

DESCRIPTION

INFORMATION PROCESSING APPARATUS, SPEECH RECOGNITION SYSTEM, AND INFORMATION PROCESSING METHOD

TECHNICAL FIELD

[0001]

The present disclosure relates to an information processing apparatus for use in a speech recognition system.

BACKGROUND ART

[0002]

What is called virtual personal assistance (VPA) services for operating devices by speech have started to spread. Patent Literature 1 discloses an information processing apparatus which predicts utterance of a user when the apparatus detected an activation word.

Citation List

Patent Literature

[0003]

PTL 1: Japanese Unexamined Patent Application Publication No. 2017-138476

SUMMARY OF THE INVENTION

TECHNICAL PROBLEM

[0004]

When main speech recognition is performed in a cloud server, for example, an information processing apparatus starts transmission of a speech signal to a speech recognition server by recognizing an activation word.

[0005]

Embodiments of the present disclosure provide an information processing apparatus capable of transmitting speech signals selectively to a plurality of speech recognition servers.

SOLUTION TO PROBLEM

[0006]

According to a first aspect of the present disclosure, there is provided an

information processing apparatus including: a speech obtainer which obtains speech of a user; a first controller which, when the first controller recognizes that the speech obtained by the speech obtainer is a first activation word, outputs a speech signal corresponding to the first activation word; and a second controller which starts a first speech transmission process of transmitting the speech signal of the speech obtained by the speech obtainer to a first speech recognition server when the second controller recognizes that the speech signal output by the first controller indicates the first activation word, wherein , when the first controller recognizes in the first speech transmission process that the speech obtained by the speech obtainer is the second activation word for causing the second controller to start a second speech transmission process, the first controller determines whether to cancel the first speech transmission process which is currently being performed and output a speech signal corresponding to a second activation word to the second controller based on a predetermined priority level, and the second speech transmission process is a process for transmitting the speech signal of the speech obtained by the speech obtainer to a second speech recognition server different from the first speech recognition server.

[0006a]

According to a second aspect of the present disclosure, there is provided a speech recognition system, including: the above information processing apparatus; the first speech recognition server; and the second speech recognition server.

[0006b]

According to a third aspect of the present disclosure, there is provided an information processing method which is executed by a computer, the information processing method including: obtaining speech of a user; when the speech obtained is recognized as a first activation word, outputting a speech signal corresponding to the first activation word; when the speech signal output is recognized as the first activation word, starting a first speech transmission process of transmitting a speech signal of the speech obtained to a first speech recognition server; and when the speech obtained in the first speech transmission process is the second activation word for starting a second speech transmission process, determining whether to cancel the first speech transmission process currently being performed and output a speech signal corresponding to the second activation word based on predetermined priority levels, wherein the second speech transmission process is a process of transmitting the speech signal of the speech obtained to a second speech recognition server different from the first

speech recognition server

[0007]

For example, these general and specific aspects of the present disclosure may be implemented using, a system, a method, an integrated circuit, a computer program, or a computer-readable recording medium such as a CD-ROM, or any combination of, systems, methods, integrated circuits, computer programs, or computer-readable recording media.

ADVANTAGEOUS EFFECT OF INVENTION

[0008]

The information processing apparatus according to embodiments of the present disclosure is capable of transmitting speech signals selectively to a plurality of speech recognition servers.

BRIEF DESCRIPTION OF DRAWINGS

[0009]

FIG. 1 is a diagram for illustrating functions of a smart speaker.

FIG. 2 is a diagram for illustrating how household electrical appliances are controlled using the smart speaker.

FIG. 3 is a diagram indicating the relationship between services and activation words.

FIG. 4 is a diagram for illustrating a case in which a user calls a service using a smartphone.

FIG. 5 is a block diagram indicating a configuration of a speech recognition system according to Embodiment 1.

FIG. 6 is a flowchart indicating operations performed by the speech recognition system according to Embodiment 1.

FIG. 7 is a diagram indicating the relationship between services and activation words according to Embodiment 1.

FIG. 8 is a flowchart indicating operations performed by the speech recognition system according to Embodiment 2.

FIG. 9 is a diagram indicating the relationship between services and activation words according to Embodiment 2.

FIG. 10 is a block diagram indicating a configuration of a speech recognition system according to Embodiment 3.

CLAIMS

1. An information processing apparatus, comprising:

a speech obtainer which obtains speech of a user;

a first controller which, when the first controller recognizes that the speech obtained by the speech obtainer is a first activation word, outputs a speech signal corresponding to the first activation word; and

a second controller which starts a first speech transmission process of transmitting the speech signal of the speech obtained by the speech obtainer to a first speech recognition server when the second controller recognizes that the speech signal output by the first controller indicates the first activation word,

wherein, when the first controller recognizes in the first speech transmission process that the speech obtained by the speech obtainer is the second activation word for causing the second controller to start a second speech transmission process, the first controller determines whether to cancel the first speech transmission process which is currently being performed and output a speech signal corresponding to a second activation word to the second controller based on a predetermined priority level, and

the second speech transmission process is a process for transmitting the speech signal of the speech obtained by the speech obtainer to a second speech recognition server different from the first speech recognition server.

2. The information processing apparatus according to claim 1,

wherein the first controller determines to output the speech signal corresponding to the second activation word to the second controller when a priority level of the first speech recognition server is lower than a priority level of a second speech recognition server.

3. The information processing apparatus according to claim 2,

wherein the first controller determines not to output the speech signal corresponding to the second activation word to the second controller when the priority level of the first speech recognition server is higher than the priority level of the second speech recognition server.

4. The information processing apparatus according to claim 1,

wherein the first controller determines to output the speech signal corresponding to the second activation word to the second controller when a priority level of a first service that is provided as a result of the first speech transmission process is lower than a priority level of a second service that is provided as a result of the second speech transmission process.

5. The information processing apparatus according to claim 4,

wherein the first controller determines not to output the speech signal corresponding to the second activation word to the second controller when the priority level of the first service is higher than the priority level of the second service.

6. A speech recognition system, comprising:

the information processing apparatus according to any one of claim 1 to claim 5;

the first speech recognition server; and

the second speech recognition server.

7. An information processing method which is executed by a computer,

the information processing method comprising:

obtaining speech of a user;

when the speech obtained is recognized as a first activation word, outputting a speech signal corresponding to the first activation word;

when the speech signal output is recognized as the first activation word, starting a first speech transmission process of transmitting a speech signal of the speech obtained to a first speech recognition server; and

when the speech obtained in the first speech transmission process is the second activation word for starting a second speech transmission process, determining whether to cancel the first speech transmission process currently being performed and output a speech signal corresponding to the second activation word based on predetermined priority levels,

wherein the second speech transmission process is a process of transmitting the speech signal of the speech obtained to a second speech recognition server different from the first speech recognition server.