

DESCRIPTION

Title of Invention

CONTENT GENERATION SYSTEM, METHOD, PROGRAM, AND STORAGE
MEDIUM

Technical Field

[0001]

The present invention relates to a content generation system, a method, a program, and a storage medium.

Background Art

[0002]

Conventionally, a method for controlling a virtual character (or "avatar") using a multimodal model is known (see Patent Literature 1).

Citation List

Patent Literature

[0003]

Patent Literature 1: JP 2022-534708 A

Summary of Invention

Technical Problem

[0004]

However, conventionally, there is no proposal for generating content constituted by frames including a

character having an expression intended by a user.

[0005]

A technical problem of the present invention is to easily generate content constituted by frames including a character having an expression intended by a user.

Solution to Problem

[0006]

[1] A content generation system according to an aspect is

a content generation system for generating content constituted by one or more frames including a moving image or a still image of a character, and includes:

a display control unit configured to allow a display unit to display a user interface including a first two-dimensional operation region and a preview image of a frame to be edited by a user;

an expression determination unit configured to determine an expression of a character on the preview image depending on a position of a pointer on the first two-dimensional operation region; and

a content generation unit configured to generate content constituted by one or more frames including the frame edited by the user.

[0007]

[2] A content generation system according to an aspect is the content generation system according to [1], in which the user interface includes a second two-dimensional operation region, and the content generation system further includes: a gaze determination unit configured to determine a gaze of a character on the preview image depending on a position of a pointer on the second two-dimensional operation region.

[0008]

[3] A content generation system according to an aspect is the content generation system according to [1] or [2], in which

the content generation system further includes: a motion determination unit configured to determine a motion of the character on the preview image depending on an operation of the user interface by the user.

[0009]

[4] A content generation system according to an aspect is the content generation system according to [3], in which the user interface includes a one-dimensional operation region, and

the motion determination unit determines a playback range of the motion depending on a position of a knob on the one-dimensional operation region.

[0010]

[5] A content generation system according to an aspect is the content generation system according to any one of [1] to [4], in which

the content generation system further includes: a motion determination unit configured to determine a motion of the character on the preview image depending on a position of a pointer on the first two-dimensional operation region.

[0011]

[6] A content generation system according to an aspect is the content generation system according to any one of [1] to [5], in which

the content generation system further includes: a character determination unit configured to determine a character to be displayed on the preview image.

[0012]

[7] A content generation system according to an aspect is the content generation system according to any one of [1] to [6], in which

the content generation system further includes: a background image determination unit configured to determine a background image of the preview image.

[0013]

[8] A content generation system according to an aspect is the content generation system according to any one of

[1] to [7], in which

the content generation system further includes: a text information determination unit configured to determine at least one of a text to be displayed on the preview image, a type of balloon to display the text, a color of the text, and a size of the text depending on an operation of the user interface by the user.

[0014]

[9] A content generation system according to an aspect is the content generation system according to any one of [1] to [8], in which

the first two-dimensional operation region is a user interface to set a feeling of the character.

[0015]

[10] A content generation system according to an aspect is the content generation system according to any one of [1] to [9], in which

the content generation system further includes: a voice generation unit where image data of a character selected by the user is input to a voice model using a character image and voice data associated with the character image as learning data to generate voice data of a character.

[0016]

[11] A content generation system according to an aspect is the content generation system according to any one of

[1] to [10], in which

the content generation system further includes: a speech data generation unit where speech data of the user is input to a language model learned by using image data and/or voice data of a character as a part of learning data to generate speech data of a character in response to speech data of the user.

[0017]

[12] A content generation system according to an aspect is the content generation system according to any one of [1] to [11], in which

the content generation system further includes: a character image generation unit where image data input by the user is input to a learned image processing model to generate a character image.

[0018]

[13] A content generation system according to an aspect is the content generation system according to any one of [1] to [12], in which

the content generation system further includes: a background image generation unit where image data input by the user is input to a learned image processing model to generate a background image.

[0019]

[14] A content generation system according to an aspect

is the content generation system according to any one of [1] to [13], in which

the character image generation unit and the background image generation unit each generate three-dimensional image data, each of which is arranged on a metaverse space.

[0020]

[15] A content generation system according to an aspect is the content generation system according to any one of 1] to [14], in which

the content generation system further includes: an NFT issuance unit configured to issue an NFT related to content generated by the content generation unit and/or image data of a character in the content.

[0021]

[16] A method according to an aspect is

a method for generating content constituted by one or more frames including a moving image or a still image of a character, and includes:

a display control step of allowing a display unit to display a user interface including a first two-dimensional operation region and a preview image of a frame to be edited by a user;

an expression determination step of determining an expression of a character on the preview image depending on a position of a pointer on the first two-dimensional

operation region; and

a content generation step of generating content constituted by one or more frames including the frame edited by the user.

[0022]

[17] A program according to an aspect is

a program for making a computer execute a method for generating content constituted by one or more frames including a moving image or a still image of a character, and the method includes:

a display control step of allowing a display unit to display a user interface including a first two-dimensional operation region and a preview image of a frame to be edited by a user;

an expression determination step of determining an expression of a character on the preview image depending on a position of a pointer on the first two-dimensional operation region; and

a content generation step of generating content constituted by one or more frames including the frame edited by the user.

[0023]

[18] A storage medium according to an aspect is

a computer-readable storage medium storing the program according to [17].

[0024]

[19] A content generation system according to an aspect is

a content generation system for generating content constituted by one or more frames including a moving image or a still image of a character, and includes:

a display control unit configured to allow a display unit to display a user interface and a preview image of a frame to be edited by a user;

a character determination unit configured to determine a character to be displayed on the preview image depending on an operation of the user interface by the user;

an expression determination unit configured to determine an expression of the character depending on an operation of the user interface by the user;

a gaze determination unit configured to determine a gaze of the character depending on an operation of the user interface by the user;

a motion determination unit configured to determine a motion of the character depending on an operation of the user interface by the user;

a background determination unit configured to determine a background image of the preview image depending on an operation of the user interface by the user;

a text information determination unit configured to

determine at least one of a text to be displayed on the preview image, a type of balloon to display the text, a color of the text, and a size of the text depending on an operation of the user interface by the user; and

a content generation unit configured to generate content constituted by one or more frames including the frame edited by the user.

Brief Description of Drawings

[0025]

[Fig. 1] Fig. 1 is a diagram illustrating an example of the system configuration of a content generation system 1 according to the embodiment.

[Fig. 2] Fig. 2 is a block diagram illustrating an example of the hardware configuration of a content generation system 1 according to the embodiment.

[Fig. 3] Fig. 3 is a block diagram illustrating an example of the functional configuration of a content generation system 1 according to the embodiment.

[Fig. 4] Fig. 4 is a flowchart illustrating an example of the operation of a content generation system 1 according to the embodiment.

[Fig. 5] Fig. 5 is a diagram illustrating an example of a content creation screen according to the embodiment.

[Fig. 6] Fig. 6 is a diagram illustrating an example of

a content creation screen according to the embodiment, and illustrating a state after a character is determined.

[Fig. 7] Fig. 7 is a diagram illustrating an example of a content creation screen according to the embodiment, and illustrating a state after the expression of a character is determined.

[Fig. 8] Fig. 8 is a diagram illustrating an example of a content creation screen according to the embodiment, and illustrating a state after the gaze of a character is determined.

[Fig. 9] Fig. 9 is a diagram illustrating an example of a content creation screen according to the embodiment, and illustrating a state where the motion of a character is being selected.

[Fig. 10] Fig. 10 is a diagram illustrating an example of a content creation screen according to the embodiment, and illustrating a state after the motion of a character is determined.

[Fig. 11] Fig. 11 is a diagram illustrating an example of a content creation screen according to the embodiment, and illustrating a state after the arrangement of a character is determined.

[Fig. 12] Fig. 12 is a diagram illustrating an example of a content creation screen according to the embodiment, and illustrating a state after a background image is

determined.

[Fig. 13] Fig. 13 is a diagram illustrating an example of a content creation screen according to the embodiment, and illustrating a state after text information is determined.

[Fig. 14] Fig. 14 is a diagram illustrating an example of an expression setting unit according to the embodiment.

[Fig. 15] Fig. 15 is a diagram illustrating an example of a gaze setting unit according to the embodiment.

[Fig. 16] Fig. 16 is a diagram illustrating an example of a motion setting unit according to the embodiment.

[Fig. 17] Fig. 17 is a diagram illustrating another example of an expression setting unit according to the embodiment.

[Fig. 18] Fig. 18 is a diagram illustrating another example of a gaze setting unit according to the embodiment.

Description of Embodiments

[0026]

Hereinafter, the embodiments of the present invention will be described in detail with reference to the accompanying drawings. Note that components having the same function in each figure are denoted by the same reference signs, and the detailed descriptions of the components with the same reference signs are not repeated.

[0027]

(Summary of Content Generation System)

The content generation system according to the embodiment is used for a user to generate content constituted by one or more frames such as comics or animations. Hereinafter, the content generation system according to the embodiment will be described in detail.

[0028]

(Configuration of Content Generation System)

Fig. 1 is a diagram illustrating the outline of the configuration of a content generation system 1 according to the embodiment. As illustrated in Fig. 1, the content generation system 1 according to the embodiment includes a user terminal 2 (terminal device) to be used by a user, and a server 3.

[0029]

The user terminal 2 and the server 3 are connected to each other via a network 4 such as the internet so that they can communicate with each other. The network 4 may be a wire circuit or a wireless circuit regardless of the type and form of the circuit. At least a part of the user terminal 2 and the server 3 is realized by computer (information processing device). Examples of the user terminal 2 include terminal devices such as a personal computer, a smartphone, and a tablet terminal. A number of

user terminals 2 are present depending on the number of users using the content generation system 1.

[0030]

(Hardware Configuration)

Next, the hardware configuration of the content generation system 1 according to the embodiment will be described. Fig. 2 is a block diagram illustrating an example of the hardware configuration of the user terminal 2 and the server 3 included in the content generation system 1 according to the embodiment.

[0031]

In the user terminal 2, a CPU 201 is a processing device to control the overall operation of the user terminal 2. A ROM 202 is a nonvolatile memory to store a control program to be executed by the CPU 201 and various data. A RAM 203 is a volatile memory to be used as the load region and work region for the program executed by the CPU 201. A storage device 204 is a storage means to store various information, and may be incorporated into the main body of the user terminal 2, or may have a detachable storage medium. An input device 205 is a device for the user of the user terminal 2 to input information, and examples thereof include a keyboard, a mouse, a touch panel, and a microphone. A display 206 is a display device to display various information (user interface or the

like). An imaging element 207 is a photoelectric transducer to image a subject. A communication I/F (interface) 208 is an interface for connection with the network 4. A bus 209 is a bus line to connect the above-described components with each other.

[0032]

In the server 3, a CPU 301 is a processing device to control the overall operation of the server 3. A ROM 302 is a nonvolatile memory to store a control program to be executed by the CPU 301 and various data. A RAM 303 is a volatile memory to be used as the load region and work region for the program executed by the CPU 301. A storage device 304 is a storage means to store various information, and may be incorporated into the main body of the server 3, or may have a detachable storage medium. A communication I/F (interface) 305 is an interface for connection with the network 4. A bus 306 is a bus line to connect the above-described components with each other.

[0033]

(Functional Configuration)

Next, the functional configuration of the content generation system 1 according to the embodiment will be described. Fig. 3 is a diagram illustrating an example of the functional configuration of the content generation system 1 according to the embodiment.

[0034]

First, the functional configuration of the user terminal 2 will be described. As illustrated in Fig. 3, the user terminal 2 has a communication unit 21; a control unit 22 that controls the overall operation of the user terminal 2; an input unit 23 with which the user inputs various information; an output unit 24 that outputs various information; an imaging unit 25 that images a subject; and a storage unit 26 that stores various information.

[0035]

The communication unit 21 is a communication interface between the user terminal 2 and the network 4. The communication unit 21 sends and receives information between the user terminal 2 and the server 3 via the network 4.

[0036]

The control unit 22 has a display control unit 22a; a frame determination unit 22b; a character determination unit 22c; an expression determination unit 22d; a gaze determination unit 22e; a motion determination unit 22f; a character arrangement determination unit 22g; a background image determination unit 22h; and a text information determination unit 22i.

[0037]

The display control unit 22a allows the output unit 24

(display unit) described below to display a content creation screen (operation screen) of a content generation application (program) according to the content generation system 1.

[0038]

Fig. 5 is a diagram illustrating an example of a content creation screen 50 (operation screen) that the display control unit 22a allows the output unit 24 (display unit) to display. As illustrated in Fig. 5, the content creation screen 50 displays a plurality of GUIs (Graphical User Interface) to edit a frame constituting the content. Specifically, the content creation screen 50 includes a tab 51 where items to be edited are switched; a frame list area 52 where thumbnail images of frames constituting the content are listed, and selection of the frame to be edited or the like is performed; a preview display area 53 where the preview image of the frame that is being edited is displayed; a background image area 53a occupying at least some region in the preview display area, where a background image is displayed; a character selection area 54 where a character is selected; an expression setting unit 55 of a character; a gaze setting unit 56 of a character; a motion setting unit 57 of a character; and a cursor C with which the user selects each GUI.

[0039]

Going back to Fig. 3, the frame determination unit 22b determines the frame to be edited by the user among one or more frames constituting the content. Specifically, the frame determination unit 22b determines the frame to be edited by the user, depending on the operation through which the user uses the cursor C to select the thumbnail image of a frame shown in the frame list area 52 (see Fig. 5). Fig. 5 illustrates a state where among the thumbnail images of the frames 1 to 4 shown in the frame list area, the frame 1 is selected by the user, and the frame determination unit 22b determines the frame 1 as the frame to be edited.

[0040]

When a "+" button 52a shown inside the frame list area 52 is selected, a frame can be added in the frame list area 52. When a "-" button 52b is selected, the selected frame (in the example illustrated in Fig. 5, the frame 1) can be deleted from the frame list area 52. When a play back button 52c is selected, content constituted by one or more frames shown in the frame list area 52 (in the example illustrated in Fig. 5, the frames 1 to 4) is generated by the content generation unit 32c described below, and the generated content is displayed in the preview display area. For example, in the example illustrated in Fig. 5, the frames 1 to 4 shown in thumbnail in the frame list area 52

are displayed in the preview display area 53 while the frames are continuously switched with a predetermined time intervals as a slide show.

[0041]

Going back to Fig. 3, the character determination unit 22c determines a character to be displayed in the frame that is being edited by the user. Specifically, the character determination unit 22c determines the character that the user selects with the cursor C as a frame displayed in the frame that is being edited by the user, among the characters shown in the character selection area 54 (see Fig. 5). The image of the character selected by the character determination unit 22c is superimposed and displayed on the preview image of the frame displayed in the preview display area 53. As in the example illustrated in Fig. 5, in a state where a "non-display" button is selected from the character selection area 54, no character is displayed in the preview display area 53. However, when any character is selected from the character selection area 54, the character determination unit 22c determines the selection of the character, and as in the example illustrated in Fig. 6, the image of a character A selected by the user is superimposed and displayed on the preview image of the frame displayed in the preview display area 53. The character A displayed in the preview display area

53 can be moved, or zoomed in or out on the preview display area 53, using the cursor C. For example, when a mouse is used as the input device 205 of the user terminal 2, the character A may be moved to any position by operating the cursor C with the mouse to drag the character A on the preview display area 53. In addition, the character A may be zoomed in or out by operating the mouse wheel during the dragging operation.

[0042]

The image of the character A displayed in the preview display area 53 may be a moving image or a still image. When the motion determination unit described below determines "do nothing" as the motion (move) of the character, the image of the character displayed in the preview display area 53 is a still image.

[0043]

The expression determination unit 22d determines the "expression" of the character A displayed in the preview display area 53, depending on the user operating the expression setting unit 55.

[0044]

Fig. 14 is a diagram illustrating an example of the expression setting unit 55. As illustrated in Fig. 14, the expression setting unit 55 includes a two-dimensional operation region 55a (first two-dimensional operation

region) and a pointer 55b that can be moved on the two-dimensional operation region 55a. The expression setting unit 55 can set the expression (happy, angry, sad, and enjoy) of the character A displayed in the preview display area 53, depending on the position of the pointer 55b on the two-dimensional operation region 55a. Specifically, as the pointer 55b is moved to the "enjoy" side (in the +X direction in Fig. 14) on the two-dimensional operation region 55a of the expression setting unit 55, the expression of the character A displayed in the preview display area 53 is changed to an expression that expresses "enjoy" more. To the contrary, as the pointer 55b is moved to the "sad" side (in the -X direction in Fig. 14) on the two-dimensional operation region 55a of the expression setting unit 55, the expression of the character A displayed in the preview display area 53 is changed to an expression that expresses "sad" more. As the pointer 55b is moved to the "happy" side (in the +Y direction in Fig. 14) on the two-dimensional operation region 55a of the expression setting unit 55, the expression of the character A displayed in the preview display area 53 is changed to an expression that expresses "happy" more. To the contrary, as the pointer 55b is moved to the "angry" side (in the -Y direction in Fig. 14) on the two-dimensional operation region 55a of the expression setting unit 55, the

expression of the character A displayed in the preview display area 53 is changed to an expression that expresses "angry" more. Thus, the expression determination unit 22d determines the "expression" of the character A displayed in the preview display area 53, depending on the position of the pointer 55b on the two-dimensional operation region 55a of the expression setting unit 55. Accordingly, the user of the content generation system 1 can set the "expression" of the character on the content with an intuitive operation.

[0045]

Fig. 7 illustrates a state after the user operates the expression setting unit 55 with the cursor C to determine the "expression" of the character A. As illustrated in Fig. 7, since the pointer 55b is moved to the "happy" side on the two-dimensional operation region 55a of the expression setting unit 55, the expression of the character A displayed in the preview display area 53 is changed to an expression that expresses "happy".

[0046]

Going back to Fig. 3, the gaze determination unit 22e determines the "gaze" of the character A displayed in the preview display area 53, depending on the user operating the gaze setting unit 56.

[0047]

Fig. 15 is a diagram illustrating an example of the gaze setting unit 56. As illustrated in Fig. 15, the gaze setting unit 56 includes a two-dimensional operation region 56a (second two-dimensional operation region) and a pointer 56b that can be moved on the two-dimensional operation region 56a. The gaze setting unit 56 can set the gaze (gaze direction) of the character A displayed in the preview display area 53, depending on the position of the pointer 56b on the two-dimensional operation region 56a. Specifically, as the pointer 56b is moved to the "right" side (in the +X direction in Fig. 15) on the two-dimensional operation region 56a of the gaze setting unit 56, the gaze of the character A displayed in the preview display area 53 turns to the right direction more (the character's eye moves to the right side on the screen). To the contrary, as the pointer 56b is moved to the "left" side (in the -X direction in Fig. 15) on the two-dimensional operation region 56a of the gaze setting unit 56, the gaze of the character A displayed in the preview display area 53 turns to the left direction more (the character's eye moves to the left side on the screen). As the pointer 56b is moved to the "up" side (in the +Y direction in Fig. 15) on the two-dimensional operation region 56a of the gaze setting unit 56, the gaze of the character A displayed in the preview display area 53 turns

to the up direction more. To the contrary, as the pointer 56b is moved to the "down" side (in the -Y direction in Fig. 14) on the two-dimensional operation region 56a of the gaze setting unit 56, the gaze of the character A displayed in the preview display area 53 turns to the down direction more. Thus, the gaze determination unit 22e determines the "gaze" of the character A displayed in the preview display area 53, depending on the position of the pointer 56b on the two-dimensional operation region 56a of the gaze setting unit 56. Accordingly, the user of the content generation system 1 can set the "gaze" of the character on the content with an intuitive operation.

[0048]

Fig. 8 illustrates a state after the user operates the gaze setting unit 56 with the cursor C to determine the "gaze" of the character A. As illustrated in Fig. 8, since the pointer 56b is moved to the "right" side on the two-dimensional operation region 56a of the gaze setting unit 56, the gaze of the character A displayed in the preview display area 53 is changed to the right direction (right side on the screen).

[0049]

Going back to Fig. 3, the motion determination unit 22f determines the motion (move) of the character A displayed in the preview display area 53, depending on the user

operating the motion setting unit 57.

[0050]

Fig. 16 is a diagram illustrating an example of the motion setting unit 57. As illustrated in Fig. 16, the motion setting unit 57 includes a pull-down menu 571 to select the motion type of the character; and a playback range setting unit 572 to set the playback range of the selected motion type. The playback range setting unit 572 includes a first shaft 572a; a first knob 572b that can be moved on the first shaft 572a; a second shaft 572c; and a second knob 572d that can be moved on the second shaft 572c.

[0051]

When the pull-down menu 571 is selected with the cursor C (not illustrated in Fig. 16) in the motion setting unit 57, the motion types that can be set are listed. Then the intended motion type is selected from the listed motion types, and thereby the motion determination unit 22f determines the motion of the character.

[0052]

The shafts (the first shaft 572a and the second shaft 572c) of the playback range setting unit 572 represent the time base of the motion set in advance, and the knobs (the first knob 572b and the second knob 572d) are moved on the shafts so that the motion can be set in the intended

playback range.

[0053]

Fig. 9 is a diagram illustrating a state where the pull-down menu 571 of the motion setting unit 57 is selected with the cursor C, and then the motion types are listed. For example, when "wave hand" is selected from the motion type list illustrated in Fig. 9, the motion determination unit 22f determines the motion of the character as "wave hand", and the character A displayed in the preview display area 53 dynamically changes so as to show a motion like "wave hand" (animation) as illustrated in Fig. 10.

[0054]

Going back to Fig. 3, the character arrangement determination unit 22g determines the arrangement of the character A displayed in the preview display area 53, depending on the user operating a character arrangement setting unit 58 (see Fig. 11). For example, as illustrated in Fig. 11, the character arrangement setting unit 58 includes setting items of "camera", "character position", "rotation", and "zoom out/in". In the setting item of "camera", the contents corresponding to the angle of view of camera, such as the "whole body", "bust", and "head" of the character, can be set. In the setting item of "character position", the display position of the character

on the preview display area 53 can be set. In the setting item of "rotation", the character A on the preview display area 53 can be rotated with a predetermined angle interval (for example, 90 degree interval) in the yaw direction or in the pitch direction. In the setting item of "zoom out/in", a knob is moved on a shaft so that the character A displayed in the preview display area 53 can be zoomed in or zoomed out.

[0055]

Going back to Fig. 3, the background image determination unit 22h determines a background image to be displayed in the preview display area 53, depending on the user operating a background image setting unit 59 (see Fig. 12).

[0056]

Fig. 12 is a diagram illustrating a state where the background image is set by the user operating the background image setting unit 59, and the background image is determined by the background image determination unit 22h. As illustrated in Fig. 12, the background image setting unit 59 includes, for example, setting items of "background", "background position", "background size", and "lighting".

[0057]

In the setting item of "background", a plurality of

background images (two-dimensional image or three-dimensional image, moving image or still image) prepared in advance can be selected, and "single color" (no background image) can be selected. Besides, when an "addition" button is selected, any background image the user has can be added. Further, when an "automatic generation" button is selected, a background image can be automatically generated by using a learned image processing model. Specifically, in response to the selection of the "automatic generation" button, a background image generation unit (not illustrated) of the control unit 32 inputs a character selected by the user, the expression, gaze, and motion of the character, the text information described below, and the like as input data into the learned image processing model, so that a background image is generated. For example, the image processing model is learned by using generative adversarial networks (GAN). A still image such as a scenery photograph or the like the user has prepared in advance may be used as input data, which is input into the image processing model to generate a background image. Thus, the user's favorite background can be automatically generated. Here, when the background image is generated as three-dimensional data (three-dimensional image), the three-dimensional data of the background image may be arranged on a metaverse space together with the three-

dimensional data of the character generated by a character image generation unit 32a.

[0058]

In the setting item of "background position", the position of the background image area 53a in the preview display area 53 can be set. Specifically, when a pointer is moved on the two-dimensional operation region, the position of the background image area 53a on the preview display area 53 can be set. In the setting item of "background size", the size of the background image area 53a on the preview display area 53 can be set. Specifically, a knob is moved on a shaft so that the size of the background image area 53a on the preview display area 53 can be zoomed in or zoomed out. In the setting item of "lighting", "daytime", "evening", "night", and the like can be selected from a pull-down menu, and the description of the background image can be set depending on time zone.

[0059]

Going back to Fig. 3, the text information determination unit 22i determines text information to be displayed in the preview display area 53, depending on the user operating a text information setting unit 60 (see Fig. 13).

[0060]

Fig. 13 is a diagram illustrating a state where text information is set by the user operating the text information setting unit 60, and the text information is determined by the text information determination unit 22i. As illustrated in Fig. 13, the text information setting unit 60 includes, for example, setting items of "balloon", "font", "text", "text color", and "letter size". In the setting item of "balloon", a plurality of balloon images prepared in advance can be selected, and "non-display" (no text information) or "only text" (no balloon image) can be selected. In the setting item of "font", the font type of the text to be displayed in the preview display area 53 can be selected from a pull-down menu. In the setting item of "text", the text to be displayed in the preview display area 53 can be input by the user's input operation via the input unit 23 described below. In the setting item of "text color", the color of the text to be displayed in the preview display area 53 can be set. In the setting item of "letter size", the size of the text to be displayed in the preview display area 53 can be set. When the text information setting unit 60 sets text information, the text information determination unit 22i determines the text information, and the text information is displayed in the preview display area 53. In the example illustrated in Fig. 13, a balloon image 61 is displayed in the preview

display area 53, the balloon image 61 being obtained by synthesizing the text "hello" that the user inputs in the setting item of "text" of the text information setting unit 60 to the balloon image that the user selects in the setting item of "balloon" of the text information setting unit 60. The balloon image 61 can be moved on the preview display area 53 by the user's operation with the cursor C (for example, dragging operation). In addition, when an icon D for zooming in and zooming out displayed in the vicinity of the balloon image 61 is operated (for example dragging operation), the size of the balloon image 61 can be changed (zoomed in or zoomed out). When the region outside the preview display area 53 is selected with the cursor C, the position and size of the balloon image 61 on the preview display area 53 are decided, and the icon D for zooming in and zooming out is eliminated.

[0061]

Going back to Fig. 3, the input unit 23 is a component for the user of the user terminal 2 to input information. Examples thereof include a keyboard, a mouse, a touch panel, a microphone, and a gesture input device.

[0062]

The output unit 24 is an interface that outputs various information (image or voice) from the user terminal 2 to the user. Examples thereof include a video display device

(display unit) such as a liquid crystal display; and a speaker. When the output unit 24 is constituted as a display unit, the display unit is allowed by the display control unit 22a to display a GUI to receive operation from the user.

[0063]

The imaging unit 25 images a subject with a camera module including an imaging element on the user terminal 2. An image imaged by the imaging unit 25 is output to the control unit 22, image-processed in various manners, and displayed in the display unit (output unit 24) by the display control unit 22a or sent by the communication unit 21 to the server 3 via the network 4.

[0064]

The storage unit 26 is, for example, a data storage such as an internal memory and an external memory (such as SD memory card). The storage unit 26 stores various data handled by the control unit 22; various information downloaded by the communication unit 21 from the server 3 via the network 4; an image imaged by the imaging unit 25; and the like. Note that the storage unit 26 is not necessarily provided inside the user terminal 2, and a part or all of the storage unit 26 may be provided inside a separate device connected to the user terminal 2 via the network 4 so that they can communicate with each other.

[0065]

Next, the functional configuration of the server 3 will be described. As illustrated in Fig. 3, the server 3 has a communication unit 31, a control unit 32, and a storage unit 33.

[0066]

The communication unit 31 is a communication interface between the server 3 and the network 4. The communication unit 31 sends and receives information between the server 3 and the user terminal 2 via the network 4.

[0067]

The control unit 32 controls the overall operation of the server 3. The control unit 32 has a character image generation unit 32a, a voice generation unit 32b, a content generation unit 32c, a speech data generation unit 32d, and an NFT issuance unit 32e.

[0068]

The character image generation unit 32a generates image data of a character (two-dimensional or three-dimensional image data) based on an image received from the user terminal 2 (for example, a face image or a whole body image of the user imaged using the imaging unit 25, and a character image prepared in advance). For example, a character image may be generated by using an image generation AI as the character image generation unit 32a.

Specifically, for example, the character image generation unit 32a may generate a character image such that image data input by the user via the user terminal 2 is input into an image processing model learned by using generative adversarial networks (GAN). The generated character image may be stored in a character image DB 33a of the storage unit 33 described below, or may be sent by the communication unit 31 to the user terminal 2 via the network 4.

[0069]

The voice generation unit 32b generates voice data corresponding to the character image (two-dimensional image or three-dimensional image) generated by the character image generation unit 32a. Specifically, the voice generation unit 32b inputs the character image generated by the character image generation unit 32a into a learned voice synthesis model to output voice data. Thereby, it is possible from a character image to automatically generate voice data with a tone matching the impression of the appearance, atmosphere, and personality of the character. Note that the voice model is learned, for example, by using a number of character images and voice data associated with the character images (words of animation characters and the like) as learning data.

[0070]

Based on the voice data generated by the voice generation unit 32b, the voice generation unit 32b performs text voice synthesis, through which voice is generated based on the generated voice data from the text determined by the text information determination unit 22i and the text generated by the speech data generation unit 32d described below. The voice generated when the voice generation unit 32b performs text voice synthesis is output as a voice file.

[0071]

The content generation unit 32c generates content (comics, animations, and the like) constituted by one or more frames edited by the user's operation using various interfaces displayed in the display unit (output unit 24) of the user terminal 2. For example, the content generation unit 32c generates content data in response to the selection of the play back button 52c in the frame list area 52 of the content creation screen 50 illustrated in Fig. 5. Here, the content data means data obtained by capsulizing one or more frames edited by the user into one file, and examples thereof include a moving image file such as MP4 and AVI and a still image file such as HEIF (High Efficiency Image File Format) and AVIF (AV1 Image File Format), which can store a plurality of images. When a voice file is generated by the voice generation unit 32b,

the voice file is included therein. The generated content data is stored in a content DB 33e of the storage unit 33 described below, and sent by the communication unit 31 to the user terminal 2 via the network 4.

[0072]

The speech data generation unit 32d generates speech data of the character in response to speech data of the user input via the input unit 23 of the user terminal 2 (voice data or text data). Thus, the character generated by the character image generation unit 32a can function as an interactive robot. Specifically, the speech data generation unit 32d converts the speech data of the user input via the input unit 23 of the user terminal 2 into text data, inputs the text data into a language model (large language model), and outputs speech data of the character in response to the speech data of the user, thereby generating speech data. Here, the language model may be learned, using image data and/or voice data of the character as a part of learning data. Thus, since image and voice data of the character is input into a language model together with the speech data of the user, speech data matching the impression of the appearance, atmosphere, and personality of the character can be generated. The speech data of the character generated by the speech data generation unit 32d is sent by the communication unit 31 to

the user terminal 2 via the network 4, and output as voice from the output unit 24 (speaker) of the user terminal 2.

[0073]

The NFT issuance unit 32e issues an NFT (Non-Fungible Token) related to the content data generated by the content generation unit 32c or the image data of the character in the content (the character image (two-dimensional image or three-dimensional image) data generated by the character image generation unit 32a). That is, the NFT issuance unit 32e creates an NFT from the content generated by the content generation unit 32c and/or the image data of the character in the content. Specifically, for example, when receiving a direction for creating an NFT of content data or character image data in response to an operation on an interface (not illustrated) shown in the display unit (output unit 24) of the user terminal 2, the NFT issuance unit 32e generates an NFTID associated with the content data or the character image data. Then, the NFT issuance unit 32e generates a code having metadata including the generated NFTID, and sends the code to a block chain network (not illustrated). Thereby, an NFT can be created from content data or character image data in the content. In addition, the user can buy or sell the NFT of content data or character image data at an NFT market, or can exchange the NFT with other NFT data.

[0074]

The storage unit 33 is, for example, a data storage such as an internal memory and an external memory (such as SD memory card). The storage unit 33 stores various data handled by the control unit 32; various information received by the communication unit 31 from the user terminal 2 via the network 4; various databases (DB); and the like. The databases include a character image DB 33a, a background image DB 33b, a balloon image DB 33c, and a content DB 33d.

[0075]

The character image DB 33a stores image data of a plurality of characters prepared in advance (including 2D or 3D model data). In addition, the character image DB 33a stores character images generated by the character image generation unit 32a of the control unit 32. The character images stored in the character image DB 33a may be associated with a plurality of motion patterns. In response to a request from the user terminal 2, a character image stored in the character image DB 33a is sent by the communication unit 31 to the user terminal 2 via the network 4, and displayed by the display control unit 22a of the user terminal 2 in the display unit (output unit 24).

[0076]

The background image DB 33b stores image data of a

plurality of background images prepared in advance. In response to a request from the user terminal 2, a background image stored in the background image DB 33b is sent by the communication unit 31 to the user terminal 2 via the network 4, and displayed by the display control unit 22a of the user terminal 2 in the display unit (output unit 24).

[0077]

The balloon image DB 33c stores image data of a plurality of balloon images prepared in advance. In response to a request from the user terminal 2, a balloon image stored in the balloon image DB 33c is sent by the communication unit 31 to the user terminal 2 via the network 4, and displayed by the display control unit 22a of the user terminal 2 in the display unit (output unit 24).

[0078]

The content DB 33d stores content data generated by the content generation unit 32c. In response to a request from the user terminal 2, content data stored in the content DB 33d is sent by the communication unit 31 to the user terminal 2 via the network 4, and displayed by the display control unit 22a of the user terminal 2 in the display unit (output unit 24).

[0079]

Note that the storage unit 33 is not necessarily

provided inside the server 3, and a part or all of the storage unit 33 may be provided inside a separate device connected to the server 3 via the network 4 so that they can communicate with each other.

[0080]

(Example of Operation)

Next, an example of the operation of the content generation system 1 will be described. Fig. 4 is a flowchart illustrating an operation in which the user of the content generation system 1 uses a user interface shown in the display unit (output unit 24) of the user terminal 2 to generate content constituted by one or more frames.

[0081]

First, in the user terminal 2, the display control unit 22a allows the display unit (output unit 24) to display a user interface (Step S1).

[0082]

Next, in the user terminal 2, depending on the user operating the user interface shown in the display unit (output unit 24), the frame determination unit 22b determines a frame to be edited (Step S2).

[0083]

Next, in the user terminal 2, depending on the user operating the user interface shown in the display unit (output unit 24), the character determination unit 22c

determines a character (Step S3).

[0084]

Next, in the user terminal 2, depending on the user operating the user interface (expression setting unit 55) shown in the display unit (output unit 24), the expression determination unit 22d determines the expression of the character (Step S4).

[0085]

Next, in the user terminal 2, depending on the user operating the user interface (gaze setting unit 56) shown in the display unit (output unit 24), the gaze determination unit 22e determines the gaze of the character (Step S5).

[0086]

Next, in the user terminal 2, depending on the user operating the user interface (motion setting unit 57) shown in the display unit (output unit 24), the motion determination unit 22f determines the motion of the character (Step S6).

[0087]

Next, in the user terminal 2, depending on the user operating the user interface (character arrangement setting unit 58) shown in the display unit (output unit 24), the character arrangement determination unit 22g determines the arrangement of the character (Step S7).

[0088]

Next, in the user terminal 2, depending on the user operating the user interface (background image setting unit 59) shown in the display unit (output unit 24), the background image determination unit 22h determines a background image (Step S8).

[0089]

Next, in the user terminal 2, depending on the user operating the user interface (text information setting unit 60) shown in the display unit (output unit 24), the text information determination unit 22i determines text information (Step S9).

[0090]

Next, in the server 3, the content generation unit 32c generates content including one or more frames (Step S10).

[0091]

As described above, the content generation system 1 according to the embodiment is a content generation system 1 for generating content constituted by one or more frames including a moving image or a still image of a character, the content generation system 1 including: a display control unit 22a that allows a display unit (output unit 24) to display a user interface including a two-dimensional operation region 55a (expression setting unit 55) and a preview image of a frame to be edited by a user (preview

display area 53); an expression determination unit 22d that determines an expression of a character on the preview image depending on a position of a pointer 55b on the two-dimensional operation region 55a; and a content generation unit 32c that generates content constituted by one or more frames including the frame edited by the user. Thereby, it is possible to easily generate content constituted by frames including a character having an expression intended by a user.

[0092]

Although the embodiment has been described with reference to an example where the expression setting unit 55 sets expressions corresponding to four feelings of "happy", "angry", "sad", and "enjoy", expressions corresponding to four or more feelings may be set. For example, expressions corresponding to eight feelings of "angry", "fear", "expect", "surprise", "happy", "sad", "trust", and "hate" (Plutchik model) may be set.

[0093]

Although the embodiment has been described with reference to an example where the motion of the character is set using the pull-down menu 571, expressions that can be set by the expression setting unit 55 may be associated with corresponding motions in advance so that a motion is automatically set corresponding to an expression set by the

expression setting unit 55.

[0094]

Although the embodiment has been described with reference to an example where the expression setting unit 55 and the motion setting unit 57 are separately provided, and the setting units each set an "expression" and a "motion", a single setting unit may set an "expression" and a "motion". For example, one feeling setting unit having a two-dimensional operation region capable of setting a plurality of feelings may be provided as a user interface, so that the feelings set by the feeling setting unit are associated with the corresponding "expression" and "motion" of the character in advance, and an "expression" and a "motion" are automatically set corresponding to a feeling set by the feeling setting unit. Here, the type of the feelings to be set by the feeling setting unit may be, for example, four feelings of "happy", "angry", "sad", and "enjoy", may be eight feelings of "angry", "fear", "expect", "surprise", "happy", "sad", "trust", and "hate" (Plutchik model), or may be more types of feelings than them.

[0095]

Although the embodiment has been described with reference to an example where the expression setting unit 55 has a two-dimensional operation region 55a having a

rectangular shape, the shape of the two-dimensional operation region is not limited thereto. For example, as illustrated in Fig. 17, the expression setting unit 55 may have a two-dimensional operation region 55c having a circular shape so that a pointer 55d having a dot shape is moved inside the two-dimensional operation region 55c having a circular shape to set each expression, such as happy, angry, sad, and enjoy, or the shape of the two-dimensional operation region may be a shape other than a rectangular shape or a circular shape so that a pointer is moved inside the shape to set each expression.

[0096]

Although the embodiment has been described with reference to an example where the gaze setting unit 56 has a two-dimensional operation region 56a having a rectangular shape, the shape of the two-dimensional operation region is not limited thereto. For example, as illustrated in Fig. 18, the gaze setting unit 56 may have a two-dimensional operation region 56c having a circular shape so that a pointer 56d having a dot shape is moved inside the two-dimensional operation region 56c having a circular shape to set a gaze in each direction, or the shape of the two-dimensional operation region may be a shape other than a rectangular shape or a circular shape so that a pointer is moved inside the shape to set a gaze in each direction.

[0097]

Any part or all of each functional unit described in the specification may be realized by a program. The program mentioned in the specification may be recorded in a non-transitory manner on a computer-readable storage medium for distribution, may be distributed via a communication circuit such as the internet (including wireless communication), or may be distributed in a state where the program is installed on any terminal.

[0098]

Based on the above descriptions, a person skilled in the art may conceive an additional effect of the present invention or various modifications, and embodiments of the present invention are not limited to the individual embodiments described above. Various additions, changes, and partial deletion are possible in a range without departing from the conceptual idea and gist of the present invention that are derived from the content provided in Claims and equivalents thereof.

[0099]

For example, those described in the specification as one device (or member, the same applies hereinafter) (including those illustrated as one device in the drawings) may be realized as a plurality of devices. To the contrary, those described in the specification as a

plurality of devices (including those illustrated as a plurality of devices in the drawings) may be realized as one device. Alternatively, a part or all of the means or functions to be included in a device (for example, a server) may be included in another device (for example, a user terminal).

[0100]

All contents described in the specification are not always essential requirements. Especially, a content described in the specification but not described in Claims can be an arbitrary additional content.

[0101]

Note that the applicant only knows the inventions described in a publication that are described as literature in the section of "Citation List" in the specification, and the present invention does not always aim to solve the problems of the inventions described in a publication. The problems to be solved by the present invention should be acknowledged in consideration of the entirety of the specification. For example, when the specification describes that a specific configuration exhibits a predetermined effect, it can be acknowledged that a problem behind the predetermined effect is solved. However, the gist is not always to essentially require such a specific configuration.

CLAIMS

[Claim 1]

A content generation system for generating content constituted by one or more frames including a moving image or a still image of a character, the content generation system comprising:

a display control unit configured to allow a display unit to display a user interface including a first two-dimensional operation region and a preview image of a frame to be edited by a user;

an expression determination unit configured to determine an expression of a character on the preview image depending on a position of a pointer on the first two-dimensional operation region; and

a content generation unit configured to generate content constituted by one or more frames including the frame edited by the user.

[Claim 2]

The content generation system according to claim 1, wherein the user interface includes a second two-dimensional operation region, and the content generation system further comprising: a gaze determination unit configured to determine a gaze of a character on the preview image depending on a position of a pointer on the second two-dimensional operation region.

[Claim 3]

The content generation system according to claim 1, further comprising: a motion determination unit configured to determine a motion of the character on the preview image depending on an operation of the user interface by the user.

[Claim 4]

The content generation system according to claim 3, wherein

the user interface includes a one-dimensional operation region, and

the motion determination unit determines a playback range of the motion depending on a position of a knob on the one-dimensional operation region.

[Claim 5]

The content generation system according to claim 1, further comprising: a motion determination unit configured to determine a motion of the character on the preview image depending on a position of a pointer on the first two-dimensional operation region.

[Claim 6]

The content generation system according to claim 1, further comprising: a character determination unit configured to determine a character to be displayed on the preview image.

[Claim 7]

The content generation system according to claim 1, further comprising: a background image determination unit configured to determine a background image of the preview image.

[Claim 8]

The content generation system according to claim 1, further comprising: a text information determination unit configured to determine at least one of a text to be displayed on the preview image, a type of balloon to display the text, a color of the text, and a size of the text depending on an operation of the user interface by the user.

[Claim 9]

The content generation system according to claim 1, wherein the first two-dimensional operation region is a user interface to set a feeling of the character.

[Claim 10]

The content generation system according to claim 1, further comprising: a voice generation unit where image data of a character selected by the user is input to a voice model using a character image and voice data associated with the character image as learning data to generate voice data of a character.

[Claim 11]

The content generation system according to claim 1, further comprising: a speech data generation unit where speech data of the user is input to a language model learned by using image data and/or voice data of a character as a part of learning data to generate speech data of a character in response to speech data of the user.

[Claim 12]

The content generation system according to claim 1, further comprising: a character image generation unit where image data input by the user is input to a learned image processing model to generate a character image.

[Claim 13]

The content generation system according to claim 12, further comprising: a background image generation unit where image data input by the user is input to a learned image processing model to generate a background image.

[Claim 14]

The content generation system according to claim 13, wherein the character image generation unit and the background image generation unit each generate three-dimensional image data, each of which is arranged on a metaverse space.

[Claim 15]

The content generation system according to claim 12, further comprising: an NFT issuance unit configured to

issue an NFT related to content generated by the content generation unit or image data of a character in the content.

[Claim 16]

A method for generating content constituted by one or more frames including a moving image or a still image of a character, the method comprising:

a display control step of allowing a display unit to display a user interface including a first two-dimensional operation region and a preview image of a frame to be edited by a user;

an expression determination step of determining an expression of a character on the preview image depending on a position of a pointer on the first two-dimensional operation region; and

a content generation step of generating content constituted by one or more frames including the frame edited by the user.

[Claim 17]

A program for making a computer execute a method for generating content constituted by one or more frames including a moving image or a still image of a character, the method comprising:

a display control step of allowing a display unit to display a user interface including a first two-dimensional

operation region and a preview image of a frame to be edited by a user;

an expression determination step of determining an expression of a character on the preview image depending on a position of a pointer on the first two-dimensional operation region; and

a content generation step of generating content constituted by one or more frames including the frame edited by the user.

[Claim 18]

A computer-readable storage medium storing the program according to claim 17.

[Claim 19]

A content generation system for generating content constituted by one or more frames including a moving image or a still image of a character, the content generation system comprising:

a display control unit configured to allow a display unit to display a user interface and a preview image of a frame to be edited by a user;

a character determination unit configured to determine a character to be displayed on the preview image depending on an operation of the user interface by the user;

an expression determination unit configured to determine an expression of the character depending on an

operation of the user interface by the user;

a gaze determination unit configured to determine a gaze of the character depending on an operation of the user interface by the user;

a motion determination unit configured to determine a motion of the character depending on an operation of the user interface by the user;

a background determination unit configured to determine a background image of the preview image depending on an operation of the user interface by the user;

a text information determination unit configured to determine at least one of a text to be displayed on the preview image, a type of balloon to display the text, a color of the text, and a size of the text depending on an operation of the user interface by the user; and

a content generation unit configured to generate content constituted by one or more frames including the frame edited by the user.