

음악 파일의 무손실 압축

김성용*, Kolesnikova Anastasiya*, 이원돈*

*충남대학교 컴퓨터 전공

e-mail : msm18e@hanmail.net

Lossless Compression of Music Files

Sung Yong Kim*, Kolesnikova Anastasiya*, Won Don Lee*

*Dept. of Computer Science & Engineering, Chungnam National University

요 약

본 논문은 현존하는 음원 파일 압축의 개선을 위해 작성되었다. 새로운 무손실 압축 알고리즘을 통한 오차값의 분포는 새로운 압축기법의 가능성을 설명한다. 몇가지 실험은 제안하는 알고리즘의 장점을 제시한다.

1. 서론

디지털 시대에 도달한 이래, 멀티미디어 데이터의 변환 기술은 빠르게 발전해 왔다. 우리가 일반적으로 사용하는 MPEG 3 음악 파일은 이런 기술중 하나이다. MPEG 3 는 사이즈는 작지만 음질은 좋지 못한 프로그램이다. 왜냐하면 원본 데이터를 손상시켜 압축하기 때문이다. 무손실 압축된 음악 파일은 크기가 크다는 단점이 있다. 하드 디스크의 저장공간이 증가하고 인터넷을 통한 파일의 전송비용이 낮아지면서, 사용자들이 더 높은 음질을 요구함에 따라 무손실 압축 기술이 중요해지고 있다. 컴퓨터에 음원 파일을 저장하는 것이 보편화 되어 있고 이 파일들은 많은 저장공간을 필요로 한다. 일반 압축 알고리즘들은 이런 파일들의 데이터 구조와 본래의 신호를 잘 처리하지 못한다. 본 논문은 현존하는 음원 파일 압축 기술의 개선을 위해 작성되었다. 일반적인 음악 파일 압축 기법의 구성에 따르면: 첫째, 데이터를 적당한 크기로 구성한다. 둘째, 선형 예측 모델을 예측한다. 마지막으로 실제 데이터와 예측된 데이터의 차이를 라이스 코딩한다. 현존하는 음원 파일 압축 기술은 네가지 다른 예측 방법을 사용한다. 본 논문에서는 새로운 예측 기능과 가능성을 3 절의 실험을 통해 제공한다.

2. 새로운 예측을 이용한 무손실 알고리즘

압축은 원본 음악 파일의 예측하는 모델을 만드는 것에 의해 달성된다. 잘 알려진 모델은 선형 압축 코딩(LPC)이다. 과정은 전의 분포의 조합으로 예측된다.

$$\hat{w}(t) = a_t w(t-1) \quad (식 1)$$

$d(t)$ 는 선형 예측 $\hat{w}(t)$ 의 추정과 신호음 $w(t)$ 의 차이로 설정했다.

$$d(t) = w(t) - \hat{w}(t) \quad (식 2)$$

Shorten 압축 알고리즘 [6]에서 유용한 선형 예측의

한정 방식이 발견되어졌다. 이 방법에서 예측 계수는 지난 데이터 포인트에 적합한 다항식을 통해 나타내진다. 예를 들면:

$$\hat{w}_0(t) = 0 \quad (식 3)$$

$$\hat{w}_1(t) = w(t-1) \quad (식 4)$$

$$\hat{w}_2(t) = 2w(t-1) - w(t-2) \quad (식 5)$$

$$\hat{w}_3(t) = 3w(t-1) - 3w(t-2) + w(t-3) \quad (식 6)$$

$w_i(t)$ 은 예측 다항식의 예러 표시에 쓰인다.

$$d_0(t) = w(t) \quad (식 7)$$

$$d_1(t) = d_0(t) - d_0(t-1) \quad (식 8)$$

$$d_2(t) = d_1(t) - d_1(t-1) \quad (식 9)$$

$$d_3(t) = d_2(t) - d_2(t-1) \quad (식 10)$$

예측 재귀 알고리즘은 오차 예측 방정식(식 7-10)을 컴퓨터에서 활용하기 위해 있다. 각 오차 구간은 이전의 차이에 의해 판단된다. 각 구간은 오직 하나의 정수 추가/삭제를 포함한다, 이것은 모든 예측과 선택시 최선의 계산을 가능하게 한다. Shorten 알고리즘은 예측의 근거에 이러한 식들을 사용한다. 그후에 라이스 코딩 기법에 의해 오차값이 코딩된다. 이런 코딩 방식으로 숫자는 신호 비트, n 개의 Low 정렬 비트 그리고 남은 High 정렬 비트로 나뉜다. High 정렬 비트는 종료비트 1 에 0 의 개수를 전달하는 정수로 간주된다. n 개의 Low 정렬 비트에 대한 예시는 다음 <표 1>에 있다.

<표 1> $n=3$ 일때 라이스 코딩의 예

Number	sign bit	lower bits	Number of '0'	full code
0	0	000	0	00001
31(11111)	0	111	3	01110001
-11(01011)	1	011	1	101101

본 논문중에 제시된건, 첫번째 예측식인 (식 4)를 (식 11)로 대체하는 것이다.

$$w_j(t) = 2w(t-1) - 2w(t-2) + w(t-3) \quad (\text{식 11})$$

(식 11)로 대체한 이유는 (식 4)보다 음악 파일의 압축 테스트 결과 성능 향상이 가능하였기 때문이다. 그 결과는 3 절의 실험에서 보도록 하겠다.

Tony Robison의 Shorten 의하면 Low 정렬 비트의 최상의 숫자는 신호의 변화에 관계되어진다. Laplacian의 정의와 같이:

$$p(x) = \frac{1}{\sqrt{2}\sigma} e^{-\frac{\sqrt{2}}{\sigma}|x|}$$

|x|는 x의 절대값이다. |x|가 주는 예외는:

$$\begin{aligned} E(|x|) &= \int_{-\infty}^{\infty} |x|p(x)dx \\ &= \int_0^{\infty} x \frac{\sqrt{2}}{\sigma} e^{-\frac{\sqrt{2}}{\sigma}x} dx \\ &= \int_0^{\infty} e^{-\frac{\sqrt{2}}{\sigma}x} dx - [xe^{-\frac{\sqrt{2}}{\sigma}x}]_0^{\infty} \\ &= \frac{\sigma}{\sqrt{2}} \end{aligned}$$

이것은 low 비트, n, ±2ⁿ 범위안에 놓여 있는 표본들의 절반 같은 숫자를 찾는데 필요하다. 이것은 n+1 비트에서 0.5 일 확률과 n+k+1에서 2^{-(n+k)} 일 확률 같은 상황에서 라이스 코딩하는 것을 안전하게 한다. 이를 최적화한 것은

$$\begin{aligned} 1/2 &= \int_{-2^n}^{2^n} p(x)dx \\ &= \int_{-2^n}^{2^n} \frac{1}{\sqrt{2}\sigma} e^{-\frac{\sqrt{2}}{\sigma}|x|} dx \\ &= -e^{-\frac{\sqrt{2}}{\sigma}2^n} + 1 \end{aligned}$$

n을 해결하면

$$\begin{aligned} n &= \log_2\left(\log(2) \frac{\sigma}{\sqrt{2}}\right) \\ &= \log_2(\log(2) E(|x|)) \end{aligned}$$

마지막 n 비트의 분포가 일정하기 않기 때문에, Laplacian에 따르면, 이미 라이스 코드된 압축 파일을 엔트로피 코딩을 더 사용하여 추가로 압축할수 있는 가능성이 있다. 3 절에서는 제시된 압축 기법을 선택해 라이스 코딩 실험의 결과와 결과에 대한 설명을 하고 있다.

3. 실험 및 결과

제시된 압축 방식 테스트를 위해 5 가지 장르의 음악 파일의 샘플을 10 개씩 선택했다. 이 음악 파일들은 제시된 기법과 종래의 Shorten 알고리즘을 각각 사용하여 압축한뒤 결과를 비교하여 <표 2-11>에 나타내었다. Shorten ratio와 Improved shorten ratio는 압축비율로써 size/원래파일*100을 계산한것이다. 그리고 결과물인 <표 2-11>은 제시된 기법으로 압축한 결과, 기존보다 더 압축할수 있다는 것을 입증한다.

<표 2> 발라드

(kb)	song1	song2	song3	song4	song5
original	11480	7709	13947	6192	6515
shorten	7415	5379	9362	4157	4346
improved shorten	7335	5324	9250	4117	4295
shorten ratio	64.59%	69.78%	67.13%	67.14%	66.71%
improved shorten ratio	63.89%	69.06%	66.32%	66.49%	65.92%

<표 3> 발라드

(kb)	song6	song7	song8	song9	song10
original	9056	18967	10622	21274	7157
shorten	4479	13732	6781	14867	5190
improved shorten	4395	13548	6620	14686	5097
shorten ratio	49.46%	72.40%	63.84%	69.88%	72.52%
improved shorten ratio	48.53%	71.43%	62.32%	69.03%	71.22%

<표 4> 클래식

(kb)	song1	song2	song3	song4	song5
original	12693	2884	4143	6314	23099
shorten	5786	1511	2167	3579	11532
improved shorten	5685	1487	2135	3530	11403
shorten ratio	45.58%	52.39%	52.31%	56.68%	49.92%
improved shorten ratio	44.79%	51.56%	51.53%	55.91%	49.37%

<표 5> 클래식

(kb)	song6	song7	song8	song9	song10
original	15082	20734	4582	13487	12500
shorten	7764	12158	1800	6051	5459
improved shorten	7664	12053	1755	5951	5379
shorten ratio	51.48%	58.64%	39.28%	44.87%	43.67%
improved shorten ratio	50.82%	58.13%	38.30%	44.12%	43.03%

<표 6> 성가

(kb)	song1	song2	song3	song4	song5
original	5042	1201	13966	3954	12766
shorten	2281	545	9270	1700	6252
improved shorten	2247	539	9191	1674	6169
shorten ratio	45.24%	45.38%	66.38%	42.99%	48.97%
improved shorten ratio	45.04%	45.30%	66.30%	42.84%	48.83%

<표 11> 메탈 & 락

(kb)	song6	song7	song8	song9	song10
original	11414	9205	6365	5583	5909
shorten	8862	7510	4560	4476	4779
improved shorten	8782	7436	4507	4420	4713
shorten ratio	77.64%	81.59%	71.64%	80.17%	80.88%
improved shorten ratio	76.94%	80.78%	70.81%	79.17%	79.76%

<표 7> 성가

(kb)	song6	song7	song8	song9	song10
original	6330	7478	2842	6322	4300
shorten	2927	3588	1661	3431	2404
improved shorten	2886	3539	1645	3397	2378
shorten ratio	46.24%	47.98%	58.44%	54.27%	55.91%
improved shorten ratio	45.59%	47.33%	57.88%	53.73%	55.30%

4. 결론

본 논문은 개선된 무손실 압축 알고리즘을 제시하고 파형 파일의 저장에 대해 기술했다. 오차값을 위한 예측의 사용은 라이스 코딩에 따른 Lalpacion 분포를 따라 발견해왔다. 오차값의 분포의 예측을 분석하고 무손실 압축에서 새로운 압축 기법의 가능성을 설명했다. 마지막으로 실험은 제시된 기법의 장점을 증명하고 실제 압축됨을 보여주도록 설계되었다.

<표 8> 경음악

(kb)	song1	song2	song3	song4	song5
original	10420	20737	7197	7407	7600
shorten	5420	7953	3793	3924	4575
improved shorten	5357	7808	3749	3881	4532
shorten ratio	52.02%	38.35%	52.70%	52.98%	60.20%
improved shorten ratio	51.41%	37.65%	52.09%	52.40%	59.63%

<표 9> 경음악

(kb)	song6	song7	song8	song9	song10
original	7809	11371	11789	7695	5315
shorten	3994	6800	5670	4606	2972
improved shorten	3945	6470	5594	4563	2943
shorten ratio	52.02%	38.35%	52.70%	52.98%	60.20%
improved shorten ratio	51.41%	37.65%	52.09%	52.40%	59.63%

<표 10> 메탈 & 락

(kb)	song1	song2	song3	song4	song5
original	10717	4395	14143	5843	11563
shorten	8704	3022	11255	4287	8338
improved shorten	8650	2986	11032	4237	8220
shorten ratio	81.22%	68.76%	79.58%	73.37%	72.11%
improved shorten ratio	80.71%	67.94%	78.00%	72.51%	71.09%

참고문헌

- [1] John Garofolo, Tony Robinson, and Jonathan Fiscus. The development of file formats for very large speech corpora: sphere and shorten. In Proc. ICASSP, volumn I, pages 113-116, 1994.
- [2] N. S. jayant and P. Noll. Digital Coding of Waveforms. Prentice Hall, Englewood Cliffs, NJ, 1984. ISBN 0-13-211913-7 01.
- [3] Van der Geer J, Hanraads JAJ, Lupton RA. The art of writing a scientific article. J Sci Commun 2000;163:51-9.
- [4] Pen-Shu Yeh, Robert Rice, and Warner Miller. On the optimality of code options for a universal noisless order. JPL Publication 91-2, Jet Propulsion Laboratories, February 1991.
- [5] Robert F.Rice. Some practical noiseless coding techniques, Part II, Module PSI14,k+. JPL Publication 91-3, Jet Propulsion Laboratories, November 1991.
- [6] Tony Robinson. SHORTEN: Simple lossless and near-lossless waveform comparession, Technical report CUED/ F-INFENG/ TR. 156. Cambridge University Engineering Department, Trumpington Street, Cambridge, CB2 1PZ, UK. December 1994.