



(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)



1-0025101

(51)⁷ G10L 19/032

(13) B

(21) 1-2016-01154

(22) 09/09/2014

(86) PCT/CN2014/086135 09/09/2014

(87) WO2015/035896 19/03/2015

(30) 61/875,690 10/09/2013 US; 14/478,839 05/09/2014 US

(45) 25/08/2020 389

(43) 27/06/2016 339A

(73) HUAWEI TECHNOLOGIES CO., LTD. (CN)

Huawei Administration Building, Bantian, Longgang, Shenzhen, Guangdong 518129, China

(72) GAO, Yang (US).

(74) Công ty Luật TNHH Phạm và Liên danh (PHAM & ASSOCIATES)

(54) PHƯƠNG PHÁP GIẢI MÃ DÒNG BIT ÂM THANH ĐƯỢC MÃ HÓA VÀ TẠO MỞ RỘNG BĂNG THÔNG TẦN SỐ Ở BỘ GIẢI MÃ, VÀ THIẾT BỊ GIẢI MÃ ĐỂ XỬ LÝ THOẠI

(57) Sáng chế đề xuất phương pháp giải mã dòng bit âm thanh được mã hóa và tạo mở rộng băng thông tần số gồm giải mã dòng bit âm thanh để tạo tín hiệu âm thanh băng thấp được giải mã và tạo phổ kích thích dải thấp tương ứng với băng tần số thấp. Khu vực băng phụ được lựa chọn từ trong băng tần số thấp bằng cách sử dụng tham số mà chỉ báo thông tin năng lượng của đường bao đặc biệt của tín hiệu âm thanh băng thấp được giải mã. Phổ kích thích băng cao được tạo cho băng tần số cao bằng cách sao chép phổ kích thích băng phụ từ khu vực băng phụ được lựa chọn đến khu vực băng phụ cao tương ứng với băng tần số cao. Bằng cách sử dụng phổ kích thích băng cao được tạo, tín hiệu âm thanh băng cao được mở rộng được tạo bằng cách áp dụng đường bao phổ băng cao. Tín hiệu âm thanh băng cao được mở rộng được thêm vào tín hiệu âm thanh băng thấp được giải mã để tạo tín hiệu đầu ra âm thanh có băng thông tần số mở rộng.



Lĩnh vực kỹ thuật được đề cập

Sáng chế chủ yếu liên quan đến lĩnh vực xử lý thoại, và cụ thể là đến việc mở rộng băng thông thích ứng và thiết bị để mở rộng.

Tình trạng kỹ thuật của sáng chế

Ở hệ thống truyền thông tín hiệu số âm thanh/thoại hiện đại, tín hiệu số được nén ở bộ mã hóa; thông tin được nén (bitstream – dòng bit) có thể được đóng gói và được gửi đến bộ giải mã thông qua kênh truyền thông theo từng khung. Hệ thống của bộ mã hóa và bộ giải mã đều được gọi là codec (encoder/decoder). Việc nén thoại/âm thanh có thể được sử dụng để giảm số lượng bit biểu thị tín hiệu thoại/âm thanh nhờ đó giảm tốc độ bit cần để truyền. Kỹ thuật nén thoại/âm thanh có thể được phân loại chung thành mã hóa miền thời gian và mã hóa miền tần số. Mã hóa miền thời gian thường được sử dụng để mã hóa tín hiệu thoại hoặc để mã hóa tín hiệu âm thanh ở các tốc độ bit thấp. Mã hóa miền tần số thường được sử dụng để mã hóa tín hiệu âm thanh hoặc để mã hóa tín hiệu thoại ở các tốc độ bit cao. BWE (Bandwidth Extension – mở rộng băng thông) có thể là một phần của mã hóa miền thời gian hoặc mã hóa miền tần số để tạo tín hiệu băng cao ở tốc độ bit rất thấp hoặc ở tốc độ bit 0.

Tuy nhiên, các bộ tạo mã thoại là các bộ tạo mã có tổn hao, tức là, tín hiệu được giải mã khác với ban đầu. Do vậy, một trong các mục đích trong việc tạo mã thoại là giảm thiểu méo (hoặc tổn hao thấy được) ở tốc độ bit cho trước, hoặc tối thiểu hóa tốc độ bit để đạt độ méo cho trước.

Việc tạo mã thoại khác với các dạng tạo mã âm thanh khác trong đó thoại là tín hiệu đơn giản hơn nhiều phần lớn các tín hiệu âm thanh khác, và thông tin thống kê nhiều hơn có sẵn về các đặc tính của thoại. Kết quả là, một số thông tin thính giác liên quan trong việc tạo mã âm thanh có thể là không cần thiết trong ngữ cảnh tạo mã thoại. Khi tạo mã thoại, tiêu chí quan trọng nhất là bảo toàn độ nghe rõ và “sự dễ chịu” của thoại, với lượng dữ liệu được truyền bị giới hạn.

Độ nghe rõ của thoại gồm, bên cạnh nội dung trực tiếp thực, còn danh tính ống nghe, cảm xúc, ngữ điệu, âm sắc v.v. đều rất quan trọng cho độ nghe rõ hoàn hảo. Khái niệm trừu tượng hơn của sự dễ chịu của thoại xuống cấp là thuộc tính khác ngoài độ nghe rõ, do có thể thoại xuống cấp hoàn toàn là nghe rõ, nhưng chủ quan gây phiền nhiễu cho người nghe.

Sự dư thừa của các dạng sóng thoại có thể được xem xét so với một số loại tín hiệu thoại khác, như các tín hiệu thoại có thanh âm và không có thanh âm. Các âm thanh có thanh âm, chẳng hạn, ‘a’, ‘b’, chủ yếu do các rung động của các dây thanh âm, và dao động. Do vậy, trên các chu kỳ thời gian ngắn, chúng được tạo hình tốt bằng các tổng của các tín hiệu tuần hoàn như các hình sin. Nói theo cách khác, đối với thoại có thanh âm, tín hiệu thoại thường tuần hoàn. Tuy nhiên, sự tuần hoàn này có thể biến thiên trong toàn bộ thời gian của đoạn thoại và dạng sóng tuần hoàn thường thay đổi dần từ đoạn này sang đoạn khác. Việc tạo mã thoại tốc độ bit thấp có thể hưởng lợi lớn từ việc khám phá tính tuần hoàn này. Chu kỳ thoại có thanh âm cũng được gọi là pitch, và việc dự báo pitch thường được gọi là LTP (Dự báo dài hạn – Dự báo dài hạn). Ngược lại, các âm không có thanh âm như ‘s’, ‘sh’, thường giống nhiễu hơn. Đây là do tín hiệu thoại không có thanh âm thường giống nhiễu ngẫu nhiên và có mức độ dự báo được nhỏ hơn.

Như đã biết, tất cả các phương pháp tạo mã thoại tham số như mã hóa miền thời gian tận dụng sự dư thừa vốn có trong tín hiệu thoại để giảm

lượng thông tin phải được gửi và ước tính các tham số các mẫu thoại của tín hiệu ở các khoảng ngắn. Sự dư thừa này chủ yếu phát sinh từ việc lặp các dạng sóng thoại ở tốc độ chuẩn tuần hoàn, và đường bao phổ thay đổi chậm của tín hiệu thoại.

Sự dư thừa của các dạng sóng thoại có thể được xem xét so với một số loại tín hiệu thoại khác nhau, như có thanh âm và không có thanh âm. Mặc dù tín hiệu thoại thường tuần hoàn đối với thoại có thanh âm, song tính tuần hoàn này có thể biến thiên trên khoảng thời gian của đoạn thoại và dạng sóng tuần hoàn thường thay đổi dần từ đoạn này sang đoạn khác. Việc tạo mã thoại tốc độ bit thấp có thể hưởng lợi lớn trong việc khám phá tính tuần hoàn này. Chu kỳ thoại có thanh âm còn được gọi là pitch, và việc dự báo pitch thường được gọi là LTP. Như đối với thoại không thanh âm, tín hiệu giống nhiều ngẫu nhiên hơn và có lượng dự báo được nhỏ hơn.

Trong trường hợp khác, việc tạo mã tham số có thể được sử dụng để giảm sự dư thừa của các đoạn thoại bằng cách tách riêng thành phần kích thích của tín hiệu thoại từ thành phần đường bao phổ. Đường bao phổ thay đổi chậm có thể được biểu thị bằng LPC (Linear Prediction Coding – Tạo mã dự báo tuyến tính) còn được gọi là STP (Dự báo ngắn hạn – Dự báo kỳ hạn ngắn). Việc tạo mã thoại tốc độ bit thấp cũng có thể hưởng lợi lớn từ việc khám phá STP này. Ưu điểm tạo mã phát sinh từ tốc độ thấp mà ở đó các tham số thay đổi. Nhưng, khá hiếm với các tham số mà khác đáng kể với các giá trị được giữ trong vài giây. Do đó, ở tốc độ lấy mẫu 8kHz, 12,8kHz hoặc 16kHz, thuật toán tạo mã thoại là sao cho khoảng thời gian khung danh định nằm trong khoảng từ mười đến ba mươi giây. Khoảng thời gian khung 20s là sự lựa chọn phổ thông nhất.

Việc tạo mã âm thanh dựa trên công nghệ dãy bộ lọc được sử dụng rộng rãi, chẳng hạn, trong việc mã hóa miền tần số. Khi xử lý tín hiệu, dãy bộ lọc là mảng bộ lọc thông băng mà tách riêng tín hiệu đầu vào

thành nhiều thành phần, mỗi thành phần mang một băng phụ tần số của tín hiệu ban đầu. Quá trình phân ra được thực hiện bởi dãy bộ lọc được gọi là *phân tích*, và đầu ra của việc phân tích dãy bộ lọc được gọi là tín hiệu băng phụ với nhiều băng phụ như các bộ lọc trong dãy bộ lọc. Quá trình tái tạo lại được gọi là *tổng hợp* dãy bộ lọc. Khi xử lý tín hiệu số, thuật ngữ dãy bộ lọc còn thường được áp dụng cho dãy bộ nhận. Sự khác biệt chính là các bộ nhận còn chuyển đổi xuống các dải phụ thành tần số trung tâm thấp mà có thể được lấy mẫu lại ở tốc độ giảm. Kết quả tương tự đôi khi có thể đạt được bằng cách lấy mẫu dưới các băng phụ thông dải. Đầu ra phân tích dãy bộ lọc có thể ở dạng các hệ số phức tạp. Mỗi hệ số phức tạp chứa *phần thực* và *phần ảo* lần lượt biểu thị *số hạng cosin* và *số hạng sin* cho mỗi băng phụ của dãy bộ lọc.

Ở các chuẩn đã biết gần đây hơn như G.723.1, G.729, G.718, EFP (Enhanced Full Rate – Tốc độ đủ tăng cường), SMV (Selectable Mode Vocoder – bộ mã hóa tiếng nói chế độ chọn được), AMR (Adaptive Multi-Rate – đa tốc độ thích ứng), VMR-WB (Variable-Rate Multimode Wideband – Băng rộng đa chế độ tốc độ biến thiên), hoặc AMR-WB (Adaptive Multi-Rate Wideband – Băng rộng đa tốc độ thích ứng), CELP (Code Excited Linear Prediction Technique – kỹ thuật dự báo tuyến tính được kích thích bằng mã) đã được sử dụng. CELP thường được hiểu là kết hợp kỹ thuật của kích thích được mã hóa (Kích thích được tạo mã – CE), LTP và STP. CELP chủ yếu được sử dụng để mã hóa tín hiệu thoại bằng cách hưởng lợi từ các đặc tính giọng người cụ thể hoặc mô hình tạo giọng nói thanh âm người. Việc tạo mã thoại CELP là nguyên lý thuật toán rất phổ biến trong khu vực nén thoại mặc dù các chi tiết về CELP cho các codec khác nhau có thể khác nhau đáng kể. Do tính phổ biến của nó, thuật toán CELP đã được sử dụng trong các chuẩn ITU-T, MPEG, 3GPP, và 3GPP2 khác nhau. Các biến thức của CELP gồm CELP đại số, CELP tích thoát, CELP độ trễ thấp và dự báo tuyến tính được kích thích

bằng tổng vector, và các tham số khác. CELP là thuật ngữ chung cho lớp thuật toán và không cho codec cụ thể.

Thuật toán CELP được dựa trên bốn ý tưởng chính. Trước hết, mô hình bộ lọc nguồn sản xuất thoại thông qua LP được sử dụng. Mô hình bộ lọc nguồn của tạo thoại sẽ tạo mô hình thoại như là kết hợp nguồn âm thanh, như các dây thanh âm, và bộ lọc âm thanh tuyến tính, bó thanh âm (và đặc tính bức xạ). Khi triển khai mô hình bộ lọc nguồn tạo thoại, nguồn âm thanh, hoặc tín hiệu kích thích, thường được tạo hình như là chuỗi xung tuần hoàn, cho thoại có thanh âm, hoặc nhiễu trắng cho thoại không có thanh âm. Thứ hai, từ điển mã cố định và thích ứng được sử dụng làm đầu vào (kích thích) của mô hình LP. Thứ ba, tìm kiếm được thực hiện trong vòng kín trong “miền có trọng số giác quan”. Thứ tư, VQ (vector quantization – lượng tử hóa vector) được áp dụng.

Bản chất kỹ thuật của sáng chế

Phương án thực hiện sáng chế mô tả phương pháp giải mã dòng bit âm thanh được mã hóa và tạo mở rộng băng thông tần số ở bộ giải mã. Phương pháp bao gồm giải mã dòng bit âm thanh để tạo tín hiệu âm thanh băng thấp được giải mã và tạo phổ kích thích dải thấp tương ứng với băng tần số thấp. Khu vực băng phụ được lựa chọn từ trong băng tần số thấp bằng cách sử dụng tham số mà chỉ báo thông tin năng lượng của đường bao đặc biệt của tín hiệu âm thanh băng thấp được giải mã. Phổ kích thích băng cao được tạo cho băng tần số cao bằng cách sao chép phổ kích thích băng phụ từ khu vực băng phụ được lựa chọn đến khu vực băng phụ cao tương ứng với băng tần số cao. Bằng cách sử dụng phổ kích thích băng cao được tạo, tín hiệu âm thanh băng cao được mở rộng được tạo bằng cách áp dụng đường bao phổ băng cao. Tín hiệu âm thanh băng cao được mở rộng được thêm vào tín hiệu âm thanh băng thấp được giải mã để tạo tín hiệu đầu ra âm thanh có băng thông tần số mở rộng.

Theo phương án thực hiện khác của sáng chế, bộ giải mã để giải mã dòng bit âm thanh được mã hóa và tạo băng thông tần số bao gồm khối giải mã băng thấp được cấu hình để giải mã dòng bit âm thanh để tạo tín hiệu âm thanh băng thấp được giải mã và để tạo phổ kích thích băng thấp tương ứng với băng tần số thấp. Bộ giải mã còn gồm khối mở rộng băng thông được ghép nối với khối giải mã băng thấp. Khối mở rộng băng thông bao gồm khối lựa chọn băng phụ và khối sao chép. Khối lựa chọn băng phụ được cấu hình để lựa chọn khu vực băng phụ từ trong băng tần số thấp bằng cách sử dụng tham số mà chỉ báo thông tin năng lượng của đường bao đặc biệt của tín hiệu âm thanh băng thấp được giải mã. Khối sao chép được cấu hình để tạo phổ kích thích băng cao cho băng tần số cao bằng cách sao chép phổ kích thích băng phụ từ khu vực băng phụ được lựa chọn đến khu vực băng phụ cao tương ứng với băng tần số cao.

Theo phương án thực hiện khác của sáng chế, bộ giải mã để xử lý thoại bao gồm bộ xử lý và vật lưu trữ máy tính đọc được lưu trữ chương trình để thực thi bởi bộ xử lý. Chương trình gồm các lệnh để giải mã dòng bit âm thanh để tạo tín hiệu âm thanh băng thấp được giải mã và tạo phổ kích thích dải thấp tương ứng với băng tần số thấp. Chương trình gồm các lệnh để lựa chọn khu vực băng phụ từ trong băng tần số thấp bằng cách sử dụng tham số mà chỉ báo thông tin năng lượng của đường bao đặc biệt của tín hiệu âm thanh băng thấp được giải mã, và tạo phổ kích thích băng cao cho băng tần số cao bằng cách sao chép phổ kích thích băng phụ từ khu vực băng phụ được lựa chọn đến khu vực băng phụ cao tương ứng với băng tần số cao. Chương trình còn gồm các lệnh để sử dụng phổ kích thích băng cao được tạo để tạo tín hiệu âm thanh băng cao được mở rộng bằng cách áp dụng đường bao phổ băng cao, và thêm tín hiệu âm thanh băng cao được mở rộng vào tín hiệu âm thanh băng thấp được giải mã để tạo tín hiệu đầu ra âm thanh có băng thông tần số mở rộng.

Phương án thực hiện khác của sáng chế mô tả phương pháp giải mã dòng bit âm thanh được mã hóa và tạo mở rộng băng thông tần số ở bộ giải mã. Phương pháp bao gồm giải mã dòng bit âm thanh để tạo tín hiệu âm thanh băng thấp được giải mã và tạo phổ băng thấp tương ứng với băng tần số thấp và lựa chọn khu vực băng phụ từ trong băng tần số thấp bằng cách sử dụng tham số mà chỉ báo thông tin năng lượng của đường bao đặc biệt của tín hiệu âm thanh băng thấp được giải mã. Phương pháp còn gồm tạo phổ băng cao bằng cách sao chép phổ băng phụ từ khu vực băng phụ được lựa chọn đến khu vực băng phụ cao, và bằng cách sử dụng phổ băng cao được tạo để tạo tín hiệu âm thanh băng cao được mở rộng bằng cách áp dụng năng lượng đường bao phổ băng cao. Phương pháp còn gồm thêm tín hiệu âm thanh băng cao được mở rộng vào tín hiệu âm thanh băng thấp được giải mã để tạo tín hiệu đầu ra âm thanh có băng thông tần số mở rộng.

Mô tả vắn tắt các hình vẽ

Để hiểu toàn diện hơn về sáng chế, và các ưu điểm của nó, phần viện dẫn được thực hiện đến các phần mô tả dưới đây có dựa vào các hình vẽ đi kèm, trong đó:

Fig.1 minh họa các hoạt động được thực hiện trong quá trình mã hóa thoại gốc bằng cách sử dụng bộ mã hóa CELP đã biết;

Fig.2 minh họa các hoạt động được thực hiện trong quá trình giải mã của thoại gốc bằng cách sử dụng bộ giải mã CELP khi triển khai các phương án thực hiện sáng chế như sẽ được mô tả thêm dưới đây;

Fig.3 minh họa các hoạt động được thực hiện trong quá trình mã hóa thoại gốc trong bộ mã hóa CELP đã biết;

Fig.4 minh họa bộ giải mã CELP cơ bản tương ứng với bộ mã hóa trên Fig.5 khi triển khai các phương án thực hiện sáng chế như sẽ được mô tả dưới đây;

Fig.5A và Fig.5B minh họa ví dụ mã hóa/giải mã với BWE (Band Width Extension – mở rộng băng thông), trong đó Fig.5A minh họa các hoạt động ở bộ mã hóa với thông tin phía BWE trong khi Fig.5B minh họa các hoạt động ở bộ giải mã với BWE;

Fig.6A và Fig.6B minh họa ví dụ khác về mã hóa/giải mã với BWE mà không truyền thông tin phụ, trong đó Fig.6A minh họa các hoạt động trong quá trình ở bộ mã hóa trong khi Fig.6B minh họa các hoạt động ở bộ giải mã;

Fig.7 minh họa ví dụ về phổ kích thích lý tưởng cho thoại có giọng nói hoặc âm nhạc điều hòa khi loại CELP của codec được sử dụng;

Fig.8 thể hiện ví dụ về BWE đã biết của phổ kích thích được giải mã cho thoại có giọng nói hoặc âm nhạc điều hòa khi loại CELP của codec được sử dụng;

Fig.9 minh họa ví dụ về BWE được áp dụng cho phổ kích thích được giải mã cho thoại có giọng nói hoặc âm nhạc điều hòa khi loại CELP của codec được sử dụng theo phương án thực hiện sáng chế;

Fig.10 minh họa các hoạt động ở bộ giải mã để triển khai dịch chuyển băng phụ hoặc sao chép cho BWE theo các phương án thực hiện sáng chế;

Fig.11 minh họa bộ giải mã để triển khai dịch chuyển băng phụ hoặc sao chép cho BWE theo phương án thực hiện khác;

Fig.12 minh họa các hoạt động được thực hiện ở bộ giải mã theo các phương án thực hiện sáng chế;

Fig.13A và Fig.13B minh họa bộ giải mã triển khai BWE theo các phương án thực hiện sáng chế;

Fig.14 minh họa hệ thống truyền thông theo phương án thực hiện sáng chế; và

Fig.15 minh họa sơ đồ khối của hệ thống xử lý có thể được sử dụng để triển khai các thiết bị và các phương pháp được bộc lộ ở đây.

Mô tả chi tiết các phương án thực hiện sáng chế

Trong hệ thống truyền thông tín hiệu số âm thanh/thoại hiện đại, tín hiệu số được nén ở bộ mã hóa, và thông tin được nén hoặc dòng bit có thể được đóng gói và được gửi đến bộ giải mã theo từng khung thông qua kênh truyền thông. Bộ giải mã tiếp nhận và giải mã thông tin được nén để thu được tín hiệu số âm thanh/thoại.

Sáng chế chủ yếu liên quan đến tạo mã tín hiệu thoại/âm thanh và mở rộng băng thông tín hiệu thoại/âm thanh. Cụ thể là, các phương án thực hiện sáng chế có thể được sử dụng để cải thiện chuẩn của bộ tạo mã thoại ITU-T AMR-WB trong lĩnh vực BWE.

Một số tần số quan trọng hơn các tần số khác. Các tần số quan trọng có thể được tạo mã bằng độ phân giải mịn. Các tần số nhỏ ở các tần số này là quan trọng và cần phương tiện tạo mã bảo tồn các khác biệt này. Nói theo cách khác, các tần số ít quan trọng hơn không cần phải chính xác. Phương tiện tạo mã thô hơn có thể được sử dụng, mặc dù một số chi tiết mịn hơn sẽ mất trong khi tạo mã. Phương tiện tạo mã thô hơn đặc trưng dựa trên khái niệm BWE. Khái niệm công nghệ này còn được gọi là HBE (High Band Extension – mở rộng băng cao), SBR (SubBand Replica – Bản sao băng phụ) hoặc SBR (Spectral Band Replication – Sao chép băng phổ). Mặc dù tên có thể khác nhau, song chúng đều có chung nghĩa mã hóa/giải mã một số băng phụ tần số (thường là băng cao) với chi phí tốc độ bit thấp (thậm chí chi phí tốc độ bit bằng 0) hoặc tốc độ bit thấp hơn đáng kể so với cách thức mã hóa/giải mã thông thường.

Ở công nghệ SBR, cấu trúc mịn phổ trong băng tần số cao được sao chép từ băng tần số thấp và một số nhiễu ngẫu nhiên có thể được thêm vào. Sau đó, đường bao phổ trong băng tần số cao được tạo dạng bằng cách sử dụng thông tin phụ được truyền từ bộ mã hóa đến bộ giải mã. Việc dịch chuyển băng tần số hoặc việc sao chép từ băng thấp đến băng cao thường là bước thứ nhất cho công nghệ BWE.

Các phương án thực hiện sáng chế sẽ được mô tả để cải thiện công nghệ BWE bằng cách sử dụng quá trình thích ứng để lựa chọn băng dịch chuyển dựa vào mức năng lượng của đường bao phổ.

Fig.1 minh họa các hoạt động được thực hiện trong quá trình mã hóa thoại gốc bằng cách sử dụng bộ mã hóa CELP đã biết.

Fig.1 minh họa bộ mã hóa CELP ban đầu đã biết trong đó lỗi có trọng số 109 giữa thoại tổng hợp 102 và thoại gốc 101 được tối thiểu thường bằng cách sử dụng cách phân tích bằng tổng hợp, nghĩa là việc mã hóa (phân tích) được thực hiện bằng cách tối ưu hóa giác quan tín hiệu được giải mã (tổng hợp) trong vòng kín.

Nguyên lý cơ bản chính là tất cả các bộ tạo mã thoại khai thác là thực tế rằng các tín hiệu thoại là các dạng sóng tương quan cao. Như minh họa, thoại có thể được biểu diễn bằng cách sử dụng mô hình tự hồi quy (autoregressive - AR) như trong phương trình (11) dưới đây.

$$X_n = \sum_{i=1}^L a_i X_{n-i} + e_n \quad (11)$$

Trong phương trình (11), mỗi mẫu được biểu thị dưới dạng kết hợp tuyến tính của các mẫu L trước cộng nhiễu trắng. Các hệ số trọng số $a_1, a_2, \dots, a_L$, được gọi là các LPC. Đối với mỗi khung, các hệ số trọng số $a_1, a_2, \dots, a_L$, được lựa chọn sao cho phổ $\{X_1, X_2, \dots, X_N\}$, được tạo bằng cách sử dụng mô hình nêu trên, gần như so khớp phổ của khung thoại đầu vào.

Theo cách khác, các tín hiệu thoại cũng có thể được biểu thị bằng cách kết hợp mô hình điều hòa và mô hình nhiễu. Phần điều hòa của mô hình là biểu diễn chuỗi Fourier của phần tuần hoàn của tín hiệu. Nói chung, đối với các tín hiệu có thanh âm, mô hình nhiễu cộng điều hòa của thoại gồm kết hợp của cả điều hòa và nhiễu. Phần điều hòa và nhiễu trong thoại có thanh âm phụ thuộc vào nhiều yếu tố gồm đặc tính ống nghe (chẳng hạn, ở mức độ nào đó giọng của ống nghe là bình thường hoặc nghe rõ cả tiếng thở); đặc trưng đoạn thoại (chẳng hạn, ở mức độ nào đó đoạn thoại

là tuần hoàn) và trên tần số. Các tần số cao hơn của thoại có thanh âm có tỷ lệ cao hơn các phần giống nhau.

Mô hình LP và mô hình nhiễu điều hòa là hai phương pháp chính để tạo hình và tạo mã các tín hiệu thoại. Mô hình LP đặc biệt tốt ở việc tạo hình đường bao phổ của thoại trong khi mô hình nhiễu điều hòa tốt ở việc tạo hình cấu trúc mịn của thoại. Hai phương pháp này có thể được kết hợp để tận dụng các điểm mạnh tương đối của chúng.

Như được nêu trước đó, trước khi tạo mã CELP, tín hiệu đầu vào micrô của tai nghe được lọc và lấy mẫu, chẳng hạn, ở tốc độ 8000 mẫu trên giây. Sau đó mỗi mẫu được lượng tử hóa, chẳng hạn, với 13 bit trên mẫu. Thoại được lấy mẫu được phân mảnh thành các đoạn hoặc khung 20ms (chẳng hạn, trong trường hợp này 160 mẫu).

Tín hiệu thoại được phân tích và mô hình LP của nó, các tín hiệu kích thích và pitch được trích rút. Mô hình LP biểu thị đường bao phổ của thoại. Nó được biến đổi sang nhóm các hệ số tần số phổ đường thẳng (line spectral frequencies - LSF), vốn là biểu diễn khác của các tham số LP, do các hệ số LSF có các đặc tính lượng tử hóa tốt. Các hệ số LSF có thể được lượng tử vô hướng hoặc hiệu quả hơn chúng có thể được lượng tử hóa vector bằng cách sử dụng các từ điển mã vector LSF được đào tạo trước đó.

Việc kích thích mã gồm từ điển mã bao gồm các vector mã, có các thành phần đều được lựa chọn độc lập sao cho mỗi vector mã có phổ xấp xỉ 'trắng'. Đối với mỗi khung phụ của thoại đầu vào, mỗi trong số các vector mã được lọc thông qua bộ lọc dự báo tuyến tính ngắn hạn 103 và bộ lọc dự báo dài hạn 105, và đầu ra được so sánh với các mẫu thoại. Ở mỗi khung phụ, vector mã có đầu ra so khớp tốt nhất thoại đầu vào (lỗi được tối thiểu) được chọn để biểu thị khung phụ đó.

Kích thích được tạo mã 108 thông thường bao gồm tín hiệu giống xung hoặc tín hiệu giống nhiễu, được tạo bằng toán học hoặc được lưu trữ

trong từ điển mã. Từ điển mã khả dụng cho cả bộ mã hóa lẫn bộ giải mã tiếp nhận. Kích thích được tạo mã 108, có thể là từ điển mã cố định hoặc ngẫu nhiên, có thể là từ điển lượng tử vector vốn được mã hóa cứng (ngầm hoặc tường minh) vào codec. Từ điển mã cố định này có thể là LP được kích thích bằng mã đại số hoặc được lưu trữ tường minh.

Vector mã từ từ điển mã được tăng bằng độ khuếch đại thích hợp để khiến năng lượng này bằng năng lượng của thoại đầu vào. Do đó, đầu ra của kích thích được tạo mã 108 được tăng bằng độ khuếch đại G_c 107 trước khi đi qua các bộ lọc tuyến tính.

Bộ lọc dự báo tuyến tính ngắn hạn 103 tạo dạng phổ ‘trắng’ của vector mã để giống phổ của thoại đầu vào. Một cách tương đương, ở miền thời gian, bộ lọc dự báo tuyến tính ngắn hạn 103 đưa vào các tương quan ngắn hạn (tương quan với các mẫu trước) trong chuỗi trắng. Bộ lọc mà tạo dạng kích thích có mô hình toàn cực của dạng $1/A(z)$ (bộ lọc dự báo tuyến tính ngắn hạn 103), trong đó $A(z)$ được gọi là bộ lọc dự báo và có thể thu được bằng cách sử dụng LP (chẳng hạn, thuật toán Levinson–Durbin). Theo một hoặc nhiều phương án thực hiện, bộ lọc toàn cực có thể được sử dụng do nó là biến đổi tốt của bó thanh âm của con người và do nó dễ tính toán.

Bộ lọc dự báo tuyến tính ngắn hạn 103 thu được bằng cách phân tích tín hiệu gốc 101 và được biểu diễn bằng nhóm hệ số:

$$A(z) = \sum_{i=1}^P 1 + a_i \cdot z^{-i}, i=1,2,\dots,P \quad (12)$$

Như được mô tả trước đó, các vùng thoại có thanh âm biểu diễn tính chu kỳ dài hạn. Chu kỳ này, được biết là pitch, được đưa vào phổ tổng hợp bằng bộ lọc pitch $1/(B(z))$. Đầu ra của bộ lọc dự báo dài hạn 105 phụ thuộc vào pitch và độ khuếch đại pitch. Theo một hoặc nhiều phương án thực hiện, pitch có thể được ước lượng từ tín hiệu gốc, tín hiệu dư, hoặc tín hiệu gốc có trọng số. Theo một phương án thực hiện, hàm dự báo dài

hạn ($B(z)$) có thể được biểu diễn bằng cách sử dụng phương trình (13) như sau.

$$B(z) = 1 - G_p \cdot z^{-Pitch} \quad (13)$$

Bộ lọc trọng số 110 được liên quan đến bộ lọc dự báo ngắn hạn nêu trên. Một trong các bộ lọc trọng số đặc trưng có thể được biểu diễn như được mô tả ở phương trình (14).

$$W(z) = \frac{A(z/\alpha)}{1 - \beta \cdot z^{-1}} \quad (14)$$

trong đó $\beta < \alpha$, $0 < \beta < 1$, $0 < \alpha \leq 1$.

Theo phương án thực hiện khác, bộ lọc trọng số $W(z)$ có thể được suy ra từ bộ lọc LPC bằng cách sử dụng BWE như được minh họa theo một phương án thực hiện trong phương trình (15) dưới đây.

$$W(z) = \frac{A(z/\gamma_1)}{A(z/\gamma_2)} \quad (15),$$

Ở phương trình (15), $\gamma_1 > \gamma_2$, vốn là các hệ số mà các cực của nó được di chuyển về gốc.

Do đó, đối với mỗi khung thoại, các LPC và pitch được tính toán và các bộ lọc được cập nhật. Đối với mỗi khung phụ thoại, vectơ mã tạo kết quả được lọc 'tốt nhất' được chọn để biểu diễn khung phụ. Giá trị độ khuếch đại được lượng tử hóa tương ứng phải được truyền đến bộ giải mã để giải mã đúng. Các LPC và các giá trị pitch cũng phải được lượng tử hóa và gửi từng khung để tái tạo các bộ lọc ở bộ giải mã. Do đó, chỉ số kích thích được tạo mã, chỉ số khuếch đại được lượng tử hóa, chỉ số tham số dự báo dài hạn được lượng tử hóa, và chỉ số tham số dự báo ngắn hạn được lượng tử hóa được truyền đến bộ giải mã.

Fig.2 minh họa các hoạt động được thực hiện trong quá trình giải mã thoại gốc bằng cách sử dụng bộ giải mã CELP khi triển khai các phương án thực hiện sáng chế như sẽ được mô tả dưới đây.

Tín hiệu thoại được tái tạo ở bộ giải mã bằng cách đưa các vector mã nhận được qua các bộ lọc tương ứng. Kết quả là, mỗi khối ngoại từ hậu xử lý có cùng định nghĩa như được mô tả trong bộ mã hóa trên Fig.1.

Dòng bit CELP được mã hóa được tiếp nhận và tháo gói 80 ở thiết bị nhận. Đối với mỗi khung phụ được nhận, chỉ số kích thích được tạo mã được nhận, chỉ số khuếch đại được lượng tử hóa, chỉ số tham số dự báo dài hạn được lượng tử hóa, và chỉ số tham số dự báo ngắn hạn được lượng tử hóa, được sử dụng để tìm các tham số tương ứng bằng cách sử dụng các bộ giải mã tương ứng, chẳng hạn, bộ giải mã độ khuếch đại 81, bộ giải mã dự báo dài hạn 82, và bộ giải mã dự báo ngắn hạn 83. Chẳng hạn, các vị trí và các dấu biên độ của các xung kích thích và vector mã đại số của kích thích mã 402 có thể được xác định từ chỉ số kích thích được tạo mã được nhận.

Như được thể hiện trên Fig.2, bộ giải mã là tổ hợp của vài khối gồm kích thích được tạo mã 201, dự báo dài hạn 203, dự báo ngắn hạn 205. Bộ giải mã ban đầu còn gồm khối hậu xử lý 207 sau thoại tổng hợp 206. Bước hậu xử lý có thể còn gồm hậu xử lý ngắn hạn và hậu xử lý dài hạn.

Fig.3 minh họa bộ mã hóa CELP đã biết.

Fig.3 minh họa bộ mã hóa CELP cơ bản bằng cách sử dụng từ điển mã thích ứng bổ sung để cải thiện dự báo tuyến tính dài hạn. Kích thích được tạo bằng cách tính tổng các phần góp từ từ điển mã thích ứng 307 và kích thích mã 308, có thể là từ điển mã cố định hoặc ngẫu nhiên như được mô tả trước đó. Các mục trong từ điển mã thích ứng bao gồm các phiên bản kích thích trễ. Điều này có thể mã hóa hiệu quả các tín hiệu tuần hoàn như các âm thanh có thanh âm.

Như được thể hiện trên Fig.3, từ điển mã thích ứng 307 bao gồm kích thích tổng hợp quá khứ 304 hoặc lặp lại chu kỳ pitch kích thích quá khứ ở quãng pitch. Độ trễ pitch có thể được mã hóa ở giá trị nguyên khi nó lớn hoặc dài. Độ trễ pitch thường được mã hóa ở giá trị phân số chính xác

hơn khi nó nhỏ hoặc ngắn. Thông tin tuần hoàn pitch được sử dụng để tạo phần thích ứng của kích thích. Sau đó, phần kích thích này được tăng lên bởi độ khuếch đại G_p 305 (còn được gọi là độ khuếch đại pitch).

Dự báo dài hạn đóng vai trò quan trọng để tạo mã thoại có thanh âm do thoại có thanh âm có tính tuần hoàn mạnh. Các chu kỳ pitch lân cận của thoại có thanh âm giống nhau, nghĩa là về mặt toán học độ khuếch đại pitch G_p trong phần biểu diễn kích thích dưới đây cao hoặc gần 1. Kích thích thu được có thể được biểu diễn dưới dạng như trong phương trình (16) dưới dạng tổ hợp các kích thích thành phần.

$$e(n) = G_p \cdot e_p(n) + G_c \cdot e_c(n) \quad (16)$$

trong đó, $e_p(n)$ là một khung phụ của các chuỗi mẫu được đánh chỉ số n , đến từ từ điển mã thích ứng 307 bao gồm kích thích quá khứ 304 thông qua vòng phản hồi (Fig.3). $e_p(n)$ có thể được lọc thông thấp thích ứng do khu vực tần số thấp thường tuần hoàn hơn hoặc điều hòa hơn khu vực tần số cao. $e_c(n)$ từ kích thích được tạo mã từ từ điển mã 308 (còn được gọi là từ điển mã cố định) vốn là phần góp kích thích hiện tại. Ngoài ra, $e_c(n)$ cũng có thể được tăng cường như bằng cách sử dụng tăng cường lọc thông cao, tăng cường pitch, tăng cường phân tán, tăng cường focman, và các tăng cường khác.

Đối với thoại có thanh âm, phần góp $e_p(n)$ từ từ điển mã thích ứng 307 có thể chi phối và độ khuếch đại pitch G_p 305 xấp xỉ giá trị bằng 1. Kích thích thường được cập nhật cho mỗi khung phụ. Kích thước khung đặc trưng là 20 giây và kích thước khung phụ đặc trưng là 5 giây.

Như được mô tả trên Fig.1, kích thích được tạo mã cố định 308 được tăng lên bởi độ khuếch đại G_c 306 trước khi đi qua các bộ lọc tuyến tính. Hai phần kích thích được tăng từ kích thích được tạo mã cố định 108 và từ từ điển mã thích ứng 307 cùng được thêm vào trước khi lọc qua bộ lọc dự báo tuyến tính ngắn hạn 303. Hai độ khuếch đại (G_p và G_c) được lượng tử hóa và được truyền đến bộ giải mã. Do đó, chỉ số kích thích được tạo mã,

chỉ số từ điển mã thích ứng, các chỉ số độ khuếch đại được lượng tử hóa, và chỉ số tham số dự báo ngắn hạn được lượng tử hóa được truyền đến thiết bị âm thanh tiếp nhận.

Dòng bit CELP được mã hóa bằng cách sử dụng thiết bị được minh họa trên Fig.3 được tiếp nhận ở thiết bị nhận. Fig.4 minh họa bộ giải mã tương ứng của thiết bị nhận.

Fig.4 minh họa bộ giải mã CELP cơ bản tương ứng với bộ mã hóa trên Fig.5. Fig.4 gồm khối hậu xử lý 408 tiếp nhận thoại tổng hợp 407 từ bộ giải mã chính. Bộ giải mã này tương tự với Fig.3 ngoại trừ từ điển mã thích ứng 307.

Đối với mỗi khung phụ được tiếp nhận, chỉ số kích thích được tạo mã được nhận, chỉ số khuếch đại kích thích được tạo mã lượng tử hóa, chỉ số pitch lượng tử hóa, chỉ số khuếch đại từ điển mã thích ứng lượng tử hóa, và chỉ số tham số dự báo ngắn hạn lượng tử hóa, được sử dụng để tìm các tham số tương ứng bằng cách sử dụng các bộ giải mã tương ứng, chẳng hạn, bộ giải mã khuếch đại 81, bộ giải mã pitch 84, bộ giải mã khuếch đại từ điển mã thích ứng 85, và bộ giải mã dự báo ngắn hạn 83.

Theo phương án thực hiện khác, bộ giải mã CELP là tổ hợp của vài khối và bao gồm kích thích được tạo mã 402, từ điển mã thích ứng 401, dự báo ngắn hạn 406, và hậu xử lý 408. Mỗi khối ngoại trừ hậu xử lý có cùng định nghĩa như được mô tả trên bộ mã hóa trên Fig.3. Hậu xử lý có thể còn gồm hậu xử lý ngắn hạn và hậu xử lý dài hạn.

Như nêu trên, CELP chủ yếu được sử dụng để mã hóa tín hiệu thoại bằng cách hưởng lợi từ các đặc tính thoại riêng của người hoặc mô hình sản xuất giọng nói thanh âm của người. Để mã hóa tín hiệu thoại hiệu quả hơn, tín hiệu thoại có thể được phân loại thành các lớp khác nhau và mỗi lớp được mã hóa theo cách khác nhau. Phân loại có thanh âm/không có thanh âm hoặc quyết định không có thanh âm có thể là phân loại cơ bản và quan trọng trong số tất cả các phân loại của các lớp khác nhau. Đối với

mỗi lớp, bộ lọc LPC hoặc STP thường được sử dụng để biểu diễn đường bao phổ. Nhưng kích thích đến bộ lọc LPC có thể là khác nhau. Các tín hiệu không có thanh âm có thể được tạo mã bằng kích thích giống nhau. Nói theo cách khác, các tín hiệu có thanh âm có thể được tạo mã bằng kích thích giống xung.

Khối kích thích mã (được đánh chỉ dẫn bằng nhãn 308 trên Fig.3 và 402 trên Fig.4) minh họa vị trí của từ điển mã cố định (Fixed codebook - FCB) để tạo mã CELP tổng quát. Vector mã được lựa chọn từ FCB được tăng lên bởi độ khuếch đại thường được ghi chỉ là G_c 306.

Fig.5A và Fig.5B minh họa ví dụ về mã hóa/giải mã bằng BWE. Fig.5A minh họa các hoạt động ở bộ mã hóa với thông tin phía BWE trong khi Fig.5B minh họa các hoạt động ở bộ giải mã với BWE.

Tín hiệu băng thấp 501 được mã hóa bằng cách sử dụng các tham số băng thấp 502. Các tham số băng thấp 502 được lượng tử hóa và chỉ số lượng tử được tạo có thể được truyền qua kênh dòng bit 503. Tín hiệu băng cao được trích rút từ tín hiệu âm thanh/thoại 504 được mã hóa với số lượng bit nhỏ bằng cách sử dụng các tham số phía băng cao 505. Các tham số phía băng cao được lượng tử hóa (chỉ số thông tin phụ) được truyền qua kênh dòng bit 506.

Như được thể hiện trên Fig.5B, ở bộ giải mã, dòng bit băng thấp 507 được sử dụng để tạo tín hiệu băng thấp được giải mã 508. Dòng bit phía băng cao 510 được sử dụng để giải mã các tham số phía băng cao 511. Tín hiệu băng cao 512 được tạo từ tín hiệu băng thấp 508 với sự giúp đỡ từ các tham số phía băng cao 511. Tín hiệu âm thanh/thoại 509 cuối cùng được tạo bằng cách kết hợp tín hiệu băng thấp 508 và tín hiệu băng cao 512.

Fig.6A và Fig.6B minh họa ví dụ khác về mã hóa/giải mã với BWE mà không truyền thông tin phụ. Fig.6A minh họa các hoạt động trong quá trình ở bộ mã hóa trong khi Fig.6B minh họa các hoạt động ở bộ giải mã.

Như được thể hiện trên Fig.6A, tín hiệu băng thấp 601 được mã hóa bằng cách sử dụng các tham số băng thấp 602. Các tham số băng thấp 602 được lượng tử hóa để tạo chỉ số lượng tử hóa, có thể được truyền qua kênh dòng bit 603.

Như được thể hiện trên Fig.6B, ở bộ giải mã, dòng bit băng thấp 604 được sử dụng để tạo tín hiệu băng thấp được giải mã 605. Tín hiệu băng cao 607 được tạo từ tín hiệu băng thấp 605 mà không trợ giúp từ truyền thông tin phụ. Tín hiệu âm thanh/thoại 606 cuối cùng được tạo bằng cách kết hợp tín hiệu băng thấp 605 và tín hiệu băng cao 607.

Fig.7 minh họa ví dụ về phổ kích thích lý tưởng cho thoại có giọng nói hoặc âm nhạc điều hòa khi loại CELP của codec được sử dụng.

Phổ kích thích lý tưởng 702 hầu hết là phẳng sau khi loại bỏ đường bao phổ LPC 704. Phổ kích thích băng thấp lý tưởng 701 có thể được sử dụng làm tham chiếu để mã hóa kích thích băng thấp. Phổ kích thích băng cao lý tưởng 703 không khả dụng ở bộ giải mã. Về mặt lý thuyết, phổ kích thích băng cao lý tưởng hoặc không lượng tử hóa có thể có mức năng lượng gần như tương tự với phổ kích thích băng thấp.

Thực tế, phổ kích thích được giải mã hoặc tổng hợp không tốt như phổ kích thích lý tưởng được thể hiện trên Fig.7.

Fig.8 thể hiện ví dụ về phổ kích thích được giải mã cho thoại có giọng nói hoặc âm nhạc điều hòa khi loại CELP của codec được sử dụng.

Phổ kích thích được giải mã 802 hầu như phẳng sau khi loại bỏ đường bao phổ LPC 804. Phổ kích thích băng thấp được giải mã 801 là khả dụng ở bộ giải mã. Chất lượng của phổ kích thích băng thấp được giải mã 801 trở nên xấu hơn hoặc bị méo hơn đặc biệt trong vùng trong đó năng lượng đường bao là thấp. Đây là do các nguyên nhân. Chẳng hạn, hai nguyên nhân chính là việc tạo mã CELP vòng kín nhấn mạnh nhiều hơn vào vùng năng lượng cao hơn vùng năng lượng thấp, và rằng dạng sóng so khớp cho tín hiệu tần số thấp dễ hơn tín hiệu tần số cao do tín hiệu tần số cao

thay đổi nhanh hơn. Đối với việc tạo mã CELP tốc độ bit thấp như AMR-WB, băng cao thường không được mã hóa nhưng được tạo trong bộ giải mã với công nghệ BWE. Trong trường hợp này, phổ kích thích băng cao 803 có thể được sao chép đơn giản từ phổ kích thích băng thấp 801 và đường bao năng lượng phổ băng cao có thể được dự báo hoặc ước lượng từ đường bao năng lượng phổ băng thấp. Theo cách truyền thống, phổ kích thích băng cao được tạo 803 sau khi 6400Hz được sao chép từ băng phụ chỉ trước 6400Hz. Điều này có thể là tốt nếu chất lượng phổ tương đương từ 0Hz đến 6400Hz. Tuy nhiên, đối với codec CELP tốc độ bit thấp, chất lượng phổ có thể biến đổi nhiều từ 0Hz đến 6400Hz. Băng phụ được sao chép từ khu vực cuối của băng tần số thấp chỉ trước 6400Hz có thể có chất lượng kém, sau đó đưa âm thanh nhiễu ngoại vào khu vực băng cao từ 6400Hz đến 8000Hz.

Băng thông của băng tần số cao mở rộng thường nhỏ hơn nhiều băng thông của băng tần số thấp được tạo mã. Do vậy, theo các phương án thực hiện khác, băng phụ tốt nhất từ băng thấp được lựa chọn và được sao chép vào khu vực băng cao.

Băng phụ chất lượng cao có thể tồn tại ở vị trí bất kỳ trong toàn bộ băng tần số thấp. Vị trí khả thi nhất của băng phụ chất lượng cao trong vùng tương ứng với khu vực năng lượng phổ cao – khu vực focman phổ.

Fig.9 minh họa ví dụ về phổ kích thích được giải mã cho thoại có giọng nói hoặc âm nhạc điều hòa khi loại CELP của codec được sử dụng.

Phổ kích thích được giải mã 902 hầu như phẳng sau khi loại bỏ đường bao phổ LPC 904. Phổ kích thích băng thấp được giải mã 901 khả dụng ở bộ giải mã nhưng không khả dụng ở băng cao 903. Chất lượng của phổ kích thích băng thấp được giải mã 901 trở nên xấu hơn hoặc bị méo nhiều hơn đặc biệt trong vùng trong đó năng lượng của đường bao phổ 904 thấp hơn.

Trong trường hợp được minh họa trên Fig.9, theo một phương án thực hiện, băng phụ chất lượng cao được đặt quanh khu vực focman thoải thứ nhất (chẳng hạn, xấp xỉ 2000Hz trong ví dụ này theo phương án thực hiện). Theo các phương án thực hiện khác, băng phụ chất lượng cao có thể được đặt ở vị trí bất kỳ giữa 0Hz và 6400Hz.

Sau khi xác định vị trí của băng phụ tốt nhất, nó được sao chép từ trong băng thấp vào băng cao, như còn được minh họa trên Fig.9. Do vậy, phổ kích thích băng cao 903 được tạo bằng cách sao chép từ băng phụ được lựa chọn. Chất lượng cảm quan của băng cao 903 trên Fig.9 có vẻ tốt hơn nhiều so với băng cao 803 trên Fig.8 do phổ kích thích được cải thiện.

Theo một hoặc nhiều phương án thực hiện, nếu đường bao phổ băng thấp khả dụng trong miền tần số ở bộ giải mã, băng phụ tốt nhất có thể được xác định bằng cách tìm kiếm năng lượng băng phụ cao nhất từ tất cả các dự bị của băng phụ.

Theo cách khác, theo một hoặc nhiều phương án thực hiện, nếu đường bao phổ miền tần số không khả dụng, thì vị trí năng lượng cao cũng có thể được xác định từ các tham số bất kỳ mà có thể phản ánh đường bao năng lượng phổ hoặc đỉnh focman phổ. Vị trí băng phụ tốt nhất cho BWE tương ứng vị trí đỉnh phổ cao nhất.

Khoảng tìm kiếm của điểm bắt đầu băng phụ tốt nhất có thể phụ thuộc vào tốc độ bit codec. Chẳng hạn, đối với tốc độ bit codec rất thấp, khoảng tìm kiếm có thể từ 0Hz đến $6400-1600=4800\text{Hz}$ (2000Hz đến 4800Hz), giả sử băng thông của băng cao bằng 1600Hz. Ở ví dụ khác, đối với tốc độ bit codec trung bình, khoảng tìm kiếm có thể từ 2000Hz đến $6400-1600=4800\text{Hz}$ (2000Hz đến 4800Hz), giả sử băng thông của băng cao bằng 1600Hz.

Do đường bao phổ thay đổi chậm từ khung này sang khung khác, điểm bắt đầu băng phụ tốt nhất tương ứng với năng lượng focman phổ cao nhất

thông thường được thay đổi chậm. Để tránh dao động hoặc thay đổi thường xuyên điểm bắt đầu băng phụ tốt nhất từ khung này sang khung khác, có thể làm nhẵn trong suốt vùng có thanh âm tương tự trong miền thời gian, trừ khi năng lượng đỉnh phổ bị thay đổi đáng kể từ khung này sang khung khác hoặc vùng có thanh âm mới xuất hiện.

Fig.10 minh họa các hoạt động ở bộ giải mã theo các phương án thực hiện sáng chế để triển khai dịch chuyển băng phụ hoặc việc sao chép cho BWE.

Tín hiệu băng thấp miền thời gian 1002 được giải mã bằng cách sử dụng dòng bit nhận được 1001. Kích thích băng thấp miền thời gian 1003 thường khả dụng ở bộ giải mã. Đôi khi, kích thích băng thấp miền tần số cũng khả dụng. Nếu không khả dụng, kích thích băng thấp miền thời gian 1003 có thể được biến đổi thành miền tần số để có kích thích băng thấp miền tần số.

Đường bao phổ của thoại có thanh âm hoặc tín hiệu âm nhạc thường được biểu thị bằng các tham số LPC. Đôi khi, đường bao phổ miền tần số trực tiếp khả dụng ở bộ giải mã. Trong trường hợp bất kỳ, thông tin phân tán năng lượng 1004 có thể được trích rút từ các tham số LPC hoặc từ đường bao phổ miền tần số trực tiếp hoặc các tham số bất kỳ như miền DFT hoặc miền FFT. Bằng cách sử dụng thông tin phân tán năng lượng băng thấp 1004, băng phụ tốt nhất từ băng thấp được lựa chọn bằng cách tìm kiếm đỉnh năng lượng tương đối cao. Sau đó, băng phụ được lựa chọn được sao chép từ khu vực băng thấp đến khu vực băng cao. Sau đó đường bao phổ băng cao được dự báo hoặc được ước lượng được áp dụng cho khu vực băng cao, hoặc kích thích băng cao miền thời gian 1005 đi qua bộ lọc băng cao được ước lượng hoặc dự báo mà biểu thị đường bao phổ băng cao. Đầu ra của bộ lọc băng cao là tín hiệu băng cao 1006. Tín hiệu đầu ra thoại/âm thanh 1007 cuối cùng thu được bằng cách kết hợp tín hiệu băng thấp 1002 và tín hiệu băng cao 1006.

Fig.11 minh họa bộ giải mã để triển khai dịch chuyển băng phụ hoặc sao chép cho BWE theo phương án thực hiện khác.

Không giống Fig.10, Fig.11 giả thiết rằng miền tần số phổ băng thấp là khả dụng. Băng phụ tốt nhất ở băng tần số thấp được lựa chọn bằng cách tìm kiếm đơn giản đỉnh năng lượng tương đối cao trong miền tần số. Sau đó, băng phụ được lựa chọn được sao chép từ băng thấp sang băng cao. Sau khi áp dụng đường bao phổ băng cao được ước lượng, phổ băng cao 1103 được tạo. Phổ miền tần số thoại/âm thanh cuối cùng thu được bằng cách kết hợp phổ băng thấp 1102 và phổ băng cao 1103. Đầu ra tín hiệu miền thời gian thoại/âm thanh cuối cùng được tạo bằng cách biến đổi phổ miền tần số thoại/âm thanh thành miền thời gian.

Khi việc phân tích và tổng hợp dãy bộ lọc là khả dụng ở bộ giải mã bao phủ khoảng phổ mong muốn, thuật toán SBR có thể hiện thực việc dịch chuyển băng tần số bằng cách sao chép các hệ số băng tần số thấp của đầu ra tương ứng với băng thấp được lựa chọn từ việc phân tích dãy bộ lọc đến khu vực băng tần số cao.

Fig.12 minh họa các hoạt động được thực hiện ở bộ giải mã theo các phương án thực hiện sáng chế.

Như được thể hiện trên Fig.12, phương pháp giải mã dòng bit âm thanh được mã hóa ở bộ giải mã gồm tiếp nhận dòng bit âm thanh được tạo mã. Theo một hoặc nhiều phương án thực hiện, dòng bit âm thanh được tiếp nhận đã được tạo mã CELP. Cụ thể là, chỉ băng tần số thấp được tạo mã bằng CELP. CELP tạo chất lượng phổ tương đối cao hơn ở khu vực năng lượng phổ cao hơn so với khu vực năng lượng phổ thấp hơn. Do đó, các phương án thực hiện sáng chế gồm giải mã dòng bit âm thanh để tạo tín hiệu âm thanh băng thấp được giải mã và phổ kích thích băng thấp tương ứng với băng tần số thấp (hộp 1210). Khu vực băng phụ được lựa chọn từ trong băng tần số thấp bằng cách sử dụng thông tin phổ của đường bao phổ của tín hiệu âm thanh băng thấp được giải mã (hộp

1220). Phổ kích thích băng cao được tạo cho băng tần số cao bằng cách sao chép phổ kích thích băng phụ từ khu vực băng phụ được lựa chọn đến khu vực băng phụ cao tương ứng với băng tần số cao (hộp 1230). Tín hiệu đầu ra âm thanh được tạo bằng cách sử dụng phổ kích thích băng cao (hộp 1240). Cụ thể là, băng cách sử dụng phổ kích thích băng cao được tạo tín hiệu âm thanh băng cao được mở rộng được tạo bằng cách áp dụng đường bao phổ băng cao. Tín hiệu âm thanh băng cao được mở rộng được thêm vào tín hiệu âm thanh băng thấp được giải mã để tạo tín hiệu đầu ra âm thanh có băng thông tần số mở rộng.

Như được mô tả trước đó bằng cách sử dụng Fig.10 và Fig.11, các phương án thực hiện sáng chế có thể được áp dụng khác nhau tùy thuộc vào liệu đường bao phổ miền tần số có khả dụng hay không. Chẳng hạn, nếu đường bao phổ miền tần số là khả dụng, băng phụ có năng lượng băng phụ cao nhất có thể được lựa chọn. Tuy nhiên, nếu đường bao phổ miền tần số không khả dụng, thì phân tán năng lượng đường bao phổ có thể được nhận diện từ các tham số LPC (linear predictive coding – mã hóa dự báo tuyến tính), miền DFT (Discrete Fourier Transform – biến đổi Fourier rời rạc), hoặc các tham số miền FFT (Fast Fourier Transform – biến đổi Fourier nhanh). Tương tự, thông tin đỉnh phổ nếu khả dụng (hoặc tính toán được) có thể được sử dụng theo một số phương án thực hiện. Nếu chỉ kích thích băng thấp miền thời gian là khả dụng, kích thích băng thấp miền tần số có thể được tính toán bằng cách biến đổi kích thích băng thấp miền thời gian sang miền tần số.

Theo các phương án thực hiện, đường bao phổ có thể được tính toán bằng cách sử dụng phương pháp đã biết bất kỳ như người có kiến thức trung bình trong lĩnh vực đã biết. Chẳng hạn, trong miền tần số, đường bao phổ có thể chỉ là nhóm năng lượng mà biểu thị các năng lượng của nhóm các băng phụ. Tương tự, ở ví dụ khác, trong miền thời gian, đường bao phổ có thể được biểu thị bằng các tham số LPC. Các tham số LPC có

thể có nhiều dạng như các hệ số phản xạ, các hệ số LPC, các hệ số LSP, các hệ số LSF theo các phương án thực hiện khác.

Fig.13A và Fig.13B minh họa bộ giải mã triển khai BWE theo các phương án thực hiện sáng chế.

Như được thể hiện trên Fig.13A, bộ giải mã để giải mã dòng bit âm thanh được mã hóa bao gồm khối giải mã băng thấp 1310 được cấu hình để giải mã dòng bit âm thanh để tạo phổ kích thích băng thấp tương ứng với băng tần số thấp.

Bộ giải mã còn gồm khối mở rộng băng thông 1320 được ghép nối với khối giải mã băng thấp 1310 và bao gồm khối lựa chọn băng phụ 1330 và khối sao chép 1340. Khối lựa chọn băng phụ 1330 được cấu hình để lựa chọn khu vực băng phụ từ trong băng tần số thấp bằng cách sử dụng thông tin phổ của đường bao phổ của dòng bit âm thanh được giải mã. Khối sao chép 1340 được cấu hình để tạo phổ kích thích băng cao cho băng tần số cao bằng cách sao chép phổ kích thích băng phụ từ khu vực băng phụ được lựa chọn đến khu vực băng phụ cao tương ứng với băng tần số cao.

Bộ tạo tín hiệu băng cao 1350 được ghép nối với khối sao chép 1340. Bộ tạo tín hiệu băng cao 1350 được cấu hình để áp dụng đường bao phổ băng cao được dự báo để tạo tín hiệu miền thời gian băng cao. Bộ tạo đầu ra được ghép nối với bộ tạo tín hiệu băng cao 1350 và khối giải mã băng thấp 1310. Bộ tạo đầu ra 1360 được cấu hình để tạo tín hiệu đầu ra âm thanh bằng cách kết hợp tín hiệu miền thời gian băng thấp thu được bằng cách giải mã dòng bit âm thanh với tín hiệu miền thời gian băng cao.

Fig.13B minh họa bộ giải mã triển khai BWE theo phương án thực hiện khác.

Tương tự với Fig.13A, bộ giải mã trên Fig.13B còn gồm khối giải mã băng thấp 1310 và khối mở rộng băng thông 1320, được ghép nối với

khối giải mã băng thấp 1310, và bao gồm khối lựa chọn băng phụ 1330 và khối sao chép 1340.

Như được thể hiện trên Fig.13B, bộ giải mã còn gồm bộ tạo phổ băng cao 1355, được ghép nối với khối sao chép 1340. Bộ tạo tín hiệu băng cao 1355 được cấu hình để áp dụng năng lượng đường bao phổ băng cao để tạo phổ băng cao cho băng tần số cao bằng cách sử dụng phổ kích thích băng cao.

Bộ tạo phổ đầu ra 1365 được ghép nối với bộ tạo phổ băng cao 1355 và khối giải mã băng thấp 1310. Bộ tạo phổ đầu ra được cấu hình để tạo phổ âm thanh miền tần số bằng cách kết hợp phổ băng thấp thu được bằng cách giải mã dòng bit âm thanh từ khối giải mã băng thấp 1310 với phổ băng cao từ bộ tạo phổ băng cao 1355.

Bộ tạo tín hiệu biến đổi ngược 1370 được cấu hình để tạo tín hiệu âm thanh miền thời gian bằng cách biến đổi ngược phổ âm thanh miền tần số thành miền thời gian.

Các thành phần khác được mô tả trên Fig.13A và Fig.13B có thể được triển khai trong phần cứng theo một hoặc nhiều phương án thực hiện. Theo một số phương án thực hiện, chúng có thể được triển khai trong phần mềm và được thiết kế để vận hành trong bộ xử lý tín hiệu.

Do đó, các phương án thực hiện sáng chế có thể được sử dụng để cải thiện BWE ở bộ giải mã mà giải mã dòng bit âm thanh được tạo mã bởi CELP.

Fig.14 minh họa hệ thống truyền thông 10 theo phương án thực hiện sáng chế.

Hệ thống truyền thông 10 có các thiết bị truy nhập âm thanh 7 và 8 được ghép nối với mạng 36 qua các liên kết truyền thông 38 và 40. Theo một phương án thực hiện, thiết bị truy nhập âm thanh 7 và 8 là các thiết bị VOIP (voice over internet protocol – thoại trên giao thức Internet) và mạng 36 là mạng diện rộng (wide area network - WAN), PSTN (public

switched telephone network – mạng điện thoại chuyển mạch công cộng) và/hoặc mạng Internet. Theo phương án thực hiện khác, các liên kết truyền thông 38 và 40 là các kết nối băng rộng có dây và/hoặc không dây. Theo phương án thực hiện khác, các thiết bị truy nhập âm thanh 7 và 8 là các điện thoại di động hoặc tế bào, các liên kết 38 và 40 là các kênh điện thoại di động không dây và mạng 36 biểu thị mạng điện thoại di động.

Thiết bị truy nhập âm thanh 7 sử dụng micrô 12 để biến đổi âm thanh, như âm nhạc hoặc giọng nói của người thành tín hiệu đầu vào âm thanh tương tự 28. Giao diện micrô 16 biến đổi tín hiệu đầu vào âm thanh tương tự 28 thành tín hiệu âm thanh số 33 để nhập vào bộ mã hóa 22 của CODEC 20. Bộ mã hóa 22 tạo tín hiệu âm thanh được mã hóa TX để truyền đến mạng 26 qua giao diện mạng 26 theo các phương án thực hiện sáng chế. Bộ giải mã 24 trong CODEC 20 tiếp nhận tín hiệu âm thanh được mã hóa RX từ mạng 36 qua giao diện mạng 26, và biến đổi tín hiệu âm thanh được mã hóa RX thành tín hiệu âm thanh số 34. Giao diện ống nghe 18 biến đổi tín hiệu âm thanh số 34 thành tín hiệu âm thanh 30 thích hợp để điều khiển loa 14.

Theo các phương án thực hiện sáng chế, trong đó thiết bị truy nhập âm thanh 7 là thiết bị VOIP, một số hoặc tất cả các thành phần trong thiết bị truy nhập âm thanh 7 được triển khai trong máy điện thoại. Theo một số phương án thực hiện, tuy nhiên, micrô 12 và loa 14 là các khối riêng rẽ, và giao diện micrô 16, giao diện ống nghe 18, CODEC 20 và giao diện mạng 26 được triển khai trong máy tính cá nhân. CODEC 20 có thể được triển khai trong phần mềm chạy trên máy tính hoặc bộ xử lý dành riêng, hoặc bằng phần cứng dành riêng, chẳng hạn, trên ASIC (application specific integrated circuit – mạch tích hợp hướng ứng dụng). Giao diện micrô 16 được triển khai bằng bộ biến đổi A/D (analog-to-digital – tương tự sang số), cũng như hệ mạch giao diện khác được đặt trong máy điện thoại và/hoặc trong máy tính. Tương tự, giao diện ống nghe 18 được triển

khai bởi bộ biến đổi D/A (digital-to-analog – tương tự sang số) và hệ mạch giao diện khác được đặt trong máy điện thoại và/hoặc trong máy tính. Theo các phương án thực hiện khác, thiết bị truy nhập âm thanh 7 có thể được triển khai và phân vùng theo cách khác đã biết.

Theo các phương án thực hiện sáng chế trong đó thiết bị truy nhập âm thanh 7 là điện thoại di động hoặc tế bào, các phần tử trong thiết bị truy nhập âm thanh 7 được triển khai trong máy điện thoại tế bào. CODEC 20 được triển khai bằng phần mềm chạy trên bộ xử lý trong máy điện thoại hoặc bằng phần cứng dành riêng. Theo các phương án thực hiện khác của sáng chế, thiết bị truy nhập âm thanh có thể được triển khai trong các thiết bị khác như các hệ thống truyền thông số không dây và có dây ngang hàng, như các intercom, và các máy điện thoại vô tuyến. Trong các ứng dụng như các thiết bị âm thanh người dùng, thiết bị truy nhập âm thanh có thể chứa CODEC với chỉ bộ mã hóa 22 hoặc bộ giải mã 24, chẳng hạn, trong hệ thống micro số hoặc thiết bị phát lại âm nhạc. Theo các phương án thực hiện khác của sáng chế, CODEC 20 có thể được sử dụng không có micro 12 và ống nghe 14, chẳng hạn, trong các trạm gốc tế bào truy nhập PTSN.

Việc xử lý thoại để cải thiện phân loại có thanh âm/không có thanh âm được mô tả theo các phương án thực hiện khác của sáng chế có thể được triển khai trong bộ mã hóa 22 hoặc bộ giải mã 24, chẳng hạn. Việc xử lý thoại để cải thiện phân loại có thanh âm/không có thanh âm có thể được triển khai trong phần cứng hoặc phần mềm theo các phương án thực hiện khác. Chẳng hạn, bộ mã hóa 22 hoặc bộ giải mã 24 có thể là một phần của vi mạch DSP (digital signal processing - xử lý tín hiệu số).

Fig.15 minh họa sơ đồ khối của hệ thống xử lý có thể được sử dụng để triển khai các thiết bị và các phương pháp được bộc lộ ở đây. Các thiết bị cụ thể có thể sử dụng tất cả các thành phần được thể hiện, hoặc chỉ nhóm con của các thành phần, và các mức tích hợp có thể biến đổi từ thiết bị

này sang thiết bị khác. Ngoài ra, thiết bị có thể chứa nhiều thể hiện của thành phần, như nhiều khối xử lý, các bộ xử lý, các bộ nhớ, các bộ truyền, các bộ tiếp nhận, v.v.. Hệ thống xử lý có thể bao gồm khối xử lý có một hoặc nhiều thiết bị nhập/xuất, như ống nghe, micrô, chuột, màn hình cảm ứng, vùng phím bấm, bàn phím, máy in, phần hiển thị, và tương tự. Khối xử lý có thể gồm CPU (central processing unit – khối xử lý trung tâm), bộ nhớ, thiết bị lưu trữ khối, bộ điều hợp video, và giao diện I/O được kết nối với đường truyền.

Đường truyền có thể là một hoặc nhiều loại bất kỳ kiến trúc đường truyền gồm đường truyền bộ nhớ hoặc bộ điều khiển bộ nhớ, đường truyền ngoại vi, đường truyền video, hoặc tương tự. CPU có thể bao gồm loại bộ xử lý dữ liệu điện tử bất kỳ. Bộ nhớ có thể bao gồm loại bộ nhớ hệ thống bất kỳ như SRAM (static random access memory – bộ nhớ truy nhập ngẫu nhiên tĩnh), DRAM (dynamic RAM – RAM động), SDRAM (synchronous DRAM – DRAM đồng bộ), ROM (read-only memory – bộ nhớ chỉ đọc), kết hợp của chúng, hoặc tương tự. Theo phương án thực hiện, bộ nhớ có thể gồm ROM được sử dụng khi khởi động, và DRAM cho chương trình và lưu trữ dữ liệu được sử dụng khi thực thi chương trình.

Thiết bị lưu trữ khối có thể bao gồm loại thiết bị lưu trữ bất kỳ được cấu hình để lưu trữ dữ liệu, các chương trình, và thông tin khác và khiến dữ liệu, các chương trình, và thông tin khác truy nhập được qua đường truyền. Thiết bị lưu trữ khối có thể bao gồm, chẳng hạn, một hoặc nhiều trong số ổ trạng thái rắn, ổ đĩa cứng, ổ đĩa từ, ổ đĩa quang, hoặc tương tự.

Bộ điều hợp video và giao diện I/O cung cấp các giao diện để ghép nối các thiết bị nhập và xuất ngoại vi với khối xử lý. Như được minh họa, các ví dụ về các thiết bị nhập và xuất gồm phần hiển thị được ghép nối với bộ điều hợp video và máy in chuột/bàn phím/máy in được ghép nối với giao diện I/O. Các thiết bị khác có thể được ghép nối với khối xử lý, và một số

thể giao diện có thể được sử dụng. Chẳng hạn, giao diện nối tiếp như USB (Universal Serial Bus – đường truyền nối tiếp đa năng) (không được thể hiện trên hình vẽ) có thể được sử dụng để cung cấp giao diện cho máy in.

Khối xử lý còn gồm một hoặc nhiều giao diện mạng, có thể bao gồm các liên kết có dây, như cáp Ethernet hoặc tương tự, và/hoặc các liên kết không dây để truy nhập các nút hoặc các mạng khác nhau. Giao diện mạng cho phép khối xử lý truyền thông với các khối từ xa qua các mạng. Chẳng hạn, giao diện mạng có thể cung cấp truyền thông không dây qua một hoặc nhiều bộ truyền/anten truyền và một hoặc nhiều bộ nhận/anten nhận. Theo phương án thực hiện, khối xử lý được ghép nối với mạng cục bộ hoặc mạng diện rộng để xử lý dữ liệu và truyền thông với các thiết bị từ xa, như các khối xử lý khác, mạng Internet, các phương tiện lưu trữ từ xa, hoặc tương tự.

Trong khi sáng chế được mô tả có dựa vào các phương án thực hiện minh họa, phần mô tả này không được nhằm hiểu theo nghĩa giới hạn. Các cải biến và kết hợp khác nhau của các phương án thực hiện minh họa, cũng như các phương án thực hiện khác của sáng chế, sẽ trở nên rõ ràng với chuyên gia trong lĩnh vực khi tham khảo phần mô tả này. Chẳng hạn, các phương án thực hiện khác được mô tả trên đây có thể được kết hợp với nhau.

Mặc dù sáng chế và các ưu điểm của nó được mô tả chi tiết, song nên hiểu rằng các thay đổi khác, các thay thế có thể được thực hiện dưới đây mà không xa rời tinh thần và phạm vi của sáng chế như được định nghĩa bởi các điểm yêu cầu bảo hộ đi kèm. Chẳng hạn, nhiều dấu hiệu và chức năng được đề cập trên đây có thể được triển khai ở phần mềm, phần cứng, hoặc phần sụn, hoặc kết hợp của chúng. Ngoài ra, phạm vi của sáng chế không được nhằm giới hạn ở các phương án thực hiện cụ thể của quá trình, máy móc, sản xuất, thành phần, phương tiện, các phương pháp và

các bước được mô tả trong sáng chế. Người có hiểu biết trung bình trong lĩnh vực sẽ hiểu rằng từ phần mô tả sáng chế, các quá trình, máy móc, sản xuất, thành phần, phương tiện, các phương pháp, hoặc các bước, có sẵn hoặc sẽ được phát triển sau này, mà thực hiện chức năng gần như tương tự hoặc đạt được kết quả gần như tương tự với các phương án thực hiện tương ứng được mô tả dưới đây có thể được sử dụng theo sáng chế. Do đó, các điểm yêu cầu bảo hộ đi kèm được bao gồm trong phạm vi của chúng như các quá trình, máy móc, sản xuất, các thành phần, phương tiện, các phương pháp, hoặc các bước thực hiện.

YÊU CẦU BẢO HỘ

1. Phương pháp giải mã dòng bit âm thanh được mã hóa và tạo mở rộng băng thông tần số ở bộ giải mã, phương pháp bao gồm các bước:

giải mã dòng bit âm thanh để tạo tín hiệu âm thanh băng thấp được giải mã và tạo phổ kích thích dải thấp tương ứng với băng tần số thấp;

lựa chọn khu vực băng phụ từ trong băng tần số thấp bằng cách sử dụng tham số mà chỉ báo thông tin năng lượng của đường bao đặc biệt của tín hiệu âm thanh băng thấp được giải mã, trong đó khu vực băng phụ được chọn trong vùng tương ứng với điểm cao nhất năng lượng của đường bao phổ;

tạo phổ kích thích băng cao cho băng tần số cao bằng cách sao chép phổ kích thích băng phụ từ khu vực băng phụ được lựa chọn đến khu vực băng phụ cao tương ứng với băng tần số cao;

tạo tín hiệu âm thanh băng cao được mở rộng theo đường bao phổ băng cao và phổ kích thích băng cao được tạo; và

thêm tín hiệu âm thanh băng cao được mở rộng vào tín hiệu âm thanh băng thấp được giải mã để tạo tín hiệu đầu ra âm thanh có băng thông tần số mở rộng.

2. Phương pháp theo điểm 1, trong đó lựa chọn khu vực băng phụ từ trong băng tần số thấp bằng cách sử dụng tham số bao gồm nhận diện băng phụ chất lượng cao nhất trong băng tần số thấp bằng cách tìm kiếm điểm năng lượng cao nhất của đường bao phổ và lựa chọn băng phụ chất lượng cao nhất được nhận diện.

3. Phương pháp theo điểm 1, trong đó lựa chọn khu vực băng phụ từ trong băng tần số thấp bằng cách sử dụng tham số bao gồm lựa chọn khu vực băng phụ tương ứng với năng lượng đường bao phổ cao nhất.

4. Phương pháp theo điểm 1, trong đó lựa chọn khu vực băng phụ từ trong băng tần số thấp bằng cách sử dụng tham số bao gồm nhận diện băng phụ từ trong băng thấp bằng cách sử dụng các tham số phản ánh năng lượng cao nhất của đường bao phổ hoặc đỉnh focman phổ và lựa chọn băng phụ được nhận diện.

5. Phương pháp theo điểm 1, trong đó phương pháp giải mã áp dụng kỹ thuật mở rộng băng thông để tạo băng tần số cao.

6. Phương pháp theo điểm 1, trong đó việc áp dụng đường bao phổ băng cao bao gồm áp dụng bộ lọc băng cao được dự đoán biểu thị đường bao phổ băng cao.

7. Phương pháp theo điểm 1, còn bao gồm bước:

tạo tín hiệu đầu ra âm thanh bằng cách biến đổi ngược phổ âm thành miền tần số thành miền thời gian.

8. Phương pháp theo điểm 1, trong đó việc sao chép phổ kích thích băng phụ từ khu vực băng phụ được lựa chọn đến khu vực băng phụ cao tương ứng với băng tần số cao bao gồm việc sao chép các hệ số băng tần số thấp của đầu ra từ việc phân tích dãy bộ lọc đến khu vực băng phụ cao.

9. Phương pháp theo điểm 1, trong đó dòng bit âm thanh bao gồm thoại có thanh âm hoặc âm nhạc điều hòa.

10. Thiết bị giải mã để xử lý thoại bao gồm:

bộ xử lý; và

vật lưu trữ máy tính đọc được được ghép nối với bộ xử lý; trong đó vật lưu trữ máy tính đọc được được tạo cấu hình để lưu trữ chương trình

thực thi được, trong đó bộ xử lý được tạo cấu hình để chạy chương trình thực thi được đề:

giải mã dòng bit âm thanh để tạo tín hiệu âm thanh băng thấp được giải mã và tạo phổ kích thích dải thấp tương ứng với băng tần số thấp,

lựa chọn khu vực băng phụ từ trong băng tần số thấp bằng cách sử dụng tham số mà chỉ báo thông tin năng lượng của đường bao đặc biệt của tín hiệu âm thanh băng thấp được giải mã, trong đó khu vực băng phụ được chọn trong vùng tương ứng với điểm cao nhất năng lượng của đường bao phổ;

tạo phổ kích thích băng cao cho băng tần số cao bằng cách sao chép phổ kích thích băng phụ từ khu vực băng phụ được lựa chọn đến khu vực băng phụ cao tương ứng với băng tần số cao,

tạo tín hiệu âm thanh băng cao được mở rộng theo đường bao phổ băng cao và phổ kích thích băng cao được tạo, và

thêm tín hiệu âm thanh băng cao được mở rộng vào tín hiệu âm thanh băng thấp được giải mã để tạo tín hiệu đầu ra âm thanh có băng thông tần số mở rộng.

11. Thiết bị giải mã theo điểm 10, trong đó bộ xử lý được tạo cấu hình để chạy chương trình thực thi được đề:

nhận diện băng phụ chất lượng cao nhất trong băng tần số thấp bằng cách tìm kiếm điểm năng lượng cao nhất của đường bao phổ; và

lựa chọn băng phụ chất lượng cao nhất được nhận diện.

12. Thiết bị giải mã theo điểm 10, trong đó bộ xử lý được tạo cấu hình để chạy chương trình thực thi được đề:

lựa chọn khu vực băng phụ tương ứng với năng lượng đường bao phổ cao nhất.

13. Thiết bị giải mã theo điểm 10, trong đó bộ xử lý được tạo cấu hình để chạy chương trình thực thi được đề:

nhận diện băng phụ từ trong băng thấp bằng cách sử dụng các tham số phản ánh năng lượng cao nhất của đường bao năng lượng phổ hoặc đỉnh focman phổ và lựa chọn băng phụ được nhận diện.

14. Thiết bị giải mã theo điểm 10, trong đó bộ xử lý được tạo cấu hình để chạy chương trình thực thi được đề:

áp dụng bộ lọc băng cao được dự đoán biểu thị đường bao phổ băng cao để tạo tín hiệu miền thời gian băng cao; và

tạo tín hiệu đầu ra âm thanh bằng cách kết hợp tín hiệu miền thời gian băng thấp thu được bằng cách giải mã dòng bit âm thanh có tín hiệu miền thời gian băng cao.

15. Thiết bị giải mã theo điểm 10, trong đó bộ xử lý được tạo cấu hình để chạy chương trình thực thi được đề:

sao chép phổ kích thích băng phụ từ khu vực băng phụ được lựa chọn đến khu vực băng phụ cao.

16. Thiết bị giải mã theo điểm 10, trong đó bộ xử lý được tạo cấu hình để chạy chương trình thực thi được đề:

áp dụng đường bao phổ băng cao được ước tính để tạo phổ băng cao cho băng cao tần sử dụng phổ kích thích băng cao; và

tạo phổ âm thanh miền tần bằng cách kết hợp phổ băng thấp thu được bằng cách giải mã dòng bit âm thanh có phổ băng cao.

17. Thiết bị giải mã theo điểm 10, trong đó bộ xử lý được tạo cấu hình để chạy chương trình thực thi được đề:

Tạo tín hiệu âm thanh miền thời gian bằng cách biến đổi ngược phổ âm thanh miền tần số thành miền thời gian.

18. Phương pháp giải mã dòng bit âm thanh được mã hóa và tạo mở rộng băng thông tần số ở bộ giải mã, phương pháp bao gồm các bước:

giải mã dòng bit âm thanh để tạo tín hiệu âm thanh băng thấp được giải mã và tạo phổ kích thích dải thấp tương ứng với băng tần số thấp;

lựa chọn khu vực băng phụ từ trong băng tần số thấp bằng cách sử dụng tham số mà chỉ báo thông tin năng lượng của đường bao đặc biệt của tín hiệu âm thanh băng thấp được giải mã; trong đó khu vực băng phụ được chọn nằm trong vùng tương ứng với điểm cao nhất năng lượng của đường bao phổ;

tạo phổ kích thích băng cao cho băng tần số cao bằng cách sao chép phổ kích thích băng phụ từ khu vực băng phụ được lựa chọn đến khu vực băng phụ cao;

sử dụng phổ kích thích băng cao được tạo để tạo tín hiệu âm thanh băng cao được mở rộng bằng cách áp dụng đường bao phổ băng cao; và

thêm tín hiệu âm thanh băng cao được mở rộng vào tín hiệu âm thanh băng thấp được giải mã để tạo tín hiệu đầu ra âm thanh có băng thông tần số mở rộng.

19. Phương pháp theo điểm 18, trong đó việc lựa chọn khu vực băng phụ từ trong băng tần số thấp sử dụng tham số bao gồm việc lựa chọn khu vực băng phụ tương ứng với năng lượng đường bao phổ cao nhất.

20. Phương pháp theo điểm 18, trong đó việc áp dụng đường bao phổ băng cao bao gồm áp dụng bộ lọc băng cao được dự báo biểu thị đường bao phổ băng cao.

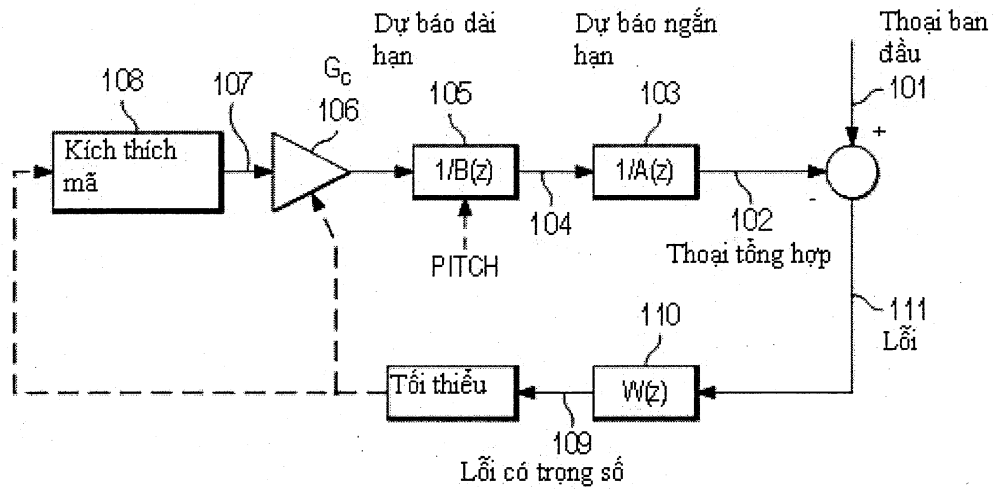


Fig.1

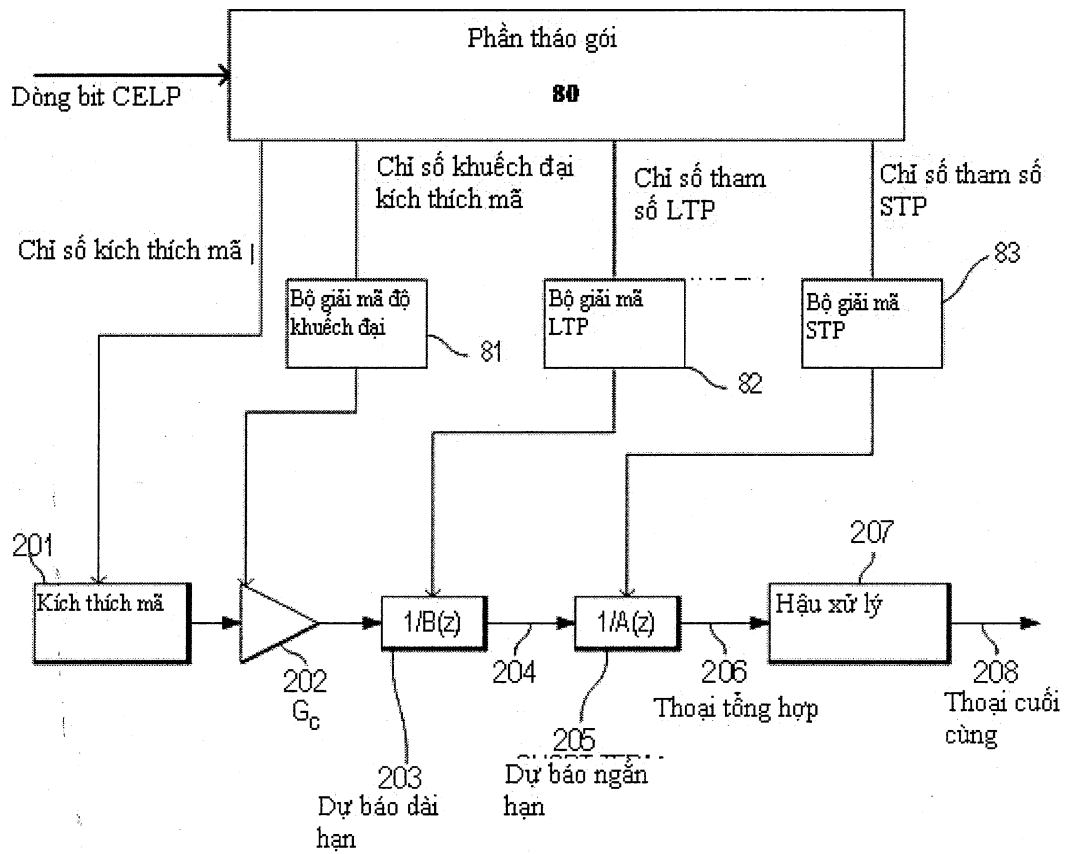


Fig.2

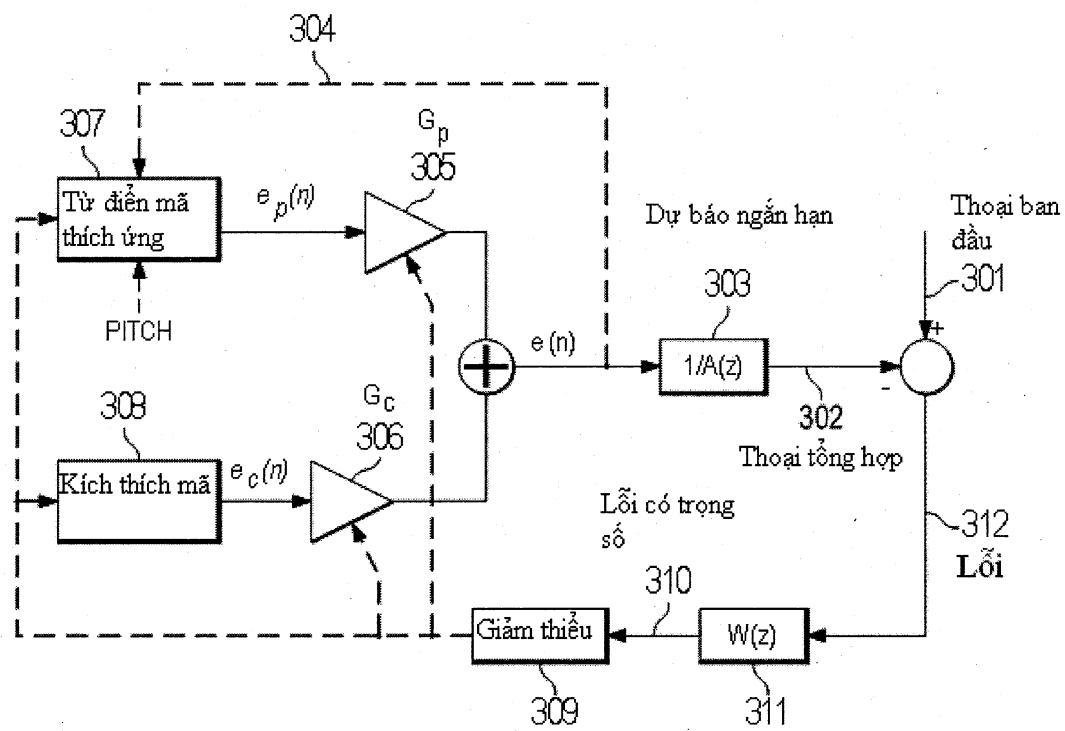


Fig.3

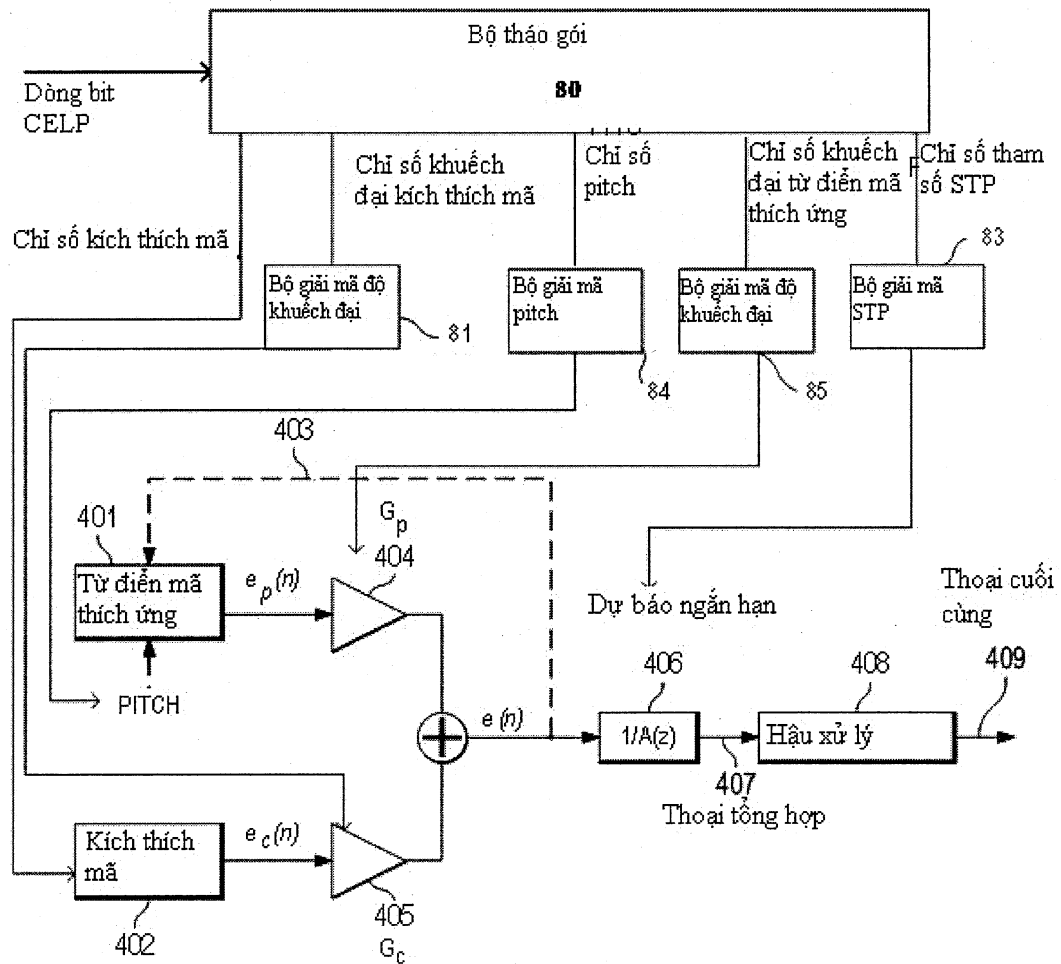


Fig.4

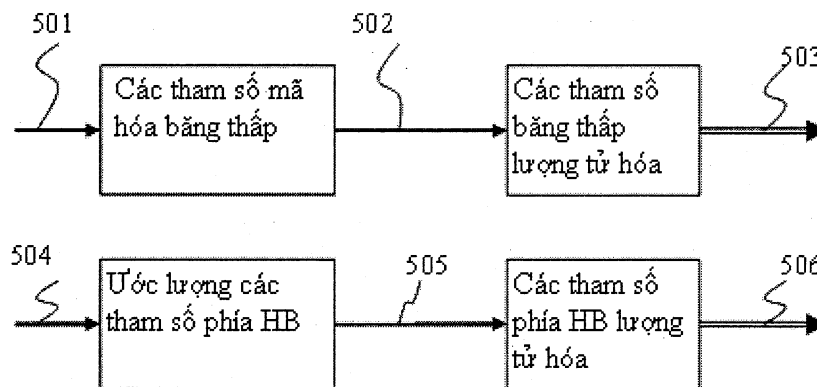


Fig.5A

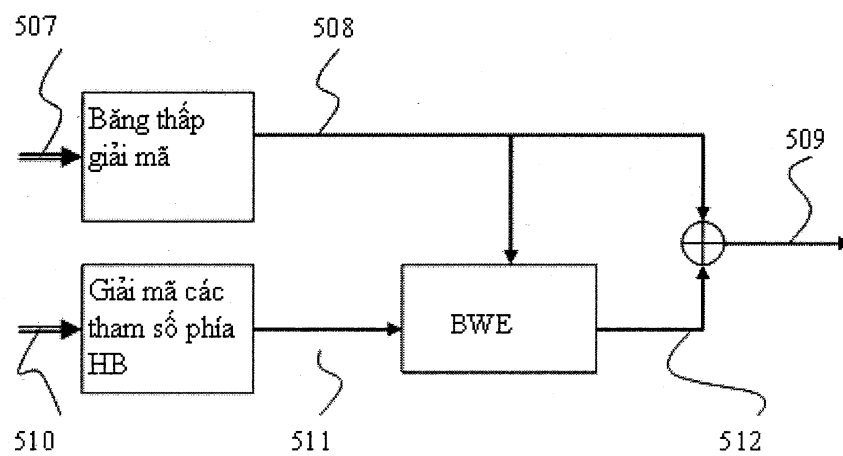


Fig.5B

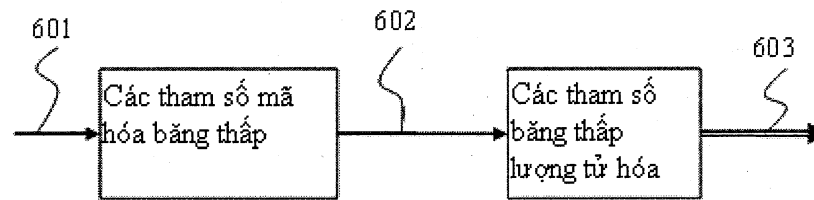


Fig. 6A

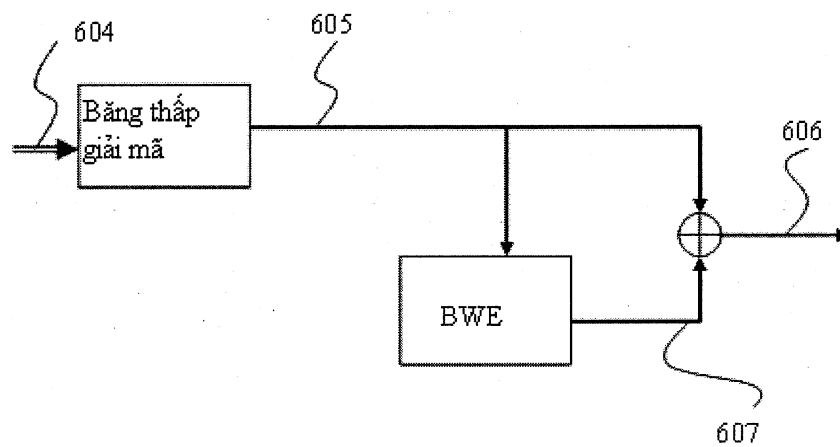


Fig. 6B

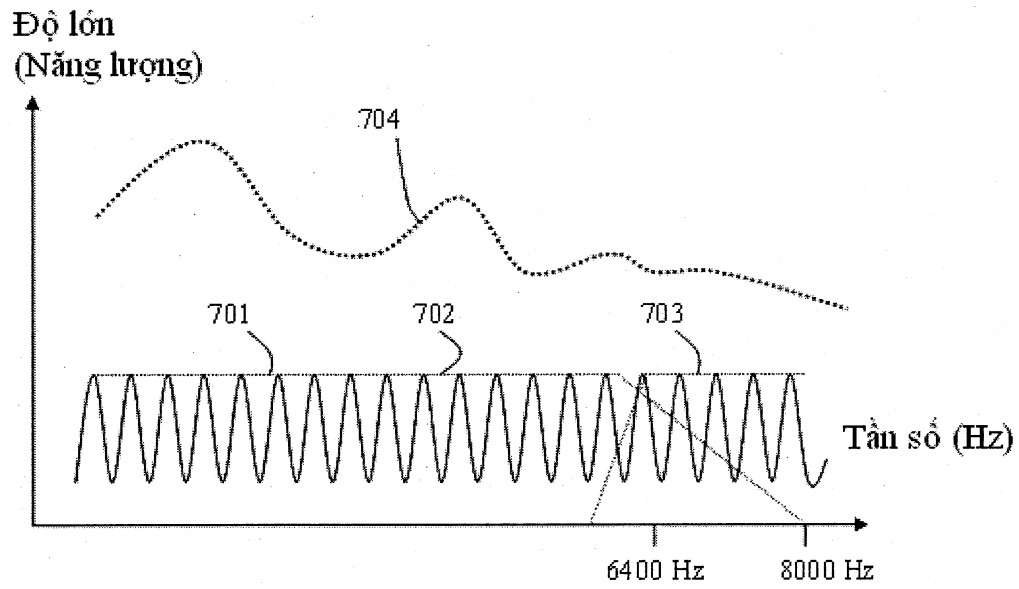


Fig. 7

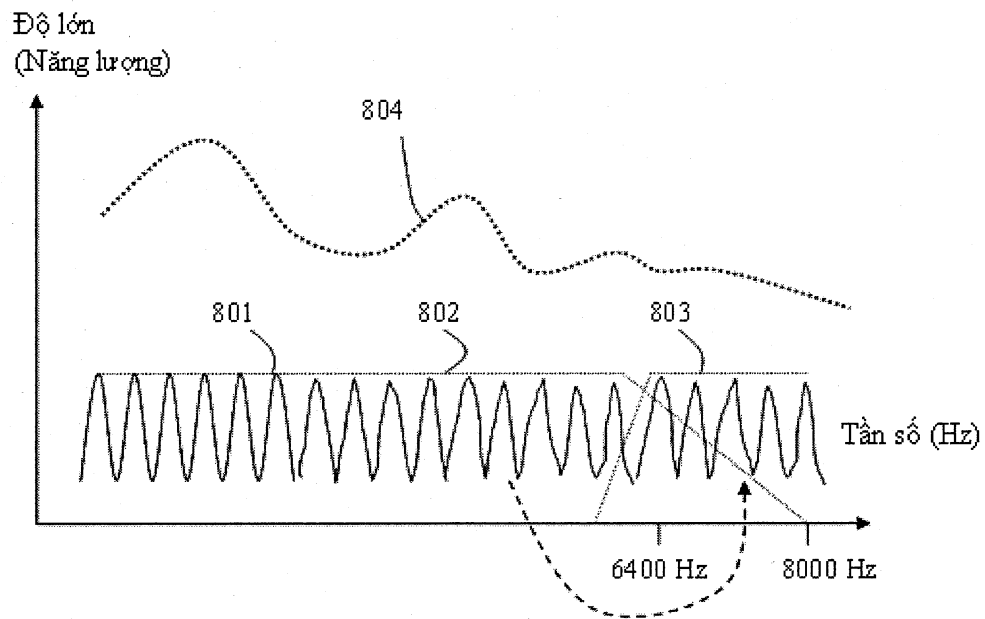


Fig.8

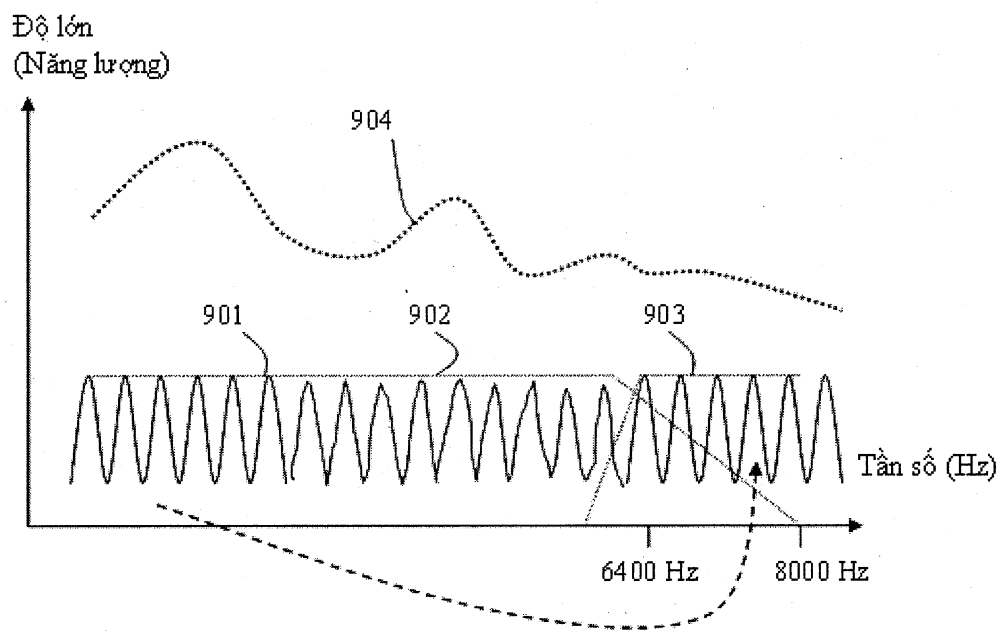


Fig.9

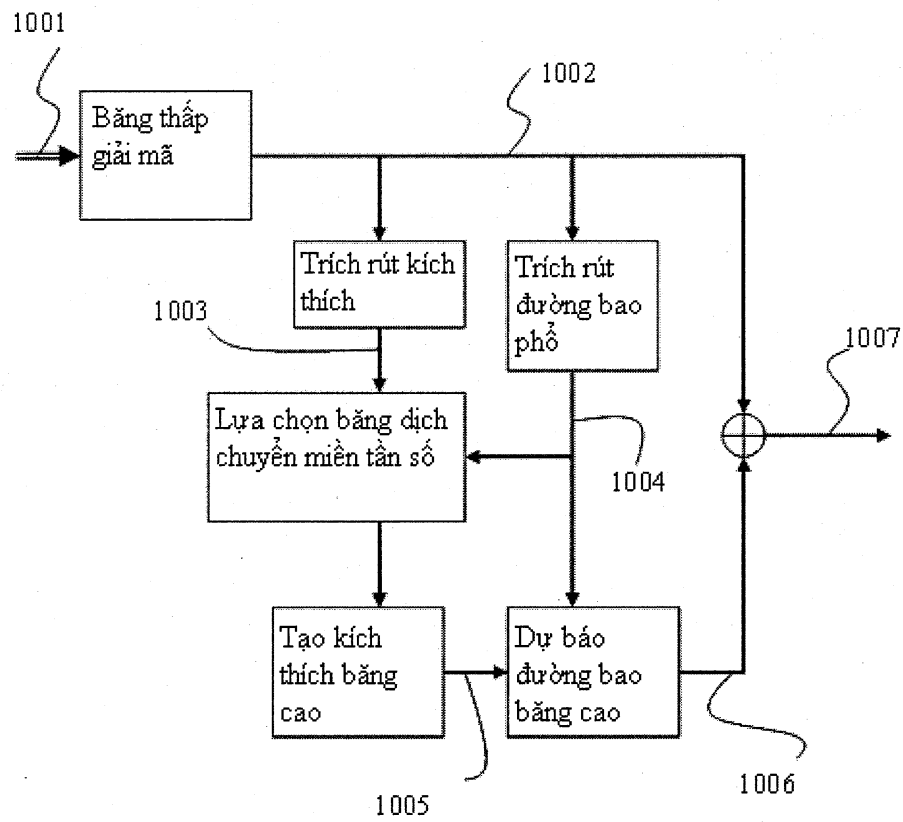


Fig.10

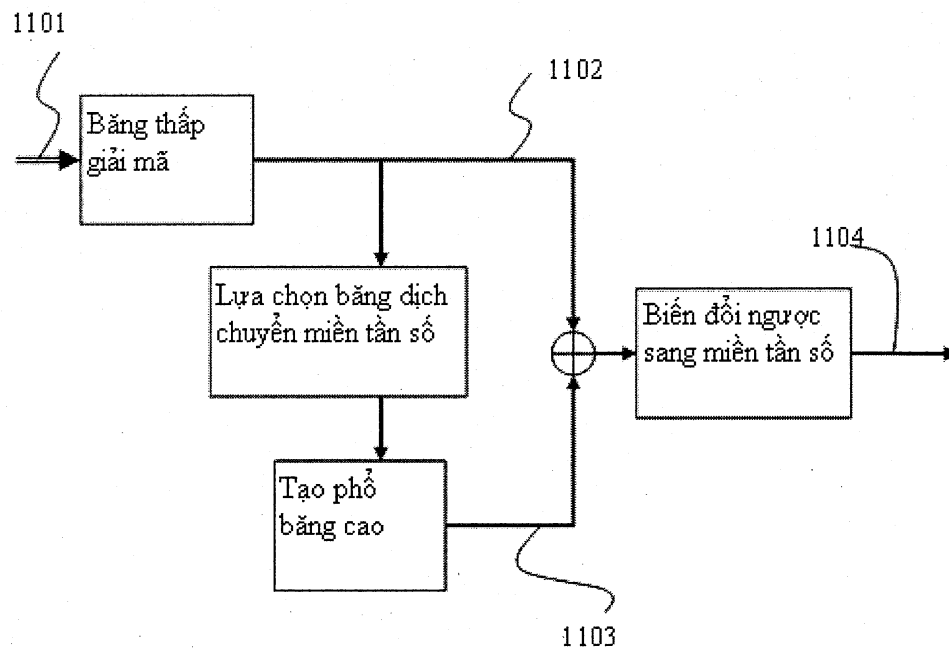


Fig.11

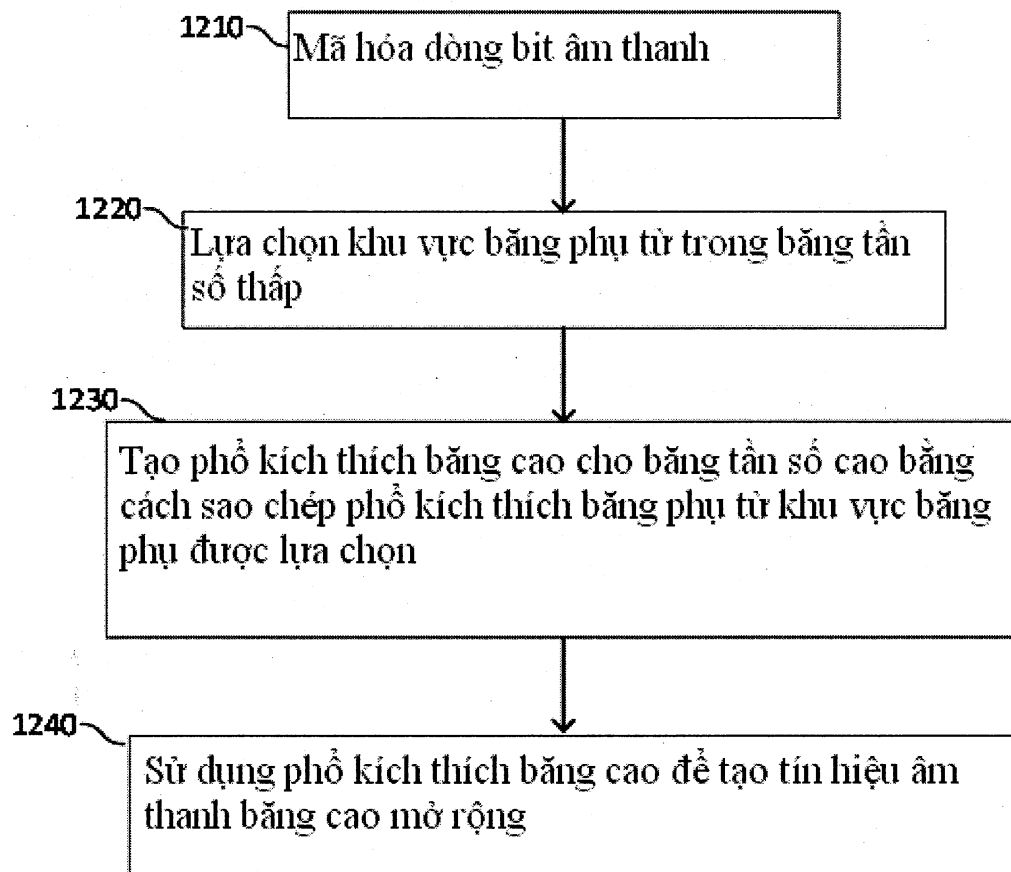


Fig.12

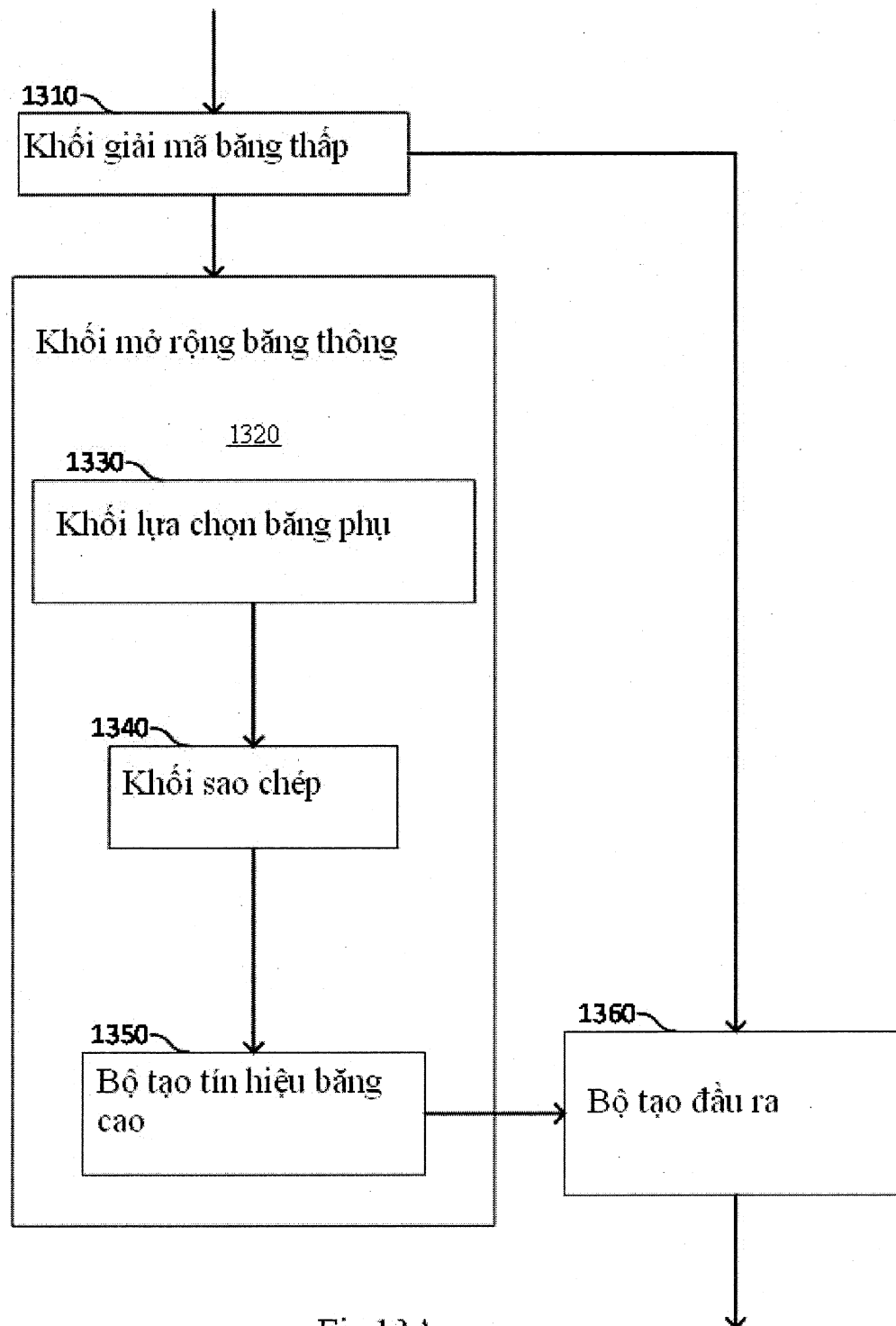


Fig.13A

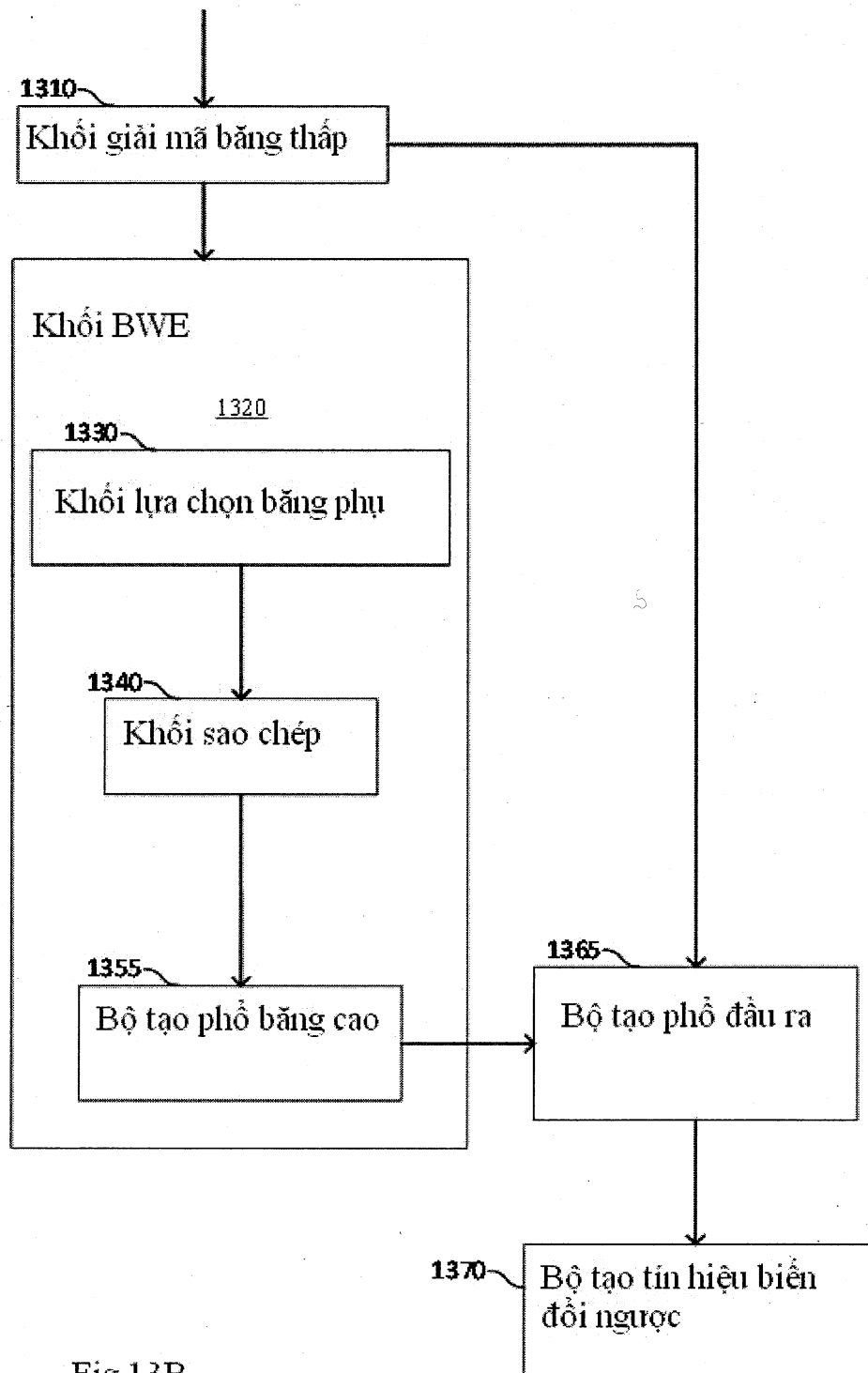


Fig.13B

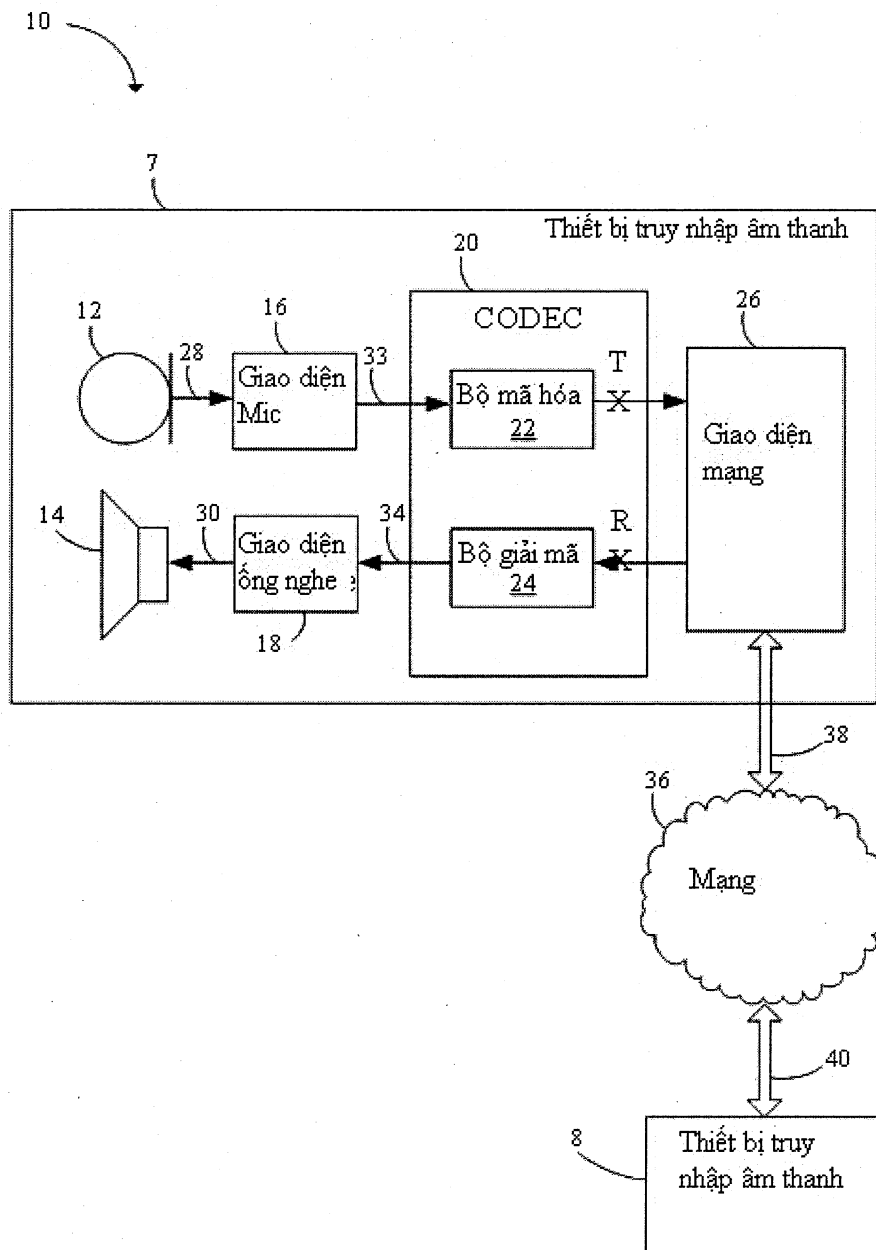


Fig.14

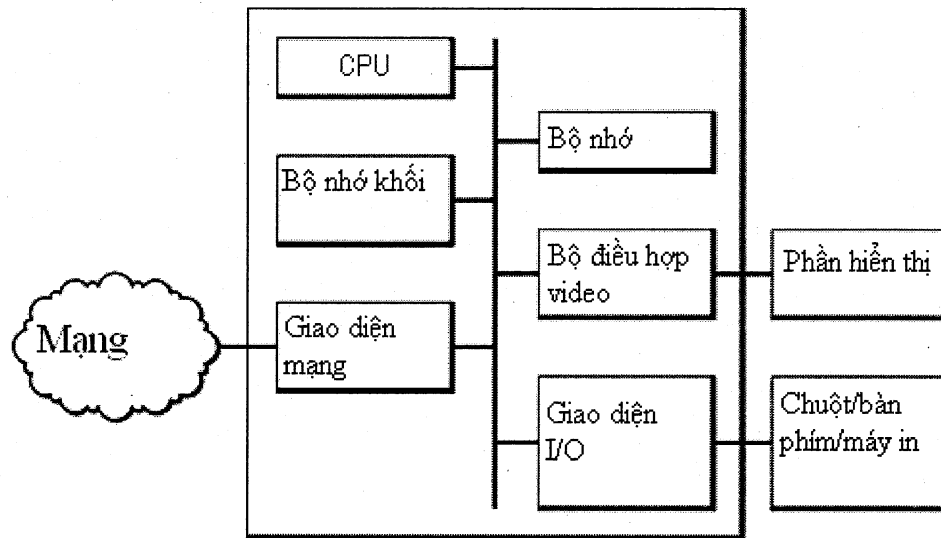


Fig.15