

# QR 코드에 음성 데이터 삽입을 위한 AMR 압축 비트열 분석

오은주 · 조현지 · 정현아 · 배정은 · 유 훈\*

상명대학교

## Analysis of AMR Compressed Bit Stream for Insertion of Voice Data in QR Code

Eun-ju Oh · Hyun-ji Cho · Hyeon-ah Jung · Joung-eun Bae · Hoon Yoo\*

SangMyung University

E-mail : hunie@smu.ac.kr

### 요 약

본 논문은 음성 데이터를 QR 코드에 입력 및 전송하는 기법을 연구하기 위해 실생활에 가장 많이 사용되는 AMR 음성 데이터를 분석한 결과를 제공한다. AMR은 HEADER와 Speech Data로 구성되어 있고, 비트 형식으로 전송되고 있으며 총 8개의 비트 전송률 모드를 갖고 있다. HEADER에는 Speech Data의 모드 정보가 포함되어 있으며 모드에 따라 Speech Data의 길이는 달라진다. 그 중 QR 코드에 삽입하기 가장 적절한 전송률 모드를 선택하고 해당 모드에 대한 분석을 제공한다. 각 모드에 대한 분석 및 실험을 통해 추후 음성 데이터에 대해 더 높은 압축률을 보이는 것이 최종 목표이다. 그럼으로써 음성 데이터를 보다 효율적으로 전송할 수 있다는 점에서 성능 개선을 보인다.

### ABSTRACT

This paper presents an analysis of the AMR speech data as a basic work to study the technique of inputting and transmitting AMR voice data which is widely used in the public cell phone. AMR consists of HEADER and Speech Data, and it is transmitted in bit format and has 8 bit-rate modes in total. HEADER contains mode information of Speech Data, and the length of Speech Data differs depending on the mode. We chose the best mode which is best to input into QR code and analyzed that mode. It is a goal to show a higher compression ratio for voice data by the analysis and experiments. This analysis shows improvement in that it can transmit voice data more effectively.

### 키워드

AMR codec, Voice Compression, QR code, Bit Stream, Speech Data

### 1. 서 론

AMR(Adaptive Multi-Rate)은 음성 부호화에 최적화된 오디오 데이터 압축 시스템으로, 다양한 전송률을 지원하여 유동적인 환경에서 유연하게 대처할 수 있다는 점에서 많은 통화 기지국들이 사용하는 국제 표준 코덱이다[1]. 총 8개의 가변 전송률을 가지고 있어 상황에 따라 가변적으로 모드를 선택하여 음성 데이터를 전송할 수 있다.

또한, 정보의 양이 많아짐에 따라 해당 정보를 담을 수 있는 공간 확보가 중요해진 만큼, QR코드의 등장은 기존의 1차원 바코드보다 훨씬 많은 데이터를 저장할 수 있다는 점에서 큰 강점을 띄고

있다[2].

본 논문은 QR 코드에 음성 데이터를 삽입하여 전송하기 위해 대중적으로 많이 사용되는 AMR 음성 코덱에 대한 분석을 제공한다. 또한, QR 코드에 음성 데이터를 삽입하기에 가장 적합한 가변 전송률을 선택하여 그에 대한 포맷 분석을 제공한다.

제안되는 분석은 추후 음성 데이터에 대해 더 높은 압축률을 보이기 위한 전초 단계로서, 음성 데이터를 보다 적은 비용으로 효율적으로 전송할 수 있다는 점에서 성능 개선을 보인다.

\* corresponding author

## II. 기본 개념

### 2.1 AMR codec 기본 개념

통상적으로 AMR codec에 사용되는 입력값은 16bit로 패킹된 13비트의 PCM데이터를 입력 데이터로 취급하고 이 입력 데이터를 전송률 4.75, 5.15, 5.90, 6.70, 7.40, 7.95, 10.2, 12.2 kbit/s로 부호화 하여 각각 95, 103, 118, 134, 148, 159, 204, 244개의 비트가 포함된 음성 신호를 송출한다[1].

위와 같은 과정을 거치고 난 후 출력되는 AMR 코덱의 구조는 다음과 같다.

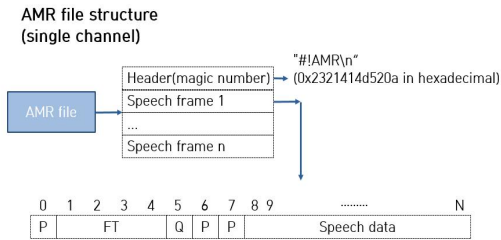


그림 1. AMR codec file 기본 구조

그림 1을 보면, AMR은 맨 앞에 magic number라 불리는 헤더 정보와 speech frame의 연속으로 이루어진다[3]. speech frame에는 해당 스피치 데이터에 대한 모드 정보와 오류없이 인코딩이 되었는지에 대한 정보가 담겨있다. Q는 해당 프레임이 아무 손상없이 인코딩이 되었을 때 1의 값을 리턴한다. FT에는 해당 프레임에 대한 모드 정보가 담겨 있는데 해당 모드에 대한 코드는 아래 표 1과 같다.

표 1. 모드별 비트 전송률

Code	Mode	Bit Stream Rate
0 (0000)	MR475	4.75 kbit/s
1 (0001)	MR515	5.15 kbit/s
2 (0010)	MR59	5.90 kbit/s
3 (0011)	MR67	6.70 kbit/s
4 (0100)	MR74	7.40 kbit/s
5 (0101)	MR795	7.95 kbit/s
6 (0110)	MR102	10.20 kbit/s
7 (0111)	MR122	12.20 kbit/s
8 (1000)	SID	comfort noise
15(1111)	no mode	

표 1과 같이 모드 별로 비트 전송률이 차이가 나는 것을 확인할 수 있다.

그 중, 본 논문에서는 가장 높은 압축률의 mode0에 대하여 분석하고 QR코드에 음성 데이터를 삽입하기 위한 포맷을 설정한다[4].

AMR 포맷에서 헤더, 모드 정보, parity를 제거 후 나타나는 비트의 평균값에 대한 그래프를 제공한다.

## III. 실험 및 분석

본 실험에서는 샘플링 주파수를 8kHz로 설정하고 부호화에 사용되는 입력 데이터는 20ms를 단위로 한 프레임당 160개의 샘플을 취급해 1분 단위로 데이터를 설정하여 총 3000프레임의 데이터를 기준으로 실험을 제공한다. 20개의 실험 데이터를 통해 AMR 코덱 전송률 중 출력 비트 4.75kbit/s로 가장 높은 압축률을 보여주는 mode0으로 부호화된 AMR 평균 비트 스트림은 다음과 같은 그래프로 나타난다.

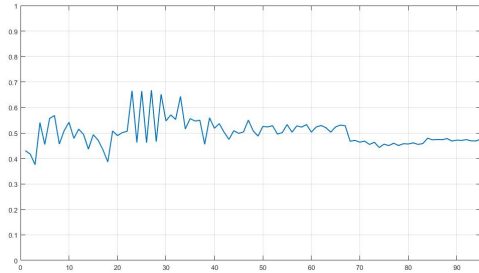


그림 2. mode 0에 대한 보이스 평균 비트열 그래프

그림 1을 보면 초반 23 비트까지 각각 8,8,7 비트에 LSF submatrix의 지수에 대한 내용이 할당된다. 뒷부분의 subframe에는 8bit의 adaptive codebook의 지수가 할당되고 9비트의 Algebraic codebook이 할당된다. 주로 앞부분에 높은 평균값을 보여준다는 점에서 해당 구간에는 엔트로피가 있으며 이 말은 곧 편차가 발생하는 구간은 보다 더 많은 압축이 가능하다는 것을 암시한다. 즉, 모드 간 비트 분석을 통해 AMR 코덱에서 제공하는 압축률 보다 더 높은 압축이 가능하다는 것을 의미한다.

## IV. 결론

본 논문은 QR코드에 음성 데이터를 삽입하기 위해 음성 데이터 압축 포맷 중 하나인 AMR 코덱 비트열에 대한 분석을 제공한다. 8가지의 비트 전송률 중 압축률이 가장 높은 모드(mode0)를 선택하여 해당 모드에 대한 분석 및 비트열 특징을 찾아내 QR코드에 삽입하기 적절한 포맷으로 변경한다. 위 실험을 통해 편차가 발생하는 구간은 압축이 더 가능하다는 것을 발견할 수 있었고 이를 통해 음성 데이터를 기존의 방법보다 저비용이지만 보다 효과적으로 전송할 수 있다는 점에서 성능의 개선을 보인다.

추후 이어지는 연구에서는 각 음성 데이터별 특징을 찾아내어 성별, 연령대별로 더 높은 압축을 이루게 하는 것이 최종 목표이다.

## Acknowledgments

This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government(MSIP)(No.2017-0-00515, Development of integraphy content generation technique for N-dimensional barcode application)

## References

- [1] “Adaptive Multi-Rate(AMR) speech codec; Transcoding functions,” *3GPP TS 26.090*, v14.0.0(2017-03)
- [2] Y. Gu, and W. Zhang, “QR code recognition based on image processing,” *Information Science and Technology, International Conference on IEEE*, Nanjing; China, pp. 733-736, May. 2011.
- [3] J. Sjoberg, M. Westerlund, A. Lakaniemi, and Q. Xie, “RTP payload format and file storage format for the Adaptive Multi-Rate (AMR) and Adaptive Multi-Rate Wideband (AMR-WB) audio codecs”, *No. RFC 4867*, April. 2007.
- [4] T. Lundberg, P. de Bruin, S. Bruhn, S. Hakansson, and S. Craig, “Adaptive thresholds for AMR codec mode selection”, *IEEE In Vehicular Technology Conference 2005. VTC 2005-Spring*, Stockholm; Sweden, Vol. 4, pp. 2325-2329, June, 2005.