

일반논문-04-09-4-09

오디오 신호를 위한 표본화율 변환 알고리즘 성능 비교

이용희^{a)*}, 김인철^{b)}

A Performance Comparison of Sampling Rate Conversion Algorithms for Audio Signal

Yong-Hee Lee^{a)*} and Rin-Chul Kim^{b)}

요 약

본 논문에서는 44.1KHz에서 48KHz로 표본화 주파수를 변환하는 알고리즘들의 성능을 비교한다. 비교 한 기법은 다위상 구조로 구현된 기본적인 기법, sinc 함수를 이용한 기법, 다위상 구조의 다단계 구현 기법, 그리고 B-spline을 이용한 기법 등이다. 먼저, 공정한 비교를 위해 이 4가지 기법을 이용한 표본화율 변환기들을 고품질 조건하에 재설계하고, 이들의 H/W 복잡도를 메모리 요구량과 계산량 측면에서 비교하였다. 그 결과, 메모리 요구량 측면에서는 B-spline을 이용한 기법이 가장 우수하였지만, 계산량 측면에서는 기본적인 기법과 sinc 함수를 이용한 기법이 가장 우수함을 확인할 수 있었다.

Abstract

In this paper we compare the performance of 4 different algorithms for converting the sampling frequency of an audio from 44.1KHz to 48KHz. The algorithms considered here include the basic polyphase method, sinc function based method, multi-stage method, and B-spline based method. For a fair comparison, the sampling rate converters using the 4 algorithms are redesigned under a high fidelity condition. Then, their H/W complexities are compared in terms of the computational complexity and the memory size. As a result, it is shown that the basic polyphase method and sinc function based method outperform the other two in terms of the computational complexity, while the B-spline based method requires less memory than the others.

Keywords : Sampling Rate Conversion, H/W Complexity, Audio

I. 서 론

Compact disc (CD), digital audio tape (DAT) 등은 20~20,000Hz의 가청 주파수 대역을 충실히 표현하는 고품질 오디오 시스템이다. 오디오 시스템에서는 오디오 표현

형식에 따라 44.1, 48, 96KHz 등의 표본화 주파수를 채택한다. 따라서, 다양한 오디오 형식을 효과적으로 수용하고, 이를 적절히 처리하기 위해서는 입력되는 오디오 신호의 표본화 주파수를 단일 표본화 주파수로 변환시킬 필요가 있다. 이 가운데, 44.1KHz와 48KHz 표본화 주파수 사이의 변환은 그 비가 간단한 정수배로 나타내어지지 않기 때문에, 지금까지 이에 대한 연구가 심도 있게 수행되었다^{[2][3][4]}. 그러나 지금까지 알려진 표본화율 변환 기법들은 서로 상이한 기준에 의해 설계되어졌기 때문에, 고품질 오디오 시스템에 적용하였을 때, 각 방법의 성능과 복잡도는 검토되지 않았다. 본 논문에서는 고품질 오디오 시스템에 적합하게 표본화

a) 정회원, LG 전자기술원 MM연구소

LG Electronics Institute of Tech., MM lab.

b) 서울시립대학교 전자전기컴퓨터공학부

Dept. of Electrical & Computer Eng., University of Seoul

※ 본 연구는 '2003년 서울시립대 연구교수 지원 프로그램'에 의한 결과입니다.

을 변환기(sampling rate converter)들을 다시 설계한다. 각 표본화율 변환기는 20~20000Hz의 가청주파수 대역에서 동적 범위(dynamic range)가 100dB 이상, 그리고 주파수 응답 리플이 $+5 \times 10^{-4}$ dB 이하로 유지될 수 있도록 설계한다. 설계된 각 변환기의 성능을 검토하고, 메모리 요구량과 계산량 측면에서 복잡도를 비교한다. 여기서 동적 범위는 시스템에서 잡음과 구별하여 표현할 수 있는 신호의 범위를 가리키고, 주파수 응답 리플은 통과 대역에서 신호의 왜곡 정도를 나타낸다. 보다 구체적으로, 동적 범위는 어떤 시스템에 0dB 입력신호를 인가했을 때 출력 신호의 최대값에서 잡음의 최대값 혹은 평균값을 뺀 값으로 측정하는데, 본 논문에서는 더 엄격한 잡음의 최대값을 적용하였다. 주파수 응답 리플은 가청 주파수 대역에서의 주파수응답이 0dB를 기준으로 편차가 어느 정도인지를 나타낸다.

44.1KHz에서 48KHz로 표본화 주파수를 변환하는 기본적인 기법은 먼저, 160배 up-sampling 하고, $\pi/160$ 의 차단 주파수를 갖는 저역통과 필터(lowpass filter)를 거쳐 147배 down-sampling 하는 것이다. 이때 저역통과 필터의 입력은 대부분 0이라는 것을 참조하여 다위상(polyphase) 구조로 구현하면 계산 복잡도를 감축시킬 수 있다. 그리고 4배 보간 후 40/147 배의 표본화율 변환을 수행하는 다단계(multi-stage) 구조를 사용한다면 보다 효과적으로 구현할 수 있다^[1,4]. 한편, 표본화율 변환은 샘플마다 비 정수 값의 가변적인 지연을 가지는 가변적인 시스템으로 볼 수 있으므로, fractional delay filter(FDF)를 이용하여 구현할 수 있다. 이러한 방법으로는 sinc 함수를 이용한 기법^[2], B-spline을 이용하는 기법^[3] 등이 있다. 본 논문에서는 이상에서 언급한 기존의 기법들에 대해서 계산량 및 메모리 요구량 등의 측면에서 비교하고, 고품질 오디오 시스템에 적합성 여부를 검토하고자 한다.

II. 기존의 기법들

1. 기본적인 기법

기본적인 기법에서는 입력을 160배 up-sampling하고 차단 주파수 $\pi/160$ 인 필터를 거친 다음, 147배 down-sampling하여 44.1KHz에서 48KHz로 표본화율을 변환한다. 그림 1의 구조에서 $H(k)$ 의 필터 계수가 대칭이라는 것을 참조하면, 필터의 길이가 L 이라고 할 때, $L/2$ words의

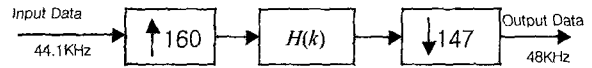


그림 1. 기본적인 기법을 이용한 구조
Fig. 1. Structure of the basic method

메모리가 필요함을 알 수 있다. 또한 160배 up-sampling 후 필터에 입력되는 신호의 대부분이 그 값이 0이라는 점을 고려하여 다위상 구조로 구현한다면, 한 샘플을 얻기 위해 곱셈 $L/160$ 회, 덧셈 $L/160 - 1$ 회의 연산이 필요함을 알 수 있다.

2. Sinc 함수를 이용한 기법

Sinc 함수를 이용한 구조^[2]는 그림 2와 같다. 이 구조에서는 입력 신호에 대해 비 정수 값 만큼의 지연이 있는 위치에서 출력 신호를 얻는다고 보고 다음과 같은 식으로 출력을 계산한다.

$$s(n) = \sum_{m=-N}^N w(m) \text{sinc}(m - \tau(n)) s(k - m) \quad (1)$$

여기서 $w(m)$ 은 Blackman-Harris 윈도우로 블록의 가장 자리에서 발생하는 블록 효과를 감축시키기 위해서 사용하였다. 식 (1)에서 출력을 얻기 위해 샘플 당 곱셈 $2N+1$ 회와 덧셈 $2N$ 회의 계산이 요구됨을 알 수 있다. 또한, sinc 함수 및 윈도우 함수의 곱은 $2N+1$ 개의 계수로 표현된다. 한편, 44.1KHz에서 48KHz로의 표본화율 변환 비가 160/147이라는 것을 고려하면, 160개의 서로 다른 $\tau(n)$ 이 존재한다. 따라서 식 (1)을 이용하여 출력을 얻기 위해 길이가 $2N+1$ 인 필터 160개로 sinc 함수와 윈도우 함수의 곱이 표현되므로, 이 값들을 저장하는데 필요한 메모리는 총 $160 \times (2N+1)$ words이다^[2]. 이러한 메모리 요구량은 식 (2)와 (3)에 보인 바와 같이

$$w(m) = w(m) \quad (2)$$

$$\text{sinc}(m - \tau(n)) = \text{sinc}(m + \tau(n)) \quad (3)$$

윈도우 함수가 우함수이고, sinc 함수 또한 우함수라는 사

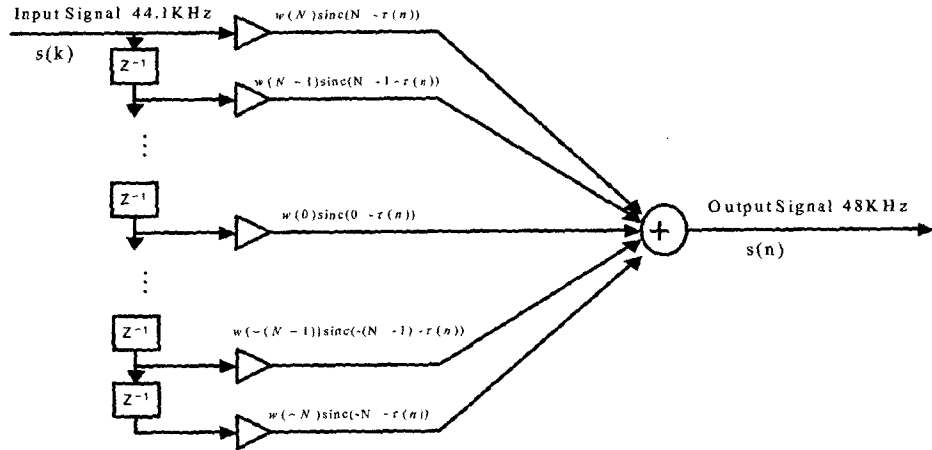


그림 2. Sinc 함수를 이용한 기법의 구조
Fig. 2. Structure of the sinc function based method

실을 참조한다면, $81 \times (2N + 1)$ words로 감축된다.

시간 $t_i < t < t_{i+1}$ 에서의 값을 다음과 같이 얻는다.

3. 다단계 구조를 이용한 기법

$$x(t) = \sum_{k=0}^N x(t_k) B_k(t) \tag{4}$$

다단계 (multi-stage) 구조^[1]를 이용하게 되면 그림 3과 같이 먼저 4배 보간하여 표본화 주파수가 176.4KHz인 신호로 변환한 다음, 40배 up-sampling과 $\pi/147$ 필터 및 147배 down-sampling을 거쳐 표본화율을 변환한다. 이 구조는 기본적인 기법에서의 up-sampling을 두 단계로 나눔으로 필터 설계에 있어서 동일한 성능 기준 하에서 기본적인 기법에 비해 전체 필터의 길이를 줄일 수 있다.

이때, $B_k(t)$ 는 6차 B-spline함수의 경우 식 (4)와 같이 구할 수 있다.

4. B-spline을 이용한 기법

$$\begin{aligned} B_0 &= \frac{1}{6!} (q)^6 \\ B_1 &= \frac{1}{6!} ((1+q)^6 - 7(q)^6) \\ B_2 &= \frac{1}{6!} ((2+q)^6 - 7(1+q)^6 + 21(q)^6) \\ B_3 &= \frac{1}{6!} ((3+q)^6 - 7(2+q)^6 + 21(1+q)^6 - 35(q)^6) \tag{5} \\ B_4 &= \frac{1}{6!} ((3-q)^6 - 7(2-q)^6 + 21(1-q)^6) \\ B_5 &= \frac{1}{6!} ((2-q)^6 - 7(1-q)^6) \\ B_6 &= \frac{1}{6!} (1-q)^6 \end{aligned}$$

B-spline^[3]을 이용한 보간은 이산 시간에서 정의되는 $n+1$ 개의 $x(t)$, $t \in (t_0, t_1, \dots, t_n)$ 을 이용해서 임의의

여기서, q 는 서로 다른 표본화율 신호의 샘플 간 차이로, $0 \leq q \leq 1$ 의 값을 갖게 된다. 그림 4와 같은 경우에 서로 다른 40개의 값이 필요하다.

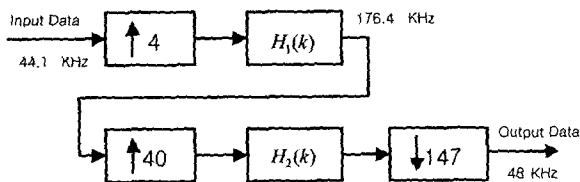


그림 3. 다단계 구조를 이용한 기법의 구조
Fig. 3. Structure of the multistage method

B-spline을 이용한 기법은 그림 4와 같이 먼저, 입력 신호를 차단 주파수가 $\pi/2$ 인 필터 및 2배 up-sampler를 반복 적용하여 4배 up-sampling한 다음, 위에서 설명한 B-spline 보간을 통하여 표본화 주파수가 48KHz인 출력 신호를 얻

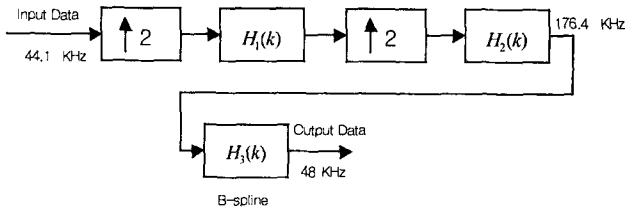


그림 4. B-spline 구조를 이용한 기법의 구조
Fig. 4. Structure of the B-spline based method.

는다. 여기서 B-spline의 감쇠가 $(\sin(x)/x)^2$ 과 같음을 고려하여 가청 주파수 대역에서 이득을 보상해주는 필터를 설계하여야 한다.^[3]

Ⅲ. 실험 및 성능 비교

지금까지 알려진 표본화율 변환 기법들은 서로 상이한 기준에 의해 그 성능들이 제시되었다. 이에 공정한 비교를 위해 동일한 성능을 얻는데 요구되는 하드웨어 복잡도를 메모리 요구량과 계산량 측면에서 검토하여, 최적의 방법을 모색하고자 한다.

실험은 44.1KHz의 표본화 주파수에서 24 비트로 표본화된 정현파 입력을 사용한다. 그리고 사용되는 필터 계수들을 모두 24비트로 표현한다. 이러한 환경 하에서, 가청 주파수 대역에서 100dB 이상의 동적 범위와 $\pm 5 \times 10^{-4}$ dB 이하의 리플 크기를 유지할 수 있도록 각 기법들을 재설계하였다. 각 기법에 사용된 모든 필터는 Matlab^[6] 상에서

Blackman-Harris 윈도우를 이용한 윈도우 기법으로 설계되었다. 그리고 각 기법의 성능 또한 Matlab 상에서 비교하고 평가하였다.

1. 기본적인 기법

그림 1의 구조에서 $H(k)$ 는 그림 5(a)와 같은 주파수 응답을 보이는 14241 taps 필터를 사용하였다. 그 결과, 그림 5(b) 및 그림 5(c)에 보인 바와 같이 가청주파수대역에서 100dB 이상의 동적 범위와 $\pm 5 \times 10^{-4}$ dB 이하 크기의 리플을 얻음을 확인하였다. 이때, 앞서 설명한 필터 계수가 대칭이라는 것을 참조하면, 7121 words의 메모리가 필요함을 알 수 있다. 또한 160배 up-sampling 후에 필터를 거쳐 대부분의 값이 0이라는 점을 고려하여 다위상 구조로 구현한다면, 한 샘플을 얻기 위해 곱셈 89회, 덧셈 88회의 연산이 필요함을 알 수 있다.

2. Sinc 함수를 이용한 기법

Park^[2]은 그림 2의 구조에서 $N=31$ 인 필터를 이용하여 약 83dB의 동적 범위를 얻었는데, 본 논문의 요구 조건을 만족시키기 위해서는 $N=44$ 인 필터를 사용하여야 함을 알 수 있었다. 이때, 그림 6에 보인 바와 같이 가청 주파수 대역에서 100dB 이상의 동적 범위와 $\pm 5 \times 10^{-4}$ dB 이하의 리플 크기를 보장함을 확인하였다. 즉, sinc 함수를 이용한 기법에서는 계수를 저장하기 위해 7209(81×89)

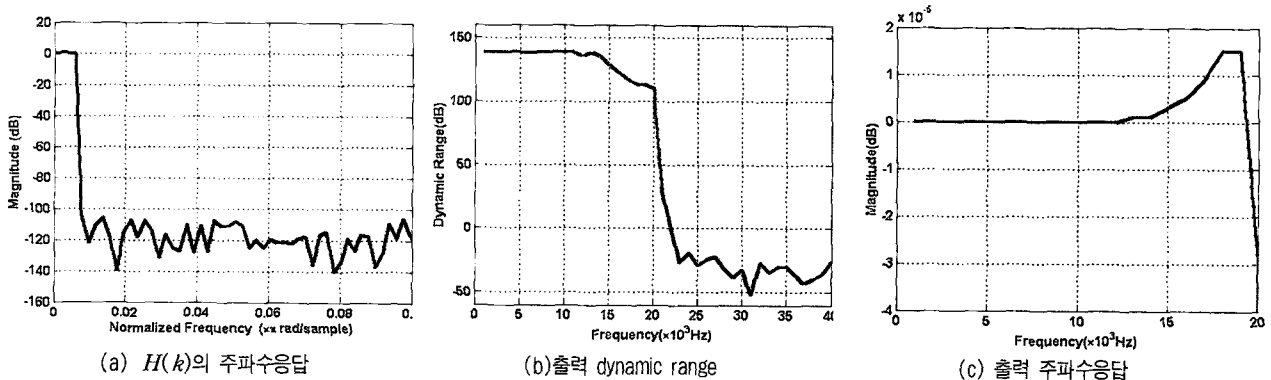
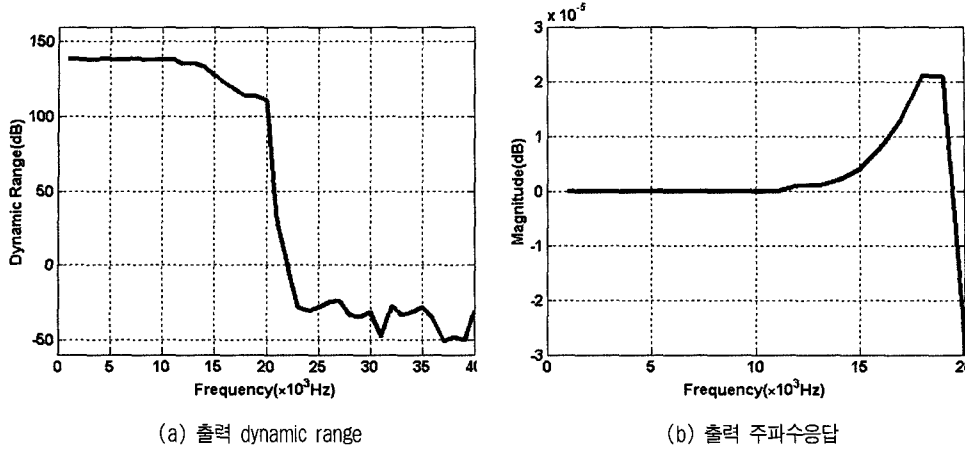


그림 5. 기본적인 기법의 성능
Fig. 5. Performance of the basic method



(a) 출력 dynamic range (b) 출력 주파수응답
 그림 6. Sinc 함수를 이용한 기법의 성능
 Fig. 6. Performance of the sinc function based method

words의 메모리가 필요하고, 출력 한 샘플 당 곱셈 89회, 덧셈 88회가 요구된다.

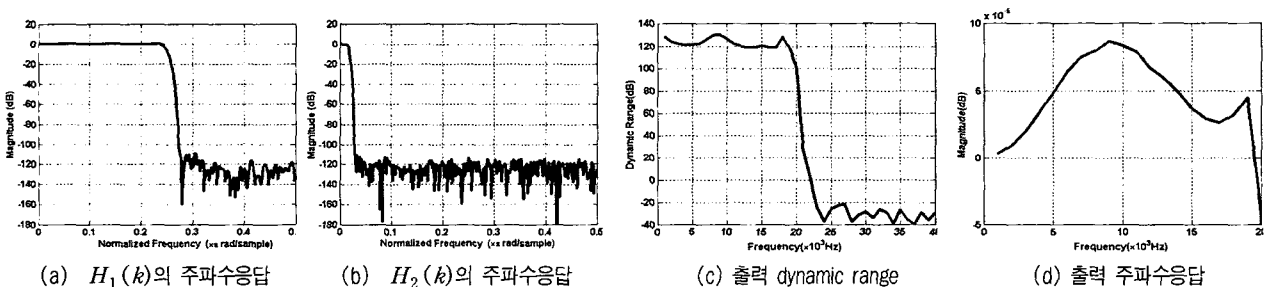
한편, 그림 1의 기본적인 기법을 다위상 구조로 나타내면, 그림 2의 sinc 함수를 이용한 기법의 성능과 매우 유사함을 알 수 있다. 기본적인 기법의 다위상 구조는 하나의 필터를 0이 아닌 데이터의 분포에 따라 다위상 필터로 분해한 것임에 반해, sinc 함수를 이용한 기법은 이 다위상 분해를 비 정수 값으로 지연된 필터군으로 보았다는 점이 다르다. 두 기법에서의 차이는 구현된 필터의 특성에 따른 차이로 보인다.

3. 다단계 구조를 이용한 기법

다단계 구조를 이용한 기법이 본 논문에서 요구하는 성능을 만족시키려면, 차단 주파수가 $\pi/4$ 인 $H_1(k)$ 은 327

taps, 차단 주파수가 $\pi/147$ 인 $H_2(k)$ 는 781 taps이라는 필터 길이가 요구된다. 두 필터의 주파수응답은 그림 7 (a)와 (b)에 제시하였고, 이때 그림 7 (c)와 (d)에 제시한 바와 같이 100dB 이상의 동적 범위와 $\pm 5 \times 10^{-4}$ dB 이하의 리플 크기라는 요구 조건을 만족시킴을 알 수 있다.

이 구조에서 $H_1(k)$ 및 $H_2(k)$ 는 대칭 필터라는 점을 고려하면, 메모리는 555 words가 소요된다. 그리고 각각을 다위상 구조로 구현하면, 147샘플의 입력에 대해 $H_1(k)$ 를 포함하는 4배 up-sampling 과정에서 (327×147)회 곱셈과 (323×147)회 덧셈을 통해 (147×4)샘플을 얻는다. 그런 다음, $H_2(k)$ 를 포함하는 40/147 재표본화 과정에서 출력 한 샘플 당 20회 곱셈과 19회 덧셈을 통해 160 샘플을 얻는다. 따라서, 다단계 구조를 이용한 기법은 출력 한 샘플 당 곱셈 약 321회와 덧셈 약 316회의



(a) $H_1(k)$ 의 주파수응답 (b) $H_2(k)$ 의 주파수응답 (c) 출력 dynamic range (d) 출력 주파수응답
 그림 7. 다단계 구조를 이용한 기법의 성능
 Fig. 7. Performance of the multistage method

표 1. 표본화율 변환 기법별 메모리 및 계산량

Table 1. Comparison of the 4 sampling rate converters in terms of memory size and computational complexity

	메모리(word) (24bit=1 word)	계산량(곱셈)
기본적인 기법	7121	89
sinc 함수	7209	89
다단계	555	321
B-spline	381	226
	101	263

연산이 필요하다. 즉, 다단계 구조를 이용한 기법은 기본적인 기법 혹은 sinc 함수를 이용하는 기법에 비해 계산량은 약 3.6배로 증가하나, 메모리 요구량은 약 7.7%만을 사용한다.

4. B-spline을 이용한 기법

Cucchi등^[3]에서는 필터의 길이가 각각 194, 24인 $H_1(k)$, $H_2(k)$ 를 이용하였지만, 그림 8(c)와 그림 8(d) 같이 가청 주파수 대역에서 100dB 이상의 동적 범위와 리플 $+5 \times 10^{-4}$ dB 이하의 성능을 얻기 위해서는 $H_1(k)$, $H_2(k)$ 각각 165, 35 taps로 구현되어야 하고, B-spline은 280(40×7)개의 계수가 필요하다. 메모리는 4배 up-sampling에서 101 words와 B-spline에서 280 words가 소요된다. 계산량은 4배 up-sampling에서 4개의 샘플을 얻는데, 곱셈 235회, 덧셈 229회 요구된다. 그리고 B-spline에서 출력 한 샘플 당 7회의 곱셈과 6회의 덧셈이 필요하다. 즉, B-spline을 이용한 기법에서는 총 381

words의 메모리와 출력 한 샘플 당 곱셈 226회, 덧셈 217회의 연산이 요구된다.

만약 B-spline에서 계수를 연산을 통해 얻는다면 곱셈 37회, 덧셈 15회가 추가로 요구되고^[3], B-spline 계수를 저장하기 위한 메모리는 부가적으로 필요 없다. 따라서 이러한 경우 메모리는 101 words, 그리고 출력 샘플 당 곱셈 263회, 덧셈 232회의 연산이 필요하다. 즉, 메모리 요구량과 계산량 등을 참조하여 B-spline 계수의 계산여부를 결정할 수 있을 것이다.

IV. 결론

이상에서 살펴본 4가지의 표본화율 변환 기법들에 대해서 메모리 요구량 및 계산량의 측면에서 비교하면 표 1과 같다. 각각의 값은 출력 한 샘플을 얻는데 필요한 메모리와 계산량이며, B-spline의 경우는 계수들을 메모리에 저장하고 있을 때를 의미한다.

각각의 기법들에 대해서 24비트를 1 word로 보았을 때, 표 1과 같이 B-spline의 경우가 가장 적은 381 words가 필요한 것을 볼 수 있다. 그러나, 계산량 측면에서는 기본적인 기법과 sinc 함수를 이용한 기법이 다단계 기법 및 B-spline을 이용한 기법에 비해 약 3.0 및 3.6배 정도로 우수한 것으로 나타났다. 이러한 점을 참조할 때, 계산량이 중요시되는 대부분의 응용에서는 다위상 형태로 구현되는 기본적인 구조 혹은 sinc 함수를 이용한 기법을 사용하여 표본화율 변환기를 구현하는 것이 유리할 것으로 판단된다. 그러나 메모리 요구량이 중요시되는 응용에서는 B-spline을 이용한 기법이 적합할 것으로 판단된다.

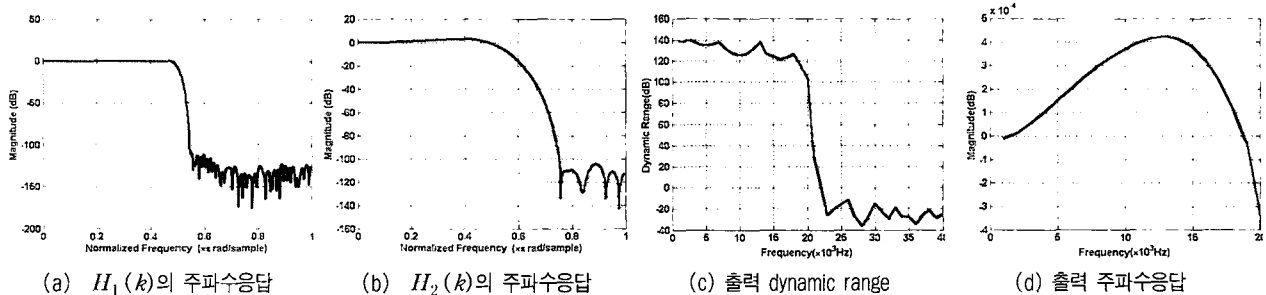


그림 8. B-spline 구조를 이용한 기법의 성능
Fig. 8. Performance of the B-spline based method

참 고 문 헌

- [1] A.V. Oppenheim, R.W. Schaffer and J.R. Buck, Discrete-Time Signal Processing - Second edition, Prentice Hall, 1998.
- [2] S. Park, "A Real-Time Method for Sample-Rate Conversion from CD to DAT," IEEE 1990 International Conference on Consumer Electronics, pp. 360~361, 1990.
- [3] S. Cucchi, F. Desinan, G. Parladori and G. Sicuranza, "DSP Implementation of Arbitrary Sampling Frequency Conversion for High Quality Sound Application," International Conference on Acoustics, Speech, and Signal Processing, vol. 5, pp. 3609~3612, 1991.
- [4] P.P. Vaidyanathan, "Multirate Digital Filters, Filter Banks, Polyphase Networks, and Applications: A Tutorial," Proceeding of the IEEE, vol. 78, no. 1, Jan. 1990.
- [5] 이용희, 김인철, "오디오 신호를 위한 표본화율 변환 알고리즘 성능 비교", 대한전자공학회 하계학술대회 논문집, 제25권, 제1호, 187~190쪽, 2002년 6월.
- [6] Release Notes for Release 13, MathWorks.

저 자 소 개



이 용 희

- 2002년 서울시립대학교 전자전기컴퓨터공학부 학사 졸업.
- 2004년 서울시립대학교 전자전기컴퓨터공학부 석사 졸업.
- 2004년 LG전자기술원 Mobile Multimedia 연구소 재직
- 주관심분야: 디지털 신호처리, 음향 신호처리



김 인 철

- 1985년 서울대학교 제어계측공학과 학사 졸업.
- 1987년 서울대학교 제어계측공학과 석사 졸업.
- 1992년 서울대학교 제어계측공학과 박사 졸업.
- 1992~1994년 대우전자(주) 영상연구소 선임연구원
- 1994~1999년 한성대학교 정보공학과 조교수
- 1999~현재 서울시립대학교 전자전기컴퓨터공학과 부교수
- 주관심분야 : 오디오 및 비디오 신호처리, 실시간 신호처리