US Patent & Trademark Office
Patent Public Search | Text View

United States Patent Application Publication	20250266050
Kind Code	A1
Publication Date	August 21, 2025
Inventor(s)	NISHIYAMA; Yukinori et al.

CREATION METHOD AND CREATION APPARATUS

Abstract

A creation method of the present disclosure includes an acquisition step of acquiring first sound data of a floating point format, based on a sound that is generated from a sound source and is collected by a sound collection device, and an accessory information creation step of creating accessory information that is attached to the first sound data and includes device information, which relates to the sound collection device, or sound source information, which relates to the sound source.

Inventors: **NISHIYAMA; Yukinori (Saitama-shi, JP), OKIYAMA; Kazuya (Saitama-shi, JP), OTSURU; Mototada (Saitama-shi, JP)**

Applicant: **FUJIFILM Corporation** (Tokyo, JP)

Family ID: **1000008631014**

Assignee: **FUJIFILM Corporation (Tokyo, JP)**

Appl. No.: **19/203156**

Filed: **May 08, 2025**

Foreign Application Priority Data

JP	2022-186836	Nov. 22, 2022

Related U.S. Application Data

parent WO continuation PCT/JP2023/037767 20231018 PENDING child US 19203156

Publication Classification

Int. Cl.: **G10L19/018** (20130101); **H03M7/30** (20060101)

U.S. Cl.:

CPC **G10L19/018** (20130101); **H03M7/3059** (20130101);

Background/Summary

CROSS-REFERENCE TO RELATED APPLICATIONS [0001] This application is a continuation application of International Application No. PCT/JP2023/037767, filed Oct. 18, 2023, the disclosure of which is incorporated herein by reference in its entirety. Further, this application claims priority from Japanese Patent Application No. 2022-186836, filed on Nov. 22, 2022, the disclosure of which is incorporated herein by reference in its entirety.

BACKGROUND
1. Technical Field
[0002] The technique of the present disclosure relates to a creation method and a creation apparatus.
2. Description of the Related Art
[0003] JP2012-073435A discloses an audio signal conversion device that samples an input analog audio signal of an L channel and an R channel with a sampling frequency of 192 kHz and a quantization bit rate of 24 bits in an A/D conversion device to generate a digital signal. A signal processing device is connected to an output side of the A/D conversion device. This signal processing device performs processing of down-sampling the frequency to ¼ (48 kHz) and processing of converting the down-sampled signal into a floating point format of a quantization bit rate of 32 bits.
[0004] JP2002-246913A discloses a data processing device that converts input data from a fixed point format to a floating point format by a conversion unit.
SUMMARY
[0005] An object of one embodiment according to the technique of the present disclosure is to provide a creation method and a creation apparatus capable of improving quality of sound data.
[0006] In order to achieve the above object, a creation method of the present disclosure comprises an acquisition step of acquiring first sound data of a floating point format, based on a sound that is generated from a sound source and is collected by a sound collection device, and an accessory information creation step of creating accessory information that is attached to the first sound data and includes device information, which relates to the sound collection device, or sound source information, which relates to the sound source.
[0007] It is preferable that the first sound data is used to create second sound data having the number of bits, which is smaller than the number of bits of the first sound data.
[0008] It is preferable that the accessory information includes the device information. In this case, it is preferable that the device information relates to gain processing used for the sound collected by the sound collection device or relates to performance of the sound collection device.
[0009] It is preferable that the accessory information includes the sound source information. In this case, it is preferable that the sound source information is associated with time information included in the first sound data.
[0010] It is preferable that the creation method further comprises an imaging step of creating, by an imaging apparatus, video data corresponding to the first sound data.
[0011] It is preferable that the sound source is a subject included in the video data.
[0012] It is preferable that the sound source is a main subject selected from a plurality of subjects included in the video data.
[0013] It is preferable that the sound source information is a type of a drive sound accompanied by

drive of the imaging apparatus.
[0014] It is preferable that the creation method further comprises a first file creation step of creating a first file including the first sound data and the accessory information.
[0015] It is preferable that the creation method further comprises an editing step of editing the first sound data based on the accessory information to create second sound data having the number of bits, which is smaller than the number of bits of the first sound data.
[0016] In the editing step, deterioration information that relates to a sound of the second sound data deteriorated by the editing may be created, and a second file including the second sound data and the deterioration information may be created. It is preferable that the second file includes the accessory information.
[0017] A creation apparatus according to the present disclosure comprises a processor, in which the processor is configured to execute an acquisition step of acquiring first sound data of a floating point format, based on a sound that is generated from a sound source and is collected by a sound collection device, and an accessory information creation step of creating accessory information that is attached to the first sound data and includes device information, which relates to the sound collection device, or sound source information, which relates to the sound source.

Description

BRIEF DESCRIPTION OF THE DRAWINGS
[0018] Exemplary embodiments according to the technique of the present disclosure will be described in detail based on the following figures, wherein:
[0019] FIG. **1** is a diagram showing an example of a configuration of an imaging apparatus according to a first embodiment,
[0020] FIG. **2** is a diagram showing an example of a configuration of a sound signal processing circuit,
[0021] FIG. **3** is a diagram conceptually showing sound signal processing,
[0022] FIG. **4** is a diagram showing an example of a functional configuration of a processor,
[0023] FIG. **5** is a diagram conceptually showing combination processing and data format conversion processing,
[0024] FIG. **6** is a diagram conceptually showing editing processing,
[0025] FIG. **7** is a flowchart showing an example of an operation of the imaging apparatus,
[0026] FIG. **8** is a diagram showing an example of a functional configuration of the processor according to a second embodiment,
[0027] FIG. **9** is a diagram showing an example of a relationship between sound source information and first sound data,
[0028] FIG. **10** is a flowchart showing an example of an operation of the imaging apparatus according to the second embodiment,
[0029] FIG. **11** is a diagram showing an example of a functional configuration of the processor according to a third embodiment,
[0030] FIG. **12** is a diagram showing an example of the relationship between the sound source information and the first sound data according to the third embodiment, and
[0031] FIG. **13** is a diagram showing an example of a sound data file created by an editing unit.
DETAILED DESCRIPTION
[0032] An example of an embodiment according to the technique of the present disclosure will be described with reference to accompanying drawings.
[0033] First, terms used in the following description will be described.
[0034] In the following description, “AF” is an abbreviation for “auto focus”. “MF” is an abbreviation for “manual focus”. “IC” is an abbreviation for “integrated circuit”. “CPU” is an

abbreviation for "central processing unit". "RAM" is an abbreviation for "random access memory". "CMOS" is an abbreviation for "complementary metal oxide semiconductor".

[0035] "FPGA" is an abbreviation for "field programmable gate array". "PLD" is an abbreviation for "programmable logic device". "ASIC" is an abbreviation for "application specific integrated circuit". "OVF" is an abbreviation for "optical view finder". "EVF" is an abbreviation for "electronic view finder". "ADC" is an abbreviation for "analog to digital converter". "LPCM" is an abbreviation for "linear pulse code modulation".

[0036] As an embodiment of an imaging apparatus, the technique of the present disclosure will be described by using a lens-interchangeable digital camera as an example. The technique of the present disclosure is not limited to the lens-interchangeable type, and can be employed for a lens-integrated digital camera.

First Embodiment

[0037] FIG. **1** shows an example of a configuration of an imaging apparatus **10** according to a first embodiment. The imaging apparatus **10** is the lens-interchangeable digital camera. The imaging apparatus **10** is configured of a housing **11** and an imaging lens **12** that is interchangeably mounted on the housing **11** and includes a focus lens **31**. The imaging lens **12** is attached to a front surface side of the housing **11** via a mount **11**A. The imaging apparatus **10** is an example of "creation apparatus" according to the technique of the present disclosure.

[0038] Further, an external microphone **13** can be attachably and detachably attached to the housing **11**. The external microphone **13** is attached to the housing **11** via a connecting part **11**B provided on an upper surface of the housing **11**. The external microphone **13** is a gun microphone, a zoom microphone, or the like. The connecting part **11**B is, for example, a hot shoe. The external microphone **13** is an example of "sound collection device" according to the technique of the present disclosure.

[0039] The housing **11** is provided with an operation unit **16** including a dial, a release button, and the like. Examples of an operation mode of the imaging apparatus **10** include a still image capturing mode, a video capturing mode, and an image display mode. The operation unit **16** is operated by a user in a case where the operation mode is set. Further, the operation unit **16** is operated by the user in a case where execution of still image capturing or video capturing is started.

[0040] Further, the operation unit **16** is operated by the user in a case where a focusing mode is selected. The focusing mode includes an AF mode and an MF mode. In the AF mode, a subject area selected by the user or a subject area automatically detected by the imaging apparatus **10** is set as a focus detection area (hereinafter referred to as AF area) to perform focusing control. In the MF mode, the user operates a focus ring (not shown) to manually perform the focusing control.

[0041] Further, the housing **11** is provided with a finder **14**. For example, the finder **14** is a hybrid finder (registered trademark). The hybrid finder refers to, for example, a finder in which an optical view finder (hereinafter referred to as "OVF") and an electronic view finder (hereinafter referred to as "EVF") are selectively used. The user can observe an optical image or live view image of a subject projected onto the finder **14** via a finder eyepiece portion (not shown).

[0042] Further, a display **15** is provided on a rear surface side of the housing **11**. The display **15** displays an image based on video data PD obtained by imaging, various menu screens, and the like. The user can also observe the live view image projected onto the display **15**, instead of the finder **14**.

[0043] Further, the housing **11** is provided with a speaker **17**. The speaker **17** outputs a sound based on second sound data AS**2** included in a moving image file **28** described below.

[0044] The housing **11** is electrically connected to the imaging lens **12** via an electrical contact **11**C provided on the mount **11**A.

[0045] The imaging lens **12** includes a focus lens **31**, a stop **32**, and a lens driving controller **33**. The lens driving controller **33** is electrically connected to a processor **25** accommodated in the housing **11**, via the electrical contact **11**C.

[0046] The lens driving controller **33** drives the focus lens **31** and the stop **32**, based on control signals transmitted from the processor **25**. The lens driving controller **33** performs drive control of the focus lens **31**, based on the control signal for the focusing control that is transmitted from the processor **25**, in order to adjust a position of the focus lens **31**.

[0047] The stop **32** has an opening with a variable opening diameter. The lens driving controller **33** performs drive control of the stop **32**, based on the control signal for stop adjustment that is transmitted from the processor **25**, in order to adjust an amount of light incident on an imaging sensor **20**.

[0048] Further, the imaging sensor **20**, an image processing circuit **21**, a built-in microphone **22**, a sound signal processing circuit **23**, the processor **25**, and a storage device **26** are provided inside the housing **11**. The processor **25** controls operations of the imaging sensor **20**, the image processing circuit **21**, the built-in microphone **22**, the sound signal processing circuit **23**, the storage device **26**, the display **15**, and the speaker **17**.

[0049] The processor **25** is configured of, for example, a CPU. The processor **25** is connected to a RAM **25**A, which is a memory for primary storage. The storage device **26** is configured of, for example, a non-volatile memory such as a flash memory. The processor **25** executes various types of processing based on a program **27** stored in the storage device **26**. The processor **25** may be configured of an assembly of a plurality of IC chips.

[0050] The imaging sensor **20** is, for example, a CMOS-type image sensor. Light (subject image) that has passed through the imaging lens **12** is incident on a light-receiving surface **20**A of the imaging sensor **20**. A plurality of pixels that generate imaging signals through photoelectric conversion are formed on the light-receiving surface **20**A. The imaging sensor **20** performs the photoelectric conversion on light incident on each pixel to generate and output the video data PD.

[0051] The image processing circuit **21** performs, on the video data PD output from the imaging sensor **20**, image processing including white balance correction, gamma correction processing, and the like.

[0052] The built-in microphone **22** is a stereo microphone including a pair of sound collection elements **22**A and **22**B. The sound collection elements **22**A and **22**B are sound sensors for a left side channel (hereinafter referred to as L channel) and a right side channel (hereinafter referred to as R channel). The sound collection elements **22**A and **22**B are sound sensors of an electrostatic type, a piezoelectric type, an electrodynamic type, or the like, and output collected sounds as sound signals AL and AR. The sound signal processing circuit **23** performs sound signal processing including gain processing, A/D conversion processing, and the like on the sound signals AL and AR output from the sound collection elements **22**A and **22**B.

[0053] The external microphone **13** includes a sound collection element **41**, an amplifier **42**, and a microphone control unit **43**. In the present embodiment, the external microphone **13** is a mono microphone having one sound collection element **41**. The sound collection element **41** is a sound sensor of an electrostatic type, a piezoelectric type, an electrodynamic type, or the like, and outputs a collected sound as the sound signal. The amplifier **42** performs the gain processing on the sound signal output from the sound collection element **41**. The microphone control unit **43** controls a gain amount of the gain processing by the amplifier **42**.

[0054] Further, the microphone control unit **43** supplies the sound signal subjected to the gain processing by the amplifier **42** to the sound signal processing circuit **23** in the housing **11** via the connecting part **11**B. A mono and analog sound signal AS is supplied from the external microphone **13** to the sound signal processing circuit **23**. The processor **25** controls the operation of the microphone control unit **43**.

[0055] Further, the storage device **26** stores device information **13**A, which is information related to the external microphone **13**. In the present embodiment, information related to the gain processing used for the collected sound using the device information **13**A is employed. Specifically, the device information **13**A is the gain amount (that is, sensitivity of the external microphone **13**) of the gain processing by the amplifier **42**.

[0056] FIG. **2** shows an example of a configuration of the sound signal processing circuit **23**. The sound signal processing circuit **23** includes a first preamplifier **51**A, a first ADC **52**A, a second preamplifier **51**B, and a second ADC **52**B.

[0057] The first preamplifier **51**A and the first ADC **52**A are processing units for L-channel that perform the gain processing and the A/D conversion processing on the sound signal AL output from the sound collection element **22**A included in the built-in microphone **22**. The second preamplifier **51**B and the second ADC **52**B are processing units for R-channel that perform the gain processing and the A/D conversion processing on the sound signal AR output from the sound collection element **22**B included in the built-in microphone **22**.

[0058] In the first preamplifier **51**A, the processor **25** controls a gain amount G**1**. In the second preamplifier **51**B, the processor **25** controls a gain amount G**2**. In a case where the gain processing is performed on the sound signals AL and AR output from the built-in microphone **22**, the processor **25** sets the gain amount G**1** and the gain amount G**2** to the same value. The first ADC **52**A and the second ADC **52**B perform sampling with, for example, a quantization bit rate of 24 bits to convert an analog sound signal into a digital signal of a 24-bit LPCM format. The LPCM format is an example of “pulse code modulation format” according to the technique of the present disclosure.

[0059] The sound signal AS output from the external microphone **13** is input to the first preamplifier **51**A and the second preamplifier **51**B. The first preamplifier **51**A performs the gain processing on the sound signal AS with the gain amount G**1**. The second preamplifier **51**B performs the gain processing on the sound signal AS with the gain amount G**2**. In a case where the gain processing is performed on the sound signal AS output from the external microphone **13**, the processor **25** sets the gain amount G**1** and the gain amount G**2** to different values. Hereinafter, the gain processing performed by the first preamplifier **51**A is referred to as first gain processing, and the gain processing performed by the second preamplifier **51**B is referred to as second gain processing.

[0060] The first ADC **52**A converts the sound signal AS subjected to the first gain processing by the first preamplifier **51**A into the digital signal. The second ADC **52**B converts the sound signal AS subjected to the second gain processing by the second preamplifier **51**B into the digital signal. Hereinafter, the sound signal AS digitized by the first ADC **52**A is referred to as modulation sound data ASH, and the sound signal AS digitized by the second ADC **52**B is referred to as modulation sound data ASL. The modulation sound data ASH and ASL are output from the sound signal processing circuit **23** to the processor **25**.

[0061] FIG. **3** conceptually shows sound signal processing of the sound signal AS by the sound signal processing circuit **23**. The sound signal AS output from the external microphone **13** is input to the processing unit for L-channel and the processing unit for R-channel. The sound signal AS input to the processing unit for L-channel is subjected to the first gain processing with the gain amount G**1**, then is converted into the digital signal, and thus, is output from the sound signal processing circuit **23** as the modulation sound data ASH. The sound signal AS input to the processing unit for R-channel is subjected to the second gain processing with the gain amount G**2**, then is converted into the digital signal, and thus, is output from the sound signal processing circuit **23** as the modulation sound data ASL. In the present embodiment, the number of bits of the modulation sound data ASH and ASL is 24 bits.

[0062] For example, the gain amount G**1** is assumed to be +48 dB, and the gain amount G**2** is assumed to be −48 dB. Since 48 dB corresponds to a volume width of 8 bits, there is a deviation of 16 bits between the modulation sound data ASH of high gain and the modulation sound data ASL of low gain, as shown in FIG. **3**. In other words, the modulation sound data ASH overlaps with the modulation sound data ASL by 8 bits.

[0063] FIG. **4** shows an example of a functional configuration of the processor **25**. The processor

25 executes the processing according to the program **27**, which is stored in the storage device **26**, to implement various functional units. Various functional units shown in FIG. **4** are implemented in the video capturing mode. As shown in FIG. **4**, for example, a main controller **60**, a combination processing unit **61**, a data format conversion unit **62**, an accessory information creation unit **63**, a sound data file creation unit **64**, an editing unit **65**, and a file creation unit **66** are implemented in the processor **25**. The editing unit **65** includes a volume range setting unit **65**A and a data extraction unit **65**B.

[0064] The main controller **60** integrally controls each unit of the imaging apparatus **10**. The main controller **60** controls the operation of the imaging apparatus **10** based on an instruction signal input from the operation unit **16**. The main controller **60** controls the imaging sensor **20** to cause the imaging sensor **20** to perform the imaging operation. The imaging sensor **20** outputs the video data PD, which is generated by performing the imaging via the imaging lens **12**. In the video capturing mode, the imaging sensor **20** outputs the video data PD for each frame cycle. The video data PD output from the imaging sensor **20** is subjected to the image processing by the image processing circuit **21** and then input to the processor **25**. In a case of the video capturing mode, the video data PD is data consisting of a plurality of frames.

[0065] Further, in the video capturing mode, in a case where the external microphone **13** is connected to the connecting part **11**B, the main controller **60** controls the external microphone **13** to perform a sound collection operation. The external microphone **13** outputs the sound signal AS to the sound signal processing circuit **23** via the connecting part **11**B while the imaging sensor **20** performs the imaging operation. The sound signal processing circuit **23** performs the above sound signal processing to output the modulation sound data ASH and ASL. The modulation sound data ASH and ASL correspond to the video data PD obtained by the imaging sensor **20** imaging the subject.

[0066] The combination processing unit **61** acquires the modulation sound data ASH and ASL output from the sound signal processing circuit **23**, and combines the modulation sound data ASH and ASL to create first sound data AS**1** having a first number of bits. The first sound data AS**1** is digital data of the LPCM format.

[0067] The data format conversion unit **62** converts a data format of the first sound data AS**1** into a floating point format. Hereinafter, the first sound data AS**1** converted into the floating point format is referred to as first sound data AS**1**F. The first sound data AS**1**F is used to create the second sound data AS**2** having the number of bits smaller than the number of bits of the first sound data AS**1**F.

[0068] The accessory information creation unit **63** reads out the device information **13**A from the storage device **26** to create accessory information SI including the device information **13**A. The accessory information creation unit **63** supplies the created accessory information SI to the sound data file creation unit **64**. The accessory information SI is so-called meta information.

[0069] The sound data file creation unit **64** creates a sound data file **67** including the first sound data AS**1**F, which is created by the data format conversion unit **62**, and the accessory information SI, which is created by the accessory information creation unit **63**. The sound data file creation unit **64** records the created sound data file **67** in the storage device **26**. The sound data file **67** corresponds to “first file” according to the technique of the present disclosure.

[0070] The editing unit **65** edits the first sound data AS**1**F included in the sound data file **67**, which is recorded in the storage device **26**, based on the accessory information SI to create the second sound data AS**2** having a second number of bits smaller than the first number of bits. For example, the second number of bits is 24 bits.

[0071] Specifically, the volume range setting unit **65**A sets a volume range VR having a width of the second number of bits for a dynamic range of the first sound data AS**1**F. In the present embodiment, the volume range setting unit **65**A sets the volume range VR based on the accessory information SI. The data extraction unit **65**B extracts data of the volume range VR set by the volume range setting unit **65**A to create the second sound data AS**2**, based on the first sound data

AS**1**F.

[0072] The file creation unit **66** creates the moving image file **28** including the video data PD, which is output from the image processing circuit **21**, and the second sound data AS**2**, which is output from the data extraction unit **65**B, and stores the moving image file **28** in the storage device **26**.

[0073] FIG. **5** conceptually shows combination processing by the combination processing unit **61** and data format conversion processing by the data format conversion unit **62**. The combination processing unit **61** performs the mixing process on the overlap portion of 8 bits between the modulation sound data ASH and ASL to combine the modulation sound data ASH and the modulation sound data ASL. The number of bits (that is, the first number of bits) of the first sound data AS**1**, which is generated by the combination processing, is 40 bits. In this manner, with the combination of the modulation sound data ASH and ASL having different gain amounts, it is possible to obtain the first sound data AS**1** with an expanded volume dynamic range.

[0074] The data format conversion unit **62** converts the first sound data AS**1** of a 40-bit fixed point format into the first sound data AS**1**F of a 32-bit floating point format (so-called 32-bit float). The 32-bit float is configured of a 1-bit sign, an 8-bit exponent part, and a 23-bit mantissa part. A known method can be used for the conversion from the fixed point format to the floating point format. In the floating point format, a wide range of numerical values can be expressed.

[0075] FIG. **6** conceptually shows editing processing by the editing unit **65**. The volume range setting unit **65**A sets the volume range VR based on the device information **13**A included in the accessory information SI. Specifically, the volume range setting unit **65**A sets the volume range VR based on the gain amount (that is, sensitivity of the external microphone **13**) of the gain processing of the amplifier **42**, which is an example of the device information **13**A. For example, the volume range setting unit **65**A sets the volume range VR to a low volume side as the gain amount is larger (that is, sensitivity is higher), and sets the volume range VR to a high volume side as the gain amount is smaller (that is, sensitivity is lower). Accordingly, it is possible to suppress sound cracking and the like.

[0076] The data extraction unit **65**B extracts the data of the volume range VR to create the second sound data AS**2**, based on the first sound data AS**1**F. In this manner, with the creation of the second sound data AS**2** based on the device information **13**A, it is possible to generate the second sound data AS**2** of the 24-bit fixed point format in which the sound cracking and the like are suppressed.

[0077] FIG. **7** is a flowchart showing an example of the operation of the imaging apparatus **10**. FIG. **7** shows an operation in a case where the video capturing mode is selected as the operation mode and the external microphone **13** is connected to the connecting part **11**B.

[0078] First, the main controller **60** determines whether or not the user issues a start instruction for the video capturing (step S**10**). In a case where the start instruction is determined to be issued (YES in step S**10**), an imaging step (step S**11**) and an acquisition step (step S**12**) are executed in parallel. In the imaging step, the imaging sensor **20** images the subject to generate the video data PD. In the acquisition step, the first sound data AS**1**F of the floating point format is acquired based on the sound collected by the external microphone **13**.

[0079] After the imaging step and the acquisition step, the main controller **60** determines whether or not the user issues an end instruction for the video capturing (step S**13**). In a case where the end instruction is determined to be not issued (NO in step S**13**), the processing returns to steps S**11** and S**12**. Steps S**11** to S**12** are repeatedly executed until the end instruction is determined to be issued in step S**13**.

[0080] In a case where the end instruction is determined to be issued (YES in step S**13**), an accessory information creation step is executed (step S**14**). In the accessory information creation step, the accessory information creation unit **63** creates the accessory information SI including the device information **13**A.

[0081] After the accessory information creation step, a sound data file creation step is executed (step S**15**). In the sound data file creation step, the sound data file creation unit **64** creates the sound data file **67** including the first sound data AS**1**F and the accessory information SI. The sound data file creation step corresponds to "first file creation step" according to the technique of the present disclosure.

[0082] After the sound data file creation step, an editing step is executed (step S**16**). In the editing step, the first sound data AS**1**F is edited based on the accessory information SI to create the second sound data AS**2**. Thereafter, the moving image file **28** including the video data PD and the second sound data AS**2** is created and recorded in the storage device **26**. The operation of the imaging apparatus **10** is ended as described above.

[0083] As described above, a creation method of the present disclosure includes the acquisition step of acquiring the first sound data of the floating point format, based on the sound that is generated from a sound source such as the subject and is collected by the sound collection device, and the creation step of creating the device information that is the accessory information attached to the first sound data and is information related to the sound collection device. Accordingly, it is possible to improve quality of the sound data.

[0084] In the above embodiment, the device information **13**A is the information related to the

[0085] sound collection device, but the device information **13**A may be information related to performance of the sound collection device. The information related to the performance of the sound collection device includes a noise level, a maximum sound pressure level, an output impedance, and the like. In this case, the volume range setting unit **65**A sets the volume range VR based on the information related to the performance of the external microphone **13**.

[0086] The noise level represents susceptibility of the external microphone **13** to noise. In a case where the performance of the sound collection device is the noise level, the volume range setting unit **65**A limits an upper limit of the volume range VR according to the noise level, for example. This is because, in a case where the noise level is small and the volume range VR is set to the high volume side, the sound from the subject as the sound source is difficult to hear due to the noise.

[0087] The maximum sound pressure level represents a sound pressure level at maximum at which the external microphone **13** can perform the sound collection. In a case where the performance of the sound collection device is the maximum sound pressure level, the volume range setting unit **65**A limits the upper limit of the volume range VR according to the maximum sound pressure level, for example. This is because distortion may occur in a sound exceeding the maximum sound pressure level.

[0088] The output impedance represents magnitude of an internal resistance of the external microphone **13**. In a case where the performance of the sound collection device is the output impedance, the volume range setting unit **65**A sets the volume range VR to the low volume side as the output impedance is smaller, for example. This is because a voltage drop is smaller as the output impedance is lower, and thus the sound signal AS output from the external microphone **13** is less deteriorated. Further, the volume range setting unit **65**A may limit the upper limit of the volume range VR according to the output impedance. This is because the voltage drop is larger as the output impedance is higher, and thus the noise included in the sound signal AS may be large. For example, the upper limit of the volume range VR is set lower as the output impedance is higher.

[0089] Further, the information related to the performance of the sound collection device may be a polar pattern. The polar pattern represents sensitivity (that is, directivity) to a sound collection direction of the external microphone **13**. The polar pattern includes single directivity, bidirectional directivity, and omnidirectional directivity. The volume range setting unit **65**A sets the volume range VR according to a type of the polar pattern, for example. Further, with consideration of the video data PD in addition to the polar pattern, it is possible to set the volume range VR to an optimal range such that the sound generated from the sound source such as the subject is included.

[0090] Further, the information related to the performance of the sound collection device may be a frequency range. The frequency range is a range of frequency of a sound that can be collected and reproduced by the external microphone **13**.

Second Embodiment

[0091] Next, a second embodiment will be described. In the first embodiment, the accessory information creation unit **63** creates the accessory information SI including the device information **13**A. In the second embodiment, the accessory information creation unit **63** creates the accessory information SI including sound source information, instead of the device information **13**A. The sound source information relates to the sound source such as the subject.

[0092] The configuration of the imaging apparatus **10** according to the second embodiment other than the processor **25** is the same as that of the first embodiment. In the following, the same reference numerals are assigned to the same components as those in the first embodiment, and the description thereof will be omitted as appropriate.

[0093] FIG. **8** shows an example of a functional configuration of the processor **25** according to the second embodiment. In the present embodiment, only the function of the accessory information creation unit **63** is different from that of the first embodiment. In the present embodiment, the accessory information creation unit **63** acquires the sound source information based on the video data PD output from the image processing circuit **21**, and creates the accessory information SI including the acquired sound source information. In the present embodiment, the sound source is a main subject selected from a plurality of subjects, which are included in the video data PD. For example, the accessory information creation unit **63** acquires a type of the main subject as the sound source information, for each frame constituting the video data PD, during the imaging operation by the imaging apparatus **10**.

[0094] The main subject is a subject determined to have high importance by the user or the main controller **60**, among the plurality of subjects included in the video data PD. In the present embodiment, the volume range setting unit **65**A sets the volume range VR based on the type of the main subject. For example, in a case where the type of the main subject is a type in which a generated sound is assumed to be high in volume, such as “airplane”, the volume range setting unit **65**A sets the volume range VR to the high volume side. On the other hand, in a case where the type of the main subject is a type in which a generated sound is assumed to be low in volume, such as “person”, the volume range setting unit **65**A sets the volume range VR to the low volume side.

[0095] Various types of processing can be employed as the processing in which the accessory information creation unit **63** selects the main subject from the plurality of subjects included in the video data PD. As an example, the accessory information creation unit **63** selects the main subject based on sizes of the plurality of subjects included in the video data PD. In this case, the accessory information creation unit **63** selects, as the main subject, a subject having a largest size among the plurality of subjects.

[0096] Further, the accessory information creation unit **63** can also select the main subject based on types of the plurality of subjects included in the video data PD. In this case, for example, the accessory information creation unit **63** determines the type of each subject and selects, as the main subject, a subject that matches a type set by the user using the operation unit **16**. For example, in a case where a person imaging mode is set, the accessory information creation unit **63** selects a person as the main subject from the plurality of subjects included in the video data PD.

[0097] Further, the accessory information creation unit **63** can also select the main subject based on positions of the plurality of subjects included in the video data PD within an angle of view. In this case, for example, the accessory information creation unit **63** obtains the position of each subject within the angle of view to select, as the main subject, a subject located at a center of the angle of view.

[0098] Further, the accessory information creation unit **63** can also select the main subject based on a focusing position of the imaging lens **12**. In this case, for example, the accessory information creation unit **63** acquires information related to the focusing position from the main controller **60** to select, as the main subject, a subject closest to the focusing position from the plurality of subjects included in the video data PD.

[0099] Further, the accessory information creation unit **63** can also select the main subject based on input information of the user. In this case, for example, the accessory information creation unit **63** selects, as the main subject, a subject located in the subject area selected by the user using the operation unit **16**. The accessory information creation unit **63** may select, as the main subject, a subject located at the AF area automatically detected by the imaging apparatus **10**.

[0100] Further, the accessory information creation unit **63** can also select the main subject based on visual line information of the user. In this case, for example, the imaging apparatus **10** has a function of detecting a visual line of the user who looks through the finder **14**. The accessory information creation unit **63** acquires the visual line information of the user to select, as the main subject, a subject present at a position of the visual line within the angle of view.

[0101] FIG. **9** shows an example of a relationship between the sound source information and the first sound data AS**1**F. As shown in FIG. **9**, the first sound data AS**1**F represents a change in volume (that is, change in amplitude) to time. The sound source information is associated with time information included in the first sound data AS**1**F. In the example shown in FIG. **9**, a main subject A is associated as the sound source information from t**1** to t**2**, and a main subject B is associated as the sound source information from t**2** to t**3**. The main subject A and the main subject B are of different types. For example, the main subject A is “person” of which a generated sound is low in volume, and the main subject B is “airplane” of which a generated sound is high in volume.

[0102] FIG. **10** is a flowchart showing an example of the operation of the imaging apparatus **10** according to the second embodiment. FIG. **10** shows an operation in a case where the video capturing mode is selected as the operation mode and the external microphone **13** is connected to the connecting part **11**B.

[0103] The operation of the imaging apparatus **10** according to the present embodiment is different from the above embodiment only in the timing at which the accessory information creation step (step S**14**) is executed. In the present embodiment, since the sound source information changes over time, the accessory information creation step is executed after the acquisition step. That is, in the present embodiment, steps S**11**, S**12**, and S**14** are repeatedly executed until the end instruction is determined to be issued in step S**13**. In a case where the end instruction is determined to be issued (YES in step S**13**), the sound data file creation step (step S**15**) and the editing step (step S**16**) are executed.

Third Embodiment

[0104] Next, a third embodiment will be described. In the second embodiment, the sound source of the sound source information, which is acquired by the accessory information creation unit **63**, is the subject. In the third embodiment, the sound source information acquired by the accessory information creation unit **63** is a type of a drive sound accompanied by drive of the imaging apparatus **10**. For example, the type of the drive sound is a drive sound of the focus lens **31**, a drive sound of the stop **32**, and the like. In a case where the imaging apparatus **10** comprises a heat radiation fan, the type of the drive sound includes a drive sound of the heat radiation fan.

[0105] The configuration of the imaging apparatus **10** according to the third embodiment other than the processor **25** is the same as that of the first embodiment. In the following, the same reference numerals are assigned to the same components as those in the first embodiment, and the description thereof will be omitted as appropriate.

[0106] FIG. **11** shows an example of a functional configuration of the processor **25** according to the third embodiment. In the present embodiment, only the function of the accessory information creation unit **63** is different from that of the first embodiment. In the present embodiment, the accessory information creation unit **63** acquires the sound source information based on the first sound data AS**1**F output from the data format conversion unit **62** to create the accessory information SI including the acquired sound source information. The accessory information creation unit **63** acquires, as sound source information, the type of the drive sound accompanied by the drive of the imaging apparatus **10** during the imaging operation of the imaging apparatus **10**. For example, the accessory information creation unit **63** analyzes the first sound data AS**1**F to determine the type of the drive sound.

[0107] The accessory information creation unit **63** is not limited to the first sound data AS**1**F, and may determine the type of the drive sound based on sound data such as the first sound data AS**1** and the modulation sound data ASH and ASL. Further, the accessory information creation unit **63** may acquire, from the main controller **60**, information representing a type of a device in operation to determine the type of the drive sound.

[0108] FIG. **12** shows an example of a relationship between the sound source information and the first sound data AS**1**F according to the third embodiment. The sound source information is associated with time information included in the first sound data AS**1**F. In the example shown in FIG. **9**, a drive sound A is associated as the sound source information from t**1** to t**2**, and a drive sound B is associated as the sound source information from t**2** to t**3**. The drive sound A and the drive sound B are of different types. For example, the drive sound A is “drive sound of stop **32**” in which a generated sound is low in volume, and the drive sound B is “drive sound of heat radiation fan” in which a generated sound is high in volume.

[0109] The operation of the imaging apparatus **10** according to the third embodiment is the same as that of the second embodiment except that the type of the drive sound is acquired as the sound source information in the accessory information creation step.

Modification Example

[0110] Next, various modification examples will be described. In each of the above embodiments, the editing unit **65** edits the first sound data AS**1**F based on the accessory information SI to create the second sound data AS**2**. The editing unit **65** may create deterioration information DI, which relates to the sound of the second sound data AS**2** deteriorated by the editing, and may create a sound data file **68** including the second sound data AS**2** and the deterioration information DI as shown in FIG. **13**. The editing unit **65** records the created sound data file **68** in the storage device **26**. The sound data file **68** corresponds to “second file” according to the technique of the present disclosure.

[0111] The deterioration means that at least a part of the sound of the sound source included in the first sound data AS**1**F is not included in the volume range of the second sound data AS**2**. The deterioration information is sound information that is lost in a case where the second sound data AS**2** having a smaller amount of information than the first sound data AS**1**F is created from the first sound data AS**1**F. The sound data file **68** may further include the accessory information SI.

[0112] The technique of the present disclosure is not limited to the digital camera and can also be employed for electronic devices such as a smartphone and a tablet terminal having an imaging function.

[0113] In each of the above embodiments, various processors shown below can be used as the hardware structure of the control unit using the processor **25** as an example. The above various processors include not only a CPU which is a general-purpose processor that functions by executing software (programs) but also a processor that has a changeable circuit configuration after manufacturing, such as an FPGA. The FPGA includes a dedicated electrical circuit that is a processor which has a dedicated circuit configuration designed to execute specific processing, such as PLD or ASIC, and the like.

[0114] The control unit may be configured by one of these various processors or a combination of two or more of the processors of the same type or different types (for example, a combination of a plurality of FPGAs or a combination of a CPU and an FPGA). Alternatively, a plurality of control units may be configured with one processor.

[0115] A plurality of examples in which a plurality of control units are configured as one processor can be considered. As a first example, there is an aspect in which one or more CPUs and software

are combined to configure one processor and the processor functions as a plurality of control units, as represented by a computer such as a client and a server. As a second example, there is an aspect in which a processor that implements the functions of the entire system, which includes a plurality of control units, with one IC chip is used, as represented by system on chip (SOC). In this manner, the control unit can be configured by using one or more of the above various processors as the hardware structure.

[0116] Furthermore, more specifically, it is possible to use an electrical circuit in which circuit elements such as semiconductor elements are combined, as the hardware structure of these various processors.

[0117] Contents described and illustrated above are for detailed description of a portion according to the technique of the present disclosure and are only an example of the technique of the present disclosure. For example, the descriptions regarding the configurations, the functions, the actions, and the effects are descriptions regarding an example of the configurations, the functions, the actions, and the effects of the part according to the technique of the present disclosure. Accordingly, in the contents described and the contents shown hereinabove, it is needless to say that removal of an unnecessary part, or addition or replacement of a new element may be employed within a range not departing from the gist of the present technique of the present disclosure. Furthermore, to avoid confusion and to facilitate understanding of a part according to the technique of the present disclosure, description relating to common technical knowledge and the like that does not require particular description to enable implementation of the technique of the present disclosure is omitted from the content of the above description and from the content of the drawings.

[0118] In a case where all of documents, patent applications, and technical standard described in the specification are built into the specification as references, to the same degree as a case where the incorporation of each of documents, patent applications, and technical standard as references is specifically and individually noted.

[0119] The following technique can be understood from the above description.

Supplementary Note 1

[0120] A creation method comprising: [0121] an acquisition step of acquiring first sound data of a floating point format, based on a sound that is generated from a sound source and is collected by a sound collection device; and [0122] an accessory information creation step of creating accessory information that is attached to the first sound data and includes device information, which relates to the sound collection device, or sound source information, which relates to the sound source.

Supplementary Note 2

[0123] The creation method according to Supplementary Note 1, [0124] wherein the first sound data is used to create second sound data having the number of bits, which is smaller than the number of bits of the first sound data.

Supplementary Note 3

[0125] The creation method according to Supplementary Note 1 or 2, [0126] wherein the accessory information includes the device information.

Supplementary Note 4

[0127] The creation method according to Supplementary Note 3, [0128] wherein the device information relates to gain processing used for the sound collected by the sound collection device or relates to performance of the sound collection device.

Supplementary Note 5

[0129] The creation method according to Supplementary Note 2, [0130] wherein the accessory information includes the sound source information.

Supplementary Note 6

[0131] The creation method according to Supplementary Note 5, [0132] wherein the sound source information is associated with time information included in the first sound data.

Supplementary Note 7
[0133] The creation method according to Supplementary Note 5 or 6, further comprising: [0134] an imaging step of creating, by an imaging apparatus, video data corresponding to the first sound data.
Supplementary Note 8
[0135] The creation method according to Supplementary Note 7, [0136] wherein the sound source is a subject included in the video data.
Supplementary Note 9
[0137] The creation method according to Supplementary Note 7, [0138] wherein the sound source is a main subject selected from a plurality of subjects included in the video data.
Supplementary Note 10
[0139] The creation method according to Supplementary Note 7, [0140] wherein the sound source information is a type of a drive sound accompanied by drive of the imaging apparatus.
Supplementary Note 11
[0141] The creation method according to any one of Supplementary Notes 1 to 10, further comprising: [0142] a first file creation step of creating a first file including the first sound data and the accessory information.
Supplementary Note 12
[0143] The creation method according to Supplementary Note 11, further comprising: [0144] an editing step of editing the first sound data based on the accessory information to create second sound data having the number of bits, which is smaller than the number of bits of the first sound data.
Supplementary Note 13
[0145] The creation method according to Supplementary Note 12, [0146] wherein, in the editing step, deterioration information that relates to a sound of the second sound data deteriorated by the editing is created, and a second file including the second sound data and the deterioration information is created.
Supplementary Note 14
[0147] The creation method according to Supplementary Note 13, [0148] wherein the second file includes the accessory information.

Claims

1. A creation method comprising: an acquisition step of acquiring first sound data of a floating point format, based on a sound that is generated from a sound source and is collected by a sound collection device; and an accessory information creation step of creating accessory information that is attached to the first sound data and includes device information, which relates to the sound collection device, or sound source information, which relates to the sound source.
2. The creation method according to claim 1, wherein the first sound data is used to create second sound data having the number of bits, which is smaller than the number of bits of the first sound data.
3. The creation method according to claim 2, wherein the accessory information includes the device information.
4. The creation method according to claim 3, wherein the device information relates to gain processing used for the sound collected by the sound collection device or relates to performance of the sound collection device.
5. The creation method according to claim 2, wherein the accessory information includes the sound source information.
6. The creation method according to claim 5, wherein the sound source information is associated with time information included in the first sound data.
7. The creation method according to claim 5, further comprising: an imaging step of creating, by an

imaging apparatus, video data corresponding to the first sound data.
8. The creation method according to claim 7, wherein the sound source is a subject included in the video data.
9. The creation method according to claim 7, wherein the sound source is a main subject selected from a plurality of subjects included in the video data.
10. The creation method according to claim 7, wherein the sound source information is a type of a drive sound accompanied by drive of the imaging apparatus.
11. The creation method according to claim 1, further comprising: a first file creation step of creating a first file including the first sound data and the accessory information.
12. The creation method according to claim 11, further comprising: an editing step of editing the first sound data based on the accessory information to create second sound data having the number of bits, which is smaller than the number of bits of the first sound data.
13. The creation method according to claim 12, wherein, in the editing step, deterioration information that relates to a sound of the second sound data deteriorated by the editing is created, and a second file including the second sound data and the deterioration information is created.
14. The creation method according to claim 13, wherein the second file includes the accessory information.
15. A creation apparatus comprising: a processor, wherein the processor is configured to execute an acquisition step of acquiring first sound data of a floating point format, based on a sound that is generated from a sound source and is collected by a sound collection device, and an accessory information creation step of creating accessory information that is attached to the first sound data and includes device information, which relates to the sound collection device, or sound source information, which relates to the sound source.