



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

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

(11)



1-0041324

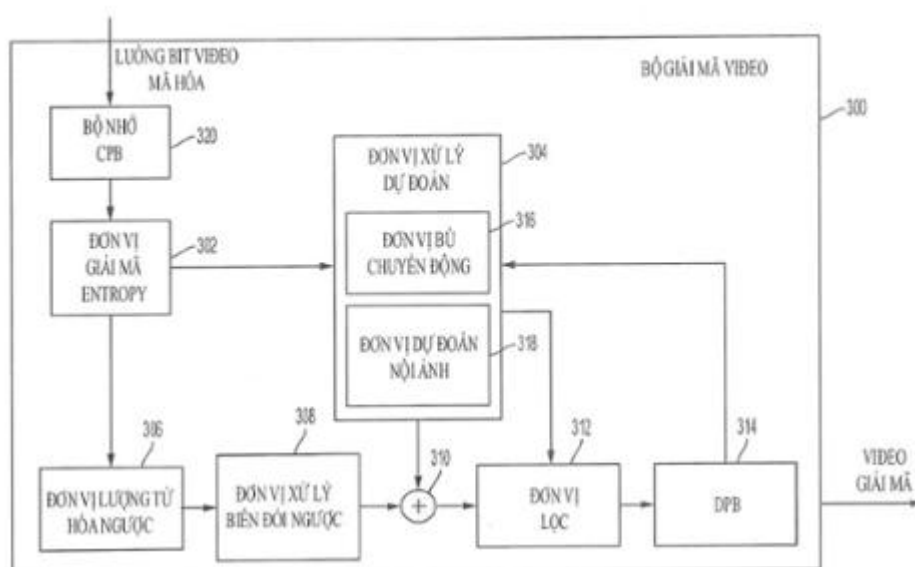
(51)^{2020.01} H04N 19/52; H04N 19/70

(13) B

- (21) 1-2020-07211 (22) 18/06/2019
(86) PCT/US2019/037722 18/06/2019 (87) WO 2019/246092 A1 26/12/2019
(30) 62/687,052 19/06/2018 US; 16/443,113 17/06/2019 US
(45) 25/10/2024 439 (43) 26/04/2021 397
(73) QUALCOMM INCORPORATED (US)
ATTN: International IP Administration, 5775 Morehouse Drive, San Diego, CA
92121-1714, United States of America
(72) HUANG, Han (CN); CHIEN, Wei-Jung (US); SEREGIN, Vadim (US);
KARCZEWICZ, Marta (US).
(74) Công ty TNHH Quốc tế D & N (D&N INTERNATIONAL CO.,LTD.)

(54) PHƯƠNG PHÁP, THIẾT BỊ GIẢI MÃ DỮ LIỆU VIDEO VÀ PHƯƠNG TIỆN LƯU TRỮ BẤT BIẾN ĐỌC ĐƯỢC BẰNG MÁY TÍNH

(57) Sáng chế đề cập đến phương pháp giải mã dữ liệu video bao gồm bước phân tích cú pháp cờ chuyển động của đơn vị dự đoán con từ dữ liệu video mã hóa nhận được, suy ra danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán con nếu cờ chuyển động của đơn vị dự đoán con hoạt động, suy ra danh sách ứng viên dự đoán chuyển động mức đơn vị dự đoán nếu cờ chuyển động của đơn vị dự đoán con không hoạt động, và giải mã dữ liệu video được mã hóa bằng cách sử dụng bộ dự đoán vector chuyển động đã lựa chọn. Sáng chế còn đề cập đến thiết bị giải mã dữ liệu video và phương tiện lưu trữ bất biến đọc được bằng máy tính.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến việc mã hóa video và giải mã video.

Tình trạng kỹ thuật của sáng chế

Các tính năng video số có thể được đưa vào áp dụng trong rất nhiều thiết bị bao gồm tivi số, hệ thống phát quảng bá trực tiếp số, hệ thống phát quảng bá không dây, thiết bị hỗ trợ 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, máy phát phương tiện số, thiết bị chơi trò chơi điện tử, bàn giao tiếp trò chơi điện tử, điện thoại di động hoặc điện thoại vô tuyến vệ tinh, thiết bị được gọi là “điện thoại thông minh”, thiết bị hội thảo truyền hình từ xa, thiết bị truyền trực tiếp video, và thiết bị tương tự. Các thiết bị video số thực hiện các kỹ thuật mã hóa video, chẳng hạn như các kỹ thuật được mô tả theo các chuẩn được quy định bởi 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), ITU-T H.265/mã hóa video hiệu suất cao (HEVC), và các phiên bản mở rộng của các chuẩn như vậy. Các thiết bị video có thể truyền, nhận, mã hóa, giải mã và/hoặc lưu trữ thông tin video kỹ thuật số hiệu quả hơn bằng cách thực hiện các kỹ thuật mã hóa video như vậy.

Các kỹ thuật mã hóa video bao gồm dự đoán không gian (nội ảnh) và/hoặc dự đoán thời gian (liên ảnh) để giảm hoặc loại bỏ phần dư vốn có trong các chuỗi video. Đối với kỹ thuật mã hóa video theo khối, lát video (ví dụ, hình ảnh video hoặc một phần hình ảnh video) có thể được chia thành các khối video, các khối này cũng có thể được gọi là đơn vị cây mã hóa (CTU - coding tree unit), đơn vị mã hóa (CU - coding unit) và/hoặc nút mã hóa. Các khối video trong lát mã hóa nội ảnh (I) của hình được mã hóa bằng cách sử dụng kỹ thuật dự đoán không gian đối với các mẫu tham chiếu trong các khối lân cận trong cùng một hình. Các khối video trong lát mã hóa liên ảnh (P hoặc B) của hình có thể sử dụng kỹ thuật dự đoán không gian đối với các mẫu tham chiếu trong các khối lân cận trong cùng một hình, hoặc dự đoán thời gian đối với các mẫu tham chiếu trong các hình tham chiếu

khác. Hình có thể được gọi là khung, và hình tham chiếu có thể được gọi là khung tham chiếu.

Bản chất kỹ thuật của sáng chế

Nói chung, sáng chế liên quan đến việc dự đoán vector chuyển động trong các bộ mã hóa-giải mã (codec) video. Ví dụ, bộ dự đoán vector chuyển động được lựa chọn một cách thích ứng từ hai danh sách ứng viên dự đoán vector chuyển động, các ứng viên này được tạo ra. Danh sách thứ nhất bao gồm các ứng viên dự đoán vector chuyển động mức PU, danh sách thứ hai bao gồm các ứng viên dự đoán vector chuyển động mức PU con.

Theo một phương án làm ví dụ, sáng chế đề xuất phương pháp giải mã dữ liệu video. Phương pháp này bao gồm bước nhận dữ liệu video mã hóa, phân tích cú pháp cơ chuyển động của đơn vị dự đoán con từ dữ liệu video mã hóa, đáp lại việc xác định cơ chuyển động của đơn vị dự đoán con hoạt động, suy ra danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán phụ, đáp lại việc xác định cơ chuyển động của đơn vị dự đoán con không hoạt động, suy ra danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán, lựa chọn bộ dự đoán vector chuyển động từ danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán con hoặc danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán, và giải mã dữ liệu video mã hóa bằng cách sử dụng bộ dự đoán vector chuyển động đã lựa chọn. Dữ liệu video được mã hóa bao gồm khối hiện thời và trong đó danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán con và danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán được suy ra từ các khối lân cận của khối hiện thời. Các khối lân cận là các lân cận về mặt không gian của khối hiện thời trong hình hiện thời hoặc các lân cận về mặt thời gian của khối hiện thời trong hình mã hóa trước đó. Danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán con hoặc danh sách ứng viên dự đoán chuyển động mức đơn vị dự đoán được sắp xếp ít nhất một phần dựa trên các lần xuất hiện dự đoán chuyển động trong các khối lân cận. Các điểm ảnh trong đơn vị dự đoán dùng chung thông tin vector chuyển động thứ nhất và các điểm ảnh trong đơn vị dự đoán con dùng chung thông tin vector chuyển động thứ hai, và trong đó thông tin vector chuyển động thứ nhất hoặc thông tin vector thứ hai được xác định từ bộ dự đoán vector chuyển động đã lựa chọn. Danh sách các ứng viên vector chuyển động mức đơn vị dự đoán bao gồm ít nhất một trong số: ứng viên lân cận không gian và ứng viên lân cận thời gian.

Danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán con bao gồm ít nhất một trong số: dự đoán vector chuyển động afin, dự đoán vector chuyển động thời gian thay thế (alternative temporal motion vector prediction - ATMVP), dự đoán vector chuyển động thời gian-không gian (spatial-temporal motion vector prediction - STMVP), dự đoán vector chuyển động phẳng, và suy ra vector chuyển động phù hợp mẫu (pattern matched motion vector derivation - PMVD). Phương pháp bao gồm bước suy ra chỉ số ứng viên hợp nhất đáp lại việc xác định cờ chuyển động của đơn vị dự đoán con hoạt động, trong đó chỉ số ứng viên hợp nhất chỉ ra bộ dự đoán vector chuyển động sẽ được lựa chọn.

Theo một phương án làm ví dụ khác, sáng chế đề xuất thiết bị giải mã dữ liệu video. Thiết bị bao gồm bộ nhớ để lưu trữ dữ liệu video mã hóa nhận được và bộ xử lý. Bộ xử lý được tạo cấu hình để phân tích cú pháp cờ chuyển động của đơn vị dự đoán con từ dữ liệu video mã hóa, đáp lại việc xác định cờ chuyển động của đơn vị dự đoán con hoạt động, suy ra danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán phụ, đáp lại việc xác định cờ chuyển động của đơn vị dự đoán con không hoạt động, suy ra danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán, lựa chọn bộ dự đoán vector chuyển động từ danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán con hoặc danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán, và giải mã dữ liệu video mã hóa bằng cách sử dụng bộ dự đoán vector chuyển động đã lựa chọn. Dữ liệu video được mã hóa bao gồm khối hiện thời và trong đó danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán con và danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán được suy ra từ các khối lân cận của khối hiện thời. Các khối lân cận là các lân cận về mặt không gian của khối hiện thời trong hình hiện thời hoặc các lân cận về mặt thời gian của khối hiện thời trong hình mã hóa trước đó. Danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán con hoặc danh sách ứng viên dự đoán chuyển động mức đơn vị dự đoán được sắp xếp ít nhất một phần dựa trên các lần xuất hiện dự đoán chuyển động trong các khối lân cận. Các điểm ảnh trong đơn vị dự đoán dùng chung thông tin vector chuyển động thứ nhất và các điểm ảnh trong đơn vị dự đoán con dùng chung thông tin vector chuyển động thứ hai, và trong đó thông tin vector chuyển động thứ nhất hoặc thông tin vector thứ hai được xác định từ bộ dự đoán vector chuyển động đã lựa chọn. Danh sách các ứng viên vector chuyển động mức đơn vị dự đoán bao gồm ít nhất một trong số: ứng viên lân cận không gian và ứng viên lân cận thời gian. Danh sách các ứng viên dự đoán chuyển động mức đơn

vị dự đoán con bao gồm ít nhất một trong số: dự đoán vector chuyển động afin, dự đoán vector chuyển động thời gian thay thế (alternative temporal motion vector prediction - ATMVP), dự đoán vector chuyển động thời gian-không gian (spatial-temporal motion vector prediction - STMVP), dự đoán vector chuyển động phẳng, và suy ra vector chuyển động phù hợp mẫu (pattern matched motion vector derivation - PMVD). Bộ xử lý còn được tạo cấu hình để suy ra chỉ số ứng viên hợp nhất đáp lại việc xác định cờ chuyển động của đơn vị dự đoán con hoạt động, trong đó chỉ số ứng viên hợp nhất chỉ ra bộ dự đoán vector chuyển động sẽ được lựa chọn.

Theo một phương án làm ví dụ khác, sáng chế đề xuất thiết bị giải mã dữ liệu video. Thiết bị bao gồm phương tiện nhớ để lưu trữ dữ liệu video mã hóa đã nhận và phương tiện xử lý. Phương tiện xử lý được tạo cấu hình để, phân tích cú pháp cờ chuyển động của đơn vị dự đoán con từ dữ liệu video mã hóa, đáp lại việc xác định rằng cờ chuyển động của đơn vị dự đoán con hoạt động, suy ra danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán phụ, đáp lại việc xác định cờ chuyển động của đơn vị dự đoán con không hoạt động, suy ra danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán, lựa chọn bộ dự đoán vector chuyển động từ danh sách ứng viên dự đoán chuyển động mức đơn vị dự đoán con hoặc danh sách ứng viên dự đoán chuyển động mức đơn vị dự đoán, và giải mã dữ liệu video mã hóa bằng cách sử dụng bộ dự đoán vector chuyển động đã lựa chọn. Dữ liệu video được mã hóa bao gồm khối hiện thời và trong đó danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán con và danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán được suy ra từ các khối lân cận của khối hiện thời. Các khối lân cận là các lân cận về mặt không gian của khối hiện thời trong hình hiện thời hoặc các lân cận về mặt thời gian của khối hiện thời trong hình mã hóa trước đó. Danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán con hoặc danh sách ứng viên dự đoán chuyển động mức đơn vị dự đoán được sắp xếp ít nhất một phần dựa trên các lần xuất hiện dự đoán chuyển động trong các khối lân cận. Các điểm ảnh trong đơn vị dự đoán dùng chung thông tin vector chuyển động thứ nhất và các điểm ảnh trong đơn vị dự đoán con dùng chung thông tin vector chuyển động thứ hai, và trong đó thông tin vector chuyển động thứ nhất hoặc thông tin vector thứ hai được xác định từ bộ dự đoán vector chuyển động đã lựa chọn. Danh sách các ứng viên vector chuyển động mức đơn vị dự đoán bao gồm ít nhất một trong số: ứng viên lân cận không gian và ứng viên lân cận thời gian. Danh sách các

ứng viên dự đoán chuyển động mức đơn vị dự đoán con bao gồm ít nhất một trong số: dự đoán vector chuyển động afin, dự đoán vector chuyển động thời gian thay thế (alternative temporal motion vector prediction - ATMVP), dự đoán vector chuyển động thời gian-không gian (spatial-temporal motion vector prediction - STMVP), dự đoán vector chuyển động phẳng, và suy ra vector chuyển động phù hợp mẫu (pattern matched motion vector derivation - PMVD). Phương tiện xử lý còn được tạo cấu hình để suy ra chỉ số ứng viên hợp nhất đáp lại việc xác định cờ chuyển động của đơn vị dự đoán con hoạt động, trong đó chỉ số ứng viên hợp nhất chỉ ra bộ dự đoán vector chuyển động sẽ được lựa chọn.

Theo một phương án làm ví dụ khác, phương tiện lưu trữ bất biến đọc được bằng máy tính có các lệnh lưu trữ trên đó mà, khi được thực thi, khiến cho một hoặc nhiều bộ xử lý thực hiện phương pháp. Phương pháp này bao gồm bước nhận dữ liệu video mã hóa, phân tích cú pháp cờ chuyển động của đơn vị dự đoán con từ dữ liệu video mã hóa, đáp lại việc xác định cờ chuyển động của đơn vị dự đoán con hoạt động, suy ra danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán phụ, đáp lại việc xác định cờ chuyển động của đơn vị dự đoán con không hoạt động, suy ra danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán, lựa chọn bộ dự đoán vector chuyển động từ danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán con hoặc danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán, và giải mã dữ liệu video mã hóa bằng cách sử dụng bộ dự đoán vector chuyển động đã lựa chọn. Dữ liệu video được mã hóa bao gồm khối hiện thời và trong đó danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán con và danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán được suy ra từ các khối lân cận của khối hiện thời. Các khối lân cận là các khối gần kề về mặt không gian của khối hiện thời trong hình hiện thời hoặc các khối gần kề về mặt thời gian của khối hiện thời trong hình mã hóa trước đó. Danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán con hoặc danh sách ứng viên dự đoán chuyển động mức đơn vị dự đoán được sắp xếp ít nhất một phần dựa trên các lần xuất hiện dự đoán chuyển động trong các khối lân cận. Các điểm ảnh trong đơn vị dự đoán dùng chung thông tin vector chuyển động thứ nhất và các điểm ảnh trong đơn vị dự đoán con dùng chung thông tin vector chuyển động thứ hai, và trong đó thông tin vector chuyển động thứ nhất hoặc thông tin vector thứ hai được xác định từ bộ dự đoán vector chuyển động đã lựa chọn. Danh sách các ứng viên vector chuyển động mức đơn vị dự đoán bao gồm ít nhất một trong số: ứng viên lân cận không

gian và ứng viên lân cận thời gian. Danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán con bao gồm ít nhất một trong số: dự đoán vector chuyển động afin, dự đoán vector chuyển động thời gian thay thế (alternative temporal motion vector prediction - ATMVP), dự đoán vector chuyển động thời gian-không gian (spatial-temporal motion vector prediction - STMVP), dự đoán vector chuyển động phẳng, và suy ra vector chuyển động phù hợp mẫu (pattern matched motion vector derivation - PMVD). Phương pháp bao gồm bước suy ra chỉ số ứng viên hợp nhất đáp lại việc xác định cờ chuyển động của đơn vị dự đoán con hoạt động, trong đó chỉ số ứng viên hợp nhất chỉ ra bộ dự đoán vector chuyển động sẽ được lựa chọn.

Chi tiết của một hoặc nhiều ví dụ được trình bày trong các hình vẽ kèm theo và phần mô tả bên dưới. Các dấu hiệu, mục đích và ưu điểm khác sẽ trở nên rõ ràng qua phần mô tả, hình vẽ, và các điểm yêu cầu bảo hộ.

Mô tả vắn tắt các hình vẽ

Fig.1 là sơ đồ khối minh họa ví dụ về hệ thống mã hóa và giải mã video mà có thể thực hiện các kỹ thuật theo sáng chế.

Các hình vẽ trên Fig.2A và Fig.2B là các sơ đồ khái niệm minh họa ví dụ về cấu trúc cây nhị phân tứ phân (quadtree binary tree - QTBT), và đơn vị cây mã hóa (coding tree unit - CTU) tương ứng.

Fig.3 là lưu đồ minh họa quy trình suy ra bộ dự đoán vector chuyển động.

Fig.4 là sơ đồ khái niệm minh họa các ứng viên vector chuyển động lân cận không gian cho chế độ hợp nhất.

Fig.5 là sơ đồ khái niệm minh họa các ứng viên vector chuyển động thời gian.

Fig.6 minh họa các kỹ thuật lựa chọn các khối vector chuyển động ứng viên.

Fig.7 là sơ đồ khái niệm minh họa dự đoán vector chuyển động thời gian-không gian (Spatial-Temporal Motion Vector Prediction - STMVP).

Fig.8 là sơ đồ khái niệm minh họa các kỹ thuật so khớp hai chiều.

Fig.9 là sơ đồ khái niệm minh họa các kỹ thuật so khớp mẫu.

Fig.10 là sơ đồ khái niệm minh họa dự đoán vector chuyển động phẳng.

Fig.11 là sơ đồ khối minh họa ví dụ về bộ mã hóa video có thể thực hiện các kỹ thuật của sáng chế.

Fig.12 là sơ đồ khối minh họa ví dụ về bộ giải mã video mà có thể thực hiện các kỹ thuật của sáng chế.

Mô tả chi tiết sáng chế

Sáng chế liên quan đến kỹ thuật suy ra vector chuyển động phía bộ giải mã (decoder-side motion vector derivation - DMVD). Các kỹ thuật suy ra vector chuyển động phía bộ giải mã được mô tả trong sáng chế này có thể được sử dụng cùng với bộ mã hóa-giải mã bất kỳ trong số các bộ mã hóa-giải mã dữ liệu video hiện có, chẳng hạn như mã hóa video hiệu quả cao (High Efficiency Video Coding - HEVC), hoặc có thể được dùng làm các kỹ thuật mã hóa cho chuẩn mã hóa video bất kỳ trong tương lai, như mã hóa video đa năng H.266 (Versatile Video Coding - VVC) và mã hóa dữ liệu video thiết yếu (Essential Video Coding - EVC).

Fig.1 là sơ đồ khối minh họa ví dụ về hệ thống mã hóa và giải mã video 100 có thể thực hiện các kỹ thuật của sáng chế. Các kỹ thuật của sáng chế này nói chung hướng đến việc mã hóa (mã hóa và/hoặc giải mã) dữ liệu video, và cụ thể, liên quan đến các kỹ thuật được mô tả ở đây. Nói chung, dữ liệu video bao gồm dữ liệu bất kỳ để xử lý video. Do đó, dữ liệu video có thể bao gồm video thô, video chưa mã hóa, video đã mã hóa, video đã giải mã (ví dụ, được tái tạo), và siêu dữ liệu video, như dữ liệu báo hiệu chẳng hạn.

Như được thể hiện trên Fig.1, hệ thống 100 bao gồm thiết bị nguồn 102 cung cấp dữ liệu video được mã hóa cần được giải mã và hiển thị bởi thiết bị đích 116 trong ví dụ này. Cụ thể, thiết bị nguồn 102 cung cấp dữ liệu video cho thiết bị đích 116 qua phương tiện đọc được bằng máy tính 110. Thiết bị nguồn 102 và thiết bị đích 116 có thể bao gồm thiết bị bất kỳ trong số rất nhiều thiết bị, bao gồm máy tính để bàn, máy tính sổ tay (tức là, máy tính xách tay), máy tính bảng, bộ giải mã tín hiệu truyền hình, các máy điện thoại cầm tay như điện thoại thông minh, ti vi, camera, thiết bị hiển thị, thiết bị phát phương tiện kỹ thuật số, bàn giao tiếp trò chơi điện tử, thiết bị truyền trực tiếp video hoặc thiết bị tương tự. Trong một số trường hợp, thiết bị nguồn 102 và thiết bị đích 116 có thể được trang bị để truyền thông không dây, và do đó có thể được gọi là các thiết bị truyền thông không dây.

Theo ví dụ trên FIG.1, thiết bị nguồn 102 bao gồm nguồn video 104, bộ nhớ 106, bộ mã hóa video 200, và giao diện đầu ra 108. Thiết bị đích 116 bao gồm giao diện đầu vào 122, bộ giải mã video 300, bộ nhớ 120, và thiết bị hiển thị 118. Theo sáng chế, bộ mã hóa video 200 của thiết bị nguồn 102 và/hoặc bộ giải mã video 300 của thiết bị đích 116 có thể được tạo cấu hình để áp dụng các kỹ thuật suy ra vectơ chuyển động phía bộ giải mã. Do đó, thiết bị nguồn 102 là ví dụ về thiết bị mã hóa video, còn thiết bị đích 116 là ví dụ về thiết bị giải mã video. Theo các ví dụ khác, thiết bị nguồn và thiết bị đích có thể bao gồm các thành phần hoặc thiết bị khác. Ví dụ, thiết bị nguồn 102 có thể nhận dữ liệu video từ nguồn video bên ngoài, như camera bên ngoài. Tương tự, thiết bị đích 116 có thể giao tiếp với thiết bị hiển thị bên ngoài, chứ không phải bao gồm thiết bị hiển thị tích hợp.

Hệ thống 100 được thể hiện trên Fig.1 chỉ là một ví dụ. Nói chung, thiết bị mã hóa và/hoặc giải mã video số bất kỳ có thể thực hiện các kỹ thuật suy ra vectơ chuyển động phía bộ giải mã. Thiết bị nguồn 102 và thiết bị đích 116 chỉ là ví dụ về các thiết bị mã hóa như vậy, trong đó thiết bị nguồn 102 tạo ra dữ liệu video mã hóa để truyền đến thiết bị đích 116. Sáng chế này đề cập đến thiết bị "mã hóa" là thiết bị thực hiện mã hóa (mã hóa và/hoặc giải mã) dữ liệu. Do đó, bộ mã hóa video 200 và bộ giải mã video 300 lần lượt là các ví dụ về thiết bị mã hóa, cụ thể là bộ mã hóa video và bộ giải mã video. Trong một số ví dụ, các thiết bị 102, 116 có thể hoạt động theo cách gần như đối xứng sao cho mỗi thiết bị trong số các thiết bị 102, 116 bao gồm các thành phần mã hóa và giải mã video. Do vậy, hệ thống 100 có thể hỗ trợ truyền video một chiều hoặc hai chiều giữa các thiết bị video 102, 116, ví dụ, để truyền trực tiếp video, phát lại video, phát quảng bá video, và/hoặc điện thoại truyền hình.

Nói chung, nguồn video 104 là nguồn dữ liệu video (tức là, dữ liệu video thô, chưa được mã hóa) và cung cấp chuỗi các hình liên tiếp (còn được gọi là "khung") của dữ liệu video cho bộ mã hóa video 200 mã hóa dữ liệu cho các hình. Nguồn video 104 của thiết bị nguồn 102 có thể bao gồm thiết bị quay video, như máy quay video, kho video chứa video thô đã quay trước đó, và/hoặc giao diện nạp video để nhận video từ nhà cung cấp nội dung video. Theo một phương án khác, nguồn video 104 có thể tạo ra dữ liệu dựa trên đồ họa máy tính dưới dạng video nguồn, hoặc kết hợp của video trực tiếp, video đã lưu trữ, và video được tạo ra bởi máy tính. Trong mỗi trường hợp, bộ mã hóa video 200 mã hóa dữ liệu video được chụp, quay trước hoặc được tạo ra bởi máy tính. Bộ mã hóa video 200 có

thể sắp xếp lại các hình từ thứ tự đã nhận (đôi khi được gọi là “thứ tự hiển thị”) thành thứ tự mã hóa để mã hóa. Bộ mã hóa video 200 có thể tạo ra luồng bit bao gồm dữ liệu video mã hóa. Thiết bị nguồn 102 sau đó có thể xuất ra dữ liệu video mã hóa qua giao diện đầu ra 108 lên phương tiện đọc được bằng máy tính 110 để nhận và/hoặc khôi phục bằng, ví dụ, giao diện đầu vào 122 của thiết bị đích 116.

Bộ nhớ 106 của thiết bị nguồn 102 và bộ nhớ 120 của thiết bị đích 116 là các bộ nhớ đa năng. Trong một số ví dụ, các bộ nhớ 106, 120 có thể lưu trữ dữ liệu video thô, ví dụ, video thô từ nguồn video 104 và dữ liệu video thô, được giải mã từ bộ giải mã video 300. Ngoài ra hoặc theo cách khác, các bộ nhớ 106, 120 có thể lưu trữ các lệnh phần mềm có thể lần lượt thực thi được bằng, ví dụ, bộ mã hóa video 200 và bộ giải mã video 300. Mặc dù được thể hiện riêng với bộ mã hóa video 200 và bộ giải mã video 300 trong ví dụ này, nhưng cần phải hiểu rằng bộ mã hóa video 200 và bộ giải mã video 300 cũng có thể bao gồm các bộ nhớ trong cho các mục đích tương tự hoặc tương đương về mặt chức năng. Hơn nữa, các bộ nhớ 106, 120 có thể lưu trữ dữ liệu video mã hóa, ví dụ, xuất ra từ bộ mã hóa video 200 và nhập vào bộ giải mã video 300. Trong một số ví dụ, các phần của các bộ nhớ 106, 120 có thể được phân bổ dưới dạng một hoặc nhiều bộ đệm video, ví dụ, để lưu trữ dữ liệu video thô, được giải mã và/hoặc được mã hóa.

Phương tiện đọc được bằng máy tính 110 có thể là loại phương tiện hoặc thiết bị bất kỳ có khả năng truyền dữ liệu video mã hóa từ thiết bị nguồn 102 đến thiết bị đích 116. Trong một số ví dụ, phương tiện đọc được bằng máy tính 110 là phương tiện truyền thông cho phép thiết bị nguồn 102 truyền dữ liệu video mã hóa trực tiếp đến thiết bị đích 116 theo thời gian thực, ví dụ, qua mạng tần số vô tuyến hoặc mạng dựa trên máy tính. Giao diện đầu ra 108 có thể điều chế tín hiệu truyền bao gồm dữ liệu video mã hóa và giao diện đầu vào 122 có thể điều chế tín hiệu truyền nhận được, theo chuẩn truyền thông, như giao thức truyền thông không dây. Phương tiện truyền thông có thể là bao gồm phương tiện truyền thông không dây hoặc có dây bất kỳ, như phổ tần số vô tuyến (Radio Frequency - RF) hay một hoặc nhiều đường truyền vật lý. Phương tiện truyền thông có thể tạo nên một phần mạng dựa vào gói, như mạng cục bộ, mạng diện rộng, hoặc mạng toàn cầu như Internet. Phương tiện truyền thông có thể bao gồm các bộ định tuyến, bộ chuyển mạch, trạm gốc, hoặc thiết bị khác bất kỳ có thể có ích để hỗ trợ truyền thông từ thiết bị nguồn 102 đến thiết bị đích 116.

Trong một số ví dụ, thiết bị nguồn 102 có thể xuất ra dữ liệu mã hóa từ giao diện đầu ra 108 đến thiết bị lưu trữ 112. Tương tự, thiết bị đích 116 có thể truy cập dữ liệu mã hóa từ thiết bị lưu trữ 112 qua giao diện đầu vào 122. Thiết bị lưu trữ 112 có thể bao gồm phương tiện bất kỳ trong số nhiều phương tiện lưu trữ dữ liệu truy cập phân tán hoặc cục bộ như ổ cứng, đĩa Blu-ray, DVD, CD-ROM, bộ nhớ nhanh, bộ nhớ khả biến hoặc không khả biến hoặc phương tiện lưu trữ số thích hợp khác bất kỳ để lưu trữ dữ liệu video mã hóa.

Trong một số ví dụ, thiết bị nguồn 102 có thể xuất ra dữ liệu video mã hóa cho máy chủ tệp 114 hoặc một thiết bị lưu trữ trung gian khác mà có thể lưu trữ video mã hóa bằng thiết bị nguồn 102. Thiết bị đích 116 có thể truy cập dữ liệu video đã lưu trữ từ máy chủ tệp 114 qua truyền trực tuyến hoặc tải xuống. Máy chủ tệp 114 có thể là loại thiết bị 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 116. Máy chủ tệp 114 có thể là máy chủ web (ví dụ, cho trang web), máy chủ giao thức truyền tệp (File Transfer Protocol - FTP), thiết bị mạng phân phối nội dung, hoặc thiết bị lưu trữ gắn với mạng (network attached storage - NAS). Thiết bị đích 116 có thể truy cập dữ liệu video mã hóa từ máy chủ tệp 114 qua kết nối dữ liệu tiêu chuẩn bất kỳ, bao gồm kết nối Internet. Kết nối này có thể bao gồm kênh không dây (ví dụ, kết nối Wi-Fi), kết nối có dây (ví dụ, đường thuê bao kỹ thuật số (Digital Subscriber Line - DSL), modem cáp, v.v.), hoặc kết hợp hai loại này phù hợp với việc truy cập dữ liệu video mã hóa được lưu trữ trên máy chủ tệp 114. Máy chủ tệp 114 và giao diện đầu vào 122 có thể được tạo cấu hình để hoạt động theo giao thức truyền trực tuyến, giao thức truyền tải xuống hoặc kết hợp của chúng.

Giao diện đầu ra 108 và giao diện đầu vào 122 có thể đại diện cho bộ phát/bộ thu không dây, modem, các thành phần mạng có dây (ví dụ, card Ethernet), các thành phần truyền thông không dây hoạt động theo bất kỳ trong số các tiêu chuẩn IEEE 802.11, hoặc các thành phần vật lý khác. Trong các ví dụ trong đó giao diện đầu ra 108 và giao diện đầu vào 122 bao gồm các thành phần không dây, giao diện đầu ra 108 và giao diện đầu vào 122 có thể được tạo cấu hình để truyền dữ liệu, chẳng hạn như dữ liệu video được mã hóa, theo chuẩn truyền thông di động, như 4G, 4G-Phát triển dài hạn (Long - Term Evolution - LTE), LTE cải tiến, 5G, hoặc tương tự. Trong một số ví dụ trong đó giao diện đầu ra 108 bao gồm bộ phát không dây, giao diện đầu ra 108 và giao diện đầu vào 122 có thể được tạo cấu

hình để truyền dữ liệu, như dữ liệu video mã hóa, theo các tiêu chuẩn không dây khác, chẳng hạn như đặc tả kỹ thuật IEEE 802.11, đặc tả kỹ thuật IEEE 802.15 (ví dụ, ZigBee™), tiêu chuẩn Bluetooth™, hoặc tương tự. Trong một số ví dụ, thiết bị nguồn 102 và/hoặc thiết bị đích 116 có thể bao gồm các thiết bị hệ thống trên chip (system-on-a-chip - SoC) tương ứng. Ví dụ, thiết bị nguồn 102 có thể bao gồm thiết bị SoC để thực hiện chức năng được quy cho bộ mã hóa video 200 và/hoặc giao diện đầu ra 108 và thiết bị đích 116 có thể bao gồm thiết bị SoC để thực hiện chức năng được gán cho bộ giải mã video 300 và/hoặc giao diện đầu vào 122.

Các kỹ thuật theo sáng chế có thể được áp dụng để mã hóa video để hỗ trợ ứng dụng đa phương tiện bất kỳ trong số nhiều ứng dụng đa phương tiện, như phát sóng truyền hình qua không gian, truyền hình cáp, truyền hình vệ tinh, truyền video trực tuyến qua mạng internet, như truyền trực tuyến thích ứng động qua giao thức HTTP (DASH - Dynamic Adaptive Streaming over HTTP), video dạng số được mã hóa trên phương tiện lưu trữ dữ liệu, giải mã video dạng số lưu trữ trên phương tiện lưu trữ dữ liệu, hoặc các ứng dụng khác.

Giao diện đầu vào 122 của thiết bị đích 116 nhận luồng bit video được mã hóa từ phương tiện đọc được bằng máy tính 110 (ví dụ, thiết bị lưu trữ bất biến 112, máy chủ tệp 114 hoặc tương tự). Phương tiện đọc được bằng máy tính có luồng bit video được mã hóa 110 có thể bao gồm thông tin báo hiệu được xác định bởi bộ mã hóa video 200, thông tin này cũng được bộ giải mã video 300 sử dụng, chẳng hạn như các phân tử cú pháp có các giá trị mô tả các đặc điểm và/hoặc xử lý các khối video hoặc các đơn vị mã hóa khác (ví dụ, lát, hình, nhóm hình, chuỗi hoặc tương tự). Thiết bị hiển thị 118 hiển thị hình giải mã của dữ liệu video giải mã cho người dùng. Thiết bị hiển thị 118 có thể là bất kỳ trong số nhiều thiết bị hiển thị như màn hình ống tia catot (Cathode Ray Tube - CRT), màn hình tinh thể lỏng (Liquid Crystal Display - LCD), màn hình plasma, màn hình điốt phát sáng hữu cơ (Organic Light Emitting Diode - OLED), hoặc một loại thiết bị hiển thị khác.

Mặc dù không được thể hiện trên Fig.1, nhưng trong một số ví dụ, mỗi bộ mã hóa video 200 và bộ giải mã video 300 có thể được tích hợp với bộ mã hóa âm thanh và/hoặc bộ giải mã âm thanh, và có thể bao gồm các đơn vị ghép kênh-giải ghép kênh (MUX-DEMUX) thích hợp, hoặc phần cứng và/hoặc phần mềm khác, để xử lý các luồng được

ghép kênh bao gồm cả âm thanh và video trong một luồng dữ liệu chung. Nếu có thể, các đơn vị MUX-DEMUX có thể tuân theo giao thức ghép kênh ITU H.223, hoặc các giao thức khác như giao thức gói dữ liệu người dùng (user datagram protocol - UDP).

Mỗi bộ mã hóa video 200 và bộ giải mã video 300 có thể được triển khai dưới dạng bất kỳ trong số nhiều mạch mã hóa và/hoặc giải mã thích hợp, như một hoặc nhiều bộ vi xử lý, bộ xử lý tín hiệu số (digital signal processor - DSP), mạch tích hợp chuyên dụng (Application Specific Integrated Circuit - ASIC), mảng cổng lập trình được theo trường (Field Programmable Gate Array - FPGA), logic rời rạc, phần mềm, phần cứng, firmware hoặc kết hợp bất kỳ của chúng. Khi các kỹ thuật này được triển khai một phần trong phần mềm, thiết bị có thể lưu trữ các lệnh cho phần mềm trên phương tiện bất biến đọc được bằng máy tính 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 của sáng chế này. Mỗi bộ mã hóa video 200 và bộ giải mã video 300 có thể được bao gồm trong một hoặc nhiều bộ mã hóa hoặc bộ giải mã, mỗi trong số chúng có thể được tích hợp như là một phần của bộ mã hóa/bộ giải mã (encoder/decoder - CODEC) kết hợp trong thiết bị tương ứng. Thiết bị bao gồm bộ mã hóa video 200 và/hoặc bộ giải mã video 300 có thể bao gồm mạch tích hợp, bộ vi xử lý và/hoặc thiết bị truyền thông không dây, như điện thoại di động.

Các tiêu chuẩn mã hóa video bao gồm ITU-T H.261, ISO/IEC MPEG-1 Visual, ITU-T H.262 hoặc ISO/IEC MPEG-2 Visual, ITU-T H.263, ISO/IEC MPEG-4 Visual và ITU-T H.264 (cũng được biết đến là ISO/IEC MPEG-4 AVC), bao gồm các phần mở rộng mã hóa video định tỷ lệ được (Scalable video Coding - SVC) và mã hóa video đa khung hình (Multiview Video Coding - MVC).

Chuẩn mã hóa dữ liệu video, có tên là chuẩn mã hóa dữ liệu video hiệu suất cao (HEVC) (còn được gọi là ITU-T H.265), bao gồm các phiên bản mở rộng dài, phiên bản mở rộng đa khung hình (MV-HEVC) và phiên bản mở rộng định tỷ lệ được (SHVC) của nó, được phát triển bởi nhóm hợp tác liên kết về mã hóa video (Joint Collaboration Team on Video Coding - JCT-VC) cũng như Nhóm cộng tác chung về phát triển mở rộng mã hóa video 3D (Joint Collaboration Team on 3D Video Coding Extension Development - JCT-3V) của nhóm chuyên gia mã hóa video (Video Coding Experts Group - VCEG) của tổ chức ITU-T và nhóm chuyên gia hình động (Motion Picture Experts Group - MPEG) của tổ chức ISO/IEC.

ITU-T VCEG (Q6/16) và ISO/IEC MPEG (JTC 1/SC 29/WG 11) hiện nay đang nghiên cứu nhu cầu tiềm năng về chuẩn hóa công nghệ mã hóa video trong tương lai với khả năng nén vượt qua khả năng nén của chuẩn HEVC hiện thời (bao gồm các phiên bản mở rộng hiện thời của nó và các phiên bản mở rộng sắp tới về mã hóa nội dung màn hình và mã hóa dải động cao). Các nhóm này đang cùng nhau làm việc về hoạt động khảo sát này với nỗ lực hợp tác chung được biết đến dưới tên nhóm khảo sát video chung (Joint Video Exploration Team - JVET) để đánh giá các thiết kế công nghệ nén được đề xuất bởi các chuyên gia trong lĩnh vực này. JVET gặp nhau lần đầu tiên từ ngày 19 đến 21/10/2015. Và phiên bản mới nhất của phần mềm tham khảo, tức là mô hình khảo sát chung 7 (Joint Exploration Model 7 - JEM 7) có thể được tải về từ:

https://jvet.hhi.fraunhofer.de/svn/svn_HMJEMSoftware/tags/HM-16.6-JEM-57.0/

Phần mô tả thuật toán của mô hình JEM7 có thể đề cập đến JVET-D1001. [JVET-G1001](#).

Bộ mã hóa video 200 và bộ giải mã video 300 có thể hoạt động theo chuẩn mã hóa video, chẳng hạn như ITU-T H.265, còn được gọi là mã hóa video hiệu suất cao (HEVC) hoặc các phần mở rộng của nó, chẳng hạn như các phần mở rộng mã hóa video đa khung hình và/hoặc định tỷ lệ được. Theo cách khác, bộ mã hóa video 200 và bộ giải mã video 300 có thể hoạt động theo các chuẩn độc quyền hoặc công nghiệp khác, như mô hình thử nghiệm khảo sát chung (Joint Exploration Test Model - JEM) hoặc ITU-T H.266, còn được gọi là mã hóa video đa năng (Versatile Video Coding - VVC). Dự thảo gần đây của chuẩn VVC được mô tả trong Bross, và các cộng sự. “Versatile Video Coding (Dự thảo 3),” Nhóm chuyên gia video chung (Joint Video Experts Team - JVET) của ITU-T SG 16 WP 3 và ISO/IEC JTC 1/SC 29/WG 11, Hội nghị lần thứ 13: Marrakech, MA, từ ngày 9 đến 18 tháng 1 năm 2019, JVET-M1001-v5 (sau đây gọi là “VVC Dự thảo 4”). Tuy nhiên, các kỹ thuật của sáng chế không bị giới hạn ở bất kỳ chuẩn mã hóa cụ thể nào.

Như sẽ được mô tả, bộ giải mã video 300 có thể được tạo cấu hình để thực hiện một hoặc nhiều kỹ thuật của sáng chế. Ví dụ, bộ giải mã video 300 có thể phân tích cú pháp cờ chuyển động của đơn vị dự đoán con từ dữ liệu video được mã hóa nhận được, suy ra danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán con nếu cờ chuyển động của đơn vị dự đoán con hoạt động, suy ra danh sách các ứng viên dự đoán chuyển động mức đơn vị dự đoán con nếu cờ chuyển động của đơn vị dự đoán con không hoạt động, và giải mã dữ liệu video được mã hóa bằng cách sử dụng bộ dự đoán vector chuyển động đã lựa chọn.

Nói chung, bộ mã hóa video 200 và bộ giải mã video 300 có thể thực hiện mã hóa các hình dựa trên khối. Thuật ngữ “khối” thường đề cập đến cấu trúc bao gồm dữ liệu cần xử lý (ví dụ, mã hóa, giải mã, hoặc được sử dụng trong quy trình mã hóa và/hoặc giải mã). Ví dụ, khối có thể bao gồm ma trận mẫu hai chiều của dữ liệu độ chói và/hoặc sắc độ. Nói chung, bộ mã hóa video 200 và bộ giải mã video 300 có thể mã hóa dữ liệu video được biểu diễn theo định dạng YUV (ví dụ, Y, Cb, Cr). Tức là, thay vì mã hóa dữ liệu màu đỏ, xanh lục và xanh lam (red, green, and blue-RGB) cho các mẫu của hình, bộ mã hóa video 200 và bộ giải mã video 300 có thể mã hóa các thành phần độ chói và sắc độ, trong đó các thành phần sắc độ có thể bao gồm cả thành phần màu đỏ và màu xanh lam. Trong một số ví dụ, bộ mã hóa video 200 chuyển đổi dữ liệu định dạng RGB nhận được thành dạng biểu diễn YUV trước khi mã hóa, và bộ giải mã video 300 chuyển đổi dạng biểu diễn YUV thành định dạng RGB. Theo cách khác, các đơn vị xử lý trước và sau (không được thể hiện trên hình vẽ) có thể thực hiện các chuyển đổi này.

Nói chung sáng chế có thể đề cập đến việc mã hóa (ví dụ, mã hóa và giải mã) các hình để bao gồm quy trình mã hóa hoặc giải mã dữ liệu của hình. Tương tự, sáng chế có thể đề cập đến việc mã hóa các khối của hình để bao gồm quy trình mã hóa hoặc giải mã dữ liệu cho các khối, ví dụ, mã hóa dự đoán và/hoặc mã hóa phần dư. Luồng bit video mã hóa nói chung bao gồm một loạt giá trị cho các phần tử cú pháp đại diện cho các quyết định mã hóa (ví dụ, chế độ mã hóa) và phân chia các hình thành các khối. Do đó, các tham chiếu đến mã hóa hình hoặc khối thường được hiểu là mã hóa các giá trị cho các phần tử cú pháp tạo thành hình hoặc khối.

HEVC định nghĩa các khối khác nhau, bao gồm các đơn vị mã hóa (coding unit - CU), đơn vị dự đoán (prediction unit - PU) và đơn vị biến đổi (transform unit - TU). Theo HEVC, bộ mã hóa video (chẳng hạn như bộ mã hóa video 200) phân chia đơn vị cây mã hóa (coding tree unit - CTU) thành các CU theo cấu trúc cây tứ phân. Tức là, bộ mã hóa video phân chia các CTU và CU thành bốn hình vuông không chồng lên nhau bằng nhau, và mỗi nút của cây tứ phân có không hoặc bốn nút con. Các nút không có nút con có thể được gọi là “nút lá” và các CU của các nút lá như vậy có thể bao gồm một hoặc nhiều PU và/hoặc một hoặc nhiều TU. Bộ mã hóa video có thể phân chia tiếp các PU và các TU. Ví dụ, trong HEVC, cây tứ phân dư (residual quadtree - RQT) biểu diễn việc phân chia các TU. Trong HEVC, các PU biểu diễn dữ liệu dự đoán liên ảnh, còn các TU biểu diễn dữ

liệu dư. Các CU được dự đoán nội ảnh chứa thông tin dự đoán nội ảnh, như chỉ báo chế độ nội ảnh.

Trong một ví dụ khác, bộ mã hóa video 200 và bộ giải mã video 300 có thể được tạo cấu hình để hoạt động theo JEM hoặc VVC. Theo JEM hoặc VVC, bộ mã hóa video (như bộ mã hóa video 200) phân chia hình thành nhiều đơn vị cây mã hóa (CTU). Bộ mã hóa video 200 có thể phân chia CTU theo cấu trúc cây, chẳng hạn như cấu trúc cây tứ phân - cây nhị phân (quadtree-binary tree - QTBT). Cấu trúc QTBT loại bỏ các khái niệm về nhiều loại phân chia, như phân chia giữa các CU, PU và TU theo tiêu chuẩn HEVC. Cấu trúc QTBT của JEM bao gồm hai mức: mức thứ nhất được chia theo kỹ thuật phân chia cây tứ phân và mức thứ hai được phân chia theo kỹ thuật phân chia cây nhị phân. Nút gốc của cấu trúc QTBT tương ứng với CTU. Các nút lá của cây nhị phân tương ứng với các đơn vị mã hóa (CU).

Trong cấu trúc phân chia MTT, các khối có thể được phân chia bằng cách sử dụng phân chia cây tứ phân (quadtree - QT), phân chia cây nhị phân (binary tree - BT), và một hoặc nhiều loại phân chia cây tam phân (triple tree - TT). Phân chia cây tam phân là phân chia trong đó khối được phân chia thành ba khối con. Trong một số ví dụ, phân chia cây tam phân chia khối thành ba khối con mà không chia khối gốc qua tâm. Các loại phân chia trong MTT (ví dụ, QT, BT, và TT), có thể là đối xứng hoặc bất đối xứng.

Trong một số ví dụ, bộ mã hóa video 200 và bộ giải mã video 300 có thể sử dụng một cấu trúc QTBT hoặc MTT để biểu diễn mỗi thành phần trong số các thành phần độ chói và sắc độ, còn trong các ví dụ khác, bộ mã hóa video 200 và bộ giải mã video 300 có thể sử dụng hai hoặc nhiều cấu trúc QTBT hoặc MTT, như cấu trúc QTBT/MTT cho thành phần độ chói và cấu trúc QTBT/MTT khác cho cả hai thành phần sắc độ (hoặc hai cấu trúc QTBT/MTT cho các thành phần sắc độ tương ứng).

Bộ mã hóa video 200 và bộ giải mã video 300 có thể được tạo cấu hình để sử dụng kỹ thuật phân chia cây tứ phân theo cấu trúc phân chia HEVC, phân chia QTBT, hoặc phân chia MTT, hoặc các cấu trúc phân chia khác. Nhằm mục đích giải thích, phần mô tả về các kỹ thuật của sáng chế được trình bày liên quan đến kỹ thuật phân chia cây QTBT. Tuy nhiên, cần phải hiểu rằng các kỹ thuật của sáng chế cũng có thể được áp dụng cho các bộ

mã hóa video được tạo cấu hình để sử dụng kỹ thuật phân chia cây tứ phân hoặc các kiểu phân chia khác.

Sáng chế có thể sử dụng “ $N \times N$ ” và “ N với N ” thay thế cho nhau để chỉ các kích thước mẫu của khối (chẳng hạn như CU hoặc khối video khác) theo chiều dọc và chiều ngang, ví dụ, 16×16 mẫu hoặc 16 với 16 mẫu. Nói chung, CU 16×16 sẽ có 16 mẫu theo chiều dọc ($y = 16$) và 16 mẫu theo chiều ngang ($x = 16$). Tương tự, CU $N \times N$ nói chung có N mẫu theo chiều dọc và N mẫu theo chiều ngang, trong đó N là giá trị nguyên không âm. Các mẫu trong CU có thể được sắp xếp thành các hàng và các cột. Hơn nữa, các CU không nhất thiết phải có số lượng mẫu theo chiều ngang bằng số lượng mẫu theo chiều dọc. Ví dụ, các CU có thể bao gồm $N \times M$ mẫu, trong đó M không nhất thiết bằng N .

Bộ mã hóa video 200 mã hóa dữ liệu video cho các CU biểu diễn thông tin dự đoán và/hoặc thông tin dư, và thông tin khác. Thông tin dự đoán chỉ báo cách thức CU sẽ được dự đoán để tạo ra khối dự đoán cho CU. Thông tin dư thường biểu diễn các chênh lệch mẫu với mẫu giữa các mẫu của CU trước khi mã hóa và khối dự đoán.

Để dự đoán CU, bộ mã hóa video 200 nói chung có thể tạo ra khối dự đoán cho CU thông qua dự đoán liên ảnh hoặc dự đoán nội ảnh. Dự đoán liên ảnh thường chỉ việc dự đoán CU từ dữ liệu của hình được mã hóa trước đó, trong khi dự đoán nội ảnh thường chỉ việc dự đoán CU từ dữ liệu được mã hóa trước đó của cùng một hình. Để thực hiện dự đoán liên ảnh, bộ mã hóa video 200 có thể tạo ra khối dự đoán bằng cách sử dụng một hoặc nhiều vector chuyển động. Bộ mã hóa video 200 có thể thường thực hiện tìm kiếm chuyển động để nhận dạng khối tham chiếu phù hợp nhất với CU, ví dụ, về các khác biệt giữa CU và khối tham chiếu. Bộ mã hóa video 200 có thể tính toán metric chênh lệch bằng cách sử dụng tổng chênh lệch tuyệt đối (sum of absolute difference - SAD), tổng chênh lệch bình phương (sum of squared differences - SSD), chênh lệch tuyệt đối trung bình (mean absolute difference - MAD), chênh lệch bình phương trung bình (mean squared differences - MSD) hoặc các tính toán chênh lệch như vậy khác để xác định xem khối tham chiếu có phù hợp nhất với CU hiện thời không. Trong một số ví dụ, bộ mã hóa video 200 có thể dự đoán CU hiện thời bằng cách sử dụng dự đoán một hướng hoặc dự đoán hai hướng.

Một số ví dụ về JEM và VVC còn cung cấp chế độ bù chuyển động afin, chế độ này có thể được coi là chế độ dự đoán liên ảnh. Trong chế độ bù chuyển động afin, bộ mã hóa

video 200 có thể xác định hai hoặc nhiều vector chuyển động đại diện cho chuyển động không tịnh tiến, chẳng hạn như phóng to hoặc thu nhỏ, xoay, chuyển động phối cảnh hoặc các loại chuyển động không đều khác.

Để thực hiện dự đoán nội ảnh, bộ mã hóa video 200 có thể lựa chọn chế độ dự đoán nội ảnh để tạo khối dự đoán. Một số ví dụ về JEM và VVC cung cấp sáu mươi bảy chế độ dự đoán nội ảnh, bao gồm các chế độ định hướng khác nhau, cũng như chế độ phẳng và chế độ DC. Nói chung, bộ mã hóa video 200 chọn chế độ dự đoán nội ảnh để mô tả các mẫu lân cận cho khối hiện thời (ví dụ, khối của CU) để dự đoán các mẫu của khối hiện thời từ đó. Các mẫu như vậy có thể thường ở trên, ở trên và bên trái hoặc bên trái của khối hiện thời trong cùng một hình với khối hiện thời, giả sử bộ mã hóa video 200 mã hóa CTU và CU theo thứ tự quét mảnh (từ trái sang phải, từ trên xuống dưới).

Bộ mã hóa video 200 mã hóa dữ liệu đại diện cho chế độ dự đoán cho khối hiện thời. Ví dụ, đối với các chế độ dự đoán liên ảnh, bộ mã hóa video 200 có thể mã hóa dữ liệu đại diện cho chế độ dự đoán liên ảnh nào được dùng trong số các chế độ dự đoán liên ảnh có sẵn, cũng như thông tin chuyển động cho chế độ tương ứng. Ví dụ, đối với dự đoán liên ảnh một chiều hoặc hai chiều, bộ mã hóa video 200 có thể mã hóa các vector chuyển động bằng cách sử dụng chế độ dự đoán vector chuyển động nâng cao (advanced motion vector prediction - AMVP) hoặc chế độ hợp nhất. Bộ mã hóa video 200 có thể sử dụng các chế độ tương tự để mã hóa các vector chuyển động cho chế độ bù chuyển động afin.

Sau dự đoán, chẳng hạn như dự đoán nội ảnh hoặc dự đoán liên ảnh một khối, bộ mã hóa video 200 có thể tính toán dữ liệu dư cho khối. Dữ liệu dư, chẳng hạn như khối dư, biểu diễn chênh lệch mẫu với mẫu giữa khối và khối dự đoán cho khối, được tạo ra bằng cách sử dụng chế độ dự đoán tương ứng. Bộ mã hóa video 200 có thể áp dụng một hoặc nhiều phép biến đổi cho khối dư, để tạo ra dữ liệu được biến đổi trong miền biến đổi thay vì miền mẫu. Ví dụ, bộ mã hóa video 200 có thể áp dụng phép biến đổi cosin rời rạc (discrete cosine transform - DCT), biến đổi số nguyên, biến đổi sóng nhỏ, hoặc biến đổi tương tự về mặt khái niệm cho dữ liệu video dư. Ngoài ra, bộ mã hóa video 200 có thể áp dụng phép biến đổi thứ cấp tiếp theo phép biến đổi thứ nhất, chẳng hạn như biến đổi thứ cấp không phân tách phụ thuộc vào chế độ (mode-dependent non-separable secondary transform - MDNSST), biến đổi phụ thuộc vào tín hiệu, biến đổi Karhunen-Loeve

(Karhunen-Loeve transform - KLT) hoặc tương tự. Bộ mã hóa video 200 tạo ra các hệ số biến đổi sau khi áp dụng một hoặc nhiều phép biến đổi.

Như nêu ở trên, tiếp theo phép biến đổi bất kỳ để tạo ra các hệ số biến đổi, bộ mã hóa video 200 có thể thực hiện phép lượng tử hóa các hệ số biến đổi. Nói chung, lượng tử hóa chỉ quy trình trong đó các hệ số biến đổi được lượng tử hóa để có thể giảm lượng dữ liệu dùng để biểu diễn các hệ số, cho phép nén thêm. Bằng cách thực hiện quy trình lượng tử hóa, bộ mã hóa video 200 có thể giảm độ sâu bit liên quan đến một số hoặc tất cả các hệ số. Ví dụ, bộ mã hóa video 200 có thể làm tròn giá trị n -bit xuống thành giá trị m -bit trong quá trình lượng tử hóa, trong đó n lớn hơn m . Trong một số ví dụ, để thực hiện lượng tử hóa, bộ mã hóa video 200 có thể thực hiện dịch sang phải theo bit giá trị cần lượng tử hóa.

Sau khi lượng tử hóa, bộ mã hóa video 200 có thể quét các hệ số biến đổi, tạo ra vector một chiều từ ma trận hai chiều bao gồm các hệ số biến đổi đã lượng tử hóa. Việc quét có thể được thiết kế để đặt các hệ số năng lượng cao hơn (và do đó tần số thấp hơn) ở phía trước vector và để đặt các hệ số biến đổi năng lượng thấp hơn (và do đó tần số cao hơn) ở phía sau của vector. Trong một số ví dụ, bộ mã hóa video 200 có thể sử dụng thứ tự quét định trước để quét các hệ số biến đổi lượng tử hóa để tạo ra vector nối tiếp, và sau đó mã hóa entropy các hệ số biến đổi lượng tử hóa của vector. Trong các ví dụ khác, bộ mã hóa video 200 có thể thực hiện quét thích ứng. Sau khi quét các hệ số biến đổi được lượng tử hóa để tạo vector một chiều, bộ mã hóa video 200 có thể mã hóa entropy vector một chiều, ví dụ, theo kỹ thuật mã hóa số học nhị phân thích ứng ngữ cảnh (context-adaptive binary arithmetic coding - CABAC). Bộ mã hóa video 200 cũng có thể mã hóa entropy các giá trị cho các phân tử cú pháp mô tả siêu dữ liệu liên quan đến dữ liệu video được mã hóa dùng bởi bộ giải mã video 300 khi giải mã dữ liệu video.

Để thực hiện CABAC, bộ mã hóa video 200 có thể gán ngữ cảnh trong mô hình ngữ cảnh cho ký hiệu cần được truyền. Ngữ cảnh có thể liên quan tới, ví dụ, việc các giá trị lân cận của ký hiệu có giá trị bằng không hay không. Việc xác định xác suất có thể dựa vào ngữ cảnh được gán cho ký hiệu.

Bộ mã hóa video 200 còn có thể tạo dữ liệu cú pháp, như dữ liệu cú pháp dựa trên khối, dữ liệu cú pháp dựa trên hình, và dữ liệu cú pháp dựa trên chuỗi, cho bộ giải mã video 300, ví dụ, trong phần đầu hình, phần đầu khối, phần đầu lát, hoặc dữ liệu cú pháp khác,

như tập tham số chuỗi (sequence parameter set - SPS), tập tham số hình (picture parameter set - PPS), hoặc tập tham số video (video parameter set - VPS). Tương tự, bộ giải mã video 300 có thể giải mã dữ liệu cú pháp như vậy để xác định cách giải mã dữ liệu video tương ứng.

Theo cách này, bộ mã hóa video 200 có thể tạo luồng bit bao gồm dữ liệu video mã hóa, ví dụ, các phần tử cú pháp mô tả việc phân chia hình thành các khối (ví dụ, các CU) và thông tin dự đoán và/hoặc thông tin dư cho các khối. Cuối cùng, bộ giải mã video 300 có thể nhận luồng bit và giải mã dữ liệu video mã hóa.

Nói chung, bộ giải mã video 300 thực hiện quy trình ngược lại với quy trình được thực hiện bởi bộ mã hóa video 200 để giải mã dữ liệu video mã hóa của luồng bit. Ví dụ, bộ giải mã video 300 có thể giải mã các giá trị cho các phần tử cú pháp của luồng bit bằng cách sử dụng CABAC theo cách về cơ bản tương tự như, mặc dù ngược với, quy trình mã hóa CABAC của bộ mã hóa video 200. Các phần tử cú pháp có thể xác định thông tin phân chia hình thành các CTU, và phân chia mỗi CTU theo cấu trúc phân chia tương ứng, chẳng hạn như cấu trúc QTBT, để xác định các CU của CTU. Các phần tử cú pháp còn có thể xác định thông tin dự đoán và thông tin dư cho các khối (ví dụ, các CU) của dữ liệu video.

Thông tin dư có thể được biểu diễn bằng, ví dụ, các hệ số biến đổi lượng tử hóa. Bộ giải mã video 300 có thể lượng tử hóa ngược và biến đổi ngược các hệ số biến đổi lượng tử hóa của khối để tái tạo khối dư cho khối. Bộ giải mã video 300 sử dụng chế độ dự đoán được báo hiệu (dự đoán nội ảnh hoặc dự đoán liên ảnh) và thông tin dự đoán liên quan (ví dụ, thông tin chuyển động cho dự đoán liên ảnh) để tạo ra khối dự đoán cho khối. Sau đó, bộ giải mã video 300 có thể kết hợp khối dự đoán và khối dư (dựa trên cơ sở từng mẫu) để tái tạo khối ban đầu. Bộ giải mã video 300 có thể thực hiện xử lý bổ sung, chẳng hạn như thực hiện quy trình tách khối để giảm các xảo ảnh trực quan dọc theo các biên của khối.

Nói chung, sáng chế có thể liên quan đến việc "báo hiệu" một số thông tin như các phần tử cú pháp. Nói chung, thuật ngữ "báo hiệu" có thể chỉ việc truyền thông giá trị các phần tử cú pháp và/hoặc dữ liệu khác được dùng để giải mã dữ liệu video mã hóa. Tức là, bộ mã hóa video 200 có thể báo hiệu các giá trị cho các phần tử cú pháp trong luồng bit. Nói chung, việc báo hiệu đề cập đến việc tạo ra giá trị trong luồng bit. Như đã lưu ý ở trên, thiết bị nguồn 102 có thể truyền luồng bit đến thiết bị đích 116 về cơ bản theo thời gian

thực, hoặc không theo thời gian thực, việc này có thể xảy ra khi lưu trữ các phần tử cú pháp vào thiết bị lưu trữ 112 để thiết bị đích 116 khôi phục sau.

Các hình vẽ trên Fig.2A và Fig.2B là sơ đồ khái niệm minh họa ví dụ về cấu trúc cây nhị phân cây tứ phân (quadtree binary tree - QTBT) 130, và đơn vị cây mã hóa (coding tree unit - CTU) 132 tương ứng. Các đường nét liền biểu diễn việc tách cây tứ phân và các đường nét đứt biểu diễn việc tách cây nhị phân. Trong mỗi nút tách (tức là, nút không phải lá) của cây nhị phân, một cờ được báo hiệu để chỉ báo kiểu tách nào (tức là, tách ngang hay tách dọc) được dùng, trong đó 0 chỉ báo tách ngang và 1 chỉ báo tách dọc trong ví dụ này. Đối với việc tách cây tứ phân, không cần thiết biểu thị kiểu tách, vì các nút cây tứ phân tách một khối theo chiều ngang và theo chiều dọc thành 4 khối con có kích thước bằng nhau. Do đó, bộ mã hóa video 200 có thể mã hóa, và bộ giải mã video 300 có thể giải mã, các phần tử cú pháp (ví dụ như thông tin tách) cho mức cây khu vực của cấu trúc QTBT 130 (tức là, các đường nét liền) và các phần tử cú pháp (ví dụ như thông tin tách) cho mức cây dự đoán của cấu trúc QTBT 130 (tức là, các đường nét đứt). Bộ mã hóa video 200 có thể mã hóa, và bộ giải mã video 300 có thể giải mã, dữ liệu video, như dữ liệu dự đoán và biến đổi, cho các CU biểu diễn bởi nút lá cuối của cấu trúc QTBT 130.

Nói chung, CTU 132 trên Fig.2B có thể được kết hợp với các tham số xác định kích thước của các khối tương ứng với các nút của cấu trúc QTBT 130 ở mức thứ nhất và thứ hai. Các tham số này có thể bao gồm kích thước CTU (đại diện cho kích thước CTU 132 trong các mẫu), kích thước cây tứ phân tối thiểu (MinQTSIZE, đại diện cho kích thước nút lá cây tứ phân tối thiểu được phép), kích thước cây nhị phân tối đa (MaxBTSIZE, đại diện cho kích thước nút gốc cây nhị phân tối đa được phép), độ sâu cây nhị phân tối đa (MaxBTDepth, đại diện cho độ sâu cây nhị phân tối đa được phép) và kích thước cây nhị phân tối thiểu (MinBTSIZE, đại diện cho kích thước nút lá cây nhị phân tối thiểu được phép).

Nút gốc của cấu trúc QTBT tương ứng với CTU có thể có bốn nút con ở mức thứ nhất của cấu trúc QTBT, mỗi nút có thể được phân chia theo kỹ thuật phân chia cây tứ phân. Tức là, các nút mức thứ nhất là các nút lá (không có nút con) hoặc có bốn nút con. Ví dụ về cấu trúc QTBT 130 đại diện cho các nút như vậy bao gồm nút cha và nút con có các đường nét liền ở các nhánh. Nếu các nút của mức thứ nhất không lớn hơn kích thước

nút gốc cây nhị phân lớn nhất được phép (MaxBTSIZE), thì chúng có thể được cây nhị phân tương ứng phân chia tiếp. Việc tách cây nhị phân của một nút có thể được lặp lại cho đến khi các nút tạo ra từ việc tách đạt đến kích thước nút lá cây nhị phân tối thiểu được phép (MinBTSIZE) hoặc độ sâu cây nhị phân tối đa được phép (MaxBTDepth). Ví dụ về cấu trúc QTBT 130 đại diện cho các nút như vậy có các đường nét đứt ở các nhánh. Nút lá cây nhị phân được gọi là đơn vị mã hóa (CU), đơn vị này được sử dụng để dự đoán (ví dụ, dự đoán nội ảnh hoặc dự đoán liên ảnh) và biến đổi, mà không cần phân chia thêm. Như mô tả ở trên, các CU cũng có thể được gọi là các “khối video” hoặc “các khối”.

Trong một ví dụ về cấu trúc phân chia QTBT, kích thước CTU được thiết lập là 128x128 (mẫu độ chói và hai mẫu sắc độ 64x64 tương ứng), MinQTSIZE được thiết lập là 16x16, MaxBTSIZE được thiết lập là 64x64, MinBTSIZE (cho cả chiều rộng lẫn chiều cao) được thiết lập là 4 và MaxBTDepth được thiết lập là 4. Phân chia cây tứ phân được áp dụng cho CTU trước tiên để tạo ra các nút lá cây tứ phân. Các nút lá cây tứ phân có thể có kích thước từ 16x16 (tức là MinQTSIZE) đến 128x128 (tức là, kích thước CTU). Nếu nút lá cây tứ phân là 128x128, nó sẽ không được tách tiếp bởi cây nhị phân bởi vì kích thước này vượt quá MaxBTSIZE (tức là 64x64 trong ví dụ này). Nếu không thì, nút lá cây tứ phân sẽ được phân chia tiếp bởi cây nhị phân. Do đó, nút lá cây tứ phân cũng là nút gốc cho cây nhị phân và có độ sâu cây nhị phân bằng 0. Khi độ sâu cây nhị phân đạt đến MaxBTDepth (bằng 4 trong ví dụ này), thì không được phép tách thêm nữa. Khi nút cây nhị phân có độ rộng bằng MinBTSIZE (bằng 4 trong ví dụ này), điều này biểu thị rằng không được phép tách ngang thêm nữa. Tương tự, nút cây nhị phân có chiều cao bằng MinBTSIZE biểu thị rằng không được phép tách dọc cho nút cây nhị phân đó thêm nữa. Như nêu trên đây, các nút lá của cây nhị phân được gọi là các CU, và được xử lý thêm theo kỹ thuật dự đoán và biến đổi mà không phân chia thêm nữa.

Trong phần này, các chuẩn mã hóa video, đặc biệt là các kỹ thuật liên quan đến dự đoán vectơ chuyển động của các chuẩn trước đó được mô tả.

Các tiêu chuẩn mã hóa video bao gồm ITU-T H.261, ISO/IEC MPEG-1 Visual, ITU-T H.262 hoặc ISO/IEC MPEG-2 Visual, ITU-T H.263, ISO/IEC MPEG-4 Visual và ITU-T H.264 (cũng được biết đến là ISO/IEC MPEG-4 AVC), bao gồm các phần mở rộng mã hóa video định tỷ lệ được (Scalable video Coding - SVC) và mã hóa video đa khung

hình (Multiview Video Coding - MVC). Bản dự thảo chung mới nhất của MVC được mô tả trong “Mã hóa video nâng cao cho các dịch vụ nghe nhìn chung,” ITU-T Khuyến nghị H.264, Tháng 3 năm 2010.

Ngoài ra, có một chuẩn mã hóa video mới được phát triển, là chuẩn mã hóa dữ liệu video hiệu suất cao (HEVC), được phát triển bởi nhóm hợp tác chung về mã hóa dữ liệu video (Joint Collaboration Team on Video Coding-JCT-VC) của nhóm chuyên gia mã hóa dữ liệu video ITU-T (Video Coding Experts Group-VCEG) và nhóm chuyên gia ảnh động ISO/IEC (Motion Picture Experts Group-MPEG). Bản dự thảo mới đây về HEVC có tại http://phenix.int-evry.fr/jct/doc_end_user/documents/12_Geneva/wg11/JCTVC-L1003-v34.zip.

Thông tin chuyển động

Đối với mỗi khối, tập hợp thông tin chuyển động có thể có sẵn. Tập hợp thông tin chuyển động chứa thông tin chuyển động cho các hướng dự đoán ngược và xuôi. Ở đây hướng dự đoán xuôi và ngược là hai hướng dự đoán tương ứng với danh sách hình tham chiếu 0 (RefPicList0) và danh sách hình tham chiếu 1 (RefPicList1) của hình hoặc lát hiện thời. Các thuật ngữ “xuôi” và “ngược” không nhất thiết có nghĩa về mặt hình học. Thay vào đó, các thuật ngữ này được sử dụng để phân biệt danh sách hình tham chiếu nào mà vector chuyển động dựa vào. Dự đoán xuôi có nghĩa là dự đoán được tạo ra dựa trên danh sách tham chiếu 0, còn dự đoán ngược có nghĩa là dự đoán được tạo ra dựa trên danh sách tham chiếu 1. Trong trường hợp cả danh sách tham chiếu 0 và danh sách tham chiếu 1 được dùng để tạo ra dự đoán cho một khối nhất định, nó được gọi là dự đoán hai hướng.

Đối với hình hoặc lát nhất định, nếu chỉ sử dụng một danh sách hình tham chiếu, thì mọi khối bên trong hình hoặc lát sẽ được dự đoán xuôi. Nếu cả hai danh sách hình tham chiếu đều được sử dụng cho một hình hoặc lát nhất định, thì khối bên trong hình hoặc lát đó có thể được dự đoán xuôi, hoặc dự đoán ngược hoặc dự đoán hai chiều.

Đối với mỗi hướng dự đoán, thông tin chuyển động chứa chỉ số tham chiếu và vector chuyển động. Chỉ số tham chiếu được dùng để nhận dạng hình tham chiếu trong danh sách hình tham chiếu tương ứng (RefPicList0 hoặc RefPicList1). Vector chuyển động có cả thành phần ngang và thành phần dọc, với mỗi thành phần chỉ báo lần lượt giá trị độ lệch theo hướng ngang và dọc. Theo một số mô tả, để đơn giản, từ "vector chuyển động" có thể được

sử dụng thay thế cho thông tin chuyển động, để chỉ báo cả vector chuyển động và chỉ số tham chiếu liên quan của nó.

POC

Đếm thứ tự hình (POC - Picture Order Count) được dùng trong các chuẩn mã hóa video để nhận dạng thứ tự hiển thị của hình. Mặc dù có trường hợp hai hình trong một chuỗi video mã hóa có thể có cùng giá trị POC, nhưng điều này thường không xảy ra trong một chuỗi video mã hóa. Khi nhiều chuỗi video mã hóa có trong một luồng bit, các hình có cùng giá trị POC có thể ở gần nhau hơn về mặt thứ tự giải mã.

Các giá trị POC của các hình thường được dùng để xây dựng danh sách hình tham chiếu, suy ra tập hợp hình tham chiếu như trong HEVC và định tỷ lệ vectơ chuyển động.

Đơn vị dự đoán

Đơn vị dự đoán (PU) đề cập đến đơn vị mẫu cơ bản dùng chung cùng thông tin dự đoán. Trong PU dự đoán liên ảnh, thông tin dự đoán có thể là cùng tập hợp thông tin chuyển động hoặc cùng phương pháp để suy ra thông tin chuyển động cho PU. Trong một ví dụ, PU có thể giống như trong cấu trúc khối HEVC, và cũng có thể là khối cơ bản trong các cấu trúc phân chia khối khác, phân chia cây tứ phân cộng cây nhị phân, cấu trúc chia cây nhiều kiểu, v.v.

PU con là khối con trong PU, trong đó các mẫu dùng chung cùng tập hợp thông tin chuyển động. Trong chuyển động mức PU con, các mẫu trong PU dùng chung cùng phương pháp/mô hình để suy ra thông tin chuyển động, nhưng PU con có thể có tập hợp thông tin chuyển động riêng của nó và có thể khác với PU con khác trong cùng PU.

Fig.3 minh họa lưu đồ ví dụ về suy ra bộ dự đoán vectơ chuyển động, như được mô tả ở đây. Hai danh sách ứng viên dự đoán vectơ chuyển động được suy ra từ thông tin của các khối lân cận trong miền ba chiều. Bộ giải mã chọn danh sách một cách thích ứng theo các phần tử cú pháp đã giải mã *is_sub_pu_motion_flag* từ luồng bit.

Trong một ví dụ, danh sách *sub_pu* chỉ có thể tồn tại nếu khối hiện thời lớn hơn giá trị định trước. Trong một ví dụ khác, khối có thể chỉ chứa duy nhất một PU con, ví dụ như khối 4x4, theo đó cờ là để chỉ báo việc tạo ra thông tin chuyển động có dựa trên phương pháp PU con hay không. Hai danh sách có thể bao gồm số lượng ứng viên dự đoán chuyển

động khác nhau, có thể được xác định trước hoặc được báo hiệu trong SPS, PPS hoặc phần đầu lát. Các khối lân cận có thể là các lân cận không gian trong hình hiện thời hoặc lân cận thời gian trong các hình mã hóa trước đó.

Như được minh họa trên Fig.3, quy trình suy ra hai danh sách sẽ được mô tả. Ở bước 300, bộ giải mã xác định xem cờ có hoạt động hay không. Nếu bộ giải mã xác định cờ không hoạt động, thì bộ giải mã chuyển sang bước 302. Nếu cờ hoạt động, thì bộ giải mã chuyển sang bước 306.

Ở bước 302, danh sách các ứng viên dự đoán chuyển động mức PU thứ nhất được suy ra. Ở bước 306, danh sách các ứng viên dự đoán chuyển động mức PU con thứ hai được suy ra. Ứng viên dự đoán chuyển động mức PU có nghĩa là tất cả các điểm ảnh trong cùng PU dùng chung cùng tập hợp thông tin chuyển động. Ứng viên dự đoán chuyển động mức PU con có nghĩa là tất cả các điểm ảnh trong cùng PU con dùng chung cùng tập hợp thông tin chuyển động, nhưng các PU con khác nhau trong PU có thể có các tập hợp thông tin chuyển động khác nhau. Tập hợp thông tin chuyển động có thể bao gồm hướng dự đoán liên ảnh, chỉ số hình tham chiếu hoặc các chỉ số hình tham chiếu nếu sử dụng nhiều tham chiếu, vector chuyển động, hoặc các vector chuyển động nếu sử dụng nhiều tham chiếu.

Ví dụ về danh sách ứng viên vector chuyển động mức PU là danh sách ứng viên hợp nhất HEVC. Ví dụ về dự đoán chuyển động mức PU con bao gồm, nhưng không bị giới hạn ở, dự đoán vector chuyển động afin (Afin), dự đoán vector chuyển động theo thời gian thay thế (Alternative Temporal Motion Vector Prediction - ATMVP), dự đoán vector chuyển động thời gian-không gian (Spatial-Temporal Motion Vector Prediction - STMVP), dự đoán vector chuyển động phẳng, và suy ra vector chuyển động phù hợp mẫu (Pattern Matched Motion Vector Derivation - PMVD), v.v.. Bảng cú pháp làm ví dụ được thể hiện trong Bảng 1: bên dưới. Phần tử cú pháp `sub_pu_motion_idx` có thể được sử dụng để chỉ báo ứng viên được chọn trong danh sách ứng viên dự đoán chuyển động mức PU con và phần tử cú pháp `pu_motion_idx` được sử dụng để chỉ báo ứng viên được chọn trong danh sách ứng viên dự đoán chuyển động mức PU.

Bảng 1:

	Mô tả
<i>is_sub_pu_motion_flag</i>	ae(v)
if <i>is_sub_pu_motion_flag</i> {	
sub_pu_motion_idx	ae(v)
}	
else {	
pu_motion_idx	ae(v)
}	

Theo một phương pháp khác, các ứng viên vector chuyển động mức PU có thể được chia thành hai nhóm. *is_sub_pu_motion_flag* được báo hiệu khi bộ giải mã nhận cú pháp chỉ báo rằng ứng viên được lựa chọn không thuộc nhóm ứng viên thông tin chuyển động mức PU thứ nhất. *is_sub_pu_motion_flag* có thể được báo hiệu ngầm trong chỉ số hợp nhất mức PU là một trong các chỉ số. Sau đó chỉ số hợp nhất mức PU con được báo hiệu thêm nếu *is_sub_pu_motion_flag* là đúng. Trong một ví dụ, thứ tự chèn danh sách ứng viên dự đoán chuyển động mức PU là $A \rightarrow B \rightarrow S \rightarrow C \rightarrow D \rightarrow E$. A, B, C, D, E biểu thị các ứng viên dự đoán chuyển động mức PU và S là chỉ báo về sub_pu_motion_flag. Nếu S được chọn, thì sub_pu_motion_flag được suy ra là đúng. Một ví dụ về A có thể là ứng viên hợp nhất không gian dưới cùng bên trái, và một ví dụ về B có thể là ứng viên hợp nhất không gian trên cùng bên phải. Lưu ý rằng thứ tự chèn khác cũng có thể được áp dụng.

Lưu ý rằng các ứng viên dự đoán chuyển động mức PU con được nhóm theo tất cả các phương pháp được mô tả ở trên.

Báo hiệu để lựa chọn ứng viên dự đoán vector chuyển động

Phần tử cú pháp *is_sub_pu_motion_flag* trước tiên được dùng để chỉ báo liệu ứng viên dự đoán chuyển động có phải là mức PU con không. *is_sub_pu_motion_flag* có thể được nhị phân hóa bằng cách sử dụng một bin (0/1) và được mã hóa bởi bộ mã hóa số học nhị phân dựa trên ngữ cảnh. Ngữ cảnh có thể phụ thuộc vào kích thước/khu vực PU hoặc

độ sâu PU trong cây phân chia khối. PU lớn hơn có thể có xu hướng lựa chọn dự đoán vector chuyển động mức PU con thường xuyên hơn so với PU nhỏ hơn. Ngưỡng cũng có thể phụ thuộc vào `sub_pu_motion_flag` từ các khối lân cận không gian/thời gian. Cơ hội của PU hiện thời sử dụng chuyển động PU con là cao hơn nếu các khối lân cận có chuyển động PU con.

Nếu `is_sub_pu_motion_flag` là đúng (“1”), thì phần tử cú pháp `sub_pu_motion_idx` được dùng để chỉ báo phương pháp suy ra ứng viên dự đoán chuyển động PU con. Tổng số phương pháp, tức là, tổng số ứng viên dự đoán chuyển động mức PU con, `num_sub_pu_motion` có thể được báo hiệu trong cú pháp mức cao. `sub_pu_motion_idx` có thể được nhị phân hóa bằng cách sử dụng mã đơn phân cắt ngắn tùy thuộc vào `num_sub_pu_motion`. Tuy nhiên, phương pháp nhị phân hóa khác cũng có thể được áp dụng.

Nếu `is_sub_pu_motion_flag` là sai (“0”), thì phần tử cú pháp `nor_pu_motion_idx` được dùng để chỉ báo phương pháp suy ra ứng viên dự đoán chuyển động mức PU. Tổng số ứng viên dự đoán vectơ chuyển động mức PU, `num_nor_pu_motion` có thể được báo hiệu trong cú pháp mức cao. `nor_pu_motion_idx` có thể được nhị phân hóa bằng cách sử dụng mã đơn phân cắt ngắn tùy thuộc vào `num_nor_pu_motion`. Tuy nhiên, phương pháp nhị phân hóa khác cũng có thể được áp dụng.

Suy ra các ứng viên dự đoán chuyển động mức PU

Ứng viên dự đoán chuyển động mức PU có thể được suy ra từ các khối mã hóa lân cận không gian hoặc thời gian một cách tương tự trong HEVC. Trong chế độ hợp nhất HEVC,

Trong HEVC, danh sách ứng viên MV chứa tối đa 5 ứng viên cho chế độ hợp nhất và chỉ 2 ứng viên cho chế độ AMVP. Ứng viên hợp nhất có thể chứa tập hợp thông tin chuyển động, ví dụ, các vectơ chuyển động tương ứng với cả danh sách hình tham chiếu (danh sách 0 và danh sách 1) và các chỉ số tham chiếu. Nếu ứng viên hợp nhất được nhận dạng bằng chỉ số hợp nhất, thì các hình tham chiếu được dùng để dự đoán các khối hiện thời, cũng như các vectơ chuyển động liên quan được xác định.

Dựa vào phân mô tả ở trên, một hoặc nhiều bộ dự đoán vectơ chuyển động được lựa chọn dựa trên chỉ số ứng viên giải mã.

Trong chế độ hợp nhất HEVC, tối đa bốn ứng viên MV không gian 402, 404, 406 và 408 của khối 400 có thể được suy ra theo thứ tự thể hiện trên Fig.4. Thứ tự này là như sau: trái (0, A1), trên (1, B1), phải trên (2, B0), trái dưới (3, A0), và trái trên (4, B2).

Các ứng viên lân cận thời gian trong HEVC sẽ được mô tả ngay sau đây. Ứng viên dự đoán vector chuyển động theo thời gian (temporal motion vector predictor - TMVP), nếu được kích hoạt và có sẵn, được bổ sung vào danh sách ứng viên MV sau các ứng viên vector chuyển động không gian. Vị trí khối chính để suy ra ứng viên TMVP là khối dưới cùng bên phải bên ngoài PU được đặt gần như thể hiện trên Fig.5 dưới dạng khối “T” 500, để bù độ lệch cho các khối bên trên và bên trái được dùng để tạo ra các ứng viên lân cận không gian. Tuy nhiên, nếu khối 502 được đặt bên ngoài hàng CTB hiện thời hoặc thông tin chuyển động là không có sẵn, thì khối được thay thế bằng khối trung tâm 504 của PU. Vector chuyển động của ứng viên TMVP được suy ra từ PU cùng vị trí của hình cùng vị trí, được chỉ báo trong mức lát. Vector chuyển động cho PU cùng vị trí được gọi là MV được đặt gần.

Suy ra các ứng viên dự đoán chuyển động mức PU con

Các ứng viên dự đoán chuyển động mức PU con có thể bao gồm, nhưng không bị giới hạn ở, dự đoán chuyển động afin, dự đoán vector chuyển động theo thời gian thay thế (ATMVP), dự đoán vector chuyển động thời gian-không gian (STMVP), dự đoán vector chuyển động phẳng, suy ra vector chuyển động phù hợp mẫu (PMVD), v.v.. Sau đây, các tác giả sáng chế sẽ minh họa các ví dụ về dự đoán chuyển động mức PU con. Tuy nhiên, một số thay đổi hoặc dự đoán chuyển động mức PU con khác cũng có thể được thêm vào.

Dự đoán chuyển động afin

Trong phương pháp dự đoán chuyển động afin 4 tham số, trường vector chuyển động của khối được mô tả bằng phương trình (1):

$$\begin{cases} v_x = \frac{v_{1x}-v_{0x}}{w}x - \frac{v_{1y}-v_{0y}}{h}y + v_{0x} \\ v_y = \frac{v_{1y}-v_{0y}}{w}x - \frac{v_{1x}-v_{0x}}{h}y + v_{0y} \end{cases} \quad (1)$$

trong đó (w, h) là kích thước của khối, và (x, y) là tọa độ. (v_{0x}, v_{0y}) là vector chuyển động của điểm điều khiển góc trên bên trái, và (v_{1x}, v_{1y}) là vector chuyển động của điểm điều khiển góc trên bên phải.

Trong phương pháp dự đoán chuyển động affin 6 tham số, trường chuyển động của khối được mô tả bằng phương trình (2):

$$\begin{cases} v_x = \frac{v_{1x}-v_{0x}}{w}x - \frac{v_{2x}-v_{0x}}{h}y + v_{0x} \\ v_y = \frac{v_{1y}-v_{0y}}{w}x - \frac{v_{2y}-v_{0y}}{h}y + v_{0y} \end{cases} \quad (2)$$

trong đó ngoài ra, (v_{2x}, v_{2y}) là vector chuyển động của điểm điều khiển góc dưới bên trái.

Trong dự đoán chuyển động affin dựa trên PU con, MV của PU con có thể được suy ra bằng cách tính toán MV tại tâm của PU con. Theo cách khác, tác giả sáng chế có thể định tỷ lệ giảm (w, h) và (x, y) theo kích thước của PU con.

Trong một phương pháp, (v_{ix}, v_{iy}) có thể được suy ra từ khối lân cận đã được mã hóa trước đó bằng chuyển động affin, coi như khối hiện thời dùng chung cùng mô hình chuyển động affin với khối lân cận được mã hóa trước đó.

Trong một phương pháp khác, (v_{ix}, v_{iy}) có thể được suy ra bởi vector chuyển động tại khối mã hóa lân cận. Ví dụ, như được thể hiện trên Fig.6, (v_{0x}, v_{0y}) có thể được suy ra từ vector chuyển động ở khối A 600, B 602, hoặc C 604, (v_{1x}, v_{1y}) có thể được suy ra từ vector chuyển động ở khối C 606 hoặc D 608, (v_{2x}, v_{2y}) có thể được suy ra từ khối E 610 hoặc F 612.

Dự đoán vector chuyển động theo thời gian thay thế

Trong phương pháp dự đoán vector chuyển động theo thời gian thay thế (ATMVP) (hoặc đôi khi được gọi là dự đoán vector chuyển động thời gian cải tiến), dự đoán vector chuyển động theo thời gian (TMVP) được sửa đổi bằng cách tìm nạp nhiều tập hợp thông tin chuyển động (bao gồm các vector chuyển động và các chỉ số tham chiếu) từ các PU con của PU hiện thời.

Dự đoán vector chuyển động thời gian - không gian

Trong phương pháp dự đoán vector chuyển động thời gian-không gian, các vector chuyển động của các PU con được suy ra theo cách đệ quy, theo thứ tự quét mảnh. Fig.7 minh họa khái niệm STMVP. Xét PU 8x8 chứa bốn PU con 4x4 A, B, C và D. Các khối 4x4 lân cận trong khung hiện thời được dán nhãn là a, b, c và d. Việc suy ra chuyển động cho CU con A bắt đầu bằng cách nhận dạng hai khối lân cận không gian của nó. Lân cận thứ nhất là khối $N \times N$ phía trên CU con A (khối c). Nếu khối c này không khả dụng hoặc

được mã hóa nội ảnh, thì các khối NxN khác phía trên CU con A được kiểm tra (từ trái sang phải, bắt đầu ở khối c). Lần cận thứ hai là khối ở bên trái của CU con A (khối b). Nếu khối b không khả dụng hoặc được mã hóa nội ảnh, thì các khối khác ở bên trái của CU con A được kiểm tra (từ trên xuống dưới, bắt đầu ở khối b). Thông tin chuyển động thu được từ các khối lân cận cho mỗi danh sách được định tỷ lệ theo khung tham chiếu thứ nhất cho danh sách đã cho. Tiếp theo, bộ dự đoán vector chuyển động thời gian (TMVP) của khối con A được suy ra bằng cách thực hiện theo cùng thủ tục suy ra TMVP như được chỉ ra trong HEVC. Thông tin chuyển động của khối được đặt gần tại vị trí D được tìm nạp và được định tỷ lệ tương ứng. Cuối cùng, sau khi khôi phục và định tỷ lệ thông tin chuyển động, tất cả các vector chuyển động khả dụng (tối đa 3) được lấy trung bình riêng cho mỗi danh sách tham chiếu. Vector chuyển động trung bình được gán là vector chuyển động của CU con hiện thời.

Suy ra vector chuyển động phù hợp mẫu

Phương pháp suy ra vector chuyển động phù hợp mẫu (Pattern matched motion vector derivation - PMMVD) dựa trên các kỹ thuật chuyển đổi tăng tốc độ khung. Quy trình suy ra chuyển động có hai bước. Tìm kiếm chuyển động mức PU được thực hiện trước tiên, sau đó là hiệu chỉnh chuyển động mức PU con. Ở mức PU, vector chuyển động ban đầu được suy ra cho toàn bộ PU dựa trên việc so khớp hai chiều hoặc so khớp mẫu. Trước tiên, danh sách các ứng viên MV được tạo ra và ứng viên mà dẫn đến hao tổn so khớp tối thiểu được lựa chọn làm điểm khởi đầu cho việc hiệu chỉnh mức PU tiếp theo. Sau đó tìm kiếm cục bộ dựa vào so khớp hai chiều hoặc so khớp mẫu quanh điểm khởi đầu được thực hiện và MV tạo ra hao tổn so khớp nhỏ nhất được lấy làm MV cho toàn bộ CU. Sau đó, thông tin chuyển động được hiệu chỉnh thêm ở mức PU con với các vector chuyển động CU được suy ra làm các điểm khởi đầu.

Khái niệm so khớp hai chiều được minh họa trên Fig.8. So khớp hai chiều được sử dụng để suy ra thông tin chuyển động bằng cách tìm kiếm so khớp gần nhất giữa hai khối dọc theo quỹ đạo chuyển động của khối hiện thời trong hai hình tham chiếu khác nhau. Theo giả thiết về quỹ đạo chuyển động liên tục, các vector chuyển động MV0 và MV1 trở đến hai khối tham chiếu phải tỷ lệ với các khoảng cách thời gian, tức là, TD0 và TD1, giữa hình hiện thời và hai hình tham chiếu. Trong trường hợp đặc biệt, khi hình hiện thời tạm thời ở giữa hai hình tham chiếu và khoảng thời gian từ hình hiện thời đến hai hình tham

chiều giống nhau, thì việc so khớp hai chiều sẽ trở thành MV hai chiều dựa vào đối xứng gương.

So khớp mẫu như được minh họa trên Fig.9 được dùng để suy ra thông tin chuyển động bằng cách tìm kiếm so khớp gần nhất giữa mẫu (các khối lân cận trên cùng và/hoặc bên trái của khối hiện thời) trong hình hiện thời và khối (cùng kích thước với mẫu) trong hình tham chiếu.

Dự đoán vector chuyển động phẳng

Dự đoán vector chuyển động phẳng đạt được bằng cách tính trung bình nội suy tuyến tính ngang và dọc trên cơ sở khối 4x4 như sau.

$$P(x, y) = (H \times P_h(x, y) + W \times P_v(x, y) + H \times W) / (2 \times H \times W)$$

W và H biểu thị chiều rộng và chiều cao của khối, (x, y) là các tọa độ của khối con hiện thời so với khối con góc trên bên trái. Tất cả khoảng cách được biểu thị bằng khoảng cách điểm ảnh chia cho 4. $P(x, y)$ là vector chuyển động của khối con hiện thời.

Dự đoán ngang $P_h(x, y)$ và dự đoán dọc $P_v(x, y)$ cho vị trí (x, y) được tính như sau:

$$P_h(x, y) = (W - 1 - x) \times L(-1, y) + (x + 1) \times R(W, y)$$

$$P_v(x, y) = (H - 1 - y) \times A(x, -1) + (y + 1) \times B(x, H)$$

trong đó $L(-1, y)$ và $R(W, y)$ là các vector chuyển động của các khối 4x4 ở bên trái và bên phải của khối hiện thời. $A(x, -1)$ và $B(x, H)$ là các vector chuyển động của các khối 4x4 ở bên trên và bên dưới của khối hiện thời, như được thể hiện trên Fig.9.

Thông tin chuyển động tham chiếu của các khối lân cận cột trái và hàng trên được suy ra từ các khối lân cận không gian với khối hiện thời.

Thông tin chuyển động tham chiếu của các khối lân cận cột bên phải và hàng dưới cùng được tìm ra như sau.

- 1) Suy ra thông tin chuyển động của khối 4x4 lân cận theo thời gian dưới cùng bên phải
- 2) Tính toán các vector chuyển động của các khối 4x4 lân cận cột bên phải, bằng cách sử dụng thông tin chuyển động đã tìm ra của khối 4x4 lân cận dưới cùng bên phải cùng với thông tin chuyển động của khối 4x4 lân cận trên bên phải, là $R(W, y) = ((H - y - 1) \times AR + (y + 1) \times BR) / H$.
- 3) Tính toán các vector chuyển động của khối 4x4 lân cận hàng dưới cùng, bằng cách sử dụng thông tin chuyển động được suy ra của khối 4x4 lân cận dưới cùng bên phải cùng

với thông tin chuyển động của khối 4x4 lân cận dưới cùng bên trái, là $B(x,H) = ((W-x-1) \times BL + (x+1) \times BR)/W$.

trong đó AR là vector chuyển động của khối 4x4 lân cận không gian phía trên bên phải, BR là vector chuyển động của khối 4x4 lân cận theo thời gian dưới cùng bên phải, và BL là vector chuyển động của khối 4x4 lân cận không gian dưới cùng bên trái, như được thể hiện trên Fig.10.

Thông tin chuyển động thu được từ các khối lân cận cho mỗi danh sách được định tỷ lệ thành hình tam chiếu thứ nhất cho danh sách đã cho.

Chèn danh sách ứng viên

Nếu có, các ứng viên dự đoán chuyển động mức PU con được chèn vào danh sách theo thứ tự mặc định sau: $Afin \rightarrow ATMVP \rightarrow STMVP \rightarrow PMMVD \rightarrow Phẳng$. Thứ tự mặc định có thể được định trước hoặc được báo hiệu dựa trên loại lát, lớp thời gian, mô hình chuyển động afin và/hoặc tính khả dụng của bộ dự đoán chuyển động thời gian. Thứ tự mặc định cũng có thể khác nhau dựa trên loại khối, hình dạng khối hoặc kích thước khối. Số lượng ứng viên dự đoán chuyển động mức PU con tối đa được xác định bằng một số được xác định trước và/hoặc tổng số ứng viên dự đoán chuyển động mức PU con khả dụng như được chỉ báo bằng cú pháp mức cao. Theo một ví dụ, nếu số định trước là 3 và tất cả các ứng viên afin, ATMVP, STMVP, PMMVD và Phẳng là khả dụng, thì số lượng tối đa là 3. Nhưng nếu chỉ có afin và ATMVP là khả dụng, thì số lượng tối đa là 2. Nếu số lớn nhất là 0, thì `is_sub_pu_motion_flag` được suy ra là sai ('0') và không được báo hiệu trong luồng bit.

Thứ tự ứng viên mặc định khác cũng có thể được sử dụng.

Theo một ví dụ, $Afin \rightarrow PMMVD \rightarrow ATMVP \rightarrow STMVP \rightarrow Phẳng$.

Theo một ví dụ khác, $PMMVD \rightarrow Afin \rightarrow ATMVP \rightarrow STMVP \rightarrow Phẳng$.

Theo một ví dụ khác, hai hoặc nhiều ứng viên afin có thể được sử dụng. $Afin1 \rightarrow Afin2 \rightarrow PMMVD \rightarrow ATMVP \rightarrow STMVP \rightarrow Phẳng$

Hoặc $Afin1 \rightarrow ATMVP \rightarrow Afin2 \rightarrow Phẳng$

Sắp xếp lại ứng viên được ưu tiên

Danh sách các ứng viên dự đoán chuyển động mức PU con mặc định có thể được sắp xếp lại dựa trên các lần xuất hiện của chúng trong các khối mã hóa lân cận. Theo một

ví dụ, ứng viên có nhiều lần xuất hiện hơn trong các khối mã hóa lân cận được đưa vào vị trí chỉ số thấp hơn trong danh sách.

Sắp xếp lại ứng viên được ưu tiên một phần

Để giảm độ phức tạp của việc sắp xếp lại các ứng viên, việc sắp xếp lại ứng viên được ưu tiên chỉ được áp dụng cho một hoặc nhiều danh sách con. Ví dụ, ứng viên 1-2, 3-4 theo thứ tự mặc định được sắp xếp lại riêng dựa trên sự xuất hiện của chúng trong các khối mã hóa lân cận.

Cắt tia

Để giảm độ phức tạp, có thể không áp dụng cắt tia hoặc cắt tia từng phần trong danh sách ứng viên PU con. Theo một ví dụ, việc cắt tia giữa ATMVP, STMVP và phẳng có thể được áp dụng, nhưng không áp dụng cắt tia giữa afin và các ứng viên PU con còn lại.

Theo một ví dụ khác, việc cắt tia chỉ được áp dụng cho số lượng PU con nhỏ hơn hoặc bằng giá trị định trước.

Các phương pháp báo hiệu thay thế của ứng viên dự đoán chuyển động PU con

Việc lựa chọn các ứng viên dự đoán chuyển động PU con có thể được báo hiệu bằng cách bật các cờ theo thứ tự chèn ứng viên thay vì chỉ số. Theo một ví dụ, nếu thứ tự chèn là Afin → PMMVD → ATMVP → STMVP → Phẳng, thì việc báo hiệu là như được thể hiện trên Bảng 2:.

Theo một phương pháp thay thế khác, các ứng viên dự đoán chuyển động PU con có thể được nhóm vào một số nhóm con. Ví dụ, Afin, Phẳng được nhóm vào nhóm con 1, ATMVP và STMVP được nhóm vào nhóm con 2, và PMMVD cũng là một nhóm con 3 khác. Báo hiệu làm ví dụ được thể hiện trên Bảng 3:.. Cú pháp `pmmvd_flag` được báo hiệu lần thứ nhất để chỉ báo đó có phải là ứng viên PMMVD hay không. Sau đó, nếu đó không phải là ứng viên PMMVD, thì một phần tử cú pháp khác `sub_group1_flag` được báo hiệu để chỉ báo đó có phải là ứng viên nhóm con 1 hay không. Nếu `sub_group1_flag` là đúng chỉ báo đó là ứng viên nhóm con 1, thì `afin_flag` được báo hiệu để chỉ báo đó có phải là ứng viên Afin hay không. Nếu đó không phải là ứng viên afin, thì `planar_mv_flag` được thiết lập là đúng để chỉ báo đó là ứng viên Phẳng. Nếu `sub_group1_flag` là sai, thì `atmvp_flag` được báo hiệu để chỉ báo liệu đó có phải là ứng viên ATMVP hay không. Nếu đó không phải là ứng viên ATMVP, thì `stmvp_flag` được thiết lập là đúng để chỉ báo đó là ứng viên STMVP.

Bảng 2:

	Mô tả
affine_flag	ae(v)
if(!affine_flag) {	
pmmvd_flag	ae(v)
if(!pmmvd_flag) {	
atmvp_flag	ae(v)
if(!atmvp_flag) {	
stmvp_flag	ae(v)
if(!stmvp_flag) {	
planar_mv_flag = true	
}	
}	
}	
}	

Bảng 3:

	Mô tả
pmmvd_flag	ae(v)
if(!pmmvd_flag) {	
sub_group1_flag	ae(v)
if(sub_group1_flag) {	
affine_flag	ae(v)
if(!affine_flag) {	
planar_mv_flag = true	

}	
} else {	
atmvp_flag	ae(v)
if(!atmvp_flag) {	
stmvp_flag = true	
}	
}	
}	

Lưu ý rằng trong Bảng 2: và Bảng 3:, việc kiểm tra độ khả dụng được bỏ qua để dễ làm ví dụ. Còn được suy ra là sai ('0') nếu ứng viên tương ứng không khả dụng. Cũng lưu ý rằng các thay đổi khác của việc nhóm cũng có thể được áp dụng.

Fig.11 là sơ đồ khối minh họa ví dụ về bộ mã hóa video 200 mà có thể thực hiện các kỹ thuật của sáng chế. Fig.11 được cung cấp cho mục đích giải thích và không nên được coi là làm giới hạn các kỹ thuật như được minh họa và được mô tả rộng rãi trong sáng chế này. Với mục đích giải thích, sáng chế này mô tả bộ mã hóa video 200 trong ngữ cảnh các tiêu chuẩn mã hóa video như tiêu chuẩn mã hóa video HEVC và tiêu chuẩn mã hóa video H.266 đang phát triển. Tuy nhiên, các kỹ thuật của sáng chế không bị giới hạn ở các tiêu chuẩn mã hóa video kể trên, và thường được áp dụng cho việc mã hóa và giải mã dữ liệu video.

Theo ví dụ trên Fig. 11, bộ mã hóa video 200 bao gồm bộ nhớ dữ liệu video 230, đơn vị lựa chọn chế độ 202, đơn vị tạo dư 204, đơn vị xử lý biến đổi 206, đơn vị lượng tử hóa 208, đơn vị lượng tử hóa ngược 202, đơn vị xử lý biến đổi ngược 212, đơn vị tái tạo 214, đơn vị lọc 216, bộ đệm hình giải mã (decoded picture buffer - DPB) 218 và đơn vị mã hóa entropy 220.

Bộ nhớ dữ liệu video 230 có thể lưu trữ dữ liệu video cần được mã hóa bởi các thành phần của bộ mã hóa video 200. Bộ mã hóa video 200 có thể nhận dữ liệu video lưu trữ trong bộ nhớ dữ liệu video 230 từ, ví dụ, nguồn video 104 (Fig.1). DPB 218 có thể hoạt

động như bộ nhớ hình tham chiếu lưu trữ dữ liệu video tham chiếu để sử dụng trong dự đoán dữ liệu video tiếp theo bằng bộ mã hóa video 200. Bộ nhớ dữ liệu video 230 và DPB 218 có thể được tạo ra bởi bất kỳ trong số nhiều thiết bị nhớ, như bộ nhớ truy cập ngẫu nhiên động (dynamic random access memory - DRAM), bao gồm DRAM đồng bộ (synchronous DRAM - SDRAM), RAM từ điện trở (magnetoresistive RAM - MRAM), RAM điện trở (resistive RAM - RRAM), hoặc các loại thiết bị nhớ khác. Bộ nhớ dữ liệu video 230 và DPB 218 có thể được cung cấp bởi cùng thiết bị nhớ hoặc các thiết bị nhớ riêng. Theo các ví dụ khác nhau, bộ nhớ dữ liệu video 230 có thể nằm trên chip cùng các bộ phận khác của bộ mã hóa video 200, như được minh họa, hoặc ngoài chip so với các bộ phận này.

Theo sáng chế, việc tham chiếu đến bộ nhớ dữ liệu video 230 không được hiểu là bị giới hạn ở bộ nhớ bên trong bộ mã hóa video 200, trừ khi được mô tả cụ thể như vậy, hoặc bộ nhớ bên ngoài bộ mã hóa video 200, trừ khi được mô tả cụ thể như vậy. Thay vào đó, tham chiếu đến bộ nhớ dữ liệu video 230 nên được hiểu là bộ nhớ tham chiếu lưu trữ dữ liệu video mà bộ mã hóa video 200 nhận để mã hóa (ví dụ, dữ liệu video cho khôi hiện thời sẽ được mã hóa). Bộ nhớ 106 trên Fig.1 có thể cũng cung cấp lưu trữ tạm thời các đầu ra từ các đơn vị khác nhau của bộ mã hóa video 200.

Các đơn vị khác nhau trên Fig.11 được minh họa để giúp hiểu được các hoạt động được thực hiện bởi bộ mã hóa video 200. Các đơn vị có thể được triển khai như các mạch chức năng cố định, các mạch lập trình được, hoặc kết hợp của chúng. Các mạch chức năng cố định chỉ các mạch có chức năng cụ thể, và thiết lập trước trên các hoạt động mà có thể được thực hiện. Các mạch lập trình được là các mạch mà có thể được lập trình để thực hiện các nhiệm vụ khác nhau, và cung cấp chức năng linh hoạt trong các hoạt động mà có thể được thực hiện. Ví dụ, các mạch lập trình được có thể thực thi phần mềm hoặc firmware khiến các mạch lập trình được hoạt động theo cách được xác định bởi các lệnh phần mềm hoặc firmware. Các mạch chức năng cố định có thể thực thi các lệnh phần mềm (ví dụ, để nhận các tham số hoặc xuất ra các tham số), nhưng các loại hoạt động mà các mạch chức năng cố định thực hiện nói chung là không thay đổi được. Trong một số ví dụ, một hoặc nhiều trong số các đơn vị có thể là các khối mạch riêng (chức năng cố định hoặc có thể lập trình được), và trong một số ví dụ, một hoặc nhiều đơn vị có thể là các mạch tích hợp

Bộ mã hóa video 200 có thể bao gồm đơn vị logic số học (arithmetic logic unit - ALU), đơn vị chức năng cơ bản (elementary function unit - EFU), mạch kỹ thuật số, mạch tương tự và/hoặc lõi có thể lập trình, được tạo ra từ các mạch lập trình được. Trong các ví dụ mà các hoạt động của bộ mã hóa video 200 được thực hiện bằng cách sử dụng phần mềm được thực thi bởi các mạch lập trình được, bộ nhớ 106 (Fig.1) có thể lưu trữ mã đối tượng của phần mềm mà bộ mã hóa video 200 nhận và thực thi, hoặc một bộ nhớ khác trong bộ mã hóa video 200 (không được thể hiện trên hình vẽ) có thể lưu trữ các lệnh như vậy.

Bộ nhớ dữ liệu video 230 được tạo cấu hình để lưu trữ dữ liệu video nhận được. Bộ mã hóa video 200 có thể khôi phục hình của dữ liệu video từ bộ nhớ dữ liệu video 230 và cung cấp dữ liệu video cho đơn vị tạo dư 204 và đơn vị lựa chọn chế độ 202. Dữ liệu video trong bộ nhớ dữ liệu video 230 có thể là dữ liệu video thô cần được mã hóa.

Đơn vị lựa chọn chế độ 202 bao gồm đơn vị ước lượng chuyển động 222, đơn vị bù chuyển động 224 và đơn vị dự đoán nội ảnh 226. Đơn vị lựa chọn chế độ 202 có thể bao gồm các đơn vị chức năng bổ sung để thực hiện dự đoán video theo các chế độ dự đoán khác. Ví dụ, đơn vị lựa chọn chế độ 202 có thể bao gồm đơn vị bảng màu, đơn vị sao chép khối nội ảnh (có thể là một phần của đơn vị ước lượng chuyển động 222 và/hoặc đơn vị bù chuyển động 224), đơn vị afin, đơn vị mô hình tuyến tính (linear model - LM), hoặc tương tự.

Đơn vị lựa chọn chế độ 202 thường phối hợp nhiều bước mã hóa để kiểm tra sự kết hợp của các tham số mã hóa và các giá trị biến dạng tốc độ thu được cho các kết hợp như vậy. Các tham số mã hóa có thể bao gồm phân chia các CTU thành các CU, các chế độ dự đoán cho các CU, các kiểu biến đổi cho dữ liệu dư của các CU, các tham số lượng tử hóa cho dữ liệu dư của các CU, v.v.. Đơn vị lựa chọn chế độ 202 cuối cùng có thể lựa chọn tổ hợp của các tham số mã hóa có giá trị biến dạng tốc độ tốt hơn các tổ hợp đã thử nghiệm khác.

Bộ mã hóa video 200 có thể phân chia hình được khôi phục từ bộ nhớ dữ liệu video 230 thành một loạt các CTU, và đóng gói một hoặc nhiều CTU trong lát. Đơn vị lựa chọn chế độ 210 có thể chia CTU của hình theo cấu trúc cây, chẳng hạn như cấu trúc QTBT hoặc cấu trúc cây tứ phân của HEVC được mô tả ở trên. Như mô tả ở trên, bộ mã hóa video

200 có thể tạo ra một hoặc nhiều CU từ việc phân chia CTU theo cấu trúc cây. CU như vậy cũng có thể được gọi chung là “khối video” hoặc “khối”.

Nói chung, đơn vị lựa chọn chế độ 202 cũng điều khiển các thành phần của chúng (ví dụ, đơn vị ước lượng chuyển động 222, đơn vị bù chuyển động 224, và đơn vị dự đoán nội ảnh 226) để tạo ra khối dự đoán cho khối hiện thời (ví dụ, CU hiện thời, hoặc trong HEVC, phần chồng lấn của PU và TU). Để dự đoán liên ảnh khối hiện thời, đơn vị ước lượng chuyển động 222 có thể thực hiện tìm kiếm chuyển động để nhận dạng một hoặc nhiều khối tham chiếu phù hợp gần nhất trong một hoặc nhiều hình tham chiếu (ví dụ, một hoặc nhiều hình được mã hóa trước đó được lưu trữ trong DPB 218). Cụ thể, đơn vị ước lượng chuyển động 222 có thể tính toán giá trị biểu diễn mức độ tương tự của khối tham chiếu tiềm năng với khối hiện thời, ví dụ, theo tổng chênh lệch tuyệt đối (sum of absolute difference - SAD), tổng chênh lệch bình phương (sum of squared differences - SSD), chênh lệch tuyệt đối trung bình (mean absolute difference - MAD), chênh lệch bình phương trung bình (mean squared difference - MSD), hoặc tương tự. Đơn vị ước lượng chuyển động 222 nói chung có thể thực hiện các tính toán này bằng cách sử dụng các chênh lệch mẫu với mẫu giữa khối hiện thời và khối tham chiếu đang được xem xét. Đơn vị ước lượng chuyển động 222 có thể nhận dạng khối tham chiếu có giá trị thấp nhất thu được từ các phép tính này, chỉ báo khối tham chiếu phù hợp gần nhất với khối hiện thời.

Đơn vị ước lượng chuyển động 222 có thể tạo ra một hoặc nhiều vector chuyển động (MV) mà xác định các vị trí của các khối tham chiếu trong hình tham chiếu so với vị trí của khối hiện thời trong hình hiện thời. Sau đó, đơn vị ước lượng chuyển động 222 có thể cung cấp các vector chuyển động cho đơn vị bù chuyển động 224. Ví dụ, đối với dự đoán liên ảnh một chiều, đơn vị ước lượng chuyển động 222 có thể cung cấp một vector chuyển động, trong khi đối với dự đoán liên ảnh hai chiều, đơn vị ước lượng chuyển động 222 có thể cung cấp hai vector chuyển động. Sau đó, đơn vị bù chuyển động 224 có thể tạo ra khối dự đoán bằng cách sử dụng các vector chuyển động. Ví dụ, đơn vị bù chuyển động 224 có thể khôi phục dữ liệu của khối tham chiếu bằng cách sử dụng vector chuyển động. Một ví dụ khác, nếu vector chuyển động có độ chính xác mẫu phân số, thì đơn vị bù chuyển động 224 có thể nội suy các giá trị cho khối dự đoán theo một hoặc nhiều bộ lọc nội suy. Hơn nữa, đối với dự đoán liên ảnh hai chiều, đơn vị bù chuyển động 224 có thể khôi phục dữ liệu cho hai khối tham chiếu được nhận dạng bởi các vector chuyển động tương ứng và kết

hợp dữ liệu được khôi phục, ví dụ, qua phép lấy trung bình từng mẫu hoặc lấy trung bình có trọng số.

Một ví dụ khác, để dự đoán nội ảnh hoặc mã hóa dự đoán nội ảnh, đơn vị dự đoán nội ảnh 226 có thể tạo ra khối dự đoán từ các mẫu lân cận với khối hiện thời. Ví dụ, đối với các chế độ định hướng, đơn vị dự đoán nội ảnh 226 thường có thể kết hợp về mặt toán học các giá trị của các mẫu lân cận và điền các giá trị được tính toán này theo hướng xác định trên khối hiện thời để tạo ra khối dự đoán. Một ví dụ khác, đối với chế độ DC, đơn vị dự đoán nội ảnh 226 có thể tính trung bình các mẫu lân cận với khối hiện thời và tạo ra khối dự đoán để bao gồm giá trị trung bình thu được cho mỗi mẫu của khối dự đoán.

Đơn vị lựa chọn chế độ 202 cung cấp khối dự đoán cho đơn vị tạo ra khối dư 204. Đơn vị tạo ra khối dư 204 nhận phiên bản thô, không được mã hóa của khối hiện thời từ bộ nhớ dữ liệu video 230 và khối dự đoán từ đơn vị lựa chọn chế độ 202. Đơn vị tạo ra khối dư 204 tính toán các chênh lệch mẫu với mẫu giữa khối hiện thời và khối dự đoán. Chênh lệch mẫu với mẫu thu được xác định khối dư cho khối hiện thời. Trong một số ví dụ, đơn vị tạo dư 204 cũng có thể xác định chênh lệch giữa các giá trị mẫu trong khối dư để tạo ra khối dư bằng cách sử dụng điều chế mã xung vi sai dư (residual differential pulse code modulation - RDPCM). Trong một số ví dụ, đơn vị tạo dư 204 có thể được ra bằng cách sử dụng một hoặc nhiều mạch trừ thực hiện phép trừ nhị phân.

Trong các ví dụ trong đó đơn vị lựa chọn chế độ 202 phân chia các CU thành các PU, mỗi PU có thể được kết hợp với một đơn vị dự đoán độ chói và các đơn vị dự đoán sắc độ tương ứng. Bộ mã hóa video 200 và bộ giải mã video 300 có thể hỗ trợ các PU có các kích thước khác nhau. Như nêu trên đây, kích thước của CU có thể chỉ kích thước khối mã hóa độ chói của CU và kích thước của PU có thể chỉ kích thước của đơn vị dự đoán độ chói của PU. Giả sử kích thước của một CU cụ thể là $2N \times 2N$, bộ mã hóa video 200 có thể hỗ trợ các kích thước PU $2N \times 2N$ hoặc $N \times N$ đối với dự đoán nội ảnh, và các kích thước PU đối xứng $2N \times 2N$, $2N \times N$, $N \times 2N$, $N \times N$, hoặc tương tự đối với dự đoán liên ảnh. Bộ mã hóa video 200 và bộ giải mã video 300 cũng có thể hỗ trợ phân chia không đối xứng các kích thước PU $2N \times nU$, $2N \times nD$, $nL \times 2N$, và $nR \times 2N$ để dự đoán liên ảnh.

Trong các ví dụ trong đó đơn vị lựa chọn chế độ 202 không phân chia thêm CU thành các PU, mỗi CU có thể được kết hợp với một khối mã hóa độ chói và các khối mã

hóa sắc độ tương ứng. Như nêu trên đây, kích thước của CU có thể là kích thước của khối mã hóa độ chói của CU. Bộ mã hóa video 200 và bộ giải mã video 300 có thể hỗ trợ các kích thước CU $2N \times 2N$, $2N \times N$ hoặc $N \times 2N$.

Đối với các kỹ thuật mã hóa video khác, như mã hóa chế độ sao chép khối nội ảnh, mã hóa chế độ afin, và mã hóa chế độ mô hình tuyến tính (linear model - LM), chẳng hạn, đơn vị lựa chọn chế độ 202, qua các đơn vị tương ứng liên quan đến các kỹ thuật mã hóa, tạo ra khối dự đoán cho khối hiện thời được mã hóa. Trong một số ví dụ, như mã hóa chế độ bảng màu, đơn vị lựa chọn chế độ 202 có thể không tạo ra khối dự đoán, và thay vào đó tạo ra các phần tử cú pháp chỉ báo cách thức tái tạo khối dựa vào bảng màu đã lựa chọn. Trong các chế độ như vậy, đơn vị lựa chọn chế độ 202 có thể cung cấp các phần tử cú pháp này cho đơn vị mã hóa entropy 220 cần được mã hóa.

Như mô tả trên đây, đơn vị tạo dư 204 nhận dữ liệu video cho khối hiện thời và khối dự đoán tương ứng. Đơn vị tạo dư 204 sau đó tạo ra khối dư cho khối hiện thời. Để tạo ra khối dư, đơn vị tạo ra khối dư 204 tính toán các chênh lệch mẫu với mẫu giữa khối dự đoán và khối hiện thời. Do đó,

Đơn vị xử lý biến đổi 206 áp dụng một hoặc nhiều phép biến đổi cho khối dư để tạo ra khối các hệ số biến đổi (ở đây được gọi là "khối hệ số biến đổi"). Đơn vị xử lý biến đổi 206 có thể áp dụng các phép biến đổi khác nhau cho khối dư để tạo ra khối hệ số biến đổi. Ví dụ, đơn vị xử lý biến đổi 206 có thể áp dụng phép biến đổi cosin rời rạc (discrete cosine transform - DCT), biến đổi có hướng, biến đổi Karhunen-Loeve (Karhunen-Loeve transform - KLT) hoặc biến đổi tương tự về mặt khái niệm cho khối dư. Trong một số ví dụ, đơn vị xử lý biến đổi 206 có thể thực hiện nhiều phép biến đổi cho khối dư, ví dụ, biến đổi sơ cấp và biến đổi thứ cấp, chẳng hạn như biến đổi quay. Trong một số ví dụ, đơn vị xử lý biến đổi 206 không áp dụng các phép biến đổi cho khối dư.

Đơn vị lượng tử hóa 208 có thể lượng tử hóa các hệ số biến đổi trong khối hệ số biến đổi, để tạo ra khối hệ số biến đổi lượng tử hóa. Đơn vị lượng tử hóa 208 có thể lượng tử hóa các hệ số biến đổi của khối hệ số biến đổi theo giá trị tham số lượng tử hóa (quantization parameter - QP) liên quan đến khối hiện thời. Bộ mã hóa video 200 (ví dụ, qua đơn vị lựa chọn chế độ 202) có thể điều chỉnh mức lượng tử hóa áp dụng cho các khối hệ số liên quan đến khối hiện thời bằng cách điều chỉnh giá trị QP liên quan đến CU. Việc

lượng tử hóa có thể làm mất thông tin, và do đó, các hệ số biến đổi lượng tử hóa có thể có độ chính xác thấp hơn các hệ số biến đổi gốc được tạo ra bởi đơn vị xử lý biến đổi 206.

Đơn vị lượng tử hóa ngược 210 và đơn vị xử lý biến đổi ngược 212 có thể áp dụng lần lượt lượng tử hóa ngược và biến đổi ngược cho khối hệ số biến đổi được lượng tử hóa để tái tạo khối dư từ khối hệ số biến đổi. Đơn vị tái tạo 214 có thể tạo ra khối tái tạo tương ứng với khối hiện thời (mặc dù có khả năng bị biến dạng ở mức độ nào đó) dựa trên khối dư được tái tạo và khối dự đoán được tạo ra bởi đơn vị lựa chọn chế độ 202. Ví dụ, đơn vị tái tạo 214 có thể thêm các mẫu của khối dư tái tạo vào các mẫu tương ứng từ khối dự đoán được tạo ra bởi đơn vị lựa chọn chế độ 202 để tạo ra khối tái tạo.

Đơn vị lọc 216 có thể thực hiện một hoặc nhiều hoạt động lọc trên các khối tái tạo. Ví dụ, đơn vị lọc 216 có thể thực hiện các hoạt động tách khối để giảm các xảo ảnh dạng khối dọc theo các cạnh của các CU. Trong một số ví dụ, các hoạt động của đơn vị lọc 216 có thể được bỏ qua.

Bộ mã hóa video 200 lưu trữ các khối tái tạo vào DPB 218. Ví dụ, trong các ví dụ trong đó không cần đến hoạt động của đơn vị lọc 216, đơn vị tái tạo 216 có thể lưu trữ các khối tái tạo vào DPB 218. Trong các ví dụ trong đó cần đến các hoạt động của đơn vị lọc 224, đơn vị lọc 216 có thể lưu trữ các khối tái tạo đã lọc vào DPB 218. Đơn vị ước lượng chuyển động 222 và đơn vị bù chuyển động 224 có thể khôi phục hình tham chiếu từ DPB 218, được tạo ra từ các khối tái tạo (và có thể được lọc), để dự đoán liên ảnh các khối của các hình được mã hóa sau đó. Ngoài ra, đơn vị dự đoán nội ảnh 226 có thể sử dụng các khối tái tạo trong DPB 218 của hình hiện thời để dự đoán nội ảnh các khối khác trong hình hiện thời.

Nói chung, đơn vị mã hóa entropy 220 có thể mã hóa entropy các phần tử cú pháp nhận được từ các thành phần chức năng khác của bộ mã hóa video 200. Ví dụ, đơn vị mã hóa entropy 220 có thể mã hóa entropy các khối hệ số biến đổi lượng tử hóa từ đơn vị lượng tử hóa 208. Trong một ví dụ khác, đơn vị mã hóa entropy 220 có thể mã hóa entropy các phần tử cú pháp dự đoán (ví dụ, thông tin chuyển động để dự đoán liên ảnh hoặc thông tin chế độ nội ảnh để dự đoán nội ảnh) từ đơn vị lựa chọn chế độ 202. Đơn vị mã hóa entropy 220 có thể thực hiện một hoặc nhiều hoạt động mã hóa entropy trên các phần tử cú pháp, mà là một ví dụ khác về dữ liệu video, để tạo ra dữ liệu được mã hóa entropy. Ví

dụ, đơn vị mã hóa entropy 220 có thể thực hiện hoạt động lập mã có độ dài thay đổi thích ứng theo ngữ cảnh (context-adaptive variable length coding - CAVLC), hoạt động CABAC, hoạt động lập mã có độ dài biến đổi đến biến đổi (variable-to-variable - V2V), hoạt động lập mã số học nhị phân thích ứng theo ngữ cảnh dựa trên cú pháp (syntax-based context-adaptive binary arithmetic coding - SBAC), hoạt động lập mã entropy phân vùng khoảng xác suất (Probability Interval Partitioning Entropy - PIPE), hoạt động mã hóa hàm mũ Golomb, hoặc loại hoạt động mã hóa entropy khác trên dữ liệu. Trong một số ví dụ, đơn vị mã hóa entropy 220 có thể hoạt động ở chế độ bỏ qua trong đó các phần tử cú pháp không được mã hóa entropy.

Bộ mã hóa video 200 có thể xuất ra luồng bit bao gồm các phần tử cú pháp mã hóa entropy cần thiết để tái tạo các khối của lát hoặc hình. Cụ thể, đơn vị mã hóa entropy 220 có thể xuất ra luồng bit

Các hoạt động mô tả trên đây được mô tả liên quan đến khối. Sự mô tả như vậy nên được hiểu là các hoạt động đối với khối mã hóa độ chói và/hoặc các khối mã hóa sắc độ. Như được mô tả trên đây, trong một số ví dụ, khối mã hóa độ chói và các khối mã hóa sắc độ là các thành phần độ chói và sắc độ của CU. Trong một số ví dụ, khối mã hóa độ chói và các khối mã hóa sắc độ là các thành phần độ chói và sắc độ của PU.

Trong một số ví dụ, các hoạt động được thực hiện đối với khối mã hóa độ chói không cần phải lặp lại đối với các khối mã hóa sắc độ. Ví dụ, các hoạt động để nhận dạng vectơ chuyển động (MV) và hình tham chiếu cho khối mã hóa độ chói không cần phải lặp lại để nhận dạng MV và hình tham chiếu cho các khối sắc độ. Thay vào đó, MV cho khối mã hóa độ chói có thể được định tỷ lệ để xác định MV cho các khối sắc độ, và hình tham chiếu có thể giống nhau. Một ví dụ khác, quy trình dự đoán nội ảnh có thể giống nhau đối với các khối mã hóa độ chói và các khối mã hóa sắc độ.

Bộ mã hóa video 200 là ví dụ về thiết bị được tạo cấu hình để mã hóa dữ liệu video bao gồm bộ nhớ được tạo cấu hình để lưu trữ dữ liệu video, và một hoặc nhiều đơn vị xử lý được triển khai trong mạch và được tạo cấu hình để suy ra mẫu chuyển đổi tăng tốc độ khung (frame rate up conversion - FRUC) sử dụng giá trị điểm ảnh dự đoán và thực hiện kỹ thuật suy ra vectơ chuyển động phía bộ giải mã bằng cách sử dụng mẫu được suy ra. Trong một ví dụ khác, bộ giải mã video 300 có thể được tạo cấu hình để xác định các vectơ

chuyển động tương ứng từ một hoặc nhiều khối dữ liệu video lân cận và suy ra vector chuyển động cho khối dữ liệu video hiện thời bằng cách sử dụng các vector chuyển động tương ứng từ một hoặc nhiều khối dữ liệu video lân cận. Trong một ví dụ khác, bộ giải mã video 300 có thể được tạo cấu hình để xác định các vector chuyển động tương ứng từ một hoặc nhiều khối dữ liệu video được đặt gần nhau, và suy ra vector chuyển động cho khối dữ liệu video hiện thời bằng cách sử dụng các vector chuyển động tương ứng từ một hoặc nhiều khối dữ liệu video đặt gần nhau.

Fig.12 là sơ đồ khối minh họa ví dụ về bộ giải mã video 300 có thể thực hiện các kỹ thuật theo sáng chế. Fig.12 được cung cấp cho mục đích giải thích và không hạn chế ở các kỹ thuật như được minh họa và mô tả rộng rãi trong sáng chế. Với mục đích giải thích, sáng chế này mô tả bộ giải mã video 300 được mô tả theo các kỹ thuật JEM, H.266 / VVC và HEVC. Tuy nhiên, các kỹ thuật của sáng chế có thể được thực hiện bởi các thiết bị mã hóa video được tạo cấu hình theo các tiêu chuẩn mã hóa video khác.

Trong ví dụ trên Fig.12, bộ giải mã video 300 bao gồm bộ nhớ đệm hình mã hóa (coded picture buffer - CPB) 320, đơn vị giải mã entropy 302, đơn vị xử lý dự đoán 304, đơn vị lượng tử hóa ngược 306, đơn vị xử lý biến đổi ngược 308, đơn vị tái tạo 310, đơn vị lọc 312 và bộ đệm hình giải mã (DPB) 314. Đơn vị xử lý dự đoán 304 bao gồm đơn vị bù chuyển động 316 và đơn vị dự đoán nội ảnh 318. Đơn vị xử lý dự đoán 304 có thể bao gồm các đơn vị bổ sung để thực hiện dự đoán theo các chế độ dự đoán khác. Ví dụ, đơn vị xử lý dự đoán 304 có thể bao gồm đơn vị bảng màu, đơn vị sao chép khối nội ảnh (có thể tạo thành một phần của đơn vị bù chuyển động 316), đơn vị afin, đơn vị mô hình tuyến tính (linear model - LM), hoặc tương tự. Trong các ví dụ khác, bộ giải mã video 300 có thể bao gồm nhiều hơn, ít hơn hoặc các thành phần chức năng khác.

Bộ nhớ CPB 320 có thể lưu trữ dữ liệu video, ví dụ như luồng bit video mã hóa, cần được giải mã bởi các thành phần của bộ giải mã video 300. Dữ liệu video lưu trữ trong bộ nhớ CPB 320 có thể thu được, ví dụ, từ phương tiện đọc được bằng máy tính 110 (Fig.1). Bộ nhớ CPB 320 có thể bao gồm CPB lưu trữ dữ liệu video mã hóa (ví dụ, các phần tử cú pháp) từ luồng bit video mã hóa. Ngoài ra, bộ nhớ CPB 320 có thể lưu trữ dữ liệu video ngoài các phần tử cú pháp của hình mã hóa, chẳng hạn như dữ liệu tạm thời thể hiện các đầu ra từ các đơn vị khác nhau của bộ giải mã video 300. DPB 314 thường lưu trữ các hình

đã giải mã mà bộ giải mã video 300 có thể xuất ra và/hoặc sử dụng làm dữ liệu video tham chiếu khi giải mã dữ liệu hoặc các hình tiếp theo của luồng bit video mã hóa. Bộ nhớ CPB 320 và DPB 314 có thể được tạo ra bởi thiết bị bất kỳ trong số nhiều thiết bị nhớ, như DRAM, bao gồm SDRAM, RAM từ trở (MRAM), RAM điện trở (RRAM), hoặc các loại thiết bị nhớ khác. Bộ nhớ CPB 320 và DPB 314 có thể được cung cấp bởi cùng một thiết bị nhớ hoặc các thiết bị nhớ riêng. Trong các ví dụ khác nhau, bộ nhớ CPB 320 có thể nằm trên chip cùng các bộ phận khác của bộ giải mã video 300, hoặc không nằm trên chip so với các bộ phận đó.

Ngoài ra hoặc theo cách khác, trong một số ví dụ, bộ giải mã video 300 có thể khôi phục dữ liệu video mã hóa từ bộ nhớ 120 (Fig.1). Tức là, bộ nhớ 120 có thể lưu trữ dữ liệu như mô tả trên đây vào bộ nhớ CPB 320. Tương tự, bộ nhớ 120 có thể lưu trữ các lệnh cần được thực thi bởi bộ giải mã video 300, khi một số hoặc tất cả các chức năng của bộ giải mã video 300 được triển khai trong phần mềm được thực thi bởi mạch xử lý của bộ giải mã video 300.

Các đơn vị khác nhau được thể hiện trên Fig.12 được minh họa để giúp hiểu được các hoạt động được thực hiện bởi bộ giải mã video 300. Các đơn vị có thể được triển khai dưới dạng các mạch chức năng cố định, các mạch lập trình được, hoặc kết hợp của chúng. Tương tự với Fig.11, các mạch chức năng cố định là các mạch có chức năng cụ thể, và được cài đặt trước trong các hoạt động mà có thể được thực hiện. Các mạch lập trình được là các mạch mà có thể lập trình để thực hiện các nhiệm vụ khác nhau, và có chức năng linh hoạt trong các hoạt động mà có thể được thực hiện. Ví dụ, các mạch lập trình được có thể thực thi phần mềm hoặc firmware khiến các mạch lập trình được hoạt động theo cách được xác định bởi các lệnh phần mềm hoặc firmware. Các mạch chức năng cố định có thể thực thi các lệnh phần mềm (ví dụ, để nhận các tham số hoặc xuất ra các tham số), nhưng các loại hoạt động mà các mạch chức năng cố định thực hiện nói chung là không thay đổi được. Trong một số ví dụ, một hoặc nhiều trong số các đơn vị có thể là các khối mạch khác nhau (chức năng cố định hoặc có thể lập trình được), và trong một số ví dụ, một hoặc nhiều đơn vị có thể là các mạch được tích hợp

Bộ giải mã video 300 có thể bao gồm các ALU, EFU, mạch kỹ thuật số, mạch tương tự và/hoặc các lõi lập trình được tạo ra từ mạch có thể lập trình. Trong các ví dụ trong đó

các hoạt động của bộ giải mã video 300 được thực hiện bởi phần mềm thực thi trên các mạch lập trình được, bộ nhớ trên chip hoặc ngoài chip có thể lưu trữ các lệnh (ví dụ, mã đối tượng) của phần mềm mà bộ giải mã video 300 nhận và thực thi.

Đơn vị giải mã entropy 302 có thể nhận dữ liệu video mã hóa từ CPB và giải mã entropy dữ liệu video để tái tạo các phần tử cú pháp. Đơn vị xử lý dự đoán 304, đơn vị lượng tử hóa ngược 306, đơn vị xử lý biến đổi ngược 308, đơn vị tái tạo 310, và đơn vị lọc 312 có thể tạo ra dữ liệu video đã giải mã dựa vào các phần tử cú pháp trích ra từ luồng bit.

Nói chung, bộ giải mã video 300 tái tạo hình trên cơ sở từng khối. Bộ giải mã video 300 có thể thực hiện hoạt động tái tạo trên từng khối riêng (trong đó khối hiện đang được tái tạo, tức là được giải mã, có thể được gọi là “khối hiện thời”).

Đơn vị giải mã entropy 302 có thể giải mã entropy các phần tử cú pháp xác định hệ số biến đổi lượng tử hóa của khối hệ số biến đổi lượng tử hóa, cũng như thông tin biến đổi, chẳng hạn như tham số lượng tử hóa (quantization parameter - QP) và/hoặc (các) chỉ báo chế độ biến đổi. Đơn vị lượng tử hóa ngược 306 có thể sử dụng QP liên quan đến khối hệ số biến đổi lượng tử hóa để xác định mức lượng tử hóa và, tương tự, mức lượng tử hóa ngược cho đơn vị lượng tử hóa ngược 306 để áp dụng. Ví dụ, đơn vị lượng tử hóa ngược 306 có thể thực hiện phép toán dịch trái theo bit để lượng tử hóa ngược các hệ số biến đổi lượng tử hóa. Do đó, đơn vị lượng tử hóa ngược 306 có thể tạo ra khối hệ số biến đổi bao gồm các hệ số biến đổi.

Sau khi đơn vị lượng tử hóa ngược 306 tạo ra khối hệ số biến đổi, đơn vị xử lý biến đổi ngược 308 có thể áp dụng một hoặc nhiều phép biến đổi ngược cho khối hệ số biến đổi để tạo ra khối dư liên quan đến khối hiện thời. Ví dụ, đơn vị xử lý biến đổi ngược 308 có thể áp dụng DCT ngược, biến đổi số nguyên ngược, biến đổi Karhunen-Loeve (Karhunen-Loeve transform - KLT) ngược, biến đổi quay ngược, biến đổi hướng ngược hoặc một biến đổi ngược khác cho khối hệ số.

Hơn nữa, đơn vị xử lý dự đoán 304 tạo ra khối dự đoán theo các phần tử cú pháp thông tin dự đoán đã được giải mã entropy bởi đơn vị giải mã entropy 302. Ví dụ, nếu các phần tử cú pháp thông tin dự đoán chỉ báo rằng khối hiện thời được dự đoán liên ảnh, thì đơn vị bù chuyển động 316 có thể tạo ra khối dự đoán. Trong trường hợp này, các phần tử cú pháp thông tin dự đoán có thể chỉ báo hình tham chiếu trong DPB 314 mà từ đó suy ra

khối tham chiếu, cũng như vector chuyển động nhận dạng vị trí của khối tham chiếu trong hình tham chiếu liên quan đến vị trí của khối hiện thời trong hình hiện thời. Đơn vị bù chuyển động 316 thường có thể thực hiện quy trình dự đoán liên ảnh theo cách thức về cơ bản tương tự như quy trình được mô tả liên quan đến đơn vị bù chuyển động 224 (Fig. 11).

Theo một ví dụ khác, nếu các phần tử cú pháp thông tin dự đoán chỉ báo rằng khối hiện thời được dự đoán nội ảnh, thì đơn vị dự đoán nội ảnh 318 có thể tạo ra khối dự đoán theo chế độ dự đoán nội ảnh được chỉ báo bởi các phần tử cú pháp thông tin dự đoán. Một lần nữa, đơn vị dự đoán nội ảnh 318 thường có thể thực hiện quy trình dự đoán nội ảnh theo cách thức về cơ bản tương tự như quy trình được mô tả liên quan đến đơn vị dự đoán nội ảnh 226 (Fig.11). Đơn vị dự đoán nội ảnh 318 có thể khôi phục dữ liệu của các mẫu lân cận cho khối hiện thời từ DPB 314.

Đơn vị tái tạo 310 có thể tái tạo khối hiện thời bằng cách sử dụng khối dự đoán và khối dư. Ví dụ, đơn vị tái tạo 310 có thể thêm các mẫu của khối dư vào các mẫu tương ứng của khối dự đoán để tái tạo khối hiện thời.

Đơn vị lọc 312 có thể thực hiện một hoặc nhiều hoạt động lọc trên các khối tái tạo. Ví dụ, đơn vị lọc 312 có thể thực hiện các hoạt động tách khối để giảm các xảo ảnh dạng khối dọc theo các cạnh của các khối tái tạo. Các hoạt động của đơn vị lọc 312 không cần thiết được thực hiện trong tất cả các ví dụ.

Bộ giải mã video 300 có thể lưu trữ các khối tái tạo vào DPB 314. Như đã mô tả trên đây, DPB 314 có thể cung cấp thông tin tham chiếu, chẳng hạn như các mẫu của hình hiện thời để dự đoán nội ảnh và các hình giải mã trước đó để bù chuyển động tiếp theo, cho đơn vị xử lý dự đoán 304. Hơn nữa, bộ giải mã video 300 có thể xuất ra hình giải mã từ DPB để trình bày sau trên thiết bị hiển thị, chẳng hạn như thiết bị hiển thị 118 trên Fig.1.

Theo cách này, bộ giải mã video 300 là ví dụ về thiết bị giải mã video bao gồm bộ nhớ được tạo cấu hình để lưu trữ dữ liệu video, và một hoặc nhiều đơn vị xử lý được triển khai trong mạch và được tạo cấu hình để triển khai các phương pháp và quy trình được mô tả ở đây.

Ví dụ, dữ liệu video mã hóa có thể bao gồm cú pháp sau:

Bảng 4. Cú pháp dữ liệu hợp nhất

merge_data(x0, y0, cbWidth, cbHeight) {	Mô tả
mmvd_flag[x0][y0]	ae(v)
if(mmvd_flag[x0][y0] == 1) {	
mmvd_merge_flag[x0][y0]	ae(v)
mmvd_distance_idx[x0][y0]	ae(v)
mmvd_direction_idx[x0][y0]	ae(v)
} else {	
if(MaxNumSubblockMergeCand > 0 && cbWidth >= 8 && cbHeight >= 8)	
merge_subblock_flag[x0][y0]	ae(v)
if(merge_subblock_flag[x0][y0] == 1) {	
if(MaxNumSubblockMergeCand > 1)	
merge_subblock_idx[x0][y0]	ae(v)
} else {	
if(sps_ciip_enabled_flag && cu_skip_flag[x0][y0] == 0 && (cbWidth * cbHeight) >= 64 && cbWidth < 128 && cbHeight < 128) {	
ciip_flag[x0][y0]	ae(v)
if(ciip_flag[x0][y0]) {	
if(cbWidth <= 2 * cbHeight cbHeight <= 2 * cbWidth)	
ciip_luma_mpm_flag[x0][y0]	ae(v)
if(ciip_luma_mpm_flag[x0][y0])	
ciip_luma_mpm_idx[x0][y0]	ae(v)
}	

}	
if(sps_triangle_enabled_flag && tile_group_type == B && cbWidth * cbHeight >= 64)	
merge_triangle_flag[x0][y0]	ae(v)
if(merge_triangle_flag[x0][y0])	
merge_triangle_idx[x0][y0]	ae(v)
else if(MaxNumMergeCand > 1)	
merge_idx[x0][y0]	ae(v)
}	
}	
}	

merge_subblock_flag[x0][y0] chỉ ra liệu các tham số dự đoán liên ảnh dựa trên khối con cho đơn vị mã hóa hiện thời có được suy ra từ các khối lân cận hay không. Các chỉ số mảng x0, y0 chỉ ra vị trí (x0, y0) của mẫu độ chói trên cùng bên trái của khối mã hóa được xem xét so với mẫu độ chói trên cùng bên trái của hình. Khi không có merge_subblock_flag [x0] [y0], nó được coi là bằng 0.

merge_subblock_idx[x0] [y0] chỉ ra chỉ số ứng viên hợp nhất của danh sách ứng viên hợp nhất dựa trên khối con trong đó x0, y0 chỉ ra vị trí (x0, y0) của mẫu độ chói trên cùng bên trái của khối mã hóa được xem xét so với mẫu độ chói trên cùng bên trái của hình.

Khi không có merge_subblock_idx [x0] [y0], nó được coi là bằng 0.

Nếu merge_subblock_flag [xCb] [yCb] bằng 1, thì quy trình suy ra các vector chuyển động và các chỉ số tham chiếu trong chế độ hợp nhất khối con như được chỉ ra trong 8.4.4.2 được gọi với vị trí khối mã hóa độ chói (xCb, yCb), chiều rộng khối mã hóa độ chói cbWidth và chiều cao khối mã hóa độ chói cbHeight là đầu vào, số lượng khối con mã hóa độ chói theo hướng ngang numSbX và theo hướng dọc numSbY, chỉ số tham chiếu

refIdxL0, refIdxL1, mảng cờ sử dụng danh sách dự đoán predFlagL0 [xSbIdx] [ySbIdx] và predFlagL1 [xSbIdx] [ySbIdx], mảng vector chuyển động khối con độ chói mvL0 [xSbIdx] [ySbIdx] và mảng vector chuyển động khối con sắc độ mvCL0 [xSbIdx] [ySbIdx] và mvCL1 [xSbIdx] [ySbIdx], với xSbIdx = 0 .. numSbX - 1, ySbIdx = 0 .. numSbY - 1 và chỉ số trọng số dự đoán hai chiều gbiIdx là đầu ra.

Cần lưu ý rằng tùy thuộc vào ví dụ, một số hoạt động hoặc sự kiện của kỹ thuật bất kỳ trong số các kỹ thuật được mô tả trong sáng chế này có thể được thực hiện theo các trình tự khác nhau, có thể được thêm, hợp nhất, hoặc bỏ 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). Ngoài ra, trong một số ví dụ, các hoạt động hoặc sự kiện có thể được thực hiện đồng thời, ví dụ, thông qua việc xử lý đa xâu chuỗi, xử lý ngắt, hoặc nhiều bộ xử lý, thay vì thực hiện tuần tự.

Trong một hoặc nhiều ví dụ, các chức năng được mô tả có thể được triển khai dưới dạng phần cứng, phần mềm, firmware, hoặc kết hợp bất kỳ của chúng. Nếu được thực hiện trong phần mềm, các chức năng có thể được lưu trữ trên hoặc được truyền dưới dạng một hoặc nhiều lệnh hoặc mã trên phương tiện đọc được bằng máy tính và được thực thi bởi đơn vị xử lý dựa vào phần cứng. Phương tiện đọc được bằng máy tính có thể bao gồm các 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 phương tiện bất kỳ hỗ trợ truyền chương trình máy tính từ nơi này đến nơi khác, ví dụ, theo giao thức truyền thông. Theo cách này, phương tiện đọc được bằng máy tính thường có thể tương ứng với (1) phương tiện lưu trữ hữu hình đọc được bằng máy tính là phương tiện bất biến hoặc (2) phương tiện truyền thông như tín hiệu hoặc sóng mang. Phương tiện lưu trữ dữ liệu có thể là phương tiện khả dụng bất kỳ mà có thể được truy cập bởi một hoặc nhiều máy tính hoặc một hoặc nhiều bộ xử lý để truy hồi các lệnh, mã và/hoặc các cấu trúc dữ liệu để triển khai các kỹ thuật được mô tả trong sáng chế. Sản phẩm chương trình máy tính có thể bao gồm phương tiện đọc được bằng máy tính.

Ví dụ, và không giới hạn, các phương tiện lưu trữ đọc được bằng máy tính như vậy có thể bao gồm RAM, ROM, EEPROM, (CD-ROM hoặc các thiết bị lưu trữ đĩa quang khác, thiết bị lưu trữ đĩa từ, hoặc các thiết bị lưu trữ từ tính khác, bộ nhớ nhanh, hoặc

phương tiện khác bất kỳ mà có thể được dùng để lưu trữ mã chương trình mong muốn ở dạng các lệnh hoặc các cấu trúc dữ liệu và có thể được truy cập bởi máy tính. Ngoài ra, kết nối bất kỳ 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 dây xoắn, đường thuê bao kỹ thuật số (digital subscriber line - DSL) hoặc các công nghệ không dây như hồng ngoại, vô tuyến, và vi ba, thì cáp đồng trục, cáp sợi quang, cáp dây xoắn, DSL, hoặc các công nghệ không dây như hồng ngoại, vô tuyến và vi ba được bao gồm trong định nghĩa về phương tiện. Tuy nhiên, cần hiểu rằng phương tiện lưu trữ đọc được bằng máy tính và phương tiện lưu dữ liệu không bao gồm kết nối, sóng mang, tín hiệu hoặc phương tiện tạm thời khác, nhưng thay vào đó được hướng đến phương tiện hữu hình, bất biến. Đĩa từ và đĩa quang, như được sử dụng ở đây, bao gồm đĩa compac (CD), đĩa laser, đĩa quang, đĩa số đa năng (DVD - Digital Versatile Disc), đĩa mềm và đĩa bluray, 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 laser. Dạng kết hợp của các loại phương tiện nêu trên cũng được bao gồm trong phạm vi của phương tiện đọc được bằng máy tính.

Các lệnh có thể được thực thi bởi một hoặc nhiều bộ xử lý, chẳng hạn như một hoặc nhiều bộ xử lý tín hiệu số (digital signal processor - DSP), các bộ vi xử lý đa năng, mạch tích hợp chuyên dụng (application specific integrated circuit - ASIC), mảng công lập trình được theo trường (field programmable gate array - FPGA), hoặc các hệ mạch logic tích hợp hoặc rời rạc tương đương khác. Do vậy, thuật ngữ “bộ xử lý”, như được sử dụng ở đây có thể chỉ cấu trúc bất kỳ trong số các cấu trúc nêu trên hoặc cấu trúc khác bất kỳ phù hợp để triển khai các kỹ thuật được mô tả ở đây. Ngoài ra, theo một số khía cạnh, chức năng nêu trong sáng chế có thể được cung cấp trong mô đun phần cứng và/hoặc phần mềm chuyên dụng được tạo cấu hình để mã hóa và giải mã, hoặc được kết hợp thành bộ mã hoá-giải mã kết hợp. Hơn nữa, các kỹ thuật có thể được triển khai hoàn toàn trong một hoặc nhiều mạch hoặc phần tử logic.

Các kỹ thuật của sáng chế có thể được triển khai trong nhiều loại máy hoặc thiết bị khác nhau, bao gồm máy cầm tay không dây, mạch tích hợp (integrated circuit - IC) hoặc bộ IC (ví dụ, bộ chip). Các thành phần, mô đun hoặc đơn vị khác nhau 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 được bộc lộ, nhưng không nhất thiết phải được thực hiện bằng các đơn vị phần cứng khác nhau. Thay vào đó, như được mô tả trên đây, các đơn vị khác nhau có thể được kết hợp trong đơn vị phần cứng bộ mã hóa-giải mã hoặc được cung cấp bởi tập hợp các đơn vị phần cứng hoạt động liên kết, bao gồm một hoặc nhiều bộ xử lý như được mô tả trên đây, kết hợp với phần mềm và/hoặc firmware thích hợp

Các ví dụ khác nhau đã được mô tả. Những ví dụ này và những ví dụ khác nằm trong phạm vi của phần yêu cầu bảo hộ sau.

YÊU CẦU BẢO HỘ

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

nhận dữ liệu video được mã hóa, dữ liệu video được mã hóa bao gồm đơn vị dự đoán (prediction unit - PU);

phân tích cú pháp, đối với PU, phân tử cú pháp cấp độ PU từ dữ liệu video được mã hóa, trong đó phân tử cú pháp cấp độ PU là cờ chuyển động PU con;

để đáp lại việc xác định cờ chuyển động PU con là hoạt động:

phân tích phân tử cú pháp chỉ số lựa chọn ứng viên dự đoán vector chuyển động PU con cho PU con trong PU,

suy ra danh sách ứng viên dự đoán vector chuyển động cấp độ PU con cho PU con,

lựa chọn bộ dự đoán vector chuyển động cho PU con từ danh sách ứng viên dự đoán vector chuyển động cấp độ PU con dựa trên giá trị được phân tích cú pháp của phân tử cú pháp chỉ số lựa chọn ứng viên dự đoán vector chuyển động PU con, và

giải mã dữ liệu video được mã hóa bao gồm giải mã PU con bằng cách sử dụng bộ dự đoán vector chuyển động đã lựa chọn; và

để đáp lại việc xác định cờ chuyển động PU con là không hoạt động:

phân tích phân tử cú pháp chỉ số lựa chọn ứng viên dự đoán vector chuyển động PU cho PU là phân tử cú pháp khác với phân tử cú pháp chỉ số lựa chọn ứng viên dự đoán vector chuyển động PU con,

suy ra danh sách ứng viên dự đoán vector chuyển động cấp độ PU cho PU,

lựa chọn bộ dự đoán vector chuyển động cho PU từ danh sách ứng viên dự đoán vector chuyển động cấp độ PU dựa trên giá trị được phân tích cú pháp của phân tử cú pháp chỉ số lựa chọn ứng viên dự đoán vector chuyển động PU, và

giải mã dữ liệu video được mã hóa bao gồm giải mã PU bằng cách sử dụng bộ dự đoán vector chuyển động đã lựa chọn;

trong đó danh sách ứng viên dự đoán vectơ chuyển động cấp độ PU khác với danh sách ứng viên dự đoán vectơ chuyển động cấp độ PU con, danh sách ứng viên dự đoán vectơ chuyển động cấp độ PU con bao gồm ứng viên dự đoán vectơ chuyển động afin.

2. Phương pháp theo điểm 1, trong đó danh sách ứng viên dự đoán vectơ chuyển động cấp độ PU con và danh sách ứng viên dự đoán vectơ chuyển động cấp độ PU được suy ra từ các khối lân cận của PU con.

3. Phương tiện theo điểm 2, trong đó các khối lân cận là các lân cận không gian của PU con trong hình hiện thời hoặc các lân cận thời gian của PU con trong hình mã hóa trước đó.

4. Phương pháp theo điểm 2, trong đó danh sách ứng viên dự đoán vectơ chuyển động cấp độ PU con hoặc danh sách ứng viên dự đoán vectơ chuyển động cấp độ PU được sắp xếp ít nhất một phần dựa trên các lần xuất hiện dự đoán chuyển động trong các khối lân cận.

5. Phương pháp theo điểm 1, trong đó các điểm ảnh trong PU dùng chung thông tin vectơ chuyển động thứ nhất và các điểm ảnh trong PU con dùng chung thông tin vectơ chuyển động thứ hai, và trong đó thông tin vectơ chuyển động thứ nhất hoặc thông tin vectơ thứ hai được xác định từ bộ dự đoán vectơ chuyển động đã lựa chọn.

6. Phương pháp theo điểm 1, trong đó danh sách ứng viên vectơ chuyển động cấp độ PU bao gồm ít nhất một trong số:

các ứng viên lân cận theo không gian và các ứng viên lân cận theo thời gian.

7. Phương pháp theo điểm 1, trong đó danh sách ứng viên dự đoán vectơ chuyển động cấp độ PU con bao gồm ít nhất một trong số:

ứng viên dự đoán vectơ chuyển động theo thời gian thay thế (alternative temporal motion vector prediction - ATMVP), ứng viên dự đoán vectơ chuyển động thời gian - không gian (spatial-temporal motion vector prediction - STMVP), ứng viên dự đoán vectơ chuyển động phẳng, và ứng viên suy ra vectơ chuyển động phù hợp mẫu (pattern matched motion vector derivation - PMVD).

8. Phương pháp theo điểm 1, phương pháp này còn bao gồm bước:

suy ra chỉ số ứng viên hợp nhất đáp lại việc xác định cờ chuyển động PU con hoạt động, trong đó chỉ số ứng viên hợp nhất chỉ ra bộ dự đoán vector chuyển động sẽ được lựa chọn.

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

bộ nhớ; và

bộ xử lý được tạo cấu hình để thực hiện phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 8.

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

phương tiện nhận và lưu trữ dữ liệu video được mã hóa, dữ liệu video được mã hóa bao gồm đơn vị dự đoán (prediction unit - PU);

phương tiện phân tích cú pháp, đối với PU, phân tử cú pháp cấp độ PU từ dữ liệu video được mã hóa, trong đó phân tử cú pháp cấp độ PU là cờ chuyển động đơn vị dự đoán con;

để đáp lại việc xác định cờ chuyển động đơn vị dự đoán con là hoạt động:

phương tiện phân tích phân tử cú pháp chỉ số lựa chọn ứng viên dự đoán vector chuyển động PU con cho PU con trong PU,

phương tiện suy ra danh sách ứng viên dự đoán vector chuyển động cấp độ PU con cho PU con,

phương tiện lựa chọn bộ dự đoán vector chuyển động cho PU con từ danh sách ứng viên dự đoán vector chuyển động cấp độ PU con dựa trên giá trị được phân tích cú pháp của phân tử cú pháp chỉ số lựa chọn ứng viên dự đoán vector chuyển động PU con, và

phương tiện giải mã dữ liệu video được mã hóa bao gồm giải mã PU con bằng cách sử dụng bộ dự đoán vector chuyển động đã lựa chọn; và

để đáp lại việc xác định cờ chuyển động đơn vị dự đoán con là không hoạt động:

phương tiện để phân tích phân tử cú pháp chỉ số lựa chọn ứng viên dự đoán vector chuyển động PU cho PU là phân tử cú pháp khác với phân tử cú pháp chỉ số lựa chọn ứng viên dự đoán vector chuyển động PU con,

phương tiện suy ra danh sách ứng viên dự đoán vector chuyển động cấp độ PU cho PU,

phương tiện lựa chọn bộ dự đoán vector chuyển động cho PU từ danh sách ứng viên dự đoán vector chuyển động cấp độ PU dựa trên giá trị được phân tích cú pháp của phần tử cú pháp chỉ số lựa chọn ứng viên dự đoán vector chuyển động PU, và

phương tiện giải mã dữ liệu video được mã hóa bao gồm giải mã PU bằng cách sử dụng bộ dự đoán vector chuyển động đã lựa chọn,

trong đó danh sách ứng viên dự đoán vector chuyển động cấp độ PU khác với danh sách ứng viên dự đoán vector chuyển động cấp độ PU con, danh sách ứng viên dự đoán vector chuyển động cấp độ PU con bao gồm ứng viên dự đoán vector chuyển động afin.

11. Thiết bị theo điểm 10, trong đó danh sách ứng viên dự đoán vector chuyển động cấp độ PU con và danh sách ứng viên dự đoán vector chuyển động cấp độ PU được suy ra từ các khối lân cận của PU con.

12. Thiết bị theo điểm 11, trong đó các khối lân cận là các lân cận không gian của PU con trong hình hiện thời hoặc các lân cận thời gian của PU con trong hình mã hóa trước đó.

13. Thiết bị theo điểm 11, trong đó danh sách ứng viên dự đoán vector chuyển động cấp độ PU con hoặc danh sách ứng viên dự đoán vector chuyển động cấp độ PU được sắp xếp ít nhất một phần dựa trên các lần xuất hiện dự đoán chuyển động trong các khối lân cận.

14. Thiết bị theo điểm 10, trong đó các điểm ảnh trong PU dùng chung thông tin vector chuyển động thứ nhất và các điểm ảnh trong PU con dùng chung thông tin vector chuyển động thứ hai, và trong đó thông tin vector chuyển động thứ nhất hoặc thông tin vector thứ hai được xác định từ bộ dự đoán vector chuyển động đã lựa chọn.

15. Thiết bị theo điểm 10, trong đó danh sách ứng viên vector chuyển động cấp độ PU bao gồm ít nhất một trong số:

các ứng viên lân cận theo không gian và các ứng viên lân cận theo thời gian.

16. Thiết bị theo điểm 10, trong đó danh sách ứng viên dự đoán vector chuyển động cấp độ PU con bao gồm ít nhất một trong số:

ứng viên dự đoán vector chuyển động theo thời gian thay thế (alternative temporal motion vector prediction - ATMVP), ứng viên dự đoán vector chuyển động thời gian - không gian (spatial-temporal motion vector prediction - STMVP), ứng viên dự đoán vector chuyển động phẳng, và ứng viên suy ra vector chuyển động phù hợp mẫu (pattern matched motion vector derivation - PMVD).

17. Thiết bị theo điểm 10, thiết bị này còn bao gồm:

phương tiện suy ra chỉ số ứng viên hợp nhất để đáp lại việc xác định cờ chuyển động của PU con là hoạt động, trong đó chỉ số ứng viên hợp nhất chỉ ra bộ dự đoán vector chuyển động sẽ được lựa chọn.

18. Phương tiện lưu trữ bất biến đọc được bằng máy tính lưu trữ trên đó các lệnh mà, khi được thực thi, khiến cho một hoặc nhiều bộ xử lý thực hiện phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 8.

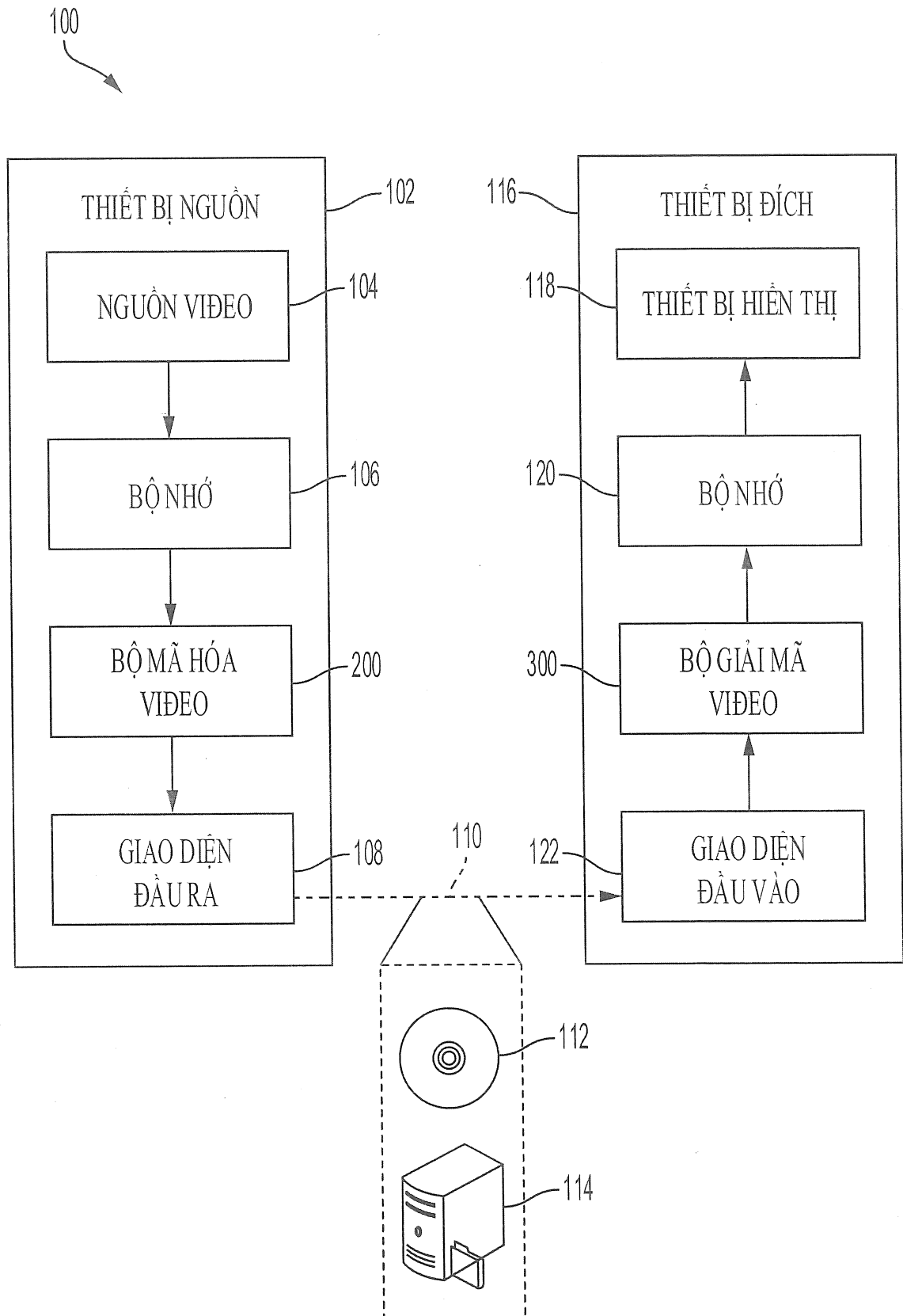


Fig.1

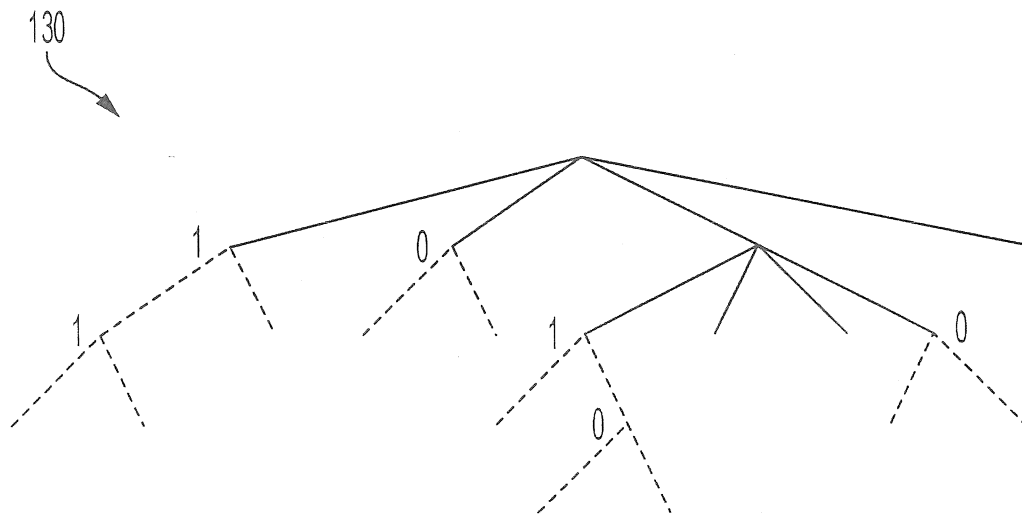


Fig.2A

132

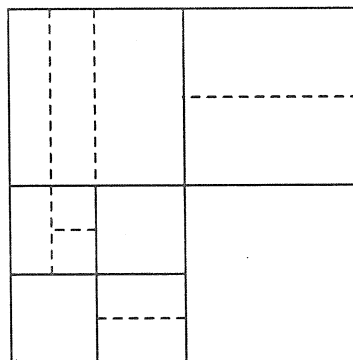


Fig.2B

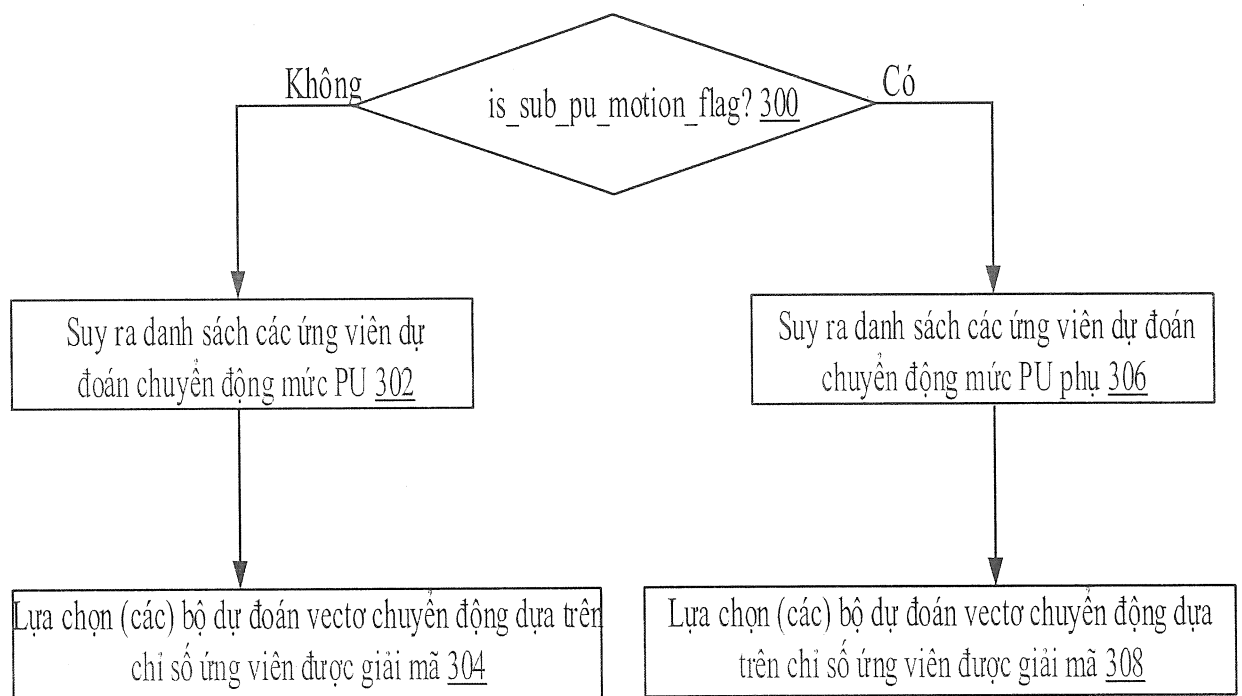


Fig.3

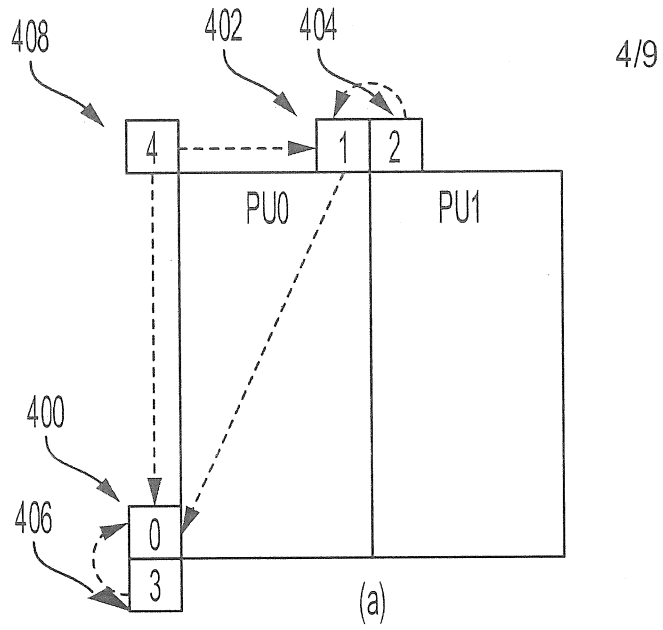
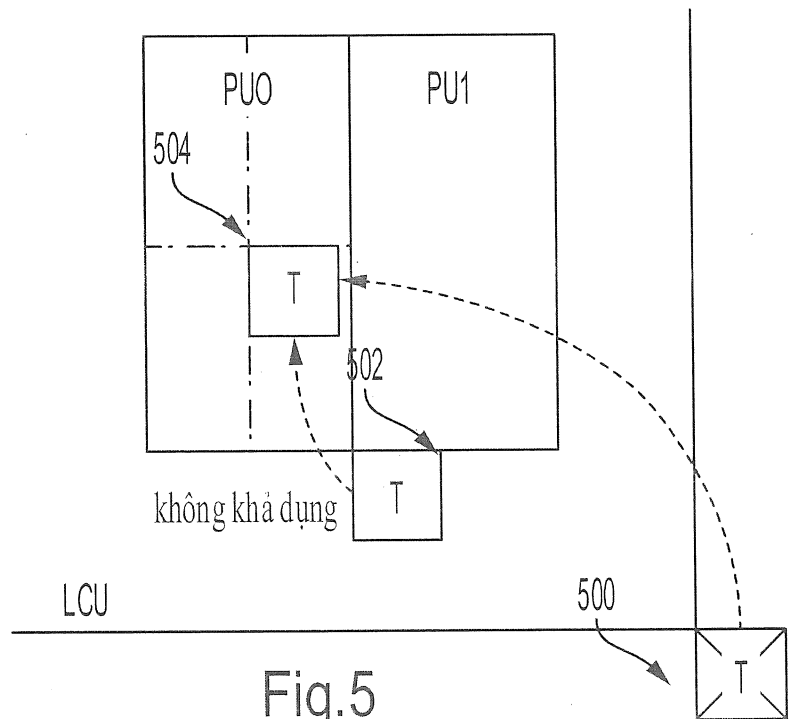


Fig.4



5/9

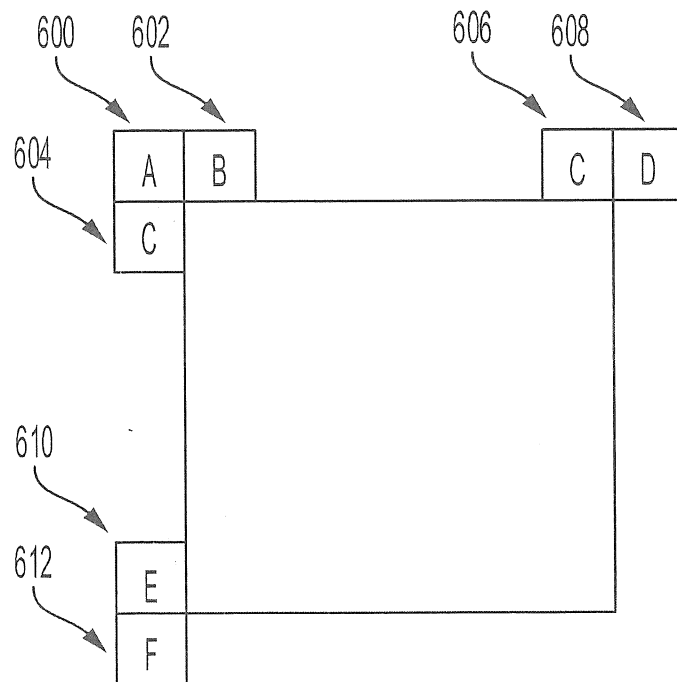


Fig.6

	c	d
b	A	B
a	C	D

Fig.7

6/9

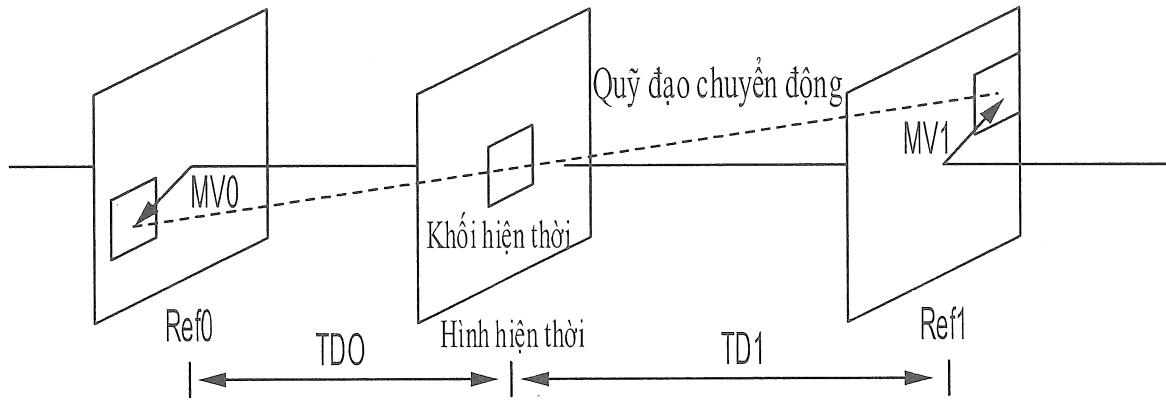


Fig.8

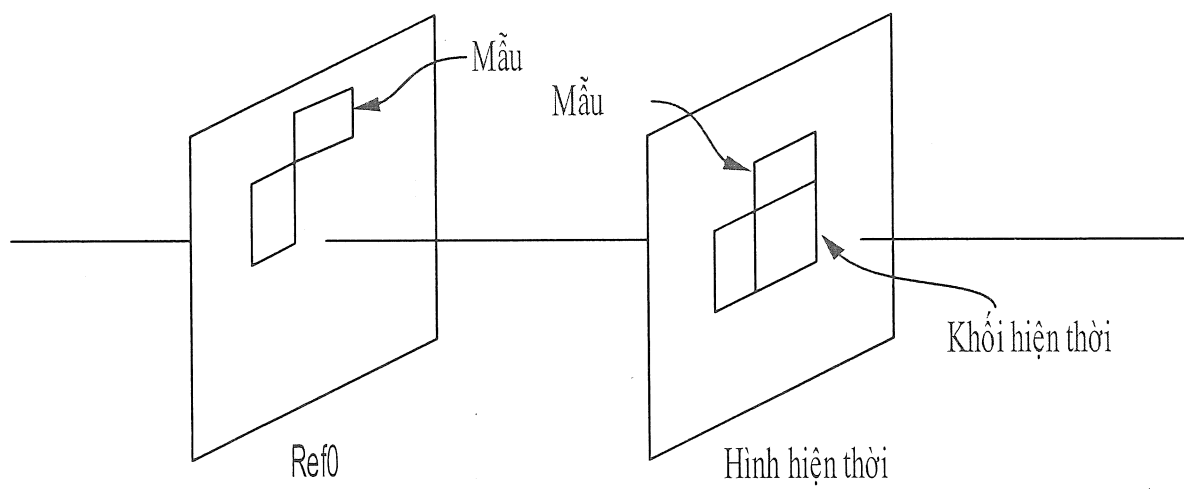


FIG. 9

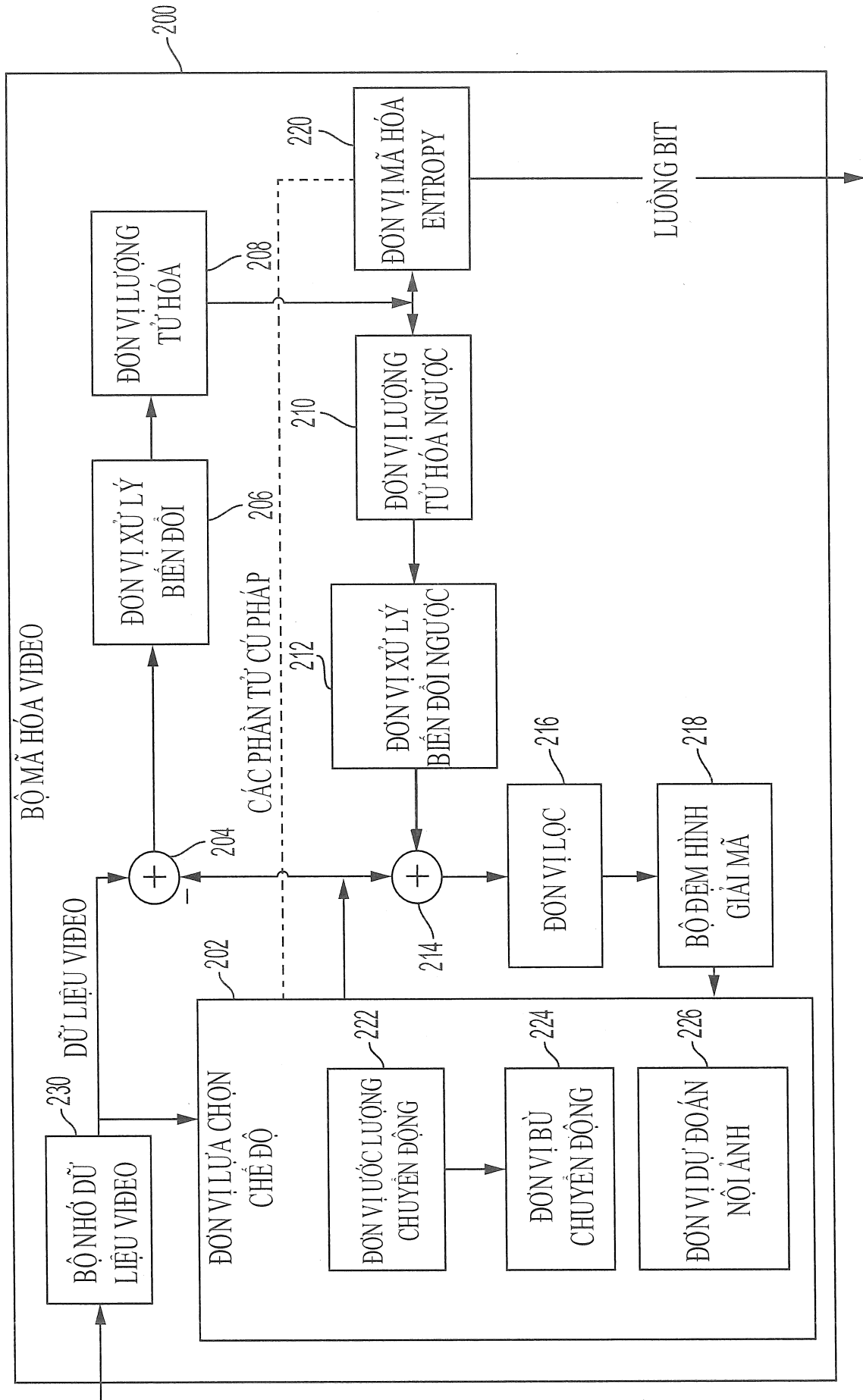


Fig.11

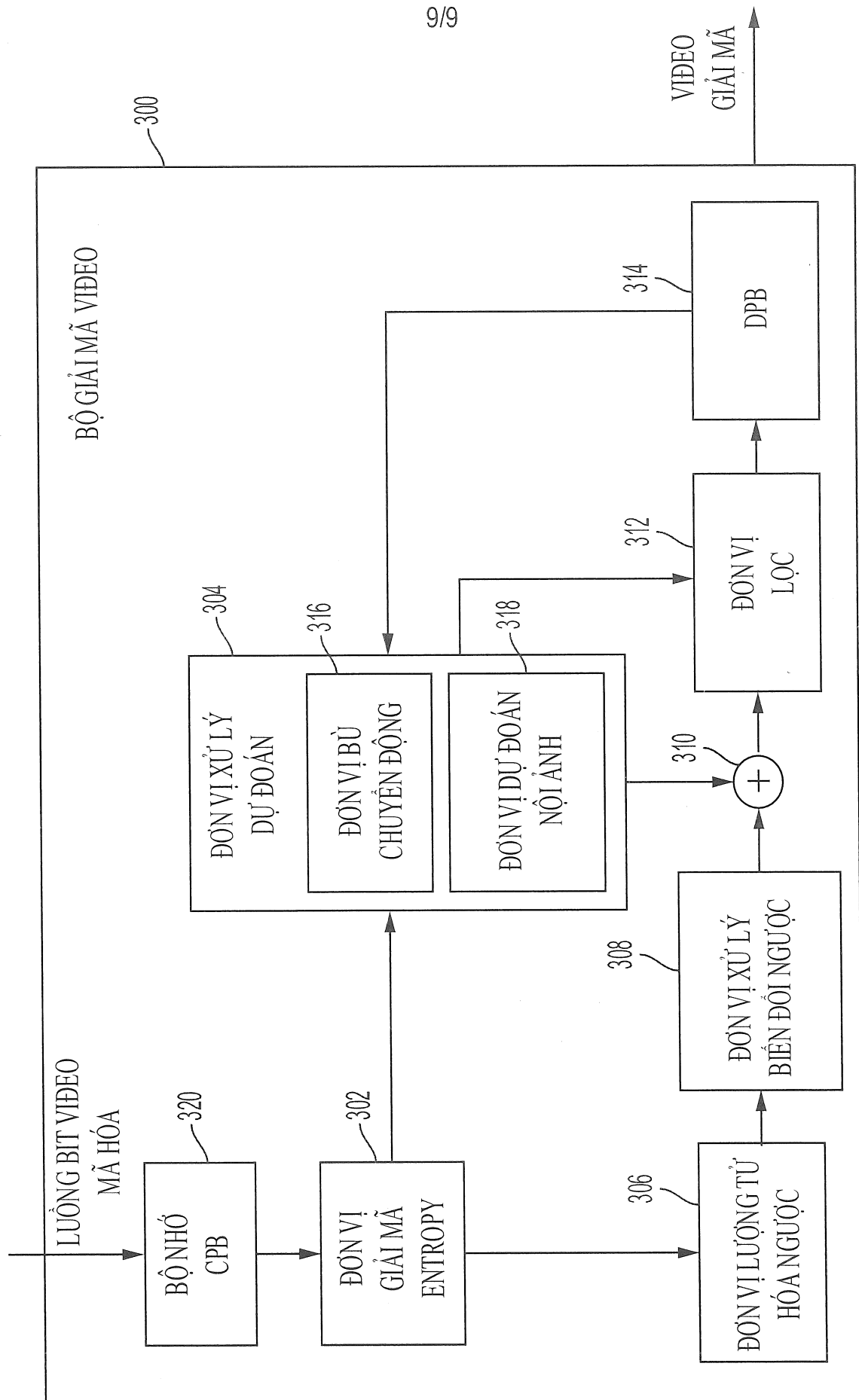


Fig.12