



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

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

(11)



1-0024395

(51)⁷ H04N 7/32

(13) B

(21) 1-2016-03594

(22) 24/04/2013

(62) 1-2015-00392

(86) PCT/JP2013/062061 24/04/2013

(87) WO2014/006959A1 09/01/2014

(30) 2012-148310 02/07/2012 JP

(45) 27/07/2020 388

(43) 27/04/2015 325A

(73) NTT DOCOMO, INC. (JP)

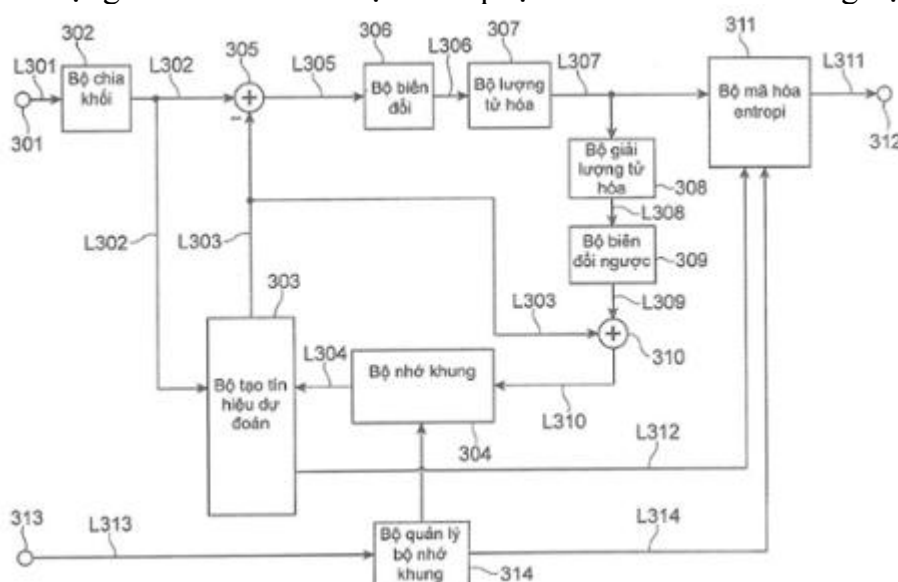
11-1, Nagatacho 2-chome, Chiyoda-ku, Tokyo 100-6150, Japan

(72) FUJIBAYASHI Akira (JP); BOON Choong Seng (MY); TAN Thiow Keng (MY).

(74) Công ty TNHH một thành viên Sở hữu trí tuệ VCCI (VCCI-IP CO.,LTD)

(54) PHƯƠNG PHÁP GIẢI MÃ DỰ ĐOÁN VIDEO VÀ THIẾT BỊ GIẢI MÃ DỰ ĐOÁN VIDEO

(57) Sáng chế đề cập đến thiết bị mã hóa và giải mã dự đoán video, trong đó thiết bị mã hóa dự đoán video bao gồm bộ lưu trữ ảnh để lưu trữ một hoặc nhiều ảnh được khôi phục làm ảnh tham chiếu để sử dụng cho việc mã hóa ảnh tiếp theo, ảnh này được khôi phục từ dữ liệu ảnh được nén và bộ điều khiển bộ nhớ để điều khiển bộ nhớ khung trong bộ lưu trữ ảnh. Khi kích thước khung của ảnh đầu vào không lớn hơn $1/(2^L)$ (L là số nguyên lớn hơn hoặc bằng một) của kích thước khung tối đa của ảnh, bộ điều khiển bộ nhớ xác định rằng (2^L) lần của số lượng tối đa của ảnh được khôi phục là có thể lưu trữ trong bộ nhớ khung.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến các thiết bị, phương pháp và chương trình mã hóa và giải mã dự đoán video, cụ thể hơn là đề cập đến các thiết bị, phương pháp và chương trình nhằm quản lý số lượng tối đa ảnh được khôi phục có thể lưu trữ được trong bộ đệm ảnh được giải mã.

Tình trạng kỹ thuật của sáng chế

Các kỹ thuật nén video được sử dụng để truyền và lưu trữ hiệu quả dữ liệu video. MPEG1-4 và H.261 tới H.264 là các chuẩn nén dữ liệu video được sử dụng rộng rãi.

Trong những chuẩn nén video này, ảnh để mã hóa được chia thành các khối, các khối này được mã hóa và rồi được giải mã. Nhằm tăng hiệu quả của việc mã hóa, mã hóa dự đoán được mô tả bên dưới đây được sử dụng. Trong dự đoán trong khung, tín hiệu dự đoán được tạo ra sử dụng tín hiệu của ảnh lân cận được khôi phục (tín hiệu được lưu trữ từ dữ liệu ảnh được nén trước đây) mà có trong khung chứa khối đích. Bằng cách loại bỏ tín hiệu dự đoán khỏi tín hiệu của khối đích, sự khác biệt giữa chúng được thu lại và được mã hóa. Trong dự đoán liên khung, tín hiệu ảnh được khôi phục, mà có trong khung không phải khung chứa khối đích, được tìm kiếm để chuyển đổi tín hiệu. Tín hiệu dự đoán được tạo ra để bù cho phần chuyển đổi. Bằng cách loại bỏ tín hiệu dự đoán từ tín hiệu của khối đích, sự khác biệt giữa chúng được thu lại và được mã hóa. Ảnh được khôi phục mà là đối tượng của việc tìm kiếm chuyển động và bù trừ chuyển động được gọi là ảnh tham chiếu.

Trong dự đoán liên khung hai chiều, không chỉ ảnh quá khứ được tham chiếu mà cả ảnh tương lai cũng được tham chiếu, ảnh này được sắp xếp để trình chiếu sau ảnh đích (ảnh tương lai cần được mã hóa và khôi phục trước khi mã hóa ảnh đích). Sau đó, các tín hiệu dự đoán thu được từ ảnh quá khứ và ảnh tương lai được lấy trung bình. Phương pháp dự đoán này hiệu quả để dự đoán một vật thể

không có trong quá khứ nhưng được đưa vào khung tương lai và để giảm nhiễu có trong hai tín hiệu dự đoán.

Thêm vào đó, trong dự đoán liên khung được xác định trong H.264, các ảnh tham chiếu mà đã được mã hóa và sau đó được khôi phục là để tìm chuyển động, và tín hiệu dự đoán với lỗi nhỏ nhất được chọn lựa làm tín hiệu dự đoán tối ưu cho khối đích. Sự khác biệt được tính toán giữa tín hiệu điểm ảnh của khối đích và tín hiệu dự đoán tối ưu và sau đó được thực hiện phép biến đổi cosin rời rạc, lượng tử hóa và mã hóa entropi. Đồng thời, thông tin được mã hóa mà xác định ảnh tham chiếu được chọn và vùng trong ảnh tham chiếu được chọn mà từ đó tín hiệu dự đoán tối ưu cho khối đích được thu (tương ứng được gọi là “danh mục tham chiếu” và “véctor chuyển động”).

Trong H.264, các ảnh được khôi phục có thể được tham chiếu. Ảnh được khôi phục được lưu trữ, làm ảnh tham chiếu để sử dụng trong dự đoán, trong bộ đệm ảnh được giải mã (DPB) là bộ nhớ đệm ảnh. Kích thước của bộ đệm ảnh được giải mã (DPB) được xác định bởi hiện trạng và cấp độ, và được xác định là một số bit, thay vì số lượng ảnh tham chiếu. Ngay cả với cùng một hiện trạng và cấp độ, số lượng ảnh tham chiếu có thể lưu trữ thay đổi tùy theo kích thước khung của ảnh. Ví dụ, trong trường hợp mà hiện trạng là chính (Main) và mức độ là 3,2, kích cỡ tối đa của bộ đệm ảnh (MaxDPBSize: Maximum Decoded Picture Buffer Size) để lưu trữ ảnh tham chiếu được sử dụng cho dự đoán được xác định là $7680,0 \times 1024$ [byte]. Chính vì vậy, số lượng ảnh khôi phục có thể lưu trữ trong bộ đệm ảnh được giải mã (DPB) là 5 nếu ảnh là 1280×720 và 4:2:0, và số lượng ảnh khôi phục tối đa có thể được lưu trữ trong bộ đệm ảnh giải mã (DPB) là 4 nếu ảnh là 1280×1024 và 4:2:0. Fig.1(a), Fig.1(b) và Fig.1(c) cho thấy ảnh được sắp xếp trong bộ đệm ảnh được giải mã mà trong đó kích cỡ khung hình của ảnh xác định số lượng tối đa ảnh được có thể được lưu trữ, cụ thể là 4 (Fig.1(a)), 5 (Fig.1(b)), hoặc 6 (Fig.1(c)). Con trỏ bộ nhớ được cung cấp thích nghi với kích thước khung hình của ảnh được khôi phục trong bộ nhớ đệm ảnh được chuẩn bị trước, theo đó việc sắp xếp bộ nhớ thích nghi được thực hiện trong bộ nhớ đệm ảnh.

Tài liệu trích dẫn

Tài liệu phi sáng chế

Tài liệu phi sáng chế 1: "H.264: Advanced video coding for generic audiovisual services – Mã hóa video nâng cao đối với các dịch vụ nghe nhìn chung," Nhóm video liên kết của ITU-T VCEG và ISO/IEC MPEG, ITU-T Rec. H.264 và ISO/IEC14496-10 (MPEG4 - phần 10), tháng 11 năm 2007.

Vấn đề kỹ thuật

Do thiết bị giải mã mà tuân thủ các yêu cầu nói trên của H.264 cần có thể giải mã ảnh với tất cả các kích thước khung được xác định bởi hiện trạng và cấp độ mà nằm trong phạm vi được hỗ trợ bởi thiết bị giải mã, sẽ không thể sắp xếp bộ nhớ trong bộ nhớ đệm ảnh và lưu trữ ảnh được khôi phục trong các bộ nhớ này một cách cố định. Với lý do này, cần phải biến đổi, theo kích thước khung của ảnh cần giải mã, con trỏ bộ nhớ mà chỉ vào vị trí của bộ nhớ lưu trữ ảnh được khôi phục, kết quả làm cho việc thực hiện và điều khiển bộ nhớ trở nên phức tạp.

Một giải pháp tạm thời cho vấn đề này là áp dụng kỹ thuật mà trong đó số lượng tối đa ảnh được khôi phục có thể lưu trữ trong bộ nhớ đệm ảnh (max_dec_pic_buffering) được cố định, không phụ thuộc vào kích thước khung của ảnh được giải mã. Theo giải pháp này, do việc sắp xếp bộ nhớ của bộ đệm ảnh được giải mã (bộ đệm ảnh được giải mã-DPB) được xác định cố định bởi kích thước khung tối đa được chỉ ra trên Fig.2(a), con trỏ bộ nhớ cũng có thể được cố định. Cụ thể, việc điều khiển bộ nhớ trở nên dễ dàng hơn so với cách sắp xếp mà trong đó con trỏ bộ nhớ cần được điều khiển một cách biến đổi. Tuy nhiên, khi kích thước khung của ảnh được giải mã nhỏ hơn so với kích thước khung tối đa hoặc khi bề rộng theo chiều dọc của ảnh bằng một nửa bề rộng theo chiều dọc của khung ảnh như trong trường hợp trường ảnh được xen kẽ, việc thực hiện các giải pháp trên có thể tạo ra các vùng bộ nhớ không sử dụng như thể hiện trong Fig.2(b) và cản trở bộ nhớ được sử dụng một cách hiệu quả. Hơn nữa, giải pháp này có thể lấy đi chỗ trống để nâng cao hiệu quả mã hóa mà có thể được thực hiện bằng cách lưu trữ nhiều ảnh được khôi phục hơn để tăng số lượng ảnh tham chiếu.

Bản chất kỹ thuật của sáng chế

Vì thế, mục đích của sáng chế là để giải quyết vấn đề nêu trên cũng như đề xuất phương pháp, thiết bị, và chương trình mã hóa và giải mã mà trong đó số lượng tối đa của ảnh được khôi phục có thể lưu trữ trong bộ nhớ đệm ảnh được xác định, theo kích thước khung của ảnh được khôi phục, qua đó cho phép sử dụng hiệu quả bộ nhớ đệm ảnh và cải thiện hơn nữa hiệu quả mã hóa.

Giải quyết vấn đề

Để đạt được mục tiêu nêu trên, thiết bị mã hóa dự đoán video theo phương án của sáng chế bao gồm các phương tiện nhập để nhập vào các ảnh tạo thành chuỗi video và phương tiện mã hóa để nén các ảnh, sử dụng dự đoán trong khung hoặc dự đoán liên khung, để tạo ra dữ liệu ảnh nén bao gồm các thông tin về kích thước khung của ảnh đích, số lượng tối đa ảnh được khôi phục có thể lưu trữ, và kích thước khung tối đa được định rõ là ràng buộc mã hóa, và mã hóa các dữ liệu ảnh nén cùng với các dữ liệu đặc tả kỹ thuật mã hóa liên quan đến các thông tin về kích thước khung của ảnh đích, số lượng tối đa ảnh được khôi phục có thể lưu trữ, và kích thước khung tối đa. Thiết bị mã hóa dự đoán video còn bao gồm phương tiện giải mã để giải mã dữ liệu ảnh được nén để khôi phục ảnh, việc lưu trữ ảnh có nghĩa là lưu trữ một hoặc nhiều ảnh được khôi phục như ảnh tham chiếu được sử dụng để mã hóa ảnh tiếp theo, và điều khiển bộ nhớ có nghĩa là điều khiển bộ nhớ đệm ảnh trong phương tiện lưu trữ ảnh, trong đó khi kích thước khung của ảnh đầu vào không vượt quá $1/(2^L)$ (L là số nguyên lớn hơn hoặc bằng một) kích thước khung tối đa của ảnh, phương tiện điều khiển bộ nhớ xác định rằng (2^L) lần của số lượng tối đa ảnh được khôi phục là có thể lưu trữ trong bộ nhớ khung.

Trong thiết bị mã hóa dự đoán video này, số lượng tối đa ảnh được khôi phục có thể lưu trữ trong bộ nhớ đệm ảnh được xác định, dựa trên mối quan hệ giữa kích thước khung hình gắn với ảnh tương ứng tạo thành các chuỗi video hoặc dữ liệu ảnh nén và kích thước khung tối đa được định rõ là ràng buộc mã hóa. Cụ thể hơn, ví dụ, khi kích thước khung của ảnh đầu vào lớn hơn $1/(2^L)$ (L là số nguyên lớn hơn hoặc bằng một) kích thước khung tối đa được định rõ là ràng buộc

mã hóa, số lượng tối đa ảnh khôi phục có thể lưu trữ được thiết lập là N (N là số nguyên lớn hơn hoặc bằng một), trong khi đó khi kích thước khung của ảnh đầu vào không vượt quá $1/(2^L)$ lần kích thước khung tối đa, số lượng bộ nhớ đệm có sẵn để lưu trữ ảnh được khôi phục được điều khiển để cho phép lưu trữ lên đến $((2^L) \times N)$ lần số ảnh được khôi phục. Trong cấu hình này, khi kích thước khung của ảnh được khôi phục không vượt quá $1/(2^L)$ lần kích thước khung tối đa được định rõ là cơ sở mã hóa, số lượng tối đa của ảnh được khôi phục có thể lưu trữ có thể được tăng lên, và do đó số lượng ảnh tham chiếu có thể lưu trữ cũng có thể tăng lên nhằm nâng cao hiệu quả mã hóa.

Trong thiết bị mã hóa dự đoán video theo phương án khác của sáng chế, phương tiện điều khiển bộ nhớ cố định các con trỏ bộ nhớ mà chỉ các vị trí lưu trữ của ảnh được khôi phục trong bộ nhớ khi xác định số lượng tối đa của ảnh được khôi phục có thể lưu trữ trong bộ nhớ khung.

Trong thiết bị mã hóa dự đoán video này, các con trỏ bộ nhớ mà chỉ các vị trí lưu trữ của ảnh được khôi phục tương ứng có thể được cố định, ví dụ, tại vị trí tương đương với kích thước khung tối đa và $1/(2^L)$ lần của kích thước đó. Vì vậy, không cần thiết phải điều khiển để biến đổi con trỏ bộ nhớ phụ thuộc vào kích thước khung hình, và việc thực hiện đơn giản hơn, điều khiển hiệu quả hơn có thể được tiến hành. Vì khoảng cách con trỏ được giới hạn là bội số của $1/(2^L)$, vị trí của con trỏ bộ nhớ có thể được tính toán với thao tác dịch chuyển, và do đó chi phí tính toán có thể được giảm.

Thiết bị giải mã dự đoán video theo phương án của sáng chế bao gồm phương tiện nhập để nhập vào dữ liệu được mã hóa bao gồm dữ liệu ảnh được nén mà có thông tin về kích thước khung của ảnh đích, số lượng tối đa ảnh được khôi phục có thể lưu trữ, và kích thước khung tối đa được định rõ là ràng buộc mã hóa, trong đó các ảnh tạo thành chuỗi video được nén bởi dự đoán trong khung hoặc dự đoán liên khung. Thiết bị giải mã dự đoán video còn bao gồm thiết bị giải mã để giải mã dữ liệu ảnh được nén và dữ liệu ảnh được nén mà tạo thành thông tin về kích thước khung của ảnh đích, số lượng tối đa của ảnh được khôi phục có thể lưu

trữ, và kích thước khung tối đa để khôi phục ảnh và thông tin về kích thước khung ảnh đích, số lượng tối đa của ảnh được khôi phục có thể lưu trữ, và kích thước khung tối đa. Thiết bị còn bao gồm phương tiện lưu trữ ảnh để lưu trữ ảnh được khôi phục làm ảnh tham chiếu được sử dụng để giải mã ảnh tiếp theo, và thiết bị điều khiển bộ nhớ để điều khiển các phương tiện lưu trữ ảnh, trong đó khi kích thước khung của ảnh đầu vào không vượt quá $1/(2^L)$ (L là số nguyên lớn hơn hoặc bằng một) lần kích thước khung tối đa của ảnh, phương tiện điều khiển bộ nhớ xác định rằng (2^L) lần của số lượng tối đa ảnh được khôi phục là có thể lưu trữ trong bộ nhớ khung.

Trong thiết bị giải mã dự đoán video này, số lượng tối đa ảnh được khôi phục có thể lưu trữ trong bộ nhớ đệm ảnh được xác định dựa trên mối quan hệ giữa kích thước khung gắn với ảnh tương ứng tạo thành các chuỗi video hoặc dữ liệu ảnh nén và kích thước khung tối đa được quy định như ràng buộc mã hóa. Cụ thể hơn, ví dụ, khi kích thước khung của ảnh đầu vào lớn hơn $1/(2^L)$ (L là số nguyên lớn hơn hoặc bằng một) lần kích thước khung tối đa được quy định làm ràng buộc mã hóa, số lượng tối đa ảnh khôi phục có thể lưu trữ được thiết lập là N (N là số nguyên lớn hơn hoặc bằng một), trong đó khi kích thước khung của ảnh đầu vào không vượt quá $1/(2^L)$ lần kích thước khung tối đa, số ảnh có thể lưu trữ được xác định lên tới $(2^L) \times N$ lần số ảnh được khôi phục. Trong cấu hình này, khi kích thước khung của ảnh được khôi phục không vượt quá $1/(2^L)$ lần kích thước khung tối đa được định rõ là cơ sở mã hóa, số lượng tối đa của ảnh được khôi phục có thể lưu trữ có thể được tăng lên, và do đó số lượng ảnh tham chiếu có thể lưu trữ cũng có thể tăng lên nhằm nâng cao hiệu quả mã hóa.

Trong thiết bị giải mã dự đoán video theo phương án khác của sáng chế, phương tiện điều khiển bộ nhớ cố định con trỏ bộ nhớ mà chỉ các vị trí lưu trữ của ảnh được khôi phục trong bộ nhớ khung, khi xác định số lượng tối đa của ảnh được khôi phục có thể lưu trữ trong bộ nhớ khung.

Trong thiết bị giải mã dự đoán video này, các con trỏ bộ nhớ mà chỉ các vị trí lưu trữ của ảnh được khôi phục tương ứng có thể được cố định, ví dụ, tại vị trí

tương đương với kích thước khung tối đa và $1/(2^L)$ lần của kích thước đó. Vì vậy, không cần thiết phải điều khiển để biến đổi con trỏ bộ nhớ phụ thuộc vào kích thước khung hình, và việc thực hiện đơn giản hơn, điều khiển hiệu quả hơn có thể được tiến hành. Vì khoảng cách con trỏ được giới hạn là bội số của $1/(2^L)$, vị trí của con trỏ bộ nhớ có thể được tính toán với thao tác dịch chuyển, và do đó chi phí tính toán có thể được giảm.

Phương pháp mã hóa dự đoán video theo phương án của sáng chế được thực hiện bởi thiết bị mã hóa dự đoán video. Phương pháp này bao gồm bước nhập để nhập các ảnh tạo thành chuỗi video, bước mã hóa để mã hóa ảnh bằng dự đoán trong khung hoặc dự đoán liên khung để tạo ra dữ liệu ảnh nén chứa thông tin về kích thước khung của ảnh đích, số lượng tối đa ảnh được khôi phục có thể lưu trữ, và kích thước khung tối đa được quy định làm ràng buộc mã hóa, trong đó dữ liệu ảnh nén được nén cùng với dữ liệu đặc tả kỹ thuật mã hóa liên quan đến thông tin về kích thước khung của ảnh đích, số lượng tối đa ảnh được khôi phục có thể lưu trữ, và kích thước khung tối đa. Phương pháp này còn bao gồm bước giải mã để giải mã dữ liệu ảnh được nén để khôi phục ảnh, bước lưu trữ ảnh để lưu trữ một hoặc nhiều ảnh được khôi phục như là ảnh tham chiếu được sử dụng để mã hóa ảnh tiếp theo, và bước điều khiển bộ nhớ để điều khiển bộ nhớ đệm ảnh trong bước lưu trữ ảnh, trong đó khi kích thước khung của ảnh đầu vào không vượt quá $1/(2^L)$ (L là số nguyên lớn hơn hoặc bằng một) lần kích thước khung tối đa của ảnh, bước điều khiển bộ nhớ xác định rằng (2^L) lần của số lượng tối đa ảnh được khôi phục là có thể lưu trữ trong bộ nhớ khung.

Chương trình mã hóa dự đoán video theo phương án của sáng chế bao gồm các môđun nhập để nhập vào các ảnh tạo thành chuỗi video và môđun mã hóa để mã hóa ảnh bằng dự đoán trong khung hoặc dự đoán liên khung để tạo ra dữ liệu ảnh nén chứa thông tin về kích thước khung của ảnh đích, số lượng tối đa ảnh được khôi phục có thể lưu trữ, và kích thước khung tối đa được quy định làm ràng buộc mã hóa, và mã hóa các dữ liệu ảnh nén cùng với các dữ liệu đặc tả kỹ thuật mã hóa liên quan đến các thông tin về kích thước khung của ảnh đích, số lượng tối đa ảnh

được khôi phục có thể lưu trữ, và kích thước khung tối đa. Chương trình này còn bao gồm môđun giải mã để giải mã dữ liệu ảnh được nén để khôi phục ảnh, môđun lưu trữ ảnh để lưu trữ một hoặc nhiều hơn các ảnh được khôi phục như là ảnh tham chiếu được sử dụng để mã hóa ảnh tiếp theo, và môđun điều khiển bộ nhớ để điều khiển bộ nhớ đệm ảnh trong môđun lưu trữ ảnh, trong đó khi kích thước khung của ảnh đầu vào không vượt quá $1/(2^L)$ (L là số nguyên lớn hơn hoặc bằng một) lần kích thước khung tối đa của ảnh, môđun điều khiển bộ nhớ xác định rằng (2^L) lần của số lượng tối đa ảnh được khôi phục là có thể lưu trữ trong bộ nhớ khung.

Phương pháp mã hóa dự đoán video và chương trình mã hóa dự đoán video có thể đạt được hiệu quả như thiết bị mã hóa dự đoán video được nhắc tới trong phần trước.

Phương pháp giải mã dự đoán video theo phương án của sáng chế được thực hiện bởi thiết bị giải mã dự đoán video. Phương pháp này bao gồm bước nhập để nhập vào dữ liệu được mã hóa bao gồm dữ liệu ảnh được nén lại mà có thông tin về kích thước khung của ảnh đích, số lượng tối đa ảnh được khôi phục có thể lưu trữ, và kích thước khung tối đa được quy định làm ràng buộc mã hóa, trong đó các ảnh tạo thành chuỗi video được nén bởi dự đoán trong khung hoặc dự đoán liên khung. Phương pháp còn bao gồm bước giải mã để giải mã dữ liệu ảnh được nén và dữ liệu ảnh được nén mà tạo thành thông tin về kích thước khung của ảnh đích, số lượng tối đa của ảnh được khôi phục có thể lưu trữ, và kích thước khung tối đa để khôi phục ảnh và thông tin về kích thước khung ảnh đích, số lượng tối đa của ảnh được khôi phục có thể lưu trữ, và kích thước khung tối đa. Phương pháp còn bao gồm bước lưu trữ ảnh để lưu trữ, trong phương tiện lưu trữ ảnh, một hoặc nhiều hơn ảnh được khôi phục như là ảnh tham chiếu được sử dụng để giải mã ảnh tiếp theo, và bước điều khiển bộ nhớ để điều khiển phương tiện lưu trữ ảnh, trong đó khi kích thước khung của ảnh đầu vào không vượt quá $1/(2^L)$ (L là số nguyên lớn hơn hoặc bằng một) lần kích thước khung tối đa của ảnh, bước điều khiển bộ nhớ xác định rằng (2^L) lần của số lượng tối đa ảnh được khôi phục là có thể lưu trữ trong bộ nhớ khung.

Chương trình giải mã dự đoán video theo phương án của sáng chế bao gồm môđun nhập để nhập vào dữ liệu được mã hóa bao gồm dữ liệu ảnh được nén mà có thông tin về kích thước khung của ảnh đích, số lượng tối đa ảnh được khôi phục có thể lưu trữ, và kích thước khung tối đa được quy định làm ràng buộc mã hóa, trong đó các ảnh tạo thành chuỗi video được nén bởi dự đoán trong khung hoặc dự đoán liên khung. Chương trình còn bao gồm môđun giải mã để giải mã dữ liệu ảnh được nén và dữ liệu ảnh được nén đó tạo thành thông tin về kích thước khung của ảnh đích, số lượng tối đa của ảnh được khôi phục có thể lưu trữ, và kích thước khung tối đa để khôi phục ảnh và thông tin về kích thước khung ảnh đích, số lượng tối đa của ảnh được khôi phục có thể lưu trữ, và kích thước khung tối đa. Chương trình còn bao gồm môđun lưu trữ ảnh để lưu trữ một hoặc nhiều hơn ảnh được khôi phục làm ảnh tham chiếu được sử dụng để giải mã ảnh tiếp theo, và môđun điều khiển bộ nhớ để điều khiển môđun lưu trữ ảnh, trong đó khi kích thước khung của ảnh đầu vào không vượt quá $1/(2^L)$ (L là số nguyên lớn hơn hoặc bằng một) lần kích thước khung tối đa của ảnh, môđun điều khiển bộ nhớ xác định rằng (2^L) lần của số lượng tối đa ảnh được khôi phục là có thể lưu trữ trong bộ nhớ khung.

Phương pháp giải mã dự đoán video và chương trình giải mã dự đoán video có thể đạt được hiệu quả như thiết bị giải mã dự đoán video được nhắc tới trong phần nêu trên.

Hiệu quả sáng chế

Thiết bị, phương pháp và chương trình mã hóa và giải mã dự đoán video theo sáng chế thực hiện việc xác định số lượng tối đa ảnh được khôi phục có thể lưu trữ trong bộ nhớ đệm ảnh, phù hợp với kích thước khung của ảnh được khôi phục, nhờ đó đạt được hiệu quả sử dụng của bộ nhớ đệm ảnh đồng thời cũng nâng cao hiệu quả mã hóa.

Mô tả vắn tắt các hình vẽ

Fig.1(a) và Fig.1(b) là các hình vẽ thể hiện bộ nhớ đệm ảnh, số lượng ảnh có thể lưu trữ, và vị trí của con trỏ bộ nhớ biến đổi.

Fig.2(a) và Fig.2(b) là các hình vẽ giải thích vấn đề của con trỏ bộ nhớ cố

định trong bộ nhớ đệm ảnh.

Fig.3 là sơ đồ khối thể hiện thiết bị mã hóa dự đoán video theo phương án của sáng chế.

Fig.4 là sơ đồ khối thể hiện thiết bị giải mã dự đoán video theo phương án của sáng chế.

Fig.5 là sơ đồ trình tự thể hiện phương pháp mã hóa và giải mã dự đoán video đầu tiên theo phương án của sáng chế.

Fig.6(a) và Fig.6(b) là hình vẽ biểu đồ để giải thích quá trình xử lý của phương pháp mã hóa và giải mã dự đoán video theo phương án của sáng chế.

Fig.7 là hình vẽ thể hiện cấu trúc phần cứng của máy tính để thực hiện chương trình được ghi lại trên vật ghi.

Fig.8 là hình phối cảnh của máy tính để thực hiện chương trình được ghi lại trên vật ghi.

Mô tả chi tiết sáng chế

Các phương án của sáng chế sẽ được mô tả dựa vào các hình vẽ từ Fig.3 đến Fig.8 dưới đây.

Fig.3 là sơ đồ khối thể hiện thiết bị mã hóa dự đoán video theo phương án của sáng chế. Số chỉ dẫn 301 biểu thị thiết bị đầu cuối nhập, 302 biểu thị bộ chia khối, 303 biểu thị bộ tạo tín hiệu dự đoán, 304 biểu thị bộ nhớ khung, 305 biểu thị bộ trừ, 306 biểu thị bộ biến đổi, 307 biểu thị bộ lượng tử hóa, 308 biểu thị bộ giải lượng tử hóa, 309 biểu thị bộ biến đổi ngược, 310 biểu thị bộ cộng, 311 biểu thị bộ mã hóa entropi, 312 biểu thị thiết bị đầu cuối đầu ra, và 314 biểu thị bộ quản lý bộ nhớ khung (hoặc bộ quản lý bộ đệm). Thiết bị đầu cuối nhập 301 cấu thành phương tiện nhập. Bộ tạo tín hiệu dự đoán 303, bộ trừ 305, bộ biến đổi 306, bộ lượng tử hóa 307, và bộ mã hóa entropi 311 cấu thành phương tiện mã hóa. Bộ giải lượng tử hóa 308, bộ biến đổi ngược 309, và bộ cộng 310 cấu thành phương tiện giải mã. Bộ nhớ khung 304 cấu thành phương tiện lưu trữ ảnh. Bộ quản lý bộ nhớ khung 314 cấu thành phương tiện điều khiển bộ nhớ.

Hoạt động của thiết bị mã hóa dự đoán video được cấu hình như mô tả nêu

trên sẽ được mô tả dưới đây. Tín hiệu video biểu diễn các ảnh được đưa vào thiết bị đầu cuối nhập 301. Ảnh được mã hóa ("ảnh đích") được chia thành các vùng bởi bộ chia khối 302. Trong phương án của sáng chế, ảnh đích được chia thành khối mà mỗi khối có 8×8 điểm ảnh, nhưng nó có thể được chia thành khối với bất kỳ kích thước hoặc hình dạng nào khác với loại ở trên. Tín hiệu dự đoán sau đó được tạo ra cho vùng cần được mã hóa (sau đây được gọi là khối đích). Phương án của sáng chế áp dụng hai loại phương thức dự đoán. Cụ thể, chúng là dự đoán liên khung và dự đoán trong khung.

Trong dự đoán liên khung, ảnh được khôi phục mà được mã hóa và sau đó được phục hồi được sử dụng để làm ảnh tham chiếu, và thông tin về chuyển động mà cung cấp tín hiệu dự đoán với lỗi nhỏ nhất từ khối đích được xác định từ ảnh tham chiếu. Quá trình này được gọi là quá trình phát hiện chuyển động. Tùy vào tình huống, có thể được phép chia khối đích thành các tiểu vùng và thực hiện dự đoán liên khung trên mỗi tiểu vùng này. Trong quá trình này, phương pháp phân chia mà đưa ra kết quả hiệu quả nhất cho tổng thể khối đích được lựa chọn từ các phương pháp phân chia khác nhau, và thông tin về chuyển động của từng tiểu vùng được xác định. Theo phương án của sáng chế, hoạt động này được thực hiện trong bộ tạo tín hiệu dự đoán 303, mà trong đó khối đích được đưa vào thông qua đường L302, và ảnh tham chiếu được đưa vào thông qua L304. Ảnh tham chiếu được sử dụng ở đây là các ảnh mà đã được mã hóa và phục hồi. Chi tiết của hoạt động này giống với kỹ thuật thông thường được áp dụng trong MPEG-2 hoặc 4 và H.264. Thông tin về chuyển động và phương pháp phân chia tiểu vùng được xác định như mô tả ở trên được đưa vào đường L321 tới bộ mã hóa entropi 311, được mã hóa tại đó và sau đó đưa ra từ thiết bị đầu cuối đầu ra 312. Thông tin (danh mục tham chiếu) xác định ảnh tham chiếu giữa các ảnh tham chiếu mà từ đó tín hiệu dự đoán được tạo ra cũng được gửi qua đường L312 tới bộ mã hóa entropi 311. Trong phương án của sáng chế, M ảnh được khôi phục (M là số nguyên bằng hoặc lớn hơn một) được lưu trữ trong bộ nhớ khung 304 và sẽ được sử dụng làm ảnh tham chiếu. Bộ tạo tín hiệu dự đoán 303 thu tín hiệu ảnh tham chiếu từ bộ nhớ khung

304, dựa vào ảnh tham chiếu và thông tin về chuyển động, phù hợp với phương pháp phân chia tiểu vùng và từng tiểu vùng, và tạo ra tín hiệu dự đoán. Tín hiệu dự đoán liên khung được tạo ra theo cách này được đưa vào đường L303 tới bộ trừ 305.

Trong dự đoán trong khung, tín hiệu dự đoán trong khung được tạo ra sử dụng giá trị điểm ảnh được khôi phục trước đó liền kề với khối đích. Đặc biệt, bộ tạo tín hiệu dự đoán 303 thu tín hiệu điểm ảnh khôi phục trước đó trong cùng một khung từ bộ nhớ khung 304 và ngoại suy các tín hiệu này để tạo ra tín hiệu dự đoán trong khung. Thông tin về phương pháp ngoại suy được đưa vào đường L321 tới bộ mã hóa entropi 311, được mã hóa tại đó và sau đó đưa ra từ thiết bị đầu cuối đầu ra 312. Tín hiệu dự đoán trong khung được tạo ra theo cách này được đưa vào bộ trừ 305. Phương pháp tạo tín hiệu dự đoán trong khung trong bộ tạo tín hiệu dự đoán 303 là kỹ thuật thông thường được áp dụng trong H.264. Một trong hai tín hiệu dự đoán liên khung và tín hiệu dự đoán trong khung được thu lại như mô tả ở trên, mà tạo ra lỗi nhỏ hơn, sẽ được chọn lựa, và tín hiệu dự đoán được chọn lựa được đưa vào bộ trừ 305.

Do không có ảnh trước ảnh đầu tiên, tất cả khối đích đó sẽ được xử lý bằng dự đoán trong khung.

Bộ trừ 305 trừ đi các tín hiệu dự đoán (đưa vào theo đường L303) từ tín hiệu của khối đích (đưa vào theo đường L302) để tạo ra tín hiệu dư. Tín hiệu dư này là đối tượng của biến đổi cosin rời rạc của bộ biến đổi 306, và sau đó hệ số biến đổi thu được được lượng tử hóa bởi bộ lượng tử hóa 307. Cuối cùng, bộ mã hóa entropi 311 mã hóa hệ số biến đổi được lượng tử hóa, và dữ liệu được mã hóa được đưa ra từ thiết bị đầu cuối đầu ra 312, cùng với thông tin về phương pháp dự đoán.

Để thực hiện dự đoán trong khung hoặc dự đoán liên khung trên khối đích tiếp theo, các tín hiệu được nén của khối đích được xử lý ngược và phục hồi. Cụ thể, hệ số biến đổi được lượng tử hóa được lượng tử hóa ngược lại bởi bộ giải lượng tử hóa 308 và sau đó được thực hiện phép biến đổi cosin rời rạc nghịch đảo

bởi bộ biến đổi ngược 309 để phục hồi tín hiệu dư. Bộ cộng 310 thêm tín hiệu dư được phục hồi vào tín hiệu dự đoán được đưa vào qua đường L303 để khôi phục tín hiệu của khối đích, mà được lưu trữ trong bộ nhớ khung 304. Mặc dù phương án này sử dụng bộ biến đổi 306 và bộ biến đổi ngược 309, có thể sử dụng loại xử lý biến đổi khác thay vì các bộ biến đổi này. Tùy vào tình huống, bộ biến đổi 306 và bộ biến đổi ngược 309 có thể bị loại bỏ.

Bộ nhớ khung 304 là bộ lưu trữ hữu hạn và không thể lưu trữ tất cả ảnh được phục hồi. Chỉ những ảnh được khôi phục để sử dụng cho mã hóa ảnh tiếp theo được lưu trữ trong bộ nhớ khung 304. Bộ quản lý bộ nhớ khung 314 là bộ mà điều khiển bộ nhớ khung 304 này. Bộ quản lý bộ nhớ khung 314 điều khiển bộ nhớ khung 304 nhằm xóa đi ảnh cũ nhất trong số M ảnh được khôi phục (M là số nguyên lớn hơn hoặc bằng một) được lưu trữ trong bộ nhớ khung 304. Thông qua thiết bị đầu cuối nhập 313, kích thước khung của từng ảnh và kích thước khung tối đa được quy định làm ràng buộc mã hóa (hiện trạng và cấp độ) được đưa vào, dựa vào đó số lượng tối đa ảnh được khôi phục có thể lưu trữ trong bộ nhớ khung được xác định, và bộ quản lý bộ nhớ khung 314 tiến hành lưu trữ ảnh được khôi phục trong phạm vi số lượng tối đa. Đồng thời, thông tin về kích thước khung của từng ảnh, kích thước khung tối đa được quy định làm ràng buộc mã hóa, và số lượng tối đa của ảnh được khôi phục có thể lưu trữ trong bộ nhớ khung được đưa vào qua đường L314 tới bộ mã hóa entropi 311, được mã hóa tại đó, và đưa ra cùng với dữ liệu ảnh nén. Kích thước khung, kích thước khung tối đa được quy định làm ràng buộc mã hóa, và số lượng tối đa ảnh được khôi phục có thể lưu trữ trong bộ nhớ khung được kết hợp với từng ảnh. Kích thước khung có thể được thể hiện bởi giá trị kích thước khung thô hoặc được thể hiện dưới dạng số mũ hoặc lũy thừa. Nếu kích thước khung tối đa được chỉ ra trước dưới dạng là thông tin hiện trạng và cấp độ, kích thước khung tối đa có thể được đưa vào sử dụng thông tin đó. Cần chú ý rằng trong phương án này, giá trị kích thước khung được trực tiếp chuyển thành dữ liệu nhị phân, và kích thước khung tối đa được định rõ là thông tin cấp độ. Cách thức bộ quản lý bộ nhớ khung 314 được điều khiển sẽ được mô tả

sau.

Tiếp theo, phương pháp giải mã dự đoán video sẽ được mô tả. Fig.4 là sơ đồ khối thể hiện thiết bị giải mã dự đoán video theo phương án của sáng chế. Số chỉ dẫn 401 biểu thị thiết bị đầu cuối nhập, 402 biểu thị bộ phân tích dữ liệu, 403 biểu thị bộ giải lượng tử hóa, 404 biểu thị bộ biến đổi ngược, 405 biểu thị bộ cộng, 408 biểu thị bộ tạo tín hiệu dự đoán, 407 biểu thị bộ nhớ khung, 406 biểu thị thiết bị đầu cuối đầu ra, và 409 biểu thị bộ quản lý bộ nhớ khung. Thiết bị đầu cuối nhập 401 cấu thành phương tiện nhập. Bộ giải lượng tử hóa 403 và bộ biến đổi ngược 404 cấu thành phương tiện giải mã. Một số loại phương tiện giải mã khác cũng có thể được áp dụng thay cho loại kể trên. Bộ nhớ khung 407 cấu thành phương tiện lưu trữ ảnh. Bộ quản lý bộ nhớ khung 409 cấu thành phương tiện điều khiển bộ nhớ. Thêm vào đó, phương tiện giải mã có thể được cấu hình mà không có bộ biến đổi ngược 404.

Hoạt động của thiết bị giải mã dự đoán video được cấu hình như mô tả ở trên sẽ được mô tả dưới đây. Dữ liệu nén được mã hóa như mô tả ở trên được đưa vào thiết bị đầu cuối nhập 401. Dữ liệu nén chứa tín hiệu dư có được từ việc mã hóa dự đoán của từng khối đích thu được thông qua việc phân chia ảnh thành các khối, và thông tin về việc tạo thành tín hiệu dự đoán. Thông tin về việc tạo thành tín hiệu dự đoán bao gồm thông tin về việc chia khối (kích thước khối), thông tin về chuyển động, và danh mục tham chiếu nhắc tới ở trên khi mà dự đoán liên khung được thực hiện, hoặc bao gồm thông tin về phương pháp ngoại suy được thực hiện trên điểm ảnh bao quanh được khôi phục khi mà dự đoán trong khung được thực hiện.

Bộ phân tích dữ liệu 402 trích ra, từ dữ liệu được nén, tín hiệu dư của khối đích, thông tin về việc tạo thành tín hiệu dự đoán, tham số lượng tử hóa, kích thước khung của ảnh, kích thước khung tối đa được định rõ là ràng buộc mã hóa, và số lượng tối đa ảnh được khôi phục có thể lưu trữ trong bộ nhớ khung. Tín hiệu dư của khối đích được lượng tử hóa ngược với tham số lượng tử hóa (đưa vào theo đường L402) bởi bộ giải lượng tử hóa 403. Kết quả được thực hiện biến đổi cosin

rời rạc ngược bởi bộ biến đổi ngược 404.

Tiếp theo, thông tin về việc tạo thành tín hiệu dự đoán được đưa vào qua đường L406b tới bộ tạo tín hiệu dự đoán 408. Bộ tạo tín hiệu dự đoán 408 truy cập bộ nhớ khung 407, dựa vào thông tin về việc tạo tín hiệu dự đoán, để thu tín hiệu tham chiếu từ các ảnh tham chiếu nhằm tạo ra tín hiệu dự đoán. Tín hiệu dự đoán này được đưa vào qua đường L408 tới bộ cộng 405. Bộ cộng 405 thêm tín hiệu dự đoán vào tín hiệu dự được phục hồi để khôi phục tín hiệu khối đích, mà được đưa ra thông qua đường L405 và đồng thời được lưu trữ vào trong bộ nhớ khung 407.

Chỉ những ảnh được khôi phục để sử dụng cho giải mã ảnh tiếp theo được lưu trữ trong bộ nhớ khung 407. Bộ quản lý bộ nhớ khung 409 điều khiển bộ nhớ khung 407. Bộ nhớ khung 407 được điều khiển để xóa đi ảnh cũ nhất trong số M ảnh khôi phục được lưu trữ (M là số nguyên) theo đó ảnh được khôi phục mới nhất được sử dụng làm ảnh tham chiếu có thể được lưu trữ. Thông tin về kích thước khung của ảnh đích, kích thước khung tối đa được định rõ là ràng buộc mã hóa, và số lượng tối đa của ảnh được khôi phục có thể lưu trữ trong bộ nhớ khung được đưa vào qua đường L406 tới bộ quản lý bộ nhớ khung 409, dựa vào đó bộ quản lý bộ nhớ khung 409 hoạt động. Cách thức bộ quản lý bộ nhớ khung 409 được điều khiển sẽ được mô tả sau.

Hoạt động của phương pháp mã hóa dự đoán video và phương pháp giải mã dự đoán video sẽ được mô tả dưới đây thông qua Fig.5 và Fig.6. Fig.5 là sơ đồ chuỗi thể hiện phương pháp mã hóa và giải mã dự đoán video theo phương án của sáng chế. Các biến ghi trong cùng bản vẽ sẽ được giải thích. Pic_width là biến biểu thị số điểm ảnh sáng của bề rộng theo chiều ngang của ảnh được khôi phục, pic_height là biến biểu thị số điểm ảnh sáng của độ rộng theo chiều dọc của ảnh được khôi phục, MaxLumaFS là biến biểu thị kích thước ảnh tối đa được xác định bởi thông tin về cấp độ, MFSBuffer là biến biểu thị số lượng bộ nhớ đệm biểu hiện bởi số lượng tối đa của ảnh được khôi phục có thể lưu trữ trong bộ đệm ảnh được giải mã (DPB-decoded picture buffer), mà được xác định là thông tin cấp độ, maxDPBsize là biến biểu thị số lượng bộ nhớ đệm biểu hiện bởi số lượng tối đa

ảnh được khôi phục có thể lưu trữ trong bộ đệm ảnh được giải mã (DPB-decoded picture buffer), `max_dec_pic_buffering` là biến biểu thị số lượng lưu trữ tối đa (số lượng bộ nhớ đệm) của ảnh được khôi phục được sử dụng trong giải mã, và `DPBpointer` là biến biểu thị con trỏ chỉ ra vị trí lưu trữ của ảnh được phục hồi trong bộ đệm ảnh được giải mã.

Trên Fig.5, đầu vào là `pic_width` biểu thị số điểm ảnh sáng của bề rộng theo chiều ngang của ảnh được khôi phục, `pic_height` biểu thị số điểm ảnh sáng của bề rộng theo chiều dọc của ảnh được khôi phục, `MaxLumaFS` biểu thị kích thước khung tối đa được định rõ là ràng buộc mã hóa, và `MFSBuffer` biểu thị số lượng bộ nhớ đệm biểu hiện bởi số lượng tối đa của ảnh được khôi phục có thể lưu trữ trong bộ nhớ đệm ảnh được giải mã (DPB). Tiếp theo, tích của `pic_width` biểu thị số lượng điểm ảnh sáng của bề rộng theo chiều ngang của ảnh được khôi phục và `pic_height` biểu thị số điểm ảnh sáng của bề rộng theo chiều dọc của ảnh được khôi phục trong dữ liệu đầu vào được so sánh với một nửa ($\text{MaxLumaFS}/2$) của `MaxLumaFS` biểu thị kích thước ảnh tối đa được định rõ là ràng buộc mã hóa (bước S502). Nếu điều kiện không được thỏa mãn (hoặc nếu tích của `pic_width` và `pic_height` lớn hơn $\text{MaxLumaFS}/2$), `maxDPBsize = MFSBuffer` được thiết lập (bước S503). Nếu điều kiện được thỏa mãn (hoặc nếu tích của `pic_width` và `pic_height` không lớn hơn $\text{MaxLumaFS}/2$), `maxDPBsize = 2*MFSBuffer` được thiết lập (bước S504).

Thao tác

```
maxDPBsize = 2*MFSBuffer (if pic_width*pic_height ≤ (MaxLumaFS >>
1))
```

```
maxDPBsize = MFSBuffer (otherwise)
```

Trong thao tác ở trên ">>" biểu hiện thao tác dịch phải và có cùng ý nghĩa với $\text{MaxLumaFS}/2$.

Sau đó, số lượng bộ nhớ đệm (`max_dec_pic_buffering`) được thể hiện bởi số lượng lưu trữ tối đa của ảnh được khôi phục được sử dụng trong giải mã được thiết lập không cao hơn `maxDPBsize`.

Thao tác

$$\text{max_dec_pic_buffering} \leq \text{maxDPBsize}$$

Nếu $\text{maxDPBsize} = \text{MFSBuffer}$, như trong Fig.6(a), $\text{DPBpointer} = 1$ được thiết lập làm mốc đầu của bộ nhớ đệm ảnh; và $\text{DPBpointer} = 2, 3$, và 4 được thiết lập lần lượt tại các vị trí cách vị trí $\text{DPBpointer} = 1$ nhân với hai lần, ba lần và bốn lần kích thước khung tối đa.

Nếu $\text{maxDPBsize} = 2 * \text{MFSBuffer}$, như thể hiện trong Fig.6(b), ngoài con trỏ DPB 1 tới 4 đã được thiết lập, con trỏ DPB xa hơn được thiết lập lần lượt tại các vị trí $\text{DPBpointer} = 1$ nhân với $1/2$ lần, $3/2$ lần, $5/2$ lần, và $7/2$ lần kích thước khung tối đa (bước S505).

Con trỏ $\text{max_dec_pic_buffering}$ và DPB đã thiết lập được sử dụng trong hoạt động điều khiển của bộ nhớ khung trong bước S508 như mô tả dưới đây. Cụ thể, trong giải mã, ảnh được khôi phục được lưu trữ trong bộ nhớ khung trong phạm vi số lượng tối đa của $\text{max_dec_pic_buffering}$, và khi được lưu trữ, ảnh được khôi phục được điều khiển để được lưu trữ dựa trên con trỏ DPB như thiết lập ở trên. Trong mã hóa, thông tin được đưa ra từ bên ngoài mà có chứa thông tin về pic_width biểu thị số điểm ảnh sáng của bề rộng theo chiều ngang của ảnh được khôi phục và pic_height biểu thị số điểm ảnh sáng của bề rộng theo chiều dọc của ảnh được khôi phục và MaxLumaFS biểu thị kích thước ảnh tối đa và MFSBuffer biểu thị số lượng bộ nhớ đệm biểu hiện bởi số lượng tối đa của ảnh được khôi phục có thể lưu trữ trong bộ nhớ đệm ảnh được giải mã (DPB). Trên Fig.3, thông tin được đưa vào qua thiết bị đầu cuối nhập 313 từ thiết bị điều khiển mà không được thể hiện.

Mặt khác, trong bước S506, ảnh đích được mã hóa và sau đó giải mã bằng phương pháp được mô tả với Fig.3. Dữ liệu ảnh được mã hóa được đưa ra bên ngoài hoặc được lưu trữ. Trong bước S507 xác định liệu ảnh đích có được sử dụng làm ảnh tham chiếu trong quá trình xử lý tiếp theo hay không. Việc xác định trong bước S507 phụ thuộc vào cách thức ảnh được mã hóa (mã hóa dự đoán trong khung, mã hóa dự đoán liên khung, hay mã hóa dự đoán hai chiều). Khi ảnh đích

không được sử dụng làm ảnh tham chiếu, quy trình đi tới bước S510. Khi ảnh đích được sử dụng làm ảnh tham chiếu, quy trình đi tới bước S508, tại đó con trỏ DPB được thiết lập tới vị trí được thể hiện trong Fig.6(a) hoặc Fig.6(b). Sau đó, ảnh được giải mã và ảnh được khôi phục được lưu trữ trong bộ nhớ khung trong bước S509, và quy trình đi tới bước S511. Nếu có ảnh kế tiếp trong bước S511, quy trình tiếp tục tới S506, tại đó nếu không có ảnh kế tiếp, quy trình sẽ kết thúc. Quy trình mã hóa được thực hiện thông qua ảnh cuối cùng trong cách này.

Các quy trình được mô tả ở trên được thực hiện trong toàn bộ thiết bị mã hóa video thể hiện trong Fig.3. Đặc biệt, các bước S502, S503, S504, S505, và S508 được thực hiện trong bộ quản lý bộ nhớ khung 314.

Quy trình được thể hiện trên Fig.5 được mô tả ở trên là phương pháp mã hóa video, nhưng có thể áp dụng cho quy trình của phương pháp giải mã video. Khi giải mã được thực hiện, trong bước S501, dữ liệu ảnh được nén và mã hóa (dòng bit) được đưa vào. Dữ liệu được trích từ dữ liệu đầu vào pic_width biểu thị số điểm ảnh sáng của bề rộng theo chiều ngang của ảnh được khôi phục, pic_height biểu thị số điểm ảnh sáng của bề rộng theo chiều dọc của ảnh được khôi phục, MaxLumaFS biểu thị kích thước khung tối đa được định rõ là ràng buộc mã hóa, MFSBuffer biểu thị số lượng bộ nhớ đệm biểu hiện bởi số lượng tối đa của ảnh được khôi phục có thể lưu trữ trong bộ nhớ đệm ảnh được giải mã (DPB), và max_dec_pic_buffering biểu thị số lượng lưu trữ tối đa (số lượng bộ nhớ đệm) của ảnh được khôi phục được sử dụng trong giải mã, và các bước S502-S505 và S508 được thực hiện như mô tả ở trên. Về mặt giải mã, trong bước S505 kiểm tra liệu max_dec_pic_buffering được lưu trữ có nhiều hơn maxDPBsize hay không. Trong bước S506, dữ liệu được nén của ảnh đích được giải mã để phục hồi ảnh. Bước S507 và các bước tiếp sau được thực hiện như mô tả ở trên. Các quy trình này được thực hiện trong thiết bị mã hóa video thể hiện trong Fig.4. Đặc biệt, các bước S502, S503, S504, S505, và S508 được thực hiện trong bộ quản lý bộ nhớ khung 409 của thiết bị giải mã video.

Fig.6(a) và Fig.6(b) là sơ đồ mạch khái lược để giải thích vị trí của con trỏ

bộ nhớ trong bộ nhớ khung được thực hiện trong phương pháp mã hóa và giải mã dự đoán video theo phương án của sáng chế. So sánh giữa các Fig.6(a) và 6(b) cho thấy vị trí của con trỏ bộ nhớ 1 tới 4 là cố định. Khi điều kiện trong bước S503 được thỏa mãn, con trỏ bộ nhớ mới 5 tới 8 được thêm vào. Cụ thể, không phụ thuộc vào kết quả xác định trong bước S503, vị trí con trỏ bộ nhớ là cố định.

Vị trí của con trỏ bộ nhớ

Trong phương án này, con trỏ bộ nhớ 1 tới 4 và 5 tới 8 được thiết lập luân phiên, nhưng con trỏ bộ nhớ có thể được gán các số gia tăng từ 1 tới 8.

Số lượng phân chia bộ nhớ

Trong phương án này, `pic_width` biểu thị số lượng điểm ảnh sáng của bề rộng theo chiều ngang của ảnh được khôi phục và `pic_height` biểu thị số điểm ảnh sáng của bề rộng theo chiều dọc của ảnh được khôi phục được so sánh với một nửa ($\text{MaxLumaFS}/2$) của `MaxLumaFS` biểu thị kích thước ảnh tối đa được định rõ là ràng buộc mã hóa. Cần chú ý rằng `pic_width` biểu thị số lượng điểm ảnh sáng của bề rộng theo chiều ngang của ảnh được khôi phục và `pic_height` biểu thị số điểm ảnh sáng của bề rộng theo chiều dọc của ảnh được khôi phục có thể được so sánh với kích thước khung tối đa $\text{MaxLumaFS}/(2^L)$ (với L là số nguyên tố lớn hơn hoặc bằng 2) được định rõ là ràng buộc mã hóa. Khi được so sánh như vậy, vị trí của con trỏ bộ nhớ chỉ ra vị trí lưu trữ của ảnh được khôi phục có thể được thiết lập để con trỏ bộ nhớ được sắp xếp tại vị trí được xác định bởi $1/(2^L)$ của `MaxLumaFS` trong khi con trỏ đã được sắp xếp là cố định, như thể hiện trong Fig.6(a) và Fig.6(b).

Định nghĩa kích thước khung

Trong phương án này, kích thước khung được thể hiện bằng `pic_width` biểu thị số lượng điểm ảnh sáng của bề rộng theo chiều ngang của ảnh được khôi phục và `pic_height` biểu thị số điểm ảnh sáng của bề rộng theo chiều dọc của ảnh được khôi phục. Tuy nhiên, cần chú ý rằng cũng có thể sử dụng giá trị biểu thị của giá trị đã được nhân lên mà có nghĩa là kích thước khung. Kích thước khung cũng có thể được tính toán bằng phương pháp khác với phương pháp kể trên.

Trong phương án này, MaxLumaFS biểu thị kích thước ảnh tối đa và MFSBuffer biểu thị số lượng bộ nhớ đệm được thể hiện bởi số lượng tối đa của ảnh được khôi phục có kích thước MaxLumaFS có thể lưu trữ trong bộ đệm ảnh được giải mã (DPB) được định rõ là thông tin về cấp độ. Tuy nhiên cần chú ý rằng chúng không cần được truyền đi như là thông tin cấp độ. Chúng có thể được thêm vào như là thông tin khác không phải thông tin cấp độ.

Trong phương án này, FSBuffer và maxDPBsize được biểu hiện bởi số lượng tối đa ảnh được khôi phục có thể lưu trữ. Tuy nhiên, nên chú ý rằng chúng có thể được biểu hiện là số lượng bộ nhớ thực tế. Nếu được biểu hiện như vậy, số lượng ảnh được khôi phục được tính toán bằng cách chia số lượng bộ nhớ cho kích thước khung.

Do số lượng tối đa của ảnh được khôi phục có thể lưu trữ trong bộ nhớ đệm ảnh được xác định, dựa vào mối quan hệ giữa kích thước khung của ảnh và kích thước khung tối đa được định rõ là ràng buộc mã hóa, khi kích thước khung thực tế nhỏ hơn kích thước khung tối đa, lãng phí trong bộ nhớ khung có thể được giảm đi, và khu vực bộ nhớ để lưu trữ ảnh tham chiếu có thể được tăng lên do đó cho phép cải thiện hơn nữa hiệu quả mã hóa.

Chương trình mã hóa dự đoán video và chương trình giải mã dự đoán video để máy tính thực hiện chức năng như là thiết bị mã hóa dự đoán video và thiết bị giải mã dự đoán video có thể được cung cấp như là chương trình được lưu trữ trong vật ghi. Ví dụ về vật ghi như vậy bao gồm vật ghi như đĩa mềm (nhãn hiệu được đăng ký), CD-ROM, DVD, và ROM, bộ nhớ bán dẫn, và những vật ghi tương tự.

Đặc biệt, chương trình mã hóa dự đoán video được cung cấp cùng với môđun nhập để thực hiện việc nhập vào các ảnh tạo thành chuỗi video, môđun mã hóa để mã hóa ảnh bởi hoặc mã hóa trong khung hoặc mã hóa liên khung để tạo ra dữ liệu ảnh nén bao gồm thông tin về kích thước khung của ảnh đích, số lượng tối đa ảnh được khôi phục có thể lưu trữ, và kích thước khung tối đa như là thiết lập mã hóa, và để mã hóa dữ liệu ảnh được nén cùng với mã hóa dữ liệu mô tả về

thông tin kích thước khung ảnh đích, số lượng tối đa ảnh được khôi phục có thể lưu trữ, và kích thước khung tối đa, môđun giải mã để giải mã dữ liệu ảnh được nén để tái tạo ảnh, môđun lưu trữ ảnh để lưu trữ một hoặc nhiều ảnh được khôi phục để sử dụng cho mã hóa ảnh tiếp theo, và môđun điều khiển bộ nhớ để điều khiển bộ nhớ đệm ảnh trong môđun lưu trữ ảnh, tại đó khi mà kích thước ảnh của ảnh đầu vào không lớn hơn $1/(2^L)$ (L là số nguyên lớn hơn hoặc bằng một) của kích thước khung tối đa của ảnh, môđun điều khiển bộ nhớ xác định rằng (2^L) lần số lượng tối đa của ảnh được khôi phục là có thể lưu trữ trong bộ nhớ khung.

Chương trình giải mã dự đoán video được bố trí cùng với môđun nhập để thực hiện việc đưa vào dữ liệu đã được mã hóa chứa dữ liệu ảnh được nén mà bao gồm thông tin về kích thước khung của ảnh đích, số lượng tối đa ảnh được khôi phục có thể lưu trữ, và kích thước khung tối đa như là thiết lập mã hóa, tại đó các ảnh tạo thành chuỗi video được mã hóa với mã hóa trong khung hoặc mã hóa liên khung, môđun giải mã để giải mã dữ liệu ảnh được nén và dữ liệu ảnh nén đó tạo thành thông tin về kích thước khung của ảnh đích, số lượng tối đa của ảnh được khôi phục có thể lưu trữ, và kích thước khung tối đa để tái tạo ảnh được khôi phục và thông tin về kích thước khung của ảnh đích, số lượng tối đa của ảnh được khôi phục có thể lưu trữ, và kích thước khung tối đa, môđun lưu trữ ảnh để lưu trữ một hoặc nhiều hơn các ảnh được khôi phục làm ảnh tham chiếu để giải mã ảnh tiếp theo, và môđun điều khiển bộ nhớ để điều khiển môđun lưu trữ ảnh, tại đó khi mà kích thước ảnh của ảnh đầu vào không lớn hơn $1/(2^L)$ (L là số nguyên lớn hơn hoặc bằng một) của kích thước khung tối đa của ảnh, môđun điều khiển bộ nhớ xác định rằng (2^L) lần số lượng tối đa của ảnh được khôi phục là có thể lưu trữ trong bộ nhớ khung.

Fig.7 là hình vẽ thể hiện cấu trúc phần cứng của máy tính thực hiện chương trình được ghi trong vật ghi và Fig.8 là hình phối cảnh của máy tính thực hiện chương trình lưu trữ trong vật ghi. Máy tính đó có ổ đọc DVD, bộ giải mã tín hiệu truyền hình, máy điện thoại cầm tay, v.v. được cung cấp cùng với CPU và được cấu hình để thực hiện các quy trình và việc điều khiển bởi phần mềm.

Như được thể hiện trên Fig.7, máy tính 30 được cung cấp cùng với thiết bị đọc 12 ví dụ như ổ đĩa mềm (nhãn hiệu được đăng ký), hoặc ổ đĩa DVD, bộ nhớ làm việc (RAM) 14 mà trên đó có hệ thống hoạt động, bộ nhớ 16 để lưu trữ chương trình được lưu trên vật ghi 10, bộ theo dõi 18 như màn hình, chuột 20 và bàn phím 22 là các thiết bị nhập, thiết bị truyền thông 24 để truyền và nhận dữ liệu hoặc tương tự, và CPU 26 để điều khiển việc thực hiện các chương trình. Khi vật ghi 10 được đặt vào thiết bị đọc 12, máy tính 30 có thể truy cập vào chương trình mã hóa và giải mã dự đoán video được lưu trữ trong vật ghi 10 thông qua thiết bị đọc 12 và có thể hoạt động như thiết bị mã hóa và giải mã dự đoán video theo như sáng chế, dựa trên chương trình mã hóa và giải mã dự đoán video.

Như được thể hiện trong Fig.8, chương trình mã hóa dự đoán video hoặc chương trình giải mã dự đoán video có thể được cung cấp dưới dạng tín hiệu dữ liệu máy tính 40 xếp chồng trên sóng mang, thông qua mạng. Trong trường hợp này, máy tính 30 có thể thực hiện chương trình mã hóa dự đoán video hoặc chương trình giải mã dự đoán video sau khi chương trình mã hóa dự đoán video hoặc chương trình giải mã dự đoán video nhận được từ thiết bị truyền thông 24 được lưu trữ trong bộ nhớ 16.

Danh sách số chỉ dẫn

301 thiết bị đầu cuối nhập; 302 bộ chia khối; 303 bộ tạo tín hiệu dự đoán; 304 bộ nhớ khung; 305 bộ trừ; 306 bộ biến đổi; 307 bộ lượng tử hóa; 308 bộ giải lượng tử hóa; 309 bộ biến đổi ngược; 310 bộ cộng; 311 bộ mã hóa entropi; 312 thiết bị đầu cuối đầu ra; 313 thiết bị đầu cuối nhập; 314 bộ quản lý bộ nhớ khung; 401 thiết bị đầu cuối nhập; 402 bộ phân tích dữ liệu; 403 bộ giải lượng tử hóa; 404 bộ biến đổi ngược; 405 bộ cộng; 406 thiết bị đầu cuối đầu ra; 407 bộ nhớ khung; 408 bộ tạo tín hiệu dự đoán; 409 bộ quản lý bộ nhớ khung.

YÊU CẦU BẢO HỘ

1. Phương pháp giải mã dự đoán video được thực hiện bởi thiết bị giải mã dự đoán video, phương pháp này bao gồm các bước:

nhập để nhập dữ liệu được mã hóa bao gồm dữ liệu ảnh nén mà chứa kích thước khung của ảnh đích, số lượng tối đa của ảnh được khôi phục có thể lưu trữ, và kích thước khung tối đa được định rõ là ràng buộc mã hóa, trong đó các ảnh tạo thành chuỗi video được mã hóa bởi dự đoán trong khung hoặc dự đoán liên khung;

giải mã để giải mã dữ liệu ảnh nén và dữ liệu ảnh nén tạo thành thông tin về kích thước khung của ảnh đích, số lượng tối đa của ảnh được khôi phục có thể lưu trữ, và kích thước khung tối đa để tái tạo các ảnh và thông tin về kích thước khung của ảnh đích, số lượng tối đa của ảnh được khôi phục có thể lưu trữ, và kích thước khung tối đa; và

lưu trữ ảnh để lưu trữ trong các phương tiện lưu trữ ảnh một hoặc nhiều hơn ảnh được khôi phục làm ảnh tham chiếu để sử dụng cho việc giải mã ảnh tiếp theo,

trong đó số lượng tối đa của ảnh được khôi phục có thể lưu trữ được xác định là bằng, bằng 2 lần hoặc 4 lần số định trước dựa vào mối quan hệ giữa kích thước khung tối đa của các ảnh và kích thước khung của các ảnh đầu vào.

2. Phương pháp giải mã dự đoán video theo điểm 1, trong đó khi kích thước khung của ảnh đầu vào không lớn hơn $1/2$ kích thước khung tối đa của các ảnh, số lượng tối đa của các ảnh được khôi phục có thể lưu trữ được xác định là bằng 2 lần số định trước.

3. Phương pháp giải mã dự đoán video theo điểm 1, trong đó khi kích thước khung của các ảnh đầu vào không lớn hơn $1/4$ kích thước khung tối đa của các ảnh, số lượng tối đa của các ảnh được khôi phục có thể lưu trữ được xác định là bằng 4 lần số định trước.

4. Thiết bị giải mã dự đoán video bao gồm:

phương tiện nhập để nhập dữ liệu được mã hóa bao gồm dữ liệu ảnh nén mà có thông tin về kích thước khung của ảnh đích, số lượng tối đa của các ảnh được khôi phục có thể lưu trữ, và kích thước khung tối đa được định rõ là ràng buộc mã hóa, trong đó các ảnh tạo thành chuỗi video được mã hóa bởi dự đoán trong khung hoặc dự đoán liên khung;

phương tiện giải mã để giải mã dữ liệu ảnh nén và dữ liệu ảnh nén tạo thành thông tin về kích thước khung của ảnh đích, số lượng tối đa của các ảnh được khôi phục có thể lưu trữ, và kích thước khung tối đa để khôi phục các ảnh và thông tin về kích thước khung của ảnh đích, số lượng tối đa của ảnh được khôi phục có thể lưu trữ, và kích thước khung tối đa; và

phương tiện lưu trữ ảnh để lưu trữ một hoặc nhiều hơn ảnh được khôi phục làm các ảnh tham chiếu được sử dụng cho việc giải mã ảnh tiếp theo,

trong đó số lượng tối đa của ảnh được khôi phục có thể lưu trữ được xác định là bằng, bằng 2 lần hoặc 4 lần một số lần định trước dựa vào mối quan hệ giữa kích thước khung tối đa của các ảnh và kích thước khung của các ảnh đầu vào.

Fig.1

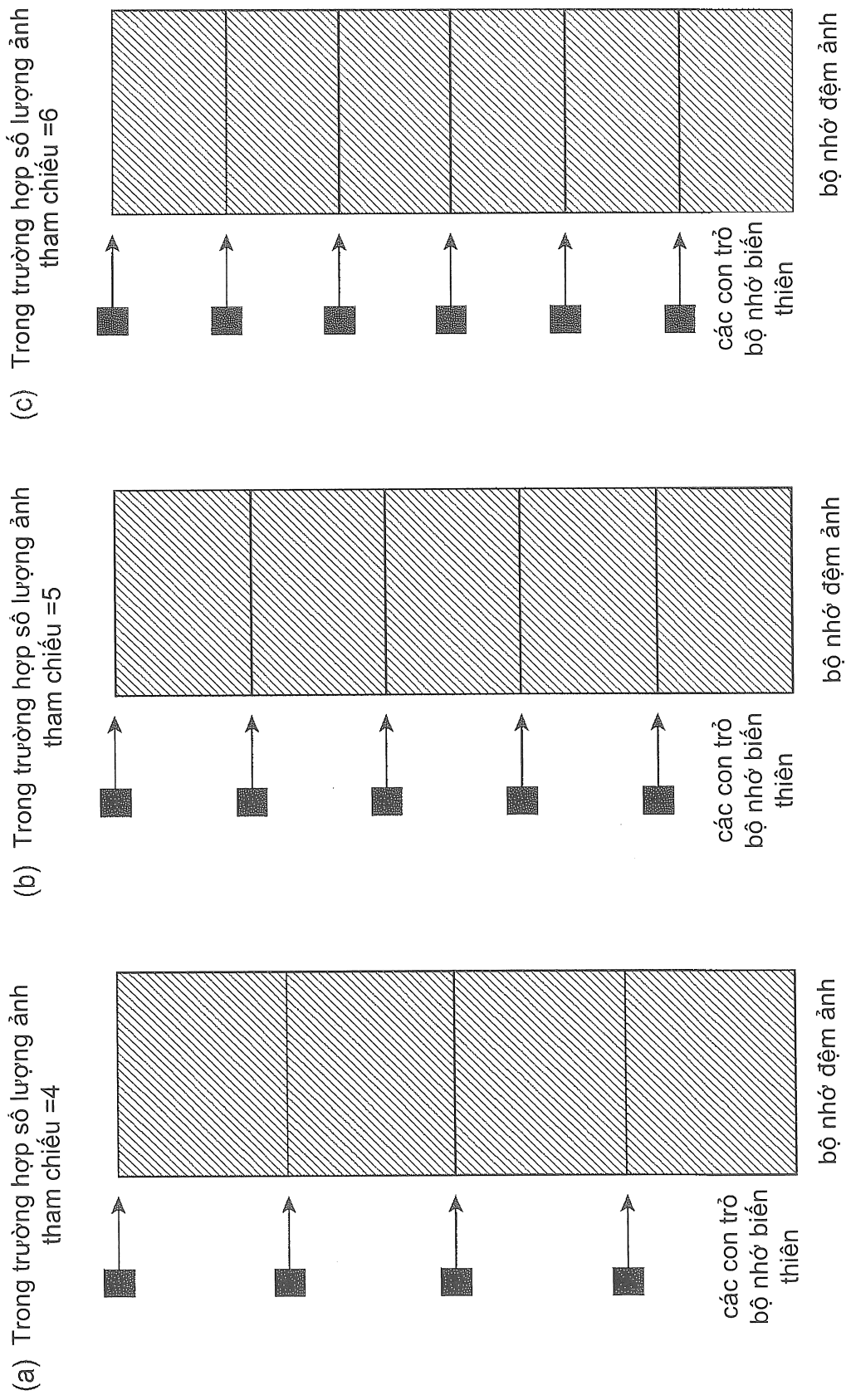
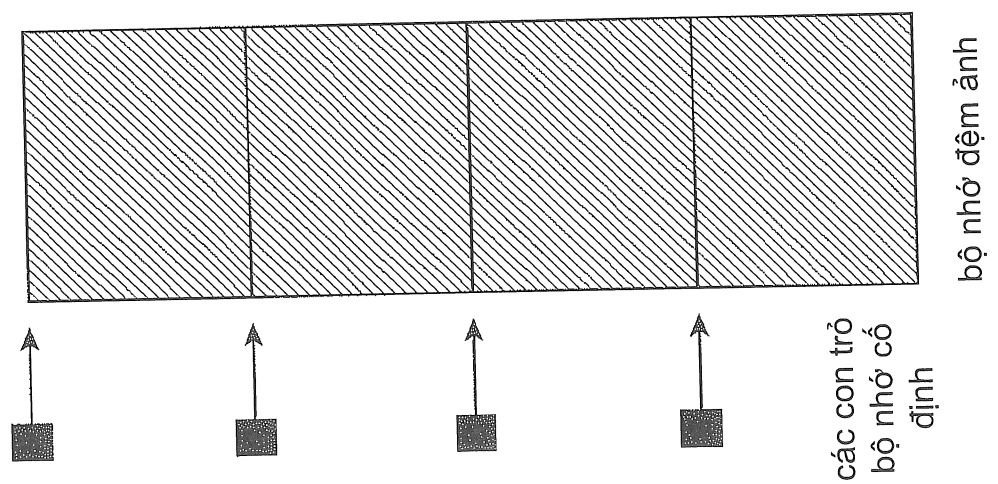


Fig.2

(a) Trong trường hợp kích thước khung =
kích thước khung tối đa



(b) Trong trường hợp kích thước khung <
kích thước khung tối đa

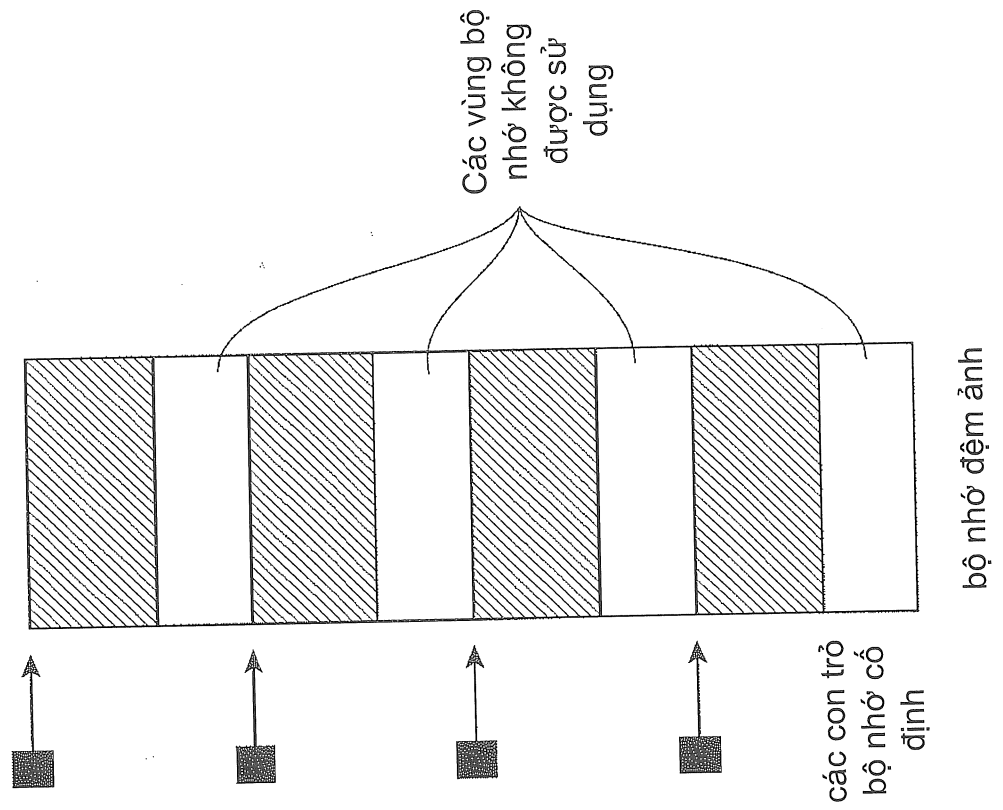


Fig.3

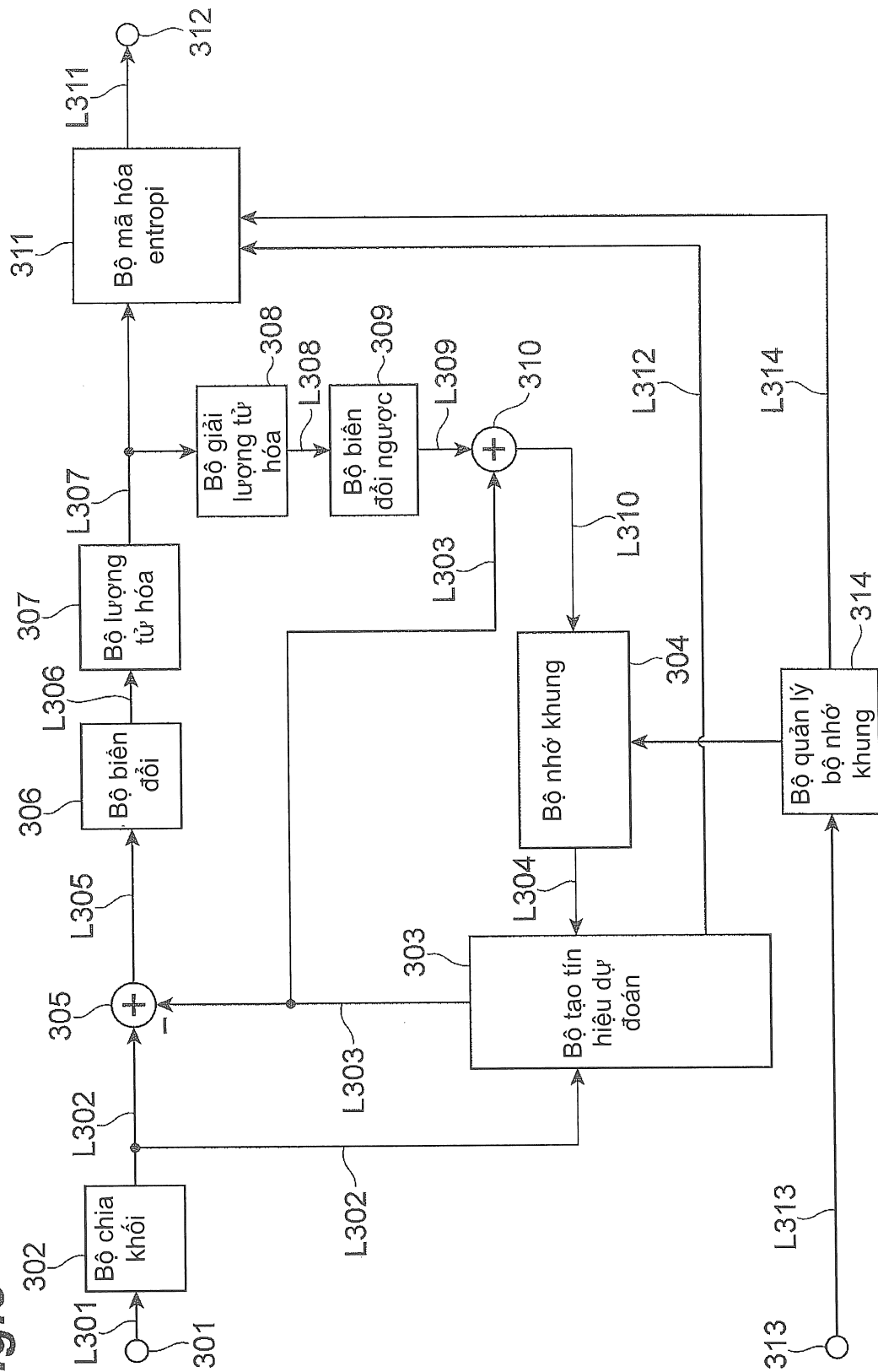


Fig.4

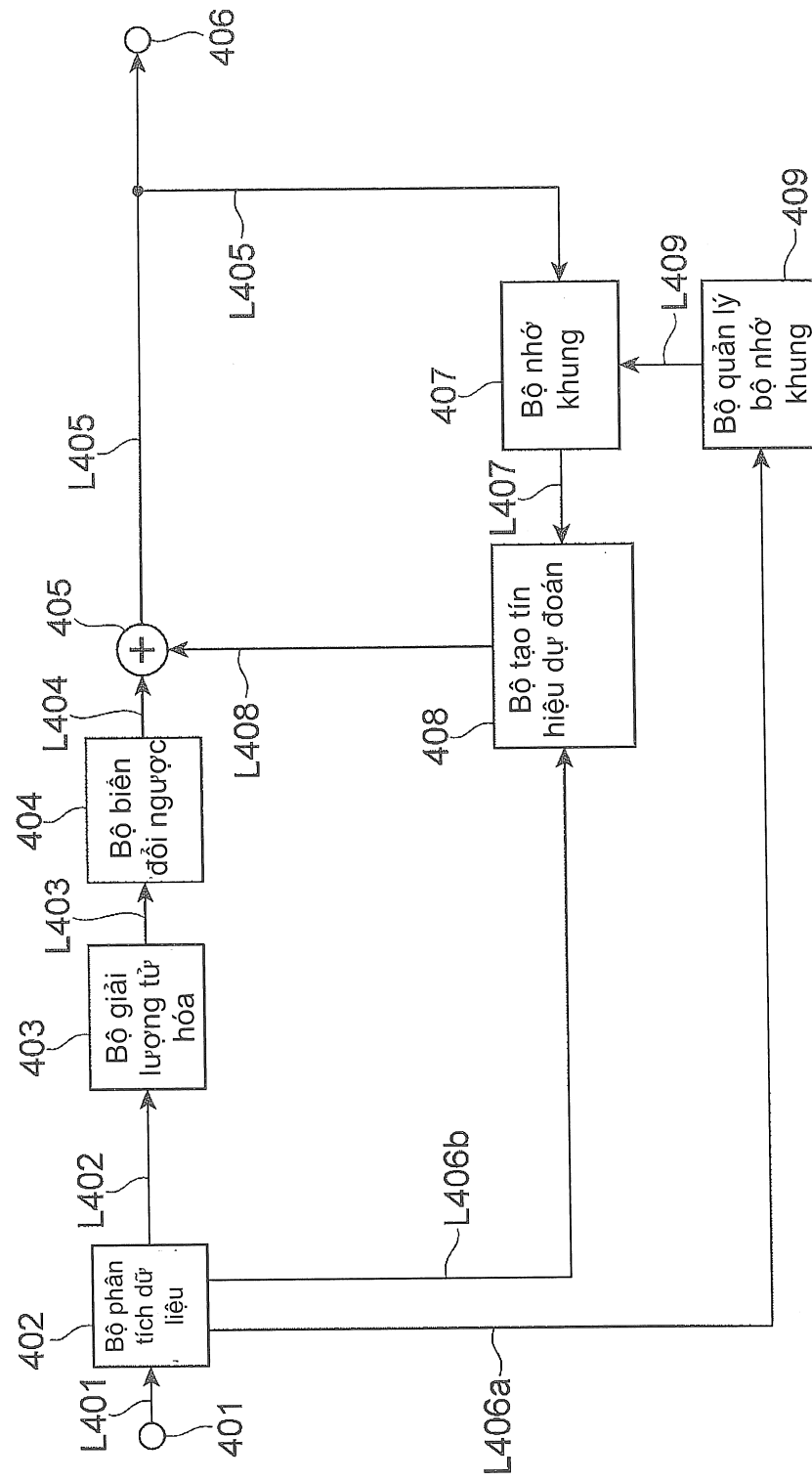


Fig. 5

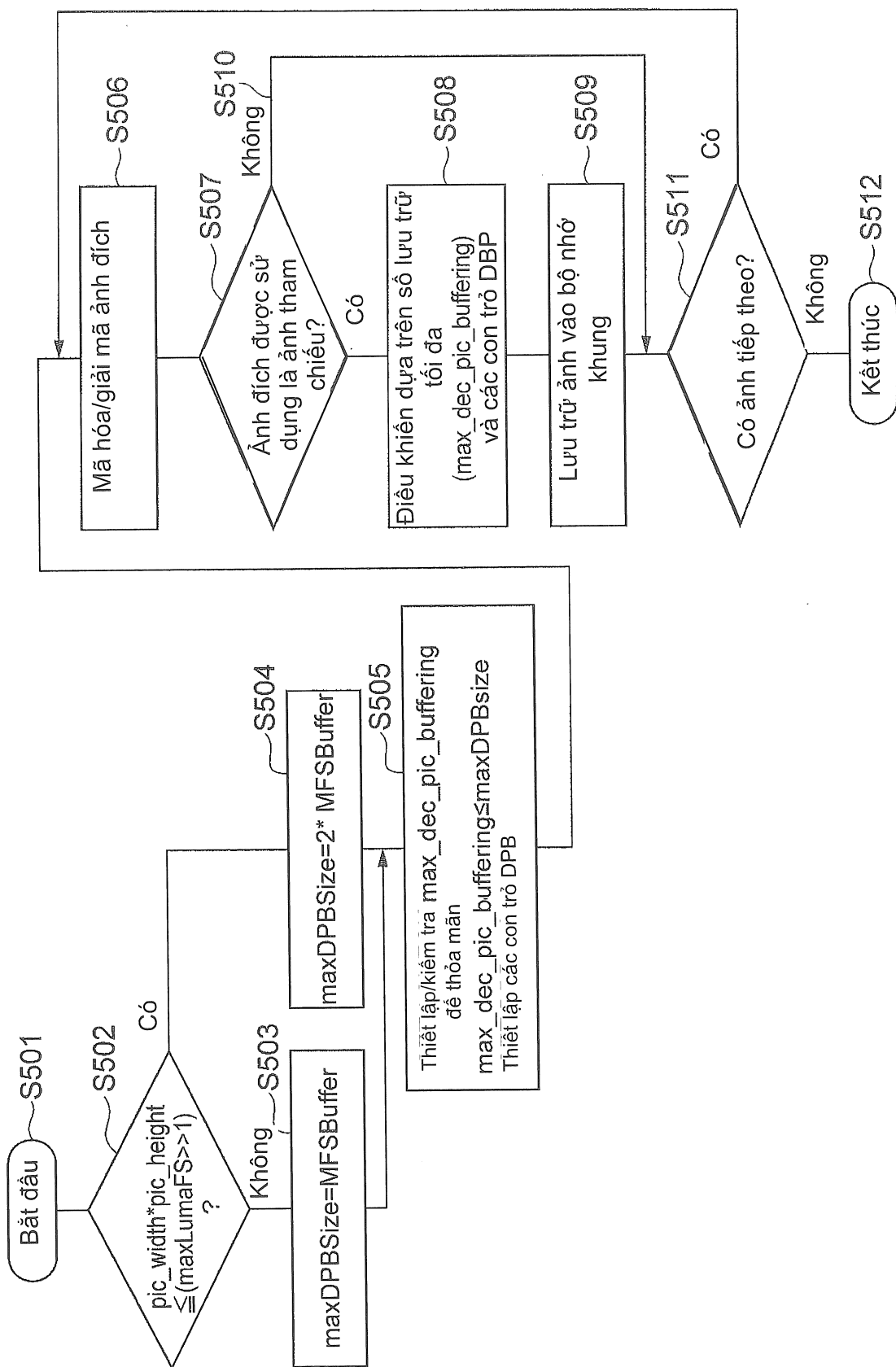


Fig.6

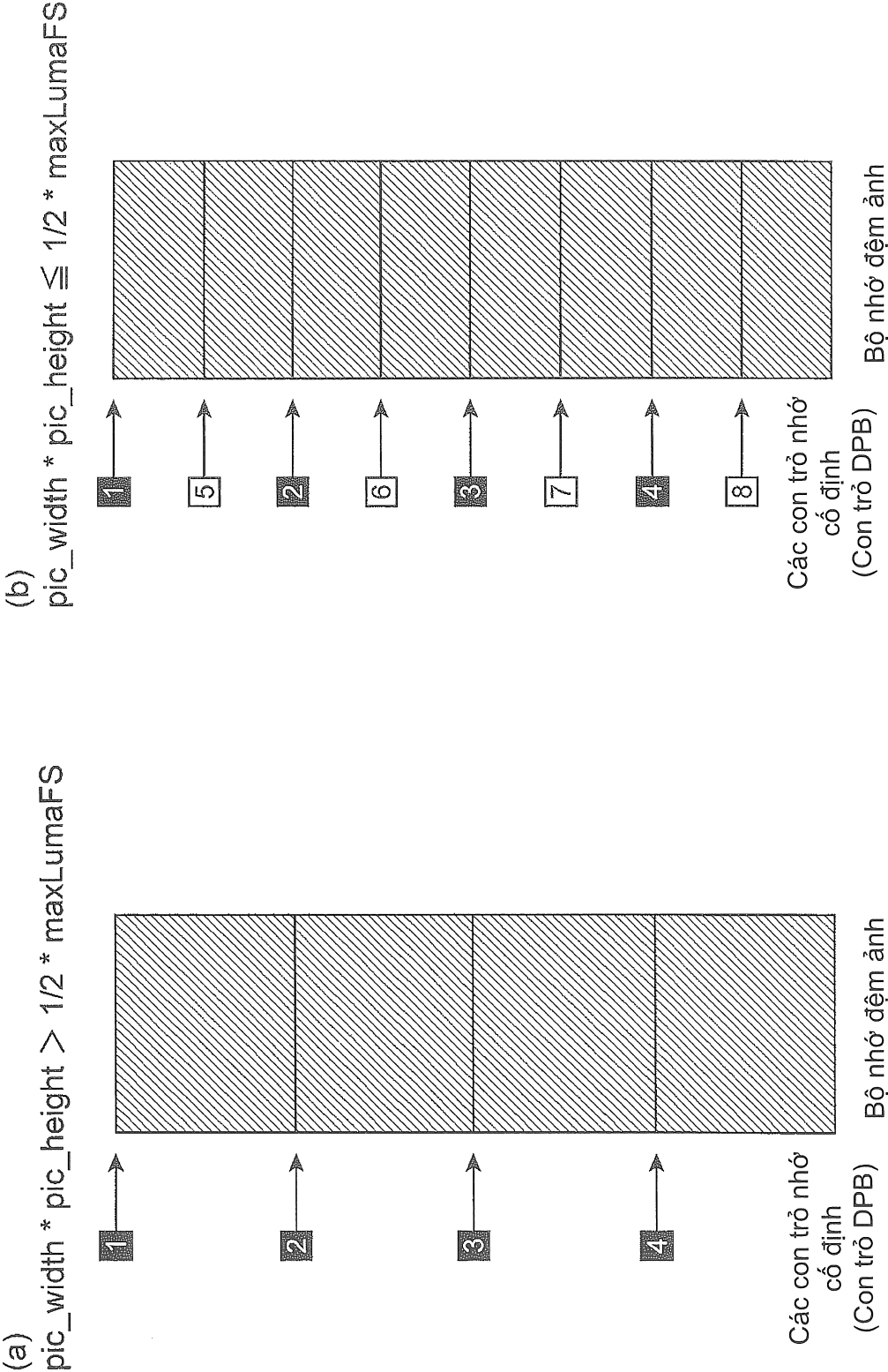


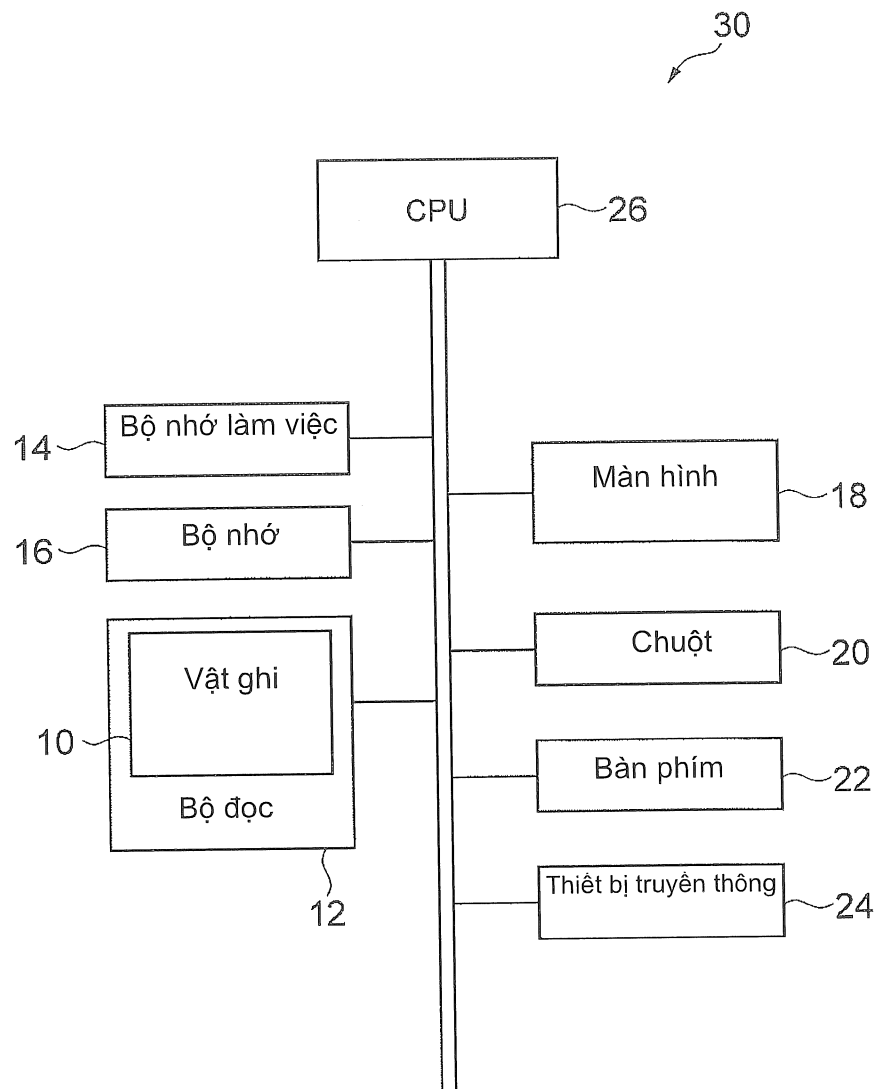
Fig.7

Fig.8