 1 and states “wherein the object is a spot on a displayed image on a display” (Ex-1001, 50:22-23).

222. Claim 8 depends from claim 1 and contains the same limitations as claim 5, but depends from claim 1. Claim 8 states “wherein the object is a subarea of a space on the displayed image on the display” (Ex-1001, 50:24-25).

223. Claim 9 depends from claim 1 and states “wherein the object is one of a plurality of objects in the displayed image on the display” (Ex-1001, 50:26-27).

224. Boncyk describes an embodiment where the user “‘points and clicks’ the terminal 202 at a television screen (*e.g.*, “sub-area of the space on the displayed image on the display”) to interact with television programming in real-time” (Ex-1005 at 23:48-55). As noted, Boncyk describes that the term “‘point and click’ refers to the operation of the user capturing imagery of the target object 200 and initiating the transfer of the image data 205 to the identification server 206” (Ex-1005 at 22:48-55). Therefore, Boncyk teaches that the object that is captured and sent to the identification server could be an object on a television screen.

225. Further, Boncyk also teaches that its system is “[c]apable of discriminating the target object 100 from any foreground or background objects or image information” (Ex-1005 at 15:30-38). Therefore, Boncyk teaches that the object could be one of many objects on a television screen, and as such, Boncyk discloses that the object being detected is a “spot on a displayed image on a display,” a “sub-area of the space on the displayed image on the display,” and “the object is one of a plurality of objects in the displayed image on the display,” as set forth in claims 5-9 of the ’812 patent.

226. Boncyk describes an embodiment where the user “‘points and clicks’ the terminal 202 at a television screen to interact with television programming in real-time.” (Ex-1005, ¶ [0273]). Boncyk describes that the term “‘point and click’ to refer to the operation of the user capturing imagery of the target object 200 and initiating the transfer of the image data 205 to the identification server 206” (Ex-1005, ¶ [0270]). Therefore, Boncyk teaches that the object is captured and sent to the identification server and can be an object on a television screen. Since Boncyk discloses that the object being detected can be on a television screen with “point and click,” Boncyk teaches that the object could be “a spot on a displayed image on a display,” a “sub-area of the space on the displayed image on the display,” and “the object is one of a plurality of objects in the displayed image on the display,” as set forth in claims 5-9 of the ’812 patent.

227. Darrell and Boncyk disclose and/or render obvious claims 5-9 because Darrell and Boncyk teach a cell phone with a digital camera for taking pictures without restriction of what pictures to take. Thus, Darrell's and Boncyk's cell phones can take a picture of a display, such as TV screen, a computer screen, and/or a camera screen. Moreover, a picture of a display includes a spot, subarea, an object in the display, and an object that is a spot on the displayed image on the display.

228. Darrell teaches that the "system allows users to stay where they are and point at a remote place in sight simply by taking photographs" (Ex-1004, ¶ [0023]). The remote place can include a display.

229. Below is an image of a modern day television screen, and the contents of what is displayed on the screen is indicative of the type of content that could have been presented on a television screen prior to September 2005. Prior to September 2005, in view of Darrell and/or Boncyk, a user could have taken a picture of any television screen to capture whatever was presented on the television screen at that time. The resulting picture could include an object that is a spot on the display (purple), a sub-area of the space on an image on a display (green), and/or one of many objects in a display (red).

Example of “Objects” in a Picture of a TV Display



6. Claim 10 is Obvious over Boncyk and Darrell

230. Boncyk discloses that “the target object 100 could be, for example, beverage [*sic*] can, a music CD box, a DVD video box, a magazine advertisement, a poster, a theatre, a store, a building, a car, or any other object that user [*sic*] is interested in or wishes to interact with” (Ex-1005, ¶ [0127]). Boncyk discloses that its system is “[c]apable of discriminating the target object 100 from any foreground or background objects or image information” (Ex-1005, ¶ [0154]). Finally, Boncyk discloses an algorithm that can “find an object or portion thereof in an input image, given a database of multiple objects with multiple views (from different angles) of each object” (Ex-1005, ¶ [0200]). Therefore, Boncyk discloses that the object is “an object in space, or near an object in space,” as required by claim 10 of the ‘812 patent.

231. Darrell discloses claim 10.

7. Claim 11 is Obvious over Boncyk and Darrell

232. Boncyk discloses that “the target object 100 could be, for example, beverage can, a music CD box, a DVD video box, a magazine advertisement, a poster, a theatre, a store, a building, a car, or any other object that user is interested in or wishes to interact with” (Ex-1005, ¶ [0127]). Boncyk discloses that its system is capable of “recognizing both 2-dimensional (i.e., flat) and 3-dimensional objects” and “Capable of discriminating the target object 100 from any foreground or background objects or image information” (Ex-1005, ¶¶ [0152]-[0154]). Finally, Boncyk discloses an algorithm that can “find an object or portion thereof in an input image, given a database of multiple objects with multiple views (from different angles) of each object.” (Ex-1005, ¶ [0200]). Since Boncyk teaches that the object could be a 3-dimensional object, such as a building or car, and that its system can discriminate the object from objects in the foreground, the target object could be a subarea of a surface due at least in part to other objects being located in the foreground of the target object. Therefore, Boncyk discloses that the object is “a subarea of a surface in space,” as required by claim 11 of the ‘812 patent.

233. Darrell discloses claim 11.

8. Claim 12 is Obvious over Boncyk and Darrell

234. Boncyk discloses that “the target object 100 could be, for example, beverage [*sic*] can, a music CD box, a DVD video box, a magazine advertisement, a

poster, a theatre, a store, a building, a car, or any other object that user [*sic*] is interested in or wishes to interact with” (Ex-1005, ¶ [0127]). Boncyk discloses that its system is “Capable of discriminating the target object 100 from any foreground or background objects or image information” (Ex-1005, ¶ [0154]). Finally, Boncyk discloses an algorithm that can “find an object or portion thereof in an input image, given a database of multiple objects with multiple views (from different angles) of each object” (Ex-1005, ¶ [0200]). Therefore, Boncyk discloses that the object is “one of a plurality of objects in space,” as required by claim 12 of the ‘812 patent.

235. Darrell discloses claim 12.

C. Obviousness over Boncyk, Darrell, and the Level of Ordinary Skill in the Art

236. Claims 5-9 are further obvious over Boncyk, Darrell, and the level of ordinary skill in the art because claims 5-9 identify what is in a picture taken with a “digital camera.” There is nothing unique about what is in a picture. For long periods of time prior to 2005, pictures have been taken of almost any objects including displays, such as television displays and computer displays. Prior to September 2005, pictures of displays (*e.g.*, television or computer screens) included an “object” on the “displayed image” on the display (*e.g.*, soccer field). Moreover, pictures of a “displayed image” on the display included an “object that is a “spot” (*e.g.*, soccer ball) on the “displayed image,” a “sub-area on the space of the displayed

image” (e.g., soccer penalty box), and “one of a plurality of objects in the displayed image” (e.g., soccer goal). Thus, claims 5-9 would have been obvious in view of the teachings of Boncyk and Darrell, as described above, and the level of ordinary skill in the art.

D. Obviousness over Boncyk, Darrell, and Fukuda

237. The ’812 patent claims do not recite a laser or a reticle. I was asked to express my opinion if the claimed pointing and identification device was construed to include a laser or a reticle. It is my opinion that Boncyk and Darrell would still render claims 1-12 obvious because Darrell discloses a reticle as described above. I was also asked to provide my opinion if Darrell is considered not to disclose a reticle (which I believe is incorrect). If Darrell is considered not to disclose a reticle, it is my opinion that claims 1-12 are obvious over Boncyk, in view of Darrell, and further in view of Fukuda.

238. Boncyk and Darrell teaches the limitations of claims 1-12 above. Fukuda discloses a reticle.

239. To the extent that Pointwise contends that the claimed “pointing and identification” device requires a laser or a reticle and that Boncyk and Darrell do not disclose a reticle, the inclusion of a reticle into the combined Boncyk and Darrell system would have been obvious in view of Fukuda.

240. Fukuda discloses reticles in the form of the “rectangular area” and the “pointer” described and illustrated in Fukuda (Ex-1006, Abstract; ¶¶ [0024], [0061], [0070], [0071], FIG. 7 and 9). Fukuda describes a “digital camera” that “has a display such as an LCD” and an “electronic viewfinder” or “EVF” that can display “a viewfinder image (preview image) of the subject before photographing” (Ex-1006, ¶¶ [0024], [0028]).

241. Fukuda describes that a rectangular area, or reticle, is presented in the viewfinder image and can be moved to allow the user to move or adjust the point of focus of the camera (Ex-1006, ¶¶ [0070], [0071], FIG. 7, FIG. 9). Fukuda describes that “a pointer” may be displayed in the viewfinder display and that the user may “move the pointer PT up, down, right and left in the screen” to set a designation position in the screen (Ex-1006, ¶ [0070]).

242. Fukuda illustrates in Figures 7 and 9 (reproduced below) and describes that “a focus point may be associated with the pointer PT” and that “the in-focus area” is defined by “a rectangular area which is set around the pointer PT” and that the rectangular area can move “interlockingly with the pointer” to allow for “focusing control” to be performed by the user (Ex-1006, ¶ [0071], FIG. 9). Therefore, Fukuda discloses that a user may move a pointer around a viewfinder display in order to move or adjust a reticle, defined by the “rectangular area,” to adjust the point of focus of the camera prior to taking a picture.

Figure 7 of Fukuda

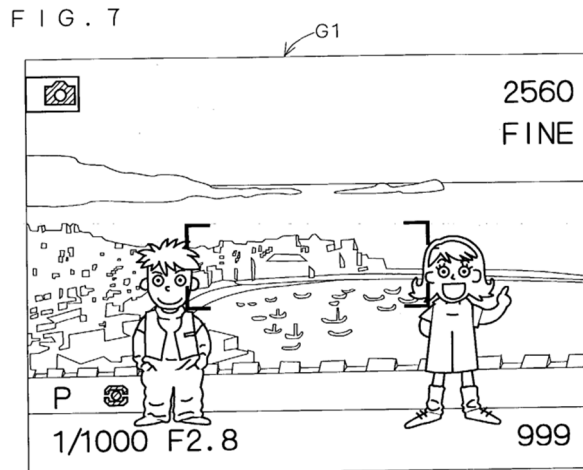
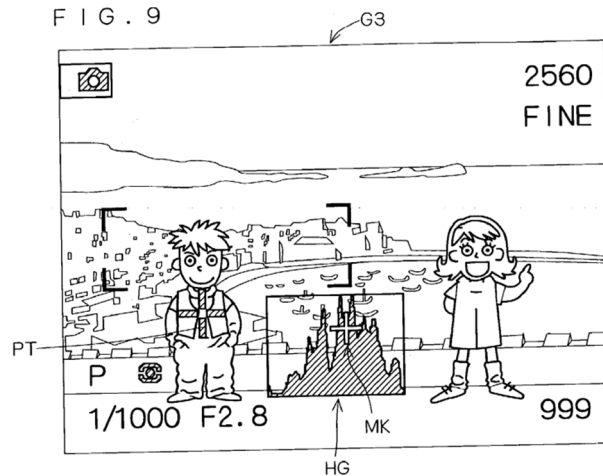


Figure 9 of Fukuda



243. It would have been obvious to a POSITA to combine Boncyk and Darrell with Fukuda to form the claimed invention. Boncyk, Darrell, and Fukuda disclose all limitations of the claims. Thus, there is no difference between the claim and the prior art.

244. Fukuda discloses the benefits of using reticles in digital cameras, since it recognizes that the reticle can be used by the user to define and illustrate the area of focus of an image before the user takes a picture. Since Boncyk and Darrell describe a handheld device that includes a similar display screen and a camera used to capture an image of an object, it would have been obvious to incorporate the reticle, or “rectangular area”, taught by Fukuda into the Boncyk and Darrell handheld devices so the user of the Boncyk and Darrell devices could more accurately focus the image on the object of interest.

245. It would have been obvious and routine to combine Fukuda's pointer and rectangular reticle with Boncyk's and Darrell's devices because the systems include a display screen and buttons that allow for the operability of the pointer and rectangular reticle by the user.

Figure 3 of Darrell

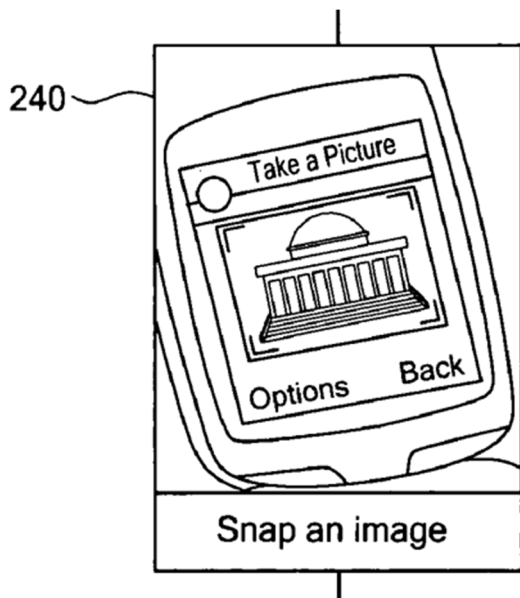
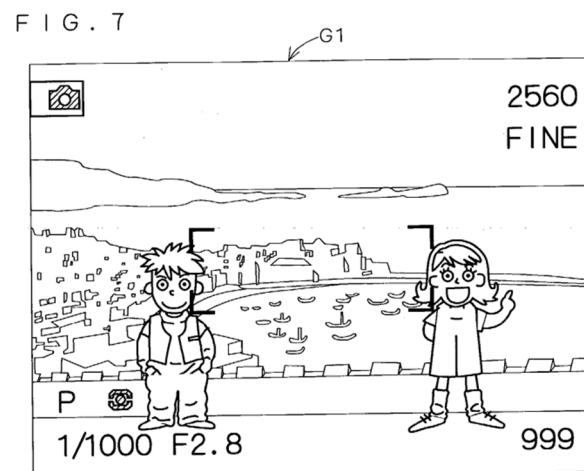


Figure 7 of Fukuda



246. Moreover, Boncyk, Darrell, and Fukuda are in the same technical field. Combining the references leads to the predictable result of sending a list of target object information rather than a single target object information. There is no technical difficulty in sending more than one target object information.

E. Secondary Considerations

247. I have not been asked to consider any secondary conditions of non-obviousness in rendering this opinion.

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

Executed on: 3/28/2025



Oded Gottesman