

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
21 October 2010 (21.10.2010)

PCT

(10) International Publication Number
WO 2010/119181 A1

(51) International Patent Classification:
G11B 27/031 (2006.01) **G06F 17/30** (2006.01)

(21) International Application Number:
PCT/FI2010/050309

(22) International Filing Date:
16 April 2010 (16.04.2010)

(25) Filing Language: Finnish

(26) Publication Language: English

(30) Priority Data:
20095415 16 April 2009 (16.04.2009) FI

(71) Applicant (for all designated States except US): **VALTION TEKNIILLINEN TUTKIMUSKESKUS** [FI/FI]; Vuorimiehentie 3, FI-02044 VTT (FI).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **JÄRVINEN, Sari** [FI/FI]; Valtion teknillinen tutkimuskeskus, PL 1100, FI-90571 Oulu (FI). **OJUTKANGAS, Onni** [FI/FI]; Valtion teknillinen tutkimuskeskus, PL 1100, FI-90571 Oulu (FI). **PELTOLA, Johannes** [FI/FI]; Valtion teknillinen tutkimuskeskus, PL 1100, FI-90571 Oulu (FI).

(74) Agent: **BERGGREN OY AB**; P.O. Box 16 (Antinkatu 3 C), FI-00101 Helsinki (FI).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, 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, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (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, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

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

(54) Title: VIDEO EDITING SYSTEM

(57) Abstract: The invention relates to a video editing system. According to the invention, the system comprises means for defining features of videos, means for automatically searching in a group of videos those videos or video sections i.e. segments which include the above defined features and means for creating a video compilation of videos and/or segments including the above defined features. In an advantageous embodiment of the invention, features to be defined include the geographical region where the video has been shot and the time and/or time period when the video has been shot and recorded.



WO 2010/119181 A1

Video editing system

The invention relates to a video editing system according to the preamble of claim 1.

- 5 The invention also relates to a method for editing videos according to the preamble of claim 16.

Video refers to live picture material consisting of successive pictures quickly played back, the duration of which can vary, recorded by means of video cameras, particularly digital video cameras, mobile phones or equivalent devices. Video
10 recorded with a mobile phone can be referred to as mobile video. The subjects of videos, particularly of personal videos, are, inter alia, family events, such as birthdays, and holiday trips, but videos are recorded as well of various everyday situations. The editing of videos refers to the compiling and processing of one or more videos such that its picture material e.g. is rearranged, sections are deleted
15 of it, or the picture material of several videos, i.e. its sections or segments, is combined to a new final video to be stored and shown.

Managing personal video files and contents is more and more difficult as the volume of the contents increases at high speed. Furthermore, ordinary users have neither time nor often skills or interest in editing their videos afterwards which
20 would improve the watching of the videos considerably later.

The libraries of digital mobile videos have expanded and watching them is considered wearisome and unexciting. Traditionally, the problem has been solved by editing a presentation of the videos by cutting and pasting them together. However, this requires special tools, time and expertise which an ordinary user
25 does not have. Furthermore, in extensive video libraries, which include videos shot by several users, the combining of videos into logical units involves additional challenges for the editing process.

A problem of the development of multimedia services in mobile devices is often the small size of their display and the limited data transfer capacity of the mobile
30 network. Services, which offer the user a possibility to search or browse video contents with a mobile phone, often show the result by means of keyframe

pictures as a list or equivalent format which the user scrolls on the phone up and down. This is awkward and time-consuming.

A problem of the compiling i.e. editing of videos is often defining the compilation basis.

- 5 An object of the invention is to eliminate problems related to known video editing systems and equivalent methods. An object of the invention is also to provide a novel improved video editing system the principles of which are simple, efficient and cost-effective to implement.

- 10 A video editing system according to the invention is characterised by what is presented in claim 1. A method for editing videos according to the invention is characterised by what is presented in claim 16. The dependent claims present some advantageous embodiments of the invention.

The video editing system according to the invention comprises:

- means for defining features of videos;
- 15 - means for automatically searching in a group of videos those videos or video sections i.e. segments which include the above defined features; and
- means for creating a video compilation of videos and/or segments including the above defined features.

- 20 In an advantageous embodiment of the invention, features to be defined of the videos include the geographical region where the video has been shot and the time and/or time period when the video has been shot and recorded.

- In an advantageous embodiment of the invention, the video editing system comprises means for selecting suitable sections, segments, of the videos, which means include a video feature analysing unit which analyses sound, picture and/or
- 25 motion features calculated of the video (i.e. content analysis metadata) based on which the segments are selected.

In an advantageous embodiment of the invention, the video editing system comprises means for creating a new video i.e. video compilation of the videos or their segments most advantageously automatically.

In an advantageous embodiment of the invention, the selection of video segments is arranged to be implemented by calculating the above features and by giving each segment points based on this and by selecting those segments in the video compilation which comply with the predetermined point limits.

- 5 In an advantageous embodiment of the invention, the video editing system comprises means for creating video scenes of the selected video segments by clustering the video segments on the basis of the time stamp and/or geographical data.

- 10 In an advantageous embodiment of the invention, the video editing system comprises means for creating a new video or video presentation i.e. video compilation of the videos or their segments based on the contents of the user and those recommended to the user i.e. the selected segments such that the location, time or content analysis metadata constitute continua between the video segments.

- 15 In an advantageous embodiment of the invention, the video editing system comprises means according to independent claim 1 for searching and retrieving videos from a multimedia database or equivalent video database.

- 20 In an advantageous embodiment of the invention, the video editing system comprises means for searching, by means of specific search criteria, desired videos in a multimedia database, of which videos a video compilation is to be created.

In an advantageous embodiment of the invention, the video editing system comprises means for combining classified video segments into a new video i.e. video compilation.

- 25 In an advantageous embodiment of the invention, the video editing system comprises a video database, such as a multivideo database, which is shared by several users and accessible through a data network, such as the Internet.

In an advantageous embodiment of the invention, the video editing system comprises means for updating the video database with new videos.

- 30 In an advantageous embodiment of the invention, the video editing system comprises means for automatically performing video editing according to predetermined features.

In an advantageous embodiment of the invention, the video editing system is arranged on a data network, advantageously on the Internet, where it is accessible for the users.

5 In an advantageous embodiment of the invention, the video editing system comprises a music database from which the user can select music to accompany the video compilation.

The method according to the invention comprises the following steps:

- defining features of videos;
- automatically searching in a group of videos those videos or video sections i.e.
10 segments which include the above defined features; and
- creating a video compilation of videos and/or segments including the above defined features.

15 In advantageous embodiments of the method according to the invention, instead of the means of system claims 2–15 are defined equivalent action steps which are implemented with the means in question and, by performing which, the method steps are realised. This principle becomes evident, inter alia, by comparing independent claims 1 and 16.

Next, the invention, its various embodiments and advantages will be described in more detail.

20 **Creating video compilation based on semantic metadata**

By means of various content analysis methods and collecting context data, it is possible to obtain low-level data on the features of video contents. It is e.g. possible to calculate momentums of the video or to save the GPS coordinates of the creation location of the video as metadata. Of this lower-level metadata, it is
25 possible to develop higher-level semantic metadata. By means of GPS coordinates, for instance, it is possible to define the city or, by utilising the coordinates together with the shooting time, to obtain data on prevailing weather from external services. Then, it is possible to automatically create a compilation of separate video clips, in which the compilation basis is semantic metadata. For
30 example, videos from a specific location shot in sunny weather are selected in the compilation.

A problem of the compiling i.e. editing of the videos is often determining the compilation basis. Metadata produced by low-level content analysis methods does not necessarily offer a sensible compilation basis for a video. It is possible to upgrade the low-level metadata into higher-level metadata which is better suited for the compilation basis of videos.

The system searches lower-level metadata related to videos in the multimedia database and upgrades it into higher-level (semantic) metadata. The system creates a video compilation based on the higher-level metadata. The video compilation is created by segmenting the videos into smaller sections (*video segmentation*), by giving points to the video segments and by selecting the best segments based on the points into the compilation. Finally, scenes are created of the video segments based on the context data of the videos. The scenes are combined into one video presentation.

Automatic compilation of personal videos by means of context data

By means of the automatic video editing system, it is possible to create a service by means of which the user can define specific features for the video compilation. The user can supply the service with the desired video length and define the geographical region and time period for videos to be selected into the compilation. Thus, the user is automatically provided with a personal video compilation in which the system selects the most interesting segments of the video contents created by the user. The user can e.g. define that "I want a video compilation of videos shot at the summer house in July".

Managing personal video files and contents is more and more difficult as the volume of the contents increases at high speed. In addition to digital video cameras, videos are shot with mobile phones, which enables context data (mostly the creation time and location of the video as GPS data). However, ordinary users have neither time nor often skills or interest in editing their videos afterwards which would improve the watching of the videos considerably later. By means of the service created, the users are offered a possibility to create easily and quickly based on context data video compilations which contain essential segments of the original videos. The user can e.g. create a compilation of videos shot in the summer at the summer house by means of the service.

The user is supplied a service by means of which it is possible to define the context (the creation region and time period of the video), based on which the video compilation of the videos shot by the user or user group is created. In the automatic video editing system is entered a set of videos selected based on context for analysis, segmentation and compilation. The system creates a video compilation defined by the user and supplies it to the user for presentation.

Semiautomatic video editor

By means of the automatic video compilation system, it is possible to provide a service which facilitates the editing of videos. The system segments and classifies videos automatically recommending interesting video contents for the user to be added into the compilation desired by the user. The user selects the desired segments and the system creates a compilation of them inserting suitable effects between the scenes. By means of the service, the editing of videos is easy also in mobile terminals.

The libraries of digital mobile videos have expanded and watching them is considered wearisome and unexciting. Traditionally, the problem has been solved by editing a presentation of the videos by cutting and pasting video clips together. However, this requires special tools, time and expertise which an ordinary user does not have. The invention automatises the hardest stages of video editing: cutting and pasting. In the semiautomatic video editor, the user only has to select the desired video segments and the system takes care of editing.

The system searches videos in the multimedia database and performs their segmentation and classification. The system gives points to the segments and recommends the user via the service the best segments to be included in the compilation. The user selects the desired segments and the system creates a video compilation of the selected segments. The system utilises the context data (location and time) of videos in creating the video compilation. The video compilation can be shown by means of the service in the user's terminal.

"On-line experience video channel"

Many social media services offer their users a possibility to collect their own contents and to recommend suitable multimedia contents for other users. By means of the automatic video compilation, it is possible to create each user his/her

own "on-line experience video channel" of video clips the contents of which are automatically updated according to new recommendation or contents. Thus, the user has no need to separately watch each video clip, but he/she is able to receive a general view of the videos by following the contents of the "on-line" channel.

- 5 The libraries of digital mobile videos have expanded and watching them is considered wearisome and unexciting. Traditionally, the problem has been solved by editing a presentation of the videos by cutting and pasting them together. However, this requires special tools, time and expertise which an ordinary user does not have. Furthermore, in extensive video libraries, which include videos
10 recorded by several users, the combining of videos into logical units involves additional challenges for the editing process. Grouping the videos without the context data of the video is extremely challenging.

The main principle is that the system segments the videos into smaller sections, selects the best of them and compiles them into a new video presentation. The
15 segmenting of videos (*video segmentation*) is implemented by analysing the sound, picture and motion features calculated of the video (*feature extraction*). The selection of the video segments is implemented by giving them points by means of supplied weighting coefficients and feature values. Video segments are compiled into scenes (*video scene*) by clustering the segments based on the video time
20 stamp (*timestamp*) and geographical data (*geotagging*). Finally, the scenes are combined into one video presentation (*video rendering*). The combining is performed based on the contents of the user and those recommended to the user such that the location, time or content analysis metadata constitute continua between the segments. At specific intervals, the programme can be changed,
25 whereby it is possible to shift to video contents related to a totally different location and time. The channel compiles the contents into a continuous video stream which can be shown on an IPTV set top box, a digital picture frame or a computer display.

Using geographical data in clustering the video segments enables the combining
30 of videos shot in the same context (location/time) into one unit (a scene in the video compilation). Without the geographical data, combining videos received from various sources based on context is difficult or impossible to automatise.

Mobile video compilation

A problem of the development of multimedia services in mobile devices is often the small size of their display. When the user searches videos with a mobile phone, the result is often shown by means of keyframe pictures as a list or equivalent format which the user scrolls on the phone up and down. By means of the automatic video compilation, it is possible to create a compilation of the videos obtained as a search result for the user in which the most interesting segments of various videos are displayed. Thus, the user is able to have a general picture of the contents of the videos without scrolling up and down, and it is also possible to offer quicker operating services as the volume of transferable data decreases.

A problem of the development of multimedia services in mobile devices is often the small size of their display and the limited data transfer capacity of the mobile network. Services, which offer the user a possibility to search or browse video contents with a mobile phone, often show the result by means of keyframe pictures as a list or equivalent format which the user scrolls on the phone up and down. By means of the automatic video compilation, it is possible to create a compilation of the videos obtained as a search result for the user in which the most interesting segments of various videos are displayed. Thus, the user is able to have a general picture of the contents of the videos without scrolling up and down, and it is also possible to offer quicker operating services as the volume of transferable data possibly decreases (depends naturally on video coding vs. the size and number of keyframe pictures).

The user searches or browses the desired contents in the video achieves with his/her mobile phone. These tasks performed by the user generate a search for video contents in the Internet-based system. The result of the search is entered in the service supplying the automatic video compilation which analyses the videos, segments them and creates a compilation of the selected segments. For this, parameters suitable for this service for the creation of the video compilation have been defined for the segmentation and video rendering.

Mobile service for compiling personal videos

Nowadays, mobile phone users shoot great numbers of pictures and videos. The editing of the videos is often omitted, because it is difficult on the phone and transferring to a computer requires extra work. By means of the automatic video

compilation system, it is still possible to create a web-based service to offer the user a possibility to create various compilations of his/her videos and to select suitable music to accompany them. The user is supplied e.g. five music pieces by means of a mobile interface and he/she can select some of his/her own videos of which the video compilation is automatically created by considering e.g. the music style, duration and tempo of the piece.

Nowadays, mobile phone users shoot great numbers of pictures and videos. The editing of the videos is often omitted, because it is difficult on the phone and transferring to a computer requires extra work. By means of the automatic video compilation system, it is still possible to create a web-based service to offer the user a possibility to create various compilations of his/her videos and to select suitable music to accompany them. The video compilation is automatically created by considering e.g. the music style, duration and tempo of the piece.

The user selects some of the videos created by him/herself for the compilation. The service offers the user some alternatives for background music one of which the user selects for the compilation. The automatic video compilation system analyses, segments and compiles videos considering the music style, piece duration etc. as parameters when creating the compilation.

Multimodal video compilation

In the video editing system according to the invention, by means of content analysis methods, it is possible to find e.g. the speech of a specific person in the video i.e. video contents or to identify video segments which include music. By combining this information with the context data (mostly location and time) of the user and the video contents, it is possible to create new kinds of video compilations for various services. It is e.g. possible to create a video compilation in which are selected video segments on a specific route taken including the speech of a specific person.

Earlier known automatically created video compilations lack factors combining video segments. This weakens the entertainment value and consistency of the video compilation. The invention enables compiling videos according to novel compilation bases. The compilation bases can be developed based on available content analysis methods and available context data.

By means of the content analysis methods, the system calculates features of the video contents which are used together with the context data of the user and the videos in selecting the videos. A video compilation is created of the selected video segments by combining them based on the context data of the videos.

5 Video compilation of user's context route

Various location-aware services have been studied and developed in recent years. The position of a mobile phone can be determined by means of the GPS technology or mobile phone networks very accurately. Yet, the position information has not been widely utilised in the management of multimedia contents. An embodiment of the invention utilises a context-aware service which saves the route taken by the user, offers the user based on this an automatically created video compilation from the route. For the video compilation are advantageously selected videos created on the route by all users and, in the creation of the compilation, it is also possible to consider e.g. the creation time of the videos (season, clock time etc.) Then, e.g. the user is supplied in winter a compilation which includes videos created by users in winter.

The invention offers the user a possibility to re-experience a previous trip in the form of the video compilation. Earlier, re-experiencing the trips has relied on the user's own recordings. By means of the invention, it is possible to enrich the video compilation with material shot by other users in the same context.

The video editing system according to the invention searches and selects in the multimedia database those videos which correspond the context route of the user (location and time data). The system creates a video compilation based on the user's videos and videos selected from the multimedia database. The video compilation is created by segmenting the videos into smaller sections (*video segmentation*), by giving points to the video segments according the predetermined criteria and by selecting the best segments based on the points i.e. having obtained most points into the compilation. Finally, scenes are created of the video segments based on the context information of the videos. The scenes are combined into one video presentation.

Automatic video compilation for location-aware mobile service

In an embodiment of the invention, a mobile service is implemented in which the user is offered a video compilation created based on the creation date and location of the videos. The user can watch with his/her mobile phone in his/her current position the video contents previously created by the users/user groups as one video compilation in which the most interesting video segments have been selected. The user is thus offered a compact view of videos, such as video clips, created in the location in question. By selecting the contents from a specific type of a video database (e.g. travel contents, historic contents, sports), it is possible to create new services by means of the same concept.

The small size of the display in mobile devices creates challenges for the development of multimedia services. New arrangements are required to display the great volume of multimedia contents in the mobile phone. By means of the automatic video compilation based on the invention, a location-aware service is created in which the user is able to view video contents produced by other users in his/her current position or in its vicinity as a video compilation in which the most interesting segments of the videos have been collected. By means of the arrangement, the usability of location-aware mobile video services can be improved.

A location-aware video compilation is automatically created by comparing the creation location of the videos to the present position of the user. The videos thus selected are analysed, segmented and compiled into one video presentation.

Sharing experiences by means of video compilation

Studies have shown that the greatest motive for recording pictures and videos is to later share experiences with others. Often, the sharing of experiences as videos suffers from the quality of the contents of the shot video when the videos are not edited. The automatic video compilation implemented with an advantageous embodiment of the invention (i.e. ultimately the video to be shown) compresses the videos shot by the user in a presentation in which are selected the most important segments of the video contents which are most essential in sharing the experience. Additional information for the video is obtained by utilising context data in compiling the video. Thus, the video segments are in the compilation in the chronological order or, if desired, e.g. classified according to location.

Studies have shown that the greatest motive for recording pictures and videos is to later share the experiences with others. This is emphasised when studying digital multimedia contents created with a mobile phone. However, the sharing of experiences as videos often suffers from the quality of the video contents shot when the videos are not edited. The automatic video compilation compresses the videos shot by the user or user group in a presentation in which are selected the most important segments of the video contents which are most essential in sharing the experience. Additional information for the video is obtained by utilising context data in compiling the video. Thus, the video segments are in the compilation in the chronological order or, if desired, e.g. classified according to location.

By means of the invention, the aim is at describing the experience of the user or user group as a video compilation. When creating video contents, the contents are provided with context data on the video (location and time), the aim is to also include other information on the event/situation. For instance, it is possible to save the contents of a calendar entry of the mobile phone as a part of context data. Furthermore, the service can save information on other users in the same location at the same time by means of user management and thus create a dynamic group for recording the experience. This more extensive context data is utilised in creating the video compilation describing the experience. Particularly by utilising the user data, it is possible to select in the automatic video compilation contents created in the same situation in order to be able to create of the experience a presentation as versatile as possible and describing it well.

Systems designed for recording and sharing experiences often include various mobile devices which contain sensors measuring the status and behaviour of the person. On the other hand, the experience is recorded as multimedia contents: pictures, videos and sound. Various collection and sharing services of multimedia contents exist (e.g. Flickr, YouTube), but they lack the active contents combination implemented by the service for describing the experience.

Automatic video compilation for dynamic group

In the invention by means of context data (inter alia, location and time) related to video contents, it is possible to create dynamic user groups for creating a video compilation. Based on the context data, users at a specific period are collected

into one group at a specific time and a compilation of the video clips of these users is created to be shared with the group.

The libraries of digital mobile videos have expanded and watching them is considered wearisome and unexciting. Traditionally, the problem has been solved
5 by editing a presentation of the videos by cutting and pasting them together. However, this requires special tools, time and expertise which an ordinary user does not have. Furthermore, in extensive video libraries, which include videos recorded by several users, the combining of videos into logical units involves additional challenges for the editing process. Grouping the videos without the
10 context data of the video is extremely challenging. Furthermore, combining contents shot by others for one's own purposes is awkward as the material is difficult to obtain and edit.

The principle of the invention is that the system segments the videos into smaller sections, selects the best of them and compiles them into a new video
15 presentation. The segmenting of videos (*video segmentation*) is implemented by analysing the sound, picture and motion features calculated of the video (*feature extraction*). The selection of the video segments is implemented by giving them points by means of supplied weighting coefficients and feature values. Video segments are compiled into scenes (*video scene*) by clustering the segments
20 based on the video time stamp (*timestamp*) and geographical data (*geotagging*). Finally, the scenes are combined into one video presentation (*video rendering*). The selection of video segments focuses on contents created by other persons and on combining them to one's own contents. The emphasis of compilation is in the contents others have created in the same location at the same time as the
25 acquirer of the video compilation. Thus, the video compilation is provided with shots by other persons who shot the same subjects as the acquirer of the compilation.

Collaborative video compilation for group

Various social media services are very popular. Different user groups create
30 multimedia contents and share it with other users and user groups. By means of the automatic video compilation implemented with an advantageous embodiment of the invention, it is possible to offer a group video compilations of video clips created by the group and thus to create a more extensive video of the experiences

of the group members to describe the collaborative experience of the group. When creating the compilation, data on the shooting location and date of the video clips is utilised in order to be able to determine the segments selected into the compilation and to arrange them in a suitable order.

- 5 The libraries of digital mobile videos have expanded and watching them is considered wearisome and unexciting. Traditionally, the problem has been solved by editing a presentation of the videos by cutting and pasting them together. However, this requires special tools, time and expertise which an ordinary user does not have. Furthermore, in extensive video libraries, which include videos
- 10 recorded by several users, the combining of videos into logical units involves additional challenges for the editing process. Grouping the videos without the context data of the video is extremely challenging. Furthermore, combining contents shot by others for one's own purposes is awkward as the material is difficult to obtain and edit.
- 15 The principle of the invention is that the system segments the videos into smaller sections, selects the best of them and compiles them into a new video presentation. The segmenting of videos (*video segmentation*) is implemented by analysing the sound, picture and motion features calculated of the video (*feature extraction*). The selection of the video segments is implemented by giving them
- 20 points by means of supplied weighting coefficients and feature values. Video segments are compiled into scenes (*video scene*) by clustering the segments based on the video time stamp (*timestamp*) and geographical data (*geotagging*). Finally, the scenes are combined into one video presentation (*video rendering*). The selection of video segments focuses on the contents created by persons
- 25 known by the acquirer of the video compilation. Acquaintances can be managed, inter alia, with a social network service, such as Facebook. The emphasis of compilation is in the contents the acquaintances have created in the same location at the same time as the acquirer of the video compilation. Thus, the video compilation is provided with shots by acquaintances who shot the same subjects
- 30 as the acquirer of the compilation.

Gesture interface for creating video compilation

In an embodiment of the invention, the creation of the video compilation, particularly in connection with mobile phones, utilises acceleration and bearing

sensors connected to the mobile station by means of which a gesture interface is implemented. By means of this, the user defines e.g. a period from the videos of which contents are selected into the video compilation. Furthermore, the user has a possibility to affect the creation of the video compilation by means of the gesture interface e.g. by changing the rhythm and tempo of the video compilation by shaking or tilting the mobile phone or equivalent.

The main principle of the invention is that the system segments the videos into smaller sections, selects the best of them and compiles them into a new video presentation. The segmenting of videos (*video segmentation*) is implemented by analysing the sound, picture and motion features calculated of the video (*feature extraction*). The selection of the video segments is implemented by giving them points by means of supplied weighting coefficients and feature values. Video segments are compiled into scenes (*video scene*) by clustering the segments based on the video time stamp (*timestamp*) and geographical data (*geotagging*). Finally, the scenes are combined into one video presentation (*video rendering*). For the creation of the compilation is used a mobile phone provided with sensors and, for combining segments, the use of sensors is utilised. The creation of video editing can be controlled by motion and bearing sensors. The sensors can e.g. limit the parameters used in the creation of the compilation: a short and fast motion creates a fast compilation including a lot of motion or limits the contents of the compilation in terms of geography or time. Equivalently, a slow and extensive motion creates a more peaceful compilation or a compilation selected more widely in terms of geography and time. During the compilation, a tap on the phone forces the segment to change and tilting can affect the effect which is created in the change of the segment.

The invention is not limited to the above exemplifying embodiment, but many variations are possible within the scope of the inventive idea defined by the claims.

Claims

1. A video editing system, **characterised** in that the system comprises:
 - means for defining features of videos;
 - means for automatically searching in a group of videos those videos or video
5 sections i.e. segments which include the above defined features; and
 - means for creating a video compilation of videos and/or segments including the above defined features.
2. A video editing system according to claim 1, **characterised** in that features to be defined include the geographical region where the video has been shot and the
10 time and/or time period when the video has been shot and recorded.
3. A video editing system according to claim 1 or 2, **characterised** in that the system comprises means for selecting suitable sections, segments, of the videos, which means include a video feature analysing unit which analyses sound, picture and/or motion features calculated of the video (i.e. content analysis metadata)
15 based on which the segments are selected.
4. A video editing system according to claim 1, 2 or 3, **characterised** in that the system comprises means for creating a new video i.e. video compilation of the videos or their segments most advantageously automatically.
5. A video editing system according to claim 3, **characterised** in that the
20 selection of video segments is arranged to be implemented by calculating the above features and by giving each segment points based on this and by selecting those segments in the video compilation which comply with the predetermined point limits.
6. A video editing system according to claim 3 or 5, **characterised** in that the
25 system comprises means for creating video scenes of selected video segments by clustering the video segments on the basis of the time stamp and/or geographical data.
7. A video editing system according to claim 3, 5 or 6, **characterised** in that the system comprises means for creating a new video or video presentation i.e. video
30 compilation of the videos or their segments based on the contents of the user and

those recommended to the user i.e. the selected segments such that the location, time or content analysis metadata constitute continua between the video segments.

8. A video editing system according to claim 1, 2 or 3, **characterised** in that the system comprises means according to claim 1 for searching videos in a multimedia database or equivalent video database.

9. A video editing system according to claim 8, **characterised** in that the system comprises means for searching, by means of specific search criteria, the desired videos in a multimedia database, of which videos a video compilation is to be created.

10. A video editing system according to claim 9, **characterised** in that the system comprises means for combining the classified video segments into a new video i.e. video compilation.

11. A video editing system according to claim 9, **characterised** in that the system comprises a video database, such as a multivideo database, which is shared by several users and accessible through a data network, such as the Internet. (KI 1.3)

12. A video editing system according to claim 11, **characterised** in that the system comprises means for updating the video database with new videos.

13. A video editing system according to claim 11 or 12, **characterised** in that the system comprises means for automatically performing video editing according to predetermined features.

14. A video editing system according to any one of preceding claims, **characterised** in that the video editing system is arranged on a data network, advantageously on the Internet, where it is accessible for the users.

15. A video editing system according to claim 14, **characterised** in that the system comprises a music database of which the user can select music to accompany the video compilation.

16. A method for editing videos, **characterised** by the method comprising the following steps:

- defining features of videos;

- automatically searching in a group of videos those videos or video sections i.e. segments which include the above defined features; and
- creating a video compilation of videos and/or segments including the above defined features.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/FI2010/050309

A. CLASSIFICATION OF SUBJECT MATTER

See extra sheet

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: G11B, G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
FI, SE, NO, DKElectronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, WPI, IEEE Xplore

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2008304806 A1 (LIN, B.) 11 December 2008 (11.12.2008) abstract, paragraphs [0017]-[0023], Figs. 1, 4, and 5, claim 1	1, 4, 16
X	EP 1241673 A2 (GATEWAY INC) 18 September 2002 (18.09.2002) abstract, par. [0017]-[0021], claim 1, Fig. 2	1, 4, 16
X	US 6697564 B1 (TOKLU, C. et al.) 24 February 2004 (24.02.2004) abstract, column 4 line 57 - col. 6 l. 8, Figs. 1 and 4, claim 1	1, 4, 16
X	WO 2006065223 A1 (MUVEE TECHNOLOGIES PTE LTD et al.) 22 June 2006 (22.06.2006) claim 3	2
X	claim 4	6
X	WO 2007082169 A2 (EYESPOT CORP et al.) 19 July 2007 (19.07.2007) par. [0079]-[0082], Figs. 2-5	8-10
X	abstract, par. [0031]-[0036], Fig. 2	14
X	par. [0033], Fig. 2	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 August 2010 (17.08.2010)

Date of mailing of the international search report

20 August 2010 (20.08.2010)

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

Facsimile No. +358 9 6939 5328

Authorized officer

Timo Laakso

Telephone No. +358 9 6939 500

INTERNATIONAL SEARCH REPORT

International application No.

PCT/FI2010/050309

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 6757027 B1 (EDWARDS, E. et al.) 29 June 2004 (29.06.2004) col. 2 l. 13 - col. 3 l. 52, claims 10 and 14, Figs. 1A-C	11-13
A	US 2006251382 A1 (VRONAY, D. et al.) 09 November 2006 (09.11.2006) the whole document	1-16
A	US 2005152666 A1 (DEMEYER, M.) 14 July 2005 (14.07.2005) the whole document	1-16
A	US 7362946 B1 (KOWALD, J.) 22 April 2008 (22.04.2008) the whole document	1-16

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/FI2010/050309

Patent document cited in search report	Publication date	Patent family members(s)	Publication date
US 2008304806 A1	11/12/2008	None	
EP 1241673 A2	18/09/2002	JP 2002359811 A US 7248778 B1	13/12/2002 24/07/2007
US 6697564 B1	24/02/2004	None	
WO 2006065223 A1	22/06/2006	GB 2437004 A US 2008101762 A1 JP 2008523539T T EP 1839434 A1 CN 101180870 A CA 2590234 A1	10/10/2007 01/05/2008 03/07/2008 03/10/2007 14/05/2008 22/06/2006
WO 2007082169 A2	19/07/2007	None	
US 6757027 B1	29/06/2004	None	
US 2006251382 A1	09/11/2006	US 2006251384 A1 US 2006251383 A1	09/11/2006 09/11/2006
US 2005152666 A1	14/07/2005	None	
US 7362946 B1	22/04/2008	JP 2000350159 A AU 2521400 A AU 745436B B2	15/12/2000 19/10/2000 21/03/2002

INTERNATIONAL SEARCH REPORT

International application No.

PCT/FI2010/050309

CLASSIFICATION OF SUBJECT MATTER

Int.Cl.

G11B 27/031 (2006.01)

G06F 17/30 (2006.01)