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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
20 August 2015 (20.08.2015)

WIPO | PCT

(10) International Publication Number
WO 2015/121715 A1

(51) International Patent Classification:
G06K 9/34 (2006.01)

(21) International Application Number:
PCT/IB2014/063970

(22) International Filing Date:
19 August 2014 (19.08.2014)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
2014106042 14 February 2014 (14.02.2014) RU

(71) Applicant *(for all designated States except AM, AZ, BY, KG, KZ, RU, TJ, TM, ZW)*: **YANDEX EUROPE AG** [CH/CH]; Werftestrasse 4, CH-6005 Luzern (CH).

(71) Applicant *(for AM, AZ, BY, KG, KZ, RU, TJ, TM only)*: **YANDEX LLC** [RU/RU]; 16 Leo Tolstoy St., Moscow, 119021 (RU).

(71) Applicant *(for ZW only)*: **YANDEX INC.** [US/US]; 299 South California Ave., Palo Alto, California 94306 (US).

(72) Inventors: **POPELO, Lidia Vladimirovna**; Bolshaya Dorogomilovskaya street, House 16, Apt. 42, Moscow, 121248 (RU). **CHUPROV, Dmitry Vladimirovich**; Govorova street, House 26, Apt. 176, Odintsovo, Moscow Region, 143005 (RU).

(74) Agent: **CUTLER, Jonathan D.**; BCF LLP, 1100 Rene-Levesque Blvd West, Suite 2500, Montreal, Québec H3B 5C9 (CA).

(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, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

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

(54) Title: METHOD OF AND SYSTEM FOR GENERATING METADATA

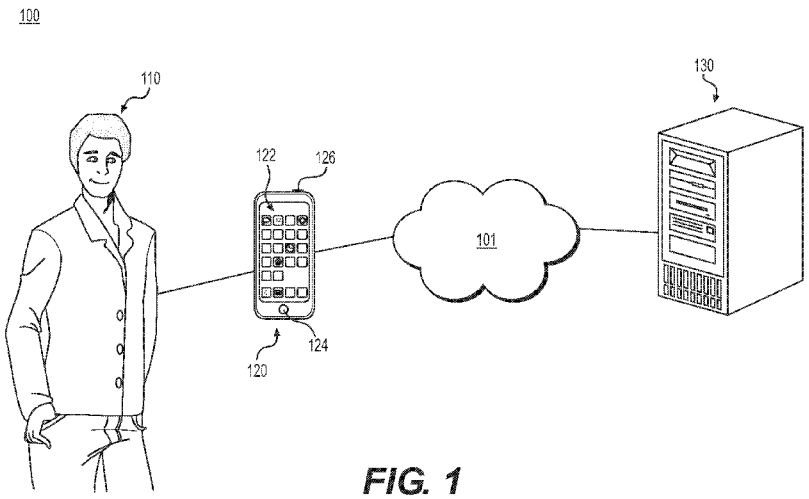

FIG. 1

(57) Abstract: Method of and system for generating image metadata, comprising, at an electronic device: receiving an indication of text to be included in an image, the text comprising at least one character, each character being encoded according to a character encoding; generating the image based at least in part on the text, the image including an image representation of the text; generating the image metadata based at least in part on the text; and associating the image metadata with the image. Method of and system for generating video metadata. Method of and system for generating audio metadata.

WO 2015/121715 A1

METHOD OF AND SYSTEM FOR GENERATING METADATA

CROSS-REFERENCE

[01] The present application claims convention priority to Russian Patent Application No. 2014106042, filed February 14, 2014, entitled "METHOD OF AND SYSTEM FOR GENERATING METADATA" which is incorporated by reference herein in its entirety.

FIELD

[02] The present technology relates to methods and systems for generating metadata in respect of images, videos, and/or audio clips.

BACKGROUND

[03] Digital content creation has become increasingly affordable, accessible, and popular in recent years, as digital cameras, scanners, and graphics software have become commonplace. As a result, the number of digital content files created has increased, and so too has the need for techniques to organize, index, and search through digital content collections, such as collections of images, videos, or audio clips.

[04] The ability to associate various metadata with images, videos, and audio clips is essential in this regard. For example, image metadata may include a title, a description, image size, camera settings, authorship and/or copyright information, creation and/or editing date and time, a thumbnail version of the image, and one or more descriptive keywords (sometimes called "tags"). Because these metadata are generally stored as computer-readable text, it is a simple matter for computers to index and/or search through the information they contain, thus enabling digital content with particular features described in the metadata to be quickly and efficiently identified from among many items in a large collection.

[05] Some digital images include an image representation of text. For example, a photograph of a movie theatre may include a movie title (e.g. "Casablanca") displayed on the theatre's marquee. While such text is often easily identifiable by a human observer, a computer may only identify an image representation of text as being text *per se* by performing an analysis of the image representation of the text, known as optical character recognition (OCR). An OCR algorithm analyzes images to detect visual patterns representative of text characters and then outputs those text characters in a definite, machine-encoded form known

as a character encoding, normally a standard character encoding such as ANSI, ASCII or Unicode. The resulting text may then be unambiguously interpreted and manipulated by computer systems.

[06] Online service Evernote™ uses OCR technology to identify text in an image uploaded by a user and associates metadata including the identified text with the image. The metadata associated with the image may then be indexed and/or searched, thus allowing the user (or another user) to find the image via a text-based search query including elements of the text as search terms. With reference to the aforementioned example, the photograph of the movie theatre may be uploaded to Evernote™, which may identify the movie title "Casablanca" in the image using OCR and consequently include the text string "Casablanca" in the image's metadata. A subsequent search for "Casablanca" may yield the image as a search result.

SUMMARY

[07] Inventors have developed embodiments of the present technology based on their appreciation of at least one shortcoming of the prior art. Notably, although generating image metadata using OCR in the manner of Evernote™ as described above may be effective in some cases, in other cases it is inconvenient due to the computational intensity and potential inaccuracy of OCR.

[08] The present technology arises from the inventors' recognition that in some cases, image metadata associated with an image may be automatically generated as part of the process of generating and/or modifying that image. More specifically, when an image is generated or modified so as to include an image representation of known text, there is an opportunity to efficiently and reliably generate metadata based on that text, rather than later performing OCR to imperfectly recover the text from its image representation in the generated image.

[09] Thus, in one aspect, implementations of the present technology provide a method of generating image metadata, the method comprising, at an electronic device:

- receiving an indication of text to be included in an image, the text comprising at least one character, each character being encoded according to a character encoding;
- generating the image based at least in part on the text, the image including an image representation of the text;

- generating the image metadata based at least in part on the text; and
- associating the image metadata with the image.

[10] In some implementations, the image to be generated is not an entirely new image, but a previously existing image modified to include an image representation of the text. Thus, in some implementations, receiving the indication of the text to be included in the image comprises receiving an indication of text with which to modify an unmodified image, and generating the image based at least in part on the text comprises generating the image based at least in part on the text and the unmodified image. Any metadata associated with the previously existing image may be preserved, updated, or otherwise processed when generating the image metadata in respect of the generated image. Thus, in some further implementations, generating the image metadata based at least in part on the text comprises generating the image metadata based at least in part on the text and existing image metadata associated with the unmodified image.

[11] In some such implementations, a screenshot of the electronic device's display may first be taken before being modified with the text to generate the image. For example, a user of a smartphone may take a screenshot while playing a game of Tetris™ and then provide text to be overlaid on the image. Thus, in some implementations, the unmodified image comprises a screenshot image of a display of the electronic device, and the method further comprises, before generating the image: receiving an instruction to generate the screenshot image from a user of the electronic device; and generating the screenshot image as the unmodified image. In other implementations, the screenshot image is that of a display of a device other than the electronic device. Thus, in some implementations, the unmodified image comprises a screenshot image of a display of a second electronic device in communication with the electronic device via a communications network; and further comprising, before generating the image, receiving the screenshot image from the second digital electronic device via the communications network.

[12] In other implementations, a digital photograph may first be taken before being modified with the text to generate the image. In such implementations, the unmodified image comprises a digital photograph, and the method further comprises, before generating the image: receiving an instruction to capture the digital photograph from a user of the electronic device; and capturing the digital photograph via a camera coupled to the electronic device as the unmodified image.

[13] In some implementations, in response to a user instructing the device to take a screenshot of a display of the electronic device, some or all of the text displayed on the display may be captured as the text to be used in the generation of the image. Thus, in some implementations, receiving the indication of the text to be included in the image comprises: receiving an instruction to generate a screenshot image of a display of the electronic device from a user of the electronic device; and capturing as the text at least some of the text displayed on the display. For example, this may be accomplished by requesting the displayed text from the one or more applications causing the text to be displayed on the display. An image including an image representation of the captured text may then be generated along with image metadata based on the captured text to be associated with the image. In some implementations, the image generated may actually be a screenshot of the display. In such implementations, generating the image based at least in part on the text comprises generating the screenshot image as the image. Thus, some implementations of the present technology allow for generation of screenshot images and association of metadata including text displayed in the screenshot images with those images, without having to perform OCR on the screenshot images. In other implementations, the image generated is not a screenshot image, though it includes an image representation of text that was displayed on the display when the instruction to take the screenshot was received. Thus, in other implementations, generating the image based at least in part on the text is generating the image based at least in part on the text without generating the screenshot image.

[14] In another aspect, various implementations of the present technology provide a method of generating image metadata, the method comprising, at an electronic device:

- receiving an indication of text with which to modify an image, the text comprising at least one character, each character being encoded according to a character encoding;
- modifying the image based at least in part on the text to include an image representation of the text;
- generating the image metadata based at least in part on the text; and
- associating the image metadata with the image.

[15] In another aspect, various implementations of the present technology provide a method of augmenting image metadata associated with an image, the method comprising, at an electronic device:

- receiving an indication of text with which to modify the image, the text comprising at least one character, each character being encoded according to a character encoding;
- modifying the image based at least in part on the text to include an image representation of the text;
- generating additional image metadata based at least in part on the text; and
- associating the additional image metadata with the image by adding the additional image metadata to the image metadata associated with the image.

[16] The image and the image metadata may be associated in a variety of ways. Some image file types (e.g. JPEG, TIFF, PNG, and others) allow metadata to be stored in the file along with the image content. Thus, in some implementations, associating the image metadata with the image comprises writing an image file including the image and the image metadata to a non-transitory computer-readable medium. In other implementations, image metadata associated with the image may be stored separately from the digital image file, and an association between the two may be maintained in a database. Thus, in other implementations, associating the image metadata with the image comprises at least one of creating and modifying an entry in a database, the entry including an indication of the image and an indication of the image metadata. In still other implementations, the image and the image metadata may be associated by virtue of being referenced in a same communication, whether a low-level communication such as a single TCP or UDP packet, or a higher level communication such as an email or a transmission of an HTML or XML document. In such implementations, associating the image metadata with the image includes sending a communication including an indication of the image and an indication of the image metadata via a communications network.

[17] In another aspect, various implementations of the present technology provide a method of generating video metadata, the method comprising, at an electronic device:

- receiving an indication of text to be included in at least one frame of a video, the text comprising at least one character, each character being encoded according to a character encoding;
- generating the video comprising the at least one frame based at least in part on the text, the at least one frame including an image representation of the text;
- generating the video metadata based at least in part on the text; and
- associating the video metadata with the video.

[18] The video and the video metadata may be associated in a variety of ways. Some video file types (e.g. various types compliant with the MPEG-7 standard) allow metadata to be stored in the file along with the video content. Thus, in some implementations, associating the video metadata with the video comprises writing a video file including the video and the video metadata to a non-transitory computer-readable medium. In other implementations, video metadata associated with the video may be stored separately from the video file, and an association between the two may be maintained in a database. Thus, in other implementations, associating the video metadata with the video comprises at least one of creating and modifying an entry in a database, the entry including an indication of the video and an indication of the video metadata. In still other implementations, the video and the video metadata may be associated by virtue of being referenced in a same communication, whether a low-level communication such as a single TCP or UDP packet, or a higher level communication such as an email or a transmission of an HTML or XML document. In such implementations, associating the video metadata with the video includes sending a communication including an indication of the video and an indication of the video metadata via a communications network.

[19] In another aspect, various implementations of the present technology provide a method of generating audio metadata, the method comprising, at an electronic device:

- receiving an indication of text to be included in an audio clip, the text comprising at least one character, each character being encoded according to a character encoding;
- generating the audio clip based at least in part on the text, the audio clip including an audio representation of the text;
- generating the audio metadata based at least in part on the text; and
- associating the audio metadata with the audio clip.

[20] The audio clip and the audio metadata may be associated in a variety of ways. Some audio file types (e.g. various types compliant with AES metadata standards, MP3 files with ID3 tags) allow metadata to be stored in the file along with the audio clip. Thus, in some implementations, associating the audio metadata with the audio clip comprises writing an audio file including the audio clip and the audio metadata to a non-transitory computer-readable medium. In other implementations, audio metadata associated with the audio clip may be stored separately from the audio file, and an association between the two may be maintained in a database. Thus, in other implementations, associating the audio metadata

with the audio clip comprises at least one of creating and modifying an entry in a database, the entry including an indication of the audio clip and an indication of the audio metadata. In still other implementations, the audio clip and the audio metadata may be associated by virtue of being referenced in a same communication, whether a low-level communication such as a single TCP or UDP packet, or a higher level communication such as an email or a transmission of an HTML or XML document. In such implementations, associating the audio metadata with the audio clip includes sending a communication including an indication of the audio clip and an indication of the audio metadata via a communications network.

[21] In various implementations of above aspects, the metadata generated may include any number of fields. In some implementations, the metadata includes a text field, and generating the metadata includes populating the text field with at least some of the text. In some cases, the character encoding of the text may differ from that to be used in the metadata, requiring translation of the text from one encoding to the other. For example, the text may be encoded according to the ASCII standard and the image metadata may be encoded according to the Unicode standard. Thus, in some implementations, the character encoding is a first character encoding, the text field conforms to a second character encoding other than the first character encoding, and populating the text field with at least some of the text comprises translating the at least some of the text from the first character encoding to the second character encoding.

[22] In various implementations of above aspects, the indication of the text is received after a user inputs the text using a keyboard, touchscreen, or other tactile device. A user may also input text using a microphone coupled to a voice recognition component implemented in hardware, software, or a combination of hardware and software. Thus, in some implementations, receiving the indication of the text comprises receiving the indication of the text from a user of the electronic device via the electronic device. As non-limiting examples, the user may type text using a physical keyboard or a virtual keyboard (perhaps displayed on a touchscreen), or speak text into a microphone to be interpreted by a voice recognition component. Thus, in some such implementations, receiving the indication of the text from the user of the electronic device comprises receiving the indication of the text via at least one of a physical keyboard of the electronic device, a virtual keyboard of the electronic device, and a voice recognition component coupled to a microphone of the electronic device. The physical keyboard, virtual keyboard, and/or microphone of the electronic device may (but need not) be part of the electronic device itself, so long as they are coupled to the electronic device – e.g.

via a wired or wireless direct link or communications network – so as to be able to relay information to the electronic device based on inputs they receive from the user.

[23] In other various implementations of above aspects, the text may be remotely communicated to the electronic device from another device via a direct link or via a communications network. Thus, in some implementations, receiving the indication of the text comprises receiving the indication of the text from a second electronic device in communication with the electronic device via at least one of a direct link and a communications network. Any suitable direct link or communications network may be used, whether wired, wireless, or a combination of wired and wireless. Suitable examples include universal serial bus (USB) cables, Ethernet cables, TOSLINK fiber optic cables, coaxial cables, IrDA wireless links, Bluetooth™ wireless links, Wi-Fi Direct™ wireless links, local area networks, cellular networks, and the Internet, though any other means of communicating an indication of text may be employed.

[24] In more general terms, the present technology allows for metadata to be generated and associated with digital content in whatever form it may take (including images, videos, audio, and other forms) as part of the process of generating and/or modifying that content to include a non-textual representation of text – that is, a representation of the text in the medium of the digital content itself. Thus, in one aspect, various implementations of the present technology provide a method of generating metadata, the method comprising, at an electronic device:

- receiving an indication of text to be represented in digital content, the text comprising at least one character, each character being encoded according to a character encoding;
- generating the digital content based at least in part on the text, the digital content including a non-textual representation of the text in at least a portion of the digital content;
- generating the metadata based at least in part on the text; and
- associating the metadata with the digital content.

[25] As described above, various approaches of associating the metadata with the digital content may be taken, such as storing them in a same file, including a reference to one in the other, including a reference to both in a same file or database entry, sending a communication including an indication of the metadata and an indication of the digital content via a communications network, or any other suitable means.

[26] In other aspects, various implementations of the present technology provide an electronic device suitable for carrying out above-described methods.

[27] In the context of the present specification, an "electronic device" is any hardware and/or software appropriate to the relevant task at hand. Thus, some (non-limiting) examples of electronic devices include computers (servers, desktops, laptops, netbooks, etc.), smartphones, and tablets, as well as network equipment such as routers, switches, and gateways.

[28] In the context of the present specification, a "display" of an electronic device is any electronic component capable of displaying an image to a user of the electronic device. Non-limiting examples include cathode ray tubes, liquid crystal displays, plasma televisions, projectors, and head-mounted displays such as Google Glass™.

[29] In the context of the present specification, the expression "information" includes information of any nature or kind whatsoever capable of being stored in a database. Thus information includes, but is not limited to audiovisual works (images, movies, sound records, presentations etc.), data (location data, numerical data, etc.), text (opinions, comments, questions, messages, etc.), documents, spreadsheets, etc.

[30] In the context of the present specification, the expression "indication of" is meant to refer to any type and quantity of information enabling identification of the object which it qualifies, whether or not that information includes the object itself. For instance, an "indication of text" refers to information enabling identification of the text in question, whether or not that information includes the text itself. Non-limiting examples of indications that do not include the object itself include hyperlinks, references, and pointers.

[31] In the context of the present specification, a character may be said to be "encoded according to a character encoding" if it may be unambiguously interpreted by appropriately programmed computer hardware and/or software as representative of that character with reference to that character encoding. The present technology is not limited to any particular character encoding, nor is it limited to standard character encodings such as ASCII or Unicode (e.g. UTF-8), as proprietary character encodings may also be used. As a counterexample, an image representation of a character is not a "character encoded according to a character encoding" because the image representation may be interpreted to represent one

of two (or more) characters, depending on particularities of the OCR algorithm employed to detect the character represented by the image representation.

[32] In the context of the present specification, "image metadata" is meant to refer to any type and quantity of information about at least one image, structured either according to a known standard or according to a proprietary structure, whether the one or more elements of that metadata are located together with the image, separately from the image, or a combination thereof.

[33] In the context of the present specification, "video metadata" is meant to refer to any type and quantity of information about at least one video, structured either according to a known standard or according to a proprietary structure, whether the one or more elements of that metadata are located together with the video, separately from the video, or a combination thereof.

[34] In the context of the present specification, "audio metadata" is meant to refer to any type and quantity of information about at least one audio clip, structured either according to a known standard or according to a proprietary structure, whether the one or more elements of that metadata are located together with the audio clip, separately from the audio clip, or a combination thereof.

[35] In the context of the present specification, the expressions "unmodified image" and "modified image" are meant to refer only to an incremental modification of an image according to the present technology. An unmodified image may well have been modified previously, whether according to the present technology or not.

[36] In the context of the present specification, a "screenshot image" of a display is meant to refer to an image substantially replicating the visual content displayed on the display at a given time (usually but not necessarily at the time generation of the screenshot image was requested).

[37] In the context of the present specification, a "database" is any structured collection of data, irrespective of its particular structure, the database management software, or the computer hardware on which the data is stored, implemented or otherwise rendered available for use. A database may reside on the same hardware as the process that stores or makes use

of the information stored in the database or it may reside on separate hardware, such as a dedicated server or plurality of servers.

[38] In the context of the present specification, the expression "component" is meant to refer either to hardware, software, or a combination of hardware and software that is both necessary and sufficient to achieve the specific function(s) being referenced. For example, a "voice recognition component" includes hardware and/or software suitable for translating a live or previously recorded audio sample of a human voice into a textual equivalent.

[39] In the context of the present specification, the expression "computer-readable medium" is intended to include media of any nature and kind whatsoever, including RAM, ROM, disks (CD-ROMs, DVDs, floppy disks, hard drivers, etc.), USB keys, solid state-drives, tape drives, etc.

[40] In the context of the present specification, the words "first", "second", "third", etc. have been used as adjectives only for the purpose of allowing for distinction between the nouns that they modify from one another, and not for the purpose of describing any particular relationship between those nouns. Thus, for example, it should be understood that, the use of the terms "first server" and "third server" is not intended to imply any particular order, type, chronology, hierarchy or ranking (for example) of/between the server, nor is their use (by itself) intended imply that any "second server" must necessarily exist in any given situation. Further, as is discussed herein in other contexts, reference to a "first" element and a "second" element does not preclude the two elements from being the same actual real-world element. Thus, for example, in some instances, a "first" server and a "second" server may be the same software and/or hardware, in other cases they may be different software and/or hardware.

[41] Implementations of the present technology each have at least one of the above-mentioned object and/or aspects, but do not necessarily have all of them. It should be understood that some aspects of the present technology that have resulted from attempting to attain the above-mentioned object may not satisfy this object and/or may satisfy other objects not specifically recited herein.

[42] Additional and/or alternative features, aspects and advantages of implementations of the present technology will become apparent from the following description, the accompanying drawings and the appended claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[43] For a better understanding of the present technology, as well as other aspects and further features thereof, reference is made to the following description which is to be used in conjunction with the accompanying drawings, where:

[44] Figure 1 is a context diagram of a networked computing environment suitable for use with implementations of the present technology described herein;

[45] Figure 2 is an example of a screenshot image of a display of an electronic device;

[46] Figure 3 is an example of a user interaction of a user with a touchscreen display displaying the image of Figure 2, wherein the user taps a portion of the screenshot image;

[47] Figure 4 shows a continuation of the user interaction of Figure 3, wherein a cursor and virtual keyboard are displayed to invite the user to add text to the image;

[48] Figure 5 shows a further continuation of the user interaction, wherein the user has partially entered text via the virtual keyboard;

[49] Figure 6 shows a modified image resulting from the completion of the user interaction, wherein text has been added to the image of Figure 2;

[50] Figure 7 shows a block diagram representing an image file according to the Portable Network Graphics (PNG) specification.

[51] Figure 8-22 show flowcharts of various embodiments of methods for generating metadata according to various implementations of the present technology.

DETAILED DESCRIPTION

[52] Referring to Fig. 1, there is shown a diagram of a simple networked computing environment 100 comprising a smartphone 120 in communication with a server 130 via a communications network 101 (e.g. the Internet). It is to be expressly understood that the various elements of networked computing environment 100 depicted herein and hereinafter described are merely intended to illustrate some possible implementations of the present technology. The description which follows is not intended to define the scope of the present technology, nor to set forth its bounds. In some cases, what are believed to be helpful

examples of modifications to computer systems 100 may also be described below. This is done merely as an aid to understanding, and, again, not to define the scope or bounds of the present technology. These modifications are not an exhaustive list, and, as a person skilled in the art would understand, other modifications are likely possible. Further, where examples of modifications are absent, the mere absence of such examples should not be interpreted to mean that no modifications are possible and/or that what is described is the sole manner of implementing that element of the present technology. As a person skilled in the art would understand, this is likely not the case. It is also to be understood that elements of the networked computing environment 100 may represent relatively simple implementations of the present technology, and that where such is the case, they have been presented in this manner as an aid to understanding. As persons skilled in the art would understand, various implementations of the present technology may be of a greater complexity.

[53] Smartphone 120 depicted in Fig. 1 is an Apple™ iPhone™ running the iOS™ operating system. In other implementations, another suitable operating system (e.g. Google Android™, Microsoft Windows Phone™, BlackBerry OS™) may be used. Moreover, because the present technology is not limited to mobile devices, smartphone 120 may be replaced by a non-mobile device in other implementations of the present technology. In the depicted implementation, smartphone 120 includes a touchscreen display 122, a home button 124, and a power button 126, and it is operated by user 110.

[54] User 110 may operate smartphone 120 to launch an application which displays visual content on touchscreen display 122. For example, user 110 may launch the "Stocks" iOS application and then operate it to display a two-year chart of shares trading under the ticker YNDX on the NASDAQ stock exchange, as depicted in Fig. 2. User 110 may then provide an instruction to smartphone 120 to capture a screenshot image of the visual content displayed on display 122, for example by simultaneously pressing home button 124 and power button 126 of smartphone 120, causing smartphone 120 to generate a screenshot image such as the screenshot image 200 depicted in Fig. 2.

[55] In some cases, the visual content displayed on display 122 when user 110 instructs smartphone 120 to capture the screenshot image may include known text (i.e. text susceptible of unambiguous interpretation by smartphone 120 based on a character encoding of the one or more characters included in that text). For example, with reference again to Fig 2., when user 110 instructed smartphone 120 to capture screenshot image 200, the visual content being

displayed by the "Stocks" app on display 122 of smartphone 120 included several known text elements, such as the text "YNDX" labeled 202. By querying the "Stocks" app to identify any text elements (e.g. "YNDX") it is displaying on display 122 when the screenshot capture instruction is received, implementations of the present technology may obtain the text elements and subsequently use them when generating image metadata to be associated with the screenshot image. Note that the text 202 "YNDX" has been arbitrarily singled out from among the many text elements shown in Fig. 2 (e.g. "YANDEX N.V.", "+0.62", "JULY", "2013", "23.55", "3M", "ROGERS", etc.), any of which could be substituted for text 202 in the following description.

[56] In some but not all implementations of the present technology, smartphone 120 takes advantage of the fact that text 202 is known unambiguously at the time screenshot image 200 is generated. More specifically, in such implementations, smartphone 120 generates image metadata based on text 202 either in parallel with or as part of the process of generating screenshot image 200, and then associates that image metadata with screenshot image 200. In some implementations, this may be as simple as copying text 202 (e.g. "YNDX") into a text field of the image metadata, and then saving that image metadata together with the image in an image file (e.g. in the iTXt chunk of a PNG image file, as described in more detail below with reference to Fig. 7).

[57] In some implementations, image metadata is generated while an unmodified image such as screenshot image 200 is modified by user 110 to include an image representation of text. An example user interaction resulting in such a modification is depicted in Figs. 3 to 6. In Fig. 3, user 110 taps a portion of screenshot image 200. This causes smartphone 120 to display cursor 204 and virtual keyboard 128 on display 122 of smartphone 120 as depicted in Fig. 4. In Fig. 5, user 110 is in the process of entering the text "A GOOD YEAR FOR YANDEX SHAREHOLDERS" by tapping the virtual keys of virtual keyboard 128. In some implementations, individual keystrokes are received and processed one by one. In others, keystrokes are buffered until the user indicates the text 206 is complete (e.g. by tapping return or a portion of the screen other than virtual keyboard 128). Once text 206 has been completely entered, smartphone 120 generates the modified image shown in Fig. 6, which includes an image representation of text 206. Smartphone 120 also generates image metadata based at least in part on text 206. For example, the image metadata may comprise a text field, and smartphone 120 may populate the text field with one or more characters included in text 206, such as "SHARE" or "SHAREHOLDERS" or "GOOD YEAR".

[58] In some implementations, the functionality of generating metadata from known text while generating a screenshot image may be combined with the functionality of generating metadata while modifying that screenshot image. For example, first image metadata may be generated based on the text 202 "YNDX" while generating screenshot image 200 as an image to be modified (i.e. an "unmodified image"), second image metadata may be generated based on text 206 "A GOOD YEAR FOR YANDEX SHAREHOLDERS", and both the first image metadata and the second image metadata may be associated with the resulting image (i.e. that shown in Fig. 6), which includes a respective image representation of each of text 202 and text 206.

[59] One means of associating the generated image and the generated image metadata is by writing an image file including them both to a computer-readable storage medium, such as a memory of smartphone 120. For the sake of compatibility, a popular image file format such as the Portable Network Graphics (PNG) file format may be used. A variety of programming libraries for creating and manipulating PNG files exist, including *libpng,* which is available as source code in the C programming language. Fig. 7 shows a block diagram of a PNG image file 300. The first eight bytes of the file (labeled 301) consist of the standard PNG file signature. A series of critical "chunks" 302 to 305 then follows. The IHDR chunk 302 contains image 300's width, height, and bit depth. The PLTE chunk 303 contains the palette or list of colors used in image 300. One or more IDAT chunks 304 contain the actual image data of image 300. Finally, the IEND chunk 305 indicates the end of the image data. According to the PNG specification, a variety of ancillary chunks may be included in a PNG image file 300. One such chunk is the iTXt chunk 310, which allows for storage of text comprising characters encoded according to the UTF-8 character encoding. Some implementations of the present technology may associate a generated image with generated image metadata comprising a text field by including in a PNG image file 300 the image data of the image in one or more IDAT chunks 304 and the text field in an iTXt chunk 310. In some implementations, characters of the text will need to be converted to UTF-8 from a character encoding other than UTF-8, according to known techniques well known to those skilled in the art.

[60] Apart from PNG files, many other image file formats are also suitable for storing image metadata along with an image. Non-limiting examples include JPEG and TIFF files, which support the EXIF (exchangeable image file format) standard commonly used in digital camera technology to store information about digital photographs.

[61] Other means of associating the image and image metadata are also possible. One such means comprises creating or modifying one or more database entries to indicate that the image metadata pertains to the image. For example, this may be indicated merely by including, in the one or more database entries, both an indication of the image and an indication of the image metadata. Another means comprises storing each of the image and the image metadata in separate files, wherein at least one of the files includes an indication of the other file (e.g. an absolute or relative link/pointer/reference to the other file).

[62] As those skilled in the art will understand, implementations of the present technology may likewise provide a method of generating metadata in respect of a video based on text to be included in one or more of the images that make up the individual frames (series of images) of the video. The video and metadata may each be generated based at least in part on the text and then associated with one another (e.g. via a video file including the video and the metadata, a database entry, a communication including the video and the metadata, or some other means of association). Similarly, implementations of the present technology may provide a method of generating metadata in respect of audio, wherein both audio which includes an audio representation of text (e.g. generated via text-to-speech technology) and metadata based at least in part on the text may be generated and associated with one another.

[63] Fig. 8 shows a flowchart of a method 400 for generating image metadata according to an embodiment of a client-server implementation of the present technology, wherein a smartphone 120 acts as a client device in communication with a server 130 as depicted in Fig. 1. At step 402, smartphone 120 captures a screenshot of its display 122 based on an instruction from user 110. At step 404, user 110 inputs text with which to modify the screenshot image. At step 406, both the screenshot image and the text are sent by smartphone 120 to server 130 via communications network 101. At step 408, server 130 receives the screenshot image and the indication of the text. At step 410, server 130 generates the image based at least in part on the text and the screenshot image, the image including an image representation of the text. At step 412, server 130 generates image metadata based at least in part on the text and existing image metadata associated with the screenshot image. This includes populating a text field of the image metadata with at least some of the text, which in turn includes translating the text from a first character encoding to a second character encoding. At step 414, server 130 associates the generated image and image metadata by sending an indication of each to smartphone 120 in a communication. At step 416, smartphone 120 receives the indications, and finally, at step 418, smartphone 120 writes an

image file including the image and the image metadata to a non-transitory computer-readable medium of smartphone 120.

[64] Fig. 9 shows a flowchart of a method 500 for generating image metadata while modifying a screenshot image according to implementations of the present technology. At step 510, an instruction to generate a screenshot image of a display of the electronic device is received from a user of the electronic device. At step 520, the screenshot image is generated as an unmodified image. At step 530, an indication of text to be included in an image is received. Step 530 comprises step 532. At step 532, an indication of text with which to modify the unmodified image is received. At steps 540/542, the image is generated based at least in part on the text and the unmodified image, the image including an image representation of the text. At steps 550/552, the image metadata is generated based at least in part on the text and existing image metadata associated with the unmodified image (the screenshot image). At step 560 the image metadata is associated with the image. Step 560 comprises step 562. At step 562, an image file is written to a non-transitory computer-readable medium.

[65] Fig. 10 shows a flowchart of a method 600 for generating image metadata while modifying a screenshot image according to implementations of the present technology. At step 610, a screenshot image is received from another electronic device via a communications network. At step 620, an indication of text to be included in an image is received. Step 620 comprises step 622. At step 622, an indication of text with which to modify an unmodified image comprising the screenshot image is received. At steps 630/632, the image is generated based at least in part on the text and the unmodified image, the image including an image representation of the text. At step 640, the image metadata is generated based at least in part on the text. At step 650, the image metadata is associated with the image. Step 650 comprises step 652. At step 652, an entry in a database is created and/or modified to include an indication of the image and an indication of the image metadata.

[66] Fig. 11 shows a flowchart of a method 700 for generating image metadata while modifying a digital photograph according to implementations of the present technology. At step 710, an instruction to capture a digital photograph is received from a user of the electronic device. At step 720, the digital photograph is captured via a camera coupled to the electronic device. At step 730, an indication of text to be included in an image is received. Step 730 comprises step 732. At step 732, an indication of text with which to modify an

unmodified image, the unmodified image comprising the digital photograph, is received. At steps 740/742, an image is generated based at least in part on the text and the unmodified image generating the image based at least in part on the text, the image including an image representation of the text. At step 750, image metadata is generated based at least in part on the text. At step 760, the image metadata is associated with the image. Step 760 includes step 762. At step 762, a communication including an indication of the image and an indication of the image metadata is sent via a communications network.

[67] Fig. 12 shows a flowchart of a method 800 for generating screenshot image metadata based on text displayed on a display when a screenshot capture instruction is received. At step 810, an indication of text to be included in an image is received. Step 810 comprises steps 812 and 814. At step 812, an instruction to generate a screenshot image of a display of the electronic device is received from a user of the electronic device. At step 814, at least some of the text displayed on the display is captured as the text to be included in the image. At steps 820/822, the screenshot image is generated based at least in part on the text, the image including an image representation of the text. At step 830, the image metadata is generated based at least in part on the text. At step 840, the image metadata is associated with the image.

[68] Fig. 13 shows a flowchart of a method 900 for generating image metadata in respect of an image other than a screenshot image when a screenshot capture instruction is received. Thus, instead of generating an actual screenshot, an image which includes at least some text displayed on the display is generated. For example, if the words “File” and “Home” are displayed on the display when the screenshot capture instruction is received, an image including an image representation of “File” may be generated, without necessarily including other text or graphics displayed on the display. At step 910, an indication of text to be included in the image is received. Step 910 comprises steps 912 and 914. At step 912, an instruction to generate a screenshot image of a display of the electronic device is received. At step 914, at least some of the text displayed on the display is captured (e.g. “File”) as the text to be included in the image. The text in question may be captured, for example, by sending a request to the application which is causing the text to be displayed to identify the text. At step 920 an image based at least in part on the text is generated, the image including an image representation of the text. At step 930, the image metadata is generated based at least in part on the text. At step 940, the image metadata is associated with the image.

[69] Fig. 14 shows a flowchart of a method 1000 for generating image metadata while modifying an image. At step 1010, an indication of text with which to modify an image is received, the text comprising at least one character, each character being encoded according to a first character encoding (e.g. ASCII). At step 1020, the image is modified based at least in part on the text to include an image representation of the text. At step 1030, the image metadata is generated based at least in part on the text, the image metadata including a text field conforming to a second character encoding (e.g. UTF-8) other than the first character encoding. Step 1030 comprises step 1032. At step 1032, the text field is populated with at least some of the text. Because the character encoding is different, step 1032 comprises step 1034, namely translating the at least some of the text from the first character encoding to the second character encoding. At step 1040, the image metadata is associated with the image.

[70] Fig. 15 shows a flowchart of a method 1100 for augmenting image metadata associated with an image when that image is modified. At step 1110, an indication of text with which to modify an image is received. Step 1110 comprises step 1112. At step 1112, the indication of the text is received from another electronic device in communication with the electronic device via a direct link and/or a communications network. At step 1120, the image is modified based at least in part on the text to include an image representation of the text. At step 1130, additional image metadata is generated based at least in part on the text. At step 1140, the additional image metadata is associated with the image by adding the additional image metadata to the image metadata associated with the image.

[71] Fig. 16 shows a flowchart of a method 1200 for generating video metadata while generating a video. At step 1210, an indication of the text to be included in at least one frame of the video is received, the text comprising at least one character, each character being encoded according to a first character encoding (e.g. UTF-8). At step 1220, the video is generated based at least in part on the text, the video comprising at least one frame including an image representation of the text. At step 1230, the video metadata is generated based at least in part on the text, the video metadata including a text field conforming to a second character encoding (e.g. UTF-16) other than the first character encoding. Step 1230 comprises step 1232. At step 1232, the text field is populated with at least some of the text. Because the character encoding is different, step 1232 comprises step 1234, namely translating the at least some of the text from the first character encoding to the second character encoding. At step 1240, the video metadata is associated with the video. Step 1240

comprises step 1242. At step 1242, a video file including the video and the video metadata is written to a non-transitory computer-readable medium.

[72] Fig. 17 shows a flowchart of a method 1300 for generating video metadata while generating a video. At step 1310, an indication of text to be included in at least one frame of the video is received. At step 1320, the video comprising the at least one frame is generated based at least in part on the text, the at least one frame including an image representation of the text. At step 1330, the video metadata is generated based at least in part on the text. At step 1340, the video metadata is associated with the video. Step 1340 comprises step 1342. At step 1342, an entry in a database is created and/or modified to include an indication of the video and an indication of the video metadata.

[73] Fig. 18 shows a flowchart of a method 1400 for generating video metadata while generating a video. At step 1410, an indication of text to be included in at least one frame of the video is received. At step 1420, the video comprising the at least one frame is generated based at least in part on the text, the at least one frame including an image representation of the text. At step 1430, the video metadata is generated based at least in part on the text. At step 1440, the video metadata is associated with the video. Step 1440 comprises step 1442. At step 1442, a communication including an indication of the video and an indication of the video metadata is sent via a communications network.

[74] Fig. 19 shows a flowchart of a method 1500 for generating audio metadata while generating an audio clip. At step 1510, an indication of text to be included in the audio clip is received, the text comprising at least one character, each character being encoded according to a first character encoding (e.g. a proprietary, non-standard character encoding). At step 1520, the audio clip is generated based at least in part on the text, the audio clip including an audio representation of the text. For example, a text-to-speech component may be used to generate spoken audio representative of the text. At step 1530, audio metadata is generated based at least in part on the text, the audio metadata including a text field conforming to a second character encoding (e.g. UTF-8) other than the first character encoding. Step 1530 comprises step 1532. At step 1532, the text field is populated with at least some of the text. Because the character encoding is different, step 1532 comprises step 1534, namely translating the at least some of the text from the first character encoding to the second character encoding. At step 1540, the audio metadata is associated with the audio clip. Step 1540 comprises step 1542.

At step 1542, an audio file including the audio clip and the audio metadata is written to a non-transitory computer-readable medium.

[75] Fig. 20 shows a flowchart of a method 1600 for generating audio metadata while generating an audio clip. At steps 1610/1612/1614, an indication of text to be included in the audio clip is received from a user of the electronic device via at least one of a physical keyboard, a virtual keyboards, and a voice recognition component coupled to a microphone of the electronic device. At step 1620, the audio clip is generated based at least in part on the text, the audio clip including an audio representation of the text. At step 1630, the audio metadata is generated based at least in part on the text. At step 1640, the audio metadata is associated with the audio clip. Step 1640 comprises step 1642. At step 1642, an entry including an indication of the audio clip and an indication of the audio metadata is created and/or modified in a database.

[76] Fig. 21 shows a flowchart of a method 1700 for generating audio metadata while generating an audio clip. At step 1710, an indication of text to be included in the audio clip is received. At step 1720, the audio is generated based at least in part on the text, the audio clip including an audio representation of the text. At step 1730, the audio metadata is generated based at least in part on the text. At step 1740, the audio metadata is associated with the audio clip. Step 1740 comprises step 1742. At step 1742, a communication including an indication of the audio clip and an indication of the audio metadata is sent via a communications network.

[77] Fig. 22 shows a flowchart of a method 1800 for generating metadata while generating digital content. At step 1810, an indication of text to be represented in digital content is received. At step 1820, the digital content is created based at least in part on the text, the digital content including a non-textual representation of the text in at least a portion of the digital content. At step 1830, the metadata is generated based at least in part on the text. At step 1840, the metadata is associated with the digital content.

[78] Modifications and improvements to the above-described implementations of the present technology may become apparent to those skilled in the art. The foregoing description is intended to be exemplary rather than limiting. The scope of the present technology is therefore intended to be limited solely by the scope of the appended claims.

CLAIMS

1. A method of generating image metadata, the method comprising, at an electronic device:

receiving an indication of text to be included in an image, the text comprising at least one character, each character being encoded according to a character encoding;

generating the image based at least in part on the text, the image including an image representation of the text;

generating the image metadata based at least in part on the text; and

associating the image metadata with the image.

2. The method of claim 1, wherein:

receiving the indication of the text to be included in the image comprises receiving an indication of text with which to modify an unmodified image; and

generating the image based at least in part on the text comprises generating the image based at least in part on the text and the unmodified image.

3. The method of claim 2, wherein generating the image metadata based at least in part on the text comprises generating the image metadata based at least in part on the text and existing image metadata associated with the unmodified image.

4. The method of any one of claims 2 and 3, wherein the unmodified image comprises a screenshot image of a display of the electronic device, and further comprising, before generating the image:

receiving an instruction to generate the screenshot image from a user of the electronic device; and

generating the screenshot image as the unmodified image.

5. The method of any one of claims 2 and 3, wherein the unmodified image comprises a screenshot image of a display of a second electronic device in communication with the electronic device via a communications network; and further comprising, before generating the image, receiving the screenshot image from the second digital electronic device via the communications network.

6. The method of any one of claims 2 and 3, wherein the unmodified image comprises a digital photograph, and further comprising, before generating the image:

receiving an instruction to capture the digital photograph from a user of the electronic device; and

capturing the digital photograph via a camera coupled to the electronic device as the unmodified image.

7. The method of claim 1, wherein receiving the indication of the text to be included in the image comprises:

receiving an instruction to generate a screenshot image of a display of the electronic device from a user of the electronic device; and

capturing as the text at least some of the text displayed on the display.

8. The method of claim 7, wherein generating the image based at least in part on the text comprises generating the screenshot image as the image.

9. The method of claim 7, wherein generating the image based at least in part on the text is generating the image based at least in part on the text without generating the screenshot image.

10. A method of generating image metadata, the method comprising, at an electronic device:

receiving an indication of text with which to modify an image, the text comprising at least one character, each character being encoded according to a character encoding;

modifying the image based at least in part on the text to include an image representation of the text;

generating the image metadata based at least in part on the text; and

associating the image metadata with the image.

11. A method of augmenting image metadata associated with an image, the method comprising, at an electronic device:

receiving an indication of text with which to modify the image, the text comprising at least one character, each character being encoded according to a character encoding;

modifying the image based at least in part on the text to include an image representation of the text;

generating additional image metadata based at least in part on the text; and

associating the additional image metadata with the image by adding the additional image metadata to the image metadata associated with the image.

12. The method of any one of claims 1 to 11, wherein the image metadata includes a text field, and generating the image metadata based at least in part on the text includes populating the text field with at least some of the text.

13. The method of claim 12, wherein the character encoding is a first character encoding, the text field conforms to a second character encoding other than the first character encoding, and populating the text field with at least some of the text comprises translating the at least some of the text from the first character encoding to the second character encoding.

14. The method of any one of claims 1 to 13, wherein associating the image metadata with the image comprises writing an image file including the image and the image metadata to a non-transitory computer-readable medium.

15. The method of any one of claims 1 to 13, wherein associating the image metadata with the image comprises at least one of creating and modifying an entry in a database, the entry including an indication of the image and an indication of the image metadata.

16. The method of any one of claims 1 to 13, wherein associating the image metadata with the image comprises sending a communication including an indication of the image and an indication of the image metadata via a communications network.

17. A method of generating video metadata, the method comprising, at an electronic device:

receiving an indication of text to be included in at least one frame of a video, the text comprising at least one character, each character being encoded according to a character encoding;

generating the video comprising the at least one frame based at least in part on the text, the at least one frame including an image representation of the text;

generating the video metadata based at least in part on the text; and

associating the video metadata with the video.

18. The method of claim 17, wherein the video metadata includes a text field, and generating the video metadata based at least in part on the text includes populating the text field with at least some of the text.

19. The method of claim 18, wherein the character encoding is a first character encoding, the text field conforms to a second character encoding other than the first character encoding, and populating the text field with at least some of the text comprises translating the at least some of the text from the first character encoding to the second character encoding.

20. The method of any one of claims 17 to 19, wherein associating the video metadata with the video comprises writing a video file including the video and the video metadata to a non-transitory computer-readable medium.

21. The method of any one of claims 17 to 19, wherein associating the video metadata with the video comprises at least one of creating and modifying an entry in a database, the entry including an indication of the video and an indication of the video metadata.

22. The method of any one of claims 17 to 19, wherein associating the video metadata with the video comprises sending a communication including an indication of the video and an indication of the video metadata via a communications network.

23. A method of generating audio metadata, the method comprising, at an electronic device:

receiving an indication of text to be included in an audio clip, the text comprising at least one character, each character being encoded according to a character encoding;

generating the audio clip based at least in part on the text, the audio clip including an audio representation of the text;

generating the audio metadata based at least in part on the text; and

associating the audio metadata with the audio clip.

24. The method of claim 23, wherein the audio metadata includes a text field, and generating the audio metadata based at least in part on the text includes populating the text field with at least some of the text.

25. The method of claim 24, wherein the character encoding is a first character encoding, the text field conforms to a second character encoding other than the first character encoding, and populating the text field with at least some of the text comprises translating the at least some of the text from the first character encoding to the second character encoding.

26. The method of any one of claims 23 to 25, wherein associating the audio metadata with the audio clip comprises writing an audio file including the audio clip and the audio metadata to a non-transitory computer-readable medium.

27. The method of any one of claims 23 to 25, wherein associating the audio metadata with the audio clip comprises at least one of creating and modifying an entry in a database, the entry including an indication of the audio clip and an indication of the audio metadata.

28. The method of any one of claims 23 to 25, wherein associating the audio metadata with the audio clip comprises sending a communication including an indication of the audio clip and an indication of the audio metadata via a communications network.

29. The method of any one of claims 1 to 28, wherein receiving the indication of the text comprises receiving the indication of the text from a user of the electronic device via the electronic device.

30. The method of claim 29, wherein receiving the indication of the text from the user of the electronic device comprises receiving the indication of the text via at least one of a physical keyboard of the electronic device, a virtual keyboard of the electronic device, and a voice recognition component coupled to a microphone of the electronic device.

31. The method of any one of claims 1 to 28, wherein receiving the indication of the text comprises receiving the indication of the text from a second electronic device in communication with the electronic device via at least one of a direct link and a communications network.

32. A method of generating metadata, the method comprising, at an electronic device:

receiving an indication of text to be represented in digital content, the text comprising at least one character, each character being encoded according to a character encoding;

generating the digital content based at least in part on the text, the digital content including a non-textual representation of the text in at least a portion of the digital content;

generating the metadata based at least in part on the text; and

associating the metadata with the digital content.

33. An electronic device for generating image metadata, the electronic device comprising:

means for receiving an indication of text; and

at least one processor operationally connected to the means for receiving the indication of text and structured and configured to:

receive, via the means for receiving, an indication of text to be included in an image, the text comprising at least one character, each character being encoded according to a character encoding;

generate the image based at least in part on the text, the image including an image representation of the text;

generate the image metadata based at least in part on the text; and

associate the image metadata with the image.

34. An electronic device for generating image metadata, the electronic device comprising:

means for receiving an indication of text; and

at least one processor operationally connected to the means for receiving the indication of text and structured and configured to:

receive, via the means for receiving, an indication of text with which to modify an image, the text comprising at least one character, each character being encoded according to a character encoding;

modify the image based at least in part on the text to include an image representation of the text;

generate the image metadata based at least in part on the text; and

associate the image metadata with the image.

35. An electronic device for augmenting image metadata associated with an image, the electronic device comprising:

means for receiving an indication of text; and

at least one processor operationally connected to the means for receiving the indication of text and structured and configured to:

receive, via the means for receiving, an indication of text with which to modify the image, the text comprising at least one character, each character being encoded according to a character encoding;

modify the image based at least in part on the text to include an image representation of the text;

generate additional image metadata based at least in part on the text; and

associate the additional image metadata with the image by adding the additional image metadata to the image metadata associated with the image.

36. An electronic device for generating video metadata, the electronic device comprising:

means for receiving an indication of text; and

at least one processor operationally connected to the means for receiving the indication of text and structured and configured to:

receive, via the means for receiving, an indication of text to be included in at least one frame of a video, the text comprising at least one character, each character being encoded according to a character encoding;

generate the video comprising the at least one frame based at least in part on the text, the at least one frame including an image representation of the text;

generate the video metadata based at least in part on the text; and

associate the video metadata with the video.

37. An electronic device for generating audio metadata, the electronic device comprising:

means for receiving an indication of text; and

at least one processor operationally connected to the means for receiving the indication of text and structured and configured to:

receive, via the means for receiving, an indication of text to be included in an audio clip, the text comprising at least one character, each character being encoded according to a character encoding;

generate the audio clip based at least in part on the text, the audio clip including an audio representation of the text;

generate the audio metadata based at least in part on the text; and

associate the audio metadata with the audio clip.

38. An electronic device for generating metadata, the electronic device comprising:

means for receiving an indication of text; and

at least one processor operationally connected to the means for receiving the indication of text and structured and configured to:

receive, via the means for receiving, an indication of text to be represented in digital content, the text comprising at least one character, each character being encoded according to a character encoding;

generate the digital content based at least in part on the text, the digital content including a non-textual representation of the text in at least a portion of the digital content;

generate the metadata based at least in part on the text; and

associate the metadata with the digital content.

39. A non-transitory computer-readable information storage medium storing program instructions that, when executed by at least one processor of an electronic device comprising means for receiving an indication of text, effect:

receiving, via the means for receiving, of an indication of text to be included in an image, the text comprising at least one character, each character being encoded according to a character encoding;

generating of the image based at least in part on the text, the image including an image representation of the text;

generating of the image metadata based at least in part on the text; and

associating of the image metadata with the image.

40. A non-transitory computer-readable information storage medium storing program instructions that, when executed by at least one processor of an electronic device comprising means for receiving an indication of text, effect:

receiving, via the means for receiving, of an indication of text with which to modify an image, the text comprising at least one character, each character being encoded according to a character encoding;

modifying of the image based at least in part on the text to include an image representation of the text;

generating of the image metadata based at least in part on the text; and

associating of the image metadata with the image.

41. A non-transitory computer-readable information storage medium storing program instructions that, when executed by at least one processor of an electronic device comprising means for receiving an indication of text, effect:

receiving, via the means for receiving, of an indication of text with which to modify the image, the text comprising at least one character, each character being encoded according to a character encoding;

modifying of the image based at least in part on the text to include an image representation of the text;

generating of additional image metadata based at least in part on the text; and

associating of the additional image metadata with the image by adding the additional image metadata to the image metadata associated with the image.

42. A non-transitory computer-readable information storage medium storing program instructions that, when executed by at least one processor of an electronic device comprising means for receiving an indication of text, effect:

receiving, via the means for receiving, of an indication of text to be included in at least one frame of a video, the text comprising at least one character, each character being encoded according to a character encoding;

generating of the video comprising the at least one frame based at least in part on the text, the at least one frame including an image representation of the text;

generating of the video metadata based at least in part on the text; and

associating of the video metadata with the video.

43. A non-transitory computer-readable information storage medium storing program instructions that, when executed by at least one processor of an electronic device comprising means for receiving an indication of text, effect:

receiving, via the means for receiving, of an indication of text to be included in an audio clip, the text comprising at least one character, each character being encoded according to a character encoding;

generating of the audio clip based at least in part on the text, the audio clip including an audio representation of the text;

generating of the audio metadata based at least in part on the text; and

associating of the audio metadata with the audio clip.

44. A non-transitory computer-readable information storage medium storing program instructions that, when executed by at least one processor of an electronic device comprising means for receiving an indication of text, effect:

receiving, via the means for receiving, of an indication of text to be represented in digital content, the text comprising at least one character, each character being encoded according to a character encoding;

generating of the digital content based at least in part on the text, the digital content including a non-textual representation of the text in at least a portion of the digital content;

generating of the metadata based at least in part on the text; and

associating of the metadata with the digital content.

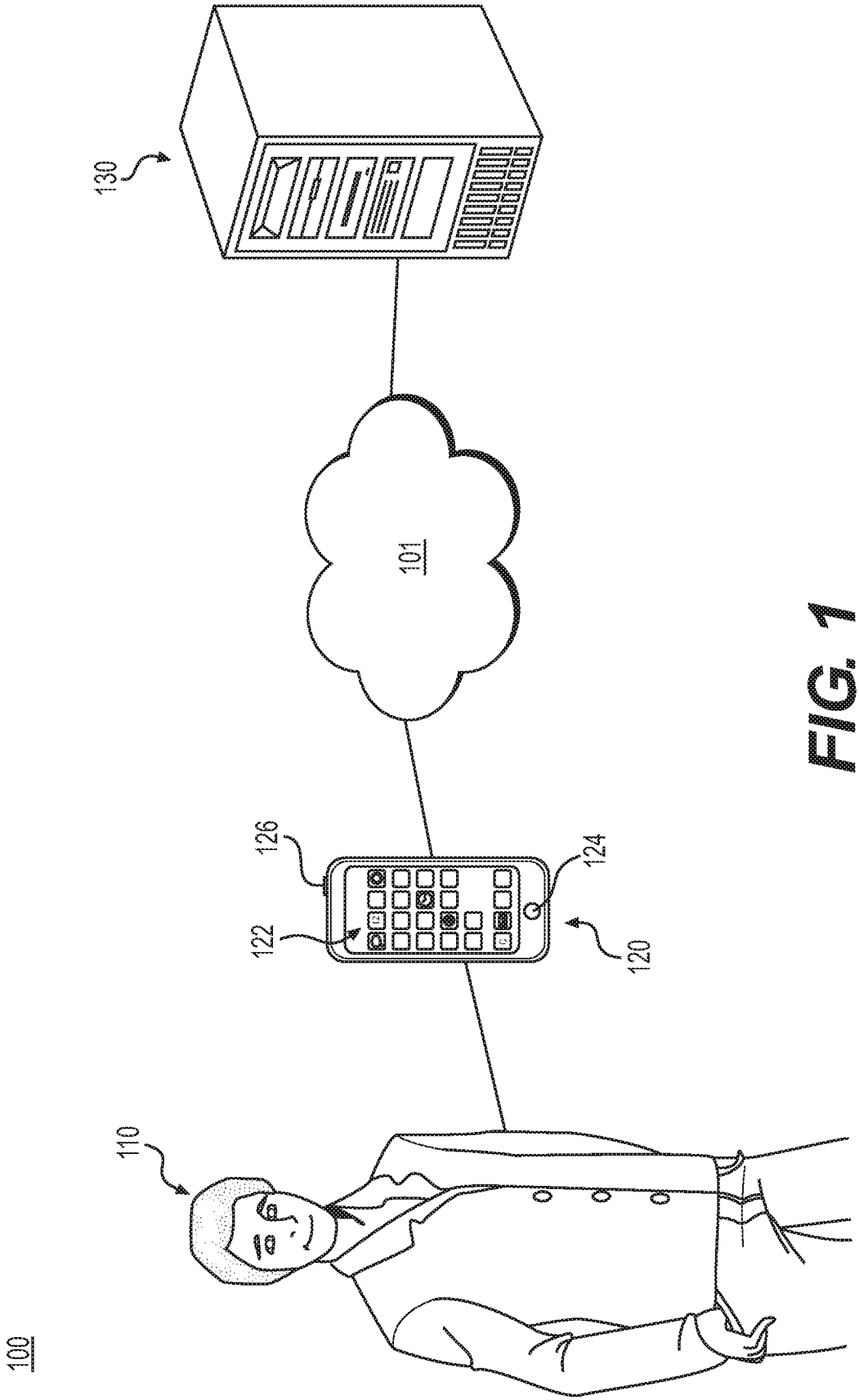

FIG. 1

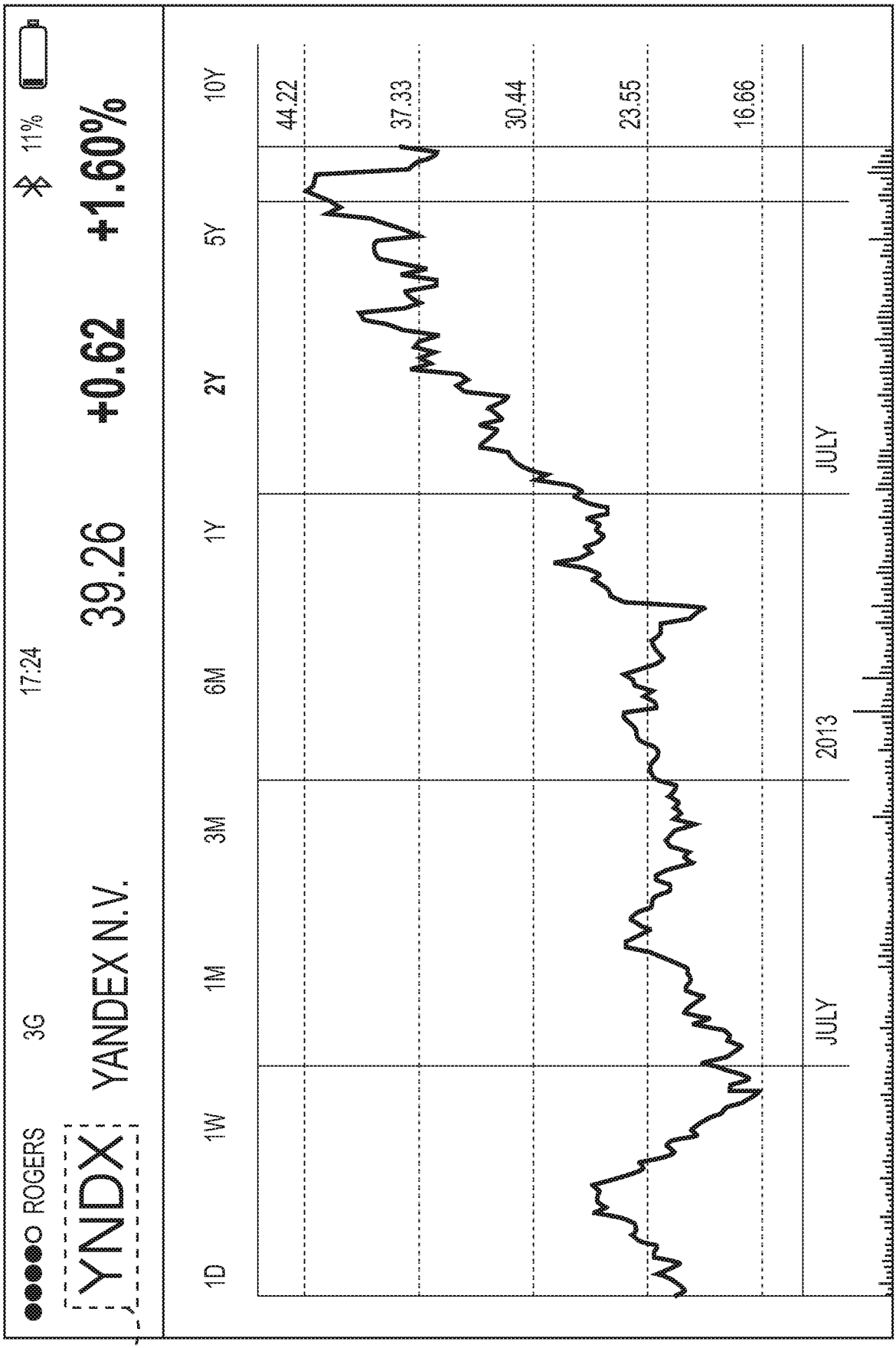
200
202
ROGERS
3G
17:24
11%
YNDX
YANDEX N.V.
39.26
+0.62
+1.60%
1D
1W
1M
3M
6M
1Y
2Y
5Y
10Y
44.22
37.33
30.44
23.55
16.66
JULY
2013
JULY

FIG. 2

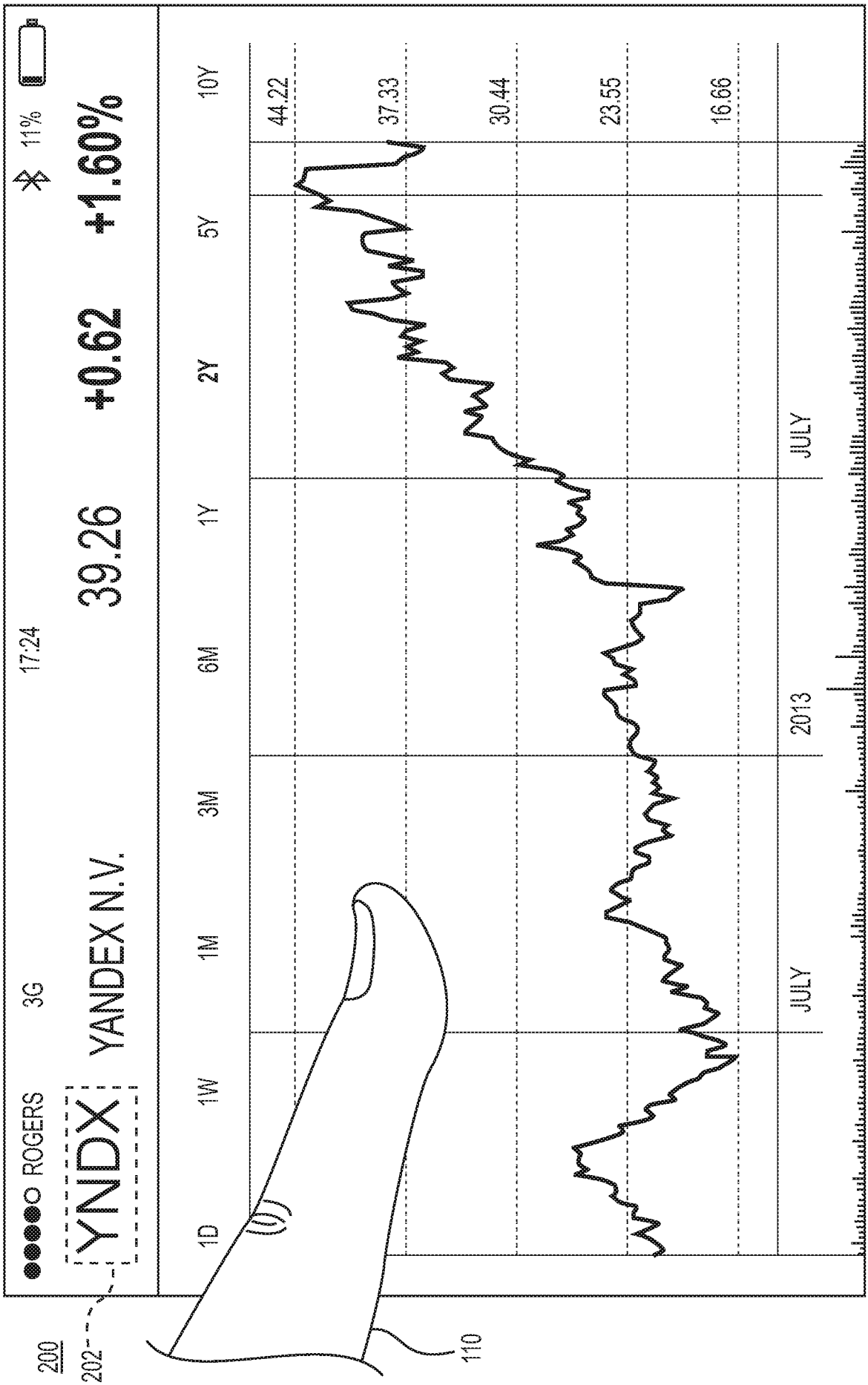
200
202
110
ROGERS
3G
17:24
11%
YNDX
YANDEX N.V.
39.26
+0.62
+1.60%
1D
1W
1M
3M
6M
1Y
2Y
5Y
10Y
44.22
37.33
30.44
23.55
16.66
JULY
2013
JULY

FIG. 3

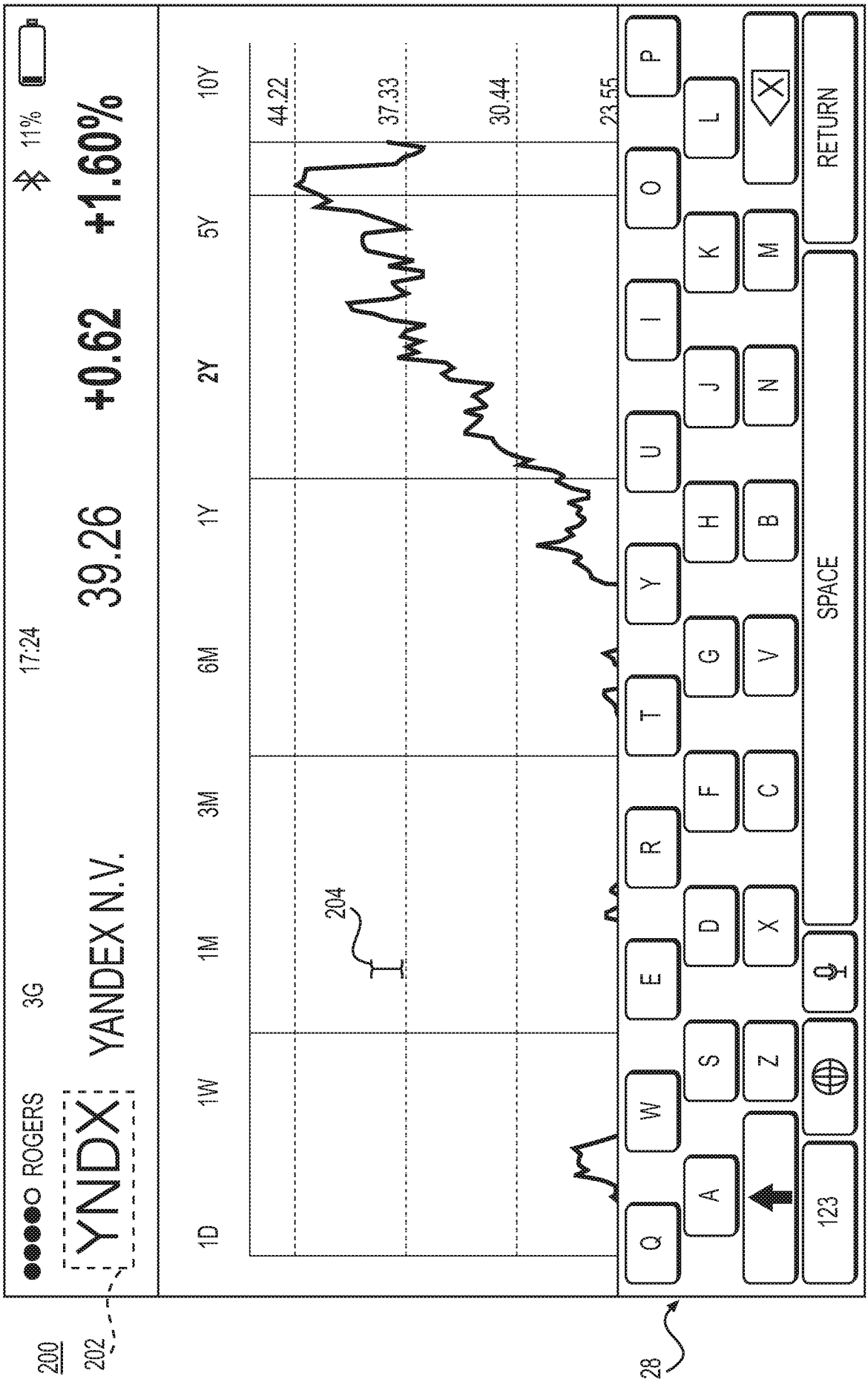
200
202
ROGERS
3G
17:24
11%
YNDX
YANDEX N.V.
39.26
+0.62
+1.60%
1D
1W
1M
3M
6M
1Y
2Y
5Y
10Y
44.22
37.33
30.44
23.55
204
128
Q
W
E
R
T
Y
U
I
O
P
A
S
D
F
G
H
J
K
L
Z
X
C
V
B
N
M
123
SPACE
RETURN

FIG. 4

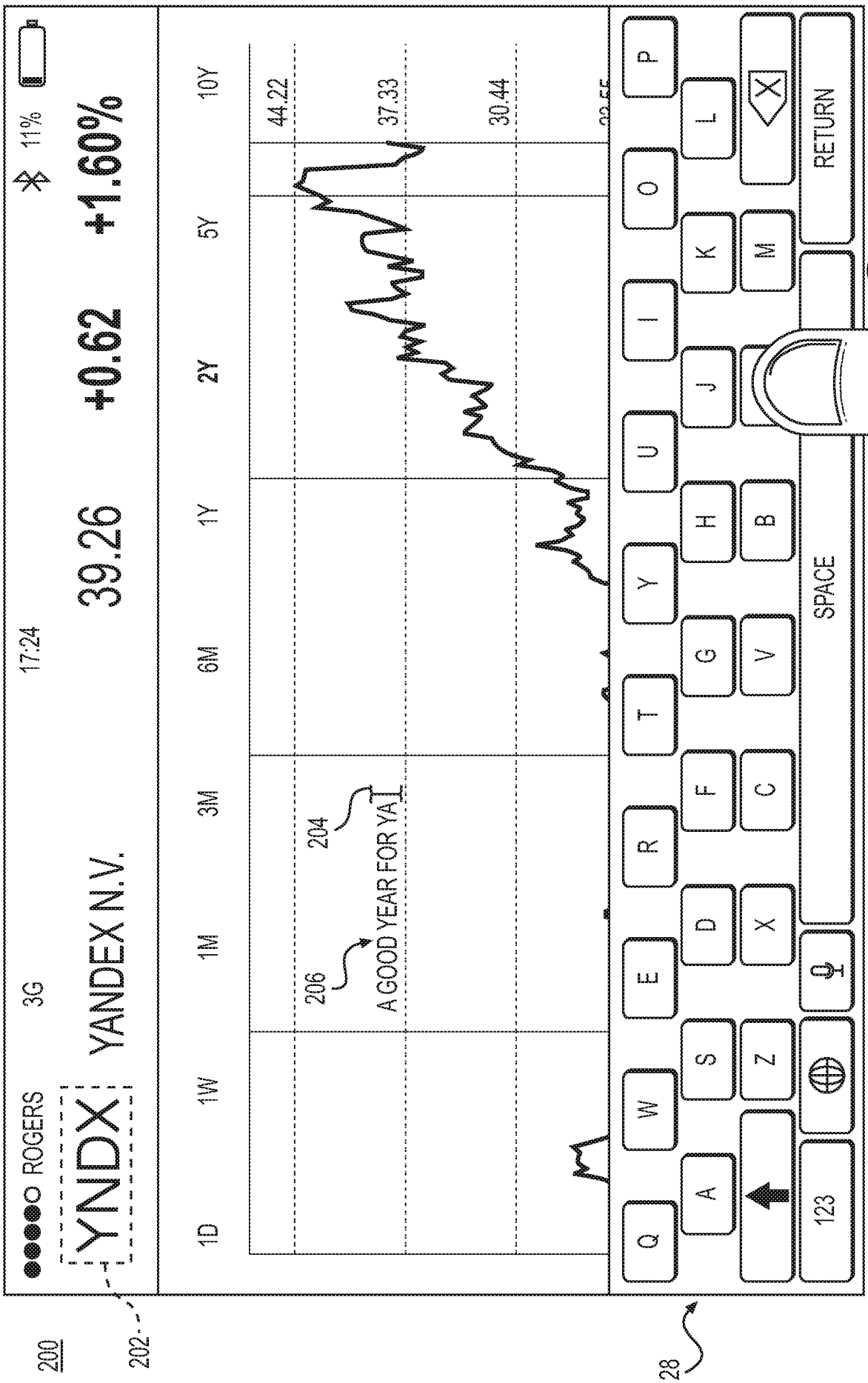
200
202
128
110
ROGERS
3G
17:24
11%
YNDX
YANDEX N.V.
39.26
+0.62
+1.60%
1D
1W
1M
3M
6M
1Y
2Y
5Y
10Y
44.22
37.33
30.44
204
206
A GOOD YEAR FOR YA
SPACE
RETURN
123

FIG. 5

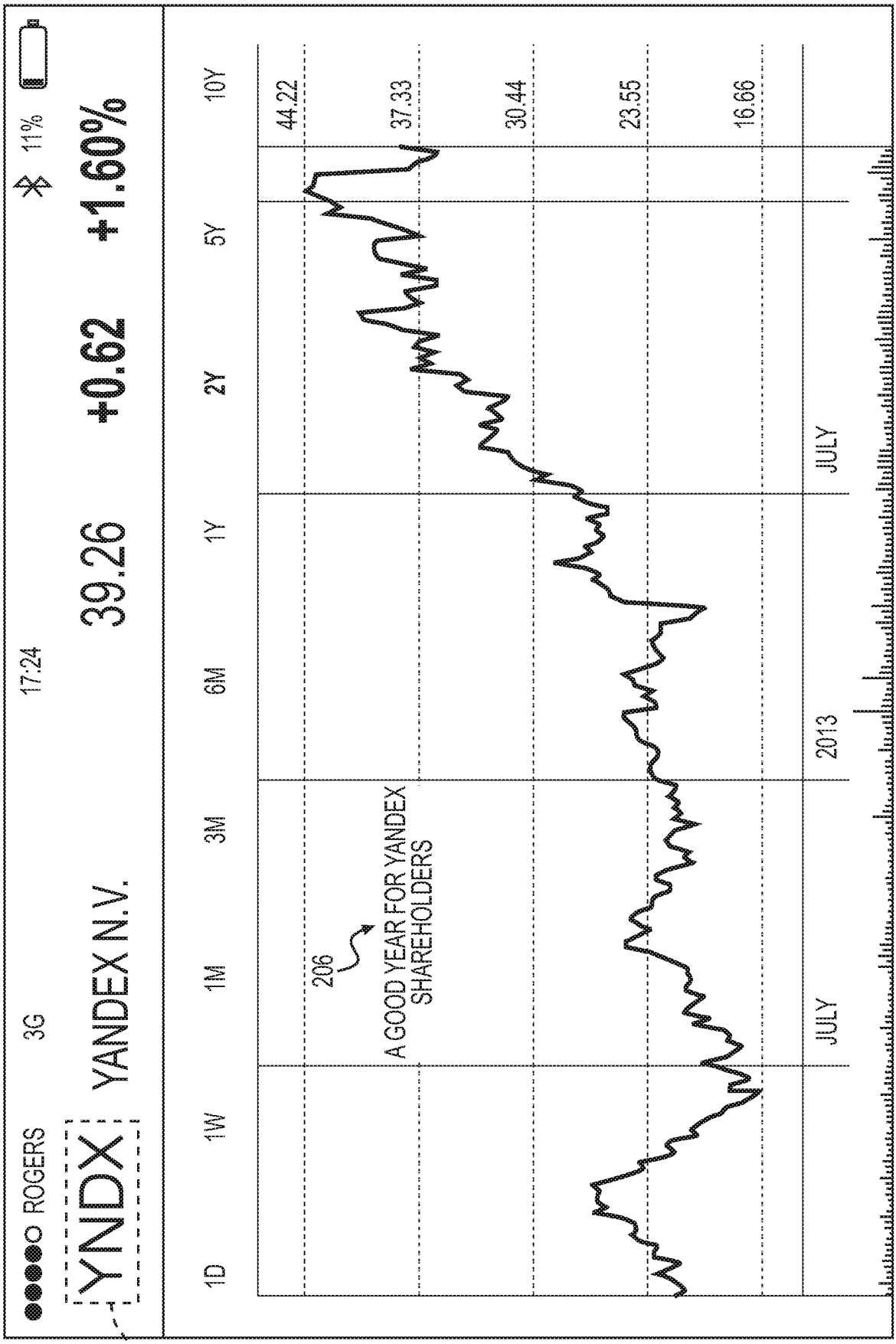
200
202
ROGERS
3G
17:24
11%
YNDX
YANDEX N.V.
39.26
+0.62
+1.60%
1D
1W
1M
3M
6M
1Y
2Y
5Y
10Y
44.22
37.33
30.44
23.55
16.66
206
A GOOD YEAR FOR YANDEX SHAREHOLDERS
JULY
2013
JULY

FIG. 6

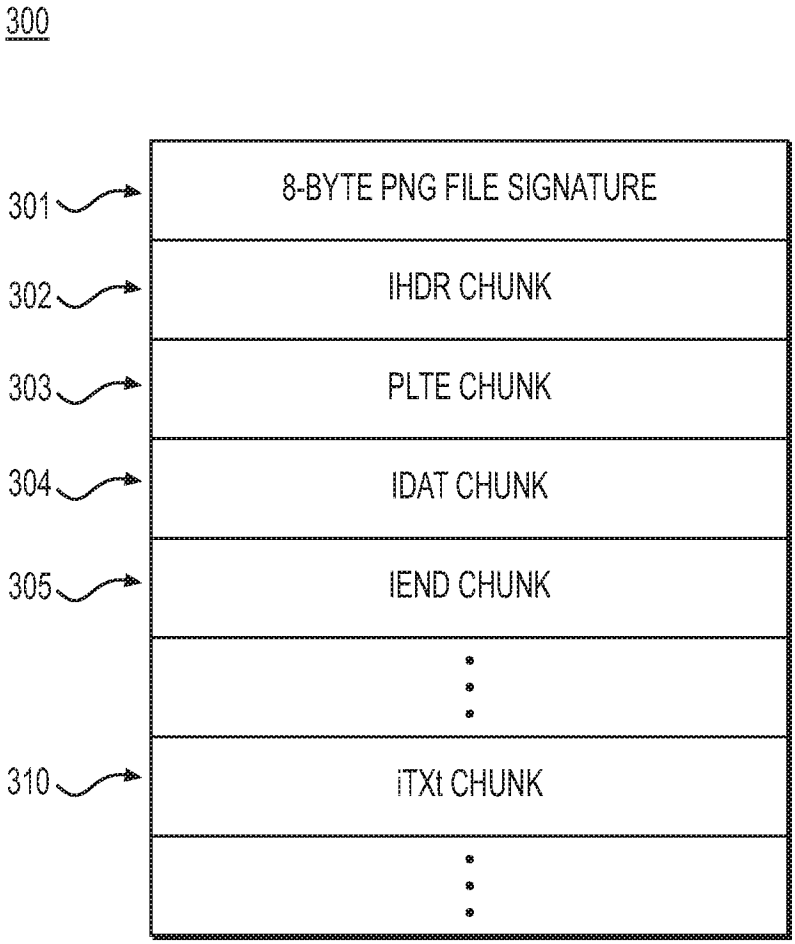
300
301
8-BYTE PNG FILE SIGNATURE
302
IHDR CHUNK
303
PLTE CHUNK
304
IDAT CHUNK
305
IEND CHUNK
310
iTXt CHUNK

FIG. 7

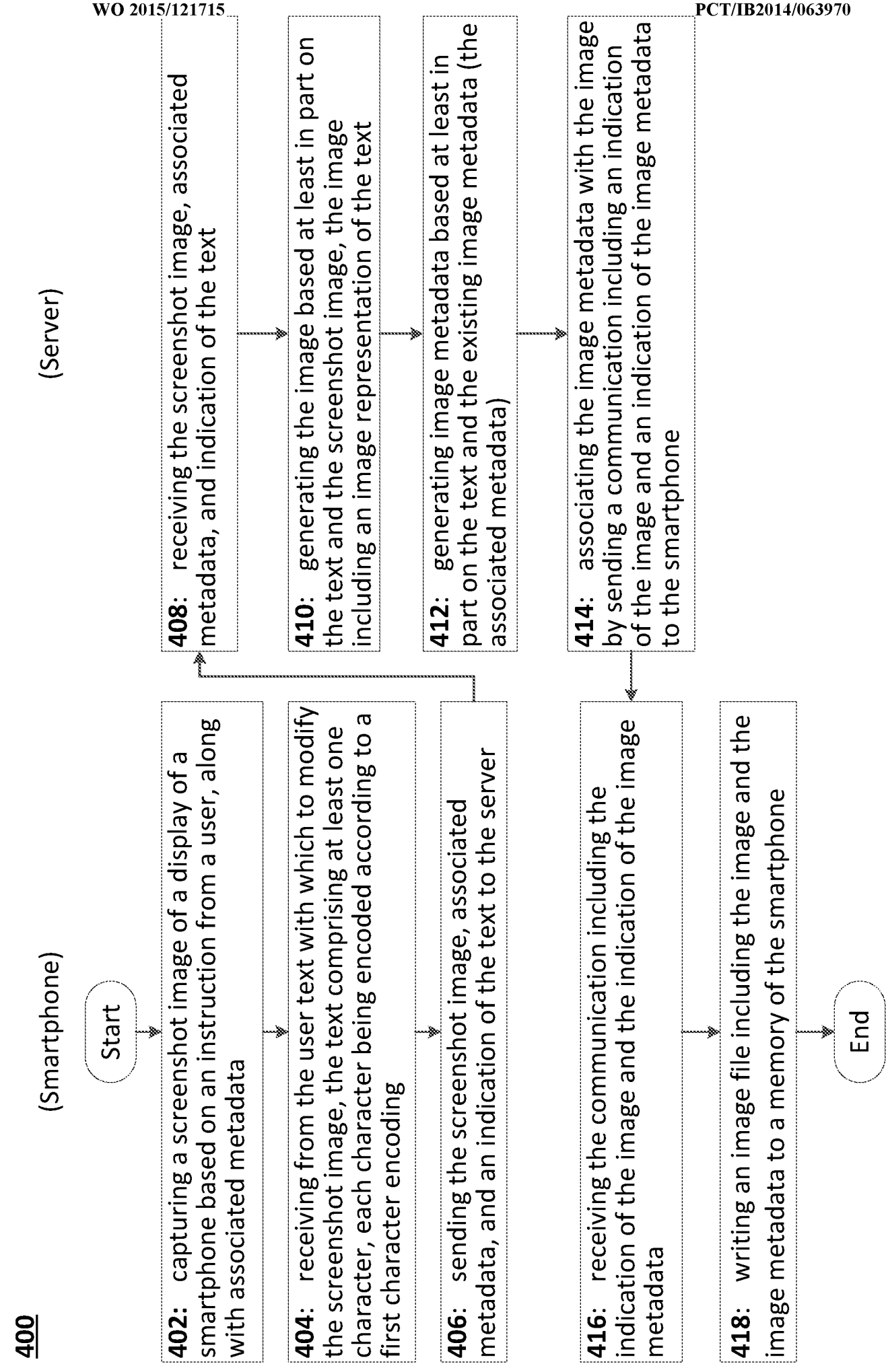
400
(Smartphone)
(Server)
Start
402: capturing a screenshot image of a display of a smartphone based on an instruction from a user, along with associated metadata
404: receiving from the user text with which to modify the screenshot image, the text comprising at least one character, each character being encoded according to a first character encoding
406: sending the screenshot image, associated metadata, and an indication of the text to the server
408: receiving the screenshot image, associated metadata, and indication of the text
410: generating the image based at least in part on the text and the screenshot image, the image including an image representation of the text
412: generating image metadata based at least in part on the text and the existing image metadata (the associated metadata)
414: associating the image metadata with the image by sending a communication including an indication of the image and an indication of the image metadata to the smartphone
416: receiving the communication including the indication of the image and the indication of the image metadata
418: writing an image file including the image and the image metadata to a memory of the smartphone
End

Fig. 8

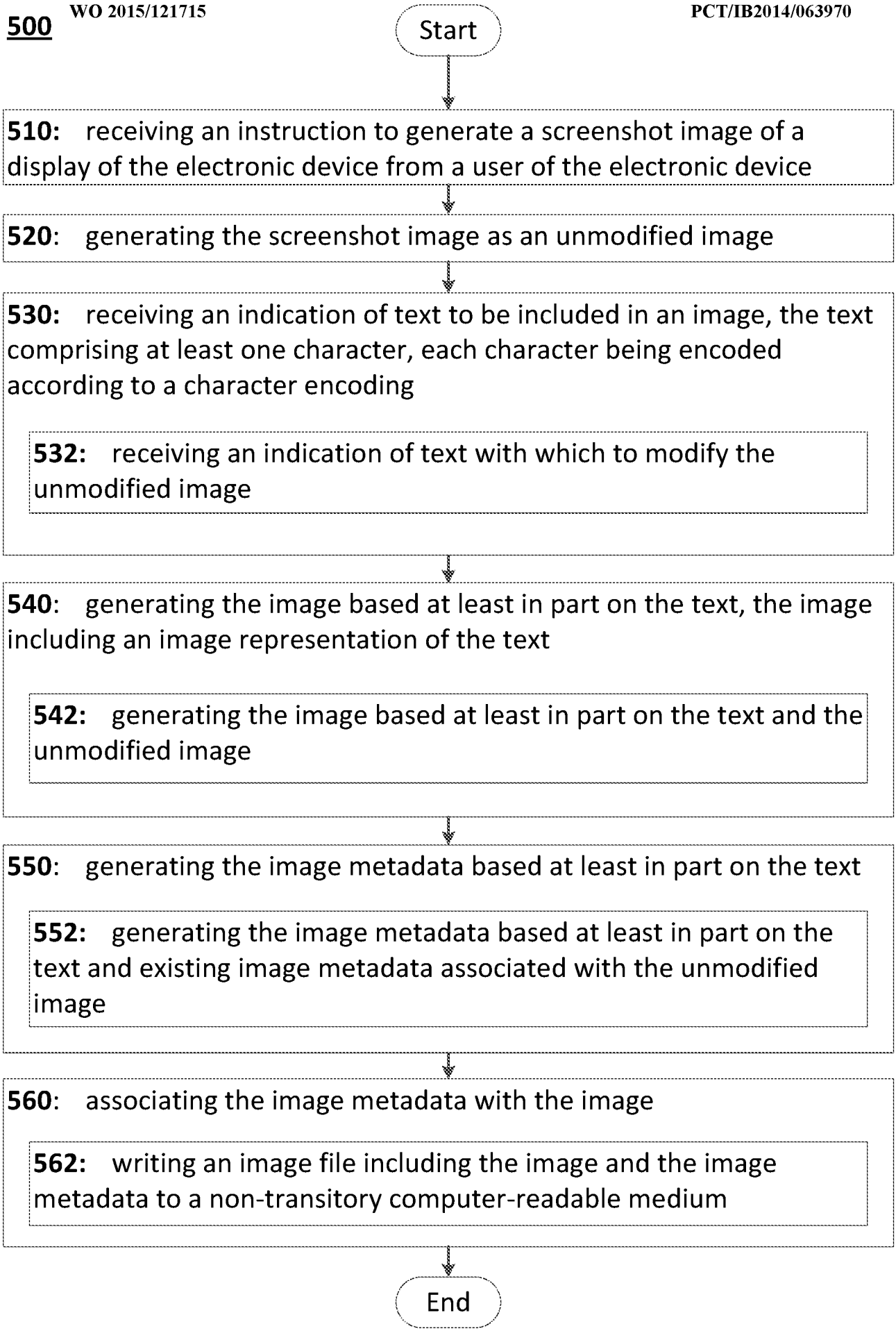
500
Start
510: receiving an instruction to generate a screenshot image of a display of the electronic device from a user of the electronic device
520: generating the screenshot image as an unmodified image
530: receiving an indication of text to be included in an image, the text comprising at least one character, each character being encoded according to a character encoding
532: receiving an indication of text with which to modify the unmodified image
540: generating the image based at least in part on the text, the image including an image representation of the text
542: generating the image based at least in part on the text and the unmodified image
550: generating the image metadata based at least in part on the text
552: generating the image metadata based at least in part on the text and existing image metadata associated with the unmodified image
560: associating the image metadata with the image
562: writing an image file including the image and the image metadata to a non-transitory computer-readable medium
End

Fig. 9

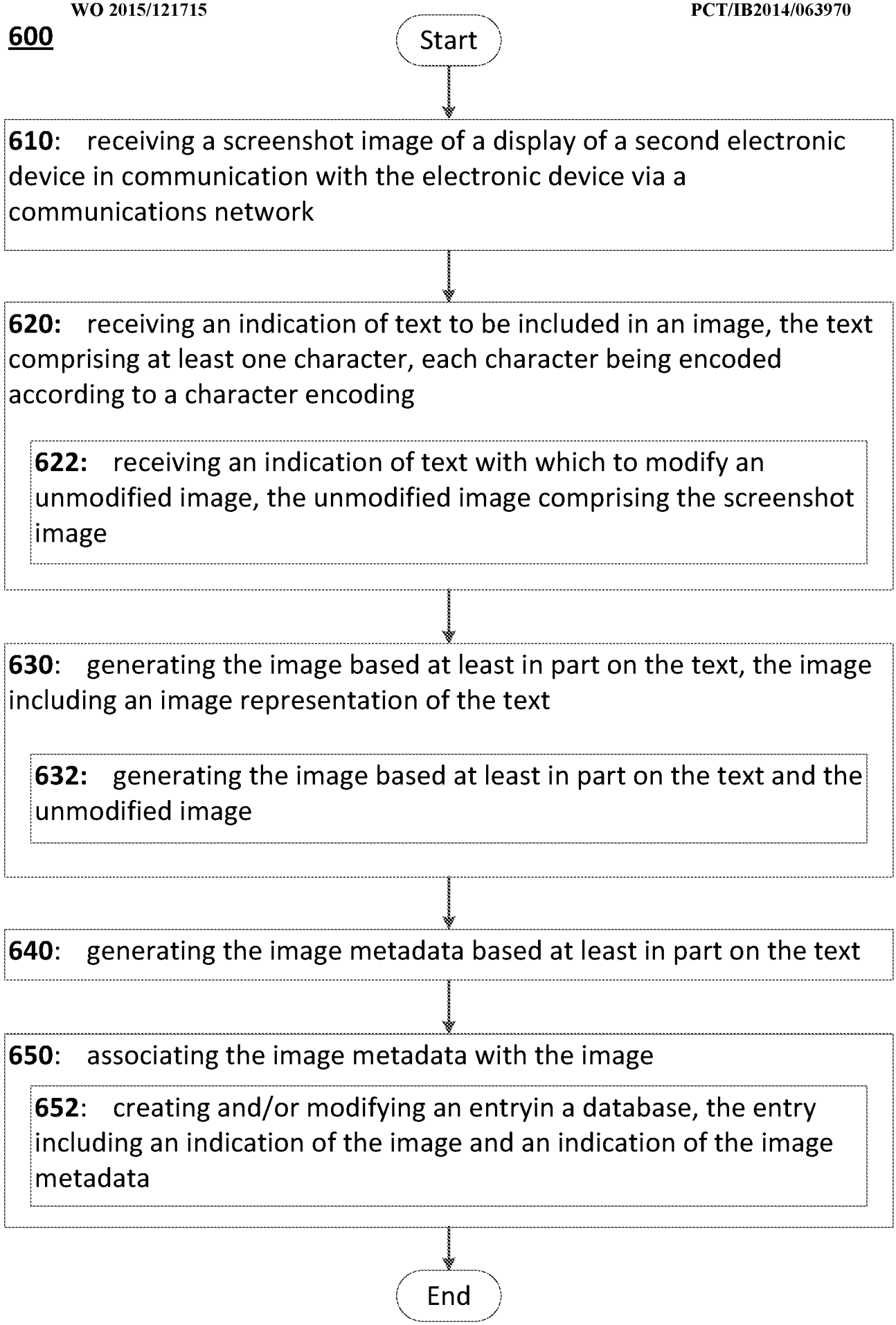
600
Start
610: receiving a screenshot image of a display of a second electronic device in communication with the electronic device via a communications network
620: receiving an indication of text to be included in an image, the text comprising at least one character, each character being encoded according to a character encoding
622: receiving an indication of text with which to modify an unmodified image, the unmodified image comprising the screenshot image
630: generating the image based at least in part on the text, the image including an image representation of the text
632: generating the image based at least in part on the text and the unmodified image
640: generating the image metadata based at least in part on the text
650: associating the image metadata with the image
652: creating and/or modifying an entryin a database, the entry including an indication of the image and an indication of the image metadata
End

Fig. 10

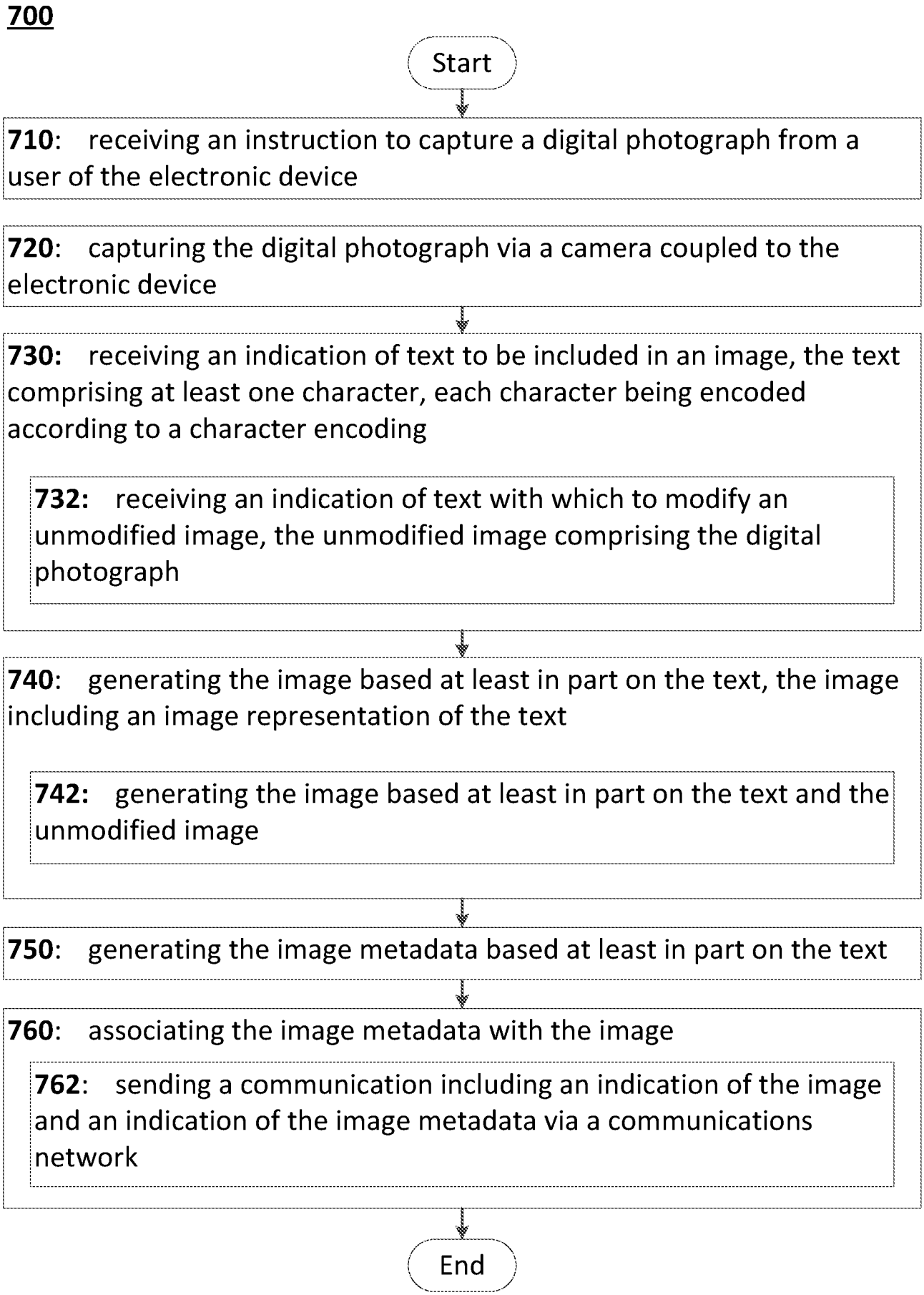
700
Start
710: receiving an instruction to capture a digital photograph from a user of the electronic device
720: capturing the digital photograph via a camera coupled to the electronic device
730: receiving an indication of text to be included in an image, the text comprising at least one character, each character being encoded according to a character encoding
732: receiving an indication of text with which to modify an unmodified image, the unmodified image comprising the digital photograph
740: generating the image based at least in part on the text, the image including an image representation of the text
742: generating the image based at least in part on the text and the unmodified image
750: generating the image metadata based at least in part on the text
760: associating the image metadata with the image
762: sending a communication including an indication of the image and an indication of the image metadata via a communications network
End

Fig. 11

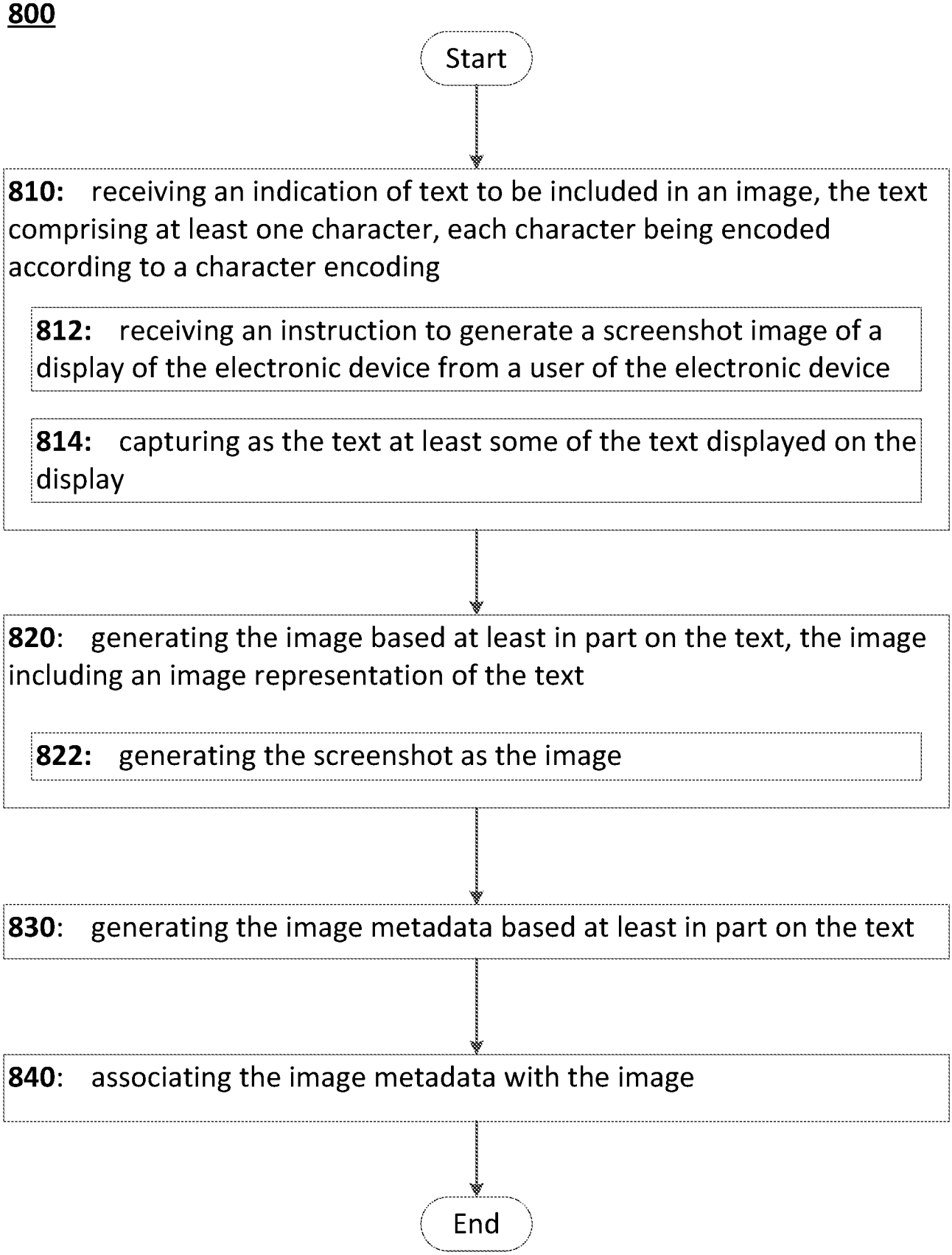
800
Start
810: receiving an indication of text to be included in an image, the text comprising at least one character, each character being encoded according to a character encoding
812: receiving an instruction to generate a screenshot image of a display of the electronic device from a user of the electronic device
814: capturing as the text at least some of the text displayed on the display
820: generating the image based at least in part on the text, the image including an image representation of the text
822: generating the screenshot as the image
830: generating the image metadata based at least in part on the text
840: associating the image metadata with the image
End

Fig. 12

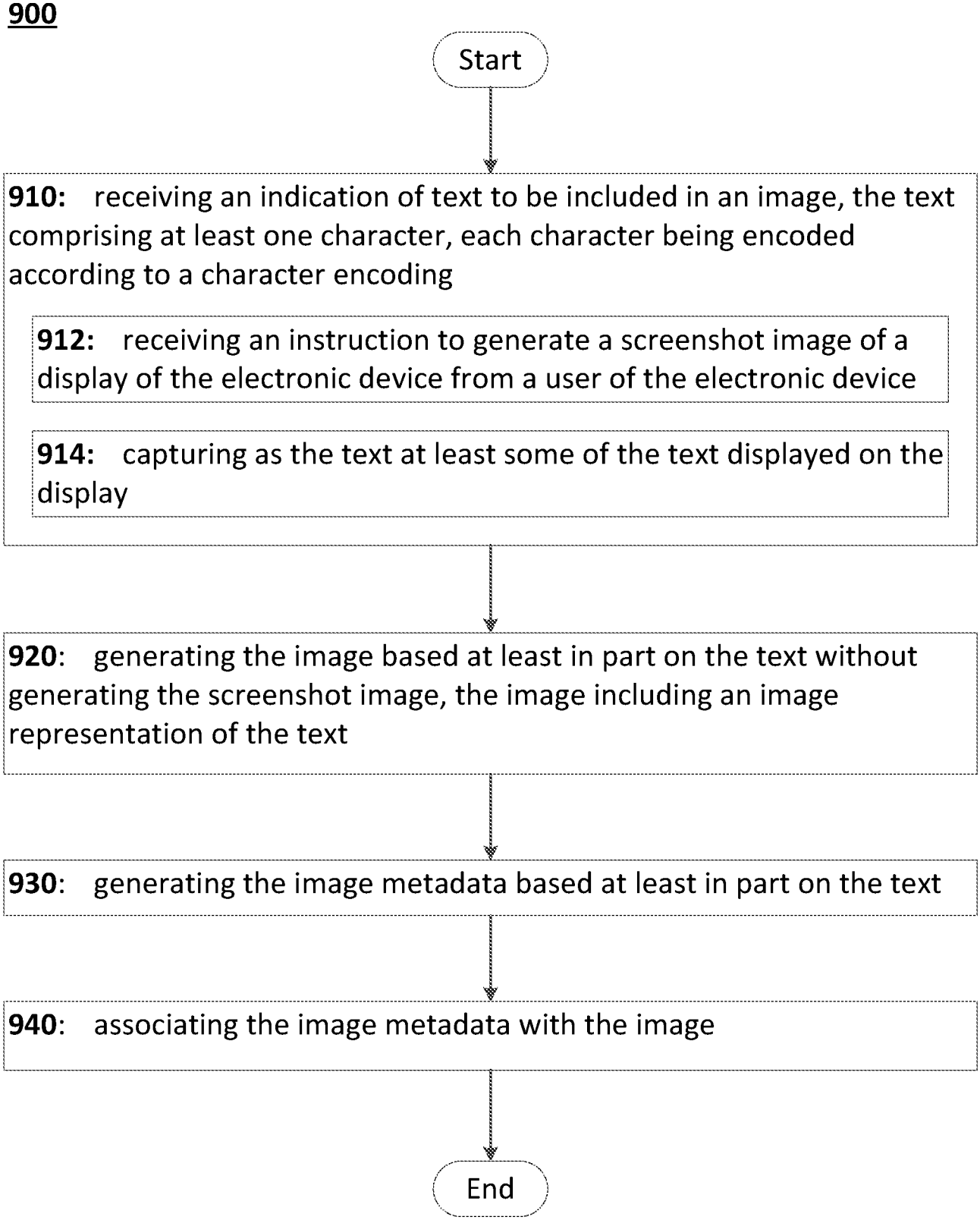
900
Start
910: receiving an indication of text to be included in an image, the text comprising at least one character, each character being encoded according to a character encoding
912: receiving an instruction to generate a screenshot image of a display of the electronic device from a user of the electronic device
914: capturing as the text at least some of the text displayed on the display
920: generating the image based at least in part on the text without generating the screenshot image, the image including an image representation of the text
930: generating the image metadata based at least in part on the text
940: associating the image metadata with the image
End

Fig. 13

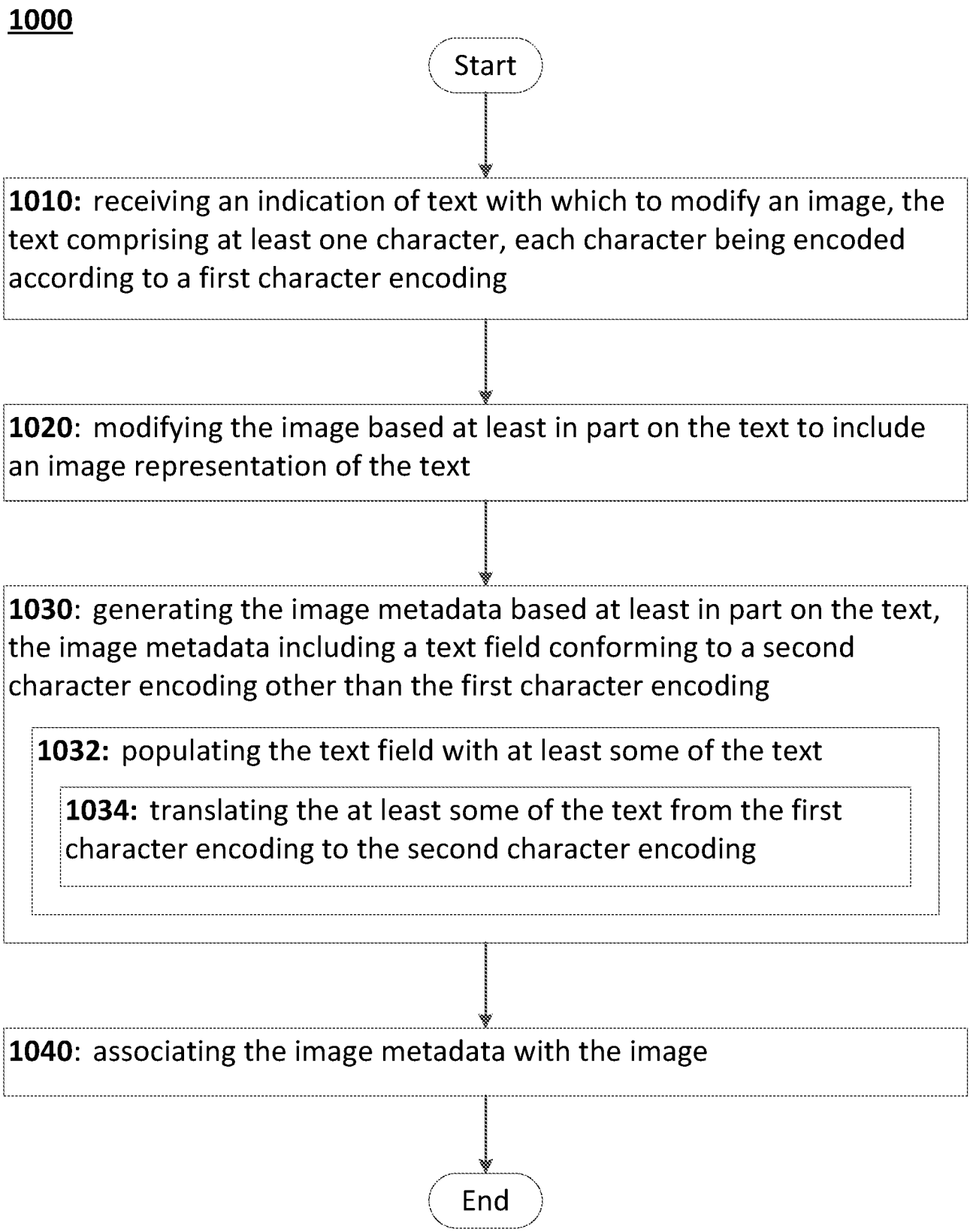
1000
Start
1010: receiving an indication of text with which to modify an image, the text comprising at least one character, each character being encoded according to a first character encoding
1020: modifying the image based at least in part on the text to include an image representation of the text
1030: generating the image metadata based at least in part on the text, the image metadata including a text field conforming to a second character encoding other than the first character encoding
1032: populating the text field with at least some of the text
1034: translating the at least some of the text from the first character encoding to the second character encoding
1040: associating the image metadata with the image
End

Fig. 14

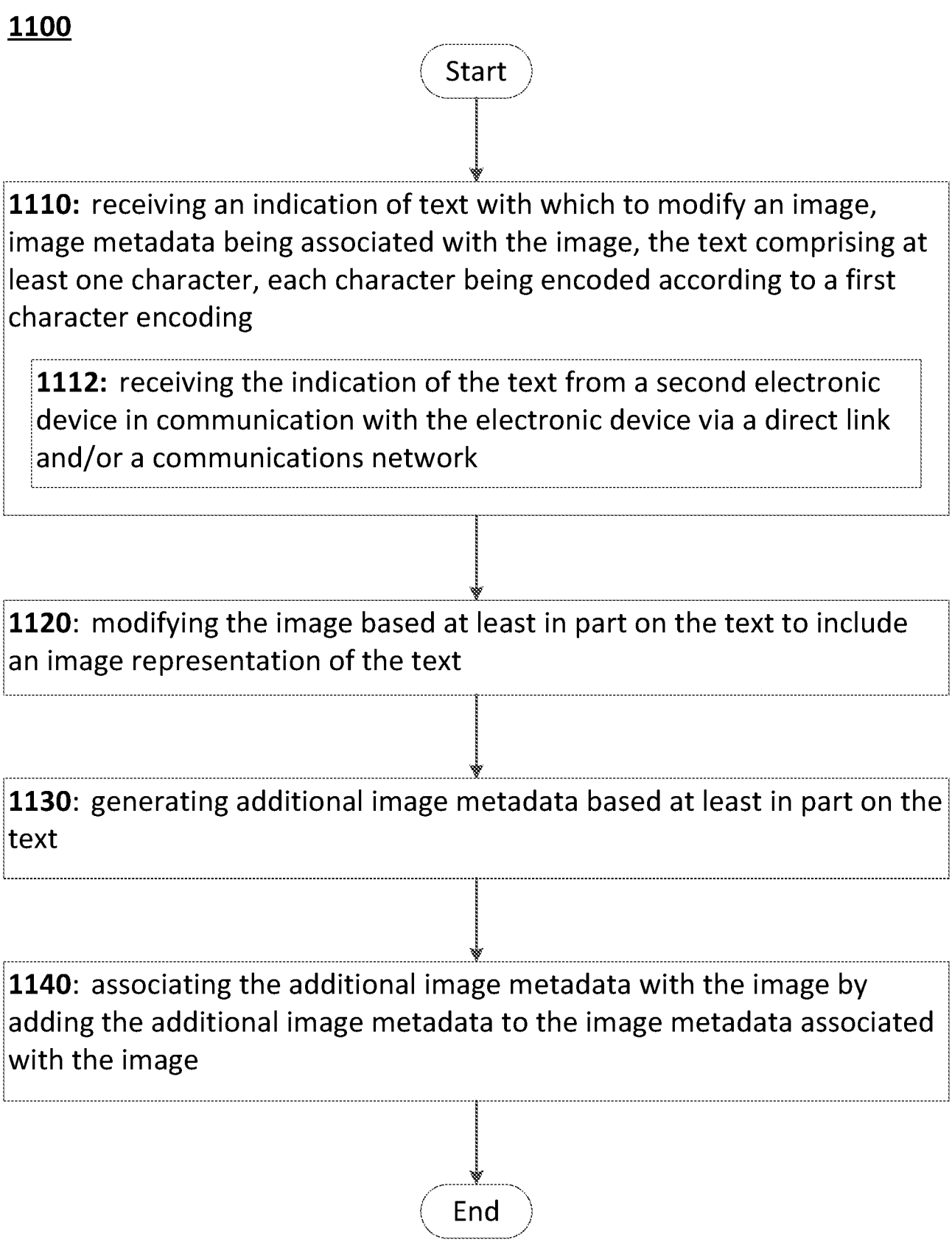
1100
Start
1110: receiving an indication of text with which to modify an image, image metadata being associated with the image, the text comprising at least one character, each character being encoded according to a first character encoding
1112: receiving the indication of the text from a second electronic device in communication with the electronic device via a direct link and/or a communications network
1120: modifying the image based at least in part on the text to include an image representation of the text
1130: generating additional image metadata based at least in part on the text
1140: associating the additional image metadata with the image by adding the additional image metadata to the image metadata associated with the image
End

Fig. 15

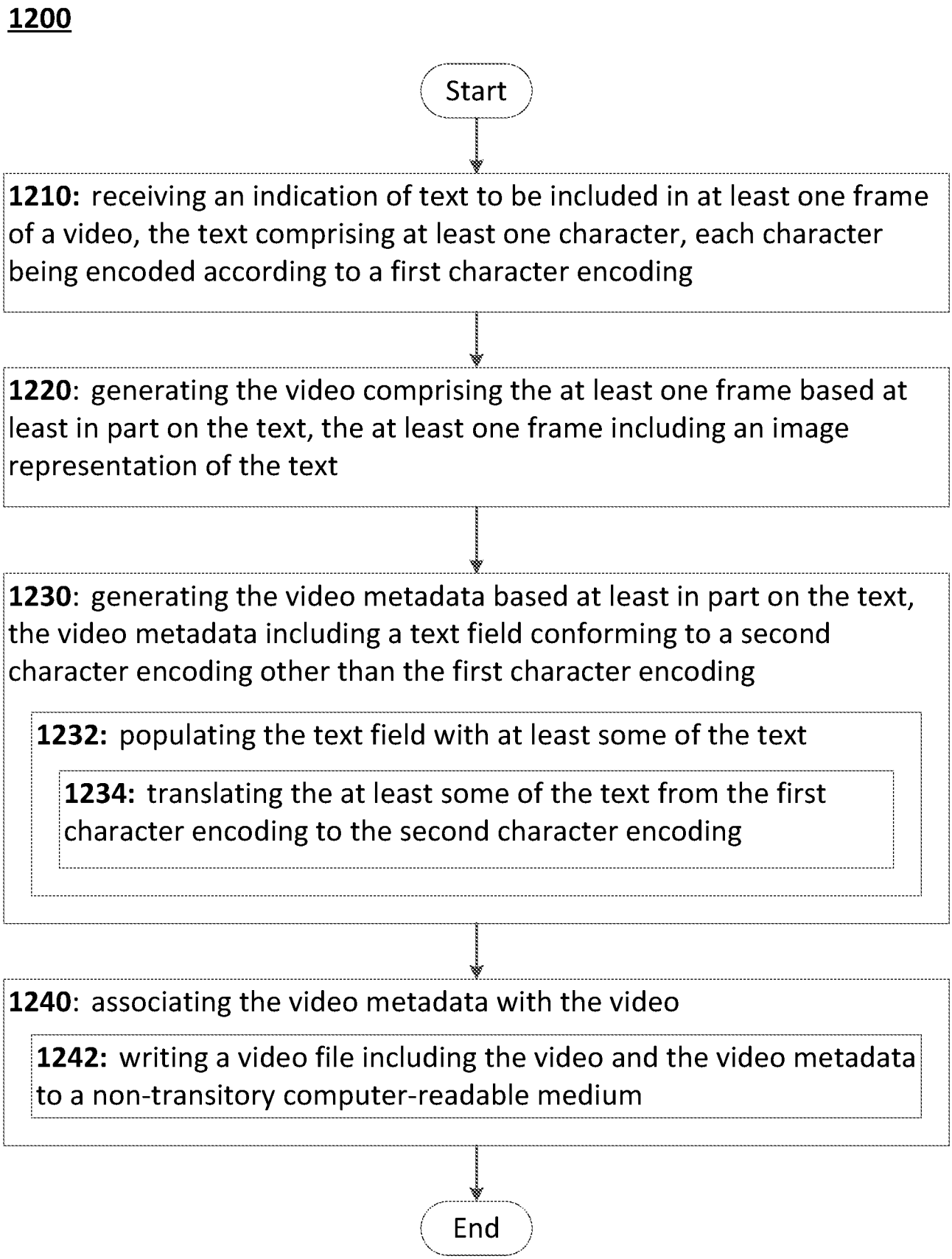
1200
Start
1210: receiving an indication of text to be included in at least one frame of a video, the text comprising at least one character, each character being encoded according to a first character encoding
1220: generating the video comprising the at least one frame based at least in part on the text, the at least one frame including an image representation of the text
1230: generating the video metadata based at least in part on the text, the video metadata including a text field conforming to a second character encoding other than the first character encoding
1232: populating the text field with at least some of the text
1234: translating the at least some of the text from the first character encoding to the second character encoding
1240: associating the video metadata with the video
1242: writing a video file including the video and the video metadata to a non-transitory computer-readable medium
End

Fig. 16

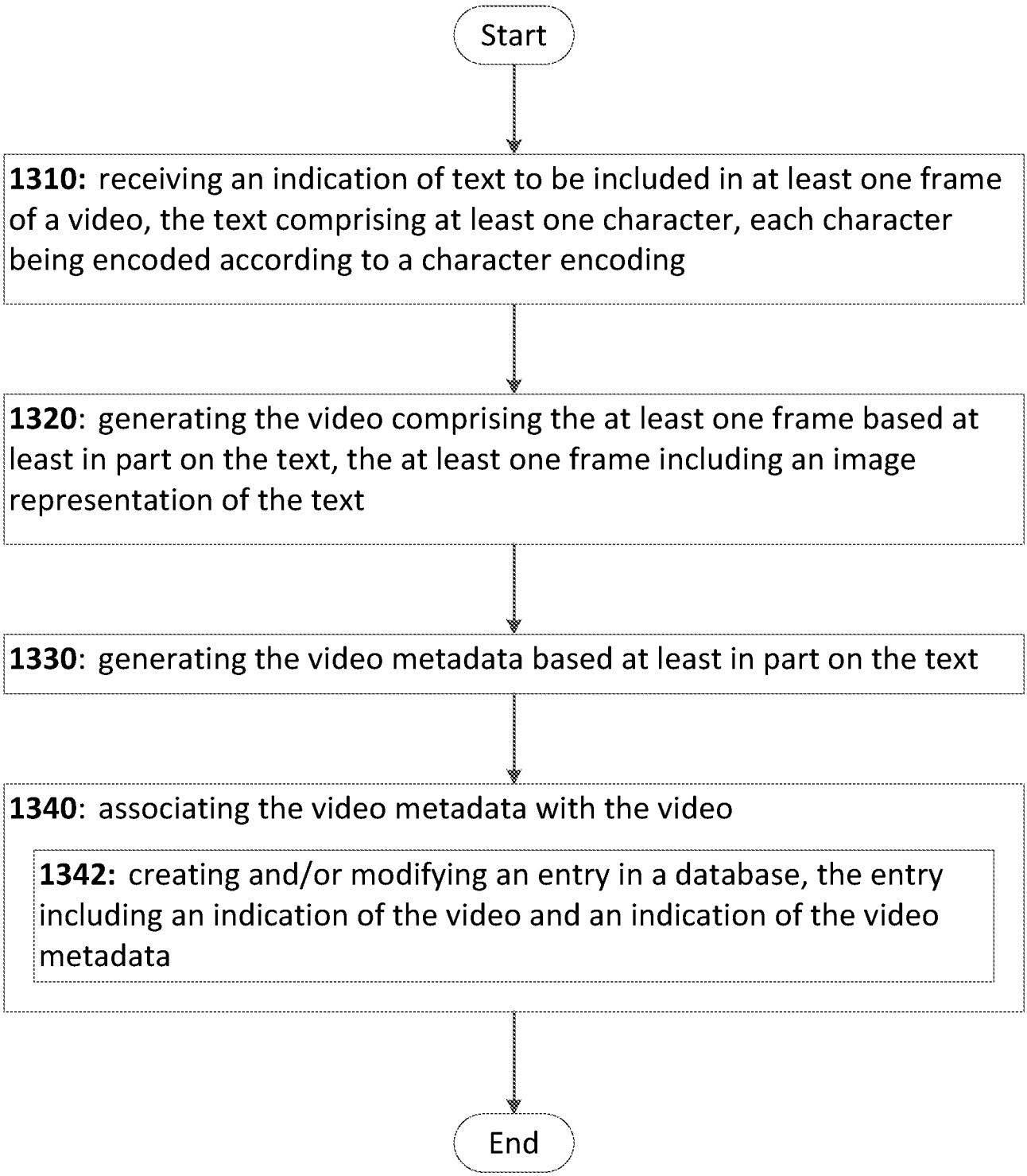
1300
Start
1310: receiving an indication of text to be included in at least one frame of a video, the text comprising at least one character, each character being encoded according to a character encoding
1320: generating the video comprising the at least one frame based at least in part on the text, the at least one frame including an image representation of the text
1330: generating the video metadata based at least in part on the text
1340: associating the video metadata with the video
1342: creating and/or modifying an entry in a database, the entry including an indication of the video and an indication of the video metadata
End

Fig. 17

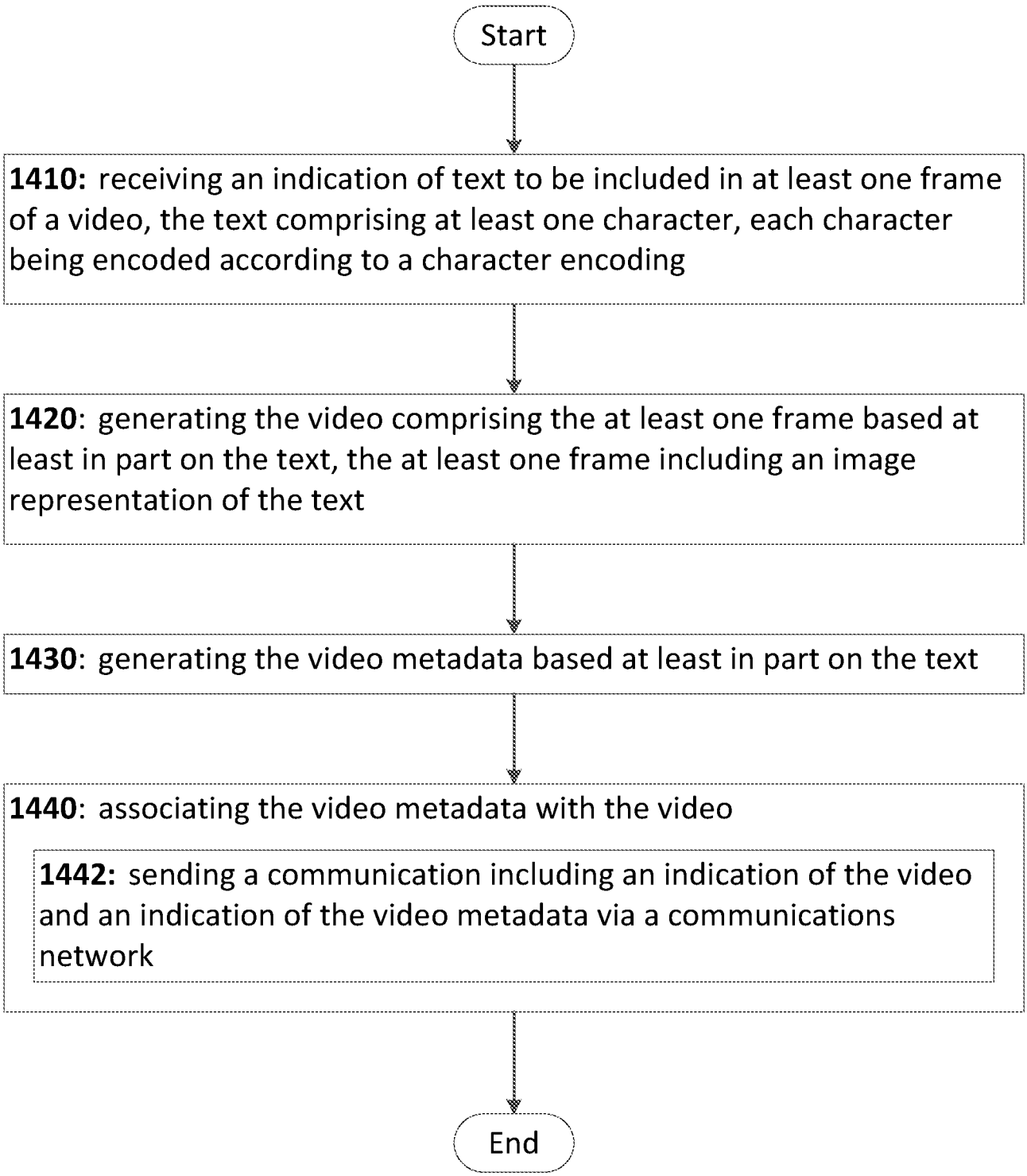
1400
Start
1410: receiving an indication of text to be included in at least one frame of a video, the text comprising at least one character, each character being encoded according to a character encoding
1420: generating the video comprising the at least one frame based at least in part on the text, the at least one frame including an image representation of the text
1430: generating the video metadata based at least in part on the text
1440: associating the video metadata with the video
1442: sending a communication including an indication of the video and an indication of the video metadata via a communications network
End

Fig. 18

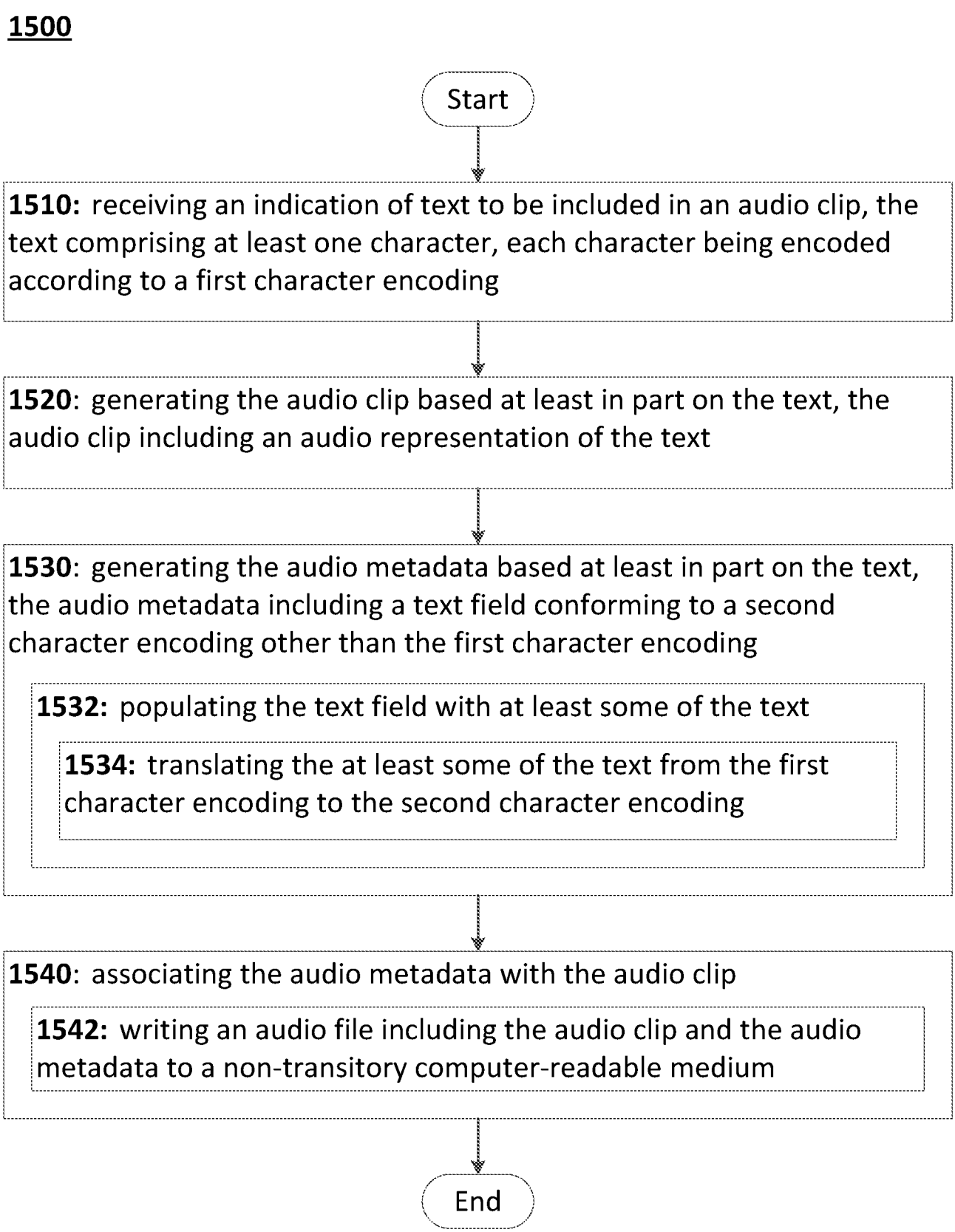
1500
Start
1510: receiving an indication of text to be included in an audio clip, the text comprising at least one character, each character being encoded according to a first character encoding
1520: generating the audio clip based at least in part on the text, the audio clip including an audio representation of the text
1530: generating the audio metadata based at least in part on the text, the audio metadata including a text field conforming to a second character encoding other than the first character encoding
1532: populating the text field with at least some of the text
1534: translating the at least some of the text from the first character encoding to the second character encoding
1540: associating the audio metadata with the audio clip
1542: writing an audio file including the audio clip and the audio metadata to a non-transitory computer-readable medium
End

Fig. 19

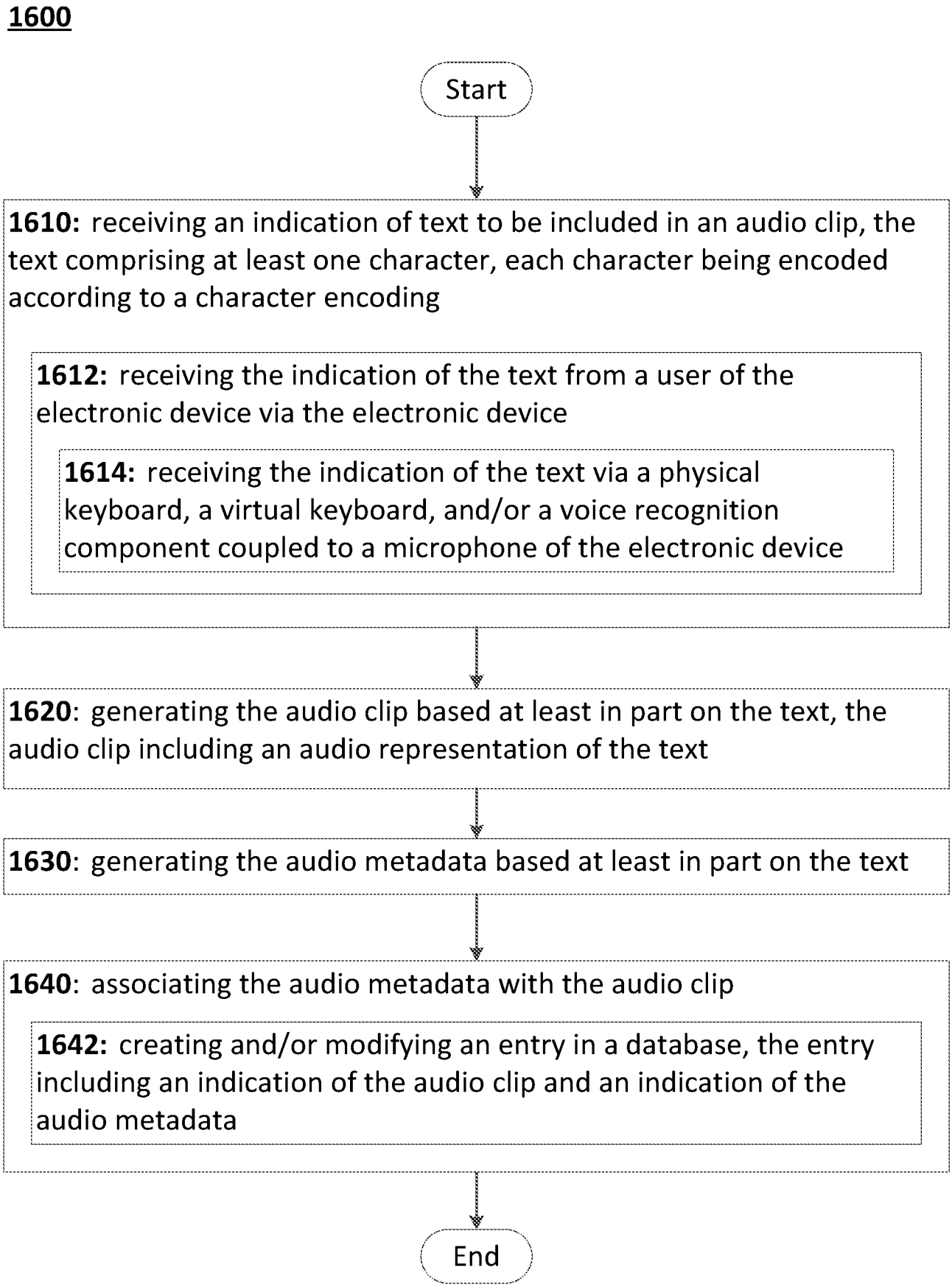
1600
Start
1610: receiving an indication of text to be included in an audio clip, the text comprising at least one character, each character being encoded according to a character encoding
1612: receiving the indication of the text from a user of the electronic device via the electronic device
1614: receiving the indication of the text via a physical keyboard, a virtual keyboard, and/or a voice recognition component coupled to a microphone of the electronic device
1620: generating the audio clip based at least in part on the text, the audio clip including an audio representation of the text
1630: generating the audio metadata based at least in part on the text
1640: associating the audio metadata with the audio clip
1642: creating and/or modifying an entry in a database, the entry including an indication of the audio clip and an indication of the audio metadata
End

Fig. 20

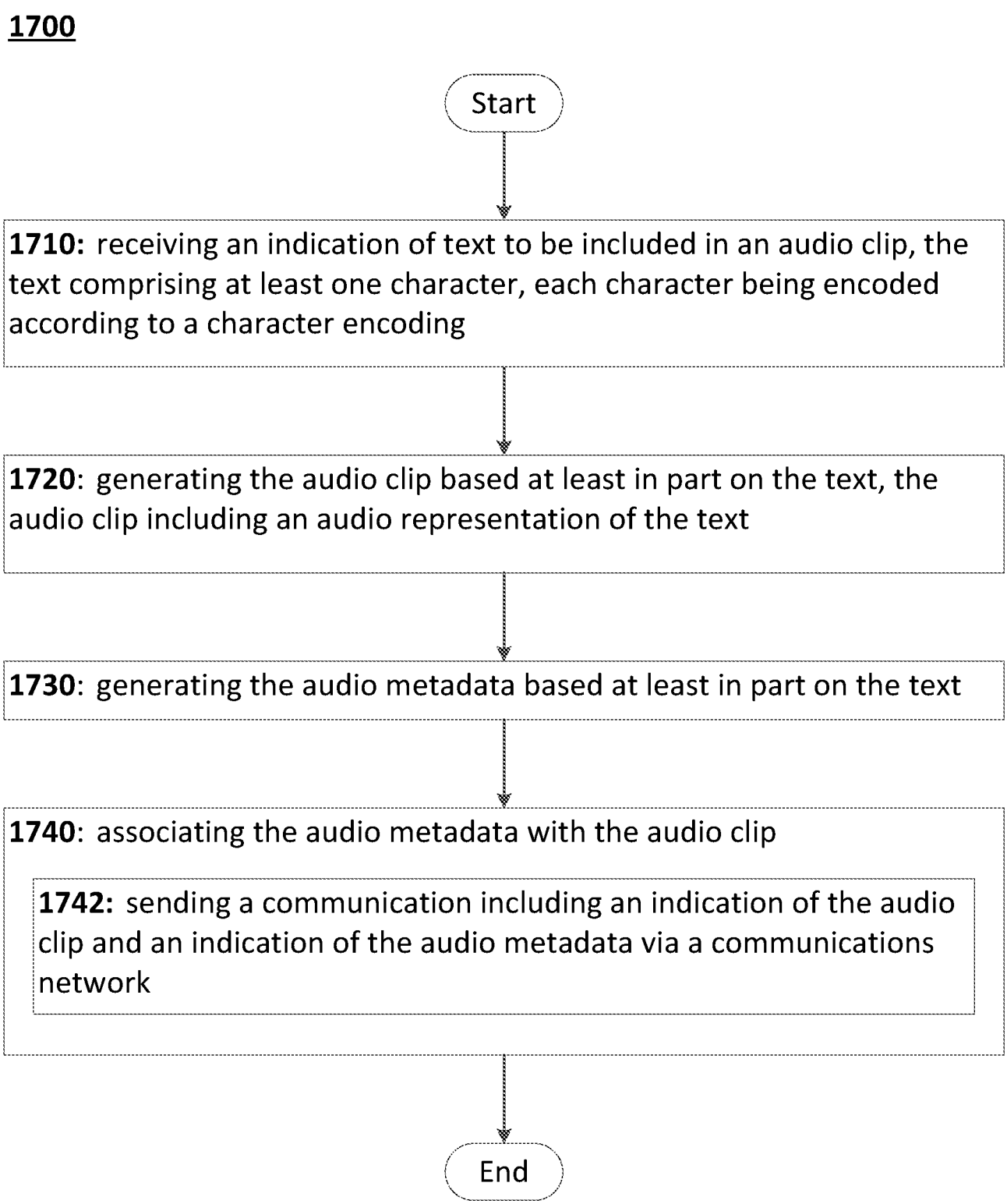
1700
Start
1710: receiving an indication of text to be included in an audio clip, the text comprising at least one character, each character being encoded according to a character encoding
1720: generating the audio clip based at least in part on the text, the audio clip including an audio representation of the text
1730: generating the audio metadata based at least in part on the text
1740: associating the audio metadata with the audio clip
1742: sending a communication including an indication of the audio clip and an indication of the audio metadata via a communications network
End

Fig. 21

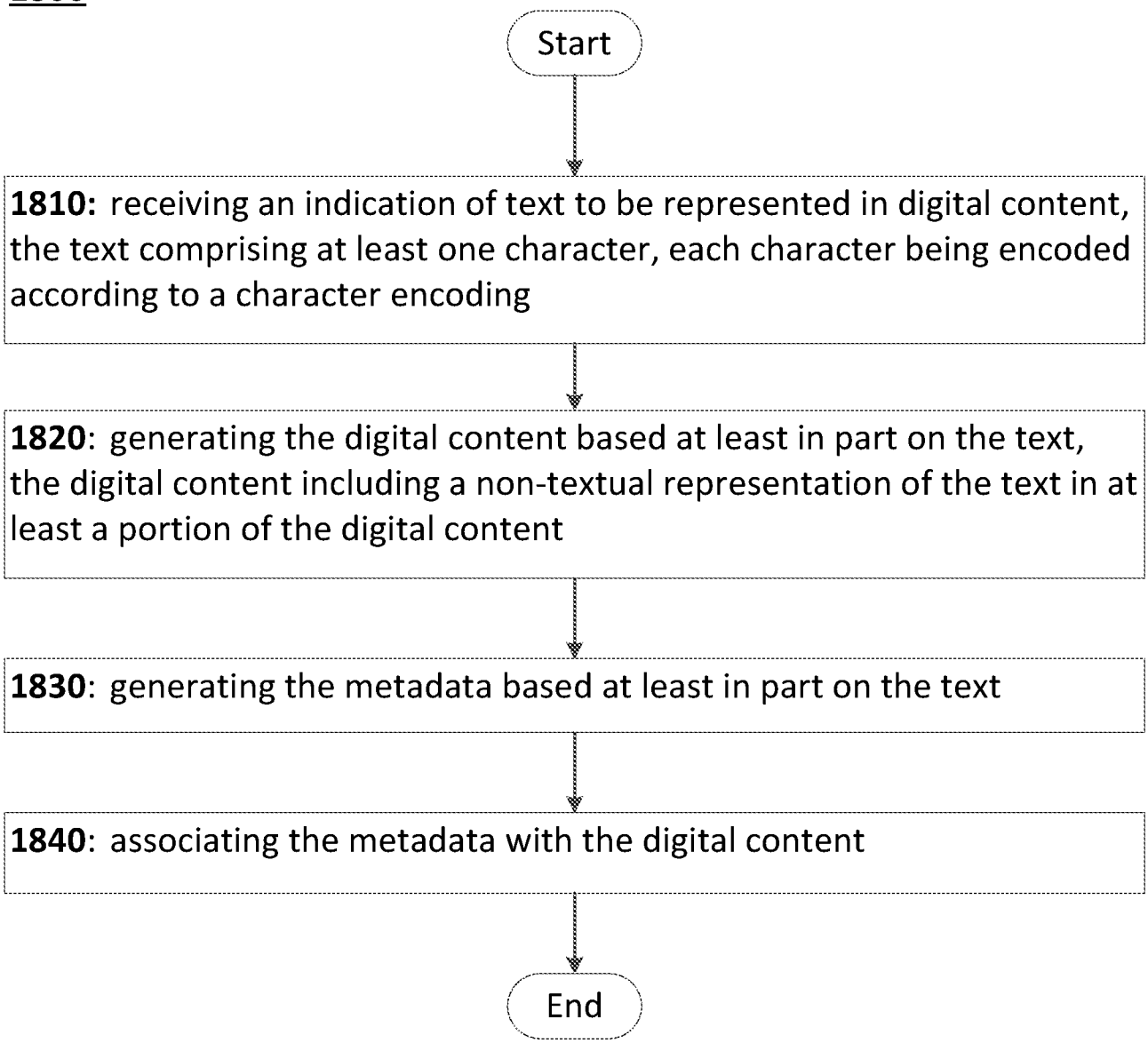
1800
Start
1810: receiving an indication of text to be represented in digital content, the text comprising at least one character, each character being encoded according to a character encoding
1820: generating the digital content based at least in part on the text, the digital content including a non-textual representation of the text in at least a portion of the digital content
1830: generating the metadata based at least in part on the text
1840: associating the metadata with the digital content
End

Fig. 22

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IB2014/063970

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - G06K 9/34 (2014.01)
CPC - G06T 3/4053 (2014.11)
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(8) - G06K 9/34, G06K 9/72, G06K 9/00, G06F 17/00, G06F 17/30 (2014.01)
USPC - 382/176, 382/165, 382/229, 1/1, 707/E17.009, 707/999.107

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
CPC - G06T 3/4053, G06K 9/36, G06K 9/3258, G06T 11/60, G06K 2209/01, G11B 27/034, G11B 27/28, G06F 17/30781, H04N 5/147, G06F 17/30253, G06F 17/3028, G06F 17/30265 (2014.11) (keyword delimited)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Orbit, Google Patents, Google Scholar, Google.
Search terms used: generate image metadata, electronic device, indication, text, character, to be included in image, encode character, image representation of text

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2004/0049734 A1 (SIMSKE) 11 March 2004 (11.03.2004), entire document	1-11, 17-28, 32-44
Y	US 2009/0164494 A1 (DODIN) 25 June 2009 (25.06.2009), entire document	1-11, 17-28, 32-44
Y	US 2012/0185066 A1 (KERN et al) 19 July 2012 (19.07.2012), entire document	4, 7-9
Y	US 2013/0219365 A1 (RAGO et al) 22 August 2013 (22.08.2013), entire document	5
Y	US 2008/0018784 A1 (BENNETT) 24 January 2008 (24.01.2008), entire document	17-22, 36, 42
Y	US 2011/0035222 A1 (SCHILLER) 10 February 2011 (10.02.2011), entire document	23-28, 37, 43
Y	US 6,549,922 B1 (SRIVASTAVA et al) 15 April 2003 (15.04.2003), entire document	21, 26, 27
A	US 2012/0114241 A1 (VINCENT et al) 10 May 2012 (10.05.2012), entire document	1-11, 17-28, 32-44
A	US 2010/0246944 A1 (YANG et al) 30 September 2010 (30.09.2010), entire document	1-11, 17-28, 32-44
A	US 2009/0006474 A1 (RICHARDSON et al) 01 January 2009 (01.01.2009), entire document	1-11, 17-28, 32-44
A	US 2008/0033983 A1 (KO) 07 February 2008 (07.02.2008), entire document	1-11, 17-28, 32-44

☐ Further documents are listed in the continuation of Box C. ☐

* 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
28 December 2014	08 JAN 2015

Name and mailing address of the ISA/US	Authorized officer:
Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Blaine R. Copenheaver PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IB2014/063970

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 12-16, 29-31
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.