(19)대한민국특허청(KR)
(12) 등록특허공보(B1)

(51) 。Int. Cl.
G10L 19/00 (2006.01)

(45) 공고일자 2006년10월16일
(11) 등록번호 10-0633720
(24) 등록일자 2006년10월04일

(21) 출원번호 10-2000-7014191
(22) 출원일자 2000년12월14일
번역문 제출일자 2000년12월14일
(86) 국제출원번호 PCT/US2000/010467
국제출원일자 2000년04월19일

(65) 공개번호 10-2001-0052855
(43) 공개일자 2001년06월25일
(87) 국제공개번호 WO 2000/63881
국제공개일자 2000년10월26일

(81) 지정국 국내특허 : 캐나다, 일본, 대한민국, 멕시코, 미국,

EP 유럽특허 : 오스트리아, 벨기에, 스위스, 독일, 덴마크, 스페인, 프랑스, 영국, 그리스, 아일랜드, 이탈리아, 룩셈부르크, 모나코, 네덜란드, 포르투칼, 스웨덴, 핀란드, 사이프러스,

(30) 우선권주장 60/130,016 1999년04월19일 미국(US)

(73) 특허권자 에이티 앤드 티 코포레이션
미국 뉴욕 10013-2412 뉴욕 애비뉴 오브 디 아메리카즈 32

(72) 발명자 캐필로우데이빗에이.
미국,뉴저지07922,버클리헤이츠,워싱톤스트리트122

(74) 대리인 이병호
정상구
신현문
이범래

심사관 : 장호근

(54) 패킷 손실 또는 프레임 삭제 은폐를 수행하는 방법 및 장치

요약

본 발명은 내장 또는 표준 FEC 처리를 갖지 않는 음성 코더(speed coder)에서 패킷(packet) 손실 또는 프레임 삭제 은폐(Frame Erasure Concealment, FEC)를 수행하기 위한 방법 및 장치에 관한 것이다. 복호기를 갖춘 수신기는 부호기로부터 전송된 압축 음성 정보의 부호화된 프레임들을 수신한다. 수신기에 있는 손실 프레임 검출기는 부호화된 프레임이 전송시 손실되거나 또는 변형되었는지, 또는 삭제되었는지의 여부를 결정한다. 부호화된 프레임이 삭제되지 않은 경우, 부호화된 프레임은 복호기에 의해 복호화되고 일시 메모리(temporary memory)는 복호기의 출력으로 업데이트된다. 미리 결정된 지연 주기가 적용되고 그 후 오디오 프레임이 출력된다. 손실 프레임 검출기가 부호화된 프레임이 삭제된 것을 결정하면, FEC 모듈은 신호에 프레임 은폐 처리를 적용한다. FEC 처리는 삭제된 프레임들에 대해 자연스럽게 들리는 합성 음성을 산출한다.

대표도

도 1

색인어

패킷 손실, 프레임 삭제 은폐 처리, 음성 코더, 손실 신호 검출기, FEC 모듈, 오버랩 가산 동작

명세서

기술분야

본 비임시 출원은 여기서 참고로 포함되고 1999년 4월 19일 출원된 미국 임시 출원 제 60/130,016 호의 이점을 주장한다. 또한, 다음의 문서는 여기서 참고로 포함된다: ITU-T Recommendation G.711 - Appendix 1, "G.711에서 패킷 손실 은폐를 위한 고품질의 저복잡성 알고리즘(A high quality low complexity algorithm for packet loss concealment with G.711)" (9/99) 및 전기 통신에 대한 미국 국제 표준 - ITU-T Recommendation G.711과 사용되는 패킷 손실 은폐 (T1.521-1999).

발명의 분야

본 발명은 패킷(packet) 손실 또는 프레임 삭제 은폐(Frame Erasure Concealment; FEC)를 수행하기 위한 기술에 관한 것이다.

배경기술

관련된 기술의 설명

프레임 삭제 은폐(FEC) 알고리즘들은 입력 음성 신호가 전송기에서 부호화 및 패킷화되고 네트워크(임의의 종류의)를 통해 전달되어, 패킷을 복호화하고 음성 출력을 재생하는 수신기에서 수신되는 음성 통신 시스템에서 전송 손실들을 감춘다. G.723.1, G.728, 및 G.729와 같이, 음성 코더들에 기초한- 많은 표준 CELP는 표준들로 제안되거나 내장된 FEC 알고리즘들을 갖는다.

FEC의 목적은 수신된 비트-스트림(bit-stream)에서 손실된 데이터를 커버하도록 합성 음성 신호를 발생하는 것이다. 이상적으로, 합성 신호는 손실 신호와 똑같은 음질 및 스펙트럼 특성들을 가지므로, 자연스럽지 않은 인위 구조들을 생성하지 않는다. 음성 신호들이 때로 부분적으로 정적이므로, 손실된 세그먼트(segment)에 적절한 근사치를 발생시키기 위해 신호들의 과거 히스토리를 사용하는 것이 가능하다. 삭제들이 그렇게 길지 않고, 신호가 신속하게 변하고 있는 영역에 삭제가 위치하지 않으면, 삭제들은 은폐 이후에 알아들을 수 없다.

이전 시스템들은 예를 들어 디.제이. 굿맨(D.J. Goodman) 등의 패킷 음성 통신들에서 손실 음성 세그먼트들을 회복하기 위한 파형 대용 기술들(Waveform Substitution Techniques for Recovering Missing Speech Segments in Packet Voice Communications), 통권. 34, 제 6 호, 음향 상태, 음성, 및 신호 처리 1440 내지 48(1996년 12월)에서 IEEE 트랜스.(IEEE Trans.) 및 오.제이.와셈(O.J. Wasem)등의 PCM 패킷 통신들의 품질에 대한 파형 대용의 효과(The Effect of Waveform Substitution on the Quality of PCM Packet Communications), 통권. 36, 제 3 호, 음향 상태, 음성, 및 신호 처리에서 IEEE 트랜잭션들(IEEE Transactions) 342-48(1988년 3월)과 같이, 프레임 삭제들을 은폐하기 위해 피치(pitch) 파형 복제 기술을 사용하였다.

표준들-기초한 (그리고 비-표준-기초한) 음성 코딩 시스템들 중 어떤 복호기들은 상태 정보를 유지한다. 즉, 그들은 미래에 복호화 동작들에 사용되는 어떤 신호들의 값들을 저장하는 메모리를 포함한다. 자체적으로 통합된 FEC 기술을 포함하는 그들의 코딩 시스템들에 대해, 상태 정보는 FEC 기술들이 손실 패킷들에 대하여 보상하기 위해 적용될 때 적절하게 조정된다.

발명의 상세한 설명

본 발명의 발명자는 음성 코딩 시스템의 복호기가 상태 정보를 유지하지만, 복호기를 사용하여 통합되지 않는 FEC 기술(예를 들어, 복호기를 수행되는 FEC 처리)을 사용하기를 시도하는 경우, 상태 정보는 FEC에 따라 적절하게 조절되지 않을 것이고 복호화된 음성 신호에 대해 문제들을 야기시킬 것을 이해했다. 본 발명은 상태 정보를 유지하기 위해 복호기들을 가지는 음성 코딩 시스템들을 위한 음성 코딩 시스템 처리를 위한 패킷 손실 또는 프레임 삭제 은폐(FEC)를 수행하기 위한

기술에 관한 것이다. 본 발명은 비유용한 패킷들에 대응하는 음성 신호의 합성에 관한 것이며 그 후 합성된 음성 신호를 부호화하는 것에 관한 것이다. 그 후 부호화 처리를 반영하는 신호들은 복호기 상태 정보를 유지하는데 사용하기 위해 복호기에 제공된다.
본 발명은 후속하는 도면들에 관련하여 자세하게 설명되며, 동일한 번호들은 동일한 요소들을 참조한다.

도면의 간단한 설명

도 1은 전형적인 오디오 전송 시스템을 도시하는 도면.

도 2는 G.711 코더 및 FEC 모듈을 갖춘 전형적인 오디오 전송 시스템을 도시하는 도면.

도 3은 FEC 기술을 사용하여 출력 오디오 신호를 도시하는 도면.

도 4는 삭제의 종료시 오버랩-가산(overlap-add; OLA) 동작을 도시하는 도면.

도 5는 G.711 코더를 사용하여 FEC를 실행하는 전형적인 처리의 순서도.

도 6은 히스토리 버퍼(history buffer)의 업데이팅 처리를 도시하는 그래프.

도 7은 제 1 프레임의 신호를 은폐하는 전형적인 처리의 순서도.

도 8은 자기-상관 관계(auto-correlation)로부터 피치(pitch) 추정을 도시하는 도면.

도 9는 미세한(fine) 피치 추정 대 거친(coarse) 피치 추정들을 도시하는 도면.

도 10은 피치 및 최종 1/4(lastquarter) 버퍼들에서 신호들을 도시하는 도면.

도 11은 단일-주기 피치 버퍼를 사용한 합성 신호 발생을 도시하는 도면.

도 12는 제 2 또는 최종 삭제된 프레임의 신호를 은폐하기 위한 전형적인 처리의 순서도.

도 13은 제 2 삭제된 프레임에 이어지는 합성된 신호들을 도시하는 도면.

도 14는 2-주기 피치 버퍼를 사용하여 합성 신호 발생을 도시하는 도면.

도 15는 제 2 삭제된 프레임의 시작에서 OLA를 도시하는 도면.

도 16은 삭제된 이후에 제 1 프레임을 처리하는 전형적인 방법의 순서도.

도 17은 3-주기 피치 버퍼를 사용하여 합성 신호 발생을 도시하는 도면.

도 18은 다른 음성 코더들로 FEC 기술들을 사용을 도시하는 블록도.

실시예

최근에는 POTS(Plain-Old-Telephony Service)를 지지하는 보장된 품질의 서비스 없이 패킷(packet) 네트워크들에서 G.711을 사용하는 것에 많은 관심이 있다. 이러한 네트워크에서 프레임 삭제들(또는 패킷 손실들)이 일어날 때, 은폐 기술들이 필요로 되거나 또는 통화의 품질이 심각하게 저하된다. 고품질의 저복잡성 프레임 삭제 은폐(Frame Erasure Concealment; FEC) 기술은 개발되어 있고, 이후에 상세히 설명된다.

FEC를 갖춘 오디오 시스템의 전형적인 블록도는 도 1에 도시된다. 도 1에서, 부호기(110)는 입력 오디오 프레임을 수신하고, 코드화된 비트-스트림(bit-stream)을 출력한다. 비트-스트림은 어떤 프레임들이 손실되었는지의 여부를 결정하는 손실 프레임 검출기(115)에 의해 수신된다. 손실 프레임 검출기(115)가 프레임들이 손실된 것을 결정하는 경우, 손실 프레임 검출기(115)는 FEC 알고리즘이나 손실 프레임들을 재구성하는 처리를 적용하도록 FEC 모듈(130)에 신호를 전한다.

그러므로, FEC 처리는 입력 신호가 전송기에서 부호화되어 패킷화되고, 네트워크를 통해 전달되고, 프레임이 손실되었음을 결정하는 손실 프레임 검출기(115)에서 수신되는, 오디오 시스템에서 전송 손실들을 감춘다. 도 1에서는 손실 프레임 검출기(115)가 기대되는 프레임이 사용되기에 도착되지 않았나, 또는 너무 늦게 도착되었나를 결정하는 방법을 갖는다고 가정된다. IP 네트워크들에서, 이는 일반적으로 전송 프레임에서 데이터에 시퀀스 번호나 시간스탬프(timestamp)를 부가함으로서 실행된다. 손실 프레임 검출기(115)는 도착 프레임들의 시퀀스 번호들을 프레임이 손실되지 않았다면 기대되는 시퀀스 번호들과 비교한다. 손실 프레임 검출기(115)는 기대될 때 프레임이 도착된 것으로 검출하면, 이는 복호기(120)에 의해 복호화되고 오디오의 출력 프레임은 출력 시스템에 주어진다. 프레임이 손실되면, FEC 모듈(130)은 대신하는 오디오의 합성 프레임의 값을 발생시킴으로서 손실 오디오 프레임을 감추는 처리를 적용한다.

G.723.1, G.728, 및 G.729와 같이, 많은 표준 ITU-T CELP-기초하는 음성 코더들이 복호기들에서 음성 재생을 모델화한다. 그러므로, 복호기들은 복호기에서 직접 FEC 처리를 통합하기에 충분한 상태 정보를 갖는다. 이들 음성 코더들은 그 표준들의 일부로 지정된 FEC 알고리즘들 또는 처리들을 갖는다.

비교하면, G.711은 음성 재생을 모델화하지 않는 샘플-대-샘플 부호화 구조이다. 코더에는 FEC를 보조하는 상태 정보가 없다. 그 결과로, G.711의 FEC 처리는 코더와 무관하다.

G.711 코더로 사용되는 시스템의 전형적인 블록도는 도 2에 도시된다. 도 1에서와 같이, G.711 부호기(210)는 비트-스트림 데이터를 부호화하여 손실 프레임 검출기(215)에 전송한다. 또, 손실 프레임 검출기(215)는 도착 프레임들의 시퀀스 번호들을 프레임들이 손실되지 않았다면 기대되는 시퀀스의 번호들과 비교한다. 프레임이 기대될 때 도착하면, 이는 복호기(220)에 의해 복호화되도록 전달되고, 그 후 히스토리 버퍼(history buffer)(240)로 출력되어 신호를 저장한다. 프레임이 손실되는 경우, 손실 프레임 검출기(215)는 대신하는 오디오의 합성 프레임의 값을 발생함으로서 손실 오디오 프레임을 감추는 처리를 적용하도록 FEC 모듈(230)에 알린다.

그러나, 손실 프레임들을 감추기 위해, FEC 모듈(230)은 어느 신호가 손실 프레임에 있어야 하는가를 추정하도록 히스토리 버퍼(240)에 의해 제공되는 복호화된 출력 신호의 과거 히스토리를 사용하는 G.711 FEC 처리를 적용한다. 부가하여, 삭제된 및 삭제되지 않은 프레임들 사이의 매끄러운 전이를 보장하도록, 지연 모듈(150)은 또한 미리 결정된 시간 주기, 예를 들면 3.75 msec만큼 시스템의 출력을 지연시킨다. 이러한 지연은 합성 삭제 신호가 삭제의 시작시 실제 출력 신호와 천천히 혼합되도록 허용한다.

FEC 모듈(230)과 히스토리 버퍼(240) 및 지연 모듈(250) 블록 각각 사이의 화살표들은 저장된 히스토리가 합성 신호를 발생하기 위해 FEC 처리에 의해 사용됨을 의미한다. 부가하여, FEC 모듈(230)의 출력은 삭제시 히스토리 버퍼(240)를 업데이트하는데 사용된다. FEC 처리는 단지 G.711의 복호화 출력에만 의존하므로, 그 처리는 음성 코더가 주어지지 않을 때와 같이 작용하게 됨을 주목하여야 한다.

입력 신호가 FEC 모듈(230)에서 FEC 처리에 의해 처리되는 방법의 그래픽 예가 도 3에 도시된다.

도면에 도시된 상단 파형은 남성 발성자로부터 나온 음성 영역에서 20 msec 삭제가 일어날 때 시스템으로의 입력을 도시한다. 그 아래 파형에서는 FEC 처리가 갭(gap)에 합성 음성을 발생함으로서 손실 세그먼트를 은폐하였다. 비교 목적들을 위해, 삭제되지 않은 원래의 입력 신호가 또한 도시된다. 이상적인 시스템에서, 은폐된 음성 사운드들은 원래 사운드와 똑같다. 도면으로부터 볼 수 있는 바와 같이, 합성 파형은 손실 세그먼트들에서 원래 파형과 매우 흡사하다. 이후에는 "입력" 파형으로부터 "은폐" 파형이 발생되는 방법이 논의된다.

FEC 모듈(230)에 의해 사용되는 FEC 처리는 히스토리 버퍼(240)에 저장된 음성에 유사한 특성을 갖는 합성 음성을 발생시킴으로서 손실 프레임을 은폐한다. 기본적인 개념은 다음과 같다. 신호가 목소리이면, 신호가 반주기적이고 국부적으로 정적이라 가정한다. 피치를 추정하고 히스토리 버퍼(240)에서 최종 피치 주기를 수회 반복한다. 그러나, 삭제가 길거나 피치가 짧으면(주파수가 높으면), 똑같은 피치 주기를 너무 여러번 반복하는 것은 자연스러운 음성과 비교해 너무 조화적인(harmonic) 출력을 만들게 된다. 삑하는 소리들 및 퉁하는 소리들과 같이 들을 수 있는 조화적인 인위 구조들을 방지하기 위해, 히스토리 버퍼(240)로부터 사용되는 피치 주기들의 수는 삭제 과정들의 길이만큼 증가된다. 짧은 삭제들은 합성 신호를 발생하는데 히스토리 버퍼(240)로부터의 최종 또는 마지막 몇 개의 피치 주기들만을 사용한다. 긴 삭제들은 또한 히스토리 버퍼(240)에서 더 뒤에 있는 피치 주기를 사용한다. 긴 삭제들에서, 히스토리 버퍼(240)로부터의 피치 주기들은 원래 음성에서 일어난 것과 똑같은 순서로 재생되지 않는다. 그러나, 테스트에서, 긴 삭제들에서 발생된 합성 음성 신호가 아직까지 자연스러운 사운드를 만드는 것으로 밝혀졌다.

삭제가 더 길면 길수록, 합성 신호가 실제 신호에서 멀어질 가능성이 더 높아진다. 특정한 종류의 사운드를 너무 오래 간직하여 발생되는 인위 구조들을 방지하기 위해, 합성 신호는 삭제가 더 길이짐에 따라 감쇄된다. 기간 10 msec 이하의 삭제들 동안에는 감쇄가 필요없다. 10 msec 이상의 삭제들에서는 합성 신호가 추가 10 msec 당 20%의 비율로 감쇄된다. 60 msec를 넘어서면, 합성 신호는 0(묵음)으로 설정된다. 이는 합성 신호가 원래 신호에 너무 비슷하지 않아 평균적으로 60 msec 이후에 손실 음성을 은폐하도록 계속 시도하는 것은 유리하기보다 불리해지기 때문이다.

다른 소스들로부터의 신호들 사이에 전이기 이루어질 때마다. 전이가 불연속성, 클릭들과 같은 가청음, 또는 부자연스러운 인위 구조들을 출력 신호에 도입시키지 않는 것이 중요하다. 이러한 전이는 여러 위치에서 일어난다:

1. 합성 신호의 시작과 최종 양호한 프레임의 끝부분 사이의 경계에서 삭제의 시작시.

2. 합성 신호와 삭제 이후 처음 양호한 프레임내 신호의 시작 사이의 경계에서 삭제의 종료시.

3. 히스토리 버퍼(240)로부터 사용되는 피치 주기들의 수가 신호 변화를 증가시키도록 변화될 때마다.

4. 히스토리 버퍼(240)의 반복부들 사이의 경계들에서.

매끄러운 전이들을 보장하기 위해, 모든 신호 경계에서는 오버랩 가산(OLA, Overlap Add)이 실행된다. OLA는 한 엣지에서 오버랩되는 두 신호들을 매끄럽게 조합하는 방법이다. 신호들이 오버랩되는 영역에서, 신호들은 윈도우(window)에 의해 가중화되고, 함께 부가(혼합)된다. 윈도우들은 특정한 샘플에서 가중치들의 합이 1이 되도록 설계된다. 즉, 신호들의 전체적인 합에는 이득이나 감쇄가 적용되지 않는다. 부가하여, 윈도우들은 신호가 좌측에서 가중치 1로 시작되어 점차 0으로 페이드 아웃(fade out)되고, 우측에서 신호가 가중치 0으로 시작되어 점차 가중치 1로 페이드 인(fade in)되도록 설계된다. 그래서, 오버랩 윈도우의 좌측 영역에서는 좌측 신호만이 주어지고, 오버랩 윈도우의 우측 영역에서는 우측 신호만이 주어진다. 오버랩핑 영역에서, 신호는 점차적으로 좌측 신호에서 우측 신호로 전이된다. FEC 처리에서, 삼각형 윈도우들은 가변 길이 윈도우들을 계산하는 복잡성을 낮게 유지하는데 사용되지만, 해닝 윈도우(Hanning window)들과 같은 다른 윈도우들이 대체 사용될 수 있다.

도 4는 삭제가 끝난 이후 시작되는 실제 음성과 OLA 처리된 20-msec 삭제의 종료시 합성 음성을 도시한다. 한 예에서, OLA 가중화 윈도우는 5.75 msec의 삼각형 윈도우이다. 상단 신호는 삭제하는 동안 발생된 합성 신호이고, 그 아래의 오버랩 신호는 삭제 이후의 실제 음성이다. OLA 가중화 윈도우는 신호 아래 도시된다. 여기서는 삭제시 실제 신호에서의 피치 변화로 인해, 합성 신호와 실제 신호의 피크가 정합되지 않고, OLA 없이 신호들을 조합하도록 시도하는 경우 도입되는 불연속성은 "OLA 없이 조합"이라 칭하여진 그래프에 도시된다. "OLA 없이 조합"된 그래프는 OLA 윈도우의 시작까지 합성 신호를 복사하고 나머지는 실제 신호를 복사함으로서 생성되었다. OLA 동작들의 결과는 경계에서의 불연속성들이 얼마나 매끄러워졌는가를 도시한다.

상기 논의는 설명된 처리가 정적 음성과 작용되는 것에 관심을 두지만, 음성이 신속하게 변하거나 목소리가 아니면, 음성은 주기적인 구조를 갖지 않을 수 있다. 그러나, 이러한 신호들은 이후 설명될 바와 같이, 똑같은 방법으로 처리된다.

먼저, 본 실시예에서 피치 측정시 허가되는 가장 작은 피치는 5 msec이고, 이는 200 Hz의 주파수에 대응한다. 일부 고주파수 여성 및 어린이 발성자들이 200 Hz 이상의 기본 주파수를 갖는 것으로 알려져 있으므로, 윈도우들이 비교적 크게 유지되도록 이를 200 Hz로 제한한다. 이 방법에서는 10 msec 삭제된 프레임내에서 선택된 피치 주기가 최대 2회 반복된다. 고주파수 발성자들인 경우에는, 피치 추정기가 실제 피치 주기의 배수를 복귀시키므로 이것이 실제로 출력을 저하시키지 않는다. 또한, 음성을 너무 자주 반복시키지 않음으로서, 처리는 비주기적 음성 이외의 합성 주기적 음성을 생성하지 않는다. 두 번째로, 삭제가 더 길어짐에 따라 합성 주기를 발생하는데 사용되는 피치 주기들의 수가 증가되므로, 긴 삭제들 동안 주기성이 주어지지 않는 신호에는 충분한 분산이 부가된다.

음성의 시간 스케일링(scaling)을 위한 WSOLA(Waveform Similarity Overlap Add) 처리는 또한 똑같은 처리가 주기적 및 비주기적 음성 신호들을 모두 시간-스케일로 사용될 수 있도록 큰 고정 크기의 OLA 윈도우들을 사용함을 주목하여야 한다.

상기에는 FEC 처리의 개요가 주어졌고, 이후에는 각 단계가 상세히 논의된다.

이 논의의 목적을 위해, 예를 들면, 프레임은 10 msec의 음성을 포함하고, 샘플링 비율은 8 kHz인 것으로 가정한다. 그래서, 삭제들은 80개 샘플(8000 * .010 = 80)의 증가들로 일어날 수 있다. FEC 처리는 다른 프레임 크기들 및 샘플링 비율들에 쉽게 적용될 수 있음을 주목하여야 한다. 샘플링 비율을 변화시키기 위해서는 단지 msec로 주어진 시간 주기를 .001로 곱하고, 그 샘플링 비율로 적절한 버퍼 크기들을 정한다. 예를 들면, 히스토리 버퍼(240)는 음성의 마지막 48.75 msec를 포함한다. 8 kHz에서, 이는 버퍼가(48.75 * .001 * 800) = 390 샘플들의 길이임을 의미한다. 16 kHz 샘플링에서는 이것이 2배, 즉 780 샘플들이 된다.

몇 가지 버퍼 크기들은 처리에서 만날 것으로 기대되는 최저 주파수를 기초로 한다. 예를 들면, 설명되는 처리에서는 8 kHz 샘플링에서 보게 될 최저 주파수가 66 2/3 Hz인 것으로 가정한다. 이는 15 msec(1/(66 2/3) = .015)의 최대 피치 주기를 만든다. 히스토리 버퍼(240)의 길이는 최저 주파수 주기의 3.25 배이다. 그래서, 히스토리 버퍼(240)는 15 * 3.25 = 48.75 msec이다. 16 kHz 샘플링에서 입력 필터들이 50 Hz(20 msec 주기) 만큼 낮은 주파수들을 허용하는 경우, 히스토리 버퍼(240)는 20 * 3.25 = 65 msec로 연장되어야 한다.

프레임 크기도 또한 변화될 수 있다; 10 msec는 G.729와 같이, 몇 가지 표준 음성 코더들에 의해 사용되는 프레임 크기이므로 디폴트(default)로 선택되었고, 또한 몇 가지 무선 시스템들에서도 사용된다. 프레임 크기를 변화시키는 것은 간단하다. 원하는 프레임 크기가 10 msec의 배수이면, 처리는 변화하지 않고 유지된다. 단순히, 삭제 처리의 프레임 크기를 10 msec로 남겨두고, 이를 프레임 당 여러 번 통화한다. 원하는 패킷 프레임 크기가 5 msec와 같이 10 msec의 약수이면, FEC 처리는 기본적으로 변하지 않고 유지된다. 그러나, 피치 버퍼에서 주기들의 수가 증가되는 비율은 10 msec내 프레임들의 수를 근거로 수정되어야 한다. 12 msec와 같이, 10 msec의 배수들 또는 약수들이 아닌 프레임 크기들도 또한 수용될 수 있다. FEC 처리는 합리적으로 피치 버퍼로부터 사용되는 피치 주기들의 수에 대한 증가 비율을 변화시키는 것을 용인하고 있다. 10 msec마다 1회 이외에 12 msec마다 1회로 주기들의 수를 증가시키는 것은 그다지 차이를 만들지 않을 것이다.

도 5는 도 2의 실시예에 의해 실행되는 FEC 처리의 블록도이다. 주요 동작 중 일부를 실시하는데 필요한 서브 단계들은 도 7, 도 12, 및 도 16에서 더 상세히 설명되고, 이후 논의된다. 다음 논의에서는, 몇 가지 변수들은 값들 및 버퍼들을 유지하는데 사용된다. 이들 변수들은 다음과 같이 요약된다:

[표 1]

변수들 및 그들의 내용들

변수	타입	설명	코멘트
B	Array	피치 버퍼	범위[-P*3.25:-1]
H	Array	히스토리 버퍼	범위[-390:-1]
L	Array	최종 1/4 버퍼	범위[-P*.25:-1]
O	Scalar	피치 버퍼내의 오프셋	
P	Scalar	피치 추정	40 <= P < 120
P4	Scalar	1/4 피치 추정	P4 = P >> 2
S	Array	합성된 음성	범위[0:79]
U	Scalar	사용된 파장	1 <= U <= 3

도 5의 순서도에 도시된 바와 같이, 처리는 단계 505에서 시작되어, 다음 프레임이 손실 프레임 검출기(215)에 의해 수신된다. 단계 510에서, 손실 프레임 검출기(215)는 프레임이 삭제되는지 여부를 결정한다. 프레임이 삭제되지 않으면, 단계 512에서, 프레임은 복호기(220)에 의해 복호화된다. 이어서, 단계 515에서, 복호화된 프레임은 FEC 모듈(230)에 의해 사용하도록 히스토리 버퍼(240)에 저장된다.

히스토리 버퍼 업데이트 단계에서, 이 버퍼(240)의 길이는 기대되는 최장 피치 주기의 길이의 3.25배이다. 8 kHz 샘플링에서, 최장 피치 주기는 15 msec 또는 120 샘플들이므로, 히스토리 버퍼(240)의 길이는 48.75 msec 또는 390 샘플들이다. 그러므로, 각 프레임이 복호기(220)에 의해 복호화된 이후에, 히스토리 버퍼(240)는 최근 음성 히스토리를 포함하도록 업데이트된다. 히스토리 버퍼(240)의 업데이트는 도 6에 도시된다. 이 도면에 도시된 바와 같이, 히스토리 버퍼(240)는 가

장 최근 음성 샘플들을 우측에 포함하고 가장 오래된 음성 샘플들을 좌측에 포함한다. 복호화 음성의 가장 최근 프레임이 수신되면, 이는 우측에서 버퍼(240)로 쉬프트되고, 가장 오래된 음성에 대응하는 샘플들은 좌측에서 버퍼를 벗어나 쉬프트된다(도 6b를 참고).

부가하여, 단계 520에서는 지연 모듈(250)이 최장 피치 주기의 1/4 만큼 음성 출력을 지연시킨다. 8 kHz 샘플링에서, 이는 120 * 1/4 = 30 샘플들 또는 3.75 msec이다. 이 지연은 FEC 모듈(230)에 의해 생성된 합성 신호와 삭제 이전의 실제 신호 사이의 매끄러운 전이를 보장하도록 삭제의 시작에서 FEC 모듈(230)이 1/4 파장 OLA를 실행하도록 허용한다. 출력은 프레임을 복호화한 이후 다음 프레임이 삭제되는가 여부를 알지 못하기 때문에 지연되어야 한다.

단계 525에서는 오디오가 출력되고, 단계 530에서는 프레임들이 더 있는가를 결정한다. 프레임들이 더 이상 없으면, 처리는 종료된다. 프레임들이 더 있으면, 처리는 다음 프레임을 구하도록 단계 505로 되돌아간다.

그러나, 단계510에서, 수신된 프레임이 삭제된 것으로 손실 프레임 검출기(215)가 결정하면, 처리는 FEC 모듈(230)이 제 1 삭제된 프레임을 은폐시키는 단계 535로 진행되고, 그 처리는 이후 도 7에서 상세히 설명된다. 제 1 프레임이 은폐된 이후에, 단계 540에서는 손실 프레임 검출기(215)가 다음 프레임을 구한다. 단계 545에서, 손실 프레임 검출기(215)는 다음 프레임이 삭제되는가 여부를 결정한다. 다음 프레임이 삭제되지 않으면, 단계 555에서, FEC 모듈(230)은 삭제 이후 제 1 프레임을 처리하고, 그 처리는 이후 도 16에서 상세히 설명된다. 제 1 프레임이 처리된 이후에, 처리는 단계 530로 복귀되어, 손실 프레임 검출기(215)는 프레임들이 더 있는가 여부를 결정한다.

단계 545에서, 다음 또는 이어지는 프레임들이 삭제되는 것으로 손실 프레임 검출기(215)가 결정하면, FEC 모듈(230)은 도 12에서 상세히 설명되는 처리에 따라 제 2 및 이어지는 프레임들을 은폐한다.

도 7은 처음 10 msec의 삭제를 은폐하도록 취해지는 단계들을 설명한다. 이후에는 그 단계가 상세히 조사된다.

도 7에서 볼 수 있는 바와 같이, 단계 705에서, 삭제의 시작시 제 1 동작은 피치를 추정하는 것이다. 이를 위해, 표준화된 자기-상관 관계(auto-correlation)가 40 샘플들 내지 120 샘플들의 탭(tap) 지연에서 20 msec(160 샘플들) 윈도우로 히스토리 버퍼(240) 신호에 실행된다. 8 kHz 샘플링에서는 이러한 지연들이 5 msec 내지 15 msec의 피치 주기 또는 200 Hz 내지 66 2/3 Hz의 기본 주파수들에 대응한다. 자기-상관 관계의 주기에서 탭은 피치 추정 P이다. H가 이 히스토리를 포함하고, -1(삭제 이전의 우측 샘플)에서 -390(삭제가 시작되기 이전에 390개 샘플들의 샘플)으로 인덱스된다고 가정하면, 탭 j에 대한 자기 상관 관계는 수학적으로 다음과 같이 표시될 수 있다:

$$Autocor(j) = \frac{\sum_{i=1}^{160} H[-i]H[-i-j]}{\sqrt{\sum_{k=1}^{160} H^2[-k-j]}}$$

그래서, 자기 상관 관계의 피크 또는 피치 추정은 다음과 같이 표시될 수 있다:

$$P = \{\max_j(Autocor(j)) \mid 40 \le j \le 120\}$$

상술된 바와 같이, 5 msec 또는 40 샘플들로 허용된 최저 피치 주기는 단일 피치 주기가 10 msec 삭제된 프레임에서 최대 2회 반복되기에 충분하게 크다. 이는 목소리가 아닌 음성에서 인위 구조들을 방지하고, 또는 높은 피치의 발성자들에서 부자연스러운 조화적인 인위 구조들을 방지한다.

도 3의 삭제에 대해 표준화된 자기-상관 관계를 계산하는 그래프의 예가 도 8에 도시된다.

"히스토리(history)"라 칭하여진 파형은 삭제 직전에 히스토리 버퍼(240)의 내용들이다. 점선 수평선은 신호의 참고 부분, 즉 삭제 직전의 20 msec 음성인 히스토리 버퍼(240) H[-1]:H[-160]를 도시한다. 직선 수평선들은 40 샘플들(상단선, 5 msec 주기, 200 Hz 주파수) 내지 120 샘플들(하단선, 15 msec 주기, 66.66 Hz 주파수)의 탭에서 지연된 20 msec 윈도우들이다. 상관 관계의 출력은 또한 윈도우들의 위치들과 정렬되어 그려진다. 상관 관계에서 점선 수직선은 곡선의 피크로, 추정된 피치를 나타낸다. 이 선은 삭제의 시작으로부터 한 주기 뒤이다. 이 경우, P는 56 샘플들로, 7 msec의 피치 주기 및 142.9 Hz의 기본 주파수에 대응한다.

자기-상관 관계의 복잡성을 낮추기 위해, 두 가지 특수한 과정들이 사용된다. 이들 쇼트컷들은 출력을 현저하게 변화시키지 않지만, 처리의 전체적인 운행 시간 복잡성에 큰 영향을 갖는다. FEC 처리에서 대부분의 복잡성은 자기-상관 관계에 있다.

먼저, 모든 탭에서 상관 관계를 계산하기보다는, 추려진 신호에 피크의 거친 추정이 먼저 결정되고, 그 후 미세한(fine) 탐색이 거친 피크 주변에서 실행된다. 거친 추정을 위해, 2:1로 간추려진 신호에 작용하고 모든 다른 탭 마다 조사만을 하는 새로운 함수로 상기의 Autocor 함수를 정정한다:

$$Autocor_{rough}(j) = \frac{\sum_{i=1}^{80} H[-2i]H[-2i-j]}{\sqrt{\sum_{k=1}^{80} H^2[-2k-j]}}$$

$$P_{rough} = 2\{\max_j(Autocor_{rough}(2j)) \mid 20 \le j \le 60\}$$

거친 추정을 사용하면, 범위 $P_{rough} - 1 \le j \le P_{rough} + 1$에서만 원래의 탐색 처리가 반복된다. j가 40 샘플들 내지 120 샘플들 사이의 원래 범위에 머물도록 확인하는데 주위한다. 샘플링 비율이 증가되면, 데시메이션 비율도 증가되어, 처리의 전체적인 복잡성이 대략 일정하게 유지됨을 주목한다. 44.1 kHz로 샘플링된 음성에 8:1의 데시메이션 요소들을 가지고 테스트들이 실행되었고 양호한 결과들을 얻었다. 도 9는 $Autocor_{rough}$의 그래프를 Autocor의 그래프와 비교한다. 도면에서 볼 수 있는 바와 같이, $Autocor_{rough}$는 Autocor에 양호하게 근접하고, 복잡성은 8 kHz 샘플링에서 거의 4의 인수 만큼 - 모든 다른 탭만이 조사되기 때문에 2의 계수, 또한 미리 결정된 탭에서 모든 다른 샘플만이 조사되기 때문에 2의 계수 만큼 - 감소된다.

제 2 절차는 Autocor 및 $Autocor_{rough}$에서 에너지 계산의 복잡성을 낮추도록 실행된다. 각 단계에서 전체적인 합을 계산하기보다는 운행되는 에너지의 합이 유지된다. 즉:

$$Energy(j) = \sum_{k=1}^{160} H^2[-k-j]$$

then:

$$Energy(j+1) = \sum_{k=1}^{160} H^2[-k-j-1] = Energy(j) + H^2[-j-161] - H^2[-j-1]$$

그래서, 제 1 에너지 항이 계산된 이후에 FEC 처리의 각 단계에서는 에너지 항을 업데이트하는데 단지 2회 곱셈들 및 2회 합산들만이 필요하다.

이제는 피치 추정 P를 갖고, 파형은 삭제 동안 발생되기 시작한다. 다시 도 7의 순서도를 참고로, 단계 710에서, 최근 3.25 파장(3.25 * P 샘플)은 히스토리 버퍼(240) H에서 피치 버퍼 B로 복사된다. 피치 버퍼의 내용들은 최근 1/4 파장을 제외하고, 삭제 기간 동안 일정하게 유지된다. 한편, 히스토리 버퍼(240)는 삭제 동안 합성 음성으로 계속하여 업데이트된다.

단계 715에서, 히스토리 버퍼(240)로부터의 최근 1/4 파장(.25 * P 샘플)은 최종 1/4 버퍼(last quater buffer), L에 저장된다. 이 1/4 파장은 수회의 OLA 동작들 동안 필요하다. 편의상, 히스토리 버퍼(240)에서와 같이, B 및 L 버퍼들을 액세스하는데는 똑같은 음의 인덱싱 구조가 사용된다. B[-1]는 삭제 도착하기 이전의 최종 샘플이고, B[-2]는 그 이전의 샘플이다. 합성 음성은 0 이상으로 인데스 처리되는 합성 버퍼 S로 배치된다. 그래서, S[0]은 제 1 합성된 샘플이고, S[1]은 제 2 합성 샘플 등이다.

도 3의 삭제 동안 피치 버퍼, B, 및 최종 1/4 버퍼, L의 내용은 도 10에 도시된다. 이전 섹션에서, 주기, P는 56 샘플들로 계산되었다. 그래서, 피치 버퍼는 3.25 * 56 = 182 샘플의 길이다. 최종 1/4 버퍼는 .25*56=14 샘플들 길이이다. 도면에서, 수직선들은 삭제의 시작으로부터 뒤로 매 P 샘플들마다 배치된다.

삭제의 처음 10 msec 동안에는, 피치 버퍼로부터 최종 피치 주기만이 사용되므로, 단계 720에서는 U = 1이다. 음성 신호가 진짜 주기적이고 피치 추정이 단지 추정이 아니라 정확한 진짜 값이면, 파형을 피치 버퍼 B에서 합성 버퍼 S로 직접 복사할 수 있고, 합성 신호는 매끄럽고 연속적이 된다. 즉, S[0] = B[-P]이고, S[1] = B[-P+ 1] 등이다. 피치가 10 msec 프레임보다 짧으면, 즉 P < 80이면, 단일 피치 주기는 삭제된 프레임에서 1회 이상 반복된다. 본 예에서는 P = 56이므로, 복사가 S[56]까지 이른다. 샘플들 56 부근에서 샘플당 복사 시퀀스는: S[54] = B[-2], S[55] = B[-1], S[56] = B[-56], S[57] = B[-55] 등이다.

실제로, 피치 추정은 정확하지 않고 신호는 진짜 주기적이지 않을 수 있다. (a) 실제 신호와 합성 신호 사이의 경계 및(b) 주기가 반복되는 경계에서, 불연속성들을 방지하기 위해서는 OLA가 요구된다. 두 경계들에 대하여, 모두 실제 음성의 종료 B[-1]로부터 한 주기 이전의 음성 B[-P]까지 매끄러운 전이를 원한다. 그러므로, 단계 725에서, 이는 B[-P] 이전의 1/4 파장을 히스토리 버퍼(240)의 최종 1/4 파장, 또는 L의 내용들과 오버랩 가산(OLA)함으로서 달성될 수 있다. 그래프에서, 이는 피치 버퍼에서 최종 1 1/4 파장을 취하여, 이를 한 파장 우측으로 쉬프트하고, 1/4 파장 오버랩핑 영역에서 OLA를 실행하는 것과 동일하다. 단계 730에서, OLA의 결과는 히스토리 버퍼(240)에서 최종 1/4 파장으로 복사된다. 합성 파형의 추가 주기들을 발생시키기 위해, 피치 버퍼는 추가 파장들만큼 쉬프트되고, 추가 OLA들이 실행된다.

도 11은 처음 2회 반복들 동안의 OLA 동작을 도시한다. 이 도면에서, 모든 파형과 교차하는 수직선은 삭제의 시작부이다. 짧은 수직선들은 피치 마커(marker)들이고, 삭제 경계로부터 P 샘플들 위치에 놓인다. 파형들 "피치 버퍼"와 "P 만큼 우측으로 쉬프트" 사이의 오버랩핑 영역은 "P 만큼 우측으로 쉬프트"와 "2P 만큼 우측으로 쉬프트" 사이의 오버랩핑 영역과 정확하게 똑같은 샘플들에 대응한다. 그러므로, 1/4 파장 OLA만이 한번 계산될 필요가 있다.

단계 735에서, OLA를 먼저 계산하고 피치 버퍼의 최종 1/4 파장에 그 결과를 배치함으로서, 진짜 주기적 신호에 대해 합성 파형을 발생하는 처리가 사용될 수 있다. 샘플 B[-P]에서 시작되어, 단순히 샘플들을 피치 버퍼로부터 합성 버퍼로 복사하고, 피치 버퍼의 끝부분에 도달된 경우, 순환하는 피치 버퍼 포인터를 다시 피치 주기의 시작으로 보낸다. 이 기술을 사용하여, 임의의 기간의 합성 파형이 발생될 수 있다. 도 11의 "OLA들과 조합된" 파형에서 삭제 시작의 좌측까지의 피치 주기는 피치 버퍼의 업데이트 내용들에 대응한다.

"OLA들과 조합된" 파형은 단일 주기의 피치 버퍼가 불연속성들 없이 주기 P인 주기적 신호를 발생함을 설명한다. 히스토리 버퍼(240)에서 단일 파장으로부터 발생되는 이 합성 음성은 처음 10 msec의 삭제를 은폐하는데 사용된다. OLA의 효과는 "피치 버퍼" 및 "OLA들과 조합된" 파형에서 삭제가 시작되기 직전의 1/4 파장을 비교함으로서 관찰될 수 있다. 단계 730에서, "OLA들과 조합된" 파형의 이러한 1/4 파장은 히스토리 버퍼(240)에서 최종 1/4 파장을 또한 대체한다.

삼각형 윈도우의 OLA 동작은 또한 수학적으로 표시될 수 있다. 먼저, 변수 P4는 샘플들에서 피치 주기의 1/4인 것으로 정의된다. 그러므로, P4 = P >> 2이다. 본 예에서, P는 56이므로, P4는 14이다. OLA 동작은 그 때 범위 1 ≤ i ≤ P4에서 다음과 같이 표시될 수 있다:

$$B[-i] = \frac{i}{P4} L[-i] + \left(\frac{P4 - i}{P4}\right) B[-i - P]$$

OLA의 결과는 피치 버퍼 및 히스토리 버퍼(240)에서 최종 1/4 파장을 모두 대체한다. 히스토리 버퍼(240)를 대체함으로서, 히스토리 버퍼(240)는 또한 3.75 msec 만큼 출력을 지연시키므로, 히스토리 버퍼(240)가 업데이트될 때, 1/4 파장 OLA 전이가 출력된다. 처음 10 msec의 삭제시 출력 파형은 도 3의 "은폐된" 파형에서 처음 두 점선들 사이의 영역에서 관찰될 수 있다.

단계 740에서, 프레임에 대한 합성 음성을 발생하는 것이 종료시, 현재 오프셋은 피치 버퍼에 변수 O로 저장된다. 이 오프셋은 다음 프레임의 실제 또는 합성 신호를 가진 OLA를 위해 합성 파형이 다음 프레임으로 계속되게 허용한다. O는 또한 삭제가 10 msec를 넘어 확장되면 신호 위상이 유지되게 허용한다. 80 샘플 프레임들 및 P = 56인 본 예에서, 삭제의 시작시 오프셋은 -56이다. 56 샘플들 이후에, 이는 다시 -56이 된다. 추가 80 - 56 = 24 샘플들 이후에, 오프셋은 -56 + 24 = -32이므로, O는 제 1 프레임의 종료시 -32이다.

단계 745에서, 합성 버퍼가 S[0] 내지 S[79]로 채워진 이후에, S는 히스토리 버퍼(240)를 업데이트하는데 사용된다. 단계 750에서, 히스토리 버퍼(240)는 또한 3.75 msec 지연을 부가한다. 히스토리 버퍼(240)의 조작은 삭제된 프레임 및 비-삭제된 프레임 동안 동일하다. 이 점에서, 도 5의 단계 535에 대한 제 1 프레임 은폐 동작은 종료되고, 처리는 도 5에서 단계 540으로 진행된다.

도 5의 단계 550에 도시된 바와 같이, FEC 모듈(230)이 10 msec를 넘어서 나중 프레임을 은폐하도록 동작하는 방법의 상세한 내용은 도 12에 상세히 도시된다. 제 2 및 나중 삭제된 프레임들 동안 합성 신호를 발생하는데 사용되는 기술은 비록 신호에 일부 변화를 부가하는데 일부 추가 작업이 행해질 필요가 있더라도 제 1 삭제된 프레임과 아주 유사하다.

단계 1205에서, 삭제 코드는 제 2 또는 제 3 프레임이 삭제되고 있는지의 여부를 결정한다. 제 2 및 제 3 삭제된 프레임들 동안, 피치 버퍼로부터 사용되는 피치 주기들의 수는 증가된다. 이는 신호에서 더 큰 변화를 도입하고, 합성 출력이 너무 조화적으로(harmonic) 발성되는 것을 방지한다. 다른 모든 전이들과 같이, OLA는 피치 주기들의 수가 증가될 때 경계를 평활화할 필요가 있다. 제 3 프레임(30 msec의 삭제)을 넘으면, 피치 버퍼는 3 파장들의 길이로 일정하게 유지된다. 이들 3 파장들은 삭제의 기간 동안 모든 합성 음성을 발생시킨다. 그래서, 도 12에서 좌측에 있는 브랜치는 단지 제 2 및 제 3 삭제된 프레임들에만 취해진다.

다음, 단계 1210에서는, 피치 버퍼에서 사용되는 파장의 수가 증가시켰다. 즉, U = U+ 1로 설정한다.

제 2 또는 제 3 삭제된 프레임이 시작시, 단계 1215에서 이전 프레임으로부터의 합성 신호가 추가 1/4 파장 동안 현재 프레임의 시작까지 계속된다. 예를 들면, 제 2 프레임의 시작시, 본 실시예의 합성 신호는 도 13에 도시된 바와 같이 나타난다. 이 1/4 파장은 피치 버퍼로부터 더 오래된 파장들을 사용하는 새로운 합성 신호와 오버랩 가산된다.

제 2 삭제된 프레임이 시작시, 파장의 수는 2로 증가되어, U = 2가 된다. 1 파장 피치 버퍼와 같이, OLA는 2-파장 피치 버퍼가 그 자체를 반복하는 경계에서 실행되어야 한다. 피치 버퍼 B의 끝에서부터 돌아와 U 파장에서 끝나는 1/4 파장인 이 시간은 단계 1220에서 최종 1/4 버퍼 L의 내용과 오버랩 부가된다. 이 OLA 연산자는 범위 1 ≤ i ≤ P4에서 다음과 같이 표현될 수 있다:

$$B[-i] = \frac{i}{P4} L[-i] + \left(\frac{P4 - i}{P4}\right) B[-i - PU]$$

이 식의 이전 버전으로부터 다른 점은 단지 우변에서 B를 인덱스 처리하는데 사용되는 상수 P가 PU로 변환된 것이다. 2-파장 피치 버퍼의 생성은 도 14에 그래프로 도시된다.

도 11에서와 같이, 삭제 시작의 좌측에 있는 "OLA들과 조합된" 파형의 영역은 2-주기 피치 버퍼의 업데이트된 내용들이다. 짧은 수직선들은 피치 주기를 나타낸다. "OLA들과 조합된" 파형에서 연속적인 피크들의 면밀한 조사는 피크가 삭제 시작 전에 되돌아와 하나 및 두 개의 파장들을 피크들로부터 교체하는 것을 나타낸다.

제 2 프레임에서 합성 출력이 시작될 때, 새로운 피치 버퍼로부터의 신호를 도 13에서 발생된 1/4 파장과 병합시켜야 한다. 새로운 피치 버퍼로부터의 합성 신호는 사용되는 버퍼 중 가장 오래된 부분으로부터 전해지는 것이 바람직하다. 그러나, 새로운 부분은 파형의 유사한 부분으로부터 주어지거나, 또는 그들을 혼합할 때, 가청의 인위 구조들이 생성될 것이라는 것을 주위하여야 한다. 다른 말로 하면, 이들을 혼합할 때, 정확한 위상을 유지하기를 원하거나 또는 파형들이 파괴적으로 간섭될 수 있다.

이는 단계 1225(도 12)에서 피치 버퍼의 사용 부분에서, 가장 오래된 파장을 나타낼 때까지, 이전 프레임, O의 종료시 저장된 오프셋에서 주기들 P를 감산함으로서 달성된다.

예를 들면, 제 1 삭제된 프레임에서, 피치 버퍼 B에 대해 유효한 인덱스는 -1 내지 -P이었다. 그래서, 제 1 삭제된 프레임으로부터 저장된 O는 이 범위내에 있어야 한다. 제 2 삭제된 프레임에서, 유효 범위는 -1 내지 -2P이다. 그래서, O가 범위 -2P ≤O < -P에 있을 때까지 O로부터 P를 감산한다. 또는, 보다 일반적으로, 범위 -UP ≤O < -(U-1)P에 있을 때까지 O로부터 P를 감산한다. 본 실시예에서는 제 1 삭제된 프레임의 종료시 P = 56이고 O = -32이다. -32에서 56을 감산하면, -88이 산출된다. 그래서, 제 2 프레임에서 제 1 합성 샘플은 B[-88]로부터 주어지고, 다음은 B[-87]로부터 주어진다.

제 2 삭제된 프레임의 시작시 하나 및 두 개의 주기 피치 버퍼들로부터 합성 신호들의 OLA 혼합하는 것이 도 15에 도시된다.

O로부터 P를 감산함으로서, 적절한 파형 위상이 유지되고, "1P 피치 버퍼" 및 "2P 피치 버퍼" 파형들에서 신호의 피크들이 정렬됨을 주목하여야 한다. "OLA 조합된" 파형은 또한 제 2 삭제된 프레임의 시작시 다른 피치 버퍼들 사이의 매끄러운 전이를 도시한다. 도 15의 "OLA 조합된" 파형에서 제 2 프레임이 출력될 수 있기 전에 동작이 1 회 더 요구된다.

단계 1230(도 12)에서, 새로운 오프셋은 피치 버퍼로부터의 1/4 파장을 임시 버퍼에 복사하는데 사용된다. 단계 1235에서, 1/4 파장은 오프셋에 부가된다. 그 후, 단계 1240에서, 임시 버퍼는 출력 버퍼의 시작과 OLA 동작되고, 그 결과는 출력 버퍼의 처음 1/4 파장에 배치된다.

단계 1245에서, 오프셋은 그 때 출력 버퍼에서 신호의 나머지를 발생시키는데 사용된다. 피치 버퍼는 10 msec 프레임의 기간 동안 출력 버퍼로 복사된다. 단계 1250에서, 현재 오프셋이 변수 O로 피치 버퍼로 저장된다.

제 2 및 나중의 삭제된 프레임 동안, 단계 1255에서는 합성 신호가 선형 기울기로 감쇄된다. 합성 신호는 60 msec를 넘어서 0으로 설정되거나 묵음이 될 때까지 점차적으로 페이드 아웃(fade out)된다. 삭제가 길어질수록, 은폐 음성은 진짜 신호로부터 발산될 가능성이 더 높다. 사운드가 짧은 시간 주기 동안 고립되어 자연스럽게 들리더라도, 특정한 종류들의 사운드들을 너무 오래 간직하면, 은폐 처리의 출력에서 부자연스러운 가청 인위 구조들이 만들어질 수 있다. 합성 신호에서 이러한 인위 구조들을 방지하기 위해서, 저속 페이드 아웃(fade out)이 사용된다. G.723.1, G.728, 및 G,729와 같이, 모든 표준 음성 코더들에서 발견되는 은폐 처리들에서 유사한 동작이 실행된다.

FEC 처리는 제 2 프레임에서 시작되어, 10 msec 프레임 당 20%로 신호를 감쇄한다. S, 즉 합성 버퍼가 감쇄 이전의 합성 신호를 포함하고, F가 연속적인 삭제된 프레임의 수이면(제 1 삭제된 프레임에 대해 F = 1, 제 2 삭제된 프레임에 대해 F = 2), 그 때 감쇄는 다음과 같이 표시될 수 있다:

$$S'[i] = [1 - .2(F-2) - \frac{.2i}{80}]S[i]$$

범위 $0 \le i \le 79$에서 $2 \le F \le 6$이다. 예를 들어, 제 2 삭제된 프레임의 시작시 샘플에서 F = 2이므로, F-2 = 0이고, .2/80 = .0025이므로, S'[0] = 1.S[0], S'[1] = 0.9975S[1], S'[2] = 0.995S[2], 또한 S'[79] = 0.8025S[79]이다. 제 6 삭제된 프레임을 넘으면, 출력은 간단히 0으로 설정된다.

단계 1255에서 합성 신호가 감쇄된 이후에, 이는 단계 1260에서 히스토리 버퍼(240)로 주어지고, 단계 1265에서는 출력이 3.75 msec 만큼 지연된다. 오프셋 포인터 O는 또한 제 2 프레임의 종료시 피치 버퍼에서 그 위치로 업데이트되므로, 합성 신호는 다음 프레임에서 계속될 수 있다. 그 후, 처리는 다음 프레임을 구하도록 단계 540로 되돌아간다.

삭제가 두 프레임들을 넘어 계속되면, 제 3 프레임상의 처리는 피치 버퍼에서 주기들의 수가 1부터 2 대신에 2부터 3으로 증가되는 것을 제외하고 제 2 프레임에서와 완전히 똑같다. 본 예의 삭제는 두 개의 프레임에서 종료되지만, 제 3 프레임 이상에서 사용되는 3-주기 피치 버퍼가 도 17에 도시된다. 제 3 프레임을 넘으면, 피치 버퍼에서 주기들의 수는 3으로 고정되어 유지되므로, 도 12의 우측에 있는 경로만이 취해진다. 이 경우, 오프셋 포인터 O는 간단히 피치 버퍼를 합성 출력에 복사하는데 사용되고, 오버랩 가산 동작들은 필요하지 않다.

삭제 이후 양호한 제 1 프레임에서 FEC 모듈(230)의 동작은 도 16에서 상술된다. 삭제의 종료시, 삭제시 발생된 합성 음성과 실제 음성 사이에는 매끄러운 전이가 필요하다. 삭제가 단지 1 프레임 길이이면, 단계 1610에서, 1/4 파장의 합성 음성이 계속되고 실제 음성과의 오버랩 가산이 수행된다.

FEC 모듈(230)이 단계 1620에서 삭제가 10 msec보다 긴 것을 결정하면, 합성 신호와 실제 신호 사이의 부정합 가능성이 더 커지므로, 단계 1630에서는 합성 음성 발생이 계속되고, OLA 윈도우는 최대 10 msec까지 삭제된 프레임 당 추가 4 msec 만큼 증가된다. 피치의 추정이 약간 벗어나거나, 실제 음성의 피치가 삭제 동안 변화되었으면, 합성 신호와 실제 신호 사이의 위상 부정합 가능성은 삭제의 길이에 따라 증가된다. 더 긴 OLA 윈도우들은 더 천천히 합성 신호가 페이드 아웃(fade out)되게 하고, 실제 음성 신호가 페이드 인(fade in)되게 한다. 삭제가 10 msec보다 더 길면, OLA가 실행되기 이전에 단계 1640에서 이전 프레임에서 신호의 레벨과 정합하여, 합성 음성을 감쇄시키는 것이 필요하다.

단계 1650에서, OLA는 새로운 입력 프레임의 시작시 출력 버퍼(합성 음성)의 내용들상에 수행된다. 입력 버퍼의 시작은 OLA의 결과로 대체된다. 상기의 예에서 삭제의 종료시 OLA는 도 4에서 관찰될 수 있다. 상기 예에 대한 은폐 처리의 완전한 출력은 도 3의 "은폐된" 파형에서 관찰될 수 있다.

단계 1660에서, 히스토리 버퍼는 입력 버퍼의 내용들로 업데이트된다. 단계 1670에서, 음성의 출력은 3.75 msec 만큼 지연되고, 처리는 다음 프레임을 구하도록 도 5의 단계 530로 복귀한다.

작은 조정으로, FEC 처리는 G.726과 같이, 샘플들 또는 프레임들 사이의 상태 정보를 유지하고 은폐를 제공하지 않는 다른 음성 코더들에 적용될 수 있다. FEC 처리는 삭제 동안 합성 파형을 발생하는데 이전 섹션에서 설명된 것과 똑같이 처리된다. 그러나, 코더의 내부 상태 변수들이 FEC 처리에 의해 발생시킨 합성 음성의 추적을 보장하도록 주위되어야 한다. 그렇지 않으면, 삭제가 끝난 이후, 복호기가 에러있는 상태를 사용하여 재시작될 때 출력에 인위 구조들 및 불연속성들이 나타나게 된다. 삭제의 종료시 OLA 윈도우가 돕지만, 더 실행되어야 한다.

더 나은 결과는 부호기(1860)의 입력으로 FEC 모듈(1830)의 합성된 출력을 사용하여 삭제의 기간 동안 복호기(1820)를 부호기(1860)로 변환함에 의해, 도 18에 도시된 바와 같이 구해질 수 있다.

이 방법으로, 복호기(1820)의 변수들 상태는 은폐된 음성을 추적할 것이다. 전형적인 부호기와 다르게, 부호기(1860)는 단지 상태 정보만을 유지하고 그 출력은 사용되지 않음을 주목하여야 한다. 그러므로, 손쉬운 방법들은 운행 시간 복잡성을 현저히 낮추도록 취해질 수 있다.

상술된 바와 같이, 본 발명에 의해 제공되는 많은 이점들 및 양태들이 있다. 특별히, 프레임 삭제가 진행됨에 따라, 합성 신호를 발생시키도록 신호 히스토리로부터 사용되는 피치 주기들의 수는 시간의 함수로 증가된다. 이는 오랜 삭제들에서 조화적인 인위 구조들을 현저히 감소시킨다. 피치 주기들이 원래 순서로 재생되지 않더라도, 출력은 여전히 자연스럽게 들린다.

G.726 및 샘플들 또는 프레임들 사이에 상태 정보를 유지하는 다른 코더들에서, 복호기는 은폐 처리의 합성된 출력의 출력에 부호기로서 작용될 수 있다. 이 방법으로, 복호기의 내부 상태 변수들이 출력을 추적할 것이며, 삭제가 끝난 이후에 복호기에서 에러가 있는 상태의 정보에 의해 발생되는 불연속성들을 방지 -적어도 감소- 시키게 된다. 부호기로부터의 출력은 결코 사용되지 않으므로(단지 상태 정보를 유지할 목적이므로), 불필요한 장비를 모두 제거한(stripped-down) 저복잡성 버전의 부호기가 사용될 수 있다.

본 실시예에서 허용되는 최소 피치 주기(40 샘플들, 또는 200 Hz)는 일부 여성 및 어린이 발성자들에 대해 기대하는 기본 주파수보다 더 크다. 그래서, 고주파수 음성자들의 경우에는 하나 이상의 피치 주기가 삭제의 시작시에도 합성 신호를 발생하는데 사용된다. 높은 기본 주파수의 발성자들인 경우, 파형들은 더 자주 반복된다. 합성 신호에서 다중 피치 주기들은 조화적인 인위 구조들을 덜 가능하게 한다. 이 기술은 또한 중단과 같은 신속한 전이의 영역들 뿐만 아니라, 음성의 목소리가 아닌 세그먼트들 동안 신호가 자연스럽게 들리는데 도움을 준다.

삭제 이후 양호한 제 1 프레임의 종료시 OLA 윈도우는 삭제의 길이와 함께 커진다. 삭제들이 더 길어지면, 다음 양호한 프레임이 도착할 때, 위상 정합이 일어날 더 많은 가능성이 있다. OLA 윈도우를 삭제 길이의 함수로 연장시키는 것은 긴 삭제에서 위상 부정합에 의해 발생되는 글리치들(glitches)을 감소시키지만, 삭제가 짧은 경우 신호가 신속하게 회복되도록 허용한다.

본 발명의 FEC 처리는 또한 1/4 파장인 추정 피치의 작은 부분이고 피치 피크들로 정렬되지 않는 가변 길이의 OLA 윈도우들을 사용한다.

본 발명의 FEC 처리는 목소리와 목소리가 아닌 음성 사이를 구별하지 않는다. 그 대신에, 다음의 두 가지 처리의 특성들 때문에 목소리가 아닌 음성을 재생을 잘 수행한다:(A) 최소 윈도우 크기는 음성 중 목소리가 아닌 영역들이 적당한 변화를 갖더라도 적절하게 크고,(B) 처리가 진행됨에 따라, 피치 버퍼의 길이가 증가되어, 또 조화적인 인위 구조들이 생기지 않도록 보장한다. 목소리 음성 및 목소리가 아닌 음성을 다르게 나타내는 조작을 방지하는데 큰 윈도우들을 사용은 또한 이미 공지된 시간-스케일링 기술 WSOLA에서 개시되었음을 주목하여야 한다.

삭제의 시작시 OLA를 허용하는 지연을 부가하는 것이 본 발명의 처리 중 바람직하지 않은 양태로서 생각될 수 있지만, 삭제의 시작시 실제 신호와 합성 신호 사이에 매끄러운 전이를 보장하는 것이 필수적이다.

본 발명은 상술된 특정 실시예들과 관련되어 설명되었지만, 많은 대안들, 변경들, 및 변화들이 종래 기술에 숙련된 자에게는 명백할 것임이 분명하다. 따라서, 상술된 본 발명의 바람직한 실시예들은 제한되는 것이 아니라, 설명하는 것으로 의도된다. 다음 청구항들에서 정의되는 바와 같이, 본 발명의 의도 및 범위에서 벗어나지 않고 다양한 변화들이 만들어질 수 있다.

(57) 청구의 범위

청구항 1.

음성 부호기로부터 코드화된 음성 정보를 포함하는 패킷들의 비유용성을 경험하는 음성 복호기에서 상태 정보를 유지하는 방법으로서, 상기 복호기의 상기 상태 정보는 수신된 부호화된 음성 정보를 반영하며, 상기 방법은:

음성 신호를 형성하기 위해 수신된 패킷들을 복호화하는 단계;

하나 이상의 패킷들이 상기 복호기에서 이용가능하지 않음을 결정하는 단계;

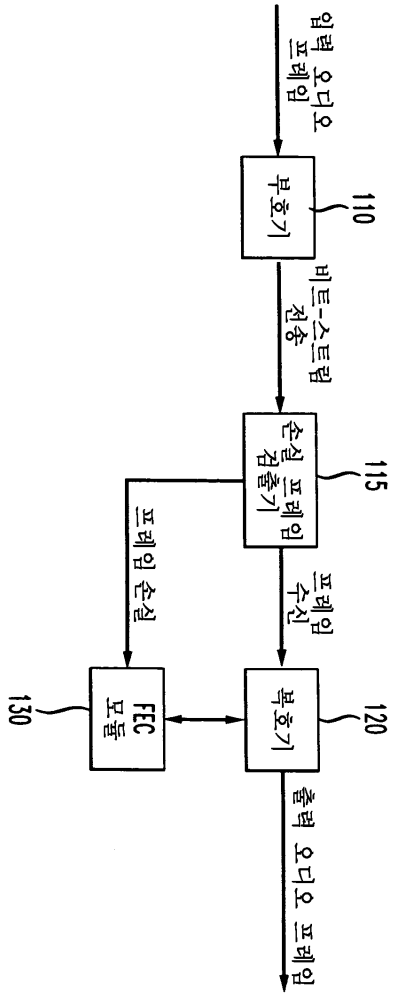
상기 비유용성 패킷들에 대응하는 음성 신호를 합성하는 단계;

상기 합성된 음성 신호를 부호화하는 단계; 및

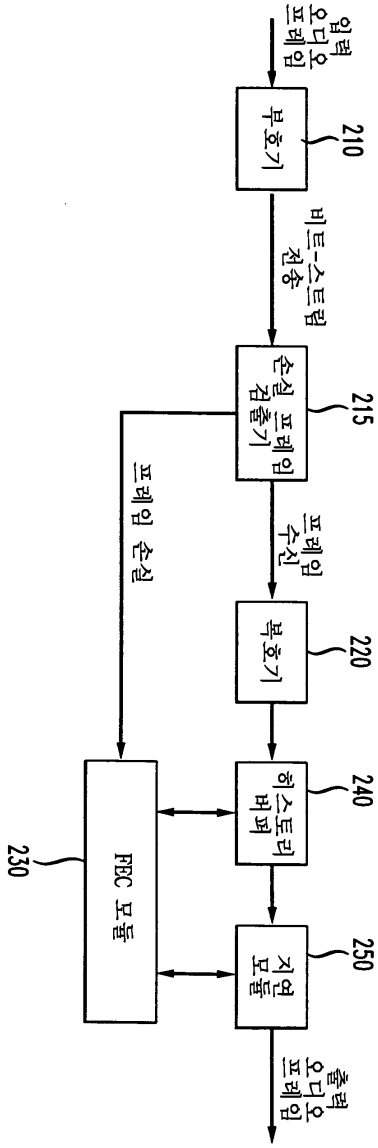
복호기 상태 정보를 유지하는데 사용하기 위해 상기 복호기에 상기 부호화된 합성된 음성을 반영하는 신호들을 제공하는 단계를 구비하는, 상태 정보 유지 방법.

도면

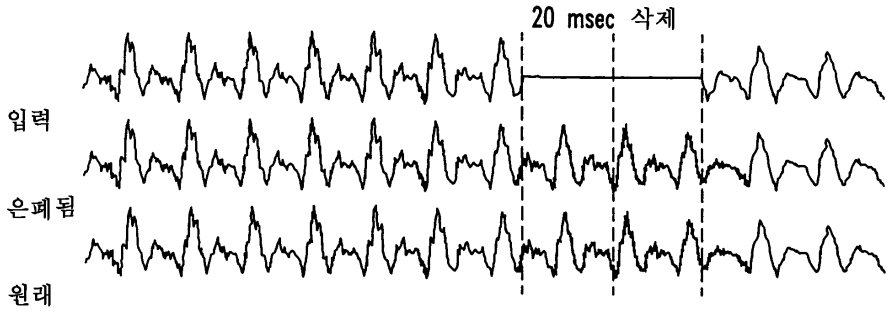
도면1


입력 오디오 프레임
부호기
110
비트-스트림 전송
손실 프레임 검출기
115
프레임 손실
프레임 수신
FEC 모듈
130
복호기
120
출력 오디오 프레임

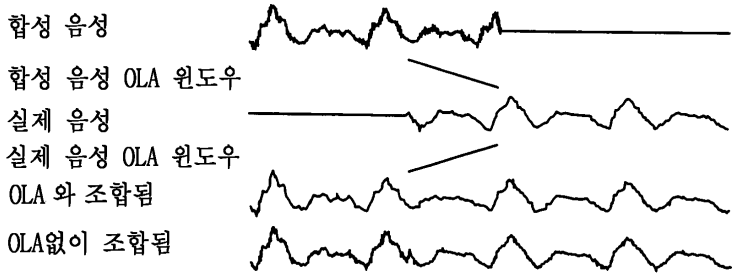
도면2


입력 오디오 프레임
부호기
210
비트-스트림 전송
손실 프레임 검출기
215
프레임 손실
프레임 수신
복호기
220
히스토리 버퍼
240
FEC 모듈
230
지연 모듈
250
출력 오디오 프레임

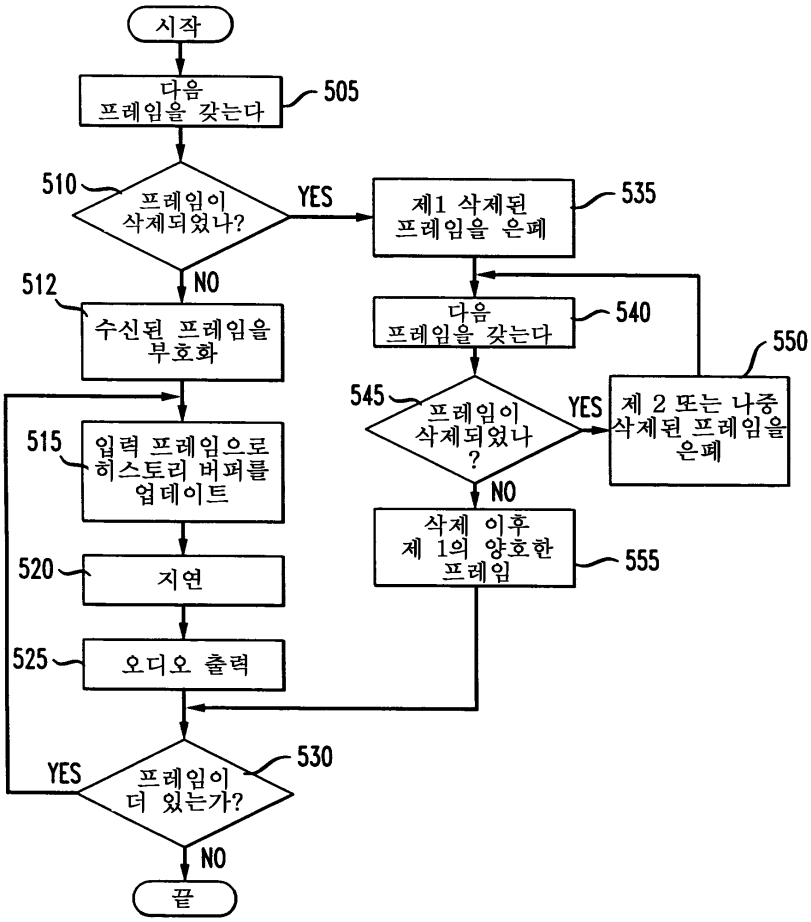
도면3


20 msec 삭제
입력
은폐됨
원래

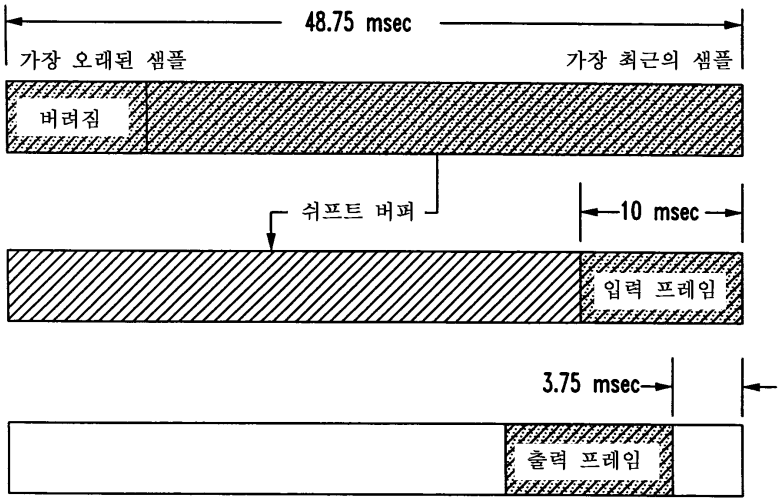
도면4


합성 음성
합성 음성 OLA 윈도우
실제 음성
실제 음성 OLA 윈도우
OLA 와 조합됨
OLA없이 조합됨

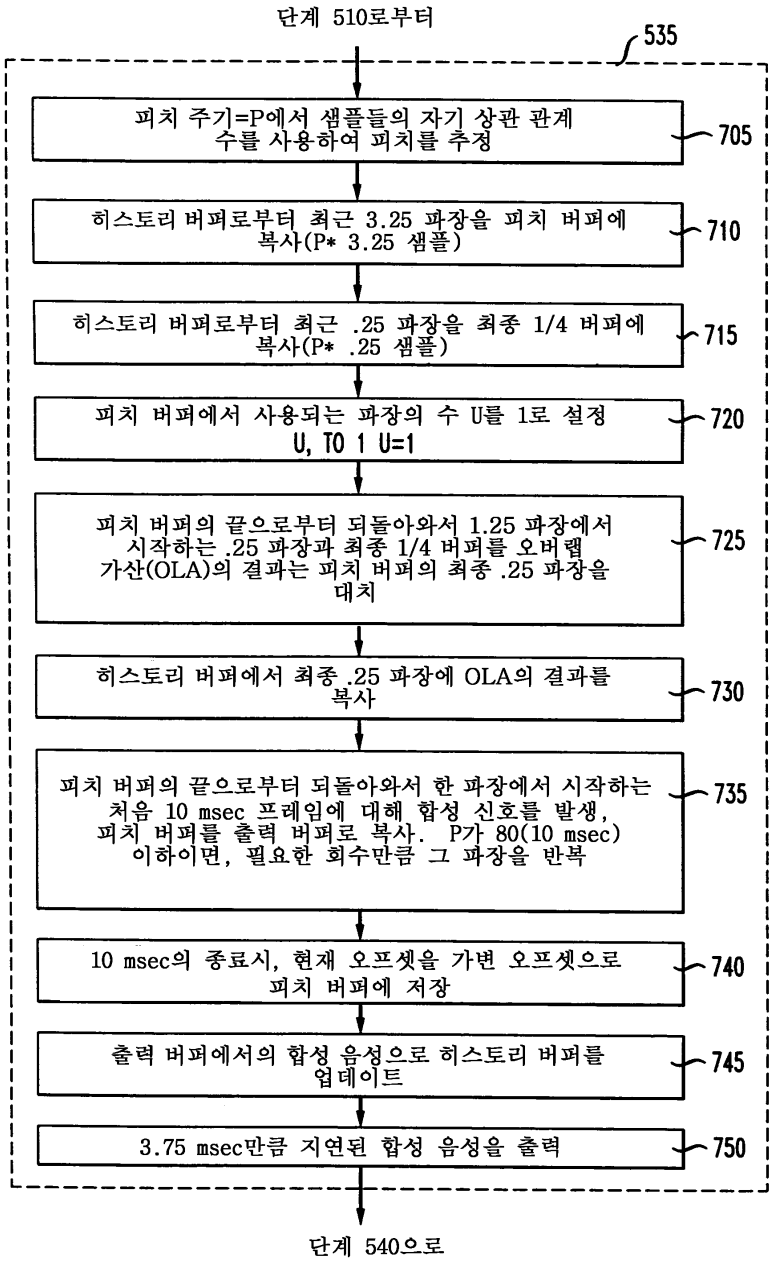
도면5


시작
다음 프레임을 갖는다
505
510
프레임이 삭제되었나?
YES
NO
제1 삭제된 프레임을 은폐
535
512
수신된 프레임을 부호화
다음 프레임을 갖는다
540
550
545
프레임이 삭제되었나 ?
YES
제 2 또는 나중 삭제된 프레임을 은폐
515
입력 프레임으로 히스토리 버퍼를 업데이트
NO
삭제 이후 제 1의 양호한 프레임
555
520
지연
525
오디오 출력
530
YES
프레임이 더 있는가?
NO
끝

도면6


48.75 msec
가장 오래된 샘플
가장 최근의 샘플
버려짐
쉬프트 버퍼
10 msec
입력 프레임
3.75 msec
출력 프레임

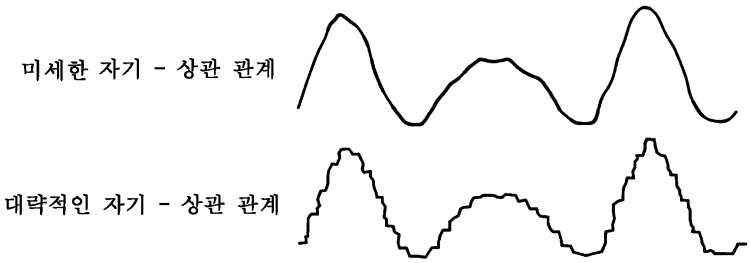
도면7


단계 510로부터
535
피치 주기=P에서 샘플들의 자기 상관 관계 수를 사용하여 피치를 추정
705
히스토리 버퍼로부터 최근 3.25 파장을 피치 버퍼에 복사(P* 3.25 샘플)
710
히스토리 버퍼로부터 최근 .25 파장을 최종 1/4 버퍼에 복사(P* .25 샘플)
715
피치 버퍼에서 사용되는 파장의 수 U를 1로 설정
U, TO 1 U=1
720
피치 버퍼의 끝으로부터 되돌아와서 1.25 파장에서 시작하는 .25 파장과 최종 1/4 버퍼를 오버랩 가산(OLA)의 결과는 피치 버퍼의 최종 .25 파장을 대치
725
히스토리 버퍼에서 최종 .25 파장에 OLA의 결과를 복사
730
피치 버퍼의 끝으로부터 되돌아와서 한 파장에서 시작하는 처음 10 msec 프레임에 대해 합성 신호를 발생, 피치 버퍼를 출력 버퍼로 복사. P가 80(10 msec) 이하이면, 필요한 회수만큼 그 파장을 반복
735
10 msec의 종료시, 현재 오프셋을 가변 오프셋으로 피치 버퍼에 저장
740
출력 버퍼에서의 합성 음성으로 히스토리 버퍼를 업데이트
745
3.75 msec만큼 지연된 합성 음성을 출력
750
단계 540으로

도면8

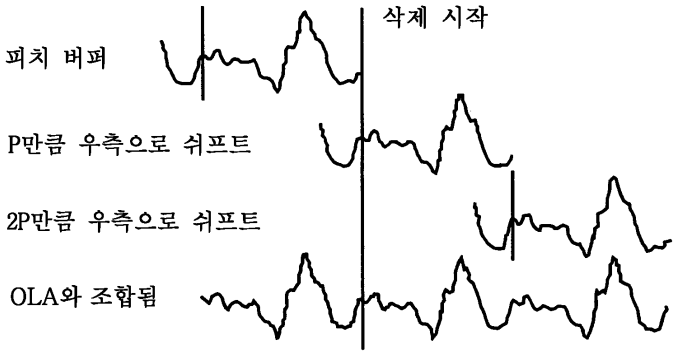
히스토리
상관 관계 윈도우들
상관 관계

도면9

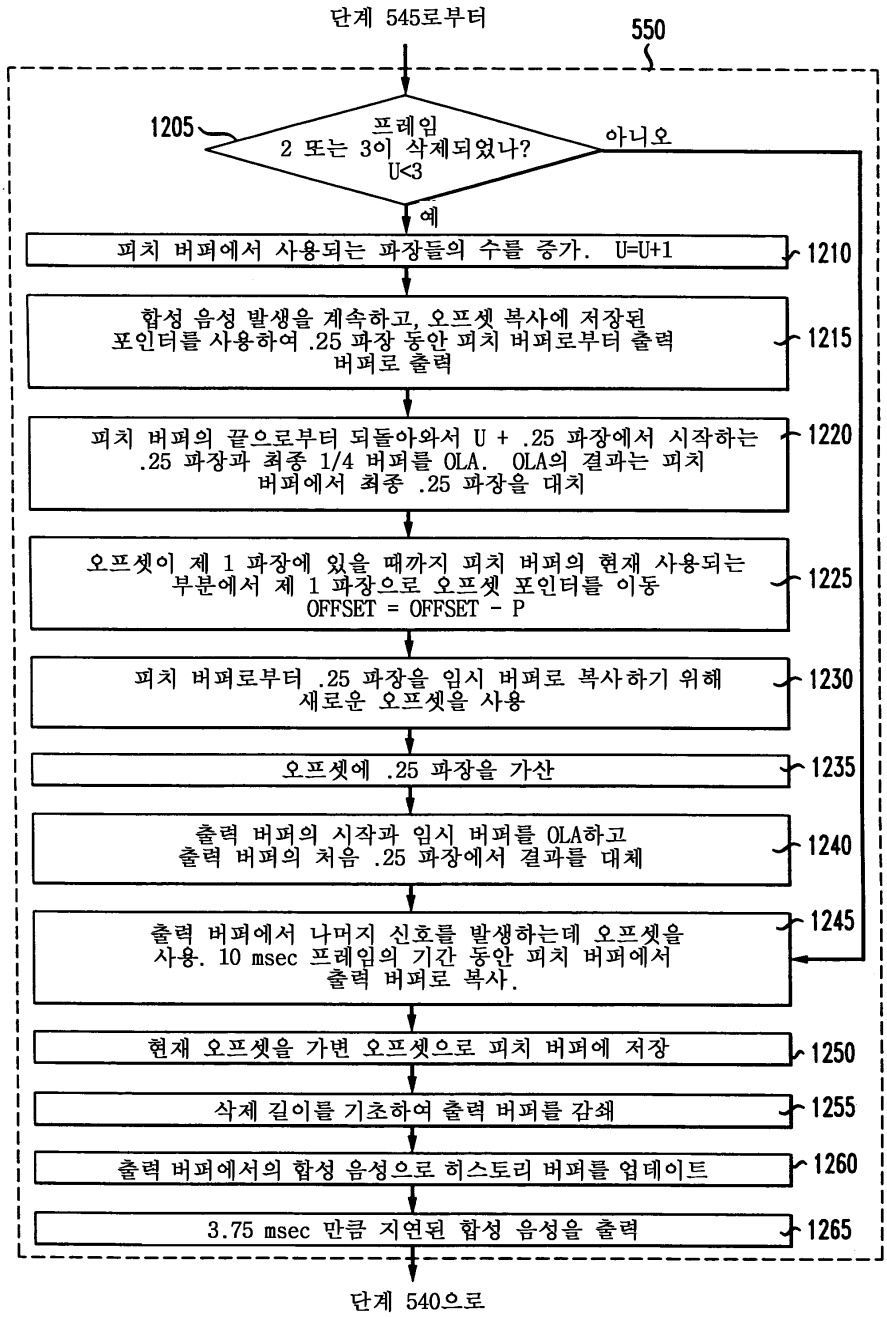


도면10

도면11



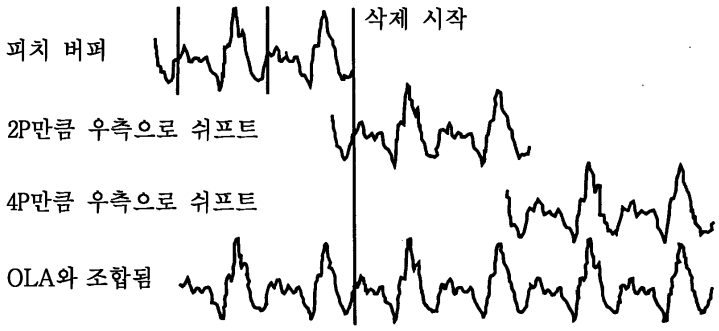
도면12


단계 545로부터
550
1205
프레임
2 또는 3이 삭제되었나?
U<3
아니오
예
피치 버퍼에서 사용되는 파장들의 수를 증가. U=U+1
1210
합성 음성 발생을 계속하고, 오프셋 복사에 저장된
포인터를 사용하여 .25 파장 동안 피치 버퍼로부터 출력
버퍼로 출력
1215
피치 버퍼의 끝으로부터 되돌아와서 U + .25 파장에서 시작하는
.25 파장과 최종 1/4 버퍼를 OLA. OLA의 결과는 피치
버퍼에서 최종 .25 파장을 대치
1220
오프셋이 제 1 파장에 있을 때까지 피치 버퍼의 현재 사용되는
부분에서 제 1 파장으로 오프셋 포인터를 이동
OFFSET = OFFSET - P
1225
피치 버퍼로부터 .25 파장을 임시 버퍼로 복사하기 위해
새로운 오프셋을 사용
1230
오프셋에 .25 파장을 가산
1235
출력 버퍼의 시작과 임시 버퍼를 OLA하고
출력 버퍼의 처음 .25 파장에서 결과를 대체
1240
출력 버퍼에서 나머지 신호를 발생하는데 오프셋을
사용. 10 msec 프레임의 기간 동안 피치 버퍼에서
출력 버퍼로 복사.
1245
현재 오프셋을 가변 오프셋으로 피치 버퍼에 저장
1250
삭제 길이를 기초하여 출력 버퍼를 감쇄
1255
출력 버퍼에서의 합성 음성으로 히스토리 버퍼를 업데이트
1260
3.75 msec 만큼 지연된 합성 음성을 출력
1265
단계 540으로

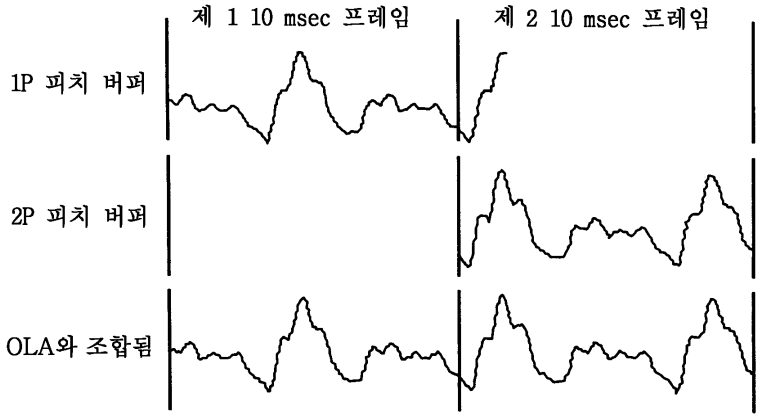
도면13

제 1 10 msec 프레임
제 2 10 msec 프레임

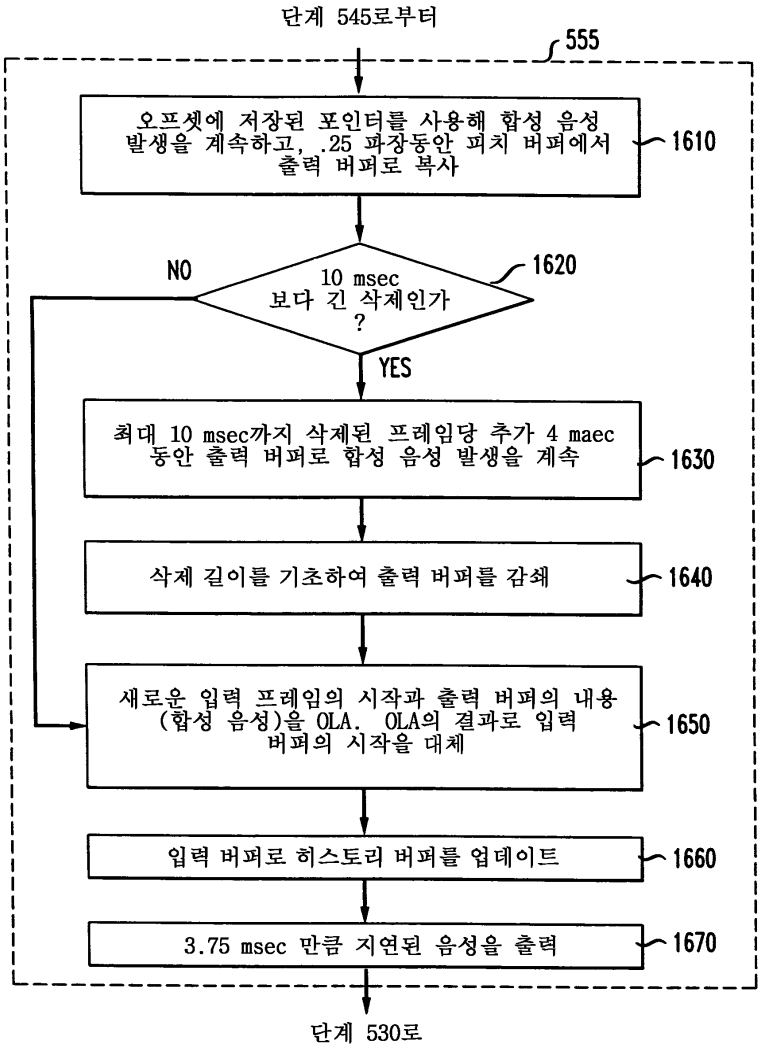
도면14


삭제 시작
피치 버퍼
2P만큼 우측으로 쉬프트
4P만큼 우측으로 쉬프트
OLA와 조합됨

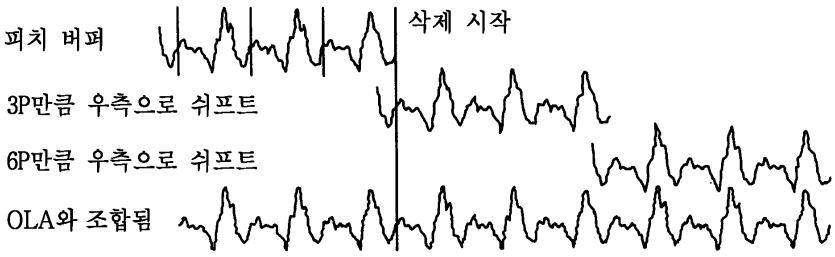
도면15


제 1 10 msec 프레임
제 2 10 msec 프레임
1P 피치 버퍼
2P 피치 버퍼
OLA와 조합됨

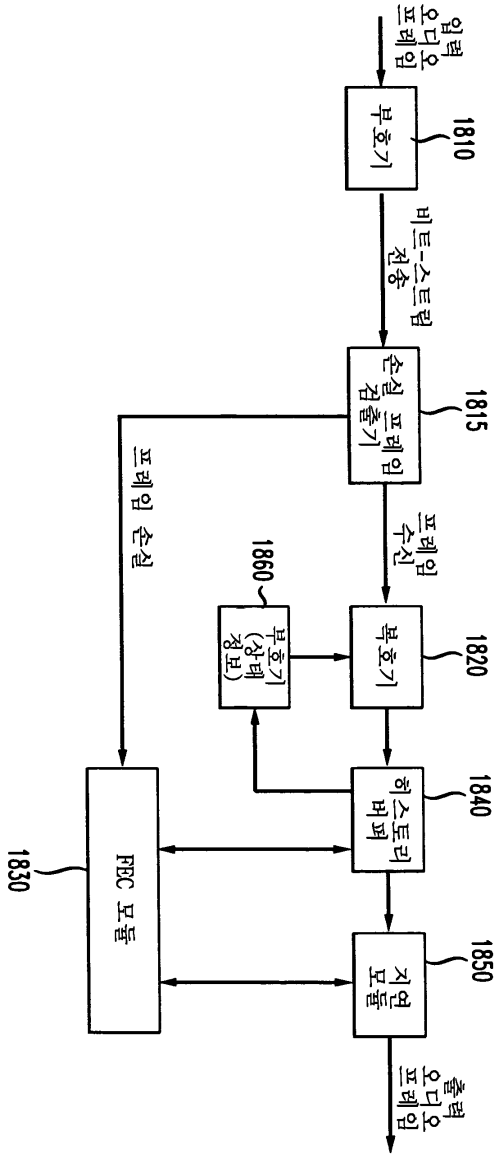
도면16


단계 545로부터
555
오프셋에 저장된 포인터를 사용해 합성 음성 발생을 계속하고, .25 파장동안 피치 버퍼에서 출력 버퍼로 복사
1610
10 msec 보다 긴 삭제인가 ?
1620
NO
YES
최대 10 msec까지 삭제된 프레임당 추가 4 maec 동안 출력 버퍼로 합성 음성 발생을 계속
1630
삭제 길이를 기초하여 출력 버퍼를 감쇄
1640
새로운 입력 프레임의 시작과 출력 버퍼의 내용 (합성 음성)을 OLA. OLA의 결과로 입력 버퍼의 시작을 대체
1650
입력 버퍼로 히스토리 버퍼를 업데이트
1660
3.75 msec 만큼 지연된 음성을 출력
1670
단계 530로

도면17


피치 버퍼
삭제 시작
3P만큼 우측으로 쉬프트
6P만큼 우측으로 쉬프트
OLA와 조합됨

도면18


입력 오디오 프레임
부호기
1810
비트-스트림 전송
손실 프레임 검출기
1815
프레임 수신
프레임 손실
복호기
1820
부호기 (상태 정보)
1860
히스토리 버퍼
1840
FEC 모듈
1830
지연 모듈
1850
출력 오디오 프레임