

Context Tree Weighting을 이용한 AMR 음성 데이터 압축 성능 개선[☆]

Improvement of AMR Data Compression Using the Context Tree Weighting Method

이 은 수¹
Eun-su Lee

오 은 주¹
Eun-ju Oh

유 훈^{2*}
Hoon Yoo

요 약

본 논문은 Context Tree Weighting (CTW) 를 이용하여 Adaptive Multi-Rate (AMR) 데이터의 압축 성능을 개선하는 알고리즘을 제안한다. AMR은 IMT-2000에서 채택된 음성부호화 표준안으로써, 무선채널의 환경변화에 대처할 수 있도록 4.75 kbit/s 에서 12.2 kbit/s 까지 8가지의 전송률을 지원한다. CTW는 산술부호화기의 일종으로, 가변 차수 마르코프 모델을 사용하는 압축기이다. 우리는 CTW가 비트단위로 수행한다는 점을 고려하여 AMR 데이터를 변환한 후 CTW로 압축하는 알고리즘을 제안한다. 제안하는 알고리즘의 유효성을 검증하기 위하여 ZIP을 포함한 기존 압축방식과 제안된 알고리즘의 압축률을 비교하는 실험을 하였다. 실험 결과, AMR 데이터의 평균 추가 압축률이 ZIP의 경우 약 3.21%, 제안된 알고리즘의 경우 약 9.10%로 나타났다. 따라서 본 논문에서 제안한 알고리즘이 AMR 데이터의 압축 성능을 약 5.89% 개선하였다.

☞ 주제어 : AMR 음성 부호화기, Context Tree Weighting, 산술부호화기, 데이터 압축

ABSTRACT

This paper proposes an algorithm to improve the compression performance of the adaptive multi-rate (AMR) speech coding using the context tree weighting (CTW) method. AMR is the voice encoding standard adopted by IMT-2000, and supports 8 transmission rates from 4.75 kbit/s to 12.2 kbit/s to cope with changes in the channel condition. CTW as a kind of the arithmetic coding, uses a variable-order Markov model. Considering that CTW operates bit by bit, we propose an algorithm that re-orders AMR data and compresses them with CTW. To verify the validity of the proposed algorithm, an experiment is conducted to compare the proposed algorithm with existing compression methods including ZIP in terms of compression ratio. Experimental results indicate that the average additional compression rate in AMR data is about 3.21% with ZIP and about 9.10% with the proposed algorithm. Thus our algorithm improves the compression performance of AMR data by about 5.89%.

☞ keyword : AMR coding, Context Tree Weighting, Arithmetic coding, Data compression

1. 서 론

현대사회에서 요구되고 있는 음성 부호화 기술은 무선

채널의 환경변화에도 일정한 품질이 유지되어야 하고, 압축률이 높으며, 송수신 지연을 최소화해야 하고 따라서 간단한 알고리즘이어야 한다. 위의 모든 조건에 부합하는 음성 부호화 기술은 IMT-2000에서 표준기술로 채택한 Adaptive Multi-Rate (AMR) 음성 부호화기이다. 무선통신 분야는 기존의 디지털 셀룰러 및 PCS를 거쳐 IMT-2000으로 발전하였으며 음성서비스 품질 향상과 무선 멀티미디어 서비스의 제공을 목표로 표준화 작업을 시행하였다. 이 중 음성부호화 표준안으로 AMR이 채택되었다. AMR 음성 부호화기는 가변 비트레이트의 협대역 신호를 부호화하는 음성 코덱으로 Adaptive Multi-Rate Narrowband (AMR-NB) 로도 불린다. 현재 3G 환경에서 기본적으로 쓰이고 있는 음성 코덱이며 협대역 음성 코덱의 대표적인

¹ Department of Computer Science, Sangmyung University, Seoul, 03015, Korea.

² Department of Electronics, Sangmyung University, Seoul, 03015, Korea.

* Corresponding author (hunie@smu.ac.kr)

[Received 3 April 2020, Reviewed 23 April 2020(R2 5 June 2020), Accepted 3 July 2020]

☆ This work was supported by the research fund of Signal Intelligence Research Center supervised by Defense Acquisition Program Administration and Agency for Defense Development of Korea.

인 기술이다. 4G 및 5G에서 쓰이는 Adaptive Multi-Rate Wideband (AMR-WB)와 Extended Adaptive Multi-Rate Wideband (AMR-WB+)는 AMR-NB 인코딩 기술에 기반해서 개발된 광대역 음성 부호화이며 협대역 환경에선 여전히 AMR-NB가 사용된다[1,2]. AMR 음성부호화기는 무선채널의 환경변화에도 유연하게 대처할 수 있도록 4.75 kbit/s에서 12.2 kbit/s 사이의 다양한 전송률을 지원한다[3-6]. 따라서 전송 환경에 따라 최적의 모드를 적응적으로 선택함으로써 최적의 데이터 압축률과 명료한 음질 전송을 동시에 해낼 수 있다는 장점을 가지고 있다. 또한 전송률을 낮추어 전송의 효율성을 증가시키기 위해 VAD(Voice Activity Detector)[7], SCR(Source Controlled Rate)[8] 등의 기술을 포함하고 있다. AMR의 입력신호는 8 kHz의 샘플링 주파수에서 20 ms 당 160 개 음성 샘플에 해당하는 프레임 단위로 CELP 모델의 파라미터를 추출하기 위해 분석된다. 추출된 파라미터 정보는 전송 모드에 따라 비트 단위로 정해진 위치에 할당된다. 따라서 프레임간의 위치별 비트 상관도는 높을 것으로 예상할 수 있다.

Context Tree Weighting (CTW)는 가변 차수 마르코프 모델의 예측을 혼합하는 방식의 무손실 압축 및 예측 알고리즘이다[9,10]. 컨텍스트 모델을 바이너리 트리로 이용하여 확률을 추정하고 탐색한 모든 노드의 가중치를 효율적인 재귀방식으로 가중한다. 이렇게 부호화된 소스는 복잡성과 중복성을 낮추기 위해 산술 부호화(Arithmetic Coding)로 다시 한번 더 인코딩 된다. CTW는 미지의 모델 및 파라미터를 갖는 트리소스에 대해서도 효율적인 압축을 진행한다는 장점이 있다. 또한 바이너리 소스를 사용하기 때문에 비트 단위로 수행한다는 특징이 있다. 위에서 언급했듯이 AMR 음성부호화기의 경우 프레임간의 위치별 비트 상관도가 높을 것으로 예상되기 때문에 이 점을 이용하여 CTW 압축에 용이하도록 AMR 데이터를 변환한다면 압축률을 더 높일 수 있을 것이다.

본 논문에서는 음성부호화 표준안인 AMR 부호화 데이터가 제안된 알고리즘인 비트 재배치 및 CTW 압축에 의해 기존의 압축 방식보다 더 높은 추가 압축률을 얻을 수 있을지에 대해 논하고자 한다. 특히, 본 연구에서 제안하는 데이터 변환 방식이 추가 압축률을 얻는 점에 있어서 성능이 어떠한지 살펴보고자 한다.

본 논문 2장에서는 AMR 음성 부호화기의 구조 및 저장 방식에 대해 설명하고 3장에서 CTW에 대해 설명하며 4장에서 CTW를 이용한 AMR 음성 데이터 압축 성능 개선을 위한 알고리즘을 제안하고 5,6장에서 실험 결과 및

(표 1) AMR 음성부호화기의 전송 모드와 비트레이트
(Table 1) Transmit mode and bit-rate of AMR codec

Code	Mode	Bit Stream Rate
0 (0000)	MR 475	4.75 kbit/s
1 (0001)	MR 515	5.15 kbit/s
2 (0010)	MR 59	5.90 kbit/s
3 (0011)	MR 67	6.70 kbit/s
4 (0100)	MR 74	7.40 kbit/s
5 (0101)	MR 795	7.95 kbit/s
6 (0110)	MR 102	10.20 kbit/s
7 (0111)	MR 122	12.20 kbit/s
8 (1000)	MR SID	1.80 kbit/s

결론을 다룬다.

2. AMR 음성 부호화기

AMR 음성 부호화기의 구조는 8개의 전송모드를 갖는 multi-rate 음성부호화기와 음성과 비음성을 구분할 수 있는 정보를 나타내는 VAD (Voice Activity Detector), 비음성 구간에서 전송률을 낮추어 전송의 효율성을 증가시키기 위한 CNG(Comfort Noise Generation), SID(Silence Descriptor) 등의 시스템을 포함하는 SCR(Source Controlled Rate) 그리고 오류나 손실포켓의 영향을 줄이기 위한 error concealment mechanism[11] 등으로 이루어져 있다. 8가지 전송모드의 음성부호화기는 ACELP (Algebraic Codebook Excited Linear Prediction)에 기반을 두고 있고 다양한 전송률을 지원한다는 면에서 MR-ACELP(Multi-Rate ACELP) 라고 하기도 한다. 입력 데이터로는 13bits의 PCM 데이터를 다루는데 보통 종단을 0 으로 채워 16 bits 로 패키징된 데이터를 다룬다. 입력 데이터는 8 kHz의 샘플링 주파수에서 20ms 당 160 개 음성 샘플에 해당하는 프레임 단위로 다뤄지기 때문에 입력 비트레이트는 128 kbit/s 라 할 수 있다. 전송률은 4.75 kbit/s에서 12.2 kbit/s 사이의 8가지 source rate와 1.8 kbit/s의 comfort noise가 있으며 모드에 따른 전송률은 표 1과 같다.

AMR에서 8개의 전송모드는 갖는 multi-rate 음성부호화기는 기본적으로 CELP(Code Excited Linear Predictive) 모델에 기반을 두고 있다. 일반적으로 CELP 부호화 방식은 LP(Linear Prediction) 분석과 적응 코드북(adaptive codebook) 및 고정 코드북(fixed codebook)을 탐색하여 여 기신호를 추정하는 과정을 포함하고 있다[12]. 이 과정은

프레임상에서 동작하는데 각각의 프레임에서 스피치 신호는 CELP 모델의 파라미터 (LP 필터 계수, 적응 및 고정 코드북의 인덱스 및 이득 값 등)를 추출하기 위해 분석된다. 추출된 파라미터 정보는 전송 모드에 따라 프레임 당 95 bits (MR 475 모드)에서 244 bits (MR 122 모드) 사이로 할당된다.

(표 2) MR 475와 MR 122의 프레임 정보 비교
(Table 2) Frame information comparison between MR 475 and MR 122

Mode	Parameter	bits per frame	parity
MR 475	LSP set	23	1
	Pitch delay	20	
	Algebraic code	36	
	Gains	16	
	Total	95	
MR 122	2 LSP set	38	0
	Pitch delay	30	
	Pitch gain	16	
	Algebraic code	140	
	Codebook gains	20	
	Total	224	

표준코드로 부호화된 AMR 포맷의 음성 데이터는 맨 앞에 헤더와 그 뒤에 다수의 스피치 프레임으로 구성된다. 헤더는 그 뒤에 따라오는 스피치 프레임들의 모드가 단일모드인지 멀티모드인지에 따라 각각 다른 magic number를 나타내며 멀티모드인 경우 파일에 포함된 오디오 채널 수를 표시하는 32 bits 의 채널 설명 필드가 헤더에 포함된다. 헤더 뒤에 따라오는 다수의 스피치 프레임들은 8 bits 에 해당하는 프레임 헤더로 시작한다. 프레임 헤더의 2 ~ 5 번째 bit 는 FT(Frame type index)로 해당 스피치 프레임의 전송모드 및 SID 모드를 표 1의 Code로 나타낸다. 6 번째 bit 는 Q(Frame quality indicator)로 해당 프레임이 심하게 손상되었다면 0으로 설정된다. 남은 3 bits 는 0으로 채워지게 된다. 예를 들어 MR 102 모드로 전송된 경우 정상적으로 전송된 프레임이라면 프레임 헤더는 "00110100"이 된다. 프레임 헤더 뒤에 존재하는 순수 스피치 데이터는 앞서 말한 CELP 모델의 파라미터 정보를 담고 있는데 전송 모드마다 파라미터 종류나 순서, 해당 파라미터 정보가 저장되는 비트 수 등이 다르다. 이는 AMR 인코딩이 전송 모드별로 특정 과정의 횟수가 추가되거나 계산범위가 줄어드는 등의 차이가 있기 때문이다. 예를 들어 LP 분석의 경우 MR 122 모드에서는 두 번 수

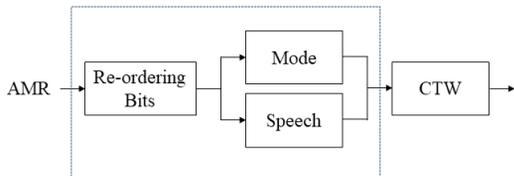
행되지만 나머지 7개 모드에선 한 번만 수행된다. 이후 양자화와 interpolation을 쉽게 수행하기 위해 LP 계수를 LSP(Line Spectrum Pair)영역으로 변환시켜 순수 스피치 데이터 앞쪽에 저장된다. 모드별로 저장되는 파라미터 종류 및 비트 수 등이 다른 예시를 MR 475와 MR 122 모드를 대표로 표 2에 나타내었다. 예를 들어 전송모드가 MR 475인 경우 추출된 파라미터 정보는 프레임 당 95 bits가 할당되고 95 bits의 1 ~ 23 번째 bit는 LSP set에 대한 정보가, 24 ~ 43 번째 bit는 Pitch delay에 대한 정보가, 44 ~ 79 번째 bit는 Algebraic code 정보가, 80 ~ 95 번째 bit는 Gains 정보가 저장된다. MR 122은 또다른 파라미터 종류와 순서, 저장되는 비트 수 등을 갖는다. 이처럼 전송모드마다 비트들이 위치함에 있어서 고유의 순서가 있음을 알 수 있다. 파일로 저장될 경우 한 프레임을 바이트 단위로 저장하기 위해 프레임 종단에 parity가 올 수 있다.

3. CTW

CTW(Context Tree Weighting)는 가변 차수 마르코프 모델(Variable Order Markov model 이하 VOM)의 예측을 혼합하는 방식의 무손실 압축 및 예측 알고리즘이다. 마르코프 모델은 예측하고자 하는 값 바로 직전의 상태 값이 이전의 정보들을 대표할 수 있는 값이라는 가정을 전제로 두고 있다. 따라서 직전의 상태 값 뿐만 아니라 더 이전의 정보까지 활용하기 위해 고 차수(high order) 마르코프 모델로 개념을 확장하였는데 차수가 증가함에 따라 계산복잡도가 지수적으로 증가하는 단점으로 인해 실질적으로 활용되기엔 어려운 개념이다[13]. 이러한 단점을 해결하는 방안으로 VOM이 활용된다[14-16]. VOM은 고 차수 마르코프 모델의 일종이지만, 다음 상태 값을 예측하기 위해 필요한 과거 상태 값의 개수를 기존 정보의 특성 및 흐름의 패턴을 이용하여 정하기 때문에 매년 다른 차수를 가질 수 있다. 따라서 계산 복잡도가 불필요하게 증가하지 않아 실질적 활용도가 높아진다. 여러 가지 VOM 모델 중 바이너리소스를 이용하여 압축하는 방식이 CTW이다. 컨텍스트 모델을 바이너리 트리로 이용하여 확률을 추정하고 탐색한 모든 노드의 가중치를 효율적인 재귀방식으로 가중한다. 이렇게 부호화된 소스는 복잡성과 중복성을 낮추기 위해 산술 부호화(Arithmetic Coding)로 다시 한번 더 인코딩 된다. CTW의 큰 장점은 알려지지 않은 모델 및 파라미터를 갖는 트리 소스에 대해서도 효율적인 압축을 진행한다는 것이다.

4. CTW를 이용한 AMR 음성 데이터 압축

3장에서 설명한 바와 같이 CTW는 바이너리 소스에 대한 순차 데이터 압축 방식으로 컨텍스트 트리를 사용하여 코딩 분포를 생성한다. 즉, 순차적으로 비트 단위로 수행하기 때문에 입력 데이터의 중복도가 비트 단위로 계산된다. 또한 2장 AMR 음성 부호화기에서 설명했듯이 표준코드로 부호화된 AMR 포맷의 음성 데이터는 표 2에 나타난 것처럼 모드별로 파라미터 할당 비트수와 위치가 정해져있다. 이 두 가지 특징을 고려하여 CTW가 압축하기 용이하도록 AMR 음성 데이터를 새로운 포맷으로 변환한다면 결과적으로 압축률이 증가할 것이다. 본 장에서는 제안된 알고리즘의 설명과 실험 방식에 대해 설명하기로 한다.

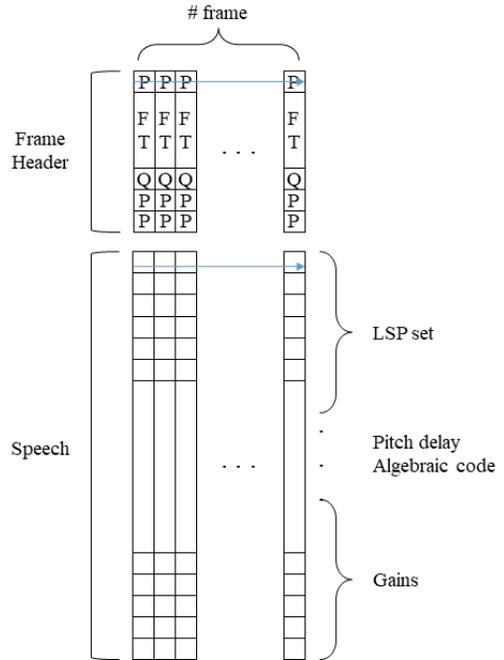


(그림 1) 제안된 알고리즘의 흐름도
(Figure 1) Flowchart of the proposed algorithm

4.1. 제안된 알고리즘

단일모드로 전송된 AMR 음성 데이터의 경우 모드 정보를 담고 있는 8 bits의 프레임 헤더는 계속해서 반복될 것이고 순수 스피치 데이터들 또한 비트들이 나타내는 파라미터 종류가 위치별로 반복될 것이다. 따라서 제안된 알고리즘은 비트의 중복도가 증가하도록 데이터를 재배치하고 CTW가 압축하기 용이하도록 모드정보가 담겨진 프레임 헤더와 순수 스피치 데이터를 따로 저장하여 압축하는 방식이다. 알고리즘의 흐름은 그림 1과 같다. 그림 1에 비트 재배치(Re-ordering Bits)는 입력 데이터를 비트 단위로 재배치 하는 과정으로 비트의 중복도가 증가하도록 그림 2와 같이 맨 앞 6 바이트에 해당하는 헤더가 제거된 AMR 음성 데이터를 한 열에 한 프레임이 오도록 재배열한다. 이 때, 순수 스피치 데이터의 중단에 위치한 parity는 제거한다. 재배열된 데이터는 비트단위로 래스터(raster) 스캔하여 파일을 생성하는데 프레임 헤더와 순수 스피치 데이터를 따로 저장한다. 마지막으로 두 파일을 CTW로 압축하여 저장한다.

Ex) MR 475



(그림 2) CTW에 용이한 AMR 포맷 변환
(Figure 2) Convert AMR format to facilitate CTW

프레임 헤더로부터 생성된 파일은 mode 파일, 순수 스피치 데이터로부터 생성된 파일은 speech 파일이라 정의하겠다. 두 파일로 나눠 저장하는 이유는 각 파일별로 비트 상관도가 높아 나눠 압축할 경우 압축률이 증가한다는 장점이 있고, mode 파일의 크기로부터 프레임 개수를 알 수 있기 때문이다.

4.2. 디코딩

제안된 알고리즘으로 인코딩된 데이터를 디코딩하기 위해선 CTW 디코딩 후 비트 재배치에 대한 디코딩을 통해 AMR 데이터로 복원시켜야 한다. AMR과 CTW의 디코딩은 기존의 방식을 따르면 된다. 프레임헤더는 압축모드에 관계없이 프레임당 1바이트가 할당되기 때문에 mode파일의 크기로부터 전체 프레임의 개수를 알 수 있다. mode파일의 데이터를 한 열에 프레임 개수만큼의 비트가 오도록 재배열한 뒤 래스터(raster) 스캔한다. 이 데이터를 1바이트씩 자르면 프레임헤더 정보가 복원되며 모드정보를 알 수 있다. 표 2에서 든 예시와 같이 모드에

따라 parity 개수가 정해져 있기 때문에 speech파일 또한 mode파일과 같은 방법으로 비트 재배열 한 뒤 parity를 종단에 붙여주면 원래의 AMR 데이터로 복원된다.

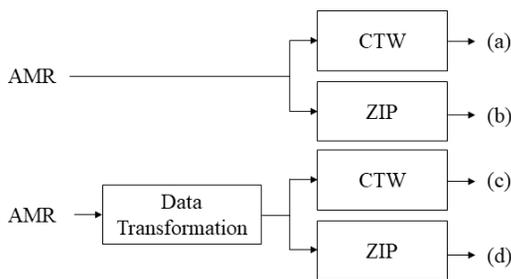
4.3. 실험 방식

실험은 총 네 가지의 경우의 수를 비교하는 방식으로 그림 3과 같이 진행하였다. 추가 압축률을 얻기 위한 기존의 이중 압축 방식 즉, 원본 AMR 파일을 다른 압축방식으로 추가 압축하는 방식과 AMR 데이터를 제안된 알고리즘에서와 같이 비트 재배치한 후 추가 압축하는 방식을 비교하는 실험을 하였고, CTW의 성능을 비교하기 위해 대증적으로 쓰이는 ZIP 압축 방식을 함께 비교하는 실험을 하였다.

입력된 음성 데이터는 전송률이 가장 낮은 MR 475 단일 모드의 AMR 음성 데이터로, 맨 앞 6 바이트에 해당하는 헤더를 제거한 상태이다. 그림 3의 (a)와 (b)는 헤더가 제거된 AMR 음성 데이터에 대해 각각 CTW와 ZIP 방식으로 압축한 결과이다. (c)는 헤더가 제거된 AMR 음성 데이터를 4장에서 제안된 알고리즘으로 압축한 결과이고, (d)는 제안된 알고리즘에서 CTW 대신 ZIP 방식을 채택한 결과이다. 압축률(CR)은 압축 전 데이터의 총 크기를 N_0 , 압축 후 데이터의 총 크기를 N 이라 할 때

$$CR = \left(1 - \frac{N}{N_0}\right) \times 100 \quad (\%)$$

로 계산하였다.



(그림 3) 실험 흐름도

(Figure 3) Flowchart of experiment

5. 실험 결과

실험에 사용된 데이터는 100초 분량의 음성 데이터로 각각 한국어 남자, 한국어 여자, 영어 남자, 영어 여자 총

(표 3) 데이터 변환 적용 및 압축기 종류에 따른 압축률 (Table 3) Compression rate according to data conversion application and compressor type

* 그림 3에서 명시된 (a),(b),(c),(d)와 같음
* Same as (a),(b),(c),(d) specified in Figure 3

test data	CR (%)			
	(a)	(b)	(c)	(d)
1	4.5938	3.1169	9.2567	8.5515
2	4.0046	2.7585	8.7130	8.2641
3	6.1262	3.7954	9.3017	8.5608
4	4.6969	3.1785	9.1262	8.6338
ave.	4.8554	3.2123	9.0994	8.5025
	4.0338		8.8010	

4개(총 400초)의 파일을 사용하였다. 헤더를 제거한 입력 AMR 파일의 크기는 65 kbytes 이고 각각 CTW와 ZIP으로 단순 추가 압축한 결과 평균 압축률은 (a)CTW일 때 4.8554%, (b)ZIP일 때 3.2123%로 CTW가 ZIP에 비해 압축률이 1.6431% 더 높은 결과를 보였다. 같은 입력 파일을 그림 2와 같이 데이터 변환을 하면 mode 파일은 5 kbytes, speech 파일은 59.375 kbytes 이며 mode 파일의 경우 CTW로 압축하였을 때 27 bytes, ZIP으로 압축하였을 때 154 bytes로 압축되었다. mode 파일과 speech 파일을 각각 CTW와 ZIP으로 압축한 결과 즉, 비트 재배치를 거친 후 추가 압축을 시행한 결과 평균 압축률은 (c)CTW일 때 9.0994%, (d)ZIP일 때 8.5025%로 CTW가 ZIP에 비해 압축률이 0.5969% 더 높은 결과를 보였다. 또한 CTW와 ZIP 모두 입력 데이터의 비트 재배치 과정으로 인해 평균 압축률이 4.0338%에서 8.8010%로 향상되었다. 결과적으로, 흔히 사용되는 ZIP에 의한 단순 추가 압축에 비해 제안된 알고리즘이 약 5.89% 높은 추가 압축률을 보였다. 실험 데이터 별 압축률 및 평균 압축률은 표 3에 정리하였다.

6. 결 론

무선통신 분야는 모바일 기기의 보급화에 따라 서비스 품질 향상 및 동영상을 포함한 멀티미디어 콘텐츠 제공을 위해 많은 발전을 기해왔다. 무선통신 분야는 음성서비스 품질 향상과 무선 멀티미디어 서비스의 제공을 목표로 표준화 작업을 시행하였으며 음성부호화의 경우 무선채널의 환경변화에도 음성품질을 유지할 수 있도록 8개의 전송모드를 갖는 AMR 음성부호화 방식을 IMT-2000 표준안으로 선정하였다. 본 논문에서는 무손실 압축기 CTW를 이용하여 AMR 음성 데이터의 압축 성능을 개

선하는 알고리즘을 제안하였다. CTW는 무손실 압축 및 예측 알고리즘으로 바이너리 소스를 사용하기 때문에 비트단위로 수행한다는 특징이 있으며 미지의 모델 및 파라미터를 갖는 트리소스에 대해서도 효율적인 압축을 진행한다는 장점이 있다. 이러한 특징과 장점을 이용해 AMR 포맷 데이터가 CTW 압축에 효과적일 수 있도록 AMR 저장 포맷을 비트 단위로 변환하는 알고리즘을 제안하였다. 실험 결과, AMR 데이터에서 ZIP보다 CTW 압축이 추가 압축률을 향상시켰고, 범용적으로 사용되는 ZIP과의 단순 이중 압축에서 평균 압축률이 3.2123%, 제안된 알고리즘을 통한 평균 압축률이 9.0994%로 약 5.89%의 압축률 향상을 본 연구에서 보여주었다.

참고문헌(Reference)

- [1] B. Bessette, R. Salami, R. Lefebvre, M. Jelinek, J. R. Pukkila, J. Vainio, H. Mikkola and K. Jarvinen, "The adaptive multirate wideband speech codec (AMR-WB)", IEEE, Vol. 10, No. 8, pp. 620-636, 2002.
<http://dx.doi.org/10.1109/TSA.2002.804299>
- [2] J. Makinen, B. Bessette, S. Bruhnm P. Ojala, R. Salami and A. Taleb, "AMR-WB+: a new audio coding standard for 3rd generation mobile audio services", IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.
<http://dx.doi.org/10.1109/ICASSP.2005.1415603>
- [3] ETSI Draft EN 301 704, Digital cellular telecommunication system(Phase 2+); Adaptive Multi-Rate(AMR) speech transcoding, 1999.
https://www.etsi.org/deliver/etsi_en/301700_301799/301704/07.01.00_40/en_301704v070100o.pdf
- [4] Adaptive Multi-Rate(AMR) speech codec; Transcoding functions, 3GPP TS 26.090, v14.0.0, 2017.
<https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=1392>
- [5] P. A. Barany, J. A. Bharatia and C. S. Bontu, "Communications using adaptive multi-rate codecs", U.S. Patent, No. 7 7072 336, 2006.
<https://patents.google.com/patent/US7072336B2/en>
- [6] B. Bessette, R. Salami, R. Lefebvre, M. Jelinek, J. Rotola-Pukkila, J. Bainio, H. Mikkola, and K. Jarvinen, "The adaptive multi-rate wideband speech codec (AMR-WB)", IEEE Transactions on Speech and Audio Processing, Vol. 10, No. 8, pp. 620-636, 2002.
<http://dx.doi.org/10.1109/TSA.2002.804299>
- [7] ETSI Draft EN 301 708, Digital cellular telecommunication system(Phase 2+); Voice Activity Detector(VAD) for adaptive Multi-Rate(AMR) speech traffic channels, 1999.
https://www.etsi.org/deliver/etsi_en/301700_301799/301708/07.01.00_40/en_301708v070100o.pdf
- [8] Mandatory Speech Codec speech precessing functions : AMR Speech Codec; Source Controlled Rate operation, 3GPP, 3G TS 26.093, 1999.
- [9] F. M. J. Willems, Y. M. Shtarkov and T. J. Tjajklens, "The context-tree weighting method: basic properties" IEEE Transactions on Information Theory, Vol. 41, no. 3, pp. 653-664, 1995.
<http://dx.doi.org/10.1109/18.382012>
- [10] F. M. J Willems, "The context-tree weighting method :extensions", IEEE Transactions on Information Theory, Vol. 44, no. 2, pp. 792-798, 1998.
<http://dx.doi.org/10.1109/18.661523>
- [11] "Digital cellular telecommunication system(Phase 2+); Substitution and mting of lost frames for Adaptive Multi-Rate(AMR) speech traffic channels", ETSI Draft EN 301 705.
https://www.etsi.org/deliver/etsi_en/301700_301799/301705/07.00.01_40/en_301705v070001o.pdf
- [12] A. M. Kondoz, Digital Speech, John Wiley & Sons, 1994.
- [13] D. Mochilhashi and E. Sumita, "The Infinite Markov Model", Advances in Neural Information Processing Systems, Vol. 20, pp. 1017-1024, 2008.
<http://mlg.eng.cam.ac.uk/zoubin/papers/ihtm.pdf>
- [14] A. Apostolico and G. Bejerano, "Optimal Amnesic probabilistic automata or how to learn and classify proteins in linear time and space", Journal of Computational Biology, Vol. 7, No. 3-4, pp. 381-393, 2000.
<http://dx.doi.org/10.1089/106652700750050844>
- [15] P. Buhlmann and A. J. Wyner, "Variable length Markov chains", The Annals of Statistics, Vol. 27, No. 2, pp. 480-513, 1999.
<https://www.stat.berkeley.edu/~binyu/212A/papers/vlmc.pdf>

- [16] F. C. Pereira, Y. Singer, and N. Tishby, "Beyond Word N-Grams", Natural Language Processing Using Very Large Corpora, Vol. 11, pp. 121-136, 1999.
<https://www.aclweb.org/anthology/W95-0108.pdf>

◎ 저 자 소개 ◎



이 은 수(Eun-su Lee)

2020년 을지대학교 안경광학과(보건학사)
2020년~현재 상명대학교 일반대학원 컴퓨터과학과 석사과정
관심분야 : 영상 처리
E-mail : dldmstr0510@naver.com



오 은 주(Eun-ju Oh)

2018년 상명대학교 미디어소프트웨어학과(이학사)
2018년~현재 상명대학교 일반대학원 컴퓨터과학과 석사과정
관심분야 : 영상 압축, 영상 처리
E-mail : essenti10@naver.com



유 훈(Hoon Yoo)

1990년~1997년 한양대학교 전자통신공학과 공학사
1997년~1999년 한양대학교 대학원 전자통신공학과 공학석사
1999년~2003년 한양대학교 대학원 전자통신공학과 공학박사
2003년~2005년 삼성전자 무선사업부 책임연구원
2005년~2008년 동서대학교 컴퓨터정보공학부 조교수
2008년~2012년 상명대학교 디지털미디어학부 조교수
2012년~2017년 상명대학교 미디어소프트웨어학과 부교수
2017년~현재 상명대학교 융합전자공학과 정교수
1997년2월26일 제3회 삼성전자 휴먼테크 논문상 대회 입상
2003년~2005년 삼성전자 위성DMB 폰, 지상파DMB 폰 최초 개발 및 양산
관심분야 : 영상 처리, 3차원 집적 영상 처리
E-mail : hunie@smu.ac.kr