



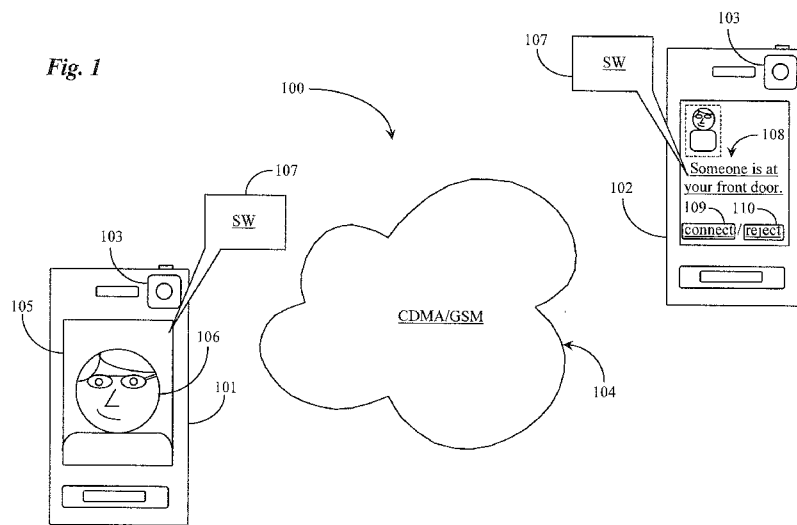
- (51) International Patent Classification:
G06K 9/36 (2006.01)
- (21) International Application Number:
PCT/US2012/049366
- (22) International Filing Date:
2 August 2012 (02.08.2012)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
13/196,304 2 August 2011 (02.08.2011) US
- (71) Applicant (for all designated States except US):
GENESYS TELECOMMUNICATIONS LABORATORIES, INC. [US/US]; 2001 Junipero Serra Boulevard, Daly City, CA 94014 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **SURTY, Aaron, S.** [CA/CA]; 21 Bobwhite Cres., Toronto, ON M2L 2E2 (CA). **LEE-KWEN, Adrian** [CA/CA]; 78 Harrison Garden Boulevard, Apt. 2502, Toronto, ON M2N 7E2 (CA). **BEILIS, David** [CA/CA]; 129 Cabernet Road, Vaughan, ON L4J 8V7 (CA).
- (74) Agent: **CHANG, Josephine, E.**; Christie, Parker & Hale, LLP, P.o. Box 29001, Glendale, CA 91309-9001 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): 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, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, 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.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, 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, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— with international search report (Art. 21(3))

(54) Title: HANDS-FREE VOICE/VIDEO SESSION INITIATION USING FACE DETECTION



(57) Abstract: A communication system includes a telecommunication appliance connected to a communication network, an image acquisition appliance coupled to the telecommunication appliance, software executing on the telecommunication appliance from a non-transitory physical medium, the software providing a first function enabling detecting that an image acquired by the camera comprises a human face in at least a portion of the image, and a second function initiating a communication event directed to a pre-programmed destination, the second function initiated by the first function detecting the human face image portion.

WO 2013/019970 A1

1 HANDS-FREE VOICE/VIDEO SESSION INITIATION USING FACE DETECTION

FIELD OF THE INVENTION

5 [0001] The present invention is in the field of telecommunications and pertains particularly to methods and apparatus for hands-free initiation of a communications session between two or more parties over a network.

DISCUSSION OF THE STATE OF THE ART

10 [0002] In the art of telecommunication, hands-free communication is an important goal for companies that provide communications appliances and mobile software applications. In typical voice communications, there are a number of regimens that seek to provide mostly hands-free initiation of telephone calls. These include voice activation whereby a user says the name of a contact and then says call, whereby the communications appliance enabled
15 with voice recognition initiates a call based on the spoken commands of the user. Earpieces are commonplace and enable communication in a hands-free fashion on many appliances once the call is in progress. Speed dialing by touching a single button to initiate a communication session is also very common.

[0003] A challenge with complete hands-free initiation of a voice/video session between
20 two communications appliances over a network is that current hands-free protocols are too cumbersome for some users, especially those who are not computer literate such as very young children or others with no prior computing experience. Moreover, some prior knowledge of the process for initiating the two-way voice/video call is required. Therefore, what is clearly needed is a system for hands-free voice/video session initiation using face
25 detection.

SUMMARY OF THE INVENTION

[0004] The problem stated above is that simplicity is desirable for initiating a telecommunications session over a digital network, but many of the conventional means for
30 initiating a telecommunications session such as selecting a destination and manually or voice dialing a number, also include complexity whereby tasks are manually initiated. The inventors therefore considered functional elements of a telephony application, looking for elements that exhibit interoperability that could potentially be harnessed to provide voice and video connection but in a manner that would not add more work for the initiator.

35 [0005] Every communications system is characterized by efficiency and simplicity, one by-product of which is an abundance of users communicating via the system. Most such telecommunications systems employ telecommunications appliances with video capability to conduct communications sessions over the prevailing network, and communications

1 appliances having connectivity to a communications network are typically a part of such
apparatus.

[0006] The present inventor realized in an inventive moment that if, at the point of
initiation of a telecommunications session, the face of the user could be detected and the fact
5 of detection could initiate a communication event, significant advantages would accrue.

[0007] In one embodiment of the invention a communication system is provided,
comprising a telecommunication appliance connected to a communication network, an image
acquisition appliance coupled to the telecommunication appliance, and software executing on
the telecommunication appliance from a non-transitory physical medium. The software
10 provides a first function enabling detecting that an image acquired by the camera comprises a
human face in at least a portion of the image, and a second function initiating a
communication event directed to a pre-programmed destination, the second function initiated
by the first function detecting the human face image portion.

[0008] Also in one embodiment the image acquisition device is coupled to the
15 telecommunication appliance by a network. Still in one embodiment the image acquisition
device is physically integrated into the telecommunication appliance. In these embodiments
telecommunication appliance may be an iPad2, an iPhone4, a computer, or an android device.

[0009] In one embodiment of the communication system the image acquisition device is a
video camera, and the communication event is a video call initiated by the telecommunication
20 appliance to the destination. IN another embodiment the telecommunication event is a voice
call initiated by the telecommunication appliance to the destination.

[0010] In yet another embodiment of the invention the telecommunication event
comprises an image of at least the portion of the acquired image determined to represent a
human face by the first function, and further comprising a third function enabling automatic
25 answering of an incoming call voice or video call at the telecommunication appliance. In yet
another, there is a fourth function for recognizing a specific facial expression in an image of
a human face acquired by the video camera, and initiating one or more specific actions
accordingly.

[0011] In another aspect of the invention a communication method is provided,
30 comprising the steps of: (a) acquiring an image by an image acquisition device; (b) providing
the image to a telecommunication appliance connected to a communication network; (c)
detecting, by the telecommunication appliance, an image of a human face in a portion of the
image acquired; (d) initiating a communication event directed to a pre-programmed
destination as a result of the detection of the image of a human face in step (c). In one
35 embodiment the image acquisition device is coupled to the telecommunication appliance by a
network. Also in one embodiment the image acquisition device is physically integrated into
the telecommunication appliance. In the integrated case the telecommunications appliance
may be an iPad2, an iPhone4, a computer, or an android device.

1 [0012] In another embodiment of the present invention the image acquisition device is a video camera, and the communication event is a video call initiated by the telecommunication appliance to the destination. In yet another, the telecommunication event is a voice call initiated by the telecommunication appliance to the destination.

5 [0013] In still another embodiment the communication event comprises an image of at least the portion of the acquired image determined to represent a human face by the first function, and further comprising a third function enabling automatic answering of an incoming call voice or video call at the telecommunication appliance. In yet another, a fourth function is provided for recognizing a specific facial expression in an image of a human face
10 acquired by the video camera, and initiating one or more specific actions accordingly.

BRIEF DESCRIPTION OF THE DRAWINGS

[0014] Fig. 1 is an architectural view of a communications network supporting two-way video and voice communication initiated through face detection according to an embodiment
15 of the present invention.

[0015] Fig. 2 is an architectural view of a sales network supporting face-detect initiation of a service call from an enabled kiosk according to an embodiment of the present invention.

[0016] Fig. 3 is a process flow chart illustrating steps for initiating two-way voice and video communication using face detection according to the embodiment of Fig. 1.

20 [0017] Fig. 4 is a process flow chart illustrating steps for initiating a service communication using face detection according to the embodiment of Fig. 2.

[0018] Fig. 5 is an architectural view of a communications network supporting two-way communication using face detection software to initiate a session and face expression recognition to control the session once established.

25

DETAILED DESCRIPTION

[0019] The inventors provide a communications system for enabling voice and video calls through facial detection of a human operator. The present invention is described in enabling detail using the following examples, which may describe more than one
30 embodiment falling within the scope of the present invention.

[0020] Fig. 1 is an architectural view of a communications network 100 supporting two-way video and voice communication initiated through face detection according to an embodiment of the present invention. Communications network 100 may include a wireless carrier network 104. Wireless network 104 may include a code division multiple access
35 (CDMA) network, a global system for mobile communications (GSM) network, another wireless carrier network type or a combination of these. Examples are WiFi, WiMax or LTE, essentially any 3G or 4G network.

1 [0021] Carrier network 104 supports voice and video communications between mobile
communications appliances. A communications appliance 101 is illustrated in this example
and has connectivity to network 104. Appliance 101 may be an android, a cellular telephone,
a smart hand-held personal computer (PC), or any other communications appliance. In a
5 preferred embodiment, communications appliance 101 has a forward facing camera lens 103.
By forward facing, it is meant to mean that the camera aperture is on the same side of the
appliance as the display. Appliance 101 may be a touch screen appliance without departing
from the spirit and scope of the present invention.

[0022] Appliance 101 includes a visual display window 105. In this example, camera
10 101 is displaying the face 106 of a user who wishes to initiate a voice/video communication
session with another appliance over network 104. To initiate communication between the
instant appliance 101 and another appliance over the network, a face-detection software (SW)
107 is provided to reside on and execute from a non-transitory medium coupled to appliance
101. Face detection SW 107 is adapted to monitor for the presence a human face, preferably
15 the face of the appliance owner when the user looks into the forward-facing camera lens
while SW 107 is running on the device.

[0023] SW 107 then uses the camera function to acquire a still picture of the user and
packages the photo into a session initiation request that is propagated from the initiating
appliance to a targeted appliance over network 104. A second communications appliance 102
20 is illustrated in this example and has access to wireless network 104. Appliance 102 may be
an android, a cellular telephone, a smart hand-held personal computer (PC), or any other
communications appliance. In a preferred embodiment, communications appliance 102 has a
forward facing camera lens 103. In the present example of the invention, communications
appliance 101 is initiating a request for a video/voice telephony session from appliance 102.

25 [0024] Appliance 102 may be a same or similar appliance as appliance 101. In one
embodiment the initiating and receiving appliances may be a fixed appliances instead of
mobile appliances without departing from the spirit and scope of the present invention.
Appliance 102 has a forward facing camera 103 and has an instance of face-detection SW
107 residing and executing from a non-transitory medium on the appliance. Once appliance
30 101 running SW 107 has taken a photo 106 of the requestor, the photo along with a request
for communication is automatically propagated over network 104 to, in this case appliance
102.

[0025] Appliance 102 displays a thumbnail picture received photo 106 in the right upper
corner of the display when the request for a session is received over network 104. A text
35 message 108 informs the recipient "Someone is at your front door". Message 108 may be
worded differently than the message illustrated without departing from the spirit and scope of
the present invention. In this example, the recipient has an option of connecting to the
requesting user in a voice and video session by hitting a connect button 109. In one

1 embodiment the recipient of photo 106 simply clicks on or activates the photo in order to
launch a video/voice communication session. The recipient may reject the request by
clicking on a reject button 110.

5 [0026] When the recipient accepts the communication request, a voice/video session is
launched and may be answered as an incoming communication at the requestor's end
(appliance 101) by having an automatic answering function turned on. This makes the
session initiation from the point of the requestor 100 percent hands free except perhaps what
it takes to "look into" the forward facing camera. In one embodiment, the requesting
10 appliance 101 is a stationary appliance or is mounted in a stationary fashion with
speakerphone and video enabled.

[0027] Face detection enables a very young person, perhaps of three or four years of age,
to request a voice/video session with a pre-programmed recipient such as a parent or
guardian. The system of the invention enables disabled persons who otherwise cannot
operate complex electronic devices to request a voice/video session with a preprogrammed
15 recipient. In one embodiment, voice recognition enables a user operating appliance 101 to
call out a contact that has a number listed on the appliance or to call out a number before the
video/voice request is generated. In one embodiment, a list of numbers can be utilized for
multiple requests of a video/voice session and the appliance may auto-answer the first
responder.

20 [0028] Fig. 2 is an architectural view of a sales network 200 supporting face-detect
initiation of a service call from an enabled kiosk according to an embodiment of the present
invention. Sales network 200 includes a call-in center 207 that is accessible over a data
network 201, that may include an automated transaction machine (ATM) network, a time
division multiplexing (TDM) network, or an Internet protocol (IP) network or a combination
25 of these. In this example, a kiosk, or an ATM terminal 202 is provided for customers to
make contact with a sales or service representative of a store hosting the kiosk and network.
Terminal 202 enables a customer to communicate with a rep and purchase a product.

[0029] An instance of face detection SW 217 is provided to reside on and execute from a
non-transitory medium coupled to terminal 202. SW 217 enables initiation and launch of a
30 voice/video request. Terminal 202 includes two speakers 203, a microphone 206 and a
forward facing camera 205. Terminal 202 includes a large display screen 204. When a user
steps in front of the camera, the likeness of the user is displayed on screen and an image of
the user is acquired and propagated over network 201 to call-in center 207.

[0030] Call-in center 207 represents any local or remote group of service representatives
35 enabled for telephony communications. Call-in center 207 includes a local area network
(LAN) 208 that is enabled for Internet Protocol (IP) telephony. A service center may take
incoming calls through a computer telephony integration (CTI)-enabled private branch
exchange (PBX) switch 209. Switch 209 is connected to LAN 208 and to carrier network

1 201. In one embodiment, incoming calls are intercepted by an interactive voice response (IVR) unit, not illustrated but assumed present in this example. CTI SW 212 resides on a non-transitory medium coupled to switch 209 and provides intelligent routing routines for agent level routing of incoming calls.

5 [0031] LAN 208 supports a collection of agent workstations 216 (1-n). Workstations 216 (1-n) are adapted as communications terminals manned by live agents or knowledge workers. Each station is represented herein by a personal computer (PC) that is directly connected to LAN 208. Telephone handsets or headsets may also be assumed present in this example though they are not specifically illustrated.

10 [0032] LAN 208 supports an interaction routing (IR) server 210. Server 210 is aided by software 213 residing on and executable from a non-transitory medium on the server. Interaction routing server 210 also includes a message queue 214. Message queue 214 is adapted to contain communications requests that are waiting to be routed to a live agent. Communications requests are also held in PBX 209. In this example, the call-in center
15 differentiates between normal incoming traffic and incoming requests resulting from kiosk interaction at host storefronts. LAN 208 supports a statistics (Stat) server 211. Statistics server 211 includes software 215 resident on and executable from a non-transitory medium that aids in compiling and reporting call-center statistics to requesting entities.

[0033] If a user wishes to request a communications session with a live agent from call-in
20 center 207, the user simply looks into the camera in front of the terminal. SW 217 aids in generating a communications request including the image acquired at the time of the request. The communications request may be propagated over network 201 to switch 209. IVR SW may intercept the call and may determine if it is a call incoming from the network or one incoming from a host terminal such as terminal 202. Workstations 216 (1), 216 (2), and 216
25 (n) are busy handling voice calls. In one embodiment, the destination party is a video/voiceXML-automated script and the call is launched from the mobile telecommunications appliance upon detection of a face.

[0034] Workstation 216 (3) has received a routed message with a thumbnail picture attached indicating it is from a Kiosk interaction where the operator wishes to engage in a
30 voice/video session with a live agent. The agent operating station 216 (3) may answer the request by double clicking the thumbnail, which immediately launches a voice /video session between the agent and a user operating the kiosk. In another embodiment, the operator of terminal 202 steps in front of the kiosk, looks into the camera, and the terminal camera acquires an image and sends a normal voice/video call to call-in center 207 for further agent
35 level routing as an incoming voice/video call.

[0035] In one embodiment, the present invention may be deployed as an intercom system that connects visitors to a home with a remote homeowner that is not physically on the premises. In this case the display screen with a forward-facing camera is a fixed

1 telecommunications device or terminal that the visitors may step in front of to automatically
initiate a voice /video session. In one embodiment, a public event may have a terminal that
performs, using the present invention, as a telecommunications device for interviewing and
registering members of the public whereby media uploads and video/voice confirmation may
5 be practiced.

[0036] Fig. 3 is a process flow chart 300 illustrating steps for initiating two-way voice
and video communication using face detection according to the embodiment of Fig. 1. In this
example, it is assumed that the initiator of the communications session is operating a
telecommunications appliance having a display and forward-facing camera.

10 **[0037]** At step 301 the initiator or third party selects one or more numbers for automatic
request propagation as a result of face detection. In one embodiment one number may be
preprogrammed for association to a video/voice communications session request. In another
embodiment, more than one number might be preprogrammed for initiation of a video/voice
communications request. The latter case may be important if one or more parties are un-
15 reachable and the session initiation is due to an emergency of some type. This embodiment
may also assume that the person initiating the session request may be unable to operate a
hand-held and therefore uses 100 % hands-free face detection SW to launch the request. The
number or numbers preprogrammed for session initiation may be preprogrammed using voice
or by a third party.

20 **[0038]** At step 301, the initiator or a third party selects one or more telephone numbers
for voice/video hands-free session initiation as previously described above. At step 302, the
initiator or third party activates the face detection application. The face detection application
may have a sleeping mode that uses minimal power. Motion in front of the forward-facing
camera may trigger face detection.

25 **[0039]** At step 303, the initiator makes a decision whether or not to initiate a voice/video
session through a face detect operation. If the initiator determines not to make a request at
step 303, the application stays asleep at step 304. If the initiator accidentally invokes the face
detect application by motion, a few seconds of still posing might be required before a session
request is generated and sent. This would enable an initiator the option of opting out of a
30 session initiation operation before the window time or “pose” time is finished. The process
resolves back to step 303 from step 304.

[0040] If at step 303, the initiator determines to make a voice/video communications
session request, then the initiator looks into the appliance camera in order to activate face
detection at step 305. The initiator may be required to handle appliance if it is not stationary.
35 In one embodiment, the appliance is fixed or otherwise mounted in a stationary position that
is accessible to the initiator. In this case the initiator may simply move in front of the camera
space in order for the application to wake up and notice the user. The appliance may be set to
speakerphone or the user may wear a headset.

1 [0041] At step 306, the appliance may acquire an image of the initiator and may generate
and send a request message containing the image to one or more of the preset numbers
telephone numbers (from list). At step 306, if there is more than one preset number, the
request may be sent to all of the potential recipients as a request for a video/voice session. If
5 there is only one number in the list, only one recipient will get a request.

[0042] At step 307, a message with a thumbnail picture displays on the recipient's
appliance. In this case recipients are assumed to be human operators of communications
appliances. At step 308 the recipient of a message requesting a voice/video session may
make a decision relative to accepting or rejecting the proposed interaction. If at step 308, the
10 recipient determines to reject the request, a call rejection notice may be sent back at step 309
in lieu of a voice/video session. The process would then resolve back to step 304.

[0043] If the recipient of a message request for a voice/video session decides to accept
the request at step 308, the recipient might interact with the thumbnail representation of the
initiator, or may select a preordained button to launch a voice/video call at step 310 to the
15 initiator of the request message of step 306. At step 311, the video session is launched. At
the original request initiator's end, the initiator may determine if auto answer is set to on at
step 312. If auto answer is not on at step 312, then the initiator of the request may manually
answer the video/voice call at step 313. At step 312, if the auto answer is on, the process
moves to step 313 where the video session is answered and in progress. After manual answer
20 at step 313, the process resolves to step 314. This example follows the embodiment of Fig. 1.

[0044] Fig. 4 is a process flow chart 400 illustrating steps for initiating a service
communication using face detection according to the embodiment of Fig. 2. Flow chart 400
follows the embodiment of Fig. 2 where it is assumed that the initiator of a video/voice
session is a shopper operating a kiosk type terminal. At step 401, the shopper has a question
25 about a product he or she is considering for purchase. At step 402, the shopper moves in
front of the video-enabled kiosk analogous to terminal 202 of Fig. 2 with or without the
product in question. In one embodiment, the shopper has the product in hand and the kiosk
asks the shopper to enter or scan in the product code at step 403.

[0045] At step 404, the kiosk acquires an image of the shopper and sends a request
30 including the picture and the product code to customer service as a request for a voice/video
session. At step 405, the sales representative that receives the request, which may have been
routed to a first available agent, clicks on the thumbnail picture of the shopper to initiate or
launch a video/voice call. The product code is also forwarded to the sales rep along with any
other data that may be known about the shopper. At step 406, the video/voice session
35 requested is established at the kiosk for the user. After the session, the product may be
purchased through the terminal using standard keypad and touch screen payment options.

[0046] In one embodiment the act of step 404 may, instead of being a request for a
session, may be a session launch that is routed to a next available service representative as an

1 incoming video/voice call. In this case, in step 405 the sales rep would simply accept the call
(pick-up). In one embodiment, step 403 is not required and step 405 is not required. In this
case, step 406 would call out the session at the terminal of the sales representative picked up
as an incoming interaction from the switch or queue.

5 **[0047]** Fig. 5 is an architectural view of a communications network 500 supporting two-
way communication initiated through face detection and that may be controlled during the
session using facial expression recognition according to an embodiment of the present
invention. Communications network 500 includes CDMA/GSM network 104 and previously
described communications appliances 101 and 102 aided by face detection SW 107. Similar
10 to the embodiment described with respect to Fig. 1, face detection SW 107 is utilized on
appliance 101 to initiate and send a message to appliance 102 requesting a voice/video
communications session.

[0048] In this example, a gateway router 505 is illustrated within network 104 for
enabling Internet Protocol (IP) based brokerage of the pending communications session.
15 Gateway router 505 may be a multimedia gateway connected to the Internet network. The
Internet network is represented herein by a network backbone 504. Backbone 504 represents
all of the lines, equipment, and access points that make up the Internet network as a whole.
Therefore, there are no geographic limitations to the practice of the present invention.
Gateway router 505 is connected to backbone 504 to enable Internet connection for
20 appliances 101 and 102.

[0049] A third party server 501 is illustrated and is connected to Internet backbone 504.
Server 501 includes a non-transitory digital medium adapted with all of the software and data
required to enable function as a communications session broker for sessions engaged in
between appliances connected to the network like appliances 101 and 102. Server 501 has an
25 instance of expression recognition SW 503 resident and executable from the non-transitory
medium coupled to the server.

[0050] Expression recognition SW 503 is adapted with the aid of a connected facial
recognition database 502 to enable server 501 to broker communications session established
between appliances 101 and 102 that have been established through face detection and
30 subsequent acceptance of a session request. In this example communication is brokered by
server 501 through gateway router 505. Examples of facial expression recognition may
include but are not limited to nods, head shaking horizontally or vertically, winks, frowns,
smiles, and the like. In one embodiment of the present invention facial expressions may help
control the session in terms of session termination, direction of session content, destination
35 redirect, and session request routing in case of multiple possible destinations. In an
embodiment with multiple routing destinations or departments, a frown may result in a
prompt to be routed to customer service department whereas a smile may result in a proactive

1 sales offer. In this case a video-enable IVR (not illustrated) may be employed to detect facial expressions of a session initiator for routing is completed.

[0051] In one embodiment, the session requesting appliance is a stationary appliance or otherwise fixed in a stationary position. In this embodiment the face detection application
5 107 may be left open. Colors may be used to indicate different states of the face detection application. When someone walks inside the field of view of the camera the face detection application may activate by displaying a camera with a colored border. The colors may be an indication of the state of the application. For example, red may indicate the camera does not detect anyone. If the application is left in the red state for a long enough period, the
10 application may go back to sleep. Yellow may indicate that someone is detected and the application is preparing an image from the camera to be sent to the other party. Green may indicate that a notification (request for a voice/video session) was sent to the other party. At this point the other party may choose to launch the video call or the face detection application may back to sleep. A simple motion detection module may be used to determine if someone
15 is in the field of view of the camera. In the embodiment where facial expression recognition is provided, face recognition may also be used to recognize users who have a likeness of them in the form of a frontal image stored in a database. In this way users are recognized and identified by the system.

[0052] It will be apparent to one with skill in the art that the communications system of
20 the invention may be provided using some or all of the mentioned features and components without departing from the spirit and scope of the present invention. It will also be apparent to the skilled artisan that the embodiments described above are specific examples of a single broader invention that may have greater scope than any of the singular descriptions taught. There may be many alterations made in the descriptions without departing from the spirit and
25 scope of the present invention.

30

35

1 WHAT IS CLAIMED IS:

1. A communication system comprising:
a telecommunication appliance connected to a communication network;
an image acquisition appliance coupled to the telecommunication appliance; and
5 software executing on the telecommunication appliance from a non-transitory physical medium, the software providing:
a first function enabling detecting that an image acquired by the camera comprises a human face in at least a portion of the image; and
a second function initiating a communication event directed to a pre-programmed
10 destination, the second function initiated by the first function detecting the human face image portion.
2. The communication system of claim 1 wherein the image acquisition device is coupled to the telecommunication appliance by a network.
15
3. The communication system of claim 1 wherein the image acquisition device is physically integrated into the telecommunication appliance.
4. The communication system of claim 3, wherein the telecommunication
20 appliance is an iPad2, an iPhone4, a computer, or an android device.
5. The communication system of claim 1 wherein the image acquisition device is a video camera, and the communication event is a video call initiated by the telecommunication appliance to the destination.
25
6. The communication system of claim 1 wherein the communication event is a voice call initiated by the telecommunication appliance to the destination.
7. The communication system of claim 1 wherein the communication event
30 comprises an image of at least the portion of the acquired image determined to represent a human face by the first function, and further comprising a third function enabling automatic answering of an incoming call voice or video call at the telecommunication appliance.
8. The communication system of claim 2, further comprising a fourth function
35 for recognizing a specific facial expression in an image of a human face acquired by the video camera, and initiating one or more specific actions accordingly.

1 9. A communication method comprising the steps of:
 (a) acquiring an image by an image acquisition device;
 (b) providing the image to a telecommunication appliance connected to a
communication network;
5 (c) detecting, by the telecommunication appliance, an image of a human face in a
portion of the image acquired;
 (d) initiating a communication event directed to a pre-programmed destination as a
result of the detection of the image of a human face in step (c).

10 10. The communication method of claim 9 wherein the image acquisition device
is coupled to the telecommunication appliance by a network.

 11. The communication method of claim 9 wherein the image acquisition device
is physically integrated into the telecommunication appliance.

15 12. The communication method of claim 11, wherein the telecommunications
appliance is an iPad2, an iPhone4, a computer, or an android device.

 13. The communication method of claim 9 wherein the image acquisition device
20 is a video camera, and the communication event is a video call initiated by the
telecommunication appliance to the destination.

 14. The communication method of claim 9 wherein the telecommunication event
is a voice call initiated by the telecommunication appliance to the destination.

25 15. The communication method of claim 9 wherein the communication event
comprises an image of at least the portion of the acquired image determined to represent a
human face by the first function, and further comprising a third function enabling automatic
answering of an incoming call voice or video call at the telecommunication appliance.

30 16. The communication method of claim 13, further comprising a fourth function
for recognizing a specific facial expression in an image of a human face acquired by the
video camera, and initiating one or more specific actions accordingly.

35

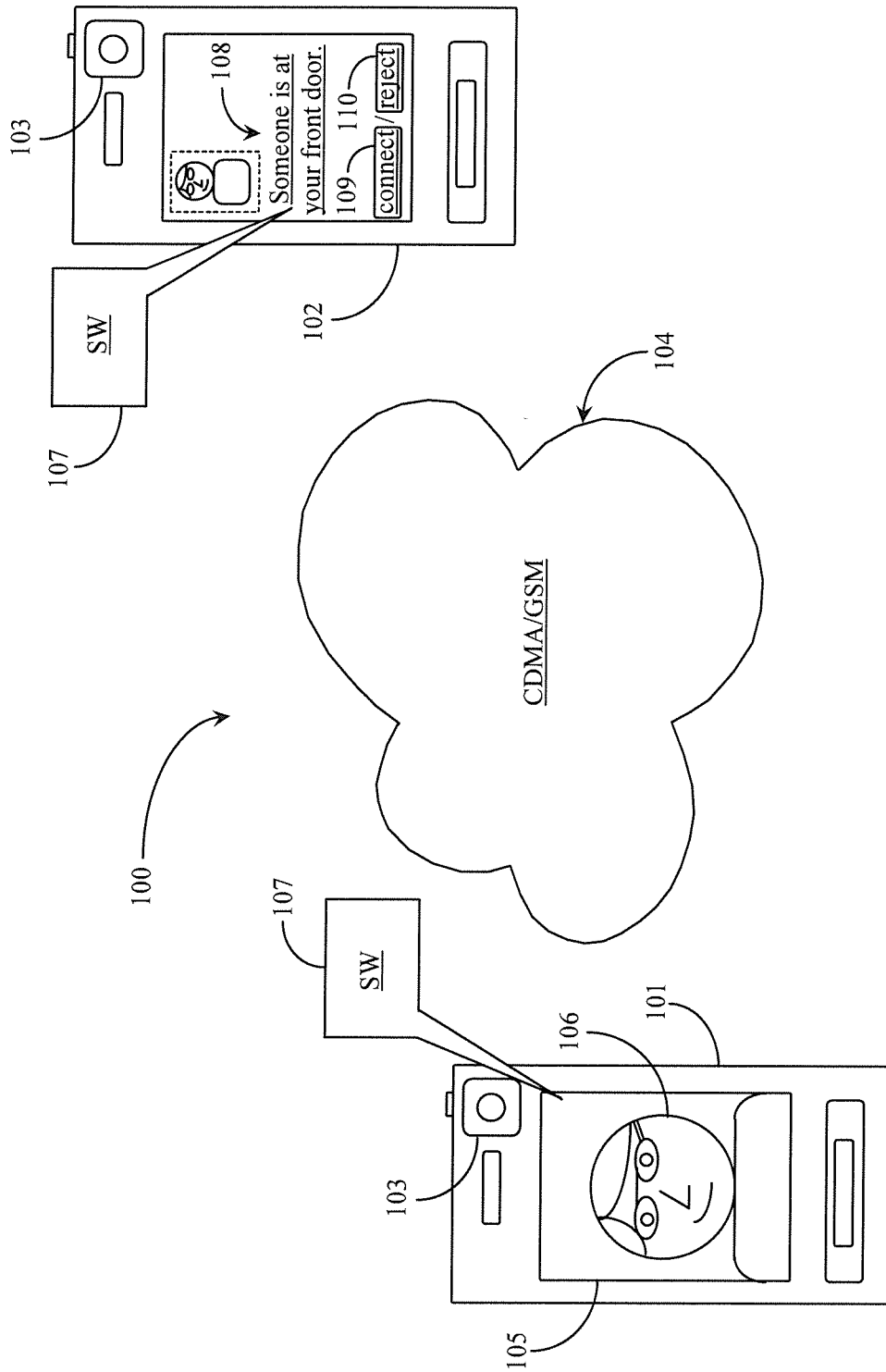


Fig. 1

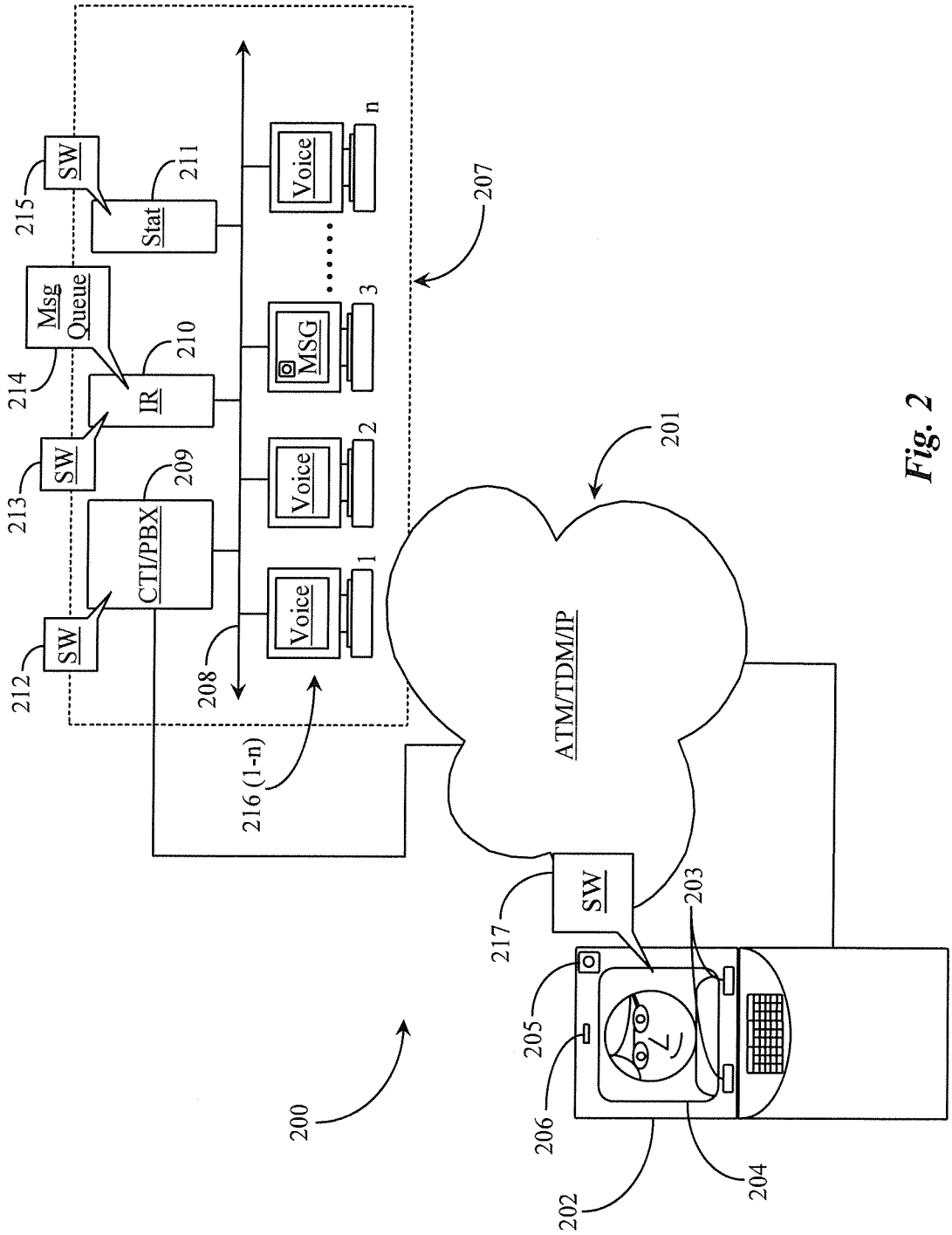


Fig. 2

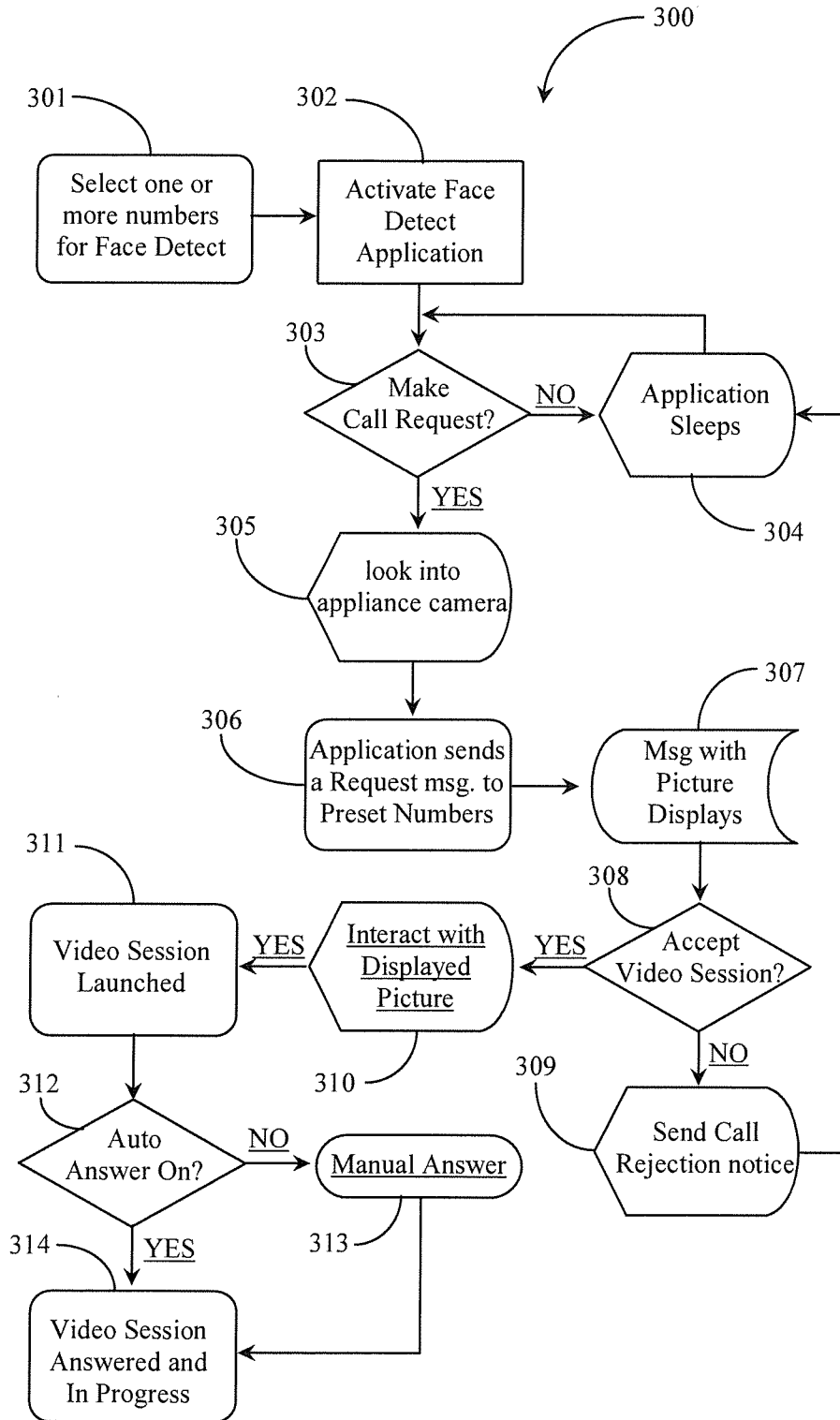


Fig. 3

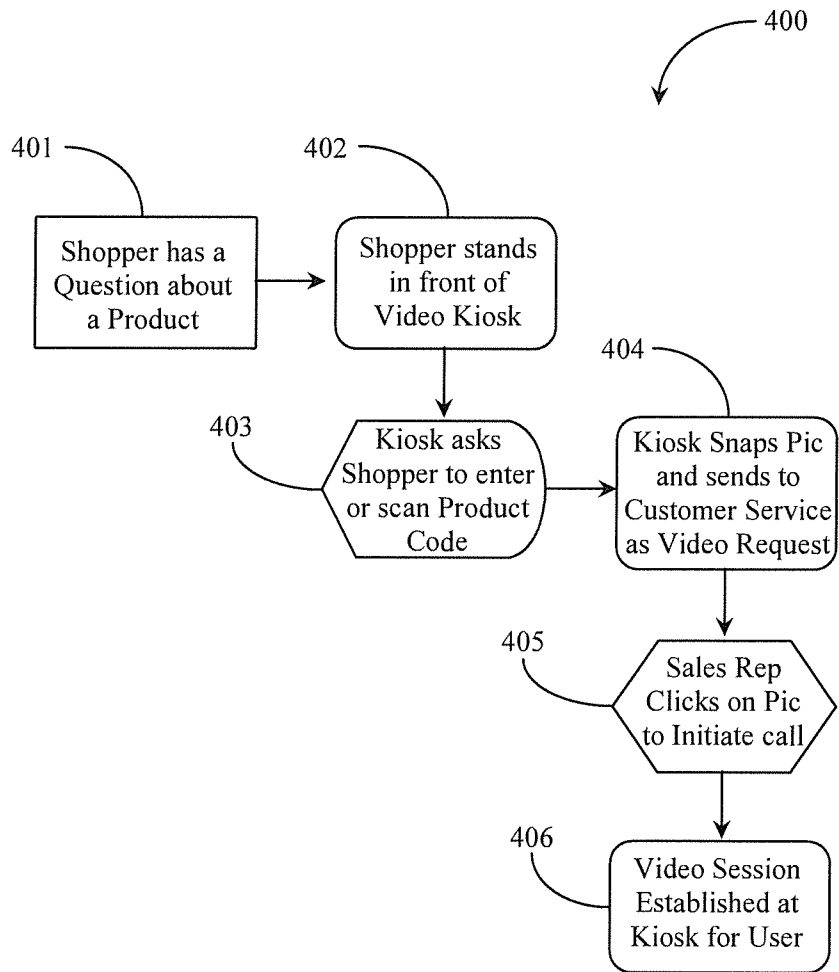


Fig. 4

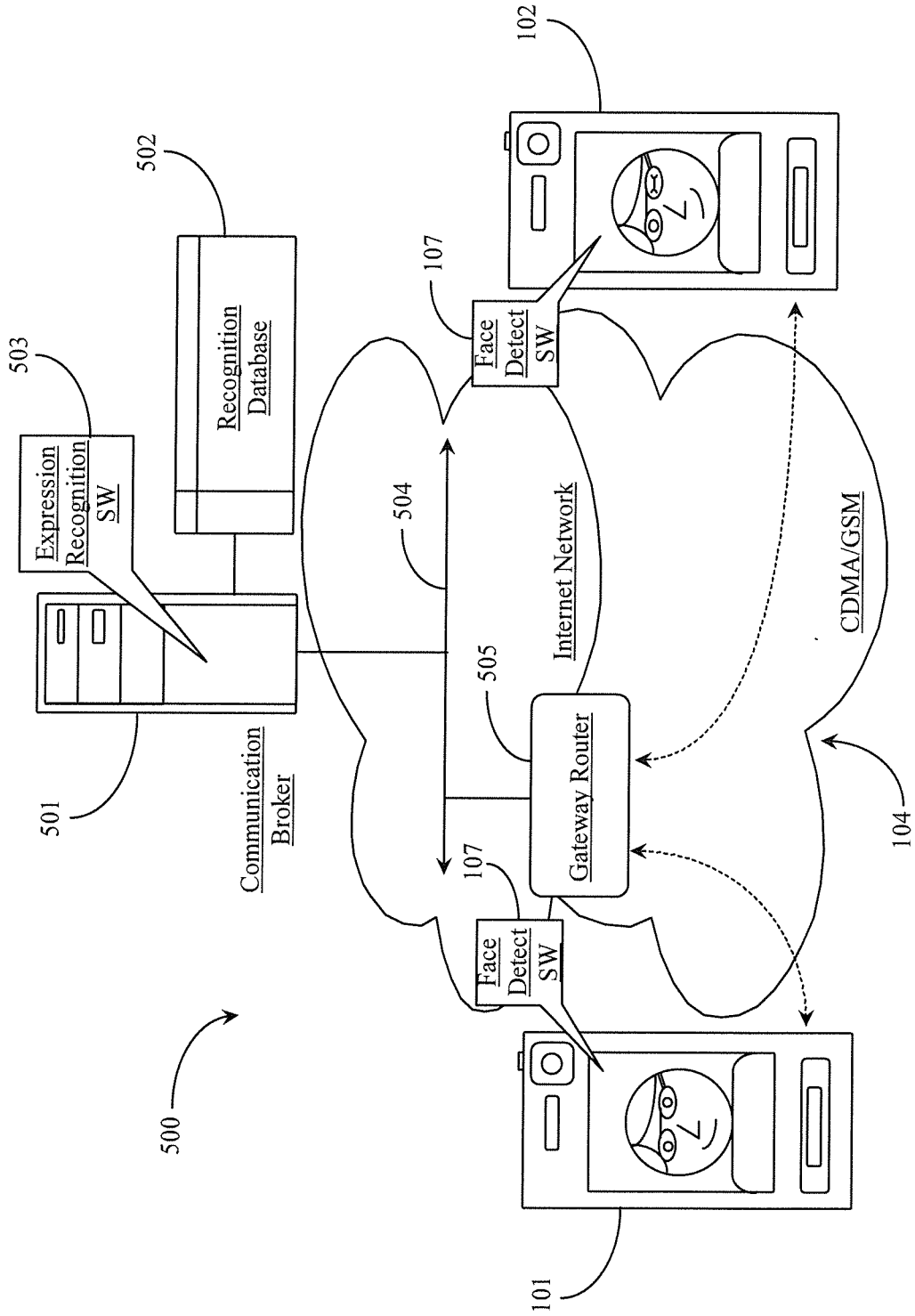


Fig. 5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2012/049366

A. CLASSIFICATION OF SUBJECT MATTER IPC(8) - G06K9/36 (2012.01) USPC - 382/118 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) IPC(8) - G06K9/36, G06F17/30, H04B 7/00 (2012.01) USPC - 382/118,181, 276; 379/420.02, 420.03 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) Orbit.com, Google Patents, Google Scholar		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2002/0176565 A1 (DECROUY et al) 28 November 2002 (28.11.2002) entire document	1-16
Y	US 2009/0160609 AI (LIN et al) 25 June 2009 (25.06.2009) entire document	1-16
Y	US 2010/0149303 A1 (THORNE et al) 17 June 2010 (17.06.2010) entire document	2, 5, 8,10, 13, 16
Y	US 7,688,967 B2 (PATEL et al) 30 March 2010 (30.03.2010) entire document	5, 13, 16
Y	US 2001/0028703 A1 (KATSEFF et al) 11 October 2001 (11.10.2001) entire document	7, 8, 15, 16
Y	US 2009/0110246 A1 (OLSSON et al) 30 April 2009 (30.04.2009) entire document	8,16
A	US 2009/0046900 A1 (OGAWA) 19 February 2009 (19.02.2009) entire document	1-16
A	US 2009/0170562 A1 (KUJALA et al) 02 July 2009 (02.07.2009) entire document	1-16
A	US 6,018,668 A (SCHMIDT) 25 January 2000 (25.01.2000) entire document	1-16
A	US 7,130,454 B1 (BERUBE et al) 31 October 2006 (31.10.2006) entire document	1-16
A	US 7,404,001 B2 (CAMPBELL et al) 22 July 2008 (22.07.2008) entire document	1-16
A	US 7,930,555 B2 (BELL) 19 April 2011 (19.04.2011) entire document	1-16
A	US 2009/0116684 A1(ANDREASSON) 07 May 2009 (07.05.2009) entire document	1-16
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/>		
* 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 28 September 2012		Date of mailing of the international search report 05 NOV 2012
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201		Authorized officer: Blaine R. Copenheaver PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774