

상 기법(sender-based repair technique)과 수신측 보상 기법(receiver-based repair technique)이 있다[3].

2.1 송신측 보상기법

Forward Error Correction(FEC) 기법[5]은 하나의 스트림에 손실된 패킷을 복구할 수 있는 정보를 추가 시켜서 전송하는 기법이다. 이러한 기법은 전송되는 미디어 종류와는 무관하게 적용시킬 수 있고 상대적으로 적은 계산시간으로 손실 복구 패킷을 생성할 수 있지만 스트림을 구성하는 n개의 패킷 중 하나의 패킷 손실에 대해서만 복구가 가능하고 전송 지연과 필요한 대역폭이 증가되는 단점이 있다.

Interleaving 기법[6]은 오디오 데이터내의 패킷들을 일정한 간격으로 재배치하여 전송한 후 수신측에서 다시 원래의 순서대로 복원시키는 기법이다. 이 기법은 패킷 손실 효과를 분산시켜준다. Interleaved stream 으로부터 한 패킷의 손실은 재구성된 스트림내에서 다수의 적은 겂(gap)을 나타내게 된다. 이 기법은 대역폭의 증가가 필요하지 않은 장점이 있지만 상대적으로 높은 지연시간(latency)을 지니는 단점이 있다.

2.2 수신측 보상기법

수신측 보상기법은 손실된 패킷에 대해 원래의 패킷과 유사한 패킷을 만들어서 대체시키는 기법이다. 이것은 오디오 신호의 경우 서로 이웃한 패킷들은 유사한 특성을 지니고 있다는 점을 이용한 것이다. 이 기법은 크게 네가지로 나눌 수 있다[3].

삽입(Insertion)방법은 손실된 패킷을 묵음이나 잡음, 또는 손실 직전의 패킷으로 대체하는 기법이다. 이 기법은 구현은 간단하나 음질이 상당히 떨어진다.

보간(Interpolation) 기법은 손실된 패킷과 유사하다고 예상되는 대체 패킷을 생성하기 위해 손실 주위에 있는 패킷과 패턴비교나 보간을 수행한다. 이러한 기법은 삽입방법보다 계산량은 많지만 신호의 변화 특성을 고려할 수 있기 때문에 삽입방법보다 나은 음질을 가진다. 본 논문에서 적용한 SOLA 알고리즘은 이 부류에 속한다.

재생성(Regeneration)방법은 손실된 패킷의 앞과 뒤 패킷의 정보를 이용하여 손실 패킷의 대체 패킷을 생성한다. 이 기법은 삽입이나 보간 기법보다 계산량은 더 많이 요구되지만 상대적으로 더 나은 음질을 얻을 수 있다.

이 밖에도 Adaptive playout buffer algorithm[4]이 있는데 이 기법은 각 talkspurt에 대한 playout delay D를 선택하여 각 talkspurt의 playout 시점을 $t_i + D$ 로 조정하는 기법이다. 여기서 t_i 는 패킷의 생성시간, D는 하나의 talkspurt내의 모든 패킷들의 생성시간과 playout 시간간의 차이를 말한다. D가 증가하면 패킷들이 늦게 도착하여 손실로 처리되는 부분은 줄어들지만 중단간에 발생하는 지연시간은 커지게 된다. 따라서 이러한 알고리즘은 손실과 지연시간사이에 최적으로 동작할 수 있는 버퍼링 시간을 계산하여 최적 지점에 가깝게 playout delay를 조정한다.

3. SOLA를 이용한 음질향상 기법

SOLA 기법[7]은 음성신호의 시간축 변환을 위해

프레임 단위의 신호들을 상호 상관 함수(normalized cross-correlation function)를 이용하여 축소 또는 확장하고 동기를 맞춘 후 중첩 가산(overlap and add)하는 방법을 사용한다.

입력 프레임 $x(m)$ 을 변환 비율 a 에 의해 변환시켜 시간축으로 확장 또는 축소 변환된 프레임 $y(m)$ 을 얻을 경우 수행되는 과정은 그림 1과 같다. 첫번째 과정은 입력 프레임 $x(m)$ 을 다시 Sa 간격으로 시간축 변환에 이용할 패킷들로 나눈다. 두번째 과정은 Sa 간격으로 나누어진 패킷 중 첫번째 패킷을 $y(m)$ 으로 가져온다. 세번째 과정은 $x(m)$ 의 다음 패킷을 가져와서 $y(m)$ 에 Ss 간격으로 패킷들을 위치시킨 후 서로간의 동기가 최대한 일치하는 지점을 얻기 위해 상호 상관 함수를 이용하여 동기가 일치하는 지점 Km 을 찾는다. 마지막으로 동기가 일치하는 지점으로 재배치한 후 중첩되는 구간(Lm)을 중첩가산한다. 합성과정은 매 패킷당 이루어지고 새로운 패킷이 이전 과정에서 합성된 프레임 $y(m)$ 에 더해진다. 이 때 Sa 와 Ss 사이에는 $Ss = a Sa$ 라는 관계가 성립한다. 따라서 a 가 1보다 클 경우 확장의 의미를 가지고 1보다 적을 경우에는 축소한다는 의미를 지닌다.

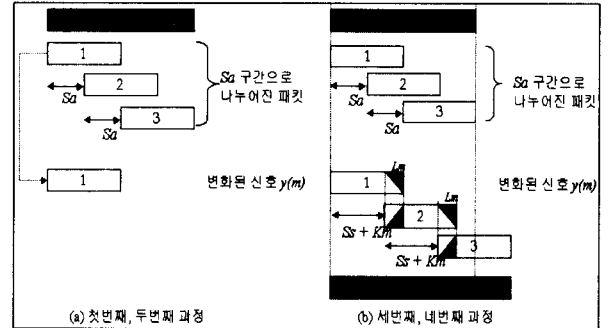


그림 1. SOLA 수행 과정

본 논문에서는 전송지연으로 인한 패킷 손실을 줄이기 위해 수신측에 전송된 패킷에 변환비율 $a = 1.25$ 로 SOLA 알고리즘을 적용하여 패킷을 확장하였다. 적용 알고리즘은 그림 2와 같다.

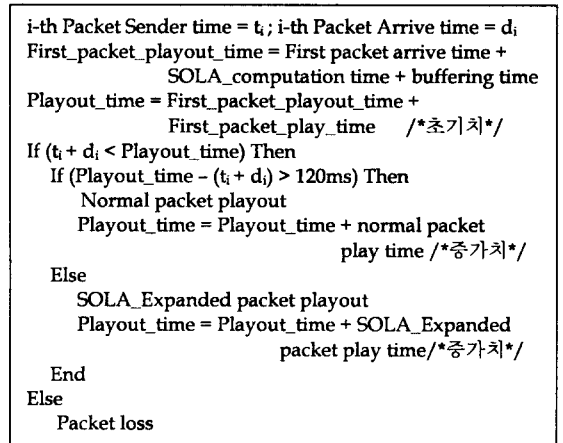


그림 2. SOLA 기법을 이용한 playout 시간 보상 알고리즘

송신측에서 40ms단위로 샘플링된 패킷을 전송하면 수신측에서는 수신패킷에 SOLA 알고리즘을 적용하여 수신패킷을 사람이 인지할 수 있을 정도까지 확장한다. 확장 후 발생하는 여유시간을 이용하여 전송지연으로 인한 패킷손실을 감소시킨다. 원음에 비해 과다하게 확장되는 것을 방지하기 위해 여유시간의 상태에 따라(Playout_time - (t+d_i) > 120ms) 패킷에 대한 SOLA 적용 여부를 결정한다.

그림 3에서는 각 패킷이 수신측에 수신된 후 playout되는 과정을 나타내었다. 도착즉시 playout할 경우 5번, 9번, 10번 패킷이 손실되지만 2 packet delay 시킨 후 playout할 경우 손실되는 패킷은 없다. 하지만 이 경우 2 packet size만큼 중단간에 지연이 발생된다. 하지만 SOLA를 적용할 경우 지연도 줄이면서 음성의 통화품질을 높일 수 있다.

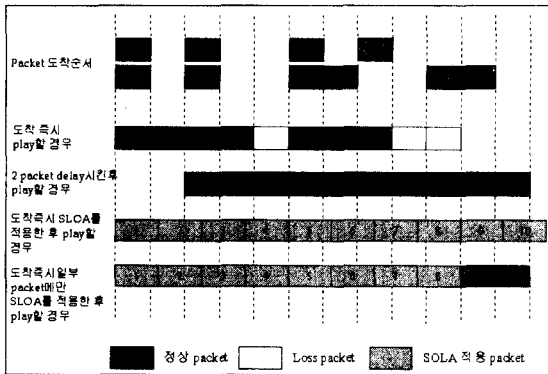


그림 3. SOLA를 이용한 Playout 과정 예

4. 시뮬레이션 및 고찰

본 논문에서 제시한 알고리즘을 평가하기 위한 시뮬레이터는 그림 4와 같다.

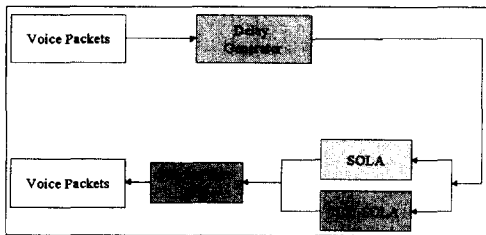


그림 4. 시뮬레이터 구성도

Voice Packet에서는 8KHz 8-bit 모노로 샘플링 된 패킷을 40ms 간격으로 나누어서 전송한다. Delay Generator에서는 40ms 간격으로 분할된 패킷을 데이터로 사용하여 Ping 프로그램을 이용한다. 수신된 데이터와 시간정보를 이용하여 수신된 패킷에 SOLA를 적용할 것인지 아닌지를 여유시간의 상태에 따라 결정한 후 패킷의 playout 시간을 조절한다.

본 논문에서는 SOLA 알고리즘을 이용하여 전송지연으로 인한 패킷 손실을 최소화하는 방법으로 기존의 묵음대체 기법 및 직전 패킷 대체 기법과 음질 비교를 수행하였다.

청취평가는 비교대상이 되는 음성을 어떤 기법이 사람이 듣기에 더 명료한 소리를 나타내는가를 판단하는 형태로 진행되었다. 30명의 청취자가 실험에 참여하였으며, 청취자 자신이 어떤 음성 데이터가 어떠한 방법으로 복원된 신호인지를 알 수 없도록 음성 데이터 순서를 무작위로 제시하였다. 그림 5는 각 복원 기법으로 복원된 음성 데이터를 청취한 결과를 MOS로 나타낸 것인데 기존 기법보다 우수한 음질을 나타내고 있다.

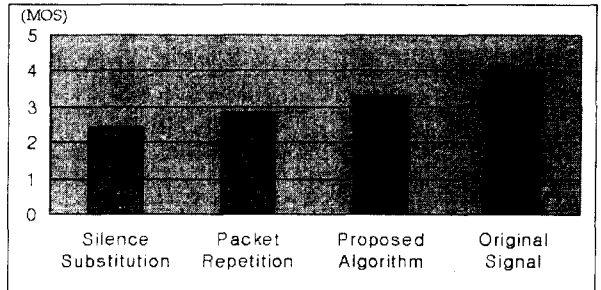


그림 5. 각 기법들의 Mean Opinion Score(MOS)

5. 결론

인터넷폰에서 전송지연에 의한 음성패킷 손실은 음질에 매우 심각한 영향을 야기시킨다. 본 논문에서는 수신단에서 전송지연으로 인한 패킷 손실을 줄이기 위해 수신된 패킷에 SOLA 알고리즘을 적용하여 각 패킷의 playout 시간을 사람이 인지할 수 있는 범위내에서 확장시켰다. 이에 따라 확장된 시간 여유만큼 전송지연으로 인한 패킷 손실을 최대한 줄일 수 있게 된다. 또한 여유시간의 양에 따라 SOLA 기법의 적용 여부를 결정하기 때문에 원음에 비해 음질은 약간 떨어지지만 인지하는 데는 큰 영향을 주지 않는다.

앞으로는 네트워크 트래픽 상황에 따라 변환비율을 가변적으로 적용하여 음질을 보다 더 개선시키는 방식에 대해 계속 연구할 계획이다.

6. 참고 문헌

- [1] Thomas J. Kostas, Michael S. Borella, Ikhlaz Sidhu, Guido M. Schuster, Jacek Grabciec, and Jerry Mahler, "Real-Time Voice Over Packet-Switched Networks", IEEE Network, Jan/Feb 1998
- [2] J.C.Bolot, "Characterizing End-to-End Packet Delay and Loss in the Internet", Journal of High-Speed Networks, vol.2, No.3, pp.305-323, Dec. 1993
- [3] C.S. Perkins, O.Hodson and V.Hardman, "A Survey of Packet Loss Recovery Techiques for Streaming Audio", IEE Network Magazine, Sep./Oct. 1998
- [4] R. Ramjee, J. Kurose, D. Towsley, H. Schulzrinne, "Adaptive Playout Mechanisms for Packetized Audio Applications in Wide-Area Networks", IEEE Infocom, June 1994
- [5] J. Rosenberg and H. Schulzrinne, "An RTP payload format for generic forward error correction", IETF Audio/Video Transport WG, July 1998
- [6] J.L.Ramsey, "Realization of optimum interleavers", IEEE Trans. Info. Theory, Vol. IT-16, May 1970
- [7] S.Roucos and A.M.Wilgud, "High Quality Time-Scale Modification for Speech", Proc. ICASSP, pp. 493-496, Apr. 1986