Webology (ISSN: 1735-188X)
Volume 17, Number 4, 2020

Arrangement And Technique For Exploring An Appeal To Deliver A Customized Voice Recognition Service Towards The User: Snow Boy Smart Speaker

Laxmi Shankar Awasthi[1, 2, 3], Karuna Shankar Awasthi[1] and Abhay S. Pandey[4*]

[1] Associate Professor, Department of Computer Science, Lucknow Public College of Professional Studies, Lucknow, India.

[2] Head, Artificial Intelligence Research Laboratory, Lucknow Public College of Professional Studies, Lucknow, India.

[3] Dean (Academics), Lucknow Public College of Professional Studies, Lucknow, India.

[4] Assistant Professor (Physics), Department of Science, Lucknow Public College of Professional Studies, Lucknow, India.

[*]Corresponding Author: abhaypandey.liquidcrystal@gmail.com

Abstract: The present disclosure envisages a method and a system for providing a customized service to user. The method comprises the step of receiving a request of each of a plurality of users from a user terminal, wherein the request corresponds to at least an instruction generated by the user or a reminder pertaining to said instruction of the user. The method further comprises the step of analyzing the request to identify a type of customized service in the instruction, wherein the type of customized service corresponds to updating a pre-stored set of data. The method also comprises the step of generating a suggestive report depicting an actionable intelligence based on said analysis, wherein the suggestive report comprises an output pertaining to the customized service. Further, the present disclosure relates to the field of personalized devices. More specifically, the invention belongs to the field of trainable personalized devices which involve customization of instructions according to different requirements of an individual.

Keywords: Artificial Intelligence; Computer Science and Engineering; Automatic Speech Recognition (ASR); Automatic Speaker Verification (ASV); Snowboy Smart Speaker; Smartphone Application; Smart Speakers; Amazon Alexa; and Google Home.

Introduction: Automatic Speech Recognition (ASR) and Automatic Speaker Verification (ASV) systems are often used for accepting voice commands to learn to associate the acoustic input with the corresponding sounds ('phone labels') of the language. However, conventional ASR systems continued to show-difficulty in processing speech in multi-speaker and high noise environments despite the availability of substantially increased processing power. To achieve high-performance speech recognition in multi-speaker and high noise environments, a speech recognition system that uses the same cues for recognition and noise robustness that human beings do is needed. Such a system should be based on detailed neurobiological and psychoacoustic knowledge of human auditory function, accomplishing noise robustness via auditory stream separation and by using noise-robust phonetic cues. The simplified block diagram of Voice Assistant: Snowboy Smart Speaker is shown in **Figure 1**.

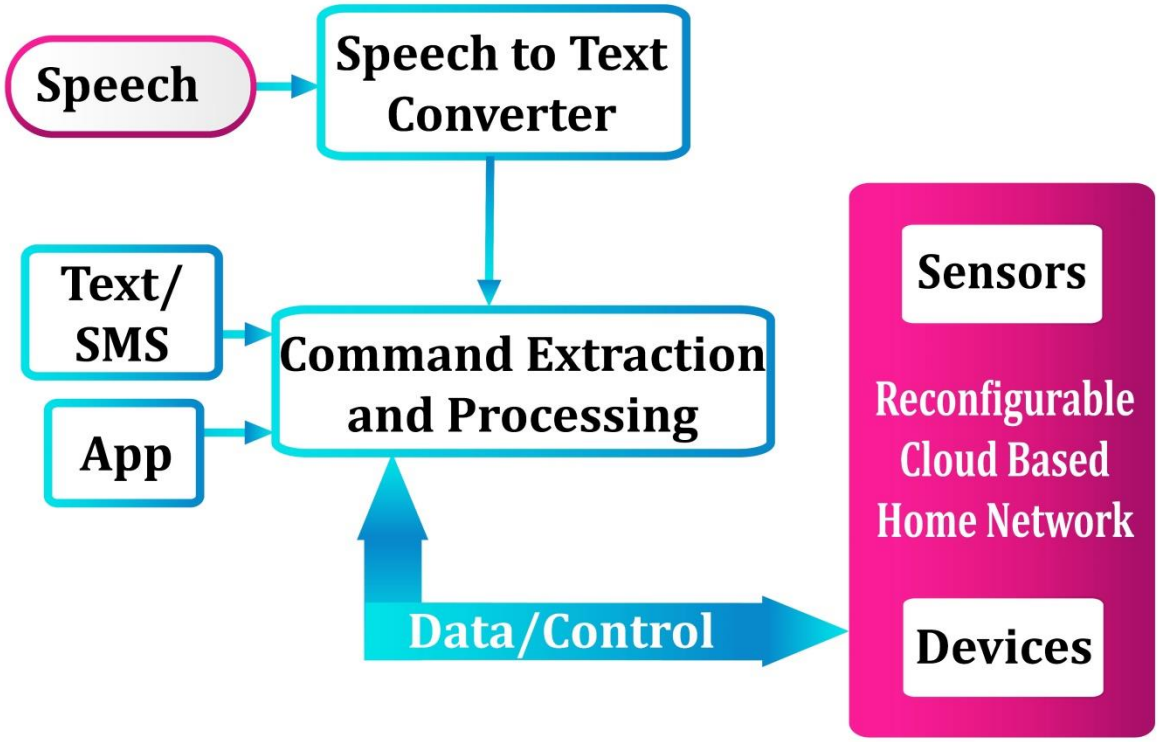

Figure 1: The simplified block diagram of Voice Assistant: Snowboy Smart Speaker.

The human approach to noise robustness is based on a high-resolution spectral analysis, followed by intelligent groupings of fine-grained sound features that can be ascribed to a common source. By contrast, conventional speech recognition systems achieve the groupings of fine-grained sound features by indiscriminately blurring them together in few millisecond frames. However, such an approach works passably in quiet environments, but is the major limiting factor preventing

conventional recognizers from achieving noise robustness, once the signal features have been blurred in with the other sounds, they can never be recovered. Other aspects and features of the present invention will become apparent to those ordinarily skilled in the art upon review of the following description of specific embodiments of the invention in conjunction with the accompanying figures.

Explanation and Description of the Present Research Work: Embodiments are provided so as to thoroughly and fully convey the scope of the present disclosure to the person skilled in the art. Numerous details are set forth relating to specific components, and methods, to provide a complete understanding of embodiments of the present disclosure. It will be apparent to the person skilled in the art that the details provided in the embodiments should not be construed to limit the scope of the present disclosure. In some embodiments, well-known processes, well-known apparatus structures, and well-known techniques are not described in detail. As used in the following, the terms “have”, “comprise” or “include” or any arbitrary grammatical variations thereof are used in a non-exclusive way. Thus, these terms may both refer to a situation in which, besides the feature introduced by these terms, no further features are present in the entity described in this context and to a situation in which one or more further features are present. As an example, the expressions “A has B”, “A comprises B” and “A includes B” may both refer to a situation in which, besides B, no other element is present in A (i.e., a situation in which A solely and exclusively consists of B) and to a situation in which, besides B, one or more further elements are present in entity A, such as element C, elements C and D or even further elements. Further, as used in the following, the terms “preferably”, “more preferably”, “particularly”, “more particularly”, “specifically”, “more specifically” or similar terms are used in conjunction with optional features, without restricting alternative possibilities. Thus, features introduced by these terms are optional features and are not intended to restrict the scope of the claims in any way. The invention may, as the skilled person will recognize, be performed by using alternative features. Similarly, features introduced by "in an embodiment of the invention" or similar expressions are intended to be optional features, without any restriction regarding alternative embodiments of the invention, without any restrictions regarding the scope of the invention, and without any restriction regarding the possibility of combining the features introduced in such way with other optional or non-optional features of the invention. The block diagram of the Voice Recognition System: Snowboy Smart Speaker is shown in **Figure 2.**

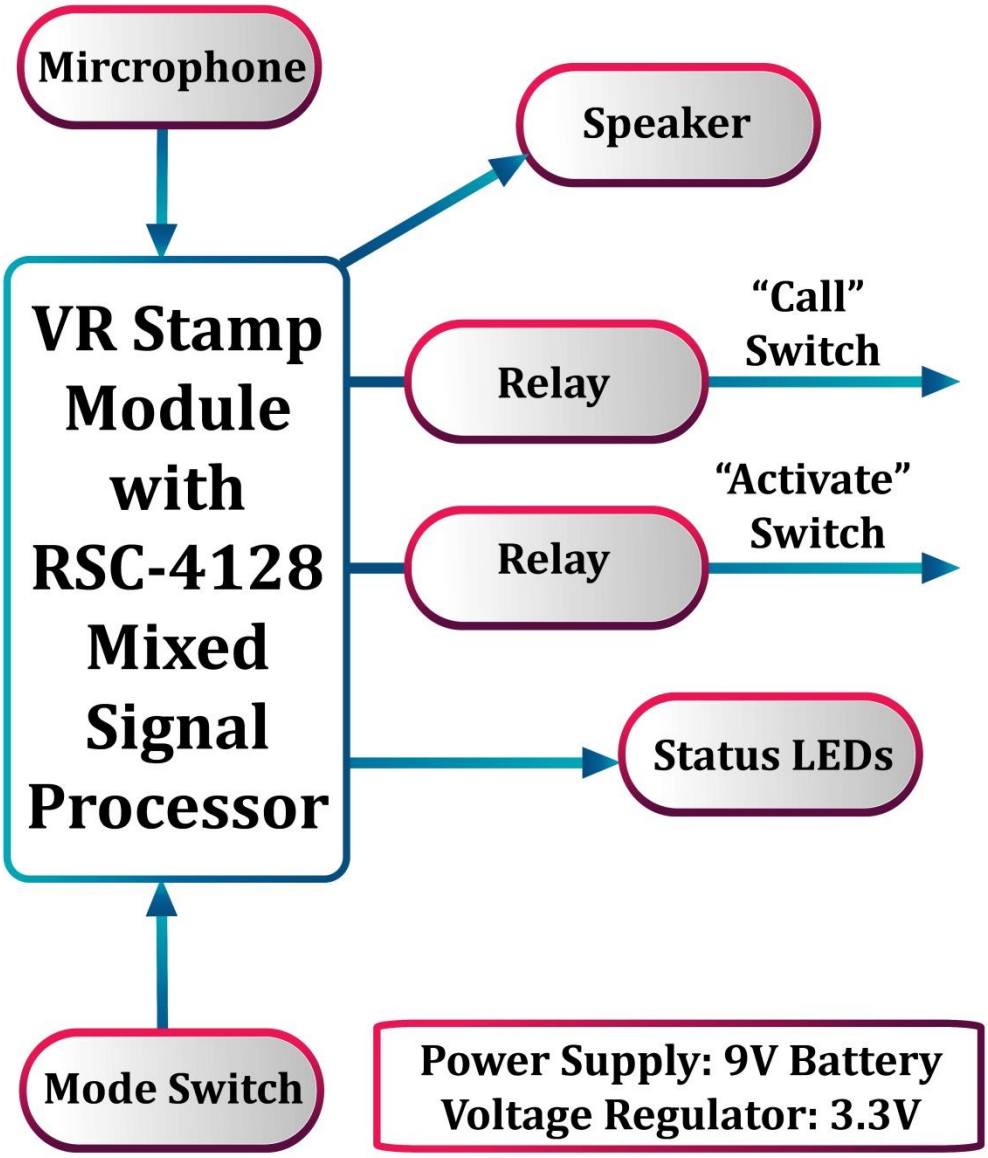

Figure 2: The block diagram of the Voice Recognition System: Snowboy Smart Speaker.

The following detailed description illustrates the invention by way of example, and not by way of limitation. This description will clearly enable one skilled in the art to make and use the invention and describes several embodiments, adaptions, variations, alternatives, and uses of the invention, including what we presently believe is the best mode of carrying out the invention. The General Architecture of a Voice Assistant: Snowboy Smart Speaker is shown in **Figure 3**.

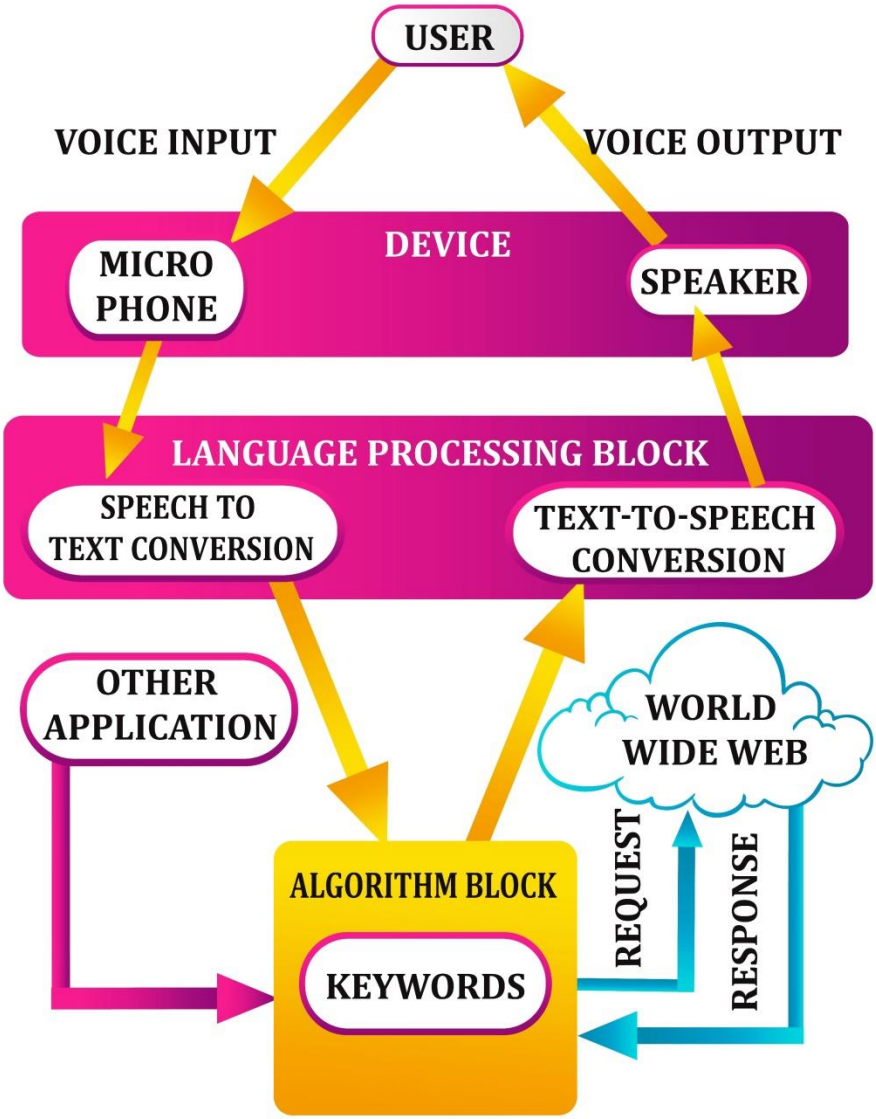

Figure 3: The General Architecture of a Voice Assistant: Snowboy Smart Speaker.

The present invention envisages a system and a method for providing a customized service to user. The present invention envisages an exemplary environment for placing a request corresponding to one or more users' queries. In one case, the system is capable of storing personalized questions and accordingly providing customized answers to the user. A service providing server receives the request from a user via a user terminal. In one case, the user terminal may correspond to a terminal installed in a space in which a user lives, such as a living room of a home, and includes a microphone and a speaker integrally formed therewith, or a display device, communicatively coupled with a wireless communication terminal such as a smartphone carried by the service providing server or the user. Further the system comprises the user terminal based on a Raspberry Pi4 platform running RASPIAN Linux and supporting WIFI, BLUETOOTH, ZWAVE and

ZIGBEE dongles, may be used to host the user terminal application. Further, the speaker instilled in the user terminal may correspond to 3W twin speaker. In an embodiment, the user may generate a request by a voice command saying, "What is the current weather?". In such a case, the system may process the voice command and accordingly provide a solution by fetching it from a cloud server. In an embodiment, the user may generate a request by a voice command saying, "Please play the latest song?". Upon providing the solution, the user may further recommend "Please recommend another song", and accordingly the response is provided to the user based upon requested message. It is to be noted that the user terminal outputs the customized answers from the service providing server to the user. In another embodiment, the user may be interested in knowing to-do list of the user. In one case, the instructions pertaining to to-do list of the user corresponds to already save data by the user. For an instance, the user terminal may remind the user to take medicines at the fixed time duration, previously stored by the user. In yet another embodiment, the user may generate a request for operating home appliances, or the device for the Internet of Things (IoT) such as the lighting device, heating device, and air conditioner, which are operated in conjunction with the service providing server, and the received request by the user terminal is stored in storage unit of the service providing server. Method comprises the step of receiving a request of each of a plurality of users from a user terminal, wherein the request corresponds to at least an instruction generated by the user or a reminder pertaining to said instruction of the user. The method further includes the step of analyzing the request to identify a type of customized service in the instruction, wherein the type of customized service corresponds to updating a pre-stored set of data. The method furthermore includes the step of generating a suggestive report depicting an actionable intelligence based on said analysis, wherein the suggestive report comprises an output pertaining to the customized service.

Technical Advancements of the Present Research Work: The present disclosure described herein above has several technical advantages including, but not limited to, the realization of a system and a method for supervising content items displayed on a signage device that:

1. Personalize questions and customize answers according to user's requirement.

2. Results into inexpensive setup.

3. Performs analysis based on Artificial Intelligence and is easy to use; economical; and employs machine learning for accuracy.

The embodiments herein and the various features and advantageous details thereof are explained with reference to the non-limiting embodiments in the following description. Descriptions of well-known components and processing techniques are omitted so as to not unnecessarily obscure the embodiments herein. The examples used herein are intended merely to facilitate an understanding of ways in which the embodiments herein may be practiced and to further enable those of skill in the art to practice the embodiments herein. Accordingly, the examples should not be construed as

limiting the scope of the embodiments herein. The use of the expression "at least" or "at least one" suggests the use of one or more elements or ingredients or quantities, as the use may be in the embodiment of the disclosure to achieve one or more of the desired objects or results. The foregoing description of the specific embodiments will so fully reveal the general nature of the embodiments herein that others can, by applying current knowledge, readily modify and/or adapt for various applications such specific embodiments without departing from the generic concept, and, therefore, such adaptations and modifications should and are intended to be comprehended within the meaning and range of equivalents of the disclosed embodiments. It is to be understood that the phraseology or terminology employed herein is for the purpose of description and not of limitation. Therefore, while the embodiments herein have been described in terms of preferred embodiments, those skilled in the art will recognize that the embodiments herein can be practiced with modification within the spirit and scope of the embodiments as described herein. The above description is considered that of the preferred embodiment(s) only. Modifications of these embodiments will occur to those skilled in the art and to those who make or use the illustrated embodiments. Therefore, it is understood that the embodiment(s) described above are merely exemplary and not intended to limit the scope of this disclosure, which is defined by the following claims as interpreted according to the principles of patent law, including the doctrine of equivalents.

Conclusions: Following points summarize the investigational results with key objectives of the present research article "Arrangement and Technique for Exploring an Appeal to Deliver a Customized Voice Recognition Service towards the User: Snowboy Smart Speaker":

1. An object of the present disclosure is to provide a system and a method for analyzing a request to provide a customized service to user.

2. Another object of the present disclosure is to provide a system and a method that involves flexibility to modify the systems as per user requirement. Yet object of the present disclosure is to provide a customized voice recognition service considering unique information for each individual user when there is a plurality of users using the same voice recognition service.

3. Still another object of the present disclosure is to provide a system and a method that is capable of working in online (with internet connectivity) and offline environment. It controls all the home electrical appliances using smartphone application and smart speakers like Amazon Alexa, Google Home.

4. Yet another object of the present disclosure is to provide a system and a method that generates a suggestive report depicting an actionable intelligence based on said analysis, wherein the suggestive report comprises an output pertaining to the customized service.

5. A further object of the present disclosure is to provide a system and a method that is inexpensive and easy to use. Yet another object of the present disclosure is to provide a system and a method that is economical.

6. Another object of the present disclosure is to provide a system and a method that performs analysis based on artificial intelligence. A still further object of the present disclosure is to provide a system and a method that employs machine learning for accuracy.

Acknowledgements: The authors thank Lucknow Public Educational Society (LPES), 'A' Block, Rajajipuram, Lucknow for financial assistance under a research grant no. 2017/MRP/02/LPES.

References:

[1]. A guide to the security of voice-activated smart speakers an ISTR special report, https://www.symantec.com/content/dam/symantec/docs/security-center/white-papers/istr-security-voice-activated-smart-speakers-en.pdf.

[2]. Picovoice benchmark, https://github.com/Picovoice/wakeword-benchmark.

[3]. Home - sensory (2014), https://www.sensory.com/.

[4]. Trulysecure - sensory (2014), https://www.sensory.com/products/technologies/trulysecure/.

[5]. Amazon.com help: About Alexa voice profiles (2019), https://www.amazon.com/gp/help/customer/display.html?nodeId=202199440.