



(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)



1-0026167

(51)⁷ G10L 19/107

(13) B

(21) 1-2015-01535

(22) 31/07/2013

(86) PCT/EP2013/066074 31/07/2013

(87) WO/2014/053261 10/04/2014

(30) 61/710,137 05/10/2012 US

(45) 25/11/2020 392

(43) 25/08/2015 329A

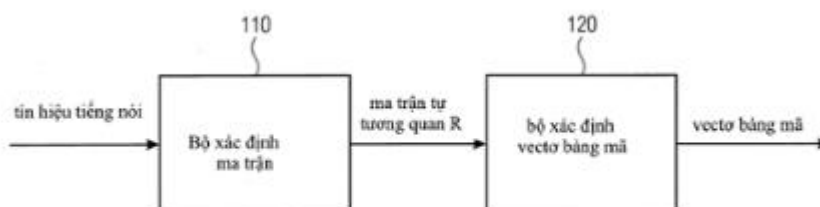
(73) Fraunhofer-Gesellschaft zur Foerderung der angewandten Forschung e.V. (DE)
Hansastrasse 27c, 80686 Muenchen, Germany

(72) BAECKSTROEM, Tom (FI); MULTRUS, Markus (DE); FUCHS, Guillaume (FR);
HELMRICH, Christian (DE); DIETZ, Martin (DE).

(74) CÔNG TY LUẬT TRÁCH NHIỆM HỮU HẠN AMBYS HÀ NỘI (AMBYS
HANOI)

(54) THIẾT BỊ VÀ PHƯƠNG PHÁP MÃ HÓA TÍN HIỆU TIẾNG NÓI, HỆ THỐNG
VÀ PHƯƠNG PHÁP MÃ HÓA VÀ GIẢI MÃ TÍN HIỆU TIẾNG NÓI

(57) Sáng chế đề cập đến thiết bị và phương pháp mã hóa tín hiệu tiếng nói, hệ thống và phương pháp mã hóa và giải mã tín hiệu tiếng nói. Thiết bị bao gồm bộ xác định ma trận (110) để xác định ma trận tự tương quan R và bộ xác định vectơ bằng mã (120) để xác định vectơ bằng mã phụ thuộc vào ma trận tự tương quan R. Bộ xác định ma trận (110) được tạo cấu hình để xác định ma trận tự tương quan R bằng cách xác định các hệ số vectơ của vectơ r, trong đó ma trận tự tương quan R gồm có nhiều hàng và nhiều cột, trong đó vectơ r chỉ ra một trong số các cột hoặc một trong số các hàng của ma trận tự tương quan R, trong đó $R(i, j) = r(|i - j|)$, trong đó $R(i, j)$ chỉ các hệ số của ma trận tự tương quan R, trong đó i là chỉ số thứ nhất chỉ ra một trong số nhiều hàng của ma trận tự tương quan R, và trong đó j là chỉ số thứ hai chỉ ra một trong số nhiều cột của ma trận tự tương quan R.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến mã hóa tín hiệu âm thanh, và, cụ thể là đề cập đến thiết bị mã hóa tín hiệu tiếng nói sử dụng ACELP trong miền tự tương quan.

Tình trạng kỹ thuật của sáng chế

Trong mã hóa tiếng nói bởi dự báo tuyến tính kích thích bằng mã (Code-Excited Linear Prediction - CELP), đường bao phổ (hoặc tương đương, cấu trúc thời gian ngắn hạn) của tín hiệu tiếng nói được mô tả bởi mô hình dự báo tuyến tính (linear predictive - LP) và phần dư dự báo được mô hình hóa bằng bộ dự báo thời gian dài (long-time predictor - LTP, cũng được biết đến như bảng mã thích ứng) và tín hiệu dư được trình bày bởi bảng mã (cũng được biết đến như bảng mã cố định). Gần đây, bảng mã cố định được ứng dụng phổ biến như bảng mã đại số, mà bảng mã này được trình bày bởi công thức đại số hoặc thuật toán, do đó không cần lưu trữ toàn bộ bảng mã, mà chỉ thuật toán, trong khi cho phép đồng thời cho thuật toán tìm kiếm nhanh. Bộ mã hóa - giải mã CELP ứng dụng bảng mã đại số cho phần dư được biết đến như bộ mã hóa - giải mã dự báo tuyến tính kích thích bằng mã đại số (Algebraic Code-Excited Linear Prediction - ACELP) (xem các tài liệu [1], [2], [3], [4]).

Trong mã hóa tiếng nói, sử dụng bảng mã phần dư đại số là phương pháp lựa chọn trong bộ mã hóa - giải mã dòng chính như trong các tài liệu [17], [13], [18]. ACELP được dựa trên việc mô hình hoá đường bao phổ bởi bộ lọc dự báo tuyến tính (LP), tần số cơ sở của các âm hữu thanh bằng bộ dự báo thời gian dài (LTP) và phần dư dự báo bằng bảng mã đại số. LTP và các tham số bảng mã đại số được tối ưu hóa bằng thuật toán bình phương tối thiểu trong miền cảm nhận, mà miền cảm nhận được thiết lập bởi bộ lọc.

Phần tính toán phức tạp nhất của các thuật toán loại ACELP, phần thất nghẽn, là sự tối ưu hóa của bảng mã phần dư. Chỉ thuật toán tối ưu được biết đến hiện nay có thể là sự tìm kiếm toàn bộ của không gian kích thước N^p cho mọi khung con, mà tại mọi điểm, sự đánh giá của mức độ phức tạp $O(N^2)$ được yêu cầu. Vì các giá trị đặc trưng là độ dài khung con $N = 64$ (tức là 5mili giây (mili second – ms)) với $p = 8$ xung, điều này hàm ý nhiều hơn 10^{20} phép toán mỗi giây. Rõ ràng điều này không phải

là lựa chọn khả thi. Để lưu lại trong các giới hạn phức tạp cài đặt bởi các yêu cầu phần cứng, các phương pháp tối ưu hóa bằng mã phải hoạt động với các thuật toán lặp không tối ưu. Nhiều thuật toán và sự cải tiến như vậy để quá trình tối ưu hóa đã có mặt trước đây, ví dụ trong các tài liệu [17], [19], [20], [21], [22].

Rõ ràng, sự tối ưu hóa ACELP được dựa trên sự mô tả tín hiệu tiếng nói $x(n)$ như đầu ra của mô hình dự báo tuyến tính sao cho tín hiệu tiếng nói ước lượng là

$$\hat{x}(n) = -\sum_{k=1}^m a(k) \hat{x}(n-k) + \hat{e}(k) \quad (1)$$

trong đó $a(k)$ là các hệ số LP và $\hat{e}(k)$ là tín hiệu dư. Trong dạng vector, phương trình này có thể được thể hiện như

$$\hat{x} = H\hat{e} \quad (2)$$

trong đó ma trận H được xác định như ma trận tích chập Toeplitz tam giác dưới với đường chéo $h(0)$ và các đường chéo dưới $h(1), \dots, h(39)$ và vector $h(k)$ là đáp ứng xung của mô hình LP. Cần được chú ý rằng trong ký hiệu này mô hình cảm nhận (thường phù hợp với mô hình trọng số LP) được bỏ qua, nhưng được giả định rằng mô hình cảm nhận được bao gồm trong đáp ứng xung $h(k)$. Sự bỏ qua này không ảnh hưởng đến tính tổng quát của các kết quả, nhưng đơn giản hóa ký hiệu. Sự bao gồm mô hình cảm nhận được ứng dụng như trong [1].

Sự phù hợp của mô hình được đo bởi sai số bình phương. Đó là,

$$\epsilon^2 = \sum_{k=1}^N (x(k) - \hat{x}(k))^2 = (e - \hat{e})^H H^H H (e - \hat{e}). \quad (3)$$

Sai số bình phương này được sử dụng để tìm các tham số mô hình tối ưu. Tại đây, được giả định rằng cả LTP và bảng mã xung đều được sử dụng để mô hình hóa vector e . Ứng dụng thực tế có thể được tìm thấy trong các công bố có liên quan (xem các tài liệu [1-4]).

Trong thực tế, việc đo sự phù hợp như trên có thể được đơn giản hóa như sau. Để cho ma trận $B = H^T H$ bao gồm sự tương quan của $h(n)$, để cho c_k là vector bảng mã cố định k 'th và cài đặt $\hat{e} = g c_k$, trong đó g là thừa số khuếch đại. Giả định rằng g được chọn một cách tối ưu, sau đó bảng mã được tìm kiếm bằng cách tối đa hóa tiêu chuẩn tìm kiếm

$$\frac{C_k^2}{E_k} = \frac{(x^T H c_k)^2}{c_k^T B c_k} = \frac{(d^T c_k)^2}{c_k^T B c_k} \quad (4)$$

trong đó $d = H^T x$ là vector gồm sự tương quan giữa vector mục tiêu và đáp ứng xung $h(n)$ và số mũ T biểu thị sự chuyển vị. Vector d và ma trận B được tính trước khi tìm kiếm bảng mã. Công thức này được sử dụng phổ biến trong sự tối ưu của cả LTP và bảng mã xung.

Nhiều nghiên cứu đã được đầu tư trong sự tối ưu hóa cách sử dụng của công thức trên. Ví dụ,

- 1) Chỉ các phần tử đó của ma trận B được tính toán mà được truy cập thực tế bởi thuật toán tìm kiếm. Hoặc:
- 2) Thuật toán thử và sai của tìm kiếm xung được giảm bớt để thử chỉ các vector bảng mã đó có xác suất thành công cao, dựa trên sự sàng lọc trước (xem ví dụ [1, 5]).

Chi tiết thực tế của thuật toán ACELP được đề cập đến khía cạnh của đáp ứng xung bằng không (zero impulse response - ZIR). Khía cạnh xuất hiện khi xét đến tín hiệu tổng hợp miền nguyên bản trong sự so sánh với phần dư được tổng hợp. Phần dư được mã hóa trong các khối tương ứng với kích thước khung hoặc khung con. Tuy nhiên, khi tổng hợp tín hiệu miền nguyên bản với mô hình LP của phương trình 1, phần dư độ dài cố định sẽ có "mục cuối" độ dài vô hạn, phù hợp với đáp ứng xung của bộ lọc LP. Đó là, dù vector bảng mã phần dư có độ dài hữu hạn, sẽ có ảnh hưởng đến tín hiệu tổng hợp quá xa khung hiện tại hoặc khung con. Ảnh hưởng của khung đến sau này có thể được tính toán bằng cách mở rộng vector bảng mã với các số không và tính toán đầu ra tổng hợp của phương trình 1 cho tín hiệu mở rộng này. Sự mở rộng này của tín hiệu tổng hợp được biết đến như đáp ứng xung bằng không. Sau đó, để tính đến ảnh hưởng của các khung trước trong mã hóa khung hiện tại, ZIR của khung trước được trừ từ đích của khung hiện tại. Trong mã hóa khung hiện tại, do đó, chỉ phần đó của tín hiệu được xét đến, mà không được mô hình hóa bởi khung trước đó.

Trong thực tế, ZIR được tính đến như sau: khi khung (con) $N-1$ đã được mã hóa, phần dư lượng tử hóa mở rộng với các số không đến độ dài của khung (con) tiếp theo N . Phần dư lượng tử hóa mở rộng được lọc bởi LP để thu được ZIR của tín hiệu

lượng tử hóa. ZIR của tín hiệu lượng tử hóa được trừ sau đó từ tín hiệu nguyên bản (không lượng tử hóa) và tín hiệu được thay đổi này tạo thành tín hiệu đích khi mã hóa khung (con) N. Cách này, tất cả các sai số lượng tử hóa tạo thành trong khung (con) N-1 sẽ được tính đến khi lượng tử hóa khung (con) N. Thực tế này cải thiện đáng kể chất lượng cảm nhận của tín hiệu đầu ra.

Tuy nhiên, sẽ được đánh giá cao nếu các khía cạnh được cải thiện hơn nữa để mã hóa âm thanh sẽ được cung cấp.

Tài liệu tham khảo

- [1] Salami, R. and Laflamme, C. and Bessette, B. and Adoul, J.P., "ITU-T G. 729 Annex A: reduced complexity 8 kb/s CS-ACELP codec for digital simultaneous voice and data", Communications Magazine, IEEE, vol 35, no 9, pp 56-63, 1997.
- [2] 3GPP TS 26.190 V7.0.0 , "Adaptive Multi-Rate (AMR-WB) speech codec", 2007.
- [3] ITU-T G.718, "Frame error robust narrow-band and wideband embedded variable bit-rate coding of speech and audio from 8-32 kbit/s", 2008.
- [4] Schroeder, M. and Atal, B., "Code-excited linear prediction (CELP): High-quality speech at very low bit rates", Acoustics, Speech, and Signal Processing, IEEE Int Conf, pp 937-940, 1985.
- [5] Byun, K.J. and Jung, H.B. and Hahn, M. and Kim, K.S., "A fast ACELP codebook search method", Signal Processing, 2002 6th International Conference on, vol 1, pp 422-425, 2002.
- [6] G. H. Golub and C. F. van Loan, "Matrix Computations", 3rd Edition, John Hopkins University Press, 1996.
- [7] Boley, D.L. and Luk, F.T. and Vandevoorde, D., "Vandermonde factorization of a Hankel matrix", Scientific computing, pp 27-39, 1997.
- [8] Bäckström, T. and Magi, C., "Properties of line spectrum pair polynomials - A review", Signal processing, vol. 86, no. 11, pp. 3286-3298, 2006.
- [9] A. Härmä, M. Karjalainen, L. Savioja, V. Välimäki, U. Laine, and J. Huopaniemi, "Frequencywarped signal processing for audio applications," J. Audio Eng. Soc, vol. 48, no. 11, pp. 1011–1031, 2000.

- [10] T. Laakso, V. Välimäki, M. Karjalainen, and U. Laine, “Splitting the unit delay [FIR/all pass filters design],” *IEEE Signal Process. Mag.*, vol. 13, no. 1, pp. 30–60, 1996.
- [11] J. Smith III and J. Abel, “Bark and ERB bilinear transforms,” *IEEE Trans. Speech Audio Process.*, vol. 7, no. 6, pp. 697–708, 1999.
- [12] R. Schappelle, “The inverse of the confluent Vandermonde matrix,” *IEEE Trans. Autom. Control*, vol. 17, no. 5, pp. 724–725, 1972.
- [13] B. Bessette, R. Salami, R. Lefebvre, M. Jelinek, J. Rotola-Pukkila, J. Vainio, H. Mikkola, and K. Jarvinen, “The adaptive multirate wideband speech codec (AMR-WB),” *Speech and Audio Processing, IEEE Transactions on*, vol. 10, no. 8, pp. 620–636, 2002.
- [14] M. Bosi and R. E. Goldberg, *Introduction to Digital Audio Coding and Standards*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 2003.
- [15] B. Edler, S. Disch, S. Bayer, G. Fuchs, and R. Geiger, “A time-warped MDCT approach to speech transform coding,” in *Proc 126th AES Convention*, Munich, Germany, May 2009.
- [16] J. Makhoul, “Linear prediction: A tutorial review,” *Proc. IEEE*, vol. 63, no. 4, pp. 561–580, April 1975.
- [17] J.-P. Adoul, P. Mabillean, M. Delprat, and S. Morissette, “Fast CELP coding based on algebraic codes,” in *Acoustics, Speech, and Signal Processing, IEEE Int Conf (ICASSP’87)*, April 1987, pp. 1957–1960.
- [18] ISO/IEC 23003-3:2012, “MPEG-D (MPEG audio technologies), Part 3: Unified speech and audio coding,” 2012.
- [19] F.-K. Chen and J.-F. Yang, “Maximum-take-precedence ACELP: a low complexity search method,” in *Acoustics, Speech, and Signal Processing, 2001. Proceedings.(ICASSP’01). 2001 IEEE International Conference on*, vol. 2. IEEE, 2001, pp. 693–696.
- [20] R. P. Kumar, “High computational performance in code excited linear prediction speech model using faster codebook search techniques,” in *Proceedings of the International Conference on Computing: Theory and Applications*. IEEE Computer Society, 2007, pp. 458–462.

- [21] N. K. Ha, “A fast search method of algebraic codebook by reordering search sequence,” in Acoustics, Speech, and Signal Processing, 1999. Proceedings., 1999 IEEE International Conference on, vol. 1. IEEE, 1999, pp. 21-24.
- [22] M. A. Ramirez and M. Gerken, “Efficient algebraic multipulse search,” in Telecommunications Symposium, 1998. ITS’98 Proceedings. SBT/IEEE International. IEEE, 1998, pp. 231-236.
- [23] ITU-T Recommendation G.191, “Software tool library 2009 user’s manual,” 2009.
- [24] ITU-T Recommendation P.863, “Perceptual objective listening quality assessment,” 2011.
- [25] T. Thiede, W. Treurniet, R. Bitto, C. Schmidmer, T. Sporer, J. Beerends, C. Colomes, M. Keyhl, G. Stoll, K. Brandenburg et al., “PEAQ – the ITU standard for objective measurement of perceived audio quality,” Journal of the Audio Engineering Society, vol. 48, 2012.
- [26] ITU-R Recommendation BS.1534-1, “Method for the subjective assessment of intermediate quality level of coding systems,” 2003.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là đề xuất các khía cạnh được cải thiện để mã hóa mục tiêu âm thanh. Mục đích của sáng chế được giải quyết bằng thiết bị theo điểm 1, bằng phương pháp mã hóa theo điểm 15, bằng bộ giải mã theo điểm 16, bằng phương pháp giải mã theo điểm 17, bằng hệ thống theo điểm 18, bằng phương pháp theo điểm 19 và bằng vật ghi đọc được bằng máy tính theo điểm 20.

Thiết bị mã hóa tín hiệu tiếng nói bằng cách xác định vector bảng mã của thuật toán mã hóa tiếng nói được đề xuất. Thiết bị gồm có bộ xác định ma trận để xác định ma trận tự tương quan R , và bộ xác định vector bảng mã để xác định vector bảng mã phụ thuộc vào ma trận tự tương quan R . Bộ xác định ma trận được tạo cấu hình để xác định ma trận tự tương quan R bằng cách xác định các hệ số của vector r , trong đó ma trận tự tương quan R gồm có nhiều hàng và nhiều cột, trong đó vector r chỉ ra một trong số các cột hoặc một trong số các hàng của ma trận tự tương quan R , trong đó $R(i, j) = r(i - j)$, trong đó $R(i, j)$ chỉ các hệ số của ma trận tự tương quan R , trong đó i là chỉ số thứ nhất chỉ ra một trong số nhiều hàng của ma trận tự tương quan R , và trong đó j là chỉ số thứ hai chỉ ra một trong số nhiều cột của ma trận tự tương quan R .

Thiết bị được tạo cấu hình để sử dụng vector bảng mã để mã hóa tín hiệu tiếng nói. Ví dụ, thiết bị có thể phát ra tín hiệu tiếng nói được mã hóa sao cho tín hiệu tiếng nói được mã hóa bao gồm nhiều hệ số dự báo tuyến tính, sự biểu thị tần số cơ sở của các âm hữu thanh (ví dụ, các tham số tần số cơ bản), và sự biểu thị vector bảng mã, ví dụ, chỉ số của vector bảng mã.

Ngoài ra, bộ giải mã để giải mã tín hiệu tiếng nói được mã hóa được mã hóa bằng thiết bị theo phương án được mô tả ở trên để thu được tín hiệu tiếng nói được giải mã được đề xuất.

Hơn nữa hệ thống mã hóa và giải mã tín hiệu tiếng nói đầu vào được đề xuất. Hệ thống bao gồm thiết bị theo phương án được mô tả ở trên để mã hóa tín hiệu tiếng nói đầu vào để thu được tín hiệu tiếng nói được mã hóa. Ngoài ra, hệ thống bao gồm bộ giải mã theo phương án được mô tả ở trên để giải mã tín hiệu tiếng nói được mã hóa để thu được tín hiệu tiếng nói được giải mã.

Các khía cạnh cải tiến cho hàm mục tiêu của thuật toán mã hóa tiếng nói ACELP được đề xuất, mà tính đến không chỉ ảnh hưởng của đáp ứng xung của khung trước đến khung hiện tại, mà còn cả ảnh hưởng của đáp ứng xung của khung hiện tại vào khung tiếp theo, khi tối ưu hóa các tham số của khung hiện tại. Một số phương án đề cập ở đây cải tiến bằng sự thay đổi ma trận tương quan, là trung tâm để sự tối ưu hóa ACELP quy ước đến ma trận tự tương quan, mà có cấu trúc Hermit Toeplitz. Bằng cách sử dụng cấu trúc này có thể tạo ra sự tối ưu hóa ACELP hiệu quả hơn theo quan điểm của cả độ phức tạp tính toán cũng như các yêu cầu bộ nhớ. Đồng thời, ứng dụng mô hình cảm nhận cũng trở nên nhất quán hơn và sự phụ thuộc liên khung có thể được tránh để cải thiện hiệu suất dưới ảnh hưởng của tiêu hao gói.

Mã hóa tiếng nói với hệ biến hóa ACELP được dựa trên thuật toán bình phương tối thiểu trong miền cảm nhận, miền cảm nhận được thiết lập bởi bộ lọc. Theo các phương án, độ phức tạp tính toán của độ rõ quy ước của bài toán bình phương tối thiểu có thể được giảm đi bằng cách tính đến tác động của đáp ứng xung bằng không trong khung tiếp theo. Các thay đổi được đề xuất giới thiệu cấu trúc Toeplitz để ma trận tương quan xuất hiện trong hàm mục tiêu, mà đơn giản hóa cấu trúc và giảm sự tính toán. Các khía cạnh đã đưa ra làm giảm độ phức tạp tính toán lên đến 17% mà không làm giảm chất lượng cảm nhận.

Các phương án được dựa trên việc tìm ra rằng bằng sự thay đổi nhỏ của hàm mục tiêu, độ phức tạp trong sự tối ưu hóa của bảng mã phần dư có thể giảm đi hơn nữa. Sự giảm đi này trong mức độ phức tạp đến mà không làm giảm chất lượng cảm nhận. Như một sự thay thế, vì sự tối ưu hóa phần dư ACELP được dựa trên các thuật toán tìm kiếm lặp, với sự thay đổi hiện tại, có thể tăng số lượng lặp mà không tăng mức độ phức tạp, và theo cách này thu được chất lượng cảm nhận được cải thiện.

Cả nhận thức mô hình hàm mục tiêu quy ước cũng như được biến đổi và nỗ lực để giảm thiểu sự biến dạng về mặt cảm nhận. Tuy nhiên, giải pháp tối ưu cho phương pháp quy ước là không nhất thiết tối ưu đối với các hàm mục tiêu thay đổi và ngược lại. Riêng điều này không có nghĩa là một phương pháp sẽ tốt hơn phương pháp khác, nhưng các lập luận phân tích làm cho sự thể hiện hàm mục tiêu thay đổi phù hợp hơn. Cụ thể, trong sự tương phản với hàm mục tiêu quy ước, các khía cạnh được đề xuất xử lý tương đương tất cả mẫu trong khung con, với cảm nhận phù hợp và dễ xác định và các mô hình tín hiệu.

Trong các phương án, các thay đổi đã đưa ra có thể được ứng dụng sao cho chỉ thay đổi sự tối ưu hóa của bảng mã phần dư. Do đó không thay đổi cấu trúc dòng bit và có thể được ứng dụng theo cách tương thích ngược lại đến bộ mã hóa - giải mã hiện tại.

Ngoài ra, phương pháp mã hóa tín hiệu tiếng nói bằng cách xác định vector bảng mã của thuật toán mã hóa tiếng nói được đề xuất. Phương pháp bao gồm:

- Xác định ma trận tự tương quan R. Và:
- Xác định vector bảng mã phụ thuộc vào ma trận tự tương quan R.

Xác định ma trận tự tương quan R gồm có các hệ số vector xác định của vector r. Ma trận tự tương quan R gồm nhiều hàng và nhiều cột. Vector r chỉ ra một trong số các cột hoặc một trong số các hàng của ma trận tự tương quan R, trong đó

$$R(i, j) = r(|i - j|).$$

$R(i, j)$ là các hệ số của ma trận tự tương quan R, trong đó i là chỉ số thứ nhất chỉ ra một trong số nhiều hàng của ma trận tự tương quan R, và trong đó j là chỉ số thứ hai chỉ ra một trong số nhiều cột của ma trận tự tương quan R.

Hơn nữa, phương pháp giải mã tín hiệu tiếng nói được mã hóa được mã hóa theo phương pháp mã hóa tín hiệu tiếng nói theo phương án được mô tả ở trên để thu được tín hiệu tiếng nói được giải mã được đề xuất.

Ngoài ra, phương pháp mã hóa và giải mã tín hiệu tiếng nói đầu vào được đề xuất. Phương pháp gồm có:

- Mã hóa tín hiệu tiếng nói đầu vào theo phương pháp mô tả ở trên để mã hóa tín hiệu tiếng nói để thu được tín hiệu tiếng nói được mã hóa. Và:
- Giải mã tín hiệu tiếng nói được mã hóa để thu được tín hiệu tiếng nói được giải mã theo phương pháp được mô tả ở trên để mã hóa tín hiệu tiếng nói.

Hơn nữa, sáng chế còn đề xuất các chương trình máy tính để thực hiện các phương pháp mô tả ở trên khi được chạy trên máy tính hoặc bộ xử lý tín hiệu.

Các phương án được ưu tiên sẽ được cung cấp trong các điểm yêu cầu bảo hộ phụ thuộc.

Mô tả văn tắt hình vẽ

Sau đây, các phương án của sáng chế được mô tả chi tiết hơn với sự tham chiếu đến các hình vẽ, trong đó:

Fig.1 là hình vẽ minh họa thiết bị mã hóa tín hiệu tiếng nói bằng cách xác định vector bảng mã của thuật toán mã hóa tiếng nói theo phương án.

Fig.2 là hình vẽ minh họa bộ giải mã theo phương án và bộ giải mã, và

Fig.3 là hình vẽ minh họa hệ thống bao gồm thiết bị mã hóa tín hiệu tiếng nói theo phương án và bộ giải mã.

Mô tả chi tiết sáng chế

Fig.1 minh họa thiết bị mã hóa tín hiệu tiếng nói để xác định vector bảng mã của thuật toán mã hóa tiếng nói theo phương án.

Thiết bị bao gồm bộ xác định ma trận (110) để xác định ma trận tự tương quan R , và bộ xác định vector bảng mã (120) để xác định vector bảng mã phụ thuộc vào ma trận tự tương quan R .

Bộ xác định ma trận (110) được tạo cấu hình để xác định ma trận tự tương quan R bằng cách xác định các hệ số vector của vector r .

Ma trận tự tương quan R bao gồm nhiều hàng và nhiều cột, trong đó vector r chỉ ra một trong số nhiều hàng hoặc một trong số nhiều cột của ma trận tự tương quan R , trong đó $R(i, j) = r(|i - j|)$.

$R(i, j)$ là các hệ số của ma trận tự tương quan R , trong đó i là chỉ số thứ nhất chỉ ra một trong số nhiều hàng của ma trận tự tương quan R , và trong đó j là chỉ số thứ hai chỉ ra một trong số nhiều cột của ma trận tự tương quan R .

Thiết bị được tạo cấu hình để sử dụng vector bảng mã để mã hóa tín hiệu tiếng nói. Ví dụ, thiết bị có thể phát ra tín hiệu tiếng nói được mã hóa sao cho tín hiệu tiếng nói được mã hóa bao gồm nhiều hệ số dự báo tuyến tính, sự biểu thị của tần số cơ sở của âm hữu thanh (ví dụ các tham số tần số cơ bản), và sự biểu thị của vector bảng mã.

Ví dụ, theo phương án cụ thể để mã hóa tín hiệu tiếng nói, thiết bị có thể được tạo cấu hình để xác định nhiều hệ số dự báo tuyến tính ($a(k)$) phụ thuộc vào tín hiệu tiếng nói. Ngoài ra, thiết bị được tạo cấu hình để xác định tín hiệu dư phụ thuộc vào nhiều hệ số dự báo tuyến tính ($a(k)$). Hơn nữa, bộ xác định ma trận 110 có thể được tạo cấu hình để xác định ma trận tự tương quan R phụ thuộc vào tín hiệu dư.

Sau đây, một số phương án nữa của sáng chế được mô tả.

Quay lại phương trình 3 và 4, trong đó phương trình 3 xác định sai số bình phương chỉ ra sự phù hợp của mô hình cảm nhận như:

$$\epsilon^2 = \sum_{k=1}^N (x(k) - \hat{x}(k))^2 = (e - \hat{e})^H H^H H (e - \hat{e}), \quad (3)$$

và trong đó phương trình 4

$$\frac{C_k^2}{E_k} = \frac{(x^T H c_k)^2}{c_k^T B c_k} = \frac{(d^T c_k)^2}{c_k^T B c_k} \quad (4).$$

chỉ ra tiêu chuẩn tìm kiếm, mà được tối đa hóa.

Thuật toán ACELP được xoay quanh phương trình 4, mà lần lượt được dựa vào phương trình 3.

Các phương án được dựa vào việc tìm ra rằng sự phân tích của các phương trình này bộc lộ rằng các giá trị phần dư lượng tử hóa $e(k)$ có ảnh hưởng rất khác đến năng lượng lỗi ϵ^2 phụ thuộc vào chỉ số k . Ví dụ, khi xét đến các chỉ số $k=1$ và $k=N$, nếu chỉ giá trị khác không của bảng mã phần dư sẽ xuất hiện tại $k=1$, sau đó năng lượng lỗi ϵ^2 dẫn đến kết quả:

$$\epsilon_1^2 = \sum_{k=1}^N (x(k) - e(1)h(k))^2 \quad (5)$$

trong khi cho $k=N$, năng lượng lỗi ϵ^2 dẫn đến kết quả:

$$\epsilon_N^2 = (x(N) - e(N)h(1))^2 + \sum_{k=1}^{N-1} (x(k))^2. \quad (6)$$

Nói cách khác, $e(1)$ là trọng số với đáp ứng xung $h(k)$ trong phạm vi từ 1 đến N , trong khi $e(N)$ là trọng số với chỉ $h(1)$. Theo quan điểm của trọng số phổ, điều này có nghĩa là mỗi $e(k)$ là trọng số với hàm trọng số phổ khác, sao cho, trong cực biên, $e(N)$ là trọng số tuyến tính. Từ góc nhìn mô hình hóa cảm nhận, có thể tạo ra độ nhạy để ứng dụng trọng số theo cảm nhận tương tự cho tất cả các mẫu trong khung. Phương trình 3 do đó nên được mở rộng sao cho tính đến ZIP trong khung tiếp theo. Nên được chú ý tại đây, không kể những cái khác, sự khác biệt với tình trạng kỹ thuật đã có là cả ZIR từ khung trước và ZIR trong khung tiếp theo cũng được tính đến.

Để cho $e(k)$ là nguyên bản, phần dư không được lượng tử hóa và $\hat{e}(k)$ phần dư lượng tử hóa. Hơn nữa, để cho cả hai phần dư là khác không trong phạm vi 1 đến N và bằng không ở vị trí khác. Sau đó

$$\begin{aligned} x(n) &= -\sum_{k=1}^m a(k)x(n-k) + e(n) = \sum_{k=1}^{\infty} e(n-k)h(k) \\ \hat{x}(n) &= -\sum_{k=1}^m a(k)\hat{x}(n-k) + \hat{e}(n) = \sum_{k=1}^{\infty} \hat{e}(n-k)h(k) \end{aligned} \quad (7)$$

Một cách tương đương, các mối quan hệ tương tự trong dạng ma trận có thể được thể hiện là:

$$\begin{aligned} x &= \tilde{H}e \\ \hat{x} &= \tilde{H}\hat{e} \end{aligned} \quad (8)$$

trong đó $\tilde{H}$ là ma trận tích chập kích thước vô hạn tương ứng với đáp ứng xung $h(k)$. Chèn vào trong phương trình 3 mang lại

$$\epsilon^2 = \|\tilde{H}e - \tilde{H}\hat{e}\|^2 = (e - \hat{e})^T \tilde{H}^T \tilde{H} (e - \hat{e}) = (e - \hat{e})^T R (e - \hat{e}) \quad (9)$$

trong đó $R = \tilde{H}^T \tilde{H}$ là kích thước cố định, ma trận Hermit Toeplitz tương ứng với sự tự tương quan của $h(n)$. Bằng sự suy ra tương tự phương trình 4, hàm mục tiêu thu được:

$$\frac{(e^T R \hat{e})^2}{(\hat{e}^T R \hat{e})} = \frac{(d^T e)^2}{(\hat{e}^T R \hat{e})}. \quad (10)$$

Hàm mục tiêu này rất giống với phương trình 4. Sự khác biệt chính là thay vì ma trận tương quan B , ở đây ma trận Hermit Toeplitz R là mẫu số.

Như giải thích ở trên, công thức mới này có lợi ích mà tất cả các mẫu của phần dư trong khung sẽ nhận trọng số phần dư tương tự. Tuy nhiên, một cách quan trọng, công thức này đưa ra lợi ích đáng kể cho độ phức tạp tính toán và yêu cầu bộ nhớ tốt.

Vì R là ma trận Hermit Toeplitz, cột thứ nhất $r(0) \dots r(N-1)$ xác định hoàn toàn ma trận. Nói cách khác, thay vì lưu trữ hoàn toàn ma trận $N \times N$, đủ để lưu trữ chỉ $N \times 1$ vector $r(k)$, do đó mang lại sự tiết kiệm đáng kể trong sự phân bổ bộ nhớ. Ngoài ra, độ phức tạp tính toán cũng được giảm vì không cần thiết để xác định tất cả $N \times N$ phần tử, nhưng chỉ cột thứ nhất $N \times 1$. Sự chỉ số hóa trong ma trận cũng đơn giản, vì phần tử (i,j) có thể tìm thấy bằng $R(i, j) = r(i-j)$.

Vì hàm mục tiêu trong phương trình 10 cũng giống với phương trình 4, cấu trúc của ACELP tổng quát có thể được tiếp tục dùng. Một cách cụ thể, bất kỳ phép toán nào sau đây có thể được thực hiện với hàm mục tiêu kia, với chỉ các thay đổi nhỏ đến thuật toán:

1. Sự tối ưu hóa của độ trễ LTP (bảng mã thích ứng)
2. Sự tối ưu hóa của bảng mã xung để mô hình hóa phần dư (bảng mã cố định)
3. Sự tối ưu hóa của độ khuếch đại của LTP và các xung, hoặc riêng rẽ hoặc phối hợp
4. Sự tối ưu hóa của các tham số khác bất kỳ mà hiệu suất của nó có thể được đo bằng sai số bình phương của phương trình 3

Chỉ phần mà phải được thay đổi trong các ứng dụng ACELP quy ước là sự xử lý của ma trận tương quan B , mà được thay thế bằng ma trận R , cũng như đích, mà phải bao gồm ZIR trong khung sau đây.

Một số phương án sử dụng các khái niệm của sáng chế bởi, bất cứ vị trí nào trong thuật toán ACELP, mà ma trận tương quan B xuất hiện, được thay thế bởi các ma trận tự tương quan R . Nếu tất cả các trường hợp của ma trận B được bỏ qua, sau đó việc tính toán giá trị của nó có thể được tránh.

Ví dụ, ma trận tự tương quan R được xác định bằng cách xác định các hệ số của cột thứ nhất $r(0), \dots, r(N-1)$ của ma trận tự tương quan R .

Ma trận R được xác định trong phương trình 9 bằng $R=H^TH$, do đó các phần tử của nó $R_{ij}=r(i-j)$ có thể được tính toán thông qua

$$r(k) = h(k) * h(-k) = \sum_l h(l)h(l-k) \quad (9a):$$

Tức là, chuỗi $r(k)$ là sự tự tương quan của $h(k)$.

Tuy nhiên, thông thường $r(k)$ có thể thu được bằng các phương tiện thậm chí còn hiệu quả hơn. Một cách cụ thể, trong các tiêu chuẩn mã hóa tiếng nói như AMR và

G.718, chuỗi $h(k)$ là đáp ứng xung của bộ lọc dự báo tuyến tính $A(z)$ được lọc bởi hàm trọng số theo cảm nhận $W(z)$, mà được tính đến bao gồm các chỉnh tăng. Nói cách khác, $h(k)$ chỉ ra đáp ứng xung trọng số theo cảm nhận của mô hình dự báo tuyến tính.

Bộ lọc $A(z)$ luôn được dự tính từ sự tự tương quan của tín hiệu tiếng nói $r_x(k)$, đó là, $r_x(k)$ đã được biết đến. Vì $H(z) = A^{-1}(z)W(z)$, cho phép chuỗi tự tương quan chuỗi $r(k)$ có thể được xác định bằng cách tính toán sự tự tương quan của $w(k)$ bằng

$$r_w(k) = w(k) * w(-k) = \sum_l w(l)w(l-k) \quad (9b)$$

do đó sự tự tương quan của $h(k)$ là

$$r(k) = r_x(k) * r_w(k) = \sum_l r_x(l)r_x(l-k). \quad (9c)$$

Trong một số phương án, phụ thuộc vào thiết kế của toàn bộ hệ thống, các phương trình này có thể được thay đổi một cách phù hợp

Vector bảng mã của bảng mã có thể sau đó, ví dụ, được xác định dựa trên ma trận tự tương quan R . Cụ thể, theo một số phương án, phương trình 10 có thể được sử dụng để xác định vector bảng mã của bảng mã.

Trong trường hợp này, phương trình 10 xác định hàm mục tiêu trong dạng

$$f(\hat{e}) = \frac{(d^T \hat{e})^2}{\hat{e}^T R \hat{e}} \text{ mà dạng tương tự khác như trong các tiêu chuẩn mã hóa tiếng nói AMR}$$

và G.718 nhưng sao cho ma trận R hiện tại có cấu trúc Toeplitz đối xứng. Hàm mục tiêu là sự tương quan định mức cơ bản giữa vector đích d và vector bảng mã $\hat{e}$ và vector bảng mã có thể nhất là cái đó, cái mà cho giá trị cao nhất cho sự tương quan định mức $f(\hat{e})$, ví dụ, tối đa hóa sự tương quan định mức $f(\hat{e})$.

Các vector bảng mã do đó có thể tối ưu hóa với các phương pháp tương tự như trong các tiêu chuẩn được đề cập. Một cách cụ thể, ví dụ, thuật toán rất đơn giản để tìm kiếm bảng mã đại số tối ưu (tức là bảng mã cố định) vector $\hat{e}$ cho phần dư có thể được ứng dụng, như đã mô tả ở trên. Tuy nhiên, cần được lưu ý, nỗ lực đáng kể đã được đầu tư trong thiết kế các thuật toán tìm kiếm có hiệu quả (so sánh với AMR và G.718), và thuật toán tìm kiếm này chỉ là ví dụ minh họa của ứng dụng.

1. Xác định vector bảng mã ban đầu $\hat{e}_0 = [0, 0 \dots 0]^T$ và cài đặt số lượng của các xung để $p = 0$.
2. Cài đặt sự đo chất lượng bảng mã ban đầu để $f_0 = 0$.
3. Cài đặt sự đo chất lượng bảng mã tạm thời để $\hat{f}_p = f_{p-1}$.
4. Cho mỗi vị trí k trong vector bảng mã
 - (i) Tăng p thêm một.
 - (ii) Nếu vị trí k đã chứa xung âm, tiếp tục bước vii.
 - (iii) Tạo vector bảng mã tạm thời $\varepsilon_p^+ = \hat{e}_{p-1}$ và bổ sung xung âm tại vị trí k .
 - (iv) Giá trị ước tính của vector bảng mã tạm thời bằng $f(\varepsilon_p^+)$.
 - (v) Nếu vector bảng mã tạm thời tốt hơn những vector trước bất kỳ, $f(\varepsilon_p^+) > \hat{f}_p$, sau đó lưu vector bảng mã này, cài đặt $\hat{f}_p = f(\varepsilon_p^+)$ và tiếp tục đến bước lặp tiếp theo.
 - (vi) Nếu vị trí k đã chứa xung âm, tiếp tục đến bước lặp tiếp theo.
 - (vii) Tạo vector bảng mã tạm thời $\varepsilon_p^- = \hat{e}_{p-1}$ và bổ sung xung âm tại vị trí k .
 - (viii) Đánh giá chất lượng của vector bảng mã tạm thời bằng $f(\varepsilon_p^-)$.
 - (ix) Nếu vector bảng mã tạm thời tốt hơn những vector trước bất kỳ, $f(\varepsilon_p^-) > \hat{f}_p$, sau đó lưu vector bảng mã này, cài đặt $\hat{f}_p = f(\varepsilon_p^-)$ và tiếp tục đến bước lặp tiếp theo.
5. Xác định vector bảng mã $\hat{e}_p$ để là cái cuối cùng (đó là, tốt nhất) của các vector bảng mã đã lưu.
6. Nếu số lượng của các xung p đã đạt đến số lượng của các xung, sau đó xác định vector đầu ra như $\hat{e} = \hat{e}_p$, và dừng lại. Mặt khác, tiếp tục với bước 4.

Như đã chỉ ra, so sánh với với các ứng dụng ACELP quy ước, trong một số phương án, đích được thay đổi sao cho bao gồm ZIR trong khung sau đây.

Phương trình 1 mô tả mô hình dự báo tuyến tính được sử dụng trong bộ mã hóa-giải mã loại ACELP. Đáp ứng xung bằng không (ZIR, đôi khi cũng được biết đến như đáp ứng đầu vào bằng không), đề cập đến đầu ra của các mô hình dự báo tuyến tính khi phần dư của khung hiện tại (và tất cả các khung hình trong tương lai) được cài đặt đến không. ZIR có thể đã được tính bằng cách xác định phần dư mà là không từ vị trí N về phía trước như

$$e_K'(n) = \begin{cases} e(n) & \text{for } n < K \\ 0 & \text{for } n \geq K \end{cases} \quad (10a):$$

do đó ZIR có thể được xác định là

$$ZIR_K(n) = \sum_{k=0}^N h(k) e_K'(n-k). \quad (10b)$$

Bằng cách trừ ZIR này từ tín hiệu đầu vào, tín hiệu thu được mà phụ thuộc vào phần dư chỉ từ khung hiện tại về phía trước.

Một cách tương đương, ZIR có thể được xác định bằng cách lọc các tín hiệu đầu vào trước như sau:

$$ZIR_K(n) = \begin{cases} x(n) & \text{for } n < K \\ -\sum_{k=1}^m a(k) ZIR_K(n-k) & \text{for } n \geq K. \end{cases} \quad (10c)$$

Tín hiệu đầu vào mà ZIR đã được loại bỏ thường được biết như đích và có thể được xác định cho khung mà bắt đầu tại vị trí K như $d(n) = x(n) - ZIR_K(n)$. Đích này là theo nguyên tắc đúng bằng đích trong các tiêu chuẩn ARM và G.718. Khi lượng tử hóa tín hiệu, tín hiệu lượng tử hóa $\tilde{d}(n)$ được so sánh với $d(n)$ khoảng thời gian của khung $K \leq n < K + N$.

Ngược lại, phần dư của khung hiện tại có ảnh hưởng đến các khung sau đây, do đó là hữu ích để xem xét ảnh hưởng của nó khi lượng tử hóa tín hiệu, đó là, một người do đó có thể muốn đánh giá sự khác biệt $\tilde{d}(n) - d(n)$ cũng vượt ra khỏi khung hiện tại, $n > K + N$. Tuy nhiên, để làm điều đó, một người có thể muốn xem xét sự ảnh hưởng của tín hiệu của khung hiện tại chỉ bằng cách cài đặt các phần dư của khung sau đây đến không. Do đó, ZIR của $\tilde{d}(n)$ trong khung tiếp theo có thể được so sánh. Nói cách khác, đích thay đổi thu được:

$$d'(n) = \begin{cases} 0 & n < K \\ d(n) & K \leq n < K + N \\ -\sum_{k=1}^m a(k) d'(n-k) & n > K + N. \end{cases} \quad (10d)$$

Một cách tương đương, sử dụng đáp ứng xung $h(n)$ của $A(z)$, sau đó

$$d'(n) = \sum_{k=K}^{K+N-1} e(k)h(n-k). \quad (10e)$$

Công thức này có thể được viết trong dạng ma trận thuận tiện bằng $d' = He$ mà H và e được xác định như trong phương trình 2. Có thể thấy rằng mục tiêu thay đổi đổi là đúng x của phương trình 2.

Trong sự tính toán của ma trận R , chú ý rằng trong lý thuyết, đáp ứng xung $h(k)$ là chuỗi vô hạn, mà không thể thực hiện được trong hệ thống thực tế.

Tuy nhiên, cũng không

- 1) cắt bỏ hoặc tạo cửa sổ đáp ứng xung cho một độ dài hữu hạn và xác định sự tự tương quan của đáp ứng xung bị cắt bỏ, hoặc
- 2) tính phổ năng lượng của đáp ứng xung sử dụng phổ Fourier của LP liên quan và các bộ lọc theo cảm nhận, và thu được sự tự tương quan bằng sự biến đổi Fourier nghịch đảo

là có khả năng.

Hiện nay, việc sử dụng LTP mở rộng được mô tả.

Bộ dự báo thời gian dài (LTP) cũng thực sự là bộ dự báo tuyến tính.

Theo phương án, bộ xác định ma trận R 110 có thể được tạo cấu hình để xác định ma trận tự tương quan R phụ thuộc vào bộ dự báo tuyến tính có trọng số theo cảm nhận, ví dụ, phụ thuộc vào bộ dự báo thời gian dài.

LP và LTP có thể được chập vào trong một bộ dự báo ghép, mà bao gồm cả hai dạng đường bao phổ cũng như cấu trúc sóng hài. Đáp ứng xung của bộ dự báo như vậy sẽ rất dài, do đó mà thậm chí còn khó khăn hơn để xử lý với tình trạng kỹ thuật trước đây. Tuy nhiên, nếu sự tự tương quan của bộ dự báo tuyến tính đã được biết đến, sau đó sự tự tương quan của bộ dự báo ghép có thể được tính toán bằng cách lọc một cách đơn giản sự tự tương quan với LTP phía trước và phía sau, hoặc với quá trình tương tự trong miền tần số.

Chú ý rằng các phương pháp trước đây sử dụng LTP có vấn đề khi độ trễ LTP ngắn hơn độ dài khung, vì LTP sẽ gây ra vòng phản hồi trong khung. Lợi ích của sự

bao gồm LTP trong hàm mục tiêu là như vậy khi độ trễ của LTP ngắn hơn độ dài khung, sau đó sự phản hồi này được tính đến một cách rõ ràng trong sự tối ưu hóa.

Trong phần tiếp theo, sự mở rộng cho sự tối ưu hóa nhanh trong miền không tương quan được mô tả.

Thách thức chủ yếu trong thiết kế của các hệ thống ACELP là sự giảm độ phức tạp tính toán. Các hệ thống ACELP phức tạp vì việc lọc bởi LP gây ra các mối tương quan phức tạp giữa các mẫu phân dư, mà được mô tả bằng ma trận B hoặc trong ngữ cảnh hiện tại bằng ma trận R. Vì các mẫu của $e(n)$ được tương quan, không thể chỉ lượng tử hóa $e(n)$ với độ chính xác mong muốn, nhưng nhiều sự kết hợp của các sự lượng tử hóa khác nhau với phương pháp thử và sai đã được thử nghiệm, để tìm sự lượng tử hóa tốt nhất đối với hàm mục tiêu của phương trình 3 hoặc 10.

Bằng sự đưa ra ma trận R, góc nhìn mới về các tương quan này thu được. Cụ thể là, vì R có cấu trúc Hermit Toeplitz, một số sự phân tích ma trận hiệu quả có thể được ứng dụng, như sự phân tích giá trị đơn, sự phân tích Cholesky hoặc sự phân tích Vandermonde của các ma trận Hankel (các ma trận Hankel là các ma trận Toeplitz đảo ngược lên, do đó, các phân tích tương tự có thể được ứng dụng cho các ma trận Toeplitz và Hankel (xem [6] và [7])). Để cho $R = E D E^H$ là sự phân tích của R sao cho D là ma trận đường chéo của ma trận có kích thước và vị trí tương tự như R. Phương trình 9 sau đó có thể được thay đổi như sau:

$$\epsilon^2 = (e - \hat{e})^H R (e - \hat{e}) = (e - \hat{e})^H E D E^H (e - \hat{e}) = (f - \hat{f}) D (f - \hat{f}) \quad (11)$$

trong đó $\hat{f} = E^H \hat{e}$. Vì D là đường chéo, sai số cho mỗi mẫu của $f(k)$ là độc lập với các mẫu $f(i)$ khác. Trong phương trình 10, được giả định rằng các vector bằng mã được thu nhỏ lại bởi độ khuếch đại tối ưu, do đó hàm mục tiêu mới là

$$\frac{(f^H D \hat{f})^2}{\hat{f}^H D \hat{f}}. \quad (12)$$

Ở đây, các mẫu được tương quan lại (vì sự thay đổi sự lượng tử hóa của một dòng thay đổi độ khuếch đại tối ưu cho tất cả các dòng), nhưng trong sự so sánh với phương trình 10, hiệu quả của sự tương quan được giới hạn ở đây. Tuy nhiên, thậm chí nếu sự tương quan được tính đến, sự tối ưu hóa của hàm mục tiêu này là đơn giản hơn nhiều sự tối ưu hóa của phương trình 3 hoặc 10.

Sử dụng phương pháp phân tích này, có thể

1. để ứng dụng bất kỳ đại lượng vô hướng quy ước hoặc kỹ thuật lượng tử hóa vector với độ chính xác mong muốn, hoặc
2. để sử dụng phương trình 12 như hàm mục tiêu với bất kỳ thuật toán tìm kiếm xung ACELP quy ước.

Cả hai phương pháp cho sự lượng tử gần như tối ưu đối với phương trình 12. Vì kỹ thuật lượng tử quy ước thường không yêu cầu bất kỳ phương pháp thô nào (đối với ngoại lệ của vòng lặp tỷ lệ có khả năng), và bởi vì ma trận D đơn giản hơn B hoặc R , cả hai phương pháp lượng tử hóa là ít phức tạp hơn các thuật toán tìm kiếm xung ACELP quy ước. Nguồn chính của độ phức tạp tính toán trong phương pháp này là do sự tính toán của sự phân tích ma trận.

Một số phương án sử dụng phương trình 12 để xác định vector bảng mã của bảng mã.

Ví dụ, một số phép tìm thừa số ma trận cho R của dạng $R = E^H D E$ tồn tại. Ví dụ,

- (a) Sự phân tích giá trị riêng có thể được tính toán cho ví dụ bằng cách sử dụng thư viện khoa học GNU (http://www.gnu.org/software/gsl/manual/html_node/Real-Symmetric-Matrices.html). Các ma trận R là có thật và đối xứng (cũng như Toeplitz), do đó chức năng "gls_eigen_symm ()" có thể được sử dụng để xác định các ma trận E và D . Sự bổ sung khác của sự phân tích giá trị riêng tương tự đã có sẵn trong tài liệu [6].
- (b) Phép tìm thừa số Vandermonde của các ma trận Toeplitz [7] có thể được sử dụng bằng cách sử dụng thuật toán mô tả trong tài liệu [8]. Thuật toán này trả về các ma trận E và D sao cho E là ma trận Vandermonde, mà tương đương với sự biến đổi Fourier riêng biệt với sự phân bố tần số đồng đều.

Sử dụng phép tìm thừa số như vậy, vector phần dư e có thể được biến đổi thành miền biến đổi bằng $f = E^H e$ hoặc $f' = D^{1/2} E^H e$. Bất kỳ phương pháp lượng tử hóa thông thường nào cũng có thể được áp dụng miền này, ví dụ,

1. Vector f' có thể được lượng tử hóa chính xác bằng bảng mã đại số như trong các hệ thống xử lý thông thường của ACELP. Tuy nhiên, vì các phần tử của f' là không tương quan, hàm tìm kiếm phức tạp như trong ACELP là không cần thiết, nhưng thuật toán đơn giản có thể được ứng dụng, như

- (a) Thiết lập độ khuếch đại ban đầu đến $g=1$
 - (b) Lượng tử hóa f' bằng $\hat{f}' = \text{round}(gf')$.
 - (c) Nếu số xung trong f' lớn hơn số lượng xác định trước p , $\|\hat{f}'\|_1 > p$, sau đó độ khuếch đại g tăng và quay lại bước b.
 - (d) Mặt khác, nếu số xung trong $\hat{f}'$ nhỏ hơn số lượng xác định trước p , $\|\hat{f}'\|_1 < p$, sau đó độ khuếch đại g giảm và quay lại bước b.
 - (e) Mặt khác, số xung trong $\hat{f}'$ bằng số lượng xác định trước p , $\|\hat{f}'\|_1 = p$, và quá trình có thể được dừng lại.
2. Bộ mã hóa số học có thể được sử dụng tương tự để được sử dụng như vậy trong sự lượng tử hóa của các vạch phổ trong TCX trong các tiêu chuẩn AMR-WB+ hoặc MPEG USAC.

Cần được chú ý rằng vì các phần tử của f là trực giao (như có thể được xem từ phương trình 12) và chúng có trọng số tương tự trong hàm mục tiêu của phương trình 12, chúng có thể được lượng tử hóa một cách riêng rẽ, và với cùng kích thước bước lượng tử hóa. Sự lượng tử hóa này sẽ tự động tìm kiếm giá trị tối ưu (lớn nhất) của hàm mục tiêu trong phương trình 12, mà có thể với độ chính xác lượng tử hóa đó. Nói cách khác, các thuật toán lượng tử hóa đã trình bày ở trên, cả hai sẽ quay lại sự lượng tử hóa tối ưu đối với phương trình 12.

Thuận lợi này của sự tối ưu được gắn liền với thực tế mà các phần tử của f có thể được xử lý riêng rẽ. Nếu phương pháp bảng mã có thể được sử dụng, mà các vector bảng mã c_k là không bình thường (có nhiều hơn một phần tử khác không), sau đó các vector bảng mã sẽ không có các phần tử độc lập nữa và thuận lợi của phép tìm thừa số ma trận bị mất.

Chú ý rằng phép tìm thừa số Vandermonde của ma trận Toeplitz có thể được chọn sao cho ma trận Vandermonde là ma trận biến đổi Fourier nhưng với các tần số phân bố không đều. Nói cách khác, ma trận Vandermonde tương ứng với sự biến đổi Fourier biến dạng tần số. Tiếp theo mà trong trường hợp này vector f tương ứng với sự trình bày miền tần số của tín hiệu dư trên tỷ lệ tần số biến dạng (xem "đặc tính thay đổi gốc" trong [8]).

Một cách quan trọng, chú ý rằng hệ quả này không được biết đến. Trong thực tế, kết quả này chỉ ra rằng nếu tín hiệu x được lọc với ma trận tích chập C , sau đó

$$\|C x\|^2 = \|D V x\|^2 \quad (13)$$

Trong đó V là (ví dụ, biến dạng) biến đổi Fourier (là ma trận Vandermonde với các phần tử trên vòng tròn đơn vị) và D ma trận đường chéo. Đó là, nếu được mong muốn để đo năng lượng của tín hiệu được lọc, năng lượng của tín hiệu tần số biến dạng có thể được đo một cách tương ứng. Ngược lại, sự đánh giá bất kỳ mà sẽ được thực hiện trong miền Fourier bị biến dạng, có thể được thực hiện một cách tương ứng trong miền thời gian được lọc. Do tính hai mặt của thời gian và tần số, sự tương ứng giữa tạo cửa sổ trong miền thời gian và thời gian biến dạng cũng tồn tại. Tuy nhiên, vấn đề thực tế là, sự tìm kiếm ma trận tích chập C mà đáp ứng các mối quan hệ trên là một vấn đề nhạy cảm về số lượng, do đó thường dễ dàng hơn để tìm kiếm sự thay thế các giải pháp gần đúng $\hat{C}$.

Quan hệ $\|C x\|^2 = \|D V x\|^2$ có thể được sử dụng để xác định vector bảng mã của bảng mã.

Đối với điều này, đầu tiên cần được lưu ý rằng ở đây, bằng H , ma trận tích chập tương tự phương trình 2 sẽ được ký hiệu thay vì C . Nếu, sau đó, một người muốn giảm thiểu nhiễu do lượng tử hóa $e = Hx - H\hat{x}$, năng lượng của nó có thể được đo:

$$\begin{aligned} \varepsilon^2 &= \|Hx - H\hat{x}\|^2 = \|H(x - \hat{x})\|^2 = (x - \hat{x})^T H^T H (x - \hat{x}) = (x - \hat{x})^T R (x - \hat{x}) \\ &= (x - \hat{x})^T V^H D V (x - \hat{x}) = \|D^{1/2} V (x - \hat{x})\|^2 = \|D^{1/2} V (x - \hat{x})\|^2 = \\ &= \|D^{1/2} (f - \hat{f})\|^2 = \|f' - \hat{f}'\|^2. \end{aligned} \quad (13a)$$

Hiện nay, sự mở rộng cho sự độc lập với khung được mô tả.

Khi tín hiệu tiếng nói được mã hóa được truyền tải trên đường dây truyền tải không hoàn thiện như sóng vô tuyến, các gói của dữ liệu đôi khi sẽ bị mất. Nếu các khung phụ thuộc vào nhau, sao cho gói N là cần thiết để giải mã hoàn thiện $N-1$, sau đó sự thiếu hụt gói $N-1$ sẽ tạo lỗi tổng hợp của cả hai gói $N-1$ và N . Mặt khác, nếu các khung là độc lập, sau đó sự thiếu hụt gói $N-1$ sẽ tạo lỗi tổng hợp của chỉ gói $N-1$. Do đó là quan trọng để các phương pháp thiết bị mà được tự do từ các phụ thuộc liên khung.

Trong các hệ thống ACELP quy ước, nguồn chính của phụ thuộc liên khung là LTP và đến mức độ nào đó cũng là LP. Một cách cụ thể, vì cả hai là các bộ lọc đáp ứng xung vô hạn (infinite impulse response - IIR), khung bị tạo lỗi sẽ gây ra mục cuối

"vô hạn" của các mẫu bị tạo lỗi. Trong thực tế, mục cuối này có thể kéo dài vài khung, mà là gây nhiễu về mặt cảm nhận.

Sử dụng khuôn khổ của sáng chế, đường dẫn qua mà sự phụ thuộc liên khung được phát ra có thể được lượng tử hóa bằng ZIR từ khung hiện vào trong khung tiếp theo được thực hiện. Để tránh sự phụ thuộc liên khung này, ba sự thay đổi để ACELP quy ước cần được tạo ra.

1. Khi tính toán AIR từ khung trước đó vào trong khung (con) hiện tại, cần được tính toán từ phần dư nguyên bản (không lượng tử hóa) được mở rộng với các số không, không phải từ phần dư lượng tử hóa. Theo cách này, các lỗi lượng tử hóa từ khung (con) trước sẽ không lan truyền vào trong khung (con) hiện tại.
2. Khi lượng tử hóa khung hiện tại, sai số trong ZIR vào trong khung tiếp theo giữa tín hiệu nguyên bản và lượng tử hóa phải được tính đến. Điều này có thể thực hiện bằng cách thay thế ma trận tương quan B với ma trận tự tương quan R, như được giải thích ở trên. Điều này đảm bảo rằng lỗi trong ZIR vào trong khung tiếp theo được tối thiểu hóa cùng với lỗi trong khung hiện tại.
3. Vì sự lan truyền lỗi là do cả LP và LTP, cả hai thành phần phải được chứa trong ZIR. Tức là, trong sự khác biệt với phương pháp quy ước mà ZIR được tính toán cho chỉ LP.

Nếu các lỗi lượng tử hóa của khung trước mà lượng tử hóa khung hiện tại không được tính đến, hiệu quả trong chất lượng cảm nhận của đầu ra bị mất. Do đó, có thể chọn để tính đến các lỗi trước khi không có nguy cơ của sự lan truyền lỗi. Ví dụ, hệ thống ACELP quy ước ứng dụng khung mà mọi khung 20ms được chia nhỏ thành 4 hoặc 5 khung con. LTP và phần dư được lượng tử hóa và được mã hóa riêng rẽ cho mỗi khung con, nhưng toàn bộ khung được truyền tải như một khối dữ liệu. Do đó, các khung con riêng lẻ không thể bị mất, mà chỉ các khung hoàn chỉnh. Sau đó chúng được yêu cầu để sử dụng các ZIR độc lập với khung chỉ ở các biên khung, nhưng các ZIR có thể được sử dụng với các phụ thuộc liên khung giữa các khung con còn lại.

Các phương án thay đổi các thuật toán ACELP quy ước bằng sự bao hàm của ảnh hưởng của đáp ứng xung của khung hiện tại vào trong khung tiếp theo, vào trong hàm mục tiêu của khung hiện tại. Trong hàm mục tiêu của bài toán tối ưu hóa, sự thay đổi này tương ứng với sự thay thế ma trận tương quan với ma trận tự tương quan mà có cấu trúc Hermit Toeplitz. Sự thay đổi này có các lợi ích sau:

1. Độ phức tạp tính toán và các yêu cầu bộ nhớ được giảm do cấu trúc Hermit Toeplitz được bổ sung của ma trận tự tương quan.
2. Mô hình cảm nhận tương tự sẽ được ứng dụng trên tất cả các mẫu, làm ra thiết kế và điều chỉnh mô hình cảm nhận đơn giản hơn, và ứng dụng của nó hiệu quả và phù hợp hơn.
3. Các tương quan liên khung có thể được tránh hoàn toàn trong sự lượng tử hóa của khung hiện tại, bằng cách tính đến chỉ đáp ứng xung không lượng tử hóa từ khung trước và đáp ứng xung lượng tử hóa vào trong khung tiếp theo. Sự cải thiện mạnh mẽ này của hệ thống mà tiêu hao gói được dự kiến.

Fig.2 minh họa bộ giải mã 220 để giải mã tín hiệu tiếng nói được mã hóa mà được mã hóa bằng thiết bị theo phương án được mô tả ở trên để thu được tín hiệu tiếng nói được giải mã. Bộ giải mã 220 được tạo cấu hình để nhận tín hiệu tiếng nói được mã hóa, trong đó tín hiệu tiếng nói được mã hóa bao gồm sự biểu thị của các vector bảng mã, được xác định bằng thiết bị mã hóa tín hiệu tiếng nói theo một trong số các phương án được mô tả ở trên, ví dụ, chỉ số của vector bảng mã được xác định. Hơn nữa, bộ giải mã 220 được tạo cấu hình để giải mã tín hiệu tiếng nói được mã hóa để thu được tín hiệu tiếng nói được giải mã phụ thuộc vào vector bảng mã.

Fig.3 minh họa hệ thống theo phương án. Hệ thống bao gồm thiết bị 210 theo một trong số các phương án được mô tả ở trên để mã hóa tín hiệu tiếng nói đầu vào để thu được tín hiệu tiếng nói được mã hóa. Tín hiệu tiếng nói được mã hóa bao gồm sự biểu thị của vector bảng mã xác định được xác định bằng thiết bị 210 để mã hóa tín hiệu tiếng nói, ví dụ, bao gồm chỉ số của vector bảng mã. Ngoài ra, hệ thống bao gồm bộ giải mã 220 theo phương án được mô tả ở trên để giải mã tín hiệu tiếng nói được mã hóa để thu được tín hiệu tiếng nói được giải mã. Bộ giải mã 220 được tạo cấu hình để nhận tín hiệu tiếng nói được mã hóa. Ngoài ra, bộ giải mã 220 được tạo cấu hình để giải mã tín hiệu tiếng nói được mã hóa để thu được tín hiệu tiếng nói được giải mã phụ thuộc vào vector bảng mã được xác định.

Mặc dù một vài khía cạnh đã được mô tả trong ngữ cảnh của thiết bị, các khía cạnh này cũng thể hiện sự mô tả của phương pháp tương ứng, mà khối hoặc thiết bị tương ứng với bước phương pháp hoặc đặc trưng của bước phương pháp. Tương tự, các khía cạnh được mô tả trong ngữ cảnh của bước phương pháp cũng thể hiện sự mô tả của khối hoặc mục tương ứng hoặc đặc điểm của thiết bị tương ứng.

Tín hiệu được phân tích của sáng chế có thể được lưu trữ trên vật ghi lưu trữ số hoặc có thể được truyền trên vật ghi truyền như vật ghi truyền không dây hoặc vật ghi truyền có dây như internet.

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 thi có thể được thực hiện sử dụng vật ghi lưu trữ số, ví dụ, đĩa mềm, DVD, CD, ROM, PROM, EPROM, EEPROM hoặc bộ nhớ FLASH, có các tín hiệu điều khiển có thể đọc được bằng điện được lưu trữ trên đó, mà kết hợp (hoặc có thể kết hợp) với hệ thống máy tính có thể lập trình được sao cho phương pháp tương ứng được thực hiện.

Một số phương án theo sáng chế gồm có sóng mang dữ liệu không tạm thời có các tín hiệu điều khiển đọc được nhờ điện tử, chúng có khả năng kết hợp với hệ thống máy tính lập trình được, do đó một trong số 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 số các phương pháp khi sản phẩm chương trình máy tính chạy trên máy tính. Mã chương trình có thể, ví dụ, được lưu trữ trên sóng mang máy đọc được.

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

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

Do đó, phương án khác của các phương pháp theo sáng chế là, sóng mang dữ liệu (hoặc vật ghi lưu trữ số, hoặc vật ghi máy tính đọc được) gồm có, đã được ghi lại trên đó, chương trình máy tính để thực hiện một trong số các phương pháp được mô tả ở đây.

Do đó, phương án khác của các phương pháp theo sáng chế là, dòng dữ liệu hoặc chuỗi các tín hiệu thể hiện chương trình máy tính để thực hiện một trong số các phương pháp đã được mô tả ở đây. Ví dụ có thể tạo cấu hình dòng dữ liệu hoặc chuỗi các tín hiệu để được truyền thông qua sự kết nối truyền thông dữ liệu, ví dụ thông qua internet.

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

Phương án khác gồm có máy tính đã cài trong đó chương trình máy tính để thực hiện một trong số các phương pháp được mô tả ở đây.

Trong một số phương án, thiết bị logic lập trình được (ví dụ, mảng cổng 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. Theo 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 số các phương pháp được mô tả ở đây. Thông thường, các phương pháp được thực hiện một cách ưu tiên bởi thiết bị phần 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 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, sáng chế chỉ được giới hạn bởi phạm vi của các điểm yêu cầu bảo hộ và không được giới hạn bởi các chi tiết cụ thể được thể hiện trong bản mô tả và phần giải thích của các phương án trong bản mô tả.

YÊU CẦU BẢO HỘ

1. Thiết bị mã hóa tín hiệu tiếng nói bằng cách xác định vector bảng mã của thuật toán mã hóa tiếng nói, trong đó thiết bị bao gồm:

bộ xác định ma trận (110) để xác định ma trận tự tương quan R , và

bộ xác định vector bảng mã (120) để xác định vector bảng mã phụ thuộc vào ma trận tự tương quan R ,

trong đó bộ xác định ma trận (110) được tạo cấu hình để xác định ma trận tự tương quan R bằng cách xác định các hệ số vector của vector r , trong đó ma trận tự tương quan R bao gồm nhiều hàng và nhiều cột, trong đó vector r chỉ ra một trong số các cột hoặc một trong số các hàng của ma trận tự tương quan R , trong đó

$$R(i, j) = r(|i - j|),$$

trong đó $R(i, j)$ chỉ ra các hệ số của ma trận tự tương quan R , trong đó i là chỉ số thứ nhất chỉ ra một trong số nhiều hàng của ma trận tự tương quan R , và trong đó j là chỉ số thứ hai chỉ ra một trong số nhiều cột của ma trận tự tương quan R ,

trong đó bộ xác định vector bảng mã (120) được tạo cấu hình để xác định vector bảng mã bằng cách áp dụng công thức

$$f(\hat{e}) = \frac{(d^T \hat{e})^2}{\hat{e}^T R \hat{e}}$$

trong đó R là ma trận tự tương quan, và trong đó $\hat{e}$ là một trong số các vector bảng mã của thuật toán mã hóa tiếng nói, trong đó $f(\hat{e})$ là sự tương quan được chuẩn hóa, và trong đó d^T được xác định theo

$$\frac{(e^T R \hat{e})^2}{(\hat{e}^T R \hat{e})} = \frac{(d^T e)^2}{(\hat{e}^T R \hat{e})},$$

trong đó e là tín hiệu dư ban đầu, không được lượng tử hóa.

2. Thiết bị theo điểm 1,

trong đó bộ xác định vector bảng mã (120) được tạo cấu hình để xác định vector bảng mã $\hat{e}$ đó của thuật toán mã hóa tiếng nói mà tối đa hóa sự tương quan được chuẩn hóa

$$f(\hat{e}) = \frac{(d^T \hat{e})^2}{\hat{e}^T R \hat{e}}.$$

3. Thiết bị theo điểm 1 hoặc 2,

trong đó bộ xác định ma trận (110) được tạo cấu hình để xác định các hệ số vector của vector r bằng cách áp dụng công thức:

$$r(k) = h(k) * h(-k) = \sum_l h(l)h(l-k)$$

trong đó $h(k)$ chỉ ra đáp ứng xung có trọng số theo cảm nhận của mô hình dự báo tuyến tính, và trong đó k là chỉ số là số nguyên, và trong đó l là chỉ số là số nguyên.

4. Thiết bị theo một trong số các điểm nêu trên,

trong đó bộ xác định ma trận (110) được tạo cấu hình để xác định ma trận tự tương quan R phụ thuộc vào bộ dự báo tuyến tính có trọng số theo cảm nhận.

5. Thiết bị theo một trong số các điểm nêu trên, trong đó bộ xác định vector bằng mã (120) được tạo cấu hình để phân tích ma trận tự tương quan R bằng cách thực hiện sự phân tích ma trận.

6. Thiết bị theo điểm 5, trong đó bộ xác định vector bằng mã (120) được tạo cấu hình để thực hiện sự phân tích ma trận để xác định ma trận đường chéo D cho việc xác định vector bằng mã.

7. Thiết bị theo điểm 6,

trong đó bộ xác định vector bằng mã (120) được tạo cấu hình để xác định vector bằng mã bằng cách sử dụng

$$\frac{(f^H D \hat{f})^2}{\hat{f}^H D \hat{f}},$$

trong đó D là ma trận đường chéo, trong đó f là vector thứ nhất, và trong đó $\hat{f}$ là vector thứ hai.

8. Thiết bị theo điểm 6 hoặc 7, trong đó bộ xác định vector bằng mã (120) được tạo cấu hình để thực hiện phép tìm thừa số Vandermonde trên ma trận tự tương quan R để phân tích ma trận tự tương quan R để thực hiện sự phân tích ma trận để xác định ma trận đường chéo D cho việc xác định vector bằng mã.

9. Thiết bị theo một trong số các điểm từ 6 đến 8, trong đó bộ xác định vector bằng mã (120) được tạo cấu hình để sử dụng phương trình

$$\|Cx\|^2 = \|DVx\|^2$$

để xác định vector bảng mã, trong đó C chỉ ra ma trận tích chập, trong đó V chỉ ra biến đổi Fourier, và trong đó x chỉ ra tín hiệu tiếng nói.

10. Thiết bị theo một trong số các điểm từ 6 đến 9, trong đó bộ xác định vector bảng mã (120) được tạo cấu hình để thực hiện sự phân tích giá trị đơn trên ma trận tự tương quan R để phân tích ma trận tự tương quan R để thực hiện sự phân tích ma trận để xác định ma trận đường chéo D cho việc xác định vector bảng mã.

11. Thiết bị theo một trong số các điểm từ 6 đến 9, trong đó bộ xác định vector bảng mã (120) được tạo cấu hình để thực hiện sự phân tích Cholesky trên ma trận tự tương quan R để phân tích ma trận tự tương quan R để thực hiện sự phân tích ma trận để xác định ma trận đường chéo D cho việc xác định vector bảng mã.

12. Thiết bị theo một trong số các điểm nêu trên, trong đó bộ xác định vector bảng mã (120) được tạo cấu hình để xác định vector bảng mã phụ thuộc vào đáp ứng xung bằng không của tín hiệu tiếng nói.

13. Thiết bị theo một trong số các điểm nêu trên,

trong đó thiết bị là bộ mã hóa để mã hóa tín hiệu tiếng nói bằng cách sử dụng mã hóa tiếng nói dự báo tuyến tính kích thích mã đại số, và

trong đó bộ xác định vector bảng mã (120) được tạo cấu hình để xác định vector bảng mã dựa trên ma trận tự tương quan R như vector bảng mã của bảng mã đại số.

14. Hệ thống mã hóa tín hiệu tiếng nói để thu được tín hiệu tiếng nói được mã hóa và giải mã tín hiệu tiếng nói được mã hóa, thiết bị bao gồm:

thiết bị (210) theo một trong số các điểm từ 1 đến 13 để mã hóa tín hiệu tiếng nói đầu vào để thu được tín hiệu tiếng nói được mã hóa, và

bộ giải mã (220) để giải mã tín hiệu tiếng nói được mã hóa để thu được tín hiệu tiếng nói được giải mã,

trong đó bộ giải mã (220) được tạo cấu hình để nhận tín hiệu tiếng nói được mã hóa, trong đó tín hiệu tiếng nói được mã hóa bao gồm sự biểu thị vector bảng mã, được xác định bởi thiết bị theo một trong số các điểm từ 1 đến 13,

trong đó bộ giải mã (220) được tạo cấu hình để giải mã tín hiệu tiếng nói được mã hóa để thu được tín hiệu tiếng nói được giải mã phụ thuộc vào vector bảng mã.

15. Phương pháp mã hóa tín hiệu tiếng nói bằng cách xác định vector bảng mã của thuật toán mã hóa tiếng nói, trong đó phương pháp bao gồm các bước:

xác định ma trận tự tương quan R , và

xác định vector bảng mã phụ thuộc vào ma trận tự tương quan R ,

trong đó việc xác định ma trận tự tương quan R bao gồm xác định các hệ số vector của vector r , trong đó ma trận tự tương quan R bao gồm nhiều hàng và nhiều cột, trong đó vector r chỉ ra một trong số các cột hoặc một trong số các hàng của ma trận tự tương quan R , trong đó

$$R(i, j) = r(|i - j|),$$

trong đó $R(i, j)$ chỉ ra các hệ số của ma trận tự tương quan R , trong đó i là chỉ số thứ nhất chỉ ra một trong số nhiều hàng của ma trận tự tương quan R , và trong đó j là chỉ số thứ hai chỉ ra một trong số nhiều cột của ma trận tự tương quan R ,

trong đó việc xác định vector bảng mã được thực hiện bằng cách sử dụng công thức

$$f(\hat{e}) = \frac{(d^T \hat{e})^2}{\hat{e}^T R \hat{e}}$$

trong đó R là ma trận tự tương quan, và trong đó $\hat{e}$ là một trong số các vector bảng mã của thuật toán mã hóa tiếng nói, trong đó $f(\hat{e})$ là sự tương quan được chuẩn hóa, và trong đó d^T được xác định theo

$$\frac{(e^T R \hat{e})^2}{(\hat{e}^T R \hat{e})} = \frac{(d^T e)^2}{(\hat{e}^T R \hat{e})},$$

trong đó e là tín hiệu dư ban đầu, không được lượng tử hóa.

16. Phương pháp mã hóa tín hiệu tiếng nói đầu vào để thu được tín hiệu tiếng nói được mã hóa và giải mã tín hiệu tiếng nói được mã hóa, trong đó phương pháp bao gồm:

mã hóa tín hiệu tiếng nói đầu vào theo phương pháp theo điểm 15 để thu được tín hiệu tiếng nói được mã hóa, trong đó tín hiệu tiếng nói được mã hóa bao gồm sự biểu thị vector bảng mã, và

giải mã tín hiệu tiếng nói được mã hóa để thu được tín hiệu tiếng nói được giải mã phụ thuộc vào vector bảng mã.

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 chạy trên máy tính hoặc bộ xử lý tín hiệu.

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

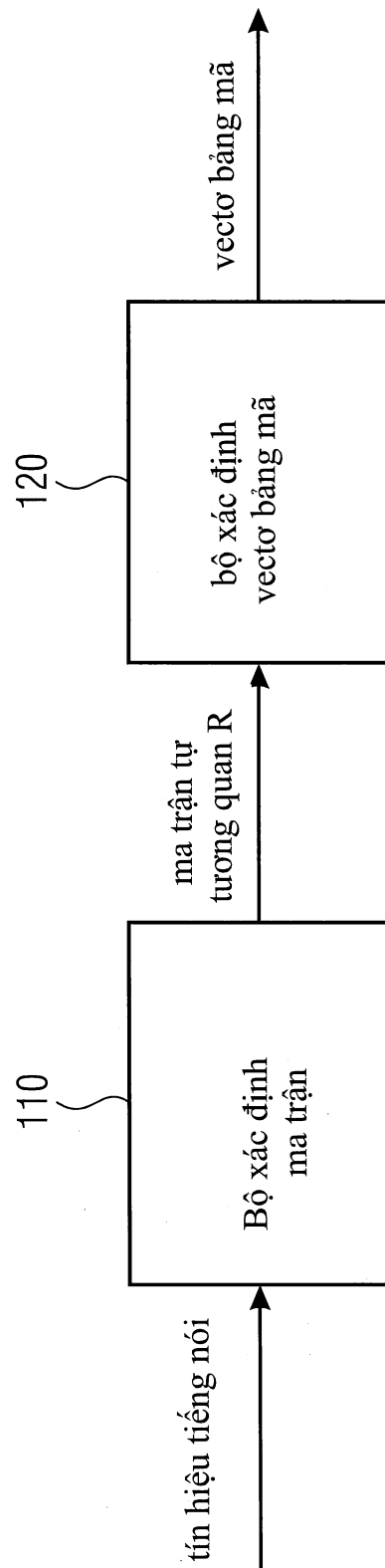


FIG 1

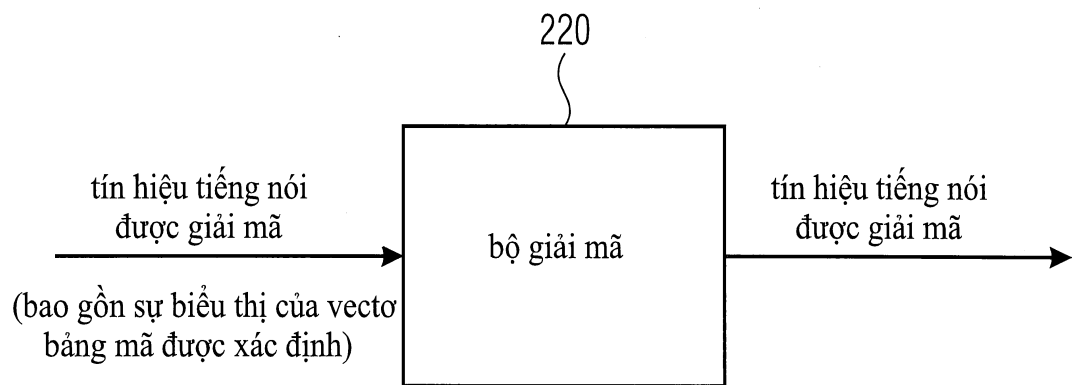


FIG 2

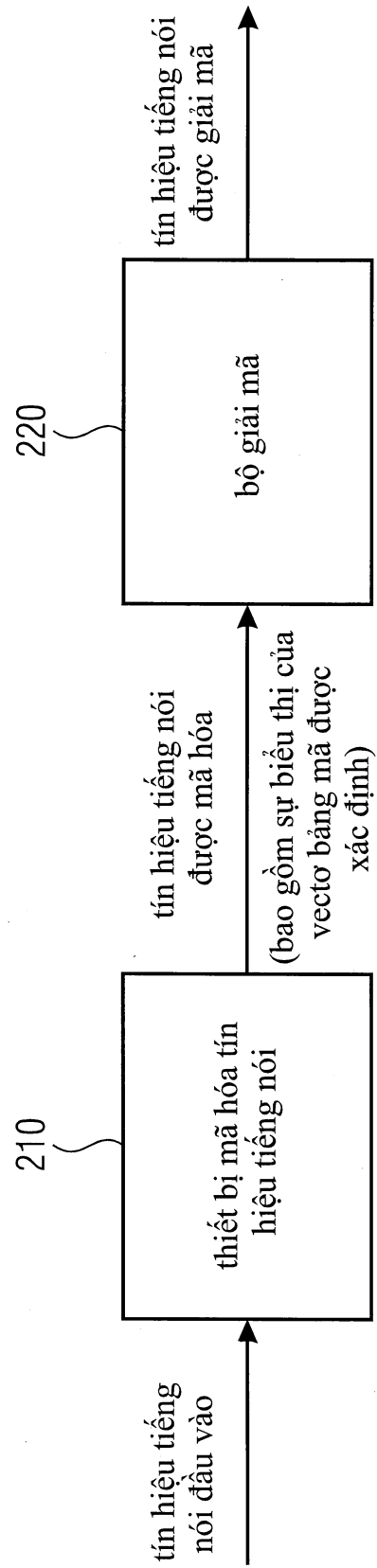


FIG 3