



(51) International Patent Classification:

H04N 5/265 (2006.01) H04N 21/858 (2011.01)
H04N 21/8549 (2011.01) H04N 5/262 (2006.01)

(21) International Application Number:

PCT/FI2012/051007

(22) International Filing Date:

22 October 2012 (22.10.2012)

(25) Filing Language:

English

(26) Publication Language:

English

(71) Applicant: NOKIA CORPORATION [FI/FI]; Keilalahdentie 4, FI-02150 Espoo (FI).

(72) Inventor: OJANPERÄ, Juha Petteri; Luuppikatu 3, FI-37120 Nokia (FI).

(74) Agents: NOKIA CORPORATION et al.; IPR Department, Jussi Jaatinen, Keilalahdentie 4, FI-02150 Espoo (FI).

(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, 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, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— of inventorship (Rule 4.17(iv))

Published:

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

(54) Title: PERSONALIZED MEDIA REMIX

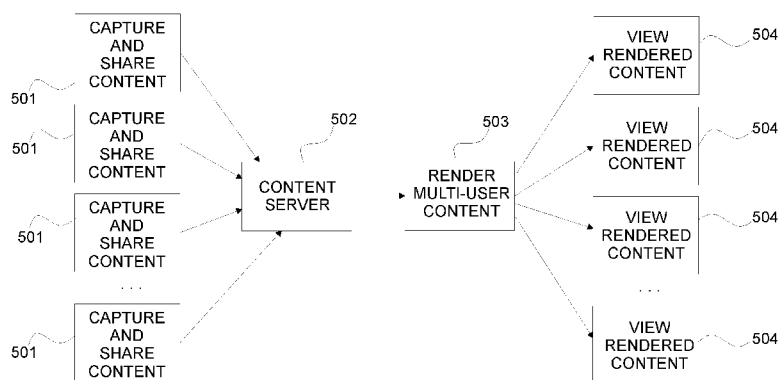


Fig. 5

(57) Abstract: An embodiment of the invention relates to a method comprising receiving media content from at least one recording device, wherein at least one media content received from said at least one recording device is complemented with personating data and creating remixed media content of the media content being received with said at least one personating data. In addition an embodiment of the invention relates to a method comprising capturing media content by a recording device; monitoring the capture of the media content by logging personating data to the recording device and transmitting at least part of the captured media content to a server, which at least part of the captured media is complemented with the personating data. Embodiments of the present invention also relate to a technical equipment for executing the methods.

PERSONALIZED MEDIA REMIX

Technical field

- 5 The present solution relates generally to a method and a technical equipment for creating media remix of a media being recorded by multiple recording devices.

Background

- 10 Multimedia capturing capabilities have become common features in portable devices. Thus, many people tend to record or capture an event, such as a music concert or a sport event, they are attending.

- Media remixing is an application where multiple media recordings are combined in order to obtain a media mix that contains some segments selected from the plurality of media recordings. Video remixing, as such, is one of the basic manual video editing applications, for which various software products and services are already available. Some automatic video remixing systems depend only on the recorded content, while others are capable of utilizing environmental context data that is recorded together with the video content. The context data may be, for example, sensor data received from a compass, an accelerometer, or a gyroscope, and/or location data.

Summary

- 25 Now there has been invented an improved method and technical equipment implementing the method, by which the media remix of a multicaptured media can be personalized for a particular user. Various aspects of the invention include methods, apparatuses, a system and a computer readable medium comprising a computer program stored therein, which are characterized by what is stated in the independent claims. Various embodiments of the invention are disclosed in the dependent claims.

- 35 According to a first aspect, the method comprises receiving media content from at least one recording device, wherein at least one media content received from said at least one recording device is complemented with personating data; creating

remixed media content of the media content being received with said at least one personating data.

5 According to a second aspect, an apparatus comprises at least one processor, memory including computer program code, the memory and the computer program code configured to, with the at least one processor, cause the apparatus to perform at least the following: receive media content from at least one recording device, wherein at least one media content received from said at least one recording device is complemented with personating data; create remixed media
10 content of the media content being received from with said at least one personating data.

15 According to a third aspect, an apparatus comprises at least means for processing, memory means including computer program code, means for receiving media content from at least one recording device, wherein at least one media content from said at least one recording device is complemented with personating data; means for creating remixed media content of the media content being received with said at least one personating data.

20 According to a fourth aspect, a computer program product embodied on a non-transitory computer readable medium, comprising computer program code configured to, when executed on at least one processor, cause an apparatus or a system to: receive media content from at least one recording device, wherein at least one media content received from said at least one recording device is
25 complemented with personating data; create remixed media content of the media content being received with said at least one personating data.

30 According to a fifth aspect, a computer program product embodied on a non-transitory computer readable medium comprising computer program code for user with a computer, the computer program code comprising code for receiving media content from at least one recording device, wherein at least one media content received from said at least one recording device is complemented with personating data; code for creating remixed media content of the media content being received with said at least one personating data.

35

According to an embodiment, a request from a user is received to provide a remixed media content to said user.

5 According to an embodiment, a mood of the user is analyzed by means of the received face image.

10 According to an embodiment the received media content is at least partly video content, wherein video content received from multiple recording devices is examined to find such content that comprises data corresponding to the face image.

According to an embodiment, a cluster is created for recording devices sharing a common grouping factor.

15 According to an embodiment, for examining the video content received from multiple recording devices to find such content that comprises data corresponding to the face image, such video content is selected from the video content received from multiple recording devices that has been recorded by recording devices belonging to a same cluster with the recording device having provided the face
20 image.

According to an embodiment, the personating data is the personating data of the requesting user.

25 According to an embodiment, the personating data is data on user activities during media capture.

30 According to an embodiment, the personating data is data on activities of the recording device during media capture.

According to an embodiment, the personating data includes a face image of the user of the recording device.

35 According to an embodiment, the grouping factor is an audio, whereby the cluster is created for recording devices sharing a common audio timeline.

According to an embodiment, the grouping factor is a location, whereby the cluster is created for recording devices sharing being close to each other.

5 According to a sixth aspect, a method comprises capturing media content by a recording device; monitoring the capture of the media content by logging personating data to the recording device; transmitting at least part of the captured media content to a server, which at least part of the captured media is complemented with the personating data.

10 According to a seventh aspect, a recording apparatus comprises at least one processor, memory including computer program code, the memory and the computer program code configured to, with the at least one processor, cause the apparatus to perform at least the following: capture media content; monitor the capture of the media content by logging personating data to the recording
15 apparatus; transmit at least part of the captured media content to a server, which at least part of the captured media is complemented with the personating data.

According to an embodiment, the personating data is data on user activities during media capture.

20 According to an embodiment, the personating data is data on activities of the recording device during media capture.

25 According to an embodiment, the personating data includes a face image of the user of the recording device.

According to an embodiment, a media remix is requested from a server with at least said personating data.

30 According to an eighth aspect, a system comprises at least one processor, memory including computer program code, the memory and the computer program code configured to, with the at least one processor, cause the system to perform at least the following: receive media content from at least one recording device, wherein at least one media content received from said at least one
35 recording device is complemented with personating data; create remixed media

content of the media content being received with said at least one personating data.

Description of the Drawings

5

In the following, various embodiments of the invention will be described in more detail with reference to the appended drawings, in which

10

Fig. 1 shows a system and device according to an embodiment;

Fig. 2 shows an apparatus according to an embodiment;

Fig. 3 shows a layout of an apparatus according to an embodiment;

15

Fig. 4 shows a server according to an embodiment;

Fig. 5 shows an embodiment of a media remixing arrangement;

20

Fig. 6 shows a block diagram of an embodiment for a recording device;

Fig. 7a,b show block diagrams of an alternative embodiments for a server;

Fig. 8 shows example of media highlight segments for a media in a timeline;

25

Fig. 9 shows a block diagram of another embodiment for the server;

Fig. 10 shows a block diagram for locating a specified user segments according to an embodiment;

30

Fig. 11 shows an example for Fig. 10;

Fig. 12 shows an example of user positions and capturing direction;

35

Fig. 13 shows a block diagram of an embodiment for creating clusters; and

Fig. 14 shows an embodiment for applying Fig's 13 analysis to media remix.

Description of Example Embodiments

- 5 In the following, several embodiments of the invention will be described in the context of capturing media by multiple devices. In addition, the present embodiments provide a solution to create a media presentation of the recorded media, which presentation is personalized for a certain user.
- 10 As is generally known, many portable devices, such as mobile phones, cameras, and tablets, are provided with high quality cameras, which enable to capture high quality video files and still images. The recorded media content can be transmitted to a specific server configured to perform remixing of such content.
- 15 The media content to be used in media remixing services may comprise at least video content including 3D video content, still images (i.e. pictures), and audio content including multi-channel audio content. The embodiments disclosed herein are mainly described from the viewpoint of creating a video remix from video and audio content of source videos, but the embodiments are not limited to video and
- 20 audio content of source videos, but they can be applied generally to any type of media content.

Figure 1 shows a system and devices according to an embodiment. In Fig. 1, the different devices may be connected via a fixed network 210 such as the Internet

25 or a local area network; or a mobile communication network 220 such as the Global System for Mobile communications (GSM) network, 3rd Generation (3G) network, 3.5th Generation (3.5G) network, 4th Generation (4G) network, Wireless Local Area Network (WLAN), Bluetooth®, or other contemporary and future networks. Different networks are connected to each other by means of a

30 communication interface 280. The networks comprise network elements such as routers and switches to handle data (not shown), and communication interfaces such as the base stations 230 and 231 in order for providing access for the different devices to the network, and the base stations 230, 231 are themselves connected to the mobile network 220 via a fixed connection 276 or a wireless

35 connection 277.

There may be a number of servers connected to the network, and in the example of Fig. 1 are shown servers 240, 241 and 242, each connected to the mobile network 220, which servers may be arranged to operate as computing nodes (i.e. to form a cluster of computing nodes or a so-called server farm) for the automatic video remixing service. Some of the above devices, for example the computers 240, 241, 242 may be such that they are arranged to make up a connection to the Internet with the communication elements residing in the fixed network 210.

There are also a number of end-user devices such as mobile phones and smart phones 251, Internet access devices (Internet tablets) 250, personal computers 260 of various sizes and formats, televisions and other viewing devices 261, video decoders and players 262, as well as video cameras 263 and other encoders. These devices 250, 251, 260, 261, 262 and 263 can also be made of multiple parts. The various devices may be connected to the networks 210 and 220 via communication connections such as a fixed connection 270, 271, 272 and 280 to the internet, a wireless connection 273 to the internet 210, a fixed connection 275 to the mobile network 220, and a wireless connection 278, 279 and 282 to the mobile network 220. The connections 271-282 are implemented by means of communication interfaces at the respective ends of the communication connection.

Figures 2—4 show devices for video remixing according to an example embodiment. As shown in Fig. 4, the server 240 contains memory 245, one or more processors 246, 247, and computer program code 248 residing in the memory 245 for implementing, for example, video remixing. The different servers 241, 242 of Fig. 1 may contain at least these elements for employing functionality relevant to each server.

Similarly, the apparatus 151 shown in Figure 2 contains memory 152, at least one processor 153 and 156, and computer program code 154 residing in the memory 152. The apparatus may also have one or more cameras 155 and 159 for capturing image data, for example stereo video. The apparatus may also contain one, two or more microphones 157 and 158 for capturing sound. The apparatus may also contain sensor for generating sensor data relating to the apparatus' relationship to the surroundings. The apparatus may also comprise a display 160 for viewing single-view, stereoscopic (2-view) or multiview (more-than-2-view) images. The display 160 may be extended at least partly on the back cover of the

apparatus. The apparatus 151 may also comprise an interface means (e.g. a user interface) which allows a user to interact with the apparatus. The user interface means may be implemented using the display 160, a keypad 161, voice control, or other structures. The apparatus may also be connected to another device e.g. by means of a communication block (not shown in Fig. 2) able to receive and/or transmit information.

Figure 3 shows a layout of an apparatus according to an example embodiment. The electronic device 50 may for example be a mobile terminal (e.g. mobile phone, a smart phone, a camera device, a tablet device) or user equipment of a wireless communication system. However, it would be appreciated that embodiments of the invention may be implemented within any electronic device or apparatus which are capable of recording media and transmitting the recorded media to another device, e.g. a server device.

The apparatus 50 may comprise a housing 30 for incorporating and protecting the device. The apparatus 50 further may comprise a display 32 in the form of e.g. a liquid crystal display. In other embodiments of the invention the display may be any suitable display technology suitable to display an image or video. The apparatus 50 may further comprise a keypad 34. In other embodiments of the invention any suitable data or user interface mechanism may be employed. For example the user interface may be implemented as a virtual keyboard or data entry system as part of a touch-sensitive display. The apparatus may comprise a microphone 36 or any suitable audio input which may be a digital or analogue signal input. The apparatus 50 may further comprise an audio output device which in embodiments of the invention may be any one of: an earpiece 38, speaker, or an analogue audio or digital audio output connection. The apparatus 50 may also comprise a battery 40 (or in other embodiments of the invention the device may be powered by any suitable mobile energy device such as solar cell, fuel cell or clockwork generator). The apparatus may further comprise an infrared port 42 for short range line of sight communication to other devices. In other embodiments the apparatus 50 may further comprise any suitable short range communication solution such as for example a Bluetooth wireless connection or a USB/firewire wired connection. The apparatus 50 may also comprise one or more camera capable of recording or detecting individual frames which are then passed to the codec or controller for processing. In some embodiments of the invention, the

apparatus may receive the video image data for processing from another device prior to transmission and/or storage. In some embodiments of the invention, the apparatus 50 may receive the image either wirelessly or by a wired connection.

- 5 Figure 5 illustrates an embodiment of a media remixing arrangement. The arrangement comprises more than one users (501) that are arbitrarily positioned within the space to capture content from a scene. The users have recording devices, for example mobile terminals shown in Figure 2. The content may be audio only, audio and video, only video, still images or combination of these four.
- 10 The captured content is transmitted (or alternatively stored for later consumption) to a content server (502), such as the one shown in Figure 4, comprising rendering means (503) which provides remixed media signals to end users (504). The remixed media leverages the best media segments from multiple contributing users (501) to provide the best user experience of the multi-user rendered content.
- 15 End users (504) may be users (501) who uploaded content to the server or some other users who just want to view multi-user rendered content from an event. End user may have any electronic device capable of at least receiving media data and playing the media. Examples of such a device are illustrated in Figure 1 (250; 251; 260; 261; 262; 263)
- 20 The present embodiments propose personalizing the media remix such that each contributing user is able to obtain such a media remix where his/her captured media has preference. The personalized media remix can be created to contain such media segments which that are important for the user. These segments
- 25 typically relate to such a situation where the user has experienced strong emotions. Therefore one of the purposes of the present embodiments is to propose an enabler that makes it possible to personalize media remix according to a specific user for the multi-user captured content.
- 30 An embodiment for personalizing media for a multi-user media remix comprises capturing and rendering methods. The capturing method is performed at the recording device, i.e. client device. The rendering method on the other hand may be performed at the server.
- 35 While the recording device is capturing the media content, the recording device is capable of logging and analyzing user activities that occur during capturing. The

user activities can be logged and analyzed by means of sensor data. The user activities may also include logging zoom level data. The user activities may also include front camera analysis of the device for detecting and analyzing user profile. The media highlights are determined for the rendering by means of the data that has been associated with the media, e.g. as metadata. The media segments comprising media highlight(s) can be determined at the recording device or at the server. The media highlights are then rendered to multi-user media remix at the server. When a user requests personalized media remix, the media preference is selected based on user identification. Therefore, a requesting user will receive such a media remix that has been created based on his/her own preferences.

Figure 6 shows a high level block diagram of an embodiment for the recording device. During media capture (610), the activities of the recording device and the user are monitored (620). The monitoring and data may be stored for later rendering and personalization purposes. The device activities can be monitored by storing sensor data (630) during capturing such as gyroscope/accelerometer and compass data. For carrying this out, the electronic device is capable of logging sensor entries at certain rate that corresponds to a time instance within the capturing activity. For example, compass data may be logged at 10Hz rate, whereby 10 compass sensor entries are obtained per second that describe the user activities during capturing.

Also other activities relating to recording may be stored, such as orientation of the device, time instances when user is zooming along with the zooming level data. The recording device may be capable of logging the zooming time instance and related data in the following format

time_instant, zduration, zlevel

where *time_instant* is the time instant of the start of the zooming measured from the start of the capturing, *zduration* is the time duration user is capturing at the specified zoom level, and *zlevel* is then the actual zoom level.

In addition, the user's moods may be analyzed (540) and in case something relevant is detected in user's mood (such as smiling, laughing, crying, cheering

etc.) those time instants are also stored for later use. The mood analysis can be carried out by analyzing image data captured by a front camera of the recording device.

- 5 The front camera analysis for monitoring and detecting user's mood may be carried out according to following steps

10 1 Take image shot using front camera or alternatively
 extract image from front camera video
 2 Is face included?
 3 Is it user's face?
 4 Detect mood
 5 Known mood detected, log time instant and mood

15

To determine whether the front camera image is user's face (step 3), the user has had to provide a reference image of user's face to the recording device. For detecting the mood any known face recognition methods can be used.

- 20 The front camera analysis may log data in the following format

time_instant, mduration, mood

25 where *time_instant* is the time instant of the start of the analyzed mood, *mduration* is the time duration of the mood, and *mood* is the actual mood that was analyzed. The number of mood to be detected may depend on implementation but e.g. smiling and laughing may indicate strong emotions within that particular time segment during capturing. In addition some other sensor modalities may be used for the detection. For example, the captured audio scene is analyzed to get better
30 confirmation that user is e.g. laughing. In such a case, the audio signal can be classified such that if sound of laughter is detected and also front camera analysis confirms this, then such data entry is logged.

35 It is also possible that the front camera image is recorded to a low resolution video and associated with the main media recording. The actual analysis of the mood may then be determined at the server side. This approach will result in improved

battery lifetime and enables more complex processing as at the server side processing capabilities may be more advanced compared to that of a mobile device side.

- 5 At some point after the media capture has ended the user selects media to be uploaded to the content server side (650).

Figure 7a illustrates a high level block diagram of an embodiment for the server performing at least the rendering functions. The server may also carry out some
10 other functions, which are described later.

At first, a common timeline is created (710) for the participating media. The participating media includes media content being received from plurality of recording devices, wherein the media content relates to a shared experience, e.g.
15 a concert, a sports event, a race, a party etc. Next, media highlights in the media for a particular user are determined (720). This means that any user who has provided media highlights together with the media content, will have his/her own media highlights at the server. The user may be determined by a user identification. For example, when a user is requesting a media remix from the
20 content server, the media preferences may also be signaled by the user. The media preferences may be all the media the user has contributed to a particular event or only a subset of that. The media highlights for the particular user are then determined according to following steps:

- 25 1. For each media in the media preference set the logging data and other associated metadata are analyzed and the time segments that seem to include important media highlights are selected. At least the following time segments are extracted for further highlight processing:

- 30 Detected mood segments (a)
 Zooming segments (b)
 My compass OOI (Orientation-Of-Interest) segments (c)
 My compass non-OOI segments (d)

35 where the orientation-of-interest (OOI) can be determined from the

compass data and it describes the OOI angles (that is, dominant interest points in the compass plane) for the captured media. The non-OOI segments are the opposite of the previous, that is, non-OOI segment describes the interest point in the compass plane that is not dominant in the overall capturing activity but still represents a segment of reasonable duration (e.g. 2-5s at minimum). The non-OOI is an indication that something has activated the user to capture from a certain (deviating) direction which typically indicates important aspect for the user.

There may be overlapping time segments which may then be handled such that certain segment events have higher priority than others. For example, gyroscope/accelerometer may override compass data in case the device is tilted down or up in which case those time segments should not be used (for example, user may be capturing his/her foot for a while which most probably is not interesting event in the users capturing activity).

Finally, the media remix is generated (730). Such a media remix combines the media highlights for at least one particular user and the general multi-user media remix.

As an alternative embodiment, shown in Figure 7b, the general multi-user media remix may be generated first and then the segments (or media views) from the remix are replaced with the media highlight segments (or media views) to personalize the media remix. The rendered media can then be provided to end user consumption.

Figure 8 illustrates the media highlight segments for a media in the timeline. The following highlight segments were identified: two mood segments (a), two zooming segments (b), two OOI segment (c), and one non-OOI segment (d). The lower part of Figure 4 shows the time segments which contain interesting highlights for the selected media. These are the segments which will be used in the media remix. Depending on the duration of the highlight segment, the segment may be used for the entire duration or only a portion of the segment is used. The user can specify how much of his/her content should be used in the media remix. Depending on this value the media remix can adjust the interval when to include the highlight media. For example, it may be possible that in some cases (depending of segment

length) every other view is from the highlight media if that media should appear regularly in the final media remix.

5 In the previous, an embodiment for personalizing media remix according to user experienced highlights was disclosed. Such a media remix can be further personalized by including such segments to the remix that includes video and/or still images of the user. Therefore, the personalized media remix includes highlights for the user but also recordings of the user experiencing the highlights. In order to carry this out, an embodiment of the present invention proposes
10 locating user segments from other user's media. This can be implemented so that front camera shots are taken by the user's recording device during the media capture. Such image shots that includes the face of the user are used as a reference image. The front camera shots can be associated with sensor data such as compass and/or gyroscope/accelerometer data. The front camera shots may
15 also have a timestamp that relates to the start of the media. Yet further, the camera shots may contain one or more still images.

The content of the reference image is searched from other media files taken by other users. The potential other media files from which the content of the reference
20 is searched can be selected by comparing capture times of the media files. The capture time may be included as metadata in a media file. When a set of potential media files are selected, the content of them is examined in order to find a corresponding content with the content of the reference image. As a result of the examination, such media files, which are captured by one or more other users and
25 which comprises a specified user as content are found. After having found media segments including video of the specified user, these media files (partly or in total) can be included in the personalized media remix.

Turning again to Figure 6 illustrating a high level block diagram of an embodiment
30 for the recording device. To utilize the further embodiment for personalization, the shots (e.g. still images) by the front camera (640) are taken at certain time intervals and those time instances along with the (optional) still images are stored for later use as a reference image.

35 Figure 9 illustrates a high level block diagram of the embodiment for the server. At first a common timeline is created (910) for the participating media, i.e. the

captured media received from plurality of recording devices. Next, media segments that includes a specified user as content are determined (920). These segments can be found by comparing the media from other users to the reference image of the specified user. If the media from other users contain the content of the reference image, such media segments are stored for remixing purposes. Further, the determined segments may be extended (930) to cover also such segments or time instances that most likely contain the specified user based on the previous (920) analysis results. Finally, the identified segments are rendered (940) to the media remix. In some situations, the user may only request as the final media remix, only the identified segments relating to the user. Therefore, the server is also capable of creating a media remix comprising only media material of the specific user.

The front camera shots can be analyzed according to following steps in order to create a reference image/video:

```
1   Take image shot using front camera or alternatively
    extract image from front camera video
20  2   Is face included?
    3       Is it user's face?
    4           Store face and timestamp
    5   Go to step 1 if media capturing still active
```

25

For step 3, the user has had to provide a reference image of user's face to the recording device. Otherwise it cannot be determined whether the face is user's.

The front camera analysis makes it sure that the user is at best position in order to be located from other user's media. In an embodiment, also such time instances may be saved, where user's face is not detected. This is because, that may indicate interesting moment for the user in question. In such a case the previous steps 2—4 would be replaced merely with step "store front camera image and timestamp".

35

The front camera may store data in the following format:

time_instant, (face_image)

where *time_instant* is the time instant of the still image with respect to the start of the media capture. The captured face (*face_image*) may be included for each log entry but there may also be only one face image that is shared by all log entries to save storage space. Alternatively, some entries may share one face image, whereas other entries may share another face. The front camera may operate continuously or image shots are taken at fixed or random intervals. It is appreciated that instead of face image (*face_image*), also some other content can be stored with time instant, as mentioned above.

Figure 10 illustrates a block diagram for locating the specified user segments from other user's media. First, the media from specified user is analyzed to see if data relating to front camera shots is included (1010). For this embodiment, the front camera data is used to define a reference image. If such data is present, the other user's media is then located (1020). Such other user's media can be located by determining the overlapping media with respect to the specified user's media, e.g. by comparing capturing times of the reference image and the other user's media. If there is not any front camera data that can be interpreted as reference image, the determination of user segments is terminated for this media. After having the identified media segments from block 1020, each identified media segment is analyzed to see whether the other user having captured the media segment in question is possibly pointing towards the specified user (1030). This can be done by utilizing sensor data being included in the metadata of the media file. If it is determined that the other user is most likely pointing towards the specified user, the final step (1040) is then to confirm this by analyzing the actual media segment and finding the specified user from the media segment. Steps 1030 and 1040 are repeated for each identified media from step 1020. In addition, steps 1010—1040 may be repeated for each media that belongs to the specified user.

Figure 11 illustrates an example for Figure 10. Let m_1 represent one of the media of the specified user. The media has one face related shot at time instant mt_1 . Next, overlapping media is determined using the common timeline and in this case the overlapping media with respect to media m_1 at time instant mt_1 are m_2 and m_3 . After this, it is determined, whether these two media m_2, m_3 are pointing towards

the specified user. For this purpose, the position and sensor data of the media are analyzed by utilizing the metadata of the media files. If the system is able to provide accurate positioning (see Figure 12), this can be used for determining whether the other user (Fig. 12: B) is pointing towards the specified user (Fig. 12: A). If, on the other hand, the positioning is not accurate enough or if the users are closely located (within few meters), the positioning data may be unreliable due to errors in the actual position. Therefore, other techniques may be used to determine the media which include the specified user in the media view. One of such techniques is to determine the direction of capturing for the specified user media and based on this value, the target direction of capturing can be determined for the other user's media. Let cx_t be the capturing direction of the specified user media at time instant mt_l . The target direction of capturing can then be determined according to

$$cDiff = \begin{cases} cx_t - 180^\circ, & cx_t - 180^\circ \geq 0 \\ 360^\circ + (cx_t - 180^\circ), & otherwise \end{cases}$$

$$cThr = cDiff \pm cDev$$

where $cDev$ is the direction angle deviation, for example $\pm 45^\circ$. It can be determined, that the other media points to the specified user if its direction of capturing cy_t at time instant mt_l satisfies the following condition:

$$cThr_{\min} \leq cy_t \leq cThr_{\max}$$

Once it has been verified that other user is pointing towards the specified user, the next step is to verify this from the captured media. This can be realized according to following steps:

- 1 Extract media view
- 2 Is face included?
- 3 Is it specified user's face?
- 4 Direction of capturing verified
- 5 Go to step 1 if media views still available

To ensure efficient operation, only the media views in the vicinity of the specified time instance can be analyzed (in Figure 11 between t_1 and t_2). After above steps have been completed, the rendering server will become aware of the media that includes the specified user.

5

The duration of the media segments including the specified user may be fixed (e.g. $\pm t$ seconds around the time instance mt_1) or determined, e.g. by using object tracking in order to determine, e.g. how long the face/head remains in the view if compass angle stays the same in both media. Furthermore, in order to improve detection robustness, all face image shots can be used until a match is found. In addition, the detection may apply different correction techniques to the uploaded face in case the face image is not exactly matching the direction of capturing in the other user's media.

10

15

It is also possible that the face detection fails to produce positive output (i.e. presence of specified user is not verified). In that case the verification may occur only at the sensor data level and this verification mode can be separately signaled to the rendering server. If the direction of capturing is valid according to above equations, even though the face is not found, the segment can still be marked as "potential face found". There can be couple of levels of potential verifications: 1) the specified user was found from the media but from different position, i.e. at some time instance the verification was successful, but at another time instant of the same media, positive output could not be produced; 2) the specified user was not found from the media at all, but the equations are valid making the chance of the specified user to be present in such media very high. The rendering may then occur such that first the segments with positive output are selected, and if it is required that certain amount of segments comprising the specified user should be present in the media remix, level 1 can be processed next and followed by level 2.

20

25

30

In the previous, a method for locating a specified user from a media being captured by other user's recording devices. In such method, media from all other users may be examined to locate the specified user, or only such media is examined that is captured by such other users that are temporally close enough to the specified user.

35

In addition to these alternatives, yet another possibility to select the media for examination is disclosed next.

5 In this embodiment, only such media is examined for locating a specified user, which is captured by recording devices belonging to same cluster with the specified user. The cluster can be determined according to a grouping factor such as a location being based on e.g. GPS (Global Positioning System), GLONASS (Global Navigation Satellite System), Galileo, Beidou, Cellular Identification (Cell-ID) or A-GPS (Assisted Global Positioning System). In the following, the
10 cluster is created according to a grouping factor being a common audio scene.

Figure 13 illustrates a high level block diagram of an embodiment. Let x_i^t represent media signals for overlapping time segment t with $0 \leq i < N$, where N is the number of signals in the segment. The steps of Figure 13 is applied for each time segment.
15 First, an alignment matrix is determined for the multi-user media (1310), i.e. the media being received from plurality of recording devices. Next, the alignment matrix is mapped to groups of media (1320), in order to find out which media belong to same group. The group structures are analyzed and media which acts as links with other media are determined (1330).

20 The purpose of the alignment matrix is to describe the relation of a signal with respect to the other signals. The audio scene status is a metric that indicates whether the audio scenes of two media are similar.

25 The steps 1310—1330 of Figure 13 are now described in more detail. The matrix entries for the alignment matrix may be determined using time alignment methods known in the art such that matrix entry '1' indicates that the signals share the audio scene that can be aligned, and matrix entry '0' indicates that the signals do not share exactly the same audio scene, that is, the signals may still be from the same
30 audio scene but due to various issues such as different capturing positions and surrounding ambience level at the actual capturing position, the signals do not align. It is realized that the alignment matrix summarizes the audio scene status of a media with respect to the other media.

35 In the following example, the main steps according to an embodiment are described. a , b , c , d and e represents the signals that are part of a time segment.

The alignment matrix after time aligning each signal pair in the group of signals may look as follows:

	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>
<i>a</i>	1	1	0	0	0
<i>b</i>	1	1	1	0	0
<i>c</i>	0	1	1	1	1
<i>d</i>	0	0	1	1	1
<i>e</i>	0	0	1	1	1

5

The signal groups (i.e. groups having aligned signals) are then

(a, b)

(a, b, c)

10

(b, c, d, e)

(c, d, e)

(c, d, e)

As a next step, it needs to be determined which groups can be the basis for the groups by analyzing whether the signal group is a subset of another group. After applying this analysis, the preliminary basis group structure is:

15

(a, b): 2 counts

(a, b, c): 1 count

20

(b, c, d, e): 1 count

(c, d, e): 2 counts

The groups which can be the basis for the groups need to have at least two count instants, whereby the final media grouping is

25

(a, b), (c, d, e)

The next step is to locate the signal that contains (or signals that contain) a link to other signal groups. The final media groups are compared against the preliminary basis groups that contain only single count instance. Thus the comparisons are

30

$$\left\{ \begin{array}{l} (a,b) \text{ vs } (a,b,c) \\ (a,b) \text{ vs } (b,c,d,e) \end{array} \right\} \text{ and } \left\{ \begin{array}{l} (c,d,e) \text{ vs } (a,b,c) \\ (c,d,e) \text{ vs } (b,c,d,e) \end{array} \right\}$$

The final media group needs to be a subset of the signal group to which it is compared against and after eliminating the non-subset groups, the final comparison is as follows:

$$(a,b) \text{ vs } (a,b,c), (c,d,e) \text{ vs } (b,c,d,e)$$

which means that signal that is linking with first group is signal c , and the signal that is linking with the second group is signal b .

The mapping data that is stored for this time segment is therefore

Media groups: (a, b) and (c, d, e)

Linking media: c and b

As a final step of Figure 13, this mapping data is stored (1340) as audio scene mapping index for rendering purposes.

Once the mapping data is available for each time segment, the media switching may take place. Figure 14 illustrates a high level block diagram of an embodiment for applying the previous analysis data to the multi-user media remix.

The first step in the media switching is to locate/determine (1410) the grouping data that contains the currently selected/viewed media. Let the grouping data be y_j with $0 \leq j < M$ where M is the number of signals in the segment. This grouping data is then used in combination with the media selection switch to determine the next media view (1420) to be examined in order to find an image of the specified user. This can be carried out by locating the media group within the grouping data and then determine the next media. To select the media for examination, the selection may follow predefined rules. For example, at certain times (time intervals) the next media view to be selected for examination can be near to the current view (1430). In such case, the media should be selected to be one of the media from the same media group (e.g. current media is a and next media is b). At certain times (time intervals), however, the next media view to be selected for examination can be

from neighbouring media group (1440). In this case, the next media may be selected in such a manner that it is one of the media from some other media group that is selected using the media links (e.g. from media *a* to media *d* where *c* is the linking media in between groups). At certain times (time intervals) the next media
5 for examination can be such that it has minimum distance to the current media view (1450). It is appreciated that other switching logics may be generated by using the audio scene mapping data.

It is also appreciated that a group contains multiple linking media to different
10 groups. The audio scene mapping data effectively clusters the signals that are present in the scene. Signals that appear to be in the vicinity of each other during capturing may get assigned to different groups. Thus, the clusters represents a virtual grouping of the media signals present in the scene and when mapping data is indexed in controlled manner, the end user experience may be better than
15 randomly selecting the media views.

The overall end-to-end framework may be traditional client-server architecture where the server resides at the network or ad-hoc type of architecture where one the capturing devices may act as a server. The previous functions may be shared
20 between the client device and the server device so that the client at least performs the media capturing and detecting the sensor data that can be utilized for giving information of the captured media. In addition, the client device may utilize the front camera to give information on user's moods and/or to provide means to detect user from other user's media. The server device can then perform the
25 rendering of the captured media from plurality of recording devices. For the rendering, the server may use the personalization data received from one or more of the recording devices, so that the media remix will contain user experienced highlights. In addition, the server may use such media that has been captured of the specific user. As a result, the media remix will contain also recording of the
30 user e.g. at the time the user is experiencing the highlights. However, in order to carry this out, the server needs to go through the media views received from other user's. To help this process, one of the present embodiments propose to create clusters by means of e.g. audio to see which users potentially could have media views of the specific user.

There is also few possibilities to create the media remix. For example, user A may request such a media remix that also comprises only such highlights that are specific for user A (i.e. provided by the user A). As an another example, user A may request such a media remix that also comprises highlights of selected users
5 B—D. Yet as an another example, user A may request such media remix that also comprises all the highlights that were obtained together with the media view. These alternatives can be completed with media views being captured of the user A. In another embodiment the user A may also request such media remix that has been created only of such media content that relates to the highlights of the user
10 A. In such a case, the media remix is a personal summary of a complete event.

The various embodiments may provide advantages. For example, personalized media remix can be thought are most valuable and important aspect when rendering multiuser content. The personalization combines different media view,
15 with personalized highlights. In addition, an embodiment of the solution provides computationally efficient personalization that is based on media groups being created according to a time scene. By means of the present embodiments, the user is able to receive personalized media remix that is based on a media being received from multiple recording devices.

The various embodiments of the invention can be implemented with the help of computer program code that resides in a memory and causes the relevant apparatuses to carry out the invention. For example, a device may comprise circuitry and electronics for handling, receiving and transmitting data, computer
20 program code in a memory, and a processor that, when running the computer program code, causes the device to carry out the features of an embodiment. Yet further, a network device like a server may comprise circuitry and electronics for handling, receiving and transmitting data, computer program code in a memory,
25 and a processor that, when running the computer program code, causes the network device to carry out the features of an embodiment.
30

It is obvious that the present invention is not limited solely to the above-presented embodiments, but it can be modified within the scope of the appended claims.

Claims:

1. A method, comprising:
 - 5 – receiving media content from at least one recording device, wherein at least one media content received from said at least one recording device is complemented with personating data;
 - creating remixed media content of the media content being received with said at least one personating data.
- 10 2. The method according to claim 1, further comprising
 - receiving a request from a user to provide a remixed media content to said user.
- 15 3. The method according to claim 2, wherein the personating data is the personating data of the requesting user.
4. The method according to any of the previous claims 1—3, wherein the personating data is data on user activities during media capture.
- 20 5. The method according to any of the previous claims 1—4, wherein the personating data is data on activities of the recording device during media capture.
6. The method according to any of the previous claims 1—5, wherein the
25 personating data includes a face image of the user of the recording device.
7. The method according to claim 6, further comprising
 - 30 - analyzing a mood of the user by means of the received face image.
8. The method according to claim 6, wherein the received media content is at least partly video content, whereby the method comprises
 - 35 - examining the video content received from multiple recording devices to find such content that comprises data corresponding to the face image.

9. The method according to any of the previous claims 1—8, further comprising
- creating a cluster for recording devices sharing a common grouping factor.

5 10. The method according to claim 9, wherein the grouping factor is an audio, whereby the cluster is created for recording devices sharing a common audio timeline.

10 11. The method according to claim 9, wherein the grouping factor is a location, whereby the cluster is created for recording devices being close to each other.

15 12. The method according to any of the claims 9—11, wherein for examining the video content received from multiple recording devices to find such content that comprises data corresponding to the face image the method further comprises

- selecting from the video content received from multiple recording devices such video content that has been recorded by recording devices belonging to a same cluster with the recording device having

20 provided the face image.

13. A method, comprising:

- capturing media content by a recording device;
- monitoring the capture of the media content by

25 ○ logging personating data to the recording device;

- transmitting at least part of the captured media content to a server, which at least part of the captured media is complemented with the personating data.

30 14. The method according to claim 13, wherein the personating data is data on user activities during media capture.

15. The method according to claim 13 or 14, wherein the personating data is data on activities of the recording device during media capture.

16. The method according to claim 13 or 14 or 15, wherein the personating data includes a face image of the user of the recording device.

17. The method according to any of the previous claims 13—16, further comprising

- requesting a media remix from a server with at least said personating data.

18. An apparatus comprising at least one processor, memory including computer program code, the memory and the computer program code configured to, with the at least one processor, cause the apparatus to perform at least the following:

- receive media content from at least one recording device, wherein at least one media content received from said at least one recording device is complemented with personating data;
- create remixed media content of the media content being received from with said at least one personating data.

19. The apparatus according to claim 18, further comprising computer program code configured to, with the processor, cause the apparatus to perform at least the following:

- receive a request from a user to provide a remixed media content to said user.

20. The apparatus according to claim 19, wherein the personating data is the personating data of the requesting user.

21. The apparatus according to any of the previous claims 18—20, wherein the personating data is data on user activities during media capture.

22. The apparatus according to any of the previous claims 18—21, wherein the personating data is data on activities of the recording device during media capture.

23. The apparatus according to any of the previous claims 18—22, wherein the personating data includes a face image of the user of the recording device.

24. The apparatus according to claim 23, further comprising computer program code configured to, with the processor, cause the apparatus to perform at least the following:

- analyze a mood of the user by means of the received face image.

25. The apparatus according to claim 23, wherein the received media content is at least partly video content, whereby the apparatus further comprises computer program code configured to, with the processor, cause the apparatus to perform at least the following:

- examine the video content received from multiple recording devices to find such content that comprises data corresponding to the face image.

26. The apparatus according to any of the previous claims 18--25, further comprising computer program code configured to, with the processor, cause the apparatus to perform at least the following:

- create a cluster for recording devices sharing a common grouping factor.

27. The apparatus according to claim 26, wherein the grouping factor is an audio, whereby the cluster is created for recording devices sharing a common audio timeline.

28. The apparatus according to claim 26, wherein the grouping factor is a location, whereby the cluster is created for recording devices being close to each other.

29. The apparatus according to any of the claims 26—28, wherein for examining the video content received from multiple recording devices to find such content that comprises data corresponding to the face image, the apparatus further comprises computer program code configured to, with the processor, cause the apparatus to perform at least the following:

- select from the video content received from multiple recording devices such video content that has been recorded by recording devices belonging to a same cluster with the recording device having provided the face image.

- 5 30. A recording apparatus comprising at least one processor, memory including computer program code, the memory and the computer program code configured to, with the at least one processor, cause the apparatus to perform at least the following:
- capture media content;
 - monitor the capture of the media content by
 - logging personating data to the recording apparatus;
 - transmit at least part of the captured media content to a server,
 - 10 which at least part of the captured media is complemented with the personating data.
- 15 31. The recording apparatus according to claim 30, wherein the personating data is data on user activities during media capture.
32. The recording apparatus according to claim 30 or 31, wherein the personating data is data on activities of the recording device during media capture.
- 20 33. The recording apparatus according to claim 30 or 31 or 32, wherein the personating data includes a face image of the user of the recording device.
- 25 34. The recording apparatus according to any of the previous claims 30—33, further comprising computer program code configured to, with the processor, cause the apparatus to perform at least the following:
- request a media remix from a server with at least said personating data.
- 30 35. A system comprising at least one processor, memory including computer program code, the memory and the computer program code configured to, with the at least one processor, cause the system to perform at least the following:
- receive media content from at least one recording device, wherein at
 - 35 least one media content received from said at least one recording device is complemented with personating data;

- create remixed media content of the media content being received with said at least one personating data.

36. An apparatus comprising at least

- 5 – means for processing,
- memory means including computer program code,
- means for receiving media content from at least one recording device, wherein at least one media content from said at least one recording device is complemented with personating data;
- 10 – means for creating remixed media content of the media content being received with said at least one personating data.

37. The apparatus according to claim 36, further comprising means for receiving a request from a user to provide a remixed media content to said user.

15

38. The apparatus according to claim 37, wherein the personating data is the personating data of the requesting user.

39. The apparatus according to any of the previous claims 36—38, wherein the personating data is data on user activities during media capture.

20

40. The apparatus according to any of the previous claims 36—39, wherein the personating data is data on activities of the recording device during media capture.

25

41. The apparatus according to any of the previous claims 36—40, wherein the personating data includes a face image of the user of the recording device.

42. The apparatus according to claim 41, further comprising means for analyzing a mood of the user by means of the received face image.

30

43. The apparatus according to claim 41, wherein the received media content is at least partly video content, whereby the apparatus further comprises means for examining the video content received from multiple recording devices to find such content that comprises data corresponding to the face image

35

44. The apparatus according to any of the previous claims 36--43, further comprising means for creating a cluster for recording devices sharing a common grouping factor.

5 45. The apparatus according to claim 44, wherein the grouping factor is an audio, whereby the cluster is created for recording devices sharing a common audio timeline.

10 46. The apparatus according to claim 44, wherein the grouping factor is a location, whereby the cluster is created for recording devices being close to each other.

15 47. The apparatus according to any of the claims 44—46, wherein for examining the video content received from multiple recording devices to find such content that comprises data corresponding to the face image, the apparatus further comprises means for selecting from the video content received from multiple recording devices such video content that has been recorded by recording devices belonging to a same cluster with the recording device having provided the face image.

20 48. A computer program product embodied on a non-transitory computer readable medium, comprising computer program code configured to, when executed on at least one processor, cause an apparatus or a system to:

- 25
- receive media content from at least one recording device, wherein at least one media content received from said at least one recording device is complemented with personating data;
 - create remixed media content of the media content being received with said at least one personating data.

30 49. A computer program product embodied on a non-transitory computer readable medium comprising computer program code for user with a computer, the computer program code comprising

- 35
- code for receiving media content from at least one recording device, wherein at least one media content received from said at least one recording device is complemented with personating data;

- code for creating remixed media content of the media content being received with said at least one personating data.

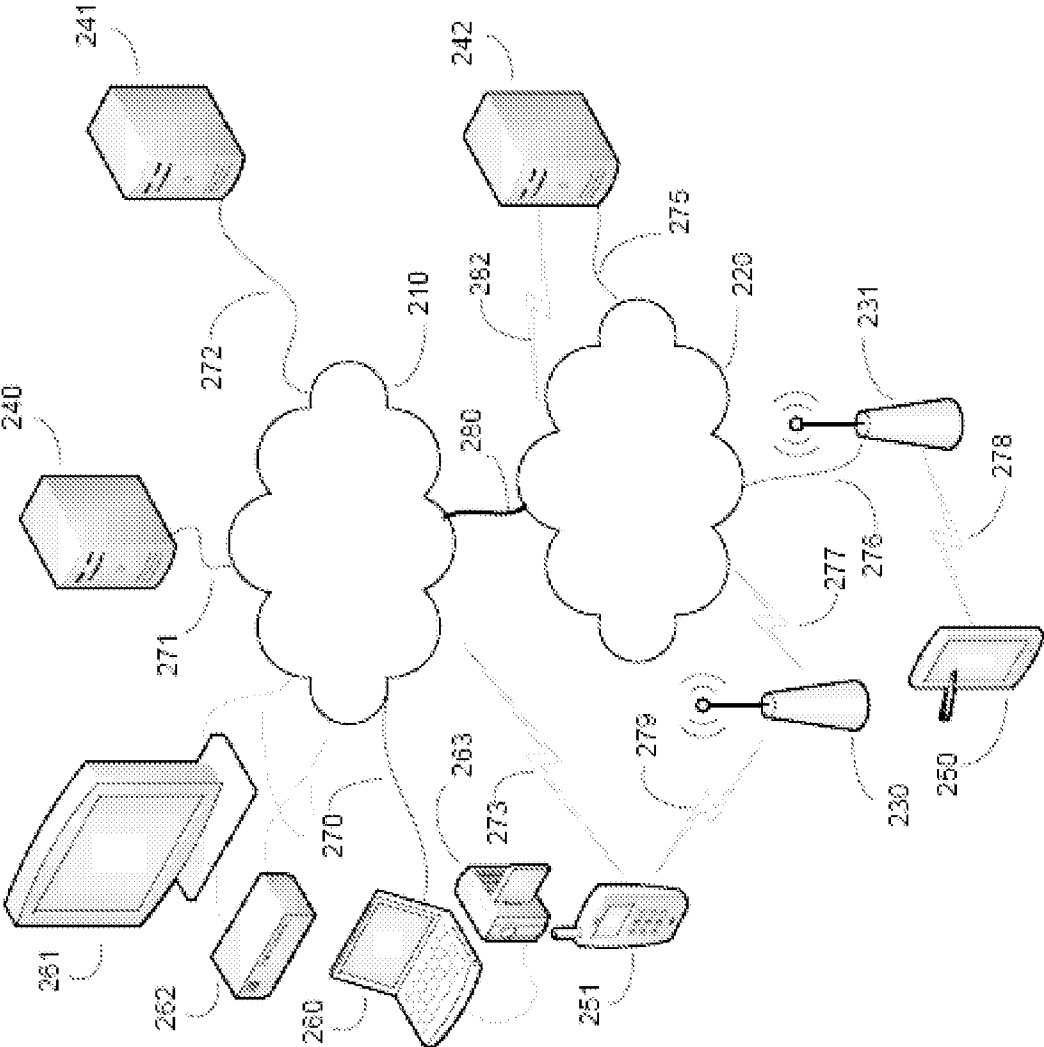


Fig. 1

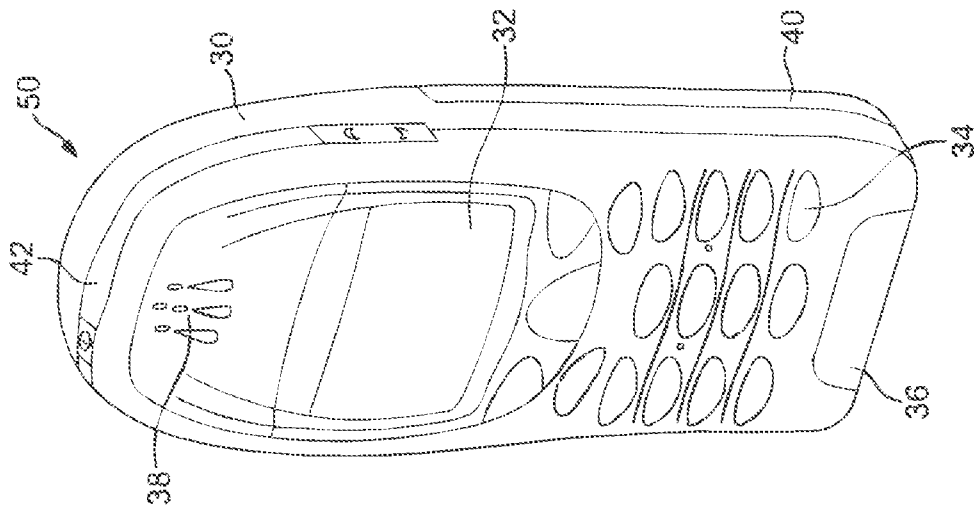


Fig. 3

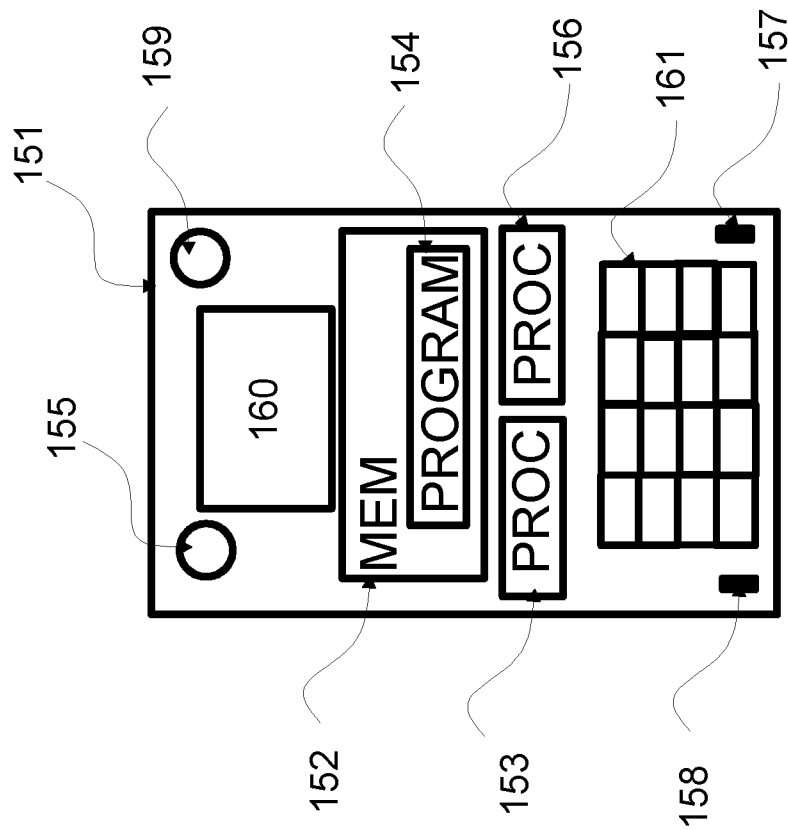


Fig. 2

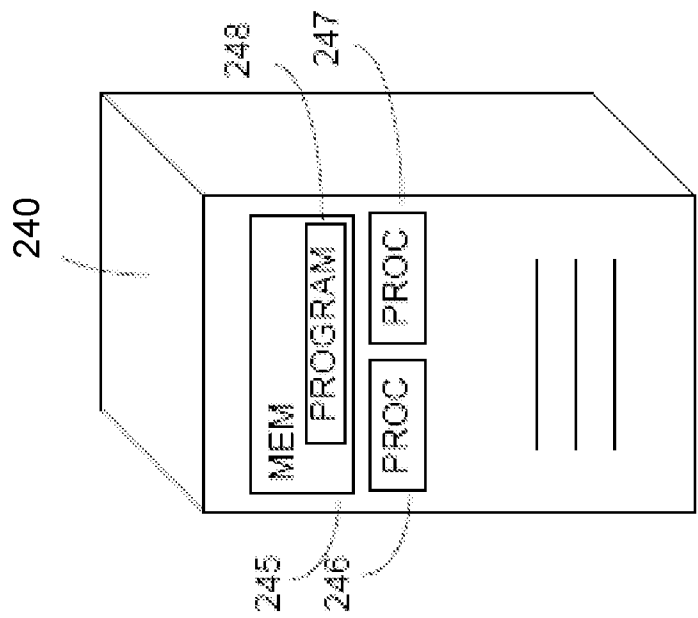


Fig. 4

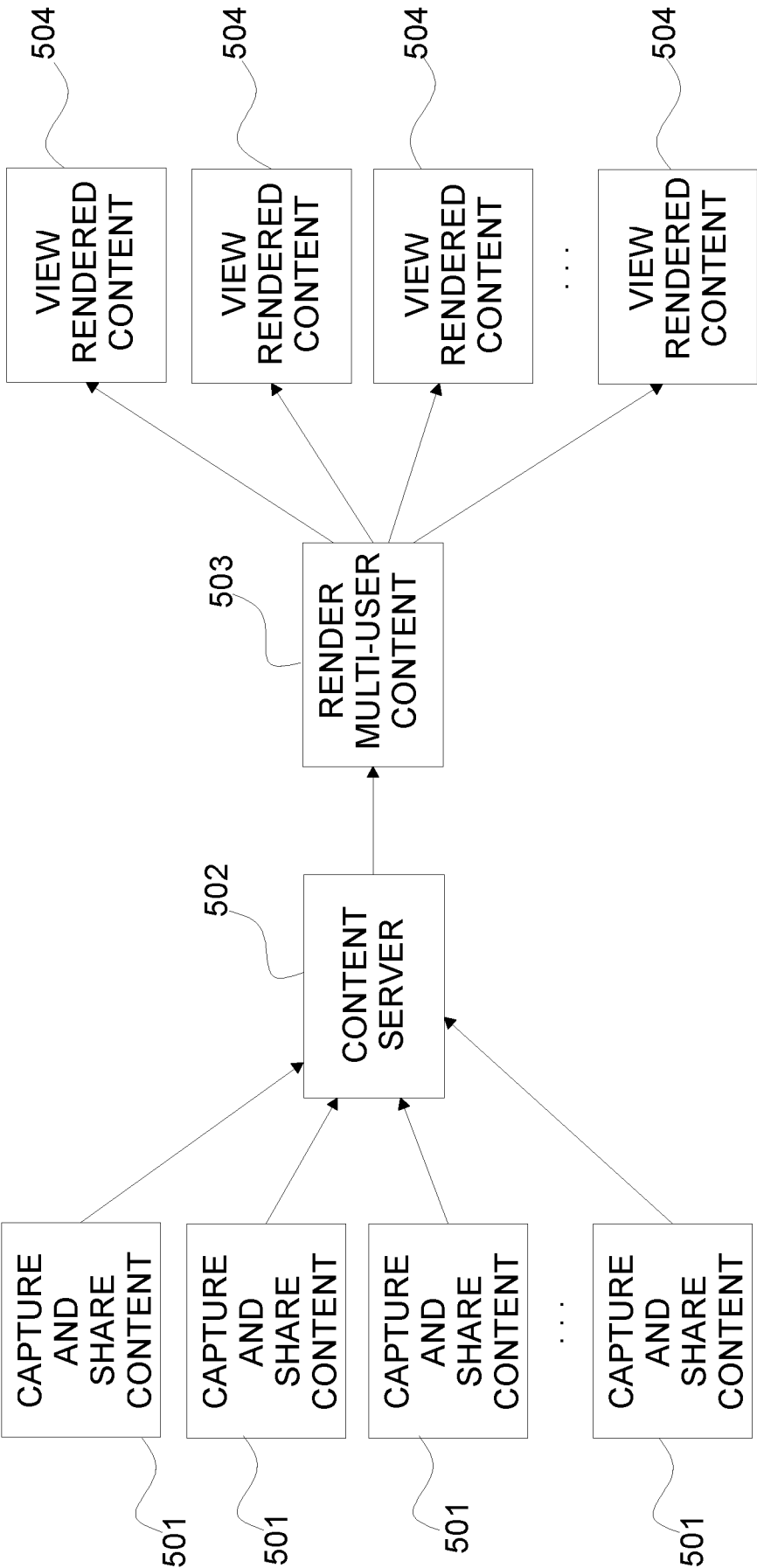


Fig. 5

Fig. 6

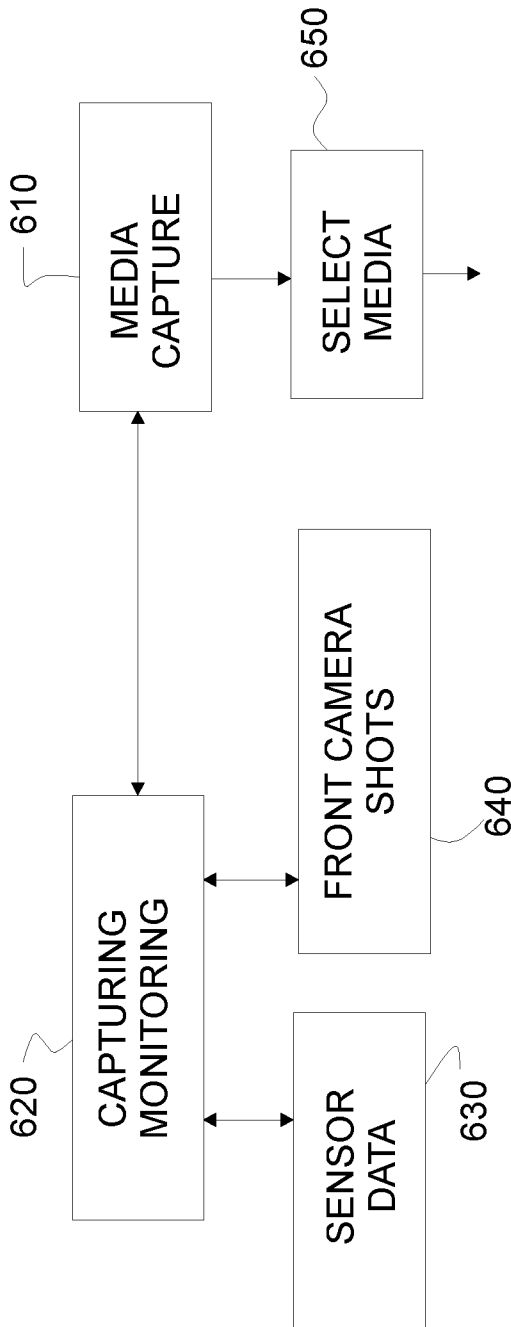


Fig. 7a

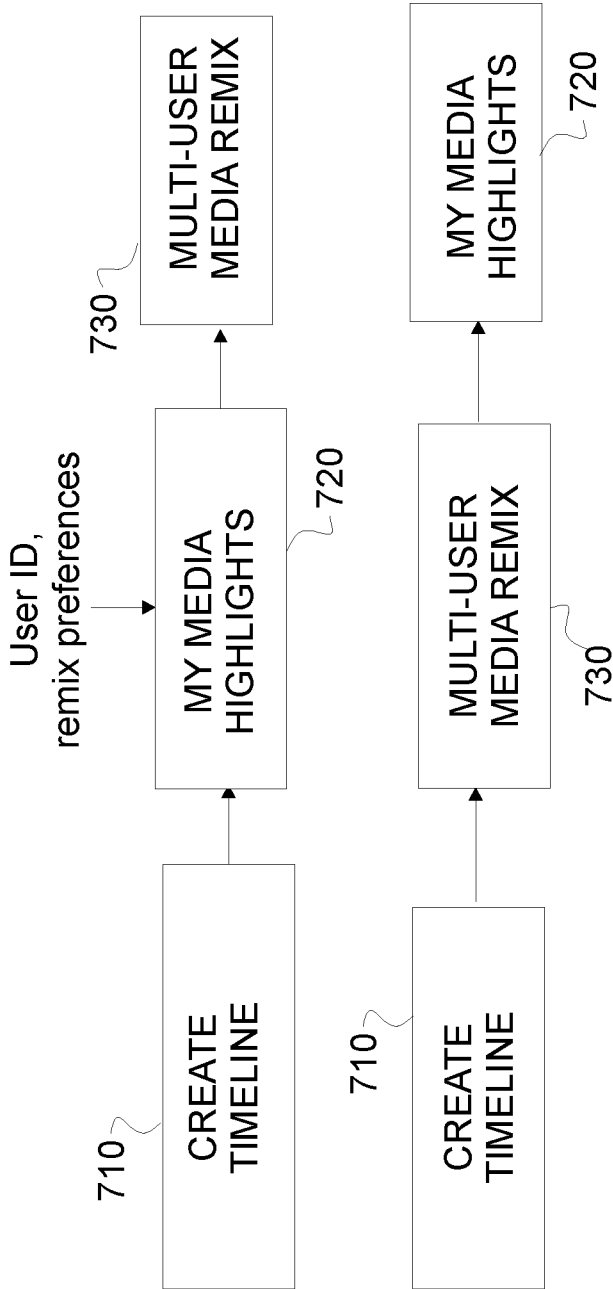


Fig. 7b

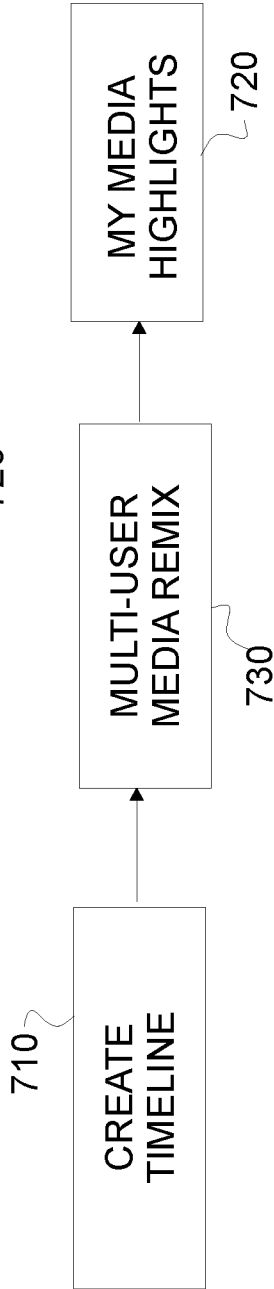
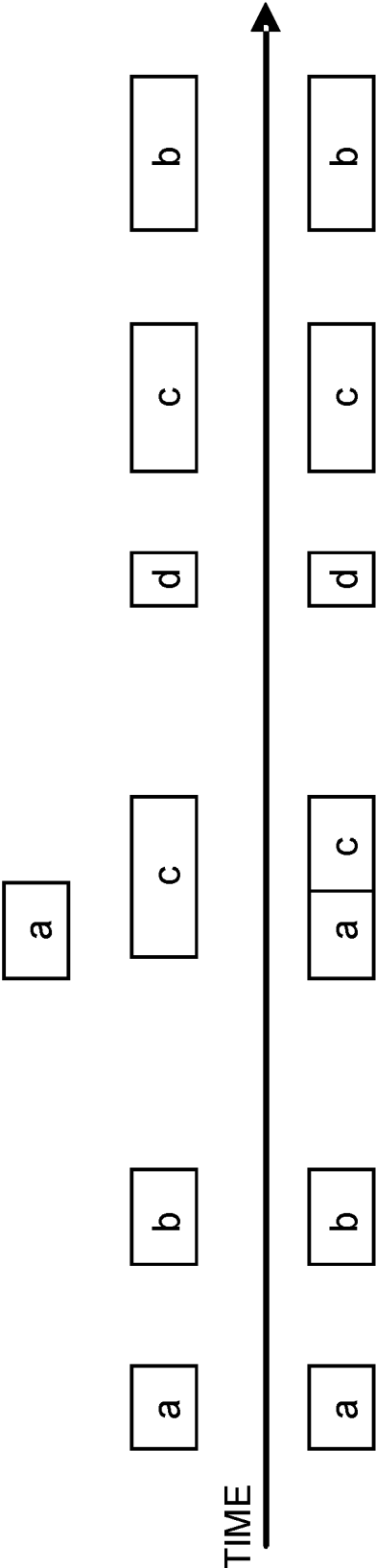


Fig. 8



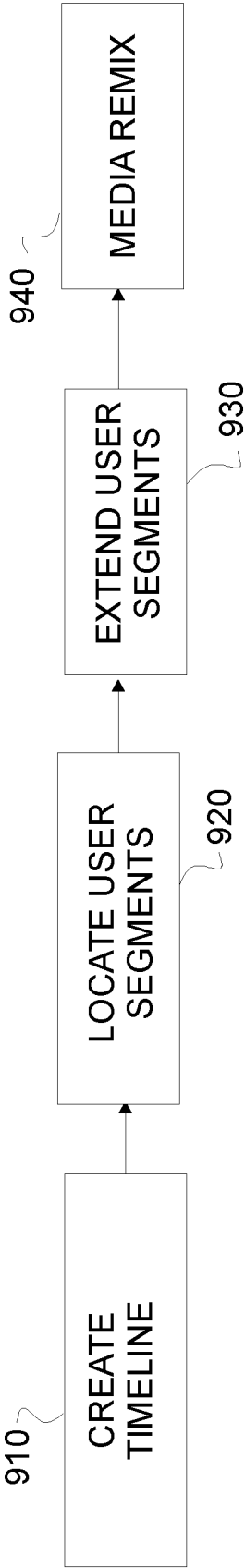
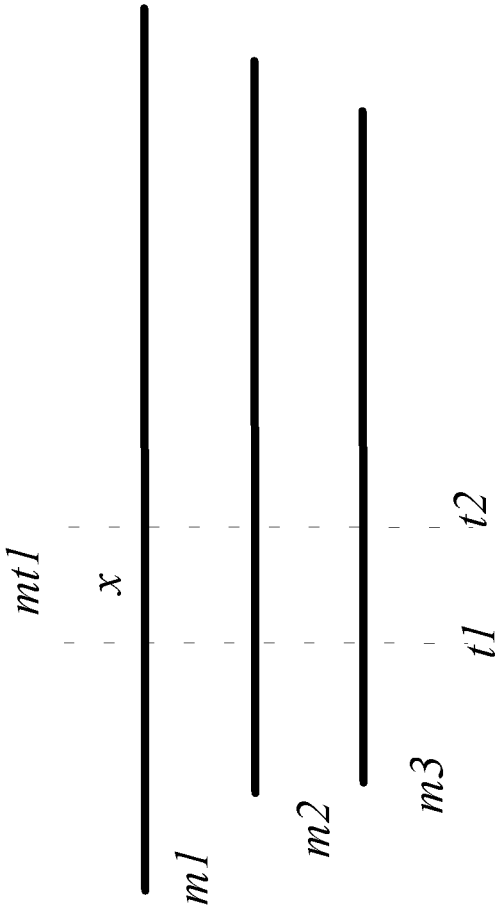
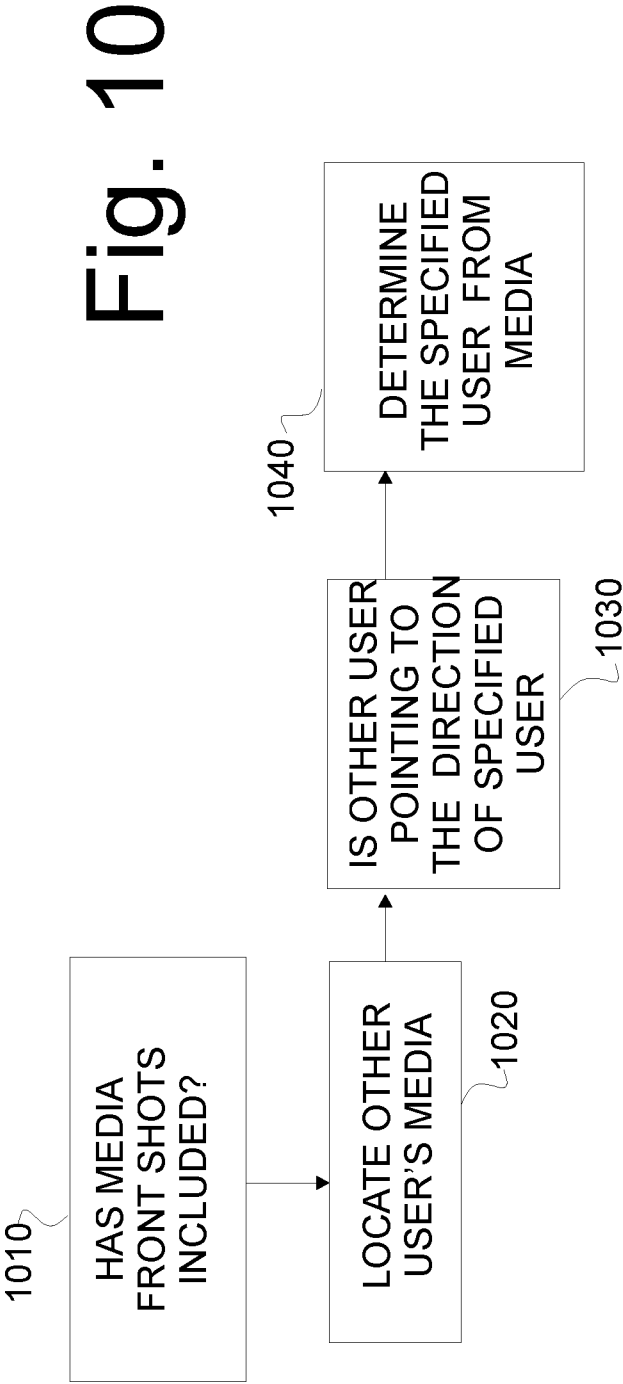
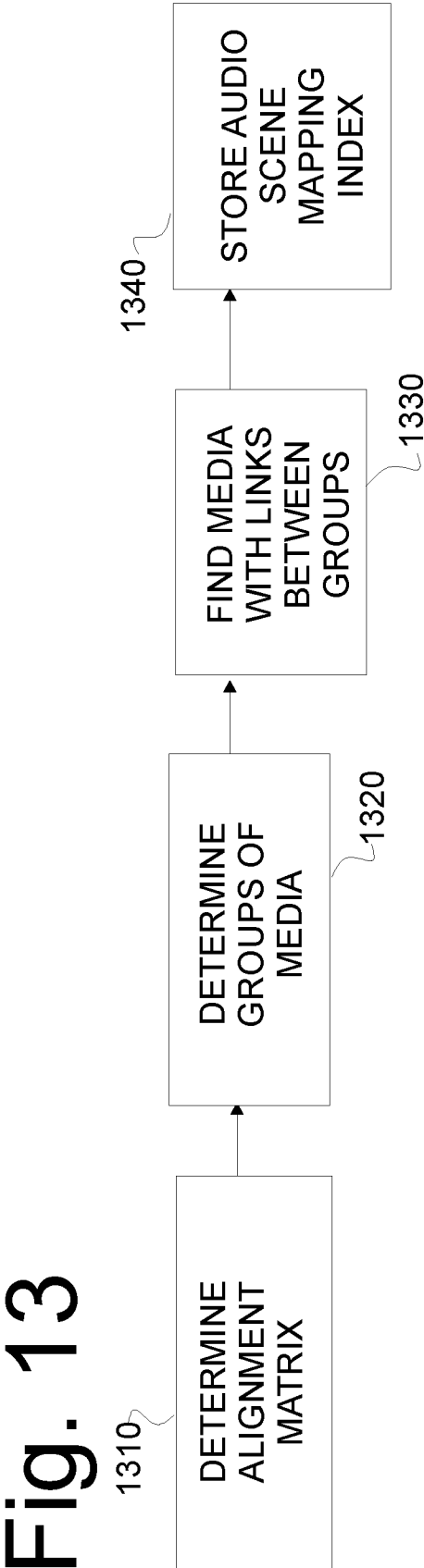
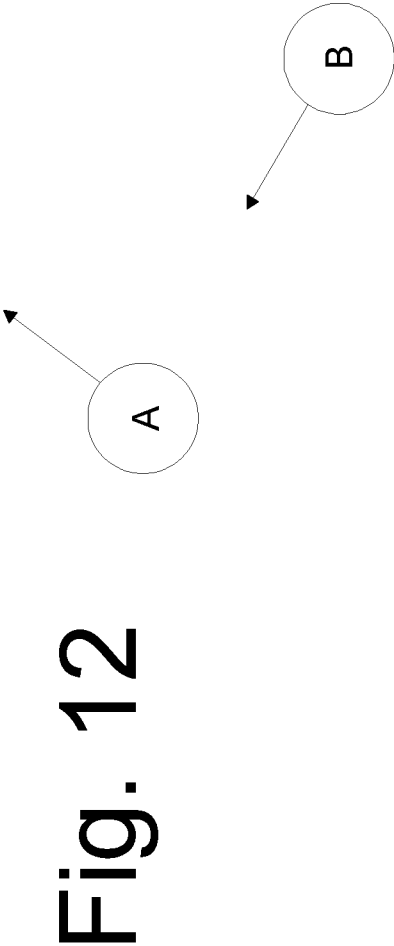
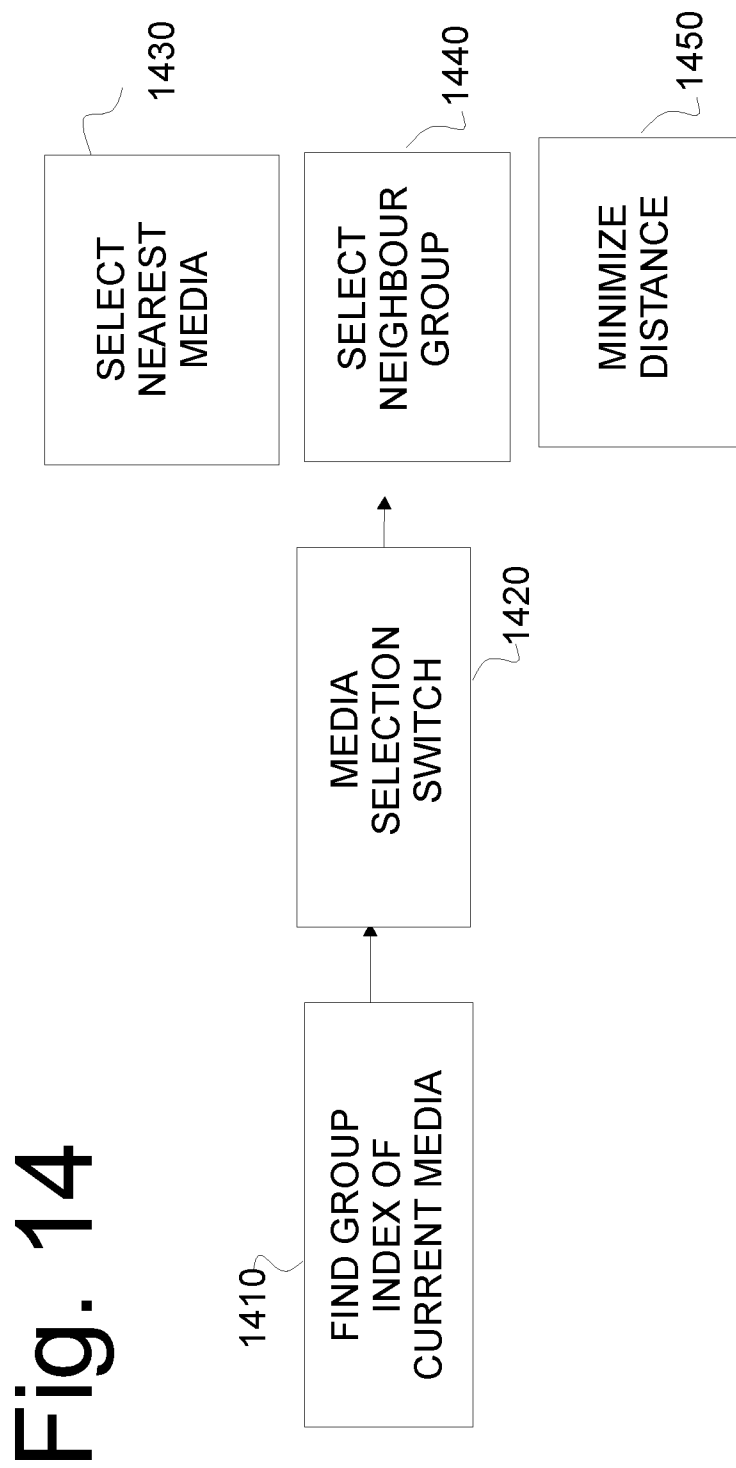


Fig. 9





10 / 10



INTERNATIONAL SEARCH REPORT

International application No.

PCT/FI2012/051007

A. CLASSIFICATION OF SUBJECT MATTER

See extra sheet

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: H04N

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

FI, SE, NO, DK

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

EPO-Internal, WPI, Google Scholar, IEEE Xplore

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2009148124 A1 (ATHSANI, A. et al.) 11 June 2009 (11.06.2009) abstract; Figs. 2A, 2B, 2C, 5; paragraphs [0008]-[0019], [0058]-[0059]; claim 1	1-49
X	WO 2008033840 A2 (EYESPOT CORP.) 20 March 2008 (20.03.2008) abstract; Figs. 2, 3, 4; par. [016]-[019], [047]-[065]	1, 13, 18, 30, 35-36, 48-49
X	US 2010204811 A1 (TRANSEAU, B.) 12 August 2010 (12.08.2010) abstract; Figs. 3, 4; par. [0006]-[0007], [0024]-[0035]; claim 1	1, 13, 18, 30, 35-36, 48-49
A	US 2011032378 A1 (KANEDA, Y.) 10 February 2011 (10.02.2011) abstract; Fig. 3	6-8, 16, 23-25, 33, 41-43

☒ 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

29 August 2013 (29.08.2013)

Date of mailing of the international search report

30 August 2013 (30.08.2013)

Name and mailing address of the ISA/FI
National Board of Patents and Registration of Finland
P.O. Box 1160, FI-00101 HELSINKI, Finland

Facsimile No. +358 9 6939 5328

Authorized officer

Timo Laakso

Telephone No. +358 9 6939 500

INTERNATIONAL SEARCH REPORT

International application No.

PCT/FI2012/051007

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MULTISILTA, J. et al.: "Mobile video stories", Int. Conf. on Digital Interactive Media in Entertainment and Arts (DIMEA'08), Sept. 10-12, 2008, Athens, Greece, pp. 401-406 the whole document	1-49
A	ZSOMBORI, V. et al.: "Automatic generation of video narratives from shared UGC", ACM Conf. on Hypertext and Media (HT'11), June 6-9, 2011, Eindhoven, The Netherlands, pp. 325-334 the whole document	1-49

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/FI2012/051007

Patent document cited in search report	Publication date	Patent family members(s)	Publication date
US 2009148124 A1	11/06/2009	US 2009089294 A1 US 8250616 B2	02/04/2009 21/08/2012
WO 2008033840 A2	20/03/2008	None	
US 2010204811 A1	12/08/2010	US 2008184868 A1 US 7935879 B2 US 2009281793 A1 US 8145496 B2 US 8452432 B2	07/08/2008 03/05/2011 12/11/2009 27/03/2012 28/05/2013
US 2011032378 A1	10/02/2011	JP 2009253848 A JP 4914398 B2 WO 2009125733 A1	29/10/2009 11/04/2012 15/10/2009

INTERNATIONAL SEARCH REPORT

International application No.
PCT/FI2012/051007

CLASSIFICATION OF SUBJECT MATTER

Int.Cl.

H04N 5/265 (2006.01)

H04N 21/8549 (2011.01)

H04N 21/858 (2011.01)

H04N 5/262 (2006.01)