



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

(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN) (11)
CỤC SỞ HỮU TRÍ TUỆ



1-0043602

(51)^{2020.01} G10L 19/008; G10L 19/06; G10L 19/02 (13) B

-
- (21) 1-2021-00274 (22) 19/06/2019
(86) PCT/EP2019/066228 19/06/2019 (87) WO2019/243434 26/12/2019
(30) 18179373.8 22/06/2018 EP
(45) 25/02/2025 443 (43) 25/05/2021 398A
(73) Fraunhofer-Gesellschaft zur Foerderung der angewandten Forschung e. V. (DE)
Hansastrasse 27c, 80686 Muenchen, Germany
(72) BÜTHE, Jan (DE); FOTOPOULOU, Eleni (GR); KORSE, Srikanth (IN); MABEN,
Pallavi (IN); MULTRUS, Markus (DE); REUTELHUBER, Franz (DE).
(74) CÔNG TY LUẬT TRÁCH NHIỆM HỮU HẠN AMBYS HÀ NỘI (AMBYS
HANOI)
-

- (54) THIẾT BỊ SO SÁNH, BỘ MÃ HÓA ĐA KÊNH, BỘ GIẢI MÃ CHO CÁC TÍN
HIỆU ÂM THANH ĐA KÊNH, VÀ PHƯƠNG PHÁP SO SÁNH

(21) 1-2021-00274

(57) Sáng chế đề cập đến thiết bị so sánh, bộ mã hóa đa kênh, bộ giải mã cho các tín hiệu âm thanh đa kênh, và phương pháp so sánh. Trong sự mã hóa âm thanh đa kênh, hiệu quả tính toán được cải thiện đạt được bằng cách tính toán các tham số so sánh cho sự bù ITD giữa hai kênh bất kỳ trong miền tần số cho bộ mã hóa âm thanh theo tham số. Điều này có thể làm giảm hiệu quả bất lợi về các ước lượng tham số bộ mã hóa.

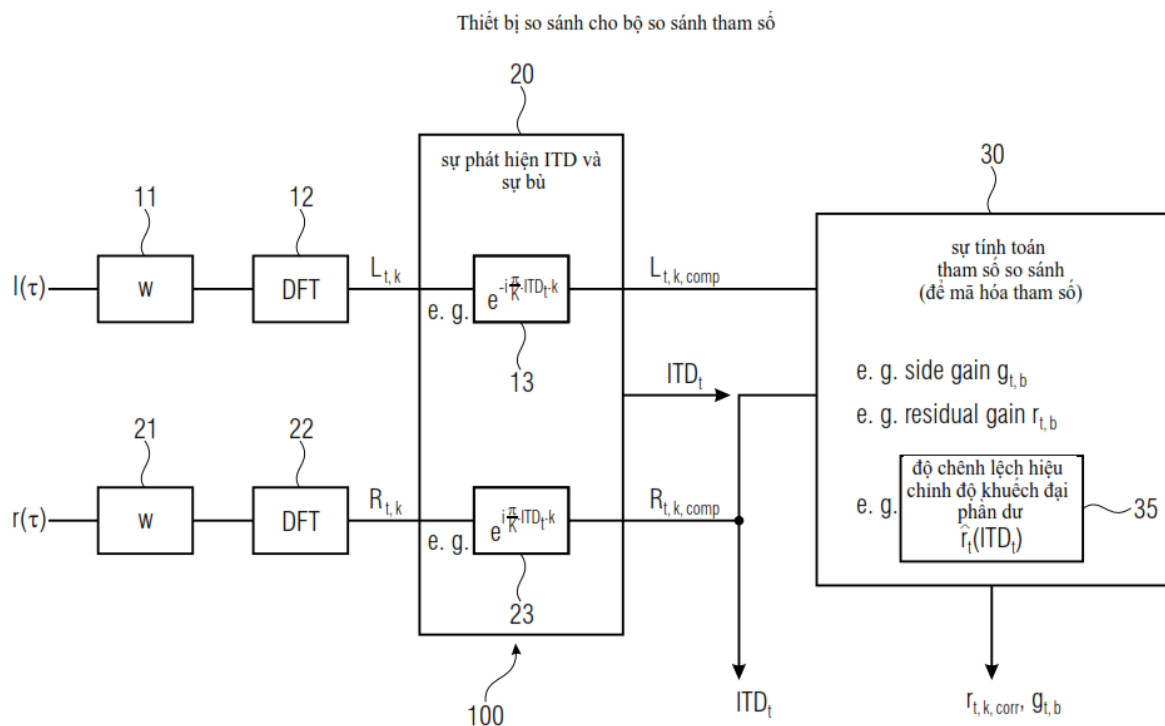


Fig. 1

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến sự mã hóa âm thanh đa kênh theo tham số.

Tình trạng kỹ thuật của sáng chế

Phương pháp theo tình trạng kỹ thuật của sáng chế cho sự mã hóa theo tham số tổng hao của các tín hiệu âm lập thể tại tốc độ bit thấp dựa trên âm lập thể theo tham số như được đặt tiêu chuẩn trong MPEG-4 Phần 3 [1]. Ý tưởng chung là để giảm số lượng các kênh của hệ thống đa kênh bằng cách tính toán tín hiệu trộn giảm từ hai kênh đầu vào sau khi trích các tham số âm lập thể/không gian mà được gửi như thông tin phụ đến bộ giải mã. Các tham số âm lập thể/không gian có thể thường bao gồm sự chênh lệch mức liên kênh *ILD*, sự chênh lệch pha liên kênh *IPD*, và sự nhất quán liên kênh *ICC*, mà có thể được tính toán trong các băng phụ và chụp ảnh không gian đến mức nhất định.

Tuy nhiên, phương pháp này không thể bù hoặc tổng hợp các chênh lệch thời gian liên kênh (inter-channel-time-differences - *ITD*) mà là, ví dụ, mong muốn trộn giảm hoặc tái sản xuất lời nói được ghi lại bởi hoặc tổng hợp các cảnh được kết xuất lập thể. Sự tổng hợp *ITD* đã được chỉ ra trong sự mã hóa tín hiệu điều khiển lập thể (binaural cue coding - *BCC*) [2], mà thường sử dụng các tham số *ILD* và *ICC*, trong khi các *ITD* được ước lượng và sự căn chỉnh kênh được thực hiện trong miền tần số.

Mặc các bộ ước lượng *ITD* miền thời gian tồn tại, thường ưu tiên hơn cho sự ước lượng *ITD* sử dụng sự biến đổi thời gian thành tần số, điều này cho phép lọc phổ của hàm tương quan chéo và cũng hiệu quả về mặt tính toán. Đối với các lý do phức tạp, mong muốn sử dụng sự biến đổi tương tự mà cũng được sử dụng để trích các tham số lập thể/không gian và có khả năng cho các kênh trộn giảm, mà cũng được thực hiện trong phương pháp *BCC*.

Tuy nhiên, điều này đi cùng với hạn chế: sự ước lượng chính xác các tham số âm lập thể được thực hiện một cách lý tưởng trên các kênh được căn chỉnh. Nhưng nếu các kênh được căn chỉnh trong miền tần số, ví dụ, bởi sự dịch chuyển theo vòng tròn trong miền tần số, điều này có thể gây ra sự chênh lệch trong các cửa sổ phân tích, mà ảnh

hưởng tiêu cực đến các ước lượng tham số. Trong trường hợp BCC, điều này ảnh hưởng chính đến *ICC*, trong đó các độ lệch của số tăng thậm chí đẩy giá trị *ICC* về phía 0 ngay cả khi các tín hiệu đầu vào thực sự hoàn toàn đồng nhất.

Tài liệu tham khảo

[1] MPEG-4 High Efficiency Advanced Audio Coding (HE-AAC) v2

[2] Jürgen Herre, *FROM JOINT STEREO TO SPATIAL AUDIO CODING - RECENT PROGRESS AND STANDARDIZATION*, Proc. of the 7th Int. Conference on digital Audio Effects (DAFX-04), Naples, Italy, October 5-8, 2004

[3] Christoph Tourney and Christof Faller, *Improved Time Delay Analysis/Synthesis for Parametric Stereo Audio Coding*, AES Convention Paper 6753, 2006

[4] Christof Faller and Frank Baumgarte, *Binaural Cue Coding Part II: Schemes and Applications*, IEEE Transactions on Speech and Audio Processing, Vol. 11, No. 6, November 2003

Bản chất kỹ thuật của sáng chế

Do đó, mục đích của sáng chế là cung cấp khái niệm cho sự tính toán tham số trong sự mã hóa âm thanh đa kênh mà có thể bù sự chênh lệch thời gian liên kênh mà tránh các hiệu quả bất lợi cho các ước lượng tham số không gian.

Mục đích này đạt được bởi đối tượng của các điểm yêu cầu bảo hộ độc lập.

Sáng chế dựa trên sự phát hiện rằng trong sự mã hóa âm thanh đa kênh, hiệu quả tính toán được cải thiện có thể đạt được bằng cách tính toán ít nhất một tham số so sánh cho sự bù *ITD* giữa hai kênh bất kỳ trong miền tần số sẽ được sử dụng bởi bộ mã hóa âm thanh theo tham số. Ít nhất một tham số so sánh đã nêu có thể được sử dụng bởi bộ mã hóa tham số để giảm các hiệu quả bất lợi nêu trên các ước lượng tham số không gian.

Phương án có thể bao gồm bộ mã hóa âm thanh tham số mà nhắm vào việc biểu diễn âm lập thể hoặc nội dung không gian chung bằng ít nhất một tín hiệu trộn giảm và các tham số lập thể hoặc không gian bổ sung. Trong số các tham số lập thể/không gian này có thể là các *ITD*, mà có thể được ước lượng và bù trong miền tần số, trước khi tính

toán các tham số lập thể/không gian còn lại. Quy trình này có thể nghiêng về các tham số lập thể/không gian khác, vấn đề rằng cách khác sẽ được giải quyết bằng cách tốn kém hơn bằng cách tính toán sự biến đổi tần số thành thời gian. Trong phương án đã nêu, vấn đề này có thể được giảm bằng cách áp dụng sơ đồ hiệu chỉnh về tính toán mà có thể sử dụng giá trị của *ITD* và dữ liệu nhất định của sự biến đổi cơ sở.

Phương án đề cập đến bộ giải mã âm thanh theo tham số tổn hao mà có thể dựa trên phương pháp biến đổi giữa/bên được gán trọng số, có thể sử dụng các tham số lập thể/không gian *IPD*, *ITD*, cũng như hai thừa số khuếch đại và có thể vận hành trong miền tần số. Các phương án khác có thể sử dụng sự biến đổi khác và có thể sử dụng các tham số không gian khác nhau cho phù hợp.

Trong phương án, bộ mã hóa âm thanh tham số có thể cùng có khả năng bù và tổng hợp các *ITD* trong miền tần số. Có thể tạo đặc trưng sơ đồ hiệu chỉnh độ khuếch đại hiệu quả về tính toán mà giảm các hiệu quả bất lợi của độ lệch của số đã nêu. Cũng như sơ đồ hiệu chỉnh cho bộ mã hóa BCC được gợi ý.

Mô tả vắn tắt các hình vẽ

Các sự thực thi có lợi của sáng chế là đối tượng của các điểm phụ thuộc. Các phương án được ưu tiên của sáng chế được mô tả sau đây đối với các hình vẽ, trong đó:

Fig.1 thể hiện sơ đồ khối của thiết bị so sánh cho bộ mã hóa theo tham số theo phương án của sáng chế.

Fig.2 thể hiện sơ đồ khối của bộ mã hóa theo tham số theo phương án của sáng chế.

Fig.3 thể hiện sơ đồ khối của bộ giải mã theo tham số theo phương án của sáng chế.

Mô tả chi tiết sáng chế

Fig.1 thể hiện thiết bị so sánh 100 cho tín hiệu âm thanh đa kênh. Như được thể hiện, có thể bao gồm đầu vào cho các tín hiệu âm thanh cho cặp các kênh lập thể, cụ thể tín hiệu kênh âm thanh bên trái $l(\tau)$ và tín hiệu kênh âm thanh bên phải $r(\tau)$. Các phương án khác, có thể đương nhiên bao gồm nhiều kênh để chụp lại đặc tính không gian của các nguồn âm.

Trước khi biến đổi các tín hiệu âm thanh miền thời gian miền thời gian $l(\tau)$, $r(\tau)$ thành miền tần số, giống với các hàm cửa sổ chồng lấp giống nhau 11, 21 $w(\tau)$ có thể được áp dụng cho các tín hiệu kênh đầu vào bên trái và bên phải $l(\tau)$, $r(\tau)$ một cách lần lượt. Hơn nữa, trong các phương án, lượng nhất định sự đệm số 0 có thể được thêm vào mà cho phép các sự dịch chuyển trong miền tần số. Tiếp theo, các tín hiệu âm thanh được tạo cửa sổ có thể được cung cấp đến các khối biến đổi Fourier rời rạc (discrete Fourier transform - DFT) tương ứng 12, 22 để thực hiện sự biến đổi thời gian thành tần số tương ứng. Các tín hiệu này có thể tạo ra các ngăn thời gian-tần số $L_{t,k}$ và $R_{t,k}$, $k = 0, \dots, K - 1$ cũng như các biến đổi tần số của các tín hiệu âm thanh cho cặp kênh.

Các biến đổi tần số đã nêu $L_{t,k}$ và $R_{t,k}$, có thể được cung cấp cho khối bù và phát hiện *ITD* 20. Khối bù có thể được tạo cấu hình để suy ra, để biểu diễn *ITD* giữa các tín hiệu âm thanh cho cặp các kênh, tham số *ITD*, ở đây ITD_t , sử dụng các biến đổi tần số $L_{t,k}$ và $R_{t,k}$ của các tín hiệu âm thanh của các cặp kênh trong các cửa sổ phân tích đã nêu $w(\tau)$. Các phương án khác có thể sử dụng các phương pháp khác nhau để suy ra tham số *ITD* mà có thể cũng được xác định trước các khối DFT trong miền thời gian.

Sự suy ra tham số *ITD* để tính toán *ITD* có thể bao gồm sự tính toán hàm tự tương quan hoặc tương quan chéo có thể được gán trọng số. Thông thường, điều này có thể được tính toán ở các ngăn thời gian - tần số $L_{t,k}$ và $R_{t,k}$ bằng cách áp dụng sự biến đổi Fourier rời rạc nghịch đảo (inverse discrete Fourier transform - IDFT) đến số hạng $(L_{t,k}R_{t,k}^*\omega_{t,k})_k$.

Cách thích hợp để bù *ITD* được đo để thực hiện sự căn chỉnh kênh trong miền thời gian và sau đó áp dụng sự biến đổi thời gian thành tần số tương tự một lần nữa đến (các) kênh được dịch chuyển để thu được các ngăn thời gian tần số được bù *ITD*. Tuy nhiên, để tiết kiệm độ phức tạp, quy trình này có thể được làm gần đúng bằng cách thực hiện sự dịch chuyển theo hình tròn trong miền tần số. Tương ứng, sự bù *ITD* có thể được thực hiện bởi khối bù và phát hiện *ITD* 20 trong miền tần số, ví dụ, bằng cách thực hiện các sự dịch chuyển hình tròn bằng các khối dịch chuyển hình tròn 13 và 23 tương ứng để tạo ra

$$L_{t,k,comp} \leftarrow e^{-i\frac{\pi}{K}ITD_t k} L_{t,k} \quad (1)$$

và

$$R_{t,k,comp} \leftarrow e^{i\frac{\pi}{K}ITD_t k} R_{t,k} \quad (2),$$

trong đó ITD_t có thể biểu thị ITD cho khung t trong các mẫu.

Trong phương án, điều này có thể tiến tới kênh trễ và có thể làm trễ kênh trễ bởi các mẫu $ITD_t/2$. Tuy nhiên, trong các phương án - nếu độ trễ là nghiêm trọng - có thể có lợi ích đối với việc tiến tới kênh trễ bằng các mẫu ITD_t , mà không tăng độ trễ của hệ thống.

Kết quả là, khôi phục và phát hiện ITD 20 có thể bù ITD cho cặp các kênh trong miền tần số bằng (các) dịch chuyển hình tròn sử dụng tham số ITD ITD_t để tạo ra cặp các biến đổi tần số được bù ITD $L_{t,k,comp}$, $R_{t,k,comp}$ tại đầu ra của mình. Hơn nữa, khôi phục và phát hiện ITD 20 có thể phát ra tham số ITD được suy ra, cụ thể ITD_t , ví dụ, đối với sự truyền dẫn bởi bộ mã hóa theo tham số.

Như được thể hiện trên Fig.1, khối so sánh và tính toán tham số không gian 30 có thể thu tham số ITD ITD_t và cặp các biến đổi tần số được bù ITD $L_{t,k,comp}$, $R_{t,k,comp}$ như các tín hiệu đầu vào của chúng. Khối so sánh và tính toán tham số không gian 30 có thể sử dụng một số hoặc tất cả các tín hiệu đầu vào của chúng để trích các tham số lập thể/không gian của tín hiệu âm thanh đa kênh như sự chênh lệch liên pha IPD .

Hơn nữa, khối so sánh và tính toán tham số không gian 30 có thể tạo ra - dựa trên tham số ITD ITD_t và cặp các biến đổi tần số được bù ITD $L_{t,k,comp}$, $R_{t,k,comp}$ - ít nhất một tham số so sánh, ở đây hai thừa số khuếch đại $g_{t,b}$ và $r_{t,b,corr}$, cho bộ mã hóa tham số. Các phương án khác có thể sử dụng một cách bổ sung hoặc thay thế các biến đổi tần số $L_{t,k}$, $R_{t,k}$ và/hoặc các tham số không gian/lập thể được trích trong khối so sánh và tính toán tham số không gian 30 để tạo ra ít nhất một tham số so sánh.

Ít nhất một tham số so sánh có thể phục vụ như phần của sơ đồ hiệu chỉnh hiệu quả về tính toán để giảm các hiệu quả bất lợi của độ lệch nêu trên trong các cửa sổ phân tích $w(\tau)$ trên tham số không gian/lập thể cho bộ mã hóa tham số, độ lệch đã nêu được gây ra bởi sự căn chỉnh các kênh bằng các dịch chuyển hình tròn trong miền DFT trong khối phát hiện và bù ITD 20. Trong phương án, ít nhất một tham số so sánh có thể được

tính toán để khôi phục các tín hiệu âm thanh của cặp các kênh tại bộ giải mã, ví dụ, từ tín hiệu trộn giảm.

Fig.2 thể hiện phương án của bộ mã hóa tham số 200 cho các tín hiệu âm thanh lập thể mà thiết bị so sánh 100 của Fig.1 có thể được sử dụng để cung cấp tham số ITD ITD_t , cặp các biến đổi tần số được bù ITD $L_{t,k,comp}$, $R_{t,k,comp}$ và các tham số bù $r_{t,b,corr}$ và $g_{t,b}$.

Bộ mã hóa tham số 200 có thể tạo ra tín hiệu trộn giảm $DMX_{t,k}$ trong khối trộn giảm 40 cho các tín hiệu kênh đầu vào bên trái và bên phải $l(\tau)$, $r(\tau)$ sử dụng các biến đổi tần số được bù ITD $L_{t,k,comp}$, $R_{t,k,comp}$ như đầu vào. Các phương án khác có thể sử dụng bổ sung hoặc thay thế các biến đổi tần số $L_{t,k}$, $R_{t,k}$ để tạo ra tín hiệu trộn giảm $DMX_{t,k}$.

Bộ mã hóa tham số 200 có thể tính toán các tham số lập thể - như, ví dụ, IPD - trên cơ sở khung trong khối so sánh và tính toán tham số không gian 30. Các phương án khác có thể xác định các tham số âm lập thể/không gian bổ sung. Quy trình mã hóa bộ mã hóa tham số 200 phương án trên Fig.2 có thể tuân theo nghiêm ngặt các bước sau, mà được mô tả chi tiết như sau.

1. Sự biến đổi thời gian thành tần số các tín hiệu đầu vào sử dụng các DFT được tạo cửa sổ

trong cửa sổ và các khối DFT 11, 12, 21, 22

2. Sự ước lượng ITD và bù trong miền tần số

trong khối phát hiện và bù ITD 20

3. Sự trích tham số âm lập thể và sự tính toán tham số so sánh

trong khối so sánh và tính toán tham số không gian 30

4. Trộn giảm

trong khối trộn giảm 40

5. Sự biến đổi tần số thành thời gian theo sau bằng sự tạo cửa sổ và chồng lấp cộng

trong khối IDFT 50

Phương án bộ mã hóa âm thanh theo tham số 200 trên Fig.2 có thể được dựa trên sự biến đổi giữa/bên được gán trọng số của các kênh đầu vào trong miền tần số sử dụng các biến đổi tần số được bù ITD $L_{t,k,comp}$, $R_{t,k,comp}$ cũng như ITD là đầu vào. Có thể còn tính toán các tham số âm lập thể/không gian, như IPD , cũng như hai thừa số khuếch đại chụp ảnh âm lập thể. Điều này có thể giảm các hiệu quả bất lợi của độ lệch cửa sổ đề cập trên đây.

Đối với sự trích tham số không gian trong khối so sánh và tính toán tham số không gian 30, ITD các ngăn thời gian - tần số được bù $L_{t,k,comp}$ và $R_{t,k,comp}$ có thể được nhóm trong các băng phụ, và đối với mỗi băng phụ, độ chênh lệch liên pha IPD và hai thừa số khuếch đại có thể được tính toán. Để I_b biểu thị các chỉ số của các ngăn tần số trong băng phụ b . Khi đó IPD có thể được tính toán như

$$IPD_{t,b} = \arg(\sum_{k \in I_b} L_{t,k,comp} R_{t,k,comp}^*) \quad (3).$$

Hai thừa số khuếch đại nêu trên có thể liên quan đến các biến đổi giữa/bên được bù pha theo băng của cặp các biến đổi tần số được bù ITD $L_{t,k,comp}$ và $R_{t,k,comp}$ được cho bởi phương trình (4) và (5) như

$$M_{t,k} = L_{t,k,comp} + e^{iIPD_{t,b}} R_{t,k,comp} \quad (4)$$

và

$$S_{t,k} = L_{t,k,comp} - e^{iIPD_{t,b}} R_{t,k,comp} \quad (5)$$

đối với $k \in I_b$.

Thừa số khuếch đại thứ nhất $g_{t,b}$ trong số các thừa số khuếch đại đã nêu có thể được đề cập đến như độ khuếch đại dự báo tối ưu cho sự dự báo theo băng của sự biến đổi tín hiệu bên S_t từ sự biến đổi tín hiệu giữa M_t trong phương trình (6):

$$S_{t,k} = g_{t,b} M_{t,k} + \rho_{t,k} \quad (6)$$

sao cho năng lượng của phần dư dự báo $\rho_{t,k}$ trong phương trình (6) như được cho bởi phương trình (7) như

$$\sum_{k \in I_b} |\rho_{t,k}|^2 \quad (7)$$

là cực tiểu. Thừa số khuếch đại thứ nhất $g_{t,b}$ có thể được đề cập đến như độ khuếch đại bên.

Thừa số khuếch đại thứ hai $r_{t,b}$ mô tả tỉ lệ của năng lượng của phần dư dự báo $\rho_{t,k}$ liên quan đến năng lượng của sự biến đổi tín hiệu giữa $M_{t,k}$ được cho bởi phương trình (8) như

$$r_{t,b} = \left(\frac{\sum_{k \in I_b} |\rho_{t,k}|^2}{\sum_{k \in I_b} |M_{t,k}|^2} \right)^{1/2} \quad (8)$$

và có thể được đề cập đến như độ khuếch đại phần dư. Độ khuếch đại phần dư $r_{t,b}$ có thể được sử dụng tại bộ giải mã như bộ giải mã theo phương án trên Fig.3 để tạo hình sự thay thế phù hợp cho phần dư dự báo $\rho_{t,k}$ của biến đổi giữa/bên.

Trong bộ mã hóa theo phương án được thể hiện trên Fig.2, cả hai thừa số khuếch đại $g_{t,b}$ và $r_{t,b}$ có thể được tính toán như các tham số so sánh trong khối so sánh và tính toán tham số không gian 30 sử dụng các năng lượng $E_{L,t,b}$ và $E_{R,t,b}$ của các biến đổi tần số được bù ITD $L_{t,k,comp}$ và $R_{t,k,comp}$ được cho trong phương trình (9) như

$$E_{L,t,b} = \sum_{k \in I_b} |L_{t,k,comp}|^2 \quad \text{và} \quad E_{R,t,b} = \sum_{k \in I_b} |R_{t,k,comp}|^2 \quad (9)$$

và giá trị tuyệt đối của sản phẩm bên trong

$$X_{L/R,t,b} = \left| \sum_{k \in I_b} L_{t,k,comp} R_{t,k,comp}^* \right| \quad (10)$$

được cho trong phương trình (10).

Trên cơ sở năng lượng đã nêu $E_{L,t,b}$ và $E_{R,t,b}$ cùng với sản phẩm bên trong $X_{L/R,t,b}$, thừa số khuếch đại bên $g_{t,b}$ có thể được tính toán sử dụng phương trình (11) như

$$g_{t,b} = \frac{E_{L,t,b} - E_{R,t,b}}{E_{L,t,b} + E_{R,t,b} + 2X_{L/R,t,b}} \quad (11).$$

Hơn nữa, thừa số khuếch đại phần dư $r_{t,b}$ có thể được tính toán trên cơ sở các năng lượng $E_{L,t,b}$ và $E_{R,t,b}$ cùng với sản phẩm bên trong $X_{L/R,t,b}$ và thừa số khuếch đại bên $g_{t,b}$ sử dụng phương trình (12) như

$$r_{t,b} = \left(\frac{(1-g_{t,b})E_{L,t,b} + (1+g_{t,b})E_{R,t,b} - 2X_{L/R,t,b}}{E_{L,t,b} + E_{R,t,b} + 2X_{L/R,t,b}} \right)^{1/2} \quad (12).$$

Trong các phương án khác, các phương pháp khác và/hoặc các phương trình khác có thể được sử dụng để tính toán thừa số khuếch đại bên $g_{t,b}$ và thừa số khuếch đại phần dư $r_{t,b}$ và/hoặc các tham số so sánh khác nếu phù hợp.

Như nêu trên, sự bù *ITD* trong miền tần số thường tiết kiệm độ phức tạp nhưng - không có các ước số khác - sẽ đi kèm với hạn chế. Một cách lý tưởng, đối với tiếng nói không có tiếng vọng được ghi bằng thiết lập micro AB, tín hiệu kênh bên trái $l(\tau)$ về cơ bản là phiên bản được làm trễ (bằng d trễ) và định tỉ lệ (bằng độ khuếch đại c) của kênh bên phải $r(\tau)$. Tình huống này có thể được biểu diễn bởi phương trình (13) sau, trong đó

$$l(\tau) = c r(\tau - d) \quad (13).$$

Sau sự bù *ITD* phù hợp của các tín hiệu âm thanh kênh đầu vào không được tạo cửa sổ $l(\tau)$ và $r(\tau)$, sự ước lượng cho thừa số khuếch đại bên $g_{t,b}$ sẽ được cho trong phương trình (14) như

$$g_{t,b} = \frac{c-1}{c+1} \quad (14)$$

với thừa số khuếch đại phần dư biến mất $r_{t,b}$ được cho như

$$r_{t,b} = 0 \quad (15).$$

Tuy nhiên, nếu sự căn chỉnh kênh được thực hiện trong miền tần số như trong phương án trên Fig.2 bởi khối phát hiện và bù *ITD* 20 sử dụng lần lượt các khối dịch chuyển hình tròn 13 và 23, các cửa sổ phân tích DFT $w(\tau)$ tương ứng cũng được quay. Do đó, sau khi bù các *ITD* trong miền tần số, biến đổi tần số được bù *ITD* cho kênh phải $R_{t,k,comp}$ có thể được xác định ở dạng các ngăn thời gian - tần số bởi DFT của

$$w(\tau)r(\tau) \quad (16),$$

trong đó biến đổi tần số được bù *ITD* $L_{t,k,comp}$ cho kênh bên trái có thể được xác định ở dạng các ngăn thời gian - tần số như DFT của

$$w(\tau + ITD_t)r(\tau) \quad (17),$$

trong đó w là hàm cửa sổ phân tích DFT.

Các tác giả sáng chế quan sát được rằng sự căn chỉnh kênh trong miền tần số chủ yếu ảnh hưởng thừa số khuếch đại dự báo phần dư $r_{t,b}$, mà trở nên lớn hơn khi tăng

ITD_t . Không có bất kỳ ước số nào khác, sự căn chỉnh kênh trong miền tần số do đó sẽ thêm môi trường xung quanh bổ sung vào tín hiệu âm thanh đầu ra tại bộ giải mã như được thể hiện trên Fig.3. Môi trường xung quanh bổ sung này là không mong muốn, cụ thể khi tín hiệu âm thanh sẽ được mã hóa chứa tiếng nói sạch, vì môi trường xung quanh nhân tạo làm suy yếu tính dễ hiểu của tiếng nói.

Tiếp theo, hiệu quả nêu trên có thể được giảm bằng cách hiệu chỉnh thừa số khuếch đại phần dư (dự báo) $r_{t,b}$ trong sự có mặt của $ITDs$ khác 0 sử dụng tham số so sánh khác.

Trong phương án, điều này có thể được thực hiện bằng cách tính toán độ lệch độ khuếch đại cho độ khuếch đại phần dư $r_{t,b}$, mà nhằm mục đích kết hợp tín hiệu phần dư mong muốn $e(\tau)$ khi tín hiệu nhất quán và phẳng theo thời gian. Trong trường hợp này, các tác giả sáng chế mong muốn độ khuếch đại dự báo toàn cục $\hat{g}$ được cho bởi phương trình (18) như

$$\hat{g} = \frac{c+1}{c-1} \quad (18)$$

và IPD toàn cục biến mất được cho bởi $IPD = 0$. Tiếp theo, tín hiệu phần dư mong muốn $e(\tau)$ có thể được xác định sử dụng phương trình (19) như

$$e(\tau) = \frac{2c}{1+c} (w(\tau) - w(\tau + ITD_t)) r(\tau) \quad (19).$$

Trong phương án, tham số so sánh khác thừa số khuếch đại bên $g_{t,b}$ và thừa số khuếch đại phần dư $r_{t,b}$ có thể được tính toán dựa trên tín hiệu phần dư mong muốn $e(\tau)$ trong khối so sánh và tính toán tham số không gian 3D sử dụng tham số ITD ITD_t và hàm bằng hoặc làm gần đúng hàm tự tương quan $W_X(n)$ của hàm cửa sổ phân tích w được cho trong phương trình (20) như

$$W_X(n) = \sum_{\tau} w(\tau) w(\tau + n) \quad (20).$$

Nếu M_r biểu thị giá trị trung bình thời hạn ngắn của $r^2(\tau)$ năng lượng của tín hiệu phần dư mong muốn $e(\tau)$ có thể được tính toán gần đúng bởi phương trình (21) như

$$\frac{8c^2}{(1+c)^2} (W_X(0) - W_X(ITD_t)) M_r \quad (21).$$

Với tín hiệu giữa được tạo cửa sổ được cho bởi phương trình (22) như

$$m_t(\tau) = (w_t(\tau) + c w_t(\tau + ITD_t))r(\tau) \quad (22),$$

năng lượng của tín hiệu giữa được tạo cửa sổ $m_t(\tau)$ có thể được làm gần đúng bằng phương trình (23) như

$$[(1 + c^2)W_X(0) + 2 c W_X(ITD_t)]M_r \quad (23).$$

Trong phương án, hàm nêu trên được sử dụng trong sự tính toán của tham số so sánh trong khối so sánh và tính toán tham số không gian 30 bằng hoặc làm gần đúng phiên bản được chuẩn hóa $\widehat{W}_X(n)$ của hàm tự tương quan $W_X(n)$ của cửa sổ phân tích như được cho trong phương trình (23a) như

$$\widehat{W}_X(n) = W_X(n)/W_X(0) \quad (23a).$$

Dựa trên hàm tự tương quan được chuẩn hóa $\widehat{W}_X(n)$ này, tham số so sánh khác $\hat{r}_t$ đã nêu có thể được tính toán sử dụng phương trình (24) như

$$\hat{r}_t = \frac{2c}{c+1} \sqrt{2 \frac{1 - \widehat{W}_X(ITD_t)}{1 + c^2 + 2c \widehat{W}_X(ITD_t)}} \quad (24)$$

để cung cấp tham số hiệu chỉnh được ước lượng cho độ khuếch đại phần dư $r_{t,b}$. Trong phương án, tham số so sánh $\hat{r}_t$ có thể được sử dụng như ước lượng cho các độ khuếch đại phần dư cục bộ $r_{t,b}$ trong các băng phụ b . Trong phương án khác, sự hiệu chỉnh các độ khuếch đại phần dư $r_{t,b}$ có thể bị ảnh hưởng bằng cách sử dụng tham số so sánh $\hat{r}_t$ như độ lệch. Tức là, các giá trị của độ khuếch đại phần dư $r_{t,b}$ có thể được thay thế bằng độ khuếch đại phần dư được hiệu chỉnh $r_{t,b,corr}$ như đã cho trong phương trình (25) như

$$r_{t,b,corr} \leftarrow \max\{0, r_{t,b} - \hat{r}_t\} \quad (25).$$

Do đó, trong phương án, tham số so sánh khác được tính trong khối so sánh và tính toán tham số không gian 30 có thể bao gồm độ khuếch đại phần dư được hiệu chỉnh $r_{t,b,corr}$ mà tương ứng với độ khuếch đại phần dư $r_{t,b}$ được hiệu chỉnh bởi tham số hiệu chỉnh độ khuếch đại phần dư $\hat{r}_t$ như được cho tổng phương trình (24) ở dạng độ lệch được định nghĩa trong phương trình (25).

Do đó, phương án khác đề cập đến sự mã hóa âm thanh theo tham số sử dụng DFT được tạo cửa sổ và [tập con gồm] các tham số IPD theo phương trình (3), độ khuếch đại

bên $g_{t,b}$ theo phương trình (11), độ khuếch đại phần dư $r_{t,b}$ theo phương trình (12) và $ITDs$, trong đó độ khuếch đại phần dư $r_{t,b}$ được điều chỉnh theo phương trình (25).

Trong sự đánh giá thực nghiệm, ước lượng độ khuếch đại phần dư $\hat{r}_t$ có thể được thử nghiệm với các lựa chọn khác nhau cho tín hiệu âm thanh kênh bên phải $r(\tau)$ trong phương trình (13). Đối với các tín hiệu đầu vào nhiễu âm trắng $r(\tau)$, mà thỏa mãn giả thiết độ phẳng theo thời gian, các ước lượng độ khuếch đại phần dư $\hat{r}_t$ hơi gần với trung bình của độ khuếch đại phần dư $r_{t,b}$ được đo trong các bảng phụ như có thể thấy từ bảng 1 dưới đây.

$ITD \setminus c$	1	2	4	8	16	32
mili giây	0,0893	0,0793	0,0569	0,0351	0,0196	0,0104
	(0,0885)	(0,0785)	(0,0565)	(0,0349)	(0,0195)	(0,0104)
mili giây	0,1650	0,1460	0,1045	0,0640	0,0357	0,0189
	(0,1631)	(0,1458)	(0,1039)	(0,0640)	(0,0357)	(0,0189)
mili giây	0,2348	0,2073	0,1472	0,0896	0,0498	0,0263
	(0,2327)	(0,2062)	(0,1473)	(0,0904)	(0,0504)	(0,0267)
mili giây	0,3005	0,2644	0,1862	0,1125	0,0621	0,0327
	(0,2992)	(0,2627)	(0,1885)	(0,1151)	(0,0641)	(0,0339)

Bảng 1 Trung bình của các độ khuếch đại phần dư được đo $r_{t,b}$ cho nhiễu âm trắng được quét với ITD và độ khuếch đại phần dư ước lượng $\hat{r}_t$ (được biểu thị trong ngoặc đơn).

Đối với các tín hiệu tiếng nói $r(\tau)$, giả định độ phẳng theo thời gian thường xuyên bị vi phạm, mà thường làm tăng độ trung bình của các độ khuếch đại phần dư $r_{t,b}$ (xem bảng 2 dưới đây so với bảng 1 trên đây). Phương pháp điều chỉnh độ khuếch đại phần dư hoặc hiệu chỉnh theo phương trình (25) do đó có thể được xem như bảo toàn. Tuy nhiên, có thể vẫn loại bỏ môi trường xung quanh không mong muốn cho sự ghi tiếng nói rõ ràng hơn.

$ITD \setminus c$	1	2	4
mili giây	0,1055	0,1022	0,0874
	(0,0885)	(0,0785)	(0,0565)
mili giây	0,1782	0,1634	0,1283
	(0,1631)	(0,1458)	(0,1039)
mili giây	0,2435	0,2191	0,1657
	(0,2327)	(0,2062)	(0,1473)
mili giây	0,3050	0,2720	0,2014
	(0,2992)	(0,2627)	(0,1885)

Bảng 2 Trung bình của các độ khuếch đại phần dư được đo $r_{t,b}$ cho tiếng nói đơn được quét với ITD và độ khuếch đại phần dư ước lượng $\hat{r}_t$ (được biểu thị trong ngoặc đơn).

Hàm tự tương quan được chuẩn hóa $\hat{W}_X$ được cho trong phương trình (23a) có thể được xem xét để độc lập của chỉ số khung t trong trường hợp của cửa sổ phân tích đơn w được sử dụng. Hơn nữa, hàm tự tương quan được chuẩn hóa $\hat{W}_X$ có thể được xem là thay đổi rất chậm đối với hàm cửa sổ phân tích điển hình w . Do đó, $\hat{W}_X$ có thể được nội suy một cách chính xác từ bảng các giá trị nhỏ, mà khiến sơ đồ hiệu chỉnh này hiệu quả về độ phức.

Do đó, trong các phương án này, hàm để xác định các ước lượng độ khuếch đại phần dư hoặc độ lệch hiệu chỉnh độ khuếch đại phần dư $\hat{r}_t$ như tham số so sánh trong khối 30 có thể thu được bằng cách nội suy phiên bản được chuẩn hóa $\hat{W}_X$ của hàm tự tương quan của cửa sổ phân tích được lưu trữ trong bảng tìm kiếm. Trong phương án khác, các phương pháp khác để nội suy hàm tự tương quan được chuẩn hóa $\hat{W}_X$ có thể được sử dụng một cách phù hợp.

Đối với BCC, như được mô tả trong tài liệu [2], vấn đề tương tự phát sinh khi tính toán sự nhất quán liên kênh ICC trong các băng phụ. Trong phương án, $ICC_{t,b}$ tương ứng có thể được ước lượng bởi phương trình (26) sử dụng các năng lượng $E_{L,t,b}$ và $E_{R,t,b}$ của phương trình (9) và sản phẩm bên trong của phương trình (10) như

$$ICC_{t,b} = \frac{X_{L/R,t,b}}{\sqrt{E_{L,t,b} \cdot E_{R,t,b}}} (26).$$

Bằng sự định nghĩa, ICC được đo sau khi bù các ITD . Tuy nhiên, các hàm cửa sổ không phù hợp w có thể nghiêng về sự đo ICC . Trong thiết lập tiếng nói không có tiếng vọng trong được đề cập trên đây bởi phương trình (13), ICC sẽ là 1 nếu được tính toán trên các kênh đầu vào được căn chỉnh một cách phù hợp.

Tuy nhiên, độ lệch - bị gây ra bởi sự quay của các hàm cửa sổ phân tích $w(\tau)$ trong miền tần số khi bù ITD của ITD_t trong miền tần số bởi (các) dịch chuyển hình tròn - có thể nghiêng về sự đo của ICC về phía $\hat{ICC}_t$ như đã cho trong phương trình (27) như

$$\hat{ICC}_t = \hat{W}_X(ITD_t) \quad (27).$$

Trong phương án, sự nghiêng của ICC có thể được hiệu chỉnh trong cách tương tự so với sự hiệu chỉnh của độ khuếch đại phần dư $r_{t,b}$ trong phương trình (25), cụ thể bằng cách thay thế như đã cho trong phương trình (28) như

$$ICC_{b,t} \leftarrow 1 + \min\{ICC_{b,t} - \hat{ICC}_t, 0\} \quad (28).$$

Do đó, phương án khác liên quan đến sự mã hóa âm thanh theo tham số sử dụng DFT được tạo của sổ và [tập con gồm] các tham số IPD theo phương trình (3), ILD , ICC theo phương trình (26) và $ITDs$, trong đó ICC được điều chỉnh theo phương trình (28).

Trong phương án của bộ mã hóa theo tham số 200 được thể hiện trên Fig.2, khối trộn giảm 40 có thể giảm số lượng các kênh của hệ thống đa kênh, ở đây là âm lập thể bằng cách tính toán tín hiệu trộn giảm $DMX_{t,k}$ được cho bởi phương trình (29) trong miền tần số. Trong phương án, tín hiệu trộn giảm $DMX_{t,k}$ có thể được tính toán sử dụng các biến đổi tần số được bù ITD $L_{t,k,comp}$ và $R_{t,k,comp}$ theo

$$DMX_{t,k} = \frac{e^{-i\beta L_{t,k,comp}} + e^{i(IPD_{t,b} - \beta) R_{t,k,comp}}}{\sqrt{2}} (29).$$

Trong phương trình (29), β có thể là tham số tính toán pha tuyệt đối thực được tính toán từ các tham số âm lập thể/không gian. Trong các phương án khác, sơ đồ mã hóa như được thể hiện trên Fig.2 có thể cũng làm việc với bất kỳ phương pháp trộn giảm khác. Các phương án khác có thể sử dụng các biến đổi tần số $L_{t,k}$ và $R_{t,k}$ và các tham số khác để xác định tín hiệu trộn giảm $DMX_{t,k}$.

Trong phương án bộ mã hóa của Fig.2, khối biến đổi Fourier rời rạc nghịch đảo (inverse discrete Fourier transform - IDFT) 50 có thể thu tín hiệu trộn giảm miền tần số $DMX_{t,k}$ từ khối trộn giảm 40. Khối IDFT 50 có thể biến đổi các ngăn thời gian-tần số trộn giảm $DMX_{t,k}$, $k = 0, \dots, K - 1$, từ miền tần số thành miền thời gian để tạo ra tín hiệu trộn giảm miền thời gian $dmx(\tau)$. Trong các phương án, cửa sổ tổng hợp $w_s(\tau)$ có thể được áp dụng và bổ sung vào tín hiệu trộn giảm miền thời gian $dmx(\tau)$.

Hơn nữa, như trong phương án trên Fig.2, bộ mã hóa lõi 60 có thể thu tín hiệu trộn giảm miền $dmx(\tau)$ để mã hóa tín hiệu âm thanh kênh đơn theo MPEG-4 Phần 3 [1] hoặc bất kỳ thuật toán mã hóa âm thanh phù hợp nếu thích hợp. Trong phương án của Fig.2, tín hiệu trộn giảm miền thời gian được mã hóa lõi $dmx(\tau)$ có thể được tổ hợp với tham số ITD ITD_t , độ khuếch đại bên $g_{t,b}$ và độ khuếch đại phần dư hiệu chỉnh $r_{t,b,corr}$ được xử lý một cách phù hợp và/hoặc được mã hóa để truyền dẫn đến bộ giải mã.

Fig.3 thể hiện phương án của bộ giải mã đa kênh. Bộ giải mã có thể thu tín hiệu được tổ hợp bao gồm tín hiệu đầu vào đơn/trộn giảm $dmx(\tau)$ trong miền thời gian và sự so sánh và/hoặc các tham số không gian như thông tin phụ trên cơ sở khung. Bộ giải mã như được thể hiện trên Fig.3 có thể thực hiện các bước sau đây, mà được mô tả cụ thể sau đây.

1. Sự biến đổi thời gian thành tần số đầu vào sử dụng các DFT được tạo cửa sổ trong khối DFT 80
2. Sự dự báo phần dư thiếu sót trong miền tần số trong khối trộn tăng và phục hồi không gian 90
3. Sự trộn tăng trong miền tần số trong khối trộn tăng và phục hồi không gian 90
4. Phân tích ITD trong miền tần số trong khối tổng hợp ITD 100
5. Sự biến đổi miền tần số thành thời gian, sự tạo cửa sổ và chồng lấp cộng trong khối IDFT 112, 122 và các khối cửa sổ 111, 121

Sự biến đổi thời gian thành tần số của tín hiệu đầu vào tín hiệu đơn/trộn giảm $dmx(\tau)$ có thể được hoàn thành theo cách tương tự như đối với các tín hiệu âm thanh đầu vào của bộ mã hóa trên Fig.2. Trong các phương án nhất định, lượng phù hợp các vùng đệm số 0 được thêm vào cho sự khôi phục ITD trong miền tần số. Quy trình này có thể tạo ra biến đổi tần số của tín hiệu trộn giảm ở dạng các ngăn thời gian - tần số $DMX_{t,k}$, $k = 0, \dots, K - 1$.

Để khôi phục các đặc tính không gian của tín hiệu trộn giảm $DMX_{t,k}$, tín hiệu thứ hai, phụ thuộc vào tín hiệu trộn giảm được truyền dẫn $DMX_{t,k}$ có thể được cần. Sao cho tín hiệu có thể, ví dụ, được khôi phục (lại) trong khối trộn tăng và khôi phục không gian 90 sử dụng độ khuếch đại phần dư được hiệu chỉnh $r_{t,b,corr}$ như tham số so sánh - được truyền dẫn bởi bộ mã hóa như bộ mã hóa trên Fig.2 – và các ngăn thời gian - tần số được làm trễ thời gian của tín hiệu trộn giảm $DMX_{t,k}$ như được cho trong phương trình (30):

$$\hat{\rho}_{t,k} = r_{t,b,corr} \sqrt{\frac{\sum_{k \in I_b} |DMX_{t,k}|^2}{\sum_{k \in I_b} |DMX_{t-d_b,k}|^2}} DMX_{t-d_b,k} \quad (30)$$

đối với $k \in I_b$.

Trong các phương án khác, các phương pháp và phương trình khác nhau có thể được sử dụng để khôi phục các đặc tính không gian của tín hiệu trộn giảm $DMX_{t,k}$ dựa trên ít nhất một tham số so sánh được truyền dẫn.

Hơn nữa, khối khôi phục trộn tăng và không gian 90 có thể thực hiện sự trộn tăng bằng cách áp dụng sự nghịch đảo lên sự biến đổi giữa/bên tại bộ mã hóa sử dụng tín hiệu trộn giảm $DMX_{t,k}$ và độ khuếch đại bên $g_{t,b}$ như được truyền dẫn bởi bộ mã hóa cũng như tín hiệu phần dư được khôi phục lại $\hat{\rho}_{t,k}$. Điều này có thể tạo ra các biến đổi tần số được bù ITD được giải mã $\hat{L}_{t,k}$ và $\hat{R}_{t,k}$ được cho bởi các phương trình (31) và (32) như

$$\hat{L}_{t,k} = \frac{e^{i\beta(DMX_{t,k}(1+g_{t,b})+\hat{\rho}_{t,k})}}{\sqrt{2}} \quad (31)$$

và

$$\hat{R}_{t,k} = \frac{e^{i(\beta-IPD_b)(DMX_{t,k}(1-g_{t,b})-\hat{\rho}_{t,k})}}{\sqrt{2}} \quad (32)$$

đối với $k \in I_b$, trong đó β là tham số quay pha tuyệt đối như trong quy trình trộn giảm trong phương trình (29).

Hơn nữa, như được thể hiện trên Fig.3, các biến đổi tần số được bù ITD được giải mã $\hat{L}_{t,k}$ và $\hat{R}_{t,k}$ có thể được thu bởi khối tổng hợp/khử bù ITD 100. Khối sau có thể áp dụng tham số ITD ITD_t trong miền tần số bằng cách quay $\hat{L}_{t,k}$ và $\hat{R}_{t,k}$ như được cho trong phương trình (33) và (34) để tạo ra các biến đổi tần số được giải mã được khử bù ITD $\hat{L}_{t,k,decomp}$ và $\hat{R}_{t,k,decomp}$:

$$\hat{L}_{t,k,decomp} \leftarrow e^{i\frac{\pi}{K}ITD_t k} \hat{L}_{t,k} \quad (33)$$

và

$$\hat{R}_{t,k,decomp} \leftarrow e^{-i\frac{\pi}{K}ITD_t k} \hat{R}_{t,k}, \quad (34).$$

Trên Fig.3, sự biến đổi miền tần số đến thời gian của các biến đổi tần số được giải mã được khử bù ITD ở dạng các ngăn thời gian - tần số $\hat{L}_{t,k,decomp}$ và $\hat{R}_{t,k,decomp}$, $k = 0, \dots, K-1$ có thể được thực hiện lần lượt bởi các khối IDFT 112 và 122. Các tín hiệu miền thời gian thu được sau đó có thể được tạo cửa sổ lần lượt bằng các khối cửa sổ 111 và 121 và được thêm vào các tín hiệu âm thanh đầu ra miền thời gian được khôi phục $\hat{l}(\tau)$ và $\hat{r}(\tau)$ của kênh âm thanh bên trái và bên phải.

Các phương án được mô tả bên trên chỉ mang tính minh họa cho các nguyên lý của sáng chế. Cần hiểu rằng các cải biên và biến đổi của các phương án và các chi tiết được mô tả ở đây sẽ là rõ ràng đối với người có trình độ trong lĩnh vực kỹ thuật tương ứng. Do đó, mục đích của sáng chế là chỉ được giới hạn bởi phạm vi của các điểm yêu cầu bảo hộ sắp đưa ra dưới đây và không bị giới hạn bởi các chi tiết cụ thể được biểu diễn bằng bản mô tả và giải thích của các phương án ở đây.

YÊU CẦU BẢO HỘ

1. Thiết bị so sánh cho tín hiệu âm thanh đa kênh được tạo cấu hình để:

suy ra, cho sự chênh lệch thời gian liên kênh (inter-channel time difference - ITD) giữa các tín hiệu âm thanh cho ít nhất một cặp kênh, ít nhất một tham số ITD (ITD_t) của các tín hiệu âm thanh của ít nhất một cặp kênh trong cửa sổ phân tích ($w(\tau)$),

bù ITD cho ít nhất một cặp kênh trong miền tần số bằng sự dịch chuyển vòng tròn sử dụng ít nhất một tham số ITD để tạo ra ít nhất một cặp biến đổi tần số được bù ITD ($L_{t,k,comp}; R_{t,k,comp}$),

tính toán, trên cơ sở ít nhất một tham số ITD và ít nhất một cặp biến đổi tần số được bù ITD , ít nhất một tham số so sánh ($\hat{r}_t, IC_t$).

2. Thiết bị so sánh theo điểm 1, thiết bị này còn được tạo cấu hình để sử dụng các biến đổi tần số ($L_{t,k}; R_{t,k}$) của các tín hiệu âm thanh của ít nhất một cặp kênh trong cửa sổ biến đổi ($w(\tau)$) để suy ra ít nhất một tham số ITD (ITD_t).

3. Thiết bị so sánh theo điểm 1 hoặc điểm 2, thiết bị này còn được tạo cấu hình để:

tính toán ít nhất một tham số so sánh sử dụng hàm bằng hoặc làm gần đúng hàm tự tương quan ($W_X(n) = \sum_{\tau} w(\tau) w(\tau + n)$) của cửa sổ phân tích và ít nhất một tham số ITD .

4. Thiết bị so sánh theo điểm 3, trong đó

hàm bằng hoặc làm gần đúng phiên bản được chuẩn hóa của hàm tự tương quan ($\hat{W}_X(n) = W_X(n) / W_X(0)$) của cửa sổ phân tích.

5. Thiết bị so sánh theo điểm 4, thiết bị này còn được tạo cấu hình để:

thu được hàm bằng phép nội suy phiên bản được chuẩn hóa của hàm tự tương quan của cửa sổ phân tích được chứa trong bảng tìm kiếm.

6. Thiết bị so sánh theo điểm bất kỳ trong số các điểm từ 1 đến 5, trong đó:

ít nhất một tham số so sánh bao gồm ít nhất một độ khuếch đại bên ($g_{t,b}$) của ít nhất một cặp biến đổi giữa/bên ($M_{t,k}; S_{t,k}$) của ít nhất một cặp biến đổi tần số được bù $ITD (L_{t,k,comp}; R_{t,k,comp})$, ít nhất một độ khuếch đại bên là độ khuếch đại dự báo ($S_{t,k} = g_{t,b}M_{t,k} + \rho_{t,k}$) của biến đổi bên ($S_{t,k}$) từ biến đổi giữa ($M_{t,k}$) của ít nhất một cặp biến đổi giữa/bên.

7. Thiết bị so sánh theo điểm 6, trong đó

ít nhất một tham số so sánh bao gồm ít nhất một độ khuếch đại phần dư được hiệu chỉnh ($r_{t,b,corr}$) tương ứng với ít nhất một độ khuếch đại phần dư ($r_{t,b}$) được hiệu chỉnh bởi tham số hiệu chỉnh độ khuếch đại phần dư ($\hat{r}_t$), ít nhất một độ khuếch đại phần dư ($r_{t,b}$) là hàm của năng lượng phần dư ($\rho_{t,k}$) trong sự dự báo biến đổi bên ($S_{t,k}$) từ biến đổi giữa ($M_{t,k}$) so với năng lượng của biến đổi giữa

$$\left(r_{t,b} = \left(\frac{\sum_{k \in I_b} |\rho_{t,k}|^2}{\sum_{k \in I_b} |M_{t,k}|^2} \right)^{1/2} \right).$$

8. Thiết bị so sánh theo điểm 7, thiết bị này còn được tạo cấu hình để:

tính toán ít nhất một độ khuếch đại bên và ít nhất một độ khuếch đại phần dư sử dụng năng lượng và sản phẩm bên trong của ít nhất một cặp biến đổi tần số được bù $ITD (L_{t,k,comp}; R_{t,k,comp})$.

9. Thiết bị so sánh theo điểm bất kỳ trong số các điểm từ 7 đến 8, thiết bị này còn được tạo cấu hình để:

hiệu chỉnh ít nhất một độ khuếch đại phần dư bởi độ lệch tương ứng với tham số hiệu chỉnh độ khuếch đại phần dư $\hat{r}_t$ được tính toán như $\hat{r}_t = \frac{2c}{c+1} \sqrt{2 \frac{1 - \widehat{W}_X(ITD_t)}{1 + c^2 + 2c\widehat{W}_X(ITD_t)}}$, trong đó c là độ khuếch đại định tỉ lệ giữa các tín hiệu âm thanh của ít nhất một cặp kênh và $\widehat{W}_X(n)$ là hàm làm gần đúng phiên bản được chuẩn hóa của hàm tự tương quan của cửa sổ phân tích.

10. Thiết bị so sánh theo điểm bất kỳ trong số các điểm từ 1 đến 9, trong đó:

ít nhất một tham số so sánh bao gồm ít nhất một tham số hiệu chỉnh (ICC_t) đồng nhất liên kênh (inter-channel coherence - ICC) để hiệu chỉnh sự ước lượng ($ICC_{b,t}$) của ICC - được xác định trong miền tần số - của ít nhất một cặp tín hiệu kênh âm thanh trên cơ sở ít nhất một tham số ITD .

11. Thiết bị so sánh theo điểm bất kỳ trong số các điểm từ 1 đến 10, trong đó:

tạo ra ít nhất một tín hiệu trộn giảm cho các tín hiệu âm thanh của ít nhất một cặp kênh, trong đó ít nhất một tham số so sánh ($\hat{r}_t, ICC_t$) được tính toán để lưu trữ các tín hiệu âm thanh của ít nhất một cặp kênh từ ít nhất một tín hiệu trộn giảm.

12. Thiết bị so sánh theo điểm bất kỳ trong số các điểm từ 1 đến 11, thiết bị còn được tạo cấu hình để:

tạo ra ít nhất một tín hiệu trộn giảm dựa trên ít nhất một cặp biến đổi tần số được bù ITD .

13. Bộ mã hóa đa kênh bao gồm thiết bị so sánh theo điểm 11 hoặc điểm 12, bộ mã hóa đa kênh này còn được tạo cấu hình để

mã hóa ít nhất một tín hiệu trộn giảm, ít nhất một tham số ITD và ít nhất một tham số so sánh để truyền dẫn đến bộ giải mã.

14. Bộ giải mã cho các tín hiệu âm thanh đa kênh được tạo cấu hình để

giải mã ít nhất một tín hiệu trộn giảm, ít nhất một tham số chênh lệch thời gian liên kênh (inter-channel time difference - ITD) và ít nhất một tham số so sánh ($\hat{r}_t, ICC_t$) thu được từ bộ mã hóa,

trộn tăng ít nhất một tín hiệu trộn giảm để khôi phục các tín hiệu âm thanh của ít nhất một cặp kênh từ ít nhất một tín hiệu trộn giảm sử dụng ít nhất một tham số so sánh để tạo ra ít nhất một cặp biến đổi tần số được bù ITD được giải mã ($\hat{L}_{t,k}; \hat{R}_{t,k}$),

khử bù ITD cho ít nhất một cặp biến đổi tần số được bù ITD được giải mã ($\hat{L}_{t,k}; \hat{R}_{t,k}$) của ít nhất một cặp kênh trong miền tần số bởi sự dịch chuyển vòng tròn sử dụng ít nhất một tham số ITD để tạo ra ít nhất một cặp biến đổi tần số được giải

mã được khử bù ITD để khôi phục ITD của các tín hiệu âm thanh của ít nhất một cặp trong miền thời gian,

biến đổi tần số nghịch đảo ít nhất một cặp biến đổi tần số được giải mã được khử bù ITD để tạo ra ít nhất một cặp tín hiệu âm thanh được giải mã của ít nhất một cặp kênh.

15. Phương pháp so sánh cho tín hiệu âm thanh đa kênh bao gồm các bước:

suy ra, cho sự chênh lệch thời gian liên kênh (ITD) giữa các tín hiệu âm thanh cho ít nhất một cặp kênh, ít nhất một tham số ITD (ITD_t) của các tín hiệu âm thanh của ít nhất một cặp kênh trong cửa sổ phân tích ($w(\tau)$),

bù ITD cho ít nhất một cặp kênh trong miền tần số bằng sự dịch chuyển vòng tròn sử dụng ít nhất một tham số ITD để tạo ra ít nhất một cặp biến đổi tần số được bù ITD ($L_{t,k,comp}$; $R_{t,k,comp}$),

tính toán, trên cơ sở ít nhất một tham số ITD và ít nhất một cặp biến đổi tần số được bù ITD , ít nhất một tham số so sánh ($\hat{r}_t$, ICC_t).

Thiết bị so sánh cho bộ so sánh tham số

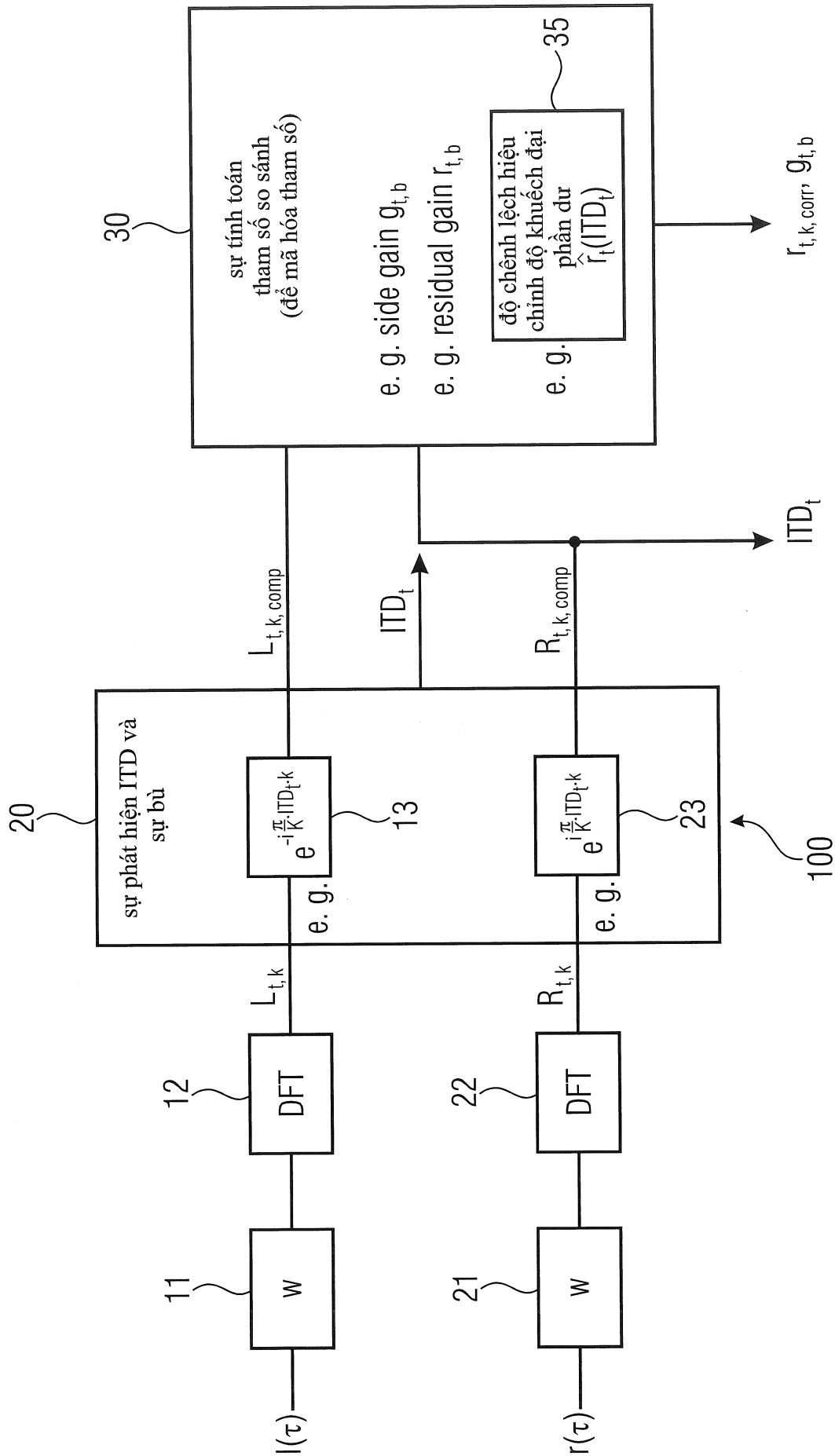


Fig. 1

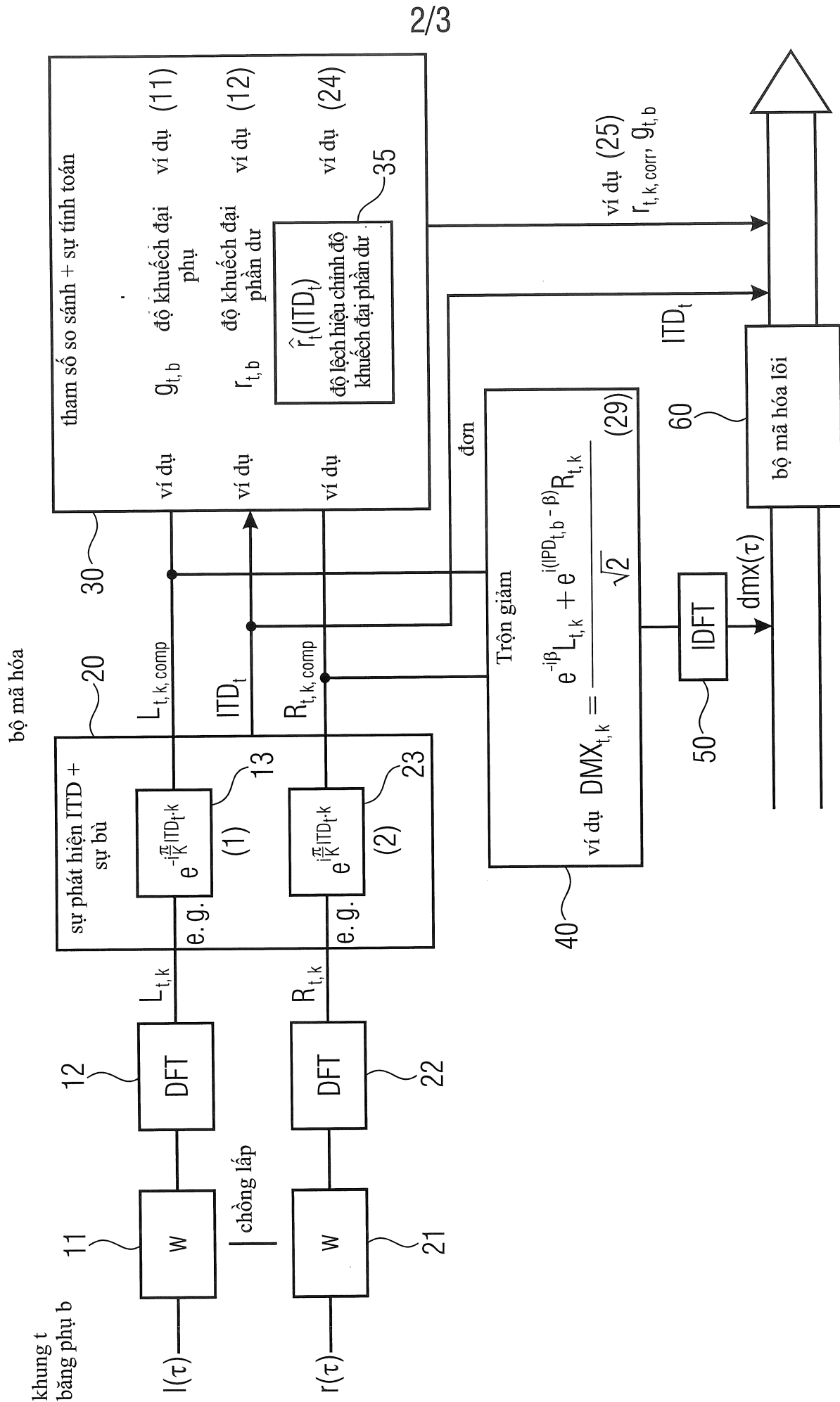


Fig. 2

