

음성회의 시스템에서의 동적인 에러복구 기법

윤성덕, 안종석
동국대학교 컴퓨터공학과

A Dynamic Error Correction Technique in Audio Conference Systems

Sung-Duck Yun and Jung-Suk Ahn
Dept. of Computer Engineering, Dongguk University

요 약

최근 Mbone을 통한 다양한 회의 시스템들이 등장하고 있다. 이들 회의 시스템은 주로 영상회의 시스템과 음성회의 시스템으로 이루어져 있으며, 사람들은 영상보다 음성에 더 민감하기 때문에 음성품질이 영상품질보다 더 중요하다. 현재 인터넷은 QoS(Quality Of Service)를 제공하지 않기 때문에, 인터넷을 사용하는 음성회의 시스템들도 근본적으로 품질을 보장할 수 없다. 따라서 음성품질을 향상시키기 위한 다양한 방법들이 연구되었고, 대표적인 기법으로 중복정보(redundant information) 복구기법이 있다. 본 논문에서는 기존의 중복정보 복구기법을 개선한 동적 중복정보 복구기법을 제안한다. 이 기법은 네트워크의 상태에 따라 중복정보의 양을 동적으로 조절하기 때문에 중복정보의 효율을 향상시킬 수 있다. 실제 인터넷을 통한 실험에서 평균 9%의 성능향상이 있었다.

1. 서 론

패킷 스위치 네트워크를 통한 음성 데이터의 전송에 관한 연구는 70년대 말부터 80년대 초까지 활발히 연구가 되었다[1]. 이러한 연구들은 대부분 단일 네트워크상에서 패킷 스위칭을 통한 음성과 문자 데이터의 전송에 관계되는 것이었다. 현재 인터넷의 발달에 의한 증가된 대역폭, 하드웨어적인 지원으로 음성 데이터 전송에 관한 연구는 다시 활발히 진행되고 있다. 특히, Mbone을 통한 회의시스템이 광범위하게 사용되고 있다.

회의 시스템은 주로 영상회의 시스템과 음성회의 시스템으로 이루어져 있으며, 사람들은 영상보다 음성에 더 민감하기 때문에 음성품질이 영상품질보다 더 중요하다. 현재 vat[2]과 Nevot[3]등과 같은 음성회의 시스템이 개발되었고, 진화 정도의 품질(toll-quality)을 제공하려 하고 있다. 그러나 현재 인터넷은 시간에 따라 패킷손실과 지연이 변하는 동적인 서비스를 제공하기 때문에 인터넷을 사용하는 음성회의 시스템들은 근본적으로 품질을 보장할 수 없다.

따라서 음성품질을 향상시키기 위한 연구가 활발히 진행되었다.

음성품질은 패킷손실(packet loss)과 지터(jitter)에 의해 크게 좌우되며, 특히 음성회의 시스템은 실시간성을 제공해야 하므로 지연(delay)의 발생도 최소화시켜야 한다. 음성품질을 향상시키는 대표적인 기법으로 중복정보 복구기법이 있다. 이 기법은 손실된 패킷을 중복정보를 사용하여 복구하는 것이다. 그러나 기존의 이 기법은 네트워크의 동적인 특성을 반영하지 않았기 때문에 불필요하게 대역폭을 낭비하였다. 즉 패킷손실이 많은 구간과 패킷손실이 적은 구간을 동일하게 처리하여 대역폭을 낭비하고, 대역폭의 낭비는 다시 패킷손실을 유발시킨다. 따라서 본 논문에서는 네트워크의 상태에 따라 중복정보의 양을 동적으로 조절하여 중복정보의 효율을 향상시키는 기법을 제안한다.

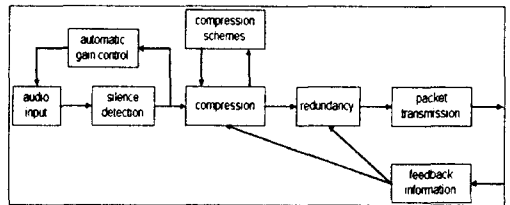
본 논문의 구성은 다음과 같다. 2장에서는 음성회의 시스템의 구조에 대하여 설명하고, 3장에서는 음성회의 시스템의 패킷손실 복구기법들에 대하여 설명한다. 4장에서는 패킷손실의 특징을 분석한다. 그리고 5장에서는 동적 중복정보 복구기법에 대하여 제안하고, 기존의 중복정보 복구기법과 성능을 비교한다. 6장에서는 결론과 향후 연구과제에 대하여 기술한다.

2. 음성회의 시스템의 구조

2.1 송신자 부분

음성회의 시스템은 그림 1과 같은 송신자(Sender) 구조를 가지고 있다. 하드웨어를 통해 들어온 음성 데이터는 침묵탐지(silence detection)를 거친다. 침묵탐지는 사용자가 대화하고 있을 경우에만 음성 데이터를 생성시켜 불필요한 데이터를 전송하지 않게 한다. 침묵탐지에 사용되는 기준값은 자동 입력값 조절(automatic gain control)에서 음성 입력값의 크기를 반영하여 결정한다. 불필요한 데이터를 제거하더라도 음성 데이터는 크기가 크므로 다양한 압축방식(PCM, ADM, GSM, LPC)으로 압축되어 전송된다. 일반적으로 압축률이 높을수록 음성품질이 저하되는데 표 1

은 압축방식에 따른 사용 대역폭이다[4]. 중복(redundancy)을 통해 이전 패킷에 대한 중복정보를 첨가하여 전송하고, 송신자는 수신자가 보내주는 피드백 정보(feedback information)를 받아 네트워크의 상태를 파악하여 압축방식과 중복방식을 결정한다. 음성 데이터는 실시간 전송 프로토콜인 RTP(Real-time Transport Protocol)를 사용하여 전송되며, RTP는 헤더부분에 시간 정보와 순서정보를 갖고 있다. 수신자는 순서정보를 사용하여 패킷이 역순으로 전달되더라도 올바른 순서로 정렬할 수 있으며, 시간정보를 사용하여 영상 데이터와의 동기화와 재생시간을 결정할 수 있다.



[그림 1] 음성회의 시스템의 송신자 부분

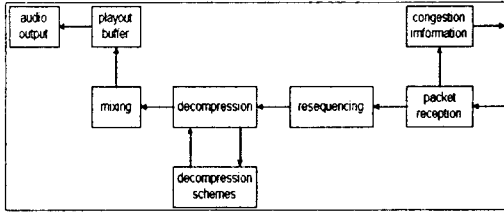
압축방식	사용 대역폭
PCM	64kbps
ADM4	32kbps
GSM	13kbps
LPC	4.8kbps

[표 1] 압축 방식에 따른 대역폭

2.2 수신자 부분

음성회의 시스템은 그림 2와 같은 수신자(Receiver) 구조를 가지고 있다. 네트워크를 통해 전달된 RTP패킷은 순서정보를 가지고 있어서, 수신자는 패킷손실을 판단할 수 있다. 그리고 수신자는 일정한 간격으로 RTCP(RTP Control Protocol)를 이용하여 송신자에게 이러한 혼잡정보(congestion information)를 전달한다. 수신된 패킷은 순서맞춤(resequencing)을 통해 올바른 순서로 정렬되며, 패킷손실이 있으면 중복정보를 사용하여 복구한다. 압축방식에 따라 음성 데이터를 복호(decompression)하여 원신호로 복원한 후, 믹싱(mixing)을 하게 된다. 믹싱은 여러

명의 송신자가 있을 경우, 동시에 같이 이야기를 하는 것처럼 들려준다.



[그림 2] 음성회의 시스템의 수신자 부분

3. 음성회의 시스템의 패킷손실 복구기법들

음성회의 시스템에서 음성 데이터의 손실을 복구하기 위한 방법으로 다음과 같은 3가지 방법들이 연구되었다.

첫째, ARQ(Automatic Repeat Request)는 패킷손실이 발생한 경우, 수신자의 재전송 요청에 의해 송신자가 재전송을 하는 방법이다. 그러나 이 방법은 지연을 증가시켜 실시간성을 저해하기 때문에 음성회의 시스템에 적용할 수 없으며, 멀티캐스팅 환경에 적용하기 어렵다는 문제점이 있다.

둘째, FEC(Forward Error Correction)는 복구정보를 사용하는 방법이다. 이 방식은 n개의 패킷을 전송할 때마다 n+1번째 패킷에 n개의 패킷을 exclusive-or한 내용을 전송하는 방법이다. 이것은 간단한 방법이지만 n+1번째 패킷은 복구만을 위한 패킷이므로 전송율을 1/n만큼 감소시킨다. 그리고 1개의 패킷손실에 대해서만 복구할 수 있다. 그러나 더욱 중요한 문제는 n+1번째 복구패킷이 도착한 후에야 복구가 가능하여 지연을 너무 증가시킨다는 것이다.

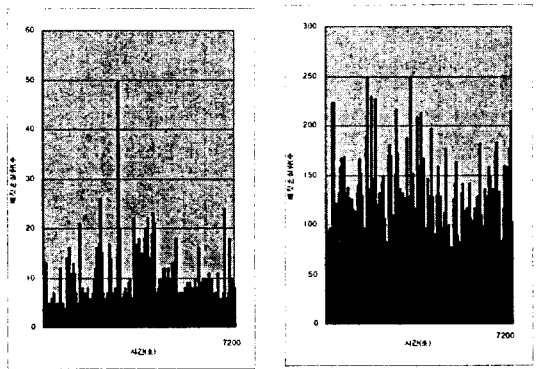
셋째, 중복정보 복구기법은 패킷을 전송할 때마다 이전 패킷에 대한 중복정보를 첨가하여 전송하는 방법이다. 음성회의 시스템에서는 음성 데이터를 다양한 압축방식을 사용하기 압축하기 때문에, 중복정보도 다양한 압축방식을 사용하여 만들어질 수 있다. 보통 대역폭 절약을 위해 현재

내용은 고품질의 압축방식(PCM, ADM)을 사용하여 압축하고, 중복정보는 저품질의 압축방식(GSM, LPC)을 사용하여 압축한 후, 하나의 패킷으로 전송한다.

4. 패킷손실 분석

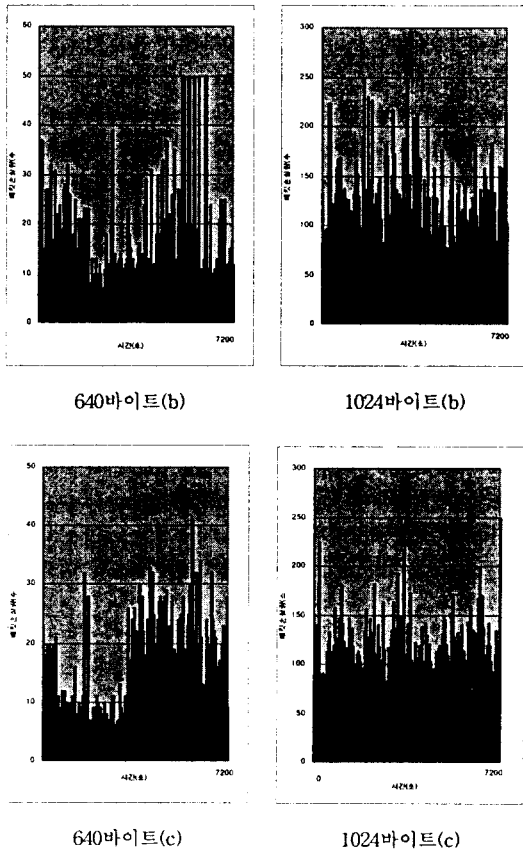
네트워크의 특성을 파악하는 연구는 네트워크 응용프로그램을 효율적으로 구현하기 위한 기초가 된다. 본 논문에서는 시간에 따른 패킷손실량과 패킷손실의 발생주기를 분석하여 음성 데이터의 복구를 위한 기초를 제공하고자 한다.

패킷손실에 대한 특성을 분석하기 위한 실험환경은 다음과 같다. 부산대에 있는 SUN UltraSparc2(CPU 200MHz, 256M 주메모리)에서 송신자 프로그램을 실행시키고, 동국대에 있는 SUN Sparc5에서 수신자 프로그램을 실행시켰다. 동시에 동국대 SUN Sparc5에서 tcpdump를 사용하여 패킷을 받아 분석하였다. 서로 다른 패킷크기와 전송률이 미치는 영향을 알아보기 위해, RTP에 640바이트의 데이터를 100ms간격으로 전송한 실험과 1024바이트의 데이터를 20ms간격으로 전송한 실험을 하였다. 또한 시간에 따른 네트워크의 상태변화가 미치는 영향을 알아보기 위해 오전 9~11시(a), 오후 1~3시(b), 오후 7~9시(c)에 실험을 하였다. 다음 그림은 5초간격으로 패킷손실량을 측정하여 나타낸 것이다.



640바이트(a)

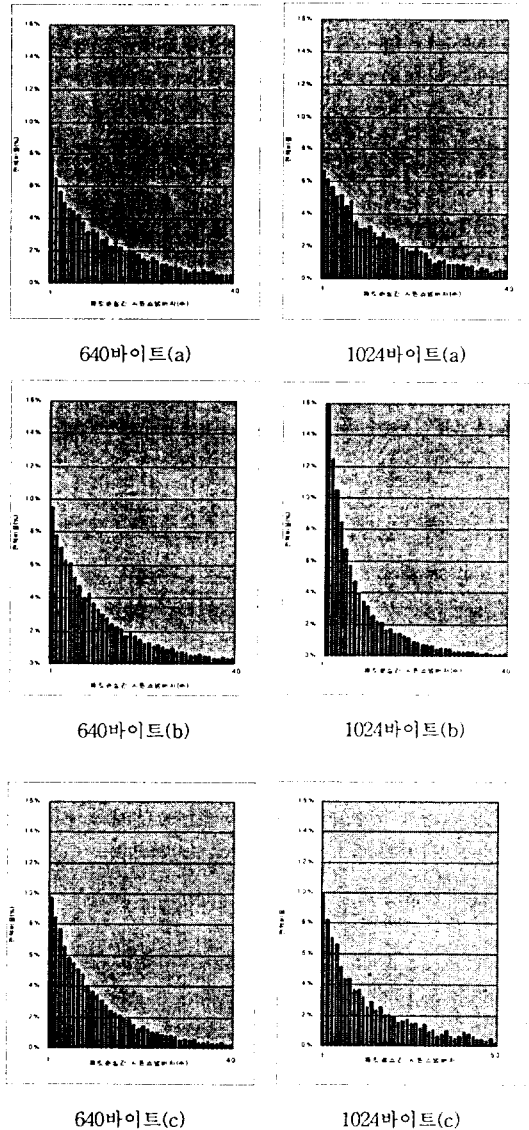
1024바이트(a)



[그림 3] 패킷손실량 측정실험

실험에서 640바이트는 5초당 50개의 패킷을 전송하고 1024바이트는 250개의 패킷을 전송한다. 그런데 전송량을 초과한 패킷손실도 극소수 일어난다. 즉 5초 간격의 측정 결과 각각 50, 250개 이상의 패킷손실이 발생한 것이다. 이는 라우터의 버퍼를 초과하여 다량의 패킷이 손실된 경우인 것 같다. 그림 3을 보면 패킷손실량이 급격하게 변하는 경우도 있으나 대부분 이전에 발생한 패킷손실량과 비슷하다는 것을 알 수 있다. 따라서 패킷손실량이 적은 적은 구간과 패킷손실량이 많은 구간의 특성을 이용하면 효율적인 복구방안을 구현할 수 있을 것이다.

다음은 패킷손실의 주기, 즉 패킷손실 후 다음번 패킷손실이 얼마나 떨어져서 발생하는지를 알아보기 위해 패킷손실간의 시퀀스번호를 분석하여 비율로 나타낸 것이다.



[그림4] 패킷손실간의시퀀스번호차

그림 4는 패킷손실의 발생주기를 나타낸 것으로, 패킷손실이 발생한 직후에 패킷손실이 많이 일어난다는 것을 나타낸다. 즉 패킷손실이 발생한 후 연속된 10개 이내에서 전체 패킷손실의 50%, 20개 이내에 70%가 발생한다. 따라서 전송자는 패킷손실이 발생하면 중복정보를 많이 첨가하고, 패킷손실이 발생하지 않으면 중복정보를 적게 첨가하여 효율적인 복구를 할 수 있다.

5. 동적인 에러복구 기법과 성능평가

중복 데이터를 사용하면 음의 품질은 향상되나 대역폭, CPU사용률, 지연 등이 늘어난다. 따라서 대역폭과 CPU사용률, 음의 품질향상 등을 고려한 효율적인 중복 데이터를 사용해야 한다. 그러나 같은 네트워크라 하더라도 시간에 따라 상태가 다양하게 변하기 때문에 기존의 정적인 중복정보 복구기법은 불필요한 대역폭을 사용할 수 있었다. 그리고 불필요한 대역폭의 사용은 네트워크를 혼잡상태로 만들 수 있다는 점에서 악영향을 끼친다.

패킷손실의 특성을 이용하여 패킷손실이 심한 구간은 중복정보를 많이 사용하여 전송하고, 손실이 적은 구간은 중복정보를 적게 사용하면, 효율적인 복구와 대역폭의 낭비를 막을 수 있다. 본 논문에서는 다음과 같은 3가지 방식으로 동적 중복정보 복구기법을 실험하였다.

첫째, NACK를 이용한 방법이다. 패킷손실이 발생하면 수신자는 즉시 송신자에게 NACK를 전송하고, 송신자는 NACK를 받아 10개의 패킷에는 중복정보를 2개 간격으로, 10~20개의 패킷에는 3개 간격으로, 그 이후로는 중복정보를 사용하지 않는다.

둘째, ACK를 이용한 방법이다. 10개의 연속된 패킷에서 패킷손실이 발생하지 않은 경우에만 수신자가 송신자에게 ACK를 전송하고, 송신자는 10개의 패킷을 전송하는 동안 ACK를 받지 못하면 중복정보를 2개 간격으로, ACK를 받으면 중복정보를 3개 간격으로 사용한다. 연속으로 ACK를 받으면 더 이상 중복정보를 사용하지 않는다.

셋째, 실시간 회의 시스템에서 널리 사용되는 RTCP를 이용한 방법이다. 수신자는 RTCP를 이용하여 네트워크 상태를 5초 간격으로 송신자에게 전달한다. 송신자는 RTCP정보 중 패킷손실율(packetloss)을 이용하여 다음과 같은 로우패스필터(low-pass filter)를 이용하여 중복정보의 사용주기를 조절한다. 이 방식은 패킷손실율을 누적하여 일시적인 패킷손실율 급변에 의한 정보주기의 변경을 막는다.

$rate = (0.75 * rate) + (0.25 * packetloss);$
 rate가 5% 이하면 중복정보를 사용하지 않고, 5~10%면 5개 간격으로, 10~15%면 4개 간격으로, 15~20%면 3개 간격으로, 20% 이상이면 2개 간격으로 중복정보를 사용하였다.

음성패킷은 PCM 데이터 640바이트를 기본정보로 하고, GSM 데이터 132바이트를 중복정보로 사용하였다. 아래의 표는 각 방식의 성능을 비교한 표이다.

	중복율	패킷분실율	복구율	중복율/복구율
오전9~11	25%	6.0%	18.2%	0.73
오후1~3	25%	13.1%	22.0%	0.88
오후7~9	25%	9.5%	17.7%	0.71

[표 2] 주기가 4인 정적인 중복정보 사용방법

	중복율	패킷분실율	복구율	중복율/복구율
오전9~11	20.2%	5.3%	16.2%	0.80
오후1~3	35.3%	14.2%	31.9%	0.90
오후7~9	33.4%	9.8%	24.4%	0.73

[표 3] NACK를 이용한 동적인 중복정보 사용방법

	중복율	패킷분실율	복구율	중복율/복구율
오전9~11	24.6%	5.7%	20.9%	0.85
오후1~3	35.8%	13.0%	33.4%	0.93
오후7~9	36.1%	9.6%	26.2%	0.73

[표 4] ACK를 이용한 동적인 중복정보 사용방법

	중복율	패킷분실율	복구율	중복율/복구율
오전9~11	10.8%	5.9%	9.0%	0.83
오후 1~3	25.1%	14.4%	24.7%	0.98
오후 7~9	24.8%	10.0%	19.2%	0.77

[표 5] RTCP를 이용한 동적인 중복정보 사용방법

표 5에서 성능을 분석하기 위해 복구율/중복율을 계산하였다. 이 값이 클수록 적은 중복정보로 많은 수의 패킷손실을 복구할 수 있다. 위 표에서 모든 동적인 방법이 정적인 방법보다 효율이 높은 것으로 나타났다. 기존의 정적 복구기법에 비해 NACK를 이용한 기법은 4%, ACK를 이용한 기법은 6% 그리고 RTCP를 이용한 기법은 9% 정도의 성능향상이 있었다. NACK와 ACK를 이용한 기법이 RTCP를 이용한 기법보다 성능이 좋지 않는 이유는 NACK와 ACK의 손실과 지연에 의해 패킷손실을 전송자가 즉각적으로 알기 어렵다는 것과 패킷손실에 너무 민감하게 반응하는 것 때문이다.

6. 결론 및 향후 연구과제

회의 시스템에서 음성품질이 영상품질보다 중요한 이유는 사람들이 음성에 더 민감하기 때문이다. 따라서 음성패킷의 손실을 복구하는 방법들이 연구되었고, 대표적인 기법으로 중복정보 복구기법이 있다. 그런데 중복정보는 패킷손실이 발생한 경우에만 사용되는 정보로서, 패킷손실이 없는 경우에는 대역폭을 낭비하며, 이러한 낭비는 오히려 네트워크를 혼잡상태로 만들 수 있다.

동일한 네트워크라 하더라도 시간에 따라서 상태는 다양하게 변하기 때문에 기존의 정적 중복정보 복구기법은 불필요하게 대역폭을 낭비한다. 따라서 네트워크의 상황에 따라 중복정보량을 조절하면 불필요한 대역폭 사용을 막고, 패킷손실을 효율적으로 복구할 수 있다. 본 논문에서는 NACK 사용기법과 ACK 사용기법, RTCP 사용기법으로 네트워크의 상태에 따라 중복정보량을 동적으로 조절하였다. 기존의 정적 복구기법에 비해 NACK 사용기법은 4%, ACK 사용기법은 6%, RTCP 사용기법은 9%의 성능향상이 있었다. 따라서 현재 음성회의 시스템에 널리 사용되는 RTCP를 이용하여 동적으로 중복율을 조절하면 성능을 향상시킬 수 있다.

본 논문에서는 5초간격의 RTCP정보를 이용하고 있으나 수신자가 네트워크 상태를 파악하여 송신자에게 전달하기 위한 시간으로는 너무 큰 시간이다. 따라서 RTCP전송

시간도 동적으로 변화시켜 효율을 향상시키는 연구를 계속할 것이다. 또한 NACK와 ACK를 이용한 방식에서 패킷손실에 너무 민감하게 반응하지 않으면서 NACK와 ACK의 수를 많이 줄이는 연구를 계속할 것이다.

[참고문헌]

- [1] C. Weinstein, J. Forgie, "A dynamic rate control mechanism for source in packet networks", IEEE JSAC, vol. 1, no. 6, pp 963-980, Dec. 1983
- [2] V. Jacobson, S. MacCanne, "vat", Maunal Pages, Lawrence Laboratory, University of California, Berkeley, CA.
- [3] H. Schulzrinne, "Voice Communication Across the Internet : A Network Voice Terminal", Research Report, Dept. of Electrical Engineering, University of Massachusetts at Amherst, July 92.
- [4] Jean-Chrysostome Bolot, Andres Vega-Garcia, "Control Mechanisms for Packet Audio in the Internet"
- [5] J-C. Bolot, "End-to-end packet delay and loss behavior in the Internet", Proc. ACM Sigcomm '93, San Francisco, CA, pp 189-199, Aug. 1993.
- [6] J-C. Bolot, A. Vega Garcia, "The case for FEC-based error control for packet audio in the Internet" to appear in ACM Multimedia Systems.
- [7] F. Bonomi, D. Mitra, J. B. Seery, "Adaptive algorithms for feedback-based flow control in high speed, wide area ATM networks", IEEE JSAC, vol. 13, no. 7, pp. 1267-1283, Sept. 1995.