

DESCRIPTION

TRANSMISSION DEVICE, TRANSMISSION METHOD, RECEPTION DEVICE, AND
RECEPTION METHOD

TECHNICAL FIELD

[0001]

The present technology relates to a transmission device, a transmission method, a reception device, and a reception method, and particularly relates to a transmission device, a transmission method, a reception device, and a reception method that can enhance the accessibility for people with visual disabilities.

BACKGROUND ART

[0002]

In the field of digital broadcasting, the accessibility for people with visual disabilities is required (for example, see Patent Document 1).

[0003]

Especially, the 21st Century Communications and Video Accessibility Act of 2010 (CVAA), or the so-called American accessibility act was enacted in the United States. Federal Communications Commission (FCC) has published various regulations on the accessibility based on the act.

CITATION LIST

PATENT DOCUMENT

[0004]

Patent Document 1: Japanese Patent Application Laid-Open No.
2009-204711

SUMMARY OF THE INVENTION

PROBLEMS TO BE SOLVED BY THE INVENTION

[0005]

By the way, when a user interface (UI), for example, for program information is presented to people with visual disabilities, the text information of the program information is read aloud with a Text To Speech (TTS) engine. This is a typical way to enhance the

accessibility.

[0006]

However, when the TTS engine is used, the text information is not necessarily read aloud as the producer of the program information intends. It is uncertain whether people with visual disabilities can receive the information similar to the information that able-bodied people receive. In light of the foregoing, a technology with which utterance as the producer intends is surely performed so that people with visual disabilities can receive the information similar to the information that able-bodied people receive.

[0007]

In light of the foregoing, embodiments of the present invention enable the enhancement in the accessibility for people with visual disabilities by surely performing utterance as the producer intends.

SOLUTIONS TO PROBLEMS

[0008]

According to a first aspect of the present invention, there is provided a transmission device, comprising: one or more processors configured to: generate metadata about speech phonetic information of display information, wherein text-to-speech utterance of the speech phonetic information is generated based on the metadata; generate a service guide that includes the metadata; and transmit the service guide to a reception device that displays the display information, wherein the metadata includes an embedded speech information object which is described using a markup language associated with the embedded speech information object, wherein the markup language is indicated by the metadata, and/or the metadata includes an address information to acquire a speech information object which is described using a markup language associated with the speech information object referenced by the address information, wherein the markup language is indicated by the metadata.

[0009]

The metadata can include information about utterance of a character string of which pronunciation is not uniquely determined, or a character string difficult to pronounce.

[0010]

The display information can include information or an icon of a content.

[0011]

A content acquisition unit that acquires the content can be further included, and the transmission unit can transmit the electronic program information together with the content using a digital broadcast signal.

[0012]

The electronic program information can comply with Electronic Service Guide (ESG) prescribed in Open Mobile Alliance - Mobile Broadcast Services Enabler Suite (OMA-BCAST), the metadata can be described in a Speech Synthesis Markup Language (SSML) format, and a predetermined fragment included in the ESG can include address information indicating an address at which a file of the metadata described in the SSML format is acquired, or content of the metadata described in the SSML format.

[0013]

The transmission device can be an independent device, or can be an internal block included in a device.

[0014]

According to a second aspect of the present invention, there is provided a transmission method of a transmission device, the method comprising: generating, by one or more processors of the transmission device, metadata about speech phonetic information of display information, wherein text-to-speech utterance of the speech phonetic information is generated based on the metadata; generating, by the one or more processors of the transmission device, a service guide including the metadata; and transmitting the service guide to a reception device capable of displaying the display information, wherein the metadata includes an embedded speech information object which is described using a markup language associated with the embedded speech information object, wherein the markup language is indicated by the metadata, and/or the metadata includes an address information to acquire a speech information object which is described using a markup language associated with the speech information object referenced by the address information,

wherein the markup language is indicated by the metadata.

[0015]

In the transmission device and transmission method according to the above aspects of the present invention, the metadata about the text-to-speech utterance that the producer intends is generated for display information, and the electronic program information including the metadata is generated. Then, the electronic program information is transmitted to a reception device capable of displaying the display information.

[0016]

According to a third aspect of the present invention, there is provided a reception device, comprising: one or more processors configured to: receive a service guide from a transmission device, wherein the service guide includes metadata about speech phonetic information of display information, and wherein a text-to-speech utterance of the speech phonetic information is generated based on the metadata; acquire the metadata included in the service guide; and read the display information aloud based on the metadata, wherein the metadata includes an embedded speech information object which is described using a markup language associated with the embedded speech information object, wherein the markup language is indicated by the metadata, and/or the metadata includes an address information to acquire a speech information object which is described using a markup language associated with the speech information object referenced by the address information, wherein the markup language is indicated by the metadata.

[0017]

The metadata can include information about utterance of a character string of which pronunciation is not uniquely determined, or a character string difficult to pronounce.

[0018]

The display information can include information or an icon of a content.

[0019]

The reception unit can receive the electronic program information transmitted together with the content as a digital broadcast signal.

[0020]

The electronic program information can comply with ESG prescribed in OMA-BCAST, the metadata can be described in an SSML format, a predetermined fragment included in the ESG can include address information indicating an address at which a file of the metadata described in the SSML format is acquired, or content of the metadata described in the SSML format, and the metadata acquisition unit can acquire the file of the metadata in accordance with the address information, or acquires the metadata from the fragment.

[0021]

The reception device can be an independent device, or can be an internal block included in a device.

[0022]

According to a fourth aspect of the present invention, there is provided a reception method of a reception device, the method comprising: receiving a service guide from a transmission device, wherein the service guide includes metadata about speech phonetic information of display information, and wherein text-to-speech utterance of the speech phonetic information is generated based on the metadata; acquiring the metadata included in the service guide; and reading the display information aloud based on the metadata, wherein the metadata includes an embedded speech information object which is described using a markup language associated with the embedded speech information object, wherein the markup language is indicated by the metadata, and/or the metadata includes an address information to acquire a speech information object which is described using a markup language associated with the speech information object referenced by the address information, wherein the markup language is indicated by the metadata.

[0023]

In the reception device and reception method according to the above aspects of the present invention, the electronic program information, which includes the metadata about the text-to-speech utterance of the display information as the producer intends and is transmitted from the transmission device, is received. Then,

the metadata included in the electronic program information is acquired, and the display information is read aloud on the basis of the metadata.

EFFECTS OF THE INVENTION

[0024]

According to the above aspects of the present invention, the accessibility for people with visual disabilities can be enhanced.

[0025]

Note that the effects of embodiments of the present invention are not necessarily limited to the effect described above, and can be any one of the effects described herein.

BRIEF DESCRIPTION OF DRAWINGS

[0026]

Fig. 1 is a diagram of an example in which program information and a title are read aloud.

Fig. 2 is a diagram of an example in which an icon is read aloud.

Fig. 3 is an explanatory diagram of an example in which text information is read aloud with an existing TTS engine.

Fig. 4 is an explanatory diagram of an example in which text information is read aloud with an existing TTS engine.

Fig. 5 is an explanatory diagram of an example in which text information is read aloud with a TTS engine to which the present technology is applied.

Fig. 6 is an explanatory diagram of an example in which text information is read aloud with a TTS engine to which the present technology is applied.

Fig. 7 is a diagram of an exemplary configuration of a broadcasting system to which the present technology is applied.

Fig. 8 is a diagram of an exemplary configuration of a transmission device to which the present technology is

CLAIMS

1. A transmission device, comprising:

one or more processors configured to:

generate metadata about speech phonetic information of display information, wherein text-to-speech utterance of the speech phonetic information is generated based on the metadata;

generate a service guide that includes the metadata; and

transmit the service guide to a reception device that displays the display information, wherein

the metadata includes an embedded speech information object which is described using a markup language associated with the embedded speech information object, wherein the markup language is indicated by the metadata, and/or

the metadata includes an address information to acquire a speech information object which is described using a markup language associated with the speech information object referenced by the address information, wherein the markup language is indicated by the metadata.

2. The transmission device according to claim 1, wherein the metadata includes information about utterance of one of a character string of which pronunciation is not uniquely determined, or a character string that includes a proper noun.

3. The transmission device according to claim 2, wherein the display information includes one of information of content or an icon of the content.

4. The transmission device according to claim 3, wherein the one or more processors are further configured to:

acquire the content, and

transmit the service guide together with the content based on a digital broadcast signal.

5. The transmission device according to claim 1, wherein the service guide complies with Electronic Service Guide (ESG) prescribed in Open Mobile Alliance - Mobile Broadcast Services Enabler Suite (OMA-BCAST),

the embedded speech information object and/or the speech information object is described in a Speech Synthesis Markup Language (SSML) format, and

a fragment included in the ESG includes the address information to acquire the speech information object described in the SSML format, and/or the embedded speech information object described in the SSML format.

6. A transmission method of a transmission device, the method comprising:

generating, by one or more processors of the transmission device, metadata about speech phonetic information of display information, wherein text-to-speech utterance of the speech phonetic information is generated based on the metadata;

generating, by the one or more processors of the transmission device, a service guide including the metadata; and

transmitting the service guide to a reception device capable of displaying the display information, wherein

the metadata includes an embedded speech information object which is described using a markup language associated with the embedded speech information object, wherein the markup language is indicated by the metadata, and/or

the metadata includes an address information to acquire a speech information object which is described using a markup language associated with the speech information object referenced by the

address information, wherein the markup language is indicated by the metadata.

7. A reception device, comprising:
one or more processors configured to:

receive a service guide from a transmission device, wherein the service guide includes metadata about speech phonetic information of display information, and wherein a text-to-speech utterance of the speech phonetic information is generated based on the metadata;

acquire the metadata included in the service guide; and
read the display information aloud based on the metadata,
wherein

the metadata includes an embedded speech information object which is described using a markup language associated with the embedded speech information object, wherein the markup language is indicated by the metadata, and/or

the metadata includes an address information to acquire a speech information object which is described using a markup language associated with the speech information object referenced by the address information, wherein the markup language is indicated by the metadata.

8. The reception device according to claim 7, wherein the metadata includes information about utterance of one of a character string of which pronunciation is not uniquely determined, or a character string that includes a proper noun.

9. The reception device according to claim 8, wherein the display information includes one of information of content or an icon of the content.

10. The reception device according to claim 9, wherein one or more processors are further configured to receive the service guide transmitted together with the content as a digital broadcast signal.

11. The reception device according to claim 7, wherein
the service guide complies with Electronic Service Guide (ESG) prescribed in Open Mobile Alliance - Mobile Broadcast Services Enabler Suite (OMA-BCAST),

the metadata is described in a Speech Synthesis Markup Language (SSML) format,

a fragment included in the ESG includes the address information to acquire the speech information object described in the SSML format, and /or the embedded speech information object described in the SSML format, and

the one or more processors are further configured to acquire the speech information object, or acquire the embedded speech information object from the fragment.

12. A reception method of a reception device, the method comprising:

receiving a service guide from a transmission device, wherein the service guide includes metadata about speech phonetic information of display information, and wherein text-to-speech utterance of the speech phonetic information is generated based on the metadata;

acquiring the metadata included in the service guide; and
reading the display information aloud based on the metadata, wherein

the metadata includes an embedded speech information object which is described using a markup language associated with the embedded speech information object, wherein the markup language is indicated by the metadata, and/or

the metadata includes an address information to acquire a speech information object which is described using a markup language associated with the speech information object referenced by the address information, wherein the markup language is indicated by the metadata.