

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
4 January 2007 (04.01.2007)

PCT

(10) International Publication Number
WO 2007/000706 A2

(51) International Patent Classification:
G06F 17/30 (2006.01)

(21) International Application Number:
PCT/IB2006/052062

(22) International Filing Date: 23 June 2006 (23.06.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
05105832.9 29 June 2005 (29.06.2005) EP

(71) Applicant (for all designated States except US): **KONINKLIJKE PHILIPS ELECTRONICS N.V.** [NL/NL];
Groenewoudseweg 1, NL-5621 BA Eindhoven (NL).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **VAN GASSEL, Jozef, P.** [NL/NL]; C/o Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL). **MESUT, Ozcan** [NL/NL]; C/o Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).

(74) Agents: **GROENENDAAL, Antonius, W., M.** et al.;
Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).

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

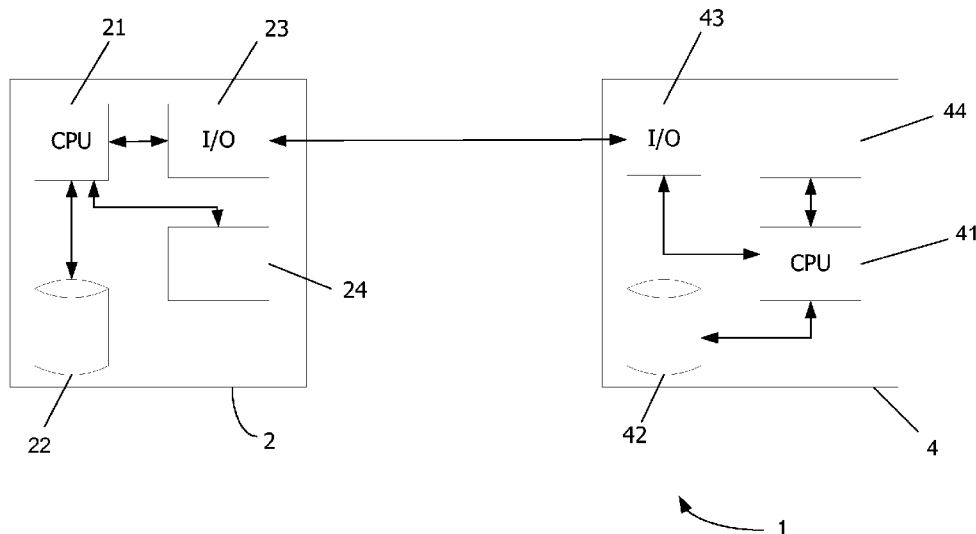
(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: EFFICIENT DATA TRANSFER



(57) Abstract: A method for transferring data blocks from a first device to a second device is described, in which the second device stores signatures relating to respective data blocks stored on the second device, each signature relating to contents of the associated data block. For each data block to be transferred from the first device to the second device a signature for the data block concerned is transferred from the first device to the second device, it is determined whether the received signature is stored on the second device, and if the received signature block is not stored on the second device, the data block is transferred to the second device, and the received signature block and the data block are stored on the second device.

WO 2007/000706 A2

Efficient data transfer

The present invention relates to efficient data transfer, and, in particular but not exclusively, to transfer of large audio/visual (A/V) collections, between a server and a portable A/V device.

5

Mobile storage devices, particularly audio/visual (A/V) devices, are becoming increasingly popular. A/V devices are used to store audio and video data for the user to enjoy on the move. Such multimedia data is typically transferred to the mobile device from a personal computer (PC) or a Home Entertainment Hub (e.g. by ripping, broadcast recording or purchasing). In order to provide a superior user experience, it is desirable that the transfer process between the stationary and mobile devices is as fast as possible.

Very often, data that to be transferred from or to the mobile device is merely an updated version of data that was already available on the target device. This is especially true when embedded metadata (e.g. artist, title and album information in an MP3 file) is updated, or when camcorder edits or play lists have been compiled. Also, within single multimedia files, some level of redundancy can be present (e.g. sequence headers in an MPEG file that do not change over time). Previously-considered transfer techniques make only limited use of such characteristics to enhance the speed of transfer of data.

Another problem occurs when so-called virtual play lists are being used. A virtual play list consists of a sequence of clips defined by in- and out- points of one or more audio / visual files. It is very likely that multiple play lists refer to the same (section within a) file (for instance in the case of commercial skipping of broadcast recordings or different camcorder edits of the same original content). This is illustrated in Figure 1 of the accompanying drawings.

When a virtual play list is transferred between devices as a single multimedia stream, valuable information that can be used to identify redundancy (i.e. duplicate data fragments) typically is lost. In that case, the complete data sequence must be transferred (as one file) to the target device. This leads to inefficient use of the (limited) storage capacity of the mobile device due to duplication of data, and also to increased transfer time.

According to one aspect of the present invention, there is provided a method for transferring data blocks from a first device to a second device, the method comprising:

5 storing retained signatures relating to respective data blocks stored on the second device, each stored signature relating to contents of the related data block;
for each data block to be transferred from the first device to the second device;
determining whether a transfer signature relating to the data block concerned corresponds to a retained signature relating to a data block stored on the second device; and
10 if the transfer signature does not correspond to a retained signature, transferring the data block to the second device, and storing the transfer signature and the data block on the second device.

According to another aspect of the present invention, there is provided a method for receiving data blocks from a first device at a second device, the method
15 comprising:

storing retained signatures relating to respective data blocks stored on the second device, each stored signature relating to the contents of the related data block;
for each data block to be received from the first device:
receiving a transfer signature for the data block concerned;
20 determining whether the transfer signature corresponds to a retained signature relating to a data block stored on the second device; and
if the transfer signature does not correspond to a retained signature, receiving the data block from the first device, storing data block on the second device, and storing the transfer signature as a retained signature.

25 According to another aspect of the present invention, there is provided a method for transmitting data blocks from a first device to a second device, the method comprising:

30 for each data block to be transmitted:
transmitting, from a first device, a transfer signature for the data block concerned to a second device;

generating a comparison signal indicative as to whether the transfer signature corresponds to a retained signature relating to a data block stored on the second device; and

5 if the comparison signal indicates that the transfer signature does not correspond to a retained signature, transmitting the data block from the first device to the second device.

According to another aspect of the present invention, there is provided a system for transferring data blocks between first and second electronic devices, the system comprising:

10 comparison means operable to compare a transfer signature relating to a data block to be transferred with a retained signature relating to a data block stored on a second device, and to generate a comparison signal indicative as to whether such a transfer signature corresponds to such a retained signature; and

15 a controller operable to initiate transfer of the data block from a first device to a second device in dependence upon the comparison signal.

According to another aspect of the present invention, there is provided an electronic device comprising:

a storage device operable to store data blocks and corresponding respective retained signatures;

20 an input unit operable to receive, from a transmitting device a transfer signature relating to a data block to be transferred;

a comparison unit operable to compare such a transfer signature with stored retained signatures; and

25 a controller operable, if the transfer signature does not match a retained signature, to cause the storage device to store the transfer signature therein as a retained signature and to store the data block therein.

According to another aspect of the present invention, there is provided an electronic device comprising:

a storage device operable to store data blocks;

30 an output unit operable to output, to a receiver device, a transfer signature relating to a data block for transfer to the receiver device;

an input unit operable to receive a comparison signal, the comparison signal being indicative as to whether the transfer signature corresponds to a retained signature relating to a data block stored on the receiver device; and

a controller operable, if such a comparison signal indicates that the transfer signature does not correspond to such a retained signature, to transfer the data block to the receiver device.

In one example, at least one retained signature is stored on the second device.

5 At least one transfer signature may be stored on the first device.

In another example, at least one transfer signature is calculated when a related data block is to be transferred from the first device to the second device. At least one retained signature may be calculated when a data block is to be transferred from the first device to the second device.

10 Respective signatures may be stored for all data blocks for all files on the second device.

The data blocks may be defined in dependence upon the format of the associated file. For example, when the format is MPEG, the data blocks may be defined by MPEG start codes.

15 According to another aspect of the present invention, there is provided a method for synchronising data between first and second electronic devices, the method comprising storing a signature for each data block stored on the first and second devices, comparing stored signatures with a signature of a data block to be synchronised between the first and second devices, and only synchronising the data block if the signature is not stored
20 on each of the first and second devices.

According to another aspect of the present invention, there is provided a device for synchronising data blocks between first and second electronic devices, the device comprising a comparison unit for comparing signatures relating to data blocks stored on first and second electronic devices, and operable to synchronise data blocks between the first and
25 second devices when such signatures are not stored on each of the first and second devices.

Embodiments of the present invention serve to increase transfer speed of data between devices, by avoiding unnecessary transfers of data blocks that are already present on the target device. Embodiments can make use of defined data block boundaries, and, in particular, to boundaries defined by an appropriate data compression and storage standard
30 such as MPEG. Advantageously, the data blocks are ordered in a hierarchy, such that a super-block higher in the hierarchy contains sub-blocks that are lower in the hierarchy. Each of the super- and sub-blocks have respective transfer signatures and respective retain signatures. Comparing these signatures is performed top-down in the hierarchy. So first the highest level retain signature is compared to the corresponding transfer signature. If these are

identical, no data block needs to be transferred at all. If the signatures differ, the signatures on the next level are compared. This way the hierarchy is recursively traversed until all data blocks match.

5

Fig. 1 illustrates a server and mobile device according to one aspect of the present invention;

Fig. 2 is a flowchart illustrating steps in a method embodying an aspect of the present invention; and

10

Fig. 3 illustrates data transfer according to another aspect of the present invention.

Figure 1 illustrates a system 1 embodying one aspect of the present invention.

15

The system 1 includes a server 2 and a mobile device 4. The server 2 comprises a CPU (central processing unit) 21, a storage device 22, input/output interface device 23, and MMI (man machine interface) devices 24. The MMI devices 24 can be provided, in known manner, by a keyboard, a mouse, a display device and loud speakers. The server 2, the mobile device 4, and the method of transferring data (both transmitting and receiving) relate to different respective aspects of the present invention.

20

The server 2 operates as a stationary multimedia server, and multimedia files are stored on the storage device 22 under the control of the CPU 21.

The mobile device 4 includes a mobile CPU 41, a mobile storage device 42, a mobile I/O interface 43 and mobile MMI devices 44. The mobile device can typically be an MP3 player, handheld video player, mobile phone or any suitable device. The mobile storage device 42 operates to store multimedia files for use by a user of the device on the move. As is well known, data is transferred between the mobile device 4 and the server 2 in order to update the files stored on the mobile storage device 42. In such a way, a user is able to download a selected number of audio visual files for use later away from the server.

25

30

Embodiments of the present invention are concerned with the transfer of data between the server 2 and the mobile device 4, and in particular are concerned with efficient transfer of data between those two devices. The server 2 can be considered a first device, and the mobile device can be considered a second device, such that embodiments of the present

invention are concerned with transfer, or synchronisation, of data between a first and a second device.

Known techniques, such as the “rsync” tool as used in Linux™ / Unix™ allow two sets of data to be synchronized without having prior knowledge of the content stored on the other side of a link. This is achieved by transferring two types of signature (a weak and a strong signature) relating to blocks of data between the two devices. A signature is a short representative identifier based on the original data (e.g. CRC32 which 32 bits or MD5, for example see RFC 1321: The MD5 Message-Digest Algorithm. R. Rivest. April 1992.

<http://www.ietf.org/rfc/rfc1321.txt>) or SHA which is 128 bits [RFC 3174: US Secure Hash

Algorithm 1 (SHA1 D. Eastlake 3rd, P. Jones. September 2001,

<http://www.ietf.org/rfc/rfc3174.txt>). The present invention exploits knowledge of the structure of the stored data in such a way as to improve on the previously-considered techniques. Such improvements result in increased synchronization speed and elimination of redundancy on the (relatively small) storage device 42 of the mobile device 4.

Audio/video data is structured in small blocks, due to the nature of audio / video formats (e.g. MPEG). Examples of such blocks are Access Units (frames in a video stream) and blocks separated by start codes (for example slices or sequence headers in MPEG video). Editing of video and audio will often take place on such elementary block boundaries (e.g. on a frame by frame basis). Typically these block boundaries coincide with in- and out points of audio / video edits. The data blocks can be random in size (this is dictated by the structure of the stream). Usually starts end endings of such blocks in an encoded stream are signalled by so-called start codes (e.g. in MPEG a sequence starting with the hexadecimal sequence 0x00 0x00 0x01). This allows for very low-complexity block boundary detection.

In a method of the present invention, these block boundaries are used to avoid the need to calculate signatures for all possible byte boundaries of the file. Such simplification means that a single signature can be used. Using single signatures enables the signatures to be stored in a table in the mobile storage device 2 of the mobile device 4. Such static boundaries and storage means that the signatures need only be calculated once. Since A/V devices typically have an embedded database that stores Characteristic Point Information (e.g. for trick play purposes), this will only add one extra column with signatures.

When Access Unit (frame) boundaries as suggested above are used, the overhead for storing the block signatures can be calculated as follows. Assuming 128 bit signatures (e.g. MD5) and a video frame rate of 25 fps, this leads to $25 \times 128 = 3.2\text{kbps}$ (or

400 Bps) for the signature stream. Hence, the extra storage capacity required for storage of these signatures in case of a Mbps video stream is less than 1% compared to the size of the original video file (assuming a video bit-rate of 400 kbps).

In case blocks (or access units) are used multiple times in the same file or over different files there are two different options:

1. A block reference counting mechanism is required. This administration information is preferably part of the metadata that is stored in the metadata database or the file system, and keeps track of at which location in particular files the stored blocks are being used.

2. The data is duplicated in the target file at the receiving end. The original block does not need to be sent from the server 2 to the mobile device 4 but can be copied from the local storage and is appended in the new file.

Although the referencing mechanism leads to much more efficient usage of the available storage space, it is also more complicated to implement. At minimum it leads to extra functionality in the database but it might also require extra features in the file system implementation.

One practical embodiment of the present invention will now be described with reference to the flow chart of Figure 2. Transfer of data blocks from the server 2 to the mobile device 4 is performed using the following steps, as illustrated in Figure 2:

A. The server 2 sends the signature of the first data block of the file to be transferred to the mobile device 4 (if the corresponding block is repeated at multiple positions in the file being transferred, then these subsequent indexes of such occurrences can be sent as well).

B1/B2. The mobile device 4 checks whether the incoming signature occurs in the look-up table of its database / file system to check whether it already stores this data block On the storage device 42. The mobile device 4 determines whether the signature is stored on the storage device 42 of the mobile device 4, by comparing the incoming signature with stored signatures.

C. If YES: the mobile device 4 sends a comparison signal to the server 2 in order to indicate that the data block is not required since the signature therefor already is stored on the storage device 42. A reference counter referring to the already existing block is increased or the block is copied locally in a later stage (note that the former may require special provisions in the file system) to the new file.

D. If NOT: the mobile device sends a comparison signal to the server 2. the comparison is indicative of the fact that the signature does not match a stored signature. Accordingly, the data block is transferred in its entirety from the server 2 to the mobile device where it will be stored on the storage device 42. Furthermore its associated signature is added to the local database / file system tables.

E. The sequence repeats itself from A until the complete file is transferred. These steps are repeated until the entire object (for example, file, play list or virtual edit) is transferred to the mobile device 4 from the server 2.

In a possible extension of this technique, the efficiency of data transfer could be improved further by organising the blocks and the associated signatures in a hierarchical way. MPEG streams are, for example, divided into Groups of Pictures (or GOP), Pictures and slices (part of a single picture). When sending the data blocks to detect differences, the transfer technique may start by sending a signature of the higher hierarchical level (e.g. GOPs) and only when differences have been detected on the GOP level would lower levels be searched. Such a technique would require fewer control messages to be sent over the line between the server 2 and the mobile device 4. Another advantage is that the search converges logarithmically. Many more hierarchical levels are conceivable (e.g. scenes, macroblocks, etc. etc.).

The blocks may form one or more hierarchies depending on the semantic meaning of the block: chapter, playlist, scene, GOP, etcetera. In case of comparing signatures, the method starts by comparing the signature of the aggregated higher level block. If this is identical than no copying needs to be done. However, if they differ, than all signatures of all subblocks (that are aggregated by the superblock) need to be examined in a similar way. This is a recursive process that needs to be carried out until we end up at the leaves of the tree.

In order to increase the possibility of data blocks already being available at the receiving end, one possibility is to de-multiplex the streams into elementary audio, video and private streams before the file is split up in the individual data blocks. Such an approach enables even greater efficiency, since, for example, only the audio track may have changed, and so the video track need not be downloaded.

It should be noted that the above-mentioned embodiments illustrate rather than limit the invention, and that those skilled in the art will be capable of designing many alternative embodiments without departing from the scope of the invention as defined by the appended claims. In the claims, any reference signs placed in parentheses shall not be

construed as limiting the claims. The words "comprising" and "comprises", and the like, do not exclude the presence of elements or steps other than those listed in any claim or the specification as a whole. The singular reference of an element does not exclude the plural reference of such elements and vice-versa. The invention may be implemented by means of

5 hardware comprising several distinct elements, and by means of a suitably programmed computer, if appropriate. In a 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.

CLAIMS:

1. A method for transferring data blocks from a first device (2) to a second device (4), the method comprising:

storing retained signatures relating to respective data blocks stored on the second device (4), each stored signature relating to contents of the related data block;

5 for each data block to be transferred from the first device (2) to the second device (4);

determining whether a transfer signature relating to the data block concerned corresponds to a retained signature relating to a data block stored on the second device (4);
and

10 if the transfer signature does not correspond to a retained signature,
transferring the data block to the second device (4), and storing the transfer signature and the data block on the second device (4).

2. A method for receiving data blocks from a first device (2) at a second device (4), the method comprising:

15 storing retained signatures relating to respective data blocks stored on the second device (4), each stored signature relating to the contents of the related data block;

for each data block to be received from the first device (2):

receiving a transfer signature for the data block concerned;

20 determining whether the transfer signature corresponds to a retained signature relating to a data block stored on the second device (4); and

if the transfer signature does not correspond to a retained signature,
receiving the data block from the first device (2), storing data block on the second device (4),
and storing the transfer signature as a retained signature.

25 3. A method for transmitting data blocks from a first device (2) to a second device (4), the method comprising:

for each data block to be transmitted:

transmitting, from a first device (2), a transfer signature for the data block concerned to a second device (4);

generating a comparison signal indicative as to whether the transfer signature corresponds to a retained signature relating to a data block stored on the second

5 device (4); and

if the comparison signal indicates that the transfer signature does not correspond to a retained signature, transmitting the data block from the first device (2) to the second device (4).

10 4. A method as claimed in claim 1, 2 or 3, wherein at least one retained signature is stored on the second device.

5. A method as claimed in any one of the preceding claims, wherein at least one transfer signature is stored on the first device.

15

6. A method as claimed in any one of claims 1 to 4, wherein at least one transfer signature is calculated when a related data block is to be transferred from the first device to the second device.

20 7. A method as claimed in any one of the preceding claims, wherein at least one retained signature is calculated when a data block is to be transferred from the first device to the second device.

8. A method as claimed in any one of the preceding claims, comprising storing
25 respective signatures for all data blocks for all files on the second device (4).

9. A method as claimed in any one of the preceding claims, wherein the data blocks are defined in dependence upon the format of the associated file.

30 10. A method as claimed in claim 9, wherein the format is MPEG, and the data blocks are defined by MPEG start codes.

11. A computer program product comprising computer program code items for carrying out a method as claimed in any one of claims 1 to 10.

12. A system for transferring data blocks between first and second electronic devices, the system comprising:

comparison means (21, 41) operable to compare a transfer signature relating to a data block to be transferred with a retained signature relating to a data block stored on a second device, and to generate a comparison signal indicative as to whether such a transfer signature corresponds to such a retained signature; and

a controller (21, 41) operable to initiate transfer of the data block from a first device to a second device in dependence upon the comparison signal.

13. An electronic device comprising:

a storage device (42) operable to store data blocks and corresponding respective retained signatures;

an input unit (43, 41) operable to receive, from a transmitting device (2) a transfer signature relating to a data block to be transferred;

a comparison unit (41) operable to compare such a transfer signature with stored retained signatures; and

a controller (41) operable, if the transfer signature does not match a retained signature, to cause the storage device (42) to store the transfer signature therein as a retained signature and to store the data block therein.

14. A device as claimed in claim 13, wherein the controller is operable to cause the storage device to store therein respective retained signatures for all data blocks for all files on the second device.

15. A device as claimed in claim 13 or 14 wherein the controller (21) is operable to output a comparison signal to the transmitting device (2) indicative as to whether the transfer signature corresponds to a retained signature.

16. A device as claimed in claim 15, wherein, if the comparison signal indicates that the transfer signature does not match a retained signature, the input unit (43, 41) is operable to receive the data block corresponding to the incoming signature from the transmitting device.

17. A device as claimed in any one of claims 19 to 22, wherein the controller (41) is operable to calculate at least one retained signature value when a data block is to be received from a transmitting device (2).

5 18. An electronic device comprising:

a storage device (22) operable to store data blocks;

an output unit (21, 23) operable to output, to a receiver device (4), a transfer signature relating to a data block for transfer to the receiver device (4);

an input unit (23, 21) operable to receive a comparison signal, the comparison signal being

10 indicative as to whether the transfer signature corresponds to a retained signature relating to a data block stored on the receiver device (4); and

a controller (21) operable, if such a comparison signal indicates that the transfer signature does not correspond to such a retained signature, to transfer the data block to the receiver device (4).

15 19. A device as claimed in claim 18, wherein the controller (21) is operable to calculate at least one transfer signature when a data block is to be transferred.

20. A device as claimed in claim 18, wherein the storage device (22) is operable to
20 store transfer signatures for data blocks stored on the storage device (22).

21. A device as claimed in any one of claims 13 to 20, wherein the data blocks are defined in dependence upon the format of the corresponding file.

25 22. A device as claimed in claim 13, wherein the format is MPEG, and the data blocks are defined by MPEG start codes.

23. A method for synchronising data between first and second electronic devices, the method comprising storing a signature for each data block stored on the first and second
30 devices, comparing stored signatures with a signature of a data block to be synchronised between the first and second devices, and only synchronising the data block if the signature is not stored on each of the first and second devices.

24. A device for synchronising data blocks between first and second electronic devices, the device comprising a comparison unit for comparing signatures relating to data blocks stored on first and second electronic devices, and operable to synchronise data blocks between the first and second devices when such signatures are not stored on each of the first and second devices.
- 5

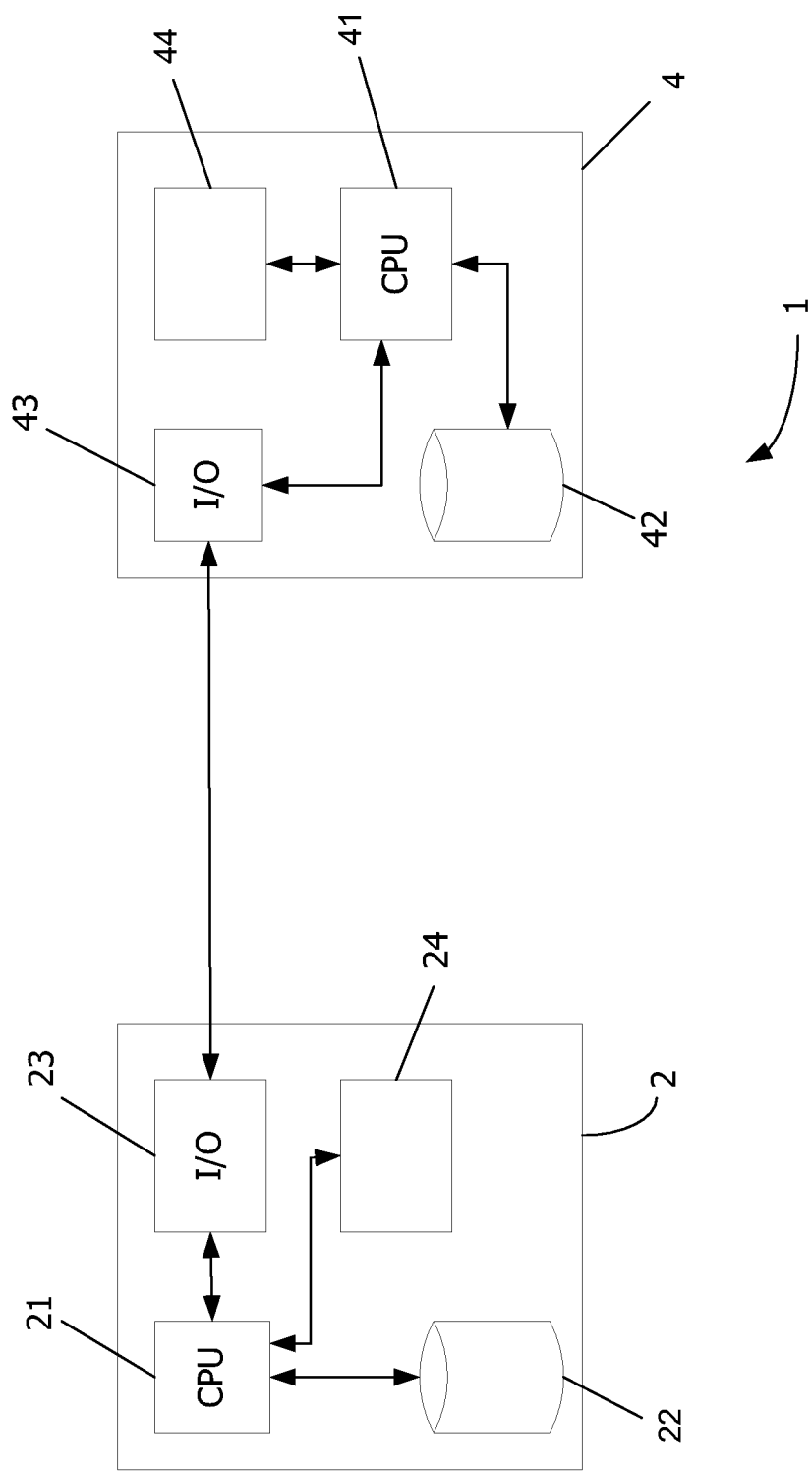


FIG. 1

2/3

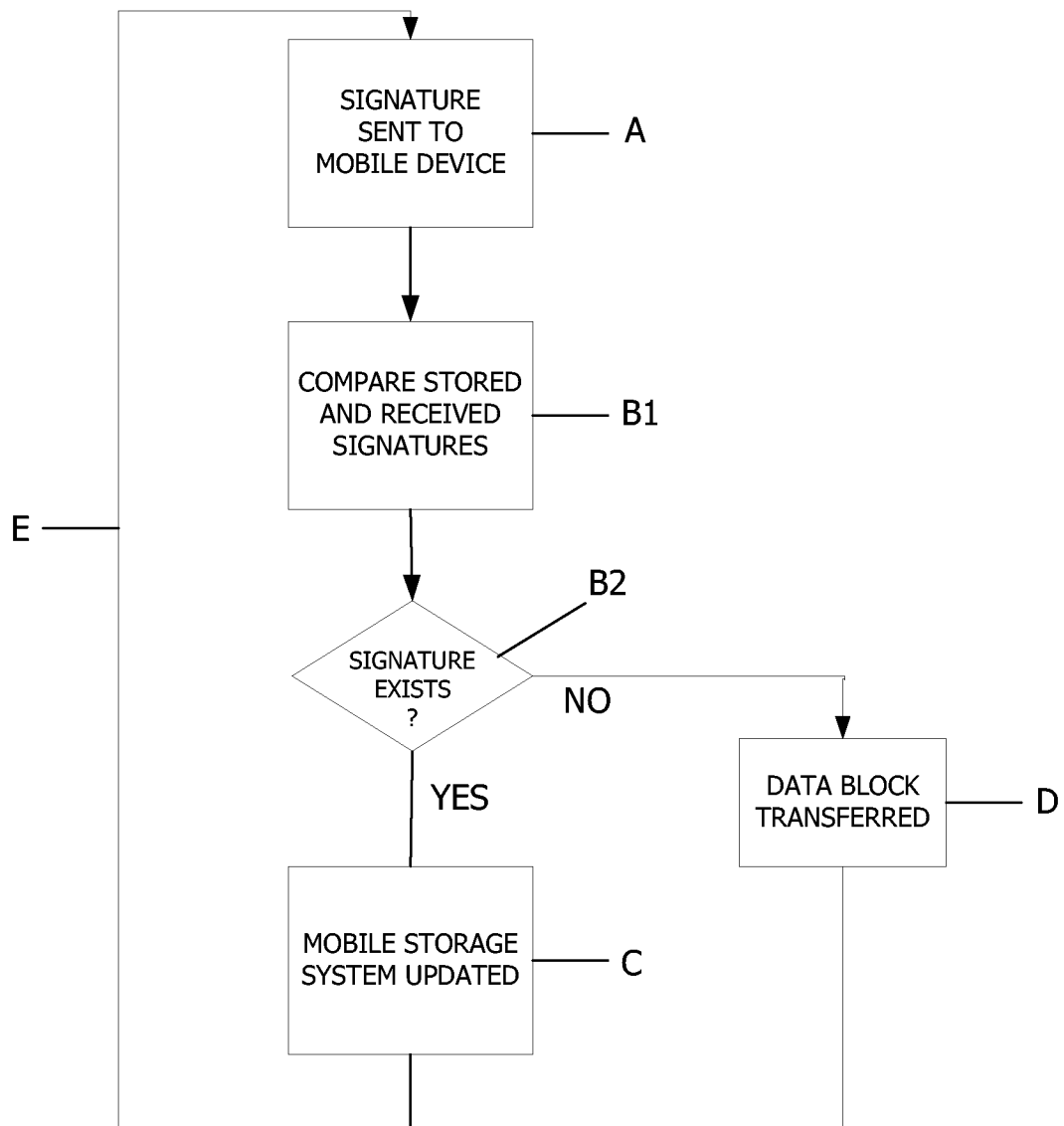


FIG. 2

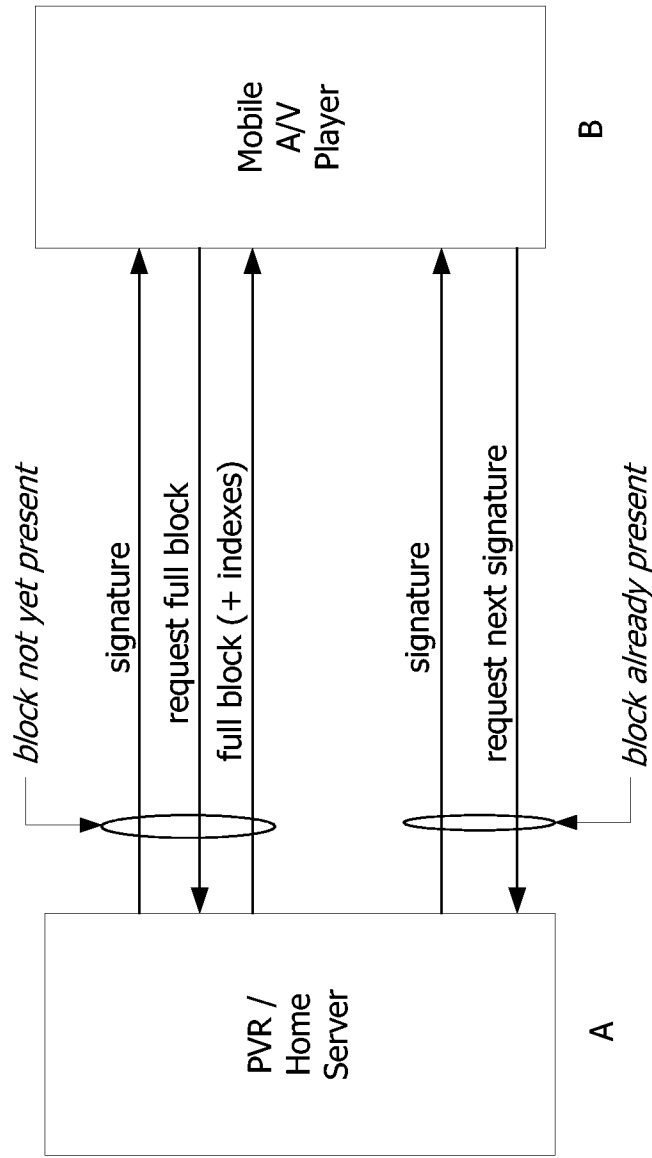


FIG. 3