



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

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

(11)



1-0042535

(51)⁷ H04N 19/48

(13) B

(21) 1-2017-00962

(22) 20/08/2014

(86) PCT/EP2014/067757 20/08/2014

(87) WO 2016/026526 25/02/2016

(45) 27/01/2025 442

(43) 25/09/2017 354A

(73) FRAUNHOFER-GESELLSCHAFT ZUR FOERDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (DE)

Hansastrasse 27c, 80686 Muenchen, Germany

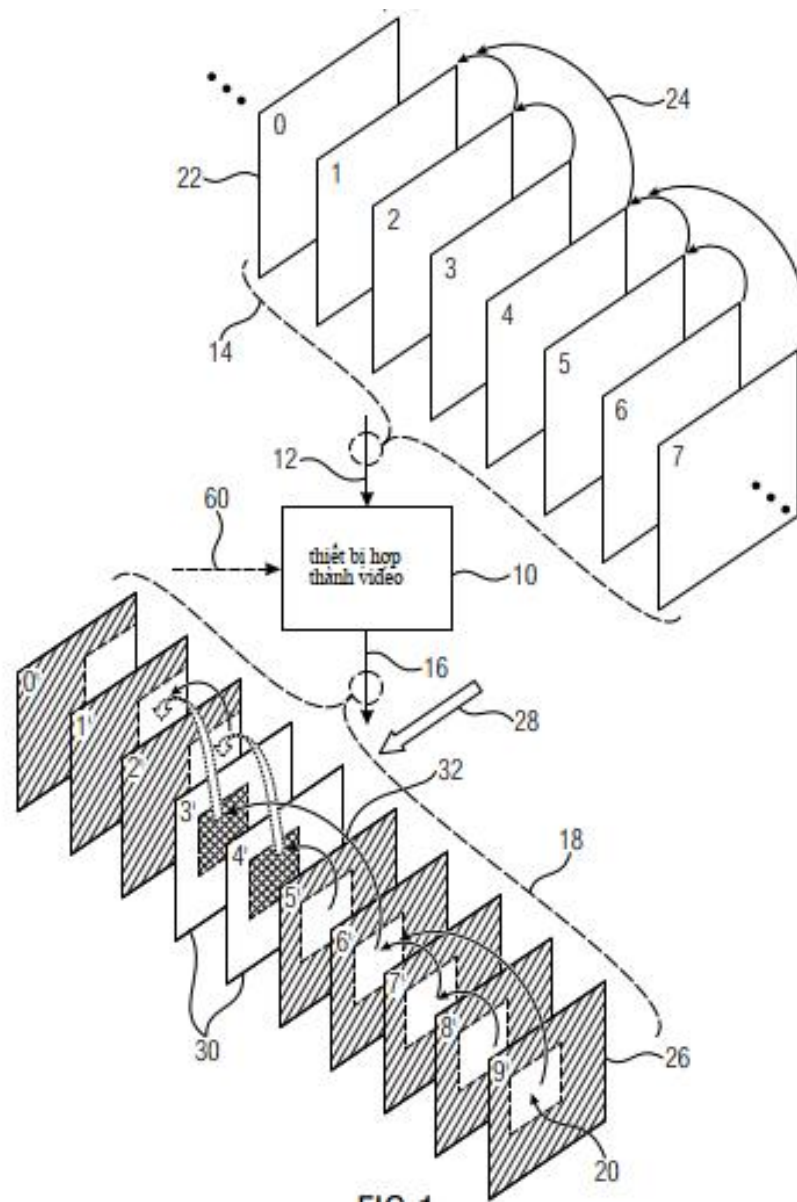
(72) SANCHEZ DE LA FUENTE, Yago (ES); SKUPIN, Robert (DE); SCHIERL,
Thomas (DE).

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

(54) HỆ THỐNG HỢP THÀNH VIDEO

(21) 1-2017-00962

(57) Sáng chế đề cập đến thiết bị hợp thành video, hệ thống của thiết bị hợp thành video và phương pháp hợp thành video. Việc hợp thành video được thực hiện bằng cách điền đầy một phần của dòng dữ liệu video được hợp thành bằng cách thu thập và sao chép, và một phần khác được thêm theo cách tổng hợp bằng cách tham chiếu, qua phép dự đoán theo thời gian được bù chuyển động, phần khác của dòng dữ liệu video được hợp thành, thu được bằng cách thu thập và sao chép.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến việc hợp thành video, như việc hợp thành dòng dữ liệu video được hợp thành bằng cách sử dụng một hoặc nhiều dòng dữ liệu video đầu vào.

Tình trạng kỹ thuật của sáng chế

Có nhiều ứng dụng và trường hợp sử dụng như hội thảo video, giám sát video, ứng dụng y khoa, truyền theo dòng toàn cảnh, chèn quảng cáo, hiển thị ảnh trên ảnh hoặc chồng lấp video trong đó một số dòng bit video dành riêng được giải mã đồng thời và được hiển thị cho người dùng ở dạng được hợp thành. Ví dụ minh họa cho các ứng dụng như vậy là hệ thống giám sát giao thông với nhiều nguồn video được hiển thị cho người dùng. Vấn đề đối với các ứng dụng như vậy là nhiều thiết bị chỉ kết hợp bộ giải mã video phần cứng hoặc cách khác giới hạn ở tài nguyên tính toán, công suất và/hoặc các tài nguyên khác. Các ví dụ của các thiết bị như vậy là các đầu thu giải mã tín hiệu (Set-Top-Box - STB), các bộ TV giá rẻ hoặc các thiết bị di động vận hành bằng pin.

Để cho phép các ứng dụng và các trường hợp sử dụng cho các thiết bị như vậy, dòng bit video đơn tích hợp một số dòng bit video dành riêng phải được tạo ra trước. Để đạt được dòng bit đơn như vậy, việc xử lý video miền điểm ảnh (ví dụ, hợp thành như vát, ghép hoặc trộn) thường được áp dụng, trong đó các dòng bit video khác nhau được chuyển mã thành dòng bit đơn. Việc chuyển mã có thể được thực hiện bằng cách sử dụng bộ giải mã và bộ mã hóa video theo tầng, dẫn đến việc giải mã các dòng bit đến, hợp thành video mới từ các dòng bit đầu vào trong miền điểm ảnh và mã hóa video mới vào dòng bit đơn. Phương pháp này cũng có thể được gọi là chuyển mã đầy đủ truyền thống bao gồm xử lý trong miền không nén. Tuy nhiên, phương pháp này có một số hạn chế. Thứ nhất, việc mã hóa lặp lại thông tin video có thể dẫn đến suy giảm chất lượng tín hiệu hơn nữa do mã hóa các xáo ảnh. Thứ hai và quan trọng hơn, việc chuyển mã đầy đủ giải mã và mã hóa nhiều lần các dòng bit video đi vào và đi ra phức tạp về tính toán và do đó không định tỉ lệ tốt.

Do đó, một phương pháp khác được trình bày trong tài liệu [1], trong đó phương pháp vát video được thực hiện trong miền được nén. Ý tưởng chính trong [1] là thiết lập

các ràng buộc tại các bộ mã hóa, ví dụ, không cho phép vector chuyển động nào đó cũng như phép dự đoán vector chuyển động tại các đường biên ảnh, cho phép dòng bit ít phức tạp viết lại quy trình có thể được áp dụng cho các dòng bit khác để tạo ra dòng bit đơn chứa tất cả các vector dự định sẽ được trộn. Phương pháp này cũng ít phức tạp về tính toán hơn so với phương pháp chuyển mã đầy đủ và không làm suy giảm chất lượng tín hiệu.

Ví dụ minh họa cho hệ thống như vậy được thể hiện trên Fig.23 cho hệ thống giám sát video sử dụng cơ sở hạ tầng máy chủ đám mây. Như có thể thấy, nhiều dòng bit video 900a-d được gửi bởi các bộ gửi 902a-d khác nhau và được vá trong bộ trộn đám mây 904 để tạo ra dòng bit video đơn 906.

Mô tả chi tiết hơn về các kỹ thuật của quy trình và được áp dụng có thể tìm thấy trong tài liệu [1].

Việc xử lý miền được nén có thể được áp dụng cho nhiều ứng dụng và các trường hợp sử dụng để cho phép xử lý video ít phức tạp, tiết kiệm tuổi thọ pin và/hoặc chi phí thực hiện. Tuy nhiên, các đặc điểm của mỗi ứng dụng lại có các vấn đề riêng khi xử lý video miền được nén. Tương tự các đặc điểm và đặc trưng của tiêu chuẩn/sơ đồ nén video có thể được sử dụng để cho phép xử lý miền được nén ít phức tạp cho các ứng dụng mới.

Các vấn đề không được nêu ra đầy đủ bằng sơ đồ và miền được mã hóa trên Fig.23 xảy ra, ví dụ, nếu cách hợp thành dòng bit video đơn 906 trong số các dòng bit video đi vào 900a-d sẽ phải chịu các thay đổi như, ví dụ, sự sắp xếp lại các dòng bit video đi vào nằm trong dòng bit video được hợp thành 906, sự chuyển dịch không gian của dòng bit video đầu vào nhất định nằm trong khu vực ảnh hoặc tương tự của dòng bit video được hợp thành 906. Đối với các trường hợp này, sơ đồ cấu thành trên Fig.23 không hoạt động một cách chính xác do phép dự đoán bù chuyển động theo thời gian kết nối các ảnh đơn lẻ của các dòng bit video đi vào từ 900a đến 900d với nhau theo thời gian để, trong sắp xếp lại của dòng bit video đi vào mà không sử dụng đường vòng qua miền được giải mã/không được nén, bị cấm ngoại trừ đối với các điểm truy cập ngẫu nhiên của dòng bit video đi vào được biểu diễn bởi các ảnh bên trong không sử dụng bất kỳ phép dự đoán được bù chuyển động theo thời gian, mà dẫn đến việc tăng tốc độ bit và

các đỉnh bằng thông tức thời không mong muốn. Do đó, không có bất kỳ nỗ lực bổ sung, sự tự do khi biến thiên sự hợp thành dòng bit video đầu ra 906 mà không rời khỏi miền được nén, sẽ bị giới hạn để đơn thuần xảy ra tại các điểm nhất định trong thời gian bằng các điểm truy cập ngẫu nhiên của dòng bit video đi vào không sử dụng bất kỳ sự dự đoán được bù chuyển động theo thời gian nào. Tần số cao của các điểm truy cập ngẫu nhiên như vậy nằm trong các dòng bit video đi vào 900a-900d, tuy nhiên, gồm tốc độ nén thấp hơn do việc thiếu các bộ dự đoán theo thời gian trong các ảnh được dự đoán.

Tài liệu tham khảo

[1] Yago Sánchez de la Fuente, Ralf Globisch, Thomas Schierl, và Thomas Wiegand, "Low Complexity Cloud-video-Mixing Using HEVC", Proceedings of IEEE Consumer Communications and Networking Conference, Las Vegas, NV, USA, Tháng 1 năm 2014.

[2] Sullivan, Gary J., và cộng sự "Overview of the high efficiency video coding (HEVC) standard." Circuits and Systems for Video Technology, IEEE Transactions ngày 22.12 (2012): 1649-1668.

Bản chất kỹ thuật của sáng chế

Phương án có thể có thiết bị hợp thành video được tạo cấu hình để hợp thành dòng dữ liệu video được hợp thành nhờ sử dụng ít nhất một dòng dữ liệu video đầu vào, dòng dữ liệu video được hợp thành và ít nhất một dòng dữ liệu video đầu vào được mã hóa nhờ sử dụng phép dự đoán theo thời gian được bù chuyển động, thiết bị hợp thành video được tạo cấu hình để hợp thành dòng dữ liệu video được hợp thành bằng cách tạo ra chuỗi các ảnh kế thừa của dòng dữ liệu video được hợp thành với việc điền đầy phần không gian của các ảnh kế thừa bằng cách thu thập và sao chép thông tin phụ bù chuyển động và dữ liệu phần dư dự đoán của ít nhất một dòng dữ liệu video đầu vào, trong đó vị trí không gian của phần không gian thay đổi theo thời gian từ một ảnh kế thừa đến ảnh tiếp theo trong chuỗi các ảnh kế thừa, và chèn ít nhất một ảnh không xuất ra vào dòng dữ liệu video được hợp thành giữa các ảnh kế thừa của chuỗi các ảnh kế thừa, ít nhất một ảnh không xuất ra tham chiếu, qua phép dự đoán theo thời gian được bù chuyển động, tập con của các ảnh kế thừa.

Phương án khác có thể có thiết bị hợp thành video được tạo cấu hình để hợp

thành dòng dữ liệu video được hợp thành trong số nhiều dòng dữ liệu video đầu vào, dòng video được hợp thành và nhiều dòng dữ liệu video đầu vào được mã hóa bằng cách sử dụng phép dự đoán theo thời gian được bù chuyển động, thiết bị hợp thành video được tạo cấu hình để hợp thành dòng dữ liệu video được hợp thành bằng cách ghép kênh nhiều dòng dữ liệu video đầu vào vào phần thứ nhất gồm các ảnh của dòng dữ liệu video được hợp thành bằng cách thu thập và sao chép thông tin phụ bù chuyển động và dữ liệu phần dư dự đoán của nhiều dòng dữ liệu video đầu vào vào phần thứ nhất của dòng dữ liệu video được hợp thành, và điền đầy phần thứ hai gồm các ảnh của dòng dữ liệu video được hợp thành với video đối chiếu bằng cách tham chiếu, qua phép dự đoán theo thời gian được bù chuyển động, phần thứ nhất gồm dòng dữ liệu video được hợp thành, trong đó phần thứ nhất được báo hiệu để không được xuất ra và phần thứ hai được báo hiệu để được xuất ra.

Theo phương án khác, hệ thống có thể có ít nhất một bộ mã hóa video được tạo cấu hình để cung cấp ít nhất một dòng dữ liệu video đầu vào; và thiết bị hợp thành video theo sáng chế được tạo cấu hình để hợp thành dòng dữ liệu video được hợp thành bằng cách sử dụng ít nhất một dòng dữ liệu video đầu vào, dòng dữ liệu video được hợp thành và ít nhất một dòng dữ liệu video đầu vào được mã hóa bằng cách sử dụng phép dự đoán theo thời gian được bù chuyển động.

Theo phương án khác, hệ thống có thể có nhiều bộ mã hóa video, mỗi bộ mã hóa được tạo cấu hình để cung cấp dòng dữ liệu tương ứng trong số nhiều dòng dữ liệu video đầu vào; và thiết bị hợp thành video theo sáng chế được tạo cấu hình để hợp thành dòng dữ liệu video được hợp thành trong số nhiều dòng dữ liệu video đầu vào, dòng dữ liệu video được hợp thành và nhiều dòng dữ liệu video đầu vào được mã hóa bằng cách sử dụng phép dự đoán theo thời gian được bù chuyển động.

Theo phương án khác, phương pháp hợp thành video để hợp thành dòng dữ liệu video được hợp thành bằng cách sử dụng ít nhất một dòng dữ liệu video đầu vào, dòng dữ liệu video được hợp thành và ít nhất một dòng dữ liệu video đầu vào được mã hóa bằng cách sử dụng phép dự đoán theo thời gian được bù chuyển động, phương pháp hợp thành video có thể bao gồm các bước: tạo ra chuỗi các ảnh kế thừa của dòng dữ liệu video được hợp thành với việc điền đầy phần không gian của các ảnh kế thừa bằng cách thu thập và sao chép thông tin phụ bù chuyển động và dữ liệu phần dư dự đoán

của ít nhất một dòng dữ liệu video đầu vào, trong đó vị trí không gian của phần không gian thay đổi theo thời gian từ một ảnh kế thừa đến ảnh tiếp theo trong chuỗi các ảnh kế thừa, và chèn các ảnh không xuất ra vào dòng dữ liệu video được hợp thành giữa các ảnh kế thừa gồm chuỗi các ảnh kế thừa, các ảnh không xuất ra tham chiếu, qua phép dự đoán theo thời gian được bù chuyển động, tập con gồm các ảnh kế thừa.

Theo một phương án khác, phương pháp hợp thành video để hợp thành dòng dữ liệu video được hợp thành trong số nhiều dòng dữ liệu video đầu vào, dòng dữ liệu video được hợp thành và nhiều dòng dữ liệu video đầu vào được mã hóa bằng cách sử dụng phép dự đoán theo thời gian được bù chuyển động, phương pháp hợp thành video có thể bao gồm các bước: ghép kênh nhiều dòng dữ liệu video đầu vào vào phần thứ nhất gồm các ảnh của dòng dữ liệu video được hợp thành bằng cách thu thập và sao chép thông tin phụ bù chuyển động và dữ liệu phần dư dự đoán của nhiều dòng dữ liệu video đầu vào vào phần thứ nhất của dòng dữ liệu video được hợp thành, và điền đầy phần thứ hai gồm các ảnh của dòng dữ liệu video được hợp thành với video đối chiếu bằng cách tham chiếu, qua phép dự đoán theo thời gian được bù chuyển động, phần thứ nhất gồm dòng dữ liệu video được hợp thành, trong đó phần thứ nhất được báo hiệu để không được xuất ra và phần thứ hai được báo hiệu để được xuất ra.

Một phương án khác có thể có dòng dữ liệu video được mã hóa theo thời gian theo thứ bậc được mã hóa bằng cách sử dụng phép dự đoán theo thời gian được bù chuyển động và phép dự đoán thông tin phụ bù chuyển động, trong đó dòng dữ liệu video được mã hóa theo thời gian theo thứ bậc báo hiệu ID lớp thứ bậc theo thời gian phạm vi ảnh đối với mỗi ảnh của dòng dữ liệu video được mã hóa theo thời gian theo thứ bậc và bao gồm cú pháp mức cao với bộ phân biệt ID lớp thứ bậc theo thời gian vượt qua phạm vi ảnh đảm bảo cho một hoặc nhiều ID lớp thứ bậc theo thời gian rằng các ảnh của một hoặc nhiều ID mức thứ bậc theo thời gian tương ứng không được tham chiếu, qua phép dự đoán thông tin phụ bù chuyển động theo thời gian, bằng bất kỳ ảnh đầu vào khác.

Phương án khác có thể có dòng dữ liệu video được mã hóa bằng cách sử dụng phép dự đoán theo thời gian được bù chuyển động và phép dự đoán thông tin phụ bù chuyển động, trong đó dòng dữ liệu video phân biệt, bằng cách báo hiệu theo ảnh, các ảnh thứ nhất mà tại đó không có tập hợp ảnh tham chiếu hiện thời nào được tham chiếu, qua phép dự đoán thông tin phụ bù chuyển động theo thời gian, bằng bất kỳ ảnh đứng

sau nào của dòng dữ liệu video, và các ảnh thứ hai mà tại đó một trong số tập hợp ảnh tham chiếu hiện thời - hoặc là ứng viên đề - được tham chiếu, qua phép dự đoán thông tin phụ bù chuyển động theo thời gian, bằng bất kỳ ảnh đứng sau nào của dòng dữ liệu video.

Phương án khác có thể có phương tiện lưu trữ số bất biến có chương trình máy tính được lưu trữ trên đó để thực hiện phương pháp hợp thành video để hợp thành dòng dữ liệu video được hợp thành bằng cách sử dụng ít nhất một dòng dữ liệu video đầu vào, dòng dữ liệu video được hợp thành và ít nhất một dòng dữ liệu video đầu vào được mã hóa bằng cách sử dụng phép dự đoán theo thời gian được bù chuyển động, phương pháp hợp thành video bao gồm các bước: tạo ra chuỗi các ảnh kế thừa của dòng dữ liệu video được hợp thành với việc điền đầy phần không gian của các ảnh kế thừa bằng cách thu thập và sao chép thông tin phụ bù chuyển động và dữ liệu phần dư dự đoán của ít nhất một dòng dữ liệu video đầu vào, trong đó vị trí không gian của phần không gian thay đổi theo thời gian từ một ảnh kế thừa đến ảnh tiếp theo trong chuỗi các ảnh kế thừa, và chèn các ảnh không xuất ra vào dòng dữ liệu video được hợp thành giữa các ảnh kế thừa của chuỗi các ảnh kế thừa, các ảnh không xuất ra tham chiếu, qua phép dự đoán theo thời gian được bù chuyển động, tập con gồm các ảnh kế thừa, khi chương trình máy tính nêu trên được chạy bởi máy tính.

Phương án khác có thể có phương tiện lưu trữ số bất biến có chương trình máy tính được lưu trữ trên đó để thực hiện phương pháp hợp thành video để hợp thành dòng dữ liệu video được hợp thành trong số nhiều dòng dữ liệu video đầu vào, dòng dữ liệu video được hợp thành và nhiều dòng dữ liệu video đầu vào được mã hóa bằng cách sử dụng phép dự đoán theo thời gian được bù chuyển động, phương pháp hợp thành video bao gồm các bước: ghép kênh nhiều dòng dữ liệu video đầu vào vào phần thứ nhất gồm các ảnh của dòng dữ liệu video được hợp thành bằng cách thu thập và sao chép thông tin phụ bù chuyển động và dữ liệu phần dư dự đoán của nhiều dòng dữ liệu video đầu vào vào phần thứ nhất của dòng dữ liệu video được hợp thành, và điền đầy phần thứ hai gồm các ảnh của dòng dữ liệu video được hợp thành với video đối chiếu bằng cách tham chiếu, qua phép dự đoán theo thời gian được bù chuyển động, phần thứ nhất của dòng dữ liệu video được hợp thành, trong đó phần thứ nhất được báo hiệu để không được xuất ra và phần thứ hai được báo hiệu để được xuất ra, khi chương trình máy tính nêu

trên được chạy bởi máy tính.

Theo khía cạnh thứ nhất của sáng chế, dòng dữ liệu video được hợp thành được hợp thành bằng cách sử dụng ít nhất một dòng dữ liệu video đầu vào, trong đó dòng dữ liệu video được hợp thành và ít nhất một dòng dữ liệu video đầu vào được mã hóa sử dụng phép dự đoán thời gian được bù chuyển động, và sự hợp thành video được thực hiện bằng cách tạo ra chuỗi các ảnh kế thừa của dòng dữ liệu video được hợp thành với sự điền đầy phần không gian của các ảnh kế thừa bằng cách thu thập vào sao chép thông tin phụ được bù chuyển động và dữ liệu phần dư dự đoán của ít nhất một dữ liệu video đầu vào, trong đó vị trí không gian của phần không gian thay đổi theo thời gian từ một ảnh kế thừa đến ảnh tiếp theo trong chuỗi các ảnh kế thừa, và chèn ít nhất một ảnh không xuất ra vào dòng dữ liệu video được hợp thành giữa các ảnh kế thừa của chuỗi các ảnh kế thừa, ít nhất một ảnh không xuất ra tham chiếu, thông qua phép dự đoán theo thời gian được bù chuyển động, tập con các ảnh kế thừa. Theo khía cạnh này, sự thay đổi theo thời gian của vị trí không gian của phần không gian được làm giảm, theo nghĩa độ phức tạp tính toán, bằng cách chèn ít nhất một ảnh không xuất ra giữa các ảnh kế thừa. Bằng cách sử dụng tương tự, tránh được sự cần thiết rời miền được nén và có thể tránh được việc “khuyết tham chiếu” giữa các ảnh của ít nhất một dòng video đầu vào với ít nhất một ảnh không xuất ra, sử dụng cái mà có khả năng để “ghi lại” các ảnh của ít nhất một dòng dữ liệu video đầu vào lần lượt trước và sau sự thay đổi vị trí không gian.

Theo khía cạnh khác của sáng chế, khái niệm hợp thành dòng dữ liệu video được hợp thành trong số nhiều dòng dữ liệu video đầu vào bao gồm ghép kênh nhiều dòng dữ liệu video đầu vào vào phần thứ nhất gồm các ảnh của dòng dữ liệu video được hợp thành bằng cách thu thập và sao chép thông tin phụ bù chuyển động và dữ liệu phần dư dự đoán của nhiều dòng dữ liệu video đầu vào vào dòng dữ liệu video được hợp thành và điền đầy phần thứ hai gồm các ảnh của dòng dữ liệu video được hợp thành với video đối chiếu bằng cách tham chiếu, thông qua phép dự đoán theo thời gian được bù chuyển động, phần thứ nhất gồm dòng dữ liệu video được hợp thành, trong đó phần thứ nhất được báo hiệu để không được xuất ra và phần thứ hai được báo hiệu để được xuất ra. Phù hợp với ý tưởng của khía cạnh thứ nhất, sự hợp thành video theo khía cạnh thứ hai của sáng chế phân chia giữa phần của dòng dữ liệu video được hợp thành được điền đầy

bằng cách thu thập và sao chép, và phần khác được thêm theo cách tổng hợp bằng cách tham chiếu, thông qua phép dự đoán thời gian được bù chuyển động, phần khác của dòng dữ liệu video được hợp thành thu được bằng cách thu thập và sao chép. Tuy nhiên, theo khía cạnh thứ hai, phần thứ hai tạo ra video được thu thập thực sự và được báo hiệu để được xuất ra. Phần thứ nhất đơn giản biểu diễn loại kho dự trữ nội dung ảnh tham chiếu mà trên cơ sở đó nội dung ảnh có thể được hợp thành theo cách chọn lọc bằng cách tham chiếu các phần con nhất định của phần thứ nhất của dòng dữ liệu video được hợp thành bằng cách sử dụng phép dự đoán theo thời gian được bù chuyển động.

Vì vậy, sự hợp thành video được thực hiện trên cơ sở của một hoặc nhiều dòng dữ liệu video đầu vào dẫn đến dòng dữ liệu video được hợp thành có thể được giải mã bởi bộ giải mã video tiêu chuẩn mà không cần phải có khả năng sắp xếp lại dòng dữ liệu video đi vào qua khu vực trình diễn video hoặc không cần phải có khả năng giải mã nhiều dòng dữ liệu video đầu vào theo cách song song. Hơn nữa, sự hợp thành video giả định tác vụ chuyển mã, và sự chuyển mã này tránh được việc rời khỏi miền được nén/được mã hóa so với một hoặc nhiều dòng dữ liệu video đầu vào bằng cách giữ các nỗ lực tính toán ở lượng hợp lý.

Mô tả vắn tắt các hình vẽ

Các khía cạnh có lợi của sáng chế theo các điểm phụ thuộc và các phương án ưu tiên của sáng chế được mô tả dưới đây ở các hình vẽ, trong đó:

Fig.1 thể hiện sơ đồ khối dạng giản lược của thiết bị hợp thành video cùng với dạng biểu diễn giản lược của ít nhất một dòng dữ liệu video đầu vào và dòng dữ liệu video được hợp thành theo phương án;

Fig.2 thể hiện sơ đồ khối dạng giản lược của ảnh kế thừa cùng với ảnh của dòng dữ liệu video đi vào để minh họa quy trình điền đầy phần không gian của ảnh kế thừa bằng cách thu thập và sao chép thông tin phụ bù chuyển động và dữ liệu phần dư dự đoán của ảnh của dòng dữ liệu video đầu vào theo phương án;

Fig.3 thể hiện sơ đồ khối dạng giản lược minh họa ảnh không xuất ra được chèn cùng với ảnh kế thừa mang trong phần không gian của nó ảnh tham chiếu, các phần thể ảnh không xuất ra, và dòng dữ liệu video được hợp thành để minh họa khả năng xây dựng/mã hóa phần không gian của ảnh không xuất ra theo các phương án của sáng chế;

Fig.4 thể hiện sơ đồ khối của ảnh kế thừa tham chiếu, từ phần không gian của nó, qua đường vòng của ảnh không xuất ra của dòng dữ liệu video được hợp thành, phần không gian của ảnh tham chiếu trước;

Fig.5 thể hiện sơ đồ giản lược của hệ thống sử dụng thiết bị hợp thành video và còn bao gồm một hoặc nhiều bộ mã hóa video 70, hệ thống được minh họa khi xuất ra dòng dữ liệu video được hợp thành đến bộ giải mã;

Fig.6 thể hiện ứng dụng truyền theo dòng toàn cảnh;

Fig.7 thể hiện sự tham chiếu theo thời gian thất bại do thay đổi vị trí;

Fig.8 thể hiện sự minh họa của khái niệm ảnh tham chiếu được tạo ra;

Fig.9 thể hiện GRP trong bối cảnh hội nghị video;

Fig.10 thể hiện toàn cảnh 360 độ từ phương tiện được điều khiển từ xa;

Fig.11 thể hiện sơ đồ khối dạng giản lược của thiết bị hợp thành video cùng với dạng biểu diễn giản lược của nhiều dòng dữ liệu video đầu vào và dòng dữ liệu video được hợp thành theo phương án;

Fig.12 thể hiện sơ đồ dạng giản lược của SLCP của phần video được đối chiếu của dòng dữ liệu video được hợp thành cùng với mũi tên nét đứt minh họa các vectơ chuyển động/dự đoán sao chép từ phần tham chiếu của dòng dữ liệu video được hợp thành;

Fig.13 thể hiện sự minh họa của phương thức tham chiếu liên tục;

Fig.14 thể hiện sự minh họa của phương thức tham chiếu được vá;

Fig.15 thể hiện sơ đồ khối của SLCP với các tham chiếu được vá của các chiều không gian khác nhau bằng cách sử dụng phương pháp không xuất ra;

Fig.16 thể hiện sự minh họa việc sử dụng giả để lấy mẫu lại theo thời gian;

Fig.17 thể hiện sự minh họa của SLCP với các tham chiếu B2B;

Fig.18a và Fig.18b thể hiện sự minh họa của SLCP với các tham chiếu được vá;

Fig.19 thể hiện SLCP với kỹ thuật pha trộn alpha qua dự đoán được gán trọng số;

Fig.20 thể hiện SLCP sử dụng phép lấy mẫu lại qua các vectơ dự đoán biến thiên;

Fig.21 thể hiện ví dụ của SLCP đa bước để lấy mẫu lại lặp lại;

Fig.22 mô tả minh họa việc báo hiệu tương ứng dưới dạng bản tin SEI; và

Fig.23 thể hiện cấu trúc hệ thống diễn hình để trộn video, ở đây thể hiện dưới dạng minh họa để giám sát video giao thông.

Mô tả chi tiết sáng chế

Fig.1 thể hiện thiết bị hợp thành video theo một phương án của sáng chế. Thiết bị hợp thành video trên Fig.1 thực hiện hợp thành video theo khía cạnh của sáng chế mà theo đó các hình ảnh không xuất ra được chèn vào dòng dữ liệu video được hợp thành, do đó cho phép sự thay đổi theo thời gian trong vị trí không gian của phần không gian, phần này được điền đầy bằng cách thu thập và sao chép thông tin phụ được bù chuyển động và dữ liệu phần dư dự đoán của dữ liệu video đầu vào nhất định, bằng việc chèn các ảnh không xuất ra thay thế các ảnh tham chiếu hiển thị vị trí không gian của phần không gian trước sự thay đổi theo thời gian với các ảnh tham chiếu hiển thị vị trí không gian của phần không gian sau sự thay đổi theo thời gian trong bộ đệm ảnh được giải mã của bộ giải mã thực hiện giải mã dòng dữ liệu video được hợp thành.

Thiết bị hợp thành video trên Fig.1 được biểu thị chung sử dụng ký hiệu tham chiếu 10 và bao gồm đầu vào 12 để nhận dòng dữ liệu video đầu vào 14 và đầu ra 16 để xuất ra dòng dữ liệu video được hợp thành 18. Dòng dữ liệu video được hợp thành 18 và dòng dữ liệu video đầu vào 14 được mã hóa bằng cách sử dụng phép dự đoán theo thời gian được bù chuyển động. Các dòng dữ liệu video này có thể, ví dụ, được mã hóa nhờ sử dụng kỹ thuật HEVC, VP9 hoặc nhờ sử dụng sơ đồ mã hóa video khác nào đó. Như sẽ được nêu chi tiết dưới đây, thiết bị hợp thành video 10 có khả năng tạo ra hoặc hợp thành dòng dữ liệu video được hợp thành 18 với nỗ lực tính toán thấp và duy trì trong miền được nén/mã hóa miễn là dòng dữ liệu video đầu vào 14 có liên quan.

Bằng cách được nêu chi tiết hơn dưới đây, thiết bị hợp thành video trên Fig.1 có khả năng tạo ra dòng dữ liệu video được hợp thành 18 sao cho khi giải mã dòng dữ liệu video được hợp thành, video được biểu diễn bởi dòng dữ liệu video đầu vào 14 được hiển thị tại phần không gian 20, thay đổi theo thời gian vị trí không gian của mình trong các ảnh của dòng dữ liệu video được hợp thành 18. Ví dụ, các ảnh của dòng dữ liệu video đầu vào 14 có kích thước thứ nhất, như rộng $n \times m$ mẫu, trong khi các ảnh của dòng dữ liệu video được hợp thành 18 có kích thước thứ hai lớn hơn kích thước thứ

nhất, như rộng $N \times M$ mẫu, với $n < N$ và/hoặc $m < M$, với $n < N$ và $m < M$ được minh họa trên Fig.1. Phần không gian 20 mà nội dung của dòng dữ liệu video đầu vào 14 nằm trong đó, theo ví dụ trên Fig.1, có kích thước thứ nhất, tức là, rộng $n \times m$, mặc dù ví dụ trên Fig.1 có thể được sửa đổi ở điểm thiết bị hợp thành video 10 mô tả trong phần không gian 20 của dòng dữ liệu video được hợp thành 18 đơn thuần là phần phụ không gian cố định nhất định của dòng dữ liệu video đầu vào 14.

Để dễ dàng hiểu chế độ vận hành của thiết bị hợp thành video 10 trên Fig.1, Fig.1 thể hiện phần làm ví dụ của chuỗi các ảnh 22 của dòng dữ liệu video đầu vào 14. Các ảnh 22 được thể hiện minh họa trên Fig.1 được đánh số từ 0 đến 7 phù hợp với thứ tự mã hóa của chúng trong dòng dữ liệu 14 và sử dụng các mũi tên nét liền 24. Fig.1 minh họa sự phụ thuộc lẫn nhau giữa các ảnh 22 bằng phép dự đoán theo thời gian được bù chuyển động, tức là, ảnh 22 nào tham chiếu ảnh nào qua phép dự đoán theo thời gian được bù chuyển động. Tức là, bất kỳ ảnh 22 nào mà mũi tên 24 chỉ đến, biểu diễn ảnh tham chiếu của ảnh mà từ đó mũi tên tương ứng bắt đầu. Ảnh 22 với số 1, ví dụ, biểu diễn ảnh tham chiếu của các ảnh 22 với các số 2 và 4. Phép dự đoán theo thời gian được bù chuyển động được điều khiển nằm trong dòng dữ liệu video đầu vào 14 bằng thông tin phụ bù chuyển động tương ứng thêm vào đó dòng dữ liệu video đầu vào 14 bao gồm dữ liệu phần dư dự đoán để hiệu chỉnh phép dự đoán theo phép dự đoán theo thời gian được bù chuyển động. Ví dụ, thông tin phụ được bù chuyển động có thể bao gồm các vector chuyển động cho các khối của ảnh tham chiếu tương ứng, và dữ liệu phần dư dự đoán có thể cũng được mã hóa theo khối bằng cách sử dụng, ví dụ, mã hóa biến đổi, ví dụ, sử dụng phép biến đổi phân tích phổ, chẳng hạn như DCT hoặc tương tự, để dữ liệu phần dư dự đoán có thể, ví dụ, bao gồm các hệ số biến đổi. Thông tin phụ bù chuyển động và dữ liệu phần dư dự đoán có trong dòng dữ liệu video đầu vào 14 cho mỗi ảnh 22 có thể có trong dòng dữ liệu dưới dạng, ví dụ, được mã hóa entropy bằng cách sử dụng, ví dụ, mã hóa độ dài biến thiên và/hoặc mã hóa số học. Cụ thể, thông tin phụ bù chuyển động và dữ liệu dự đoán cho ảnh 22 nhất định có thể có trong dòng dữ liệu 14 dưới dạng của một hoặc nhiều gói của các lát mà mỗi ảnh 22 có thể được chia nhỏ theo không gian, cùng với, ví dụ, thứ tự mã hóa nào đó ngang qua ảnh 22 tương ứng.

Miễn là dòng dữ liệu video được hợp thành 18 có liên quan, Fig.1 minh họa một đoạn của chuỗi ảnh 26 trong số dòng dữ liệu được hợp thành 18 này mà phần tương ứng

theo thời gian của dòng dữ liệu video đầu vào 14 được hiển thị nằm trong đó được tạo nên bởi các ảnh 22 với các số từ 0 đến 7 được thể hiện trên Fig.1. Cho các mục đích minh họa, Fig.1 thể hiện rằng thiết bị hợp thành video 10 thay đổi một lần trong khoảng thời gian của phần không gian 20, mà trong đó, nội dung của dòng dữ liệu video 14 được hiển thị, về vị trí không gian của nó, nhưng nên lưu ý rằng thiết bị hợp thành video 10 không chỉ giới hạn ở việc thay đổi theo thời gian vị trí không gian của phần không gian 20 chỉ một lần, và cũng không liên quan đến toàn bộ dòng dữ liệu video được hợp thành 18 hoặc bất kỳ chuỗi con nào của nó.

Để tránh các vấn đề được nêu tóm lược ở trên, trong việc thay đổi vị trí không gian của phần không gian 20 trong các ảnh 26 của dòng dữ liệu 18, thiết bị hợp thành video 10 hoạt động như sau.

Trước khi thay đổi trong vị trí không gian của phần không gian 20, trong đó thời điểm được biểu thị minh họa bằng mũi tên 28 trên Fig.1, thiết bị hợp thành video 10 tạo ra các ảnh 26 của dòng dữ liệu video được hợp thành 18 đơn giản bằng cách kế thừa dữ liệu từ các ảnh 22 tương ứng của dòng dữ liệu video đầu vào 14. Để chính xác hơn, trong ví dụ trên Fig.1, ba ảnh 22 đầu tiên được thể hiện trên Fig.1 của dòng dữ liệu video đầu vào 14 đứng trước sự thay đổi 28 trong vị trí không gian của phần không gian 20, và theo đó ba ảnh đầu tiên của dòng dữ liệu được hợp thành 18 là "các ảnh kế thừa", có phần không gian 20 được điền đầy bởi thiết bị hợp thành video 10 bằng cách thu thập và sao chép thông tin phụ bù chuyển động và dữ liệu phần dư dự đoán của dòng dữ liệu video đầu vào 14. Để chính xác hơn, phần không gian 20 của ảnh 26 với số 0' trong dòng dữ liệu video được hợp thành 18 được điền đầy, ví dụ, bằng cách thu thập và sao chép thông tin phụ bù chuyển động và dữ liệu phần dư dự đoán của ảnh 22 tương ứng của video 14 có số 0, và tương tự, phần không gian 20 của ảnh 26 với số 1' được điền đầy bằng cách thu thập và sao chép thông tin phụ bù chuyển động và dữ liệu phần dư dự đoán của ảnh 22 tương ứng với số 1 và phần không gian 20 của ảnh 26 với số 2' được điền đầy phù hợp bằng cách sử dụng dữ liệu của ảnh 22 với số 2 tương ứng. Tức là, phần không gian 20 của mỗi ảnh kế thừa 26 được điền đầy bằng cách sử dụng một ảnh liên quan số các ảnh 22 với việc duy trì thứ tự ảnh của các ảnh 22.

Các ảnh 26 của dòng dữ liệu video được hợp thành 18, phần không gian 20 của dòng này sau đó được điền đầy bằng cách thu thập vào sao chép dữ liệu tương ứng của

các ảnh 22 tương ứng của dòng dữ liệu video đầu vào 14, như được mô tả, được gọi là "các ảnh kế thừa" và được biểu thị trên Fig.1 sử dụng tô bóng đơn giản trong một vùng của các ảnh 26 bên ngoài phần không gian 20. Vùng này bao quanh phần không gian 20 có thể được điền đầy bởi thiết bị hợp thành video 10 một cách nhân tạo, ví dụ, có thể đơn giản là đen hoặc có thể thể hiện ảnh tĩnh khác nào đó hoặc tương tự. Trong thiết bị hợp thành video 10, cũng có thể khả thi để thể hiện bên trong vùng bao quanh/còn lại của các ảnh 26 bên ngoài phần không gian 20 video khác được biểu diễn bởi dòng dữ liệu video đầu vào khác. Các phương án được mô tả dưới đây còn biểu diễn ví dụ về việc mở rộng như vậy của thiết bị hợp thành video 10. Như được mô tả tại đó, nhiều hơn một dòng dữ liệu đầu vào với các phần không gian liên quan trong các ảnh của dòng dữ liệu video được hợp thành có thể được xử lý song song. Các phần không gian của dòng dữ liệu đầu vào có thể, ví dụ, cùng thay đổi vị trí của chúng tại thời điểm 28. Chi tiết hơn, tham khảo đến phần mô tả được đưa ra dưới đây.

Thiết bị hợp thành video 10 tiếp tục với việc tạo ra chuỗi các ảnh 26 kế thừa theo sau sự thay đổi vị trí không gian của phần không gian 20 trong các ảnh 26. Theo đó, ảnh 26 với số 5' là ảnh kế thừa, phần không gian 20 của ảnh này được thay thế không gian so với các phần không gian 20 của các ảnh 26 được đề cập trước đó với các số từ 0' đến 2' và phần không gian 20 của nó được điền đầy bằng cách thu thập và sao chép thông tin phụ bù chuyển động và dữ liệu phần dư dự đoán của ảnh 22 với số 3 của dòng dữ liệu video đầu vào 14, trong đó các ảnh 26 tiếp theo với các số 6' đến 9' của dòng dữ liệu video được hợp thành 18 được điền đầy trong phần không gian 20 tương tự cách thu thập và sao chép dữ liệu tương ứng của các ảnh sau đó 22 với các số từ 4 đến 7 của dòng dữ liệu video đầu vào 14.

Do đó, các ảnh 26 với các số 0', 1', 2', 5' đến 9' tạo ra chuỗi các ảnh kế thừa được điền đầy, như được mô tả, bằng cách thu thập và sao chép thông tin phụ bù chuyển động và dữ liệu phần dư dự đoán tương ứng của các ảnh tương ứng 22 của dòng dữ liệu video đầu vào 14 và theo cách khác điền đầy vùng bên ngoài khác phần không gian 20. Tuy nhiên, do sự thay đổi 28 trong vị trí không gian của phần không gian 20, nội dung của dòng dữ liệu video được hợp thành 18 nằm trong phần không gian 20 sẽ không chính xác khi sự có mặt của bất kỳ biện pháp nào theo sau thời điểm 28 do sự di chuyển hoặc dịch chuyển của phần không gian 20 tại thời điểm 28. Ví dụ, xem ảnh 26 với số 5': vì

ảnh 26 này được điền đầy trong phần không gian 20 bằng cách thu thập và sao chép thông tin phụ bù chuyển động và dữ liệu phần dư dự đoán tương ứng của ảnh 22 với số 3 tương ứng được tham chiếu, thông qua phép dự đoán theo thời gian được bù chuyển động, đến ảnh đứng ngay trước mà có thể, khi không có bất kỳ biện pháp nào, là ảnh kế thừa đứng trước 2'. Tuy nhiên, ảnh 26 sau cùng có nội dung ảnh kế thừa từ ảnh tham chiếu của ảnh 22 với số 3, cụ thể là ảnh 22 với số 2, được chèn vào tại vị trí được thay thế, và theo đó thiết bị hợp thành video 10 sửa vấn đề này theo cách sau đây.

Cụ thể, thiết bị hợp thành video 10 trên Fig.1 chèn ảnh không xuất ra 30 vào dòng dữ liệu video được hợp thành 18 giữa các ảnh kế thừa 26, trong đó các ảnh không xuất ra 30 tham chiếu, thông qua phép dự đoán theo thời gian được bù chuyển động, tập con gồm các ảnh kế thừa.

Để minh họa cách để các ảnh không xuất ra giải quyết vấn đề vừa được nêu về việc khuyết tham chiếu, Fig.1 minh họa sự phụ thuộc lẫn nhau giữa các ảnh 26 và 30 của dòng dữ liệu video được hợp thành 18 dưới dạng kết quả của thông tin phụ bù chuyển động được thu thập và sao chép sử dụng các mũi tên nét liền 32, trong khi sự phụ thuộc lẫn nhau được thêm vào bằng các ảnh không xuất ra 30 và thông tin phụ bù chuyển động của chúng được tạo ra một cách nhân tạo bằng thiết bị hợp thành video 10 theo cách được mô tả chi tiết dưới đây được minh họa sử dụng các mũi tên nét đứt 34. Như có thể thấy, thiết bị hợp thành video 10 đã chèn các ảnh không xuất ra 30 tại thời điểm 38, tại đó phần không gian 20 đã thay đổi về không gian vị trí không gian của nó. Ảnh 26 “gây vấn đề” được mô tả trước đó là, ví dụ, ảnh 5'. Tuy nhiên, ảnh phía sau của ảnh đứng ngay trước trong dòng dữ liệu 18 bây giờ là ảnh không xuất ra 30 với số 4'. Như có thể thấy, thiết bị hợp thành video 10 điền đầy phần không gian 20 của ảnh không xuất ra 30 sao cho 1) phần không gian 20 của ảnh không xuất ra 30 được đặt vị trí tại vị trí mới của phần không gian 20 được đặt cùng vị trí với phần không gian 20 của ảnh 5', và sao cho 2) vùng của phần không gian 20 của ảnh không xuất ra 30 với số 4' được điền đầy, qua phép dự đoán được bù chuyển động, bằng cách sao chép nội dung phần không gian 20 của ảnh kế thừa 2'. Ảnh kế thừa sau cùng, cụ thể ảnh 2' là, do đó, ảnh tham chiếu của ảnh 5' - nếu các ảnh không xuất ra 30 không được chèn vào dòng dữ liệu 18. Nói cách khác, ảnh không xuất ra 30 với số 4' sao chép nội dung của phần không gian 20 của ảnh kế thừa 2' vào vị trí mới của phần không gian 20 là hợp lệ theo sau để

thay đổi thời điểm 28, và vì vậy giả định, miễn là phần không gian 20 của nó có liên quan, vai trò của ảnh 2 miễn là vai trò của ảnh 2 là ảnh tham chiếu đối với ảnh 3 có liên quan. Để chính xác hơn, ảnh không xuất ra 4' là ảnh tham chiếu của ảnh 5' như ảnh 2' là ảnh tham chiếu của ảnh 3 bằng cách sử dụng thông tin phụ được bù chuyển động và thông tin phần dư dự đoán mà phần không gian 20 của ảnh kế thừa 5' được điền đầy. Bây giờ, nội dung ảnh tham chiếu cho phần không gian 20 của ảnh kế thừa 5' ở vị trí đúng: nó nằm trong ảnh đứng ngay trước hoặc dòng dữ liệu 18, cụ thể ảnh không xuất ra với số 4', và nó ở ngay vị trí không gian đúng, tức là, tại vị trí tương ứng về không gian chính xác với phần không gian 20 của ảnh 5'.

Như được thể hiện trên Fig.1, thiết bị hợp thành video 10 cũng chèn ảnh không xuất ra 30 tại vị trí của thời điểm 28, thay thế ảnh tham chiếu kế thừa trước đó 1': tức là, thiết bị hợp thành video 10 điền đầy ảnh không xuất ra 30 tại phần không gian 20 cùng vị trí tại phần không gian 20 của các ảnh kế thừa 26 sau thời điểm 28 với thông tin phụ bù chuyển động mà tham chiếu, và nhờ đó sao chép, nội dung ảnh của ảnh kế thừa 1' nằm trong phần không gian 20 của nó. Bằng phương pháp này, ảnh không xuất ra 30 này, gọi là ảnh 3', đóng vai trò chính xác là ảnh tham chiếu cho ảnh kế thừa 6' trong đó nó được điền đầy trong phần không gian 20 với nội dung ảnh của ảnh 22 với số tham chiếu 1, là ảnh tham chiếu trước đó của ảnh 4 bằng cách sử dụng thông tin phụ được bù chuyển động và dữ liệu phần dư dự đoán mà phần không gian 20 của ảnh kế thừa 6' được điền đầy chúng. Nói cách khác, thiết bị hợp thành video 10 trên Fig.1 chèn nhiều ảnh không xuất ra 30 tại thời điểm 28 vì có các ảnh tham chiếu trong số các ảnh 22 nằm trong dòng dữ liệu video 14 đứng trước thời điểm 28, là các ảnh tham chiếu cho các ảnh 22 của dòng dữ liệu video đầu vào 14 sau thời điểm 28. Nói theo cách khác nữa, thiết bị hợp thành video 10, tại thời điểm 28, chèn nhiều ảnh không xuất ra 30 do có các ảnh kế thừa 26 nằm trong dòng dữ liệu 18, mà có thể, khi không chèn, biểu diễn các ảnh tham chiếu cho các ảnh kế thừa 26 của dòng dữ liệu 18 sau thời điểm 28.

Thiết bị hợp thành video 10 có thể, khi thực hiện việc thu thập và sao chép thông tin phụ bù chuyển động và dữ liệu phần dư dự đoán của các ảnh 22 của dòng dữ liệu video đầu vào 14, trả về một cách sơ bộ từ, ví dụ, miễn được mã hóa entropy đến miền mức cú pháp, tức là, thực hiện việc giải mã entropy của thông tin phụ bù chuyển động và dữ liệu phần dư dự đoán của các ảnh 22 của dòng dữ liệu video đầu vào 14 để thu

được các phần tử cú pháp mô tả thông tin phụ bù chuyển động và dữ liệu phần dư dự đoán của các ảnh 22, bằng cách sử dụng thông tin phụ bù chuyển động và dữ liệu phần dư dự đoán được thu thập và được sao chép để điền đầy ảnh kế thừa 26 tương ứng trong phần không gian 20, bằng cách thực hiện phép mã hóa entropy tương tự để tạo ra dòng dữ liệu video được hợp thành. Thông tin phụ bù chuyển động và dữ liệu phần dư dự đoán của các ảnh 22 của dòng dữ liệu video đầu vào 14 có thể, bằng phương pháp này, thậm chí được phân loại lại hoặc được sắp xếp thứ tự lại về không gian để tính đến, ví dụ, thứ tự mã hóa khác khi đi qua phần không gian 20 trong ảnh kế thừa 26 so với thứ tự mã hóa được sử dụng để mã hoá ảnh 22 tương ứng. Tuy nhiên, mặc dù cách thực hiện việc thu thập và sao chép này tránh được sự cần thiết cho thiết bị hợp thành video 10 thực hiện sự tìm kiếm bù chuyển động một lần nữa, thậm chí sự chệch hướng giải mã/mã hóa entropy qua mức phần tử cú pháp có thể được tránh bởi thiết bị hợp thành video 10 theo phương án. Nói cách khác, việc thu thập và sao chép có thể được thực hiện theo cách giữ trong miền được mã hóa entropy. Để minh họa khả năng này, tham chiếu đến Fig.2.

Fig.2 minh họa theo biểu diễn ảnh kế thừa 5' và ảnh đầu vào 22 với số 3 tương ứng, thông tin phụ bù chuyển động và dữ liệu phần dư dự đoán của ảnh này được sử dụng để điền đầy, bằng cách thu thập và sao chép, phần không gian 20 của ảnh 5'. Fig.2 cũng minh họa cách ảnh 22 với số 3 được biểu diễn trong dòng dữ liệu 14: theo sau thứ tự mã hóa được định trước, như thứ tự quét raster 36, nội dung ảnh 22 được mã hóa tuần tự vào dòng dữ liệu 14 theo đơn vị một hoặc nhiều lát 38. Hai lát được minh họa làm ví dụ trên Fig.2. Mỗi lát 38 bao gồm, ví dụ, nằm trong dòng dữ liệu 14 phần đầu lát 40 biểu thị, ví dụ, các lát tương ứng được đặt trong ảnh 22 được biểu thị, ví dụ, so với góc phía trên bên trái của ảnh 22.

Để có thể kế thừa các lát 38 vào dòng dữ liệu video được hợp thành 18 về cơ bản vì chúng không rời khỏi miền mã hóa entropy ít nhất miễn là phần tải tin 42 của chúng có liên quan, mà bao gồm, được mã hóa bằng cách sử dụng mã hóa entropy, thông tin phụ bù chuyển động và dữ liệu phần dư dự đoán, để mã hóa nội dung của ảnh 26 miễn là phần không gian 20 có liên quan, thiết bị hợp thành video chia nhỏ ảnh 26 theo phương án trên Fig.2 thành các ô 44, việc chia nhỏ các ô được minh họa bằng cách sử dụng các đường kẻ - chấm và có sẵn, ví dụ, trong HEVC. Bằng cách chia nhỏ ảnh 26

thành các ô 44, thứ tự mã hóa, sử dụng ảnh 26 nào được mã hóa trong dòng dữ liệu 18, được dẫn hướng để nằm trong phần không gian 20 của ảnh 26, thứ tự mã hóa trùng khớp với thứ tự mã hóa 36 sử dụng ảnh 22 được mã hóa thành các lát 38. Để chính xác hơn, như được mô tả ở trên, ảnh 26 lớn hơn ảnh 22. Phần không gian 20 có cùng kích thước như ảnh 22. Việc phân chia ô của ảnh 26 thành các ô 44 được thực hiện sao cho phần không gian 20 trùng khớp chính xác với một trong các ô 44. Thứ tự mã hóa của ảnh 26 được sửa đổi bởi việc phân chia ô vì thứ tự mã hóa 44 của ảnh 26 đi ngang qua các ô 44 phù hợp với thứ tự ô đi ngang qua ảnh 26 nằm trong một ô trước tiên trước khi tiếp tục với ô tiếp theo trong thứ tự ô. Bằng phương pháp này, thiết bị hợp thành video 10 có thể trực tiếp sao chép các lát 38 vào dòng dữ liệu 18, cụ thể vào khối truy cập 46 thành ảnh 26 được mã hóa trong dòng dữ liệu 18. Các lát 38 này sau đó mã hóa ô 44 tương ứng biểu diễn phần không gian 20.

Thiết bị hợp thành video 10 có thể, khi sử dụng/áp dụng khái niệm trên Fig.2 sửa đổi dữ liệu nào đó trong phần đầu lát 40. Ví dụ, mỗi phần đầu lát 40 có thể bao gồm địa chỉ lát biểu thị vị trí bắt đầu của lát tương ứng (thứ tự mã hóa 36) so với góc phía trên bên trái của ảnh tương ứng. Theo đó, để tính đến vị trí mới của lát nằm trong ảnh kế thừa 26, cụ thể, nằm trong phần không gian 20, thiết bị hợp thành video 10 có thể thay đổi địa chỉ lát để đo vị trí mới của lát bên trong phần không gian 20 so với góc phía trên bên trái của ảnh kế thừa 26.

Ngoài ra hoặc theo cách khác, mỗi phần đầu lát 40 có thể bao gồm dữ liệu đếm thứ tự ảnh. Việc đếm thứ tự ảnh có thể sắp đặt các ảnh trong dòng dữ liệu tương ứng. Như được biểu thị trên đây, vì thiết bị hợp thành video 10 chèn các ảnh không xuất ra 30 vào dòng dữ liệu video được hợp thành 18, thiết bị hợp thành video 10 có thể thay đổi dữ liệu đếm thứ tự ảnh trong phần đầu lát 40 khi điền đầy phần không gian 20 của các ảnh kế thừa 26. Ví dụ, phần đầu lát 40 trên Fig.2 có thể biểu thị 3 là số đếm thứ tự ảnh, trong khi phần đầu lát của các lát 38 nằm trong dòng dữ liệu 18 có thể biểu thị 5 để tính đến việc chèn hai ảnh không xuất ra 30.

Ngoài ra hoặc theo cách khác, phần đầu lát 40 có thể bao gồm các trị số chênh lệch đếm thứ tự ảnh tham chiếu, tức là, các trị số biểu thị ảnh tham chiếu của ảnh 22 hiện thời theo nghĩa tương đối, tức là, ảnh thứ x đứng trước hoặc đứng sau ảnh 22 hiện thời. Trong phương án được mô tả dựa vào Fig.1, sự thay đổi trị số chênh lệch đếm thứ

tự ảnh tham chiếu có thể không cần thiết, vì các ảnh không xuất ra 30 được định vị tại vị trí đúng trong dòng dữ liệu video 18, nhưng để phù hợp với ví dụ khác, thiết bị hợp thành video 10 có thể thay đổi các trị số chênh lệch đếm thứ tự ảnh tham chiếu như vậy trong các phần đầu lát 40.

Hơn nữa, và ngoài ra hoặc cách khác, phần đầu lát 40 có thể bao gồm tham chiếu đến tập hợp các trị số chênh lệch đếm thứ tự ảnh tham chiếu. Sự tham chiếu này có thể, ví dụ, đề cập đến tập hợp tham số được truyền tải trong dòng dữ liệu video đầu vào 14 và sự tham chiếu có thể được thay đổi trong phần đầu lát 40 khi điền đầy phần không gian 20 sử dụng các lát 38. Tương tự, bản thân các tập hợp tham số có thể được chấp nhận từ dòng dữ liệu video đầu vào mà có hoặc không có sửa đổi.

Ngoài ra, có thể khả thi rằng ngoài ra hoặc cách khác, phần đầu lát 40 bao gồm tham số lượng tử hóa được mã hóa vi sai. Tức là, trong dòng dữ liệu 14, ví dụ, tham số lượng tử hóa được mã hóa vi sai của phần đầu lát 40 có thể được mã hóa vi sai so với tham số lượng tử hóa được truyền tải trong dòng dữ liệu 40 cho ảnh 22 hoặc thậm chí cho chuỗi các ảnh gồm có ảnh 22. Trong khi xây dựng dòng dữ liệu video được hợp thành 18, thiết bị hợp thành video 10 có thể cũng chọn tham số lượng tử hóa nhất định cho ảnh 26 hoặc thậm chí chuỗi các ảnh gồm có ảnh 26 nằm trong dòng dữ liệu 18 và nó có thể, ví dụ, khác với tham số lượng tử hóa đóng vai trò làm cơ sở của tham số lượng tử hóa được mã hóa vi sai nằm trong phần đầu lát 40. Theo đó, phần đầu lát 40 của các lát 38 như được truyền thành dòng dữ liệu 18 có khả năng biến thiên bởi thiết bị hợp thành video 10 để tính đến sự thay đổi trong các tham số lượng tử hóa tham chiếu được báo hiệu ở phần khác trong dòng dữ liệu 18 bằng thiết bị 10.

Fig.3 minh họa khả năng về cách thiết bị hợp thành video 10 sẽ tạo ra một cách nhân tạo các ảnh không xuất ra. Cụ thể, Fig.3 thể hiện ảnh không xuất ra 30 có số 4' và ảnh kế thừa 26 có số 2', tức là, ảnh kế thừa 26 thay thế bởi ảnh không xuất ra 30 và tham chiếu qua phép dự đoán theo thời gian được bù chuyển động. Cụ thể, Fig.3 minh họa trường hợp trong đó các dòng dữ liệu 14 và 18 được mã hóa bằng cách sử dụng phép dự đoán theo thời gian được bù chuyển động theo cách dựa vào khối. Tức là, các ảnh tương ứng được phân chia thành các khối, một số khối này được dự đoán bằng cách sử dụng phép dự đoán theo thời gian được bù chuyển động, và đối với mỗi khối trong số các khối này, vectơ chuyển động riêng được biểu thị bởi thông tin phụ bù chuyển động.

Thiết bị 10 xác định thông tin phụ bù chuyển động cho phần không gian 20 của ảnh không xuất ra 30 sao cho phần 20 của ảnh không xuất ra 30, bằng cách tịnh tiến, được sao chép từ phần không gian 20 của ảnh kế thừa 26 với cả hai phần 20 có cùng kích thước. Tức là, mỗi mẫu của phần 20 của ảnh không xuất ra 30 được sao chép từ mẫu tương ứng nằm trong ảnh 26 được thay thế bằng cách sử dụng một và cùng vector chuyển động 50 tương đối về vị trí với ảnh 26 được cùng định vị đến mẫu tương ứng của phần 20 của ảnh không xuất ra 30. Tuy nhiên, vì vector chuyển động 50 là giống nhau đối với tất cả các khối 52, nên theo phương án, thiết bị 10 khai thác cơ chế dự đoán không gian để mã hóa thông tin phụ bù chuyển động cho phần 20 của ảnh không xuất ra 30 nếu khả dụng bởi sơ đồ mã hóa nằm dưới các dòng dữ liệu 14 và 18. Trong trường hợp này, ví dụ, vector chuyển động 50 sẽ được mã hóa rõ ràng trong dòng dữ liệu 18 đơn thuần cho một khối trong số các khối 52 của phần không gian 20 của ảnh không xuất ra 30, trong khi đối với các khối 52 khác của phần 20, vector chuyển động sẽ được chấp nhận/dự đoán bằng phép dự đoán không gian. Ví dụ, chế độ bỏ qua có thể được sử dụng cho các khối 52 thay vì chế độ để vector chuyển động 50 được mã hóa rõ ràng. Bằng cách sử dụng chế độ bỏ qua, ví dụ, sẽ báo hiệu cho mỗi khối 52 rằng vector chuyển động 52 được chấp nhận hoặc được dự đoán theo không gian, và không có dữ liệu phần dư dự đoán nào cho khối 52 tương ứng. Dữ liệu phần dư dự đoán không được mã hóa bởi thiết bị 10 cho bất kỳ phần không gian 20 nào của ảnh không xuất ra 30, ví dụ, để sao chép nội dung ảnh của phần không gian 20 của ảnh kế thừa 26, mà ảnh không xuất ra 30 thay thế, không được sửa đổi.

Tham chiếu nhanh trở lại Fig.1, Fig.1 minh họa khả năng thiết bị hợp thành video 10 trên Fig.1 có thể được tạo cấu hình để điều khiển sự thay đổi theo thời gian của vị trí không gian của phần không gian 20 để đáp lại tín hiệu bên ngoài 60, tức là, các yêu cầu bên ngoài cho sự thay đổi. Nếu là vậy, thiết bị hợp thành video 10 có thể không thực thi yêu cầu bên ngoài cho sự thay đổi ngay lập tức và không rõ ràng. Hơn nữa, thiết bị hợp thành video 10 có thể, ngay khi nhận yêu cầu 60, kiểm tra tuần tự các ảnh đầu vào 22 để xác định ảnh đầu tiên trong số các ảnh không được tham chiếu, qua phép dự đoán thông tin phụ bù chuyển động theo thời gian, bằng bất kỳ ảnh đầu vào 22 khác. Các chi tiết và các lý do làm vậy được nêu ra dưới đây một cách cụ thể hơn.

Ví dụ, xem Fig.4. Fig.4 minh họa sự thay thế của ảnh 2' bằng ảnh không xuất ra 4'

so với ảnh 5'. Tức là, ảnh không xuất ra 4' đã được chèn vào bởi thiết bị 10 sao chép nội dung ảnh của vị trí của phần không gian 20 hợp lệ trước khi thay đổi của vị trí các phần không gian để được chèn vào trong ảnh không xuất ra 4' tại vị trí mới của phần không gian 20 để ảnh 5', phần không gian 20 của ảnh này đã được điền đầy bằng cách thu thập và sao chép từ một trong số các ảnh đầu vào, cụ thể ảnh đầu vào 3, có khả năng dự đoán nội dung ảnh bên trong phần không gian 20 bằng cách sử dụng thông tin phụ bù chuyển động của ảnh đầu vào 3, được minh họa sử dụng một số mũi tên 62 trên Fig.4, từ phần không gian 20 cùng vị trí của ảnh không xuất ra 4'. Nói cách khác, thông tin phụ bù chuyển động được thu thập và sao chép từ ảnh đầu vào 3 vào phần không gian 20 của ảnh 5' có thể, ví dụ, mã hóa một vector chuyển động trên mỗi khối liên dự đoán trong phần không gian 20.

Như được giải thích với Fig.3, phép dự đoán theo thời gian có thể là một lựa chọn để giảm tốc độ bit sẽ được dùng cho thông tin bù chuyển động, như thông tin phụ bù chuyển động trong các ảnh đầu vào 22. Tuy nhiên, phép dự đoán theo thời gian thông tin phụ bù chuyển động của phần không gian 20 của ảnh 5' từ ảnh tham chiếu sẽ, ngay lúc này, dẫn đến các sai số do lý do sau: ảnh không xuất ra 4' thay thế ảnh 2' do chức năng của ảnh 2' là ảnh tham chiếu cho ảnh 5'. Việc này hoạt động so với nội dung ảnh, tức là, nội dung ảnh cuối cùng được tái tạo được sao chép vào phần không gian 20 của ảnh 5' bằng thông tin phụ bù chuyển động 62. Tuy nhiên, như rõ ràng từ phần mô tả trên Fig.3, thông tin phụ bù chuyển động được truyền tải trong dòng dữ liệu cho ảnh không gian 20 của ảnh không xuất ra 4' không trùng với thông tin phụ bù chuyển động được báo hiệu cho phần không gian 20 của ảnh 2' như thông tin phụ bù chuyển động của phần không gian 20 của ảnh không xuất ra 4' đơn thuần, bằng cách tịnh tiến, sao chép phần không gian 20 của ảnh 2' vào phần không gian 20 của ảnh không xuất ra 4'. Nói cách khác, các vector chuyển động của phần không gian 20 của phần không xuất ra 4' đơn thuần là các vector chuyển động nhân tạo trong số toàn bộ phần không gian 20 trong khi các vector chuyển động được báo hiệu cho phần không gian 20 của ảnh 2' biểu diễn chuyển động của các nội dung ảnh trong cảnh video.

Theo đó, theo phương án của sáng chế, thiết bị hợp thành video 10 được tạo cấu hình để thực hiện yêu cầu bên ngoài 60 cho sự thay đổi không trực tiếp nhưng khi gặp cơ hội tuần tự tiếp theo hoặc tại hoặc thời điểm có sự thay thế các ảnh tham chiếu trong

số các ảnh kế thừa đã qua - so với cơ hội hoặc thời điểm đó - có thể không dẫn đến phép dự đoán thông tin phụ bù chuyển động theo thời gian sai, bằng bất kỳ ảnh đầu vào 22 theo sau - so với cơ hội hoặc thời điểm đó. Ví dụ, tưởng tượng rằng yêu cầu 60 sẽ đến thiết bị 10 tại thời điểm mà phần không gian 20 của ảnh 1' cần được điền đầy bằng cách thu thập vào sao chép dữ liệu tương ứng của ảnh đầu vào 1 trên Fig.1. Thiết bị 10 sẽ kiểm tra liệu bất kỳ trong số các ảnh tham chiếu khả dụng hiện có, tức là, bất kỳ ảnh trong số các ảnh tham chiếu có thể là ứng viên để được thay thế bởi các ảnh không xuất ra tương ứng hay không nếu thời điểm 28 ở ngay trước ảnh 1 - được sử dụng để dự đoán theo thời gian để dự đoán thông tin phụ bù chuyển động. Nếu có, thiết bị sẽ hoãn lại việc thực thi thay đổi của vị trí không gian của phần không gian 20. Sau đó, thiết bị 10 cũng sẽ kiểm tra ảnh đầu vào 2, ví dụ. Tức là, thiết bị sẽ kiểm tra liệu các ảnh tham chiếu tại thời điểm đó, tức là, các ứng viên cần được thay thế bởi ảnh không xuất ra, có được tham chiếu cho thông tin phụ bù chuyển động hay không. Trong trường hợp trên Fig.1, ví dụ, ảnh 3 có thể là ảnh đầu vào 22 đầu tiên mà từ đó không có ảnh nào sử dụng phép dự đoán theo thời gian để dự đoán thông tin phụ bù chuyển động từ bất kỳ ảnh trong số các ảnh tham chiếu sẽ được thay thế tại thời điểm đó, tức là, sau đó tất cả các ảnh tham chiếu khả dụng sẽ không được sử dụng để dự đoán theo thời gian thông tin phụ bù chuyển động và theo đó thiết bị 10 thực thi các yêu cầu 60 giữa các ảnh 2 và 3 của các ảnh đầu vào 22.

Một cách để báo hiệu một cách hiệu quả đến thiết bị hợp thành video 10 có khả năng chỉ ra về thời gian tại đó yêu cầu 60 được thực thi, tức là, cách hiệu quả để phát hiện các ảnh đầu vào tại đó yêu cầu 60 có thể được thực thi bằng cách chèn ảnh tương ứng trước các ảnh không xuất ra 30, để tạo ra video 14 sao cho các ảnh của một hoặc nhiều lớp thứ bậc theo thời gian nhất định được đảm bảo để không được sử dụng cho phép dự đoán theo thời gian thông tin phụ bù chuyển động. Tức là, trong khi các ảnh 22 của (các) ID lớp theo thời gian nhất định chẳng hạn như khi chúng vượt qua ID lớp thứ bậc theo thời gian ngưỡng nhất định, được phép để được tham chiếu, qua phép dự đoán theo thời gian cho thông tin phụ bù chuyển động, từ các ảnh 22 khác của dòng dữ liệu 14, thì dòng dữ liệu 14 có thể báo hiệu đến thiết bị hợp thành video 10 rằng các ảnh 22 của ID lớp thứ bậc theo thời gian khác một hoặc nhiều ID lớp theo thời gian nhất định, như ID đứng sau ngưỡng đó được đảm bảo để không được sử dụng làm tham chiếu

trong phép dự đoán theo thời gian của thông tin phụ bù chuyển động. Trong trường hợp này, thiết bị 10 có thể phát hiện việc bắt gặp ảnh tiếp theo tuần tự trong chuỗi các ảnh đầu vào 22 - từ thời điểm đến của yêu cầu 60 - tại thời điểm mã hóa mà không ảnh nào trong số một hoặc nhiều ảnh tham chiếu đã được thay thế bởi các ảnh không xuất ra nếu thời điểm mã hóa đó được xác định là thời điểm chuyển thích hợp 28 - là ID lớp theo thời gian chứa các ảnh được cho phép để được tham chiếu cho thông tin phụ bù chuyển động. Thiết bị có thể thực hiện việc kiểm tra này chủ yếu trên cơ sở ID lớp thứ bậc theo thời gian phạm vi ảnh của các ảnh đầu vào và bộ phân biệt lớp thứ bậc theo thời gian vượt qua phạm vi ảnh được bao gồm bởi cú pháp mức cao của dòng dữ liệu vào trong 14. ID lớp thứ bậc theo thời gian phạm vi ảnh của ảnh đầu vào 22 có thể, ví dụ, được chứa trong phần đầu khối truy cập được kết hợp riêng với mỗi ảnh 22, hoặc nằm trong phần đầu lát 40 của các lát của ảnh đầu vào 22 tương ứng. Cú pháp mức cao bao gồm bộ phân biệt lớp thứ bậc theo thời gian vượt qua phạm vi ảnh đã nêu có thể được bao gồm trong bản tin SEI của dòng dữ liệu 14 như sẽ được nêu ra dưới đây đối với ví dụ về cú pháp bản tin SEI cụ thể. Theo cách khác, thiết bị 10 có thể phát hiện sự vắng mặt hoàn toàn của phép dự đoán thông tin phụ bù chuyển động theo thời gian trong dòng dữ liệu 14 (được kết hợp với sự mất mát về hiệu quả mã hóa) qua cú pháp mức cao tương ứng và do đó thực thi yêu cầu 60 độc lập với ID lớp thứ bậc theo thời gian phạm vi ảnh. Theo cách khác nữa, dòng dữ liệu 14 đi vào có thể bao gồm các chỉ báo, ví dụ, ở dạng bản tin SEI hoặc loại đơn vị NAL đặc biệt, biểu thị sự thích hợp của ảnh hiện thời để thực thi yêu cầu 60 bằng vị trí của chỉ báo nằm trong dòng bit. Theo một cách khác nữa, thiết bị 10 có thể, theo mặc định, kỳ vọng sự vắng mặt hoàn toàn của phép dự đoán thông tin phụ bù chuyển động theo thời gian trong dòng dữ liệu 14 mà không kiểm chứng rõ ràng thực tế. Các dòng dữ liệu video đầu vào sẽ được cung cấp cho thiết bị 10 theo đó, tức là, tuân theo các ràng buộc tương ứng.

Đối với các ảnh không xuất ra 30 được chèn một cách nhân tạo, lưu ý rằng, để giảm tốc độ dữ liệu cho ảnh này, thiết bị có thể sử dụng phép dự đoán theo thời gian của thông tin phụ bù chuyển động giữa một ảnh không xuất ra 30 và một ảnh khác sao cho giữa chúng, ảnh khác được chèn tại một thời điểm 28 nhất định hoặc thậm chí giữa chúng ảnh được chèn vào tại các thời điểm khác nhau. Bằng phương pháp này, thậm chí chi phí mã hóa để mã hóa một vector chuyển động được mã hóa rõ ràng tạo ra hạt

giống cho phép dự đoán thông tin phụ bù chuyển động không gian cho các khối khác trong phần không gian của ảnh không xuất ra được giảm vì chỉ phần dư của phần dự đoán theo thời gian của ảnh được mã hóa.

Fig.5 minh họa thiết bị hợp thành video 10 trong khung làm việc cùng với bộ mã hóa video 70 và bộ giải mã video 72 nhận dòng dữ liệu video được hợp thành 16. Ở đây, thiết bị hợp thành video 10 tạo ra cùng với bộ mã hóa video 70 hệ thống 74, các ví dụ của hệ thống này được thể hiện dưới đây đối với các phương án sử dụng nhiều hơn một bộ mã hóa video 70 như, ví dụ, hệ thống giám sát giao thông cung cấp cho một hoặc nhiều máy khách các sự pha trộn của các cảnh camera giao thông có sẵn, hệ thống thực tế ảo thực hiện ứng dụng video toàn cảnh cung cấp cho một hoặc nhiều máy khách phần con của video toàn cảnh, hoặc hệ thống hội nghị điện thoại cung cấp cho một hoặc nhiều máy khách với sự pha trộn cảnh của những người tham gia ở bên thứ ba.

Thiết bị hợp thành video 10 nhận dòng dữ liệu video đầu vào 14 từ bộ mã hóa video 70 và hiển thị dòng dữ liệu này trong phần không gian 20 của các ảnh của dòng dữ liệu video được hợp thành 16 bằng cách được nêu trên. Bộ giải mã video 72 chỉ cần giải mã dòng dữ liệu video được hợp thành đầu vào 16 làm đầu ra bởi thiết bị 10. Tại đầu ra của bộ giải mã 72, chuỗi ảnh được tái tạo sẽ được hiển thị làm đầu ra, được biểu thị sử dụng ký hiệu tham chiếu 76 trên Fig.5. Bên trong, bộ giải mã 72 được biểu thị một cách minh họa như được hợp thành từ chuỗi bộ đệm ảnh được mã hóa 78, theo sau bởi máy giải mã 80 một lần nữa theo sau bởi bộ đệm ảnh được giải mã 82. Dòng dữ liệu video được hợp thành 18 đến đi vào bộ đệm 78. Máy giải mã 80 giải mã tuần tự các ảnh 26 và 30 của dòng dữ liệu 18 và chèn kết quả của việc giải mã các ảnh này vào bộ đệm ảnh giải mã 82. Như được biểu thị bởi mũi tên 84, đầu ra của bộ đệm ảnh giải mã 82 cũng được nạp trở lại vào máy giải mã 80 để các ảnh được giải mã trong bộ đệm 82 có thể đóng vai trò làm các ảnh tham chiếu cho các ảnh được giải mã sau như đã được nêu trên đối với phép dự đoán theo thời gian được bù chuyển động.

Khi xuất ra các ảnh được giải mã, một mặt, bộ giải mã 72 phân biệt ảnh kế thừa 26 và mặt còn lại là các ảnh không xuất ra 30: các ảnh không xuất ra 30 được báo hiệu trong dòng dữ liệu 18 là các ảnh không được xuất ra, tức là, không phải là một phần của video 76 cần được hiển thị. Ví dụ, dòng dữ liệu 18 sẽ bao gồm, đối với mỗi ảnh 26 và 30 của dòng dữ liệu 18, cờ báo hiệu liệu ảnh tương ứng sẽ được xuất ra hay không. Tuy

nhiên, cờ có thể cũng được báo hiệu trên cơ sở từng lát. Tức là, tất cả các lát thuộc về ảnh không xuất ra 30 sẽ báo hiệu nội dung ảnh tương ứng sẽ không được hiển thị. Trong HEVC, ví dụ, cờ `pic_output_flag` có thể được sử dụng để đạt được mục đích này. Nên lưu ý rằng các ảnh của các dòng dữ liệu video đầu vào có thể là tất cả các loại ảnh đầu ra, nhưng cách khác, nó có thể là các dòng dữ liệu này đã được đặt rải rác bằng cách ảnh không xuất ra.

Mặc dù không nêu cụ thể ở trên, bộ mã hóa video 70 có thể được tạo cấu hình để tuân theo một số ràng buộc mã hóa được đặt ra bởi chính bộ mã hóa-giải mã video. Ví dụ, khi điền đầy phần không gian 20 của các ảnh kế thừa 26 trên cơ sở dữ liệu tương ứng của các ảnh đầu vào 22, rõ ràng rằng các đường biên ảnh trước đó của các ảnh đầu vào 22 trở thành đường biên bên trong của phần không gian 20 sau khi điền đầy phần không gian 20. Tuy nhiên, sự thay đổi tình huống này có thể thay đổi phép dự đoán được bù chuyển động theo thời gian: các vector chỉ các vùng của các ảnh tham chiếu vượt ra ngoài đường biên ảnh có thể dẫn ra một số xử lý đặc biệt đối với một phần của vùng được sao chép của ảnh tham chiếu nằm bên ngoài ảnh tham chiếu chẳng hạn như phép ngoại suy hoặc tương tự. Tuy nhiên, sự dẫn ra như vậy có thể không xảy ra tại đường biên của phần không gian 20 vì đường biên đứng sau có thể, ví dụ, nằm bên trong ảnh kế thừa 26 tương ứng. Theo đó, bộ mã hóa video 70 có thể hạn chế thông tin phụ bù chuyển động gần các đường biên của ảnh 22 trong phạm vi mà thông tin phụ bù chuyển động không sao chép các ảnh tham chiếu tại các vùng vượt ra ngoài đường biên của ảnh 22. Ngoài ra, bộ mã hóa video 70 có thể bị ràng buộc để tránh phép nội suy dưới điểm ảnh được yêu cầu cho các vector chuyển động với độ chính xác dưới điểm ảnh gần với các đường biên của ảnh đầu vào 22. Ví dụ, các vector chuyển động với độ chính xác dưới điểm ảnh của các khối nằm trong vùng của ảnh đầu vào 22 có thể dẫn ra quy trình bộ lọc đáp ứng xung hữu hạn cho phép nội suy của các trị số mẫu độ chói hoặc sắc độ. Khi vị trí dưới điểm ảnh vector chuyển động gần về không gian với đường biên ảnh không gian của ảnh đầu vào 22, nhân bộ lọc có thể chồng lấn với các vùng vượt ra ngoài các đường biên của ảnh 22 trong phép nội suy các trị số mẫu dưới điểm ảnh. Trong trường hợp như vậy, phép xử lý đường biên ảnh đặc biệt nêu trên như phép ngoại suy hoặc tương tự có thể được dẫn ra. Việc điền đầy phần không gian 20 của các ảnh kế thừa 26 trên cơ sở dữ liệu tương ứng của các ảnh đầu vào 22, các đường biên

ảnh của các ảnh đầu vào 22 có thể trở thành các đường biên bên trong của phần không gian 20 của các ảnh kế thừa 26 sau khi điền đầy phần không gian 20 và việc dẫn ra như vậy có thể không xảy ra. Theo đó, bộ mã hóa video 70 có thể hạn chế việc sử dụng các vector chuyển động chính xác dưới điểm ảnh gần các đường biên của ảnh 22 đến mức mà quy trình nội suy dưới điểm ảnh không sử dụng các vùng điểm ảnh tham chiếu vượt ra ngoài các đường biên của ảnh 22. Ngoài ra, bộ mã hóa 70 có thể, khi thực hiện phép dự đoán theo thời gian của thông tin phụ bù chuyển động cho các khối được định vị không gian tại các đường biên không gian của ảnh đầu vào, tạo điều kiện cho các vector chuyển động của các khối xung quanh các khối được sắp xếp bên trong các ảnh tham chiếu nếu các khối như vậy tồn tại. Như được chỉ ra trước đó, việc điền đầy phần không gian 20 của các ảnh kế thừa 26 trên cơ sở dữ liệu tương ứng của các ảnh đầu vào 22 có thể biến đường biên ảnh của các ảnh đầu vào 22 thành các đường biên bên trong của phần không gian 20 của ảnh kế thừa 26. Do đó, quy trình dự đoán theo thời gian của thông tin phụ vector chuyển động của khối đã cho gần về không gian với đường biên bên trong của phần 20 có thể truy cập các khối nằm trong các ảnh tham chiếu lân cận các khối được sắp xếp tương ứng và do đó có thể truy cập các khối không khả dụng với bộ mã hóa 70 dẫn đến việc không khớp dự đoán. Theo đó, bộ mã hóa video 70 có thể hạn chế phép dự đoán theo thời gian của thông tin phụ bù chuyển động gần các đường biên của ảnh 22 trong phạm vi mà quy trình dự đoán không sử dụng thông tin từ các vùng ảnh tham chiếu vượt ra ngoài các đường biên của ảnh 22. Tương tự, các đường biên bên trong của ảnh đầu vào 22 có thể trở thành các đường biên ảnh bên trong ảnh 26 và bộ mã hóa 70 có thể hạn chế phép dự đoán theo thời gian của thông tin phụ vector chuyển động theo đó đối với đường biên bên trong của ảnh đầu vào 22. Lọc trong vòng lại, các bộ mã hóa video 70 có thể được thiết lập để sử dụng hoặc không sử dụng việc lọc trong vòng khi cung cấp các dòng dữ liệu video đầu vào. Nếu việc lọc trong vòng được báo hiệu để được sử dụng trong các dòng dữ liệu video dữ liệu đầu vào, thì thiết bị 10 có thể chấp nhận việc lọc trong vòng của các phần không gian tương ứng của các ảnh kế thừa 22 và nếu được báo hiệu để không được sử dụng trong các dòng dữ liệu video dữ liệu đầu vào, thì thiết bị 10 có thể không áp dụng việc lọc trong vòng của các phần không gian tương ứng của các ảnh kế thừa 22. Tuy nhiên, nếu lọc trong vòng được sử dụng, thiết bị không kích hoạt việc lọc trong vòng các ảnh kế thừa 22 theo cách để

việc lọc trong vòng đi qua đường biên của phần không gian 20 trong các ảnh kế thừa. Việc lọc trong vòng bổ sung không được kích hoạt bằng thiết bị 10 trong các ảnh không xuất ra 30, ví dụ, để không thay đổi nội dung ảnh tham chiếu.

Hơn nữa, trong trường hợp sử dụng nhiều hơn một bộ mã hóa video 70, có thể có lợi để đồng bộ hóa các bộ mã hóa video 70 này đối với cấu trúc GOP dự đoán theo thời gian được minh họa trên Fig.1 sử dụng các mũi tên 24 và/hoặc các tham số và công cụ mã hóa.

Hơn nữa, bộ mã hóa video 70 có thể, như được nêu trên đối với Fig.4, sử dụng khái niệm mã hóa thứ bậc theo thời gian khi tạo ra dòng dữ liệu 14 và đối với tập con gồm các mức thứ bậc theo thời gian, bộ mã hóa video 70 có thể, trên cơ sở tình nguyện, không sử dụng phép dự đoán thông tin phụ bù chuyển động theo thời gian, tức là, không sử dụng các ảnh của tập con tương ứng gồm các mức thứ bậc theo thời gian như tham chiếu cho TMVP, với sự đảm bảo rằng thiết bị 10 không sử dụng phép dự đoán này bằng cách báo hiệu đảm bảo nêu trên qua cú pháp mức cao tương ứng của dòng dữ liệu 14 sao cho thiết bị 10, sau đó, có thể nhận dạng các ảnh 22 này của dòng dữ liệu đi vào 14 mà tại đó sự thay đổi của vị trí không gian của phần không gian 20 xảy ra.

Như cũng sẽ trở nên rõ ràng từ phần mô tả của các kịch bản ứng dụng cụ thể được mô tả đối với các hình từ Fig.6 đến Fig.10, (các) bộ mã hóa video 70 tạo ra (các) dòng dữ liệu video đầu vào 14 có thể được tạo cấu hình để mã hóa các video được thu bởi camera video tương ứng, trong đó việc mã hóa video này có thể diễn ra lần lượt ở trên không hoặc theo thời gian thực. Bộ mã hóa video 70 có thể được tích hợp trong camera như vậy. Thiết bị 10 có thể nằm trong máy chủ trong khi bộ giải mã 72 có thể được tích hợp trong máy khách của máy chủ đó. Tuy nhiên, thậm chí cách khác, thiết bị 10 cũng được triển khai tại phía máy khách nhờ đó mở rộng bằng cách không đắt, khả năng của bộ giải mã 72 (tiêu chuẩn) để đạt được sự tự do hợp thành được nêu trên đơn thuần bằng cách kết nối theo chuỗi thiết bị 10 ngược dòng đến bộ giải mã 72. Các phương án được mô tả dưới đây, ví dụ, liên quan đến hệ thống hội nghị viễn tuyến trong đó, ví dụ, bộ giải mã 72 là một phần của máy khách của người tham gia. Cách khác, bộ giải mã 72 có thể là máy khách sao cho sự kiện hiển thị được gắn trên đầu, truy hồi phần con video toàn cảnh ra khỏi mảng video toàn cảnh với sự hợp thành được thực hiện bởi thiết bị hợp thành video 10 theo cách được nêu ra cụ thể dưới đây. Chính thiết bị hợp thành

video 10 này có thể được triển khai dưới dạng phần mềm chạy trên máy tính hoặc tương tự, trong khi bộ giải mã 72 có thể được triển khai bằng cách sử dụng phần mềm, phần cứng, hoặc phần cứng khả trình trên thiết bị di động.

Mặc dù không được minh họa trên Fig.5, có thể là yêu cầu 60 đến thiết bị hợp thành video 10 có thể bắt nguồn từ phía giải mã. Cách khác, yêu cầu 60 được tạo ra thủ công tại điểm điều khiển nào đó, như thực thể giám sát.

Dưới đây, các phương án được mô tả để theo đó, các phương án trên các hình vẽ từ Fig.1 đến Fig.5 được sử dụng đồng thời đối với nhiều hơn một dòng dữ liệu video. Trong phần mô tả dưới đây, các ảnh không xuất ra 30 được gọi là các ảnh tham chiếu được tạo ra (generated-reference-picture - GRP). Chúng thay thế ảnh tham chiếu đối với nhiều dòng dữ liệu video đầu vào. Như được mô tả trên đây, các GRP là các ảnh được hợp thành được chèn vào dòng bit/dòng dữ liệu video được hợp thành không được xuất ra và nhằm mục đích thay thế các ảnh tham chiếu ban đầu tại vị trí nhất định ở thời điểm khi giải mã dòng dữ liệu video được hợp thành. Các chi tiết khác về các GRP, ngoài phần mô tả đã được thể hiện ở trên, được trình bày dưới đây sử dụng một số ứng dụng mẫu, nhưng cũng có thể áp dụng cho các ứng dụng khác. Cụ thể, các chi tiết như vậy sẽ có thể được áp dụng riêng lẻ cho phần mô tả trên đây.

Fig.6 biểu diễn ứng dụng truyền theo dòng toàn cảnh trên cơ sở ô. Cách truyền theo dòng toàn cảnh được thực hiện bằng cách có máy chủ riêng theo vết vùng quan tâm (region of interest - ROI) của máy khách, tức là, vùng nội dung hiển thị trên thiết bị máy khách, và chuyển mã video toàn cảnh thành video chỉ chứa ROI cho máy khách đã cho. Phương pháp như vậy có hạn chế là không định tỉ lệ tốt. Tuy nhiên, với HEVC, ví dụ, dòng bit video có thể được chia thành các phần nhỏ hơn gọi là các ô. Các ô là các phần chia ở dạng hình chữ nhật nhỏ gồm ảnh, có thể được mã hóa theo cách sự phụ thuộc về thời gian trong số các ô khác nhau được loại bỏ [2]. Tuy nhiên, nếu các ô được giải mã độc lập, bộ giải mã video có thể không được sử dụng và do đó cần sử dụng quy trình vá ô tương tự như kỹ thuật được thể hiện trong tài liệu [1].

Hai thời điểm, tức là, $T = 0$ và $T = 1$, cùng với chuyển động màn hình máy khách được mô tả trên Fig.6, trong đó $T = 1$ biểu diễn điểm chuyển đổi để truyền theo dòng tương tác tại đó phía máy khách thay đổi vị trí của các ô được trình bày. Trong kịch bản

truyền theo dòng toàn cảnh, các máy khách thường điều hướng trên toàn cảnh bằng cách lựa chọn ROI thích ứng theo thời gian, ví dụ, bằng tương tác người dùng hoặc quy trình tự động như sự nhận biết ROI.

Điều này có nghĩa là vị trí của các ô nhận/tải đối với toàn bộ chuỗi toàn cảnh thay đổi theo thời gian. Bất kỳ ô tại vị trí mới được nhận trong quy trình truyền theo dòng tại $T = 1$ (tức là, không được nhận trước đó tại $T = 0$) yêu cầu sự truy cập ngẫu nhiên tại nội dung video được giải nén của ô này vì thời điểm trước $T < 1$ là không khả dụng với bộ giải mã để tham chiếu.

Tuy nhiên, tập hợp thứ hai gồm các ô của các vị trí được nhận trước không cần thiết yêu cầu sự truy cập ngẫu nhiên vì thông tin cho các ô trong tập hợp này đã được nhận. Các ô này được biểu thị bằng cách sử dụng tô bóng trên Fig.6 và Fig.7. Đối với các ô này, chỉ vị trí trong ảnh đầu ra khác với vị trí trong ảnh đã được giải mã là khả dụng để tham chiếu. Do đó, kỹ thuật được mô tả trong [1] không thể được sử dụng như vậy. Phép dự đoán theo thời gian từ các khung tham chiếu thông thường trong bộ đệm bộ giải mã video thất bại đối với tập hợp thứ hai gồm các ô này, vì các ảnh tham chiếu có thể có thông tin khác với ảnh được bắt gập trên phía bộ mã hóa.

Một lưu ý phụ nhỏ, rằng sự phân chia chín ô của video được hợp thành được chọn trên Fig.6 và Fig.7 chỉ dành cho mục đích minh họa và dĩ nhiên sự phân chia khác nào đó có thể cũng được sử dụng. Các ô được biểu thị sử dụng ký hiệu tham chiếu 90. Như trở nên rõ ràng hơn từ phần mô tả dưới đây, các ô của dòng dữ liệu video được hợp thành biểu diễn các vị trí không gian có thể của phần không gian trong đó dòng dữ liệu video đầu vào tương ứng có thể được hiển thị. Trong kịch bản ứng dụng toàn cảnh được mô tả trên Fig.6, nhiều dòng dữ liệu video đầu vào khả dụng: trong ví dụ trên Fig.6, 11 x 5 dòng dữ liệu video đầu vào 92 bao phủ cảnh video toàn cảnh trong đó tất cả các dòng dữ liệu video đầu vào 92 thu được phần không gian khác nhau của cảnh toàn cảnh, các phần không gian tiếp giáp với nhau, ví dụ, không có khoảng cách và được phân bố về không gian theo mảng với các cột và các hàng được mô tả trên Fig.6, để các video 92 được kết hợp với các phần cảnh toàn cảnh được sắp xếp lần lượt theo các cột và hàng. Tại mỗi thời điểm, dòng dữ liệu video được hợp thành mô tả trong các ô 3 x 3 của nó đơn thuần là ma trận con 3 x 3 trong 11 x 5 dòng dữ liệu video đầu vào 92.

Fig.7 minh họa rằng do vị trí mới của các ô tại thời điểm $T = 1$, sự tham chiếu được sử dụng trong video được vá, tức là, dòng dữ liệu video được hợp thành, mà sự tham chiếu được biểu thị tại 94 trên Fig.7, sẽ khác với dòng dữ liệu ban đầu trên phía bộ mã hóa được minh họa sử dụng các khối 96 trên Fig.7, dẫn đến độ lệch mà không cho phép sử dụng phép dự đoán theo thời gian cho các phần không được làm mới được tạo nét gạch bằng các đường trên Fig.6 và Fig.7. Do đó, các ô đã được nhận và được tái định vị ở vị trí mới trong tín hiệu đầu ra được hợp thành/được vá yêu cầu sự truy cập ngẫu nhiên nên phép dự đoán theo thời gian không được phép, dẫn đến tốc độ bit cao hơn của dòng bit không được hợp thành nhận được. Khái niệm của GRP giải quyết vấn đề này.

Fig.8 minh họa khái niệm GRP giải quyết vấn đề nêu trên đối với Fig.6 và Fig.7. GRP là ảnh thực hiện sự thay thế nội dung của ảnh tham chiếu thông thường để các ảnh sau có thể sử dụng phép dự đoán theo thời gian của các ô (tức là, các vùng của video) đã được nhận trước đó và được thay thế đến vị trí mới trên video được hợp thành/được vá tại và sau điểm chuyển đổi truyền dòng được biểu thị là 28 trên Fig.1. Các GRP chỉ được sử dụng để tham chiếu và không được xuất ra.

Trước khi xử lý với cảnh của ứng dụng tiếp theo, trong đó các phương án của các hình vẽ từ Fig.1 đến Fig.5 được áp dụng cho trường hợp sử dụng các dòng dữ liệu video đầu vào, phương án vừa được đề cập được giải thích tóm tắt bằng cách kết hợp các hình vẽ từ Fig.1 đến Fig.8. Cụ thể, theo các phương án được mô tả như trên đối với các hình vẽ từ Fig.6 đến Fig.8, hệ thống trên Fig.5, ví dụ, có thể bao gồm một bộ mã hóa video 70 trên mỗi video phân toàn cảnh 92, tức là, 11×5 trong phương án minh họa trên Fig.6. Thiết bị hợp thành video 10 sẽ vá mảng con của, ví dụ, các dòng dữ liệu video đi vào 3×3 vào trong dòng dữ liệu video được hợp thành. Khi nào mảng con di chuyển qua mảng của các dòng dữ liệu video đầu vào 92, thiết bị hợp thành video 10 thực hiện việc chèn các ảnh không xuất ra hoặc các GRP, trong đó các GRP được chèn 30 tại thời điểm 28 đó sẽ bao gồm thông tin phụ bù chuyển động nhân tạo tại các vị trí tương ứng với các dòng dữ liệu video đầu vào đã là một phần của mảng con đứng trước thời điểm 28. Trong trường hợp chuyển động đường chéo của mảng con, có 4 dòng dữ liệu, trong khi chuyển động ngang hoặc dọc cùng sở hữu 6 dòng dữ liệu trước và sau mỗi thời điểm chuyển đổi. Trong trường hợp trên Fig.6, ví dụ, bốn dòng dữ liệu video đầu vào là một

phần của mảng con sau thời điểm, gọi là $T = 1$, đã là một phần của mảng con đứng trước thời điểm, cụ thể tại $T = 0$, cụ thể là mảng con được biểu thị sử dụng sự bóng. GRP hoặc các GRP được chèn vào tại thời điểm 28, tức là, giữa các ảnh kế thừa 26, trên Fig.8, bằng cách tịnh tiến, sẽ sao chép nội dung của các ô trong đó bốn dòng dữ liệu video đầu vào này đã được định vị trước đó, đến các vị trí ô mới của bốn dòng dữ liệu video đầu vào này. (Các) GRP có thể thực hiện điều này, do đó, đối với một số dòng dữ liệu video đầu vào 14 song song. Trong trường hợp của kịch bản ứng dụng toàn cảnh này, thông tin phụ bù chuyển động nhân tạo sẽ tạo nấc chuyển động tịnh tiến cho phần không gian tương ứng với tất cả các dòng dữ liệu video đầu vào còn lại trong dòng dữ liệu video được hợp thành được hiển thị hiện thời. Trong kịch bản ứng dụng được mô tả tiếp theo, điều này có thể khác, tức là, thông tin phụ bù chuyển động được tạo ra nhân tạo cho các GRP có thể tịnh tiến đối với một dòng dữ liệu video đầu vào, nhưng hướng chuyển động có thể khác đối với các dòng dữ liệu video đầu vào khác nhau được xử lý song song.

Cụ thể, khi khái niệm GRP được mô tả trên ở trong ngữ cảnh của ứng dụng truyền theo dòng toàn cảnh đối với các hình vẽ từ Fig.5 đến Fig.8, ví dụ khác được thể hiện dưới đây đối với Fig.9. Như được đề cập trên đây, GRP cũng cho phép các ứng dụng khác, ví dụ, hội nghị video. Trong hệ thống hội nghị video, các dòng bit video của tất cả các thành viên tham gia được hợp thành/vá bằng cách cơ bản tương tự vào trong dòng bit video đơn. Khi thay đổi của sự bố trí thành phần loa, ví dụ, qua sự thay đổi loa hoặc sự dao động của người tham gia, các GRP được thêm vào dòng bit để cho phép phép dự đoán theo thời gian đúng cho các dòng bit video của người tham gia đã thay đổi vị trí trong thành phần. Trong ứng dụng như vậy, các vectơ chuyển động trong GRP tương ứng không cần thiết không đổi trên toàn bộ GRP, nhưng ít nhất đối với vùng được bao phủ bởi dòng bit video của mỗi người tham gia như được minh họa trên Fig.9, trong đó hai thành phần khác nhau được cho đối với $T = 0$ và $T = 1$ và các vectơ chuyển động của GRP cũng được minh họa. Do đó, trong trường hợp này, nhiều lát hoặc ô có thể được sử dụng để tương ứng với sơ đồ bố trí sao cho đối với mỗi loa, một khối mã hóa đơn có thông tin vectơ chuyển động và phần còn lại của các khối mã hóa cho loa đó được mã hóa được bỏ qua. Nói cách khác, GRP 30 trên Fig.9 có thể được mã hóa theo cách tương tự phần mô tả được đưa ra phía trên đối với Fig.3: ảnh 30 có thể được chia

nhỏ để dẫn đến kết quả một phần không gian 20 trên mỗi dòng dữ liệu video đầu vào thể hiện trước và sau thời điểm chuyển đổi 28, là ví dụ 3 trong trường hợp trên Fig.9, và bằng cách sử dụng phép dự đoán không gian, sự thay thế của phần không gian cho mỗi dòng trong số ba dòng dữ liệu video đầu vào có thể chỉ được mã hóa một lần cho mỗi dòng trong số ba dòng dữ liệu video đầu vào này, tức là, đối với chỉ một khối bên trong phần không gian tương ứng, trong khi phép dự đoán không gian có thể được sử dụng để tránh báo hiệu chuyển động tịnh tiến của mỗi dòng dữ liệu video đầu vào cho các khối còn lại trong cùng một phần không gian, một cách tương ứng.

Ứng dụng đích khác cho khái niệm GRP truyền theo dòng toàn cảnh đến các màn hình gắn trên đầu thường có trong các trường hợp sử dụng thực tế ảo hoặc hiện diện từ xa bằng cách thay đổi tốc độ truy cập ngẫu nhiên và điều khiển truyền theo dòng trong các dòng đầu vào 14. Trong ứng dụng đề cập, các dòng video khả dụng gần như bao phủ góc nhìn 360 độ tại vị trí (các) camera. Tương tự, phần lớn hơn của video góc nhìn khả dụng được thể hiện đồng thời cho người dùng để cung cấp tầm nhìn ngoại vi. Hơn nữa, sự điều chỉnh của vùng được hiển thị của video toàn cảnh do sự chuyển động phần đầu có thể xảy ra tại các khoảng thời gian ngắn và tốc độ cao hơn nhiều so với, ví dụ, các hệ thống đầu vào dựa vào cảm ứng.

Trong một số trường hợp, vùng quan tâm (region of interest - ROI) trong dòng video được ưu tiên một cách rõ ràng, ví dụ, hướng chuyển động trên phương tiện được điều khiển từ xa hoặc tương tự như được minh họa trên Fig.10. Các vùng ảnh nằm trong vùng có thể được mã hóa với tốc độ truy cập ngẫu nhiên thô hơn (hoặc bằng 0) như video tham chiếu được dự định là khả dụng tại phía bộ giải mã (tức là, luôn luôn được giải mã) thậm chí không được hiển thị để cung cấp truy cập tức thời nếu được biểu thị. Các góc nhìn khác (khu vực ngoại vi) cung cấp tốc độ truy cập ngẫu nhiên tương đối tốt để cung cấp để đôi khi thay đổi về hướng nhìn. Video được vá trong kịch bản này luôn luôn chứa ROI và các phần tùy chọn của các vùng ngoại vi. Các GRP sau đó được tạo ra như được mô tả trước khi phụ thuộc vào khu vực video được trình diễn.

So với kỹ thuật được nêu trong [1], khái niệm GRP được nêu trên có thể bao gồm một số phần mở rộng và danh sách dưới đây nêu rõ các ràng buộc và đặc điểm mà các dòng dữ liệu video đầu vào 14 và các GRP có thể lần lượt tuân theo và bao gồm.

Phép dự đoán vector chuyển động theo thời gian: Một số ràng buộc dòng bit mà trên cơ sở đó dòng dữ liệu video đầu vào 14 được tạo ra đã được mô tả ở trên, như một số hạn chế liên quan đến các video chuyển động tại các phần rìa của các ảnh 22. Ngoài ra, phép dự đoán vector chuyển động theo thời gian (temporal motion vector prediction - TMVP), tức là, phép dự đoán theo thời gian của thông tin phụ bù chuyển động, có thể được hạn chế để không có ảnh nào có thể được thay thế bằng các GRP được sử dụng cho TMVP. Thông thường, trong các ứng dụng minh họa được liệt kê, sẽ có các điểm chuyển đổi được xác định được sử dụng để thay đổi ROI hoặc sự bố trí loa từ thời điểm này về sau. Ví dụ, nếu khả năng định tỉ lệ theo thời gian dự đoán thứ bậc được sử dụng, các điểm chuyển đổi có thể, ví dụ, được lựa chọn tại các ảnh mức 0 theo thời gian hoặc các ảnh của tập con khác nào đó gồm các lớp theo thời gian khả dụng. Trong trường hợp này, không có ảnh nào với mức 0 theo thời gian được lựa chọn cho TMVP, vì ảnh này có khả năng được thay đổi thành GRP. Cách khác, TMVP có thể được vô hiệu hóa cho tất cả các mức theo thời gian tại các điểm chuyển đổi cho lượng ảnh/thời gian được xác định bởi cấu trúc dự đoán. Cấu trúc GOP có thể được duy trì không đổi qua dòng bit để các ảnh được vá có chung trị số của chỉ báo mức theo thời gian và quy trình viết lại được đơn giản hóa.

Các tập hợp ảnh tham chiếu: để tránh tăng bộ nhớ được yêu cầu tại bộ nhận/bộ giải mã và giảm kích thước bộ đệm ảnh được giải mã (decoded picture buffer - DPB) cần thiết, sẽ tốt hơn là các dòng dữ liệu video đầu vào sử dụng toàn bộ cấu trúc dự đoán giống nhau, để lượng và kích thước của tập hợp ảnh tham chiếu (reference picture set - RPS) có thể được giữ tối thiểu khi các dòng dữ liệu video đầu vào (tức là, các ô) được trộn/vá/được hợp thành bằng cách đồng bộ, đặc biệt khi một số dòng dữ liệu có lát I và một số lát P hoặc B. Các RPS thống nhất có thể được lựa chọn, ví dụ, khi HEVC IDR hoặc CRA của một dòng dữ liệu video đầu vào được chuyển đổi thành ảnh theo sau với lát I trong dòng dữ liệu video được hợp thành, RPS không trống có thể được lựa chọn phù hợp với RPS của ô, mà ô lát I được trộn/vá với, và đối với các lát sau chỉ sử dụng lát I làm tham chiếu RPS phù hợp (với các ô khác) có thể được biểu thị, và các cấu trúc cú pháp nhận dạng lát I dưới dạng ảnh tham chiếu độc nhất, các cấu trúc cú pháp cho danh sách ảnh tham chiếu và chỉ số ảnh cần được thêm vào các lát.

Đếm thứ tự ảnh (Picture order count - POC): sự chèn vào/thêm các GRP 30 có thể

đi kèm với thiết bị 10 thay đổi các trị số POC khác khi so sánh các trị số POC của các ảnh đầu vào 22 và mặt khác là các ảnh kế thừa tương ứng 26. Sai phân POC của các GRP duy trì tương tự như các ảnh ban đầu được thay thế, có thể đạt được bằng cách xác định delta POC, lớn hơn hoặc bằng sai phân POC lớn nhất trong số tất cả các ảnh trong DPB cộng một. Delta POC này được sử dụng cho các GRP để tính toán trị số POC bằng cách cộng vào POC của ảnh được thay thế trong DPB. Tổng của tất cả các delta từ IDR cuối cùng có thể được sử dụng để cộng vào trị số POC được suy ra từ phần đầu lát. Ngoài ra, các cấu trúc cú pháp mức cao khác có thể yêu cầu sự thích ứng, ví dụ, trong trường hợp sử dụng HEVC, vui_poc_proportional_to_timing_flag trong VPS có thể được bỏ thiết lập trong dòng đầu ra.

Các bộ lọc trong vòng: ngoài ra, để tránh sự lệch dự đoán các ảnh theo sau các GRP 30 do sự sửa đổi các ảnh dòng video đầu vào ban đầu 22 khi được tái định vị không gian qua phép dự đoán (các mũi tên 34), các bộ lọc trong vòng trong các GRP 30, chẳng hạn như tách khối và bộ lọc độ lệch thích ứng mẫu trong HEVC nên được vô hiệu hóa, ví dụ, trong PPS, tức là, không có bộ lọc trong vòng bổ sung nào ở trên đỉnh của sự lọc trong vòng được thực hiện bởi bộ mã hóa video 70 có thể được áp dụng cho các GRP 30.

Các ô và các lát: để giảm lượng dữ liệu được chèn, thiết bị 10 có thể giữ cấu trúc các GRP qua các ô và các lát tại mức tối thiểu vì việc này tốn chi phí báo hiệu không cần thiết. Tuy nhiên, việc thiết lập ô/lát tương tự với các ảnh được vá hoặc bất kỳ ảnh khác có thể là phương án thực hiện được yêu cầu/ưu tiên hoặc theo video nguồn.

Báo hiệu đầu ra: như được mô tả trên đây, bộ mã hóa-giải mã video nằm dưới dòng dữ liệu video được hợp thành cho phép báo hiệu của đặc tính đầu ra cho các GRP, tức là, các GRP không được xuất ra và chỉ được sử dụng để tham chiếu, ví dụ, qua output_flag trong các phần đầu lát HEVC. Sự có mặt của các phân tử cú pháp như vậy trong phần đầu lát có thể yêu cầu PPS bổ sung để báo hiệu sự có mặt của nó trong lát tham chiếu đến PPS đó.

Chèn tập hợp tham số: Các GRP không nhất thiết dùng chung tất cả các đặc tính được báo hiệu trong tập hợp các tham số tương ứng của chúng với các ảnh khác của dòng. Do đó, có thể có lợi để chèn các tập hợp tham số bổ sung vào dòng bit đầu ra cho

các GRP để tham chiếu đến.

(Các) khối GRP NAL: mã hóa các GRP bằng cách sử dụng các khối mã hóa (Coding Unit - CU) hoặc các khối có kích thước khả thi lớn nhất được ưu tiên để sản xuất ít bit nhất có thể. CU hoặc khối thứ nhất có thể được mã hóa, như nêu trên đối với Fig.3, biểu thị sự thay thế cần thiết của ảnh, tức là, bao nhiêu điểm ảnh và (các) ô tương ứng được chuyển đến hướng nào. Phần còn lại của các CU hoặc các khối được mã hóa trong chế độ bỏ qua hoặc tương tự, vì chuyển động là giống nhau cho tất cả các CU của phần không gian tương ứng. Tuy nhiên, các kích thước CU của các dòng đầu vào hoặc chuyển động đơn lẻ trên mỗi vùng ảnh có thể thúc đẩy các kích thước CU nhỏ hơn. Lát GRP có thể gồm có RPS mới hoặc chỉ số đến các RPS trong SPS đánh dấu các ảnh được thay thế là không cần thiết cho việc tham chiếu, khi được yêu cầu. Chỉ số có nghĩa là một khi các ảnh được thay thế bằng các GRP và các GRP khác được bao gồm, các RPS của các GRP khác không gồm có sự tham chiếu đến các ảnh ban đầu để giữ yêu cầu bộ nhớ DPB thấp.

Sự nhất quán SPS: Các IDR có thể kích hoạt SBS mới với các trị số khác với SPS hoạt động trước đó. Tuy nhiên, để có thể vá với dòng bit khác nhau trong dòng bit đơn, yêu cầu đó là các SPS của các dòng khác nhau là thống nhất. Để viết lại IDR vào lát I của ảnh theo sau và được vá với các lát không phải I khác, không được kích hoạt SPS với các phần tử cú pháp khác với SPS hoạt động trước đó.

Đối với sự mô tả trên, nên lưu ý rằng các GRP 30, dưới dạng dữ liệu lát hoặc phần tải tin lát 42, ví dụ, có thể được tiền mã hoá để chèn. Tức là, thiết bị 10 có thể có các tập hợp gồm các GRP được tiền mã hoá 30 phù hợp với các cấu hình dòng dữ liệu đầu vào được hình dung. Bằng cách này, các GRP như vậy có thể được chèn vào các dòng video được hợp thành 18 vì nội dung tải tin lát của chúng phụ thuộc chỉ vào các tham số mức cao như các kích thước ảnh, sự thay thế hoặc cấu trúc ô. Điều này cho phép, ví dụ, phương án thực hiện không có các bộ mã hóa entropy thực tế như máy mã hóa số học nhị phân thích ứng ngữ cảnh (context-adaptive binary arithmetic coding - CABAC) đối với, ví dụ, nội dung được mã hóa H.264/AVC hoặc HEVC.

Trong các phương án được kể trên, có các phương án hợp thành một số dòng dữ liệu video đầu vào vào một dòng dữ liệu video được hợp thành. Dưới đây, các phương

án được mô tả để đạt được sự hợp thành dòng dữ liệu video được hợp thành hoặc nhiều dòng dữ liệu video đầu vào sử dụng khái niệm khác một chút. Fig.11 thể hiện thiết bị 100 với đầu vào 102 để nhận nhiều 104 gồm các dòng dữ liệu video đầu vào $105_1, 105_2 \dots 105_N$ và đầu ra 106 để xuất ra dòng dữ liệu video được hợp thành 108. Chỉ số của các dòng dữ liệu video đầu vào đôi khi bị bỏ qua trong phần mô tả sau đây. Các dòng dữ liệu video đầu vào 105 và dòng dữ liệu video được hợp thành 108 được mã hóa bằng cách sử dụng phép dự đoán theo thời gian được bù chuyển động.

Bằng cách được nêu cụ thể dưới đây, thiết bị hợp thành video 100 trên Fig.11 có thể hợp thành video trong số các dòng dữ liệu video đầu vào đi vào 105 với tự do hợp thành tăng so với các phương án được mô tả trên đây so với các hình vẽ từ Fig.1 đến Fig.10, nhưng chi phí cho việc mở rộng số lượng ảnh đến khoảng không kích thước không gian được quét bởi dòng dữ liệu video được hợp thành 100. Nói chung, thiết bị hợp thành video 100 trên Fig.11 "ẩn" các dòng dữ liệu video đầu vào 105 trong phần tham chiếu của dòng dữ liệu video được hợp thành 100, để không xuất ra, trong khi phần khác của dòng dữ liệu video được hợp thành, được tạo ra theo cách tổng hợp bởi thiết bị hợp thành video 100, bao gồm video được đối chiếu bằng cách tham chiếu, qua phép dự đoán theo thời gian được bù chuyển động, nhiều vùng trong số phần tham chiếu. Phần sau của nội dung được tạo ra theo cách tổng hợp là một phần của dòng dữ liệu được hợp thành 100 để được xuất ra thực tế tại phía bộ giải mã.

Nói cách khác, theo khái niệm trên Fig.11, số lượng dòng dữ liệu video đầu vào 105 được xử lý trong miền được nén bằng thiết bị 100 để tạo ra dòng dữ liệu video đầu ra/dòng bit đơn 108 mà khi được nạp vào bộ giải mã đơn tạo ra sự hợp thành không gian của nhiều dòng dữ liệu video đầu vào/dòng bit 105. Fig.12 minh họa trường hợp sử dụng minh họa thứ nhất trong đó nội dung của một dòng bit đầu vào (dòng 1) 105_1 được phủ lên các phần của dòng bit đầu vào khác (dòng 2) 105_2 để đạt được hợp thành sự hợp thành dự định 110 bằng cách giải mã dòng bit đầu ra 108. Bằng cách được nêu ra cụ thể dưới đây, để đạt được mục đích này, dòng dữ liệu đầu ra 108 bao gồm phần tham chiếu 112 mang nội dung video của các dòng dữ liệu đầu vào 105_1 và 105_2 và phần hợp thành sẽ được xuất ra thực tế, tham chiếu, qua phép dự đoán theo thời gian được bù chuyển động, phần tham chiếu 112. Trên Fig.12, các mũi tên được tô bóng khác nhau minh họa các vùng tham chiếu được sử dụng trong số các dòng dữ liệu video

đầu vào 105. Nói cách khác, những mũi tên tương tự sẽ minh họa các vector dự đoán để tạo lập nội dung video được hợp thành trong phần 114. Các chi tiết khác được thể hiện dưới đây.

Đề cập đến Fig.11, thiết bị hợp thành video 100 ghép kênh nhiều dòng dữ liệu video đầu vào 105 vào phần tham chiếu 112 của các ảnh 116 của dòng dữ liệu video được hợp thành 108 bằng cách thu thập và sao chép thông tin phụ bù chuyển động và dữ liệu phần dư dự đoán của nhiều 104 dòng dữ liệu video đầu vào 105 vào trong phần thứ nhất 112. Phần thứ hai 114 của các ảnh 116 của dòng dữ liệu video được hợp thành 108 được điền đầy với video được đối chiếu được tạo ra theo cách tổng hợp bằng cách tham chiếu, qua phép dự đoán theo thời gian được bù chuyển động, phần tham chiếu 112. Trong khi phần tham chiếu được báo hiệu trong dòng dữ liệu 108 để không được xuất ra, nhưng phần thứ hai 114 được báo hiệu để được xuất ra.

Như được nêu ra cụ thể dưới đây, có một số khả năng về cách để ghép kênh các dòng dữ liệu video đầu vào 105 vào phần tham chiếu 112. Gọi "lượng video" được truyền tải bởi dòng dữ liệu video đầu vào 105_i nhất định ký hiệu, ví dụ, số lượng mẫu trên mỗi ảnh 118 của dòng dữ liệu video đầu vào tương ứng 105_i là $n_i \times m_i$ lần số lượng ảnh trên mỗi giây như t_i , và để còn ký hiệu "lượng video" của phần thứ hai 114 là số lượng mẫu trên mỗi ảnh của video được đối chiếu là $n_o \times m_o$ lần số lượng ảnh trên mỗi giây, khi đó dòng dữ liệu video được đối chiếu 108 bao gồm, ví dụ, ít nhất $n_o \times m_o \times t_o + \sum n_i \times m_i \times t_i$ mẫu trên mỗi giây. Các thay đổi về kích thước ảnh trong số các dòng dữ liệu đầu vào video có thể dẫn đến các vùng được điền đầy dữ liệu giả được thêm vào kích thước tối thiểu sẽ được giải thích dựa vào Fig.15. Các cách khác nhau hiện có bằng cách nào các dòng dữ liệu video đầu vào 105 có thể được "ẩn" hoặc được ghép kênh vào phần tham chiếu 112. Để đạt được mục đích này, ví dụ, phần tham chiếu 112 có thể bao gồm các ảnh không xuất ra và/hoặc các vùng ảnh gồm các ảnh đầu ra của dòng dữ liệu video được hợp thành 108 sẽ được cắt bỏ. Các chi tiết khác được mô tả dưới đây.

Nói cách khác, khái niệm trên Fig.11 sử dụng một số dòng bit đầu vào 105 để tạo ra phần 112 của dòng bit đầu ra đơn 108 mới qua việc ghép kênh. Các ảnh 118 hoặc các phần của chúng gồm ảnh của các dòng bit đầu vào 105 được dự định để tạo ra sự hợp thành sẽ được xuất ra tại thời điểm đơn được tham chiếu đến như các ảnh trong tập hợp

các ảnh tham chiếu (picture in a set of reference pictures - PSR) dưới đây.

Đối với Fig.13 và Fig.14, hai sự thay thế cho phép ghép kênh các dòng bit đầu vào 105 vào trong phần tham chiếu 112 được mô tả dưới đây cụ thể hơn. Sự thay thế thứ nhất được biểu diễn và được minh họa dựa vào Fig.13. Fig.13 minh họa ví dụ trong đó hai dòng dữ liệu video đầu vào 105 đều sử dụng IPPP tham chiếu cấu trúc GOP, nhưng điều này được chọn trên Fig.13 đơn thuần cho mục đích minh họa. Hai dòng dữ liệu video đầu vào minh họa 105₁ và 105₂ trên Fig.13 được ghép kênh sử dụng kỹ thuật ghép kênh phân chia theo thời gian, vào phần không gian tĩnh về không gian 118 của ít nhất tập con gồm các ảnh 116 của dòng dữ liệu video được hợp thành 108. Tức là, trong trường hợp trên Fig.13, các ảnh 118 được sử dụng thay thế để điền đầy ít nhất một tập con của các ảnh 116 của dòng dữ liệu 108 để, như được minh họa trên Fig.13, cặp các ảnh 116 liên tiếp của dòng dữ liệu 108 có phần không gian tĩnh về không gian 119 được điền đầy với một ảnh 118 của dòng dữ liệu đầu vào 105₁ và một ảnh 118 của dòng dữ liệu video đầu vào 105₂ khác, cả hai được dự định để tạo ra video hợp thành tại thời điểm giống nhau và thuộc về thời điểm tương tự trong hai dòng dữ liệu video đầu vào 105₁ và 105₂, chẳng hạn. Ví dụ, các ảnh 118 của dòng dữ liệu video đầu vào 105₁ có thể có cùng kích thước như các ảnh 118 của dòng dữ liệu video đầu vào 105₂, tức là, $n_1 = n_2$ và $m_1 = m_2$, để phần không gian tĩnh về không gian 119 tương ứng có chung kích thước trong ít nhất một tập con gồm các ảnh 116 của dòng dữ liệu video được hợp thành 108 được điền đầy bằng cách thu thập và sao chép thông tin phụ bù chuyển động và dữ liệu phần dư dự đoán của các ảnh 118 tương ứng của các dòng dữ liệu video đầu vào 105 này. Tuy nhiên, mặc dù được mô tả trên Fig.13, nhưng các ảnh 118 của các dòng dữ liệu video đầu vào 105₁ và 105₂ khác có thể có kích thước khác. Do đó, trong phần tham chiếu 112 của dòng dữ liệu video đầu ra, các ảnh 118 của các dòng dữ liệu video đầu vào 105₁ và 105₂ theo sau lẫn nhau theo cách được gọi là liên tục (back-to-back - B2B) sau đây. Nói cách khác, các ảnh được dự định tạo ra sự hợp thành tại thời điểm nhất định theo sau nhau trong dòng dữ liệu video được hợp thành dưới dạng các ảnh riêng lẻ với các trị số POC khác biệt. Do sự xen kẽ theo thời gian của các ảnh 118 của các dòng dữ liệu video đầu vào 105₁ và 105₂ trong dòng dữ liệu video đầu ra 108, thiết bị 10 có thể sửa đổi các trị số chênh lệch đếm thứ tự ảnh tham chiếu hoặc các tham chiếu đến các tập hợp của các trị số chênh lệch đếm thứ tự ảnh tham chiếu trong các

phần đầu lát trong các lát của các ảnh 118 của các dòng dữ liệu video đầu vào 105₁ và 105₂ để tính đến sự thay đổi các trị số POC của các ảnh 116 mà nội dung của các ảnh 118 được chấp nhận.

Ví dụ, trên Fig.13, hai ảnh của dòng dữ liệu video đầu vào 105₁ được minh họa để được sử dụng để điền đầy phần 119 gồm các ảnh đầu ra 116 của dòng dữ liệu 108. Do cấu trúc tham chiếu IPPP của dòng dữ liệu video đầu vào ban đầu 105₁, ảnh thứ nhất trong số các ảnh này tạo ra ảnh tham chiếu của ảnh thứ hai trong số các ảnh này như được biểu thị sử dụng mũi tên 120. Để duy trì sự tham chiếu này trong dòng dữ liệu video được hợp thành 108, thiết bị 10 có thể tính đến thực tế là mặc dù các ảnh tương ứng trong dòng dữ liệu video đầu vào 105₁ có, ví dụ, sự chênh lệch POC 1 như được minh họa bởi các số nhỏ “1” và “2” trong các góc phía trên bên trái của các ảnh 116, nhưng sự chênh lệch POC nằm giữa các ảnh 116 của dòng dữ liệu video đầu ra của dòng dữ liệu video đầu ra 108 có phần không gian 119 được điền đầy với các ảnh đầu vào 118 tương ứng bây giờ có sự chênh lệch POC là 2, tức là, 3 – 1. Cách tương tự áp dụng đối với phép dự đoán theo thời gian 122 nằm giữa các ảnh của dòng dữ liệu video đầu vào 105₂ như được biểu thị với dấu tham chiếu 122.

Do đó, mặc dù thiết bị 10 có thể thực hiện việc điều chỉnh đối với thông tin cú pháp mức cao như POC và RPS, nhưng sự tự do trong việc tạo ra các dòng dữ liệu video đầu vào 105₁ đến 105_N bởi các bộ mã hóa video có thể tăng so với các phương án được mô tả trên dựa vào các hình vẽ từ Fig.1 đến Fig.10. Ví dụ, vì phương pháp liên tục trên Fig.13 không bao gồm kỹ thuật vá các ảnh đầu vào 118 sử dụng các ô, nên việc không phù hợp của bộ mã hóa - bộ giải mã có thể không xảy ra.

Khả năng thứ hai của việc ghép kênh nội dung ảnh của các ảnh 118 của các dòng dữ liệu video đầu vào 105 vào phần tham chiếu 112 của dòng dữ liệu video được hợp thành 108 được mô tả trên Fig.14. Ở đây, phép ghép kênh phân chia theo không gian được sử dụng để ghép kênh các ảnh 118 của các dòng dữ liệu video đầu vào 105 vào dòng dữ liệu 108. Các dòng dữ liệu video đầu vào 105₁ và 105₂ chiếm giữ các ô khác nhau 122 và 124 của các ảnh 116 của dòng dữ liệu video được hợp thành 108. Cụ thể, Fig.14 minh họa việc sử dụng minh họa của cấu trúc tham chiếu tương tự như trong trường hợp trên Fig.13. Ở đây, các PSR được vá với nhau như được mô tả trong [1] và/hoặc như được mô tả trên đây dựa vào các hình vẽ từ Fig.6 đến Fig.10. Ở đây, so với

Fig.13, kích thước ảnh của dòng bit đầu ra, tức là, kích thước của các ảnh 116, tăng so với các dòng bit đầu vào riêng lẻ phụ thuộc vào các kích thước không gian của các PSR được vá. Tức là, trong trường hợp trên Fig.14, thiết bị 10 điền đầy, ví dụ, một ảnh 116 nằm trong ô thứ nhất 122 với ảnh 118 của dòng dữ liệu video đầu vào thứ nhất 105₁ và ô 124 khác của cùng ảnh 116 với ảnh được căn chỉnh theo thời gian 118 dòng dữ liệu video 105₂ và làm tương tự với các ảnh 116 khác. Ví dụ, ảnh 116 tiếp theo của dòng dữ liệu video được hợp thành 108 được điền đầy bằng cách sử dụng phương pháp thu thập và sao chép được nêu trên dựa vào các hình vẽ từ Fig.1 đến Fig.13, tại ô 122 bằng cách sử dụng ảnh 118 tiếp theo của dòng dữ liệu video đầu vào 105₁ và tại ô 124 bằng cách sử dụng ảnh được căn chỉnh theo thời gian 118 của dòng dữ liệu video đầu vào 105₂. Theo đó, các ảnh được sắp hàng theo thời gian 118 của các dòng khác nhau trong số các dòng dữ liệu video đầu vào 115 được chấp nhận vào các ô khác nhau của các ảnh 116 của dòng dữ liệu video được hợp thành 108 và, theo đó, được kết hợp trong dòng dữ liệu video 108 với một trị số POC chung được biểu thị tại t(1) cho các ảnh thứ nhất 118 của các dòng dữ liệu video đầu vào 105₁ và 105₂ được thể hiện trên Fig.14, và t(2) cho các ảnh thứ hai của các dòng dữ liệu video đầu vào 105₁ và 105₂ này. Phụ thuộc vào cách phân thứ hai 114 được thêm vào dòng dữ liệu 108 như được mô tả chi tiết hơn dưới đây, thiết bị 10 có thể có hoặc có thể không sửa đổi các trị số chênh lệch đếm thứ tự ảnh tham chiếu và/hoặc các tham chiếu đến các tập hợp gồm các trị số chênh lệch đếm thứ tự ảnh tham chiếu trong dòng dữ liệu video được hợp thành 108 so với các dòng dữ liệu video đầu vào 105. Ví dụ, như được mô tả cụ thể hơn dưới đây, có thể là phần video được đối chiếu 114 được truyền tải trong dòng dữ liệu video được hợp thành 108 bằng cách gán không gian với phần tham chiếu 112, và trong trường hợp đó, dòng dữ liệu video được hợp thành 108 có thể không bao gồm bất kỳ ảnh 116 nào vượt ra khỏi các ảnh có các ô 122 và 124 được điền đầy bằng cách sử dụng các dòng dữ liệu video đầu vào 105 để, ví dụ, sự chênh lệch POC tương ứng tham chiếu đến phép dự đoán theo thời gian 120 và 122 vẫn duy trì không đổi. Nếu việc chèn của phần video được đối chiếu 114 vào dòng dữ liệu video được hợp thành 108 dẫn đến việc chèn các ảnh 116 bổ sung vào dòng dữ liệu video được hợp thành 108, không được thể hiện trên Fig.14, thì thiết bị 10 có thể sửa đổi các cấu trúc cú pháp mức cao như trị số đếm thứ tự ảnh, các trị số chênh lệch đếm thứ tự ảnh tham chiếu hoặc các tham chiếu đến các tập

hợp gồm các trị số chênh lệch đếm thứ tự ảnh tham chiếu một cách phù hợp.

Do đó, trong trường hợp trên Fig.14, phần tham chiếu 112 tiêu thụ, ví dụ, $2 \times n \times m$ mẫu về không gian trong số các ảnh 116 của dòng dữ liệu video được hợp thành 108 thuộc về phần tham chiếu 112 khi, như được minh họa trên Fig.14, các ảnh 118 của các dòng dữ liệu video đầu vào có kích thước bằng nhau, ví dụ, $n \times m$.

Giả thiết sau rằng các ảnh 118 của các dòng dữ liệu video đầu vào 105 khác nhau có cùng kích thước không cần thiết được hoàn thành. Trong cả hai trường hợp, các ảnh 118 của các dòng dữ liệu video đầu vào 105 khác nhau có thể có kích thước khác nhau. Trong trường hợp đó, một số vùng của dòng dữ liệu video được hợp thành 108 có thể được điền đầy với nội dung giả, như được mô tả sau đây dựa vào Fig.15.

Fig.15 thể hiện theo sơ đồ cách tạo ra video được đối chiếu như Fig.12 đã làm, nhưng bây giờ minh họa trường hợp trong đó các ảnh 118 của dòng dữ liệu video đầu vào 105₂ có kích thước nhỏ hơn các ảnh của dòng dữ liệu video đầu vào 105₁. Trong phương pháp tham chiếu được vá đã được thể hiện trên đây dựa vào Fig.14, điều này có thể dẫn đến, ví dụ, trong phần tham chiếu 112 tiêu thụ không gian của vùng có $(n_1 + n_2) \times m_1$ mẫu trên mỗi ảnh 116 của dòng dữ liệu đầu ra 108 miễn là thuộc về phần tham chiếu 112, giả thiết rằng kích thước của các ảnh của dòng dữ liệu video đầu vào thứ nhất 105₁ là $n_1 \times m_1$ và kích thước của các ảnh của dòng dữ liệu video đầu vào thứ hai 105₂ là $n_2 \times m_2$. Trong trường hợp đó, vùng hình chữ nhật nằm trong mỗi ảnh 116 của phần tham chiếu 112 có thể, ví dụ, được điền đầy với thông tin giả 130. Tất nhiên, các ảnh của dòng dữ liệu video đầu vào 105₁ và 105₂ có thể được vá luân phiên theo chiều dọc với nhau thay vì theo chiều ngang như được minh họa trên Fig.14 và Fig.15. Giải thích phương pháp vừa được nêu ra cho phương pháp B2B được minh họa ở trên dựa vào Fig.13 có thể có nghĩa là, ví dụ, mỗi ảnh 116 của dòng dữ liệu video được hợp thành có phần không gian 119 được điền đầy với ảnh 118 của dòng dữ liệu video đầu vào 105₂ có phần không được điền đầy là $(n_1 \times m_1) - (n_2 \times m_2)$ mẫu được điền đầy với dữ liệu giả. Ví dụ, các ảnh 118 của các dòng dữ liệu video đầu vào 105₁ và 105₂ có thể tất cả được ghi vào góc phía trên bên trái của phần không gian tĩnh 119 để phần giả sẽ được điền đầy có thể có dạng chữ L dọc theo phía dưới và bên phải của phần 119.

Thực tế là Fig.15 minh họa kích thước của phần video được đối chiếu 114 của

dòng dữ liệu video được hợp thành 108 trùng với kích thước của ảnh lớn hơn trong số các ảnh của các dòng dữ liệu video đầu vào thứ nhất 105₁ và thứ hai 105₂ trên Fig.15 đơn thuần được chọn cho mục đích minh họa.

Do đó, Fig.15 đã minh họa rằng các dòng dữ liệu video đầu vào có thể được đi kèm với các lát giả với nội dung tùy ý khi được giải mã và không được sử dụng để tham chiếu bởi phần video được đối chiếu hoặc SLCP 114. Các lát giả có thể được sử dụng để căn chỉnh cả hai chiều ảnh không gian của tất cả các dòng bit đầu vào 105 trong phương pháp B2B hoặc cho phép vá bằng cách căn chỉnh ít nhất một chiều ảnh nếu cần, như được thể hiện trên Fig.15.

Lưu ý rằng để cho mục đích minh họa, tất cả phân giải thích sau đây sử dụng phương pháp tham chiếu được vá, nhưng tất cả các giải thích sau đây cũng có thể được thực hiện một cách thay thế bằng cách sử dụng phương pháp B2B.

Trong khi Fig.15 minh họa khả năng về cách giải quyết các chênh lệch giữa các dòng dữ liệu video đầu vào về chiều không gian, Fig.16 minh họa khả năng thiết bị 10 giải quyết các dòng dữ liệu video đầu vào 105 gồm các tốc độ khung khác nhau. Các dòng dữ liệu đầu vào 105 gồm các tốc độ khung khác nhau có thể được ghép kênh bằng thiết bị 10 vào phần tham chiếu 112 bằng cách lấy mẫu lại các dòng bit đầu vào 105 có các tốc độ khung thấp hơn đến tốc độ khung tối đa được bắt gộp trong số tất cả các dòng bit đầu vào 105. Một cách để căn chỉnh các tốc độ khung trong dòng dữ liệu video 108 là thêm theo cách chọn lọc các lát giả không được sử dụng cho việc tham chiếu và sử dụng cùng ảnh (vùng) tham chiếu trong nhiều SLCP liên tiếp như được minh họa trên Fig.16 trong đó tốc độ khung của dòng dữ liệu video đầu vào 105₂ được minh họa là bằng một nửa tốc độ khung của dòng dữ liệu video đầu vào 105₁ và các SLCP được tạo ra tại tốc độ khung khả dụng tối đa, tức là, tốc độ khung của dòng dữ liệu video đầu vào 105₁. Tương tự, tốc độ khung của các SLCP có thể, ví dụ, thấp hơn tốc độ khung tối đa trong số các dòng dữ liệu video đầu vào, bằng cách không sử dụng, ví dụ, một số ảnh tham chiếu của (các) dòng dữ liệu video đầu vào tốc độ khung cao hơn để hợp thành.

Các RPS nên được thiết lập theo cách nhờ thiết bị 10 để các tham chiếu cần đến bởi các PSR gồm có tất cả các ảnh tham chiếu cần thiết cho tất cả các dòng được ghép kênh. Đối với phương pháp B2B, trong đó các tham chiếu thuộc về cùng thời điểm

không chia sẻ POC chung, điều này chắc chắn dẫn đến sự tăng kích thước của mỗi RPS tương ứng. Đối với phương pháp tham chiếu được vá và khi RPS được căn chỉnh (theo nghĩa RPS thực tế và các tham chiếu đến RPS), thì kích thước hoặc lượng tăng là tối thiểu.

Phương án thực hiện dễ dàng có thể, ví dụ, báo hiệu RPS mới trong các phần đầu lát của các ảnh tham chiếu và/hoặc các ảnh SLCP, mặc dù điều này có thể dẫn đến chi phí báo hiệu. Tuy nhiên, điều này là không đáng kể khi quy trình được thực hiện trên thiết bị cuối mà không có sự truyền dẫn tiếp theo. Khi quy trình được thực hiện trên thực thể từ xa, chẳng hạn như các máy chủ đám mây ở xa thiết bị cuối, và sau đó được truyền đến thiết bị cuối, thì điều này có thể có lợi để điều chỉnh các RPS trong các tập hợp tham số thích hợp để giảm thiểu chi phí báo hiệu.

Sau khi hợp nhất/ghép kênh các dòng đầu vào 104 để f tham chiếu vào phần tham chiếu của dòng dữ liệu video đầu ra 108 để dẫn đến các PSR tương ứng cho mỗi thời điểm, việc thêm dữ liệu lát SLCP chứa sự hợp thành bằng phương tiện dự đoán theo thời gian tham chiếu đến B2B hoặc phần ảnh tham chiếu được vá 112 được tạo ra và được bổ sung dưới dạng phần 114 đến dòng dữ liệu 108. SLCP được hợp thành này được dự định để xuất ra bởi bộ giải mã và/hoặc màn hình đến người dùng cuối. SLCP có thể bao gồm thông tin phụ bù chuyển động, chẳng hạn như các vectơ dự đoán/chuyển động, chỉ đến các vị trí điểm ảnh trong các PSR để tạo ra sự hợp thành các nội dung ảnh qua phép dự đoán mẫu.

Có nhiều khả năng để thiết bị 10 định vị phần 114 so với phần tham chiếu 112 trong không gian truy cập không gian đếm thứ tự ảnh, một số trong số đó được trình bày dưới đây dựa vào Fig.17, Fig.18a và Fig.18b.

Vị trí của các SLCP trong dòng bit đầu ra 108 có thể được chọn phụ thuộc vào phương pháp tham chiếu/ghép kênh được áp dụng. Đối với phương pháp tham chiếu B2B được trình bày trên đây dựa vào Fig.13, ví dụ, dữ liệu lát SLCP có thể được ghép kênh vào dòng bit đầu ra 108 dưới dạng các ảnh đơn với các POC khác biệt với ảnh 116 của dòng bit đầu ra 108 mà các ảnh đầu vào 118 được ghép kênh vào đó. Cụ thể, trong trường hợp này, phần 114 có thể bao gồm các ảnh đầu ra rải rác bên trong các ảnh tham chiếu được thể hiện trên Fig.13, các ảnh đầu ra được sắp xếp theo thời gian tiếp sau các

PSR được tham chiếu. Điều này được minh họa ví dụ trên Fig.17, trong phạm vi mở rộng phương pháp B2B như được mô tả trên Fig.13 bởi phần 114.

Tức là, theo Fig.17, thiết bị 10 ghép kênh, bằng cách ghép kênh phân chia theo thời gian, các ảnh 118 của các dòng bit video đầu vào đi vào 105₁ và 105₂ vào các ảnh 116 tương ứng của dòng bit video được hợp thành 108 để tạo ra các phần tham chiếu 112 của dòng bit 108, và rải rác giữa các ảnh này là các ảnh 116 khác được điền đầy với dữ liệu tham chiếu, qua phép dự đoán theo thời gian bù chuyển động 130, các ảnh 116 tạo ra phần tham chiếu 112. Trong ví dụ trên Fig.17, toàn bộ vùng của ảnh 116 thuộc về phần 114 có thể được dành riêng để được hiển thị/xuất ra. Cách khác, đơn thuần phần phụ của nó có thể được dành riêng để được xuất ra với các phần còn lại được cắt. Cụ thể, thiết bị 10 có thể thay thế ảnh 116 giữa các ảnh 116 thuộc về phần tham chiếu 112 theo cách được ghi theo thời gian với nhau để các ảnh 116 của phần tham chiếu 112 thuộc về một thời điểm và do đó tạo ra tập hợp các ảnh tham chiếu liên tiếp ngay với nhau, và các ảnh 116 là một phần của phần 114 được rải rác giữa các ảnh 116 như vậy của phần tham chiếu 112 đã được điền đầy bằng cách thu thập và sao chép từ các ảnh của các dòng dữ liệu video đầu vào 105 hoặc thời điểm bằng nhau. Như đã được ký hiệu ở trên, tốc độ khung của các ảnh 116 thuộc về phần 114 có thể giống với tốc độ khung của các ảnh của các dòng bit video đầu vào 105. Trong trường hợp trên Fig.17, các ảnh 116 của dòng dữ liệu đầu ra 108 có thể được biểu thị bởi thiết bị 10 là các ảnh không xuất ra của dòng bit 108, trong khi các ảnh 116 tạo ra phần 114 sẽ được báo hiệu đến các ảnh xuất ra, tức là, các ảnh sẽ được xuất ra tại phía bộ giải mã. Nên lưu ý rằng Fig.17 biểu thị rằng sự thay thế các ảnh 116 thuộc về phần 114 ở giữa các ảnh 116 tạo ra phần tham chiếu 112 bằng cách điền đầy với các ảnh đầu vào 118 bằng cách sử dụng phép ghép kênh phân chia theo thời gian dẫn đến tốc độ tăng POC cao hơn so với việc không chèn các ảnh 116 thuộc về phần 114, mà thực tế được giải quyết một cách phù hợp bởi thiết bị 10 trong việc quản lý sự tham chiếu dự đoán theo thời gian, tức là, sự sửa đổi các trị số chênh lệch đếm thứ tự ảnh tham chiếu và các phần đầu lát và/hoặc tham chiếu đến tập hợp các trị số chênh lệch đếm thứ tự ảnh tham chiếu. Chính xác hơn, thiết bị 10 có thể, ví dụ, điền đầy mỗi ảnh 116 của dòng dữ liệu 108, thuộc về phần tham chiếu 112, bằng cách chấp nhận một hoặc nhiều lát vào ảnh đầu vào 118 tương ứng của dòng dữ liệu đầu vào 105 tương ứng được mã hóa theo cách tương ứng với phần mô tả

Fig.2, trong khi đồng thời xem lại phần đầu lát 40 để tính đến sự thay đổi chênh lệch POC của các ảnh 116 mà các ảnh của dòng bit video đầu vào giống nhau được ghép kênh vào đó. Các ảnh 116 thuộc về phần tham chiếu 112 không được xuất ra, như được nêu trên. Chúng là các ảnh không xuất ra. Video được đối chiếu được xác định bởi các ảnh đầu ra 116 thuộc về phần 114. Thiết bị 10 có thể tạo ra video được đối chiếu đơn giản bằng cách mã hóa thông tin phụ bù chuyển động cho các vùng khác nhau của các ảnh 116 thuộc về phần 114.

Trên Fig.12 và Fig.15, ví dụ, các ảnh 116 đã được minh họa là một phần của phần 114, tức là, các SLCP, được chia vào các vùng trong đó trong mỗi vùng, các vector chuyển động theo thời gian là không đổi trong vùng tương ứng, nhưng thông tin phụ bù chuyển động theo thời gian khác nhau giữa các vùng riêng lẻ. Trên Fig.17, ví dụ, vùng con 132 của ảnh 116 thuộc về phần 114, ví dụ, được mã hóa bằng cách sử dụng thông tin phụ bù chuyển động mà sao chép, bằng cách tịnh tiến, phần tương ứng có cùng kích thước từ các PSR 116 thuộc về dòng dữ liệu video đầu vào 105₁ có dòng bit đầu ra POC 1, ví dụ, khi vùng còn lại 134 của SLCP 116 này được mã hóa bằng cách sử dụng thông tin phụ bù chuyển động mà sao chép, ví dụ, bằng cách tịnh tiến phần tương ứng có cùng kích thước và hình dạng từ PSR 116 của dòng bit video đầu vào 105₂ khác. SLCP 116 của thời điểm tiếp theo, không được thể hiện trên Fig.17, có thể được mã hóa bằng cách sử dụng cùng thông tin phụ bù chuyển động và chia nhỏ thành các vùng 132 và 134 hoặc sử dụng thiết lập khác. Ví dụ, tưởng tượng việc sử dụng cùng phép mã hóa của các SLCP liên tiếp. Hơn nữa, tưởng tượng rằng một trong số các vùng 134 sử dụng vector chuyển động bằng không, tức là, nó đơn giản sao chép về không gian các phần được cùng vị trí của dòng dữ liệu video đầu vào tương ứng. Trong trường hợp này, dòng dữ liệu video được hợp thành như vậy sẽ dẫn đến sự trình bày/hiển thị của một trong số các dòng dữ liệu video đầu vào với một dòng dữ liệu video đầu vào khác được biểu diễn hoặc được chồng lấp tại vùng nhất định của dòng dữ liệu video đầu vào. Ví dụ này được thể hiện trên Fig.12 và Fig.14. Trong trường hợp sử dụng phép sao chép tịnh tiến nội dung ảnh của các PSR, thiết bị 10 có thể mã hóa các vùng 132 và 134 của các SLCP 116 theo cách được mô tả ở trên dựa vào Fig.3, cụ thể là bằng cách sử dụng phép dự đoán không gian và/hoặc sử dụng chế độ bỏ qua đối với bất kỳ khối nào theo sau khối thứ nhất của vùng là 132 và 134 tương ứng, ví dụ. Thiết bị 10 có thể thậm chí mã hóa

các SLCP bằng cách sử dụng dữ liệu phân dự đoán hoặc có thể mã hóa các vùng khác ngoài các vùng 132 và 134 được dự đoán theo thời gian trong các PSR, bằng cách sử dụng phép mã hóa ảnh tĩnh hoặc tương tự.

Khi sử dụng phương pháp tham chiếu được vá để ghép kênh các dòng dữ liệu video đầu vào như đã nêu trên Fig.14 và Fig.16, một vài khả năng của các vị trí dòng bit SLCP tồn tại, tức là, để định vị phần 114 của dòng bit video được hợp thành 108 với hai khả năng được minh họa trên Fig.18a và Fig.18b. Fig.18b thể hiện khả năng đã được mô tả trên Fig.16: các dòng dữ liệu video đầu vào 105 được vá theo không gian cùng nhau bởi thiết bị 10 cho đến khi các ảnh 116 đầu tiên có một ô cho mỗi dòng dữ liệu video đầu vào, và các SLCP được nằm rải rác giữa chúng và được hợp thành, bằng cách sử dụng phép dự đoán được bù chuyển động, trên cơ sở các ảnh tham chiếu được vá. Các SLCP 116 có kích thước tăng tương ứng với việc vá không gian của các dòng bit video đầu vào khác nhau, nhưng việc cắt một phần của các SLCP có thể được sử dụng để giảm kích thước của các ảnh sẽ được hiển thị và theo đó, kích thước của video được đối chiếu do đó được xác định bởi thiết bị 10.

Trong khi Fig.18b nhờ đó đan xen theo thời gian các ảnh 116 thuộc về phần 114 và các ảnh thuộc phần tham chiếu 112, theo Fig.18a, các SLCP, tức là, phần video được đối chiếu 114, được gắn theo không gian với các ảnh 116 của dòng bit video được hợp thành 108. Đáng lưu ý rằng vì phép dự đoán được bù chuyển động theo thời gian được sử dụng để hợp thành nội dung ảnh của các SLCP, độ trễ thời gian xảy ra giữa video được đối chiếu được xác định bằng phần 114, tức là, các SLCP, và nội dung được tham chiếu của các dòng dữ liệu video đầu vào 105.

Do đó, trong trường hợp trên Fig.18a, thiết bị 10 có thể hợp thành các ảnh 116 của các dòng dữ liệu video được hợp thành 108 để bao gồm một ô trên mỗi dòng dữ liệu video đầu vào 105, và ô bổ sung khác để truyền tải phần 114. Nếu một hoặc nhiều dòng dữ liệu video đầu vào đã được phân chia thành nhiều ô, thì một ô trên mỗi ô của các dòng dữ liệu video đầu vào có thể được biểu diễn trong dòng dữ liệu video được hợp thành 108. Khi một mặt sử dụng các tốc độ khung khác nhau của các dòng dữ liệu video đầu vào 105 và mặt khác là các SLCP, thì một số ô khi đó có thể được điền đầy với dữ liệu giả, điều này là đúng đối với bất kỳ dòng dữ liệu video đầu vào và các SLCP nào, có tốc độ khung thấp hơn.

Do đó, Fig.18a và Fig.18b thể hiện rằng dữ liệu SLCP có thể được chèn vào bởi thiết bị 10 dưới dạng các ảnh riêng lẻ, ví dụ với các trị số POC riêng biệt theo sau các PSR được tham chiếu hoặc các SLCP có thể được vá với dữ liệu của các PSR theo sau các PSR được tham chiếu.

Khi sử dụng phương pháp tham chiếu được vá, SLCP có thể được cắt để trở lại kích thước ảnh đầu ra được hình dung, ví dụ, kích thước ảnh là, ví dụ, một trong số các dòng đầu vào riêng lẻ.

Tóm tắt các hình vẽ từ Fig..17 đến Fig.18b, thiết bị 100 có thể hoạt động như sau:

Theo Fig.17, thiết bị 100 có thể điền đầy các ảnh $P_{i(N+K)+n}^{\text{composed}}$ của dòng dữ liệu video được hợp thành bằng cách thu thập và sao chép thông tin phụ bù chuyển động và dữ liệu phần dư dự đoán của các ảnh P_i^n của chỉ số i của các dòng dữ liệu video đầu vào n với $0 < n \leq N$ (N là số lượng dòng dữ liệu video đầu vào và K là số lượng ảnh - các ảnh trung gian của phần trung gian hoặc các ảnh thuộc về phần video (hoặc đầu ra) được đối chiếu - được chèn vào bởi thiết bị 100 mỗi N ảnh P_i^n có cùng chỉ số i), và hợp thành video được đối chiếu bằng cách tham chiếu, qua phép dự đoán theo thời gian được bù chuyển động, từ ảnh $P_{i(N+K)+N+k}^{\text{composed}}$, $0 < k \leq K$ (k chỉ các chuỗi của các ảnh được chèn) của dòng dữ liệu video được hợp thành, một hoặc nhiều ảnh $P_{i(N+K)+1}^{\text{composed}} \dots P_{i(N+K)+N}^{\text{composed}}$ của dòng dữ liệu video được hợp thành ngoài ra gồm có cho các ảnh của chuỗi $k > 1$ (nếu có), một cách tùy chọn, các ảnh $P_{i(N+K)+N+1}^{\text{composed}} \dots P_{i(N+K)+N+k-1}^{\text{composed}}$, tạo ra "tổng vùng được tham chiếu" cho ảnh $P_{i(N+K)+N+k}^{\text{composed}}$. Tham số K có thể là một, như được biểu thị trên Fig.17, nhưng để cho phép tốc độ khung SLCP cao hơn - K chuỗi có thể gồm có hai hoặc nhiều hơn hai chuỗi các ảnh đầu ra tạo ra phần đầu ra 114 - hoặc để đạt được phương pháp tham chiếu đa bước như được nêu sau đây - K chuỗi có thể bao gồm ít nhất một chuỗi các ảnh trung gian và ít nhất một chuỗi các ảnh đầu ra.

Cách khác, theo Fig.18a, thiết bị 100 điền đầy N ô T_n của ảnh $P_{(1+K)i}^{\text{composed}}$ của dòng dữ liệu video được hợp thành bằng cách thu thập và sao chép thông tin phụ bù chuyển động của các ảnh P_i^n của các dòng dữ liệu video đầu vào n , một cách lần lượt, và hợp thành video được đối chiếu bằng cách tham chiếu, qua phép dự đoán theo thời gian

được bù chuyển động, từ các ảnh $P_{(1+K)i+1}^{\text{composed}} \dots P_{(1+K)i+K}^{\text{composed}}$ của dòng dữ liệu video được hợp thành, một hoặc nhiều ô $T_1 \dots T_N$ của các ảnh $P_{(1+K)i}^{\text{composed}}$, ngoài ra gồm có cho các ảnh của chuỗi $k>1$ (nếu có), ngoài ra, các ảnh $P_{(1+K)i+1}^{\text{composed}} \dots P_{(1+K)i+k-1}^{\text{composed}}$, tạo ra "tổng vùng được tham chiếu" cho ảnh $P_{(1+K)i+k}^{\text{composed}}$. Tham số K có thể là một, như được biểu thị trên Fig.18a, nhưng để cho phép tốc độ khung SLCP cao hơn – K chuỗi có thể cũng gồm có hai hoặc nhiều hơn hai chuỗi các ảnh đầu ra tạo ra phần đầu ra 114 - hoặc để đạt được phương pháp tham chiếu đa bước như được nêu sau đây - K chuỗi có thể bao gồm ít nhất một chuỗi các ảnh trung gian và ít nhất một chuỗi các ảnh đầu ra.

Theo cách khác, thiết bị 10 điền đầy các ô T_N của ảnh P_i^{composed} của dòng dữ liệu video được hợp thành bằng cách thu thập và sao chép thông tin phụ bù chuyển động và dữ liệu phần dư dự đoán của ảnh P_i^n của dòng dữ liệu video đầu vào n và hợp thành video được đối chiếu bằng cách tham chiếu, qua phép dự đoán theo thời gian được bù chuyển động, từ ô T_{N+1} của ảnh P_i^{composed} của dòng dữ liệu video được hợp thành, các ô $T_1 \dots T_N$ của các ảnh $P_{i-1}^{\text{composed}}$ của dòng dữ liệu video được hợp thành, tạo ra "tổng vùng được tham chiếu" cho ô T_{N+1} của ảnh P_i^{composed} . Theo cách tương tự với trường hợp được ký hiệu bởi $K>1$ trên đây, nhiều hơn một ô có thể được chèn vào mỗi ảnh P_i^{composed} cho phép dự đoán đa bước như được mô tả cụ thể hơn dưới đây, tức là, để tạo ra phần trung gian của dòng dữ liệu video được hợp thành.

Sự tham chiếu trong tiến trình hợp thành video được đối chiếu có thể được thực hiện bởi thiết bị 100 bằng cách phân chia vùng đầu ra A của ảnh hoặc ô tương ứng của dòng dữ liệu video được hợp thành thành J vùng A_j , mỗi vùng gồm ít nhất một tập con của các vùng A_j này, $0 < j \leq J$, tham chiếu một hoặc hai (thậm chí nhiều hơn) phần trong số "tổng vùng được tham chiếu" tương ứng bằng cách sử dụng tập hợp $m(A_j)$ gồm một hoặc nhiều video chuyển động không đổi (với các ảnh tham chiếu kết hợp trong số "tổng vùng được tham chiếu"), tức là, $m(p)=m(q)$ cho tất cả các khối p và q nằm trong A_j , hoặc tập hợp gồm một hoặc nhiều vector chuyển động dần thay đổi theo chiều ngang và/hoặc chiều dọc, tức là $|m(p)-m(q)| < \text{ngưỡng}$ đối với tất cả các khối xung quanh p và q nằm trong A_j , để đạt được các hiệu quả phóng to hoặc thu nhỏ như được

mô tả dưới đây. Do đó, các phần phụ không gian trong số "tổng vùng được tham chiếu" có thể, trong trường hợp thứ nhất, được sao chép theo phép tịnh tiến và, trong trường hợp thứ hai, tịnh tiến theo nghĩa nội khối nhưng phóng to và/hoặc thu nhỏ theo nghĩa liên khối. Các vùng của ít nhất tập con của các vùng A_j có thể được hợp thành gồm các khối được mã hóa theo thời gian. Để mã hóa "vùng đầu ra", kích thước khối khả thi nhỏ nhất khả dụng để mã hóa thông tin phụ bù chuyển động có thể được chọn cho trường hợp thay đổi dần dần. Ít nhất một tập con của các vùng, về các đặc điểm và về số lượng của chúng, là không đổi theo thời gian cho chuỗi lần lượt gồm các ảnh $P_{i(N+K)+N+k}^{composed}$, các ảnh $P_{(l+K)j+k}^{composed}$ và ô T_{N+1} của các ảnh $P_i^{composed}$, hoặc có thể thay đổi dọc theo chuỗi (có chỉ số i). Vùng đầu ra có thể trùng với vùng ảnh trong số chuỗi các ảnh $P_{i(N+K)+N+k}^{composed}$, vùng ảnh của các ảnh $P_{(l+K)j+k}^{composed}$ hoặc vùng ô của ô T_{N+1} của các ảnh $P_i^{composed}$, và có thể là một phần của chúng với phần còn lại được cắt. Có thể có một vùng có vector chuyển động bằng không. Như được mô tả, một hoặc nhiều khu vực trong số các vùng A_j có thể có tập hợp $m(A_j)$ gồm nhiều hơn một vector chuyển động, tức là $m(A_j) = \{(m_1, \Delta P_1), \dots, (m_M, \Delta P_M)\}$ với m_l ($0 < l < M$) biểu thị ảnh tham chiếu vector chuyển động $P_{current-\Delta P_l}^{composed}$ là ảnh thứ ΔP_l^{th} của dòng dữ liệu video được hợp thành ngược dòng ảnh hiện thời bao gồm vùng hiện thời A_j . Trong trường hợp có hai vector chuyển động, tức là, $M = 2$, điều này có thể gọi là phép dự đoán đôi. Đối với mỗi vector chuyển động, thông tin phụ bù chuyển động có thể cũng bao gồm hệ số dự đoán/gán trọng số α_l dẫn đến phép dự đoán được bù chuyển động theo "tổng các bản sao của $P_{current-\Delta P_l}^{composed}$ tại vị trí được dịch chuyển tương đối m_l được gán trọng số α_l để thu được A_j ". Bằng cách này, sự chồng lấp giữa hai hoặc nhiều hơn hai dòng dữ liệu video đầu vào có thể đạt được tại vùng A_j . Bằng cách thay đổi dần dần các trọng số α_l theo thời gian trong chuỗi các vùng đầu ra cho A_j , sự pha trộn giữa hai dòng dữ liệu video đầu vào có thể đạt được.

Việc chồng lấp hoặc nằm ngoài tập con các vùng vừa được mô tả, có thể có phần không gian của "vùng đầu ra" mà trong đó phần video được đối chiếu 114 được mã hóa bằng cách sử dụng dữ liệu phần dư dự đoán. Trong phần không gian như vậy, nội dung đặc biệt như ảnh không động, biểu tượng dịch vụ hoặc tương tự, có thể được mã hóa.

Trong các vùng bên ngoài các vùng được bù chuyển động A_j , các chế độ dự đoán không gian có thể được sử dụng để giảm tiêu thụ bit cho việc mã hóa vùng đầu ra của phần video được đối chiếu 114. Do đó, có thể có vùng hoàn toàn được mã hóa nội ảnh nằm trong vùng đầu ra.

Để làm giảm tốc độ dữ liệu để mã hóa vùng đầu ra của phần video được đối chiếu 114, thiết bị 100 có thể sử dụng phép dự đoán theo thời gian của thông tin phụ bù chuyển động, cụ thể giữa các ảnh liên tiếp của phần video được đối chiếu 114, tức là, lần lượt giữa các ảnh liên tiếp trong số chuỗi các ảnh $P_{i(N+K)+N+k}^{\text{composed}}$, các ảnh liên tiếp trong số chuỗi các ảnh $P_{(l+K)+k}^{\text{composed}}$ và các ảnh liên tiếp của T_{N+1} trong số chuỗi các ảnh P_i^{composed} .

Sự điền đầy bằng cách thu thập và sao chép có thể được thực hiện bằng thiết bị 100 theo cách được nêu trên dựa vào Fig.3. Thiết bị 100, ví dụ, ngoài làm lại sự tham chiếu đến các ảnh tham chiếu trong các phần đầu lát khi truyền dẫn các lát từ các ảnh được sao chép của dòng dữ liệu video đầu vào đến các ô hoặc các ảnh tương ứng của dòng dữ liệu video được hợp thành, và ví dụ, điền đầy các vùng phần dư không được điền đầy trong ô hoặc ảnh tương ứng của dòng dữ liệu video được hợp thành bằng cách thu thập và sao chép dữ liệu giả như được mô tả trên đây dựa vào Fig.15.

Ngoài ra, trong trường hợp trên Fig.17, thiết bị 10 có thể báo hiệu rằng các ảnh $P_{i(N+K)+1}^{\text{composed}} \dots P_{i(N+K)+N}^{\text{composed}}$ là các ảnh không xuất ra trong khi các ảnh $P_{i(N+K)+N+k}^{\text{composed}}$ là các ảnh xuất ra hoặc các ảnh trung gian, trong đó các ảnh với $k=K$ có thể là các ảnh xuất ra. Trong trường hợp trên Fig.18a, thiết bị 10 có thể báo hiệu rằng các ảnh $P_{(l+K)i}^{\text{composed}}$ là các ảnh không xuất ra trong khi các ảnh $P_{(l+K)i+k}^{\text{composed}}$ là các ảnh xuất ra hoặc các ảnh trung gian, trong đó các ảnh với $k=K$ có thể là các ảnh xuất ra, và trong trường hợp trên Fig.18b, thiết bị 10 có thể báo hiệu rằng tất cả các dòng dữ liệu video được hợp thành là các ảnh xuất ra, tuy nhiên với việc báo hiệu rằng tất cả trừ ô T_{N+1} sẽ được cắt bỏ khi hiển thị/tái tạo dòng dữ liệu video được hợp thành. Việc báo hiệu cắt có thể cũng được áp dụng cho thiết bị 10 đối với Fig.18a: đơn thuần một phần trong số các ảnh đầu ra trong số các ảnh $P_{(l+K)i+k}^{\text{composed}}$ có thể được sử dụng để tạo ra đầu ra của dòng dữ liệu video được hợp thành, tức là, phần khác của các ảnh đầu ra có thể được cắt.

Như sẽ được nêu ra tiếp theo đối với Fig.21, thiết bị hợp thành video có thể được tạo cấu hình để điền đầy phần trung gian của các ảnh của dòng dữ liệu video được hợp thành bằng cách tham chiếu, qua phép dự đoán theo thời gian được bù chuyển động, phần tham chiếu 112 của dòng dữ liệu video được hợp thành, trong đó video được đối chiếu tạo ra phần 108 của các ảnh của dòng dữ liệu video được hợp thành được điền đầy với video được đối chiếu bằng cách tham chiếu, qua phép dự đoán theo thời gian được bù chuyển động, phần thứ nhất của dòng dữ liệu video được hợp thành một cách gián tiếp qua phần trung gian. Ví dụ, bên cạnh các ô T_1 đến T_{N+1} đã được nêu, các ô khác như các ô T_{N+2} có thể được tạo ra bởi thiết bị 100 cùng với ô T_{N+1} của ảnh P_i^{composed} của dòng dữ liệu video được hợp thành tham chiếu một hoặc nhiều ô $T_1 \dots T_N$ của ảnh $P_{i-2}^{\text{composed}}$ của dòng dữ liệu video được hợp thành một cách gián tiếp qua ô T_{N+2} của ảnh $P_{i-1}^{\text{composed}}$ của dòng dữ liệu video được hợp thành một cách lần lượt, tham chiếu một hoặc nhiều ô trong số các ô $T_1 \dots T_N$ của ảnh $P_{i-2}^{\text{composed}}$ của dòng dữ liệu video được hợp thành.

Nói cách khác, vì chỉ SLCP của phần 114 sẽ được xuất ra bởi bộ giải mã và/hoặc được hiển thị trong khi các PSR được sử dụng chỉ để tham chiếu bởi SLCP và - nếu có - phần trung gian của dòng dữ liệu video được hợp thành và không được xuất ra, bộ mã hóa-giải mã được áp dụng làm cơ sở cho các dòng dữ liệu video đầu vào 105 cũng như dòng dữ liệu video được hợp thành 108 nên hỗ trợ khái niệm các ảnh khả dụng để tham chiếu mà ví dụ, không được xuất ra hoặc bất kỳ cơ cấu có thể so sánh được nên được áp dụng trên lớp hệ thống. Cách khác, đối với bộ mã hóa-giải mã mà thiếu đặc điểm này, như H.264/AVC, ví dụ, phương pháp tham chiếu được và có thể được sử dụng thay thế và dữ liệu của SLCP và các PSR theo sau các PSR được tham chiếu có thể được và vào các ảnh đơn như được mô tả trên đây.

Vì chỉ SLCP của phần 114 được dự định để xuất ra bởi bộ giải mã, với bất kỳ phương pháp được trình bày và tổ hợp của chúng, có thể mong muốn cắt các ảnh được giải mã sẽ được xuất ra bởi bộ giải mã, hoặc để loại bỏ nội dung ảnh của các PSR, hoặc để điều chỉnh các kích thước không gian của SLCP được giải mã đã được mang dưới dạng ảnh đơn lẻ như được mô tả trên đây.

Nhiều khía cạnh khác nhau thúc đẩy việc tạo cấu trúc SLCP theo ô hoặc lát. Đầu tiên, SLCP có thể tuân theo cấu trúc ô/lát được sử dụng trong một hoặc nhiều dòng đầu

vào của cấu trúc của các ảnh tham chiếu được vá, nhờ đó không yêu cầu sự khởi tạo mới của các cấu trúc hạ tầng giải mã song song cho mỗi SLCP, tức là, đối với mỗi ảnh, ít nhất phần không gian của ảnh thuộc về phần 114. Thứ hai, việc nén hiệu quả dữ liệu lát SLCP có thể thúc đẩy cấu trúc ô hoặc lát theo sau sự phân bố các vector dự đoán để các vùng là bằng nhau, tương tự hoặc các vector dự đoán liên quan có thể được mã hóa một cách hiệu quả thành ô và/hoặc lát đơn.

Có các trường hợp sử dụng trong đó việc pha trộn hai video đầu vào (theo các phần hoặc toàn bộ) cho SLCP thu được (một lần nữa, theo các phần hoặc toàn bộ) có thể được mong muốn. Trường hợp sử dụng minh họa được mô tả trên đây, chẳng hạn như tham chiếu hai phần trong số "tổng vùng được tham chiếu" tương ứng được cho trên Fig.19 với ứng dụng đầu thu giải mã tín hiệu chồng lấp giao diện đồ họa người dùng (Graphical User Interface - GUI) trên nội dung video trong miền được nén. Trong ví dụ này, dòng đầu vào 1 được chồng lấp với GUI dưới dạng dòng đầu vào 2 khác để hợp thành.

Hình tam giác nét đứt ở dưới trong sự hợp thành trên Fig.19 được đề cập đến như vùng được pha trộn anpha sau đây. Quy trình dự đoán được gán trọng số được sử dụng để tạo ra hiệu quả mong muốn cho các vùng được pha trộn anpha của SLCP. Các trị số mẫu trong vùng được pha trộn anpha được dự đoán từ nhiều ảnh tham chiếu (khi sử dụng phương pháp tham chiếu B2B) hoặc nhiều vùng không gian của ảnh tham chiếu đơn (khi sử dụng phương pháp tham chiếu được vá), mỗi vùng có trọng số riêng. Tổ hợp một số vùng được pha trộn anpha với SLCP từ các dòng đầu vào giống hoặc khác nhau là khả thi, như sự chồng lấp GUI và kênh, trạm hoặc biểu tượng dịch vụ.

Dấu hiệu khác của các SLCP được mô tả trên đây là việc thay đổi dần dần các vector chuyển động được minh họa trên Fig.20 dựa trên các ứng dụng ảnh trong ảnh. Khái niệm này kết hợp sự định lại kích cỡ không gian của nội dung ảnh dòng đầu vào để hợp thành. Điều này đạt được bằng cách cung cấp các vector dự đoán trên khối mã hóa khả thi nhỏ nhất và/hoặc độ chi tiết của khối đơn vị dự đoán. Các tọa độ vector dự đoán được điều chỉnh trong tiến trình đi qua vùng tương ứng nằm trong SLCP để lấy mẫu lại dòng bit đầu vào đích. Kết quả trong SLCP được giải mã là dạng biểu diễn khác về không gian của một phần hoặc toàn bộ nội dung ảnh dòng bit đầu vào tương ứng.

Đối với mỗi vị trí mẫu (tức là, khối mã hóa và/hoặc đơn vị dự đoán) của vùng được lấy mẫu lại trong SLCP, phép dự đoán được gán trọng số từ nhiều khối mã hóa và/hoặc đơn vị dự đoán (ví dụ, xung quanh) của các ảnh đầu vào có thể được áp dụng để tăng chất lượng của vùng ảnh được lấy mẫu lại tương ứng.

Thông thường, nên tránh xử lý SLCP trong bộ lọc trong vòng như bộ lọc tách khối để không dẫn đến các chênh lệch so với các ảnh của các dòng bit đầu vào. Tuy nhiên, để giảm sự suy giảm chất lượng có thể gây ra bởi quy trình lấy mẫu lại, bộ lọc trong vòng bổ sung, như bộ lọc tách khối hoặc bộ lọc SAO trong HEVC, có thể được sử dụng trên toàn bộ SLCP hoặc các vùng được lấy mẫu con khi thích hợp.

Trong một số tình huống, sự tạo ra sự hợp thành mong muốn từ các dòng bit đầu vào trong dòng bit đầu ra đơn có thể yêu cầu một số bước dự đoán, tức là, một số ảnh không xuất ra bổ sung sẽ được thêm vào dòng bit đầu ra. Các ảnh trung gian được dự đoán bằng cách sử dụng các PSR và/hoặc các ảnh trung gian được tạo ra trước đó cho đến khi sự hợp thành mong muốn cuối cùng có thể được tạo ra trong SLCP tham chiếu một hoặc nhiều ảnh trung gian này và/hoặc các PSR khi chỉ có SLCP được dự định để xuất ra.

Ví dụ thứ nhất cho các tình huống như vậy là việc lấy mẫu con lặp lại, trong đó phần trung gian thứ nhất của dòng dữ liệu video được hợp thành, nội dung ảnh dòng bit đầu vào được lấy mẫu con theo không gian theo hướng nằm ngang sử dụng phép dự đoán đôi được gán trọng số và trong bước thứ hai, nội dung của ảnh trung gian này được lấy mẫu con theo không gian theo hướng dọc sử dụng phép dự đoán đôi được gán trọng số một lần nữa và sự hợp thành SLCP cuối cùng sẽ được tạo ra dưới dạng đầu ra như được minh họa trên Fig.21.

Ví dụ khác của trường hợp này là việc sử dụng phép dự đoán đa hợp trong VP9 để tạo ra hiệu quả pha trộn anpha và sử dụng ảnh trung gian thu được như tham chiếu cho SLCP để được xuất ra.

Như các phần trước đã thể hiện, nhiều ứng dụng và trường hợp sử dụng có thể được thực hiện bằng cách điều chỉnh các vector dự đoán của SLCP phù hợp. Ví dụ khác ngoài hợp thành nội dung video điển hình là cung cấp các chữ cái trong bảng chữ cái hoặc tập hợp của chúng dưới dạng nội dung ảnh của (các) dòng bit đầu vào để soạn thảo

các bản tin trong SLCP bằng cách sử dụng các vector dự đoán đúng đến vùng ảnh dòng bit đầu vào mô tả các chữ cái riêng lẻ.

Trong khi tất cả các ví dụ đã cho sử dụng cấu trúc tham chiếu cụ thể, các khái niệm được trình bày là áp dụng được cho nhiều cấu trúc tham chiếu. Cũng cần đề cập rằng khái niệm SLCP được mô tả như vậy cho phép xử lý nhiều hơn hai dòng bit video đầu vào.

Tương tự đối với các GRP, dữ liệu lát SLCP có thể được tiền mã hóa để chèn vào các dòng bit tùy ý vì nội dung của chúng phụ thuộc vào các tham số mức cao như các chiều ảnh và sự thay thế.

Như đã nêu trên đối với thiết bị hợp thành video trên các hình vẽ từ Fig.1 đến Fig.10, thiết bị hợp thành video 100 được mô tả trên đây đối với các hình vẽ từ Fig.11 đến Fig.21, có thể được sử dụng trong hệ thống như được thể hiện trên Fig.5, tuy nhiên, trong đó nhiều bộ mã hóa video có thể, như được mô tả trên đây, có nhiều tự do hơn trong việc tạo ra các dòng dữ liệu video đầu vào như được mô tả đối với các hình vẽ từ Fig.11 đến Fig.21. Yêu cầu 60 có thể được xử lý bởi thiết bị 100 như các yêu cầu thay đổi các SLCP, tức là, sự hợp thành các dòng dữ liệu video đầu vào tham chiếu đến video được đối chiếu cuối cùng. Các yêu cầu có thể bắt nguồn từ máy khách cũng bao gồm bộ giải mã hoặc bộ vận hành máy chủ bao gồm thiết bị 100.

Việc báo hiệu trong băng hoặc ngoài băng của các ràng buộc mã hóa được thỏa mãn để cần thiết để vá, việc chèn các GRP và/hoặc sự hợp thành qua các SLCP như được mô tả cho phép hệ thống thanh toán và/hoặc phát hiện rằng các dòng đầu vào có thể được vá cho quy trình khác như được mô tả trên đây. Do đó, việc báo hiệu, ví dụ, trong SDP để truyền theo dòng RTP hoặc trong MPD để truyền theo dòng trên cơ sở DASH hoặc ngoài ra có thể được sử dụng như mô tả trên đây. Fig.22 mô tả minh họa việc báo hiệu tương ứng dưới dạng bản tin SEI.

composition_enabled_type biểu thị loại hợp thành (xử lý) có thể được áp dụng cho dòng bit hiện thời để sử dụng cho việc hợp thành dòng bit đầu ra trong đó dòng bit hiện thời là dòng bit phụ hoặc một phần (ô/lát) của dòng bit được hợp thành **composition_enabled_type** với trị số bằng 0 biểu thị rằng dòng bit hiện thời có thể được vá như được mô tả trong [1] với các dòng bit khác thỏa mãn các ràng buộc được mô tả

trong [1] với sự căn chỉnh IDR. `composition_enabled_type` với trị số bằng 1 biểu thị rằng dòng bit hiện thời có thể được vá như được mô tả trong [1] với các dòng bit khác thỏa mãn các ràng buộc được mô tả trong [1] nhưng các ảnh IDR có thể được chuyển đổi thành các lát I của ảnh kế tiếp và có thể được hợp nhất với các lát không phải I, tức là, `composition_enabled_type` với trị số bằng 1 biểu thị rằng các IDR hoặc IRAP không kích hoạt SPS với các trị số khác ngoài SPS hoạt động trước đó nếu sự kích hoạt tập hợp tham số xảy ra. `composition_enabled_type` với trị số là 2 ngoài ra biểu thị rằng TVMP còn bị hạn chế và rằng các ảnh có `temporal_id_plus1` nhỏ hơn hoặc bằng `max_temporal_id_plus1_TMVP_disabled` không được sử dụng làm tham chiếu cho TMVP.

`max_temporal_id_plus1_TMVP_disabled` biểu thị trị số đường biên của `temporal_id_plus1` của các ảnh sẽ không được sử dụng làm tham chiếu cho TMVP.

Cùng với bản tin SEI được đề xuất, các SEI hiện có có thể được sử dụng để thỏa thuận khiến việc xử lý dòng bit dễ dàng hơn. Một ví dụ là `structure_of_pictures_info` SEI trong đó cấu trúc GO của các dòng bit khác nhau được tóm tắt và điều này có thể được sử dụng để nhận dạng các điểm chuyển đổi để chèn GRP. Bất kỳ ảnh mà tại thời điểm giải mã danh sách `RefPicSetStCurrBefore`, `RefPicSetStCurrAfter`, hoặc `RefPicSetLtCurr`, nói cách khác, bộ đệm ảnh tham chiếu bộ giải mã, chỉ chứa các ảnh có `TemporalID` thấp hơn hoặc bằng trị số của **`max_temporal_id_plus1_TMVP_disabled`** trừ 1 có thể phục vụ như điểm chuyển đổi để chèn GRP.

Theo cách khác, các bản tin SEI nằm trong các dòng dữ liệu đầu vào có thể được sử dụng để báo hiệu các ảnh mà có thể đóng vai trò như các điểm chuyển đổi để chèn các GRP. Trên dòng bit có SEI cụ thể này, thì sẽ không có ảnh nào nằm trong `RefPicSetStCurrBefore`, `RefPicSetStCurrAfter`, hoặc `RefPicSetLtCurr` được sử dụng làm tham chiếu cho TMVP của các ảnh theo sau trong thứ tự giải mã.

Thậm chí cách khác, trị số biểu thị loại đơn vị có thể được sử dụng để báo hiệu cho ràng buộc nêu trên theo cách tương tự việc báo hiệu đơn vị NAL STSA. Trên dòng bit xuất hiện ảnh A của trị số biểu thị loại đơn vị NAL cụ thể này, sẽ không có ảnh nào nằm trong `RefPicSetStCurrBefore`, `RefPicSetStCurrAfter`, hoặc `RefPicSetLtCurr` được

sử dụng làm tham chiếu cho TMVP của các ảnh theo sau trong thứ tự giải mã và gồm có ảnh A.

Một cách tương tự, SDP hoặc MPD hoặc bất kỳ dạng khác của cơ cấu báo hiệu có thể gồm có việc báo hiệu tương đương như SEI được đề xuất để truyền tải thông tin này ra khỏi băng.

Mặc dù một vài khía cạnh đã được mô tả trong ngữ cảnh của thiết bị, rõ ràng rằng các khía cạnh này cũng thể hiện mô tả của phương pháp tương ứng, trong đó khối hoặc thiết bị tương ứng với bước xử lý của phương pháp hoặc đặc điểm của bước xử lý của phương pháp. Tương tự, các khía cạnh được mô tả trong ngữ cảnh của bước phương pháp cũng biểu diễn sự mô tả của khối hoặc mục hoặc dấu hiệu tương ứng của thiết bị tương ứng. Một vài hoặc tất cả các bước phương pháp có thể được thực thi bởi (hoặc sử dụng) thiết bị phần cứng, tương tự ví dụ, bộ vi xử lý, máy tính có thể lập trình hoặc mạch điện tử. Theo một số phương án, một số một hoặc nhiều bước phương pháp quan trọng nhất có thể được thực hiện bởi thiết bị như vậy.

Phụ thuộc vào các yêu cầu thực hiện nhất định, các phương án của sáng chế có thể được thực hiện trong phần cứng hoặc trong phần mềm. Phương án có thể được thực hiện sử dụng phương tiện lưu trữ số, ví dụ, đĩa mềm, DVD, Blu-Ray, CD, ROM, PROM, EPROM, EEPROM hoặc bộ nhớ FLASH, có các tín hiệu điều khiển có thể đọc được bằng điện được lưu trữ trên đó, kết hợp (hoặc có thể kết hợp) với hệ thống máy tính lập trình được sao cho phương pháp tương ứng được thực hiện. Do đó, phương tiện lưu trữ số có thể đọc được bằng máy tính.

Một số phương án theo sáng chế bao gồm sóng mang dữ liệu có các tín hiệu điều khiển có thể đọc được bằng điện tử, mà có khả năng kết hợp với hệ thống máy tính có thể lập trình được, sao cho một trong số các phương pháp được mô tả ở đây được thực hiện.

Nói chung, các phương án của sáng chế có thể được thực thi dưới dạng sản phẩm chương trình máy tính với mã chương trình, mã chương trình có thể hoạt động để thực hiện một trong các phương pháp khi sản phẩm chương trình máy tính chạy trên máy tính. Mã chương trình có thể ví dụ được lưu trữ trên sóng mang máy đọc được.

Các phương án khác bao gồm chương trình máy tính để thực hiện một trong các

phương pháp được mô tả ở đây, được lưu trữ trên sóng mang có thể đọc được bằng máy.

Nói cách khác, do đó, phương án của phương pháp theo sáng chế là, chương trình máy tính có mã chương trình để thực hiện một trong số các phương pháp được mô tả ở đây, khi chương trình máy tính chạy trên máy tính.

Phương án khác của các phương pháp theo sáng chế là, sóng mang dữ liệu (hoặc phương tiện lưu trữ số, hoặc phương tiện đọc được bằng máy tính) gồm có, đã được ghi lại trên đó, chương trình máy tính để thực hiện một trong các phương pháp được mô tả ở đây. Sóng mang dữ liệu, phương tiện lưu trữ số hoặc phương tiện được ghi thường là hữu hình và/hoặc bất biến.

Do đó, phương án khác của phương pháp theo sáng chế là dòng dữ liệu hoặc chuỗi các tín hiệu biểu diễn chương trình máy tính để thực hiện một trong số các phương pháp được mô tả ở đây. Ví dụ, có thể tạo cấu hình dòng dữ liệu hoặc chuỗi tín hiệu để được truyền thông qua sự kết nối truyền thông dữ liệu, ví dụ thông qua mạng Internet.

Phương án khác gồm có phương tiện xử lý, ví dụ máy tính, hoặc thiết bị logic lập trình được, được tạo cấu hình để hoặc được làm thích ứng để thực hiện một trong số các phương pháp được mô tả ở đây.

Phương án khác bao gồm máy tính có chương trình máy tính được cài đặt trên máy tính để thực hiện một trong số các phương pháp được mô tả ở đây.

Phương án khác nữa theo sáng chế bao gồm thiết bị hoặc hệ thống được tạo cấu hình để chuyển (ví dụ, bằng điện tử hoặc quang học) chương trình máy tính để thực hiện một trong các phương pháp được mô tả ở đây đến bộ thu. Bộ thu có thể là, ví dụ, máy tính, thiết bị di động, thiết bị nhớ hoặc tương tự. Thiết bị hoặc hệ thống có thể, ví dụ, bao gồm máy chủ tệp tin để truyền tải chương trình máy tính đến bộ thu.

Theo một số phương án, thiết bị logic lập trình được (ví dụ, mảng công lập trình được theo trường) có thể được sử dụng để thực hiện một số hoặc tất cả các chức năng của các phương pháp được mô tả ở đây. Trong một số phương án, mảng công lập trình được theo trường có thể kết hợp với bộ vi xử lý để thực hiện một trong số các phương pháp được mô tả ở đây. Nói chung, các phương pháp ưu tiên được thực hiện bởi thiết bị phần cứng bất kỳ.

Các phương án được mô tả bên trên chỉ mang tính minh họa cho các nguyên lý của sáng chế. Cần hiểu rằng các sửa đổi và biến đổi của các phương án và các chi tiết được mô tả ở đây sẽ là rõ ràng đối với người có hiểu biết trung bình trong lĩnh vực kỹ thuật tương ứng. Do đó, mục đích của sáng chế là chỉ giới hạn bởi phạm vi của các điểm yêu cầu bảo hộ sắp đưa ra dưới đây và không bị giới hạn bởi các chi tiết cụ thể được biểu diễn bằng bản mô tả và giải thích của các phương án ở đây.

YÊU CẦU BẢO HỘ

1. Hệ thống hợp thành video bao gồm:

ít nhất một bộ mã hóa video được tạo cấu hình để cung cấp ít nhất một dòng dữ liệu video đầu vào; và

thiết bị hợp thành video được tạo cấu hình để hợp thành dòng dữ liệu video được hợp thành bằng cách sử dụng ít nhất một dòng dữ liệu video đầu vào, dòng dữ liệu video được hợp thành và ít nhất một dòng dữ liệu video đầu vào được mã hóa bằng cách sử dụng phép dự đoán theo thời gian được bù chuyển động,

trong đó thiết bị hợp thành video được tạo cấu hình để hợp thành dòng dữ liệu video được hợp thành bằng cách:

tạo ra chuỗi các ảnh kế thừa của dòng dữ liệu video được hợp thành với việc điền đầy phần không gian của các ảnh kế thừa bằng cách thu thập và sao chép thông tin phụ bù chuyển động và dữ liệu phần dư dự đoán của ít nhất một dòng dữ liệu video đầu vào, trong đó vị trí không gian của phần không gian thay đổi theo thời gian từ một ảnh kế thừa đến ảnh tiếp theo trong chuỗi các ảnh kế thừa, và

chèn ít nhất một ảnh không xuất ra vào dòng dữ liệu video được hợp thành giữa các ảnh kế thừa trong số chuỗi các ảnh kế thừa, ít nhất một ảnh không xuất ra tham chiếu, qua phép dự đoán theo thời gian được bù chuyển động, tập con của các ảnh kế thừa.

2. Hệ thống theo điểm 1, trong đó ít nhất một bộ mã hóa video được tạo cấu hình để cung cấp ít nhất một dòng dữ liệu video đầu vào bằng việc hạn chế thông tin phụ bù chuyển động của ít nhất một dòng dữ liệu video đầu vào sao cho các ảnh của ít nhất một dòng dữ liệu video đầu vào không, qua phép dự đoán theo thời gian được bù chuyển động, đứng trước các ảnh của ít nhất một dòng dữ liệu video đầu vào tại các phần mở rộng vượt qua đường biên của vùng không gian tĩnh được định trước của ít nhất một dòng dữ liệu video đầu vào.

3. Hệ thống theo điểm 1, trong đó hệ thống bao gồm nhiều bộ mã hóa video bao gồm ít

nhất một bộ mã hóa video và ít nhất một bộ mã hóa video được tạo cấu hình để cung cấp ít nhất một dòng dữ liệu video đầu vào sử dụng cấu trúc GOP dự đoán theo thời gian được đồng bộ hóa trong số nhiều bộ mã hóa video.

4. Hệ thống theo điểm 1, trong đó ít nhất một bộ mã hóa video được tạo cấu hình để:

cung cấp ít nhất một dòng dữ liệu video đầu vào bằng cách sử dụng mã hóa thứ bậc theo thời gian và phép dự đoán thông tin phụ bù chuyển động theo thời gian,

cung cấp ít nhất một dòng dữ liệu video đầu vào sao cho ID lớp thứ bậc theo thời gian phạm vi ảnh được kết hợp với mỗi ảnh trong số các ảnh của ít nhất một dòng dữ liệu video đầu vào, và

hạn chế cho một hoặc nhiều ID mức thứ bậc theo thời gian mà các ảnh của một hoặc nhiều ID mức thứ bậc theo thời gian tương ứng, phép dự đoán thông tin phụ bù chuyển động theo thời gian đến mức các ảnh của một hoặc nhiều ID mức thứ bậc theo thời gian tương ứng không được tham chiếu, qua phép dự đoán thông tin phụ bù chuyển động theo thời gian, bằng bất kỳ ảnh khác; và

cung cấp cú pháp mức cao của ít nhất một dòng dữ liệu video đầu vào với bộ phân biệt ID lớp thứ bậc theo thời gian vượt qua phạm vi ảnh đảm bảo cho một hoặc nhiều ID mức thứ bậc theo thời gian mà các ảnh đầu vào của một hoặc nhiều ID mức thứ bậc theo thời gian tương ứng không được tham chiếu, qua phép dự đoán thông tin phụ bù chuyển động theo thời gian, bằng bất kỳ ảnh khác.

5. Hệ thống theo điểm 1, trong đó thiết bị hợp thành video được tạo cấu hình để chèn một hoặc nhiều ảnh không xuất ra tại thời điểm tại đó phần không gian thay đổi về không gian sao cho mỗi ảnh trong số một hoặc nhiều ảnh không xuất ra thay thế ảnh tham chiếu trong số các ảnh kế thừa trước thời điểm được tham chiếu, qua phép dự đoán theo thời gian được bù chuyển động, bởi thông tin phụ bù chuyển động bằng cách thu thập và sao chép mà phần không gian của bất kỳ ảnh trong số các ảnh kế thừa sau thời điểm được điền đầy.

6. Hệ thống theo điểm 5, trong đó thiết bị hợp thành video được tạo cấu hình để chèn một hoặc nhiều ảnh không xuất ra sao cho mỗi ảnh không xuất ra tham chiếu, qua phép

dự đoán theo thời gian được bù chuyển động, tại phần không gian tương ứng về không gian với phần không gian mà phần không gian thay đổi tại thời điểm, phần không gian của ảnh tham chiếu được thay thế bởi ảnh không xuất ra tương ứng.

7. Hệ thống theo điểm 5, trong đó thiết bị hợp thành video được tạo cấu hình để chèn tại thời điểm tại đó phần không gian thay đổi về không gian từ ảnh thứ nhất trong số các ảnh kế thừa đến ảnh thứ hai trong số các ảnh kế thừa, đứng ngay sau ảnh thứ nhất trong số các ảnh kế thừa trong chuỗi các ảnh kế thừa, nhiều ảnh không xuất ra vì có các ảnh tham chiếu trong số các ảnh kế thừa bao gồm và đứng trước ảnh thứ nhất trong số các ảnh kế thừa, được tham chiếu, qua phép dự đoán theo thời gian được bù chuyển động, bằng thông tin phụ bù chuyển động bằng cách thu thập và sao chép mà phần không gian của bất kỳ ảnh trong số các ảnh kế thừa bao gồm và đứng sau ảnh thứ hai trong số các ảnh kế thừa được điền đầy.

8. Hệ thống theo điểm 1, trong đó thiết bị hợp thành video được tạo cấu hình để ghép kênh nhiều dòng dữ liệu video đầu vào, gồm có ít nhất một dòng dữ liệu video đầu vào, thành chuỗi các ảnh thứ nhất sao cho chuỗi các ảnh kế thừa được chia nhỏ theo không gian thành các ô, và sao cho mỗi ô được tạo ra từ một trong số nhiều dòng dữ liệu video đầu vào, được kết hợp với ô tương ứng, bằng cách thu thập và sao chép thông tin phụ bù chuyển động và dữ liệu phần dư dự đoán của dòng dữ liệu video đầu vào được kết hợp, trong đó phần không gian được xác định bởi ô mà ít nhất một dòng dữ liệu video đầu vào được kết hợp với ô đó và vị trí không gian của phần không gian thay đổi theo thời gian từ một ảnh kế thừa sang ảnh tiếp theo trong chuỗi các ảnh kế thừa đáp lại sự thay đổi theo thời gian trong sự kết hợp giữa một bên là các ô và một bên là nhiều dòng dữ liệu video đầu vào.

9. Hệ thống theo điểm 8, trong đó thiết bị hợp thành video được tạo cấu hình để chèn các ảnh không xuất ra tại các thời điểm mà tại đó sự kết hợp giữa một bên là các ô và một bên là nhiều dòng dữ liệu video đầu vào thay đổi, các ảnh không xuất ra bao gồm thông tin phụ bù chuyển động di chuyển vùng ảnh của ô được kết hợp với một dòng được định trước trong số nhiều dòng dữ liệu video đầu vào trước sự thay đổi tương ứng đến vùng ảnh của ô được kết hợp với dòng dữ liệu video đầu vào được định trước sau sự thay đổi.

10. Hệ thống theo điểm 1, trong đó thiết bị hợp thành video được tạo cấu hình để điều khiển sự thay đổi theo thời gian của vị trí không gian của phần không gian thay đổi theo thời gian đáp lại các yêu cầu bên ngoài để thay đổi.

11. Hệ thống theo điểm 1, hệ thống được tạo cấu hình để tạo ra chuỗi các ảnh kế thừa của dòng dữ liệu video được hợp thành bằng cách điền đầy phần không gian của các ảnh kế thừa sao cho phần không gian của mỗi ảnh trong số các ảnh kế thừa được điền đầy bằng cách thu thập và sao chép thông tin phụ bù chuyển động và dữ liệu phần dư dự đoán của chính xác một ảnh trong số chuỗi các ảnh đầu vào của ít nhất một dòng dữ liệu video đầu vào.

12. Hệ thống theo điểm 11, trong đó thiết bị hợp thành video được tạo cấu hình để thực thi yêu cầu bên ngoài để thay đổi khi bắt gặp ảnh tiếp theo liên tiếp trong số chuỗi các ảnh đầu vào mà tại đó không có tập hợp hiện thời gồm các ảnh tham chiếu nào được tham chiếu, qua phép dự đoán thông tin phụ bù chuyển động theo thời gian, bằng bất kỳ ảnh đầu vào kế tiếp nào.

13. Hệ thống theo điểm 12, trong đó thiết bị hợp thành video được tạo cấu hình để phát hiện sự bắt gặp ảnh tiếp theo liên tiếp trong số chuỗi các ảnh đầu vào mà tại đó không có tập hợp các ảnh tham chiếu nào được tham chiếu, qua phép dự đoán thông tin phụ bù chuyển động theo thời gian, bằng bất kỳ ảnh đầu vào kế tiếp nào trên cơ sở ID lớp thứ bậc theo thời gian phạm vi ảnh của các ảnh đầu vào và bộ phân biệt ID lớp thứ bậc theo thời gian vượt qua phạm vi ảnh bao gồm cú pháp mức cao của ít nhất một dòng dữ liệu video đầu vào đảm bảo cho một hoặc nhiều ID mức thứ bậc theo thời gian rằng các ảnh đầu vào của một hoặc nhiều ID mức thứ bậc theo thời gian tương ứng không được tham chiếu, qua phép biểu diễn thông tin phụ bù chuyển động theo thời gian, bằng bất kỳ ảnh đầu vào khác.

14. Hệ thống theo điểm 1, trong đó, thông tin phụ bù chuyển động và dữ liệu phần dư dự đoán được mã hóa entropy, và việc thu thập và sao chép được thực hiện mà không có sự giải mã entropy.

15. Hệ thống theo điểm 1, trong đó việc tạo ra chuỗi các ảnh kế thừa của dòng dữ liệu video được hợp thành bao gồm bước thu thập và sao chép các lát của ít nhất một dòng

dữ liệu video đầu vào với việc để dữ liệu tải tin của các lát không thay đổi và sửa dữ liệu trong phần đầu lát.

16. Hệ thống theo điểm 15, trong đó dữ liệu trong phần đầu lát bao gồm địa chỉ lát và/hoặc dữ liệu đếm thứ tự ảnh và/hoặc các trị số chênh lệch đếm thứ tự ảnh tham chiếu và/hoặc tham chiếu đến tập hợp các trị số chênh lệch đếm thứ tự ảnh tham chiếu và/hoặc tham số lượng tử hóa được mã hóa vi sai.

17. Hệ thống theo điểm 1, trong đó ít nhất một ảnh không xuất ra là không có dữ liệu phần dư dự đoán.

18. Hệ thống theo điểm 1, trong đó thiết bị hợp thành video được tạo cấu hình để chen ít nhất một ảnh không xuất ra sao cho ít nhất một ảnh không xuất ra tham chiếu, qua phép dự đoán theo thời gian được bù chuyển động, tập con gồm các ảnh kế thừa sao cho phần không gian của các ảnh đầu vào trong số chuỗi các ảnh đầu vào của ít nhất một dòng dữ liệu video đầu vào được sao chép, bằng cách tịnh tiến, vào phần không gian của ít nhất một ảnh không xuất ra.

19. Hệ thống theo điểm 18, trong đó dòng dữ liệu video được hợp thành và ít nhất một dòng dữ liệu video đầu vào được mã hóa bằng cách sử dụng phép dự đoán theo thời gian được bù chuyển động theo mức độ chi tiết khối và trong đó thiết bị hợp thành video được tạo cấu hình sao cho ít nhất một phần không gian của ảnh không xuất ra gồm số lượng các khối mà vectơ chuyển động của chúng bằng nhau và được mã hóa bên trong dòng dữ liệu video được hợp thành cho chỉ một khối trong số nhiều khối và sử dụng phép dự đoán không gian cho khối trong số nhiều khối khác một khối đó.

20. Hệ thống theo điểm 19, trong đó thiết bị hợp thành video được tạo cấu hình sao cho bất kỳ khối nào khác một khối đó được mã hóa bằng cách sử dụng chế độ bỏ qua báo hiệu sự vắng mặt của dữ liệu phần dư dự đoán và sự chấp nhận thông tin phụ bù chuyển động từ khối xung quanh.

21. Hệ thống theo điểm 1, trong đó thiết bị hợp thành video được tạo cấu hình để xây dựng các tập hợp tham số mới cho dòng dữ liệu video được hợp thành trên cơ sở các tập hợp tham số của ít nhất một dòng dữ liệu đầu vào video.

22. Phương tiện lưu trữ số bất biến lưu trữ dòng dữ liệu video được mã hóa theo thời gian theo thứ bậc được mã hóa bằng cách sử dụng phép dự đoán theo thời gian được bù chuyển động và phép dự đoán thông tin phụ bù chuyển động, trong đó dòng dữ liệu video được mã hóa theo thời gian theo thứ bậc báo hiệu ID lớp thứ bậc theo thời gian phạm vi ảnh đối với mỗi ảnh của dòng dữ liệu video được mã hóa theo thời gian theo thứ bậc và bao gồm cú pháp mức cao với bộ phân biệt ID lớp thứ bậc theo thời gian vượt qua phạm vi ảnh đảm bảo cho một hoặc nhiều ID lớp thứ bậc theo thời gian rằng các ảnh của một hoặc nhiều ID mức thứ bậc theo thời gian tương ứng không được tham chiếu, qua phép dự đoán thông tin phụ bù chuyển động theo thời gian, bằng bất kỳ ảnh khác.

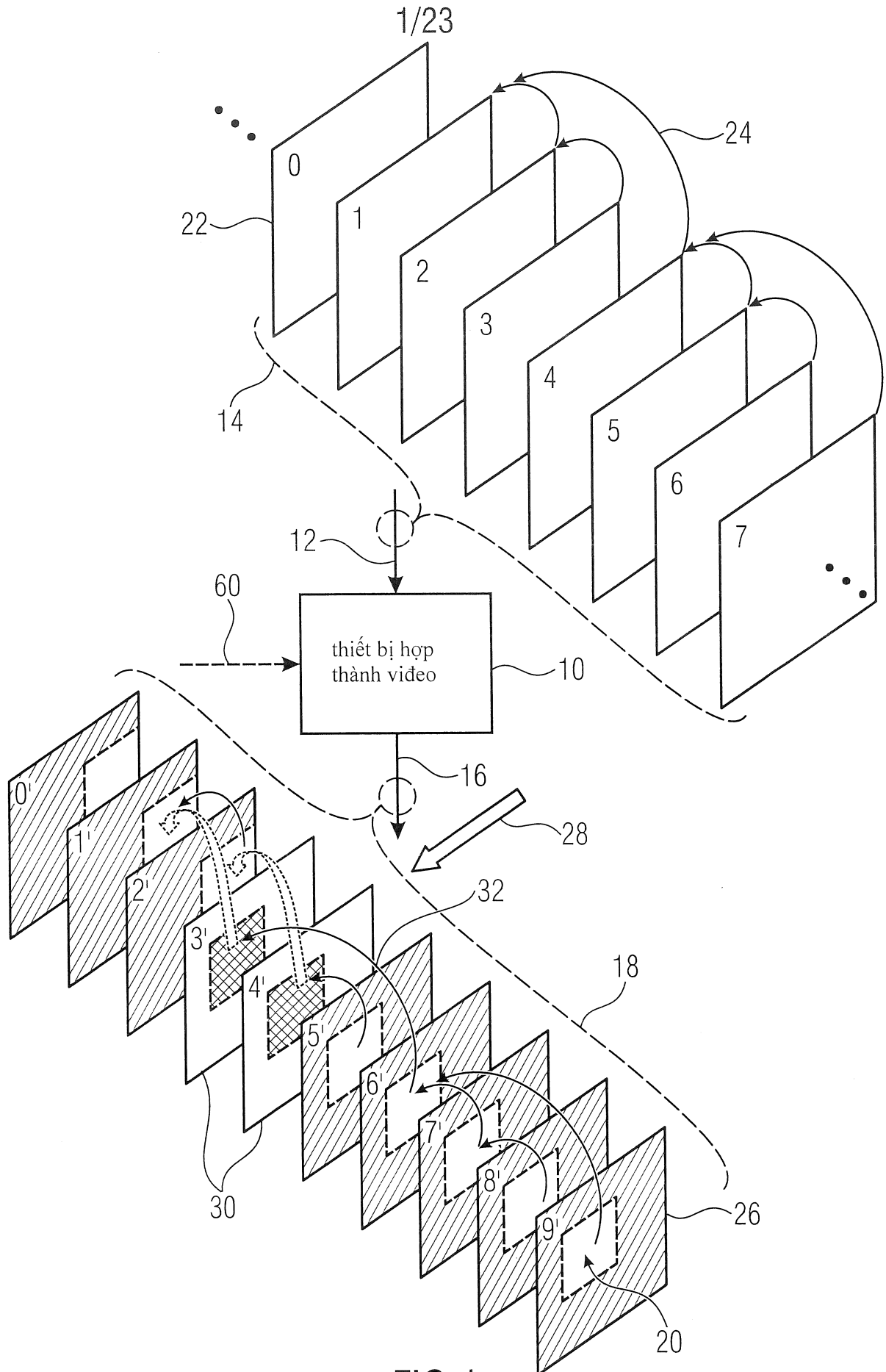
23. Phương tiện lưu trữ số bất biến theo điểm 22, trong đó cú pháp mức cao nằm trong bản tin thông tin nâng cao bổ sung (supplemental enhancement information - SEI) của dòng dữ liệu video được mã hóa theo thời gian theo thứ bậc.

24. Phương tiện lưu trữ số bất biến theo điểm 22, trong đó ID lớp thứ bậc theo thời gian phạm vi ảnh được báo hiệu trong dòng dữ liệu video được mã hóa theo thời gian theo thứ bậc, đối với mỗi ảnh trong phần đầu đơn vị truy cập được kết hợp với ảnh tương ứng 22, hoặc trong phần đầu lát gồm các lát của ảnh tương ứng.

25. Phương tiện lưu trữ số bất biến lưu trữ dòng dữ liệu video được mã hóa bằng cách sử dụng phép dự đoán theo thời gian được bù chuyển động và phép dự đoán thông tin phụ bù chuyển động, trong đó dòng dữ liệu video phân biệt, bằng cách báo hiệu theo ảnh,

các ảnh thứ nhất mà tại đó không có tập hợp hiện thời gồm các ảnh tham chiếu được tham chiếu, qua phép dự đoán thông tin phụ bù chuyển động, bằng bất kỳ ảnh đứng sau nào của dòng dữ liệu video, và

các ảnh thứ hai mà tại đó một trong số tập hợp hiện thời gồm các ảnh tham chiếu - hoặc là ứng viên để - được tham chiếu, qua phép dự đoán thông tin phụ bù chuyển động, bằng bất kỳ ảnh kế tiếp nào của dòng dữ liệu video.



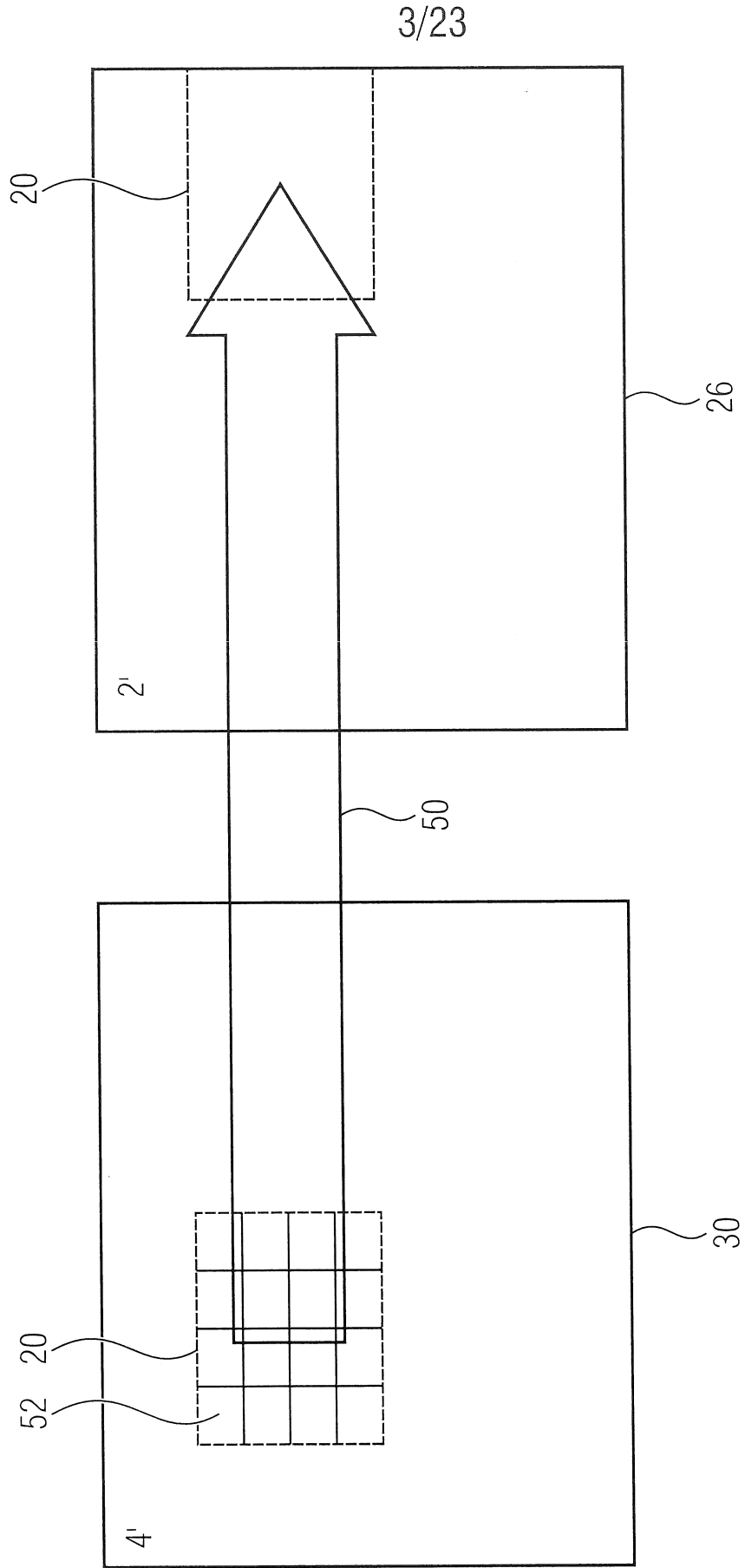
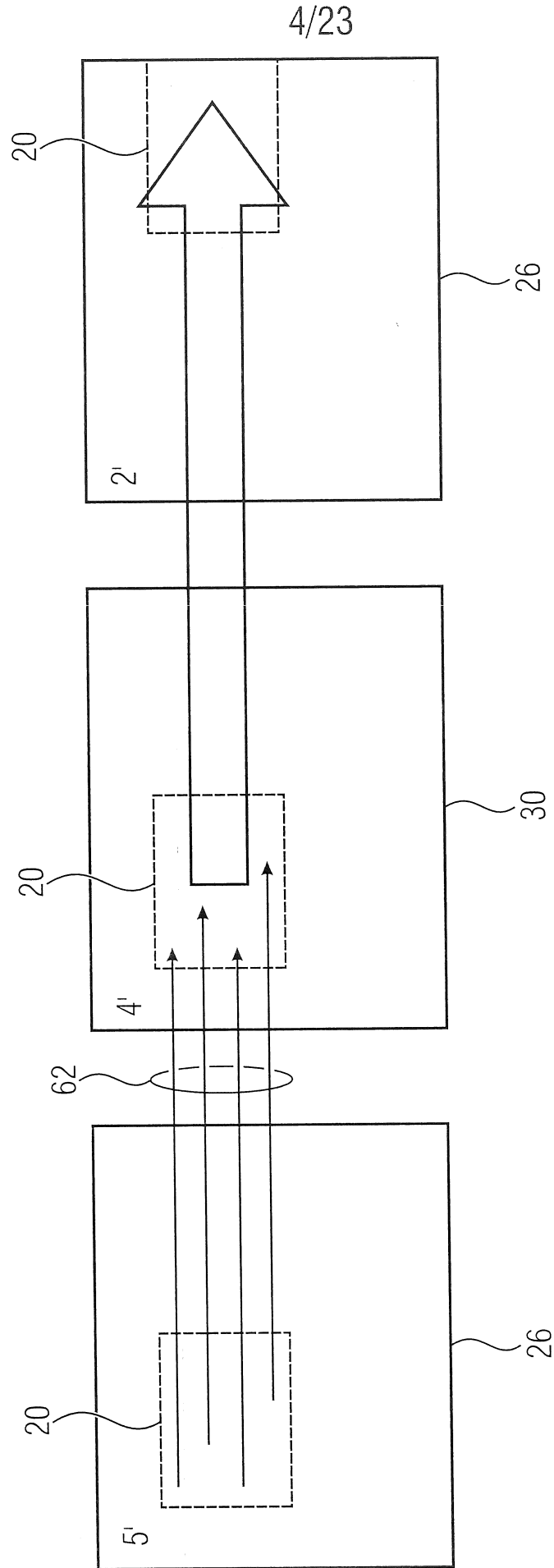


FIG 3



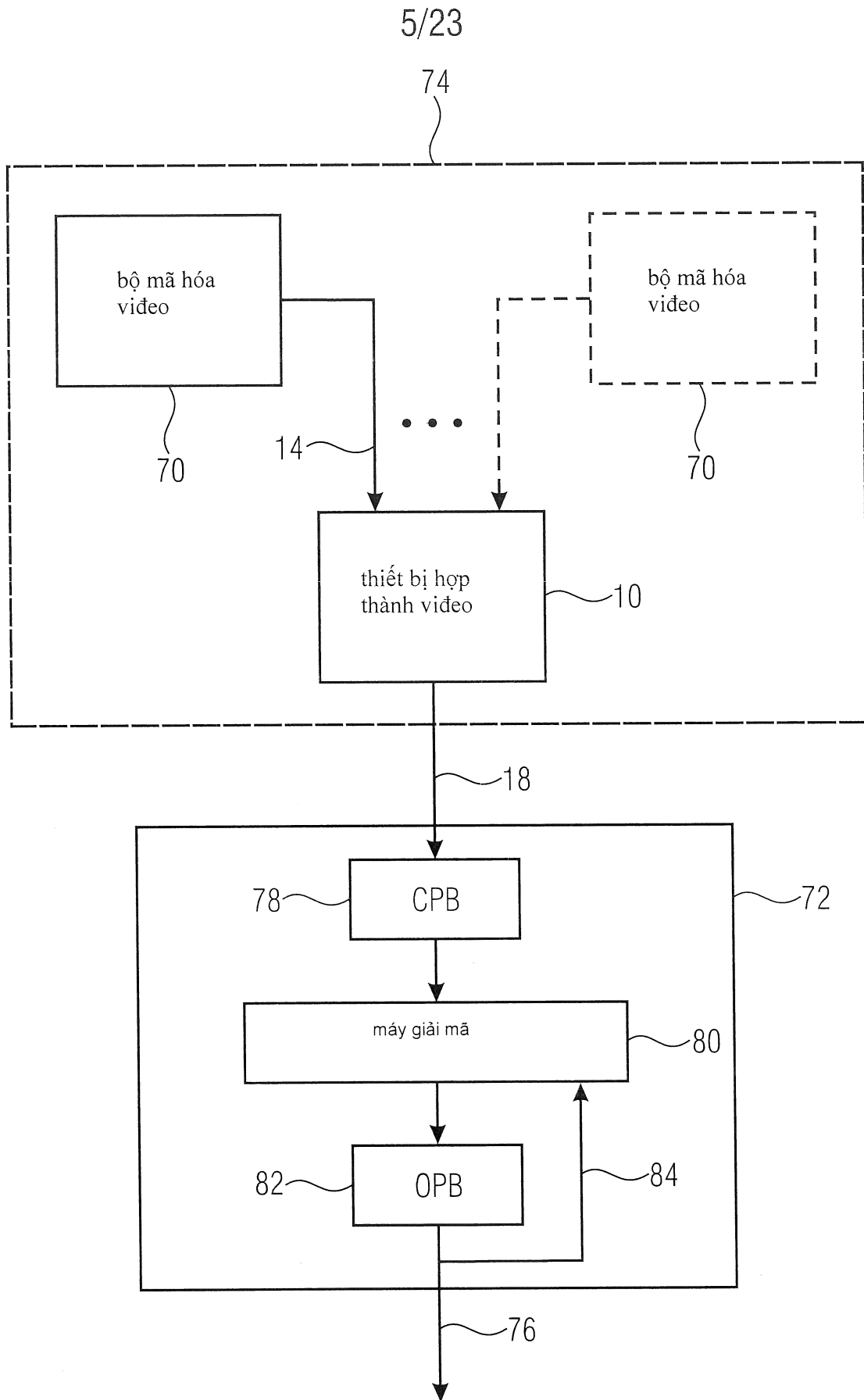


FIG 5

6/23

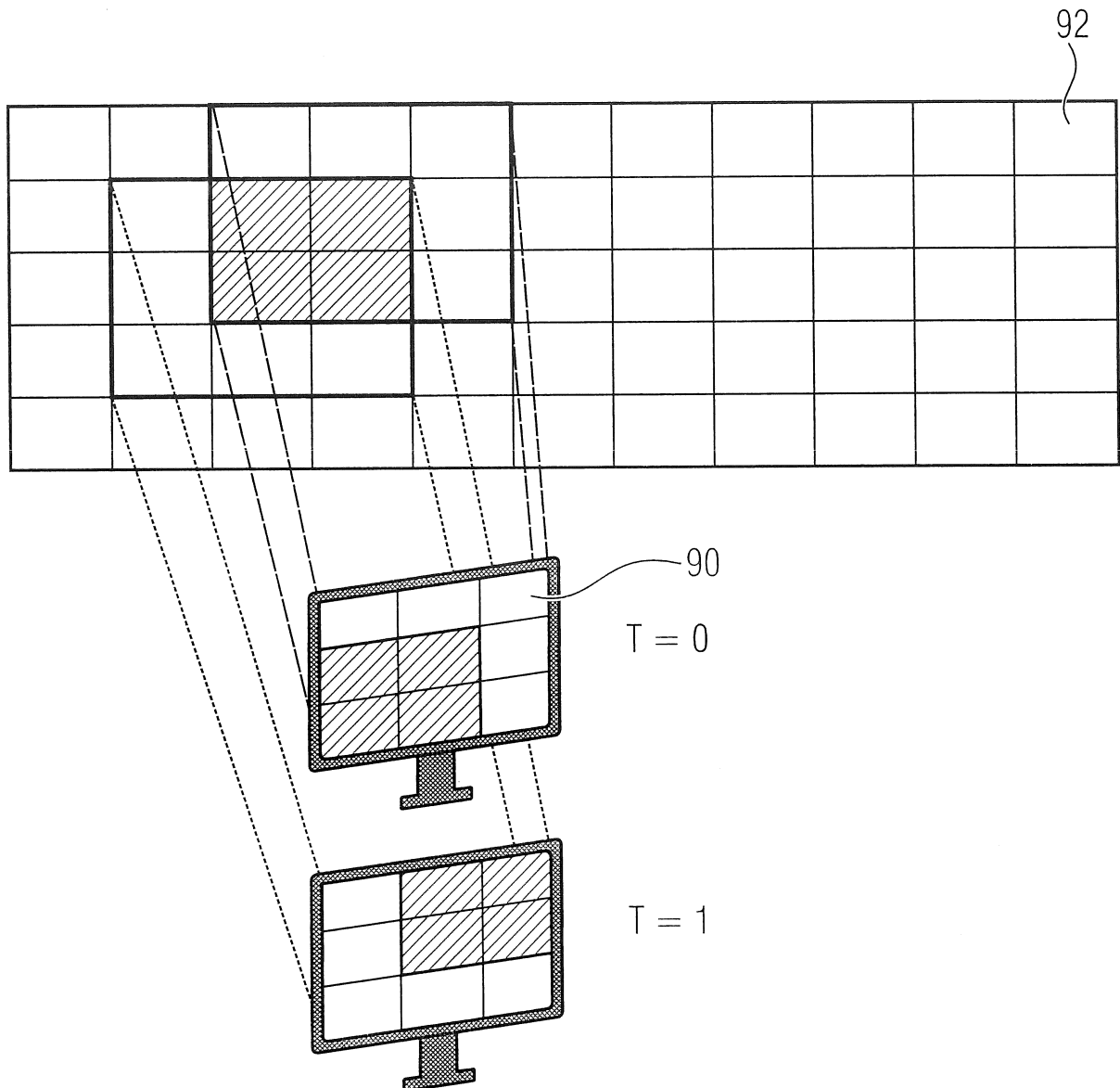


FIG 6

7/23

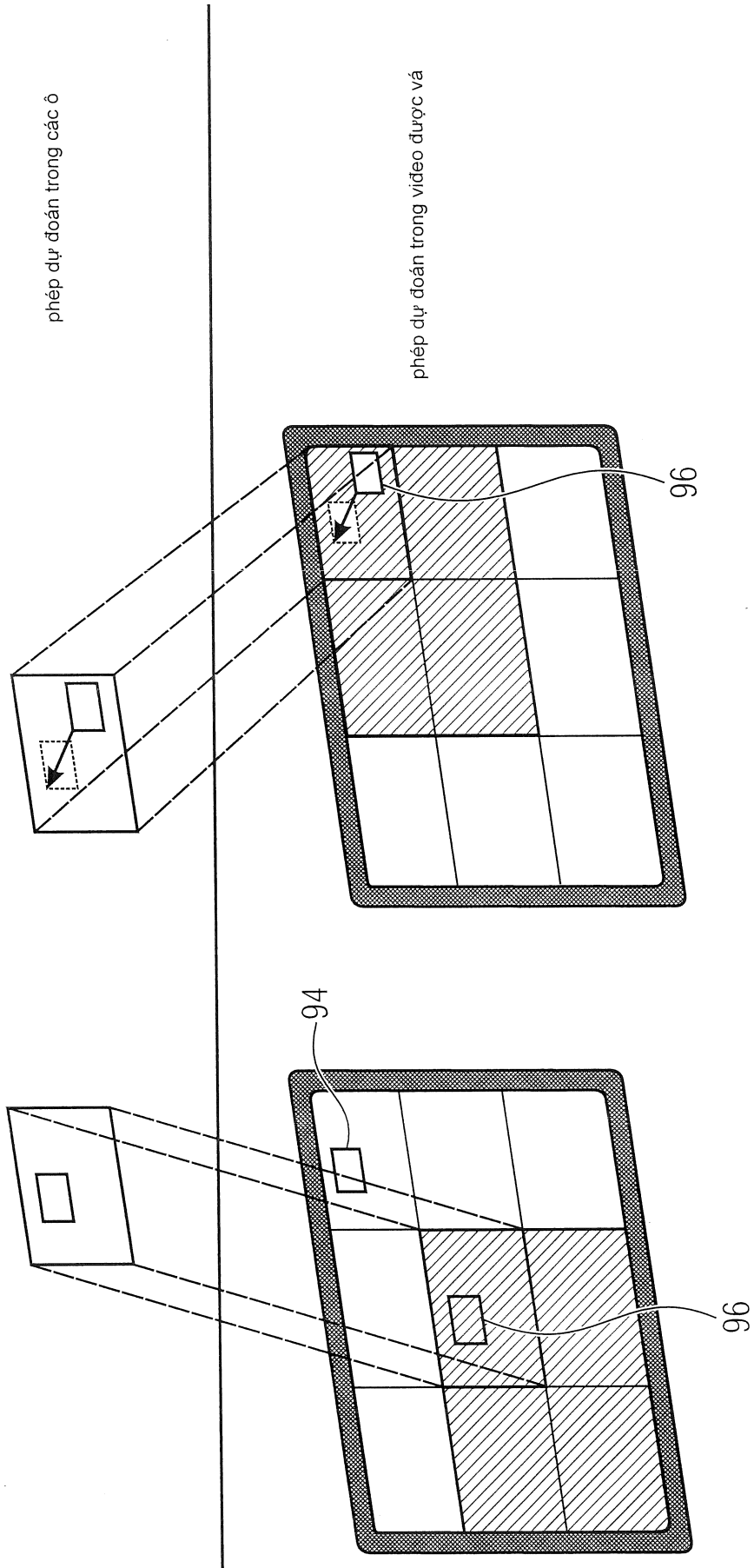


FIG 7

8/23

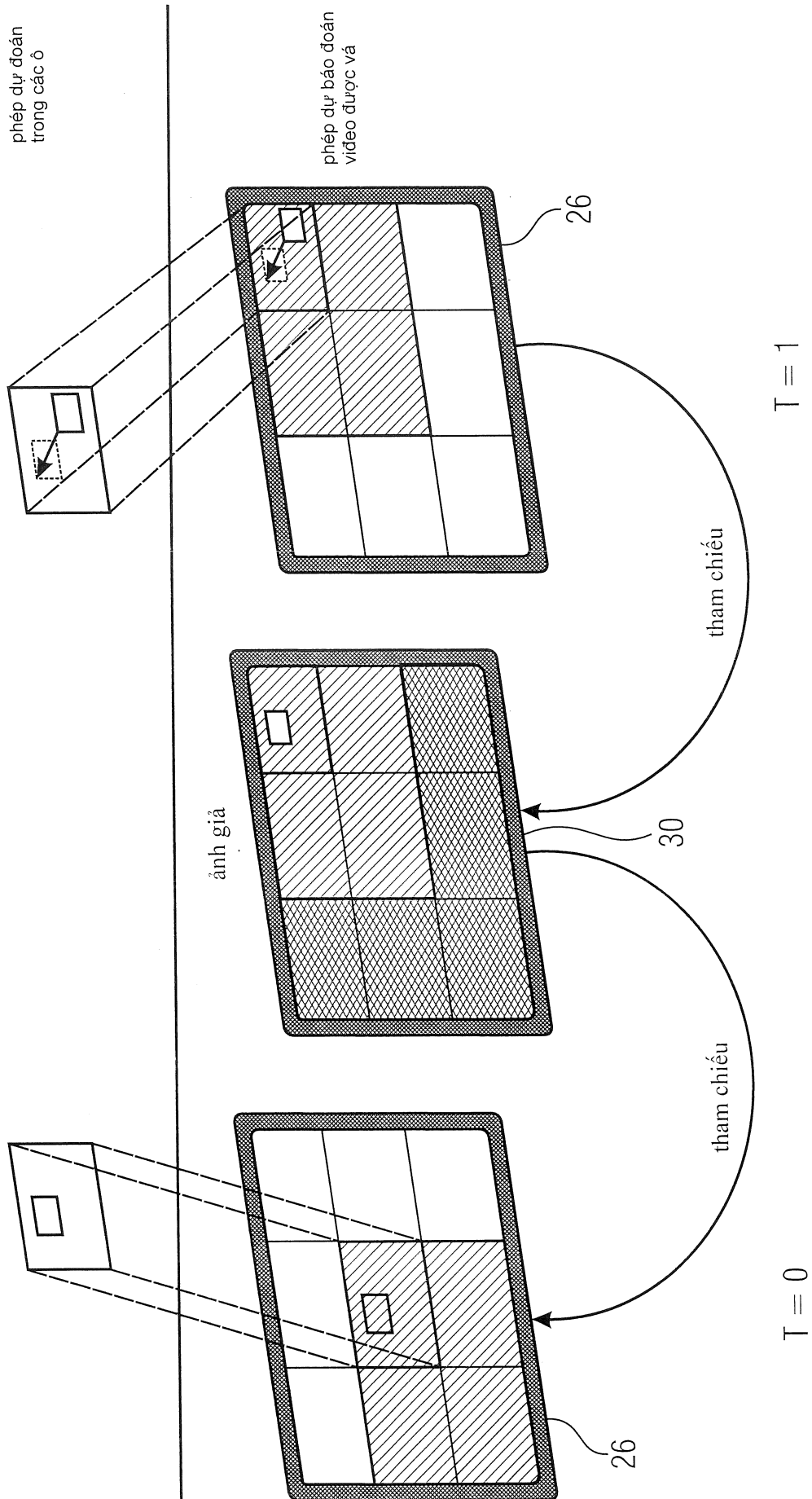
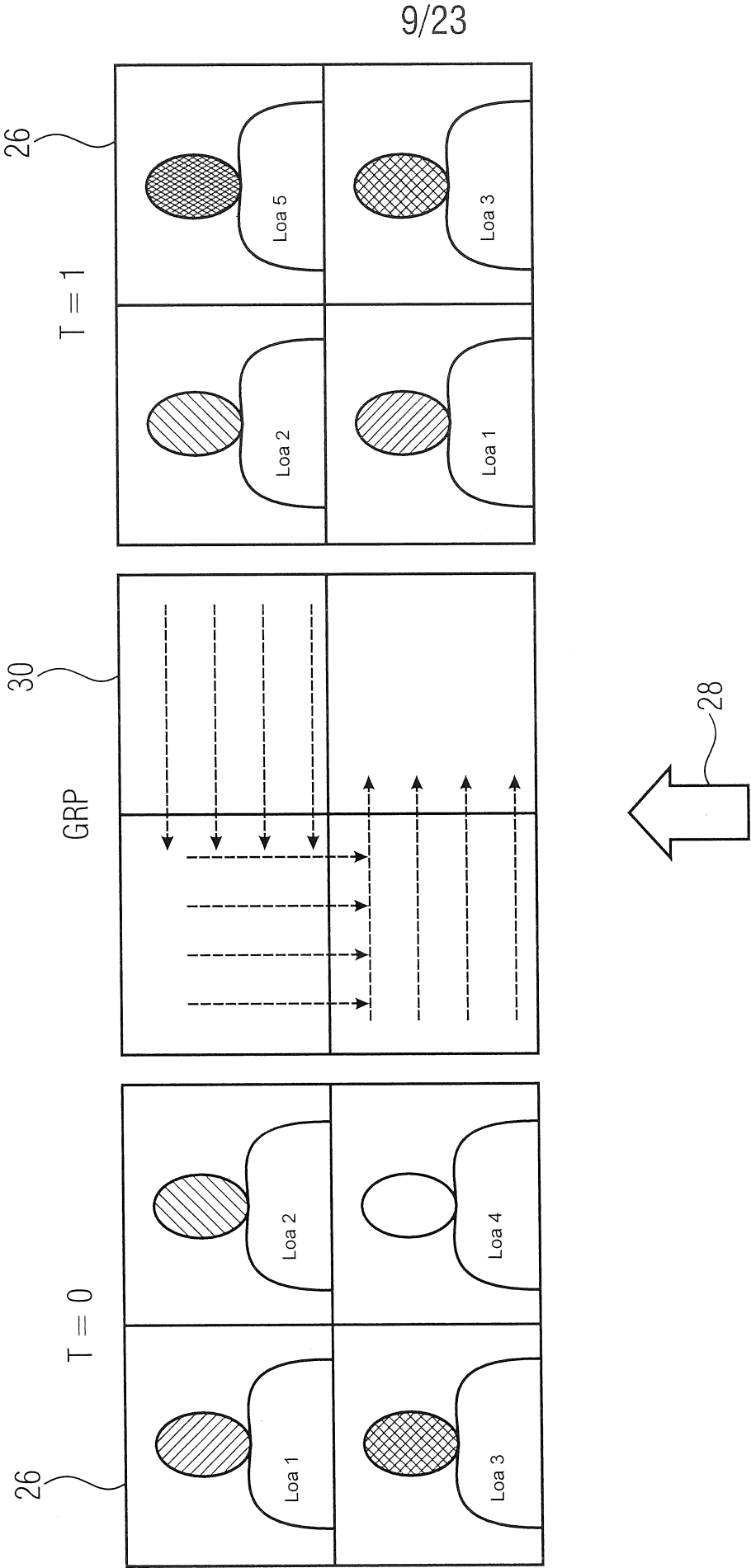


FIG 8



10/23

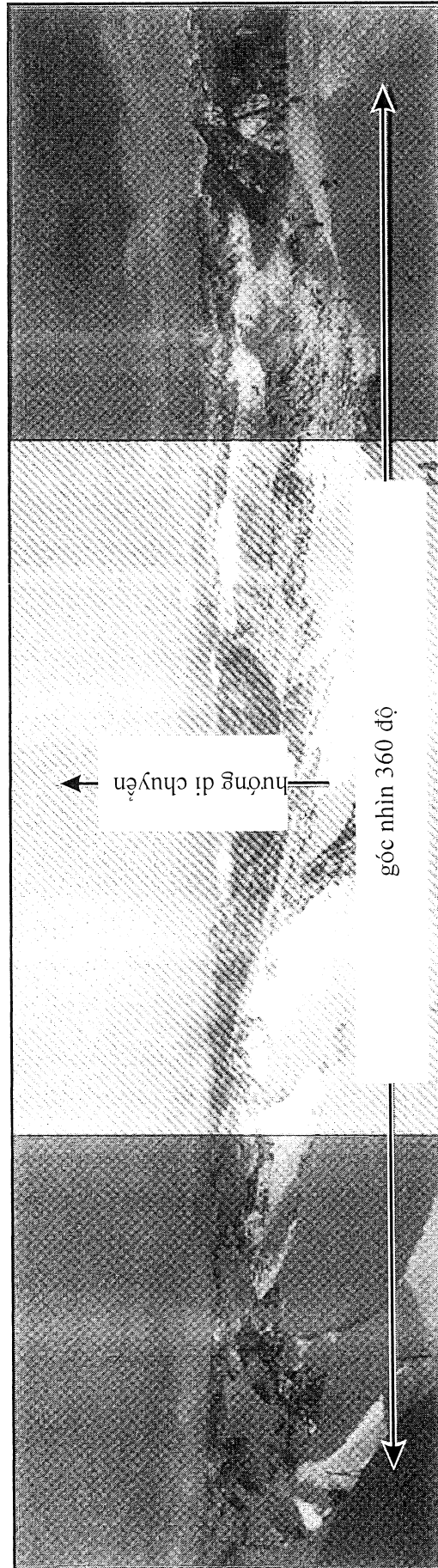


FIG 10

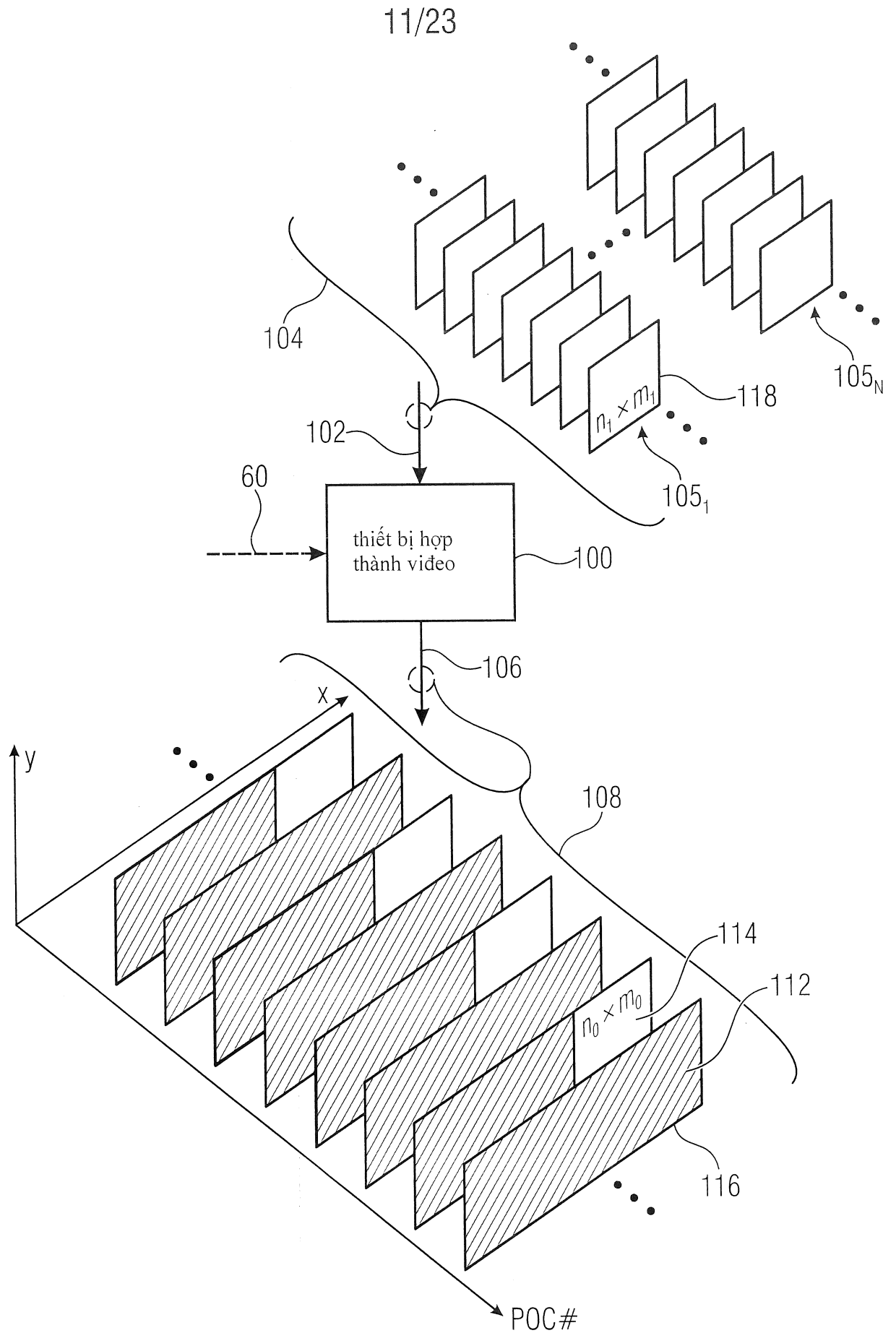


FIG 11

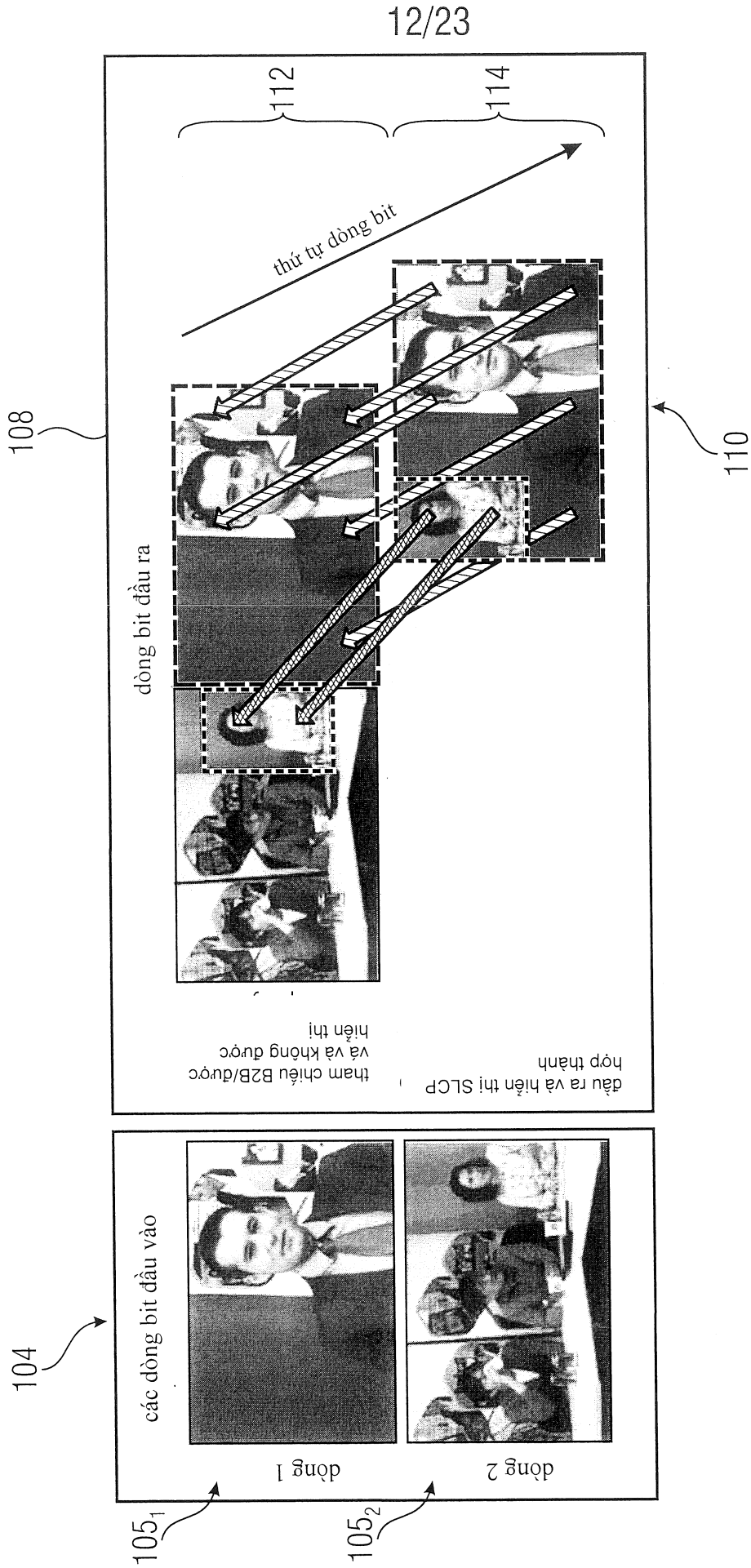


FIG 12

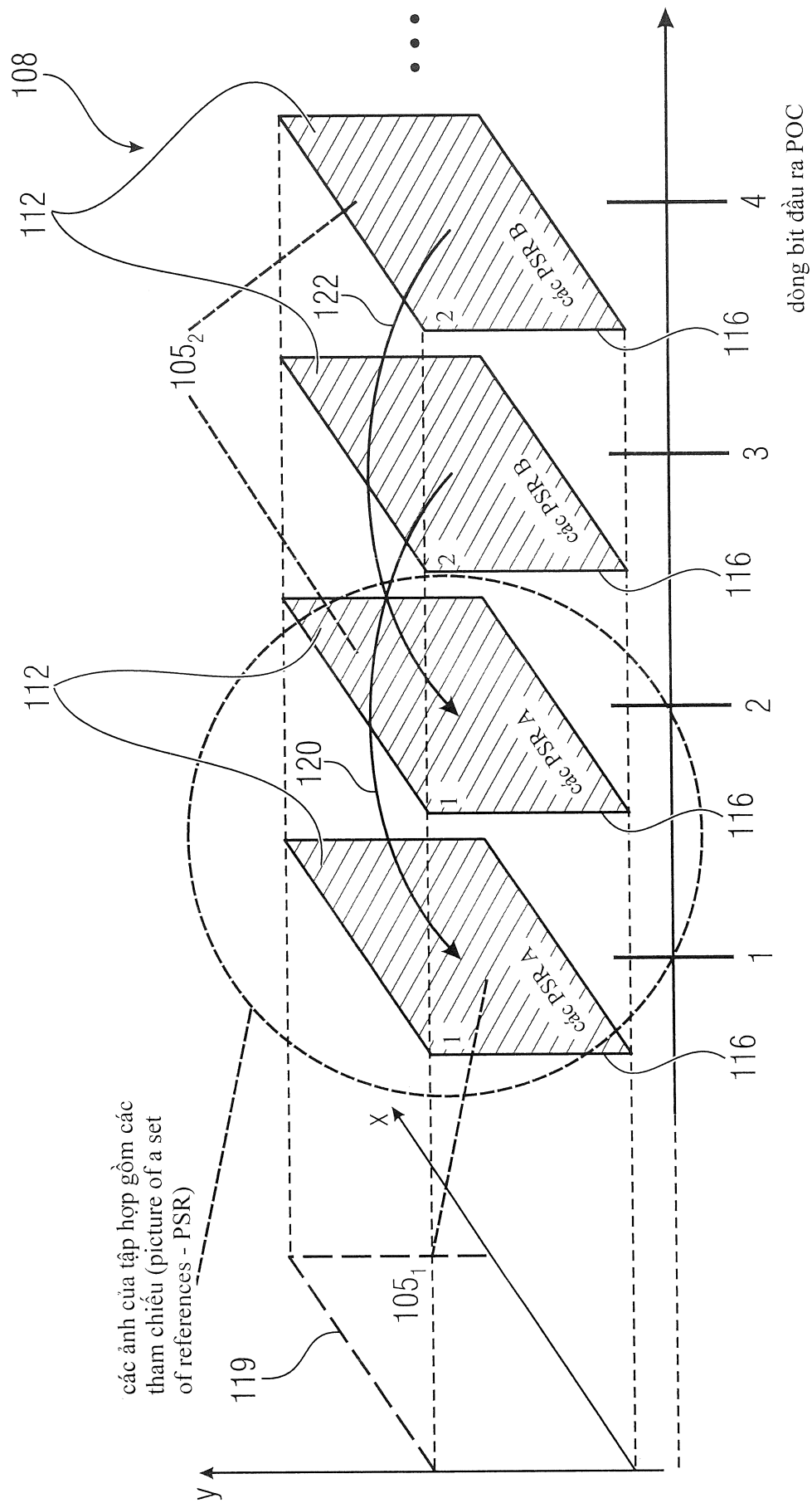


FIG 13

14/23

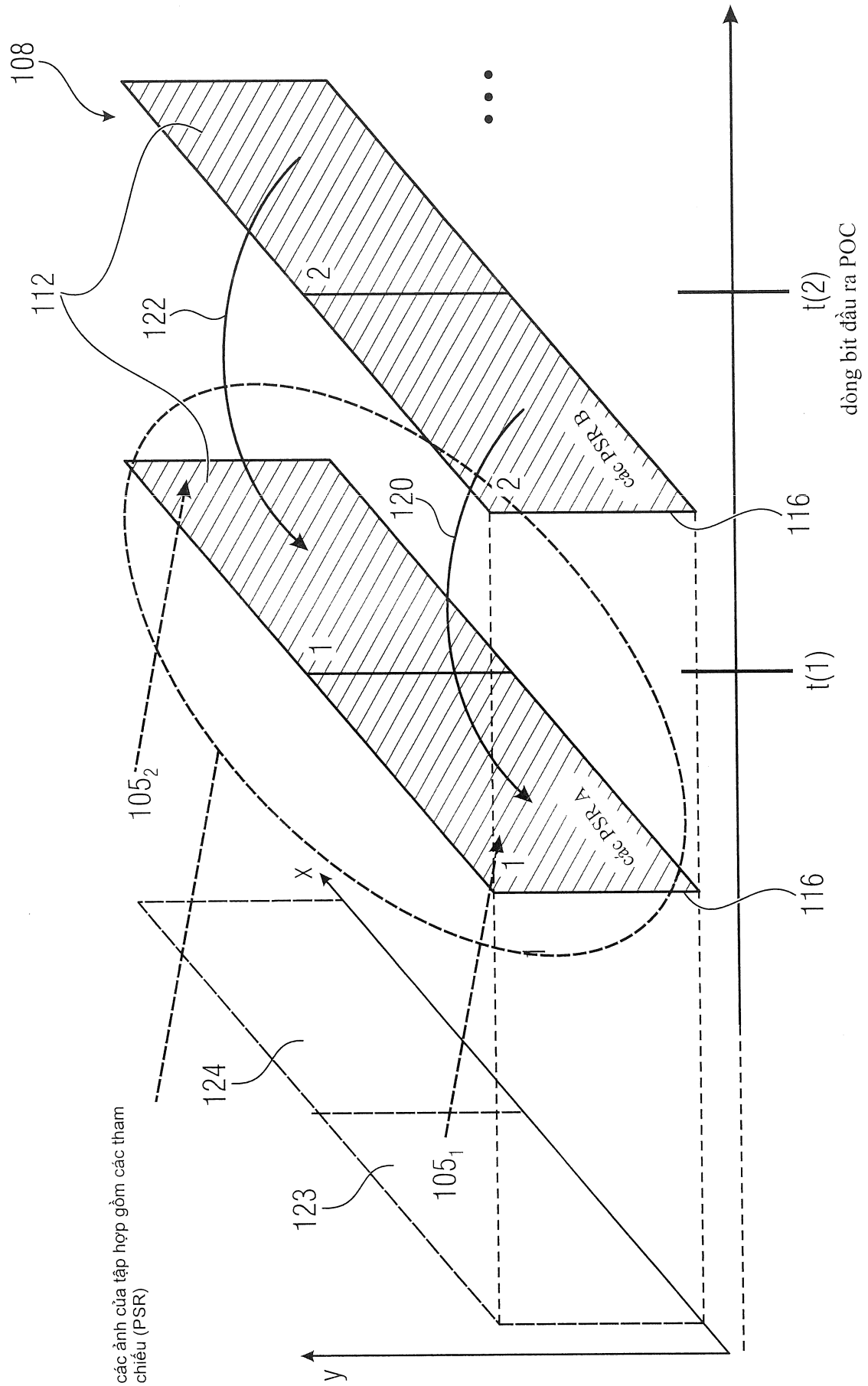


FIG 14

15/23

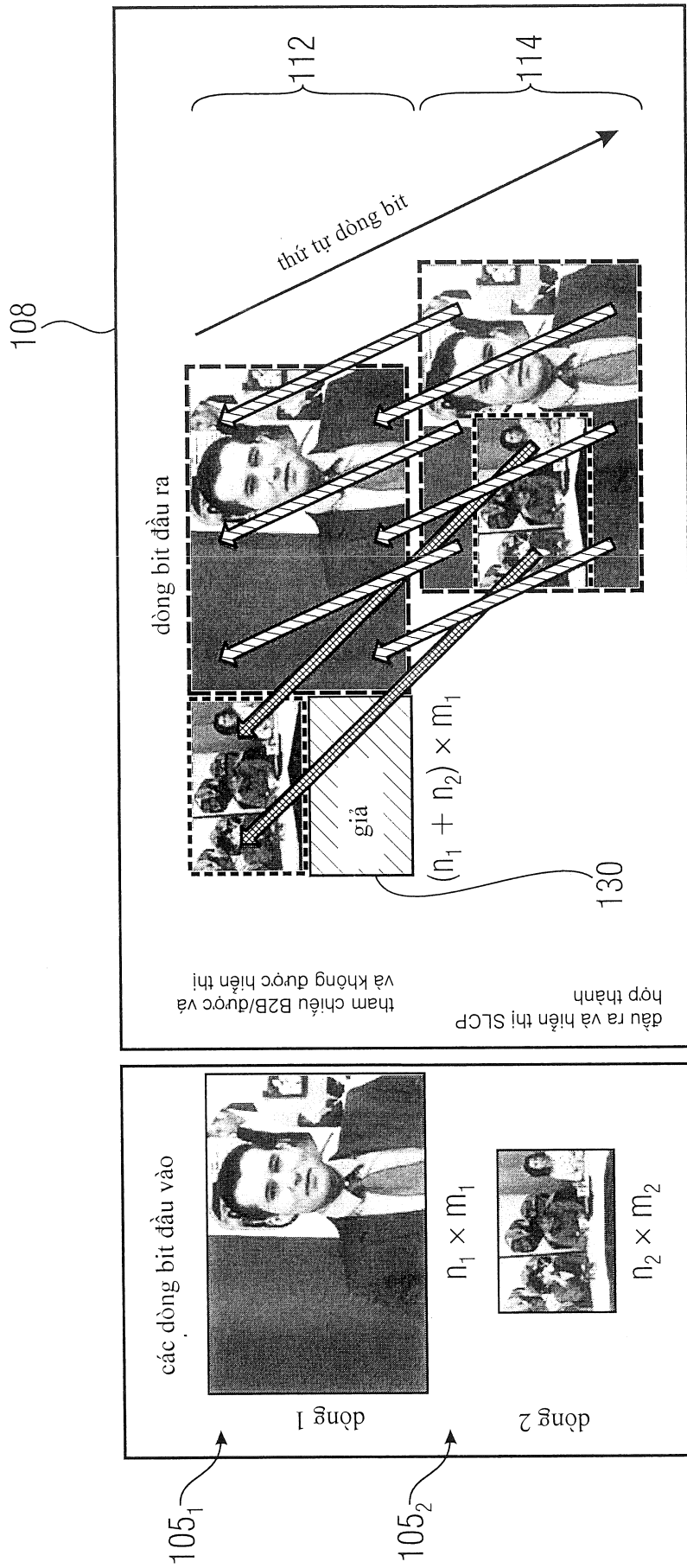


FIG 15

17/23

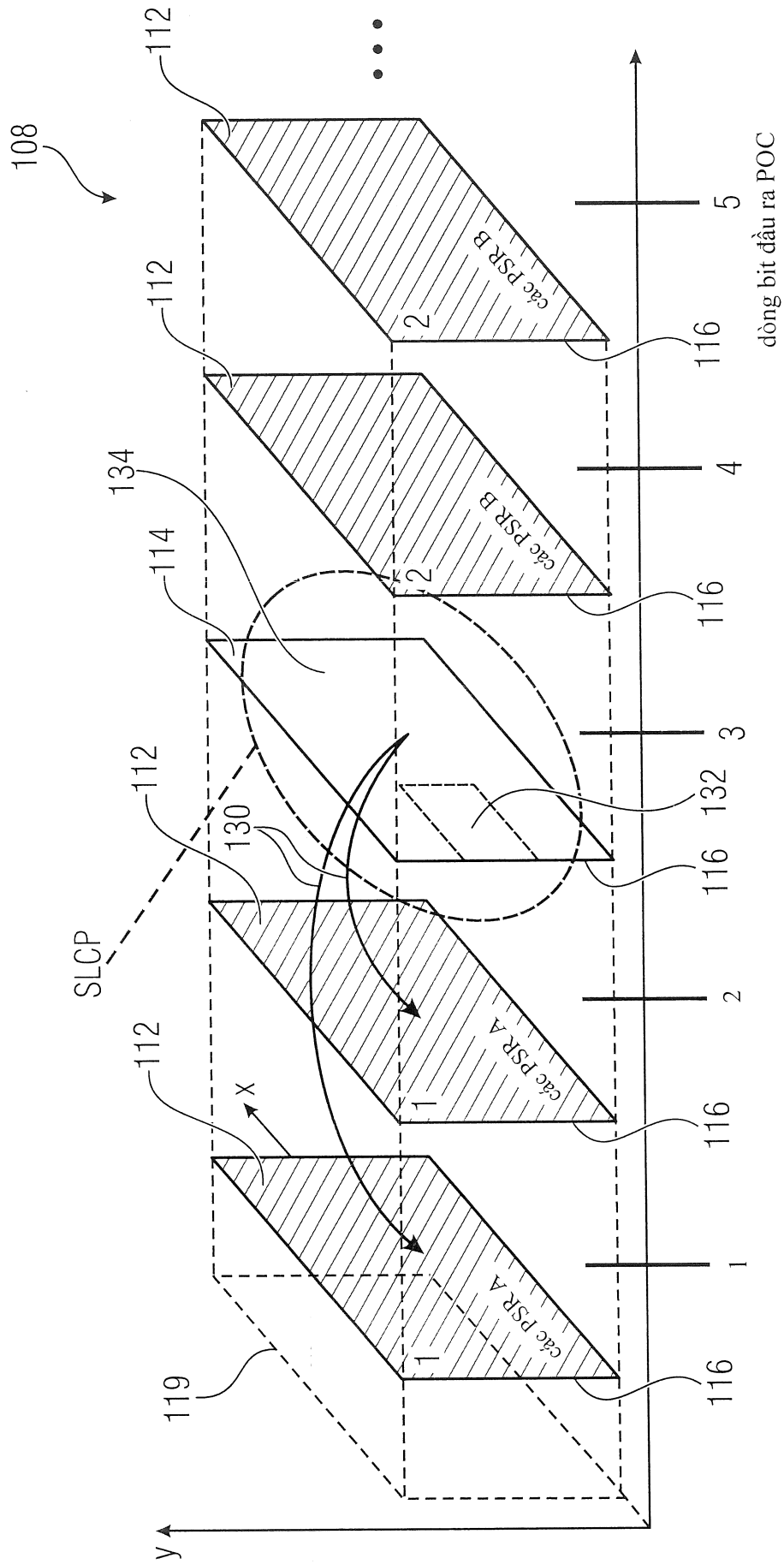


FIG 17

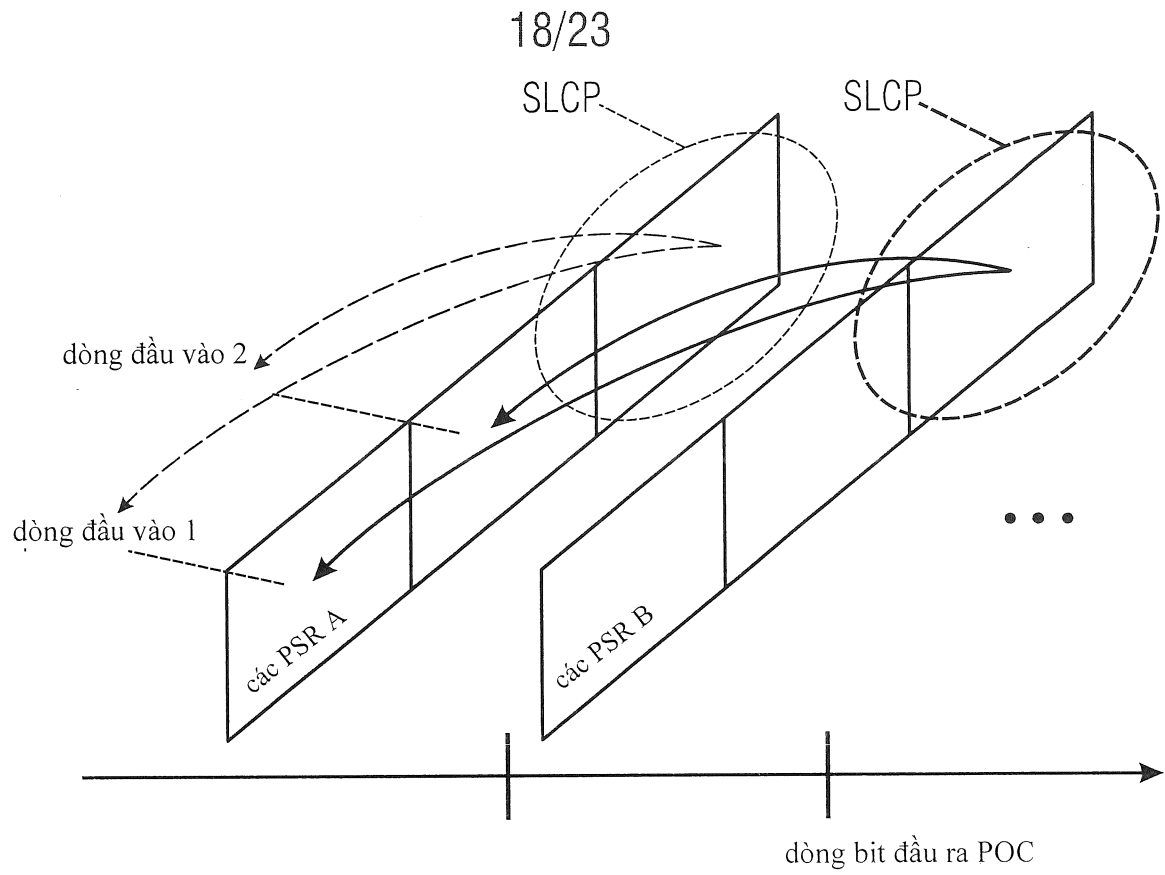


FIG 18A

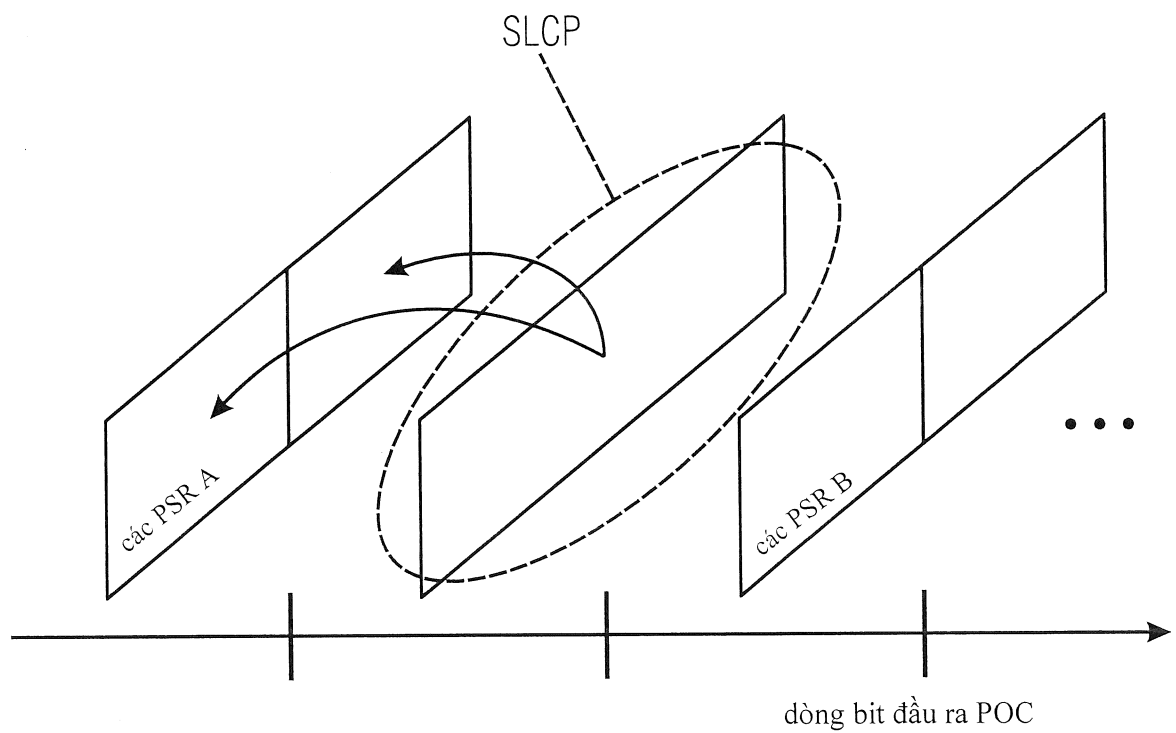


FIG 18B

19/23

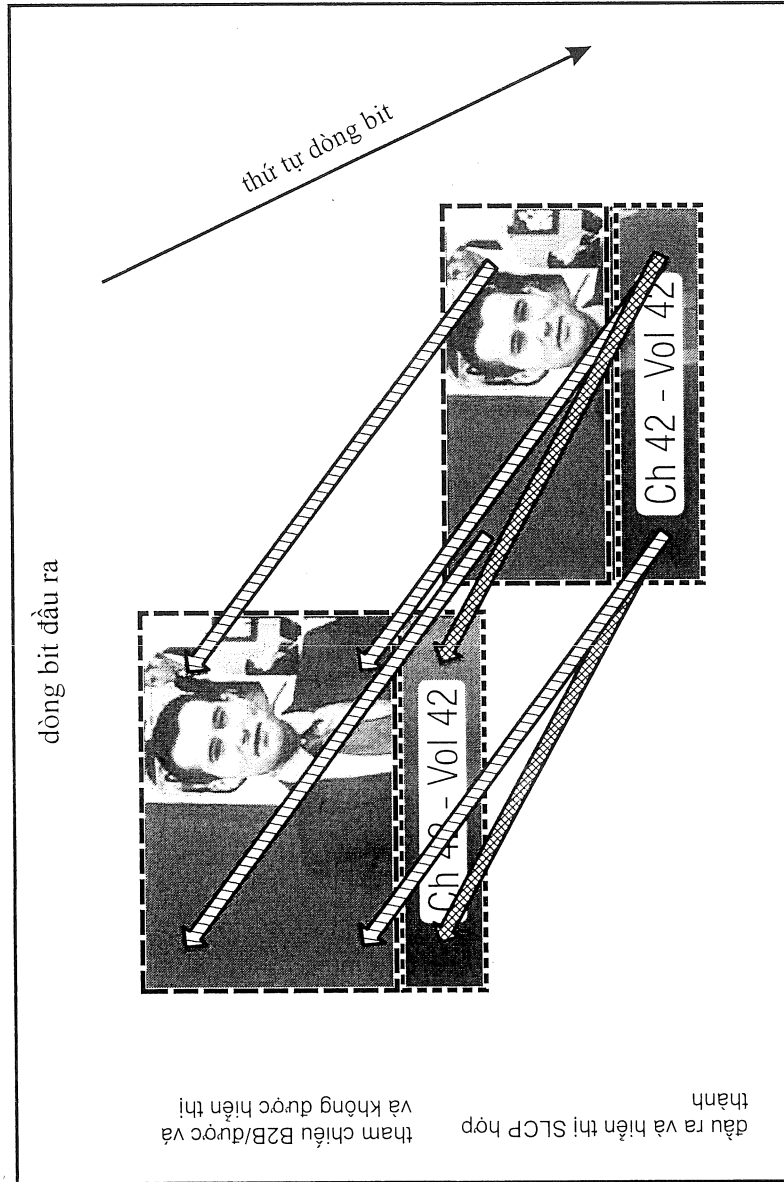
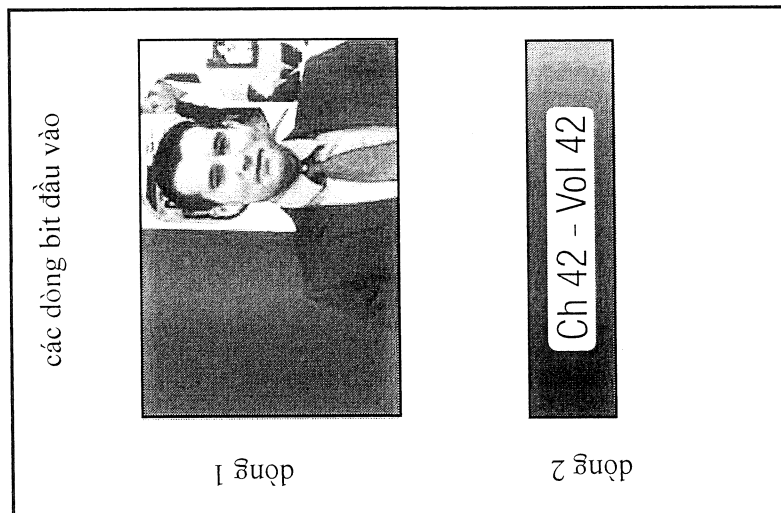


FIG 19



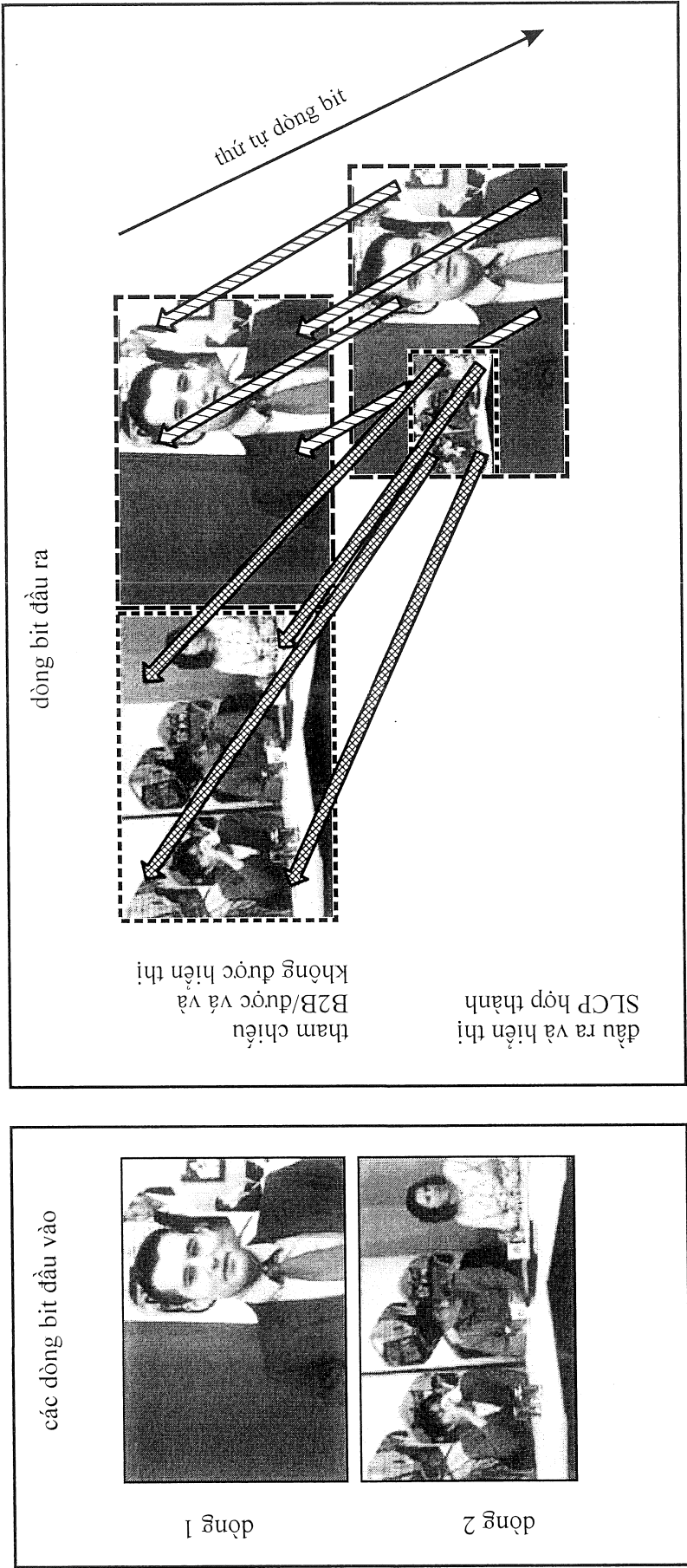


FIG 20

21/23

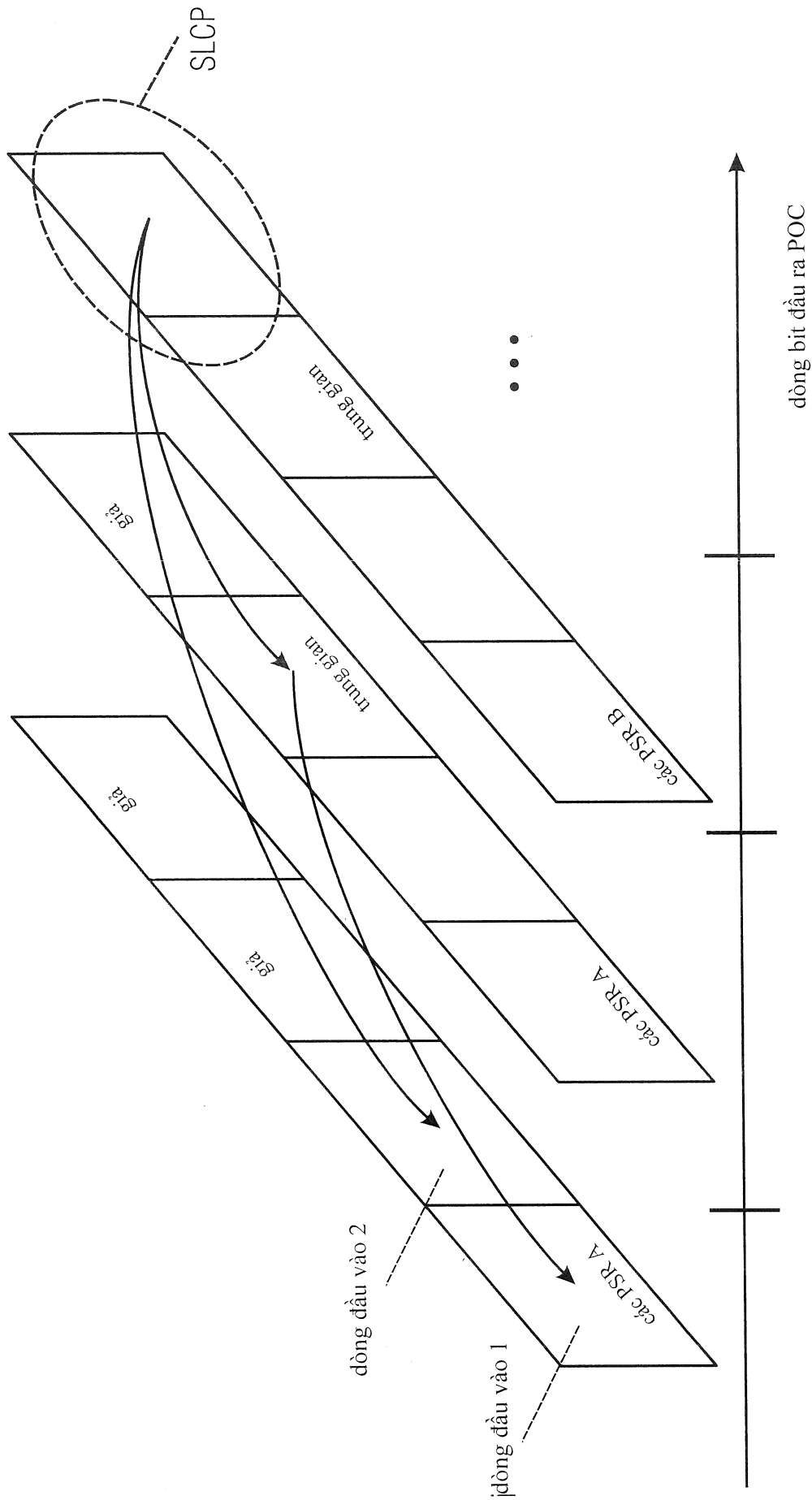


FIG 21

22/23

bitstream_composition_characteristics(payloadSize) {	Bộ mô tả
composition_enabled_type	ue(v)
if(composition_enabled_type==2){	
max_temporal_id_plus1_TMVP_disabled	u(3)
}	

FIG 22

23/23

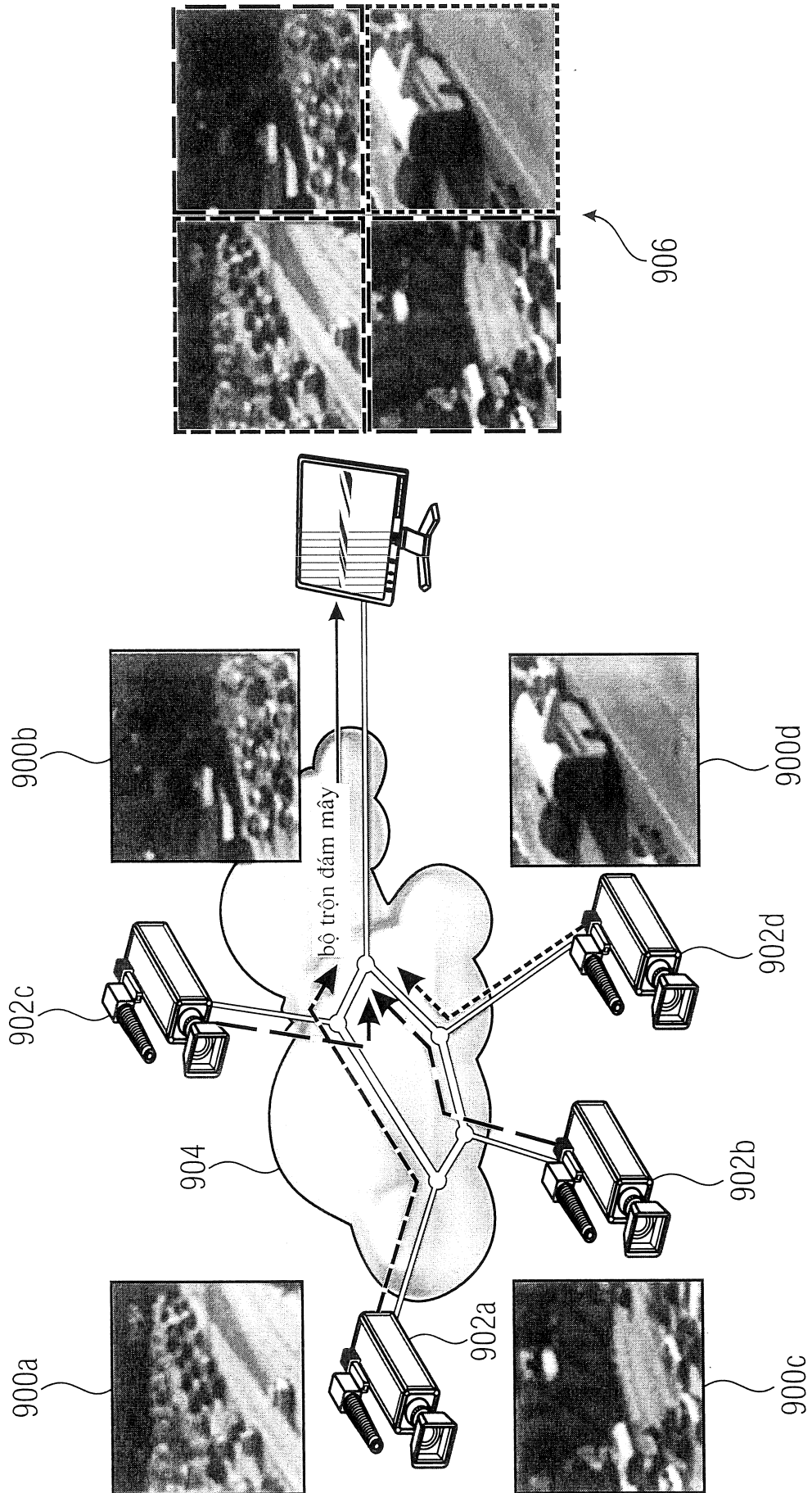


FIG 23