



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

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

(11)



1-0023701

(51)⁷ H04N 7/36; H04N 7/26

(13) B

(21) 1-2015-00048

(22) 14/05/2013

(86) PCT/US2013/040938 14/05/2013

(87) WO2013/184305 12/12/2013

(30) 61/656,877 07/06/2012 US; 13/828,354 14/03/2013 US

(45) 25/05/2020 386

(43) 25/05/2015 326A

(73) Qualcomm Incorporated (US)

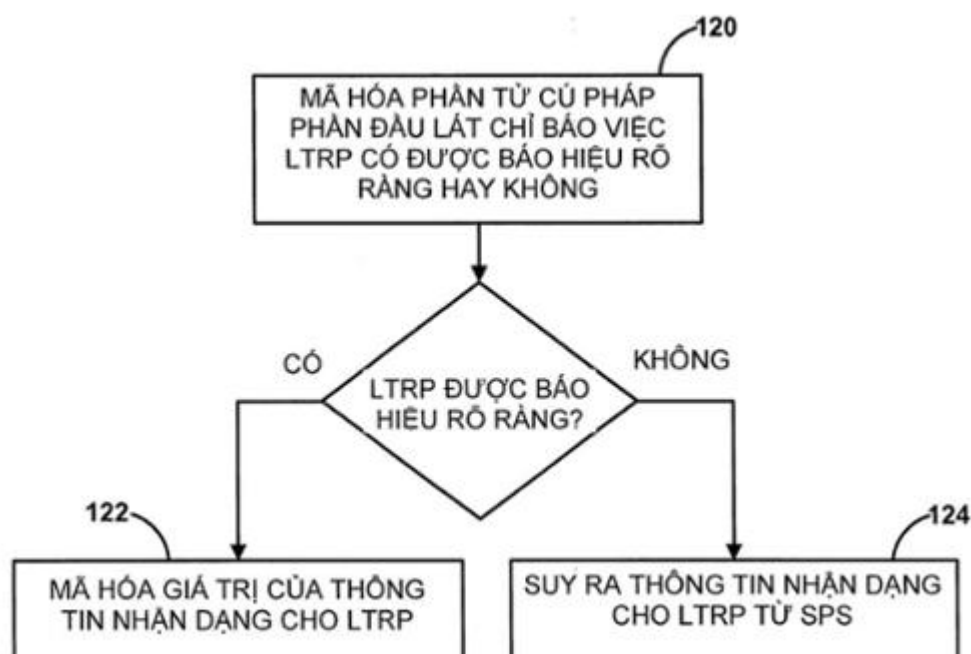
5775 Morehouse Drive, San Diego, California 92121, United States of America

(72) RAMASUBRAMONIAN, Adarsh Krishnan (IN); WANG, Ye-Kui (CN); CHEN, Ying (CN)

(74) Công ty TNHH Quốc tế D & N (D&N INTERNATIONAL CO.,LTD.)

(54) PHƯƠNG PHÁP MÃ HÓA VÀ GIẢI MÃ DỮ LIỆU VIDEO, VÀ THIẾT BỊ MÃ HÓA DỮ LIỆU VIDEO

(57) Sáng chế đề cập đến thiết bị mã hóa dữ liệu video, phương pháp mã hóa và giải mã dữ liệu video. Thiết bị mã hóa dữ liệu video mã hóa tiêu đề lát cho lát dữ liệu video. Tiêu đề lát bao gồm phần tử cú pháp chứa thông tin nhận dạng cho hình ảnh tham chiếu dài hạn, trong đó thông tin nhận dạng này được báo hiệu chính xác trong tiêu đề lát hoặc thu được từ bộ tham số trình tự tương ứng với lát này. Khi phần tử cú pháp này chỉ ra rằng thông tin nhận dạng cho hình ảnh tham chiếu dài hạn được báo hiệu chính xác, để mã hóa tiêu đề lát, thiết bị mã hóa video được định cấu hình thêm để mã hóa trị số thông tin nhận dạng cho hình ảnh tham chiếu dài hạn trong tiêu đề lát.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến mã hóa video, cụ thể là phương pháp và thiết bị mã hóa và giải mã dữ liệu video.

Tình trạng kỹ thuật của sáng chế

Khả năng video số có thể được đưa vào nhiều loại thiết bị, bao gồm truyền hình số, hệ thống phát rộng số trực tiếp, hệ thống phát rộng không dây, thiết bị số hỗ trợ cá nhân (personal digital assistant: PDA), máy tính xách tay hoặc máy tính để bàn, máy tính bảng, máy đọc sách điện tử, camera số, thiết bị ghi âm số, trình phát phương tiện số, thiết bị trò chơi video, bàn điều khiển trò chơi video, điện thoại di động hoặc điện thoại vô tuyến vệ tinh, còn gọi là “điện thoại thông minh”, thiết bị hội nghị truyền hình từ xa, thiết bị truyền liên tục video, và tương tự. Các thiết bị video số thực hiện các kỹ thuật mã hóa video, như các kỹ thuật được mô tả trong các chuẩn được xác định bởi MPEG-2, MPEG-4, ITU-T H.263, ITU-T H.264/MPEG-4, Phần 10, chuẩn mã hóa video tiên tiến (Advanced Video Coding: AVC), chuẩn mã hóa video hiệu quả cao (High Efficiency Video Coding: HEVC) hiện đang được phát triển, và các phiên bản mở rộng của các chuẩn này như mã hóa video mở rộng được (scalable video coding: SVC) và mã hóa video đa khung hình (multiview video coding: MVC). Các thiết bị video này có thể truyền, thu, mã hóa, giải mã và/hoặc lưu trữ thông tin video số hiệu quả hơn bằng cách thực hiện các kỹ thuật mã hóa video này.

Dự thảo hiện nay về chuẩn HEVC sắp tới, được gọi là “Dự thảo làm việc HEVC 6” hoặc “WD6” được mô tả trong tài liệu JCTVC-H1003, Bross và các cộng sự, “High efficiency video coding (HEVC) text specification draft 6”, Nhóm hợp tác chung về mã hóa video (Joint Collaborative Team on Video Coding: JCT-VC) của ITU-T SG16 WP3 và ISO/IEC JTC1/SC29/WG11, Hội nghị lần thứ 8: San Jose, California, USA, tháng 2/2012, kể từ ngày 07/6/2012 có thể tải về từ http://phenix.int-evry.fr/jct/doc_end_user/documents/8_San_Jose/wg11/JCTVC-H1003-v22.zip. Dự thảo khác về chuẩn HEVC sắp tới, còn được gọi là “Dự thảo làm việc 7” hoặc “WD7”, được mô tả trong tài liệu JCTVC-H1003, Bross và các cộng sự, “High Efficiency Video Coding (HEVC) Text specification Draft 7”, Nhóm hợp tác chung về mã hóa video

(JCT-VC) của ITU-T SG16 WP3 và ISO/IEC JTC1/SC29/WG11, Hội nghị lần thứ 9: Geneva, Thụy Sĩ, từ 27/4/2012 đến 07/5/2012, kể từ ngày 07/6/2012 có thể tải về từ http://phenix.it-sudparis.eu/jct/doc_end_user/documents/9_Geneva/wg11/JCTVC-I1003-v3.zip.

Các kỹ thuật mã hóa video bao gồm dự báo không gian (nội ảnh) và/hoặc dự báo thời gian (liên ảnh) để làm giảm hoặc loại bỏ dư thừa vốn có trong chuỗi video. Để mã hóa video dựa trên khối, lát video (ví dụ, khung video hoặc một phần của khung video) có thể được chia thành các khối video, các khối này còn có thể được gọi là khối cây, đơn vị cây mã hóa, đơn vị mã hóa (coding unit: CU) và/hoặc nút mã hóa. Các khối video trong lát mã hóa nội cấu trúc (I) của hình ảnh được mã hóa bằng cách sử dụng dự báo không gian đối với các mẫu tham chiếu trong các khối lân cận trong cùng hình ảnh. Các khối video trong lát mã hóa liên cấu trúc (P hoặc B) của hình ảnh có thể sử dụng dự báo không gian đối với các mẫu tham chiếu trong các khối lân cận trong cùng hình ảnh n hoặc dự báo thời gian đối với các mẫu tham chiếu trong các hình ảnh tham chiếu khác. Các hình ảnh có thể được gọi là các khung, và hình ảnh tham chiếu có thể được gọi là khung tham chiếu.

Dự báo không gian hoặc thời gian tạo ra khối dự báo cho khối cần được mã hóa. Dữ liệu dư biểu diễn chênh lệch điểm ảnh giữa khối ban đầu cần được mã hóa và khối dự báo. Khối mã hóa liên cấu trúc được mã hóa theo vectơ chuyển động mà trở đến khối các mẫu tham chiếu tạo ra khối dự báo, và dữ liệu dư chỉ báo chênh lệch giữa khối được mã hóa và khối dự báo. Khối mã hóa nội cấu trúc được mã hóa theo chế độ mã hóa nội cấu trúc và dữ liệu dư. Để nén thêm, dữ liệu dư có thể được biến đổi từ miền điểm ảnh sang miền biến đổi, tạo ra các hệ số biến đổi dư, mà sau đó có thể được lượng tử hóa. Các hệ số biến đổi được lượng tử hóa, ban đầu được sắp xếp trong mảng 2 chiều, có thể được quét để tạo ra vectơ một chiều của các hệ số biến đổi, và kỹ thuật mã hóa entropy có thể được áp dụng để nén được nhiều hơn nữa.

Bản chất kỹ thuật của sáng chế

Nói chung, sáng chế mô tả các kỹ thuật liên quan đến việc suy ra tập hợp hình ảnh tham chiếu (reference picture set: RPS) và báo hiệu các hình ảnh tham chiếu dài hạn (long term reference picture: LTRP) được bao gồm trong RPS này, ví dụ, trong tập hợp tham số chuỗi (sequence parameter set: SPS), tập hợp tham số hình ảnh (picture

parameter set: PPS) hoặc trong phần đầu lát. Các kỹ thuật này có thể bao gồm đảm bảo rằng giá trị của các bit có trọng số nhỏ nhất (least significant bit: LSB) của giá trị số đếm thứ tự hình ảnh (picture order count: POC) cho LTRP là không giảm hoặc không tăng. Ngoài ra hoặc theo cách khác, các kỹ thuật này có thể bao gồm báo hiệu dữ liệu cho giá trị POC của LTRP trong SPS hoặc trong phần đầu lát, và cung cấp thông tin chỉ báo việc dữ liệu cho giá trị POC được báo hiệu trong SPS hay trong phần đầu lát.

Trong một ví dụ, phương pháp mã hóa dữ liệu video bao gồm bước mã hóa phần đầu lát cho lát dữ liệu video, trong đó phần đầu lát bao gồm phần tử cú pháp chỉ báo việc thông tin nhận dạng cho hình ảnh tham chiếu dài hạn được báo hiệu rõ ràng trong phần đầu lát hay được suy ra từ tập hợp tham số chuỗi tương ứng với lát này, và khi phần tử cú pháp này chỉ báo rằng thông tin nhận dạng cho hình ảnh tham chiếu dài hạn được báo hiệu rõ ràng, bước mã hóa phần đầu lát còn bao gồm mã hóa giá trị của thông tin nhận dạng cho hình ảnh tham chiếu dài hạn trong phần đầu lát.

Trong ví dụ khác, thiết bị mã hóa dữ liệu video bao gồm phương tiện mã hóa phần đầu lát cho lát dữ liệu video, trong đó phần đầu lát bao gồm phần tử cú pháp chỉ báo việc thông tin nhận dạng cho hình ảnh tham chiếu dài hạn được báo hiệu rõ ràng trong phần đầu lát hay được suy ra từ tập hợp tham số chuỗi tương ứng với lát này, và phương tiện mã hóa giá trị của thông tin nhận dạng cho hình ảnh tham chiếu dài hạn trong phần đầu lát khi phần tử cú pháp chỉ báo rằng thông tin nhận dạng cho hình ảnh tham chiếu dài hạn được báo hiệu rõ ràng.

Trong ví dụ khác, thiết bị mã hóa dữ liệu video, thiết bị này bao gồm bộ mã hóa video có thể được tạo cấu hình để: mã hóa phần đầu lát cho lát dữ liệu video, trong đó phần đầu lát bao gồm phần tử cú pháp chỉ báo việc thông tin nhận dạng cho hình ảnh tham chiếu dài hạn được báo hiệu rõ ràng trong phần đầu lát hay được suy ra từ tập hợp tham số chuỗi tương ứng với lát này, và khi phần tử cú pháp chỉ báo rằng thông tin nhận dạng cho hình ảnh tham chiếu dài hạn được báo hiệu rõ ràng, để mã hóa phần đầu lát, thiết bị này còn được tạo cấu hình để mã hóa giá trị của thông tin nhận dạng cho hình ảnh tham chiếu dài hạn trong phần đầu lát.

Trong ví dụ khác, vật ghi bất biến đọc được bằng máy tính đã lưu trữ trên đó các lệnh mã, khi được thực hiện, khiến cho bộ xử lý của thiết bị thực hiện mã hóa video, trong đó các lệnh này khiến cho bộ xử lý: mã hóa phần đầu lát cho lát dữ liệu video, trong đó phần đầu lát bao gồm phần tử cú pháp chỉ báo việc thông tin nhận

dạng cho hình ảnh tham chiếu dài hạn được báo hiệu rõ ràng trong phần đầu lát hay được suy ra từ tập hợp tham số chuỗi tương ứng với lát này, và khi phần tử cú pháp chỉ báo rằng thông tin nhận dạng cho hình ảnh tham chiếu dài hạn được báo hiệu rõ ràng, để mã hóa phần đầu lát, thiết bị này còn được tạo cấu hình để mã hóa giá trị của thông tin nhận dạng cho hình ảnh tham chiếu dài hạn trong phần đầu lát.

Chi tiết về một hoặc nhiều ví dụ được đưa ra trong các hình vẽ kèm theo và phần mô tả dưới đây. Các đặc điểm, mục tiêu, ưu điểm khác sẽ trở nên rõ ràng từ phần mô tả và hình vẽ, và phần yêu cầu bảo hộ.

Mô tả vắn tắt các hình vẽ

Fig.1 là sơ đồ khối minh họa hệ thống mã hóa và giải mã video làm ví dụ có thể sử dụng các kỹ thuật để báo hiệu dữ liệu cho các hình ảnh tham chiếu dài hạn (LTRP) trong tập hợp tham số chuỗi (SPS) hoặc phần đầu lát.

Fig.2 là sơ đồ khối minh họa ví dụ về bộ mã hóa video có thể thực hiện các kỹ thuật để báo hiệu dữ liệu cho LTRP trong SPS hoặc phần đầu lát.

Fig.3 là sơ đồ khối minh họa ví dụ về bộ giải mã video có thể thực hiện các kỹ thuật để báo hiệu dữ liệu cho LTRP trong SPS hoặc phần đầu lát.

Fig.4 là lưu đồ minh họa quy trình mã hóa dữ liệu cho các hình ảnh tham chiếu theo một hoặc nhiều ví dụ được mô tả trong bản mô tả này.

Fig.5 là lưu đồ minh họa quy trình báo hiệu dữ liệu cho LTRP trong SPS hoặc phần đầu lát theo một hoặc nhiều ví dụ được mô tả trong bản mô tả này.

Mô tả chi tiết sáng chế

Chuỗi video thường được biểu diễn dưới dạng một chuỗi các hình ảnh. Thông thường, các kỹ thuật mã hóa theo khối được sử dụng để mã hóa mỗi trong số các hình ảnh riêng biệt này. Nghĩa là, mỗi hình ảnh được chia thành các khối, và mỗi khối được mã hóa riêng. Việc mã hóa khối dữ liệu video thường bao gồm tạo ra giá trị dự báo cho khối này và mã hóa giá trị dư, nghĩa là chênh lệch giữa khối ban đầu và giá trị được dự báo. Cụ thể là, khối dữ liệu video ban đầu bao gồm ma trận các giá trị điểm ảnh, và giá trị được dự báo bao gồm ma trận các giá trị điểm ảnh được dự báo. Giá trị dư tương ứng với chênh lệch theo từng điểm ảnh giữa các giá trị điểm ảnh của khối ban đầu và các giá trị điểm ảnh được dự báo.

Các kỹ thuật dự báo cho khối dữ liệu video thường được phân loại thành dự báo nội cấu trúc và dự báo liên cấu trúc. Dự báo nội cấu trúc, hoặc dự báo không gian, thường bao gồm dự báo khối từ các giá trị điểm ảnh của các khối lân cận đã được mã hóa trước đó. Dự báo liên cấu trúc, hoặc dự báo thời gian, thường bao gồm dự báo khối từ giá trị điểm ảnh của các hình ảnh đã được mã hóa trước đó.

Các hình ảnh đã được mã hóa trước đó có thể biểu diễn các hình ảnh được hiển thị sớm hơn hoặc muộn hơn so với hình ảnh đang được mã hóa. Nói cách khác, thứ tự hiển thị cho các hình ảnh không nhất thiết giống như thứ tự giải mã của các hình ảnh này, và vì vậy, các hình ảnh được hiển thị sớm hơn hoặc muộn hơn so với hình ảnh hiện thời đang được mã hóa có thể được sử dụng làm tham chiếu để mã hóa các khối của hình ảnh hiện thời.

Ví dụ, bộ mã hóa video (là thuật ngữ có thể chỉ một trong hai hoặc cả hai bộ mã hóa video và bộ giải mã video) bao gồm bộ đệm hình ảnh được giải mã (decoded picture buffer: DPB). DPB lưu trữ các hình ảnh tham chiếu, là các hình ảnh có thể được sử dụng để dự báo liên cấu trúc hình ảnh. Nói cách khác, bộ mã hóa video có thể dự báo hình ảnh dựa vào một hoặc nhiều hình ảnh tham chiếu được lưu trữ trong DPB. Nói chung, các kỹ thuật được mô tả cung cấp nhiều phương pháp khác nhau để suy ra tập hợp hình ảnh tham chiếu (RPS) và báo hiệu các hình ảnh tham chiếu dài hạn (LTRP) được bao gồm trong RPS của hình ảnh được mã hóa để sử dụng khi thực hiện các khía cạnh dự báo liên cấu trúc của kỹ thuật mã hóa video.

Bộ giải mã video cũng có thể được giao nhiệm vụ thiết lập danh sách hình ảnh tham chiếu chỉ báo các hình ảnh tham chiếu nào được sử dụng cho mục đích dự báo liên cấu trúc. Hai trong số các danh sách hình ảnh tham chiếu này được gọi lần lượt là Danh sách 0 và Danh sách 1. Trước hết, bộ giải mã video sử dụng các kỹ thuật thiết lập mặc định để thiết lập Danh sách 0 và Danh sách 1 (ví dụ, các sơ đồ thiết lập được tạo cấu hình trước để thiết lập Danh sách 0 và Danh sách 1). Tùy ý, sau khi Danh sách 0 và Danh sách 1 ban đầu được thiết lập, bộ giải mã có thể giải mã các phần tử cú pháp, nếu có mặt, mà lệnh cho bộ giải mã video sửa đổi Danh sách 0 và Danh sách 1 ban đầu.

Bộ mã hóa video có thể báo hiệu các phần tử cú pháp chỉ báo các ký hiệu nhận dạng của các hình ảnh tham chiếu trong DPB, và bộ mã hóa video cũng có thể báo hiệu các phần tử cú pháp bao gồm các chỉ số, trong Danh sách 0, Danh sách 1, hoặc cả

hai Danh sách 0 và Danh sách 1, chỉ báo hình ảnh tham chiếu hoặc các hình ảnh tham chiếu nào sử dụng để giải mã khối đã được mã hóa của hình ảnh hiện thời. Bộ giải mã video sử dụng kí hiệu nhận dạng thu được để nhận dạng giá trị chỉ số hoặc các giá trị chỉ số cho hình ảnh tham chiếu hoặc các hình ảnh tham chiếu được liệt kê trong Danh sách 0, Danh sách 1, hoặc cả hai Danh sách 0 và Danh sách 1. Từ (các) giá trị chỉ số cũng như (các) ký hiệu nhận dạng của hình ảnh tham chiếu hoặc các hình ảnh tham chiếu, bộ mã hóa video khôi phục lại hình ảnh tham chiếu hoặc các hình ảnh tham chiếu từ DPB, và giải mã khối đã được mã hóa của hình ảnh hiện thời.

Trong tập hợp tham số hình ảnh (PPS) hoặc tập hợp tham số chuỗi (SPS) gắn với phần đầu lát của hình ảnh hiện thời, bộ mã hóa video có thể báo hiệu RPS. RPS của hình ảnh hiện thời bao gồm thông tin nhận dạng cho các hình ảnh tham chiếu có thể được sử dụng để dự báo hình ảnh hiện thời và các hình ảnh có thể được sử dụng để dự báo các hình ảnh sau hình ảnh hiện thời theo thứ tự giải mã. Chỉ các hình ảnh tham chiếu trong RPS có thể được bao gồm trong Danh sách 0 hoặc Danh sách 1.

Trong dòng bit video được mã hóa, bộ mã hóa video (nghĩa là, bộ mã hóa video hoặc bộ giải mã video) có thể sử dụng hai loại giá trị để tham chiếu đến các hình ảnh: giá trị số đếm thứ tự hình ảnh (POC), thường tương ứng với thứ tự xuất ra của các hình ảnh này, và giá trị số khung (frame num), thường tương ứng với thứ tự giải mã của các hình ảnh. Nói chung, thuật ngữ “khung” và “hình ảnh” có thể được sử dụng thay thế nhau. Do đó, giá trị số khung tương ứng với thứ tự giải mã của các hình ảnh.

Thông tin nhận dạng cho các hình ảnh tham chiếu có thể bao gồm một hoặc nhiều giá trị số đếm thứ tự hình ảnh (POC). Giá trị POC chỉ báo thứ tự trong đó các hình ảnh trong chuỗi video được mã hóa được xuất ra hoặc hiển thị (nghĩa là, thứ tự hiển thị của các hình ảnh). Ví dụ, hình ảnh có giá trị POC thấp hơn được hiển thị sớm hơn hình ảnh có giá trị POC cao hơn trong cùng một chuỗi video được mã hóa.

Trong quy trình mã hóa video, dữ liệu cho các hình ảnh, hoặc lát của các hình ảnh, thường được bọc trong các đơn vị lớp mạng trừu tượng (Network abstraction layer: NAL). Các đơn vị NAL có thể bao gồm dữ liệu lớp mã hóa video (video coding layer: VCL), như dữ liệu được mã hóa cho các hình ảnh hoặc lát, còn được gọi là các đơn vị NAL VCL, hoặc dữ liệu không VCL, như các tập hợp tham số và các thông báo thông tin nâng cao phụ (supplemental enhancement information: SEI), còn được gọi là các đơn vị NAL không VCL. Các đơn vị NAL còn bao gồm dữ liệu phần đầu mô tả

loại dữ liệu được bao gồm trong các đơn vị NAL tương ứng. Ví dụ, các đơn vị NAL có thể bao gồm `nal_ref_flag` chỉ báo việc đơn vị NAL có bao gồm dữ liệu cho hình ảnh tham chiếu hay không. Theo cách này, hình ảnh tham chiếu có thể được định nghĩa là hình ảnh có `nal_ref_flag` bằng “1”. Hình ảnh tham chiếu có thể chứa các mẫu (nghĩa là, giá trị điểm ảnh) có thể được sử dụng để dự báo liên cấu trúc trong quá trình giải mã các hình ảnh tiếp theo theo thứ tự giải mã.

Trong HEVC, có thể có nhiều hơn một loại hình ảnh tham chiếu. Theo một ví dụ, HEVC định nghĩa hình ảnh tham chiếu dài hạn, mà có thể được định nghĩa là hình ảnh tham chiếu được đánh dấu là “được sử dụng để tham chiếu dài hạn”. Thêm vào đó, HEVC cũng có thể bao gồm các hình ảnh tham chiếu ngắn hạn, mặc dù các kỹ thuật theo sáng chế hướng tới việc báo hiệu các hình ảnh tham chiếu dài hạn. Các hình ảnh tham chiếu ngắn hạn nằm ngoài phạm vi của bản mô tả này.

HEVC mô tả các kỹ thuật báo hiệu và sử dụng các hình ảnh tham chiếu dài hạn (LTRP). Ví dụ về việc báo hiệu thông thường của LTRP được mô tả trong HEVC WD6. WD6 cũng mô tả các kỹ thuật được chọn liên quan đến báo hiệu LTRP, bao gồm các thay đổi trong việc báo hiệu các bit có trọng số nhỏ nhất (least significant bit: LSB) của LTRP, ngữ nghĩa của phần tử cú pháp liên quan đến LTRP, và việc suy ra tập hợp hình ảnh tham chiếu. Việc thực hiện dự thảo làm việc HEVC cũng bao gồm các kỹ thuật mô tả việc báo hiệu LTRP trong các tập hợp tham số chuỗi (SPS). Việc thực hiện liên quan đến báo hiệu các LTRP trong SPS được mô tả trong JCTVC-I0340r2 như được đưa vào HEVC WD7, được mô tả trong Ramasubramanian và các cộng sự, “Signaling of Long term Reference Pictures in the SPS”, tài liệu JCTVC-I0340r2, hội nghị lần thứ 9, JCT-VC của ITU-T SG16 WP3 và ISO/IEC JTC1/SC29/WG11, Geneva, Thụy Sĩ, từ 27/4/2012 đến 07/5/2012, sẵn có từ http://phenix.int-evry.fr/jct/doc_end_user/documents/9_Geneva/wg11/JCTVC-I0340-v3.zip. Chi tiết về việc báo hiệu các LTRP trong SPS được tóm tắt dưới đây.

Trong các kỹ thuật được mô tả trong JCTVC-I0340r2, SPS có thể bao gồm phần tử cú pháp chỉ báo việc các LTR có được báo hiệu (nghĩa là bộ mã hóa video có sử dụng LTRP cho dự báo liên cấu trúc hay không) cho chuỗi video được mã hóa hay không. Nếu các LTRP được báo hiệu cho chuỗi video được mã hóa, các phần tử cú pháp của SPS hoạt động hoặc phần đầu lát cho chuỗi video được mã hóa có thể tham chiếu đến các LTRP sử dụng số lượng LSB của POC cho LTRP này. Việc chỉ sử

dùng các LSB của POC để tham chiếu đến các LTRP có thể cải thiện hiệu quả mã hóa và làm giảm tính phức tạp của dòng bit video được mã hóa. Bộ mã hóa video báo hiệu số lượng LSB LTRP được sử dụng để tham chiếu mỗi LTRP trong SPS trong phần tử cú pháp thứ nhất, sau đó là danh sách các LSB của các giá trị POC của các LTRP này.

Trong chuỗi video được mã hóa, phần đầu lát có thể kế thừa hoặc tham chiếu một hoặc nhiều LTRP của SPS này. Trong phần đầu lát, bộ mã hóa video báo hiệu phần tử cú pháp danh sách bao gồm các giá trị LSB của các giá trị POC cho các LTRP cần được kế thừa từ SPS. Ngoài các LTRP được kế thừa từ SPS, phần đầu lát cũng có thể báo hiệu rõ ràng các LSB của các LTRP bổ sung mà không được kế thừa từ SPS. Phần đầu lát chứa phần tử cú pháp bao gồm danh sách các giá trị LSB của các giá trị POC cho các LTRP được báo hiệu rõ ràng trong phần đầu lát.

Trong một số trường hợp, hai LTRP có thể có cùng chuỗi LSB POC, tạo ra sự không rõ ràng là phần tử cú pháp LSB POC cụ thể tham chiếu đến LTRP nào. Để làm rõ phần tử cú pháp tham chiếu đến LTRP nào, phần đầu lát hoặc SPS có thể báo hiệu một số trong số các bit có trọng số lớn nhất (most significant bit: MSB) cho LTRP nếu nhiều hơn một hình ảnh tham chiếu trong bộ đệm hình ảnh được giải mã (DPB) có cùng LSB. Phần tử cú pháp cờ được sử dụng để chỉ báo việc các MSB cho LTRP có được báo hiệu không, và việc báo hiệu thông tin MSB cho LTRP, nếu có, kế tiếp ngay sau cờ này. Cuối cùng, bộ mã hóa video có thể mã hóa cờ để chỉ báo việc LTRP được báo hiệu có thể được sử dụng bởi hình ảnh hiện thời làm hình ảnh tham chiếu hay không. Cú pháp và ngữ nghĩa của các phần tử cú pháp liên quan trong SPS và phần đầu lát được mô tả dưới đây. Đối với các phần tử cú pháp được bao gồm trong bảng cú pháp nhưng không được cung cấp ngữ nghĩa thì ngữ nghĩa là giống như trong HEVC WD6.

JCTVC-I0340r2 cung cấp cú pháp và ngữ nghĩa của SPS được mô tả dưới đây trong Bảng 1:

Bảng 1

seq_parameter_set_rbsp() {	Mô tả
...	
num_short_term_ref_pic_sets	ue(v)
for(i = 0; i < num_short_term_ref_pic_sets; i++)	
short_term_ref_pic_set(i)	
long_term_ref_pics_present_flag	u(1)
if(long_term_ref_pics_present_flag) {	
num_long_term_ref_pics_sps	ue(v)
for(i = 0; i < num_long_term_ref_pics_sps; i++)	
lt_ref_pic_poc_lsb_sps[i]	u(v)
}	
...	
}	

Ngữ nghĩa của các phần tử cú pháp của SPS sửa đổi bởi JCTVC- I0340r2 được mô tả dưới đây:

num_long_term_ref_pics_sps xác định số lượng hình ảnh tham chiếu dài hạn được xác định trong tập hợp tham số chuỗi. Giá trị của **num_long_term_ref_pics_sps** nằm trong khoảng từ 0 đến 32, bao gồm cả các giá trị đầu mút.

lt_ref_pic_poc_lsb_sps[i] xác định các bit có trọng số nhỏ nhất của số đếm thứ tự hình ảnh của hình ảnh tham chiếu dài hạn thứ i được xác định trong tập hợp tham số chuỗi, số lượng bit được sử dụng để biểu diễn **lt_ref_pic_poc_lsb_sps[i]** bằng $\log_2_max_pic_order_cnt_lsb_minus4 + 4$.

JCTVC-I0340r2 cũng sửa đổi cú pháp và ngữ nghĩa của phần đầu lát, như được mô tả dưới đây trong Bảng 2:

Bảng 2

slice_header() {	Mô tả
...	
if(IdrPicFlag) {	
idr_pic_id	ue(v)
no_output_of_prior_pics_flag	u(1)
} else {	
pic_order_cnt_lsb	u(v)
short_term_ref_pic_set_sps_flag	u(1)
if(!short_term_ref_pic_set_sps_flag)	
short_term_ref_pic_set(num_short_term_ref_pic_sets)	
else	
short_term_ref_pic_set_idx	u(v)
if(long_term_ref_pics_present_flag) {	
num_long_term_pics	ue(v)
if(num_long_term_ref_pics_sps)	
num_long_term_sps	ue(v)
for(i = 0; i < num_long_term_sps + num_long_term_pics; i++) {	
if(i < num_long_term_sps)	
long_term_idx_sps[i]	ue(v)
else	
poc_lsb_lt[i]	u(v)
delta_poc_msb_present_flag[i]	u(1)
if(delta_poc_msb_present_flag[i])	
delta_poc_msb_cycle_lt[i]	ue(v)
used_by_curr_pic_lt_flag[i]	u(1)
}	
}	
}	
...	
}	

Ngữ nghĩa của các phần tử cú pháp của phần đầu lát sửa đổi bởi JCTVC-I0340r2 được mô tả dưới đây:

num_long_term_pics xác định số lượng hình ảnh tham chiếu dài hạn được bao gồm trong tập hợp hình ảnh tham chiếu dài hạn của hình ảnh hiện thời và được báo hiệu trực tiếp trong phần đầu lát. Giá trị của **num_long_term_pics** nằm trong khoảng từ 0 đến $\text{sps_max_dec_pic_buffering}[\text{sps_max_temporal_layers_minus1}] - \text{NumNegativePics}[\text{StRpsIdx}] - \text{NumPositivePics}[\text{StRpsIdx}] - \text{num_long_term_sps}$, bao gồm cả các giá trị đầu mút. Khi không có mặt, bộ mã hóa video suy ra giá trị của **num_long_term_pics** bằng 0.

num_long_term_sps xác định số lượng của các hình ảnh tham chiếu dài hạn được xác định trong tập hợp tham số chuỗi hoạt động và được bao gồm trong tập hợp hình ảnh tham chiếu dài hạn của hình ảnh hiện thời. Nếu **num_long_term_sps** không

có mặt, bộ mã hóa video suy ra giá trị này bằng 0. Giá trị của `num_long_term_sps` nằm trong khoảng từ 0 đến $\text{Min}(\text{num_long_term_ref_pics_sps}, \text{max_dec_pic_buffering}[\text{max_temporal_layers_minus1}] - \text{NumNegativePics}[\text{StRpsIdx}] - \text{NumPositivePics}[\text{StRpsIdx}] - \text{num_long_term_pics})$, bao gồm cả các giá trị đầu mút.

`long_term_idx_sps[i]` xác định chỉ số, đến danh sách các hình ảnh tham chiếu dài hạn được xác định trong tập hợp tham số chuỗi hoạt động, của hình ảnh tham chiếu dài hạn thứ *i* được kế thừa từ tập hợp tham số chuỗi được tham chiếu đến tập hợp hình ảnh tham chiếu dài hạn của hình ảnh hiện thời. Giá trị của `long_term_idx_sps[i]` nằm trong khoảng từ 0 đến `num_long_term_ref_pics_sps - 1`, bao gồm cả các giá trị đầu mút.

`poc_lsb_lt[i]` xác định giá trị của các bit có trọng số nhỏ nhất của giá trị số đếm thứ tự hình ảnh của hình ảnh tham chiếu dài hạn thứ *i* được bao gồm trong tập hợp hình ảnh tham chiếu dài hạn của hình ảnh hiện thời. Độ dài của phần tử cú pháp `poc_lsb_lt[i]` là $\log_2 \text{max_pic_order_cnt_lsb_minus4} + 4$ bit. Đối với giá trị bất kỳ là *j* và *k* trong khoảng từ `num_long_term_sps` đến `num_long_term_pics + num_long_term_sps - 1`, bao gồm cả các giá trị đầu mút, nếu *j* nhỏ hơn *k*, `poc_lsb_lt[j]` sẽ không nhỏ hơn `poc_lsb_lt[k]`.

Biến `PocLsbLt[i]` được suy ra như sau.

```
if( i < num_long_term_sps )
```

```
    PocLsbLt[i] = lt_ref_pic_poc_lsb_sps[ long_term_idx_sps[i] ]
```

```
else
```

```
    PocLsbLt[i] = poc_lsb_lt[i].
```

Nghĩa là, đối với các chỉ số từ 0 đến `num_long_term_sps`, giá trị `PocLsbLt[i]` thu được bằng cách sử dụng chỉ số trong `lt_ref_pic_poc_lsb_sps`, được báo hiệu trong phần đầu lát. Đối với các chỉ số `PocLsbLt` lớn hơn `num_long_term_sps`, giá trị `PocLsbLt[i]` là LSB POC của LTRP được báo hiệu rõ ràng trong phần đầu lát.

`delta_poc_msb_present_flag[i]` bằng 1 xác định là phần tử cú pháp `delta_poc_msb_cycle_lt[i]` có mặt. `delta_poc_msb_present_flag[i]` bằng 0 xác định là `delta_poc_msb_cycle_lt[i]` không có mặt, và do đó, không có MSB nào được báo hiệu cho phần đầu lát vì chỉ có một hình ảnh tham chiếu trong DPB có các LSB bằng `PocLsbLt[i]`. `delta_poc_msb_present_flag[i]` bằng 1 khi có nhiều hơn một hình ảnh

tham chiếu trong bộ đệm hình ảnh được giải mã với các bit có trọng số nhỏ nhất của giá trị số đếm thứ tự hình ảnh bằng $PocLsbLt[i]$.

$\Delta_{poc_msb_cycle_lt}[i]$ được sử dụng để xác định giá trị của các bit có trọng số lớn nhất của giá trị số đếm thứ tự hình ảnh của hình ảnh tham chiếu dài hạn thứ i được bao gồm trong tập hợp hình ảnh tham chiếu dài hạn của hình ảnh hiện thời. Bộ mã hóa video sử dụng giá trị $\Delta_{poc_msb_cycle_lt}[i]$ để xác định các MBS của LTRP thứ i .

Biến $\Delta_{poc_msb_cycle_lt}[i]$ được suy ra như được mô tả trong giả mã sau đây:

```
if( i == 0 || PocLsbLt[ i - 1 ] != PocLsbLt[ i ] )
    DeltaPocMSBCycleLt[ i ] = delta_poc_msb_cycle_lt[ i ]
else
    DeltaPocMSBCycleLt[ i ] = delta_poc_msb_cycle_lt[ i ] +
        DeltaPocMSBCycleLt[ i - 1 ]
```

Giá trị của $\Delta_{poc_msb_cycle_lt}[i] * MaxPicOrderCntLsb + pic_order_cnt_lsb - PocLsbLt[i]$ nằm trong khoảng từ 1 đến $2^{24} - 1$, bao gồm cả các giá trị đầu mút, đối với các kỹ thuật tuân theo JCTVC-I0340r2.

$used_by_curr_pic_lt_flag[i]$ bằng 0 xác định là hình ảnh tham chiếu dài hạn thứ i được bao gồm trong tập hợp hình ảnh tham chiếu dài hạn của hình ảnh hiện thời không được sử dụng để tham chiếu bởi hình ảnh hiện thời. Vì vậy, dựa vào các giá trị của $used_by_curr_pic_lt_flag[i]$, tập con các hình ảnh tham chiếu được kế thừa từ SPS hoặc được báo hiệu rõ ràng trong phần đầu lát có thể được sử dụng để tham chiếu bởi hình ảnh hiện thời.

JCTVC-I0340r2 cũng sửa đổi quy trình giải mã cho tập hợp hình ảnh tham chiếu. Theo JCTVC-I0340r2, quy trình này được gọi ra một lần cho mỗi hình ảnh, sau khi giải mã phần đầu lát nhưng trước khi giải mã đơn vị mã hóa bất kỳ và trước quy trình giải mã để thiết lập danh sách hình ảnh tham chiếu của lát này như được xác định trong mục 8.3.3 của HEVC WD6. Quy trình này có thể dẫn đến việc đánh dấu một hoặc nhiều hình ảnh tham chiếu là “không được sử dụng để tham chiếu”. JCTVC-I0340r2 bao gồm các sửa đổi sau đây đối với HEVC WD6:

```

for( i = 0, j = 0, k = 0; i < NumNegativePics[ StRpsIdx ]; i++ )
    if( UsedByCurrPicS0[ StRpsIdx ][ i ] )
        PocStCurrBefore[ j++ ] = PicOrderCntVal + DeltaPocS0[ StRpsIdx ][ i ]
    else
        PocStFoll[ k++ ] = PicOrderCntVal + DeltaPocS0[ StRpsIdx ][ i ]
NumPocStCurrBefore = j

for( i = 0, j = 0; i < NumPositivePics[ StRpsIdx ]; i++ )
    if( UsedByCurrPicS1[ StRpsIdx ][ i ] )
        PocStCurrAfter[ j++ ] = PicOrderCntVal + DeltaPocS1[ StRpsIdx ][ i ]
    else
        PocStFoll[ k++ ] = PicOrderCntVal + DeltaPocS1[ StRpsIdx ][ i ]
NumPocStCurrAfter = j
NumPocStFoll = k

```

(8-5)

```

for( i = 0, j = 0, k = 0; i < num_long_term_sps + num_long_term_pics; i++ )
    if( delta_poc_msb_present_flag[ i ] )
        if( used_by_curr_pic_lt_flag[ i ] )
            PocLtCurr[ j++ ] = PicOrderCntVal - DeltaPocMSBCycleLt[ i ] *
MaxPicOrderCntLsb -
            pic_order_cnt_lsb + PocLsbLt[ i ]
        else
            PocLtFoll[ k++ ] = PicOrderCntVal - DeltaPocMSBCycleLt[ i ] *
MaxPicOrderCntLsb -
            pic_order_cnt_lsb +
PocLsbLt[ i ]
        else
            if( used_by_curr_pic_lt_flag[ i ] )
                PocLtCurr[ j++ ] = PocLsbLt[ i ]
            else
                PocLtFoll[ k++ ] = PocLsbLt[ i ]
NumPocLtCurr = j
NumPocLtFoll = k

```

Một trong số các vấn đề tiềm tàng đi kèm theo thiết kế hiện có để báo hiệu các LTRP liên quan đến ngữ nghĩa của cờ `delta_poc_msb_present_flag[ i ]`. Trong các ngữ nghĩa gần đây nhất như được mô tả trên đây, `delta_poc_msb_present_flag[ i ]` là bằng

1 khi có nhiều hơn một hình ảnh tham chiếu trong bộ đệm hình ảnh được giải mã với các bit có trọng số nhỏ nhất của giá trị số đếm thứ tự hình ảnh bằng $\text{PocLsbLt}[i]$. Giá trị của $\text{delta_poc_msb_present_flag}[i]$ vì vậy phụ thuộc vào các hình ảnh tham chiếu trong DPB. Ngay cả khi có các hình ảnh tham chiếu trong DPB được đánh dấu là “không được sử dụng để tham chiếu”, giá trị của $\text{delta_poc_msb_present_flag}[i]$ cũng sẽ phụ thuộc vào các hình ảnh đó. Điều này có thể không được mong muốn, vì các hình ảnh trong DPB đánh dấu là “không được sử dụng để tham chiếu” có mặt và chỉ đợi để được xuất ra và không còn được sử dụng để tham chiếu. Việc các hình ảnh này có mặt trong DPB hay không thường phụ thuộc vào lịch xuất ra của bộ giải mã.

Cho trước dòng bit, các hệ thống và ứng dụng có thể áp dụng các lịch xuất ra khác nhau dựa vào bộ nhớ sẵn có mà có thể được sử dụng để lưu trữ các hình ảnh được giải mã. Vì vậy, việc đếm các hình ảnh này không thể chỉ làm tăng một cách không cần thiết cơ hội đòi hỏi giá trị của $\text{delta_poc_msb_present_flag}[i]$ bằng 1, vì vậy đòi hỏi nhiều bit hơn để báo hiệu LTRP, nhưng cũng có thể tạo ra các vấn đề về sự thích nghi dòng bit và tương kết. Ví dụ, khi bộ giải mã thực sự sử dụng nhiều bộ nhớ DPB hơn lượng tối thiểu cần thiết, dòng bit thích nghi có thể xuất hiện dưới dạng không thích nghi với bộ giải mã do nhiều hình ảnh chỉ đợi để được xuất ra, và không được sử dụng để tham chiếu nữa, có thể có mặt trong DPB (so với bộ giải mã sử dụng chính xác lượng tối thiểu bộ nhớ DPB cần thiết) và do đó, bộ giải mã thậm chí có thể không giải mã chính xác dòng bit này.

Một vấn đề tiềm tàng khác với các ngữ nghĩa được mô tả trên đây liên quan đến việc báo hiệu $\text{long_term_idx_sps}[i]$. Các ngữ nghĩa như được mô tả trên đây không giới hạn thứ tự trong đó các LTRP kế thừa từ SPS được báo hiệu. Việc thiếu đi sự giới hạn thứ tự cũng có thể là không hiệu quả, vì bộ mã hóa video có thể sử dụng nhiều bit hơn cần thiết để báo hiệu thông tin, như các MSB và LSB của các LTRP. Ví dụ, giả định là MaxPicOrderCntLsb bằng 256, và hai LSB được báo hiệu trong SPS - 0 và 10. Giả định đối với hình ảnh có POC 2560, là bộ mã hóa video sẽ báo hiệu ba LTRP, có lần lượt POC 0, 10, và 256. Đối với các hình ảnh có POC 0 và 256, bộ mã hóa video tuân theo với các kỹ thuật báo hiệu LTRP được mô tả trên đây sẽ phải báo hiệu các chu kỳ MSB cho cả hai LTRP. Tuy nhiên, các kỹ thuật theo sáng chế có thể cải thiện hiệu quả của việc báo hiệu này. Cụ thể là, sự thiếu hiệu quả nhất định được thể hiện bằng bảng 3 dưới đây. Nghĩa là, Bảng 3 đưa ra ví dụ về việc báo hiệu thiếu hiệu quả

các LTRP được kế thừa từ SPS tương ứng. Trong ví dụ này, “i” mô tả hình ảnh tham chiếu dài hạn cụ thể, LTRP POC mô tả giá trị POC của hình ảnh tham chiếu dài hạn tương ứng, PocLsbLt mô tả các bit có trọng số nhỏ nhất (LSB) của giá trị POC cho hình ảnh tham chiếu dài hạn được kế thừa từ SPS, delta_poc_msb_present_flag chỉ báo việc delta_poc_msb_cycle_lt có được báo hiệu hay không, và delta_poc_msb_cycle_lt được sử dụng để xác định giá trị của các bit có trọng số lớn nhất (MSB) của giá trị POC cho hình ảnh tham chiếu dài hạn, khi được báo hiệu.

Bảng 3

i	0	1	2
LTRP POC	0	10	256
PocLsbLt	0	10	0
delta_poc_msb_present_flag	1	0	1
delta_poc_msb_cycle_lt	10	-	9

Sáng chế mô tả các kỹ thuật để báo hiệu các thông tin nhận dạng cho các hình ảnh tham chiếu dài hạn, có thể có nhiều ưu điểm so với các kỹ thuật báo hiệu được mô tả trên đây trong một số trường hợp. Trong một số ví dụ, các kỹ thuật theo sáng chế có thể làm giảm số lượng bit được sử dụng để báo hiệu các phần tử cú pháp bao gồm MSB của các LTRP. Một cách tốt hơn để báo hiệu các hình ảnh có thể là xác định thứ tự mã hóa hình ảnh của các LTRP, nghĩa là thứ tự mà các LTRP được mã hóa trong poc_lsb_lt [i], sao cho các LSB của các LTRP là không giảm (hoặc với hiệu quả thực hiện gần như tương tự, không tăng), như được mô tả trong bảng 4. Việc mã hóa các LSB của các LTRP theo thứ tự không giảm hoặc không tăng có thể khiến cho bộ mã hóa video báo hiệu ít bit MSB hơn trong một số ví dụ. Bằng cách xếp thứ tự các LTRP theo thứ tự LSB không tăng hoặc không giảm, bộ mã hóa video sẽ báo hiệu liên tục các LTRP bất kỳ với các LSB bằng nhau theo thứ tự mã hóa hình ảnh của phần tử cú pháp poc_lsb_lt.

Nếu hai LTRP có giá trị LSB bằng nhau, bộ mã hóa video có thể mã hóa dữ liệu biểu diễn MSB của ít nhất một trong hai LTRP này (ví dụ, phần tử cú pháp delta_poc_msb_cycle_lt). Bộ giải mã video có thể mã hóa dữ liệu biểu diễn MSB dưới dạng chênh lệch giữa giá trị MSB của LTRP thứ hai và MSB của LTRP thứ nhất. Để xác định giá trị MSB của LTRP từ các giá trị được báo hiệu của delta_poc_msb_cycle_lt[i], bộ mã hóa video có thể tính toán biến

$\Delta\text{PocMSBCycleLt}[i]$ bằng tổng của $\text{delta_poc_msb_cycle_lt}[i]$ và $\text{delta_poc_msb_cycle_lt}[i-1]$. Trong trường hợp không có LTRP trước đó theo thứ tự mã hóa hình ảnh, bộ mã hóa video có thể báo hiệu chênh lệch giữa MSB của LTRP và MSB của hình ảnh hiện thời trong giá trị của phần tử cú pháp $\text{delta_poc_msb_cycle_lt}$.

Ví dụ, bằng cách sử dụng các kỹ thuật trước đây để báo hiệu MSB LTRP được minh họa trong bảng 3, bộ mã hóa video có thể báo hiệu LTRP với giá trị POC bằng 0 là LTRP thứ nhất theo thứ tự đếm hình ảnh ($i = 0$), và giá trị của $\text{delta_poc_msb_cycle_lt}$ là 10 đối với POC 0. Bộ mã hóa video có thể báo hiệu LTRP với giá trị POC 256 là LTRP thứ 3 theo thứ tự đếm hình ảnh (nghĩa là, $i = 2$), và giá trị $\text{delta_poc_msb_cycle_lt}$ là 9 đối với LTRP.

Mặc dù các kỹ thuật mã hóa video được mô tả đối với bảng 3 báo hiệu 10 là giá trị của $\text{delta_poc_msb_cycle_lt}$ đối với POC 0, bộ mã hóa ví dụ tuân theo các kỹ thuật của sáng chế có thể báo hiệu dữ liệu mà có thể được sử dụng để sao chép MSB của LTRP như là phần bù mà có thể là chênh lệch giữa giá trị chu kỳ MSB của POC của LTRP hiện thời và MSB của POC của LTRP trước đó. Như được minh họa trong bảng 4, bộ mã hóa video có thể giới hạn thứ tự đếm hình ảnh ở thứ tự không giảm của các LSB của mỗi LTRP được kế thừa từ SPS. Kết quả của việc xếp thứ tự này là các LTRP với POC 256, POC 0, mà đều có LSB bằng 0, được xếp thứ tự thứ nhất và thứ hai, và LTRP với POC 10 được xếp thứ ba theo thứ tự đếm hình ảnh. Bộ mã hóa video báo hiệu giá trị chu kỳ MSB của LTRP với POC 256 là 9 vì nó là LTRP thứ nhất theo thứ tự đếm hình ảnh, và giá trị chu kỳ MSB của LTRP thứ hai là chênh lệch giữa MSB LTRP hiện thời (10) và MSB LTRP trước đó (9) nghĩa là, $10 - 9 = 1$, do đó bộ mã hóa video báo hiệu 1 là giá trị cho $\text{delta_poc_msb_cycle_lt}[i]$. Việc báo hiệu giá trị 1 có thể đòi hỏi ít bit hơn so với việc báo hiệu giá trị 10, giúp cải thiện hiệu quả mã hóa để báo hiệu của phần tử cú pháp $\text{delta_poc_msb_cycle_lt}$ trong ví dụ này. Do đó, các kỹ thuật theo sáng chế, được mô tả liên quan đến Bảng 4 có thể làm giảm số lượng bit cần thiết để báo hiệu các MSB cho các giá trị POC của các LTRP được kế thừa từ SPS tương ứng cho chuỗi video được mã hóa.

Bảng 4

i	0	1	2
LTRP POC	256	0	10
PocLsbLt	0	0	10
delta_poc_msb_present_flag	1	1	0
delta_poc_msb_cycle_lt	9	1	-

Như được thảo luận trên đây, để khắc phục các vấn đề được đưa ra trên đây liên đến việc báo hiệu các MSB của các LTRP, sáng chế mô tả nhiều kỹ thuật khác nhau để báo hiệu các hình ảnh tham chiếu dài hạn (LTRP) được bao gồm trong SPS của hình ảnh được mã hóa và lập chỉ số các LTRP trong phần đầu lát của chuỗi hình ảnh được mã hóa một cách rõ ràng bằng cách báo hiệu LSB của giá trị POC của LTRP trong phần đầu lát, hoặc bằng cách báo hiệu trong phần đầu lát chỉ số cho mục nhập trong danh sách của các LSB của LTRP trong SPS mà xác định LSB của POC của LTRP. Sáng chế cũng mô tả các thay đổi về ngữ nghĩa của `delta_poc_msb_present_flag[ i ]` mà có thể được áp dụng để đảm bảo rằng giá trị của cờ chỉ phụ thuộc vào các hình ảnh được đánh dấu là “được sử dụng để tham chiếu”. Các kỹ thuật này và các kỹ thuật khác được mô tả chi tiết hơn dưới đây.

Fig.1 là sơ đồ khối minh họa hệ thống mã hóa và giải mã video 10 làm ví dụ có thể sử dụng các kỹ thuật để báo hiệu dữ liệu cho LTRP trong SPS hoặc phần đầu lát. Như được thể hiện trên Fig.1, hệ thống 10 bao gồm thiết bị nguồn 12 cung cấp dữ liệu video mã hóa sẽ được giải mã sau đó bởi thiết bị đích 14. Cụ thể là, thiết bị nguồn 12 cung cấp dữ liệu video cho thiết bị đích 14 qua vật ghi đọc được bằng máy tính 16. Thiết bị nguồn 12 và thiết bị đích 14 có thể bao gồm bất kỳ trong số nhiều loại thiết bị, bao gồm máy tính để bàn, máy tính sổ tay (tức là máy tính xách tay), máy tính bảng, hộp giải mã, điện thoại như điện thoại thông minh, bảng “thông minh”, ti vi, camera, thiết bị hiển thị, trình phát phương tiện số, bàn điều khiển trò chơi video, thiết bị truyền liên tục video, hoặc tương tự. Trong một số trường hợp, thiết bị nguồn 12 và thiết bị đích 14 có thể được trang bị để truyền thông không dây.

Thiết bị đích 14 có thể thu dữ liệu video mã hóa sẽ được giải mã qua vật ghi đọc được bằng máy tính 16. Vật ghi đọc được bằng máy tính 16 có thể bao gồm loại vật ghi hoặc thiết bị bất kỳ có khả năng dịch chuyển dữ liệu video mã hóa từ thiết bị nguồn 12 đến thiết bị đích 14. Trong một ví dụ, vật ghi đọc được bằng máy tính 16 có thể bao gồm phương tiện truyền thông để cho phép thiết bị nguồn 12 truyền dữ liệu

video mã hóa trực tiếp đến thiết bị đích 14 theo thời gian thực. Dữ liệu video mã hóa có thể được điều biến theo chuẩn truyền thông, như giao thức truyền thông không dây, và được truyền đến thiết bị đích 14. Phương tiện truyền thông này có thể bao gồm phương tiện không dây hoặc có dây bất kỳ, như phổ tần số vô tuyến (radio frequency: RF) hoặc một hoặc nhiều đường truyền vật lý. Phương tiện truyền thông này có thể tạo thành một phần của mạng dựa trên chuyển mạch gói, như mạng cục bộ, mạng diện rộng, hoặc mạng toàn cầu như Internet. Phương tiện truyền thông này có thể bao gồm bộ định tuyến, bộ chuyển mạch, trạm cơ sở, hoặc thiết bị bất kỳ khác có thể hữu ích để tạo điều kiện cho việc truyền thông từ thiết bị nguồn 12 đến thiết bị đích 14.

Trong một số ví dụ, dữ liệu được mã hóa có thể được xuất từ giao diện đầu ra 22 đến thiết bị lưu trữ. Tương tự, dữ liệu được mã hóa có thể được truy cập từ thiết bị lưu trữ bởi giao diện đầu vào. Thiết bị lưu trữ này có thể bao gồm bất kỳ trong số nhiều loại phương tiện lưu trữ dữ liệu được truy cập cục bộ hoặc phân tán như ổ cứng, đĩa Blu-ray, DVD, CD-ROM, bộ nhớ nhanh, bộ nhớ khả biến hoặc không khả biến, hoặc phương tiện lưu trữ số thích hợp bất kỳ khác để lưu trữ dữ liệu video mã hóa. Trong ví dụ khác, thiết bị lưu trữ này có thể tương ứng với máy chủ tập tin hoặc thiết bị lưu trữ trung gian khác mà có thể lưu trữ video đã mã hóa được tạo ra bởi thiết bị nguồn 12. Thiết bị đích 14 có thể truy cập dữ liệu video được lưu trữ từ thiết bị lưu trữ bằng cách truyền liên tục hoặc tải về. Máy chủ tập tin có thể là loại máy chủ bất kỳ có khả năng lưu trữ dữ liệu video mã hóa và truyền dữ liệu video mã hóa đó đến thiết bị đích 14. Các máy chủ tập tin làm ví dụ bao gồm máy chủ web (ví dụ, đối với website), máy chủ FTP, thiết bị lưu trữ gắn mạng (network attached storage: NAS), hoặc ổ đĩa cục bộ. Thiết bị đích 14 có thể truy cập dữ liệu video mã hóa qua kết nối dữ liệu chuẩn bất kỳ, bao gồm kết nối Internet. Kết nối này có thể bao gồm kênh không dây (ví dụ, kết nối Wi-Fi), kết nối có dây (ví dụ, DSL, modem cáp, v.v...), hoặc tổ hợp của cả hai loại kết nối mà thích hợp để truy cập dữ liệu video mã hóa được lưu trữ trên máy chủ tập tin. Cuộc truyền dữ liệu video mã hóa từ thiết bị lưu trữ này có thể là cuộc truyền liên tục, cuộc truyền tải về, hoặc tổ hợp của chúng.

Các kỹ thuật theo sáng chế không nhất thiết bị giới hạn ở các ứng dụng hoặc thiết lập không dây. Các kỹ thuật này có thể được áp dụng cho quy trình mã hóa video để hỗ trợ bất kỳ trong số nhiều loại ứng dụng đa phương tiện, như truyền hình qua vô tuyến, truyền hình cáp, truyền hình vệ tinh, truyền video liên tục qua internet, như

truyền liên tục thích ứng động qua HTTP (dynamic adaptive streaming over HTTP: DASH), video số được mã hóa lên phương tiện lưu trữ dữ liệu, quy trình giải mã video số được lưu trữ trên phương tiện lưu trữ dữ liệu, hoặc các ứng dụng khác. Trong một số ví dụ, hệ thống 10 có thể được tạo cấu hình để hỗ trợ cuộc truyền video một chiều hoặc hai chiều để hỗ trợ các ứng dụng như truyền liên tục video, phát lại video, phát rộng video, và/hoặc điện thoại video.

Trong ví dụ trên Fig.1, thiết bị nguồn 12 bao gồm nguồn video 18, bộ mã hóa video 20, và giao diện đầu ra 22. Thiết bị đích 14 bao gồm giao diện đầu vào 28, bộ giải mã video 30 và thiết bị hiển thị 32. Theo sáng chế, bộ mã hóa video 20 của thiết bị nguồn 12 có thể được tạo cấu hình để áp dụng các kỹ thuật để báo hiệu dữ liệu cho các LTRP trong SPS. Trong các ví dụ khác, thiết bị nguồn và thiết bị đích có thể bao gồm các thành phần khác hoặc các cách sắp xếp khác. Ví dụ, thiết bị nguồn 12 có thể thu dữ liệu video từ nguồn video ngoài 18, như camera ngoài. Tương tự, thiết bị đích 14 có thể giao tiếp với thiết bị hiển thị ngoài, hơn là bao gồm thiết bị hiển thị tích hợp.

Hệ thống đã được minh họa trên Fig.1 chỉ là một ví dụ. Các kỹ thuật để báo hiệu dữ liệu cho các LTRP trong SPS có thể được thực hiện bằng thiết bị mã hóa và/hoặc giải mã video số bất kỳ. Nói chung, mặc dù các kỹ thuật theo sáng chế được thực hiện bằng thiết bị mã hóa video, các kỹ thuật này cũng có thể được thực hiện bằng bộ mã hóa/bộ giải mã video, thường được gọi là “CODEC”. Ngoài ra, các kỹ thuật theo sáng chế cũng có thể được thực hiện bằng bộ xử lý video. Thiết bị nguồn 12 và thiết bị đích 14 chỉ là các ví dụ về các thiết bị mã hóa này trong đó thiết bị nguồn 12 tạo ra dữ liệu video mã hóa để truyền đến thiết bị đích 14. Trong một số ví dụ, các thiết bị 12, 14 có thể hoạt động theo cách gần như đối xứng sao cho mỗi thiết bị 12, 14 bao gồm các thành phần mã hóa và giải mã video. Do đó, hệ thống 10 có thể hỗ trợ cuộc truyền video một chiều hoặc hai chiều giữa các thiết bị video 12, 14, ví dụ, để truyền video liên tục, phát lại video, phát rộng video, hoặc điện thoại video.

Nguồn video 18 của thiết bị nguồn 12 có thể bao gồm thiết bị ghi video, như camera video, bộ lưu trữ video chứa video đã được ghi trước đó, và/hoặc giao diện nạp video để thu video từ nhà cung cấp nội dung video. Theo phương án khác, nguồn video 18 có thể tạo ra dữ liệu dựa trên đồ họa máy tính làm video nguồn, hoặc tổ hợp của video trực tiếp, video được lưu trữ, và video được tạo ra bởi máy tính. Trong một số trường hợp, nếu nguồn video 18 là camera video, thiết bị nguồn 12 và thiết bị đích

14 có thể tạo ra thiết bị gọi là điện thoại camera hoặc điện thoại video. Tuy nhiên, như được đề cập trên đây, các kỹ thuật được mô tả trong tài liệu này có thể được ứng dụng cho mã hóa video nói chung, và có thể được áp dụng cho các ứng dụng không dây và/hoặc có dây. Trong mỗi trường hợp, video được tạo ra bằng máy tính, được ghi, được ghi trước đó có thể được mã hóa bởi bộ mã hóa video 20. Sau đó, thông tin video được mã hóa có thể được xuất ra bởi giao diện đầu ra 22 lên vật ghi đọc được bằng máy tính 16.

Vật ghi đọc được bằng máy tính 16 có thể bao gồm phương tiện tạm thời, như phát rộng không dây hoặc truyền qua mạng nối dây, hoặc phương tiện lưu trữ (nghĩa là, phương tiện lưu trữ bất biến), như đĩa cứng, ổ flash, đĩa nén, đĩa video số, đĩa Blu-ray, hoặc vật ghi đọc được bằng máy tính. Trong một số ví dụ, máy chủ mạng (không được thể hiện trên hình vẽ) có thể thu dữ liệu video mã hóa từ thiết bị nguồn 12 và cung cấp dữ liệu video mã hóa đến thiết bị đích 14, ví dụ, qua truyền mạng. Tương tự, thiết bị máy tính của cơ sở sản xuất vật ghi, như cơ sở dập đĩa có thể thu dữ liệu video mã hóa từ thiết bị nguồn 12 và tạo ra đĩa chứa dữ liệu video mã hóa. Do đó, vật ghi đọc được bằng máy tính 16 có thể được hiểu là bao gồm một hoặc nhiều vật ghi đọc được bằng máy tính thuộc nhiều dạng khác nhau, trong nhiều ví dụ khác nhau.

Giao diện đầu vào 28 của thiết bị đích 14 thu thông tin từ vật ghi đọc được bằng máy tính 16. Thông tin của vật ghi đọc được bằng máy tính 16 có thể bao gồm thông tin cú pháp được xác định bởi bộ mã hóa video 20, cũng được sử dụng bởi bộ giải mã video 30, bao gồm các phần tử cú pháp mô tả các đặc điểm và/hoặc quy trình xử lý các khối và các đơn vị được mã hóa khác, ví dụ, GOP. Thiết bị hiển thị 32 hiển thị dữ liệu video được giải mã cho người dùng, và có thể bao gồm bất kỳ trong số nhiều loại thiết bị hiển thị như ống tia catôt (cathode ray tube: CRT), màn hình tinh thể lỏng (liquid crystal display: LCD), màn hình plasma, màn hình điốt phát quang hữu cơ (organic light emitting diode: OLED), hoặc loại thiết bị hiển thị khác.

Bộ mã hóa video 20 và bộ giải mã video 30 có thể hoạt động theo chuẩn mã hóa video, như chuẩn mã hóa video hiệu suất cao (High Efficiency Video Coding: HEVC) hiện đang được phát triển, và có thể tuân theo HEVC Test Model (HM). Ngoài ra, bộ mã hóa video 20 và bộ giải mã video 30 có thể hoạt động theo các chuẩn công nghiệp hoặc độc quyền khác, như chuẩn ITU-T H.264, ngoài ra còn được gọi là MPEG-4, phần 10, chuẩn mã hóa video tiên tiến (Advanced Video Coding: AVC), hoặc các

phiên bản mở rộng của các chuẩn này. Tuy nhiên, các kỹ thuật theo sáng chế không bị giới hạn ở chuẩn mã hóa cụ thể bất kỳ. Các ví dụ khác về các chuẩn mã hóa video bao gồm MPEG-2 và ITU-T H.263. Mặc dù không được thể hiện trên Fig.1, theo một số khía cạnh, mỗi bộ mã hóa video 20 và bộ giải mã video 30 có thể được tích hợp với bộ mã hóa và bộ giải mã âm thanh, và có thể bao gồm các thiết bị MUX-DEMUX thích hợp, hoặc phần cứng và phần mềm khác, để thực hiện mã hóa cả âm thanh và video trong chuỗi dữ liệu chung hoặc các chuỗi dữ liệu riêng. Nếu có thể áp dụng, các thiết bị MUX-DEMUX có thể tuân theo giao thức ghép kênh ITU H.223, hoặc các giao thức khác như giao thức gói dữ liệu người dùng (user datagram protocol: UDP).

Chuẩn ITU-T H.264/MPEG-4 (AVC) được đề ra bởi Nhóm chuyên gia mã hóa video (Video Coding Experts Group: VCEG) ITU-T cùng với Nhóm chuyên gia hình ảnh động (Moving Picture Experts Group: MPEG) ISO/IEC là sản phẩm của nhóm hợp tác tập thể có tên là Đội ngũ video kết hợp (Joint Video Team: JVT). Theo một số khía cạnh, các kỹ thuật được mô tả trong tài liệu này có thể được áp dụng cho các thiết bị mà thường tuân theo chuẩn H.264. Chuẩn H.264 được mô tả trong Khuyến nghị ITU-T H.264, mã hóa video tiên tiến cho các dịch vụ nghe nhìn chung bởi Nhóm nghiên cứu ITU-T, vào tháng 3/2005, có thể được gọi là chuẩn H.264 hoặc đặc tả H.264, hoặc chuẩn hoặc đặc tả H.264/AVC. Đội ngũ video kết hợp (JVT) tiếp tục làm việc trên các phiên bản mở rộng của H.264/MPEG-4 AVC.

Bộ mã hóa video 20 và bộ giải mã video 30 có thể được cài đặt dưới dạng mạch bất kỳ trong số nhiều loại mạch mã hóa thích hợp, như một hoặc nhiều bộ vi xử lý, bộ xử lý tín hiệu số (digital signal processor: DSP), mạch tích hợp chuyên dụng (application specific integrated Circuit: ASIC), mảng cổng lập trình được dạng trường (field programmable gate array: FPGA), logic rời rạc, phần mềm, phần cứng, phần sụn hoặc tổ hợp bất kỳ của chúng. Khi các kỹ thuật này được thực hiện một phần trong phần mềm, thiết bị có thể lưu trữ các lệnh cho phần mềm trong vật ghi bất biến đọc được bằng máy tính thích hợp, và thực hiện các lệnh trong phần cứng bằng cách sử dụng một hoặc nhiều bộ xử lý để thực hiện các kỹ thuật theo sáng chế. Mỗi bộ mã hóa video 20 và bộ giải mã video 30 có thể được bao gồm trong một hoặc nhiều bộ mã hóa hoặc giải mã, một trong số chúng có thể được tích hợp như là một phần của bộ mã hóa/bộ giải mã kết hợp (CODEC) trong thiết bị tương ứng.

JCT-VC tiếp tục phát triển chuẩn HEVC. Các nỗ lực chuẩn hóa HEVC là dựa trên mô hình phát triển của thiết bị mã hóa video được gọi là Mô hình thử nghiệm HEVC (HEVC Test Model: HM). HM giả định một số tính năng bổ sung của thiết bị mã hóa video so với các thiết bị hiện nay theo, ví dụ, ITU-T H.264/ AVC. Ví dụ, trong khi H.264 cung cấp 9 chế độ mã hóa dự báo nội cấu trúc, HM có thể cung cấp đến 33 chế độ mã hóa dự báo nội cấu trúc.

Nói chung, mô hình làm việc của HM mô tả là khung video hoặc hình ảnh có thể được chia thành chuỗi khối cây hoặc các đơn vị mã hóa lớn nhất (LCU) (còn được gọi là “đơn vị cây mã hóa”) bao gồm cả mẫu độ chói và màu sắc. Dữ liệu cú pháp trong dòng bit có thể xác định kích thước cho LCU, là đơn vị mã hóa lớn nhất về số lượng điểm ảnh. Lát bao gồm một số khối cây liên tiếp theo thứ tự mã hóa. Khung video hoặc hình ảnh có thể được chia thành một hoặc nhiều lát. Mỗi khối cây có thể được chia thành các đơn vị mã hóa (coding unit: CU) theo cây tứ phân. Nói chung, cấu trúc dữ liệu tứ phân bao gồm một nút mỗi CU, với nút rễ tương ứng với khối cây. Nếu CU được tách thành 4 CU con, nút tương ứng với CU bao gồm 4 nút lá, mỗi nút tương ứng với một trong số các CU con.

Theo đó, bộ mã hóa video 20 có thể mã hóa một loạt các LCU tạo ra lát của hình ảnh. Ngoài ra, bộ mã hóa video 20 có thể sử dụng các kỹ thuật theo sáng chế để mã hóa phần đầu lát cho lát này. Phần đầu lát có thể bao gồm dữ liệu có thể được sử dụng trong quá trình giải mã lát này. Ví dụ, phần đầu lát có thể bao gồm dữ liệu chỉ báo các hình ảnh tham chiếu mà có thể được sử dụng trong quá trình giải mã dữ liệu video được bao gồm trong lát. Theo các kỹ thuật của sáng chế, bộ mã hóa video 20 có thể mã hóa dữ liệu của phần đầu lát mà chỉ báo việc dữ liệu biểu diễn thông tin nhận dạng cho hình ảnh tham chiếu dài hạn cụ thể có được mã hóa rõ ràng trong phần đầu lát hay không, hoặc việc thông tin nhận dạng này có được kế thừa từ tập hợp tham số chuỗi (sequence parameter set: SPS) tương ứng với chuỗi hình ảnh bao gồm hình ảnh trong đó lát này xuất hiện hay không.

Tương tự, bộ giải mã video 30 có thể sử dụng các kỹ thuật theo sáng chế để giải mã phần đầu lát, mà có thể bao gồm giải mã dữ liệu chỉ báo việc thông tin nhận dạng cho hình ảnh tham chiếu dài hạn cụ thể có được mã hóa rõ ràng trong phần đầu lát hay không, hoặc việc thông tin nhận dạng này có được kế thừa từ SPS tương ứng với lát này hay không. Nếu phần đầu lát chỉ báo là thông tin nhận dạng này được mã hóa rõ

ràng, bộ giải mã video 30 có thể giải mã thông tin nhận dạng cho hình ảnh tham chiếu dài hạn từ phần đầu lát. Tuy nhiên, nếu phần đầu lát chỉ báo rằng thông tin nhận dạng được kế thừa từ SPS, bộ giải mã video 30 có thể khôi phục được thông tin nhận dạng cho SPS từ SPS này. Các kỹ thuật để mã hóa phần đầu lát được giải thích chi tiết hơn dưới đây.

Mỗi nút của cấu trúc dữ liệu cây tứ phân có thể cung cấp dữ liệu cú pháp cho CU tương ứng. Ví dụ, nút trong cây tứ phân có thể bao gồm cờ tách, chỉ báo việc CU tương ứng với nút này có được tách thành các CU con hay không. Các phần tử cú pháp cho CU có thể được xác định đệ quy, và có thể phụ thuộc vào việc CU có được tách thành các CU con hay không. Nếu CU không được tách thêm, nó được gọi là CU lá. Trong bản mô tả này, 4 CU con của CU lá cũng sẽ được gọi là CU lá ngay cả khi không có sự tách rõ ràng CU lá gốc. Ví dụ, nếu CU kích thước 16x16 không được tách thêm, 4 CU con kích thước 8x8 cũng sẽ được gọi là CU lá mặc dù CU kích thước 16x16 không bao giờ tách.

CU có mục đích tương tự như khối macro của chuẩn H.264, ngoại trừ là CU không có sự khác biệt về kích thước. Ví dụ, khối cây có thể được tách thành 4 nút con (còn được gọi là CU con) và mỗi nút con này lại có thể là nút cha và được tách thành 4 nút con khác. Nút con không tách cuối cùng, gọi là nút lá của cây tứ phân, bao gồm nút mã hóa, còn được gọi là CU lá. Dữ liệu cú pháp liên quan đến dòng bit được mã hóa có thể xác định số lần tối đa mà khối cây có thể được tách, được gọi là độ sâu CU tối đa, và cũng có thể xác định kích thước tối thiểu của nút mã hóa. Theo đó, dòng bit cũng có thể xác định đơn vị mã hóa nhỏ nhất (smallest coding unit: SCU). Sáng chế sử dụng thuật ngữ “khối” để chỉ bất kỳ trong số CU, PU, hoặc TU, trong ngữ cảnh HEVC, hoặc cấu trúc dữ liệu tương tự trong ngữ cảnh của các chuẩn khác (ví dụ, khối macro và các khối con của nó trong chuẩn H.264/AVC).

CU bao gồm nút mã hóa và đơn vị dự báo (prediction unit: PU) và đơn vị biến đổi (transform unit: TU) gắn với nút mã hóa. Kích thước của CU tương ứng với kích thước của nút mã hóa và phải có hình vuông. Kích thước của CU này có thể nằm trong khoảng từ 8x8 điểm ảnh đến kích thước của khối cây có tối đa là 64x64 điểm ảnh hoặc lớn hơn. Mỗi CU có thể chứa một hoặc nhiều PU và một hoặc nhiều TU. Dữ liệu cú pháp gắn với CU có thể mô tả, ví dụ, việc chia CU thành một hoặc nhiều PU. Các chế độ chia có thể khác nhau giữa việc CU được mã hóa theo chế độ bỏ qua hay chế độ

trước tiếp, được mã hóa theo chế độ dự báo nội cấu trúc hay chế độ dự báo liên cấu trúc. Các PU có thể được chia để thành hình khác hình vuông. Dữ liệu cú pháp gắn với CU cũng có thể mô tả, ví dụ, việc chia CU thành một hoặc nhiều TU theo cây tứ phân. TU có thể có dạng hình vuông hoặc không vuông (ví dụ, hình chữ nhật).

Chuẩn HEVC cho phép biến đổi theo TU, mà có thể khác nhau đối với các CU khác nhau. Các TU thường được định kích thước dựa vào kích thước của các PU trong CU đã cho được xác định cho LCU đã chia, mặc dù không phải luôn như vậy. Các TU thường có cùng kích thước hoặc nhỏ hơn so với các PU. Trong một số ví dụ, các mẫu dư tương ứng với CU có thể được chia thành thêm các đơn vị nhỏ hơn bằng cách sử dụng cấu trúc cây tứ phân còn được gọi là “cây tứ phân dư” (residual quad tree: RQT). Nút lá của RQT có thể được gọi là các đơn vị biến đổi (transform unit: TU). Các giá trị chênh lệch điểm ảnh gắn với các TU có thể được biến đổi để tạo ra các hệ số biến đổi, mà có thể được lượng tử hóa.

CU lá có thể bao gồm một hoặc nhiều đơn vị dự báo (prediction unit: PU). Nói chung, PU biểu diễn vùng không gian tương ứng với tất cả hoặc một phần của CU tương ứng, và có thể bao gồm dữ liệu để phục hồi mẫu tham chiếu cho PU này. Ngoài ra, PU bao gồm dữ liệu liên quan đến dự báo. Ví dụ, khi PU được mã hóa theo chế độ nội cấu trúc, dữ liệu cho PU này có thể được bao gồm trong cây tứ phân dư (RQT), mà có thể bao gồm dữ liệu mô tả chế độ dự báo nội cấu trúc cho TU tương ứng với PU này. Theo ví dụ khác, khi PU được mã hóa liên cấu trúc, PU này có thể bao gồm dữ liệu xác định một hoặc nhiều vector chuyển động cho PU này. Dữ liệu xác định vector chuyển động cho PU có thể mô tả, ví dụ, thành phần ngang của vector chuyển động, thành phần dọc của vector chuyển động, độ phân giải của vector chuyển động (ví dụ, độ chính xác 1/4 điểm ảnh hoặc độ chính xác 1/8 điểm ảnh), hình ảnh tham chiếu mà vector chuyển động này trỏ đến, và/hoặc danh sách hình ảnh tham chiếu (ví dụ, Danh sách 0, Danh sách 1, hoặc Danh sách C) cho vector chuyển động này.

CU lá có một hoặc nhiều PU cũng có thể bao gồm một hoặc nhiều đơn vị biến đổi (TU). Các đơn vị biến đổi này có thể được xác định bằng cách sử dụng RQT (còn được gọi là cấu trúc cây tứ phân TU), như được thảo luận trên đây. Ví dụ, cờ tách có thể chỉ báo việc CU lá có được tách thành 4 đơn vị biến đổi hay không. Sau đó, mỗi đơn vị biến đổi có thể được tách thêm thành các TU con. Khi TU không được tách thêm, nó có thể được gọi là TU lá. Thông thường, để mã hóa nội cấu trúc, toàn bộ các

TU lá thuộc CU lá dùng chung cùng một chế độ dự báo nội cấu trúc. Nghĩa là, cùng một chế độ dự báo nội cấu trúc thường được áp dụng để tính các giá trị được dự báo cho toàn bộ các TU của CU lá. Để mã hóa nội cấu trúc, bộ mã hóa video có thể tính toán giá trị dư cho mỗi TU lá bằng cách sử dụng chế độ dự báo nội cấu trúc, dưới dạng chênh lệch giữa phần của CU tương ứng với TU và khối ban đầu. TU không nhất thiết bị giới hạn ở kích thước của PU. Do đó, các TU có thể lớn hơn hoặc nhỏ hơn so với PU. Để mã hóa nội cấu trúc, PU có thể được sắp xếp cùng một vị trí với TU lá tương ứng cho cùng một CU. Trong một số ví dụ, kích thước tối đa của CU lá có thể tương ứng với kích thước của CU lá tương ứng.

Ngoài ra, TU của CU lá cũng có thể được gắn với cấu trúc dữ liệu cây tứ phân tương ứng, gọi là cây tứ phân dư (RQT). Nghĩa là, CU lá có thể bao gồm cây tứ phân chỉ báo cách thức mà CU lá được chia thành các TU. Nút rễ của cây tứ phân TU thường tương ứng với CU lá, trong khi nút rễ của cây tứ phân CU thường tương ứng với khối cây (hoặc LCU). Các TU của RQT mà không bị tách được gọi là các TU lá. Nói chung, bản mô tả này sử dụng các thuật ngữ CU và TU lần lượt để chỉ CU lá và TU lá, trừ khi được lưu ý khác đi.

Chuỗi video thường bao gồm chuỗi khung video hoặc hình ảnh. Nhóm hình ảnh (GOP) thường bao gồm chuỗi của một hoặc nhiều hình ảnh video. GOP có thể bao gồm dữ liệu cú pháp trong phần đầu của GOP, phần đầu của một hoặc nhiều hình ảnh, hoặc phần khác, mà mô tả số lượng hình ảnh được bao gồm trong GOP. Mỗi lát của hình ảnh có thể bao gồm dữ liệu cú pháp lát mô tả chế độ mã hóa cho lát tương ứng. Bộ mã hóa video 2D thường thao tác trên các khối video trong các lát video riêng lẻ để mã hóa dữ liệu video. Khối video có thể tương ứng với nút mã hóa trong CU. Các khối video này có thể có kích thước cố định hoặc thay đổi, và có thể khác nhau về kích thước theo chuẩn mã hóa xác định.

Ví dụ, HM hỗ trợ dự báo trong các kích thước PU khác nhau. Giả định là kích thước của CU cụ thể là $2N \times 2N$, HM hỗ trợ dự báo nội cấu trúc trong các kích thước PU là $2N \times 2N$ hoặc $N \times N$, và dự báo liên cấu trúc trong các kích thước PU đối xứng là $2N \times 2N$, $2N \times N$, $N \times 2N$, hoặc $N \times N$. HM còn hỗ trợ việc chia không đối xứng để dự báo liên cấu trúc trong các kích thước PU là $2N \times nU$, $2N \times nD$, $nL \times 2N$, và $nR \times 2N$. Trong chia không đối xứng, một chiều của CU không được chia, trong khi đó chiều còn lại được chia thành 25% và 75%. Phần CU tương ứng với phần chia 25% được chỉ báo

bởi “n” sau đó là chỉ báo “Lên”, “Xuống”, “Trái” hoặc “Phải”. Do đó, ví dụ, “2Nx_nU” chỉ CU kích thước 2Nx2N được chia theo chiều ngang với PU 2Nx0,5N ở trên và PU 2Nx1,5N ở dưới.

Trong bản mô tả này, “NxN” và “N nhân N” có thể được sử dụng thay thế cho nhau để chỉ các kích thước điểm ảnh của khối video theo chiều ngang và dọc, ví dụ 16x16 điểm ảnh hoặc 16 nhân 16 điểm ảnh. Nói chung, khối 16x16 sẽ có 16 điểm ảnh theo chiều dọc ($y = 16$) và 16 điểm ảnh theo chiều ngang ($x = 16$). Tương tự, khối NxN thường có N điểm ảnh theo chiều dọc và N điểm ảnh theo chiều ngang, trong đó N biểu diễn số nguyên không âm. Các điểm ảnh trong khối có thể được sắp xếp thành hàng và cột. Ngoài ra, các khối không cần thiết có số lượng điểm ảnh theo chiều ngang và theo chiều dọc giống nhau. Ví dụ, khối có thể bao gồm NxM điểm ảnh, trong đó M không nhất thiết bằng N.

Như nêu trên, bộ mã hóa video 20 và bộ giải mã video 30 có thể được tạo cấu hình để dự báo liên cấu trúc hoặc nội cấu trúc các PU của CU. Nói chung, mã hóa liên cấu trúc bao gồm dự báo liên quan đến một hoặc nhiều hình ảnh tham chiếu. Hình ảnh tham chiếu có thể là hình ảnh trước đó theo thứ tự thời gian, hình ảnh tương lai, hoặc tổ hợp của các dự báo từ hai hoặc nhiều hình ảnh đã được mã hóa trước đó. Bộ mã hóa video 20 hoặc bộ giải mã video 30 có thể lưu trữ các hình ảnh tham chiếu trong bộ đệm hình ảnh được giải mã (DPB). Bộ mã hóa video 20 có thể báo hiệu giá trị đơn vị NAL mà chỉ báo rằng hình ảnh sẽ được sử dụng làm hình ảnh tham chiếu bằng cách đánh dấu hình ảnh này là hình ảnh tham chiếu ngắn hạn hoặc hình ảnh tham chiếu dài hạn. Bộ mã hóa video 20 hoặc bộ giải mã video 30 lưu trữ các hình ảnh tham chiếu ngắn hạn và dài hạn trong danh sách hình ảnh tham chiếu. Bộ mã hóa video 20 hoặc bộ mã hóa video 20 có thể lưu trữ các hình ảnh tham chiếu dài hạn trong DPB trong khoảng thời gian dài hơn so với các hình ảnh tham chiếu ngắn hạn, và do đó có thể được sử dụng làm các hình ảnh tham chiếu cho nhiều hình ảnh hơn so với các hình ảnh tham chiếu ngắn hạn.

Bộ mã hóa video 20 hoặc bộ giải mã video 30 có thể loại bỏ các hình ảnh tham chiếu khỏi DPB theo một số cách. Khi các hình ảnh tham chiếu không còn cần thiết, bộ mã hóa video 20 hoặc bộ giải mã video 30 có thể đánh dấu hình ảnh tham chiếu này là không còn cần thiết cho việc tham chiếu và cuối cùng có thể loại bỏ các hình ảnh không còn cần thiết cho việc tham chiếu khỏi DPB, ví dụ, bằng cách xuất ra hình ảnh

tham chiếu để hiển thị, bằng cách loại bỏ hình ảnh tham chiếu nếu hình ảnh tham chiếu đã được xuất ra để hiển thị. Nếu DPB đã đầy, bộ mã hóa video 20 hoặc bộ giải mã video 30 có thể loại bỏ hình ảnh tham chiếu ngắn hạn cũ nhất.

Ngoài việc lưu trữ các hình ảnh tham chiếu trong DPB, bộ mã hóa video 20 hoặc bộ giải mã video 30 lưu trữ các hình ảnh tham chiếu trong một trong số hai danh sách hình ảnh tham chiếu, được gọi là Danh sách 0 và Danh sách 1. Các hình ảnh tham chiếu được lưu trữ trong Danh sách 0 và Danh sách 1 được sử dụng để dự báo hình ảnh cụ thể có thể được gọi là tập hợp hình ảnh tham chiếu (RPS). Các hình ảnh được lưu trữ trong Danh sách 0 và Danh sách 1 có thể đến từ nhiều loại nguồn khác nhau, như SPS, và PPS. Như được mô tả trên đây, phần đầu lát cũng có thể báo hiệu rõ ràng các hình ảnh được bao gồm trong RPS. RPS được xác định là tập hợp hình ảnh tham chiếu gắn với hình ảnh, bao gồm tất cả các hình ảnh tham chiếu trước hình ảnh gắn kèm theo thứ tự giải mã, có thể được sử dụng để dự báo liên cấu trúc các khối trong hình ảnh gắn kèm hoặc hình ảnh bất kỳ sau hình ảnh gắn kèm theo thứ tự giải mã, ví dụ, đến khi hình ảnh làm mới giải mã tức thời (instantaneous decoding refresh: IDR) tiếp theo, hoặc hình ảnh truy cập đường liên kết gãy (broken link access: BLA). Nói cách khác, các hình ảnh tham chiếu trong tập hợp hình ảnh tham chiếu có thể cần các đặc điểm sau đây: (1) tất cả các hình ảnh này đều trước hình ảnh hiện thời theo thứ tự giải mã, và (2) chúng có thể được sử dụng để dự báo liên cấu trúc hình ảnh hiện thời và/hoặc dự báo liên cấu trúc hình ảnh bất kỳ sau hình ảnh hiện thời theo thứ tự giải mã, và trong một số ví dụ, đến khi hình ảnh IDR tiếp theo hoặc hình ảnh BLA tiếp theo.

Bộ mã hóa video 20 hoặc bộ giải mã video 30 có thể suy ra KPS, và sau khi suy ra, có thể thiết lập danh sách hình ảnh tham chiếu, ví dụ, Danh sách 0 và Danh sách 1. Bộ mã hóa video 20 hoặc bộ giải mã video 30 có thể sử dụng các hình ảnh tham chiếu từ Danh sách 0 khi dự báo lát p, và các hình ảnh tham chiếu từ Danh sách 0 hoặc Danh sách 1 khi dự báo lát được dự báo hai chiều. Chỉ các hình ảnh tham chiếu trong RPS có thể là các hình ảnh tham chiếu ứng viên được sử dụng để thiết lập danh sách hình ảnh tham chiếu.

Để thiết lập tập hợp hình ảnh tham chiếu, bộ mã hóa video có thể thiết lập nhiều tập con hình ảnh tham chiếu. WD 9 mô tả 5 tập con hình ảnh tham chiếu mà bộ giải mã video thiết lập. Tổ hợp của các tập con hình ảnh tham chiếu có thể cùng tạo ra tập

hợp hình ảnh tham chiếu. Ví dụ, bộ mã hóa video 20 có thể báo hiệu rõ ràng, trong dòng bit được mã hóa, các giá trị cho phép bộ giải mã video xác định các ký hiệu nhận dạng cho các hình ảnh tham chiếu được bao gồm trong tập hợp hình ảnh tham chiếu này. Ví dụ, các ký hiệu nhận dạng của các hình ảnh tham chiếu có thể là số đếm thứ tự hình ảnh (POC). Mỗi hình ảnh được gắn với một giá trị POC. Giá trị POC của hình ảnh chỉ báo thứ tự xuất ra hoặc thứ tự hiển thị của hình ảnh tương ứng so với hình ảnh IDR trước đó theo thứ tự giải mã, và trong một số phương án khác, chỉ báo vị trí của hình ảnh gắn kèm theo thứ tự xuất ra so với vị trí thứ tự xuất ra của các hình ảnh còn lại trong cùng một chuỗi video được mã hóa. Ví dụ, trong chuỗi video được mã hóa, hình ảnh có giá trị POC nhỏ hơn được xuất ra hoặc hiển thị sớm hơn so với hình ảnh có giá trị POC lớn hơn.

Như được mô tả trên đây, các hình ảnh tham chiếu được lưu trữ trong DPB, và có thể được báo hiệu trong SPS, PPS hoặc phần đầu lát. PPS có thể kế thừa tập con hình ảnh tham chiếu từ SPS, và có thể báo hiệu các hình ảnh tham chiếu bổ sung mà có thể được sử dụng để báo hiệu các hình ảnh mà bộ mã hóa video 20 hoặc bộ giải mã video 30 có thể sử dụng trong quá trình dự báo liên cấu trúc hình ảnh hiện thời. Bộ mã hóa video có thể xác định ký hiệu nhận dạng cho hình ảnh tham chiếu, ví dụ, các giá trị POC hoặc các ký hiệu nhận dạng khác và từ các ký hiệu nhận dạng này thiết lập các tập con hình ảnh tham chiếu.

Có thể có nhiều cách khác nhau trong đó bộ giải mã video 30 có thể xác định các ký hiệu nhận dạng cho các hình ảnh thuộc tập hợp hình ảnh tham chiếu. Nói chung, bộ mã hóa video 20 có thể báo hiệu các giá trị mà từ đó bộ giải mã video 30 có thể xác định các ký hiệu nhận dạng cho các hình ảnh, bao gồm các hình ảnh thuộc tập hợp hình ảnh tham chiếu. Các ký hiệu nhận dạng của các hình ảnh có thể là POC cho mỗi hình ảnh này. Như được mô tả trên đây, giá trị POC có thể chỉ báo thứ tự hiển thị hoặc thứ tự xuất ra của hình ảnh, trong đó các hình ảnh có giá trị POC nhỏ hơn được hiển thị sớm hơn so với các hình ảnh có giá trị POC lớn hơn. Giá trị POC của hình ảnh đã cho có thể là so với hình ảnh IDR trước đó. Ví dụ, PicOrderCnt (nghĩa là giá trị POC) cho hình ảnh IDR có thể là 0, giá trị POC cho hình ảnh sau hình ảnh IDR theo thứ tự hiển thị hoặc thứ tự xuất ra có thể là 1, giá trị POC cho hình ảnh sau hình ảnh có giá trị POC 1 theo thứ tự hiển thị hoặc thứ tự xuất ra có thể là 2, v.v..

Từ các tập con hình ảnh tham chiếu này, bộ giải mã video có thể suy ra tập hợp hình ảnh tham chiếu, như được mô tả cụ thể hơn dưới đây. Trong một số ví dụ, mỗi tập con trong số các tập con hình ảnh tham chiếu bao gồm các hình ảnh tham chiếu khác nhau, trong đó không có sự chồng lấp các hình ảnh tham chiếu trong các tập con hình ảnh tham chiếu. Theo cách này, mỗi trong số các hình ảnh tham chiếu này có thể chỉ nằm trong một trong số các tập con hình ảnh tham chiếu, và không nằm trong tập con hình ảnh tham chiếu khác.

Sau khi xác định các ký hiệu nhận dạng (ví dụ, giá trị POC) của các hình ảnh tham chiếu trong tập hợp hình ảnh tham chiếu hoặc các tập con của nó, bộ giải mã video có thể thiết lập các tập con hình ảnh tham chiếu. Như được mô tả chi tiết hơn dưới đây, bộ giải mã video có thể thiết lập năm tập con hình ảnh tham chiếu, mặc dù bộ giải mã video có thể thiết lập nhiều hơn hoặc ít hơn tập con hình ảnh tham chiếu.

Năm tập con hình ảnh tham chiếu này được gọi là: RefPicSetStCurrBefore, RefPicSetStCurrAfter, RefPicSetStFoll, RefPicSetLtCurr và RefPicSetLtFoll.

Các tập con hình ảnh tham chiếu RefPicSetStCurrBefore, RefPicSetStCurrAfter, và RefPicSetStFoll có thể nhận dạng các hình ảnh tham chiếu ngắn hạn. Trong một số ví dụ, các tập con hình ảnh tham chiếu này có thể nhận dạng các hình ảnh tham chiếu ngắn hạn dựa vào việc các hình ảnh tham chiếu ngắn hạn này là sớm hơn theo thứ tự hiển thị hay muộn hơn theo thứ tự hiển thị so với hình ảnh hiện thời đang được mã hóa, cũng như việc các hình ảnh tham chiếu ngắn hạn có thể được sử dụng cho việc dự báo liên cấu trúc hình ảnh hiện thời và các hình ảnh sau hình ảnh hiện thời theo thứ tự giải mã, hay có thể có khả năng được sử dụng để chỉ dự báo liên cấu trúc các hình ảnh sau hình ảnh hiện thời theo thứ tự giải mã.

Ví dụ, tập con hình ảnh tham chiếu RefPicSetStCurrBefore có thể bao gồm, và có thể chỉ bao gồm thông tin nhận dạng, như giá trị POC, của tất cả các hình ảnh tham chiếu ngắn hạn mà có thứ tự hiển thị hoặc thứ tự xuất ra sớm hơn so với hình ảnh hiện thời, và có thể có khả năng được sử dụng để tham chiếu trong dự báo liên cấu trúc hình ảnh hiện thời, và có thể có khả năng được sử dụng để tham chiếu trong dự báo liên cấu trúc một hoặc nhiều hình ảnh sau hình ảnh hiện thời theo thứ tự giải mã. Tập con hình ảnh tham chiếu RefPicSetStCurrAfter có thể bao gồm, và có thể chỉ bao gồm, thông tin nhận dạng của tất cả các hình ảnh tham chiếu ngắn hạn mà có thứ tự xuất ra hoặc hiển thị muộn hơn so với hình ảnh hiện thời và có thể có khả năng được sử dụng

để tham chiếu trong dự báo liên cấu trúc hình ảnh hiện thời, và có thể có khả năng được sử dụng để tham chiếu trong dự báo liên cấu trúc một hoặc nhiều hình ảnh sau hình ảnh hiện thời theo thứ tự giải mã.

Tập con hình ảnh tham chiếu RefPicSetStFoll có thể bao gồm, và có thể chỉ bao gồm, thông tin nhận dạng của tất cả các hình ảnh tham chiếu ngắn hạn có thứ tự xuất ra hoặc thứ tự hiển thị sớm hơn so với hình ảnh hiện thời, có thể có khả năng được sử dụng để tham chiếu trong dự báo liên cấu trúc một hoặc nhiều hình ảnh sau hình ảnh hiện thời theo thứ tự giải mã, và không thể được sử dụng để tham chiếu trong dự báo liên cấu trúc hình ảnh hiện thời.

Tập con hình ảnh tham chiếu RefPicSetLtCurr và RefPicSetLtFoll có thể nhận dạng các hình ảnh tham chiếu dài hạn. Trong một số ví dụ, các tập con hình ảnh tham chiếu này có thể nhận dạng các hình ảnh tham chiếu dài hạn dựa vào việc các hình ảnh tham chiếu dài hạn là sớm hơn theo thứ tự hiển thị hay muộn hơn theo thứ tự hiển thị so với hình ảnh hiện thời đang được mã hóa.

Ví dụ, tập con hình ảnh tham chiếu RefPicSetLtCurr có thể bao gồm, và có thể chỉ bao gồm, thông tin nhận dạng của tất cả các hình ảnh tham chiếu dài hạn có thể có khả năng được sử dụng để tham chiếu trong dự báo liên cấu trúc hình ảnh hiện thời, và có thể có khả năng được sử dụng để tham chiếu trong dự báo liên cấu trúc một hoặc nhiều hình ảnh sau hình ảnh hiện thời theo thứ tự giải mã. Tập con hình ảnh tham chiếu RefPicSetLtFoll có thể bao gồm, và có thể chỉ bao gồm, thông tin nhận dạng của tất cả các hình ảnh tham chiếu dài hạn mà có thể có khả năng được sử dụng để tham chiếu trong dự báo liên cấu trúc một hoặc nhiều hình ảnh sau hình ảnh hiện thời theo thứ tự giải mã, và không thể được sử dụng để tham chiếu trong dự báo liên cấu trúc hình ảnh hiện thời.

Trong một số ví dụ, các tập con RefPicSetStCurrBefore, RefPicSetStCurrAfter, và RefPicSetLtCurr có thể bao gồm tất cả các hình ảnh tham chiếu mà có thể được sử dụng trong dự báo liên cấu trúc đối với khối trong hình ảnh hiện thời và có thể được sử dụng trong dự báo liên cấu trúc một hoặc nhiều hình ảnh sau hình ảnh hiện thời theo thứ tự giải mã. Các tập con RefPicSetStFoll, và RefPicSetLtFoll có thể bao gồm tất cả các hình ảnh tham chiếu mà không được sử dụng trong dự báo liên cấu trúc đối với khối trong hình ảnh hiện thời, nhưng có thể được sử dụng trong dự báo liên cấu trúc một hoặc nhiều hình ảnh sau hình ảnh hiện thời theo thứ tự giải mã.

Để thiết lập tập con hình ảnh tham chiếu RefPicSetLtCurr và RefPicSetLtAfter, bộ giải mã video thao tác lặp lại qua DPB và bổ sung các LTRP được lưu trữ trong DPB vào tập con hình ảnh tham chiếu RefPicSetLtCurr và RefPicSetLtAfter theo giả mã sau đây:

```

for( i = 0; i < NumPocLtCurr; i++ )
    if( !CurrDeltaPocMsbPresentFlag[ i ] )
        if( there is a long term reference picture picX in the DPB with
            pic_order_cnt_lsb equal to PocLtCurr[ i ] )
            RefPicSetLtCurr[ i ] = picX
        else if( there is a short-term reference picture picY in the DPB
            with pic_order_cnt_lsb equal to PocLtCurr[ i ] )
            RefPicSetLtCurr[ i ] = picY
        else
            RefPicSetLtCurr[ i ] = "no reference picture"
    else
        if( there is a long term reference picture picX in the DPB
            with PicOrderCntVal equal to PocLtCurr[ i ] )
            RefPicSetLtCurr[ i ] = picX
        else if( there is a short-term reference picture picY in the DPB
            with pic_order_cnt_lsb equal to PocLtCurr[ i ] )
            RefPicSetLtCurr[ i ] = picY
for( i = 0; i < NumPocLtFoll; i++ )
    if( !FollDeltaPocMsbPresentFlag[ i ] )
        if( there is a long term reference picture picX in the DPB
            with pic_order_cnt_lsb equal to PocLtFoll[ i ] )
            RefPicSetLtFoll[ i ] = picX
        else if( there is a short-term reference picture picY in the DPB
            with pic_order_cnt_lsb equal to PocLtFoll[ i ] )
            RefPicSetLtFoll[ i ] = picY
        else
            RefPicSetLtFoll[ i ] = "no reference picture"
    else
        if( there is a long term reference picture picX in the DPB
            with PicOrderCntVal to PocLtFoll[ i ] )
            RefPicSetLtFoll[ i ] = picX

```

```

else if( there is a short-term reference picture picY in the
        DPB with PicOrderCntVal equal to PocLtFoll[ i ] )
    RefPicSetLtFoll[ i ] = picY
else
    RefPicSetLtFoll[ i ] = "no reference picture"

```

Sau khi xây dựng các danh sách hình ảnh tham chiếu RefPicSetLtCurr, và RefPicSetLtFoll, bộ giải mã video 30 đánh dấu mỗi số các hình ảnh trong các hình ảnh được tham chiếu là “được sử dụng để tham chiếu dài hạn”.

Để thiết lập các tập hợp hình ảnh tham chiếu RefPicSetStCurrBefore, RefPicSetStCurrAfter, và RefPicSetLtCurr, bộ giải mã video 30 có thể thực hiện các giả mã sau đây, quy trình này lặp lại qua DPB, và bổ sung các hình ảnh từ DPB với các giá trị POC vào tập hợp hình ảnh tham chiếu tương ứng:

```

for( i = 0; i < NumPocStCurrBefore; i++ )
    if( there is a short-term reference picture picX in the DPB
        with PicOrderCntVal equal to PocStCurrBefore[ i ] )
        RefPicSetStCurrBefore[ i ] = picX
    else
        RefPicSetStCurrBefore[ i ] = "no reference picture"

for( i = 0; i < NumPocStCurrAfter; i++ )
    if( there is a short-term reference picture picX in the DPB
        with PicOrderCntVal equal to PocStCurrAfter[ i ] )
        RefPicSetStCurrAfter[ i ] = picX
    else
        RefPicSetStCurrAfter[ i ] = "no reference picture" (8-7)

for( i = 0; i < NumPocStFoll; i++ )
    if( there is a short-term reference picture picX in the DPB
        with PicOrderCntVal equal to PocStFoll[ i ] )
        RefPicSetStFoll[ i ] = picX
    else
        RefPicSetStFoll[ i ] = "no reference picture"

```

Bộ giải mã video 30 đánh dấu các hình ảnh tham chiếu bất kỳ mà không được bao gồm trong 1 trong 5 tập con hình ảnh tham chiếu là “không được sử dụng để tham chiếu”.

Nếu hình ảnh hiện thời cần được giải mã là hình ảnh IDR, bộ giải mã video 30 có thể đặt các tập con hình ảnh tham chiếu RefPicSetStCurrBefore, RefPicSetStCurrAfter, RefPicSetStFoll, RefPicSetLtCurr, và RefPicSetLtFoll là rỗng. Điều này có thể là vì hình ảnh IDR không thể được dự báo liên cấu trúc và không có hình ảnh nào sau hình ảnh IDR theo thứ tự giải mã có thể sử dụng hình ảnh 'bất kỳ trước hình ảnh IDR theo thứ tự giải mã để tham chiếu. Nếu không (ví dụ, khi hình ảnh hiện thời là hình ảnh không IDR), bộ giải mã video 30 có thể thiết lập các tập con hình ảnh tham chiếu ngắn hạn và các tập con hình ảnh tham chiếu dài hạn bằng cách thực hiện giả mã sau đây.

Khi bộ giải mã video 30 suy ra tập hợp hình ảnh tham chiếu từ nhiều tập con hình ảnh tham chiếu, bộ giải mã video có thể thiết lập các danh sách hình ảnh tham chiếu (ví dụ, Danh sách 0 và Danh sách 1) từ tập hợp hình ảnh tham chiếu. Ví dụ, việc thiết lập các danh sách hình ảnh tham chiếu có thể bao gồm bước khởi tạo và có thể là bước sửa đổi.

Trong một số ví dụ, bộ giải mã video có thể được tạo cấu hình để thực hiện kỹ thuật thiết lập danh sách tham chiếu mặc định, trong đó bộ giải mã video sử dụng tập con hình ảnh tham chiếu để thiết lập các danh sách hình ảnh tham chiếu ban đầu. Sau đó, nếu việc sửa đổi danh sách hình ảnh tham chiếu là không cần thiết, thì danh sách hình ảnh tham chiếu cuối cùng có thể giống như danh sách hình ảnh tham chiếu ban đầu, mà không cần bước sắp xếp lại thêm bất kỳ đối với các danh sách hình ảnh tham chiếu.

Trong quá trình dự báo liên cấu trúc PU được dự báo một chiều, bộ mã hóa video 20 hoặc bộ giải mã video 30 sử dụng các hình ảnh tham chiếu của danh sách hình ảnh tham chiếu 0 để dự báo các giá trị hiện thời và các giá trị độ chói-màu sắc của khối hiện thời. Trong quá trình dự báo liên cấu trúc PU được dự báo hai chiều, bộ mã hóa video 20 hoặc bộ giải mã video 30 sử dụng các hình ảnh tham chiếu của cả danh sách hình ảnh tham chiếu 0 và danh sách hình ảnh tham chiếu 1 để dự báo các giá trị của các khối độ chói và màu sắc hiện thời.

Bộ giải mã video 30 có thể gửi các hình ảnh tham chiếu, như LTRP và STRP bằng ghi chỉ số vào một trong hai danh sách hình ảnh tham chiếu, được gọi là RefPicList0, và RefPicList1, RefPicList0 và RefPicList1 bao gồm các hình ảnh tham chiếu từ 5 tập con hình ảnh tham chiếu, RefPicSetStCurrBefore, RefPicSetStCurrAfter, RefPicSetStFoll, RefPicSetLtCurr và RefPicSetLtFoll. Khi giải mã PU được dự báo một chiều, bộ giải mã video 30 có thể sử dụng một trong số các hình ảnh được lưu trữ trong RefPicList0 để dự báo các giá trị của các khối độ chói và màu sắc hiện thời. Khi giải mã PU được dự báo hai chiều, bộ giải mã video 30 có thể sử dụng hình ảnh của RefPicList0 và hình ảnh của RefPicList1 để dự báo các CU của lát hiện thời. Khi bắt đầu giải mã lát P hoặc B, bộ giải mã video 30 có thể thiết lập danh sách hình ảnh tham chiếu tạm thời, RefPicList0, theo giả mã sau đây:

```

rIdx = 0
while( rIdx < NumRpsCurrTempList0 ) {
    for( i = 0; i < NumPocStCurrBefore && rIdx < NumRpsCurrTempList0;
        rIdx++, i++ )
        RefPicListTemp0[ rIdx ] = RefPicSetStCurrBefore[ i ]
    for( i = 0; i < NumPocStCurrAfter && rIdx < NumRpsCurrTempList0;
        rIdx++, i++ )
        RefPicListTemp0[ rIdx ] = RefPicSetStCurrAfter[ i ]
    for( i = 0; i < NumPocLtCurr && rIdx < NumRpsCurrTempList0;
        rIdx++, i++ )
        RefPicListTemp0[ rIdx ] = RefPicSetLtCurr[ i ]
}

```

trong đó RefPicListTemp0 là biến tạm thời bao gồm các nội dung là RefPicSetStCurrBefore, sau đó là RefPicSetStCurrAfter, và RefPicSetLtCurr.

Trong một số trường hợp, dòng bit có thể báo hiệu là các hình ảnh trong RefPicList0 nên được xếp thứ tự lại. Bộ giải mã video 30 có thể xếp thứ tự lại RefPicListTemp0 dựa vào ref_pic_list_modification_flag_l0. Sau khi thiết lập RefPicListTemp0, bộ giải mã video 30 thiết lập RefPicList0 theo giả mã sau đây:

```

for( rIdx = 0; rIdx <= num_ref_idx_l0_active_minus1; rIdx++ )
    RefPicList0[ rIdx ] = ref_pic_list_modification_flag_l0 ?
    RefPicListTemp0[ list_entry_l0[ rIdx ] ] :
    RefPicListTemp0[ rIdx ]

```

Khi PU hiện thời được dự báo hai chiều, bộ giải mã video 30 thiết lập phiên bản tạm thời của RefPicList1, được gọi là RefPicListTemp1 theo giả mã sau đây trong quy trình tương tự được sử dụng để thiết lập RefPicListTemp0:

```

rIdx = 0
while( rIdx < NumRpsCurrTempList1 ) {
    for( i = 0; i < NumPocStCurrAfter && rIdx < NumRpsCurrTempList1;
        rIdx++, i++ )
        RefPicListTemp1[ rIdx ] = RefPicSetStCurrAfter[ i ]
    for( i = 0; i < NumPocStCurrBefore && rIdx < NumRpsCurrTempList1;
        rIdx++, i++ )
        RefPicListTemp1[ rIdx ] = RefPicSetStCurrBefore[ i ]
    for( i = 0; i < NumPocLtCurr && rIdx < NumRpsCurrTempList1;
        rIdx++, i++ )
        RefPicListTemp1[ rIdx ] = RefPicSetLtCurr[ i ]
}

```

Bộ giải mã video 30 có thể sắp xếp lại thứ tự RefPicListTemp1 dựa vào ref_pic_list_modification_flag_l1 theo giả mã sau đây:

```

for( rIdx = 0; rIdx <= num_ref_idx_l1_active_minus1; rIdx++ )
    RefPicList1[ rIdx ] = ref_pic_list_modification_flag_l1 ?
    RefPicListTemp1[ list_entry_l1[ rIdx ] ] :
    RefPicListTemp1[ rIdx ]

```

Như được mô tả trên đây, dòng bit HEVC bao gồm các đơn vị mã hóa (CU) bao gồm các đơn vị dự báo (PU) và các đơn vị biến đổi (TU). Đối với các PU được dự báo liên cấu trúc, mỗi đơn vị dự báo có thể bao gồm thông tin dự báo chuyển động, như chế độ dự báo liên cấu trúc cụ thể của PU, một hoặc nhiều giá trị chênh lệch vector chuyển động (motion vector difference value: MVD), và các hình ảnh tham chiếu sẽ được sử dụng khi dự báo các giá trị điểm ảnh của CU này. Có thể có một hoặc hai hình ảnh tham chiếu được báo hiệu cho mỗi PU, một hình ảnh tham chiếu nếu CU được dự báo một chiều, và hai hình ảnh tham chiếu nếu CU được dự báo hai chiều.

PU được dự báo liên cấu trúc của lát có thể bao gồm các phần tử cú pháp, ref_idx_10, và ref_idx_11. Giá trị của mỗi ref_idx_10 và ref_idx_11 lần lượt là các chỉ số trong các danh sách hình ảnh tham chiếu, RefPicList0, và RefPicList1. ref_idx_10 và ref_idx_11 báo hiệu các hình ảnh tham chiếu mà bộ giải mã video sử dụng kết hợp

với các giá trị chênh lệch vector chuyển động để xác định tập hợp các mẫu dự báo mà bộ giải mã video 30 sử dụng để dự báo các giá trị màu sắc và độ chói của khối hiện thời.

Trong trường hợp khi PU được dự báo một chiều hoặc hai chiều, bộ giải mã video 30 sử dụng các giá trị MVD của PU để khôi phục một hoặc nhiều vector chuyển động màu sắc và vector chuyển động độ chói. Bộ giải mã video 30 có thể khôi phục các vector chuyển động độ chói và màu sắc bằng cách xác định biến độc lập vector chuyển động cho CU và kết hợp biến độc lập vector chuyển động với MVD để tạo ra vector chuyển động cho CU. Bộ giải mã video 30 có thể xác định biến độc lập vector chuyển động từ CU khác được mã hóa trước đó, ví dụ, CU lân cận, được mã hóa trước. Trong một số trường hợp, PU có thể được mã hóa bằng cách sử dụng chế độ hợp nhất hoặc dự báo vector chuyển động tiên tiến (advanced motion vector prediction: AMVP), trong đó bộ giải mã video 30 xác định biến độc lập vector chuyển động cho PU hiện thời chỉ dựa vào một hoặc nhiều vector chuyển động của các CU được mã hóa trước đó. Trong trường hợp này, PU hiện thời không bao gồm các giá trị MVD được báo hiệu bất kỳ.

Khi bộ giải mã video 30 đã khôi phục các vector chuyển động màu sắc và độ chói cho PU hiện thời, bộ giải mã video 30 xác định dãy mẫu dự báo độ chói, dãy mẫu dự báo màu sắc, và dãy mẫu màu sắc dựa vào một hoặc nhiều hình ảnh tham chiếu được báo hiệu bởi các chỉ số của chúng trong PU. Để tính toán các mẫu dự báo độ chói (ví dụ, các giá trị dự báo) cho khối độ chói hiện thời, bộ giải mã video 30 tính toán phần nguyên và phần phân số của vector chuyển động độ chói. Bộ giải mã video 30 bổ sung các tọa độ x và y của góc trái trên của khối độ chói hiện thời với vector chuyển động độ chói để xác định độ dịch nguyên và phân số tương ứng cho các mẫu độ chói của hình ảnh tham chiếu được chỉ báo. Dựa vào độ dịch này, bộ mã hóa video 20 có thể nội suy các giá trị của các vị trí phụ phân số dựa vào các giá trị điểm ảnh nguyên vẹn của các mẫu độ chói của hình ảnh tham chiếu để xác định các mẫu dự báo độ chói cho khối độ chói hiện thời. Ví dụ, chênh lệch vector chuyển động độ chói, khi kết hợp với biến độc lập vector chuyển động, có thể tham chiếu vị trí $\frac{1}{4}$ điểm ảnh cụ thể của các mẫu dự báo độ chói. Bộ giải mã video 30 có thể nội suy các giá trị $\frac{1}{4}$ điểm ảnh dựa vào một hoặc nhiều bộ lọc nội suy. Các bộ lọc nội suy có thể sử dụng các tập hợp điểm ảnh nguyên vẹn hỗ trợ từ các mẫu tham chiếu độ chói để tạo ra các giá trị $\frac{1}{4}$ điểm ảnh.

Bộ giải mã video 30 có thể sử dụng các giá trị độ chói $\frac{1}{4}$ điểm ảnh đã tạo ra làm giá trị dự báo của khối độ chói hiện thời.

Bộ giải mã video 30 có thể thực hiện quy trình tương tự để xác định các mẫu dự báo cho khối màu sắc hiện thời. Bộ giải mã video 30 khôi phục ít nhất một vector chuyển động màu sắc, có thể khác với vector chuyển động độ chói, và sử dụng vector chuyển động màu sắc để xác định độ dịch trong các mẫu màu sắc của hình ảnh tham chiếu. Ngoài ra, trong khi vector chuyển động độ chói có thể có độ chính xác tới $\frac{1}{4}$ điểm ảnh, vector chuyển động màu sắc có thể có độ chính xác tới $\frac{1}{8}$ điểm ảnh. Dựa vào độ dịch trong các mẫu màu sắc của hình ảnh tham chiếu, bộ giải mã video 30 có thể nội suy các giá trị màu sắc dưới điểm ảnh dự báo làm biến độc lập của các giá trị màu sắc của khối màu sắc hiện thời.

Trong các trường hợp khi PU hiện thời được dự báo hai chiều, bộ giải mã video 30 xác định 2 dãy mẫu dự báo độ chói, một cho mỗi trong số 2 vector chuyển động độ chói, và hai dãy mẫu dự báo màu sắc, một cho mỗi trong số 2 vector chuyển động màu sắc. Bộ giải mã video 30 kết hợp hai mẫu dự báo độ chói này và hai dãy mẫu dự báo màu sắc để tạo ra một dãy đơn các mẫu dự báo. Kỹ thuật được sử dụng để dự báo mỗi dãy mẫu độ chói và màu sắc phụ thuộc vào hai phần tử cú pháp cờ dự báo, `predFlagL0` và `predFlagL1`, chỉ báo bộ giải mã video 30 thực hiện quy trình nào trong số hai quy trình dự báo mẫu có trọng số. Trọng số riêng được gán cho mỗi một trong hai dãy mẫu dự báo màu sắc và độ chói được xác định bởi bảng cú pháp dự báo có trọng số, mà có thể được báo hiệu trong phần đầu đoạn lát.

Sau khi mã hóa dự báo nội cấu trúc hoặc liên cấu trúc bằng cách sử dụng các PU của CU, bộ mã hóa video 20 hoặc bộ giải mã video 30 có thể tính toán dữ liệu dư cho các TU của CU. Các PU này có thể bao gồm dữ liệu cú pháp mô tả phương pháp hoặc chế độ tạo ra dữ liệu điểm ảnh dự báo trong miền không gian (còn được gọi là miền điểm ảnh) và các TU có thể bao gồm các hệ số trong miền biến đổi sau khi áp dụng biến đổi, ví dụ, biến đổi cosin rời rạc (discrete cosine transform: DCT), biến đổi số nguyên, biến đổi sóng con, hoặc biến đổi tương tự về khái niệm đối với dữ liệu video dư. Dữ liệu dư có thể tương ứng với chênh lệch điểm ảnh giữa các điểm ảnh của hình ảnh chưa được mã hóa và các giá trị dự báo tương ứng với các PU. Bộ mã hóa video 20 hoặc bộ giải mã video 30 có thể tạo ra các TU bao gồm dữ liệu dư cho CU, và sau đó biến đổi các TU để tạo ra hệ số biến đổi cho CU.

Sau biến đổi bất kỳ để tạo ra các hệ số biến đổi, bộ mã hóa video 20 hoặc bộ giải mã video 30 có thể thực hiện lượng tử hóa các hệ số biến đổi. Việc lượng tử hóa thường chỉ quy trình trong đó các hệ số biến đổi được lượng tử hóa để có thể làm giảm lượng dữ liệu được sử dụng để biểu diễn các hệ số, tạo ra sự nén thêm. Quá trình lượng tử hóa có thể làm giảm độ sâu bit liên quan đến một số hoặc toàn bộ các hệ số. Ví dụ, giá trị n bit có thể được làm tròn xuống giá trị m bit trong quá trình lượng tử hóa, trong đó n lớn hơn m .

Sau khi lượng tử hóa, bộ mã hóa video có thể quét các hệ số biến đổi, tạo ra vector một chiều từ ma trận hai chiều bao gồm các hệ số biến đổi được lượng tử hóa. Việc quét này có thể được thiết kế để đặt các hệ số năng lượng cao hơn (và do đó tần số thấp hơn) ở trước dãy này và đặt các hệ số năng lượng thấp hơn (và do đó tần số cao hơn) ở sau dãy này. Trong một số ví dụ, bộ mã hóa video 20 hoặc bộ giải mã video 30 có thể sử dụng kỹ thuật quét xác định trước để quét các hệ số biến đổi được lượng tử hóa để tạo ra vector nối tiếp có thể được mã hóa entropy. Trong các ví dụ khác, bộ mã hóa video 20 hoặc bộ giải mã video 30 có thể thực hiện quét thích ứng. Sau khi quét, hệ số biến đổi được lượng tử hóa để tạo ra vector một chiều, bộ mã hóa video 20 hoặc bộ giải mã video 30 có thể mã hóa entropy vector một chiều, ví dụ, theo phương pháp mã hóa độ dài biến đổi thích ứng ngữ cảnh (context-adaptive variable length coding: CAVLC), mã hóa số học nhị phân thích ứng ngữ cảnh (context-adaptive binary arithmetic coding: CABAC), mã hóa số học nhị phân thích ứng ngữ cảnh dựa vào cú pháp (syntax-based context-adaptive binary arithmetic coding: SBAC), mã hóa entropy phân chia khoảng xác suất (Probability Interval Partitioning Entropy: PIPE) hoặc phương pháp mã hóa entropy khác. Bộ mã hóa video 20 cũng có thể mã hóa entropy các phần tử cú pháp liên quan đến dữ liệu video mã hóa để sử dụng bởi bộ giải mã video 30 trong việc giải mã dữ liệu video.

Để thực hiện CABAC, bộ mã hóa video 20 có thể gán ngữ cảnh trong mô hình ngữ cảnh cho ký hiệu được truyền. Ngữ cảnh này có thể liên quan đến, ví dụ, việc các giá trị lân cận của ký hiệu này có phải khác 0 hay không. Để thực hiện CAVLC, bộ mã hóa video 20 có thể chọn mã độ dài biến đổi cho ký hiệu được truyền. Các từ mã trong VLC có thể được thiết lập sao cho các mã tương đối ngắn hơn tương ứng với các ký hiệu có xác suất cao hơn, trong khi các mã dài hơn tương ứng với các ký hiệu có xác suất thấp hơn. Theo cách này, việc sử dụng VLC có thể đạt được sự tiết kiệm bit so

với, ví dụ, việc sử dụng các từ mã có độ dài bằng nhau cho mỗi ký hiệu được truyền. Việc xác định xác suất có thể dựa vào ngữ cảnh được gán cho ký hiệu này.

Bộ mã hóa video 20 còn có thể gửi dữ liệu cú pháp, như dữ liệu cú pháp dựa vào khối, dữ liệu cú pháp dựa vào khung, và dữ liệu cú pháp dựa vào GOP, cho bộ giải mã video 30, ví dụ, trong phần đầu khung, phần đầu khối, phần đầu lát, hoặc phần đầu GOP. Dữ liệu cú pháp GOP có thể mô tả số lượng khung trong GOP tương ứng, và dữ liệu cú pháp khung này có thể mô tả chế độ mã hóa/dự báo được sử dụng để mã hóa khung tương ứng.

Kỹ thuật theo sáng chế hướng tới các kỹ thuật để cải thiện việc báo hiệu các hình ảnh tham chiếu trong dòng bit video mã hóa. Nói chung, để tiết kiệm bit khi báo hiệu các tham chiếu đến các giá trị POC của các hình ảnh tham chiếu, WD6 chỉ báo rằng chỉ một số lượng nhất định các bit có trọng số nhỏ nhất (least-significant bit: LSB) của giá trị POC được báo hiệu để tham chiếu đến các hình ảnh tham chiếu dài hạn, trong khi đó các hình ảnh tham chiếu ngắn hạn được nhận dạng bởi giá trị POC đầy đủ.

Sáng chế đề xuất kỹ thuật liên quan đến báo hiệu các hình ảnh tham chiếu dài hạn mà bộ mã hóa video 20 và bộ giải mã video 30 có thể thực hiện được. Cụ thể là, sáng chế mô tả các kỹ thuật báo hiệu thông tin nhận dạng cho các hình ảnh tham chiếu dài hạn, ví dụ, trong phần đầu lát. Cụ thể là, các kỹ thuật theo sáng chế hướng tới các kỹ thuật cải thiện việc báo hiệu các bit có trọng số lớn nhất (most significant bit: MSB) và bit có trọng số nhỏ nhất của các hình ảnh tham chiếu dài hạn trong phần đầu lát của chuỗi video được mã hóa. Các kỹ thuật theo sáng chế cũng hướng tới kỹ thuật báo hiệu, trong phần đầu lát, các hình ảnh tham chiếu bổ sung mà không được bao gồm trong tập hợp các hình ảnh tham chiếu được bao gồm trong tập hợp tham số chuỗi hoặc tập hợp tham số hình ảnh hoạt động cho hình ảnh mã hóa hiện thời.

Khía cạnh thứ nhất của kỹ thuật theo sáng chế đề cập đến việc báo hiệu các hình ảnh tham chiếu dài hạn, và cụ thể hơn là báo hiệu các bit có trọng số lớn nhất của bit tham chiếu dài hạn trong phần đầu lát của chuỗi video được mã hóa. Nói chung, các kỹ thuật theo sáng chế có thể thực hiện cú pháp phần đầu lát như đã được mô tả liên quan đến Bảng 2 trên đây. Theo Bảng 2, phần đầu lát bao gồm các phần tử cú pháp `long_term_idx_sps[ i ]` và `poc_lsb_lt`. Đối với giá trị chỉ số cụ thể *i*, giá trị của `long_term_idx_sps[ i ]` chỉ rõ chỉ số vào danh sách các hình ảnh tham chiếu dài hạn

ứng viên từ tập hợp tham số chuỗi hoạt động, và giá trị của $poc_lsb_lt[i]$ chỉ rõ bit có trọng số nhỏ nhất cho hình ảnh tham chiếu dài hạn thứ i .

Theo cách này, các phần tử cú pháp $long_term_idx_sps$ và poc_lsb_lt nhận dạng tập con các hình ảnh tham chiếu dài hạn mà được kế thừa từ tập hợp tham số chuỗi này, và bộ giải mã video 30 sử dụng nó để dự báo các CU của lát hiện thời. Các kỹ thuật theo sáng chế giới hạn thứ tự của các giá trị LSB POC LTRP được kế thừa từ SPS trong phần tử cú pháp $long_term_idx_sps[i]$. Việc giới hạn thứ tự của các giá trị LSB POC LTRP của $long_term_idx_sps[i]$ theo các kỹ thuật của sáng chế có thể làm giảm độ phức tạp của dòng bit video được mã hóa. Theo đó, ngữ nghĩa của các phần tử cú pháp phần đầu lát của Bảng 2 có thể được sửa đổi như sau, so với ngữ nghĩa được mô tả liên quan đến Bảng 2 trên đây.

Đối với chỉ số cụ thể i , giá trị của $long_term_idx_sps[i]$ chỉ rõ chỉ số, đến danh sách các hình ảnh tham chiếu dài hạn được xác định trong tập hợp tham số chuỗi hoạt động, của hình ảnh tham chiếu dài hạn thứ nhất được kế thừa từ tập hợp tham số chuỗi được đề cập đến tập hợp hình ảnh tham chiếu dài hạn của hình ảnh hiện thời. Giá trị của $long_term_idx_sps[i]$ có thể nằm trong khoảng từ 0 đến $num_long_term_ref_pics_sps - 1$, bao gồm cả các giá trị đầu mút.

Các kỹ thuật theo sáng chế còn sửa đổi việc xếp thứ tự các giá trị của poc_lsb_lt , sao cho đối với các giá trị bất kỳ của j và k trong khoảng từ 0 đến $num_long_term_sps - 1$, bao gồm cả các giá trị đầu mút, nếu j nhỏ hơn k , $lt_ref_pic_poc_lsb_sps[long_term_idx_sps[j]]$ không nhỏ hơn $lt_ref_pic_poc_lsb_sps[long_term_idx_sps[k]]$. Trong ví dụ này, các giá trị POC LSB của $long_term_idx_sps[j]$ không nhỏ hơn các giá trị POC LSB của $long_term_idx_sps[k]$, đối với tất cả $j < k$.

Phần đầu lát cũng có thể bao gồm phần tử cú pháp $delta_poc_msb_present_flag[i]$. $delta_poc_msb_present_flag[i]$ chỉ rõ việc số lượng MSB nhất định có được báo hiệu cho LTRP thứ i , nghĩa là, LTRP được báo hiệu bởi $long_term_idx_sps[i]$ hoặc $poc_lsb_lt[i]$ hay không. Nếu chỉ số thứ i của $delta_poc_msb_present_flag[1]$ bằng 1, giá trị $delta_poc_msb_cycle_lt[i]$ có mặt. Nếu giá trị của $delta_poc_msb_present_flag[i]$ bằng 0, $delta_poc_msb_cycle_lt[i]$ không có mặt cho LTRP thứ i được báo hiệu. Các kỹ thuật theo sáng chế làm sửa đổi việc báo hiệu $delta_poc_msb_present_flag[i]$ sao cho $delta_poc_msb_present_flag[i]$ chỉ

được báo hiệu bằng 1 khi có nhiều hơn một hình ảnh trong bộ đệm hình ảnh được giải mã, và hình ảnh tham chiếu được đánh dấu là “được sử dụng để tham chiếu” và có bit có trọng số nhỏ nhất của giá trị số đếm thứ tự hình ảnh bằng $PocLsbLt[i]$. Các kỹ thuật trước đây có thể báo hiệu $\Delta poc_msb_present_flag[i]$ bằng 1, và các giá trị $\Delta poc_msb_lt[i]$ ngay cả khi hình ảnh này không được đánh dấu là “được sử dụng để tham chiếu”, nghĩa là, khi hình ảnh này không còn cần để tham chiếu. Bằng cách giới hạn việc báo hiệu $\Delta poc_msb_present_flag[i]$ ở các hình ảnh được đánh dấu là “được sử dụng để tham chiếu”, các kỹ thuật theo sáng chế có thể làm giảm tần suất báo hiệu các MSB của hình ảnh tham chiếu, mà có thể làm giảm độ phức tạp của dòng bit video được mã hóa.

Phần đầu lát cũng có thể bao gồm phần tử cú pháp $\Delta poc_msb_cycle_lt[i]$. Bộ giải mã video 30 có thể sử dụng giá trị $\Delta poc_msb_cycle_lt[i]$ để xác định giá trị của bit có trọng số lớn nhất của giá trị số đếm thứ tự hình ảnh của hình ảnh tham chiếu dài hạn thứ i mà được bao gồm trong tập hợp hình ảnh tham chiếu dài hạn của hình ảnh hiện thời. Để xác định giá trị MSB của LTRP thứ i , bộ giải mã video có thể xác định giá trị của biến $\Delta poc_msb_cycle_lt[i]$ chỉ báo MSB của LTRP thứ i .

Bộ giải mã video có thể suy ra biến $\Delta poc_msb_cycle_lt[i]$ theo giả mã sau đây:

```

if( i == 0 || i == num_long_term_sps || PocLsbLt[ i - 1 ] != PocLsbLt[ i
] )
    DeltaPocMSBCycleLt[ i ] = delta_poc_msb_cycle_lt[ i ]
else
    DeltaPocMSBCycleLt[ i ] = delta_poc_msb_cycle_lt[ i ] +
        DeltaPocMSBCycleLt[ i - 1 ]

```

Giá trị của $\Delta poc_msb_cycle_lt[i] * MaxPicOrderCntLsb + pic_order_cnt_lsb - PocLsbLt[i]$ có thể nằm trong khoảng từ 1 đến $2^{24} - 1$, bao gồm cả các giá trị đầu mút.

Trong một số trường hợp, bộ giải mã video 30 xác định $\Delta poc_msb_cycle_lt[i]$ bằng cách cộng giá trị của $\Delta poc_msb_cycle_lt[i-1]$, giá trị Δpoc_msb của LTRP trước đó, vào giá trị $\Delta poc_msb_cycle_lt[i]$ của LTRP hiện thời.

So với kỹ thuật trước đó để xác định $\Delta\text{PocMSBCycleLt}[i]$, $\Delta\text{PocMSBCycleLt}[i]$ được xác định là $\text{delta_poc_msb_cycle_lt}[i]$ nếu i bằng số lượng LTRP được chỉ rõ trong SPS hoạt động và được bao gồm trong tập hợp các LTRP cho hình ảnh hiện thời.

Trong ví dụ khác theo các kỹ thuật của sáng chế (có thể được kết hợp với ví dụ được thảo luận trên đây), cú pháp phần đầu lát có thể gần như giống trong bảng 2. Tuy nhiên, phần đầu lát cũng có thể bao gồm phần tử cú pháp, " $\text{lt_pic_from_sps_flag}[i]$ ", như được mô tả dưới đây. Thứ tự của các giá trị LSB POC được kế thừa từ các giá trị SPS và LSB POC được báo hiệu trực tiếp trong phần đầu lát có thể được giới hạn chung qua ràng buộc trên các giá trị của dải $\text{PocLsbLt}[i]$. Giới hạn trên $\text{delta_poc_msb_present_flag}[i]$ có thể là giống như trong ví dụ được mô tả trên đây. Bảng 5 cung cấp ví dụ về cú pháp cho phần đầu lát theo ví dụ này:

Bảng 5

slice_header() {	Descriptor
...	
if(IdrPicFlag) {	
idr_pic_id	ue(v)
no_output_of_prior_pics_flag	u(1)
} else {	
pic_order_cnt_lsb	u(v)
short_term_ref_pic_set_sps_flag	u(1)
if(!short_term_ref_pic_set_sps_flag)	
short_term_ref_pic_set(num_short_term_ref_pic_sets)	
else	
short_term_ref_pic_set_idx	u(v)
if(long_term_ref_pics_present_flag) {	
num_long_term_pics	ue(v)
if(num_long_term_ref_pics_sps)	
num_long_term_sps	ue(v)
for(i = 0; i < num_long_term_sps + num_long_term_pics; i++) {	
if(num_long_term_sps)	
lt_pic_from_sps_flag[i]	u(1)
if(lt_pic_from_sps_flag[i])	
long_term_idx_sps[i]	ue(v)
else	
poc_lsb_lt[i]	u(v)
delta_poc_msb_present_flag[i]	u(1)
if(delta_poc_msb_present_flag[i])	
delta_poc_msb_cycle_lt[i]	ue(v)
used_by_curr_pic_lt_flag[i]	u(1)
}	
}	
}	
...	
}	

Như được nêu trên đây, bảng 5 cung cấp phần tử cú pháp bổ sung, “lt_pic_from_sps_flag[i]”. Bộ giải mã video 30 có thể sử dụng lt_pic_from_sps_flag[i] để xác định việc LTRP thứ i được báo hiệu trong phần đầu lát được suy ra từ SPS hay được báo hiệu rõ ràng trong phần đầu lát. Khi cờ bằng 1, LTRP có thể được kế thừa từ SPS và tương ứng với chỉ số được báo hiệu. Khi cờ bằng 0, LSB của hình ảnh tham chiếu dài hạn có thể được báo hiệu rõ ràng. Khi không có mặt, bộ mã hóa video có thể suy ra lt_pic_from_sps_flag[i] bằng 0.

Theo cách này, lt_pic_from_sps_flag[i] thể hiện ví dụ về phần tử cú pháp, mà có thể được mã hóa trong phần đầu lát, mà chỉ báo việc thông tin nhận dạng cho hình ảnh tham chiếu dài hạn được báo hiệu rõ ràng trong phần đầu lát hay được suy ra từ

tập hợp tham số chuỗi bao gồm chuỗi hình ảnh, trong đó lát này xuất hiện (nghĩa là, tập hợp tham số chuỗi tương ứng với lát này).

Trong ví dụ này, `long_term_idx_sps[ i ]` có thể chỉ rõ chỉ số, đến danh sách các hình ảnh tham chiếu dài hạn được xác định trong tập hợp tham số chuỗi hoạt động, của hình ảnh tham chiếu dài hạn thứ `i` đến tập hợp hình ảnh tham chiếu dài hạn của hình ảnh hiện thời. Giá trị của `long_term_idx_sps[ i ]` có thể nằm trong khoảng từ 0 đến `num_long_term_ref_pics_sps - 1`, bao gồm cả các giá trị đầu mút.

Trong ví dụ này, `poc_lsb_lt[i]` có thể chỉ rõ giá trị của bit có trọng số nhỏ nhất của giá trị số đếm thứ tự hình ảnh của hình ảnh tham chiếu dài hạn thứ `i` mà được bao gồm trong tập hợp hình ảnh tham chiếu dài hạn của hình ảnh hiện thời. Độ dài của phần tử cú pháp `poc_lsb_lt[i]` có thể là `log2_max_pic_order_cnt_lsb_minus4 + 4` bit. Biến `PocLsbLt[i]` có thể chỉ rõ các LSB cho LTRP thứ `i`.

Biến `PocLsbLt[ i ]` có thể được suy ra như sau trong giả mã sau đây:

```
if(lt_pic_from_sps_flag[ i ])
    PocLsbLt[ i ] = lt_ref_pic_poc_lsb_sps [ long_term_idx_sps[ i ] ]
else
    PocLsbLt[ i ] = poc_lsb_lt[ i ]
```

Với các giá trị bất kỳ của `j` và `k` trong khoảng từ 0 đến `num_long_term_pics + num_long_term_pics - 1`, bao gồm cả các giá trị đầu mút, nếu `j` nhỏ hơn `k`, `PocLsbLt[ j ]` bị giới hạn sao cho không nhỏ hơn `PocLsbLt[ k ]`. Do đó, `PocLsbLt[ i ]` được xác định là các LSB của hình ảnh tham chiếu được kế thừa từ SPS nếu LTRP được kế thừa từ SPS, và LSB được báo hiệu rõ ràng trong phần đầu lát nếu LTRP không được kế thừa từ SPS. Thêm vào đó, các LSB được giới hạn theo thứ tự không tăng dần.

Các thay đổi về ngữ nghĩa của `delta_poc_msb_present_flag[i]` và `delta_poc_msb_cycle_lt[i]` có thể là giống như được mô tả trong ví dụ trên đây, và việc suy ra biến `DeltaPocMSBCycleLt[i]` có thể là giống như phương pháp liên quan đến Bảng 2 trên đây.

Ngoài ra, `num_long_term_pics` có thể chỉ rõ tổng số LTRP được báo hiệu bao gồm các LTRP được kế thừa từ SPS và LTRP được báo hiệu rõ ràng trong phần đầu lát, và bộ giải mã video không cần báo hiệu `num_long_term_sps`.

Bộ mã hóa video 20 và bộ giải mã video 30 có thể được cấu hình để mã hóa dữ liệu của phần đầu lát và/hoặc tập hợp tham số chuỗi theo các ví dụ khác nhau về cú

pháp và ngữ nghĩa được mô tả trên đây. Theo cách này, bộ mã hóa video 20 và bộ giải mã video 30 có thể thực hiện tiết kiệm bit, ví dụ, bằng cách báo hiệu bit có trọng số lớn nhất với số lượng ít hơn cho một số hình ảnh tham chiếu dài hạn.

Do đó, bộ mã hóa video 20 và bộ giải mã video 30 biểu diễn ví dụ về bộ giải mã video được tạo cấu hình để mã hóa phần đầu lát cho lát dữ liệu video, trong đó phần đầu lát bao gồm phần tử cú pháp chỉ báo việc thông tin nhận dạng cho hình ảnh tham chiếu dài hạn được báo hiệu rõ ràng trong phần đầu lát hay được suy ra từ tập hợp tham số chuỗi tương ứng với lát này, và khi phần tử cú pháp chỉ báo là thông tin nhận dạng cho hình ảnh tham chiếu dài hạn được báo hiệu rõ ràng, mã hóa giá trị cho LSB cho hình ảnh tham chiếu dài hạn trong phần đầu lát.

Mỗi bộ mã hóa video 20 và bộ giải mã video 30 có thể được thực hiện dưới dạng mạch bất kỳ trong số nhiều mạch mã hóa hoặc giải mã thích hợp, nếu áp dụng được, chẳng hạn như một hoặc nhiều bộ vi xử lý, bộ xử lý tín hiệu số (digital signal processor: DSP), mạch tích hợp chuyên dụng (application specific integrated circuits: ASICs), mảng cổng lập trình được dạng trường (field programmable gate array: FPGA), mạch logic rời rạc, phần mềm, phần cứng, phần sụn hoặc tổ hợp của chúng. Mỗi bộ mã hóa video 20 và bộ giải mã video 30 có thể được bao gồm trong một hoặc nhiều bộ mã hóa hoặc bộ giải mã, một trong số chúng có thể được tích hợp làm một phần của bộ mã hóa/giải mã video kết hợp (CODEC). Thiết bị bao gồm bộ mã hóa video 20 và/hoặc bộ giải mã video 30 có thể bao gồm mạch tích hợp, bộ vi xử lý, và/hoặc thiết bị truyền thông không dây, như điện thoại di động.

Fig.2 là sơ đồ khối minh họa ví dụ về bộ mã hóa video 20 có thể thực hiện các kỹ thuật để báo hiệu dữ liệu cho LTRP trong SPS hoặc phần đầu lát. Bộ mã hóa video 20 có thể thực hiện mã hóa nội cấu trúc hoặc mã hóa liên cấu trúc các khối video trong lát video. Việc mã hóa nội cấu trúc dựa vào dự báo không gian để làm giảm hoặc loại bỏ dư thừa không gian trong video trong khung video hoặc hình ảnh đã cho. Việc mã hóa liên cấu trúc dựa vào dự báo thời gian để làm giảm hoặc loại bỏ dư thừa thời gian trong video trong các khung hoặc hình ảnh liên kế của chuỗi video. Chế độ nội cấu trúc (chế độ I) có thể dùng để chỉ bất kỳ trong số vài chế độ mã hóa dựa vào không gian. Chế độ liên cấu trúc, như dự báo một chiều (chế độ P) hoặc dự báo hai chiều (chế độ B), có thể chỉ bất kỳ trong số vài chế độ mã hóa dựa vào thời gian.

Như được thể hiện trên Fig.2, bộ mã hóa video 20 nhận khối video hiện thời trong khung video cần được mã hóa. Trong ví dụ trên Fig.2, bộ mã hóa video 20 bao gồm bộ phận chọn chế độ 40, bộ nhớ hình ảnh tham chiếu 64, bộ cộng 50, bộ phận xử lý biến đổi 52, bộ phận lượng tử hóa 54 và bộ phận mã hóa entropy 56. Bộ phận chọn chế độ 40 lại bao gồm bộ phận bù chuyển động 44, bộ phận ước lượng chuyển động 42, bộ phận dự báo nội cấu trúc 46, và bộ phận chia 48. Để khôi phục khối video, bộ mã hóa video 20 cũng bao gồm bộ phận lượng tử hóa ngược 58, bộ phận biến đổi ngược 60 và bộ cộng 62. Bộ lọc tách khối (không được thể hiện trên Fig.2) cũng có thể được bao gồm vào các biên khối bộ lọc để loại bỏ các thành phần lạ dạng khối khối video được khôi phục. Nếu muốn, bộ lọc tách khối thường sẽ lọc đầu ra của bộ cộng 62. Các bộ lọc bổ sung (trong vòng hoặc sau vòng) cũng có thể được sử dụng ngoài bộ lọc tách khối. Các bộ lọc này không được thể hiện cho ngắn gọn, nhưng nếu muốn, có thể lọc đầu ra của bộ cộng 50 (như bộ lọc trong vòng).

Trong quá trình mã hóa, bộ mã hóa video 20 nhận khung video hoặc lát video cần được mã hóa. Khung hoặc lát có thể được chia thành nhiều khối video. Bộ phận ước lượng chuyển động 42 và bộ phận bù chuyển động 44 thực hiện mã hóa dự báo liên cấu trúc khối video thu được liên quan đến một hoặc nhiều khối trong một hoặc nhiều khung tham chiếu để cung cấp dự báo thời gian. Theo cách khác, bộ phận dự báo nội cấu trúc 46 có thể thực hiện mã hóa dự báo nội cấu trúc khối video nhận được liên quan đến một hoặc nhiều khối lân cận trong cùng một khung hoặc lát với khối cần được mã hóa để cung cấp dự báo không gian. Bộ mã hóa video 20 có thể thực hiện nhiều quy trình mã hóa, ví dụ để chọn chế độ mã hóa thích hợp cho mỗi khối dữ liệu video.

Ngoài ra, bộ phận chia 48 có thể chia các khối dữ liệu video thành các khối con, dựa vào đánh giá sơ đồ chia trước đó trong các quy trình mã hóa trước đó. Ví dụ bộ phận chia 48 ban đầu có thể chia khung hoặc lát thành các LCU, và chia mỗi LCU thành các CU con dựa vào phân tích tốc độ - méo (ví dụ, tối ưu hóa tốc độ - méo). Bộ phận chọn chế độ 40 còn có thể tạo ra cấu trúc dữ liệu cây tứ phân chỉ báo việc chia LCU thành các CU con. Các CU nút lá của cây tứ phân có thể bao gồm một hoặc nhiều PU và một hoặc nhiều TU.

Bộ phận chọn chế độ 40 có thể chọn một trong số các chế độ mã hóa, nội cấu trúc hoặc liên cấu trúc, ví dụ dựa vào các kết quả sai số, và cung cấp khối được mã hóa

nội cấu trúc hoặc liên cấu trúc kết quả cho bộ cộng 50 để tạo ra dữ liệu khối dư và cho bộ cộng 62 để khôi phục khối được mã hóa để sử dụng làm khung tham chiếu. Bộ phận chọn chế độ 40 cũng cung cấp các phần tử cú pháp, như vector chuyển động, chỉ báo chế độ nội cấu trúc, thông tin chia, và thông tin cú pháp khác tương tự, cho bộ phận mã hóa entropy 56.

Bộ phận ước lượng chuyển động 42 và bộ phận bù chuyển động 44 có thể được tích hợp cao nhưng được minh họa riêng với mục đích giải thích khái niệm. Ước lượng chuyển động, được thực hiện bởi bộ phận ước lượng chuyển động 42, là quy trình tạo vector chuyển động để ước lượng chuyển động cho khối video. Vector chuyển động, ví dụ, có thể chỉ báo sự dịch chuyển của PU của khối video trong khung hoặc hình ảnh video hiện thời so với khối dự báo trong khung tham chiếu (hoặc đơn vị được mã hóa khác) so với khối hiện thời đang được mã hóa trong khung hiện thời (hoặc đơn vị được mã hóa khác). Khối dự báo là khối được phát hiện ra là phù hợp nhất với khối được mã hóa, về độ chênh lệch điểm ảnh, mà có thể được xác định bằng tổng chênh lệch tuyệt đối (SAD), tổng chênh lệch bình phương (SSD), hoặc các số đo chênh lệch khác. Trong một số ví dụ, bộ mã hóa video 20 có thể tính toán các giá trị cho các vị trí điểm ảnh dưới số nguyên của hình ảnh tham chiếu được lưu trữ trong bộ nhớ hình ảnh tham chiếu 64. Ví dụ, bộ mã hóa video 20 có thể nội suy các giá trị của các vị trí điểm ảnh $1/4$, vị trí điểm ảnh $1/8$, hoặc các vị trí điểm ảnh phân số khác của hình ảnh tham chiếu. Do đó, bộ phận ước lượng chuyển động 42 có thể thực hiện tìm kiếm chuyển động liên quan đến các vị trí điểm ảnh đầy đủ và các vị trí điểm ảnh phân số và xuất ra vector chuyển động với độ chính xác điểm ảnh phân số.

Bộ phận ước lượng chuyển động 42 tính toán vector chuyển động cho PU của khối video trong lát được mã hóa nội cấu trúc bằng cách so sánh vị trí của PU này với vị trí của khối dự báo của hình ảnh tham chiếu. Hình ảnh tham chiếu có thể được lựa chọn từ danh sách hình ảnh tham chiếu thứ nhất (Danh sách 0) hoặc danh sách hình ảnh tham chiếu thứ hai (Danh sách 1), mỗi danh sách này nhận dạng một hoặc nhiều hình ảnh tham chiếu được lưu trữ trong bộ nhớ hình ảnh tham chiếu 64. Bộ phận ước lượng chuyển động 42 gửi vector chuyển động tính toán được đến bộ phận mã hóa entropy 56 và bộ phận bù chuyển động 44.

Quy trình bù chuyển động, được thực hiện bởi bộ phận bù chuyển động 44, có thể bao gồm tìm nạp hoặc tạo ra khối dự báo dựa vào vector chuyển động được xác

định bởi bộ phận ước lượng chuyển động 42. Ngoài ra, bộ phận ước lượng chuyển động 42 và bộ phận bù chuyển động 44 có thể được tích hợp chức năng, trong một số ví dụ. Khi nhận vector chuyển động cho PU của khối video hiện thời, bộ phận bù chuyển động 44 có thể định vị khối dự báo mà vector chuyển động trở đến trong một trong số các danh sách hình ảnh tham chiếu. Bộ cộng 50 tạo ra khối video dư bằng cách lấy các giá trị điểm ảnh của khối video hiện thời đang được mã hóa trừ đi các giá trị điểm ảnh của khối dự báo, tạo ra giá trị chênh lệch điểm ảnh, như được thảo luận dưới đây. Nói chung, bộ phận ước lượng chuyển động 42 thực hiện ước lượng chuyển động liên quan đến các thành phần độ chói, và bộ phận bù chuyển động 44 sử dụng vector chuyển động được tính toán dựa vào các thành phần độ chói này cho cả thành phần màu sắc và độ chói. Bộ phận chọn chế độ 40 cũng có thể tạo ra các phần tử cú pháp gắn với các khối video và lát video để sử dụng bởi bộ giải mã video 30 trong việc giải mã khối video của lát video.

Bộ phận dự báo nội cấu trúc 46 có thể dự báo nội cấu trúc khối hiện thời, như là một phương án thay thế cho việc dự báo liên cấu trúc được thực hiện bằng bộ phận ước lượng chuyển động 42 và bộ phận bù chuyển động 44, như được mô tả trên đây. Cụ thể là, bộ phận dự báo nội cấu trúc 46 có thể xác định chế độ dự báo nội cấu trúc sử dụng để mã hóa khối hiện thời. Trong một số ví dụ, bộ phận dự báo nội cấu trúc 46 có thể mã hóa khối hiện thời bằng cách sử dụng các chế độ dự báo nội cấu trúc khác nhau, ví dụ, trong các quy trình mã hóa riêng biệt, và bộ phận dự báo nội cấu trúc 46 (hoặc bộ phận chọn chế độ 40, trong một số ví dụ) có thể lựa chọn chế độ dự báo nội cấu trúc thích hợp để sử dụng từ các chế độ được kiểm tra.

Ví dụ, bộ phận dự báo nội cấu trúc 46 có thể tính toán các giá trị tốc độ - méo bằng cách sử dụng phân tích tốc độ - méo cho nhiều chế độ dự báo nội cấu trúc được kiểm tra khác nhau, và lựa chọn chế độ dự báo nội cấu trúc có các đặc điểm tốc độ - méo tốt nhất trong số các chế độ được kiểm tra. Phân tích tốc độ - méo thường xác định định lượng méo (hoặc sai số) giữa khối mã hóa và khối chưa được mã hóa ban đầu mà được mã hóa để tạo ra khối mã hóa, cũng như tốc độ bit (nghĩa là số lượng bit) được sử dụng để tạo ra khối mã hóa. Bộ phận dự báo nội cấu trúc 46 có thể tính toán các tỉ lệ từ các giá trị tốc độ và méo cho nhiều khối mã hóa khác nhau để xác định chế độ dự báo nội cấu trúc nào thể hiện giá trị tốc độ - méo tốt nhất cho khối.

Sau khi lựa chọn chế độ dự báo nội cấu trúc cho khối, bộ phận dự báo nội cấu trúc 46 có thể cung cấp thông tin chỉ báo chế độ dự báo nội cấu trúc được chọn cho khối này đến bộ phận mã hóa entropy 56. Bộ phận mã hóa entropy 56 có thể mã hóa thông tin chỉ báo chế độ dự báo nội cấu trúc được lựa chọn. Bộ mã hóa video 20 có thể bao gồm, trong dòng bit được truyền, dữ liệu cấu hình, dữ liệu này có thể bao gồm nhiều bảng chỉ số chế độ dự báo nội cấu trúc và nhiều bảng chỉ số chế độ dự báo nội cấu trúc sửa đổi (còn được gọi là bảng ánh xạ từ mã), định nghĩa về các ngữ cảnh mã hóa cho nhiều khối khác nhau và chỉ báo về chế độ dự báo nội cấu trúc có xác suất cao nhất, bảng chỉ số chế độ dự báo nội cấu trúc, và bảng chỉ số chế độ dự báo nội cấu trúc sửa đổi để sử dụng cho mỗi ngữ cảnh.

Bộ mã hóa video 20 tạo ra khối video dư bằng cách trừ khối video ban đầu đang được mã hóa cho dữ liệu dự báo từ bộ phận chọn chế độ 40. Bộ cộng 50 biểu diễn thành phần hoặc các thành phần thực hiện phép toán trừ này. Bộ phận xử lý biến đổi 52 áp dụng biến đổi, như biến đổi cosin rời rạc (DCT) hoặc biến đổi tương tự về khái niệm, cho khối dư, tạo ra khối video chứa các giá trị hệ số biến đổi dư. Bộ phận xử lý biến đổi 52 có thể thực hiện các biến đổi khác mà tương tự về khái niệm với DCT. Các biến đổi sóng con, các biến đổi số nguyên, các biến đổi dải phụ hoặc các loại biến đổi khác cũng có thể được sử dụng. Trong trường hợp bất kỳ, bộ phận xử lý biến đổi 52 áp dụng biến đổi cho khối dư, tạo ra khối hệ số biến đổi dư. Phép biến đổi có thể chuyển đổi thông tin dư từ miền giá trị điểm ảnh sang miền biến đổi, như miền tần số. Bộ phận xử lý biến đổi 52 có thể gửi các hệ số biến đổi kết quả đến bộ phận lượng tử hóa 54. Bộ phận lượng tử hóa 54 lượng tử hóa các hệ số biến đổi để làm giảm thêm tốc độ bit. Quá trình lượng tử hóa có thể làm giảm độ sâu bit liên quan đến một số hoặc toàn bộ hệ số này. Mức độ lượng tử hóa có thể được sửa đổi bằng cách điều chỉnh tham số lượng tử hóa. Trong một số ví dụ, bộ phận lượng tử hóa 54 có thể thực hiện quét ma trận bao gồm các hệ số biến đổi được lượng tử hóa. Ngoài ra, bộ phận mã hóa entropy 56 có thể thực hiện việc quét này.

Sau khi lượng tử hóa, bộ phận mã hóa entropy 56 mã hóa entropy các hệ số biến đổi được lượng tử hóa. Ví dụ, bộ phận mã hóa entropy 56 có thể thực hiện mã hóa độ dài biến đổi thích ứng ngữ cảnh (context adaptive variable length coding: CAVLC), mã hóa số học nhị phân thích ứng ngữ cảnh (context adaptive binary arithmetic coding: CABAC), mã hóa số học nhị phân thích ứng ngữ cảnh dựa vào cú

pháp (syntax-based context-adaptive binary arithmetic coding: SBAC), mã hóa entropy phân chia khoảng xác suất (probability interval partitioning entropy: PIPE) hoặc các kỹ thuật mã hóa entropy khác. Trong trường hợp mã hóa entropy dựa vào ngữ cảnh, ngữ cảnh có thể là dựa vào các khối lân cận. Sau khi mã hóa entropy bằng bộ phận mã hóa entropy 56, dòng bit được mã hóa có thể được truyền đến thiết bị khác (ví dụ, bộ giải mã video 30) hoặc được lưu trữ để truyền sau này hoặc thu hồi.

Bộ phận lượng tử hóa ngược 58 và bộ phận biến đổi ngược 60 lần lượt áp dụng lượng tử hóa ngược và biến đổi ngược để khôi phục khối dư trong miền điểm ảnh, ví dụ, để sử dụng sau này làm khối tham chiếu. Bộ phận bù chuyển động 44 có thể tính toán khối tham chiếu bằng cách cộng khối dư vào khối dự báo của một trong số các khung của bộ nhớ hình ảnh tham chiếu 64. Bộ phận bù chuyển động 44 cũng có thể áp dụng một hoặc nhiều bộ lọc nội suy cho khối dư được khôi phục để tính toán các giá trị điểm ảnh dưới số nguyên để sử dụng trong ước lượng chuyển động. Bộ cộng 62 cộng khối dư được khôi phục vào khối dự báo được bù chuyển động được tạo ra bởi bộ phận bù chuyển động 44 để tạo ra khối video được khôi phục để lưu trữ trong bộ nhớ hình ảnh tham chiếu 64. Khối video được khôi phục này có thể được sử dụng bởi bộ phận ước lượng chuyển động 42 và bộ phận bù chuyển động 44 làm khối tham chiếu để mã hóa liên cấu trúc khối trong khung video tiếp theo.

Bộ phận ước lượng chuyển động 42 có thể xác định một hoặc nhiều hình ảnh tham chiếu, mà bộ mã hóa video 20 có thể sử dụng để dự báo các giá trị điểm ảnh của một hoặc nhiều PU mà được dự báo liên cấu trúc. Bộ phận ước lượng chuyển động 42 có thể báo hiệu mỗi hình ảnh tham chiếu dưới dạng LTRP hoặc hình ảnh tham chiếu ngắn hạn. Bộ phận ước lượng chuyển động 42 có thể lưu trữ hình ảnh tham chiếu trong bộ đệm hình ảnh được giải mã (DPB) đến khi các hình ảnh này được đánh dấu là không được sử dụng để tham chiếu.

Bộ phận chọn chế độ 40 của bộ giải mã video 20 có thể mã hóa nhiều phân tử cú pháp khác nhau mà bao gồm thông tin nhận dạng cho một hoặc nhiều hình ảnh tham chiếu. Trong quá trình mã hóa hình ảnh được dự báo liên cấu trúc, bộ phận chọn chế độ 40 có thể mã hóa thông tin nhận dạng của một hoặc nhiều LTRP cho hình ảnh hiện thời trong tập hợp tham số chuỗi. Bộ phận chọn chế độ 40 cũng có thể mã hóa thông tin nhận dạng cho một hoặc nhiều hình ảnh tham chiếu ngắn hạn được sử dụng để dự báo hình ảnh hiện thời trong phần đầu lát của hình ảnh hiện thời hoặc tập hợp

tham số hình ảnh cho hình ảnh hiện thời. Trong một số ví dụ, các phần tử cú pháp có thể chỉ báo việc phần đầu lát có kế thừa các hình ảnh tham chiếu dài hạn từ tập hợp tham số chuỗi hoạt động hay không, hay việc hình ảnh tham chiếu dài hạn có được báo hiệu rõ ràng trong phần đầu lát hay không.

Theo cách này, bộ phận chọn chế độ 40 của bộ mã hóa video 20 trên Fig.2 biểu diễn ví dụ về bộ mã hóa video được tạo cấu hình để mã hóa phần đầu lát cho lát dữ liệu video, trong đó phần đầu lát bao gồm phần tử cú pháp chỉ báo việc thông tin nhận dạng cho hình ảnh tham chiếu dài hạn được báo hiệu rõ ràng trong phần đầu lát hay được suy ra từ tập hợp tham số chuỗi tương ứng với lát này, và, khi phần tử cú pháp này chỉ báo là thông tin nhận dạng cho hình ảnh tham chiếu dài hạn được báo hiệu rõ ràng, mã hóa giá trị của thông tin nhận dạng cho hình ảnh tham chiếu dài hạn trong phần đầu lát.

Trong một số ví dụ, bộ mã hóa video 20 có thể báo hiệu thông tin nhận dạng của các hình ảnh tham chiếu dài hạn bằng cách chỉ báo hiệu một số trong số các LSB của các giá trị POC của các hình ảnh tham chiếu dài hạn trong SPS hoạt động hoặc phần đầu lát. Tuy nhiên, trong một số trường hợp, hai LTRP có thể có cùng các LSB POC, vì vậy bộ mã hóa video 20 có thể báo hiệu thêm một số trong số các MSB của các giá trị POC của các LTRP.

Trong ví dụ khác về các kỹ thuật theo sáng chế, bộ mã hóa video 20 có thể báo hiệu danh sách các LTRP được sử dụng để dự báo các giá trị điểm ảnh của lát hiện thời trong các phần tử cú pháp của phần đầu lát hiện thời. Bộ mã hóa video 20 có thể mã hóa dữ liệu cho nhiều hình ảnh theo thứ tự mã hóa. Dữ liệu này có thể chỉ báo là mỗi hình ảnh trong số nhiều hình ảnh này là sẵn có để sử dụng làm LTRP. Bộ mã hóa video 20 có thể mã hóa phần tử cú pháp trong phần đầu lát của hình ảnh hiện thời mà bao gồm các giá trị cho bit có trọng số nhỏ nhất (LSB) của các giá trị số đếm thứ tự hình ảnh (POC) tương ứng của nhiều hình ảnh sao cho các giá trị cho LSB là không giảm hoặc không tăng theo thứ tự mã hóa hình ảnh.

Dựa trên hình ảnh tham chiếu được báo hiệu trong SPS, PPS, và phần đầu lát, bộ mã hóa video 20 thiết lập các danh sách hình ảnh tham chiếu, như các danh sách hình ảnh tham chiếu RefPicSetStCurrBefore, RefPicSetStCurrAfter, RefPicSetStFoll, RefPicSetLtCurr và RefPicSetLtFoll. Bộ mã hóa video 20 có thể thiết lập danh sách hình ảnh tham chiếu dựa trên các giá trị POC của các hình ảnh tham chiếu. Dựa vào 5

danh sách hình ảnh tham chiếu, bộ mã hóa video 20 cũng có thể thiết lập các danh sách hình ảnh tham chiếu RefPicList0 và RefPicList1. Bộ mã hóa video 20 có thể sử dụng hình ảnh tham chiếu được bao gồm trong RefPicList0 và RefPicList1 để dự báo các giá trị điểm ảnh của khối được dự báo hai chiều, và hình ảnh tham chiếu của RefPicList0 để dự báo các giá trị điểm ảnh của khối được dự báo một chiều.

Fig.3 là sơ đồ khối minh họa ví dụ về bộ giải mã video 30 có thể thực hiện các kỹ thuật để báo hiệu dữ liệu cho các LTRP trong SPS hoặc phần đầu lát. Trong ví dụ trên Fig.3, bộ giải mã video 30 bao gồm thiết bị giải mã entropy 70, bộ phận bù chuyển động 72, bộ phận dự báo nội cấu trúc 74, bộ phận lượng tử hóa ngược 76, bộ phận biến đổi ngược 78, bộ nhớ hình ảnh tham chiếu 82 và bộ cộng 80. Bộ giải mã video 30 có thể, trong một số ví dụ, thực hiện quy trình giải mã thường là ngược với quy trình mã hóa được mô tả liên quan đến bộ mã hóa video 20 (Fig.2). Bộ phận bù chuyển động 72 có thể tạo ra dữ liệu dự báo dựa vào các vector chuyển động nhận được từ bộ phận giải mã entropy 70, trong khi bộ phận dự báo nội cấu trúc 74 có thể tạo ra dữ liệu dự báo dựa vào các chỉ báo chế độ dự báo nội cấu trúc nhận được từ bộ phận giải mã entropy 70.

Trong quá trình giải mã, bộ giải mã video 30 nhận dòng bit video được mã hóa biểu diễn các khối video của lát video được mã hóa và các phần tử cú pháp gắn kèm từ bộ mã hóa video 20. Bộ phận giải mã entropy 70 của bộ giải mã video 30 giải mã entropy dòng bit để tạo ra các hệ số được lượng tử hóa, vector chuyển động hoặc các chỉ báo chế độ dự báo nội cấu trúc, và các phần tử cú pháp khác. Bộ phận giải mã entropy 70 chuyển tiếp vector chuyển động và các phần tử cú pháp khác đến bộ phận bù chuyển động 72. Bộ giải mã video 30 có thể thu các phần tử cú pháp ở mức lát video và/hoặc mức khối video.

Khi lát video được mã hóa dưới dạng lát được mã hóa nội cấu trúc (I), bộ phận dự báo nội cấu trúc 74 có thể tạo ra dữ liệu dự báo cho khối video của lát video hiện thời dựa vào chế độ dự báo nội cấu trúc được báo hiệu và dữ liệu từ các khối được giải mã trước đó của khung hoặc hình ảnh hiện thời. Khi khung video được mã hóa dưới dạng lát được mã hóa liên cấu trúc (nghĩa là, B, P hoặc GPB), bộ phận bù chuyển động 72 tạo ra các khối dự báo cho khối video của lát video hiện thời dựa vào vector chuyển động và các phần tử cú pháp khác nhận được từ bộ phận giải mã entropy 70. Các khối dự báo có thể được tạo ra từ một trong số các hình ảnh tham chiếu trong một trong số

các danh sách hình ảnh tham chiếu. Bộ giải mã video 30 có thể thiết lập danh sách khung tham chiếu, Danh sách 0 và Danh sách 1, sử dụng các kỹ thuật thiết lập mặc định dựa vào các hình ảnh tham chiếu được lưu trữ trong bộ nhớ hình ảnh tham chiếu 82.

Bộ phận bù chuyển động 72 xác định thông tin dự báo cho khối video của lát video hiện thời bằng cách phân tích cú pháp vector chuyển động và các phần tử cú pháp khác, và sử dụng thông tin dự báo để tạo ra các khối dự báo cho khối video hiện thời đang được mã hóa. Ví dụ, bộ phận bù chuyển động 72 sử dụng một số trong số các phần tử cú pháp nhận được để xác định chế độ dự báo (ví dụ, nội cấu trúc hoặc liên cấu trúc) được sử dụng để mã hóa khối video của lát video, loại lát dự báo liên cấu trúc (ví dụ, lát B, lát P, hoặc lát GPB), thông tin thiết lập cho một hoặc nhiều danh sách hình ảnh tham chiếu cho lát, các vector chuyển động cho mỗi khối video được mã hóa liên cấu trúc của lát, trạng thái dự báo liên cấu trúc cho mỗi khối video được mã hóa liên cấu trúc của lát này và thông tin khác để giải mã các khối video trong lát video hiện thời.

Bộ phận bù chuyển động 72 cũng có thể thực hiện nội suy dựa vào các bộ lọc nội suy. Bộ phận bù chuyển động 72 có thể sử dụng các bộ lọc nội suy như được sử dụng bởi bộ mã hóa video 20 trong quá trình mã hóa các khối video để tính toán các giá trị được nội suy cho các điểm ảnh dưới số nguyên của các khối tham chiếu. Trong trường hợp này, bộ phận bù chuyển động 72 có thể xác định các bộ lọc nội suy được sử dụng bởi bộ mã hóa video 20 từ các phần tử cú pháp thu được và sử dụng các bộ lọc nội suy để tạo ra các khối dự báo.

Bộ phận lượng tử hóa ngược 76 lượng tử hóa ngược, tức là, giải lượng tử hóa, các hệ số biến đổi được lượng tử hóa được cung cấp trong dòng bit và được giải mã bởi bộ phận giải mã entropy 70. Quá trình lượng tử hóa ngược này có thể bao gồm việc sử dụng tham số lượng tử hóa QP_Y được tính toán bằng bộ giải mã video 30 cho mỗi khối video trong lát video để xác định mức độ lượng tử hóa và tương tự, mức độ lượng tử hóa ngược nên được sử dụng.

Bộ phận biến đổi ngược 78 áp dụng kỹ thuật biến đổi ngược, ví dụ, DCT ngược, biến đổi số nguyên ngược, hoặc quy trình biến đổi ngược tương tự về khái niệm, cho hệ số biến đổi để tạo ra các khối dư trong miền điểm ảnh.

Sau khi bộ phận bù chuyển động 72 tạo ra khối dự báo cho khối video hiện thời dựa vào các vectơ chuyển động và các phần tử cú pháp khác, bộ giải mã video 30 tạo ra khối video được giải mã bằng cách cộng các khối dự từ bộ phận biến đổi ngược 78 với các khối dự báo tương ứng được tạo ra bởi bộ phận bù chuyển động 72. Bộ cộng 80 biểu diễn thành phần hoặc các thành phần thực hiện phép cộng này. Nếu muốn, bộ lọc giải khối cũng có thể được áp dụng để lọc các khối được giải mã để loại bỏ các thành phần lạ dạng khối. Các bộ lọc vòng khác (cả trong vòng mã hóa hoặc sau vòng mã hóa) cũng có thể được sử dụng để làm trơn các chuyển tiếp điểm ảnh, hoặc cải thiện chất lượng video. Các khối video được giải mã trong khung hoặc hình ảnh đã cho sau đó được lưu trữ trong bộ nhớ hình ảnh tham chiếu 82, mà lưu trữ các hình ảnh tham chiếu được sử dụng để bù chuyển động tiếp theo. Bộ nhớ hình ảnh tham chiếu 82 cũng lưu trữ video được giải mã để trình diễn sau trên thiết bị hiển thị, như thiết bị hiển thị 32 trên Fig.1.

Như được mô tả trên đây, trong quá trình dự báo liên cấu trúc, bộ phận bù chuyển động 72 có thể xác định một hoặc nhiều hình ảnh tham chiếu, mà bộ giải mã video 30 có thể sử dụng để tạo ra các khối video dự báo cho khối hiện thời đang được giải mã. Bộ phận bù chuyển động 72 có thể xác định xem các hình ảnh tham chiếu là các hình ảnh tham chiếu dài hạn hay các hình ảnh tham chiếu ngắn hạn dựa vào các phần tử cú pháp của dòng bit video được mã hóa, mà chỉ báo việc hình ảnh tham chiếu được đánh dấu là để tham chiếu dài hạn hay tham chiếu ngắn hạn. Bộ phận bù chuyển động 72 có thể lưu trữ hình ảnh tham chiếu trong bộ đệm hình ảnh được giải mã (DPB) đến khi các hình ảnh này được đánh dấu là không được sử dụng để tham chiếu.

Bộ phận bù chuyển động 72 của bộ giải mã video 30 có thể giải mã các phần tử cú pháp khác nhau bao gồm thông tin nhận dạng cho một hoặc nhiều hình ảnh tham chiếu được sử dụng để tạo ra các khối dự báo cho khối giải mã hiện thời. Trong quá trình giải mã PU được dự báo liên cấu trúc, bộ phận bù chuyển động 72 có thể giải mã thông tin nhận dạng của một hoặc nhiều LTRP cho hình ảnh hiện thời được báo hiệu trong tập hợp tham số chuỗi hoạt động. Bộ phận bù chuyển động 72 cũng có thể giải mã thông tin nhận dạng cho một hoặc nhiều hình ảnh tham chiếu ngắn hạn được sử dụng để dự báo hình ảnh hiện thời trong phần đầu lát của hình ảnh hiện thời hoặc tập hợp tham số hình ảnh cho hình ảnh hiện thời. Trong một số ví dụ, các phần tử cú pháp này có thể chỉ báo việc phần đầu lát có kế thừa các hình ảnh tham chiếu dài hạn từ tập

hợp tham số chuỗi hoạt động hay không, hoặc việc thông tin nhận dạng cho hình ảnh tham chiếu dài hạn được báo hiệu rõ ràng trong phần đầu lát hay không.

Theo cách này, bộ phận bù chuyển động 72 của bộ giải mã video 30 trên Fig.3 biểu diễn ví dụ về bộ giải mã video được tạo cấu hình để mã hóa phần đầu lát cho lát dữ liệu video, trong đó phần đầu lát bao gồm phần tử cú pháp chỉ báo việc thông tin nhận dạng cho hình ảnh tham chiếu dài hạn được báo hiệu rõ ràng trong phần đầu lát hay được suy ra từ tập hợp tham số chuỗi tương ứng với lát này, và, khi phần tử cú pháp chỉ báo là thông tin nhận dạng hình ảnh tham chiếu dài hạn được báo hiệu rõ ràng, mã hóa giá trị của thông tin nhận dạng cho hình ảnh tham chiếu dài hạn trong phần đầu lát.

Trong một số ví dụ, bộ giải mã video 30 có thể giải mã thông tin nhận dạng của các hình ảnh tham chiếu dài hạn chỉ dựa vào một số trong số các LSB của các giá trị POC của các hình ảnh tham chiếu dài hạn được báo hiệu trong SPS hoạt động hoặc phần đầu lát. Tuy nhiên, trong một số trường hợp, hai LTRP có thể có các LSB POC giống nhau, do đó bộ mã hóa video 20 có thể giải mã thêm một số trong các MSB của các giá trị POC của các LTRP để gộp các LTRP mà có cùng các LSB POC.

Trong ví dụ khác của các kỹ thuật theo sáng chế, bộ giải mã video 30 có thể mã hóa danh sách LTRP được sử dụng để dự báo các giá trị điểm ảnh của lát hiện thời trong phần tử cú pháp của phần đầu lát hiện thời. Bộ giải mã video 30 có thể mã hóa dữ liệu cho nhiều hình ảnh theo thứ tự mã hóa. Dữ liệu này có thể chỉ báo là mỗi hình ảnh trong số nhiều hình ảnh là sẵn có để sử dụng làm LTRP. Bộ giải mã video 30 có thể mã hóa phần tử cú pháp trong phần đầu lát của hình ảnh hiện thời mà bao gồm các giá trị cho bit có trọng số nhỏ nhất (LSB) của các giá trị số đếm thứ tự hình ảnh tương ứng (POC) của nhiều hình ảnh sao cho các giá trị cho các LSB này là không tăng hoặc không giảm theo thứ tự mã hóa hình ảnh.

Dựa vào hình ảnh tham chiếu được báo hiệu trong SPS, PPS, và phần đầu lát, bộ giải mã video 30 thiết lập tập con hình ảnh tham chiếu, chẳng hạn như tập con hình ảnh tham chiếu RefPicSetStCurrBefore, RefPicSetStCurrAfter, RefPicSetStFoll, RefPicSetLtCurr và RefPicSetLtFoll. Bộ giải mã video 30 có thể thiết lập các danh sách hình ảnh tham chiếu dựa trên các giá trị POC của các hình ảnh tham chiếu. Dựa vào 5 danh sách hình ảnh tham chiếu, bộ giải mã video 30 cũng có thể thiết lập danh sách hình ảnh tham chiếu RefPicList0 và RefPicList1. Bộ giải mã video 30 có thể sử

dụng hình ảnh tham chiếu được bao gồm trong RefPicList0 và RefPicList1 để dự báo các giá trị điểm ảnh của khối được dự báo hai chiều, và các hình ảnh tham chiếu của RefPicList0 để dự báo các giá trị điểm ảnh của khối được dự báo một chiều.

Fig.4 là lưu đồ minh họa quá trình mã hóa dữ liệu cho các hình ảnh tham chiếu theo một hoặc nhiều ví dụ được mô tả trong tài liệu này. Chỉ với mục đích minh họa, phương pháp trên Fig.4 có thể được thực hiện bằng bộ mã hóa video tương ứng với bộ mã hóa video 20 hoặc bộ giải mã video 30. Trong phương pháp trên Fig.4, bộ mã hóa video có thể mã hóa thông tin nhận dạng cho các hình ảnh tham chiếu dài hạn trong tập hợp tham số chuỗi hoạt động (SPS) (100). Trong một số ví dụ, thông tin nhận dạng này có thể bao gồm các LSB của các giá trị POC cho các LTRP.

Bộ giải mã video cũng có thể mã hóa một hoặc nhiều phần tử cú pháp mà bao gồm thông tin nhận dạng cho một hoặc nhiều hình ảnh tham chiếu ngắn hạn trong SPS và phần đầu lát cho hình ảnh hiện đang được mã hóa (102).

Ngoài việc mã hóa thông tin nhận dạng cho các LTRP trong SPS, bộ giải mã video cũng có thể mã hóa thông tin nhận dạng cho các LTRP trong phần đầu lát hiện thời của hình ảnh mã hóa hiện thời (104). Phần đầu lát có thể bao gồm phần tử cú pháp chứa danh sách các LTRP cho lát hiện thời. Trong một số ví dụ, ít nhất một số trong các LTRP trong danh sách được kế thừa từ SPS hoạt động, và một số trong các LTRP trong danh sách này được mã hóa trong phần đầu lát hiện thời. Quy trình mã hóa SPS được kế thừa được mô tả chi tiết hơn dưới đây với Fig.5.

Sau khi mã hóa các LTRP và STRP, bộ giải mã video có thể thiết lập tập con hình ảnh tham chiếu. Trong một số ví dụ, tập con hình ảnh tham chiếu, như tập con hình ảnh tham chiếu RefPicSetStCurrBefore, RefPicSetStCurrAfter, RefPicSetStFoll, RefPicSetLtCurr và RefPicSetLtFoll (106). Dựa vào tập con hình ảnh tham chiếu, bộ mã hóa video thiết lập danh sách hình ảnh tham chiếu, RefPicList0, và RefPicList1 (108). Bộ mã hóa video có thể sử dụng các kỹ thuật thiết lập mặc định để thiết lập các danh sách hình ảnh tham chiếu. Bộ mã hóa video cũng có thể sửa đổi thứ tự của các hình ảnh tham chiếu trong RefPicList0 và RefPicList1 dựa vào các phần tử cú pháp mà chỉ rõ thứ tự khác nhau cho các hình ảnh tham chiếu.

Sau khi thiết lập các danh sách hình ảnh tham chiếu, bộ giải mã video có thể tạo ra các khối dự báo cho hình ảnh mã hóa hiện thời dựa vào các hình ảnh tham chiếu của các danh sách hình ảnh tham chiếu (110). Trong quá trình mã hóa, bộ giải mã video có

thể dự báo một phần dữ liệu video của lát sử dụng hình ảnh tham chiếu dài hạn và mã hóa phần dữ liệu video này dựa vào việc dự báo (112).

Fig.5 là lưu đồ minh họa quy trình mã hóa dữ liệu cho các LTRP trong SPS hoặc phần đầu lát theo một hoặc nhiều ví dụ được mô tả trong tài liệu này. Chỉ với mục đích minh họa, phương pháp trên Fig.5 có thể được thực hiện bởi bộ mã hóa video tương ứng với bộ mã hóa video 20 hoặc bộ giải mã video 30. Trong phương pháp trên Fig.5, bộ mã hóa video có thể mã hóa phần đầu lát cho lát dữ liệu video (120). Trong một số ví dụ, phần đầu lát có thể là phần đầu lát cho hình ảnh mã hóa hiện thời.

Phần đầu lát bao gồm phần tử cú pháp chỉ báo việc thông tin nhận dạng cho LTRP được báo hiệu rõ ràng trong phần đầu lát hay được suy ra từ tập hợp tham số chuỗi tương ứng với lát này. Nếu phần tử cú pháp này chỉ báo là thông tin nhận dạng cho LTRP được báo hiệu rõ ràng, bộ mã hóa video có thể mã hóa giá trị của thông tin nhận dạng cho hình ảnh tham chiếu dài hạn trong phần đầu lát (122). Nếu phần tử cú pháp này không chỉ báo là thông tin nhận dạng cho LTRP được báo hiệu rõ ràng, bộ mã hóa video có thể suy ra thông tin nhận dạng cho LTRP từ tập hợp tham số chuỗi khi phần tử cú pháp này chỉ báo là thông tin nhận dạng cho hình ảnh tham chiếu dài hạn không được báo hiệu rõ ràng (124).

Trong một số ví dụ, thiết bị mã hóa video được tạo cấu hình để thực hiện quy trình trên Fig.5 có thể còn được tạo cấu hình để lưu trữ hình ảnh được mã hóa thứ nhất tương ứng với hình ảnh thứ nhất trong bộ đệm hình ảnh được giải mã. Khi bộ đệm hình ảnh được giải mã chứa nhiều hơn một hình ảnh tham chiếu mà được đánh dấu là “được sử dụng để tham chiếu” và có cùng giá trị của bit có trọng số nhỏ nhất (LSB) của giá trị số đếm thứ tự hình ảnh (POC) với hình ảnh thứ nhất, để mã hóa thông tin nhận dạng, thiết bị này còn được tạo cấu hình để mã hóa phần tử cú pháp thứ nhất bằng một trong trong phần đầu lát của lát của hình ảnh thứ hai. Khi bộ đệm hình ảnh được giải mã không chứa nhiều hơn một hình ảnh mà được đánh dấu là “được sử dụng để tham chiếu” và có cùng giá trị LSB POC với hình ảnh thứ nhất, để mã hóa thông tin nhận dạng, thiết bị này có thể được tạo cấu hình để mã hóa phần tử cú pháp thứ nhất bằng 0 trong phần đầu lát. Khi phần tử cú pháp thứ nhất bằng 1, để mã hóa thông tin nhận dạng, thiết bị này còn được tạo cấu hình để mã hóa phần tử cú pháp thứ hai chỉ

báo các bit có trọng số lớn nhất (MSB) của giá trị POC của hình ảnh thứ nhất và mã hóa hình ảnh thứ hai sử dụng hình ảnh thứ nhất làm hình ảnh tham chiếu dài hạn.

Trong một số ví dụ, bộ mã hóa video được tạo cấu hình để thực hiện quy trình trên Fig.5 có thể còn được tạo cấu hình để mã hóa dữ liệu cho nhiều hình ảnh theo thứ tự mã hóa hình ảnh. Dữ liệu này có thể chỉ báo là nhiều hình ảnh này là sẵn có để sử dụng làm các hình ảnh tham chiếu dài hạn. Bộ mã hóa video cũng có thể mã hóa các giá trị cho bit có trọng số nhỏ nhất (LSB) của các giá trị số đếm thứ tự hình ảnh (POC) của nhiều hình ảnh sao cho các giá trị của các LSB là không tăng hoặc không giảm theo thứ tự mã hóa hình ảnh.

Trong ví dụ khác, để mã hóa các giá trị cho các LSB, bộ giải mã video có thể được tạo cấu hình để mã hóa các giá trị cho các LSB này cho các giá trị bất kỳ trong số j và k , trong đó j và k là số nguyên, nằm trong khoảng từ 0 đến số lượng hình ảnh tham chiếu dài hạn trừ đi 1, các LSB của giá trị POC của hình ảnh thứ nhất trong số nhiều hình ảnh ở vị trí j trong danh sách các hình ảnh tham chiếu là lớn hơn hoặc bằng, hoặc nhỏ hơn hoặc bằng, các LSB của giá trị POC của hình ảnh thứ hai trong số nhiều hình ảnh ở vị trí k trong danh sách các hình ảnh tham chiếu, Bộ giải mã video có thể còn được tạo cấu hình để mã hóa SPS bao gồm thông tin chỉ báo số lượng LTRP.

Trong ví dụ khác nữa, dữ liệu cho nhiều hình ảnh có thể còn chỉ báo việc nhiều hình ảnh này có được sử dụng để tham chiếu hay không và đối với mỗi hình ảnh trong nhiều hình ảnh mà được chỉ báo là được sử dụng để tham chiếu, dữ liệu này còn có thể bao gồm dữ liệu để xác định xem giá trị của các bit có trọng số lớn nhất (MSB) của các giá trị POC cho mỗi hình ảnh có được mã hóa hay không.

Trong ví dụ khác, bộ giải mã video có thể xác định giá trị của biến `DeltaPocMSBCycleLt`, trong đó giá trị `DeltaPocMSBCycleLt` của một trong số nhiều hình ảnh này là bằng giá trị cho phần tử cú pháp `delta_poc_msb_cycle_lt` cho một trong số nhiều hình ảnh khi một trong số nhiều hình ảnh này có giá trị thứ tự mã hóa bằng số lượng hình ảnh tham chiếu dài hạn được xác định trong tập hợp tham số chuỗi.

Cần hiểu rằng, tùy theo ví dụ, một số trường hợp hoặc sự kiện của kỹ thuật bất kỳ trong số các kỹ thuật được mô tả ở đây có thể được thực hiện theo trình tự khác, có thể được bổ sung, hợp nhất hoặc loại bỏ cùng nhau (ví dụ, không phải tất cả các trường hợp hay sự kiện được mô tả đều cần thiết để thực hiện các kỹ thuật này). Ngoài ra, trong một số ví dụ, các trường hợp hoặc sự kiện có thể được thực hiện đồng thời, qua

nhiều bước xử lý đa luồng, xử lý gián đoạn hoặc nhiều bộ xử lý hơn là thực hiện lần lượt.

Trong một hoặc nhiều ví dụ, các chức năng được mô tả có thể được thực hiện trong phần cứng, phần mềm, hoặc phần sụn hoặc tổ hợp bất kỳ của chúng. Nếu được thực hiện trong phần mềm, các chức năng này có thể được lưu trữ trên hoặc được truyền qua dưới dạng một hoặc nhiều lệnh hoặc mã trên vật ghi đọc được bằng máy tính và được thực hiện bằng bộ phận xử lý dựa vào phần cứng. Vật ghi đọc được bằng máy tính này có thể bao gồm phương tiện lưu trữ đọc được bằng máy tính, mà tương ứng với phương tiện hữu hình như phương tiện lưu trữ dữ liệu, hoặc phương tiện truyền thông bao gồm phương tiện bất kỳ mà tạo thuận lợi cho việc truyền chương trình máy tính từ nơi này đến nơi khác, ví dụ, theo giao thức truyền thông. Theo cách này, vật ghi đọc được bằng máy tính thường có thể tương ứng với (1) phương tiện lưu trữ hữu hình đọc được bằng máy tính mà bất biến hoặc (2) phương tiện truyền thông như tín hiệu hoặc sóng mang. Phương tiện lưu trữ dữ liệu có thể là phương tiện sẵn có bất kỳ mà có thể được truy cập bởi một hoặc nhiều máy tính hoặc một hoặc nhiều bộ xử lý để phục hồi các lệnh, mã và/hoặc cấu trúc dữ liệu để thực hiện các kỹ thuật được mô tả trong tài liệu này. Sản phẩm chương trình máy tính có thể bao gồm vật ghi đọc được bằng máy tính.

Ví dụ, và không giới hạn, vật ghi đọc được bằng máy tính có thể bao gồm RAM, ROM, EEPROM, CD-ROM hoặc ổ đĩa quang khác, ổ đĩa từ hoặc thiết bị lưu trữ từ tính khác, bộ nhớ nhanh hoặc phương tiện khác bất kỳ mà có thể được sử dụng để lưu trữ mã chương trình mong muốn ở dạng các lệnh hoặc cấu trúc dữ liệu và có thể được truy cập bằng máy tính. Ngoài ra, kết nối bất kỳ được gọi chính xác là vật ghi đọc được bằng máy tính. Ví dụ, nếu các lệnh được truyền từ website, máy chủ, hoặc nguồn từ xa khác sử dụng cáp đồng trục, cáp sợi quang, cáp xoắn đôi, đường dây thuê bao số (DSL), hoặc các công nghệ không dây như hồng ngoại, vô tuyến, và vi sóng, thì cáp đồng trục, cáp sợi quang, cáp xoắn đôi, DSL hoặc các công nghệ không dây như hồng ngoại, vô tuyến, và vi sóng được bao gồm trong khái niệm về vật ghi này. Tuy nhiên, nên hiểu là, vật ghi đọc được bằng máy tính và phương tiện lưu trữ dữ liệu không bao gồm các kết nối, sóng mang, tín hiệu hoặc phương tiện khả biến khác, mà thay vào đó hướng đến phương tiện lưu trữ bất biến, hữu hình. Đĩa từ và đĩa quang, như được sử dụng ở đây, bao gồm đĩa nén (CD), đĩa laze, đĩa quang, đĩa số đa năng

(DVD), đĩa mềm và đĩa Blu-ray, trong đó các đĩa từ thường sao chép dữ liệu bằng từ tính, trong khi đó các đĩa quang sao chép dữ liệu quang bằng laze. Dạng kết hợp của các phương tiện trên đây cũng được bao gồm trong phạm vi của vật ghi đọc được bằng máy tính.

Các lệnh có thể được thực hiện bằng một hoặc nhiều bộ xử lý, như một hoặc nhiều bộ xử lý tín hiệu số (DSP), bộ vi xử lý đa năng, mạch tích hợp chuyên dụng (ASIC), mảng công lập trình được dạng trường (FPGA), hoặc các mạch tích hợp tương đương hoặc mạch logic rời rạc khác. Theo đó, thuật ngữ “bộ xử lý” như được sử dụng ở đây dùng để chỉ bất kỳ trong số các cấu trúc trên đây hoặc cấu trúc khác bất kỳ thích hợp để thực hiện các kỹ thuật được mô tả ở đây. Thêm vào đó, theo một số khía cạnh, chức năng được mô tả ở đây có thể được cung cấp trong các môđun phần cứng và/hoặc phần mềm chuyên dụng được tạo cấu hình để mã hóa và giải mã, hoặc kết hợp vào bộ mã hóa – bộ giải mã (codec) kết hợp. Ngoài ra, các kỹ thuật này có thể được thực hiện đầy đủ trong một hoặc nhiều mạch hoặc phần tử logic.

Các kỹ thuật theo sáng chế có thể thực hiện trong nhiều thiết bị hoặc phương tiện, bao gồm điện thoại không dây, mạch tích hợp (IC) hoặc tập hợp các IC (ví dụ, bộ chip). Nhiều thành phần, môđun hoặc bộ phận khác nhau được mô tả trong tài liệu này để nhấn mạnh khía cạnh chức năng của thiết bị được tạo cấu hình để thực hiện các kỹ thuật được mô tả, nhưng không nhất thiết đòi hỏi việc thực hiện bởi các bộ phận phần cứng khác nhau. Như được mô tả trên đây, nhiều bộ phận khác nhau có thể được kết hợp trong bộ phận phần cứng codec hoặc được cung cấp bởi tập hợp các bộ phận phần cứng hoạt động tương tác, bao gồm 1 hoặc nhiều bộ xử lý như được mô tả trên đây, kết hợp với phần mềm, và/hoặc phần sụn thích hợp.

Nhiều ví dụ khác nhau được mô tả. Các ví dụ này và các ví dụ khác nằm trong phạm vi của phần yêu cầu bảo hộ sau đây.

Yêu cầu bảo hộ

1. Phương pháp giải mã dữ liệu video, phương pháp này bao gồm các bước:

giải mã dữ liệu phần tử cú pháp cho nhiều hình ảnh theo thứ tự mã hóa hình ảnh, trong đó dữ liệu phần tử cú pháp chỉ báo rằng mỗi hình ảnh trong số nhiều hình ảnh này là sẵn có để sử dụng làm hình ảnh tham chiếu dài hạn;

giải mã phần đầu lát cho lát của dữ liệu video, trong đó phần đầu lát bao gồm phần tử cú pháp chỉ báo số lượng hình ảnh tham chiếu dài hạn được suy ra từ tập hợp tham số chuỗi;

xác định xem hình ảnh tham chiếu dài hạn trong nhiều hình ảnh này được báo hiệu rõ ràng trong phần đầu lát hay được suy ra từ tập hợp tham số chuỗi tương ứng với lát này;

khi hình ảnh tham chiếu dài hạn được báo hiệu rõ ràng, bước giải mã phần đầu lát còn bao gồm giải mã giá trị của thông tin nhận dạng cho hình ảnh tham chiếu dài hạn trong phần đầu lát; và

xác định rằng giá trị cho $\Delta POC_{MSBCycleLt}[i]$ của hình ảnh tham chiếu dài hạn có giá trị chỉ số i luôn bằng giá trị cho phần tử cú pháp $\Delta poc_msb_cycle_lt[i]$ của hình ảnh tham chiếu dài hạn có giá trị chỉ số i khi hình ảnh tham chiếu dài hạn có giá trị chỉ số i bằng với số lượng hình ảnh tham chiếu dài hạn được suy ra từ tập hợp tham số chuỗi.

2. Phương pháp theo điểm 1, trong đó hình ảnh tham chiếu dài hạn bao gồm hình ảnh thứ nhất, phương pháp này còn bao gồm các bước:

lưu trữ hình ảnh được giải mã thứ nhất tương ứng với hình ảnh thứ nhất trong bộ đệm hình ảnh được giải mã;

khi bộ đệm hình ảnh được giải mã chứa nhiều hơn một hình ảnh tham chiếu mà được đánh dấu là “được sử dụng để tham chiếu” và có cùng giá trị của bit có trọng số nhỏ nhất (least significant bit - LSB) của số đếm thứ tự hình ảnh (picture order count - POC) với hình ảnh thứ nhất, trong đó bước giải mã giá trị cho thông tin nhận dạng còn bao gồm: giải mã phần tử cú pháp thứ nhất bằng một trong phần đầu lát của lát của hình ảnh thứ hai;

khi bộ đệm hình ảnh được giải mã không chứa nhiều hơn một hình ảnh được đánh dấu là “được sử dụng để tham chiếu” và có cùng giá trị LSB POC với hình ảnh thứ nhất, trong đó bước giải mã giá trị cho thông tin nhận dạng còn bao gồm: giải mã phần tử cú pháp thứ nhất bằng 0 trong phần đầu lát; và

khi phần tử cú pháp thứ nhất bằng 1, bước giải mã giá trị cho thông tin nhận dạng còn bao gồm: giải mã phần tử cú pháp thứ hai chỉ báo bit có trọng số lớn nhất (most significant bit - MSB) của giá trị POC của hình ảnh thứ nhất; và

giải mã hình ảnh thứ hai bằng cách sử dụng hình ảnh thứ nhất làm hình ảnh tham chiếu dài hạn.

3. Phương pháp theo điểm 1, trong đó phương pháp này còn bao gồm bước:

suy ra hình ảnh tham chiếu dài hạn từ tập hợp tham số chuỗi dựa trên việc xác định rằng hình ảnh tham chiếu dài hạn được suy ra.

4. Phương pháp theo điểm 1, trong đó phương pháp này còn bao gồm bước:

giải mã phần tử cú pháp trong phần đầu lát của hình ảnh hiện thời, trong đó phần tử cú pháp này bao gồm các giá trị cho bit có trọng số nhỏ nhất (least significant bit - LSB) của giá trị số đếm thứ tự hình ảnh (picture order count - POC) tương ứng của nhiều hình ảnh sao cho các giá trị cho các LSB này là không tăng hoặc không giảm theo thứ tự mã hóa hình ảnh.

5. Phương pháp theo điểm 4, trong đó bước giải mã các giá trị cho LSB bao gồm giải mã các giá trị của LSB sao cho đối với giá trị bất kỳ trong số j và k , trong đó j và k là các số nguyên nằm trong khoảng 0 đến số lượng hình ảnh tham chiếu dài hạn được suy ra từ tập hợp tham số chuỗi trừ đi một, các LSB của giá trị POC của hình ảnh thứ nhất trong số nhiều hình ảnh ở vị trí j trong danh sách các hình ảnh tham chiếu được báo hiệu trong phần đầu lát là lớn hơn hoặc bằng các LSB của giá trị POC của hình ảnh thứ hai trong số nhiều hình ảnh ở vị trí k trong danh sách các hình ảnh tham chiếu.

6. Phương pháp theo điểm 5, trong đó phương pháp này còn bao gồm bước giải mã tập hợp tham số chuỗi bao gồm thông tin chỉ báo số lượng hình ảnh tham chiếu dài hạn.

7. Phương pháp theo điểm 4, trong đó bước giải mã các giá trị cho LSB bao gồm giải mã các giá trị của LSB sao cho đối với giá trị bất kỳ trong số j và k , trong đó j và k là các số nguyên trong khoảng 0 đến số lượng hình ảnh tham chiếu dài hạn được suy ra từ tập hợp tham số chuỗi trừ đi một, các LSB của giá trị POC của hình ảnh thứ nhất trong số nhiều hình ảnh ở vị trí j trong danh sách các hình ảnh tham chiếu được báo hiệu trong phần đầu lát là nhỏ hơn hoặc bằng các LSB của giá trị POC của hình ảnh thứ hai trong số nhiều hình ảnh ở vị trí k trong danh sách các hình ảnh tham chiếu.

8. Phương pháp theo điểm 7, trong đó phương pháp này còn bao gồm bước giải mã tập hợp tham số chuỗi bao gồm thông tin chỉ báo số lượng hình ảnh tham chiếu dài hạn.

9. Phương pháp theo điểm 1, trong đó dữ liệu video chỉ báo các giá trị của bit có trọng số lớn nhất (most significant bit - MSB) của các giá trị POC của nhiều hình ảnh này.

10. Phương pháp theo điểm 1, trong đó hình ảnh tham chiếu dài hạn bao gồm hình ảnh thứ nhất, trong đó bước giải mã còn bao gồm giải mã hình ảnh thứ hai bằng cách sử dụng hình ảnh thứ nhất làm hình ảnh tham chiếu dài hạn.

11. Phương pháp theo điểm 1, trong đó phương pháp này còn bao gồm bước dự báo một phần của dữ liệu video của lát này bằng cách sử dụng hình ảnh tham chiếu dài hạn và giải mã phần dữ liệu video này dựa vào việc dự báo.

12. Phương pháp mã hóa dữ liệu video, phương pháp này bao gồm các bước:

mã hóa dữ liệu phần tử cú pháp cho nhiều hình ảnh theo thứ tự mã hóa hình ảnh, trong đó dữ liệu phần tử cú pháp chỉ báo rằng mỗi hình ảnh trong số nhiều hình ảnh này là sẵn có để sử dụng làm hình ảnh tham chiếu dài hạn;

mã hóa phần đầu lát cho lát của dữ liệu video trong đó phần đầu lát bao gồm phần tử cú pháp chỉ báo số lượng hình ảnh tham chiếu dài hạn được suy ra từ tập hợp tham số chuỗi;

xác định xem hình ảnh tham chiếu dài hạn trong nhiều hình ảnh này được báo hiệu rõ ràng trong phần đầu lát hay được suy ra từ tập hợp tham số chuỗi tương ứng với lát này;

khi hình ảnh tham chiếu dài hạn được báo hiệu rõ ràng, trong đó bước mã hóa phần đầu lát còn bao gồm bước mã hóa giá trị cho thông tin nhận dạng cho hình ảnh tham chiếu dài hạn trong phần đầu lát; và

xác định rằng giá trị cho $\Delta POC_{MSBCycleLt}[i]$ của hình ảnh tham chiếu dài hạn có giá trị chỉ số i luôn bằng giá trị cho phần tử cú pháp $\Delta poc_msb_cycle_lt[i]$ của hình ảnh tham chiếu dài hạn có giá trị chỉ số i khi hình ảnh tham chiếu dài hạn có giá trị chỉ số i bằng với số lượng hình ảnh tham chiếu dài hạn được suy ra từ tập hợp tham số chuỗi.

13. Phương pháp theo điểm 12, trong đó hình ảnh tham chiếu dài hạn bao gồm hình ảnh thứ nhất, phương pháp này còn bao gồm các bước:

lưu trữ hình ảnh được giải mã thứ nhất tương ứng với hình ảnh thứ nhất trong bộ đệm hình ảnh được giải mã;

khi bộ đệm hình ảnh được giải mã chứa nhiều hơn một hình ảnh tham chiếu mà được đánh dấu là “được sử dụng để tham chiếu” và có cùng giá trị của bit có trọng số nhỏ nhất (least significant bit - LSB) của số đếm thứ tự hình ảnh (picture order count - POC) với hình ảnh thứ nhất, trong đó bước mã hóa giá trị cho thông tin nhận dạng còn bao gồm: mã hóa phần tử cú pháp thứ nhất bằng một trong phần đầu lát của lát của hình ảnh thứ hai;

khi bộ đệm hình ảnh được giải mã không chứa nhiều hơn một hình ảnh được đánh dấu là “được sử dụng để tham chiếu” và có cùng giá trị LSB POC với hình ảnh thứ nhất, trong đó bước mã hóa giá trị cho thông tin nhận dạng còn bao gồm: mã hóa phần tử cú pháp thứ nhất bằng 0 trong phần đầu lát; và

khi phần tử cú pháp thứ nhất bằng 1, bước mã hóa giá trị cho thông tin nhận dạng còn bao gồm: mã hóa phần tử cú pháp thứ hai chỉ báo bit có trọng số lớn nhất (most significant bit - MSB) của giá trị POC của hình ảnh thứ nhất; và

mã hóa hình ảnh thứ hai bằng cách sử dụng hình ảnh thứ nhất làm hình ảnh tham chiếu dài hạn.

14. Phương pháp theo điểm 12, trong đó phương pháp này còn bao gồm bước:

suy ra hình ảnh tham chiếu dài hạn từ tập hợp tham số chuỗi dựa trên việc xác định rằng hình ảnh tham chiếu dài hạn được suy ra.

15. Phương pháp theo điểm 12, trong đó phương pháp này còn bao gồm bước:

mã hóa phần tử cú pháp trong phần đầu lát của hình ảnh hiện thời, trong đó phần tử cú pháp này bao gồm các giá trị cho bit có trọng số nhỏ nhất (least significant bit - LSB) của giá trị số đếm thứ tự hình ảnh (picture order count - POC) tương ứng của nhiều hình ảnh sao cho các giá trị cho các LSB này là không tăng hoặc không giảm theo thứ tự mã hóa hình ảnh.

16. Phương pháp theo điểm 15, trong đó bước mã hóa các giá trị cho LSB bao gồm mã hóa các giá trị của LSB sao cho đối với giá trị bất kỳ trong số j và k , trong đó j và k là các số nguyên trong khoảng 0 đến số lượng hình ảnh tham chiếu dài hạn được suy ra từ tập hợp tham số chuỗi trừ đi một, các LSB của giá trị POC của hình ảnh thứ nhất trong số nhiều hình ảnh ở vị trí j trong danh sách các hình ảnh tham chiếu được báo hiệu trong phần đầu lát là lớn hơn hoặc bằng các LSB của giá trị POC của hình ảnh thứ hai trong số nhiều hình ảnh ở vị trí k trong danh sách các hình ảnh tham chiếu.

17. Phương pháp theo điểm 16, trong đó phương pháp này còn bao gồm bước mã hóa tập hợp tham số chuỗi bao gồm thông tin chỉ báo số lượng hình ảnh tham chiếu dài hạn.

18. Phương pháp theo điểm 15, trong đó bước mã hóa các giá trị cho LSB bao gồm mã hóa các giá trị của LSB sao cho đối với giá trị bất kỳ trong số j và k , trong đó j và k là các số nguyên trong khoảng 0 đến số lượng hình ảnh tham chiếu dài hạn được suy ra từ tập hợp tham số chuỗi trừ đi một, các LSB của giá trị POC của hình ảnh thứ nhất trong số nhiều hình ảnh ở vị trí j trong danh sách các hình ảnh tham chiếu được báo hiệu trong phần đầu lát là nhỏ hơn hoặc bằng các LSB của giá trị POC của hình ảnh thứ hai trong số nhiều hình ảnh ở vị trí k trong danh sách các hình ảnh tham chiếu.

19. Phương pháp theo điểm 18, trong đó phương pháp này còn bao gồm bước mã hóa tập hợp tham số chuỗi bao gồm thông tin chỉ báo số lượng hình ảnh tham chiếu dài hạn.

20. Phương pháp theo điểm 12, trong đó dữ liệu video còn chỉ báo các giá trị của bit có trọng số lớn nhất (most significant bit - MSB) của các giá trị POC của nhiều hình ảnh này.

21. Phương pháp theo điểm 12, trong đó hình ảnh tham chiếu dài hạn bao gồm hình ảnh thứ nhất, trong đó bước mã hóa còn bao gồm mã hóa hình ảnh thứ hai bằng cách sử dụng hình ảnh thứ nhất làm hình ảnh tham chiếu dài hạn.

22. Phương pháp theo điểm 12, trong đó phương pháp này còn bao gồm bước dự báo một phần của dữ liệu video của lát này bằng cách sử dụng hình ảnh tham chiếu dài hạn và mã hóa phần dữ liệu video này dựa vào việc dự báo.

23. Thiết bị mã hóa video, thiết bị này bao gồm:

phương tiện mã hóa dữ liệu phần tử cú pháp cho nhiều hình ảnh theo thứ tự mã hóa hình ảnh, trong đó dữ liệu phần tử cú pháp chỉ báo rằng mỗi hình ảnh trong số nhiều hình ảnh này là sẵn có để sử dụng làm hình ảnh tham chiếu dài hạn;

phương tiện mã hóa phần đầu lát cho lát của dữ liệu video, trong đó phần đầu lát bao gồm phần tử cú pháp chỉ báo số lượng hình ảnh tham chiếu dài hạn được suy ra từ tập hợp tham số chuỗi;

phương tiện xác định xem hình ảnh tham chiếu dài hạn trong nhiều hình ảnh này được báo hiệu rõ ràng trong phần đầu lát hay được suy ra từ tập hợp tham số chuỗi tương ứng với lát này;

phương tiện mã hóa phần đầu lát còn bao gồm phương tiện mã hóa giá trị cho thông tin nhận dạng cho hình ảnh tham chiếu dài hạn trong phần đầu lát khi hình ảnh tham chiếu dài hạn được báo hiệu rõ ràng; và

phương tiện xác định rằng giá trị cho $\Delta POC_{MSBCycleLt}[i]$ của hình ảnh tham chiếu dài hạn có giá trị chỉ số i luôn bằng giá trị cho phần tử cú pháp $\Delta POC_{MSB_Cycle_Lt}[i]$ của hình ảnh tham chiếu dài hạn có giá trị chỉ số i khi hình ảnh tham chiếu dài hạn có giá trị chỉ số i bằng với số lượng hình ảnh tham chiếu dài hạn được suy ra từ tập hợp tham số chuỗi.

24. Thiết bị theo điểm 23, trong đó thiết bị này còn bao gồm:

phương tiện suy ra hình ảnh tham chiếu dài hạn từ tập hợp tham số chuỗi dựa trên việc xác định rằng hình ảnh tham chiếu dài hạn được suy ra.

25. Thiết bị theo điểm 23, trong đó thiết bị này còn bao gồm:

phương tiện mã hóa phần tử cú pháp trong phần đầu lát của hình ảnh hiện thời, trong đó phần tử cú pháp này bao gồm các giá trị cho bit có trọng số nhỏ nhất (least significant bit - LSB) của giá trị số đếm thứ tự hình ảnh (picture order count - POC) tương ứng của nhiều hình ảnh sao cho các giá trị cho các LSB này là không tăng hoặc không giảm theo thứ tự mã hóa hình ảnh.

26. Thiết bị theo điểm 23, trong đó phương tiện mã hóa các giá trị cho LSB bao gồm phương tiện mã hóa các giá trị cho LSB sao cho đối với giá trị bất kỳ trong số j và k , trong đó j và k là các số nguyên trong khoảng 0 đến số lượng hình ảnh tham chiếu dài hạn được suy ra từ tập hợp tham số chuỗi trừ đi một, các LSB của giá trị POC của hình ảnh thứ nhất trong số nhiều hình ảnh ở vị trí j trong danh sách các hình ảnh tham chiếu được báo hiệu trong phần đầu lát là lớn hơn hoặc bằng các LSB của giá trị POC của hình ảnh thứ hai trong số nhiều hình ảnh ở vị trí k trong danh sách các hình ảnh tham chiếu.

27. Thiết bị theo điểm 26, trong đó thiết bị này còn bao gồm phương tiện mã hóa tập hợp tham số chuỗi bao gồm thông tin chỉ báo số lượng hình ảnh tham chiếu dài hạn.

28. Thiết bị theo điểm 25, trong đó phương tiện mã hóa các giá trị cho LSB bao gồm phương tiện mã hóa các giá trị cho LSB sao cho đối với giá trị bất kỳ trong số j và k , trong đó j và k là các số nguyên trong khoảng 0 đến số lượng hình ảnh tham chiếu dài hạn được suy ra từ tập hợp tham số chuỗi trừ đi một, các LSB của giá trị POC của hình ảnh thứ nhất trong số nhiều hình ảnh ở vị trí j trong danh sách các hình ảnh tham chiếu được báo hiệu trong phần đầu lát là nhỏ hơn hoặc bằng các LSB của giá trị POC của hình ảnh thứ hai trong số nhiều hình ảnh ở vị trí k trong danh sách các hình ảnh tham chiếu.

29. Thiết bị theo điểm 28, trong đó thiết bị này còn bao gồm phương tiện mã hóa tập hợp tham số chuỗi bao gồm thông tin chỉ báo số lượng hình ảnh tham chiếu dài hạn.

30. Thiết bị theo điểm 23, trong đó dữ liệu video còn chỉ báo các giá trị của bit có trọng số lớn nhất (most significant bit - MSB) của các giá trị POC của nhiều hình ảnh này.

31. Thiết bị theo điểm 23, trong đó hình ảnh tham chiếu dài hạn bao gồm hình ảnh thứ nhất, trong đó phương tiện mã hóa bao gồm phương tiện giải mã hình ảnh thứ hai bằng cách sử dụng hình ảnh thứ nhất làm hình ảnh tham chiếu dài hạn.

32. Thiết bị theo điểm 23, trong đó hình ảnh tham chiếu dài hạn bao gồm hình ảnh thứ nhất, trong đó phương tiện mã hóa bao gồm phương tiện mã hóa hình ảnh thứ hai bằng cách sử dụng hình ảnh thứ nhất làm hình ảnh tham chiếu dài hạn.

33. Thiết bị theo điểm 23, trong đó thiết bị này còn bao gồm phương tiện dự báo một phần của dữ liệu video của lát này bằng cách sử dụng hình ảnh tham chiếu dài hạn và phương tiện mã hóa phần dữ liệu video này dựa vào việc dự báo.

34. Thiết bị mã hóa dữ liệu video, thiết bị này bao gồm:

bộ nhớ được tạo cấu hình để lưu trữ dữ liệu video; và
một hoặc nhiều bộ xử lý được tạo cấu hình để:

mã hóa dữ liệu phần tử cú pháp cho nhiều hình ảnh theo thứ tự mã hóa hình ảnh, trong đó dữ liệu phần tử cú pháp chỉ báo rằng mỗi hình ảnh trong số nhiều hình ảnh này là sẵn có để sử dụng làm hình ảnh tham chiếu dài hạn;

mã hóa phần đầu lát cho lát của dữ liệu video, trong đó phần đầu lát bao gồm phần tử cú pháp chỉ báo số lượng hình ảnh tham chiếu dài hạn được suy ra từ tập hợp tham số chuỗi;

xác định xem hình ảnh tham chiếu dài hạn trong nhiều hình ảnh này được báo hiệu rõ ràng trong phần đầu lát hay được suy ra từ tập hợp tham số chuỗi tương ứng với lát này;

khi hình ảnh tham chiếu dài hạn được báo hiệu rõ ràng, giải mã giá trị cho thông tin nhận dạng cho hình ảnh tham chiếu dài hạn trong phần đầu lát; và xác định rằng giá trị cho $\Delta POC_{MSBCycleLt}[i]$ của hình ảnh tham chiếu dài hạn có giá trị chỉ số i luôn bằng giá trị cho phần tử cú pháp $\Delta poc_msb_cycle_lt[i]$ của hình ảnh tham chiếu dài hạn có giá trị chỉ số i khi hình ảnh tham chiếu dài hạn có giá trị chỉ số i bằng với số lượng hình ảnh tham chiếu dài hạn được suy ra từ tập hợp tham số chuỗi.

35. Thiết bị theo điểm 34, trong đó hình ảnh tham chiếu dài hạn bao gồm hình ảnh thứ nhất, trong đó một hoặc nhiều bộ xử lý còn được tạo cấu hình để:

lưu trữ hình ảnh được giải mã thứ nhất tương ứng với hình ảnh thứ nhất trong bộ đệm hình ảnh được giải mã;

khi bộ đệm hình ảnh được giải mã chứa nhiều hơn một hình ảnh tham chiếu mà được đánh dấu là “được sử dụng để tham chiếu” và có cùng giá trị của bit có trọng số nhỏ nhất (least significant bit - LSB) của số đếm thứ tự hình ảnh (picture order count - POC) với hình ảnh thứ nhất, mã hóa giá trị cho thông tin nhận dạng, một hoặc nhiều bộ xử lý còn được tạo cấu hình để: mã hóa phần tử cú pháp thứ nhất bằng một trong phần đầu lát của lát của hình ảnh thứ hai;

khi bộ đệm hình ảnh được giải mã không chứa nhiều hơn một hình ảnh được đánh dấu là “được sử dụng để tham chiếu” và có cùng giá trị LSB POC với hình ảnh thứ nhất, mã hóa giá trị cho thông tin nhận dạng, một hoặc nhiều bộ xử lý còn được tạo cấu hình để:

mã hóa phần tử cú pháp thứ nhất bằng 0 trong phần đầu lát; và

khi phần tử cú pháp thứ nhất bằng một, mã hóa giá trị cho thông tin nhận dạng, một hoặc nhiều bộ xử lý còn được tạo cấu hình để:

mã hóa phần tử cú pháp thứ hai chỉ báo bit có trọng số lớn nhất (most significant bit - MSB) của giá trị POC của hình ảnh thứ nhất; và

mã hóa hình ảnh thứ hai bằng cách sử dụng hình ảnh thứ nhất làm hình ảnh tham chiếu dài hạn.

36. Thiết bị theo điểm 34, trong đó một hoặc nhiều bộ xử lý còn được tạo cấu hình để:

suy ra hình ảnh tham chiếu dài hạn từ tập hợp tham số chuỗi dựa trên việc xác định rằng hình ảnh tham chiếu dài hạn được suy ra.

37. Thiết bị theo điểm 34, trong đó một hoặc nhiều bộ xử lý còn được tạo cấu hình để:

mã hóa giá trị cho bit có trọng số nhỏ nhất (least significant bit - LSB) của giá trị số đếm thứ tự hình ảnh (picture order count - POC) của nhiều hình ảnh sao cho các giá trị cho các LSB này là không tăng hoặc không giảm theo thứ tự mã hóa hình ảnh.

38. Thiết bị theo điểm 37, trong đó để mã hóa các giá trị cho LSB, một hoặc nhiều bộ xử lý được tạo cấu hình để mã hóa các giá trị cho LSB sao cho đối với giá trị bất kỳ trong số j và k , trong đó j và k là các số nguyên, trong khoảng 0 đến số lượng hình ảnh tham chiếu dài hạn được suy ra từ tập hợp tham số chuỗi trừ đi một, các LSB của giá trị POC của hình ảnh thứ nhất trong số nhiều hình ảnh ở vị trí j trong danh sách các hình ảnh tham chiếu được báo hiệu trong phần đầu lát là lớn hơn hoặc bằng các LSB của giá trị POC của hình ảnh thứ hai trong số nhiều hình ảnh ở vị trí k trong danh sách các hình ảnh tham chiếu.

39. Thiết bị theo điểm 38, trong đó một hoặc nhiều bộ xử lý còn được tạo cấu hình để mã hóa tập hợp tham số chuỗi bao gồm thông tin chỉ báo số lượng hình ảnh tham chiếu dài hạn.

40. Thiết bị theo điểm 37, trong đó để mã hóa các giá trị cho LSB, một hoặc nhiều bộ xử lý được tạo cấu hình để mã hóa các giá trị cho LSB sao cho đối với giá trị bất kỳ trong số j và k , trong đó j và k là các số nguyên, trong khoảng 0 đến số lượng hình ảnh tham chiếu dài hạn được suy ra từ tập hợp tham số chuỗi trừ đi một, các LSB của giá trị POC của hình ảnh thứ nhất trong số nhiều hình ảnh ở vị trí j trong danh sách các hình ảnh tham chiếu được báo hiệu trong phần đầu lát là nhỏ hơn hoặc bằng các LSB của giá trị POC của hình ảnh thứ hai trong số nhiều hình ảnh ở vị trí k trong danh sách các hình ảnh tham chiếu được báo hiệu trong phần đầu lát.

41. Thiết bị theo điểm 40, trong đó một hoặc nhiều bộ xử lý còn được tạo cấu hình để mã hóa tập hợp tham số chuỗi bao gồm thông tin chỉ báo số lượng hình ảnh tham chiếu dài hạn.

42. Thiết bị theo điểm 34, trong đó dữ liệu video còn chỉ báo các giá trị của bit có trọng số lớn nhất (most significant bit - MSB) của các giá trị POC của nhiều hình ảnh này.

43. Thiết bị theo điểm 34, trong đó hình ảnh tham chiếu dài hạn bao gồm hình ảnh thứ nhất, trong đó để mã hóa, một hoặc nhiều bộ xử lý được tạo cấu hình để giải mã hình ảnh thứ hai bằng cách sử dụng hình ảnh thứ nhất làm hình ảnh tham chiếu dài hạn.

44. Thiết bị theo điểm 34, trong đó hình ảnh tham chiếu dài hạn bao gồm hình ảnh thứ nhất, trong đó để mã hóa, một hoặc nhiều bộ xử lý được tạo cấu hình để mã hóa hình ảnh thứ hai bằng cách sử dụng hình ảnh thứ nhất làm hình ảnh tham chiếu dài hạn.

45. Thiết bị theo điểm 34, trong đó thiết bị này bao gồm ít nhất một trong số:

mạch tích hợp;

bộ vi xử lý; hoặc

thiết bị truyền thông không dây.

46. Thiết bị theo điểm 34, trong đó một hoặc nhiều bộ xử lý còn được tạo cấu hình để dự báo một phần của dữ liệu video của lát này bằng cách sử dụng hình ảnh tham chiếu dài hạn và mã hóa phần dữ liệu video này dựa vào việc dự báo.

47. Vật ghi bất biến đọc được bằng máy tính trên đó lưu trữ các lệnh, mà khi được thực thi, khiến cho bộ xử lý:

mã hóa dữ liệu phần tử cú pháp cho nhiều hình ảnh theo thứ tự mã hóa hình ảnh, trong đó dữ liệu phần tử cú pháp chỉ báo số lượng hình ảnh tham chiếu dài hạn được suy ra từ tập hợp tham số chuỗi;

mã hóa phần đầu lát cho lát của dữ liệu video, trong đó phần đầu lát bao gồm phần tử cú pháp bao gồm thông tin nhận dạng cho hình ảnh tham chiếu dài hạn của nhiều hình ảnh;

xác định xem hình ảnh tham chiếu dài hạn trong số nhiều hình ảnh này được báo hiệu rõ ràng trong phần đầu lát hay được suy ra từ tập hợp tham số chuỗi tương ứng với lát này;

khi phần tử cú pháp này chỉ báo rằng thông tin nhận dạng cho hình ảnh tham chiếu dài hạn được báo hiệu rõ ràng, để mã hóa phần đầu lát, các lệnh này khiến cho bộ xử lý mã hóa giá trị cho thông tin nhận dạng cho hình ảnh tham chiếu dài hạn trong phần đầu lát; và

xác định rằng giá trị cho $\Delta POC_{MSBCycleLt}[i]$ của hình ảnh tham chiếu dài hạn có giá trị chỉ số i luôn bằng giá trị cho phần tử cú pháp $\Delta POC_{MSBCycleLt}[i]$ của hình ảnh tham chiếu dài hạn có giá trị chỉ số i khi hình ảnh tham chiếu dài hạn có giá trị chỉ số i bằng với số lượng hình ảnh tham chiếu dài hạn được suy ra từ tập hợp tham số chuỗi.

48. Vật ghi đọc được bằng máy tính theo điểm 47, trong đó vật ghi này còn bao gồm các lệnh khiến cho bộ xử lý:

suy ra hình ảnh tham chiếu dài hạn từ tập hợp tham số chuỗi khi giá trị cho thông tin nhận dạng này chỉ báo rằng hình ảnh tham chiếu dài hạn không được báo hiệu rõ ràng.

49. Vật ghi đọc được bằng máy tính theo điểm 47, trong đó vật ghi này còn bao gồm các lệnh khiến cho bộ xử lý:

mã hóa giá trị cho bit có trọng số nhỏ nhất (least significant bit - LSB) của giá trị số đếm thứ tự hình ảnh (picture order count - POC) của nhiều hình ảnh sao cho các giá trị cho các LSB này là không tăng hoặc không giảm theo thứ tự mã hóa hình ảnh.

50. Vật ghi đọc được bằng máy tính theo điểm 49, trong đó các lệnh khiến cho bộ xử lý mã hóa các giá trị cho LSB bao gồm các lệnh khiến cho bộ xử lý mã hóa các giá trị cho LSB sao cho đối với giá trị bất kỳ trong số j và k , trong đó j và k là các số nguyên, trong khoảng 0 đến số lượng hình ảnh tham chiếu dài hạn được suy ra từ tập hợp tham

số chuỗi trừ đi một, các LSB của giá trị POC của hình ảnh thứ nhất trong số nhiều hình ảnh ở vị trí j trong danh sách các hình ảnh tham chiếu được báo hiệu trong phần đầu lát là lớn hơn hoặc bằng các LSB của giá trị POC của hình ảnh thứ hai trong số nhiều hình ảnh ở vị trí k trong danh sách các hình ảnh tham chiếu.

51. Vật ghi đọc được bằng máy tính theo điểm 50, trong đó vật ghi này còn bao gồm các lệnh khiến cho bộ xử lý mã hóa tập hợp tham số chuỗi bao gồm thông tin chỉ báo số lượng hình ảnh tham chiếu dài hạn.

52. Vật ghi đọc được bằng máy tính theo điểm 49, trong đó các lệnh khiến cho bộ xử lý mã hóa các giá trị cho LSB bao gồm các lệnh khiến cho bộ xử lý mã hóa các giá trị cho LSB sao cho đối với giá trị bất kỳ trong số j và k , trong đó j và k là các số nguyên, trong khoảng 0 đến số lượng hình ảnh tham chiếu dài hạn được suy ra từ tập hợp tham số chuỗi trừ đi một, các LSB của giá trị POC của hình ảnh thứ nhất trong số nhiều hình ảnh ở vị trí j trong danh sách các hình ảnh tham chiếu được báo hiệu trong phần đầu lát là nhỏ hơn hoặc bằng các LSB của giá trị POC của hình ảnh thứ hai trong số nhiều hình ảnh ở vị trí k trong danh sách các hình ảnh tham chiếu.

53. Vật ghi đọc được bằng máy tính theo điểm 52, trong đó vật ghi này còn bao gồm các lệnh mà, khi được thực thi, khiến cho bộ xử lý mã hóa tập hợp tham số chuỗi bao gồm thông tin chỉ báo số lượng hình ảnh tham chiếu dài hạn.

54. Vật ghi đọc được bằng máy tính theo điểm 47, trong đó dữ liệu video còn bao gồm các giá trị của bit có trọng số lớn nhất (most significant bit - MSB) của các giá trị POC của nhiều hình ảnh này.

55. Vật ghi đọc được bằng máy tính theo điểm 47, trong đó hình ảnh tham chiếu dài hạn bao gồm hình ảnh thứ nhất, trong đó các lệnh khiến cho bộ xử lý mã hóa bao gồm các lệnh khiến cho bộ xử lý giải mã hình ảnh thứ hai bằng cách sử dụng hình ảnh thứ nhất làm hình ảnh tham chiếu dài hạn.

56. Vật ghi đọc được bằng máy tính theo điểm 47, trong đó hình ảnh tham chiếu dài hạn bao gồm hình ảnh thứ nhất, trong đó các lệnh này còn khiến cho bộ xử lý mã hóa hình ảnh thứ hai bằng cách sử dụng hình ảnh thứ nhất làm hình ảnh tham chiếu dài hạn.

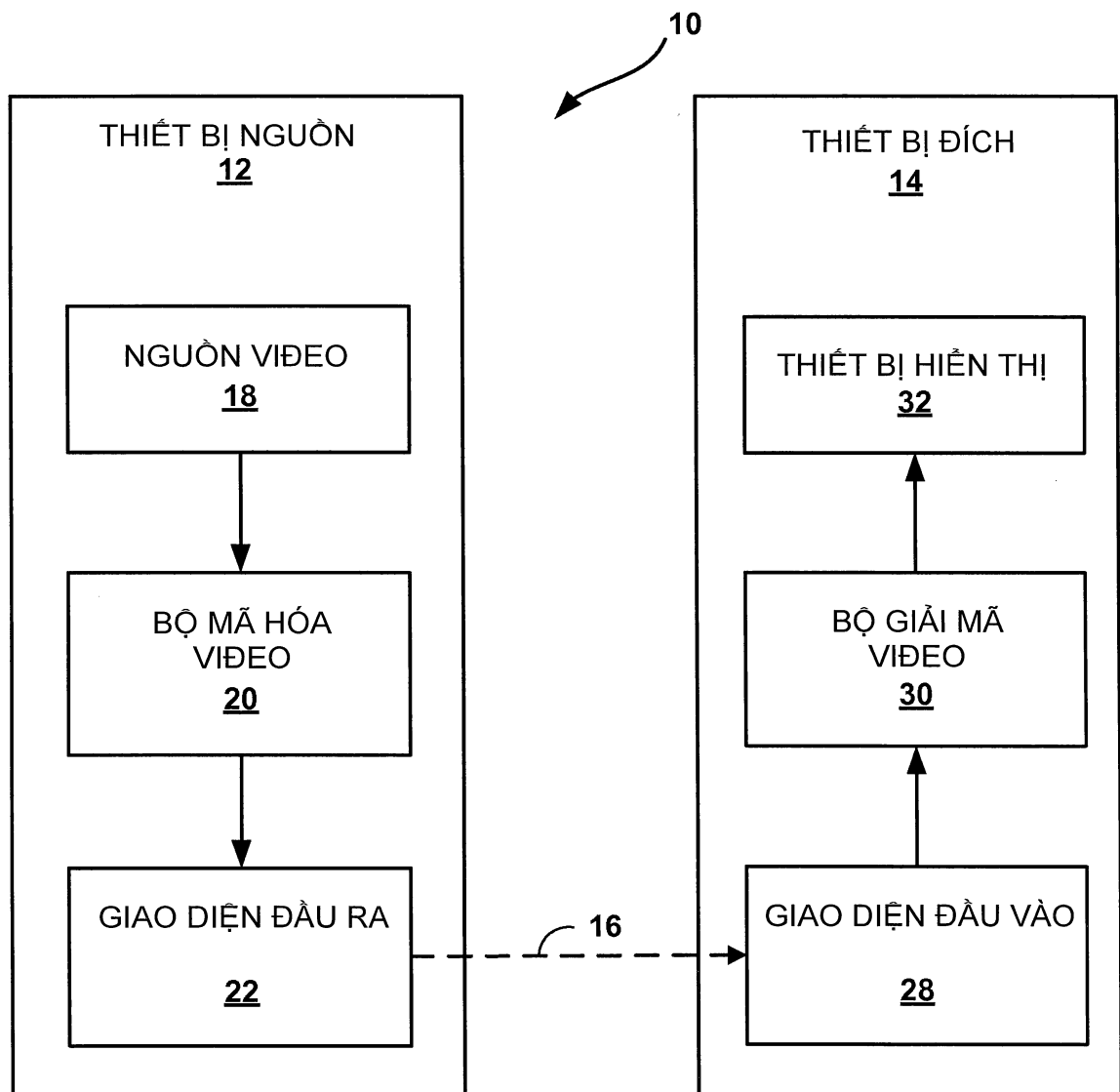
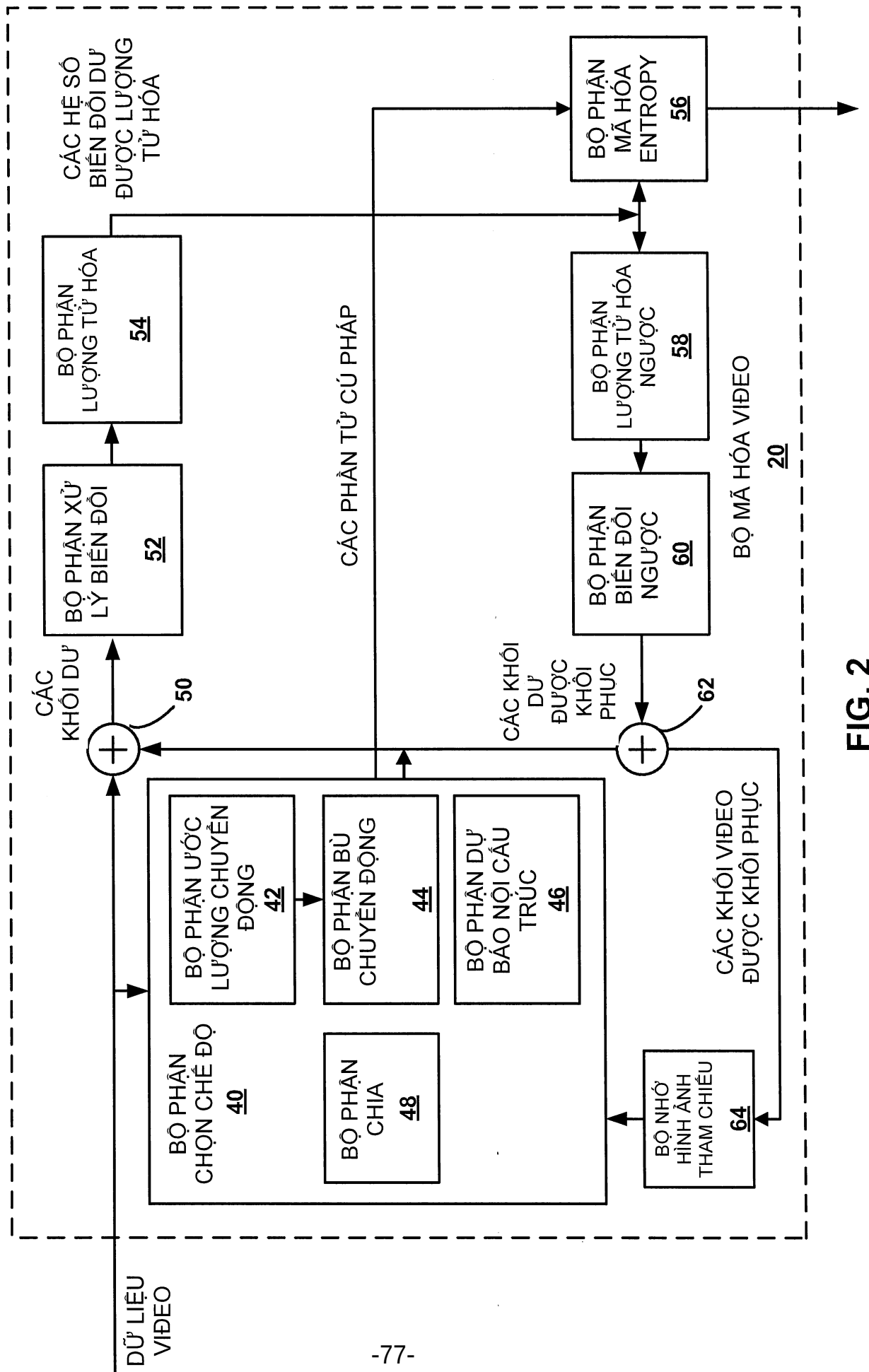


FIG. 1



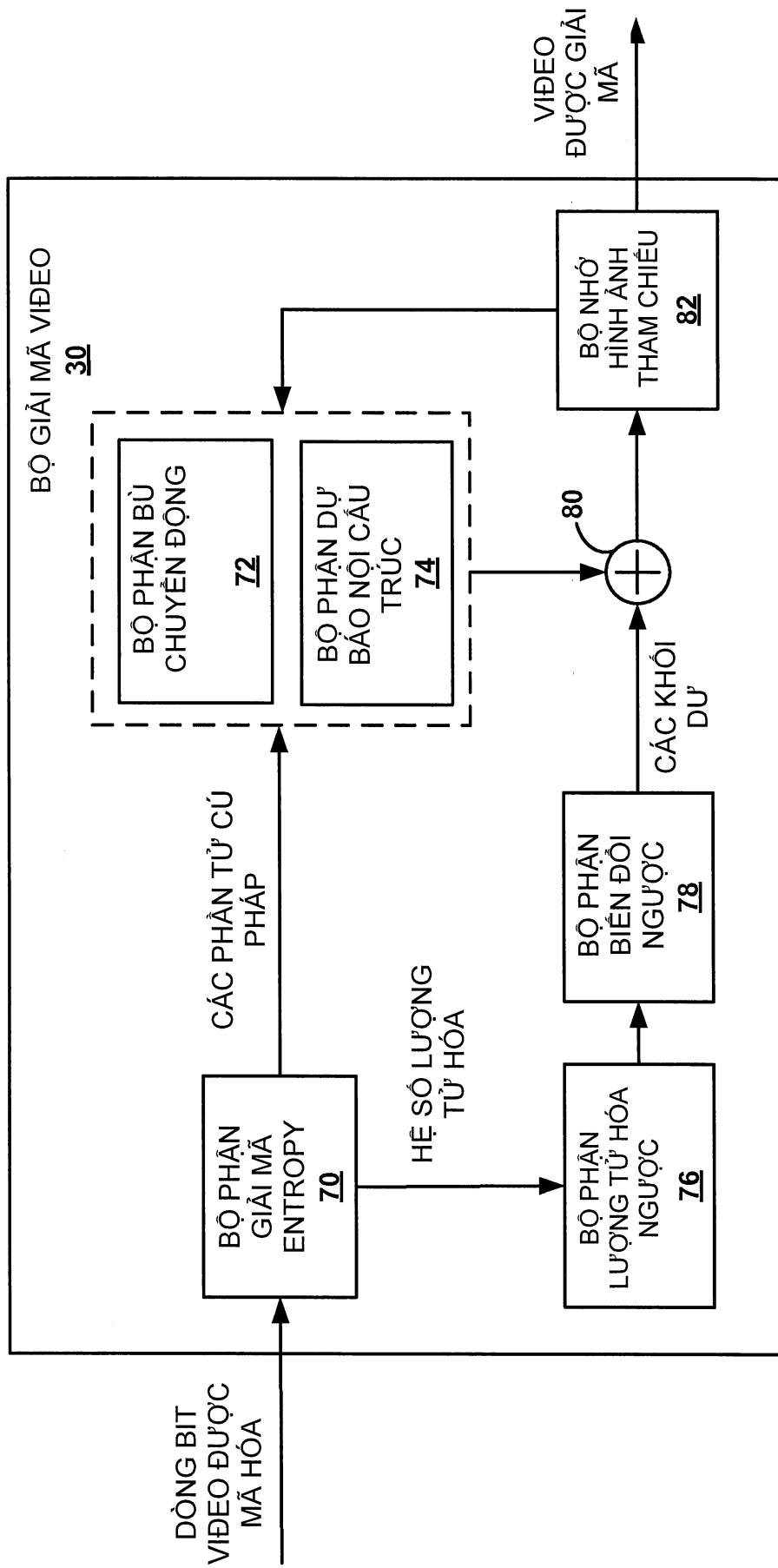


FIG. 3

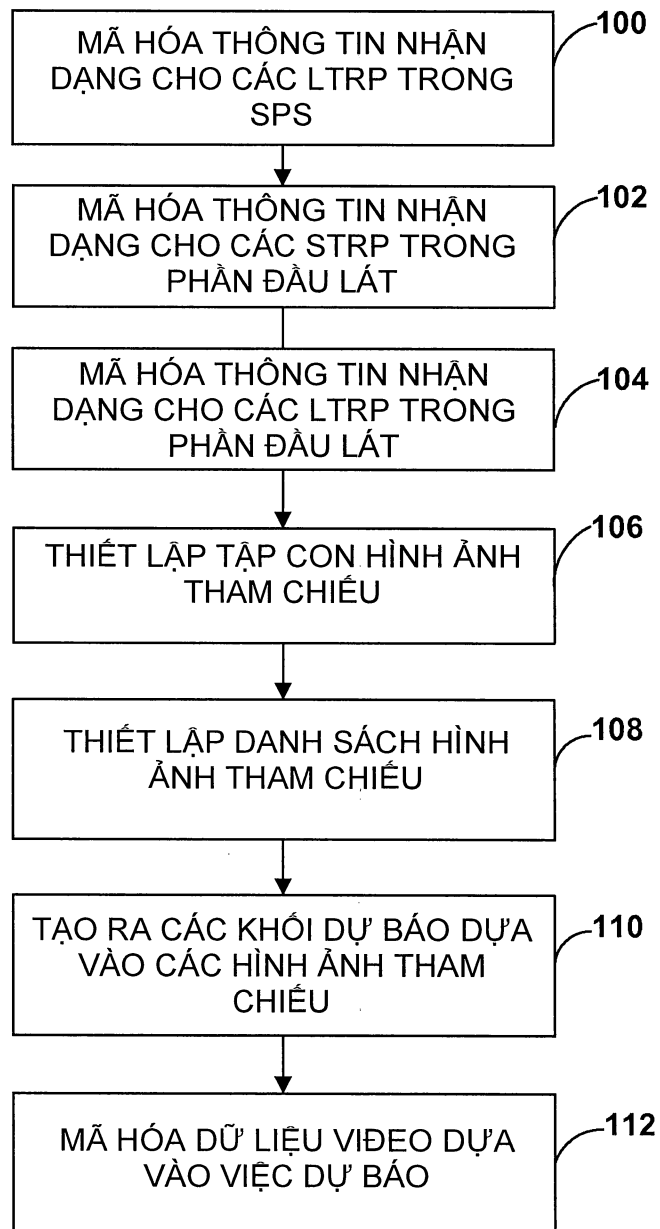


FIG. 4

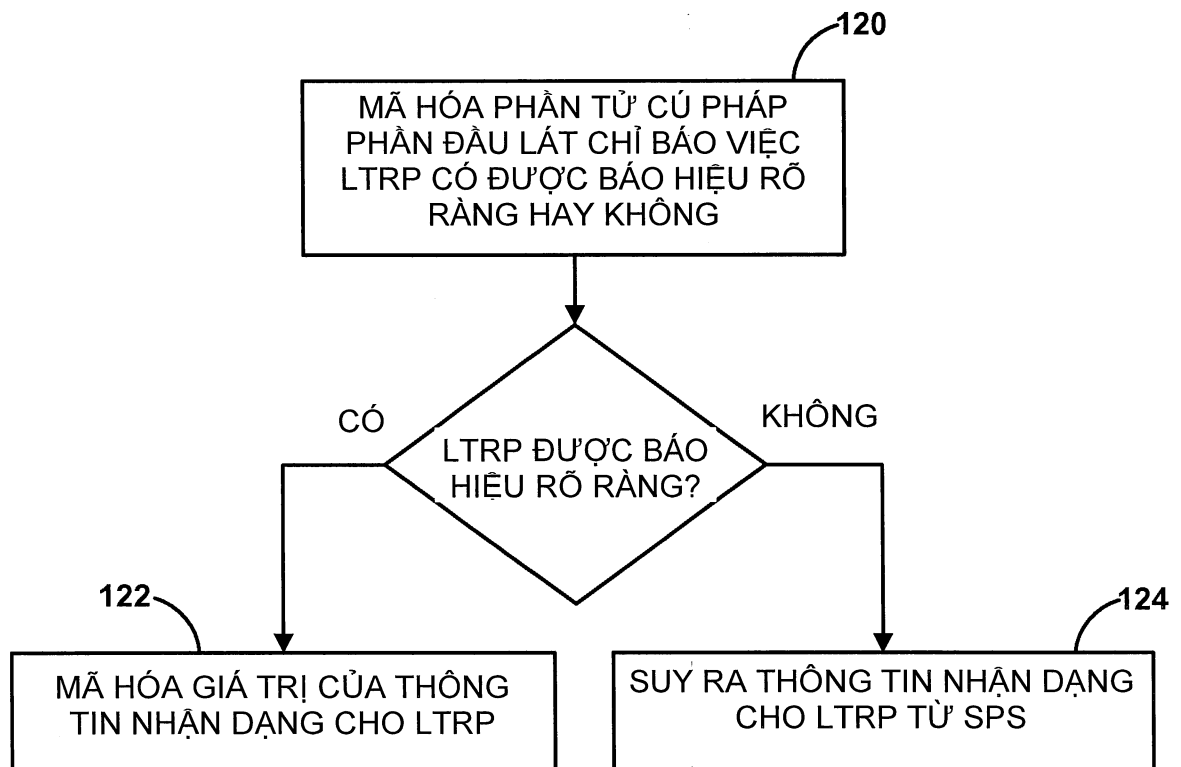


FIG. 5