Technical Disclosure Commons

Defensive Publications Series

September 05, 2018

Alternate Phrase Suggestions from Voice Input

Joshua Morton

Christopher Halaska

Follow this and additional works at: https://www.tdcommons.org/dpubs_series

Alternate Phrase Suggestions from Voice Input

ABSTRACT

Techniques to improve the process of correcting text transcription of a voice input are described. Words or phrases of the voice input that are not recognized with sufficient confidence are identified and displayed in a manner that distinguishes such phrases from the rest of the transcribed text. Alternatives for the low confidence terms are determined and displayed as suggestions. The user can view and select one or more of the alternative suggestions to replace the low confidence terms.

KEYWORDS

- Speech transcription

- Voice input

- Voice UI

- Speech API

- Phrase correction

BACKGROUND

A voice transcription system uses techniques, such as speech recognition, to recognize and transcribe voice input such as user speech into text. However, the transcribed text may have errors if certain words or phrases of the voice input are not recognized. The user may need to manually modify the transcribed text to add, replace, or delete such phrases in the transcribed text. In some scenarios, correcting the errors may require the user to repeat the voice input. The burden on the user to make corrections manually or to repeat the input slows down the process of voice transcription and degrades user experience.

DESCRIPTION

Techniques of this disclosure identify words or phrases in voice input that are recognized with low confidence (low confidence terms). When the transcribed text is displayed on a user interface, the low confidence terms are displayed in a manner that distinguishes such terms from the rest of the transcribed text. For example, the low confidence terms are highlighted with a dotted underline in the transcribed text.

The visual distinction indicates to a user that alternative suggestions are available for the low confidence terms. Such alternative suggestions can be selected by the user to replace the low confidence terms. For instance, when the user interacts with any of the low confidence terms, e.g., a user taps or clicks on a low confidence term, alternative suggestions for the low confidence term are displayed. Upon selection of an alternative suggestion from the displayed alternative suggestions, the transcribed text is modified to incorporate the suggested alternative.

The techniques disclosed herein are implemented by use of algorithms that identify the low confidence terms and determine alternative suggestions. For example, with user permission, a machine learning model trained based on prior user selections of alternative suggestions of the low confidence terms may be utilized.

The user interface displays the low confidence terms that have alternative suggestions as visually distinctive from rest of the transcribed text. Upon selection of any of the low confidence terms by the user, a menu is shown on the user interface, anchored just below the low confidence term. The menu includes alternative suggestions for the selected low confidence term. Alternatively, drop-down menu with alternative suggestions for the entire transcribed text can also be displayed to enable the user to select any of the suggested alternatives.
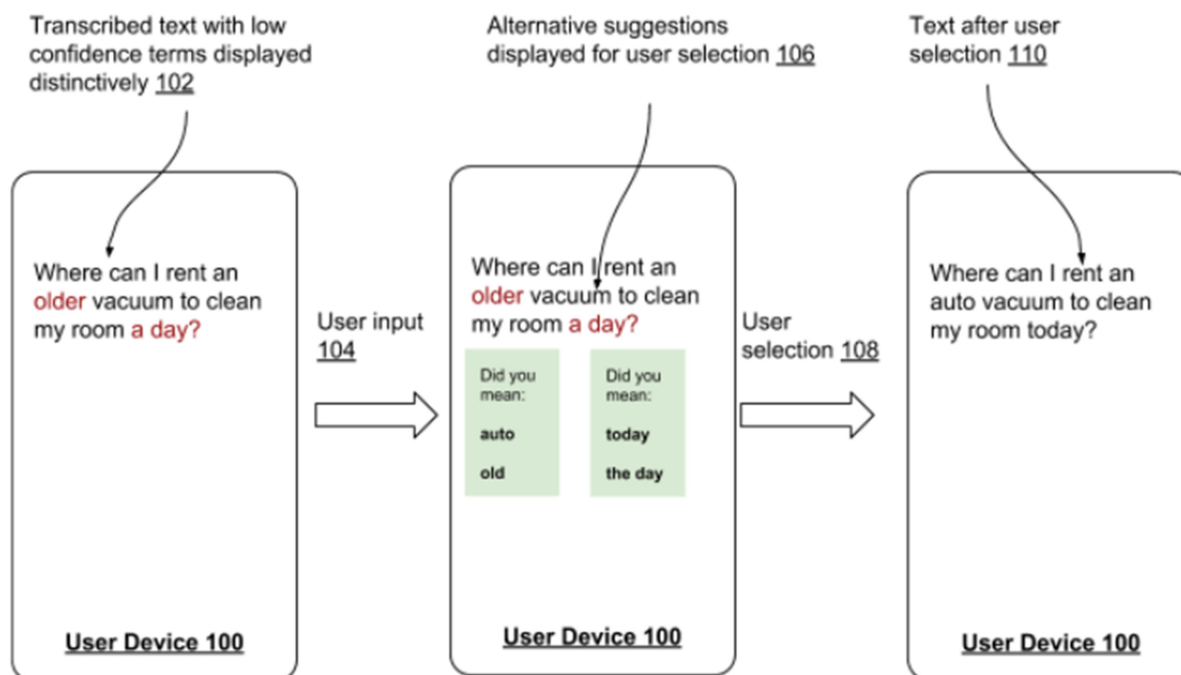
Fig. 1: Enabling correction of transcribed text

Fig. 1 illustrates an example of a user device (100) with a user interface per the techniques described herein. The user device performs speech recognition to transcribe a voice input from a user into text "Where can I rent an older vacuum to clean my room a day?" and displays the transcribed text. For example, the speech recognition may be performed locally, or by calling a server-based speech recognition service, e.g., via an API.

In this example, the word "older" and the phrase "a day" of the transcribed text are associated with low confidence. Alternative suggestions for these words are determined. When the transcribed text is displayed, the low confidence terms "older" and "a day" are displayed visually distinctive from rest of the transcribed text, e.g., in a different color than the rest of the transcribed text (102). The visually distinctive low confidence terms "older" and "a day" indicate to the user that these terms have alternative suggestions available.

Upon the display of the transcribed text, user input (104) is received. For example, the user may provide such input by selecting the low confidence terms "older" and/or "a day" to see the alternative suggestions for the low confidence terms "older" and/or "a day". The selection may be performed using tap, touch, or other input mechanism.

Responsive to the user input the alternative suggestions (106) for the low confidence terms "older" and/or "a day" are displayed for user selection. As illustrated in Fig. 1, the words "auto" and "old" are suggested for the word "older," and the word "today" and the phrase "the day" are suggested for the phrase "a day." Upon user selection of the alternative suggestion "auto" for the word "older" and the alternative suggestion "today" for the phrase "a day" the corrected text "Where can I rent an auto vacuum to clean my room today ?" (110) is displayed.

In this manner, the techniques disclosed herein alleviate the need for the user to manually correct the transcribed text or repeat speech input.

CONCLUSION

Techniques to improve the process of correcting text transcription of a voice input are described. Words or phrases of the voice input that are not recognized with sufficient confidence are identified and displayed in a manner that distinguishes such phrases from the rest of the transcribed text. Suggested alternatives for the low confidence terms are determined and displayed. The user is enabled to view and select one or more of the alternative suggestions to replace the low confidence terms.