



(43) International Publication Date
29 September 2016 (29.09.2016)

- | | |
|--|--|
| <p>(51) International Patent Classification:
 G06K 9/00 (2006.01)</p> <p>(21) International Application Number:
 PCT/US2016/021780</p> <p>(22) International Filing Date:
 10 March 2016 (10.03.2016)</p> <p>(25) Filing Language:
 English</p> <p>(26) Publication Language:
 English</p> <p>(30) Priority Data:
 14/664,099 20 March 2015 (20.03.2015) US</p> <p>(71) Applicant: GOOGLE INC. [US/US]; 1600 Amphitheatre Parkway, Mountain View, CA 94043 (US).</p> <p>(72) Inventors: RAYNAUD, Daniel; 1600 Amphitheatre Parkway, Mountain View, CA 94043 (US). BLUNTSCHLI, Boris; Brandschenkestrasse 110, 8002 Zurich (CH). COTTING, Daniel; Brandschenkestrasse 110, 8002 Zurich (CH).</p> | <p>(74) Agents: RICHER, Natalie S. et al.; Lerner, David, Littenberg, Krumholz & Mentlik, LLP, 600 South Avenue West, Westfield, NJ 07090 (US).</p> <p>(81) Designated States <i>(unless otherwise indicated, for every kind of national protection available):</i> AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.</p> <p>(84) Designated States <i>(unless otherwise indicated, for every kind of regional protection available):</i> ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK,</p> |
|--|--|

[Continued on next page]

- (54) Title:** DETECTING THE LOCATION OF A MOBILE DEVICE BASED ON SEMANTIC INDICATORS

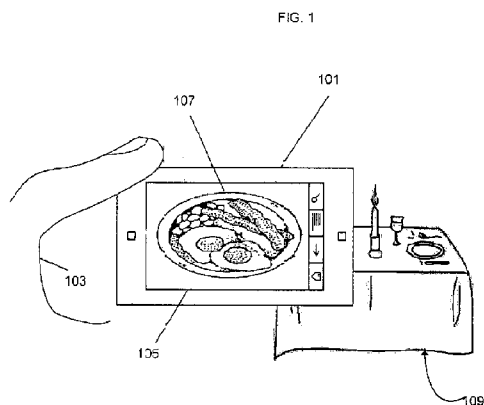


FIG. 1

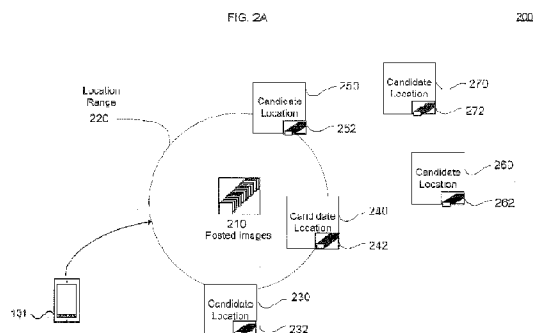


FIG. 2A

(57) Abstract: A system and computer implemented method for detecting the location of a mobile device using semantic indicators is provided. The method includes receiving, using one or more processors, a plurality of images (210) captured by a mobile device at an area. The area is associated with a set of candidate locations (230-270). Using the one or more processors, one or more feature indicators associated with the plurality of images (210) are detected. These feature indicators include semantic features related to the area. The semantic features are compared with a plurality of stored location features (232-272) for the set of candidate locations. In accordance with the comparison, a location from the set of candidate locations is selected to identify an estimated position of the mobile device.



SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

Published:

— *with international search report (Art. 21(3))*

DETECTING THE LOCATION OF A MOBILE DEVICE BASED ON SEMANTIC INDICATORS

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] The present application is a continuation of U.S. Patent Application No. 14/664,099, filed March 20, 2015, the disclosure of which is hereby incorporated herein by reference.

BACKGROUND

[0002] Modern mobile devices are equipped with location-based features. These devices use signals from Global Positioning Service (GPS) satellites to identify a location, determine a direction of motion, and other functions. Typically, in locations where the GPS satellite signals are weak (e.g., indoors) GPS may not function well or at all. To provide indoor navigation, some devices may use other information such as wireless network signals (WiFi), Bluetooth signals, cellular, compass data, and accelerometer data as well as existing floor plans and pre-generated databases or indices of measurement. However, in certain situations, it may be difficult to ascertain the location of the mobile device, such as in high density commercial areas where there are a lot of internal locations for the device to be located within.

BRIEF SUMMARY

[0003] Aspects of the disclosure may be advantageous for providing a computer implemented method of detecting the location of a mobile device based on captured semantic indicators. In one example, the method includes receiving, using one or more processors, a plurality of images captured by a mobile device at an area, the area being associated with a set of candidate locations. The one or more processors are further used to detect one or more feature indicators associated with the plurality of images, the feature indicators including semantic features related to the area, and compare the semantic features with a plurality of stored location features for the set of candidate locations. A location is selected, using the one or more processors, from the set of candidate locations according to the comparison to identify an estimated position of the mobile device. In one example, position information associated with the mobile device overlaps with position information associated with the plurality of stored location features for the set of candidate locations. In some examples, the comparing further includes determining whether the semantic features correspond to stored location features for a given location in the set. Such semantic features may include visual features of a food image and the selected location is associated with a menu corresponding to

the food image, audio data of one or more objects at the selected location, a logo associated with the selected location, a time of day and the selected location is associated with a business open at that time of day, textures of objects located within the selected location, sensor data associated with the plurality of images and the sensor data is captured by the mobile device, etc.

[0004] The disclosure further provides a non-transitory computer readable medium storing instructions that, when executed by one or more processors, cause the one or more processors to perform a method. Such method includes receiving a plurality of images captured by a mobile device at an area, the area being associated with a set of candidate locations, detecting one or more feature indicators associated with the plurality of images, the feature indicators including semantic features related to the area, comparing the semantic features with a plurality of stored location features for the set of candidate locations, and selecting a location from the set of candidate locations according to the comparison to identify an estimated position of the mobile device.

[0005] The disclosure further provides a system comprising a memory storing a plurality of location features for a set of candidate locations, and one or more processors in communication with the memory. The one or more processors are configured to receive a plurality of images captured by a mobile device at an area, the area being associated with a set of candidate locations, detect one or more feature indicators associated with the plurality of images, the feature indicators including semantic features related to the area, compare the semantic features with a plurality of location features stored in memory, and select a location from the set of candidate locations according to the comparison to identify an estimated position of the mobile device.

BRIEF DESCRIPTION OF THE DRAWINGS

[0006] FIG. 1 is an image capture process in accordance with aspects of the disclosure.

[0007] FIG. 2A is a diagram of an area including a number of candidate locations in accordance with aspects of the disclosure.

[0008] FIG. 2B illustrates an image processing system in accordance with aspects of the disclosure.

[0009] FIG. 3 is an example of a system in accordance with aspects of the disclosure.

[0010] FIG. 4 is another example of a system in accordance with aspects of the disclosure.

[0011] FIG. 5 is an example of semantic indicators in accordance with aspects of the disclosure.

[0012] FIG. 6 is another example of semantic indicators in accordance with aspects of the disclosure.

[0013] FIG. 7 is yet another example of semantic indicators in accordance with aspects of the disclosure.

[0014] FIG. 8 is a flow diagram depicting an example of a method in accordance with aspects of the disclosure.

DETAILED DESCRIPTION

[0015] The present disclosure generally relates to techniques for detecting a location of a mobile computing device based on certain semantic indicators captured by the device. The techniques described herein may be useful in collecting semantic indicators as well as other type of attributes associated with a certain area and inferring a particular location within the area that indicates where the semantic indicators were collected. In some embodiments, a system may be provided that can analyze image pixels, audio data, as well as other captured information from still and/or video images in order to detect semantics features that can be used to determine the location of a mobile device.

[0016] To determine the location of the user device, the system may detect or otherwise extract the one or more feature indicators associated with a plurality of images captured by the user device. These feature indicators may include semantic features related to a certain location at a particular area, such as the location of a store inside of a shopping mall. The semantic features are compared to a plurality of stored location features for a set of candidate locations known to be within range of the particular area. In accordance with the comparison, a location is selected from the set of candidate locations within range of the particular area to identify an estimated position of the mobile device.

[0017] In one illustrative example, a user device may "check-in" to a social networking site by using a food image taken at a restaurant. In this example, the system may detect the type of food in the photo in order to determine that the user device is at a particular type of restaurant in the area which serves this type of food. In response to receiving the food image as well as other kinds of feature indicators, the system may compare these feature indicators to indicators associated with different food images stored in a database. For example, the stored food images may be retrieved from different food menus for various restaurants. A match to a particular location may be determined when a number of matching features

between the detected indicators and the stored indicators meet a predetermined threshold level.

[0018] Turning to a second example, the user device may capture a video inside of a location. Based on the detected feature indicators related to the type of sounds of various objects in the video, such the sound of an espresso machine or loud music, the system may estimate that device is in a particular café. In some embodiments, a timestamp may be captured indicating a time when the video was taken. Based on this additional semantic information, the system may further determine that the device is at a particular café open at that time.

[0019] In other examples, the system may detect other kinds of feature indicators to determine the location of the user device, such as detecting textures of objects in the images, or detecting logos of a particular business, school/university, or other type of organizations associated with a specific location. For example, the textures may include bright colors or unique color patterns. Still further, many other types of feature indicators may be captured along with the images (for example, by sensors associated with the mobile device) and used to further narrow down the position of the user device. Examples of such other feature indicators include motion detection indicators (e.g., indicating dancing), barometric indicators, humidity indicators, radiation indicators, etc.

[0020] For situations in which the subject matter described herein collects information about users, or may make use of user-related information, the users may be provided with an opportunity to control whether programs or features collect user information (e.g., information about a user's social network, social actions or activities, profession, preferences or a user's current location), or to control whether and/or how to receive information that may be of interest to the user. In addition, certain data may be treated in one or more ways before it is stored or used, so that personally identifiable information is removed. For example, a user's identity may be treated so that no personally identifiable information can be determined for the user, or a user's current and historical location may be generalized where location information is obtained (such as to a city, ZIP code or state level), so that a particular location of a user cannot be determined. A user may also be provided with the opportunity to decide whether, and to control how, such information (such as pictures and location history) is used and stored on a user's device and by servers with which the device communicates.

[0021] Aspects, features and advantages of the disclosure will be appreciated when considered with reference to the following description of embodiments and accompanying figures. The same reference numbers in different drawings may identify the same or similar

elements. Furthermore, the following description is not limiting; the scope of the present technology is defined by the appended claims and equivalents. While certain processes in accordance with example embodiments are shown in the figures as occurring in a linear fashion, this is not a requirement unless expressly stated herein. Different processes may be performed in a different order or concurrently. Steps may also be added or omitted unless otherwise stated.

[0022] FIG. 1 is an image capture process 100. For example, a user 103 may activate a camera associated with a mobile computing device 101, such as a mobile phone, for capturing images at different locations. In this regard, the mobile computing device 101 may include a camera (not shown) for capturing an image 105 taken at certain location. A typical captured image, such as image 105, may include features of objects present at the location when the images were taken. For example, as shown in FIG 1, image 105 includes features associated with a food image 107 that may have been taken at a particular restaurant 109. Along with the food image 107 certain types of metadata may be captured by the mobile computing device 101. For example, this metadata may include data from sensors, such as an accelerometer, gyroscope, motion and audio sensors or other types of sensors of the mobile computing device 101, and stored in the memory of the device 101 as well as with the captured images.

[0023] Once the camera has captured the image information (e.g., images and other metadata), this information may be stored locally at the mobile computing device 101 and/or transmitted to a server. In some aspects, user 103 may utilize an application associated with the mobile computing device 101 for uploading images to a social media web site or an image hosting service. In some embodiments, the user may have to install the application and/or select a service in order to obtain the benefits of the techniques described herein.

[0024] The application may be downloaded onto mobile computing device 101. For example, the user 103 may elect to download the application from a service associated with an online server. The mobile computing device 101 may transmit a request for the application over a network and in response, receive the application from the service. The application may be installed locally at the mobile computing device 101. Alternatively, the application can be stored at the service and may be accessed through the mobile computing device 101, for example, via a mobile web browser. By using the application, user 103 may keep the image information private, or publically post or otherwise share images online. In some embodiments, semantic indicators associated with the image information can be used to

estimate a certain location of the mobile device 101 relatively near to where the images were captured.

[0025] FIG 2A is a diagram of an area 200 including a number of candidate locations 230-270. In some examples, the mobile computing device 101 may include geographic positioning features (not shown) that may indicate the user was in a certain area at a particular time when images 210 were captured. For example, the user may have parked their car in the area 200 sometime before taking images 210. Once an initial location is identified, the mobile computing device 101 may store this area information. For example, the area information can be stored at set time intervals (e.g., every 30 seconds) or on demand by the user. The area location information may include one or more data fields, which include positioning information such as a global coordinate (e.g., latitude/longitude) and a timestamp indicating a time when the user of device 101 was at that coordinate.

[0026] To estimate an actual location of the mobile computing device 101 in the area 200, several blocks of potential candidate locations 230-270 may be selected. For example, a system may identify a number of candidate locations located in the area 200. In this example, system may use additional information to refine the selection of viable candidate locations. For example, the system may select a subset of candidate locations that are within a predetermined range, such as location range 220. In this example, the system may select candidate locations 230-250 as possible locations of the mobile device 101 when posted images 210 associated with these candidate locations are within location range 220. This may be determined by comparing location data stored with the mobile computing device 101 with coordinates of the candidate locations 230-270 for any overlap. In this regard, the predetermined range may be expressed in a unit of measurement, such as feet, meters, inches, miles, etc., that is within reach of the mobile computing device 101.

[0027] Each block of candidate locations 230-270 may be associated with a particular set of stored location features 232-272 that are related to objects at the respective candidate locations 230-270. In some example, the system may compare semantic features from the posted images 210 with the stored location features 232-272 in order to find a closest matching candidate location. For example, a match to a particular candidate location may be determined when a number of matching semantic features between the posted images 210 and the stored location features 232-272 meet a predetermined threshold level. Further aspects of this system are discussed below with respect to FIG. 2B.

[0028] FIG 2B illustrates an image processing system 280 for receiving a number of images 210, analyzing the imagery for semantic feature information and detecting a particular

location depicted in the images 308. As shown, the system 300 may have a number of modules to identify location features associated with the images 210 and compare those features with stored semantic features of several candidate locations. For example, the modules may include a feature indicator detector 282 to identify location features associated with an input image and a stored semantic Features filter 284 to compare features associated with a set of candidate locations to the location features. In some examples, other modules (not shown) may also be employed.

[0029] The modules may employ different techniques for analyzing images. For instance, the location feature detector 282 may include software for analyzing pixels of objects depicted in an image. The modules are shown as being connected to one or more database 218. This database may store semantic features associated with a location that may correspond to where the images 308 were captured. These modules may operate in series or parallel fashion. Another module 286 of system 300 may use the results of other modules 282 and 284, for example, to detect the location associated with image 210.

[0030] FIG. 3 is an example of a system 300 including computing devices for performing aspects of the present disclosure. As shown, system 300 depicts various computing devices that can be used alone or in a networked configuration. For example, this figure illustrates a computer network 395 having a plurality of computers 310 and 320 as well as other types of mobile devices like mobile computing device 101 from FIG. 1, a PDA 330, laptop/netbook 340 and tablet computing device 350. These various devices may be interconnected via a local bus or direct connection 318 and/or may be coupled via a communications network 395 such as a LAN, WAN, the Internet, etc. and which may be wired or wireless.

[0031] Each computing device may include, for example, user input devices such as a keyboard 312 and mouse 314 and/or various other types of input devices such as pen-inputs, joysticks, buttons, touch screens, etc., as well as a display 316, which could include, for instance, a CRT, LCD, plasma screen monitor, TV, projector, etc. Each computer 310 and 320 may be a personal computer, application server, etc. By way of example only, computer 310 may be a personal computer while computer 320 may be a server. Databases, such as database 288, are accessible to one or more of the servers or other devices of system 300.

[0032] FIG. 4 is another example of a system 400, which may be used to detect the location of a mobile device based on captured semantic indicators as described herein. As shown, the system 400 includes a number of client devices, such as client devices 101 and 410, capable of communicating with a server 420 over network 395. The server 420 may

include one or more processors 422, memory 424, and other components typically present in general purpose computers.

[0033] Memory 424 of server 420 may store information that is accessible by the processors 422, including instructions 426 that may be executed by the processors 422, and data 428. The memory 424 may be of a type of memory operative to store information accessible by the processors 422, including a non-transitory computer-readable medium, or other medium that stores data that may be read with the aid of an electronic device, such as a hard-drive, memory card, read-only memory ("ROM"), random access memory ("RAM"), digital versatile disc ("DVD") or other optical disks, as well as other write-capable and read-only memories. The subject matter disclosed herein may include different combinations of the foregoing, whereby different portions of the instructions 426 and data 428 are stored on different types of media.

[0034] Although FIG. 4 functionally illustrates the processors 422 and memory 424 as being within the same block, the processors 422 and memory 424 may actually include multiple processors and memories that may or may not be stored within the same physical housing. For example, some of the instructions 426 and data 428 may be stored on removable CD-ROM and others within a read only computer chip. Some or all of the instructions and data may be stored in a location physically remote from, yet still accessible by, the processors 422. Similarly, the processors 422 may actually comprise a collection of processors, which may or may not operate in parallel.

[0035] Data 918 may be retrieved, stored or modified by processors 422 in accordance with the instructions 426. For instance, although the present disclosure is not limited by a particular data structure, the data 428 may be stored in computer registers, in a relational database as a table having a plurality of different fields and records, XML documents, or flat files. The data 428 may also be formatted in a computer-readable format such as, but not limited to, binary values, ASCII or Unicode. By further way of example only, the data 912 may be stored as bitmaps comprised of pixels that are stored in compressed or uncompressed, or various image formats (e.g., JPEG), vector-based formats (e.g., SVG) or computer instructions for drawing graphics. Moreover, the data 428 may comprise information sufficient to identify the relevant information, such as numbers, descriptive text, proprietary codes, pointers, references to data stored in other memories (including other network locations) or information that is used by a function to calculate the relevant data. In this example, the data 428 includes database 427 for storing location information 429 relevant to a number of different locations.

[0036] Database 427 may store the location information 429 in association with other information, such as images, semantic features, location descriptions, etc. In this regard, as information such as an image is received from the client device 410, location information such as geographic coordinates, map information, or the name of a business, city, or the like may be returned. While the database 427 is shown within the server 420 in Fig. 4, it should be understood that the server may additionally or alternatively be located remotely from the server 420. For example, the server 420 may access multiple databases in various locations. In some examples, the different databases may be store different types of information.

[0037] The server 420 may be at one node of network 395 and capable of directly and indirectly communicating with other nodes of the network 395. For example, the server 420 may include a web server that may be capable of communicating with client devices 101 and 410 via network 395 such that it uses the network 395 to transmit and display information to a user on a display 413 of the client device 410. Server 420 may also include a plurality of computers, e.g., a load balanced server farm, that exchange information with different nodes of a network for the purpose of receiving, processing and transmitting data to client devices 101 and 410. In this instance, the client devices 101 and 410 will typically still be at different nodes of the network 395 than the computers comprising server 420.

[0038] Network 395, and intervening nodes, may include various configurations and protocols including the Internet, World Wide Web, intranets, virtual private networks, wide area networks, local networks, private networks using communication protocols proprietary to one or more companies, Ethernet, WiFi (e.g., 802.11, 802.11b, g, n, or other such standards), HTTP, and various combinations of the foregoing. Such communication may be facilitated by a device capable of transmitting data to and from other computers, such as modems (e.g., dial-up, cable or fiber optic) and wireless interfaces.

[0039] Although certain advantages are obtained when information is transmitted or received as noted above, other aspects of the subject matter disclosed herein are not limited to a particular manner of transmission of information. For example, in some aspects, information may be sent via a medium such as a disk, tape or CD ROM. Yet further, although some functions are indicated as taking place on a single server having a single processor, various aspects may be implemented by a plurality of servers, for example, communicating information to client devices 101 and 410 over network 395.

[0040] Each client device 101 and 410 may be configured similarly to the server 420, with one or more processors 412, memory 414, instructions 416, data 418 and all of the internal components normally found in a personal computer. By way of example only, the

client device 930 may include a central processing unit (CPU), display device 413 (for example, a monitor having a screen, a projector, a touch-screen, a small LCD screen, a television, or another device such as an electrical device that is operable to display information processed by the processors 412), CD ROM, hard drive, user input devices 939, such as a keyboard, mouse, touch screen or microphone, speakers, modem and/or network interface device (telephone, cable or otherwise) and all of the components used for connecting these elements to one another.

[0041] Client device 410 may be a computing device. For example, client device 410 may be a laptop computer, a netbook, a desktop computer, and a portable personal computer such as a wireless-enabled PDA, a tablet PC or another type of computing device capable of obtaining information via a network like the Internet. Although aspects of the disclosure generally relate to a single client device 410, the client device 410 may be implemented as multiple devices with both portable and non-portable components (e.g., software executing on a rack-mounted server with an interface for gathering location information).

[0042] Although the client device 410 may include a full-sized personal computer, the subject matter of the present disclosure may also be used in connection with mobile devices capable of wirelessly exchanging data. For example, client device 410 may be a wireless-enabled mobile device, such as a Smartphone, or an Internet-capable cellular phone. In either regard, the user may input information using a small keyboard, a keypad, a touch screen or other means of user input. In various aspects, the client devices and computers described herein may comprise a device capable of processing instructions and transmitting data to and from humans and other devices and computers.

[0043] Sensors 415 may include one or more of a variety of sensors. For example, thermometers, barometers, motion detectors, heat sensors, lasers, light sensors, radiation sensors, microphones, etc. may be used alone or in combination to detect various objects, information, or activities. As one example, a motion detector may detect motion at a particular rate and/or direction which suggests people walking by. A relatively (for the candidate geography) low level of humidity and temperature may also be detected (possibly implying the use of air conditioning), and music and/or crowd noise at a particular decibel level may be detected by the microphone. In combination, such information may indicate that the client device 410 is at an indoor shopping mall.

[0044] Image capture module 417 may be a camera, camcorder, or any other device capable of capturing still or motion pictures.

[0045] Geographic positioning component 419 may be used to determine the geographic location of the device 410. For example, the position component may include a GPS receiver to determine the device's latitude, longitude and/or altitude position. Other location systems such as laser-based localization systems, inertial-aided GPS, or camera-based localization may also be used to identify the location of the device 410. The location of the device 410 may include an absolute geographical location, such as latitude, longitude, and altitude as well as relative location information, such as location relative to other client devices or other objects.

[0046] Instructions 416 and 426 of the client device 410 and server 420 respectively may be a set of instructions to be executed directly (such as machine code) or indirectly (such as scripts) by the processor. In that regard, the terms "instructions," "steps" and "programs" may be used interchangeably herein. The instructions 416 and 426 may be stored in object code format for direct processing by the processor, or in another computer language including scripts or collections of independent source code modules that are interpreted on demand or compiled in advance. Functions, methods and routines of the instructions are explained in more detail below.

[0047] In order to facilitate operations of system 900 for detecting the location of a mobile device, such as client devices 101 and 410, the server 910 may further include a feature indicator detector 423 for identifying location features associated with an image and a location detector for comparing those identified indicators with the semantic features of various locations stored in database 429. The functionality of these modules can exist in a fewer or greater number of modules than what is shown, with such modules residing at one or more computing devices, which may be geographically dispersed. The modules may be operable in conjunction with the client devices 101 and 410 from which it may receive images captured at locations as well as other semantic information regarding those images. Thereupon, the images and other semantic information may be used to ascertain the location of client devices 101 and 410.

[0048] FIG. 5 is an example of semantic indicators 500. In this example, device 101 may capture an image 105 of food at a restaurant. The system may detect the type of food in the photo in order to determine that the user device is at a particular type of restaurant in the area which serves this type of food. For example, the system may compare the captured image 105 with stored images 501 in database 427. For example, the stored food images 501 may be retrieved from different food menus 515 for various restaurants, or may have been uploaded by restaurant owners or patrons. The stored images 501 may each be associated

with a particular restaurant, type of restaurant, menu, or the like. The stored images 501 may be categorized in any way, such as by type of food, restaurant, etc. In addition to comparing the image 105 and the image 501, the system may compare feature indicators, such as sounds, colors, etc., received from the computing device 101 with stored feature indicators associated with particular restaurants. A match to a particular location may be determined when at least one of a number of objects in the image 105 and a number of detected features received from the device 101 match objects in stored images or stored features to a predetermined threshold level. For example, stored image 510 may be a closest match to the image 105, and the location associated with the stored image 510 may therefore be determined to be the location of the device 101. Such matching may enable a user device to, for example, "check-in" to a social networking site by using a food image taken at a restaurant.

[0049] FIG. 6 is another example of semantic indicators 600. In this example, the user device 101 may capture a video 602 inside of a location. Based on the detected feature indicators related to the type of sounds of various objects in the video, such the sound of an espresso machine 604 or loud music, the system may estimate that the device 101 is in a particular café. Objects in the video 602, such as a logo 608, tables, décor, etc., may also be analyzed against images 501 and other semantic features 612 in the database 427. In some embodiments, a timestamp 609 may be captured indicating a time when the video was taken. Based on this additional semantic information, the system select a match 610 and determine based on the selected match 610 that the device is at a particular café open at that time.

[0050] FIG. 7 is yet another example of semantic indicators 700. In this example, the system may detect other kinds of feature indicators to determine the location of the user device 101, such as textures of objects in the images, colors, patterns, architecture, or the like. For example, the captures image 105 may include a carpet 708 having a particular texture, pattern, and/or color which may be relatively distinctive. Walls 702 in the image 105 may also have a particular color or pattern, and in some instances may include décor 704 which is also distinctive. For example, detected décor in mainly a particular color or combination of colors may indicate a particular business or school/university or other type of organization associated with a specific location. As in the example above, a timestamp 706 may also be used to match against stored feature indicators 712 associated with establishments open during certain hours. Still further, many other types of feature indicators, such as motion detection indicators (e.g., to detect people walking by or some other type of movement) barometric, humidity and radiation indicators, etc., that were captured along with the images (for example, by sensors associated with the mobile device) can be matched against the

stored feature indicators 712 to identify a particular data set 710 and further narrow down the position of the user device.

[0051] To better aid in understanding an example of some of the aspects described above, reference is now made to FIG. 8, which is a flow diagram 800. As previously discussed, the following operations do not have to be performed in the precise order described below. Rather, as mentioned above, various operations can be handled in a different order or simultaneously, and operations may be added or omitted.

[0052] In block 810, one or more images captured by a mobile device at an area associated with a set of candidate locations may be received. For example, a mobile device may be within range of a number of different businesses, landmarks, or other locations. The mobile device may capture images of or within a particular location send the images to a server or database, for example, by uploading the images to a website.

[0053] In block 820, one or more feature indicators associated with the plurality of images may be detected. These feature indicators may include semantic features related to the area. For example, indicators such as objects, colors, textures, or architecture may be detected from the images. In an example where the captured images are video images, the feature indicators may further include sounds, motions, variations in brightness, etc. In some examples, feature indicators captured by other sensors may be detected. For example, indicia such as humidity, temperature, time, radiation, etc. may be captured by one or more sensors coupled to the device 101 and sent to the server or database.

[0054] In block 830, the semantic features may be compared with a plurality of stored features for the set of candidate locations. For example, an image received by the mobile device may be compared with stored images. Elements within the images, such as objects, colors, and textures, may be analyzed in performing the comparison. Additionally, other features, such as sound, temperature, motion, etc., received from the mobile device may be compared with stored features associated with particular locations. In one example, scores may be calculated for possible locations based on the comparison. A location with a higher number of features matching the features received by the mobile device may be assigned a higher score. In some examples, the scores may be weighted based on the matching features. For example, a matching object in an image may receive more points than a matching sound.

[0055] In block 840, a location from the set of locations may be selected according to the comparison. The selected location may identify an estimated position of the mobile device. For example, a best match to the received images and feature indicators may be determined based on the comparison. The location associated with the best match may be identified as

the estimated position of the mobile device. The best match may be selected based on a highest score or any other convention.

[0056] Aspects, features and advantages of the disclosure will be appreciated when considered with reference to the following description of embodiments and accompanying figures. The same reference numbers in different drawings may identify the same or similar elements. Furthermore, the following description is not limiting; the scope of the present technology is defined by the appended claims and equivalents. While certain processes in accordance with example embodiments are shown in the figures as occurring in a linear fashion, this is not a requirement unless expressly stated herein. Different processes may be performed in a different order or concurrently. Steps may also be added or omitted unless otherwise stated.

CLAIMS

1. A computer implemented method, comprising:
receiving, using one or more processors, a plurality of images captured by a mobile device at an area, the area being associated with a set of candidate locations;
detecting, using the one or more processors, one or more feature indicators associated with the plurality of images, the feature indicators including semantic features related to the area;
comparing, using the one or more processors, the semantic features with a plurality of stored location features for the set of candidate locations;
selecting, using the one or more processors, a location from the set of candidate locations based at least in part on the comparison; and
estimating, using the one or more processors, a position of the mobile device based at least in part on the selected location.
2. The method of claim 1, further comprising comparing position information associated with the mobile device with position information associated with the plurality of stored location features for the set of candidate locations.
3. The method of claim 1, wherein the comparing further comprises determining whether the semantic features correspond to stored location features for a given location in the set.
4. The method of claim 1, wherein the semantic features include visual features of a food image and the selected location is associated with a menu corresponding to the food image.
5. The method of claim 1, wherein the semantic features include audio data of one or more objects at the selected location.
6. The method of claim 1, wherein the semantic features include a logo associated with the selected location.
7. The method of claim 1, wherein the semantic features include a time of day and the selected location is associated with a business open at that time of day.

8. The method of claim 1, wherein the semantic features include textures of objects located within the selected location.

9. The method of claim 1, wherein the semantic features include sensor data associated with the plurality of images and the sensor data is captured by the mobile device.

10. A non-transitory computer readable medium storing instructions that, when executed by one or more processors, cause the one or more processors to perform a method, the method comprising:

receiving a plurality of images captured by a mobile device at an area, the area being associated with a set of candidate locations;

detecting one or more feature indicators associated with the plurality of images, the feature indicators including semantic features related to the area;

comparing the semantic features with a plurality of stored location features for the set of candidate locations;

selecting a location from the set of candidate locations based at least in part on the comparison; and

estimating a position of the mobile device based at least in part on the selected location.

11. The non-transitory computer readable medium of claim 10, the method further comprising comparing position information associated with the mobile device with position information associated with the plurality of stored location features for the set of candidate locations.

12. A system comprising:

a memory storing a plurality of location features for a set of candidate locations; and

one or more processors in communication with the memory, the one or more processors being configured to:

receive a plurality of images captured by a mobile device at an area, the area being associated with a set of candidate locations;

detect one or more feature indicators associated with the plurality of images, the feature indicators including semantic features related to the area;

compare the semantic features with a plurality of location features stored in memory;
select a location from the set of candidate locations based at least in part on the comparison; and
estimate a position of the mobile device based at least in part on the selected location.

13. The system of claim 12, wherein the one or more processors are further configured to compare position information associated with the mobile device with position information associated with the plurality of stored location features for the set of candidate locations.

14. The system of claim 12, wherein the comparing further comprises determining whether the semantic features correspond to stored location features for a given location in the set.

15. The system of claim 12, wherein the semantic features include visual features of a food image and the selected location is associated with a menu corresponding to the food image.

16. The system of claim 12, wherein the semantic features include audio data of one or more objects at the selected location.

17. The system of claim 12, wherein the semantic features include a logo associated with the selected location.

18. The system of claim 12, wherein the semantic features include a time of day and the selected location is associated with a business open at that time of day.

19. The system of claim 12, wherein the semantic features include textures of objects located within the selected location.

20. The system of claim 12, wherein the semantic features include sensor data associated with the plurality of images and the sensor data is captured by the mobile device.

100

FIG. 1

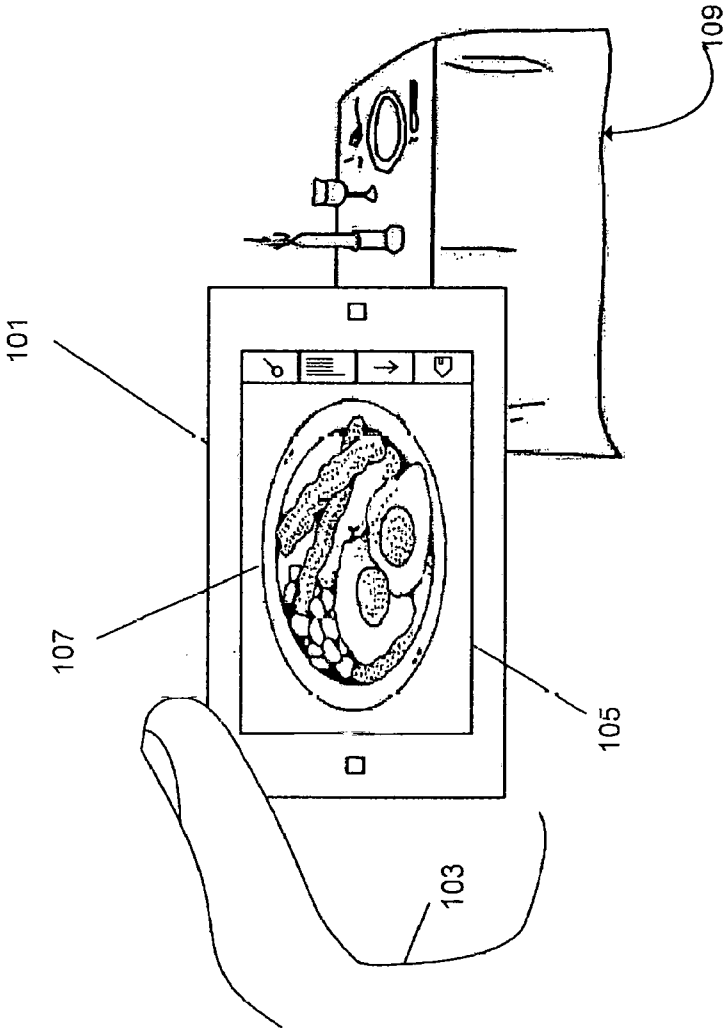


FIG. 2A

200

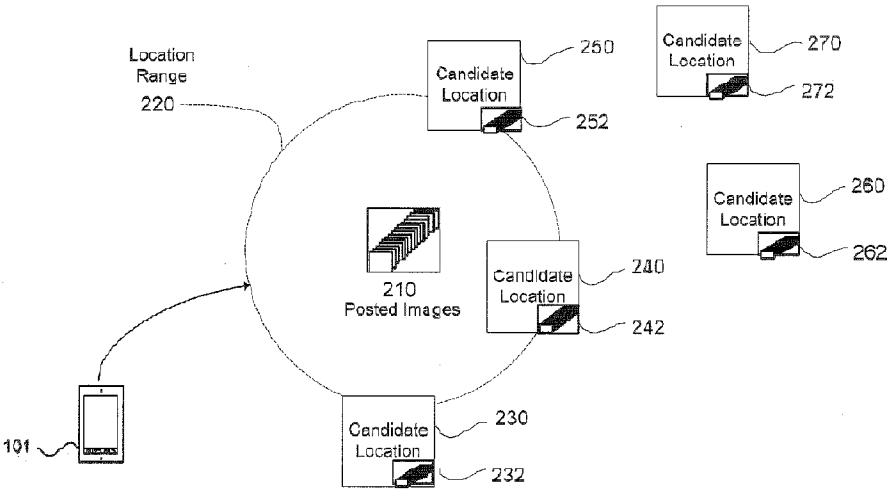


FIG. 2B

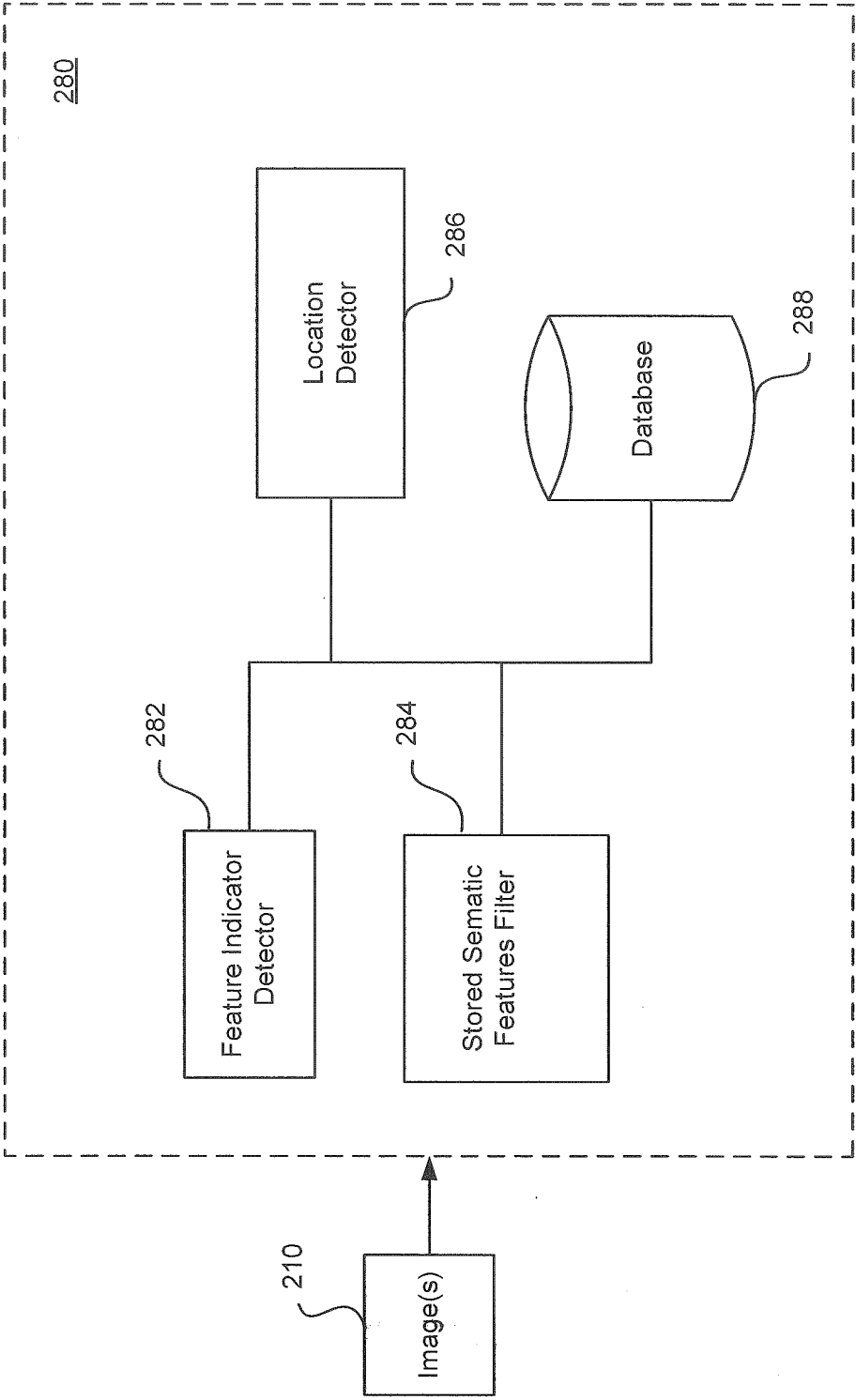
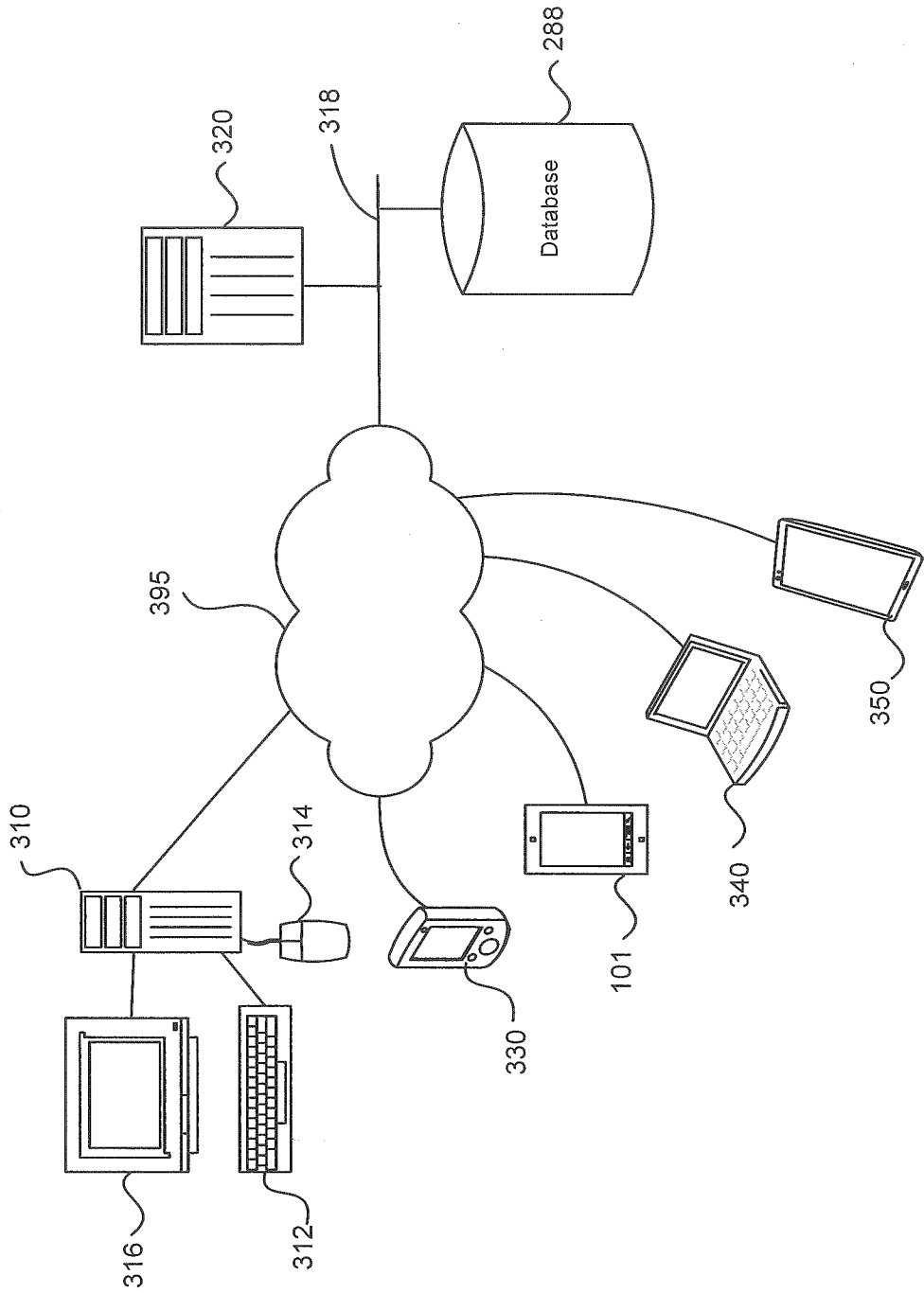
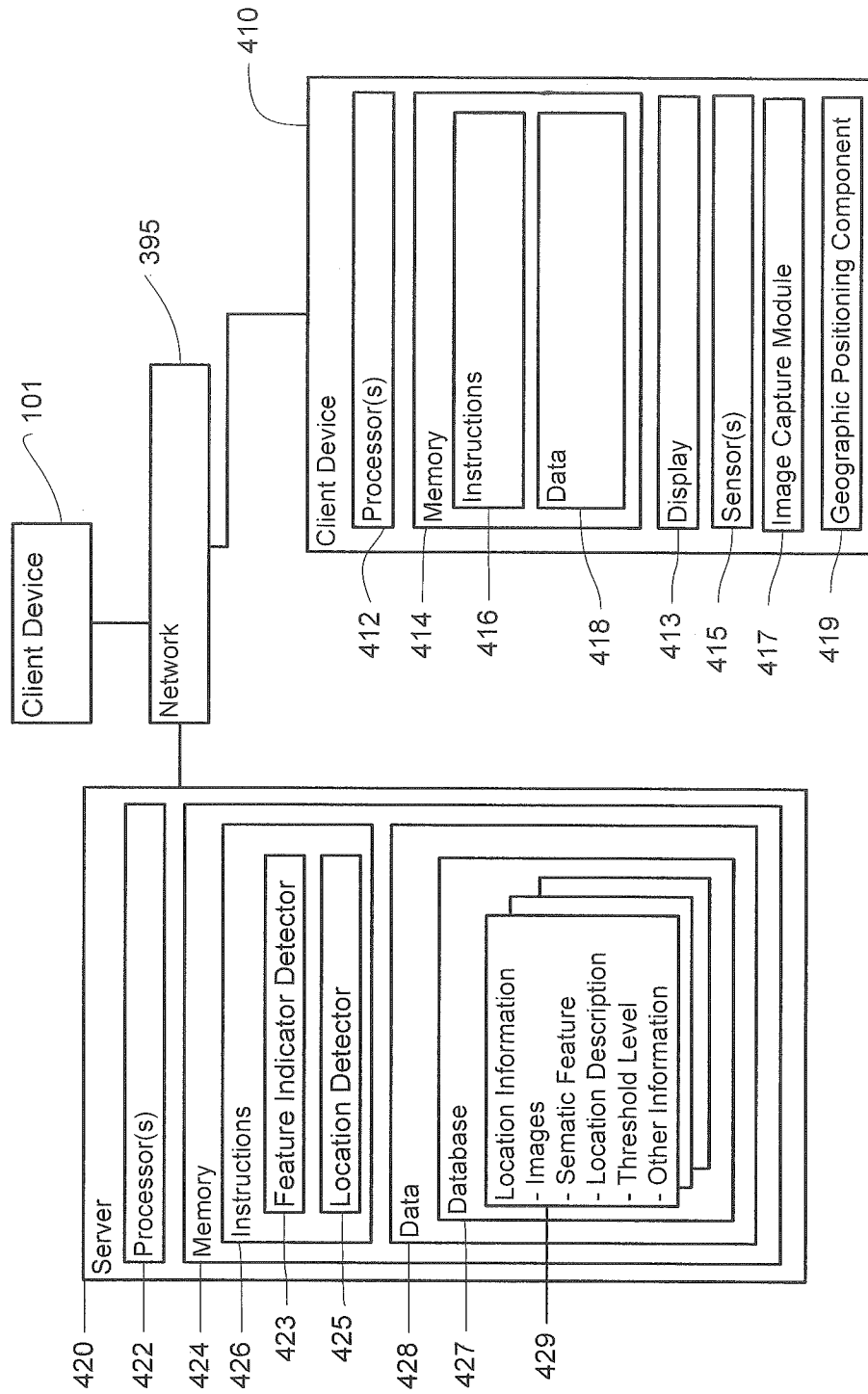


FIG. 3



400

FIG. 4



500

FIG. 5

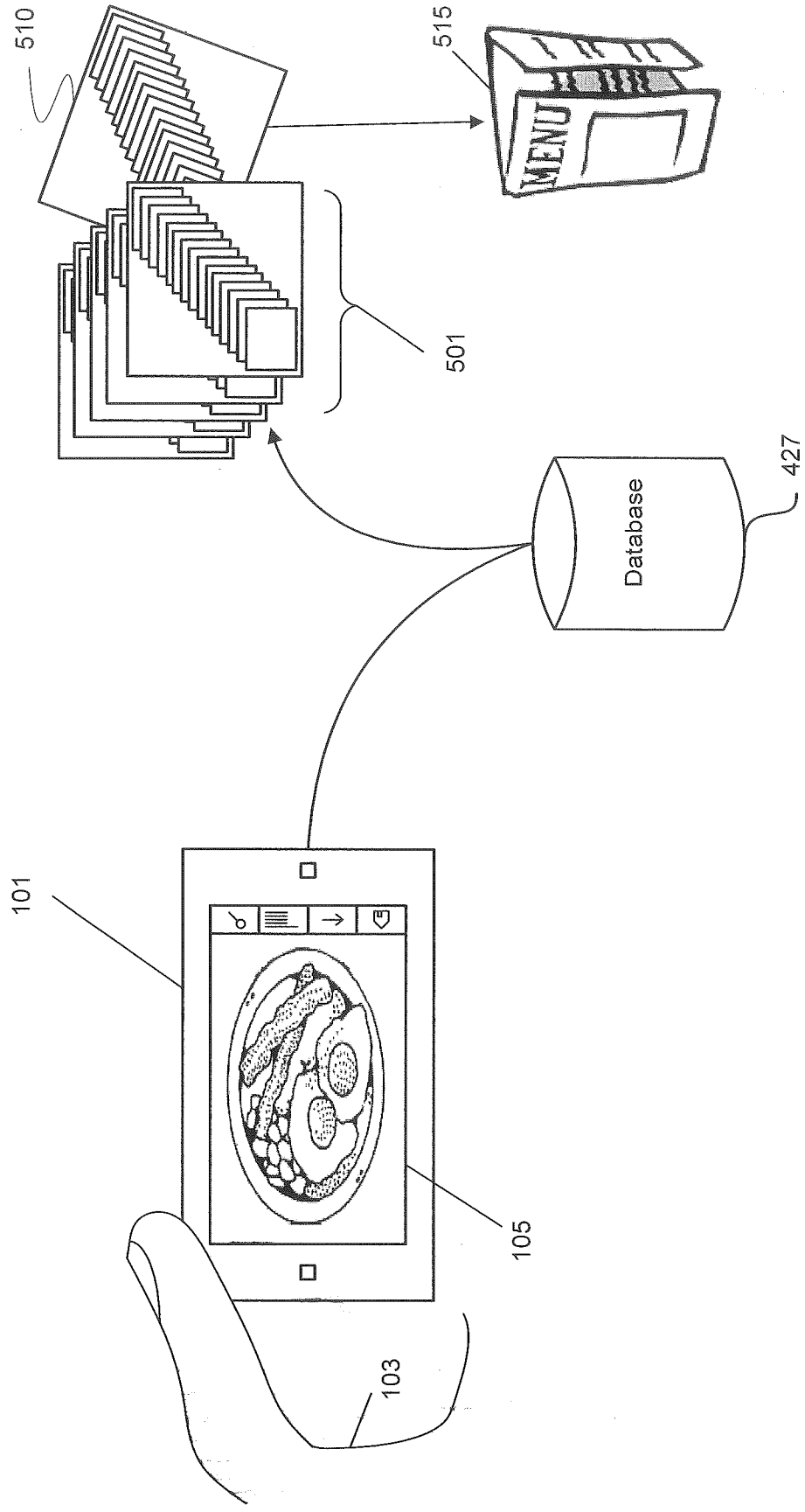
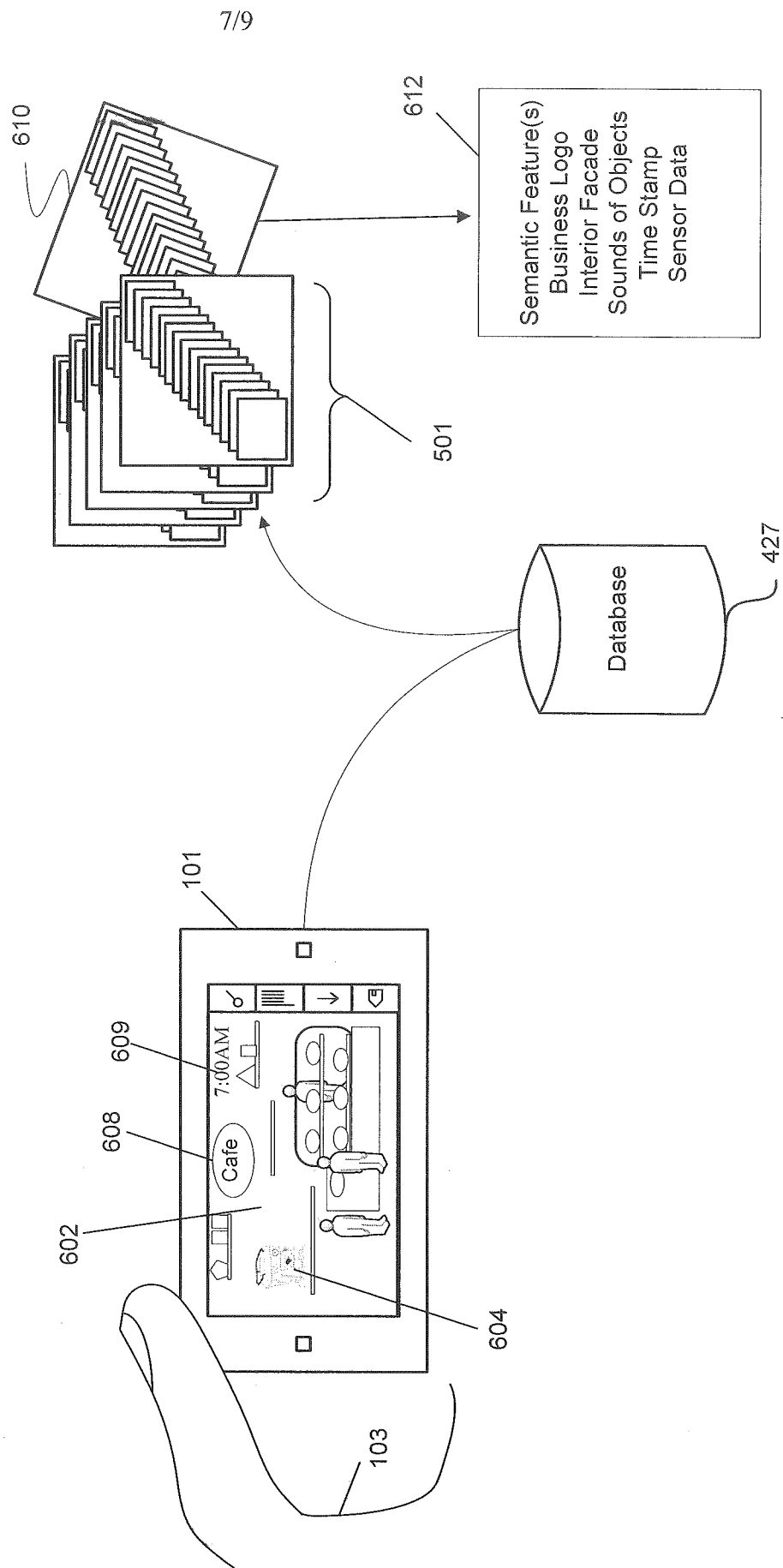


FIG. 6



700

FIG. 7

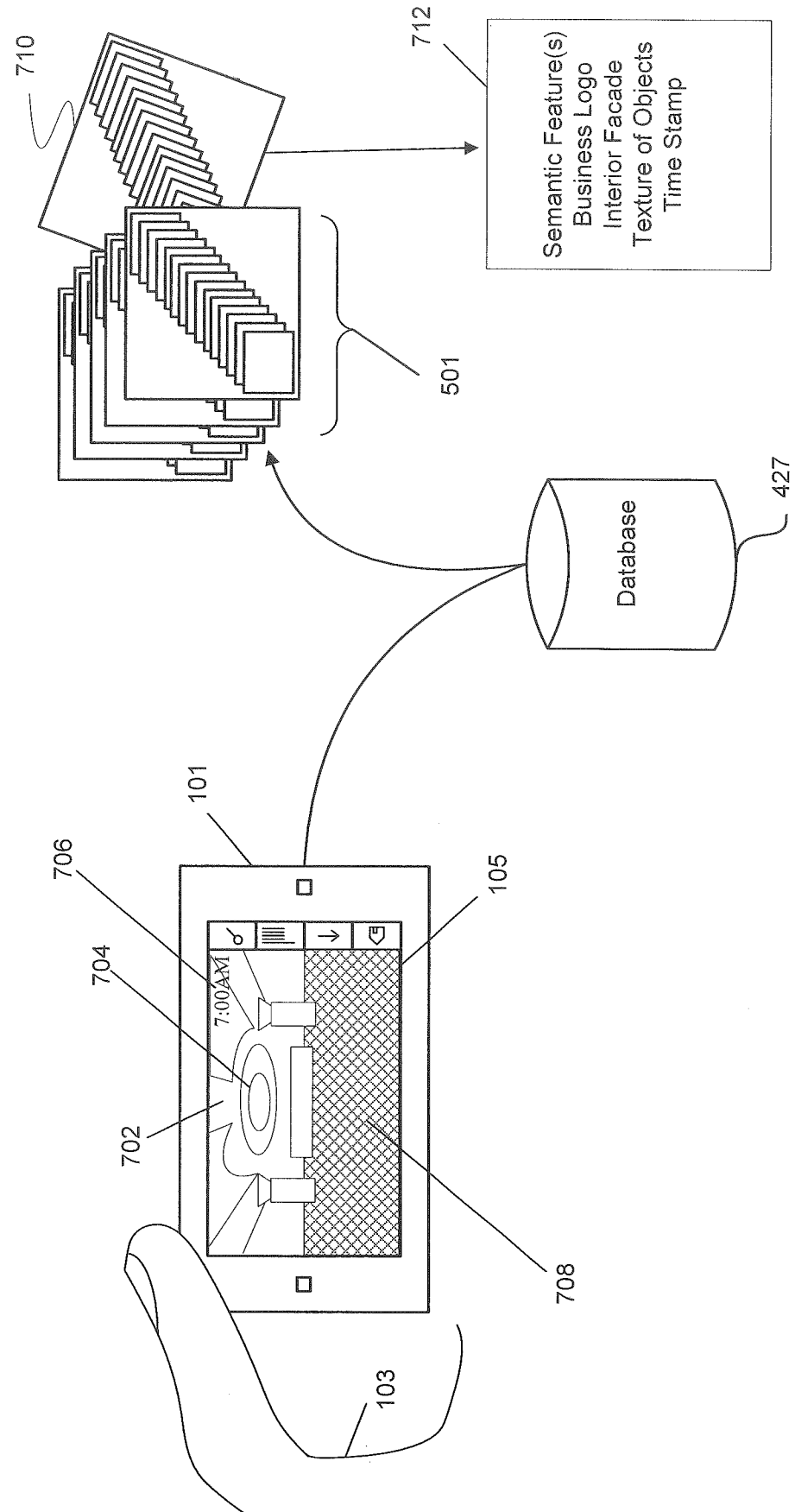
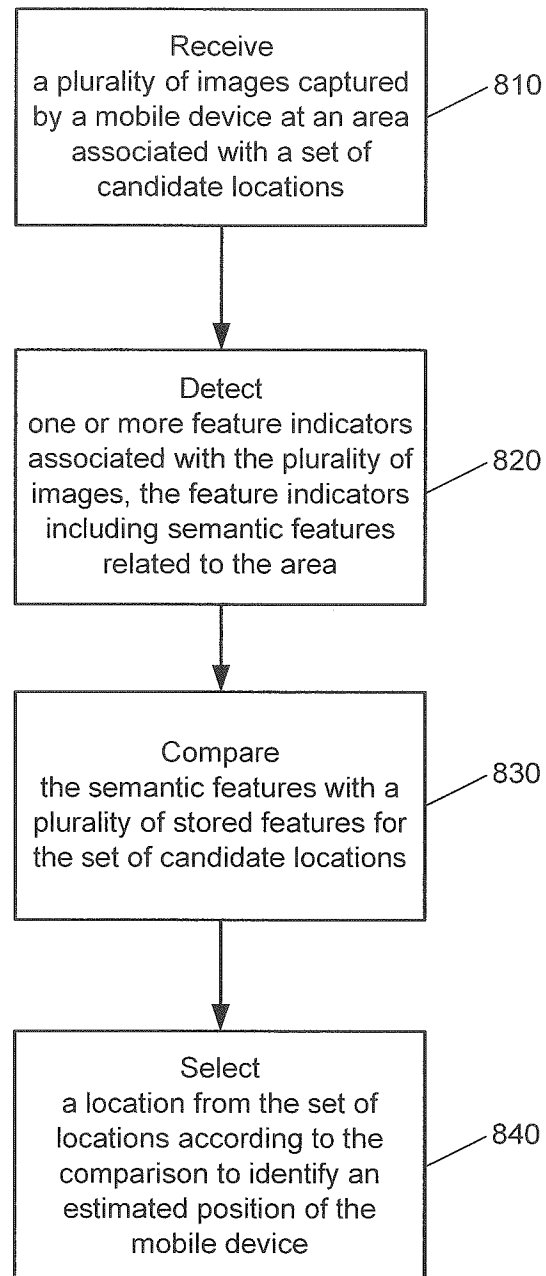


FIG. 8

800

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2016/021780

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06K9/00
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EP0-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	BEIJBOM OSCAR ET AL: "Menu-Match: Restaurant-Specific Food Logging from Images", 2015 IEEE WINTER CONFERENCE ON APPLICATIONS OF COMPUTER VISION, IEEE, 5 January 2015 (2015-01-05), pages 844-851, XP032738297, DOI: 10.1109/WACV.2015.117 [retrieved on 2015-02-19]	1-4, 7-15, 18-20
Y	the whole document ----- -/--	5,6,16, 17

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

10 June 2016

Date of mailing of the international search report

17/06/2016

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Neubüser, Bernhard

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2016/021780

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Selina Chu ET AL: "Content Analysis for Acoustic Environment Classification in Mobile Robots", AAAI Fall Symposium, Aurally Informed Performance: Integrating Machine Listening and Auditory Presentation in Robotic Systems, 1 January 2006 (2006-01-01), XP055279763, Retrieved from the Internet: URL:http://sail.usc.edu/aigaion2/index.php/attachments/single/197 [retrieved on 2016-06-10] abstract	5,16
Y	----- US 2012/183172 A1 (STROILA MATEI [US]) 19 July 2012 (2012-07-19) abstract -----	6,17

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2016/021780

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2012183172 A1	19-07-2012	NONE	
