



(51) International Patent Classification:

H05B 37/02 (2006.01) *H04N 21/44* (2011.01)
A63J 17/00 (2006.01) *G06F 16/683* (2019.01)
H04N 21/41 (2011.01) *G06F 16/60* (2019.01)
H04N 21/439 (2011.01)

(21) International Application Number:

PCT/EP2019/081814

(22) International Filing Date:

19 November 2019 (19.11.2019)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

18209056.3 29 November 2018 (29.11.2018) EP

(71) Applicant: **SIGNIFY HOLDING B.V.** [NL/NL]; High Tech Campus 48, 5656 AE Eindhoven (NL).

(72) Inventors: **ALIAKSEYEU, Dzmitry, Viktorovich**; c/o High Tech Campus 7, 5656 AE Eindhoven (NL). **SIRAJ, Muhammad, Mohsin**; c/o High Tech Campus 7, 5656 AE Eindhoven (NL). **MAES, Jérôme, Eduard**; c/o High Tech Campus 7, 5656 AE Eindhoven (NL).

(74) Agent: **VAN EEUWIJK, Alexander, Henricus, Walterus** et al.; Signify Netherlands B.V., High Tech Campus 7, 5656 AE Eindhoven (NL).

(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, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP,

(54) Title: A CONTROLLER FOR PROVIDING LIGHT SCRIPTS FOR MEDIA FRAGMENTS AND A METHOD THEREOF

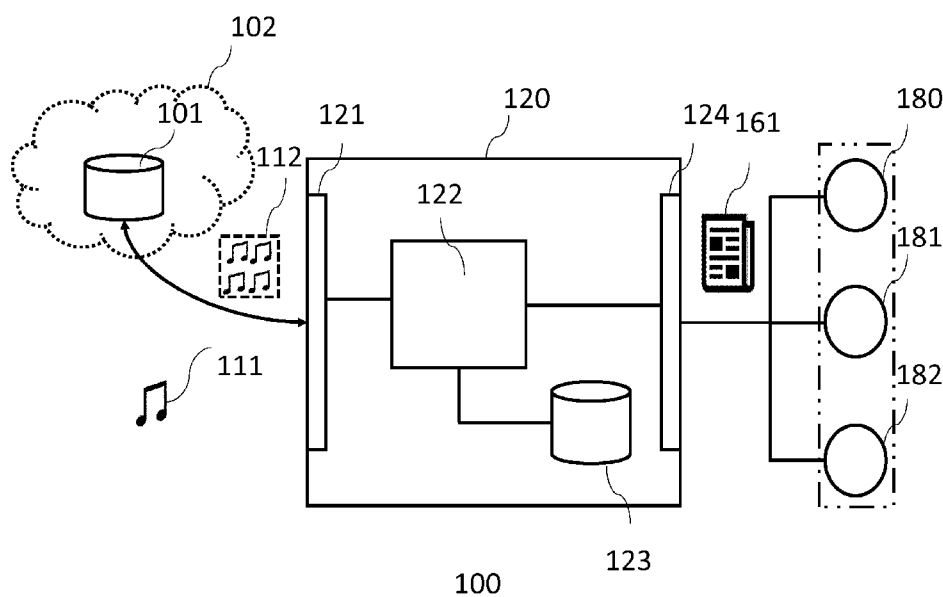


Fig. 1

(57) Abstract: A method of providing light scripts for media fragments is presented, wherein the method comprising: obtaining a first media fragment and a plurality of second media fragments, wherein the first media fragment comprises a first audio fragment; and wherein the plurality of second media fragments comprises second audio fragments, and wherein each of the plurality of second media fragments is associated with respective light scripts, analyzing at least one audio characteristic of the first media fragment, analyzing audio characteristics of the plurality of second media fragments, comparing the analyzed at least one audio characteristic of the first media fragment with the analyzed audio characteristics of the plurality of second media fragments, selecting one or more second media fragments from the plurality of second media fragments based on the comparison, providing a light script based on light scripts associated with the selected one or more second media fragments, associating the (provided) light script with the first media

KR, KW, 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).

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))*

A controller for providing light scripts for media fragments and a method thereof

FIELD OF THE INVENTION

The invention relates to a method of providing light scripts for media fragments. The invention further relates to a controller for providing light scripts for media fragments.

5

BACKGROUND

Light can be used to enhance entertainment experiences such as audio-visual media. A well-known add-on of light to video content is technology which augments the video experience for a user/viewer by controlling nearby luminaires to create lighting effects which are perceivable by the user and appear to match the video (e.g. the same color as the overall image at a point in time). These effects can be dynamic. To provide such effects, some hardware such as TVs are also being fitted with built-in light. Similarly, audio fragments, e.g. with a song, can be augmented with the light effects to enhance the listening experience.

10

15

Light effects to accompany entertainment experiences are commonly specified in “light scripts”. A light script (also called a “lighting script”, or just “script”) is a data structure defining particular lighting effects to be rendered by one or more luminaires over a time period. The light script is accessed by the lighting system to “play” it alongside the entertainment experience, by interpreting it to control the luminaires of the lighting system in accordance with the effects defined in the lighting script.

20

US 2018/061438 A1 discloses a system and method for predictively generating visual experiences based on streaming audio. More specifically, the method is directed to analyzing streaming audio and predictively mapping the information in the stream to a sequence of visual patterns generated by a lighting system in a manner that induces a perceptual association between the streaming audio and visual patterns.

25

US 2014/072272A1 discloses a method of generating data for controlling a rendering system includes obtaining data representative of a recording of at least intervals of an event, the recording having at least two components obtainable through different respective modalities. The data is analyzed to determine at least a dependency between a first

and a second of the components. At least the dependency is used to provide settings for a system for rendering in perceptible form at least one output through a first modality in dependence on at least the settings and on at least one signal for rendering in perceptible form through a second modality.

5

SUMMARY OF THE INVENTION

The inventors have realized that not all audio or video fragments have a light script associated with them. Currently, software algorithms are used to analyze these audio/video fragments and generate light scripts based on certain characteristics of the audio/video fragments. Generating light scripts by analyzing the characteristics of an audio/video fragment may result in unwanted light effects, thereby deteriorating a user's entertainment experience.

It is therefore an object of the present invention to provide a method which improves the user's entertainment experience by providing light scripts for those audio/video fragments, e.g. songs, videos etc., not having associated light scripts. It is a further object to minimize the risk of generating unwanted light effects.

According to a first aspect, the object is achieved by a method of providing light scripts for media fragments, the method comprising: obtaining a first media fragment and a plurality of second media fragments, wherein the first media fragment comprises a first audio fragment; and wherein the plurality of second media fragments comprises second audio fragments, and wherein each of the plurality of second media fragments is associated with respective light scripts, analyzing at least one audio characteristic of the first media fragment, analyzing audio characteristics of the plurality of second media fragments, comparing the analyzed at least one audio characteristic of the first media fragment with the analyzed audio characteristics of the plurality of second media fragments, selecting one or more second media fragments from the plurality of second media fragments based on the comparison, providing a light script based on light scripts associated with the selected one or more second media fragments, associating the (provided) light script with the first media fragment, such that when the first media fragment is being rendered by a media fragment rendering device, one or more lights are controlled according to one or more lighting instructions defined by the (provided) light script.

The method provides a light script for the first media fragment which does not have a light script associated with it. The first and the plurality of the second media fragment can be audio fragments, e.g. a song, and/or be audio fragments comprised in video fragments,

e.g. a movie. The method comprises analyzing at least one audio characteristic of the first media fragment, and further comprises analyzing audio characteristics of the plurality of second media fragments. The method may further comprise analyzing at least one audio characteristic of the first audio fragment, and further comprises analyzing audio

characteristics of the plurality of second audio fragments. Each of the plurality of second media fragments is associated with respective light scripts. A second media fragment is selected based on a comparison of the at least one audio characteristic of the first media fragment and the audio characteristics of the second media fragments. The comparison may be based, for example, on determining a similarity measure. The selection of the second media fragment is then based on the determined similarity measure.

The method may further comprise providing a light script associated with the selected second media fragment and associating the light script with the first media fragment. When the first media fragment, e.g. a song, is being rendered by a media fragment rendering device, e.g. a speaker, one or more lights are controlled according to one or more lighting instructions defined by the light script. Therefore, the method provides light scripts for those songs and/or for videos which do not have associated light scripts, wherein the method is based on analyzing and comparing the audio characteristic of the song/video with audio characteristics of the plurality of songs/videos having a respective associated light script, the method then selects a song/video from the plurality of songs/videos of which the audio characteristics substantially matches the audio characteristics of the song/video of interest. Hence, by rendering the light script of the best matched song/video when the song/video of interest is played, the user's entertainment experience is improved and the risk of generating unwanted light effects is reduced.

The light scripts can be typically categorized in at least 3 different categories: 'compiled scripts', which comprise control commands or instructions intended for one of more lights; 'high-level light scripts', which are independent of a lighting infrastructure of an environment and describe light effects in a more abstract way, such that they can then be compiled depending on the user's setup into a compiled script; and 'light scripts describing content' which do not comprise control commands or instructions, but rather comprise all the information about the content of the light script sufficient to create either a compiled or a high-level light script.

The at least one audio characteristic of the first media fragment and the audio characteristics of the plurality of second media fragments may comprise one or more of: beat, timbre, pitch, intensity, rhythm, major and minor key.

Additionally, or alternatively the at least one audio characteristic of the first media fragment and the audio characteristics of the plurality of second media fragments may comprise audio features. For example, the audio features may comprise, e.g. direct mood, valence and arousal/energy. Direct mood may be estimated using set of mood labels. A combination of valence and arousal may be used for defining mood. In an embodiment, new features may be defined based on the first media fragment and the plurality of second media fragments.

Additionally, the method may further comprise: clustering a plurality of audio characteristics of the first media fragment into a plurality of clusters based on a similarity criterion, analyzing the plurality of clusters, comparing the analyzed plurality of clusters with the analyzed audio characteristics of the plurality of second media fragments, selecting a subset of the plurality of second media fragments based on the comparison, providing the light script by generating the light script based on light scripts associated with the subset of the plurality of second media fragments, and associating the generated light script with the first media fragment.

The first media fragment, e.g. an audio fragment, may have distinct sets or clusters of properties. For example, one cluster of properties may describe the energy and dynamism of the audio fragment. These clusters may then be compared with the audio characteristics of the plurality of second media fragments, wherein for each set of properties, i.e. for each cluster, different media fragments from the plurality of second media fragments can be selected. The light script is then generated from multiple light scripts from the selected subset of the plurality of second media fragments, wherein different features of the script are selected from different scripts based on the comparison with clusters of the first media fragment.

As different clusters represent different sets of properties, e.g. energy, valence etc., of an audio fragment, different weights may be assigned to these clusters; wherein the selection of the subset may be based on these weights. The weights may be assigned based, for example, on a user preference via a user interface device such as a smart phone or on a learning model, wherein the learning model is developed based on historic data.

Alternatively, the method may further comprise: identifying substantially similar audio characteristics of the plurality of second media fragments, clustering the plurality of second media fragments into a plurality of clusters based on the identification, generating a plurality of master light scripts based on the plurality of clusters, and storing the plurality of master light scripts.

According to this embodiment, the method may further comprise: comparing the analyzed at least one audio characteristic of the first media fragment with the plurality of master light scripts, selecting at least one master light script from the plurality of master light scripts based on the comparison, and associating the selected at least one master light script with the first media fragment.

The audio characteristics of the plurality of second media fragments may be clustered according to a similarity criterion, and a 'master script' may be generated based on these clusters. A benefit of applying the master light script is that, from a user's perspective, the light effects for similar songs (e.g. of the same artist) are consistent. Another benefit is that the user is not confronted with a lot of options for selecting a light script for the first audio fragment.

The step of comparing may be performed by using similarity learning, wherein the similarity learning uses at least one of: regression similarity learning, classification similarity learning, ranking similarity learning and locality sensitive hashing.

In order to improve the reliability of the comparison step, machine learning algorithms such as similar learning can be used at the comparison step. Other methods known in the art for comparing audio characteristics may also be used.

The second media fragment may be selected if a difference between the analyzed at least one audio characteristic of the first media fragment and an analyzed at least one audio characteristic of the selected second media fragment exceeds a threshold. The threshold may be defined as a value such that when the difference is lower than, equal to or higher than the threshold, the second media fragment is selected. Alternatively, the threshold may be defined as a range.

The method may further comprise: modifying the light script based on a difference between the analyzed at least one audio characteristic of the first media fragment and the analyzed at least one audio characteristic of the selected second media fragment.

The modification of light script may at least comprise: modifying the color palette of the light script, increasing/decreasing the dynamic range of the light script, i.e. limiting or extending a range between the lowest and the highest brightness of the light effects, replacing or excluding some of the light effects in the light script, e.g. brightness, intensity, color etc.

The comparison of the analyzed at least one audio characteristic of the first media fragment with the analyzed audio characteristics of the plurality of second media fragments may be based on a user-based subset of the plurality of second media fragments,

wherein the user-based subset is based on a user playlist. The plurality of second media fragments may be based on the user's playlist which, e.g. indicates user choices, psychological or physiological state of the user, e.g. mood, and activity. This allows the selected second media fragment and the associated light script to match the user preference.

5 In an embodiment, when a group of second media fragments has been selected based on the comparison, the step of selecting the second media fragment may further comprise: selecting a second media fragment from the group based on one or more of: a user preference, a lighting setup of a user, a psychological or physiological state of a user, a previous light script used by a user, the time of day, an activity of a user.

10 When the step of comparing, for instance, results in more than one to-be-selected second media fragment, and the selection of, e.g. a single second media fragment and the light script thereof, may become relevant problem. This embodiment improves the selection process and further allows the other aspects such as the lighting setup, psychological or physiological state, e.g. mood, and/or preference of user etc. to be
15 considered for an optimal selection of the second media fragment.

The first media fragment may be an audio fragment and the plurality of second media fragments may comprise audio fragments.

Alternatively, the first media fragment may be a video fragment comprising an audio fragment, and the plurality of second media fragments may comprise audio fragments.

20 According to another aspect of the present invention, the object is achieved by a controller for providing light scripts for media fragments, the controller comprising: an input configured to obtain a first media fragment and a plurality of second media fragments, wherein the first media fragment comprises a first audio and/or a first video fragment, and wherein the plurality of second media fragments comprises second audio fragments and/or
25 second video fragments, and wherein each of the plurality of second media fragments is associated with respective light scripts, a processor configured to execute steps of the method, an output configured to output information about the association.

According to another aspect, a computer program product is provided which comprises instructions configured to execute the steps of the method.

30 It should be understood that the computer program product and the controller may have similar and/or identical embodiments and advantages as the above-mentioned methods.

According to another aspect, the object is achieved by a method of providing light scripts for media fragments, the method comprising: obtaining audio characteristics of a first media fragment and a plurality of second media fragments, respectively, wherein the

first media fragment comprises a first audio fragment; and wherein the plurality of second media fragments comprises second audio fragments, and wherein each of the plurality of second media fragments is associated with respective light scripts; selecting one or more second media fragments from the plurality of second media fragments, wherein selection of the one or more second media fragments are based on a similarity criterion between the first media fragment and the plurality of second media fragments; providing a light script based on light scripts associated with the selected one or more second media fragments; associating the (provided) light script with the first media fragment, such that when the first media fragment is being rendered by a media fragment rendering device, one or more lights are controlled according to one or more lighting instructions defined by the (provided) light script.

In the context of the present invention the term “audio characteristic” is to be understood as any characteristic related to auditory/audible properties of the media fragment. For example, audio characteristic may be the beat, timbre, pitch, intensity, rhythm, major and minor key of the media fragment.

BRIEF DESCRIPTION OF THE DRAWINGS

The above, as well as additional objects, features and advantages of the disclosed systems, devices and methods will be better understood through the following illustrative and non-limiting detailed description of embodiments of devices and methods, with reference to the appended drawings, in which:

Fig. 1 shows schematically a system comprising a controller for providing light scripts for media fragments,

Fig. 2 shows schematically a flowchart illustrating an embodiment of a method of providing light scripts for media fragments,

Fig. 3 shows schematically a flowchart illustrating another embodiment of a method of providing light scripts for media fragments,

Fig. 4 shows schematically a flowchart illustrating another embodiment of a method of providing light scripts for media fragments, and

Fig. 5 shows schematically a controller for providing light scripts for media fragments.

All the figures are schematic, not necessarily to scale, and generally only show parts which are necessary in order to elucidate the invention, wherein other parts may be omitted or merely suggested.

DETAILED DESCRIPTION OF EMBODIMENTS

Fig. 1 shows schematically a system 100 comprising a controller 120 for providing light scripts for media fragments. The controller 120 comprises a processor 122, an input 121 and an output 124. Optionally, the controller 120 may comprise a memory 123. The controller is configured to obtain a first media fragment 111 and a plurality of second media fragments 112.

The first media fragment 111 may be an audio fragment, e.g. a song, and the plurality of second media fragments 112 may comprise audio fragments, e.g. a collection of songs. Each of the plurality of second media fragments 112 is associated with respective light scripts. The first media fragment 111 may be a song from a singer and the plurality of second media fragments 112 may be a collection of songs from the same singer or from the same album. The first media fragment 111 may be a song from a specific genre, e.g. pop, disco etc., or from an era, e.g. 70s, 80s music etc., and the plurality of second media fragments 112 may be a collection of songs from the same genre or from the same era. In another example, the first media fragment 111 may be a movie with a sound track belonging to a genre, e.g. horror, romantic etc., and the plurality of second media fragments 112 may be a collection of movies with different sound tracks belonging to the same genre. Any other combination, for instance, a song from a genre as the first media fragment 111 and the collection of songs as the second media fragments 112 from other genre are also possible. The media fragments in the plurality of second media fragments 112 may be part of a user playlist. Alternatively, these media fragments may be selected based, for instance, on a user preference via a user interface (not shown), e.g. a smart phone, a tablet pc, etc., or automatically selected, e.g. based on a learning algorithm.

The controller 120 may obtain the first and/or the plurality of second media fragments 111-112 from a server 102. The controller 120 may also obtain the first and/or the plurality of second media fragments from a memory 101 which is external to the controller 120. The memory 101 may be located in the server 102 or at least external to the controller 120. The first and/or the plurality of second media fragments 111-112 may also be stored in the controller internal memory 123. The controller 120 may obtain the first media fragment 111 from the memory 123 and may obtain the plurality of second media fragments externally from the controller 120, e.g. from the memory 101 or the server 102. The controller 120 may obtain the first and/or the plurality of second media fragments 111-112 from a user (not shown) via a user interface device (not shown). The media fragments may be obtained from a

music streaming or video sharing platform such as Spotify, YouTube etc. Any other ways to obtain the first and the plurality of second media fragments 111-112 may also be possible.

The processor 122 may obtain the first and the plurality of second media fragments 111-112 from the input 121 and/or from the memory 123. The processor 122 analyzes the at least one audio characteristic of the first media fragment 111 and the audio characteristics of the second media fragments 112. In an embodiment, the analyzed at least one audio characteristic of the first media fragment 111 and the audio characteristics of the plurality of second media fragments 112 may comprise one or more of: beat, timbre, pitch, intensity, rhythm, major and minor key. For the analysis, the processor 122 may use algorithms from the field of machine learning, such as supervised learning and recognition. The processor 122 may also use other known methods to analyze the audio characteristics of media fragments.

The processor 122 is further configured to compare the analyzed at least one audio characteristic of the first media fragment 111 with the analyzed audio characteristics of the plurality of second media fragments 112. The processor 122 may be configured to perform comparison, for example, by determining a similarity measure or similarity function based on the analyzed at least one audio characteristic of the first media fragment 111 and the analyzed audio characteristics of the plurality of second media fragments 112. In an embodiment, the first media fragment 111 may be an audio fragment and the second media fragments 112 may comprise audio fragments. Alternatively, the first media fragment 111 may be a video fragment comprising an audio fragment, and the second media fragments 112 may comprise audio fragments.

The processor 122 is further configured to select a second media fragment based on the comparison. The processor 122 may use a threshold and the second media fragment is selected if a difference between the analyzed at least one audio characteristic of the first media fragment 111 and an analyzed at least one audio characteristic of the selected second media fragment exceeds the threshold. In an example, the selected second media fragment may be the same first media fragment 111, for instance the first media fragment 111 is a song and the second media fragment is also the same song but obtained from a different media streaming or video sharing platform such as Spotify, YouTube etc., wherein the song in the other platform has an associated light script. The threshold may be provided by the user or set by a learning algorithm based on historic data.

The processor 122 is further configured to provide a light script 161 associated with the selected second media fragment. The processor 122 is further configured to associate

the light script 161 with the first media fragment 111. The processor 122 is further configured to output the associated light script 161, e.g. to a plurality of light sources 180, 181, 182, such that when the first media fragment 111 is being rendered by a media fragment rendering device, one or more lights, in this example light sources 180, 181, 182, are controlled according to one or more lighting instructions defined by the light script 161. The one or more lighting sources 180, 181, 182 may be any type of lighting devices arranged for receiving lighting control commands. The lighting devices may comprise an LED light source, an incandescent light source, a fluorescent light source, a high-intensity discharge light source, etc. In an example, when the selected light script is not temporally synchronized with the first media fragment 111, the processor 122 is further configured to temporally synchronized with the first media fragment 111. The selected light script 161 may be is stored in, e.g. the server 102, internal 123 or external memory 101, such that anytime when the first media fragment 111 is being rendered by a media fragment rendering device, the selected light script 161 can be retrieved.

Fig. 2 shows schematically a flowchart illustrating an embodiment of a method 200 of providing light scripts for media fragments. In an obtaining step 205, the first media fragment 111, e.g. a song, and the plurality of second media fragments 112, e.g. a collection of songs, may be obtained, for instance, from a server 102 or from an internal 123 or external memory 101 or from a user via a user interface device, e.g. a smart phone. Other combinations of obtaining the first media fragment 111 and the plurality of second media fragments 112 are also possible.

The method 200 may further comprise the step of analyzing 210 the at least one audio characteristic of the first media fragment 111. The method 200 may further comprise the step of analyzing 220 the audio characteristics of the second media fragments 112. Audio characteristics are related to a listening experience of a user. These audio characteristics may comprise one or more of: beat, timbre, pitch, intensity, rhythm, major and minor key of the media fragments. Other characteristics of the media fragments associated with the listening experience of a user are not excluded. Different algorithms, e.g. from the field of machine learning or from audio processing, known in the art can be used for the analysis steps 210, 220. The analysis 210, 220 may be focused on identifying the dominant audio characteristics of the media fragments, e.g. to determine if a song has a high pitch, or if a song is played in a major key or in a minor key. The analysis 210, 220 may be based, for example on individual properties, e.g. intensity, pitch etc, or on a group of properties

representing a particular aspect, e.g. energy, of the media fragment. The steps of analyzing 210, 220 may be performed in any order.

The method 200 may further comprise the step of comparing 230 the analyzed audio characteristics of the media fragments. For example, the dominant identified audio characteristics, e.g. a song with a high pitch, is compared with the collection of songs with high pitches. The step of comparing 230 may be performed by using similarity learning, and wherein the similarity learning uses at least one of: regression similarity learning, classification similarity learning, ranking similarity learning and locality sensitive hashing. The to-be-compared audio characteristic, e.g. intensity of the audio of media fragments, are denoted as (x_i^1, x_i^2) with their similarity $y_i \in R$, wherein x_i^1 is the i^{th} data point from the first media fragment 111 and x_i^2 is the i^{th} data point from the second media fragment 112. The goal of these learning algorithms is to learn a similarity function based on the data (x_i^1, x_i^2) .

In the regression learning, the goal is to learn a function that approximates $f(x_i^1, x_i^2) \sim y_i$ for every new labeled triplet example (x_i^1, x_i^2, y_i) . This is typically achieved by minimizing a regularized loss $\min_w \sum_i \text{loss}(w; x_i^1, x_i^2, y_i) + \text{reg}(w)$.

In the classification similarity learning, given are pairs of similar objects (x_i, x_i^+) and non-similar objects (x_i, x_i^-) . An equivalent formulation is that every pair (x_i^1, x_i^2) is given together with a binary label $y_i \in \{0,1\}$ that determines if the two objects are similar or not.

The goal is again to learn a classifier that can decide if a new pair of objects is similar or not.

In the ranking similarity learning, given are triplets of objects (x_i, x_i^+, x_i^-) whose relative similarity obey a predefined order: x_i is known to be more similar to x_i^+ than to x_i^- . The goal is to learn a function f such that for any new triplet of objects (x, x^+, x^-) , it obeys $f(x, x^+) > f(x, x^-)$. This setup assumes a weaker form of supervision than in regression, because instead of providing an exact measure of similarity, one only has to provide the relative order of similarity.

In the Locality sensitive hashing (LSH), input items are hashed so that similar items map to the same "buckets" in memory with high probability (the number of buckets being much smaller than the universe of possible input items).

Other algorithms known for comparing audio characteristics, e.g. from audio processing, are not excluded. The comparison of the analyzed at least one audio characteristic of the first media fragment 111 with the analyzed audio characteristics of the plurality of

second media fragments 112 may be based on a user-based subset of the plurality of second media fragments, wherein the user-based subset is based on a user playlist.

The method 200 may further comprise the step of selecting 240. The second media fragment may be selected if a difference between the analyzed at least one audio characteristic of the first media fragment 111 and an analyzed at least one audio characteristic of the selected second media fragment exceeds a threshold. For example, if certain characteristics of the audio fragments do not match (e.g. 140 bpm of the first audio fragment vs. 142 bpm of the second audio fragment) the second audio fragment may still be selected if the difference is within a threshold range. The threshold may be defined such that when the difference is lower than, equal to or higher than the threshold, the second media fragment is selected. Alternatively, the threshold may be defined as a range. In case, when a group of second media fragments has been selected based on the comparison 230, e.g. for a high pitch song, a group of second songs has been selected all having high pitch, such that the difference between the pitch level of the song and the selected group of songs does not exceed the threshold, the step of selecting 240 the second media fragment 112 may further comprise: selecting a second media fragment from the group based on one or more of: a user preference, a lighting setup of a user, a psychological or physiological state, e.g. mood, of a user, a previous light script used by a user, the time of day, an activity of a user. For example, different users have different lighting setup, e.g. different number/type of light sources, and not all the light script suits the lighting setup of the user. Therefore, in case of selection of the second media fragment from the group, this embodiment allows to consider the lighting setup. The physiological or psychological state, e.g. the mood, and/or activity may also be considered such that the selected second media fragment is according to the current mood and/or activity of the user. The physiological or psychological state of the user can be implicitly measured using wearable personal devices or any other devices, or explicitly indicated by the user. Furthermore, based on the historic data of user selection, a learning model may be developed which can act as a recommendation function to provide recommendation for the selection of a second media fragment from the group of selected second media fragment.

The method 200 may further comprise the step of providing 250 a light script 161 associated with the selected second media fragment or generating a light script 161 based on the selected second media fragment. The step of providing 250 may further comprise modifying a light script based, e.g. on a difference between the analyzed at least one audio

characteristic of the first media fragment 111 and the analyzed at least one audio characteristic of the selected second media fragment.

The method 200 may further comprise the step of associating 260 the light script 161 with the first media fragment 111, such that when the first media fragment 111 is being rendered by a media fragment rendering device, one or more lights 181-183 are controlled according to one or more lighting instructions defined by the light script 161.

Fig. 3 shows schematically a flowchart illustrating another embodiment of a method 300 of providing light scripts for media fragments. In an obtaining step 305, the first media fragment 111 and the plurality of second media fragments 112 may be obtained, for instance, from a server 102 or from an internal 123 or external memory 101 or from a user via a user interface device, e.g. a smart phone. Other combinations of obtaining the first media fragment 111 and the plurality of second media fragments 112 are also possible.

The method 300 may further comprise the step of analyzing 310 the at least one audio characteristic of the first media fragment 111 and the step of analyzing 320 the audio characteristics of the plurality of second media fragments 112, wherein the step of analyzing 310 may be divided into the clustering step 310a and the analyzing step 310b. In the clustering 310a step, a plurality of audio characteristics of the first media fragment 111 is clustered into a plurality of clusters based on a similarity criterion. The first media fragment 111, e.g. a song, may have distinct sets or clusters of properties. For example, a cluster, may describe how energetic and dynamic the song is, while the other set of properties may describe the valence of the song. Valence of a song may be defined, as a measure from 0.0 to 1.0 describing the musical positiveness conveyed by a song. Songs with high valence sound more positive (e.g. happy, cheerful, euphoric), while songs with low valence sound more negative (e.g. sad, depressed, angry). Valence of the song may be estimated in multiple ways, the most common way is using classification approach where a set of songs is classified by valence for example by music experts.

Different algorithms for clustering, e.g. from machine learning, known in the art may be used. Given a set of data points $(x_1, x_2, \dots, x_n)$, a clustering algorithm may classify each data point into a specific group. In theory, data points that are in the same group should have similar properties and/or features, while data points in different groups should have dissimilar properties and/or features. The step of clustering 310a may be performed by using clustering algorithms using machine learning, and wherein the clustering uses at least one of: K-mean clustering, mean-shift clustering, Density-Based Spatial Clustering of Applications with Noise (DBSCAN), Expectation–Maximization (EM)

Clustering using Gaussian Mixture Models (GMM). These algorithms are well-known in the art and therefore are not further discussed here.

The analyzing step 310b comprises analyzing the clusters. The clusters may be analyzed based on an average value of the clusters, or on extreme values (minimum or maximum values at the cluster edges), or on the center of clusters; wherein in some clustering algorithms, the average value of the cluster is the center of the cluster.

The method 300 may further comprise the step of comparing 330 the analyzed plurality of clusters with the analyzed audio characteristics of the plurality of second media fragments 112, wherein for each set of properties, i.e. for each cluster, different second media fragments can be identified. For example, a set of common features and relations may be identified, then these features may be used to, measure, e.g. the valence or energy, of all other songs. The step of comparing 330 may be performed by using similarity learning as discussed before or any other algorithm known in the art.

The method 300 may further comprise the step of selecting 340 a subset of the plurality of second media fragments 112 based on the comparison 330. The subset of second media fragments may be selected if a difference between the analyzed clusters of the first media fragment 111 and an analyzed at least one audio characteristic of the second media fragment exceeds a threshold. In an embodiment, weights may be assigned to the plurality of clusters and the selection 340 of the subset may be based on these weights. These weights may be defined by a user via a user interface device, e.g. smart phone, or be set by a learning algorithm based on historic data.

The method 300 may further comprise the step of providing 350, wherein the light script 161 is provided by generating the light script 161 based on light scripts associated with the subset of the plurality of second media fragments 112. For example, if during analysis 310, 320, two media fragments may be identified, wherein one has a similar energy and another has similar valence properties, the light script may be generated such that the overall palette and set of light effects are taken from the “valence” light script while the speed of changes and brightness of effects are taken from the “energy” light script. This embodiment is dependent on a format wherein scripts are stored, e.g. if scripts only describe set of commands to be sent to lights this approach may not work, while if scripts include description of light effects and colour palettes this embodiment may be realised.

Alternatively, the first media fragment 111 may be splitted based on the audio analysis, for example by splitting a song based on detected song sections, and each part may be separately compared with the plurality of second media fragments 112 that have an associated light

script. The light script for the first media fragment 111 may then be generated from the light scripts selected based on the first media fragment 111 parts comparison with other media fragments 112 that have a script. Additional processing may be required for such a “stitched” light script, for example by normalizing the colour palette or adjusting brightness.

5 The method 300 may further comprise the step of associating 360 the generated light script 161 with the first media fragment 111.

Fig. 4 shows schematically a flowchart illustrating another embodiment of a method 400 of providing light scripts for media fragments. In an obtaining step 405, the first media fragment 111, e.g. a song, and the plurality of second media fragments 112, e.g. a
10 collection of songs, may be obtained, for instance, from a server 102 or from an internal 123 or external memory 101 or from a user via a user interface device, e.g. a smart phone. Other combinations of obtaining the first media fragment 111 and the plurality of second media fragments 112 are also possible.

The method 400 may further comprise the step of analyzing 410 at least one
15 characteristic of the first media fragment 111 and the step of analyzing 420 audio characteristics of the plurality of second media fragments 112, wherein the step of analyzing 420 is further divided into identifying 420a substantially similar audio characteristics of the plurality of second media fragments 112, and clustering 420b the plurality of second media fragments 112 into a plurality of clusters based on the identification 420a.

20 The step of identifying 420a comprises identifying media fragments similar characteristics in the plurality of second media fragments 112, for example a group of songs with a high pitch or a group of songs played in minor or major key. In the step of clustering 420b, the identified media fragments based on the similar characteristics are clustered into a plurality of clusters. The step of clustering 310a may be performed by using clustering
25 algorithms, which are known in the art.

The method 400 may further comprise the step of providing 450 a plurality of master light scripts by generating the plurality of master light scripts based on the plurality of clusters. A master light script may be composed of light scripts associated with a plurality of media fragments representing a particular set of properties, e.g. energy or valence. A benefit
30 of applying the master light script is that, from a user’s perspective, the light effects for similar songs (e.g. of the same artist) are consistent. Another benefit is that the user is not confronted with a lot of options for selecting a light script for the first audio fragment 111. For example, if the first audio fragment 111 is played in minor or major key, the master light script representing the minor/major is used for the comparison with the first media fragment

111. The master light scripts are associated to respective plurality of clusters. The method 400 may further comprise the step of storing 470 the plurality of master light scripts.

The method 400 may further comprise the step of comparing 430 the analyzed at least one audio characteristic of the first media fragment 111 with the plurality of master light scripts. The step of comparing 430 may comprise comparing 430 the analyzed at least one audio characteristic of the first media fragment 111 with the plurality of clusters each associated with the of master light scripts. The method 400 may further comprise the step of selecting 440 and associating 460, wherein at least one master light script from the plurality of master light scripts is selected based on the comparison, and the selected at least one master light script is associated with the first media fragment 111.

Fig. 5 shows schematically a controller 520 for providing light scripts for media fragments. The controller 520 comprising: an input 521 configured to obtain a first media fragment 111 and a plurality of second media fragments 112, wherein the first media fragment 111 comprises a first audio and/or a first video fragment, and wherein the plurality of second media fragments 112 comprises second audio fragments and/or second video fragments, and wherein each of the plurality of second media fragments 112 is associated with respective light scripts.

The controller 520 further comprises a processor comprising: an analyzing unit 510 configured to analyze at least one audio characteristic of the first media fragment 111, an analyzing unit 520 configured to analyze audio characteristics of the plurality of second media fragments 112, a comparing unit 530 configured to compare the analyzed at least one audio characteristic of the first media fragment with the analyzed audio characteristics of the plurality of second media fragments, a selecting unit 540 configured to select a second media fragment based on the comparison, a providing unit 550 configured to provide a light script (161) based on the selected second media fragment, an associating unit 560 to associate the light script (161) with the first media fragment, such that when the first media fragment is being rendered by a media fragment rendering device, one or more lights are controlled according to one or more lighting instructions defined by the light script 161. Furthermore, the associated information, e.g. the light script 161, is output by the output 524 to, e.g. light sources 180-182.

It will be understood that the processor 122 or processing system or circuitry referred to herein may in practice be provided by a single chip or integrated circuit or plural chips or integrated circuits, optionally provided as a chipset, an application-specific

integrated circuit (ASIC), field-programmable gate array (FPGA), digital signal processor (DSP), graphics processing units (GPUs), etc.

The controller 520 may be implemented in a unit, such as wall panel, desktop computer terminal, in the bridge, or even a portable terminal such as a laptop, tablet or smartphone. Further, the controller 520 may be implemented remotely (e.g. on a server); and the controller 520 may be implemented in a single unit or in the form of distributed functionality distributed amongst multiple separate units (e.g. a distributed server comprising multiple server units at one or more geographical sites, or a distributed control function distributed amongst the light sources 180-182. Furthermore, the controller 520 may be implemented in the form of software stored on a memory (comprising one or more memory devices) and arranged for execution on a processor (comprising one or more processing units), or the controller 520 may be implemented in the form of dedicated hardware circuitry, or configurable or reconfigurable circuitry such as a PGA or FPGA, or any combination of these.

The method 200, 300, 400 may be executed by computer program code of a computer program product when the computer program product is run on a processing unit of a computing device, such as the processor 122 of the system 100.

It should be noted that the above-mentioned embodiments illustrate rather than limit the invention, and that those skilled in the art will be able to design many alternative embodiments without departing from the scope of the appended claims.

In the claims, any reference signs placed between parentheses shall not be construed as limiting the claim. Use of the verb "comprise" and its conjugations does not exclude the presence of elements or steps other than those stated in a claim. The article "a" or "an" preceding an element does not exclude the presence of a plurality of such elements. The invention may be implemented by means of hardware comprising several distinct elements, and by means of a suitably programmed computer or processing unit. In the device claim enumerating several means, several of these means may be embodied by one and the same item of hardware. The mere fact that certain measures are recited in mutually different dependent claims does not indicate that a combination of these measures cannot be used to advantage.

Aspects of the invention may be implemented in a computer program product, which may be a collection of computer program instructions stored on a computer readable storage device which may be executed by a computer. The instructions of the present invention may be in any interpretable or executable code mechanism, including but not

limited to scripts, interpretable programs, dynamic link libraries (DLLs) or Java classes. The instructions can be provided as complete executable programs, partial executable programs, as modifications to existing programs (e.g. updates) or extensions for existing programs (e.g. plugins). Moreover, parts of the processing of the present invention may be distributed over multiple computers or processors or even the 'cloud'.

Storage media suitable for storing computer program instructions include all forms of nonvolatile memory, including but not limited to EPROM, EEPROM and flash memory devices, magnetic disks such as the internal and external hard disk drives, removable disks and CD-ROM disks. The computer program product may be distributed on such a storage medium, or may be offered for download through HTTP, FTP, email or through a server connected to a network such as the Internet.

CLAIMS:

1. A method of providing light scripts for media fragments, the method comprising:

- obtaining (205,305,405) a first media fragment (111) and a plurality of second media fragments (112),

5 wherein the first media fragment (111) comprises a first audio fragment; and wherein the plurality of second media fragments (112) comprises second audio fragments, wherein the second audio fragments are comprised in video fragments; and wherein each of the plurality of second media fragments (112) is associated with respective light scripts,

10 - analyzing (210,310,410) at least one audio characteristic of the first media fragment (111),

- analyzing (220,320,420) audio characteristics of the plurality of second media fragments (112),

15 - comparing (230,330,430) the analyzed at least one audio characteristic of the first media fragment (111) with the analyzed audio characteristics of the plurality of second media fragments (112),

- selecting (240,340,440) one or more second media fragments from the plurality of second media fragments based on the comparison,

20 - providing (250,350,450) a light script (161) based on light scripts associated with the selected one or more second media fragments,

- associating (260,360,460) the light script (161) with the first media fragment (111), such that when the first media fragment (111) is being rendered by a media fragment rendering device, one or more lights are controlled according to one or more lighting instructions defined by the light script (161).

25

2. The method of claim 1, wherein the at least one audio characteristic of the first media fragment (111) and the audio characteristics of the plurality of second media fragments (112) comprises one or more of: beat, timbre, pitch, intensity, rhythm, major and minor key.

3. The method of any preceding claim, wherein the method further comprises:
- clustering (310a) a plurality of audio characteristics of the first media fragment (111) into a plurality of clusters based on a similarity criterion,
 - 5 - analyzing (310b) the plurality of clusters,
 - comparing (330) the analyzed plurality of clusters with the analyzed audio characteristics of the plurality of second media fragments,
 - selecting (340) a subset of the plurality of second media fragments (112) based on the comparison,
 - 10 - wherein the step of providing (350) the light script (161) further comprises
 - generating the light script (161) based on light scripts associated with the subset of the plurality of second media fragments (112),
 - associating (360) the generated light script with the first media fragment (111).

15 4. The method of claim 3, wherein weights have been assigned to the plurality of clusters and wherein the selection (340) of the subset is based on these weights.

5. The method according to claim 1 or 2, wherein the method further comprises:
- identifying (420a) audio characteristics of the plurality of second media
 - 20 fragments (112) based on a similarity criterion,
 - clustering (420b) the plurality of second media fragments (112) into a plurality of clusters based on the identification (420a),
 - providing (450) a plurality of master light scripts by generating the plurality of master light scripts based on the plurality of clusters, and
 - 25 - storing (470) the plurality of master light scripts.

6. The method according to claim 5, wherein the method further comprises:
- comparing (430) the analyzed at least one audio characteristic of the first media fragment (111) with the plurality of clusters each associated with the plurality of
 - 30 master light scripts,
 - selecting (440) at least one master light script from the plurality of master light scripts based on the comparison, and
 - associating (460) the selected at least one master light script with the first media fragment (111).

7. The method of any preceding claim, wherein the step of comparing (230,330,430) is performed by using similarity learning, and wherein the similarity learning uses at least one of: regression similarity learning, classification similarity learning, ranking
5 similarity learning and locality sensitive hashing.

8. The method of any preceding claim, wherein the second media fragment is selected if a difference between the analyzed at least one audio characteristic of the first media fragment (111) and an analyzed at least one audio characteristic of the selected second
10 media fragment exceeds a threshold.

9. The method of any preceding claim, wherein the method further comprises: modifying the light script (161) based on a difference between the analyzed at least one audio characteristic of the first media fragment (111) and the analyzed at least one audio
15 characteristic of the selected second media fragment.

10. The method of any preceding claim, wherein the comparison of the analyzed at least one audio characteristic of the first media fragment (111) with the analyzed audio characteristics of the plurality of second media fragments (112) is based on a user-based
20 subset of the plurality of second media fragments, wherein the user-based subset is based on a user playlist.

11. The method of claim 1, wherein, when the one or more of second media fragments have been selected based on the comparison (230, 330, 430), the step of selecting
25 (240, 340, 440) the second media fragment further comprises: selecting a second media fragment from the group based on one or more of: a user preference, a lighting setup of a user, a psychological or physiological state of a user, a previous light script used by a user, the time of day, an activity of a user.

30 12. The method of any preceding claim, wherein the first media fragment (111) comprises an audio fragment and wherein the plurality of second media fragments (112) comprises audio fragments.

13. The method of claim 12, wherein the first media fragment (111) comprises an audio fragment comprised in a video fragment.

14. A computer-readable storage medium comprising instructions, which when
5 executed by a computer, cause the computer to execute the steps of the method of claim 1-13.

15. A controller (520) for providing light scripts for media fragments, the controller (520) comprising:

- an input (521) configured to obtain a first media fragment (111) and a plurality
10 of second media fragments (112),
wherein the first media fragment (111) comprises a first audio fragment, and wherein the plurality of second media fragments (112) comprises second audio fragments, wherein the second audio fragments are comprised in video fragments; and
wherein each of the plurality of second media fragments (112) is associated with respective
15 light scripts,
- a processor (522) configured to execute steps of the method according to claim 1,
- an output (524) configured to output information about association of the light script (161) with the first media fragment.

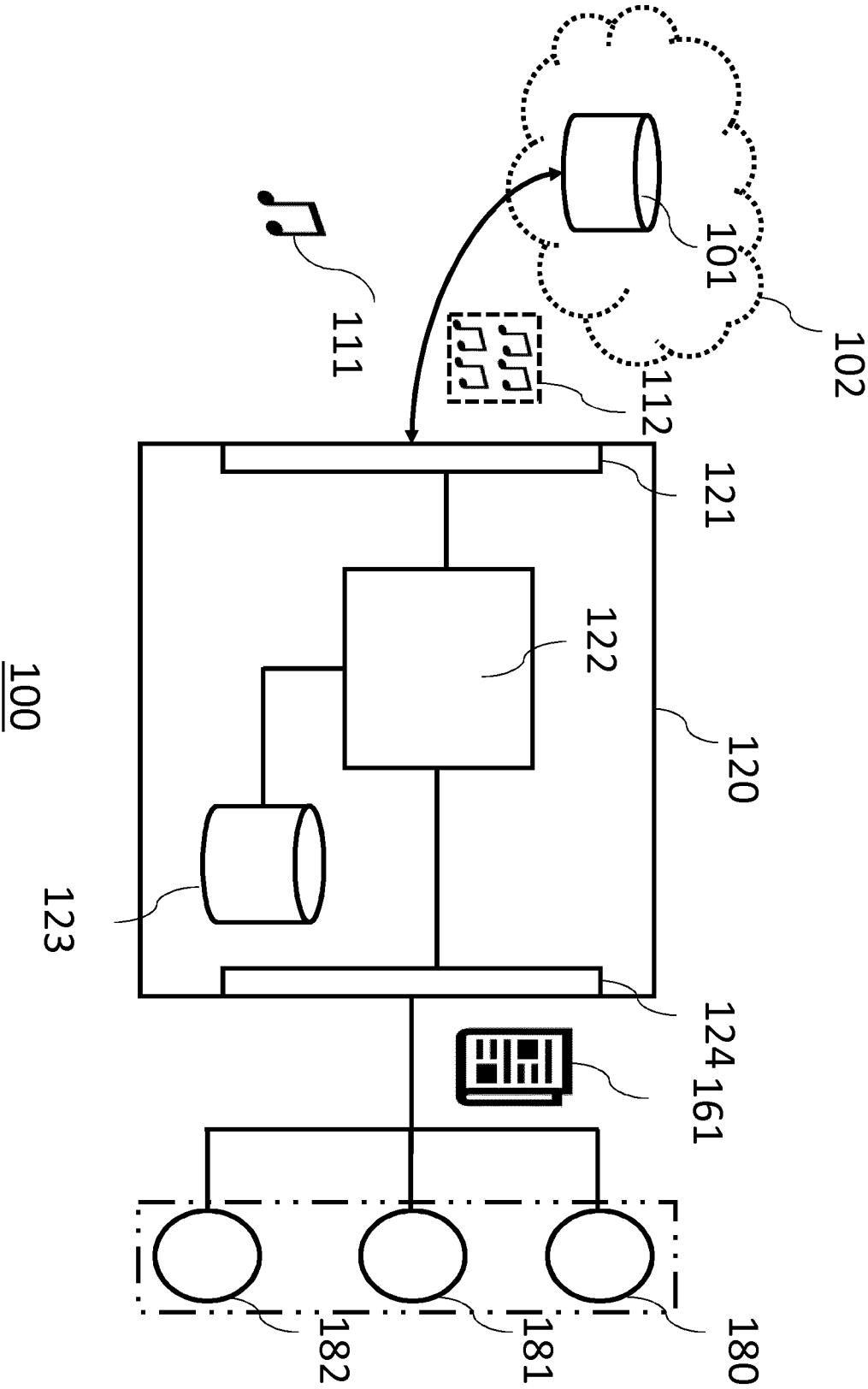


Fig. 1

2/5

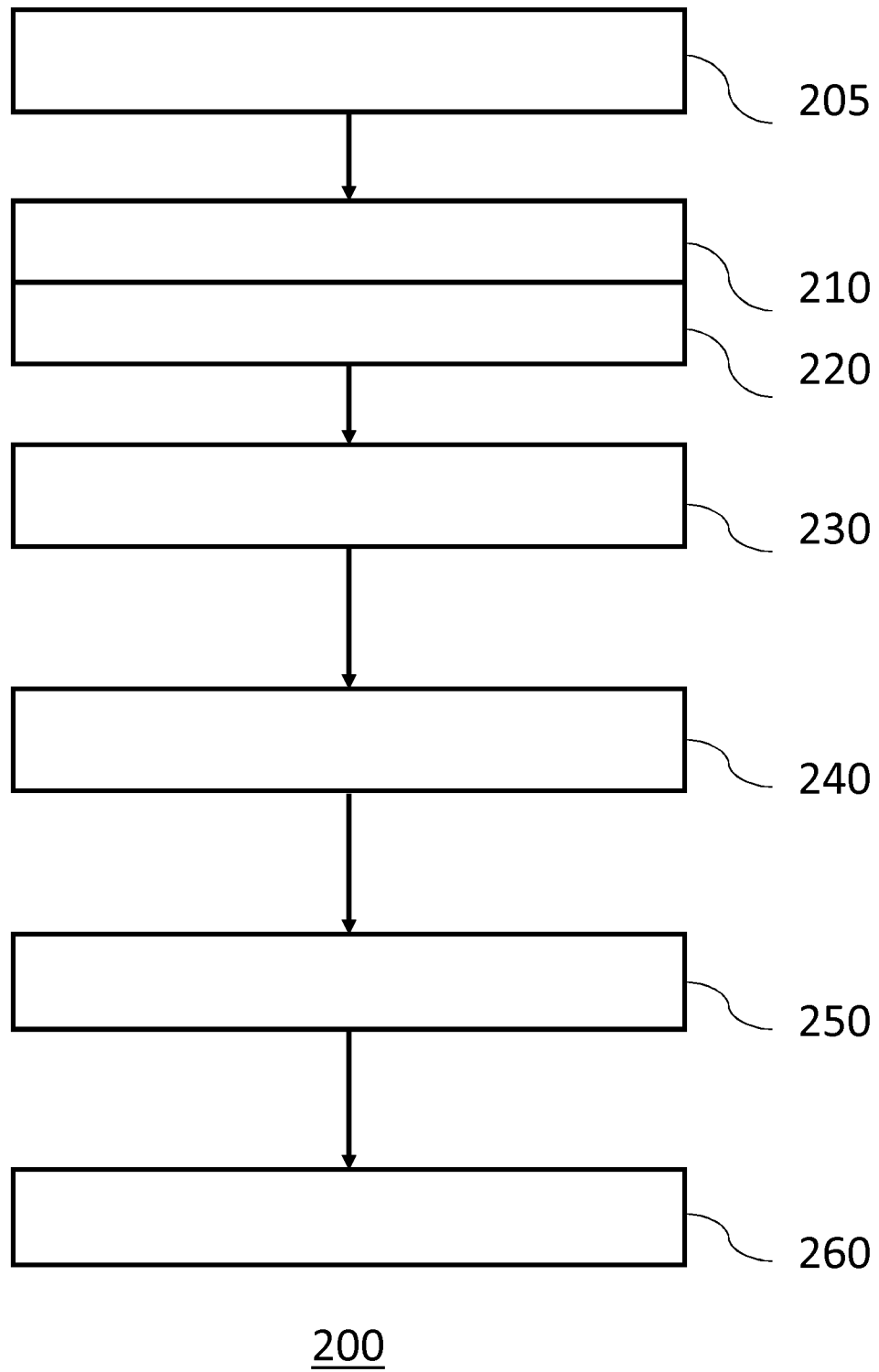


Fig. 2

3/5

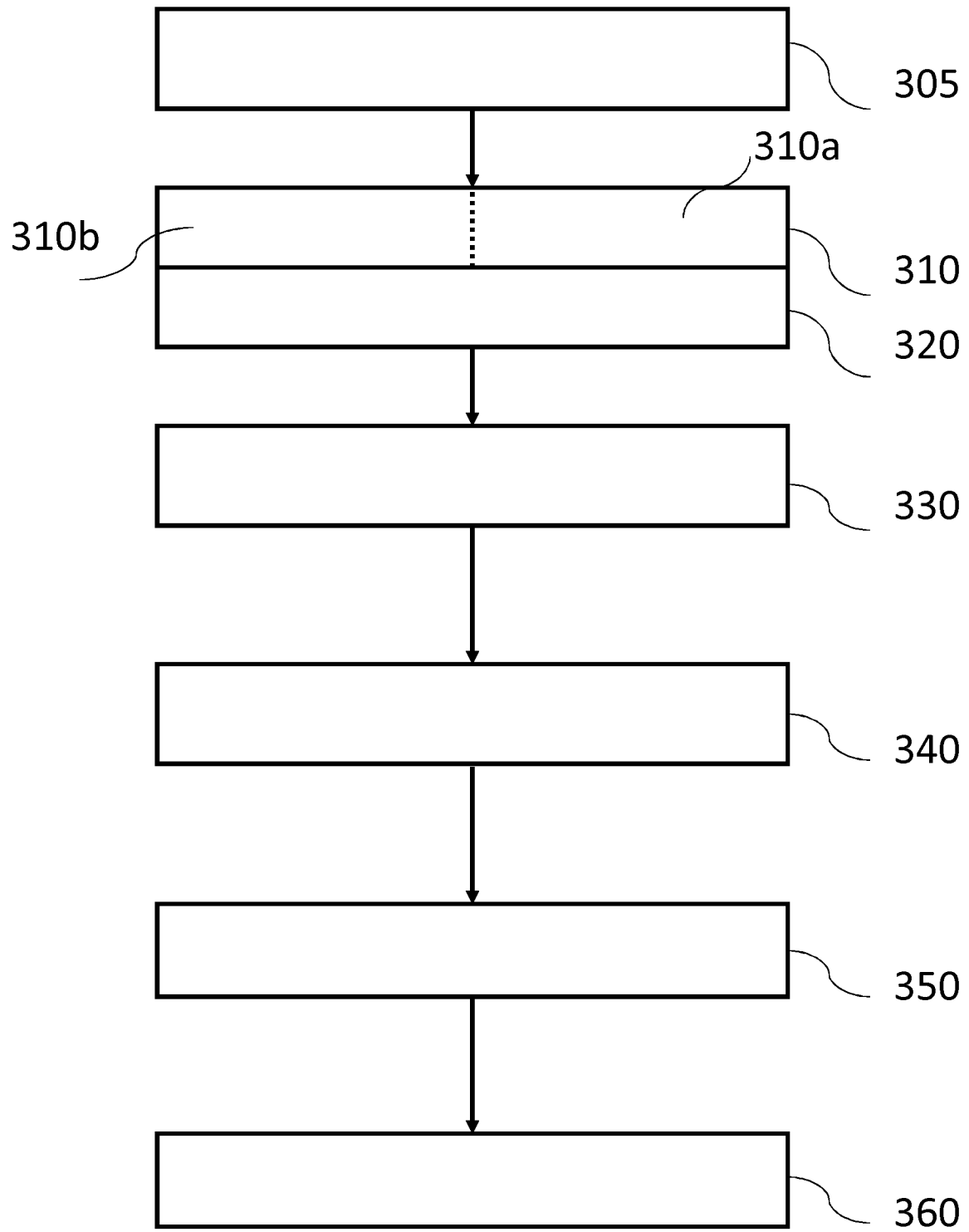
300

Fig. 3

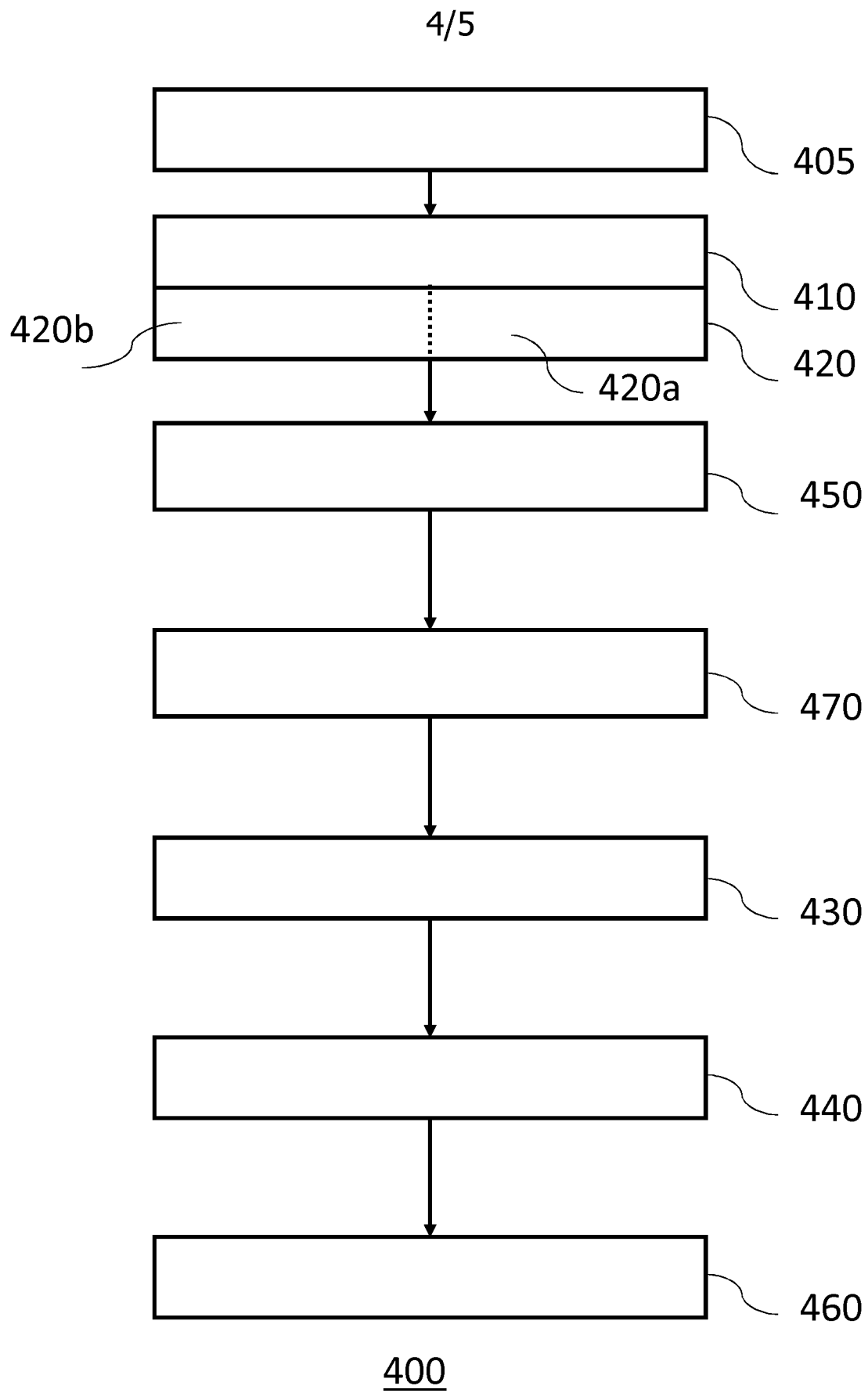


Fig. 4

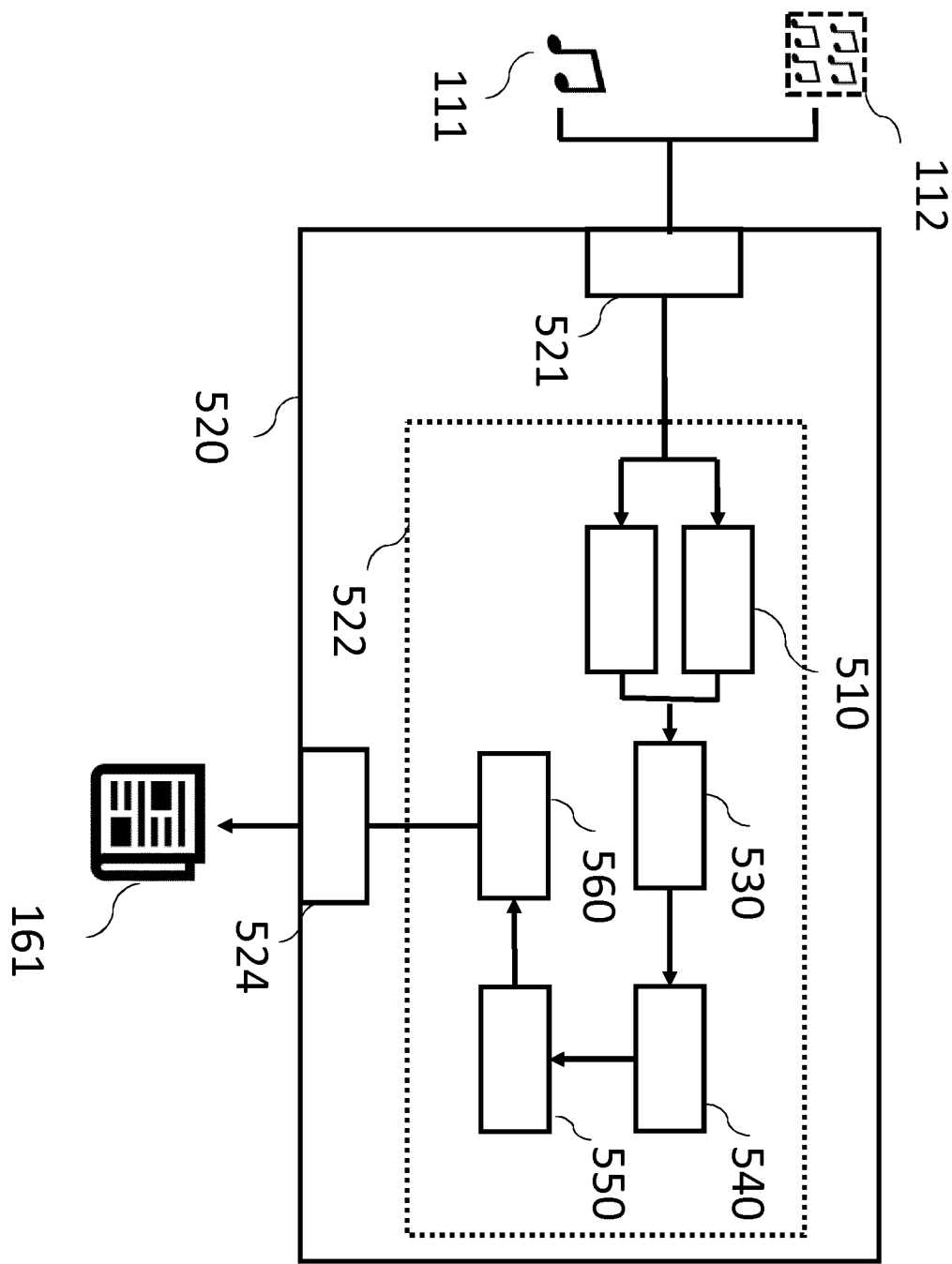


Fig. 5

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2019/081814

A. CLASSIFICATION OF SUBJECT MATTER

INV. H05B37/02 A63J17/00 H04N21/41 H04N21/439 H04N21/44
G06F16/683 G06F16/60

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)

H05B H04N G06F A63J

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 2018/061438 A1 (LOVE JUSTIN THOMAS [CA] ET AL) 1 March 2018 (2018-03-01)	1,2,7,8, 10,12, 14,15
Y	paragraphs [0015] - [0036]; claims 1-8 paragraphs [0049] - [0096]; figures 1-8 -----	3-6
X	US 2014/072272 A1 (DE WAELE STIJN [US] ET AL) 13 March 2014 (2014-03-13) paragraphs [0051] - [0064]; claim 20; figures 2-3 -----	1,2,8,9, 12-15
X	WO 2017/162469 A1 (PHILIPS LIGHTING HOLDING BV [NL]) 28 September 2017 (2017-09-28) page 5, line 20 - page 14, line 20; claims 1-13; figures 1-6 ----- -/-	1,8, 10-12, 14,15



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

17 December 2019

Date of mailing of the international search report

07/01/2020

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Ferla, Monica

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2019/081814

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2012/078824 A1 (BONET ANTONIO TRIAS [ES] ET AL) 29 March 2012 (2012-03-29) the whole document -----	3,4
Y	US 2012/254363 A1 (MARTIN KEITH DANA [US] ET AL) 4 October 2012 (2012-10-04) the whole document -----	5,6

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2019/081814

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2018061438	A1	01-03-2018	NONE
US 2014072272	A1	13-03-2014	CN 102245274 A 16-11-2011
		EP 2376207 A1 19-10-2011	
		ES 2548139 T3 14-10-2015	
		JP 5475799 B2 16-04-2014	
		JP 2012511792 A 24-05-2012	
		KR 20110110147 A 06-10-2011	
		US 2011245941 A1 06-10-2011	
		US 2014072272 A1 13-03-2014	
		WO 2010067268 A1 17-06-2010	
WO 2017162469	A1	28-09-2017	EP 3434073 A1 30-01-2019
		US 2019104593 A1 04-04-2019	
		WO 2017162469 A1 28-09-2017	
US 2012078824	A1	29-03-2012	NONE
US 2012254363	A1	04-10-2012	NONE