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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
27 March 2014 (27.03.2014)

WIPO | PCT

(10) International Publication Number
WO 2014/044503 A1

(51) International Patent Classification:
A61B 5/00 (2006.01) *A61B 5/024* (2006.01)
A61B 5/053 (2006.01) *A61B 5/0488* (2006.01)
A61B 5/16 (2006.01) *G06F 17/30* (2006.01)
H04H 60/33 (2008.01)

(21) International Application Number:
PCT/EP2013/067717

(22) International Filing Date:
27 August 2013 (27.08.2013)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
12306134.3 19 September 2012 (19.09.2012) EP

(71) Applicant: **THOMSON LICENSING** [FR/FR]; 1 rue Jeanne d'Arc, F-92443 Issy-les-Moulineaux (FR).

(72) Inventors: **GUILLOTEL, Philippe**; c/o Technicolor R&D France, 975 avenue des Champs-Blancs, CS17616, F-35576 Cesson Sévigné (FR). **FLEUREAU, Julien**; c/o Technicolor R&D France, 975 avenue des Champs-Blancs, CS17616, F-35576 Cesson Sévigné (FR). **ORLAC, Izabela**; c/o Technicolor R&D France, 975 avenue des Champs-Blancs, CS17616, F-35576 Cesson Sévigné (FR).

(74) Agents: **LABELLE, Lilian** et al.; Technicolor, 1 rue Jeanne d'Arc, F-92443 Issy-les-Moulineaux (FR).

(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,

[Continued on next page]

(54) Title: METHOD AND DEVICE OF DETECTING THE EMOTIONAL IMPACT OF AN AUDIENCE WATCHING A TIME-STAMPED MOVIE

(57) Abstract: A method of detecting the emotional impact of an audience watching a time-stamped movie comprises the following steps: • - detecting the positive variations of a captured biological signal, and • - averaging said detected positive variations over a temporal window.

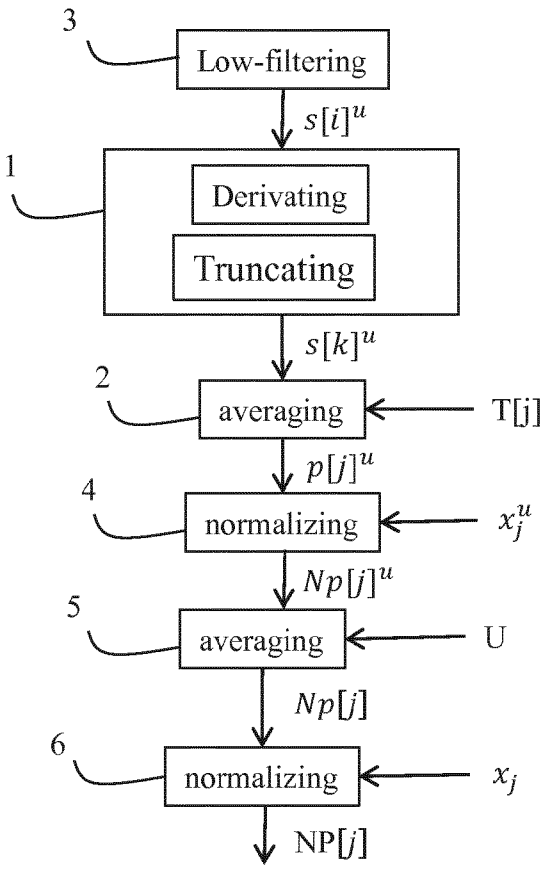


Fig. 1

WO 2014/044503 A1

DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, 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, 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).

Declarations under Rule 4.17:

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

Published:

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

Method and device of detecting the emotional impact of an audience watching a time-stamped movie.

1. Field of invention.

The present invention generally relates to the technical domain of physiological response of an audience watching a movie. More precisely it relates to a method of detecting the emotional impact of an audience watching a movie and reporting a corresponding emotional profile for this movie.

2. Technical background.

Catching the emotion of an audience member watching a movie clip may have many potential applications in a video content creation and distribution context. An advertising agency, a studio or a movie director trying to evaluate the emotional impact of its creation may be interested by such information. The creator maybe interested to validate his artistically choices and/or adapt the editing to minimize/maximize the emotional "level". A precise affective profile of a movie may be useful for predicting the potential impact of the movie and associated potential business. It may also help selecting the appropriate countries or population where to promote and distribute the movie.

Many biological signals are known to vary with the emotional state of a subject: Skin Conductivity (SC) measured through Galvanic Skin Response (GSR), SKin Temperature (SKT), Heart Rate Variability (HRV) measured by the means of a PhotoPlethysmoGraphy (PPG) devices or facial surface ElectroMyoGram (EMG) are some exemplary signals linked to emotional states. The link of such signals with the emotional state is described, for instance, in P. Lang, "The emotion probe", American Psychologist, vol. 50, no. 5, pp. 372–385, 1995; J. Wagner and E. Andre.

Nevertheless, obtaining the emotional state of an audience in an objective and low-intrusive manner is not an easy task.

One direct method would be to collect a direct self-assessment of each viewer minute per minute. This approach is clearly too intrusive, quite

subjective and may biased the reporting by distracting the subject from the movie.

Face analysis through an adapted camera could be an alternative, but the associated algorithms may be very sensitive to the user environment.

Correlating the user's emotion with the movie is another approach which tries to classify each segment of a movie into an emotional class.

All these methods do not provide any means which reflects the significant excitation variations of a whole audience over the time. They do not also provide any information to quantify the significance of highlights of a movie and means to compare the highlights of two movies one to each other.

3. Summary of the invention.

The present invention aims at alleviating some of the inconveniences of prior art.

More precisely, the invention relates to a method of detecting the emotional impact of an audience watching a timestamped movie. The method further comprises the following steps:

- detecting the positive variations of a captured biological signal, and
- averaging or cumulating said detected positive variations over a temporal window.

Thanks to the timestamps, the movie and the captured biological signal may be temporally synchronized together. Then, selecting multiple temporal windows to cover the whole duration of the movie, the method gets the mean reaction of a whole audience over the time. This mean reaction is much related to the study of the arousal fluctuations of this same audience during the show when considering a bi-dimensional "Valence/Arousal" representation of the emotion. Thus, the method leads to a time-variant affective signal (profile) which reflects the significant excitation variations of the whole audience over the time.

4. List of figures.

More advantages of the invention will appear through the description of particular, non-restricting embodiments of the invention.

The embodiments will be described with reference to the following figures:

- **Figure 1** shows a diagram of the steps of the method of detecting the emotional impact of an audience watching a time-stamped movie according to the invention.
- **Figure 2** shows an example of a temporal window which gathers samples of a biological signal captured from a user watching a movie.
- **Figure 3** shows an example of two overlapped temporal windows.
- **Figure 4** illustrates a method according to the present invention.
- **Figure 5** shows a device that comprises means configured to implement the method of the invention.

5. Detailed description of the invention.

Figure 1 shows a diagram of the steps of the method of detecting the emotional impact of an audience watching a time-stamped movie according to the invention.

Capturing a biological signal and synchronizing it with the timestamps of a movie are well-known techniques which are no more detailed in the following. For more details, refers to Julien Fleureau, Philippe Guillotel, Quan Huynh-Thu, "Physiological-Based Affect Event Detector for Entertainment Video Applications," *IEEE Transactions on Affective Computing*, Vol. 3, No 2, 2012.

N samples $s[i]^u$ are obtained from a biological signal captured from a user **u**.

According to an embodiment, the samples $s[i]^u$ are obtained from a raw Electro Dermal Activity (EDA) signal sampled at a predefined frequency, for example 32Hz.

An EDA signal measures the local variations of the skin electrical conductance in the palm area which is known to be highly correlated with the user affective arousal. This signal embeds high-level human-centered information and might be used to provide a continuous and unconscious feedback about the level of excitation of the end-user. More specifically, each shift in the arousal flow of a given user is correlated with slow (around 2s) phasic changes in his/her EDA consisting in a fast increasing step until a maximum value (the higher the emotional impact is, the higher the peak is), followed by a slow return to the initial state.

The detecting method comprises a step 1 of detecting the positive variations of the captured biological signal and a step 2 of averaging or cumulating said detected positive variations over a temporal window T[j].

Then, selecting multiple temporal windows to cover the whole duration of the movie, the highlights of the movie are temporally located in the temporal windows with the highest mean values of the detected positive variations computed over these temporal windows.

According to an embodiment of the detecting step, the biological signal is first derivate and truncated to positive values in order to highlight the relevant phasic changes. The truncation, for example, removes the detected negative variations.

Such an embodiment is illustrated by the **Figure 2** which is just given as an illustrative example and does not limit the scope of the invention.

Figure 2 shows an example of a temporal window T[j] which gathers N samples $s[i]^u$ of the biological signal captured from a user **u** watching a movie. First derivative of the biological signal quantifies the variations between the amplitudes of two consecutive samples $s[i-1]^u$ and $s[i]^u$. Next, only variations which are greater than the zero value are considered. Here, the variations d[1] and d[2] which have positive values are considered in the following and the variation d[3] which has a negative value is truncated (not considered in the following). By this way, K positive variations d[k], $k \in \{1; K\}$ are obtained for each temporal window T[j]. Each of the positive variation d[k] is relative to a sample $s[k]^u$ of the biological signal. For instance, the sample

$s[i]^u$ is relative to the positive variation d[1] on **Figure 2**. This sample defines the temporal instant of the beginning of a highlight of the movie.

According to the invention and as illustrated in **Figure 4**, a mean value $p[j]^u$ of positive detected variations over a temporal window T[j] is computed by equation (1):

$$p[j]^u = \frac{1}{K}\sum_{k=1}^{K} s[k]^u \quad (1)$$

with K the number of positive variations detected in the temporal window T[j].

M means values $p[j]^u$ are then obtained for the total duration T of the movie where M is the number of temporal windows which cover the total duration of the movie.

According to a variant, each mean value $p[j]^u$ is normalized (step 4) by a scaling factor x_j^u in order to remove the user-independent part related to the amplitude of the biological signal which may vary from one subject to another.

The normalized mean value $Np[j]^u$ is done, for example, by equation (2):

$$Np[j]^u = p[j]^u / x_j^u \quad \text{with} \quad x_j^u = \frac{1}{M}\sum_{j=1}^{M} p[j]^u \quad (2)$$

The normalized mean value $Np[j]^u$ may then be interpreted as a probability of reaction over the time of a current audience member. More precisely, the user-dependent part related to the amplitude of $p[j]^u$ has been removed and may allow a comparison between multiple users.

According to an embodiment, two consecutive temporal windows overlap one to each other.

A illustrated on **Figure 3**, a temporal window T[j+1] starts at the mid-point of a previous temporal window T[j]. If the duration of a temporal window equals 60s then the next window starts at 30s. This is sufficient to analyze the emotional flow in a movie. Other overlapping can also be used (not necessarily at the mid-point).

According to an embodiment, the method further comprises a step 3 of low-filtering the captured biological signal to remove possible artifacts. The cut frequency of the low-filter is 1 HZ for example. The positive variations of

the biological signal are thus detected from samples of a low-filtered version of the raw biological signal.

The mean values $p[j]^u$ (or their normalized versions $Np[j]^u$) only reflect the reaction of one given user **u**. Thus, these mean values are much related to the user's sensitivity and may be also affected by user-specific noises such as motions artifacts or user's external reactions for instance.

To obtain more relevant information concerning the arousal variations of a global audience during a movie, the method further comprises a step 5 of averaging the normalized mean values $Np[j]^u$ for all the users of an audience.

Mathematically speaking, the average value Np_j of the normalized mean values $Np[j]^u$ is given by equation (3):

$$Np[j] = \frac{1}{U}\sum_{u=1}^{U} Np[j]^u \qquad (3)$$

with U the number of users of the audience, $Np[j]^u$ the normalized mean value of the positive detected variations over a same temporal window T[j].

According to a variant the average value $Np[j]$ is scaled (step 6) by a factor x_j according to equation (6):

$$NP[j] = x_j * Np[j] \quad (4)$$

with $x_j = \frac{1}{U}\sum_{u=1}^{U} x_j^u$ the scaling factor.

The normalized average value $NP[j]$ takes into account the arousal fluctuations of the whole audience over the time.

Thus, by observing the variations of the normalized average values $NP[j]$ computed for temporal window which cover the wole duration T of the movie, one can affectively compare different time instants of a movie and thus identify highlights of a movie. But it is also possible to compare time instants of different movies to compare quantitatively their highlights.

As soon as the number of users of the audience is small, a strong inter-subject variability makes the interpretation, by an expert, of the quantity $NP[j]$ more tricky.

According to an embodiment, a criterion is defined to quantify the significance of the quantity $NP[j]$ in term of affective reaction.

According to a variant, such a criterion is mathematically defined as follows: for each temporal window, a random variable $X_{N_p}[j]$ distributed according to a probability density the sample of which is given by $\{Np[j]^1, \ldots, Np[j]^u, Np[j]^{u+1}, \ldots, Np[j]^U\}$ is defined.

A subset B of these M random variables is considered as being a background noise in terms of affective reaction. This subset B may be obtained by considering a selection criterion.

According to a variant, the selection criterion is keep the variables with the lowest average values $Np[j]$, for example the first decile.

For each $X_{N_p}[j]$ a relevancy value E[j] is computed by averaging the p-value of the bilateral Mann-Whitney-Wilcoxon test (also known as the U-test) performed between $X_{N_p}[j]$ and each element b of the subset B. A bilateral Mann-Whitney-Wilcoxon test is defined, for example by Kruskal, William H. (September 1957), "Historical Notes on the Wilcoxon Unpaired Two-Sample Test", Journal of the American Statistical Association 52 (279): 356–360.

Each $X_{N_p}[j]$ with an associated E[j] lower than a given threshold is considered as significantly different from the background noise in terms of affective reaction.

Figure 5 shows a device 500 that comprises means configured to implement the method of the invention. The device comprises the following components, interconnected by a digital data- and address bus 501:

- a processing unit 503 (or CPU for Central Processing Unit);
- a memory 505 ;
- a network interface 504, for interconnection of device 500 to other devices connected in a network via connection 502.

Processing unit 503 can be implemented as a microprocessor, a custom chip, a dedicated (micro-) controller, and so on. Memory 505 can be implemented in any form of volatile and/or non-volatile memory, such as a RAM (Random Access Memory), hard disk drive, non-volatile random-access memory, EPROM (Erasable Programmable ROM), and so on. Device 500 is suited for implementing a data processing device according to the method of the invention. The data processing device 500 and the memory 505 are

configured to detect the positive variations of a captured biological signal and to average or cumulate said detected positive variations over a temporal window.

According to an embodiment, the data processing device 500 and the memory 505 are also configured to first derivate the biological signal and to truncate the first derivate to positive values in order to highlight the relevant phasic changes.

According to an embodiment, the data processing device 500 and the memory 505 are also configured to compute a mean value $p[j]^u$ of positive detected variations over a temporal window T[j].

According to a variant, the data processing device 500 and the memory 505 are also configured to normalize a mean value $p[j]^u$ by a scaling factor x_j^u.

According to an embodiment, the data processing device 500 and the memory 505 are also configured to low-filtering the captured biological signal to remove possible artifacts.

According to an embodiment, the data processing device 500 and the memory 505 are also configured to average the normalized mean values $Np[j]^u$ for all the users of an audience.

According to a variant, the data processing device 500 and the memory 505 are also configured to scale the average value $Np[j]$ by a factor x_j.

According to an embodiment, the data processing device 500 and the memory 505 are also configured to define a criterion to quantify the significance of the quantity $NP[j]$ in term of affective reaction.

According to a particular embodiment, the invention is entirely implemented in hardware, for example as a dedicated component (for example as an ASIC, FPGA or VLSI) (respectively « Application Specific Integrated Circuit », « Field-Programmable Gate Array » and « Very Large Scale Integration ») or according to another variant embodiment, as distinct electronic components integrated in a device or, according to yet another embodiment, in a form of a mix of hardware and software.

CLAIMS

1. Method of detecting the emotional impact of an audience watching a time-stamped movie, the method comprises the following steps:

- detecting the positive variations of a captured biological signal, and
- averaging or cumulating said detected positive variations over a temporal window.

2. Method according to the claim 1, wherein multiple temporal windows are considered to cover the whole duration of the movie and said detected positive variations is averaged or cumulated over each of them.

3. Method according to the claim 2, wherein two consecutive temporal windows overlap one to each other.

4. Method according to any one of the preceding claims, wherein the captured biological signal is first derivate and truncated to positive values in order to highlight the relevant phasic changes.

5. Method according to any one of the preceding claims, wherein the captured biological signal is low-filtered.

6. Method according to any one of the preceding claims, wherein the averaged or cumulative value of the detected positive variations computed over a temporal window for each user of the audience is averaged for all the users of the audience.

7. Method according to any one of the preceding claims, wherein a criterion is defined to quantify the significance of the averaged or cumulative value of the detected positive variations computed over a temporal window in term of affective reaction.

8. Apparatus comprising means configured for implementing a method according to one of the claims 1-7.

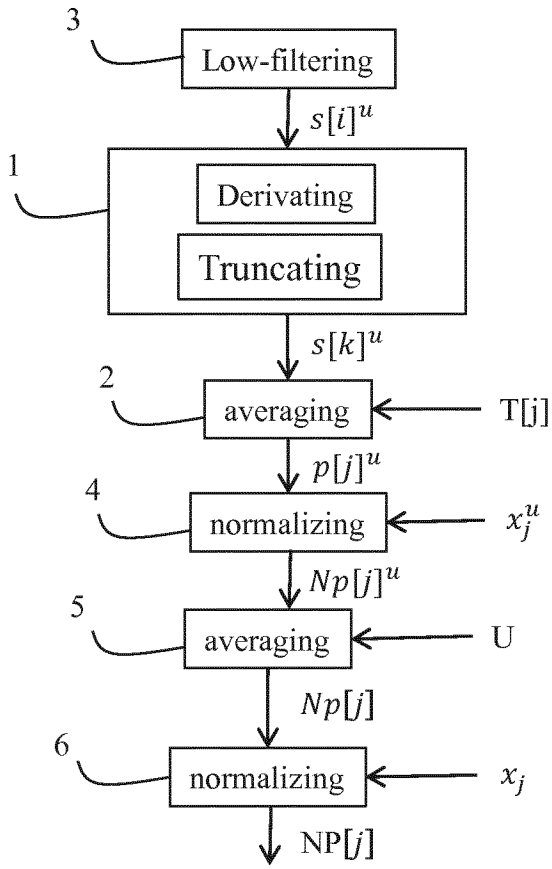


Fig. 1

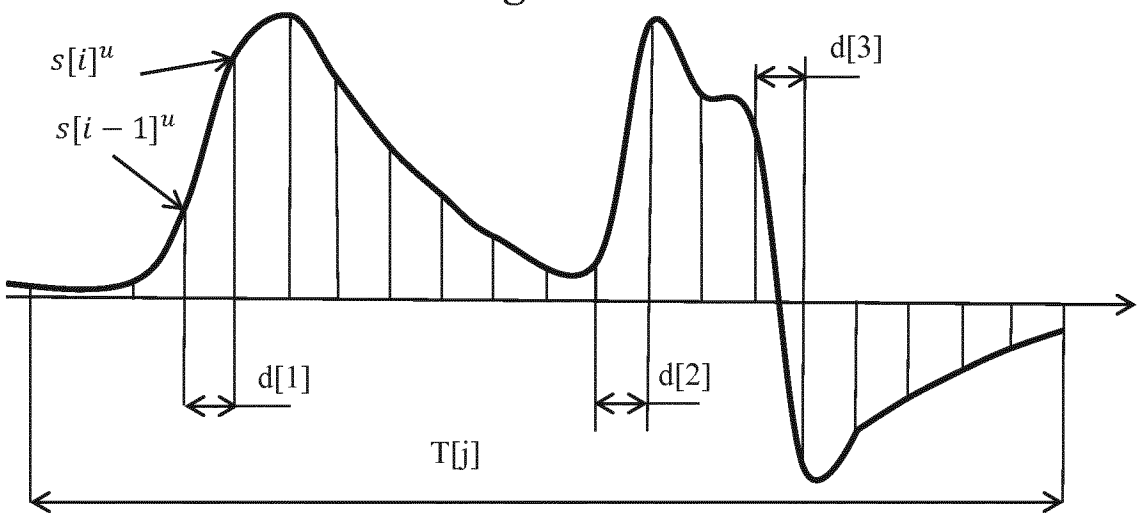


Fig. 2

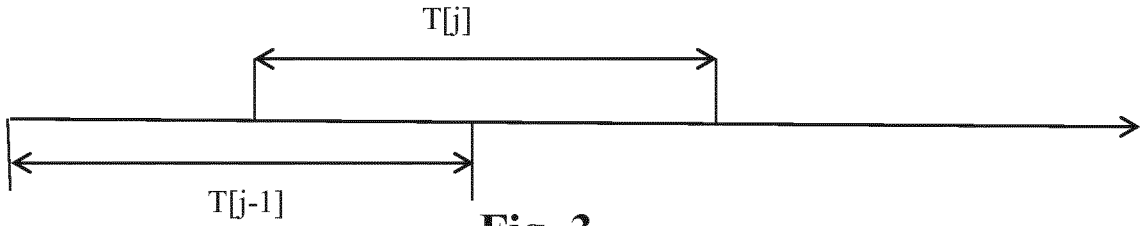


Fig. 3

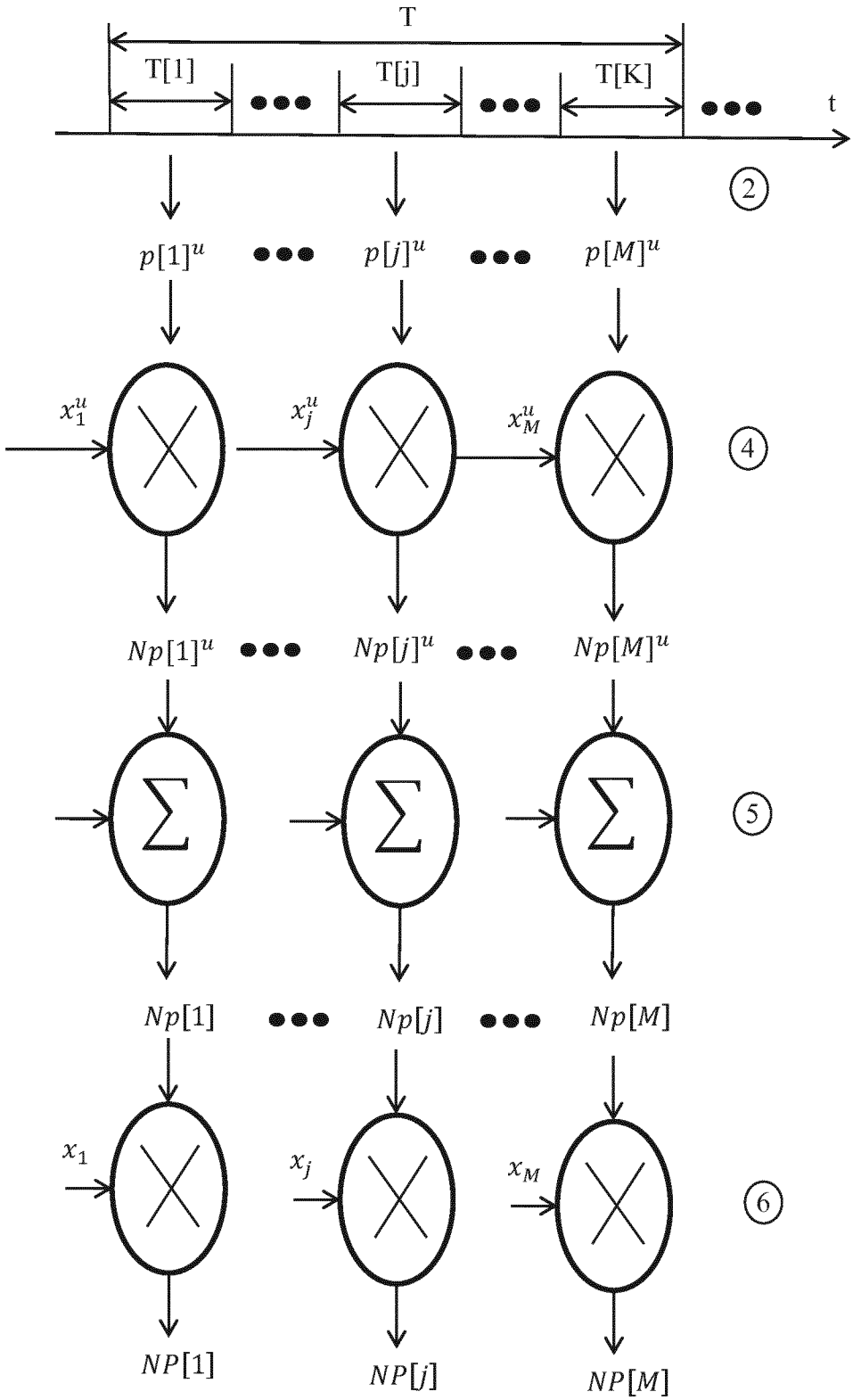

T
T[1]
T[j]
T[K]
t
2
$p[1]^u$
$p[j]^u$
$p[M]^u$
x_1^u
x_j^u
x_M^u
4
$Np[1]^u$
$Np[j]^u$
$Np[M]^u$
5
$Np[1]$
$Np[j]$
$Np[M]$
x_1
x_j
x_M
6
$NP[1]$
$NP[j]$
$NP[M]$

Fig. 4

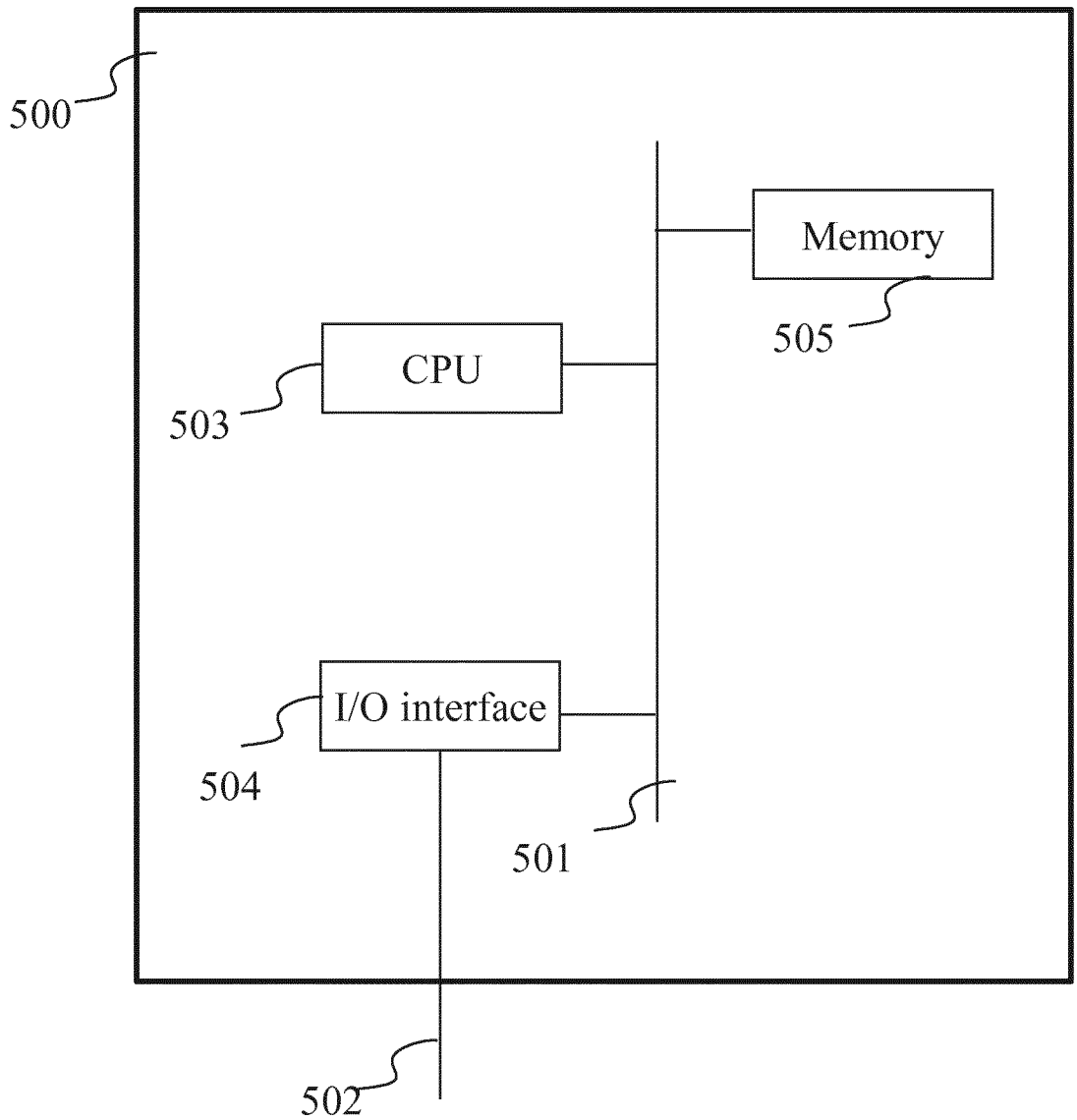


Fig. 5

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2013/067717

A. CLASSIFICATION OF SUBJECT MATTER
INV. A61B5/00 A61B5/053 A61B5/16 H04H60/33
ADD. A61B5/024 A61B5/0488 G06F17/30

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)
A61B G06F H04H

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	MOHAMMAD SOLEYMANI ET AL: "Continuous emotion detection in response to music videos", AUTOMATIC FACE&GESTURE RECOGNITION AND WORKSHOPS (FG 2011), 2011 IEEE INTERNATIONAL CONFERENCE ON, IEEE, 21 March 2011 (2011-03-21), pages 803-808, XP031869355, DOI: 10.1109/FG.2011.5771352 ISBN: 978-1-4244-9140-7 abstract table I chapter II.A chapter II.B chapter III chapter IV ----- -/--	1-8

[X] 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	Date of mailing of the international search report
30 October 2013	07/11/2013

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	Albrecht, Ronald

1

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SANDER KOELSTRA ET AL: "Single Trial Classification of EEG and Peripheral Physiological Signals for Recognition of Emotions Induced by Music Videos", 28 August 2010 (2010-08-28), BRAIN INFORMATICS, SPRINGER BERLIN HEIDELBERG, BERLIN, HEIDELBERG, PAGE(S) 89 - 100, XP019149400, ISBN: 978-3-642-15313-6 chapter 1 chapter 2.4 chapter 2.2 chapter 3.3 chapter 3.4	1-8

X	SOLEYMANI M ET AL: "Affective Characterization of Movie Scenes Based on Multimedia Content Analysis and User's Physiological Emotional Responses", MULTIMEDIA, 2008. ISM 2008. TENTH IEEE INTERNATIONAL SYMPOSIUM ON, IEEE, PISCATAWAY, NJ, USA, 15 December 2008 (2008-12-15), pages 228-235, XP031403541, ISBN: 978-0-7695-3454-1 abstract table 2 chapter 3.3.2 chapter 3.4.2	1-8

A	JULIEN FLEUREAU ET AL: "Physiological-Based Affect Event Detector for Entertainment Video Applications", IEEE TRANSACTIONS ON AFFECTIVE COMPUTING, IEEE, USA, vol. 3, no. 3, 1 July 2012 (2012-07-01), pages 379-385, XP011466977, ISSN: 1949-3045, DOI: 10.1109/T-AFFC.2012.2 cited in the application chapter 3.1 chapter 4.3	2,3,7

1