QR Code Based Text To Speech Conversion

Deepak Moorty, Munazza Ruman, Nauman Irfan, Aishwarya K.R
Electronics and Communication Engineering, VVIET, Mysuru, Karnataka, India

***Abstract*- QR codes, developed by a Japanese company, have been around for over fifteen years. With the advent of smart and Web capable mobile devices, we witness a steady growth of interesting commercial applications using QR codes. The main objective of our project is to provide people with a technology that can enable them to obtain information about anything by simply scanning the QR code. The scanning result in obtaining the text information of a particular specimen and this information is further converted into speech using specialized software.**

***Keywords*— Quick Response (QR) Code, Barcodes, Mobile Phone, Text to Speech Converter**

I. INTRODUCTION

A barcode is an optical machine-readable exemplification of data relating to the object to which it is committed. Primitively barcodes represented data by varying the widths and spacings of parallel lines, and may be referred to as linear or one-dimensional codes. Later they evolved into rectangles, dots, hexagons and other geometric patterns in two dimensions. QR code stands for Quick Response Code, which is the trademark for the type of matrix barcode which was invented by the Japanese corporation Denso Wave. QR code has a number of features such as large capacity data encoding, dirt and damage resistant, high speed reading, small print out size, 360 degree reading and structural flexibility of application.

1.1 Motivation

In certain situation when a person goes on a trip to a tourist place, say a cultural heritage, then in such situation, the person has to take help of a tourist guide to get to know about the cultural aspects of that place. This sometimes results in overcrowding of people around the crowd. Consequently, the person cannot complete his visit peacefully. Moreover, if the person wants to explore the place by himself/herself, then he/she has to buy a guide and has to read out the information contained in it. This will in turn prove to be time consuming. Hence, we have come up with an idea that textual information can be best conveyed orally.

1.2 Objective

The main objective of this project is to provide the person with the independence of exploring things by giving a means to access information at his/her fingertips. Our project basically deals with the conversion of a piece of text information into a QR code and back to text in a smartphone. This text is further converted into speech. Let us assume that information regarding any historic place say, Mysore Palace is encoded into a QR code. This QR code is scanned by a tourist using a smartphone with a QR code scanner and is converted back into text in the smartphone. Now the user gets the information about Mysore Palace in his phone. Then using specialized software, the text is converted into speech (basically electronic speech). Now the tourist can listen to the information of the palace using earphones, plugged into his/her ears and continue exploring other sites of the historic place as well. Hence, within a short span of time, the user has all information at his/her disposal. The user can also read the text information.

1.3 Barcode v/s Qr-Code

Barcodes and QR codes are similar in the fact that they are data collection technologies, meaning they automate the process of collecting data. However they also differ significantly in many areas.

Attribute	Barcode	QR code
Line of sight	Required	Required
Read range	Several inches to feet	Several inches to feet
Identification	Most barcodes identify only one type of item	QR codes can identify items uniquely
Information Capacity	Very less	Less
Reliability	Wrinkled and smeared tags won‘t work	Wrinkled tags may work 30% data recoverable
Data capacity	<20 characters	Upto 7,089 characters
Orientation Dependent	Yes	No
Marginal Cost	0.01$	0.05$

Table 1: Comparison between QR code and Barcode

1.4 Understanding QR Code

QR codes have already overtaken the popularity of classical barcode in many areas because of several advantages like increase in capacity, reduced size, etc. the variety of applications it can offer makes the use of QR code more prominent than that of the barcodes. Statistically, QR codes can accommodate the same of data in approximately one tenth the space of a traditional barcode. Information such as URL, SMS, contact information and plain text can be embedded into the two dimensional matrix. Also, the explosive increment in the use of smartphones has played an important role in the popularizing of QR codes.

1.4.1 Architecture and Encoding

QR code is a two dimensional matrix with a cell architecture arranged in a square. Figure 1 shows the QR code architecture. QR codes include different areas that are reserved for specific purposes. Finder, separator, timing patterns and alignment patterns are the function patterns. Function patterns are not used for the data encoding. The finder patterns are located at the three corners of the symbol and are intended to assist in easy location of its position, size and inclination.

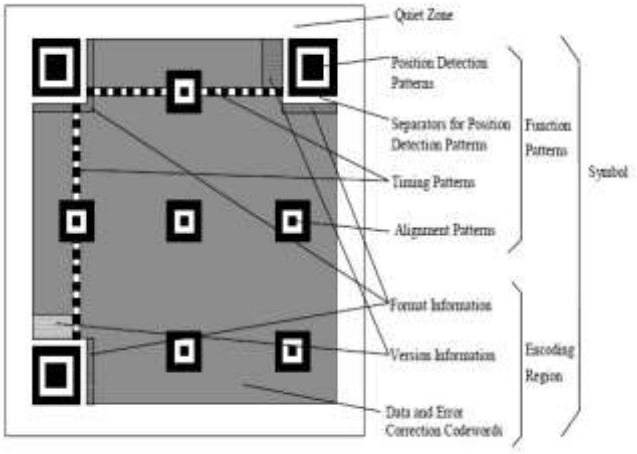


Figure 1: QR code architecture.

The procedure used for encoding of the QR Code includes the following steps: Initially the input data is encoded according to most efficient mode and bit streams are obtained. The bit streams are then divided into codewords. Further, the codewords are divided into blocks, and error correction is performed. All these codewords are put into a matrix and are masked with mask pattern. Finally function patterns are added into the QR symbol. And thus, we get a QR Code symbol.

1.4.2 Basic characteristics

1) Encodable character set:

1.	Numeric data (Digits 0-9)
2.	Alphanumeric data (Digits 0 - 9; upper case letters A -Z; nine other characters: space, $ % * + - . / :)
3.	8-bit byte data
4.	Kanji characters

Table 2: Encodable character set

2) Representation of data: A dark module is a binary one and a light module is a binary zero.

3) Symbol size (not including quiet zone): 21 modules to 177 modules. Versions 1 to 40, increasing in steps of 4 modules per side.

4) Data characters per symbol: Maximum allowable data capacity for maximum symbol size version 40 and minimum error correcting level L.

1.	Numeric data	7,089 characters
2.	Alphanumeric data	4,296 characters
3.	8-bit byte data	2,953 characters
4.	Kanji data	1,817 characters

Table 3: Data characters allowed per symbol

1.5 Text To Speech Converter

Speech synthesis is the artificial production of human speech. A computer system used for this purpose is called a speech synthesizer, and can be implemented in software or hardware. A text-to-speech (TTS) system converts normal language text into speech; other systems render symbolic linguistic representations like phonetic transcriptions into speech. Text-to-speech (TTS) convention transforms linguistic information stored as data or text into speech. It is widely used in audio reading devices for blind people in recent times. In the last few years however, the use of text-to-speech conversion technology has grown far beyond the disabled community to become a major adjunct to the rapidly growing use of digital voice storage for voice mail and voice response systems. Also developments in Speech synthesis technology for various languages have already taken place.

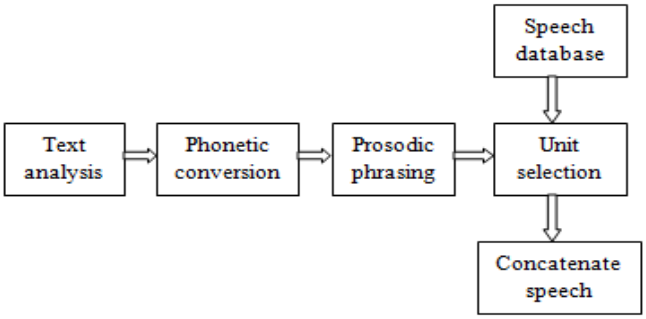


Figure 2: Block diagram of Text to Speech conversion

A text-to-speech system (or "engine") is composed of two parts: a front-end and a back-end. The front-end has two major tasks. First, it converts raw text containing symbols like numbers and abbreviations into the equivalent of written-out words. This process is often called text normalization, pre- processing, or tokenization. The front-end then assigns phonetic transcriptions to each word, and divides and marks the text into prosodic units, like phrases, clauses, and sentences. The process of assigning phonetic transcriptions to words is called text-to-phoneme or grapheme-to-phoneme conversion. Phonetic transcriptions and prosody information together make up the symbolic linguistic representation that is output by the front-end. The back-end—often referred to as

the synthesizer—then converts the symbolic linguistic representation into sound.

II. METHODOLOGY

The methodology has two stages namely, generation of QR code and obtaining information from it by scanning it and secondly, conversion of the information into electronic speech.

2.1 Generating Codes

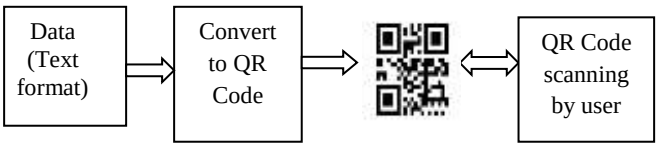
QR codes can be created for free using any number of web-based forms. It is as simple as copying and pasting the desired link URL into a text box and hitting "enter." You may have noticed that goo.gl even generates QR codes on the display page; this image can be saved to your computer and used. However, goog.gl only generates very small codes that are unsuitable for most printed materials. For that reason we used the QR code generator located on the Kaywa website, a site which provides services for the mobile web. It is located at http://qrcode.kaywa.com/.

Following steps were carried out in the generation of the QR code in Kaywa website:

1. Copy the short code for the content that you generated with goo.gl
2. Open http://qrcode.kaywa.com/ and paste the address in the field labeled URL.
3. Select the preferred size of the Qr code that you want to generate and click on "generate" option. The QR code will appear in the left hand side of the screen.
4. The generated code is saved by doing a right click on it using the mouse and saving it on the computer.

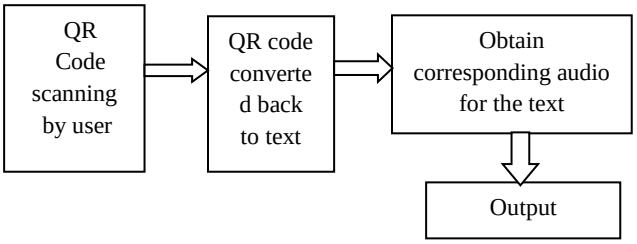
2.2 Block Diagram

Stage 1



Firstly, a QR code is generated for the given information and the QR code is stuck on the billboard. This stage is completed with the user scanning the QR code. Here the QR code is generated using Matlab.

Stage 2



Stage two is a continuation of stage one where the user scans the QR code. Further, the corresponding embedded text is obtained. Simultaneously, corresponding electronic audio is also generated. The text information is available to the user's disposal which can either be read or can be heard using speakers.

III. SPEECH SYNTHESIZER

Text-to-speech synthesis takes place in several steps. The TTS systems get 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.

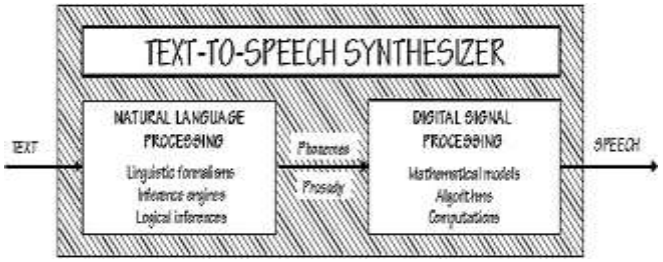


Figure 3: Speech synthesizer

The structure of the text-to-speech synthesizer can be broken down into major modules:

Natural Language Processing (NLP) module:

The NLP module takes a series of text input and produces a phonetic transcription together with the desired intonation and prosody (rhythm) that is ready to pass on the DSP module. The NLP module is composed of three major components: text analyzer, letter-to-sound (LTS), and prosody generator. Text analysis is a language-dependent component in TTS system. It is invoked to analyze the input text.

The major operations of the NLP module are as follows:

Text Analysis:

First the text is segmented into tokens. The token-to-word conversion creates the orthographic form of the token. For the token "Mr." the orthographic form "Mister" is formed by expansion, the token "12" gets the orthographic form "twelve" and "1997" is transformed to "nineteen ninety seven".

Application of Pronunciation Rules:

After the text analysis has been completed, pronunciation rules can be applied. Letters cannot be transformed 1:1 into phonemes because correspondence is not always parallel. In certain environments, a single letter can correspond to either no phoneme (for example, "h" in "caught") or several phonemes.

Digital Signal Processing (DSP) module:

It transforms the symbolic information it receives from NLP into audible and intelligible speech. The operations involved in the DSP module are the computer analogue of dynamically controlling the articulartory muscles and the vibratory frequency of the vocal folds so that the output signal matches the input requirements. This can be basically achieved in two ways: Explicitly, in the form of a series of rules which formally describe the influence of phonemes on one another and Implicitly, by storing examples of phonetic transitions and co-articulations into a speech segment database, and using them just as they are, as ultimate acoustic units .

Figure 4: TTS module

Pin No.	Pin Name	Function
1	GND	Ground pin
2	+5V	Provide regulated +5V DC voltage to this pin for board operation. Board has LM1117-3.3V regulator to power all parts.
3	TX-OUT	Outgoing Serial data for 3-5V UART 9600bps. Typically connected to RXD pin of microcontroller or MAX232 or USB-TTL. It outputs text that the module is speaking.
4	RX-IN	Incoming serial data at 5V or 3V level. If your data source is microcontroller, then you can connect its TXD pin to RX-IN of the board directly. If your data is at RS232 level data is use a MAX232 to convert that data to 5V UART suitable for board. It has to be 9600bps - No Parity - 1 stop bit Maximum string it can accept is 256 characters at a time. Once it's done speaking you can send another string. You can monitor its TX-OUT pin for hash # character to see when it has completed the talking and then you can send another string.

Table no.4: Interface/Power Pins

IV. EXPECTED OUTCOME

Our project tries to accomplish some of the important constraints

1. Our project distinguishes it from the RFID, giving it a new idea and concept.
2. We are able to convert the QR Code into a text message.
3. Easy information available by scanning the QR Code.
4. Text message is later converted into voice message.
5. Both text and voice messages are achieved.

V. ADVANTAGES

1. Fast and error-free process.
2. User can easily save the data on their smartphones
3. QR codes are 2 dimensional codes, capable of storing data horizontally and vertically but in bar codes they are one dimensional numeric codes.
4. Can carry up to some hundred times the amount of information when compared to conventional bar codes.
5. Can be read in 360 degrees from any direction.

VI. APPLICATIONS

1. Conveying useful information regarding any event in an educational institute.
2. Displaying a QR code on a bill board in hospitals can help in directing people to the wards or the rooms.
3. QR codes in library can help finding the location of a particular book.
4. Tagging an antique or an artifact with QR code in a Museum will help in getting the information of the object.
5. Class notes can be encrypted in QR codes so students can pay more attention during the lecture.

REFERENCES

[1]. C. Law and S. So, "QR Codes in Education", Journal Of Educational Technology Development and Exchange, Vol. 3, No. 1, PP. (85-100), (2010).
[2]. "Recognition of QR code by Mobile phones", Yue Liu, Ju Yang and Mingjun Liu 2008 Chinese Control and Decision Conference, 2008
[3]. "Design and Implementation of Text To Speech Conversion for Visually Impaired People", Itunuoluwa Isewon, Jelili Oyelade and Olufunke Oladipupo.
[4]. "QR Code Technologies in Museum Exhibitions", Sandra Medić, Nataša Pavlović* Received: July 2014 |Accepted: September 2014
[5]. "Design and Generation of Speech Generation System Using Matlab" International Journal of Engineering and Innovative Technology (IJEIT) Volume 4, Issue 6, December 2014
[6]. A Survey on QR Codes: in context of Research and Application Kinjal H. Pandya, Hiren J. Galiyawala, International Journal of merging Technology and Advanced Engineering Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 4, Issue 3, March 2014)
[7]. "Evaluating the Use of Quick Response (QR) Code at Sulaimani University Libraries", Ako Muhammad Abdullah*, Roza Hikmat Hama Aziz
[8]. Implementation of Text to Speech Conversion, Chaw Su Thu Thu, Theingi Zin, Department of Electronic Engineering, Mandalay Technological University, Mandalay, International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181 Vol. 3 Issue 3, March - 2014
[9]. "QR code, Wikipedia", http://en.wikipedia.org/wiki/QR_code [Online] [Retrieved 2102-02-09]
[10]. H. Hanks, "Successful Scanning: A Guide to QR Code Best Practices", (2012), The Agency Inside, PP. (1-18), [Online] Available: http://www.*hartehanks*.com/static/ base/ *qr code*-wp.pdf