



(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ
(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN) (11) 
CỤC SỞ HỮU TRÍ TUỆ
1-0052152
(51)^{2021.01} H04N 19/46; H04N 19/169; H04N 19/174 (13) B

(21) 1-2022-06111 (22) 05/03/2021
(86) PCT/KR2021/002772 05/03/2021 (87) WO2021/177791 10/09/2021
(30) 62/985,333 05/03/2020 US
(45) 27/10/2025 451 (43) 25/11/2022 416A
(73) LG ELECTRONICS INC. (KR)
128, Yeoui-daero, Yeongdeungpo-gu Seoul 07336, Republic of Korea
(72) HENDRY, Hendry (ID).
(74) Công ty Luật TNHH T&G (TGVN)

(54) PHƯƠNG PHÁP MÃ HÓA VÀ GIẢI MÃ ẢNH, VÀ PHƯƠNG PHÁP TRUYỀN
LUỒNG BIT

(21) 1-2022-06111

(57) Sáng chế đề cập đến phương pháp và thiết bị mã hóa/giải mã ảnh. Phương pháp giải mã ảnh này bao gồm các bước thu, từ luồng bit, thông tin loại đơn vị lớp trừu tượng mạng (network abstraction layer, NAL) lớp lập mã video (video coding layer, VCL) của hình ảnh hiện tại và thông tin cờ thứ nhất chỉ rõ liệu hình ảnh con được chứa trong hình ảnh hiện tại có được xử lý như một hình ảnh hay không, xác định loại đơn vị NAL của mỗi lát trong số nhiều lát được chứa trong hình ảnh hiện tại, dựa trên thông tin loại đơn vị NAL VCL thu được, và giải mã nhiều lát này dựa trên loại đơn vị NAL được xác định và thông tin cờ thứ nhất. Hình ảnh hiện tại có thể bao gồm hai hay hơn hai hình ảnh con dựa trên việc ít nhất một số lát trong số nhiều lát đã nêu đang có các loại đơn vị NAL khác nhau, và thông tin cờ thứ nhất có thể có giá trị được xác định trước chỉ rõ rằng mỗi hình ảnh con trong số hai hay hơn hai hình ảnh con này được xử lý như một hình ảnh dựa trên việc ít nhất một số lát trong số nhiều lát đã nêu đang có các loại đơn vị NAL khác nhau.

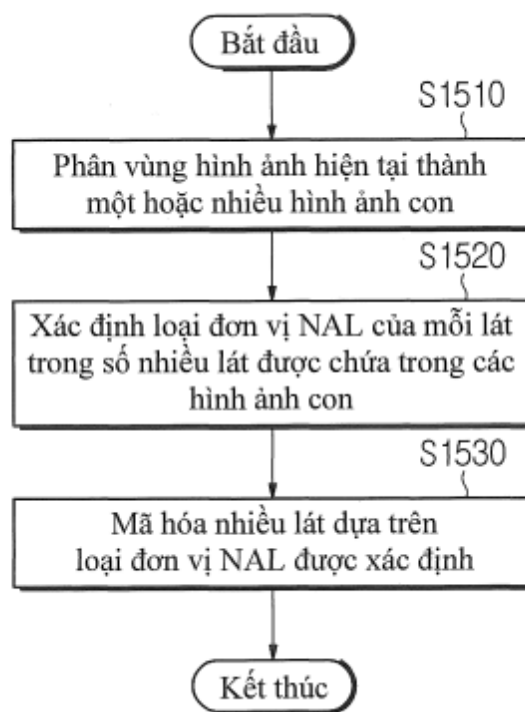


FIG.15

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến phương pháp và thiết bị mã hóa/giải mã ảnh, và, cụ thể hơn là, đề cập đến phương pháp và thiết bị mã hóa và giải mã ảnh dựa trên loại đơn vị NAL được trộn, và phương tiện ghi để lưu trữ luồng bit được tạo ra bởi phương pháp/thiết bị mã hóa ảnh của sáng chế.

Tình trạng kỹ thuật của sáng chế

Gần đây, nhu cầu đối với các ảnh độ phân giải cao và chất lượng cao, chẳng hạn như các ảnh độ phân giải cao (High Definition, HD) và các ảnh độ phân giải siêu cao (Ultra High Definition, UHD), đang ngày càng gia tăng trong các lĩnh vực khác nhau. Khi độ phân giải và chất lượng của dữ liệu ảnh được cải thiện, lượng thông tin hoặc bit được truyền tăng tương đối so với dữ liệu ảnh hiện có. Sự gia tăng về lượng thông tin hoặc bit được truyền làm gia tăng về chi phí truyền và chi phí lưu trữ.

Theo đó, cần có công nghệ nén ảnh hiệu quả cao để truyền, lưu trữ, và tái tạo một cách hiệu quả thông tin về các ảnh độ phân giải cao và chất lượng cao.

Bản chất kỹ thuật của sáng chế

Vấn đề kỹ thuật

Một mục đích của sáng chế là để đề xuất phương pháp và thiết bị mã hóa/giải mã ảnh có hiệu suất mã hóa/giải mã được cải thiện.

Một mục đích khác của sáng chế là để đề xuất phương pháp và thiết bị mã hóa/giải mã ảnh dựa trên loại đơn vị NAL được trộn.

Một mục đích khác của sáng chế là để đề xuất phương pháp và thiết bị mã hóa/giải mã ảnh dựa trên hai hay hơn hai hình ảnh con có các loại đơn vị NAL khác nhau.

Một mục đích khác của sáng chế là để đề xuất phương tiện lưu trữ phi chuyển tiếp đọc được bằng bộ giải mã để lưu trữ luồng bit được tạo ra bởi phương pháp hoặc thiết bị mã hóa ảnh theo sáng chế.

Một mục đích khác của sáng chế là để đề xuất phương tiện ghi lưu trữ luồng bit được nhận, được giải mã và được sử dụng để xây dựng lại ảnh bằng thiết bị giải mã ảnh

theo sáng chế.

Một mục đích khác của sáng chế là để đề xuất phương pháp truyền luồng bit được tạo ra bởi phương pháp hoặc thiết bị mã hóa ảnh theo sáng chế.

Các vấn đề kỹ thuật được giải quyết bởi sáng chế này không bị giới hạn vào các vấn đề kỹ thuật trên đây và các vấn đề kỹ thuật khác không được mô tả ở đây sẽ trở nên rõ ràng với những người có hiểu biết trong lĩnh vực tương ứng nhờ phần mô tả sau đây.

Giải pháp kỹ thuật của sáng chế

Phương pháp giải mã ảnh theo một khía cạnh của sáng chế bao gồm các bước thu, từ luồng bit, thông tin loại đơn vị lớp trừu tượng mạng (network abstraction layer, NAL) lớp lập mã video (video coding layer, VCL) của hình ảnh hiện tại và thông tin cờ thứ nhất chỉ rõ liệu hình ảnh con được chứa trong hình ảnh hiện tại có được xử lý như một hình ảnh hay không, xác định loại đơn vị NAL của mỗi lát trong số nhiều lát được chứa trong hình ảnh hiện tại, dựa trên thông tin loại đơn vị NAL VCL thu được, và giải mã nhiều lát này dựa trên loại đơn vị NAL được xác định và thông tin cờ thứ nhất. Hình ảnh hiện tại có thể bao gồm hai hay hơn hai hình ảnh con dựa trên việc ít nhất một số lát trong số nhiều lát đã nêu đang có các loại đơn vị NAL khác nhau, và thông tin cờ thứ nhất có thể có giá trị được xác định trước chỉ rõ rằng mỗi hình ảnh con trong số hai hay hơn hai hình ảnh con này được xử lý như một hình ảnh dựa trên việc ít nhất một số lát trong số nhiều lát đã nêu đang có các loại đơn vị NAL khác nhau.

Thiết bị giải mã ảnh theo một khía cạnh khác của sáng chế bao gồm bộ nhớ và ít nhất một bộ xử lý. Ít nhất một bộ xử lý có thể thu, từ luồng bit, thông tin loại đơn vị lớp trừu tượng mạng (NAL) lớp lập mã video (VCL) của hình ảnh hiện tại và thông tin cờ thứ nhất chỉ rõ liệu hình ảnh con được chứa trong hình ảnh hiện tại có được xử lý như một hình ảnh hay không, xác định loại đơn vị NAL của mỗi lát trong số nhiều lát được chứa trong hình ảnh hiện tại, dựa trên thông tin loại đơn vị NAL VCL thu được, và giải mã nhiều lát này dựa trên loại đơn vị NAL được xác định và thông tin cờ thứ nhất. Hình ảnh hiện tại có thể bao gồm hai hay hơn hai hình ảnh con dựa trên việc ít nhất một số lát trong số nhiều lát đã nêu đang có các loại đơn vị NAL khác nhau, và thông tin cờ thứ nhất có thể có giá trị được xác định trước chỉ rõ rằng mỗi hình ảnh con trong số hai hay hơn hai hình ảnh con này được xử lý như một hình ảnh dựa trên việc ít nhất một số lát trong số nhiều lát đã nêu đang có các loại đơn vị NAL khác nhau.

Phương pháp mã hóa ảnh theo một khía cạnh khác của sáng chế bao gồm các bước phân vùng hình ảnh hiện tại thành một hoặc nhiều hình ảnh con, xác định loại đơn vị lớp trừu tượng mạng (NAL) của mỗi lát trong số nhiều lát được chứa trong một hoặc nhiều hình ảnh con này, và mã hóa nhiều lát này dựa trên loại đơn vị NAL được xác định. Hình ảnh hiện tại có thể được phân vùng thành hai hay hơn hai hình ảnh con dựa trên việc ít nhất một số lát trong số nhiều lát đã nêu đang có các loại đơn vị NAL khác nhau, và mỗi hình ảnh con trong số hai hay hơn hai hình ảnh con này có thể được xử lý như một hình ảnh dựa trên việc ít nhất một số lát trong số nhiều lát đã nêu đang có các loại đơn vị NAL khác nhau.

Thêm vào đó, phương tiện ghi có thể đọc được bằng máy tính theo một khía cạnh khác của sáng chế có thể lưu trữ luồng bit được tạo ra bởi phương pháp mã hóa ảnh hoặc thiết bị mã hóa ảnh của sáng chế này.

Thêm vào đó, phương pháp truyền theo một khía cạnh khác của sáng chế có thể truyền luồng bit được tạo ra bởi thiết bị mã hóa ảnh hoặc phương pháp mã hóa ảnh của sáng chế này.

Các dấu hiệu được tóm tắt ngắn gọn trên đây đối với sáng chế này chỉ đơn thuần là các khía cạnh làm ví dụ của phần mô tả chi tiết dưới đây của sáng chế, và không làm giới hạn phạm vi bảo hộ của sáng chế.

Hiệu quả của sáng chế

Theo sáng chế, có thể đề xuất phương pháp và thiết bị mã hóa/giải mã ảnh với hiệu suất mã hóa/giải mã được cải thiện.

Ngoài ra, theo sáng chế, có thể đề xuất phương pháp và thiết bị mã hóa/giải mã ảnh dựa trên loại đơn vị NAL được trộn.

Ngoài ra, theo sáng chế, có thể đề xuất phương pháp và thiết bị mã hóa/giải mã ảnh dựa trên hai hay hơn hai hình ảnh con có các loại đơn vị NAL được trộn khác nhau.

Ngoài ra, theo sáng chế, có thể đề xuất phương pháp truyền luồng bit được tạo ra bởi phương pháp hoặc thiết bị mã hóa ảnh theo sáng chế.

Thêm vào đó, theo sáng chế, có thể đề xuất phương tiện ghi lưu trữ luồng bit được tạo ra bởi phương pháp hoặc thiết bị mã hóa ảnh theo sáng chế.

Thêm vào đó, theo sáng chế, có thể đề xuất phương tiện ghi lưu trữ luồng bit được nhận, được giải mã và được sử dụng để xây dựng lại ảnh bởi thiết bị giải mã ảnh theo

sáng chế.

Những người có hiểu biết trong lĩnh vực kỹ thuật tương ứng sẽ hiểu rằng các hiệu quả có thể đạt được thông qua sáng chế này không bị giới hạn vào những gì đã được mô tả một cách cụ thể trên đây và các ưu điểm khác của sáng chế sẽ được hiểu rõ ràng hơn từ phần mô tả chi tiết.

Mô tả vắn tắt các hình vẽ

Fig.1 là hình minh họa sơ lược hệ thống lập mã video, mà phương án của sáng chế có thể áp dụng được với nó.

Fig.2 là hình minh họa sơ lược thiết bị mã hóa ảnh, mà phương án của sáng chế có thể áp dụng được với nó.

Fig.3 là hình minh họa sơ lược thiết bị giải mã ảnh, mà phương án của sáng chế có thể áp dụng được với nó.

Fig.4 là lưu đồ minh họa thủ tục giải mã ảnh, mà phương án của sáng chế có thể áp dụng được với nó.

Fig.5 là lưu đồ minh họa thủ tục mã hóa ảnh, mà phương án của sáng chế có thể áp dụng được với nó.

Fig.6 là hình minh họa ví dụ về cấu trúc lớp cho ảnh/video được lập mã.

Fig.7 là hình minh họa tập tham số hình ảnh (picture parameter set, PPS) theo một phương án của sáng chế.

Fig.8 là hình minh họa tiêu đề lát theo một phương án của sáng chế.

Fig.9 là hình minh họa ví dụ về hình ảnh con.

Fig.10 là hình minh họa ví dụ về hình ảnh có loại đơn vị NAL được trộn.

Fig.11 là hình minh họa ví dụ về tập tham số hình ảnh (PPS) theo một phương án của sáng chế.

Fig.12 là hình minh họa ví dụ về tập tham số trình tự (sequence parameter set, SPS) theo một phương án của sáng chế.

Fig.13 là hình minh họa thứ tự giải mã và thứ tự xuất ra đối với mỗi loại hình ảnh.

Fig.14 là lưu đồ minh họa phương pháp xác định loại đơn vị NAL của hình ảnh hiện tại theo một phương án của sáng chế.

Fig.15 là lưu đồ minh họa phương pháp mã hóa ảnh theo một phương án của sáng chế.

Fig.16 là lưu đồ minh họa phương pháp giải mã ảnh theo một phương án của sáng chế.

Fig.17 là hình minh họa hệ thống tạo luồng nội dung, mà phương án của sáng chế có thể áp dụng được với nó.

Mô tả chi tiết sáng chế

Ở dưới đây, các phương án của sáng chế sẽ được mô tả chi tiết với tham chiếu đến các hình vẽ kèm theo để dễ dàng được thi hành bởi những người có hiểu biết trong lĩnh vực tương ứng. Tuy nhiên, sáng chế có thể được thi hành dưới nhiều dạng khác nhau, và không bị giới hạn ở các phương án được mô tả ở đây.

Khi mô tả sáng chế này, nếu xác định được rằng phần mô tả chi tiết của chức năng hoặc cấu trúc đã biết có liên quan khiến cho phạm vi của sáng chế này trở nên mơ hồ một cách không cần thiết thì phần mô tả chi tiết của chúng sẽ được lược bỏ. Trên các hình vẽ, các phần không liên quan đến phần mô tả của sáng chế được lược bỏ, và các số tham chiếu tương tự nhau được gắn với các phần tương tự nhau.

Trong sáng chế, khi một thành phần được “kết nối”, “ghép nối” hoặc “liên kết” với thành phần khác, điều này có thể bao gồm không chỉ mối quan hệ kết nối trực tiếp mà cả mối quan hệ kết nối gián tiếp trong đó có mặt thành phần xen giữa. Thêm vào đó, khi một thành phần “bao gồm” hoặc “có” các thành phần khác, điều này có nghĩa là các thành phần khác có thể được đưa vào thêm, chứ không phải loại trừ các thành phần khác, trừ khi được nêu rõ khác đi.

Trong sáng chế, các thuật ngữ thứ nhất, thứ hai, v.v. có thể được sử dụng chỉ nhằm mục đích phân biệt một thành phần với các thành phần khác, và không làm giới hạn thứ tự hoặc tầm quan trọng của các thành phần trừ khi được nêu rõ khác đi. Theo đó, trong phạm vi của sáng chế, thành phần thứ nhất trong một phương án có thể được gọi là thành phần thứ hai trong phương án khác, và tương tự, thành phần thứ hai trong một phương án có thể được gọi là thành phần thứ nhất trong phương án khác.

Trong sáng chế, các thành phần mà chúng được phân biệt với nhau được dự tính để mô tả rõ ràng mỗi đặc điểm, và không có nghĩa là các thành phần này nhất thiết phải tách rời nhau. Tức là, nhiều thành phần có thể được hợp nhất và được thi hành trong một đơn vị phần cứng hoặc phần mềm, hoặc một thành phần có thể được phân tán và được thi hành trong nhiều đơn vị phần cứng hoặc phần mềm. Do đó, kể cả nếu không được

nêu rõ khác đi thì các phương án như vậy, mà trong đó các thành phần được hợp nhất hoặc thành phần được phân tán, cũng được bao gồm trong phạm vi của sáng chế.

Trong sáng chế, các thành phần được mô tả trong các phương án khác nhau không nhất thiết là để chỉ các thành phần thiết yếu, và một số thành phần có thể là các thành phần tùy chọn. Theo đó, phương án bao gồm tập con của các thành phần được mô tả trong một phương án cũng được bao gồm trong phạm vi của sáng chế. Thêm vào đó, các phương án bao gồm các thành phần khác, ngoài các thành phần được mô tả trong các phương án khác nhau, được bao gồm trong phạm vi của sáng chế.

Sáng chế đề cập đến việc mã hóa và giải mã ảnh, và các thuật ngữ được sử dụng trong sáng chế này có thể có nghĩa chung được sử dụng phổ biến trong lĩnh vực kỹ thuật, mà sáng chế thuộc về lĩnh vực kỹ thuật này, trừ khi được định nghĩa mới trong sáng chế.

Trong sáng chế, “hình ảnh” nhìn chung đề cập đến đơn vị biểu diễn một ảnh trong khoảng thời gian cụ thể, và lát/ô xếp là đơn vị lập mã cấu thành nên một phần của hình ảnh, và một hình ảnh có thể gồm một hoặc nhiều lát/ô xếp. Thêm vào đó, lát/ô xếp có thể bao gồm một hoặc nhiều đơn vị cây lập mã (coding tree unit, CTU).

Trong sáng chế, “điểm ảnh” (“pixel”) hoặc “pel” có thể nghĩa là đơn vị nhỏ nhất cấu thành nên một hình ảnh (hoặc ảnh). Thêm vào đó, “mẫu” có thể được sử dụng như thuật ngữ tương ứng với điểm ảnh. Mẫu nhìn chung có thể biểu diễn điểm ảnh hoặc giá trị của điểm ảnh, và có thể biểu diễn chỉ điểm ảnh/giá trị điểm ảnh của thành phần độ chói hoặc chỉ điểm ảnh/giá trị điểm ảnh của thành phần sắc độ.

Trong sáng chế, “đơn vị” có thể biểu diễn đơn vị cơ sở của việc xử lý ảnh. Đơn vị có thể bao gồm ít nhất một thành phần trong số vùng cụ thể của hình ảnh và thông tin liên quan đến vùng này. Đơn vị có thể được sử dụng thay thế được cho nhau với các thuật ngữ chẳng hạn như “mảng mẫu”, “khối” hoặc “khu vực” trong một số trường hợp. Trong trường hợp chung, khối $M \times N$ có thể bao gồm các mẫu (hoặc các mảng mẫu) hoặc tập hợp (hoặc mảng) các hệ số biến đổi gồm M cột và N hàng.

Trong sáng chế, “khối hiện tại” có thể nghĩa là một thành phần trong số “khối lập mã hiện tại”, “đơn vị lập mã hiện tại”, “khối mục tiêu lập mã”, “khối mục tiêu giải mã” hoặc “khối mục tiêu xử lý”. Khi việc dự đoán được thực hiện, “khối hiện tại” có thể nghĩa là “khối dự đoán hiện tại” hoặc “khối mục tiêu dự đoán”. Khi việc biến đổi (biến đổi ngược)/lượng tử hóa (giải lượng tử hóa) được thực hiện, “khối hiện tại” có thể nghĩa

là “khối biến đổi hiện tại” hoặc “khối mục tiêu biến đổi”. Khi việc lọc được thực hiện, “khối hiện tại” có thể nghĩa là “khối mục tiêu lọc”.

Thêm vào đó, trong sáng chế này, “khối hiện tại” có thể nghĩa là khối bao gồm cả khối thành phần độ chói và khối thành phần sắc độ hoặc “khối độ chói của khối hiện tại” trừ khi được nêu tường minh là khối sắc độ. Khối thành phần độ chói của khối hiện tại có thể được biểu diễn bằng cách bao gồm phần mô tả tường minh về khối thành phần độ chói này chẳng hạn như “khối độ chói” hoặc “khối độ chói hiện tại”. Thêm vào đó, “khối thành phần sắc độ của khối hiện tại” có thể được biểu diễn bằng cách bao gồm phần mô tả tường minh về khối thành phần sắc độ, chẳng hạn như “khối sắc độ” hoặc “khối sắc độ hiện tại”.

Trong sáng chế này, thuật ngữ “/” và “,” cần được hiểu là để thể hiện “và/hoặc.” Ví dụ, cách trình bày “A/B” và “A, B” có thể có nghĩa là “A và/hoặc B.” Hơn nữa, “A/B/C” và “A/B/C” có thể có nghĩa là “ít nhất một thành phần trong số A, B, và/hoặc C.”

Trong sáng chế này, thuật ngữ “hoặc” cần được diễn giải để thể hiện “và/hoặc.” Ví dụ, cách trình bày “A hoặc B” có thể bao gồm 1) chỉ “A”, 2) chỉ “B”, và/hoặc 3) cả “A và B”. Nói cách khác, trong sáng chế này, thuật ngữ “hoặc” cần được diễn giải là để thể hiện “thêm vào đó hoặc thay vào đó.”

Tổng quan về hệ thống lập mã video

Fig.1 là hình minh họa một cách sơ lược hệ thống lập mã video theo sáng chế.

Hệ thống lập mã video theo một phương án có thể bao gồm thiết bị mã hóa 10 và thiết bị giải mã 20. Thiết bị mã hóa 10 có thể phân phát thông tin hoặc dữ liệu video và/hoặc ảnh được mã hóa đến thiết bị giải mã 20 dưới dạng tệp tin hoặc tạo luồng qua phương tiện lưu trữ kỹ thuật số hoặc mạng.

Thiết bị mã hóa 10 theo một phương án có thể bao gồm bộ tạo ra nguồn video 11, đơn vị mã hóa 12 và bộ truyền 13. Thiết bị giải mã 20 theo một phương án có thể bao gồm bộ nhận 21, đơn vị giải mã 22 và bộ kết xuất 23. Đơn vị mã hóa 12 có thể được gọi là đơn vị mã hóa video/ảnh, và đơn vị giải mã 22 có thể được gọi là đơn vị giải mã video/ảnh. Bộ truyền 13 có thể được bao gồm trong đơn vị mã hóa 12. Bộ nhận 21 có thể được bao gồm trong đơn vị giải mã 22. Bộ kết xuất 23 có thể bao gồm bộ hiển thị và bộ hiển thị này có thể được tạo cấu hình như thiết bị tách biệt hoặc thành phần bên

ngoài.

Bộ tạo ra nguồn video 11 có thể thu nhận video/ảnh thông qua quy trình chụp, tổng hợp hoặc tạo ra video/ảnh. Bộ tạo ra nguồn video 11 có thể bao gồm thiết bị chụp video/ảnh và/hoặc thiết bị tạo ra video/ảnh. Thiết bị chụp video/ảnh có thể bao gồm, ví dụ, một hoặc nhiều camera, các phần lưu trữ video/ảnh chứa các video/ảnh đã được chụp trước đó, và dạng tương tự. Thiết bị tạo ra video/ảnh có thể bao gồm, ví dụ, các máy tính, máy tính bảng, và điện thoại thông minh, và có thể tạo ra các video/ảnh (theo cách điện tử). Ví dụ, video/ảnh ảo có thể được tạo ra qua máy tính hoặc dạng tương tự. Trong trường hợp này, quy trình chụp video/ảnh có thể được thay thế bởi quy trình tạo ra dữ liệu liên quan.

Đơn vị mã hóa 12 có thể mã hóa video/ảnh đầu vào. Đơn vị mã hóa 12 có thể thực hiện chuỗi các thủ tục như dự đoán, biến đổi, và lượng tử hóa để đạt được hiệu quả nén và lập mã. Đơn vị mã hóa 12 có thể xuất ra dữ liệu được mã hóa (thông tin video/ảnh được mã hóa) dưới dạng luồng bit.

Bộ truyền 13 có thể truyền thông tin hoặc dữ liệu video/ảnh được mã hóa được xuất ra dưới dạng luồng bit đến bộ nhận 21 của thiết bị giải mã 20 thông qua phương tiện lưu trữ kỹ thuật số hoặc mạng dưới dạng tệp tin hoặc tạo luồng. Phương tiện lưu trữ kỹ thuật số có thể bao gồm các phương tiện lưu trữ khác nhau như USB, SD, CD, DVD, Blu-ray, HDD, SSD, và loại tương tự. Bộ truyền 13 có thể bao gồm phần tử để tạo ra tệp tin phương tiện thông qua định dạng tệp tin được xác định trước và có thể bao gồm phần tử để truyền qua mạng phát quảng bá/truyền thông. Bộ nhận 21 có thể trích xuất/nhận luồng bit từ phương tiện lưu trữ hoặc mạng và truyền luồng bit đến đơn vị giải mã 22.

Đơn vị giải mã 22 có thể giải mã video/ảnh bằng cách thực hiện chuỗi các thủ tục như giải lượng tử hóa, biến đổi ngược, và dự đoán tương ứng với hoạt động của đơn vị mã hóa 12.

Bộ kết xuất 23 có thể kết xuất video/ảnh được giải mã. Video/ảnh được kết xuất có thể được hiển thị qua bộ hiển thị.

Tổng quan về thiết bị mã hóa ảnh

Fig.2 là hình minh họa sơ lược thiết bị mã hóa ảnh, mà phương án của sáng chế có thể áp dụng được với nó.

Như được thể hiện trên Fig.2, thiết bị mã hóa ảnh 100 có thể bao gồm bộ phân vùng ảnh 110, bộ trừ 115, bộ biến đổi 120, bộ lượng tử hóa 130, bộ giải lượng tử hóa 140, bộ biến đổi ngược 150, bộ cộng 155, bộ lọc 160, bộ nhớ 170, bộ dự đoán liên 180, bộ dự đoán nội 185 và bộ mã hóa entropy 190. Bộ dự đoán liên 180 và bộ dự đoán nội 185 có thể được gọi chung là “bộ dự đoán”. Bộ biến đổi 120, bộ lượng tử hóa 130, bộ giải lượng tử hóa 140 và bộ biến đổi ngược 150 có thể được bao gồm trong bộ xử lý phần dư. Bộ xử lý phần dư có thể còn bao gồm bộ trừ 115.

Tất cả hoặc ít nhất một số thành phần trong số nhiều thành phần tạo cấu hình thiết bị mã hóa ảnh 100 có thể được tạo cấu hình bởi một thành phần phần cứng (ví dụ, bộ mã hóa hoặc bộ xử lý) trong một số phương án. Thêm vào đó, bộ nhớ 170 có thể bao gồm bộ đệm hình ảnh được giải mã (decoded picture buffer, DPB) và có thể được tạo cấu hình bằng phương tiện lưu trữ kỹ thuật số.

Bộ phân vùng ảnh 110 có thể phân vùng ảnh (hoặc hình ảnh hoặc khung) đầu vào được nhập vào thiết bị mã hóa ảnh 100 thành một hoặc nhiều đơn vị xử lý. Ví dụ, đơn vị xử lý có thể được gọi là đơn vị lập mã (coding unit, CU). Đơn vị lập mã có thể được thu nhận bằng cách phân vùng theo cách đệ quy đơn vị cây lập mã (coding tree unit, CTU) hoặc đơn vị lập mã lớn nhất (largest coding unit, LCU) theo cấu trúc cây tứ phân cây nhị phân cây tam phân (quad-tree binary-tree ternary-tree, QT/BT/TT). Ví dụ, một đơn vị lập mã có thể được phân vùng thành nhiều đơn vị lập mã có độ sâu sâu hơn dựa trên cấu trúc cây tứ phân, cấu trúc cây nhị phân, và/hoặc cấu trúc tam phân. Để phân vùng đơn vị lập mã, cấu trúc cây tứ phân có thể được áp dụng đầu tiên và cấu trúc cây nhị phân và/hoặc cấu trúc tam phân có thể được áp dụng sau. Thủ tục lập mã theo sáng chế có thể được thực hiện dựa trên đơn vị lập mã cuối cùng mà nó không được phân vùng thêm nữa. Đơn vị lập mã lớn nhất có thể được sử dụng làm đơn vị lập mã cuối cùng, hoặc đơn vị lập mã có độ sâu sâu hơn thu được bằng cách phân vùng đơn vị lập mã lớn nhất có thể được sử dụng làm đơn vị lập mã cuối cùng. Ở đây, thủ tục lập mã có thể bao gồm thủ tục dự đoán, biến đổi, và xây dựng lại, mà chúng sẽ được mô tả sau. Như là một ví dụ khác, đơn vị xử lý của thủ tục lập mã có thể là đơn vị dự đoán (prediction unit, PU) hoặc đơn vị biến đổi (transform unit, TU). Đơn vị dự đoán và đơn vị biến đổi có thể được chia tách hoặc phân vùng từ đơn vị lập mã cuối cùng. Đơn vị dự đoán có thể là đơn vị của việc dự đoán mẫu, và đơn vị biến đổi có thể là đơn vị để dẫn

ra hệ số biến đổi và/hoặc đơn vị để dẫn ra tín hiệu phần dư từ hệ số biến đổi này.

Bộ dự đoán (bộ dự đoán liên 180 hoặc bộ dự đoán nội 185) có thể thực hiện việc dự đoán trên khối cần được xử lý (khối hiện tại) và tạo ra khối được dự đoán bao gồm các mẫu dự đoán cho khối hiện tại. Bộ dự đoán có thể xác định xem liệu việc dự đoán nội hay việc dự đoán liên được áp dụng trên cơ sở khối hiện tại hay CU. Bộ dự đoán có thể tạo ra thông tin khác nhau liên quan đến việc dự đoán khối hiện tại và truyền thông tin được tạo ra đến bộ mã hóa entropy 190. Thông tin về việc dự đoán có thể được mã hóa trong bộ mã hóa entropy 190 và được xuất ra dưới dạng luồng bit.

Bộ dự đoán nội 185 có thể dự đoán khối hiện tại bằng cách tham chiếu đến các mẫu trong hình ảnh hiện tại. Các mẫu được tham chiếu có thể được đặt trong vùng lân cận của khối hiện tại hoặc có thể được đặt tách ra tùy theo chế độ dự đoán nội và/hoặc kỹ thuật dự đoán nội. Các chế độ dự đoán nội có thể bao gồm nhiều chế độ không hướng và nhiều chế độ có hướng. Các chế độ không hướng có thể bao gồm, ví dụ, chế độ DC và chế độ phẳng. Chế độ có hướng có thể bao gồm, ví dụ, 33 chế độ dự đoán có hướng hoặc 65 chế độ dự đoán có hướng tùy theo mức độ chi tiết của hướng dự đoán. Tuy nhiên, điều này chỉ đơn thuần là ví dụ, nhiều hoặc ít chế độ dự đoán có hướng hơn có thể được sử dụng phụ thuộc vào việc cài đặt. Bộ dự đoán nội 185 có thể xác định chế độ dự đoán được áp dụng cho khối hiện tại bằng cách sử dụng chế độ dự đoán được áp dụng cho khối lân cận.

Bộ dự đoán liên 180 có thể dẫn ra khối được dự đoán cho khối hiện tại dựa trên khối tham chiếu (mảng mẫu tham chiếu) được chỉ rõ bởi vector chuyển động trên hình ảnh tham chiếu. Trong trường hợp này, để làm giảm lượng thông tin chuyển động được truyền trong chế độ dự đoán liên, thông tin chuyển động có thể được dự đoán theo các đơn vị là các khối, các khối con, hoặc các mẫu dựa trên sự tương quan của thông tin chuyển động giữa khối lân cận và khối hiện tại. Thông tin chuyển động có thể bao gồm vector chuyển động và chỉ số hình ảnh tham chiếu. Thông tin chuyển động có thể còn bao gồm thông tin hướng dự đoán liên (dự đoán L0, dự đoán L1, song dự đoán (Bi prediction), v.v.). Trong trường hợp dự đoán liên, khối lân cận có thể bao gồm khối lân cận theo không gian có mặt trong hình ảnh hiện tại và khối lân cận theo thời gian có mặt trong hình ảnh tham chiếu. Hình ảnh tham chiếu chứa khối tham chiếu và hình ảnh tham chiếu chứa khối lân cận theo thời gian có thể là giống nhau hoặc khác nhau. Khối lân

cận theo thời gian có thể được gọi là khối tham chiếu được đặt cùng vị trí, CU được đặt cùng chỗ (co-located CU, colCU), và dạng tương tự. Hình ảnh tham chiếu chứa khối lân cận theo thời gian có thể được gọi là hình ảnh được đặt cùng vị trí (collocated picture, colPic). Ví dụ, bộ dự đoán liên 180 có thể tạo cấu hình danh sách ứng viên thông tin chuyển động dựa trên các khối lân cận và tạo ra thông tin chỉ rõ ứng viên nào được sử dụng để dẫn ra vector chuyển động và/hoặc chỉ số hình ảnh tham chiếu của khối hiện tại. Việc dự đoán liên có thể được thực hiện dựa trên các chế độ dự đoán khác nhau. Ví dụ, trong trường hợp của chế độ bỏ qua (skip) và chế độ hợp nhất (merge), bộ dự đoán liên 180 có thể sử dụng thông tin chuyển động của khối lân cận làm thông tin chuyển động của khối hiện tại. Trong trường hợp của chế độ bỏ qua, không giống chế độ hợp nhất, tín hiệu phần dư có thể không được truyền. Trong trường hợp chế độ dự đoán vector chuyển động (motion vector prediction, MVP), vector chuyển động của khối lân cận có thể được sử dụng làm bộ dự đoán vector chuyển động, và vector chuyển động của khối hiện tại có thể được báo hiệu bằng cách mã hóa chênh lệch vector chuyển động và chỉ dấu cho bộ dự đoán vector chuyển động. Chênh lệch vector chuyển động có thể nghĩa là chênh lệch giữa vector chuyển động của khối hiện tại và bộ dự đoán vector chuyển động.

Bộ dự đoán có thể tạo ra tín hiệu dự đoán dựa trên các phương pháp dự đoán và các kỹ thuật dự đoán khác nhau được mô tả dưới đây. Ví dụ, bộ dự đoán có thể không chỉ áp dụng việc dự đoán nội hoặc dự đoán liên mà còn áp dụng đồng thời cả việc dự đoán nội và dự đoán liên để dự đoán khối hiện tại. Phương pháp dự đoán áp dụng đồng thời cả việc dự đoán nội và việc dự đoán liên cho việc dự đoán khối hiện tại có thể được gọi là việc dự đoán liên và nội kết hợp (combined inter and intra prediction, CIIP). Thêm vào đó, bộ dự đoán có thể thực hiện việc sao chép khối nội (Intra block copy, IBC) cho việc dự đoán khối hiện tại. Việc sao chép khối nội có thể được sử dụng cho việc lập mã ảnh/video nội dung của trò chơi hoặc dạng tương tự, ví dụ, lập mã nội dung màn (screen content coding, SCC). IBC là phương pháp dự đoán hình ảnh hiện tại nhờ sử dụng khối tham chiếu được xây dựng lại trước đó trong hình ảnh hiện tại ở vị trí đặt cách khối khối hiện tại một khoảng cách được xác định trước. Khi IBC được áp dụng, vị trí của khối tham chiếu trong hình ảnh hiện tại có thể được mã hóa dưới dạng vector (vector khối) tương ứng với khoảng cách được xác định trước này. IBC về cơ bản thực hiện việc dự đoán trong hình ảnh hiện tại, nhưng có thể được thực hiện tương tự với việc dự đoán

liên ở chỗ khối tham chiếu được dẫn ra bên trong hình ảnh hiện tại. Tức là, IBC có thể sử dụng ít nhất một trong các kỹ thuật dự đoán liên được mô tả trong sáng chế này.

Tín hiệu dự đoán được tạo ra bởi bộ dự đoán có thể được sử dụng để tạo ra tín hiệu được xây dựng lại hoặc để tạo ra tín hiệu phần dư. Bộ trừ 115 có thể tạo ra tín hiệu phần dư (khối phần dư hoặc mảng mẫu phần dư) bằng cách lấy tín hiệu ảnh đầu vào (khối gốc hoặc mảng mẫu gốc) trừ đi tín hiệu dự đoán (khối được dự đoán hoặc mảng mẫu dự đoán) được xuất ra từ bộ dự đoán. Tín hiệu phần dư được tạo ra có thể được truyền đến bộ biến đổi 120.

Bộ biến đổi 120 có thể tạo ra các hệ số biến đổi bằng cách áp dụng kỹ thuật biến đổi cho tín hiệu phần dư. Ví dụ, kỹ thuật biến đổi có thể bao gồm ít nhất một biến đổi trong số biến đổi cosin rời rạc (discrete cosine transform, DCT), biến đổi sin rời rạc (discrete sine transform, DST), biến đổi Karhunen-loève (Karhunen-loève transform, KLT), biến đổi dựa trên đồ thị (graph-based transform, GBT), hoặc biến đổi phi tuyến có điều kiện (conditionally non-linear transform, CNT). Ở đây, GBT nghĩa là việc biến đổi thu được từ đồ thị khi thông tin mối quan hệ giữa các điểm ảnh được biểu diễn bằng đồ thị. CNT đề cập đến việc biến đổi thu được dựa trên tín hiệu dự đoán được tạo ra nhờ sử dụng tất cả các điểm ảnh được xây dựng lại trước đó. Thêm vào đó, quy trình biến đổi có thể được áp dụng cho các khối điểm ảnh hình vuông có cùng kích thước hoặc có thể được áp dụng cho các khối có kích thước biến thiên thay vì hình vuông.

Bộ lượng tử hóa 130 có thể lượng tử hóa các hệ số biến đổi và truyền chúng đến bộ mã hóa entropy 190. Bộ mã hóa entropy 190 có thể mã hóa tín hiệu được lượng tử hóa (thông tin về các hệ số biến đổi được lượng tử hóa) và xuất ra luồng bit. Thông tin về các hệ số biến đổi được lượng tử hóa có thể được gọi là thông tin phần dư. Bộ lượng tử hóa 130 có thể sắp xếp lại các hệ số biến đổi được lượng tử hóa ở loại khối thành dạng vector một chiều dựa trên thứ tự quét hệ số và tạo ra thông tin về các hệ số biến đổi được lượng tử hóa dựa trên các hệ số biến đổi được lượng tử hóa ở dạng vector một chiều này.

Bộ mã hóa entropy 190 có thể thực hiện các phương pháp mã hóa khác nhau ví dụ chẳng hạn như Golomb hàm số mũ, lập mã chiều dài biến đổi thích ứng ngữ cảnh (context-adaptive variable length coding, CAVLC), lập mã số học nhị phân thích ứng ngữ cảnh (context-adaptive binary arithmetic coding, CABAC), và dạng tương tự. Bộ

mã hóa entropy 190 có thể mã hóa thông tin cần thiết cho việc xây dựng lại video/ảnh ngoài các hệ số biến đổi được lượng tử hóa (ví dụ, các giá trị của các phần tử cú pháp, v.v.) cùng nhau hoặc tách biệt. Thông tin được mã hóa (ví dụ, thông tin video/ảnh được mã hóa) có thể được truyền hoặc được lưu trữ theo các đơn vị là các lớp trừu tượng mạng (network abstraction layer, NAL) dưới dạng luồng bit. Thông tin video/ảnh có thể còn bao gồm thông tin về các tập tham số khác nhau như tập tham số thích ứng (adaptation parameter set, APS), tập tham số hình ảnh (picture parameter set, PPS), tập tham số trình tự (sequence parameter set, SPS), hoặc tập tham số video (video parameter set, VPS). Thêm vào đó, thông tin video/ảnh có thể còn bao gồm thông tin ràng buộc chung. Thông tin được báo hiệu, thông tin được truyền và/hoặc các phần tử cú pháp được mô tả trong sáng chế này có thể được mã hóa qua thủ tục mã hóa được mô tả trên đây và được đưa vào trong luồng bit.

Luồng bit có thể được truyền qua mạng hoặc có thể được lưu trữ trong phương tiện lưu trữ kỹ thuật số. Mạng này có thể bao gồm mạng phát quảng bá và/hoặc mạng truyền thông, và phương tiện lưu trữ kỹ thuật số này có thể bao gồm các phương tiện lưu trữ khác nhau như USB, SD, CD, DVD, Blu-ray, HDD, SSD, và dạng tương tự. Bộ truyền (không được thể hiện) truyền tín hiệu được xuất ra từ bộ mã hóa entropy 190 và/hoặc đơn vị lưu trữ (không được thể hiện) lưu trữ tín hiệu này có thể được đưa vào làm phần tử bên trong/bên ngoài của thiết bị mã hóa ảnh 100. Theo cách khác, bộ truyền có thể được bố trí như thành phần của bộ mã hóa entropy 190.

Các hệ số biến đổi được lượng tử hóa được xuất ra từ bộ lượng tử hóa 130 có thể được sử dụng để tạo ra tín hiệu phần dư. Ví dụ, tín hiệu phần dư (khối phần dư hoặc các mẫu phần dư) có thể được xây dựng lại bằng cách áp dụng việc giải lượng tử hóa và biến đổi ngược đối với các hệ số biến đổi được lượng tử hóa qua bộ giải lượng tử hóa 140 và bộ biến đổi ngược 150.

Bộ cộng 155 cộng tín hiệu phần dư được xây dựng lại vào tín hiệu dự đoán được xuất ra từ bộ dự đoán liên 180 hoặc bộ dự đoán nội 185 để tạo ra tín hiệu được xây dựng lại (hình ảnh được xây dựng lại, khối được xây dựng lại, mảng mẫu được xây dựng lại). Nếu không có phần dư nào cho khối cần được xử lý, như trường hợp mà trong đó chế độ bỏ qua được áp dụng, thì khối được dự đoán có thể được sử dụng làm khối được xây dựng lại. Bộ cộng 155 có thể được gọi là bộ xây dựng lại hoặc bộ tạo ra khối được xây

dựng lại. Tín hiệu được xây dựng lại được tạo ra có thể được sử dụng cho việc dự đoán nội khối tiếp theo cần được xử lý trong hình ảnh hiện tại và có thể được sử dụng cho việc dự đoán liên hình ảnh tiếp theo qua việc lọc như được mô tả dưới đây.

Bộ lọc 160 có thể cải thiện chất lượng ảnh chủ quan/khách quan bằng cách áp dụng việc lọc cho tín hiệu được xây dựng lại. Ví dụ, bộ lọc 160 có thể tạo ra hình ảnh được xây dựng lại được cải biến bằng cách áp dụng các phương pháp lọc khác nhau đối với hình ảnh được xây dựng lại và lưu trữ hình ảnh được xây dựng lại được cải biến trong bộ nhớ 170, cụ thể, DPB của bộ nhớ 170. Các phương pháp lọc khác nhau này có thể bao gồm, ví dụ, lọc khử khối, phân bù thích ứng mẫu, bộ lọc vòng thích ứng, bộ lọc song phương, và dạng tương tự. Bộ lọc 160 có thể tạo ra thông tin khác nhau liên quan đến việc lọc và truyền thông tin được tạo ra đến bộ mã hóa entropy 190 như sẽ được mô tả sau trong phần mô tả về mỗi phương pháp lọc. Thông tin liên quan đến việc lọc có thể được mã hóa bởi bộ mã hóa entropy 190 và xuất ra dưới dạng luồng bit.

Hình ảnh được xây dựng lại được cải biến được truyền đến bộ nhớ 170 có thể được sử dụng làm hình ảnh tham chiếu trong bộ dự đoán liên 180. Khi việc dự đoán liên được áp dụng thông qua thiết bị mã hóa ảnh 100, việc không khớp dự đoán giữa thiết bị mã hóa ảnh 100 và thiết bị giải mã ảnh có thể tránh được và hiệu suất mã hóa có thể được cải thiện.

DPB của bộ nhớ 170 có thể lưu trữ hình ảnh được xây dựng lại được cải biến để sử dụng làm hình ảnh tham chiếu trong bộ dự đoán liên 180. Bộ nhớ 170 có thể lưu trữ thông tin chuyển động của khối mà từ đó thông tin chuyển động trong hình ảnh hiện tại được dẫn ra (hoặc được mã hóa) và/hoặc thông tin chuyển động của các khối trong hình ảnh mà chúng đã được xây dựng lại. Thông tin chuyển động được lưu trữ có thể được truyền đến bộ dự đoán liên 180 và được sử dụng làm thông tin chuyển động của khối lân cận theo không gian hoặc thông tin chuyển động của khối lân cận theo thời gian. Bộ nhớ 170 có thể lưu trữ các mẫu được xây dựng lại của các khối được xây dựng lại trong hình ảnh hiện tại và có thể chuyển các mẫu được xây dựng lại đến bộ dự đoán nội 185.

Tổng quan về thiết bị giải mã ảnh

Fig.3 là hình minh họa sơ lược thiết bị giải mã ảnh, mà phương án của sáng chế có thể áp dụng được với nó.

Như được thể hiện trên Fig.3, thiết bị giải mã ảnh 200 có thể bao gồm bộ giải mã

entropy 210, bộ giải lượng tử hóa 220, bộ biến đổi ngược 230, bộ cộng 235, bộ lọc 240, bộ nhớ 250, bộ dự đoán liên 260 và bộ dự đoán nội 265. Bộ dự đoán liên 260 và bộ dự đoán nội 265 có thể được gọi chung là “bộ dự đoán”. Bộ giải lượng tử hóa 220 và bộ biến đổi ngược 230 có thể được bao gồm trong bộ xử lý phân dư.

Tất cả hoặc ít nhất một số thành phần trong số nhiều thành phần tạo cấu hình thiết bị giải mã ảnh 200 có thể được tạo cấu hình bằng thành phần phần cứng (ví dụ, bộ giải mã hoặc bộ xử lý) theo một phương án. Thêm vào đó, bộ nhớ 250 có thể bao gồm bộ đệm hình ảnh được giải mã (decoded picture buffer, DPB) hoặc có thể được tạo cấu hình bằng phương tiện lưu trữ kỹ thuật số.

Thiết bị giải mã ảnh 200, mà nó đã nhận luồng bit bao gồm thông tin video/ảnh, có thể xây dựng lại ảnh bằng cách thực hiện quy trình tương ứng với quy trình được thực hiện bởi thiết bị mã hóa ảnh 100 trên Fig.2. Ví dụ, thiết bị giải mã ảnh 200 có thể thực hiện việc giải mã sử dụng đơn vị xử lý được áp dụng trong thiết bị mã hóa ảnh. Do đó, đơn vị xử lý của việc giải mã, ví dụ, có thể là đơn vị lập mã. Đơn vị lập mã này có thể được thu nhận bằng cách phân vùng đơn vị cây lập mã hoặc đơn vị lập mã lớn nhất. Tín hiệu ảnh được xây dựng lại được giải mã và được xuất ra qua thiết bị giải mã ảnh 200 có thể được tái tạo qua thiết bị tái tạo (không được thể hiện).

Thiết bị giải mã ảnh 200 có thể nhận tín hiệu được xuất ra từ thiết bị mã hóa ảnh trên Fig.2 dưới dạng luồng bit. Tín hiệu nhận được có thể được giải mã qua bộ giải mã entropy 210. Ví dụ, bộ giải mã entropy 210 có thể phân tích cú pháp luồng bit để dẫn ra thông tin (ví dụ, thông tin video/ảnh) cần thiết cho việc xây dựng lại ảnh (hoặc xây dựng lại hình ảnh). Thông tin video/ảnh có thể còn bao gồm thông tin về các tập tham số khác nhau như tập tham số thích ứng (adaptation parameter set, APS), tập tham số hình ảnh (picture parameter set, PPS), tập tham số trình tự (sequence parameter set, SPS), hoặc tập tham số video (video parameter set, VPS). Thêm vào đó, thông tin video/ảnh có thể còn bao gồm thông tin ràng buộc chung. Thiết bị giải mã ảnh có thể còn giải mã hình ảnh dựa trên thông tin về tập tham số và/hoặc thông tin ràng buộc chung. Thông tin và/hoặc các phần tử cú pháp được báo hiệu/được nhận được mô tả trong sáng chế này có thể được giải mã qua thủ tục giải mã và được thu từ luồng bit. Ví dụ, bộ giải mã entropy 210 giải mã thông tin trong luồng bit dựa trên phương pháp lập mã như lập mã Golomb hàm số mũ, CAVLC, hoặc CABAC, và các giá trị được xuất ra của các phần

tử cú pháp cần thiết cho việc xây dựng lại ảnh và các giá trị được lượng tử hóa của các hệ số biến đổi cho phần dư. Cụ thể hơn, phương pháp giải mã entropy CABAC có thể nhận ngăn (bin) tương ứng với mỗi phần tử cú pháp trong luồng bit, xác định mô hình ngữ cảnh sử dụng thông tin về phần tử cú pháp mục tiêu giải mã, thông tin giải mã của khối lân cận và khối mục tiêu giải mã hoặc thông tin về ký hiệu/ngăn được giải mã trong giai đoạn trước, và thực hiện việc giải mã số học trên ngăn này bằng cách dự đoán xác suất xuất hiện ngăn theo mô hình ngữ cảnh xác định được, và tạo ra ký hiệu tương ứng với giá trị của mỗi phần tử cú pháp. Trong trường hợp này, phương pháp giải mã entropy CABAC có thể cập nhật mô hình ngữ cảnh bằng cách sử dụng thông tin về ký hiệu/ngăn được giải mã cho mô hình ngữ cảnh của ký hiệu/ngăn tiếp theo sau khi xác định mô hình ngữ cảnh. Thông tin liên quan đến việc dự đoán trong số thông tin được giải mã bởi bộ giải mã entropy 210 có thể được cung cấp cho bộ dự đoán (bộ dự đoán liên 260 và bộ dự đoán nội 265), và giá trị phần dư mà việc giải mã entropy được thực hiện trên đó trong bộ giải mã entropy 210, tức là, các hệ số biến đổi được lượng tử hóa và thông tin tham số liên quan, có thể được cung cấp cho bộ giải lượng tử hóa 220. Thêm vào đó, thông tin về việc lọc trong số thông tin được giải mã bởi bộ giải mã entropy 210 có thể được cung cấp cho bộ lọc 240. Trong khi đó, bộ nhận (không được thể hiện) để nhận tín hiệu được xuất ra từ thiết bị mã hóa ảnh có thể còn được tạo cấu hình như phần tử bên trong/bên ngoài của thiết bị giải mã ảnh 200, hoặc bộ nhận này có thể là thành phần của bộ giải mã entropy 210.

Trong khi đó, thiết bị giải mã ảnh theo sáng chế có thể được gọi là thiết bị giải mã video/ảnh/hình ảnh. Thiết bị giải mã ảnh có thể được phân loại thành bộ giải mã thông tin (bộ giải mã thông tin video/ảnh/hình ảnh) và bộ giải mã mẫu (bộ giải mã mẫu video/ảnh/hình ảnh). Bộ giải mã thông tin có thể bao gồm bộ giải mã entropy 210. Bộ giải mã mẫu có thể bao gồm ít nhất một trong số bộ giải lượng tử hóa 220, bộ biến đổi ngược 230, bộ cộng 235, bộ lọc 240, bộ nhớ 250, bộ dự đoán liên 160 hoặc bộ dự đoán nội 265.

Bộ giải lượng tử hóa 220 có thể giải lượng tử hóa các hệ số biến đổi được lượng tử hóa và xuất ra các hệ số biến đổi. Bộ giải lượng tử hóa 220 có thể sắp xếp lại các hệ số biến đổi được lượng tử hóa dưới dạng khối hai chiều. Trong trường hợp này, việc sắp xếp lại có thể được thực hiện dựa trên thứ tự quét hệ số được thực hiện trong thiết bị mã

hóa ảnh. Bộ giải lượng tử hóa 220 có thể thực hiện việc giải lượng tử hóa trên các hệ số biến đổi được lượng tử hóa bằng cách sử dụng tham số lượng tử hóa (ví dụ, thông tin kích thước bước lượng tử hóa) và thu các hệ số biến đổi.

Bộ biến đổi ngược 230 có thể biến đổi ngược các hệ số biến đổi để thu tín hiệu phần dư (khối phần dư, mảng mẫu phần dư).

Bộ dự đoán có thể thực hiện việc dự đoán trên khối hiện tại và tạo ra khối được dự đoán bao gồm các mẫu dự đoán cho khối hiện tại. Bộ dự đoán có thể xác định xem liệu việc dự đoán nội hay việc dự đoán liên được áp dụng cho khối hiện tại dựa trên thông tin về việc dự đoán được xuất ra từ bộ giải mã entropy 210 và có thể xác định chế độ dự đoán (kỹ thuật dự đoán) nội/liên cụ thể.

Điều này là giống như được mô tả trong bộ dự đoán của thiết bị mã hóa ảnh 100 ở chỗ bộ dự đoán có thể tạo ra tín hiệu dự đoán dựa trên các phương pháp (các kỹ thuật) dự đoán khác nhau mà chúng sẽ được mô tả sau.

Bộ dự đoán nội 265 có thể dự đoán khối hiện tại bằng cách tham chiếu đến các mẫu trong hình ảnh hiện tại. Phần mô tả của bộ dự đoán nội 185 được áp dụng tương đương với bộ dự đoán nội 265.

Bộ dự đoán liên 260 có thể dẫn ra khối được dự đoán cho khối hiện tại dựa trên khối tham chiếu (mảng mẫu tham chiếu) được chỉ rõ bởi vectơ chuyển động trên hình ảnh tham chiếu. Trong trường hợp này, để làm giảm lượng thông tin chuyển động được truyền trong chế độ dự đoán liên, thông tin chuyển động có thể được dự đoán theo các đơn vị là các khối, các khối con, hoặc các mẫu dựa trên sự tương quan của thông tin chuyển động giữa khối lân cận và khối hiện tại. Thông tin chuyển động có thể bao gồm vectơ chuyển động và chỉ số hình ảnh tham chiếu. Thông tin chuyển động có thể còn bao gồm thông tin hướng dự đoán liên (dự đoán L0, dự đoán L1, song dự đoán (Bi prediction), v.v.). Trong trường hợp dự đoán liên, khối lân cận có thể bao gồm khối lân cận theo không gian có mặt trong hình ảnh hiện tại và khối lân cận theo thời gian có mặt trong hình ảnh tham chiếu. Ví dụ, bộ dự đoán liên 260 có thể tạo cấu trúc danh sách ứng viên thông tin chuyển động dựa trên các khối lân cận và dẫn ra vectơ chuyển động của khối hiện tại và/hoặc chỉ số hình ảnh tham chiếu dựa trên thông tin chọn ứng viên nhận được. Việc dự đoán liên có thể được thực hiện dựa trên các chế độ dự đoán khác nhau, và thông tin về việc dự đoán có thể bao gồm thông tin chỉ rõ chế độ dự đoán liên

cho khối hiện tại.

Bộ cộng 235 có thể tạo ra tín hiệu được xây dựng lại (hình ảnh được xây dựng lại, khối được xây dựng lại, mảng mẫu được xây dựng lại) bằng cách cộng tín hiệu phần dư thu được vào tín hiệu dự đoán (khối được dự đoán, mảng mẫu được dự đoán) được xuất ra từ bộ dự đoán (bao gồm bộ dự đoán liên 260 và/hoặc bộ dự đoán nội 265). Nếu không có phần dư nào cho khối cần được xử lý, như khi chế độ bỏ qua được áp dụng, thì khối được dự đoán có thể được sử dụng làm khối được xây dựng lại. Phần mô tả của bộ cộng 155 là có thể áp dụng tương đương được với bộ cộng 235. Bộ cộng 235 có thể được gọi là bộ xây dựng lại hoặc bộ tạo ra khối được xây dựng lại. Tín hiệu được xây dựng lại được tạo ra có thể được sử dụng cho việc dự đoán nội khối tiếp theo cần được xử lý trong hình ảnh hiện tại và có thể được sử dụng cho việc dự đoán liên hình ảnh tiếp theo qua việc lọc như được mô tả dưới đây.

Bộ lọc 240 có thể cải thiện chất lượng ảnh chủ quan/khách quan bằng cách áp dụng việc lọc cho tín hiệu được xây dựng lại. Ví dụ, bộ lọc 240 có thể tạo ra hình ảnh được xây dựng lại được cải biến bằng cách áp dụng các phương pháp lọc khác nhau đối với hình ảnh được xây dựng lại và lưu trữ hình ảnh được xây dựng lại được cải biến trong bộ nhớ 250, cụ thể, DPB của bộ nhớ 250. Các phương pháp lọc khác nhau này có thể bao gồm, ví dụ, lọc khử khối, phần bù thích ứng mẫu, bộ lọc vòng thích ứng, bộ lọc song phương, và dạng tương tự.

Hình ảnh được xây dựng lại (được cải biến) được lưu trữ trong DPB của bộ nhớ 250 có thể được sử dụng làm hình ảnh tham chiếu trong bộ dự đoán liên 260. Bộ nhớ 250 có thể lưu trữ thông tin chuyển động của khối mà thông tin chuyển động trong hình ảnh hiện tại được dẫn ra (hoặc được giải mã) từ đó và/hoặc thông tin chuyển động của các khối trong hình ảnh mà chúng đã được xây dựng lại sẵn. Thông tin chuyển động được lưu trữ có thể được truyền đến bộ dự đoán liên 260 sao cho nó được sử dụng làm thông tin chuyển động của khối lân cận theo không gian hoặc thông tin chuyển động của khối lân cận theo thời gian. Bộ nhớ 250 có thể lưu trữ các mẫu được xây dựng lại của các khối được xây dựng lại trong hình ảnh hiện tại và chuyển các mẫu được xây dựng lại này đến bộ dự đoán nội 265.

Trong sáng chế này, các phương án được mô tả trong bộ lọc 160, bộ dự đoán liên 180, và bộ dự đoán nội 185 của thiết bị mã hóa ảnh 100 có thể được áp dụng tương

đương hoặc tương ứng với bộ lọc 240, bộ dự đoán liên 260, và bộ dự đoán nội 265 của thiết bị giải mã ảnh 200.

Thủ tục lập mã ảnh/video nói chung

Trong việc lập mã ảnh/video, hình ảnh tạo cấu hình ảnh/video có thể được mã hóa/giải mã theo thứ tự giải mã. Thứ tự hình ảnh tương ứng với thứ tự xuất ra của hình ảnh được giải mã có thể được thiết lập khác với thứ tự giải mã, và, dựa trên điều này, không chỉ việc dự đoán thuận chiều mà cả việc dự đoán ngược chiều có thể được thực hiện trong quá trình dự đoán liên.

Fig.4 là lưu đồ minh họa thủ tục giải mã ảnh, mà phương án (các phương án) của sáng chế có thể áp dụng được với nó.

Mỗi thủ tục được thể hiện trên Fig.4 đều có thể được thực hiện bởi thiết bị giải mã ảnh trên Fig.3. Ví dụ, bước S410 có thể được thực hiện bởi bộ giải mã entropy 210, bước S420 có thể được thực hiện bởi bộ dự đoán bao gồm các bộ dự đoán 265 và 260, bước S430 có thể được thực hiện bởi bộ xử lý phần dư 220 và 230, bước S440 có thể được thực hiện bởi bộ cộng 235, và bước S450 có thể được thực hiện bởi bộ lọc 240. Bước S410 có thể bao gồm thủ tục giải mã thông tin được mô tả trong sáng chế này, bước S420 có thể bao gồm thủ tục dự đoán liên/nội được mô tả trong sáng chế này, bước S430 có thể bao gồm thủ tục xử lý phần dư được mô tả trong sáng chế này, bước S440 có thể bao gồm thủ tục xây dựng lại khối/hình ảnh được mô tả trong sáng chế này, và bước S450 có thể bao gồm thủ tục lọc trong vòng được mô tả trong sáng chế này.

Tham chiếu đến Fig.4, thủ tục giải mã hình ảnh có thể sơ lược bao gồm thủ tục (S410) để thu thông tin ảnh/video (qua việc giải mã) từ luồng bit, thủ tục xây dựng lại hình ảnh (S420 đến S440) và thủ tục lọc trong vòng (S450) đối với hình ảnh được xây dựng lại. Thủ tục xây dựng lại hình ảnh có thể được thực hiện dựa trên các mẫu dự đoán và các mẫu phần dư được thu qua việc dự đoán liên/nội (S420) và việc xử lý phần dư (S430) (việc giải lượng tử hóa và biến đổi ngược hệ số biến đổi được lượng tử hóa) được mô tả trong sáng chế này. Hình ảnh được xây dựng lại được cải biến có thể được tạo ra qua thủ tục lọc trong vòng đối với hình ảnh được xây dựng lại được tạo ra qua thủ tục xây dựng lại hình ảnh. Trong trường hợp này, hình ảnh được xây dựng lại được cải biến có thể được xuất ra dưới dạng hình ảnh được giải mã, được lưu trữ trong bộ đệm hình ảnh được giải mã (DPB) của bộ nhớ 250 và được sử dụng làm hình ảnh tham chiếu trong

thủ tục dự đoán liên khi sau này giải mã hình ảnh này. Thủ tục lọc trong vòng (S450) có thể được lược bỏ. Trong trường hợp này, hình ảnh được xây dựng lại có thể được xuất ra dưới dạng hình ảnh được giải mã, được lưu trữ trong DPB của bộ nhớ 250, và được sử dụng làm hình ảnh tham chiếu trong thủ tục dự đoán liên khi sau này giải mã hình ảnh này. Thủ tục lọc trong vòng (S450) có thể bao gồm thủ tục lọc khử khối, thủ tục phân bù thích ứng mẫu (sample adaptive offset, SAO), thủ tục bộ lọc vòng thích ứng (adaptive loop filter, ALF) và/hoặc thủ tục bộ lọc song phương (bi-lateral), như được mô tả trên đây, mà một vài hoặc tất cả trong số chúng có thể được lược bỏ. Thêm vào đó, một hoặc một số thành phần trong số thủ tục lọc khử khối, thủ tục phân bù thích ứng mẫu (SAO), thủ tục bộ lọc vòng thích ứng (ALF) và/hoặc thủ tục bộ lọc song phương có thể được áp dụng tuần tự hoặc tất cả trong số chúng có thể được áp dụng tuần tự. Ví dụ, sau khi thủ tục lọc khử khối được áp dụng cho hình ảnh được xây dựng lại, thủ tục SAO có thể được thực hiện. Theo cách khác, sau khi thủ tục lọc khử khối được áp dụng cho hình ảnh được xây dựng lại, thủ tục ALF có thể được thực hiện. Việc này có thể được thực hiện tương tự kể cả trong thiết bị mã hóa.

Fig.5 là lưu đồ minh họa thủ tục mã hóa ảnh, mà phương án của sáng chế có thể áp dụng được với nó.

Mỗi thủ tục được thể hiện trên Fig.5 đều có thể được thực hiện bởi thiết bị mã hóa ảnh trên Fig.2. Ví dụ, bước S510 có thể được thực hiện bởi các bộ dự đoán 185 và 180, bước S520 có thể được thực hiện bởi bộ xử lý phần dư 115, 120 và 130, và bước S530 có thể được thực hiện trong bộ mã hóa entropy 190. Bước S510 có thể bao gồm thủ tục dự đoán liên/nội được mô tả trong sáng chế này, bước S520 có thể bao gồm thủ tục xử lý phần dư được mô tả trong sáng chế này, và bước S530 có thể bao gồm thủ tục mã hóa thông tin được mô tả trong sáng chế này.

Tham chiếu đến Fig.5, thủ tục mã hóa hình ảnh có thể sơ lược bao gồm không chỉ thủ tục để mã hóa và xuất ra thông tin cho việc xây dựng lại hình ảnh (ví dụ, thông tin dự đoán, thông tin phần dư, thông tin phân vùng, v.v.) dưới dạng luồng bit mà cả thủ tục để tạo ra hình ảnh được xây dựng lại cho hình ảnh hiện tại và thủ tục (tùy chọn) để áp dụng việc lọc trong vòng cho hình ảnh được xây dựng lại. Thiết bị mã hóa có thể dẫn ra các mẫu phần dư (được cải biến) từ hệ số biến đổi được lượng tử hóa qua bộ giải lượng tử hóa 140 và bộ biến đổi ngược 150, và tạo ra hình ảnh được xây dựng lại dựa

trên các mẫu dự đoán mà chúng là đầu ra của bước S510 và các mẫu phần dư (được cải biến) này. Hình ảnh được xây dựng lại được tạo ra theo cách này có thể là bằng với hình ảnh được xây dựng lại được tạo ra trong thiết bị giải mã. Hình ảnh được xây dựng lại được cải biến có thể được tạo ra qua thủ tục lọc trong vòng đối với hình ảnh được xây dựng lại này. Trong trường hợp này, hình ảnh được xây dựng lại được cải biến có thể được lưu trữ trong bộ đệm hình ảnh được giải mã của bộ nhớ 170, và có thể được sử dụng làm hình ảnh tham chiếu trong thủ tục dự đoán liên khi sau này mã hóa hình ảnh, tương tự với thiết bị giải mã. Như được mô tả trên đây, trong một số trường hợp, một phần hoặc toàn bộ thủ tục lọc trong vòng có thể được lược bỏ. Khi thủ tục lọc trong vòng được thực hiện, thông tin (tham số) liên quan đến việc lọc (trong vòng) có thể được mã hóa trong bộ mã hóa entropy 190 và được xuất ra dưới dạng luồng bit, và thiết bị giải mã có thể thực hiện thủ tục lọc trong vòng nhờ sử dụng cùng phương pháp với thiết bị mã hóa dựa trên thông tin liên quan đến việc lọc này.

Qua thủ tục lọc trong vòng như vậy, tiếng ồn xảy ra trong việc lập mã ảnh/video, chẳng hạn như thành phần lạ tạo khối và thành phần lạ tiếng vọng, có thể được làm giảm và chất lượng trực quan chủ quan/khách quan có thể được cải thiện. Thêm vào đó, qua việc thực hiện thủ tục lọc trong vòng trong cả thiết bị mã hóa và thiết bị giải mã, thiết bị mã hóa và thiết bị giải mã có thể dẫn ra cùng kết quả dự đoán, độ tin cậy của việc lập mã hình ảnh có thể được tăng lên và lượng dữ liệu cần được truyền cho việc lập mã hình ảnh có thể được giảm xuống.

Như được mô tả trên đây, thủ tục xây dựng lại hình ảnh có thể được thực hiện không chỉ trong thiết bị giải mã ảnh mà cả trong thiết bị mã hóa ảnh. Khối được xây dựng lại có thể được tạo ra dựa trên việc dự đoán nội/dự đoán liên theo các đơn vị là các khối, và hình ảnh được xây dựng lại bao gồm các khối được xây dựng lại có thể được tạo ra. Khi hình ảnh/lát/nhóm ô xếp hiện tại là hình ảnh/lát/nhóm ô xếp I, các khối được chứa trong hình ảnh/lát/nhóm ô xếp hiện tại này có thể được xây dựng lại chỉ dựa trên việc dự đoán nội. Mặt khác, khi hình ảnh/lát/nhóm ô xếp hiện tại là hình ảnh/lát/nhóm ô xếp P hoặc B, các khối được chứa trong hình ảnh/lát/nhóm ô xếp hiện tại này có thể được xây dựng lại dựa trên việc dự đoán nội hoặc việc dự đoán liên. Trong trường hợp này, việc dự đoán liên có thể được áp dụng cho một số khối trong hình ảnh/lát/nhóm ô xếp hiện tại này và việc dự đoán nội có thể được áp dụng cho các khối còn lại. Thành

phần màu của hình ảnh có thể bao gồm thành phần độ chói và thành phần sắc độ và các phương pháp và các phương án của sáng chế là có thể áp dụng được với cả thành phần độ chói lẫn thành phần sắc độ trừ khi bị giới hạn một cách tường minh trong sáng chế này.

Ví dụ về cấu trúc lớp lập mã

Video/ảnh được lập mã của sáng chế có thể được xử lý, ví dụ, theo lớp và cấu trúc lập mã sẽ được mô tả sau.

Fig.6 là hình minh họa cấu trúc lớp cho ảnh được lập mã.

Ảnh được lập mã được phân loại thành lớp lập mã video (VCL) cho quy trình giải mã ảnh và để xử lý bản thân nó, hệ thống bên dưới để truyền và lưu trữ thông tin được mã hóa, và lớp trùu tượng mạng (NAL) có mặt giữa VCL và hệ thống bên dưới này và chịu trách nhiệm về chức năng thích ứng mạng.

Trong VCL, dữ liệu VCL bao gồm dữ liệu ảnh (dữ liệu lát) được nén có thể được tạo ra hoặc tập tham số bao gồm thông tin, chẳng hạn như tập tham số hình ảnh (PPS), tập tham số trình tự (SPS) hoặc tập tham số video (VPS), hoặc tin nhắn thông tin tăng cường bổ sung (supplemental enhancement information, SEI) được yêu cầu thêm cho quy trình giải mã ảnh có thể được tạo ra.

Trong NAL, thông tin tiêu đề (tiêu đề đơn vị NAL) có thể được thêm vào phần tải hữu ích trình tự byte thô (raw byte sequence payload, RBSP) được tạo ra trong VCL để tạo ra đơn vị NAL. Trong trường hợp này, RBSP chỉ dữ liệu lát, tập tham số, tin nhắn SEI được tạo ra trong VCL. Tiêu đề đơn vị NAL có thể bao gồm thông tin loại đơn vị NAL được chỉ rõ tùy theo dữ liệu RBSP được chứa trong đơn vị NAL tương ứng.

Như được thể hiện trên Fig.6, đơn vị NAL có thể được phân loại thành đơn vị NAL VCL và đơn vị NAL phi VCL (non-VCL NAL unit) tùy theo loại RBSP được tạo ra trong VCL này. Đơn vị NAL VCL có thể nghĩa là đơn vị NAL chứa thông tin về ảnh (dữ liệu lát), và đơn vị NAL phi VCL có thể nghĩa là đơn vị NAL chứa thông tin (tập tham số hoặc tin nhắn SEI) cần thiết để giải mã ảnh.

Đơn vị NAL VCL và đơn vị NAL phi VCL có thể được gắn với thông tin tiêu đề và được truyền qua mạng theo định dạng dữ liệu của hệ thống bên dưới. Ví dụ, đơn vị NAL có thể được cải biến thành dữ liệu có định dạng dữ liệu được xác định trước, chẳng hạn như định dạng tệp tin H.266/VVC, RTP (giao thức truyền thời gian thực, Real-time

Transport Protocol), hoặc TS (luồng truyền, Transport Stream), và được truyền qua các mạng khác nhau.

Như được mô tả trên đây, trong đơn vị NAL, loại đơn vị NAL có thể được chỉ rõ theo cấu trúc dữ liệu RBSP được chứa trong đơn vị NAL tương ứng, và thông tin về loại đơn vị NAL có thể được lưu trữ trong tiêu đề đơn vị NAL và được báo hiệu. Ví dụ, việc này có thể được phân loại khái quát thành loại đơn vị NAL VCL và loại đơn vị NAL phi VCL dựa trên việc liệu đơn vị NAL này có chứa thông tin về ảnh (dữ liệu lát) hay không. Loại đơn vị NAL VCL có thể được phân loại theo đặc tính và loại của hình ảnh, và loại đơn vị NAL phi VCL có thể được phân loại theo loại của tập tham số.

Ví dụ về loại đơn vị NAL được chỉ rõ theo loại của tập tham số/thông tin được chứa trong loại đơn vị NAL phi VCL sẽ được liệt kê dưới đây.

- Loại đơn vị NAL (NAL unit type, NUT) DCI (Decoding capability information, thông tin khả năng giải mã): loại dành cho đơn vị NAL chứa DCI
- NUT VPS (tập tham số video): loại dành cho đơn vị NAL chứa VPS
- NUT SPS (tập tham số trình tự): loại dành cho đơn vị NAL chứa SPS
- NUT PPS (tập tham số hình ảnh): loại dành cho đơn vị NAL chứa PPS
- NUT APS (tập tham số thích ứng): loại dành cho đơn vị NAL chứa APS
- NUT PH (Picture header, tiêu đề hình ảnh): loại dành cho đơn vị NAL chứa PH

Các loại đơn vị NAL được mô tả trên đây có thể có thông tin cú pháp về các loại đơn vị NAL, và thông tin cú pháp này có thể được lưu trữ trong tiêu đề đơn vị NAL và được báo hiệu. Ví dụ, thông tin cú pháp này có thể bao gồm `nal_unit_type`, và các loại đơn vị NAL có thể được chỉ rõ bởi giá trị của `nal_unit_type`.

Trong khi đó, như được mô tả trên đây, một hình ảnh có thể bao gồm nhiều lát, và một lát có thể bao gồm tiêu đề lát và dữ liệu lát. Trong trường hợp này, một tiêu đề hình ảnh có thể còn được thêm vào nhiều lát (tiêu đề lát và tập dữ liệu lát) trong một hình ảnh. Tiêu đề hình ảnh (cú pháp tiêu đề hình ảnh) có thể bao gồm thông tin/các tham số áp dụng được chung cho hình ảnh này. Tiêu đề lát (cú pháp tiêu đề lát) có thể bao gồm thông tin/các tham số áp dụng được chung cho lát này. APS (cú pháp APS) hoặc PPS (cú pháp PPS) có thể bao gồm thông tin/các tham số áp dụng được chung cho một hoặc nhiều lát hoặc hình ảnh. SPS (cú pháp SPS) có thể bao gồm thông tin/các tham số áp dụng được chung cho một hoặc nhiều trình tự. VPS (cú pháp VPS) có thể bao gồm thông

tin/các tham số áp dụng được chung cho nhiều lớp. DCI (cú pháp DCI) có thể bao gồm thông tin/các tham số liên quan đến khả năng giải mã.

Trong sáng chế này, cú pháp mức cao (high level syntax, HLS) có thể bao gồm ít nhất một cú pháp trong số cú pháp APS, cú pháp PPS, cú pháp SPS, cú pháp VPS, cú pháp DCI, cú pháp tiêu đề hình ảnh hoặc cú pháp tiêu đề lát. Thêm vào đó, trong sáng chế này, cú pháp mức thấp (low level syntax, LLS), ví dụ, có thể bao gồm cú pháp dữ liệu lát, cú pháp CTU, cú pháp đơn vị lập mã, cú pháp đơn vị biến đổi, v.v.

Trong sáng chế này, thông tin ảnh/video được mã hóa bởi thiết bị mã hóa và được báo hiệu đến thiết bị giải mã dưới dạng luồng bit có thể bao gồm không chỉ thông tin liên quan đến việc phân vùng trong hình ảnh, thông tin dự đoán nội/liên, thông tin phần dư, thông tin việc lọc trong vòng, mà cả thông tin về tiêu đề lát, thông tin về tiêu đề hình ảnh, thông tin về APS, thông tin về PPS, thông tin về SPS, thông tin về VPS và/hoặc thông tin về DCI. Thêm vào đó, thông tin ảnh/video có thể còn bao gồm thông tin ràng buộc chung và/hoặc thông tin về tiêu đề đơn vị NAL.

Tổng quan về việc báo hiệu các điểm nhập vào (entry point)

Như được mô tả trên đây, đơn vị NAL VCL có thể bao gồm dữ liệu lát dưới dạng RBSP (phần tải hữu ích trình tự byte thô). Dữ liệu lát này có thể được canh chỉnh trong đơn vị NAL VCL tính bằng byte, và có thể bao gồm một hoặc nhiều tập con. Ít nhất một điểm nhập vào cho việc truy cập ngẫu nhiên (random access, RA) có thể được định ra cho các tập con này, và việc xử lý song song có thể được thực hiện dựa trên điểm nhập vào này.

Tiêu chuẩn lập mã video vạn năng (versatile video coding, VVC) (ITU-T Rec. H.266) hỗ trợ việc xử lý song song đầu sóng (wavefront parallel processing, WPP) mà nó là một trong số nhiều kỹ thuật xử lý song song khác nhau. Nhiều lát trong hình ảnh có thể được lập mã/giải mã song song dựa trên WPP.

Để kích hoạt năng lực xử lý song song, thông tin điểm nhập vào có thể được báo hiệu. Thiết bị giải mã ảnh có thể truy cập trực tiếp điểm bắt đầu của đoạn dữ liệu được chứa trong đơn vị NAL dựa trên thông tin điểm nhập vào. Ở đây, điểm bắt đầu của đoạn dữ liệu có thể nghĩa là điểm bắt đầu của các ô xếp trong lát hoặc điểm bắt đầu của các hàng CTU trong lát.

Thông tin điểm nhập vào có thể được báo hiệu trong cú pháp mức cao, ví dụ, tập

tham số hình ảnh (PPS) và/hoặc tiêu đề lát.

Fig.7 là hình minh họa tập tham số hình ảnh (PPS) theo một phương án của sáng chế, và Fig.8 là hình minh họa tiêu đề lát theo một phương án của sáng chế.

Trước hết, tham chiếu đến Fig.7, tập tham số hình ảnh (PPS) có thể bao gồm `entry_point_offsets_present_flag` làm phần tử cú pháp chỉ ra liệu thông tin điểm nhập vào có được báo hiệu hay không.

`entry_point_offsets_present_flag` có thể chỉ ra liệu việc báo hiệu thông tin điểm nhập vào có có mặt trong tiêu đề lát đang tham chiếu đến tập tham số hình ảnh (PPS) hay không. Ví dụ, `entry_point_offsets_present_flag` có giá trị thứ nhất (ví dụ, 0) có thể chỉ ra rằng việc báo hiệu thông tin điểm nhập vào cho ô xếp hoặc các hàng CTU cụ thể trong ô xếp là không có mặt trong tiêu đề lát. Ngược lại, `entry_point_offsets_present_flag` có giá trị thứ hai (ví dụ, 1) có thể chỉ ra rằng việc báo hiệu thông tin điểm nhập vào cho ô xếp hoặc các hàng CTU cụ thể trong ô xếp là có mặt trong tiêu đề lát.

Trong khi đó, mặc dù trường hợp trong đó `entry_point_offsets_present_flag` được chứa trong tập tham số hình ảnh (PPS) được minh họa trên Fig.7, đây chỉ là ví dụ và do đó các phương án của sáng chế không bị giới hạn vào đó. Ví dụ, `entry_point_offsets_present_flag` có thể được chứa trong tập tham số trình tự (SPS).

Tiếp theo, tham chiếu đến Fig.8, tiêu đề lát có thể bao gồm `offset_len_minus1` và `entry_point_offset_minus1[ i ]` làm các phần tử cú pháp để nhận dạng điểm nhập vào.

`offset_len_minus1` có thể chỉ ra giá trị thu được bằng cách trừ 1 khỏi chiều dài bit của `entry_point_offset_minus1[ i ]`. Giá trị của `offset_len_minus1` có thể có khoảng từ 0 đến 31. Trong một ví dụ, `offset_len_minus1` có thể được báo hiệu dựa trên biến số `NumEntryPoints` chỉ ra tổng số lượng điểm nhập vào. Ví dụ, `offset_len_minus1` có thể được báo hiệu chỉ khi `NumEntryPoints` lớn hơn 0. Thêm vào đó, `offset_len_minus1` có thể được báo hiệu dựa trên `entry_point_offsets_present_flag` được mô tả trên đây với tham chiếu đến Fig.7. Ví dụ, `offset_len_minus1` có thể được báo hiệu chỉ khi `entry_point_offsets_present_flag` có giá trị thứ hai (ví dụ, 1) (tức là, việc báo hiệu thông tin điểm nhập vào là có mặt trong tiêu đề lát).

`entry_point_offset_minus1[ i ]` có thể chỉ ra phần bù (offset) của điểm nhập vào thứ `i` tính theo byte, và có thể được thể hiện bằng cách thêm các bit 1 vào

offset_len_minus1. Dữ liệu lát trong đơn vị NAL có thể bao gồm cùng số lượng tập con với giá trị thu được bằng cách cộng 1 vào NumEntryPoints, và giá trị chỉ số chỉ rõ mỗi tập con trong số các tập con này có thể có khoảng từ 0 đến NumEntryPoints. Byte thứ nhất của dữ liệu lát trong đơn vị NAL có thể được thể hiện bởi byte 0.

Khi entry_point_offset_minus1[i] được báo hiệu, các byte ngăn chặn cạnh tranh (emulation prevention) được chứa trong dữ liệu lát trong đơn vị NAL có thể được đếm như một phần của dữ liệu lát để nhận dạng các tập con. Tập con 0 mà nó là tập con thứ nhất của dữ liệu lát có thể có cấu hình từ byte 0 đến entry_point_offset_minus1[0]. Tương tự, tập con k mà nó là tập con thứ k của dữ liệu lát có thể có cấu hình từ byte firstByte[k] đến lastByte[k]. Ở đây, firstByte[k] có thể được dẫn ra như được thể hiện trên phương trình 1 dưới đây, lastByte[k] có thể được dẫn ra như được thể hiện trên phương trình 2 dưới đây.

Phương trình 1

$$\text{firstByte}[k] = \sum_{n=1}^k (\text{sh_entry_point_offset_minus1}[n - 1] + 1)$$

Phương trình 2

$$\text{lastByte}[k] = \text{firstByte}[k] + \text{sh_entry_point_offset_minus1}[k]$$

Trong phương trình 1 và phương trình 2, k có thể có khoảng từ 1 đến giá trị được thu bằng cách trừ 1 khỏi NumEntryPoints.

Tập con cuối cùng (tức là, tập con thứ NumEntryPoints) của dữ liệu lát có thể gồm các byte còn lại của dữ liệu lát.

Trong khi đó, nếu quy trình đồng bộ được xác định trước cho các biến số ngữ cảnh không được thực hiện trước việc giải mã CTU đang bao gồm CTB thứ nhất của hàng CTB trong mỗi ô xếp (ví dụ, sps_entropy_coding_sync_enabled_flag == 0) và lát bao gồm một hoặc nhiều ô xếp hoàn chỉnh thì mỗi tập con của tiêu đề lát có thể gồm tất cả các bit được lập mã cho tất cả các CTU trong cùng ô xếp. Trong trường hợp này, tổng số lượng tập con của dữ liệu lát có thể bằng với tổng số lượng ô xếp trong lát.

Ngược lại, khi quy trình đồng bộ được xác định trước này không được thực hiện và lát chứa một tập con cho các hàng CTU trong ô xếp đơn lẻ, NumEntryPoints có thể là 0. Trong trường hợp này, một tập con của dữ liệu lát có thể gồm tất cả các bit được lập mã cho tất cả các CTU trong lát.

Ngược lại, khi quy trình đồng bộ được xác định trước này được thực hiện (ví dụ, `sps_entropy_coding_sync_enabled_flag == 1`), mỗi tập con có thể gồm tất cả các bit được lập mã cho tất cả các CTU của một hàng CTU trong một ô xếp. Trong trường hợp này, tổng số lượng tập con của dữ liệu lát có thể bằng với tổng số lượng hàng CTU của mỗi ô xếp trong lát.

Tổng quan về loại đơn vị NAL được trộn

Nói chung, một loại đơn vị NAL có thể được thiết lập cho một hình ảnh. Như được mô tả trên đây, thông tin cú pháp chỉ rõ loại đơn vị NAL có thể được lưu trữ trong tiêu đề đơn vị NAL của đơn vị NAL và được báo hiệu. Ví dụ, thông tin cú pháp này có thể bao gồm `nal_unit_type`, và các loại đơn vị NAL có thể được chỉ rõ bởi giá trị của `nal_unit_type`.

Ví dụ về các loại đơn vị NAL, mà các phương án của sáng chế là có thể áp dụng được với chúng, được thể hiện trong bảng 1 dưới đây.

Bảng 1

nal_unit_type	Tên của nal_unit_type	Nội dung của đơn vị NAL và cấu trúc cú pháp RBSP	Đơn vị NAL Phân loại của loại
0	TRAIL_NUT	Lát được lập mã của hình ảnh kéo theo <code>slice_layer_rbsp()</code>	VCL
1	STSA_NUT	Lát được lập mã của hình ảnh STSA <code>slice_layer_rbsp()</code>	VCL
2	RADL_NUT	Lát được lập mã của hình ảnh RADL <code>slice_layer_rbsp()</code>	VCL
3	RASL_NUT	Lát được lập mã của hình ảnh RASL <code>slice_layer_rbsp()</code>	VCL
4..6	RSV_VCL_4.. RSV_VCL_6	Các loại đơn vị NAL VCL phi-IRAP được dự trữ	VCL
7 8	IDR_W_RADL IDR_N_LP	Lát được lập mã của hình ảnh IDR <code>slice_layer_rbsp()</code>	VCL
9	CRA_NUT	Lát được lập mã của hình ảnh CRA <code>silce_layer_rbsp()</code>	VCL
10	GDR_NUT	Lát được lập mã của hình ảnh GDR <code>slice_layer_rbsp()</code>	VCL
11 12	RSV_IRAP_11 RSV_IRAP_12	Các loại đơn vị NAL VCL IRAP được dự trữ	VCL
13	DCI_NUT	Thông tin khả năng giải mã <code>decoding_capability_information_rbsp()</code>	phi-VCL

14	VPS_NUT	Tập tham số video video_parameter_set_rbsp()	phi-VCL
15	SPS_NUT	Tập tham số trình tự seq_parameter_set_rbsp()	phi-VCL
16	PPS_NUT	Tập tham số hình ảnh pic_parameter_set_rbsp()	phi-VCL
17 18	PREFIX_APS_ NUT SUFFIX_APS_ NUT	Tập tham số thích ứng adaptation_parameter_set_rbsp()	phi-VCL
19	PH_NUT	Tiêu đề hình ảnh picture_header_rbsp()	phi-VCL
20	AUD_NUT	Dấu phân cách AU access_unit_delimiter_rbsp()	phi-VCL
21	EOS_NUT	Kết thúc trình tự end_of_seq_rbsp()	phi-VCL
22	EOB_NUT	Kết thúc luồng bit end_of_bitstream_rbsp()	phi-VCL
23 24	PREFIX_SEI_ NUT SUFFIX_SEI_ NUT	Thông tin tăng cường bổ sung sei_rbsp()	phi-VCL
25	FD_NUT	Dữ liệu điền đầy filler_data_rbsp()	phi-VCL
26 27	RSV_NVCL_2 6 RSV_NVCL_2 7	Các loại đơn vị NAL phi-VCL được dự trữ	phi-VCL
28..31	UNSPEC_28.. UNSPEC_31	Các loại đơn vị NAL phi-VCL không được chỉ rõ	phi-VCL

NUT: loại đơn vị NAL (NAL unit type)

STSA: truy cập chuyển đổi lớp con theo thời gian từng bước (Step-wise Temporal sub-layer Switching Access)

RADL: dẫn có thể giải mã được truy cập ngẫu nhiên (Random Access Decodable Leading)

RASL: dẫn được bỏ qua truy cập ngẫu nhiên (Random Access Skipped Leading)

IDR: làm tươi giải mã tức thời (Instantaneous Decoding Refresh)

LP: hình ảnh dẫn (Leading picture)

_W_RADL: cùng với RADL (With RADL)

_N_LP: không LP (No LP), không có LP

CRA: truy cập ngẫu nhiên sạch (Clean Random Access)

GDR: làm tươi giải mã dần dần (Gradual Decoding Refresh)

- IRAP: điểm truy nhập ngẫu nhiên nội (Intra Random Access Point)

Tham chiếu đến bảng 1, loại đơn vị NAL VCL có thể được phân loại thành các loại đơn vị NAL 0 đến 12 tùy theo thuộc tính và loại hình ảnh. Thêm vào đó, loại đơn

vị NAL phi VCL có thể được phân loại thành các loại đơn vị NAL 13 đến 31 tùy theo loại của tập tham số.

Các ví dụ chi tiết về loại đơn vị NAL VCL là như sau.

- Loại đơn vị NAL (NUT) IRAP (điểm truy cập ngẫu nhiên nội): loại dành cho đơn vị NAL của hình ảnh IRAP, mà nó được thiết lập vào khoảng từ IDR_W_RADL đến CRA_NUT.

- NUT IDR (làm tươi giải mã tức thời): loại dành cho đơn vị NAL của hình ảnh IDR, mà nó được thiết lập thành IDR_W_RADL hoặc IDR_N_LP.

- NUT CRA (truy cập ngẫu nhiên sạch): loại dành cho đơn vị NAL dành cho hình ảnh CRA, mà nó được thiết lập thành CRA_NUT.

- NUT RADL (dẫn có thể giải mã được truy cập ngẫu nhiên): loại dành cho đơn vị NAL dành cho hình ảnh RADL, mà nó được thiết lập thành RADL_NUT.

- NUT RASL (dẫn được bỏ qua truy cập ngẫu nhiên): loại dành cho đơn vị NAL dành cho hình ảnh RASL, mà nó được thiết lập thành RASL_NUT.

- NUT kéo theo (Trailing NUT): loại dành cho đơn vị NAL dành cho hình ảnh kéo theo, mà nó được thiết lập thành TRAIL_NUT.

- NUT GDR (làm tươi giải mã dần dần): loại dành cho đơn vị NAL dành cho hình ảnh GDR, mà nó được thiết lập thành GDR_NUT.

- NUT STSA (truy cập lớp con theo thời gian từng bước): loại dành cho đơn vị NAL dành cho hình ảnh STSA, mà nó được thiết lập thành STSA_NUT.

Trong khi đó, tiêu chuẩn VVC cho phép một hình ảnh bao gồm nhiều lát có các loại đơn vị NAL khác nhau. Ví dụ, một hình ảnh có thể bao gồm ít nhất một lát thứ nhất có loại đơn vị NAL thứ nhất và ít nhất một lát thứ hai có loại đơn vị NAL thứ hai khác với loại đơn vị NAL thứ nhất. Trong trường hợp này, loại đơn vị NAL của hình ảnh này có thể được gọi là loại đơn vị NAL được trộn. Vì tiêu chuẩn VVC hỗ trợ loại đơn vị NAL được trộn, nhiều hình ảnh có thể được xây dựng lại/tổng hợp một cách dễ dàng hơn trong quy trình tổng hợp nội dung, quy trình mã hóa/giải mã, v.v.

Tuy nhiên, theo sơ đồ hiện có đối với loại đơn vị NAL được trộn, tồn tại giới hạn đó là chỉ hai loại đơn vị NAL được phép được trộn với nhau. Thêm vào đó, khi một hình ảnh có loại đơn vị NAL được trộn, tồn tại giới hạn đó là một hoặc nhiều đơn vị NAL VCL của hình ảnh này sẽ phải có loại đơn vị NAL trong khoảng từ IDR_W_RADL đến

CRA_NUT, và các đơn vị NAL VCL còn lại của hình ảnh này sẽ phải có loại đơn vị NAL trong khoảng từ TRAIL_NUT đến RSV_VCL_6. Kết quả là, mặc dù loại đơn vị NAL được trộn rất hữu ích trong thủ tục xử lý ảnh, tồn tại vấn đề đó là loại đơn vị NAL được trộn thường không được sử dụng.

Để giải quyết vấn đề như vậy, theo các phương án của sáng chế, hai hay hơn hai loại đơn vị NAL có thể được cho phép được trộn và thêm các loại đơn vị NAL được trộn khác nhau có thể được đề xuất.

Ở dưới đây, các phương án của sáng chế sẽ được mô tả chi tiết.

Khi một hình ảnh bao gồm hai hay hơn hai lát và các loại đơn vị NAL của các lát này là khác nhau, hình ảnh này có thể được giới hạn để bao gồm hai hay hơn hai hình ảnh con. Tức là, khi một hình ảnh có loại đơn vị NAL được trộn, hình ảnh này có thể bao gồm hai hay hơn hai hình ảnh con.

Hình ảnh con có thể bao gồm một hoặc nhiều lát và xây dựng nên vùng hình chữ nhật trong hình ảnh. Mỗi hình ảnh con được chứa trong hình ảnh có thể có kích thước khác nhau. Mặt khác, đối với tất cả các hình ảnh thuộc về một trình tự, các kích thước và các vị trí của các hình ảnh con riêng rẽ cụ thể có thể được đặt cho bằng nhau.

Fig.9 là hình minh họa ví dụ về hình ảnh con.

Tham chiếu đến Fig.9, một hình ảnh có thể được phân vùng thành tổng cộng 18 ô xếp. 12 ô xếp trong số 18 ô xếp này có thể được sắp xếp trên phía bên trái của hình ảnh và mỗi ô xếp trong số 12 ô xếp này có thể bao gồm một lát bao gồm các CTU 4x4. Thêm vào đó, 6 ô xếp có thể được sắp xếp trên phía bên phải của hình ảnh và mỗi ô xếp trong số 6 ô xếp này có thể bao gồm hai lát mà mỗi lát bao gồm các CTU 2x2 được xếp chồng theo hướng dọc. Kết quả là, hình ảnh này có thể bao gồm 24 hình ảnh con và 24 lát và mỗi hình ảnh con có thể bao gồm một lát.

Trong một phương án, mỗi hình ảnh con trong một hình ảnh có thể được xử lý như một hình ảnh để hỗ trợ loại đơn vị NAL được trộn. Khi hình ảnh con được xử lý như một hình ảnh, hình ảnh con này có thể được lập mã/giải mã một cách độc lập bất kể kết quả của việc mã hóa/giải mã hình ảnh con khác. Ở đây, việc mã hóa/giải mã độc lập có thể nghĩa là cấu trúc phân vùng khối (ví dụ, cấu trúc cây đơn, cấu trúc cây kép, v.v.), loại chế độ dự đoán (ví dụ, việc dự đoán nội, việc dự đoán liên, v.v.), thứ tự giải mã, v.v. của hình ảnh con là khác với cấu trúc phân vùng khối, loại chế độ dự đoán, thứ tự giải

mã, v.v. của hình ảnh con khác. Ví dụ, khi hình ảnh con thứ nhất được lập mã/giải mã dựa trên chế độ dự đoán nội, hình ảnh con thứ hai, liền kề với hình ảnh con thứ nhất này và được xử lý như một hình ảnh, có thể được lập mã/giải mã dựa trên chế độ dự đoán liên.

Khi một hình ảnh bao gồm hai hay hơn hai hình ảnh con độc lập và các hình ảnh con này có các loại đơn vị NAL khác nhau, hình ảnh này có thể có loại đơn vị NAL được trộn.

Fig.10 là hình minh họa ví dụ về hình ảnh có loại đơn vị NAL được trộn.

Tham chiếu đến Fig.10, một hình ảnh 1000 có thể bao gồm các hình ảnh con thứ nhất đến thứ ba 1010 đến 1030. Mỗi hình ảnh con trong số các hình ảnh con thứ nhất và thứ ba 1010 và 1030 có thể bao gồm hai lát. Ngược lại, hình ảnh con thứ hai 1020 có thể bao gồm bốn lát.

Khi mỗi hình ảnh con trong số các hình ảnh con thứ nhất đến thứ ba 1010 đến 1030 được xử lý như một hình ảnh, các hình ảnh con thứ nhất đến thứ ba 1010 đến 1030 có thể được lập mã một cách độc lập để xây dựng nên các luồng bit khác nhau. Ví dụ, dữ liệu lát được lập mã của hình ảnh con thứ nhất 1010 có thể được đóng gói thành một hoặc nhiều đơn vị NAL có cùng loại đơn vị NAL là RASL_NUT để xây dựng nên luồng bit thứ nhất (Luồng bit 1, Bitstream 1). Thêm vào đó, dữ liệu lát được lập mã của hình ảnh con thứ hai 1020 có thể được đóng gói thành một hoặc nhiều đơn vị NAL có cùng loại đơn vị NAL là RADL_NUT để xây dựng nên luồng bit thứ hai (Luồng bit 2, Bitstream 2). Thêm vào đó, dữ liệu lát được lập mã của hình ảnh con thứ ba 1030 có thể được đóng gói thành một hoặc nhiều đơn vị NAL có cùng loại đơn vị NAL là RASL_NUT để xây dựng nên luồng bit thứ ba (Luồng bit 3, Bitstream 3). Kết quả là, một hình ảnh 1000 có thể có loại đơn vị NAL được trộn mà RASL_NUT và RADL_NUT được trộn trong đó.

Trong một phương án, tất cả các lát được chứa trong mỗi hình ảnh con trong hình ảnh có thể được giới hạn để có cùng loại đơn vị NAL. Ví dụ, cả hai lát được chứa trong hình ảnh con thứ nhất 1010 đều có thể có cùng loại đơn vị NAL là RASL_NUT. Thêm vào đó, tất cả bốn lát được chứa trong hình ảnh con thứ hai 1020 đều có thể có cùng loại đơn vị NAL là RADL_NUT. Thêm vào đó, cả hai lát được chứa trong hình ảnh con thứ ba 1030 đều có thể có cùng loại đơn vị NAL là RASL_NUT.

Thông tin về hình ảnh con có thể được báo hiệu trong cú pháp mức cao, ví dụ, tập tham số hình ảnh (PPS) và tập tham số trình tự (SPS). Thêm vào đó, thông tin chỉ rõ liệu có áp dụng loại đơn vị NAL được trộn hay không có thể được báo hiệu trong cú pháp mức cao, ví dụ, tập tham số hình ảnh (PPS).

Fig.11 là hình minh họa ví dụ về tập tham số hình ảnh (PPS) theo một phương án của sáng chế, và Fig.12 là hình minh họa ví dụ về tập tham số trình tự (SPS) theo một phương án của sáng chế.

Trước hết, tham chiếu đến Fig.11, tập tham số hình ảnh (PPS) có thể bao gồm `pps_mixed_nalu_types_in_pic_flag` làm phần tử cú pháp chỉ ra liệu có áp dụng loại đơn vị NAL được trộn hay không.

`pps_mixed_nalu_types_in_pic_flag` có thể chỉ ra liệu hình ảnh hiện tại có loại đơn vị NAL được trộn hay không. Ví dụ, `pps_mixed_nalu_types_in_pic_flag` có giá trị thứ nhất (ví dụ, 0) có thể chỉ ra rằng hình ảnh hiện tại không có loại đơn vị NAL được trộn. Trong trường hợp này, hình ảnh hiện tại có thể có cùng loại đơn vị NAL đối với tất cả các đơn vị NAL VCL, ví dụ, cùng loại đơn vị NAL với các đơn vị NAL của lát được lập mã. Ngược lại, `pps_mixed_nalu_types_in_pic_flag` có giá trị thứ hai (ví dụ, 1) có thể chỉ ra rằng hình ảnh hiện tại có loại đơn vị NAL được trộn.

Trong một phương án, đơn vị NAL VCL của hình ảnh hiện tại có thể được giới hạn để không có cùng loại đơn vị NAL như là GDR_NUT, khi hình ảnh hiện tại có loại đơn vị NAL được trộn (ví dụ, `pps_mixed_nalu_types_in_pic_flag == 1`).

Trong một phương án, khi hình ảnh hiện tại có loại đơn vị NAL được trộn (ví dụ, `pps_mixed_nalu_types_in_pic_flag == 1`), nếu một đơn vị NAL VCL của hình ảnh hiện tại có cùng loại đơn vị NAL (loại đơn vị NAL A) như là IDR_W_RADL, IDR_N_LP hoặc CRA_NUT thì các đơn vị NAL VCL khác của hình ảnh hiện tại có thể được giới hạn để có cùng loại đơn vị NAL như là IDR_W_RADL, IDR_N_LP, CRA_NUT hoặc TRAIL_NUT.

Khi hình ảnh hiện tại có loại đơn vị NAL được trộn (ví dụ, `pps_mixed_nalu_types_in_pic_flag == 1`), mỗi hình ảnh con trong hình ảnh hiện tại đều có thể có một loại bất kỳ trong số các loại đơn vị NAL VCL được mô tả trên đây với tham chiếu đến bảng 1. Ví dụ, khi hình ảnh con trong hình ảnh hiện tại là hình ảnh con IDR, hình ảnh con này có thể có cùng loại đơn vị NAL như là IDR_W_RADL hoặc

IDR_N_LP. Theo cách khác, khi hình ảnh con trong hình ảnh hiện tại là hình ảnh con kéo theo, hình ảnh con này có thể có cùng loại đơn vị NAL như là TRAIL_NUT.

pps_mixed_nalu_types_in_pic_flag có giá trị thứ hai (ví dụ, 1) có thể chỉ ra rằng các hình ảnh đang tham chiếu đến tập tham số hình ảnh (PPS) có thể bao gồm các lát có các loại đơn vị NAL khác nhau. Ở đây, các hình ảnh này có thể được bắt nguồn từ hoạt động hợp nhất luồng bit hình ảnh con trong đó bộ mã hóa đảm bảo việc so khớp của cấu trúc luồng bit và việc canh chỉnh giữa các tham số của các luồng bit gốc. Để làm ví dụ cho việc canh chỉnh này, khi các phần tử cú pháp danh sách hình ảnh tham chiếu (reference picture list, RPL) cho lát, đang có cùng loại đơn vị NAL như là IDR_W_RADL hoặc IDR_N_LP, không có mặt trong tiêu đề lát (ví dụ, sps_idr_rpl_present_flag == 0) và hình ảnh hiện tại đang chứa lát này có loại đơn vị NAL được trộn (ví dụ, pps_mixed_nalu_types_in_pic_flag == 1), hình ảnh hiện tại có thể được giới hạn để không bao gồm các lát đang có cùng loại đơn vị NAL như là IDR_W_RADL hoặc IDR_N_LP.

Trong khi đó, khi tồn tại giới hạn đó là loại đơn vị NAL được trộn không được áp dụng cho tất cả các hình ảnh trong tập lớp đầu ra (output layer set, OLS) (ví dụ, gci_no_mixed_nalu_types_in_pic_constraint_flag == 1), pps_mixed_nalu_types_in_pic_flag có thể có giá trị thứ nhất (ví dụ, 0).

Thêm vào đó, tập tham số hình ảnh (PPS) có thể bao gồm pps_no_pic_partition_flag làm phần tử cú pháp chỉ ra liệu có áp dụng việc phân vùng hình ảnh hay không.

pps_no_pic_partition_flag có thể chỉ ra liệu việc phân vùng hình ảnh có áp dụng được cho hình ảnh hiện tại hay không. Ví dụ, pps_no_pic_partition_flag có giá trị thứ nhất (ví dụ, 0) có thể chỉ ra rằng hình ảnh hiện tại không thể được phân vùng. Ngược lại, pps_no_pic_partition_flag có giá trị thứ hai (ví dụ, 1) có thể chỉ ra rằng hình ảnh hiện tại có thể được phân vùng thành hai hay hơn hai ô xếp hoặc lát. Khi hình ảnh hiện tại có loại đơn vị NAL được trộn (ví dụ, pps_mixed_nalu_types_in_pic_flag == 1), hình ảnh hiện tại này có thể được giới hạn để có hai hay hơn hai ô xếp hoặc lát (ví dụ, pps_no_pic_partition_flag = 1).

Thêm vào đó, tập tham số hình ảnh (PPS) có thể bao gồm pps_num_subpics_minus1 làm phần tử cú pháp chỉ ra số lượng hình ảnh con.

`pps_num_subpics_minus1` có thể chỉ ra giá trị thu được bằng cách trừ 1 khỏi số lượng hình ảnh con được chứa trong hình ảnh hiện tại. `pps_num_subpics_minus1` có thể được báo hiệu chỉ khi việc phân vùng hình ảnh là áp dụng được cho hình ảnh hiện tại (ví dụ, `pps_no_pic_partition_flag == 1`). Khi `pps_num_subpics_minus1` không được báo hiệu, giá trị của `pps_num_subpics_minus1` có thể được suy ra là 0. Trong khi đó, phần tử cú pháp chỉ ra số lượng hình ảnh con có thể được báo hiệu trong cú pháp mức cao khác với tập tham số hình ảnh (PPS), ví dụ, tập tham số trình tự (SPS).

Trong một phương án, khi hình ảnh hiện tại bao gồm chỉ một hình ảnh con (ví dụ, `pps_num_subpics_minus1 == 0`), hình ảnh hiện tại này có thể được giới hạn để không có loại đơn vị NAL được trộn (ví dụ, `pps_mixed_nalu_types_in_pic_flag == 0`). Tức là, khi hình ảnh hiện tại có loại đơn vị NAL được trộn (ví dụ, `pps_mixed_nalu_types_in_pic_flag == 1`), hình ảnh hiện tại này có thể được giới hạn để bao gồm hai hay hơn hai hình ảnh con (ví dụ, `pps_num_subpics_minus1 > 0`).

Tiếp theo, tham chiếu đến Fig.12, tập tham số trình tự (SPS) có thể bao gồm `sps_subpic_treated_as_pic_flag[ i ]` làm phần tử cú pháp liên quan đến việc xử lý hình ảnh con trong quá trình mã hóa/giải mã.

`sps_subpic_treated_as_pic_flag[ i ]` có thể chỉ ra liệu mỗi hình ảnh con trong hình ảnh hiện tại có được xử lý như một hình ảnh hay không. Ví dụ, `sps_subpic_treated_as_pic_flag[ i ]` có giá trị thứ nhất (ví dụ, 0) có thể chỉ ra rằng hình ảnh con thứ *i* trong hình ảnh hiện tại không được xử lý như một hình ảnh. Ngược lại, `sps_subpic_treated_as_pic_flag[ i ]` có giá trị thứ hai (ví dụ, 1) có thể chỉ ra rằng hình ảnh con thứ *i* trong hình ảnh hiện tại được xử lý như một hình ảnh trong quy trình mã hóa/giải mã ngoại trừ các hoạt động lọc trong vòng. Khi `sps_subpic_treated_as_pic_flag[ i ]` không được báo hiệu, `sps_subpic_treated_as_pic_flag[ i ]` có thể được suy ra là có giá trị thứ hai (ví dụ, 1).

Trong một phương án, khi hình ảnh hiện tại bao gồm hai hay hơn hai hình ảnh con (ví dụ, `pps_num_subpics_minus1 > 0`) và ít nhất một hình ảnh con trong số các hình ảnh con này không được xử lý như một hình ảnh (ví dụ, `sps_subpic_treated_as_pic_flag[ i ] == 0`), hình ảnh hiện tại có thể được giới hạn để không có loại đơn vị NAL được trộn (ví dụ, `pps_mixed_nalu_types_in_pic_flag = 0`). Tức là, khi hình ảnh hiện tại có loại đơn vị NAL được trộn (ví dụ, `pps_mixed_nalu_types_in_pic_flag == 1`), tất cả các hình ảnh

con trong hình ảnh hiện tại này có thể được giới hạn để được xử lý như một hình ảnh (ví dụ, `sps_subpic_treated_as_pic_flag[i] = 1`).

Ở dưới đây, loại đơn vị NAL theo các phương án của sáng chế sẽ được mô tả chi tiết đối với mỗi loại hình ảnh.

(1) Hình ảnh IRAP (điểm truy cập ngẫu nhiên nội)

Hình ảnh IRAP là hình ảnh có thể truy cập được một cách ngẫu nhiên và có thể có cùng loại đơn vị NAL chẳng hạn như `IDR_W_RADL`, `IDR_N_LP` hoặc `CRA_NUT` như được mô tả trên đây với tham chiếu đến bảng 1. Hình ảnh IRAP này có thể không tham chiếu đến hình ảnh khác ngoài hình ảnh IRAP cho việc dự đoán liên trong quy trình giải mã. Hình ảnh IRAP này có thể bao gồm hình ảnh IDR (làm tươi giải mã tức thời) và hình ảnh CRA (truy cập ngẫu nhiên sạch).

Hình ảnh thứ nhất trong luồng bit theo thứ tự của bộ giải mã có thể được giới hạn thành hình ảnh IRAP hoặc hình ảnh GDR (làm tươi giải mã dần dần). Đối với luồng bit lớp đơn lẻ, khi tập tham số thiết yếu cần được tham chiếu đến là khả dụng, mặc dù các hình ảnh đang đứng trước hình ảnh IRAP theo thứ tự giải mã hoàn toàn không được giải mã, hình ảnh IRAP này và tất cả các hình ảnh phi RASL theo sau hình ảnh IRAP này theo thứ tự giải mã có thể được giải mã một cách đúng đắn.

Trong một phương án, hình ảnh IRAP có thể không có loại đơn vị NAL được trộn. Tức là, đối với hình ảnh IRAP, `pps_mixed_nalu_types_in_pic_flag` có thể có giá trị thứ nhất (ví dụ, 0), và tất cả các lát trong hình ảnh IRAP này có thể có cùng loại đơn vị NAL trong khoảng từ `IDR_W_RADL` đến `CRA_NUT`. Kết quả là, khi lát thứ nhất cần được giải mã trong hình ảnh có loại đơn vị NAL trong khoảng từ `IDR_W_RADL` đến `CRA_NUT`, hình ảnh này có thể được xác định là hình ảnh IRAP.

(2) Hình ảnh CRA (truy cập ngẫu nhiên sạch)

Hình ảnh CRA là một hình ảnh trong số các hình ảnh IRAP và có thể có cùng loại đơn vị NAL như là `CRA_NUT` như được mô tả trên đây với tham chiếu đến bảng 1. Hình ảnh CRA này có thể không tham chiếu đến hình ảnh khác ngoài hình ảnh CRA cho việc dự đoán liên trong quy trình giải mã.

Hình ảnh CRA có thể là hình ảnh thứ nhất trong luồng bit theo thứ tự giải mã hoặc hình ảnh theo sau hình ảnh thứ nhất này. Hình ảnh CRA có thể liên quan đến các hình ảnh RADL hoặc RASL.

Khi NoIncorrectPicOutputFlag có giá trị thứ hai (ví dụ, 1) cho hình ảnh CRA, các hình ảnh RASL liên quan đến hình ảnh CRA này tham chiếu đến các hình ảnh mà chúng không có mặt trong luồng bit và do đó có thể không được giải mã và, kết quả là, có thể không được xuất ra bởi thiết bị giải mã ảnh. Ở đây, NoIncorrectPicOutputFlag có thể chỉ ra liệu các hình ảnh đứng trước hình ảnh điểm phục hồi (recovery point) theo thứ tự giải mã có được xuất ra trước hình ảnh điểm phục hồi này hay không. Ví dụ, NoIncorrectPicOutputFlag có giá trị thứ nhất (ví dụ, 0) có thể chỉ ra rằng các hình ảnh đứng trước hình ảnh điểm phục hồi theo thứ tự giải mã có được xuất ra trước hình ảnh điểm phục hồi này. Trong trường hợp này, hình ảnh CRA có thể không là hình ảnh thứ nhất trong luồng bit hoặc hình ảnh thứ nhất theo sau đơn vị NAL kết thúc trình tự (end of sequence, EOS) theo thứ tự giải mã, mà điều này có thể nghĩa là việc truy cập ngẫu nhiên không xảy ra. Ngược lại, NoIncorrectPicOutputFlag có giá trị thứ hai (ví dụ, 1) có thể chỉ ra rằng các hình ảnh đứng trước hình ảnh điểm phục hồi theo thứ tự giải mã không thể được xuất ra trước hình ảnh điểm phục hồi này. Trong trường hợp này, hình ảnh CRA có thể là hình ảnh thứ nhất trong luồng bit hoặc hình ảnh thứ nhất theo sau đơn vị NAL kết thúc trình tự (EOS) theo thứ tự giải mã, mà điều này có thể nghĩa là việc truy cập ngẫu nhiên có xảy ra. Trong khi đó, NoIncorrectPicOutputFlag có thể được gọi là NoOutputBeforeRecoveryFlag trong một số phương án.

Đối với tất cả các đơn vị hình ảnh (picture unit, PU) theo sau hình ảnh hiện tại theo thứ tự giải mã trong CLVS (coded layer video sequence, trình tự video lớp được lập mã), danh sách hình ảnh tham chiếu 0 (ví dụ, RefPicList[0]) và danh sách hình ảnh tham chiếu 1 (ví dụ, RefPicList[1]) cho một lát được chứa trong hình ảnh con CRA thuộc về các đơn vị hình ảnh (các PU) này có thể được giới hạn để không bao gồm hình ảnh nào bất kỳ đứng trước hình ảnh chứa hình ảnh con CRA này theo thứ tự giải mã bên trong mục nhập hoạt động (active entry). Ở đây, đơn vị hình ảnh (PU) có thể nghĩa là tập đơn vị NAL chứa nhiều đơn vị NAL mà chúng liên quan tương hỗ theo quy tắc phân loại được xác định trước và nằm liên tục theo thứ tự giải mã đối với một hình ảnh được lập mã.

(3) Hình ảnh IDR (làm tươi giải mã tức thời)

Hình ảnh IDR là một hình ảnh trong số các hình ảnh IRAP, và có thể có cùng loại đơn vị NAL như là IDR_W_RADL hoặc IDR_N_LP như được mô tả trên đây với tham

chiếu đến bảng 1. Hình ảnh IDR này có thể không tham chiếu đến hình ảnh khác ngoài hình ảnh IDR cho việc dự đoán liên trong quy trình giải mã.

Hình ảnh IDR có thể là hình ảnh thứ nhất trong luồng bit theo thứ tự giải mã hoặc có thể là hình ảnh theo sau hình ảnh thứ nhất này. Mỗi hình ảnh IDR có thể là hình ảnh thứ nhất của CVS (Coded Video Sequence, trình tự video được lập mã) theo thứ tự giải mã.

Khi hình ảnh IDR có cùng loại đơn vị NAL như là IDR_W_RADL cho mỗi đơn vị NAL, hình ảnh IDR có thể có các hình ảnh RADL liên quan. Ngược lại, khi hình ảnh IDR có cùng loại đơn vị NAL như là IDR_N_LP cho mỗi đơn vị NAL, hình ảnh IDR có thể không có các hình ảnh dẫn liên quan. Trong khi đó, hình ảnh IDR có thể không liên quan đến các hình ảnh RASL.

Đối với tất cả các đơn vị hình ảnh (PU) theo sau hình ảnh hiện tại theo thứ tự giải mã trong CLVS (trình tự video lớp được lập mã), danh sách hình ảnh tham chiếu 0 (ví dụ, RefPicList[0]) và danh sách hình ảnh tham chiếu 1 (ví dụ, RefPicList[1]) cho một lát được chứa trong hình ảnh con IDR thuộc về các đơn vị hình ảnh (các PU) này có thể được giới hạn để không bao gồm hình ảnh nào bất kỳ đứng trước hình ảnh chứa hình ảnh con IDR này theo thứ tự giải mã bên trong mục nhập hoạt động.

(4) Hình ảnh RADL (dẫn có thể giải mã được truy cập ngẫu nhiên)

Hình ảnh RADL là một hình ảnh trong số các hình ảnh dẫn và có thể có cùng loại đơn vị NAL như là RADL_NUT như được mô tả trên đây với tham chiếu đến bảng 1.

Hình ảnh RADL có thể không được sử dụng làm hình ảnh tham chiếu trong quy trình giải mã hình ảnh kéo theo đang có cùng hình ảnh IRAP liên quan. Khi `field_seq_flag` có giá trị thứ nhất (ví dụ, 0) đối với hình ảnh RADL, hình ảnh RADL này có thể đứng trước tất cả các hình ảnh phi dẫn có cùng hình ảnh IRAP liên quan theo thứ tự giải mã. Ở đây, `field_seq_flag` có thể chỉ ra liệu CLVS (trình tự video lớp được lập mã) truyền tải hình ảnh chỉ ra các trường hay hình ảnh chỉ ra các khung. Ví dụ, `field_seq_flag` có giá trị thứ nhất (ví dụ, 0) có thể chỉ ra rằng CLVS truyền tải hình ảnh chỉ ra các khung. Ngược lại, `field_seq_flag` có giá trị thứ hai (ví dụ, 1) có thể chỉ ra rằng CLVS truyền tải hình ảnh chỉ ra các trường.

(5) Hình ảnh RASL (dẫn được bỏ qua truy cập ngẫu nhiên)

Hình ảnh RASL là một hình ảnh trong số các hình ảnh dẫn và có thể có cùng loại

đơn vị NAL như là RASL_NUT như được mô tả trên đây với tham chiếu đến bảng 1.

Trong một ví dụ, tất cả các hình ảnh RASL có thể đều là các hình ảnh dẫn của hình ảnh CRA liên quan. Khi NoIncorrectPicOutputFlag có giá trị thứ hai (ví dụ, 1) cho hình ảnh CRA, các hình ảnh RASL tham chiếu đến các hình ảnh mà chúng không có mặt trong luồng bit và do đó có thể không được giải mã và, kết quả là, có thể không được xuất ra bởi thiết bị giải mã ảnh.

Hình ảnh RASL có thể không được sử dụng làm hình ảnh tham chiếu trong quy trình giải mã các hình ảnh phi RASL. Tuy nhiên, khi hình ảnh RADL, mà nó thuộc về cùng lớp với hình ảnh RASL và liên quan đến cùng hình ảnh CRA, có mặt, hình ảnh RASL này có thể được sử dụng làm hình ảnh tham chiếu được đặt cùng vị trí cho việc dự đoán liên của hình ảnh con RADL được chứa trong hình ảnh RADL này.

Khi field_seq_flag có giá trị thứ nhất (ví dụ, 0) đối với hình ảnh RASL, hình ảnh RASL này có thể đứng trước tất cả các hình ảnh phi dẫn của hình ảnh CRA liên quan đến hình ảnh RASL này theo thứ tự giải mã.

(6) Hình ảnh kéo theo

Hình ảnh kéo theo là hình ảnh phi IRAP theo sau hình ảnh IRAP hoặc hình ảnh GDR liên quan theo thứ tự xuất ra và có thể không phải là hình ảnh STSA. Thêm vào đó, hình ảnh kéo theo có thể theo sau hình ảnh IRAP liên quan theo thứ tự giải mã. Tức là, hình ảnh kéo theo mà nó theo sau hình ảnh IRAP liên quan theo thứ tự xuất ra nhưng đứng trước hình ảnh IRAP liên quan này theo thứ tự giải mã có thể không được phép.

(7) Hình ảnh GDR (làm tươi giải mã dần dần)

Hình ảnh GDR là hình ảnh truy cập được một cách ngẫu nhiên và có thể có cùng loại đơn vị NAL như là GDR_NUT như được mô tả trên đây với tham chiếu đến bảng 1.

(8) Hình ảnh STSA (truy cập lớp con theo thời gian từng bước)

Hình ảnh STSA là hình ảnh truy cập được một cách ngẫu nhiên và có thể có cùng loại đơn vị NAL như là STSA_NUT như được mô tả trên đây với tham chiếu đến bảng 1.

Hình ảnh STSA có thể không tham chiếu đến các hình ảnh có cùng TemporalId với hình ảnh STSA này cho việc dự đoán liên. Ở đây, TemporalId có thể là bộ nhận dạng chỉ ra lớp theo thời gian, ví dụ, lớp con theo thời gian trong việc lập mã video định

tỷ lệ được. Trong một phương án, các hình ảnh STSA có thể được giới hạn để có TemporalId lớn hơn 0.

Các hình ảnh có cùng TemporalId với hình ảnh STSA và theo sau hình ảnh STSA này theo thứ tự giải mã có thể không tham chiếu đến các hình ảnh có cùng TemporalId với hình ảnh STSA này và đứng trước hình ảnh STSA này theo thứ tự giải mã, cho việc dự đoán liên. Hình ảnh STSA có thể kích hoạt việc chuyển đổi lên từ lớp con ngay dưới lớp con hiện tại, mà hình ảnh STSA này thuộc về lớp con này, sang lớp con hiện tại.

Fig.13 là hình minh họa thứ tự giải mã và thứ tự xuất ra đối với mỗi loại hình ảnh.

Nhiều hình ảnh có thể được phân loại thành hình ảnh I, hình ảnh P hoặc hình ảnh B tùy theo phương pháp dự đoán. Hình ảnh I có thể chỉ hình ảnh mà chỉ việc dự đoán nội là áp dụng được với nó và có thể được giải mã mà không tham chiếu đến hình ảnh khác. Hình ảnh I có thể được gọi là hình ảnh nội và có thể bao gồm hình ảnh IRAP được mô tả trên đây. Hình ảnh P có thể chỉ hình ảnh mà việc dự đoán nội và việc dự đoán liên đơn hướng là áp dụng được với nó, và có thể được giải mã bằng cách tham chiếu đến một hình ảnh khác. Hình ảnh B có thể chỉ hình ảnh mà việc dự đoán nội và việc dự đoán liên song hướng/đơn hướng là áp dụng được với nó và có thể được giải mã nhờ sử dụng một hoặc hai hình ảnh khác. Hình ảnh P và hình ảnh B có thể được gọi là hình ảnh liên và có thể bao gồm hình ảnh RADL, hình ảnh RASL và hình ảnh kéo theo được mô tả trên đây.

Hình ảnh liên có thể còn được phân loại thành hình ảnh dẫn (leading picture, LP) hoặc hình ảnh phi dẫn (non-leading picture, NLP) tùy theo thứ tự giải mã và thứ tự xuất ra. Hình ảnh dẫn có thể chỉ hình ảnh mà nó theo sau hình ảnh IRAP theo thứ tự giải mã và đứng trước hình ảnh IRAP này theo thứ tự xuất ra, và có thể bao gồm hình ảnh RADL và hình ảnh RASL được mô tả trên đây. Hình ảnh phi dẫn có thể chỉ hình ảnh mà nó theo sau hình ảnh IRAP theo thứ tự giải mã và thứ tự xuất ra, và có thể bao gồm hình ảnh kéo theo được mô tả trên đây.

Trên Fig.13, tên của mỗi hình ảnh có thể chỉ loại hình ảnh của nó. Ví dụ, hình ảnh I5 có thể là hình ảnh I và các hình ảnh B0, B2, B3, B4 và B6 có thể là các hình ảnh B, và các hình ảnh P1 và P7 có thể là các hình ảnh P. Thêm vào đó, trên Fig.13, mỗi mũi tên có thể biểu thị hướng tham chiếu giữa các hình ảnh. Ví dụ, hình ảnh B0 có thể được giải mã bằng cách tham chiếu đến hình ảnh P1.

Tham chiếu đến Fig.13, hình ảnh I5 có thể là hình ảnh IRAP, ví dụ, hình ảnh CRA. Khi việc truy cập ngẫu nhiên đối với hình ảnh I5 xảy ra, hình ảnh I5 có thể là hình ảnh thứ nhất theo thứ tự giải mã.

Các hình ảnh B0 và P1 có thể đứng trước hình ảnh I5 theo thứ tự giải mã và xây dựng nên trình tự video riêng biệt so với hình ảnh I5. Các hình ảnh B2, B3, B4, B6 và P7 có thể theo sau hình ảnh I5 theo thứ tự giải mã, và xây dựng nên một trình tự video cùng với hình ảnh I5.

Các hình ảnh B2, B3 và B4 theo sau hình ảnh I5 theo thứ tự giải mã và đứng trước hình ảnh I5 theo thứ tự xuất ra và do đó có thể được phân loại thành hình ảnh dẫn. Hình ảnh B2 có thể được giải mã bằng cách tham chiếu đến hình ảnh P1 đang đứng trước hình ảnh I5 theo thứ tự giải mã. Theo đó, khi việc truy cập ngẫu nhiên xảy ra từ hình ảnh I5, hình ảnh B2 có thể không được giải mã một cách đúng đắn bởi việc tham chiếu đến hình ảnh P1 mà nó không có mặt trong luồng bit. Loại hình ảnh như là hình ảnh B2 có thể được gọi là hình ảnh RASL. Ngược lại, hình ảnh B3 có thể được giải mã bằng cách tham chiếu đến hình ảnh I5 đang đứng trước hình ảnh B3 theo thứ tự giải mã. Theo đó, khi việc truy cập ngẫu nhiên xảy ra từ hình ảnh I5, hình ảnh B3 có thể được giải mã một cách đúng đắn bởi việc tham chiếu đến hình ảnh I5 đã được giải mã trước. Thêm vào đó, B4 có thể được giải mã bằng cách tham chiếu đến hình ảnh I5 và hình ảnh B3 đang đứng trước hình ảnh B4 theo thứ tự giải mã. Theo đó, khi việc truy cập ngẫu nhiên xảy ra từ hình ảnh I5, hình ảnh B4 có thể được giải mã một cách đúng đắn bởi việc tham chiếu đến hình ảnh I5 và hình ảnh B3 đã được giải mã trước. Loại hình ảnh như là các hình ảnh B3 và B4 có thể được gọi là hình ảnh RADL.

Trong khi đó, các hình ảnh B6 và P7 theo sau hình ảnh I5 theo thứ tự giải mã và thứ tự xuất ra và do đó có thể được phân loại thành hình ảnh phi dẫn. Hình ảnh B6 có thể được giải mã bằng cách tham chiếu đến hình ảnh I5 và hình ảnh P7 đang đứng trước theo thứ tự giải mã. Theo đó, khi việc truy cập ngẫu nhiên xảy ra từ hình ảnh I5, hình ảnh B6 có thể được giải mã một cách đúng đắn bởi việc tham chiếu đến hình ảnh I5 và hình ảnh P7 đã được giải mã trước.

Trong một trình tự video, quy trình giải mã và quy trình xuất ra có thể được thực hiện theo các thứ tự khác nhau dựa trên loại hình ảnh. Ví dụ, khi một trình tự video bao gồm hình ảnh IRAP, hình ảnh dẫn và hình ảnh phi dẫn, quy trình giải mã có thể được

thực hiện theo thứ tự là hình ảnh IRAP, hình ảnh dẫn và hình ảnh phi dẫn và quy trình xuất ra có thể được thực hiện theo thứ tự là hình ảnh dẫn, hình ảnh IRAP và hình ảnh phi dẫn.

Fig.14 là lưu đồ minh họa phương pháp xác định loại đơn vị NAL của hình ảnh hiện tại theo một phương án của sáng chế.

Tham chiếu đến Fig.14, thiết bị mã hóa ảnh có thể xác định xem liệu hình ảnh hiện tại có bao gồm hai hay hơn hai hình ảnh con hay không (S1410). Thông tin phân vùng của hình ảnh hiện tại có thể được báo hiệu nhờ sử dụng một hoặc nhiều phần tử cú pháp trong cú pháp mức cao. Ví dụ, qua tập tham số hình ảnh (PPS) được mô tả trên đây với tham chiếu đến Fig.7, `no_pic_partition_flag` chỉ ra liệu hình ảnh hiện tại có được phân vùng hay không và `pps_num_subpics_minus1` chỉ ra số lượng hình ảnh con được chứa trong hình ảnh hiện tại có thể được báo hiệu. Khi hình ảnh hiện tại bao gồm hai hay hơn hai hình ảnh con, `no_pic_partition_flag` có thể có giá trị thứ nhất (ví dụ, 0) và `pps_num_subpics_minus` có thể có giá trị lớn hơn 0.

Khi hình ảnh hiện tại không bao gồm hai hay hơn hai hình ảnh con ('Không' của S1410), thiết bị mã hóa ảnh có thể xác định rằng hình ảnh hiện tại có loại đơn vị NAL đơn lẻ (S1450). Tức là, tất cả các lát trong hình ảnh hiện tại có thể có một loại bất kỳ trong số các loại đơn vị NAL được mô tả trên đây với tham chiếu đến bảng 1.

Ngược lại, khi hình ảnh hiện tại bao gồm hai hay hơn hai hình ảnh con ('Có' của S1410), thiết bị mã hóa ảnh có thể xác định xem liệu mỗi hình ảnh con trong hình ảnh hiện tại có được xử lý như một hình ảnh hay không (S1420). Thông tin chỉ ra liệu hình ảnh con có được xử lý như một hình ảnh hay không có thể được báo hiệu nhờ sử dụng phần tử cú pháp được xác định trước trong cú pháp mức cao, ví dụ, qua tập tham số trình tự (SPS) được mô tả trên đây với tham chiếu đến Fig.12, `sps_subpic_treated_as_pic_flag[ i ]` chỉ ra liệu hình ảnh con có được xử lý như một hình ảnh hay không có thể được báo hiệu. Khi mỗi hình ảnh con trong hình ảnh hiện tại được xử lý như một hình ảnh, `sps_subpic_treated_as_pic_flag[ i ]` có thể có giá trị thứ hai (ví dụ, 1).

Khi mỗi hình ảnh con trong hình ảnh hiện tại không được xử lý như một hình ảnh ('Không' của S1420), thiết bị mã hóa ảnh có thể xác định rằng hình ảnh hiện tại có loại đơn vị NAL đơn lẻ (S1450).

Ngược lại, khi mỗi hình ảnh con trong hình ảnh hiện tại được xử lý như một hình ảnh ('Có' của S1420), thiết bị mã hóa ảnh có thể xác định xem liệu ít nhất một số lát trong số nhiều lát trong hình ảnh hiện tại có có các loại đơn vị NAL khác nhau hay không (S1430).

Khi ít nhất một số lát trong số nhiều lát này trong hình ảnh hiện tại có các loại đơn vị NAL khác nhau ('Có' trong S1430), thiết bị mã hóa ảnh có thể xác định rằng hình ảnh hiện tại có loại đơn vị NAL được trộn (S1440).

Ngược lại, khi tất cả các lát trong số nhiều lát này trong hình ảnh hiện tại đều có cùng loại đơn vị NAL ('Không' trong S1430), thiết bị mã hóa ảnh có thể xác định rằng hình ảnh hiện tại có loại đơn vị NAL đơn lẻ (S1450).

Khi hình ảnh hiện tại có loại đơn vị NAL được trộn, thiết bị mã hóa ảnh có thể tạo ra luồng bit con cho mỗi hình ảnh con được chứa trong hình ảnh hiện tại. Trong trường hợp này, nhiều luồng bit con được tạo ra cho hình ảnh hiện tại có thể cấu thành nên một luồng bit.

Thông tin chỉ ra liệu hình ảnh hiện tại có có loại đơn vị NAL được trộn hay không có thể được báo hiệu nhờ sử dụng phần tử cú pháp được xác định trước trong cú pháp mức cao. Ví dụ, qua tập tham số hình ảnh (PPS) được mô tả trên đây với tham chiếu đến Fig.11, `pps_mixed_nalu_types_in_pic_flag` chỉ ra liệu hình ảnh hiện tại có có loại đơn vị NAL được trộn hay không có thể được báo hiệu. Trong trường hợp này, thiết bị giải mã ảnh có thể xác định xem liệu hình ảnh hiện tại có có loại đơn vị NAL được trộn hay không dựa trên `pps_mixed_nalu_types_in_pic_flag`. Ví dụ, khi `pps_mixed_nalu_types_in_pic_flag` có giá trị thứ nhất (ví dụ, 0), thiết bị giải mã ảnh có thể xác định rằng hình ảnh hiện tại có loại đơn vị NAL đơn lẻ. Ngược lại, khi `pps_mixed_nalu_types_in_pic_flag` có giá trị thứ hai (ví dụ, 1), thiết bị giải mã ảnh có thể xác định rằng hình ảnh hiện tại có loại đơn vị NAL được trộn.

Trong khi đó, tất cả các lát trong mỗi hình ảnh con được chứa trong hình ảnh hiện tại có thể có cùng loại đơn vị NAL. Tức là, mỗi hình ảnh con có thể có chỉ một loại đơn vị NAL đơn lẻ. Ngoài ra, khi hình ảnh hiện tại có loại đơn vị NAL được trộn, ít nhất một số hình ảnh con trong số các hình ảnh con được chứa trong hình ảnh hiện tại có thể có các loại đơn vị NAL khác nhau. Ví dụ, hình ảnh con thứ nhất được chứa trong hình ảnh hiện tại có thể có loại đơn vị NAL thứ nhất, và hình ảnh con thứ hai được chứa trong

hình ảnh hiện tại có thể có loại đơn vị NAL thứ hai khác với loại đơn vị NAL thứ nhất.

Loại của mỗi hình ảnh con được chứa trong hình ảnh hiện tại có thể được xác định dựa trên loại đơn vị NAL của mỗi hình ảnh con. Ví dụ, như được mô tả trên đây với tham chiếu đến bảng 1, loại của hình ảnh con có loại đơn vị NAL chẳng hạn như TRAIL_NUT có thể được xác định là hình ảnh con kéo theo. Ngoài ra, loại của hình ảnh con có loại đơn vị NAL chẳng hạn như STSA_NUT có thể được xác định là hình ảnh con STSA, và loại của hình ảnh con có loại đơn vị NAL chẳng hạn như GDR_NUT có thể được xác định là hình ảnh con GDR. Ngoài ra, loại của hình ảnh con có loại đơn vị NAL chẳng hạn như RADL_NUT có thể được xác định là hình ảnh con RADL, và loại của hình ảnh con có loại đơn vị NAL chẳng hạn như RASL_NUT có thể được xác định là hình ảnh con RASL. Thêm vào đó, loại của hình ảnh con có loại đơn vị NAL chẳng hạn như IDR_W_RADL hoặc IDR_N_LP có thể được xác định là hình ảnh con IDR, và loại của hình ảnh con có loại đơn vị NAL chẳng hạn như CRA_NUT có thể được xác định là hình ảnh con CRA.

Theo các phương án của sáng chế, khi hình ảnh hiện tại bao gồm hai hay hơn hai hình ảnh con và mỗi hình ảnh con được xử lý như một hình ảnh, hình ảnh hiện tại có thể có loại đơn vị NAL được trộn. Do đó, trong các ứng dụng hoặc trường hợp sử dụng khác nhau, hiệu suất mã hóa/giải mã ảnh có thể được cải thiện hơn nữa.

Ở dưới đây, phương pháp mã hóa/giải mã ảnh theo một phương án của sáng chế sẽ được mô tả chi tiết với tham chiếu đến Fig.15 và Fig.16.

Fig.15 là lưu đồ minh họa phương pháp mã hóa ảnh theo một phương án của sáng chế.

Phương pháp mã hóa ảnh trên Fig.15 có thể được thực hiện bởi thiết bị mã hóa ảnh trên Fig.2. Ví dụ, bước S1510 có thể được thực hiện bởi bộ phân vùng ảnh 110, và các bước S1520 và S1530 có thể được thực hiện bởi bộ mã hóa entropy 190.

Tham chiếu đến Fig.15, thiết bị mã hóa ảnh có thể phân vùng hình ảnh hiện tại thành hai hay hơn hai hình ảnh con (S1510).

Trong một phương án, nếu hình ảnh hiện tại có loại đơn vị NAL được trộn thì hình ảnh hiện tại có thể được giới hạn để bao gồm hai hay hơn hai hình ảnh con. Tức là, dựa trên việc ít nhất một số lát trong số nhiều lát được chứa trong hình ảnh hiện tại có các loại đơn vị NAL khác nhau, hình ảnh hiện tại có thể được phân vùng thành hai hay hơn

hai hình ảnh con.

Thông tin phân vùng của hình ảnh hiện tại có thể được báo hiệu nhờ sử dụng một hoặc nhiều phần tử cú pháp trong cú pháp mức cao. Ví dụ, qua tập tham số hình ảnh (PPS) được mô tả trên đây với tham chiếu đến Fig.7, thông tin cờ `no_pic_partition_flag` chỉ ra liệu hình ảnh hiện tại có được phân vùng hay không và thông tin cờ `pps_num_subpics_minus1` chỉ ra số lượng hình ảnh con được chứa trong hình ảnh hiện tại có thể được mã hóa/báo hiệu. Trong trường hợp này, khi hình ảnh hiện tại được phân vùng thành hai hay hơn hai hình ảnh con, `no_pic_partition_flag` có thể có giá trị thứ nhất (ví dụ, 0) và `pps_num_subpics_minus` có thể có giá trị lớn hơn 0.

Trong một phương án, khi hình ảnh hiện tại có loại đơn vị NAL được trộn, mỗi hình ảnh con được chứa trong hình ảnh hiện tại có thể được xử lý như một hình ảnh. Ví dụ, hình ảnh hiện tại có loại đơn vị NAL được trộn có thể bao gồm hình ảnh con thứ nhất và hình ảnh con thứ hai, và mỗi hình ảnh con trong số hình ảnh con thứ nhất và hình ảnh con thứ hai này có thể được mã hóa/giải mã một cách độc lập ngoại trừ hoạt động lọc trong vòng.

Thông tin về việc liệu hình ảnh con được chứa trong hình ảnh hiện tại có được xử lý như một hình ảnh hay không có thể được mã hóa/báo hiệu nhờ sử dụng phần tử cú pháp được xác định trước trong cú pháp mức cao hơn. Ví dụ, thông tin cờ `sps_subpic_treated_as_pic_flag[ i ]` chỉ ra liệu hình ảnh con thứ *i* được chứa trong hình ảnh hiện tại có được xử lý như một hình ảnh hay không có thể được mã hóa/báo hiệu qua tập tham số trình tự (SPS) được mô tả trên đây với tham chiếu đến Fig.12. Trong trường hợp này, khi hình ảnh con thứ *i* được chứa trong hình ảnh hiện tại không được xử lý như một hình ảnh, `sps_subpic_treated_as_pic_flag[ i ]` có thể có giá trị thứ nhất (ví dụ, 0). Mặt khác, khi hình ảnh con thứ *i* được chứa trong hình ảnh hiện tại được xử lý như một hình ảnh, `sps_subpic_treated_as_pic_flag[ i ]` có thể có giá trị thứ hai (ví dụ, 1).

Thiết bị mã hóa ảnh có thể xác định loại đơn vị NAL của mỗi lát trong số nhiều lát được chứa trong hình ảnh hiện tại (S1520).

Trong một phương án, tất cả các lát được chứa trong mỗi hình ảnh con trong hình ảnh hiện tại có thể có cùng loại đơn vị NAL. Ví dụ, tất cả các lát được chứa trong hình ảnh con thứ nhất của hình ảnh hiện tại có thể có loại đơn vị NAL thứ nhất, và tất cả các

lát được chứa trong hình ảnh con thứ hai của hình ảnh hiện tại có thể có loại đơn vị NAL thứ hai. Trong trường hợp này, khi hình ảnh hiện tại có loại đơn vị NAL được trộn, loại đơn vị NAL thứ hai có thể khác với loại đơn vị NAL thứ nhất. Theo cách khác, khi hình ảnh hiện tại có loại đơn vị NAL đơn lẻ, loại đơn vị NAL thứ hai có thể giống với loại đơn vị NAL thứ nhất.

Trong một phương án, việc liệu ít nhất một số lát trong số nhiều lát được chứa trong hình ảnh hiện tại có các loại đơn vị NAL khác nhau hay không (tức là, việc liệu hình ảnh hiện tại có loại đơn vị NAL được trộn hay không) có thể được mã hóa/báo hiệu nhờ sử dụng phần tử cú pháp. Ví dụ, thông tin cờ `pps_mixed_nalu_types_in_pic_flag` chỉ ra liệu hình ảnh hiện tại có các loại đơn vị NAL được trộn hay không có thể được mã hóa/báo hiệu qua tập tham số hình ảnh (PPS) được mô tả trên đây với tham chiếu đến Fig.11. Trong trường hợp này, khi hình ảnh hiện tại có loại đơn vị NAL đơn lẻ, `pps_mixed_nalu_types_in_pic_flag` có thể có giá trị thứ nhất (ví dụ, 0). Ngược lại, khi hình ảnh hiện tại có loại đơn vị NAL được trộn, `pps_mixed_nalu_types_in_pic_flag` có thể có giá trị thứ hai (ví dụ, 1).

Trong một phương án, giá trị của `pps_mixed_nalu_types_in_pic_flag` có thể được xác định dựa trên việc liệu loại đơn vị NAL được trộn có áp dụng được đối với tất cả các hình ảnh trong tập lớp đầu ra (OLS) hay không. Ví dụ, khi loại đơn vị NAL được trộn không áp dụng được đối với tất cả các hình ảnh trong tập lớp đầu ra (OLS), `pps_mixed_nalu_types_in_pic_flag` có thể được giới hạn để có giá trị thứ nhất (ví dụ, 0). Ngược lại, khi loại đơn vị NAL được trộn áp dụng được đối với tất cả các hình ảnh trong tập lớp đầu ra (OLS), `pps_mixed_nalu_types_in_pic_flag` có thể có giá trị thứ nhất (ví dụ, 0) hoặc giá trị thứ hai (ví dụ, 1).

Trong khi đó, việc liệu loại đơn vị NAL được trộn có áp dụng được đối với tất cả các hình ảnh trong tập lớp đầu ra (OLS) hay không có thể được mã hóa/báo hiệu nhờ sử dụng phần tử cú pháp được xác định trước trong cú pháp mức cao hơn. Ví dụ, thông tin cờ `gci_no_mixed_nalu_types_in_pic_constraint_flag` chỉ ra liệu loại đơn vị NAL được trộn có được giới hạn hay không có thể được mã hóa/báo hiệu qua `general_constraints_info` chứa thông tin ràng buộc chung. Trong trường hợp này, khi loại đơn vị NAL được trộn áp dụng được đối với tất cả các hình ảnh trong tập lớp đầu ra (OLS), `gci_no_mixed_nalu_types_in_pic_constraint_flag` có thể có giá trị thứ nhất

(ví dụ, 0). Ngược lại, khi loại đơn vị NAL được trộn không áp dụng được đối với tất cả các hình ảnh trong tập lớp đầu ra (OLS), `gci_no_mixed_nalu_types_in_pic_constraint_flag` có thể có giá trị thứ hai (ví dụ, 1).

Trong một phương án, giá trị của `pps_mixed_nalu_types_in_pic_flag` có thể được xác định dựa trên việc liệu hình ảnh hiện tại có bao gồm chỉ một hình ảnh con hay không. Ví dụ, khi hình ảnh hiện tại chỉ bao gồm một hình ảnh con, `pps_mixed_nalu_types_in_pic_flag` có thể được giới hạn để có giá trị thứ nhất (ví dụ, 0). Mặt khác, khi hình ảnh hiện tại bao gồm hai hay hơn hai hình ảnh con, `pps_mixed_nalu_types_in_pic_flag` có thể có giá trị thứ nhất (ví dụ, 0) hoặc giá trị thứ hai (ví dụ, 1).

Trong một phương án, giá trị của `pps_mixed_nalu_types_in_pic_flag` có thể được xác định dựa trên việc liệu mỗi hình ảnh con trong số tất cả các hình ảnh con được chứa trong hình ảnh hiện tại có được xử lý như một hình ảnh hay không. Ví dụ, khi ít nhất một hình ảnh con trong số các hình ảnh con được chứa trong hình ảnh hiện tại không được xử lý như một hình ảnh, `pps_mixed_nalu_types_in_pic_flag` có thể được giới hạn để có giá trị thứ nhất (ví dụ, 0). Ngược lại, khi mỗi hình ảnh con trong số tất cả các hình ảnh con được chứa trong hình ảnh hiện tại đều được xử lý như một hình ảnh, `pps_mixed_nalu_types_in_pic_flag` có thể có giá trị thứ nhất (ví dụ, 0) hoặc giá trị thứ hai (ví dụ, 1).

Thiết bị mã hóa ảnh có thể mã hóa các hình ảnh con trong hình ảnh hiện tại dựa trên loại đơn vị NAL được xác định trong S1520 (S1530). Tại thời điểm này, quy trình mã hóa cho mỗi lát có thể được thực hiện trong các đơn vị lập mã (các CU) dựa trên chế độ dự đoán được xác định trước, như được mô tả trên đây. Trong khi đó, mỗi hình ảnh con có thể được lập mã một cách độc lập để xây dựng nên các luồng bit (hình ảnh con) khác nhau. Ví dụ, luồng bit con thứ nhất chứa thông tin mã hóa của hình ảnh con thứ nhất có thể được xây dựng và luồng bit con thứ hai chứa thông tin mã hóa của hình ảnh con thứ hai có thể được xây dựng.

Fig.16 là lưu đồ minh họa phương pháp giải mã ảnh theo một phương án của sáng chế.

Phương pháp giải mã ảnh trên Fig.16 có thể được thực hiện bởi thiết bị giải mã ảnh trên Fig.3. Ví dụ, các bước S1610 và S1620 có thể được thực hiện bởi bộ giải mã

entropy 210, và bước S1630 có thể được thực hiện bởi bộ giải lượng tử hóa 220 đến bộ dự đoán nội 265.

Tham chiếu đến Fig.16, thiết bị giải mã ảnh có thể thu, từ luồng bit, thông tin loại đơn vị NAL VCL của hình ảnh hiện tại và thông tin cờ thứ nhất chỉ rõ liệu hình ảnh con được chứa trong hình ảnh hiện tại có được xử lý như một hình ảnh hay không (S1610).

Thông tin loại đơn vị NAL của hình ảnh hiện tại có thể bao gồm giá trị loại đơn vị NAL của đơn vị NAL VCL bao gồm dữ liệu ảnh được mã hóa (ví dụ, dữ liệu lát) của hình ảnh hiện tại. Giá trị loại đơn vị NAL này có thể được thu bằng cách phân tích cú pháp phần tử cú pháp `nal_unit_type` được chứa trong tiêu đề đơn vị NAL của đơn vị NAL VCL.

Thông tin cờ thứ nhất có thể được thu bằng cách phân tích cú pháp cú pháp mức cao hơn, ví dụ như `sps_subpic_treated_as_pic_flag[ i ]` được chứa trong tập tham số trình tự (SPS). Trong trường hợp này, khi hình ảnh con thứ *i* được chứa trong hình ảnh hiện tại không được xử lý như một hình ảnh, `sps_subpic_treated_as_pic_flag[ i ]` có thể có giá trị thứ nhất (ví dụ, 0). Mặt khác, khi hình ảnh con thứ *i* được chứa trong hình ảnh hiện tại được xử lý như một hình ảnh, `sps_subpic_treated_as_pic_flag[ i ]` có thể có giá trị thứ hai (ví dụ, 1).

Thiết bị giải mã ảnh có thể xác định loại đơn vị NAL của mỗi lát trong số nhiều lát trong hình ảnh hiện tại dựa trên thông tin loại đơn vị NAL VCL của hình ảnh hiện tại (S1620).

Trong một phương án, khi ít nhất một số lát trong số nhiều lát trong hình ảnh hiện tại có các loại đơn vị NAL khác nhau (cụ thể là, khi hình ảnh hiện tại có loại đơn vị NAL được trộn), hình ảnh hiện tại bao gồm hai hay hơn hai hình ảnh con.

Trong một phương án, khi ít nhất một số lát trong số nhiều lát trong hình ảnh hiện tại có các loại đơn vị NAL khác nhau, mỗi hình ảnh con trong số tất cả các hình ảnh con được chứa trong hình ảnh hiện tại có thể được xử lý như một hình ảnh. Trong trường hợp này, đối với tất cả các hình ảnh con được chứa trong hình ảnh hiện tại, thông tin cờ thứ nhất (ví dụ, `sps_subpic_treated_as_pic_flag[ i ]`) có thể có giá trị thứ hai (ví dụ, 1).

Trong một phương án, tất cả các lát được chứa trong mỗi hình ảnh con trong hình ảnh hiện tại có thể có cùng loại đơn vị NAL. Ví dụ, tất cả các lát được chứa trong hình ảnh con thứ nhất của hình ảnh hiện tại có thể có loại đơn vị NAL thứ nhất, và tất cả các

lát được chứa trong hình ảnh con thứ hai của hình ảnh hiện tại có thể có loại đơn vị NAL thứ hai. Trong trường hợp này, khi hình ảnh hiện tại có loại đơn vị NAL được trộn, loại đơn vị NAL thứ hai có thể khác với loại đơn vị NAL thứ nhất. Ngược lại, khi hình ảnh hiện tại có loại đơn vị NAL đơn lẻ, loại đơn vị NAL thứ hai có thể giống với loại đơn vị NAL thứ nhất.

Trong một phương án, việc liệu ít nhất một số lát trong số nhiều lát trong hình ảnh hiện tại có các loại đơn vị NAL khác nhau hay không có thể được xác định dựa trên cú pháp mức cao hơn, ví dụ như thông tin cờ thứ hai (ví dụ, `pps_mixed_nalu_types_in_pic_flag`) được thu từ tập tham số hình ảnh (PPS). Ví dụ, khi thông tin cờ thứ hai có giá trị thứ nhất (ví dụ, 0), tất cả các lát trong hình ảnh hiện tại có thể có cùng loại đơn vị NAL. Ngược lại, khi thông tin cờ thứ hai có giá trị thứ hai (ví dụ, 1), ít nhất một số lát trong số nhiều lát trong hình ảnh hiện tại có thể có các loại đơn vị NAL khác nhau.

Trong một phương án, thông tin cờ thứ hai (ví dụ, `pps_mixed_nalu_types_in_pic_flag`) có thể có giá trị được xác định trước dựa trên việc liệu loại đơn vị NAL được trộn có áp dụng được đối với tất cả các hình ảnh trong tập lớp đầu ra (OLS) hay không. Ví dụ, khi loại đơn vị NAL được trộn không áp dụng được đối với tất cả các hình ảnh trong tập lớp đầu ra (OLS), thông tin cờ thứ hai có thể có giá trị thứ nhất (ví dụ, 0). Trong trường hợp này, thiết bị giải mã ảnh có thể suy ra giá trị của thông tin cờ thứ hai là giá trị thứ nhất (ví dụ, 0) mà không phân tích cú pháp một cách riêng rẽ thông tin cờ thứ hai này. Ngược lại, khi loại đơn vị NAL được trộn áp dụng được đối với tất cả các hình ảnh trong tập lớp đầu ra (OLS), thông tin cờ thứ hai có thể có giá trị thứ nhất (ví dụ, 0) hoặc giá trị thứ hai (ví dụ, 1).

Trong một phương án, thông tin cờ thứ hai (ví dụ, `pps_mixed_nalu_types_in_pic_flag`) có thể có giá trị được xác định trước dựa trên việc liệu hình ảnh hiện tại có bao gồm chỉ một hình ảnh con hay không. Ví dụ, khi hình ảnh hiện tại chỉ bao gồm một hình ảnh con, thông tin cờ thứ hai có thể chỉ có giá trị thứ nhất (ví dụ, 0). Trong trường hợp này, thiết bị giải mã ảnh có thể suy ra giá trị của thông tin cờ thứ hai là giá trị thứ nhất (ví dụ, 0) mà không phân tích cú pháp một cách riêng rẽ thông tin cờ thứ hai này. Ngược lại, khi hình ảnh hiện tại bao gồm hai hay hơn hai hình ảnh con, thông tin cờ thứ hai có thể có giá trị thứ nhất (ví dụ, 0) hoặc giá trị thứ hai (ví

dụ, 1).

Trong một phương án, thông tin cờ thứ hai (ví dụ, `pps_mixed_nalu_types_in_pic_flag`) có thể có giá trị được xác định trước dựa trên việc liệu mỗi hình ảnh con trong số tất cả các hình ảnh con được chứa trong hình ảnh hiện tại có được xử lý như một hình ảnh hay không. Ví dụ, khi ít nhất một hình ảnh con trong số các hình ảnh con được chứa trong hình ảnh hiện tại không được xử lý như một hình ảnh, `pps_mixed_nalu_types_in_pic_flag` có thể chỉ có giá trị thứ nhất (ví dụ, 0). Trong trường hợp này, thiết bị giải mã ảnh có thể suy ra giá trị của thông tin cờ thứ hai là giá trị thứ nhất (ví dụ, 0) mà không phân tích cú pháp một cách riêng rẽ thông tin cờ thứ hai này. Ngược lại, khi mỗi hình ảnh con trong số tất cả các hình ảnh con được chứa trong hình ảnh hiện tại đều được xử lý như một hình ảnh, `pps_mixed_nalu_types_in_pic_flag` có thể có giá trị thứ nhất (ví dụ, 0) hoặc giá trị thứ hai (ví dụ, 1).

Trong khi đó, loại của mỗi hình ảnh con được chứa trong hình ảnh hiện tại có thể được xác định dựa trên loại đơn vị NAL của mỗi hình ảnh con. Ví dụ, như được mô tả trên đây với tham chiếu đến bảng 1, loại của hình ảnh con có loại đơn vị NAL chẳng hạn như `TRAIL_NUT` có thể được xác định là hình ảnh con kéo theo. Ngoài ra, loại của hình ảnh con có loại đơn vị NAL chẳng hạn như `STSA_NUT` có thể được xác định là hình ảnh con STSA, và loại của hình ảnh con có loại đơn vị NAL chẳng hạn như `GDR_NUT` có thể được xác định là hình ảnh con GDR. Ngoài ra, loại của hình ảnh con có loại đơn vị NAL chẳng hạn như `RADL_NUT` có thể được xác định là hình ảnh con RADL, và loại của hình ảnh con có loại đơn vị NAL chẳng hạn như `RASL_NUT` có thể được xác định là hình ảnh con RASL. Thêm vào đó, loại của hình ảnh con có loại đơn vị NAL chẳng hạn như `IDR_W_RADL` hoặc `IDR_N_LP` có thể được xác định là hình ảnh con IDR, và loại của hình ảnh con có loại đơn vị NAL chẳng hạn như `CRA_NUT` có thể được xác định là hình ảnh con CRA.

Thiết bị giải mã ảnh có thể giải mã nhiều lát trong hình ảnh hiện tại dựa trên thông tin cờ thứ nhất được thu trong bước S1610 và loại đơn vị NAL được xác định trong S1620 (S1630). Tại thời điểm này, quy trình giải mã của mỗi lát có thể được thực hiện trong các đơn vị lập mã (CU) dựa trên chế độ dự đoán được xác định trước như được mô tả trên đây.

Theo một phương án của sáng chế, hình ảnh hiện tại có thể có hai hay hơn hai loại

đơn vị NAL dựa trên cấu trúc của hình ảnh con.

Tên của phần tử cú pháp được mô tả trong sáng chế này có thể bao gồm thông tin về vị trí nơi phần tử cú pháp tương ứng được báo hiệu. Ví dụ, phần tử cú pháp bắt đầu bằng “sps_” có thể nghĩa là phần tử cú pháp tương ứng được báo hiệu trong tập tham số trình tự (SPS). Thêm vào đó, phần tử cú pháp bắt đầu bằng “pps_”, “ph_”, “sh_” có thể nghĩa là phần tử cú pháp tương ứng được báo hiệu lần lượt trong tập tham số hình ảnh (PPS), tiêu đề hình ảnh, và tiêu đề lát.

Trong khi các phương pháp làm ví dụ của sáng chế được mô tả trên đây được biểu diễn như chuỗi các hoạt động nhằm làm rõ phần mô tả, điều này không được dự định để giới hạn thứ tự theo đó các bước được thực hiện, và các bước có thể được thực hiện đồng thời hoặc theo thứ tự khác nếu cần. Để thi hành phương pháp theo sáng chế, các bước được mô tả có thể còn bao gồm các bước khác, có thể bao gồm các bước còn lại ngoại trừ một số bước, hoặc có thể bao gồm các bước bổ sung khác ngoại trừ một số bước.

Trong sáng chế này, thiết bị mã hóa ảnh hoặc thiết bị giải mã ảnh mà nó thực hiện hoạt động (bước) được xác định trước có thể thực hiện hoạt động (bước) để khẳng định điều kiện hoặc tình huống thực hiện hoạt động (bước) tương ứng. Ví dụ, nếu có được mô tả rằng hoạt động được xác định trước được thực hiện khi điều kiện được xác định trước được thỏa mãn thì thiết bị mã hóa ảnh hoặc thiết bị giải mã ảnh có thể thực hiện hoạt động được xác định trước này sau khi xác định xem liệu điều kiện được xác định trước này có được thỏa mãn hay không.

Các phương án khác nhau của sáng chế không phải là danh sách của tất cả những sự kết hợp khả dĩ và được nhằm để mô tả các khía cạnh đại diện của sáng chế, và các vấn đề được mô tả theo các phương án khác nhau này có thể được áp dụng theo cách độc lập hoặc theo sự kết hợp của hai hay hơn hai phương án.

Các phương án khác nhau của sáng chế có thể được thi hành trong phần cứng, phần sụn, phần mềm hoặc sự kết hợp của chúng. Trong trường hợp thi hành sáng chế bằng phần cứng, sáng chế có thể được thi hành bằng các mạch tích hợp chuyên dụng (application specific integrated circuit, ASIC), các bộ xử lý tín hiệu số (digital signal processor, DSP), các thiết bị xử lý tín hiệu số (digital signal processing device, DSPD), các thiết bị logic lập trình được (programmable logic device, PLD), các mảng công lập

trình được dạng trường (field programmable gate array, FPGA), các bộ xử lý chung, các bộ điều khiển, các bộ vi điều khiển, các bộ vi xử lý, v.v.

Thêm vào đó, thiết bị giải mã ảnh và thiết bị mã hóa ảnh, mà các phương án của sáng chế được áp dụng với chúng, có thể được chứa trong thiết bị truyền và nhận phát quảng bá đa phương tiện, thiết bị đầu cuối truyền thông di động, thiết bị video rạp chiếu phim gia đình, thiết bị video rạp chiếu phim số, camera giám sát, thiết bị trò chuyện video, thiết bị truyền thông theo thời gian thực chẳng hạn như truyền thông video, thiết bị tạo luồng di động, phương tiện lưu trữ, máy quay video, thiết bị cung cấp dịch vụ video theo yêu cầu (Video on demand, VoD), thiết bị video trên cùng (OTT, over the top), thiết bị cung cấp dịch vụ tạo luồng Internet, thiết bị video ba chiều (three-dimensional, 3D), thiết bị video điện thoại video, thiết bị video y tế, và loại tương tự, và có thể được sử dụng để xử lý các tín hiệu video hoặc các tín hiệu dữ liệu. Ví dụ, các thiết bị video OTT có thể bao gồm bảng điều khiển trò chơi, máy đọc đĩa Blu-ray, TV truy cập Internet, hệ thống rạp hát gia đình, điện thoại thông minh, PC dạng bảng, đầu ghi video số (Digital video recorder, DVR), hoặc loại tương tự.

Fig.17 là hình minh họa hệ thống tạo luồng nội dung, mà phương án của sáng chế có thể áp dụng được với nó.

Như được thể hiện trên Fig.17, hệ thống tạo luồng nội dung mà phương án của sáng chế này được áp dụng với nó, có thể phần lớn bao gồm máy chủ mã hóa, máy chủ tạo luồng, máy chủ web, bộ lưu trữ phương tiện, thiết bị người dùng, và thiết bị đầu vào đa phương tiện.

Máy chủ mã hóa nén nội dung được nhập vào từ các thiết bị đầu vào đa phương tiện như điện thoại thông minh, camera, máy quay video, v.v. thành dữ liệu số để tạo ra luồng bit và truyền luồng bit này tới máy chủ tạo luồng. Như là một ví dụ khác, khi các thiết bị đầu vào đa phương tiện như điện thoại thông minh, camera, máy quay video, v.v. trực tiếp tạo ra luồng bit, máy chủ mã hóa có thể được lược bỏ.

Luồng bit có thể được tạo ra bởi phương pháp mã hóa ảnh hoặc thiết bị mã hóa ảnh mà phương án của sáng chế này được áp dụng với chúng, và máy chủ tạo luồng có thể lưu trữ tạm thời luồng bit trong quy trình truyền hoặc nhận luồng bit.

Máy chủ tạo luồng truyền dữ liệu đa phương tiện tới thiết bị người dùng dựa trên yêu cầu của người dùng qua máy chủ web, và máy chủ web đóng vai trò như là phương

tiện để thông báo cho người dùng về dịch vụ. Khi người dùng yêu cầu dịch vụ mong muốn từ máy chủ web, máy chủ web có thể phân phát nó đến máy chủ tạo luồng, và máy chủ tạo luồng có thể truyền dữ liệu đa phương tiện đến người dùng. Trong trường hợp này, hệ thống tạo luồng nội dung có thể bao gồm máy chủ điều khiển riêng biệt. Trong trường hợp này, máy chủ điều khiển này đóng vai trò điều khiển mệnh lệnh/phản hồi giữa các thiết bị trong hệ thống tạo luồng nội dung.

Máy chủ tạo luồng có thể nhận nội dung từ bộ lưu trữ phương tiện và/hoặc máy chủ mã hóa. Ví dụ, khi nội dung được nhận từ máy chủ mã hóa, nội dung này có thể được nhận theo thời gian thực. Trong trường hợp này, để cung cấp dịch vụ tạo luồng tron tru, máy chủ tạo luồng có thể lưu trữ luồng bit trong thời gian được xác định trước.

Các ví dụ về thiết bị người dùng có thể bao gồm điện thoại di động, điện thoại thông minh, máy tính xách tay, thiết bị đầu cuối phát quảng bá kỹ thuật số, thiết bị trợ giúp số cá nhân (personal digital assistant, PDA), thiết bị trình diễn đa phương tiện cầm tay (portable multimedia player, PMP), thiết bị định vị, PC dạng phiên, các PC dạng bảng, các máy tính siêu di động (ultrabook), các thiết bị đeo được (ví dụ, các đồng hồ thông minh, các kính thông minh, các thiết bị hiển thị gắn trên đầu), các TV kỹ thuật số, các máy tính để bàn, bảng ký hiệu kỹ thuật số, và dạng tương tự.

Mỗi máy chủ trong hệ thống tạo luồng nội dung có thể được vận hành dưới dạng máy chủ phân tán, mà trong trường hợp đó dữ liệu được nhận từ mỗi máy chủ có thể được phân tán.

Phạm vi của sáng chế bao gồm phần mềm hoặc các mệnh lệnh có thể thực thi được bằng máy (ví dụ, hệ điều hành, ứng dụng, phần sụn, chương trình, v.v.) để cho phép các hoạt động theo các phương pháp của các phương án khác nhau được thực thi trên thiết bị hoặc máy tính, phương tiện phi chuyên tiếp đọc được bằng máy tính có phần mềm hoặc các mệnh lệnh như vậy được lưu trữ trên đó và có thể thực thi được trên thiết bị hoặc máy tính này.

Khả năng ứng dụng công nghiệp

Các phương án của sáng chế có thể được sử dụng để mã hóa hoặc giải mã ảnh.

YÊU CẦU BẢO HỘ

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

thu, từ luồng bit, thông tin loại đơn vị lớp trừu tượng mạng (network abstraction layer, NAL) lớp lập mã video (video coding layer, VCL) của hình ảnh hiện tại và thông tin cờ thứ nhất chỉ rõ liệu hình ảnh con được chứa trong hình ảnh hiện tại có được xử lý như một hình ảnh hay không;

xác định loại đơn vị NAL của mỗi lát trong số nhiều lát được chứa trong hình ảnh hiện tại, dựa trên thông tin loại đơn vị NAL VCL thu được; và

giải mã nhiều lát này dựa trên loại đơn vị NAL được xác định và thông tin cờ thứ nhất,

trong đó hình ảnh hiện tại bao gồm hai hay hơn hai hình ảnh con dựa trên việc ít nhất một số lát trong số nhiều lát đã nêu đang có các loại đơn vị NAL khác nhau, và

trong đó thông tin cờ thứ nhất có giá trị được xác định trước chỉ rõ rằng mỗi hình ảnh con trong số hai hay hơn hai hình ảnh con này được xử lý như một hình ảnh dựa trên việc ít nhất một số lát trong số nhiều lát đã nêu đang có các loại đơn vị NAL khác nhau,

trong đó tất cả các lát trong mỗi hình ảnh con được chứa trong hình ảnh hiện tại đều có cùng loại đơn vị NAL.

2. Phương pháp giải mã ảnh theo điểm 1, trong đó việc liệu ít nhất một số lát trong số nhiều lát đã nêu có hay không có các loại đơn vị NAL khác nhau được xác định dựa trên thông tin cờ thứ hai được thu từ tập tham số hình ảnh (picture parameter set, PPS).

3. Phương pháp giải mã ảnh theo điểm 2, trong đó thông tin cờ thứ hai có giá trị được xác định trước chỉ rõ rằng tất cả các lát trong số nhiều lát đã nêu đều có cùng loại đơn vị NAL, dựa trên việc loại đơn vị NAL được trộn đang là không áp dụng được cho tất cả các hình ảnh trong tập lớp đầu ra (output layer set, OLS).

4. Phương pháp giải mã ảnh theo điểm 2, trong đó thông tin cờ thứ hai có giá trị được xác định trước chỉ rõ rằng tất cả các lát trong số nhiều lát đã nêu đều có cùng loại đơn vị NAL, dựa trên việc hình ảnh hiện tại chứa chỉ một hình ảnh con.

5. Phương pháp giải mã ảnh theo điểm 2, trong đó thông tin cờ thứ hai có giá trị được

xác định trước chỉ rõ rằng tất cả các lát trong số nhiều lát đã nêu đều có cùng loại đơn vị NAL, dựa trên việc ít nhất một hình ảnh con trong số các hình ảnh con được chứa trong hình ảnh hiện tại đang không được xử lý như một hình ảnh.

6. Phương pháp giải mã ảnh theo điểm 1, trong đó loại hình ảnh con của mỗi hình ảnh con được chứa trong hình ảnh hiện tại được xác định dựa trên loại đơn vị NAL của mỗi hình ảnh con.

7. Phương pháp mã hóa ảnh được thực hiện bởi thiết bị mã hóa ảnh, phương pháp mã hóa ảnh này bao gồm các bước:

phân vùng hình ảnh hiện tại thành một hoặc nhiều hình ảnh con;

xác định loại đơn vị lớp trừu tượng mạng (NAL) của mỗi lát trong số nhiều lát được chứa trong một hoặc nhiều hình ảnh con này; và

mã hóa nhiều lát này dựa trên loại đơn vị NAL được xác định,

trong đó hình ảnh hiện tại được phân vùng thành hai hay hơn hai hình ảnh con dựa trên việc ít nhất một số lát trong số nhiều lát đã nêu đang có các loại đơn vị NAL khác nhau, và

trong đó mỗi hình ảnh con trong số hai hay hơn hai hình ảnh con này được xử lý như một hình ảnh dựa trên việc ít nhất một số lát trong số nhiều lát đã nêu đang có các loại đơn vị NAL khác nhau,

trong đó tất cả các lát trong mỗi hình ảnh con được chứa trong hình ảnh hiện tại đều có cùng loại đơn vị NAL.

8. Phương pháp mã hóa ảnh theo điểm 7, trong đó việc liệu ít nhất một số lát trong số nhiều lát đã nêu có hay không có các loại đơn vị NAL khác nhau được mã hóa dựa trên thông tin cờ được xác định trước được chứa trong tập tham số hình ảnh (PPS).

9. Phương pháp mã hóa ảnh theo điểm 8, trong đó thông tin cờ đã nêu có giá trị được xác định trước chỉ rõ rằng tất cả các lát trong số nhiều lát đã nêu đều có cùng loại đơn vị NAL, dựa trên việc loại đơn vị NAL được trộn đang là không áp dụng được cho tất cả các hình ảnh trong tập lớp đầu ra (OLS).

10. Phương pháp mã hóa ảnh theo điểm 8, trong đó thông tin cờ đã nêu có giá trị được xác định trước chỉ rõ rằng tất cả các lát trong số nhiều lát đã nêu đều có cùng loại đơn vị NAL, dựa trên việc hình ảnh hiện tại chứa chỉ một hình ảnh con.

11. Phương pháp mã hóa ảnh theo điểm 8, trong đó thông tin cờ đã nêu có giá trị được

xác định trước chỉ rõ rằng tất cả các lát trong số nhiều lát đã nêu đều có cùng loại đơn vị NAL, dựa trên việc ít nhất một hình ảnh con trong số các hình ảnh con được chứa trong hình ảnh hiện tại đang không được xử lý như một hình ảnh.

12. Phương pháp truyền luồng bit, phương pháp này bao gồm các bước:

thu luồng bit được tạo ra bởi phương pháp mã hóa ảnh; và

truyền luồng bit này,

trong đó phương pháp mã hóa ảnh này bao gồm các bước:

phân vùng hình ảnh hiện tại thành một hoặc nhiều hình ảnh con;

xác định loại đơn vị lớp trừu tượng mạng (NAL) của mỗi lát trong số nhiều lát được chứa trong một hoặc nhiều hình ảnh con này; và

mã hóa nhiều lát này dựa trên loại đơn vị NAL được xác định,

trong đó hình ảnh hiện tại được phân vùng thành hai hay hơn hai hình ảnh con dựa trên việc ít nhất một số lát trong số nhiều lát đã nêu đang có các loại đơn vị NAL khác nhau, và

trong đó mỗi hình ảnh con trong số hai hay hơn hai hình ảnh con này được xử lý như một hình ảnh dựa trên việc ít nhất một số lát trong số nhiều lát đã nêu đang có các loại đơn vị NAL khác nhau,

trong đó tất cả các lát trong mỗi hình ảnh con được chứa trong hình ảnh hiện tại đều có cùng loại đơn vị NAL.

FIG.1

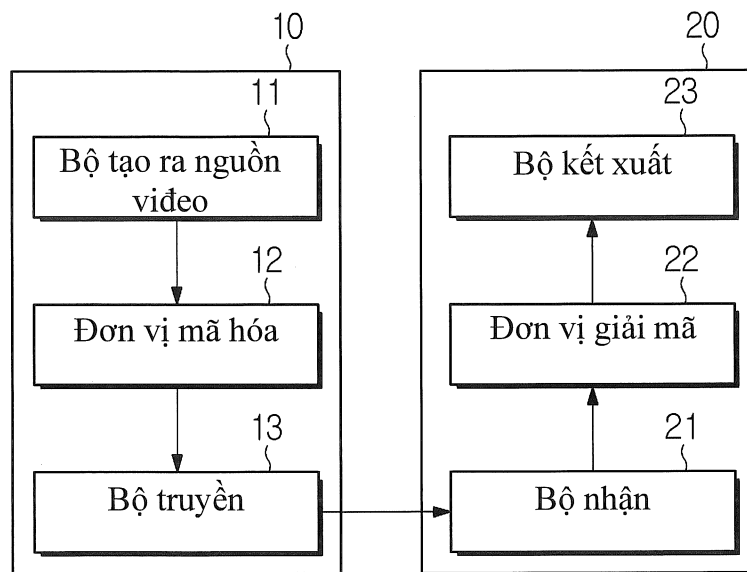


FIG. 2

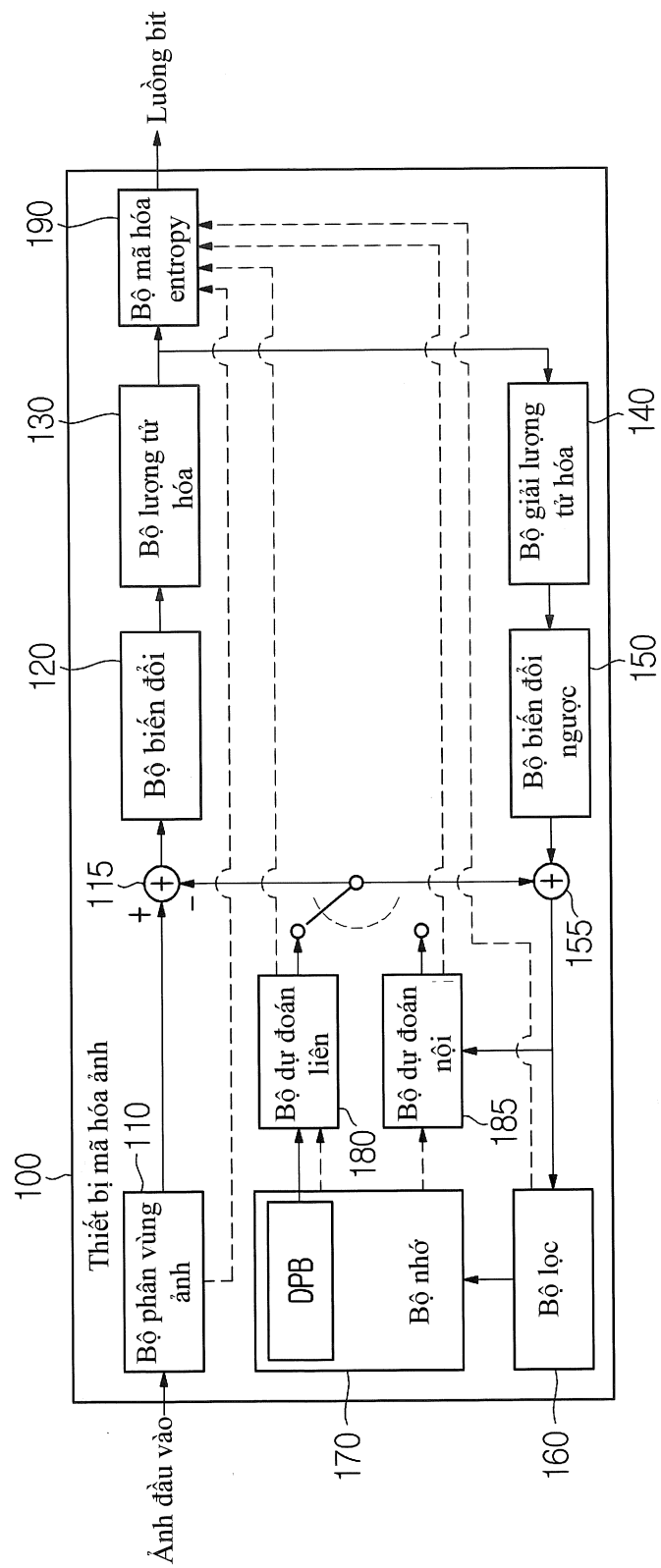


FIG.3

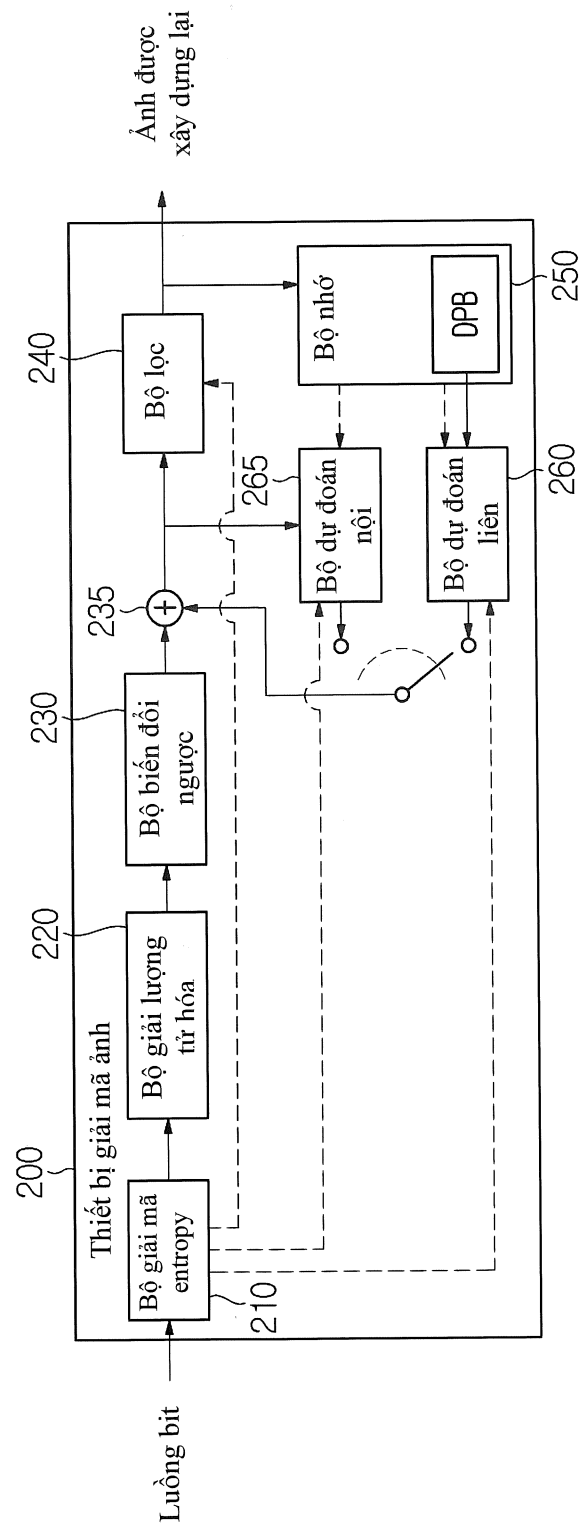


FIG.4

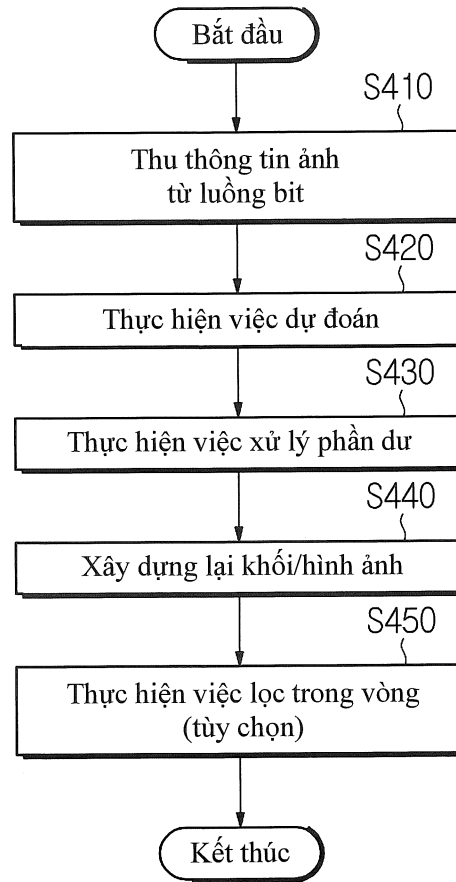


FIG.5

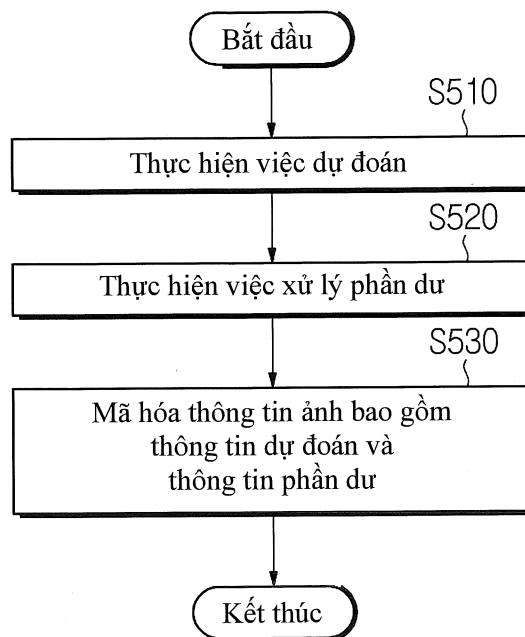


FIG.6

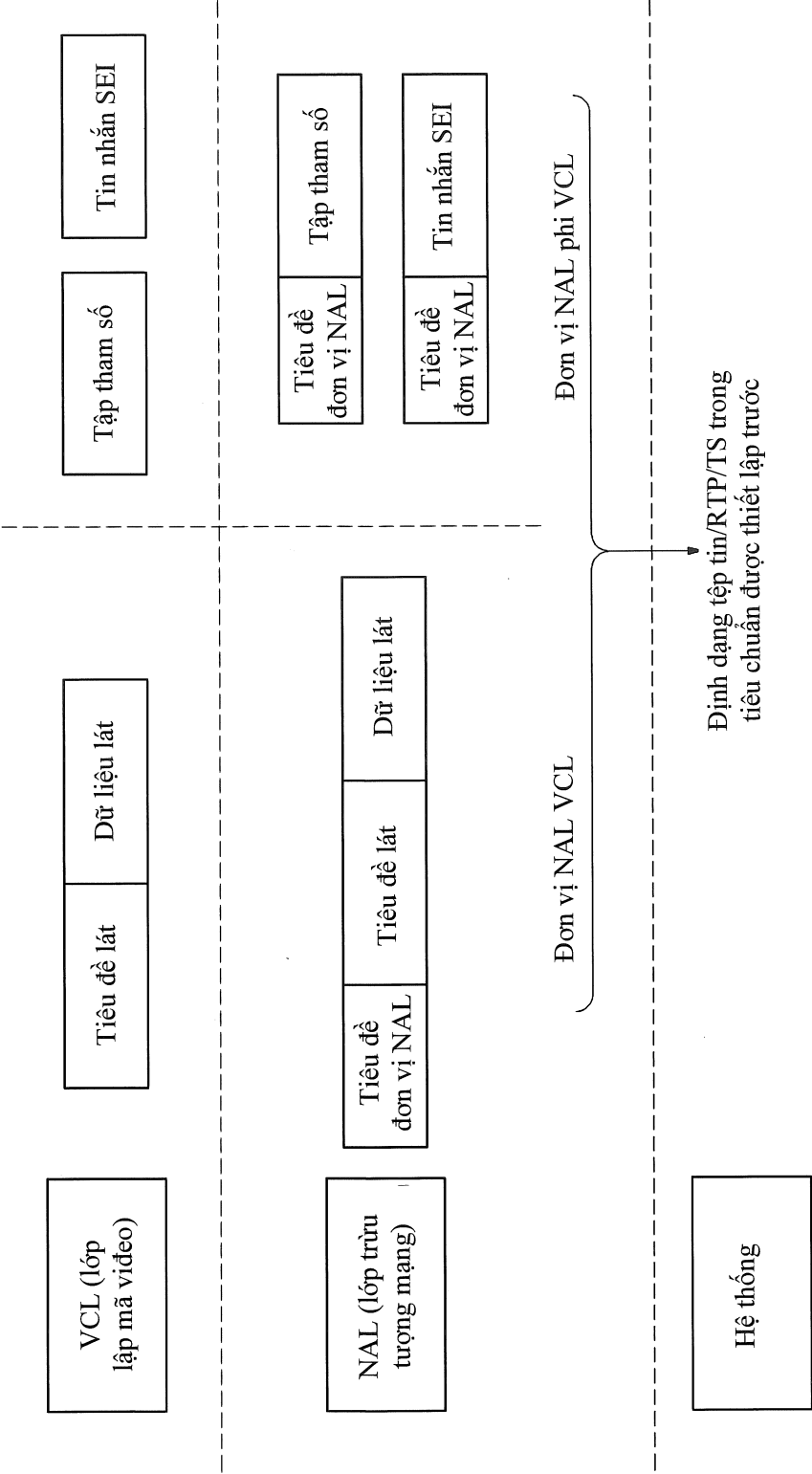


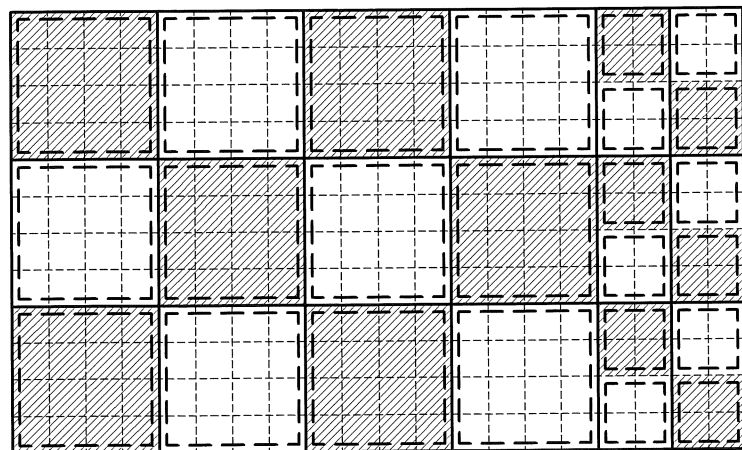
FIG.7

pic parameter set rbsp() {	Descriptor
...	
no pic partition flag	u(1)
if(!no pic partition flag) {	
...	
}	
entropy coding sync enabled flag	u(1)
if(!no pic partition flag entropy coding sync enabled flag)	
entry point offsets present flag	u(1)
...	u(1)
}	

FIG.8

slice_header() {	Descriptor
...	
if(entry point offsets present flag && NumEntryPoints > 0) {	
offset_len_minus1	ue(v)
for(i = 0; i < NumEntryPoints; i++)	
entry_point_offset_minus1[i]	u(v)
}	
...	
}	

FIG.9



CTU

Ô xép

Hình ảnh con/lát

FIG.10

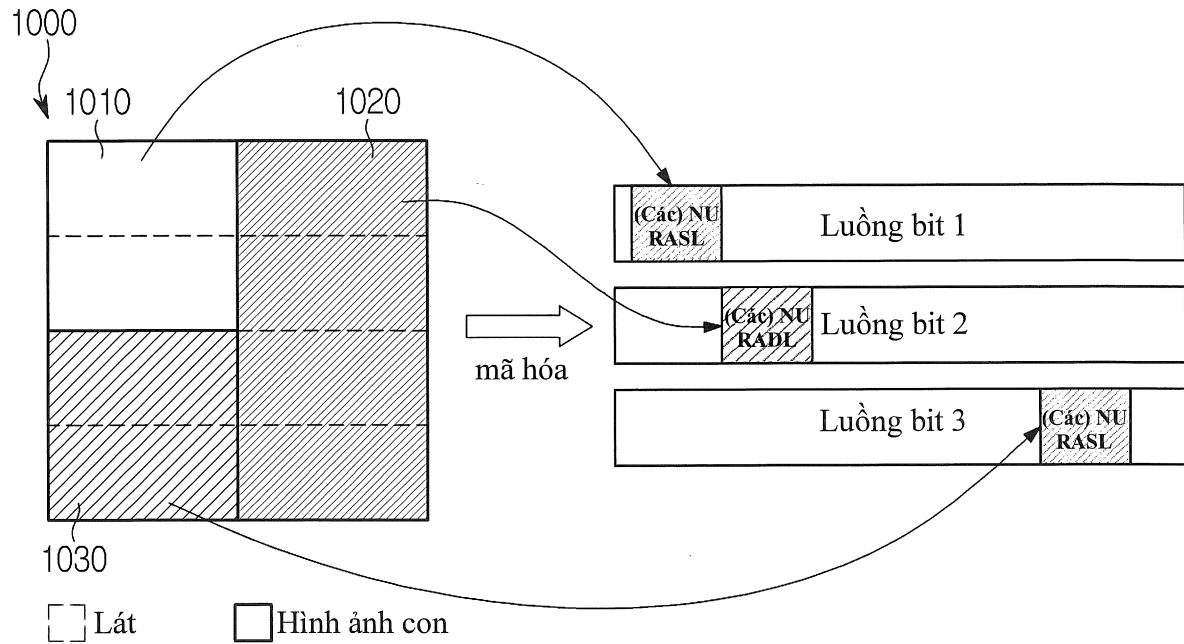


FIG.11

pic parameter set rbsp() {	Descriptor
...	
pps mixed nalu types in pic flag	
...	
pps no pic partition flag	u(1)
pps subpic id mapping present flag	u(1)
if(pps subpic id mapping present flag) {	
if(!pps no pic partition flag)	
pps num subpics minus1	ue(v)
pps subpic id len minus1	ue(v)
for(i=0; i <= pps num subpics minus1; i++)	
pps subpic id[i]	u(v)
}	
...	

FIG.12

seq parameter set rbsp() {	Descriptor
...	
sps_subpic_info_present_flag	u(1)
if(sps_subpic_info_present_flag) {	
sps_num_subpics_minus1	ue(v)
if(sps_num_subpics_minus1 > 0) {	
sps_independent_subpics_flag	u(1)
sps_subpic_same_size_flag	u(1)
}	
...	
if(!sps_independent_subpics_flag) {	
sps_subpic_treated_as_pic_flag[i]	u(1)
...	

FIG.13

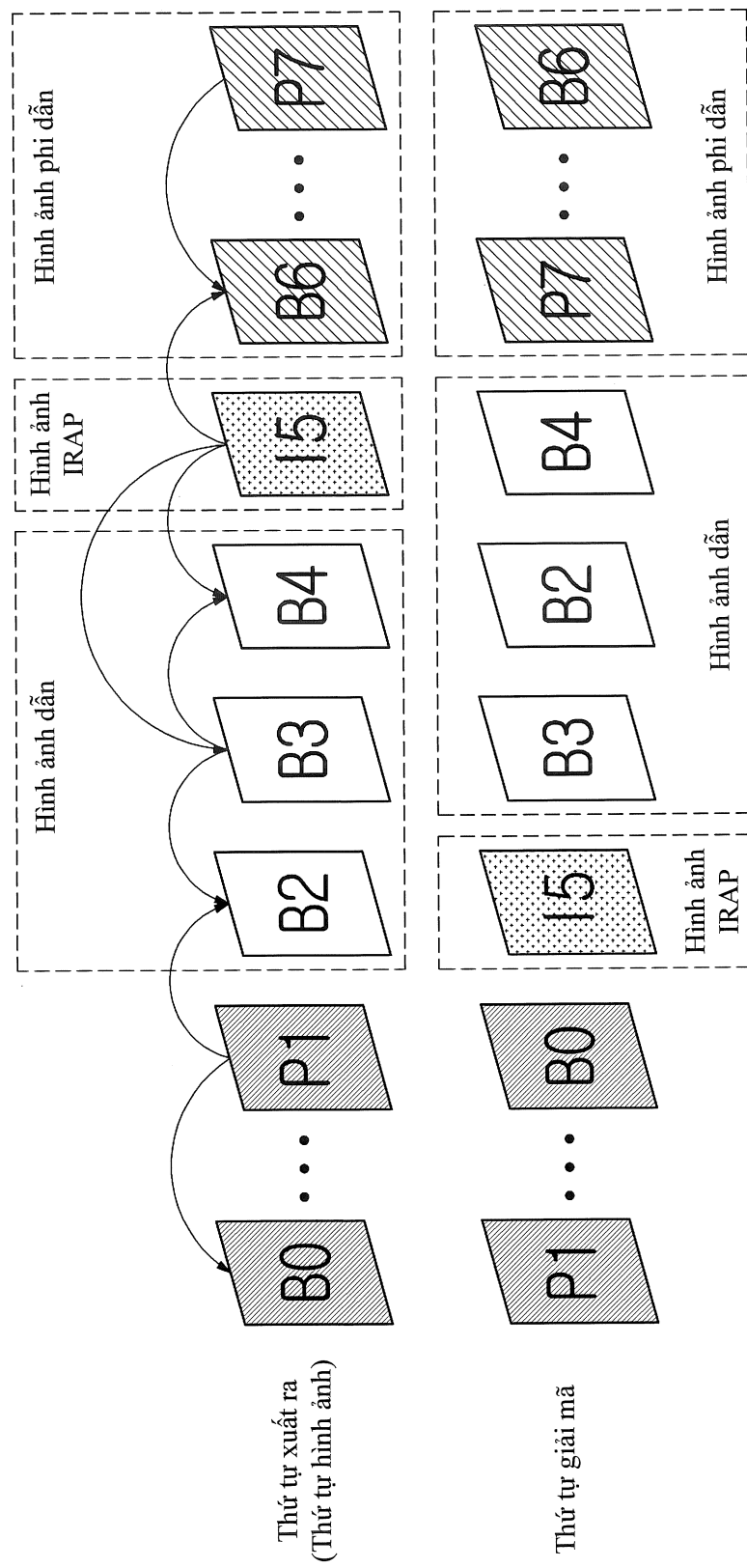


FIG.14

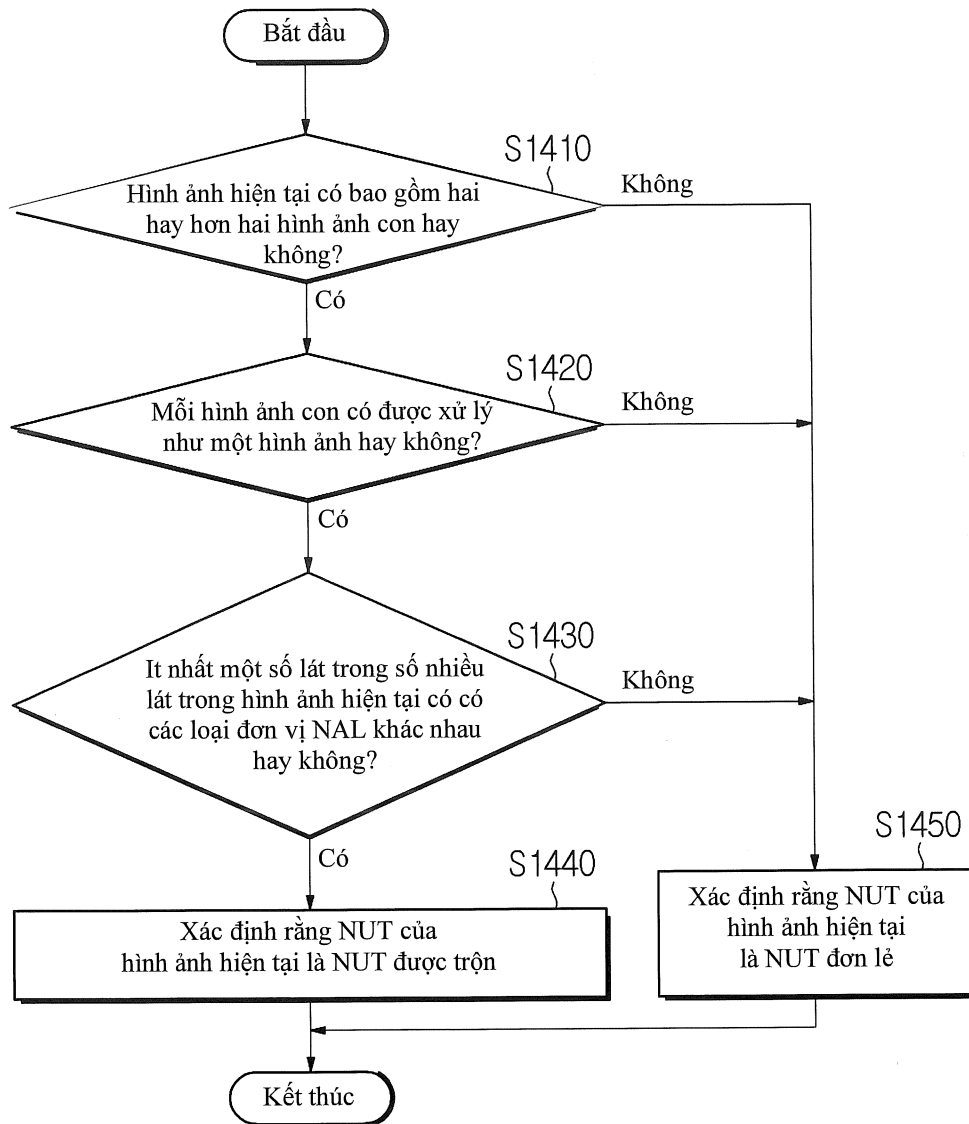


FIG.15

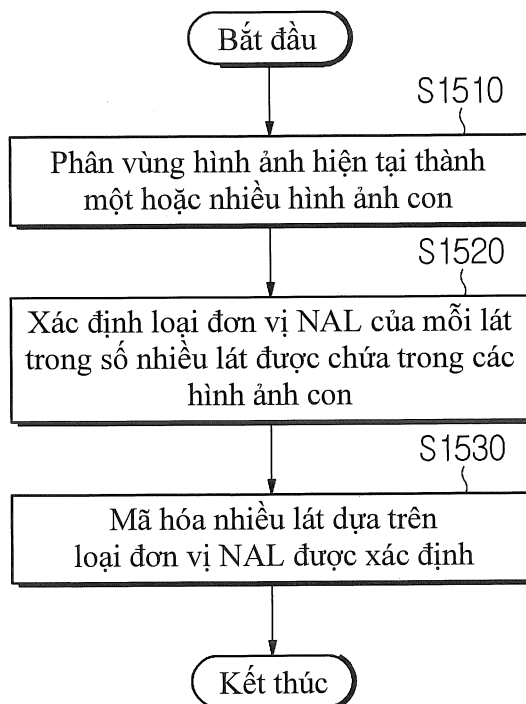


FIG.16

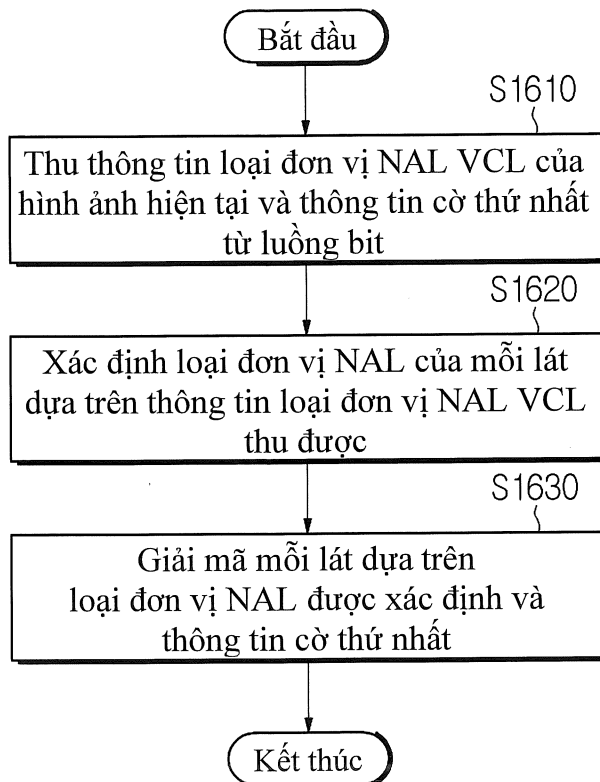


FIG.17

