# Turnitin Originality Report

Check by Shriram Vasudevan

From Paper (CSE Papers)

Processed on 18-Feb-2015 4:27 PM
SGT
ID: 506667704
Word Count: 1499

| Similarity Index | Similarity by Source | |
|---|---|---|
| 4% | Internet Sources: | N/A |
| | Publications: | 4% |
| | Student Papers: | N/A |

**sources:**

1 3% match (publications)
Angelo Ciaramella. "ICA based identification of dynamical systems generating synthetic and real world time series", Soft Computing, 05/2006

2 1% match (publications)
Ozdogan, M.. "The spatial distribution of crop types from MODIS data: Temporal unmixing using Independent Component Analysis", Remote Sensing of Environment, 20100615

3 1% match (publications)
N. Fakhfakh. "3D Objects Localization Using Fuzzy Approach and Hierarchical Belief Propagation: Application at Level Crossings", EURASIP Journal on Image and Video Processing, 2011

**paper text:**

Voice based Interactive Systems for Visually Challenged and physically disabled Vamsee Krishna Kiran, Assistant Professor, Department of computer science engineering, Amrita school of engineering, Amrita Vishwa Vidyapeetham (University), Ettimadai, Coimbatore -641112. mk_vamseekrishna@cb.amrita.edu Shriram K Vasudevan, Assistant Professor (Sr), Department of computer science engineering, Amrita school of engineering, Amrita Vishwa Vidyapeetham (University), Ettimadai, Coimbatore -641112. Kv_shriram@cb.amrita.edu Archanaa Rajendran, Assistant Professor, Department of computer science engineering, Amrita school of engineering, Amrita Vishwa Vidyapeetham (University), Ettimadai, Coimbatore -641112. r_archanaa@cb.amrita.edu Abstract Voice is one of the five senses that plays a major role in human life. Specially for those people who are visually challenged and physically disabled, voice will play a phenomenal role in identifying persons, communication etc. Voice based systems are not new to computer science domain, since the first speech system that is integrated into an operating system named atari[1], there are many other speech systems that are competing in today's world. The idea of this paper is not in enhancing or optimizing the speech systems rather it concentrates on application of speech systems so as to facilitate the visually challenged people for performing some simple day- to-day tasks on a computer. After all humans will not use computers just for programming there are many other uses of a computer for visually challenged and physically disabled people. Keywords: Voice Recognition, speech synthesis, visually challenged, Speech applications, Independent Component Analysis. Introduction In computer science, voice based systems are those that accepts voice commands as input and performs respective actions as output. This output need not be always an action it can be a simple voice response from the system too. In general voice based systems consists of two major components: one voice recognition and second voice synthesis. Voice recognition is the translation of spoken words into text, also called "speech to text" (STT). Speech synthesis is the translation of "text to speech" (TTS). Speech

synthesis is artificial production of human voice. The system that helps doing this is called as speech synthesizer. Stephen Hawking is one of the most popular persons who uses speech synthesis to communicate. Both speech recognition and speech synthesis in current systems still persist to have some problems. Most of the modern day speech recognition systems either use “speaker independent speech recognition systems” or training where a user reads sections of text into the speech recognition system. These speech systems are not only confined to traditional computers but also has wide range of uses in mobile platforms too. In current market we have three major competitors who are working in speech recognition and synthesis mainly focussing the mobile platforms. First one is Google Now, second being Apple’s siri and the latest being the Microsoft’s Cortana. Although there is a huge debate going on in the industry to decide which among them performs better? most of the polls are closer to Microsoft’s cortana[4][5]. Our focus is not in exploring the various varieties of speech recognition systems rather we are more interested in using anyone of these available speech systems and developing an application that can facilitate and ease the access of computer for visually challenged people. Proposed System There are ‘n’ number of situations where visually challenged or physically disabled person want to use a computer. For example: listening to music, accessing youtube, accessing online news, typing a small notepad, saving files on to cloud, retrieving files, controlling system with voice etc. To facilitate this kind of access to a computer we have developed our Speech recognition application and we call it as “Swaram” as shown in Figure-1. Our speech recognition application runs on top of the Operating system and runs as an automatic startup process. This application runs as a background process while the user is working on something else. Our aim of the proposed system is to control a computer just by speaking out some basic commands. We use speech synthesis to read out the contents in a screen so that it will help to visually challenged persons to select the necessary item they need on the screen. Speech synthesis is also used to speak out some alerts to the user notifying the person with the current operations that are running on the system. Our aim of the project is to build such a speech recognition application that will run on top of OS even performing the system related functions taking a precedence over the other applications. Figure-1: User Interface of Swaram Current application Swaram is capable of doing a number of tasks like answering some basic questions, controlling system by locking desktop, shutting down, closing applications etc. Swaram is a customisable application where a command and action need to be performed can be added dynamically to the system as shown in Figure-2. Swaram can perform a web search just by giving commands, user can open text editor, write content and can save it as shown in Figure- 3. Swaram is also capable of indexing all the music files and can play it for visually and physically challenged people. Swaram is also capable of accessing cloud storage of a user where the user can upload his files using simple voice commands. Swaram is just a prototype that is built to address some simple problems that are faced by visually and physically challenged people. Figure-2: Adding commands to Swaram Figure-3: Typing notepad and saving it using swaram Challenges The major challenge that comes with speech recognition is noise. Speech recognition applications will not interpret the commands properly when the environment is noisy. In order to enhance the accuracy of speech recognition we have adopted Independent Component Analysis (ICA) for noise separation [2]. ICA is another version of Blind Source Separation, this is used in our application so as to differentiate the commands from the external noise. Implementing ICA in the application has improved the efficiency of speech recognition from 72% to 88% provided the system is operated in a less noisy environment rather than an open environment. Independent Component Analysis for Noise removal Independent Component Analysis

3**is a special case of Blind Source Separation**

problem and

2**superficially related to Principal Component Analysis(PCA)[** 3]. **ICA is a**

statistical and computational method for separating multivariate signals into additive subcomponents. In ICA ,

1**the data variables are assumed to be linear mixtures of some unknown latent variables and the mixing system is also unknown.** These **latent variables are assumed non-gaussian and mutually independent,** thus **they are called Independent Components** or Source factors **of observed data** which **can be** found by ICA **as**

shown in Figure-4. Figure-4: Independent Component Analysis Implementation Swaram is mainly developed for Windows platform so as to target large number of audience. It is built as Windows Desktop application using Windows speech in .NET framework. In this application we have also adopted Independent Component Analysis for noise removal which is implemented combining .NET class libraries with Matlab. We have used windows speech package of .NET and application is developed by inheriting its basic speech recognition and synthesis libraries. Our application is just like any other application running on Operating System. It will be in listening mode once it starts during the windows startup. Inputs are processed as and when received. Our application maintains a commands file in which commands and respective actions are stored. If the inputs are not matching with any of the commands they are just ignored. If the inputs received are matched with the commands then respective action is performed. The system flow is clearly explained in Figure-5. Figure-5: System flow Although there are different speech recognition softwares available for the desktops, no speech recognition software is highly customizable. Windows operating system itself will come with a speech recognition software with which we can operate the system. But customizing the software is a drawback. Our implementation is not a complete product but it forms a good prototype which can be customized to a good extent easily and facilitating the physically challenged people. Conclusion and future scope In this paper we have introduced a new speech recognition application called Swaram that mainly targets the physically and visually challenged people and facilitates them in using a computer with simple voice based commands. This prototype can also be deployed into embedded devices like washing machines, coffee vending machines etc. where we can use voice synthesis to guide a user step by step on how to operate the machines by accepting voice commands from him. In future this system can be integrated with neural networks so as to train the system on different voices in different environments. This can probably improve the accuracy of speech recognition. References [1] Atari speech System, Available at: http://en.wikipedia.org/wiki/Atari_ST, accessed on 24 January, 2015 [2] Lee, Te-Won. Independent component analysis. Springer US, 1998. [3] Wold, Svante, Kim Esbensen, and Paul Geladi. "Principal component analysis." Chemometrics and intelligent laboratory systems 2.1 (1987): 37-52. [4] Google Now vs Siri vs Cortana: showdown, http://www.phonearena.com/news/Google-Now- vs-Siri-vs-Cortana-showdown_id59877, accessed on 22 January, 2015. [5] Cortana vs. Siri vs. Google Now: An early look at how Cortana stacks up, http://www.cnet.com/news/cortana-vs-siri-vs-google-now/, accessed on 22 January, 2015.