



(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)



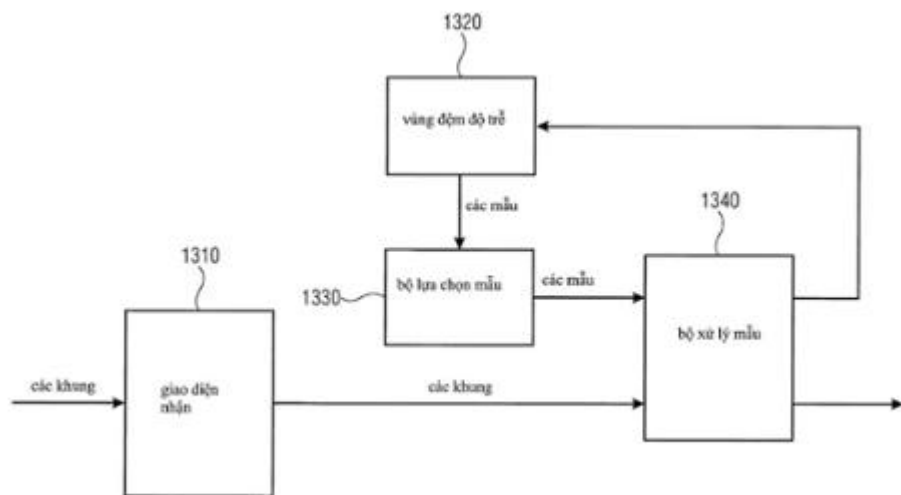
1-0026214

(51)<sup>7</sup> G10L 19/09; G10L 25/90; G10L 19/005 (13) B

- (21) 1-2016-00232 (22) 23/06/2014  
(86) PCT/EP2014/063176 23/06/2014 (87) WO 2014/202789 A1 24/12/2014  
(30) 13 173 154.9 21/06/2013 EP; 14 166 998.6 05/05/2014 EP  
(45) 25/11/2020 392 (43) 25/04/2016 337A  
(73) Fraunhofer-Gesellschaft zur Foerderung der angewandten Forschung e.V. (DE)  
Hansastrasse 27c, 80686 Muenchen, Germany  
(72) SCHNABEL, Michael (DE); MARKOVIC, Goran (RS); SPERSCHNEIDER, Ralph  
(DE); LECOMTE, Jérémie (FR); HELMRICH, Christian (DE).  
(74) CÔNG TY LUẬT TRÁCH NHIỆM HỮU HẠN AMBYS HÀ NỘI (AMBYS  
HANOI)

(54) THIẾT BỊ VÀ PHƯƠNG PHÁP GIẢI MÃ TÍN HIỆU ÂM THANH ĐƯỢC MÃ HÓA ĐỂ THU ĐƯỢC TÍN HIỆU ÂM THANH ĐƯỢC KHÔI PHỤC

(57) Sáng chế đề cập đến thiết bị và phương pháp giải mã tín hiệu âm thanh được mã hóa để thu được tín hiệu âm thanh được khôi phục. Thiết bị bao gồm giao diện nhận (1310) để nhận nhiều khung, vùng đệm trễ (1020; 1320) để lưu trữ các mẫu tín hiệu âm thanh của tín hiệu âm thanh được giải mã, bộ lựa chọn mẫu (1030; 1330) để lựa chọn nhiều mẫu tín hiệu âm thanh được lựa chọn từ các mẫu tín hiệu âm thanh được lưu trữ trong vùng đệm trễ (1020; 1320), và bộ xử lý mẫu (1040; 1340) để xử lý các mẫu tín hiệu âm thanh được lựa chọn để thu được các mẫu tín hiệu âm thanh được khôi phục của tín hiệu âm thanh được khôi phục. Bộ lựa chọn mẫu (1030; 1330) được tạo cấu hình để lựa chọn, nếu khung hiện thời được nhận bởi giao diện nhận (1310) và nếu khung hiện thời được nhận bởi giao diện nhận (1310) không bị sai lệch, nhiều mẫu tín hiệu âm thanh được lựa chọn từ các mẫu tín hiệu âm thanh được lưu trữ trong vùng đệm trễ (1020; 1320) phụ thuộc vào thông tin độ trễ bước được bao gồm bởi khung hiện thời. Ngoài ra, bộ lựa chọn mẫu (1030; 1330) được tạo cấu hình để lựa chọn, nếu khung hiện thời không được nhận bởi giao diện nhận (1310) hoặc nếu khung hiện thời được nhận bởi giao diện nhận (1310) bị sai lệch, nhiều mẫu tín hiệu âm thanh được lựa chọn từ các mẫu tín hiệu âm thanh được lưu trữ trong vùng đệm trễ (1020; 1320) phụ thuộc vào thông tin độ trễ bước được bao gồm bởi khung khác được nhận trước đó bởi giao diện nhận (1310).



### **Lĩnh vực kỹ thuật được đề cập**

Sáng chế đề cập đến việc mã hóa, xử lý và giải mã tín hiệu âm thanh, và cụ thể, đề cập đến thiết bị và phương pháp cho sự giảm dần cường độ tín hiệu đã được cải thiện cho các hệ thống mã hóa âm thanh được chuyển đổi trong quá trình che giấu lỗi.

### **Tình trạng kỹ thuật của sáng chế**

Sau đây, tình trạng kỹ thuật được mô tả liên quan đến các bộ mã hóa-giải mã âm thanh và tiếng nói trong sự che giấu mất gói (packet loss concealment - PLC). Sự giải thích liên quan đến tình trạng kỹ thuật với các bộ mã hóa-giải mã ITU-T của chuỗi G (G.718, G.719, G.722, G.722.1, G.729, G.729.1), được theo sau bởi các bộ mã hóa-giải mã 3GPP (AMR, AMR-WB, AMR-WB+) và một bộ mã hóa-giải mã IETF (OPUS), và kết thúc với hai bộ mã hóa-giải mã MPEG (HE-AAC, HILN) (Liên minh viễn thông quốc tế (International Telecommunication Union - ITU); Dự án hợp tác thế hệ thứ ba (3rd Generation Partnership Project - 3GPP); Đa tốc độ thích ứng (Adaptive Multi-Rate - AMR); Băng rộng (Wideband - WB); Lực lượng chuyên trách kỹ thuật liên mạng (Internet Engineering Task Force - IETF). Tiếp theo, tình trạng kỹ thuật liên quan đến việc tạo vết mức độ tạp âm nền được phân tích, theo sau là bản tóm tắt cung cấp cái nhìn tổng quan.

Đầu tiên, G.718 được xem xét. G.718 là bộ mã hóa-giải mã tiếng nói băng hẹp và băng rộng, mà hỗ trợ DTX/CNG (Các hệ thống nhà hát kỹ thuật số (Digital Theater Systems - DTX); Thế hệ nhiễu âm nhân tạo (Comfort Noise Generation - CNG)). Vì các phương án đề cập một cách cụ thể đến sự mã hóa độ trễ thấp, chế độ phiên bản độ trễ thấp sẽ được mô tả cụ thể hơn trong tài liệu này.

Xem xét ACELP (Lớp 1) (dự báo tuyến tính xuất phát từ mã đại số (Algebraic Code Excited Linear Prediction - ACELP)), ITU-T giới thiệu cho G.718 [ITU08a, phần 7.11] sự giảm dần cường độ thích ứng trong miền dự báo tuyến tính để điều khiển tốc độ giảm dần cường độ. Nhìn chung, sự che giấu đi theo nguyên lý sau:

Theo G.718, trong trường hợp xóa khung, chiến lược che giấu có thể được tổng

kết như sự hội tụ của năng lượng tín hiệu và đường bao phổ đến các tham số được ước lượng của tập âm nền. Chu kỳ của tín hiệu được hội tụ về 0. Tốc độ của sự hội tụ phụ thuộc vào các tham số của khung được nhận chính xác cuối cùng và số lượng các khung đã xóa liên tiếp, và được kiểm soát bởi hệ số suy giảm,  $\alpha$ . Hệ số suy giảm  $\alpha$ , còn phụ thuộc vào độ ổn định,  $\theta$ , của bộ lọc dự báo tuyến tính (Linear Prediction - LP) cho các khung VÔ THANH. Nhìn chung, sự hội tụ sẽ chậm nếu khung được nhận tốt cuối cùng nằm trong phân đoạn ổn định và sẽ nhanh đến khung nằm trong phân đoạn chuyển tiếp.

Hệ số suy giảm  $\alpha$  phụ thuộc vào loại tín hiệu tiếng nói, mà được suy ra bởi sự phân loại tín hiệu được mô tả trong [ITU08a, phần 6.8.1.3.1 và 7.11.1.1]. Hệ số ổn định  $\theta$  được tính toán dựa trên số đo khoảng cách giữa các bộ lọc ISF liên kề (Immittance Spectral Frequency) [ITU08a, phần 7.1.2.4.2].

Bảng 1 thể hiện sơ đồ tính toán của  $\alpha$ :

| Khung được nhận tốt cuối cùng | Số lượng các khung được xóa liên tiếp | $\alpha$                 |
|-------------------------------|---------------------------------------|--------------------------|
| Sự bắt đầu nhân tạo           |                                       | 0,6                      |
| Sự bắt đầu, hữu thanh         | $\leq 3$                              | 1,0                      |
|                               | $> 3$                                 | 0,4                      |
| Sự chuyển tiếp hữu thanh      |                                       | 0,4                      |
| Sự chuyển tiếp vô thanh       |                                       | 0,8                      |
| Vô thanh                      | $= 1$                                 | $0,2 \cdot \theta + 0,8$ |
|                               | $= 2$                                 | 0,6                      |
|                               | $> 2$                                 | 0,4                      |

Bảng 1 Các giá trị của hệ số suy giảm  $\alpha$ , giá trị  $\theta$  là hệ số ổn định được tính toán từ số đo khoảng cách giữa các bộ lọc LP liên kề. [ITU08a, phần 7.1.2.4.2].

Hơn nữa, G.718 đề xuất phương pháp tăng giảm cường độ để thay đổi đường bao phổ. Ý kiến chung là để hội tụ các tham số ISF cuối cùng về phía vector trung bình ISF thích ứng. Đầu tiên, vector ISF trung bình được tính toán từ 3 vector ISF đã biết cuối cùng. Sau đó, vector ISF trung bình lần nữa được tính trung bình với vector ISF dài hạn

được tạo chuỗi ngoại tuyến (mà là véctor bất biến) [ITU08a, phần 7.11.1.2].

Hơn nữa, G.718 đề xuất phương pháp tăng giảm cường độ để kiểm soát động thái dài hạn và do đó tương tác với tạp âm nền, trong đó năng lượng kích thích bước (và theo đó chu kỳ kích thích) được hội tụ về 0, trong khi năng lượng kích thích ngẫu nhiên hội tụ về năng lượng kích thích CNG [ITU08a, phần 7.11.1.6]. Sự suy giảm độ khuếch đại đổi mới được tính toán như

$$g_s^{[1]} = \alpha g_s^{[0]} + (1 - \alpha)g_n \quad (1)$$

trong đó  $g_s^{[1]}$  là độ khuếch đại đổi mới tại lúc bắt đầu của khung tiếp theo,  $g_s^{[0]}$  là độ khuếch đại tại lúc bắt đầu của khung hiện thời,  $g_n$  là độ khuếch đại của sự kích thích được sử dụng trong quá trình tạo ra nhiễu âm nền nhân tạo và hệ số suy giảm  $\alpha$ .

Tương tự với sự suy giảm kích thích chu kỳ, độ khuếch đại được suy giảm một cách tuyến tính qua khung trên cơ sở từng mẫu bắt đầu với,  $g_s^{[0]}$ , và đạt đến  $g_s^{[1]}$  tại thời điểm bắt đầu khung tiếp theo.

Fig.2 phác thảo cấu trúc bộ giải mã của G.718. Cụ thể, Fig.2 minh họa kết cấu bộ giải mã mức cao G.718 cho PLC, kết hợp với bộ lọc thông cao.

Bằng phương pháp được mô tả như trên của G.718, độ khuếch đại đổi mới  $g_s$  hội tụ về độ khuếch đại được sử dụng trong quá trình tạo ra nhiễu âm nền nhân tạo  $g_n$  cho các khối tín hiệu truyền dài của các sự mất gói. Như được mô tả trong [ITU08a, phần 6.12.3], độ khuếch đại nhiễu âm nền nhân tạo  $g_n$  được cho như là căn bậc hai của năng lượng  $\tilde{E}$ . Các điều kiện của sự cập nhật  $\tilde{E}$  không được mô tả chi tiết. Tiếp theo sự thực hiện tham chiếu (mã C dấu chấm động, `stat_noise_uv_mod.c`),  $\tilde{E}$  được đạo hàm như sau:

```
if(unvoiced_vad == 0){
    if( unv_cnt > 20 ){
        ftmp = lp_gainc * lp_gainc;
        lp_ener = 0.7f * lp_ener + 0.3f * ftmp;
```

```

    }

    else{

        unv_cnt++;

    }

}

else{

    unv_cnt = 0;

}

```

trong đó `unvoiced_vad` duy trì việc phát hiện hoạt động tiếng nói, trong đó `unv_cnt` duy trì số lượng khung vô thanh trong hàng, trong đó `lp_gainc` duy trì độ khuếch đại được tạo thông thấp của bảng mã cố định, và trong đó `lp_ener` duy trì sự ước lượng năng lượng CNG được tạo thông thấp  $\tilde{E}$ , nó được thiết lập ban đầu với 0.

Hơn nữa, G.718 cung cấp bộ lọc thông cao, được dẫn vào trong đường dẫn tín hiệu của sự kích thích vô thanh, nếu tín hiệu của khung tốt cuối cùng được phân loại khác với Vô thanh, xem Fig.2, và xem [ITU08a, phần 7.11.1.6]. Bộ lọc này có đặc điểm ngăn thấp với đáp ứng tần số tại DC là khoảng 5dB thấp hơn tần số Nyquist.

Hơn nữa, G.718 đề xuất chu trình phản hồi dự báo dài hạn (Long-Term Prediction - LTP): Trong quá trình hoạt động bình thường, chu trình phản hồi cho bảng mã thích ứng được cập nhật theo khung phụ ([ITU08a, phần 7.1.2.1.4]) dựa trên sự kích thích đầy đủ. Trong quá trình che giấu chu trình đáp ứng này được cập nhật theo khung (xem [ITU08a, các phần 7.11.1.4, 7.11.2.4, 7.11.1.6, 7.11.2.6; `dec_GV_exc@dec_gen_voic.c` và `syn_bfi_post@syn_bfi_pre_post.c`]) chỉ dựa trên sự kích thích hữu thanh. Với phương pháp này, bảng mã thích ứng không bị "ô nhiễm" với nhiễu âm có nguồn gốc từ sự kích thích đổi mới được lựa chọn ngẫu nhiên.

Đối với sự biến đổi đã mã hóa các lớp nâng cao (3-5) của G.718, trong quá trình che giấu, bộ giải mã vận hành về sự giải mã lớp cao tương tự thao tác bình thường, chỉ

là phổ MDCT được thiết lập về 0. Không có động thái giảm dần cường độ đặc biệt nào được áp dụng trong quá trình che giấu.

Đối với CNG, trong G.718, sự tổng hợp CNG được thực hiện theo thứ tự sau đây. Đầu tiên, các tham số của khung nhiễu âm nền nhân tạo được giải mã. Sau đó, khung nhiễu âm nền nhân tạo được tổng hợp. Sau đó, vùng đệm bước được đặt lại. Sau đó, sự tổng hợp cho sự phân loại phục hồi lỗi khung (Frame Error Recovery - FER) được lưu lại. Sau đó, sự chỉnh giảm phổ được tiến hành. Sau đó, sự lọc sau tần số thấp được tiến hành. Sau đó, các biến CNG được cập nhật.

Trong trường hợp che giấu, chính xác là việc tương tự được thực hiện, ngoại trừ các tham số CNG không được giải mã từ dòng bit. Điều này có nghĩa là các tham số không được cập nhật trong quá trình mất khung, mà các tham số được giải mã từ khung mô tả việc chèn im lặng (Silence Insertion Descriptor - SID) tốt cuối cùng được sử dụng.

Đầu tiên, G.719 được xem xét. G.719, mà dựa trên Siren 22, là bộ mã hóa-giải mã âm thanh băng đầy đủ trên cơ sở biến đổi. ITU-T giới thiệu cho G.719 sự giảm dần cường độ với việc lặp lại khung trong miền phổ [ITU08b, phần 8.6]. Theo G.719, cơ cấu che giấu sự xóa khung được kết hợp vào trong bộ giải mã. Khi khung được nhận một cách chính xác, các hệ số biến đổi được khôi phục được lưu trữ trên vùng đệm. Nếu bộ giải mã được thông báo rằng khung đã bị mất hoặc khung đó bị sai lệch, các hệ số biến đổi được khôi phục trong khung được nhận gần nhất được định tỉ lệ giảm với thừa số 0,5 và sau đó được sử dụng như các hệ số biến đổi được khôi phục cho khung hiện thời. Bộ giải mã xử lý bằng cách biến đổi chúng đến miền thời gian và thực hiện thao tác tạo cửa sổ-chồng lắp-cộng.

Dưới đây, G.722 được mô tả. G.722 là hệ thống mã hóa từ 50 đến 7000 Hz mà sử dụng sự điều biến Mã Xung chên lệch thích ứng băng phụ (subband adaptive differential pulse code modulation - SB-ADPCM) nằm trong tốc độ truyền dữ liệu lên đến 64kbit/s. Tín hiệu được chia thành băng phụ cao hơn và thấp hơn, sử dụng sự phân tích bộ lọc gương cầu phương (Quadrature Mirror Filter - QMF). Hai băng thu được được mã hóa điều biến Mã Xung chên lệch thích ứng (Adaptive Differential Pulse Code Modulation - ADPCM).

Đối với G.722, thuật toán phức tạp cao cho sự che giấu mất gói được nêu cụ thể trong Phụ lục III [ITU06a] và thuật toán phức tạp thấp cho sự che giấu mất gói được nêu cụ thể trong Phụ lục IV [ITU07]. G.722 - Phụ lục III ([ITU06a, phần III.5]) đề xuất sự chặn tiếng được thực hiện dần dần, bắt đầu sau 20ms của sự mất khung, được hoàn thành sau 60ms của sự mất khung. Hơn nữa, G.722 - Phụ lục IV đề xuất công nghệ giảm dần cường độ mà áp dụng "đến mỗi thừa số khuếch đại mà được tính toán và được thích ứng từng mẫu" [ITU07, phần IV.6.1.2.7].

Trong G.722, quy trình chặn tiếng diễn ra trong miền băng phụ chỉ trước sự tổng hợp QMF và như bước cuối cùng của môđun PLC. Sự tính toán của thừa số chặn tiếng được thực hiện sử dụng thông tin loại từ bộ phân loại tín hiệu mà cũng là một phần của môđun PLC. Sự khác biệt được thực hiện giữa các loại TRANSIENT, UV\_TRANSITION và các loại khác. Hơn nữa, sự khác biệt được thực hiện giữa những sự mất đơn của khung 10ms và các trường hợp khác (nhiều sự mất của khung 10ms và nhiều sự mất của một/nhiều khung 20ms).

Điều này được minh họa bởi Fig.3. Cụ thể, Fig.3 mô tả cảnh, trong đó thừa số giảm dần cường độ của G.722, phụ thuộc vào thông tin loại và trong đó 80 mẫu tương đương với 10ms.

Theo G.722, môđun PLC tạo ra tín hiệu cho khung mất và một số tín hiệu bổ sung (10ms) mà được cho là được tăng giảm chéo với khung tốt tiếp theo. Sự chặn tiếng cho tín hiệu bổ sung cũng theo luật tương tự. Trong sự che giấu băng cao của G.722, sự tăng giảm chéo cũng không diễn ra.

Dưới đây, G.722.1 được xem xét. G.722.1, mà dựa trên Siren 7, là bộ mã hóa-giải mã âm thanh băng rộng trên cơ sở biến đổi với chế độ mở rộng băng siêu rộng, được đề cập đến như G.722.1C. Bản thân G. 722.1C được dựa trên Siren 14. ITU-T giới thiệu cho G.722.1 sự lặp lại khung với sự chặn tiếng theo sau trong miền phổ [ITU05, phần 4.7]. Nếu bộ giải mã được thông báo, bằng cách cơ cấu tín hiệu bên ngoài không được xác định trong khuyến nghị này, rằng khung đã bị mất hoặc bị sai lạc, nó lặp lại các hệ số biến đổi được lắp được điều biến (Modulated Lapped Transform - MLT) được mã hóa của khung trước. Nó xử lý bằng cách biến đổi chúng đến miền thời gian, và thực hiện thao tác chồng lắp và thêm với thông tin được giải mã

của khung trước và khung tiếp theo. Nếu khung trước cũng bị mất hoặc bị sai lạc, khi đó bộ giải mã thiết lập tất cả các hệ số MLT khung hiện thời đến 0.

Đầu tiên, G.729 được xem xét. G.729 là thuật toán nén dữ liệu âm thanh cho âm mà nén âm kỹ thuật số trong các gói có độ dài 10mili giây. Nó được mô tả một cách chính thức như sự mã hóa tiếng nói tại tốc độ 8kbit/giây sử dụng sự mã hóa tiếng nói dự báo tuyến tính được kích thích mã (CS-ACELP) [ITU12].

Như được nêu trong [CPK08], G.729 giới thiệu sự giảm dần cường độ trong miền LP. Thuật toán PLC được sử dụng trong tiêu chuẩn G.729 khôi phục tín hiệu tiếng nói cho khung hiện thời dựa trên thông tin tiếng nói được nhận trước đó. Nói cách khác, thuật toán PLC thay thế sự kích thích bị mất với đặc điểm tương đương của khung được nhận trước đó, mặc dù năng lượng kích thích dần dần giảm, cuối cùng, các hệ số của các bảng mã thích ứng và cố định được suy giảm bởi thừa số không thay đổi.

Độ khuếch đại bảng mã cố định được suy giảm được đưa ra bởi:

$$g_c^{(m)} = 0,98 \cdot g_c^{(m-1)}$$

với  $m$  là chỉ số khung phụ.

Độ khuếch đại bảng mã thích ứng được dựa trên phiên bản suy giảm của độ khuếch đại bảng mã thích ứng trước đó:

$$g_p^{(m)} = 0,9 \cdot g_p^{(m-1)} \quad , \text{ với điều kiện } g_p^{(m)} < 0,9$$

Nam in Park và cộng sự gợi ý cho G.729, bộ điều khiển biên độ tín hiệu sử dụng sự dự báo bằng cách hồi quy tuyến tính [CPK08, PKJ+11]. Nó đề cập đến sự mất gói khôi tín hiệu truyền và sử dụng sự hồi quy tuyến tính như kỹ thuật cốt lõi. Sự hồi quy tuyến tính được dựa trên mô hình tuyến tính như

$$g'_i = a + bi \quad (2)$$

trong đó  $g'_i$  là biên độ hiện thời được dự báo mới nhất,  $a$  và  $b$  là các hệ số cho hàm tuyến tính thứ nhất, và  $i$  là chỉ số của khung. Để tìm ra các hệ số được tối ưu hóa  $a^*$  và  $b^*$ , tổng sai số dự báo được bình phương được tối thiểu hóa:

$$\epsilon = \sum_{j=i-4}^{i-1} (g_j - g'_j)^2 \quad (3)$$

$\epsilon$  là sai số được bình phương,  $g_j$  là biên độ thứ  $j$  quá khứ ban đầu. Để tối thiểu hóa sai số này, đơn giản đạo hàm liên quan đến  $a$  và  $b$  được thiết lập là 0. Bằng cách sử dụng các tham số được tối ưu hóa  $a^*$  và  $b^*$ , sự ước lượng của mỗi  $g_i^*$  được biểu thị bởi

$$g_i^* = a^* + b^*i \quad (4)$$

Fig.4 thể hiện sự dự báo biên độ, cụ thể, sự dự báo của biên độ  $g_i^*$ , bằng cách sử dụng sự hồi quy tuyến tính.

Để thu được biên độ  $A'_i$  của gói đã mất  $i$ , tỉ lệ  $\sigma_i$

$$\sigma_i = \frac{g_i^*}{g_{i-1}} \quad (5)$$

được nhân với thừa số định tỉ lệ  $S_i$ :

$$A'_i = S_i * \sigma_i \quad (6)$$

trong đó thừa số định tỉ lệ  $S_i$  phụ thuộc vào số lượng khung được che giấu liên tiếp  $l(i)$ :

$$S_i = \begin{cases} 1,0, & \text{nếu } l(i) = 1, 2 \\ 0,9, & \text{nếu } l(i) = 3, 4 \\ 0,8, & \text{nếu } l(i) = 5, 6 \\ 0, & \text{nếu không thì} \end{cases} \quad (7)$$

Trong [PKJ+11], sự định tỉ lệ khác nhau được đề xuất.

Theo G.729, sau đó,  $A'_i$  sẽ được làm nhẵn để ngăn ngừa sự suy giảm rời rạc tại các đường biên khung. Cuối cùng, biên độ được làm nhẵn  $A_i(n)$  được nhân với sự kích thích, thu được từ các thành phần PLC trước đó.

Dưới đây, G.729.1 được xem xét. G.729.1 là bộ mã hóa tốc độ bit biến thiên được nhúng trên cơ sở G.729: Dòng bit bộ mã hóa băng rộng có thể định tỉ lệ 8-

32kbit/giây có thể liên vận hành với G.729 [ITU06b].

Theo G.729.1, như trong G.718 (xem bên trên), sự giảm dần cường độ thích ứng được đề xuất, mà phụ thuộc vào sự ổn định của các đặc điểm tín hiệu ([ITU06b, section 7.6.1]). Trong quá trình che giấu, tín hiệu thường được suy giảm dựa trên hệ số suy giảm  $\alpha$  mà phụ thuộc vào các tham số của loại khung được nhận tốt cuối cùng và số lượng của các khung bị xóa liên tiếp. Hệ số suy giảm  $\alpha$ , còn phụ thuộc vào độ ổn định của bộ lọc dự báo tuyến tính (Linear Prediction - LP) cho các khung VÔ THANH. Nhìn chung, sự suy giảm sẽ chậm nếu khung được nhận tốt cuối cùng nằm trong phân đoạn ổn định và sẽ nhanh nếu khung nằm trong phân đoạn chuyển tiếp.

Hơn nữa, hệ số suy giảm  $\alpha$  phụ thuộc vào độ khuếch đại bước trung bình mỗi khung  $\bar{g}_p$  ([ITU06b, eq. 163, 164]):

$$\bar{g}_p = 0,1g_p^{(0)} + 0,2g_p^{(1)} + 0,3g_p^{(2)} + 0,4g_p^{(3)} \quad (8)$$

trong đó  $g_p^{(i)}$  là độ khuếch đại bước trong khung phụ  $i$ .

Bảng 2 thể hiện sơ đồ tính toán của  $\alpha$ , trong đó

$$\beta = \sqrt{\bar{g}_p} \quad \text{với} \quad 0,85 \geq \beta \geq 0,98 \quad (9)$$

Trong suốt quy trình che giấu,  $\alpha$  được sử dụng trong công cụ che giấu sau:

| khung được nhận tốt cuối cùng | Số lượng các khung được xóa liên tiếp | $\alpha$    |
|-------------------------------|---------------------------------------|-------------|
| Hữu thanh                     | 1                                     | $\beta$     |
|                               | 2,3                                   | $\bar{g}_p$ |
|                               | > 3                                   | 0,4         |
| Sự bắt đầu                    | 1                                     | $0,8 \beta$ |
|                               | 2,3                                   | $\bar{g}_p$ |
|                               | > 3                                   | 0,4         |
| Sự bắt đầu nhân tạo           | 1                                     | $0.6 \beta$ |
|                               | 2,3                                   |             |

|                          |          |                    |
|--------------------------|----------|--------------------|
|                          | $> 3$    | $\bar{g}_p$<br>0,4 |
| Sự chuyển tiếp hữu thanh | $\leq 2$ | 0,8                |
|                          | $> 2$    | 0,2                |
| Sự chuyển tiếp vô thanh  |          | 0,88               |
| Vô thanh                 | 1        | 0,95               |
|                          | 2,3      | $0,6 \theta + 0,4$ |
|                          | $> 3$    | 0,4                |

Bảng 2 Các giá trị của hệ số suy giảm  $\alpha$ , giá trị  $\theta$  là hệ số ổn định được tính toán từ số đo khoảng cách giữa các bộ lọc LP liền kề. [ITU06b, phần 7.6.1].

Theo G.729.1, về sự tái đồng bộ hóa xung âm tắc, như xung cuối cùng của sự kích thích của khung trước đó được sử dụng cho sự xây dựng phần chu kỳ, độ khuếch đại của nó là xấp xỉ đúng tại lúc bắt đầu của khung được che giấu và có thể thiết lập là 1. Độ khuếch đại sau đó bị suy giảm tuyến tính trong suốt khung trên cơ sở từng mẫu để đạt được giá trị của  $\alpha$  tại lúc kết thúc khung. Sự tiến triển năng lượng của các phân đoạn hữu thanh được ngoại suy bằng cách sử dụng các giá trị độ khuếch đại kích thích bước của mỗi khung phụ của khung tốt cuối cùng. Nhìn chung, nếu các độ khuếch đại này lớn hơn 1, năng lượng tín hiệu sẽ tăng, nếu chúng nhỏ hơn 1, năng lượng sẽ giảm.  $\alpha$  do đó được thiết lập đến  $\beta = \sqrt{\bar{g}_p}$  như mô tả ở trên, xem [ITU06b, eq. 163, 164]. Giá trị của  $\beta$  được cắt giữa 0,98 và 0,85 để tránh những sự tăng và giảm năng lượng mạnh, xem [ITU06b, phần 7.6.4].

Liên quan đến việc xây dựng các phần ngẫu nhiên của sự kích thích, theo G.729.1, tại lúc bắt đầu khối bị xóa, độ khuếch đại đổi mới  $g_s$  được thiết lập ban đầu bằng cách sử dụng các độ khuếch đại kích thích đổi mới của mỗi khung phụ của khung tốt cuối cùng:

$$g_s = 0,1g^{(0)} + 0,2g^{(1)} + 0,3g^{(2)} + 0,4g^{(3)}$$

trong đó  $g^{(0)}$ ,  $g^{(1)}$ ,  $g^{(2)}$  và  $g^{(3)}$  là bảng mã cố định, hoặc độ khuếch đại đổi mới của bốn khung phụ của khung được nhận chính xác cuối cùng. Sự suy giảm độ khuếch đại

đổi mới được thực hiện như:

$$g_s^{(1)} = \alpha \cdot g_s^{(0)}$$

trong đó  $g_s^{(1)}$  là độ khuếch đại đổi mới tại lúc bắt đầu của khung tiếp theo,  $g_s^{(0)}$  là độ khuếch đại đổi mới tại lúc bắt đầu của khung hiện thời, và  $\alpha$  như được xác định trong bảng 2 nêu trên. Tương tự với sự suy giảm kích thích chu kỳ, độ khuếch đại do đó được suy giảm tuyến tính trong taonf khung trên cơ sở từng mẫu bắt đầu với  $g_s^{(0)}$  và đến giá trị  $g_s^{(1)}$  mà sẽ đạt được tại lúc bắt đầu khung tiếp theo.

Theo G.729.1, nếu khung tốt cuối cùng là VÔ THANH, chỉ sự kích thích đổi mới được sử dụng và nó còn được suy giảm bởi hệ số là 0,8. Trong trường hợp này, vùng đệm kích thích quá khứ được cập nhật với sự kích thích đổi mới như không có phần chu kỳ của sự kích thích có sẵn, xem [ITU06b, phần 7.6.6].

Dưới đây, AMR được xem xét. 3GPP AMR [3GP12b] là bộ mã hóa-giải mã tiếng nói sử dụng thuật toán ACELP. AMR có thể mã hóa tiếng nói với tốc độ lấy mẫu là 8000 mẫu/giây và tốc độ truyền dữ liệu giữa 4,75 và 12,2 kbit/giây và hỗ trợ việc tạo tín hiệu các khung mô tả sự im lặng (DTX/CNG).

Trong AMR, trong quá trình che giấu lỗi (xem [3GP12a]), nó được phân biệt giữa các khung mà dễ bị sai số (các sai số bit) và các khung, mà hoàn toàn bị mất (không có dữ liệu nào).

Đối với sự che giấu ACELP, AMR dẫn đến máy trạng thái mà ước lượng chất lượng của kênh: Giá trị của bộ đếm trạng thái càng lớn, chất lượng kênh càng kém. Hệ thống bắt đầu ở trạng thái 0. Mỗi lần khung xấu được phát hiện, bộ đếm trạng thái mà được tăng lên bởi một và được bão hòa khi đạt đến 6. Mỗi lần khung tiếng nói tốt được phát hiện, bộ đếm trạng thái được đặt lại về 0, ngoại trừ khi trạng thái là 6, trong đó bộ đếm trạng thái được thiết lập là 5. Dòng kiểm soát của máy trạng thái có thể được mô tả bởi mã C sau (BFI là bộ chỉ thị khung xấu, State là biến trạng thái):

```
if (BFI != 0 ) {
    State = State + 1;
}
```

```

else if(State == 6) {
    State = State + 5;
}
else {
    State = 0;
}
if(State > 6 ) {
    State = 6;
}

```

Ngoài máy trạng thái này, trong AMR, các cờ hiệu khung xấu từ các khung hiện thời và khung trước đó được kiểm tra (prevBFI).

Ba tổ hợp khác nhau là khả thi:

Tổ hợp thứ nhất trong số ba tổ hợp là BFI = 0, prevBFI = 0, State = 0: Không có sai số được phát hiện trong khung tiếng nói được nhận hoặc trong khung tiếng nói được nhận trước đó. Các tham số tiếng nói được nhận được sử dụng bằng cách thông thường trong sự tổng hợp tiếng nói. Khung hiện thời của các tham số tiếng nói được lưu.

Tổ hợp thứ hai trong số ba tổ hợp là BFI = 0, prevBFI = 1, State = 0 hoặc 5: Không có sai số được phát hiện trong khung tiếng nói được nhận, nhưng khung tiếng nói được nhận trước đó là xấu. Độ khuếch đại LTP và độ khuếch đại bằng mã cố định được hạn chế dưới các giá trị được sử dụng cho khung phụ tốt được nhận cuối cùng:

$$g_p = \begin{cases} g_p, & g_p \leq g_p(-1) \\ g_p(-1), & g_p > g_p(-1) \end{cases} \quad (10)$$

trong đó  $g_p$  = độ khuếch đại LTP được giải mã hiện thời,  $g_p(-1)$  = độ khuếch đại LTP được sử dụng cho khung phụ tốt cuối cùng, và

(BFI = 0), và

$$g_c = \begin{cases} g_c, & g_c \leq g_c(-1) \\ g_c(-1), & g_c > g_c(-1) \end{cases} \quad (11)$$

trong đó  $g_c$  = độ khuếch đại bảng mã cố định được giải mã hiện thời, và  $g_c(-1)$  = độ khuếch đại bảng mã cố định được sử dụng cho khung phụ tốt cuối cùng (BFI = 0).

Phần còn lại của các tham số tiếng nói được nhận thường được sử dụng trong sự tổng hợp tiếng nói. Khung hiện thời của các tham số tiếng nói được lưu.

Tổ hợp thứ ba trong số ba tổ hợp là BFI = 1, prevBFI = 0 hoặc 1, State = 1...6: Sai số được phát hiện trong khung tiếng nói được nhận và sự thay thế và quy trình chặn tiếng được bắt đầu. Độ khuếch đại LTP và độ khuếch đại bảng mã cố định được thay thế bằng các giá trị được suy giảm từ các khung phụ trước đó.

$$g_p = \begin{cases} P(\text{state}) \cdot g_p(-1), & g_p(-1) \leq \text{median5}(g_p(-1), \dots, g_p(-5)) \\ P(\text{state}) \cdot \text{median5}(g_p(-1), \dots, g_p(-5)) & g_p(-1) > \text{median5}(g_p(-1), \dots, g_p(-5)) \end{cases} \quad (12)$$

trong đó  $g_p$  biểu thị độ khuếch đại LTP được giải mã hiện thời và  $g_p(-1), \dots, g_p(-n)$  biểu thị các độ khuếch đại LTP được sử dụng cho các khung phụ  $n$  cuối cùng và median5() biểu thị thao tác trung bình 5 điểm và

$P(\text{state})$  = hệ số suy giảm,

trong đó ( $P(1) = 0,98, P(2) = 0,98, P(3) = 0,8, P(4) = 0,3, P(5) = 0,2, P(6) = 0,2$ ) và state = số trạng thái (state number), và

$$g_c = \begin{cases} C(\text{state}) \cdot g_c(-1), & g_c(-1) \leq \text{median5}(g_c(-1), \dots, g_c(-5)) \\ C(\text{state}) \cdot \text{median5}(g_c(-1), \dots, g_c(-5)) & g_c(-1) > \text{median5}(g_c(-1), \dots, g_c(-5)) \end{cases} \quad (13)$$

trong đó  $g_c$  biểu thị độ khuếch đại bảng mã cố định được giải mã hiện thời và  $g_c(-1), \dots, g_c(-n)$  biểu thị độ khuếch đại bảng mã cố định được sử dụng cho các khung phụ  $n$  cuối cùng và median5() biểu thị thao tác trung bình 5 điểm và  $C(\text{state})$  = hệ số suy giảm, trong đó ( $C(1) = 0,98, C(2) = 0,98, C(3) = 0,98, C(4) = 0,98, C(5) = 0,98, C(6) = 0,7$ ) và state = số trạng thái.

Trong AMR, các giá trị độ trễ dự báo dài hạn (Long-Term Prediction - LTP)

được thay thế bởi giá trị quá khứ từ khung phụ thứ 4 của khung phía trước (chế độ 12.2) hoặc các giá trị được thay đổi nhẹ dựa trên giá trị được nhận một cách chính xác cuối cùng (tất cả các chế độ khác).

Theo AMR, các xung đổi mới của bảng mã cố định được nhận từ các khung lỗi được sử dụng trong trạng thái mà chúng được nhận khi dữ liệu sai lạc được nhận. Trong trường hợp không có dữ liệu nào là các chỉ số bảng mã cố định ngẫu nhiên nên được sử dụng.

Đối với CNG trong AMR, theo [3GP12a, phần 6.4], mỗi khung SID bị mất thứ nhất được thay thế bằng cách sử dụng thông tin SID từ các khung SID có giá trị được nhận và quy trình cho các khung SID có giá trị được áp dụng. Đối với các khung SID bị mất liên tiếp, công nghệ suy giảm được áp dụng cho nhiều âm nền nhân tạo mà sẽ giảm dần mức đầu ra. Do đó, nó sẽ được kiểm tra xem liệu cập nhật SID cuối cùng trước đây là hơn 50 khung ( $=1\text{ s}$ ), nếu có, đầu ra sẽ được chặn tiếng (sự suy giảm mức là  $-6/8\text{dB}$  mỗi khung [3GP12d, `dtx_dec{}@sp_dec.c`] mà sản xuất  $37,5\text{dB}$  mỗi giây). Lưu ý rằng sự giảm dần cường độ được áp dụng cho CNG được thực hiện trong miền LP.

Dưới đây, AMR-WB được xem xét. Đa tốc độ thích ứng (Adaptive Multirate - WB) [ITU03, 3GP09c] là bộ mã hóa-giải mã tiếng nói, ACELP, dựa trên AMR (xem phần 1.8). Nó sử dụng sự mở rộng băng thông theo tham số và cũng hỗ trợ DTX/CNG. Trong bản mô tả của tiêu chuẩn [3GP12g] có các giải pháp ví dụ về sự che giấu được cho mà giống đối với AMR [3GP12a] với các độ lệch nhỏ. Do đó, chỉ các khác biệt với AMR được mô tả ở đây. Đối với sự mô tả tiêu chuẩn, xem phần mô tả trên đây.

Đối với ACELP, trong AMR-WB, sự giảm dần cường độ ACELP được thực hiện dựa trên mã nguồn tham chiếu [3GP12c] bằng cách thay đổi độ khuếch đại bước  $g_p$  (đối với AMR bên trên được đề cập đến như độ khuếch đại LTP) và bằng cách thay đổi độ khuếch đại bảng mã  $g_c$ .

Trong trường hợp khung bị mất, độ khuếch đại bước  $g_p$  cho khung phụ thứ nhất là giống với khung tốt cuối cùng, ngoại trừ việc nó bị giới hạn giữa 0,95 và 0,5. Đối với các khung phụ thứ hai, thứ ba và tiếp theo đây, độ khuếch đại bước  $g_p$  giảm đi thừa số là 0,95 và lại bị giới hạn.

AMR-WB đề xuất rằng trong khung được che giấu,  $g_c$  được dựa trên  $g_{c, \text{cuối}}$  cùng:

$$g_{c, \text{current}} = g_{c, \text{past}} * (1,4 - g_{p, \text{past}}) \quad (14)$$

$$g_c = g_{c, \text{current}} * g_{c, \text{inov}} \quad (15)$$

$$g_{c, \text{inov}} = \frac{1,0}{\sqrt{\frac{\text{ener}_{\text{inov}}}{\text{subframe\_size}}}} \quad (16)$$

$$\text{ener}_{\text{inov}} = \sum_{i=0}^{\text{subframe\_size}-1} \text{code}[i] \quad (17)$$

Đối với việc che giấu các độ trễ LTP, trong AMR-WB, lịch sử 5 độ trễ LTP và các độ khuếch đại LTP tốt cuối cùng được sử dụng để tìm kiếm phương pháp tốt nhất để cập nhật, trong trường hợp mất khung. Trong trường hợp khung được nhận với các sai số bit, sự dự báo được thực hiện, liệu độ trễ LTP có thể được sử dụng hay không [3GP12g].

Đối với CNG, trong AMR-WB, nếu khung được nhận chính xác cuối cùng là khung SID và khung được phân loại là đã mất, nó có thể được thay thế bởi thông tin khung SID có giá trị và quy trình cho các khung SID có giá trị nên được áp dụng.

Đối với các khung SID bị mất liên tiếp, AMR-WB đề xuất áp dụng công nghệ suy giảm cho nhiều âm nền nhân tạo mà sẽ giảm dần mức đầu ra. Do đó, nó sẽ được kiểm tra xem liệu cập nhật SID cuối cùng trước đây là lớn hơn 50 khung ( $=1$  s), nếu có, đầu ra sẽ bị chặn tiếng (sự suy giảm mức là  $-3/8$ dB mỗi khung [3GP12f, `dtx_dec{}@dtx.c`] mà tạo ra 18,75 dB mỗi giây). Lưu ý rằng sự giảm dần cường độ được áp dụng cho CNG được thực hiện trong miền LP.

Bây giờ, AMR-WB+ được xem xét. Đa tốc độ thích ứng - WB+ [3GP09a] là bộ mã hóa-giải mã được chuyển sử dụng ACELP và sự kích thích được mã hóa biến đổi (Transform Coded Excitation - TCX) như các bộ mã hóa-giải mã lõi. Nó sử dụng sự mở rộng băng thông theo tham số và cũng hỗ trợ DTX/CNG.

Trong AMR-WB+, logic phép ngoại suy chế độ được áp dụng để ngoại suy các

chế độ của các khung bị mất trong siêu khung bị méo. Phép ngoại suy chế độ này được dựa trên thực tế rằng tồn tại phần dư trong sự xác định các bộ chỉ thị chế độ. Logic quyết định (được đưa ra trong [3GP09a, hình 18]) được đề xuất bởi AMR-WB+ là như sau:

- Chế độ véctor,  $(m_{-1}, m_0, m_1, m_2, m_3)$ , được xác định, trong đó  $m_{-1}$  biểu thị chế độ của khung cuối cùng của siêu khung trước đó và  $m_0, m_1, m_2, m_3$  biểu thị các chế độ của các khung trong siêu khung hiện thời (được giải mã từ dòng bit), trong đó  $m_k = -1, 0, 1, 2$  hoặc  $3$  ( $-1$ : bị mất,  $0$ : ACELP,  $1$ : TCX20,  $2$ : TCX40,  $3$ : TCX80), và trong đó số khung bị mất  $n_{loss}$  có thể nằm giữa  $0$  và  $4$ .

- Nếu  $m_{-1} = 3$  và hai trong số các bộ chỉ thị chế độ của các khung  $0 - 3$  là bằng ba, tất cả các bộ chỉ thị được thiết lập về ba vì sau đó chắc chắn rằng khung TCX80 đã được biểu thị trong siêu khung.

- Nếu chỉ một bộ chỉ thị của các khung  $0 - 3$  là ba (và số lượng các khung bị mất  $n_{loss}$  là ba), chế độ sẽ được thiết lập đến  $(1, 1, 1, 1)$ , vì sau đó 3/4 trong số phổ đích TCX80 bị mất và có khả năng rằng độ khuếch đại TCX toàn cục bị mất.

- Nếu chế độ biểu thị  $(x, 2, -1, x, x)$  hoặc  $(x, -1, 2, x, x)$ , nó sẽ được ngoại suy đến  $(x, 2, 2, x, x)$ , biểu thị khung TCX40. Nếu chế độ biểu thị  $(x, x, x, 2, -1)$  hoặc  $(x, x, -1, 2)$  nó sẽ được ngoại suy đến  $(x, x, x, 2, 2)$ , cũng biểu thị khung TCX40. Nên lưu ý rằng  $(x, [0, 1], 2, 2, [0, 1])$  là các cấu hình không có giá trị.

- Sau đó, đối với mỗi khung mà bị mất ( $mode = -1$ ), chế độ được thiết lập về ACELP ( $mode = 0$ ) nếu khung trước đó là ACELP và chế độ được thiết lập về TCX20 ( $mode = 1$ ) cho tất cả các trường hợp khác.

Đối với ACELP, theo AMR-WB+, nếu chế độ các khung bị mất dẫn đến kết quả  $m_k = 0$  sau phép ngoại suy chế độ, phương pháp tương tự như trong [3GP12g] được áp dụng cho khung này (xem trên).

Trong AMR-WB+, phụ thuộc vào số lượng các khung bị mất và chế độ được ngoại suy, các phương pháp che giấu liên quan TCX sau đây được phân biệt:

- Nếu toàn bộ khung bị mất, thì sự che giấu như ACELP được áp dụng: Sự kích thích cuối cùng được lặp lại và các hệ số ISF được che giấu (chuyển dịch dần dần về

phía giá trị trung bình thích ứng của chúng) được sử dụng để tổng hợp tín hiệu miền thời gian. Ngoài ra, thừa số giảm dần cường độ là 0,7 mỗi khung (20ms) [3GP09b, `dec_tcxc.c`] được nhân trong miền dự báo tuyến tính, ngay trước sự tổng hợp mã hóa dự báo tuyến tính (Linear Predictive Coding - LPC).

- Nếu chế độ cuối cùng là TCX80 cũng như chế độ được ngoại suy của siêu khung (bị mất một phần( là TCX80 ( $nloss = [1, 2]$ ,  $mode = (3, 3, 3, 3, 3)$ ), sự che giấu được thực hiện trong miền FFT, sử dụng pha và phép ngoại suy biên độ, xem xét khung được nhận chính xác cuối cùng. Phương pháp ngoại suy của thông tin pha là không được quan tâm ở đây (không liên quan đến chiến lược tăng giảm cường độ) và do đó không được mô tả. Chi tiết hơn, xem [3GP09a, phần 6.5.1.2.4]. Đối với sự thay đổi biên độ của AMR-WB+, phương pháp được thực hiện đối với sự che giấu TCX bao gồm các bước sau [3GP09a, phần 6.5.1.2.3]:

- Phổ tuyệt đối khung trước đó được tính toán:

$$oldA[k] = |old\hat{X}[k]|$$

- Phổ tuyệt đối khung hiện thời được tính toán:

$$A[k] = |\hat{X}[k]|$$

- Sự chênh lệch độ khuếch đại của năng lượng của các hệ số phổ không bị mất giữa khung trước và khung hiện thời được tính toán:

$$gain = \sqrt{\frac{\sum A[k]^2}{\sum oldA[k]^2}}$$

- Biên độ của các hệ số phổ bị mất được sử dụng ngoại suy:

$$\text{nếu}(lost[k]) \quad A[k] = gain \cdot oldA[k]$$

- Trong từng trường hợp khác của khung bị mất với  $m_k = [2, 3]$ , đích TCX (FFT nghịch đảo của phổ được mã hóa cùng với sự lấp đầy nhiễu âm (sử dụng mức nhiễu âm được giải mã từ dòng bit)) được tổng hợp sử dụng tất cả thông tin có sẵn (bao gồm độ khuếch đại TCX toàn cục). Không có sự giảm dần cường độ được áp dụng trong trường hợp này.

Đối với CNG trong AMR-WB+, phương pháp tương tự như trong AMR-WB được sử dụng (xem ở trên).

Dưới đây, OPUS được xem xét. OPUS [IET12] kết hợp công nghệ từ hai bộ mã hóa-giải mã: SILK hướng tiếng nói (được biết đến như bộ mã hóa-giải mã Skype) và phép biến đổi được lắp miễn cưỡng năng lượng (Constrained-Energy Lapped Transform - CELT) thời gian chờ thấp. Opus có thể được điều chỉnh liên mạch giữa các tốc độ truyền dữ liệu thấp và cao, và bên trong, nó chuyển đổi giữa bộ mã hóa-giải mã dự báo tuyến tính tại tốc độ truyền dữ liệu thấp (SILK) và bộ mã hóa-giải mã biến đổi tại tốc độ truyền dữ liệu cao hơn (CELT) cũng như phép lai cho sự chồng lắp ngắn.

Đối với việc nén và giải nén dữ liệu âm thanh SILK, trong OPUS, có một số tham số mà được suy giảm trong quá trình che giấu trong đoạn chương trình giải mã SILK. Tham số độ khuếch đại LTP được suy giảm bằng cách nhân tất cả các hệ số LPC với hoặc 0,99, 0,95 hoặc 0,90 mỗi khung, phụ thuộc vào số lượng các khung bị mất liên tiếp, trong đó sự kích thích được xây dựng sử dụng chu trình bước cuối cùng từ sự kích thích của khung trước. Tham số độ trễ bước tăng rất chậm trong quá trình mất liên tiếp. Đối với những sự mất đơn, nó được giữ không thay đổi so với khung trước. Hơn nữa, tham số độ khuếch đại kích thích được suy giảm theo hàm mũ với  $0,99^{lost_{cnt}}$  mỗi khung, để tham số độ khuếch đại kích thích là 0,99 cho tham số độ khuếch đại kích thích thứ nhất, để tham số độ khuếch đại kích thích là 0,992 đối với tham số độ khuếch đại kích thích thứ hai, và tiếp tục như vậy. Sự kích thích được tạo ra sử dụng bộ tạo số ngẫu nhiên mà tạo ra nhiễu âm trắng bằng phần tràn biến thiên. Hơn nữa, các hệ số LPC được ngoại suy/được tính trung bình trên cơ sở tập hợp các hệ số được nhận chính xác cuối cùng. Sau khi tạo ra vectơ kích thích được suy giảm, các hệ số LPC được che giấu được sử dụng trong OPUS để tổng hợp tín hiệu đầu ra miền thời gian.

Bây giờ, trong ngữ cảnh của OPUS, CELT được xem xét. CELT là bộ mã hóa-giải mã trên cơ sở biến đổi. Sự che giấu của CELT mô tả nét nổi bật của phương pháp PLC dựa trên bước, mà được áp dụng cho lên đến năm khung bị mất liên tiếp. Bắt đầu với khung 6, phương pháp che giấu tương tự nhiễu âm được áp dụng, mà tạo ra tạp âm nền, mà đặc điểm được cho là phát ra âm như tạp âm nền đứng trước.

Fig.5 minh họa động thái mất khối tín hiệu truyền của CELT. Cụ thể, Fig.5 mô tả ảnh phổ (trục x: thời gian; trục y: tần số) của phân đoạn tiếng nói được che giấu CELT. Hộp xám nhạt biểu thị 5 khung bị mất liên tiếp thứ nhất, trong đó phương pháp PLC dựa trên bước được áp dụng. Hơn nữa, sự che giấu như nhiễu âm được thể hiện. Nên lưu ý rằng sự chuyển đổi được thực hiện một cách ngay lập tức, nó không chuyển tiếp một cách trôi chảy.

Đối với sự che giấu trên cơ sở bước, trong OPUS, sự che giấu trên cơ sở bước bao gồm việc tìm kiếm chu kỳ trong tín hiệu được giải mã bằng phép tự tương quan và lặp lại dạng sóng được tạo cửa sổ (trong miền kích thích sử dụng sự phân tích và tổng hợp LPC) sử dụng sự chênh lệch bước (độ trễ bước). Dạng sóng được tạo cửa sổ được chồng lấp theo cách để bảo toàn sự triệt tiêu sai số lấy mẫu trong miền thời gian với khung trước đó và khung tiếp theo [IET12]. Ngoài ra, hệ số giảm dần cường độ được đạo hàm và được áp dụng bởi mã sau:

```
opus_val32 E1=1, E2=1;

int period;

if (pitch_index <= MAX_PERIOD/2) {
    period = pitch_index;
}

else {
    period = MAX_PERIOD/2;
}

for (i=0;i<period;i++)
{
    E1 += exc[MAX_PERIOD- period+i] * exc[MAX_PERIOD-
period+i];

    E2 += exc[MAX_PERIOD-2*period+i] * exc[MAX_PERIOD-
2*period+i];
}
```

```

if (E1 > E2) {
    E1 = E2;
}

decay = sqrt(E1/E2);

attenuation = decay;

```

Trong mã này, exc chứa tín hiệu kích thích lên đến các mẫu MAX\_PERIOD trước sự mất.

Tín hiệu kích thích sau đó được nhân với sự suy giảm, sau đó được tổng hợp và được phát ra qua sự tổng hợp LPC.

Thuật toán tăng giảm cường độ cho phương pháp miền thời gian có thể được tóm tắt như sau:

- Tìm năng lượng đồng bộ bước của chu kỳ bước cuối cùng trước sự mất.
- Tìm năng lượng đồng bộ bước của chu kỳ bước cuối cùng thứ hai trước sự mất.
- Nếu năng lượng đang tăng, giới hạn nó để duy trì ở mức không thay đổi: sự suy giảm = 1
- Nếu năng lượng đang giảm, tiếp tục với sự suy giảm tương tự trong quá trình che giấu.

Đối với sự che giấu như nhiều âm, theo OPUS, đối với khung thứ 6 và các khung bị mất liên tiếp dưới đây, phương pháp thay thế nhiều âm trong miền MDCT được thực hiện, để tái tạo tập âm nền nhân tạo.

Đối với việc tạo vết mức và hình dạng tập âm nền, trong OPUS, sự ước lượng tập âm nền được thực hiện như sau: Sau khi phân tích MDCT, căn bậc hai của năng lượng băng MDCT được tính toán mỗi băng tần số, trong đó nhóm các ngăn MDCT tiếp sau sự định tỉ lệ Bark theo [IET12, Bảng 55]. Sau đó căn bậc hai của các năng lượng được biến đổi thành miền  $\log_2$  băng:

$$\text{bandLogE}[i] = \log_2(e) \cdot \log_e(\text{bandE}[i] - eMeans[i]) \quad \text{đối với } i = 0 \dots 21 \quad (18)$$

trong đó  $e$  là số lượng của Euler,  $\text{bandE}$  là căn bậc hai của băng MDCT và

$eMeans$  là véctor của các hằng số (cần thiết để lấy được kết quả trung bình 0, mà dẫn đến độ khuếch đại mã hóa được tăng cường).

Trong OPUS, tạp âm nền được đăng nhập trên phía bộ giải mã giống như [IET12, amp2Log2 và log2Amp @ quant\_bands.c]:

$$backgroundLogE[i] = \min(backgroundLogE[i] + 8 \cdot 0,001, bandLogE[i])$$

$$\text{đối với } i = 0 \dots 21 \quad (19)$$

Năng lượng tối thiểu được tạo vệt về cơ bản được xác định bởi căn bậc hai của năng lượng của băng của khung hiện thời, nhưng sự tăng lên từ một khung đến khung tiếp theo được giới hạn là 0,05dB.

Đối với việc áp dụng mức và hình dạng tạp âm nền, theo OPUS, nếu PLC như nhiều âm được áp dụng,  $backgroundLogE$  như được đạo hàm trong khung tốt cuối cùng được sử dụng và được chuyển đổi trở về miền tuyến tính:

$$backgroundLogE[i] = \min(backgroundLogE[i] + 8 \cdot 0,001, bandLogE[i]) \quad (20)$$

trong đó  $e$  là số Euler và  $eMeans$  là véctor tương tự của các hằng số như đối với sự biến đổi "tuyến tính về log".

Quy trình che giấu hiện thời là để lấp đầy khung MDCT với nhiễu âm trắng được sản xuất bởi bộ tạo số ngẫu nhiên, và định tỉ lệ nhiễu âm trắng này theo cách mà hợp theo băng với năng lượng của  $bandE$ . Theo đó, MDCT nghịch đảo được áp dụng mà dẫn đến kết quả trong tín hiệu miền thời gian. Sau sự bổ sung và chỉnh giảm chồng lấp (như trong việc giải mã thông thường), nó sẽ được xuất ra.

Dưới đây, MPEG-4 HE-AAC được xem xét (MPEG là Nhóm các chuyên gia hình ảnh động; HE-AAC là sự mã hóa âm thanh tiên tiến hiệu suất cao). Sự mã hóa âm thanh tiên tiến hiệu suất cao bao gồm bộ mã hóa-giải mã âm thanh trên cơ sở biến đổi (AAC), được bổ sung bởi sự mở rộng băng thông theo tham số (parametric bandwidth extension - SBR).

Đối với sự mã hóa âm thanh tiên tiến (Advanced Audio Coding - AAC), liên

doanh DAB chỉ rõ rằng đối với AAC trong DAB+, sự giảm dần cường độ về 0 trong miền tần số [EBU10, phần A1.2] (Phát sóng âm thanh kỹ thuật số (Digital Audio Broadcasting - DAB)). Động thái giảm dần cường độ, ví dụ, đường dốc suy giảm, có thể là cố định hoặc có thể được điều chỉnh bởi người sử dụng. Các hệ số phổ từ bộ phận truy cập (Access Unit - AU) cuối cùng được suy giảm bởi hệ số tương ứng với các đặc điểm giảm dần cường độ và sau đó được chuyển qua việc ánh xạ tần số đến thời gian. Phụ thuộc vào đường dốc suy giảm, sự che giấu chuyển đến việc chặn tiếng sau một số các AU không giá trị liên tiếp, điều này có nghĩa là phổ hoàn thiện sẽ được thiết lập về 0.

Liên doanh quản lý bản quyền số (Digital Rights Management - DRM) chỉ rõ đối với AAC trong DRM sự giảm dần cường độ trong miền tần số [EBU12, phần 5.3.3]. Sự che giấu làm việc trên dữ liệu phổ ngay trước khi chuyển đổi tần số đến thời gian cuối cùng. Nếu nhiều khung bị sai lệch, sự che giấu thực hiện thứ nhất là sự giảm dần cường độ trên các giá trị phổ được thay đổi nhẹ từ khung có giá trị cuối cùng. Hơn nữa, tương tự với DAB+, động thái giảm dần cường độ, ví dụ, đường dốc suy giảm, có thể là cố định hoặc có thể được điều chỉnh bởi người sử dụng. Các hệ số phổ từ khung cuối cùng được suy giảm bởi hệ số tương ứng với các đặc điểm giảm dần cường độ và sau đó được chuyển qua việc ánh xạ tần số đến thời gian. Phụ thuộc vào đường dốc suy giảm, sự che giấu chuyển đến việc chặn tiếng sau một số các khung không giá trị liên tiếp, điều này có nghĩa là phổ hoàn thiện sẽ được thiết lập về 0.

3GPP giới thiệu cho AAC trong aacPlus nâng cao sự giảm dần cường độ trong miền tần số tương tự DRM [3GP12e, phần 5.1]. Sự che giấu làm việc trên dữ liệu phổ ngay trước khi chuyển đổi tần số đến thời gian cuối cùng. Nếu nhiều khung bị sai lệch, sự che giấu thực hiện thứ nhất là sự giảm dần cường độ trên các giá trị phổ được thay đổi nhẹ từ khung tốt cuối cùng. Sự giảm dần cường độ hoàn thiện chiếm 5 khung. Các hệ số phổ từ khung tốt cuối cùng được sao chép và được suy giảm bởi hệ số:

$$fadeOutFac = 2^{-(nFadeOutFrame/2)}$$

với  $nFadeOutFrame$  như bộ đếm khung từ khung tốt cuối cùng. Sau năm khung giảm dần cường độ, sự che giấu chuyển về việc chặn tiếng, điều này có nghĩa là phổ hoàn thiện sẽ được thiết lập về 0.

Lauber và Sperschneider giới thiệu cho AAC sự giảm dần cường độ theo khung của phổ MDCT, dựa trên phép ngoại suy năng lượng [LS01, phần 4.4]. Việc tạo năng lượng phổ đứng trước sẽ được sử dụng để ngoại suy hình dạng của phổ được ước lượng. Phép ngoại suy năng lượng có thể được thực hiện độc lập với các công nghệ che giấu như loại che giấu trước.

Đối với AAC, sự tính toán năng lượng được thực hiện trên cơ sở băng hệ số tỉ lệ để được gần với các băng giới hạn của hệ thống thính giác con người. Các giá trị năng lượng riêng lẻ được giảm trên cơ sở từng khung để giảm âm lượng một cách trôi chảy, ví dụ, để giảm dần cường độ tín hiệu. Điều này trở nên cần thiết vì xác suất, mà các giá trị được ước lượng biểu diễn tín hiệu hiện thời, giảm nhanh theo thời gian.

Để tạo ra phổ, họ đề nghị lặp lại khung hoặc thay thế nhiều âm [LS01, các phần 3.2 và 3.3].

Quackenbusch và Driesen gợi ý cho AAC sự giảm dần cường độ theo khung hàm mũ về 0 [QD03]. Việc lặp lại tập hợp các hệ số thời gian/tần số liền kề được đề xuất, trong đó mỗi lần lặp lại có sự suy giảm tăng theo hàm mũ, do đó sự tăng giảm cường độ dần dần về sự chặn tiếng trong trường hợp trạng thái ngừng hoạt động được kéo dài.

Đối với sự tái tạo băng phổ (Spectral Band Replication = SBR) trong MPEG-4 HE-AAC, 3GPP gợi ý cho SBR trong aacPlus nâng cao để đệm dữ liệu đường bao được giải mã và, trong trường hợp mất khung, để tái sử dụng các năng lượng được đệm của dữ liệu đường bao được truyền và để giảm chúng bằng tỉ lệ không thay đổi là 3dB cho từng khung được che giấu. Kết quả được đưa vào quy trình giải mã thông thường trong đó bộ điều chỉnh đường bao sử dụng nó để tính toán các độ khuếch đại, được sử dụng để điều chỉnh các băng cao được đắp và được tạo bởi bộ tạo HF. Sự giải mã SBR sau đó được diễn ra như bình thường. Hơn nữa, các giá trị sàn nhiễu và mức sin được mã hóa delta đang bị xóa. Vì không có sự khác biệt với thông tin trước đó duy trì có sẵn, sàn nhiễu và các mức sin được giải mã vẫn tỉ lệ thuận với năng lượng của tín hiệu được tạo HF [3GP12e, phần 5.2].

Liên doanh DRM đã chỉ rõ cho SBR trong sự liên kết với AAC kỹ thuật tương tự như 3GPP [EBU12, phần 5.6.3.1]. Hơn nữa, liên doanh DAB chỉ rõ cho SBR trong

DAB+ kỹ thuật tương tự như 3GPP [EBU10, phần A2].

Sau đây, CELP MPEG-4 và Mã hóa kích thích hài hòa MPEG-4 (Harmonic Vector Excitation Coding - HVXC) được xem xét. Liên doanh DRM chỉ rõ cho SBR liên kết với CELP và HVXC [EBU12, phần 5.6.3.2] rằng sự che giấu yêu cầu tối thiểu cho SBR cho các bộ mã hóa-giải mã tiếng nói là để áp dụng tập hợp được định trước các giá trị dữ liệu, bất cứ khi nào khung SBR bị sai lệch đã được phát hiện. Các giá trị mang lại đường bao phổ băng cao tĩnh tại mức phát lại tương đối thấp, biểu diễn sự suy giảm về phía tần số cao hơn. Mục tiêu chỉ đơn giản là đảm bảo rằng không có âm thanh có khả năng gây ồn ào, phát ra âm thanh lớn đến tai người nghe, bằng cách chèn "nhiều âm nền nhân tạo" (như tương phản với sự chặn tiếng nghiêm ngặt). Điều này trong thực tế không có sự giảm dần cường độ thực mà hơn là nhảy đến mức năng lượng nhất định để chèn một số loại nhiễu âm nền nhân tạo.

Tiếp theo, sự thay thế được đề cập [EBU12, phần 5.6.3.2] mà tái sử dụng dữ liệu được giải mã một cách chính xác cuối cùng và chậm chậm tăng giảm cường độ các mức (levels - L) về 0, tương tự với trường hợp AAC + SBR.

Bây giờ, các đường thêm nhiễu âm hài hòa và đơn lẻ (Harmonic and Individual Lines plus Noise = HILN) MPEG-4 được xem xét. Meine và cộng sự giới thiệu sự giảm dần cường độ cho bộ mã hóa-giải mã HILN MPEG-4 theo tham số [ISO09] trong miền theo tham số [MEP01]. Đối với các thành phần sóng hài được tiếp tục, động thái mặc định tốt để thay thế các tham số được mã hóa một cách khác nhau bị sai lệch là để giữ tần số không thay đổi, để giảm biên độ bằng hệ số suy giảm (ví dụ, -6dB), và để cho đường bao phổ hội tụ về hướng đó của đặc điểm thông thấp trung bình. Sự thay thế cho đường bao phổ có thể được giữ không đổi. Đối với các biên độ và các đường bao phổ, các thành phần nhiễu âm có thể được xử lý theo cách tương tự như các thành phần sóng hài.

Sau đây, sự tạo vết của mức tạp âm nền trong tình trạng kỹ thuật được xem xét. Rangachari và Loizou [RL06] cung cấp cái nhìn tổng quan tốt của một số phương pháp và thảo luận một số hạn chế của chúng. Các phương pháp tạo vết mức tạp âm nền là, ví dụ, quy trình theo dấu cực tiểu [RL06] [Coh03] [SFB00] [Dob95], trên cơ sở phát hiện hoạt động tiếng (voice activity detection - VAD); sự lọc Kalman [Gan05]

[BJH06], sự phân tích khoảng trống phụ [BP06] [HJH08]; Quyết định mềm [SS98] [MPC89] [HE95], và các thống kê tối thiểu.

Phương pháp thống kê tối thiểu được lựa chọn để được sử dụng trong phạm vi cho Mã hóa âm thanh và tiếng nói hợp nhất 2, (USAC = Unified Speech and Audio Coding) và được liệt kê tiếp theo đây một cách chi tiết.

Sự ước lượng mật độ phổ năng lượng nhiều âm dựa trên sự làm nhấn tối ưu và sự thống kê tối thiểu [Mar01] giới thiệu bộ ước lượng nhiều âm, mà có thể làm việc độc lập với tín hiệu là tiếng nói thực sự hoặc tạp âm nền. Trái ngược với các phương pháp, thuật toán thống kê tối thiểu không sử dụng bất kỳ ngưỡng rõ ràng nào để phân biệt giữa hoạt động tiếng nói và tạm dừng tiếng nói và do đó liên quan chặt chẽ hơn với các phương pháp quyết định mềm so với các phương pháp phát hiện hoạt động tiếng truyền thống. Tương tự các phương pháp quyết định mềm, nó cũng có thể cập nhật mật độ phổ công suất (Power Spectral Density - PSD) trong quá trình hoạt động tiếng.

Phương pháp thống kê tối thiểu phụ thuộc vào hai sự quan sát cụ thể là tiếng nói và nhiều âm mà thường độc lập theo số liệu và công suất của tín hiệu tiếng nói thường suy giảm về mức công suất của nhiều âm. Do đó có thể đạo hàm số đo mật độ phổ công suất (PSD = power spectral density) nhiều âm chính xác bằng cách theo dấu cực tiểu của PSD tín hiệu nhiều. Vì cực tiểu là nhỏ hơn (hoặc trong các trường hợp khác bằng) giá trị trung bình, phương pháp theo dấu cực tiểu yêu cầu sự bù độ lệch.

Độ lệch là hàm của sự biến thiên của PSD tín hiệu được làm nhấn và điều này phụ thuộc vào tham số làm nhấn của bộ ước lượng PSD. Trái ngược với nghiên cứu trước đây về việc theo dấu cực tiểu, mà sử dụng tham số làm nhấn không thay đổi và sự hiệu chỉnh độ lệch tối thiểu không thay đổi, việc làm nhấn PSD phụ thuộc thời gian và tần số được sử dụng, mà cũng yêu cầu sự bù độ lệch phụ thuộc thời gian và tần số.

Việc sử dụng theo dấu cực tiểu cung cấp sự ước lượng sơ bộ về công suất nhiều âm. Tuy nhiên, có một số thiếu sót. Việc làm nhấn với tham số làm nhấn cố định giúp mở rộng các đỉnh của hoạt động tiếng nói của bộ ước lượng PSD được làm nhấn. Điều này sẽ dẫn đến những sự ước lượng nhiều âm không chính xác vì của sổ trượt cho việc tìm kiếm cực tiểu có thể trượt vào trong các đỉnh rộng. Do đó, các tham số làm nhấn gần với 1 không thể được sử dụng, và, kết quả là, số đo nhiều âm sẽ có sự biến thiên

tương đối lớn. Hơn nữa, số đo nhiễu âm được nghiêng về các giá trị thấp hơn. Hơn nữa, trong trường hợp tăng công suất nhiễu âm, việc theo dấu cực tiểu bị tụt lại phía sau.

Sự theo dấu PSD nhiễu âm trên cơ sở MMSE với độ phức tạp thấp [HHJ10] đưa ra phương pháp PSD nhiễu âm nên sử dụng việc tìm kiếm MMSE được sử dụng trên phổ biến đổi Fourier rời rạc (Discrete Fourier Transform - DFT). Thuật toán bao gồm các bước xử lý sau:

- Bộ ước lượng khả năng cực đại được tính toán dựa trên PSD nhiễu âm của khung trước.
- Bộ ước lượng bình phương trung bình cực tiểu được tính toán.
- Bộ ước lượng khả năng cực đại được tính toán sử dụng phương pháp có hướng quyền định [EM84].
- Hệ số nghịch đảo được tính toán thừa nhận rằng các hệ số DFT tiếng nói và nhiễu âm là được phân bố theo kiểu Gaussian.
- Mật độ phổ công suất nhiễu âm được ước lượng được làm nhẵn.

Cũng có phương pháp lưới an toàn được áp dụng để tránh việc dừng hẳn hoàn toàn của thuật toán.

Việc theo dấu của nhiễu âm không ổn định dựa trên ước lượng công suất nhiễu âm đệ quy được điều khiển bằng dữ liệu [EH08] giới thiệu phương pháp ước lượng biến thiên phổ nhiễu âm từ các tín hiệu tiếng nói bị ô nhiễm bởi các nguồn nhiễu âm không ổn định cao. Phương pháp này cũng sử dụng sự làm nhẵn theo hướng thời gian/tần số.

Thuật toán ước lượng nhiễu âm có độ phức tạp thấp dựa trên việc làm nhẵn ước lượng công suất nhiễu âm và hiệu chỉnh độ lệch ước lượng [Yu09] tăng cường cách tiếp cận được giới thiệu trong [EH08]. Sự khác biệt chính là, hàm khuếch đại phổ cho sự ước lượng công suất nhiễu âm được tìm thấy bởi phương pháp điều khiển bằng dữ liệu lặp lại.

Các phương pháp thống kê để tăng cường tiếng nói nhiễu [Mar03] kết hợp

phương pháp thống kê tối thiểu được đưa ra trong [Mar01] bằng cách thay đổi độ khuếch đại quyết định mềm [MCA99], bằng sự ước lượng SNR tiên nghiệm [MCA99], bằng cách giới hạn độ khuếch đại thích ứng [MC99] và bằng bộ ước lượng biên độ phổ bản ghi vết MMSE [EM85].

Sự giảm dần cường độ là mối quan tâm đặc biệt đối với nhiều bộ mã hóa-giải mã tiếng nói và âm thanh, cụ thể, AMR (xem [3GP12b]) (bao gồm ACELP và CNG), AMR-WB (xem [3GP09c]) (bao gồm ACELP và CNG), AMR-WB+ (xem [3GP09a]) (bao gồm ACELP, TCX và CNG), G.718 (xem [ITU08a]), G.719 (xem [ITU08b]), G.722 (xem [ITU07]), G.722.1 (xem [ITU05]), G.729 (xem [ITU12, CPK08, PKJ+11]), MPEG-4 HE-AAC / aacPlus nâng cao (xem [EBU10, EBU12, 3GP12e, LS01, QD03]) (bao gồm AAC và SBR), MPEG-4 HILN (xem [ISO09, MEP01]) và OPUS (xem [IET12]) (bao gồm SILK và CELT).

Phụ thuộc vào bộ mã hóa-giải mã, sự giảm dần cường độ được thực hiện trên các miền khác nhau:

Đối với các bộ mã hóa-giải mã mà sử dụng LPC, sự giảm dần cường độ được thực hiện trong miền dự báo tuyến tính (cũng được biết đến như miền kích thích). Điều này là đúng cho các bộ mã hóa-giải mã mà được dựa trên ACELP, ví dụ, AMR, AMR-WB, lõi ACELP của AMR-WB+, G.718, G.729, G.729.1, lõi SILK trong OPUS; các bộ mã hóa mà còn xử lý tín hiệu kích thích sử dụng sự biến đổi thời gian - tần số, ví dụ, lõi TCX của AMR-WB+, lõi CELT trong OPUS; và đối với các sơ đồ tạo nhiễu âm nền nhân tạo (comfort noise generation - CNG), mà hoạt động trong miền dự báo tuyến tính, ví dụ, CNG trong AMR, CNG trong AMR-WB, CNG trong AMR-WB+.

Đối với các bộ mã hóa-giải mã mà trực tiếp biến đổi tín hiệu thời gian vào miền tần số, sự giảm dần cường độ được thực hiện trong miền phổ/băng phụ. Điều này là đúng cho các bộ mã hóa-giải mã mà dựa trên MDCT hoặc sự biến đổi tương tự, như AAC trong MPEG-4, HE-AAC, G.719, G.722 (miền băng phụ) và G.722.1.

Đối với các bộ mã hóa-giải mã theo tham số, sự giảm dần cường độ được áp dụng trong miền tham số. Điều này là đúng cho HILN MPEG-4.

Đối với tốc độ giảm dần cường độ và đường cong giảm dần cường độ, sự giảm

dân cường độ thường được nhận ra bằng cách áp dụng hệ số suy giảm, mà được áp dụng cho biểu diễn tín hiệu trong miền thích hợp. Kích thước của hệ số suy giảm kiểm soát tốc độ giảm dân cường độ và đường cong giảm dân cường độ. Trong hầu hết các trường hợp, hệ số suy giảm được áp dụng theo khung, nhưng cũng là sự áp dụng theo mẫu được sử dụng, ví dụ, xem G.718 và G.722.

Hệ số suy giảm cho phân đoạn tín hiệu nhất định có thể được cung cấp theo hai cách, tuyệt đối và tương đối.

Trong trường hợp mà hệ số suy giảm được cung cấp một cách tuyệt đối, mức tham chiếu luôn là một trong những khung được nhận cuối cùng. Các hệ số suy giảm tuyệt đối thường bắt đầu với giá trị gần 1 cho phân đoạn tín hiệu ngay sau khung tốt cuối cùng và sau đó giảm nhanh hơn hoặc chậm hơn về phía 0. Đường con giảm dân cường độ phụ thuộc trực tiếp vào các hệ số này. Đây là, ví dụ, trường hợp cho sự che giấu được mô tả trong Phụ lục IV của G.722 (xem, cụ thể, [ITU07, hình IV.7]), trong đó các đường con giảm dân cường độ là tuyến tính hoặc dần dần tuyến tính. Xem xét hệ số độ khuếch đại  $g(n)$ , trong đó  $g(0)$  biểu diễn hệ số độ khuếch đại của khung tốt cuối cùng, hệ số suy giảm tuyệt đối  $\alpha_{abs}(n)$ , hệ số độ khuếch đại của bất kỳ khung bị mất tiếp theo có thể được đạo hàm như sau

$$g(n) = \alpha_{abs}(n) \cdot g(0) \quad (21)$$

Trong trường hợp mà hệ số suy giảm được cung cấp một cách tương đối, mức tham chiếu là một khung từ khung trước đó. Điều này có lợi ích trong trường hợp quy trình che giấu đệ quy, ví dụ, nếu tín hiệu đã được suy giảm còn được xử lý và được suy giảm lần nữa.

Nếu hệ số suy giảm được áp dụng một cách đệ quy, sau đó điều này có thể là giá trị cố định phụ độc lập khỏi số các khung bị mất liên tiếp, ví dụ, 0,5 cho G.719 (xem bên trên); giá trị cố định tương ứng với số lượng các khung bị mất liên tiếp, ví dụ, như được đề xuất cho G.729 trong [CPK08]: 1,0 cho hai khung thứ nhất, 0,9 cho hai khung tiếp theo, 0,8 cho các khung 5 và 6, và 0 cho tất cả các khung tiếp theo (xem bên trên); hoặc giá trị mà tương đối với số lượng các khung bị mất liên tiếp và điều này phụ thuộc vào đặc điểm tín hiệu, ví dụ, sự giảm dân cường độ nhanh hơn cho tín hiệu

không ổn định và sự giảm dần cường độ chậm hơn cho tín hiệu ổn định, ví dụ, G.718 (xem phần trên và [ITU08a, bảng 44]);

Giả sử hệ số giảm dần cường độ tương đối  $0 \leq \alpha_{rel}(n) \leq 1$ , trong đó  $n$  là số lượng khung bị mất ( $n \geq 1$ ); hệ số độ khuếch đại của bất kỳ khung tiếp theo nào có thể được đạo hàm như sau

$$g(n) = \alpha_{rel}(n) \cdot g(n-1) \quad (22)$$

$$g(n) = \left( \prod_{m=1}^n \alpha(m) \right) \cdot g(0) \quad (23)$$

$$g(n) = \alpha_{rel}^n \cdot g(0) \quad (24)$$

dẫn đến sự tăng giảm cường độ hàm mũ.

Đối với quy trình giảm dần cường độ, thông thường, hệ số suy giảm được chỉ rõ, nhưng trong một số tiêu chuẩn áp dụng (DRM, DAB+), hệ số suy giảm sẽ để lại cho nhà sản xuất.

Nếu các phần tín hiệu khác nhau được tăng giảm cường độ riêng biệt, các hệ số suy giảm khác nhau có thể được áp dụng, ví dụ, để tăng giảm cường độ các thành phần như nhiều âm với tốc độ khác (ví dụ, AMR, SILK).

Thông thường, độ khuếch đại nhất định được áp dụng cho toàn bộ khung. Khi sự tăng giảm cường độ được thực hiện trong miền phổ, đây là cách khả thi duy nhất. Tuy nhiên, nếu sự tăng giảm cường độ được thực hiện trong miền thời gian hoặc trong miền dự báo tuyến tính, sự tăng giảm cường độ dạng hạt là khả thi. Nhiều sự tăng giảm cường độ dạng hạt như vậy được áp dụng trong G.718, trong đó các thừa số khuếch đại riêng lẻ được đạo hàm từ mỗi mẫu bằng sự nội suy giữa thừa số khuếch đại của khung cuối cùng và thừa số khuếch đại của khung hiện thời.

Đối với các bộ mã hóa-giải mã với thời gian khung biến thiên, hệ số suy giảm không thay đổi, tương đối dẫn đến tốc độ giảm dần cường độ khác nhau phụ thuộc vào độ dài khung. Điều này là, ví dụ, trường hợp đối với AAC, trong đó độ dài khung phụ thuộc vào tốc độ lấy mẫu.

Để thừa nhận đường cong tăng giảm cường độ được áp dụng cho hình dạng theo thời gian của tín hiệu được nhận cuối cùng, các thừa số giảm dần cường độ (tính) có thể còn được điều chỉnh. Sự điều chỉnh động nữa như vậy, ví dụ, được áp dụng cho AMR trong đó trung bình của năm thừa số khuếch đại trước được xem xét (xem [3GP12b] và phần 1.8.1). Trước khi thực hiện bất kỳ sự suy giảm nào được thực hiện, độ khuếch đại hiện thời được thiết lập về trung bình, nếu trung bình nhỏ hơn độ khuếch đại cuối cùng, nếu không độ khuếch đại cuối cùng được sử dụng. Hơn nữa, sự điều chỉnh động như vậy, ví dụ, được áp dụng cho G.729, trong đó biên độ được dự báo bằng phép hồi quy tuyến tính của các thừa số khuếch đại trước đó (xem [CPK08, PKJ+11] và phần 1.6). Trong trường hợp này, thừa số khuếch đại thu được cho các khung được che giấu thứ nhất có thể vượt qua thừa số khuếch đại của khung được nhận cuối cùng.

Đối với mức đích của sự giảm dần cường độ, với sự loại trừ G.718 và CELT, mức đích là 0 đối với tất cả các bộ mã hóa-giải mã được phân tích, bao gồm sự tạo nhiễu âm nền nhân tạo (comfort noise generation - CNG) của các bộ mã hóa-giải mã đó.

Trong G.718, sự tăng giảm cường độ của sự kích thích bước (biểu diễn các thành phần âm) và sự tăng giảm cường độ của sự kích thích ngẫu nhiên (biểu diễn các thành phần tương tự nhiễu âm) được thực hiện một cách riêng lẻ. Trong khi thừa số khuếch đại bước được tăng giảm cường độ về 0, thừa số khuếch đại đổi mới được tăng giảm cường độ về năng lượng kích thích CNG.

Giả sử các hệ số suy giảm tương đối được cho, điều này dẫn đến - dựa trên công thức (23) - hệ số suy giảm tuyệt đối sau:

$$g(n) = \alpha_{rel}(n) \cdot g(n-1) + (1 - \alpha_{rel}(n)) \cdot g_n \quad (25)$$

với  $g_n$  là độ khuếch đại của sự kích thích được sử dụng trong quá trình tạo nhiễu âm nền nhân tạo. Công thức này tương đương với công thức (23), khi  $g_n = 0$ .

G.718 không thực hiện sự giảm dần cường độ nào trong trường hợp DTX/CNG.

Trong CELT không có sự tăng giảm cường độ nào về mức đích, nhưng sau 5 khung che giấu âm (bao gồm cả sự giảm dần cường độ) mức được chuyển ngay đến

mức đích tại khung bị mất liên tiếp thứ 6. Mức được đạo hàm theo băng sử dụng công thức (19).

Đối với hình dạng phổ đích của sự giảm dần cường độ, tất cả các bộ mã hóa biến đổi dựa trên sự biến đổi đơn thuần được tổng hợp (AAC, G.719, G.722, G.722.1) cũng như SBR đơn giản kéo dài hình dạng phổ của khung tốt cuối cùng trong quá trình giảm dần cường độ.

Nhiều bộ mã hóa-giải mã tiếng nói khác nhau tăng giảm cường độ hình dạng phổ thành trung bình bằng cách sử dụng tổng hợp LPC. Giá trị trung bình có thể là tĩnh (AMR) hoặc thích ứng (AMR-WB, AMR-WB+, G.718), trong đó số trung bình thích ứng được đạo hàm từ số trung bình tĩnh và số trung bình ngắn hạn (được lấy bằng cách tính trung bình  $n$  tập hợp hệ số dự báo tuyến tính (Linear Prediction - LP) cuối cùng).

Tất cả các môđun CNG trong các bộ mã hóa-giải mã được thảo luận AMR, AMR-WB, AMR-WB+, G.718 kéo dài hình dạng phổ của khung tốt cuối cùng trong quá trình giảm dần cường độ.

Đối với sự tạo vết mức tạp âm nền, có năm phương pháp khác nhau được biết đến từ các tài liệu:

- Trên cơ sở phát hiện hoạt động tiếng: dựa trên SNR/VAD, nhưng rất khó để tinh chỉnh và khó để sử dụng cho tiếng nói SNR thấp.

- Sơ đồ quyết định mềm: Phương pháp sơ đồ quyết định mềm xem xét xác suất của sự có mặt tiếng nói [SS98] [MPC89] [HE95].

- Thống kê cực tiểu: Cực tiểu của PSD được theo dấu giữ lượng các giá trị nhất định theo thời gian trong vòng đệm, do đó cho phép tìm ra nhiều âm cực tiểu từ các mẫu quá khứ [Mar01] [HHJ10] [EH08] [Yu09].

- Lọc Kalman: Thuật toán sử dụng chuỗi các phép đo được quan sát trong suốt thời gian, chứa nhiều âm (sự biến thiên ngẫu nhiên), và tạo ra các ước lượng về PSD nhiều âm mà có xu hướng chính xác hơn so với các ước lượng dựa trên chỉ một phép đo đơn. Bộ lọc Kalman hoạt động theo cách đệ quy trên các dòng dữ liệu đầu vào nhiều âm để tạo ra ước lượng tối ưu theo thống kê về trạng thái hệ thống [Gan05] [BJH06].

- Sự phân tích không gian phụ: Phương pháp này cố phân tích tín hiệu như nhiều âm thành tín hiệu tiếng nói sạch và phần nhiễu âm, ví dụ sử dụng sự biến đổi Karhunen-Loève (Karhunen-Loève transform - KLT) (còn được gọi là sự phân tích thành phần chính) và/hoặc phép biến đổi Fourier thời gian rời rạc (Discrete Time Fourier Transform - DFT). Sau đó, các vectơ riêng/ các giá trị riêng có thể được tạo vệt sử dụng thuật toán làm nhẵn bất kỳ [BP06] [HJH08].

*Tài liệu tham khảo*

[3GP09a] 3GPP; Technical Specification Group Services and System Aspects, *Extended adaptive multi-rate - wideband (AMR-WB+) codec*, 3GPP TS 26.290, 3rd Generation Partnership Project, 2009.

[3GP09b] *Extended adaptive multi-rate - wideband (AMR-WB+) codec; floating-point ANSI-C code*, 3GPP TS 26.304, 3rd Generation Partnership Project, 2009.

[3GP09c] *Speech codec speech processing functions; adaptive multi-rate - wideband (AMRWB) speech codec; transcoding functions*, 3GPP TS 26.190, 3rd Generation Partnership Project, 2009.

[3GP12a] *Adaptive multi-rate (AMR) speech codec; error concealment of lost frames (release 11)*, 3GPP TS 26.091, 3rd Generation Partnership Project, Sep 2012.

[3GP12b] *Adaptive multi-rate (AMR) speech codec; transcoding functions (release 11)*, 3GPP TS 26.090, 3rd Generation Partnership Project, Sep 2012.

[3GP12c] , *ANSI-C code for the adaptive multi-rate - wideband (AMR-WB) speech codec*, 3GPP TS 26.173, 3rd Generation Partnership Project, Sep 2012.

[3GP12d] *ANSI-C code for the floating-point adaptive multi-rate (AMR) speech codec (release11)*, 3GPP TS 26.104, 3rd Generation Partnership Project, Sep 2012.

[3GP12e] *General audio codec audio processing functions; Enhanced aacPlus general audio codec; additional decoder tools (release 11)*, 3GPP TS 26.402, 3rd Generation Partnership Project, Sep 2012.

[3GP12f] *Speech codec speech processing functions; adaptive multi-rate - wideband (amr-wb) speech codec; ansi-c code*, 3GPP TS 26.204, 3rd Generation Partnership Project, 2012.

[3GP12g] *Speech codec speech processing functions; adaptive multi-rate - wideband (AMR-WB) speech codec; error concealment of erroneous or lost frames*, 3GPP TS 26.191, 3rd Generation Partnership Project, Sep 2012.

[BJH06] I. Batina, J. Jensen, and R. Heusdens, *Noise power spectrum estimation for speech enhancement using an autoregressive model for speech power spectrum dynamics*, in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. 3 (2006), 1064–1067.

[BP06] A. Borowicz and A. Petrovsky, *Minima controlled noise estimation for klt-based speech enhancement*, CD-ROM, 2006, Italy, Florence.

[Coh03] I. Cohen, *Noise spectrum estimation in adverse environments: Improved minima controlled recursive averaging*, IEEE Trans. Speech Audio Process. 11 (2003), no. 5, 466–475.

[CPK08] Choong Sang Cho, Nam In Park, and Hong Kook Kim, *A packet loss concealment algorithm robust to burst packet loss for celp- type speech coders*, Tech. report, Korea Electronics Technology Institute, Gwang Institute of Science and Technology, 2008, The 23rd International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC 2008).

[Dob95] G. Doblinger, *Computationally efficient speech enhancement by spectral minima tracking in subbands*, in Proc. Eurospeech (1995), 1513–1516.

[EBU10] EBU/ETSI JTC Broadcast, *Digital audio broadcasting (DAB); transport of advanced audio coding (AAC) audio*, ETSI TS 102 563, European Broadcasting Union, May 2010.

[EBU12] *Digital radio mondiale (DRM); system specification*, ETSI ES 201 980, ETSI, Jun 2012.

[EH08] Jan S. Erkelens and Richards Heusdens, *Tracking of Nonstationary Noise Based on Data-Driven Recursive Noise Power Estimation*,

Audio, Speech, and Language Processing, IEEE Transactions on 16 (2008), no. 6, 1112–1123.

[EM84] Y. Ephraim and D. Malah, *Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator*, IEEE Trans. Acoustics, Speech and Signal Processing 32 (1984), no. 6, 1109–1121.

[EM85] *Speech enhancement using a minimum mean-square error log-spectral amplitude estimator*, IEEE Trans. Acoustics, Speech and Signal Processing 33 (1985), 443–445.

[Gan05] S. Gannot, *Speech enhancement: Application of the kalman filter in the estimate-maximize (em framework)*, Springer, 2005.

[HE95] H. G. Hirsch and C. Ehrlicher, *Noise estimation techniques for robust speech recognition*, Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing, no. pp. 153-156, IEEE, 1995.

[HHJ10] Richard C. Hendriks, Richard Heusdens, and Jesper Jensen, *MMSE based noise PSD tracking with low complexity*, Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on, Mar 2010, pp. 4266 – 4269.

[HJH08] Richard C. Hendriks, Jesper Jensen, and Richard Heusdens, *Noise tracking using dft domain subspace decompositions*, IEEE Trans. Audio, Speech, Lang. Process. 16 (2008), no. 3, 541–553.

[IET12] IETF, *Definition of the Opus Audio Codec*, Tech. Report RFC 6716, Internet Engineering Task Force, Sep 2012.

[ISO09] ISO/IEC JTC1/SC29/WG11, *Information technology – coding of audio-visual objects – part 3: Audio*, ISO/IEC IS 14496-3, International Organization for Standardization, 2009.

[ITU03] ITU-T, *Wideband coding of speech at around 16 kbit/s using adaptive multi-rate wideband (amr-wb)*, Recommendation ITU-T G.722.2, Telecommunication Standardization Sector of ITU, Jul 2003.

[ITU05] *Low-complexity coding at 24 and 32 kbit/s for hands-free operation in systems with low frame loss*, Recommendation ITU-T G.722.1, Telecommunication Standardization Sector of ITU, May 2005.

[ITU06a] *G.722 Appendix III: A high-complexity algorithm for packet loss concealment for G.722*, ITU-T Recommendation, ITU-T, Nov 2006.

[ITU06b] *G.729.1: G.729-based embedded variable bit-rate coder: An 8-32 kbit/s scalable wideband coder bitstream interoperable with g.729*, Recommendation ITU-T G.729.1, Telecommunication Standardization Sector of ITU, May 2006.

[ITU07] *G.722 Appendix IV: A low-complexity algorithm for packet loss concealment with G.722*, ITU-T Recommendation, ITU-T, Aug 2007.

[ITU08a] *G.718: Frame error robust narrow-band and wideband embedded variable bit-rate coding of speech and audio from 8-32 kbit/s*, Recommendation ITU-T G.718, Telecommunication Standardization Sector of ITU, Jun 2008.

[ITU08b] *G.719: Low-complexity, full-band audio coding for high-quality, conversational applications*, Recommendation ITU-T G.719, Telecommunication Standardization Sector of ITU, Jun 2008.

[ITU12] *G.729: Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear prediction (cs-acelp)*, Recommendation ITU-T G.729, Telecommunication Standardization Sector of ITU, June 2012.

[LS01] Pierre Lauber and Ralph Sperschneider, *Error concealment for compressed digital audio*, Audio Engineering Society Convention 111, no. 5460, Sep 2001.

[Mar01] Rainer Martin, *Noise power spectral density estimation based on optimal smoothing and minimum statistics*, IEEE Transactions on Speech and Audio Processing 9 (2001), no. 5, 504–512.

[Mar03] *Statistical methods for the enhancement of noisy speech*, International Workshop on Acoustic Echo and Noise Control (IWAENC2003), Technical University of Braunschweig, Sep 2003.

[MC99] R. Martin and R. Cox, *New speech enhancement techniques for low bit rate speech coding*, in Proc. IEEE Workshop on Speech Coding (1999), 165–167.

[MCA99] D. Malah, R. V. Cox, and A. J. Accardi, *Tracking speech-presence uncertainty to improve speech enhancement in nonstationary noise environments*, Proc. IEEE Int. Conf. on Acoustics Speech and Signal Processing (1999), 789–792.

[MEP01] Nikolaus Meine, Bernd Edler, and Heiko Purnhagen, *Error protection and concealment for HILN MPEG-4 parametric audio coding*, Audio Engineering Society Convention 110, no. 5300, May 2001.

[MPC89] Y. Mahieux, J.-P. Petit, and A. Charbonnier, *Transform coding of audio signals using correlation between successive transform blocks*, Acoustics, Speech, and Signal Processing, 1989. ICASSP-89., 1989 International Conference on, 1989, pp. 2021–2024 vol.3.

[NMR+12] Max Neuendorf, Markus Multrus, Nikolaus Rettelbach, Guillaume Fuchs, Julien Robilliard, Jérémie Lecomte, Stephan Wilde, Stefan Bayer, Sascha Disch, Christian Helmrich, Roch Lefebvre, Philippe Gournay, Bruno Bessette, Jimmy Lapierre, Kristopfer Kjörling, Heiko Purnhagen, Lars Villemoes, Werner Oomen, Erik Schuijers, Kei Kikuri, Toru Chinen, Takeshi Norimatsu, Chong Kok Seng, Eunmi Oh, Miyoung Kim, Schuyler Quackenbush, and Berndhard Grill, *MPEG Unified Speech and Audio Coding - The ISO / MPEG Standard for High-Efficiency Audio Coding of all Content Types*, Convention Paper 8654, AES, April 2012, Presented at the 132nd Convention Budapest, Hungary.

[PKJ+11] Nam In Park, Hong Kook Kim, Min A Jung, Seong Ro Lee, and Seung Ho Choi, *Burst packet loss concealment using multiple codebooks and comfort noise for celp-type speech coders in wireless sensor networks*, Sensors 11 (2011), 5323–5336.

[QD03] Schuyler Quackenbush and Peter F. Driessen, *Error mitigation in MPEG-4 audio packet communication systems*, Audio Engineering Society

Convention 115, no. 5981, Oct 2003.

[RL06] S. Rangachari and P. C. Loizou, *A noise-estimation algorithm for highly non-stationary environments*, Speech Commun. 48 (2006), 220–231.

[SFB00] V. Stahl, A. Fischer, and R. Bippus, *Quantile based noise estimation for spectral subtraction and wiener filtering*, in Proc. IEEE Int. Conf. Acoust., Speech and Signal Process. (2000), 1875–1878.

[SS98] J. Sohn and W. Sung, *A voice activity detector employing soft decision based noise spectrum adaptation*, Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing, no. pp. 365-368, IEEE, 1998.

[Yu09] Rongshan Yu, *A low-complexity noise estimation algorithm based on smoothing of noise power estimation and estimation bias correction*, Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on, Apr 2009, pp. 4421–4424.

### **Bản chất kỹ thuật của sáng chế**

Mục đích của sáng chế là đề xuất các khái niệm được cải thiện cho các hệ thống mã hóa âm thanh. Mục đích của sáng chế được giải quyết bởi thiết bị theo điểm 1, bởi phương pháp mã hóa theo điểm 12 và bởi chương trình máy tính theo điểm 13.

Thiết bị giải mã tín hiệu âm thanh được mã hóa để thu được tín hiệu âm thanh được khôi phục được đề xuất. Thiết bị bao gồm giao diện nhận để nhận nhiều khung, vùng đệm trễ để lưu trữ các mẫu tín hiệu âm thanh của tín hiệu âm thanh được giải mã, bộ chọn lọc mẫu để chọn lọc nhiều mẫu tín hiệu âm thanh được lựa chọn từ các mẫu tín hiệu âm thanh được lưu trữ trong vùng đệm trễ, và bộ xử lý mẫu để xử lý các mẫu tín hiệu âm thanh được lựa chọn để thu được các mẫu tín hiệu âm thanh được khôi phục của tín hiệu âm thanh được khôi phục. Bộ lựa chọn mẫu được tạo cấu hình để lựa chọn, nếu khung hiện thời được nhận bởi giao diện nhận và nếu khung hiện thời được nhận bởi giao diện nhận không bị sai lệch, nhiều mẫu tín hiệu âm thanh được lựa chọn từ các mẫu tín hiệu âm thanh được lưu trữ trong vùng đệm trễ phụ thuộc vào thông tin độ trễ được bao gồm bởi khung hiện thời. Hơn nữa, bộ lựa chọn mẫu được tạo cấu hình để lựa chọn, nếu khung hiện thời không được nhận bởi giao diện nhận hoặc nếu

khung hiện thời được nhận bởi giao diện nhận bị sai lạc, nhiều mẫu tín hiệu âm thanh được lựa chọn, từ các mẫu tín hiệu âm thanh được lưu trữ trong vùng đệm trễ phụ thuộc vào thông tin độ trễ bước được bao gồm trong khung khác được nhận trước đó bởi giao diện nhận.

Theo phương án, bộ xử lý mẫu có thể, ví dụ, được tạo cấu hình để thu được các mẫu tín hiệu âm thanh được khôi phục, nếu khung hiện thời được nhận bởi giao diện nhận và nếu khung hiện thời được nhận bởi giao diện nhận không bị sai lạc, bằng cách định tỉ lệ lại các mẫu tín hiệu âm thanh được lựa chọn phụ thuộc vào thông tin độ khuếch đại được bao gồm bởi khung hiện thời. Hơn nữa, bộ lựa chọn mẫu có thể, ví dụ, được tạo cấu hình để thu được các mẫu tín hiệu âm thanh được khôi phục, nếu khung hiện thời không được nhận bởi giao diện nhận hoặc nếu khung hiện thời được nhận bởi giao diện nhận bị sai lạc, bằng cách định tỉ lệ lại các mẫu tín hiệu âm thanh được lựa chọn phụ thuộc vào thông tin độ khuếch đại được bao gồm bởi khung khác đã nêu được nhận trước đó bởi giao diện nhận.

Trong phương án, bộ xử lý mẫu có thể, ví dụ, được tạo cấu hình để thu được các mẫu tín hiệu âm thanh được khôi phục, nếu khung hiện thời được nhận bởi giao diện nhận và nếu khung hiện thời được nhận bởi giao diện nhận không bị sai lạc, bằng cách nhân các mẫu tín hiệu âm thanh được lựa chọn và giá trị phụ thuộc vào thông tin độ khuếch đại được bao gồm bởi khung hiện thời. Hơn nữa, bộ lựa chọn mẫu được tạo cấu hình để thu được các mẫu tín hiệu âm thanh được khôi phục, nếu khung hiện thời không được nhận bởi giao diện nhận hoặc nếu khung hiện thời được nhận bởi giao diện nhận bị sai lạc, bằng cách nhân các mẫu tín hiệu âm thanh được lựa chọn phụ thuộc vào thông tin độ khuếch đại được bao gồm bởi khung khác đã nêu được nhận trước đó bởi giao diện nhận.

Theo phương án, bộ xử lý mẫu có thể, ví dụ, được tạo cấu hình để lưu trữ các mẫu tín hiệu âm thanh được khôi phục vào trong vùng đệm độ trễ.

Trong phương án, bộ xử lý mẫu có thể, ví dụ, được tạo cấu hình để lưu trữ các mẫu tín hiệu âm thanh được khôi phục vào trong vùng đệm độ trễ trước khi khung tiếp theo được nhận bởi giao diện nhận.

Theo phương án, bộ xử lý mẫu có thể, ví dụ, được tạo cấu hình để lưu trữ các

mẫu tín hiệu âm thanh được khôi phục vào trong vùng đệm độ trễ sau khi khung tiếp theo được nhận bởi giao diện nhận.

Trong phương án, bộ xử lý mẫu có thể, ví dụ, được tạo cấu hình để định tỉ lệ lại các mẫu tín hiệu âm thanh được lựa chọn phụ thuộc vào thông tin độ khuếch đại để thu được các mẫu tín hiệu âm thanh được định tỉ lệ lại và bằng cách tổ hợp các mẫu tín hiệu âm thanh được định tỉ lệ với các mẫu tín hiệu âm thanh đầu vào để thu được các mẫu tín hiệu âm thanh được xử lý.

Theo phương án, bộ xử lý mẫu có thể, ví dụ, được tạo cấu hình để lưu trữ các mẫu tín hiệu âm thanh được xử lý, biểu thị sự tổ hợp của các mẫu tín hiệu âm thanh được định tỉ lệ lại và các mẫu tín hiệu âm thanh đầu vào vào trong vùng đệm độ trễ, và không lưu trữ các mẫu tín hiệu âm thanh được định tỉ lệ vào vùng đệm trễ, nếu khung hiện thời được nhận bởi giao diện nhận và nếu khung hiện thời được nhận bởi giao diện nhận không bị sai lệch. Hơn nữa, bộ xử lý mẫu được tạo cấu hình để lưu trữ các mẫu tín hiệu âm thanh được định tỉ lệ vào trong vùng đệm trễ và không lưu trữ các mẫu tín hiệu âm thanh vào vùng đệm trễ, nếu khung hiện thời không được nhận bởi giao diện nhận hoặc nếu khung hiện thời được nhận bởi giao diện nhận bị sai lệch.

Theo phương án khác, bộ xử lý mẫu có thể, ví dụ, được tạo cấu hình để lưu trữ các mẫu tín hiệu âm thanh được xử lý vào trong vùng đệm trễ, nếu khung hiện thời không được nhận bởi giao diện nhận hoặc nếu khung hiện thời được nhận bởi giao diện nhận bị sai lệch.

Trong phương án, bộ lựa chọn mẫu có thể, ví dụ, được tạo cấu hình để thu được các mẫu tín hiệu âm thanh được khôi phục bằng cách định tỉ lệ lại các mẫu tín hiệu âm thanh được lựa chọn phụ thuộc vào độ khuếch đại được thay đổi, trong đó độ khuếch đại được thay đổi được xác định theo công thức:

$$\text{gain} = \text{gain\_past} * \text{damping};$$

trong đó  $\text{gain}$  là độ khuếch đại được thay đổi, trong đó bộ lựa chọn mẫu có thể, ví dụ, được tạo cấu hình để thiết lập  $\text{gain\_past}$  thành  $\text{gain}$  sau  $\text{gain}$  và đã được tính toán, và trong đó  $\text{damping}$  là giá trị thực.

Theo phương án, bộ lựa chọn mẫu có thể, ví dụ, được tạo cấu hình để tính toán

độ khuếch đại được thay đổi.

Trong phương án, damping có thể, ví dụ, được xác định theo:  $0 \leq \text{damping} \leq 1$ .

Theo phương án, độ khuếch đại được thay đổi gain có thể, ví dụ, được thiết lập về 0, nếu ít nhất số các khung được định trước chưa được nhận bởi giao diện nhận vì khung cuối cùng đã được nhận bởi giao diện nhận.

Hơn nữa, phương pháp giải mã tín hiệu âm thanh được mã hóa để thu được tín hiệu âm thanh được khôi phục được đề xuất. Phương pháp gồm có:

- Nhận nhiều khung.
- Lưu trữ các mẫu tín hiệu âm thanh của tín hiệu âm thanh được giải mã.
- Lựa chọn nhiều mẫu tín hiệu âm thanh từ các mẫu tín hiệu âm thanh được lưu trữ trong vùng đệm trễ. Và:
- Xử lý các mẫu tín hiệu âm thanh được lựa chọn để thu được các mẫu tín hiệu âm thanh được khôi phục của tín hiệu âm thanh được khôi phục.

Nếu khung hiện thời được nhận và nếu khung hiện thời được nhận không bị sai lạc, bước lựa chọn nhiều mẫu tín hiệu âm thanh được lựa chọn từ các mẫu tín hiệu âm thanh được lưu trữ trong vùng đệm trễ được thực hiện phụ thuộc vào thông tin độ trễ bước được bao gồm bởi khung hiện thời. Hơn nữa, nếu khung hiện thời không được nhận hoặc nếu khung hiện thời được nhận bị sai lạc, bước lựa chọn nhiều mẫu tín hiệu âm thanh được lựa chọn từ các mẫu tín hiệu âm thanh được lưu trữ trên vùng đệm trễ được thực hiện phụ thuộc vào thông tin độ trễ bước được bao gồm bởi khung khác được nhận trước đó bởi giao diện nhận.

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

Các phương án sử dụng sự dự báo dài hạn kích thích được mã hóa biến đổi (Transform Coded Excitation Long-Term Prediction - TCX LTP). Trong suốt thao tác thông thường, bộ nhớ TCX LTP được cập nhật với tín hiệu được tổng hợp chứa nhiều âm và các thành phần âm được khôi phục.

Thay vì vô hiệu hóa TCX LTP trong quá trình che giấu, thao tác thông thường có thể được tiếp tục trong quá trình che giấu với các tham số được nhận trong khung tốt cuối cùng. Điều này bảo toàn hình dạng phổ của tín hiệu, cụ thể các thành phần âm này mà được tạo mẫu bởi bộ lọc LTP.

Hơn nữa, các phương án khử ghép vòng phản hồi TCX LTP. Sự tiếp tục đơn giản của thao tác TCX LTP thông thường giới thiệu nhiều âm bổ sung, vì với mỗi bước cập nhật còn được tạo ra một cách ngẫu nhiên nhiều âm từ sự kích thích LTP được dẫn. Các thành phần âm do đó dần bị méo nhiều hơn theo thời gian bởi nhiều âm được thêm vào.

Để vượt qua điều này, chỉ vùng đệm TCX LTP được cập nhật có thể được phản hồi (mà không thêm nhiều âm), để không làm ô nhiễm thông tin âm với nhiều âm ngẫu nhiên không mong muốn.

Hơn nữa, theo các phương án, độ khuếch đại TCX LTP được tăng giảm cường độ về 0.

Các phương án này dựa trên việc tìm ra rằng sự tiếp tục TCX LTP giúp bảo toàn các đặc điểm tín hiệu trong thời hạn ngắn, nhưng có bất lợi trong thời hạn dài: Tín hiệu được phát hết trong quá trình che giấu sẽ bao gồm thông tin hữu thanh/âm có mặt trước khi mất. Đặc biệt đối với tiếng nói sạch hoặc tiếng nói trên tạp âm nền, nó đặc biệt không có khả năng rằng âm hoặc sóng hài sẽ suy tàn rất chậm suốt thời gian rất dài. Bằng cách tiếp tục thao tác TCX LTP trong quá trình che giấu, cụ thể nếu sự cập nhật bộ nhớ LTP được khử ghép (chỉ các thành phần âm được phản hồi và không có phần được xáo trộn ký hiệu), thông tin hữu thanh/âm sẽ xuất hiện trong tín hiệu được che giấu trong quá trình mất, được suy giảm chỉ bằng cách giảm dần cường độ toàn bộ về mức nhiễu âm nền nhân tạo. Hơn nữa, không thể đạt được đường bao nhiễu âm nền nhân tạo trong quá trình mất gói khối tín hiệu truyền, nếu TCX LTP được áp dụng trong quá trình mất khối tín hiệu truyền mà không bị suy giảm theo thời gian, vì tín hiệu sau đó sẽ luôn kết hợp với thông tin của LTP.

Do đó, độ khuếch đại TCX LTP được tăng giảm cường độ về 0, sao cho các thành phần âm được biểu diễn bởi LTP sẽ được tăng giảm cường độ về 0, cùng lúc đó tín hiệu được tăng giảm cường độ về mức và hình dạng tín hiệu nền, và sao cho sự

giảm dần cường độ đạt đến đường bao nền phỏ (nhiều âm nền nhân tạo) mà không kết hợp với các thành phần âm không mong muốn.

Trong các phương án, tốc độ tăng giảm cường độ giống nhau được sử dụng cho sự tăng giảm cường độ độ khuếch đại LTP như cho sự tăng giảm cường độ nhiễu âm trắng.

Ngược lại, trong tình trạng kỹ thuật, không có bộ mã hóa-giải mã biến đổi được biết đến mà sử dụng LTP trong quá trình che giấu. Đối với MPEG-4 LTP [ISO09] không có phương pháp che giấu nào tồn tại trong tình trạng kỹ thuật. Bộ mã hóa-giải mã trên cơ sở MDCT khác của tình trạng kỹ thuật mà tận dụng LTP là CELT, nhưng bộ mã hóa-giải mã này sử dụng sự che giấu như ACELP cho năm khung thứ nhất, và đối với các khung tiếp theo tạp âm nền được tạo ra, mà không sử dụng LTP. Bất lợi của tình trạng kỹ thuật không sử dụng TCX LTP là, tất cả các thành phần âm được tạo mẫu với LTP biến mất một cách đột ngột. Hơn nữa, trong các bộ mã hóa-giải mã trên cơ sở ACELP của tình trạng kỹ thuật, thao tác LTP được kéo dài trong quá trình che giấu, và độ khuếch đại của bảng mã thích ứng được tăng giảm cường độ về 0. Đối với thao tác vòng phản hồi, tình trạng kỹ thuật sử dụng hai phương pháp, hoặc toàn bộ sự kích thích, ví dụ, tổng của sự kích thích đổi mới và thích ứng, được nạp trở lại (AMR-WB); hoặc chỉ có sự kích thích thích ứng được cập nhật, ví dụ, các phần tín hiệu âm, được phản hồi (G.718). Các phương án nêu trên khắc phục các nhược điểm của tình trạng kỹ thuật.

Hơn nữa, thiết bị giải mã tín hiệu âm thanh được đề xuất.

Thiết bị bao gồm giao diện nhận. Giao diện nhận được tạo cấu hình để nhận nhiều khung, trong đó giao diện nhận được tạo cấu hình để nhận khung thứ nhất trong số nhiều khung, khung thứ nhất đã nêu bao gồm phần tín hiệu âm thanh thứ nhất, phần tín hiệu âm thanh thứ nhất đã nêu được biểu diễn trong miền thứ nhất, và trong đó giao diện nhận được tạo cấu hình để nhận khung thứ hai trong số nhiều khung, khung thứ hai đã nêu bao gồm phần tín hiệu âm thanh thứ hai của tín hiệu âm thanh.

Ngoài ra, thiết bị bao gồm bộ phận biến đổi để biến đổi phần tín hiệu âm thanh thứ hai hoặc giá trị hoặc tín hiệu được suy ra từ phần tín hiệu âm thanh thứ hai từ miền thứ hai đến miền tạo vết để thu được thông tin phần tín hiệu thứ hai, trong đó miền thứ

hai khác với miền thứ nhất, trong đó miền tạo vết khác với miền thứ hai, và trong đó miền tạo vết có thể bằng hoặc khác với miền thứ nhất.

Hơn nữa, thiết bị bao gồm bộ phận tạo vết mức nhiều âm, trong đó bộ phận tạo vết mức nhiều âm được tạo cấu hình để nhận thông tin phân tín hiệu thứ nhất được biểu diễn trong miền tạo vết, trong đó thông tin phân tín hiệu thứ nhất phụ thuộc vào phân tín hiệu âm thanh thứ nhất. Bộ phận tạo vết mức nhiều âm được tạo cấu hình để nhận phân tín hiệu thứ hai được biểu diễn trong miền tạo vết, và trong đó bộ phận tạo vết mức nhiều âm được tạo cấu hình để xác định thông tin mức nhiều âm phụ thuộc vào thông tin phân tín hiệu thứ nhất được biểu diễn trong miền tạo vết và phụ thuộc vào thông tin phân tín hiệu thứ hai được biểu diễn trong miền tạo vết.

Ngoài ra, thiết bị bao gồm bộ phận khôi phục để khôi phục phân tín hiệu âm thanh thứ ba của tín hiệu âm thanh phụ thuộc vào thông tin mức nhiều âm, nếu khung thứ ba trong số nhiều khung không được nhận bằng giao diện nhận nhưng bị sai lạc.

Tín hiệu âm thanh có thể, ví dụ, là tín hiệu tiếng nói, hoặc tín hiệu âm nhạc, hoặc tín hiệu mà bao gồm tiếng nói và âm nhạc, v.v.

Sự trình bày rằng thông tin phân tín hiệu thứ nhất phụ thuộc vào phân tín hiệu âm thanh thứ nhất có nghĩa là thông tin phân tín hiệu thứ nhất hoặc là phân tín hiệu âm thanh thứ nhất, hoặc là thông tin phân tín hiệu thứ nhất đã thu được/được tạo ra phụ thuộc vào phân tín hiệu âm thanh thứ nhất hoặc bằng một số cách khác phụ thuộc vào phân tín hiệu âm thanh thứ nhất. Ví dụ, phân tín hiệu âm thanh thứ nhất có thể đã được biến đổi từ một miền đến miền khác để thu được thông tin phân tín hiệu thứ nhất.

Tương tự, sự trình bày rằng thông tin phân tín hiệu thứ hai phụ thuộc vào phân tín hiệu âm thanh thứ hai có nghĩa là thông tin phân tín hiệu thứ hai hoặc là phân tín hiệu âm thanh thứ hai, hoặc là thông tin phân tín hiệu thứ hai đã thu được/được tạo ra phụ thuộc vào phân tín hiệu âm thanh thứ hai hoặc bằng một số cách khác phụ thuộc vào phân tín hiệu âm thanh thứ hai. Ví dụ, phân tín hiệu âm thanh thứ hai có thể đã được biến đổi từ một miền đến miền khác để thu được thông tin phân tín hiệu thứ hai.

Trong phương án, phân tín hiệu âm thanh thứ nhất có thể, ví dụ, được biểu diễn trong miền thời gian như miền thứ nhất. Hơn nữa, bộ phận biến đổi có thể, ví dụ, được

tạo cấu hình để biến đổi phần tín hiệu âm thanh thứ hai hoặc giá trị được suy ra từ phần tín hiệu âm thanh thứ hai từ miền kích thích là miền thứ hai đến miền thời gian là miền tạo vết. Hơn nữa, bộ phận tạo vết mức nhiều âm có thể, ví dụ, được tạo cấu hình để nhận thông tin phần tín hiệu thứ nhất được biểu diễn trong miền thời gian như miền tạo vết. Ngoài ra, bộ phận tạo vết mức nhiều âm có thể, ví dụ, được tạo cấu hình để nhận phần tín hiệu thứ hai được biểu diễn trong miền thời gian như miền tạo vết.

Theo phương án, phần tín hiệu âm thanh thứ nhất có thể, ví dụ, được biểu diễn trong miền kích thích như miền thứ nhất. Hơn nữa, bộ phận biến đổi có thể, ví dụ, được tạo cấu hình để biến đổi phần tín hiệu âm thanh thứ hai hoặc giá trị được suy ra từ phần tín hiệu âm thanh thứ hai từ miền thời gian là miền thứ hai đến miền kích thích là miền tạo vết. Hơn nữa, bộ phận tạo vết mức nhiều âm có thể, ví dụ, được tạo cấu hình để nhận thông tin phần tín hiệu thứ nhất được biểu diễn trong miền kích thích như miền tạo vết. Ngoài ra, bộ phận tạo vết mức nhiều âm có thể, ví dụ, được tạo cấu hình để nhận phần tín hiệu thứ hai được biểu diễn trong miền kích thích như miền tạo vết.

Trong phương án, phần tín hiệu âm thanh thứ nhất có thể, ví dụ, được biểu diễn trong miền kích thích như miền thứ nhất, trong đó bộ phận tạo vết mức nhiều âm có thể, ví dụ, được tạo cấu hình để nhận thông tin phần tín hiệu thứ nhất trong đó thông tin phần tín hiệu thứ nhất được biểu diễn trong miền FFT, là miền tạo vết, và trong đó thông tin phần tín hiệu thứ nhất phụ thuộc vào phần tín hiệu âm thanh thứ nhất đã nêu được biểu diễn trong miền kích thích, trong đó bộ phận biến đổi có thể, ví dụ, được tạo cấu hình để biến đổi phần tín hiệu âm thanh thứ hai hoặc giá trị được suy ra từ phần tín hiệu âm thanh thứ hai từ miền thời gian là miền thời gian thứ hai đến miền FFT là miền tạo vết, và trong đó bộ phận tạo vết mức nhiều âm có thể, ví dụ, được tạo cấu hình để nhận phần tín hiệu âm thanh thứ hai được biểu diễn trong miền FFT.

Trong phương án, thiết bị có thể, ví dụ, còn bao gồm bộ phận gộp thứ nhất để xác định giá trị được gộp thứ nhất phụ thuộc vào phần tín hiệu âm thanh thứ nhất. Hơn nữa, thiết bị có thể, ví dụ, còn bao gồm bộ phận gộp thứ hai để xác định, phụ thuộc vào phần tín hiệu âm thanh thứ hai, giá trị được gộp thứ hai như giá trị được suy ra từ

phần tín hiệu âm thanh thứ hai. Hơn nữa, bộ phận tạo vết mức nhiều âm có thể, ví dụ, được tạo cấu hình để nhận giá trị được gộp thứ nhất như thông tin phần tín hiệu thứ nhất được biểu diễn trong miền tạo vết, trong đó bộ phận tạo vết mức nhiều âm có thể, ví dụ, được tạo cấu hình để nhận giá trị được gộp thứ hai như thông tin phần tín hiệu thứ hai được biểu diễn trong miền tạo vết, và trong đó bộ phận tạo vết mức nhiều âm có thể, ví dụ, được tạo cấu hình để xác định thông tin mức nhiều âm phụ thuộc vào giá trị được gộp thứ nhất được biểu diễn trong miền tạo vết và phụ thuộc vào giá trị được gộp thứ hai được biểu diễn trong miền tạo vết.

Theo phương án, bộ phận gộp thứ nhất có thể, ví dụ, được tạo cấu hình để xác định giá trị được gộp thứ nhất sao cho giá trị được gộp thứ nhất biểu thị căn bậc hai trung bình của phần tín hiệu âm thanh thứ nhất hoặc tín hiệu được suy ra từ phần tín hiệu âm thanh thứ nhất. Hơn nữa, bộ phận gộp thứ hai có thể, ví dụ, được tạo cấu hình để xác định giá trị được gộp thứ hai sao cho giá trị được gộp thứ hai biểu thị căn bậc hai trung bình của phần tín hiệu âm thanh thứ hai hoặc tín hiệu được suy ra từ phần tín hiệu âm thanh thứ hai.

Trong phương án, bộ phận biến đổi có thể, ví dụ, được tạo cấu hình để biến đổi giá trị được suy ra từ phần tín hiệu âm thanh thứ hai từ miền thứ hai đến miền tạo vết bằng cách áp dụng giá trị độ khuếch đại trên giá trị được suy ra từ phần tín hiệu âm thanh thứ hai.

Theo các phương án, giá trị độ khuếch đại có thể, ví dụ, biểu thị độ khuếch đại được dẫn bởi sự tổng hợp mã hóa dự báo tuyến tính, hoặc giá trị độ khuếch đại có thể, ví dụ, biểu thị độ khuếch đại được dẫn bởi sự tổng hợp mã hóa dự báo tuyến tính và sự chỉnh giảm.

Trong phương án, bộ phận theo dấu mức nhiều âm có thể, ví dụ, được tạo cấu hình để xác định thông tin mức nhiều âm bằng cách áp dụng phương pháp thống kê tối thiểu.

Theo phương pháp, bộ phận tạo vết mức nhiều âm có thể, ví dụ, được tạo cấu hình để xác định mức nhiều âm như thông tin mức nhiều âm. Bộ phận khôi phục có thể, ví dụ, được tạo cấu hình để khôi phục phần tín hiệu âm thanh thứ ba phụ thuộc vào thông tin mức nhiều âm, nếu khung thứ ba nêu trên trong số nhiều khung không

được nhận bởi giao diện nhận hoặc khung thứ ba đã nêu được nhận bởi giao diện nhận nhưng bị sai lạc.

Trong phương án, bộ phận tạo vết mức nhiều âm có thể, ví dụ, được tạo cấu hình để xác định mức nhiều âm nền nhân tạo như thông tin mức nhiều âm được lấy từ phổ mức nhiều âm, trong đó phổ mức nhiều âm nêu trên thu được bằng cách áp dụng phương pháp thống kê tối thiểu. Bộ phận khôi phục có thể, ví dụ, được tạo cấu hình để khôi phục phần tín hiệu âm thanh thứ ba phụ thuộc vào nhiều hệ số dự báo tuyến tính, nếu khung thứ ba đã nêu trong số nhiều khung không được nhận bởi giao diện nhận hoặc khung thứ ba đã nêu được nhận bởi giao diện nhận nhưng bị sai lạc.

Theo phương án khác, bộ phận tạo vết mức nhiều âm có thể, ví dụ, được tạo cấu hình để xác định nhiều hệ số dự báo tuyến tính biểu thị mức nhiều âm nền nhân tạo như thông tin mức nhiều âm, và bộ phận khôi phục có thể, ví dụ, được tạo cấu hình để khôi phục phần tín hiệu âm thanh thứ ba phụ thuộc vào nhiều hệ số dự báo tuyến tính.

Trong phương án, bộ phận tạo vết mức nhiều âm được tạo cấu hình để xác định nhiều hệ số FFT biểu thị mức nhiều âm nền nhân tạo như thông tin mức nhiều âm, và bộ phận khôi phục thứ nhất được tạo cấu hình để khôi phục phần tín hiệu âm thanh thứ ba phụ thuộc vào mức nhiều âm nền nhân tạo được suy ra từ các hệ số FFT đã nêu, nếu khung thứ ba đã nêu trong số nhiều khung không được nhận bởi giao diện nhận hoặc nếu khung thứ ba đã nêu được nhận bởi giao diện nhận nhưng bị sai lạc.

Trong phương án, bộ phận khôi phục có thể, ví dụ, được tạo cấu hình để khôi phục phần tín hiệu âm thanh thứ ba phụ thuộc vào thông tin mức nhiều âm và phụ thuộc vào phần tín hiệu âm thanh thứ nhất, nếu khung thứ ba đã nêu trong số nhiều khung không được nhận bởi giao diện nhận hoặc khung thứ ba đã nêu được nhận bởi giao diện nhận nhưng bị sai lạc.

Theo phương án, bộ phận khôi phục có thể, ví dụ, được tạo cấu hình để khôi phục phần tín hiệu âm thanh thứ ba bằng cách suy giảm hoặc khuếch đại tín hiệu được suy ra từ phần tín hiệu âm thanh thứ nhất hoặc thứ hai.

Trong phương án, thiết bị có thể, ví dụ, còn bao gồm bộ phận dự báo dài hạn bao gồm bộ đệm độ trễ. Hơn nữa, bộ phận dự báo dài hạn có thể, ví dụ, còn được tạo cấu

hình để tạo ra tín hiệu được xử lý phụ thuộc vào phần tín hiệu âm thanh thứ nhất hoặc thứ hai, phụ thuộc vào vùng đệm trễ và phụ thuộc vào độ khuếch đại dự báo dài hạn. Hơn nữa, bộ phận dự báo dài hạn có thể, ví dụ, được tạo cấu hình để tăng giảm cường độ độ khuếch đại dự báo dài hạn về 0, nếu khung thứ ba đã nêu trong số nhiều khung không được nhận bởi giao diện nhận hoặc khung thứ ba đã nêu được nhận bởi giao diện nhận nhưng bị sai lạc.

Theo phương án, bộ phận dự báo dài hạn có thể, ví dụ, được tạo cấu hình để tăng giảm cường độ độ khuếch đại dài hạn về 0, trong đó tốc độ mà với tốc độ này độ khuếch đại dự báo dài hạn được giảm dần cường độ về 0 phụ thuộc vào thừa số giảm dần cường độ.

Trong phương án, bộ phận dự báo dài hạn có thể, ví dụ, được tạo cấu hình để cập nhật đầu vào vùng đệm trễ bằng cách lưu trữ tín hiệu được xử lý được tạo ra trong vùng đệm trễ, nếu khung thứ ba đã nêu trong số nhiều khung không được nhận bởi giao diện nhận hoặc khung thứ ba đã nêu được nhận bởi giao diện nhận nhưng bị sai lạc.

Theo phương án, bộ phận biến đổi có thể, ví dụ, là bộ phận biến đổi thứ nhất, và bộ phận khôi phục là bộ phận khôi phục thứ nhất. Thiết bị còn bao gồm bộ phận biến đổi thứ hai và bộ phận khôi phục thứ hai. Thiết bị biến đổi thứ hai có thể, ví dụ, được tạo cấu hình để biến đổi thông tin mức nhiều âm từ miền tạo vết thành miền thứ hai, nếu khung thứ tư trong số nhiều khung không được nhận bởi giao diện nhận hoặc nếu khung thứ tư đã nêu được nhận bởi giao diện nhận nhưng bị sai lạc. Hơn nữa, bộ phận khôi phục thứ hai có thể, ví dụ, được tạo cấu hình để khôi phục phần tín hiệu âm thanh thứ tư của tín hiệu âm thanh phụ thuộc vào thông tin mức nhiều âm được biểu diễn trong miền thứ hai nếu khung thứ tư đã nêu trong số nhiều khung không được nhận bởi giao diện nhận hoặc nếu khung thứ tư đã nêu được nhận bởi giao diện nhận nhưng bị sai lạc.

Theo phương án, bộ phận khôi phục thứ hai có thể, ví dụ, được tạo cấu hình để khôi phục phần tín hiệu âm thanh thứ tư phụ thuộc vào thông tin mức nhiều âm và phụ thuộc vào phần tín hiệu âm thanh thứ hai.

Theo phương án, bộ phận khôi phục thứ hai có thể, ví dụ, được tạo cấu hình để

khôi phục phần tín hiệu âm thanh thứ tư bằng cách suy giảm hoặc khuếch đại tín hiệu được suy ra từ phần tín hiệu âm thanh thứ nhất hoặc thứ hai.

Hơn nữa, phương pháp giải mã tín hiệu âm thanh được đề xuất.

Phương pháp gồm có:

- Nhận khung thứ nhất trong số nhiều khung, khung thứ nhất đã nêu bao gồm phần tín hiệu âm thanh thứ nhất, phần tín hiệu âm thanh thứ nhất đã nêu được biểu diễn trong miền thứ nhất.

- Nhận khung thứ hai trong số nhiều khung, khung thứ hai đã nêu bao gồm phần tín hiệu âm thanh thứ hai của tín hiệu âm thanh.

- Biến đổi phần tín hiệu âm thanh thứ hai hoặc giá trị hoặc tín hiệu được suy ra từ phần tín hiệu âm thanh thứ hai từ miền thứ hai đến miền tạo vết để thu được thông tin phần tín hiệu thứ hai, trong đó miền thứ hai khác với miền thứ nhất, trong đó miền tạo vết khác với miền thứ hai, và trong đó miền tạo vết có thể bằng hoặc khác với miền thứ nhất.

- Xác định thông tin mức nhiễu âm phụ thuộc vào thông tin phần tín hiệu thứ nhất được biểu diễn trong miền tạo vết, và phụ thuộc vào thông tin phần tín hiệu thứ hai được biểu diễn trong miền tạo vết, trong đó thông tin phần tín hiệu thứ nhất phụ thuộc vào phần tín hiệu âm thanh thứ nhất. Và:

- Khôi phục phần tín hiệu âm thanh thứ ba của tín hiệu âm thanh phụ thuộc vào thông tin mức nhiễu âm được biểu diễn trong miền tạo vết, nếu khung thứ ba trong số nhiều khung không được nhận hoặc nếu khung thứ ba đã nêu được nhận nhưng bị sai lạc.

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

Một số phương án của sáng chế cung cấp tham số làm nhấn biến thiên theo thời gian sao cho khả năng theo dấu của chu kỳ đồ được làm nhấn và biến thiên của nó được cân bằng tốt hơn, để phát triển thuật toán cho sự bù độ lệch, và nhìn chung để tăng tốc sự theo dấu.

Các phương án của sáng chế dựa trên việc tìm ra rằng đối với sự giảm dần cường độ, các tham số sau đây được quan tâm: Miền giảm dần cường độ; tốc độ giảm dần cường độ, hoặc, chung hơn nữa, đường cong giảm dần cường độ; mức đích của sự giảm dần cường độ; hình dạng phổ đích của sự giảm dần cường độ; và/hoặc sự tạo vết mức tạp âm nền. Trong bối cảnh này, các phương án được dựa trên việc tìm ra rằng tình trạng kỹ thuật có hạn chế đáng kể.

Thiết bị và phương pháp cho sự giảm dần cường độ tín hiệu được cải thiện cho các hệ thống mã hóa âm thanh được chuyển trong quá trình che giấu được đề xuất.

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

Các phương án nhận ra sự giảm dần cường độ đến mức nhiễu âm nền nhân tạo. Theo các phương án, sự tạo vết mức nhiễu âm nền nhân tạo chung trong miền kích thích được nhận ra. Mức nhiễu âm nền nhân tạo được tạo đích trong quá trình mất gói khối tín hiệu truyền sẽ là giống nhau, không xét đến bộ mã hóa lõi (ACELP/TCX) đang sử dụng, và nó sẽ luôn được cập nhật. Không có tình trạng kỹ thuật đã biết, trong đó sự tạo vết mức nhiễu âm là cần thiết. Các phương án đề xuất sự tăng giảm cường độ của bộ mã hóa-giải mã được chuyển đến tín hiệu tương tự nhiễu âm nền nhân tạo trong quá trình mất gói khối tín hiệu truyền.

Hơn nữa, các phương án nhận ra rằng toàn bộ độ phức tạp sẽ được hạ thấp so với hai mô đun tạo vết mức nhiễu âm độc lập, vì các hàm (PROM) và bộ nhớ có thể được chia sẻ.

Trong các phương án, phép đạo hàm mức trong miền kích thích (so với phép đạo hàm mức trong miền thời gian) cung cấp nhiều cực tiểu trong quá trình tiếng nói hoạt động, vì phần của thông tin tiếng nói được che bởi các hệ số LP.

Trong trường hợp ACELP, theo các phương án, phép đạo hàm mức diễn ra trong miền kích thích. Trong trường hợp của TCX, trong các phương án, mức được đạo hàm trong miền thời gian, và độ khuếch đại của sự tổng hợp LPC và sự chỉnh giảm được áp dụng như hệ số hiệu chỉnh để làm mẫu mức năng lượng trong miền kích thích. Tạo vết mức trong miền kích thích, ví dụ, trước FDNS, về lý thuyết cũng sẽ khả thi, nhưng sự

bù mức giữa miền kích thích TCX và miền kích thích ACELP được cho là khá phức tạp.

Không có tình trạng kỹ thuật kết hợp sự tạo vết mức nền thông thường trong các miền khác nhau. Các công nghệ của tình trạng kỹ thuật không có sự tạo vết mức nhiều âm nền nhân tạo thông thường như vậy, ví dụ, trong miền kích thích, trong hệ thống bộ mã hóa-giải mã được chuyển. Do đó, các phương án có lợi thế hơn so với tình trạng kỹ thuật, vì đối với các công nghệ của tình trạng kỹ thuật, mức nhiều âm nền nhân tạo mà được lấy mục tiêu trong quá trình mất gói có thể khác, phụ thuộc vào chế độ mã hóa đứng trước (ACELP/TCX), trong đó mức được tạo vết; như trong tình trạng kỹ thuật, sự tạo vết mà được tách cho mỗi chế độ mã hóa sẽ gây ra sự phức tạp tính toán bổ sung và phía trước không cần thiết; và như trong tình trạng kỹ thuật, không có mức nhiều âm nền nhân tạo hiện đại nào có thể sẵn có trong lõi do sự chuyển gần đây đến lõi này.

Theo một số phương án, sự tạo vết mức được thực hiện trong miền kích thích, nhưng sự giảm dần cường độ TCX được thực hiện trong miền thời gian. Bằng cách tăng giảm cường độ trong miền thời gian, các thất bại của TDAC có thể được ngăn chặn, mà sẽ gây ra sai số lấy mẫu. Điều này trở nên được quan tâm đặc biệt khi các thành phần tín hiệu âm được che giấu. Hơn nữa, sự chuyển mức giữa miền kích thích ACELP và miền phổ MDCT được ngăn chặn và do đó, ví dụ, các nguồn tính toán được lưu lại. Bởi vì sự chuyển giữa miền kích thích và miền thời gian, sự điều chỉnh mức được yêu cầu giữa miền kích thích và miền thời gian. Điều này được giải quyết bởi phép đạo hàm của độ khuếch đại mà sẽ được giới thiệu bằng sự tổng hợp LPC và sự chỉnh tăng và để sử dụng giá trị này như hệ số hiệu chỉnh để chuyển đổi mức giữa hai miền.

Ngược lại, các công nghệ trong lĩnh vực kỹ thuật không tiến hành tạo vết mức trong miền kích thích và sự giảm dần cường độ TCX trong miền thời gian. Đối với các bộ mã hóa biến đổi trên cơ sở biến đổi của tình trạng kỹ thuật, hệ số suy giảm được áp dụng hoặc trong miền kích thích (đối với miền thời gian/các phương pháp che giấu như ACELP), xem [3GP09a]) hoặc trong miền tần số (đối với các phương pháp miền tần số như sự lặp khung hoặc sự thay thế nhiều âm, xem [LS01]). Nhược điểm của

phương pháp của tình trạng kỹ thuật là để áp dụng hệ số suy giảm trong miền tần số mà sai số lấy mẫu sẽ bị gây ra trong vùng chồng lấp-cộng trong miền thời gian. Điều này sẽ là trường hợp cho khung liên kề mà các hệ số suy giảm khác được áp dụng cho khung này, vì quy trình tăng giảm cường độ gây ra sự triệt tiêu sai số lấy mẫu trong miền thời gian (time domain alias cancellation - TDAC) để thất bại. Điều này liên quan đặc biệt khi các thành phần tín hiệu âm được che giấu. Các phương án nêu trên do đó là có lợi so với tình trạng kỹ thuật.

Các phương án bù sự ảnh hưởng của bộ lọc thông cao trên hệ số tổng hợp LPC. Theo các phương án, để bù cho sự thay đổi độ khuếch đại không mong muốn của sự tổng hợp và sự chỉnh tần LPC được gây ra bởi sự kích thích vô thanh được lọc thông cao, hệ số hiệu chỉnh được đạo hàm. Hệ số hiệu chỉnh xem xét sự thay đổi độ khuếch đại không mong muốn này và thay đổi mức nhiễu âm nền nhân tạo đích trong miền kích thích sao cho mức đích đúng đạt được trong miền thời gian.

Ngược lại, trong tình trạng kỹ thuật, ví dụ, G.718 [ITU08a], giới thiệu bộ lọc thông cao vào đường dẫn tín hiệu của sự kích thích vô thanh, như được thể hiện trên Fig.2, nếu tín hiệu của khung tốt cuối cùng không được phân loại là VÔ THANH. Bằng cách này, các công nghệ tình trạng kỹ thuật gây ra các hiệu ứng phụ không mong muốn, vì độ khuếch đại của sự tổng hợp LPC phụ thuộc vào các đặc điểm tín hiệu, mà được thay đổi bởi bộ lọc thông cao này. Vì mức nền được tạo vết và được áp dụng trong miền kích thích, thuật toán dựa vào độ khuếch đại sự tổng hợp LPC, mà lại một lần nữa phụ thuộc vào đặc điểm của tín hiệu kích thích. Nói cách khác: Sự thay đổi các đặc điểm tín hiệu của sự kích thích do bộ lọc thông cao, như được thực hiện bởi tình trạng kỹ thuật, có thể dẫn đến độ khuếch đại được thay đổi (thường là được giảm) của sự tổng hợp LPC. Điều này dẫn đến mức đầu ra sai mặc dù mức kích thích đúng.

Các phương án khắc phục các bất lợi của tình trạng kỹ thuật.

Cụ thể, các phương án nhận ra hình dạng phổ thích ứng của nhiễu âm nền nhân tạo. Trái ngược với G.718, bằng cách tạo vết hình dạng phổ của tạp âm nền, và bằng cách áp dụng (tăng giảm cường độ) hình dạng này trong quá trình mất gói khôi tín hiệu truyền, đặc điểm nhiễu âm của tạp âm nền sẽ được so khớp, dẫn đến đặc điểm nhiễu âm dễ chịu của nhiễu âm nền nhân tạo. Điều này tránh việc không so khớp gây trở

ngại cho hình dạng phổ mà có thể được dẫn bằng cách sử dụng đường bao phổ mà được đạo hàm bởi chuỗi ngoại tuyến và/hoặc hình dạng phổ của các khung được nhận cuối cùng.

Hơn nữa, thiết bị giải mã tín hiệu âm thanh được đề xuất. Thiết bị bao gồm giao diện nhận, trong đó giao diện nhận được tạo cấu hình để nhận khung thứ nhất bao gồm phần tín hiệu âm thanh thứ nhất của tín hiệu âm thanh, và trong đó giao diện nhận được tạo cấu hình để nhận khung thứ hai bao gồm phần tín hiệu âm thanh thứ hai của tín hiệu âm thanh.

Hơn nữa, thiết bị bao gồm bộ phận tạo vết mức nhiều âm, trong đó bộ phận tạo vết mức nhiều âm được tạo cấu hình để xác định thông tin mức nhiều âm phụ thuộc vào ít nhất một trong số phần tín hiệu âm thanh thứ nhất và phần tín hiệu âm thanh thứ hai (điều này có nghĩa là: phụ thuộc vào phần tín hiệu âm thanh thứ nhất và/hoặc phần tín hiệu âm thanh thứ hai), trong đó thông tin mức nhiều âm được biểu diễn trong miền tạo vết.

Hơn nữa, thiết bị bao gồm bộ phận khôi phục thứ nhất để khôi phục, trong miền khôi phục thứ nhất, phần tín hiệu âm thanh thứ ba của tín hiệu âm thanh phụ thuộc vào thông tin mức nhiều âm, nếu khung thứ ba trong số nhiều khung không được nhận bởi giao diện nhận hoặc nếu khung thứ ba đã nêu được nhận bởi giao diện nhận nhưng bị sai lạc, trong đó miền khôi phục thứ nhất khác với hoặc bằng miền tạo vết.

Hơn nữa, thiết bị bao gồm bộ phận biến đổi để biến đổi thông tin mức nhiều âm từ miền tạo vết đến miền khôi phục thứ hai, nếu khung thứ tư trong số nhiều khung không được nhận bởi giao diện nhận hoặc nếu khung thứ tư đã nêu được nhận bởi giao diện nhận nhưng bị sai lạc, trong đó, miền khôi phục thứ hai khác với miền tạo vết, và trong đó miền khôi phục thứ hai khác với miền khôi phục thứ nhất, và

Hơn nữa, thiết bị bao gồm bộ phận khôi phục thứ hai để khôi phục, trong miền khôi phục thứ hai, phần tín hiệu âm thanh thứ tư của tín hiệu âm thanh phụ thuộc vào thông tin mức nhiều âm được biểu diễn trong miền khôi phục thứ hai, nếu khung thứ tư đã nêu trong số nhiều khung không được nhận bởi giao diện nhận hoặc nếu khung thứ tư đã nêu được nhận bởi giao diện nhận nhưng bị sai lạc.

Theo một số phương án, miền tạo vết có thể, ví dụ, trong đó miền tạo vết là miền thời gian, miền phổ, miền FFT, miền MDCT, hoặc miền kích thích. Miền khôi phục thứ nhất có thể, ví dụ, là miền thời gian, miền phổ, miền FFT, miền MDCT, hoặc miền kích thích. Miền khôi phục thứ hai có thể, ví dụ, là miền thời gian, miền phổ, miền FFT, miền MDCT, hoặc miền kích thích.

Trong phương án, miền tạo vết có thể, ví dụ, là miền FFT, miền khôi phục thứ nhất có thể, ví dụ, là miền thời gian, và miền khôi phục thứ hai có thể, ví dụ, là miền kích thích.

Trong phương án khác, miền tạo vết có thể, ví dụ, là miền thời gian, miền khôi phục thứ nhất có thể, ví dụ, là miền thời gian, và miền khôi phục thứ hai có thể, ví dụ, là miền kích thích.

Theo phương án, phần tín hiệu âm thanh thứ nhất đã nêu có thể, ví dụ, được biểu diễn trong miền đầu vào thứ nhất, và phần tín hiệu âm thanh thứ hai có thể, ví dụ, được biểu diễn trong miền đầu vào thứ hai. Bộ phận biến đổi có thể, ví dụ, là bộ phận biến đổi thứ hai. Thiết bị có thể, ví dụ, còn bao gồm bộ phận biến đổi thứ nhất để biến đổi phần tín hiệu âm thanh thứ hai hoặc giá trị hoặc tín hiệu được suy ra từ phần tín hiệu âm thanh thứ hai từ miền đầu vào thứ hai đến miền tạo vết để thu được thông tin phần tín hiệu thứ hai. Bộ phận tạo vết mức nhiều âm có thể, ví dụ, được tạo cấu hình để nhận thông tin phần tín hiệu thứ nhất được biểu diễn trong miền tạo vết, trong đó thông tin phần tín hiệu thứ nhất phụ thuộc vào phần tín hiệu âm thanh thứ nhất, trong đó bộ phận tạo vết mức nhiều âm được tạo cấu hình để nhận phần tín hiệu thứ hai được biểu diễn trong miền tạo vết, và trong đó bộ phận tạo vết mức nhiều âm được tạo cấu hình để xác định thông tin mức nhiều âm phụ thuộc vào thông tin phần tín hiệu thứ nhất được biểu diễn trong miền tạo vết và phụ thuộc vào thông tin phần tín hiệu thứ hai được biểu diễn trong miền tạo vết.

Theo phương án, miền đầu vào thứ nhất có thể, ví dụ, là miền kích thích, và miền đầu vào thứ hai có thể, ví dụ, là miền MDCT.

Trong phương án khác, miền đầu vào thứ nhất có thể, ví dụ, là miền MDCT, và trong đó miền đầu vào thứ hai có thể, ví dụ, là miền MDCT.

Theo phương án, bộ phận khôi phục thứ nhất có thể, ví dụ, được tạo cấu hình để khôi phục phần tín hiệu âm thanh thứ ba bằng cách thực hiện sự tăng giảm cường độ đến phổ như nhiều âm. Bộ phận khôi phục thứ hai có thể, ví dụ, được tạo cấu hình để khôi phục phần tín hiệu âm thanh thứ tư bằng cách thực hiện sự tăng giảm cường độ thứ hai đến phổ như nhiều âm và/hoặc sự tăng giảm cường độ thứ hai của độ khuếch đại LTP. Hơn nữa, bộ phận khôi phục thứ nhất và bộ phận khôi phục thứ hai có thể, ví dụ, được tạo cấu hình để thực hiện sự tăng giảm cường độ thứ nhất và sự tăng giảm cường độ thứ hai đến phổ như nhiều âm và/hoặc sự tăng giảm cường độ thứ hai của độ khuếch đại LTP với tốc độ tăng giảm cường độ tương tự.

Trong phương án, thiết bị có thể, ví dụ, còn bao gồm bộ phận gộp thứ nhất để xác định giá trị được gộp thứ nhất phụ thuộc vào phần tín hiệu âm thanh thứ nhất. Hơn nữa, thiết bị còn có thể, ví dụ, bao gồm bộ phận gộp thứ hai để xác định, phụ thuộc vào phần tín hiệu âm thanh thứ hai, giá trị được gộp thứ hai như giá trị được suy ra từ phần tín hiệu âm thanh thứ hai. Bộ phận tạo vết mức nhiều âm có thể, ví dụ, được tạo cấu hình để nhận giá trị được gộp thứ nhất như thông tin phần tín hiệu thứ nhất được biểu diễn trong miền tạo vết, trong đó bộ phận tạo vết mức nhiều âm có thể, ví dụ, được tạo cấu hình để nhận giá trị được gộp thứ hai như thông tin phần tín hiệu thứ hai được biểu diễn trong miền tạo vết, và trong đó bộ phận tạo vết mức nhiều âm được tạo cấu hình để xác định thông tin mức nhiều âm phụ thuộc vào giá trị được gộp thứ nhất được biểu diễn trong miền tạo vết và phụ thuộc vào giá trị được gộp thứ hai được biểu diễn trong miền tạo vết.

Theo phương án, bộ phận gộp thứ nhất có thể, ví dụ, được tạo cấu hình để xác định giá trị được gộp thứ nhất sao cho giá trị được gộp thứ nhất biểu thị căn bậc hai trung bình của phần tín hiệu âm thanh thứ nhất hoặc tín hiệu được suy ra từ phần tín hiệu âm thanh thứ nhất. Bộ phận gộp thứ hai được tạo cấu hình để xác định giá trị được gộp thứ hai sao cho giá trị được gộp thứ hai biểu thị căn bậc hai trung bình của phần tín hiệu âm thanh thứ hai hoặc tín hiệu được suy ra từ phần tín hiệu âm thanh thứ hai.

Trong phương án, bộ phận biến đổi thứ nhất có thể, ví dụ, được tạo cấu hình để biến đổi giá trị được suy ra từ phần tín hiệu âm thanh thứ hai từ miền đầu vào thứ hai

đến miền tạo vết bằng cách áp dụng giá trị độ khuếch đại trên giá trị được suy ra từ phần tín hiệu âm thanh thứ hai.

Theo phương án, giá trị độ khuếch đại có thể, ví dụ, biểu thị độ khuếch đại được dẫn bởi sự tổng hợp mã hóa dự báo tuyến tính, hoặc trong đó giá trị độ khuếch đại biểu thị độ khuếch đại được dẫn bởi sự tổng hợp mã hóa dự báo tuyến tính và sự chỉnh giảm.

Trong phương án, bộ phận theo dấu mức nhiều âm có thể, ví dụ, được tạo cấu hình để xác định thông tin mức nhiều âm bằng cách áp dụng phương pháp thống kê tối thiểu.

Theo phương pháp, bộ phận tạo vết mức nhiều âm có thể, ví dụ, được tạo cấu hình để xác định mức nhiều âm như thông tin mức nhiều âm. Bộ phận khôi phục có thể, ví dụ, được tạo cấu hình để khôi phục phần tín hiệu âm thanh thứ ba phụ thuộc vào thông tin mức nhiều âm, nếu khung thứ ba đã nêu trong số nhiều khung không được nhận bởi giao diện nhận hoặc khung thứ ba đã nêu được nhận bởi giao diện nhận nhưng bị sai lạc.

Trong phương án, bộ phận tạo vết mức nhiều âm có thể, ví dụ, được tạo cấu hình để xác định mức nhiều âm nền nhân tạo như thông tin mức nhiều âm được suy ra từ phổ mức nhiều âm, trong đó phổ mức nhiều âm đã nêu thu được bằng cách áp dụng phương pháp thống kê tối thiểu. Bộ phận khôi phục có thể, ví dụ, được tạo cấu hình để khôi phục phần tín hiệu âm thanh thứ ba phụ thuộc vào nhiều hệ số dự báo tuyến tính, nếu khung thứ ba đã nêu trong số nhiều khung không được nhận bởi giao diện nhận hoặc khung thứ ba đã nêu được nhận bởi giao diện nhận nhưng bị sai lạc.

Theo phương án, bộ phận khôi phục thứ nhất có thể, ví dụ, được tạo cấu hình để khôi phục phần tín hiệu âm thanh thứ ba phụ thuộc vào thông tin mức nhiều âm và phụ thuộc vào phần tín hiệu âm thanh thứ nhất, nếu khung thứ ba đã nêu trong số nhiều khung không được nhận bởi giao diện nhận hoặc khung thứ ba đã nêu được nhận bởi giao diện nhận nhưng bị sai lạc.

Trong phương án, bộ phận khôi phục thứ nhất có thể, ví dụ, được tạo cấu hình để khôi phục phần tín hiệu âm thanh thứ ba bằng cách suy giảm hoặc khuếch đại phần tín

hiệu âm thanh thứ nhất.

Theo phương án, bộ phận khôi phục thứ hai có thể, ví dụ, được tạo cấu hình để khôi phục phần tín hiệu âm thanh thứ tư phụ thuộc vào thông tin mức nhiễu âm và phụ thuộc vào phần tín hiệu âm thanh thứ hai.

Trong phương án, bộ phận khôi phục thứ hai có thể, ví dụ, được tạo cấu hình để khôi phục phần tín hiệu âm thanh thứ tư bằng cách suy giảm hoặc khuếch đại phần tín hiệu âm thanh thứ hai.

Theo phương án, thiết bị có thể, ví dụ, còn bao gồm bộ phận dự báo dài hạn bao gồm vùng đệm trễ, trong đó bộ phận dự báo dài hạn có thể, ví dụ, được tạo cấu hình để tạo ra tín hiệu được xử lý phụ thuộc vào phần tín hiệu âm thanh thứ hai, phụ thuộc vào đầu vào vùng đệm trễ được lưu trữ trong vùng đệm trễ và phụ thuộc vào độ khuếch đại dự báo dài hạn, và trong đó bộ phận dự báo dài hạn được tạo cấu hình để tăng giảm độ khuếch đại dự báo dài hạn về 0, nếu khung thứ ba đã nêu trong số nhiều khung không được nhận bởi giao diện nhận hoặc nếu khung thứ ba đã nêu được nhận bởi giao diện nhận nhưng bị sai lạc.

Trong phương án, bộ phận dự báo dài hạn có thể, ví dụ, được tạo cấu hình để tăng giảm cường độ độ khuếch đại dài hạn về 0, trong đó tốc độ mà độ khuếch đại dự báo dài hạn được giảm dần cường độ về 0 phụ thuộc vào thừa số giảm dần cường độ.

Trong phương án, bộ phận dự báo dài hạn có thể, ví dụ, được tạo cấu hình để cập nhật đầu vào vùng đệm trễ bằng cách lưu trữ tín hiệu được xử lý được tạo ra trong vùng đệm trễ, nếu khung thứ ba đã nêu trong số nhiều khung không được nhận bởi giao diện nhận hoặc khung thứ ba đã nêu được nhận bởi giao diện nhận nhưng bị sai lạc.

Hơn nữa, phương pháp giải mã tín hiệu âm thanh được đề xuất. Phương pháp gồm có:

- Nhận khung thứ nhất bao gồm phần tín hiệu âm thanh thứ nhất của tín hiệu âm thanh, và nhận khung thứ hai bao gồm phần tín hiệu âm thanh thứ hai của tín hiệu âm thanh.

- Xác định thông tin mức nhiễu âm phụ thuộc vào ít nhất một trong số phần tín

hiệu âm thanh thứ nhất và phần tín hiệu âm thanh thứ hai, trong đó thông tin mức nhiều âm được biểu diễn trong miền tạo vết.

- Khôi phục, trong miền khôi phục thứ nhất, phần tín hiệu âm thanh thứ ba của tín hiệu âm thanh phụ thuộc vào thông tin mức nhiều âm, nếu khung thứ ba trong số nhiều khung không được nhận hoặc nếu khung thứ ba đã nêu được nhận nhưng bị sai lạc, trong đó miền khôi phục thứ nhất khác với hoặc bằng miền tạo vết.

- Biên đổi thông tin mức nhiều âm từ miền tạo vết đến miền khôi phục thứ hai, nếu khung thứ tư trong số nhiều khung không được nhận hoặc nếu khung thứ tư đã nêu được nhận nhưng bị sai lạc, trong đó, miền khôi phục thứ hai khác với miền tạo vết, và trong đó miền khôi phục thứ hai khác với miền khôi phục thứ nhất. Và:

- Khôi phục, trong miền khôi phục thứ hai, phần tín hiệu âm thanh thứ tư của tín hiệu âm thanh phụ thuộc vào thông tin mức nhiều âm được biểu diễn trong miền khôi phục thứ hai, nếu khung thứ tư đã nêu trong số nhiều khung không được nhận hoặc nếu khung thứ tư đã nêu được nhận nhưng bị sai lạc.

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

Hơn nữa, thiết bị giải mã tín hiệu âm thanh được mã hóa để thu được tín hiệu âm thanh được khôi phục được đề xuất. Thiết bị bao gồm giao diện nhận để nhận một hoặc nhiều khung, bộ tạo hệ số, và bộ khôi phục tín hiệu. Bộ tạo hệ số được tạo cấu hình để xác định, nếu khung hiện thời trong số một hoặc nhiều khung được nhận bởi giao diện nhận và nếu khung hiện thời được nhận bởi giao diện nhận không bị sai lạc, một hoặc nhiều hệ số tín hiệu âm thanh thứ nhất, được bao gồm bởi khung hiện thời, trong đó một hoặc nhiều hệ số tín hiệu âm thanh thứ nhất đã nêu biểu thị đặc điểm của tín hiệu âm thanh được mã hóa, và một hoặc nhiều hệ số nhiều âm biểu thị tập âm nền của tín hiệu âm thanh được mã hóa. Hơn nữa, bộ tạo hệ số được tạo cấu hình để tạo ra một hoặc nhiều hệ số tín hiệu âm thanh thứ hai, phụ thuộc vào một hoặc nhiều hệ số tín hiệu âm thanh thứ nhất và phụ thuộc vào một hoặc nhiều hệ số nhiều âm, nếu khung hiện thời không được nhận bởi giao diện nhận hoặc nếu khung hiện thời được nhận bởi giao diện nhận bị sai lạc. Bộ khôi phục tín hiệu âm thanh được tạo cấu hình để khôi phục phần thứ nhất của tín hiệu âm thanh được khôi phục phụ thuộc vào một

hoặc nhiều hệ số tín hiệu âm thanh thứ nhất, nếu khung hiện thời được nhận bởi giao diện nhận và nếu khung hiện thời được nhận bởi giao diện nhận không bị sai lạc. Hơn nữa, bộ khôi phục tín hiệu âm thanh được tạo cấu hình để khôi phục phần thứ hai của tín hiệu âm thanh được khôi phục phụ thuộc vào một hoặc nhiều hệ số tín hiệu âm thanh thứ hai, nếu khung hiện thời không được nhận bởi giao diện nhận hoặc nếu khung hiện thời được nhận bởi giao diện nhận bị sai lạc.

Trong một số phương án, một hoặc nhiều hệ số tín hiệu âm thanh thứ nhất có thể, ví dụ, là một hoặc nhiều hệ số bộ lọc dự báo tuyến tính của tín hiệu âm thanh được mã hóa. Trong một số phương án, một hoặc nhiều hệ số tín hiệu âm thanh thứ nhất có thể, ví dụ, là một hoặc nhiều hệ số bộ lọc dự báo tuyến tính của tín hiệu âm thanh được mã hóa.

Theo phương án, một hoặc nhiều hệ số nhiễu âm có thể, ví dụ, là một hoặc nhiều hệ số bộ lọc dự báo tuyến tính biểu thị tạp âm nền của tín hiệu âm thanh được mã hóa. Trong phương án, một hoặc nhiều hệ số bộ lọc dự báo tuyến tính có thể, ví dụ, biểu diễn hình dạng phổ của tạp âm nền.

Trong phương án, bộ tạo hệ số có thể, ví dụ, được tạo cấu hình để xác định một hoặc nhiều phần tín hiệu âm thanh sao cho một hoặc nhiều phần tín hiệu âm thanh thứ hai là một hoặc nhiều hệ số bộ lọc dự báo tuyến tính của tín hiệu âm thanh được khôi phục, hoặc sao cho một hoặc nhiều hệ số tín hiệu âm thanh thứ nhất là một hoặc nhiều cặp phổ điện vào của tín hiệu âm thanh được khôi phục.

Theo phương án, bộ tạo hệ số có thể, ví dụ, được tạo cấu hình để tạo ra một hoặc nhiều hệ số tín hiệu âm thanh thứ hai bằng cách áp dụng công thức:

$$f_{current}[i] = \alpha \cdot f_{last}[i] + (1 - \alpha) \cdot pt_{mean}[i]$$

trong đó  $f_{current}[i]$  biểu thị một trong số một hoặc nhiều các hệ số tín hiệu âm thanh thứ hai, trong đó  $f_{last}[i]$  biểu thị một trong số một hoặc nhiều hệ số tín hiệu âm thanh thứ nhất, trong đó  $pt_{mean}[i]$  là một trong số một hoặc nhiều hệ số nhiễu âm, trong đó  $\alpha$  là số thực với  $0 \leq \alpha \leq 1$ , và trong đó  $i$  là chỉ số. Trong phương án,  $0 < \alpha < 1$ .

Theo phương án,  $f_{last}[i]$  biểu thị hệ số bộ lọc dự báo tuyến tính của tín hiệu âm thanh được mã hóa, và trong đó  $f_{current}[i]$  biểu thị hệ số bộ lọc dự báo tuyến tính của tín

hiệu âm thanh được khôi phục.

Trong phương án,  $pt_{mean}[i]$  có thể, ví dụ, biểu thị tập âm nền của tín hiệu âm thanh được mã hóa.

Trong phương án, bộ tạo hệ số có thể, ví dụ, được tạo cấu hình để xác định, nếu khung hiện thời của một hoặc nhiều khung được nhận bởi giao diện nhận và nếu khung hiện thời được nhận bởi giao diện nhận không bị sai lạc, một hoặc nhiều hệ số nhiễu âm bằng cách xác định phổ nhiễu âm của tín hiệu âm thanh được mã hóa.

Theo phương án, bộ tạo hệ số có thể, ví dụ, được tạo cấu hình để xác định các hệ số LPC biểu diễn tập âm nền bằng cách sử dụng phương pháp thống kê tối thiểu trên phổ tín hiệu để xác định phổ tạp âm nền và bằng cách tính toán các hệ số LPC biểu diễn hình dạng tạp âm nền từ phổ tạp âm nền.

Hơn nữa, phương pháp giải mã tín hiệu âm thanh được mã hóa để thu được tín hiệu âm thanh được khôi phục được đề xuất. Phương pháp gồm có:

- Nhận một hoặc nhiều khung.
- Xác định, nếu khung hiện thời trong số một hoặc nhiều khung được nhận và nếu khung hiện thời được nhận không bị sai lạc, một hoặc nhiều hệ số tín hiệu âm thanh thứ nhất, được bao gồm bởi khung hiện thời, trong đó một hoặc nhiều hệ số tín hiệu âm thanh thứ nhất đã nêu biểu thị đặc điểm của tín hiệu âm thanh được mã hóa, và một hoặc nhiều hệ số nhiễu âm biểu thị tạp âm nền của tín hiệu âm thanh được mã hóa.
- Tạo ra một hoặc nhiều hệ số tín hiệu âm thanh thứ hai, phụ thuộc vào một hoặc nhiều hệ số tín hiệu âm thanh thứ nhất và phụ thuộc vào một hoặc nhiều hệ số nhiễu âm, nếu khung hiện thời không được nhận hoặc nếu khung hiện thời được nhận bị sai lạc.
- Khôi phục phần thứ nhất của tín hiệu âm thanh được khôi phục phụ thuộc vào một hoặc nhiều hệ số tín hiệu âm thanh thứ nhất, nếu khung hiện thời được nhận và nếu khung hiện thời được nhận không bị sai lạc. Và:
- Khôi phục phần thứ hai của tín hiệu âm thanh được khôi phục phụ thuộc vào một hoặc nhiều hệ số tín hiệu âm thanh thứ hai, nếu khung hiện thời không được nhận và nếu khung hiện thời được nhận bị sai lạc.

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

Các phương tiện thông thường để tạo vết và áp dụng hình dạng phổ của nhiều âm nền nhân tạo trong quá trình giảm dần cường độ có một số ưu điểm sau. Bằng cách tạo vết và áp dụng hình dạng phổ sao cho nó có thể được thực hiện tương tự cho cả các bộ mã hóa-giải mã lỗi cho phép phương pháp chung đơn giản. CELT dẫn dắt rằng chỉ sự tạo vết theo băng của các năng lượng trong miền phổ và sự tạo thành theo băng của hình dạng phổ trong miền phổ, mà không khả thi cho lỗi CELP.

Ngược lại, trong tình trạng kỹ thuật, hình dạng phổ của nhiều âm nền nhân tạo được dẫn trong quá trình mất khối tín hiệu truyền hoặc hoàn toàn tĩnh, hoặc tĩnh một phần và thích ứng một phần với loại hình dạng phổ ngắn hạn (như được nhận ra trong G.718 [ITU08a]), và sẽ thường không so khớp với tạp âm nền trong tín hiệu trước sự mất gói. Sự không so khớp này của các đặc điểm nhiều âm nền nhân tạo có thể là nhiễu loạn. Theo tình trạng kỹ thuật, tạo hình tạp âm nền (tĩnh) được tạo chuỗi ngoại tuyến có thể được sử dụng mà có thể là dễ chịu đối với các tín hiệu cụ thể, nhưng sẽ ít dễ chịu đối với các tín hiệu khác, ví dụ, tiếng ồn của xe hoàn toàn khác với tiếng ồn văn phòng.

Hơn nữa, trong tình trạng kỹ thuật, sự thích ứng với loại hình dạng phổ ngắn hạn của các khung được nhận trước đó có thể được sử dụng mà sẽ mang các đặc điểm tín hiệu gần đến tín hiệu được nhận trước đó, nhưng không nhất thiết đến các đặc điểm tạp âm nền. Trong tình trạng kỹ thuật, sự tạo vết hình dạng phổ theo băng trong miền phổ (như được nhận ra trong CELT [IET12]) là không thể áp dụng cho bộ mã hóa-giải mã được chuyển sử dụng không chỉ lỗi trên cơ sở miền MDCT (TCX) mà cũng dựa trên lỗi trên cơ sở ACELP. Các phương án nêu trên do đó là có lợi so với tình trạng kỹ thuật.

Hơn nữa, thiết bị giải mã tín hiệu âm thanh được mã hóa để thu được tín hiệu âm thanh được khôi phục được đề xuất. Thiết bị giải mã tín hiệu âm thanh được mã hóa để thu được tín hiệu âm thanh được khôi phục được đề xuất. Thiết bị bao gồm giao diện nhận để nhận một hoặc nhiều khung bao gồm thông tin trên nhiều mẫu tín hiệu âm thanh của phổ tín hiệu âm thanh của tín hiệu âm thanh được mã hóa, và bộ mã hóa để

tạo ra tín hiệu âm thanh được khôi phục. Bộ xử lý được tạo cấu hình để tạo ra tín hiệu âm thanh được khôi phục bằng cách tăng giảm cường độ phổ được thay đổi đến phổ đích, nếu khung hiện thời không được nhận bởi giao diện nhận hoặc nếu khung hiện thời được nhận bởi giao diện nhận nhưng bị sai lạc, trong đó phổ được thay đổi bao gồm nhiều mẫu phổ tín hiệu được thay đổi, trong đó, đối với mỗi mẫu tín hiệu được thay đổi của phổ được thay đổi, giá trị tuyệt đối của mẫu tín hiệu được thay đổi đã nêu bằng với giá trị tuyệt đối của một trong số các mẫu tín hiệu âm thanh của phổ tín hiệu âm thanh. Hơn nữa, bộ xử lý được tạo cấu hình để không làm tăng giảm cường độ phổ được thay đổi đến phổ đích, nếu khung hiện thời của một hoặc nhiều khung được nhận bởi giao diện nhận và nếu khung hiện thời được nhận bởi giao diện nhận không bị sai lạc.

Theo phương án, phổ đích có thể, ví dụ, là phổ như nhiều âm.

Trong phương án, phổ như nhiều âm có thể, ví dụ, biểu diễn nhiều âm trắng.

Theo phương án, phổ đích có thể, ví dụ, được tạo hình.

Theo phương án, hình dạng của phổ như nhiều âm có thể, ví dụ, phụ thuộc vào phổ tín hiệu âm thanh của tín hiệu âm thanh được nhận lúc trước.

Theo phương án, phổ như nhiều âm có thể, ví dụ, được tạo hình dựa trên hình dạng của phổ tín hiệu âm thanh.

Theo phương án, bộ xử lý có thể, ví dụ, sử dụng hệ số nghiêng để tạo hình phổ như nhiều âm.

Theo phương án, bộ xử lý có thể, ví dụ, sử dụng công thức

$$\text{shaped\_noise}[i] = \text{noise} * \text{power}(\text{tilt\_factor}, i/N)$$

trong đó N biểu thị số lượng mẫu, trong đó i là chỉ số, trong đó  $0 \leq i < N$ , với  $\text{tilt\_factor} > 0$ , và trong đó power là hàm lũy thừa.

$\text{power}(x, y)$  biểu thị  $x^y$

$\text{power}(\text{tilt\_factor}, i/N)$  biểu thị  $\text{tilt\_factor}^{\frac{i}{N}}$

Nếu  $\text{tilt\_factor}$  là nhỏ hơn 1 điều này có nghĩa là sự suy giảm với i tăng.

Nếu `tilt_factor` là lớn hơn 1 có nghĩa là sự khuếch đại với  $i$  tăng.

Theo phương án khác, bộ xử lý có thể, ví dụ, sử dụng công thức

$$\text{shaped\_noise}[i] = \text{noise} * (1 + i / (N-1) * (\text{tilt\_factor}-1))$$

trong đó  $N$  biểu thị số lượng mẫu, trong đó  $i$  là chỉ số, trong đó  $0 \leq i < N$ , với  $\text{tilt\_factor} > 0$ .

Nếu `tilt_factor` là nhỏ hơn 1 điều này có nghĩa là sự suy giảm với  $i$  tăng.  
Nếu `tilt_factor` là lớn hơn 1 có nghĩa là sự khuếch đại với  $i$  tăng.

Theo phương án, bộ xử lý có thể, ví dụ, được tạo cấu hình để tạo ra phổ được thay đổi, bằng cách thay đổi ký hiệu của một hoặc nhiều mẫu trong số các mẫu tín hiệu âm thanh của phổ tín hiệu âm thanh, nếu khung hiện thời không được nhận bởi giao diện nhận hoặc nếu khung hiện thời được nhận bởi giao diện nhận bị sai lạc.

Theo phương án, mỗi mẫu tín hiệu âm thanh của phổ tín hiệu âm thanh có thể, ví dụ, được biểu diễn bởi số thực nhưng không phải bởi số ảo.

Theo phương án, các mẫu tín hiệu âm thanh của phổ tín hiệu âm thanh có thể, ví dụ, được biểu diễn trong miền biến đổi cosin rời rạc cải biên.

Theo phương án khác, các mẫu tín hiệu âm thanh của phổ tín hiệu âm thanh có thể, ví dụ, được biểu diễn trong miền biến đổi sin rời rạc cải biên.

Theo phương án khác, bộ xử lý có thể, ví dụ, được tạo cấu hình để tạo ra phổ được thay đổi bằng cách sử dụng hàm ký hiệu ngẫu nhiên mà ngẫu nhiên hoặc giả ngẫu nhiên phát ra hoặc giá trị thứ nhất hoặc giá trị thứ hai.

Trong phương án, bộ xử lý có thể, ví dụ, được tạo cấu hình để tăng giảm cường độ phổ được thay đổi đến phổ đích bằng cách sau đó giảm hệ số suy giảm.

Theo phương án, bộ xử lý có thể, ví dụ, được tạo cấu hình để tăng giảm cường độ phổ được thay đổi đến phổ đích bằng cách sau đó tăng hệ số suy giảm.

Trong phương án, nếu khung hiện thời không được nhận bởi giao diện nhận hoặc nếu khung hiện thời được nhận bởi giao diện nhận bị sai lạc, bộ xử lý có thể, ví dụ, được tạo cấu hình để tạo ra tín hiệu âm thanh được khôi phục bằng cách sử dụng công

thức:

$$x[i] = (1 - \text{cum\_damping}) * \text{noise}[i] + \text{cum\_damping} * \text{random\_sign}() * x_{\text{old}}[i]$$

trong đó  $i$  là chỉ số, trong đó  $x[i]$  biểu thị mẫu của tín hiệu âm thanh được khôi phục, trong đó  $\text{cum\_damping}$  là hệ số suy giảm, trong đó  $x_{\text{old}}[i]$  biểu thị một trong số các mẫu tín hiệu âm thanh của phổ tín hiệu âm thanh của tín hiệu âm thanh được mã hóa, trong đó  $\text{random\_sign}()$  quay về 1 hoặc -1, và trong đó  $\text{noise}$  là vectơ ngẫu nhiên biểu thị phổ đích.

Trong phương án, vectơ ngẫu nhiên đã nêu  $\text{noise}$  có thể, ví dụ, được định tỉ lệ sao cho trung bình bình phương tương tự với trung bình bình phương của phổ của tín hiệu âm thanh được mã hóa được bao gồm một trong số các khung được nhận cuối cùng bởi giao diện nhận.

Theo phương án chung, bộ xử lý có thể, ví dụ, được tạo cấu hình để tạo ra tín hiệu âm thanh được khôi phục, bằng cách sử dụng vectơ ngẫu nhiên mà được định tỉ lệ sao cho trung bình bình phương của nó tương tự với trung bình bình phương của phổ của tín hiệu âm thanh được mã hóa được bao gồm một trong số các khung được nhận cuối cùng bởi giao diện nhận.

Hơn nữa, phương pháp giải mã tín hiệu âm thanh được mã hóa để thu được tín hiệu âm thanh được khôi phục được đề xuất. Phương pháp bao gồm:

- Nhận một hoặc nhiều khung bao gồm thông tin trên nhiều mẫu tín hiệu âm thanh của phổ tín hiệu âm thanh của tín hiệu âm thanh được mã hóa. Và:
- Tạo ra tín hiệu âm thanh được khôi phục.

Tạo ra tín hiệu âm thanh được khôi phục được thực hiện bằng cách tăng giảm cường độ phổ được thay đổi đến phổ đích, nếu khung hiện thời không được nhận hoặc nếu khung hiện thời được nhận nhưng bị sai lạc, trong đó phổ được thay đổi bao gồm nhiều mẫu phổ tín hiệu được thay đổi, trong đó, đối với mỗi mẫu tín hiệu được thay đổi của phổ được thay đổi, giá trị tuyệt đối của mẫu tín hiệu được thay đổi đã nêu bằng với giá trị tuyệt đối của một trong số các mẫu tín hiệu âm thanh của phổ tín hiệu âm thanh. Phổ được thay đổi không được tăng giảm cường độ đến phổ nhiễu âm trắng,

nếu khung hiện thời trong số một hoặc nhiều khung được nhận và nếu khung hiện thời được nhận không bị sai lạc.

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

Các phương án nhận ra sự tăng giảm cường độ phổ MDCT đến nhiễu âm trắng trước ứng dụng thay thế nhiễu âm miền tần số (Frequency Domain Noise Substitution - FDNS).

Theo tình trạng kỹ thuật, trong các bộ mã hóa-giải mã trên cơ sở ACELP, bảng mã đổi mới được thay thế bằng vectơ ngẫu nhiên (ví dụ, bằng nhiễu âm). Trong các phương án, phương pháp ACELP, mà bao gồm việc thay thế bảng mã đổi mới bằng vectơ ngẫu nhiên (ví dụ, bằng nhiễu âm) được thừa nhận cho cấu trúc bộ giải mã TCX. Ở đây, sự tương đương của bảng mã đổi mới là phổ MDCT thường được nhận trong dòng bit và được đưa vào trong FDNS.

Phương pháp che giấu MDCT cổ điển sẽ để lại một cách đơn giản phổ này như là hoặc để áp dụng quy trình ngẫu nhiên hóa nhất định, mà về cơ bản kéo dài hình dạng phổ của khung được nhận cuối cùng [LS01]. Điều này có hạn chế rằng hình dạng phổ ngắn hạn được kéo dài, thường dẫn đến âm thanh lặp đi lặp lại, kim loại mà không giống tạp âm nền, và do đó không thể được sử dụng như nhiễu âm nền nhân tạo.

Việc sử dụng phương pháp được đề xuất là tạo hình phổ ngắn hạn được thực hiện bởi FDNS và TCX LTP, việc tạo hình phổ về lâu dài được thực hiện chỉ bởi FDNS. Việc tạo hình bởi FDNS được tăng giảm cường độ từ hình dạng phổ ngắn hạn đến hình dạng phổ dài hạn được tạo vết, và TCX LTP được tăng giảm cường độ về 0.

Việc tăng giảm cường độ các hệ số FDNS đến các hệ số tạp âm nền được tạo vết dẫn đến việc có sự chuyển tiếp trôi chảy giữa đường bao phổ tốt cuối cùng và đường bao nền phổ mà nên được nhắm đến về lâu dài, để đạt được tạp âm nền vừa ý trong trường hợp mất khung khối tín hiệu truyền dài.

Ngược lại, theo tình trạng kỹ thuật, đối với các bộ mã hóa biến đổi trên cơ sở biến đổi, nhiễu âm như sự che giấu được thực hiện bằng sự lặp lại khung hoặc sự thay thế nhiễu âm trong miền tần số [LS01]. Trong tình trạng kỹ thuật, sự thay thế nhiễu âm

thường được thực hiện bằng cách xáo trộn ký hiệu của các ngăn phổ. Nếu trong tình trạng kỹ thuật sự xáo trộn ký hiệu TCX (miền tần số) được sử dụng trong quá trình che giấu, các hệ số MDCT được nhận cuối cùng được tái sử dụng và mỗi ký hiệu được ngẫu nhiên hóa trước phổ được biến đổi một cách nghịch đảo về miền thời gian. Hạn chế của quy trình này của tình trạng kỹ thuật là, rằng đối với các khung liên tiếp bị mất, phổ tương tự được sử dụng lặp đi lặp lại với các sự ngẫu nhiên hóa ký hiệu khác nhau và sự suy giảm toàn cục. Khi tìm đường bao phổ theo thời gian trên lưới thời gian sơ bộ, có thể thấy rằng đường bao là xấp xỉ không thay đổi trong quá trình mất khung liên tiếp, bởi vì các năng lượng băng được giữ không thay đổi so với nhau trong khung và vừa được suy giảm một cách toàn cục. Trong hệ thống mã hóa được sử dụng, theo tình trạng kỹ thuật, các giá trị phổ được xử lý sử dụng FDNS, để khôi phục phổ ban đầu. Điều này có nghĩa là, rằng nếu một người muốn tăng giảm cường độ phổ MDCT đến đường bao phổ nhất định (sử dụng các hệ số FDNS, ví dụ, mô tả tạp âm nền hiện thời), kết quả không chỉ phụ thuộc vào các hệ số FDNS, mà còn phụ thuộc vào phổ được mã hóa trước đó mà bị xáo trộn ký hiệu. Các phương án nêu trên khắc phục các bất lợi của tình trạng kỹ thuật.

Các phương án được dựa trên việc tìm ra rằng cần thiết để tăng giảm cường độ phổ được sử dụng cho việc làm xáo trộn ký hiệu đến nhiều âm trắng, trước khi đưa nó vào sự xử lý FDNS. Mặt khác, phổ được phát ra sẽ không bao giờ kết hợp với đường bao được nhắm mục tiêu được sử dụng để xử lý FDNS.

Trong các phương án, tốc độ tăng giảm cường độ giống nhau được sử dụng cho sự tăng giảm cường độ độ khuếch đại LTP như cho sự tăng giảm cường độ nhiễu âm trắng.

### **Mô tả vắn tắt các hình vẽ**

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

Fig.1a mô tả thiết bị để giải mã tín hiệu âm thanh theo phương án,

Fig.1b mô tả thiết bị để giải mã tín hiệu âm thanh theo phương án khác,

Fig.1c mô tả thiết bị để giải mã tín hiệu âm thanh theo phương án khác, trong đó

thiết bị còn bao gồm bộ phận gộp thứ nhất và bộ phận gộp thứ hai,

Fig.1d minh họa thiết bị để giải mã tín hiệu âm thanh theo phương án khác, trong đó thiết bị còn bao gồm bộ dự báo dài hạn bao gồm vùng đệm độ trễ,

Fig.2 minh họa cấu trúc bộ giải mã của G.718,

Fig.3 mô tả cảnh, trong đó thừa số giảm dần cường độ của G.722, phụ thuộc vào thông tin loại,

Fig.4 thể hiện phương pháp cho sự dự báo biên độ sử dụng sự hồi quy tuyến tính

Fig.5 minh họa trạng thái mất khối tín hiệu truyền của phép biến đổi được lắp năng lượng miễn cưỡng (Constrained-Energy Lapped Transform - CELT),

Fig. 6 thể hiện sự tạo vết tạp âm nền theo phương án trong bộ giải mã trong suốt chế độ vận hành không lỗi,

Fig.7 minh họa đạo hàm khuếch đại của sự chỉnh giảm và sự tổng hợp LPC theo sáng chế,

Fig.8 mô tả ứng dụng mức nhiễu âm nền nhân tạo trong quá trình mất gói theo phương án,

Fig.9 mô tả bộ bù khuếch đại thông cao trong quá trình che giấu ACELP theo phương án,

Fig.10 minh họa sự tách của các vòng lặp phản hồi LTP trong quá trình che giấu theo phương án,

Fig.11 mô tả thiết bị để giải mã tín hiệu âm thanh được mã hóa để thu được tín hiệu âm thanh được khôi phục theo phương án,

Fig. 12 thể hiện thiết bị để giải mã tín hiệu âm thanh được mã hóa để thu được tín hiệu âm thanh được tái tạo theo phương án, và

Fig. 13 minh họa thiết bị để giải mã tín hiệu âm thanh được mã hóa để thu được tín hiệu âm thanh được tái tạo theo phương án khác, và

Fig. 14 minh họa thiết bị để giải mã tín hiệu âm thanh được mã hóa để thu được tín hiệu âm thanh được tái tạo theo phương án khác.

### **Mô tả chi tiết sáng chế**

Fig.1a mô tả thiết bị để giải mã tín hiệu âm thanh theo phương án.

Thiết bị bao gồm giao diện nhận 110. Giao diện nhận được tạo cấu hình để nhận nhiều khung, trong đó giao diện nhận 110 được tạo cấu hình để nhận khung thứ nhất trong số nhiều khung, khung thứ nhất đã nêu bao gồm phần tín hiệu âm thanh thứ nhất, phần tín hiệu âm thanh thứ nhất đã nêu được biểu diễn trong miền thứ nhất. Ngoài ra, giao diện nhận 110 được tạo cấu hình để nhận khung thứ hai trong số nhiều khung, khung thứ hai đã nêu bao gồm phần tín hiệu âm thanh thứ hai của tín hiệu âm thanh.

Ngoài ra, thiết bị bao gồm bộ phận biến đổi 120 để biến đổi phần tín hiệu âm thanh thứ hai hoặc giá trị hoặc tín hiệu lấy được từ phần tín hiệu âm thanh thứ hai từ miền thứ hai đến miền tạo vết để thu được phần thông tin phụ tín hiệu thứ hai, trong đó miền thứ hai khác với miền thứ nhất, trong đó miền tạo vết khác với miền thứ hai, và trong đó miền tạo vết bằng hoặc khác với miền thứ nhất.

Ngoài ra, thiết bị bao gồm bộ phận tạo vết mức nhiễu âm 130, trong đó bộ phận tạo vết mức nhiễu âm được tạo cấu hình để nhận thông tin phần tín hiệu thứ nhất được biểu diễn trong miền tạo vết, trong đó thông tin phần tín hiệu thứ nhất phụ thuộc vào phần tín hiệu âm thanh thứ nhất, trong đó bộ phận tạo vết mức nhiễu âm được tạo cấu hình để nhận phần tín hiệu thứ hai được biểu diễn trong miền tạo vết, và trong đó bộ phận tạo vết mức nhiễu âm được tạo cấu hình để xác định thông tin mức nhiễu âm phụ thuộc vào thông tin phần tín hiệu thứ nhất được biểu diễn trong miền tạo vết và phụ thuộc vào thông tin phần tín hiệu thứ hai được biểu diễn trong miền tạo vết.

Hơn nữa, thiết bị bao gồm bộ phận khôi phục để khôi phục phần tín hiệu âm thanh thứ ba của tín hiệu âm thanh phụ thuộc vào thông tin mức nhiễu âm, nếu khung thứ ba trong số nhiều khung không được nhận bằng giao diện nhận mà bị sai lạc.

Đối với phần tín hiệu thứ nhất và/hoặc thứ hai, ví dụ, phần tín hiệu thứ nhất và/hoặc thứ hai có thể, ví dụ, được nạp vào một hoặc nhiều bộ phận xử lý (không được thể hiện) để tạo ra một hoặc nhiều tín hiệu loa phát thanh cho một hoặc nhiều loa phát thanh, vì vậy thông tin âm thanh được nhận bao gồm bởi phần tín hiệu âm thanh thứ nhất và/hoặc thứ hai có thể được phát lại.

Ngoài ra, tuy nhiên, phần tín hiệu âm thanh thứ nhất và thứ hai cũng được sử dụng cho việc che giấu, ví dụ, trong trường hợp các khung tiếp theo không tới tại bộ nhận hoặc trong trường hợp các khung tiếp theo là không chính xác.

Không kể những cái khác, sáng chế dựa trên việc tìm kiếm mà việc tạo vết mức nhiều âm nên được tiến hành trong miền chung, ở đây gọi là "miền tạo vết". Miền tạo vết, có thể, ví dụ, là miền kích thích, ví dụ, miền mà trong đó tín hiệu được biểu diễn bằng các hệ số tiên đoán tuyến tính (Linear Predictive Coefficient - LPC) hoặc bằng các cặp phổ mạch nạp ISP (Immittance Spectral Pair - ISP) như được mô tả trong AMR-WB và AMR-WB+ (xem [3GP12a], [3GP12b], [3GP09a], [3GP09b], [3GP09c]). Việc tạo vết mức nhiều âm trong miền đơn có hiệu quả không kể tới những cái khác rằng hiệu ứng sai số lấy mẫu được ngăn ngừa khi tín hiệu chuyển giữa sự biểu diễn thứ nhất trong miền thứ nhất và sự biểu diễn thứ hai trong miền thứ hai (ví dụ, khi sự biểu diễn tín hiệu chuyển từ ACELP đến TCX hoặc ngược lại).

Liên quan đến bộ phận biến đổi 120, cái được biến đổi hoặc là bản thân phần tín hiệu âm thanh thứ hai, hoặc tín hiệu lấy từ phần tín hiệu âm thanh thứ hai (ví dụ, phần tín hiệu âm thanh thứ hai được xử lý để thu được tín hiệu được lấy), hoặc giá trị lấy từ phần tín hiệu âm thanh thứ hai (ví dụ, phần tín hiệu âm thanh thứ hai đã được xử lý để thu được giá trị được lấy).

Liên quan đến phần tín hiệu âm thanh thứ nhất, trong một số phương án, phần tín hiệu âm thanh thứ nhất có thể được xử lý và/hoặc biến đổi thành miền tạo vết.

Tuy nhiên, trong các phương án khác, phần tín hiệu âm thanh thứ nhất có thể đã được biểu diễn trong miền tạo vết.

Trong một số phương án, thông tin phần tín hiệu thứ nhất đồng nhất với phần tín hiệu âm thanh thứ nhất. Trong các phương án khác, thông tin phần tín hiệu thứ nhất là, ví dụ, giá trị được gộp phụ thuộc vào phần tín hiệu âm thanh thứ nhất.

Ngay bây giờ, đầu tiên, việc giảm dần cường độ đến mức nhiều âm nên nhân tạo được xem xét chi tiết hơn.

Phương pháp giảm dần cường độ được mô tả có thể, ví dụ, được thực hiện trong phiên bản độ trễ thấp của xHE-AAC [NMR+12] (xHE-AAC = hiệu suất cao được mở

rộng AAC), mà có khả năng chuyển đổi liền mạch giữa ACELP (tiếng nói) và MDCT (âm nhạc/ nhiều âm) mã hóa trên cơ sở mỗi khung.

Liên quan đến việc tạo vết mức thông thường trong miền tạo vết, ví dụ, miền kích thích, như để áp dụng sự giảm dần cường độ làm nhẵn lên mức nhiều âm nền nhân tạo thích hợp trong sự mất gói, như mức độ nhiều âm nền nhân tạo cần được nhận ra trong suốt quá trình xử lý giải mã bình thường. Nó có thể, ví dụ, được giả sử, rằng mức nhiều âm tương tự như tạp âm nền là dễ chịu nhất. Do đó, mức tạp âm nền có thể lấy được và cập nhật liên tục trong quá trình giải mã thông thường.

Sáng chế dựa trên việc tìm kiếm rằng khi có bộ mã hóa-giải mã lỗi được chuyển đổi (ví dụ, ACELP và TCX), xem xét mức tạp âm nền thông thường độc lập khỏi bộ mã hóa-giải mã được chọn là đặc biệt phù hợp.

Fig. 6 mô tả sự tạo vết tạp âm nền theo phương án được ưu tiên trong bộ giải mã trong suốt chế độ vận hành không lỗi, ví dụ, trong quá trình giải mã bình thường.

Việc tạo vết bản thân nó có thể, ví dụ, được thực hiện sử dụng phương pháp thống kê tối thiểu (xem [Mar01]).

Mức tạp âm nền được tạo vết này có thể, ví dụ, được xem xét như là thông tin mức nhiều âm được nói tới ở trên.

Ví dụ, việc ước lượng nhiều âm thống kê tối thiểu được trình bày trong tài liệu: “Rainer Martin, *Noise power spectral density estimation based on optimal smoothing and minimum statistics*, IEEE Transactions on Speech and Audio Processing 9 (2001), no. 5, 504–512” [Mar01] có thể được dùng cho việc tạo vết mức tạp âm nền.

Do đó, trong một số phương án, bộ phận tạo vết mức nhiều âm 130 được tạo cấu hình để xác định thông tin mức nhiều âm bằng cách áp dụng sự ước lượng nhiều âm thống kê tối thiểu của [Mar01].

Tiếp theo, một số cân nhắc và các chi tiết của phương pháp tạo vết này được mô tả.

Liên quan đến sự tạo vết mức, nền được coi là giống nhiều âm. Do đó, ưu tiên hơn để thực hiện việc tạo vết mức trong miền kích thích để tránh tạo vết các thành phần âm nền phía trước mà được đưa ra ngoài bằng LPC. Ví dụ, sự làm đầy nhiều âm

ACELP có thể cũng dùng mức nhiễu âm nền trong miền kích thích. Với sự tạo vết trong miền kích thích, chỉ một sự tạo vết đơn của mức tạp âm nền có thể phục vụ hai mục đích, mà có thể tiết kiệm độ phức tạp tính toán. Trong phương án được ưu tiên, sự tạo vết được thực hiện trong miền kích thích ACELP.

Fig.7 minh họa đạo hàm khuếch đại của sự tổng hợp và sự chỉnh giảm LPC theo sáng chế.

Liên quan đến đạo hàm mức, đạo hàm mức có thể, ví dụ, được tiến hành trong miền thời gian hoặc trong miền kích thích, hoặc trong bất kỳ miền thích hợp khác. Nếu các miền cho đạo hàm mức và sự tạo vết mức khác nhau, sự bù khuếch đại có thể, ví dụ, cần thiết.

Trong phương án được ưu tiên, đạo hàm mức cho ACELP được thực hiện trong miền kích thích. Do đó, không đòi hỏi sự bù khuếch đại.

Đối với TCX, sự bù khuếch đại có thể, ví dụ, cần thiết để điều chỉnh mức được lấy ra đến miền kích thích ACELP.

Trong phương án được ưu tiên, đạo hàm mức cho TCX diễn ra trong miền thời gian. Sự bù khuếch đại có thể kiểm soát được tìm thấy trong phương án này: Độ khuếch đại được dẫn bằng sự tổng hợp và sự chỉnh giảm LPC được lấy ra như được thể hiện trên Fig.7 và mức được lấy được chia cho độ khuếch đại.

Ngoài ra, đạo hàm mức cho TCX có thể được thực hiện trong miền kích thích TCX. Tuy nhiên, sự bù khuếch đại giữa miền kích thích TCX và miền kích thích ACELP được cho là quá phức tạp.

Do đó, quay trở lại Fig.1a, trong một số phương án, phần tín hiệu âm thanh thứ nhất được biểu diễn trong miền thời gian như miền thứ nhất Bộ phận biến đổi 120 được tạo cấu hình để biến đổi phần tín hiệu âm thanh thứ hai hoặc giá trị được lấy từ phần tín hiệu âm thanh thứ hai từ miền kích thích là miền thứ hai đến miền thời gian là miền tạo vết. Trong các phương án này, bộ phận tạo vết mức nhiễu âm 130 được tạo cấu hình để nhận thông tin phần tín hiệu thứ nhất được biểu diễn trong miền thời gian như miền tạo vết. Ngoài ra, bộ phận tạo vết mức nhiễu âm 130 được tạo cấu hình để nhận phần tín hiệu thứ hai được biểu diễn trong miền thời gian như miền tạo vết.

Trong các phương án khác, phần tín hiệu âm thanh thứ nhất được biểu diễn trong miền kích thích như miền thứ nhất. Bộ phận biến đổi 120 được tạo cấu hình để biến đổi phần tín hiệu âm thanh thứ hai hoặc giá trị được lấy từ phần tín hiệu âm thanh thứ hai từ miền thời gian là miền thứ hai đến miền kích thích là miền tạo vết. Trong các phương án này, bộ phận tạo vết mức nhiễu âm 130 được tạo cấu hình để nhận thông tin phần tín hiệu thứ nhất được biểu diễn trong miền thời gian như miền tạo vết. Hơn nữa, bộ phận tạo vết mức nhiễu âm 130 được tạo cấu hình để nhận phần tín hiệu thứ hai được biểu diễn trong miền kích thích như miền tạo vết.

Trong phương án, phần tín hiệu âm thanh thứ nhất có thể, ví dụ, được biểu diễn trong miền kích thích như miền thứ nhất, trong đó bộ phận tạo vết mức nhiễu âm 130 có thể, ví dụ, được tạo cấu hình để nhận thông tin phần tín hiệu thứ nhất trong đó thông tin phần tín hiệu thứ nhất đã nêu được biểu diễn trong miền FFT, là miền tạo vết, và trong đó thông tin phần tín hiệu thứ nhất đã nêu phụ thuộc vào phần tín hiệu âm thanh thứ nhất đã nêu được biểu diễn trong miền kích thích, trong đó bộ phận biến đổi 120 có thể, ví dụ, được tạo cấu hình để biến đổi phần tín hiệu âm thanh thứ hai hoặc giá trị được lấy từ phần tín hiệu âm thanh thứ hai từ miền thời gian là miền thứ hai đến miền FFT là miền tạo vết, và trong đó bộ phận tạo vết mức nhiễu âm 130 có thể, ví dụ, được tạo cấu hình để nhận phần tín hiệu âm thanh thứ hai được biểu diễn trong miền FFT.

Fig. 1b mô tả thiết bị theo như phương án khác. Trên Fig.1b, bộ phận biến đổi 120 của Fig.1a là bộ phận biến đổi thứ nhất và bộ phận khôi phục 140 của Fig.1a là bộ phận khôi phục thứ nhất 140. Thiết bị còn bao gồm bộ phận biến đổi thứ hai 121 và bộ phận khôi phục thứ hai 141.

Bộ phận biến đổi thứ hai 121 được tạo cấu hình để biến đổi thông tin mức nhiễu âm từ miền tạo vết đến miền thứ hai, nếu khung thứ tư trong số nhiều khung không được nhận bằng giao diện nhận hoặc khung thứ tư đã nêu được nhận bằng giao diện nhận nhưng bị sai lạc.

Hơn nữa, bộ phận khôi phục thứ hai 141 được tạo cấu hình để khôi phục phần tín hiệu âm thanh thứ tư của tín hiệu âm thanh phụ thuộc vào thông tin mức nhiễu âm được biểu diễn trong miền thứ hai nếu khung thứ tư đã nêu trong số nhiều khung

không được nhận bằng giao diện nhận hoặc nếu khung thứ tư đã nêu được nhận bằng giao diện nhận nhưng bị sai lạc.

Fig.1c mô tả thiết bị để giải mã tín hiệu âm thanh theo phương án khác, Thiết bị còn bao gồm bộ phận gộp thứ nhất 150 để xác định giá trị được gộp thứ nhất phụ thuộc vào phần tín hiệu âm thanh thứ nhất. Ngoài ra, thiết bị trên Fig.1c còn bao gồm bộ phận gộp 160 để xác định giá trị được gộp như giá trị lấy được từ phần tín hiệu âm thanh thứ hai phụ thuộc phần vào phần tín hiệu âm thanh thứ hai. Trong phương án trên Fig.1c, bộ phận tạo vết mức nhiều âm 130 được tạo cấu hình để nhận giá trị được gộp thứ nhất như thông tin phần tín hiệu thứ nhất được biểu diễn trong miền tạo vết, trong đó bộ phận tạo vết mức nhiều âm 130 được tạo cấu hình để nhận giá trị được gộp thứ hai như thông tin phần tín hiệu thứ hai được biểu diễn trong miền tạo vết. Bộ phận tạo vết mức nhiều âm 130 được tạo cấu hình để xác định thông tin mức nhiều âm phụ thuộc vào giá trị được gộp thứ nhất được biểu diễn trong miền tạo vết và phụ thuộc vào giá trị được gộp thứ hai được biểu diễn trong miền tạo vết.

Theo phương án, bộ phận gộp thứ nhất 150 được tạo cấu hình để xác định giá trị được gộp thứ nhất sao cho giá trị được gộp thứ nhất biểu thị căn bậc hai trung bình của phần tín hiệu âm thanh thứ nhất hoặc của tín hiệu được lấy từ phần tín hiệu âm thanh thứ nhất. Hơn nữa, bộ phận gộp thứ hai 160 được tạo cấu hình để xác định giá trị được gộp thứ hai sao cho giá trị được gộp thứ hai biểu thị căn bậc hai trung bình của phần tín hiệu âm thanh thứ hai hoặc của tín hiệu được lấy từ phần tín hiệu âm thanh thứ hai.

Fig.6 minh họa thiết bị để giải mã tín hiệu âm thanh theo phương án khác.

Trên Fig.6, bộ phận tạo vết mức nền 630 thực thi bộ phận tạo vết mức nhiều âm 130 theo Fig.1a.

Ngoài ra, trên Fig.6, bộ phận căn bậc hai trung bình 650 (root mean square - RMS) là bộ phận gộp thứ nhất và bộ phận RMS 660 là bộ phận gộp thứ hai.

Theo một số phương án, bộ phận biến đổi (thứ nhất) 120 của Fig.1a, Fig.1b và Fig.1c được tạo cấu hình để biến đổi giá trị lấy được từ phần tín hiệu âm thanh thứ hai từ miền thứ hai đến miền tạo vết bằng cách áp dụng giá trị khuếch đại (x) trên giá trị được lấy từ phần tín hiệu âm thanh thứ hai, ví dụ, bằng cách chia giá trị lấy được từ

phần tín hiệu âm thanh thứ hai cho giá trị khuếch đại ( $x$ ). Trong các phương án khác, giá trị khuếch đại có thể, ví dụ, được nhân lên.

Trong các phương án khác, giá trị khuếch đại ( $x$ ) có thể, ví dụ, biểu thị sự khuếch đại được đưa vào bởi sự tổng hợp mã hóa dự báo tuyến tính, hoặc giá trị khuếch đại ( $x$ ) có thể, ví dụ, biểu thị sự khuếch đại bằng sự tổng hợp và sự chỉnh giảm mã hóa dự báo tuyến tính.

Trên Fig.6, bộ phận 622 cung cấp giá trị ( $x$ ) biểu thị độ khuếch đại được dẫn bởi sự tổng hợp và sự chỉnh giảm mã hóa dự báo tuyến tính. Bộ phận 622 sau đó chia giá trị, được cung cấp bởi bộ phận gộp thứ hai 660, mà là giá trị được lấy từ phần tín hiệu âm thanh thứ hai, cho giá trị khuếch đại được cung cấp ( $x$ ) (ví dụ, hoặc bằng cách chia cho  $x$ , hoặc nhân giá trị  $1/x$ ). Do đó, bộ phận 620 của Fig.6 mà bao gồm các bộ phận 621 và 622 thực thi bộ phận biến đổi thứ nhất của Fig.1a, Fig.1b hoặc Fig.1c.

Thiết bị của Fig.6 nhận khung thứ nhất với phần tín hiệu âm thanh thứ nhất là sự kích thích hữu thanh và/hoặc sự kích thích vô thanh và được biểu diễn trong miền tạo vết, trong Fig.6 miền (ACELP) LPC. Phần tín hiệu âm thanh thứ nhất được nạp vào bộ phận tổng hợp và chỉnh giảm LPC 671 cho sự xử lý để thu được đầu ra phần tín hiệu âm thanh thứ nhất miền thời gian. Ngoài ra, phần tín hiệu âm thanh thứ nhất được nạp vào môđun RMS 650 để thu được giá trị thứ nhất biểu thị căn bậc hai trung bình của phần tín hiệu âm thanh thứ nhất. Giá trị thứ nhất này (giá trị RMS thứ nhất) được biểu diễn trong miền tạo vết. Giá trị RMS thứ nhất, được biểu diễn trong miền tạo vết, sau đó được nạp vào bộ phận tạo vết mức nhiễu âm 630.

Ngoài ra, thiết bị của Fig.6 nhận khung thứ hai với phần tín hiệu âm thanh thứ hai bao gồm phổ MDCT và được biểu diễn trong miền MDCT. Sự lấp đầy âm được tiến hành bằng môđun lấp đầy âm 681, sự tạo hình nhiễu âm miền tần số được tiến hành bằng môđun tạo hình nhiễu âm miền tần số 682, sự biến đổi đến miền thời gian được tiến hành bằng môđun iMDCT/OLA 683 (OLA = chồng lấp-cộng) và dự báo dài hạn được thực hiện bằng bộ phận dự báo dài hạn 684. Bộ phận dự báo dài hạn có thể, ví dụ, bao gồm vùng đệm độ trễ (không được thể hiện trên Fig.6).

Tín hiệu được lấy từ phần tín hiệu âm thanh thứ hai sau đó được nạp vào môđun RMS 660 để thu giá trị thứ hai biểu thị căn bậc hai trung bình của tín hiệu được lấy từ

phần tín hiệu âm thanh thứ hai được thu. Giá trị thứ hai này (giá trị RMS thứ hai) được biểu diễn trong miền thời gian. Bộ phận 620 sau đó biến đổi giá trị RMS thứ hai từ miền thời gian đến miền tạo vết, ở đây, miền LPC (ACELP). Giá trị RMS thứ hai, được biểu diễn trong miền tạo vết, sau đó được nạp vào bộ phận tạo vết mức nhiễu âm 630.

Theo một số phương án, sự tạo vết mức được thực hiện trong miền kích thích, nhưng sự giảm dần cường độ TCX được thực hiện trong miền thời gian.

Trong khi đó trong quá trình giải mã thông thường, mức tạp âm nền được tạo vết, nó có thể, ví dụ, được sử dụng trong sự mất gói như là bộ chỉ thị của mức nhiễu âm nền nhân tạo thích hợp mà tín hiệu được nhận cuối cùng được tăng giảm cường độ theo mức một cách trôi chảy

Sự lấy mức để tạo vết và sự áp dụng lên sự giảm dần cường độ mức là độc lập với nhau và có thể được thực hiện trong các miền khác nhau. Trong phương án được ưu tiên, việc áp dụng mức được thực hiện trong những miền giống nhau như đạo hàm mức, dẫn đến các lợi ích giống nhau cho ACELP, sự bù khuếch đại là không cần thiết, và cho TCX, sự bù khuếch đại nghịch đảo cho đạo hàm mức (xem Fig.6) là cần thiết và do đó đạo hàm khuếch đại giống như vậy có thể được sử dụng, như được minh họa bằng Fig.7

Tiếp theo, sự bù tác động của bộ lọc thông cao trên độ khuếch đại tổng hợp LPC theo các phương án được mô tả.

Fig.8 tóm tắt phương pháp này. Cụ thể, Fig.8 minh họa sự áp dụng mức nhiễu âm nền nhân tạo trong quá trình mất gói.

Trên Fig.8, bộ phận bộ lọc khuếch đại thông cao 643, bộ phận nhân 644, bộ phận tăng giảm cường độ 645, bộ phận bộ lọc thông cao 646, bộ phận tăng giảm cường độ 647 và bộ phận tổ hợp 648 cùng với nhau tạo ra bộ phận khôi phục thứ nhất.

Ngoài ra, trên Fig.8, bộ phận cung cấp mức nền 631 cung cấp thông tin mức nhiễu âm. Ví dụ, bộ phận cung cấp mức nền 631 có thể được thực hiện như bộ phận tạo vết mức nền 630 của Fig.6.

Ngoài ra, trên Fig.8, bộ phận khuếch đại tổng hợp và chỉnh giảm LPC 649 và bộ

phận nhân 641 cùng nhau cho bộ phận biến đổi thứ hai 640.

Ngoài ra, trên Fig.8, bộ phận tăng giảm cường độ 642 biểu diễn bộ phận khôi phục thứ hai.

Theo phương án của Fig.8, sự kích thích hữu thanh và vô thanh được tăng giảm cường độ riêng biệt: Sự kích thích hữu thanh được tăng giảm cường độ về 0, nhưng sự kích thích vô thanh được tăng giảm cường độ về mức nhiễu âm nền nhân tạo. Fig.8 còn mô tả bộ lọc thông cao, mà được dẫn vào chuỗi tín hiệu của sự kích thích vô thanh để triệt thành phần tần số thấp cho tất cả các trường hợp ngoại trừ tín hiệu được phân loại là vô thanh.

Để mô hình hóa ảnh hưởng của bộ lọc thông cao, mức sau khi tổng hợp và chỉnh giảm LPC được tính toán một lần khi có và một lần khi không có bộ lọc thông cao. Sau đó tỉ lệ của hai mức này được lấy ra và sử dụng để thay đổi mức nền được áp dụng.

Điều này được minh họa bằng Fig.9. Cụ thể là, Fig.9 mô tả bộ bù khuếch đại thông cao trong quá trình che giấu theo phương án.

Thay vì tín hiệu kích thích hiện tại, chỉ xung đơn được sử dụng làm đầu vào cho tính toán này. Điều này cho phép giảm độ phức tạp, do xung lượng phản hồi phân rã một cách nhanh chóng và do đạo hàm RMS có thể thực hiện trên khung thời gian ngắn hơn. Thực tế, chỉ một khung con được sử dụng thay vì toàn bộ khung.

Theo phương án, bộ phận tạo vết mức nhiễu âm 130 được tạo cấu hình để xác định mức nhiễu âm nền nhân tạo như thông tin mức nhiễu âm. Bộ phận khôi phục 140 được tạo cấu hình để khôi phục phần tín hiệu âm thanh thứ ba phụ thuộc vào thông tin mức nhiễu âm, nếu khung thứ ba đã nêu trong số nhiều khung không được nhận bởi giao diện nhận 110 hoặc khung thứ ba đã nêu được nhận bởi giao diện nhận 110 nhưng bị sai lạc.

Theo phương án, bộ phận tạo vết mức nhiễu âm 130 được tạo cấu hình để xác định mức nhiễu âm nền nhân tạo như thông tin mức nhiễu âm. Bộ phận khôi phục 140 được tạo cấu hình để khôi phục phần tín hiệu âm thanh thứ ba phụ thuộc vào thông tin mức nhiễu âm, nếu khung thứ ba đã nêu trong số nhiều khung không được nhận bởi

giao diện nhận 110 hoặc khung thứ ba đã nêu được nhận bởi giao diện nhận 110 nhưng bị sai lạc.

Trong phương án, bộ phận tạo vết mức nhiều âm 130 được tạo cấu hình để xác định mức nhiều âm nền nhân tạo như thông tin mức nhiều âm được lấy từ phổ mức nhiều âm, trong đó phổ mức nhiều âm thu được nhờ áp dụng phương pháp thống kê tối thiểu. Bộ phận khôi phục 140 được tạo cấu hình để khôi phục phần tín hiệu âm thanh thứ ba phụ thuộc vào nhiều hệ số dự báo tuyến tính, nếu khung thứ ba đã nêu trong số nhiều khung không được nhận bởi giao diện nhận 110 hoặc khung thứ ba đã nêu được nhận bởi giao diện nhận 110 nhưng bị sai lạc.

Trong phương án, bộ phận khôi phục (thứ nhất và/hoặc thứ hai) 140, 141 có thể, ví dụ, được tạo cấu hình để khôi phục phần tín hiệu âm thanh thứ ba phụ thuộc vào thông tin mức nhiều âm và phụ thuộc vào phần tín hiệu âm thanh thứ nhất, nếu khung thứ ba (thứ tư) đã nêu trong số nhiều khung không được nhận bởi giao diện nhận 110 hoặc nếu khung thứ ba (thứ tư) đã nêu được nhận bởi giao diện nhận 110 nhưng bị sai lạc.

Theo phương án, bộ phận khôi phục (thứ nhất và/hoặc thứ hai) 140, 141 có thể, ví dụ, được tạo cấu hình để khôi phục phần tín hiệu âm thanh thứ ba (hoặc thứ tư) bằng cách giảm nhẹ hoặc tăng cường phần tín hiệu âm thanh thứ nhất.

Fig.14 mô tả thiết bị để giải mã tín hiệu âm thanh. Thiết bị bao gồm giao diện nhận 110, trong đó giao diện nhận 110 được tạo cấu hình để nhận khung thứ nhất bao gồm phần tín hiệu âm thanh thứ nhất của tín hiệu âm thanh, và trong đó giao diện nhận 110 được tạo cấu hình để nhận khung thứ hai bao gồm phần tín hiệu âm thanh thứ hai của tín hiệu âm thanh.

Ngoài ra, thiết bị bao gồm bộ phận tạo vết mức nhiều âm 130, trong đó bộ phận tạo vết mức nhiều âm 130 được tạo cấu hình để xác định thông tin mức nhiều âm phụ thuộc vào ít nhất một trong số phần tín hiệu âm thanh thứ nhất và phần tín hiệu âm thanh thứ hai (nghĩa là: phụ thuộc vào phần tín hiệu âm thanh thứ nhất và/hoặc phần tín hiệu âm thanh thứ hai), trong đó thông tin mức nhiều âm được biểu diễn trong miền tạo vết.

Ngoài ra, thiết bị còn bao gồm bộ phận khôi phục thứ nhất 140 cho việc khôi phục, trong miền khôi phục thứ nhất, phần tín hiệu âm thanh thứ ba của tín hiệu âm thanh phụ thuộc vào thông tin mức nhiễu âm, nếu khung thứ ba trong số nhiều khung không được nhận bằng giao diện nhận 110 hoặc nếu khung thứ ba đã nêu được nhận bằng giao diện nhận 110 nhưng bị sai lệch, trong đó miền khôi phục thứ nhất khác hoặc bằng miền tạo vết.

Ngoài ra, thiết bị còn bao gồm bộ phận biến đổi 121 cho việc biến đổi thông tin mức nhiễu âm từ miền tạo vết tới miền khôi phục thứ hai, nếu khung thứ tư trong số nhiều khung không được nhận bằng giao diện nhận 110 hoặc nếu khung thứ tư đã nêu được nhận bằng giao diện nhận 110 nhưng bị sai lệch, trong đó miền khôi phục thứ hai khác với miền tạo vết, và trong đó miền khôi phục thứ hai khác với miền khôi phục thứ nhất, và

Ngoài ra, thiết bị bao gồm bộ phận khôi phục thứ hai 141 cho việc khôi phục, trong miền khôi phục thứ hai, phần tín hiệu âm thanh thứ tư của tín hiệu âm thanh phụ thuộc vào thông tin mức nhiễu âm được biểu diễn trong miền khôi phục thứ hai, nếu khung thứ tư đã nêu trong số nhiều khung không được nhận bằng giao diện nhận 110 hoặc nếu khung thứ tư đã nêu được nhận bằng giao diện nhận 110 nhưng bị sai lệch.

Theo một số phương án, miền tạo vết có thể, ví dụ, trong đó miền tạo vết là miền thời gian, miền phổ, miền FFT, miền MDCT, hoặc miền kích thích. Miền khôi phục thứ nhất có thể, ví dụ, là miền thời gian, miền phổ, miền FFT, miền MDCT, hoặc miền kích thích. Miền khôi phục thứ hai có thể, ví dụ, là miền thời gian, miền phổ, miền FFT, miền MDCT, hoặc miền kích thích.

Trong phương án, miền tạo vết có thể, ví dụ, là miền FFT, miền khôi phục thứ nhất có thể, ví dụ, là miền thời gian, và miền khôi phục thứ hai có thể, ví dụ, là miền kích thích.

Trong phương án khác, miền tạo vết có thể, ví dụ, là miền thời gian, miền khôi phục thứ nhất có thể, ví dụ, là miền thời gian, và miền khôi phục thứ hai có thể, ví dụ, là miền kích thích.

Theo phương án, phần tín hiệu âm thanh thứ nhất đã nêu có thể, ví dụ, được biểu

diễn trong miền đầu vào thứ nhất, và phần tín hiệu âm thanh thứ hai có thể, ví dụ, được biểu diễn trong miền đầu vào thứ hai. Bộ phận biến đổi có thể, ví dụ, là bộ phận biến đổi thứ hai. Thiết bị có thể, ví dụ, còn bao gồm bộ phận biến đổi thứ nhất để biến đổi phần tín hiệu âm thanh thứ hai hoặc giá trị hoặc tín hiệu được lấy từ phần tín hiệu âm thanh thứ hai từ miền đầu vào thứ hai đến miền tạo vết để thu được thông tin phần tín hiệu thứ hai. Bộ phận tạo vết mức nhiều âm có thể, ví dụ, được tạo cấu hình để nhận thông tin phần tín hiệu thứ nhất được biểu diễn trong miền tạo vết, trong đó thông tin phần tín hiệu thứ nhất phụ thuộc vào phần tín hiệu âm thanh thứ nhất, trong đó bộ phận tạo vết mức nhiều âm được tạo cấu hình để nhận phần tín hiệu thứ hai được biểu diễn trong miền tạo vết, và trong đó bộ phận tạo vết mức nhiều âm được tạo cấu hình để xác định thông tin mức nhiều âm phụ thuộc vào thông tin phần tín hiệu thứ nhất được biểu diễn trong miền tạo vết và phụ thuộc vào thông tin phần tín hiệu thứ hai được biểu diễn trong miền tạo vết.

Theo phương án, miền đầu vào thứ nhất có thể, ví dụ, là miền kích thích, và miền đầu vào thứ hai có thể, ví dụ, là miền MDCT.

Trong phương án khác, miền đầu vào thứ nhất có thể, ví dụ, là miền MDCT, và trong đó miền đầu vào thứ hai có thể, ví dụ, là miền MDCT.

Nếu, ví dụ, tín hiệu được biểu diễn trong miền thời gian, nó có thể, ví dụ, được biểu diễn bằng các mẫu miền thời gian của tín hiệu. Hoặc, ví dụ, nếu tín hiệu được biểu diễn trong miền phổ, nó có thể, ví dụ, được biểu diễn bằng các mẫu phổ của phổ của tín hiệu.

Trong phương án, miền tạo vết có thể, ví dụ, là miền FFT, miền khôi phục thứ nhất có thể, ví dụ, là miền thời gian, và miền khôi phục thứ hai có thể, ví dụ, là miền kích thích.

Trong phương án khác, miền tạo vết có thể, ví dụ, là miền thời gian, miền khôi phục thứ nhất có thể, ví dụ, là miền thời gian, và miền khôi phục thứ hai có thể, ví dụ, là miền kích thích.

Trong một số phương án, các bộ phận được minh họa trên Fig.14 có thể, ví dụ, được tạo cấu hình như được miêu tả cho Fig.1a, Fig.1b, Fig.1c và Fig.1d.

Liên quan đến các phương án cụ thể, trong, ví dụ, chế độ tốc độ thấp, thiết bị theo phương án có thể, ví dụ, nhận các khung ACELP làm đầu vào, mà được biểu diễn trong miền kích thích, và mà sau đó được biến đổi đến miền thời gian qua sự tổng hợp LPC. Ngoài ra, ở chế độ tốc độ thấp, thiết bị theo phương án có thể, ví dụ, nhận các khung TCX như là đầu vào, mà được biểu diễn trong miền MDCT, và sau đó biến đổi thành miền thời gian qua MDCT nghịch đảo.

Sự tạo vết sau đó được dẫn trong miền FFT, trong đó tín hiệu FFT lấy được từ tín hiệu miền thời gian bằng cách tiến hành phép biến đổi Fourier nhanh (Fast Fourier Transform - FFT). Sự tạo vết có thể, ví dụ, được tiến hành bằng cách tiến hành phương pháp thống kê tối thiểu, phân tách cho tất cả các đường phổ được để thu phổ nhiễu âm nền nhân tạo.

Sự che giấu sau đó được tiến hành bằng việc tiến hành đạo hàm mức dựa trên phổ nhiễu âm nền nhân tạo. Đạo hàm mức được tiến hành dựa trên phổ nhiễu âm nền nhân tạo. Sự chuyển đổi mức thành miền thời gian được tiến hành cho FD, TCX, PLC. Sự tăng giảm cường độ trong miền thời gian được tiến hành. Đạo hàm mức thành miền kích thích được tiến hành cho ACELP, PLC và cho TD TCX PLC (tương tự ACELP). Sự tăng giảm cường độ trong miền kích thích được tiến hành.

Danh sách dưới đây tóm tắt tất điều này:

tốc độ thấp:

- đầu vào:

- acelp (miền kích thích -> miền thời gian, qua sự tổng hợp LPC)
- tcx (miền mdct -> miền thời gian, qua MDCT nghịch đảo)

- sự tạo vết:

- miền fft, được lấy từ miền thời gian qua FFT
- thống kê tối thiểu, phân tách tất cả các đường phổ -> phổ nhiễu âm nền nhân tạo

nhân tạo

- sự che giấu:

- đạo hàm mức được dựa trên phổ nhiễu âm nền nhân tạo

- sự chuyển đổi mức thành miền thời gian cho
  - FD TCX PLC
    - > sự giảm cường độ trong miền thời gian
- sự chuyển đổi mức thành miền kích thích cho
  - ACELP PLC
  - TD TCX PLC (tương tự ACELP)
    - > sự tăng giảm cường độ trong miền kích thích.

Ví dụ, trong chế độ tốc độ cao, có thể, ví dụ, nhận các khung TCX làm đầu vào, mà được biểu diễn trong miền MDCT, và sau đó biến đổi thành miền thời gian qua MDCT nghịch đảo.

Sự tạo vết sau đó được tiến hành trong miền thời gian. Sự tạo vết có thể, ví dụ, được tiến hành bằng cách tiến hành phương pháp thống kê tối thiểu dựa trên mức năng lượng để thu mức nhiễu âm nền nhân tạo.

Đối với sự che giấu, cho FD TCX PLC, mức có thể được sử dụng như hiện tại và chỉ có sự tăng giảm cường độ trong miền thời gian có thể được tiến hành. Đối với TD TCX PLC (tương tự ACELP), sự chuyển đổi mức thành miền kích thích và sự tăng giảm cường độ trong miền kích thích được tiến hành.

Danh sách dưới đây tóm tắt điều này:

tốc độ cao:

• đầu vào:

- tcx (miền mdct -> miền thời gian, qua MDCT nghịch đảo)

• sự tạo vết:

- miền thời gian
- thống kê tối thiểu trên mức năng lượng -> mức nhiễu âm nền nhân tạo.

• sự che giấu:

- sử dụng mức "như là"
  - FD TCX PLC

-> sự tăng giảm cường độ trong miền thời gian

- sự chuyển đổi mức thành miền kích thích cho

- TD TCX PLC (tương tự ACELP)

-> sự tăng giảm cường độ trong miền kích thích

Miền FFT và miền MDCT đều là các miền phổ, trong đó miền kích thích là loại miền thời gian.

Theo phương án, bộ phận khôi phục thứ nhất 140 có thể, ví dụ, được tạo cấu hình để khôi phục phần tín hiệu âm thanh thứ ba bằng việc tiến hành sự tăng giảm cường độ thứ nhất thành phổ như nhiều âm. Bộ phận khôi phục thứ hai 141 có thể, ví dụ, được tạo cấu hình để khôi phục phần tín hiệu âm thanh thứ tư bằng việc tiến hành sự tăng giảm cường độ thành phổ như nhiều âm và/hoặc sự tăng giảm cường độ của độ khuếch đại LTP. Ngoài ra, bộ phận khôi phục thứ nhất 140 và bộ phận khôi phục thứ hai 141 có thể, ví dụ, được tạo cấu hình để tiến hành sự tăng giảm cường độ thứ nhất và sự tăng giảm cường độ thứ hai thành phổ như nhiều âm và/hoặc sự tăng giảm cường độ thứ hai của độ khuếch đại LTP với cùng tốc độ tăng giảm cường độ.

Bây giờ sự tạo hình phổ thích ứng của nhiều âm nền nhân tạo được xem xét.

Để đạt được tạo hình thích ứng để giảm nhiều âm nền nhân tạo trong quá trình mất gói khôi tín hiệu truyền, như bước thứ nhất, tìm kiếm các hệ số LPC thích hợp mà biểu diễn tạp âm nền có thể được tiến hành. Các hệ số LPC có thể được lấy trong quá trình tiếng nói hoạt động sử dụng phương pháp thống kê tối thiểu để tìm phổ tạp âm nền và sau đó tính toán các hệ số LPC từ nó bằng cách sử dụng thuật toán tùy ý cho đạo hàm LPC đã biết từ tài liệu. Một số phương án, ví dụ, có thể chuyển đổi một cách trực tiếp phổ tạp âm nền thành sự biểu diễn mà có thể được sử dụng trực tiếp cho FDNS trong miền MDCT.

Sự tăng giảm cường độ thành nhiều âm nền nhân tạo có thể được thực hiện trong miền ISF (cũng có thể áp dụng được trong miền LSF; tần số phổ đường LSF):

$$f_{current}[i] = \alpha \cdot f_{last}[i] + (1 - \alpha) \cdot pt_{mean}[i] \quad i = 0...16 \quad (26)$$

bằng việc thiết lập  $pt_{mean}$  thành các hệ số LP thích hợp mô tả nhiều âm nền nhân

tạo.

Liên quan tới việc tạo hình phổ thích ứng của nhiễu âm nền nhân tạo, phương án chung hơn được minh họa bằng Fig.11.

Fig.11 mô tả thiết bị để giải mã tín hiệu âm thanh được mã hóa để thu được tín hiệu âm thanh được khôi phục theo phương án.

Thiết bị bao gồm giao diện nhận 1110 để nhận một hoặc nhiều khung, bộ tạo hệ số 1120, và bộ khôi phục tín hiệu 1130.

Bộ tạo hệ số 1120 được tạo cấu hình để xác định, nếu khung hiện thời trong số một hoặc nhiều khung được nhận bằng giao diện nhận 1110 và nếu khung hiện thời được nhận bằng giao diện nhận 1110 mà không bị sai lạc/lỗi, một hoặc nhiều hệ số tín hiệu âm thanh thứ nhất, được bao gồm bởi khung hiện thời, trong đó một hoặc nhiều hệ số tín hiệu âm thanh thứ nhất đã nêu chỉ ra đặc điểm của tín hiệu âm thanh được mã hóa, và một hoặc nhiều nhiều các hệ số biểu thị tập âm nền của tín hiệu âm thanh được mã hóa. Ngoài ra, bộ tạo hệ số 1120 được tạo cấu hình để tạo ra một hoặc nhiều hệ số tín hiệu âm thanh thứ hai, phụ thuộc vào một hoặc nhiều hệ số tín hiệu âm thanh thứ nhất và phụ thuộc vào một hoặc nhiều hệ số nhiễu âm, nếu khung hiện thời không được nhận bằng giao diện nhận 1110 hoặc nếu khung hiện thời được nhận bằng giao diện nhận 1110 bị sai lạc/lỗi.

Bộ khôi phục tín hiệu âm thanh 1130 được tạo cấu hình để khôi phục phần thứ nhất của tín hiệu âm thanh được khôi phục phụ thuộc vào một hoặc nhiều hệ số tín hiệu âm thanh thứ nhất, nếu khung hiện thời được nhận bằng giao diện nhận 1110 và nếu khung hiện thời được nhận bằng giao diện nhận 1110 mà không bị sai lạc. Ngoài ra, bộ khôi phục tín hiệu âm thanh 1130 được tạo cấu hình để khôi phục phần thứ hai của tín hiệu âm thanh được khôi phục phụ thuộc vào một hoặc nhiều hệ số tín hiệu âm thanh thứ hai, nếu khung hiện thời không được nhận bằng giao diện nhận 1110 và nếu khung hiện thời được nhận bằng giao diện nhận 1110 mà bị sai lạc.

Việc xác định tập âm nền được biết đến trong tình trạng kỹ thuật (xem, ví dụ [Mar01]: Rainer Martin, *Noise power spectral density estimation based on optimal smoothing and minimum statistics*, IEEE Transactions on Speech and Audio

Processing 9 (2001), no. 5, 504 –512), và trong phương án, thiết bị xử lý theo đó.

Trong một số phương án, một hoặc nhiều hệ số tín hiệu âm thanh thứ nhất có thể, ví dụ, là một hoặc nhiều hệ số bộ lọc dự báo tuyến tính của tín hiệu âm thanh được mã hóa. Trong một số phương án, một hoặc nhiều hệ số tín hiệu âm thanh thứ nhất có thể, ví dụ, là một hoặc nhiều hệ số bộ lọc dự báo tuyến tính của tín hiệu âm thanh được mã hóa.

Trong lĩnh vực kỹ thuật, người ta đã biết đến rộng rãi cách làm thế nào để khôi phục tín hiệu âm thanh, ví dụ, tín hiệu tiếng nói, từ bộ lọc dự báo tuyến tính hoặc từ các cặp phổ dẫn (xem, ví dụ, [3GP09c]: *Speech codec speech processing functions; adaptive multi-rate - wideband (AMRWB) speech codec; transcoding functions*, 3GPP TS 26.190, 3rd Generation Partnership Project, 2009), và trong phương án, bộ khôi phục tín hiệu xử lý tương ứng.

Theo phương án, một hoặc nhiều hệ số nhiễu âm có thể, ví dụ, là một hoặc nhiều hệ số bộ lọc dự báo tuyến tính biểu thị tạp âm nền của tín hiệu âm thanh được mã hóa. Trong phương án, một hoặc nhiều hệ số bộ lọc dự báo tuyến tính có thể, ví dụ, biểu diễn hình dạng phổ của tạp âm nền.

Theo phương án, bộ tạo hệ số 1120 có thể, ví dụ, được tạo cấu hình để xác định một hoặc nhiều phần tín hiệu âm thanh thứ hai sao cho một hoặc nhiều phần tín hiệu âm thanh thứ hai là một hoặc nhiều hệ số dự báo tuyến tính của tín hiệu âm thanh được khôi phục, hoặc sao cho một hoặc nhiều hệ số tín hiệu âm thanh thứ nhất là một hoặc nhiều cặp phổ dẫn của tín hiệu âm thanh được khôi phục.

Theo phương án, bộ tạo hệ số 1120 có thể, ví dụ, được tạo cấu hình để tạo một hoặc nhiều hệ số tín hiệu âm thanh thứ hai bằng việc áp dụng công thức:

$$f_{current}[i] = \alpha \cdot f_{last}[i] + (1 - \alpha) \cdot pt_{mean}[i]$$

trong đó  $f_{current}[i]$  biểu thị một trong số một hoặc nhiều hệ số tín hiệu âm thanh thứ hai, trong đó  $f_{last}[i]$  biểu thị một trong số một hoặc nhiều hệ số tín hiệu âm thanh thứ nhất, trong đó  $pt_{mean}[i]$  là một trong số một hoặc nhiều hệ số nhiễu âm, trong đó  $\alpha$  là số thực với  $0 \leq \alpha \leq 1$ , và trong đó  $i$  là chỉ số.

Theo phương án,  $f_{last}[i]$  biểu thị hệ số bộ lọc dự báo tuyến tính của tín hiệu âm

thanh được mã hóa, và trong đó  $f_{current}[i]$  biểu thị hệ số bộ lọc dự báo tuyến tính của tín hiệu âm thanh được khôi phục.

Trong phương án,  $pt_{mean}[i]$  có thể, ví dụ, biểu thị tạp âm nền của tín hiệu âm thanh được mã hóa.

Theo phương án, bộ tạo hệ số 1120 có thể, ví dụ, được tạo cấu hình để tạo ít nhất các hệ số tín hiệu âm thanh 10 giây như một hoặc nhiều hệ số tín hiệu âm thanh thứ hai.

Theo phương án, bộ tạo hệ số 1120 có thể, ví dụ, được tạo cấu hình để xác định, nếu khung hiện thời trong số một hoặc nhiều khung được nhận bằng giao diện nhận 1110 và nếu khung hiện thời được nhận bằng giao diện nhận 1110 không bị sai lệch, một hoặc nhiều hệ số nhiễu âm bằng việc xác định phổ nhiễu âm của tín hiệu âm thanh được mã hóa.

Sau đây, sự giảm cường độ phổ MDCT thành nhiễu âm trắng trước ứng dụng FDNS được xem xét.

Thay vì thay đổi một cách ngẫu nhiên dấu hiệu của ngăn MDCT (xáo trộn dấu hiệu), phổ hoàn thiện được lấp đầy bằng nhiễu âm trắng, được tạo hình sử dụng FDNS. Để tránh việc thay đổi tức thì về các đặc điểm phổ, sự tăng giảm chéo giữa sự xáo trộn dấu hiệu và sự lấp đầy nhiễu âm được áp dụng. Sự tăng giảm chéo có thể thực hiện như sau:

```
for(i=0; i<L_frame; i++) {
    if (old_x[i] != 0) {
        x[i] = (1 - cum_damping)*noise[i] + cum_damping
* random_sign() * x_old[i];
    }
}
```

trong đó:

$cum\_damping$  là hệ số suy giảm (tuyệt đối)- nó mô tả sự suy giảm từ khung đến khung, bắt đầu từ 1 và giảm xuống 0.

$x_{old}$  là phổ của khung được nhận cuối cùng.

$random\_sign$  trở về 1 hoặc -1

$noise$  chứa véctor ngẫu nhiên (nhiều âm trắng) mà được chia tỉ lệ mà trung bình bình phương của nó (RMS) tương tự phổ tốt cuối cùng.

Thuật ngữ  $random\_sign() * old\_x[i]$  đặc trưng cho quá trình xáo trộn tín hiệu để ngẫu nhiên hóa các pha và tránh sự lặp lại sóng hài.

Sau đây, sự chuẩn hóa khác của mức năng lượng có thể được thực hiện sau sự tăng giảm chéo để đảm bảo rằng việc tính tổng năng lượng không lệch do sự tương quan của hai véctor.

Theo các phương án, bộ khôi phục thứ nhất 140 có thể, ví dụ, được tạo cấu hình để khôi phục phần tín hiệu âm thanh thứ ba phụ thuộc vào thông tin mức nhiễu âm và phụ thuộc vào phần tín hiệu âm thanh thứ nhất. Theo phương án cụ thể, bộ phận khôi phục thứ nhất 140 có thể, ví dụ, được tạo cấu hình để khôi phục phần tín hiệu âm thanh thứ ba bằng cách giảm nhẹ hoặc tăng cường phần tín hiệu âm thanh thứ nhất.

Theo các phương án, bộ phận khôi phục thứ hai 141 có thể, ví dụ, được tạo cấu hình để khôi phục phần tín hiệu âm thanh thứ tư phụ thuộc vào thông tin mức nhiễu âm và phụ thuộc vào phần tín hiệu âm thanh thứ hai. Theo phương án cụ thể, bộ phận khôi phục thứ hai 141 có thể, ví dụ, được tạo cấu hình để khôi phục phần tín hiệu âm thanh thứ tư bằng việc giảm nhẹ hoặc tăng cường phần tín hiệu âm thanh thứ hai.

Liên quan tới sự tăng giảm cường độ phổ MDCT được mô tả ở trên đến nhiễu âm trắng trước sự ứng dụng FDNS, phương án chung hơn được minh họa bằng Fig.12.

Fig.12 minh họa thiết bị để giải mã tín hiệu âm thanh được mã hóa để thu được tín hiệu âm thanh được tái tạo theo phương án.

Thiết bị bao gồm giao diện nhận 1210 để nhận một hoặc nhiều khung bao gồm thông tin về nhiều mẫu tín hiệu âm thanh của phổ tín hiệu âm thanh của tín hiệu âm thanh được mã hóa, và bộ xử lý 1220 để tạo ra tín hiệu âm thanh được khôi phục.

Bộ xử lý 1220 được tạo cấu hình để tạo ra tín hiệu âm thanh được khôi phục bằng cách tăng giảm cường độ phổ được thay đổi thành phổ đích, nếu khung hiện thời

không được nhận bởi giao diện nhận 1210 hoặc nếu khung hiện thời được nhận bằng giao diện nhận 1210 nhưng bị sai lạc, trong đó, phổ được thay đổi bao gồm nhiều mẫu tín hiệu được thay đổi, trong đó, với mỗi mẫu tín hiệu được thay đổi của phổ được điều chỉnh, giá trị tuyệt đối của mẫu tín hiệu được thay đổi đã nói bằng giá trị tuyệt đối của một trong các mẫu tín hiệu âm thanh của phổ tín hiệu âm thanh.

Hơn nữa, bộ xử lý 1220 được tạo cấu hình để không làm giảm cường độ phổ được điều chỉnh đến phổ đích, nếu khung hiện thời của một hoặc nhiều khung được nhận bởi giao diện nhận 1210 và nếu khung hiện tới được nhận bởi giao diện nhận 1210 không bị sai lạc.

Theo phương án, phổ đích là phổ như nhiều âm.

Trong phương án, phổ như nhiều âm biểu diễn nhiều âm trắng.

Theo phương án, phổ đích được tạo hình.

Theo phương án, hình dạng của phổ như nhiều âm phụ thuộc vào phổ tín hiệu âm thanh của tín hiệu âm thanh được nhận trước đó.

Theo phương án, phổ như nhiều âm được tạo hình phụ thuộc vào hình dạng của phổ tín hiệu âm thanh.

Theo phương án, bộ xử lý 1220 sử dụng hệ số nghiêng để tạo hình phổ như nhiều âm.

Theo phương án, bộ xử lý 1220 sử dụng công thức

$$\text{shaped\_noise}[i] = \text{noise} * \text{power}(\text{tilt\_factor}, i/N)$$

trong đó  $N$  biểu thị số lượng mẫu,

trong đó  $i$  là chỉ số,

trong đó  $0 \leq i < N$ , với  $\text{tilt\_factor} > 0$ ,

trong đó  $\text{power}$  là hàm mũ.

Nếu  $\text{tilt\_factor}$  là nhỏ hơn 1 điều này có nghĩa là sự suy giảm với  $i$  tăng. Nếu  $\text{tilt\_factor}$  là lớn hơn 1 có nghĩa là sự khuếch đại với  $i$  tăng.

Theo phương án khác, bộ xử lý 1220 có thể sử dụng công thức

```
shaped_noise[i] = noise * (1 + i / (N-1) *
(tilt_factor-1))
```

trong đó N biểu thị số lượng mẫu,

trong đó i là chỉ số, trong đó  $0 \leq i < N$ ,

với  $\text{tilt\_factor} > 0$ .

Theo phương án, bộ xử lý 1220 được tạo cấu hình để tạo ra phổ được điều chỉnh, bằng cách thay đổi dấu của một hoặc nhiều mẫu trong số các mẫu tín hiệu âm thanh của phổ tín hiệu âm thanh, nếu khung hiện thời không được nhận bởi giao diện nhận 1210 hoặc nếu khung hiện thời được nhận bởi giao diện nhận 1210 bị sai lạc.

Theo phương án, mỗi mẫu tín hiệu âm thanh của phổ tín hiệu âm thanh được biểu diễn bởi số thực mà không phải bởi số ảo.

Theo phương án, các mẫu tín hiệu âm thanh của phổ tín hiệu âm thanh được biểu diễn trong miền biến đổi cosin rời rạc cải biên.

Theo phương án khác, các mẫu tín hiệu âm thanh của phổ tín hiệu âm thanh được biểu diễn trong miền biến đổi sin rời rạc cải biên.

Theo phương án, bộ xử lý 1220 được tạo cấu hình để tạo ra phổ được điều chỉnh bằng cách sử dụng hàm dấu ngẫu nhiên mà ngẫu nhiên hoặc giả ngẫu nhiên phát ra hoặc giá trị thứ nhất hoặc giá trị thứ hai.

Trong phương án, bộ xử lý 1220 được tạo cấu hình để giảm cường độ phổ được điều chỉnh đến phổ đích bằng cách sau đó giảm hệ số suy giảm.

Theo phương án, bộ xử lý 1220 được tạo cấu hình để tăng giảm cường độ phổ được thay đổi đến phổ đích bằng cách sau đó tăng hệ số suy giảm.

Trong phương án, nếu khung hiện thời không được nhận bởi giao diện nhận 1210 hoặc nếu khung hiện thời được nhận bởi giao diện nhận 1210 bị sai lạc, bộ xử lý 1220, được tạo cấu hình để tạo ra tín hiệu âm thanh được khôi phục bằng cách sử dụng công thức:

```
x[i] = (1-cum_damping) * noise[i] + cum_damping *
random_sign() * x_old[i]
```

trong đó  $i$  là chỉ số, trong đó  $x[i]$  biểu thị mẫu của tín hiệu âm thanh được khôi phục, trong đó  $\text{cum\_damping}$  là hệ số suy giảm, trong đó  $x\_old[i]$  biểu thị một trong số các mẫu tín hiệu âm thanh của phổ tín hiệu âm thanh của tín hiệu âm thanh được mã hóa, trong đó  $\text{random\_sign}()$  trở về 1 hoặc -1, và trong đó  $\text{noise}$  là véc tơ ngẫu nhiên biểu thị phổ đích.

Một số phương án tiếp tục thao tác TCX LTP. Trong các phương án này, thao tác TCX, LTP được tiếp tục trong quá trình che giấu với các tham số LTP (độ trễ LTP và độ khuếch đại LTP) được lấy từ khung tốt cuối cùng.

Các thao tác LTP có thể được tóm tắt là:

- Nạp vùng đệm trễ độ trễ LTP dựa trên đầu ra được lấy ra trước đó.
- Dựa trên độ trễ LTP: chọn phần tín hiệu thích hợp khỏi vùng đệm độ trễ LTP mà được sử dụng như phần góp LTP để tạo hình tín hiệu hiện thời.
- Thay đổi tỉ lệ phần góp LTP sử dụng độ khuếch đại LTP.
- Bổ sung phần góp LTP được thay đổi tỉ lệ vào tín hiệu đầu vào LTP để tạo tín hiệu đầu ra LTP.

Các phương pháp khác nhau có thể được cân nhắc với sự liên quan đến thời gian, khi sự cập nhật vùng đệm độ trễ được thực hiện.

Như thao tác LTP thứ nhất trong khung  $n$  sử dụng đầu ra từ khung cuối  $n-1$ . Các cập nhật vùng đệm độ trễ LTP này trong khung  $n$  được sử dụng trong quá trình xử lý LTP trong khung  $n$ .

Như vận hành LTP cuối cùng trong khung  $n$  sử dụng đầu ra khung hiện thời  $n$ . Các cập nhật vùng đệm LTP trong khung  $n$  được sử dụng trong quá trình xử lý LTP trong khung  $n + 1$ .

Trong phần tiếp theo, việc tách các vòng lặp phản hồi TCX LTP được xem xét.

Việc tách vòng lặp phản hồi TCX LTP tránh việc đưa ra các nhiễu âm bổ sung (do sự thay thế nhiễu âm được áp dụng cho tín hiệu đầu ra LPT) trong suốt mỗi vòng lặp phản hồi của bộ giải mã LTP khi trong chế độ che giấu.

Fig.10 minh họa việc tách này. Cụ thể, Fig.10 mô tả sự tách của các vòng lặp

phản hồi LTP trong quá trình che giấu ( $bfi=1$ ).

Fig.10 minh họa vùng đệm độ trễ 1020, bộ lựa chọn mẫu 1030, và bộ xử lý mẫu 1040 (bộ xử lý mẫu 1040 được chỉ bằng cách đường nét đứt).

Đối với thời gian, khi sự cập nhật vùng đệm độ trễ LTP 1022 được thực hiện, một số phương án được tiến hành như sau:

- Đối với thao tác thông thường: Cập nhật vùng đệm độ trễ LTP 1020 như sự vận hành LTP thứ nhất có thể được ưu tiên, do tín hiệu đầu ra được cộng thường được lưu trữ liên tục. Với phương pháp này, vùng đệm chuyên dụng có thể được bỏ qua.

- Đối với thao tác được tách: Để cập nhật vùng đệm độ trễ 1020 như sự vận hành LTP cuối cùng có thể được ưu tiên, do phần góp LTP với tín hiệu thường chỉ được lưu trữ tạm thời. Với phương pháp này, tín hiệu góp LTP tạm thời được bảo toàn. Sự thực hiện khôn ngoan vùng đệm phần góp LTP chỉ có thể được thực hiện liên tục.

Giả sử rằng phương pháp sau được sử dụng trong bất kỳ trường hợp (vận hành bình thường và che giấu), các phương án, có thể, ví dụ, thực hiện như sau:

- Trong quá trình thao tác thông thường: Đầu ra của tín hiệu miền thời gian của bộ giải mã LTP sau khi bổ sung vào tín hiệu đầu vào LTP được sử dụng để nạp vùng đệm độ trễ LTP.

- Trong quá trình che giấu: Đầu ra của tín hiệu miền thời gian của bộ giải mã LTP trước khi bổ sung vào tín hiệu đầu vào LTP được sử dụng để nạp vùng đệm độ trễ LTP.

Một số phương án tăng giảm cường độ độ khuếch đại TCX LTP về 0. Trong phương án này, độ khuếch đại TCX LTP có thể, ví dụ, được tăng giảm cường độ đến 0 với hệ số giảm dần cường độ thích ứng tín hiệu. Điều này có thể, ví dụ, được làm lặp lại, ví dụ, theo dòng mã giả sau:

```
gain = gain_past * damping;
```

```
[...]
```

```
gain_past = gain;
```

trong đó:

gain là bộ giải mã khuếch đại TCX LTP được áp dụng trên khung hiện thời;

gain\_past là bộ giải mã khuếch đại TCX LTP được áp dụng vào khung trước đó;

damping là hệ số giảm dần cường độ (tương ứng).

Fig. 1d miêu tả thiết bị theo phương án khác, trong đó thiết bị còn bao gồm bộ phận dự báo dài hạn 170 bao gồm vùng đệm độ trễ 180. Bộ phận dự báo dài hạn 170 được tạo cấu hình để tạo ra tín hiệu được xử lý phụ thuộc vào phần tín hiệu âm thanh thứ hai, phụ thuộc vào đầu vào vùng đệm độ trễ được dự trữ trong vùng đệm độ trễ 180 và phụ thuộc vào độ khuếch đại dự báo dài hạn. Hơn nữa, bộ phận dự báo dài hạn có thể, ví dụ, được tạo cấu hình để giảm cường độ độ khuếch đại dự báo dài hạn về 0, nếu khung thứ ba đã nêu trong số nhiều khung không được nhận bởi giao diện nhận hoặc khung thứ ba đã nêu được nhận bởi giao diện nhận nhưng bị sai lạc.

Trong các phương án (không được thể hiện), bộ phận dự báo dài hạn có thể, ví dụ, còn được tạo cấu hình để tạo ra tín hiệu được xử lý phụ thuộc vào phần tín hiệu âm thanh thứ nhất, phụ thuộc vào vùng đệm độ trễ và phụ thuộc vào độ khuếch đại dự báo dài hạn.

Trên Fig.1d, bộ khôi phục thứ nhất 140 có thể, ví dụ, tạo ra phần tín hiệu âm thanh thứ ba ngoài ra phụ thuộc vào tín hiệu được xử lý.

Trong phương án, bộ phận dự báo dài hạn 170 có thể, ví dụ, được tạo cấu hình để giảm cường độ khuếch đại dài hạn về 0, trong đó tốc độ mà với tốc độ này độ khuếch đại dự báo dài hạn được tăng giảm cường độ về 0 phụ thuộc vào thừa số giảm dần cường độ.

Thay thế hoặc bổ sung, bộ phận dự báo dài hạn 170 có thể, ví dụ, được tạo cấu hình để cập nhật đầu vào vùng đệm trễ 180 bằng cách lưu trữ tín hiệu được xử lý được tạo ra trong vùng đệm độ trễ 180 nếu khung thứ ba đã nêu trong số nhiều khung không được nhận bởi giao diện nhận 110 hoặc khung thứ ba đã nêu được nhận bởi giao diện nhận 110 nhưng bị sai lạc.

Liên quan tới việc sử dụng TCX LTP miêu tả ở trên, sáng chế thông thường hơn được minh họa bằng Fig.13

Fig.13 minh họa thiết bị để giải mã tín hiệu âm thanh được mã hóa để thu được tín hiệu âm thanh được khôi phục.

Thiết bị bao gồm giao diện nhận 1310 để nhận nhiều khung, vùng đệm trễ 1320 để lưu trữ các mẫu tín hiệu âm thanh của tín hiệu âm thanh được giải mã, bộ chọn lọc mẫu 1330 để chọn lọc nhiều mẫu tín hiệu âm thanh được lựa chọn từ các mẫu tín hiệu âm thanh được lưu trữ trong vùng đệm trễ 1320, và bộ xử lý mẫu 1340 để xử lý các mẫu tín hiệu âm thanh được lựa chọn để thu được các mẫu tín hiệu âm thanh được khôi phục của tín hiệu âm thanh được khôi phục.

Bộ lựa chọn mẫu 1330 được tạo cấu hình để lựa chọn, nếu khung hiện thời được nhận bởi giao diện nhận 1310 và nếu khung hiện thời được nhận bởi giao diện nhận 1310 không bị sai lệch, nhiều mẫu tín hiệu âm thanh được lựa chọn từ các mẫu tín hiệu âm thanh được lưu trữ trong vùng đệm trễ 1320 phụ thuộc vào thông tin độ trễ được bao gồm bởi khung hiện thời. Hơn nữa, bộ lựa chọn mẫu 1320 được tạo cấu hình để lựa chọn, nếu khung hiện thời không được nhận bởi giao diện nhận 1310 hoặc nếu khung hiện thời được nhận bởi giao diện nhận 1310 bị sai lệch, nhiều mẫu tín hiệu âm thanh được lựa chọn, từ các mẫu tín hiệu âm thanh được lưu trữ trong vùng đệm độ trễ 1320 phụ thuộc vào thông tin độ trễ bước được bao gồm trong khung khác được nhận trước đó bởi giao diện nhận 1310.

Theo phương án, bộ xử lý mẫu 1340 có thể, ví dụ, được tạo cấu hình để thu được các mẫu tín hiệu âm thanh được khôi phục, nếu khung hiện thời được nhận bởi giao diện nhận 1310 và nếu khung hiện thời được nhận bởi giao diện nhận 1310 không bị sai lệch, bằng cách định tỉ lệ lại các mẫu tín hiệu âm thanh được lựa chọn phụ thuộc vào thông tin độ khuếch đại được bao gồm bởi khung hiện thời. Hơn nữa, bộ lựa chọn mẫu 1330 có thể, ví dụ, được tạo cấu hình để thu được các mẫu tín hiệu âm thanh được khôi phục, nếu khung hiện thời không được nhận bởi giao diện nhận 1310 hoặc nếu khung hiện thời được nhận bởi giao diện nhận 1310 bị sai lệch, bằng cách định tỉ lệ lại các mẫu tín hiệu âm thanh được lựa chọn phụ thuộc vào thông tin độ khuếch đại được bao gồm bởi khung khác đã nêu được nhận trước đó bởi giao diện nhận 1310.

Trong phương án, bộ xử lý mẫu 1340 có thể, ví dụ, được tạo cấu hình để thu được các mẫu tín hiệu âm thanh được khôi phục, nếu khung hiện thời được nhận bởi

giao diện nhận 1310 và nếu khung hiện thời được nhận bởi giao diện nhận 1310 không bị sai lạc, bằng cách nhân các mẫu tín hiệu âm thanh được lựa chọn và giá trị phụ thuộc vào thông tin độ khuếch đại được bao gồm bởi khung hiện thời. Hơn nữa, bộ lựa chọn mẫu 1330 được tạo cấu hình để thu được các mẫu tín hiệu âm thanh được khôi phục, nếu khung hiện thời không được nhận bởi giao diện nhận 1310 hoặc nếu khung hiện thời được nhận bởi giao diện nhận 1310 bị sai lạc, bằng cách nhân các mẫu tín hiệu âm thanh được lựa chọn phụ thuộc vào thông tin độ khuếch đại được bao gồm bởi khung khác đã nêu được nhận trước đó bởi giao diện nhận 1310.

Theo phương án, bộ xử lý mẫu 1340 có thể, ví dụ, được tạo cấu hình để lưu trữ các mẫu tín hiệu âm thanh được khôi phục vào trong vùng đệm độ trễ 1320.

Trong phương án, bộ xử lý mẫu 1340 có thể, ví dụ, được tạo cấu hình để lưu trữ các mẫu tín hiệu âm thanh được khôi phục vào trong vùng đệm độ trễ trước khi khung tiếp theo được nhận bởi giao diện nhận 1310.

Theo phương án, bộ xử lý mẫu 1340 có thể, ví dụ, được tạo cấu hình để lưu trữ các mẫu tín hiệu âm thanh được khôi phục vào trong vùng đệm độ trễ 1320 sau khi khung tiếp theo được nhận bởi giao diện nhận 1310.

Trong phương án, bộ xử lý mẫu 1340 có thể, ví dụ, được tạo cấu hình để định tỉ lệ lại các mẫu tín hiệu âm thanh được lựa chọn phụ thuộc vào thông tin độ khuếch đại để thu được các mẫu tín hiệu âm thanh được định tỉ lệ lại và bằng cách tổ hợp các mẫu tín hiệu âm thanh được định lại tỉ lệ với các mẫu tín hiệu âm thanh đầu vào để thu được các mẫu tín hiệu âm thanh được xử lý.

Theo phương án, bộ xử lý mẫu 1340 có thể, ví dụ, được tạo cấu hình để lưu trữ các mẫu tín hiệu âm thanh được xử lý, biểu thị sự tổ hợp của các mẫu tín hiệu âm thanh được định tỉ lệ lại và các mẫu tín hiệu âm thanh đầu vào vào trong vùng đệm độ trễ 1320, và không lưu trữ các mẫu tín hiệu âm thanh được định lại tỉ lệ vào vùng đệm độ trễ 1320, nếu khung hiện thời được nhận bởi giao diện nhận 1310 và nếu khung hiện thời được nhận bởi giao diện nhận 1310 không bị sai lạc. Hơn nữa, bộ xử lý mẫu 1340 được tạo cấu hình để lưu trữ các mẫu tín hiệu âm thanh được định lại tỉ lệ vào trong vùng đệm độ trễ 1320 và không lưu trữ các mẫu tín hiệu âm thanh vào vùng đệm độ trễ 1320, nếu khung hiện thời không được nhận bởi giao diện nhận 1310 hoặc nếu

khung hiện thời được nhận bởi giao diện nhận 1310 bị sai lạc.

Theo phương án khác, bộ xử lý mẫu 1340 có thể, ví dụ, được tạo cấu hình để lưu trữ các mẫu tín hiệu âm thanh được xử lý vào trong vùng đệm trễ 1320, nếu khung hiện thời không được nhận bởi giao diện nhận 1310 hoặc nếu khung hiện thời được nhận bởi giao diện nhận 1310 bị sai lạc.

Trong phương án, bộ lựa chọn mẫu 1330 có thể, ví dụ, được tạo cấu hình để thu được các mẫu tín hiệu âm thanh được khôi phục bằng cách định tỉ lệ lại các mẫu tín hiệu âm thanh được lựa chọn phụ thuộc vào độ khuếch đại được thay đổi, trong đó độ khuếch đại được thay đổi được xác định theo công thức:

$$\text{gain} = \text{gain\_past} * \text{damping};$$

trong đó  $\text{gain}$  là độ khuếch đại được thay đổi, trong đó bộ lựa chọn mẫu 1330 có thể, ví dụ, được tạo cấu hình để thiết lập  $\text{gain\_past}$  thành  $\text{gain}$  sau  $\text{gain}$  và đã được tính toán, và trong đó  $\text{damping}$  là giá trị thực.

Theo phương án, bộ lựa chọn mẫu 1330 có thể, ví dụ, được tạo cấu hình để tính toán độ khuếch đại được thay đổi.

Trong phương án,  $\text{damping}$  có thể, ví dụ, được xác định theo:  $0 < \text{damping} < 1$ .

Theo phương án, độ khuếch đại được thay đổi  $\text{gain}$  có thể, ví dụ, được thiết lập về 0, nếu ít nhất số các khung được định trước chưa được nhận bởi giao diện nhận 1310 vì khung cuối cùng đã được nhận bởi giao diện nhận 1310.

Dưới đây, tốc độ giảm dần cường độ được xem xét. Một vài môđun che giấu mà áp dụng trên loại nhất định của sự giảm dần cường độ. Trong khi tốc độ của sự giảm dần cường độ có thể được lựa chọn khác nhau ngang qua các môđun này, nó có lợi để sử dụng cùng tốc độ giảm dần cường độ cho tất cả các môđun che giấu cho một lõi (ACELP hoặc TCX). Ví dụ,

Đối với ACELP, tốc độ giảm dần cường độ giống nhau nên được sử dụng, cụ thể là, đối với bảng mã thích ứng (bằng việc thay đổi độ khuếch đại), và/hoặc đối với bảng mã đổi mới (bằng việc thay đổi độ khuếch đại).

Ngoài ra, đối với TCX, tốc độ giảm dần cường độ giống nhau nên được sử dụng, cụ thể là đối với tín hiệu miền thời gian, và/hoặc đối với độ khuếch đại LTP (tăng giảm cường độ về 0), và/hoặc đối với trọng số LPC (tăng giảm cường độ về 1), và/hoặc đối với các hệ số LP (giảm dần cường độ về dạng phổ nền), và/hoặc đối với sự tăng giảm chéo về nhiễu âm trắng.

Có thể còn ưu tiên sử dụng cùng tốc độ giảm dần cường độ đối với ACELP và TCX, nhưng do bản chất khác nhau của các lỗi nó có thể cũng được chọn để sử dụng các tốc độ giảm dần cường độ khác nhau.

Tốc độ giảm dần cường độ này có thể tĩnh, nhưng được ưu tiên thích ứng với các đặc điểm của tín hiệu. Ví dụ, tốc độ giảm dần cường độ có thể, ví dụ, phụ thuộc vào thừa số ổn định LPC (TCX) và/hoặc vào sự phân loại, và/hoặc trên số lượng các khung bị mất liên tiếp.

Tốc độ giảm dần cường độ có thể, ví dụ, được xác định phụ thuộc vào thừa số suy giảm, mà có thể đưa ra một cách tuyệt đối hoặc tương đối, và có thể thay đổi theo thời gian trong sự giảm dần cường độ nhất định.

Trong các phương án, tốc độ tăng giảm cường độ giống nhau được sử dụng cho sự tăng giảm cường độ khuếch đại LTP như cho sự tăng giảm cường độ nhiễu âm trắng.

Thiết bị, phương pháp và chương trình máy tính để tạo ra tín hiệu nhiễu âm nền nhân tạo như được mô tả ở trên được đề xuất

Mặc dù một số khía cạnh đã được mô tả trong bối cảnh của thiết bị, rõ ràng rằng các khía cạnh này cũng thể hiện sự mô tả của phương pháp tương ứng, trong đó khối hoặc thiết bị tương ứng với bước phương pháp hoặc đặc tính của bước phương pháp. Tương tự, các khía cạnh được mô tả trong ngữ cảnh của bước phương pháp cũng biểu diễn sự mô tả của khối hoặc mục hoặc dấu hiệu tương ứng của thiết bị tương ứng.

Tín hiệu đã được tách ra có thể được cất trữ trên môi trường lưu trữ số hoặc có thể được truyền trên môi trường truyền như môi trường truyền dẫn không dây hoặc môi trường truyền dẫn 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. Phương án có thể được thực hiện sử dụng vật 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 tử được lưu trữ trên đó, mà kết hợp (hoặc có thể kết hợp) với hệ thống máy tính khả trình sao cho phương pháp tương ứng được thực hiện.

Một số phương án theo sáng chế gồm có vật 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 các phương pháp được mô tả ở đây được thực hiện.

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

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

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

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

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

Phương án khác gồm có 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 các

phương pháp được mô tả trong tài liệu này.

Phương án nữa bao gồm máy tính có chương trình máy tính được cài đặt trên máy tính để thực hiện một trong các phương pháp được mô tả trong tài liệu này.

Theo một số phương án, thiết bị logic lập trình được (ví dụ, mảng công 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 trong tài liệu này. Theo một số phương án, mảng công 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 trong tài liệu này. Thông thường, các phương pháp ưu tiên được thực hiện bởi thiết bị phần cứng bất kỳ.

Các phương án được mô tả bên trên chỉ mang tính minh họa cho các nguyên lý của sáng chế. Cần hiểu rằng các sự biến thể và biến đổi của các phương án và các chi tiết được mô tả ở đây sẽ là rõ ràng đối với người có trình độ trong lĩnh vực kỹ thuật tương ứng. Do đó, mục đích chỉ bị 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 thể hiện bằng cách mô tả và giải thích của các phương án ở đây.

## YÊU CẦU BẢO HỘ

1. Thiết bị để giải mã tín hiệu âm thanh được mã hóa để thu tín hiệu âm thanh được khôi phục, trong đó thiết bị được tạo cấu hình để nhận nhiều khung, và trong đó thiết bị bao gồm:

môđun biến đổi cosin rời rạc cải biên nghịch đảo (683) để giải mã nhiều khung bằng cách tiến hành phép biến đổi cosin rời rạc cải biên nghịch đảo để thu được các mẫu tín hiệu âm thanh của tín hiệu âm thanh được giải mã, và

bộ phận dự báo dài hạn (684) để tiến hành sự dự báo dài hạn, bao gồm:

vùng đệm trễ (1020) để lưu trữ các mẫu tín hiệu âm thanh của tín hiệu âm thanh được giải mã,

bộ lựa chọn mẫu (1030) để lựa chọn nhiều mẫu tín hiệu âm thanh được lựa chọn từ các mẫu tín hiệu âm thanh được lưu trữ trong vùng đệm trễ (1020), và

bộ xử lý mẫu (1040) để xử lý các mẫu tín hiệu âm thanh được lựa chọn để thu được các mẫu tín hiệu âm thanh được khôi phục của tín hiệu âm thanh được khôi phục,

trong đó bộ lựa chọn mẫu (1030) được tạo cấu hình để lựa chọn, nếu khung hiện thời được nhận bằng thiết bị và hoặc khung hiện thời được nhận bằng thiết bị không bị sai lạc, nhiều mẫu tín hiệu âm thanh được lựa chọn từ các mẫu tín hiệu âm thanh được lưu trữ trong vùng đệm trễ (1020) phụ thuộc vào thông tin độ trễ bước được bao gồm bởi các khung hiện thời, và

trong đó bộ lựa chọn mẫu (1030) được tạo cấu hình để lựa chọn, nếu khung hiện thời không được nhận bằng thiết bị hoặc nếu khung hiện thời được nhận bằng thiết bị bị sai lạc, nhiều mẫu tín hiệu âm thanh được lựa chọn từ các mẫu tín hiệu âm thanh được lưu trữ trong vùng đệm trễ (1020) phụ thuộc vào thông tin độ trễ bước được bao gồm bởi khung khác được nhận trước đó bằng thiết bị,

trong đó bộ lựa chọn mẫu (1030) được tạo cấu hình để thu các mẫu tín hiệu âm thanh được khôi phục bằng việc định lại tỉ lệ các mẫu tín hiệu âm thanh được lựa chọn phụ thuộc vào độ khuếch đại được thay đổi, trong đó độ khuếch đại được thay đổi được xác định theo công thức:

`gain = gain_past * damping;`

trong đó `gain` là độ khuếch đại được thay đổi,

trong đó bộ lựa chọn mẫu (1030) được tạo cấu hình để thiết lập `gain_past` thành `gain` sau khi `gain` được tính toán, và

trong đó `damping` là giá trị thực với  $0 \leq \text{damping} \leq 1$ .

## 2. Thiết bị theo điểm 1,

trong đó bộ xử lý mẫu (1040) được tạo cấu hình để thu các mẫu tín hiệu âm thanh được khôi phục, nếu khung hiện thời được nhận bằng thiết bị và nếu khung hiện thời được nhận bằng thiết bị không bị sai lạc, bằng việc định lại tỉ lệ các mẫu tín hiệu âm thanh được lựa chọn phụ thuộc vào thông tin độ khuếch đại được bao gồm bởi khung hiện thời, và

trong đó bộ lựa chọn mẫu (1030) được tạo cấu hình để thu các mẫu tín hiệu âm thanh được khôi phục, nếu khung hiện thời được nhận bằng thiết bị hoặc nếu khung hiện thời được nhận bằng thiết bị bị sai lạc, bằng sự định lại tỉ lệ các mẫu tín hiệu âm thanh được lựa chọn phụ thuộc vào thông tin độ khuếch đại được bao gồm bằng khung khác đã nêu được nhận trước đó bằng thiết bị.

## 3. Thiết bị theo điểm 2,

trong đó bộ xử lý mẫu (1040) được tạo cấu hình để thu các mẫu tín hiệu âm thanh được khôi phục, nếu khung hiện thời được nhận bằng thiết bị và nếu khung hiện thời được nhận bằng thiết bị không bị sai lạc, bằng việc nhân các mẫu tín hiệu âm thanh được lựa chọn và giá trị phụ thuộc vào thông tin độ khuếch đại được bao gồm bởi khung hiện thời, và

trong đó bộ lựa chọn mẫu (1030) được tạo cấu hình để thu các mẫu tín hiệu âm thanh được khôi phục, nếu khung hiện thời không được nhận bằng thiết bị hoặc nếu khung hiện thời được nhận bằng thiết bị bị sai lạc, bằng việc nhân các mẫu tín hiệu âm thanh được lựa chọn và giá trị phụ thuộc vào thông tin độ khuếch đại được bao gồm bởi khung khác đã nêu được nhận trước đó bằng thiết bị.

4. Thiết bị theo một trong số các điểm đã nêu, trong đó bộ xử lý mẫu (1040) được tạo cấu hình để lưu trữ các mẫu tín hiệu âm thanh được khôi phục vào trong vùng đệm trễ (1020).

5. Thiết bị theo điểm 4, trong đó bộ xử lý mẫu (1040) được tạo cấu hình để lưu trữ các mẫu tín hiệu âm thanh được khôi phục vào trong vùng đệm trễ (1020) trước khi khung tiếp theo được nhận bằng thiết bị.

6. Thiết bị theo điểm 4, trong đó bộ xử lý mẫu (1040) được tạo cấu hình để lưu trữ các mẫu tín hiệu âm thanh được khôi phục vào trong vùng đệm trễ (1020) sau khi khung tiếp theo được nhận bởi thiết bị.

7. Thiết bị theo điểm 1, trong đó bộ xử lý mẫu (1040) được tạo cấu hình để định lại tỉ lệ các mẫu tín hiệu âm thanh được lựa chọn phụ thuộc vào thông tin độ khuếch đại để thu các mẫu tín hiệu âm thanh được định lại tỉ lệ và bằng việc tổ hợp các mẫu tín hiệu âm thanh được định lại tỉ lệ với các mẫu tín hiệu âm thanh đầu vào để thu các mẫu tín hiệu âm thanh được xử lý.

8. Thiết bị theo điểm 7,

trong đó bộ xử lý mẫu (1040) được tạo cấu hình để lưu trữ các mẫu tín hiệu âm thanh được xử lý, biểu thị sự tổ hợp của các mẫu tín hiệu âm thanh được định lại tỉ lệ và các mẫu tín hiệu âm thanh đầu vào, vào trong vùng đệm trễ (1020), và để không lưu trữ các tín hiệu âm thanh được định lại tỉ lệ vào trong vùng đệm trễ (1020), nếu khung hiện thời được nhận bằng thiết bị và nếu khung hiện thời được nhận bằng thiết bị không bị sai lạc, và

trong đó bộ xử lý mẫu (1040) được tạo cấu hình để lưu trữ các mẫu tín hiệu âm thanh được định lại tỉ lệ vào trong vùng đệm trễ (1020) và để không lưu trữ các mẫu tín hiệu âm thanh được xử lý vào trong vùng đệm trễ (1020), nếu khung hiện thời không được nhận bằng thiết bị hoặc nếu khung hiện thời được nhận bằng thiết bị bị sai lạc.

9. Thiết bị theo điểm 7, trong đó bộ xử lý mẫu (1040) được tạo cấu hình để lưu trữ các mẫu tín hiệu âm thanh được xử lý vào trong vùng đệm trễ (1020), nếu khung hiện thời

không được nhận bằng thiết bị hoặc nếu khung hiện thời được nhận bằng thiết bị bị sai lạc.

10. Thiết bị theo một trong số các điểm đã nêu, trong đó bộ lựa chọn mẫu (1030) được tạo cấu hình để tính toán độ khuếch đại được biến đổi.

11. Thiết bị theo một trong số các điểm nêu trên, trong đó độ khuếch đại được thay đổi *gain* được thiết lập về 0, nếu ít nhất nhiều khung được định trước không được nhận bằng thiết bị do khung cuối cùng được nhận bằng thiết bị.

12. Phương pháp giải mã tín hiệu âm thanh được mã hóa để thu được tín hiệu âm thanh được khôi phục, trong đó phương pháp bao gồm:

nhận nhiều khung,

giải mã nhiều khung bằng cách tiến hành phép biến đổi cosin rời rạc cải biên nghịch đảo để thu được các mẫu tín hiệu âm thanh của tín hiệu âm thanh được mã hóa,

tiến hành việc dự đoán dài hạn bằng cách:

lưu trữ các mẫu tín hiệu âm thanh của tín hiệu âm thanh được giải mã,

lựa chọn nhiều mẫu tín hiệu âm thanh từ các mẫu tín hiệu âm thanh được lưu trữ trong vùng đệm trễ (1020), và

xử lý các mẫu tín hiệu âm thanh được lựa chọn để thu được các mẫu tín hiệu âm thanh được khôi phục của tín hiệu âm thanh được khôi phục,

trong đó, nếu khung hiện thời được nhận và nếu khung hiện thời được nhận không bị sai lạc, bước lựa chọn nhiều mẫu tín hiệu âm thanh được lựa chọn từ các mẫu tín hiệu âm thanh được lưu trữ trong vùng đệm trễ (1020) được tiến hành phụ thuộc vào thông tin độ trễ bước được bao gồm bởi khung hiện thời, và

trong đó, nếu khung hiện thời không được nhận hoặc nếu khung hiện thời được nhận bị sai lạc, bước lựa chọn nhiều mẫu tín hiệu âm thanh được lựa chọn từ các mẫu tín hiệu âm thanh được lưu trữ trong vùng đệm trễ (1020) được tiến hành phụ thuộc vào thông tin độ trễ bước được bao gồm bởi khung khác được nhận trước đó,

trong đó phương pháp còn bao gồm bước định lại tỉ lệ các mẫu tín hiệu âm thanh được lựa chọn phụ thuộc vào độ khuếch đại được thay đổi, trong đó độ khuếch đại được thay đổi được xác định theo công thức:

$$\text{gain} = \text{gain\_past} * \text{damping};$$

trong đó  $\text{gain}$  là độ khuếch đại được thay đổi,

trong đó  $\text{gain\_past}$  được thiết lập thành  $\text{gain}$  sau khi  $\text{gain}$  được tính toán, và

trong đó  $\text{damping}$  là giá trị thực với  $0 \leq \text{damping} \leq 1$ .

13. Vật ghi có thể đọc được bằng máy tính bao gồm chương trình máy tính để thực hiện phương pháp theo điểm 12 khi chạy trên máy tính hoặc bộ xử lý tín hiệu.

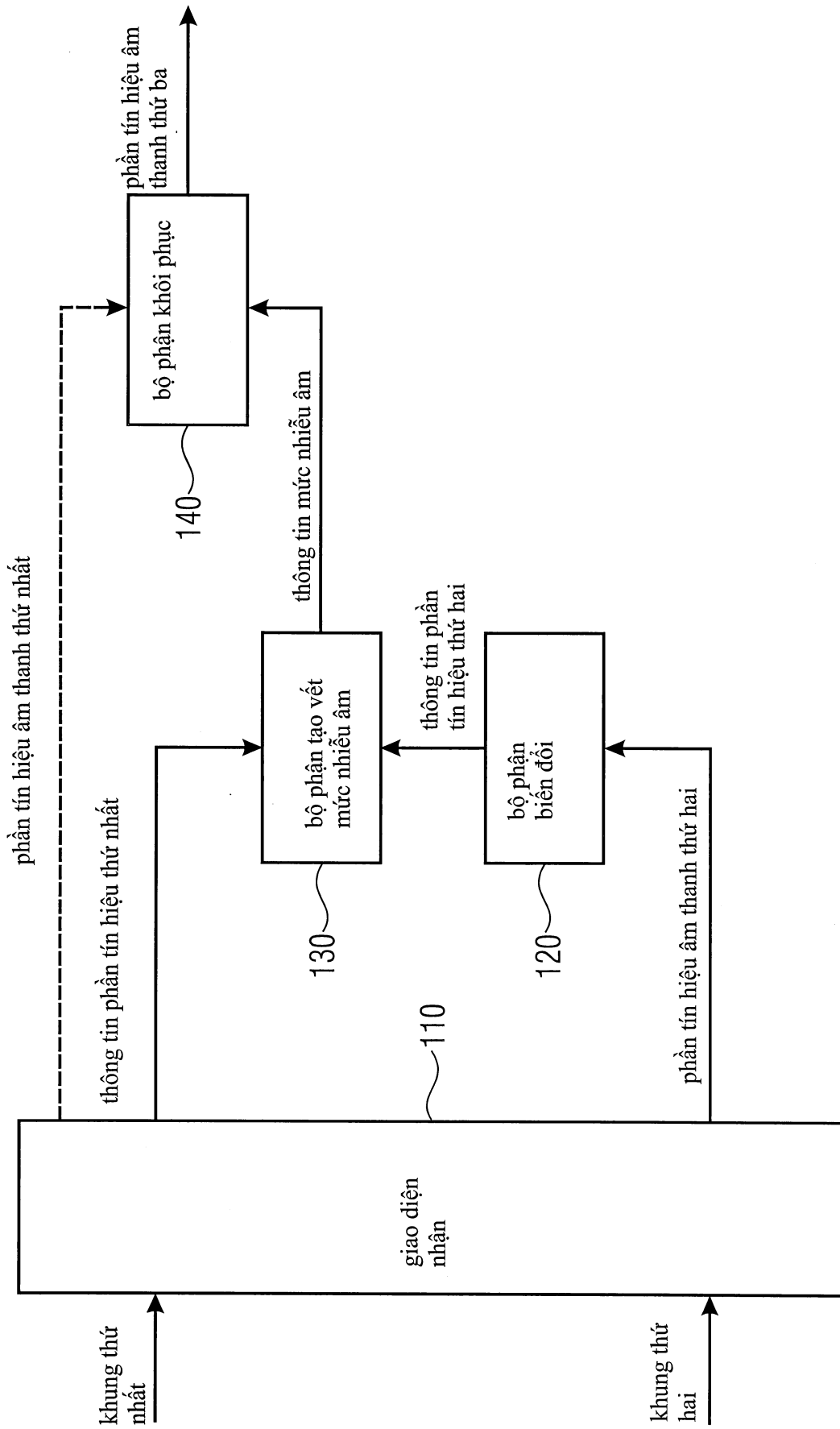


FIG 1A

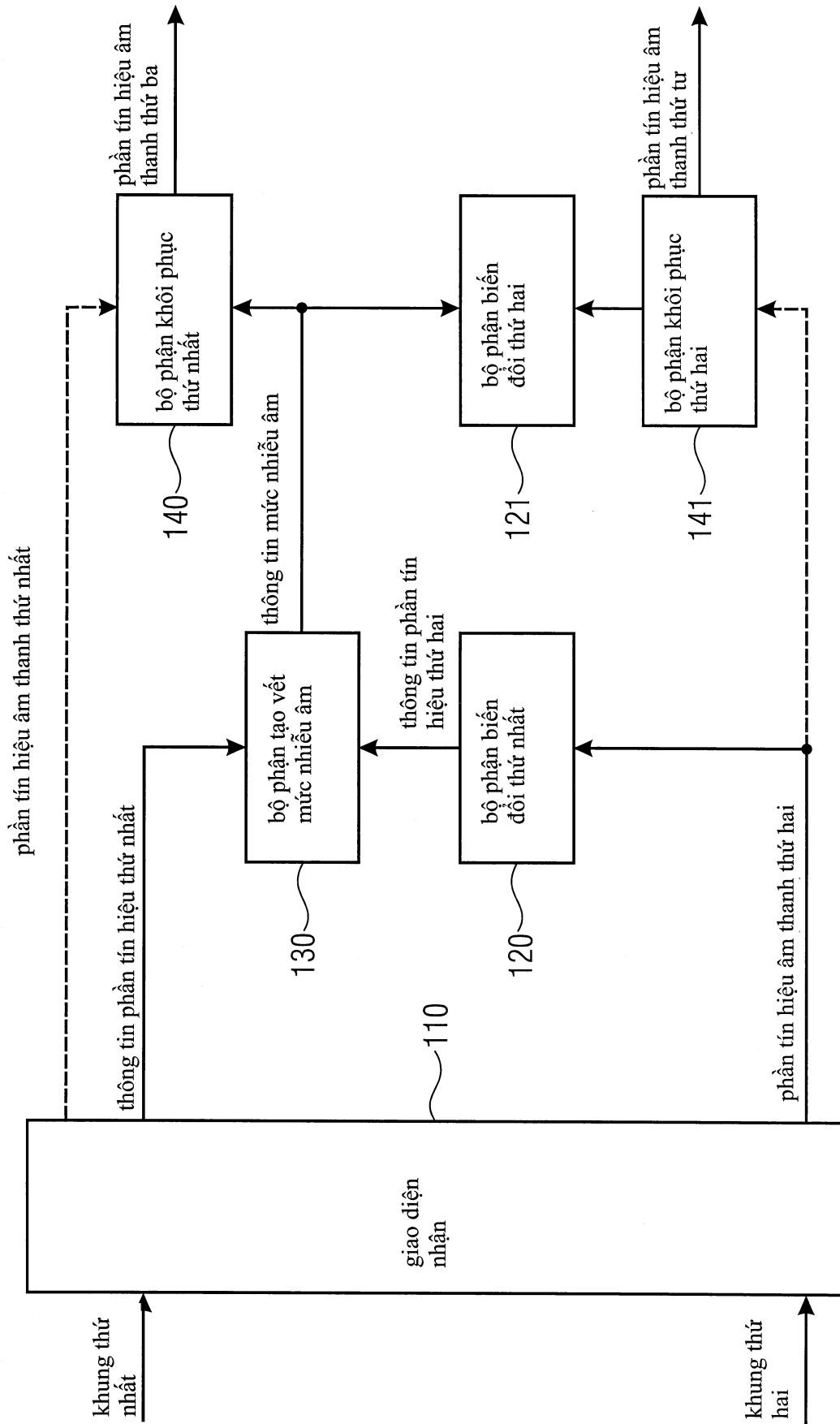


FIG 1B

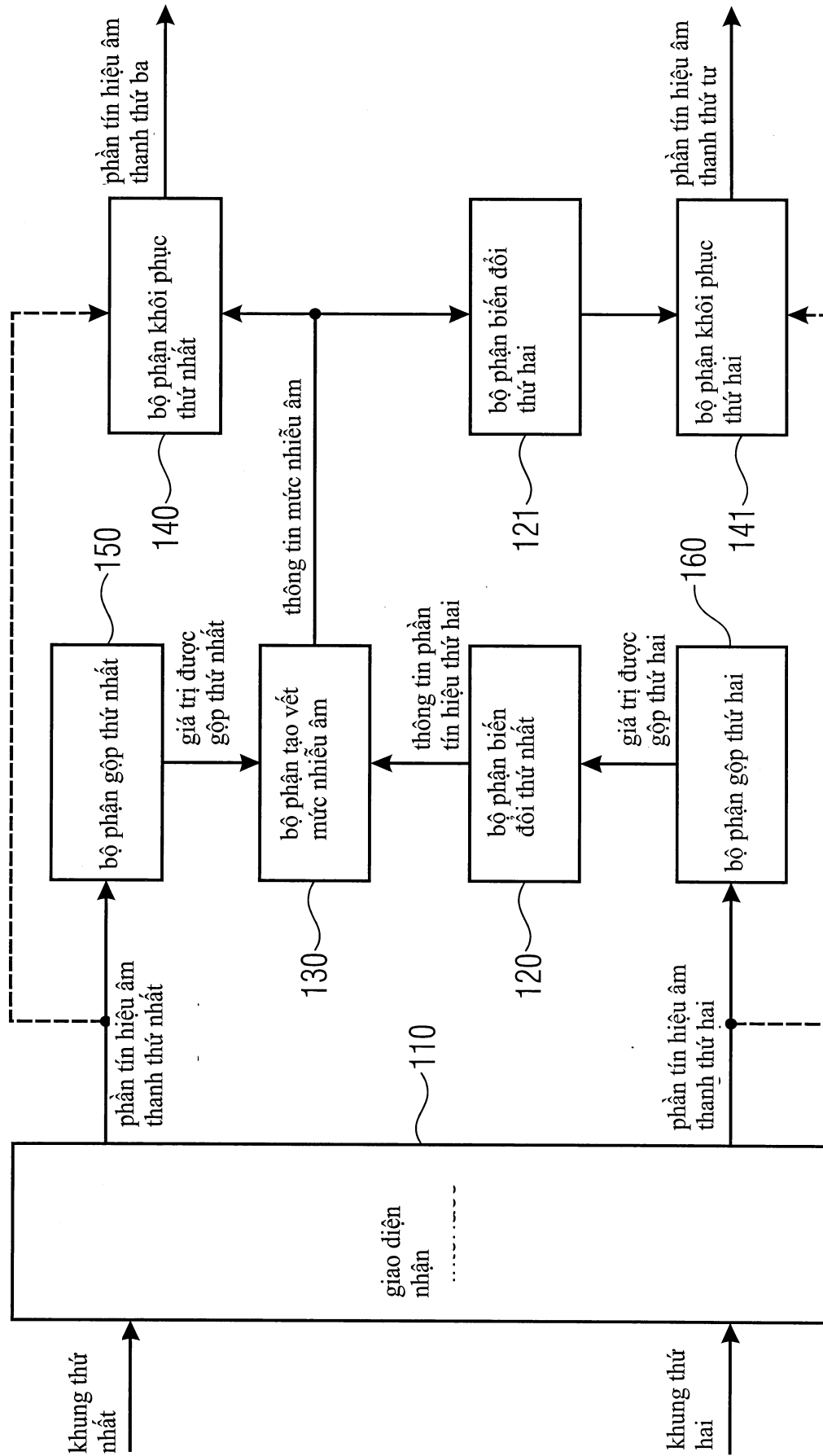


FIG 1C

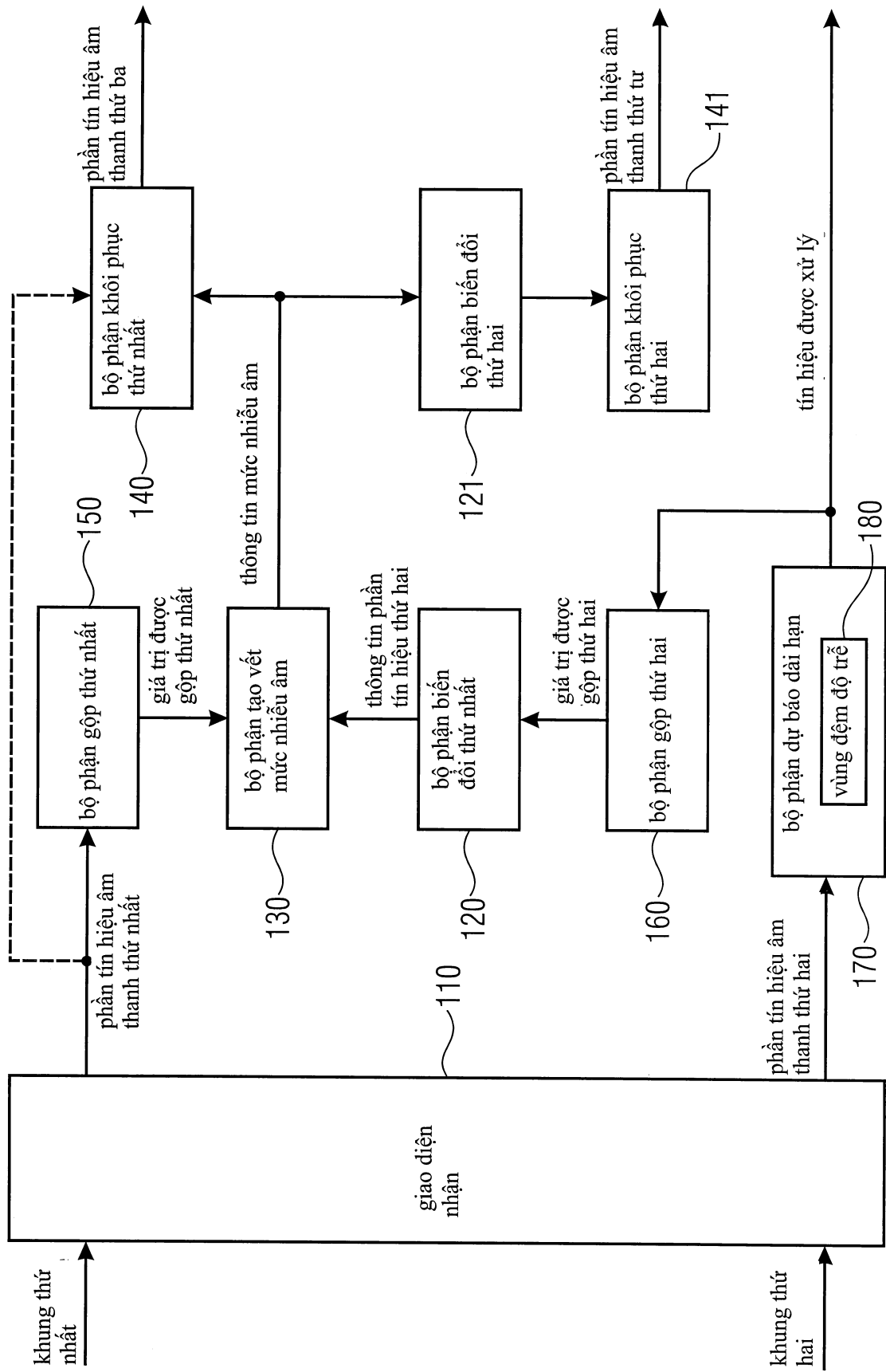


FIG 1D

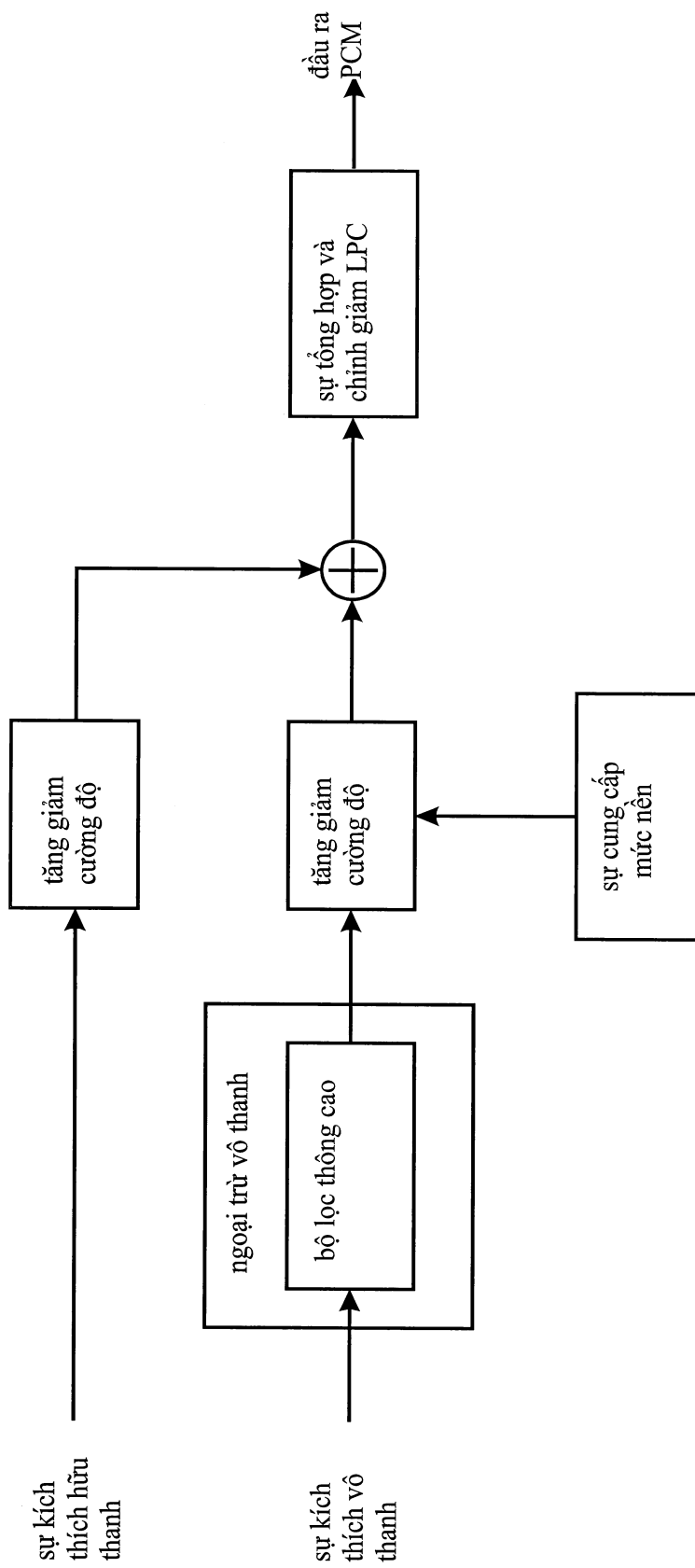
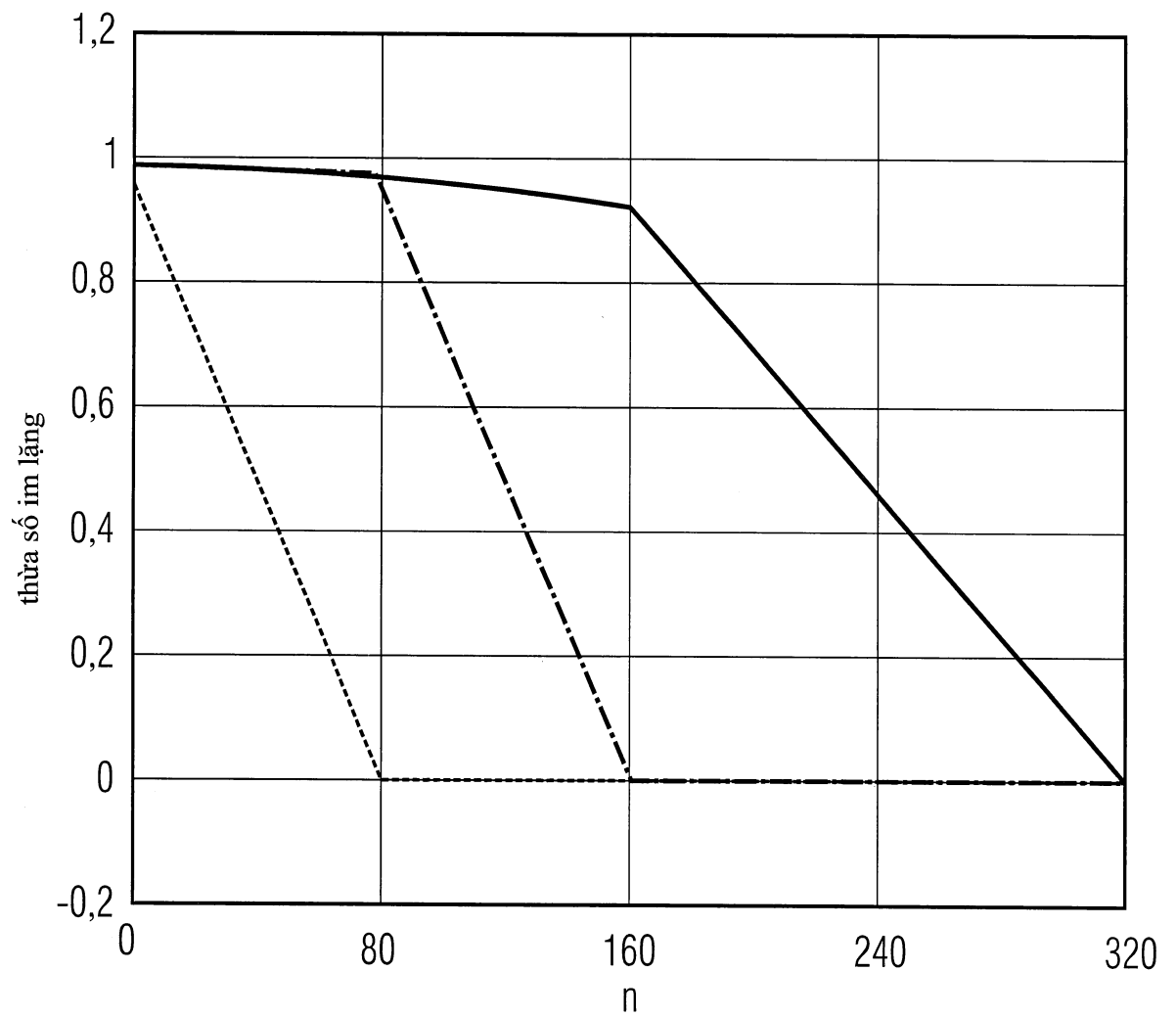


FIG 2



G.722APP.V\_F07

..... TRANSIENT  
 -.-.-.- VUV\_TRANSITION  
 ————— các loại khác

FIG 3

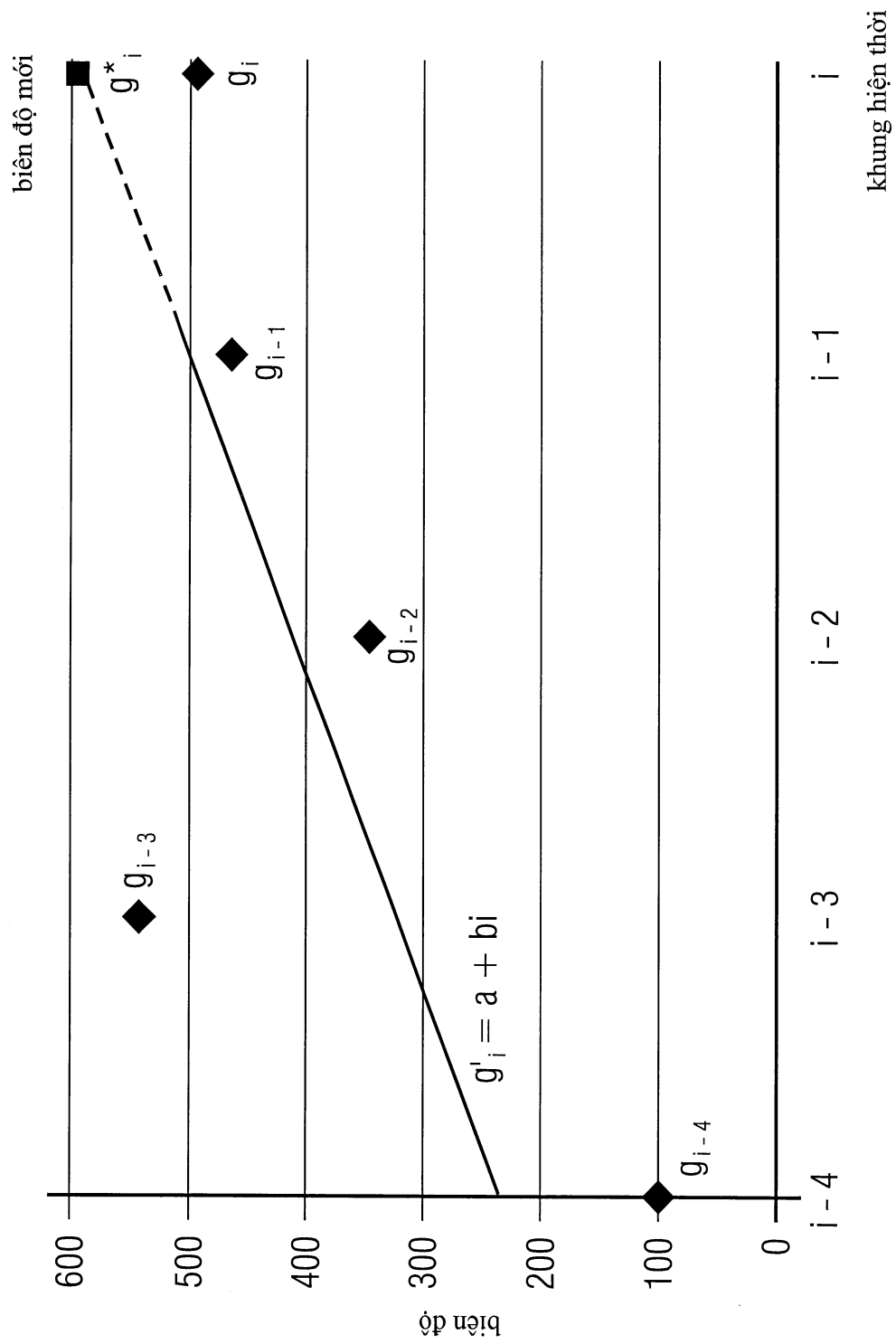


FIG 4

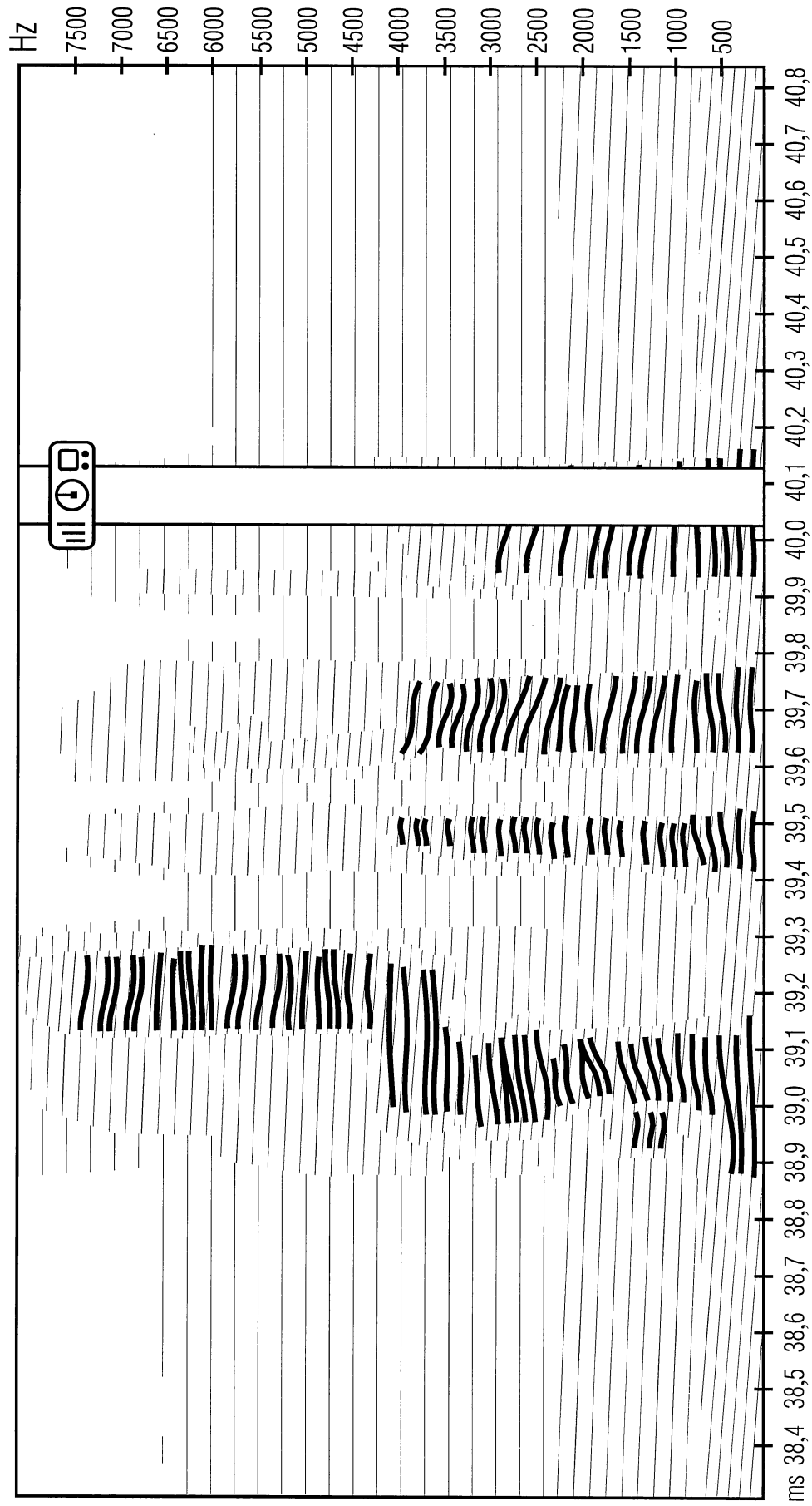


FIG 5

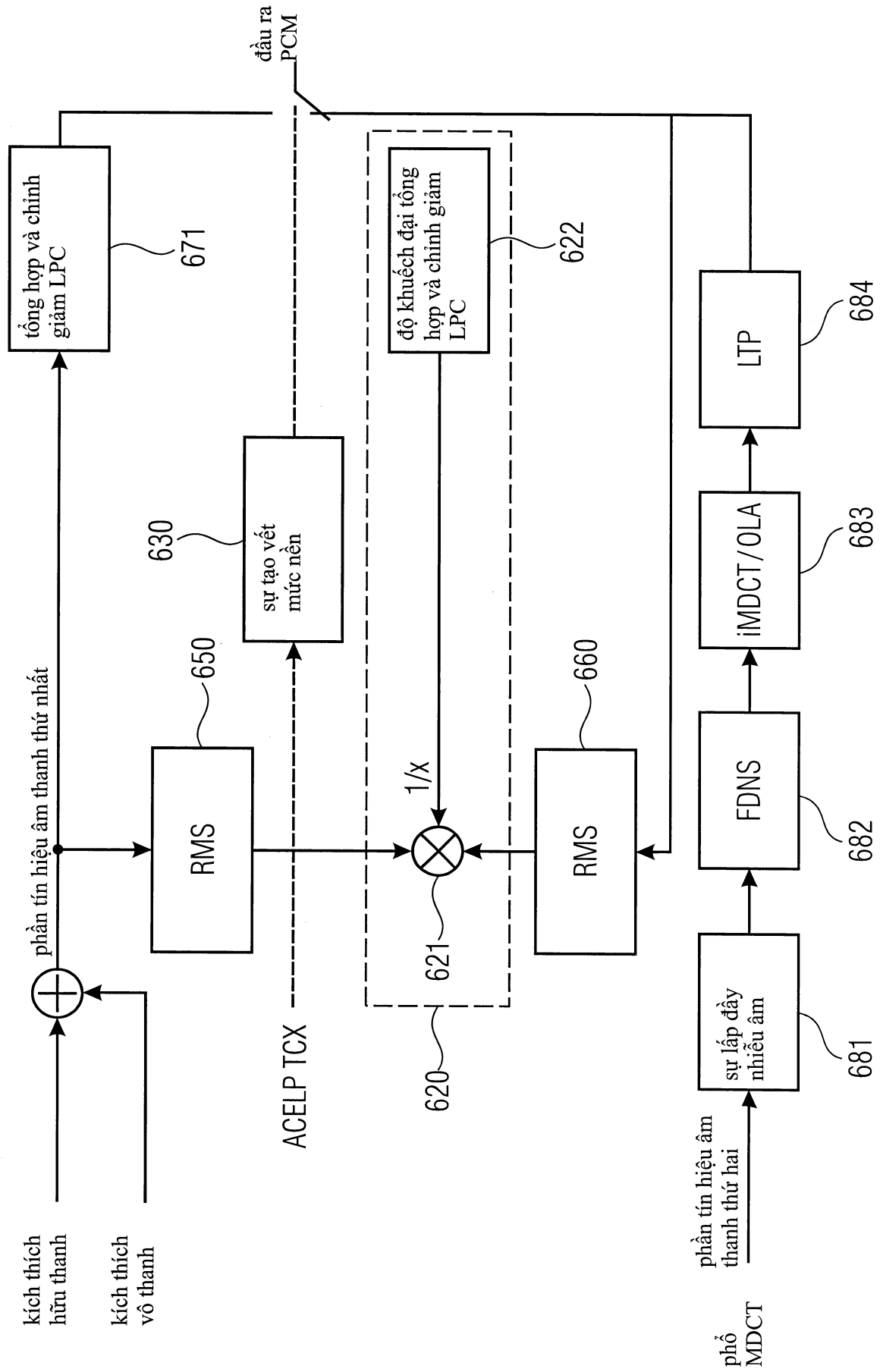


FIG 6

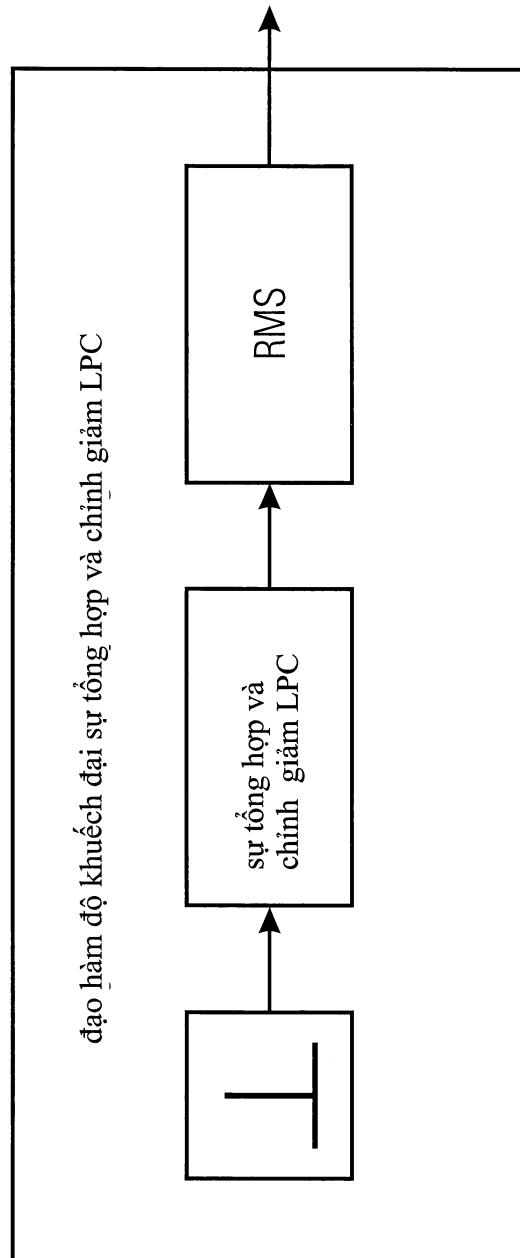


FIG 7

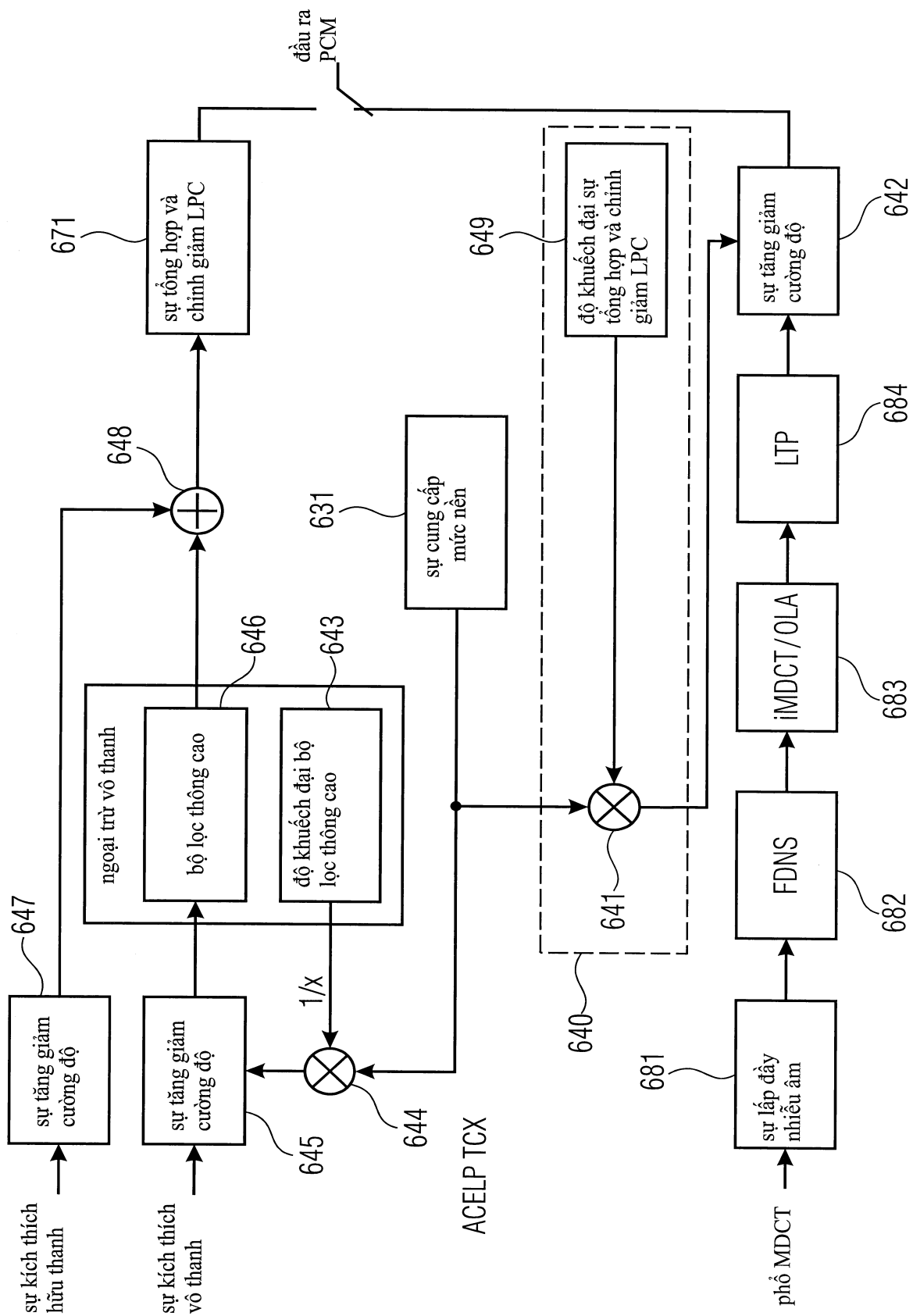


FIG 8

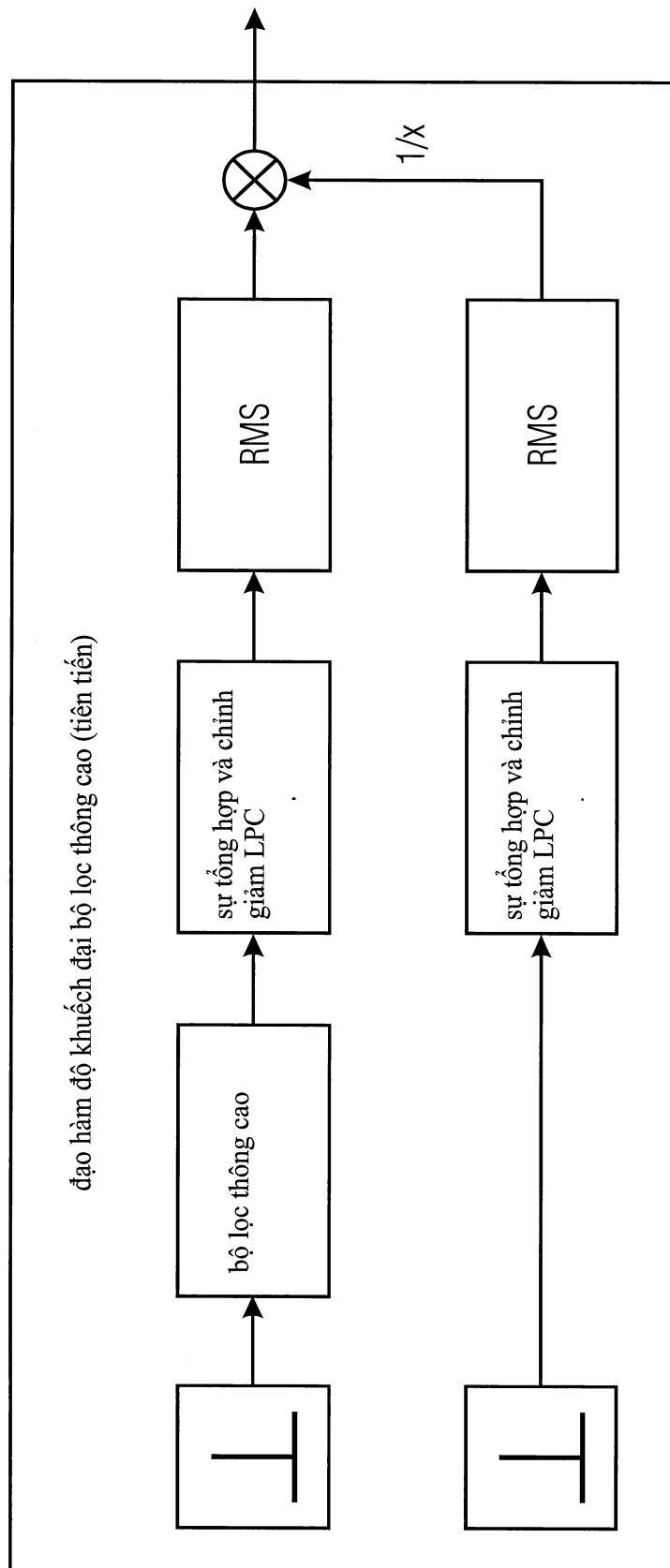


FIG 9

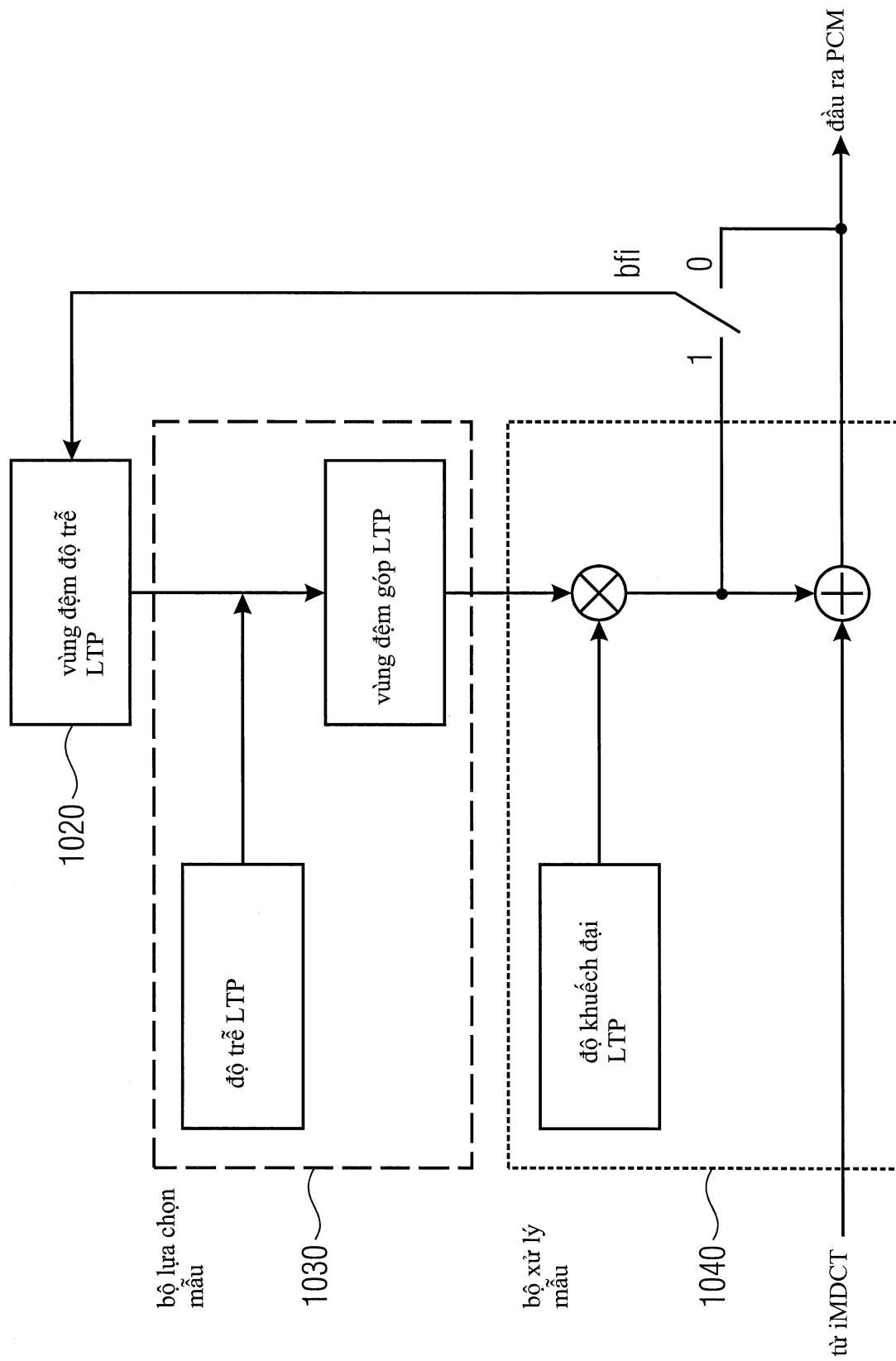


FIG 10

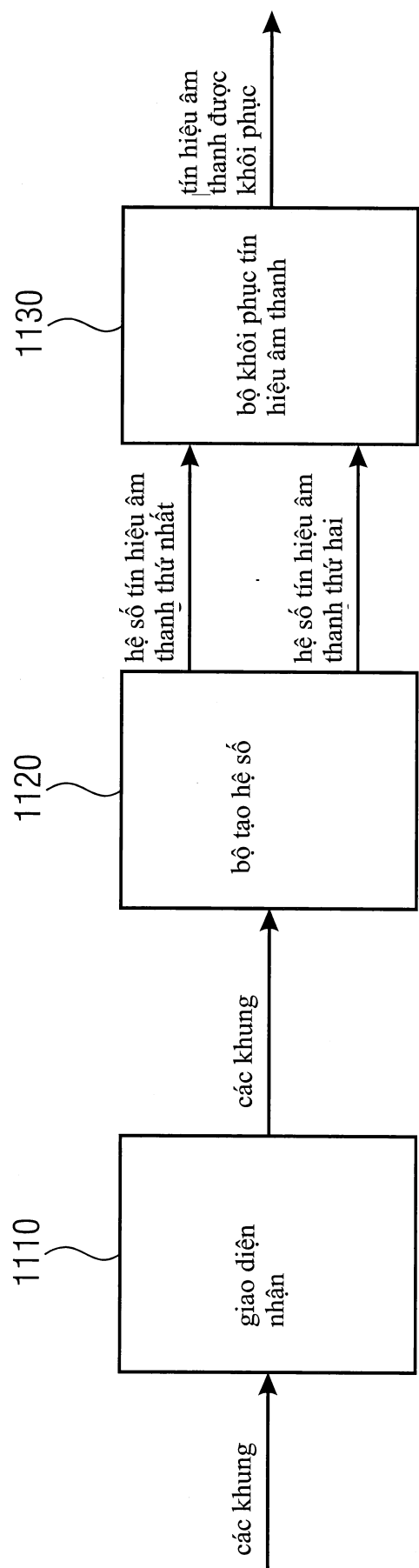


FIG 11

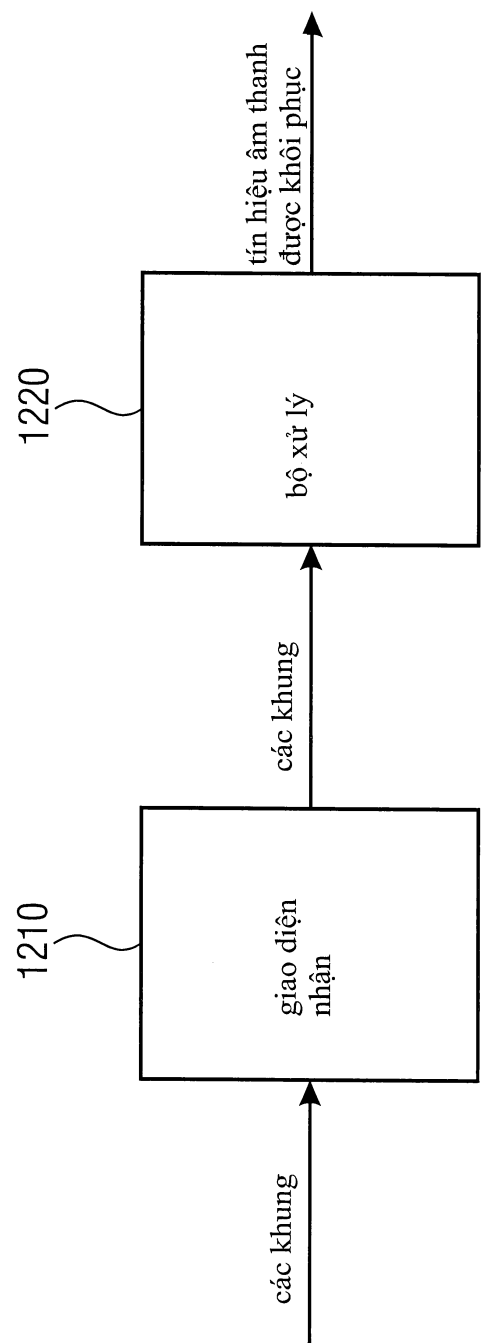


FIG 12

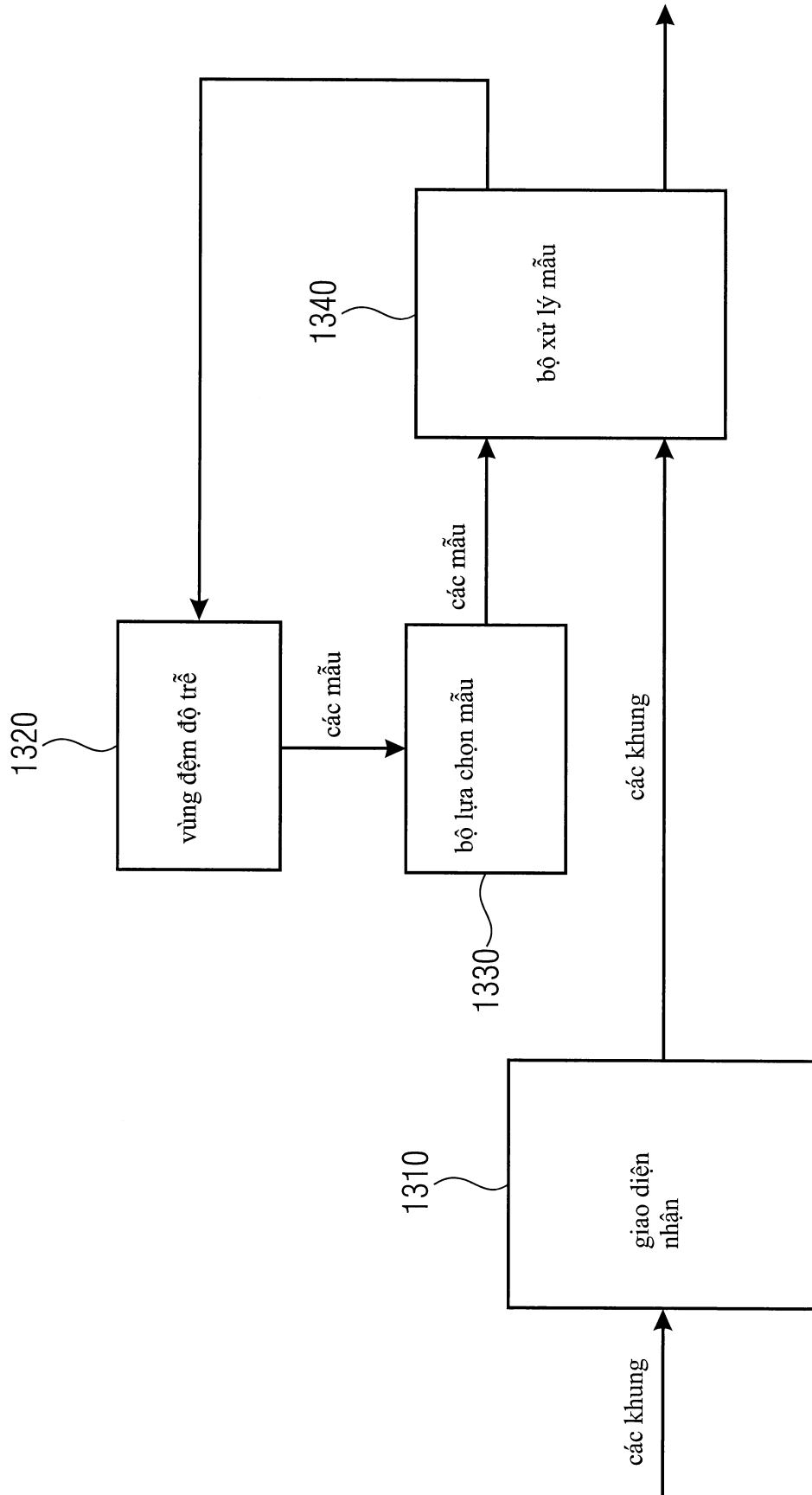


FIG 13

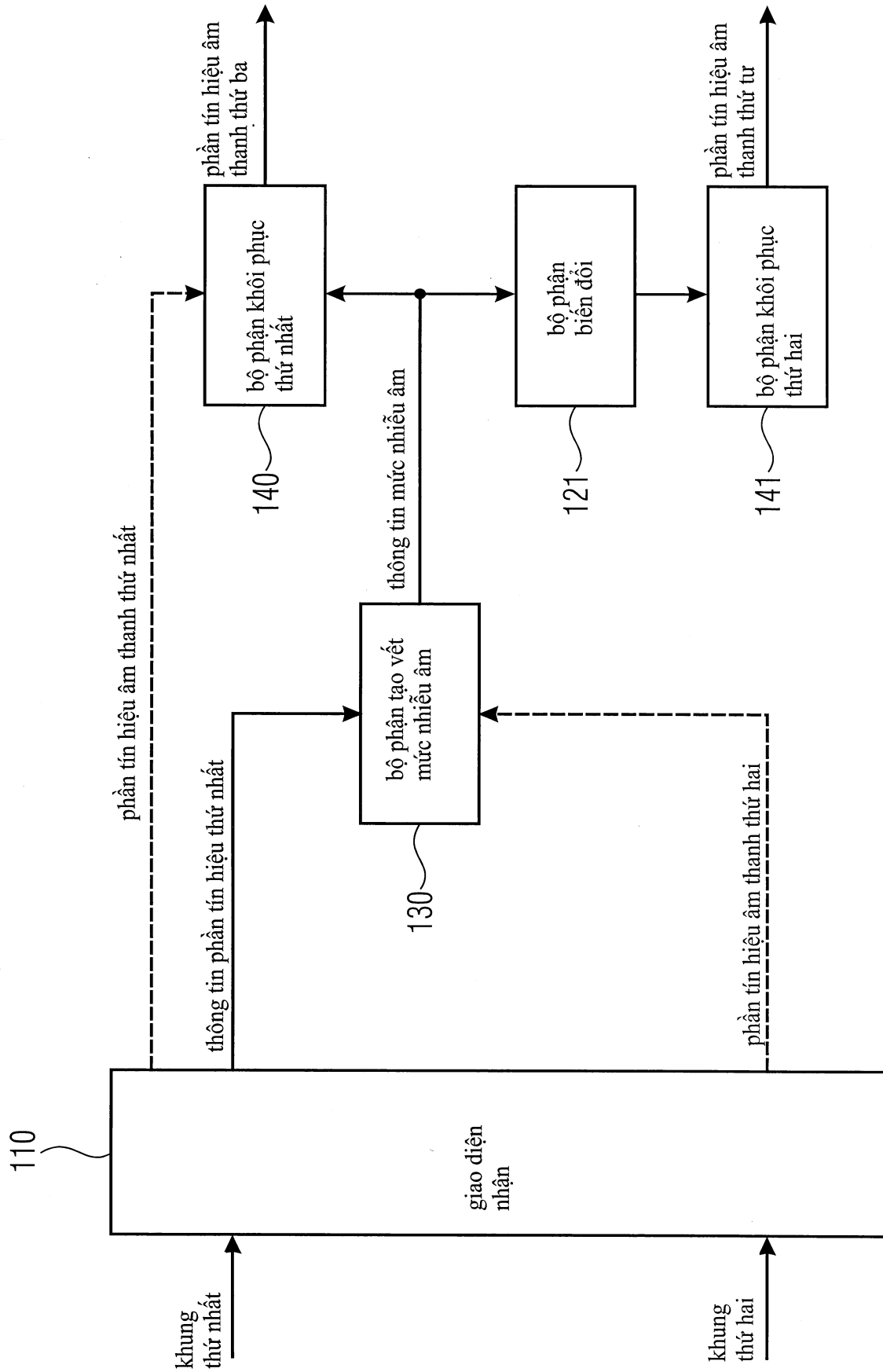


FIG 14