International Journal of Engineering Research and Technology. ISSN 0974-3154 Volume 10, Number 1 (2017)

Personal Assistant with Voice Recognition Intelligence

Dr. Kshama V. Kulhalli
Vice-Principal & HOD IT
D.Y. Patil College of Engineering and Technology,
Kolhapur-416006

Dr.Kotrappa Sirbi
Department of Computer Science and engineering,
KLE's Dr. M.S.Sheshagiri College of Engineering & Technology,
Belgaum, India

Mr. Abhijit J. Patankar
Research Scholar, Computer Science and Engineering,
VTU, Belgaum, Karnataka.

Abstract
The Most famous application of iPhone is "SIRI" which helps the end user to communicate end user mobile with voice and it also responds to the voice commands of the user. Same kind of application is also developed by the Google that is "Google Voice Search" which is used for in Android Phones. But this Application mostly works with Internet Connections. But our Proposed System has capability to work with and without Internet Connectivity. It is named as Personal Assistant with Voice Recognition Intelligence, which takes the user input in form of voice or text and process it and returns the output in various forms like action to be performed or the search result is dictated to the end user. In addition, this proposed system can change the way of interactions between end user and the mobile devices. The system is being designed in such a way that all the services provided by the mobile devices are accessible by the end user on the user's voice commands.

Keywords: SIRI, Google Voice Search, Mobile Device, Internet

Introduction

Nowadays the Mobile Technology is being very famous for the User Experience, because it is very easy to access the applications and services from anywhere of your Geo-location. Android, Apple, Windows, Blackberry, etc. are various famous and commonly used Mobile Operating Systems. All the Operating Systems provides plenty of applications and services for users.

For an instance, the Contacts Applications is used to store the contact details of the user's contact and also helps user to connect a call or send an SMS to other person using the contents stored in this application. We can get similar types of application all around the world via Apple Store, Play Store, etc. All this features gives birth to various kinds of sensors or functionalities to be implemented in the mobile devices.

The Most famous application of iPhone is "SIRI" which helps the end user to communicate end user to mobile with voice and it also responds to the voice commands of the user. Same kind of application is also developed by the Google that is "Google Voice Search" which is used for in Android Phones. But this Application mostly works with Internet Connections. But our Proposed System has capability to work with and without Internet Connectivity.

It's named as Personal Assistant with Voice Recognition Intelligence, which takes the user input in form of voice or text and process it and returns the output in various forms like action to be performed or the search result is dictated to the end user.

Literature Survey

Speech recognition has a long history with several waves of major innovations. Speech recognition for dictation, search, and voice commands has become a standard feature on smartphones and wearable devices.

Design of a compact large vocabulary speech recognition system that can run efficiently on mobile devices, accurately and with low latency. [1] This is achieved by using a CTC-based LSTM acoustic model which predicts context-independent phones and is compressed to a tenth of its original size using a combination of SVD-based compression and quantization. Quantized deep neural networks (DNNs) and on-the-fly language model rescoring to achieve real-time performance on modern smartphones.

The ASR and Search components perform speech recognition and search tasks. In addition to ASR and Search, we also integrate a query parsing module between ASR and Search for a number of reasons. [3] Set of techniques for improving the performance of automated voice search services intended for mobile users accessing these services over a range of portable devices. Voice search is implemented as a two stage search procedure where string candidates generated by an automatic speech recognition (ASR) system are re-scored in order to identify the best matching entry from a potentially very large application specific database. Study provides a good example of how additional domain specific knowledge sources can be used with a domain independent ASR system to facilitate voice access to online search indices.

As more data becomes available for a given speech recognition task, the natural way to improve recognition

accuracy is to train larger acoustic models. There are a non-parametric empirical model that exploits abundant training data to directly learn pronunciation variation. Interpolating the empirical model with a parametric model yields the best performance, with a relative improvement of 5.2% in WER over the baseline. [2] There are a number of ways in which this work could be extended. First, closer integration with acoustic model training is likely to yield sharper distributions and a tighter fit to the data. Second, estimating word-pronunciation co-occurrence counts in semi-supervised fashion (e.g. through word recognition instead of forced alignment) would broaden its applicability to a wide range of speech genres and tasks.

Table 1: Search on Dictation and Commands

LM Setup	Dictation WER	Commands WER
Linear Interpolation	12.9	10.0
Bayesian Interpolation	12.3	8.9
Bayesian + Rescoring	13.5	9.0

Finally, it would be of interest to modify our models to factor out the distinct phenomena that affect pronunciation (e.g. accent, dialect, recognition errors). This paper focuses on the development and advances in automatic speech recognition for the AT&T Speak4it R voice search application [5]. With Speak4it as real-life example, we show the effectiveness of acoustic model (AM) and language model (LM) estimation (adaptation and training) on relatively small amounts of application field-data.

Methods are investigated for identifying the most likely database entry associated with the decoded utterance. An experimental study is presented describing the characteristics of actual user utterances obtained from a prototype voice search service. The impact of these methods on word error rate is presented.

Proposed Architecture

PARI has various branches of the services, but the main feature of PARI is Voice Recognition Engine which has an ability to work without internet connection i.e. Offline Voice Recognition.

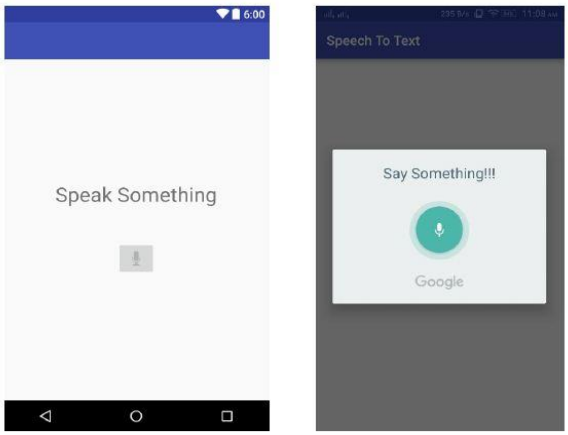

Figure 1: Voice Recognizer Activity

This Voice Recognizer works offline and performs various operations as per the user commands and requirements. This is the first activity that's opens whenever we starts PARI and it directly can be opened by pressing Power Button of the mobile devices.

This feature is specially design for Blind Persons who wish to use the Android Smart Phones but are unable to connect this technology. Also Native user who barely knows to unlock their smart phones can easily open this application and using voice commands in their local languages as per need.

PARI responds to basic commands like, Open Applications, Close Applications, Connect Calls To respective person/contacts, send text SMS to respective person or contact, Capture Photos from camera (Front Camera/ Back Camera), Add/Delete/Update Contacts, Run any media file, Start various services like Hotspot, Wi-Fi, Bluetooth, and various Services from the respective Notification Panel. All this can be performed on the voice commands of the end user without internet connectivity.

Operations such as Browsing or Searching for any topic, using Applications that need internet connections example "Send message to ABC, Hi I'm PQR sending message to you. Using WhatsApp". All these basic operations are performed by the voice commands of the end user.

Google did quietly enable offline recognition in that Search update, but there is no API or additional parameters available within the Speech Reconginzer Class. The functionality is available with no additional coding, however the user's device will need to be configured correctly for it to begin working and this is where the problem.

Also, Google have restricted certain Jelly Bean devices from using the offline recognition due to hardware constraints. Which devices this applies to is not documented, in fact, nothing is documented, so configuring the capabilities for the user has proved to be a matter of trial and error.

Steps to Start Offline Voice Recognizer in Latest Android Smart Phones-

1. Make sure the default Android Voice Recognizer is set to Google
2. Uninstall any offline recognition files you already have installed from the Google Voice Search Settings
3. Go to your Android Application Settings and see if you can uninstall the updates for the Google Search and Google Voice Search applications.
4. If you can't do the above, go to the Play Store see if you have the option there.
5. Reboot (if you achieved 2, 3 or 4)
6. Update Google Search and Google Voice Search from the Play Store (if you achieved 3 or 4 or if an update is available anyway).
7. Reboot (if you achieved 6)
8. Install English UK offline language files
9. Reboot
10. Use Utter with a connection
11. Switch to aero plane mode and give it a try
12. Once it is working, the offline recognition of other languages, such as English US should start working too.

Language is not a Barrier for PARI to understand the user voice commands for performing respective operations. Hence PARI uses its Intelligence to store the voice commands detected by the end user with respect to the Default Commands stored in PARI.

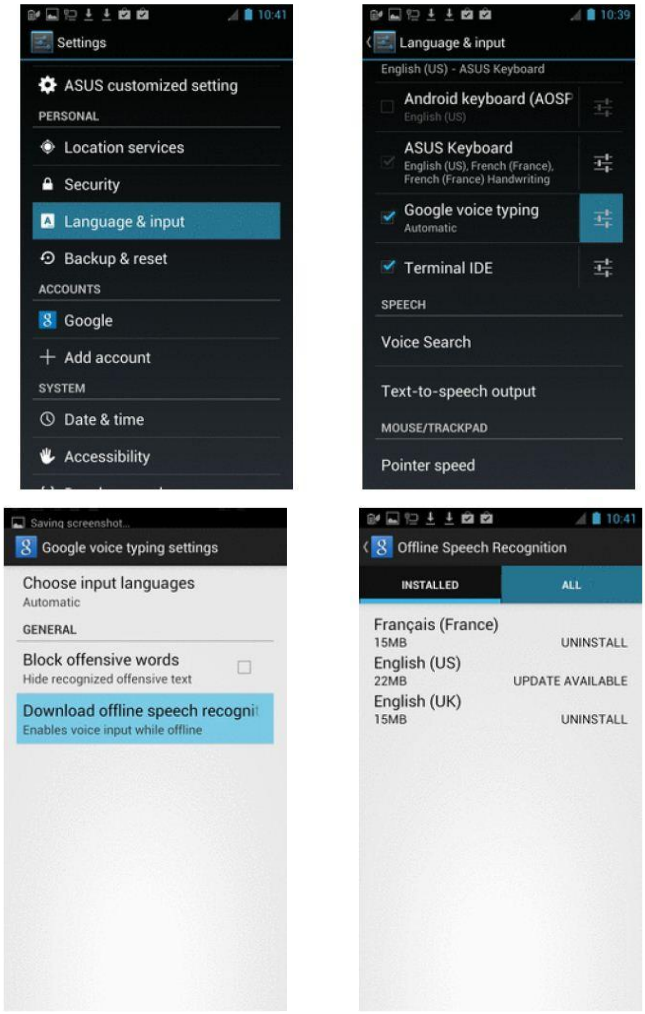

Figure 2: Selecting Various Languages

Extra Features

We have also added some extra features. These are:

1. Sync up your storage

Your computer has local storage, as does your phone -- two separate, unconnected virtual vats of space. But with a clever cloud-embracing app and a few minutes of configuration, the devices' drives can act as if they're one.

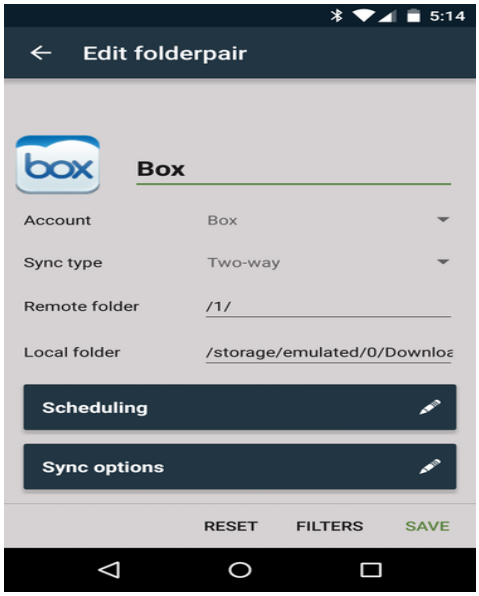

Figure 3: Sync Up Storage

The Android app FolderSync allows you to sync up storage between your desktop PC and Android smartphone. The secret resides in FolderSync, an Android utility that costs a mere (There's a free version, but it's peppered with ads and offers limited functionality.) FolderSync works with a ton of cloud storage providers, including Amazon, Box, Dropbox, Google Drive, and OneDrive. As long as the provider you select offers a companion program on the PC side (all of the ones I mentioned do), you'll be good to go.

To get everything up and running, first install the app on your phone and follow the prompts to connect it to the cloud storage service of your choice. Set up "pairs" for any folders you want to keep linked with your computer -- a folder containing your documents or downloads, for instance -- and create new folders in your cloud storage to match.

Be sure to set the pair to use two-way sync. You can either opt to sync instantly, if you want everything to be kept up-to-date at all times, or you can go for a more battery-friendly setup like syncing once a day if you don't mind a little bit of latency.

Now install the desktop app for whatever cloud service you're using. Open the app, find the paired folder you created, and get it ready for use. If you paired a Documents folder, for example, you might place a shortcut on your desktop and dump all of your existing documents into it.

Then treat that folder as if it were a regular local resource. Anything you do in it will be saved onto your hard drive and synced into the cloud, where it'll automatically find its way onto your Android device. Any changes on your phone will make their way back to your computer in the same manner.

2. Find and secure your phone

Can't find your phone in its usual place between couch cushions? Or, worse, get home from a day of travel only to realize you lost your device somewhere along the way? No need to panic: Your computer can tell you exactly where your mobile buddy is. It can even remotely lock it down and erase it if need be.

All you have to do is get on your computer (or any computer, really), navigate togoogle.com in the browser, and type "Find

my phone" into the search box. (You'll have to sign into Google first -- which you'll probably want to do in an incognito window if it isn't your own PC.) Within seconds, Google will give you a detailed map showing your smartphone's last logged location. Clicking on it will bring up the full Android Device Manager interface, where you'll find options to ring, lock, or fully erase your phone right then and there.

There's one catch: Your phone has to be configured beforehand to allow all of that to happen. Take two minutes right now, while your phone is safe and sound, and open up the app called Google Settings on your device. (Note that we're talking about Google Settings, which isn't the same as your regular system settings.)

Head into the Security menu and make sure both "Remotely locate this device" and "Allow remote lock and erase" are activated. Now take a deep breath, relax, and think of something else to worry about.

3. Access your entire computer from your phone

We couldn't talk about connecting your computer with your Android phone without mentioning the most direct connection of all: Being able to access yourentire computer from your smartphone's screen. Thanks to Google's free Chrome Remote Desktop app, it's easier than ever to do.

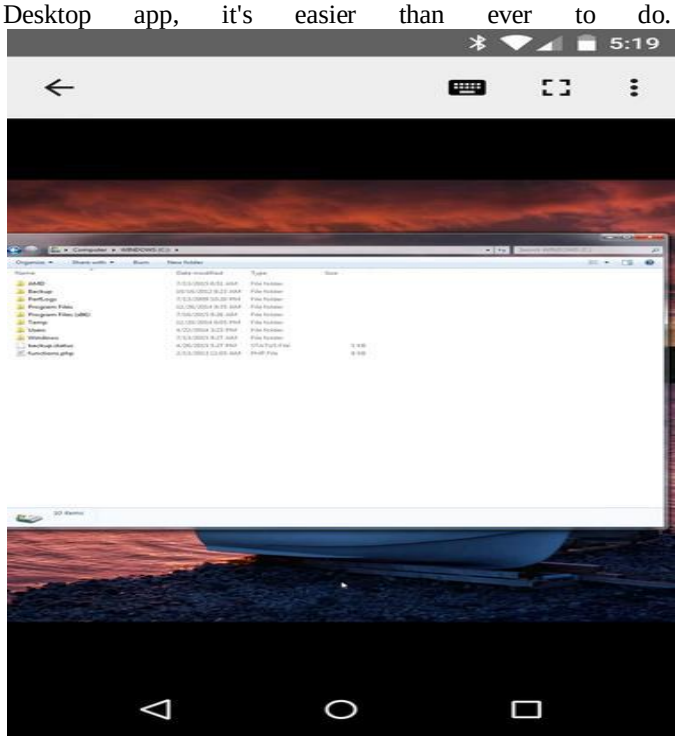

Figure 4: Access entire computer from phone

Accessing your desktop from your phone is easy, thanks to Chrome Remote Desktop. First, get the app on your phone and put the companion desktop appon your computer. (The desktop app will work on any platform where the Chrome browser can run.) Then pull up the Android app the next time you need to hop onto your computer remotely. You'll be able to move around your desktop, manipulate files, and even run programs from the palm of your hand.

Conclusion

PARI is Designed to help Native and especially for Blind persons which works on their Voice Commands. PARI also has the capability of recognizing the voice commands without internet connection.

PARI has various functionalities of mobile devices like network connection and managing various applications on just the voice commands. Contains key features like Voice Pattern Detection, Keyword Learning, etc. which helpful for end user to use various functionalities and services of the mobile devices. Hence, PARI is language barrier independent which actively responds to user's voice commands faster than the Online Voice Search applications.

References

[1] "Accurate and compact large vocabulary speech recognition on mobile devices," in INTERSPEECH. 2013, pp. 662–665, ISCA.

[2] "Connectionist temporal classification: Labelling unsegmented sequence data with recurrent neural networks," in ICML, 2006, pp. 369–376.

[3] "Lattice-based optimization of sequence classification criteria for neural-network acoustic modeling".

[4] "Language modeling for What-with Where on GOOG-411," Proc. of INTERSPEECH, 2009.

[5] "DFirect construction of compact context-dependency transducers from data," Proc. of INTERSPEECH, pp. 218–221, 2010.

[6] "Automatic question generation for decision tree based state tying," Proc. of ICASSP, 1998.

[7] "Speech recognition with flat direct models," IEEE Journal of Selected Topics in Signal Processing, 2010.