



(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) 
1-0020732

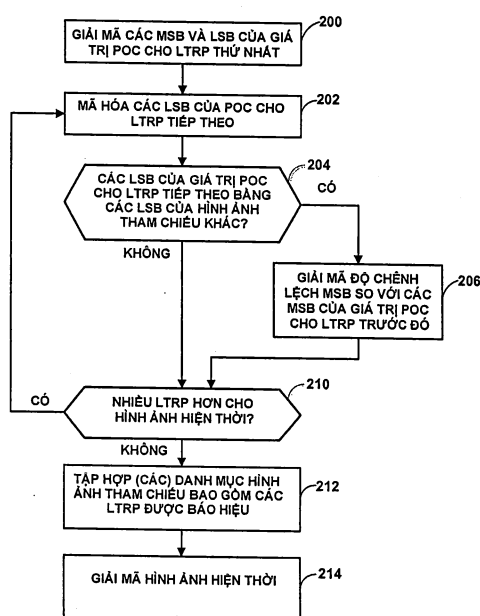
(51)⁷ H04N 7/36, 7/26

(13) B

- (21) 1-2014-04394 (22) 24.06.2013
(86) PCT/US2013/047367 24.06.2013 (87) WO2014/004391 03.01.2014
(30) 61/665,784 28.06.2012 US
13/924,016 21.06.2013 US
(45) 25.04.2019 373 (43) 25.12.2015 333
(73) Qualcomm Incorporated (US)
5775 Morehouse Drive, San Diego, California 92121, United States of America
(72) RAMASUBRAMONIAN, Adarsh Krishnan (US), WANG, Ye-Kui (US), JOSHI,
Rajan Laxman (US), CHEN, Ying (US)
(74) Công ty TNHH Quốc tế D & N (D&N INTERNATIONAL CO.,LTD.)

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

(57) Sáng chế đề cập đến bộ giải mã dữ liệu video có thể được tạo cấu hình để giải mã giá trị thứ nhất biểu diễn độ chênh lệch giữa giá trị các bit có giá trị lớn nhất (MSBs) cơ sở của giá trị số đếm thứ tự hình ảnh (POC) của hình ảnh hiện thời của dữ liệu video và giá trị MSBs thứ nhất của giá trị POC thứ nhất của hình ảnh tham chiếu dài hạn thứ nhất của dữ liệu video, giải mã giá trị thứ hai biểu diễn độ chênh lệch giữa giá trị MSBs thứ hai của giá trị POC thứ hai của hình ảnh tham chiếu dài hạn thứ hai của dữ liệu video và giá trị MSBs thứ nhất, trong đó giá trị POC thứ nhất và giá trị POC thứ hai có các giá trị các bit có giá trị nhỏ nhất khác nhau, và giải mã ít nhất một phần của hình ảnh hiện thời của dữ liệu video dựa vào ít nhất một trong số hình ảnh tham chiếu dài hạn thứ nhất và hình ảnh tham chiếu dài hạn thứ hai.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến việc mã hóa dữ liệu video.

Tình trạng kỹ thuật của sáng chế

Các tính 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 trợ giúp số cá nhân (personal digital assistant - PDA), máy tính xách tay hoặc máy tính để bàn, máy tính bảng, thiết bị đọc sách điện tử, camera số, thiết bị ghi số, thiết bị đọc đa phương tiện số, thiết bị trò chơi điện tử, bàn điều khiển trò chơi điện tử, điện thoại di động hoặc điện thoại vô tuyến vệ tinh, gọi là “điện thoại thông minh”, thiết bị hội thảo truyền hình, thiết bị truyền video liên tục, và thiết bị 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 theo MPEG-2, MPEG-4, ITU-T H.263, ITU-T H.264/MPEG-4, phần 10, mã hóa video nâng cao (Advanced Video Coding - AVC), chuẩn mã hóa video hiệu suất cao (High Efficiency Video Coding - HEVC) hiện đang được phát triển, và các phiên bản mở rộng của các chuẩn này. Dự thảo mới đây của chuẩn HEVC sắp tới, được gọi là “Dự thảo HEVC 7,” hoặc “WD7,” được mô tả trong tài liệu ICTVC-I1003, tác giả Bross và các tác giả khác, “High efficiency video coding (HEVC) text specification draft 7”, 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 9: San Jose, California, Hoa Kỳ, từ ngày 27/4 đến 7/5/2012, mà kể từ ngày 13/5/2013 có thể được tải về từ địa chỉ http://phenix.int-evey.fr/jct/doc_end_user/documents/8_San%20Jose/wg11/JCTVC-H1003-v22.zip. Các thiết bị mã hóa video 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.

Các kỹ thuật mã hóa video bao gồm thao tác dự báo không gian (nội hình ảnh) và/hoặc thao tác dự báo thời gian (liên hình ảnh) để làm giảm hoặc loại bỏ dữ liệu dư vốn có trong các chuỗi video. Để mã hóa dữ liệu video dựa vào khối, lát video (ví dụ, khung video hoặc một phần của khung video) có thể được phân tách thành các khối

video, cũng có thể được gọi là khối cấu trúc cây, đơn vị mã hóa (coding unit - CU) và/hoặc nút mã hóa. Các khối video trong lát hình ảnh mã hóa nội cấu trúc (intra-coded) (I) được mã hóa bằng cách sử dụng kỹ thuật dự báo không gian đối với các mẫu tham chiếu trong các khối lân cận trong cùng một hình ảnh. Các khối video trong lát hình ảnh mã hóa liên cấu trúc (inter-coded) (P hoặc B) có thể sử dụng kỹ thuật dự báo không gian đối với các mẫu tham chiếu trong các khối lân cận trong cùng một hình ảnh hoặc kỹ thuật dự báo thời gian đối với các mẫu tham chiếu trong các hình ảnh tham chiếu khác. Hình ảnh có thể được gọi là khung, và hình ảnh tham chiếu có thể được gọi là khung tham chiếu.

Quy trình dự báo thời gian hoặc không gian tạo ra khối dự báo cho khối cần được mã hóa. Dữ liệu dư biểu diễn độ chênh lệch điểm ảnh giữa khối gốc cần được mã hóa và khối dự báo. Khối mã hóa liên cấu trúc được mã hóa theo vectơ chuyển động chỉ đến khối gồm các mẫu tham chiếu tạo thành khối dự báo, và dữ liệu dư chỉ báo độ chênh lệch giữa khối 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 hơn nữa, dữ liệu dư có thể được biến đổi từ miền điểm ảnh sang miền biến đổi, tạo ra các hệ số biến đổi dư, sau đó có thể được lượng tử hóa. Các hệ số biến đổi đã lượng tử hóa, trước tiên được sắp xếp theo mảng hai chiều, có thể được quét để tạo ra vectơ một chiều gồm các hệ số biến đổi, và quy trình 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 báo hiệu các hình ảnh tham chiếu dài hạn để mã hóa dữ liệu video. Hình ảnh tham chiếu thường tương ứng với hình ảnh có thể được sử dụng làm tham chiếu để mã hóa dự báo thời gian dữ liệu của các hình ảnh khác. Nói chung, các hình ảnh tham chiếu dài hạn được lưu trữ trong bộ nhớ đệm hình ảnh đã giải mã lâu hơn so với các hình ảnh tham chiếu khác, ví dụ, các hình ảnh tham chiếu ngắn hạn. Sáng chế mô tả nhiều kỹ thuật khác nhau có liên quan đến việc báo hiệu các hình ảnh tham chiếu dài hạn, ví dụ báo hiệu các hình ảnh tham chiếu dài hạn ở các phần đầu lát của các lát hình ảnh. Các kỹ thuật theo sáng chế có thể cải thiện các phương pháp báo hiệu các hình ảnh tham chiếu dài hạn ở phần đầu lát, ví dụ, về hiệu quả và/hoặc khả năng áp dụng.

Theo một ví dụ, phương pháp bao gồm các bước: giải mã giá trị thứ nhất biểu diễn độ chênh lệch giữa giá trị các bit có giá trị lớn nhất (Most Significant Bit - MSB) cơ sở của giá trị số đếm thứ tự hình ảnh (picture order count - POC) của hình ảnh hiện thời của dữ liệu video và giá trị MSBs thứ nhất của giá trị POC thứ nhất của hình ảnh tham chiếu dài hạn thứ nhất của dữ liệu video, giải mã giá trị thứ hai biểu diễn độ chênh lệch giữa giá trị MSBs thứ hai của giá trị POC thứ hai của hình ảnh tham chiếu dài hạn thứ hai của dữ liệu video và giá trị MSBs thứ nhất, trong đó giá trị POC thứ nhất và giá trị POC thứ hai có các giá trị bit có giá trị nhỏ nhất (least significant bit - LSB) khác nhau, và giải mã ít nhất một phần hình ảnh hiện thời của dữ liệu video dựa vào ít nhất một trong số hình ảnh tham chiếu dài hạn thứ nhất sử dụng giá trị thứ nhất và hình ảnh tham chiếu dài hạn thứ hai sử dụng giá trị thứ nhất và giá trị thứ hai.

Theo một ví dụ khác, thiết bị giải mã dữ liệu video bao gồm bộ giải mã dữ liệu video được tạo cấu hình để giải mã giá trị thứ nhất biểu diễn độ chênh lệch giữa giá trị các bit có giá trị lớn nhất (MSB) cơ sở của giá trị số đếm thứ tự hình ảnh (POC) của hình ảnh hiện thời của dữ liệu video và giá trị MSBs thứ nhất của giá trị POC thứ nhất của hình ảnh tham chiếu dài hạn thứ nhất của dữ liệu video, giải mã giá trị thứ hai biểu diễn độ chênh lệch giữa giá trị MSBs thứ hai của giá trị POC thứ hai của hình ảnh tham chiếu dài hạn thứ hai của dữ liệu video và giá trị MSBs thứ nhất, trong đó giá trị POC thứ nhất và giá trị POC thứ hai có các giá trị bit có giá trị nhỏ nhất (LSB) khác nhau, và giải mã ít nhất một phần hình ảnh hiện thời của dữ liệu video dựa vào ít nhất một trong số hình ảnh tham chiếu dài hạn thứ nhất sử dụng giá trị thứ nhất và hình ảnh tham chiếu dài hạn thứ hai sử dụng giá trị thứ nhất và giá trị thứ hai.

Theo một ví dụ khác, thiết bị bao gồm phương tiện giải mã giá trị thứ nhất biểu diễn độ chênh lệch giữa giá trị các bit có giá trị lớn nhất (MSB) cơ sở của giá trị số đếm thứ tự hình ảnh (POC) của hình ảnh hiện thời của dữ liệu video và giá trị MSBs thứ nhất của giá trị POC thứ nhất của hình ảnh tham chiếu dài hạn thứ nhất của dữ liệu video, phương tiện giải mã giá trị thứ hai biểu diễn độ chênh lệch giữa giá trị MSBs thứ hai của giá trị POC thứ hai của hình ảnh tham chiếu dài hạn thứ hai của dữ liệu video và giá trị MSBs thứ nhất, trong đó giá trị POC thứ nhất và giá trị POC thứ hai có các giá trị bit có giá trị nhỏ nhất (LSB) khác nhau, và phương tiện giải mã ít

nhất một phần hình ảnh hiện thời của dữ liệu video dựa vào ít nhất một trong số hình ảnh tham chiếu dài hạn thứ nhất sử dụng giá trị thứ nhất và hình ảnh tham chiếu dài hạn thứ hai sử dụng giá trị thứ nhất và giá trị thứ hai.

Theo một ví dụ khác, vật ghi đọc được bằng máy tính lưu trữ trên đó các lệnh mà, khi được thực thi, khiến cho bộ xử lý giải mã giá trị thứ nhất biểu diễn độ chênh lệch giữa giá trị các bit có giá trị lớn nhất (MSB) cơ sở của giá trị số đếm thứ tự hình ảnh (POC) của hình ảnh hiện thời của dữ liệu video và giá trị MSBs thứ nhất của giá trị POC thứ nhất của hình ảnh tham chiếu dài hạn thứ nhất của dữ liệu video, giải mã giá trị thứ hai biểu diễn độ chênh lệch giữa giá trị MSBs thứ hai của giá trị POC thứ hai của hình ảnh tham chiếu dài hạn thứ hai của dữ liệu video và giá trị MSBs thứ nhất, trong đó giá trị POC thứ nhất và giá trị POC thứ hai có các giá trị bit có giá trị nhỏ nhất (LSB) khác nhau, và giải mã ít nhất một phần hình ảnh hiện thời của dữ liệu video dựa vào ít nhất một trong số hình ảnh tham chiếu dài hạn thứ nhất sử dụng giá trị thứ nhất và hình ảnh tham chiếu dài hạn thứ hai sử dụng giá trị thứ nhất và giá trị thứ hai.

Theo một ví dụ khác, phương pháp mã hóa dữ liệu video bao gồm các bước: mã hóa ít nhất một phần thứ nhất của hình ảnh hiện thời của dữ liệu video dựa vào hình ảnh tham chiếu dài hạn thứ nhất và ít nhất một phần thứ hai của hình ảnh hiện thời dựa vào hình ảnh tham chiếu dài hạn thứ hai, mã hóa giá trị thứ nhất biểu diễn độ chênh lệch giữa giá trị các bit có giá trị lớn nhất (MSB) cơ sở của giá trị số đếm thứ tự hình ảnh (POC) của hình ảnh hiện thời của dữ liệu video và giá trị MSBs thứ nhất của giá trị POC thứ nhất của hình ảnh tham chiếu dài hạn thứ nhất của dữ liệu video, và mã hóa giá trị thứ hai biểu diễn độ chênh lệch giữa giá trị MSBs thứ hai của giá trị POC thứ hai của hình ảnh tham chiếu dài hạn thứ hai của dữ liệu video và giá trị MSBs thứ nhất, trong đó giá trị POC thứ nhất và giá trị POC thứ hai có các giá trị bit có giá trị nhỏ nhất (LSB) khác nhau.

Theo một ví dụ khác, thiết bị mã hóa dữ liệu video bao gồm bộ mã hóa dữ liệu video được tạo cấu hình để mã hóa ít nhất một phần thứ nhất của hình ảnh hiện thời của dữ liệu video dựa vào hình ảnh tham chiếu dài hạn thứ nhất và ít nhất một phần thứ hai của hình ảnh hiện thời dựa vào hình ảnh tham chiếu dài hạn thứ hai, mã hóa giá trị thứ nhất biểu diễn độ chênh lệch giữa giá trị các bit có giá trị lớn nhất (MSB) cơ sở của giá trị số đếm thứ tự hình ảnh (POC) của hình ảnh hiện thời của dữ liệu

video và giá trị MSBs thứ nhất của giá trị POC thứ nhất của hình ảnh tham chiếu dài hạn thứ nhất của dữ liệu video, và mã hóa giá trị thứ hai biểu diễn độ chênh lệch giữa giá trị MSBs thứ hai của giá trị POC thứ hai của hình ảnh tham chiếu dài hạn thứ hai của dữ liệu video và giá trị MSBs thứ nhất, trong đó giá trị POC thứ nhất và giá trị POC thứ hai có các giá trị bit có giá trị nhỏ nhất (LSB) khác nhau.

Theo một ví dụ khác, thiết bị mã hóa dữ liệu video bao gồm phương tiện mã hóa ít nhất một phần thứ nhất của hình ảnh hiện thời của dữ liệu video dựa vào hình ảnh tham chiếu dài hạn thứ nhất và ít nhất một phần thứ hai của hình ảnh hiện thời dựa vào hình ảnh tham chiếu dài hạn thứ hai, phương tiện mã hóa giá trị thứ nhất biểu diễn độ chênh lệch giữa giá trị các bit có giá trị lớn nhất (MSB) cơ sở của giá trị số đếm thứ tự hình ảnh (POC) của hình ảnh hiện thời của dữ liệu video và giá trị MSBs thứ nhất của giá trị POC thứ nhất của hình ảnh tham chiếu dài hạn thứ nhất của dữ liệu video, và thiết bị mã hóa giá trị thứ hai biểu diễn độ chênh lệch giữa giá trị MSBs thứ hai của giá trị POC thứ hai của hình ảnh tham chiếu dài hạn thứ hai của dữ liệu video và giá trị MSBs thứ nhất, trong đó giá trị POC thứ nhất và giá trị POC thứ hai có các giá trị bit có giá trị nhỏ nhất (LSB) khác nhau.

Theo một ví dụ khác, vật ghi đọc được bằng máy tính lưu trữ trên đó các lệnh mà, khi được thực thi, khiến cho bộ xử lý mã hóa ít nhất một phần thứ nhất của hình ảnh hiện thời của dữ liệu video dựa vào hình ảnh tham chiếu dài hạn thứ nhất và ít nhất một phần thứ hai của hình ảnh hiện thời dựa vào hình ảnh tham chiếu dài hạn thứ hai, mã hóa giá trị thứ nhất biểu diễn độ chênh lệch giữa các bit có giá trị lớn nhất (MSB) cơ sở của số đếm thứ tự hình ảnh (POC) của hình ảnh hiện thời của dữ liệu video và giá trị MSBs thứ nhất của giá trị POC thứ nhất của hình ảnh tham chiếu dài hạn thứ nhất của dữ liệu video, và mã hóa giá trị thứ hai biểu diễn độ chênh lệch giữa giá trị MSBs thứ hai của giá trị POC thứ hai của hình ảnh tham chiếu dài hạn thứ hai của dữ liệu video và giá trị MSBs thứ nhất, trong đó giá trị POC thứ nhất và giá trị POC thứ hai có các giá trị bit có giá trị nhỏ nhất (LSB) khác nhau.

Chi tiết về một hoặc nhiều ví dụ được nêu trong các hình vẽ kèm theo và phần mô tả dưới đây. Các dấu hiệu, mục đích và ưu điểm khác sẽ được nêu rõ ràng hơn trong phần mô tả và hình vẽ, và yêu cầu bảo hộ.

Mô tả vắn tắt các hình vẽ

Fig.1 là sơ đồ khối minh họa hệ thống giải mã và mã hóa dữ liệu video làm ví dụ có thể sử dụng các kỹ thuật để báo hiệu tốt hơn các hình ảnh tham chiếu dài hạn ở phần đầu lát.

Fig.2 là sơ đồ khối minh họa ví dụ về bộ mã hóa dữ liệu video có thể thực hiện các kỹ thuật để báo hiệu tốt hơn các hình ảnh tham chiếu dài hạn ở phần đầu lát.

Fig.3 là sơ đồ khối minh họa ví dụ về bộ giải mã dữ liệu video có thể thực hiện các kỹ thuật để báo hiệu tốt hơn các hình ảnh tham chiếu dài hạn ở phần đầu lát.

Fig.4 là sơ đồ khái niệm minh họa chuỗi hình ảnh video được mã hóa.

Fig.5 là lưu đồ minh họa phương pháp làm ví dụ để mã hóa các giá trị số đếm thứ tự hình ảnh (POC) của các hình ảnh tham chiếu dài hạn theo các kỹ thuật của sáng chế.

Fig.6 là lưu đồ minh họa phương pháp làm ví dụ để giải mã các giá trị POC của các hình ảnh tham chiếu dài hạn theo các kỹ thuật của sáng chế.

Mô tả chi tiết sáng chế

Nói chung, dữ liệu video được biểu diễn bởi chuỗi các hình ảnh được chụp hoặc được hiển thị theo trình tự nhanh. Cần hiểu rằng trong một vài ví dụ, hình ảnh hoặc các phần của hình ảnh có thể được tạo ra, ví dụ, bằng cách sử dụng đồ họa máy tính, thay vì (hoặc ngoài việc) được chụp. Thứ tự trong đó các hình ảnh được hiển thị (nói chung có thể giống với thứ tự trong đó các hình ảnh được chụp hoặc được tạo ra) có thể khác với thứ tự trong đó các hình ảnh được mã hóa. Thứ tự hiển thị của hình ảnh thường được biểu diễn bởi các giá trị số đếm thứ tự hình ảnh (POC), trong khi đó thứ tự mã hóa của hình ảnh thường được biểu diễn bởi các giá trị số khung (frame_num).

Mã hóa hình ảnh thường bao gồm tận dụng dữ liệu dư xuất hiện trong hình ảnh. Ví dụ, mã hóa không gian, cũng được gọi là mã hóa dự báo nội cấu trúc, tận dụng các dữ liệu dư xuất hiện giữa các khối lân cận trong không gian của các giá trị điểm ảnh của hình ảnh chung. Theo một ví dụ khác, mã hóa thời gian, cũng được gọi là mã hóa dự báo liên cấu trúc, tận dụng các dữ liệu dư xuất hiện giữa các hình ảnh khác nhau được chụp, được tạo ra, hoặc được hiển thị ở các thời điểm khác nhau. Cụ thể hơn, hình ảnh có thể được chia thành các lát, mà có thể được chỉ định để dự báo nội cấu trúc hoặc liên cấu trúc. Hơn nữa, mỗi lát có thể được chia thành các khối (ví

dự, đơn vị mã hóa lớn nhất (largest coding unit - LCU)), và mỗi khối có thể được phân tách tiếp hoặc được mã hóa dựa vào việc lát tương ứng được chỉ định để dự báo nội cấu trúc hay liên cấu trúc.

Liên quan đến ví dụ về mã hóa dự báo liên cấu trúc, các khối hình ảnh hiện đang được mã hóa (cũng được gọi là “hình ảnh hiện thời”) có thể được dự báo từ hình ảnh tham chiếu. WD7 xác định “hình ảnh tham chiếu” là hình ảnh có nal_ref_flag bằng 1. nal_ref_flag là phần tử cú pháp của đơn vị lớp trừu tượng mạng (network abstraction layer - NAL) chỉ báo việc dữ liệu có trong đơn vị NAL có được coi là hình ảnh tham chiếu hay không. WD7 cũng đề xuất rằng hình ảnh tham chiếu chứa các mẫu (tức là, các giá trị điểm ảnh) mà có thể được sử dụng để dự báo liên cấu trúc trong quy trình giải mã các hình ảnh tiếp theo theo thứ tự giải mã.

Hơn nữa, WD7 phân biệt các hình ảnh tham chiếu dài hạn với các hình ảnh tham chiếu ngắn hạn. Ví dụ, WD7 xác định hình ảnh tham chiếu dài hạn là hình ảnh tham chiếu được đánh dấu là “được sử dụng để làm tham chiếu dài hạn”. WD7 đề xuất rằng cờ trong tập thông số chuỗi (sequence parameter set - SPS) chỉ báo việc các hình ảnh tham chiếu dài hạn có được báo hiệu cho chuỗi video được mã hóa hay không. Theo WD7, các bộ mã hóa dữ liệu video báo hiệu các bit có giá trị nhỏ nhất (least significant bit - LSB) của các giá trị POC của các hình ảnh tham chiếu dài hạn ở phần đầu lát của các lát. Việc báo hiệu các LSB có thể dẫn đến tiết kiệm bit, so với báo hiệu giá trị POC đầy đủ. Bộ mã hóa dữ liệu video, theo WD7, có thể báo hiệu bit có giá trị lớn nhất (MSB) của các giá trị POC của các hình ảnh tham chiếu dài hạn nếu có nhiều hơn một hình ảnh tham chiếu trong bộ nhớ đệm hình ảnh đã giải mã (DPB) có cùng số bit LSB như hình ảnh tham chiếu dài hạn. Hơn nữa, theo WD7, các bộ mã hóa dữ liệu video có thể sử dụng cờ để chỉ báo việc các hình ảnh tham chiếu dài hạn được báo hiệu có được sử dụng bởi hình ảnh hiện thời để làm tham chiếu hay không.

Các bộ mã hóa dữ liệu video có thể giữ các MSB của hình ảnh hiện thời, ở đây được gọi là “giá trị MSB cơ sở,” và báo hiệu các MSB của hình ảnh tham chiếu dài hạn i bằng cách sử dụng các giá trị độ lệch so với các giá trị MSB cơ sở, ví dụ, bằng cách sử dụng phần tử cú pháp `delta_poc_msb_cycle_lt[i]` cho hình ảnh i . Phần tử cú pháp này dưới đây được gọi là giá trị “chu kỳ MSB”. Nói cách khác, để tính toán giá trị POC đối với một hình ảnh cụ thể, bộ mã hóa dữ liệu video có thể tính

toán: (MSB cơ sở + chu kỳ MSB đối với hình ảnh này) được nối kết với (LSB đối với hình ảnh này).

Trong WD7, giá trị chu kỳ MSB được mã hóa dưới dạng độ chênh lệch giữa giá trị MSB cơ sở và các MSB đối với hình ảnh mà chu kỳ MSB được báo hiệu, tức là hình ảnh i . Chỉ khi hai hoặc nhiều hình ảnh tham chiếu dài hạn có cùng giá trị LSB được báo hiệu, thì chu kỳ MSB của hình ảnh tham chiếu dài hạn thứ nhất sẽ được báo hiệu như vậy và các giá trị chu kỳ MSB của phần còn lại của các hình ảnh tham chiếu dài hạn được báo hiệu dưới dạng các độ lệch tương đối của nhau. Tuy nhiên, việc báo hiệu các hình ảnh tham chiếu dài hạn này, theo WD7, có thể có một hoặc nhiều thiếu sót. Ví dụ, nếu có hai hình ảnh tham chiếu dài hạn với các giá trị LSB khác nhau, nhưng cả hai hình ảnh này đều có $\text{delta_poc_msb_cycle_lt}[i]$ được báo hiệu, thì cú pháp hiện thời sẽ báo hiệu $\text{delta_poc_msb_cycle_lt}[i]$ như được nêu trên đây, tức là, dưới dạng độ lệch so với giá trị MSB cơ sở. Tuy nhiên, sáng chế cho rằng việc báo hiệu độ chênh lệch giữa giá trị chu kỳ MSB cho các hình ảnh có thể hiệu quả hơn.

Việc báo hiệu dữ liệu xác định của các hình ảnh tham chiếu dài hạn ở phần đầu lát, theo WD7, có thể có một hoặc nhiều thiếu sót. Ví dụ, nếu hình ảnh hiện thời có giá trị POC bằng $\text{MaxPicOrderCntLsb} * N - 1$, và hình ảnh mà có giá trị POC bằng $\text{MaxPicOrderCntLsb} * N$ là hình ảnh tham chiếu dài hạn thứ nhất mà $\text{delta_poc_msb_cycle_lt}[i]$ được báo hiệu, sau đó WD7 không đề xuất các kỹ thuật báo hiệu $\text{delta_poc_msb_cycle_lt}[i]$. Giá trị của $\text{delta_poc_msb_cycle_lt}[i]$, trong trường hợp này, nên là, nhưng không thể là, -1 (âm một).

Theo một ví dụ khác, nếu có hai hình ảnh tham chiếu dài hạn với các giá trị LSB khác nhau, nhưng cả hai đều có $\text{delta_poc_msb_cycle_lt}[i]$ được báo hiệu, cú pháp hiện thời sẽ báo hiệu $\text{delta_poc_msb_cycle_lt}[i]$ như vậy. Tuy nhiên, sáng chế cho rằng việc báo hiệu độ chênh lệch cho một trong số các hình ảnh sẽ hiệu quả hơn.

Theo một ví dụ khác, khi có nhiều hơn một hình ảnh tham chiếu trong DPB mà có LSB bằng LSB của hình ảnh tham chiếu dài hạn được báo hiệu, WD7 cho rằng chu kỳ MSB của LTRP cũng được báo hiệu bằng cách sử dụng phần tử cú pháp $\text{delta_poc_msb_cycle_lt}[i]$. Tuy nhiên, sáng chế cho rằng, nếu có một hình ảnh trong DPB mà được đánh dấu là hình ảnh tham chiếu ngắn hạn bởi hình ảnh hiện thời, thì việc đưa hình ảnh tham chiếu ngắn hạn vào tập hợp dự phòng hình ảnh tham chiếu dài hạn có thể kết thúc việc báo hiệu số bit MSB nhiều hơn cần thiết.

Theo một ví dụ khác, khi có n hình ảnh tham chiếu trong DPB mà có LSB POC bằng $\text{poc_lsb_lt}[i]$, và trong đó k hình ảnh tham chiếu dài hạn, trong đó k nhỏ hơn hoặc bằng n , được báo hiệu trong hình ảnh hiện thời có LSB bằng $\text{poc_lsb_lt}[i]$, cú pháp được xác định trong WD7 cho rằng tất cả k hình ảnh sẽ có $\text{delta_poc_msb_present_flag}[i]$ được đặt bằng 1. Tuy nhiên, nếu $k = n$, sáng chế mô tả rằng bộ mã hóa dữ liệu video không cần phải đặt $\text{delta_poc_msb_present_flag}[i]$ bằng 1 đối với tất cả k hình ảnh tham chiếu dài hạn, nhưng cần thiết phải đặt $\text{delta_poc_msb_present_flag}[i]$ bằng 1 đối với $n - 1$ hình ảnh. Các kỹ thuật theo sáng chế, như được mô tả chi tiết hơn liên quan các hình vẽ dưới đây, có thể khắc phục bất kỳ hoặc tất cả các thiếu sót này. Nói chung, sáng chế mô tả các kỹ thuật khác nhau để có thể cải thiện việc báo hiệu các hình ảnh tham chiếu dài hạn ở phần đầu lát.

Fig.1 là sơ đồ khối minh họa hệ thống giải mã và mã hóa dữ liệu video 10 làm ví dụ có thể sử dụng các kỹ thuật để báo hiệu tốt hơn các hình ảnh tham chiếu dài hạn ở phần đầu lát. Như được thể hiện trên Fig.1, hệ thống 10 bao gồm thiết bị nguồn 12 để cung cấp dữ liệu video mã hóa cần được giải mã sau đó bằng thiết bị đích 14. Cụ thể là, thiết bị nguồn 12 cung cấp dữ liệu video cho thiết bị đích 14 thông qua vật ghi đọc được bằng máy tính 16. Thiết bị nguồn 12 và thiết bị đích 14 có thể bao gồm thiết bị bất kỳ trong số nhiều thiết bị, bao gồm máy tính để bàn, máy tính sổ tay (tức là, máy tính xách tay), máy tính bảng, hộp giải mã truyền hình, thiết bị điện thoại cầm tay như thiết bị gọi là máy điện thoại “thông minh”, máy tính bảng “thông minh”, máy thu hình, camera, thiết bị hiển thị, thiết bị đọc đa phương tiện số, bàn điều khiển trò chơi điện tử, thiết bị truyền video liên tục, hoặc thiết bị tương tự. Trong một vài 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 cần được giải mã thông qua vật ghi đọc được bằng máy tính 16. Vật ghi đọc được bằng máy tính 16 có thể bao gồm loại phương tiện hoặc thiết bị bất kỳ có khả năng chuyển dữ liệu video mã hóa từ thiết bị nguồn 12 đến thiết bị đích 14. Theo một ví dụ, vật ghi đọc được bằng máy tính 16 có thể bao gồm phương tiện truyền thông cho phép thiết bị nguồn 12 truyền trực tiếp dữ liệu video mã hóa đến thiết bị đích 14 trong 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 có thể bao

gồm phương tiện truyền thông không dây hoặc nối dây bất kỳ, như phổ tần số vô tuyến (radio frequency - RF) hoặc một hoặc nhiều đường truyền vật lý. Phương tiện truyền thông 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 có thể bao gồm bộ định tuyến, bộ chuyển mạch, trạm cơ sở, hoặc thiết bị khác bất kỳ để tạo điều kiện truyền thông từ thiết bị nguồn 12 đến thiết bị đích 14.

Theo một vài ví dụ, dữ liệu được mã hóa có thể được xuất ra từ giao diện xuất 22 đến thiết bị lưu trữ. Tương tự, dữ liệu được mã hóa có thể được truy nhập từ thiết bị lưu trữ bằng giao diện nhập. Thiết bị lưu trữ có thể bao gồm phương tiện bất kỳ trong số nhiều phương tiện lưu trữ dữ liệu truy nhập cục bộ hoặc phân tán khác nhau như ổ cứng, đĩa Blu-ray, đĩa DVD, đĩa CD-ROM, bộ nhớ tác động nhanh, bộ nhớ khả biến hoặc không khả biến, hoặc phương tiện lưu trữ số thích hợp khác bất kỳ để lưu trữ dữ liệu video mã hóa. Theo một ví dụ khác, thiết bị lưu trữ có thể tương ứng với máy chủ tệp hoặc thiết bị lưu trữ trung gian khác để lưu trữ video mã hóa được tạo ra bởi thiết bị nguồn 12.

Thiết bị đích 14 có thể truy nhập dữ liệu video lưu trữ từ thiết bị lưu trữ bằng cách truyền liên tục hoặc tải xuống. Máy chủ tệp 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. Máy chủ tệp làm ví dụ bao gồm máy chủ mạng (dùng cho website chẳng hạn), máy chủ giao thức chuyển tệp (File Transfer Protocol - FTP), thiết bị lưu trữ kết nối với mạng (network attached storage - NAS), hoặc ổ đĩa cục bộ. Thiết bị đích 14 có thể truy nhập dữ liệu video mã hóa qua kết nối dữ liệu tiêu chuẩn bất kỳ, kể cả kết nối Internet. Kiểu kết nối dữ liệu này bao gồm kênh không dây (ví dụ, kết nối Wi-Fi), kết nối dây (ví dụ, đường thuê bao số (Digital Subscriber Line - DSL, môdem cáp, v.v.), hoặc kết hợp cả hai kiểu này thích hợp để truy nhập dữ liệu video mã hóa lưu trữ trong máy chủ tệp. Việc truyền dữ liệu video mã hóa từ thiết bị lưu trữ có thể là truyền liên tục, tải xuống, hoặc kết hợp cả hai kiểu này.

Các kỹ thuật theo sáng chế không nhất thiết giới hạn ở các ứng dụng hoặc các thiết lập không dây. Các kỹ thuật này có thể được áp dụng cho việc mã hóa video để hỗ trợ cho ứng dụng đa phương tiện bất kỳ trong nhiều loại ứng dụng đa phương tiện khác nhau, như truyền hình vô tuyến, truyền hình cáp, truyền hình vệ tinh, truyền video liên tục qua Internet, truyền thích ứng động qua HTTP (dynamic adaptive

streaming over HTTP - DASH), dữ liệu video số được mã hóa trên phương tiện lưu trữ dữ liệu, giải mã dữ liệu video số lưu trữ trong phương tiện lưu trữ dữ liệu, hoặc các ứng dụng khác. Theo một số ví dụ, hệ thống 10 có thể được tạo cấu hình để hỗ trợ truyền video một chiều hoặc hai chiều nhằm hỗ trợ cho các ứng dụng như truyền video liên tục, phát lại video, phát rộng video, và/hoặc điện thoại truyền hình.

Theo ví dụ nêu trên Fig.1, thiết bị nguồn 12 bao gồm nguồn video 18, bộ mã hóa dữ liệu video 20, và giao diện xuất 22. Thiết bị đích 14 bao gồm giao diện nhập 28, bộ giải mã dữ liệu video 30, và thiết bị hiển thị 32. Theo sáng chế, bộ mã hóa dữ liệu video 20 của thiết bị nguồn 12 có thể được tạo cấu hình để áp dụng các kỹ thuật để báo hiệu tốt hơn các hình ảnh tham chiếu dài hạn ở phần đầu lát. Theo các ví dụ khác, thiết bị nguồn và thiết bị đích có thể bao gồm bộ phận hoặc cơ cấu khác. Ví dụ, thiết bị nguồn 12 có thể thu dữ liệu video từ nguồn dữ liệu video bên ngoài 18, như camera bên ngoài. Mặt khác, thiết bị đích 14 có thể giao diện với thiết bị hiển thị bên ngoài, mà không có thiết bị hiển thị tích hợp.

Hệ thống 10 được thể hiện trên Fig.1 chỉ là một ví dụ. Các kỹ thuật để báo hiệu tốt hơn các hình ảnh tham chiếu dài hạn ở phần đầu lát có thể được thực hiện bởi thiết bị giải mã và/hoặc mã hóa video số. Mặc dù thông thường các kỹ thuật theo sáng chế được thực hiện bởi thiết bị mã hóa dữ liệu video, các kỹ thuật này cũng có thể được thực hiện bởi bộ mã hóa/bộ giải mã dữ liệu video, thường được gọi là “CODEC”. Hơn nữa, các kỹ thuật theo sáng chế cũng có thể được thực hiện bởi bộ xử lý video. Thiết bị nguồn 12 và thiết bị đích 14 chỉ là 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 đều có các bộ phận mã hóa và giải mã dữ liệu video. Vì vậy, hệ thống 10 có thể hỗ trợ truyền dữ liệu video một chiều hoặc hai chiều giữa các thiết bị video 12, 14, ví dụ, để truyền liên tục dữ liệu video, phát lại dữ liệu video, phát rộng dữ liệu video, hoặc điện thoại truyền hình.

Nguồn dữ liệu video 18 của thiết bị nguồn 12 có thể là thiết bị ghi video, như camera ghi hình, phương tiện lưu trữ dữ liệu video có dữ liệu video đã được ghi từ trước, và/hoặc giao diện nạp dữ liệu video để thu dữ liệu video từ nhà cung cấp nội dung video. Theo phương án khác, nguồn dữ liệu video 18 có thể tạo ra dữ liệu đồ họa máy tính dùng làm nguồn dữ liệu video, hoặc kết hợp dữ liệu video trực tiếp, dữ

liệu video đã được ghi từ trước và dữ liệu video được tạo ra trên máy tính. Trong một số trường hợp, nếu nguồn dữ liệu video 18 là camera ghi hình, thì thiết bị nguồn 12 và thiết bị đích 14 có thể tạo nên thiết bị được gọi là máy điện thoại có camera hoặc máy điện thoại có truyền hình. Tuy nhiên, như đã nêu trên, các kỹ thuật nêu trong sáng chế có thể áp dụng để mã hóa dữ liệu video nói chung, và có thể được áp dụng cho các ứng dụng không dây và/hoặc nối dây. Trong mỗi trường hợp, dữ liệu video được ghi, được ghi từ trước hoặc được tạo ra trên máy tính có thể được mã hóa bằng bộ mã hóa dữ liệu video 20. Sau đó, thông tin video đã mã hóa có thể được xuất ra bằng giao diện xuất 22 trên vật ghi đọc được bằng máy tính 16.

Vật ghi đọc được bằng máy tính 16 có thể là phương tiện khả biến, như phương tiện phát rộng không dây hoặc phương tiện truyền qua mạng nối dây, hoặc phương tiện lưu trữ (tức là, phương tiện lưu trữ bất biến), như ổ đĩa cứng, ổ đĩa tác động nhanh, đĩa compac, đĩa video kỹ thuật số, đĩa Blu-ray, hoặc vật ghi đọc được bằng máy tính khác. 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 này cho thiết bị đích 14, ví dụ, thông qua tín hiệu truyền trên mạng. Tương tự, thiết bị tính toán của hãng sản xuất vật ghi, như hãng sản xuất đĩa, có thể thu dữ liệu video mã hóa từ thiết bị nguồn 12 và tạo ra đĩa chứa dữ liệu video mã hóa. Vì vậy, 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 có nhiều dạng khác nhau, trong nhiều ví dụ khác nhau.

Giao diện nhập 28 của thiết bị đích 14 thu thông tin từ vật ghi đọc được bằng máy tính. Thông tin từ vật ghi đọc được bằng máy tính 16 có thể bao gồm thông tin cú pháp được xác định bởi bộ mã hóa dữ liệu video 20, thông tin cú pháp này cũng được sử dụng bởi bộ giải mã dữ liệu video 30, chứa các phần tử cú pháp mô tả các đặc trưng và/hoặc cách xử lý các khối và các đơn vị mã hóa khác, ví dụ, các nhóm hình ảnh (Group of Picture - GOP). Thiết bị hiển thị 32 hiển thị dữ liệu video đã giải mã cho người dùng xem, và có thể bao gồm thiết bị hiển thị bất kỳ trong số nhiều thiết bị hiển thị như màn hình 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 dữ liệu video 20 và bộ giải mã dữ liệu 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 mô hình thử nghiệm HEVC (HEVC Test Model - HM). Theo cách khác, bộ mã hóa dữ liệu video 20 và bộ giải mã dữ liệu video 30 có thể hoạt động theo các chuẩn công nghiệp hoặc chuẩn riêng khác, như chuẩn ITU-T H.264, theo cách khác được gọi là MPEG-4, phần 10, mã hóa video nâng cao (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 chuẩn mã hóa video bao gồm MPEG-2 và ITU-T H.263. Mặc dù không được thể hiện trên Fig.1, theo một vài khía cạnh, mỗi bộ mã hóa dữ liệu video 20 và bộ giải mã dữ liệu video 30 có thể được tích hợp với bộ mã hóa và bộ giải mã âm thanh, và có thể bao gồm các đơn vị dồn kênh-phân kênh MUX-DEMUX thích hợp, hoặc phần cứng và phần mềm khác, để xử lý việc mã hóa cả âm thanh và video trong dòng dữ liệu chung hoặc các dòng dữ liệu riêng. Nếu áp dụng được, các đơn vị MUX-DEMUX có thể tuân theo giao thức dồn kênh ITU H.223, hoặc các giao thức khác như giao thức bó dữ liệu người dùng (user datagram protocol - UDP).

Chuẩn ITU-T H.264/MPEG-4 (AVC) được tạo ra bởi nhóm chuyên gia mã hóa dữ liệu video (Video Coding Experts Group - VCEG) của tổ chức ITU-T kết hợp với nhóm chuyên gia ảnh động (Moving Picture Experts Group - MPEG) của tổ chức ISO/IEC tạo thành một nhóm hợp tác chung gọi là nhóm hợp tác liên kết mã hóa dữ liệu video (Joint Video Team - JVT). Theo một số khía cạnh, các kỹ thuật nêu trong sáng chế có thể được áp dụng cho các thiết bị thường tuân theo chuẩn H.264. Chuẩn H.264 được mô tả trong Khuyến nghị ITU-T H.264, Mã hóa video nâng cao áp dụng cho các dịch vụ nghe nhìn chung, của nhóm nghiên cứu ITU-T, vào tháng 3 năm 2005, trong sáng chế, chuẩn này có thể được gọi là chuẩn H.264 hoặc tài liệu kỹ thuật H.264, hay chuẩn hoặc tài liệu kỹ thuật H.264/AVC. Nhóm hợp tác liên kết mã hóa dữ liệu video (JVT) tiếp tục làm việc trên các phiên bản mở rộng của chuẩn H.264/MPEG-4 AVC.

Mỗi bộ mã hóa dữ liệu video 20 và bộ giải mã dữ liệu video 30 có thể được thực hiện dưới dạng mạch mã hóa bất kỳ trong số rất nhiều mạch mã hóa phù 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

cửa lập trình được bằng trường (Field Programmable Gate Array - FPGA), mạch logic rời rạc, phần mềm, phần cứng, phần sụn hoặc mọi dạng kết hợp của các loại trên. Khi các kỹ thuật này được thực hiện một phần ở phần mềm, thì thiết bị có thể lưu trữ các lệnh cho phần mềm trong vật ghi đọc được bằng máy tính bất biến, thích hợp và thực thi 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 dữ liệu video 20 và bộ giải mã dữ liệu video 30 có thể nằm trong một hoặc nhiều bộ mã hóa hoặc bộ giải mã, mỗi bộ phận này có thể được tích hợp làm một phần của bộ mã hóa/giải mã kết hợp (CODEC) trong thiết bị tương ứng.

Nhóm hợp tác chung về mã hóa video (Joint Collaboration Team on Video Coding - JCT-VC) tiếp tục làm việc để phát triển chuẩn HEVC. Nỗ lực tiêu chuẩn hóa HEVC dựa trên mô hình cải tiến của thiết bị mã hóa dữ liệu video được gọi là mô hình thử nghiệm HEVC (HEVC Test Model - HM). Mô hình HM cho là các thiết bị mã hóa video có thêm vài khả năng so với các thiết bị hiện có, ví dụ, theo chuẩn ITU-T H.264/AVC. Ví dụ, trong khi chuẩn H.264 cung cấp chín chế độ mã hóa dự báo nội cấu trúc, thì mô hình HM có tới ba mươi ba chế độ mã hóa dự báo nội cấu trúc.

Nói chung, phiên bản dự thảo làm việc của mô hình HM mô tả khung video hoặc hình ảnh video có thể được chia ra thành một chuỗi gồm các khối cấu trúc cây hoặc các đơn vị mã hóa lớn nhất (Largest Coding Unit - LCU) chứa cả mẫu độ chói lẫn các mẫu màu. Dữ liệu cú pháp trong dòng bit có thể xác định kích thước của đơn vị LCU, đó là đơn vị mã hóa lớn nhất tính theo số lượng điểm ảnh. Một lát gồm nhiều khối cấu trúc cây liên tiếp theo thứ tự mã hóa. Khung video hoặc hình ảnh video có thể được phân tách thành một hoặc nhiều lát. Mỗi khối cấu trúc cây có thể được phân tách thành các đơn vị mã hóa (CU) theo cấu trúc cây tứ phân. Thông thường, cấu trúc dữ liệu cây tứ phân có một nút cho mỗi CU, với nút gốc tương ứng với khối cấu trúc cây. Nếu CU được phân tách thành bốn CU con, nút này tương ứng với CU gồm bốn nút lá, mỗi nút lá tương ứng với một trong số các CU con.

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 cấu trúc cây tứ phân có thể có cờ phân tách, chỉ báo việc CU tương ứng với nút đó có được phân 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 phân tách thành các CU con hay không. Nếu CU không được phân tách

nữa, thì CU đó được gọi là CU lá. Theo sáng chế, bốn CU con của một CU lá được gọi là CU lá ngay cả khi không có sự phân tách rõ ràng của CU lá ban đầu. Ví dụ, nếu CU có kích thước 16x16 điểm ảnh không được phân tách nữa, thì bốn CU con có kích thước 8x8 điểm ảnh cũng sẽ được gọi là các CU lá mặc dù CU có kích thước 16x16 điểm ảnh chưa từng được phân tách.

CU có mục đích tương tự như khối macro theo chuẩn H.264, ngoại trừ việc CU không có sự khác biệt về kích thước. Ví dụ, khối cấu trúc cây có thể được phân tách thành bốn nút con (còn được gọi là CU con), và mỗi nút con có thể lại là nút cha và được phân tách thành bốn nút con khác. Nút con cuối cùng không phân tách nữa, được gọi là nút lá của cấu trúc cây tứ phân, bao gồm nút mã hóa, còn được gọi là CU lá. Dữ liệu cú pháp liên quan đến dòng bit mã hóa có thể xác định số lần tối đa có thể phân tách khối cấu trúc cây, được gọi là độ sâu CU tối đa, và cũng có thể xác định kích thước tối thiểu của các nút mã hóa. Do 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 đơn vị bất kỳ trong số các CU, PU hoặc TU trong ngữ cảnh của chuẩn HEVC, hoặc các cấu trúc dữ liệu tương tự trong ngữ cảnh của các chuẩn khác (ví dụ, khối macro và các khối con của khối macro theo chuẩn H.264/AVC).

CU có một nút mã hóa và các đơn vị dự báo (Prediction Unit - PU) và các đơn vị biến đổi (Transform Unit-TU) liên quan đến nút mã hóa đó. Kích thước của CU tương ứng với kích thước của nút mã hóa và phải có dạng hình vuông. Kích thước của CU có thể nằm trong khoảng từ 8x8 điểm ảnh đến kích thước của khối cấu trúc cây có tối đa là 64x64 điểm ảnh hoặc lớn hơn. Mỗi CU có thể có một hoặc nhiều PU và một hoặc nhiều TU. Dữ liệu cú pháp liên quan đến CU có thể mô tả, ví dụ, việc phân tách CU thành một hoặc nhiều PU. Các chế độ phân tách có thể khác nhau tùy thuộc vào việc CU được mã hóa ở chế độ bỏ qua hay chế độ trực tiếp, được mã hóa ở chế độ dự báo nội cấu trúc hay được mã hóa ở chế độ dự báo liên cấu trúc. Các PU có thể được phân tách thành khối có dạng không phải hình vuông. Dữ liệu cú pháp liên quan đến CU cũng có thể mô tả, ví dụ, việc phân tách CU thành một hoặc nhiều TU theo cấu trúc cây tứ phân. TU có thể có dạng hình vuông hoặc dạng không phải hình vuông (ví dụ, dạng hình chữ nhật).

Chuẩn HEVC cho phép thực hiện các quy trình biến đổi theo TU, các quy trình biến đổi đó có thể khác nhau với các CU khác nhau. Các TU thường có kích

thước dựa vào kích thước của các PU trong một CU cho trước được xác định cho đơn vị LCU đã phân tách, tuy nhiên có thể không phải lúc nào cũng như vậy. Các TU thường có kích thước nhỏ hơn hoặc bằng kích thước của các PU. Trong một số ví dụ, các mẫu dữ liệu dư tương ứng với CU có thể được chia thành các đơn vị nhỏ hơn bằng cách sử dụng cấu trúc cây tứ phân gọi là “cấu trúc cây tứ phân dữ liệu dư” (Residual QuadTree - RQT). Các nút lá của cấu trúc RQT có thể được gọi là các đơn vị biến đổi (TU). Giá trị chênh lệch điểm ảnh liên quan đến các TU có thể được biến đổi để tạo ra các hệ số biến đổi, các hệ số biến đổi này có thể được lượng tử hóa.

CU lá có thể có một hoặc nhiều đơn vị dự báo (prediction unit - PU). Nói chung, PU biểu diễn vùng không gian tương ứng với toàn bộ hoặc một phần CU tương ứng, và có thể chứa dữ liệu để tìm mẫu tham chiếu cho PU. Ngoài ra, PU còn chứa dữ liệu liên quan đến việc dự báo. Ví dụ, khi PU được mã hóa dự báo nội cấu trúc, thì dữ liệu cho PU có thể được đưa vào trong cấu trúc cây tứ phân dữ liệu dư (RQT), dữ liệu này có thể là dữ liệu mô tả chế độ dự báo nội cấu trúc cho TU tương ứng với PU. Ví dụ khác, khi PU được mã hóa ở chế độ liên cấu trúc, thì PU có thể chứa dữ liệu xác định một hoặc nhiều vector chuyển động cho PU. Dữ liệu xác định vector chuyển động cho PU có thể mô tả, ví dụ, thành phần theo phương nằm ngang của vector chuyển động, thành phần theo phương thẳng đứng của vector chuyển động, độ phân giải cho vector chuyển động (ví dụ, độ chính xác bằng một phần tư điểm ảnh hoặc độ chính xác bằng một phần tám điểm ảnh), hình ảnh tham chiếu mà vector chuyển động chỉ đến đó, và/hoặc danh mục hình ảnh tham chiếu (ví dụ, danh mục 0, danh mục 1 hoặc danh mục C) cho vector chuyển động.

CU lá có một hoặc nhiều PU cũng có thể có một hoặc nhiều đơn vị biến đổi (transform unit-TU). Các đơn vị biến đổi có thể được xác định bằng cách sử dụng cấu trúc RQT (còn được gọi là cấu trúc cây tứ phân cho TU), như đã nêu trên. Ví dụ, cờ phân tách có thể chỉ báo việc CU lá có được phân tách thành bốn đơn vị biến đổi hay không. Sau đó, mỗi đơn vị biến đổi có thể được phân tách tiếp thành các TU con khác. Khi TU không được phân tách nữa, thì TU đó có thể được gọi là TU lá. Thông thường, đối với mã hóa nội cấu trúc, tất cả các TU lá thuộc về một CU lá đều dùng chung một chế độ dự báo nội cấu trúc. Có nghĩa là, cùng một chế độ dự báo nội cấu trúc thường được áp dụng để tính các giá trị dự báo cho tất cả các TU của CU lá. Đối với chế độ dự báo nội cấu trúc, bộ mã hóa dữ liệu video có thể tính toán giá trị dư cho

mỗi TU lá bằng cách sử dụng chế độ dự báo nội cấu trúc, dùng làm giá trị chênh lệch giữa phần CU tương ứng với TU và khối ban đầu. TU không nhất định chỉ giới hạn ở kích thước của PU. Do đó, các TU có thể lớn hơn hoặc nhỏ hơn so với PU. Đối với mã hóa nội cấu trúc, PU có thể có cùng vị trí với TU lá tương ứng cho cùng một CU đó. Trong một số ví dụ, kích thước lớn nhất của TU lá có thể tương ứng với kích thước của CU lá tương ứng.

Ngoài ra, các TU của các CU lá cũng có thể có liên quan đến các cấu trúc dữ liệu cây tứ phân tương ứng, được gọi là cấu trúc cây tứ phân dữ liệu dư (RQT). Có nghĩa là, CU lá có thể có cấu trúc cây tứ phân chỉ báo cách phân tách CU lá thành các TU. Nút gốc của cấu trúc cây tứ phân cho TU thường tương ứng với CU lá, trong khi đó nút gốc của cấu trúc cây tứ phân cho CU thường tương ứng với khối cấu trúc cây (hoặc LCU). Các TU trong cấu trúc RQT không được phân tách nữa thì được gọi là TU lá. Nói chung, sáng chế sử dụng các thuật ngữ CU và TU để chỉ lần lượt CU lá và TU lá, trừ trường hợp có quy định khác.

Chuỗi dữ liệu video thường là một chuỗi khung video hoặc hình ảnh video. Nhóm hình ảnh (group of picture - GOP) thường là một chuỗi gồm một hoặc nhiều hình ảnh video. Nhóm GOP có thể chứa dữ liệu cú pháp ở phần đầu của nhóm GOP, ở phần đầu của một hoặc nhiều hình ảnh, hoặc ở vị trí khác, để mô tả số lượng hình ảnh có trong nhóm GOP. Mỗi lát hình ảnh có thể chứa dữ liệu cú pháp lát để mô tả chế độ mã hóa cho lát tương ứng. Bộ mã hóa dữ liệu video 20 thường hoạt động trên các khối video trong các lát video riêng biệt để mã hóa dữ liệu video. Khối video có thể tương ứng với nút mã hóa trong CU. Các khối video có thể có kích thước cố định hoặc thay đổi, và có thể có kích thước khác nhau theo chuẩn mã hóa đã xác định.

Ví dụ, mô hình HM hỗ trợ dự báo cho nhiều kích thước PU. Giả sử kích thước của một CU cụ thể là $2N \times 2N$, thì mô hình HM hỗ trợ dự báo nội cấu trúc cho PU có kích thước $2N \times 2N$ hoặc $N \times N$, và dự báo liên cấu trúc cho PU đối xứng có kích thước $2N \times 2N$, $2N \times N$, $N \times 2N$ hoặc $N \times N$. Mô hình HM còn hỗ trợ phân tách không đối xứng để dự báo liên cấu trúc cho PU có kích thước $2N \times nU$, $2N \times nD$, $nL \times 2N$ và $nR \times 2N$. Khi phân tách không đối xứng, một chiều của CU không được phân tách, trong khi chiều còn lại được phân tách thành 25% và 75%. Phần CU tương ứng với phần 25% có ký hiệu “n” sau thông tin chỉ báo “trên”, “dưới”, “trái” hoặc “phải”. Do đó, ví dụ, “ $2N \times nU$ ” dùng để chỉ CU có kích thước $2N \times 2N$ được phân tách theo chiều ngang

thành PU có kích thước $2N \times 0,5N$ ở trên và PU có kích thước $2N \times 1,5N$ ở dưới.

Theo sáng chế, “ $N \times N$ ” và “ N nhân N ” có thể được sử dụng hoán đổi lẫn nhau để chỉ kích thước điểm ảnh của khối video tính theo chiều dọc và chiều ngang, ví dụ, 16×16 điểm ảnh hoặc 16 nhân 16 điểm ảnh. Thông thường, khối 16×16 sẽ có 16 điểm ảnh theo chiều dọc ($y = 16$) và 16 điểm ảnh theo chiều ngang ($x = 16$). Tương tự, khối $N \times N$ thường có N điểm ảnh theo chiều dọc và N điểm ảnh theo chiều ngang, trong đó N là giá trị nguyên không âm. Các điểm ảnh trong khối có thể được sắp xếp thành hàng và cột. Ngoài ra, các khối không nhất thiết phải có số lượng điểm ảnh theo chiều ngang bằng số lượng điểm ảnh theo chiều dọc. Ví dụ, các khối có thể có $N \times M$ điểm ảnh, trong đó M không nhất thiết phải bằng N .

Sau khi mã hóa dự báo nội cấu trúc hoặc mã hóa dự báo liên cấu trúc sử dụng các PU của CU, bộ mã hóa dữ liệu video 20 có thể tính dữ liệu dư cho các TU của CU. PU có thể chứa dữ liệu cú pháp mô tả phương pháp hoặc chế độ tạo ra dữ liệu điểm ảnh dự báo ở miền không gian (còn được gọi là miền điểm ảnh) và các TU có thể có các hệ số ở miền biến đổi sau khi áp dụng một quy trình biến đổi, ví dụ, biến đổi cosin rời rạc (Discrete Cosine Transform - DCT), biến đổi số nguyên, biến đổi dạng sóng, hoặc biến đổi tương tự về mặt khái niệm, đối với dữ liệu video dư. Dữ liệu dư có thể tương ứng với các giá trị 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 dữ liệu video 20 có thể tạo ra các TU chứa dữ liệu dư cho CU, và sau đó biến đổi các TU để tạo ra các hệ số biến đổi cho CU.

Sau khi biến đổi để tạo ra các hệ số biến đổi, bộ mã hóa dữ liệu video 20 có thể thực hiện thao tác lượng tử hóa cho các hệ số biến đổi. Lượng tử hóa thường là quy trình trong đó các hệ số biến đổi được lượng tử hóa để có thể giảm bớt lượng dữ liệu dùng để biểu diễn các hệ số, nhằm đạt được hiệu quả nén cao hơn. Quy trình lượng tử hóa có thể giảm độ sâu bit liên quan đến một số hoặc tất cả các hệ số. Ví dụ, giá trị n -bit có thể được làm tròn xuống thành giá trị m -bit trong quy trình lượng tử hóa, trong đó n lớn hơn m .

Sau khi lượng tử hóa, bộ mã hóa dữ liệu video có thể quét các hệ số biến đổi, tạo ra vectơ một chiều từ ma trận hai chiều chứa các hệ số biến đổi đã lượng tử hóa. Quy trình quét này được thiết kế để đặt các hệ số có năng lượng cao hơn (và do đó có tần số thấp hơn) ở phía trước của ma trận và đặt các hệ số có năng lượng thấp hơn (và

do đó có tần số cao hơn) ở phía sau của ma trận. Trong một số ví dụ, bộ mã hóa dữ liệu video 20 có thể áp dụng thứ tự quét định trước để quét các hệ số biến đổi đã lượng tử hóa tạo ra vector tuyến tính có thể được mã hóa entropy. Theo các phương án khác làm ví dụ, bộ mã hóa dữ liệu video 20 có thể thực hiện kỹ thuật quét thích ứng. Sau khi quét các hệ số biến đổi đã lượng tử hóa để tạo ra vector một chiều, bộ mã hóa dữ liệu video 20 có thể mã hóa entropy vector một chiều, ví dụ, theo phương pháp mã hóa độ dài thay đổi thích ứng với ngữ cảnh (Context-Adaptive Variable Length Coding - CAVLC), mã hóa số học nhị phân thích ứng với ngữ cảnh (Context-Adaptive Binary Arithmetic Coding - CABAC), mã hóa số học nhị phân thích ứng với ngữ cảnh theo 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 dữ liệu video 20 cũng có thể mã hóa entropy cho các phần tử cú pháp liên quan đến dữ liệu video mã hóa để cho bộ giải mã dữ liệu video 30 sử dụng khi giải mã dữ liệu video.

Để thực hiện phương pháp CABAC, bộ mã hóa dữ liệu video 20 có thể gán một ngữ cảnh trong mô hình ngữ cảnh cho ký hiệu cần truyền. Ngữ cảnh có thể liên quan đến việc, ví dụ, các giá trị lân cận của ký hiệu có khác không hay không. Để thực hiện phương pháp CAVLC, bộ mã hóa dữ liệu video 20 có thể chọn mã độ dài thay đổi cho ký hiệu cần truyền. Các từ mã trong phương pháp mã hóa độ dài thay đổi (Variable Length Coding - VLC) có thể được thiết lập sao cho các từ mã tương đối ngắn sẽ tương ứng với các ký hiệu có xác suất cao hơn, còn các từ mã dài sẽ tương ứng với các ký hiệu có xác suất thấp hơn. Theo cách này, việc áp dụng phương pháp VLC có thể tiết kiệm bit so với việc, ví dụ, sử dụng các từ mã có độ dài bằng nhau cho mỗi ký hiệu cần truyền. Việc xác định xác suất có thể dựa vào ngữ cảnh được gán cho ký hiệu.

Theo các kỹ thuật của sáng chế, bộ mã hóa dữ liệu video 20 và bộ giải mã dữ liệu video 30 có thể được tạo cấu hình để sử dụng các kỹ thuật nâng cao để báo hiệu dữ liệu để xác định các hình ảnh tham chiếu dài hạn, ví dụ, ở phần đầu lát, tập thông số (bao gồm tập thông số hình ảnh (picture parameter set - PPS) và tập thông số chuỗi (sequence parameter set - SPS)), hoặc tương tự. Nói chung, bộ mã hóa dữ liệu video 20 và bộ giải mã dữ liệu video 30 có thể được tạo cấu hình để thực hiện kỹ thuật bất

kỳ hoặc tất cả các kỹ thuật theo sáng chế, riêng lẻ hoặc kết hợp. Theo một vài ví dụ, bộ mã hóa dữ liệu video 20 và bộ giải mã dữ liệu video 30 có thể được tạo cấu hình để thực hiện các phương pháp báo hiệu các giá trị âm của chu kỳ MSB POC delta của các hình ảnh tham chiếu dài hạn. Theo một vài ví dụ, bộ mã hóa dữ liệu video 20 và bộ giải mã dữ liệu video 30 có thể được tạo cấu hình để thực hiện các phương pháp để báo hiệu hiệu quả hơn các giá trị chu kỳ MSB POC delta của các hình ảnh tham chiếu dài hạn với các giá trị LSB khác nhau bằng cách sử dụng mã hóa vi phân.

Theo một vài ví dụ, hình ảnh tham chiếu mà được báo hiệu dưới dạng các hình ảnh tham chiếu ngắn hạn cho hình ảnh hiện thời (hoặc được đưa vào trong tập hợp hình ảnh tham chiếu ngắn hạn của hình ảnh hiện thời) có thể được loại ra khỏi các hình ảnh tham chiếu được sử dụng để quyết định xem có truyền các giá trị chu kỳ MSB POC delta cho các hình ảnh tham chiếu dài hạn hay không, để tránh truyền các giá trị chu kỳ MSB POC delta không cần thiết. Theo một vài ví dụ, khi có n hình ảnh tham chiếu (bao gồm hoặc không bao gồm các hình ảnh tham chiếu ngắn hạn) trong DPB mà có LSB POC bằng $poc_lsb_lt[i]$, và khi n hình ảnh tham chiếu dài hạn được báo hiệu trong hình ảnh hiện thời mà có các giá trị LSB bằng $poc_lsb_lt[i]$, bộ mã hóa dữ liệu video 20 có thể chỉ truyền, và bộ giải mã dữ liệu video 30 có thể chỉ thu, các giá trị chu kỳ MSB POC delta cho $n - 1$ hình ảnh tham chiếu dài hạn, mặc dù có thể là dữ liệu này được truyền cho tất cả n hình ảnh tham chiếu dài hạn.

Bất kỳ hoặc tất cả các kỹ thuật khác nhau trên đây có thể được thực hiện theo phương thức thực hiện làm ví dụ dưới đây. Phần tử cú pháp trong WD7 mô tả chu kỳ MSB của hình ảnh tham chiếu dài hạn có thể được cải biến, như được thể hiện dưới đây. Sáng chế cũng mô tả phương pháp làm ví dụ để tạo ra dãy $\Delta poc_msb_cycle_lt[]$. Sáng chế còn mô tả quy trình giải mã làm ví dụ đối với tập hợp hình ảnh tham chiếu. Hơn nữa, sáng chế mô tả ngữ nghĩa của các phần tử cú pháp $\Delta poc_msb_present_flag[i]$ và $poc_lsb_lt[i]$. Cú pháp và ngữ nghĩa của các phần tử cú pháp khác của WD7 có thể vẫn không thay đổi.

Bảng 1 biểu diễn tập hợp cú pháp làm ví dụ đối với phần đầu lát. Phần tử cú pháp được gạch chân biểu diễn các thay đổi của cú pháp phần đầu lát so với WD7. Ngữ nghĩa của các phần tử cú pháp được thay đổi, và ngữ nghĩa được thay đổi của các phần tử cú pháp khác, được nêu dưới đây.

Bảng 1

slice_header() {	Bộ mô tả
...	
short_term_ref_pic_set_idx	u(v)
if(long_term_ref_pics_present_flag) {	
num_long_term_pics	ue(v)
for(i = 0; i < num_long_term_pics; i++) {	
poc_lsb_lt[i]	u(v)
delta_poc_msb_present_flag[i]	u(1)
if(delta_poc_msb_present_flag[i])	
<u>delta_poc_msb_cycle_lt_plus1[i]</u>	<u>ue(v)</u>
used_by_curr_pic_lt_flag[i]	u(1)
}	
}	
...	
byte_alignment()	
}	

Theo ví dụ trong bảng 1, cú pháp phần đầu lát bao gồm phần tử cú pháp bổ sung `delta_poc_msb_cycle_lt_plus1[i]`. Nghĩa của phần tử cú pháp này được mô tả dưới đây. Theo một vài ví dụ, `poc_lsb_lt[ i ]` có thể chỉ rõ giá trị của các bit có giá trị 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 $\log_2 \text{max_pic_order_cnt_lsb_minus4} + 4$ bit.

Theo một vài ví dụ, `delta_poc_msb_present_flag[ i ]` bằng 1 có thể chỉ rõ rằng `delta_poc_msb_cycle_lt_plus1[ i ]` là có mặt. `delta_poc_msb_present_flag[ i ]` bằng 0 có thể chỉ rõ rằng `delta_poc_msb_cycle_lt_plus1[ i ]` không có mặt. `delta_poc_msb_present_flag[ i ]` có thể được đặt bằng 1 khi có j hình ảnh tham chiếu, trong đó j là một giá trị số nguyên lớn hơn 1, trong bộ nhớ đệm hình ảnh đã giải mã, ngoại trừ các hình ảnh được đánh dấu là “được sử dụng làm tham chiếu ngắn hạn” bởi hình ảnh hiện thời, với các bit có giá trị nhỏ nhất của giá trị số đếm thứ tự hình ảnh bằng `poc_lsb_lt[ i ]`, trong khi hình ảnh tham chiếu dài hạn thứ i không phải là

hình ảnh tham chiếu dài hạn thứ j với bit có giá trị nhỏ nhất của giá trị số đếm thứ tự hình ảnh bằng $\text{poc_lsb_lt}[i]$ mà được báo hiệu bởi hình ảnh hiện thời.

Theo một vài ví dụ, $\text{delta_poc_msb_cycle_lt_plus1}[i]$ có thể được sử dụng để xác định giá trị của các bit có giá trị 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. Khi không được báo hiệu, $\text{delta_poc_msb_cycle_lt_plus1}[i]$ có thể lấy giá trị 1.

Biến $\text{DeltaPocMSBCycleLt}[i]$ có thể được tạo ra theo thuật toán sau, trong đó công thức (7-37) được dùng để chỉ một phần của HEVC, tức là, phần được thay đổi so với WD7. Lưu ý rằng giá trị thu được từ công thức làm ví dụ dưới đây có nghĩa là các chu kỳ MSB được truyền theo thứ tự tăng dần giá trị của chúng.

$$\begin{aligned} \text{DeltaPocMSBCycleLt}[i] &= \text{delta_poc_msb_present_flag}[i] * \\ &\quad (\text{delta_poc_msb_cycle_lt_plus1}[i] - 1); \\ \text{if}(i \neq 0) \end{aligned} \tag{7-37}$$

$$\text{DeltaPocMSBCycleLt}[i] += \text{DeltaPocMSBCycleLt}[i - 1];$$

Giá trị của $\text{DeltaPocMSBCycleLt}[i]$ * $\text{MaxPicOrderCntLsb} + \text{pic_order_cnt_lsb} - \text{poc_lsb_lt}[i]$ có thể nằm trong khoảng bao hàm từ 1 đến $2^{24} - 1$. Theo cách khác, chu kỳ MSB có thể được truyền bằng cách báo hiệu $\text{delta_poc_msb_cycle_lt_plusN}[i]$ dưới dạng $\text{ue}(v)$ (tức là, dưới dạng phân tử cú pháp được mã hóa Exp-Golomb nguyên không dấu với bit bên trái trước), thay vì $\text{delta_poc_msb_cycle_lt_plus1}[i]$, trong đó N là giá trị lớn hơn 1. Theo cách khác, chu kỳ MSB có thể được truyền bằng cách báo hiệu $\text{delta_poc_msb_cycle_lt}[i]$ dưới dạng $\text{se}(v)$ (tức là, dưới dạng phân tử cú pháp được mã hóa Exp-Golomb nguyên có dấu với bit bên trái trước), thay vì $\text{delta_poc_msb_cycle_lt_plus1}[i]$.

Nói cách khác, theo các kỹ thuật của sáng chế, giá trị biểu diễn chu kỳ MSB của hình ảnh i (ví dụ, $\text{delta_poc_msb_cycle_lt_plus1}[i]$, mà thực tế được báo hiệu trong dòng bit) có thể được tính toán so với chu kỳ MSB của hình ảnh tham chiếu dài hạn được mã hóa trước, ví dụ, hình ảnh $i-1$, hơn là so với giá trị MSB cơ sở. Giá trị này là tương đương về mặt toán học để tính toán giá trị biểu diễn chu kỳ MSB của hình ảnh i dưới dạng độ chênh lệch giữa các MSB của giá trị POC của hình ảnh i và các MSB của giá trị POC của hình ảnh $i-1$, giả sử cuối cùng giá trị này được bổ sung vào giá trị chu kỳ MSB của hình ảnh $i-1$ để biểu diễn giá trị chu kỳ MSB thực của

hình ảnh i (mà là độ chênh lệch giữa MSB của hình ảnh i và MSB cơ sở). Điều này có thể giúp tiết kiệm bit trong dòng bit, bởi vì các giá trị biểu diễn độ chênh lệch giữa các giá trị chu kỳ MSB liên tiếp có thể nhỏ hơn độ chênh lệch giữa các MSB của các hình ảnh tham chiếu dài hạn và MSB cơ sở.

Theo một vài ví dụ, bộ mã hóa dữ liệu video 20 và bộ giải mã dữ liệu video 30 có thể thực hiện quy trình tạo ra tập hợp hình ảnh tham chiếu và đánh dấu hình ảnh theo các bước theo thứ tự sau, trong đó DPB được dùng để chỉ bộ nhớ đệm hình ảnh đã giải mã như được mô tả trong phụ lục C của WD7. Lưu ý rằng thứ tự của hình ảnh tham chiếu ngắn hạn và tập con hình ảnh tham chiếu dài hạn thu được trong ví dụ dưới đây được đảo ngược so với thứ tự trong WD7. Các việ dẫn trong dấu ngoặc đơn gần lề bên phải được dự định là dùng để chỉ các mục của WD7. Các thay đổi này và thay đổi khác so với thuật toán nêu trong WD7 được lưu ý bằng cách sử dụng chữ viết được gạch chân trong ví dụ dưới đây:

1. Sau đây áp dụng:

```

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"

```

2. Tất cả các hình ảnh tham chiếu được đưa vào trong RefPicSetStCurrBefore, RefPicSetStCurrAfter và RefPicSetStFoll được đánh dấu là “được sử dụng làm tham chiếu ngắn hạn”.

3. Sau đây áp dụng:

```

for( i = 0; i < NumPocLtCurr; i++ ) {
    if( ! delta_poc_msb_present_flag[ i ] ) {
        if( there is a long-term reference picture picX in the DPB
            with pic_order_cnt_lsb equal to PocLtCurr[ i ] and is not
            included in any of RefPicSetStCurrBefore,
            RefPicSetStCurrAfter and RefPicSetStFoll)
            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 ] and is not
            included in any of RefPicSetStCurrBefore,
            RefPicSetStCurrAfter and RefPicSetStFoll)
            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 ] and is not
            included in any of RefPicSetStCurrBefore,
            RefPicSetStCurrAfter and RefPicSetStFoll)
            RefPicSetLtCurr[ i ] = picX
        else if( there is a short-term reference picture picY in the DPB
            with PicOrderCntVal equal to PocLtCurr[ i ] and is not
            included in any of RefPicSetStCurrBefore,
            RefPicSetStCurrAfter and RefPicSetStFoll)
            RefPicSetLtCurr[ i ] = picY
        else
            RefPicSetLtCurr[ i ] = “no reference picture”
    }
}

```

```

    }
}
(8-6)
for( i = 0; i < NumPocLtFoll; i++ ) {
    if( !delta_poc_msb_present_flag[ i ] ) {
        if( there is a long-term reference picture picX in the DPB
            with pic_order_cnt_lsb equal to PocLtFoll[ i ] and is not
            included in any of RefPicSetStCurrBefore,
            RefPicSetStCurrAfter and RefPicSetStFoll)
            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 ] and is not
            included in any of RefPicSetStCurrBefore,
            RefPicSetStCurrAfter and RefPicSetStFoll)
            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 ] and is not included
            in any of RefPicSetStCurrBefore, RefPicSetStCurrAfter
            and RefPicSetStFoll)
            RefPicSetLtFoll[ i ] = picX
        else if( there is a short-term reference picture picY in the DPB
            with PicOrderCntVal equal to PocLtFoll[ i ]
            and is not included in any of RefPicSetStCurrBefore,
            RefPicSetStCurrAfter and RefPicSetStFoll)
            RefPicSetLtFoll[ i ] = picY
        else
            RefPicSetLtFoll[ i ] = “no reference picture”
    }
}

```

4. Tất cả các hình ảnh tham chiếu được đưa vào trong RefPicSetLtCurr

và RefPicSetLtFoll được đánh dấu là “được sử dụng làm tham chiếu dài hạn”

5. Tất cả các hình ảnh tham chiếu trong bộ nhớ đệm hình ảnh đã giải mã mà không được đưa vào trong RefPicSetLtCurr, RefPicSetLtFoll, RefPicSetStCurrBefore, RefPicSetStCurrAfter hoặc RefPicSetStFoll được đánh dấu là “không được sử dụng làm tham chiếu”.

Có thể có một hoặc nhiều hình ảnh tham chiếu được đưa vào trong tập hợp hình ảnh tham chiếu nhưng không có mặt trong bộ nhớ đệm hình ảnh đã giải mã. Các mục nhập trong RefPicSetStFoll hoặc RefPicSetLtFoll mà tương đương với “không phải hình ảnh tham chiếu” nên được bỏ qua. Trừ khi một trong hai điều kiện sau là đúng, việc mất ảnh không chủ ý nên được suy ra cho mỗi mục nhập trong RefPicSetStCurrBefore, RefPicSetStCurrAfter và RefPicSetLtCurr mà tương đương với “không phải hình ảnh tham chiếu”: a) hình ảnh mã hóa đầu tiên trong dòng bit là hình ảnh CRA và hình ảnh mã hóa hiện thời là hình ảnh TFD liên quan đến hình ảnh mã hóa đầu tiên trong dòng bit; và b) hình ảnh RAP trước đó đứng trước hình ảnh mã hóa hiện thời theo thứ tự giải mã là hình ảnh BLA và hình ảnh mã hóa hiện thời là hình ảnh TFD liên quan đến hình ảnh BLA.

Do đó, theo các kỹ thuật của sáng chế, bộ mã hóa dữ liệu video (ví dụ, bộ mã hóa dữ liệu video 20 và/hoặc bộ giải mã dữ liệu video 30) có thể tính toán các giá trị chu kỳ MSB cho các hình ảnh có các giá trị LSB khác nhau sử dụng mã hóa vi phân. Ngoài ra, bộ mã hóa dữ liệu video này có thể mã hóa hình ảnh hiện thời dựa vào một trong số các hình ảnh tham chiếu dài hạn, ví dụ, hình ảnh $i-1$ hoặc hình ảnh i , bằng cách sử dụng các giá trị biểu diễn các chu kỳ MSB. Ví dụ, bộ mã hóa dữ liệu video này có thể xác định các giá trị POC cho các hình ảnh tham chiếu bằng cách sử dụng công thức được mô tả trên đây (công thức 7-37) và/hoặc thuật toán được mô tả trên đây.

Bộ mã hóa dữ liệu video 20 và bộ giải mã dữ liệu video 30, do đó, biểu diễn các ví dụ về bộ mã hóa dữ liệu video được tạo cấu hình để mã hóa giá trị thứ nhất biểu diễn độ chênh lệch giữa giá trị các bit có giá trị lớn nhất (MSBs) cơ sở của giá trị số đếm thứ tự hình ảnh (POC) của hình ảnh hiện thời của dữ liệu video và giá trị MSBs thứ nhất của giá trị POC thứ nhất của hình ảnh tham chiếu dài hạn thứ nhất của dữ liệu video, mã hóa giá trị thứ hai biểu diễn độ chênh lệch giữa giá trị MSBs thứ

hai của giá trị POC thứ hai của hình ảnh tham chiếu dài hạn thứ hai của dữ liệu video và giá trị MSBs thứ nhất, trong đó giá trị POC thứ nhất và giá trị POC thứ hai có các giá trị bit có giá trị nhỏ nhất (LSB) khác nhau, và mã hóa ít nhất một phần hình ảnh hiện thời của dữ liệu video dựa vào ít nhất một trong số hình ảnh tham chiếu dài hạn thứ nhất sử dụng giá trị thứ nhất và hình ảnh tham chiếu dài hạn thứ hai sử dụng giá trị thứ nhất và giá trị thứ hai.

Bộ mã hóa dữ liệu video 20, ví dụ, có thể xác định xem có hai hoặc nhiều hình ảnh tham chiếu dài hạn, cho hình ảnh hiện thời, mà các giá trị chu kỳ MSB (còn được gọi là các giá trị MSBs) được báo hiệu hay không, thậm chí khi các hình ảnh tham chiếu dài hạn có các giá trị LSB khác nhau. Trong trường hợp có hai (hoặc nhiều) hình ảnh tham chiếu dài hạn có các giá trị LSB khác nhau, mà các giá trị chu kỳ MSB được báo hiệu, bộ mã hóa dữ liệu video 20 có thể mã hóa dữ liệu biểu diễn giá trị chu kỳ MSB cho hình ảnh tham chiếu dài hạn thứ hai dưới dạng độ chênh lệch (hoặc delta) so với giá trị chu kỳ MSB cho hình ảnh tham chiếu dài hạn thứ nhất. Tương tự, bộ mã hóa dữ liệu video 20 có thể mã hóa dữ liệu biểu diễn giá trị chu kỳ MSB cho hình ảnh tham chiếu dài hạn thứ nhất dưới dạng độ chênh lệch so với giá trị MSB cơ sở của giá trị POC cho hình ảnh hiện thời. Bộ mã hóa dữ liệu video 20 có thể mã hóa tiếp ít nhất một phần của hình ảnh hiện thời (ví dụ, một hoặc nhiều khối) dựa vào ít nhất một trong số hình ảnh tham chiếu dài hạn thứ nhất và hình ảnh tham chiếu dài hạn thứ hai.

Do đó, khi bộ giải mã dữ liệu video 30 thu dữ liệu chỉ báo các giá trị chu kỳ MSB cho hình ảnh tham chiếu dài hạn, bộ giải mã dữ liệu video 30 có thể giải mã giá trị biểu diễn độ chênh lệch giữa các MSB cơ sở của giá trị POC cho hình ảnh tham chiếu dài hạn thứ nhất, và tính toán các MSB của giá trị POC cho hình ảnh tham chiếu dài hạn thứ nhất bằng cách bổ sung giá trị vào các MSB cơ sở. Bộ giải mã dữ liệu video 30 có thể cũng giải mã giá trị biểu diễn độ chênh lệch giữa các MSB của giá trị POC cho hình ảnh tham chiếu dài hạn thứ hai và các MSB của giá trị POC cho hình ảnh tham chiếu dài hạn thứ nhất, và tính toán các MSB của giá trị POC cho hình ảnh tham chiếu dài hạn thứ hai bằng cách bổ sung giá trị này vào các MSB của giá trị POC cho hình ảnh tham chiếu dài hạn thứ nhất. Tương tự như vậy, bộ giải mã dữ liệu video 30 có thể giải mã ít nhất một phần của hình ảnh hiện thời ít nhất một trong số hình ảnh tham chiếu dài hạn thứ nhất và hình ảnh tham chiếu dài hạn thứ hai, bằng

cách sử dụng các MSB của các giá trị POC cho các hình ảnh tham chiếu dài hạn.

Mỗi bộ mã hóa dữ liệu video 20 và bộ giải mã dữ liệu video 30 có thể được thực hiện dưới dạng mạch mã hóa bất kỳ trong số rất nhiều mạch mã hóa hoặc giải mã phù hợp, nếu áp dụng được, 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 cửa lập trình được bằng trường (field programmable gate array - FPGA), mạch logic rời rạc, phần mềm, phần cứng, phần sụn hoặc mọi dạng kết hợp bất kỳ của các loại trên. Mỗi bộ mã hóa dữ liệu video 20 và bộ giải mã dữ liệu video 30 có thể được nằm trong một hoặc nhiều bộ mã hóa hoặc bộ giải mã, mỗi bộ phận này có thể được tích hợp làm một phần của bộ mã hóa/bộ giải mã dữ liệu video kết hợp (CODEC). Thiết bị bao gồm bộ mã hóa dữ liệu video 20 và/hoặc bộ giải mã dữ liệu video 30 có thể chứa 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 dữ liệu video 20 có thể thực hiện các kỹ thuật để báo hiệu tốt hơn các hình ảnh tham chiếu dài hạn ở phần đầu lát. Bộ mã hóa dữ liệu video 20 có thể thực hiện mã hóa nội cấu trúc và liên cấu trúc các khối video trong các lát video. Quy trình mã hóa nội cấu trúc dựa vào kỹ thuật dự báo không gian để giảm hoặc loại bỏ sự dư thừa không gian trong video trong khung hoặc hình ảnh video cho trước. Quy trình mã hóa liên cấu trúc dựa vào kỹ thuật dự báo thời gian để giảm hoặc loại bỏ sự dư thừa thời gian trong video trong các khung hoặc các hình ảnh kế nhau của chuỗi dữ liệu video. Chế độ nội cấu trúc (chế độ I) có thể chỉ chế độ bất kỳ trong số một vài chế độ mã hóa dựa trên 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 chế độ bất kỳ trong số một vài chế độ mã hóa dựa trên thời gian.

Như được thể hiện trên Fig.2, bộ mã hóa dữ liệu video 20 thu khối video hiện thời trong khung video được mã hóa. Trong ví dụ trên Fig.2, bộ mã hóa dữ liệu video 20 bao gồm bộ chọn chế độ 40, bộ nhớ hình ảnh tham chiếu 64, bộ cộng 50, bộ xử lý biến đổi 52, bộ lượng tử hóa 54, và bộ mã hóa entropy 56. Bộ chọn chế độ 40, lần lượt, bao gồm bộ bù chuyển động 44, bộ ước tính chuyển động 42, bộ dự báo nội cấu trúc 46, và bộ phân tách 48. Để khôi phục khối video, bộ mã hóa dữ liệu video 20 cũng bao gồm bộ lượng tử hóa ngược 58, bộ biến đổi ngược 60, và bộ cộng 62. Bộ lọc tách khối (không được thể hiện trên Fig.2) có thể cũng được đưa vào để lọc các

biên khối để loại bỏ các thành phần lại dạng khối ra khỏi dữ liệu 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 lặp hoặc sau vòng lặp) có thể cũng được sử dụng ngoài bộ lọc tách khối. Các bộ lọc này không được thể hiện để cho ngắn gọn, nhưng nếu muốn, có thể lọc đầu ra của bộ cộng 50 (như là bộ lọc trong vòng lặp).

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

Hơn nữa, bộ phân tách 48 có thể phân tách các khối dữ liệu video thành các khối con, dựa trên sự đánh giá các sơ đồ phân tách trước đó trong các lượt mã hóa trước đó. Ví dụ, bộ phân tách 48 có thể trước hết phân tách khung hoặc lát thành các LCU, và phân tách mỗi LCU này thành các CU con dựa trên phân tích tốc độ-méo (ví dụ, tối ưu hóa tốc độ-méo). Hơn nữa, bộ chọn chế độ 40 có thể tạo ra cấu trúc dữ liệu cây tứ phân chỉ báo sự phân tách 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 chọn chế độ 40 có thể chọn một trong số các chế độ mã hóa, liên cấu trúc hoặc nội cấu trúc, ví dụ, dựa trên các kết quả sai số, và cung cấp khối được mã hóa liên cấu trúc hoặc nội cấu trúc thu được cho bộ cộng 50 để tạo ra dữ liệu khối dư và cho bộ cộng 62 để khôi phục khối mã hóa này để sử dụng làm khung tham chiếu. Bộ chọn chế độ 40 cũng cung cấp các phần tử cú pháp, như các vector chuyển động, các bộ chỉ báo chế độ nội cấu trúc, thông tin phân tách, và thông tin cú pháp khác, cho bộ mã hóa entropy 56. Theo các kỹ thuật của sáng chế, bộ chọn chế độ 40 có thể cung cấp thông tin xác định hình ảnh tham chiếu, như việc hình ảnh tham chiếu nên được coi là hình ảnh tham chiếu dài hạn hay hình ảnh tham chiếu ngắn hạn, cũng như các bit có giá trị nhỏ nhất (LSB) của các giá trị số đếm thứ tự hình ảnh (POC) cho hình ảnh tham chiếu dài hạn, cho bộ mã hóa entropy 56. Bộ mã hóa entropy 56, hoặc đơn

vị khác của bộ mã hóa dữ liệu video 20, có thể bao gồm thông tin xác định hình ảnh tham chiếu này trong phần đầu lát của lát, theo các kỹ thuật của sáng chế. Theo cách khác, bộ mã hóa entropy 56 có thể bao gồm thông tin xác định hình ảnh tham chiếu (ví dụ, để chỉ báo hình ảnh tham chiếu nào là hình ảnh tham chiếu dài hạn) trong tập thông số, như tập thông số hình ảnh (picture parameter set -PPS) hoặc tập thông số chuỗi (sequence parameter set - SPS).

Như được giải thích trên đây, hình ảnh tham chiếu dài hạn có thể được báo hiệu bằng cách sử dụng các LSB của các giá trị POC cho hình ảnh tham chiếu dài hạn. Ví dụ, bộ mã hóa entropy 56 có thể chỉ báo rằng hình ảnh tham chiếu là hình ảnh tham chiếu dài hạn bằng cách báo hiệu các LSB của giá trị POC cho hình ảnh tham chiếu này trong phần đầu lát, PPS, hoặc SPS. Các hình ảnh tham chiếu khác không được báo hiệu theo phương thức này có thể được coi là các hình ảnh tham chiếu ngắn hạn.

Bằng cách báo hiệu thông tin xác định cho hình ảnh tham chiếu dài hạn, bộ mã hóa entropy 56 cung cấp dữ liệu chỉ báo hình ảnh tham chiếu nào cần được lưu trữ trong bộ nhớ đệm hình ảnh đã giải mã (decoded picture buffer - DPB) của bộ giải mã dữ liệu video (ví dụ, bộ giải mã dữ liệu video 30) trong khoảng thời gian tương đối dài hạn, và hình ảnh tham chiếu nào có thể được loại bỏ khỏi bộ nhớ đệm hình ảnh đã giải mã trước đó. Nhìn chung, hình ảnh tham chiếu dài hạn có thể được sử dụng là hình ảnh tham chiếu cho các hình ảnh mã hóa có thời gian thứ tự giải mã tương đối muộn. Do đó, việc báo hiệu thông tin xác định cho hình ảnh tham chiếu dài hạn này cung cấp thông tin cho bộ giải mã dữ liệu video, ví dụ, bộ giải mã dữ liệu video 30, sao cho bộ giải mã dữ liệu video có thể xác định xem hình ảnh tham chiếu có thể được xóa bỏ khỏi DPB của chính nó hay không.

Bộ ước tính chuyển động 42 và bộ bù chuyển động 44 có thể được tích hợp cao, nhưng được minh họa tách biệt cho dễ phục vụ mục đích giải thích khái niệm. Quy trình ước tính chuyển động, được thực hiện bằng bộ ước tính chuyển động 42, là quy trình tạo ra các vector chuyển động, để ước tính chuyển động cho các khối video. Vector chuyển động, ví dụ, có thể chỉ báo sự dịch chuyển PU của khối video trong khung hoặc hình ảnh video hiện thời so với khối dự báo trong khung tham chiếu (hoặc đơn vị mã hóa khác) so với khối hiện thời được mã hóa trong khung hiện thời (hoặc đơn vị mã hóa khác). Khối dự báo là khối được nhận thấy là phù hợp nhất với

khối được mã hóa, xét về độ chênh lệch điểm ảnh, mà có thể được xác định bằng tổng hiệu số tuyệt đối (sum of absolute difference - SAD), tổng hiệu số bình phương (sum of square difference - SSD), hoặc các số đo độ chênh lệch khác. Trong một số ví dụ, bộ mã hóa dữ liệu video 20 có thể tính toán các giá trị cho các vị trí điểm ảnh dưới số nguyên của các hình ảnh tham chiếu được lưu trữ trong bộ nhớ hình ảnh tham chiếu 64. Ví dụ, bộ mã hóa dữ liệu video 20 có thể nội suy giá trị của các vị trí điểm ảnh một phần tư, các vị trí điểm ảnh một phần tám, các vị trí điểm ảnh phân số khác của hình ảnh tham chiếu này. Do đó, bộ ước tính chuyển động 42 có thể thực hiện tìm kiếm chuyển động so với các vị trí điểm ảnh đầy đủ và các vị trí điểm ảnh phân số và xuất ra vector chuyển động có độ chính xác điểm ảnh phân số.

Bộ ước tính chuyển động 42 tính toán vector chuyển động cho PU của khối video trong lát được mã hóa liên 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 này có thể được chọn từ danh mục hình ảnh tham chiếu thứ nhất (Danh mục 0) hoặc danh mục hình ảnh tham chiếu thứ hai (Danh mục 1), mỗi danh mục này xác định một hoặc nhiều hình ảnh tham chiếu được lưu trữ trong bộ nhớ hình ảnh tham chiếu 64. Bộ nhớ hình ảnh tham chiếu 64 có thể lưu trữ cả hình ảnh tham chiếu ngắn hạn và dài hạn. Bộ ước tính chuyển động 42 truyền vector chuyển động được tính toán này cho bộ mã hóa entropy 56 và bộ bù chuyển động 44.

Quy trình bù chuyển động, được thực hiện bởi bộ bù chuyển động 44, có thể bao gồm tìm nạp hoặc tạo ra khối dự báo dựa trên vector chuyển động này được xác định bằng bộ ước tính chuyển động 42. Hơn nữa, bộ ước tính chuyển động 42 và bộ bù chuyển động 44 có thể được tích hợp về mặt chức năng, trong một số ví dụ. Khi thu được vector chuyển động cho PU của khối video hiện thời, bộ bù chuyển động 44 có thể định vị khối dự báo mà vector chuyển động chỉ đến trong một trong số các danh mục hình ảnh tham chiếu. Bộ cộng 50 tạo ra khối video dự bằng cách trừ đi các giá trị điểm ảnh của khối dự báo từ các giá trị điểm ảnh của khối video hiện thời đang được mã hóa, tạo ra các giá trị chênh lệch điểm ảnh, như được thảo luận dưới đây. Nhìn chung, bộ ước tính chuyển động 42 thực hiện ước tính chuyển động so với các thành phần độ chói, và bộ bù chuyển động 44 sử dụng các vector chuyển động được tính toán dựa trên các thành phần độ chói này cho cả thành phần màu và thành phần độ chói.

Bộ chọn chế độ 40 cũng có thể tạo ra các phần tử cú pháp liên quan đến các khối video và lát video để sử dụng bởi bộ giải mã dữ liệu video 30 để giải mã các khối video của lát video. Ví dụ, sau khi tính toán vector chuyển động cho khối dữ liệu video (trong đó vector chuyển động chỉ đến khối tham chiếu trong hình ảnh tham chiếu), bộ chọn chế độ 40 có thể tạo ra các giá trị cho các thông số chuyển động, như danh mục hình ảnh tham chiếu trong đó hình ảnh tham chiếu được lưu trữ và chỉ mục trong danh mục hình ảnh tham chiếu tương ứng với vị trí của hình ảnh tham chiếu trong danh mục hình ảnh tham chiếu. Các thông số chuyển động này và các thông số chuyển động khác có thể được mã hóa dựa trên chế độ mã hóa vector chuyển động. Ví dụ, bằng cách sử dụng kỹ thuật dự báo vector chuyển động nâng cao (advanced motion vector prediction -AMVP), bộ mã hóa entropy 56 có thể xác định bộ dự báo vector chuyển động (tương ứng với khối lân cận theo thời gian hoặc không gian với khối hiện thời), và báo hiệu rõ ràng danh mục hình ảnh tham chiếu, chỉ mục hình ảnh tham chiếu, và các giá trị chênh lệch vector chuyển động ngang và dọc. Theo ví dụ khác, bằng cách sử dụng chế độ hợp nhất, bộ mã hóa entropy 56 có thể xác định bộ dự báo vector chuyển động sử dụng chỉ mục hợp nhất, và danh mục hình ảnh tham chiếu và chỉ mục tham chiếu có thể được kế thừa từ bộ dự báo vector chuyển động (ví dụ, thông tin chuyển động của khối lân cận, dù là khối lân cận theo không gian hay khối lân cận theo thời gian).

Bộ mã hóa dữ liệu video 20 báo hiệu thông tin chỉ báo hình ảnh tham chiếu nào là hình ảnh tham chiếu dài hạn vừa để chỉ báo rằng các hình ảnh này không được loại bỏ khỏi DPB vừa để cung cấp thông tin có thể được sử dụng trong quá trình thiết lập danh mục hình ảnh tham chiếu. Theo phương thức này, bộ giải mã dữ liệu video, như bộ giải mã dữ liệu video 30, có thể có khả năng tái tạo chính xác danh mục hình ảnh tham chiếu, sao cho chỉ mục tham chiếu này được sử dụng làm thông tin chuyển động cho khối được dự báo liên cấu trúc đề cập đến hình ảnh tham chiếu đúng.

Như được lưu ý trên đây, hình ảnh tham chiếu dài hạn có thể được báo hiệu bằng cách sử dụng các giá trị POC của các hình ảnh tham chiếu dài hạn. Các giá trị POC thường tương ứng với dữ liệu chỉ báo thứ tự hiển thị (còn được gọi là thứ tự xuất) hình ảnh, mà không cần thiết là giống như thứ tự giải mã (còn được gọi là thứ tự dòng bit). Các giá trị POC cho hình ảnh tham chiếu dài hạn có thể được báo hiệu bằng cách chỉ sử dụng các bit có giá trị nhỏ nhất (LSB). Dữ liệu biểu diễn các MSB

có thể đôi khi được báo hiệu, ví dụ, khi hai hoặc nhiều hình ảnh tham chiếu dài hạn có các LSB giống nhau. Trong một số trường hợp, bộ mã hóa dữ liệu video 20 có thể xác định rằng dữ liệu biểu diễn các MSB cho hai (hoặc nhiều) ảnh cần được báo hiệu, thậm chí khi ít nhất hai trong số các hình ảnh mà dữ liệu MSBs được báo hiệu có các LSB khác nhau.

Theo các kỹ thuật của sáng chế, bộ mã hóa dữ liệu video 20 có thể báo hiệu dữ liệu biểu diễn các MSB dưới dạng độ chênh lệch, ví dụ, độ chênh lệch giữa một tập hợp MSBs này và tập hợp MSBs khác. Ví dụ, như được thảo luận trên đây liên quan đến Bảng 1, bộ mã hóa dữ liệu video 20 có thể tính toán độ chênh lệch giữa các giá trị MSB cơ sở và giá trị MSB cho hình ảnh tham chiếu dài hạn. Bộ mã hóa entropy 56 có thể báo hiệu độ chênh lệch này là giá trị biểu diễn các MSB cho hình ảnh tham chiếu dài hạn. Các giá trị MSB cơ sở có thể tương ứng với các MSB của giá trị POC của hình ảnh hiện thời. Theo cách khác, các giá trị MSB cơ sở có thể tương ứng với các MSB của giá trị POC của hình ảnh khác, như hình ảnh làm mới giải mã tức thì (instantaneous decoder refresh - IDR).

Sau khi mã hóa giá trị chênh lệch cho một hình ảnh tham chiếu dài hạn, bộ mã hóa entropy 56 có thể tính toán độ chênh lệch giữa các MSB của giá trị POC cho hình ảnh tham chiếu dài hạn khác và các MSB của giá trị POC cho hình ảnh tham chiếu dài hạn trước đó. Tương tự, bộ mã hóa entropy 56 có thể báo hiệu giá trị chênh lệch thứ hai này cho hình ảnh tham chiếu dài hạn sau đó, sao cho bộ giải mã dữ liệu video, như bộ giải mã dữ liệu video 30, có thể khôi phục các MSB của giá trị POC cho hình ảnh tham chiếu dài hạn sau đó. Theo phương thức này, bộ mã hóa dữ liệu video 20 có thể được tạo cấu hình để mã hóa dữ liệu biểu diễn các MSB của các giá trị POC cho hình ảnh tham chiếu dài hạn dưới dạng các giá trị chênh lệch, ví dụ, so với giá trị MSB cơ sở hoặc so với MSB được mã hóa trước đó cho các giá trị POC của hình ảnh tham chiếu dài hạn khác. Hơn nữa, bộ mã hóa dữ liệu video 20 có thể báo hiệu các giá trị chênh lệch này thậm chí khi LSB của các giá trị POC là khác nhau.

Như được thảo luận trên đây, trong một số ví dụ, bộ mã hóa dữ liệu video 20 có thể được tạo cấu hình để mã hóa các giá trị nguyên âm cho các giá trị khác nhau biểu diễn độ chênh lệch giữa các MSB của các giá trị POC cho một hoặc nhiều hình ảnh tham chiếu dài hạn và giá trị MSB cơ sở (hoặc các MSB tham chiếu khác). Ví dụ, khi bộ mã hóa entropy 56 (hoặc đơn vị khác của bộ mã hóa dữ liệu video 20) xác

định rằng hình ảnh hiện thời có giá trị POC bằng $\text{MaxPicOrderCntLsb} * N - 1$, và hình ảnh có giá trị POC bằng $\text{MaxPicOrderCntLsb} * N$ là LTRP thứ nhất mà $\text{delta_poc_msb_cycle_lt}[i]$ được báo hiệu, bộ mã hóa entropy 56 có thể mã hóa giá trị của -1 cho các MSB của giá trị POC của LTRP này.

Ngoài ra, hoặc theo phương án khác, bộ mã hóa dữ liệu video 20 có thể được tạo cấu hình để mã hóa dữ liệu biểu diễn các giá trị POC cho hình ảnh tham chiếu dài hạn dựa ít nhất một phần vào tổng số hình ảnh tham chiếu hiện đang được lưu trữ trong tập hợp hình ảnh tham chiếu (hoặc bộ nhớ đệm hình ảnh đã giải mã) và/hoặc một số hình ảnh tham chiếu ngắn hạn hiện đang được lưu trữ trong tập hợp hình ảnh tham chiếu (hoặc bộ nhớ đệm hình ảnh đã giải mã). Ví dụ, bộ mã hóa dữ liệu video 20 có thể được tạo cấu hình để xác định xem có báo hiệu giá trị chu kỳ MSB POC cho hình ảnh tham chiếu dài hạn hay không dựa trên việc có hình ảnh tham chiếu dài hạn khác bất kỳ trong tập hợp hình ảnh tham chiếu (hoặc bộ nhớ đệm hình ảnh đã giải mã) có giá trị LSB POC giống như hình ảnh tham chiếu dài hạn hay không. Do đó, khi chỉ có một hoặc nhiều hình ảnh tham chiếu ngắn hạn trong tập hợp hình ảnh tham chiếu có các LSB POC giống như hình ảnh tham chiếu dài hạn, bộ mã hóa dữ liệu video 20 có thể tránh mã hóa giá trị chu kỳ MSB POC cho hình ảnh tham chiếu dài hạn.

Ngoài ra, hoặc theo cách khác, bộ mã hóa dữ liệu video 20 có thể xác định không mã hóa giá trị chu kỳ MSB POC delta cho hình ảnh tham chiếu dài hạn khi có N hình ảnh tham chiếu dài hạn trong bộ nhớ đệm hình ảnh đã giải mã mà có các LSB giống như giá trị POC cho hình ảnh tham chiếu dài hạn và khi bộ mã hóa dữ liệu video 20 đã mã hóa các giá trị chu kỳ MSB POC delta cho N-1 hình ảnh tham chiếu dài hạn.

Bộ dự báo nội cấu trúc 46 có thể dự báo nội cấu trúc khối hiện thời, theo phương án khác với dự báo liên cấu trúc được thực hiện bởi bộ ước tính chuyển động 42 và bộ bù chuyển động 44, như được mô tả trên đây. Cụ thể là, bộ dự báo nội cấu trúc 46 có thể xác định chế độ dự báo nội cấu trúc dùng để mã hóa khối hiện thời. Trong một số ví dụ, bộ 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 lượt mã hóa riêng, và bộ dự báo nội cấu trúc 46 (hoặc bộ chọn chế độ 40, trong một số ví dụ) có thể chọn chế độ dự báo nội cấu trúc thích hợp để sử dụng từ các chế độ thử nghiệm này.

Ví dụ, bộ 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 các chế độ dự báo nội cấu trúc thử nghiệm khác nhau, và chọn chế độ dự báo nội cấu trúc có các đặc tính tốc độ-méo tốt nhất trong số các chế độ thử nghiệm. Phân tích tốc độ-méo thường xác định lượng méo (hoặc sai số) giữa khối mã hóa và khối ban đầu chưa mã hóa mà đã 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 dự báo nội cấu trúc 46 có thể tính toán các tỷ lệ từ các giá trị méo và tốc độ cho các 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 chọn chế độ dự báo nội cấu trúc cho khối, bộ 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 cho bộ mã hóa entropy 56. Bộ mã hóa entropy 56 có thể mã hóa thông tin chỉ báo chế độ dự báo nội cấu trúc được chọn. Bộ mã hóa dữ liệu video 20 có thể bao gồm trong dữ liệu cấu hình dòng bit được truyền, mà có thể bao gồm các bảng chỉ mục chế độ dự báo nội cấu trúc và các bảng chỉ mục chế độ dự báo nội cấu trúc hiệu chỉnh (còn được gọi là các bảng ánh xạ từ mã), các định nghĩa về các ngữ cảnh mã hóa cho các khối khác nhau, và các chỉ báo về chế độ dự báo nội cấu trúc có xác suất cao nhất, bảng chỉ mục chế độ dự báo nội cấu trúc, và bảng chỉ mục chế độ dự báo nội cấu trúc hiệu chỉnh để sử dụng cho mỗi ngữ cảnh này.

Bộ mã hóa dữ liệu video 20 tạo ra khối video dư bằng cách trừ đi dữ liệu dự báo từ bộ chọn chế độ 40 từ khối video ban đầu được mã hóa. Bộ cộng 50 biểu diễn thành phần hoặc các thành phần thực hiện phép toán trừ này. Bộ xử lý biến đổi 52 áp dụng quy trình biến đổi, như biến đổi cosin rời rạc (discrete cosine transform - DCT) hoặc biến đổi tương tự về mặt khái niệm, cho khối dư này, tạo ra khối video chứa các giá trị hệ số biến đổi dư. Bộ xử lý biến đổi 52 có thể thực hiện các biến đổi khác mà tương tự về mặt khái niệm với DCT. Các biến đổi dạng sóng, biến đổi số nguyên, biến đổi dải phụ hoặc các loại biến đổi khác có thể cũng được sử dụng. Trong mọi trường hợp, bộ xử lý biến đổi 52 áp dụng biến đổi này cho khối dư, tạo ra khối các hệ số biến đổi dư. Quy trình biến đổi này có thể chuyển đổi thông tin dư từ miền giá trị điểm ảnh sang miền biến đổi, như miền tần số. Bộ xử lý biến đổi 52 có thể truyền các hệ số biến đổi thu được này đến bộ lượng tử hóa 54. Bộ lượng tử hóa 54 lượng tử hóa các hệ số biến đổi này để giảm tiếp tốc độ bit. Quy trình lượng tử hóa này có thể giảm

độ sâu bit liên quan đến một số hoặc tất cả các hệ số. Mức lượng tử hóa có thể được hiệu chỉnh bằng cách điều chỉnh thông số lượng tử hóa. Trong một số ví dụ, sau đó bộ lượng tử hóa 54 có thể thực hiện quét ma trận chứa các hệ số biến đổi được lượng tử hóa. Theo cách khác, bộ mã hóa entropy 56 có thể thực hiện quy trình quét.

Sau khi lượng tử hóa, bộ mã hóa entropy 56 mã hóa entropy các hệ số biến đổi được lượng tử hóa. Ví dụ, bộ mã hóa entropy 56 có thể thực hiện mã hóa độ dài thay đổi thích ứng với ngữ cảnh (context adaptive variable length coding - CAVLC), mã hóa số học nhị phân thích ứng với ngữ cảnh (context adaptive binary arithmetic coding - CABAC), mã hóa số học nhị phân thích ứng với ngữ cảnh theo cú pháp (syntax-based context-adaptive binary arithmetic coding - SBAC), mã hóa entropy phân chia khoảng xác suất (probability interval partitioning entropy - PIPE) hoặc kỹ thuật mã hóa entropy khác. Trong trường hợp mã hóa entropy dựa trên ngữ cảnh, ngữ cảnh có thể được dựa trên các khối lân cận. Sau khi mã hóa entropy bằng bộ mã hóa entropy 56, dòng bit được mã hóa này có thể được truyền đến thiết bị khác (ví dụ, bộ giải mã dữ liệu video 30) hoặc lưu trữ để truyền hoặc khôi phục sau đó.

Bộ lượng tử hóa ngược 58 và bộ biến đổi ngược 60 lần lượt áp dụng quy trình lượng tử hóa ngược và biến đổi ngược, để khôi phục khối dư trong miền điểm ảnh, ví dụ, để sử dụng sau đó làm khối tham chiếu. Bộ bù chuyển động 44 có thể tính toán khối tham chiếu bằng cách bổ sung khối dư này vào khối dự báo của một trong số các khung của bộ nhớ hình ảnh tham chiếu 64. Bộ 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 khi ước tính chuyển động. Bộ cộng 62 bổ sung khối dư được khôi phục vào khối dự báo bù chuyển động được tạo ra bởi bộ 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 có thể được sử dụng bởi bộ ước tính chuyển động 42 và bộ bù chuyển động 44 làm khối tham chiếu để mã hóa nội cấu trúc khối trong khung video tiếp theo này.

Bộ mã hóa dữ liệu video 20 trên Fig.2 biểu diễn ví dụ về bộ mã hóa dữ liệu video có thể được tạo cấu hình để mã hóa dữ liệu cho hình ảnh hiện thời của dữ liệu video, trong đó dữ liệu được mã hóa biểu diễn giá trị nguyên âm cho chu kỳ các bit có giá trị lớn nhất (MSB) của số đếm thứ tự hình ảnh (POC) cho hình ảnh tham chiếu dài hạn cho hình ảnh hiện thời, và mã hóa ít nhất một phần hình ảnh hiện thời dựa

vào hình ảnh tham chiếu dài hạn dựa ít nhất một phần vào giá trị nguyên âm cho chu kỳ MSB.

Bộ mã hóa dữ liệu video 20 trên Fig.2 cũng biểu diễn ví dụ về bộ mã hóa dữ liệu video mà có thể, ngoài ra hoặc theo cách khác, được tạo cấu hình để mã hóa ít nhất phần thứ nhất của hình ảnh hiện thời của dữ liệu video dựa vào hình ảnh tham chiếu dài hạn thứ nhất và ít nhất phần thứ hai của hình ảnh hiện thời dựa vào hình ảnh tham chiếu dài hạn thứ hai, mã hóa giá trị thứ nhất biểu diễn độ chênh lệch giữa giá trị bit có giá trị lớn nhất (MSB) cơ sở của giá trị số đếm thứ tự hình ảnh (POC) của hình ảnh hiện thời của dữ liệu video và giá trị MSBs thứ nhất của giá trị POC thứ nhất của hình ảnh tham chiếu dài hạn thứ nhất của dữ liệu video, và mã hóa giá trị thứ hai biểu diễn độ chênh lệch giữa giá trị MSBs thứ hai của giá trị POC thứ hai của hình ảnh tham chiếu dài hạn thứ hai của dữ liệu video và giá trị MSBs thứ nhất, trong đó giá trị POC thứ nhất và giá trị POC thứ hai có các giá trị bit có giá trị nhỏ nhất (LSB) khác nhau.

Bộ mã hóa dữ liệu video 20 trên Fig.2 cũng biểu diễn ví dụ về bộ mã hóa dữ liệu video mà có thể, ngoài ra hoặc theo cách khác, được tạo cấu hình để xác định số hình ảnh tham chiếu được đánh dấu là hình ảnh tham chiếu ngắn hạn, trong tổng số hình ảnh tham chiếu sẵn có, cho hình ảnh hiện thời, mã hóa giá trị bit có giá trị lớn nhất (MSB) cho các hình ảnh tham chiếu dài hạn cho hình ảnh hiện thời dựa trên tổng số hình ảnh tham chiếu và số hình ảnh tham chiếu được đánh dấu là hình ảnh tham chiếu ngắn hạn, và mã hóa ít nhất một phần hình ảnh hiện thời ít nhất một trong số các hình ảnh tham chiếu sẵn có.

Bộ mã hóa dữ liệu video 20 trên Fig.2 cũng biểu diễn ví dụ về bộ mã hóa dữ liệu video mà có thể, ngoài ra hoặc theo cách khác, được tạo cấu hình để xác định số N hình ảnh tham chiếu cho hình ảnh hiện thời mà có các giá trị số đếm thứ tự hình ảnh (POC) với các giá trị bit có giá trị thấp nhất (LSB) tương đương, xác định rằng tập hợp hình ảnh tham chiếu của hình ảnh hiện thời có thể bao gồm N hình ảnh tham chiếu dài hạn, và mã hóa các giá trị chu kỳ bit có giá trị lớn nhất (MSB) của POC delta cho nhiều nhất N-1 hình ảnh tham chiếu dài hạn dựa trên các bước xác định nêu trên.

Fig.3 là sơ đồ khối minh họa ví dụ về bộ giải mã dữ liệu video 30 có thể thực hiện các kỹ thuật để báo hiệu tốt hơn các hình ảnh tham chiếu dài hạn ở phần đầu lát.

Trong ví dụ trên Fig.3, bộ giải mã dữ liệu video 30 bao gồm bộ giải mã entropy 70, bộ bù chuyển động 72, bộ dự báo nội cấu trúc 74, bộ lượng tử hóa ngược 76, bộ biến đổi ngược 78, bộ nhớ hình ảnh tham chiếu 82 và bộ cộng 80. Bộ giải mã dữ liệu video 30, trong một số ví dụ, có thể thực hiện lượt giải mã thường ngược với lượt mã hóa được mô tả đối với bộ mã hóa dữ liệu video 20 (Fig.2). Bộ bù chuyển động 72 có thể tạo ra dữ liệu dự báo dựa trên các vector chuyển động được thu được từ bộ giải mã entropy 70, trong khi bộ dự báo nội cấu trúc 74 có thể tạo ra dữ liệu dự báo dựa trên các chỉ báo chế độ dự báo nội cấu trúc thu được từ bộ giải mã entropy 70.

Trong quá trình giải mã, bộ giải mã dữ liệu video 30 thu dòng bit dữ liệu video mã hóa biểu diễn khối video của lát video mã hóa và các phần tử cú pháp liên quan từ bộ mã hóa dữ liệu video 20. Bộ giải mã entropy 70 của bộ giải mã dữ liệu video 30 entropy giải mã dòng bit này để tạo ra các hệ số lượng tử hóa, các vector chuyển động hoặc các chỉ báo chế độ dự báo nội cấu trúc, và các phần tử cú pháp khác. Bộ giải mã entropy 70 chuyển tiếp các vector chuyển động và các phần tử cú pháp khác đến bộ bù chuyển động 72. Bộ giải mã dữ liệu video 30 có thể thu được các phần tử cú pháp này 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 dự báo nội cấu trúc 74 có thể tạo ra dữ liệu dự báo cho khối video của lát video hiện thời dựa trên 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, lát B, P hoặc GPB), bộ bù chuyển động 72 tạo ra các khối dự báo cho khối video của lát video hiện thời dựa trên vector chuyển động này và các phần tử cú pháp khác thu được từ bộ giải mã entropy 70. Khối dự báo có thể được tạo ra từ một trong số các hình ảnh tham chiếu trong một trong số các danh mục hình ảnh tham chiếu. Bộ giải mã dữ liệu video 30 có thể thiết lập các danh mục hình ảnh tham chiếu, danh mục 0 và danh mục 1, bằng cách sử dụng các kỹ thuật thiết lập mặc định dựa trên các hình ảnh tham chiếu được lưu trữ trong bộ nhớ hình ảnh tham chiếu 82.

Bộ giải mã dữ liệu video 30 có thể xác định tập hợp hình ảnh tham chiếu, mà từ đó bộ giải mã dữ liệu video 30 thiết lập các danh mục hình ảnh tham chiếu. Trong một số ví dụ, ví dụ, khi bộ giải mã dữ liệu video 30 tuân theo chuẩn HEVC, bộ giải mã dữ liệu video 30 có thể thiết lập các danh mục hình ảnh tham chiếu theo phần

8.3.3 của HEVC WD7. Tập hợp hình ảnh tham chiếu thường tương ứng với các hình ảnh tham chiếu hiện đang được lưu trữ trong bộ nhớ đệm hình ảnh đã giải mã, ví dụ, trong bộ nhớ hình ảnh tham chiếu 82. Các hình ảnh tham chiếu nhất định của tập hợp hình ảnh tham chiếu có thể được đánh dấu là các hình ảnh tham chiếu ngắn hạn, trong khi các hình ảnh tham chiếu khác của tập hợp hình ảnh tham chiếu có thể được đánh dấu là hình ảnh tham chiếu dài hạn. Nhìn chung, hình ảnh tham chiếu dài hạn được đánh dấu là “được sử dụng làm tham chiếu dài hạn,” ví dụ, trong phần đầu đơn vị lớp trừu tượng hóa mạng (network abstraction layer - NAL) của đơn vị NAL mà bao gói dữ liệu video được mã hóa cho hình ảnh tham chiếu này.

Theo các kỹ thuật của sáng chế, bộ giải mã dữ liệu video 30 có thể được tạo cấu hình để giải mã các ký hiệu nhận dạng cho hình ảnh tham chiếu dài hạn bằng cách sử dụng quy trình mã hóa vi phân. Ví dụ, với hình ảnh tham chiếu dài hạn thứ nhất (ví dụ, hình ảnh tham chiếu dài hạn có giá trị chỉ mục là 0 trong chuỗi hình ảnh), bộ giải mã dữ liệu video 30 có thể giải mã giá trị chênh lệch chỉ báo độ chênh lệch giữa các MSB POC cho hình ảnh tham chiếu dài hạn và các MSB POC cho hình ảnh hiện thời. Các MSB POC cho hình ảnh hiện thời có thể được gọi là “MSB cơ sở.” Với các hình ảnh tham chiếu dài hạn khác (ví dụ, hình ảnh tham chiếu dài hạn có giá trị chỉ mục lớn hơn 0 trong chuỗi hình ảnh), bộ giải mã dữ liệu video 30 có thể giải mã giá trị chênh lệch chỉ báo độ chênh lệch giữa các MSB POC cho hình ảnh tham chiếu dài hạn và các MSB POC cho hình ảnh tham chiếu dài hạn trước đó. Giả sử hình ảnh tham chiếu dài hạn cho các MSB POC hiện đang được xác định có giá trị chỉ mục i , bộ giải mã dữ liệu video 30 có thể dùng các MSB POC của hình ảnh tham chiếu dài hạn có giá trị chỉ mục là giá trị bất kỳ nằm trong khoảng từ 0 đến $i-1$, bao gồm cả giá trị đầu mút, để xác định các MSB POC. Nghĩa là, bộ giải mã dữ liệu video 30 có thể dùng các MSB POC của hình ảnh tham chiếu dài hạn có giá trị chỉ mục j , trong đó j nằm trong khoảng từ 0 đến $i-1$. Giá trị chênh lệch này có thể trong một số trường hợp có giá trị nguyên âm.

Sau khi thu được giá trị chênh lệch mô tả độ chênh lệch giữa các MSB POC cho hình ảnh tham chiếu dài hạn, bộ giải mã dữ liệu video 30 có thể xác định giá trị MSB POC mà giá trị chênh lệch này áp dụng (nghĩa là, giá trị chênh lệch được thêm vào đó để khôi phục các MSB POC cho hình ảnh tham chiếu dài hạn). Ví dụ, nếu hình ảnh tham chiếu dài hạn có giá trị chỉ mục là 0 (chỉ báo rằng hình ảnh tham chiếu

dài hạn là hình ảnh tham chiếu dài hạn thứ nhất trong chuỗi các hình ảnh), bộ giải mã dữ liệu video 30 có thể sử dụng giá trị MSB POC của hình ảnh hiện thời làm giá trị MSB cơ sở, và bổ sung giá trị chênh lệch được báo hiệu cho hình ảnh tham chiếu dài hạn vào giá trị MSB cơ sở. Theo ví dụ khác, nếu hình ảnh tham chiếu dài hạn có giá trị chỉ mục lớn hơn 0 (chỉ báo rằng hình ảnh tham chiếu dài hạn không phải là hình ảnh tham chiếu dài hạn thứ nhất trong chuỗi các hình ảnh), giả sử giá trị chỉ mục này là i , bộ giải mã dữ liệu video 30 có thể bổ sung giá trị chênh lệch được báo hiệu vào các MSB POC cho hình ảnh tham chiếu dài hạn có giá trị chỉ mục $i-1$.

Theo phương thức này, bộ giải mã dữ liệu video 30 có thể giải mã giá trị thứ nhất biểu diễn độ chênh lệch giữa giá trị bit có giá trị lớn nhất (MSB) cơ sở của giá trị số đếm thứ tự hình ảnh (POC) của hình ảnh hiện thời của dữ liệu video và giá trị MSBs thứ nhất của giá trị POC thứ nhất của hình ảnh tham chiếu dài hạn thứ nhất của dữ liệu video, và giải mã giá trị thứ hai biểu diễn độ chênh lệch giữa giá trị MSBs thứ hai của giá trị POC thứ hai của hình ảnh tham chiếu dài hạn thứ hai của dữ liệu video và giá trị MSBs thứ nhất, trong đó giá trị POC thứ nhất và giá trị POC thứ hai có các giá trị bit có giá trị nhỏ nhất (LSB) khác biệt. Tương tự, bộ giải mã dữ liệu video 30 có thể tính toán giá trị chu kỳ MSB thứ nhất cho hình ảnh tham chiếu dài hạn thứ nhất là $\text{DeltaPocMSBCycleLt}[i - 1]$ sử dụng giá trị thứ nhất (nghĩa là, giá trị chênh lệch thứ nhất), và tính toán giá trị chu kỳ MSB thứ hai cho hình ảnh tham chiếu dài hạn thứ hai là $\text{DeltaPocMSBCycleLt}[i]$, trong đó việc tính toán giá trị chu kỳ MSB thứ hai bao gồm tính toán $\text{DeltaPocMSBCycleLt}[i]$ dựa trên $\text{DeltaPocMSBCycleLt}[i - 1]$ và giá trị thứ hai (nghĩa là, giá trị chênh lệch thứ hai).

Ngoài ra, bộ giải mã dữ liệu video 30 có thể giải mã các LSB cho các giá trị POC của hình ảnh tham chiếu dài hạn. Các LSB có thể được báo hiệu trong phần đầu lát của lát cho hình ảnh hiện thời, tập thông số chuỗi (sequence parameter set - SPS) cho chuỗi bao gồm ảnh hiện thời, tập thông số hình ảnh (picture parameter set - PPS) cho hình ảnh hiện thời, hoặc trong phần khác. Bộ giải mã dữ liệu video 30 có thể nối kết MSB và LSB cho giá trị POC của hình ảnh tham chiếu dài hạn để xác định giá trị POC đầy đủ của hình ảnh tham chiếu dài hạn. Sau đó, bộ giải mã dữ liệu video 30 có thể kết hợp danh mục hình ảnh tham chiếu bao gồm hình ảnh tham chiếu dài hạn được xác định bởi giá trị POC. Theo phương thức này, bộ giải mã dữ liệu video 30 có thể sử dụng thông tin chuyển động (ví dụ, dữ liệu xác định danh mục hình ảnh tham

chiều và chỉ mục tham chiếu) để thực hiện dự báo liên cấu trúc. Ví dụ, như được giải thích dưới đây, bộ bù chuyển động 72 có thể xác định khối tham chiếu cho khối của hình ảnh hiện thời dựa ít nhất một phần vào thông tin chuyển động xác định hình ảnh tham chiếu, tương ứng với một trong số các hình ảnh trong một trong số các danh mục hình ảnh tham chiếu. Hình ảnh tham chiếu này có thể tương ứng với một trong số các hình ảnh tham chiếu dài hạn, hoặc với hình ảnh tham chiếu ngắn hạn.

Bộ bù chuyển động 72 xác định thông tin dự báo cho khối video của lát video hiện thời bằng cách phân tích cú pháp các vector chuyển động và các phần tử cú pháp khác, và sử dụng thông tin dự báo này để tạo ra các khối dự báo cho khối video hiện thời đang được giải mã. Ví dụ, bộ bù chuyển động 72 sử dụng một số phần tử cú pháp thu được để xác định chế độ dự báo (ví dụ, dự báo nội cấu hoặc liên cấu trúc) được sử dụng để mã hóa các khối video của lát video, kiểu lát dự báo liên cấu trúc (ví dụ, lát B, lát P, hoặc lát GPB), thông tin thiết lập cho một hoặc nhiều danh mục hình ảnh tham chiếu cho lát, các vector chuyển động cho mỗi khối video được mã hóa liên cấu trúc của lát, tình trạng 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 bù chuyển động 72 cũng có thể thực hiện phép nội suy dựa trên các bộ lọc nội suy. Bộ bù chuyển động 72 có thể sử dụng các bộ lọc nội suy như được sử dụng bởi bộ mã hóa dữ liệu 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 bù chuyển động 72 có thể xác định các bộ lọc nội suy được sử dụng bởi bộ mã hóa dữ liệu 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 lượng tử hóa ngược 76 lượng tử hóa ngược, nghĩa là, khử lượng tử, các hệ số biến đổi được lượng tử hóa có trong dòng bit và được giải mã bằng bộ giải mã entropy 70. Quy trình lượng tử hóa ngược có thể bao gồm sử dụng thông số lượng tử hóa QPY được tính toán bởi bộ giải mã dữ liệu video 30 cho mỗi khối video trong lát video để xác định mức lượng tử hóa và tương tự, mức lượng tử hóa ngược nên được áp dụng. Bộ biến đổi ngược 78 áp dụng quy trình biến đổi ngược, ví dụ, DCT ngược, biến đổi số nguyên ngược, hoặc quy trình biến đổi ngược tương tự về mặt khái niệm, cho các hệ số biến đổi để tạo ra các khối dư trong miền điểm ảnh.

Sau khi bộ bù chuyển động 72 tạo ra khối dự báo cho khối video hiện thời dựa

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

Theo phương thức này, bộ giải mã dữ liệu video 30 trên Fig.3 biểu diễn ví dụ về bộ giải mã dữ liệu video có thể được tạo cấu hình để giải mã giá trị thứ nhất biểu diễn độ chênh lệch giữa giá trị bit có giá trị lớn nhất (MSB) cơ sở của giá trị số đếm thứ tự hình ảnh (POC) của hình ảnh hiện thời của dữ liệu video và giá trị MSBs thứ nhất của giá trị POC thứ nhất của hình ảnh tham chiếu dài hạn thứ nhất của dữ liệu video, giải mã giá trị thứ hai biểu diễn độ chênh lệch giữa giá trị MSBs thứ hai của giá trị POC thứ hai của hình ảnh tham chiếu dài hạn thứ hai của dữ liệu video và giá trị MSBs thứ nhất, trong đó giá trị POC thứ nhất và giá trị POC thứ hai có giá trị bit có giá trị nhỏ nhất (LSB) khác nhau, và giải mã ít nhất một phần hình ảnh hiện thời của dữ liệu video dựa vào ít nhất một trong số các hình ảnh tham chiếu dài hạn thứ nhất sử dụng giá trị thứ nhất và hình ảnh tham chiếu dài hạn thứ hai sử dụng giá trị thứ nhất và giá trị thứ hai.

Bộ giải mã dữ liệu video 30 trên Fig.3 cũng biểu diễn ví dụ về bộ giải mã dữ liệu video có thể, ngoài ra hoặc theo cách khác, được tạo cấu hình để giải mã dữ liệu cho hình ảnh hiện thời của dữ liệu video, trong đó dữ liệu được giải mã biểu diễn giá trị nguyên âm cho chu kỳ các bit có giá trị lớn nhất (MSB) của số đếm thứ tự hình ảnh (POC) cho hình ảnh tham chiếu dài hạn cho hình ảnh hiện thời, và giải mã ít nhất một phần hình ảnh hiện thời dựa vào hình ảnh tham chiếu dài hạn dựa ít nhất một phần vào giá trị nguyên âm cho chu kỳ MSB.

Bộ giải mã dữ liệu video 30 trên Fig.3 cũng biểu diễn ví dụ về bộ giải mã dữ

liệu video mà có thể, ngoài ra hoặc theo cách khác, được tạo cấu hình để xác định số hình ảnh tham chiếu được đánh dấu là hình ảnh tham chiếu ngắn hạn, trong tổng số hình ảnh tham chiếu sẵn có, cho hình ảnh hiện thời, giải mã giá trị các bit có giá trị lớn nhất (MSB) cho các hình ảnh tham chiếu dài hạn cho hình ảnh hiện thời dựa trên tổng số hình ảnh tham chiếu và số hình ảnh tham chiếu được đánh dấu là hình ảnh tham chiếu ngắn hạn, và giải mã ít nhất một phần hình ảnh hiện thời dựa vào ít nhất một trong số các hình ảnh tham chiếu sẵn có này.

Bộ giải mã dữ liệu video 30 trên Fig.3 cũng biểu diễn ví dụ về bộ giải mã dữ liệu video mà có thể, ngoài ra hoặc theo cách khác, được tạo cấu hình để xác định số N của hình ảnh tham chiếu cho hình ảnh hiện thời mà có các giá trị số đếm thứ tự hình ảnh (POC) với các giá trị bit có giá trị nhỏ nhất (LSB) tương đương, xác định rằng tập hợp hình ảnh tham chiếu này có thể bao gồm N hình ảnh tham chiếu dài hạn, và giải mã các giá trị chu kỳ bit có giá trị lớn nhất (MSB) của POC delta cho nhiều nhất N-1 hình ảnh tham chiếu dài hạn dựa trên các bước xác định nêu trên.

Fig.4 là sơ đồ khái niệm minh họa chuỗi hình ảnh video được mã hóa từ 100 đến 132. Các hình ảnh này được tô màu khác nhau để chỉ báo các vị trí trong cấu trúc dự báo phân cấp. Ví dụ, các hình ảnh 100, 116, và 132 là tô màu đen để biểu diễn rằng các hình ảnh 100, 116, 132 là ở trên cùng của cấu trúc dự báo phân cấp. Các hình ảnh 100, 116, 132 có thể bao gồm, ví dụ, các hình ảnh mã hóa nội cấu trúc hoặc các hình ảnh mã hóa liên cấu trúc được dự báo từ các hình ảnh khác theo một chiều (ví dụ, các hình ảnh P). Khi được mã hóa nội cấu trúc, các hình ảnh 100, 116, 132 được dự báo chỉ từ dữ liệu trong cùng hình ảnh. Khi được mã hóa liên cấu trúc, hình ảnh 116, ví dụ, có thể được mã hóa dựa vào dữ liệu của hình ảnh 100, như được chỉ báo bởi mũi tên nét đứt từ hình ảnh 116 đến hình ảnh 100. Các hình ảnh 116, 132 lần lượt tạo thành các hình ảnh chính của nhóm hình ảnh (group of picture - GOP) 134, 136.

Các hình ảnh 108, 124 được tô màu tối để chỉ báo rằng chúng đứng tiếp theo trong phân cấp mã hóa sau các hình ảnh 100, 116, và 132. Các hình ảnh 108, 124 có thể bao gồm các hình ảnh được mã hóa dự báo chế độ liên cấu trúc, hai chiều. Ví dụ, hình ảnh 108 có thể được dự báo từ dữ liệu của hình ảnh 100 và 116, trong khi hình ảnh 124 có thể được dự báo từ hình ảnh 116 và 132. Các hình ảnh 104, 112, 120, và 128 được tô màu sáng để chỉ báo rằng chúng đứng tiếp theo trong phân cấp mã hóa

sau các hình ảnh 108 và 124. Các hình ảnh 104, 112, 120, và 128 cũng có thể bao gồm các hình ảnh được mã hóa dự báo chế độ liên cấu trúc, hai chiều. Ví dụ, hình ảnh 104 có thể được dự báo từ các hình ảnh 100 và 108, hình ảnh 112 có thể được dự báo từ các hình ảnh 108 và 116, hình ảnh 120 có thể được dự báo từ hình ảnh 116 và 124, và hình ảnh 128 có thể được dự báo từ hình ảnh 124 và 132. Nhìn chung, các hình ảnh ở thấp hơn trong thứ bậc có thể được mã hóa từ hình ảnh tham chiếu bất kỳ ở cao hơn trong thứ bậc này, giả sử các hình ảnh tham chiếu vẫn được đệm trong bộ nhớ đệm hình ảnh đã giải mã, và giả sử các hình ảnh tham chiếu được mã hóa trước hình ảnh hiện đang được mã hóa.

Cuối cùng, các hình ảnh 102, 106, 110, 114, 118, 122, 126, và 130 được tô màu trắng để chỉ báo rằng các hình ảnh này đứng cuối cùng trong phân cấp mã hóa. Các hình ảnh 102, 106, 110, 114, 118, 122, 126, và 130 có thể là các hình ảnh được mã hóa dự báo chế độ liên cấu trúc, hai chiều. Hình ảnh 102 có thể được dự báo từ các hình ảnh 100 và 104, hình ảnh 106 có thể được dự báo từ các hình ảnh 104 và 108, hình ảnh 110 có thể được dự báo từ các hình ảnh 108 và 112, hình ảnh 114 có thể được dự báo từ các hình ảnh 112 và 116, hình ảnh 118 có thể được dự báo từ hình ảnh 116 và 120, hình ảnh 122 có thể được dự báo từ các hình ảnh 120 và 124, hình ảnh 126 có thể được dự báo từ các hình ảnh 124 và 128, và hình ảnh 130 có thể được dự báo từ các hình ảnh 128 và 132. Hơn nữa, nên hiểu rằng hình ảnh ở thấp hơn trong phân cấp mã hóa có thể được mã hóa từ các hình ảnh khác ở cao hơn trong phân cấp mã hóa. Ví dụ, ngoài ra hoặc theo cách khác, bất kỳ hoặc tất cả các hình ảnh 102, 106, 110, hoặc 114 có thể được dự báo dựa vào hình ảnh bất kỳ trong số các hình ảnh 100, 116, hoặc 108.

Các hình ảnh từ 100-132 được minh họa theo thứ tự hiển thị. Nghĩa là, sau khi giải mã, hình ảnh 100 được hiển thị trước hình ảnh 102, hình ảnh 102 được hiển thị trước hình ảnh 104, v.v.. Như được thảo luận trên đây, các giá trị POC thường mô tả thứ tự hiển thị cho các hình ảnh, mà hầu như cũng giống như thứ tự mà các hình ảnh thô được chụp hoặc được tạo ra trước khi được mã hóa. Tuy nhiên, do phân cấp mã hóa, các hình ảnh 100-132 có thể được giải mã theo thứ tự khác. Hơn nữa, trong khi được mã hóa, các hình ảnh 100-132 có thể được sắp xếp theo thứ tự giải mã trong dòng bit chứa dữ liệu được mã hóa cho các hình ảnh 100-132. Ví dụ, hình ảnh 116 có thể được hiển thị cuối cùng trong số các hình ảnh của GOP 134. Tuy nhiên, do phân

cấp mã hóa, hình ảnh 116 có thể được giải mã đầu tiên trong GOP 134. Nghĩa là, để giải mã thích hợp hình ảnh 108, ví dụ, hình ảnh 116 có thể cần được giải mã đầu tiên, để đóng vai trò là hình ảnh tham chiếu cho hình ảnh 108. Tương tự, hình ảnh 108 có thể đóng vai trò là hình ảnh tham chiếu cho các hình ảnh 104, 106, 110, và 112, và do đó có thể cần được giải mã trước các hình ảnh 104, 106, 110, và 112.

Hơn nữa, các hình ảnh nhất định có thể được coi là hình ảnh tham chiếu dài hạn, trong khi các hình ảnh khác có thể được coi là hình ảnh tham chiếu ngắn hạn. Giả sử, ví dụ, các hình ảnh 100 và 116 biểu diễn hình ảnh tham chiếu dài hạn, trong khi các hình ảnh 108, 104, và 112 biểu diễn các hình ảnh tham chiếu ngắn hạn. Đây có thể là trường hợp, trong ví dụ này, các hình ảnh 102 và 106 có thể được dự báo dựa vào hình ảnh bất kỳ trong số các hình ảnh 100, 116, 108, hoặc 104, nhưng các hình ảnh 110 và 114 có thể được dự báo dựa vào hình ảnh bất kỳ trong số các hình ảnh 100, 116, 108, hoặc 112. Nói cách khác, hình ảnh 104 có thể không sẵn có để dùng làm tham chiếu khi mã hóa các hình ảnh 110 và 114. Trong ví dụ khác, giả sử các hình ảnh 100 và 116 biểu diễn hình ảnh tham chiếu dài hạn và các hình ảnh 108, 104, và 112 biểu diễn các hình ảnh tham chiếu ngắn hạn, các hình ảnh 108, 104, và 112 có thể không sẵn có để dùng làm tham chiếu khi mã hóa các hình ảnh 118, 122, 126, và 130.

Theo các kỹ thuật của sáng chế, dữ liệu về hình ảnh tham chiếu dài hạn có thể được báo hiệu ở phần đầu lát của các lát cho hình ảnh bất kỳ hoặc tất cả các hình ảnh từ 100-132. Theo phương án khác, dữ liệu này có thể được báo hiệu trong các PPS, SPS, hoặc các cấu trúc dữ liệu khác.

Lại giả sử các hình ảnh 100 và 116 biểu diễn hình ảnh tham chiếu dài hạn, bộ mã hóa dữ liệu video 20 có thể mã hóa các MSB POC cho hình ảnh 100 dưới dạng độ chênh lệch so với giá trị MBS cơ sở, ví dụ, các MSB POC của hình ảnh hiện thời, như hình ảnh 102. Nghĩa là, bộ mã hóa dữ liệu video 20 có thể tính toán độ chênh lệch giữa các MSB POC cho hình ảnh 100 và các MSB POC cho hình ảnh 102, và mã hóa giá trị chênh lệch cho hình ảnh 100. Tương tự như vậy, bộ mã hóa dữ liệu video 20 có thể tính toán độ chênh lệch giữa các MSB POC cho hình ảnh 116 và các MSB POC cho hình ảnh 100, và mã hóa giá trị chênh lệch này cho hình ảnh 116. Bộ mã hóa dữ liệu video 20 có thể cũng mã hóa các giá trị LSB cho các hình ảnh 100 và 116.

Do đó, bộ giải mã dữ liệu video 30 có thể khôi phục các giá trị POC cho các hình ảnh 100 và 116 bằng cách giải mã giá trị chênh lệch cho hình ảnh 100 và bổ sung giá trị chênh lệch này vào giá trị MSB cơ sở cho hình ảnh 102. Tương tự như vậy, bộ giải mã dữ liệu video 30 có thể giải mã giá trị chênh lệch cho hình ảnh 116 và bổ sung giá trị chênh lệch này vào các MSB POC cho hình ảnh 100. Bộ giải mã dữ liệu video 30 có thể cũng giải mã giá trị LSB POC cho các hình ảnh 100 và 116. Bộ giải mã dữ liệu video 30 có thể nối các MSB POC cho hình ảnh 100 với các LSB POC cho hình ảnh 100 để tái tạo giá trị POC cho hình ảnh 100. Tương tự như vậy, bộ giải mã dữ liệu video 30 có thể nối các MSB POC cho hình ảnh 116 với các LSB POC cho hình ảnh 116 để tái tạo giá trị POC cho hình ảnh 116.

Theo phương thức này, cả bộ mã hóa dữ liệu video 20 và bộ giải mã dữ liệu video 30 có thể được tạo cấu hình để mã hóa giá trị thứ nhất biểu diễn độ chênh lệch giữa giá trị bit có giá trị lớn nhất (MSB) cơ sở của giá trị số đếm thứ tự hình ảnh (POC) của hình ảnh hiện thời của dữ liệu video và giá trị MSBs thứ nhất của giá trị POC thứ nhất của hình ảnh tham chiếu dài hạn thứ nhất của dữ liệu video, mã hóa giá trị thứ hai biểu diễn độ chênh lệch giữa giá trị MSBs thứ hai của giá trị POC thứ hai của hình ảnh tham chiếu dài hạn thứ hai của dữ liệu video và giá trị MSBs thứ nhất, trong đó giá trị POC thứ nhất và giá trị POC thứ hai có các giá trị bit có giá trị nhỏ nhất (LSB), và mã hóa ít nhất một phần hình ảnh hiện thời của dữ liệu video dựa vào ít nhất một trong số hình ảnh tham chiếu dài hạn thứ nhất sử dụng giá trị thứ nhất và hình ảnh tham chiếu dài hạn thứ hai sử dụng giá trị thứ nhất và giá trị thứ hai.

Cụ thể hơn, bộ mã hóa dữ liệu video 20 có thể mã hóa dữ liệu biểu diễn các giá trị POC cho hình ảnh tham chiếu dài hạn được đưa vào trong danh mục hình ảnh tham chiếu cho hình ảnh hiện thời, ví dụ, hình ảnh 102. Nhìn chung, bộ mã hóa dữ liệu video 20 có thể chỉ báo rằng hình ảnh tham chiếu nên được đưa vào trong danh mục hình ảnh tham chiếu khi ít nhất một phần hình ảnh hiện thời được mã hóa dựa vào hình ảnh tham chiếu này. Một phần hình ảnh có thể được coi là được mã hóa dựa vào hình ảnh tham chiếu khi khối hình ảnh này được mã hóa liên cấu trúc dựa vào hình ảnh tham chiếu, sao cho vector chuyển động cho khối này chỉ đến khối tham chiếu của hình ảnh tham chiếu này. Thông tin chuyển động cho khối này có thể bao gồm ký hiệu nhận dạng danh mục hình ảnh tham chiếu (ví dụ, danh mục 0 hoặc danh mục 1), cũng như chỉ mục tham chiếu tương ứng với hình ảnh tham chiếu trong danh

mục hình ảnh tham chiếu được xác định bằng ký hiệu nhận dạng danh mục hình ảnh tham chiếu.

Do đó, bộ mã hóa dữ liệu video 20 có thể mã hóa một khối của hình ảnh 102 dựa vào khối tham chiếu của hình ảnh 100. Nghĩa là, bộ mã hóa dữ liệu video 20 có thể thực hiện tìm kiếm chuyển động cho khối này và xác định rằng khối tham chiếu của hình ảnh 100 tạo ra giá trị sai số chấp nhận được, so với các khối tham chiếu khác. Tương tự, bộ mã hóa dữ liệu video 20 có thể mã hóa khối khác của hình ảnh 102 dựa vào khối tham chiếu của hình ảnh 116.

Fig.5 là lưu đồ minh họa phương pháp làm ví dụ để mã hóa các giá trị POC cho các hình ảnh tham chiếu dài hạn theo các kỹ thuật của sáng chế. Để giải thích, phương pháp trên Fig.5 được giải thích liên quan đến bộ mã hóa dữ liệu video 20 trên Fig.1 và 2 và các thành phần của nó. Tuy nhiên, nên hiểu rằng các thiết bị mã hóa video khác có thể được tạo cấu hình để thực hiện phương pháp trên Fig.5.

Ban đầu, bộ mã hóa dữ liệu video 20 có thể mã hóa dữ liệu của hình ảnh hiện thời bằng cách sử dụng một hoặc nhiều hình ảnh tham chiếu (150). Bộ mã hóa dữ liệu video 20 sau đó có thể xác định các hình ảnh tham chiếu được sử dụng để mã hóa hình ảnh hiện thời (152). Cụ thể hơn, bộ mã hóa dữ liệu video 20 có thể xác định hình ảnh tham chiếu nào trong tập hợp các hình ảnh tham chiếu sẵn có thực tế được sử dụng làm tham chiếu khi mã hóa hình ảnh hiện thời. Trong một số ví dụ, các bước 150 và 152 có thể được tích hợp và được thực hiện hầu như đồng thời. Các hình ảnh tham chiếu sẵn có này có thể bao gồm hình ảnh tham chiếu dài hạn (LTRP) và hình ảnh tham chiếu ngắn hạn (STRP). Bộ mã hóa dữ liệu video 20 có thể xác định xem thiết lập hình ảnh tham chiếu là hình ảnh tham chiếu dài hạn hay hình ảnh tham chiếu ngắn hạn dựa trên các kết quả mã hóa thu được qua một số lượt mã hóa thử nghiệm khác nhau. Ví dụ, khi các kết quả mã hóa (ví dụ, các số đo tối ưu hóa tốc độ-méo (rate-distortion optimization - RDO)) tốt hơn khi hình ảnh tham chiếu cụ thể được coi là hình ảnh tham chiếu dài hạn (độ dịch tiềm tàng bằng dung lượng bộ nhớ được sử dụng để lưu trữ hình ảnh tham chiếu dài hạn), bộ mã hóa dữ liệu video 20 có thể coi hình ảnh tham chiếu này là hình ảnh tham chiếu dài hạn.

Trong mọi trường hợp, sau khi xác định tập hợp các hình ảnh tham chiếu và các hình ảnh tham chiếu được coi là hình ảnh tham chiếu dài hạn, bộ mã hóa dữ liệu video 20 có thể mã hóa thông tin chỉ báo hình ảnh nào được đưa vào trong các danh

mục hình ảnh tham chiếu cho ảnh hiện thời dưới dạng hình ảnh tham chiếu dài hạn. Thông tin được mã hóa có thể bao gồm dữ liệu biểu diễn các giá trị POC cho các hình ảnh tham chiếu dài hạn. Bộ mã hóa dữ liệu video 20 có thể mã hóa thông tin này trong phần đầu lát của lát của hình ảnh hiện thời, PPS tương ứng với hình ảnh hiện thời, SPS cho chuỗi bao gồm hình ảnh hiện thời, hoặc trong phần khác.

Bộ mã hóa dữ liệu video 20 có thể mã hóa dữ liệu chỉ báo các MSB của giá trị POC cho hình ảnh tham chiếu dài hạn thứ nhất, cũng như dữ liệu chỉ báo các LSB của giá trị POC cho hình ảnh tham chiếu dài hạn thứ nhất (154). Để mã hóa các MSB, bộ mã hóa dữ liệu video 20 có thể tính toán độ chênh lệch giữa các MSB của giá trị POC cho hình ảnh tham chiếu dài hạn thứ nhất và các MSB của giá trị POC cho hình ảnh hiện thời, và mã hóa giá trị chênh lệch được tính toán này.

Bộ mã hóa dữ liệu video 20 sau đó có thể xác định xem các LSB của giá trị POC cho LTRP sau đó có bằng các LSB của giá trị POC cho hình ảnh tham chiếu khác cho hình ảnh hiện thời (156) hay không. Khi bộ mã hóa dữ liệu video 20 xác định rằng các LSB của giá trị POC cho LTRP sau đó bằng các LSB của giá trị POC cho hình ảnh tham chiếu khác cho hình ảnh hiện thời (nhánh “CÓ” của 156), bộ mã hóa dữ liệu video 20 có thể mã hóa dữ liệu biểu diễn độ chênh lệch giữa các MSB của giá trị POC cho LTRP sau đó và các MSB của giá trị POC cho LTRP trước đó (nghĩa là, LTRP gần đây nhất mà thông tin MSB POC được mã hóa) (158). Các LSB của giá trị POC cho LTRP gần đây nhất không cần thiết giống với các LSB của giá trị POC cho LTRP đang được mã hóa hiện thời. Như được thảo luận trên đây, để tính toán giá trị chênh lệch, giả sử LTRP sau đó có giá trị chỉ mục i , bộ mã hóa dữ liệu video 20 có thể tính toán độ chênh lệch giữa các MSB của giá trị POC cho LTRP i và các MSB của giá trị POC cho LTRP j nếu LTRP j là LTRP trước đó mà thông tin MSB POC được mã hóa. Bộ mã hóa dữ liệu video 20 sau đó có thể mã hóa giá trị chênh lệch này. Giá trị chênh lệch này có thể tương ứng với $\text{delta_poc_msb_cycle_lt}[i]$ (hoặc $\text{delta_poc_msb_cycle_lt_plus1}[i]$).

Sau khi mã hóa giá trị chênh lệch này, hoặc khi các LSB của giá trị POC cho LTRP không bằng các LSB của hình ảnh tham chiếu khác (nhánh “KHÔNG” của 156), bộ mã hóa dữ liệu video 20 có thể mã hóa các LSB của giá trị POC cho LTRP (160). Bộ mã hóa dữ liệu video 20 sau đó có thể xác định xem có nhiều LTRP hơn cho hình ảnh hiện thời (162) hay không. Khi có nhiều LTRP hơn cho hình ảnh hiện

thời (nhánh “CÓ” của 162), bộ mã hóa dữ liệu video 20 có thể tiếp tục mã hóa các LSB của các giá trị POC cho các LTRP còn lại, và, khi cần thiết, MSB sử dụng mã hóa vi phân như được thảo luận trên đây. Tuy nhiên, khi không có nhiều LTRP hơn cho hình ảnh hiện thời (nhánh “KHÔNG” của 162), bộ mã hóa dữ liệu video 20 có thể xuất dữ liệu được mã hóa (164), dữ liệu này có thể bao gồm các khối được mã hóa cho hình ảnh hiện thời, các giá trị POC được mã hóa cho hình ảnh tham chiếu dài hạn cho hình ảnh hiện thời (bao gồm các LSB và, khi các LSB không phải là duy nhất, các giá trị khác nhau cho MSB), và tương tự.

Theo phương thức này, phương pháp trên Fig.5 biểu diễn ví dụ về phương pháp bao gồm các bước: mã hóa ít nhất phần thứ nhất của hình ảnh hiện thời của dữ liệu video dựa vào hình ảnh tham chiếu dài hạn thứ nhất và ít nhất phần thứ hai của hình ảnh hiện thời dựa vào hình ảnh tham chiếu dài hạn thứ hai, mã hóa giá trị thứ nhất biểu diễn độ chênh lệch giữa giá trị các bit có giá trị lớn nhất (MSB) cơ sở của giá trị số đếm thứ tự hình ảnh (POC) của hình ảnh hiện thời của dữ liệu video và giá trị MSBs thứ nhất của giá trị POC thứ nhất của hình ảnh tham chiếu dài hạn thứ nhất của dữ liệu video, và mã hóa giá trị thứ hai biểu diễn độ chênh lệch giữa giá trị MSBs thứ hai của giá trị POC thứ hai của hình ảnh tham chiếu dài hạn thứ hai của dữ liệu video và giá trị MSBs thứ nhất, trong đó giá trị POC thứ nhất và giá trị POC thứ hai có các giá trị các bit có giá trị nhỏ nhất (LSB) khác nhau.

Fig.6 là lưu đồ minh họa phương pháp làm ví dụ để giải mã các giá trị POC cho các hình ảnh tham chiếu dài hạn theo các kỹ thuật của sáng chế. Để giải thích, phương pháp trên Fig.6 được giải thích liên quan đến bộ giải mã dữ liệu video 30 trên Fig.1 và Fig.3 và các thành phần của nó. Tuy nhiên, nên hiểu rằng các thiết bị giải mã dữ liệu video khác có thể được tạo cấu hình để thực hiện phương pháp trên Fig.6.

Ban đầu, bộ giải mã dữ liệu video 30 có thể giải mã các MSB và các LSB của giá trị POC cho hình ảnh tham chiếu dài hạn (LTRP) thứ nhất của hình ảnh hiện thời (200). Ví dụ, bộ giải mã dữ liệu video 30 có thể giải mã giá trị chênh lệch biểu diễn độ chênh lệch giữa các MSB của giá trị POC cho LTRP thứ nhất và các MSB của giá trị POC cho hình ảnh hiện thời. Bộ giải mã dữ liệu video 30 sau đó có thể bổ sung giá trị chênh lệch này vào các MSB của giá trị POC cho hình ảnh hiện thời, để thu được các MSB của LTRP thứ nhất. Bộ giải mã dữ liệu video 30 cũng có thể giải mã các LSB của giá trị POC cho LTRP thứ nhất và nối các MSB và các LSB của giá trị POC

thứ nhất cho LTRP thứ nhất để tái tạo giá trị POC cho LTRP. Bộ giải mã dữ liệu video 30 có thể cũng giải mã các LSB của giá trị POC chỉ cho LTRP thứ nhất này, khi các MSB không được truyền, và có thể sử dụng chỉ các giá trị LSB để xác định xem nếu không có các hình ảnh tham chiếu khác với cùng giá trị LSB POC.

Bộ giải mã dữ liệu video 30 sau đó có thể giải mã các LSB của giá trị POC cho LTRP tiếp theo (202). Trong ví dụ trên Fig.6, bộ giải mã dữ liệu video 30 xác định xem các LSB của giá trị POC cho LTRP tiếp theo có bằng các LSB của LTRP khác cho hình ảnh hiện thời (204) hay không. Trong các ví dụ khác, ví dụ, như được giải thích cho bảng 1 trên đây, bộ giải mã dữ liệu video 30 có thể xác định xem giá trị có chỉ báo rằng các MSB được báo hiệu cho giá trị POC của LTRP tiếp theo hay không. Trong mọi trường hợp, bộ giải mã dữ liệu video 30 có thể xác định xem các MSB có được báo hiệu cho giá trị POC của LTRP tiếp theo hay không. Khi các MSB được báo hiệu cho giá trị POC của LTRP tiếp theo (ví dụ, khi các LSB của giá trị POC cho LTRP tiếp theo bằng các LSB của hình ảnh tham chiếu khác (nhánh “CÓ” của 204)), bộ giải mã dữ liệu video 30 có thể giải mã giá trị chênh lệch của các MSB so với các MSB của giá trị POC cho LTRP trước đó (206). Nghĩa là, bộ giải mã dữ liệu video 30 có thể giải mã giá trị chênh lệch được báo hiệu cho LTRP i , và bổ sung giá trị chênh lệch này vào các MSB của giá trị POC cho LTRP j trong đó LTRP j là LTRP trước đó mà thông tin MSB được giải mã. Bộ giải mã dữ liệu video 30 có thể còn tập hợp giá trị POC của LTRP tiếp theo bằng cách nối các MSB (cho dù được tính toán hay được xác định ngầm) với các LSB cho LTRP tiếp theo.

Sau khi xác định rằng các LSB của giá trị POC cho LTRP tiếp theo không bằng các LSB của hình ảnh tham chiếu khác (nhánh “KHÔNG” của 204), hoặc sau khi giải mã giá trị chênh lệch cho các MSB của LTRP tiếp theo, bộ giải mã dữ liệu video 30 có thể xác định xem có nhiều LTRP hơn cho hình ảnh hiện thời (210) hay không. Nếu có nhiều LTRP hơn cho hình ảnh hiện thời (nhánh “CÓ” của 210), bộ giải mã dữ liệu video 30 có thể tiếp tục giải mã các giá trị POC cho LTRP sau đó, ví dụ, như được thảo luận trên đây. Khi không có nhiều LTRP hơn cho hình ảnh hiện thời (nhánh “KHÔNG” của 210), bộ giải mã dữ liệu video 30 có thể tập hợp một hoặc nhiều danh mục hình ảnh tham chiếu bao gồm các LTRP được báo hiệu (212). Bộ giải mã dữ liệu video 30 sau đó có thể giải mã hình ảnh hiện thời bằng cách sử dụng (các) danh mục hình ảnh tham chiếu này (214).

Ví dụ, khi khối được dự báo liên cấu trúc, khối này có thể bao gồm chỉ báo về danh mục hình ảnh tham chiếu, và chỉ mục tham chiếu trong danh mục hình ảnh tham chiếu để chỉ báo hình ảnh tham chiếu cho khối này. Khối này còn có thể bao gồm dữ liệu để khôi phục vector chuyển động cho khối này. Do đó, bộ giải mã dữ liệu video 30 có thể tìm kiếm khối tham chiếu từ hình ảnh tham chiếu bằng cách sử dụng vector chuyển động. Bộ giải mã dữ liệu video 30 còn có thể giải mã giá trị sai số cho khối này, ví dụ, bằng cách giải mã, lượng tử hóa ngược, và biến đổi ngược các hệ số biến đổi cho khối này. Bộ giải mã dữ liệu video 30 sau đó có thể kết hợp khối tham chiếu với khối dự được khôi phục để giải mã khối này.

Theo phương thức này, phương pháp trên Fig.6 biểu diễn ví dụ về phương pháp bao gồm các bước: giải mã giá trị thứ nhất biểu diễn độ chênh lệch giữa giá trị có giá trị bit có giá trị lớn nhất (MSB) cơ sở của giá trị số đếm thứ tự hình ảnh (POC) của hình ảnh hiện thời của dữ liệu video và giá trị MSBs thứ nhất của giá trị POC thứ nhất của hình ảnh tham chiếu dài hạn thứ nhất của dữ liệu video, giải mã giá trị thứ hai biểu diễn độ chênh lệch giữa giá trị MSBs thứ hai của giá trị POC thứ hai của hình ảnh tham chiếu dài hạn thứ hai của dữ liệu video và giá trị MSBs thứ nhất, trong đó giá trị POC thứ nhất và giá trị POC thứ hai có các giá trị bit có giá trị nhỏ nhất (LSB) khác nhau, và giải mã ít nhất một phần hình ảnh hiện thời của dữ liệu video dựa vào ít nhất một trong số hình ảnh tham chiếu dài hạn thứ nhất sử dụng giá trị thứ nhất và hình ảnh tham chiếu dài hạn thứ hai sử dụng giá trị thứ nhất và giá trị thứ hai.

Cần lưu ý rằng tùy thuộc vào ví dụ, các hoạt động hoặc sự kiện nhất định của kỹ thuật bất kỳ trong số các kỹ thuật được mô tả ở đây có thể được thực hiện theo chuỗi khác nhau, có thể được bổ sung, hợp nhất, hoặc tách ra hoàn toàn (ví dụ, không phải tất cả các hoạt động hoặc sự kiện được mô tả là cần thiết để thực hiện các kỹ thuật này). Hơn nữa, trong các ví dụ nhất định, các hoạt động hoặc các sự kiện có thể được thực hiện đồng thời, ví dụ, thông qua quy trình xử lý đa xâu chuỗi, xử lý ngắt quãng, hoặc nhiều bộ xử lý, hơn là thực hiện tuần tự.

Trong một hoặc nhiều ví dụ, các chức năng đã mô tả có thể được thực hiện bằng phần cứng, phần mềm, phần sụn, hoặc dạng kết hợp bất kỳ của chúng. Nếu được thực hiện bằng phần mềm, thì các chức năng này có thể được lưu trữ hoặc được truyền dưới dạng một hoặc nhiều lệnh hoặc mã trên vật ghi đọc được bằng máy tính và được thi hành bằng bộ phận xử lý dựa trên phần cứng. Vật ghi đọc được bằng máy

tính có thể là phương tiện lưu trữ đọc được bằng máy tính, tương ứng với phương tiện hữu hình như phương tiện lưu trữ dữ liệu, hoặc phương tiện truyền thông bao gồm mọi phương tiện hỗ trợ 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 thuộc loại không khả biến hoặc (2) phương tiện truyền thông như tín hiệu hoặc sóng mang. Phương tiện lưu trữ dữ liệu có thể là mọi loại phương tiện có sẵn có thể truy nhập được bằng một hoặc nhiều máy tính hoặc một hay nhiều bộ xử lý để tìm kiếm các lệnh, mã và/hoặc cấu trúc dữ liệu để thực hiện các kỹ thuật nêu trong sáng chế. Sản phẩm chương trình máy tính có thể bao gồm vật ghi đọc được bằng máy tính.

Ví dụ, phương tiện lưu trữ đọc được bằng máy tính có thể bao gồm bộ nhớ truy nhập ngẫu nhiên (Random Access Memory - RAM), bộ nhớ chỉ đọc (Read Only Memory - ROM), bộ nhớ chỉ đọc lập trình được xóa được bằng điện (Electrically Erasable Programmable Read Only Memory - EEPROM)), đĩa compac-bộ nhớ chỉ đọc (Compact Disc-Read Only Memory - CD-ROM) hoặc đĩa quang khác, đĩa từ hoặc thiết bị nhớ từ tính khác, thiết bị nhớ truy nhập nhanh, hoặc mọi phương tiện khác có thể dùng để lưu trữ mã chương trình mong muốn dưới dạng lệnh hoặc cấu trúc dữ liệu và có thể truy nhập được bằng máy tính, nhưng không chỉ giới hạn ở đó. Ngoài ra, mọi loại kết nối được gọi theo cách thích hợp là phương tiện đọc được bằng máy tính. Ví dụ, nếu các lệnh được truyền từ trang web, máy chủ hoặc nguồn từ xa khác bằng cách sử dụng cáp đồng trục, cáp sợi quang, cáp xoắn đôi, đường thuê bao số (Digital Subscriber Line - DSL), hoặc sử dụng công nghệ không dây như hồng ngoại, vô tuyến và vi ba, thì cáp đồng trục, cáp sợi quang, cáp xoắn đôi, DSL, hoặc công nghệ không dây như hồng ngoại, vô tuyến và vi ba đó cũng nằm trong định nghĩa phương tiện. Tuy nhiên, cần phải hiểu rằng, phương tiện lưu trữ đọc được bằng máy tính và phương tiện lưu trữ dữ liệu không bao gồm các loại kết nối, sóng mang, tín hiệu hoặc phương tiện khả biến khác, mà chỉ bao gồm phương tiện lưu trữ hữu hình, không khả biến. Đĩa từ và đĩa quang, như được sử dụng ở đây, bao gồm đĩa compac (Compact Disc - CD), đĩa laser, đĩa quang, đĩa số đa năng (Digital Versatile Disc - DVD), đĩa mềm và đĩa blu-ray, trong đó đĩa từ thường tái tạo dữ liệu bằng phương pháp từ tính, còn đĩa quang thì tái tạo dữ liệu bằng phương pháp quang học

sử dụng laze. Dạng kết hợp của các loại phương tiện nêu trên cũng được coi là nằm trong phạm vi vật ghi đọc được bằng máy tính.

Các lệnh có thể được thi hành 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 cửa lập trình được bằng trường (FPGA), hoặc mạch logic rời rạc hoặc mạch tích hợp tương đương khác. Do đó, thuật ngữ “bộ xử lý”, như được sử dụng trong sáng chế, có thể dùng để chỉ mọi cấu trúc nêu trên hoặc mọi cấu trúc khác phù hợp để thực hiện các kỹ thuật nêu trong sáng chế. Ngoài ra, theo một số khía cạnh, chức năng nêu trong sáng chế có thể được thực hiện bằng môđun phần cứng và/hoặc phần mềm chuyên dụng được tạo cấu hình để mã hóa và giải mã, hoặc được kết hợp thành một bộ mã hóa-giải mã kết hợp. Đồng thời, các kỹ thuật có thể được thực hiện hoàn toàn trong một hoặc nhiều mạch hoặc phần tử logic.

Kỹ thuật theo sáng chế có thể được áp dụng trong rất nhiều thiết bị, trong đó có tổ hợp thiết bị không dây, mạch tích hợp (Integrated Circuit - IC) hoặc bộ IC (ví dụ, bộ chip). Các thành phần, môđun, hoặc bộ phận được mô tả trong sáng chế để nhấn mạnh các khía cạnh chức năng của các thiết bị được tạo cấu hình để thực hiện các kỹ thuật đã mô tả, nhưng không nhất thiết phải được thực hiện bằng các bộ phận phần cứng khác nhau. Thực ra, như đã nêu trên, các bộ phận khác nhau có thể được kết hợp lại thành một bộ phận phần cứng mã hóa-giải mã hoặc được thực hiện bởi một tập hợp các bộ phận phần cứng tương tác với nhau, có một hoặc nhiều bộ xử lý như đã nêu trên, kết hợp với phần mềm và/hoặc phần sụn phù hợp.

Nhiều phương án làm ví dụ đã được mô tả. Các phương án làm ví dụ này và các phương án khác làm ví dụ đều nằm trong phạm vi của sáng chế như được xác định theo các điểm yêu cầu bảo hộ dưới đây.

Yêu cầu bảo hộ

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

giải mã giá trị thứ nhất biểu diễn độ chênh lệch giữa giá trị các bit có giá trị lớn nhất (most significant bits - MSBs) cơ sở của giá trị số đếm thứ tự hình ảnh (picture order count - POC) của hình ảnh hiện thời của dữ liệu video và giá trị MSBs thứ nhất của giá trị POC thứ nhất của hình ảnh tham chiếu dài hạn thứ nhất của dữ liệu video;

giải mã giá trị thứ hai biểu diễn độ chênh lệch giữa giá trị MSBs thứ hai của giá trị POC thứ hai của hình ảnh tham chiếu dài hạn thứ hai của dữ liệu video và giá trị MSBs thứ nhất, trong đó giá trị POC thứ nhất và giá trị POC thứ hai có các giá trị các bit có giá trị nhỏ nhất (least significant bits - LSBs) khác nhau và giá trị MSBs thứ nhất khác với giá trị MSBs thứ hai; và

giải mã ít nhất một phần của hình ảnh hiện thời của dữ liệu video liên quan đến ít nhất một trong số hình ảnh tham chiếu dài hạn thứ nhất sử dụng giá trị thứ nhất hoặc hình ảnh tham chiếu dài hạn thứ hai sử dụng giá trị thứ nhất và giá trị thứ hai.

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

tính toán giá trị chu kỳ MSB thứ nhất cho hình ảnh tham chiếu dài hạn thứ nhất là DeltaPocMSBCycleLt bằng cách sử dụng giá trị thứ nhất; và

tính toán giá trị chu kỳ MSB thứ hai cho hình ảnh tham chiếu dài hạn thứ hai là DeltaPocMSBCycleLt, trong đó bước tính toán giá trị chu kỳ MSB thứ hai bao gồm tính toán DeltaPocMSBCycleLt dựa trên DeltaPocMSBCycleLt và giá trị thứ hai.

3. Phương pháp theo điểm 1, trong đó hình ảnh tham chiếu dài hạn thứ nhất tương ứng với chỉ mục j, trong đó hình ảnh tham chiếu dài hạn thứ hai tương ứng với chỉ mục i, trong đó giá trị thứ nhất bao gồm delta_poc_msb_cycle_lt_plus1, trong đó giá trị thứ hai bao gồm delta_poc_msb_cycle_lt_plus1, và trong đó j nằm trong khoảng từ 0 đến i-1, bao gồm cả giá trị đầu mút.

4. Phương pháp theo điểm 1, trong đó ít nhất một trong số giá trị thứ nhất hoặc giá trị

thứ hai bao gồm giá trị nguyên âm.

5. Phương pháp theo điểm 1, trong đó phương pháp này còn bao gồm bước xác định ít nhất một trong số giá trị POC thứ nhất hoặc giá trị POC thứ hai dựa ít nhất một phần vào số hình ảnh tham chiếu được đánh dấu là hình ảnh tham chiếu ngắn hạn cho hình ảnh hiện thời.

6. Thiết bị giải mã dữ liệu video, thiết bị này bao gồm:

bộ nhớ được tạo cấu hình để lưu trữ dữ liệu video; và

bộ giải mã dữ liệu video được tạo cấu hình để:

giải mã giá trị thứ nhất biểu diễn độ chênh lệch giữa giá trị các bit có giá trị lớn nhất (most significant bits - MSBs) cơ sở của giá trị số đếm thứ tự hình ảnh (picture order count - POC) của hình ảnh hiện thời của dữ liệu video và giá trị MSBs thứ nhất của giá trị POC thứ nhất của hình ảnh tham chiếu dài hạn thứ nhất của dữ liệu video,

giải mã giá trị thứ hai biểu diễn độ chênh lệch giữa giá trị MSBs thứ hai của giá trị POC thứ hai của hình ảnh tham chiếu dài hạn thứ hai của dữ liệu video và giá trị MSBs thứ nhất, trong đó giá trị POC thứ nhất và giá trị POC thứ hai có các giá trị bit có giá trị nhỏ nhất (least significant bits - LSBs) khác nhau và giá trị MSBs thứ nhất khác với giá trị MSBs thứ hai, và

giải mã ít nhất một phần của hình ảnh hiện thời của dữ liệu video liên quan đến ít nhất một trong số hình ảnh tham chiếu dài hạn thứ nhất sử dụng giá trị thứ nhất hoặc hình ảnh tham chiếu dài hạn thứ hai sử dụng giá trị thứ nhất và giá trị thứ hai.

7. Thiết bị theo điểm 6, trong đó bộ giải mã dữ liệu video này còn được tạo cấu hình để tính toán giá trị chu kỳ MSB thứ nhất cho hình ảnh tham chiếu dài hạn thứ nhất là $\Delta\text{PocMSBCycleLt}[i - 1]$ bằng cách sử dụng giá trị thứ nhất, và tính toán giá trị chu kỳ MSB thứ hai cho hình ảnh tham chiếu dài hạn thứ hai là $\Delta\text{PocMSBCycleLt}$, trong đó việc tính toán giá trị chu kỳ MSB thứ hai bao gồm tính toán $\Delta\text{PocMSBCycleLt}$ dựa trên $\Delta\text{PocMSBCycleLt}$ và giá trị thứ hai.

8. Thiết bị theo điểm 6, trong đó hình ảnh tham chiếu dài hạn thứ nhất tương ứng với chỉ mục $i-1$, trong đó hình ảnh tham chiếu dài hạn thứ hai tương ứng với chỉ mục i , trong đó giá trị thứ nhất bao gồm $\text{delta_poc_msb_cycle_lt_plus1}$, và trong đó giá trị thứ hai bao gồm $\text{delta_poc_msb_cycle_lt_plus1}$.

9. Thiết bị theo điểm 6, trong đó ít nhất một trong số giá trị thứ nhất hoặc giá trị thứ hai bao gồm giá trị nguyên âm.

10. Thiết bị theo điểm 6, trong đó bộ giải mã dữ liệu video này còn được tạo cấu hình để xác định ít nhất một trong số giá trị POC thứ nhất hoặc giá trị POC thứ hai dựa ít nhất một phần vào số hình ảnh tham chiếu được đánh dấu là hình ảnh tham chiếu ngắn hạn cho hình ảnh hiện thời.

11. Thiết bị giải mã dữ liệu video, thiết bị này bao gồm:

phương tiện giải mã giá trị thứ nhất biểu diễn độ chênh lệch giữa giá trị các bit có giá trị lớn nhất (most significant bits - MSBs) cơ sở của giá trị số đếm thứ tự hình ảnh (picture order count - POC) của hình ảnh hiện thời của dữ liệu video và giá trị MSBs thứ nhất của giá trị POC thứ nhất của hình ảnh tham chiếu dài hạn thứ nhất của dữ liệu video;

phương tiện giải mã giá trị thứ hai biểu diễn độ chênh lệch giữa giá trị MSBs thứ hai của giá trị POC thứ hai của hình ảnh tham chiếu dài hạn thứ hai của dữ liệu video và giá trị MSBs thứ nhất, trong đó giá trị POC thứ nhất và giá trị POC thứ hai có các giá trị các bit có giá trị nhỏ nhất (least significant bits - LSBs) khác nhau và giá trị MSBs thứ nhất khác với giá trị MSBs thứ hai; và

phương tiện giải mã ít nhất một phần của hình ảnh hiện thời của dữ liệu video liên quan đến ít nhất một trong số hình ảnh tham chiếu dài hạn thứ nhất sử dụng giá trị thứ nhất hoặc hình ảnh tham chiếu dài hạn thứ hai sử dụng giá trị thứ nhất và giá trị thứ hai.

12. Thiết bị theo điểm 11, trong đó thiết bị này còn bao gồm:

phương tiện tính toán giá trị chu kỳ MSB thứ nhất cho hình ảnh tham chiếu dài hạn thứ nhất là $\text{DeltaPocMSBCycleLt}$ bằng cách sử dụng giá trị thứ nhất; và

phương tiện tính toán giá trị chu kỳ MSB thứ hai cho hình ảnh tham chiếu dài hạn thứ hai là DeltaPocMSBCycleLt, trong đó việc tính toán giá trị chu kỳ MSB thứ hai bao gồm tính toán DeltaPocMSBCycleLt dựa trên DeltaPocMSBCycleLt và giá trị thứ hai.

13. Thiết bị theo điểm 11, trong đó hình ảnh tham chiếu dài hạn thứ nhất tương ứng với chỉ mục j, trong đó hình ảnh tham chiếu dài hạn thứ hai tương ứng với chỉ mục i, trong đó giá trị thứ nhất bao gồm delta_poc_msb_cycle_lt_plus1, trong đó giá trị thứ hai bao gồm delta_poc_msb_cycle_lt_plus1, và trong đó j nằm trong khoảng từ 0 đến i-1, bao gồm cả giá trị đầu mút.

14. Thiết bị theo điểm 11, trong đó ít nhất một trong số giá trị thứ nhất hoặc giá trị thứ hai bao gồm giá trị nguyên âm.

15. Thiết bị theo điểm 11, thiết bị này còn bao gồm phương tiện xác định ít nhất một trong số giá trị POC thứ nhất hoặc giá trị POC thứ hai dựa ít nhất một phần vào số hình ảnh tham chiếu được đánh dấu là hình ảnh tham chiếu ngắn hạn cho hình ảnh hiện thời.

16. 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:

giải mã giá trị thứ nhất biểu diễn độ chênh lệch giữa giá trị các bit có giá trị lớn nhất (most significant bits - MSBs) cơ sở của giá trị số đếm thứ tự hình ảnh (picture order count - POC) của hình ảnh hiện thời của dữ liệu video và giá trị MSBs thứ nhất của giá trị POC thứ nhất của hình ảnh tham chiếu dài hạn thứ nhất của dữ liệu video;

giải mã giá trị thứ hai biểu diễn độ chênh lệch giữa giá trị MSBs thứ hai của giá trị POC thứ hai của hình ảnh tham chiếu dài hạn thứ hai của dữ liệu video và giá trị MSBs thứ nhất, trong đó giá trị POC thứ nhất và giá trị POC thứ hai có các giá trị các bit có giá trị nhỏ nhất (least significant bits - LSBs) khác nhau và giá trị MSBs thứ nhất khác với giá trị MSBs thứ hai; và

giải mã ít nhất một phần của hình ảnh hiện thời của dữ liệu video dựa vào ít

nhất một trong số hình ảnh tham chiếu dài hạn thứ nhất sử dụng giá trị thứ nhất hoặc hình ảnh tham chiếu dài hạn thứ hai sử dụng giá trị thứ nhất và giá trị thứ hai.

17. Vật ghi bất biến đọc được bằng máy tính theo điểm 16, trong đó vật ghi này còn bao gồm các lệnh khiến cho bộ xử lý:

tính toán giá trị chu kỳ MSB thứ nhất cho hình ảnh tham chiếu dài hạn thứ nhất là DeltaPocMSBCycleLt bằng cách sử dụng giá trị thứ nhất; và

tính toán giá trị chu kỳ MSB thứ hai cho hình ảnh tham chiếu dài hạn thứ hai là DeltaPocMSBCycleLt, trong đó việc tính toán giá trị chu kỳ MSB thứ hai bao gồm tính toán DeltaPocMSBCycleLt dựa trên DeltaPocMSBCycleLt và giá trị thứ hai.

18. Vật ghi bất biến đọc được bằng máy tính theo điểm 16, trong đó hình ảnh tham chiếu dài hạn thứ nhất tương ứng với chỉ mục j, trong đó hình ảnh tham chiếu dài hạn thứ hai tương ứng với chỉ mục i, trong đó giá trị thứ nhất bao gồm delta_poc_msb_cycle_lt_plus1, trong đó giá trị thứ hai bao gồm delta_poc_msb_cycle_lt_plus1, và trong đó j nằm trong khoảng từ 0 đến i-1, bao gồm cả giá trị đầu mút.

19. Vật ghi bất biến đọc được bằng máy tính theo điểm 16, trong đó ít nhất một trong số giá trị thứ nhất hoặc giá trị thứ hai bao gồm giá trị nguyên âm.

20. Vật ghi bất biến đọc được bằng máy tính theo điểm 16, trong đó vật ghi này còn bao gồm các lệnh khiến cho bộ xử lý xác định ít nhất một trong số giá trị POC thứ nhất hoặc giá trị POC thứ hai dựa ít nhất một phần vào số hình ảnh tham chiếu được đánh dấu là hình ảnh tham chiếu ngắn hạn cho hình ảnh hiện thời.

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

mã hóa ít nhất phần thứ nhất của hình ảnh hiện thời của dữ liệu video liên quan đến hình ảnh tham chiếu dài hạn thứ nhất và ít nhất phần thứ hai của hình ảnh hiện thời liên quan đến hình ảnh tham chiếu dài hạn thứ hai;

mã hóa giá trị thứ nhất biểu diễn độ chênh lệch giữa giá trị các bit có giá trị lớn nhất (most significant bits - MSBs) cơ sở của giá trị số đếm thứ tự hình ảnh

(picture order count - POC) của hình ảnh hiện thời của dữ liệu video và giá trị MSBs thứ nhất của giá trị POC thứ nhất của hình ảnh tham chiếu dài hạn thứ nhất của dữ liệu video; và

mã hóa giá trị thứ hai biểu diễn độ chênh lệch giữa giá trị MSBs thứ hai của giá trị POC thứ hai của hình ảnh tham chiếu dài hạn thứ hai của dữ liệu video và giá trị MSBs thứ nhất, trong đó giá trị POC thứ nhất và giá trị POC thứ hai có các giá trị các bit có giá trị nhỏ nhất (least significant bits - LSBs) khác nhau và giá trị MSBs thứ nhất khác với giá trị MSBs thứ hai.

22. Phương pháp theo điểm 21, trong đó phương pháp này còn bao gồm các bước:

tính toán giá trị thứ nhất là độ chênh lệch giữa giá trị chu kỳ MSB thứ nhất cho hình ảnh tham chiếu dài hạn thứ nhất là $\Delta POC_{MSBCycleLt}$ và giá trị MSBs cơ sở; và

tính toán giá trị thứ hai là độ chênh lệch giữa giá trị chu kỳ MSB thứ hai cho hình ảnh tham chiếu dài hạn thứ hai là $\Delta POC_{MSBCycleLt}$ và giá trị MSBs thứ nhất.

23. Phương pháp theo điểm 21, trong đó hình ảnh tham chiếu dài hạn thứ nhất tương ứng với chỉ mục j , trong đó hình ảnh tham chiếu dài hạn thứ hai tương ứng với chỉ mục i , trong đó giá trị thứ nhất bao gồm $\Delta poc_msb_cycle_lt_plus1$, trong đó giá trị thứ hai bao gồm $\Delta poc_msb_cycle_lt_plus1$, và trong đó j nằm trong khoảng từ 0 đến $i-1$, bao gồm cả giá trị đầu mút.

24. Phương pháp theo điểm 21, trong đó ít nhất một trong số giá trị thứ nhất hoặc giá trị thứ hai bao gồm giá trị nguyên âm.

25. Phương pháp theo điểm 21, trong đó bước mã hóa giá trị thứ nhất và mã hóa giá trị thứ hai bao gồm mã hóa giá trị thứ nhất và mã hóa giá trị thứ hai dựa ít nhất một phần vào số hình ảnh tham chiếu được đánh dấu là hình ảnh tham chiếu ngắn hạn cho hình ảnh hiện thời.

26. 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
bộ mã hóa dữ liệu video được tạo cấu hình để:

mã hóa ít nhất phần thứ nhất của hình ảnh hiện thời của dữ liệu video liên quan đến hình ảnh tham chiếu dài hạn thứ nhất và ít nhất phần thứ hai của hình ảnh hiện thời liên quan đến hình ảnh tham chiếu dài hạn thứ hai,

mã hóa giá trị thứ nhất biểu diễn độ chênh lệch giữa giá trị các bit có giá trị lớn nhất (most significant bits - MSBs) cơ sở của giá trị số đếm thứ tự hình ảnh (picture order count - POC) của hình ảnh hiện thời của dữ liệu video và giá trị MSBs thứ nhất của giá trị POC thứ nhất của hình ảnh tham chiếu dài hạn thứ nhất của dữ liệu video, và

mã hóa giá trị thứ hai biểu diễn độ chênh lệch giữa giá trị MSBs thứ hai của giá trị POC thứ hai của hình ảnh tham chiếu dài hạn thứ hai của dữ liệu video và giá trị MSBs thứ nhất, trong đó giá trị POC thứ nhất và giá trị POC thứ hai có các giá trị các bit có giá trị nhỏ nhất (least significant bits - LSBs) khác nhau và giá trị MSBs thứ nhất khác với giá trị MSBs thứ hai.

27. Thiết bị theo điểm 26, trong đó bộ mã hóa dữ liệu video được tạo cấu hình để tính toán giá trị thứ nhất là độ chênh lệch giữa giá trị chu kỳ MSB thứ nhất cho hình ảnh tham chiếu dài hạn thứ nhất là $\Delta POC_{MSB_CycleLt}$ và giá trị MSBs cơ sở, và tính toán giá trị thứ hai là độ chênh lệch giữa giá trị chu kỳ MSB thứ hai cho hình ảnh tham chiếu dài hạn thứ hai là $\Delta POC_{MSB_CycleLt}$ và giá trị MSBs thứ nhất.

28. Thiết bị theo điểm 26, trong đó hình ảnh tham chiếu dài hạn thứ nhất tương ứng với chỉ mục j , trong đó hình ảnh tham chiếu dài hạn thứ hai tương ứng với chỉ mục i , trong đó giá trị thứ nhất bao gồm $\Delta POC_{MSB_CycleLt_plus1}$, trong đó giá trị thứ hai bao gồm $\Delta POC_{MSB_CycleLt_plus1}$, và trong đó j nằm trong khoảng từ 0 đến $i-1$, bao gồm cả giá trị đầu mút.

29. Thiết bị theo điểm 26, trong đó ít nhất một trong số giá trị thứ nhất hoặc giá trị thứ hai bao gồm giá trị nguyên âm.

30. Thiết bị theo điểm 26, trong đó để mã hóa giá trị thứ nhất và giá trị thứ hai, bộ mã

hóa dữ liệu video được tạo cấu hình để mã hóa giá trị thứ nhất và giá trị thứ hai dựa ít nhất một phần vào số hình ảnh tham chiếu được đánh dấu là hình ảnh tham chiếu ngắn hạn cho hình ảnh hiện thời.

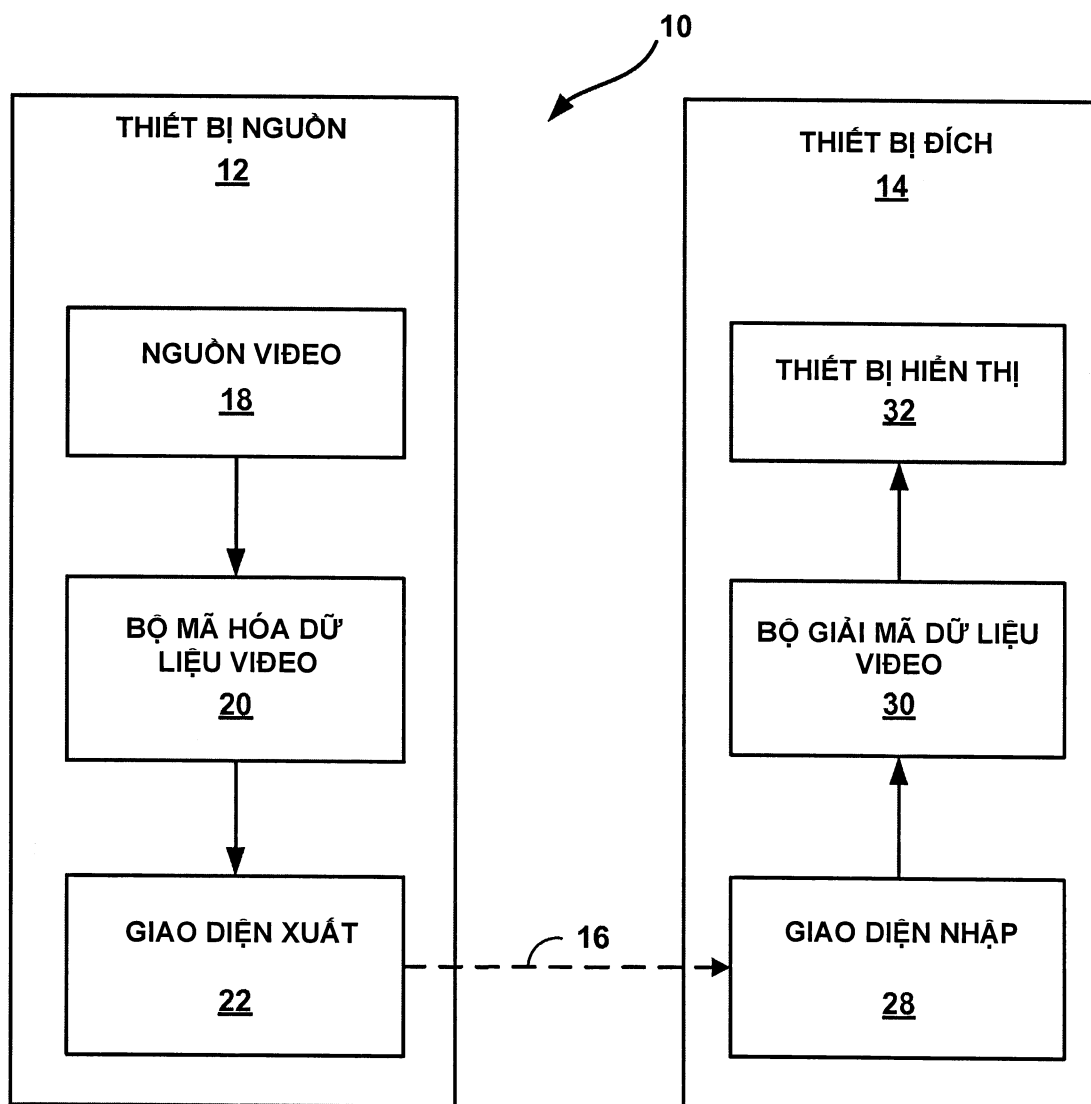


FIG. 1

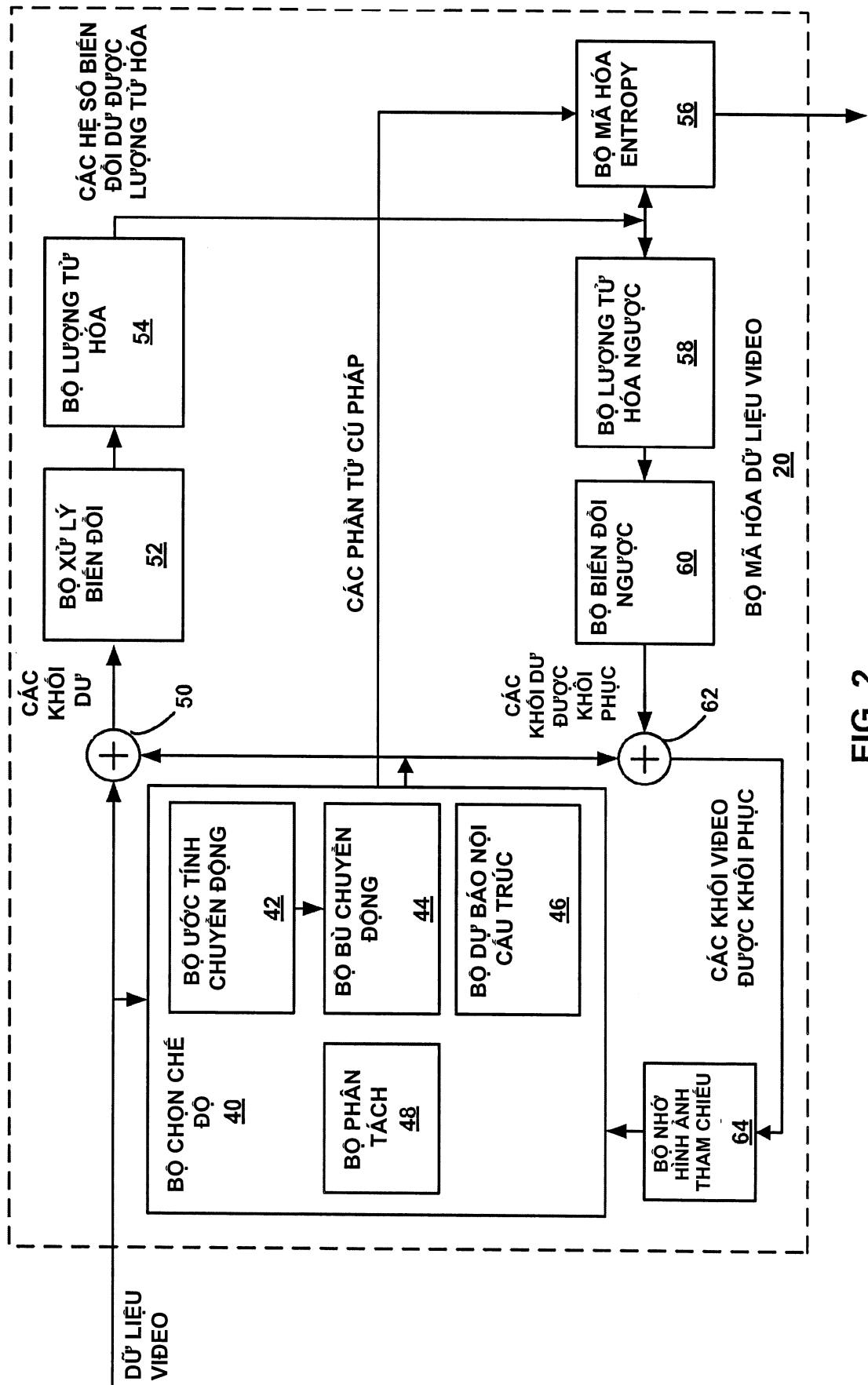


FIG. 2

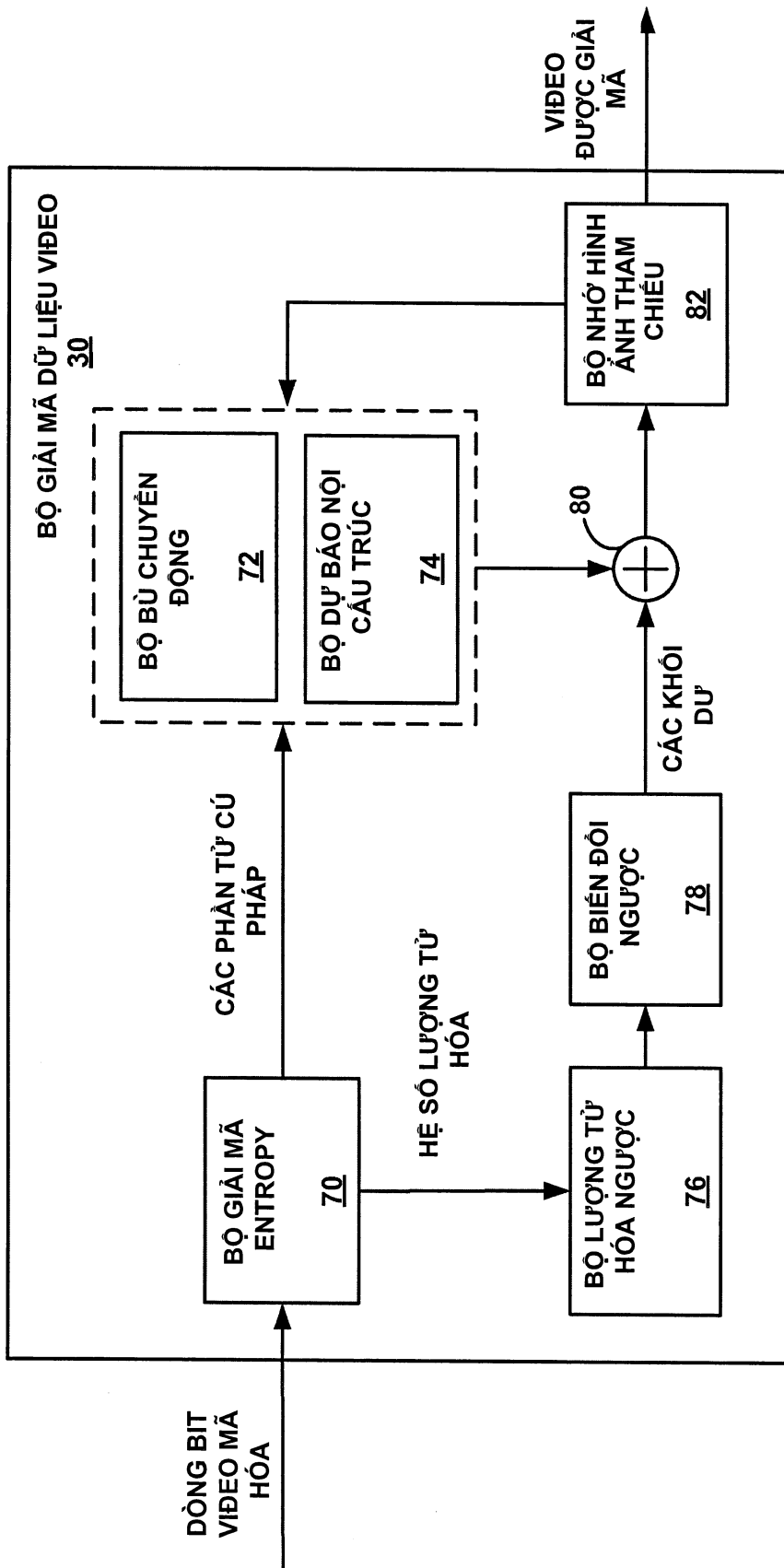


FIG. 3

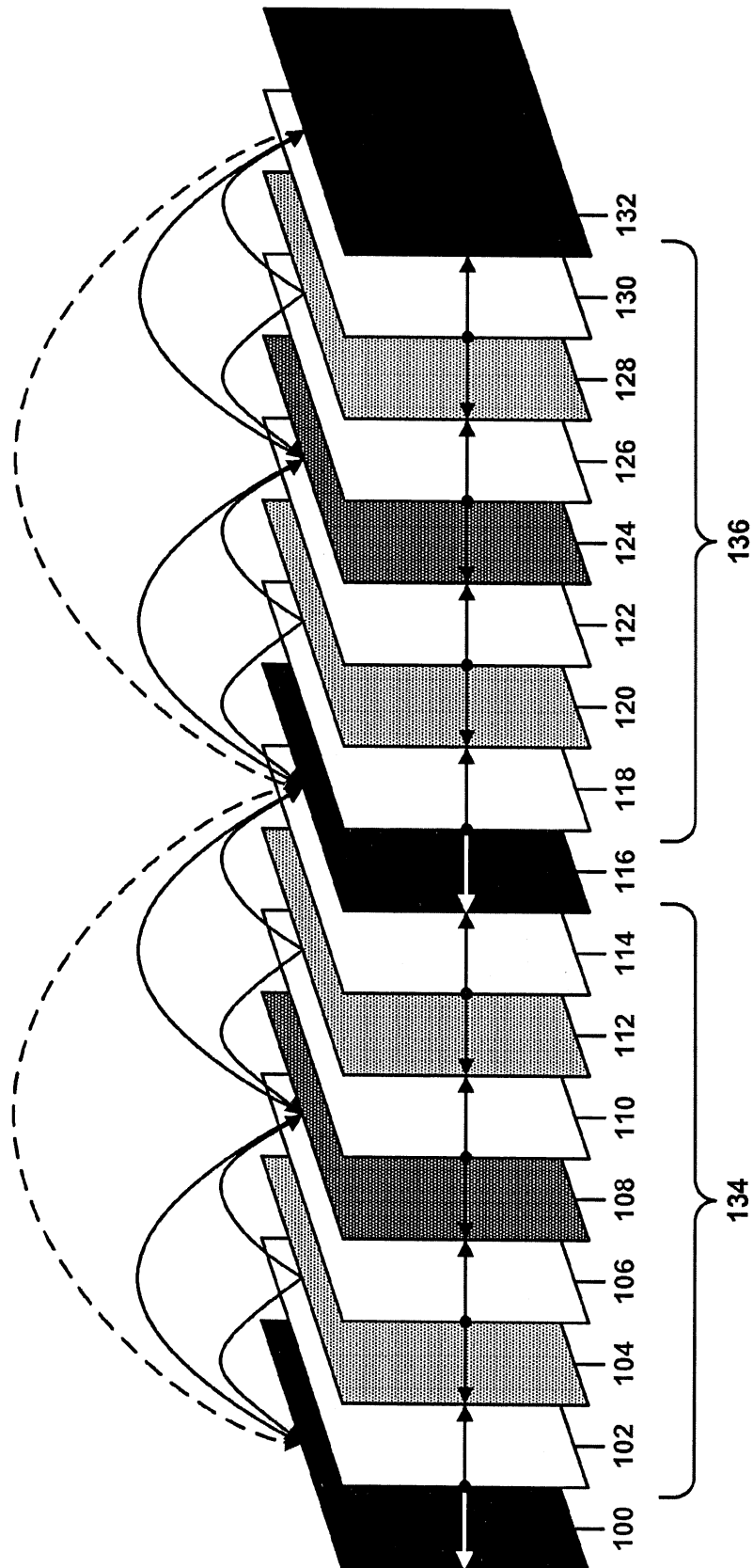


FIG. 4

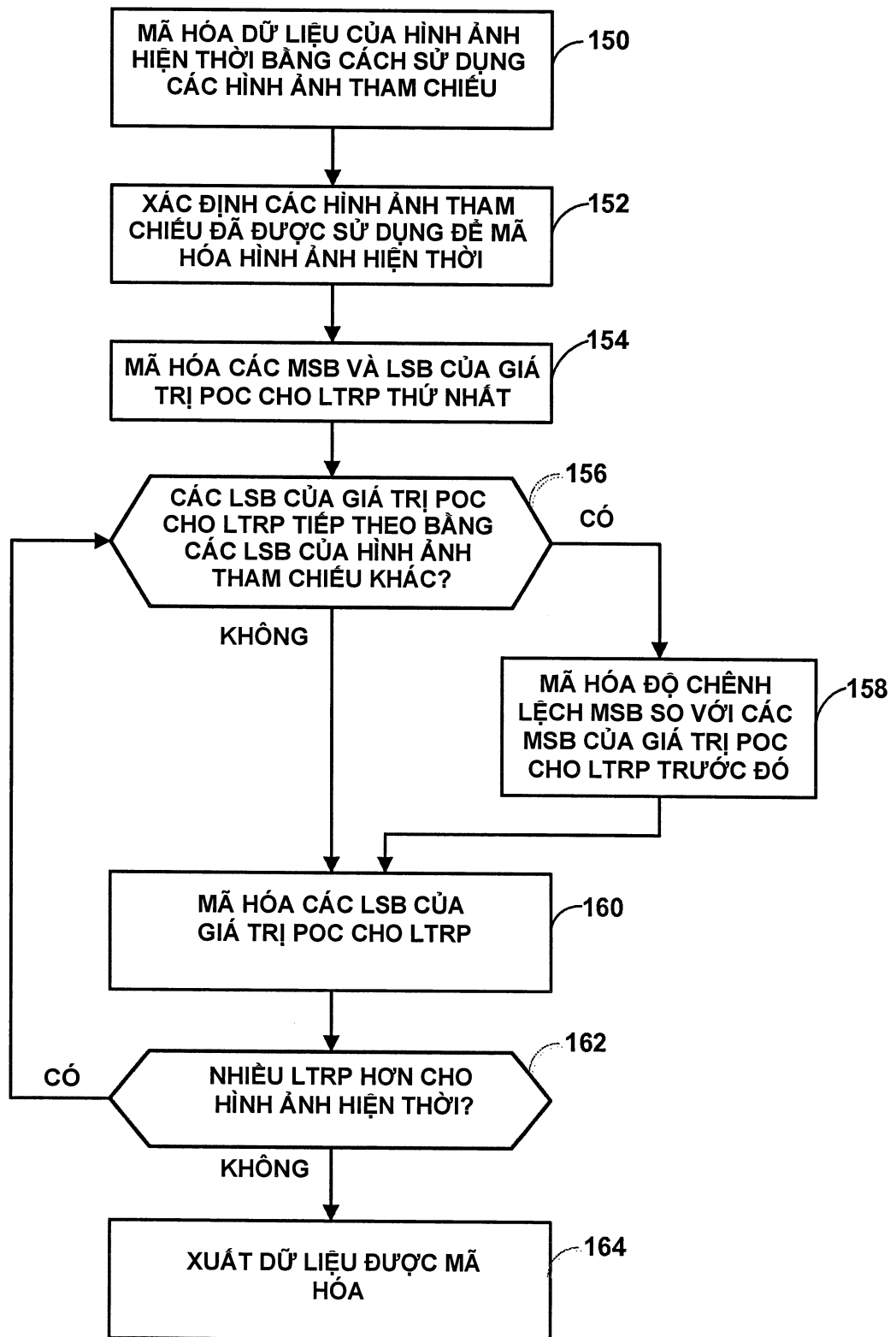


FIG. 5

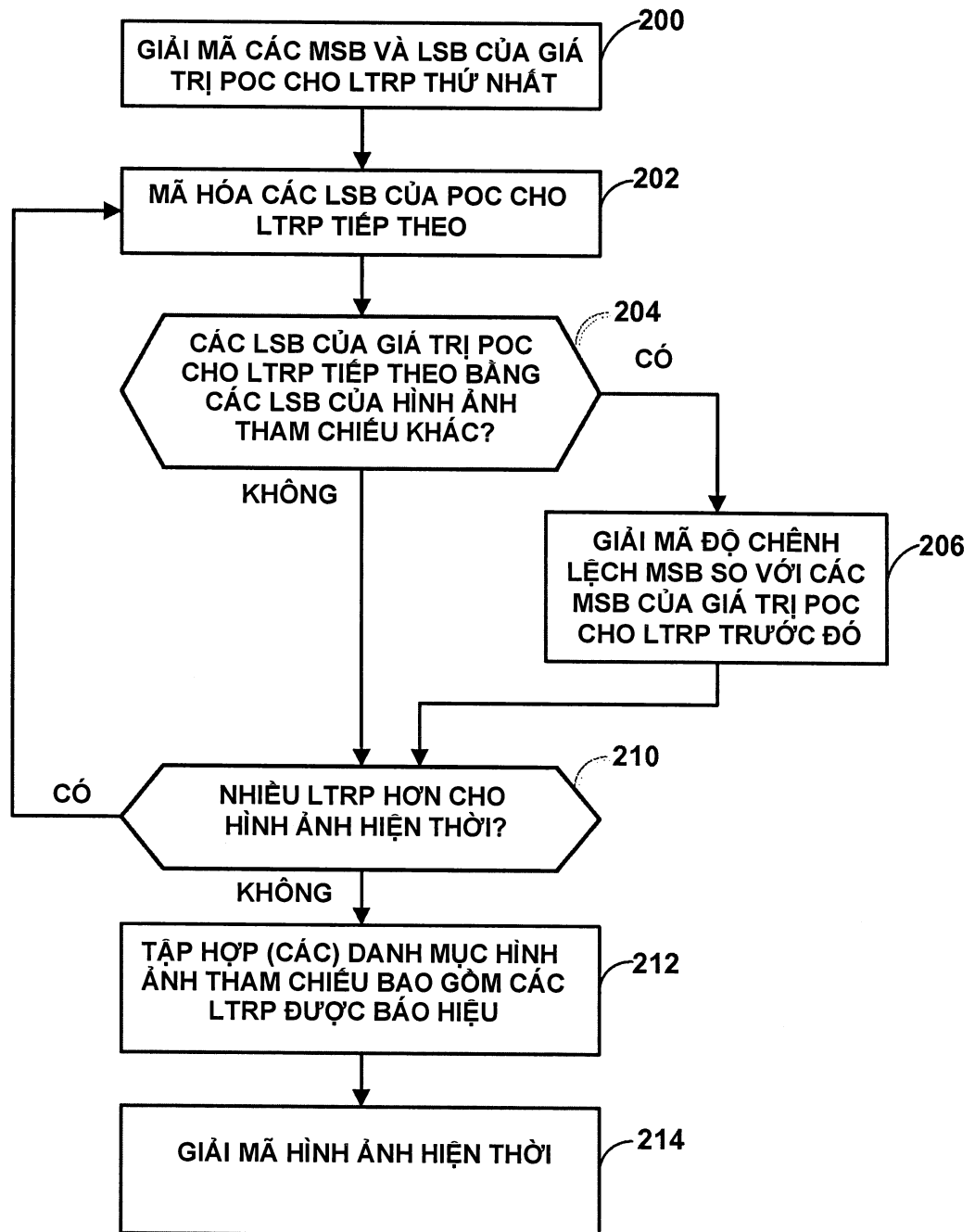


FIG. 6