



- (51) **International Patent Classification:**
H04N 7/15 (2006.01)
- (21) **International Application Number:**
PCT/EP2013/065999
- (22) **International Filing Date:**
30 July 2013 (30.07.2013)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
12305965.1 3 August 2012 (03.08.2012) EP
- (71) **Applicant:** ALCATEL LUCENT [FR/FR]; 3, avenue Octave Gréard, F-75007 Paris (FR).
- (72) **Inventors:** LIEVENS, Sammy; Leeuweriklaan 7, B-2930 Brasschaat (BE). TYTGAT, Donny; Evarist De Buckstraat 15, B-9040 Sint-Amandsberg (BE). AERTS, Maarten; Leurshoek 8, B-9120 Beveren-Waas (BE). NAM-BOODIRI, Vinay; Paul Lebrunstraat 26, Bus 201, B-3000 Leuven (BE). SIX, Erwin; Centrumwijk 47, B-9270 Kalken (BE).
- (74) **Agent:** ALU ANTW PATENT ATTORNEYS; Copernicuslaan 50, B-2018 Antwerp (BE).
- (81) **Designated States** (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM,

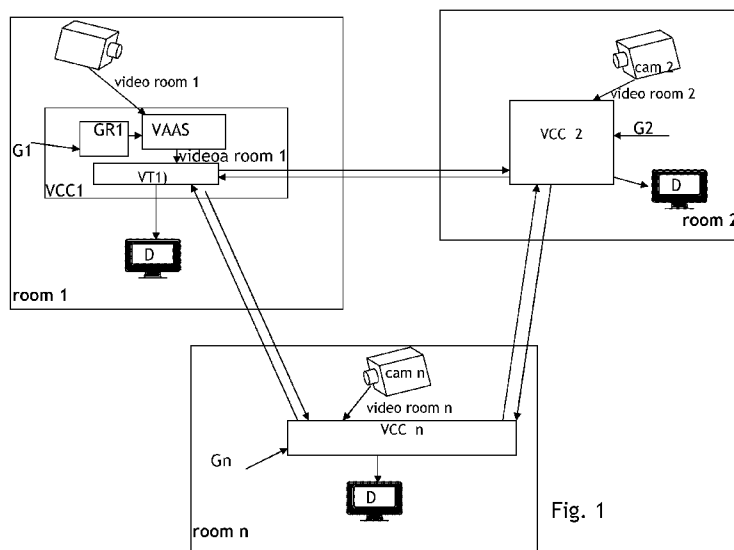
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, 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, KM, ML, MR, NE, SN, TD, TG).

Published:

- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))

(54) **Title:** METHOD AND APPARATUS FOR ENABLING VISUAL MUTE OF A PARTICIPANT DURING VIDEO CONFERENCING



(57) **Abstract:** A method for adapting video data (video room1) recorded by a camera on a location (room1) during a video conference such as to hide the presence of a participant at said location of said video conference, comprises a step of registering a pre-defined gesture possibly to be performed by any participant of said video conference at said location (room1), a step of detecting said gesture, and upon detection thereof, identifying the at least one participant having performed said gesture at said location (room1), adapting said video data such as to eliminate data relating to said at least one participant having performed said gesture from said video data, thereby generating adapted video data (video room1), for being transmitted to other participants of said video conference on other locations (room2,...,roomn).

**METHOD AND APPARATUS FOR ENABLING VISUAL MUTE OF A
PARTICIPANT DURING VIDEO CONFERENCING**

The present invention relates to a method and apparatus to enable users participating to a video conference, for having control whether they will be visible or not to other remote users of the video conference system. This relates to
5 providing a possibility to these users to put themselves on "visual mute".

A simple solution to solve this problem is to simply leave the room. But this isn't always adequate because the user may still want to follow the meeting passively, without actively participating to it.

10 Another solution is that a global human meeting leader manually controls the visual state for each particular participant to the video conference call. This is however also not a feasible solution if many participants are present, or many (de) mute-requests to the system are made.

It is therefore an object of embodiments of the present invention to
15 present a method of the known type but which does not show the aforementioned disadvantages.

According to embodiments of the present invention this object is achieved by a method for adapting video data recorded by a camera on a location during a video conference such as to hide the presence of a participant
20 at said location of said video conference, said method comprising a step of registering a predefined gesture possibly to be performed by any participant of said video conference at said location, a step of detecting said gesture, and upon detection thereof, identifying the at least one participant having performed said gesture at said location , adapting said video data such as to eliminate data
25 relating to said at least one participant having performed said gesture from said video data, thereby generating adapted video data, for being transmitted to other participants of said video conference on other locations.

-2-

In this way, an automated and simple solution is provided enabling visual muting of a participant upon detection of this participant performing a predefined gesture, previously registered.

5 In a variant said predefined gesture is detected by analyzing said video data.

In this variant gesture recognition can be performed via video analysis techniques such as image recognition or the like of the video data itself.

10 In another variant said predefined gesture is detected by means of receiving a trigger detection signal from and transmitted by an object on which said predefined gesture is performed by said at least one participant.

15 This variant allows to detect the gesture in an alternative way, by e.g. receiving a signal from an object present at the conference location, such object generally being adapted to communicate with a video conferencing client, and transmit a signal indicative of the gesture being performed by a participant. Upon receipt of such a signal, the video data may be further analyzed for recognizing the conference participant having performed this gesture. Alternatively, in other embodiments, the trigger detection signal itself may already comprise information with respect to the participant having performed the gesture such that the identification of said at least one participant having performed said gesture at said location is performed by analyzing said trigger
20 detection signal from said object.

In an embodiment the video data may be adapted by replacing video data pertaining to said at least one participant with background video data.

This presents a simple way for visual muting of the participant.

25 The present invention relates as well to embodiments of a video analysis and adaptation device for adapting video data recorded by a camera at a location during a video conference, said video analysis and adaptation device being adapted to receive said video data, to analyze said video data for

-3-

detecting at least one participant of said video conference in said location having performed a predefined gesture and to, upon detecting of said at least one participant having performed said predefined gesture, perform a step of adapting said video data such as to eliminate data relating to said at least one participant from said video data, thereby generating adapted video data and to provide said adapted video data on an output of said video adaptation device.

In a variant the video adaptation device is able to adapt said video data by replacing video data pertaining to said at least one participant with background video data.

10 In another variant the video analysis and adaptation device is further adapted to receive a trigger signal indicative of the presence of said predefined gesture, and to upon receipt of said trigger signal, start detecting said at least one participant having performed said predefined gesture.

15 In another embodiment the video analysis and adaptation device is further adapted to analyze said video data for detecting said predefined gesture, and to upon detection of said predefined gesture, start detecting said at least one participant performing said predefined gesture.

20 The present invention relates as well to embodiments of a video conferencing client adapted to receive video data from a camera recording a video conference at a location, characterized in that said video conferencing client further comprises a video analysis and adaptation device in accordance to any of the claims 6 to 8, said video conferencing client further being adapted to transmit the adapted video data towards at least one other video conferencing client serving other participants of said video conference at at least one other location.

25 In an embodiment the video conferencing client further comprises registration means being adapted to receive and store user information related to said predefined gesture.

-4-

In a variant said user information comprises gesture information performed by a human on an object.

In another embodiment the video conferencing client is further adapted to receive a trigger detection signal from and transmitted by an object
5 on which said predefined gesture is performed by said at least one participant.

In these variants the video conferencing client is adapted to communicate with such an object such as to detect and recognize said predefined gesture.

The present invention relates as well to embodiments of an object
10 being adapted to detect a predefined gesture performed by at least one participant of a video conference at a location, said object further being adapted to generate a trigger detection signal upon detection of said predefined gesture, and to provide said trigger detection signal to a video conferencing client in said location .

15 In an embodiment the object is further adapted to generate and transmit a registration request related to said predefined gesture to said video conferencing client .

In yet other embodiments said object is a portable communication device, such that said video conferencing client is able to receive from said
20 portable communication device a registration request for providing said user information related to said predefined gesture possibly performed by the participant of said video conference handling said portable communication device during the time of the video conference.

In some variants the object comprises a communication unit and a
25 movement detector or touch sensor.

In an embodiment said object may be a portable communication device such as a mobile phone, a game console, a tablet computer, a laptop etc.

-5-

Alternatively any tangible commodity object, e.g. a toaster, a coffee machine,... equipped with a small communication unit and a touch sensor, can be used in such a location such as a meeting room.

The present invention relates as well to embodiments of a video conferencing server, comprising a video analysis and adaptation device in accordance with any of the previous claims 6-8.

The video conferencing server may further be able to receive from respective video conferencing clients the video data of the conference participants in these respective locations.

10 In a variant these video conferencing clients can also transmit the predefined gesture information towards the server. In other embodiments these predefined gestures can be the same for all clients, and can be centrally stored within the server. In some embodiments the conferencing server communicates said predefined gesture information towards the video conferencing clients in the
15 different locations. For these embodiments, embodiments of the video conferencing clients are adapted to, upon detection of the predefined gesture being performed by a participant, to transmit a signal to the video conferencing server, indicative of the gesture being performed. The server can then further analyze the video images received from this particular client, and adapt the
20 video data accordingly.

The present invention relates as well to a computer program product comprising software adapted to perform the method steps in accordance to any of the claims 1 to 5, when executed on a data-processing apparatus.

It is to be noticed that the term 'coupled', used in the claims, should
25 not be interpreted as being limitative to direct connections only. Thus, the scope of the expression 'a device A coupled to a device B' should not be limited to devices or systems wherein an output of device A is directly connected to an input

-6-

of device B. It means that there exists a path between an output of A and an input of B which may be a path including other devices or means.

It is to be noticed that the term 'comprising', used in the claims, should not be interpreted as being limitative to the means listed thereafter. Thus,
5 the scope of the expression 'a device comprising means A and B' should not be limited to devices consisting only of components A and B. It means that with respect to the present invention, the only relevant components of the device are A and B.

The above and other objects and features of the invention will become
10 more apparent and the invention itself will be best understood by referring to the following description of an embodiment taken in conjunction with the accompanying drawings wherein:

Fig. 1 shows an exemplary situation of a video conference taking place between n locations, with a distributed video conferencing client
15 architecture

Fig. 2 shows a detailed embodiment of a method for visual muting for a distributed video conferencing client architecture,

Fig. 3 shows a detailed embodiment of a method for visual muting for a client-server video conferencing architecture,

20 Fig. 4 shows an exemplary situation of a video conference taking place between n locations, with a client/server video conferencing architecture,

Fig. 5 illustrates a first variant for registration of a gesture performed by a conference participant at a particular location,

25 Fig. 6 illustrates another variant for registration of a gesture performed by m conference participants at a particular location ,

Fig. 7 shows an embodiment of a video conferencing server in the architecture depicted in Fig. 4.

-7-

Fig. 1 depicts a situation of a video conference taking place between n different locations, indicated by room1, room2 to room n. Despite the fact that these locations are denoted as rooms, such a location can as well be somewhere outdoors, or in an indoor location which may be different from a "room", e.g. a working floor, a theatre, a garage, etc.. Each location is equipped by one or more video cameras for recording the different participants of the video conference at that particular location. In figure 1 a situation is depicted where each room has one camera, respectively denoted cam1, cam2 to camn. However in other situations several cameras may be used for recording the different participants in one location.

Each of the locations furthermore is equipped with a video conferencing client device, respectively denoted VCC1, VCC2 to VCCn, for the respective locations room 1, room 2 to room n. Such a video conferencing client is adapted to receive the video data recorded by the respective cameras coupled to it, at its respective location, to process the video data e.g. by compressing and encoding the data, or by mixing the data into one composed video in case several cameras are coupled to one video conferencing client, followed by encoding this resulting composed video. The encoded video from one location is then transmitted by the video conferencing client to the other video conferencing clients in the other locations. This also means that a particular video conferencing client, e.g. VCC1, will thus also receive the processed video from locations room 2 to room n and provide these to a display unit D, e.g. a screen at the particular location.

Embodiments of the present invention aim to provide a method for adapting video data recorded by one or more cameras during a video conference at a location such as to hide the presence of a participant of this video conference on this location, upon receiving a trigger by this participant, by means of this participant performing a predetermined gesture. This gesture can

-8-

be anything, e.g. waving with a left hand, turning a wallet upside down, pushing a chair, turning over or rotating a mobile device equipped with some movement detection and communication capabilities such as a game console, a tablet computer, a cell phone, etc. In one embodiment, where only one gesture is to be used for any participant of a certain location to mute himself/herself, this particular gesture is agreed upon by all participants within a certain location, and is registered by the video conferencing client at this location. In these embodiments each location can have a separate location-specific agreed upon gesture which will trigger the visual mute. Alternatively a video conferencing system with a more centralized approach, may use an initially defined or preconfigured gesture which is the same across all locations. In such situations each of the video conferencing clients may be preconfigured to be triggered by this gesture. In this case the step of registering a predefined gesture possibly to be performed by a participant of the video conference on the particular location, can merely reduce to e.g. the storage of this gesture within the different video conferencing clients. In the earlier described situation, where the participants may initially agree upon the gesture to be recognized per location, the registration of this gesture is to be performed in an initializing step.

In other, more complex embodiments several gestures may be registered, enabling each person to mute him/herself based on a different gesture. In these embodiments the registration procedure will however be more complex.

In the embodiment on Fig. 1 where only one common gesture per location is used, the registration of these different gestures is schematically denoted by means of different inputs G1, G2 to Gn, denoting respective gesture information per respective location room1, room2 to roomn, to the respective video conferencing clients VCC1, VCC2 to VCCn in these respective locations. This may comprise recording an act of performing this gesture by any participant

by a camera at said location. It is to be remarked that during the registration, only the gesture is to be registered, not the person performing it, unless each person will be enabled to perform a different gesture for his/her muting. In the embodiment of only one gesture enabling the visual mute per location, the recorded video information with a person performing this location-specific gesture, will be provided by the camera to the associated video conferencing client of this location. This is for instance depicted in Fig. 5, where the gesture of waving a hand will be recorded and registered. G1 is then the video information of any person waving his/her hand. In this situation the gesture information G1, for room1, is thus first provided by any participant performing the gesture, and recorded by the camera cam1, which will then forward the images containing the gesture information input to the video conferencing client VCC1.

For the more complex embodiments where each person may use a different gesture for being muted, each of these persons then has to register with his/her particular gesture.

Within the video conferencing client a gesture registration means, denoted GR1 for VCC1, can be present to receive and store such information related to the predefined gesture G1. In the aforementioned embodiments this registration means can thus also receive video information from the camera with the gesture to be recognized later on. However in other embodiments such a separate gesture registration means is not necessary and its functionality can be incorporated within a memory or processing unit of the VCC itself.

In these embodiments, depicted in Figs 1 and 5, with only one common registered gesture per location, which gesture is communicated to the system by means of a video of the predefined gesture, the video conference will start after the registration. The camera will continuously record the video conference, and the images will be continuously analyzed within the VCC, for detecting the particular gesture. Upon detection thereof, the video data is further

-10-

analyzed to recognize the participant or participants of said video conference at the particular location who are actually performing this gesture. Upon detecting the person(s) performing the particular gesture, visual muting of these people will take place. To accomplish this visual mute, the video data recorded by the camera is adapted such as to eliminate or hide in the recorded video this or these participants. This adaptation may comprise replacing the image pixel data pertaining to these participants with appropriate background video data .

The analysis and adaptation of the conference video data, denoted video room1 and recorded by the camera, denoted cam 1, in Fig. 1, is performed within a video analysis and adaptation device, denoted VAAS. In the situation of Fig. 1 this video analysis and adaptation device is comprised within the respective video conferencing clients. However in other situations, such as the one depicted in Fig. 4, this video analysis and adaptation device can as well be present in a central video conferencing server. This will be explained more into detail in a later paragraph.

The video analysis and adaptation device, denoted VAAS, is generating adapted video data, for possibly being encoded and being transmitted to other participants on other locations participating to said video conference. In the embodiment depicted in fig. 1, the adapted video is denoted videoa room 1, generated by the VAAS.

For the situation depicted in Fig. 1, upon having adapted the video data of room 1, VAAS will provide the thus adapted video denoted videoa room1, to a transmission unit VT1 of the video conferencing client VCC1. VT1 may perform compression and encoding, and transmit the possibly encoded adapted video data towards the other video conferencing clients VCC2 to VCCn serving other participants of the video conference at the other locations .

Depending upon the gesture itself, its detection can thus be performed by a mere analysis of the conference video data . This is for instance the case if

-11-

the gesture relates to a gesture to be performed solely by the human body itself, e.g. waving with a hand (as was depicted in Fig. 5), bowing the head etc., or with relation to an object which has no further means of communication such as ticking a part of a wooden table, turning a leather wallet...

5 Alternatively the predefined gesture may be detected by the presence of a signal transmitted by an object upon being touched by a participant of the video conference. This may for instance be the case if the gesture relates to touching a particular button on a device with some motion sensing and communication capabilities such as a mobile phone, laptop,.. turning a mobile
10 phone or tablet computer, touching a watch with touch screen and communication capabilities such as bluetooth, etc.

 In these situations the predefined gesture can be simply detected by the object itself, e.g. in case this object is equipped with a movement sensor. Upon detection of this gesture, the object will then generate and transmit a
15 particular trigger detection signal to the video conferencing client. In some cases e.g. when each video conferencing client comprises a video analysis and adaptation device VAAS, this object can also directly transmit this trigger detection signal to this VAAS. In such situations the registration of the gesture will then reduce to an initial transmission of such a signal by this particular object to
20 the VCC or VAAS, which will then accordingly store this. This situation is shown in Fig. 6 where turning respective mobile phones O11 to O1m, upside down by these respective participants 1 to m, will generate respective trigger detection signals G11 to G1m by these respective mobile phones, which will all be initially communicated from these mobile phones to the VCCs. Alternatively each video
25 conferencing client already may be pre-configured such as to immediately recognize such a signal being transmitted by such an object. In this situation, as each participant 1 to m is linked to his/her particular mobile phone, the trigger detection signal will also comprise information about the person having

-12-

performed the predefined gesture. This situation is different from the case where e.g. a smart coffee machine can also send a trigger detection signal upon being touched by any participant at the conference location. In these cases a separate detection of the person having performed the particular gesture, is to be performed.

The trigger detection signal can thus be a signal transmitted by the object such as a cell phone, which is generated by this object upon detecting the particular movement by this object (e.g. turning over or turning upside down). In another example it can be signal transmitted by an object e.g. a watch upon detecting, by this watch, that a particular button is pushed, etc. Of course all types of objects having these capabilities can be used to this purpose.

For all such cases, where a gesture detection signal is generated by an object, the video analysis and adaptation device will then only analyze and adapt the video from the camera, upon being triggered by a trigger signal. This trigger signal can either be transmitted from the VCC to the VAAS, which signal is itself generated by the VCC upon having received a trigger detection signal by a touched upon object. Alternatively the trigger signal can be directly received from this object by the VAAS. As previously mentioned, in a decentralized situation, such as depicted in fig. 1, the object can also directly communicate with a VAAS, present in each VCC, and thus directly trigger a VAAS. In more centralized situations, such as depicted in Fig. 4, the object can only communicate with a VCC as this is still decentralized and present per location. VCC will then accordingly provide an additional trigger signal to a VAAS in a server.

In a particular implementation, the object communicating with the video conferencing client can thus be a portable communication device, such as a mobile phone, a laptop, a game console etc. Such devices are then adapted to generate, e.g. before the start of the video conference, a registration request for

-13-

providing user information related to the predefined gesture possibly performed by a participant, e.g. a gesture of turning this device upside down. The generation of such a registration request then implies a detection of this particular movement by this device, followed by the generation of a particular trigger detection signal to the video conferencing client at the particular location.

In case every participant at that location wants to enjoy this feature, all these devices then need to send their registration signal to the conferencing client. This was schematically indicated in fig. 6, where participant 1 to participant m can have his/her cell phone turned upside down, which is followed by the generation of m trigger signals G11 to G1m.

In this embodiment, VAAS will thus immediately recognize the person having performed the gesture, as each person is linked to his/her personal object, with a specific trigger.

In the other embodiments, where there is no direct registration of which mobile phone belongs to which user, and any gesture detection signal from any mobile phone may lead to a visual muting of the person having touched this mobile phone, a detection of the particular person having performed the gesture is still needed, by means of the analysis of the video conferencing data.

Fig. 2 schematically shows a high level embodiment of the method, for the distributed configuration of fig. 1.

For a more centralized situation, where the VAAS is part of a video conferencing server which is depicted in fig. 4, an embodiment of the method can have the steps as depicted in Fig. 3.

In other embodiments somehow mixed architectures may exist where on one location a video conferencing client will also incorporate the features of a video conferencing server, for the video conferencing clients at the other locations.

-14-

An more detailed embodiment of another embodiment of a video conferencing server is depicted in fig. 7. The video conference server comprises a central gesture registration module, denoted GR, a central gesture detector module GD, a visual mute region selector VMRS, and an output composition module OC. The gesture registration module GR accepts and stores the registered gestures in a central database or memory, denoted by the cylinder block "gesture information" GI. The gesture detection module in the embodiment depicted in Fig. 7 is adapted to received separate gesture input signals, e.g. coming from gyroscopic detectors, coupled to communication units in objects. GD will, upon receipt thereof check whether they correspond with the previously registered gesture information from the memory or database. If it turns out that a particular gesture is present, enabling visual mute of the person having performed this gesture, GD will subsequently analyze this signal, and/or the received video for checking which person actually performed this gesture, if this was not clear from the gesture input signal itself. Once the participant is recognized, this information will be forwarded to the visual mute region selector VMRS, which module will mute the detected participant, from the camera video input received from a particular location. Within module OC the muted or non-muted video inputs from all cameras are again put together to make a general output composed video, which will then again be distributed to the different individual locations.

In yet other embodiments it may even be possible to enable visual muting of a particular person, by another person performing a predefined gesture. In such cases the gesture registration database will contain cross-reference information about which gesture will trigger which person to be muted.

It should be appreciated by those skilled in the art that any block diagrams herein represent conceptual views of illustrative circuitry embodying the principles of the invention. Similarly, it will be appreciated that any flow charts,

-15-

flow diagrams, state transition diagrams, pseudo code, and the like represent various processes which may be substantially represented in computer readable medium and so executed by a computer or processor, whether or not such computer or processor is explicitly shown.

5 While the principles of the invention have been described above in connection with specific apparatus, it is to be clearly understood that this description is made only by way of example and not as a limitation on the scope of the invention, as defined in the appended claims. In the claims hereof any element expressed as a means for performing a specified function is intended to
10 encompass any way of performing that function. This may include, for example, a combination of electrical or mechanical elements which performs that function or software in any form, including, therefore, firmware, microcode or the like, combined with appropriate circuitry for executing that software to perform the function, as well as mechanical elements coupled to software controlled circuitry,
15 if any. The invention as defined by such claims resides in the fact that the functionalities provided by the various recited means are combined and brought together in the manner which the claims call for, and unless otherwise specifically so defined, any physical structure is of little or no importance to the novelty of the claimed invention. Applicant thus regards any means which can provide those
20 functionalities as equivalent as those shown herein.

- 16 -

CLAIMS

1. A method for adapting video data (video room1) recorded by a camera on a location (room1) during a video conference such as to hide the presence of a participant at said location of said video conference, said method
5 comprising a step of registering a predefined gesture possibly to be performed by any participant of said video conference at said location (room1) , a step of detecting said gesture, and upon detection thereof, identifying the at least one participant having performed said gesture at said location (room1), adapting
10 said video data such as to eliminate data relating to said at least one participant having performed said gesture from said video data, thereby generating adapted video data (videoa room1), for being transmitted to other participants of said video conference on other locations (room2,...,roomn) .

2. Method according to claim 1 wherein said predefined gesture (G1)
15 is detected by analyzing said video data (video room1) .

3. Method according to claim 1 wherein said predefined gesture is detected by means of receiving a trigger detection (G11, G1m) signal from and transmitted by an object (O11, O1m) on which said predefined gesture is
20 performed by said at least one participant.

4. Method according to any of the claims 1-3 wherein the identification of said at least one participant having performed said gesture at said location (room1) is performed by analyzing said video data (video room1) of
25 said video conference at said location.

5. Method according to claim 3 wherein the identification of said at least one participant having performed said gesture at said location (room1) is performed by analyzing said trigger detection signal from said object (O11,
30 O1m).

- 17 -

6. Video analysis and adaptation device (VAAS) for adapting video data (video room1) recorded by a camera (cam1) at a location (room1) during a video conference, said video analysis and adaptation device (VAAS) being adapted to receive said video data (video room 1), to analyze said video data (video room 1) for detecting at least one participant of said video conference in said location having performed a predefined gesture and to, upon detecting of said at least one participant having performed said predefined gesture, perform a step of adapting said video data (video room 1) such as to eliminate video data relating to said at least one participant from said video data, thereby generating adapted video data (videoa room1) and to provide said adapted video data (videoa room 1) on an output of said video adaptation device.

7. Video analysis and adaptation device (VAAS) according to claim 6 further being adapted to receive a trigger signal indicative of the presence of said predefined gesture, and to upon receipt of said trigger signal, start detecting said at least one participant having performed said predefined gesture.

8. Video analysis and adaptation device (VAAS) according to claim 6 further being adapted to analyze said video data (video room1) for detecting said predefined gesture (G1), and to upon detection of said predefined gesture, start detecting said at least one participant performing said predefined gesture.

9. Video conferencing client (VCC1) adapted to receive video data (video room 1) from a camera (cam1) recording a video conference at a location (room1), characterized in that said video conferencing client (VCC1) further comprises a video analysis and adaptation device (VAAS) in accordance to any of the claims 6 to 8 and , said video conferencing client (VCC1) further being adapted to transmit the adapted video data (videoa room 1) towards at least one

- 18 -

other video conferencing client (VCC2; VCCn) serving other participants of said video conference at at least one other location (room2; roomn).

10. Video conferencing client (VCC1) according to claim 9 further
5 comprising registration means (GR1) being adapted to receive and store user information related to said predefined gesture .

11. Video conferencing client (VCC1) according to claims 9 or 10,
further being adapted to receive a trigger detection (G11, G1m) signal from and
10 transmitted by an object (O11, O1m) on which said predefined gesture is performed by said at least one participant.

12. Object (O11, O1m) being adapted to detect a predefined
gesture performed by at least one participant of a video conference at a location
15 (room1), said object further being adapted to generate a trigger detection signal (G11, G1m) upon detection of said predefined gesture, and to provide said trigger detection signal to a video conferencing client (VCC1) in said location (room1).

20 13. Object (O11, O1m) according to claim 12 further being adapted to generate and transmit a registration request related to said predefined gesture to said video conferencing client (VCC1).

14. Video conferencing server, characterized in that it comprises a
25 video adaptation device according to any of the previous claims 6-8.

15. A computer program product comprising software adapted to perform the method steps in accordance to any of the claims 1 to 5, when executed on a data-processing apparatus.

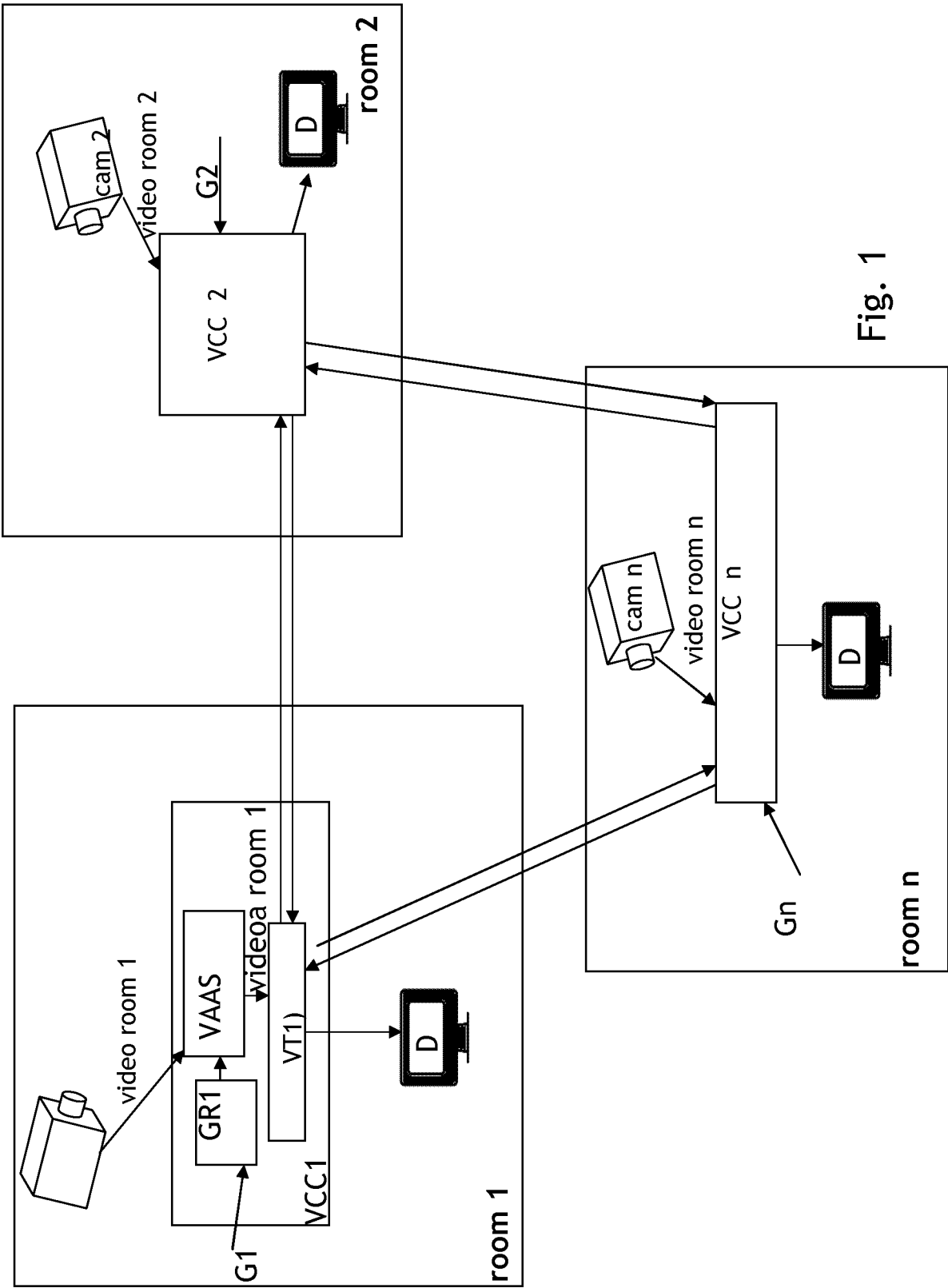


Fig. 1

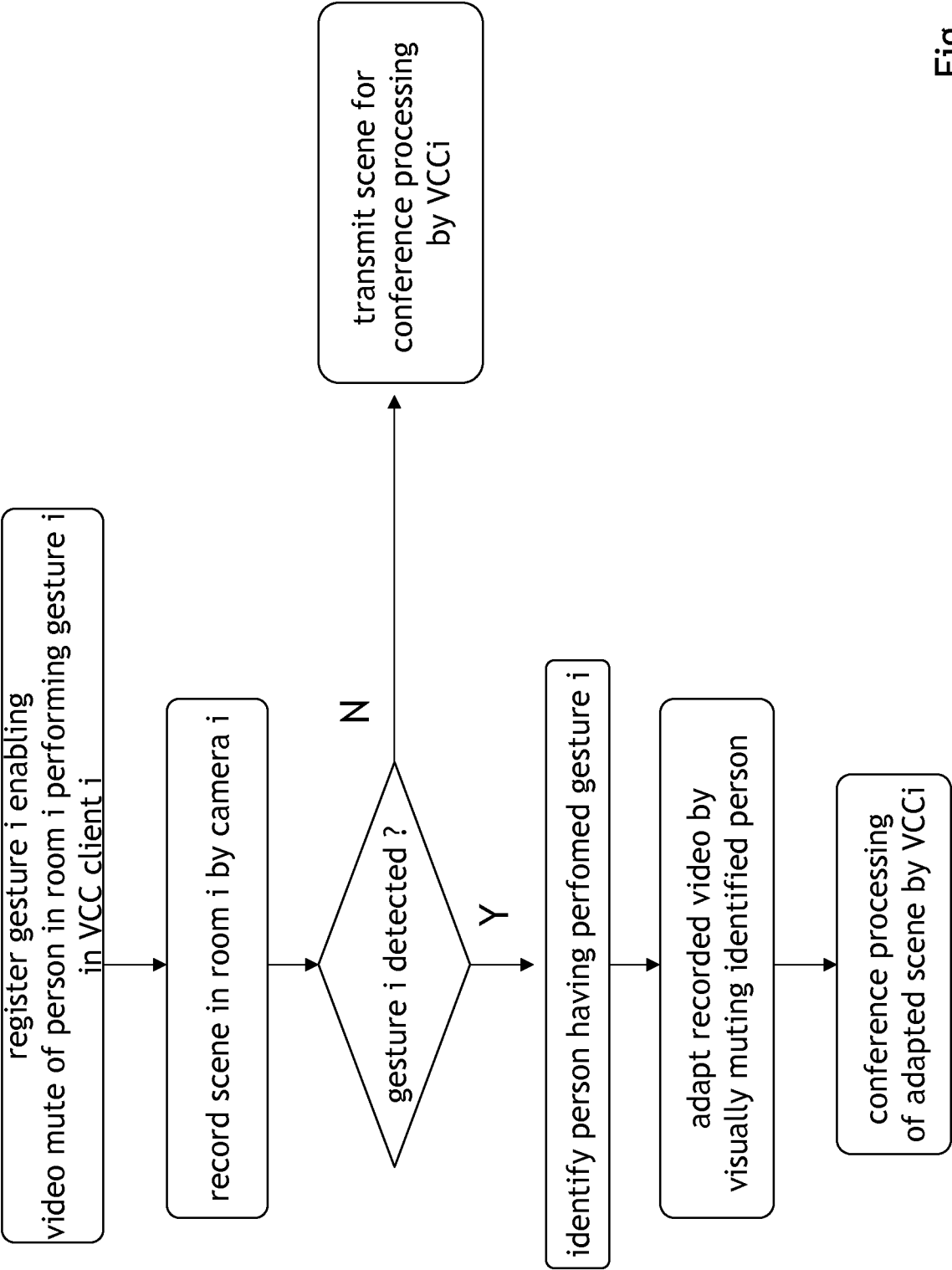


Fig. 2

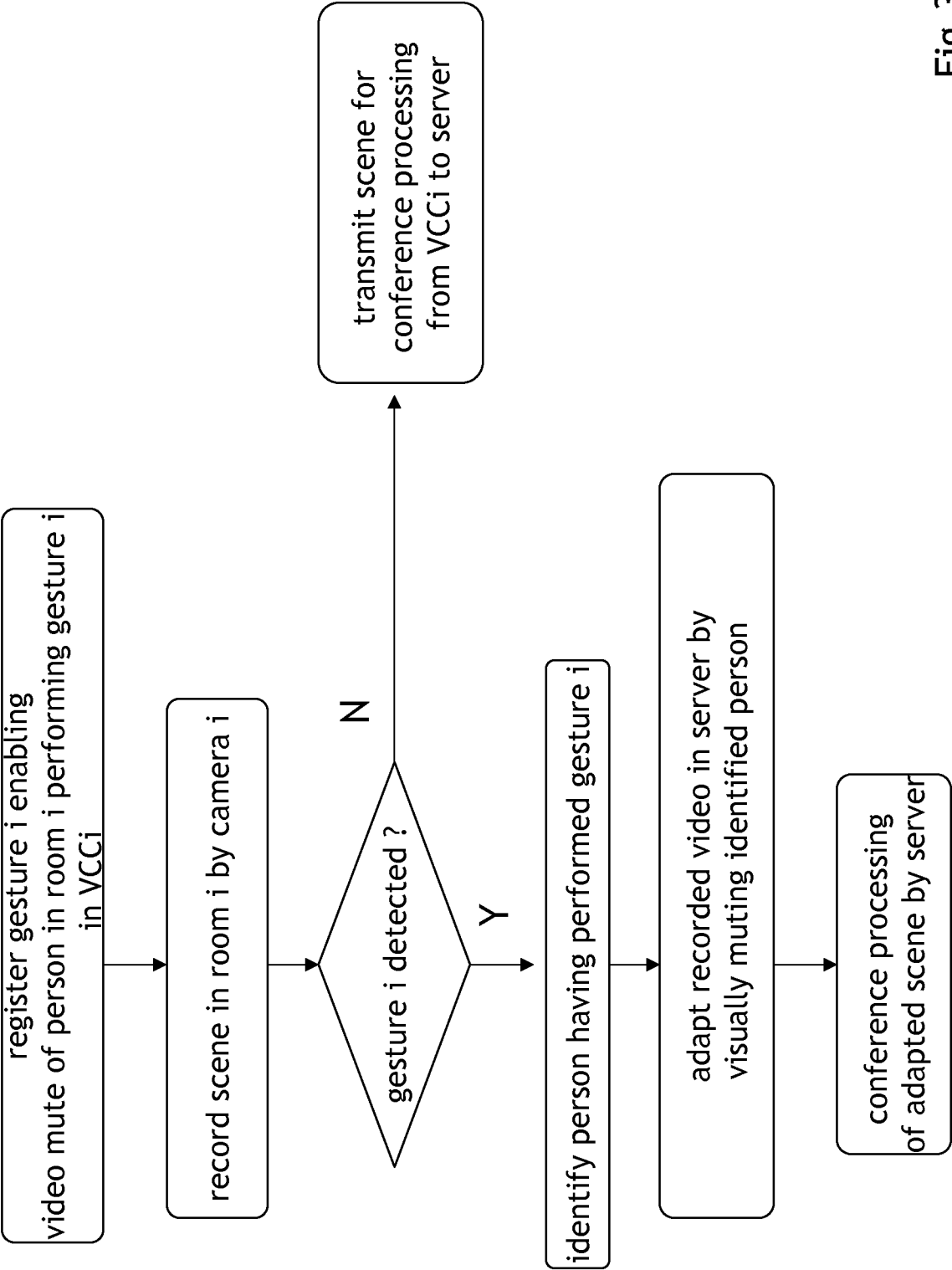


Fig.3

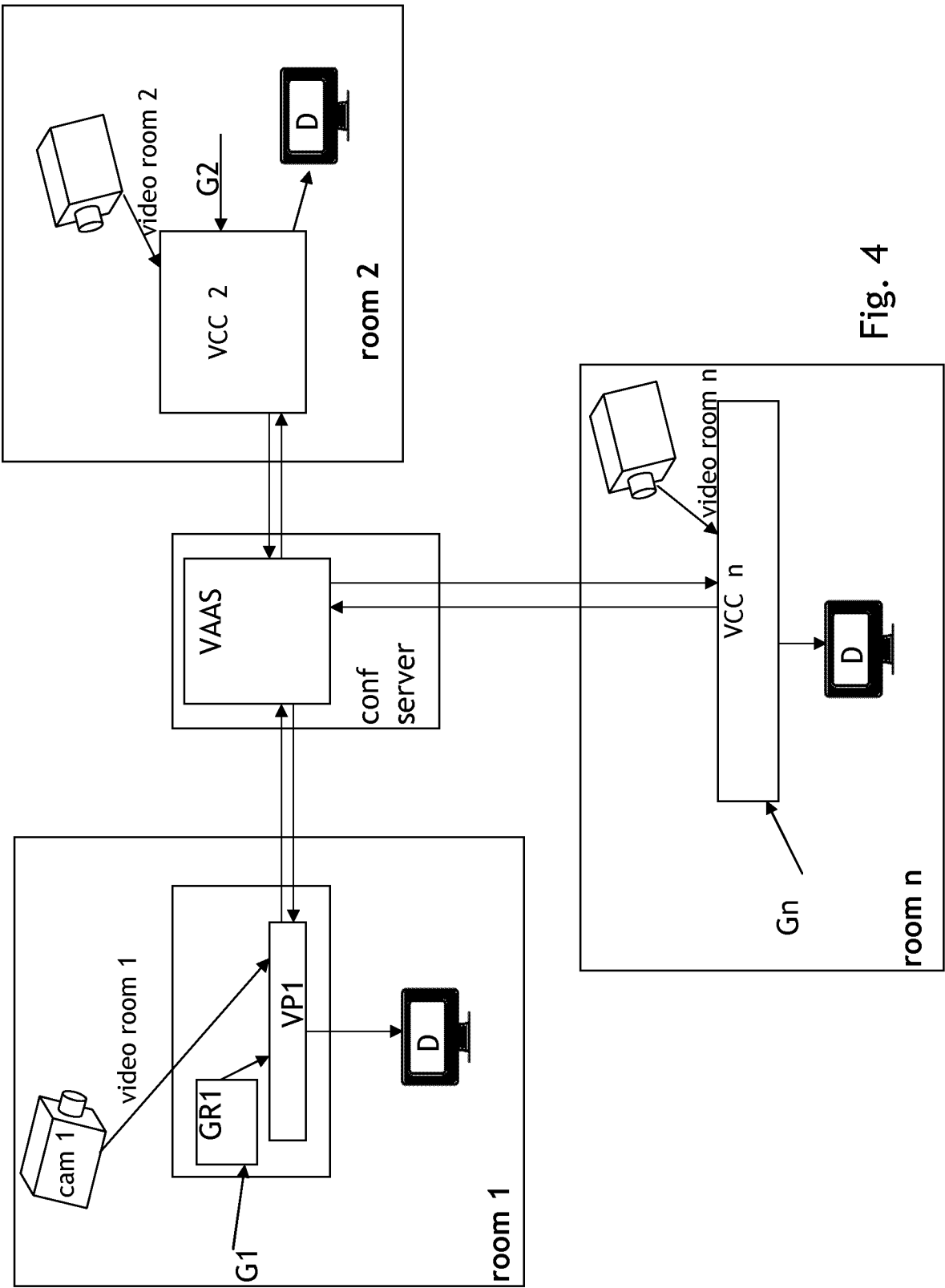


Fig. 4

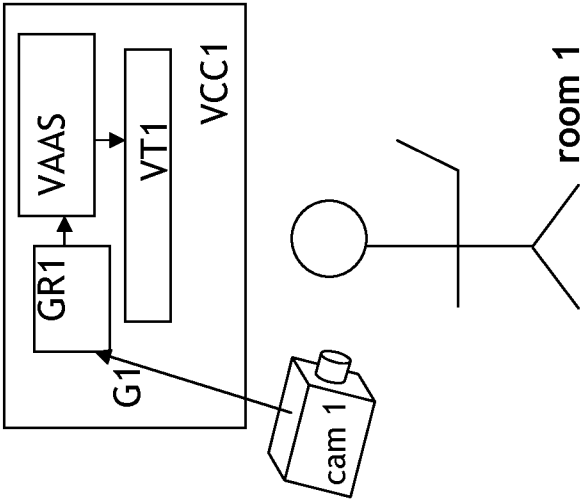


Fig. 5

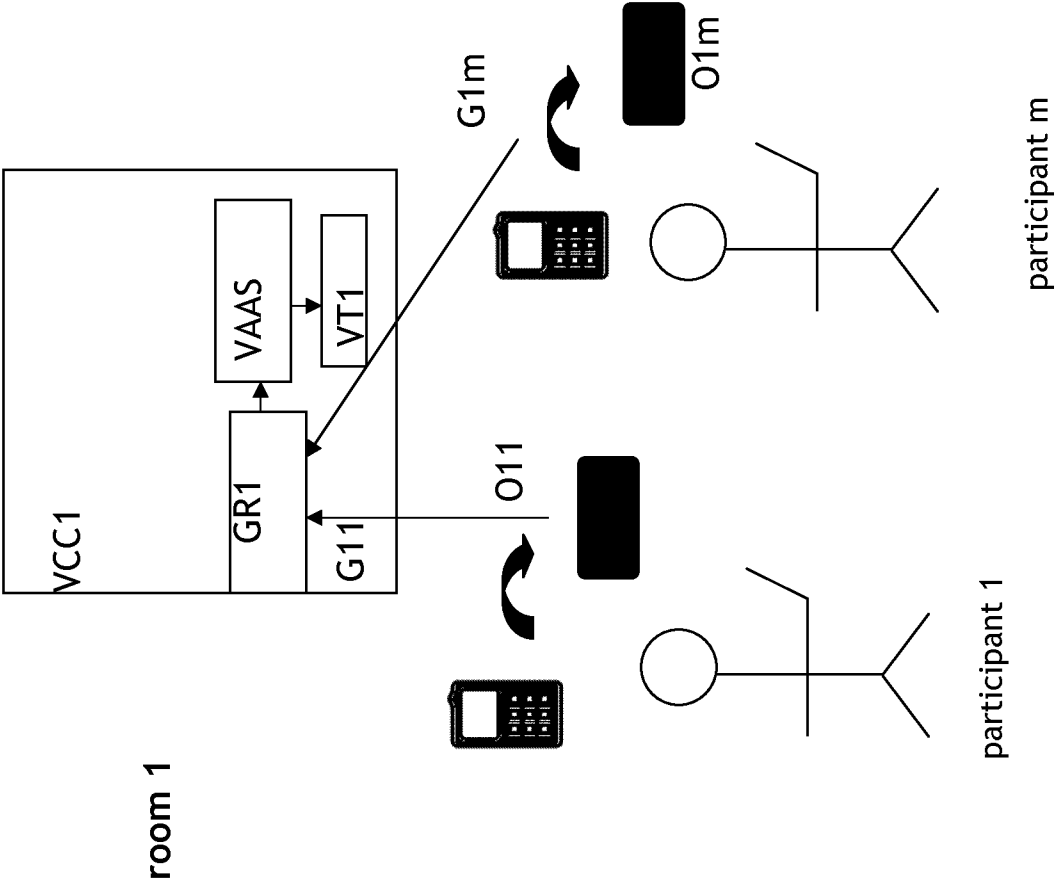


Fig. 6

-7/7-

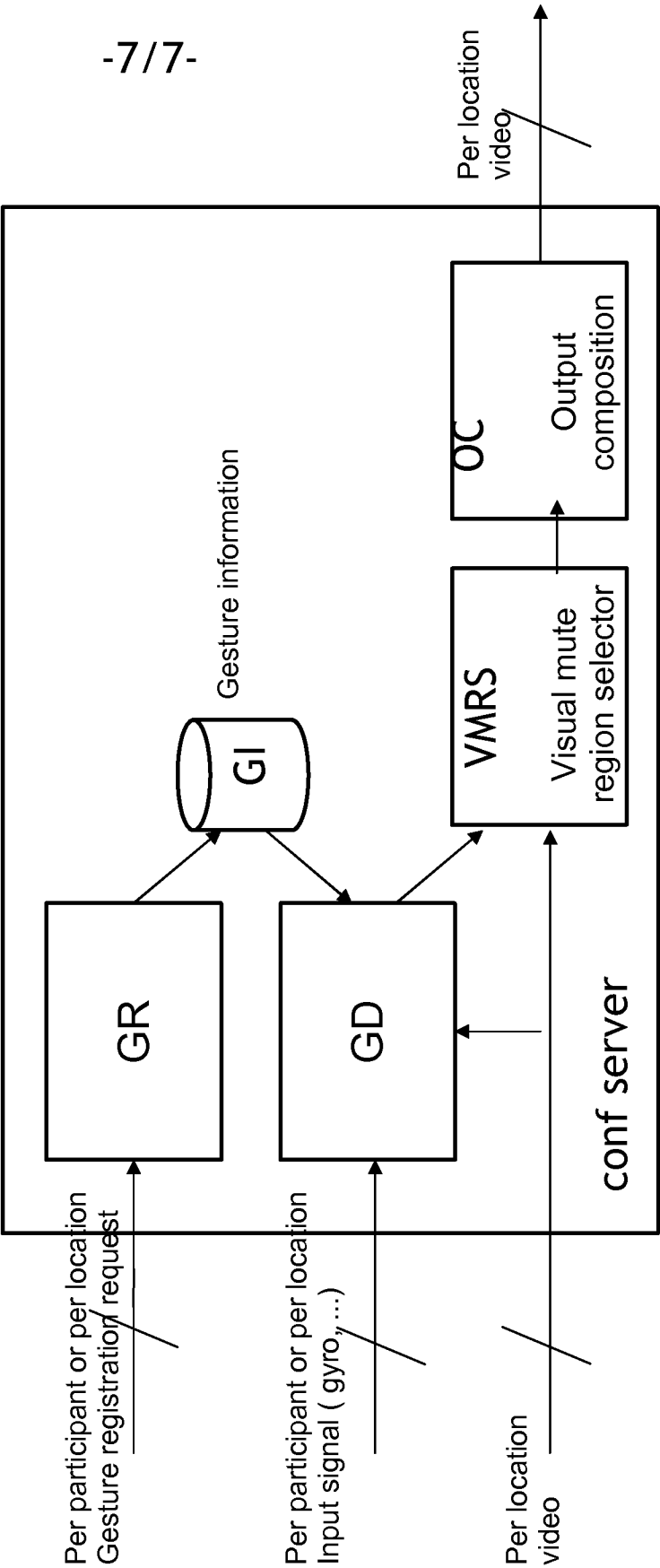


Fig. 7

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2013/065999

A. CLASSIFICATION OF SUBJECT MATTER

INV. H04N7/15
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)

H04N G06F

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)

EPO-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.
A	WO 2007/074410 A2 (FRANCE TELECOM [FR]; WALLINGTON JOHN [US]; LAMADON THIBAUT [US]; CHES) 5 July 2007 (2007-07-05) page 12, line 14 - line 22 page 18, line 21 - line 27 page 31, line 19 - page 32, line 24; figures 3,3a,4 page 35, line 17 - line 28 -----	1-11,14, 15
A	US 7 564 476 B1 (COUGHLAN MARC WILLIAM JOSEPH [AU] ET AL) 21 July 2009 (2009-07-21) column 8, line 4 - line 60; figures 3,4 -----	1-11,14, 15
A	US 2005/078172 A1 (HARVILLE MICHAEL [US] ET AL) 14 April 2005 (2005-04-14) paragraph [0019] - paragraph [0020] paragraph [0066] - paragraph [0075] ----- -/-	1-11,14, 15



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

3 October 2013

Date of mailing of the international search report

08/01/2014

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

Beaudoin, Olivier

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2013/065999

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2009/067670 A1 (GESTURETEK INC [US]; HILDRETH EVAN [CA]) 28 May 2009 (2009-05-28) paragraph [0153] -----	1-11,14, 15
A	US 2010/188473 A1 (KING KEITH C [US] ET AL) 29 July 2010 (2010-07-29) paragraph [0050] - paragraph [0051]; figure 4 paragraph [0062] -----	1-11,14, 15

INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP2013/065999

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-11, 14, 15

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-11, 14, 15

Method, devices, server and computer program product providing a visual mute of a participant of a video conference in response to the detection of a predefined gesture performed by this participant.

2. claims: 12, 13

Object providing a trigger detection signal to a video conferencing client in response to the detection of a predefined gesture.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2013/065999

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2007074410 A2	05-07-2007	US 2007153091 A1 WO 2007074410 A2	05-07-2007 05-07-2007
US 7564476 B1	21-07-2009	NONE	
US 2005078172 A1	14-04-2005	NONE	
WO 2009067670 A1	28-05-2009	CN 101925916 A CN 103369391 A JP 2011504710 A US 2009138805 A1 US 2013329966 A1 WO 2009067670 A1	22-12-2010 23-10-2013 10-02-2011 28-05-2009 12-12-2013 28-05-2009
US 2010188473 A1	29-07-2010	US 2010188473 A1 US 2011254912 A1	29-07-2010 20-10-2011