



(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ

(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN)
CỤC SỞ HỮU TRÍ TUỆ

(11)



1-0042121

(51)^{2020.01} G10K 11/00

(13) B

(21) 1-2020-01143

(22) 28/02/2020

(45) 25/12/2024 441

(43) 25/08/2020 389

(73) Tập đoàn Công nghiệp - Viễn thông Quân Đội (VN)

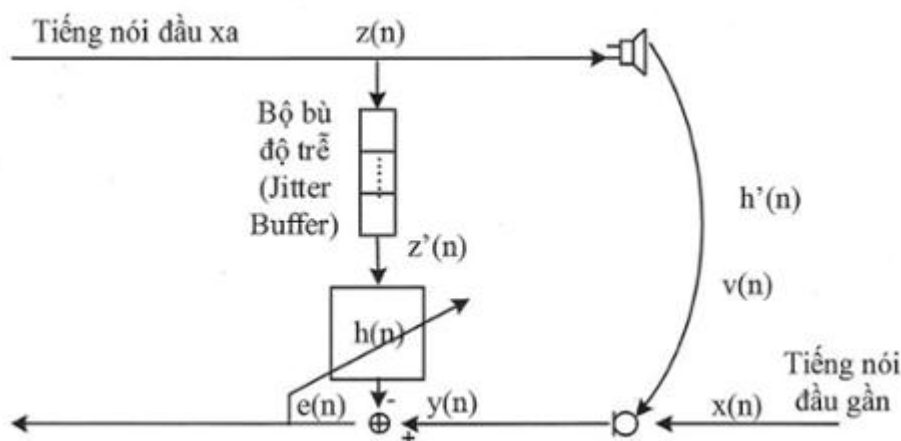
Lô D26 Khu đô thị mới Cầu Giấy, Phường Yên Hoà, Quận Cầu Giấy, Thành phố Hà Nội

(72) ĐINH VĂN PHONG (VN); ĐINH VIỆT QUÂN (VN); NGUYỄN THẾ HIẾU (VN); NGUYỄN HUY BÌNH (VN).

(74) Công ty TNHH Tư vấn Quốc Dân (NACILAW)

(54) PHƯƠNG PHÁP KHỬ TIẾNG VỌNG ĐỘ TRỄ LỚN

(57) Phương pháp khử tiếng vọng độ trễ lớn được đề xuất bao gồm ba bước, cụ thể: bước 1: Thu tiếng nói đầu xa; Bước 2: làm trễ tín hiệu tiếng nói đầu xa bằng bộ bù độ trễ; Bước 3: khử tiếng vọng độ trễ lớn. Phương pháp cốt lõi cho việc xử lý tiếng vọng trong sáng chế này là sử dụng bộ bù độ trễ kết hợp với bộ lọc thích nghi với thuật toán bình phương trung bình tối thiểu - LMS (least mean square) để khử hiệu quả các tiếng vọng độ trễ lớn mà không cần tăng kích thước bộ lọc.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề xuất phương pháp khử tiếng vọng độ trễ lớn. Cụ thể, phương pháp khử tiếng vọng độ trễ lớn được sử dụng trong lĩnh vực công nghệ thông tin – viễn thông, xử lý tín hiệu, xử lý tiếng nói.

Tình trạng kỹ thuật của sáng chế

Tiếng vọng là một vấn đề nghiêm trọng trong viễn thông, được định nghĩa là một phiên bản bị trễ và méo của tiếng nói gốc, sinh ra do sự phản xạ ngược trở về nguồn phát của tiếng nói gốc. Có hai cơ chế sinh ra tiếng vọng trong các hệ thống viễn thông. Thứ nhất, tham chiếu hình 1 – tiếng vọng gây ra do đường truyền, trong truyền dẫn tín hiệu điện thoại cố định (PSTN), tiếng vọng sinh ra do lệch phối hợp trở kháng giữa cuộn sơ cấp và thứ cấp trong bộ chuyển đổi hai dây sang bốn dây. Loại thứ hai sinh ra trên các thiết bị đầu cuối, do sự phản xạ tiếng nói từ loa tới chính microphone đặt gần đó (tham chiếu hình 2).

Nếu sử dụng một bộ phân tích tín hiệu giữa tiếng nói đầu xa và tiếng vọng phản xạ sẽ thu được dạng tín hiệu như trong hình 3. Trong đó tiếng vọng phản xạ xuất hiện là một phiên bản bị trễ và nhỏ hơn của tiếng nói gốc đầu xa.

Các phương pháp khử tiếng vọng trên hệ thống viễn thông đã được phát triển từ rất sớm với các công bố khoa học của Mr. Sondhi – Phòng thí nghiệm Bell. Ý tưởng chính của việc khử tiếng vọng là tổng hợp ra một bản sao của chính tiếng vọng đó (bằng cách cho tín hiệu đầu xa đi qua một bộ lọc thích nghi sử dụng thuật toán bình phương trung bình tối thiểu (Least Mean Square - LMS) và sau đó lấy tín hiệu bị vọng trừ đi tiếng vọng tổng hợp. Cụ thể được mô tả như trong hình 4.

Như trong hình 4, tiếng nói tại microphone $y(n)$ là sự tổng hợp của tiếng nói đầu gần $x(n)$ và tiếng vọng từ đầu xa $v(n)$.

$$y(n) = x(n) + v(n) = x(n) + z(n) * h'(n)$$

Ở đây $h'(n)$ là đáp ứng xung của môi trường. Để triệt tiêu hoàn toàn giá trị $v(n)$ trong tín hiệu tại microphone, cho tín hiệu tiếng nói đầu xa $z(n)$ đi qua bộ lọc thích nghi $h(n)$. Tín hiệu sau bộ lọc được trừ bởi tín hiệu tiếng nói tại microphone $y(n)$.

$$e(n) = y(n) - z(n) * h(n) = x(n) + z(n) * h'(n) - z(n) * h(n)$$

Trong điều kiện lý tưởng, chúng ta mong muốn $e(n) = x(n)$, tức là $z(n) * h'(n) - z(n) * h(n) = 0$. Để làm được điều này, các thông số bộ lọc thích nghi $h(n)$ cần phải hội tụ về $h'(n)$.

Người ta đã phát triển nhiều thuật toán để đưa $h(n)$ hội tụ về $h'(n)$ như: Bình phương trung bình tối thiểu - Least Mean Square (LMS), bình phương trung bình tối thiểu được chuẩn hóa - Normalized least mean square (NLMS), bình phương trung bình tối thiểu được

chuẩn hóa theo tỷ lệ - proportionate Normalized Least Mean Squares (PNLMS), bình phương tối thiểu theo đệ quy - recursive least square (RLS). Bảng 1 mô tả tóm tắt thuật toán bình phương trung bình tối thiểu (LMS).

Bảng 1. Tóm tắt thuật toán bình phương trung bình tối thiểu

Định nghĩa tham số ban đầu	$L = \text{filter length}$ (chiều dài bộ lọc) $\mu = \text{step size}$ (bước nhảy của thuật toán)
Khởi tạo các giá trị ban đầu	$h(0) = \text{zeros}(L)$ (các thành phần bộ lọc được khởi tạo = 0)
Vòng lặp xử lý và cập nhật giá trị bộ lọc	For $n = 0, 1, 2, \dots$ $\mathbf{z}(n) = [z(n), z(n-1), \dots, z(n-L+1)]^T$ $e(n) = y(n) - \mathbf{h}^H(n) \mathbf{z}(n)$ $\mathbf{h}(n+1) = \mathbf{h}(n) + \frac{\mu e^*(n) \mathbf{z}(n)}{\mathbf{z}^H(n) \mathbf{z}(n)}$

Theo các thuật toán này, độ trễ của tiếng vọng cần khử tỷ lệ thuận với chiều dài bộ lọc thích nghi $h(n)$. Có nghĩa, nếu độ trễ của tiếng vọng càng lớn, thì cần bộ lọc thích nghi $h(n)$ có chiều dài lớn tương ứng (tham chiếu hình 5).

Giả sử độ trễ tiếng vọng đo được là D (ms), với tần số lấy mẫu tiếng nói thông thường 8000 Hz, thì chiều dài bộ lọc được tính bằng $L = D \times 8$. Về lý thuyết, bất kỳ độ trễ D (ms) nào cũng tính được chiều dài bộ lọc L tương ứng. Tuy nhiên, khi bộ lọc có kích thước càng dài thì độ phức tạp tính toán càng lớn, ảnh hưởng đến hiệu năng và sự hiệu quả của thuật toán khi triển khai. Như thể hiện trong bảng 1, với mỗi bước lặp, thuật toán cần tới $3 \times L$ phép nhân và $2 \times L$ phép cộng. Với tín hiệu tiếng nói được lấy mẫu 8000 Hz/s, số lượng phép tính cần thực hiện trong 01 giây được tính như sau:

$$M = 3L \times 8000 \text{ phép nhân/s, và}$$

$$A = 2L \times 8000 \text{ phép cộng/s.}$$

Trong đó M là số lượng phép nhân và A là số lượng phép cộng.

Ví dụ, giả sử tiếng vọng có độ trễ 512 (ms), do đó cần thiết kế bộ lọc có kích thước $L = 512 \times 8 = 4096$, với bộ lọc này sẽ cần tới 98×10^6 phép nhân/s và 65×10^6 phép cộng/s. Do số lượng tính toán tăng nhanh như vậy khi chiều dài bộ lọc L tăng, nhiều hãng trên thế giới đã giới hạn chiều dài tối đa của bộ lọc thích nghi $h(n)$ mà họ có thể đáp ứng, đồng nghĩa với việc giới hạn độ trễ tiếng vọng tối đa mà sản phẩm của họ có thể khử được.

Ví dụ:

- Phần mềm khử tiếng vọng của hãng Adaptive Digital (Mỹ) cho phép khử tiếng vọng ở mức trễ tối đa 512ms.

- Phần mềm khử tiếng vọng của hãng Vocal cho phép khử tiếng vọng ở mức trễ tối đa 128ms.

Với các độ trễ lớn hơn nữa thì hiện chưa có phương pháp xử lý, vì vậy, sáng chế này tập trung giải quyết vấn đề này.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là cải tiến hiệu quả khử tiếng vọng đã biết ở trên. Sáng chế tập trung vào việc khử tiếng vọng độ trễ lớn dựa trên việc xây dựng bộ khử tiếng vọng độ trễ lớn và rất lớn, vượt qua các mức trễ tối đa mà các hãng như Adaptive Digital, Vocal đạt được ($> 512\text{ms}$).

Phương pháp khử tiếng vọng độ trễ lớn bao gồm ba bước, cụ thể như sau: bước 1: thu tiếng nói đầu xa; bước 2: làm trễ tín hiệu tiếng nói đầu xa bằng bộ bù độ trễ; bước 3: khử tiếng vọng độ trễ lớn.

Mô tả vắn tắt các hình vẽ

Hình 1 là hình vẽ minh họa tiếng vọng gây ra do lệch phối hợp trở kháng giữa đường truyền hai dây và bốn dây trên đường truyền điện thoại cố định PSTN;

Hình 2 là hình vẽ minh họa tiếng vọng gây ra ở thiết bị đầu cuối do phản xạ từ loa đến microphone đặt gần;

Hình 3 là hình vẽ phân tích, biểu diễn tín hiệu tiếng nói gốc và tiếng vọng phản xạ;

Hình 4 là hình vẽ minh họa tiếng vọng sử dụng bộ lọc thích nghi;

Hình 5 là hình vẽ biểu thị mối quan hệ giữa độ trễ và chiều dài bộ lọc L;

Hình 6 là hình vẽ biểu thị tín hiệu tiếng vọng độ trễ lớn;

Hình 7 là hình vẽ minh họa việc tìm giải pháp giảm độ trễ giữa tín hiệu tiếng nói đầu xa và tín hiệu tiếng vọng đầu gần;

Hình 8 là sơ đồ phương pháp khử tiếng vọng độ trễ lớn, kết hợp giữa bộ lọc thích nghi và bộ bù độ trễ (Jitter Buffer);

Hình 9 là hình vẽ mô tả một bộ bù độ trễ;

Hình 10 là hình vẽ mô hình triển khai khử tiếng vọng độ trễ lớn áp dụng trên mạng viễn thông Viettel;

Hình 11 là hình vẽ kết quả đo kiểm tỷ số tín hiệu/nhiều (Signal to Noise Ratio – SNR) chứng minh hiệu quả của sáng chế so với phương pháp cũ.

Mô tả chi tiết sáng chế

Như đã trình bày trong phần trước, với độ trễ tiếng vọng 512 ms, theo cách làm thông thường cần thiết kế bộ lọc có chiều dài lên tới 4096 (hệ số), điều này làm tăng cao độ phức tạp thuật toán với việc cần tới ~98 triệu phép nhân/giây và ~65 triệu phép cộng/giây. Trong khi năng lực tính toán của phần cứng máy tính là hữu hạn, độ phức tạp thuật toán cao sẽ dẫn

đến giảm số lượng cuộc gọi có thể xử lý đồng thời của hệ thống. Vấn đề sẽ càng nghiêm trọng nếu độ trễ tiếng vọng tiếp tục tăng lên 1024 ms, 2048 ms...

Tham chiếu hình 6, hình 7, nếu không muốn tăng chiều dài bộ lọc thích nghi L , cần thiết phải làm giảm độ trễ giữa tín hiệu tiếng nói đầu xa và tiếng vọng đầu gần về mức mà bộ lọc chiều dài L có thể xử lý tốt được.

Để đạt được mục đích đề ra, phương pháp khử tiếng vọng độ trễ lớn bao gồm ba bước, nội dung các bước cụ thể như sau:

- Bước 1: thu tiếng nói đầu xa.

Tiếng nói đầu xa là nguyên liệu đầu vào và là đối tượng tạo ra tiếng vọng. Tiếng nói đầu xa được thu bởi ống thu thanh (microphone) của điện thoại dưới dạng tín hiệu điện, và lưu trữ trong hệ thống dưới dạng một chuỗi giá trị nhị phân.

- Bước 2: làm trễ tín hiệu tiếng nói đầu xa bằng bộ bù độ trễ.

Để giảm độ trễ giữa tín hiệu tiếng nói đầu xa và tín hiệu tiếng vọng đầu gần, chúng tôi thực hiện bằng việc làm trễ tín hiệu tiếng nói đầu xa trước khi thực hiện thuật toán bình phương trung bình tối thiểu - LMS. Tại bước này, sử dụng một bộ bù độ trễ (jitter buffer) đặt phía trước bộ lọc thích nghi $h(n)$, tạo thành cấu trúc như hình 8.

Với cấu trúc này, chỉ cần đo kiểm được độ trễ D giữa tín hiệu tiếng nói đầu xa và tiếng vọng đầu gần, chúng ta hoàn toàn có thể tính toán kích thước bộ bù độ trễ hợp lý để đưa độ trễ về khoảng giới hạn mà chúng ta mong muốn để bộ lọc thích nghi $h(n)$ có thể xử lý hiệu quả được.

Theo đó, tiếng nói đầu xa được đóng gói thành các khung có chiều dài T ms, trước khi đưa qua bộ lọc thích nghi $h(n)$, được đưa qua bộ bù độ trễ (Jitter Buffer) theo nguyên tắc vào trước ra trước – First In First Out (FIFO). Bộ bù độ trễ có tác dụng làm chậm lại luồng tín hiệu tiếng nói đầu xa $z(n)$, thành tín hiệu $z'(n)$.

~~Một khung dữ liệu tiếng nói đầu xa $z(n)$ (dài T ms) được truyền đi ở thời điểm t_0 , và tiếng vọng đầu xa $v(n)$ quay trở lại hệ thống ở thời điểm t_1 , độ trễ tiếng vọng được tính bằng:~~

$$D = t_1 - t_0$$

Độ trễ này có thể đo kiểm được trong các hệ thống viễn thông thực tế.

Để xử lý với bộ lọc thích nghi $h(n)$ có chiều dài L , thì kích thước của bộ bù độ trễ được tính toán theo công thức như sau:

$$S = \frac{\left(\frac{D \times 8000}{10^3} \right) - L}{\left(\frac{T \times 8000}{10^3} \right)}$$

- Bước 3: Khử tiếng vọng độ trễ lớn. Sau khi được bù độ trễ, độ lệch giữa tín hiệu đưa vào bộ lọc thích nghi sử dụng thuật toán bình phương trung bình tối thiểu – LMS và tiếng vọng thực tế đã ở trong ngưỡng nhỏ mà ở đó bộ lọc thích nghi chiều dài L đã có thể loại bỏ được tiếng vọng.

Ví dụ thực hiện sáng chế

Tham chiếu hình 10 và hình 11, sáng chế đã được áp dụng thành công trong việc khử tiếng vọng trên đường truyền điện thoại cố định (PSTN) trong mạng lưới Viettel. Các tiếng vọng đo được có độ trễ rất lớn >512 ms, đã được khử hoàn toàn với mô hình sáng chế này mà chỉ cần dùng bộ lọc thích nghi có chiều dài 128ms, qua đó đảm bảo hiệu năng hệ thống.

Các phương pháp đo đạc thực tế sử dụng phần mềm đánh giá chất lượng thoại theo chuẩn ITU P.863 cho thấy tỷ số tín hiệu/nhiều (Signal to Noise Ratio - SNR) khi sử dụng bộ khử tiếng vọng có bộ bù độ trễ cho kết quả tốt hơn rất nhiều so với giải pháp không áp dụng bộ bù độ trễ.

Yêu cầu bảo hộ

1. Phương pháp khử tiếng vọng độ trễ lớn bao gồm ba bước, nội dung các bước cụ thể như sau:

bước 1: thu tiếng nói đầu xa; tiếng nói đầu xa là nguyên liệu đầu vào và là đối tượng tạo ra tiếng vọng; tiếng nói đầu xa được thu bởi ống thu thanh (microphone) của điện thoại dưới dạng tín hiệu điện, và lưu trữ trong hệ thống dưới dạng một chuỗi giá trị nhị phân;

bước 2: làm trễ tín hiệu tiếng nói đầu xa bằng bộ bù độ trễ; để giảm độ trễ giữa tín hiệu tiếng nói đầu xa và tín hiệu tiếng vọng đầu gần, cần thực hiện bằng việc làm trễ tín hiệu tiếng nói đầu xa trước khi thực hiện thuật toán bình phương trung bình tối thiểu - LMS; tại bước này, sử dụng một bộ bù độ trễ (jitter buffer) đặt phía trước bộ lọc thích nghi; với cấu trúc này, chỉ cần đo kiểm được độ trễ D giữa tín hiệu tiếng nói đầu xa và tín hiệu tiếng vọng đầu gần, chúng ta hoàn toàn có thể tính toán kích thước bộ bù độ trễ hợp lý để đưa độ trễ về khoảng giới hạn mà chúng ta mong muốn để bộ lọc $h(n)$ có thể xử lý hiệu quả được; theo đó, tiếng nói đầu xa được đóng gói thành các khung có chiều dài Tms , trước khi đưa qua bộ lọc thích nghi $h(n)$, được đưa qua bộ bù độ trễ; bộ bù độ trễ có tác dụng làm chậm lại luồng tín hiệu đầu xa $z(n)$, thành tín hiệu $z'(n)$; một khung dữ liệu tiếng nói đầu xa $z(n)$ (dài Tms) được truyền đi ở thời điểm t_0 , và tiếng vọng đầu xa $v(n)$ quay trở lại hệ thống ở thời điểm t_1 , độ trễ tiếng vọng được tính bằng:

$$D = t_1 - t_0$$

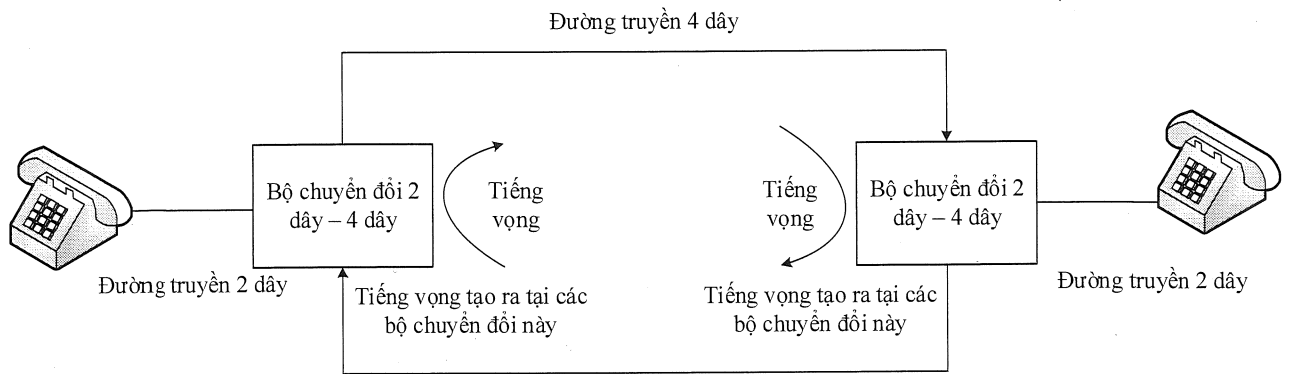
độ trễ này có thể đo kiểm được trong các hệ thống viễn thông thực tế;

để xử lý với bộ lọc thích nghi $h(n)$ có chiều dài L , thì kích thước của bộ bù độ trễ được tính toán theo công thức như sau:

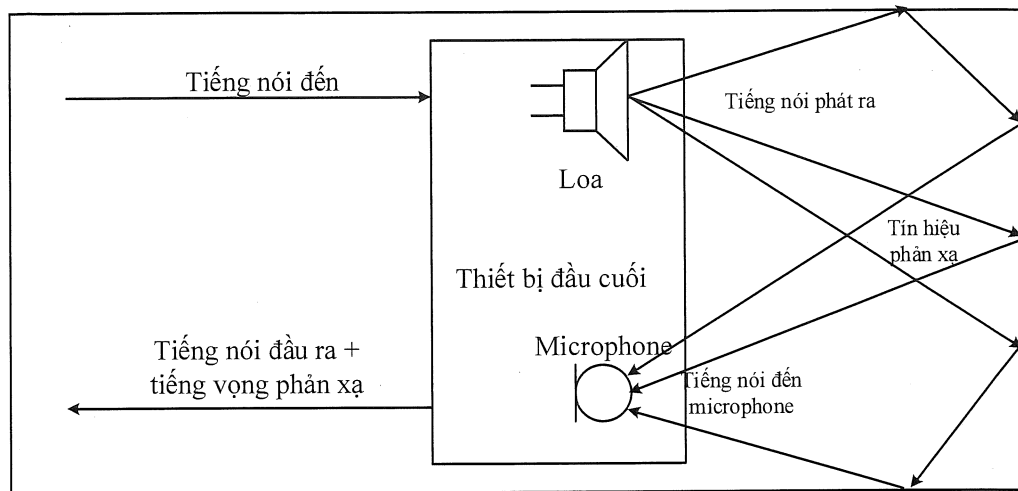
$$S = \frac{\left(\frac{D \times 8000}{10^3} \right) - L}{\left(\frac{T \times 8000}{10^3} \right)}$$

bước 3: khử tiếng vọng độ trễ lớn; sau khi được bù độ trễ, độ lệch giữa tín hiệu đưa vào bộ lọc thích nghi sử dụng thuật toán bình phương trung bình tối thiểu - LMS và tiếng vọng thực tế đã ở trong ngưỡng nhỏ mà ở đó bộ lọc thích nghi chiều dài L đã có thể loại bỏ được tiếng vọng.

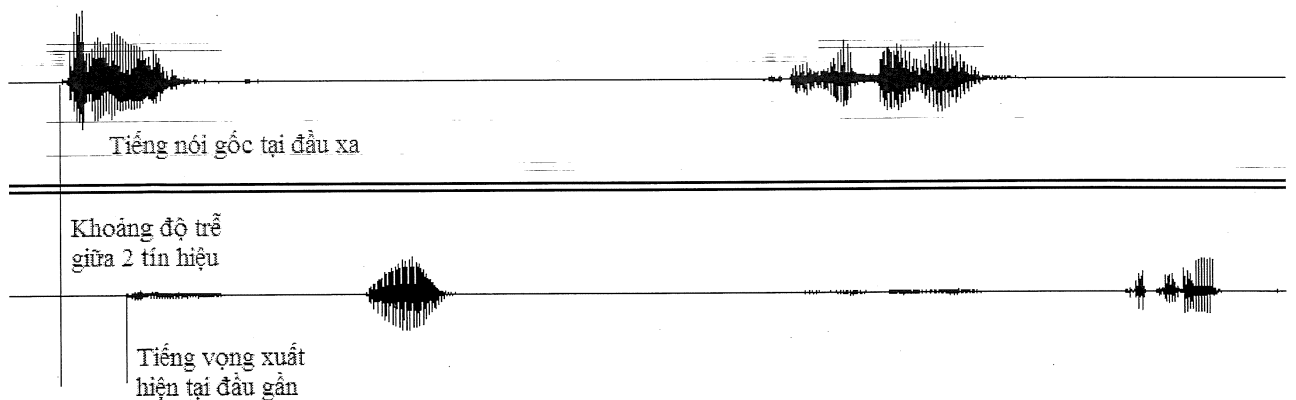
Hình vẽ



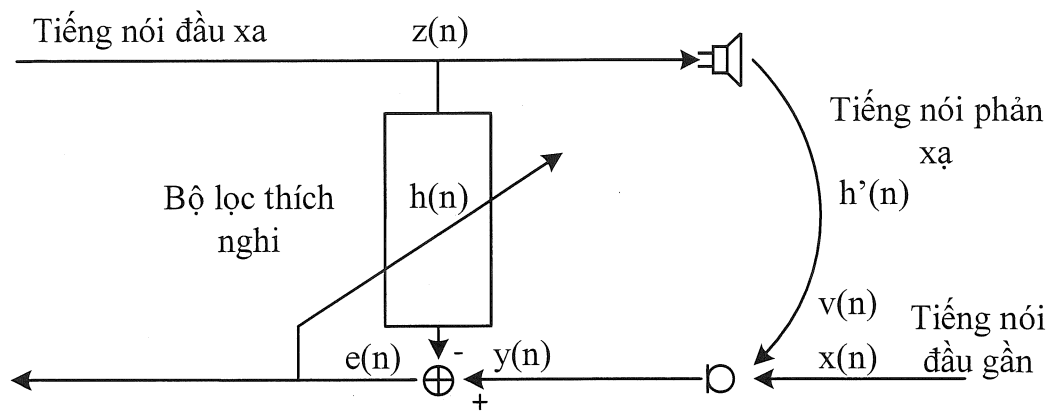
Hình 1



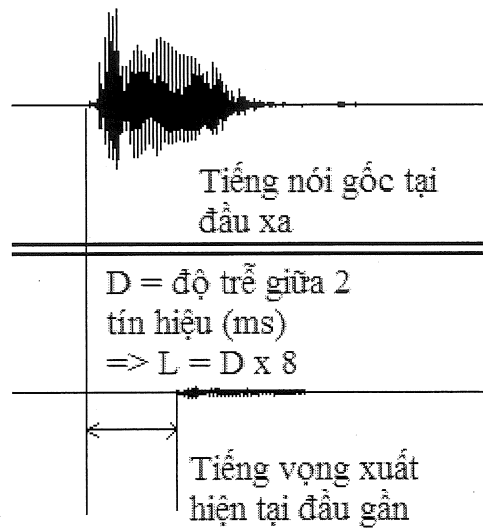
Hình 2



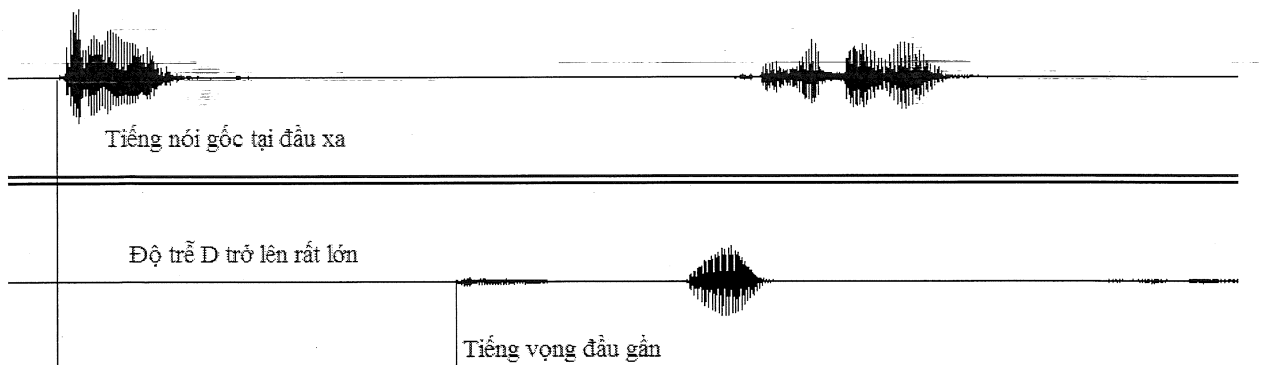
Hình 3



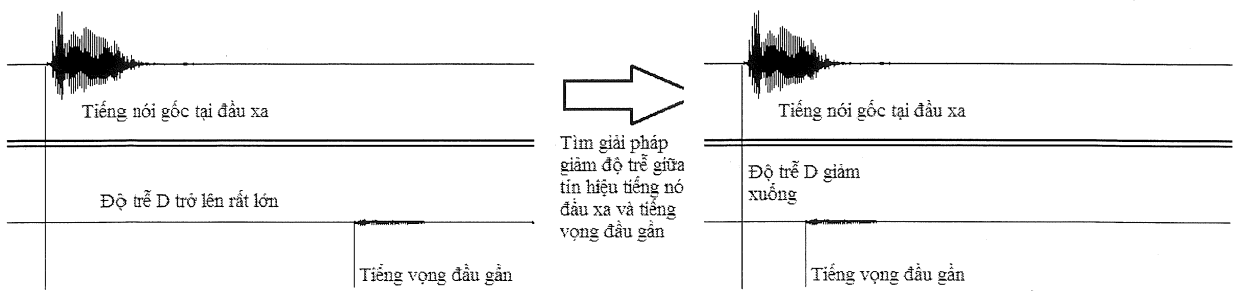
Hình 4



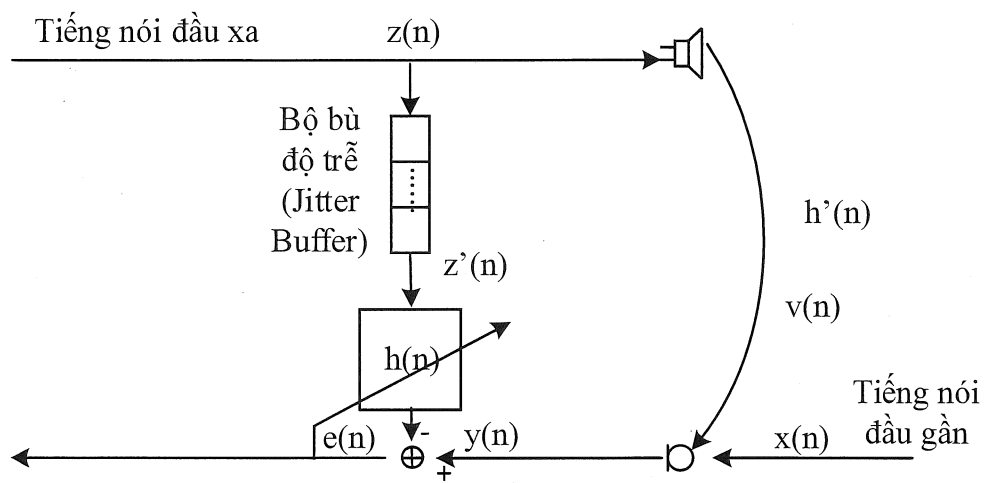
Hình 5



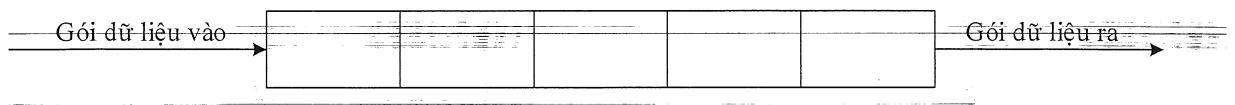
Hình 6



Hình 7

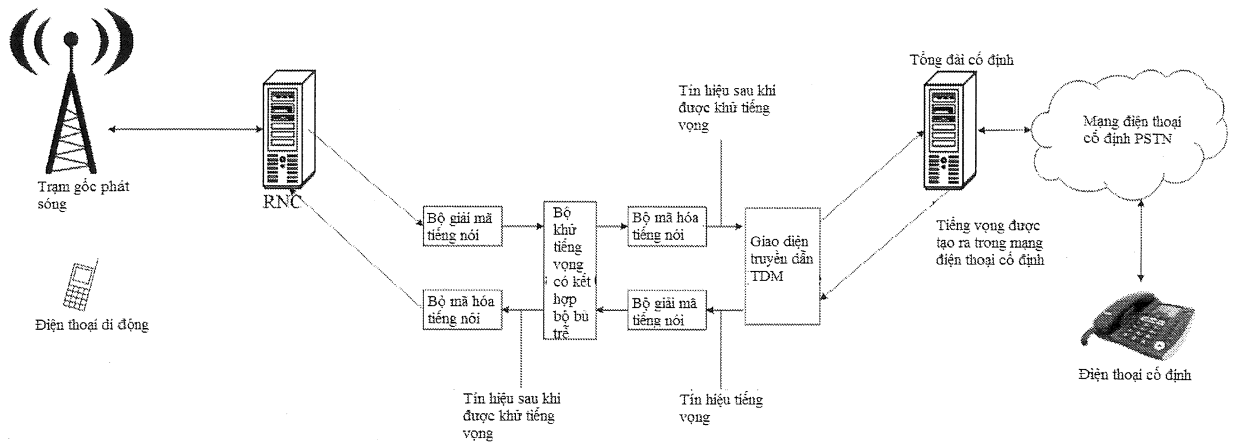


Hình 8

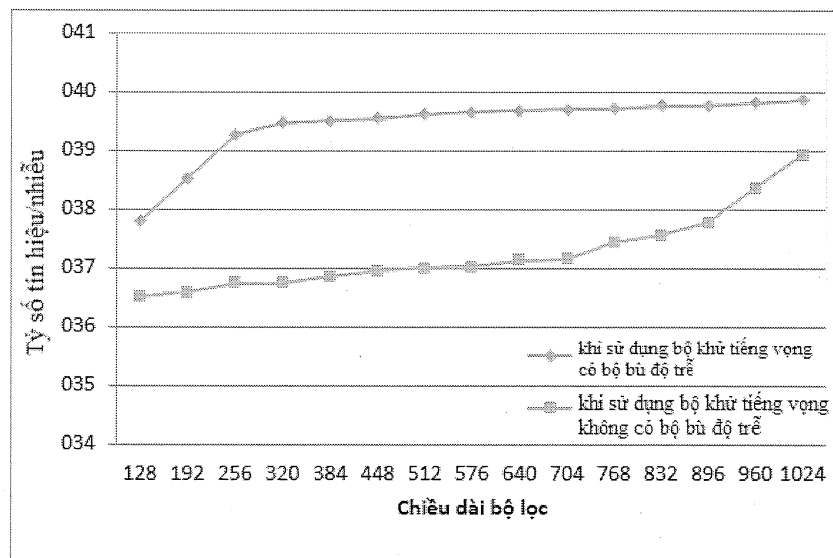


Các gói dữ liệu được đưa
vào bộ bù độ trễ (Jitter
Buffer) theo nguyên tắc
vào trước ra trước

Hình 9



Hình 10



Hình 11