

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

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
11 April 2013 (11.04.2013)

(10) International Publication Number
WO 2013/050276 A1

(51) International Patent Classification:

G11B 27/10 (2006.01) *G06K 9/64* (2006.01)
G11B 27/28 (2006.01) *G06T 7/00* (2006.01)
G06F 17/30 (2006.01)

Belle Fontaine, CS17616, F-35576 Cesson-Sévigné (FR).
GENDROT, Rémy [FR/FR]; c/o Technicolor R&D France, 1 Avenue de Belle Fontaine, CS17616, F-35576 Cesson-Sévigné (FR).

(21) International Application Number:

PCT/EP2012/068909

(74) Agents: **LORETTÉ, Anne** et al.; 1-5 rue Jeanne d'Arc, F-92130 Issy-Les-Moulineaux (FR).

(22) International Filing Date:

26 September 2012 (26.09.2012)

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

11306284.8 4 October 2011 (04.10.2011) EP

(71) Applicant: **THOMSON LICENSING** [FR/FR]; 1-5 rue Jeanne d'Arc, F-92130 Issy-les-Moulineaux (FR).

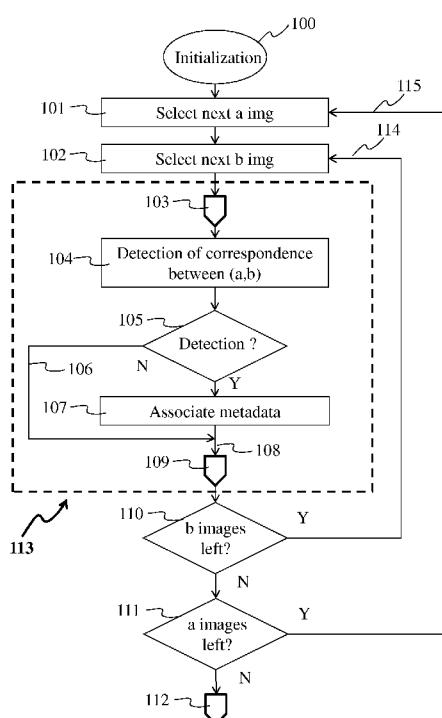
(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ,

(72) Inventors; and

(71) Applicants (for US only): **MONTALVO, Luis** [EC/FR]; c/o Technicolor R&D France, 1 Avenue de Belle Fontaine, CS17616, F-35576 Cesson-Sévigné (FR). **STRAUB, Gilles** [FR/FR]; c/o Technicolor R&D France, 1 Avenue de

[Continued on next page]

(54) Title: METHOD OF AUTOMATIC MANAGEMENT OF A COLLECTION OF IMAGES AND CORRESPONDING DEVICE



(57) Abstract: The present invention relates to the field of management of image data in data storage. In particular, the present invention relates to a method and device for automatic detection of correspondence between images in data storage and corresponding device, which method and device are particularly efficient with regard to the automatic management of large amounts of dispersed image data.

Declarations under Rule 4.17:

UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

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

Published:

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

Method of automatic management of a collection of images and corresponding device.

1. Field of invention.

5 The present invention relates to the field of management of image data in data storage. In particular, the present invention relates to a method and device for automatic detection of duplicate images in data storage and corresponding device, which method and device are particularly efficient with regard to the automatic management of large amounts of dispersed image
10 data.

2. Technical background.

15 The proliferation of digital devices that comprise a photo camera has favored an explosion of the volume of image data stored by a user, and it is quite easy for a user to end up with many image duplicates in the user's image library.

20 This situation can be even worse in the case of a home network environment, where several users can add images to an image library, the library possibly being physically distributed on several, dispersed storage devices, for example on hard drives of different PCs, on a NAS (Network Attached Storage), on USB keys, etc.).

25 The reasons why an image library can end up by containing many duplicate images are diverse. Unintentional duplicate images are produced through copy actions. For example, a user who organizes photos in different directories does not move the photos, which would have been appropriate, but unintentionally rather copies them; a user that wishes to transfer photos via e-mail adapts the photo resolution for including them in his e-mail but unintentionally keeps the low-resolution copies; a user that views images with a viewer application modifies these by rotation, or modification of color and
30 contrast and unintentionally keeps the unmodified copy in addition to the modified copy. Other copy actions are intentional and are due to the fact that the user has no longer an overview of the data that he has stored, a situation that is getting worse when the user has multiple storage devices and many images, and gets even worse when multiple users add and copy data to the
35 multitude of images stored. The user, knowing that he does not have a clear overview of the images stored, worsens this situation by preferring finally to copy rather than to move or replace images, by fear of deleting them. This

creates a situation where the user no longer knows which images are disposable copies and which are not.

In all these scenarios, a duplicate detection tool can be necessary, or at least useful, to assist the user with the cleanup or management tasks of the user's image library.

Prior-art detection of image duplicates detects duplicates according to criteria such as checksum data, creation data, file name, file size, and image format. Such criteria allows only detection of identical copies of an original image but not the copies that have been slightly or largely modified in order to enhance the visual perception of the image on a particular display.

Moreover, if more than one criterion for the duplicates detection is specified, duplicates are detected that comply with any of the selected criteria and user intervention is needed to determine if the user wishes to delete the detected duplicates from the image library. Other duplicate detection methods are capable to detect near-duplicate images by comparing image pixel data. A user is required to specify a matching percentage of pixel data of two images to mark and detect an image as being a duplicate image. The detection then detects these near-duplicates as it detects strictly identical images without distinction.

Thus, prior art solutions can still be optimized with regard to detection of duplicate images in data storage. Notably, a method is needed that reduces user intervention to the strictly needed.

3. Summary of the invention.

The invention reduces the complexity of maintaining a collection of images.

When near duplicates are found according to the method of the invention, associated metadata resumes the reason why these are considered as being duplicates (exact, near, far), for example low resolution copy, zoom, copy stored in a backup zone, etc. A set of management action rules that are associated to this metadata then allows managing the image library automatically according to these management action rules. Automatic management actions are taken for images which are considered as corresponding by the method of the invention based on the discussed metadata and associated rules. These management actions are for example delete, keep, or replace-by-link to the original image. The latter option can be

required when the existence of identical copies is to be avoided for reasons of efficiency.

The discussed advantages and other advantages not mentioned here, 5 that make the device and method of the invention advantageously well suited for automatic management of a collection of images, will become clear through the detailed description of the invention that follows.

In order to automatically manage a collection of images, the invention 10 proposes a method comprising a step of detection of correspondence between a first image and at least a second image in the collection of images according to at least one criterion for correspondence between the first image and the at least one second image, and a step of association of metadata to the at least a second image when the correspondence is detected, the 15 metadata being representative of a relation between the first image and the at least one second image and comprising the at least one criterion for correspondence between the first image and the at least one second image which has lead to the detection of correspondence.

20 According to a variant embodiment of the invention, the method further comprises a step of automatic determination and application of one of a set of predetermined actions for processing of the at least one second image according to the associated metadata.

25 According to a variant embodiment of the invention, the metadata further comprises information representative of a level of correspondence between the first image and the at least one second image.

30 According to a variant embodiment of the invention, the set of predetermined actions comprise an action of replacement of the at least one second image with a link to the first image.

35 According to a variant embodiment of the invention, the set of predetermined actions comprise an action of deletion of the at least one second image.

According to a variant embodiment of the invention, the set of predetermined actions comprise an action of transferring the at least one second image to a storage.

5 According to a variant embodiment of the invention, the set of predetermined actions comprise an action of renaming the at least one second image.

10 The invention also concerns a device for automatic management of a collection of images, the device comprising means for detection of correspondence between a first image and at least a second image in the collection of images according to at least one criterion for correspondence between the first image and the at least one second image, and means for association of metadata to the at least one second image when the 15 correspondence is detected, the metadata being representative of a relation between the first image and the at least one second image and comprising the at least one criterion for correspondence between the first image and the at least one second image which has lead to the detection of correspondence.

20

4. List of figures.

More advantages of the invention will appear through the description of particular, non-restricting embodiments of the invention. The embodiments will be described with reference to the following figures:

25 **Figure 1** shows a method for association of metadata to one or more images when a correspondence is detected.

Figure 2 shows a detection of correspondence according to a variant embodiment of the invention.

30 **Figure 3** shows an application of actions associated to the detection according to the invention.

Figure 4 illustrates the notion of normalized fingerprint distance (NFD) between two images and the relation between NFD and discussed thresholds.

35 **Figure 5** shows an example device implementing a variant embodiment of the invention.

5. Detailed description of the invention.

Figure 1 shows a method for association of metadata to one or more images when a correspondence is detected.

In a first initialization step 100, variables are initialized for the functioning of the method. When the method is implemented in a device such as device 400 of figure 4, this may comprise copying of data from non-volatile memory to volatile memory and initialization of memory. In a next step 101, a first image, "a", is fetched in data storage. In a next step 102, a second image, "b", is fetched from data storage. In a step 104, a detection of correspondence between the two images is done according to different criteria. In a test step 105, it is determined if such detection is affirmative (i.e. a correspondence between the two images is detected according to one or more of the criteria for correspondence) or not. If the detection is affirmative, metadata is associated to the second image in a step 107. This metadata comprises information on a relation between the second and the first image and information about the one or more criteria of correspondence that has lead to the detection of the correspondence. When no correspondence can be detected, step 107 is not executed. In a next step 110, it is verified if there are any second ("b") image left that have not been compared to the "a" image. If so, a next second ("b") image is selected in step 102 and the steps of detection (dotted rectangle 113) are repeated. If not, it is verified in a step 111 if there are any first images ("a") left which have not yet been processed by the detection method of the invention. If so, a next first ("a") image is selected in step 101, a next second "b" image is selected in step 102 and the steps of detection (113) are repeated. If all first "a" images of the image collection in data storage have been processed by the method of the invention, point 112 is reached, which links the detection method of the invention to an automatic determination and application of one of a set of predetermined actions for processing of all second images ("b" images) according to the associated metadata, described by means of figure 3. According to the described embodiment, the images in the data storage are first completely processed by the detection method, before being processed by the automatic determination and application of predetermined actions. According to a variant embodiment, the latter automatic determination is done immediately following the discussed association of metadata. This variant has an advantage in processing time because the mass of images to be processed by the detection method can be reduced; each time when

images that are deleted by delete actions the mass of images to process is reduced. An intelligent selection method for first ("a") and second "b" images can further reduce the processing time needed. For example, a step is added to the method that excludes detection of correspondence of two images that 5 have already been passed through the detection process. According to this variant, the detection method associates metadata to each image that has been completely processed by the detection process which metadata indicates that the image has already been processed as a first "a" image, and in each next iteration of the detection method where a next first "a" image is 10 selected, already processed first "a" images are not being processed in the detection method again, i.e. they are not selected as second ("b") images.

Figure 2 shows a detection of correspondence according to a particular embodiment of the invention. The method starts at point 103 and 15 ends at point 109 and corresponds to a detailed view of the detection method 113 of figure 1. Variant embodiments of the method of the invention are possible. Notably the steps of the method can be executed in a different order; more or less criteria (and thus tests) can be added to /removed from the detection of correspondence while still using the method of automatic 20 management of a collection of images according to the invention.

In a first test step 200, it is determined if a checksum calculated over the first ("a") image is the same as a checksum calculated over the second ("b") image. Checksum calculation is done through known methods, such as SHA (Secure Hash Algorithm) or MD5 (Message Digest 5). If the calculated 25 checksum is the same, the two images are considered as being identical and a decisional step 201 is done, in which it is determined if the location where the second ("b") image is stored is a location for storage of backup. If so, metadata is added in step 203 to the identical second ("b") image that indicates that the second image is a backup copy of the first image. If not, 30 metadata is added in step 202 to the identical second image that indicates that the second image is an identical copy. As will be handled further on, it is possible to automatically delete identical images that are not backup copies by execution of actions associated to metadata. If, as an outcome of test step 200 it is on the contrary determined that the checksums of the first and 35 the second images are different, a test step 204 is executed, in which it is determined if a normalized distance d between fingerprints of the first "a" image fp(a) and of the second "b" image fp(b) is below a first threshold th2

$d(fp(a),fp(b)) < th2$; $th2$ is a threshold that is chosen such that if $d(fp(a),fp(b)) < th2$, the second image "b" can be considered as being a modified copy of the first image "a". If $d(fp(a),fp(b))$ is not inferior to $th2$, the first and the second images are considered as being different by the method of the invention and the method continues with step 109. But if $d(fp(a),fp(b))$ is inferior to $th2$, we are dealing with a modified copy and it can be determined in following steps how the difference between the two images can be characterized. Notably, in a next step 205, the previously calculated normalized fingerprint distance is compared with a next threshold $th1$. If 5 $d(fp(a),fp(b))$ is superior to $th1$, the second image "b" is characterized in a step 206 as being a largely modified copy of the first image "a" and corresponding metadata is associated to the second image for example according to table 1, first row (LMC, <path>/a). If on the contrary 10 $d(fp(a),fp(b))$ is inferior to $th1$, a test step 207 is executed, in which it is verified if the first image ("a") has the same resolution as the second image 15 "b". Image resolution can be compared based on prior-art file metadata that is present in prior-art file systems, such as EXIF (Exchangeable Image File Format). If the image resolutions differ, a step 208 is executed in which metadata is associated to the second image that indicates that the second 20 image is a different resolution copy of the first image; e.g. a tag 'DRC' is added to metadata associated to image b together with the storage path of image a: (DRC, <path>/a). If on the contrary the resolution of the first image differs from that of the second image, a next test step 209 is executed, in which the encoding methods of the two images are compared. This 25 comparison is done according to known methods as for example by comparing file extensions (e.g. *.jpg, *.tiff). If the two images are encoded with a different encoding method, a step 210 is executed in which corresponding metadata is associated to the second image, e.g. a tag 'DEC' is added to image b together with the storage path of image a: (DEC, <path>/a). If on the contrary the two images are encoded with different 30 encoding methods, step 211 is executed in which metadata (SMC, <path/a>) is associated to the second image. After steps 202, 203, 206, 208, 210 and 211, step 109 is executed, returning to figure 1, where the steps of the method are iterated until all images have been processed. **Table 1** hereunder 35 resumes example types of metadata tags, their meaning and their means of determination.

| Tag | Meaning | Means for determination |
|-----|---|--|
| IDC | Image 'b' is an Identical Copy of image 'a' | Checksum |
| BC | Image 'b' is a Backup Copy of image 'a' | Checksum and storage location |
| LMC | Image 'b' is a Largely Modified Copy of image 'a' | Normalized image fingerprint distance ($th1 < NFD < th2$) |
| DRC | Image 'b' is a Different Resolution Copy of image 'a' | Image resolution |
| DEC | Image 'b' is a Different Encoding Copy of image 'a' | Image encoding method |
| SMC | Image 'b' is a Slightly Modified Copy of image 'a' | Normalized image fingerprint distance ($NFD < th1$) |

Table 1

5 **Figure 3** shows an application of actions associated to the detection according to an example embodiment of the invention. According to a variant embodiment of the invention as illustrated in figure 3, these actions are done following the execution of the detection steps of figs 1 and 2 (see pointer 112 in figs 1 and 3). According to yet another variant embodiment, actions are 10 executed as soon as metadata has been associated to an image, which is advantageous in terms of resources used for execution of the method. This latter variant is possible for actions that are not delete actions, such as actions creating a link, the creation of the link reduces the amount of data to be processed by subsequent iterations of the method.

15 In a first step 300, a next second image is chosen ("b" image). Its associated metadata is read in step 301 and in a step 302 an action is determined for the associated metadata, for example, according to the actions as defined in table 3. In a test 303, it is determined if the action associated to the metadata is the creation of a file link. If so, a file link is 20 created in a step 306, from the second image to the first image. The metadata remains associated to the link, so that for future iterations of the method of the invention, a trace is kept. If the action is not a create link, it is verified in a test 304 if the action is a delete image; if so, the second image is

deleted in a step 307. If the action is not a delete image action neither, it is verified in a test 305 if the action is an ask action, and if so, the second image is transferred to a temporary storage in a step 308, where images are stored for which a user decision is needed. If not, the action steps are 5 repeated with a selection of a next second image in step 300. This is also the case after steps 306, 307 and 308. The processing ends when all images have been processed.

Variant embodiments of the discussed application of actions are possible. Notably the steps of the method can be executed in a different 10 order; more or less actions (and thus tests) can be added/removed.

The method of the invention can be applied as a background task or as a clean-up tool that is more or less regularly executed. The method can be enhanced with a monitoring feature that monitors creation, deletion and 15 copying of images so as to keep the metadata updated as soon as a creation, deletion or copying is executed.

Table 2 hereunder illustrates an example lookup table for looking up actions that are associated to a tag type. The tags types are those of the example implementation illustrated by means of figures 1 and 2. For a tag 20 type 'IDC' (Identical Copy) the associated action executed by the method of the invention is to replace the second image ("b") by a link to the first image ("a"). When a second image has a metadata tag BC or LMC, no action is associated since it is wished to keep the second image. When the second image has a tag 'DRC', the associated action is to delete the second image only when the second image has a lower resolution than the first image. When the second image has an associated metadata tag 'DEC', the 25 associated action is to delete the second image only if the first image is of the 'png' encoding type. When the second image has an associated tag 'SMC', the associated action is to ask the user to decide what to do. According to a variant embodiment of the invention, images with associated action 'Ask' are 30 grouped in temporary storage and the user is only bothered once for a review of all images in this temporary storage with associated action 'Ask' for which the user's decision is required. Such a review can for example be done through a visual presentation of the corresponding first and second images 35 image pair and with a possibility for un-checking a 'keep' checkbox related to each second image of the image pair.

According to a variant embodiment of the invention, multiple metadata tags can be associated to a single image. For example, a same image can have both DRC and DEC tags, meaning that the image is a different resolution copy but also a different encoding copy. In this case, the steps of 5 the method are not executed as depicted in figure 2, but in parallel or in a different order. This variant embodiment has the advantage to allow a more extensive association of metadata, which is advantageous for fine-tuning of the associated actions. Using the previous example of an image that has both DRC and DEC tags and referring to table 2, an associated action is to 10 only delete the image if both action conditions apply, i.e. to delete the second image the resolution of the second must be lesser than that of the first image AND the first image is encoded according to the PNG (Portable Network Graphics) encoding method.

According to a variant embodiment of the invention, the actions are 15 user-configurable.

| Tag | Action |
|-----|------------------------------|
| IDC | Replace b by link to a |
| BC | None |
| LMC | None |
| DRC | Delete b if res(b) < res (a) |
| DEC | Delete b if enc(a)=*.png |
| SMC | Ask |

Table 2

Figure 4 illustrates the notion of normalized fingerprint distance (NFD) 20 between two images and the relation between NFD and discussed thresholds. To be able to classify an image by its differences with another image, NFD is one of the 'tools' used by the method. Two fixed thresholds (th1 and th2) are used representing certain values of normalized distances between fingerprint vectors of the second ('b') image and the first ('a') image. 25 This normalized distance can be expressed as:

$$\Delta(a, b) = \frac{\|a - b\|}{\|a + b\|}$$

Where $\|\cdot\|$ represents an L2 norm of a vector, i.e. its Euclidian distance.

An image fingerprint, constructed according to known prior-art methods, can be represented as an n-dimensional vector. "n" can have a 5 value of hundred or even thousand. In our example and for simplicity of illustration, we assume that n=2. The center of fig.4 (400) represents the image fingerprint of image 'a', i.e. fp(a). The zone 401 in the first circle around fp(a) corresponds to the fingerprints of the 'b' images whose distance to the fingerprint of the 'a' image is lower than the first threshold th1 (402), 10 and represents 'b' images that have been slightly modified with regard to the 'a' image. Zone 403 in the second circle around fp(a) corresponds to fingerprints of 'b' images whose distance to the fingerprint of the 'a' image is higher than the first threshold th1 (402) but lower than the second threshold th2 (404), and represents 'b' images that have been largely modified with 15 regard to the 'a' image. The zone 405 outside of the second circle corresponds to fingerprints of 'b' images whose distance to the fingerprint of the 'a' image is higher than the second threshold th2 (404), and represents 'b' images that can be considered to be different with regard to the 'a' image.

20 **Figure 5** shows an example device 500 that implements a variant of the method of the invention. The device 500 comprises the following components, interconnected by a digital data- and address bus 514:

25

- a processing unit 511 (or CPU for Central Processing Unit);
- a non-volatile memory NVM 510 ;
- a volatile memory VM 520 ;
- a clock 512, providing a reference clock signal for synchronization of operations between the components of the device 500 and for timing purposes;
- a network interface 513, for interconnection of device 500 to 30 other devices connected in a network via connection 515.

35 It is noted that the word "register" used in the description of memories 510 and 520 designates in each of the mentioned memories, a low-capacity memory zone capable of storing some binary data, as well as a high-capacity memory zone, capable of storing an executable program, or a whole data set.

Processing unit 511 can be implemented as a microprocessor, a custom chip, a dedicated (micro-) controller, and so on. Non-volatile memory

NVM 510 can be implemented in any form of non-volatile memory, such as a hard disk, non-volatile random-access memory, EPROM (Erasable Programmable ROM), and so on.

The Non-volatile memory NVM 510 comprises notably a register 5201 that holds a program representing an executable program comprising the method according to the invention. When powered up, the processing unit 511 loads the instructions comprised in NVM register 5101, copies them to VM register 5201, and executes them.

The VM memory 520 comprises notably:

- 10 - a register 5201 comprising a copy of the program 'prog' of NVM register 5101 ;
- a register 5202 comprising an iterator variable allowing an iteration over the first ('a') images;
- a register 5203 comprising an iterator variable allowing an iteration over the second ('b') images;
- 15 - a register 5204 for storing of one or more references to backup storage, so that the method can recognize when a storage location is a backup storage location;
- a register 5205 for storing of the first and the second thresholds for NFD calculation;
- 20 - a register 5206 holding a table of metadata tags that is to be used for association of metadata.

A device such as device 500 is suited for implementing the method of the invention of automatic management of a collection of images, the device comprising

- 25 - means for detection (CPU 511, VM register 5205) of correspondence between a first image and a second image(s) in said collection of images according to a criterion(s) for correspondence between said first image and the second image(s);
- 30 - means for association of metadata (CPU 511, register 5206) to the second image(s) when said correspondence is detected, the metadata being representative of a relation between the first image and the second image(s) and comprising the criterion(s) for correspondence between the first image and the second image(s) which has lead to the detection of the correspondence.

Other device architectures than illustrated by fig 5 are possible and compatible with the method of the invention. Notably, according to variant embodiments, the invention is implemented as a pure hardware implementation, for example in the form of a dedicated component (for 5 example in an ASIC, FPGA or VLSI, respectively meaning Application Specific Integrated Circuit, Field-Programmable Gate Array and Very Large Scale Integration), or in the form of multiple electronic components integrated in a device or in the form of a mix of hardware and software components, for example a dedicated electronic card in a personal computer.

CLAIMS

1. A method of automatic management of a collection of images, characterized in that it comprises the following steps:

5 detection (104) of correspondence between a first image and at least a second image in said collection of images according to at least one criterion for correspondence between said first image and said at least one second image;

10 association of metadata (107) to said at least a second image when said correspondence is detected, said metadata being representative of a reason of detection of said correspondence between said first image and said at least one second image and comprising said at least one criterion for correspondence between said first image and said at least one second image which has lead to said detection of correspondence; and

15 application of one of a set of predetermined management action rules for processing of said at least one second image according to said reason of detection of said correspondence comprised in said associated metadata.

2. The method according to claim 1, said action being one of: keep said at least one second image, replace said at least one second image by a link to said first image or deletion of said at least one second image.

3. The method according to claim 1 or 2, wherein the criterion of correspondence comprises a level of modification between said first image and said at least one second image, determined according to a fingerprint distance between said first image and said at least one second image.

4. The method according to any of claims 1 to 3, said reason of detection being one of a different resolution copy, a different encoding copy, a largely modified copy, or a slightly modified copy.

5. The method according to any of Claims 1 to 4, said metadata further comprising information representative of a level of correspondence between said first image and said at least one second image.

6. The method according to any of Claims 1 to 5, said set of predetermined actions comprising an action of transferring said at least one second image to a storage.

5 7. The method according to any of claims 1 to 6, said set of predetermined actions comprising an action of renaming said at least one second image.

8. A device for automatic management of a collection of images, the
10 device being characterized in that it comprises:

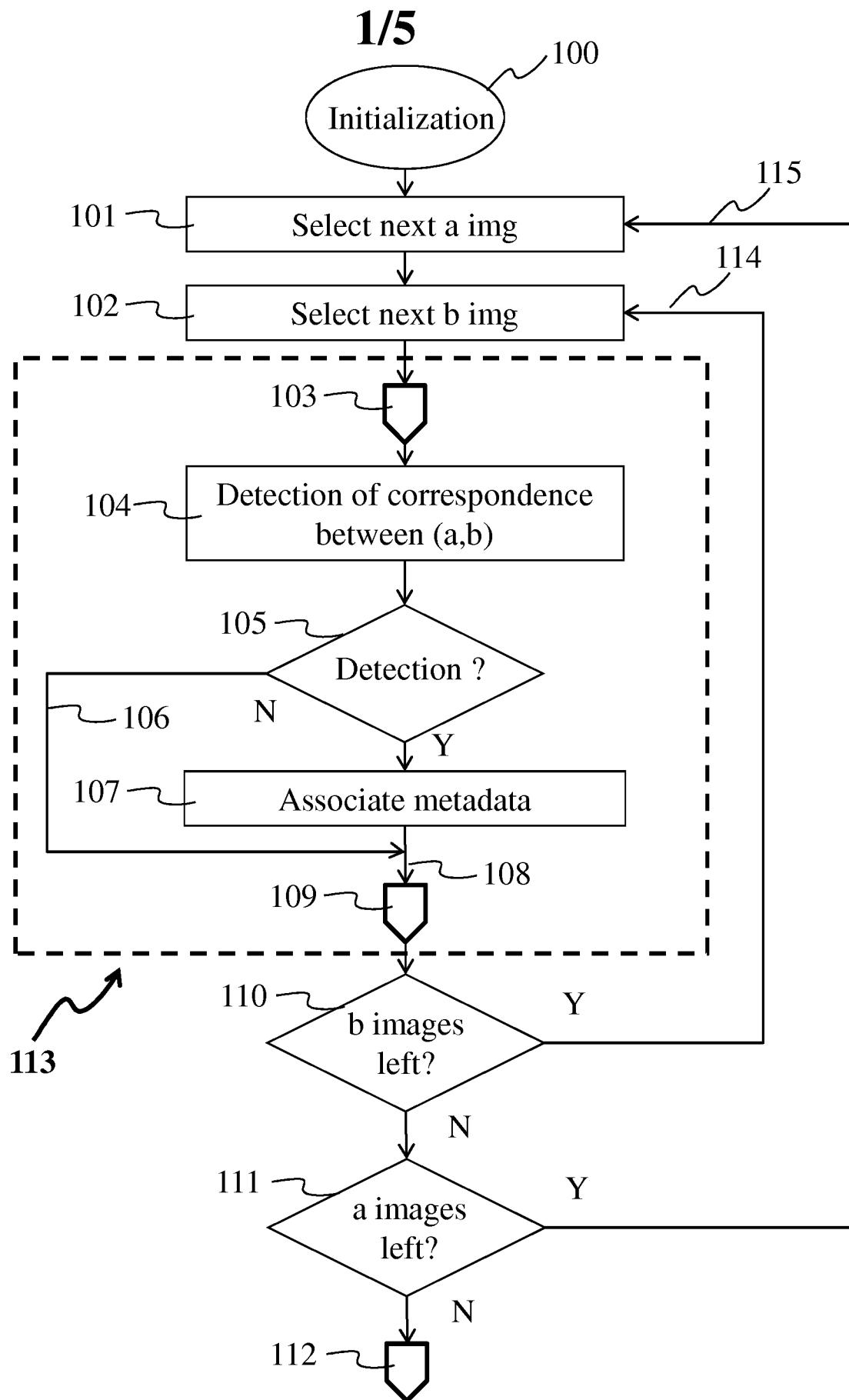
means for detection (511, 5205) of correspondence between a first image and at least a second image in said collection of images according to at least one criterion for correspondence between said first image and said at least one second image;

15 means for association of metadata (511, 5206) to said at least one second image when said correspondence is detected, said metadata being representative of a reason of correspondence between said first image and said at least one second image and comprising the at least one criterion for correspondence between said first image and said at least one second image
20 which has lead to said detection of correspondence; and

means for application of one of a set of predetermined management action rules for processing of said at least one second image according to said reason of detection of said correspondence comprised in said associated metadata.

25

30

**Fig. 1**

2/5

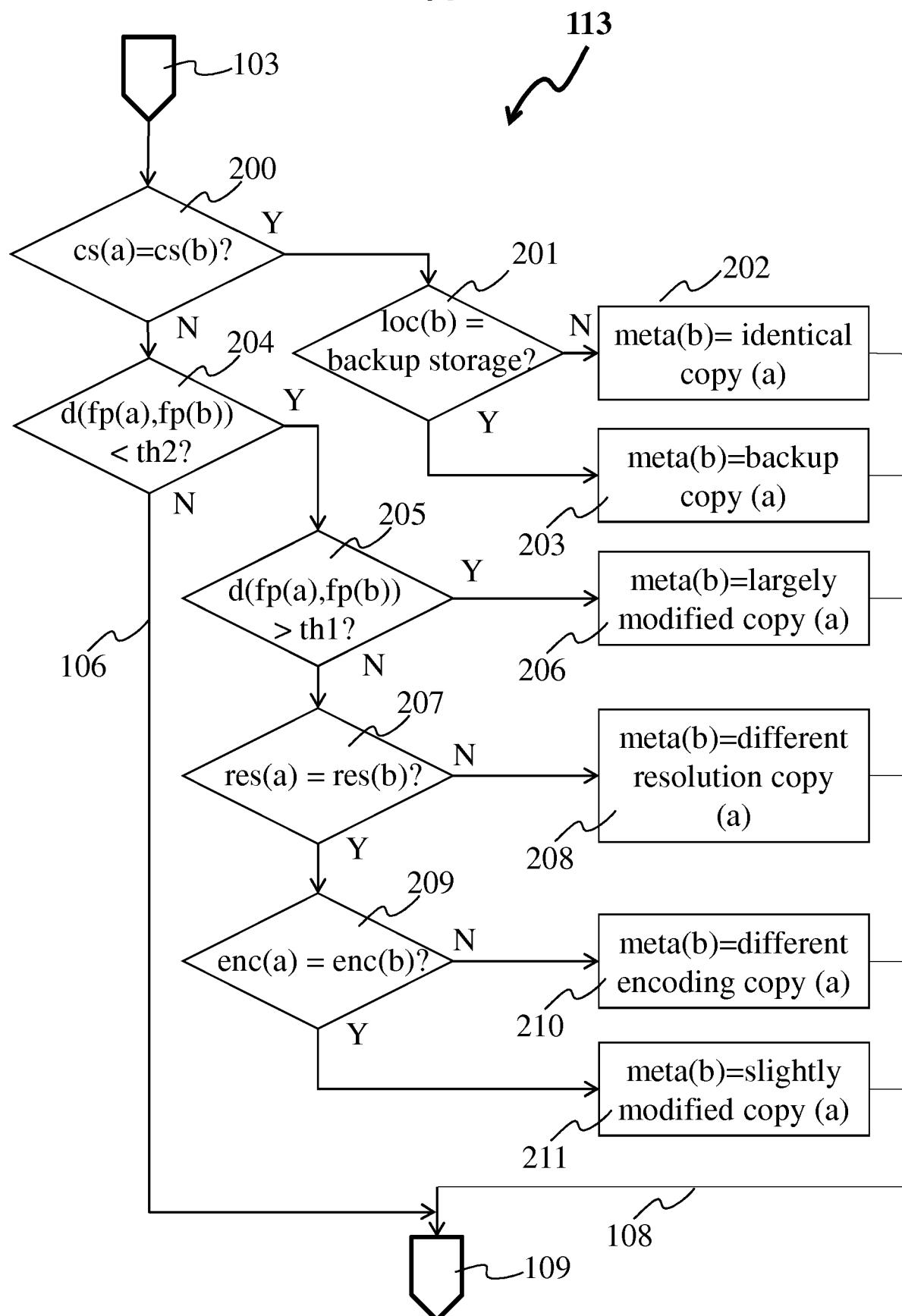


Fig. 2

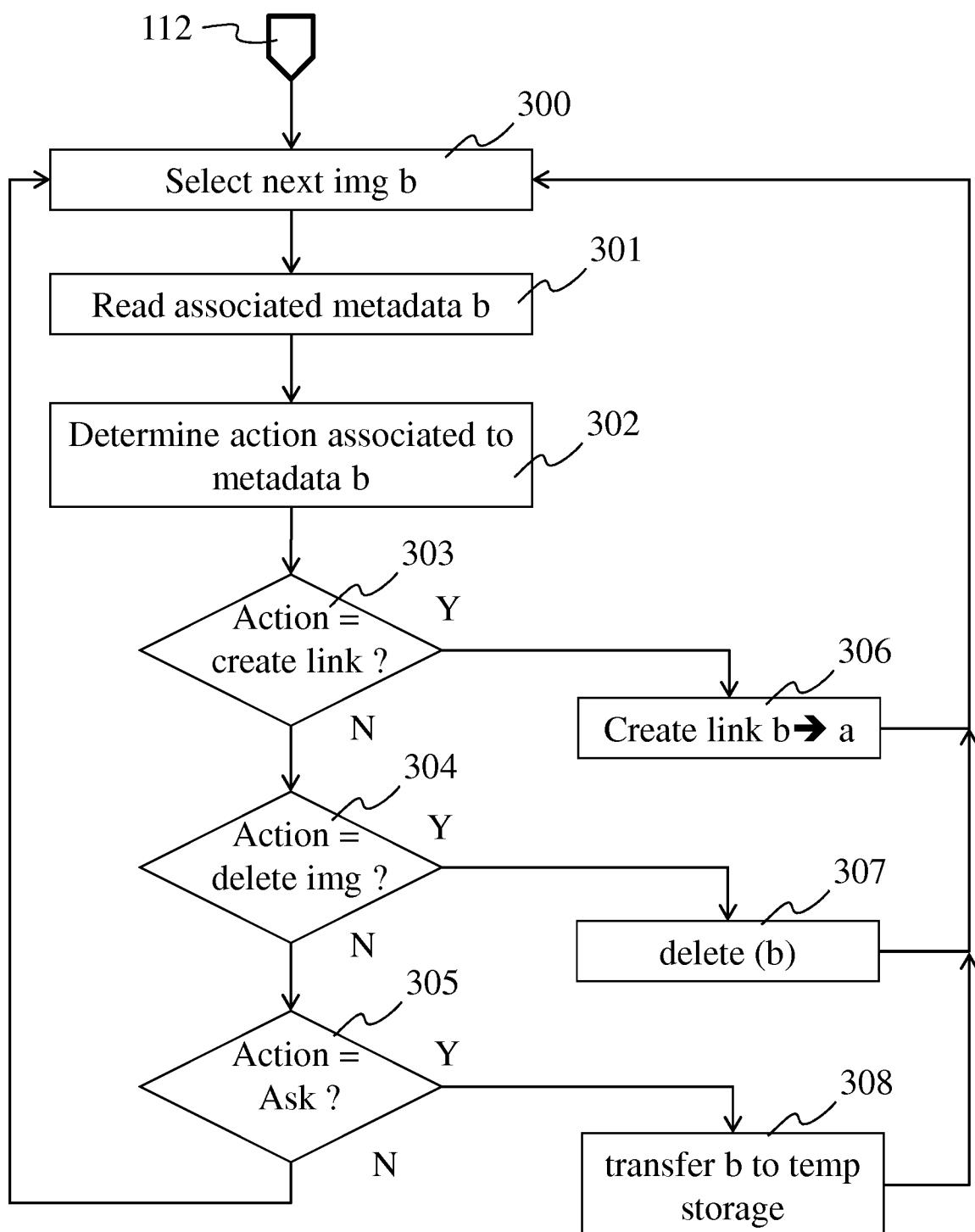


Fig. 3

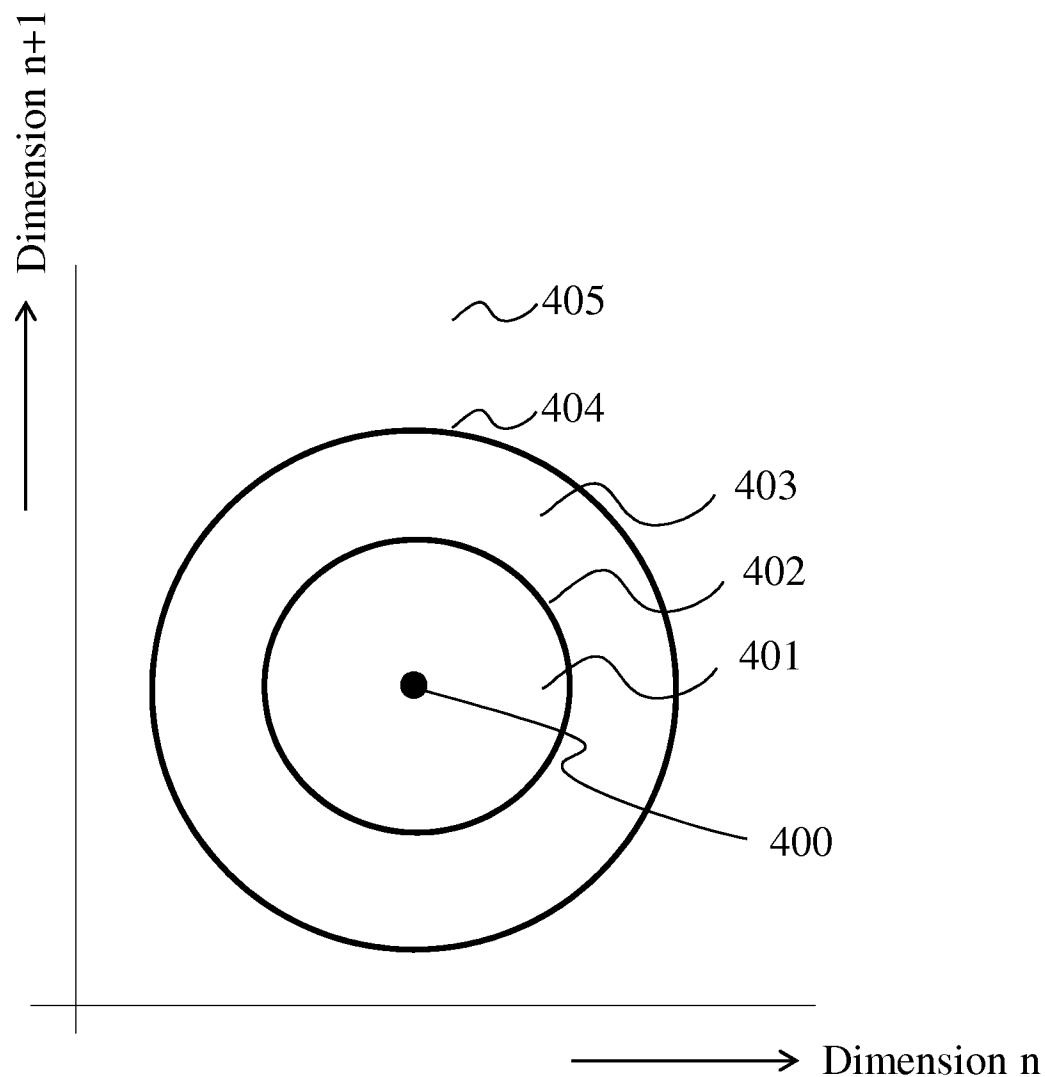


Fig. 4

5/5

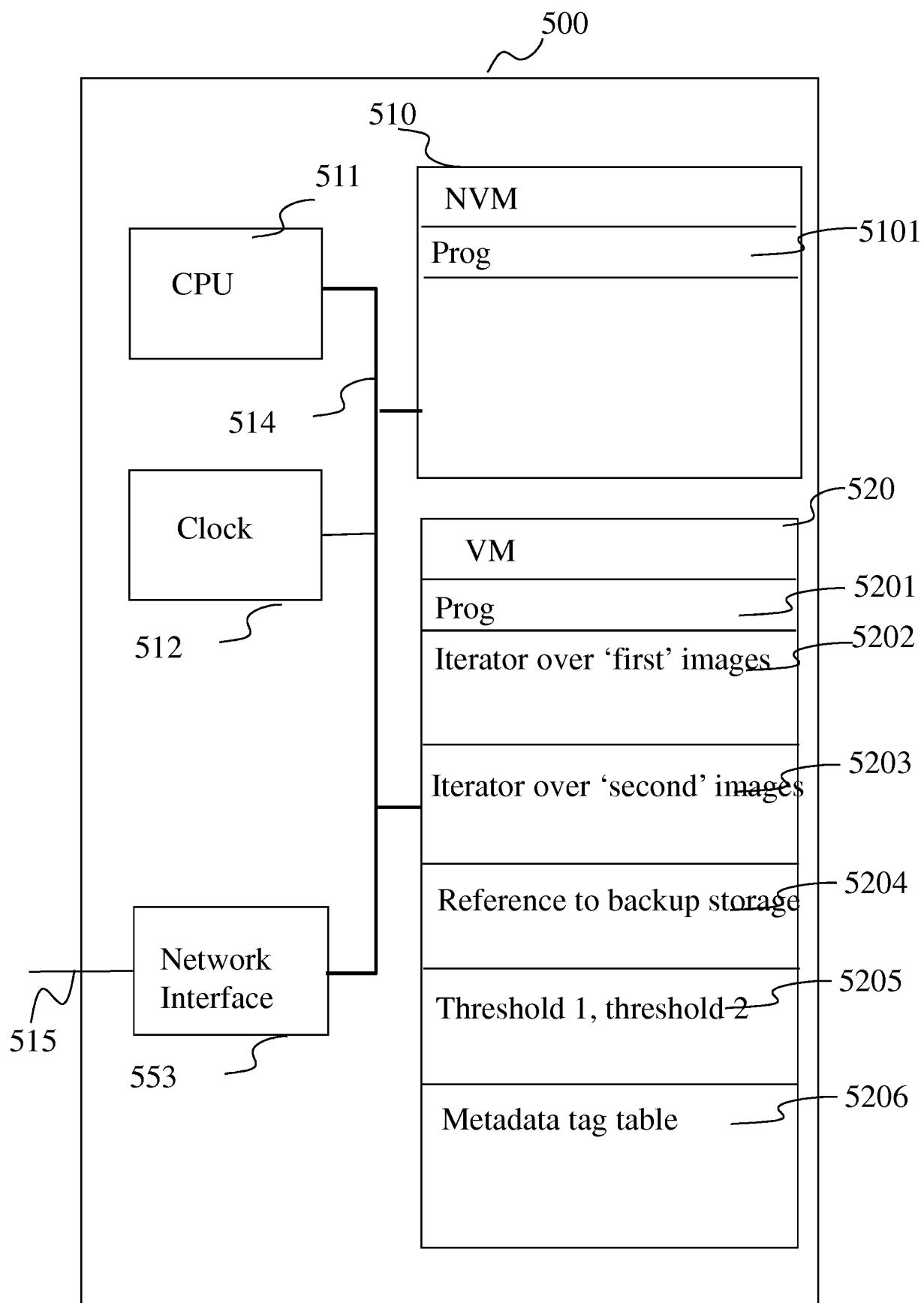


Fig. 5

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2012/068909

| | | | | |
|-------------------------------------|-----------|-----------|-----------|----------|
| A. CLASSIFICATION OF SUBJECT MATTER | | | | |
| INV. | G11B27/10 | G11B27/28 | G06F17/30 | G06K9/64 |
| ADD. | | | | |

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G11B G06F G06K G06T

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|-----------|---|-----------------------|
| X | US 2007/065045 A1 (IWASAKI MASAJIRO [JP]) 22 March 2007 (2007-03-22) figure 5 paragraphs [0050] - [0058] paragraphs [0068] - [0071] paragraph [0085] ----- US 2007/260639 A1 (TOBIN KENNETH W [US] ET AL TOBIN KENNETH WILLIAM [US] ET AL) 8 November 2007 (2007-11-08) paragraphs [0022], [0031], [0040], [0041] ----- EP 2 120 162 A1 (RICOH KK [JP]) 18 November 2009 (2009-11-18) figures 5-11 paragraphs [0025], [0034] - [0041] paragraphs [0047] - [0050] ----- -/- | 1,8 |
| X | | 1,8 |
| A | | 1-8 |
| | | |

Further documents are listed in the continuation of Box C.

See patent family annex.

* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier application or patent but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

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

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

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

"&" document member of the same patent family

| | |
|--|--|
| Date of the actual completion of the international search | Date of mailing of the international search report |
| 6 November 2012 | 12/11/2012 |
| 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 Doyle, Walter |

INTERNATIONAL SEARCH REPORT

| |
|------------------------------|
| International application No |
| PCT/EP2012/068909 |

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|-----------|---|-----------------------|
| X | WO 2005/065283 A2 (WALKER DIGITAL LLC [US]; WALKER JAY S [US]; JORASCH JAMES A [US]; SAMM) 21 July 2005 (2005-07-21) figures 9,10,11,15,16 paragraphs [0119], [0129], [0257] paragraphs [0282], [0330] - [0335] paragraphs [0343] - [0346] ----- | 1-8 |
| X | US 2007/216709 A1 (KOJIMA TAMAKI [JP] ET AL) 20 September 2007 (2007-09-20) figures 15-19, 33 paragraphs [0115] - [0116] paragraphs [0124] - [0125] paragraphs [0200], [0215] ----- | 1,8 |
| X | US 2010/166321 A1 (SAWANT NEELA [IN] ET AL) 1 July 2010 (2010-07-01) paragraphs [0002], [0004], [0007], [0037] ----- | 1,8 |
| A | WO 2006/030878 A1 (CANON KK [JP]; SHIMADA BUNGO [JP]) 23 March 2006 (2006-03-23) the whole document ----- | 1-8 |

INTERNATIONAL SEARCH REPORT

Information on patent family members

| |
|------------------------------|
| International application No |
| PCT/EP2012/068909 |

| Patent document cited in search report | Publication date | Patent family member(s) | | | Publication date |
|--|------------------|--|---|--|--|
| US 2007065045 | A1 22-03-2007 | JP JP US | 4700452 B2 2007080210 A 2007065045 A1 | | 15-06-2011 29-03-2007 22-03-2007 |
| US 2007260639 | A1 08-11-2007 | | NONE | | |
| EP 2120162 | A1 18-11-2009 | EP JP US | 2120162 A1 2009277155 A 2009285493 A1 | | 18-11-2009 26-11-2009 19-11-2009 |
| WO 2005065283 | A2 21-07-2005 | AU CA EP JP JP US US US WO | 2004311841 A1 2554135 A1 1704710 A2 2007520934 A 2010016847 A 2008192129 A1 2011128414 A1 2012242844 A1 2005065283 A2 | | 21-07-2005 21-07-2005 27-09-2006 26-07-2007 21-01-2010 14-08-2008 02-06-2011 27-09-2012 21-07-2005 |
| US 2007216709 | A1 20-09-2007 | CN JP KR US | 101013433 A 2007206919 A 20070079330 A 2007216709 A1 | | 08-08-2007 16-08-2007 06-08-2007 20-09-2007 |
| US 2010166321 | A1 01-07-2010 | | NONE | | |
| WO 2006030878 | A1 23-03-2006 | JP JP US US WO | 4541811 B2 2006081119 A 2009002744 A1 2011255132 A1 2006030878 A1 | | 08-09-2010 23-03-2006 01-01-2009 20-10-2011 23-03-2006 |