



(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)



1-0024064

(51)⁷ G10L 21/0388

(13) B

(21) 1-2015-03123

(22) 28/01/2014

(86) PCT/EP2014/051591 28/01/2014

(87) WO2014/118155 07/08/2014

(30) 61/758,092 29/01/2013 US

(45) 25/06/2020 387

(43) 25/12/2015 333A

(73) Fraunhofer-Gesellschaft zur Foerderung der angewandten Forschung e.V. (DE)
Hansastraße 27c, 80686 Muenchen, Germany

(72) NAGEL, Frederik (DE); DISCH, Sascha (DE); NIEDERMEIER, Andreas (DE).

(74) Công ty Luật TNHH AMBYS Hà Nội (AMBYS HANOI)

(54) BỘ GIẢI MÃ VÀ PHƯƠNG PHÁP TẠO RA TÍN HIỆU ÂM THANH ĐƯỢC
NÂNG CAO TẦN SỐ, BỘ MÃ HÓA VÀ PHƯƠNG PHÁP TẠO RA TÍN HIỆU
ĐƯỢC MÃ HÓA

(57) Sáng chế đề cập đến bộ giải mã và phương pháp để tạo ra tín hiệu âm thanh được nâng cao tần số, bộ mã hóa và phương pháp để tạo ra tín hiệu được mã hóa. Bộ giải mã để tạo ra tín hiệu âm thanh được nâng cao tần số (120), bao gồm: bộ trích xuất đặc tính (104) để trích xuất ra đặc tính từ tín hiệu lỗi (100); bộ trích xuất thông tin phụ (110) để trích xuất ra thông tin phụ lựa chọn được liên kết với tín hiệu lỗi; bộ tạo tham số (108) để tạo ra biểu diễn theo tham số để ước lượng vùng quang phổ của tín hiệu âm thanh được nâng cao tần số (120) không được xác định bởi tín hiệu lỗi (100), trong đó bộ tạo tham số (108) được tạo cấu hình để cung cấp số lượng lựa chọn biểu diễn theo tham số (702, 704, 706, 708) đáp ứng với đặc tính (112), và trong đó bộ tạo tham số (108) được tạo cấu hình để lựa chọn một trong số các lựa chọn biểu diễn theo tham số như biểu diễn theo tham số đáp ứng với thông tin phụ lựa chọn (712 đến 718); và bộ ước lượng tín hiệu (118) để ước lượng tín hiệu âm thanh được nâng cao tần số (120) sử dụng biểu diễn theo tham số được lựa chọn.

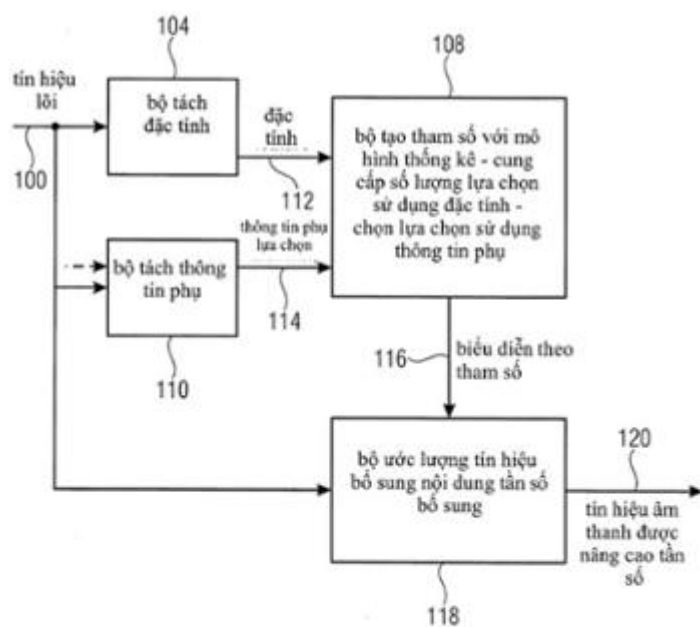
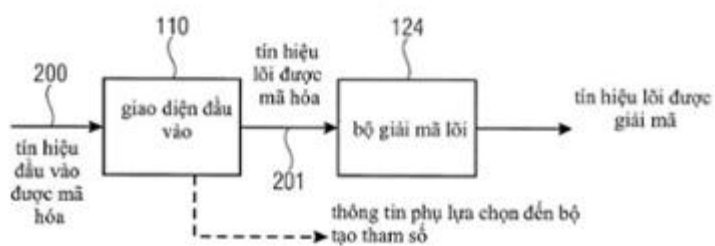


FIG 1



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến mã hóa âm thanh và, cụ thể đề cập đến mã hóa âm thanh trong phạm vi nâng cao tần số, tức là tín hiệu đầu ra bộ giải mã có số dải tần số cao hơn so với tín hiệu được mã hóa. Các quy trình đó bao gồm mở rộng băng thông, lặp lại quang phổ hoặc làm đầy khe hở thông minh.

Tình trạng kỹ thuật của sáng chế

Các hệ thống mã hóa tiếng nói hiện thời có khả năng mã hóa nội dung âm thanh dạng số băng rộng (wideband - WB), đó là, các tín hiệu với các tần số lên đến 7-8 kilôhec (Kilohertz – kHz), tại tốc độ bit thấp tới 6 kilobit trên giây (kilobit per second - kbit/s). Các ví dụ được thảo luận rộng rãi nhất là các khuyến nghị ITU-T G.722.2, tài liệu [1] cũng như G.718, tài liệu [4, 10] được phát triển gần đây và mã hóa tiếng nói và âm thanh hợp nhất (Unified Speech and Audio Coding - USAC) MPEG-D, tài liệu [8]. Cả G.722.2, cũng được biết đến như AMR-WB, và G.718 sử dụng các kỹ thuật mở rộng băng thông (bandwidth extension - BWE) trong khoảng giữa 6,4 và 7kHz để cho phép bộ mã hóa lõi ACELP nằm dưới tập trung trên các tần số thấp hơn thích hợp hơn về nhận thức (đặc biệt các tần số mà hệ thống thính lực con người nhạy cảm), và bằng cách đó đạt được đủ chất lượng đặc biệt tại các tốc độ bit thấp. Trong biên dạng mã hóa âm thanh nâng cao hiệu quả cao được mở rộng (eXtended High Efficiency Advanced Audio Coding - xHE-AAC) USAC, sự lặp lại dải quang phổ được nâng cao (enhanced spectral band replication - eSBR) được sử dụng để mở rộng băng thông âm thanh vượt qua băng thông bộ mã hóa lõi mà thường dưới 6kHz tại 16kbit/s. Các xử lý BWE theo giải pháp kỹ thuật đã biết hiện nay có thể thường được chia ra thành hai phương thức tiếp cận khái niệm:

- BWE mù hay BWE nhân tạo, trong đó các thành phần tần số cao (high-frequency - HF) được khôi phục chỉ từ tín hiệu mã hóa lõi tần số thấp (low-frequency - LF) được giải mã, tức là, không yêu cầu thông tin phụ được truyền từ bộ mã hóa. Sơ đồ này được sử dụng bởi AMR-WB và G.178 tại và dưới

16kbit/s, cũng như một số bộ xử lý sau BWE tương thích theo chiều ngược hoạt động trên tiếng nói điện thoại băng hẹp truyền thống, tài liệu [5, 9, 12] (ví dụ: Fig.15).

- BWE dẫn, mà khác với BWE mà ở chỗ một số tham số được sử dụng cho sự khôi phục nội dung HF được truyền đến bộ giải mã như thông tin phụ thay vì được ước lượng từ tín hiệu lỗi được giải mã. AMR-WB, G.718, xHE-AAC, cũng như một số bộ mã hóa-giải mã khác, các tài liệu [2, 7, 11] sử dụng phương thức tiếp cận này, nhưng không sử dụng tại các tốc độ bit thấp (Fig.16).

Fig.15 minh họa việc mở rộng băng thông mà hoặc mở rộng băng thông nhân tạo như được mô tả trong công bố của Bernd Geiser, Peter Jax, and Peter Vary: “ROBUST WIDEBAND ENHANCEMENT OF SPEECH BY COMBINED CODING AND ARTIFICIAL BANDWIDTH EXTENSION”, diễn ra tại hội thảo quốc tế về điều khiển tiếng ồn và vọng âm thanh (IWAENC), 2005. Thuật toán mở rộng băng thông riêng-tiêu chuẩn được minh họa trên Fig.15 bao gồm phép nội suy 1500, bộ lọc phân tích 1600, mở rộng kích thích 1700, bộ lọc tổng hợp 1800, quy trình trích xuất đặc tính 1510, quy trình ước lượng đường bao 1520 và mô hình thống kê 1530. Sau phép nội suy của tín hiệu dải hẹp thành tốc độ mẫu dải rộng, vector đặc tính được tính toán. Sau đó, bằng cách của mô hình Markov ẩn thống kê được đào tạo trước (hidden Markov model - HMM), việc ước lượng cho đường bao quang phổ băng rộng được xác định trong giới hạn của các hệ số dự báo tuyến tính (linear prediction - LP). Các hệ số dự báo tuyến tính băng rộng được sử dụng để lọc phân tích tín hiệu băng hẹp được nội suy. Sau khi mở rộng kích thích thu được, bộ lọc tổng hợp nghịch đảo được áp dụng. Lựa chọn mở rộng kích thích mà không thay đổi băng hẹp là trong suốt đối với các thành phần băng hẹp.

Fig.16 minh họa mở rộng băng thông với thông tin phụ như được mô tả trong công bố được nêu ở trên, mở rộng băng thông bao gồm dải giữa điện thoại 1620, khối trích xuất ra thông tin phụ 1610, bộ mã hóa (nội) 1630, bộ giải mã 1640 và khối mở rộng băng thông 1650. Hệ thống này để nâng cao băng thông của tín hiệu tiếng nói băng sai số bằng cách mã hóa được tổ hợp và mở rộng băng thông được minh họa trên Fig.16. Tại thiết bị đầu cuối truyền dẫn, đường bao quang phổ băng cao của tín hiệu

đầu vào băng thông được phân tích và thông tin phụ được xác định. Kết quả là thông điệp m được mã hóa hoặc riêng rẽ hoặc kết hợp với tín hiệu tiếng nói băng hẹp. Tại bộ nhận, thông tin phụ bộ giải mã được sử dụng để hỗ trợ việc ước lượng đường bao băng rộng trong thuật toán mở rộng băng thông. Thông điệp m thu được bằng một số thủ tục. Biểu diễn quang phổ của các tần số trong khoảng từ 3,4kHz đến 7kHz được trích xuất ra từ tín hiệu băng thông có sẵn chỉ tại phía gửi.

Đường bao băng phụ này được tính toán bởi dự báo tuyến tính lựa chọn, tức là việc tính toán của phổ năng lượng băng thông được theo sau bởi IDFT của các thành phần băng cao hơn của nó và sự đệ quy Levinson-Durbin tiếp theo của thứ tự 8. Các hệ số LPC băng phụ thu được được chuyển đổi sang miền cepstral và được lượng tử hóa cuối cùng bởi bộ lượng tử hóa vector với bảng mã kích thước $M = 2^N$. Với chiều dài khung là 20 phần nghìn giây, kết quả này trong tốc độ dữ liệu thông tin phụ là 300bit/giây. Phép xấp xỉ ước lượng được tổ hợp mở rộng tính toán của xác xuất sau và đưa lại sự phụ thuộc trên đặc tính dải hẹp. Vì vậy, dạng được cải thiện của chỗ ẩn sai số thu được mà sử dụng nhiều hơn một nguồn thông tin cho việc ước lượng tham số của nó.

Lưỡng đề dilem chất lượng nhất định trong bộ mã hóa-giải mã WB có thể được quan sát tại các tốc độ bit thấp, thông thường dưới 10kbit/giây. Một mặt, các tốc độ này đã là quá thấp để xác minh sự truyền dẫn của thậm chí lượng vừa phải của dữ liệu BWE, đường sinh ngoài các hệ thống BWE được dẫn thông thường với tốc độ 1kbit/giây hoặc cao hơn của thông tin phụ. Mặt khác, BWE mù khả thi được tìm thấy âm kém hơn đáng kể trên ít nhất một vài dạng nguyên liệu tiếng nói hoặc âm nhạc do sự không có khả năng dự báo tham số thích hợp từ tín hiệu lỗi. Điều này đặc biệt đúng đối với một số âm thanh phát âm như phụ âm sát với tương quan thấp giữa HF và LF. Do đó mong muốn để giảm tốc độ thông tin phụ của sơ đồ BWE được dẫn tới mức đến dưới 1kbit/giây, mà sẽ cho phép sự chấp nhận của nó thậm chí trong mã hóa tốc độ bit rất thấp.

Các phương thức tiếp cận BWE đa dạng được đưa ra dẫn chứng tài liệu trong những năm gần đây theo các tài liệu [từ 1 đến 10]. Nói chung, tất cả các phương thức tiếp cận này hoặc là mù hoàn toàn hoặc được dẫn hoàn toàn tại điểm hoạt động xác

định, bắt chước các đặc điểm tức thời của tín hiệu đầu vào. Hơn nữa, nhiều hệ thống BWE mà theo các tài liệu [1, 3, 4, 5, 9, 10] được tối ưu hóa cụ thể cho các tín hiệu tiếng nói thay vì cho âm nhạc và có thể do đó có thể mang lại các kết quả không thỏa đáng cho âm nhạc. Đây cũng là trường hợp cho hệ thống được mô tả trong tài liệu [13]. Cuối cùng, hầu hết những thực hiện BWR là tương đối phức tạp về tính toán, sử dụng các biến đổi Fourier, các tính toán bộ lọc LPC, hoặc lượng tử hóa vector của thông tin phụ (mã hóa vector dự báo trong MPEG-D USAC theo tài liệu [8]). Điều này có thể là một bất lợi trong việc áp dụng công nghệ mã hóa mới trong thị trường viễn thông di động, do phần lớn các thiết bị di động cung cấp năng lượng tính toán và dung lượng pin rất hạn chế.

Một phương thức tiếp cận mà mở rộng BWE mà bằng thông tin phụ nhỏ được trình bày trong tài liệu [12] và được minh họa trên Fig.16. Tuy nhiên, thông tin phụ "m" bị giới hạn trong việc truyền đường bao quang phổ của phạm vi tần số mở rộng bằng thông.

Vấn đề khác của quy trình được minh họa trên Fig.16 là cách ước lượng đường bao rất phức tạp bằng cách một mặt sử dụng đặc tính băng thấp và mặt khác sử dụng thông tin phụ đường bao bổ sung. Cả hai yếu tố đầu vào, tức là, đặc tính băng thấp và đường bao băng cao bổ sung ảnh hưởng đến mô hình thống kê. Điều này dẫn đến việc thực hiện phía bộ giải mã phức tạp, đặc biệt là vấn đề đối với các thiết bị di động do mức tiêu thụ năng lượng tăng. Hơn nữa, mô hình thống kê thậm chí còn khó cập nhật hơn do thực tế là nó không chỉ bị ảnh hưởng bởi dữ liệu đường bao băng cao bổ sung.

Tài liệu tham khảo

[1] B. Bessette *et al.*, "The Adaptive Multi-rate Wideband Speech Codec (AMR-WB)," *IEEE Trans. on Speech and Audio Processing*, Vol. 10, No. 8, Nov. 2002.

[2] B. Geiser *et al.*, "Bandwidth Extension for Hierarchical Speech and Audio Coding in ITU-T Rec. G.729.1," *IEEE Trans. on Audio, Speech, and Language Processing*, Vol. 15, No. 8, Nov. 2007.

[3] B. Iser, W. Minker, and G. Schmidt, *Bandwidth Extension of Speech Signals*, Springer Lecture Notes in Electrical Engineering, Vol. 13, New York, 2008.

[4] M. Jelínek and R. Salami, “Wideband Speech Coding Advances in VMR-WB Standard,” *IEEE Trans. on Audio, Speech, and Language Processing*, Vol. 15, No. 4, May 2007.

[5] I. Katsir, I. Cohen, and D. Malah, “Speech Bandwidth Extension Based on Speech Phonetic Content and Speaker Vocal Tract Shape Estimation,” in *Proc. EUSIPCO 2011*, Barcelona, Spain, Sep. 2011.

[6] E. Larsen and R. M. Aarts, *Audio Bandwidth Extension: Application of Psychoacoustics, Signal Processing and Loudspeaker Design*, Wiley, New York, 2004.

[7] J. Mäkinen *et al.*, “AMR-WB+: A New Audio Coding Standard for 3rd Generation Mobile Audio Services,” in *Proc. ICASSP 2005*, Philadelphia, USA, Mar. 2005.

[8] M. Neuendorf *et al.*, “MPEG Unified Speech and Audio Coding – The ISO/MPEG Standard for High-Efficiency Audio Coding of All Content Types,” in *Proc. 132nd Convention of the AES*, Budapest, Hungary, Apr. 2012. Also to appear in the Journal of the AES, 2013.

[9] H. Pulakka and P. Alku, “Bandwidth Extension of Telephone Speech Using a Neural Network and a Filter Bank Implementation for Highband Mel Spectrum,” *IEEE Trans. on Audio, Speech, and Language Processing*, Vol. 19, No. 7, Sep. 2011.

[10] T. Vaillancourt *et al.*, “ITU-T EV-VBR: A Robust 8-32 kbit/s Scalable Coder for Error Prone Telecommunications Channels,” in *Proc. EUSIPCO 2008*, Lausanne, Switzerland, Aug. 2008.

[11] L. Miao *et al.*, “G.711.1 Annex D and G.722 Annex B: New ITU-T Superwideband codecs,” in *Proc. ICASSP 2011*, Prague, Czech Republic, May 2011.

[12] Bernd Geiser, Peter Jax, and Peter Vary:: “ROBUST WIDEBAND ENHANCEMENT OF SPEECH BY COMBINED CODING AND ARTIFICIAL

BANDWIDTH EXTENSION”, Proceedings of International Workshop on Acoustic Echo and Noise Control (IWAENC), 2005

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là đề xuất khái niệm cải tiến về mã hóa/giải mã âm thanh.

Mục đích này đạt được nhờ bộ giải mã phù hợp với điểm 1, bộ mã hóa phù hợp với điểm 12, phương pháp tạo tín hiệu âm thanh được nâng cao tần số theo điểm 14, phương pháp tạo tín hiệu mã hóa theo điểm 15, vật ghi có thể đọc được bằng máy tính chứa chương trình máy tính theo điểm 16 và/hoặc điểm 17.

Sáng chế dựa trên sự phát hiện rằng để giảm hơn nữa lượng thông tin phụ, và ngoài ra, để làm cho toàn bộ bộ mã hóa/bộ giải mã không quá phức tạp, việc mã hóa tham số theo giải pháp kỹ thuật đã biết của phần băng cao đã được thay thế hoặc ít nhất là được nâng cao bởi thông tin phụ lựa chọn thực sự liên quan đến mô hình thống kê được sử dụng cùng với bộ trích xuất đặc tính trên bộ giải mã nâng cao tần số. Do thực tế rằng việc trích xuất đặc tính kết hợp với mô hình thống kê cung cấp các lựa chọn biểu diễn theo tham số mà có các phần không rõ ràng đặc biệt cho các phần tiếng nói nhất định, các tác giả sáng chế đã phát hiện ra rằng việc điều khiển thực tế mô hình thống kê trong bộ tạo tham số trên phía bộ giải mã, các lựa chọn được cung cấp sẽ là lựa chọn tốt nhất, vượt trội hơn so với mã hóa theo tham số thực tế đặc điểm nhất định của tín hiệu đặc biệt trong các ứng dụng tốc độ bit thấp tại đó thông tin phụ để mở rộng băng thông bị giới hạn.

Do đó, BWE mù được cải thiện, trong đó khai thác mô hình nguồn cho tín hiệu được mã hóa, bằng việc mở rộng với thông tin phụ bổ sung nhỏ, đặc biệt nếu bản thân tín hiệu không cho phép khôi phục nội dung HF tại mức chất lượng theo cảm giác có thể chấp nhận được. Do đó quy trình này kết hợp các tham số của mô hình nguồn, mà được tạo ra từ nội dung bộ mã hóa lõi được mã hóa, bằng thông tin bổ sung. Điều này đặc biệt có lợi để nâng cao chất lượng cảm giác của các âm thanh khó mã hóa trong

mô hình nguồn này. Những âm thanh như vậy thường thể hiện mối tương quan thấp giữa nội dung HF và LF.

Sáng chế đề cập đến những vấn đề BWE thông thường trong mã hóa âm thanh tốc độ bit rất thấp và những thiếu sót của các kỹ thuật BWE hiện có thuộc tính trạng kỹ thuật. Giải pháp cho lưỡng đề dilem chất lượng được mô tả ở trên được cung cấp bằng cách đề xuất BWE được dẫn tiêu biểu như sự kết hợp tín hiệu thích ứng của BWE mù và BWE được dẫn. BWE theo sáng chế bổ sung một số thông tin phụ nhỏ cho tín hiệu mà cho phép sự phân biệt rõ hơn các âm được mã hóa khó giải quyết khác. Trong mã hóa tiếng nói, điều này áp dụng đặc biệt cho âm xuýt hoặc phụ âm sát.

Các tác giả sáng chế đã phát hiện ra rằng, trong bộ mã hóa-giải mã WB, đường bao quang phổ của vùng HF trên vùng mã hóa lõi biểu diễn hầu hết dữ liệu tiêu chuẩn cần thiết để thực hiện BWE với chất lượng cảm giác có thể chấp nhận được. Tất cả các tham số khác, chẳng hạn như cấu trúc tinh quang phổ và đường bao theo thời gian, có thể thường được suy ra từ tín hiệu lõi được giải mã một cách khá chính xác hoặc là ít có tầm quan trọng về cảm giác. Tuy nhiên, phụ âm sát thường thiếu sự tái tạo thích đáng trong tín hiệu BWE. Thông tin phụ do đó có thể bao gồm thông tin bổ sung để phân biệt giữa các âm xuýt hoặc các phụ âm sát khác nhau như "f", "s", "ch" và "sh".

Thông tin âm thanh khó giải quyết khác cho mở rộng băng thông, khi có xuất hiện các âm bật hoặc âm tắc sát như "t" hoặc "tsch".

Sáng chế cho phép để chỉ sử dụng thông tin phụ này và thực sự truyền thông tin phụ này khi cần thiết và không truyền thông tin phụ này khi không có sự không rõ ràng được mong đợi trong mô hình thống kê.

Hơn nữa, các phương án được ưu tiên của sáng chế chỉ sử dụng lượng rất nhỏ thông tin phụ như ba bit hoặc ít hơn số bit trên mỗi khung, sự phát hiện hoạt động giọng nói/phát hiện tiếng nói/phát hiện không tiếng nói được kết hợp để điều khiển bộ ước lượng tín hiệu, các mô hình thống kê khác nhau được xác định bởi bộ phân lớp tín hiệu hoặc các lựa chọn biểu diễn theo tham số không chỉ đề cập đến ước lượng đường bao mà còn đề cập đến các công cụ mở rộng băng thông khác hoặc cải thiện các tham

số mở rộng băng thông hoặc bổ sung các tham số mới cho các tham số mở rộng băng thông hiện có và các tham số mở rộng băng thông được truyền thực tế.

Mô tả vắn tắt các hình vẽ

Phương án được ưu tiên của sáng chế được thảo luận sau đây trong phạm vi của các hình vẽ đi kèm và cũng được đặt ra trong các yêu cầu bảo hộ phụ thuộc.

Fig.1 minh họa bộ giải mã để tạo ra tín hiệu âm thanh được nâng cao tần số;

Fig.2 minh họa việc thực hiện được ưu tiên trong phạm vi của bộ trích xuất thông tin phụ của Fig.1;

Fig.3 minh họa bảng liên quan đến số bit của thông tin phụ lựa chọn liên quan đến số lượng lựa chọn biểu diễn theo tham số;

Fig.4 minh họa quy trình được ưu tiên được thực hiện trong bộ tạo tham số;

Fig.5 minh họa việc thực hiện được ưu tiên của bộ ước lượng tín hiệu được điều khiển bởi bộ phát hiện độ hoạt động giọng nói hoặc bộ phát hiện tiếng nói/không tiếng nói;

Fig.6 minh họa việc thực hiện được ưu tiên của bộ tạo tham số được điều khiển bởi bộ phân lớp tín hiệu;

Fig.7 minh họa ví dụ cho kết quả của mô hình thống kê và thông tin phụ lựa chọn được liên kết;

Fig.8 minh họa ví dụ tín hiệu được mã hóa bao gồm tín hiệu lỗi được mã hóa và thông tin phụ được liên kết;

Fig.9 minh họa sơ đồ xử lý tín hiệu mở rộng băng thông cho việc cải thiện ước lượng đường bao;

Fig.10 minh họa việc thực hiện khác của bộ giải mã trong phạm vi của các thủ tục lặp lại dải quang phổ;

Fig.11 minh họa phương án khác của bộ giải mã trong phạm vi của thông tin phụ được truyền một cách bổ sung;

Fig.12 minh họa phương án của bộ mã hóa tạo ra tín hiệu được mã hóa;

Fig.13 minh họa việc thực hiện của bộ tạo thông tin phụ lựa chọn của Fig.12;

Fig.14 minh họa việc thực hiện khác của bộ tạo thông tin phụ lựa chọn của Fig.12;

Fig.15 minh họa thuật toán mở rộng băng thông độc lập theo giải pháp kỹ thuật đã biết; và

Fig.16 minh họa tổng quan hệ thống truyền dẫn với tin bổ sung.

Mô tả chi tiết sáng chế

Fig.1 minh họa bộ giải mã để tạo ra tín hiệu âm thanh được nâng cao tần số 120. Bộ giải mã bao gồm bộ trích xuất đặc tính 104 để trích xuất ra (ít nhất) đặc tính từ tín hiệu lõi 100. Thông thường, bộ trích xuất đặc tính có thể trích xuất đặc tính đơn lẻ hoặc nhiều đặc tính, tức là hai hoặc nhiều hơn hai các đặc tính, và thậm chí được ưu tiên rằng nhiều đặc tính được trích xuất bởi bộ trích xuất đặc tính. Điều đó áp dụng không chỉ cho bộ trích xuất đặc tính trong bộ giải mã mà còn áp dụng cho bộ trích xuất đặc tính trong bộ mã hóa.

Hơn nữa, bộ trích xuất thông tin phụ 110 để trích xuất thông tin phụ lựa chọn 114 được liên kết với tín hiệu lõi 100 được cung cấp. Ngoài ra, bộ tạo tham số 108 được kết nối với bộ trích xuất đặc tính 104 thông qua đường truyền 112 và với bộ trích xuất thông tin phụ 110 thông qua thông tin phụ lựa chọn 114. Bộ tạo tham số 108 được tạo cấu hình để tạo ra sự biểu diễn theo tham số để ước lượng vùng quang phổ của tín hiệu âm thanh được nâng cao tần số không được xác định bởi tín hiệu lõi. Bộ tạo tham số 108 được tạo cấu hình để cung cấp số lượng lựa chọn biểu diễn theo tham số đáp ứng với các đặc tính 112 và lựa chọn một trong số các lựa chọn biểu diễn theo tham số như biểu diễn theo tham số đáp ứng với thông tin phụ lựa chọn 114. Bộ giải mã còn bao gồm bộ ước lượng tín hiệu 118 để ước lượng tín hiệu âm thanh được nâng cao tần số sử dụng biểu diễn theo tham số được lựa chọn bởi bộ lựa chọn, tức là, biểu diễn theo tham số 116.

Cụ thể, bộ trích xuất đặc tính 104 có thể được thực hiện để có thể trích xuất từ tín hiệu lõi được giải mã như được minh họa trên Fig.2. Sau đó, giao diện đầu vào 110 được tạo cấu hình để nhận tín hiệu đầu vào được mã hóa 200. Tín hiệu đầu vào được

mã hóa 200 này là đầu vào vào giao diện 110 và giao diện đầu vào 110 sau đó trích xuất thông tin phụ lựa chọn từ tín hiệu lỗi được mã hóa. Do đó, giao diện đầu vào 110 hoạt động như bộ trích xuất thông tin phụ 110 trên Fig.1. Tín hiệu lỗi được mã hóa 201 được xuất ra bởi giao diện đầu vào 110 sau đó là đầu vào vào bộ giải mã lỗi 124 để cung cấp tín hiệu lỗi được giải mã mà có thể là tín hiệu lỗi 100.

Tuy nhiên, ngoài ra, bộ trích xuất đặc tính có thể cũng hoạt động hoặc trích xuất đặc tính từ tín hiệu lỗi được mã hóa. Thông thường, tín hiệu lỗi được mã hóa bao gồm sự biểu diễn các hệ số tỉ lệ cho các dải tần số hoặc sự biểu diễn khác bất kỳ của thông tin âm thanh. Phụ thuộc vào loại trích xuất đặc tính, sự biểu diễn được mã hóa của tín hiệu âm thanh là sự biểu diễn cho tín hiệu lỗi được giải mã và, do đó các đặc tính có thể được trích xuất. Bằng cách thay thế hoặc bổ sung, đặc tính có thể được trích xuất không chỉ từ tín hiệu lỗi được giải mã hoàn toàn mà còn từ tín hiệu lỗi được giải mã một phần. Trong mã hóa miền tần số, tín hiệu được mã hóa đang biểu diễn biên độ miền tần số bao gồm chuỗi các khung quang phổ. Do đó, tín hiệu lỗi được mã hóa có thể chỉ được giải mã một phần để thu được sự biểu diễn được giải mã của chuỗi các khung quang phổ, trước khi thực hiện thực tế chuyển đổi phổ-thời gian. Do đó, bộ trích xuất đặc tính 104 có thể trích xuất các đặc tính hoặc từ tín hiệu lỗi được mã hóa hoặc tín hiệu lỗi được giải mã một phần hoặc tín hiệu lỗi được giải mã hoàn toàn. Bộ trích xuất đặc tính 104 có thể được thực hiện, đối với các đặc tính được trích xuất như đã biết trong kỹ thuật và bộ trích xuất đặc tính có thể, ví dụ, được thực hiện như trong đầu tay âm thanh hoặc các kỹ thuật ID âm thanh.

Tốt hơn là, thông tin phụ lựa chọn 114 bao gồm số N các bit mỗi khung của tín hiệu lỗi. Fig.3 minh họa bảng cho các lựa chọn khác nhau. Số bit cho thông tin phụ lựa chọn hoặc là được cố định hoặc được lựa chọn phụ thuộc vào số thay số biểu diễn theo tham số được cung cấp bởi mô hình thống kê đáp ứng với đặc tính được trích xuất. Một bit của thông tin phụ lựa chọn là đủ khi chỉ hai lựa chọn biểu diễn theo tham số được cung cấp bởi mô hình thống kê đáp ứng với đặc tính. Khi số cực đại của bốn lựa chọn biểu diễn được cung cấp bởi mô hình thống kê, sau đó hai bit là cần thiết cho thông tin phụ lựa chọn. Ba bit thông tin phụ lựa chọn cho phép giá trị cực đại của tám lựa chọn biểu diễn theo tham số đồng thời. Bốn bit thông tin phụ lựa chọn thực tế cho

phép 16 lựa chọn biểu diễn theo tham số và năm bit thông tin phụ lựa chọn cho phép 32 lựa chọn biểu diễn theo tham số đồng thời. Ưu tiên chỉ sử dụng ba hoặc ít hơn ba bit thông tin phụ lựa chọn mỗi khung thu được trong tốc độ thông tin phụ của 150 bit mỗi giây khi một giây được chia thành 50 khung. Tốc độ thông tin phụ này có thể thậm chí được giảm do thực tế thông tin phụ lựa chọn chỉ cần thiết khi mô hình thống kê thực tế cung cấp các lựa chọn biểu diễn. Do đó, khi mô hình thống kê chỉ cung cấp lựa chọn đơn lẻ cho đặc tính, sau đó bit thông tin phụ lựa chọn là không cần thiết ở tất cả. Một mặt, khi mô hình thống kê chỉ cung cấp bốn lựa chọn biểu diễn theo tham số, sau đó chỉ hai bit hơn là ba bit thông tin phụ lựa chọn là cần thiết. Do đó, trong các trường hợp điển hình, tốc độ thông tin phụ bổ sung thậm chí có thể được giảm dưới 150 bit mỗi giây.

Hơn nữa, bộ tạo tham số được tạo cấu hình để cung cấp, tại hầu hết, số lượng lựa chọn biểu diễn theo tham số là bằng 2^N . Mặt khác, khi bộ tạo tham số 108 cung cấp, ví dụ, chỉ năm lựa chọn biểu diễn theo tham số, sau đó nhờ thế mà ba bit của thông tin phụ lựa chọn được yêu cầu.

Fig.4 minh họa việc thực hiện được ưu tiên của bộ tạo tham số 108. Cụ thể, bộ tạo tham số 108 được tạo cấu hình để đặc tính 112 của Fig.1 là đầu vào vào mô hình thống kê như được phác họa trong bước 400. Sau đó như được phác họa trong bước 400, nhiều lựa chọn biểu diễn theo tham số được cung cấp bởi mô hình.

Hơn nữa, bộ tạo tham số 108 được tạo cấu hình để truy lại thông tin phụ lựa chọn 114 từ bộ trích xuất thông tin phụ như được phác họa trong bước 404. Sau đó, trong bước 406, lựa chọn biểu diễn theo tham số cụ thể được lựa chọn sử dụng thông tin phụ lựa chọn 114. Cuối cùng, trong bước 408, lựa chọn biểu diễn theo tham số được lựa chọn là đầu ra đến bộ ước lượng tín hiệu 118.

Tốt hơn, bộ tạo tham số 108 được tạo cấu hình để sử dụng, khi lựa chọn một trong số các lựa chọn biểu diễn theo tham số, thứ tự được xác định trước của các lựa chọn biểu diễn theo tham số hoặc, ngoài ra, thứ tự tín hiệu bộ mã hóa của các lựa chọn biểu diễn. Để đạt được điều này, tham chiếu đến Fig.7. Fig.7 minh họa kết quả của mô hình thống kê cung cấp bốn lựa chọn biểu diễn theo tham số 702, 704, 706, 708. Mã thông tin phụ lựa chọn tương ứng cũng được minh họa. Lựa chọn 702 tương ứng với

mẫu bit 712. Lựa chọn 704 tương ứng với mẫu bit 714. Lựa chọn 706 tương ứng với mẫu bit 716 và lựa chọn 708 tương ứng với mẫu bit 718. Do đó, khi bộ tạo tham số 108 hoặc, ví dụ, bước 402 truy tìm bốn lựa chọn 702 đến 708 theo thứ tự được minh họa trên Fig.7, sau đó thông tin phụ lựa chọn có mẫu bit 716 sẽ nhận biết duy nhất lựa chọn biểu diễn theo tham số 3 (số tham chiếu 706) và bộ tạo tham số 108 sau đó sẽ được lựa chọn lựa chọn thứ ba này. Tuy nhiên, khi mẫu bit thông tin phụ lựa chọn là mẫu bit 712, sau đó lựa chọn thứ nhất 702 sẽ được lựa chọn.

Do đó, thứ tự được xác định trước của các lựa chọn biểu diễn theo tham số có thể là thứ tự trong đó mô hình thống kê phân phối thực tế các lựa chọn đáp ứng với đặc tính được trích xuất. Ngoài ra, nếu lựa chọn riêng được liên kết các xác suất khác nhau mà là, tuy nhiên, khá gần nhau, sau đó thứ tự được xác định trước có thể là thứ tự biểu diễn theo tham số xác suất cao nhất đến thứ nhất và vân vân. Ngoài ra, thứ tự có thể được tín hiệu hóa ví dụ bởi bit đơn lẻ, nhưng thứ tự thậm chí lưu bit này, thứ tự được xác định trước được ưu tiên.

Tiếp theo, tham chiếu đến các hình vẽ từ Fig.9 đến Fig.11.

Trong phương án theo Fig.9, sáng chế đặc biệt thích hợp cho các tín hiệu tiếng nói, như mô hình nguồn tiếng nói dành riêng được khai thác cho việc trích xuất tham số. Tuy nhiên, sáng chế không giới hạn mã hóa tiếng nói. Các phương án khác nhau cũng có thể sử dụng các mô hình nguồn khác.

Cụ thể, thông tin phụ lựa chọn 114 cũng được gọi là "thông tin phụ âm xát", vì thông tin phụ lựa chọn này phân biệt giữa các âm xuyết hoặc các phụ âm xát khó giải quyết chẳng hạn như "f", "s" hoặc "sh". Do đó, thông tin phụ lựa chọn cung cấp định nghĩa rõ ràng của một trong ba lựa chọn khó giải quyết mà, ví dụ, được cung cấp bởi mô hình thống kê 904 trong việc xử lý ước lượng đường bao 902 mà cả hai được thực hiện trong bộ tạo tham số 108. Các kết quả ước lượng đường bao trong biểu diễn theo tham số của đường bao quang phổ của các phần quang phổ không được chứa trong tín hiệu lỗi.

Do đó, khối 104 có thể tương ứng với khối 1510 của Fig.15. Hơn nữa, khối 1530 của Fig.15 có thể tương ứng với mô hình thống kê 904 của Fig.9.

Hơn nữa, ưu tiên bộ ước lượng tín hiệu 118 bao gồm bộ lọc phân tích 910, khối mở rộng kích thích 112 và bộ lọc tổng hợp 940. Do đó, các khối 910, 912, 914 có thể tương ứng với các khối 1600, 1700 và 1800 của Fig.15. Cụ thể, bộ lọc phân tích 910 là bộ lọc phân tích LPC. Khối ước lượng đường bao 902 điều khiển các hệ số bộ lọc của bộ lọc phân tích 910 để kết quả của khối 910 là tín hiệu kích thích bộ lọc. Tín hiệu kích thích bộ lọc này được mở rộng đối với tần số để thu được tín hiệu kích thích tại đầu ra của khối 912 mà không chỉ có vùng tần số của bộ giải mã 120 cho tín hiệu đầu ra mà còn có vùng quang phổ hoặc tần số không được xác định bởi bộ mã hóa lõi và/hoặc vượt quá vùng quang phổ của tín hiệu lõi. Do đó, tín hiệu âm thanh 909 tại đầu ra của bộ giải mã được lấy mẫu lên và được nội suy bởi bộ nội suy 900 và, sau đó tín hiệu được nội suy phải trải qua xử lý trong bộ ước lượng tín hiệu 118. Do đó, bộ nội suy 900 trên Fig.9 có thể tương ứng với bộ nội suy 1500 của Fig.15. Tuy nhiên, tốt hơn, trái với Fig.15, việc trích xuất đặc tính 104 được thực hiện sử dụng tín hiệu không được nội suy hơn là trên tín hiệu được nội suy như được minh họa trên Fig.15. Điều này là lợi thế trong đó bộ trích xuất đặc tính 104 hoạt động hiệu quả hơn do thực tế tín hiệu âm thanh không được nội suy 909 có số mẫu nhỏ hơn so với phần thời gian nhất định của tín hiệu âm thanh so với tín hiệu được nội suy và được lấy mẫu lên tại đầu ra của khối 900.

Fig.10 minh họa phương án khác của sáng chế. Trái ngược với Fig.9, Fig.10 có mô hình thống kê 904 không chỉ cung cấp ước lượng đường bao như trên Fig.9 mà còn cung cấp các biểu diễn theo tham số bổ sung bao gồm thông tin để tạo ra các âm khuyết 1080 hoặc thông tin để lọc nghịch đảo 1040 hoặc thông tin trên sàn nhiễu 1020 được bổ sung. Các khối 1020, 1040, tạo ra đường bao quang phổ 1060 và các âm khuyết 1080 các thủ tục được mô tả trong Tiêu chuẩn MPEG-4 trong phạm vi Mã hóa âm thanh được nâng cao hiệu quả cao (High Efficiency Advanced Audio Coding - HE-AAC).

Do đó, các tín hiệu khác khác tiếng nói có thể cũng được mã hóa như được minh họa trên Fig.10. Trong trường hợp đó, có thể không đủ để chỉ mã hóa đường bao quang phổ 1060, mà còn mã hóa thông tin phụ khác như âm điệu 1040, mức nhiễu âm

1020 hoặc sóng hình sin khuyết 1080 như được hoàn thành trong kỹ thuật lặp lại dải quang phổ (spectral band replication - SBR) được minh họa trong tài liệu [6].

Các phương án khác được minh họa trên Fig.11, tại thông tin phụ 114, tức là, thông tin phụ lựa chọn được sử dụng trong sự bổ sung vào thông tin phụ SBR được minh họa tại 1100. Do đó, thông tin phụ lựa chọn bao gồm, ví dụ, thông tin liên quan đến các âm tiếng nói được phát hiện được bổ sung vào thông tin phụ SBR để lại 1100. Điều đó giúp tái tạo một cách thực tế hơn nội dung tần số cao cho âm tiếng nói như các âm xuýt chứa các phụ âm xát, các âm bật hoặc các nguyên âm. Do đó, thủ tục được minh họa trên Fig.11 có ưu điểm là thông tin phụ lựa chọn được truyền một cách bổ sung 114 hỗ trợ phân lớp (âm vị) phía bộ giải mã theo thứ tự để cung cấp sự thích ứng phía bộ giải mã của các tham số (mở rộng băng thông) SBR hoặc BWE. Do đó, trái với Fig.10, phương án Fig.11 ngoài việc cung cấp thông tin phụ lựa chọn còn cung cấp thông tin phụ SBR để lại.

Fig.8 minh họa việc biểu diễn ví dụ của tín hiệu đầu vào được mã hóa. Tín hiệu đầu vào được mã hóa chứa các khung tiếp theo 800, 806, 812. Mỗi khung có tín hiệu lõi được mã hóa. Ví dụ, khung 800 có tiếng nói như tín hiệu lõi được mã hóa. Khung 806 có âm nhạc như tín hiệu lõi được mã hóa và khung 812 một lần nữa có tiếng nói như tín hiệu lõi được mã hóa. Khung 800 có, ví dụ, như thông tin phụ chỉ có thông tin phụ lựa chọn nhưng không có thông tin phụ SBR. Do đó, khung 800 tương ứng với Fig.9 hoặc Fig.10. Ví dụ, khung 806 bao gồm thông tin SBR nhưng không chứa bất kỳ thông tin phụ lựa chọn nào. Hơn nữa, khung 812 bao gồm tín hiệu tiếng nói được mã hóa và, trái lại với khung 800, khung 812 không chứa bất kỳ thông tin phụ lựa chọn nào. Điều này là do thực tế rằng thông tin phụ lựa chọn là không cần thiết, vì bất kỳ sự không rõ ràng trong xử lý mô hình trích xuất/thống kê đặc tính không được tìm thấy trên phía bộ mã hóa.

Tiếp theo, Fig.5 được mô tả. Bộ phát hiện hoạt động giọng nói hoặc bộ phát hiện tiếng nói/không tiếng nói 500 hoạt động trên tín hiệu lõi được sử dụng để quyết định, hoặc bằng thông sáng chế hoặc kỹ thuật nâng cao tần số cần được sử dụng hoặc kỹ thuật mở rộng băng thông khác. Do đó, khi bộ phát hiện hoạt động giọng nói hoặc bộ phát hiện tiếng nói/không tiếng nói phát hiện giọng nói hoặc tiếng nói, sau đó kỹ

thuật mở rộng băng thông thứ nhất BW EXT.1 được minh họa tại 511 được sử dụng mà hoạt động, ví dụ như được thảo luận trên các hình vẽ Fig.1, 9, 10, 11. Do đó, bộ chuyển mạch 502, 504 được thiết lập theo cách mà các tham số từ bộ tạo tham số từ đầu vào 512 được giữ và bộ chuyển mạch 504 kết nối các tham số này tới khối 511. Tuy nhiên, khi trạng thái được phát hiện bởi bộ phát hiện 500 mà không thể hiện bất kỳ tín hiệu tiếng nói nào nhưng, ví dụ, thể hiện tín hiệu âm nhạc, sau đó các tham số mở rộng băng thông 514 từ dòng bit tốt hơn là đầu vào vào thủ tục kỹ thuật mở rộng băng thông khác 513. Do đó, bộ phát hiện 500 phát hiện, hoặc kỹ thuật mở rộng băng thông theo sáng chế 511 sẽ được sử dụng hoặc không. Với tín hiệu không tiếng nói, bộ mã hóa có thể chuyển từ kỹ thuật mở rộng băng thông khác được minh họa bởi khối 513 chẳng hạn được đề cập trong các tài liệu [6, 8]. Do đó, bộ ước lượng tín hiệu 118 của Fig.5 được tạo cấu hình để chuyển qua thủ tục mở rộng băng thông khác và/hoặc để sử dụng các tham số khác được trích xuất từ tín hiệu được mã hóa, khi bộ phát hiện 500 phát hiện hoạt động không giọng nói hoặc tín hiệu không tiếng nói. Với kỹ thuật mở rộng băng thông khác 513 này, thông tin phụ lựa chọn tốt hơn là không có mặt trong dòng bit và cũng không được sử dụng mà được ký hiệu trên Fig.15 bằng cách tắt bộ chuyển mạch 502 tới đầu vào 514.

Fig.6 minh họa việc thực hiện khác của bộ tạo tham số 108. Bộ tạo tham số 108 tốt hơn là có nhiều mô hình thống kê chẳng hạn như mô hình thống kê thứ nhất 600 và mô hình thống kê thứ hai 602. Hơn nữa, bộ lựa chọn 604 được cung cấp mà được điều khiển bởi thông tin phụ lựa chọn để cung cấp lựa chọn biểu diễn theo tham số chính xác. Mà mô hình thống kê là hoạt động được điều khiển bởi bộ phân lớp tín hiệu bổ sung 606 nhận, tại đầu vào của nó, tín hiệu lỗi, tức là tín hiệu tương tự như đầu vào vào bộ trích xuất đặc tính 104. Do đó, mô hình thống kê trên Fig.10 hoặc các hình vẽ khác bất kỳ có thể thay đổi với nội dung được mã hóa. Với tiếng nói, mô hình thống kê mà biểu diễn mô hình nguồn tạo ra tiếng nói được sử dụng, trong khi với tín hiệu khác chẳng hạn như tín hiệu âm nhạc như, ví dụ, được phân lớp bởi bộ phân lớp tín hiệu 606 mô hình khác được sử dụng mà được đào tạo trên bộ dữ liệu âm nhạc rộng. Các mô hình thống kê khác hữu ích theo cách bổ sung cho các ngôn ngữ khác nhau v.v.

Như được thảo luận trước, Fig.7 minh họa nhiều lựa chọn như thu được bởi mô hình thống kê chẳng hạn như mô hình thống kê 600. Do đó, đầu ra của khối 600, ví dụ, cho các lựa chọn khác nhau như được minh họa tại đường song song 605. Theo cách tương tự, mô hình thống kê thứ hai 602 có thể cũng xuất ra nhiều lựa chọn chẳng hạn như các lựa chọn được minh họa tại đường 606. Phụ thuộc vào mô hình thống kê cụ thể, ưu tiên rằng chỉ các lựa chọn có xác suất hoàn toàn cao đối với bộ trích xuất đặc tính 104 là đầu ra. Do đó, mô hình thống kê cung cấp, đáp ứng với đặc tính, nhiều biểu diễn theo tham số lựa chọn, trong đó mỗi biểu diễn theo tham số lựa chọn có xác suất là đồng nhất với các xác suất của các biểu diễn theo tham số lựa chọn khác hoặc khác với các xác suất của các biểu diễn theo tham số lựa chọn khác nhỏ hơn 10%. Do đó, trong phương án, chỉ biểu diễn theo tham số có xác suất cao nhất và số biểu diễn theo tham số lựa chọn khác mà tất cả có xác suất chỉ là 10% nhỏ hơn xác suất của lựa chọn khớp nhất được xuất ra.

Fig.12 minh họa bộ mã hóa để tạo ra tín hiệu được mã hóa 1212. Bộ mã hóa bao gồm bộ mã hóa lõi 1200 để mã hóa tín hiệu ban đầu 1206 để thu được tín hiệu âm thanh lõi được mã hóa 1208 có thông tin trên số nhỏ hơn của dải tần số so với tín hiệu ban đầu 1206. Hơn nữa, bộ tạo thông tin phụ lựa chọn 1202 để tạo ra thông tin phụ lựa chọn 1210 (thông tin phụ lựa chọn SSI) được cung cấp. Thông tin phụ lựa chọn 1210 chỉ ra lựa chọn biểu diễn theo tham số được xác định được cung cấp bởi mô hình thống kê đáp ứng với đặc tính được trích xuất từ tín hiệu ban đầu 1206 hoặc từ tín hiệu âm thanh được mã hóa 1208 hoặc từ bản được giải mã của tín hiệu âm thanh được giải mã. Hơn nữa, bộ mã hóa bao gồm giao diện đầu ra 1204 để xuất ra tín hiệu được mã hóa 1212. Tín hiệu được mã hóa 1212 bao gồm tín hiệu âm thanh được mã hóa 1208 và thông tin phụ lựa chọn 1210. Tốt hơn là, bộ tạo thông tin phụ lựa chọn 1202 được thực hiện như được minh họa trên Fig.13. Để đạt được điều này, bộ tạo thông tin phụ lựa chọn 1202 bao gồm bộ giải mã lõi 1300. Bộ trích xuất đặc tính 1302 được cung cấp mà hoạt động trên đầu ra tín hiệu lõi được giải mã bởi khối 1300. Đặc tính là đầu vào vào bộ xử lý mô hình thống kê 1304 để tạo ra số lựa chọn biểu diễn theo tham số để ước lượng vùng quang phổ của tín hiệu được nâng cao tần số không được xác định bởi đầu ra tín hiệu lõi được giải mã bởi khối 1300. Các lựa chọn biểu diễn theo tham số 1305 này là tất cả đầu vào vào bộ ước lượng tín hiệu 1306 để ước lượng tín hiệu âm

thanh được nâng cao tần số 1307. Các tín hiệu âm thanh được nâng cao tần số được ước lượng 1307 này sau đó là đầu vào vào bộ so sánh 1308 để so sánh các tín hiệu âm thanh được nâng cao tần số 1307 với tín hiệu ban đầu 1206 của Fig.12. Bộ tạo thông tin phụ lựa chọn 1202 được tạo cấu hình bổ sung để thiết lập thông tin phụ lựa chọn 1210 sao cho thông tin phụ lựa chọn xác định duy nhất lựa chọn biểu diễn theo tham số thu được trong tín hiệu âm thanh được nâng cao tần số khớp nhất với tín hiệu ban đầu dưới tiêu chuẩn tối ưu hóa. Tiêu chuẩn tối ưu hóa có thể là sai số bình phương trung bình cực tiểu (minimum means squared error - MMSE) dựa trên tiêu chuẩn, tiêu chuẩn cực tiểu hóa sự chênh lệch theo mẫu hoặc tốt hơn tiêu chuẩn tâm thính học cực tiểu hóa biến dạng nhận thấy được hoặc bất kỳ tiêu chuẩn tối ưu hóa khác đã biết với người có trình độ trong lĩnh vực kỹ thuật.

Trong khi Fig.13 minh họa chu trình đóng hoặc thủ tục phân tích bằng tổng hợp, Fig.14 minh họa việc thực hiện lựa chọn của thông tin phụ lựa chọn 1202 đơn giản hơn thủ tục chu trình mở. Trong phương án trên Fig.14, tín hiệu ban đầu 1206 bao gồm siêu thông tin được liên kết cho bộ tạo thông tin phụ lựa chọn 1202 mô tả chuỗi thông tin âm thanh (ví dụ các chú thích) cho các chuỗi mẫu của tín hiệu âm thanh ban đầu. Bộ tạo thông tin phụ lựa chọn 1202 bao gồm, trong phương án này, bộ trích xuất siêu dữ liệu 1400 để trích xuất chuỗi của siêu thông tin và, thêm vào đó, bộ dịch siêu dữ liệu, thường nhận ra trên mô hình thống kê được sử dụng trên phía bộ giải mã để dịch chuỗi của siêu thông tin thành chuỗi thông tin phụ lựa chọn 1210 được liên kết với tín hiệu âm thanh ban đầu. Siêu dữ liệu được trích xuất bởi bộ trích xuất siêu dữ liệu 1400 được loại bỏ trong bộ mã hóa và không được truyền trong tín hiệu được mã hóa 1212. Thay vào đó, thông tin phụ lựa chọn 1210 được truyền trong tín hiệu được mã hóa cùng với tín hiệu âm thanh được mã hóa 1208 được tạo bởi bộ mã hóa lõi mà có nội dung tần số khác nhau và, thông thường, nội dung tần số nhỏ hơn được so sánh với tín hiệu được giải mã được tạo ra cuối cùng hoặc được so sánh với tín hiệu ban đầu 1206.

Thông tin phụ lựa chọn 1210 được tạo ra bởi bộ tạo thông tin phụ lựa chọn 1202 có thể có bất kỳ đặc điểm được thảo luận trong phạm vi của các hình vẽ trước.

Mặc dù sáng chế được mô tả trong phạm vi của các sơ đồ khối tại đó các khối biểu diễn các thành phần phần cứng thực tế hoặc theo logic, sáng chế cũng có thể được thực hiện bởi phương pháp được thực hiện bởi máy tính. Trong các trường hợp sau, các khối biểu diễn các bước phương pháp tương ứng trong đó các bước này thay cho các chức năng được thực hiện bởi các khối theo logic hoặc các khối phần cứng vật chất tương ứng.

Mặc dù một vài khía cạnh đã được mô tả trong phạm vi của thiết bị, rõ ràng rằng các khía cạnh này cũng biểu diễn sự mô tả của phương pháp tương ứng, mà khối hoặc thiết bị tương ứng với bước phương pháp hoặc đặc tính của bước phương pháp. Tương tự, các khía cạnh được mô tả trong phạm vi của bước phương pháp cũng biểu diễn sự mô tả của khối hoặc mục hoặc đặc tính tương ứng của thiết bị tương ứng. Một vài hoặc tất cả các bước phương pháp có thể được thực hiện bởi (hoặc sử dụng) thiết bị phần cứng, tương tự ví dụ, bộ vi xử lý, máy tính có thể lập trình hoặc mạch điện tử. Theo một số phương án, một số một hoặc nhiều bước phương pháp quan trọng nhất có thể được thực hiện bởi thiết bị như vậy.

Tín hiệu được truyền hoặc được mã hóa theo sáng chế có thể được lưu trữ trên vật ghi lưu trữ dạng số hoặc có thể được truyền trên môi trường truyền dẫn chẳng hạn như môi trường truyền dẫn không dây hoặc môi trường truyền dẫn có dây chẳng hạn như liên mạng.

Phụ thuộc vào các yêu cầu thực hiện nhất định, các phương án của sáng chế có thể được thực hiện trong phần cứng hoặc trong phần mềm. Việc thực hiện có thể được thực hiện sử dụng vật ghi lưu trữ dạng số, ví dụ, đĩa mềm, đĩa DVD, Blu-Ray, đĩa CD, ROM, PROM, và EPROM, EEPROM hoặc bộ nhớ FLASH, có tín hiệu điều khiển có thể đọc được bằng điện tử được lưu trữ trên đó, trong đó kết hợp (hoặc có khả năng kết hợp) với hệ thống máy tính có thể lập trình sao cho các phương pháp tương ứng được thực hiện. Do đó, vật ghi lưu trữ dạng số có thể là máy tính có thể đọc được.

Một số phương án theo sáng chế bao gồm vật mang dữ liệu có các tín hiệu điều khiển có thể đọc được nhờ điện tử, mà có khả năng kết hợp với hệ thống máy tính có thể lập trình được, để một trong các phương pháp được mô tả ở đây được thực hiện.

Thông thường, các phương án của sáng chế có thể được thực hiện như sản phẩm chương trình máy tính với mã chương trình, mã chương trình hoạt động để thực hiện một trong các phương pháp khi sản phẩm của chương trình máy tính chạy trên máy tính. Ví dụ, mã chương trình có thể được lưu trữ trên vật mang có thể đọc được bằng máy.

Các phương án khác bao gồm chương trình máy tính để thực hiện một trong các phương pháp được mô tả ở đây, được lưu trữ trên vật mang có thể đọc được bằng máy.

Nói cách khác, do đó, phương án của phương pháp theo sáng chế là, chương trình máy tính có mã chương trình để thực hiện một trong các phương pháp được mô tả ở đây, khi chương trình máy tính chạy trên máy tính.

Do đó, phương án khác của sáng chế là, vật mang dữ liệu (hoặc vật ghi lưu trữ không tạm thời chẳng hạn như vật ghi lưu trữ dạng số, hoặc vật ghi có thể đọc được bằng máy tính) bao gồm chương trình máy tính được ghi trên đó để thực hiện một trong các phương pháp được mô tả trong sáng chế. Vật mang dữ liệu, vật ghi lưu trữ dạng số hoặc vật ghi được ghi dạng hữu hình điển hình và/hoặc không tạm thời.

Do đó, phương án nữa của phương pháp theo sáng chế là, dòng dữ liệu hoặc chuỗi các tín hiệu biểu diễn chương trình máy tính để thực hiện một trong các phương pháp được mô tả ở đây. Ví dụ, dòng dữ liệu hoặc chuỗi các tín hiệu được tạo cấu hình để được chuyển thông qua sự kết nối truyền thông dữ liệu, ví dụ, thông qua liên mạng.

Phương án nữa bao gồm phương tiện xử lý, ví dụ, máy tính, hoặc thiết bị logic có thể lập trình được, được tạo cấu hình để, hoặc được làm thích ứng để, thực hiện một trong các phương pháp được mô tả ở đây.

Phương án nữa bao gồm máy tính được cài đặt trên đó chương trình máy tính để thực hiện một trong các phương pháp được mô tả ở đây.

Phương án nữa theo sáng chế bao gồm thiết bị hoặc hệ thống được tạo cấu hình để chuyển (ví dụ, bằng điện tử hoặc quang học) chương trình máy tính để thực hiện một trong các phương pháp được mô tả ở đây đến bộ thu. Ví dụ, bộ thu có thể là máy tính, thiết bị di động, thiết bị nhớ hoặc tương tự. Ví dụ, thiết bị hoặc hệ thống có thể bao gồm máy chủ để chuyển chương trình máy tính đến bộ thu.

Trong một số phương án, thiết bị logic có thể lập trình được (ví dụ, mảng cổng có thể lập trình được dạng trường) có thể được sử dụng để thực hiện một số hoặc tất cả các chức năng của các phương pháp được mô tả ở đây. Trong một số phương án, mảng cổng lập trình được dạng trường có thể kết hợp với bộ vi xử lý để thực hiện một trong các phương pháp được mô tả ở đây. Thông thường, các phương pháp tốt hơn là được thực hiện bởi thiết bị phân cứng bất kỳ.

Các phương án được mô tả ở trên chỉ mang tính minh họa cho các nguyên lý của sáng chế. Cần được hiểu rằng các sự biến thể và biến đổi của các phương án và các chi tiết được mô tả ở đây sẽ là rõ ràng đối với người có hiểu biết trung bình trong lĩnh vực kỹ thuật tương ứng. Do đó, mục đích của sáng chế chỉ bị giới hạn bởi phạm vi của các điểm yêu cầu bảo hộ sau đây và không bị giới hạn bởi các chi tiết cụ thể được thể hiện bằng cách mô tả và giải thích của các phương án ở đây.

YÊU CẦU BẢO HỘ

1. Bộ giải mã tạo ra tín hiệu âm thanh được nâng cao tần số (120), bao gồm:

bộ trích xuất đặc tính (104) để trích xuất ra đặc tính từ tín hiệu lỗi (100);

bộ trích xuất thông tin phụ (110) để trích xuất ra thông tin phụ lựa chọn được liên kết với tín hiệu lỗi;

bộ tạo tham số (108) để tạo ra sự biểu diễn theo tham số để ước lượng vùng phổ của tín hiệu âm thanh được nâng cao tần số (120) không được xác định bởi tín hiệu lỗi (100), trong đó bộ tạo tham số (108) được tạo cấu hình để cung cấp số lượng lựa chọn biểu diễn theo tham số (702, 704, 706, 708) đáp ứng với đặc tính (112), và trong đó bộ tạo tham số (108) được tạo cấu hình để lựa chọn một trong số các lựa chọn biểu diễn theo tham số như biểu diễn theo tham số đáp ứng với thông tin phụ lựa chọn (712 đến 718);

bộ ước lượng tín hiệu (118) để ước lượng tín hiệu âm thanh được nâng cao tần số (120) sử dụng biểu diễn theo tham số được lựa chọn; và

bộ phân lớp tín hiệu (606) để phân lớp khung tín hiệu lỗi (100),

trong đó bộ tạo tham số (108) được tạo cấu hình để sử dụng mô hình thống kê thứ nhất (600), khi khung tín hiệu được phân lớp thuộc về lớp thứ nhất của các tín hiệu và để sử dụng mô hình thống kê khác thứ hai (602), khi khung được phân lớp thành lớp khác thứ hai của các tín hiệu.

2. Bộ giải mã theo điểm 1, còn bao gồm:

giao diện đầu vào (110) để nhận tín hiệu đầu vào được mã hóa (200) bao gồm tín hiệu lỗi được mã hóa (201) và thông tin phụ lựa chọn (114); và

bộ giải mã lỗi (124) để giải mã tín hiệu lỗi được mã hóa để thu được tín hiệu lỗi (100).

3. Bộ giải mã theo một điểm trong số các điểm trên, trong đó bộ tạo tham số (108) được tạo cấu hình để sử dụng, khi lựa chọn một trong số các lựa chọn biểu diễn theo

tham số, thứ tự được xác định trước của các lựa chọn biểu diễn theo tham số hoặc thứ tự bộ giải mã được tín hiệu hóa của các lựa chọn biểu diễn theo tham số.

4. Bộ giải mã theo một điểm trong số các điểm trên, trong đó bộ tạo tham số (108) được tạo cấu hình để cung cấp sự biểu diễn đường bao như sự biểu diễn theo tham số,

trong đó thông tin phụ lựa chọn (114) chỉ ra một trong số nhiều âm xuýt hoặc phụ âm xát khác nhau, và

trong đó bộ tạo tham số (108) được tạo cấu hình để cung cấp biểu diễn đường bao được chỉ ra bởi thông tin phụ lựa chọn.

5. Bộ giải mã theo một điểm trong số các điểm trên,

trong đó bộ ước lượng tín hiệu (118) bao gồm bộ nội suy (900) để nội suy tín hiệu lỗi (100), và

trong đó bộ trích xuất đặc tính (104) được tạo cấu hình để trích xuất ra đặc tính từ tín hiệu lỗi (100) không được nội suy.

6. Bộ giải mã theo một điểm trong số các điểm trên,

trong đó bộ ước lượng tín hiệu (118), bao gồm:

bộ lọc phân tích (910) để phân tích tín hiệu lỗi hoặc tín hiệu lỗi được nội suy để thu được tín hiệu kích thích;

khởi mở rộng kích thích (912) để tạo ra tín hiệu kích thích được nâng cao có vùng quang phổ không được chứa trong tín hiệu lỗi (100); và

bộ lọc tổng hợp (914) để lọc tín hiệu kích thích được mở rộng;

trong đó bộ lọc phân tích (910) hoặc bộ lọc tổng hợp (914) được xác định bởi biểu diễn theo tham số được lựa chọn.

7. Bộ giải mã theo một điểm trong số các điểm trên,

trong đó bộ ước lượng tín hiệu (118) bao gồm bộ xử lý mở rộng băng thông quang phổ để tạo ra dải quang phổ được mở rộng tương ứng với vùng quang phổ không được chứa trong tín hiệu lỗi sử dụng ít nhất là dải quang phổ của tín hiệu lỗi và biểu diễn theo tham số,

trong đó biểu diễn theo tham số bao gồm các tham số cho ít nhất một trong số tham số điều chỉnh đường bao quang phổ (1060), bổ sung sản nhiễu (1020), bộ lọc nghịch đảo (1040) và bổ sung các âm khuyết (1080),

trong đó bộ tạo tham số được tạo cấu hình để cung cấp, cho đặc tính, nhiều lựa chọn biểu diễn theo tham số, mỗi lựa chọn biểu diễn theo tham số có các tham số cho ít nhất một trong số tham số điều chỉnh đường bao quang phổ (1060), bổ sung sản nhiễu (1020), lọc nghịch đảo (1040), và bổ sung các âm khuyết (1080).

8. Bộ giải mã theo một điểm trong số các điểm trên, còn bao gồm:

bộ phát hiện hoạt động giọng nói hoặc bộ phân biệt tiếng nói/không tiếng nói (500),

trong đó bộ ước lượng tín hiệu (118) được tạo cấu hình để ước lượng tín hiệu được nâng cao tần số sử dụng biểu diễn theo tham số chỉ khi bộ phát hiện hoạt động giọng nói hoặc bộ phát hiện tiếng nói/không tiếng nói (500) chỉ ra hoạt động giọng nói hoặc tín hiệu tiếng nói.

9. Bộ giải mã theo điểm 8,

trong đó bộ ước lượng tín hiệu (118) được tạo cấu hình để chuyển mạch (502, 504) từ khối thủ tục nâng cao tần số (511) thành khối thủ tục nâng cao tần số khác (513) hoặc để sử dụng các tham số khác nhau (514) được trích xuất từ tín hiệu được mã hóa, khi bộ phát hiện hoạt động giọng nói hoặc bộ phát hiện tiếng nói/không tiếng nói (500) chỉ ra tín hiệu không tiếng nói hoặc tín hiệu không có hoạt động giọng nói.

10. Bộ giải mã theo một điểm trong số các điểm trên,

trong đó mô hình thống kê được tạo cấu hình để cung cấp, đáp ứng với đặc tính, nhiều lựa chọn của các biểu diễn theo tham số (702 đến 708),

trong đó mỗi biểu diễn theo tham số lựa chọn có xác suất đồng nhất với xác suất của biểu diễn theo tham số lựa chọn khác hoặc khác với xác suất của biểu diễn theo tham số lựa chọn nhỏ hơn 10% của xác suất cao nhất.

11. Bộ giải mã theo một điểm trong số các điểm trên,

trong đó thông tin phụ lựa chọn chỉ được chứa trong khung (800) của tín hiệu được mã hóa, khi bộ tạo tham số (108) cung cấp nhiều lựa chọn biểu diễn theo tham số, và

trong đó thông tin phụ lựa chọn không được chứa trong khung khác (812) của tín hiệu âm thanh được mã hóa trong đó bộ tạo tham số (108) chỉ cung cấp lựa chọn biểu diễn theo tham số đơn lẻ đáp ứng với đặc tính (112).

12. Bộ mã hóa tạo ra tín hiệu được mã hóa (1212), bao gồm:

bộ mã hóa lõi (1200) để mã hóa tín hiệu ban đầu (1206) để thu được tín hiệu âm thanh được mã hóa (1208) có thông tin về số lượng nhỏ hơn của các dải tần số so với tín hiệu ban đầu (1206);

bộ tạo thông tin phụ lựa chọn (1202) để tạo ra thông tin phụ lựa chọn (1210) chỉ ra lựa chọn biểu diễn theo tham số được xác định (từ 702 đến 708) được cung cấp bởi mô hình thống kê đáp ứng với đặc tính (112) được trích xuất ra từ tín hiệu ban đầu (1206) hoặc từ tín hiệu âm thanh được mã hóa (1208) hoặc từ bản được giải mã của tín hiệu âm thanh được mã hóa (1208);

giao diện đầu ra (1204) để xuất ra tín hiệu được mã hóa (1212), tín hiệu được mã hóa bao gồm tín hiệu âm thanh được mã hóa (1208) và thông tin phụ lựa chọn (1210);

bộ giải mã lõi (1300) để giải mã tín hiệu âm thanh được mã hóa (1208) để thu được tín hiệu lõi được giải mã,

trong đó bộ tạo thông tin phụ lựa chọn (1202) bao gồm:

bộ trích xuất đặc tính (1302) để trích xuất ra đặc tính từ tín hiệu lõi được giải mã;

bộ xử lý mô hình thống kê (1304) để tạo ra số lượng lựa chọn biểu diễn theo tham số (702 đến 708) để ước lượng vùng quang phổ của tín hiệu được nâng cao tần số không được xác định bởi tín hiệu lõi được giải mã;

bộ ước lượng tín hiệu (1306) để ước lượng các tín hiệu âm thanh được nâng cao tần số cho các lựa chọn biểu diễn theo tham số (1305); và

bộ so sánh (1038) để so sánh các tín hiệu âm thanh được nâng cao tần số (1307) với tín hiệu ban đầu (1206),

trong đó bộ tạo thông tin phụ lựa chọn (1202) được tạo cấu hình để thiết lập thông tin phụ lựa chọn (1210) sao cho thông tin phụ lựa chọn xác định duy nhất lựa chọn biểu diễn theo tham số thu được trong tín hiệu âm thanh được nâng cao tần số khớp nhất với tín hiệu ban đầu (1206) dưới tiêu chuẩn tối ưu hóa.

13. Bộ mã hóa theo điểm 12,

trong đó giao diện đầu ra (1204) được tạo cấu hình để chỉ chứa thông tin phụ lựa chọn (1210) thành tín hiệu được mã hóa (1212), khi nhiều lựa chọn biểu diễn theo tham số được cung cấp bởi mô hình thống kê và không chứa bất kỳ thông tin phụ lựa chọn thành khung cho tín hiệu âm thanh được mã hóa (1208), trong đó mô hình thống kê có tác dụng để chỉ cung cấp biểu diễn theo tham số đơn lẻ đáp ứng với đặc tính.

14. Phương pháp tạo ra tín hiệu âm thanh được nâng cao tần số (120), bao gồm:

trích xuất ra (104) đặc tính từ tín hiệu lỗi (100);

trích xuất ra (110) thông tin phụ lựa chọn được liên kết với tín hiệu lỗi;

tạo ra (108) biểu diễn theo tham số để ước lượng vùng quang phổ của tín hiệu âm thanh được nâng cao tần số (120) không được xác định bởi tín hiệu lỗi (100), trong đó số lượng lựa chọn biểu diễn theo tham số (702, 704, 706, 708) được cung cấp đáp ứng với đặc tính (112), và trong đó một trong số những lựa chọn biểu diễn theo tham số được lựa chọn làm biểu diễn theo tham số đáp ứng với thông tin phụ lựa chọn (712 đến 718);

ước lượng (118) tín hiệu âm thanh được nâng cao tần số (120) sử dụng biểu diễn theo tham số được lựa chọn; và

phân lớp (606) khung của tín hiệu lỗi (100),

trong đó việc tạo (108) sử dụng mô hình thống kê thứ nhất (600), khi khung tín hiệu được phân lớp thuộc về lớp tín hiệu thứ nhất, và sử dụng mô hình thống kê khác thứ hai (602), khi khung được phân lớp thành lớp tín hiệu khác thứ hai.

15. Phương pháp tạo ra tín hiệu được mã hóa (1212), bao gồm:

mã hóa (1200) tín hiệu ban đầu (1206) để thu được tín hiệu âm thanh được mã hóa (1208) có thông tin về số lượng nhỏ hơn của các dải tần số so với tín hiệu ban đầu (1206);

tạo ra (1202) thông tin phụ lựa chọn (1210) chỉ ra lựa chọn biểu diễn theo tham số được xác định (từ 702 đến 708) được cung cấp bởi mô hình thống kê đáp ứng với đặc tính (112) được trích xuất ra từ tín hiệu ban đầu (1206) hoặc từ tín hiệu âm thanh được mã hóa (1208) hoặc từ bản được giải mã của tín hiệu âm thanh được mã hóa (1028);

xuất ra (1204) tín hiệu được mã hóa (1212), tín hiệu được mã hóa bao gồm tín hiệu âm thanh được mã hóa (1208) và thông tin phụ lựa chọn (1210); và

phân lớp (606) khung của tín hiệu lỗi (100),

trong đó việc tạo (108) sử dụng mô hình thống kê thứ nhất (600), khi khung tín hiệu được phân lớp thuộc về lớp tín hiệu thứ nhất, và sử dụng mô hình thống kê khác thứ hai (602), khi khung được phân lớp thành lớp tín hiệu khác thứ hai.

16. Vật ghi có thể đọc được bằng máy tính bao gồm chương trình máy tính để thực hiện phương pháp theo điểm 14, khi chương trình máy tính chạy trên máy tính hoặc bộ xử lý.

17. Vật ghi có thể đọc được bằng máy tính bao gồm chương trình máy tính để thực hiện phương pháp theo điểm 15, khi chương trình máy tính chạy trên máy tính hoặc bộ xử lý.

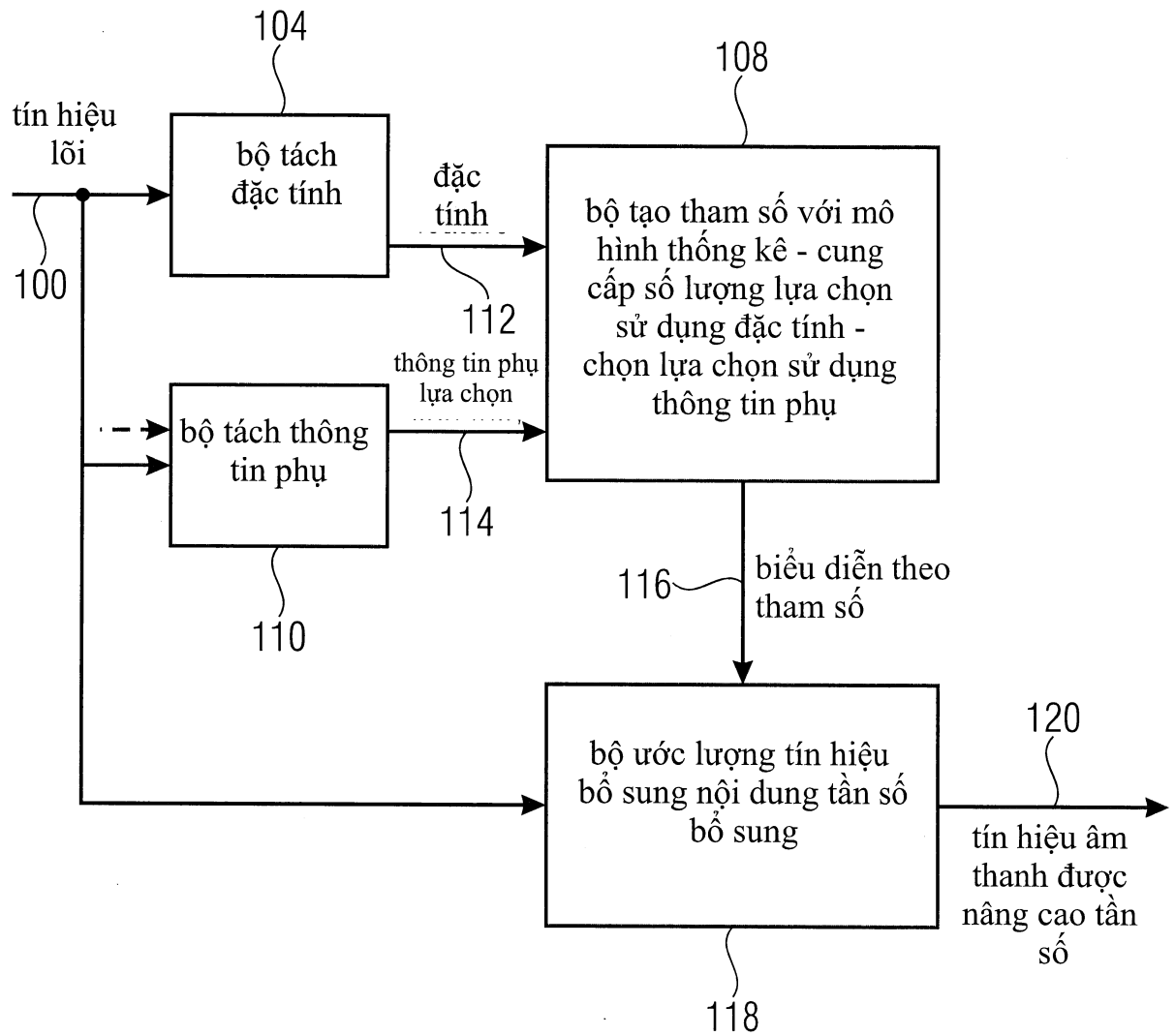


FIG 1

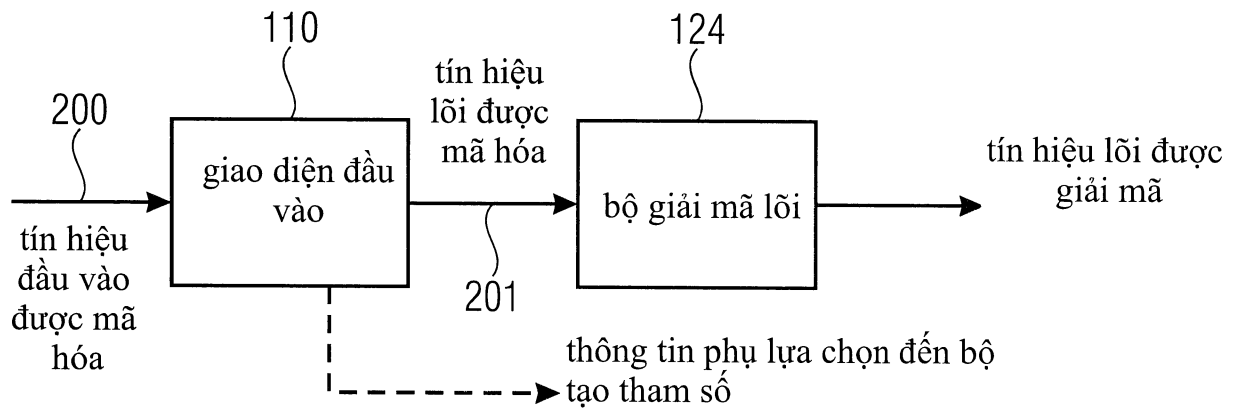


FIG 2

số bit của thông tin phụ lựa chọn	số lựa chọn biểu diễn theo tham số (cực đại)
1	2
2	4
3	8
4	16
5	32

FIG 3

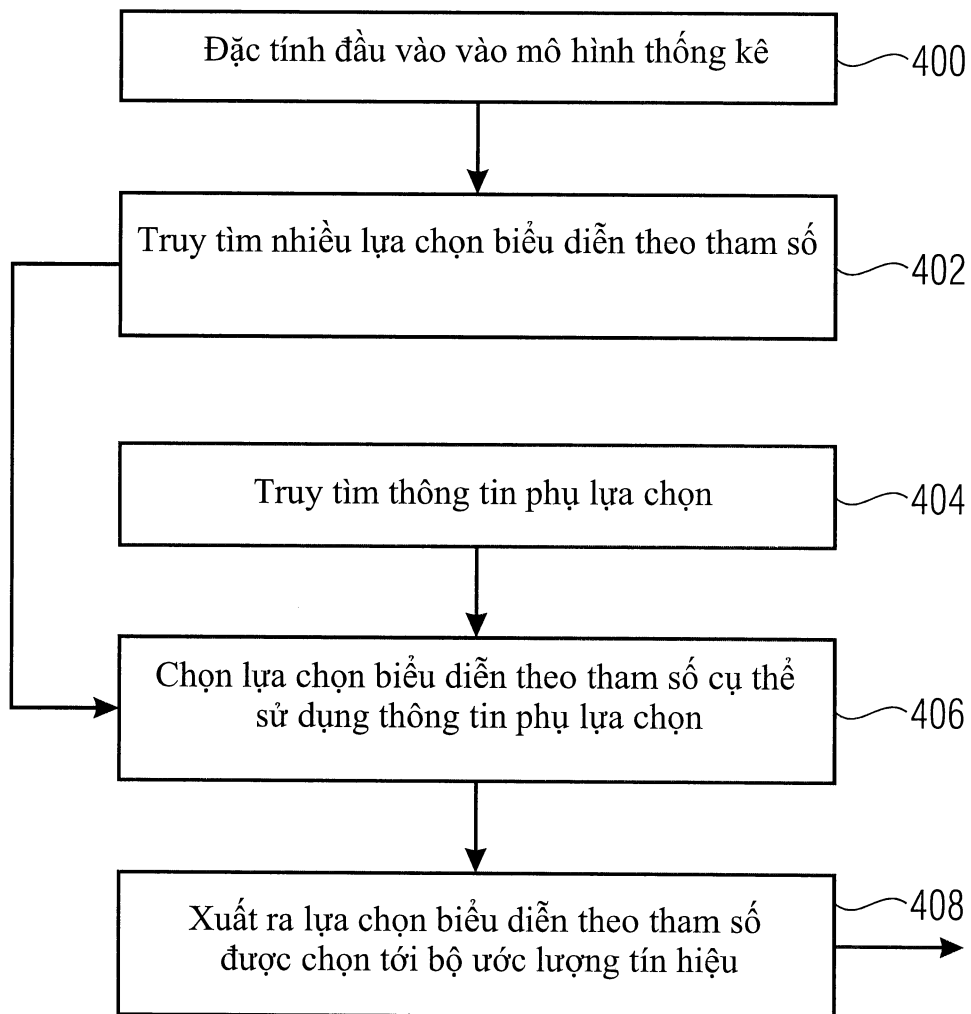


FIG 4

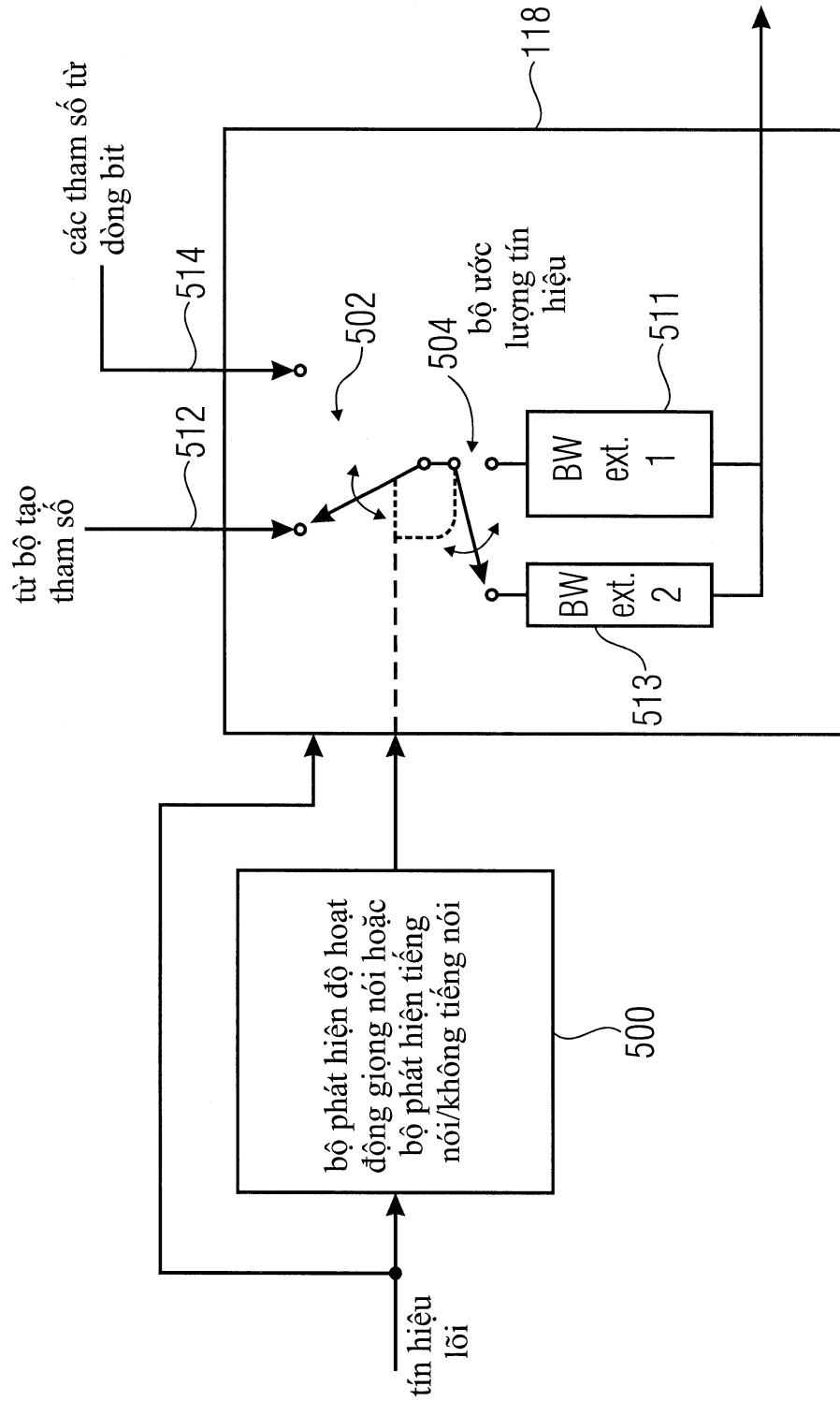


FIG 5

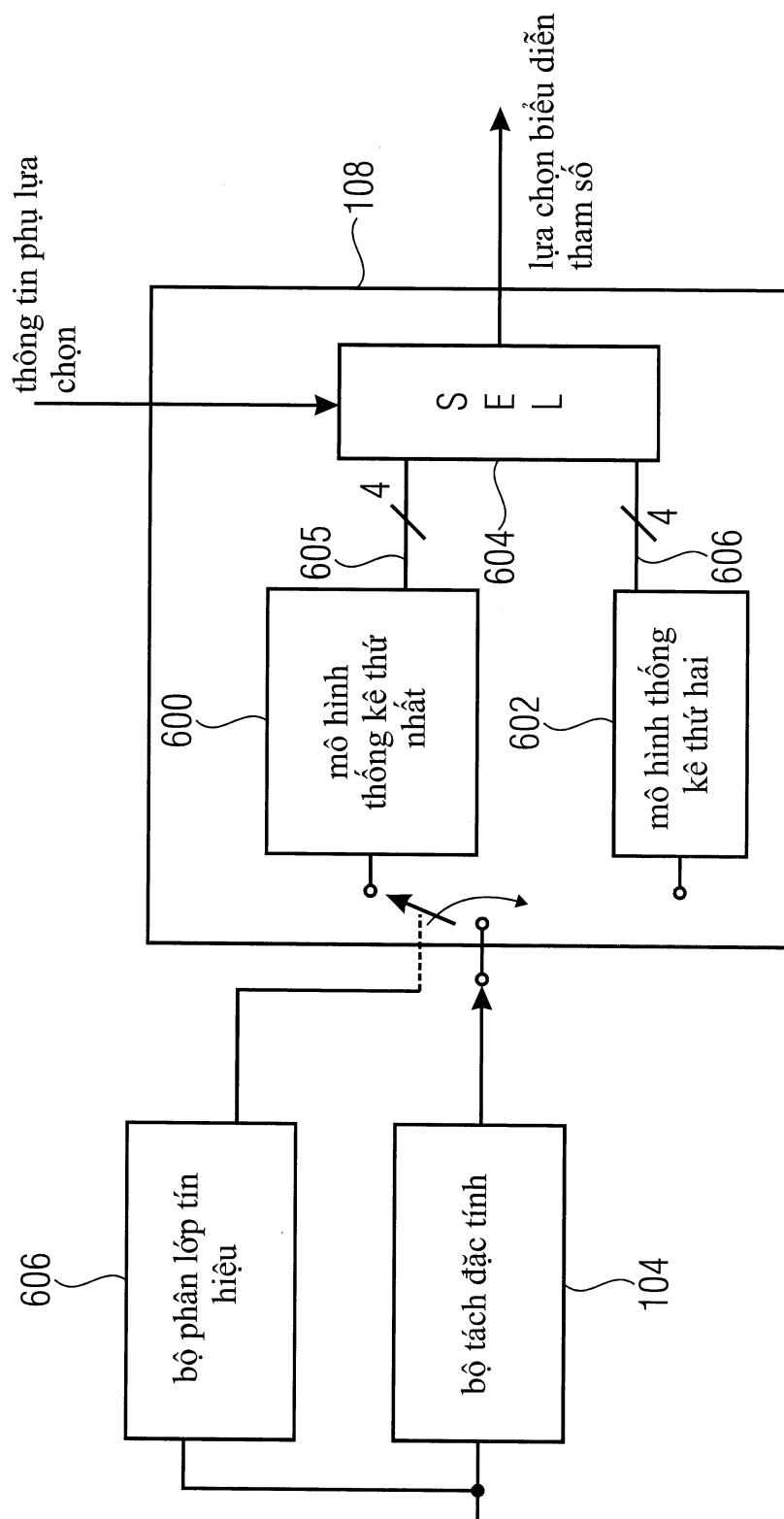


FIG 6

kết quả của mô hình thống kê

thông tin phụ lựa chọn

lựa chọn 1	702	00	712
lựa chọn 2	704	01	714
lựa chọn 3	706	10	716
lựa chọn 4	708	11	718

FIG 7

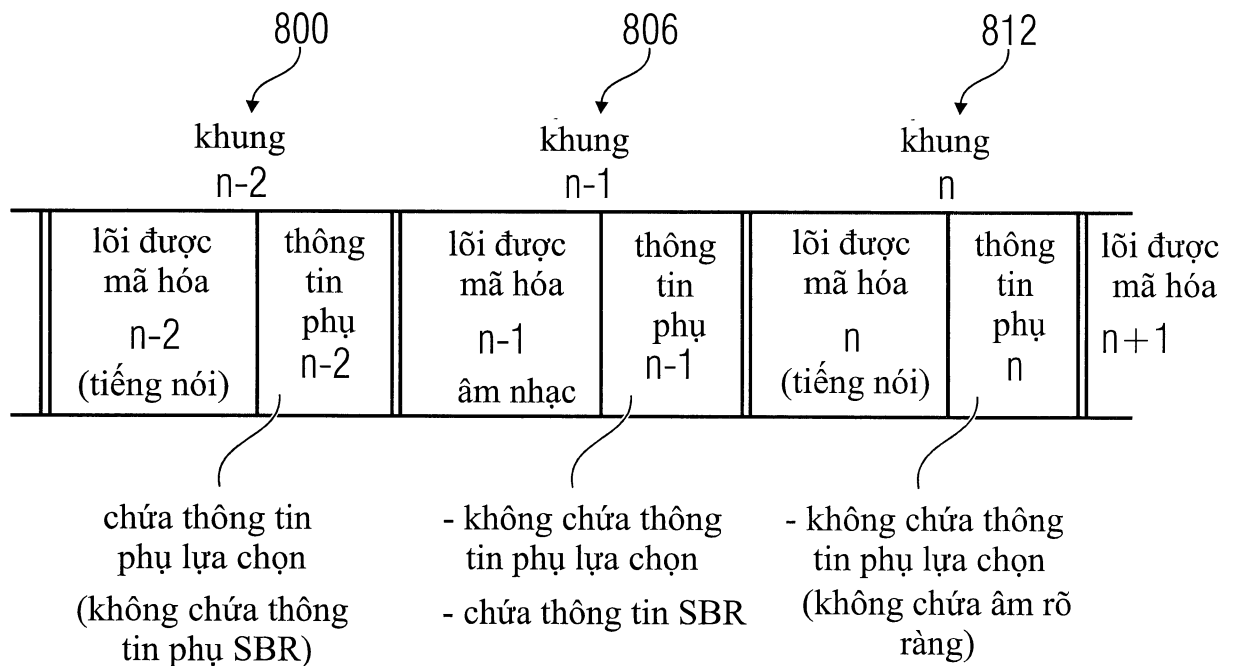


FIG 8

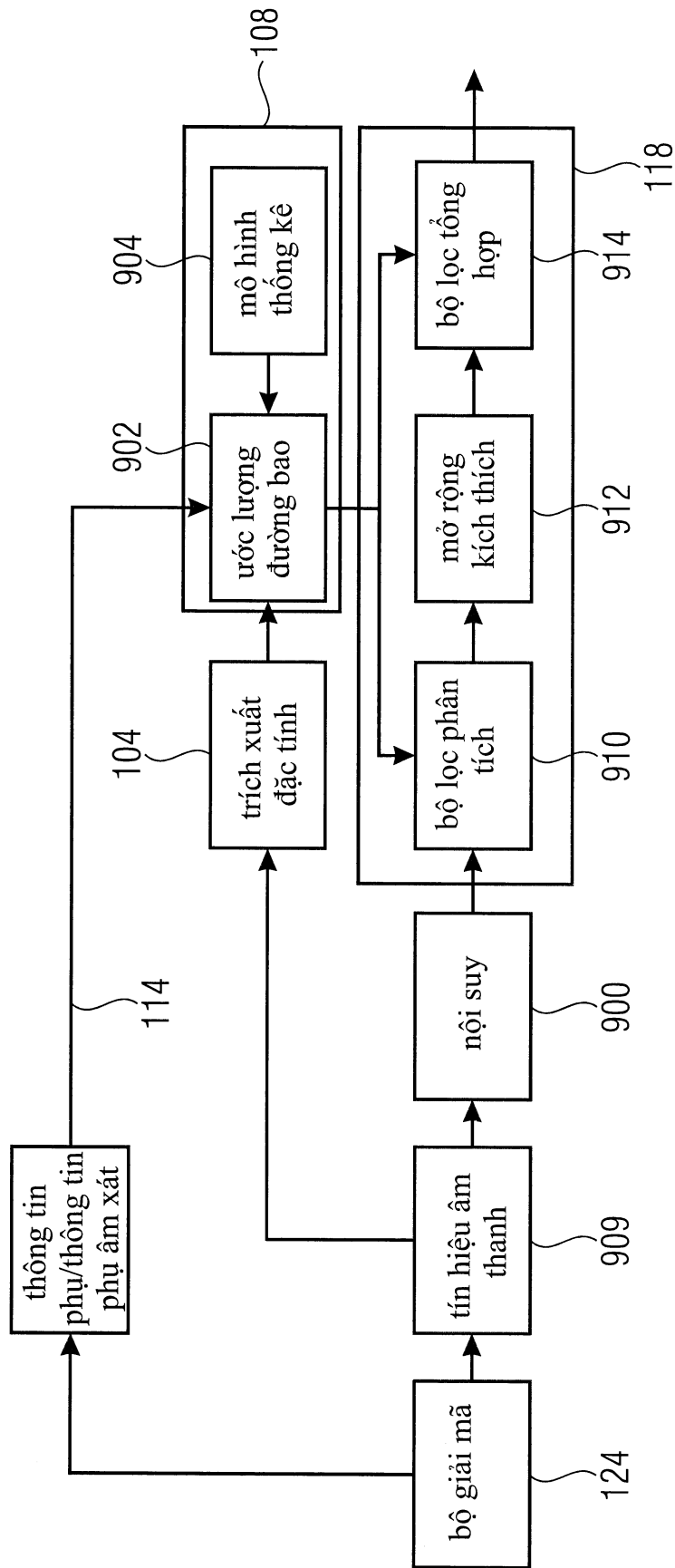


FIG 9

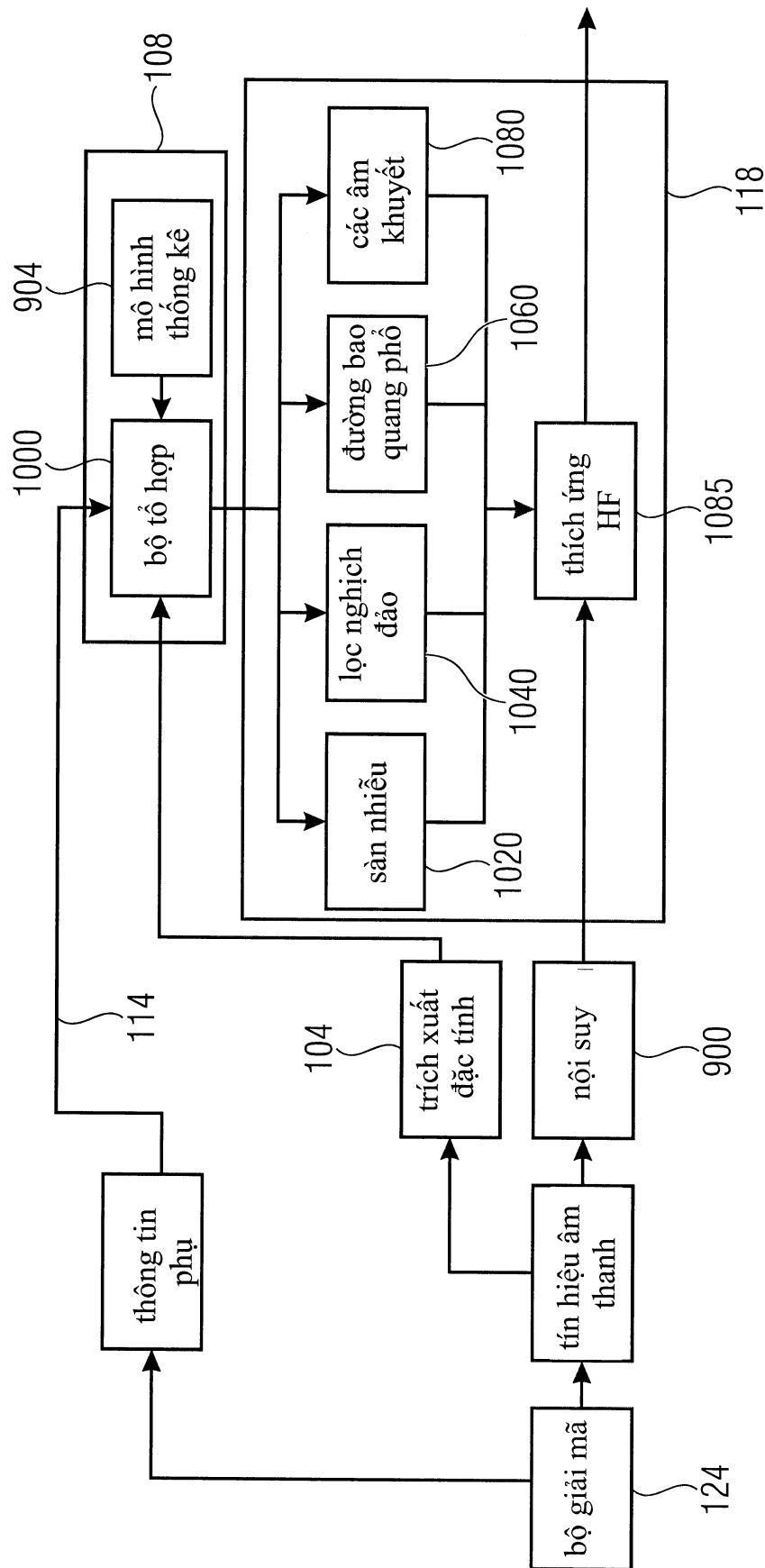


FIG 10

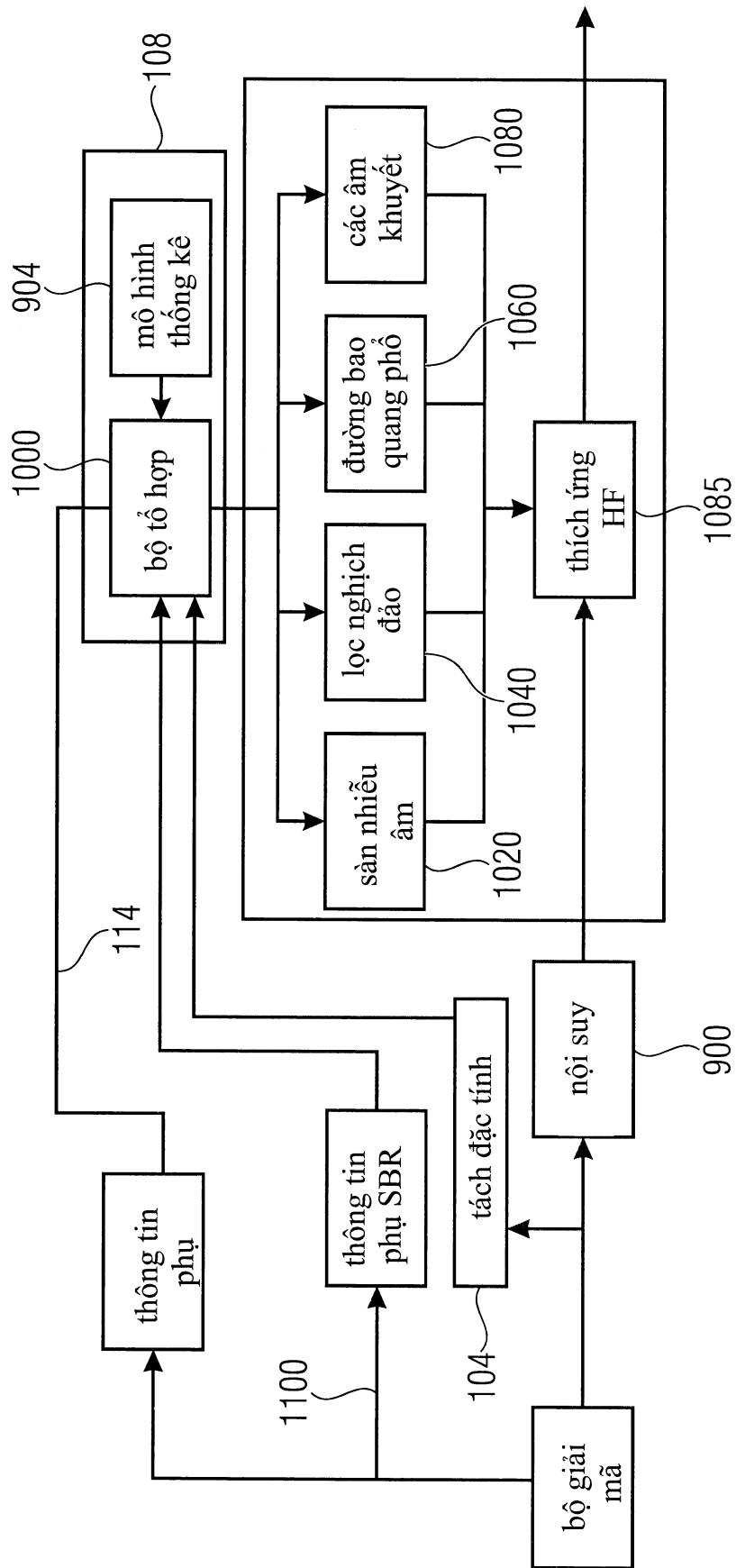


FIG 11

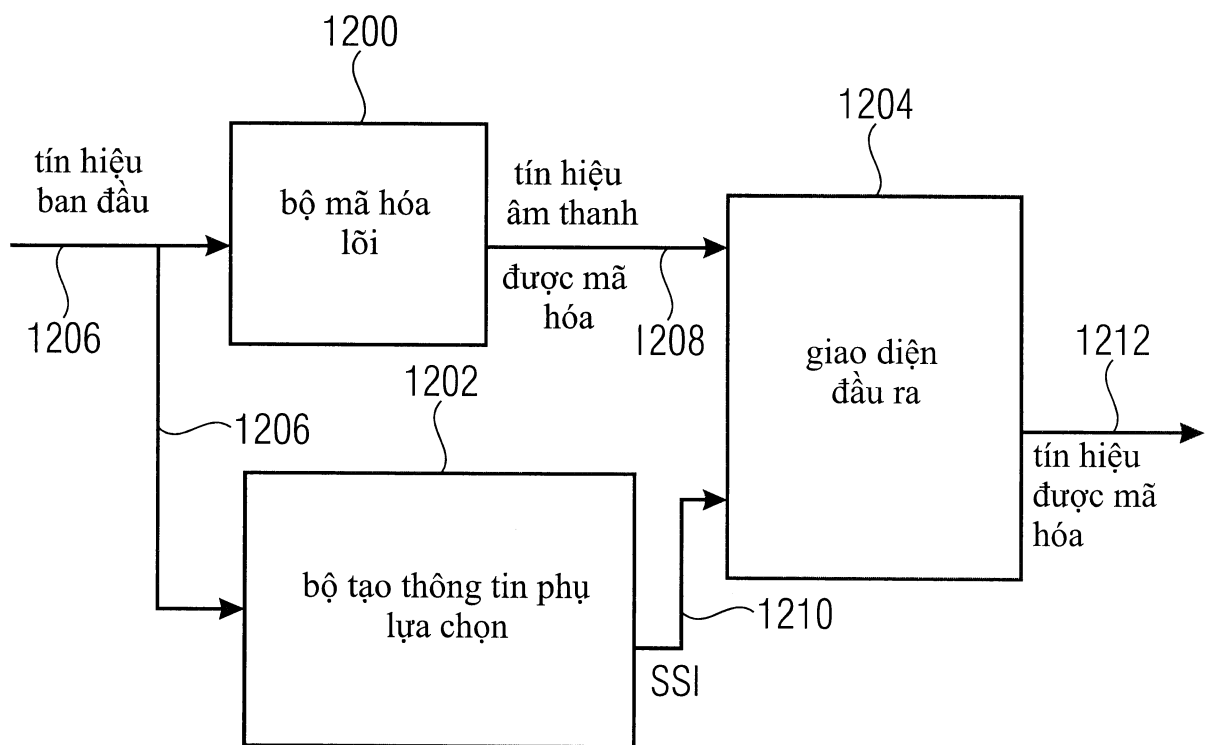


FIG 12

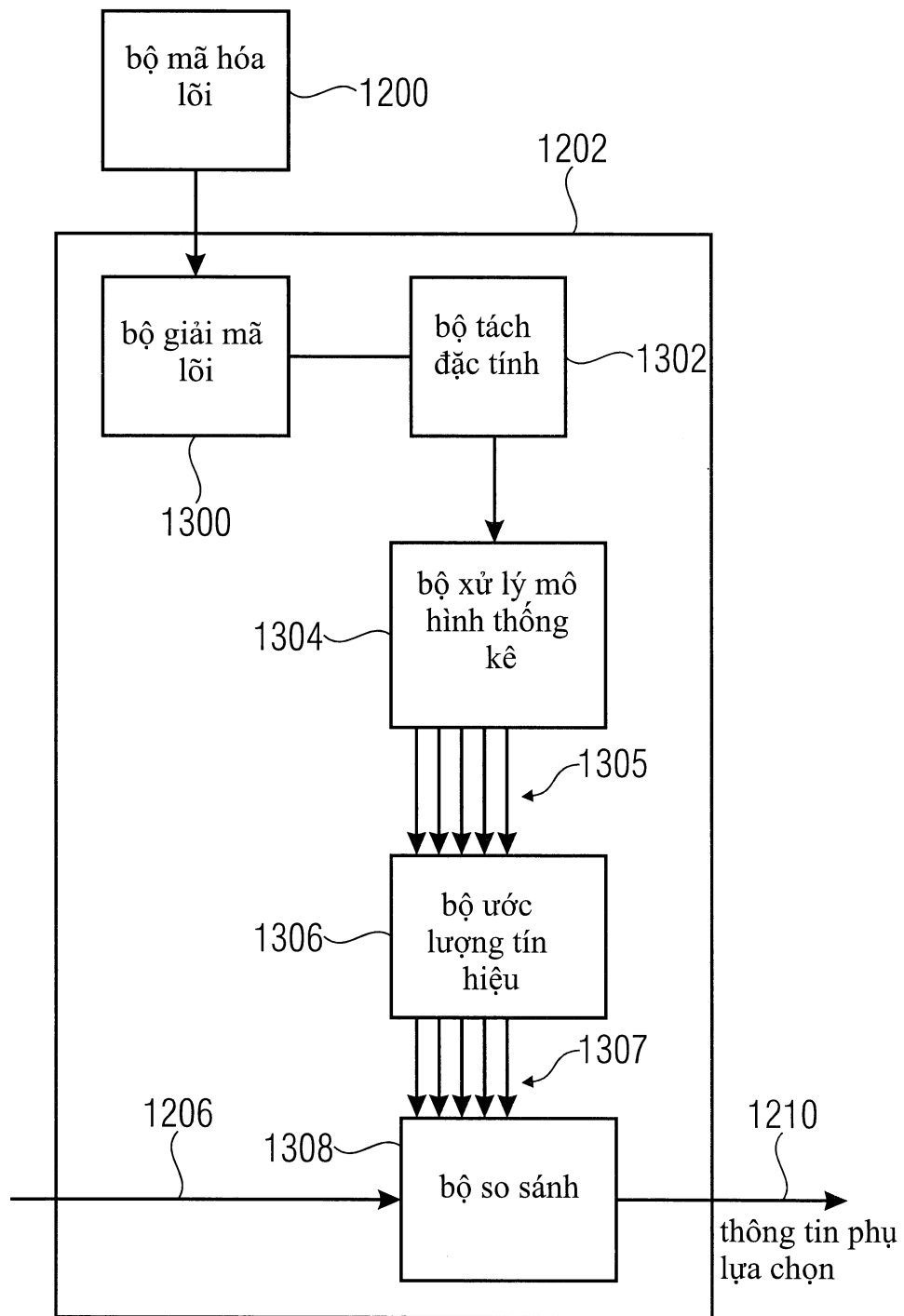


FIG 13

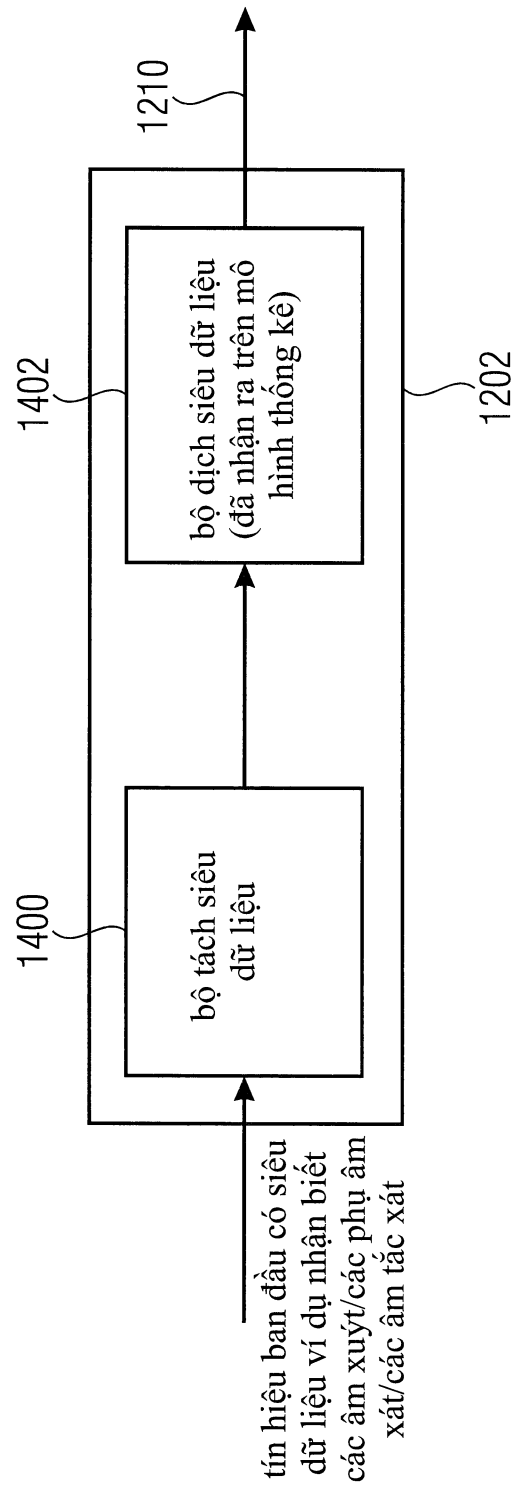


FIG 14

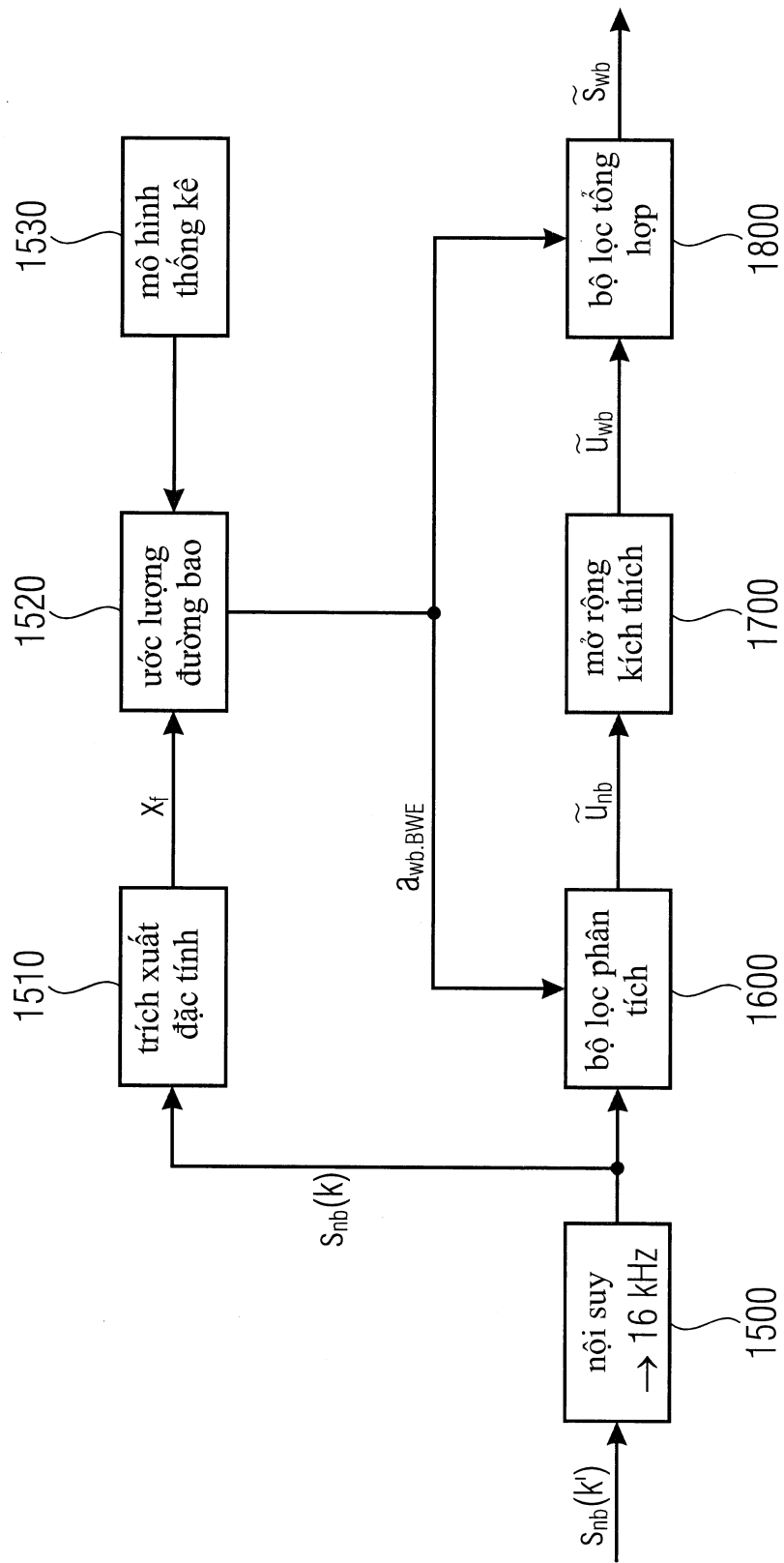


FIG 15
(Tình trạng kỹ thuật)

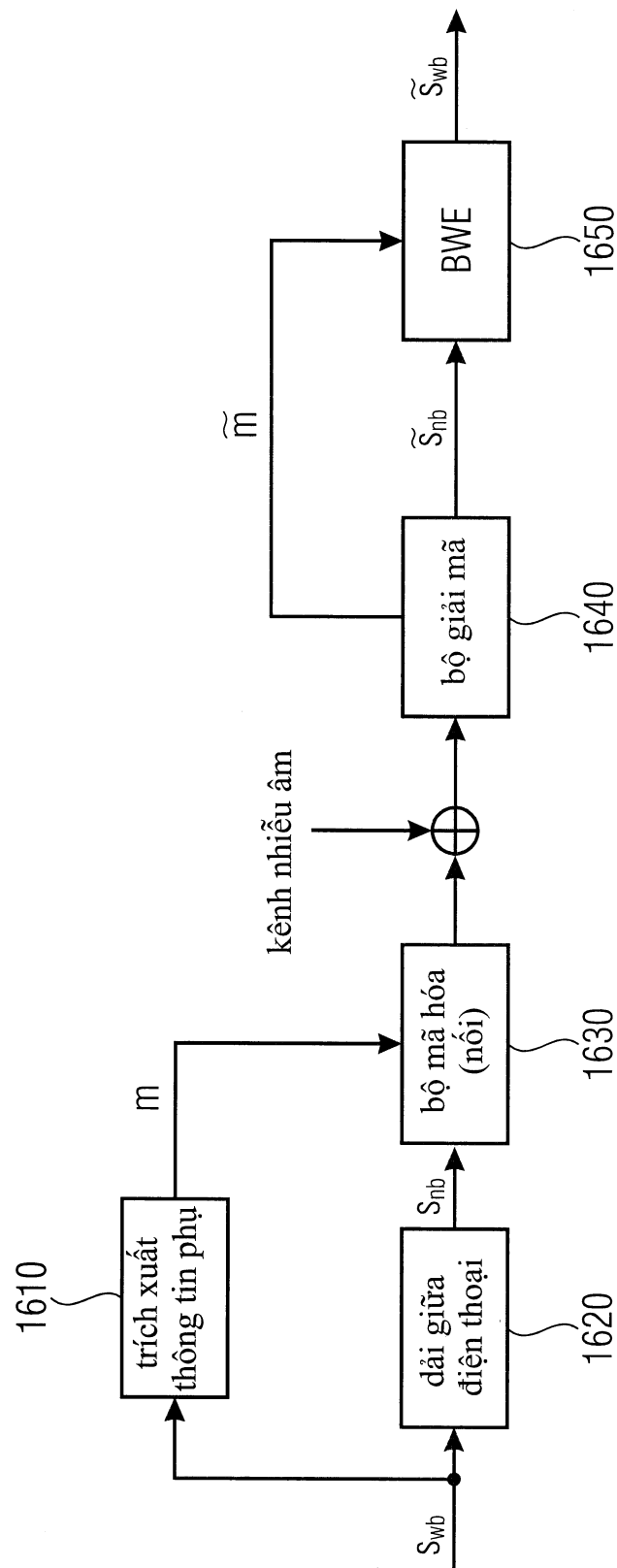


FIG 16
(Tình trạng kỹ thuật)