A Distinctive Approach for Translating Text from Quick Response Code into Gujarati Speech

Dr Dikshan N Shah[1] **Dr Nirali C Dave[2]** **Dr Hemangini S Patel[3]**

Assistant Professor[1][2][3], Vanita Vishram Women's University,
Athwagate, Surat – 395 001, Gujarat, India

Abstract

Objectives : Text-to-Speech is available in many Indian languages. Text/Information extracted, converted into native language speech from QR code printed in Food packet wrapper.

Methods : As this is a novel approach, there is no such dataset is available. In this research, first step is to create dataset. Food products which have printed QR Code contains details such as manufacturing date, batch no, expiry date or best before etc.. By suggesting a novel approach by using Machine Learning algorithm, text will be extracted from QR code which then converted to native language of Gujarat state i.e. Gujarati speech.

Findings: Using QR codes for daily soap items, this study proposes a revolutionary way to extract Text-to-Speech from a printed QR code. A QR Code will be scanned for translating text into Gujarati voice by using speech synthesizer that can be heard by blind or visually impaired person. The text is subsequently transformed into Gujarati and produced as speech. This proposed research will be valuable for persons who are blind, either completely or partially, researchers, manufacturers, and customers.

Novelty : The proposed research develop a novel approach for extracting information printed in QR Code and translate into native language speech i.e. in Gujarati language.

Keywords *– QR Code, Text to Speech, Novel approach, Native language, Indian Language*

1. Introduction

Visual impairments are divided into four categories: 1) By birth; 2) Because of age; and 3) By accident 4) Because of the corporate culture. Around 253 million people worldwide suffer from partial or total visual impairment. According to the World Health Organization's (WHO) 2021 census, over 36 million people are blind, while 217 million are moderately impaired. According to globe Health Organization estimates, contaminated food kills 420000 people per year, causes 600 million cases of foodborne disease, and costs the globe 33 million healthy years. [1] The visually challenged cannot identify this sort of foodborne sickness, which is caused by tainted or expired food on the market. Retailers sell expiry dates over food, and the visually handicapped cannot read the best-before date.

The QR Code, or Quick Response Code, was created in 1994 by Toyota's Japanese subsidiary Denso Wave. The ISO/IEC 18004 industrial standard defines QR codes. [2] Originally designed to transmit data, barcodes were one-dimensional or linear codes with different parallel line widths and spacings. Then, several two-dimensional geometric patterns, such as dots, hexagons, and rectangles, were created from these lines.[3] This two-dimensional matrix can also be used to store data such as plain text, SMS, batch numbers, production and expiration dates, emails, movies, and URLs. The differences between the barcode and QR code listed in Table 1 are as follows.

Table 1. Comparison between Barcode and QR Code

Barcode	QR Code
Barcode was developed in 1952.	The QR Code was developed.in 1994.
The inventor of the barcode was Norman Joseph Woodland.	Masahiro Hara invented the QR code.
Barcode Can be 1D (linear) or 2D.	Only 2D QR codes are possible.
Barcodes allow data to be stored in a printed format and retrieved through scanning.	A kind of 2D barcode that is scannable to retrieve data.
Barcode Used for rental car monitoring, airline baggage tracking, hospital patient tracking, retail item tracking, etc.	QR-Code used to share contacts, pictures, movies, and other documents; also used for cashless payments.
Information is stored horizontally in barcodes.	QR Codes store data both vertically and horizontally.
Usually, a vertical arrangement of parallel lines makes up a barcode.	A QR code is made up of square dots arranged in square grids.
Compared to QR codes, barcodes store less data.	QR codes stores multimedia data in addition to more information than barcodes.

1.1 Structure of QR Code

The 2D matrix that makes up the QR code has square-shaped cell layout. Figure – 1 below depict the terms "finder," "separator," "alignment patterns" and "timing patterns," refer to areas within QR codes that are reserved for particular purposes. Located in each of the three corners of the symbol, the finder patterns are designed to make it easier to determine the position, size, and inclination of the symbol.[2]

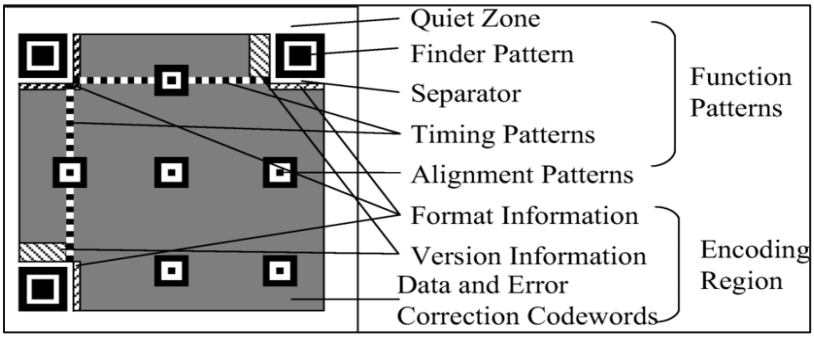


Figure 1: Structure of 2D matrix QR Code

The following varied information about a square-shaped QR code was shown in figure 1 above.

- **Finder Pattern:** The finder pattern, which appears in all but the bottom right corner of the QR code, consists of three identical structures. Every pattern starts with a black 3x3 matrix module surrounded by white modules, which are surrounded by more black modules. The decoding program finds the QR code and determines its correct orientation more easily.
- **Separators:** The one-pixel-wide white separators aid in increasing the finder patterns' recognizability by helping to separate them from the real data.
- **Timing Pattern:** By alternating black and white modules in the timing pattern, the decoder software can determine the width of a module.
- **Alignment Patterns:** Alignment Patterns help the decoder software compensate for small amounts of visual distortion. The larger the code, the more alignment patterns there are.
- **Format Information:** The selection of the masking pattern and the error correction level of the QR code are described in the format information section, which is made up of 15 bits adjacent to the separators.
- **Data:** The data portion converts the data into a bit stream and saves the data in 8 bit chunks called codewords. [5]

- **Error Correction:** Similar to data codes, error correction codes are stored in the error correction area in codewords that are eight bits long. [5]
- **Remainder Bits:** If the data and error correction bits cannot be uniformly divided into 8-bit codewords, it is composed of empty bits.

2 Methodology

The methodology of this approach is divided into sections such as 1) Corpus Development 2) Scan QR images 3) Pre-Processing of QR images, 4) Image Conversion, and 5) Speech Synthesizer.

2.1 Corpus Development

In the course of collecting data, we have taken sample QR code images for a variety of daily soap products, including packed fruits, wafers, crackers, and tetra packs of cold drinks, milk etc. Numerous distinct product categories exist as mentioned in figure - 2. In order to have 20 QR codes for a training set, we have chosen food-related ones.

Figure 2: Sample Corpus Data

For training reasons, a corpus of variably categorized product wrappers—including those for dairy products, beverages, nuts and seeds, soft drinks, etc.—has been collected.

2.2 Scan QR

A mobile smartphone or another QR image reader will be used to read QR codes from product packaging as its image input. The QR image is a two-dimensional matrix barcode since the data is kept in a matrix format. The most often utilized analytical graphics modules in two-dimensional matrix barcodes are hexagonal, square, and circular shapes.

2.3 Pre-Processing of QR Images

The Pre-Processing task of the QR Code includes the following steps: 1) Bit streams are created once the input data has been encoded using the most effective mode. 2) Codewords are then created from the bit streams. 3) The codewords are also broken up into blocks, and error checking is done. 4) These codewords are all compiled into a matrix and covered in a masking pattern. And 5) The QR symbol now includes function patterns as a final addition. Consequently, a QR code symbol appears with character set.

2.3.1 Encodable Character set

As Table 1 illustrates, The following encodable characters are utilized to construct a bit stream after the QR code has been input.

Table – 2 List of Encodable Characters

Sr No	Characters
1	(0-9) Numeric Data
2	(0 – 9 Digits, Upper case letters A-Z, other nine characters such as space, $, %, *, +, -, ., /, :,) -- Alphanumeric data
3	8-bits (Byte) Data
4	Kanji Characters

2.3.2 Data Representation: One module, called Dark, shows a binary one, and the other, called Light, shows a binary zero.

2.3.3 Symbol size: Versions 1 through 40 and a range of 21 to 177 modules are available, with increases of 4 modules each per side.

2.3.4 Data characters per Symbol: For every symbol, there is a maximum data capacity of 7089 characters for numeric data, 4296 characters for alphanumeric data, 2953 characters for 8-bits data, and 1817 characters for kanji data. These values are allowed for maximum symbol size version 40 and minimum error correcting level L.

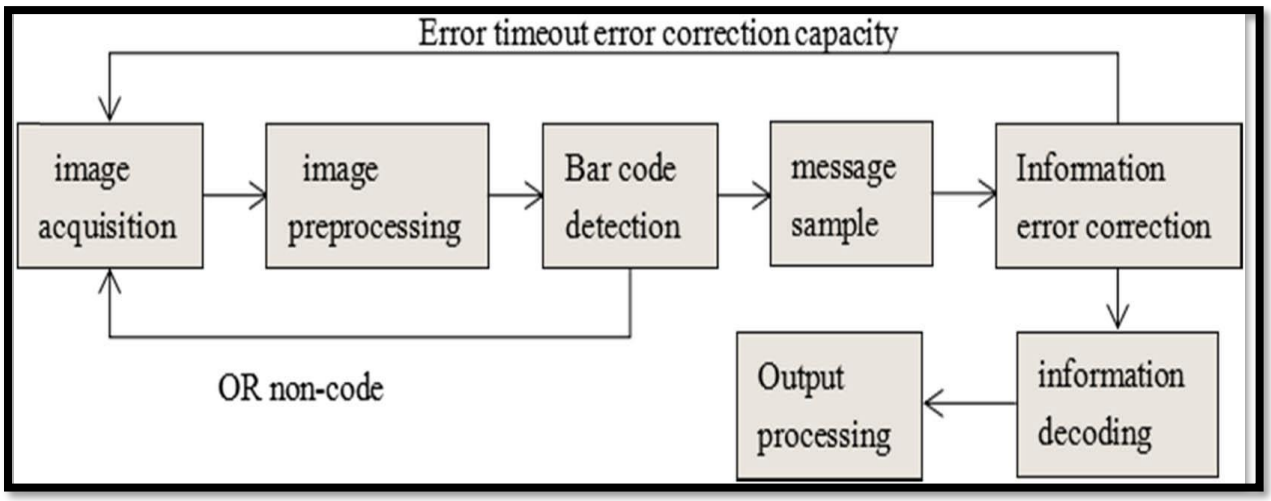


Figure – 4 Preprocessing steps for QR Code image

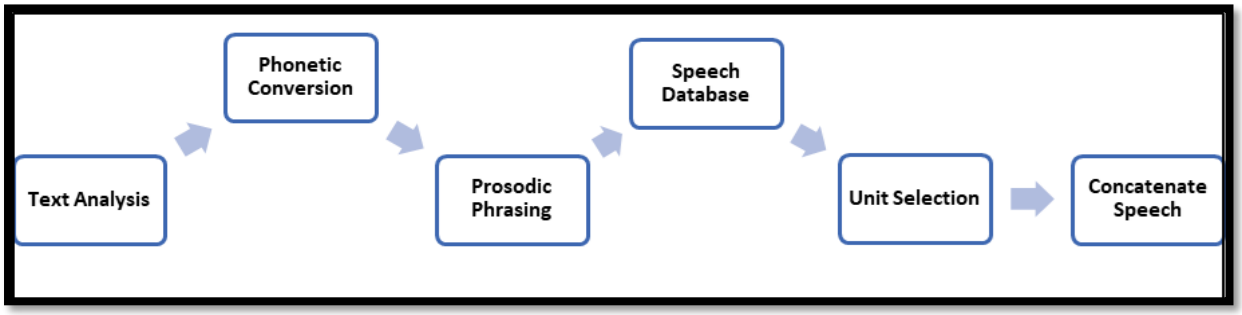
In order to identify scanned QR codes, the above figure - 4 described the stages involved in preprocessing the acquired pictures, barcode identification, information sampling, information decoding, information error correction, and results compilation.

2.3.5 Image Conversion

After the image has been preprocessed, machine learning and neural network algorithms can be used to decode the QR picture. By decoding the QR image, its content can be viewed. The details contained in QR codes include use-by or expiration dates, advised dietary allowances, lot or batch numbers, vegetarian food, date of manufacturing, date of packaging, and other information. [5]

2.3.6 Speech Synthesizer

Two approaches use to synthesize speech are as : 1) During the text-to-speech process, text analysis first removes the noise from the text and then divides the text into chunks and sends them to the Natural Language Processing (NLP) module for phonetic transcription where various Gujarati language specific handcrafted rules will be applied to identify words and entities[3]. Identified words will be then processed for signal generation for speech. 2) A statistical Hidden Markov Model, can also be used in speech recognition as speech signals can be viewed as a part of a stationary signals. By using Viterbi algorithm in decoding, we can find the state sequence which maximizes the probability of observation sequence. The decoded text input will be fed into the voice synthesizer, which will convert it into speech, as seen in Figure 5.



[Figure – 4 Block diagram for Text-to-Speech Conversion]

After decoding the observation sequences, adjust model parameters to maximize the probability of observed sequences using Forward-Backward algorithm. Emission Probability, State Probability and Transition probability is generated. The NLP module includes a prosody generator, LTS, and morphological text analyzer for language-dependent analysis.[5] The Digital Signal Processing module interprets symbolic data sent by the Natural Language Processing module into understandable speech that may be heard by those with partially or fully vision impairments.[6]

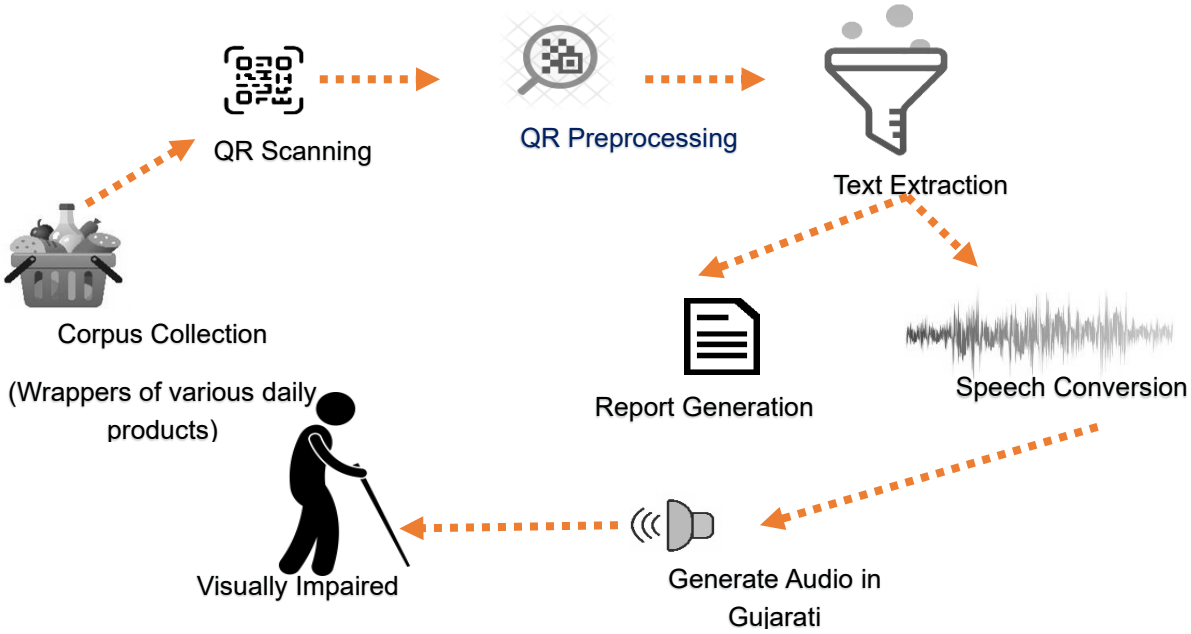


Figure 5: Text-to-Speech using QR code

The above figure – 3 depicts the stages for the graphical text-to-speech technology employing QR codes. The list of hardware required to carry out this technique is indicated in Table - 3.

Table – 3 Hardware Tools required with usage for Implementation

Sr No	Hardware Details	Hardware used for experimentation
1	Premium-grade material eyewear	high-quality polycarbonate lightweight, strong and corrosion-resistance eyeglasses used.
2	Raspberry Pi 5 Model 8GB	Raspberry Pi 5 Model 8GB offers extensive computing power in a compact form with 8GB of RAM.
3	27W USB-C PD Power Supply for Raspberry Pi 5 – Black	Provides stable and effectual power delivery to Raspberry Pi with the 27W USB-C PD Power Supply.
4	Sandisk 32 Gb Memory Card / Official Raspberry Pi Micro SD Card 32GB	Use to host the operating system, store data, and run various applications.
5	Micro-HDMI (Male) to Standard HDMI (Male) Cable for Raspberry Pi	To connect Raspberry Pi 5 with the micro-HDMI to standard HDMI cable.
6	3.5mm Audio Jack	Used to connect a pair of headphones with mobile phone.
7	Slide Switch	An electrical toggle switch which operates by a sliding lever or actuator.
8	Micro QR Reader / HC-SR04 - Ultrasonic Sensor	Used to measure non-contact distance.

3 Results and Discussion

In this research, we have discussed about QR codes—their invention, uses, processing advantages, disadvantages and its applications. The QR Code may be instantaneously scanned by anyone to get information. In this research we extract information from QR code and turn into a text message. [7] After generating text messages, it is converted into speech. Converted speech can be heard as audio message in their native language i.e. in Gujarati. [8] As it paves the way for clean meals, this approach will enhance human health for those with low vision, visual impairment, or other disorders. There will undoubtedly be more opportunities for research in implementing with Artificial Intelligence

and to look into the security of employing QR codes for businesses and their customers, even though they are still updated occasionally.

4 Conclusion

By inventing a revolutionary method for collecting information from QR codes and converting it into text messages that can then be translated into human audible speech in their native language i.e. in Gujarati. We can implement more languages with improved accuracy and data storage.

References

[1] National Blindness And Visual Impairment Survey 2015-2019 Available from: https://npcbvi.mohfw.gov.in/writeReadData/mainlinkFile/File341.pdf

[2] Mathivanan P, Ganesh AB, Venkatesan R. QR code–based ECG signal encryption/decryption algorithm. Cryptologia. 2019 May 4;43(3):233-53. Available from : https://doi.org/10.1080/01611194.2018.1549122.

[3] Shah DN, Bhadka H. Noise Removal As Pre Processing Task And Its Implementation For Gujarati Named Entity Recognition. In ICT for Competitive Strategies 2020 May 5 (pp. 275-282). CRC Press. Available from : https://doi.org/10.1201/9781003052098.

[4] Harish, B.S., Rangan, R.K. (2020). A comprehensive survey on Indian regional language processing. SN Applied Sciences.2. 1204. https://doi.org/10.1007/s42452-020-2983-x

[5] Huo L, Zhu J, Singh PK, Pavlovich PA. Research on QR image code recognition system based on artificial intelligence algorithm. Journal of Intelligent Systems. 2021 Jul 16;30(1):855-67. Available from : https://doi.org/10.1515/jisys-2020-0143

[6] N. Aasofwala, S. Verma and K. Patel, "NLP based model to convert English speech to Gujarati text for deaf & dumb people," *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Delhi, India, 2023, pp. 1-6, doi: 10.1109/ICCCNT56998.2023.10308284.

[7] Narvani, V., Arolkar, H. (2024). Text-to-Speech Conversion Using Concatenative Approach for Gujarati Language. In: Senjyu, T., So–In, C., Joshi, A. (eds) Smart Trends in Computing and Communications. SmartCom 2024. Lecture Notes in Networks and Systems, vol 949. Springer, Singapore. https://doi.org/10.1007/978-981-97-1313-4_16.

[8] Dua, M., Bhagat, B., Dua, S. et al. A review on Gujarati language based automatic speech recognition (ASR) systems. Int J Speech Technol 27, 133–156 (2024). https://doi.org/10.1007/s10772-024-10087-8.