



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2010-0009936
(43) 공개일자 2010년01월29일

(51) Int. Cl.

G10L 11/00 (2010.01) G10L 21/02 (2010.01)

(21) 출원번호 10-2008-0070775

(22) 출원일자 2008년07월21일

심사청구일자 없음

(71) 출원인

삼성전자주식회사

경기도 수원시 영통구 매탄동 416

고려대학교 산학협력단

서울 성북구 안암동5가1 고려대학교 내

(72) 발명자

김현수

경기도 용인시 수지구 상현동 만현마을 두산위브
아파트 803동 501호

고한석

서울특별시 용산구 서빙고동 신동아아파트 16-703

(뒷면에 계속)

(74) 대리인

박상수

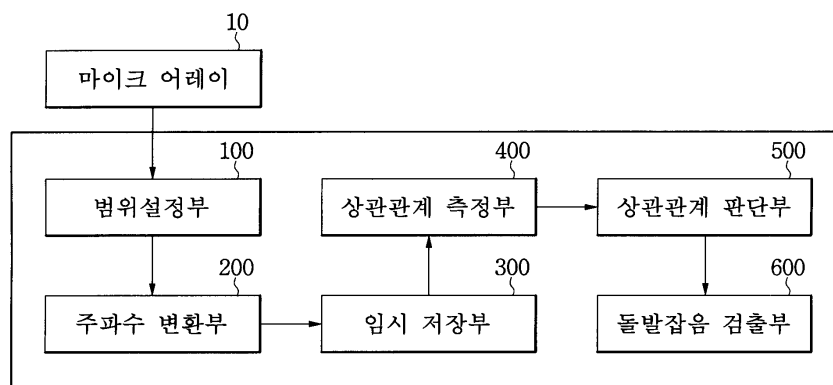
전체 청구항 수 : 총 16 항

(54) 음원 검출 시스템에서 돌발잡음 추정/제거 장치 및 방법

(57) 요약

본 발명에 따른 음원 검출 시스템에서 돌발잡음 추정/제거 장치는 잡음 섞인 음성 신호는 마이크 어레이를 통해 입력되며 hanning 윈도우를 씌워 프레임으로 구분하는 범위설정부; 범위설정부를 통해 프레임별로 구분된 신호를 주파수축으로 변환하는 주파수 변환부; 상기 주파수 변환부를 통해 변환된 프레임별 FFT 값을 임시 저장하는 임시 저장부; 현재 입력시점의 프레임과 기 설정된 시간만큼 떨어진 이후의 프레임간의 상관관계 정도를 측정하는 상관관계 측정부; 상기 상관관계 측정부를 통해 측정된 상관관계 값이 기 설정된 임계값을 초과하는지의 여부를 판단하는 상관관계 판단부; 및 상기 상관관계 판단부를 통해 상관관계 값이 기 설정된 임계값을 초과하면 돌발잡음을 검출하는 돌발잡음 검출부를 포함하는 것으로, 돌발잡음에 보다 효과적인 잡음추정을 할 수 있는 이점이 있다.

대표도 - 도3



(72) 발명자

배정훈

서울특별시 중구 신당6동 58-44

이택진

서울특별시 관악구 신림4동 472-33

특허청구의 범위

청구항 1

입력된 음성신호로부터 잡음을 제거하는 음원 검출 시스템에서의 돌발잡음 추정/제거 장치에 있어서,

음성신호를 수신하는 마이크 어레이부;

수신된 신호에 설정된 범위가 겹쳐지도록 윈도우를 씌우는 범위설정부;

상기 윈도우된 신호를 주파수축으로 변환하기 위한 푸리에 변환을 적용하는 주파수 변환부;

상기 주파수 변환부를 통해 변환된 푸리에 변환값과 이전 값의 상관값을 비교하고 상기 윈도우되어 푸리에 변환된 값의 전력(energy)를 계산하는 상관관계 측정부; 및

상관 비교값과 전력값을 이용해 잡음을 검출하는 돌발잡음 검출부를 포함하는 음원 검출 시스템에서 돌발잡음 추정/제거 장치.

청구항 2

제 1항에 있어서,

상기 마이크 어레이부를 통해 입력된 아날로그 음성신호를 디지털 음성 신호로 변환하는 A/D변환부를 더 포함하는 음원 검출 시스템에서 돌발잡음 추정/제거 장치.

청구항 3

제 2항에 있어서,

상기 주파수 변환부 뒷단에 주파수 변환부를 통해 변환된 푸리에 변환값을 저장하는 버퍼를 더 포함하는 음원 검출 시스템에서 돌발잡음 추정/제거 장치.

청구항 4

제 3항에 있어서,

상기 범위설정부는,

32ms 길이를 갖는 hanning 윈도우임을 특징으로 하는 음원 검출 시스템에서 돌발잡음 추정/제거 장치.

청구항 5

제 4항에 있어서,

상기 상관관계 측정부는,

상관 비교 값이 기 설정된 임계값을 초과하는지의 여부를 판단하는 상관관계 판단부와 상기 상관관계 판단부를 통해 상관관계 값이 기 설정된 임계값을 초과하면 잡음으로 인식하는 잡음 검출부를 더 포함하는 음원 검출 시스템에서 돌발잡음 추정/제거 장치.

청구항 6

제 5항에 있어서,

상기 상관관계 판단부는,

크로스 파워 스펙트럼을 이용하여 스펙트럼의 magnitude 값을 제공한 후 주파수 전 영역에서 합하여 해당 프레임의 에너지($\gamma(s)$)로 정의하고,

크로스 파워 스펙트럼을 통해 에너지를 검출한 프레임과, 임의의 주파수에서 로컬 에너지와 최소 통계치를 기반의 추정된 잡음과의 비($S_r(s,k)$)로 정의하며,

해당 프레임의 에너지($\gamma(s)$)와 추정된 잡음과의 비($S_r(s,k)$)에 임계값을 주어서 $\gamma(s) \geq \gamma_{th}$ 인 경우 해당 임계값보

다 작고 $S_{\mu}(s, k)$ 이 주어진 임계값보다 클 경우 돌발잡음이 존재한다고 판단하는 것을 특징으로 하는 음원 검출 시스템에서 돌발잡음 추정/제거 장치.

청구항 7

제 6항에 있어서,

상기 돌발잡음 검출부는,

기존의 MCRA 잡음 추정기법에 돌발잡음을 검출하는 파라메타를 적용시켜 이하 [수학식 8], [수학식 9], [수학식 10]을 통해 돌발잡음을 구하는 것을 특징으로 하는 음원 검출 시스템에서 돌발잡음 추정/제거 장치.

[수학식 8]

$$\lambda(k, \ell + 1) = \alpha(k, \ell) \hat{\lambda}(k, \ell + 1) + (1 - \alpha(k, \ell)) |Y(k, \ell)|^2$$

여기서, $\hat{\lambda}(k, \ell + 1)$ 는 추정된 잡음, K는 주파수 인덱스, L은 프레임 인덱스 이다.

[수학식 9]

$$\alpha(k, \ell) = \tilde{\alpha}(k, \ell) + (1 - \tilde{\alpha}(k, \ell)) p(k, \ell) (1 - I(k, \ell))$$

여기서, p(k, l)는 음성존재 확률, K는 주파수 인덱스, L은 프레임 인덱스 이다.

[수학식 10]

$$\tilde{\alpha}(k, \ell) = \alpha_{ds} + (\alpha_{dt} - \alpha_{ds}) I(k, \ell)$$

여기서, $\alpha_{ds}=0.95$, $\alpha_{dt}=0.05$ 로서 각각 stationary 잡음, 돌발 잡음 구간에서의 update 계수이다.

청구항 8

제 1항에 있어서,

상기 돌발잡음 검출부는,

돌발잡음이 검출되지 않으면 고정(stationary) 잡음으로 추정하는 것을 특징으로 하는 음원 검출 시스템에서 돌발잡음 추정/제거 장치.

청구항 9

입력된 음성신호로부터 잡음을 제거하는 음원 검출 시스템에서의 돌발잡음 추정/제거 방법에 있어서,

음성신호를 수신하는 단계;

수신된 신호에 설정된 범위가 겹쳐지도록 윈도우를 씌우는 단계;

윈도우된 신호를 주파수축으로 변환하기 위한 푸리에 변환을 적용하는 단계;

변환된 푸리에 변환값과 이전 값의 상관값을 비교하고 윈도우되어 푸리에 변환된 값의 전력(energy)를 계산하는 단계; 및

상관 비교값과 전력값을 이용해 잡음을 검출하는 단계를 포함하는 음원 검출 시스템에서 돌발잡음 추정/제거 방법.

청구항 10

제 9항에 있어서,

입력된 아날로그 음성신호를 디지털 음성 신호로 변환하는 단계를 더 포함하는 음원 검출 시스템에서 돌발잡음 추정/제거 방법.

청구항 11

제 9항에 있어서,

푸리에 변환을 적용하는 단계 이후 변환된 푸리에 변환값을 저장하는 단계를 더 포함하는 음원 검출 시스템에서 돌발잡음 추정/제거 방법.

청구항 12

제 9항에 있어서,

상기 윈도우를 씌우는 단계는,

32ms 길이를 갖는 hanning 윈도우임을 특징으로 하는 음원 검출 시스템에서 돌발잡음 추정/제거 방법.

청구항 13

제 9항에 있어서,

상기 윈도우되어 푸리에 변환된 값의 전력(energy)를 계산하는 단계는,

상관 비교 값이 기 설정된 임계값을 초과하는지의 여부를 판단하는 단계와 상기 판단 단계에서 상관관계 값이 기 설정된 임계값을 초과하면 잡음으로 인식하는 단계를 더 포함하는 음원 검출 시스템에서 돌발잡음 추정/제거 방법.

청구항 14

제 13항에 있어서,

상기 상관관계 판단 단계는,

크로스 파워 스펙트럼을 이용하여 스펙트럼의 magnitude 값을 제공한 후 주파수 전 영역에서 합하여 해당 프레임의 에너지($\gamma(s)$)로 정의하는 단계;

크로스 파워 스펙트럼을 통해 에너지를 검출한 프레임과, 임의의 주파수에서 로컬 에너지와 최소 통계치를 기반의 추정된 잡음과의 비($S_r(s,k)$)로 정의하는 단계;

이후, 해당 프레임의 에너지($\gamma(s)$)가 기 설정된 임계값 보다 큰지의 여부를 판단하는 단계; 및

상기 프레임의 에너지가 기 설정된 임계값 보다 큰지의 여부를 판단하는 단계에서 상기 프레임의 에너지($\gamma(s)$)가 기 설정된 임계값 보다 작으면, 추정된 잡음과의 비($S_r(s,k)$)가 기 설정된 임계값 보다 큰지의 여부를 판단하는 단계를 포함하는 음원 검출 시스템에서 돌발잡음 추정/제거 방법.

청구항 15

제 14항에 있어서,

상기 돌발잡음 검출 단계는,

기존의 MCRA 잡음 추정기법에 돌발잡음을 검출하는 파라메타를 적용시켜 이하 [수학식 8], [수학식 9], [수학식 10]을 통해 돌발잡음을 구하는 것을 특징으로 하는 음원 검출 시스템에서 돌발잡음 추정/제거 방법.

[수학식 8]

$$\hat{\lambda}(k, \ell + 1) = \alpha(k, \ell) \hat{\lambda}(k, \ell + 1) + (1 - \alpha(k, \ell)) |Y(k, \ell)|^2$$

여기서, $\hat{\lambda}(k, \ell + 1)$ 는 추정된 잡음, K는 주파수 인덱스, L은 프레임 인덱스 이다.

[수학식 9]

$$\alpha(k, \ell) = \tilde{\alpha}(k, \ell) + (1 - \tilde{\alpha}(k, \ell)) p(k, \ell) (1 - I_f(k, \ell))$$

여기서, p(k,l)는 음성존재 확률, K는 주파수 인덱스, L은 프레임 인덱스 이다.

[수학식 10]

$$\tilde{\alpha}(k, \ell) = \alpha_{ds} + (\alpha_{dt} - \alpha_{ds})I_f(k, \ell)$$

여기서, $\alpha_{ds}=0.95$, $\alpha_{dt}=0.05$ 로서 각각 stationary 잡음, 돌발 잡음 구간에서의 update 계수이다.

청구항 16

제 9항에 있어서,

상기 돌발잡음 검출 단계는,

돌발잡음이 검출되지 않으면 고정(stationary) 잡음으로 추정하는 것을 특징으로 하는 음원 검출 시스템에서 돌발잡음 추정/제거 방법.

명세서

발명의 상세한 설명

기술분야

[0001] 본 발명은 음성통신/인식에 필요한 음성 강화(speech enhancement)를 위한 어쿠스틱 잡음제거에 있어서 필수적인 잡음추정이다. 여기서 어쿠스틱 잡음이란 입력장치(마이크)에 음성과 같이 입력되는 주변 환경잡음을 뜻한다.

배경기술

[0002] 종래 마이크 어레이를 이용한 방향성 잡음 제거 시스템은 도 1에 도시된 바와 같이 적어도 하나 이상의 마이크 어레이(10), 마이크 어레이(10)에 각각 연결된 단구간 분석부(20), 에코 제거부(30), 방향성 잡음 제거와, 정면 소리 유/무를 기반으로 필터 weight update on/off하는 적응 빔포밍(adaptive Beamforming) 처리부(40), 마이크간의 신호의 상관관계를 이용하여 정면소리를 정면소리 검출부(50), 잔여잡음 제거와 정면소리 유/무를 기반으로 잔여 잡음을 제어하는 포스트 필터링(Post-filtering)부(60) 및 Overlap and add 처리부(70)를 포함한다.

[0003] 이러한 종래 마이크 어레이를 이용한 방향성 잡음 제거 시스템에서 잡음을 추정하는 방법은 잡음의 에너지가 음성의 에너지보다 작다는 가정을 기반으로 하고, 음성과 비음성의 구분 대신 음성이 존재할 확률을 이용하여 잡음을 추정하는 방식이다.

[0004] 즉, 잡음 섞인 음성신호가 마이크 어레이(10)를 통해 입력된다. 이 신호는 windowing과정과 푸리에 변환(Fourier transform)을 거쳐 주파수축으로 바뀌게 된다.

[0005] 주파수축으로 변환된 신호를 가지고 노이즈 스피치(noisy speech)의 로컬 에너지(local energy)를 [수학식 1]을 통해 구한다.

[0006] [수학식 1]

$$S_f(k, \ell) = \sum_{i=-w}^w b(i) |Y(k-i, \ell)|^2,$$

[0008] 여기서, $|Y()|^2$ 는 입력된 노이즈 스피치의 파워 스펙트럼(power spectrum), k 는 주파수 인덱스, l 은 frame 인덱스를 의미하며, b =window function, window length=2w+1이다.

[0009] [수학식 2]

$$S(k, s) = \alpha_s S(k, s-1) + (1 - \alpha_s) S_f(k, s), 0 < \alpha_s < 1 = \text{smoothing parameter}$$

[0011] 여기서, k 는 주파수 인덱스, l 은 frame 인덱스를 의미하며, b =window function, window length=2w+1이다.

[0012] 이어서, 로컬 에너지의 최소값을 [수학식 3]을 통해 구한다.

[0013] [수학식 3]

$$S_{\min}(k,s)=\min \{S_{\min}(k,s-1),S(k,s)\}$$

앞에서 구한 두 값을 가지고 노이즈 스펙치의 로컬 에너지와 최소값 사이의 비를 [수학식 4]를 통해 구한다.

[수학식 4]

$$S_r(k,s)AS(k,s)/S_{\min}(k,s)$$

한편, 임계 값 ' δ '을 정하여 " $S_r(k,s)>\delta$ "이면 음성이 존재한다고 판단하고 그렇지 않은 경우 음성이 없다고 판단하며, 이는 [수학식 5]와 같이 나타낼 수 있다.

[수학식 5]

$$I(k,s)=1 \text{ if } S_r(k,s)>\delta \text{ and } I(k,s)=0 \text{ otherwise}$$

위에서 구한 음성 존재 유무를 판단하는 파라메타를 이용하여 [수학식 6]을 참조하여 음성이 존재할 확률 값을 구한다.

[수학식 6]

$$\hat{p}(k,\ell)=\alpha_p \hat{p}(k,\ell-1)+(1-\alpha_p)I(k,\ell), \text{ where } \alpha_p(0<\alpha_p<1) \text{ is smoothing parameter}$$

이어서, 상기 음성이 존재할 확률 값을 이용하여 [수학식 7]과 같이 노이즈 파워를 추정한다.

[수학식 7]

$$\hat{\lambda}_d(k,\ell+1)=\hat{\lambda}_d(k,\ell)\hat{p}(k,\ell)+[\alpha_d\hat{\lambda}_d(k,\ell)+(1-\alpha_d)|Y(k,\ell)|^2](1-\hat{p}(k,\ell))$$

$$=\widetilde{\alpha}_d(k,\ell)\hat{\lambda}_d(k,\ell)+[1-\widetilde{\alpha}_d(k,\ell)]|Y(k,\ell)|^2$$

$$\text{where } \widetilde{\alpha}_d(k,\ell)\equiv\alpha_d+(1-\alpha_d)\hat{p}(k,\ell)$$

여기서, ' $\hat{\lambda}_d$ '는 추정된 잡음이다.

상기 [수학식 7]에서 알 수 있듯이 음성이 존재하는 경우 그전의 추정된 잡음 값을 사용하고 존재하지 않는 경우 그 전의 추정된 잡음 값과 들어온 신호의 값을 weight를 줘서 더해줌으로써 업데이트된 잡음의 파워를 구한다.

즉, 들어오는 신호의 음성존재유무를 판단하여 음성이 없는 곳(즉, 잡음구간)에서 잡음을 추정하는 방법이 MCRA이다.

두 번째 잡음 제거 방법은 Spectral subtraction based on minimum statistic으로 spectral subtraction은 노이즈 파워 추정(noise power estimation)이 가장 중요하다.

먼저, 입력 신호를 받아서 주파수 변환을 한 후 magnitude와 phase로 각각 분리한다.

분리된 값 중 phase 값은 그대로 유지하고 magnitude 값만 이용한다.

노이즈만 있는 구간에서의 magnitude값을 추정하여 그 값을 입력 신호의 magnitude값에 빼준다.

이 값과 phase 값을 이용하여 신호를 복원시키면 잡음이 제거된 신호를 얻을 수 있다.

노이즈만 있는 구간을 추정하는 것은 잡음 섞인 신호의 단시간 보조대역 파워 추정(short time subband power estimate)를 이용한다.

단시간 보조대역 파워 추정 값을 구해보면 도 2에 도시된 바와 같이 피크(peaks)와 밸리(valleys)가 생기게 된다.

여기서, Peaks 값은 음향 활동(speech activity) 구간으로 볼 수 있기 때문에 valleys가 있는 구간을 추정함으로써 노이즈 파워를 구할 수 있다.

- [0038] 이렇게 구한 노이즈 부분을 이용하여 앞에서 얘기한 spectral subtraction 방법을 이용하여 잡음을 제거하는 방식이 spectral subtraction based on minimum statistic이다.
- [0039] 그러나 종래 잡음 제거 방식은 돌발 잡음의 변화를 감지하여 적절히 잡음 추정에 반영하지 못하는 문제점이 있다. 즉, 실내 환경에서 발생하는 발자국 소리, 키보드 타이핑 소리 등 지속시간이 짧으나 에너지의 크기가 음성의 그것만큼 큰 잡음에 대해서는 효과가 적었다.
- [0040] 따라서 이러한 부정확한 잡음추정으로 인하여 잔여 잡음이 남는 문제점이 있으며, 잔여잡음은 음성통신에서 사용자에게 듣기 불편함을 주거나 음성인식기의 오작동을 일으켜 음성인식 제품의 성능을 하락시키는 문제점이 있었다.
- [0041] 이러한 이유는 음성/비음성의 구분의 경우 에너지의 크기 혹은 신호 대 잡음비(signal to noise ratio, SNR)를 기준으로 이 값들이 큰 경우 음성구간으로 인식하고 이 값들이 작을 경우 비 음성구간으로 간주하기 때문에 만약 음성과 비슷한 크기의 에너지를 가지는 주변잡음이 입력될 경우 잡음 추정 및 갱신을 하지 않기 때문이다.

발명의 내용

해결 하고자하는 과제

- [0042] 따라서 본 발명은 상기와 같은 종래의 문제점을 해결하기 위한 것으로서, 본 발명의 목적은 돌발잡음 환경에서 돌발잡음을 효과적으로 제거할 수 있는 음원 검출 시스템에서 돌발잡음 추정/제거 장치 및 방법을 제공하는 데 있다.

과제 해결수단

- [0043] 상기한 목적을 달성하기 위한 본 발명에 따른 음원 검출 시스템에서 돌발잡음 추정/제거 장치의 일 측면에 따르면, 음성신호를 수신하는 마이크 어레이부; 수신된 신호에 설정된 범위가 겹쳐지도록 윈도우를 씌우는 범위설정부; 상기 윈도우된 신호를 주파수축으로 변환하기 위한 푸리에 변환을 적용하는 주파수 변환부; 상기 주파수 변환부를 통해 변환된 푸리에 변환값과 이전 값의 상관값을 비교하고 상기 윈도우되어 푸리에 변환된 값의 전력(energy)를 계산하는 상관관계 측정부; 및 상관 비교값과 전력값을 이용해 잡음을 검출하는 돌발잡음 검출부를 포함한다.
- [0044] 상기 마이크 어레이부를 통해 입력된 아날로그 음성신호를 디지털 음성 신호로 변환하는 A/D변환부를 더 포함한다.
- [0045] 그리고 상기 주파수 변환부 뒷단에 주파수 변환부를 통해 변환된 푸리에 변환값을 저장하는 버퍼를 더 포함한다.
- [0046] 여기서, 상기 범위설정부는, 32ms 길이를 갖는 hanning 윈도우이다.
- [0047] 한편, 상기 상관관계 측정부는, 상관 비교 값이 기 설정된 임계값을 초과하는지의 여부를 판단하는 상관관계 판단부와 상기 상관관계 판단부를 통해 상관관계 값이 기 설정된 임계값을 초과하면 잡음으로 인식하는 잡음 검출부를 더 포함한다.
- [0048] 또한 상기 상관관계 판단부는 크로스 파워 스펙트럼을 이용하여 스펙트럼의 magnitude 값을 제공한 후 주파수 전 영역에서 합하여 해당 프레임의 에너지($\gamma(s)$)로 정의하고, 크로스 파워 스펙트럼을 통해 에너지를 검출한 프레임과, 임의의 주파수에서 로컬 에너지와 최소 통계치를 기반의 추정된 잡음과의 비($S_r(s,k)$)로 정의하며, 해당 프레임의 에너지($\gamma(s)$)와 추정된 잡음과의 비($S_r(s,k)$)에 임계값을 주어서 $\gamma(s)$ 이 주어진 해당 임계값보다 작고 $S_r(s,k)$ 이 주어진 임계값보다 클 경우 돌발잡음이 존재한다고 판단한다.
- [0049] 상기 돌발잡음 검출부는 기존의 MCRA 잡음 추정기법에 돌발잡음을 검출하는 파라메타를 적용시켜 이하 [수학식 8], [수학식 9], [수학식 10]을 통해 돌발잡음을 구한다.

- [0050] [수학식 8]
- [0051]
$$\lambda(k, \ell + 1) = \alpha(k, \ell) \hat{\lambda}(k, \ell + 1) + (1 - \alpha(k, \ell)) |Y(k, \ell)|^2$$
- [0052] 여기서, $\hat{\lambda}(k, \ell + 1)$ 는 추정된 잡음, K는 주파수 인덱스, L은 프레임 인덱스 이다.
- [0053] [수학식 9]
- [0054]
$$\alpha(k, \ell) = \tilde{\alpha}(k, \ell) + (1 - \tilde{\alpha}(k, \ell)) p(k, \ell) (1 - I_f(k, \ell))$$
- [0055] 여기서, p(k,l)는 음성존재 확률, K는 주파수 인덱스, L은 프레임 인덱스 이다.
- [0056] [수학식 10]
- [0057]
$$\tilde{\alpha}(k, \ell) = \alpha_{ds} + (\alpha_{dt} - \alpha_{ds}) I_f(k, \ell)$$
- [0058] 여기서, $\alpha_{ds}=0.95$, $\alpha_{dt}=0.05$ 로서 각각 stationary 잡음, 돌발 잡음 구간에서의 update 계수이다.
- [0059] 상기 돌발잡음 검출부는, 돌발잡음이 검출되지 않으면 고정(stationary) 잡음으로 추정한다.
- [0060] 본 발명에 따른 음원 검출 시스템에서 돌발잡음 추정/제거 방법의 일 측면에 따르면, 음성신호를 수신하는 단계; 수신된 신호에 설정된 범위가 겹쳐지도록 윈도우를 씌우는 단계; 윈도우된 신호를 주파수축으로 변환하기 위한 푸리에 변환을 적용하는 단계; 변환된 푸리에 변환값과 이전 값의 상관값을 비교하고 윈도우되어 푸리에 변환된 값의 전력(energy)를 계산하는 단계; 및 상관 비교값과 전력값을 이용해 잡음을 검출하는 단계를 포함한다.
- [0061] 입력된 아날로그 음성신호를 디지털 음성 신호로 변환하는 단계를 더 포함한다.
- [0062] 그리고, 푸리에 변환을 적용하는 단계 이후 변환된 푸리에 변환값을 저장하는 단계를 더 포함한다.
- [0063] 여기서, 상기 윈도우를 씌우는 단계는, 32ms 길이를 갖는 hanning 윈도우이다.
- [0064] 상기 윈도우되어 푸리에 변환된 값의 전력(energy)를 계산하는 단계는, 상관 비교 값이 기 설정된 임계값을 초과하는지의 여부를 판단하는 단계와 상기 판단 단계에서 상관관계 값이 기 설정된 임계값을 초과하면 잡음으로 인식하는 단계를 더 포함한다.
- [0065] 상기 상관관계 판단 단계는, 크로스 파워 스펙트럼을 이용하여 스펙트럼의 magnitude 값을 제공한 후 주파수 전 영역에서 합하여 해당 프레임의 에너지($\gamma(s)$)로 정의하는 단계; 크로스 파워 스펙트럼을 통해 에너지를 검출한 프레임과, 임의의 주파수에서 로컬 에너지와 최소 통계치를 기반의 추정된 잡음과의 비($S_{\mu}(s, k)$)로 정의하는 단계; 및 해당 프레임의 에너지($\gamma(s)$)와 추정된 잡음과의 비($S_{\mu}(s, k)$)에 임계값을 주어서 ' $\gamma(s)$ '이 주어진 해당 임계값보다 작고 ' $S_{\mu}(s, k)$ '이 주어진 임계값보다 클 경우 돌발잡음이 존재한다고 판단하는 단계를 포함한다.
- [0066] 상기 돌발잡음 검출 단계는, 기존의 MCRA 잡음 추정기법에 돌발잡음을 검출하는 파라메타를 적용시켜 이하 [수학식 8], [수학식 9], [수학식 10]을 통해 돌발잡음을 구한다.
- [0067] [수학식 8]
- [0068]
$$\lambda(k, \ell + 1) = \alpha(k, \ell) \hat{\lambda}(k, \ell + 1) + (1 - \alpha(k, \ell)) |Y(k, \ell)|^2$$
- [0069] 여기서, $\hat{\lambda}(k, \ell + 1)$ 는 추정된 잡음, K는 주파수 인덱스, L은 프레임 인덱스 이다.
- [0070] [수학식 9]
- [0071]
$$\alpha(k, \ell) = \tilde{\alpha}(k, \ell) + (1 - \tilde{\alpha}(k, \ell)) p(k, \ell) (1 - I_f(k, \ell))$$
- [0072] 여기서, p(k,l)는 음성존재 확률, K는 주파수 인덱스, L은 프레임 인덱스 이다.

[0073] [수학식 10]

$$\hat{\alpha}(k, \ell) = \alpha_{ds} + (\alpha_{dt} - \alpha_{ds})I(k, \ell)$$

[0075] 여기서, $\alpha_{ds}=0.95$, $\alpha_{dt}=0.05$ 로서 각각 stationary 잡음, 돌발 잡음 구간에서의 update 계수이다.

[0076] 반면에, 돌발잡음이 검출되지 않으면 돌발잡음 검출 단계는 고정(stationary) 잡음으로 추정한다.

효 과

[0077] 상술한 바와 같이 본 발명에 의한 음원 검출 시스템에서 돌발잡음 추정/제거 장치 및 방법에 의하면, 돌발 잡음 환경에서도 효과적으로 잡음을 추정하고 제거할 수 있는 뛰어난 효과가 있다.

발명의 실시를 위한 구체적인 내용

[0078] 이하, 본 발명에 따른 음원 검출 시스템에서 돌발잡음 추정/제거 장치 및 방법에 대한 바람직한 실시 예에 대하여 첨부한 도면을 참조하여 상세하게 살펴보기로 한다. 이 때, 아래에서 설명하는 시스템 구성은 본 발명의 설명을 위해서 인용한 시스템으로써 아래 시스템으로 본 발명을 한정하지 않음을 이 분야의 통상의 지식을 가진 자라면 이해해야할 것이다.

[0079] 도 3은 본 발명에 따른 음원 검출 시스템에서 돌발잡음 추정/제거 장치의 구성을 나타낸 도면으로서, 본 발명에 따른 음원 검출 시스템에서 돌발잡음 추정/제거 장치는 범위설정부(100), 주파수 변환부(200), 임시 저장부(300), 상관관계 측정부(400), 상관관계 판단부(500) 및 돌발잡음 검출부(600)를 포함한다.

[0080] 범위설정부(100)는 잡음 섞인 음성 신호는 마이크 어레이(10)를 통해 입력되며 hanning 윈도우를 씌워 프레임으로 구분한다.

[0081] 주파수 변환부(200)는 범위설정부(100)를 통해 프레임별로 구분된 신호를 주파수축으로 변환한다.

[0082] 임시 저장부(300)는 상기 주파수 변환부(200)를 통해 변환된 프레임별 FFT 값을 임시 저장한다.

[0083] 또한 상관관계 측정부(400)는 현재 입력시점의 프레임과 기 설정된 시간만큼 떨어진 이후의 프레임간의 상관관계 정도를 측정한다.

[0084] 그리고 상관관계 판단부(500)는 상관관계 측정부(400)를 통해 측정된 상관관계 값이 기 설정된 임계값을 초과하는지의 여부를 판단한다. 여기서, 상기 상관관계 판단부(500)는, 크로스 파워 스펙트럼을 이용하여 스펙트럼의 magnitude 값을 제공한 후 주파수 전 영역에서 합하여 해당 프레임의 에너지($\gamma(s)$)로 정의하고, 크로스 파워 스펙트럼을 통해 에너지를 검출한 프레임과, 임의의 주파수에서 로컬 에너지와 최소 통계치를 기반의 추정된 잡음과의 비($S_{\mu}(s, k)$)로 정의하며, 해당 프레임의 에너지($\gamma(s)$)와 추정된 잡음과의 비($S_{\mu}(s, k)$)에 임계값을 주어서 $\gamma(s)$ 이 주어진 해당 임계값보다 작고 $S_{\mu}(s, k)$ 이 주어진 임계값보다 클 경우 돌발잡음이 존재한다고 판단한다.

[0085] 또한 돌발잡음 검출부(600)는 상관관계 판단부(500)를 통해 상관관계 값이 기 설정된 임계값을 초과하면 돌발잡음을 검출한다. 이때, 상기 돌발잡음 검출부(600)는, 기존의 MCRA 잡음 추정기법에 돌발잡음을 검출하는 파라메타를 적용시켜 이하 [수학식 8], [수학식 9], [수학식 10]을 통해 돌발잡음을 구한다.

[0086] [수학식 8]

$$\hat{\lambda}(k, \ell + 1) = \alpha(k, \ell)\hat{\lambda}(k, \ell + 1) + (1 - \alpha(k, \ell))|Y(k, \ell)|^2$$

[0088] 여기서, $\hat{\lambda}(k, s+1)$ 는 추정된 잡음, K는 주파수 인덱스, L은 프레임 인덱스 이다.

[0089] [수학식 9]

$$\alpha(k, \ell) = \hat{\alpha}(k, \ell) + (1 - \hat{\alpha}(k, \ell))p(k, \ell)(1 - I(k, \ell))$$

[0091] 여기서, p(k, l)는 음성존재 확률, K는 주파수 인덱스, L은 프레임 인덱스 이다.

- [0092] [수학식 10]
- $$\tilde{\alpha}(k, \ell) = \alpha_{ds} + (\alpha_{dt} - \alpha_{ds})I(k, \ell)$$
- [0093]
- [0094] 여기서, $\alpha_{ds}=0.95$, $\alpha_{dt}=0.05$ 로서 각각 stationary 잡음, 돌발 잡음 구간에서의 update 계수이다.
- [0095] 만약, 돌발잡음이 검출되지 않으면 돌발잡음 검출부(600)는 고정(stationary) 잡음으로 추정한다.
- [0096] 상기한 각 구성들에 대한 일반적인 기능 및 각각의 상세한 동작에 대하여는 그 설명을 생략하고, 본 발명에 상응하는 동작 위주로 그 동작들을 설명하기로 한다.
- [0097] 먼저, 범위설정부(100)는 잡음 섞인 음성 신호가 마이크 어레이(10)를 통해 입력되며 32ms 길이의 hanning 윈도우를 씌워 프레임으로 구분한다. 이때 이동구간은 16ms이다.
- [0098] 이어서, 주파수 변환부(200)는 범위설정부(100)를 통해 프레임별로 구분된 신호를 주파수축으로 변환한다.
- [0099] 주파수 변환부(200)를 통해 변환된 프레임별 FFT 값은 임시 저장부(300)에 임시 저장된다.
- [0100] 이후, 상관관계 측정부(400)는 현재 입력시점의 프레임(1)과 기 설정된 시간(N)만큼 떨어진 이후의 프레임(1+N)간의 상관관계 정도를 측정한다. 여기서, N은 최소 100ms 이상의 구간에 해당하는 프레임의 개수이다.
- [0101] 이어서, 상관관계 판단부(500)는 상관관계 측정부(400)를 통해 측정된 상관관계 값이 기 설정된 임계값을 초과하는지의 여부를 판단한다.
- [0102] 이러한 상기 상관관계 판단부(500)는, 크로스 파워 스펙트럼을 이용하여 스펙트럼의 magnitude 값을 제공한 후 주파수 전 영역에서 합하여 해당 프레임의 에너지($\gamma(s)$)로 정의하고, 크로스 파워 스펙트럼을 통해 에너지를 검출한 프레임과, 임의의 주파수에서 로컬 에너지와 최소 통계치를 기반의 추정된 잡음과의 비($S_r(s, k)$)로 정의하며, 해당 프레임의 에너지($\gamma(s)$)와 추정된 잡음과의 비($S_r(s, k)$)에 임계값을 주어서 $\gamma(s)$ 이 주어진 해당 임계값보다 작고 $S_r(s, k)$ 이 주어진 임계값보다 클 경우 돌발잡음이 존재한다고 판단한다.
- [0103] 만약, 상관관계 판단부(500)를 통해 상관관계 값이 기 설정된 임계값을 초과하면 돌발잡음 검출부(600)는 돌발잡음을 검출한다.
- [0104] 이러한 돌발잡음 검출부(600)는, 기존의 MCRA 잡음 추정기법에 돌발잡음을 검출하는 파라메타를 적용시켜 이하 [수학식 8], [수학식 9], [수학식 10]을 통해 돌발잡음을 구한다.
- [0105] [수학식 8]
- $$\hat{\lambda}(k, \ell + 1) = \alpha(k, \ell)\hat{\lambda}(k, \ell + 1) + (1 - \alpha(k, \ell))|Y(k, \ell)|^2$$
- [0106]
- [0107] 여기서, $\hat{\lambda}(k, \ell + 1)$ 는 추정된 잡음, K는 주파수 인덱스, L은 프레임 인덱스 이다.
- [0108] [수학식 9]
- $$\alpha(k, \ell) = \tilde{\alpha}(k, \ell) + (1 - \tilde{\alpha}(k, \ell))p(k, \ell)(1 - I(k, \ell))$$
- [0109]
- [0110] 여기서, p(k,1)는 음성존재 확률, K는 주파수 인덱스, L은 프레임 인덱스 이다.
- [0111] [수학식 10]
- $$\tilde{\alpha}(k, \ell) = \alpha_{ds} + (\alpha_{dt} - \alpha_{ds})I(k, \ell)$$
- [0112]
- [0113] 여기서, $\alpha_{ds}=0.95$, $\alpha_{dt}=0.05$ 로서 각각 stationary 잡음, 돌발 잡음 구간에서의 update 계수이다.
- [0114] 반면에 돌발잡음이 검출되지 않으면 돌발잡음 검출부(600)는, 고정(stationary) 잡음으로 추정한다.
- [0115] 그러면, 상기와 같은 구성을 가지는 본 발명에 따른 음원 검출 시스템에서 돌발잡음 추정/제거 방법에 대해 도 4를 참조하여 설명하기로 한다.
- [0116] 잡음 섞인 음성 신호는 마이크 어레이를 통해 입력되며 hanning 윈도우를 씌워 프레임으로 구분한다(S1).

- [0117] 이어서, 프레임별로 구분된 신호를 주파수축으로 변환한다(S2).
- [0118] 이후, 변환된 프레임별 FFT 값을 임시 저장한다(S3).
- [0119] 이렇게 저장된 프레임별 FFT 값을 통해 현재 입력시점의 프레임과 기 설정된 시간만큼 떨어진 이후의 프레임간의 상관관계 정도를 측정한다(S4).
- [0120] 이후, 측정된 상관관계 값이 기 설정된 임계값을 초과하는지의 여부를 판단한다(S5).
- [0121] 이하, 하기에서는 상관관계 판단 단계(S5)의 세부 단계에 대하여 도 5를 참조하여 설명한다.
- [0122] 먼저, 크로스 파워 스펙트럼을 이용하여 스펙트럼의 magnitude 값을 제공한 후 주파수 전 영역에서 합하여 해당 프레임의 에너지($\gamma(s)$)로 정의한다(S51).
- [0123] 이어서, 크로스 파워 스펙트럼을 통해 에너지를 검출한 프레임과, 임의의 주파수에서 로컬 에너지와 최소 통계치를 기반의 추정된 잡음과의 비($S_{\lambda(s,k)}$)로 정의한다(S52).
- [0124] 이후, 해당 프레임의 에너지($\gamma(s)$)가 기 설정된 임계값 보다 큰지의 여부를 판단한다(S53).
- [0125] 상기 프레임의 에너지($\gamma(s)$)가 기 설정된 임계값 보다 작으면, 추정된 잡음과의 비($S_{\lambda(s,k)}$)가 기 설정된 임계값 보다 큰지의 여부를 판단한다(S54).
- [0126] 한편, 상관관계 판단 단계(S5)를 통해 상관관계 값이 기 설정된 임계값을 초과하면 돌발잡음을 검출하여 제거한다(S6). 이러한, 상기 돌발잡음 검출 단계(S6)는 기존의 MCRA 잡음 추정기법에 돌발잡음을 검출하는 파라메타를 적용시켜 이하 [수학식 8], [수학식 9], [수학식 10]을 통해 돌발잡음을 구한다.
- [0127] [수학식 8]
- [0128]
$$\lambda(k, \ell + 1) = \alpha(k, \ell) \lambda(k, \ell) + (1 - \alpha(k, \ell)) |Y(k, \ell)|^2$$
- [0129] 여기서, $\lambda(k, \ell + 1)$ 는 추정된 잡음, K는 주파수 인덱스, L은 프레임 인덱스 이다.
- [0130] [수학식 9]
- [0131]
$$\alpha(k, \ell) = \tilde{\alpha}(k, \ell) + (1 - \tilde{\alpha}(k, \ell)) p(k, \ell) (1 - I_f(k, \ell))$$
- [0132] 여기서, p(k,l)는 음성존재 확률, K는 주파수 인덱스, L은 프레임 인덱스 이다.
- [0133] [수학식 10]
- [0134]
$$\tilde{\alpha}(k, \ell) = \alpha_{ds} + (\alpha_{dt} - \alpha_{ds}) I_f(k, \ell)$$
- [0135] 여기서, $\alpha_{ds}=0.95$, $\alpha_{dt}=0.05$ 로서 각각 stationary 잡음, 돌발 잡음 구간에서의 update 계수이다.
- [0136] 상기 프레임의 에너지($\gamma(s)$)가 기 설정된 임계값 보다 크거나, 추정된 잡음과의 비가 기 설정된 임계값 보다 작아 돌발잡음이 검출되지 않으면, 고정(stationary) 잡음으로 추정한다(S7).
- [0137] 이상에서 본 발명은 기재된 구체적인 실시 예에 대해서만 상세히 설명하였지만 본 발명의 기술 사상 범위 내에서 다양한 변형 및 수정이 가능함은 당업자에게 있어서 명백한 것이며, 이러한 변형 및 수정이 첨부된 특허청구 범위에 속함은 당연한 것이다.

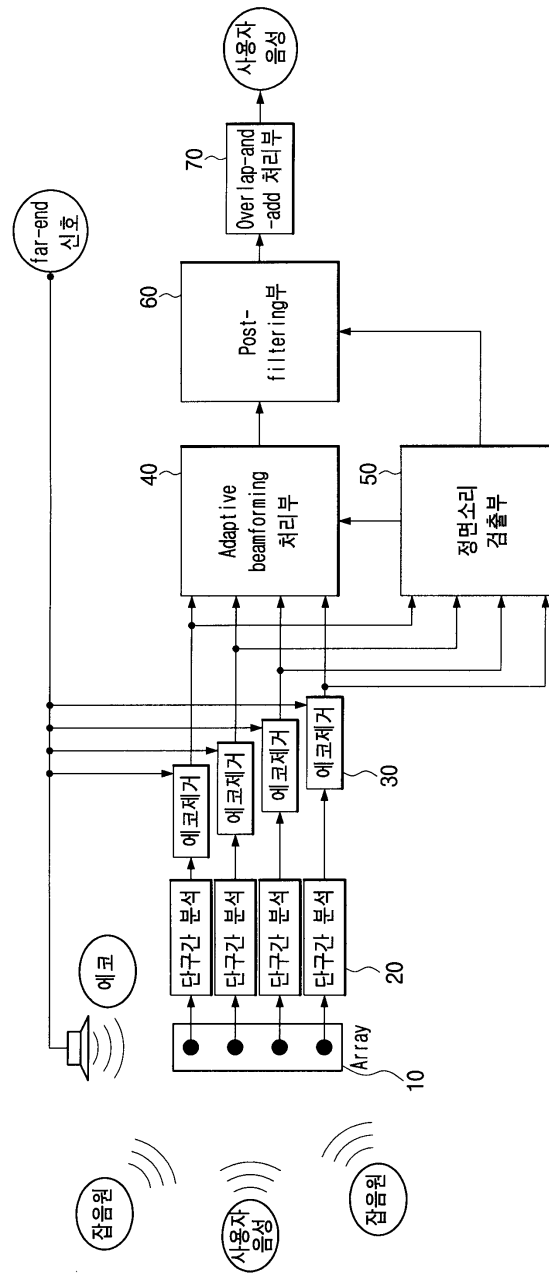
도면의 간단한 설명

- [0138] 도 1은 종래 마이크 어레이를 이용한 방향성 잡음 제거 시스템의 구성을 나타낸 기능블록도.
- [0139] 도 2는 종래 마이크 어레이를 이용한 방향성 잡음 제거 시스템에서 단시간 보조대역 파워 추정(short time subband power estimate) 값을 나타낸 도면.
- [0140] 도 3은 본 발명에 따른 음원 검출 시스템에서 돌발잡음 추정/제거 장치의 구성을 나타낸 기능블록도.
- [0141] 도 4는 본 발명에 따른 음원 검출 시스템에서 돌발잡음 추정/제거 방법을 나타낸 순서도.

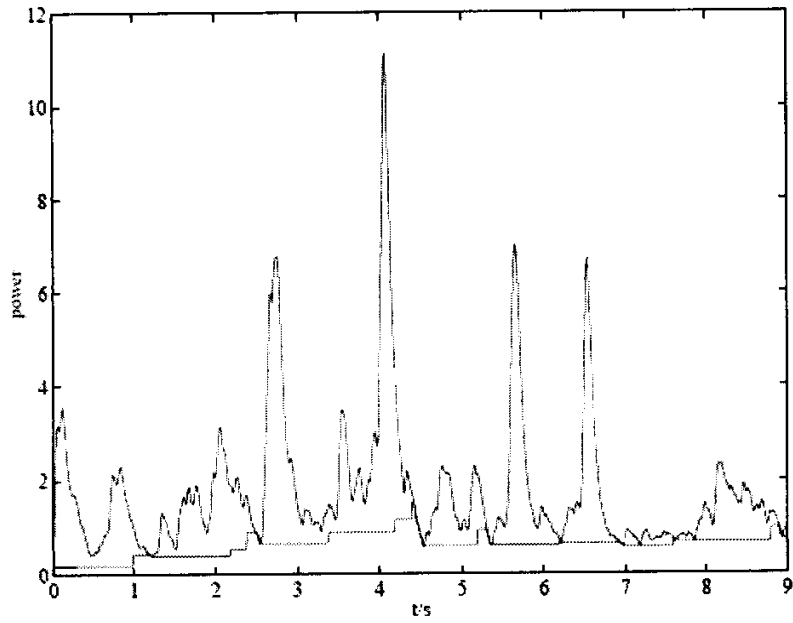
- | | | |
|--------|--|-------------------------------|
| [0142] | 도 5는 도 4에 따른 음원 검출 시스템에서 돌발잡음 추정/제거 방법에서 상관관계 판단 단계(S5)의 세부 단계를 나타낸 순서도이다. | |
| [0143] | <도면의 주요 부분에 대한 부호의 설명> | |
| [0144] | 10 : 적어도 하나 이상의 마이크 어레이 | 20 : 단구간 분석부 |
| [0145] | 30 : 에코 제거부 | 40 : adaptive beamforming 처리부 |
| [0146] | 50 : 정면소리 검출부 | 60 : Post-filtering부 |
| [0147] | 70 : Overlap and add 처리부 | |
| [0148] | 100 : 범위설정부 | 200 : 주파수 변환부 |
| [0149] | 300 : 임시 저장부 | 400 : 상관관계 측정부 |
| [0150] | 500 : 상관관계 판단부 | 600 : 돌발잡음 검출부 |

도면

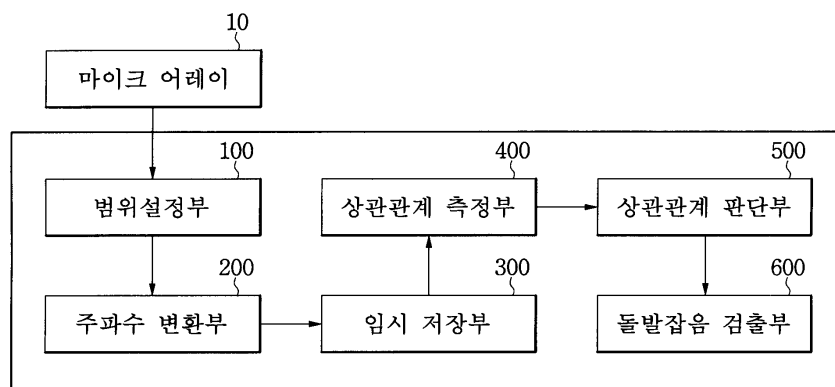
도면1



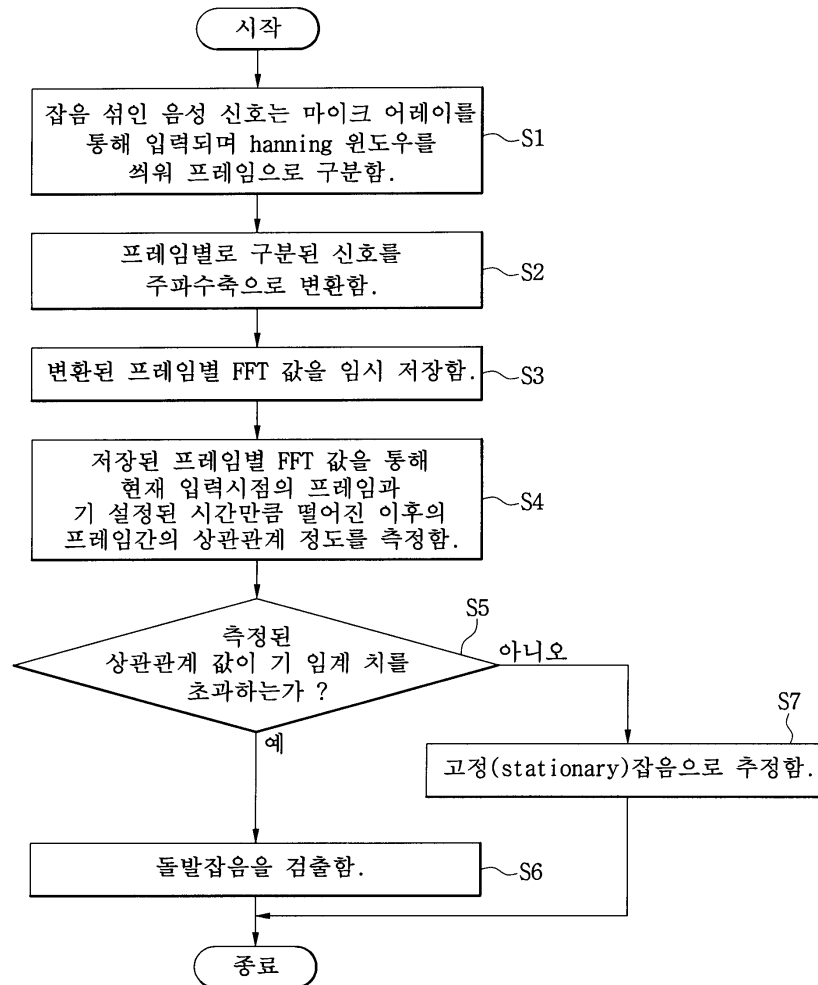
도면2



도면3



도면4



도면5

