(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
6 July 2017 (06.07.2017)

WIPO | PCT

(10) International Publication Number
WO 2017/114573 A1

(51) International Patent Classification:
G11B 27/22 (2006.01) *H04H 60/58* (2008.01)
G11B 27/28 (2006.01) *H04H 60/59* (2008.01)
G06T 7/00 (2006.01) *H04N 5/14* (2006.01)
H04H 20/04 (2008.01) *H04N 5/225* (2006.01)

(21) International Application Number:
PCT/EP2015/081414

(22) International Filing Date:
30 December 2015 (30.12.2015)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant: **TELECOM ITALIA S.P.A.** [IT/IT]; Via Gaetano Negri, 1, 20123 Milano (IT).

(72) Inventors: **MOGGIO, Fabrizio**; TELECOM ITALIA S.p.A., Via G. Reiss Romoli, 274, 10148 Torino (IT). **REALE, Nicola**; TELECOM ITALIA S.p.A., Via G. Reiss Romoli, 274, 10148 Torino (IT). **VARESIO, Andrea**; TELECOM ITALIA S.p.A., Via G. Reiss Romoli, 274, 10148 Torino (IT). **VECCHIETTI, Marco**; TELECOM ITALIA S.p.A., Via G. Reiss Romoli, 274, 10148 Torino (IT).

(74) Agents: **COLOMBO, Stefano Paolo** et al.; Via Vittor Pisani, 13, 20124 Milano (IT).

(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, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

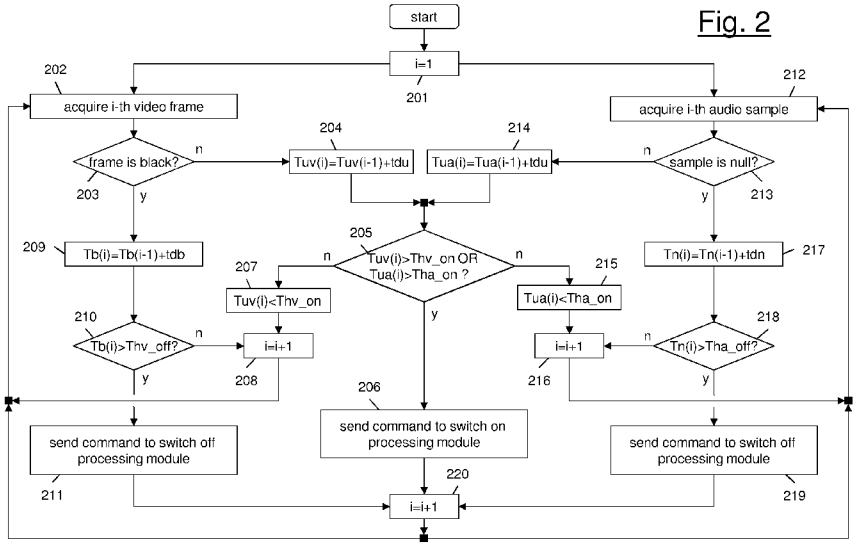
(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, 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))*

(54) Title: CONTROL OF PROCESSING OF MULTIMEDIA CONTENT

Fig. 2



(57) Abstract: It is disclosed a method for controlling processing of multimedia content captured by a capturing device, the multimedia content being provided over a signal comprising a video signal or an audio signal. The method comprises: detecting in the signal a transition between absence of a useful signal and presence of the useful signal, wherein the useful signal is a portion of the signal comprising the multimedia content; and determining a change of an operating state of the capturing device and controlling the processing of multimedia content on the basis of the detected transition.

WO 2017/114573 A1

CONTROL OF PROCESSING OF MULTIMEDIA CONTENT

Technical field

The present invention relates to the field of systems and methods for the acquisition of multimedia content. In particular, the present invention relates to a method for controlling the processing of multimedia content, for instance audio and video information, acquired by a capturing device such as a camcorder.

Background art

A system used for the acquisition of multimedia content may comprise a capturing device, such as a camcorder, and a portable processing device (such as a personal computer, a tablet or the like) connected to the capturing device. The processing device allows processing the multimedia content, in particular audio and video information, captured by the capturing device and transmitting the multimedia content to a remote station, for, e.g., broadcasting the multimedia content over a television network or over the Internet. The processing of the multimedia content is typically performed by a dedicated software application installed on the portable processing device.

Typically, to start the processing operation, it is necessary to start the software application and to switch on the capturing device. The expression “switch on the capturing device” relates to the operation of start recording (or, capturing), by the capturing device, the multimedia content from the surrounding environment. On the other hand, the expression “switch off the capturing device” relates to the operation of stop recording the multimedia content from the surrounding environment. These operations shall be both performed by the human operator holding the capturing device, who may also hold the portable processing device inside, e.g., a bag or a backpack.

In the situation described above, where the portable processing

device is easily accessible by the human operator, starting the software application may be performed, as known, by using a standard control tool, such as a key, a mouse or a touch screen. Other tools may be envisaged such as a switch connected to the portable processing device either wirelessly or by means of a cable, such as a USB cable.

US20080313686 discloses a handheld camcorder or camcorder accessory and designated webpage combination is provided, wherein the camcorder records onto a designated medium and the camcorder or camcorder accessory has an integrated wireless transmitter and receiver unit, an integrated web browser, and wherein the camcorder or camcorder accessory is pre-programmed, by default, to allow for webcasting to the designated web page; and wherein the designated webpage is programmed to receive and display the video transmission from the camcorder. The handheld camcorder or camcorder accessory having wireless internet access allows for a pre-programmed, one button webcasting, whereby a user that is familiar only with the basic operation of the camcorder can effectively webcast.

Summary of the invention

The Applicant has noticed that when known systems for the multimedia content acquisition are used, the human operator shall manually use the control tool in order to start and stop processing the captured multimedia content. In particular, the operation of starting processing the captured multimedia content involves leaving the capturing device by the human operator, extracting the portable processing device from the bag or backpack where it is stored, activating the processing operation by using the control tool and putting the processing device back in the backpack. Only after all these operations, the human operator may take back the capturing

device and start capturing the multimedia content. The operations above imply that the operator shall divert his/her attention from the camcorder and focus on the processing device and the software application to be activated for the processing of the multimedia content. This may be uneasy and may delay the transmission of the multimedia content. Moreover, manually using a control tool to activate the processing device may not be feasible in all those cases where processing of the multimedia content is performed in a device which is not accessible (or not easily accessible) by the human operator because, for instance, it is located at a remote position with respect to the capturing device.

The process for wireless webcasting of US20080313686 provides for integrating a wireless transmitter, a wireless receiver and a web browser in the handheld camcorder to allow for webcasting the multimedia content to a web page. The camcorder of US20080313686 comprises a dedicated accessory comprising, for instance, a toggle switch, for toggling the webcasting capabilities on and off of the camcorder accessory. Therefore, the method of US20080313686 may be applied by using a dedicated handheld camcorder or at least it requires modifying a standard handheld camcorder in order to integrate the necessary components. Further, the camcorder operator shall (manually) activate the toggle switch for activating the webcasting of the images captured by the camcorder.

In view of the above, the Applicant has tackled the problem of providing a method for controlling the acquisition of multimedia content which allows overcoming the drawbacks underlined above. In particular, the Applicant has tackled the problem of providing a method for controlling processing of multimedia content acquired by a handheld capturing device, such as a camcorder, which does not require altering the capturing device or its manner of use, and which does not require specific actions to be performed by the human

operator (other than those typically performed by the operator for using the camcorder) in order to start and stop processing and transmitting the captured multimedia content.

In the following description and in the claims, the expression "operating state of the capturing device" will indicate either a state according to which the capturing device is put in condition to capture multimedia content from the surrounding environment, or a state according to which the capturing device is put in condition not to capture multimedia content from the surrounding environment. The operating state of the capturing device may change by switching on the capturing device or by switching off the capturing device.

According to a first aspect, the present invention provides a method for controlling processing of multimedia content captured by a capturing device, the multimedia content being provided over a signal comprising a video signal or an audio signal, the method comprising the following steps:

a) detecting in the signal a transition between absence of a useful signal and presence of the useful signal, wherein the useful signal is a portion of the signal comprising the multimedia content; and

b) determining a change of an operating state of the capturing device and controlling the processing of multimedia content on the basis of the detected transition.

Preferably, the video signal is a digital video signal comprising a sequence of video frames, and wherein detecting comprises checking whether a current video frame of the sequence of video frames is a black frame.

Preferably, checking comprises checking whether the value of the luma and the values of the chrominance components of a selected set of pixels within the video frame are comprised within, respectively, a luma interval of black color and a chrominance

interval of black color.

In particular, preferably, the luma interval of black color is comprised between 0 and 16, and wherein the chrominance interval of black color is comprised between 127 and 129.

Preferably, the pixels of the selected set of pixels are arranged along an horizontal axis and a vertical axis of the video frame subdividing the video frame in four quadrants, and wherein the selected set of pixels also comprises at least four further pixels, each located at the center of a respective quadrant.

Preferably, the method further comprises, in case the current video frame of the sequence of video frames is not a black frame:

- incrementing a useful video frame time interval;
- checking whether the useful video frame time interval is greater than a first threshold time interval; and
- in the affirmative, determining that the transition occurred from absence of a useful signal to presence of the useful signal,

and wherein the method further comprises, in case the current video frame is a black frame,

- incrementing a black frame time interval;
- checking whether the black frame time interval is greater than a second threshold time interval; and
- in the affirmative, determining that the transition occurred from presence of a useful signal to absence of the useful signal.

Preferably, the value of the first threshold time interval and of the second threshold time interval is equal to 250 ms.

Preferably, the audio signal is a digital audio signal comprising a sequence of audio frames, each audio frame comprising a sequence of audio samples, and wherein detecting comprises checking whether a current audio frame of the sequence of audio frames is a null audio frame indicating silence.

Preferably, checking comprises checking whether the value of

each audio sample of the current audio frame is comprised within a null audio interval of values indicating silence.

In particular, preferably, the values of the null audio interval range between -16 and 16.

Preferably, the method further comprises, in case the current audio frame of the sequence of audio frames is not a null audio frame:

- incrementing a useful audio frame time interval;
- checking whether the useful audio frame time interval is greater than a third threshold time interval; and
- in the affirmative, determining that the transition occurred from absence of a useful signal to presence of the useful signal,

and wherein the method further comprises, in case the current audio frame is a null audio frame:

- incrementing a null audio frame time interval;
- checking whether the null audio frame time interval is greater than a fourth threshold time interval; and
- in the affirmative, determining that the transition occurred from presence of a useful signal to absence of the useful signal.

Preferably, the value of the third threshold time interval and of the fourth threshold time interval is equal to 250 ms.

Preferably, controlling comprises, in case the transition from absence of a useful signal to presence of the useful signal is determined, switching on the processing the multimedia content.

Preferably, controlling comprises, in case the transition from presence of a useful signal to absence of the useful signal is determined, switching off the processing the multimedia content.

According to a second aspect, the present invention provides an analysis module for a processing device configured to process a multimedia content captured by a capturing device, the multimedia content being provided to the analysis module over a signal

comprising a video signal or an audio signal, the analysis module being configured to :

- detect in the signal a transition between absence of a useful signal and presence of the useful signal, wherein the useful signal is a portion of the signal comprising the multimedia content; and
- determine a change of an operating state of the capturing device and control the processing of multimedia content on the basis of the detected transition.

Brief description of the drawings

The present invention will become clearer from the following detailed description, given by way of example and not of limitation, to be read with reference to the accompanying drawings, wherein:

- Figure 1 schematically shows an exemplary system for the acquisition of multimedia content according to an embodiment of the present invention;
- Figure 2 is a flow chart of the method for controlling processing of multimedia content according to an embodiment of the present invention;
- Figure 3 shows an exemplary video frame in which the pixels selected for analysis are evidenced, according to an embodiment of the present invention; and
- Figures 4a and 4b show alternative ways of selecting a set of pixels for the analysis according to embodiments of the present invention.

Detailed description of preferred embodiments of the invention

Figure 1 schematically shows an exemplary system 1 for the acquisition of multimedia content suitable for implementing the method for controlling processing of multimedia content acquired by a capturing device according to an embodiment of the present

invention.

The system 1 preferably comprises a capturing device 11, such as a camcorder, suitable to capture multimedia (e.g. audio and video) content, and a processing device 12 connected to the capturing device 11. For sake of simplicity, in the following description reference will be made to a (handheld) camcorder as capturing device. This is not limiting since the method for controlling the acquisition of multimedia content according to the present invention may be applied to any type of capturing device. For sake of clarifying example, the camcorder 11 may be used by a human operator for capturing in real-time the audio and video content of a live broadcasting event such as a live sport event, which may be transmitted over an IP (Internet Protocol) network as an audiovisual streaming content.

The processing device 12 may be a portable personal computer, a tablet, a Mini PC (e.g. Intel® NUC), a Single Board Computer (e.g. Raspberry Pi) or the like. The processing device 12 may be stored into a bag or a backpack by the human operator. The processing device 12 is preferably connected to the camcorder 11 by means of a wireless connection or a wired connection.

The camcorder 11 may have digital output ports and/or analog output ports. As known, the camcorder 11 may output a digital audiovisual signal which is formatted according to a digital video format, such as DV (Digital Video). In this case, the digital output ports of the camcorder 11 may comprise one or more of: a USB (Universal Serial Bus) output port , a FireWire output port, .a HDMI (High Definition Multimedia Interface) output port, a SDI (Serial Digital Interface) output port. Corresponding USB, FireWire, HDMI, or SDI input ports are present in the processing device 12 for receiving the digital audiovisual signal though a USB, FireWire, HDMI or SDI cable. For the FireWire connectivity, the processing device 12 is

typically equipped with a FireWire card for the acquisition of the digital audiovisual signal by the processing device through a FireWire input port.

In case the camcorder 11 outputs an analog audiovisual signal through an analog output port (for instance, an RCA output port or an S-Video output port), the processing device 12 may typically receive this signal though an analog input port connected to a video capture card which provides for converting the received signal into a digital audiovisual signal.

In the following description, as a non limiting example, it will be assumed that the camcorder 11, once switched on, outputs a digital audiovisual signal.

The system 1 further comprises a communication network 13 and a remote station 14. The communication network 13 is preferably a wireless communication network connecting the processing device 12 and the remote station 14. The communication network 13 may be an IP communication network. The remote station 14 may be a server.

The camcorder 11 is configured to capture multimedia content in the form of an audiovisual data flow, output data in the form of a digital audiovisual signal and transmit it to the processing device 12. The processing device 12 is configured to process the digital audiovisual signal, for instance by applying a coding operation, as it will be described herein after, and transmit it to the remote station 14 through the communication network 13. The remote station 14 may be configured to receive the audiovisual signal, to further process it (for instance, by applying a decoding operation), and to broadcast it over, e.g., a television network or over the Internet. The operation of the remote station 14 is not relevant to the present invention and hence it will not be further described herein after.

According to the present invention, the processing device 12

comprises an analysis module 121 and a processing module 122. Both the analysis module 121 and the processing module 122 receive the digital audiovisual signal comprising the multimedia content captured by the camcorder 11. The analysis module 121 is preferably configured to analyse the digital audiovisual signal and, on the basis of a result of the analysis, to control the activation of the processing module 122 or the deactivation thereof, as it will be described in detail herein after. The processing module 122 is preferably configured to, once activated by the analysis module 121, process the digital audiovisual signal in order to, for instance, apply an encoding operation to the digital video signal and the digital audio signal comprised within the digital audiovisual signal. In particular, the processing module 122 may apply a H.264/MPEG-4 AVC (Advanced Video Coding) to the digital video signal, and a AAC (Advanced Audio Coding) encoding to the digital audio signal. Alternatively, it may apply a H.265 (HEVC - High Efficiency Video Coding) encoding. The encoded signals may then be transmitted by the processing device 12 over the communication network 13 towards the remote station 14. The operations of the processing module 122 will not be described in greater detail herein after as they are not relevant to the present description.

Both the analysis module 121 and the processing module 122 may be implemented as software modules running over the hardware of the processing device 12.

The method for controlling processing of the multimedia content acquired by the capturing device according to the present invention will be now described in detail with reference to the flow chart of Figure 2.

It is assumed that the processing device 12 is already active when the operator switches on the camcorder 11. When the camcorder 11 is switched on, the camcorder 11 captures the multimedia content

from the surrounding environment, generates a corresponding digital audiovisual signal and transmits it to the processing device 12. The processing device 12 receives the digital audiovisual signal through a digital input port. The signal is forwarded to the analysis module 121 and, in parallel, to the processing module 122. At the analysis module 121, the digital audiovisual signal received from the input port is preferably split into a digital video signal and a digital audio signal. It is to be noticed that the digital audiovisual signal at the input of the analysis module 122, which carries the captured multimedia content, generally comprises portions of a "useful audiovisual signal" actually comprising the multimedia content, and portions indicating absence of such a useful signal, as it will be described in greater detail herein after.

In particular, the digital video signal is acquired by the analysis module 121 as a stream of video frames by sampling the digital video signal at a predetermined frequency of acquisition of the video signal or video capture frequency, which may be equal to 50 Hz in Europe and 60 Hz in America and Asia. Each video frame, as known, is composed of a number of lines of picture elements (pixel). The pixels of a video frame are usually associated with a color model. According to an embodiment of the present invention, each pixel may be associated with a brightness information (also called "luma") and a color information (or "chrominance"), the latter being usually composed of two color-difference components, U (blue-luma) and V (red-luma). According to an embodiment of the present invention, the brightness information and the components of the color information have a 8-bit representation. In this case, the luma information may have a value within the range from 0 to 255, the value 0 identifying the absence of light. Moreover, each of the chrominance components may have a value within the range from -128 to 127. Usually, a value of 128 is added to the chrominance components, which may be

positive or negative, so that these components assume only positive values. The black color may be hence associated with a luma value equal to 0 and values of the chrominance components equal to 127. According to the present invention, the black color may be associated with a luma value ranging within a given interval (indicated in the following lines as "luma interval of black color") around value 0: the minimum value of the interval may be equal to 0 and the maximum value of the interval may be equal to 16. Moreover, the black color may be associated with values of the chrominance components ranging within a further given interval (indicated in the following lines as "chrominance interval of black color") around value 127: the minimum value of the interval may be equal to 127 and the maximum value of the interval may be equal to 129.

The audio signal captured by the camcorder 11 may comprise, as known, one or two channels (left channel, L, and right channel, R). It is usually an analog audio signal which may be then digitalized. The digital audio signal is acquired by the analysis module 121 as one stream of audio samples per each channel L, R. The samples of a given channel are obtained by sampling the corresponding audio signal at a predetermined frequency that may be equal to 48 kHz and quantizing the value of each sample. According to an embodiment of the present invention, each sample may have a value within a discrete numeric range of relative numbers with a sign. Silence is represented by the central value of the numeric range used for quantizing the audio signal. For instance, silence may be represented by a value nearly equal to 0. The audio samples of each channel L, R are preferably grouped in audio frames. According to the present invention, an audio frame of a channel preferably comprises a number of audio samples of the channel spread over a given period which corresponds to the temporal length of the video

frame that is defined by the video capture frequency. An audio frame comprising audio samples with values indicating silence will be indicated in the following description as a “null audio frame”. According to an embodiment of the present invention, an audio sample indicates silence when its value is comprised within a null audio interval of values indicating silence, which may range between -16 and 16. According to an embodiment of the present invention, a single audio channel, for instance the left one, is analysed by the analysis module 121.

According to the method of the present invention, the digital video signal and the digital audio signal are independently analysed at the analysis module 121, for detecting a change in the operating state of the capturing device and trigger accordingly the processing of the captured multimedia content. The change in the operating state of the capturing device is determined by detecting transitions from the absence of a useful video/audio signal and the presence of the useful video/audio signal in each stream, and viceversa. The expression “useful video/audio signal” will indicate the portion of the video/audio stream containing the captured multimedia content. The absence of the useful video signal means that black frames are detected for a given interval of time within the stream of video frames. Similarly, the absence of the useful audio signal means that null audio frames are detected for a given interval of time within the stream of audio frames. Detection of a transition involving either the digital video signal or the digital audio signal triggers the activation or deactivation of the signal processing at the processing module 122, as it will be described in detail herein after.

The stream of video frames of the digital video signal and the stream of audio frames of the digital audio signal are preferably analysed in parallel by the analysis module 121. The analysis of the stream of video frames of the digital video signal will be described

first in the following lines.

Upon acquisition of a video frame of the stream of video frames by the analysis module 121 (steps 201 and 202), a check is performed by the analysis module 121 for determining whether the video frame is a black frame (step 203). The generic video frame is indicated in the flow chart of Figure 2 as the i-th video frame, wherein i is an integer indexing number indicating the position of the video frame within the stream of video frames at the analysis module 121.

In order to determine whether the i-th video frame is a black frame, the analysis module 121 preferably checks the value of the luma and the values of the chrominance components of a selected set of pixels within the video frame (step 203). The set of pixels to be analysed may be selected as follows. Reference is made to Figure 3, where the generic i-th video frame is represented as subdivided in pixels. According to an embodiment of the present invention, the video frame is split into four sections of equal area by an horizontal axis X and a vertical axis Y. The four sections will be referred to as quadrants. The set of pixels which are analysed in the analysis module 121 at step 203 comprises the pixels of the i-th video frame which are positioned on the horizontal axis X and the vertical axis Y that divide the video frame into quadrants. Moreover, the set of pixels to be analysed further comprises a number of pixels for each quadrant, which are positioned at the center of each quadrant. This number may be four. In Figure 3, the pixels belonging to the selected set of pixels to be analysed by the analysis module 121 are represented as grey pixels.

Other pixels may be chosen for performing the check of step 203. Two other exemplary sets of pixels that may be analysed at step 203 are the grey pixels in Figures 4a and 4b.

At step 203, the check is performed as follows: for each pixel of the selected set, the value of the luma and the values of the

chrominance components are checked. In particular, the analysis module 121 preferably checks whether the luma value of each selected pixel is comprised within the luma interval of black color described above, for instance between 0 and 16. Moreover, the analysis module 121 preferably checks whether the value of each chrominance component of each selected pixel is comprised within the chrominance interval of black color described above, for instance between 127 and 129. If, for each pixel of the selected set, the luma value is not comprised within the luma interval of black color and/or the values of the chrominance components are not comprised within the chrominance interval of black color, the analysis module 121 preferably determines that the i-th video frame is not a black frame. If, for each pixel of the selected set, the luma value is comprised within the luma interval of black color and the values of the chrominance components are comprised within the chrominance interval of black color, the analysis module 121 preferably determines that the i-th video frame is a black frame.

In the former case, preferably, the analysis module 121 increments a useful video frame time interval Tuv (step 204) by a given amount tdu, according to the following equation:

$$Tuv(i) = Tuv(i-1) + tdu \qquad [2]$$

wherein Tuv(i) is the useful video frame time interval for the i-th video frame, Tuv(i-1) is the useful video frame time interval in correspondence of the preceding video frame and tdu is the amount of the increment. For instance, the increment tdu may be equal to 40 ms, which corresponds to the duration of a single video frame in case the frequency of acquisition of the video signal is 50 Hz (25 frames per second). In the following description, the expression “useful video frame” will indicate a video frame in which the pixels of the selected set have luma and chrominance values outside the luma and chrominance intervals of black color.

At step 205, the analysis module 121 preferably checks whether the useful video frame time interval Tuv(i) computed at step 204 for the current i-th video frame is greater than a first threshold time interval Thv_on. The first threshold time interval Thv_on may be equal to 250 ms. In case the useful video frame time interval Tuv(i) is greater than the first threshold time interval Thv_on, the analysis module 121 preferably determines that, in correspondence of the i-th video frame, a transition is occurred from a condition of absence of a useful video signal to a condition of presence of the useful video signal. In this case, at step 208, the analysis module 121 preferably generates and sends to the processing module 122 a command in order to switch on the processing module 122. In case at step 205 the analysis module 121 determines that the useful video frame time interval Tuv(i) computed at step 206 is lower than the first threshold time interval Thv_on (step 207), the analysis module 121 preferably acquires the next video frame (steps 208 and 202) and repeats step 205.

According to the present invention, at step 205, the analysis module 121 determines that a transition is occurred from a condition of absence of the useful video signal to a condition of presence of the useful video signal only when the analysis module 121 acquires a given number of consecutive useful video frames, this number being determined by the frequency of acquisition of the video signal by the camcorder 11 and the first threshold time interval Thv_on. This situation arises when the operator switches on the camcorder 11 and starts capturing a video signal, so that the analysis module 121 starts acquiring a sequence of useful video frames. Moreover, when the camcorder 11 is switched on, the analysis module 121 may acquire a number of black frames (due, for instance, to the presence of an acquisition card in the processing device) before acquiring the useful video frames of the captured multimedia content and hence a

transition may be detected when a 250 ms sequence of useful video frames is analysed at the analysis module 121 after the last black frame.

In case, at step 203, the analysis module 121 determines that the current video frame is a black frame, it preferably increments a black frame time interval Tb (step 209) by a given amount tdb, according to the following equation:

$$Tb(i) = Tb(i-1) + tdb \quad [2]$$

wherein Tb(i) is the black frame time interval for the i-th video frame, Tb(i-1) is the black frame time interval in correspondence of the preceding video frame and tdb is the amount of the increment. For instance, the increment tdb may be equal to 40 ms, which corresponds to the duration of a single video frame in case the frequency of acquisition of the video signal is 50 Hz (25 frames per second).

At step 210, the analysis module 121 preferably checks whether the black frame time interval Tb(i) computed at step 209 for the current i-th video frame is greater than a second threshold time interval Thv_off. The second threshold time interval Thv_off may have the same value of the first threshold time interval Thv_on, and it may be equal to e.g. 250 ms. In case the black frame time interval Tb(i) is greater than the second threshold time interval Thv_off, the analysis module 121 preferably determines that, in correspondence of the i-th video frame, a transition is occurred from a condition of presence of the useful video signal to a condition of absence of the useful video signal. In this case, at step 211, the analysis module 121 preferably generates and sends to the processing module 122 a command in order to switch off the processing module 122. In case at step 210 the analysis module 121 determines that the black video frame time interval Tb(i) computed at step 209 is lower than the second threshold time interval Thv_off, the analysis module 121

preferably acquires the next video frame (steps 208 and 202) and repeats step 203.

According to the present invention, at step 210, the analysis module 121 determines that a transition is occurred from a condition of presence of the useful video signal to a condition of absence of the useful video signal only when it has acquired a given number of consecutive black video frames, the number being determined by the frequency of acquisition of the video signal by the camcorder 11 and the second threshold time interval Thv_off. This situation arises when the operator switches off the camcorder 11 and stops capturing a video signal.

As mentioned above, the analysis of the audio digital signal by the analysis module 121 proceeds in parallel with respect to the analysis of the digital video signal.

Upon acquisition of an audio frame by the analysis module 121 (steps 201 and 212), a check is performed by the analysis module 121 for determining whether the audio frame is a null audio frame (step 213). The generic audio frame is indicated in the flow chart of Figure 2 as the i-th audio frame, wherein i is the index already used for the video frames and indicating the position of the audio frame within the stream of audio frames at the input of the analysis module 121.

In order to determine whether the i-th audio frame is a null audio frame, the analysis module 121 preferably checks the values of the audio samples within the audio frame (step 213). In particular, at step 213, the analysis module 121 preferably checks whether the value of each audio sample is comprised within the null audio interval of values indicating silence, for instance between -16 and 16. If, for each audio sample of the current audio frame, the value of the audio sample is not comprised within the null audio interval of values indicating silence, the analysis module 121 preferably determines

that the i-th audio frame is not a null audio frame. If, for each audio sample of the current audio frame, the value of the audio sample is comprised within the null audio interval of values indicating silence, the analysis module 121 preferably determines that the i-th audio frame is a null audio frame.

In the former case, preferably, the analysis module 121 increments a useful audio frame time interval Tua (step 214) by the given amount tdu, according to the following equation:

$$Tua(i) = Tua(i-1) + tdu \quad [3]$$

wherein Tua(i) is the useful video frame time interval for the i-th audio frame, Tua(i-1) is the useful audio frame time interval in correspondence of the preceding audio frame and tdu is the amount of the increment, which is preferably the same amount used to increment the useful video frame time interval according to equation [2]. In the following description, the expression “useful audio frame” will indicate an audio whose audio samples have values outside the null audio interval.

At step 205, the analysis module 121 preferably checks whether the useful audio frame time interval Tua(i) computed at step 214 for the current i-th audio frame is greater than a third threshold time interval Tha_on. The third threshold time interval Tha_on may be equal to the first threshold time interval Thv_on (e.g. equal to 250 ms). In case the useful audio frame time interval Tua(i) is greater than the threshold time interval, the analysis module 121 preferably determines that, in correspondence of the i-th video frame, a transition is occurred from a condition of absence of the audio signal to a condition of presence of the audio signal. In this case, at step 208, the analysis module 121 preferably generates and sends to the processing module 122 a command in order to switch on the processing module 122. In case at step 205 the analysis module 121 determines that the useful audio frame time interval Tua(i) computed

at step 206 is lower than the third threshold time interval Tha_on (step 215), the analysis module 121 preferably acquires the next audio frame (steps 216 and 212) and repeats step 213.

According to the present invention, at step 205, the analysis module 121 determines that a transition is occurred from a condition of absence of a useful audio signal to a condition of presence of the useful audio signal only when it has acquired a given number of consecutive useful audio frames, the number being determined by the frequency of acquisition of the audio signal by the camcorder 11, which, as described above, is assumed to be equal to the frequency of acquisition of the video signal by the camcorder 11 and the third threshold time interval Tha_on. This situation arises when the operator switches on the camcorder 11 and starts capturing an audio signal.

In case, at step 213, the analysis module 121 determines that the current audio frame is a null audio frame, it preferably increments a null audio frame time interval Tn (step 217) by a given amount tdn, according to the following equation:

$$Tn(i) = Tn(i-1) + tdn \qquad [4]$$

wherein Tn(i) is the null audio frame time interval for the i-th audio frame, Tn(i-1) is the null audio frame time interval in correspondence of the preceding audio frame and tdn is the amount of the increment. For instance, the increment tdn may be equal to 40 ms, which corresponds to the duration of a single video frame in case the frequency of acquisition of the video signal is 50 Hz (25 frames per second).

At step 218, the analysis module 121 preferably checks whether the null audio frame time interval Tn(i) computed at step 217 for the current i-th audio frame is greater than a fourth threshold time interval Tha_off. The fourth threshold time interval Tha_off may have the same value of the third threshold time interval Tha_on, for

instance 250 ms. In case the null audio frame time interval Tb(i) is greater than the fourth threshold time interval Tha_off, the analysis module 121 preferably determines that, in correspondence of the i-th audio frame, a transition is occurred from a condition of presence of the useful audio signal to a condition of absence of the useful audio signal. In this case, at step 219, the analysis module 121 preferably generates and sends to the processing module 122 a command in order to switch off the processing module 122. In case at step 218 the analysis module 121 determines that the null audio frame time interval Tn(i) computed at step 217 is lower than the fourth threshold time interval Tha_off, the analysis module 121 preferably acquires the next audio frame (steps 216 and 212) and repeats step 213.

According to the present invention, at step 218, the analysis module 121 determines that a transition is occurred from a condition of presence of the useful audio signal to a condition of absence of the useful audio signal only when it has acquired a given number of consecutive null audio frames, the number being determined by the frequency of acquisition of the audio signal by the camcorder 11 and the fourth threshold time interval Tha_off. This situation arises when the operator switches off the camcorder 11 and stops capturing an audio signal.

When the processing module 122 is activated (step 206), it preferably process the digital video signal and the digital audio signal and transmits them over the communication network 13 towards the remote station 14. As mentioned above, processing by the processing module 122 may comprise encoding the digital video signal according to the H.264/MPEG-4 AVC encoding technique and encoding the digital audio signal according to the AAC encoding technique. Alternatively, the H.265 (HEVC) encoding technique may be used.

Once the processing module 122 has been switched on according

to the command generated at step 206, the camcorder 11 may continue capturing a useful video signal and/or a useful audio signal until the operator switches off the camcorder 11. Therefore, the analysis module 121 cyclically acquires a video/audio frame (step 220) and repeats steps 202-211 and 212-219. At step 203 or 213 the check is repeatedly negative and the useful video frame time interval (and/or the useful audio frame time interval) gets incremented on a frame by frame basis. The check at step 205 is hence repeatedly positive. In this situation, the analysis module 121 at step 206 continues generating the command above, which is however ignored at the processing device 122. Similarly, according to the method of the present invention, once the processing module 122 has been switched off according to the command generated at step 211 or 219, the analysis module 121 cyclically acquires a video/audio frame (step 220) and repeats steps 202-211 and 212-219. In this case, if the check at step 203 or 213 continues to be positive, as well as the check at step 210 or 218, respectively, the analysis module 121 at step 211 or 219 continues generating the command to switch off the processing module 122, which is however ignored at the processing device 122. Otherwise, if, for instance, the digital video signal restart comprising useful video frames, the processing module 122 is switched on again.

According to the preceding description, the check at step 205 is a check involving an “OR” logical operation. In particular, according to the preceding description, the processing device 122 is switched on and off by detecting a transition within any one of the digital video signal and the digital audio signal. However, other logical operations may be considered to perform the check of step 205, for instance, an AND operation. In this case, the check at step 205 is positive only when a transition is detected within both the digital video signal and the digital audio signal. Furthermore, the method of the present

invention applies also to situations in which only one of a digital video signal and a digital audio signal is available. In this case, steps 201-211 and 220 are performed for analyzing a digital video signal or, alternatively, steps 201, 212-220 are performed for analyzing a digital audio signal.

In the light of the description above, according to the present invention, once the human operator switches on the camcorder 11 (provided that the processing device 12 is already active), the analysis module 121 analyses the digital video signal and the digital audio signal and according to an outcome of the analysis, it provides for switching on the processing module 122, which is in charge of transmitting the digital video signal and the digital audio signal over the network. On the other side, once the human operator switches off the camcorder 11, the analysis module 121, which continues analysing the digital video signal and the digital audio signal, provides for switching off the processing module 122. According to the present invention, every time the human operator switches on the camcorder 11, the analysis module 121 activates the processing module 122, and every time the human operator switches off the camcorder 11, the analysis module 121 deactivates the processing module 122. The analysis module 121 is capable of detecting the presence of a useful signal by determining whether a transition occurred from absence of signal to presence of signal and viceversa, for the digital video signal and/or the digital audio signal. This allows giving the human operator the possibility to voluntarily trigger the start and stop of the operations of processing and transmitting the multimedia content captured by the camcorder by merely using the camcorder in its normal mode of operation, in particular by merely changing the operating state of the camcorder.

According to the present invention, the human operator may connect the processing device to the camcorder, switch on the

processing device and put it on a backpack, and then he/she can decide to start and stop the processing and the transmission of the audiovisual signal captured by the camcorder simply switching on and off the camcorder itself. Alternatively, the operator may switch on the processing device, put it on a backpack and connect it to the camcorder before switching on the camcorder. It is to be noticed that the application of the method of the present invention is independent from "how" and "when" the camcorder is connected to the processing device. Indeed, switching on and off the camcorder when it is connected to the processing device or connect and disconnect the camcorder when it is switched on are, according to the present invention, equivalent actions that determine the presence or absence of an incoming audiovisual signal into the processing device.

Advantageously, the method of the present invention may be implemented without altering the capturing device or its manner of use. Indeed, the method of the present invention may be applied to standard capturing devices as it does not require a specific control tool to be activated for starting and stopping the processing of the multimedia content. Moreover, the method of the present invention does not require specific actions to be performed by the human operator neither on the capturing device nor on the processing device to start and stop processing the multimedia content. Thanks to the present invention, the human operator may completely focus on his/her main activity, i.e. using the camcorder to capture the multimedia content, without the need to manage the processing and transmission of the captured content. This allows avoiding any delay in transmitting the captured multimedia content.

Advantageously, the method of the present invention allows controlling the processing of multimedia content captured by a capturing device in all those cases in which the capturing device and the processing device (or, at least, the processing module

responsible for the processing of the multimedia content) are located in remote positions or in those cases in which the processing module is not easily accessible. Indeed, as already widely discussed above, the control may be advantageously triggered only by switching on and off the capturing device.

According to the preceding description, the change of the operating state of the camcorder 11 is triggered by the switching on and off of the camcorder 11 itself. However the method of the present invention may be similarly applied to other situations in which the change of the operating state of the capturing device 11 is determined by other actions/events.

According to an alternative embodiment, if the capturing device is a camcorder with a lens cover, the start and stop of the processing of the multimedia content may be triggered by the absence/presence of the lens cover on the camcorder, which actually determines a change in the operating state of the camcorder. In this case, the method of the present invention provides for detecting a transition from absence of a useful video signal to presence of a useful video signal when the human operator takes away the lens cover from the camcorder, and, on the contrary, it provides for detecting a transition from presence of a useful video signal to absence of the useful video signal when the human operator put the lens cover on the camcorder. Also in this case, the method of the present invention does not require altering the capturing device or its manner of use, and does not require specific actions to be performed by the human operator (other than those typically performed by the operator for using the camcorder) in order to start and stop processing (and transmitting) the captured multimedia content.

CLAIMS

1. A method for controlling processing of multimedia content captured by a capturing device (11), said multimedia content being provided over a signal comprising a video signal or an audio signal, said method comprising the following steps:
   a) detecting in said signal a transition between absence of a useful signal and presence of said useful signal, wherein said useful signal is a portion of said signal comprising said multimedia content; and
   b) determining a change of an operating state of said capturing device (11) and controlling said processing of multimedia content on the basis of said detected transition.

2. The method according to claim 1, wherein said video signal is a digital video signal comprising a sequence of video frames, and wherein said detecting comprises checking whether a current video frame of said sequence of video frames is a black frame.

3. The method according to claim 2, wherein said checking comprises checking whether the value of the luma and the values of the chrominance components of a selected set of pixels within said video frame are comprised within, respectively, a luma interval of black color and a chrominance interval of black color.

4. The method according to claim 3, wherein said luma interval of black color is comprised between 0 and 16, and wherein said chrominance interval of black color is comprised between 127 and 129.

5. The method according to claim 3, wherein said pixels of said selected set of pixels are arranged along an horizontal axis and a vertical axis of said video frame subdividing the video frame in

four quadrants, and wherein said selected set of pixels also comprises at least four further pixels, each located at the center of a respective quadrant.

6. The method according to any of claims 2 to 5, wherein it further comprises, in case said current video frame of said sequence of video frames is not a black frame:
   - incrementing a useful video frame time interval (Tuv(i));
   - checking whether said useful video frame time interval (Tuv(i)) is greater than a first threshold time interval (Thv_on); and
   - in the affirmative, determining that said transition occurred from absence of a useful signal to presence of said useful signal,

   and wherein the method further comprises, in case said current video frame is a black frame,
   - incrementing a black frame time interval (Tb(i));
   - checking whether said black frame time interval (Tb(i)) is greater than a second threshold time interval (Thv_off); and
   - in the affirmative, determining that said transition occurred from presence of a useful signal to absence of said useful signal.

7. The method according to claim 6, wherein the value of said first threshold time interval (Thv_on) and of said second threshold time interval (Thv_off) is equal to 250 ms.

8. The method according to claim 1, wherein said audio signal is a digital audio signal comprising a sequence of audio frames, each audio frame comprising a sequence of audio samples, and wherein said detecting comprises checking whether a current audio frame of said sequence of audio frames is a null audio frame indicating silence.

9. The method according to claim 8, wherein said checking

comprises checking whether the value of each audio sample of said current audio frame is comprised within a null audio interval of values indicating silence.

10. The method according to claim 9, wherein said values of said null audio interval range between -16 and 16.

11. The method according to any of claims 8 to 10, wherein it further comprises, in case said current audio frame of said sequence of audio frames is not a null audio frame:
- incrementing a useful audio frame time interval (Tua(i));
- checking whether said useful audio frame time interval (Tua(i)) is greater than a third threshold time interval (Tha_on); and
- in the affirmative, determining that said transition occurred from absence of a useful signal to presence of said useful signal,

and wherein the method further comprises, in case said current audio frame is a null audio frame:
- incrementing a null audio frame time interval (Tn(i));
- checking whether said null audio frame time interval (Tn(i)) is greater than a fourth threshold time interval (Tha_off); and
- in the affirmative, determining that said transition occurred from presence of a useful signal to absence of said useful signal.

12. The method according to claim 11, wherein the value of said third threshold time interval (Tha_on) and of said fourth threshold time interval (Tha_off) is equal to 250 ms.

13. The method according to claim 6 or 11, wherein said controlling comprises, in case said transition from absence of a useful signal to presence of said useful signal is determined, switching on said processing said multimedia content.

14. The method according to claim 6 or 11, wherein said controlling comprises, in case said transition from presence of a useful signal to absence of said useful signal is determined, switching off said processing said multimedia content.

15. An analysis module (122) for a processing device (12) configured to process a multimedia content captured by a capturing device (11), said multimedia content being provided to the analysis module (122) over a signal comprising a video signal or an audio signal, said analysis module (122) being configured to :

- detect in said signal a transition between absence of a useful signal and presence of said useful signal, wherein said useful signal is a portion of said signal comprising said multimedia content; and
- determine a change of an operating state of said capturing device (11) and control said processing of multimedia content on the basis of said detected transition.

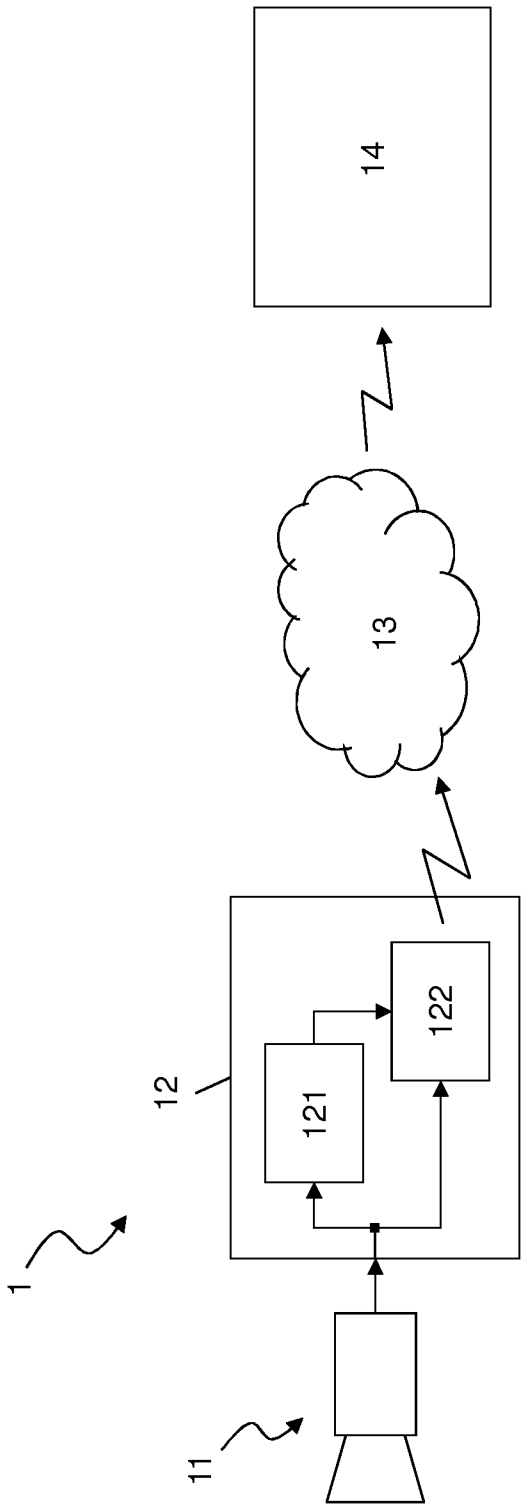
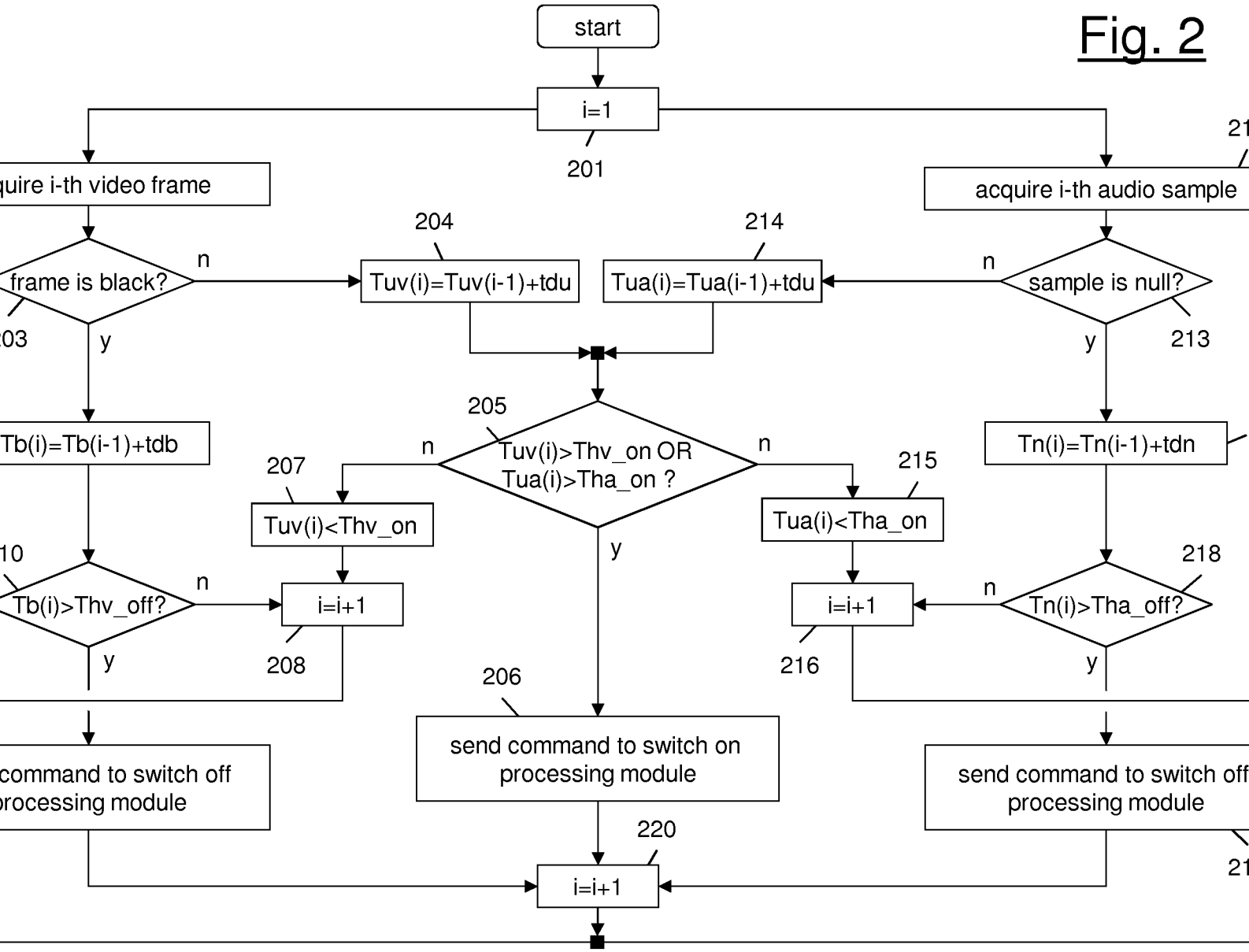


Fig. 1

Fig. 2


start
i=1
201
quire i-th video frame
acquire i-th audio sample
frame is black?
n
y
204
Tuv(i)=Tuv(i-1)+tdu
214
Tua(i)=Tua(i-1)+tdu
sample is null?
213
Tb(i)=Tb(i-1)+tdb
Tn(i)=Tn(i-1)+tdn
205
Tuv(i)>Thv_on OR
Tua(i)>Tha_on ?
207
Tuv(i)<Thv_on
215
Tua(i)<Tha_on
Tb(i)>Thv_off?
208
i=i+1
216
218
Tn(i)>Tha_off?
206
send command to switch on
processing module
command to switch off
processing module
send command to switch off
processing module
220
i=i+1

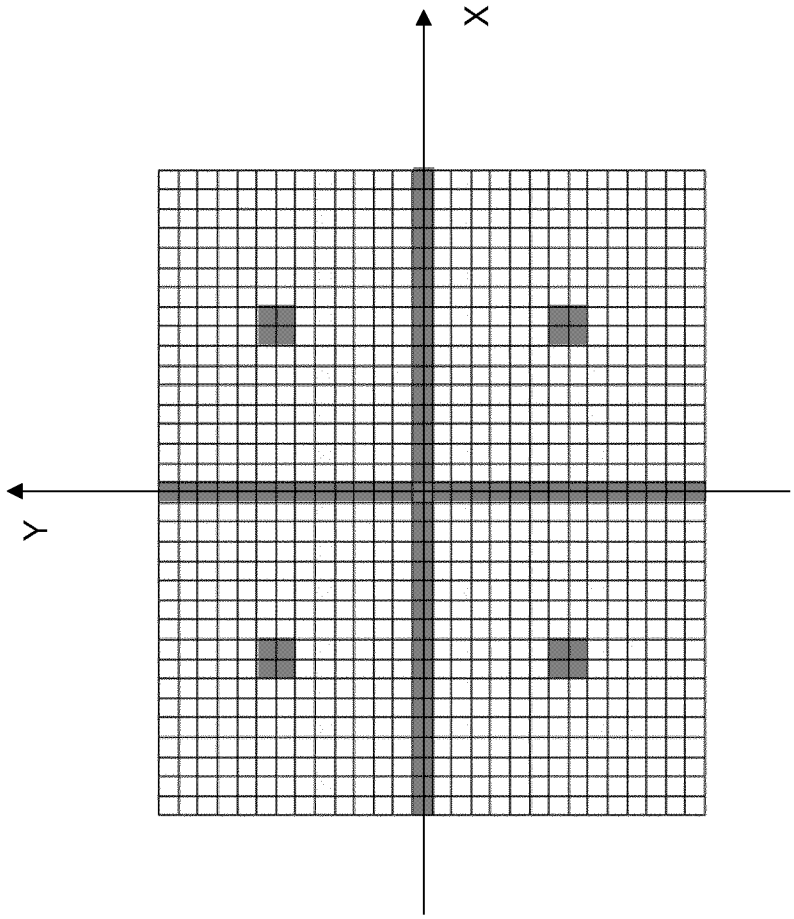


Fig. 3

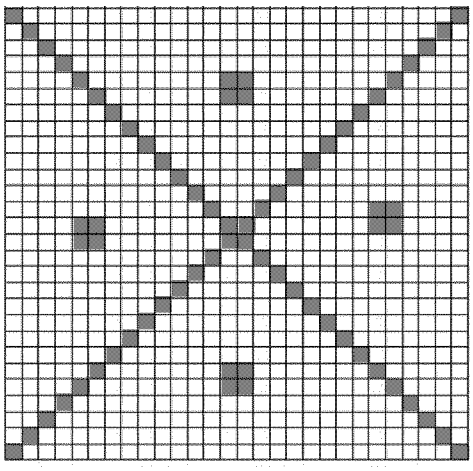


Fig. 4b

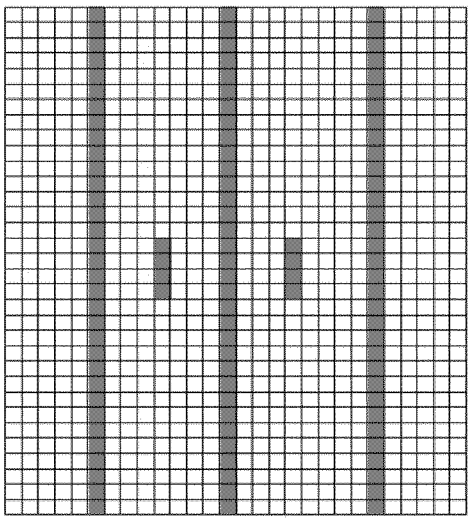


Fig. 4a

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2015/081414

A. CLASSIFICATION OF SUBJECT MATTER
INV. G11B27/22 G11B27/28 G06T7/00 H04H20/04 H04H60/58
H04H60/59 H04N5/14 H04N5/225
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)
G11B G06T H04H H04N

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.
X	US 2015/229980 A1 (REISNER SAMUEL JOSEPH [US]) 13 August 2015 (2015-08-13) the whole document	1,2,6-15
X	US 2006/230414 A1 (ZHANG TONG [US]) 12 October 2006 (2006-10-12)	1-3,6-15
Y	paragraph [0017] - paragraph [0066]	4,5
X	US 2012/095579 A1 (JUNG EDWARD K Y [US] ET AL) 19 April 2012 (2012-04-19)	1,8-10, 15
A	paragraph [0032] - paragraph [0048]	11
X	US 2004/095377 A1 (SALANDRO JEROME R [US]) 20 May 2004 (2004-05-20)	1,2,15
A	paragraph [0038] - paragraph [0041] paragraph [0062]	6,7,13, 14

-/--

[X] Further documents are listed in the continuation of Box C.

[X] 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	Date of mailing of the international search report
3 June 2016	13/06/2016

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Gil Zamorano, Gunnar

1

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2015/081414

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GB 2 243 969 A (BRITISH BROADCASTING CORP [GB]) 13 November 1991 (1991-11-13) page 3 - page 6	1,2,8,15
X	US 2002/030738 A1 (MOREINIS JOSEPH DANIEL [US] ET AL) 14 March 2002 (2002-03-14) paragraph [0024] - paragraph [0042]	1,15
Y	CN 102 006 499 B (DAYANG TECHNOLOGY DEV INC) 24 October 2012 (2012-10-24) paragraph [0032] - paragraph [0036]; figure 3	4,5
A	WO 00/07367 A2 (KONINKL PHILIPS ELECTRONICS NV [NL]; PHILIPS SVENSKA AB [SE]) 10 February 2000 (2000-02-10) page 10, line 26 - page 12, line 17; figure 10	2-5

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2015/081414

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2015229980 A1	13-08-2015	NONE	
US 2006230414 A1	12-10-2006	NONE	
US 2012095579 A1	19-04-2012	NONE	
US 2004095377 A1	20-05-2004	NONE	
GB 2243969 A	13-11-1991	NONE	
US 2002030738 A1	14-03-2002	NONE	
CN 102006499 B	24-10-2012	NONE	
WO 0007367 A2	10-02-2000	AT 395785 T	15-05-2008
		EP 1040656 A2	04-10-2000
		JP 2002521977 A	16-07-2002
		JP 2010246161 A	28-10-2010
		US 6100941 A	08-08-2000
		WO 0007367 A2	10-02-2000