International Journal of Advances in Science Engineering and Technology, ISSN(p): 2321 –8991, ISSN(e): 2321 –9009
Volume-6, Issue-1, Jan.-2018, http://iraj.in

REAL TIME VISUAL TRANSLATOR

[1]ROHINI MOHAN, [2]IYER ARUN SANKAR

[1,2]SRM University,Ramapuram,Chennai.
E-mail: [1]reach.rohinimohan@gmail.com, [2]arun.i1997@gmail.com

Abstract - Due to digitalization, and common governmental procedures, medical records being online, illiterate individuals require assistance at almost every step. Language barrier further makes it difficult to read sign boards, posters in another language especially for tourists and students. This led our team to come up with the idea of real time visual translator. The Real time visual Translator is an Android powered application that uses Computer vision to convert image to text and text to speech output and Utilizes OCR.

Keywords - OCR,TTS, Speech Synthesis, Real Time visual translator.

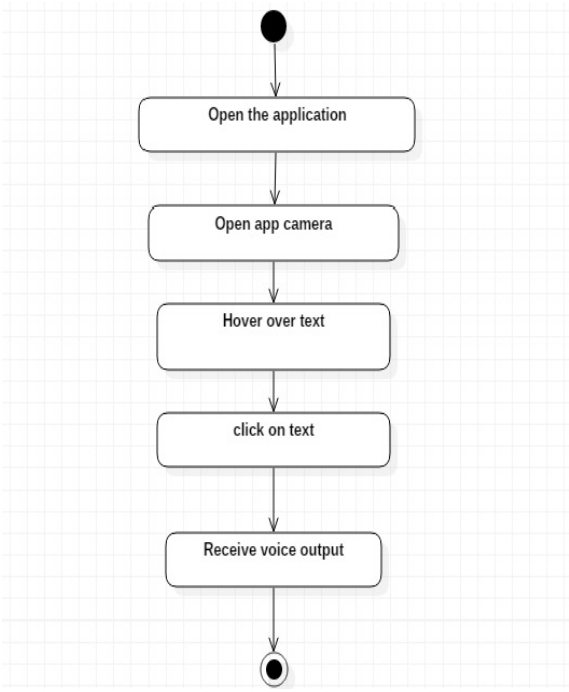
I. INTRODUCTION

Voice is the easiest interface to communicate and the advent of speech synthesis has opened a world of opportunities relating to voice recognition and speech made it convenient to deliver information. In addition to this, text detectors have brought in a wave of revolution helping individuals overcome language barriers. There have been various text detectors and voice detectors in the form of apps and paid software to aid this purpose but the thought of incorporating both these concepts gave a start to the real time visual translator that can convert text in real time to speech output. We intend to implement this concept using Optical Character recognition (OCR).

Though there are various detectors for English, there have been very few instances where a native or Indian language is taken up because Devanagari and other Indian scripts are regarded as one of the most challenging steps in the digitization of Indian literature. OCR refers to the process by which scanned images are electronically read. The objective here is to convert the text image into an editable text form. Text document scanned using the scanner is turned into bitmap files. OCR software identifies the bitmap to corresponding alphabets and numbers. Once recognized, the characters are converted into ASCII/UNICODE.

II. OVERVIEW

The workflow of the application is simple. It can be This would require us an Android device running minimum 4.2 Jellybean with 5mp Camera, 1 GB RAM, internet. The app utilizes a seamless UI for individuals who find it difficult to read. The app gives a choice of language preference followed by a camera option. The camera detects the text when hovered over it. The required text can then be selected and that text will be converted to speech format and read out loud.
Text-to-speech synthesis takes place in several steps. The TTS system gets a text as input, which it first must analyze and then transform into a phonetic description. Then in a further step it generates the prosody. From the information now available, it can produce a speech signal.



III. TEXT CONVERSION

First the text is segmented into tokens. The token-to-word conversion creates the orthographic form of the token. For the token "Nr." the orthographic form "Nummer" is formed by expansion, the token "12" gets the orthographic form "twelve" and "1995" is transformed to "nineteen ninety five". This expansion is sometimes not so easy, as can be seen with the example of the number "1": It has to be expanded differently depending on what it denotes: in a street address to "eins", in "1 Kilogramm" to "ein"; in the expression "1 Katze jagt 1 Hund", first to "eine" and then to "einen".

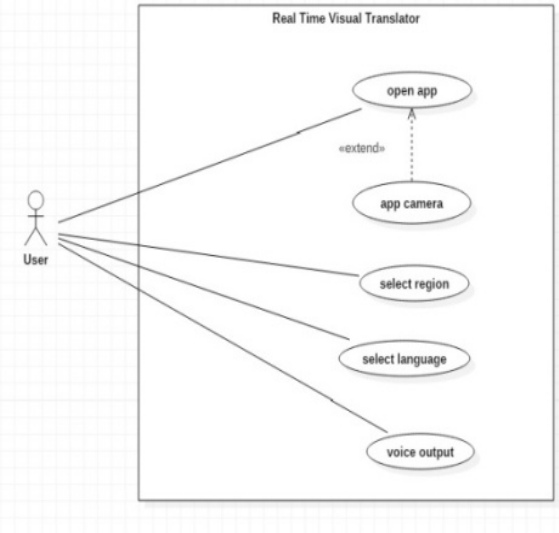


Figure 2: use case diagram

IV. PRELIMINARY INVESTIGATION

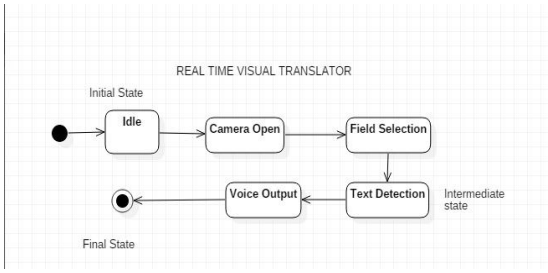
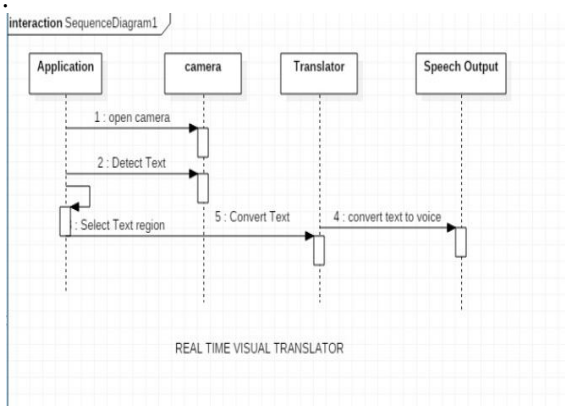


Figure 3: State chart diagram

Language technologies can provide solutions in the form natural interfaces so that digital content can reach to the masses and facilitate the exchange of information across different people speaking different languages. There are already many speech synthesizers existing for English. But there is no standard speech synthesizer for other indian languages. Government of India has made several efforts to make village people to take advantages of advances in the field of IT. But still there exists a digital divide.

V. SAMPLE CODE

.




```
 @Override
public void onRequestPermissionsResult(int requestCode,
                                       @NonNull String[] permissions,
                                       @NonNull int[] grantResults) {
    if (requestCode != RC_HANDLE_CAMERA_PERM) {
        Log.d(TAG, "Got unexpected permission result: " + requestCode);

super.onRequestPermissionsResult(requestCode, permissions, grantResults);
        return;
    }
   if (grantResults.length != 0 && grantResults[0] == PackageManager.PERMISSION_GRANTED) {
        Log.d(TAG, "Camera permission granted - initialize the camera source");
        // We have permission, so create the camerasource
        boolean autoFocus = getIntent().getBooleanExtra(AutoFocus,false);
        boolean useFlash = getIntent().getBooleanExtra(UseFlash, false);
        createCameraSource(autoFocus, useFlash);
        return;
    }

    Log.e(TAG, "Permission not granted: results len = " + grantResults.length +
            " Result code = " + (grantResults.length > 0 ? grantResults[0] : "(empty)"));

    DialogInterface.OnClickListener listener = new DialogInterface.OnClickListener() {
        public void onClick(DialogInterface dialog, int id) {
            finish();
        }
    };

    AlertDialog.Builder builder = new AlertDialog.Builder(this);
    builder.setTitle("Multitracker sample")

.setMessage(R.string.no_camera_permission)
            .setPositiveButton(R.string.ok, listener)
            .show();
  }

  /**
```

VI. IMPLEMENTATION

Real time visual translator is an application that converts text into spoken word, by analyzing and processing the text using OCR and then using

Digital Signal Processing (DSP) technology to convert this processed text into synthesized speech representation of the text. Here, we developed a useful text-to-speech synthesizer in the form of a simple application that converts inputted text into synthesized speech and reads out to the user . The development of a text to speech synthesizer will be of great help to people with visual impairment and make making through large volume of text easier

CONCLUSION

Text to speech synthesis is a rapidly growing aspect of computer technology and is increasingly playing a more important role in the way we interact with the system and interfaces across a variety of platforms. We have identified the various operations and processes involved in text to speech synthesis. We have also developed a very simple and attractive graphical user interface which allows the user to type in his/her text provided in the text field in the application. In future, we plan to make efforts to create engines for localized Indian language so as to make text to speech technology more accessible to a wider range of Indian languages. Another area of further work is the implementation of a text to speech system on other platforms, such as telephony systems, ATM machines, video games and any other platforms where text to speech technology would be an added advantage and increase functionality.

REFERENCES

[1] https://en.wikipedia.org/wiki/Optical_character_recognition.
[2] linguatec text to speech technology
[3] https://developers.google.com/vision/text_overview

★★★