



(51) International Patent Classification:

G06N 20/00 (2019.01) G06T 7/00 (2017.01)
G06N 7/00 (2006.01) G06T 7/174 (2017.01)
G06T 7/10 (2017.01)

(21) International Application Number:

PCT/US2019/022688

(22) International Filing Date:

18 March 2019 (18.03.2019)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/816,315 11 March 2019 (11.03.2019) US

(71) Applicant (for all designated States except US): **PRE-FERRED NETWORKS, INC.** [JP/JP]; Otemachi Bldg., 1-6-1 Otemachi, Chiyoda-ku, Tokyo 1000004 (JP).

(72) Inventors; and

(71) Applicants (for US only): **JIN, Yanghua** [CN/JP]; c/o Preferred Networks, Inc., Otemachi Bldg., 1-6-1 Otemachi, Chiyoda-ku, Tokyo 1000004 (JP). **ZHU, Huachun** [CN/JP]; c/o Preferred Networks, Inc., Otemachi Bldg., 1-6-1 Otemachi, Chiyoda-ku, Tokyo 1000004 (JP). **TIAN, Yingtao** [CN/US]; c/o Preferred Networks America, Inc., 2150 Shattuck Ave., Suite 200, Berkeley, California 94704 (US).

(74) Agent: **PARIS, Herman**; 1054 31st Street, N.W. Suite 400, Washington, District of Columbia 20007 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN,

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

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

Declarations under Rule 4.17:

— as to non-prejudicial disclosures or exceptions to lack of novelty (Rule 4.17(v))

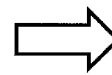
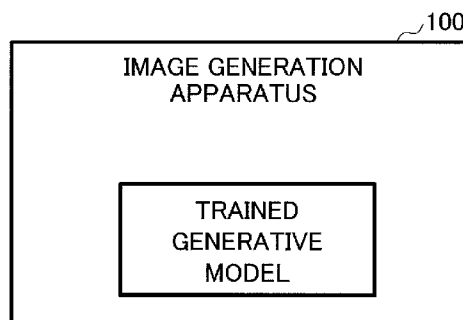
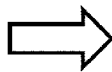
Published:

— with international search report (Art. 21(3))
— before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))
— upon request of the applicant, before the expiration of the time limit referred to in Article 21(2)(a)

(54) Title: IMAGE GENERATION METHOD, IMAGE GENERATION APPARATUS, IMAGE GENERATION SYSTEM AND PROGRAM

FIG.5

TO-BE-FUSED
LATENT
INFORMATION



FUSION
IMAGE

(57) Abstract: A novel image generation technique is disclosed. One aspect of the present disclosure relates to an image generation method including acquiring first latent information of a first image and second latent information of a second image, fusing the first latent information and the second latent information to generate fusion latent information, and supplying the fusion latent information to a trained generative model to generate a fusion image.



[DESCRIPTION]

[Title]

IMAGE GENERATION METHOD, IMAGE GENERATION APPARATUS,
IMAGE GENERATION SYSTEM AND PROGRAM

[Technical Field]

The present application claims priority to U.S. Provisional Patent Application No. 62/816,315 filed on March 11, 2019, the contents of which are incorporated herein by reference in their entirety.

The present disclosure relates to an image generation method, an image generation apparatus, an image generation system and a program.

[Background Art]

In networks such as the Internet, sites or platforms for providing various tools are provided. For example, some platforms are operated on the Internet for providing tools to enable users to generate digital images, such as an avatar image and a character image, by means of provided parts and to edit, release or exchange the generated digital images.

For example, in CryptoKitties, users are allowed to create a new kitten whose characteristics are

a recombination of those of its parents and trade users' possessed kitten images with use of cryptocurrencies.

However, the kitten images created at CryptoKitties (see <https://www.cryptokitties.co/>) are combinations of parts (for example, eyes, ears and so on) of the parent cats, and there are few unique characteristics and variations.

[Summary]

One object of the present disclosure is to provide a novel image generation technique.

In light of the above object, one aspect of the present disclosure relates to an image generation method that includes acquiring first latent information of a first image and second latent information of a second image, fusing the first latent information and the second latent information to generate fusion latent information, and supplying the fusion latent information to a trained generative model to generate a fusion image.

[Brief Description of Drawings]

Fig. 1 is a schematic diagram that depicts an image generation operation according to one embodiment of the present disclosure.

Fig. 2 is a diagram that depicts a digital

image according to one embodiment of the present disclosure.

Fig. 3 is a diagram that depicts a digital image according to one embodiment of the present disclosure.

Fig. 4 is a diagram that depicts a digital image according to one embodiment of the present disclosure.

Fig. 5 is a schematic diagram that depicts an image generation apparatus according to one embodiment of the present disclosure.

Fig. 6 is a block diagram that depicts a functional arrangement of an image generation apparatus according to one embodiment of the present disclosure.

Fig. 7 is a diagram that depicts image information on a platform according to one embodiment of the present disclosure.

Fig. 8 is a diagram that depicts image information on a platform according to one embodiment of the present disclosure.

Fig. 9 is a diagram that depicts fusion of latent information according to one embodiment of the

present disclosure.

Fig. 10 is a diagram that depicts a fusion image according to one embodiment of the present disclosure.

Fig. 11 is a diagram that depicts fusion of latent information in accordance with a genetic operation according to one embodiment of the present disclosure.

Figs. 12A to 12C are diagrams that depicts fusion of latent information according to another embodiment of the present disclosure.

Figs. 13A to 13C are diagrams that depicts fusion of latent information according to another embodiment of the present disclosure.

Fig. 14 is a flowchart that depicts an image generation operation according to one embodiment of the present disclosure.

Fig. 15 is a flowchart that depicts a training operation in accordance with GANs (Generative Adversarial Networks) according to one embodiment of the present disclosure.

Fig. 16 is a schematic diagram that depicts an image generation operation according to another

embodiment of the present disclosure.

Fig. 17 is a flowchart that depicts an image generation operation according to another embodiment of the present disclosure.

Fig. 18 is a schematic diagram that depicts an image generation system according to one embodiment of the present disclosure.

Fig. 19 is a block diagram that depicts a hardware arrangement of an image generation apparatus, a training apparatus and a user apparatus according to one embodiment of the present disclosure.

[Description of Embodiments]

Embodiments of the present disclosure are described below with reference to the drawings. In the following embodiments, an image generation apparatus and an image generation method for generating digital images are described.

Outline of Present Disclosure

An image generation apparatus according to embodiments of the present disclosure fuses latent information (for example, genetic codes, attributes or the like) of two images to be fused in accordance with a

predetermined operation (for example, a genetic operation or the like) to generate fusion latent information and supplies the generated fusion latent information to a trained generative model to generate a fusion image of the two images.

Specifically, a character image as illustrated in FIG. 1 is characterized by the latent information including a code (for example, a genetic code, which may be referred to as a noise) and an attribute (for example, a hair style, a hair color, an eye color, a skin color, an expression, an attachment such as glasses or a hat, or the like, all or a part of which may be used). For example, an image generation apparatus according to the embodiments described below may fuse the latent information of two to-be-fused character images as illustrated in FIGS. 2 and 3 and supply a machine learning model such as a generator trained in accordance with GANs (Generative Adversarial Networks) to generate a fusion image as illustrated in FIG. 4.

According to the present disclosure, the image generation apparatus can generate a variety of unique fusion images with succession of codes and attributes of the latent information of both the images rather than simple combinations of parts of the incoming images.

Also, totally harmonious fusion images can be generated with use of the generative model.

Image Generation Apparatus

First of all, an image generation apparatus according to one embodiment of the present disclosure will be described with reference to FIGS. 5 to 11. FIG. 5 is a schematic diagram for illustrating an image generation apparatus according to one embodiment of the present disclosure.

As illustrated in FIG. 5, an image generation apparatus 100 has a trained generative model. Upon acquiring latent information of two to-be-fused images, the image generation apparatus 100 fuses the acquired latent information, in accordance with a predetermined operation, to generate fusion latent information. Then, the image generation apparatus 100 supplies the generated fusion latent information to the generative model to generate a fusion image.

Specifically, the image generation apparatus 100 may be for implementing a platform for image generation and provide the platform via a website, for example. In one example, when a user logging in the platform indicates two to-be-fused images, the image

generation apparatus 100 acquires latent information of the user's indicated two images and fuses the acquired latent information in accordance with a predetermined operation. For example, the operation may be a predetermined combination operation on the latent information, for example, a genetic operation such as crossover, mutation and selection, an arithmetic operation, a logical operation or the like. Then, the image generation apparatus 100 may use the trained generative model to generate a fusion image from the fusion latent information and provide the generated fusion image on the platform.

Note that although the trained generative model is provided in the image generation apparatus 100 in the illustrated embodiment, the present disclosure is not limited to it, and the trained generative model may be located at an external apparatus that is communicatively connected to the image generation apparatus 100, for example. In this case, the image generation apparatus 100 may send the fusion latent information to the external apparatus and acquire the fusion image generated by the trained generative model from the external apparatus.

FIG. 6 is a block diagram for illustrating a

functional arrangement of an image generation apparatus according to one embodiment of the present disclosure. As illustrated in FIG. 6, the image generation apparatus 100 includes a latent information acquisition unit 110, a latent information fusion unit 120 and a fusion image generation unit 130.

The latent information acquisition unit 110 acquires respective latent information of two to-be-fused images. For example, the to-be-fused image may be a bitmap image such as an avatar image or a character image and be associated with the latent information for characterizing an object, such as an avatar, a character, or the like. The latent information used herein may be information that may be put into a latent variable and inferred from observation data (for example, image data) through a model or the like.

Specifically, the latent information may include a code and an attribute. Here, the code may represent a characteristic (for example, a body skeleton, a shape of a facial part or the like) specific to an object of interest, and the attribute may represent a variable characteristic (for example, a hair style, an expression or the like) of the object. For example, as illustrated in FIGS. 7 and 8, when a user chooses two

images (#596506 and #529690) as to-be-fused images via a user interface where a code represented as a barcode are displayed together with various attributes and sub attributes (that is, attributes provided in the image, although they do not appear in the image) in association with the object image, the latent information acquisition unit 110 acquires codes (which are illustrated as $(Xc1, Xc2, \dots, Xc(n-1), Xcn)$ and $(Yc1, Yc2, \dots, Yc(n-1), Ycn)$ in the illustrated example) and attributes (which are illustrated as $(Xa1, Xa2, \dots, Xa(m-1), Xam)$ and $(Ya1, Ya2, \dots, Ya(m-1), Yam)$ in the illustrated example) associated with these two images and provides respective latent information of the images, which is composed of the acquired code and attributes, to the latent information fusion unit 120.

The latent information fusion unit 120 fuses the latent information of the two images to generate fusion latent information. For example, upon acquiring the codes $(Xc1, Xc2, \dots, Xc(n-1), Xcn)$ and $(Yc1, Yc2, \dots, Yc(n-1), Ycn)$ and the attributes $(Xa1, Xa2, \dots, Xa(m-1), Xam)$ and $(Ya1, Ya2, \dots, Ya(m-1), Yam)$ from the latent information acquisition unit 110, the latent information fusion unit 120 performs a predetermined fusion operation on the acquired codes $(Xc1, Xc2, \dots,$

$X_{c(n-1)}, X_{cn}$) and $(Y_{c1}, Y_{c2}, \dots, Y_{c(n-1)}, Y_{cn})$ and attributes $(X_{a1}, X_{a2}, \dots, X_{a(m-1)}, X_{am})$ and $(Y_{a1}, Y_{a2}, \dots, Y_{a(m-1)}, Y_{am})$ to generate fusion latent information composed of $(Z_{c1}, Z_{c2}, \dots, Z_{c(n-1)}, Z_{cn})$ and $(Z_{a1}, Z_{a2}, \dots, Z_{a(m-1)}, Z_{am})$, as illustrated in FIG. 9.

Here, the fusion operation may be performed on a per-code basis and a per-attribute basis as illustrated. Alternatively, in other embodiments, the fusion operation may be performed on a per-latent information basis. Then, the latent information fusion unit 120 provides the generated fusion latent information to the fusion image generation unit 130. Also, the to-be-fused images may be unnecessarily two different images and may be the same image or three or more different images.

The fusion image generation unit 130 supplies the fusion latent information to the trained generative model to generate a fusion image. For example, upon acquiring the fusion latent information $(Z_{c1}, Z_{c2}, \dots, Z_{c(n-1)}, Z_{cn})$ and $(Z_{a1}, Z_{a2}, \dots, Z_{a(m-1)}, Z_{am})$ from the latent information fusion unit 120, the fusion image generation unit 130 supplies the acquired fusion latent information to the trained generative model to generate the fusion image as illustrated in FIG. 10 as its output. The generated fusion image is associated with the fusion

latent information composed of the code ($Z_{c1}, Z_{c2}, \dots, Z_{c(n-1)}, Z_{cn}$) and the attribute ($Z_{a1}, Z_{a2}, \dots, Z_{a(m-1)}, Z_{am}$). Also, the fusion image may be displayed in association with images of its two origins having a parent-child relationship. In addition, images of ancestors of the fusion image may be also displayed in the form of a family tree.

In one embodiment, the trained generative model may be a generator trained in accordance with GANs (Generative Adversarial Networks). For example, a training apparatus 200 may pre-train a generator and a discriminator, which are implemented as neural networks, as the GANs and provide them to the image generation apparatus 100.

In one exemplary training operation, a dataset of training images is provided to the training apparatus 200, and the training apparatus 200 supplies a code such as a random number to the generator to acquire an image as its output. Next, the training apparatus 200 supplies either the image generated by the generator or a training image in the dataset to the discriminator to acquire, as its output, a discrimination result indicative of whether the incoming image is the image generated by the generator or the training image. Then, the training

apparatus 200 updates parameters for the discriminator so that the discriminator may make correct discrimination. Also, the training apparatus 200 updates parameters for the generator so that the discriminator may make incorrect discrimination. If a predetermined termination condition such as completion of the above-stated operation on a predetermined number of incoming data is satisfied, the training apparatus 200 provides the finally acquired generator to the image generation apparatus 100 as the trained generative model.

Also, in another exemplary training operation, a dataset of training images with attributes is provided to the training apparatus 200, and the training apparatus 200 supplies a code such as a vector of random numbers and a list of attributes to the generator to acquire an image as its output. Next, the training apparatus 200 supplies either the image generated by the generator or a training image in the dataset to the discriminator to acquire as its output a discrimination result indicative of not only whether the incoming image is the image generated by the generator or the training image but also what the attributes are. Then, the training apparatus 200 updates parameters for the discriminator so that the discriminator may make correct discrimination. In

addition, the training apparatus 200 updates parameters for the generator so that the discriminator may make incorrect discrimination. If a predetermined termination condition such as completion of the above-stated operation on a predetermined number of incoming data is satisfied, the training apparatus 200 provides the finally acquired generator to the image generation apparatus 100 as the trained generative model.

The trained generative model according to the present disclosure is not limited to the above-stated generator or neural network trained in accordance with the GANs and may be any type of machine learning model trained in accordance with any other appropriate training scheme.

In one embodiment, the latent information fusion unit 120 may perform a genetic operation on the acquired latent information of the two images to generate the fusion latent information. For example, as illustrated in FIG. 11, the latent information fusion unit 120 may generate a fusion code and a fusion attribute through crossover of respective codes and attributes. In the example as illustrated, code Xc and attribute Xa in the to-be-fused latent information X are "0.2", "1.4", "0.4", ... and "blond hair", "with glasses",

"smile", ..., respectively, and code Yc and attribute Ya in the to-be-fused latent information Y are "-0.2", "0.9", "0.8", ... and "green hair", "without glasses", "blush", ..., respectively.

Then, the latent information fusion unit 120 may perform crossover operations on elements of the codes Xc and Yc to generate code Zc "-0.2", "1.4", "0.4", ... in the fusion latent information Z, as illustrated. Also, the latent information fusion unit 120 may perform crossover operations on elements of the attributes Xa and Ya to generate attribute Za "green hair", "with glasses", "smile", ..., as illustrated. Here, the above-stated crossover operations are performed on a per-element basis, but the crossover operations according to the present embodiment are not limited to the above and may be performed per combination of an arbitrary number of elements.

Note that the genetic operations according to the present embodiment may include, but not limited to the crossover, other genetic operations such as mutation and selection. For example, in the mutation, the latent information fusion unit 120 may set values other than those for elements of the to-be-fused latent information X and Y to one or more elements of the fusion latent

information Z. Introduction of the mutation to the fusion operation can increase variations of the generated fusion latent information.

Also, in other embodiments, the latent information fusion unit 120 may generate the fusion latent information by averaging the acquired latent information of two images. For example, for elements of the latent information that are represented as numerical values, for example, "0.2" and "-0.2" as illustrated in FIG. 1, the corresponding elements of the fusion latent information may be determined as an average value of the corresponding elements of two images. In this case, for elements having alternative values (presence of attachments), for example, "with glasses" and "without glasses", the corresponding elements of the fusion latent information may be randomly determined as one of the two.

Also, in still further embodiments, the latent information fusion unit 120 may perform a logical operation or an arithmetic operation on the acquired latent information of two images to generate the fusion latent information. For example, the logical operation may include a logical OR operation, a logical AND operation, a logical exclusive OR operation, a logical NOT operation or the like. For example, the arithmetic

operation may include addition, subtraction, multiplication, division or the like. For example, combinations of the arithmetic operations such as an arithmetic mean and a harmonic mean may be included. In addition, conversion (for example, an exponential operation or a logarithmic operation) or fluctuation (for example, a random noise) may be applied to the latent information of the two images, and then the arithmetic operation may be performed on the applied latent information. Alternatively, the arithmetic operation may be performed on the latent information of the two images, and then the conversion or the fluctuation may be applied (for example, a geometric mean or an exponential mean).

Also, in still further embodiments, the latent information fusion unit 120 may perform a combination of any one or more of the above-stated operations on the acquired latent information of the two images to generate the fusion latent information.

In one embodiment, the latent information may further include a sub attribute. The above-stated attribute (referred to as a main attribute hereinafter) is an attribute having a higher likelihood of its occurrence in a fusion image whereas the sub attribute is an attribute that does not appear in the origin image but

is likely to appear in the fusion image with a lower occurrence likelihood. For example, the sub attribute may have similar contents (for example, a hair style, an expression or the like). Either the main attribute or the sub attribute may be stochastically selected and used to generate the fusion latent information. As illustrated in FIGS. 7, 8 and 10, the sub attribute may be also displayed on a user interface.

For example, as illustrated in FIG. 12A, it is assumed that latent information X and latent information Y of two origins are composed of code Xc, main attribute Xa_main and sub attribute Xa_sub and code Yc, main attribute Ya_main and sub attribute Ya_sub, respectively. Here, selection probability Pmain of the main attribute and selection probability Psub of the sub attribute may be preset (in the illustrated example, but not limited to, Pmain = 75% and Psub = 25%), and one of the attributes for use in generating the fusion latent information may be determined in accordance with the preset probabilities.

For example, in the example as illustrated in FIG. 12B, the main attribute Xa_main is selected for the latent information X, and the sub attribute Ya_sub is selected for the latent information Y. Here, a probability value at selection of the attributes may be

set to different values for respective elements of the attributes. Also, either the main attribute or the sub attribute may be selected for respective elements of the attributes depending on the preset probability values.

The latent information fusion unit 120 generates the fusion latent information Z as illustrated in FIG. 12C based on the selected latent information X (= $X_c + X_{a_main}$) and the selected latent information Y (= $Y_c + Y_{a_sub}$).

As illustrated, the fusion latent information Z may also include main attribute Z_{a_main} and sub attribute Z_{a_sub} . For example, in crossover of the to-be-fused attributes X_{a_main} and Y_{a_sub} , the latent information fusion unit 120 may form the sub attribute Z_{a_sub} from elements that have not been selected for the main attribute Z_{a_main} in the fusion latent information Z in the crossover.

Here, if the fusion latent information Z includes code Z_c , main attribute Z_{a_main} and sub attribute Z_{a_sub} , the generative model may be pre-trained such that the incoming main attribute Z_{a_main} can be weighted with a higher weight than the sub attribute Z_{a_sub} . Also, the main attribute and the sub attribute may be fused with weights and be supplied to the trained

generative model. Alternatively, the fusion image generation unit 130 may generate the fusion image based on the code Zc and the main attribute Za_main without supplying the sub attribute Za_sub to the generative model.

In this manner, the introduction of the sub attribute can increase variations of the generated fusion latent information and can accordingly increase variations of the fusion images. The number of the sub attributes may be arbitrarily selected.

Similarly, the latent information may further include a sub code. The above-stated code (referred to as a main code hereinafter) is a code having a higher likelihood of its occurrence in fusion images. On the other hand, the sub code is a code that does not appear in the origin image but is likely to appear in the fusion image with a lower occurrence likelihood. The sub code may be also displayed as a barcode or the like on a user interface.

For example, as illustrated in FIG. 13A, it is assumed that latent information X and latent information Y of two origins are composed of main code Xc_main, sub code Xc_sub, main attribute Xa_main and sub attribute Xa_sub and main code Yc_main, sub code Yc_sub, main

attribute Y_{a_main} and sub attribute Y_{a_sub} , respectively. Here, selection probability P_{main} of the main code and selection probability P_{sub} of the sub code may be preset (in the illustrated example, but not limited to, $P_{main} = 75\%$ and $P_{sub} = 25\%$), and one of the codes for use in generating the fusion latent information may be determined in accordance with the preset probabilities.

For example, in the example as illustrated in FIG. 13B, the sub code X_{c_sub} is selected for the latent information X, and the main code Y_{c_main} is selected for the latent information Y. Here, a probability value at selection of the codes may be set to different values for different elements of the codes. Also, either the main code or the sub code may be selected for different elements of the codes depending on the preset probability values.

The latent information fusion unit 120 generates the fusion latent information Z as illustrated in FIG. 13C based on the selected latent information X (= $X_{c_sub} + X_{a_main}$) and the selected latent information Y (= $Y_{c_main} + Y_{a_sub}$).

As illustrated, the fusion latent information Z may also include main code Z_{c_main} and sub code Z_{c_sub} . For example, at crossover of the to-be-fused codes X_{c_sub}

and Yc_main , the latent information fusion unit 120 may form the sub code Zc_sub from elements that have not been selected for the main code Zc_main in the fusion latent information Z in the crossover.

Here, similar to the main attribute Za_main and the sub attribute Za_sub , the generative model may be pre-trained such that the incoming main code Zc_main can be weighted with a higher weight than the sub code Zc_sub . Also, the main code and the sub code may be fused with weights, and the fused codes may be supplied to the trained generative model. Alternatively, the fusion image generation unit 130 may generate the fusion image based on the main code Zc_main and the main attribute Za_main without supplying the sub code Zc_sub to the generative model.

In this manner, the introduction of the sub code can increase variations of the generated fusion latent information and can accordingly increase variations of the fusion images.

Image Generation Operation and Training Operation

Next, an image generation operation and a training operation according to one embodiment of the present disclosure are described with reference to FIGS.

14 and 15. FIG. 14 is a flowchart for illustrating an image generation operation according to one embodiment of the present disclosure. The image generation operation may be executed by the image generation apparatus 100, particularly a processor in the image generation apparatus 100.

As illustrated in FIG. 14, at step S101, the latent information acquisition unit 110 acquires latent information of two to-be-fused origins. Typically, the latent information of these origins is possessed by a user on a platform, and when the user indicates images of the to-be-fused origins, the latent information acquisition unit 110 acquires the latent information of the indicated origins from the platform and supplies it to the latent information fusion unit 120.

At step S102, the latent information fusion unit 120 fuses the acquired latent information. Specifically, the latent information fusion unit 120 may perform a genetic operation such as crossover on the acquired latent information to generate fusion latent information.

At step S103, the fusion image generation unit 130 supplies the acquired fusion latent information to a trained generative model. The generative model is pre-

trained by the training apparatus 200 and may be a generator for GANs, for example. The trained generative model may be trained to generate images from the incoming latent information.

At step S104, the fusion image generation unit 130 generates a fusion image from the trained generative model. The generated fusion image is stored together with the latent information associated with the fusion image, that is, the incoming fusion latent information supplied to the trained generative model, in user's folder or the like on the platform.

FIG. 15 is a flowchart for illustrating a training operation in accordance with GANs according to one embodiment of the present disclosure. The training operation is executed by the training apparatus 200, particularly a processor in the training apparatus 200, for a generator and a discriminator for GANs. In the training operation, a dataset of training images are provided to the training apparatus 200.

As illustrated in FIG. 15, at step S201, the training apparatus 200 supplies latent information, which is composed of a code (noise) such as a vector of random numbers and a list of random attributes, to the generator and acquires an image as its output.

At step S202, the training apparatus 200 supplies the acquired image or a training image in the dataset to the discriminator and acquires as its output a discrimination result indicative of whether the incoming image is the image generated by the generator or the training image. Here, an attribute may be supplied to the discriminator and may be outputted as the discrimination result.

At step S203, the training apparatus 200 adjusts parameters for the generator and the discriminator based on the discrimination result in accordance with a parameter update procedure for GANs. Specifically, the training apparatus 200 updates the parameters for the discriminator to cause the discriminator to yield correct discrimination results and updates the parameters for the generator to cause the discriminator to yield incorrect discrimination results.

At step S204, the training apparatus 200 determines whether a termination condition is satisfied, and if the termination condition is satisfied (S204: YES), the training apparatus 200 terminates the training operation. On the other hand, if the termination condition is not satisfied (S204: NO), the training apparatus 200 returns to step S201 and repeats the above-

stated steps S201 to S204 for the next training data. For example, the termination condition may be completion of the training operation for a predetermined number of incoming data.

The training apparatus 200 provides the generator that has been acquired after completion of the training operation to the image generation apparatus 100 as the trained generative model.

Blockchain Compliant Code

Next, an image generation operation according to another embodiment of the present disclosure is described with reference to FIGS. 16 and 17. FIG. 16 is a schematic diagram for illustrating an image generation operation according to another embodiment of the present disclosure. In this embodiment, the latent information is provided in the form of a code compliant with a blockchain, and as a result, the latent information can be exchanged among users via a blockchain based distribution network.

As illustrated in FIG. 16, the latent information may be represented as a blockchain compliant code that is in compliance with any standard for blockchain based distribution networks such as Ethereum.

For example, in Ethereum, the latent information is described in 320 bits. Specifically, a code is described in 256 bits, and an attribute is described in 64 bits. In other words, the blockchain compliant code can represent the latent information with a smaller data amount, compared with a typical data amount (for example, 0.65 KB) of the latent information that is not represented in the form of the blockchain compliant code.

Correspondence between elements of the latent information that is represented in the form of a blockchain compliant code and the latent information that is not represented in the form of the blockchain compliant code may be defined, and the mutual relationship can be uniquely determined. For example, the image generation apparatus 100 may use the latent information in the form of the blockchain compliant code that a user has purchased from another user together with user's possessed another latent information to generate a fusion image from the purchased latent information and the possessed latent information.

Note that the blockchain compliant code is not limited to the above-stated Ethereum compliant code and may be codes in compliance with any other appropriate

standard based on blockchain techniques. Throughout the specification, the terminology "latent information" may include not only the latent information that is not represented in the form of a blockchain compliant code but also the latent information that is represented in the form of the blockchain compliant code as described in the present embodiment.

FIG. 17 is a flowchart for illustrating an image generation operation according to another embodiment of the present disclosure. The image generation operation may be executed by the image generation apparatus 100, particularly a processor in the image generation apparatus 100.

As illustrated in FIG. 17, at step S301, the latent information acquisition unit 110 acquires blockchain compliant codes indicative of the latent information of two to-be-fused origins. For example, the blockchain compliant codes may be acquired from other users via a blockchain based platform. Here, some conversion operations of the latent information that is not represented in the form of a blockchain compliant code into the latent information that is represented in the form of the blockchain compliant code may be performed.

At step S302, the latent information fusion unit 120 fuses the acquired blockchain compliant codes. For example, the latent information fusion unit 120 may perform a genetic operation such as crossover on the acquired blockchain compliant codes to generate a fusion blockchain compliant code. Here, the fusion operation (including main-sub selection) described with reference to FIGS. 5 to 14 for the case where the blockchain compliant code is not used may be applied to the fusion operation on the blockchain compliant codes. The latent information represented in the form of the blockchain compliant code may at least include any one of the sub attribute and the sub code.

At step S303, the latent information fusion unit 120 converts the acquired fusion blockchain compliant code into fusion latent information. The conversion may be fulfilled in accordance with correspondence information indicative of correspondence between elements of the blockchain compliant codes and elements of the latent information that is not represented in the form of the blockchain compliant code.

At step S304, the fusion image generation unit 130 supplies the acquired fusion latent information to the trained generative model. The generative model is

pre-trained by the training apparatus 200 and may be a generator for GANs. The trained generative model is trained to generate images from the incoming latent information.

At step S305, the fusion image generation unit 130 generates a fusion image from the trained generative model. The generated fusion image may be stored together with the fusion latent information in user's folder on the platform.

Note that although the fusion operation is performed on the latent information represented in the form of the blockchain compliant code in the present embodiment, the present disclosure is not limited to it, and to-be-fused blockchain compliant codes may be first converted into the latent information that is not represented in the form of the blockchain compliant code, and then the fusion operation may be performed on the converted latent information.

In this manner, users can possess the latent information in the form of the blockchain compliant code having a smaller data amount, instead of the latent information having a larger data amount that is not represented in the form of the blockchain compliant code. Also, there is an advantage that the value and ownership

of digital assets can be secured with blockchain and smart contract technology. Utilization of characteristics of the generative model allows two images in accordance with smart contract to be fused, and since results of the fusion are traceable and unpredictable, it is considered that games may be more interesting to users.

Next, an image generation system according to one embodiment of the present disclosure is described with reference to FIG. 18. In the present embodiment, all or a part of functions of the image generation apparatus 100 are provided as cloud services in the image generation system 10, and a user may generate and acquire fusion images via a user apparatus 300 such as a personal computer. For example, the image generation apparatus 100 may be located at a site different from the user apparatus 300.

As illustrated in FIG. 18, the image generation system 10 includes the image generation apparatus 100 and the user apparatus 300 communicatively coupled with the image generation apparatus 100.

When the user apparatus 300 indicates to-be-fused images to the image generation apparatus 100 via a network, the image generation apparatus 100 fuses latent information of the to-be-fused images based on the

received indication to generate fusion latent information. Then, the image generation apparatus 100 supplies the generated fusion latent information to a trained generative model to generate a fusion image and sends the generated fusion image to the user apparatus 300. The user apparatus 300 acquires the generated fusion image from the image generation apparatus 100 via the network. Note that the to-be-fused latent information indicated by the user may be the latent information that is represented in the form of the blockchain compliant code or the latent information that is not represented in the form of the blockchain compliant code.

Also, the to-be-fused latent information is sent from the user apparatus 300 to the image generation apparatus 100 in FIG. 18, but information relating to two origin images indicated by the user may be sent from the user apparatus 300 to the image generation apparatus 100. In this case, the image generation apparatus 100 may acquire the to-be-fused latent information based on the user's indication.

Although the above-stated embodiments are focused on image generation, the present disclosure is not limited to it and may be applied to generation of any type of information that can be generated from the latent

information, for example, a video, a sound or the like. In addition, the above-stated image generation technique may be used in games to generate new characters or the like.

Hardware Arrangement of Image Generation Apparatus,
Training Apparatus and User Apparatus

In the image generation apparatus 100, the training apparatus 200 and the user apparatus 300 of the embodiments, respective functions may be implemented in a circuit that is formed of an analog circuit, a digital circuit or an analog-digital mixture circuit. Also, a control circuit for controlling the respective functions may be provided. The circuits may be implemented in an ASIC (Application Specific Integrated Circuit), a FPGA (Field Programmable Gate Array) or the like.

In all the above-stated embodiments, at least a part of the image generation apparatus 100, the training apparatus 200 and the user apparatus 300 may be arranged with hardware items. Also, if they are arranged with software items, a CPU (Central Processing Unit) or the like may implement them through information processing of the software items. In the case where they are arranged with software items, programs for implementing the image generation apparatus 100, the training apparatus 200, the

user apparatus 300 and functions of at least a portion thereof are stored in a storage medium and may be loaded into a computer for execution. The storage medium is not limited to a removable storage medium such as a magnetic disk (for example, a flexible disk) or an optical disk (for example, a CD-ROM or a DVD-ROM) and may be a fixed type of storage medium such as a SSD (Solid State Drive) using a hard disk device or a memory device. In other words, the information processing with software items may be some specific implementations using hardware resources. In addition, processing with software items may be implemented in a circuit such as a FPGA and may be executed with hardware resources. Jobs may be executed by using an accelerator such as a GPU (Graphics Processing Unit), for example.

For example, by a computer reading dedicated software items stored in a computer-readable storage medium, the computer can be embodied as the above implementations. The type of storage medium is not limited to any specific one. By installing the dedicated software items downloaded via a communication network into a computer, the computer can serve as the above implementations. In this manner, information processing with the software items can be concretely implemented

with hardware resources.

FIG. 19 is a block diagram for illustrating one exemplary hardware arrangement according to one embodiment of the present disclosure. Each of the image generation apparatus 100, the training apparatus 200 and the user apparatus 300 can be implemented as a computing device including a processor 101, a main memory device 102, an auxiliary storage device 103, a network interface 104 and a device interface 105, which are coupled via a bus 106.

Note that each of the image generation apparatus 100, the training apparatus 200 and the user apparatus 300 in FIG. 14 includes respective components singly, but the same component may be plurally provided. Also, although the image generation apparatus 100, the training apparatus 200 and the user apparatus 300 are singly illustrated, software items may be installed in multiple computers, and each of the multiple image generation apparatuses 100, the multiple training apparatuses 200 and the user apparatuses 300 may perform different portions of software operations. In this case, each of the multiple image generation apparatuses 100, the multiple training apparatuses 200 and the user apparatuses 300 may communicate with each other via the

network interface 104 or the like.

The processor 101 is an electronic circuit (a processing circuit or a processing circuitry) including a controller and an arithmetic unit of the image generation apparatus 100, the training apparatus 200 and the user apparatus 300. The processor 101 performs arithmetic operations based on incoming data and programs from respective internal devices in the image generation apparatus 100, the training apparatus 200 and the user apparatus 300 and supplies operation results and control signals to the respective internal devices or the like. Specifically, the processor 101 runs operating systems (OS), applications or the like in the image generation apparatus 100, the training apparatus 200 and the user apparatus 300 to control respective components of the image generation apparatus 100, the training apparatus 200 and the user apparatus 300. The processor 101 is not particularly limited to any certain one and may be any other implementation that can perform the above operations. The image generation apparatus 100, the training apparatus 200, the user apparatus 300 and respective components thereof may be implemented with the processor 101. Here, the processing circuit may be one or more electric circuits disposed on a single chip or on

two or more chips or devices. If the multiple electronic circuits are used, the respective electronic circuits may communicate with each other in a wired or wireless manner.

The main memory device 102 is a memory device for storing various data and instructions for execution by the processor 101, and information stored in the main memory device 102 is directly read by the processor 101. The auxiliary storage device 103 includes storage devices other than the main memory device 102. Note that the memory device and the storage device mean arbitrary electronic parts capable of storing electronic information and may serve as memories or storages. Also, the memory device may be any of a volatile memory and a non-volatile memory. The memory device for storing various data in the image generation apparatus 100, the training apparatus 200 and the user apparatus 300 may be implemented with the main memory device 102 or the auxiliary storage device 103, for example. As one example, at least a portion of the memory device may be implemented in the main memory device 102 or the auxiliary storage device 103. As another example, if an accelerator is provided, at least a portion of the above-stated memory device may be implemented in memory device within the accelerator.

The network interface 104 is an interface for connecting to the communication network 108 in a wired or wireless manner. The network interface 104 may be compliant with any of existing communication standards. Information may be exchanged with the external apparatus 109A communicatively coupled via the communication network 108.

The external apparatus 109A may include a camera, a motion capture, an output device, an external sensor, an input device and so on, for example. Also, the external apparatus 109A may be an apparatus having a part of functions of components in the image generation apparatus 100, the training apparatus 200 and the user apparatus 300. Then, the image generation apparatus 100, the training apparatus 200 and the user apparatus 300 may receive a part of processing results of the image generation apparatus 100, the training apparatus 200 and the user apparatus 300 via the communication network 108 as in cloud services.

The device interface 105 is an interface such as a USB (Universal Serial Bus) directly coupled with the external apparatus 109B. The external apparatus 109B may be an external storage medium or a storage device. The memory device may be implemented with the external

apparatus 109B.

The external apparatus 109B may be an output device. The output device may be a display device for displaying images or an output device for sounds or the like, for example. For example, the output device may be, but not limited to, a LCD (Liquid Crystal Display), a CRT (Cathode Ray Tube), a PDP (Plasma Display Panel), an organic EL (ElectroLuminescence) display, a speaker or the like.

Note that the external apparatus 109B may be an input device. The input device may include a device such as a keyboard, a mouse, a touch panel, a microphone or the like, and incoming information from these devices is provided to the image generation apparatus 100, the training apparatus 200 and the user apparatus 300. Signals from the input device are supplied to the processor 101.

For example, the latent information acquisition unit 110, the latent information fusion unit 120 and the fusion image generation unit 130 or the like in the image generation apparatus 100 according to the present embodiments may be implemented with one or more processors 101. Also, memory devices in the image generation apparatus 100, the training apparatus 200 and

the user apparatus 300 may be implemented with the main memory device 102 or the auxiliary storage device 103. Also, the image generation apparatus 100, the training apparatus 200 and the user apparatus 300 may include one or more memory devices.

In the specification, the representation "at least one of a, b and c" may include not only combinations a, b, c, a-b, a-c, b-c and a-b-c but also combinations of a plurality of the same elements a-a, a-b-b, a-a-b-b-c-c or the like. Also, the representation may cover arrangements including elements other than a, b and c such as the combination a-b-c-d.

Similarly, in the specification, the representation "at least one of a, b or c" may include not only combinations a, b, c, a-b, a-c, b-c and a-b-c but also combinations of a plurality of the same elements a-a, a-b-b, a-a-b-b-c-c or the like. Also, the representation may cover arrangements including elements other than a, b and c such as the combination a-b-c-d.

Although certain embodiments of the present disclosure have been described in detail, the present disclosure is not limited to the above-stated certain embodiments, and various modifications can be made within the spirit of the present disclosure as defined by claims.

[CLAIMS]

[Claim 1]

An image generation method comprising:
acquiring first latent information of a first image and second latent information of a second image;
fusing the first latent information and the second latent information to generate fusion latent information; and

supplying the fusion latent information to a trained generative model to generate a fusion image.

[Claim 2]

The image generation method as claimed in claim 1, wherein the fusing the first latent information and the second latent information comprises performing a genetic operation on the first latent information and the second latent information to generate the fusion latent information.

[Claim 3]

The image generation method as claimed in claim 1 or 2, wherein the first latent information includes a code and an attribute for the first image, and the second latent information includes a code and an attribute for the second image.

[Claim 4]

The image generation method as claimed in any one of claims 1 to 3, wherein the first latent information and the second latent information are represented in blockchain compliant codes.

[Claim 5]

The image generation method as claimed in any one of claims 1 to 4, wherein the trained generative model is a generator trained in accordance with GANs (Generative Adversarial Networks).

[Claim 6]

An image generation apparatus comprising:
a memory device configured to store first latent information of a first image and second latent information of a second image; and

one or more processors coupled to the memory device and configured to:

fuse the first latent information and the second latent information to generate fusion latent information; and

supply the fusion latent information to a trained generative model to generate a fusion image.

[Claim 7]

The image generation apparatus as claimed in claim 6, wherein the one or more processors are further configured to perform a genetic operation on the first

latent information and the second latent information to generate the fusion latent information.

[Claim 8]

The image generation apparatus as claimed in claim 6 or 7, wherein the first latent information includes a code and an attribute for the first image, and the second latent information includes a code and an attribute for the second image.

[Claim 9]

The image generation apparatus as claimed in any one of claims 6 to 8, wherein the first latent information and the second latent information are represented in blockchain compliant codes.

[Claim 10]

The image generation apparatus as claimed in any one of claims 6 to 9, wherein the trained generative model is a generator trained in accordance with GANs (Generative Adversarial Networks).

[Claim 11]

An image generation system comprising;
a first computer configured to receive an indication of a first image and a second image; and
a second computer configured to fuse first latent information of the first image and second latent information of the second image based on the indication

received at the first computer to generate fusion latent information and supply the fusion latent information to a trained generative model to generate a fusion image,

wherein the first computer is further configured to acquire the generated fusion image from the second computer.

[Claim 12]

The image generation system as claimed in claim 11, wherein the second computer is configured to perform a genetic operation on the first latent information and the second latent information to generate the fusion latent information.

[Claim 13]

The image generation system as claimed in claim 11 or 12, wherein the first latent information includes a code and an attribute for the first image, and the second latent information includes a code and an attribute for the second image.

[Claim 14]

The image generation system as claimed in any one of claims 11 to 13, wherein the first latent information and the second latent information are represented in blockchain compliant codes.

[Claim 15]

The image generation system as claimed in any

one of claims 11 to 14, wherein the trained generative model is a generator trained in accordance with GANs (Generative Adversarial Networks).

[Claim 16]

A program for causing a computer to perform operations comprising:

acquiring first latent information of a first image and second latent information of a second image;

fusing the first latent information and the second latent information to generate fusion latent information; and

supplying the fusion latent information to a trained generative model to generate a fusion image.

[Claim 17]

The program as claimed in claim 16, wherein the fusing the first latent information and the second latent information comprises performing a genetic operation on the first latent information and the second latent information to generate the fusion latent information.

[Claim 18]

The program as claimed in claim 16 or 17, wherein the first latent information includes a code and an attribute for the first image, and the second latent information includes a code and an attribute for the second image.

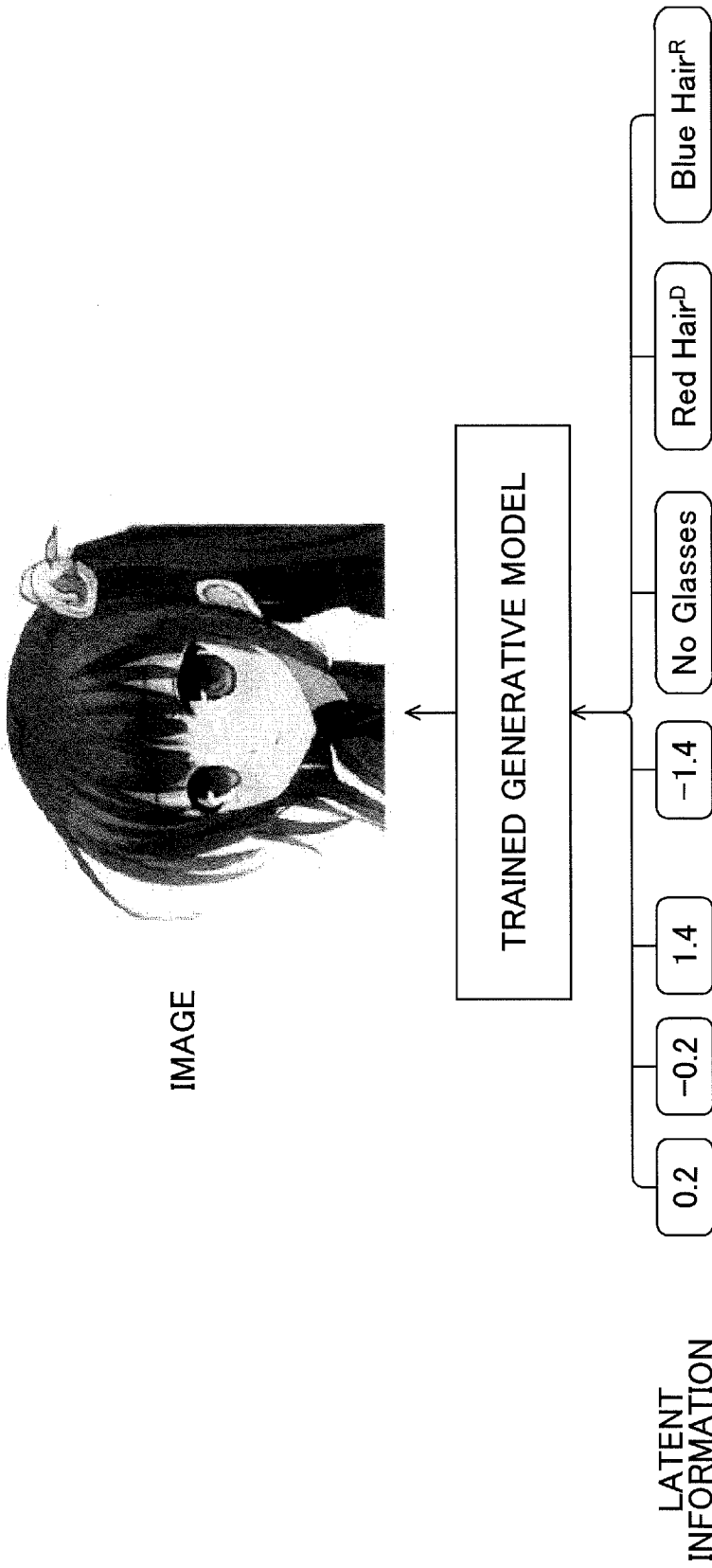
[Claim 19]

The program as claimed in any one of claims 16 to 18, wherein the first latent information and the second latent information are represented in blockchain compliant codes.

[Claim 20]

The program as claimed in any one of claims 16 to 19, wherein the trained generative model is a generator trained in accordance with GANs (Generative Adversarial Networks).

FIG.1



2/18

FIG.2



FIG.3



3/18

FIG.4



FIG.5

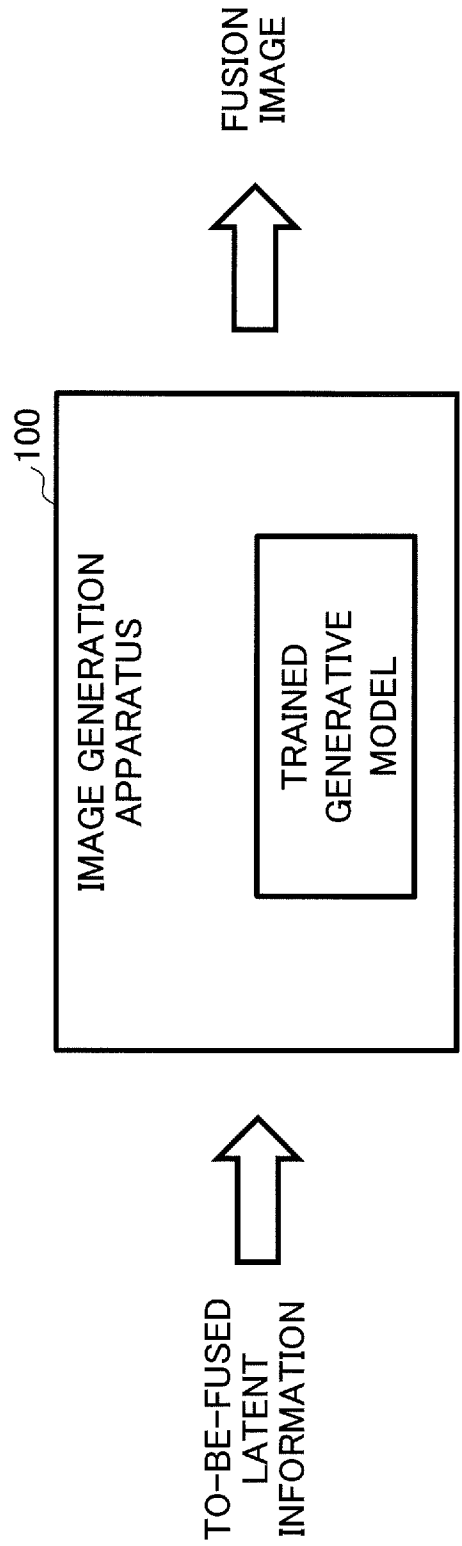
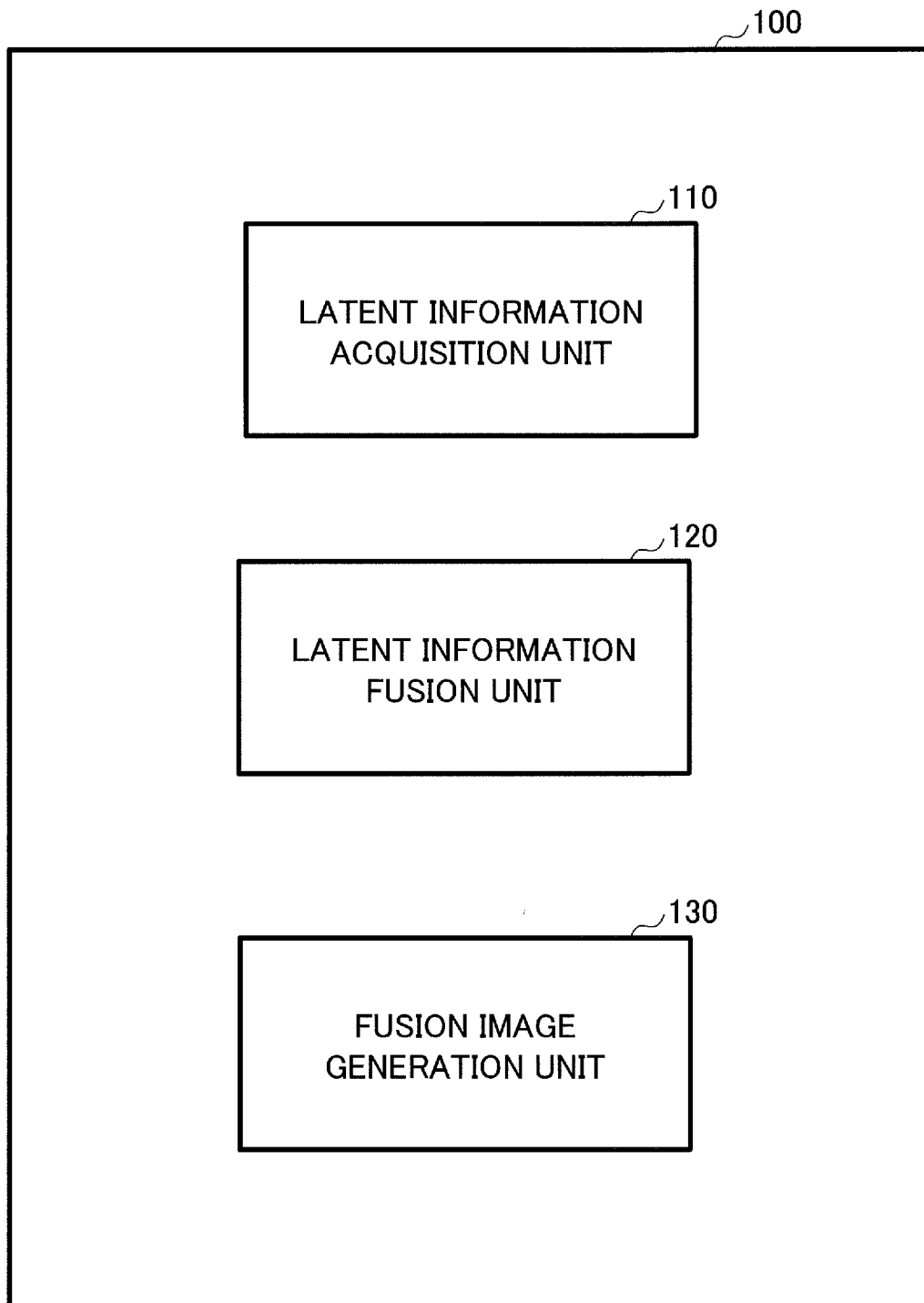


FIG.6



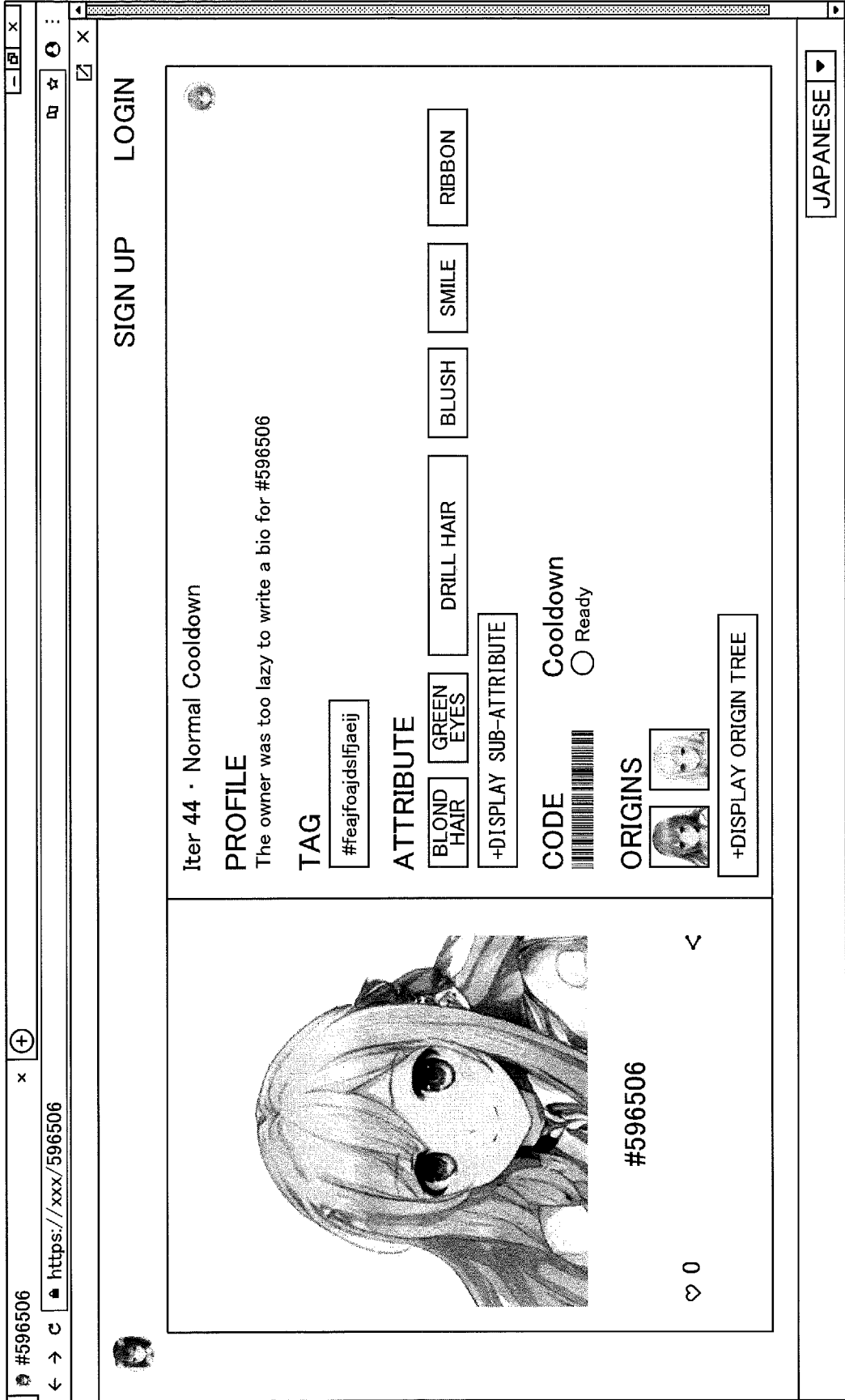
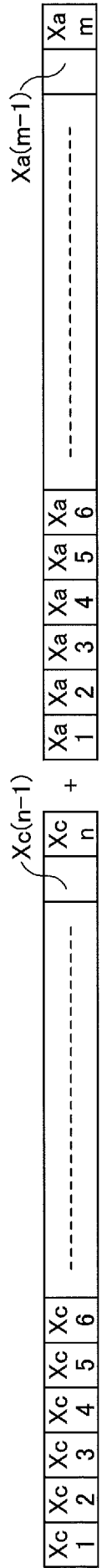


FIG.7



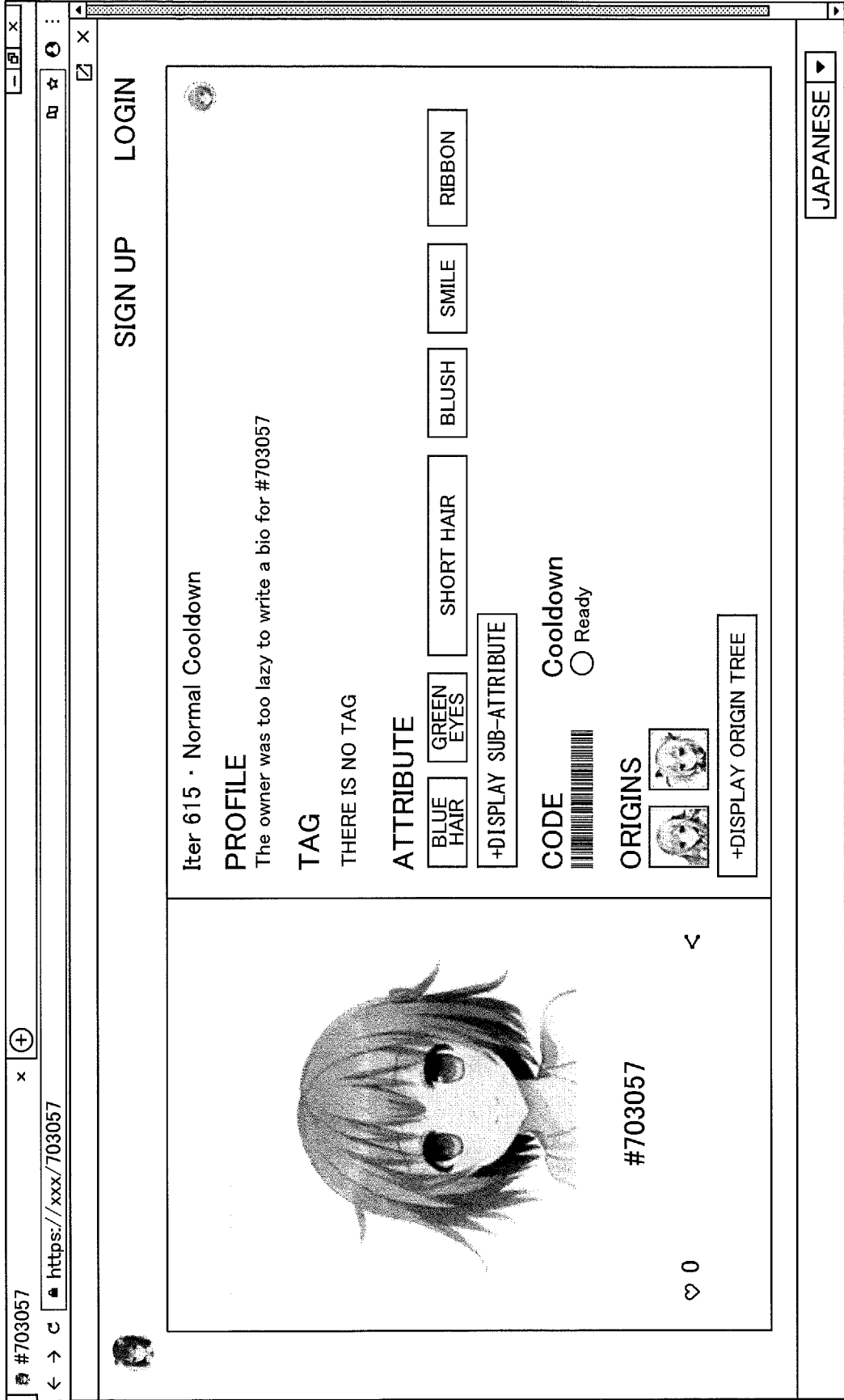


FIG.10

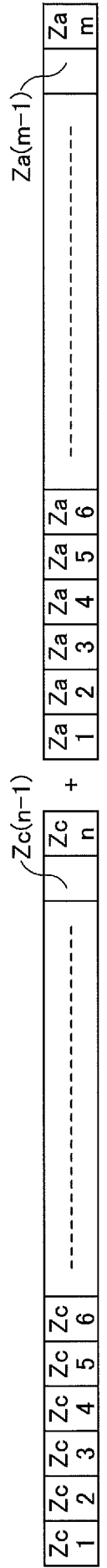


FIG.11

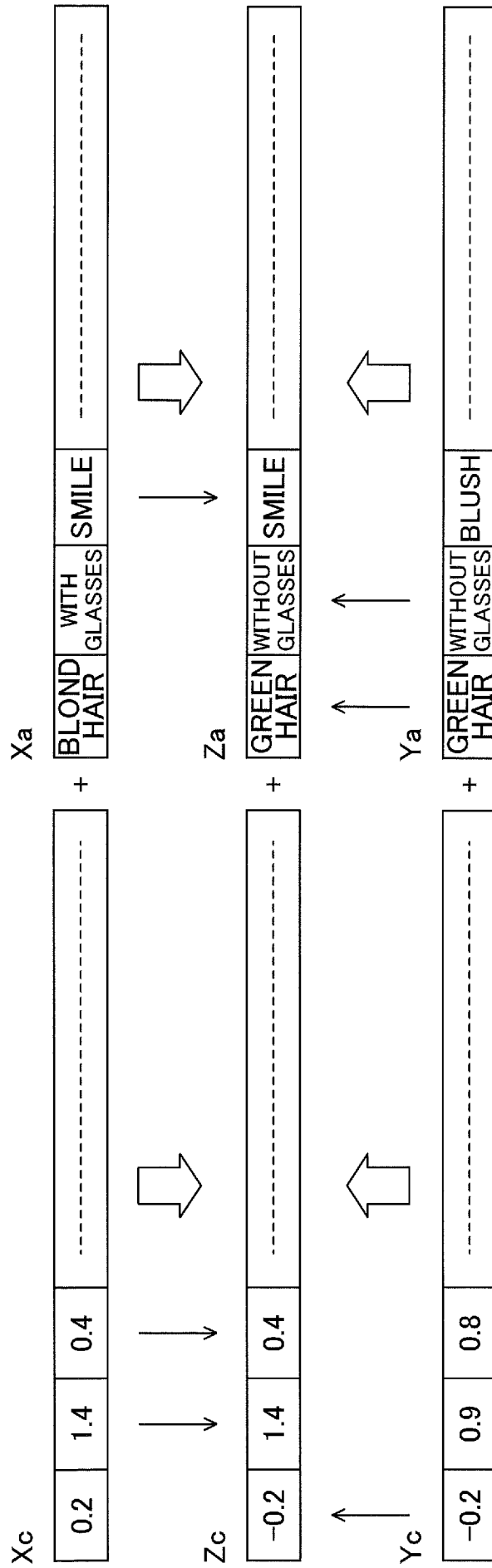


FIG.12A

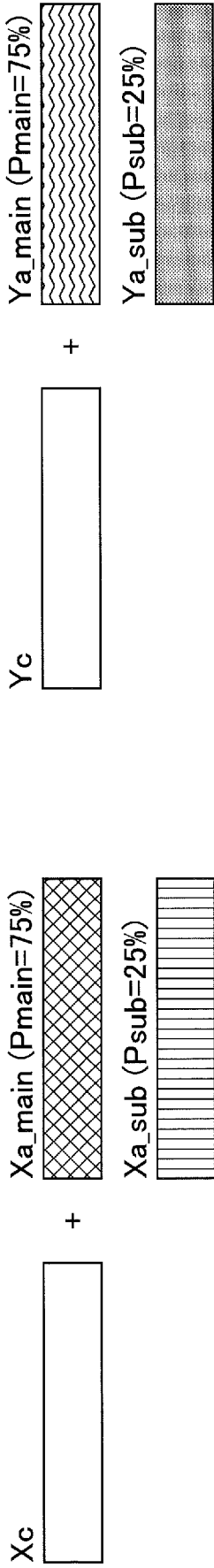


FIG.12B



FIG.12C

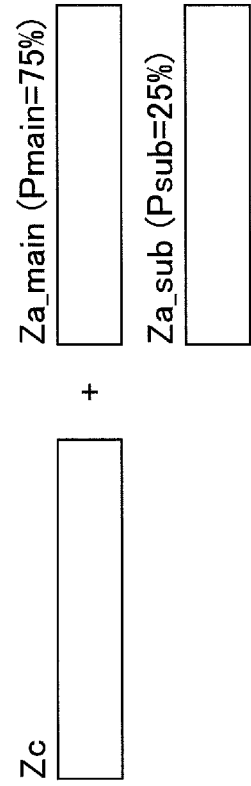


FIG.13A

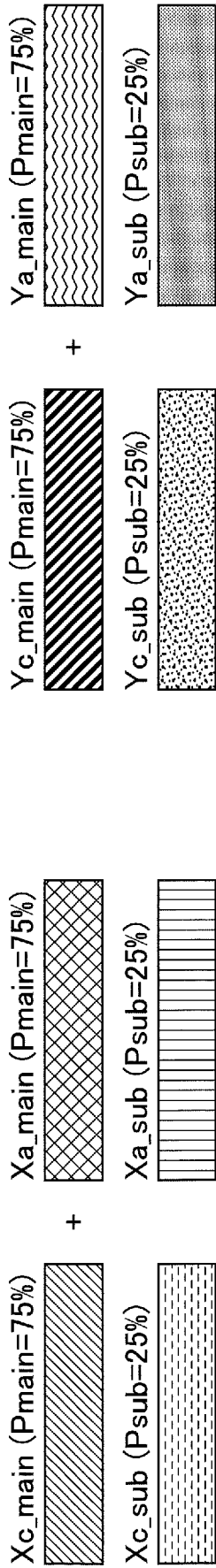


FIG.13B



FIG.13C

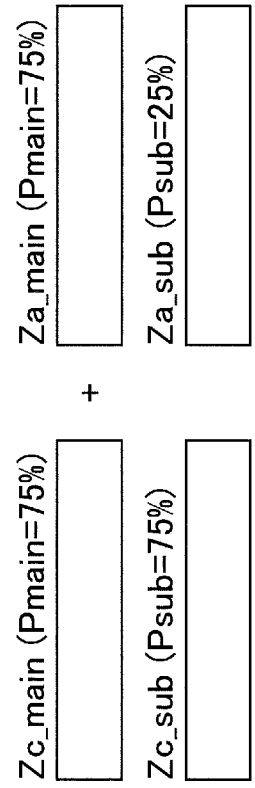


FIG.14

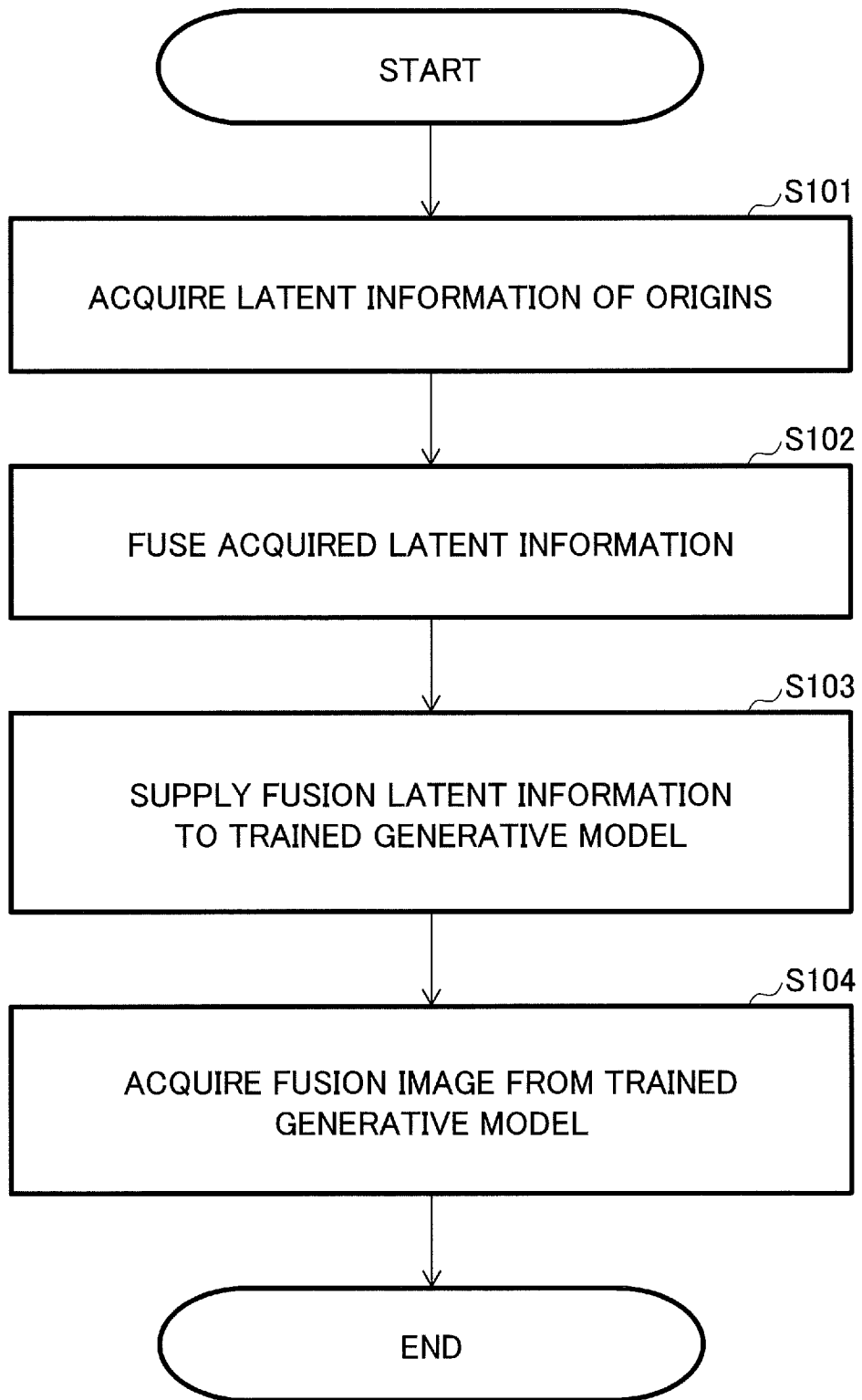


FIG.15

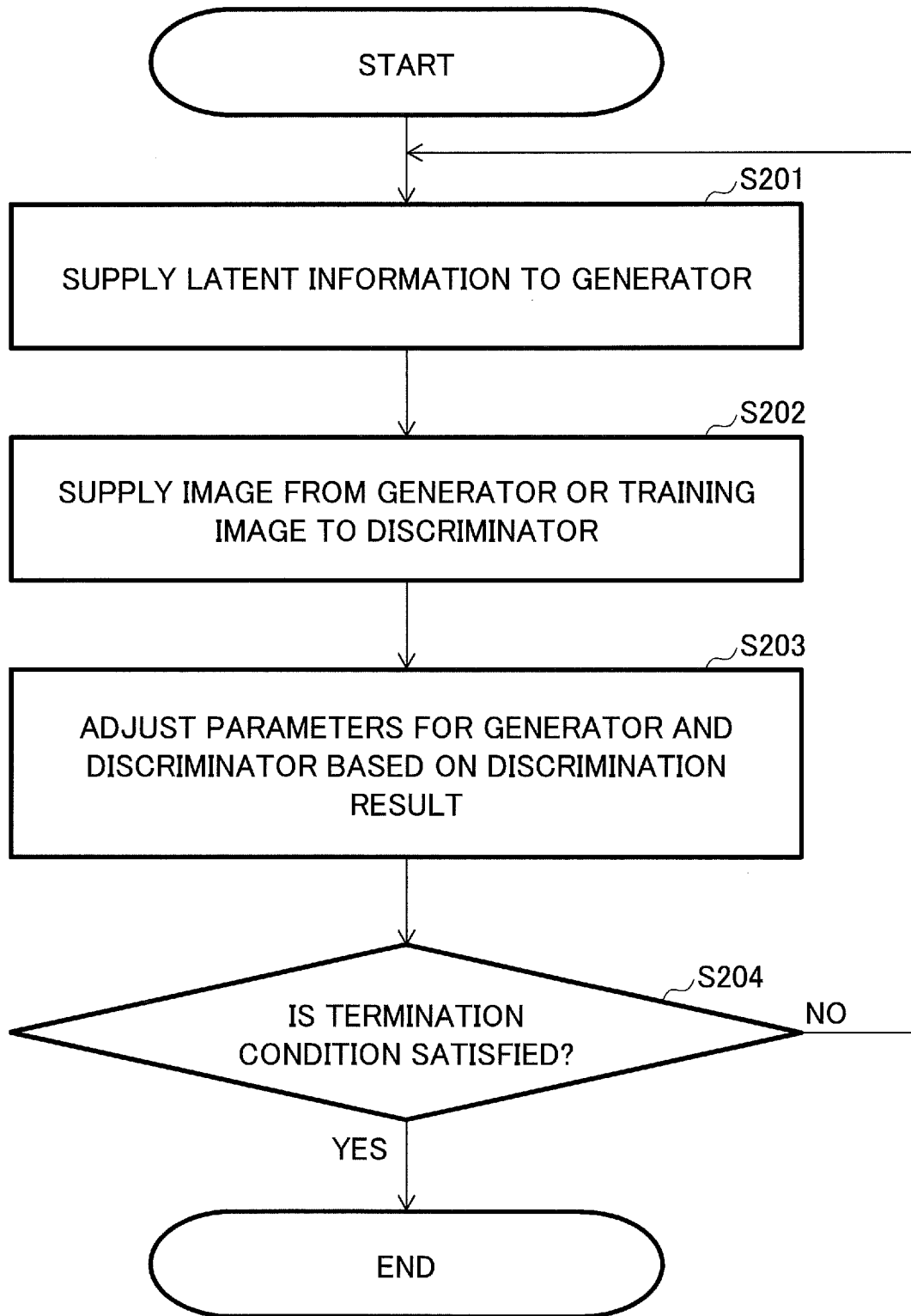


FIG.16

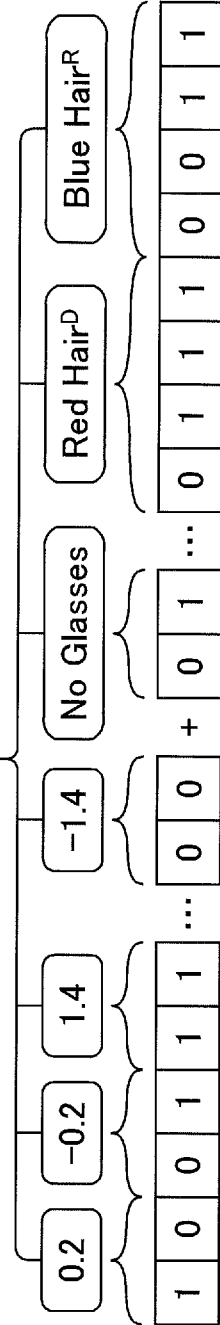


IMAGE

TRAINED GENERATIVE MODEL

LATENT INFORMATION

BLOCKCHAIN COMPLIANT CODE



16/18

FIG.17

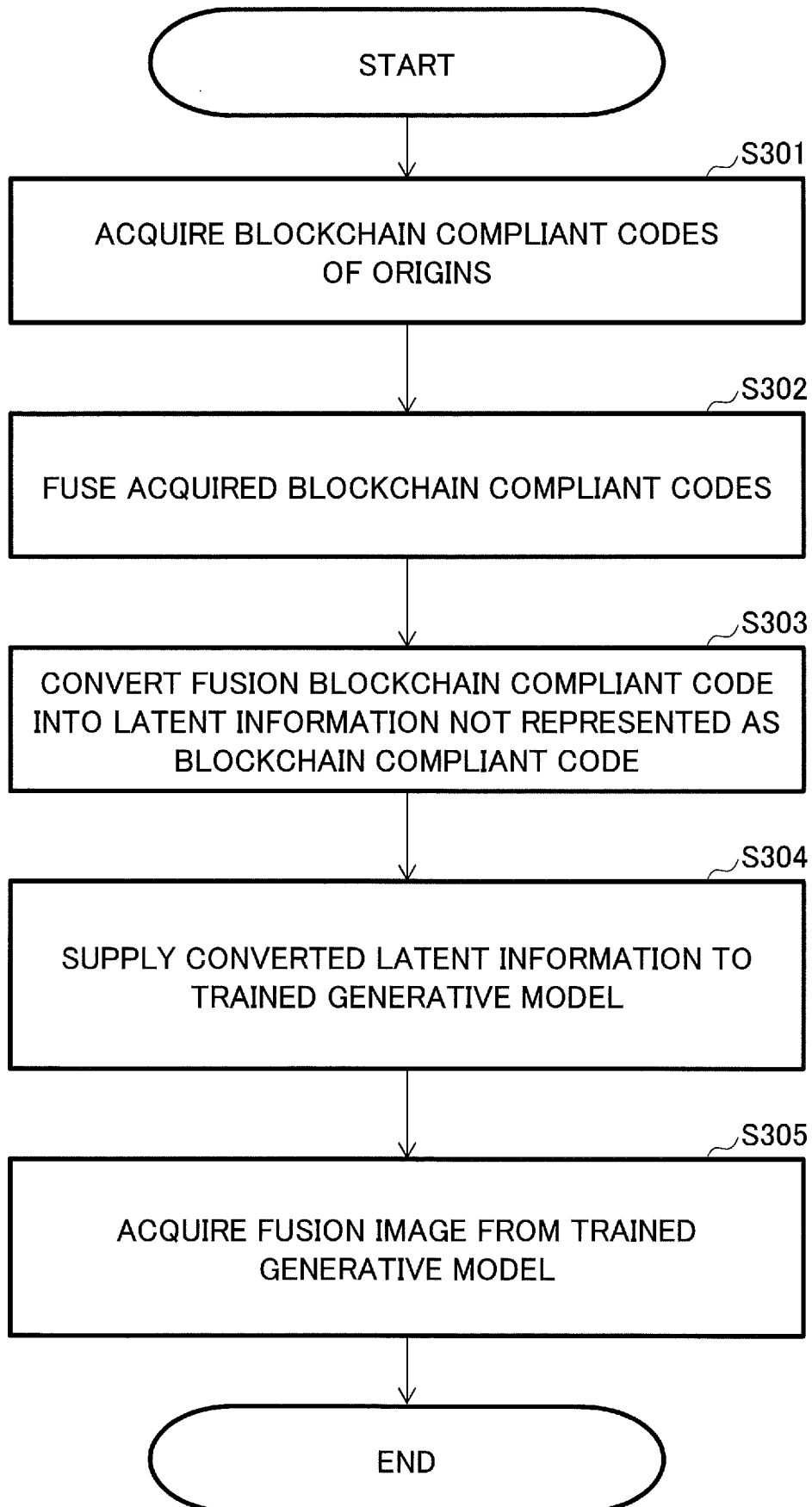


FIG.18

10

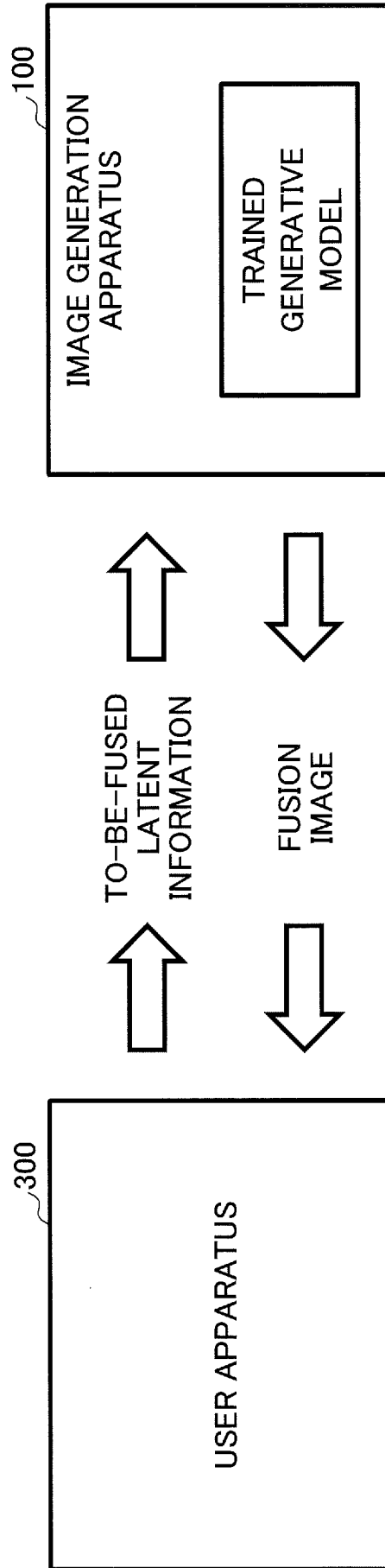
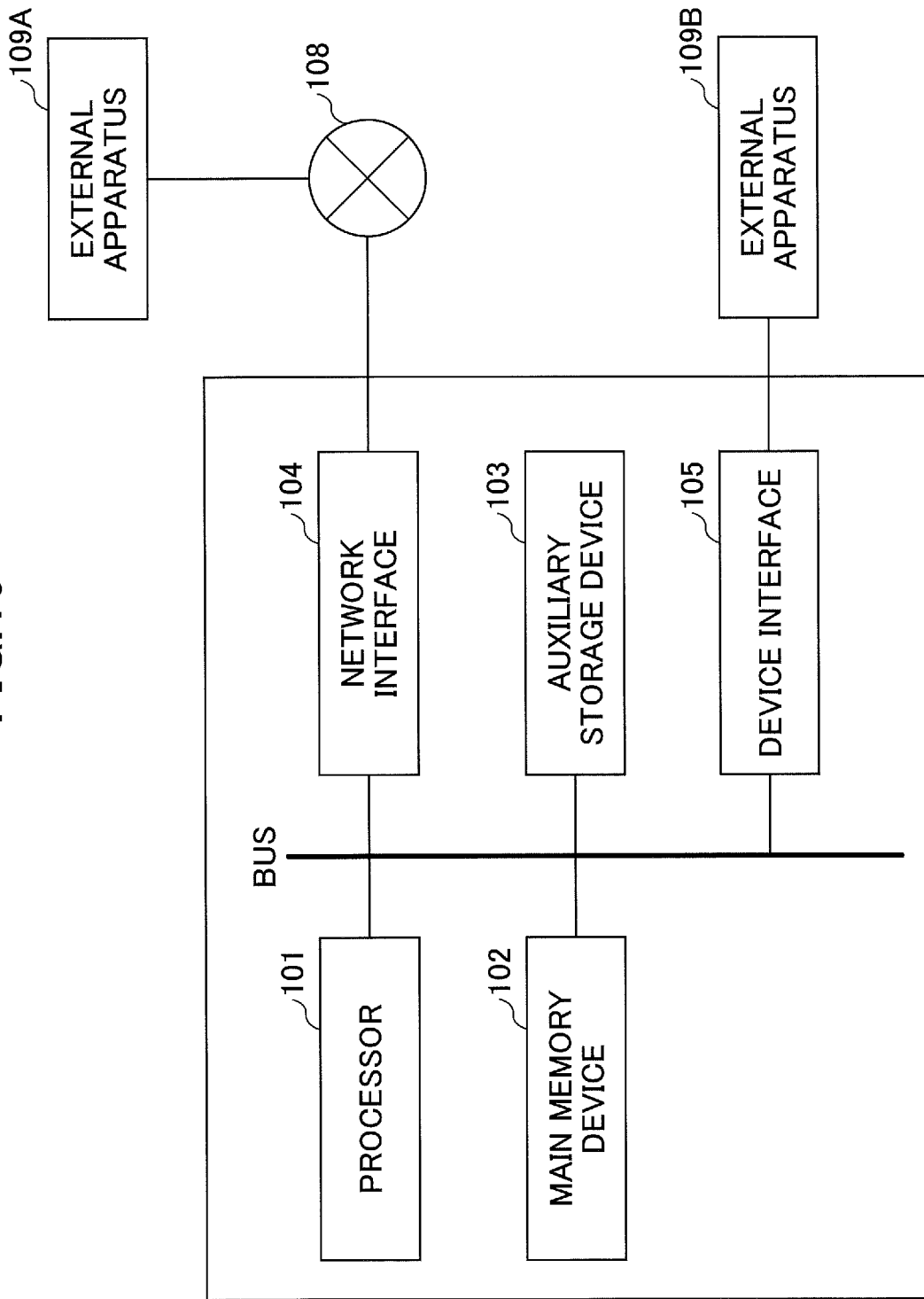


FIG.19



INTERNATIONAL SEARCH REPORT

International application No.

PCT/US19/22688

A. CLASSIFICATION OF SUBJECT MATTER
 IPC - G06N 20/00, 7/00; G06T 7/00, 7/10, 7/174 (2019.01)
 CPC - G06N 20/00, 7/00, 7/005; G06T 7/00, 7/10, 7/174

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)
 See Search History document

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
 See Search History document

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	US 2016/0026926 A1 (SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN) 28 January 2016; entire document, in particular figures 2, 21; paragraphs [0045]-[0046], [0054]-[0062], [0097]-[0098], [0102]-[0103], [0134]-[0135]; claims 19-20	1, 3/1, 11, 13/11, 16, 18/16 --- 2, 3/2, 6-7, 8/6-7, 12, 13/12, 17, 18/17
Y	US 2016/0217387 A1 (PREFERRED NETWORKS, INC.) 28 July 2016; figure 1; paragraphs [0018], [0044]-[0045]; claim 8	2, 3/2, 6-7, 8/6-7, 12, 13/12, 17, 18/17

Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents:
 "A" document defining the general state of the art which is not considered to be of particular relevance
 "E" earlier application or patent but published on or after the international filing date
 "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
 "O" document referring to an oral disclosure, use, exhibition or other means
 "P" document published prior to the international filing date but later than the priority date claimed
 "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
 "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
 "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
 "&" document member of the same patent family

Date of the actual completion of the international search
 24 April 2019 (24.04.2019)

Date of mailing of the international search report
17 MAY 2019

Name and mailing address of the ISA/
 Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
 P.O. Box 1450, Alexandria, Virginia 22313-1450
 Facsimile No. 571-273-8300

Authorized officer
 Shane Thomas
 PCT Helpdesk: 571-272-4300
 PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US19/22688

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.: 4-5, 9-10, 14-15, 19-20
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.