



(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)



1-0027503

(51)⁷ H04N 19/597; H04N 19/56; H04N
19/70; H04N 19/52; H04N 19/577

(13) B

(21) 1-2016-01254

(22) 12/09/2014

(86) PCT/US2014/055456 12/09/2014

(87) WO/2015/038937 19/03/2015

(30) 61/877,793 13/09/2013 US; 61/881,383 23/09/2013 US; 14/483,983 11/09/2014 US

(45) 25/02/2021 395

(43) 27/06/2016 339A

(73) QUALCOMM INCORPORATED (US)

ATTN: International IP Administration, 5775 Morehouse Drive, San Diego, CA
92121-1714, United States of America

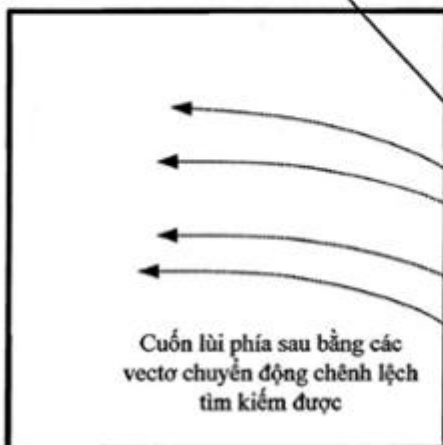
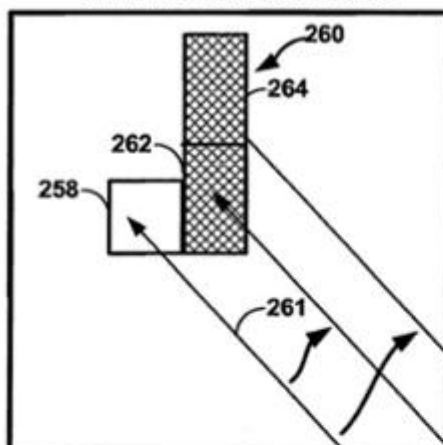
(72) CHEN, Ying (CN); ZHANG, Li (CN).

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

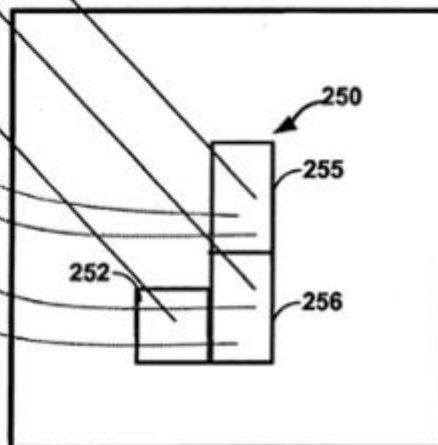
(54) PHƯƠNG PHÁP MÃ HÓA VÀ GIẢI MÃ DỮ LIỆU VIDEO VÀ THIẾT BỊ ĐƯỢC
TẠO CẤU HÌNH ĐỂ GIẢI MÃ DỮ LIỆU VIDEO

(57) Sáng chế đề xuất kỹ thuật để giải mã dữ liệu video bao gồm thu dữ liệu dự tương ứng với khối dữ liệu video, trong đó khối dữ liệu video được mã hóa bằng cách sử dụng phân chia chuyển động không đối xứng, được dự báo một chiều sử dụng dự báo tổng hợp cảnh nhìn ngược (backward view synthesis prediction - BVSP), và có kích thước 16x12, 12x16, 16x4 hoặc 4x16, phân chia khối dữ liệu video thành các khối con, mỗi khối con có kích thước 8x4 hoặc 4x8, suy ra vectơ chuyển động chênh lệch tương ứng cho mỗi khối con từ khối có chiều sâu tương ứng trong hình ảnh có chiều sâu tương ứng với hình ảnh tham chiếu, tổng hợp khối tham chiếu tương ứng cho mỗi khối con bằng cách sử dụng vectơ chuyển động chênh lệch suy ra được tương ứng và giải mã khối dữ liệu video bằng cách thực hiện bù chuyển động trên mỗi khối con sử dụng dữ liệu dự và khối tham chiếu tương ứng được tổng hợp. Sáng chế đề cập đến các phương pháp mã hóa và giải mã dữ liệu thông tin video và thiết bị được tạo cấu hình để giải mã dữ liệu video.

Hình ảnh chiều sâu tham chiếu



Hình ảnh kết cấu tham chiếu



Hình ảnh kết cấu phụ thuộc

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến việc mã hóa video, tức là, mã hóa hoặc giải mã dữ liệu video.

Tình trạng kỹ thuật của sáng chế

Tính năng video số có thể được đưa vào nhiều thiết bị khác nhau, bao gồm ti vi số, hệ thống truyền hình trực tiếp kỹ thuật số, hệ thống phát sóng không dây, thiết bị hỗ trợ số cá nhân (PDAs), máy tính xách tay hoặc máy tính để bàn, máy tính bảng, thiết bị đọc sách điện tử, máy hình ảnh số, thiết bị ghi âm số, thiết bị nghe nhạc kỹ thuật số, thiết bị chơi trò chơi video, bàn giao tiếp trò chơi video, điện thoại di động hoặc điện thoại phát sóng qua vệ tinh, được gọi là “điện thoại thông minh,” thiết bị hội nghị truyền hình, thiết bị truyền video, và thiết bị tương tự. Thiết bị video số thực thi kỹ thuật lập mã video, như các kỹ thuật được mô tả theo các chuẩn được định nghĩa bởi MPEG-2, MPEG-4, ITU-T H.263, ITU-T H.264/MPEG-4, Part 10, chuẩn mã hóa video được nâng cấp (Advanced Video Coding-AVC), chuẩn mã hóa video hiệu suất cao (High Efficiency Video Coding -HEVC) đang được phát triển hiện nay và các dạng mở rộng của các chuẩn này. Thiết bị video có thể truyền, nhận, mã hóa, giải mã và/hoặc lưu giữ thông tin video số hiệu quả hơn bằng cách thực thi kỹ thuật lập mã video này.

Kỹ thuật lập mã video bao gồm việc dự báo không gian (trong hình ảnh) và/hoặc thời gian (liên hình ảnh) để làm giảm hoặc loại bỏ phần thừa vốn có trong chuỗi video. Đối với việc lập mã video trên cơ sở khối, lát video (ví dụ, khung video hoặc một phần của khung video) có thể được phân chia thành các khối video, còn có thể được gọi là khối cây, đơn vị mã hóa (coding unit-CU) và/hoặc nút mã hóa. Các khối video trong lát được mã hóa trong hình ảnh (I) của một hình ảnh được mã hóa bằng cách sử dụng sự dự báo không gian liên quan đến các mẫu tham chiếu trong khối lân cận trong cùng một hình ảnh. Khối video trong lát video được mã hóa liên hình ảnh (P hoặc B) của một hình ảnh có thể sử dụng phép dự báo không gian liên quan đến mẫu tham chiếu trong khối lân cận trong cùng một hình ảnh hoặc dự báo thời gian

liên quan đến mẫu tham chiếu trong hình ảnh tham chiếu khác. Các hình ảnh có thể được gọi là các khung và hình ảnh tham chiếu có thể được gọi là khung tham chiếu.

Dự báo không gian hoặc thời gian dẫn đến khối dự báo cho khối cần mã hóa. Dữ liệu dư thể hiện độ lệch điểm ảnh giữa khối gốc cần mã hóa và khối dự báo. Khối được lập mã liên hình ảnh được mã hóa theo vector chuyển động hướng đến một khối của mẫu tham chiếu tạo ra khối dự báo, và dữ liệu dư chỉ báo độ lệch giữa khối được mã hóa và khối dự báo. Khối được lập mã trong hình ảnh được mã hóa theo chế độ lập mã trong hình ảnh và dữ liệu dư. Để nén thêm, dữ liệu dư có thể được biến đổi từ miền điểm ảnh thành miền biến đổi, thu được hệ số biến đổi dư, sau đó hệ số này có thể được lượng tử hóa. Hệ số biến đổi lượng tử hóa, ban đầu được sắp xếp theo mảng hai chiều, có thể được quét để suy ra vector một chiều của hệ số biến đổi và mã hóa entropi có thể được áp dụng để nén nhiều hơn nữa.

Bản chất kỹ thuật của sáng chế

Nói chung, sáng chế đề cập đến việc mã hóa video ba chiều (3D) dựa vào bộ mã hóa-giải mã cải tiến, bao gồm, theo một số ví dụ, kỹ thuật mã hóa chiều sâu. Bản mô tả này mô tả kỹ thuật mã hóa dự báo tổng hợp khung hình, bao gồm việc xác định kích thước khối, khi sử dụng cùng với việc phân chia chuyển động không đối xứng. Bản mô tả này cũng mô tả kỹ thuật dự báo chuyển động nâng cao khi được sử dụng cùng với việc phân chia chuyển động không đối xứng.

Theo một ví dụ của bản mô tả, phương pháp giải mã dữ liệu video bao gồm thu dữ liệu dư tương ứng với khối dữ liệu video, trong đó khối dữ liệu video này được giải mã bằng cách sử dụng phân chia chuyển động không đối xứng, được dự báo một chiều sử dụng dự báo tổng hợp cảnh nhìn ngược (backward view synthesis prediction - BVSP), và có kích thước 16x12, 12x16, 16x4 hoặc 4x16, phân chia khối dữ liệu video thành các khối con, mỗi khối con có kích thước 8x4 hoặc 4x8, suy ra vector chuyển động chênh lệch tương ứng cho mỗi khối con từ khối có chiều sâu tương ứng trong hình ảnh có chiều sâu tương ứng với hình ảnh tham chiếu, tổng hợp khối tham chiếu tương ứng cho mỗi khối con bằng cách sử dụng vector chuyển động chênh lệch suy ra được tương ứng và giải mã khối dữ liệu video bằng cách thực hiện bù chuyển động trên mỗi khối con sử dụng dữ liệu dư và khối tham chiếu tương ứng được tổng hợp.

Theo ví dụ khác của sáng chế, phương pháp mã hóa dữ liệu video bao gồm việc tạo ra khối dữ liệu video bằng cách sử dụng sự phân chia chuyển động không đối xứng, trong đó khối dữ liệu video được dự báo một chiều sử dụng dự báo tổng hợp cảnh nhìn ngược (BVSP), và có kích thước 16x12, 12x16, 16x4 hoặc 4x16, phân chia khối dữ liệu video thành các khối con, mỗi khối con có kích thước 8x4 hoặc 4x8, suy ra vector chuyển động chênh lệch tương ứng cho mỗi khối con từ khối có chiều sâu tương ứng trong hình ảnh có chiều sâu tương ứng với hình ảnh tham chiếu, tổng hợp khối tham chiếu tương ứng cho mỗi khối con bằng cách sử dụng vector chuyển động chênh lệch suy ra được tương ứng và mã hóa khối dữ liệu video bằng cách thực hiện bù chuyển động trên mỗi khối con sử dụng khối tham chiếu tương ứng tổng hợp được.

Theo ví dụ khác của sáng chế, thiết bị được tạo cấu hình để giải mã dữ liệu video bao gồm bộ nhớ video được tạo cấu hình để lưu giữ thông tin tương ứng với khối dữ liệu video và một hoặc nhiều bộ xử lý được tạo cấu hình để nhận dữ liệu dư tương ứng với khối dữ liệu video, trong đó khối dữ liệu video này được mã hóa bằng cách sử dụng sự phân chia chuyển động không đối xứng, được dự báo một chiều sử dụng dự báo tổng hợp cảnh nhìn ngược (BVSP), và có kích thước 16x12, 12x16, 16x4 hoặc 4x16, phân chia khối dữ liệu video thành các khối con, mỗi khối con có kích thước 8x4 hoặc 4x8, suy ra vector chuyển động chênh lệch tương ứng cho mỗi khối con từ khối có chiều sâu tương ứng trong hình ảnh có chiều sâu tương ứng với hình ảnh tham chiếu, tổng hợp khối tham chiếu tương ứng cho mỗi khối con bằng cách sử dụng vector chuyển động chênh lệch suy ra được tương ứng và mã hóa khối dữ liệu video bằng cách thực hiện bù chuyển động trên mỗi khối con sử dụng khối tham chiếu tương ứng tổng hợp được.

Theo ví dụ khác của sáng chế, thiết bị được tạo cấu hình để giải mã dữ liệu video bao gồm phương tiện để nhận dữ liệu dư tương ứng với khối dữ liệu video, trong đó khối dữ liệu video được mã hóa bằng cách sử dụng phân chia chuyển động không đối xứng, được dự báo một chiều sử dụng dự báo tổng hợp cảnh nhìn ngược (BVSP), và có kích thước 16x12, 12x16, 16x4 hoặc 4x16, phân chia khối dữ liệu video thành các khối con, mỗi khối con có kích thước 8x4 hoặc 4x8, phương tiện suy ra vector chuyển động chênh lệch tương ứng cho mỗi khối con từ khối có chiều sâu tương ứng trong hình ảnh có chiều sâu tương ứng với hình ảnh tham chiếu, phương tiện để tổng hợp khối tham chiếu tương ứng cho mỗi khối con bằng cách sử dụng

vector chuyển động chênh lệch suy ra được tương ứng và phương tiện để mã hóa khối dữ liệu video bằng cách thực hiện bù chuyển động trên mỗi khối con sử dụng khối tham chiếu tương ứng tổng hợp được.

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

Mô tả vắn tắt các hình vẽ

Fig.1 là sơ đồ khối minh họa hệ thống mã hóa và giải mã video ví dụ có thể sử dụng kỹ thuật dự báo liên hình ảnh theo sáng chế.

Fig.2 là sơ đồ khái niệm minh họa trình tự giải mã ví dụ cho video đa khung hình.

Fig.3 là sơ đồ khái niệm minh họa cấu trúc dự báo ví dụ cho video đa khung hình.

Fig.4 là sơ đồ khái niệm minh họa các giá trị cấu trúc và chiều sâu đối với video 3D.

Fig.5 là sơ đồ khái niệm minh họa các kiểu phân chia ví dụ.

Fig.6 là sơ đồ khái niệm minh họa các ứng cử vector chuyển động theo chế độ hợp nhất.

Fig.7 là bảng thể hiện chi tiết ví dụ về các chỉ số ứng cử hợp nhất.

Fig.8 là sơ đồ khái niệm minh họa các khối lân cận được sử dụng cho quy trình suy ra vector chênh lệch ví dụ.

Fig.9 là sơ đồ khái niệm minh họa quy trình suy ra vector chênh lệch của khối lân cận.

Fig.10 là sơ đồ khái niệm minh họa bốn điểm ảnh góc của khối có chiều sâu 8x8 .

Fig.11 là sơ đồ khái niệm minh họa việc tạo ra ứng cử vector chuyển động dự báo giữa các cảnh nhìn ví dụ đối với chế độ kết sắp/bỏ qua.

Fig.12 là bảng thể hiện chi tiết ví dụ về các chỉ số tham chiếu trong 3D-HEVC.

Fig.13 là sơ đồ khái niệm minh họa việc tạo ra ứng cử vector chuyển động kế thừa ví dụ để mã hóa chiều sâu.

Fig.14 minh họa cấu trúc dự báo cấu trúc dự báo của dự báo số dư tiên tiến (ARP) trong mã hóa video đa khung hình.

Fig.15 sơ đồ khái niệm minh họa mối liên hệ ví dụ giữa khối hiện thời, khối tham chiếu và khối bù chuyển động.

Fig.16 sơ đồ khái niệm minh họa dự báo chuyển động giữa các cảnh nhìn của đơn vị dự báo con.

Fig.17 là hình vẽ khái niệm thể hiện việc dự báo tổng hợp cảnh nhìn ngược và kỹ thuật bù chuyển động theo bản mô tả này khi sử dụng phân chia chuyển động không đối xứng.

Fig.18 sơ đồ khái niệm minh họa sự kế thừa vector chuyển động và kỹ thuật bù chuyển động cho các kích thước phân chia chuyển động không đối xứng 4x16 và 16x4.

Fig.19 là sơ đồ khối minh họa bộ mã hóa video ví dụ có thể thực hiện kỹ thuật dự báo liên hình ảnh theo sáng chế.

Fig.20 là sơ đồ khối minh họa bộ giải mã video ví dụ có thể thực hiện kỹ thuật dự báo liên hình ảnh theo sáng chế.

Fig.21 là lưu đồ minh họa phương pháp mã hóa ví dụ theo sáng chế.

Fig.22 là lưu đồ minh họa phương pháp mã hóa ví dụ khác theo sáng chế.

Fig.23 là lưu đồ minh họa phương pháp mã hóa ví dụ khác theo sáng chế.

Fig.24 là lưu đồ minh họa phương pháp giải mã ví dụ theo sáng chế.

Fig.25 là lưu đồ minh họa phương pháp giải mã ví dụ theo sáng chế.

Fig.26 là lưu đồ minh họa phương pháp giải mã ví dụ theo sáng chế.

Mô tả chi tiết sáng chế

Nói chung, sáng chế mô tả kỹ thuật liên quan đến việc mã hóa video 3D dựa trên bộ lập-giải mã tiên tiến, bao gồm mã hóa một hoặc nhiều khung hình theo khối có chiều sâu bằng cách sử dụng bộ lập-giải mã 3D-HEVC ((High Efficiency Video Coding-HEVC). Cụ thể, sáng chế mô tả kỹ thuật để chia tiếp các đơn vị dự báo

(prediction units-PU) được phân chia sử dụng kỹ thuật phân chia chuyển động không đối xứng thành các khối con nhỏ hơn. Kỹ thuật theo sáng chế bao gồm kỹ thuật tạo ra và/hoặc kế thừa các vectơ chuyển động và các vectơ chuyển động chênh lệch cho các khối con của PU được phân chia sử dụng phân chia chuyển động không đối xứng.

Fig.1 là sơ đồ khối minh họa hệ thống mã hóa và giải mã video ví dụ 10 mà sử dụng kỹ thuật theo sáng chế. Như được thể hiện trên Fig.1, hệ thống 10 bao gồm thiết bị nguồn 12 cung cấp dữ liệu video mã hóa được mã hóa ở thời điểm sau đó bởi thiết bị đích 14. Cụ thể, thiết bị nguồn 12 có thể cung cấp dữ liệu video đến thiết bị đích 14 qua vật ghi đọc được bằng máy tính 16. Thiết bị nguồn 12 và thiết bị đích 14 có thể bao gồm thiết bị bất kỳ trong nhiều loại thiết bị, bao gồm máy tính để bàn, máy tính notebook (tức là, máy tính xách tay), máy tính bảng, bộ giải mã, các máy điện thoại cầm tay như điện thoại "thông minh", điện thoại bảng "thông minh", ti vi, camera, thiết bị hiển thị, thiết bị phát phương tiện kỹ thuật số, bàn giao tiếp trò chơi video, thiết bị truyền video hoặc thiết bị tương tự. Theo một số trường hợp, thiết bị nguồn 12 và thiết bị đích 14 có thể được trang bị để truyền thông không dây.

Thiết bị đích 14 có thể nhận dữ liệu video mã hóa để giải mã qua vật ghi đọc được bằng máy tính 16. Vật ghi đọc được bằng máy tính 16 có thể bao gồm loại vật ghi hoặc thiết bị bất kỳ có khả năng di chuyển dữ liệu video đã mã hóa từ thiết bị nguồn 12 đến thiết bị đích 14. Theo một ví dụ, vật ghi đọc được bằng máy tính 16 có thể bao gồm vật ghi truyền thông làm cho thiết bị nguồn 12 có thể truyền dữ liệu video mã hóa trực tiếp đến thiết bị đích 14 theo thời gian thực. Dữ liệu video mã hóa có thể được điều biến theo chuẩn truyền thông, như giao thức truyền thông không dây và được truyền đến thiết bị đích 14. Phương tiện truyền thông có thể bao gồm phương tiện truyền thông không dây hoặc có dây bất kỳ, như phổ tần số vô tuyến (radio frequency-RF) hoặc một hoặc nhiều đường truyền vật lý. Phương tiện truyền thông có thể tạo thành một phần của mạng dựa trên 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, trạm gốc, hoặc thiết bị bất kỳ khác có thể là hữu dụng để tạo điều kiện thuận lợi cho việc truyền thông từ thiết bị nguồn 12 đến thiết bị đích 14.

Theo một số ví dụ, dữ liệu mã hóa có thể được xuất từ giao diện xuất 22 đến thiết bị lưu trữ. Tương tự, dữ liệu mã hóa có thể được truy cập từ thiết bị lưu trữ bởi giao diện nhập. Thiết bị lưu trữ có thể bao gồm thiết bị bất kỳ trong số phương tiện

lưu trữ dữ liệu truy cập phân bố hoặc cục bộ như đĩa cứng, đĩa Blu-ray, DVD, CD-ROM, Bộ nhớ đệm nhanh, bộ nhớ khả biến hoặc không khả biến hoặc phương tiện lưu trữ số thích hợp bất kỳ khác để lưu giữ dữ liệu video mã hóa. Theo ví dụ khác, thiết bị lưu trữ có thể tương đương với máy chủ cung cấp tệp tin hoặc thiết bị lưu giữ trung gian khác có thể lưu giữ video mã hóa được tạo ra bởi thiết bị nguồn 12. Thiết bị đích 14 có thể truy cập dữ liệu video từ thiết bị lưu giữ bằng cách truyền trực tiếp hoặc tải về. Máy chủ cung cấp tệp tin có thể là loại máy chủ bất kỳ có khả năng lưu giữ dữ liệu video mã hóa và truyền dữ liệu video mã hóa đó đến thiết bị đích 14. Máy chủ cung cấp tệp tin ví dụ bao gồm máy chủ web (ví dụ, đối với website), máy chủ FTP, thiết bị lưu giữ gắn vào mạng (Network attached storage -NAS), hoặc ổ đĩa cục bộ. Thiết bị đích 14 có thể truy cập dữ liệu video mã hóa thông qua kết nối dữ liệu chuẩn bất kỳ, kể cả kết nối Internet. Kết nối này có thể bao gồm kênh không dây (ví dụ, , kết nối Wi-Fi), kết nối có dây (ví dụ, DSL, modem cáp, v.v.), hoặc tổ hợp của cả hai kết nối trên mà thích hợp để truy cập dữ liệu video mã hóa được lưu giữ trên máy chủ cung cấp tệp tin. Việc truyền dữ liệu video mã hóa từ thiết bị lưu trữ có thể là truyền truyền trực tiếp, truyền tải về hoặc tổ hợp của chúng.

Kỹ thuật theo sáng chế không nhất thiết giới hạn ở các ứng dụng hoặc thiết lập không dây. Kỹ thuật này có thể được áp dụng để mã hóa video hỗ trợ ứng dụng bất kỳ trong nhiều ứng dụng đa phương tiện, như phát sóng truyền hình qua không khí, truyền hình cáp, truyền hình vệ tinh, truyền trực video qua Internet, như truyền trực tiếp thích ứng động trên HTTP (DASH), video số được mã hóa trên vật ghi lưu trữ dữ liệu, giải mã video số được lưu giữ trên vật ghi lưu trữ dữ liệu hoặc các ứng dụng khác. Theo một số ví dụ, hệ thống 10 có thể được tạo cấu hình để hỗ trợ truyền phát video một chiều hoặc hai chiều để hỗ trợ các ứng dụng như truyền trực tiếp video, phát lại video, truyền phát video, và/hoặc điện thoại video.

Theo ví dụ của Fig.1, thiết bị nguồn 12 bao gồm nguồn video 18, đơn vị ước lượng chiều sâu 19, bộ mã hóa video 20, và giao diện xuất 22. Thiết bị đích 14 bao gồm giao diện nhập 28, bộ giải mã video 30, bộ kết xuất dựa trên hình ảnh có chiều sâu (Depth image based rendering-DIBR) 31, và thiết bị hiển thị 32. 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 12 có thể nhận dữ liệu video từ nguồn video bên ngoài 18, như

camera bên ngoài. Tương tự, thiết bị đích 14 có thể kết nối với thiết bị hiển thị bên ngoài, thay vì bao gồm thiết bị hiển thị tích hợp.

Hệ thống được minh họa 10 trên Fig.1 chỉ là một ví dụ. Kỹ thuật theo sáng chế có thể được thực hiện bởi thiết bị mã hóa và/hoặc giải mã video kỹ thuật số bất kỳ. Mặc dù nói chung kỹ thuật theo sáng chế được thực hiện bởi thiết bị mã hóa video, kỹ thuật này cũng có thể được thực hiện bởi bộ mã hóa/giải mã video, thường được gọi là “bộ lập-giải mã.” Hơn nữa, kỹ thuật theo sáng chế cũng có thể được thực hiện bởi bộ tiền xử lý video. Thiết bị nguồn 12 và thiết bị đích 14 là các ví dụ về thiết bị mã hóa như vậy trong đó thiết bị nguồn 12 tạo ra dữ liệu video mã hóa để truyền đến thiết bị đích 14. Theo một số ví dụ, thiết bị 12, 14 có thể hoạt động theo cách thức gần như đối xứng để mỗi thiết bị 12, 14 bao gồm các thành phần mã hóa và giải mã video. Do vậy, hệ thống 10 có thể hỗ trợ truyền video một chiều hoặc hai chiều giữa thiết bị video 12, 14, ví dụ, truyền trực tiếp video, phát lại video, truyền phát video, và/hoặc điện thoại video.

Nguồn video 18 của thiết bị nguồn 12 có thể bao gồm thiết bị thu video, như camera video, kho lưu trữ video chứa video đã thu trước đó, và/hoặc giao diện nạp video để nhận video từ bộ cấp dung lượng video. Theo một ví dụ khác, nguồn video 18 có thể tạo ra dữ liệu dựa trên đồ họa máy tính làm video nguồn hoặc tổ hợp video trực tiếp, video được lưu trữ và video được tạo ra bởi máy tính. Theo một số ví dụ, nếu nguồn video 18 là camera video, thiết bị nguồn 12 và thiết bị đích 14 có thể tạo thành điện thoại có camera hoặc điện thoại video. Như nêu trên, tuy nhiên, kỹ thuật được mô tả trong bản mô tả này có thể được áp dụng để mã hóa video nói chung và có thể được áp dụng cho các ứng dụng không dây và/hoặc có dây. Trong mỗi trường hợp, video được thu, thu trước hoặc được tạo ra bởi máy tính có thể được mã hóa bằng bộ mã hóa video 20. Sau đó thông tin video mã hóa được xuất bởi giao diện xuất 22 lên vật ghi đọc được bằng máy tính 16.

Nguồn video 18 có thể cung cấp một hoặc nhiều khung hình của dữ liệu video đến bộ mã hóa video 20. Ví dụ, nguồn video 18 có thể tương đương với một loạt máy quay, mỗi máy có vị trí theo chiều ngang duy nhất cân xứng với cảnh cụ thể được quay phim. Theo cách khác, nguồn video 18 có thể tạo ra dữ liệu video từ các hình phối cảnh của máy quay theo chiều ngang khác nhau, ví dụ, sử dụng đồ họa máy tính. Đơn vị ước lượng chiều sâu 19 có thể được tạo cấu hình để xác định giá trị cho điểm

ảnh có chiều sâu tương ứng với điểm ảnh trong kết cấu hình ảnh. Ví dụ, đơn vị ước lượng chiều sâu 19 có thể là đơn vị dẫn đường và đo cự ly bằng sóng âm thanh (Sound Navigation và Ranging-SONAR), đơn vị đo cự ly và phát hiện bằng ánh sáng (Light Detection và Ranging-LIDAR), hoặc đơn vị khác có khả năng xác định trực tiếp giá trị chiều sâu gần như đồng thời với việc ghi dữ liệu video của kịch bản.

Ngoài ra hoặc theo cách khác, đơn vị ước lượng chiều sâu 19 có thể được tạo cấu hình để tính toán giá trị chiều sâu một cách gián tiếp bằng cách so sánh hai hoặc nhiều hình ảnh được thu gần như đồng thời từ các hình phối cảnh của máy quay theo chiều ngang khác nhau. Bằng cách tính toán chênh lệch theo chiều ngang giữa các giá trị điểm ảnh gần như giống nhau trong các hình ảnh, đơn vị ước lượng chiều sâu 19 có thể ước lượng chiều sâu của các đối tượng khác nhau trong kịch bản. Đơn vị ước lượng chiều sâu 19 có thể được tích hợp về mặt chức năng với nguồn video 18, theo một số ví dụ. Ví dụ, khi nguồn video 18 tạo ra các hình ảnh đồ họa máy tính, đơn vị ước lượng chiều sâu 19 có thể cung cấp bản đồ chiều sâu thực cho các đối tượng đồ họa, ví dụ, sử dụng tọa độ z của điểm ảnh và đối tượng được sử dụng để kết xuất hình ảnh kết cấu.

Vật ghi đọc được bằng máy tính 16 có thể bao gồm vật ghi tạm thời, như phát rộng không dây hoặc truyền mạng có dây hoặc vật ghi lưu trữ (tức là, vật ghi lưu trữ bất biến) như đĩa cứng, ổ đĩa nhớ nhanh, đĩa compac, đĩa video số, đĩa Blu-ray hoặc vật ghi đọc được bằng máy tính khác. Theo một số ví dụ, máy chủ mạng (không được thể hiện trên hình vẽ) có thể nhận dữ liệu video mã hóa từ thiết bị nguồn 12 và cung cấp dữ liệu video mã hóa đến thiết bị đích 14, ví dụ, thông qua sự truyền mạng. Tương tự, thiết bị tính toán của phương tiện sản xuất vật ghi, như phương tiện dập đĩa, có thể nhận dữ liệu video mã hóa từ thiết bị nguồn 12 và sản xuất đĩa chứa dữ liệu video mã hóa này. Do vậy, vật ghi đọc được bằng máy tính 16 có thể được hiểu là bao gồm một hoặc nhiều vật ghi đọc được bằng máy tính thuộc các dạng khác nhau, theo các ví dụ khác nhau.

Giao diện nhập 28 của thiết bị đích 14 nhận thông tin từ vật ghi đọc được bằng máy tính 16. Thông tin của vật ghi đọc được bằng máy tính 16 có thể bao gồm thông tin cú pháp được xác định bởi bộ mã hóa video 20, mà cũng được bộ giải mã video 30 sử dụng, bao gồm phần tử cú pháp mô tả các đặc điểm và/hoặc việc xử lý các khối và các đơn vị mã hóa khác, ví dụ, GOP. Thiết bị hiển thị 32 hiển thị dữ liệu video được

giải mã cho người dùng và có thể bao gồm thiết bị bất kỳ trong số các thiết bị hiển thị khác nhau như màn hình ống phóng điện tử (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 các loại thiết bị hiển thị khác. Theo một số ví dụ, thiết bị hiển thị 32 có thể bao gồm thiết bị có khả năng hiển thị hai hoặc nhiều khung hình đồng thời hoặc gần như là đồng thời, ví dụ, để tạo hiệu ứng 3D cho người xem.

Đơn vị DIBR 31 của thiết bị đích 14 có thể kết xuất khung hình tổng hợp, sử dụng thông tin kết cấu và chiều sâu của khung hình được giải mã nhận được từ bộ giải mã video 30. Ví dụ, đơn vị DIBR 31 có thể xác định độ chênh lệch theo chiều ngang cho dữ liệu điểm ảnh của hình ảnh kết cấu dưới dạng một hàm của giá trị điểm ảnh trên sơ đồ chiều sâu tương ứng. Đơn vị DIBR 31 sau đó có thể tạo ra hình ảnh tổng hợp bằng cách dịch chuyển các điểm ảnh trong hình ảnh kết cấu sang bên trái hoặc phải một khoảng bằng độ chênh lệch theo chiều ngang đã xác định. Theo cách này, thiết bị hiển thị 32 có thể hiển thị một hoặc nhiều khung hình, tương ứng với các khung hình giải mã và/hoặc khung hình tổng hợp, theo tổ hợp bất kỳ. Theo kỹ thuật của sáng chế, bộ giải mã video 30 có thể cung cấp giá trị độ nét gốc và được cập nhật cho các khoảng chiều sâu và các thông số máy quay cho đơn vị DIBR 31, mà có thể sử dụng khoảng chiều sâu và các thông số máy quay này để tổng hợp khung hình một cách thích hợp.

Mặc dù không được thể hiện trên Fig.1, theo một số khía cạnh, mỗi bộ mã hóa video 20 và bộ giải mã video 30 có thể được tích hợp với bộ mã hóa và giải mã âm thanh và có thể bao gồm các đơn vị MUX-DEMUX thích hợp hoặc phần cứng và phần mềm khác, để xử lý mã hóa cả âm thanh và video trong một dòng dữ liệu chung hoặc trong các dòng dữ liệu riêng biệt. Nếu sử dụng, các đơn vị MUX-DEMUX có thể thực hiện theo giao thức ghép kênh ITU H.223 hoặc giao thức khác như giao thức gói dữ liệu người dùng (User datagram protocol-UDP).

Mỗi bộ mã hóa video 20 và bộ giải mã video 30 có thể được thực thi dưới dạng sơ đồ mạch bất kỳ trong nhiều sơ đồ mạch mã hóa thích hợp, như một hoặc nhiều bộ vi xử lý, bộ xử lý tín hiệu số (digital signal processor-DSP), mạch tích hợp chuyên dụng (Application specific integrated circuits-ASIC), mảng cổng lập trình được dạng trường (Field programmable gate arrays- FPGAs), logic rời rạc, phần mềm, phần

cứng, phần sụn hoặc tổ hợp bất kỳ của chúng. Khi kỹ thuật này được thực thi một phần trong phần mềm, thiết bị có thể lưu giữ các lệnh cho phần mềm trên vật ghi bất biến đọc được bằng máy tính thích hợp và thực thi các lệnh trong phần cứng sử dụng một hoặc nhiều bộ xử lý để thực hiện kỹ thuật theo sáng chế. Mỗi bộ mã hóa video 20 và bộ giải mã video 30 có thể được bao gồm trong một hoặc nhiều bộ mã hóa hoặc giải mã, một trong hai bộ này có thể được tích hợp dưới dạng một phần của bộ mã hóa/giải mã kết hợp (bộ lập-giải mã) trong thiết bị tương ứng. Thiết bị bao gồm bộ mã hóa video 20 và/hoặc bộ giải mã video 30 có thể bao gồm mạch tích hợp, bộ vi xử lý và/hoặc thiết bị truyền thông không dây, như điện thoại di động.

Bộ mã hóa video 20 và bộ giải mã video 30 có thể hoạt động theo chuẩn mã hóa video, như chuẩn mã hóa video hiệu suất cao (HEVC) hiện nay đang được phát triển và có thể theo mẫu thử nghiệm HEVC (HEVC Test Model -HM). Theo cách khác, bộ mã hóa video 20 và bộ giải mã video 30 có thể hoạt động theo chuẩn độc quyền hoặc công nghiệp khác, như chuẩn ITU-T H.264, theo cách khác được gọi là MPEG-4, Part 10, mã hóa video nâng cao (AVC), hoặc các dạng mở rộng của các chuẩn này, như dạng mở rộng MVC của ITU-T H.264/AVC. Bản dự thảo chung mới nhất của MVC được mô tả trong “Advanced video coding for generic audiovisual services,” ITU-T Recommendation H.264, Mar 2010. Cụ thể, bộ mã hóa video 20 và bộ giải mã video 30 có thể hoạt động theo chuẩn mã hóa 3D và/hoặc đa khung hình, bao gồm dạng mở rộng 3D của chuẩn HEVC (ví dụ, 3D-HEVC).

Một bản dự thảo của chuẩn HEVC, được gọi là “HEVC Working Draft 10” hoặc “WD10,” được mô tả trong tài liệu JCTVC-L1003v34, Bross et al., “High efficiency video coding (HEVC) text specification draft 10 (for FDIS & Last Call),” nhóm hợp tác chung về mã hóa video (Joint Collaborative Team on video Coding-JCT-VC) của ITU-T SG16 WP3 và ISO/IEC JTC1/SC29/WG11, cuộc họp lần thứ 12, tại Geneva, Thụy sĩ, 14-23 tháng 1, 2013, mà kể từ 22/08/2014, có thể tải từ: http://phenix.int-evry.fr/jct/doc_end_user/documents/12_Geneva/wg11/JCTVC-L1003-v34.zip.

Bản dự thảo khác của chuẩn HEVC, được gọi là “WD10 revisions” được mô tả trong Bross et al., “Editors’ đề xuất các hiệu chỉnh đối với HEVC phiên bản 1,” Joint Collaborative Team on video Coding (JCT-VC) của ITU-T SG16 WP3 và ISO/IEC

JTC1/SC29/WG11, cuộc họp lần thứ 13, Incheon, Hàn quốc, tháng 4 năm 2013, từ 22/08/ 2014, có sẵn từ:

http://phenix.int-evry.fr/jct/doc_end_user/documents/13_Incheon/wg11/JCTVC-M0432-v3.zip. Dạng mở rộng đa khung hình đối với HEVC, tức là MV-HEVC, cũng đang được phát triển bởi JCT-3V.

Hiện nay, Nhóm hợp tác chung về mã hóa video 3D (JCT-3C) theo chuẩn VCEG và MPEG đang phát triển chuẩn 3DV dựa trên HEVC, một phần nỗ lực chuẩn hóa bao gồm chuẩn hóa bộ lập-giải mã đa khung hình dựa trên HEVC (MV-HEVC) và một phần khác cho mã hóa video 3D dựa trên HEVC (3D-HEVC). Đối với MV-HEVC, cần đảm bảo rằng chỉ có những thay đổi cú pháp bậc cao (High-level syntax-HLS) trong đó, sao cho mô đun ở cấp đơn vị mã hóa/đơn vị dự báo không cần phải thiết kế lại và có thể tái sử dụng hoàn toàn cho MV-HEVC. Đối với 3D-HEVC, các công cụ mã hóa mới, bao gồm các công cụ ở cấp đơn vị mã hóa/đơn vị dự báo, đối với cả khung hình kết cấu lẫn chiều sâu có thể được bao gồm và hỗ trợ.

Một phiên bản phần mềm 3D-HTM cho 3D-HEVC có thể được tải về từ đường dẫn sau: [3D-HTM phiên bản 8.0]: https://hevc.hhi.fraunhofer.de/svn/svn_3DVCSsoftware/tags/HTM-8.0/. Bản dự thảo hoạt động của 3D-HEVC (số tài liệu: E1001) có sẵn từ: http://phenix.int-sudparis.eu/jct2/doc_end_user/current_document.php?id=1361. Bản mô tả phần mềm mới nhất (số tài liệu: E1005) có sẵn từ: http://phenix.int-sudparis.eu/jct2/doc_end_user/current_document.php?id=1360.

Phiên bản gần đây của phần mềm 3D-HTM cho 3D-HEVC có thể tải về từ đường dẫn sau: [3D-HTM phiên bản 12.0]: https://hevc.hhi.fraunhofer.de/svn/svn_3DVCSsoftware/tags/HTM-12.0/. Bản dự thảo hoạt động tương ứng của 3D-HEVC (số tài liệu: I1001) có sẵn tại: http://phenix.int-evry.fr/jct3v/doc_end_user/current_document.php?id=2299. Bản mô tả phần mềm mới nhất (số tài liệu: I1005) có sẵn từ: http://phenix.int-evry.fr/jct3v/doc_end_user/current_document.php?id=2301.

Đầu tiên, kỹ thuật mã hóa ví dụ của HEVC được thảo luận. Nỗ lực chuẩn hóa HEVC là dựa trên mô hình cải tiến của thiết bị mã hóa video được gọi là Mô hình thử

nghiệm HEVC (HM). HM giả định trước một số khả năng khác của thiết bị mã hóa video so với thiết bị đã có theo, ví dụ, chuẩn ITU-T H.264/AVC. Ví dụ, trong khi H.264 cung cấp chín chế độ mã hóa dự báo trong hình ảnh, thì HM có thể cung cấp đến ba mươi ba chế độ dự báo trong hình ảnh góc, cùng với chế độ DC và Planar.

Theo HEVC và các mô tả về mã hóa video khác, chuỗi video thường bao gồm một loạt hình ảnh. Các hình ảnh còn được gọi là “khung.” Hình ảnh có thể bao gồm ba mảng mẫu, được ký hiệu là S_L , S_{Cb} , và S_{Cr} . S_L là mảng hai chiều (tức là, khối) của khối tín hiệu chói. S_{Cb} là mảng hai chiều của các mẫu sắc độ Cb. S_{Cr} là mảng hai chiều của các mẫu sắc độ Cr. Mẫu sắc độ còn có thể được gọi là mẫu “độ màu” Trong một số trường hợp, hình ảnh có thể là đơn sắc và chỉ bao gồm mảng khối tín hiệu chói.

Để tạo ra biểu diễn mã hóa của hình ảnh, bộ mã hóa video 20 có thể tạo ra tập đơn vị cây mã hóa (Coding tree unit-CTU). Mỗi CTU có thể bao gồm khối cây mã hóa của khối tín hiệu chói, hai khối cây mã hóa tương ứng của mẫu sắc độ và cấu trúc cú pháp được sử dụng để mã hóa các mẫu của khối cây mã hóa. Trong các hình ảnh đơn sắc hoặc các hình ảnh có ba mặt màu riêng biệt, CTU có thể bao gồm khối cây mã hóa đơn và các cấu trúc cú pháp được sử dụng để mã hóa các mẫu của khối cây mã hóa. Khối cây mã hóa có thể là khối mẫu $N \times N$. CTU cũng có thể được gọi là “khối cây” hoặc “đơn vị mã hóa lớn nhất” (largest coding unit-CU). CTU của HEVC một cách tổng quát có thể giống với các khối lớn theo các chuẩn khác như H.264/AVC. Tuy nhiên, CTU không nhất thiết giới hạn ở kích thước cụ thể và có thể bao gồm một hoặc nhiều đơn vị mã hóa (CU). Lát có thể bao gồm một số nguyên các CTU được sắp xếp liên tiếp theo thứ tự quét mảnh.

Để tạo ra CTU mã hóa, bộ mã hóa video 20 có thể thực hiện một cách đệ quy sự phân chia cây tứ phân trên khối cây mã hóa của CTU để chia khối cây mã hóa thành các khối mã hóa, do đó có tên là “đơn vị cây mã hóa.” Khối mã hóa có thể là khối mẫu $N \times N$. Đơn vị mã hóa (CU) có thể bao gồm khối mã hóa của khối tín hiệu chói và hai khối mã hóa tương ứng của mẫu sắc độ của hình ảnh có mảng khối tín hiệu chói, mảng mẫu Cb và mảng mẫu Cr và cấu trúc cú pháp được sử dụng để mã hóa các mẫu của khối mã hóa. Trong các hình ảnh đơn sắc hoặc các hình ảnh có ba mặt màu riêng biệt, CU có thể bao gồm khối cây mã hóa đơn và các cấu trúc cú pháp được sử dụng để mã hóa các mẫu của khối cây mã hóa.

Bộ mã hóa video 20 có thể phân chia khối mã hóa của CU thành một hoặc nhiều khối dự báo. Khối dự báo là khối mẫu hình chữ nhật (tức là, hình vuông hoặc không phải hình vuông) trên đó cùng một dự báo được áp dụng. Đơn vị dự báo (PU) của CU có thể bao gồm khối dự báo của khối tín hiệu chói, hai khối dự báo tương ứng của mẫu sắc độ và cấu trúc cú pháp được sử dụng để dự báo khối dự báo. Trong các hình ảnh đơn sắc hoặc các hình ảnh có ba mặt màu riêng biệt, PU có thể bao gồm khối cây dự báo đơn và các cấu trúc cú pháp được sử dụng để dự báo khối dự báo. Bộ mã hóa video 20 có thể tạo ra khối tín hiệu chói Cb, và Cr dự báo cho khối dự báo độ sáng, Cb, và Cr của mỗi PU của CU.

Bộ mã hóa video 20 có thể sử dụng dự báo trong hình ảnh hoặc dự báo liên hình ảnh để tạo ra khối dự báo cho PU. Nếu bộ mã hóa video 20 sử dụng dự báo trong hình ảnh để tạo ra khối dự báo của PU, bộ mã hóa video 20 có thể tạo ra khối dự báo của PU dựa trên các mẫu được giải mã của hình ảnh liên quan đến PU. Theo một số phiên bản của HEVC, đối với thành phần độ sáng của mỗi PU, phương pháp dự báo trong hình ảnh được sử dụng có 33 chế độ dự báo góc (được đánh số từ 2 đến 34), chế độ DC (được đánh số 1) và chế độ Planar (được đánh số là 0).

Nếu bộ mã hóa video 20 sử dụng dự báo liên hình ảnh để tạo ra khối dự báo của PU, bộ mã hóa video 20 có thể tạo ra khối dự báo của PU dựa trên các mẫu được giải mã của một hoặc nhiều hình ảnh không phải hình ảnh liên quan đến PU. Dự báo liên hình ảnh có thể là dự báo liên hình ảnh một chiều (tức là, dự báo đơn) hoặc dự báo liên hình ảnh hai chiều (tức là, dự báo đôi). Để thực hiện dự báo đơn hoặc dự báo đôi, bộ mã hóa video 20 có thể tạo ra danh sách hình ảnh tham chiếu thứ nhất (RefPicList0) và danh sách hình ảnh tham chiếu thứ hai (RefPicList1) cho lát hiện thời. Mỗi danh sách hình ảnh tham chiếu này có thể bao gồm một hoặc nhiều hình ảnh tham chiếu. Khi sử dụng dự báo một chiều, bộ mã hóa video 20 có thể tìm kiếm các hình ảnh tham chiếu trong một hoặc cả hai danh sách RefPicList0 và RefPicList1 để xác định vị trí tham chiếu trong hình ảnh tham chiếu. Hơn nữa, khi sử dụng dự báo một chiều, bộ mã hóa video 20 có thể tạo ra, dựa ít nhất một phần trên các mẫu tương ứng với vị trí tham chiếu, khối mẫu dự báo cho PU. Hơn nữa, khi sử dụng dự báo một chiều, bộ mã hóa video 20 có thể tạo ra một vector chuyển động chỉ báo sự chuyển vị không gian giữa khối dự báo của PU và vị trí tham chiếu. Để chỉ báo sự chuyển vị không gian giữa khối dự báo của PU và vị trí tham chiếu, vector chuyển động có thể

bao gồm thành phần ngang biểu diễn chuyển vị ngang giữa khối dự báo của PU và vị trí tham chiếu và có thể bao gồm thành phần dọc định rõ sự chuyển vị dọc giữa khối dự báo của PU và vị trí tham chiếu.

Khi sử dụng dự báo hai chiều để mã hóa PU, bộ mã hóa video 20 có thể xác định vị trí tham chiếu thứ nhất ở hình ảnh tham chiếu trong RefPicList0 và vị trí tham chiếu thứ hai ở hình ảnh tham chiếu trong RefPicList1. Bộ mã hóa video 20 có thể tạo ra, dựa ít nhất một phần trên mẫu tương ứng với vị trí tham chiếu thứ nhất và thứ hai, khối mẫu dự báo cho PU. Hơn nữa, khi sử dụng dự báo hai chiều để mã hóa PU, bộ mã hóa video 20 có thể suy ra vector chuyển động thứ nhất chỉ báo sự chuyển vị không gian giữa khối mẫu của PU và vị trí tham chiếu thứ nhất và vector chuyển động thứ hai chỉ báo sự chuyển vị không gian giữa khối dự báo của PU và vị trí tham chiếu thứ hai.

Thông thường việc tạo ra danh sách hình ảnh tham chiếu cho danh sách hình ảnh tham chiếu thứ nhất hoặc thứ hai (ví dụ, RefPicList0 hoặc RefPicList1) của hình ảnh B bao gồm hai bước: khởi tạo danh sách hình ảnh tham chiếu và sắp xếp lại danh sách hình ảnh tham chiếu (thay đổi). Khởi tạo danh sách hình ảnh tham chiếu là một cơ chế hoàn hảo đặt các hình ảnh tham chiếu trong bộ nhớ hình ảnh tham chiếu (còn được gọi là bộ đệm hình ảnh được mã hóa) vào danh sách dựa trên thứ tự của các giá trị đếm thứ tự hình ảnh (Picture Order Count-POC, được sắp thẳng hàng với thứ tự hiển thị của hình ảnh). Cơ chế sắp xếp lại danh sách hình ảnh tham chiếu có thể thay đổi vị trí của hình ảnh được đặt trong danh sách trong quá trình khởi tạo danh sách hình ảnh tham chiếu đến vị trí mới bất kỳ hoặc đặt hình ảnh tham chiếu bất kỳ trong bộ nhớ hình ảnh tham chiếu ở vị trí bất kỳ ngay cả khi hình ảnh này không nằm trong danh sách đã khởi tạo. Một số hình ảnh sau khi sắp xếp lại danh sách hình ảnh tham chiếu (sửa đổi), có thể được đặt ở vị trí khác xa trong danh sách. Tuy nhiên, nếu một vị trí của hình ảnh vượt quá số hình ảnh tham chiếu hoạt động của danh sách, hình ảnh này không được xem là mục nhập của danh sách hình ảnh tham chiếu cuối. Số lượng hình ảnh tham chiếu hoạt động có thể được ký hiệu trong tiêu đề của lát đối với mỗi danh sách.

Sau khi danh sách hình ảnh tham chiếu được tạo ra (tức là RefPicList0 và RefPicList1, nếu có), chỉ số tham chiếu đối với danh sách hình ảnh tham chiếu có thể được sử dụng để nhận dạng hình ảnh tham chiếu bất kỳ có trong danh sách hình ảnh tham chiếu.

Sau khi bộ mã hóa video 20 tạo ra khối tín hiệu chói, Cb, và Cr dự báo cho một hoặc nhiều PU của CU, bộ mã hóa video 20 có thể tạo ra khối tín hiệu chói dư cho CU. Mỗi mẫu trong khối tín hiệu chói dư của CU chỉ ra chênh lệch giữa khối tín hiệu chói trong một trong số các khối tín hiệu chói dự báo của CU và mẫu tương ứng trong khối tín hiệu chói mã hóa gốc của CU. Ngoài ra, bộ mã hóa video 20 có thể tạo ra khối Cb dư cho CU. Mỗi mẫu trong khối Cb dư của CU có thể chỉ ra chênh lệch giữa mẫu Cb trong một trong số các khối Cb dự báo của CU và mẫu tương ứng trong khối Cb mã hóa gốc của CU. Bộ mã hóa video 20 có thể tạo ra khối Cr dư cho CU. Mỗi mẫu trong khối Cr dư của CU chỉ ra chênh lệch giữa mẫu Cr trong một trong số các khối Cr dự báo của CU và mẫu tương ứng trong khối Cr mã hóa gốc của CU.

Hơn nữa, bộ mã hóa video 20 có thể sử dụng sự phân chia cây tứ phân để phân tách khối tín hiệu chói Cb, và Cr dư của CU thành một hoặc nhiều khối biến đổi độ chói, Cb, và Cr. Khối biến đổi là khối mẫu hình chữ nhật (ví dụ, hình vuông hoặc không phải hình vuông) trên đó cùng một biến đổi được áp dụng. Đơn vị biến đổi (TU) của CU có thể bao gồm khối biến đổi của khối tín hiệu chói, hai khối biến đổi tương ứng của mẫu sắc độ và cấu trúc cú pháp được sử dụng để biến đổi mẫu khối biến đổi. Do vậy, mỗi TU của CU có thể được kết hợp với khối biến đổi độ chói, khối biến đổi Cb, và khối biến đổi Cr. Khối biến đổi độ chói kết hợp với TU có thể là khối con của khối tín hiệu chói dư của CU. khối biến đổi Cb có thể là khối con của khối Cb dư của CU. khối biến đổi Cr có thể là khối con của khối Cr dư của CU. Trong các hình ảnh đơn sắc hoặc các hình ảnh có ba mặt màu riêng biệt, TU có thể bao gồm một khối biến đổi và các cấu trúc cú pháp được sử dụng để biến đổi các mẫu của khối biến đổi.

Bộ mã hóa video 20 có thể áp dụng một hoặc nhiều biến đổi cho khối biến đổi độ chói của TU để tạo ra khối hệ số biến đổi cho TU. Khối hệ số có thể là mảng hệ số biến đổi hai chiều. Hệ số biến đổi có thể là lượng vô hướng. Bộ mã hóa video 20 có thể áp dụng một hoặc nhiều biến đổi cho khối biến đổi Cb của TU để tạo ra khối hệ số Cb cho TU. Bộ mã hóa video 20 có thể áp dụng một hoặc nhiều biến đổi cho khối biến đổi Cr của TU để tạo ra khối hệ số Cr cho TU.

Sau khi tạo ra khối hệ số (ví dụ, khối hệ số độ chói, khối hệ số Cb hoặc khối hệ số Cr), bộ mã hóa video 20 có thể lượng tử hóa khối hệ số. Lượng tử hóa thường chỉ quá trình trong đó các hệ số biến đổi được lượng tử hóa để có thể giảm lượng số liệu được sử dụng để biểu diễn các hệ số biến đổi, tạo ra sự nén thêm. Sau khi bộ mã hóa

video 20 lượng tử hóa khối hệ số, bộ mã hóa video 20 có thể mã hóa entropi phần tử cú pháp chỉ báo các hệ số biến đổi đã được lượng tử hóa. Ví dụ, bộ mã hóa video 20 có thể thực hiện Lập giải mã Số học Nhị phân Thích ứng theo Ngữ cảnh (Context-Adaptive Binary Arithmetic Coding - CABAC) trên các phần tử cú pháp chỉ báo các hệ số biến đổi đã lượng tử hóa.

Bộ mã hóa video 20 có thể xuất dòng bit mà bao gồm chuỗi các bit tạo ra sự biểu diễn các hình ảnh mã hóa và dữ liệu liên quan. Dòng bit này có thể bao gồm chuỗi các đơn vị lớp trừu tượng hóa mạng (Network abstraction layer-NAL). Đơn vị NAL là cấu trúc cú pháp chứa chỉ báo về kiểu dữ liệu trong đơn vị NAL và bit chứa dữ liệu ở dạng tải trọng chuỗi bit thô (Raw byte sequence payload-RBSP) được đặt rải rác khi cần với các bit ngăn chặn sự mô phỏng. Mỗi đơn vị NAL bao gồm tiêu đề đơn vị NAL và đóng gói RBSP. Tiêu đề đơn vị NAL có thể bao gồm phần tử cú pháp chỉ báo mã loại đơn vị NAL. Mã loại đơn vị NAL được chỉ rõ bởi tiêu đề đơn vị NAL của đơn vị NAL chỉ báo loại đơn vị NAL. RBSP có thể là cấu trúc cú pháp chứa số nguyên bit được đóng gói trong đơn vị NAL. Trong một số trường hợp, RBSP bao gồm các bit 0.

Các loại đơn vị NAL khác nhau có thể đóng gói các loại RBSP khác nhau. Ví dụ, loại đơn vị NAL thứ nhất có thể đóng gói RBSP cho tập tham số hình ảnh (picture parameter set-PPS), loại đơn vị NAL thứ hai có thể đóng gói RBSP cho lát mã hóa, loại đơn vị NAL thứ ba có thể đóng gói RBSP cho SEI, và v.v.. Đơn vị NAL đóng gói RBSP cho dữ liệu mã hóa video (như trái ngược với RBSP cho tập tham số và thông điệp SEI) có thể được gọi là đơn vị NAL lớp mã hóa video (Video coding layer-VCL).

Bộ giải mã video 30 có thể nhận dòng bit được tạo ra bởi bộ mã hóa video 20. Ngoài ra, bộ giải mã video 30 có thể phân tích dòng bit để tạo ra các phần tử cú pháp từ dòng bit này. Bộ giải mã video 30 có thể tái tạo các hình ảnh của dữ liệu video dựa ít nhất một phần vào các phần tử cú pháp thu được từ dòng bit. Quá trình tái tạo dữ liệu video có thể thường ngược với quá trình được thực hiện bởi bộ mã hóa video 20. Chẳng hạn, bộ giải mã video 30 có thể sử dụng các vector chuyển động của PU để xác định các khối dự báo cho PU của CU hiện thời. Ngoài ra, bộ giải mã video 30 có thể lượng tử hóa ngược khối hệ số kết hợp với TU của CU hiện thời. Bộ giải mã video 30 có thể thực hiện biến đổi ngược khối hệ số để tái tạo khối biến đổi kết hợp với TU của CU hiện thời. Bộ giải mã video 30 có thể tái tạo khối mã hóa của CU hiện thời bằng cách bổ sung các mẫu của khối dự báo cho các PU của CU hiện thời vào các mẫu

tương ứng của khối biến đổi của các TU của CU hiện thời. Bằng cách tái tạo các khối mã hóa cho mỗi CU của một hình ảnh, bộ giải mã video 30 có thể tái tạo hình ảnh này.

Theo một số ví dụ, bộ mã hóa video 20 có thể báo hiệu thông tin chuyển động của PU bằng cách sử dụng chế độ kết sắp hoặc chế độ dự báo vector chuyển động nâng cao (advanced motion vector prediction - AMVP). Nói cách khác, trong HEVC, có hai chế độ để dự báo các tham số chuyển động, một là chế độ kết sắp và chế độ còn lại là AMVP. Dự báo chuyển động có thể bao gồm xác định thông tin chuyển động của đơn vị video (ví dụ, PU) dựa trên thông tin chuyển động của một hoặc nhiều đơn vị video khác. Thông tin chuyển động của PU có thể bao gồm (các) vector chuyển động của PU và (các) chỉ số tham chiếu của PU.

Khi bộ mã hóa video 20 báo hiệu thông tin chuyển động của PU hiện thời bằng cách sử dụng chế độ kết sắp, bộ mã hóa video 20 tạo ra danh sách ứng cử kết sắp. Nói cách khác, bộ mã hóa video 20 có thể thực hiện quá trình tạo danh sách biến dự báo của vector chuyển động. Danh sách ứng cử kết sắp bao gồm tập ứng cử kết sắp mà chỉ báo thông tin chuyển động của PU lân cận về không gian hoặc thời gian với PU hiện thời. Tức là, theo chế độ kết sắp, danh sách ứng cử của các tham số chuyển động (ví dụ, chỉ số tham chiếu, vector chuyển động, v.v.) được tạo khi một ứng cử có thể là từ khối lân cận về mặt không gian và thời gian. Theo một số ví dụ, các ứng cử này cũng có thể bao gồm một ứng cử nhân tạo.

Hơn nữa, theo chế độ kết sắp, bộ mã hóa video 20 có thể chọn một ứng cử kết sắp từ danh sách ứng cử kết sắp và có thể sử dụng thông tin chuyển động được chỉ báo bởi ứng cử kết sắp đã chọn này làm thông tin chuyển động của PU hiện thời. Bộ mã hóa video 20 có thể báo hiệu vị trí trong danh sách ứng cử kết sắp của ứng cử kết sắp đã chọn. Chẳng hạn, bộ mã hóa video 20 có thể báo hiệu các tham số của vector chuyển động đã chọn bằng cách truyền chỉ số vào danh sách ứng cử. Bộ giải mã video 30 có thể nhận, từ dòng bit, chỉ số này vào danh sách ứng cử (tức là, chỉ số danh sách ứng cử). Ngoài ra, bộ giải mã video 30 có thể tạo ra cùng danh sách ứng cử kết sắp và có thể xác định, dựa vào chỉ báo về vị trí của ứng cử kết sắp được chọn, ứng cử kết sắp được chọn. Sau đó, bộ giải mã video 30 có thể sử dụng thông tin chuyển động của ứng cử kết sắp được chọn để tạo ra khối dự báo cho PU hiện thời. Tức là, bộ giải mã video 30 có thể xác định, ít nhất một phần dựa trên chỉ số danh sách ứng cử, ứng cử đã chọn trong danh sách ứng cử, trong đó ứng cử đã chọn chỉ rõ vector chuyển động cho PU hiện thời. Theo cách này, ở phía bộ giải mã, ngay khi chỉ số này được giải mã, tất

cả các tham số chuyển động của khối tương ứng trong đó điểm chỉ số có thể được PU kế thừa.

Chế độ bỏ qua là tương tự với chế độ kết sắp. Ở chế độ bỏ qua, bộ mã hóa video 20 và bộ giải mã video 30 tạo ra và sử dụng danh sách ứng cử kết sắp theo cách giống như bộ mã hóa video 20 và bộ giải mã video 30 sử dụng danh sách ứng cử kết sắp theo chế độ kết sắp. Tuy nhiên, khi bộ mã hóa video 20 báo hiệu thông tin chuyển động của PU hiện thời sử dụng chế độ bỏ qua, bộ mã hóa video 20 không báo hiệu dữ liệu dư bất kỳ cho PU hiện thời. Do đó, bộ giải mã video 30 có thể xác định, không cần sử dụng dữ liệu dư, khối dự báo cho PU dựa trên khối tham chiếu được chỉ báo bởi thông tin chuyển động của ứng cử được chọn trong danh sách ứng cử kết sắp.

Chế độ AMVP tương tự với chế độ kết sắp trong đó bộ mã hóa video 20 có thể tạo ra danh sách ứng cử và có thể chọn một ứng cử từ danh sách ứng cử. Tuy nhiên, khi bộ mã hóa video 20 báo hiệu thông tin chuyển động RefPicListX của PU hiện thời sử dụng chế độ AMVP, bộ mã hóa video 20 có thể báo hiệu chênh lệch vector chuyển động RefPicListX (chuyển động vector difference-MVD) cho PU hiện thời và chỉ số tham chiếu RefPicListX cho PU hiện thời ngoài việc báo hiệu cờ RefPicListX MVP cho PU hiện thời. Cờ RefPicListX MVP cho PU hiện thời có thể chỉ báo vị trí của ứng cử AMVP được chọn trong danh sách ứng cử AMVP. RefPicListX MVD cho PU hiện thời có thể chỉ báo chênh lệch giữa vector chuyển động RefPicListX của PU hiện thời và vector chuyển động của ứng cử AMVP được chọn. Theo cách này, bộ mã hóa video 20 có thể báo hiệu thông tin chuyển động RefPicListX của PU hiện thời bằng cách báo hiệu cờ dự báo vector chuyển động (MVP) RefPicListX, giá trị chỉ số tham chiếu RefPicListX và RefPicListX MVD. Nói cách khác, dữ liệu trong dòng bit biểu diễn vector chuyển động cho PU hiện thời có thể bao gồm dữ liệu biểu diễn chỉ số tham chiếu, chỉ số đối với danh sách ứng cử và MVD.

Hơn nữa, khi thông tin chuyển động của PU hiện thời được báo hiệu bằng cách sử dụng chế độ AMVP, bộ giải mã video 30 có thể nhận được, từ dòng bit, MVD cho PU hiện thời và cờ MVP. Bộ giải mã video 30 có thể tạo ra cùng một danh sách ứng cử AMVP và có thể xác định, dựa trên cờ MVP, ứng cử AMVP được chọn. Bộ giải mã video 30 có thể khôi phục vector chuyển động của PU hiện thời bằng cách bổ sung MVD vào vector chuyển động được chỉ báo bằng ứng cử AMVP đã chọn. Tức là, bộ giải mã video 30 có thể xác định, dựa trên vector chuyển động được chỉ báo bởi ứng cử

AMVP được chọn và MVD, vector chuyển động của PU hiện thời. Sau đó, bộ giải mã video 30 có thể sử dụng tvector chuyển động hoặc vector chuyển động được khôi phục của PU hiện thời để tạo ra khối dự báo cho PU hiện thời.

Khi bộ giải mã video 30 tạo ra danh sách ứng cử AMVP cho PU hiện thời, bộ giải mã video 30 có thể suy ra một hoặc nhiều ứng cử AMVP dựa trên thông tin chuyển động của PU (tức là, các PU lân cận về mặt không gian) bao gồm các vị trí lân cận về mặt không gian với PU hiện thời. PU có thể bao gồm vị trí khi khối dự báo của PU bao gồm vị trí này

Một ứng cử trong danh sách ứng cử kết sắp hoặc danh sách ứng cử AMVP mà dựa trên thông tin chuyển động của PU lân cận về thời gian với PU hiện thời (tức là, PU ở khoảng cách về thời gian khác với PU hiện thời) có thể được gọi là TMVP. Tức là, TMVP có thể được sử dụng để cải thiện hiệu quả mã hóa của HEVC và khác với các công cụ mã hóa khác, TMVP có thể cần truy cập vector chuyển động của khung trong bộ đệm hình ảnh được giải mã, cụ thể hơn trong danh sách hình ảnh tham chiếu.

Việc sử dụng TMVP được kích hoạt hoặc bất hoạt trên cơ sở từng CVS (trình tự video mã hóa), cơ sở từng phần, hoặc cơ sở khác. Phần tử cú pháp (ví dụ, `sps_temporal_mvp_enable_flag`) trong SPS có thể chỉ báo liệu việc sử dụng TMVP có được kích hoạt cho CVS hay không. Hơn thế nữa, khi việc sử dụng TMVP được kích hoạt cho CVS, việc sử dụng TMVP có thể được kích hoạt hoặc bất hoạt đối với các lát cụ thể trong CVS. Chẳng hạn, phần tử cú pháp (ví dụ, `slice_temporal_mvp_enable_flag`) trong tiêu đề lát có thể chỉ báo liệu việc sử dụng TMVP có được kích hoạt đối với lát hay không. Do vậy, trong lát được dự báo liên hình ảnh, khi TMVP được kích hoạt cho toàn bộ CVS (ví dụ, `sps_temporal_mvp_enable_flag` trong SPS được thiết lập là 1), `slice_temporal_mvp_enable_flag` được báo hiệu trong tiêu đề của lát để chỉ báo liệu TMVP có được kích hoạt cho lát hiện thời hay không.

Để xác định TMVP, đầu tiên bộ mã hóa video có thể nhận dạng hình ảnh tham chiếu bao gồm PU được đặt ở cùng vị trí với PU hiện thời. Nói cách khác, bộ mã hóa video có thể nhận dạng hình ảnh đồng vị trí. Nếu lát hiện thời của hình ảnh hiện thời là lát B (tức là, lát được phép bao gồm các PU dự báo liên hình ảnh hai chiều), bộ mã hóa video 20 có thể báo hiệu, trong tiêu đề của lát, phần tử cú pháp (ví dụ, `collocated_from_l0_flag`) chỉ báo liệu hình ảnh đồng vị trí là từ `RefPicList0` hay `RefPicList1`. Nói cách khác, khi việc sử dụng TMVP được kích hoạt cho phần hiện

thời và phần hiện thời là phần B (ví dụ, phần được phép bao gồm các PU dự báo liên hình ảnh hai chiều), bộ mã hóa video 20 có thể báo hiệu phần tử cú pháp (ví dụ, `collocated_from_10_flag`) trong tiêu đề của lát để chỉ báo liệu hình ảnh đồng vị trí là ở trong `RefPicList0` hay `RefPicList1`. Nói cách khác, để có TMVP, đầu tiên hình ảnh đồng vị trí phải được nhận dạng. Nếu hình ảnh hiện thời là lát B, `collocated_from_10_flag` được báo hiệu trong tiêu đề của lát để chỉ báo liệu hình ảnh đồng vị trí là của `RefPicList0` hay `RefPicList1`.

Sau khi bộ giải mã video 30 nhận dạng danh sách hình ảnh tham chiếu bao gồm hình ảnh đồng vị trí, bộ giải mã video 30 có thể sử dụng một phần tử cú pháp khác (ví dụ, `collocated_ref_idx`), mà có thể được báo hiệu trong tiêu đề của lát, để nhận dạng hình ảnh (tức là, hình ảnh đồng vị trí) trong danh sách hình ảnh tham chiếu đã nhận dạng. Tức là, sau khi danh sách hình ảnh tham chiếu được nhận dạng, `collocated_ref_idx`, được báo hiệu trong tiêu đề của lát được sử dụng để nhận dạng hình ảnh này trong danh sách hình ảnh tham chiếu.

Bộ mã hóa video có thể nhận dạng PU đồng vị trí bằng cách kiểm tra hình ảnh đồng vị trí. TMVP có thể chỉ báo một trong hai thông tin chuyển động của PU dưới cùng bên phải của CU chứa PU đồng vị trí hoặc thông tin chuyển động của PU dưới cùng bên phải trong các PU trung tâm của CU chứa PU này. Do vậy, chuyển động của PU dưới cùng bên phải của CU chứa PU này hoặc chuyển động của PU dưới cùng bên phải trong các PU trung tâm chứa PU này được sử dụng. PU dưới cùng bên phải của CU chứa PU đồng vị trí có thể là PU phủ vị trí ngay dưới và bên phải của mẫu dưới cùng bên phải của khối dự báo của PU này. Nói cách khác, TMVP có thể chỉ báo thông tin chuyển động của PU mà ở trong hình ảnh tham chiếu và phủ vị trí được đặt ở cùng vị trí với góc dưới cùng bên phải của PU hiện thời hoặc TMVP có thể chỉ báo thông tin chuyển động của PU trong hình ảnh tham chiếu và phủ vị trí ở cùng vị trí với tâm của PU hiện thời.

Khi vector chuyển động được nhận diện bởi quy trình nêu trên (tức là, vector chuyển động của TMVP) được sử dụng để tạo ra ứng cử chuyển động đối với chế độ kết sắp hoặc AMVP, bộ mã hóa video có thể chia tỷ lệ vector chuyển động dựa vào vị trí thời gian (được phản ánh bởi giá trị POC). Chẳng hạn, bộ mã hóa video có thể làm tăng độ lớn của vector chuyển động ở mức lớn hơn khi chênh lệch giữa giá trị POC của hình ảnh hiện thời và hình ảnh tham chiếu lớn hơn chênh lệch giữa giá trị POC của hình ảnh hiện thời và hình ảnh tham chiếu là nhỏ hơn.

Chỉ số tham chiếu đích của tất cả các danh sách hình ảnh tham chiếu có thể có đối với ứng cử kết sắp thời gian suy ra được từ TMVP có thể luôn được thiết lập là 0. Tuy nhiên, đối với AMVP, chỉ số tham chiếu đích của tất cả các danh sách hình ảnh tham chiếu có thể có được thiết lập bằng chỉ số tham chiếu được giải mã. Nói cách khác, chỉ số tham chiếu đích của tất cả các danh sách hình ảnh tham chiếu có thể có đối với ứng cử kết sắp thời gian suy ra được từ TMVP luôn được thiết lập là 0 trong khi đối với AMVP, chỉ số này được đặc bằng chỉ số tham chiếu được giải mã. Trong HEVC, SPS có thể bao gồm cờ (ví dụ, `sps_temporal_mvp_enable_flag`) và tiêu đề của lát có thể bao gồm cờ (ví dụ, `pic_temporal_mvp_enable_flag`) khi `sps_temporal_mvp_enable_flag` bằng 1. Khi cả hai `pic_temporal_mvp_enable_flag` và `temporal_id` bằng 0 cho hình ảnh cụ thể, không có vector chuyển động nào từ các hình ảnh ở trước hình ảnh cụ thể đó theo thứ tự giải mã được sử dụng làm TMVP trong quá trình giải mã hình ảnh cụ thể hoặc hình ảnh ở sau hình ảnh cụ thể theo thứ tự giải mã.

Trong phần tiếp theo, kỹ thuật mã hóa đa khung hình (ví dụ, như trong H.264/MVC) và kỹ thuật mã hóa đa khung hình cùng với chiều sâu (ví dụ, như trong 3D-HEVC) được mô tả. Trước hết, kỹ thuật MVC được mô tả. Như nêu trên, MVC là dạng mở rộng mã hóa đa khung hình của ITU-T H.264/AVC. Theo MVC, dữ liệu cho nhiều khung hình được mã hóa theo thứ tự thời gian đầu tiên và do vậy, sự sắp xếp thứ tự giải mã được gọi là mã hóa thời gian đầu tiên. Cụ thể, các thành phần khung hình (tức là, các hình ảnh) cho mỗi trong nhiều khung hình ở khoảng cách về thời gian chung có thể được mã hóa, sau đó một tập hợp các thành phần khung hình khác đối với khoảng cách thời gian khác có thể được mã hóa và v.v. Đơn vị truy cập có thể bao gồm các hình ảnh mã hóa của tất cả các khung hình cho một khoảng cách thời gian đầu ra. Cần hiểu rằng thứ tự giải mã của đơn vị truy cập không nhất thiết phải giống với thứ tự xuất (hoặc hiển thị)

Thứ tự giải mã MVC thông thường (tức là, thứ tự dòng bit) được thể hiện trên Fig.2. Sự sắp xếp thứ tự giải mã được gọi là sự mã hóa theo thời gian đầu tiên. Lưu ý rằng thứ tự giải mã của đơn vị truy cập không thể giống với thứ tự xuất hoặc hiển thị. Fig.2, mỗi S0–S7 chỉ các khung hình khác nhau của video đa khung hình. Mỗi T0–T8 biểu diễn khoảng cách thời gian xuất. Đơn vị truy cập có thể bao gồm các hình ảnh mã hóa của tất cả các khung hình cho một khoảng cách thời gian xuất. Ví dụ, đơn vị truy cập thứ nhất có thể bao gồm tất cả các khung hình S0–S7 cho khoảng cách thời gian

T0, đơn vị truy cập thứ hai có thể bao gồm tất cả các khung hình S0–S7 cho khoảng cách thời gian T1, và v.v..

Để ngắn gọn, bản mô tả này có thể sử dụng các định nghĩa sau:

thành phần khung hình: Biểu diễn mã hóa của khung hình trong một đơn vị truy cập. Khi khung hình bao gồm cả biểu diễn kết cấu và chiều sâu mã hóa, thành phần khung hình bao gồm thành phần khung hình kết cấu và thành phần khung hình chiều sâu.

thành phần khung hình kết cấu: Biểu diễn mã hóa về kết cấu của khung hình trong một đơn vị truy cập.

thành phần khung hình chiều sâu: Biểu diễn mã hóa về chiều sâu của khung hình trong một đơn vị truy cập.

Trên Fig.2, mỗi khung hình bao gồm các tập hợp hình ảnh. Ví dụ, khung hình S0 bao gồm tập các hình ảnh 0, 8, 16, 24, 32, 40, 48, 56, và 64, khung hình S1 bao gồm tập các hình ảnh 1, 9, 17, 25, 33, 41, 49, 57, và 65, và v.v.. Để mã hóa video 3D, ví dụ, 3D-HEVC, mỗi hình ảnh có thể bao gồm hai hình ảnh thành phần; một hình ảnh thành phần được gọi là thành phần khung hình kết cấu và hình ảnh thành phần khác được gọi là thành phần khung hình chiều sâu. Thành phần khung hình kết cấu và thành phần khung hình chiều sâu trong tập hình ảnh của khung hình có thể được xem là tương ứng với nhau. Ví dụ, thành phần khung hình kết cấu trong tập hình ảnh của khung hình có thể được xem là tương ứng với thành phần khung hình chiều sâu trong tập hình ảnh của khung hình và ngược lại (tức là thành phần khung hình chiều sâu tương ứng với thành phần khung hình kết cấu trong tập này và ngược lại). Như được sử dụng theo sáng chế, thành phần khung hình kết cấu tương ứng với thành phần khung hình chiều sâu có thể được xem là thành phần khung hình kết cấu và thành phần khung hình chiều sâu là một phần của cùng khung hình của một đơn vị truy cập.

Thành phần khung hình kết cấu bao gồm nội dung hình ảnh thực được hiển thị. Ví dụ, thành phần khung hình kết cấu có thể bao gồm thành phần độ chói (Y) và độ màu (Cb và Cr). Thành phần khung hình chiều sâu có thể chỉ báo chiều sâu tương đối của điểm ảnh trong thành phần khung hình kết cấu tương ứng của nó. Ví dụ, thành phần khung hình chiều sâu là hình ảnh thang độ xám chỉ bao gồm giá trị độ chói. Nói cách khác, thành phần khung hình chiều sâu không thể truyền tải nội dung hình ảnh

bất kỳ, mà cung cấp thước đo về chiều sâu tương đối của điểm ảnh trong thành phần khung hình kết cấu.

Ví dụ, điểm ảnh hoàn toàn trắng trong thành phần khung hình chiều sâu chỉ báo rằng điểm ảnh hoặc các điểm ảnh tương ứng của nó trong thành phần khung hình kết cấu tương ứng là gần hơn từ tầm nhìn của người xem và điểm ảnh hoàn toàn đen trong thành phần khung hình chiều sâu chỉ báo rằng điểm ảnh tương ứng của nó trong thành phần khung hình kết cấu tương ứng là xa hơn từ tầm nhìn của người xem. Mức tối khác nhau của màu xám ở giữa màu đen và màu trắng chỉ báo mức chiều sâu khác nhau. Chẳng hạn, điểm ảnh rất xám trong thành phần khung hình chiều sâu chỉ báo rằng điểm ảnh tương ứng của nó trong thành phần khung hình kết cấu là xa hơn điểm ảnh ít xám trong thành phần khung hình chiều sâu. Do chỉ cần thang độ xám để nhận dạng chiều sâu của điểm ảnh, thành phần khung hình chiều sâu không cần bao gồm thành phần độ màu, vì giá trị độ màu cho thành phần khung hình chiều sâu không phục vụ cho mục đích bất kỳ.

Thành phần khung hình chiều sâu chỉ sử dụng giá trị độ chói (ví dụ, giá trị cường độ) để xác định chiều sâu được đưa ra để minh họa và không được xem là làm giới hạn. Theo các ví dụ khác, kỹ thuật bất kỳ có thể sử dụng để chỉ báo chiều sâu tương đối của các điểm ảnh trong thành phần khung hình kết cấu.

Cấu trúc dự báo MVC thông thường (bao gồm cả dự báo liên hình ảnh trong mỗi dự báo khung hình và dự báo giữa các cảnh nhìn) đối với mã hóa video đa khung hình được thể hiện trên Fig.3. Chiều dự báo được chỉ ra bởi mũi tên, đối tượng hướng đến sử dụng đối tượng xuất phát làm tham chiếu dự báo. Trong MVC, dự báo giữa các cảnh nhìn được hỗ trợ bởi việc bù chuyển động chênh lệch, sử dụng cú pháp của bù chuyển động H.264/AVC, nhưng cho phép hình ảnh trong khung hình khác được sử dụng làm hình ảnh tham chiếu.

Theo ví dụ của Fig.3, tám khung hình (có ID hình ảnh từ “S0” đến “S7”) được minh họa và 12 vị trí thời gian (“T0” đến “T11”) được minh họa cho mỗi khung hình. Tức là, mỗi hàng trên Fig.3 tương ứng với một khung hình, còn mỗi cột chỉ ra vị trí thời gian.

Mặc dù MVC có khung hình gốc, có thể được giải mã bởi bộ giải mã H.264/AVC và cặp khung hình nổi cũng được hỗ trợ bởi MVC, ưu điểm của MVC là

nó có thể hỗ trợ ví dụ sử dụng nhiều hơn hai khung hình dưới dạng đầu vào 3D và giải mã video 3D này được biểu diễn bởi đa khung hình. Bộ kết xuất của máy khách có bộ giải mã MVC có thể mong đợi dung lượng video 3D với đa khung hình.

Các hình ảnh trên Fig.3 được thể hiện ở vị trí giao cắt nhau của mỗi hàng và mỗi cột. Chuẩn H.264/AVC có thể sử dụng thuật ngữ khung hình để thể hiện một phần của video. Sáng chế có thể sử dụng thuật ngữ hình ảnh và khung hình một cách hoán đổi cho nhau.

Các hình ảnh trên Fig.3 được minh họa bằng cách sử dụng khối bao gồm một chữ cái, chữ cái này ký hiệu hình ảnh tương ứng được mã hóa liên hình ảnh (tức là, hình ảnh I), hay được mã hóa liên hình ảnh theo một chiều (tức là, hình ảnh P) hoặc theo nhiều chiều (tức là, hình ảnh B). Nói chung, việc dự báo được thể hiện theo mũi tên, trong đó hình ảnh hướng đến sử dụng hình ảnh đi ra để tham chiếu dự báo. Ví dụ, hình ảnh P của khung hình S2 ở vị trí thời gian T0 được dự báo từ hình ảnh I của khung hình S0 ở vị trí thời gian T0.

Giống như mã hóa video đơn khung hình, các hình ảnh của chuỗi video mã hóa video đa khung hình có thể được mã hóa dự báo liên quan đến các hình ảnh ở các vị trí thời gian khác nhau. Ví dụ, hình ảnh b của khung hình S0 ở vị trí thời gian T1 có mũi tên hướng đến hình ảnh này từ hình ảnh I của khung hình S0 ở vị trí thời gian T0, biểu thị rằng hình ảnh b được dự báo từ hình ảnh I. Ngoài ra, tuy nhiên, trong ngữ cảnh mã hóa đa khung hình, hình ảnh có thể được dự báo giữa các cảnh nhìn. Tức là, thành phần khung hình có thể sử dụng các thành phần khung hình trong các khung hình khác để tham chiếu. Trong MVC, ví dụ, dự báo giữa các cảnh nhìn được thực hiện khi thành phần khung hình trong hình ảnh khác là tham chiếu dự báo liên hình ảnh. Tham chiếu giữa các cảnh nhìn tiềm năng được báo hiệu trong dạng mở rộng MVC tập tham số chuỗi (Sequence Parameter Set-SPS) và có thể được biến đổi bởi quy trình tạo danh sách hình ảnh tham chiếu, có thể thực hiện sắp xếp linh động các tham chiếu dự báo liên hình ảnh hoặc dự báo giữa các cảnh nhìn. Dự báo giữa các cảnh nhìn cũng là một khía cạnh của dạng mở rộng đa khung hình được đề xuất của HEVC, bao gồm 3D-HEVC (đa khung hình cùng với chiều sâu).

Fig.3 đưa ra các ví dụ khác nhau về dự báo giữa các cảnh nhìn. Các hình ảnh của khung hình S1, theo ví dụ của Fig.3, được minh họa khi được dự báo từ các hình ảnh ở các vị trí thời gian khác nhau của khung hình S1, cũng như được dự báo giữa

các cảnh nhìn từ hình ảnh của khung hình S0 và S2 ở cùng một vị trí thời gian. Ví dụ, hình ảnh b của khung hình S1 ở vị trí thời gian T1 được dự báo từ mỗi hình ảnh trong các hình ảnh B của khung hình S1 ở vị trí thời gian T0 và T2, cũng như hình ảnh b của khung hình S0 và S2 ở vị trí thời gian T1.

Theo một số ví dụ, Fig.3 có thể được quan sát vì minh họa các thành phần khung hình kết cấu. Ví dụ, các hình ảnh I, P, B và b được minh họa trên Fig.2 có thể được coi là thành phần khung hình kết cấu cho mỗi khung hình. Theo kỹ thuật được mô tả trong sáng chế, đối với mỗi thành phần khung hình kết cấu được minh họa trên Fig.3 có thành phần khung hình chiều sâu tương ứng. Theo một số ví dụ, thành phần khung hình chiều sâu có thể được dự báo theo cách thức tương tự với cách thức được minh họa trên Fig.3 đối với thành phần khung hình kết cấu tương ứng.

Việc mã hóa hai khung hình cũng có thể được hỗ trợ trong MVC. Một trong những ưu điểm của MVC là một bộ mã hóa MVC có thể nhận hơn hai khung hình dưới dạng đầu vào video 3D và bộ giải mã MVC có thể giải mã biểu diễn đa khung hình này. Như vậy, bộ kết xuất đồ họa bất kỳ với bộ giải mã MVC kỳ vọng dung lượng video 3D có nhiều hơn hai khung hình.

Theo MVC, dự báo giữa các cảnh nhìn được được chấp nhận trong số các hình ảnh trong cùng một đơn vị truy cập (tức là, với cùng một khoảng cách thời gian). Khi việc mã hóa hình ảnh trong một trong số các khung hình không phải là gốc, một hình ảnh có thể được bổ sung vào danh sách hình ảnh tham chiếu nếu nó ở trong khung hình khác, nhưng vẫn ở cùng khoảng cách thời gian. Ảnh tham chiếu giữa các cảnh nhìn có thể đặt ở vị trí bất kỳ của danh sách hình ảnh tham chiếu, giống như hình ảnh tham chiếu dự báo liên hình ảnh bất kỳ. Như được thể hiện trên Fig.3, thành phần khung hình có thể sử dụng các thành phần khung hình ở các khung hình khác để tham chiếu. Theo MVC, dự báo giữa các cảnh nhìn được chấp nhận như là thành phần khung hình trong một khung hình khác là một tham chiếu dự báo liên hình ảnh.

Theo ngữ cảnh mã hóa video đa khung hình, nói chung, có hai loại vector chuyển động. Một vector được gọi là vector chuyển động thông thường. Vector chuyển động thông thường hướng đến hình ảnh tham chiếu thời gian và dự báo liên hình ảnh thời gian tương ứng là dự báo bù chuyển động (MCP). Vector chuyển động khác là vector chuyển động chênh lệch (disparity chuyển động vector-DMV). DMV hướng đến các hình ảnh trong khung hình khác (tức là, hình ảnh tham chiếu giữa các cảnh nhìn)

và dự báo liên hình ảnh tương ứng là dự báo bù chênh lệch (Disparity-compensated prediction-DCP).

Một kiểu định dạng mã hóa video đa khung hình khác đề xuất việc sử dụng giá trị chiều sâu (ví dụ, như trong 3D-HEVC). Đối với định dạng dữ liệu video đa khung hình cộng chiều sâu (multiview-video-plus-depth-MVD), mà là phổ biến đối với tivi 3D và các video tầm nhìn rộng, hình ảnh kết cấu và ánh xạ chiều sâu có thể được mã hóa cùng với các hình ảnh kết cấu đa khung hình một cách độc lập. Fig.4 minh họa định dạng dữ liệu MVD có hình ảnh kết cấu và ánh xạ chiều sâu cho mỗi mẫu liên quan của nó. Khoảng chiều sâu có thể được giới hạn nằm trong khoảng từ khoảng cách z_{near} nhỏ nhất đến z_{far} lớn nhất từ camera cho các điểm 3D tương ứng.

Các thông số của camera và giá trị khoảng chiều sâu có thể là hữu dụng để xử lý các thành phần khung hình mã hóa trước khi kết xuất đồ họa trên màn hình hiển thị 3D. Do vậy, thông điệp thông tin nâng cao bổ sung (Supplemental enhancement information-SEI) được định nghĩa cho phiên bản gần đây của H.264/MVC, tức là, thông tin tiếp nhận đa khung hình SEI, bao gồm thông tin chỉ rõ các tham số khác nhau của môi trường tiếp nhận. Tuy nhiên, không có cú pháp được chỉ rõ trong H.264/MVC để chỉ báo thông tin liên quan đến khoảng chiều sâu.

Phân chia chuyển động không đối xứng (AMP) và kích thước khối bù chuyển động trong HEVC sẽ được thảo luận. Trong HEVC, khối mã hóa liên hình ảnh có thể được phân chia thành một, hai hoặc bốn phần. Hình dạng khác nhau của các phần này là có thể. Các khả năng phân chia ví dụ đối với các khối mã hóa được dự báo liên hình ảnh được thể hiện trên Fig.5.

Hàng phía trên của sự phân chia trên Fig.5 minh họa các phân chia không đối xứng. Phân chia $N \times N$ đơn giản là khối mã hóa không được phân chia. Phân chia $N/2 \times N$ là khối mã hóa được phân chia thành hai phần hình chữ nhật theo chiều dọc. Phân chia $N/2 \times N$ là khối mã hóa được phân chia thành hai phần hình chữ nhật theo chiều ngang. Phân chia $N/2 \times N/2$ là khối mã hóa được phân chia thành bốn phần hình chữ nhật.

Các loại phân chia nhỏ hơn bốn trên Fig.5 được gọi là các phân chia không đối xứng và có thể được sử dụng trong việc phân chia chuyển động không đối xứng (AMP) cho dự báo liên hình ảnh. Một phân chia theo chế độ AMP có chiều cao hoặc

chiều rộng $N/4$ và chiều rộng hoặc chiều cao N và phần chia còn lại bao gồm phần còn lại của CB có chiều cao hoặc chiều rộng là $3N/4$ và chiều rộng hoặc chiều cao N . Mỗi phần chia được mã hóa liên hình ảnh được gán một hoặc hai vectơ chuyển động và chỉ số hình ảnh tham chiếu (tức là, một vectơ chuyển động và chỉ số tham chiếu đối với dự báo một chiều và hai vectơ chuyển động và chỉ số tham chiếu cho dự báo hai chiều). Theo một số ví dụ, để làm giảm thiểu băng thông nhớ trong trường hợp xấu nhất, các phần chia có kích thước 4×4 là không được phép đối với dự báo liên hình ảnh và các phần có kích thước 4×8 và 8×4 bị giới hạn ở mã hóa dự báo một chiều trên một danh sách dữ liệu dự báo.

Như được thảo luận một cách chi tiết dưới đây, sáng chế mô tả kỹ thuật dùng cho AMP khi sử dụng kết hợp với kỹ thuật mã hóa 3D-HEVC, bao gồm dự báo tổng hợp cảnh nhìn ngược (BVSP).

Phần sau đây mô tả danh sách ứng cử kết sắp trong HEVC. Ví dụ, danh sách ứng cử kết sắp này có thể được tạo ra bởi các bước sau. Đối với các ứng cử kết sắp không gian, bộ mã hóa video 20 và/hoặc bộ giải mã video 30 có thể nhận tối đa bốn ứng cử vectơ chuyển động không gian từ năm khối lân cận về mặt không gian, như được minh họa trên Fig.6.

Thứ tự trong đó bộ mã hóa video 20 và bộ giải mã video 30 có thể ước lượng các khối lân cận về mặt không gian như sau: bên trái (A1), bên trên (B1), ở trên bên phải (B0), ở dưới bên trái (A0), và bên trái ở trên (B2), như được thể hiện trên Fig.6. Theo một số ví dụ, quá trình cắt bớt có thể được áp dụng để loại bỏ các ứng cử vectơ chuyển động có thông tin chuyển động giống nhau (ví dụ, vectơ chuyển động hoặc chỉ số tham chiếu). Ví dụ, vectơ chuyển động hoặc chỉ số tham chiếu của B1 có thể được so sánh với vectơ chuyển động hoặc chỉ số tham chiếu của A1, vectơ chuyển động hoặc chỉ số tham chiếu của B0 có thể được so sánh với vectơ chuyển động hoặc chỉ số tham chiếu của B1, vectơ chuyển động hoặc chỉ số tham chiếu của A0 có thể được so sánh với vectơ chuyển động hoặc chỉ số tham chiếu của A1, và vectơ chuyển động hoặc chỉ số tham chiếu của B2 có thể được so sánh với vectơ chuyển động hoặc chỉ số tham chiếu của cả B1 và A1. Một trong hai ứng cử có thông tin chuyển động giống nhau sau đó có thể được loại bỏ ra khỏi danh sách ứng cử vectơ chuyển động. Nếu đã có bốn ứng cử khả dụng sau quá trình cắt bớt, ứng cử B2 không được chèn vào danh sách ứng cử kết sắp.

Ứng cử biến dự báo vector chuyển động thời gian đồng vị trí (temporal motion vector predictor - TMVP) từ hình ảnh tham chiếu, nếu được kích hoạt và khả dụng, được bổ sung vào danh sách ứng cử vector chuyển động sau ứng cử vector chuyển động không gian.

Nếu danh sách ứng cử vector chuyển động không hoàn chỉnh (ví dụ, có ít hơn số lượng mục nhập đã định trước, một hoặc nhiều ứng cử vector chuyển động nhân tạo có thể được tạo ra và được chèn ở cuối danh sách ứng cử kết sắp. Các loại ứng cử vector chuyển động **nhân tạo** ví dụ bao gồm các ứng cử kết sắp dự báo hai chiều được kết sắp chỉ được suy ra cho các lát B và các ứng cử kết sắp vector chuyển động không nếu không có đủ các ứng cử kết sắp dự báo hai chiều (hoặc loại ứng cử vector chuyển động nhân tạo khác) để cung cấp số lượng ứng cử vector chuyển động đã định trước.

Khi lát hiện thời là lát B, quy trình suy ra các ứng cử kết sắp dự báo hai chiều kết sắp được kích hoạt. Đối với mỗi cặp ứng cử đã có trong danh sách ứng cử và có thông tin chuyển động cần thiết, ứng cử vector chuyển động dự báo hai chiều kết hợp (có chỉ số được thể hiện bởi combIdx) được suy ra bằng cách sử dụng tổ hợp vector chuyển động của ứng cử thứ nhất (có chỉ số ứng cử kết sắp là l0CandIdx) chỉ hình ảnh trong danh sách 0 (nếu có), và vector chuyển động của ứng cử thứ hai (có chỉ số ứng cử kết sắp bằng l1CandIdx) chỉ hình ảnh trong danh sách 1 (nếu có và một trong hai hình ảnh tham chiếu hoặc vector chuyển động là khác so với ứng cử thứ nhất).

Fig.7 là bảng thể hiện chi tiết ví dụ về các chỉ số l0CandIdx và l1CandIdx trong 3D-HEVC. ví dụ, Fig.7 minh họa định nghĩa về l0CandIdx và l1CandIdx tương ứng với combIdx.

Đối với combIdx bằng 0 ... 11, quy trình tạo ứng cử vector chuyển động dự báo hai chiều kết hợp được kết thúc khi một trong các điều kiện sau là đúng: (1) combIdx bằng $(\text{numOrigMergeCand} * (\text{numOrigMergeCand} - 1))$ trong đó numOrigMergeCand thể hiện số ứng cử trong danh sách kết sắp trước khi kích hoạt quá trình này; (2) tổng số ứng cử (bao gồm các ứng cử kết sắp dự báo hai chiều mới được tạo ra) trong danh sách kết sắp bằng MaxNumMergeCand.

Phần này mô tả việc suy ra các ứng cử vector chuyển động kết sắp zero. Đối với mỗi ứng cử, vector chuyển động zero và chỉ số hình ảnh tham chiếu được đặt từ 0 đến chỉ số hình ảnh tham chiếu trừ đi 1. Nếu các ứng cử vẫn ít so với số ứng cử vector

chuyển động kết sắp lớn nhất (ví dụ, được thể hiện bằng phần tử cú pháp MaxNumMergeCand), chỉ số tham chiếu không và vector chuyển động được chèn vào cho đến khi tổng số ứng cử bằng MaxNumMergeCand.

Dưới đây mô tả ràng buộc về kích thước bù chuyển động theo HEVC. Để làm giảm thiểu băng thông nhớ trong trường hợp xấu nhất, các phần có kích thước 4x4 không được được phép để dự báo liên hình ảnh và các phần có kích thước 4x8 và 8x4 bị giới hạn ở mã hóa một chiều.

Để thỏa mãn ràng buộc nêu trên, khi kích thước PU hiện thời bằng 8x4 hoặc 4x8, ứng cử kết sắp dự báo hai chiều kết sắp không gian/thời gian, nếu kết sắp với chế độ dự báo hai chiều, PU hiện thời nên được thiết lập lại để sử dụng dự báo một chiều bằng cách biến đổi hướng dự báo cho danh sách 0, và chỉ số hình ảnh tham chiếu và vector chuyển động tương ứng với RefPicList1 lần lượt là -1 và (0, 0).

Như nêu trên, 3D-HEVC đang được phát triển. 3D-HEVC có thể cải thiện hiệu quả mã hóa bằng cách sử dụng dự báo chuyển động giữa các cảnh nhìn và dự báo số dư giữa các cảnh nhìn. Nói cách khác, để cải thiện thêm hiệu quả mã hóa, hai kỹ thuật mới là “dự báo chuyển động giữa các cảnh nhìn” và “dự báo số dư giữa các cảnh nhìn” được chấp nhận trong phần mềm tham chiếu. Trong dự báo chuyển động giữa các cảnh nhìn, bộ mã hóa video (ví dụ, bộ mã hóa video 20 hoặc bộ giải mã video 30) có thể xác định (tức là, dự báo) thông tin chuyển động của PU hiện thời dựa trên thông tin chuyển động của PU trong khung hình khác PU hiện thời. Trong dự báo số dư giữa các cảnh nhìn, bộ mã hóa video có thể xác định khối dư của CU hiện thời dựa trên dữ liệu dư trong khung hình khác CU hiện thời.

Việc suy ra vector chênh lệch dựa vào khối lân cận (NBDV) trong 3D-HEVC sẽ được mô tả. Việc suy ra NBDV được sử dụng dưới dạng kỹ thuật suy ra vector chênh lệch trong 3D-HEVC do thực tế là 3D-HEVC sử dụng thứ tự mã hóa kết cấu đầu tiên cho tất cả các khung hình. Vì ánh xạ chiều sâu tương ứng là không khả dụng đối với hình ảnh kết cấu mã hóa hiện thời, trong NBDV, vector chênh lệch suy ra được từ khối lân cận. Theo một số đề xuất đối với thiết kế 3D-HEVC, vector chênh lệch suy ra được từ quy trình suy ra NBDV có thể còn được lọc bằng cách tìm kiếm dữ liệu chiều sâu tương ứng với khung hình kết cấu tham chiếu.

3D-HEVC đầu tiên áp dụng kỹ thuật suy ra NBDV được đề xuất trong JCT3V-A0097 (3D-CE5.h: Disparity vector generation results, L. Zhang, Y. Chen, M. Karczewicz (Qualcomm)). Vector chênh lệch ẩn được được bao gồm cùng với NBDV đơn giản hóa trong JCTVC-A0126 (3D-CE5.h: Simplification of disparity vector derivation for HEVC-based 3D video coding, J. Sung, M. Koo, S. Yea (LG)). Ngoài ra, trong JCT3V-B0047 (3D-CE5.h đề cập đến: Improvements for disparity vector derivation, J. Kang, Y. Chen, L. Zhang, M. Karczewicz (Qualcomm)), kỹ thuật suy ra NBDV còn được đơn giản hóa bằng cách loại bỏ vector chênh lệch ẩn được lưu trữ trong bộ đệm hình ảnh được giải mã, nhưng vẫn cải thiện độ lợi mã hóa với việc chọn hình ảnh truy cập ngẫu nhiên (random access picture-RAP). Kỹ thuật khác đối với việc suy ra NBDV đã được mô tả trong JCT3V-D0181 (CE2: CCU-based Disparity Vector Derivation in 3D-HEVC, J. Kang, Y. Chen, L. Zhang, M. Karczewicz (Qualcomm)).

Vector chênh lệch (DV) được sử dụng làm ước lượng chuyển vị giữa hai khung hình. Do các khối lân cận gần như là cùng có chung một thông tin chuyển động/chênh lệch trong quá trình mã hóa video, khối hiện thời có thể sử dụng thông tin vector chuyển động ở các khối lân cận làm bộ chỉ báo tốt. Tiếp theo ý tưởng này, quy trình suy ra NBDV sử dụng thông tin chênh lệch lân cận để ước lượng vector chênh lệch ở các khung hình khác nhau.

Để thực hiện NBDV, bộ mã hóa video 2D ban đầu định nghĩa một số khối lân cận về mặt không gian và thời gian. Bộ mã hóa video 2D sau đó kiểm tra mỗi khối lân cận theo thứ tự định trước được xác định theo sự ưu tiên của mối tương quan giữa khối hiện thời và khối ứng cử. Ngay khi vector chuyển động chênh lệch (tức là, vector chuyển động hướng đến hình ảnh tham chiếu giữa các cảnh nhìn) được tìm thấy trong các ứng cử, bộ mã hóa video 2D biến đổi vector chuyển động chênh lệch thành vector chênh lệch và chỉ số thứ tự khung hình kết hợp quan cũng được phục hồi. Hai tập khối lân cận được sử dụng. Một tập bao gồm các khối lân cận về mặt không gian và tập còn lại bao gồm các khối lân cận thời gian.

Trong các đề xuất gần đây đối với 3D-HEVC, hai khối lân cận về mặt không gian được sử dụng trong việc suy ra NBDV. Các khối lân cận về mặt không gian ở bên trái và các khối lân cận ở trên so với đơn vị mã hóa hiện thời (CU) 90, như lần lượt được thể hiện bởi A1 và B1, trên Fig 8. Cần lưu ý rằng các khối lân cận được thể

hiện trên Fig.8 có cùng vị trí với một số khối lân cận được sử dụng trong chế độ MERGE theo HEVC. Do vậy, sự truy cập vào bộ nhớ bổ sung là không cần thiết. Tuy nhiên, cần hiểu rằng các khối lân cận ở các vị trí khác so với CU hiện thời 90 cũng có thể được sử dụng.

Để kiểm tra các khối lân cận thời gian, bộ mã hóa video 20 đầu tiên thực hiện quá trình tạo ra danh sách hình ảnh ứng cử. Có đến hai hình ảnh tham chiếu từ khung hình hiện thời có thể được xử lý dưới dạng hình ảnh ứng cử. Bộ mã hóa video 20 đầu tiên bổ sung hình ảnh tham chiếu đồng vị trí vào danh sách hình ảnh ứng cử, tiếp theo là phần còn lại của hình ảnh ứng cử theo thứ tự chỉ số tham chiếu tăng dần. Khi các hình ảnh tham chiếu có cùng chỉ số tham chiếu ở cả hai danh sách hình ảnh tham chiếu là khả dụng, hình ảnh tham chiếu trong cùng một danh sách hình ảnh tham chiếu là hình ảnh đồng vị trí đứng trước hình ảnh tham chiếu khác có cùng chỉ số tham chiếu. Đối với mỗi hình ảnh ứng cử trong danh sách hình ảnh ứng cử, bộ mã hóa video 20 xác định khối của vùng đồng vị trí bao phủ vị trí trung tâm làm khối lân cận thời gian.

Khi một khối được mã hóa bởi dự báo chuyển động giữa các cảnh nhìn, vector chênh lệch có thể được suy ra để chọn khối tương ứng trong khung hình khác. Vector chênh lệch suy ra được trong quá trình dự báo chuyển động giữa các cảnh nhìn được gọi là vector chênh lệch ẩn (IDV hoặc a.k.a. vector chênh lệch suy ra được. Ngay cả khi khối này được mã hóa bởi dự báo chuyển động, vector chênh lệch suy ra được không được loại bỏ để mã hóa khối tiếp theo.

Theo một thiết kế của HTML, trong quy trình suy ra NBDV, bộ mã hóa video (ví dụ, bộ mã hóa video 20 và/hoặc bộ giải mã video 30) được tạo cấu trúc để giám sát vector chuyển động chênh lệch theo thứ tự ở các khối lân cận thời gian, vector chuyển động chênh lệch ở các khối lân cận về mặt không gian và sau đó là IDV. Ngay khi vector chuyển động chênh lệch hoặc IDV được tìm thấy, quá trình này kết thúc.

Việc lọc của quy trình suy ra NBDV (NBDV-R) bằng cách đánh giá thông tin chiều sâu sẽ được thảo luận. Khi vector chênh lệch được nhận từ quy trình suy ra NBDV, vector chênh lệch suy ra được này có thể còn được lọc bằng cách tìm kiếm dữ liệu chiều sâu từ ánh xạ chiều sâu của khung hình tham chiếu. Quy trình lọc có thể bao gồm các kỹ thuật sau:

Định vị khối có chiều sâu tương ứng bởi vector chênh lệch suy ra được ở khung hình chiều sâu tham chiếu được mã hóa trước đó, như khung hình gốc, kích thước của khối có chiều sâu tương ứng là giống với kích thước của PU hiện thời.

Chọn một giá trị chiều sâu cho bộ điểm ảnh góc của khối có chiều sâu tương ứng và biến đổi giá trị chiều sâu thành phần ngang của vector chênh lệch. đã lọc Thành phần dọc của vector chênh lệch không thay đổi.

Lưu ý, theo một số ví dụ, vector chênh lệch đã lọc có thể được sử dụng cho dự báo chuyển động giữa các cảnh nhìn còn vector chênh lệch chưa được lọc có thể được dùng cho dự báo dư giữa các cảnh nhìn. Ngoài ra, vector chênh lệch đã lọc có thể được lưu giữ dưới dạng vector chuyển động của một PU nếu PU này được mã hóa bởi chế độ dự báo tổng hợp cảnh nhìn ngược. Theo một số đề xuất đối với 3D-HEVC, thành phần khung hình chiều sâu của khung hình gốc có thể được đánh giá bất kể giá trị của chỉ số thứ tự khung hình suy ra được từ quy trình suy ra NBDV.

Kỹ thuật dự báo tổng hợp cảnh nhìn ngược trong 3D-HEVC sẽ được thảo luận. Một phương pháp BVSP ví dụ, như được đề xuất bởi D. Tian, et al. “CE1.h: Backward View Synthesis Prediction Using Neighboring Blocks,” JCT3V-C0152, có sẵn từ http://phenix.it-sudparis.eu/jct2/doc_end_user/current_document.php?id=594, được thông qua ở cuộc họp lần thứ 3 của JCT-3V. Ý tưởng cơ bản của BVSP là tương tự với dự báo tổng hợp khung hình dựa trên khối trong 3D-AVC. Cả hai kỹ thuật này đều sử dụng dự báo tổng hợp khung hình lệch về phía sau và dựa trên khối để tránh truyền đi các chênh lệch vector chuyển động và sử dụng vector chuyển động chính xác hơn. Chi tiết thực hiện của BVSP trong 3D-HEVC và 3D-AVC là khác nhau do các nền tảng khác nhau.

Trong 3D-HEVC, chế độ BVSP được hỗ trợ cho khối mã hóa liên hình ảnh mà được mã hóa theo chế độ bỏ qua hoặc chế độ kết sắp. Theo một đề xuất ví dụ đối với 3D-HEVC, chế độ BVSP không được cho phép đối với khối được mã hóa theo chế độ dự báo vector chuyển động cải tiến (AMVP). Thay vì truyền cờ để chỉ báo sử dụng chế độ BVSP, bộ mã hóa video 20 có thể được tạo cấu hình để bổ sung một ứng cử kết sắp khác (tức là, ứng cử kết sắp BVSP) vào danh sách ứng cử kết sắp và mỗi ứng cử kết sắp với một cờ BVSP. Khi chỉ số kết sắp được giải mã tương ứng với ứng cử kết sắp BVSP, chỉ số kết sắp được giải mã chỉ ra rằng đơn vị dự báo hiện thời (PU) sử dụng chế độ BVSP. Đối với mỗi khối con trong PU hiện thời, vector chuyển động

chênh lệch có thể suy ra được bằng cách biến đổi giá trị chiều sâu trong khung hình tham chiếu chiều sâu.

Việc thiết lập cờ BVSP được định nghĩa như sau. Khi khối lân cận về mặt không gian sử dụng để suy ra ứng cử kết sắp không gian được mã hóa theo chế độ BVSP, thông tin chuyển động liên quan của ứng cử kết sắp không gian được kế thừa bởi khối hiện thời, như trong chế độ kết sắp thông thường. Ngoài ra, ứng cử kết sắp không gian này được kết sắp với cờ BVSP bằng 1 (tức là, chỉ báo rằng ứng cử kết sắp không gian được mã hóa theo chế độ BVSP). Đối với ứng cử kết sắp BVSP mới được đưa vào, cờ BVSP được đặt là 1. Đối với tất cả các ứng cử kết sắp khác, cờ BVSP kết sắp được đặt là 0.

Như nêu trên, trong 3D-HEVC, bộ mã hóa video 20 có thể được tạo cấu hình để nhận và chèn ứng cử mới, tức là ứng cử kết sắp BVSP, vào danh sách ứng cử kết sắp. Chỉ số hình ảnh tham chiếu tương ứng và vector chuyển động được thiết lập bằng phương pháp sau.

Bộ mã hóa video thứ nhất 20 có thể được tạo cấu hình để thu được chỉ số khung hình được thể hiện bởi phần tử cú pháp chỉ số hình ảnh (ví dụ, `refVIdxLX` ở 3D-HEVC) của vector chênh lệch suy ra được từ quy trình suy ra NBDV. Bộ mã hóa video 20 cũng có thể được tạo cấu hình để thu được danh sách hình ảnh tham chiếu (ví dụ, `RefPicListX` (`RefPicList0` hoặc `RefPicList1`)) kết hợp với hình ảnh tham chiếu có chỉ số thứ tự khung hình bằng `refVIdxLX`. Bộ mã hóa video 20 sau đó sử dụng chỉ số hình ảnh tham chiếu tương ứng và vector chênh lệch thu được từ quy trình suy ra NBDV làm thông tin chuyển động của ứng cử kết sắp BVSP trong `RefPicListX` (tức là, `RefPicList0` hoặc `RefPicList1`).

Nếu lát hiện thời là phần B, bộ mã hóa video 20 kiểm tra tính khả dụng của hình ảnh tham chiếu liên hình ảnh có chỉ số thứ tự khung hình được thể hiện bởi `refVIdxLY` không bằng `refVIdxLX` trong danh sách hình ảnh tham chiếu không phải là `RefPicListX`, tức là `RefPicListY` với Y bằng $1-X$. Nếu một hình ảnh tham chiếu khung hình khác được tìm thấy, bộ mã hóa video 20 thực hiện dự báo tổng hợp khung hình dự báo hai chiều. Bộ mã hóa video 20 có thể còn được tạo cấu hình để sử dụng chỉ số hình ảnh tham chiếu tương ứng của hình ảnh tham chiếu giữa các cảnh nhìn khác và vector chênh lệch được chia tỷ lệ từ quy trình suy ra NBDV làm thông tin chuyển động của ứng cử kết sắp BVSP trong `RefPicListY`. Khối có chiều sâu từ

khung hình có chỉ số thứ tự khung hình là refVIdxLX được sử dụng làm thông tin chiều sâu của khối hiện thời (trong trường hợp thứ tự mã hóa kết cấu đầu tiên). Bộ mã hóa video 20 tổng hợp hai hình ảnh tham chiếu giữa các cảnh nhìn khác (một từ mỗi danh sách hình ảnh tham chiếu) bằng quy trình làm lệch về phía sau và còn lấy trọng số hình ảnh tham chiếu đã tổng hợp để thu được bộ dự báo BVSP cuối.

Đối với các loại lát khác với lát B (ví dụ, lát P), bộ mã hóa video 20 sử dụng dự báo tổng hợp khung hình một chiều với RefPicListX làm danh sách hình ảnh tham chiếu để dự báo.

Trong 3D-HTM, mã hóa kết cấu đầu tiên được áp dụng trong các điều kiện thử nghiệm thông thường. Vì thành phần kết cấu của một khung được mã hóa trước thành phần chiều sâu, thành phần chiều sâu không phải gốc tương ứng là không khả dụng khi giải mã một thành phần kết cấu không phải gốc. Do vậy, bộ giải mã video 30 có thể được tạo cấu hình để ước lượng thông tin chiều sâu và sau đó sử dụng thông tin chiều sâu đã ước lượng để thực hiện BVSP. Để ước lượng thông tin chiều sâu cho một khối, được đề xuất đầu tiên suy ra vector chênh lệch từ các khối lân cận (ví dụ, sử dụng quy trình suy ra NBDV) và sau đó sử dụng vector chênh lệch suy ra được để tạo ra khối có chiều sâu từ khung hình tham chiếu.

Fig.9 minh họa kỹ thuật ví dụ để định vị khối có chiều sâu từ khung hình tham chiếu và sau đó sử dụng khối có chiều sâu để dự báo BVSP. Đầu tiên, bộ mã hóa video 20 và/hoặc bộ giải mã video 30 có thể sử dụng vector chênh lệch 104 kết sắp với khối lân cận 102. Tức là, bộ mã hóa video 20 và/hoặc bộ giải mã video 30 có thể truy cập thông tin vector chênh lệch từ các khối lân cận đã được mã hóa (như khối lân cận 102) và tái sử dụng thông tin vector chênh lệch kết hợp bất kỳ cho khối hiện thời 100. Vector chênh lệch 104 hướng đến khối có chiều sâu 106 ở hình ảnh có chiều sâu tham chiếu. Khi vector chênh lệch 104 được tái sử dụng cho khối hiện thời 100, vector chênh lệch 104 bây giờ hướng đến khối có chiều sâu 108 ở hình ảnh có chiều sâu tham chiếu. Khối có chiều sâu 108 tương ứng với khối hiện thời 100. Bộ mã hóa video 20 và/hoặc bộ giải mã video 30 sau đó có thể sử dụng thông tin chiều sâu ở hình ảnh có chiều sâu tham chiếu 108 để tổng hợp khối ở hình ảnh kết cấu tham chiếu bằng cách sử dụng kỹ thuật làm lệch về phía sau. Hình ảnh kết cấu tổng hợp sau đó có thể được sử dụng làm hình ảnh tham chiếu để dự báo khối hiện thời 100.

Theo một ví dụ của sáng chế, đối với quy trình suy ra NBDV, let (dv_x, dv_y) thể hiện vector chênh lệch 104 được xác định bởi quy trình suy ra NBDV, và thể hiện vị trí của khối hiện thời 100 là $(block_x, block_y)$. Theo một ví dụ của BVSP dự báo một chiều, bộ mã hóa video 20 và/hoặc bộ giải mã video 30 có thể được tạo cấu hình để tìm nạp khối có chiều sâu 108 có vị trí ở trên cùng bên trái $(block_x+dv_x, block_y+dv_y)$ trong thành phần khung hình chiều sâu của khung hình tham chiếu. Bộ mã hóa video 20 và/hoặc bộ giải mã video 30 có thể được tạo cấu hình để đầu tiên phân chia khối hiện thời 100 (ví dụ, PU) thành một vài khối con, mỗi khối có kích thước bằng nhau (ví dụ, bằng $W \times H$). Đối với mỗi khối con có kích thước bằng $W \times H$, bộ mã hóa video 20 và/hoặc bộ giải mã video 30 xác định giá trị chiều sâu lớn nhất từ bốn điểm ảnh góc của khối con chiều sâu tương ứng 108 trong thành phần khung hình chiều sâu được tìm nạp, ví dụ, như được thể hiện trên Fig.10. Fig.10 là sơ đồ khái niệm minh họa bốn điểm ảnh góc của khối có chiều sâu 8×8 110. Bốn điểm ảnh góc có thể được đánh dấu là điểm ảnh trên cùng bên trái (top-left-TL), trên cùng bên phải (Top-right-TR), dưới cùng bên trái (bottom-left-BL) và điểm ảnh dưới cùng bên phải (bottom-right-BR). Bộ mã hóa video 20 và/hoặc bộ giải mã video 30 biến đổi giá trị chiều sâu lớn nhất thành vector chuyển động chênh lệch. Vector chuyển động chênh lệch suy ra được cho mỗi khối con sau đó được sử dụng để bù chuyển động.

Phần này sẽ mô tả BVSP khi thực hiện dự báo hai chiều. Khi có nhiều hình ảnh tham chiếu giữa các cảnh nhìn từ các khung hình khác nhau trong RefPicList0 và RefPicList1, bộ mã hóa video 20 và/hoặc bộ giải mã video 30 áp dụng BVSP dự báo hai chiều. Trong BVSP dự báo hai chiều, hai bộ dự báo tổng hợp khung hình (tức là, hai tổng hợp tham chiếu tổng hợp) sẽ được tạo ra từ mỗi danh sách tham chiếu, như được mô tả trên đây. Sau đó hai bộ mã hóa dự báo tổng hợp khung hình được lấy trung bình để tạo ra bộ dự báo dự báo tổng hợp khung hình cuối.

Kích thước bù chuyển động, tức là, $W \times H$ như nêu trên, có thể là 8×4 hoặc 4×8 . Theo một ví dụ, để xác định kích thước bù chuyển động, quy tắc sau được áp dụng:

Đối với mỗi khối 8×8 , bốn góc của khối có chiều sâu tương ứng 8×8 được kiểm tra và:

```

if (vdepth[TL] < vdepth[BR]?0:1) != (vdepth[TR] < vdepth[BL]?0:1)
    use 4x8partition ( W = 4, H = 8 )
else
    use 8x4partition ( W = 8, H = 4 )

```

Dưới đây mô tả quy trình suy ra ứng cử giữa các cảnh nhìn cho chế độ bỏ qua/kết sắp trong một đề xuất đối với 3D-HEVC. Dựa vào vector chênh lệch suy ra được từ quy trình suy ra NBDV, ứng cử vector chuyển động mới được gọi là ứng cử vector chuyển động dự báo giữa các cảnh nhìn (Inter-view Predicted Motion Vector Candidate-IPMVC), nếu có, có thể được bổ sung vào AMVP và danh sách ứng cử vector chuyển động theo chế độ bỏ qua/kết sắp. Vector chuyển động dự báo giữa các cảnh nhìn, nếu có, là vector chuyển động thời gian. Vì chế độ bỏ qua có quy trình suy ra vector chuyển động giống như chế độ kết sắp, tất cả các kỹ thuật được mô tả trong tài liệu này áp dụng cho cả chế độ kết sắp và bỏ qua.

Fig.11 là sơ đồ khái niệm minh họa việc tạo ra ứng cử vector chuyển động dự báo giữa các cảnh nhìn ví dụ đối với chế độ kết sắp/bỏ qua. ví dụ, Fig.11 thể hiện một ví dụ về quy trình suy ra ứng cử vector chuyển động dự báo giữa các cảnh nhìn. Đối với chế độ kết sắp/bỏ qua, ứng cử vector chuyển động dự báo giữa các cảnh nhìn được suy ra theo các bước sau. Trước hết, bộ mã hóa video 20 và/hoặc bộ giải mã video 30 định vị, sử dụng vector chênh lệch, khối tương ứng (ví dụ, khối tham chiếu) 112 của PU/CU hiện thời 114 trong khung hình tham chiếu của cùng một đơn vị truy cập. Theo ví dụ của Fig.11, khối hiện thời (PU hiện thời) 114 là ở trong khung hình V1, trong khi đó khối tham chiếu tương ứng 112 là ở trong khung hình V0. Nếu khối tham chiếu tương ứng 112 không được mã hóa trong hình ảnh và mã hóa giữa các cảnh nhìn và hình ảnh tham chiếu của nó (trong ví dụ này trong khung V0 và thời gian T1) có giá trị POC bằng giá trị của một đầu vào trong cùng danh sách hình ảnh tham chiếu của PU/CU 114 hiện thời, thông tin chuyển động (tức là, chiều dự báo, chỉ số hình ảnh tham chiếu và vector chuyển động) của khối tham chiếu tương ứng 112 được suy ra là vector chuyển động dự báo giữa các cảnh nhìn sau khi biến đổi chỉ số tham chiếu dựa vào POC của hình ảnh tham chiếu.

Khối tham chiếu tương ứng 112 có thể được xác định như sau. Trước hết là vị trí độ sáng (xP, yP) của khối tín hiệu chói trên cùng bên trái của đơn vị dự báo hiện thời so với khối tín hiệu chói trên cùng bên trái của hình ảnh hiện thời. Các biến số nPSW và nPSH lần lượt là chiều rộng và chiều dài của đơn vị dự báo hiện thời. Chỉ số

thứ tự khung hình tham chiếu được đánh dấu là refViewIdx, và vector chênh lệch là mvDisp. Vị trí của độ sáng lớp tham chiếu (xRef, yRef) nhận được bởi:

$$xRef = Clip3(0, PicWidthInSamples_L - 1, xP + ((nPSW - 1) >> 1) + ((mvDisp[0] + 2) >> 2)) \quad (H-124)$$

$$yRef = Clip3(0, PicHeightInSamples_L - 1, yP + ((nPSH - 1) >> 1) + ((mvDisp[1] + 2) >> 2)) \quad (H-125)$$

Khối tham chiếu tương ứng 112 được đặt là đơn vị dự báo bao gồm vị trí độ sáng (xRef, yRef) trong thành phần khung hình có ViewIdx bằng refViewIdx.

Ngoài ra, vector chênh lệch có thể biến đổi thành vector chuyển động chênh lệch giữa các cảnh nhìn (inter-view disparity motion vector-IDMVC), được bổ sung vào danh sách ứng cử kết sắp ở một vị trí khác với IPMVC. Vector chuyển động chênh lệch giữa các cảnh nhìn cũng có thể được bổ sung vào danh sách ứng cử AMVP ở vị trí giống như IPMVC, khi khả dụng. IPMVC hoặc IDMVC có thể được gọi là ‘ứng cử giữa các cảnh nhìn’ trong ngữ cảnh này.

Theo một ví dụ đối với chế độ kết sắp/bỏ qua, IPMVC, nếu có, được chèn vào trước tất cả các ứng cử kết sắp không gian và thời gian vào danh sách ứng cử kết sắp. IDMVC được chèn vào trước ứng cử kết sắp không gian được nhận từ A₀.

Phần dưới đây mô tả cấu trúc của danh sách ứng cử kết sắp để mã hóa kết cấu trong 3D-HEVC. Trước hết, bộ mã hóa video 20 và/hoặc bộ giải mã video 30 suy ra vector chênh lệch, ví dụ, sử dụng kỹ thuật suy ra NBDV nêu trên. Sau khi suy ra vector chênh lệch, bộ mã hóa video 20 và/hoặc bộ giải mã video 30 có thể được tạo cấu hình để thực hiện quá trình tạo ra danh sách ứng cử kết sắp trong 3D-HEVC như được mô tả dưới đây.

Bộ mã hóa video 20 và/hoặc bộ giải mã video 30 có thể nhận một hoặc nhiều IPMVC bằng cách sử dụng quy trình được mô tả dưới đây. Nếu IPMVC là khả dụng, IPMV có thể được chèn vào danh sách kết sắp.

Tiếp đó, bộ mã hóa video 20 và/hoặc bộ giải mã video 30 có thể được tạo cấu hình để suy ra các ứng cử kết sắp không gian và một hoặc nhiều IDMVC chèn trong 3D-HEVC. Để suy ra các ứng cử kết sắp không gian, bộ mã hóa video 20 và/hoặc bộ giải mã video 30 có thể được tạo cấu hình để kiểm tra thông tin chuyển động của các

PU lân cận về mặt không gian theo thứ tự sau: A_1 , B_1 , B_0 , A_0 , hoặc B_2 , ví dụ, như được thể hiện trên Fig.6.

Bộ mã hóa video 20 và/hoặc bộ giải mã video 30 có thể còn được tạo cấu hình để thực hiện việc cắt bớt ràng buộc. Để thực hiện việc cắt bớt ràng buộc, bộ mã hóa video 20 và/hoặc bộ giải mã video 30 có thể được tạo cấu hình để không chen ứng cử kết sắp không gian ở vị trí A_1 vào danh sách ứng cử kết sắp nếu A_1 và IPMVC có cùng vector chuyển động và cùng chỉ số tham chiếu. Theo cách khác, ứng cử kết sắp không gian ở vị trí A_1 được chen vào danh sách ứng cử kết sắp.

Nếu ứng cử kết sắp ở vị trí B_1 và ứng cử kết sắp ở vị trí kết sắp A_1 (hoặc IPMVC) có cùng vector chuyển động và cùng chỉ số tham chiếu, ứng cử kết sắp ở vị trí B_1 không được chen vào danh sách ứng cử kết sắp. Theo cách khác, ứng cử kết sắp ở vị trí B_1 được chen vào danh sách ứng cử kết sắp. Nếu ứng cử kết sắp ở vị trí B_0 là khả dụng (tức là, được mã hóa và có thông tin chuyển động), ứng cử kết sắp ở vị trí B_0 được bổ sung vào danh sách ứng cử. Bộ mã hóa video 20 và/hoặc bộ giải mã video 30 nhận IDMVC bằng cách sử dụng thủ tục được mô tả trên đây. Nếu IDMVC là khả dụng và thông tin chuyển động của IDMVC là khác với các ứng cử suy ra được từ A_1 và B_1 , IDMVC được chen vào danh sách ứng cử.

Nếu BVSP có thể thực hiện đối với toàn bộ hình ảnh (hoặc cho phần hiện thời), thì ứng cử kết sắp BVSP được chen vào danh sách ứng cử kết sắp. Nếu ứng cử kết sắp ở vị trí A_0 là khả dụng, nó được bổ sung vào danh sách ứng cử. Nếu ứng cử kết sắp ở vị trí B_2 là khả dụng, nó được bổ sung vào danh sách ứng cử.

Phần tiếp theo sẽ thảo luận quy trình suy ra đối với ứng cử kết sắp thời gian trong 3D-HEVC. Việc suy ra ứng cử kết sắp thời gian trong 3D-HEVC là giống với quy trình suy ra ứng cử kết sắp thời gian trong HEVC, trong đó thông tin chuyển động của PU đồng vị trí được sử dụng. Tuy nhiên, đối với 3D-HEVC, chỉ số hình ảnh tham chiếu đích của ứng cử kết sắp thời gian có thể thay đổi thay vì cố định chỉ số hình ảnh tham chiếu là 0. Khi chỉ số tham chiếu đích bằng 0 tương ứng với hình ảnh tham chiếu thời gian (trong cùng một khung hình), còn vector chuyển động của đơn vị dự báo đồng vị trí (PU) hướng đến hình ảnh tham chiếu giữa các cảnh nhìn, chỉ số tham chiếu đích được thay đổi thành chỉ số khác tương ứng với đầu vào thứ nhất của hình ảnh tham chiếu giữa các cảnh nhìn trong danh sách hình ảnh tham chiếu. Ngược lại, khi chỉ số tham chiếu đích bằng 0 tương ứng với hình ảnh tham chiếu giữa các cảnh nhìn,

còn vector chuyển động của đơn vị dự báo đồng vị trí (PU) hướng đến hình ảnh tham chiếu thời gian, chỉ số tham chiếu đích được thay đổi thành chỉ số khác tương ứng với đầu vào thứ nhất của hình ảnh tham chiếu thời gian trong danh sách hình ảnh tham chiếu.

Quy trình suy ra ví dụ đối với các ứng cử kết sắp dự báo hai chiều kết sắp trong 3D-HEVC sẽ được thảo luận. Nếu tổng số lượng ứng cử suy ra được từ hai bước nêu trên (tức là, suy ra các ứng cử kết sắp không gian và suy ra ứng cử kết sắp thời gian) nhỏ hơn số lượng ứng cử tối đa (có thể được định trước), quy trình giống như được định nghĩa trong HEVC, như nêu trên, được thực hiện. Tuy nhiên, đặc điểm của các chỉ số tham chiếu `l0CandIdx` và `l1CandIdx` là khác nhau. Fig.12 là bảng khác thể hiện đặc điểm ví dụ của các chỉ số `l0CandIdx` và `l1CandIdx` trong 3D-HEVC. Ví dụ, mối liên hệ giữa `combIdx`, `l0CandIdx` và `l1CandIdx` được xác định trong bảng được minh họa trên Fig.12.

Một quy trình suy ra ví dụ cho ứng cử kết sắp vector chuyển động không theo 3D-HEVC là thủ tục giống như được định nghĩa trong HEVC. Theo một ví dụ cho 3D-HEVC, tổng số ứng cử trong danh sách ứng cử kết sắp tối đa là 6, và phần tử cú pháp `five_minus_max_num_merge_cand` được tạo ra trong tiêu đề của lát để mô tả số ứng cử kết sắp tối đa trừ đi 6. Cần lưu ý rằng giá trị của phần tử cú pháp `five_minus_max_num_merge_cand` nằm trong khoảng từ 0 đến 5.

Dưới đây mô tả sự kế thừa vector chuyển động (motion vector inheritance-MVI) để mã hóa chiều sâu, ví dụ, trong 3D-HEVC. Kỹ thuật MVI tìm kiếm khai thác sự giống nhau về các đặc điểm chuyển động giữa các thành phần kết cấu của một hình ảnh và các thành phần khung hình chiều sâu kết hợp của chúng. Fig.13 là sơ đồ khái niệm minh họa việc tạo ra ứng cử vector chuyển động kế thừa ví dụ để mã hóa chiều sâu. ví dụ, Fig.13 thể hiện ví dụ về quy trình suy ra ứng cử MVI trong đó khối kết cấu tương ứng 120 được chọn làm khối 4x4 được đặt ở dưới cùng bên phải của tâm của PU hiện thời 122 trong hình ảnh kết cấu 124. Đối với PU hiện thời 126 trong hình ảnh có chiều sâu 128, ứng cử MVI tái sử dụng vector chuyển động và các chỉ số tham chiếu kết hợp với khối kết cấu tương ứng đã được mã hóa 120 trong hình ảnh kết cấu tương ứng 124, nếu thông tin này là khả dụng.

Cần lưu ý rằng vector chuyển động có độ chính xác là số nguyên được sử dụng trong mã hóa chiều sâu, còn vector chuyển động có độ chính xác là một phần tư được

sử dụng để mã hóa kết cấu. Do vậy, vector chuyển động của khối kết cấu tương ứng có thể được chia tỷ lệ trước khi sử dụng ứng cử MVI.

Với việc tạo ra ứng cử MVI, danh sách ứng cử kết sắp cho khung hình chiều sâu được tạo ra như sau. Để chèn MVI, MVI được suy ra bằng cách sử dụng kỹ thuật như nêu trên và nếu khả dụng được chèn vào danh sách ứng cử kết sắp.

Quy trình suy ra đối với các ứng cử kết sắp không gian và chèn IDMVC trong 3D-HEVC đối với mã hóa chiều sâu được mô tả dưới đây. Bộ mã hóa video 20 và/hoặc bộ giải mã video 30 có thể được tạo cấu hình để kiểm tra thông tin chuyển động của các PU lân cận về mặt không gian theo thứ tự sau: A_1 , B_1 , B_0 , A_0 , hoặc B_2 .

Bộ mã hóa video 20 và/hoặc bộ giải mã video 30 có thể sau đó thực hiện việc cắt bớt ràng buộc như sau. Nếu ứng cử vector chuyển động ở vị trí A_1 và ứng cử MVI có cùng vector chuyển động và cùng chỉ số tham chiếu, ứng cử vector chuyển động ở vị trí A_1 không được chèn vào danh sách ứng cử kết sắp. Nếu ứng cử vector chuyển động ở vị trí B_1 và ứng cử vector chuyển động ở vị trí A_1 /MVI có cùng vector chuyển động và cùng chỉ số tham chiếu, ứng cử vector chuyển động ở vị trí B_1 không được chèn vào danh sách ứng cử kết sắp. Nếu ứng cử vector chuyển động ở vị trí B_0 là khả dụng, ứng cử vector chuyển động ở vị trí B_0 được bổ sung vào danh sách ứng cử kết sắp. Nếu ứng cử vector chuyển động ở vị trí A_0 là khả dụng, ứng cử vector chuyển động ở vị trí A_0 được bổ sung vào danh sách ứng cử kết sắp. Nếu ứng cử vector chuyển động ở vị trí B_2 là khả dụng, ứng cử vector chuyển động ở vị trí B_2 được bổ sung vào danh sách ứng cử kết sắp.

Quy trình suy ra ứng cử kết sắp thời gian theo mã hóa chiều sâu 3D-HEVC là giống với quy trình suy ra ứng cử kết sắp thời gian trong HEVC, trong đó thông tin chuyển động của PU đồng vị trí được sử dụng. Tuy nhiên, đối với mã hóa chiều sâu 3D-HEVC, chỉ số hình ảnh tham chiếu đích của ứng cử kết sắp thời gian có thể thay đổi, như được giải thích ở trên, thay vì cố định là 0.

Quy trình suy ra đối với các ứng cử kết sắp dự báo hai chiều kết hợp theo mã hóa chiều sâu 3D-HEVC sẽ được mô tả. Nếu tổng số lượng ứng cử suy ra được từ hai bước trên là nhỏ hơn số lượng ứng cử tối đa, quy trình giống như được định nghĩa trong HEVC được thực hiện, ngoại trừ đặc điểm của $l0CandIdx$ và $l1CandIdx$. Mỗi

liên hệ giữa $combIdx$, $l0CandIdx$ và $l1CandIdx$ được xác định trong bảng được minh họa trên Fig.12.

Quy trình suy ra các ứng cử kết sắp vector chuyển động không trong mã hóa chiều sâu 3D-HEVC là giống như quá trình được mô tả trong HEVC.

Dưới đây mô tả kỹ thuật ví dụ đối với dự báo số dư cải tiến (ARP). ARP áp dụng cho các CU có chế độ phân chia bằng với $Part_2Nx2N$ (ví dụ, NxN trên Fig.5) được thông qua ở cuộc họp JCT3V lần thứ 4, như được đề xuất trong JCT3V-D0177. Tài liệu JCT3V-D0177 có tựa đề “CE4: Advanced residual prediction for multiview coding,” by Zhang et al. Tài liệu JCT3V-D0177 là có sẵn, kể từ 22/8/2014, từ http://phenix.it-sudparis.eu/jct3v/doc_end_user/current_document.php?id=862.

Fig.14 minh họa cấu trúc dự báo của dự báo số dư cải tiến (advanced residual prediction - ARP) trong mã hóa video đa khung hình. Như được thể hiện trên Fig.14, các khối sau được kích hoạt trong dự báo khối hiện thời (“Curr”) 140. Khối tham chiếu 142 trong khung hình tham chiếu/gốc 144 nhận được bởi vector chênh lệch (DV) 146 được gắn nhãn là “gốc.” Khối 148 trong cùng khung hình (khung hình V_m) giống như khối hiện thời Curr 140, suy ra được bởi vector chuyển động 150 (thời gian) (được thể hiện là TMV) của khối hiện thời 140 được gắn nhãn là “CurrTRef.” Khối 152 trong cùng khung hình giống như khối Base 142 (khung hình V_0), suy ra được bởi vector chuyển động thời gian của khối hiện thời (TMV) được gắn nhãn là “BaseTRef. Khối tham chiếu BaseTRef 152 được xác định bằng vector $TMV+DV$ 154 được so sánh với khối hiện thời Curr 140.

Bộ dự báo dư được ký hiệu là BaseTRef-Base, trong đó phép trừ áp dụng cho mỗi điểm ảnh của dãy điểm ảnh được ký hiệu. Hệ số gia trọng “w” có thể còn được nhân với bộ dự báo dư. Do vậy, bộ dự báo cuối của khối hiện thời Curr có thể được ký hiệu là $CurrTRef + w \cdot (BaseTRef - Base)$.

Lưu ý trong phần mô tả nêu trên và Fig.14, giả sử rằng dự báo một chiều được áp dụng. Khi mở rộng ARP thành dự báo hai chiều, các bước sau được áp dụng cho mỗi danh sách hình ảnh tham chiếu. Khi khối hiện thời Curr sử dụng hình ảnh tham chiếu giữa các cảnh nhìn (trong một khung hình khác) cho một trong hai danh sách hình ảnh tham chiếu, quy trình ARP không thể thực hiện được.

Dưới đây mô tả quy trình giải mã theo ARP. Trước hết, bộ giải mã video 30 nhận được vector chênh lệch (ví dụ, bằng cách sử dụng quy trình suy ra NBDV) hướng đến khung hình tham chiếu đích. Sau đó, trong hình ảnh của khung hình tham chiếu trong cùng một đơn vị truy cập, bộ giải mã video 30 định vị khối tương ứng bằng cách sử dụng vector chênh lệch.

Bộ giải mã video 30 có thể tái sử dụng thông tin chuyển động của khối hiện thời để suy ra được thông tin chuyển động cho khối tham chiếu. Bộ giải mã video 30 có thể sau đó áp dụng bù chuyển động cho khối tương ứng dựa trên cùng một vector chuyển động của khối hiện thời, và hình ảnh tham chiếu suy ra được để tạo ra khối dư.

Fig.15 là sơ đồ khái niệm minh họa mối liên hệ ví dụ giữa khối hiện thời 160, khối tham chiếu 162 và khối bù chuyển động 164 và 166. Hình ảnh tham chiếu trong khung hình tham chiếu (V_0) có giá trị POC (Tổng số thứ tự hình ảnh) giống như hình ảnh tham chiếu của khung hình hiện thời (V_m) được chọn làm hình ảnh tham chiếu của khối tương ứng 162. Bộ mã hóa video 20 và/hoặc bộ giải mã video 30 có thể áp dụng hệ số gia trọng cho khối dư tạo ra khối dư được gia trọng và bổ sung khối dư gia trọng này vào mẫu dự báo.

Dưới đây mô tả hệ số gia trọng. Ba hệ số gia trọng được sử dụng trong ARP, tức là 0, 0,5 và 1. Hệ số gia trọng làm cho phí tổn tỉ lệ méo nhỏ nhất đối với CU hiện thời được chọn làm hệ số gia trọng cuối cùng và chỉ số hệ số gia trọng tương ứng (ví dụ, 0, 1 và 2 tương ứng với hệ số gia trọng 0, 1, và 0,5) được truyền trong dòng bit ở cấp CU. Theo một ví dụ của ARP, tất cả các dự báo PU trong một CU có cùng hệ số gia trọng. Khi hệ số gia trọng bằng 0, ARP không được sử dụng cho CU hiện thời.

Dưới đây mô tả một số dạng rút gọn thêm đối với ARP. Trước hết, việc chọn hình ảnh tham chiếu bằng cách chia tỷ lệ vector chuyển động được mô tả. Thứ hai, bộ lọc nội suy được mô tả.

Để chọn hình ảnh tham chiếu bằng cách chia tỷ lệ vector chuyển động, trong JCT3V-C0049, hình ảnh tham chiếu của đơn vị dự báo được mã hóa với hệ số gia trọng khác zero có thể khác nhau đối với mỗi khối. Tài liệu JCT3V-C0049 có tựa đề “3D-CE4: Advanced residual prediction for multiview coding,” bởi Zhang et al. Tài liệu JCT3V-C0049 có sẵn, từ 23/9/ 2013, từ http://phenix.int-evry.fr/jct3v/doc_end_user/current_document.php?id=487.

Do vậy, các hình ảnh khác nhau từ khung hình tham chiếu có thể cần được truy cập để tạo ra khối bù chuyển động (ví dụ, BaseTRef trên Fig.14) của khối tương ứng. Đã có đề xuất để chia tỷ lệ vector chuyển động được giải mã của PU hiện thời đối với hình ảnh cố định trước khi thực hiện bù chuyển động cho quá trình tạo số dư khi hệ số gia trọng không bằng 0. Như được đề xuất trong JCT3V-D0177, hình ảnh cố định được định nghĩa là hình ảnh tham chiếu thứ nhất của mỗi danh sách danh sách hình ảnh tham chiếu nếu nó được lấy từ cùng một khung hình. Khi vector chuyển động được giải mã không hướng đến hình ảnh cố định, bộ giải mã video 30 đầu tiên có thể chia tỷ lệ vector chuyển động giải mã và sau đó sử dụng vector chuyển động đã chia tỷ lệ để xác định CurrTRef và BaseTRef. Ảnh tham chiếu được sử dụng cho ARP như vậy được gọi là hình ảnh tham chiếu ARP đích.

Đối với các bộ lọc nội suy, như được mô tả trong JCT3V-C0049, bộ mã hóa video 20 và/hoặc bộ giải mã video 30 có thể áp dụng bộ lọc song tuyến tính trong quá trình nội suy khối tương ứng và khối dự báo của nó. Đối với khối dự báo của PU hiện thời trong các khung hình không phải gốc, bộ lọc 8/4-tap thông thường có thể được sử dụng. Theo một ví dụ khác, như được đề xuất trong JCT3V-D0177, bộ mã hóa video 20 và/hoặc bộ giải mã video 30 có thể luôn luôn sử dụng phép lọc song tuyến tính bất kể khối này là khung hình gốc hoặc khung hình không phải gốc khi ARP được áp dụng.

Theo một hoặc nhiều ví dụ của bản mô tả, bộ mã hóa video 20 và/hoặc bộ giải mã video 30 có thể được tạo cấu hình để xác định khung hình tham chiếu bằng cách sử dụng chỉ số thứ tự khung hình được phản hồi từ quy trình suy ra NBDV. Theo một ví dụ của ARP, khi hình ảnh tham chiếu của một PU trong một danh sách hình ảnh tham chiếu lấy từ khung hình khác khung hình hiện thời, ARP được bất hoạt đối với danh sách hình ảnh tham chiếu này.

Một số kỹ thuật khác để mã hóa chiều sâu liên hình ảnh được mô tả trong đơn Mỹ tạm số 61/840,400, nộp ngày 27/6/2013, và 61/847,942, nộp ngày 18/7/2013. Theo cách ví dụ này, khi mã hóa hình ảnh có chiều sâu, vector chênh lệch được biến đổi bởi giá trị chiều sâu ước lượng được từ các mẫu lân cận của khối hiện thời.

Theo các ví dụ khác đối với ARP, các ứng cử kết sắp khác có thể nhận được ví dụ, bằng cách truy cập khối tham chiếu của hình ảnh gốc được xác định bởi vector chênh lệch.

Dưới đây mô tả kỹ thuật để định vị khối cho dự báo chuyển động giữa các cảnh nhìn. Trong 3D-HEVC, khối 4x4 tham chiếu được xác định bằng cách sử dụng hai bước tổng quát. Bước thứ nhất là để xác định điểm ảnh trong khung hình tham chiếu bằng vector chuyển động chênh lệch. Bước thứ hai là tạo ra khối 4x4 tương ứng (có tập thông tin chuyển động duy nhất lần lượt tương ứng với RefPicList0 hoặc RefPicList1) và sử dụng thông tin chuyển động để tạo ra ứng cử kết sắp.

Điểm ảnh (xRef, yRef) trong hình ảnh tham chiếu được xác định như sau:

$$xRef = Clip3(0, PicWidthInSamples_L - 1, xP + ((nPSW - 1) >> 1) + ((mvDisp[0] + 2) >> 2))) \quad (H-124)$$

$$yRef = Clip3(0, PicHeightInSamples_L - 1, yP + ((nPSH - 1) >> 1) + ((mvDisp[1] + 2) >> 2))) \quad (H-125)$$

trong đó (xP, yP) là tọa độ của mẫu trên cùng bên trái của PU hiện thời, mvDisp là vector chênh lệch và nPSWxnPSH là kích thước của PU hiện thời và PicWidthInSamples_L và PicHeightInSamples_L định nghĩa độ phân giải của hình ảnh trong khung hình tham chiếu (giống như khung hình hiện thời).

Dưới đây mô tả dự báo chuyển động giữa các cảnh nhìn cấp PU con. Trong JCT3V-E0184, có đề xuất sử dụng phương pháp dự báo chuyển động giữa các cảnh nhìn cấp PU con cho IPMVC, tức là, ứng cử suy ra được từ khối tham chiếu trong khung hình tham chiếu. JCT3V-E0184, “3D-CE3.h related: Sub-PU level inter-view motion prediction,” bởi An et al., có sẵn, từ http://phenix.it-sudparis.eu/jct2/doc_end_user/current_document.php?id=1198.

Khái niệm cơ bản của phương pháp dự báo chuyển động giữa các cảnh nhìn được mô tả trên đây (ví dụ, liên quan đến quy trình suy ra ứng cử giữa các cảnh nhìn cho chế độ bỏ qua/kết sắp), trong đó chỉ có thông tin chuyển động của khối tham chiếu liên quan đến vị trí tâm được sử dụng cho PU hiện thời trong khung hình phụ thuộc. Tuy nhiên, PU hiện thời có thể tương ứng với vùng tham chiếu (có cùng kích thước với PU hiện thời được xác định bởi vector chênh lệch) trong khung hình tham chiếu và vùng tham chiếu có thể có lượng lớn thông tin chuyển động (tức là nhiều hơn một vector chuyển động).

Do vậy, phương pháp dự báo chuyển động giữa các cảnh nhìn cấp PU con (sub-PU level inter-view motion prediction-SPIVMP) được đề xuất. Fig.16 là sơ đồ

khái niệm minh họa dự báo chuyển động giữa các cảnh nhìn của đơn vị dự báo con (PU). Như được thể hiện trên Fig.16, PU hiện thời 170 trong khung hình hiện thời V1 có thể được tách thành nhiều PU con (ví dụ, bốn PU con). Vectơ chênh lệch đối với mỗi PU con có thể được sử dụng để định vị khối tham chiếu tương ứng trong khung hình tham chiếu V0. Bộ mã hóa video 20 và/hoặc bộ giải mã video 30 có thể được tạo cấu hình để sao chép (tức là, tái sử dụng) vectơ chuyển động liên quan đến mỗi trong số các khối được sử dụng cùng với các PU con tương ứng của PU hiện thời 170.

Theo một ví dụ, ứng cử kết sắp thời gian giữa các cảnh nhìn được suy ra được như sau. Trước hết, biểu thị kích thước PU con đã gán là $N \times N$. Trước hết, chia PU hiện thời thành nhiều PU con có kích thước nhỏ hơn. Biểu thị kích thước của PU hiện thời là $nPSW \times nPSH$ và kích thước của PU con là $nPSW_{sub} \times nPSH_{sub}$.

$$nPSW_{sub} = \min(N, nPSW)$$

$$nPSH_{sub} = \min(N, nPSH)$$

Thứ hai, thiết lập vectơ chuyển động mặc định $tmvLX$ là (0, 0) và chỉ số tham chiếu $refLX$ là -1 cho mỗi danh sách hình ảnh tham chiếu. Đối với mỗi PU con theo thứ tự quét mảnh, áp dụng các bước sau. Bổ sung DV vào vị trí ở giữa của PU con hiện thời để thu được vị trí mẫu tham chiếu ($xRefSub$, $yRefSub$) là:

$$xRefSub = \text{Clip3}(0, \text{PicWidthInSamplesL} - 1, xPSub + nPSW_{sub}/2 + ((mvDisp[0] + 2) \gg 2))$$

$$yRefSub = \text{Clip3}(0, \text{PicHeightInSamplesL} - 1, yPSub + nPSH_{sub}/2 + (((mvDisp[1] + 2) \gg 2)))$$

Một khối trong khung hình tham chiếu bao gồm ($xRefSub$, $yRefSub$) được sử dụng làm khối tham chiếu cho PU con hiện thời.

Đối với khối tham chiếu đã được xác định, nếu nó được mã hóa bằng cách sử dụng vectơ chuyển động thời gian, các bước sau được áp dụng. Nếu cả hai $refL0$ và $refL1$ bằng -1, và PU con hiện thời không phải là PU con đầu tiên theo thứ tự quét mảnh, thông tin chuyển động của khối tham chiếu được kế thừa bởi tất cả các PU trước đó. Các tham số chuyển động kết hợp có thể được sử dụng làm tham số ứng cử chuyển động cho PU con hiện thời. Phần tử cú pháp $tmvLX$ và $refLX$ được cập nhật vào thông tin chuyển động của PU con hiện thời. Theo cách khác (ví dụ, nếu khối tham chiếu được mã hóa trong hình ảnh), thông tin chuyển động của PU con hiện thời

được đặt là tmvLX và refLX . Các kích thước khác nhau của khối PU con có thể được áp dụng, ví dụ, 4×4 , 8×8 , và 16×16 . Kích thước của PU con có thể được báo hiệu trong VPS.

Dưới đây mô tả sự kế thừa vector chuyển động cấp PU con để mã hóa chiều sâu. Tương tự như đối với các đề xuất cho dự báo chuyển động giữa các cảnh nhìn cấp PU con từ một khung hình kết cấu khác, đơn Mỹ tạm thời số 61/858,089, nộp ngày 24/7/2013, các kỹ thuật được đề xuất áp dụng dự báo chuyển động cấp PU con từ một khung hình kết cấu đối với khung hình chiều sâu tương ứng. Tức là, PU hiện thời có thể phân chia thành một số PU con và mỗi PU con sử dụng thông tin chuyển động của khối kết cấu đồng vị trí để bù chuyển động. Trong trường hợp này, MVI cấp PU con được hỗ trợ và vector chênh lệch được sử dụng bởi dự báo chuyển động giữa các cảnh nhìn được coi là luôn bằng zero.

Khi thiết kế hiện thời đối với BVSP trong 3D-HEVC có các vấn đề sau. Khi AMP được sử dụng và kích thước PU hiện thời là 4×16 hoặc 16×4 , và PU được dự báo một chiều, BVSP đạt được bằng cách tạo ra một vector chênh lệch cho toàn bộ PU. Tức là mỗi khối con trong PU sử dụng cùng một vector chênh lệch để tổng hợp khối tham chiếu và bù chuyển động. Như vậy, đối với kích thước khối lớn hơn, BVSP có thể kém hiệu quả hơn do việc sử dụng cùng một vector chênh lệch để tổng hợp khối tham chiếu và bù chuyển động cho tất cả các khối con có thể kém tối ưu đối với một số khối con.

Một nhược điểm khác, khi PU hiện thời được dự báo hai chiều, BVSP có thể thực hiện với kích thước khối là 4×8 và 8×4 . Tuy nhiên, theo HEVC, việc bù chuyển động cho các khối có ít hơn 64 điểm ảnh (ví dụ, các khối 4×8 hoặc 8×4) là không được chấp nhận (mặc dù việc bù chuyển động 16×4 và 4×16 được cho phép).

Khi xem xét các nhược điểm này, sáng chế đề xuất kỹ thuật đề cập đến phương pháp dự báo tổng hợp hình ảnh trong 3D-HEVC, tập trung vào các kích thước bù chuyển động BVSP. Theo kỹ thuật của sáng chế, đối với BVSP, mỗi PU có thể được phân cắt thành các khối con và mỗi khối con có thể kết hợp với một vector chuyển động chênh lệch khác nhau và được bù chuyển động một cách riêng rẽ. Theo cách này, độ chính xác của BVSP có thể được làm tăng đối với các khối được chia bằng AMP, và do vậy, hiệu quả mã hóa có thể tăng. Theo kỹ thuật của sáng chế, kích thước của các khối con khả dụng để sử dụng cùng với BVSP có thể còn được xác định như sau.

Theo một ví dụ của sáng chế, khi kích thước PU hiện thời là 16×4 (hoặc 4×16), và PU hiện thời được dự báo một chiều bộ mã hóa video 20 và/hoặc bộ giải mã video 30 có thể được tạo cấu hình để áp dụng kỹ thuật BVSP và kỹ thuật bù chuyển động cho khối con 8×4 (hoặc 4×8) của PU hiện thời. Tức là, kích thước của vùng con BVSP là 8×4 (hoặc 4×8). Mỗi khối con có thể được gán một vector chuyển động chênh lệch được biến đổi từ khối có chiều sâu.

Fig.17 là hình vẽ khái niệm thể hiện kỹ thuật BVSP và kỹ thuật bù chuyển động theo sáng chế khi sử dụng AMP. Theo ví dụ của Fig.17, bộ mã hóa video 20 và/hoặc bộ giải mã video 30 phân chia không đối xứng PU hiện thời 250 thành khối 4×16 . Lưu ý rằng sự phân chia 4×16 chỉ là một ví dụ và rằng kỹ thuật theo sáng chế một tả có tham chiếu đến Fig.17 có thể được áp dụng cho các phân chia không đối xứng khác, kể cả psuwj phân chia 16×4 . Bộ mã hóa video 20 và/hoặc bộ giải mã video 30 có thể được tạo cấu hình để chia tiếp PU 250 thành các khối con 4×8 255 và 256.

Theo ví dụ của Fig.17, bộ mã hóa video 20 và/hoặc bộ giải mã video 30 được tạo cấu hình để dự báo một chiều PU 250 bằng cách sử dụng kỹ thuật BVSP. Về mặt này, bộ mã hóa video 20 và/hoặc bộ giải mã video 30 có thể được tạo cấu hình để suy ra vector chênh lệch cho PU 250, ví dụ, sử dụng kỹ thuật suy ra NBDV. Ví dụ, vector chênh lệch 261 có thể được suy ra từ khối lân cận 252. Bộ mã hóa video 20 và/hoặc bộ giải mã video 30 sau đó có thể được tạo cấu hình để tái sử dụng vector chênh lệch 261 để định vị khối có chiều sâu tương ứng 260 trong hình ảnh có chiều sâu tham chiếu. Theo kỹ thuật của sáng chế, không sử dụng toàn bộ khối có chiều sâu 260 để suy ra vector chuyển động chênh lệch đơn cho PU 255, bộ mã hóa video 20 và/hoặc bộ giải mã video 30 có thể được tạo cấu hình để suy ra vector chuyển động chênh lệch từ khối con 4×8 264 của khối có chiều sâu 260 cho khối con 255, và suy ra vector chuyển động chênh lệch từ khối con 4×8 262 của khối có chiều sâu 26 cho khối con 256.

Bộ mã hóa video 20 và/hoặc bộ giải mã video 30 sau đó có thể tổng hợp khối tham chiếu cho mỗi khối con 255 và 256 bằng cách sử dụng vector chuyển động chênh lệch suy ra được tương ứng để thực hiện bù chuyển động với hình ảnh tham chiếu tương ứng với hình ảnh tham chiếu giữa các cảnh nhìn có thứ tự khung hình là refVIdxLX. Bằng cách tạo ra các vector chuyển động chênh lệch riêng biệt cho mỗi khối con 255 và 256, các khung hình tham chiếu chính xác hơn có thể được tổng hợp và quá trình bù chuyển động tương ứng có thể đạt được độ lợi mã hóa tăng.

Theo ví dụ khác của sáng chế, khi kích thước PU hiện thời là 16×12 (hoặc 12×16), và PU hiện thời được dự báo một chiều, bộ mã hóa video 20 và/hoặc bộ giải mã video 30 có thể được tạo cấu hình để phân chia PU hiện thời thành các khối con 8×4 (hoặc 4×8) (còn được gọi là các vùng con BVSP) và suy ra vector chuyển động chênh lệch cho mỗi khối con sử dụng kỹ thuật BVSP.

Theo ví dụ khác, kích thước của các vùng con BVS có thể được gán là 16×12 hoặc 12×16 . Theo ví dụ khác nữa, mỗi khối con 16×12 (hoặc 12×16) được phân chia tiếp thành khối con 16×8 (hoặc 8×16) và hai khối con khối con 8×4 (hoặc 4×8) liền kề với PU 16×4 (4×16) trong cùng một CU. Theo ví dụ khác, khối con 16×8 (hoặc 8×16) có thể còn được tách tiếp thành hai vùng con 8×8 hoặc bốn vùng con 4×8 (hoặc 8×4) dựa trên, ví dụ, 4 góc của khối có chiều sâu tương ứng.

Theo ví dụ khác của sáng chế, khi cả chiều cao và chiều rộng của PU hiện thời lớn hơn hoặc bằng 8, và PU được dự báo hai chiều, bộ mã hóa video 20 và/hoặc bộ giải mã video 30 được tạo cấu hình để thiết lập kích thước của vùng con BVSP là 8×8 thay vì 4×8 hoặc 8×4 , như các đề xuất trước đó đối với 3D-HEVC. Theo ví dụ khác, thay vì sử dụng BVSP dự báo hai chiều cho các PU có kích thước bằng 12×16 hoặc 16×12 , BVSP dự báo một chiều có thể được áp dụng. Trong trường hợp này, kích thước bù chuyển động size có thể còn được thiết lập là 4×16 hoặc 16×4 . Theo ví dụ khác, khi kích thước PU hiện thời là 16×4 hoặc 4×16 , và PU hiện thời được dự báo hai chiều, kích thước của vùng con BVSP được thiết lập bằng kích thước PU.

Dự báo chuyển động PU con có thể có các nhược điểm sau. Theo ngữ cảnh này, dự báo chuyển động PU con có thể bao gồm kỹ thuật dự báo chuyển động PU con được đề xuất trong JCT3V-E0184, như nêu trên, cũng như dạng mở rộng của kỹ thuật dự báo chuyển động PU con cho MVI từ khung hình kết cấu đến khung hình chiều sâu.

Một nhược điểm là khi phân chia chuyển động không đối xứng (AMP) được thực hiện, kích thước PU hiện thời bằng, ví dụ, 4×16 , 16×4 , và kích thước khối PU con ở VPS bằng 8×8 , dựa trên các đề xuất trước đó đối với dự báo chuyển động cấp PU con, các PU này sẽ được phân chia thành hai khối con 4×8 hoặc 8×4 . Đối với mỗi khối con, thông tin chuyển động từ khối tham chiếu được kế thừa. Thông tin chuyển động có thể được xác định bởi vector chênh lệch trong khung hình kết cấu tham chiếu cho dự báo chuyển động giữa các cảnh nhìn có thể được tái sử dụng từ khối kết cấu

đồng vị trí trong khung hình kết cấu tương ứng để kế thừa vector chuyển động. Theo ví dụ này, dự báo hai chiều dựa trên 4×8 hoặc 8×4 có thể được kích hoạt, dự báo này không được cho phép bởi HEVC.

Một nhược điểm khác, khi AMP được thực hiện, kích thước PU hiện thời bằng, ví dụ, 12×16 , 16×12 , và kích thước khối PU con được báo hiệu (tức là, kích thước khối con) ở VPS bằng 8×8 , dựa trên các đề xuất trước đó đối với dự báo chuyển động cấp PU con, các PU này sẽ được phân chia thành hai khối con 8×8 và hai khối $4 \times 8/8 \times 4$. Tương tự trường hợp trên, các khối con $4 \times 8/8 \times 4$ có thể sử dụng dự báo hai chiều, mà dự báo này không được HEVC chấp nhận.

Kỹ thuật liên quan đến dự báo chuyển động giữa các cảnh nhìn và sự kế thừa vector chuyển động (đối với PU chiều sâu) được đề xuất trong sáng chế. Kỹ thuật theo sáng chế có thể được áp dụng trong ngữ cảnh khi chỉ số kết sắp chỉ ra kỹ thuật dự báo chuyển động giữa các cảnh nhìn hoặc MVI. Cụ thể là, kỹ thuật dự báo chuyển động giữa các cảnh nhìn và/hoặc MVI theo sáng chế bao gồm kỹ thuật để phân chia thêm các AMP PU thành các khối con và thu được thông tin chuyển động riêng rẽ cho mỗi khối con này. Theo cách này, độ chính xác của dự báo chuyển động giữa các cảnh nhìn và/hoặc MVI có thể được cải thiện đối với mỗi khối trong số các khối và do vậy, hiệu quả mã hóa có thể tăng.

Theo một ví dụ của sáng chế, khi PU hiện thời được mã hóa bằng cách sử dụng dự báo chuyển động giữa các cảnh nhìn và/hoặc MVI và kích thước của PU hiện thời bằng 4×16 hoặc 16×4 , bộ mã hóa video 20 và/hoặc bộ giải mã video 30 có thể được tạo cấu hình để phân chia PU thành hai khối con 4×8 hoặc 8×4 . Đối với mỗi khối con, bộ mã hóa video 20 và/hoặc bộ giải mã video 30 có thể được tạo cấu hình chỉ để thu thông tin chuyển động từ khối tham chiếu tương ứng với một danh sách hình ảnh tham chiếu cụ thể (ví dụ, RefPicList0). Thông tin chuyển động tương ứng với khối tham chiếu trong RefPicList0 được kế thừa cho các khối con 4×8 hoặc 8×4 . Trong trường hợp này, các khối con được dự báo một chiều từ một hình ảnh trong RefPicList0.

Fig.18 là sơ đồ khái niệm minh họa sự kế thừa vector chuyển động và kỹ thuật bù chuyển động cho các PU được phân chia không đối xứng thành các kích thước 4×16 và 16×4 . Ví dụ, đối với PU 4×16 , bộ mã hóa video 20 và/hoặc bộ giải mã video 30 được tạo cấu hình để chia tiếp PU 4×16 thành hai khối con 4×8 300 và 302. Thông tin chuyển động cho mỗi khối con 300 và 302 thu được từ khối tham chiếu trong hình

ảnh tham chiếu nằm trong danh sách hình ảnh tham chiếu cụ thể (ví dụ, RefPicList0). Sau đó thực hiện bù chuyển động cho mỗi khối con 300 và 302 so với khối tham chiếu trong RefPicList0. Tương tự, đối với PU 16x4, bộ mã hóa video 20 và/hoặc bộ giải mã video 30 được tạo cấu hình để chia tiếp PU 16x4 thành hai khối con 8x3 304 và 306. Thông tin chuyển động cho mỗi khối con 304 và 306 thu được từ khối tham chiếu trong hình ảnh tham chiếu nằm trong danh sách hình ảnh tham chiếu cụ thể (ví dụ, RefPicList0). Sau đó thực hiện bù chuyển động cho mỗi khối con 304 và 306 so với khối tham chiếu trong RefPicList0.

Theo ví dụ khác của sáng chế, khi kích thước của PU hiện thời là một trong số 16x12, 12x16, 4x16, hoặc 16x4, bộ mã hóa video 20 và bộ giải mã video 30 được tạo cấu hình để không sử dụng dự báo đôi đối với các khối con 8x4/4x8 khi kỹ thuật dự báo chuyển động giữa các cảnh nhìn cấp PU con và/hoặc MVI (đối với chiều sâu) được áp dụng. Tức là, khi kích thước của PU hiện thời là một trong số 16x12, 12x16, 4x16, hoặc 16x4, bộ mã hóa video 20 và bộ giải mã video 30 được tạo cấu hình để chỉ sử dụng dự báo đơn đối với các khối con 8x4/4x8 khi dự báo chuyển động giữa các cảnh nhìn cấp PU con và/hoặc MVI (đối với chiều sâu) được áp dụng.

Theo ví dụ khác của sáng chế, có đề xuất rằng khi dự báo chuyển động giữa các cảnh nhìn cấp PU con và/hoặc MVI được áp dụng và kích thước PU hiện thời bằng 4x16 hoặc 16x4, PU không được phân chia thành các PU con.

Theo ví dụ khác của sáng chế, được đề xuất rằng khi dự báo chuyển động giữa các cảnh nhìn cấp PU con và/hoặc MVI được áp dụng và kích thước PU hiện thời bằng 12x16 hoặc 16x12, PU không được phân chia thành ba khối PU con có kích thước bằng 4x16 hoặc 16x4. Đối với mỗi khối PU con, thông tin chuyển động của khối tham chiếu tương ứng được kế thừa.

Theo ví dụ khác của sáng chế, khi kích thước PU hiện thời bằng 12x16 hoặc 16x12, PU được phân chia thành hai khối PU con 8x8 và một khối 4x16 hoặc 16x4, trong đó các PU con 8x8 tạo thành nửa bên trái hoặc bên trên của CU chứa PU này. Theo khía cạnh khác của ví dụ này, các khối con 4x16 và 16x4 còn được phân chia tiếp thành hai khối PU con 4x8 hoặc 8x4. Đối với mỗi PU con 4x8 hoặc 8x4, chỉ có thông tin chuyển động của khối tham chiếu tương ứng với danh sách hình ảnh tham chiếu (RefPicList0) thu được và được tái sử dụng cho PU con 4x8 hoặc 8x4. Trong

trường hợp này, các PU con được dự báo một chiều từ một hình ảnh trong RefPicList0.

Theo một ví dụ khác của sáng chế, khi BVSP được áp dụng cho PU có kích thước bằng 12x16 hoặc 16x12, PU này được chia thành ba khối PU con có kích thước bằng nhau bằng 4x16 hoặc 16x4. Bộ mã hóa video 20 và/hoặc bộ giải mã video 30 sau đó có thể suy ra vector chuyển động chênh lệch duy nhất cho PU con của từ khối có chiều sâu tương ứng.

Fig.19 là sơ đồ khối thể hiện ví dụ về bộ mã hóa video 20 mà có thể thực hiện các kỹ thuật của sáng chế. Bộ mã hóa video 20 có thể thực hiện mã hóa trong hình ảnh và liên hình ảnh (bao gồm mã hóa giữa các cảnh nhìn) khối video trong các lát video, ví dụ, các lát của cả hình ảnh kết cấu và ánh xạ chiều sâu. Thông tin kết cấu thường bao gồm thông tin độ sáng (độ sáng hoặc cường độ) và sắc độ (màu, ví dụ, màu đỏ và màu xanh). Nói chung, bộ mã hóa video 20 có thể xác định chế độ mã hóa liên quan đến các phần độ sáng và tái sử dụng thông tin dự báo từ việc mã hóa thông tin độ sáng để mã hóa thông tin sắc độ (ví dụ, bằng cách tái sử dụng thông tin phân chia, chọn chế độ dự báo trong hình ảnh, vector chuyển động hoặc thông tin tương tự). Mã hóa trong hình ảnh dựa vào dự báo không gian để làm giảm hoặc loại bỏ phần thừa không gian trong video trong khung video xác định. Mã hóa liên hình ảnh dựa vào dự báo thời gian để làm giảm hoặc loại bỏ phần thừa thời gian trong video trong các khung hoặc hình ảnh liên kế của chuỗi video. Chế độ mã hóa trong hình ảnh (chế độ I) có thể chỉ chế độ nén video dựa trên không gian bất kỳ. Chế độ liên hình ảnh, như chế độ dự báo một chiều (chế độ P) hoặc dự báo hai chiều (chế độ B), có thể chỉ chế độ mã hóa dựa vào thời gian bất kỳ.

Như được thể hiện trên Fig.19, bộ mã hóa video 20 nhận khối video hiện thời (tức là, khối dữ liệu video, như khối tín hiệu chói, khối sắc độ hoặc khối có chiều sâu) trong khung video (ví dụ, hình ảnh kết cấu hoặc ánh xạ chiều sâu) được mã hóa Theo ví dụ của Fig.19, bộ mã hóa video 20 bao gồm bộ nhớ dữ liệu video 40, đơn vị chọn chế độ 41, bộ đệm hình ảnh được mã hóa (DPB) 64, bộ cộng 50, đơn vị xử lý biến đổi 52, đơn vị lượng tử hóa 54, đơn vị lọc trong vòng lặp 63 và đơn vị mã hóa entropi 56. Đơn vị chọn chế độ 41, lần lượt, bao gồm đơn vị bù chuyển động 44, đơn vị ước lượng chuyển động 42, đơn vị xử lý dự báo trong hình ảnh 46, và đơn vị phân chia 48. Đối với việc tái tạo khối video, bộ mã hóa video 20 còn bao gồm bộ lượng tử hóa

ngược 58, bộ xử lý biến đổi ngược 60, và bộ cộng 62. Đơn vị lọc trong vòng lặp 63 có thể bao gồm bộ lọc tách khối và bộ lọc SAO để lọc các đường biên của khối để loại bỏ các thành phần lạ dạng khối lớn ra khỏi video tái tạo. Bộ lọc khác (trong vòng lặp hoặc sau vòng lặp) cũng có thể được sử dụng ngoài Bộ lọc tách khối. Các bộ lọc này không được thể hiện một cách vắn tắt, nếu muốn, có thể lọc đầu ra của bộ cộng 50 (như bộ lọc trong vòng lặp).

Bộ nhớ dữ liệu video 40 có thể lưu trữ dữ liệu video sẽ được mã hóa bởi các thành phần của bộ mã hóa video 20. Dữ liệu video được lưu trữ trong bộ nhớ dữ liệu video 40 có thể được thu, ví dụ, từ nguồn video 18. DPB 64 là bộ nhớ đệm để lưu giữ dữ liệu video tham chiếu để sử dụng để mã hóa dữ liệu video bởi bộ mã hóa video 20 (ví dụ, chế độ mã hóa trong hình ảnh hoặc liên hình ảnh, còn được gọi là chế độ lặp giải mã trong hình ảnh hoặc liên hình ảnh). Bộ nhớ dữ liệu video 40 và bộ đệm DPB 64 có thể được tạo ra bởi thiết bị nhớ bất kỳ, như bộ nhớ truy cập ngẫu nhiên động (dynamic random access memory - DRAM), bao gồm DRAM đồng bộ (synchronous DRAM - SDRAM), RAM từ trở (magnetoresistive RAM - MRAM), RAM điện trở (resistive RAM - RRAM), hoặc các loại thiết bị nhớ khác. Bộ nhớ dữ liệu video 40 và DPB 64 có thể được cung cấp bởi cùng một thiết bị nhớ hoặc các thiết bị nhớ riêng. Trong nhiều ví dụ, bộ nhớ dữ liệu video 40 có thể nằm trên chip cùng các bộ phận khác của bộ mã hóa video 20, hoặc không nằm trên chip so với các bộ phận đó.

Trong quy trình mã hóa, bộ mã hóa video 20 nhận khung hoặc lát video cần được mã hóa. Khung hoặc lát video có thể được chia thành nhiều khối video. Bộ ước tính chuyển động 42 và bộ bù chuyển động 44 thực hiện mã hóa dự báo trong hình ảnh của khối video nhận được tương ứng với một hoặc nhiều khối trong một hoặc nhiều khung tham chiếu, bao gồm các khung tham chiếu giữa các cảnh nhìn, để tạo ra sự nén theo thời gian. Bộ xử lý dự báo trong hình ảnh 46 theo cách khác có thể thực hiện mã hóa dự báo liên hình ảnh khối video nhận được tương ứng với một hoặc nhiều khối lân cận trong cùng một khung hoặc lát vì khối này được mã hóa để tạo ra sự nén theo không gian. Bộ mã hóa video 20 có thể thực hiện nhiều lần chạy mã hóa, ví dụ, để chọn một chế độ mã hóa thích hợp cho mỗi khối dữ liệu video.

Hơn nữa, đơn vị phân chia 48 có thể phân chia các khối dữ liệu video thành khối con, dựa trên sự đánh giá của sơ đồ phân chia trước đó trong các lần chạy mã hóa trước đó. Ví dụ, bộ phân chia 48 có thể ban đầu phân chia khung hoặc lát thành các

LCU và phân chia mỗi LCU thành các CU con dựa vào phân tích tỉ lệ méo (ví dụ, tối ưu hóa tối độ méo). Đơn vị chọn chế độ 41 còn tạo ra cấu trúc dữ liệu cây tứ phân thể hiện sự phân chia của LCU thành các CU con. CU nút lá của cây tứ phân có thể bao gồm một hoặc nhiều PU và một hoặc nhiều TU.

Đơn vị chọn chế độ 41 có thể chọn một trong số các chế độ mã hóa, trong hình ảnh hoặc liên hình ảnh, ví dụ, dựa trên các kết quả lỗi và đưa khối được dự báo trong hình ảnh hoặc liên hình ảnh thu được đến bộ cộng 50 để tạo ra dữ liệu khối dư và đến bộ cộng 62 để tái tạo khối được mã hóa để dùng làm khung tham chiếu. Đơn vị chọn chế độ 41 cũng tạo ra các phần tử cú pháp, như vector chuyển động, bộ chỉ báo chế độ trong hình ảnh, thông tin phân chia và các thông tin cú pháp khác đến đơn vị mã hóa entropi 56.

Đơn vị ước lượng chuyển động 42 và đơn vị bù chuyển động 44 có thể được tích hợp hoàn toàn, nhưng vẫn được minh họa riêng nhằm mục đích dễ hiểu. Việc ước lượng chuyển động, được thực hiện bởi đơn vị ước lượng chuyển động 42, là quá trình tạo ra các vector chuyển động, các vector này ước lượng chuyển động cho các khối video. Vector chuyển động, ví dụ, có thể thể hiện sự chuyển vị của PU của khối video trong khung hoặc hình ảnh video hiện thời so với khối dự báo trong khung tham chiếu (hoặc đơn vị mã hóa khác) so với khối hiện thời được mã hóa trong khung hiện thời (hoặc đơn vị mã hóa khác).

Khối dự báo có thể là khối mà gần khớp với khối cần mã hóa xét về độ lệch điểm ảnh, giá trị này có thể được xác định bằng tổng hiệu số tuyệt đối (sum of absolute difference - SAD), tổng độ lệch bình phương (SSD), hoặc các ma trận độ lệch khác. Theo một số ví dụ, bộ mã hóa video 20 có thể tính toán các giá trị cho các vị trí điểm ảnh dưới số nguyên của các hình ảnh tham chiếu được lưu giữ trong DPB 64. Ví dụ, bộ mã hóa video 20 có thể nội suy các giá trị của các vị trí điểm ảnh một phần tư, vị trí điểm ảnh một phần tám hoặc vị trí điểm ảnh phân số khác của hình ảnh tham chiếu. Do vậy, đơn vị ước lượng chuyển động 42 có thể thực hiện việc tìm kiếm chuyển động đối với các vị trí điểm ảnh đầy đủ và các vị trí điểm ảnh phân số và xuất ra vector chuyển động có độ chính xác ở mức điểm ảnh phân số.

Đơn vị ước lượng chuyển động 42 có thể tính toán vector chuyển động cho PU của khối video trong lát được mã hóa trong hình ảnh bằng cách so sánh vị trí của PU này với vị trí của khối dự báo của hình ảnh tham chiếu. Hình ảnh tham chiếu có thể

được chọn từ danh sách hình ảnh tham chiếu thứ nhất (List 0) hoặc danh sách hình ảnh tham chiếu thứ hai (List 1), mỗi danh sách này nhận dạng một hoặc nhiều hình ảnh tham chiếu được lưu giữ trong DPB 64. Các danh sách hình ảnh tham chiếu có thể được tạo ra bằng cách sử dụng kỹ thuật theo sáng chế. Đơn vị ước lượng chuyển động 42 gửi vector chuyển động tính toán được đến đơn vị mã hóa entropi 56 và đơn vị bù chuyển động 44.

Việc bù chuyển động, được thực hiện bởi đơn vị bù chuyển động 44, có thể bao gồm việc tìm nạp hoặc tạo ra khối dự báo dựa trên vector chuyển động được xác định bởi đơn vị ước lượng chuyển động 42. Một lần nữa, đơn vị bù chuyển động 42 và đơn vị bù chuyển động 44 có thể được tích hợp về mặt chức năng, theo một số ví dụ. Khi nhận vector chuyển động cho PU của khối video hiện thời, đơn vị bù chuyển động 44 có thể định vị khối dự báo mà vector chuyển động hướng đến khối này trong một trong số các hình ảnh tham chiếu. Bộ cộng 50 tạo ra khối video dư bằng cách lấy giá trị điểm ảnh của khối dự báo trừ đi giá trị điểm ảnh của khối video hiện thời được mã hóa, tạo ra giá trị độ lệch điểm ảnh, như nêu trên. Nói chung, đơn vị ước lượng chuyển động 42 thực hiện sự ước lượng chuyển động đối với các thành phần độ sáng và đơn vị bù chuyển động unit 44 sử dụng các vector chuyển động được tính toán dựa trên các thành phần độ sáng cho cả thành phần độ màu và độ sáng. Theo cách này, đơn vị bù chuyển động 44 có thể tái sử dụng thông tin chuyển động được xác định cho các thành phần độ sáng để mã hóa thành phần độ màu để không cần thực hiện sự tìm kiếm chuyển động đối với các thành phần độ màu. Đơn vị chọn chế độ 41 cũng có thể tạo ra phần tử cú pháp liên quan đến các khối video và lát video để bộ giải mã video 30 sử dụng trong việc giải mã khối video của lát video.

Bộ xử lý dự báo liên hình ảnh 46 có thể dự báo liên hình ảnh khối hiện thời, thay cho dự báo liên hình ảnh được thực hiện bởi đơn vị ước lượng chuyển động 42 và bộ bù chuyển động 44, như nêu trên. Cụ thể là, đơn vị xử lý dự báo trong hình ảnh 46 có thể xác định chế độ dự báo trong hình ảnh để sử dụng để mã hóa khối hiện thời. Theo một số ví dụ, đơn vị xử lý dự báo trong hình ảnh 46 có thể mã hóa khối hiện thời bằng cách sử dụng các chế độ dự báo trong hình ảnh khác nhau, ví dụ, trong các lần chạy mã hóa riêng biệt và đơn vị xử lý dự báo trong hình ảnh 46 (hoặc đơn vị chọn chế độ 41, theo một số ví dụ) có thể chọn chế độ dự báo trong hình ảnh thích hợp để sử dụng từ các chế độ đã thử nghiệm.

Ví dụ, đơn vị xử lý dự báo trong hình ảnh 46 có thể tính toán giá trị tỉ lệ méo bằng cách sử dụng phân tích tỉ lệ méo đối với các chế độ dự báo trong hình ảnh khác nhau và chọn chế độ dự báo trong hình ảnh có đặc điểm tỉ lệ méo tốt nhất trong số các chế độ đã thử nghiệm. Phân tích tỉ lệ méo thường xác định mức độ méo (hoặc lỗi) giữa khối mã hóa và khối gốc, chưa được mã hóa mà khối này được mã hóa để tạo ra khối mã hóa, cũng như tốc độ (tức là, số lượng bit) được sử dụng để tạo ra khối mã hóa. Đơn vị xử lý dự báo trong hình ảnh 46 có thể tính toán tỷ lệ giữa các méo và mức độ đối với các khối mã hóa khác nhau để xác định chế độ dự báo trong hình ảnh nào có giá trị tỉ lệ méo tốt nhất cho khối này.

Sau khi chọn chế độ dự báo trong hình ảnh cho một khối, đơn vị xử lý dự báo trong hình ảnh 46 có thể cung cấp thông tin chỉ báo về chế độ dự báo trong hình ảnh đã chọn cho khối đến đơn vị mã hóa entropi 56. Đơn vị mã hóa entropi 56 có thể mã hóa thông tin chỉ báo chế độ dự báo trong hình ảnh được chọn. Bộ mã hóa video 20 có thể bao gồm trên dữ liệu cấu hiện dòng bit được truyền, bao gồm nhiều bảng chỉ số chế độ dự báo trong hình ảnh và nhiều bảng chỉ số chế độ dự báo trong hình ảnh được hiệu chỉnh (còn được gọi là bảng bản đồ từ mã), các định nghĩa về ngữ cảnh mã hóa cho các khối khác nhau và các chỉ báo về chế độ dự báo trong hình ảnh có khả năng nhất, bảng chỉ số chế độ dự báo trong hình ảnh và bảng chỉ số chế độ dự báo trong hình ảnh đã hiệu chỉnh để sử dụng cho từng ngữ cảnh.

Bộ mã hóa video 20 tạo ra khối video dư bằng cách trừ đi dữ liệu dự báo của đơn vị chọn chế độ 41 từ khối video gốc được mã hóa. Bộ cộng 50 thể hiện thành phần hoặc các thành phần thực hiện phép trừ này. Đơn vị xử lý biến đổi 52 áp dụng phép biến đổi, như phép biến đổi cosin rời rạc (DCT) hoặc phép biến đổi tương tự về mặt khái niệm, cho khối dư, tạo ra khối video bao gồm các giá trị hệ số biến đổi dư. Đơn vị xử lý biến đổi 52 có thể thực hiện phép biến đổi khác tương tự với mặt khái niệm với DCT. Phép biến đổi sóng con, phép biến đổi số nguyên, phép biến đổi dải tần phụ hoặc các phép biến đổi khác cũng có thể được sử dụng. Trong trường hợp bất kỳ, đơn vị xử lý biến đổi 52 áp dụng phép biến đổi đối với khối dư, tạo ra khối có các hệ số biến đổi dư.

Phép biến đổi có thể biến đổi thông tin dư từ miền giá trị điểm ảnh thành miền biến đổi, như miền tần số. Đơn vị xử lý biến đổi 52 có thể gửi các hệ số biến đổi thu được đến đơn vị lượng tử hóa 54. Đơn vị lượng tử hóa 54 lượng tử hóa các hệ số biến

đổi để làm giảm tiếp tốc độ bit. Quy lượng tử hóa có thể làm giảm độ sâu bit liên quan đến một số hoặc tất cả các hệ số. Mức độ lượng tử hóa có thể được thay đổi bằng cách điều chỉnh tham số lượng tử hóa. Theo một số ví dụ, đơn vị lượng tử hóa 54 có thể sau đó thực hiện quét ma trận bao gồm các hệ số biến đổi đã lượng tử hóa. Theo cách khác, đơn vị mã hóa entropi 56 có thể thực hiện việc quét này.

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

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

Bộ mã hóa video 20 có thể mã hóa ánh xạ chiều sâu theo cách thức về cơ bản giống với kỹ thuật mã hóa để mã hóa các thành phần độ sáng, mặc dù không cần có các thành phần độ màu tương ứng. Ví dụ, đơn vị xử lý dự báo trong hình ảnh 46 có thể dự báo trong hình ảnh các khối ánh xạ chiều sâu, còn đơn vị ước lượng chuyển động

42 và đơn vị bù chuyển động 44 có thể dự báo liên hình ảnh các khối ánh xạ chiều sâu. Tuy nhiên, như nêu trên, trong việc dự báo liên hình ảnh các ánh xạ chiều sâu, đơn vị bù chuyển động 44 có thể thay đổi (tức là, điều chỉnh) các giá trị ánh xạ chiều sâu tham chiếu dựa trên độ chênh lệch về khoảng chiều sâu và giá trị độ chính xác cho khoảng chiều sâu này. Ví dụ, nếu giá trị chiều sâu lớn nhất trong ánh xạ chiều sâu hiện thời và ánh xạ chiều sâu tham chiếu tương ứng với cùng một chiều sâu thể giới thực, bộ mã hóa video 20 có thể điều chỉnh giá trị chiều sâu lớn nhất của ánh xạ chiều sâu tham chiếu bằng với giá trị chiều sâu lớn nhất trong ánh xạ chiều sâu hiện thời, cho mục đích dự báo. Ngoài ra hoặc theo cách khác, bộ mã hóa video 20 có thể sử dụng giá trị khoảng chiều sâu được cập nhật và giá trị độ chính xác để tạo ra hình ảnh tổng hợp khung hình để dự báo tổng hợp khung hình, ví dụ, bằng cách sử dụng kỹ thuật tương tự với kỹ thuật dự báo giữa các cảnh nhìn.

Như sẽ được thảo luận một cách chi tiết hơn dưới đây cùng với sự tham chiếu đến các Fig.21-23, bộ mã hóa video 20 có thể được tạo cấu hình để áp dụng kỹ thuật theo sáng chế được mô tả trên đây. Cụ thể, bộ mã hóa video 20 có thể được tạo cấu hình để phân chia PU thành các khối con khi các PU này được phân chia theo chế độ phân chia không đối xứng. Bộ mã hóa video 20 có thể sau đó được tạo cấu hình để kế thừa và/hoặc suy ra vector chuyển động hoặc vector chuyển động chênh lệch cho mỗi khối con.

Fig.20 là sơ đồ khối thể hiện ví dụ về bộ mã hóa video 30 mà có thể thực hiện các kỹ thuật của sáng chế. Theo ví dụ của Fig.20, bộ giải mã video 30 bao gồm bộ nhớ dữ liệu video 79, đơn vị giải mã entropi 70, đơn vị bù chuyển động 72, đơn vị xử lý dự báo trong hình ảnh 74, đơn vị lượng tử hóa ngược 76, đơn vị xử lý biến đổi ngược 78, bộ đệm hình ảnh được mã hóa (DPB) 82, đơn vị lọc trong vòng lặp 83 và bộ cộng 80. Bộ giải mã video 30 có thể, trong một số ví dụ, thực hiện một đường chạy giải mã thường ngược với đường chạy mã hóa được mô tả với bộ mã hóa video 20 (xem Fig.19). Đơn vị bù chuyển động 72 có thể tạo ra dữ liệu dự báo dựa trên vector chuyển động nhận từ đơn vị giải mã entropi 70, còn đơn vị xử lý dự báo trong hình ảnh 74 có thể tạo ra dữ liệu dự báo dựa trên bộ chỉ báo chế độ dự báo trong hình ảnh được nhận từ đơn vị giải mã entropi 70.

Bộ nhớ dữ liệu video 79 có thể lưu trữ dữ liệu video, như dòng bit video được mã hóa, cần được giải mã bởi các thành phần của bộ giải mã video 30. Dữ liệu video

được lưu trữ trong bộ nhớ dữ liệu video 79 có thể được thu từ nguồn video cục bộ, như máy quay phim, thông qua truyền thông mạng có dây hoặc không dây, dữ liệu video, hoặc bằng cách truy cập vật ghi dữ liệu vật lý. Bộ nhớ dữ liệu video 79 có thể tạo bộ nhớ đệm hình ảnh được mã hóa (coded picture buffer - CPB) mà lưu trữ dữ liệu video được mã hóa từ dòng bit video được mã hóa. DPB 82 là một ví dụ về bộ nhớ đệm để lưu giữ dữ liệu video tham chiếu để sử dụng để mã hóa dữ liệu video bởi bộ mã hóa video 30 (ví dụ, chế độ mã hóa trong hình ảnh hoặc liên hình ảnh, còn được gọi là chế độ lập giải mã trong hình ảnh hoặc liên hình ảnh). Bộ nhớ dữ liệu video 79 và bộ đệm DPB 82 có thể được tạo ra bởi thiết bị nhớ bất kỳ, như bộ nhớ truy cập ngẫu nhiên động (dynamic random access memory - DRAM), bao gồm DRAM đồng bộ (synchronous DRAM - SDRAM), RAM từ trở (magnetoresistive RAM - MRAM), RAM điện trở (resistive RAM - RRAM), hoặc các loại thiết bị nhớ khác. Bộ nhớ dữ liệu video 79 và DPB 82 có thể được cung cấp bởi cùng một thiết bị nhớ hoặc các thiết bị nhớ riêng. Trong các ví dụ, bộ nhớ dữ liệu video 79 có thể nằm trên chip cùng các bộ phận khác của bộ giải mã video 30, hoặc không nằm trên chip so với các bộ phận đó.

Trong quá trình giải mã, bộ giải mã video 30 nhận dòng bit video mã hóa biểu diễn khối video của lát video được mã hóa và các phần tử cú pháp liên quan từ bộ mã hóa video 20. Đơn vị giải mã entropi 70 của bộ giải mã video 30 giải mã entropi dòng bit này để tạo ra các hệ số lượng tử hóa, vector chuyển động hoặc các bộ chỉ báo chế độ dự báo trong hình ảnh hoặc các phần tử cú pháp khác. Đơn vị giải mã entropi 70 chuyển tiếp vector chuyển động và các phần tử cú pháp khác đến đơn vị bù chuyển động 72. Bộ giải mã video 30 có thể nhận phần tử cú pháp ở cấp lát video và/hoặc cấp khối video.

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

trong một trong các danh sách hình ảnh tham chiếu. Bộ giải mã video 30 có thể tạo ra danh sách khung tham chiếu, List 0 và List 1, bằng cách sử dụng kỹ thuật theo sáng chế dựa trên các hình ảnh tham chiếu được lưu giữ trong bộ đệm hình ảnh được mã hóa 82. Đơn vị bù chuyển động 72 xác định thông tin dự báo cho khối video của lát video hiện thời bằng cách phân tích các vector chuyển động và phân tử cú pháp và sử dụng thông tin dự báo để tạo ra khối dự báo cho khối video hiện thời được giải mã. Ví dụ, đơn vị bù chuyển động 72 sử dụng các phần tử cú pháp nhận được để xác định chế độ dự báo (ví dụ, dự báo trong hình ảnh hoặc liên hình ảnh) được sử dụng để mã hóa các khối video, loại lát dự báo liên hình ảnh (ví dụ, lát B, lát P hoặc lát GPB), thông tin cấu trúc cho một hoặc nhiều danh sách hình ảnh tham chiếu cho lát này, các vector chuyển động cho mỗi khối video được mã hóa liên hình ảnh của lát này, trạng thái dự báo liên hình ảnh đối với mỗi khối video được mã hóa liên hình ảnh của lát này và thông tin khác để giải mã các khối video trong lát video hiện thời.

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

Bộ lượng tử hóa ngược 76 lượng tử hóa ngược, tức là, giải lượng tử hóa, các hệ số biến đổi được lượng tử hóa có trong dòng bit và được giải mã bởi bộ giải mã entropi 70. Quy trình lượng tử hóa ngược có thể bao gồm việc sử dụng tham số lượng tử hóa QP_Y được tính bởi bộ giải mã video 30 cho mỗi khối video trong lát video để xác định mức lượng tử hóa và, tương tự, mức lượng tử hóa ngược cần được áp dụng.

Đơn vị xử lý biến đổi ngược 78 áp dụng phép biến đổi ngược, ví dụ, phép biến đổi DCT ngược, phép biến đổi số nguyên ngược hoặc quá trình biến đổi ngược tương tự về mặt khái niệm, đối với các hệ số biến đổi để tạo ra khối dư trong miền điểm ảnh.

Sau khi đơn vị bù chuyển động 72 tạo ra khối dự báo cho khối video hiện thời dựa vào vector chuyển động và các phần tử cú pháp khác, bộ giải mã video 30 tạo ra khối video giải mã bằng cách cộng khối dư từ đơn vị xử lý biến đổi ngược 78 với các khối dự báo tương ứng được tạo ra bởi đơn vị bù chuyển động 72. Bộ cộng 90 thể hiện

thành phần hoặc các thành phần thực hiện phép cộng này. Đơn vị lọc trong vòng lặp 63 có thể bao gồm bộ lọc tách khối và bộ lọc SAO để lọc các biên khối để loại bỏ các thành phần lạ dạng khối lớn ra khỏi video tái tạo. Bộ lọc khác (trong vòng lặp hoặc sau vòng lặp) cũng có thể được sử dụng ngoài bộ lọc tách khối. Các bộ lọc này không được thể hiện một cách vắn tắt, nếu muốn, có thể lọc đầu ra của bộ cộng 80 (như bộ lọc trong vòng lặp). Các khối video đã giải mã trong một khung hoặc hình ảnh cho trước sau đó được lưu giữ trong bộ đệm hình ảnh được mã hóa 82, lưu giữ các hình ảnh tham chiếu được sử dụng để bù chuyển động tiếp theo. Bộ đệm hình ảnh được mã hóa 82 cũng lưu giữ video đã giải mã cho lần thể hiện sau đó trên thiết bị hiển thị, như thiết bị hiển thị 32 trên Fig.1.

Như sẽ được thảo luận một cách chi tiết hơn dưới đây cùng với sự tham chiếu đến các Fig.24-26, bộ giải mã video 30 có thể được tạo cấu hình để áp dụng kỹ thuật theo sáng chế được mô tả trên đây. Cụ thể, bộ giải mã video 30 có thể được tạo cấu hình để phân chia PU thành các khối con khi các PU này được phân chia theo chế độ phân chia không đối xứng. Bộ giải mã video 30 có thể sau đó được tạo cấu hình để kế thừa và/hoặc suy ra vectơ chuyển động hoặc vectơ chuyển động chênh lệch cho mỗi khối con.

Fig.21 là lưu đồ minh họa phương pháp mã hóa ví dụ theo sáng chế. Kỹ thuật trên Fig.21 có thể được thực hiện bởi một hoặc nhiều đơn vị cấu trúc của bộ mã hóa video 20, như bởi đơn vị chọn chế độ 41, đơn vị phân chia 48 và/hoặc đơn vị bù chuyển động 44.

Theo một ví dụ của sáng chế, bộ mã hóa video 20 (ví dụ, sử dụng đơn vị chọn chế độ 41 và đơn vị phân chia 48) có thể được tạo cấu hình để tạo ra khối dữ liệu video sử dụng AMP, trong đó khối dữ liệu video được dự báo một chiều sử dụng BVSP, và có kích thước 16x12, 12x16, 16x4 hoặc 4x16 (2100). Theo một ví dụ của sáng chế, khối dữ liệu video là đơn vị dự báo.

Bộ mã hóa video 20, sử dụng đơn vị phân chia 48, còn có thể được tạo cấu hình để phân chia khối dữ liệu video thành các khối con, mỗi khối con có kích thước 8x4 hoặc 4x8 (2110), và tạo ra (ví dụ, sử dụng đơn vị bù chuyển động 44) vectơ chuyển động chênh lệch tương ứng cho mỗi khối con từ khối có chiều sâu tương ứng trong hình ảnh có chiều sâu tương ứng với hình ảnh tham chiếu (2120). Bộ mã hóa video 20, sử dụng đơn vị phân chia 48, còn có thể được tạo cấu hình để tổng hợp khối tham

chiều tương ứng cho mỗi khối con sử dụng vector chuyển động chênh lệch được tạo ra tương ứng (2130), và mã hóa (ví dụ, sử dụng đơn vị bù chuyển động 44) khối dữ liệu video bằng cách thực hiện bù chuyển động trên mỗi khối con sử dụng các khối tham chiếu tương ứng tổng hợp được (2140).

Theo ví dụ khác của sáng chế, bộ mã hóa video 20 còn có thể được tạo cấu hình để tạo ra một hoặc nhiều phần tử cú pháp chỉ báo rằng đơn vị dự báo được mã hóa bằng cách sử dụng AMP, và chỉ báo rằng đơn vị dự báo được dự báo một chiều bằng cách sử dụng BVSP, và tạo ra chỉ số ứng viên hợp nhất trở đến ứng cử BVSP.

Theo ví dụ khác của sáng chế, bộ mã hóa video 20 (ví dụ, sử dụng đơn vị bù chuyển động 44) có thể được tạo cấu hình để suy ra vector chuyển động chênh lệch tương ứng cho mỗi khối con bằng cách suy ra vector chênh lệch cho khối dữ liệu video, định vị khối có chiều sâu tương ứng cho mỗi khối con bằng cách sử dụng vector chênh lệch được tạo ra và biến đổi một giá trị chiều sâu đã chọn của khối có chiều sâu tương ứng cho mỗi khối con thành vector chuyển động chênh lệch tương ứng.

Fig.22 là lưu đồ minh họa phương pháp mã hóa ví dụ khác theo sáng chế. Kỹ thuật trên Fig.22 có thể được thực hiện bởi một hoặc nhiều đơn vị cấu trúc của bộ mã hóa video 20, bao gồm đơn vị chọn chế độ 41, đơn vị phân chia 48 và/hoặc đơn vị bù chuyển động 44.

Theo một ví dụ của sáng chế, bộ mã hóa video 20 (ví dụ, đơn vị chọn chế độ 41 và đơn vị phân chia 48) có thể được tạo cấu hình để tạo ra khối dữ liệu video thứ hai bằng cách sử dụng AMP, trong đó khối dữ liệu video thứ hai này được mã hóa bằng cách sử dụng ít nhất một dự báo chuyển động giữa các cảnh nhìn hoặc MVI, và có kích thước 16x12, 12x16, 16x4 hoặc 4x16 (2200). Bộ mã hóa video 20, sử dụng đơn vị phân chia 48, còn có thể được tạo cấu hình để phân chia khối dữ liệu video thứ hai thành các khối con, mỗi khối con có kích thước 8x4 hoặc 4x8 (2210), và tạo ra (ví dụ, sử dụng đơn vị bù chuyển động 44) thông tin chuyển động cho mỗi khối con từ một khối tham chiếu tương ứng (2220). Bộ mã hóa video 20 sau đó có thể mã hóa khối dữ liệu video thứ hai bằng cách thực hiện bù chuyển động trên mỗi khối con bằng cách sử dụng thông tin chuyển động suy ra được và một danh sách hình ảnh tham chiếu (2230).

Theo ví dụ khác của sáng chế, bộ mã hóa video 20 (ví dụ, sử dụng đơn vị bù chuyển động 44) có thể được tạo cấu hình để thực hiện bù chuyển động bằng cách thực hiện bù chuyển động một chiều đối với hình ảnh trong một danh sách danh sách hình ảnh tham chiếu.

Fig.23 là lưu đồ minh họa phương pháp mã hóa ví dụ khác theo sáng chế. Kỹ thuật trên Fig.23 có thể được thực hiện bởi một hoặc nhiều đơn vị cấu trúc của bộ mã hóa video 20, như bởi đơn vị chọn chế độ 41, đơn vị phân chia 48 và/hoặc đơn vị bù chuyển động 44.

Theo một ví dụ của sáng chế, bộ mã hóa video 20 có thể được tạo cấu hình để tạo ra 20 (ví dụ, sử dụng đơn vị chọn chế độ 41 và đơn vị phân chia 48) khối dữ liệu video thứ hai bằng cách sử dụng AMP, trong đó khối dữ liệu video thứ hai này được mã hóa bằng cách sử dụng ít nhất một trong dự báo chuyển động giữa các cảnh nhìn hoặc kế thừa vectơ chuyển động và có kích thước 16x12 hoặc 12x16 (2300), phân chia (ví dụ, sử dụng đơn vị phân chia 48) khối dữ liệu video thứ hai này thành nhiều khối con (2310), và mã hóa (ví dụ, sử dụng đơn vị bù chuyển động 44) từng khối con này bằng dự báo một chiều (2320).

Fig 24 là lưu đồ minh họa phương pháp giải mã ví dụ theo sáng chế. Kỹ thuật trên Fig.24 có thể được thực hiện bởi một hoặc nhiều đơn vị cấu trúc của bộ giải mã video 30, như bởi đơn vị bù chuyển động 72.

Theo một ví dụ của sáng chế, bộ giải mã video 30 có thể được tạo cấu hình để nhận dữ liệu dự tương ứng với khối dữ liệu video, trong đó khối dữ liệu video được mã hóa bằng cách sử dụng AMP, được mã hóa một chiều sử dụng BVSP, và có kích thước 16x12, 12x16, 16x4 hoặc 4x16 (2400). Theo một ví dụ của sáng chế, khối dữ liệu video là đơn vị dự báo. Bộ giải mã video 30, còn có thể được tạo cấu hình để phân chia khối dữ liệu video thành các khối con, mỗi khối con có kích thước 8x4 hoặc 4x8 (2410), và suy ra vectơ chuyển động chênh lệch tương ứng cho mỗi khối con từ khối có chiều sâu tương ứng trong hình ảnh có chiều sâu tương ứng với hình ảnh tham chiếu (2420).

Bộ giải mã video 30, còn có thể được tạo cấu hình để tổng hợp khối tham chiếu tương ứng cho mỗi khối con sử dụng vectơ chuyển động chênh lệch tương ứng tạo ra tương ứng (2430), và giải mã khối dữ liệu video bằng cách thực hiện bù chuyển động

trên mỗi khối con sử dụng dữ liệu dư và khối tham chiếu tương ứng tổng hợp được (2440).

Theo ví dụ khác của sáng chế, bộ giải mã video 30 còn có thể được tạo cấu hình để nhận một hoặc nhiều phần tử cú pháp chỉ báo rằng đơn vị dự báo được mã hóa bằng cách sử dụng sự phân chia chuyển động không đối xứng, và chỉ báo rằng đơn vị dự báo được dự báo một chiều bằng cách sử dụng dự báo tổng hợp cảnh nhìn ngược, và nhận chỉ số ứng viên hợp nhất trở đến ứng cử BVSP.

Theo ví dụ khác của sáng chế, bộ giải mã video 30 có thể được tạo cấu hình để suy ra vector chuyển động chênh lệch tương ứng cho mỗi khối con bằng cách suy ra vector chênh lệch cho khối dữ liệu video, định vị khối có chiều sâu tương ứng cho mỗi khối con bằng cách sử dụng vector chênh lệch được tạo ra và biến đổi một giá trị chiều sâu đã chọn của khối có chiều sâu tương ứng cho mỗi khối con thành vector chuyển động chênh lệch tương ứng.

Fig.25 là lưu đồ minh họa phương pháp giải mã ví dụ theo sáng chế. Kỹ thuật trên Fig.23 có thể được thực hiện bởi một hoặc nhiều đơn vị cấu trúc của bộ giải mã video 30, như bởi đơn vị bù chuyển động 72.

Theo một ví dụ của sáng chế, bộ giải mã video 30 có thể được tạo cấu hình để nhận dữ liệu dư tương ứng với khối dữ liệu video thứ hai, trong đó khối dữ liệu video thứ hai được mã hóa sử dụng ít nhất một trong dự báo chuyển động giữa các cảnh nhìn hoặc MVI, và có kích thước 16x4 hoặc 4x16 (2500), phân chia khối dữ liệu video thứ hai thành các khối con, mỗi khối con có kích thước 8x4 hoặc 4x8 (2510), suy ra thông tin chuyển động cho mỗi khối con từ một khối tham chiếu tương ứng (2520), và giải mã khối dữ liệu video thứ hai bằng cách thực hiện bù chuyển động trên mỗi khối con nhờ sử dụng dữ liệu dư, thông tin chuyển động suy ra được và một danh sách hình ảnh tham chiếu.

Theo ví dụ khác của sáng chế, bộ giải mã video 30 còn có thể được tạo cấu hình để thực hiện bù chuyển động bằng cách thực hiện bù chuyển động một chiều so với hình ảnh trong một danh sách hình ảnh tham chiếu.

Fig.26 là lưu đồ minh họa phương pháp giải mã ví dụ theo sáng chế. Kỹ thuật trên Fig.23 có thể được thực hiện bởi một hoặc nhiều đơn vị cấu trúc của bộ giải mã video 30, bao gồm đơn vị bù chuyển động 72.

Theo một ví dụ của sáng chế, bộ giải mã video 30 còn có thể được tạo cấu hình để nhận dữ liệu dự tương ứng với khối dữ liệu video thứ hai, trong đó khối dữ liệu video thứ hai được mã hóa sử dụng ít nhất một trong dự báo chuyển động giữa các cảnh nhìn hoặc MVI, và có kích thước 6×12 hoặc 12×16 (2600), phân chia khối dữ liệu video thứ hai thành nhiều khối con (2610), và giải mã mỗi khối con trong nhiều khối con bằng dự báo một chiều.

Như được giải thích trên đây, kỹ thuật theo sáng chế bao gồm kỹ thuật mã hóa và giải mã khi áp dụng AMP, BVSP, dự báo chuyển động giữa các cảnh nhìn, và/hoặc MVI trên các khối dữ liệu video. Cụ thể là, kỹ thuật theo sáng chế dùng để mã hóa một cách chính xác hơn bằng áp dụng kỹ thuật mã hóa cho khối con của PU được phân chia bởi AMP. Ví dụ, việc tạo ra các vector chuyển động chênh lệch cho các khối con của PU được phân chia bằng AMP khi PU này được mã hóa bằng cách sử dụng BVSP có thể làm tăng độ chính xác của việc tổng hợp và dự báo chuyển động hình ảnh, do đó làm tăng hiệu quả mã hóa. Một ví dụ khác, việc tạo ra các vector chuyển động chênh lệch cho các khối con của PU được phân chia bằng AMP khi PU này được mã hóa bằng cách sử dụng dự báo chuyển động giữa các cảnh nhìn và/hoặc MVI có thể làm tăng độ chính xác của việc dự báo chuyển động, do đó làm tăng hiệu quả mã hóa.

Cần hiểu rằng tùy thuộc vào từng ví dụ, các hành động hoặc sự kiện nhất định của kỹ thuật bất kỳ được mô tả trong bản mô tả này có thể được thực hiện theo trình tự khác nhau, có thể được bổ sung, kết hợp hoặc được bỏ lại toàn bộ (ví dụ, không phải tất cả các hành động hoặc sự kiện được mô tả đều cần thiết cho việc thực hiện các kỹ thuật này). Hơn nữa, theo các ví dụ nhất định, các hành động hoặc sự kiện có thể được thực hiện đồng thời, ví dụ, thông qua xử lý đa luồng, xử lý gián đoạn, hoặc nhiều bộ xử lý, thay vì theo trình tự.

Theo một hoặc nhiều ví dụ, các chức năng được mô tả có thể được thực hiện trong phần cứng, phần mềm, phần sụn, hoặc sự kết hợp bất kỳ của chúng. Nếu được thực hiện trong phần mềm, thì các chức năng này có thể được lưu trữ hoặc truyền qua, dưới dạng một hoặc nhiều lệnh hoặc mã trên vật ghi đọc được bởi máy tính và được thực hiện bởi bộ xử lý dựa trên phần cứng. Vật ghi đọc được bởi máy tính có thể bao gồm vật ghi lưu trữ đọc được bởi máy tính, mà tương ứng với vật ghi hữu hình chẳng hạn như vật ghi lưu trữ dữ liệu, hoặc phương tiện truyền thông bao gồm cả phương

tiện bất kỳ mà tạo thuận lợi cho việc truyền chương trình máy tính từ nơi này sang nơi khác, ví dụ theo giao thức truyền thông. Theo cách thức này, nói chung, vật ghi đọc được bởi máy tính có thể tương ứng với (1) vật ghi lưu trữ hữu hình đọc được bởi máy tính bất biến hoặc (2) phương tiện truyền thông chẳng hạn như tín hiệu hoặc sóng mang. Các phương tiện lưu trữ dữ liệu có thể là các phương tiện có sẵn bất kỳ mà có thể được truy cập bởi một hoặc nhiều máy tính hoặc một hoặc nhiều bộ xử lý để tìm kiếm các lệnh, mã và/hoặc các cấu trúc dữ liệu để thực hiện các kỹ thuật được mô tả trong tài liệu này. Sản phẩm chương trình máy tính có thể bao gồm vật ghi đọc được bởi máy tính.

Ví dụ, và không giới hạn, vật ghi lưu trữ đọc được bởi 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 sử dụng 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. Hơn nữa, bất kỳ kết nối nào cũng được gọi là vật ghi đọc được bằng máy tính. Ví dụ, nếu các lệnh được truyền từ một trang mạng, máy chủ, hoặc nguồn từ xa khác nhờ sử dụng cáp đồng trục, cáp quang dạng sợi, cáp xoắn đôi, kênh thuê bao số (digital subscriber line - DSL), hoặc các công nghệ không dây chẳng hạn như tia hồng ngoại, sóng vô tuyến, và vi sóng, thì cáp đồng trục, cáp quang dạng sợi, cáp xoắn đôi, DSL, hoặc các công nghệ không dây chẳng hạn như tia hồng ngoại, sóng vô tuyến, và vi sóng được bao gồm trong phần định nghĩa vật ghi. Tuy nhiên, cần hiểu rằng vật ghi lưu trữ đọc được bởi máy tính và các vật ghi lưu trữ dữ liệu không bao gồm các kết nối, các sóng mang, các tín hiệu, hoặc các phương tiện chuyển tiếp khác, nhưng thay vào đó là hướng đến các phương tiện lưu trữ hữu hình, bất biến. Đĩa và ổ đĩa, như được sử dụng ở đây, bao gồm đĩa nén (Compact Disc - CD), đĩa laze, đĩa quang, đĩa kỹ thuật số đa năng (Digital Versatile Disc - DVD), đĩa mềm và Đĩa Blu-ray, trong đó các đĩa từ thường sao chép dữ liệu theo cách từ tính, trong khi đó đĩa quang sao chép dữ liệu theo cách quang học bằng các tia laze. Kết hợp của cả hai loại này cũng được bao gồm trong phạm vi của phương tiện đọc được bởi máy tính.

Các lệnh có thể được thực hiện bởi một hoặc nhiều bộ xử lý, chẳng hạn như một hoặc nhiều bộ xử lý tín hiệu kỹ thuật số (Digital Signal Processors - DSP), các bộ vi xử lý đa năng, mạch tích hợp chuyên dụng (Application Specific Integrated Circuits

- ASIC), các mảng logic lập trình được dạng trường (field programmable logic arrays - FPGAs), hoặc mạch logic rời rạc hoặc tích hợp tương đương khác. Do vậy, thuật ngữ “bộ xử lý” như được sử dụng ở đây có thể chỉ cấu trúc bất kỳ trong cấu trúc nêu trên hoặc cấu trúc khác bất kỳ phù hợp cho việc thực hiện các kỹ thuật được mô tả ở đây. Ngoài ra, theo một số khía cạnh, chức năng được mô tả ở đây có thể được tạo trong các môđun phần cứng và/hoặc phần mềm chuyên dụng được tạo cấu hình để mã hóa và giải mã, hoặc được đưa vào trong bộ lập giải mã kết hợp. Ngoài ra, các kỹ thuật này có thể được thực hiện hoàn toàn trong một hoặc nhiều mạch hoặc phần tử logic.

Các kỹ thuật của sáng chế có thể được thực hiện trong nhiều loại cơ cấu 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, môđun hoặc đơn vị khác nhau được mô tả trong bản mô tả này để nhấn mạnh các khía cạnh chức năng của thiết bị được tạo cấu hình để thực hiện các kỹ thuật được bộc lộ, nhưng không nhất thiết đòi hỏi việc thực hiện bởi 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 lập giải mã hoặc được tạo ra bởi tập hợp các đơn vị phần cứng liên kết hoạt động, bao gồm một hoặc nhiều bộ xử lý như được mô tả trên đây, kết hợp với phần mềm và/hoặc phần sụn thích hợp.

Các ví dụ khác nhau đã được mô tả. Các ví dụ này và các ví dụ khác là thuộc phạm vi của bộ Yêu cầu bảo hộ sau đây.

YÊU CẦU BẢO HỘ

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

nhận, trong dòng bit video mã hóa, dữ liệu dư tương ứng với đơn vị dự báo của dữ liệu video mà được mã hóa theo chế độ phân chia chuyển động không đối xứng (AMP - asymmetric motion partitioning), được dự báo một chiều bằng cách sử dụng kỹ thuật dự báo tổng hợp cảnh nhìn ngược (backward view synthesis prediction - BVSP), và có kích thước 16x12, 12x16, 16x4 hoặc 4x16;

nhận, trong dòng bit video mã hóa, dữ liệu biểu thị rằng đơn vị dự báo được dự báo một chiều theo kỹ thuật dự báo BVSP;

thu nhận một tập hợp thông tin chuyển động duy nhất cho toàn bộ đơn vị dự báo;

suy ra một vector chênh lệch cho toàn bộ đơn vị dự báo dựa vào một tập hợp thông tin chuyển động duy nhất thu được cho toàn bộ đơn vị dự báo;

phân chia đơn vị dự báo theo chế độ AMP để tạo ra nhiều khối con, mỗi khối con có kích thước 8x4 hoặc 4x8, và mỗi khối con có tập hợp thông tin chuyển động tương ứng;

suy ra vector chuyển động chênh lệch tương ứng cho mỗi khối con của nhiều khối con từ khối có chiều sâu tương ứng trong hình ảnh có chiều sâu tương ứng với hình ảnh tham chiếu, trong đó bước suy ra được dựa trên đơn vị dự báo được dự báo một chiều theo kỹ thuật dự báo BVSP, và trong đó bước suy ra sử dụng một tập hợp thông tin chuyển động duy nhất thu được cho toàn bộ đơn vị dự báo;

tổng hợp khối tham chiếu tương ứng cho mỗi khối con trong số nhiều khối con bằng cách sử dụng vector chuyển động chênh lệch suy ra tương ứng; và

giải mã đơn vị dự báo bằng cách thực hiện bù chuyển động trên mỗi trong số các khối con bằng cách sử dụng dữ liệu dư và các khối tham chiếu tương ứng tổng hợp được.

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

nhận một hoặc nhiều phần tử cú pháp chỉ báo rằng đơn vị dự báo được mã hóa theo chế độ AMP; và

nhận chỉ số ứng viên hợp nhất trở đến ứng viên BVSP bao gồm trong danh sách ứng viên hợp nhất.

3. Phương pháp theo điểm 1, trong đó bước suy ra vector chuyển động chênh lệch tương ứng cho mỗi khối con bao gồm các bước:

định vị khối có chiều sâu tương ứng cho mỗi khối con trong số nhiều khối con bằng cách sử dụng một vector chênh lệch được suy ra cho đơn vị dự báo; và

biến đổi một giá trị chiều sâu được chọn của khối có chiều sâu tương ứng cho mỗi khối con trong số nhiều khối con thành vector chuyển động chênh lệch tương ứng.

4. Phương pháp theo điểm 1, trong đó đơn vị dự báo là đơn vị dự báo thứ nhất, và trong đó nhiều khối con là nhiều khối con thứ nhất, phương pháp này còn bao gồm các bước:

nhận dữ liệu dư tương ứng với đơn vị dự báo thứ hai, trong đó đơn vị dự báo thứ hai được mã hóa bằng cách sử dụng ít nhất một trong dự báo chuyển động giữa các cảnh nhìn hoặc kế thừa vector chuyển động, và có kích thước 16x4 hoặc 4x16;

phân chia đơn vị dự báo thứ hai thành nhiều khối con thứ hai, mỗi khối con tương ứng trong số nhiều khối con thứ hai có kích thước 8x4 hoặc 4x8;

suy ra thông tin chuyển động cho mỗi khối con trong số nhiều khối con thứ hai từ một khối tham chiếu tương ứng; và

giải mã đơn vị dự báo thứ hai bằng cách thực hiện bù chuyển động trên mỗi khối con trong số nhiều khối con thứ hai bằng cách sử dụng dữ liệu dư, thông tin chuyển động suy ra được, và một danh sách hình ảnh tham chiếu.

5. Phương pháp theo điểm 4, trong đó bước thực hiện bù chuyển động bao gồm việc thực hiện bù chuyển động một chiều đối với hình ảnh trong một danh sách hình ảnh tham chiếu.

6. Phương pháp theo điểm 1, trong đó đơn vị dự báo là đơn vị dự báo thứ nhất, và trong đó nhiều khối con là nhiều khối con thứ nhất, phương pháp này còn bao gồm các bước:

nhận dữ liệu dư tương ứng với đơn vị dự báo thứ hai, trong đó đơn vị dự báo thứ hai được mã hóa bằng cách sử dụng ít nhất một trong số dự báo chuyển động giữa các cảnh nhìn hoặc kế thừa vector chuyển động, và có kích thước 16x12 hoặc 12x16;

phân chia đơn vị dự báo thứ hai của dữ liệu video thành nhiều khối con thứ hai; và

giải mã mỗi khối con trong số nhiều khối con thứ hai bằng giá trị dự báo một dự báo.

7. Phương pháp theo điểm 1, trong đó bước thu nhận một tập hợp thông tin chuyển động duy nhất cho toàn bộ đơn vị dự báo bao gồm bước nhận một tập hợp thông tin chuyển động duy nhất trong dòng bit video mã hóa, và trong đó bước giải mã đơn vị dự báo bao gồm bước giải mã đơn vị dự báo theo chế độ hợp nhất.

8. Phương pháp theo điểm 1, trong đó bước thu nhận một tập hợp thông tin chuyển động duy nhất cho toàn bộ đơn vị dự báo bao gồm bước suy ra một tập hợp thông tin chuyển động duy nhất, và trong đó bước giải mã đơn vị dự báo bao gồm bước giải mã đơn vị dự báo theo chế độ dự báo vectơ chuyển động cải tiến (AMVP - advanced motion vector prediction).

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

tạo ra đơn vị dự báo theo chế độ AMP, sao cho đơn vị dự báo được dự báo một chiều bằng cách sử dụng kỹ thuật BVSP, và có kích thước 16x12, 12x16, 16x4 hoặc 4x16;

phân chia đơn vị dự báo theo chế độ AMP để tạo ra nhiều khối con, mỗi khối con có kích thước 8x4 hoặc 4x8, và mỗi khối con có một tập hợp thông tin chuyển động tương ứng;

suy ra vectơ chuyển động chênh lệch tương ứng, dựa vào tập hợp thông tin chuyển động tương ứng, cho mỗi khối con trong số nhiều khối con từ khối có chiều sâu tương ứng trong hình ảnh có chiều sâu tương ứng với hình ảnh tham chiếu;

tổng hợp khối tham chiếu tương ứng cho mỗi khối con trong số nhiều khối con được phân chia từ đơn vị dự báo, bằng cách sử dụng các vectơ chuyển động chênh lệch suy ra được tương ứng;

mã hóa đơn vị dự báo bằng cách thực hiện bù chuyển động trên mỗi khối con trong số nhiều khối con bằng cách sử dụng các khối tham chiếu tương ứng tổng hợp được; và

thực hiện một trong các thao tác:

báo hiệu, trong dòng bit video mã hóa, một tập hợp thông tin chuyển động duy nhất cho toàn bộ đơn vị dự báo đã mã hóa; hoặc

thừa kế, cho toàn bộ đơn vị dự báo đã mã hóa, một tập hợp thông tin chuyển động duy nhất từ một hoặc nhiều khối lân cận của đơn vị dự báo đã mã hóa.

10. Phương pháp theo điểm 9, trong đó phương pháp này còn bao gồm bước:

tạo ra một hoặc nhiều phần tử cú pháp chỉ báo rằng đơn vị dự báo được mã hóa chế độ AMP; và

tạo ra chỉ số ứng viên hợp nhất trở đến ứng viên BVSP trong danh sách ứng viên.

11. Phương pháp theo điểm 9, trong đó bước suy ra vector chuyển động chênh lệch tương ứng cho mỗi trong số các khối con bao gồm các bước:

định vị khối có chiều sâu tương ứng cho mỗi trong số các khối con bằng cách sử dụng một vector chênh lệch được suy ra cho đơn vị dự báo; và

biến đổi một giá trị chiều sâu đã chọn của khối có chiều sâu tương ứng cho mỗi trong số các khối con thành vector chuyển động chênh lệch tương ứng.

12. Phương pháp theo điểm 9, trong đó đơn vị dự báo là đơn vị dự báo thứ nhất, phương pháp này còn bao gồm các bước:

tạo ra đơn vị dự báo thứ hai theo chế độ AMP, sao cho đơn vị dự báo thứ hai được mã hóa bằng cách sử dụng ít nhất một trong dự báo chuyển động giữa các cảnh nhìn hoặc kế thừa vector chuyển động, và có kích thước 16x4 hoặc 4x16;

phân chia đơn vị dự báo thứ hai để tạo ra nhiều khối con thứ hai, mỗi khối con trong số nhiều khối con thứ hai có kích thước 8x4 hoặc 4x8;

suy ra thông tin chuyển động cho mỗi khối con trong số nhiều khối con thứ hai từ một khối tham chiếu tương ứng; và

mã hóa đơn vị dự báo thứ hai bằng cách thực hiện bù chuyển động trên mỗi khối con trong số nhiều khối con thứ hai bằng cách sử dụng thông tin chuyển động suy ra được và một danh sách hình ảnh tham chiếu.

13. Phương pháp theo điểm 12, trong đó bước thực hiện bù chuyển động bao gồm việc thực hiện bù chuyển động một chiều đối với hình ảnh trong một danh sách hình ảnh tham chiếu.

14. Phương pháp theo điểm 9, trong đó đơn vị dự báo là đơn vị dự báo thứ nhất, phương pháp này còn bao gồm các bước:

tạo ra đơn vị dự báo thứ hai theo chế độ AMP, trong đó đơn vị dự báo thứ hai được mã hóa bằng cách sử dụng ít nhất một trong dự báo chuyển động giữa các cảnh nhìn hoặc kế thừa vectơ chuyển động, và có kích thước 16×12 hoặc 12×16 ;

phân chia đơn vị dự báo thứ hai để tạo ra nhiều khối con thứ hai; và

mã hóa mỗi khối con trong số nhiều khối con thứ hai bằng dự báo một dự báo.

15. Thiết bị được tạo cấu hình để giải mã dữ liệu video, thiết bị này bao gồm:

giao diện truyền thông được tạo cấu hình để nhận dòng bit video mã hóa;

bộ nhớ gắn với giao diện truyền thông và được tạo cấu hình để lưu trữ dòng bit video mã hóa, dòng bit video mã hóa này bao gồm dữ liệu dư tương ứng với đơn vị dự báo; và

một hoặc nhiều bộ xử lý truyền thông với thiết bị nhớ, một hoặc nhiều bộ xử lý được tạo cấu hình để:

phục hồi, từ bộ nhớ dòng bit video mã hóa, dữ liệu dư tương ứng với đơn vị dự báo, trong đó đơn vị dự báo được mã hóa theo công cụ mã hóa AMP, được dự báo một chiều bằng cách sử dụng kỹ thuật BVSP, và có kích thước 16×12 , 12×16 , 16×4 hoặc 4×16 ;

phục hồi, từ dòng bit video mã hóa được lưu trữ vào bộ nhớ, dữ liệu chỉ báo rằng đơn vị dự báo được dự báo một chiều theo BVSP;

thu được một tập hợp thông tin chuyển động duy nhất cho toàn bộ đơn vị dự báo;

suy ra một vectơ chênh lệch duy nhất cho toàn bộ đơn vị dự báo dựa vào một tập hợp thông tin chuyển động duy nhất thu được cho toàn bộ đơn vị dự báo;

phân chia đơn vị dự báo theo chế độ AMP để tạo ra nhiều khối con, mỗi khối con có kích thước 8x4 hoặc 4x8, và mỗi khối con có tập hợp thông tin chuyển động tương ứng;

suy ra vectơ chuyển động chênh lệch tương ứng cho mỗi khối con trong số nhiều khối con từ khối có chiều sâu tương ứng trong hình ảnh có chiều sâu tương ứng với hình ảnh tham chiếu, trong đó để suy ra vectơ chuyển động chênh lệch tương ứng, một hoặc nhiều bộ xử lý được tạo cấu hình để suy ra vectơ chuyển động chênh lệch tương ứng dựa vào đơn vị dự báo được dự báo một chiều theo BVSP, và trong đó để suy ra vectơ chuyển động chênh lệch tương ứng, một hoặc nhiều bộ xử lý được tạo cấu hình để sử dụng một tập hợp thông tin chuyển động thu được cho toàn bộ đơn vị dự báo; tổng hợp khối tham chiếu tương ứng cho mỗi khối con trong số nhiều khối con bằng cách sử dụng các vectơ chuyển động chênh lệch suy ra được tương ứng; và

giải mã đơn vị dự báo bằng cách thực hiện bù chuyển động trên mỗi trong số các khối con bằng cách sử dụng dữ liệu dư và các khối tham chiếu tương ứng tổng hợp được.

16. Thiết bị theo điểm 15, trong đó một hoặc nhiều bộ xử lý còn được tạo cấu hình để:

nhận một hoặc nhiều phần tử cú pháp chỉ báo rằng đơn vị dự báo được mã hóa theo chế độ AMP, và

nhận chỉ số ứng viên hợp nhất trở đến ứng viên BVSP được bao gồm trong danh sách ứng viên hợp nhất.

17. Thiết bị theo điểm 15, trong đó để suy ra vectơ chuyển động chênh lệch tương ứng cho mỗi khối con, một hoặc nhiều bộ xử lý còn được tạo cấu hình để:

định vị khối có chiều sâu tương ứng cho mỗi khối con trong số nhiều khối con bằng cách sử dụng một vectơ chênh lệch được suy ra cho đơn vị dự báo; và

biến đổi một giá trị chiều sâu đã chọn của khối có chiều sâu tương ứng cho mỗi khối con trong số nhiều khối con thành vectơ chuyển động chênh lệch tương ứng.

18. Thiết bị theo điểm 15, trong đó đơn vị dự báo là đơn vị dự báo thứ nhất, trong đó nhiều khối con là nhiều khối con thứ nhất, và trong đó một hoặc nhiều bộ xử lý còn được tạo cấu hình để:

nhận dữ liệu dư tương ứng với đơn vị dự báo thứ hai, trong đó đơn vị dự báo thứ hai được mã hóa bằng cách sử dụng ít nhất một trong dự báo chuyển động giữa các cảnh nhìn hoặc kế thừa vector chuyển động, và có kích thước 16x4 hoặc 4x16;

phân chia đơn vị dự báo thứ hai để tạo ra nhiều khối con thứ hai, mỗi khối con tương ứng trong số nhiều khối con thứ hai có kích thước 8x4 hoặc 4x8;

suy ra thông tin chuyển động cho mỗi khối con trong số nhiều khối con thứ hai từ một khối tham chiếu tương ứng; và

giải mã đơn vị dự báo thứ hai bằng cách thực hiện bù chuyển động trên mỗi khối con trong số nhiều khối con thứ hai bằng cách sử dụng dữ liệu dư, thông tin chuyển động suy ra được, và một danh sách hình ảnh tham chiếu.

19. Thiết bị theo điểm 18, trong đó để thực hiện bù chuyển động, một hoặc nhiều bộ xử lý còn được tạo cấu hình để thực hiện bù chuyển động một chiều đối với hình ảnh trong một danh sách hình ảnh tham chiếu.

20. Thiết bị theo điểm 15, trong đó đơn vị dự báo là đơn vị dự báo thứ nhất, trong đó nhiều khối con là nhiều khối con thứ nhất, và trong đó một hoặc nhiều bộ xử lý còn được tạo cấu hình để:

nhận dữ liệu dư tương ứng với đơn vị dự báo thứ hai, trong đó đơn vị dự báo thứ hai được mã hóa bằng cách sử dụng ít nhất một trong dự báo chuyển động giữa các cảnh nhìn hoặc kế thừa vector chuyển động, và có kích thước 16x12 hoặc 12x16;

phân chia đơn vị dự báo thứ hai của dữ liệu video thành nhiều khối con thứ hai; và

giải mã mỗi khối con trong số nhiều khối con thứ hai bằng dự báo một dự báo.

21. Thiết bị theo điểm 15, trong đó thiết bị này còn bao gồm:

thiết bị hiển thị được tạo cấu hình để hiển thị dữ liệu video đã giải mã dựa trên đơn vị dự báo được giải mã.

22. Thiết bị theo điểm 15, trong đó bộ nhớ và một hoặc nhiều bộ xử lý tạo ra bộ giải mã video nằm trong một trong số điện thoại di động, máy tính bảng, máy tính xách tay, máy tính để bàn, hộp giải mã tín hiệu, hoặc tivi.

23. Thiết bị theo điểm 15, trong đó để thu được một tập hợp thông tin chuyển động duy nhất cho toàn bộ đơn vị dự báo, một hoặc nhiều bộ xử lý được tạo cấu hình để

nhận một tập hợp thông tin chuyển động duy nhất trong dòng bit video đã mã hóa, và trong đó để giải mã đơn vị dự báo, một hoặc nhiều bộ xử lý được tạo cấu hình để giải mã đơn vị dự báo theo chế độ hợp nhất.

24. Thiết bị theo điểm 15, trong đó để thu được một tập hợp thông tin chuyển động duy nhất cho toàn bộ đơn vị dự báo, một hoặc nhiều bộ xử lý được tạo cấu hình để suy ra một tập hợp thông tin chuyển động duy nhất, và trong đó để giải mã đơn vị dự báo, một hoặc nhiều bộ xử lý được tạo cấu hình để giải mã đơn vị dự báo theo chế độ AMVP.

25. Thiết bị được tạo cấu hình để giải mã dữ liệu video, thiết bị này bao gồm:

phương tiện để nhận, trong dòng bit video đã mã hóa, dữ liệu dự tương ứng với đơn vị dự báo của dữ liệu video mà được mã hóa theo chế độ AMP, được dự báo một chiều bằng cách sử dụng kỹ thuật BVSP, và có kích thước 16x12, 12x16, 16x4 hoặc 4x16;

phương tiện để nhận, trong dòng bit video đã mã hóa, dữ liệu biểu thị rằng đơn vị dự báo được dự báo một chiều theo BVSP;

phương tiện để thu nhận một tập hợp thông tin chuyển động duy nhất cho toàn bộ đơn vị dự báo;

phương tiện để suy ra một vector chênh lệch cho đơn vị dự báo dựa trên một tập hợp thông tin chuyển động duy nhất thu được cho toàn bộ đơn vị dự báo;

phương tiện để phân chia đơn vị dự báo theo chế độ AMP để tạo ra nhiều khối con, mỗi khối con có kích thước 8x4 hoặc 4x8, và mỗi khối con có tập hợp thông tin chuyển động tương ứng;

phương tiện để suy ra vector chênh lệch tương ứng cho mỗi khối con trong số nhiều khối con từ khối có chiều sâu tương ứng trong hình ảnh có chiều sâu tương ứng với hình ảnh tham chiếu, trong đó phương tiện để suy ra bao gồm phương tiện suy ra vector chênh lệch tương ứng dựa trên đơn vị dự báo được dự báo một chiều theo BVSP, và trong đó phương tiện để suy ra bao gồm phương tiện suy ra vector chênh lệch tương ứng sử dụng một tập hợp thông tin chuyển động duy nhất thu được cho toàn bộ đơn vị dự báo;

phương tiện để tổng hợp khối tham chiếu tương ứng cho mỗi khối con trong số nhiều khối con bằng cách sử dụng vector chuyển động chênh lệch suy ra tương ứng; và

phương tiện để giải mã đơn vị dự báo bằng cách thực hiện bù chuyển động trên mỗi trong số các khối con bằng cách sử dụng dữ liệu dư và các khối tham chiếu tương ứng tổng hợp được.

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

phương tiện nhận một hoặc nhiều phần tử cú pháp biểu thị rằng đơn vị dự báo được mã hóa theo chế độ AMP; và

phương tiện nhận chỉ số ứng viên hợp nhất trở đến ứng viên BVSP được bao gồm trong danh sách ứng viên hợp nhất.

27. Thiết bị theo điểm 25, trong đó phương tiện để thu được vector chuyển động chênh lệch tương ứng cho mỗi khối con bao gồm:

phương tiện để định vị khối có chiều sâu tương ứng cho mỗi khối con trong số nhiều khối con bằng cách sử dụng một vector chênh lệch được suy ra cho đơn vị dự báo; và

phương tiện để biến đổi một giá trị chiều sâu đã chọn của khối có chiều sâu tương ứng cho mỗi khối con trong số nhiều khối con thành vector chuyển động chênh lệch tương ứng.

28. Thiết bị theo điểm 25, trong đó đơn vị dự báo là đơn vị dự báo thứ nhất, và trong đó nhiều khối con là nhiều khối con thứ nhất, thiết bị này còn bao gồm:

phương tiện để nhận dữ liệu dư tương ứng với đơn vị dự báo thứ hai, trong đó đơn vị dự báo thứ hai được mã hóa bằng cách sử dụng ít nhất một trong dự báo chuyển động giữa các cảnh nhìn hoặc kế thừa vector chuyển động, và có kích thước 16x4 hoặc 4x16;

phương tiện để phân chia đơn vị dự báo thứ hai để tạo ra nhiều khối con thứ hai, mỗi khối con trong số nhiều khối con thứ hai có kích thước 8x4 hoặc 4x8;

phương tiện để suy ra thông tin chuyển động cho mỗi khối con trong số nhiều khối con thứ hai từ một khối tham chiếu tương ứng; và

phương tiện để giải mã đơn vị dự báo thứ hai bằng cách thực hiện bù chuyển động trên mỗi khối con trong số nhiều khối con thứ hai bằng cách sử dụng dữ liệu dư, thông tin chuyển động suy ra và một danh sách hình ảnh tham chiếu.

29. Thiết bị theo điểm 28, trong đó phương tiện để thực hiện bù chuyển động bao gồm phương tiện để thực hiện bù chuyển động một chiều đối với một hình ảnh trong một danh sách hình ảnh tham chiếu.

30. Thiết bị theo điểm 25, trong đó đơn vị dự báo là đơn vị dự báo thứ nhất, và trong đó nhiều khối con là nhiều khối con thứ nhất, thiết bị này còn bao gồm:

phương tiện để nhận dữ liệu dư tương ứng với đơn vị dự báo thứ hai, trong đó đơn vị dự báo thứ hai được mã hóa bằng cách sử dụng ít nhất một trong dự báo chuyển động giữa các cảnh nhìn hoặc kế thừa vector chuyển động, và có kích thước 16x12 hoặc 12x16;

phương tiện để phân chia đơn vị dự báo thứ hai để tạo ra nhiều khối con thứ hai; và

phương tiện để giải mã mỗi khối con trong số nhiều khối con thứ hai bằng dự báo một dự báo.

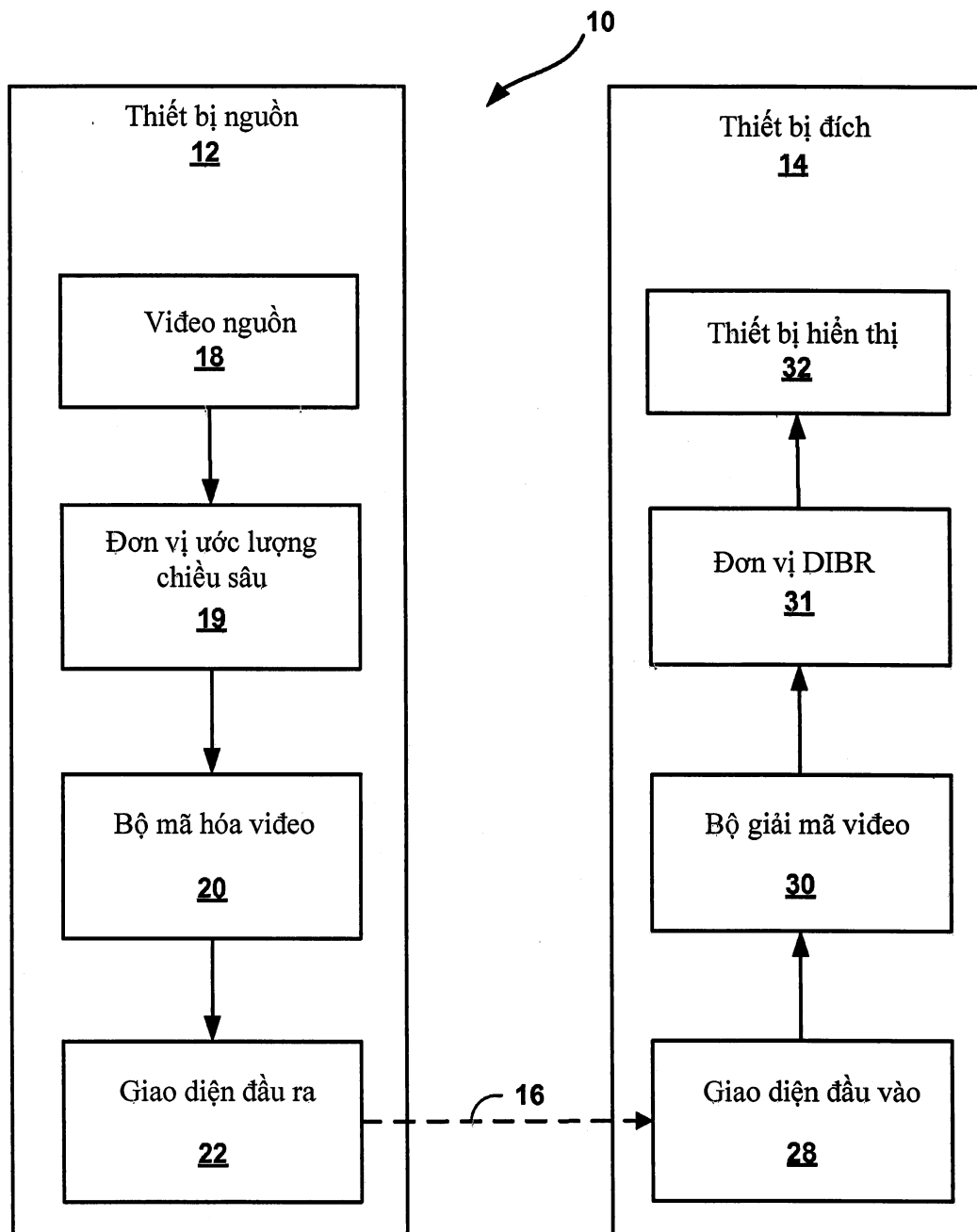


FIG. 1

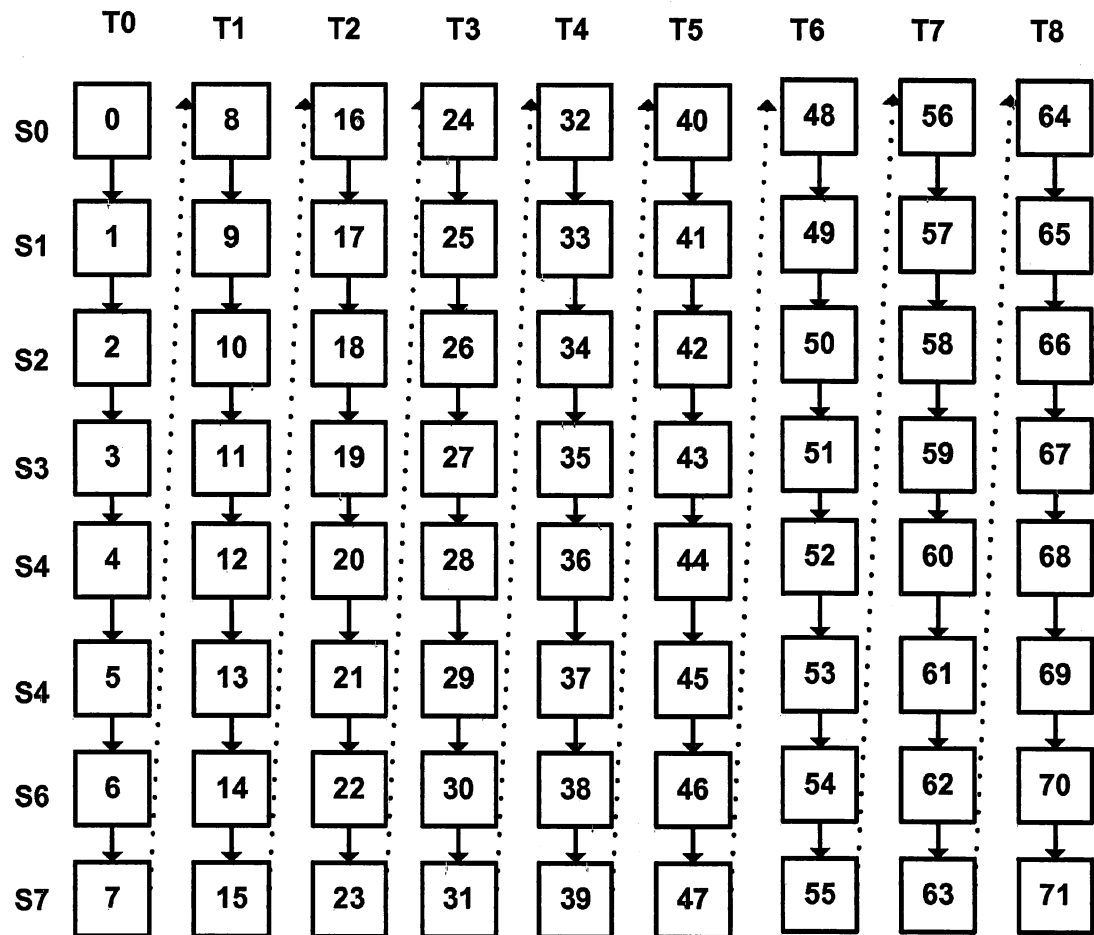


FIG. 2

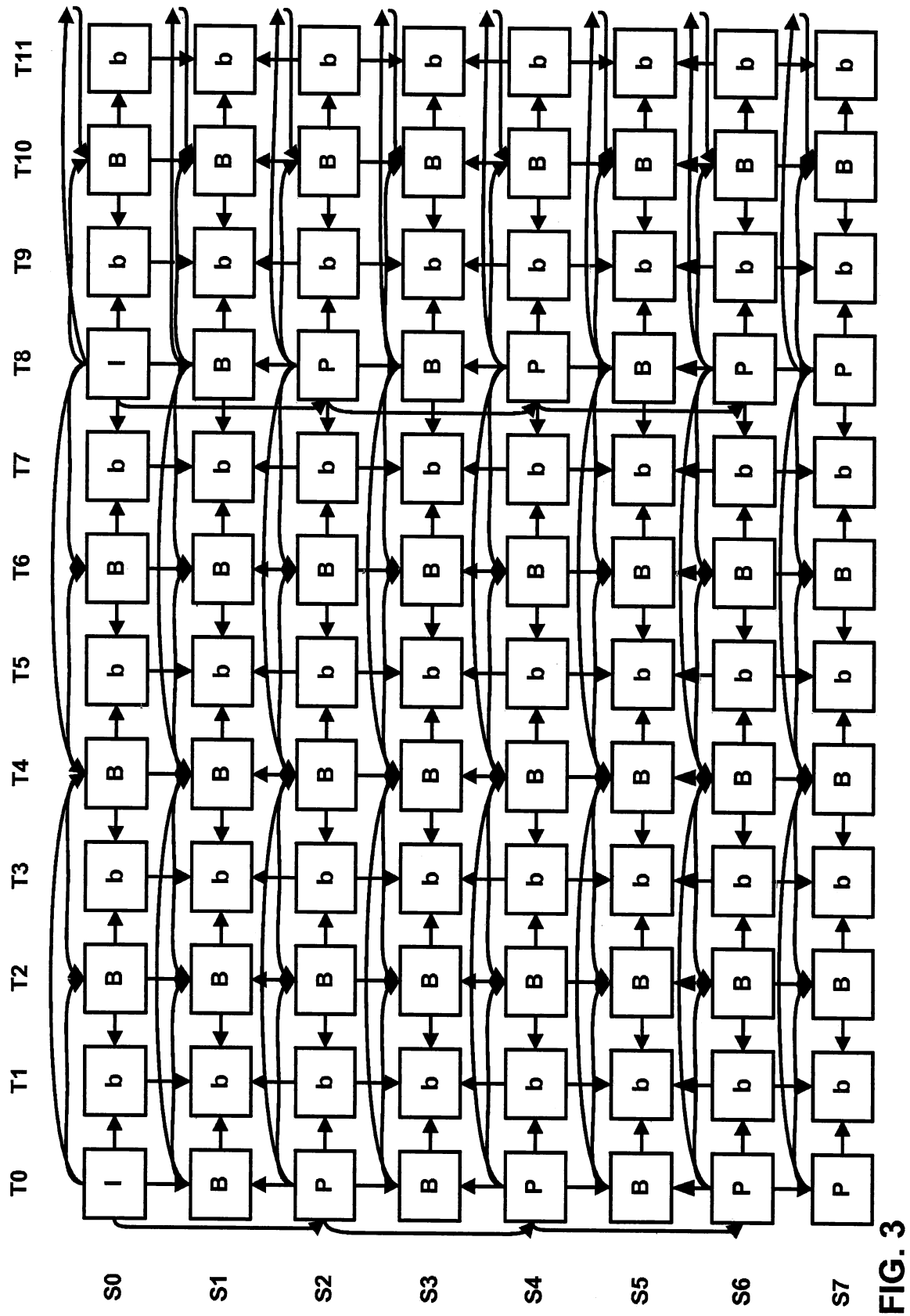


FIG. 3

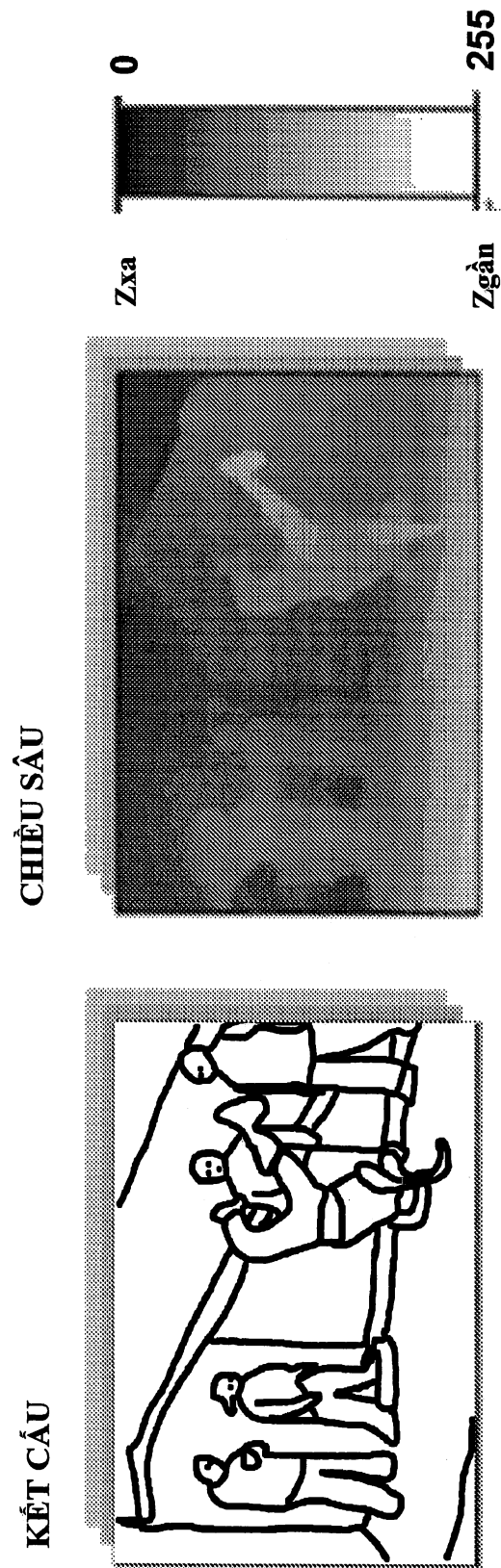
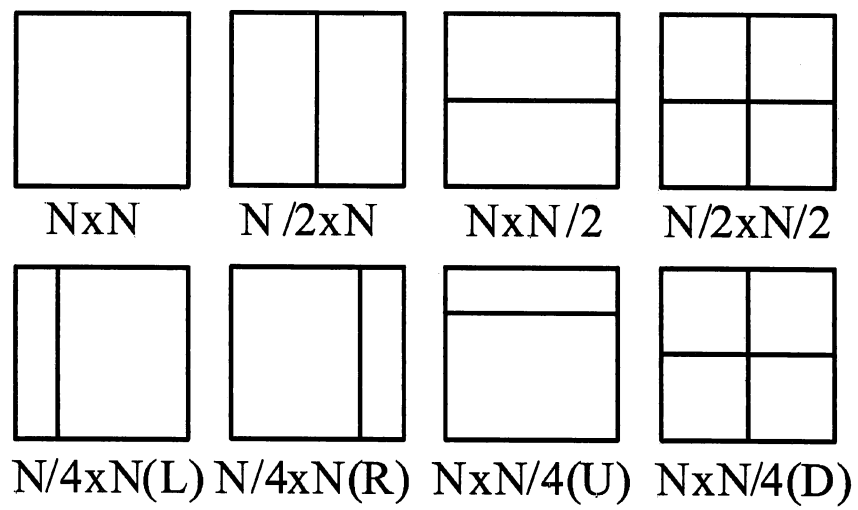


FIG. 4

**FIG. 5**

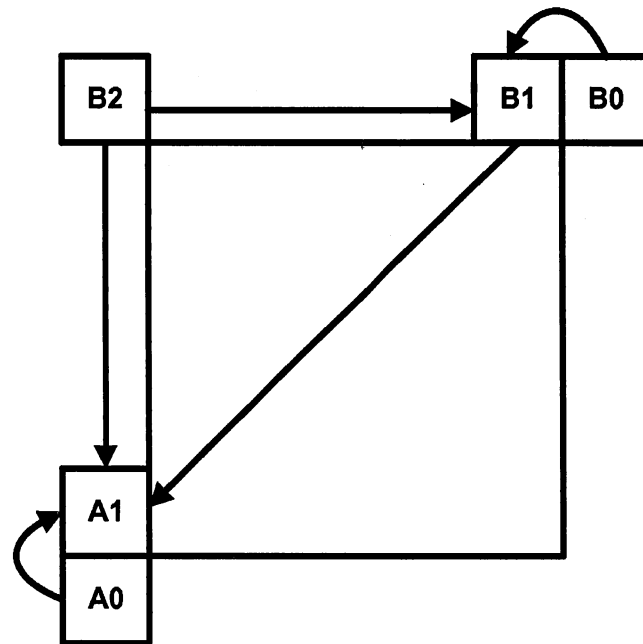
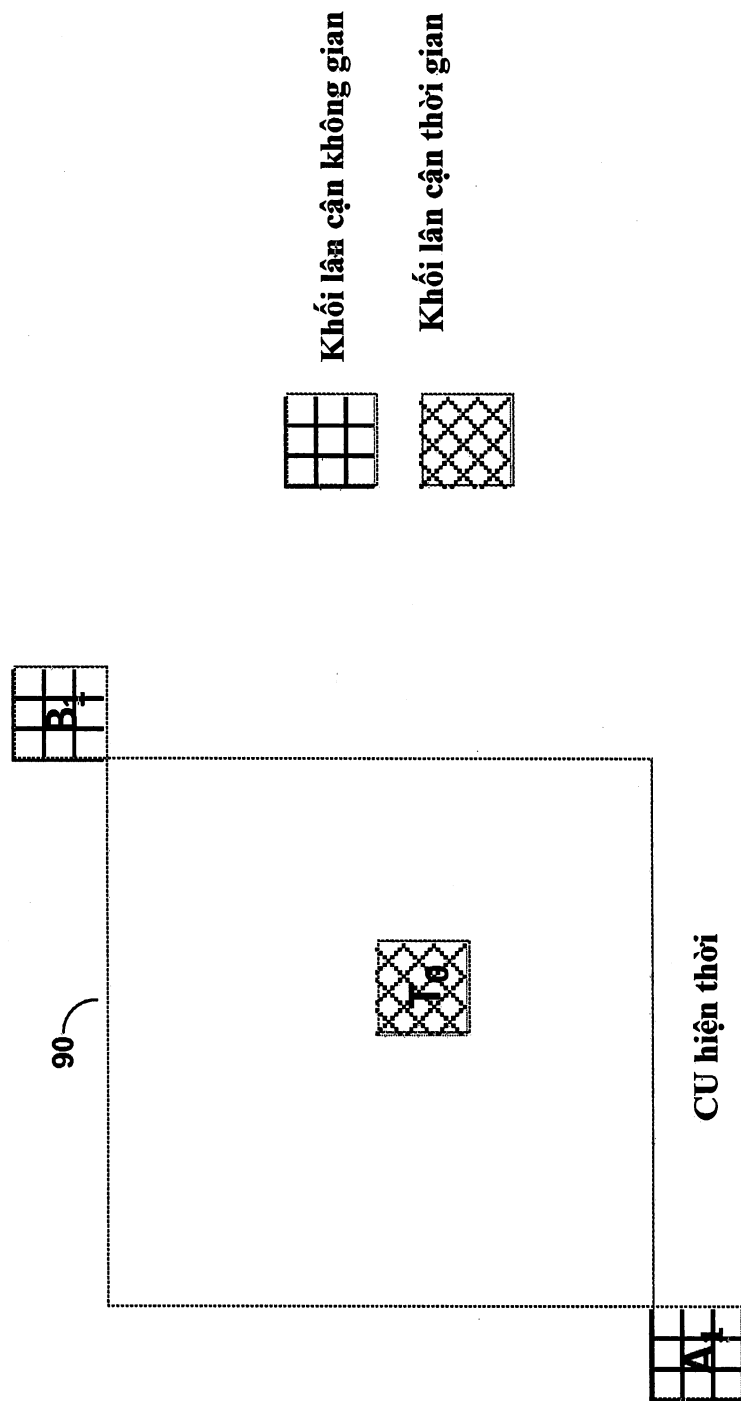


FIG. 6

comblidx	0	1	2	3	4	5	6	7	8	9	10	11
l0CandIdx	0	1	0	2	1	2	0	3	1	3	2	3
l1CandIdx	1	0	2	0	2	1	3	0	3	1	3	2

FIG. 7



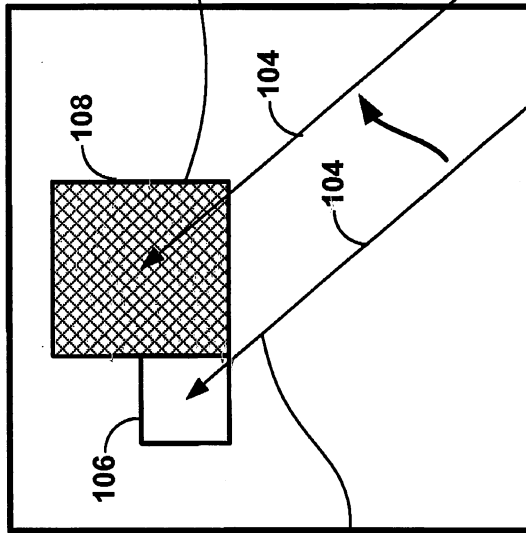
Khối lân cận không gian và thời gian so với đơn vị mã hóa hiện thời

FIG. 8

SUY RA KHỐI CHIỀU SÂU TỪ KHUNG
HÌNH THAM CHIỀU ĐỂ THỰC HIỆN DỰ
ĐOÁN BVSP

38937

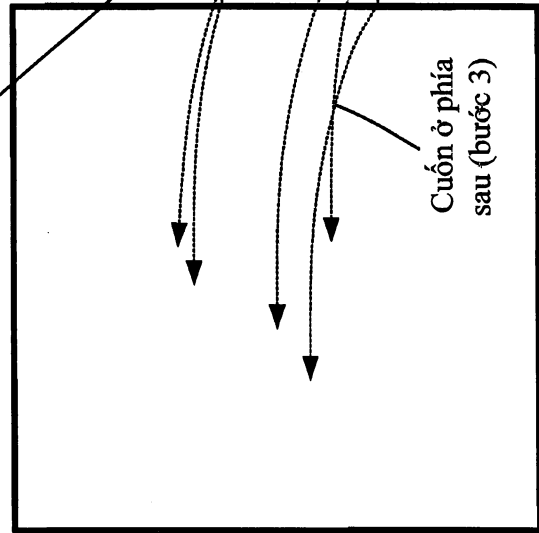
Hình ảnh chiều sâu tham
chiếu



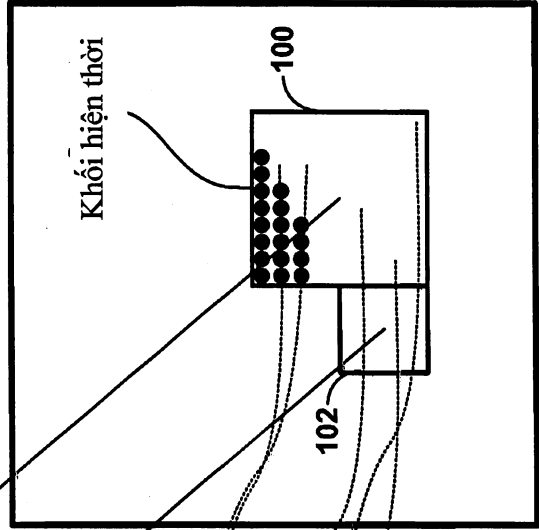
Vector chênh
lệch từ khối
lân cận
(Bước 1)

Khối chiều sâu
được sử dụng
cho BVSP

Vector chênh lệch
được sao chép từ
khối lân cận
(Bước 2)

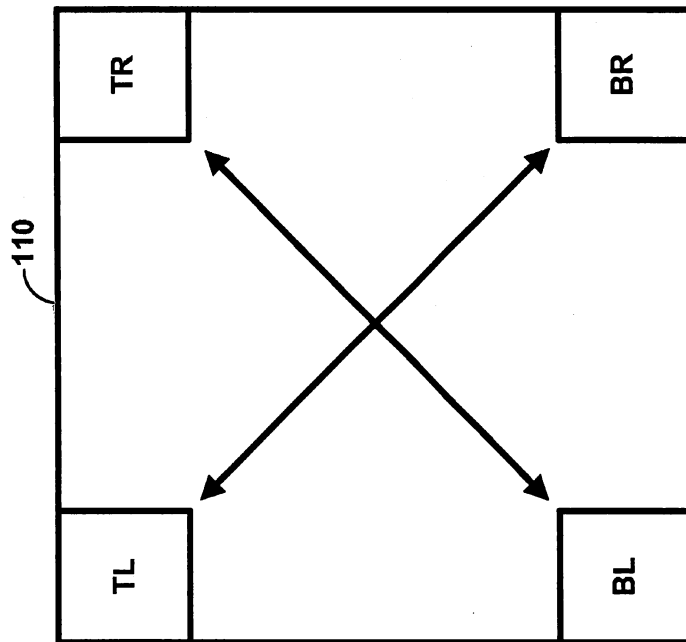


Hình ảnh kết cấu tham chiếu



Hình ảnh kết cấu phụ thuộc

FIG. 9



Bốn điểm ảnh góc của một khối chiều sâu 8x8

FIG. 10

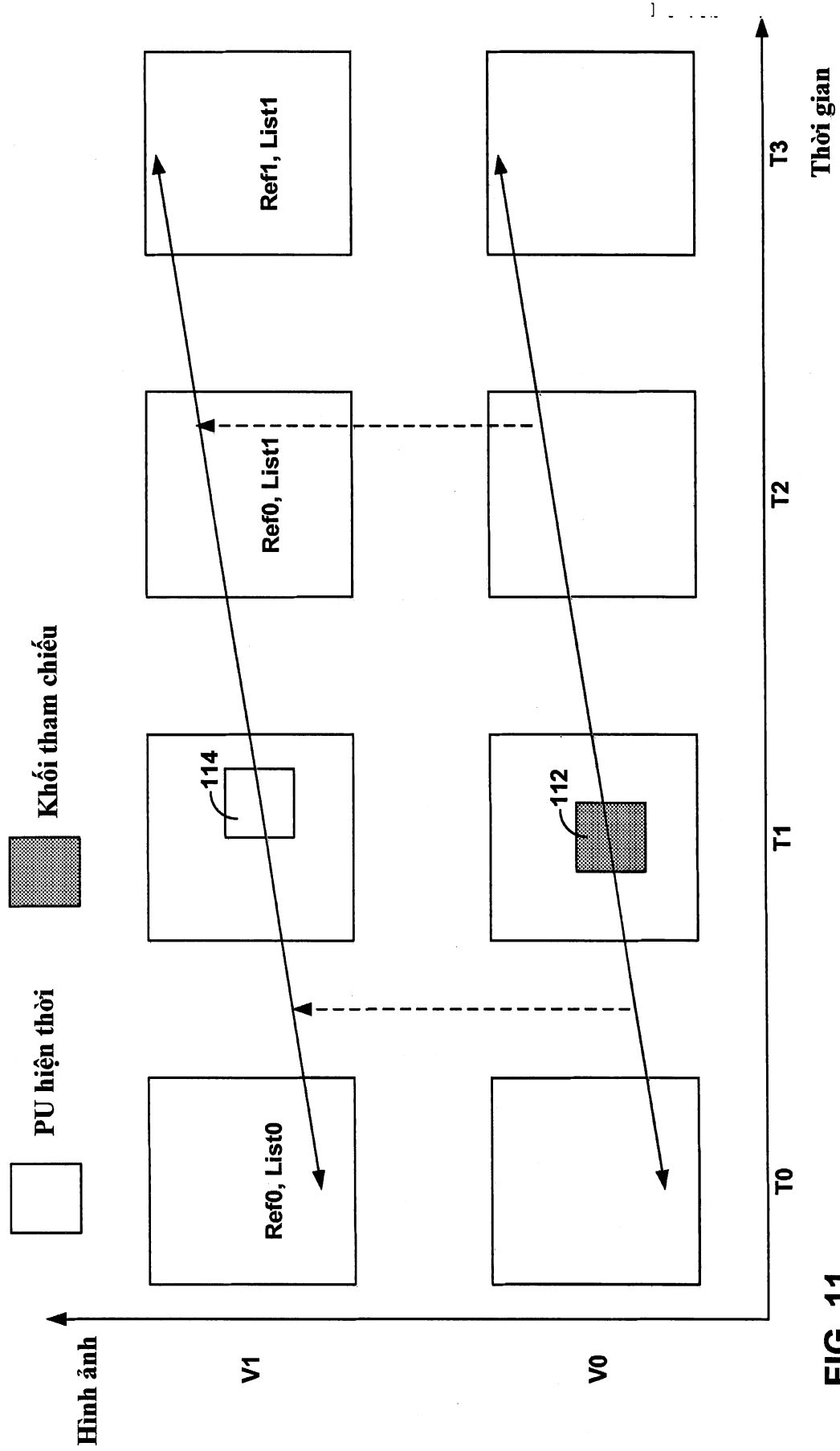


FIG. 11

comidx	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
I0Candidx	0	1	0	2	1	2	0	3	1	3	2	3	0	4	1	4	2	4	3	4
I1Candidx	1	0	2	0	2	1	3	0	3	1	3	2	4	0	4	1	4	2	4	3

FIG. 12

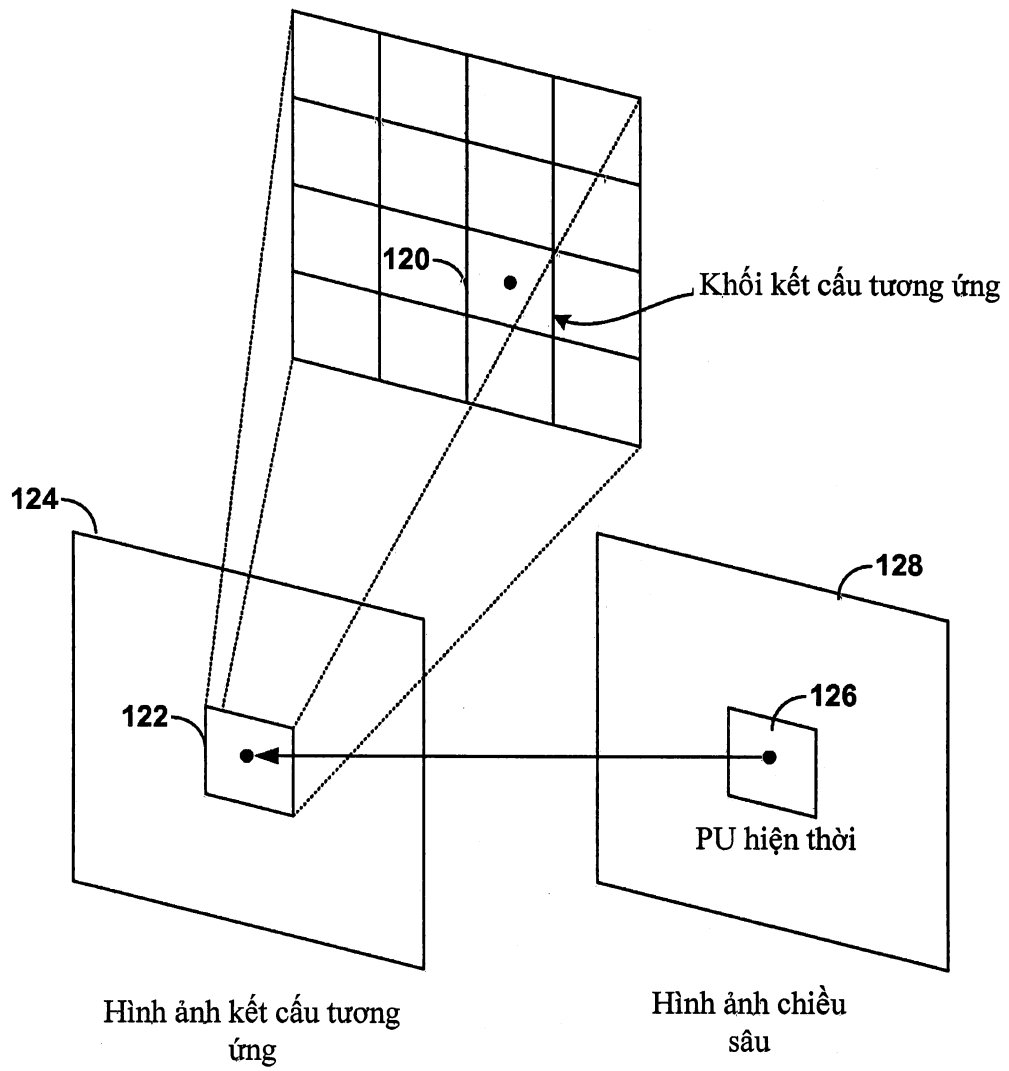


FIG. 13

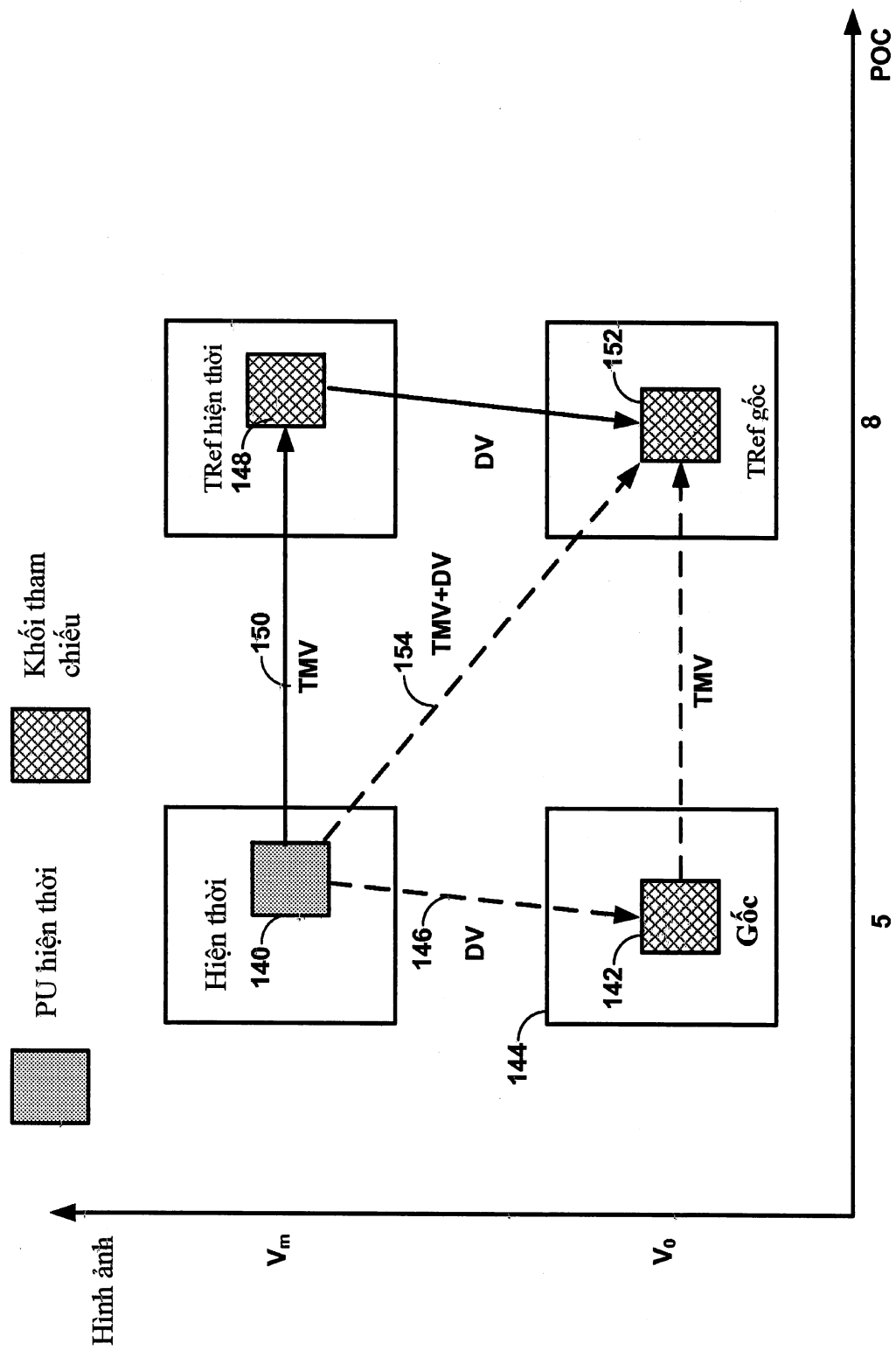


FIG. 14

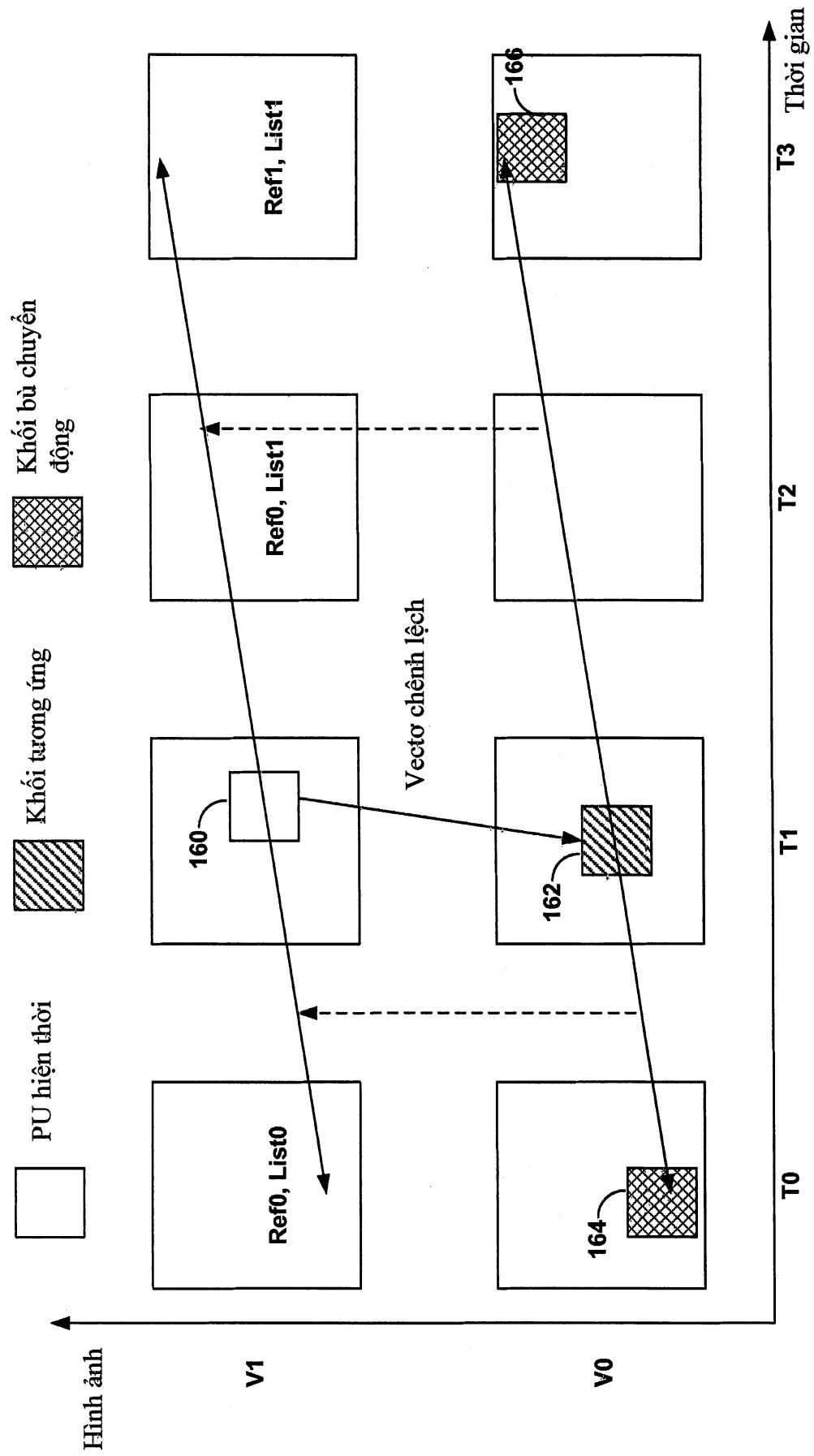


FIG. 15

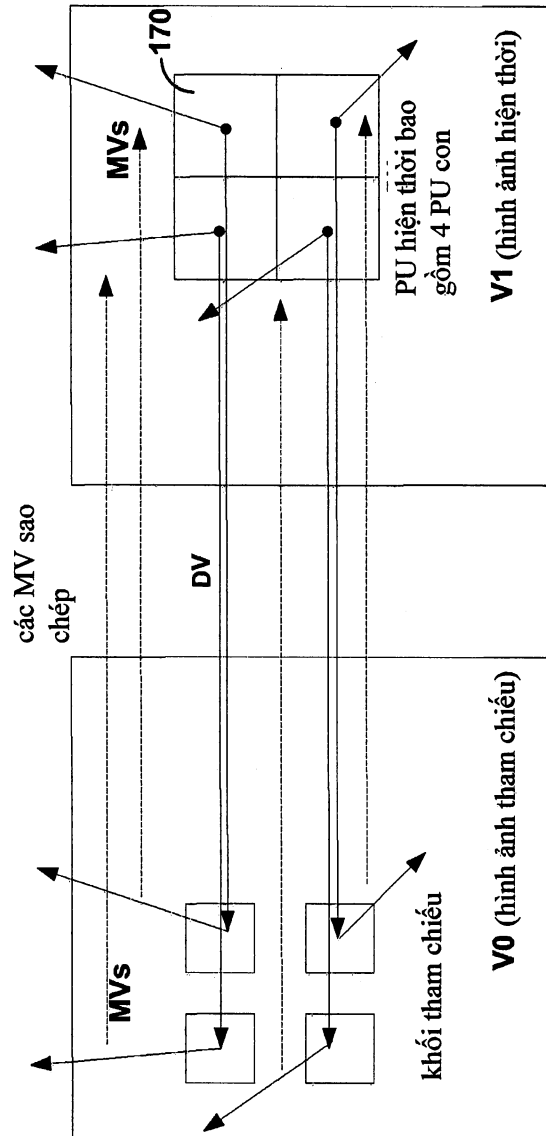


FIG. 16

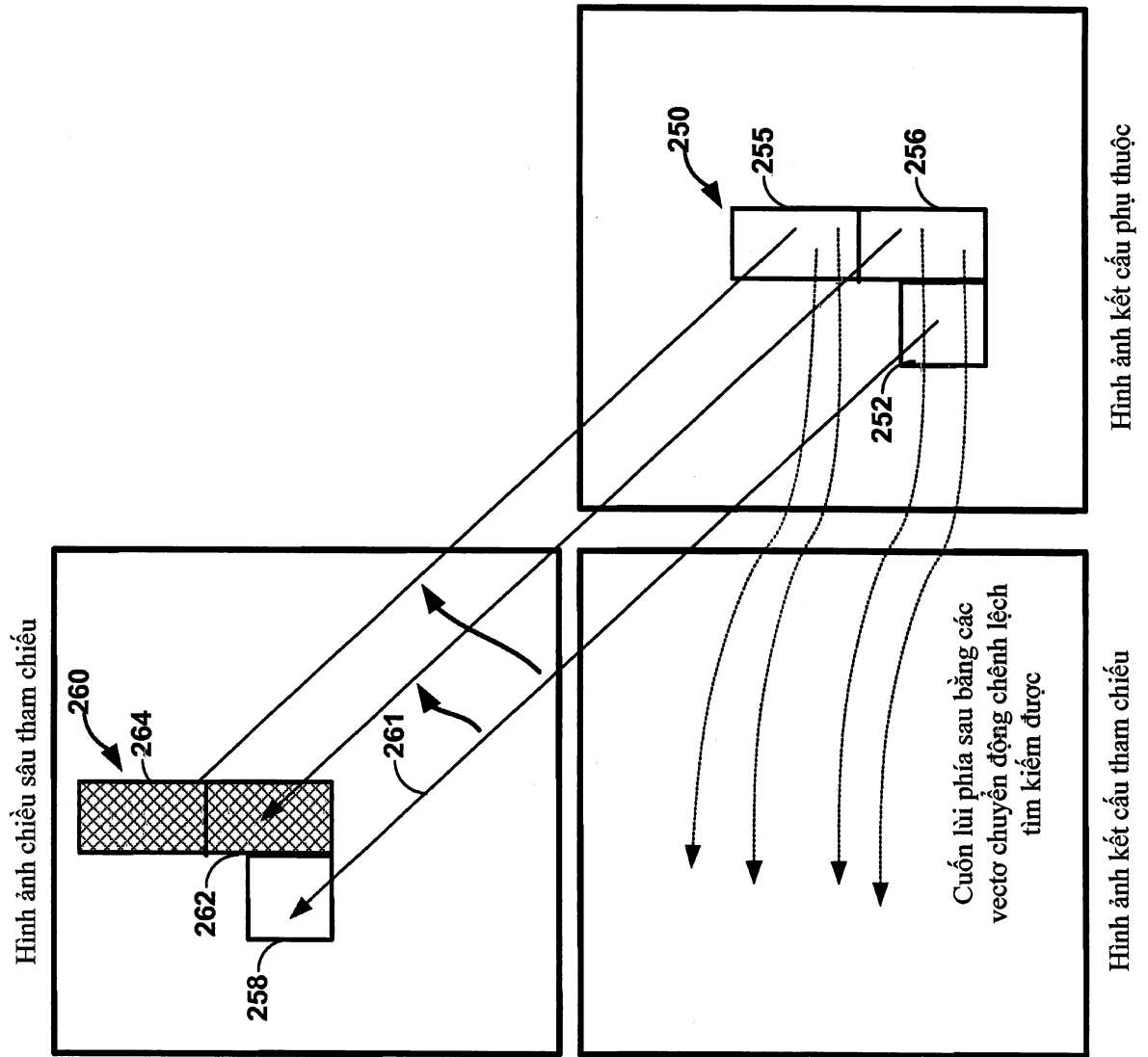


FIG. 17

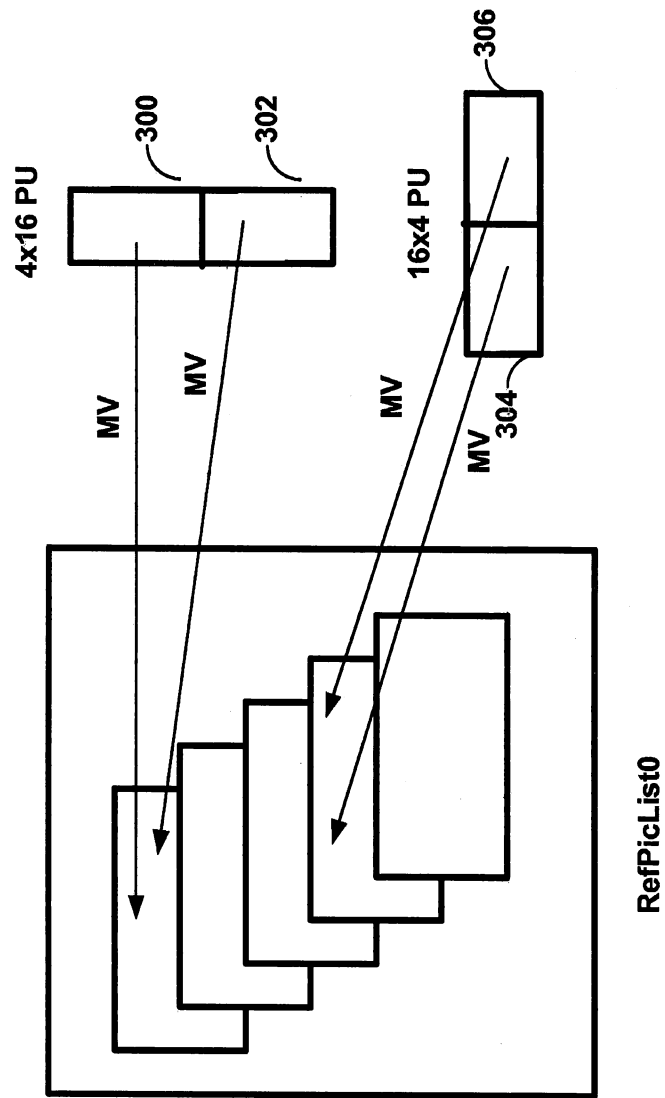


FIG. 18

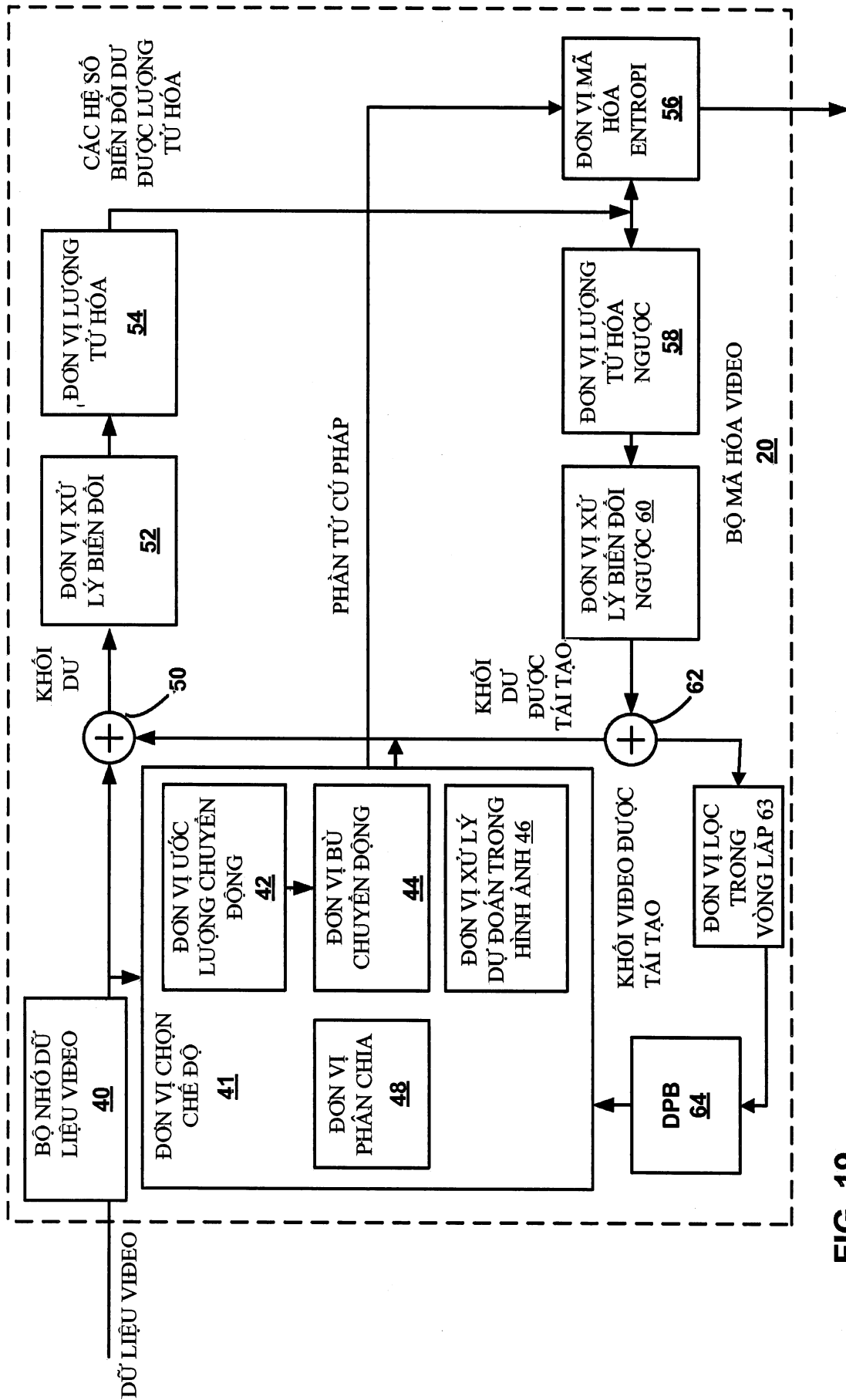


FIG. 19

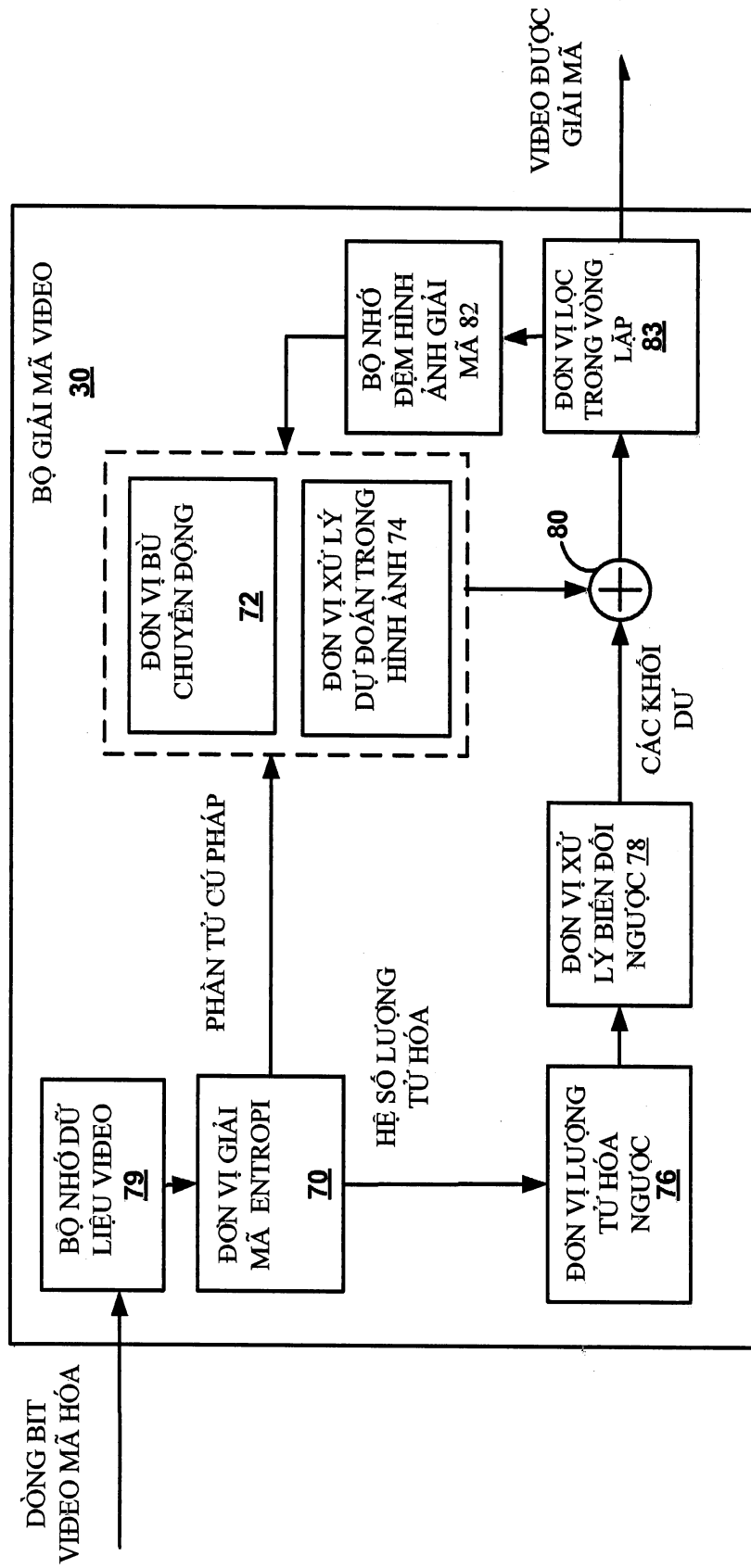


FIG. 20

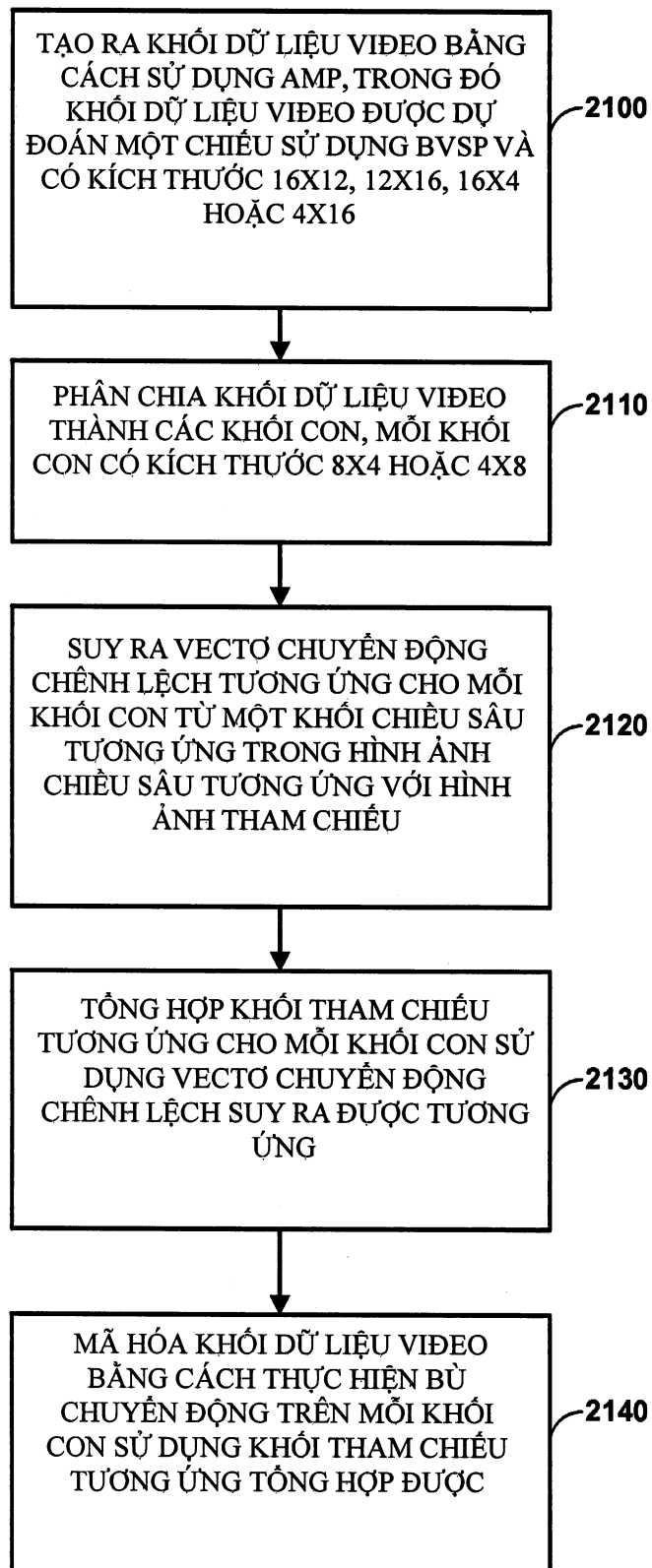


FIG. 21

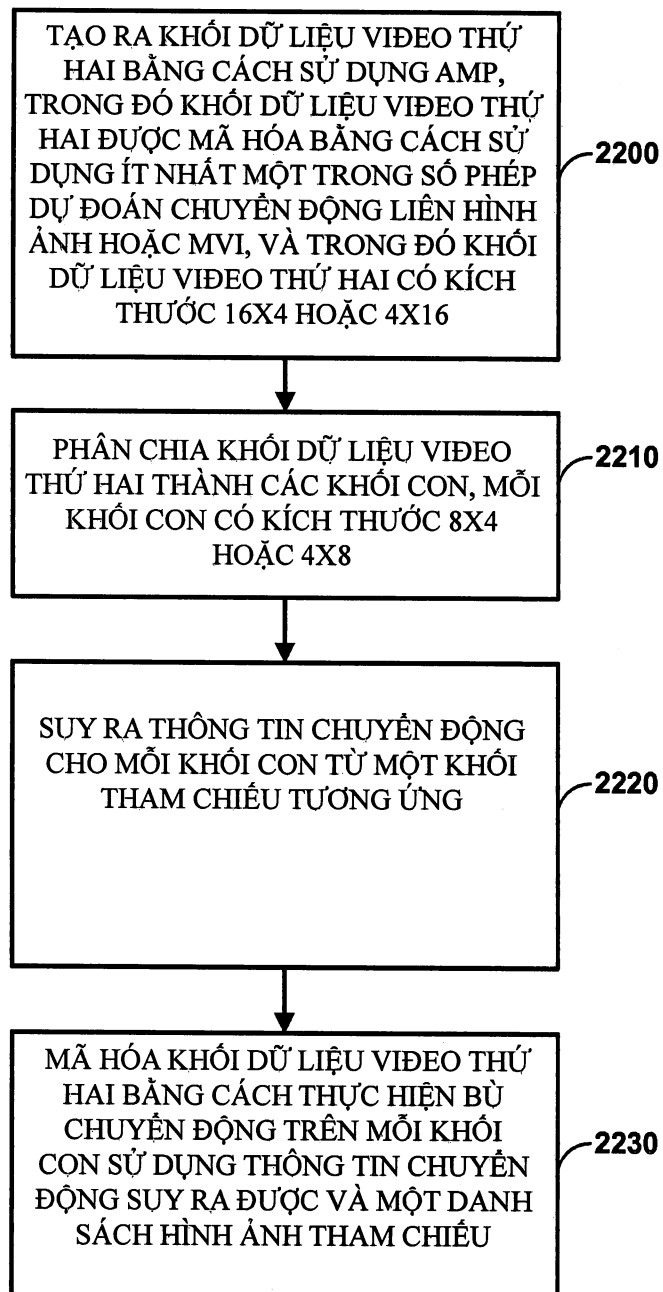


FIG. 22

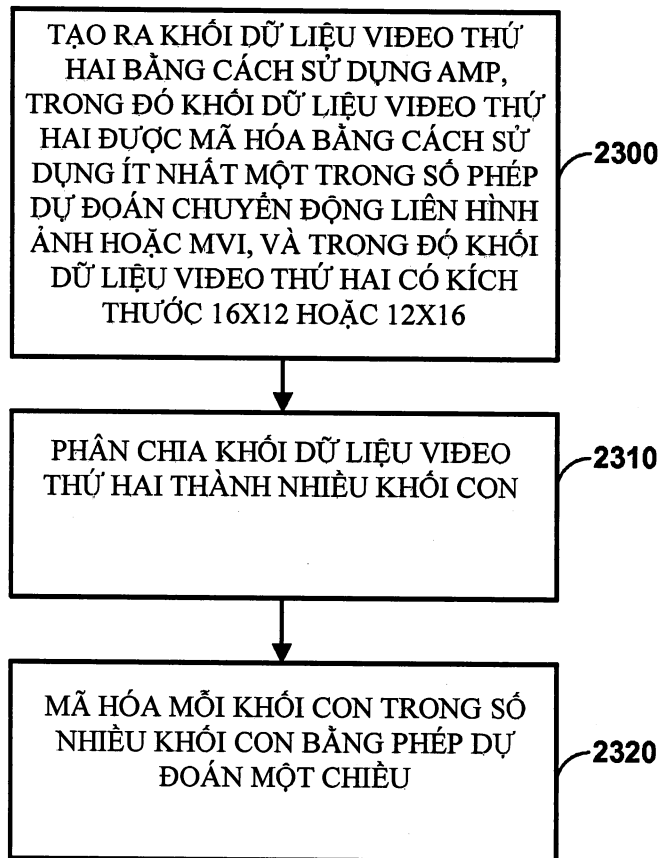


FIG. 23

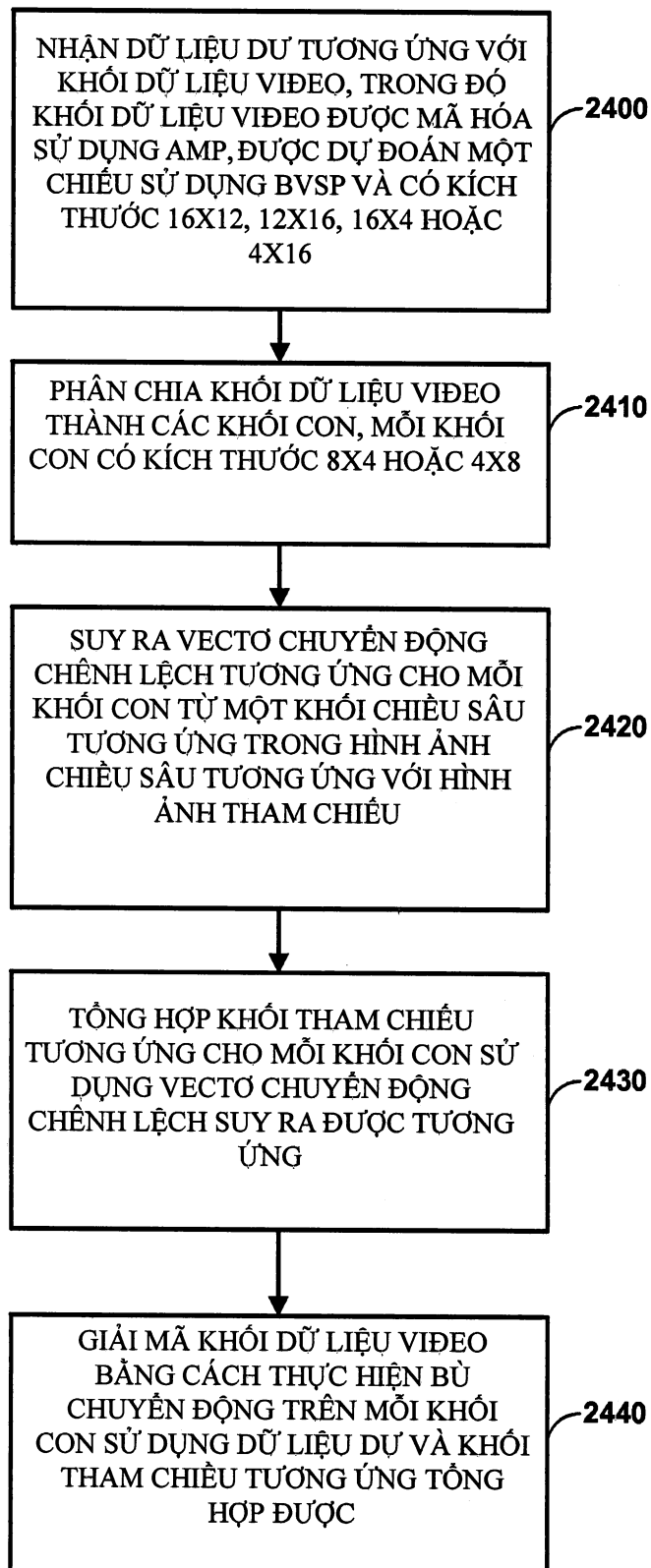


FIG. 24

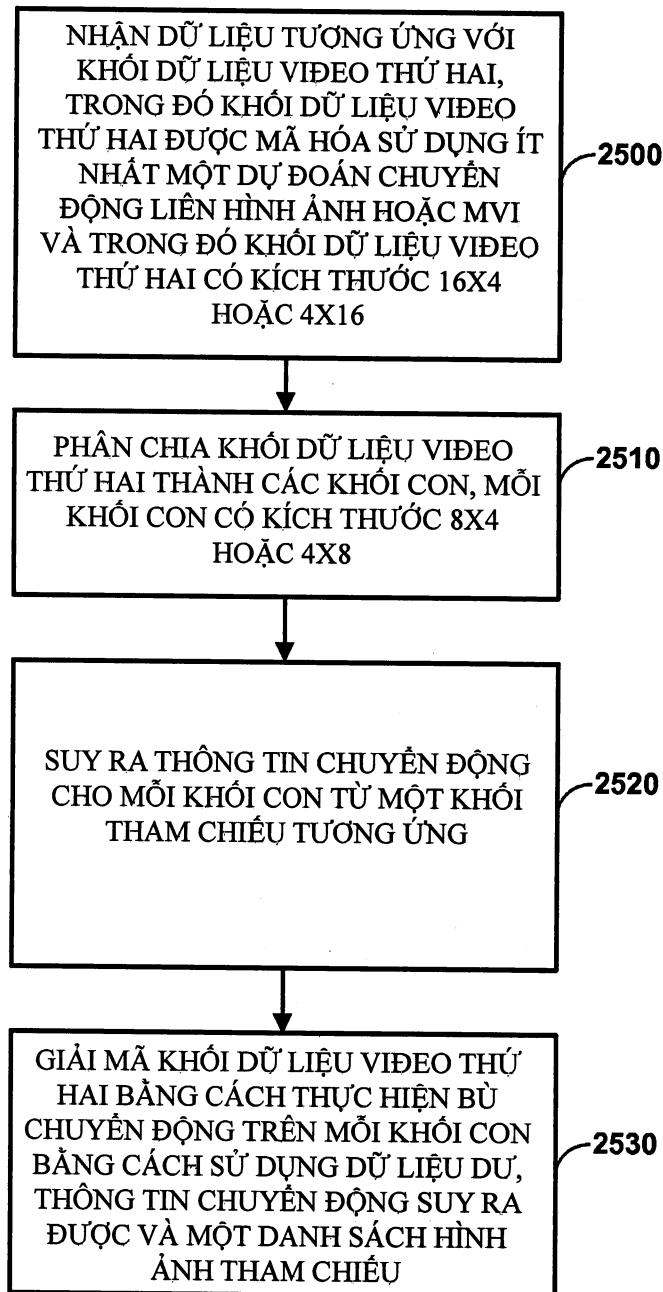


FIG. 25

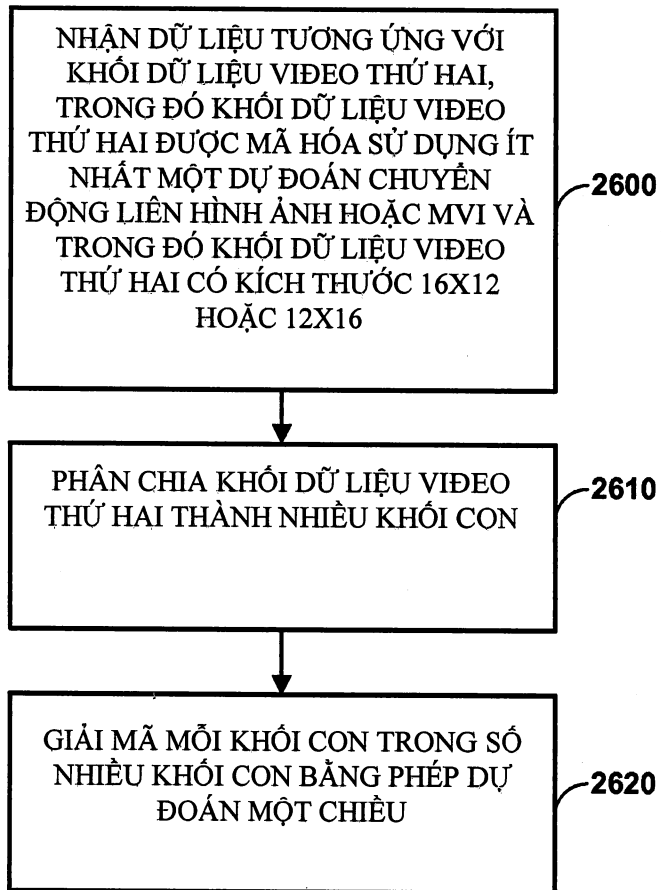


FIG. 26