



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

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

(11)



1-0052074

(51)^{2020.01} H04N 19/30; H04N 19/70

(13) B

(21) 1-2022-02353

(22) 15/09/2020

(86) PCT/US2020/050835 15/09/2020

(87) WO2021/035225 25/02/2021

(30) 62/901,552 17/09/2019 US

(45) 25/09/2025 450

(43) 27/06/2022 411A

(73) HUAWEI TECHNOLOGIES CO., LTD. (CN)

Huawei Administration Building Bantian, Longgang District Shenzhen, Guangdong
518129, China

(72) WANG, Ye-Kui (US).

(74) Công ty Luật TNHH Phạm và Liên danh (PHAM & ASSOCIATES)

(54) PHƯƠNG PHÁP GIẢI MÃ ĐƯỢC TRIỂN KHAI BỞI BỘ GIẢI MÃ VÀ THIẾT BỊ
GIẢI MÃ

(21) 1-2022-02353

(57) Sáng chế đề cập đến phương pháp giải mã video bao gồm bước nhận dòng bit bao gồm tập hợp tham số chuỗi (sequence parameter set, SPS), tập hợp tham số hình ảnh (picture parameter set, PPS), và các ảnh con được liên kết với ảnh xạ của định danh (Identifier, ID) ảnh con. Thực hiện xác định liệu cờ SPS có giá trị thứ nhất hay giá trị thứ hai. Cờ SPS có giá trị thứ nhất xác định rằng ảnh xạ của ID ảnh con được báo hiệu trong SPS, và giá trị thứ hai xác định rằng ảnh xạ của ID ảnh con được báo hiệu trong PPS. Ảnh xạ của ID ảnh con thu được từ SPS khi cờ SPS có giá trị thứ nhất và từ PPS khi cờ SPS có giá trị thứ hai hoặc cờ PPS có giá trị thứ nhất.

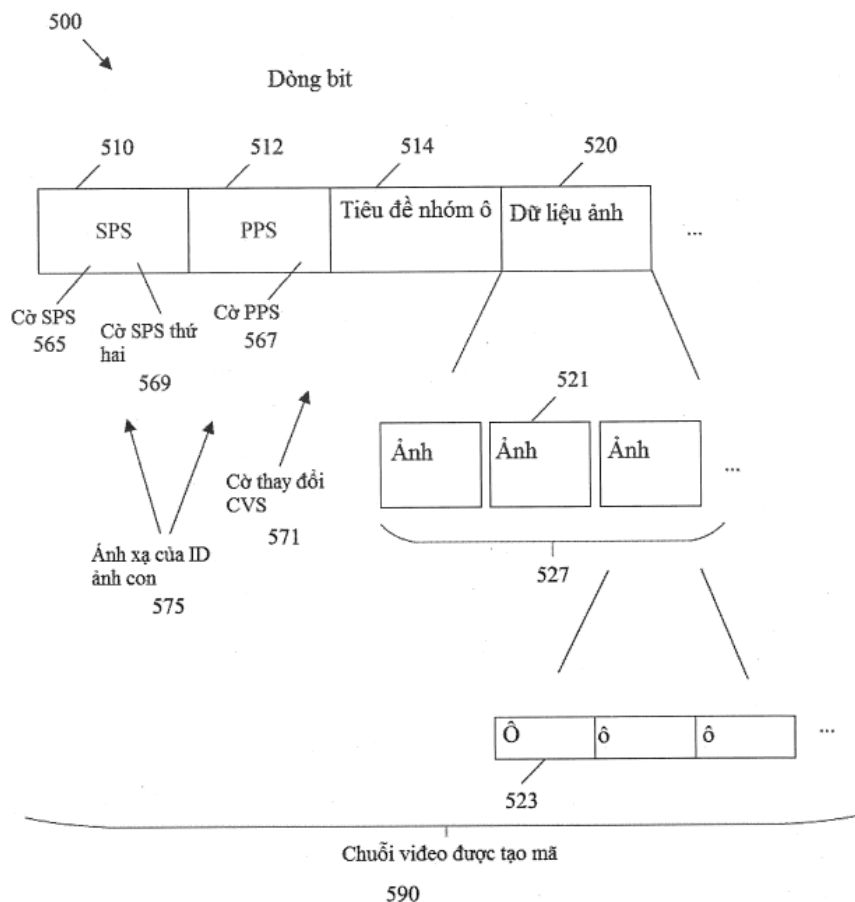


Fig.5

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập chung đến kỹ thuật tạo mã video, và cụ thể là đề cập đến báo hiệu các định danh (identifier, ID) của ảnh con trong kỹ thuật tạo mã video dựa trên ảnh con.

Tình trạng kỹ thuật của sáng chế

Lượng dữ liệu video cần mô tả thậm chí video tương đối ngắn có thể là rất lớn, có thể dẫn đến các khó khăn khi dữ liệu cần được phát trực tuyến hoặc ngược lại được truyền thông giữa mạng truyền thông có dung lượng băng thông giới hạn. Do vậy, dữ liệu video thường được nén trước khi được truyền trên các mạng viễn thông ngày nay. Kích thước của video cũng có thể là vấn đề khi video được lưu trữ trên thiết bị lưu trữ do các tài nguyên bộ nhớ có thể bị giới hạn. Các thiết bị nén video thường sử dụng phần mềm và/hoặc phần cứng ở phía nguồn để mã hóa dữ liệu video trước khi truyền hoặc lưu trữ, nhờ đó giảm lượng dữ liệu cần để biểu diễn các ảnh video số. Sau đó, dữ liệu được nén được nhận ở phía đích bằng thiết bị giải mã video mà giải mã dữ liệu video. Với các tài nguyên mạng hữu hạn và nhu cầu ngày càng tăng có chất lượng video cao hơn, mong muốn các kỹ thuật nén và giải nén được cải tiến mà cải thiện tỷ lệ nén mà không làm ảnh hưởng đến chất lượng ảnh.

Bản chất kỹ thuật của sáng chế

Khía cạnh thứ nhất đề cập đến phương pháp được triển khai bằng bộ giải mã, bao gồm nhận, bởi bộ giải mã, dòng bit bao gồm tập hợp tham số chuỗi (sequence parameter set, SPS), tập hợp tham số ảnh (picture parameter set, PPS), và các ảnh con được liên kết với ánh xạ của định danh (identifier, ID) ảnh con, trong đó SPS bao gồm cờ SPS; xác định, bởi bộ giải mã, liệu cờ SPS có giá trị thứ nhất or giá trị thứ hai, trong đó cờ SPS có giá trị thứ nhất xác định rằng ảnh

xạ của ID ảnh con được báo hiệu trong SPS, và trong đó cờ SPS có giá trị thứ hai xác định rằng ánh xạ của ID ảnh con được báo hiệu trong PPS; thu được, bởi bộ giải mã, ánh xạ của ID ảnh con từ SPS khi cờ SPS có giá trị thứ nhất và từ PPS khi cờ SPS có giá trị thứ hai; và giải mã, bởi bộ giải mã, các ảnh con nhờ sử dụng ánh xạ của ID ảnh con.

Phương pháp đề cập đến các kỹ thuật đảm bảo báo hiệu hiệu quả các ID ảnh con thậm chí khi các ID ảnh con thay đổi trong chuỗi vi đeo được tạo mã (coded video sequence, CVS) trong các kịch bản ứng dụng bao gồm trích xuất dòng bit thành phần và dòng bit thành phần. Báo hiệu hiệu quả được hoàn thành bằng cách thiết lập các cờ trong tập hợp tham số chuỗi (sequence parameter set, SPS) hoặc tập hợp tham số ảnh (picture parameter set, SPS) để chỉ báo liệu các ID ảnh con trong CVS có thể thay đổi và, nếu vậy, trong đó đặt các ID ảnh con. Điều này giảm dư thừa và tăng hiệu suất tạo mã. Do vậy, bộ mã hóa/bộ giải mã (cũng được biết đến như là “codec”) khi tạo mã video được cải thiện so với các codec hiện tại. Thực tế là, quá trình tạo mã video được cải thiện cho người dùng trải nghiệm sử dụng tốt hơn khi các video được gửi, được nhận, và/hoặc được xem.

Một cách tùy chọn, theo khía cạnh bất kỳ trong các khía cạnh trước đó, cách thực hiện khác của khía cạnh này đề cập đến việc PPS bao gồm cờ PPS, và phương pháp còn bao gồm xác định, bởi bộ giải mã, liệu cờ PPS có giá trị thứ nhất hoặc giá trị thứ hai, trong đó cờ PPS có giá trị thứ nhất xác định rằng ánh xạ của ID ảnh con được báo hiệu trong PPS, và trong đó cờ PPS có giá trị thứ hai xác định rằng ánh xạ của ID ảnh con không được báo hiệu trong PPS; và thu được, bởi bộ giải mã, ánh xạ của ID ảnh con từ PPS khi cờ PPS có giá trị thứ nhất.

Một cách tùy chọn, theo khía cạnh bất kỳ trong các khía cạnh trước đó, cách thực hiện khác của khía cạnh này đề cập đến việc cờ SPS có giá trị thứ nhất khi cờ PPS có giá trị thứ hai, và trong đó cờ SPS có giá trị thứ hai khi cờ PPS có giá trị thứ nhất.

Một cách tùy chọn, theo khía cạnh bất kỳ trong các khía cạnh trước đó, cách thực hiện khác của khía cạnh này đề cập đến việc giá trị thứ nhất bằng 1 và giá trị thứ hai bằng 0.

Một cách tùy chọn, theo khía cạnh bất kỳ trong các khía cạnh trước đó, cách thực hiện khác của khía cạnh này đề cập đến việc SPS bao gồm cờ SPS thứ hai, và trong đó cờ SPS thứ hai xác định liệu ánh xạ của ID ảnh con được báo hiệu tường minh trong SPS hoặc PPS.

Một cách tùy chọn, theo khía cạnh bất kỳ trong các khía cạnh trước đó, cách thực hiện khác của khía cạnh này đề cập đến việc dòng bit còn bao gồm cờ thay đổi CVS, và trong đó cờ thay đổi CVS chỉ báo liệu ánh xạ của ID ảnh con được phép thay đổi trong CVS của dòng bit.

Một cách tùy chọn, theo khía cạnh bất kỳ trong các khía cạnh trước đó, cách thực hiện khác của khía cạnh này đề cập đến việc dòng bit bao gồm dòng bit được hợp nhất, và trong đó ánh xạ của ID ảnh con đã thay đổi trong CVS của dòng bit.

Khía cạnh thứ hai đề cập đến phương pháp được triển khai bằng bộ mã hóa, phương pháp bao gồm mã hóa, bởi bộ giải mã, dòng bit bao gồm SPS, PPS, và các ảnh con được liên kết với ánh xạ ID ảnh con, trong đó SPS bao gồm cờ SPS; thiết lập, bởi bộ giải mã, cờ SPS bằng giá trị thứ nhất khi ánh xạ của ID ảnh con được báo hiệu trong SPS và bằng giá trị thứ hai khi ánh xạ của ID ảnh con được báo hiệu trong PPS; và lưu trữ, bởi bộ giải mã, dòng bit để truyền thông về phía bộ giải mã.

Phương pháp đề cập đến các kỹ thuật đảm bảo báo hiệu hiệu quả các ID ảnh con thậm chí khi các ID ảnh con thay đổi trong CVS trong các kịch bản ứng dụng bao gồm cả trích xuất dòng bit thành phần và hợp nhất dòng bit thành phần. Báo hiệu hiệu quả được hoàn thành bằng cách thiết lập các cờ trong SPS hoặc SPS để chỉ báo liệu các ID con trong CVS có thể thay đổi và, nếu vậy, trong đó đặt các ID ảnh con. Điều này giảm độ dư thừa và tăng hiệu suất tạo mã. Do vậy, bộ mã hóa/bộ giải mã (cũng được biết đến là “codec”) trong khi tạo mã video được cải thiện liên quan đến các codec hiện tại. Thực tế là, quá trình tạo mã video được cải tiến cho người dùng trải nghiệm sử dụng tốt hơn khi các video được gửi, được nhận, và/hoặc được xem.

Một cách tùy chọn, theo khía cạnh bất kỳ trong các khía cạnh trước đó, cách thực hiện khác của khía cạnh này đề cập đến việc bộ giải mã, cờ PPS trong PPS

đến giá trị thứ nhất khi ánh xạ của ID ảnh con được báo hiệu trong PPS và đến giá trị thứ hai khi ánh xạ của ID ảnh con không được báo hiệu trong PPS.

Một cách tùy chọn, theo khía cạnh bất kỳ trong các khía cạnh trước đó, cách thực hiện khác của khía cạnh này đề cập đến việc cờ SPS có giá trị thứ nhất khi cờ PPS có giá trị thứ hai, và trong đó cờ SPS có giá trị thứ hai khi cờ PPS có giá trị thứ nhất.

Một cách tùy chọn, theo khía cạnh bất kỳ trong các khía cạnh trước đó, cách thực hiện khác của khía cạnh này đề cập đến việc giá trị thứ nhất bằng 1 và giá trị thứ hai bằng 0.

Một cách tùy chọn, theo khía cạnh bất kỳ trong các khía cạnh trước đó, cách thực hiện khác của khía cạnh này đề cập đến việc SPS bao gồm cờ SPS thứ hai, và trong đó cờ SPS thứ hai xác định liệu việc ánh xạ của ID ảnh con được báo hiệu tường minh trong SPS hoặc PPS.

Một cách tùy chọn, theo khía cạnh bất kỳ trong các khía cạnh trước đó, cách thực hiện khác của khía cạnh này đề cập đến việc dòng bit còn bao gồm cờ thay đổi CVS, và trong đó cờ thay đổi CVS chỉ báo liệu ánh xạ của ID ảnh con có thể thay đổi trong CVS của dòng bit.

Một cách tùy chọn, theo khía cạnh bất kỳ trong các khía cạnh trước đó, cách thực hiện khác của khía cạnh này đề cập đến việc dòng bit bao gồm dòng bit được hợp nhất, và trong đó ánh xạ của ID ảnh con đã thay đổi trong CVS của dòng bit.

Khía cạnh thứ ba đề cập đến bộ phận giải mã, bao gồm bộ nhận được tạo cấu hình để nhận dòng bit video bao gồm SPS, PPS, và các ảnh con được liên kết với ánh xạ của ID ảnh con, trong đó SPS bao gồm cờ SPS; bộ nhớ được ghép nối với bộ nhận, bộ nhớ lưu trữ các lệnh; và bộ xử lý được ghép nối với bộ nhớ, bộ xử lý được tạo cấu hình để thực thi các lệnh để khiến bộ phận giải mã xác định liệu cờ SPS có giá trị thứ nhất hoặc giá trị thứ hai, trong đó cờ SPS có giá trị thứ nhất xác định rằng ánh xạ của ID ảnh con được báo hiệu trong SPS, và trong đó cờ SPS có giá trị thứ hai xác định rằng ánh xạ của ID ảnh con được báo hiệu trong PPS; thu được ánh xạ của ID ảnh con từ SPS khi cờ SPS có giá trị thứ nhất và

từ PPS khi cờ SPS có giá trị thứ hai; và giải mã các ảnh con using ánh xạ của ID ảnh con.

Bộ phận giải mã đề cập đến các kỹ thuật đảm bảo báo hiệu hiệu quả các ID ảnh con thậm chí khi các ID ảnh con thay đổi trong CVS trong các kịch bản ứng dụng bao gồm cả trích xuất dòng bit thành phần và hợp nhất dòng bit thành phần. Báo hiệu hiệu quả được hoàn thành bằng cách thiết lập các cờ trong SPS hoặc SPS để chỉ báo liệu các ID ảnh con trong CVS có thể thay đổi và, nếu vậy, trong đó đặt các ID ảnh con. Điều này giảm độ dư thừa và tăng hiệu suất tạo mã. Do vậy, bộ mã hóa/bộ giải mã (cũng được biết đến là, “codec”) trong khi tạo mã video được cải thiện liên quan đến các codec hiện tại. Thực tế là, quá trình tạo mã video được cải tiến cho người dùng trải nghiệm sử dụng tốt hơn khi các video được gửi, được nhận, và/hoặc được xem.

Một cách tùy chọn, theo khía cạnh bất kỳ trong các khía cạnh trước đó, cách thực hiện khác của khía cạnh này đề cập đến việc cờ SPS có giá trị thứ nhất khi cờ PPS có giá trị thứ hai, trong đó cờ SPS có giá trị thứ hai khi cờ PPS có giá trị thứ nhất, và trong đó giá trị thứ nhất bằng 1 và giá trị thứ hai bằng 0.

Một cách tùy chọn, theo khía cạnh bất kỳ trong các khía cạnh trước đó, cách thực hiện khác của khía cạnh này đề cập đến việc SPS bao gồm cờ SPS thứ hai, và trong đó cờ SPS thứ hai xác định liệu việc ánh xạ của ID ảnh con được báo hiệu tường minh trong SPS hoặc PPS.

Một cách tùy chọn, theo khía cạnh bất kỳ trong các khía cạnh trước đó, cách thực hiện khác của khía cạnh này đề cập đến việc dòng bit còn bao gồm cờ thay đổi CVS, và trong đó cờ thay đổi CVS chỉ báo liệu ánh xạ của ID ảnh con có thể thay đổi trong CVS của dòng bit.

Một cách tùy chọn, theo khía cạnh bất kỳ trong các khía cạnh trước đó, cách thực hiện khác của khía cạnh này đề cập đến việc dòng bit bao gồm dòng bit được hợp nhất, và trong đó ánh xạ của ID ảnh con đã thay đổi trong CVS của dòng bit.

Khía cạnh thứ tư đề cập đến bộ phận mã hóa, bao gồm bộ nhớ chứa các lệnh; bộ xử lý được ghép nối với bộ nhớ, bộ xử lý được tạo cấu hình để thực hiện các lệnh để khiến bộ phận mã hóa thực hiện mã hóa dòng bit bao gồm SPS, PPS, và

các ảnh con được liên kết với ánh xạ ID ảnh con, trong đó SPS bao gồm cờ SPS; thiết lập cờ SPS bằng giá trị thứ nhất khi ánh xạ của ID ảnh con được báo hiệu trong SPS và bằng giá trị thứ hai khi ánh xạ của ID ảnh con được báo hiệu trong PPS; thiết lập cờ PPS trong PPS bằng giá trị thứ nhất khi ánh xạ của ID ảnh con được báo hiệu trong PPS và bằng giá trị thứ hai khi ánh xạ của ID ảnh con không được báo hiệu trong PPS; và bộ truyền được ghép nối với bộ xử lý, bộ truyền được tạo cấu hình để truyền dòng bit về phía bộ giải mã video.

Bộ phận mã hóa đề cập đến các kỹ thuật đảm bảo báo hiệu hiệu quả các ID ảnh con thậm chí khi các ID ảnh con thay đổi trong CVS trong các kịch bản ứng dụng bao gồm cả trích xuất dòng bit thành phần và hợp nhất dòng bit thành phần. Báo hiệu hiệu quả được hoàn thành bằng cách thiết lập các cờ trong SPS hoặc PPS để chỉ báo liệu các ID ảnh con trong CVS có thể thay đổi và, nếu vậy, trong đó đặt các ID ảnh con. Điều này giảm độ dư thừa và tăng hiệu suất tạo mã. Do vậy, bộ mã hóa/bộ giải mã (cũng được biết đến là “codec”) trong khi tạo mã video được cải thiện liên quan đến các codec hiện tại. Thực tế là, quá trình tạo mã video được cải tiến cho người dùng trải nghiệm sử dụng tốt hơn khi các video được gửi, được nhận, và/hoặc được xem.

Một cách tùy chọn, theo khía cạnh bất kỳ trong các khía cạnh trước đó, cách thực hiện khác của khía cạnh này đề cập đến việc cờ SPS có giá trị thứ nhất khi cờ PPS có giá trị thứ hai, và trong đó cờ SPS có giá trị thứ hai khi cờ PPS có giá trị thứ nhất.

Một cách tùy chọn, theo khía cạnh bất kỳ trong các khía cạnh trước đó, cách thực hiện khác của khía cạnh này đề cập đến việc giá trị thứ nhất bằng 1 và giá trị thứ hai bằng 0.

Một cách tùy chọn, theo khía cạnh bất kỳ trong các khía cạnh trước đó, cách thực hiện khác của khía cạnh này đề cập đến việc dòng bit còn bao gồm cờ thay đổi CVS, và trong đó cờ thay đổi CVS chỉ báo liệu ánh xạ của ID ảnh con có thể thay đổi trong CVS của dòng bit.

Một cách tùy chọn, theo khía cạnh bất kỳ trong các khía cạnh trước đó, cách thực hiện khác của khía cạnh này đề cập đến việc ID PPS xác định giá trị của ID

PPS thứ hai đối với PPS đang sử dụng, và trong đó ID PPS thứ hai nhận diện PPS để tham chiếu bằng các phân tử cú pháp.

Một cách tùy chọn, theo khía cạnh bất kỳ trong các khía cạnh trước đó, cách thực hiện khác của khía cạnh này đề cập đến việc dòng bit bao gồm dòng bit được hợp nhất, và trong đó ánh xạ của ID ảnh con đã thay đổi trong CVS của dòng bit.

Khía cạnh thứ năm đề cập đến thiết bị tạo mã. Thiết bị tạo mã bao gồm bộ nhận được tạo cấu hình để nhận ảnh để mã hóa hoặc để nhận dòng bit để giải mã; bộ truyền được ghép nối với bộ nhận, bộ truyền được tạo cấu hình để truyền dòng bit đến bộ giải mã hoặc để truyền ảnh được giải mã đến bộ phận hiển thị; bộ nhớ được ghép nối với ít nhất một trong bộ nhận hoặc bộ truyền, bộ nhớ được tạo cấu hình để lưu trữ các lệnh; và bộ xử lý được ghép nối với bộ nhớ, bộ xử lý được tạo cấu hình để thực thi các lệnh được lưu trữ trong bộ nhớ để thực hiện phương pháp bất kỳ trong các phương pháp được bộc lộ ở đây.

Thiết bị tạo mã đề cập đến các kỹ thuật đảm bảo báo hiệu hiệu quả các ID ảnh con thậm chí khi các ID ảnh con thay đổi trong CVS trong các kịch bản ứng dụng bao gồm cả trích xuất dòng bit thành phần và hợp nhất dòng bit thành phần. Báo hiệu hiệu quả được hoàn thành bằng cách thiết lập các cờ trong SPS hoặc SPS để chỉ báo liệu các ID ảnh con trong CVS có thể thay đổi và, nếu vậy, trong đó đặt các ID ảnh con. Điều này giảm độ dư thừa và tăng hiệu suất tạo mã. Do vậy, bộ mã hóa/bộ giải mã (cũng được biết đến là “codec”) trong khi tạo mã video được cải thiện liên quan đến các codec hiện tại. Thực tế là, quá trình tạo mã video được cải tiến cho người dùng trải nghiệm sử dụng tốt hơn khi các video được gửi, được nhận, và/hoặc được xem.

Một cách tùy chọn, theo khía cạnh bất kỳ trong các khía cạnh trước đó, cách thực hiện khác của khía cạnh này đề cập đến việc bộ phận hiển thị được tạo cấu hình để hiển thị ảnh được giải mã.

Khía cạnh thứ sáu đề cập đến hệ thống. Hệ thống bao gồm bộ mã hóa; và bộ giải mã truyền thông với bộ mã hóa, trong đó bộ mã hóa hoặc bộ giải mã bao gồm bộ phận giải mã, bộ phận mã hóa, hoặc thiết bị tạo mã được bộc lộ ở đây.

Hệ thống đề cập đến các kỹ thuật đảm bảo bảo hiệu hiệu quả các ID ảnh con thậm chí khi các ID ảnh con thay đổi trong CVS trong các kịch bản ứng dụng bao gồm cả trích xuất dòng bit thành phần và hợp nhất dòng bit thành phần. Bảo hiệu hiệu quả được hoàn thành bằng cách thiết lập các cờ trong SPS hoặc SPS để chỉ báo liệu các ID ảnh con trong CVS có thể thay đổi và, nếu vậy, trong đó đặt các ID ảnh con. Điều này giảm độ dư thừa và tăng hiệu suất tạo mã. Do vậy, bộ mã hóa/bộ giải mã (cũng được biết đến là “codec”) trong khi tạo mã video được cải thiện liên quan đến các codec hiện tại. Thực tế là, quá trình tạo mã video được cải tiến cho người dùng trải nghiệm sử dụng tốt hơn khi các video được gửi, được nhận, và/hoặc được xem.

Khía cạnh thứ bảy đề cập đến phương tiện tạo mã. Phương tiện tạo mã bao gồm phương tiện nhận được tạo cấu hình để nhận ảnh để mã hóa hoặc để nhận dòng bit để giải mã; phương tiện truyền được ghép nối với phương tiện nhận, phương tiện truyền được tạo cấu hình để truyền dòng bit đến phương tiện giải mã hoặc để truyền ảnh được giải mã đến phương tiện hiển thị; phương tiện lưu trữ được ghép nối với ít nhất một trong phương tiện nhận hoặc phương tiện truyền, phương tiện lưu trữ được tạo cấu hình để lưu trữ các lệnh; và phương tiện xử lý được ghép nối với phương tiện lưu trữ, phương tiện xử lý được tạo cấu hình để thực thi các lệnh được lưu trữ trong phương tiện lưu trữ để thực hiện phương pháp bất kỳ trong các phương pháp được bộc lộ ở đây.

Phương tiện tạo mã đề cập đến các kỹ thuật đảm bảo bảo hiệu hiệu quả các ID ảnh con thậm chí khi các ID ảnh con thay đổi trong CVS trong các kịch bản ứng dụng bao gồm cả trích xuất dòng bit thành phần và hợp nhất dòng bit thành phần. Bảo hiệu hiệu quả được hoàn thành bằng cách thiết lập các cờ trong SPS hoặc SPS để chỉ báo liệu các ID ảnh con trong CVS có thể thay đổi và, nếu vậy, trong đó đặt các ID ảnh con. Điều này giảm độ dư thừa và tăng hiệu suất tạo mã. Do vậy, bộ mã hóa/bộ giải mã (cũng được biết đến là “codec”) trong khi tạo mã video được cải thiện liên quan đến các codec hiện tại. Thực tế là, quá trình tạo mã video được cải tiến cho người dùng trải nghiệm sử dụng tốt hơn khi các video được gửi, được nhận, và/hoặc được xem.

Để minh họa rõ ràng, phương án bất kỳ trong các phương án thực hiện nêu trên có thể được kết hợp với một hoặc nhiều phương án trong các phương án thực hiện nêu trên khác để tạo phương án thực hiện mới trong phạm vi của sáng chế.

Các dấu hiệu này và các dấu hiệu khác sẽ được hiểu rõ hơn từ phần mô tả chi tiết cùng với các hình vẽ đi kèm và các điểm yêu cầu bảo hộ.

Mô tả vắn tắt các hình vẽ

Để hiểu rõ ràng hơn sáng chế, thực hiện tham khảo phần mô tả vắn tắt sau, dựa vào các hình vẽ đi kèm và phần mô tả chi tiết, trong đó các số chỉ dẫn giống nhau biểu diễn các bộ phận giống nhau.

Fig.1 là lưu đồ của phương pháp ví dụ tạo mã tín hiệu video;

Fig.2 là sơ đồ của hệ thống tạo mã và giải mã (codec) lấy làm ví dụ để tạo mã video;

Fig.3 là sơ đồ minh họa bộ mã hóa video lấy làm ví dụ;

Fig.4 là sơ đồ minh họa bộ giải mã video lấy làm ví dụ;

Fig.5 là sơ đồ minh họa dòng bit ví dụ chứa chuỗi video được mã hóa;

Fig.6A đến Fig.6E là các hình vẽ minh họa cơ chế ví dụ để tạo rãnh trích xuất để kết hợp các ảnh con có nhiều độ phân giải từ các dòng bit khác nhau thành một ảnh để sử dụng trong các ứng dụng thực tế ảo (virtual reality, VR);

Fig.7A đến Fig.7E là các hình vẽ minh họa cơ chế ví dụ để cập nhật rãnh trích xuất dựa trên thay đổi theo hướng nhìn trong các ứng dụng VR;

Fig.8 là hình vẽ minh họa phương án thực hiện của phương pháp giải mã dòng bit video được tạo mã;

Fig.9 là hình vẽ minh họa phương án thực hiện của phương pháp mã hóa dòng bit video được tạo mã;

Fig.10 là sơ đồ của thiết bị tạo mã video; và

Fig.11 là sơ đồ phương án thực hiện phương tiện tạo mã.

Mô tả chi tiết sáng chế

Ngay từ đầu cần hiểu rằng mặc dù cách thức triển khai minh họa của một hoặc nhiều phương án thực hiện được nêu dưới đây, các hệ thống và/hoặc các phương pháp được bộc lộ có thể được thực hiện nhờ sử dụng số lượng kỹ thuật bất kỳ, liệu đang được biết hoặc hiện có. Sáng chế sẽ không bị giới hạn ở các cách thực hiện minh họa, hình vẽ, và các kỹ thuật được minh họa dưới đây, bao gồm các thiết kế và các triển khai lấy làm ví dụ được minh họa và mô tả ở đây, nhưng có thể được chỉnh sửa trong phạm vi của các điểm yêu cầu bảo hộ đi kèm cùng với toàn bộ phạm vi của các điểm tương đương.

Các cụm từ sau được định nghĩa như sau trừ khi có thông báo khác. Cụ thể là, các định nghĩa sau sẽ giúp làm rõ ràng hơn với sáng chế. Tuy nhiên, các cụm từ có thể được mô tả khác đi trong các ngữ cảnh khác nhau. Do đó, các định nghĩa sau nên được xem xét như là bổ sung và sẽ không được xem xét như là giới hạn các định nghĩa khác bất kỳ của các phần mô tả được nêu đối với các cụm từ này ở đây.

Dòng bit là chuỗi bit bao gồm dữ liệu video được nén để truyền giữa bộ mã hóa và bộ giải mã. Bộ mã hóa là thiết bị được tạo cấu hình để sử dụng các quá trình mã hóa để nén dữ liệu video thành dòng bit. Bộ giải mã là thiết bị được tạo cấu hình để sử dụng các quá trình giải mã để tái tạo dữ liệu video từ dòng bit để hiển thị. Ảnh là mảng của các mẫu độ sáng và/hoặc mảng của các mẫu sắc độ tạo khung hoặc trường của nó. Ảnh được mã hóa hoặc giải mã có thể được gọi là ảnh hiện tại để mô tả rõ ràng. Ảnh tham chiếu là ảnh chứa các mẫu tham chiếu có thể được sử dụng khi tạo mã các ảnh khác bằng tham chiếu theo dự báo ngoài và/hoặc dự báo ngoài lớp. Danh sách ảnh tham chiếu là danh sách các ảnh tham chiếu được sử dụng cho dự báo ngoài và/hoặc dự báo ngoài lớp. Một số hệ thống tạo mã video sử dụng hai danh sách ảnh tham chiếu, có thể được ký hiệu là danh sách ảnh tham chiếu 1 và danh sách ảnh tham chiếu 0. Cấu trúc danh sách ảnh tham chiếu là cấu trúc cú pháp có thể đánh địa chỉ chứa nhiều danh sách ảnh tham chiếu. Dự báo ngoài là cơ chế tạo mã các mẫu của ảnh hiện tại bằng tham chiếu cho các mẫu được chỉ báo trong ảnh tham chiếu khác với ảnh hiện tại trong đó ảnh tham chiếu và ảnh hiện tại nằm trong cùng lớp. Mục của cấu trúc danh

sách ảnh tham chiếu là vị trí có thể đánh địa chỉ trong cấu trúc danh sách ảnh tham chiếu chỉ báo ảnh tham chiếu được liên kết với danh sách ảnh tham chiếu. Tiêu đề lát là một phần của lát được tạo mã chứa các phần tử dữ liệu gắn với tất cả dữ liệu video trong ô được biểu diễn trong lát. SPS là tập hợp tham số chứa dữ liệu liên quan đến chuỗi ảnh. PPS là cấu trúc cú pháp chứa các phần tử cú pháp áp dụng cho 0 hoặc nhiều ảnh được tạo mã toàn bộ như được xác định bởi phần tử cú pháp được tìm thấy trong mỗi tiêu đề ảnh.

Cờ là biến hoặc phần tử cú pháp một bit có thể chiếm một trong hai giá trị khả thi: 0 và 1. Ảnh con là vùng chữ nhật của một hoặc nhiều lát trong ảnh. ID ảnh con là số, chữ, hoặc các dấu chỉ dẫn khác mà nhận diện duy nhất ảnh con. Các ID ảnh con (cũng được biết đến là các ID ô) được sử dụng để nhận diện các ảnh con cụ thể nhờ sử dụng chỉ số ảnh con, có thể được gọi ở đây là ánh xạ của ID ảnh con. Nói cách khác, ánh xạ của ID ảnh con là bảng tạo ánh xạ một – một giữa danh sách của các chỉ số ảnh con và các ID ảnh con. Tức là, ánh xạ của ID ảnh con tạo ID ảnh con riêng biệt cho mỗi ảnh con.

Khối truy nhập (AU) là tập hợp của một hoặc nhiều ảnh được mã hóa được liên kết với cùng thời gian hiển thị (chẳng hạn, cùng số đếm thứ tự ảnh) để xuất ra từ bộ đệm ảnh được giải mã (decoded picture buffer, DPB) (chẳng hạn, để hiển thị đến người dùng). Bộ phân cách khối truy nhập (access unit delimiter, AUD) là bộ chỉ báo hoặc cấu trúc dữ liệu được sử dụng để chỉ báo bắt đầu AU hoặc đường biên giữa các AU. Chuỗi video được giải mã là chuỗi ảnh đã được tái tạo bằng bộ giải mã đang chuẩn bị để hiển thị đến người dùng.

CVS là chuỗi AU bao gồm, theo thứ tự giải mã, AU bắt đầu cuối video được tạo mã (coded video sequence start, CVSS), tiếp theo bởi 0 hoặc nhiều AU không phải là các AU CVSS, bao gồm tất cả các AU tiếp theo có tới mà không bao gồm bất kỳ AU tiếp theo vốn là AU CVSS. AU CVSS là AU trong đó có khối dự báo (prediction unit, PU) cho mỗi lớp được xác định bởi tập hợp tham số video (video parameter set, VPS) và ảnh được tạo mã trong mỗi PU là ảnh CVSS. Theo phương án thực hiện, mỗi ảnh nằm trong AU. PU là tập hợp đơn vị lớp trừu tượng mạng (Network Abstraction Layer, NAL) được liên kết với nhau

theo quy tắc phân loại cụ thể, liên tiếp theo thứ tự giải mã, và chứa chính xác một ảnh được tạo mã.

Các từ viết tắt sau được sử dụng ở đây: bộ lọc vòng thích ứng (Adaptive Loop Filter, ALF), khối cây mã (Coding Tree Block, CTB), đơn vị cây mã (Coding Tree Unit, CTU), đơn vị mã (Coding Unit, CU), chuỗi video được mã hóa (Coded Video Sequence, CVS), bộ đệm ảnh được giải mã (Decoded Picture Buffer, DPB), làm mới giải mã tức thời (Instantaneous Decoding Refresh, IDR), điểm truy nhập ngẫu nhiên nội bộ (Intra-Random Access Point, IRAP), nhóm chuyên gia video chung (Joint Video Experts Team, JVET), bit có trọng số nhỏ nhất (Least Significant Bit, LSB), bit có trọng số lớn nhất (Most Significant Bit, MSB), tập hợp ô bị giới hạn chuyển động (Motion-Constrained Tile Set, MCTS), đơn vị truyền lớn nhất (Maximum Transfer Unit, MTU), lớp trừu tượng mạng (Network Abstraction Layer, NAL), số đếm thứ tự ảnh (Picture Order Count, POC), tập hợp tham số ảnh (Picture Parameter Set, PPS), tải tin chuỗi byte thô (Raw Byte Sequence Payload, RBSP), độ lệch mẫu thích ứng (Sample Adaptive Offset, SAO), tập hợp tham số chuỗi (Sequence Parameter Set, SPS), dự báo vectơ chuyển động thời gian (Temporal Motion Vector Prediction, TMVP), tạo mã video đa dụng (Versatile Video Coding, VVC), và phác thảo làm việc (Working Draft, WD).

Nhiều kỹ thuật nén video có thể được sử dụng để giảm kích thước của các tệp tin video có tổn hao dữ liệu nhỏ nhất. Chẳng hạn, các kỹ thuật nén video có thể bao gồm bước thực hiện dự báo không gian (chẳng hạn, trong ảnh) và/hoặc dự báo thời gian (chẳng hạn, ngoài ảnh) để giảm hoặc loại bỏ dư thừa dữ liệu trong các chuỗi video. Để tạo mã video dựa trên khối, lát video (chẳng hạn, ảnh video hoặc một phần của ảnh video) có thể được phân vùng thành các khối video, cũng có thể được gọi là các khối cây, các CTB, các CTU, các CU, và/hoặc các nút mã. Các khối video trong lát được tạo mã trong (I) của ảnh được mã hóa nhờ sử dụng dự báo không gian với các mẫu tham chiếu trong các khối lân cận trong cùng ảnh. Các khối video trong lát dự báo đơn hướng được tạo mã ngoài (P) hoặc dự báo hai hướng (B) của ảnh có thể được tạo mã nhờ sử dụng dự báo không gian với các mẫu tham chiếu trong các khối lân cận trong cùng ảnh hoặc dự báo thời

gian với các mẫu tham chiếu trong các ảnh tham chiếu khác. Các ảnh có thể được gọi là các khung và/hoặc các ảnh, và các ảnh tham chiếu có thể được gọi là khung tham chiếu và/hoặc ảnh tham chiếu. Dự báo không gian hoặc thời gian dẫn đến khối dự báo biểu diễn khối ảnh. Dữ liệu dư là các độ lệch pixel giữa khối ảnh ban đầu và khối dự báo. Do đó, khối được tạo mã ngoài được mã hóa theo vector chuyển động trở đến khối của các mẫu tham chiếu tạo khối dự báo và dữ liệu dư chỉ báo hiệu số giữa khối được tạo mã và khối dự báo. Khối được tạo mã trong được mã hóa theo chế độ tạo mã trong và dữ liệu dư. Để nén tiếp, dữ liệu dư có thể được biến đổi từ miền pixel sang miền biến đổi. Điều này dẫn đến các hệ số biến đổi dư, có thể được lượng tử. Các hệ số biến đổi được lượng tử có thể ban đầu được bố trí trong mảng hai chiều. Các hệ số biến đổi được lượng tử có thể được quét để tạo vector một chiều của các hệ số biến đổi. Kỹ thuật tạo mã entropy có thể được áp dụng để nén được nhiều hơn. Các kỹ thuật nén video này được nêu chi tiết hơn dưới đây.

Để đảm bảo video được mã hóa có thể được giải mã chính xác, video được mã hóa và giải mã theo các chuẩn tạo mã video tương ứng. Các chuẩn tạo mã video bao gồm ban chuẩn hóa liên đoàn viễn thông quốc tế ((International Telecommunication Union, ITU) Standardization Sector, ITU-T)) H.261, tổ chức quốc tế để chuẩn hóa/hội đồng cơ điện quốc tế (International Organization for Standardization/International Electrotechnical Commission, ISO/IEC) nhóm chuyên gia ảnh động (Motion Picture Experts Group, MPEG)-1 Part 2, ITU-T H.262 hoặc ISO/IEC MPEG-2 Part 2, ITU-T H.263, ISO/IEC MPEG-4 Part 2, tạo mã video cao cấp (Advanced Video Coding, AVC), cũng được biết đến là ITU-T H.264 hoặc ISO/IEC MPEG-4 Part 10, và kỹ thuật tạo mã video hiệu suất cao (High Efficiency Video Coding, HEVC), cũng được biết đến là ITU-T H.265 hoặc MPEG-H Part 2. AVC bao gồm các phần mở rộng, chẳng hạn, tạo mã video khả mở (Scalable Video Coding, SVC), tạo mã video đa khung hình (Multiview Video Coding, MVC) và MVC có chiều sâu (Multiview Video Coding plus Depth, MVC+D), và AVC ba chiều (three dimensional, 3D) (3D-AVC). HEVC bao gồm các phần mở rộng chẳng hạn, HEVC khả mở (Scalable

HEVC, SHVC), HEVC đa khung hình (Multiview HEVC, MV-HEVC), và 3D HEVC (3D-HEVC).

Cũng có chuẩn tạo mã video mới, được đặt tên là VVC, được phát triển bởi nhóm chuyên gia video chung (joint video experts team, JVET) của ITU-T và ISO/IEC. Trong khi chuẩn VVC có vài WD, một WD của VVC cụ thể là, có tên là B. Bross, J. Chen, và S. Liu, “Versatile Video Coding (Draft 5),” JVET-N1001-v3, 13th JVET Meeting, March 27, 2019 (VVC Draft 5) được viện dẫn ở đây.

HEVC bao gồm bốn sơ đồ phân vùng ảnh khác nhau, tức là các lát thường, các lát phụ thuộc, các ô, và xử lý song song mặt sóng (Wavefront Parallel Processing, WPP), có thể được áp dụng để so khớp kích thước đơn vị truyền lớn nhất (Maximum Transfer Unit, MTU), xử lý song song, và độ trễ giảm từ đầu này đến đầu kia.

Các lát thường tương tự như trong H.264/AVC. Mỗi lát thường được đóng gói trong đơn vị NAL của riêng nó, và dự báo trong ảnh (dự báo trong mẫu, dự báo thông tin chuyển động, dự báo chế độ mã) và lệ thuộc tạo mã entropy giữa các đường biên lát bị vô hiệu hóa. Do vậy, lát thường có thể được tái tạo độc lập với các lát thường khác trong cùng ảnh (mặc dù vẫn có thể có các tương thuộc do các hoạt động lọc vòng).

Lát thường là công cụ duy nhất có thể được sử dụng để tạo song song mà cũng khả dụng, ở dạng ảo giống hệt, trong chuẩn H.264/AVC. Việc song song hóa dựa trên lát thường không yêu cầu nhiều truyền thông liên bộ xử lý hoặc ngoài lõi (ngoại trừ đối với việc chia sẻ dữ liệu liên bộ xử lý hoặc ngoài lõi để bù chuyển động khi giải mã ảnh được mã hóa dự báo, thường nặng hơn nhiều chia sẻ dữ liệu ngoài bộ xử lý hoặc ngoài lõi do dự báo trong ảnh). Tuy nhiên, đối với cùng nguyên nhân, việc sử dụng các lát thường có thể phát sinh phụ thu tạo mã đáng kể do chi phí bit của tiêu đề lát và do thiếu dự báo giữa các đường biên lát. Ngoài ra, các lát thường (ngược lại với các công cụ khác được nêu dưới đây) cũng phục vụ làm cơ cấu chính để phân vùng dòng bit để khớp với các yêu cầu kích thước MTU, do sự độc lập trong ảnh với các lát thường và việc mỗi lát thường được đóng gói trong đơn vị NAL của nó. Trong nhiều trường hợp, mục

tiêu song song hóa và mục tiêu so khớp kích thước MTU đặt ra các yêu cầu ngược lại với bố trí lát trong ảnh. Việc thực hiện tình huống này dẫn đến việc phát triển các công cụ song song hóa được nêu dưới đây.

Các lát phụ thuộc có các tiêu đề lát ngăn và cho phép phân vùng dòng bit ở các đường biên khối cây mà không phá hỏng bất kỳ dự báo trong ảnh. Về cơ bản, các lát phụ thuộc phân mảnh các lát thường thành nhiều đơn vị NAL để giảm độ trễ từ đầu này đến đầu kia bằng cách cho phép một phần lát thường được gửi ra trước hoàn thành mã hóa toàn bộ lát thường.

Trong WPP, ảnh được phân vùng thành các hàng đơn của các CTB. Giải mã entropy và dự báo được phép sử dụng dữ liệu từ các CTB trong các phân vùng khác. Kỹ thuật xử lý song song có thể qua giải mã song song các hàng CTB, trong đó thời điểm bắt đầu giải mã hàng CTB bị trễ bởi hai CTB, để đảm bảo rằng dữ liệu liên quan đến CTB nêu trên và sang bên phải của CTB đối tượng là khả dụng trước khi CTB đối tượng đang được giải mã. Nhờ sử dụng điểm bắt đầu so le (xuất hiện như mặt sóng khi được biểu diễn dưới dạng đồ họa), có thể song song với nhiều bộ xử lý/lõi như ảnh chứa các hàng CTB. Do dự báo trong ảnh giữa các hàng khối cây lân cận trong ảnh được phép, truyền thông ngoài bộ xử lý/ngoài lõi để kích hoạt dự báo trong ảnh có thể là đáng kể. Việc phân vùng WPP không dẫn đến tạo thêm các đơn vị NAL bổ sung so với khi nó không được áp dụng, do vậy WPP không phải là công cụ để so khớp kích thước MTU. Tuy nhiên, nếu so khớp kích thước MTU cần thiết, các lát thường có thể được sử dụng với WPP, với phụ tải tạo mã cụ thể.

Các ô định nghĩa các đường biên ngang và dọc mà phân vùng ảnh thành các cột và các hàng ô. Thứ tự quét của các CTB được thay đổi để thành cục bộ trong ô ảnh (theo thứ tự quét mảnh CTB của ô), trước khi giải mã CTB trên bên trái của ô tiếp theo trong thứ tự quét mảnh ô của ảnh. Giống như các lát thường, các ô phá vỡ các lệ thuộc dự báo trong ảnh cũng như các lệ thuộc giải mã entropy. Tuy nhiên, cũng không cần được bao gồm trong các đơn vị NAL riêng rẽ (cũng như WPP theo khía cạnh này); do vậy các ô không thể được sử dụng để so khớp kích thước MTU. Mỗi ô có thể được xử lý bởi một bộ xử lý/lõi, và truyền thông ngoài bộ xử lý/ngoài lõi được yêu cầu để dự báo trong ảnh giữa các khối xử lý

giải mã các ô lân cận bị giới hạn ở việc vận chuyển tiêu đề lát được chia sẻ trong các trường hợp trong đó lát mở rộng nhiều hơn một ô, và chia sẻ liên quan dọc vòng của các mẫu và siêu dữ liệu được tái tạo. Khi nhiều hơn một ô hoặc đoạn WPP được bao gồm trong lát, độ lệch byte điểm đầu vào hoặc đoạn WPP khác ngoài đoạn thứ nhất trong lát được báo hiệu trong tiêu đề lát.

Để cho đơn giản, các giới hạn ở việc áp dụng bốn phương tiện phân vùng ảnh khác nhau đã được nêu trong HEVC. CVS cụ thể không thể bao gồm cả ô lẫn các mặt sóng cho phần lớn tiêu đề sử dụng được nêu trong HEVC. Đối với mỗi lát và ô, một trong hoặc cả hai điều kiện sau phải được thỏa mãn: 1) tất cả các khối cây được tạo mã trong lát thuộc cùng ô; 2) tất cả các khối cây được tạo mã trong ô thuộc cùng lát. Cuối cùng, đoạn mặt sóng chứa chính xác một hàng CTB, và khi WPP đang sử dụng, nếu lát bắt đầu trong hàng CTB, phải kết thúc trong cùng hàng CTB.

Bổ sung gần đây với HEVC được xác định trong tài liệu đầu ra JCT-VC JCTVC-AC1005, J. Boyce, A. Ramasubramonian, R. Skupin, G. J. Sullivan, A. Tourapis, Y.-K. Wang (editors), "HEVC Additional Supplemental Enhancement Information (Draft 4)," Oct. 24, 2017, công khai tại địa chỉ: http://phenix.int-evry.fr/jct/doc_end_user/documents/29_Macau/wg11/JCTVC-AC1005-v2.zip.

Với sửa đổi được bao gồm, HEVC xác định ba thông điệp thông tin tăng cường bổ sung (supplemental enhancement information, SEI) liên quan MCTS, tức là thông điệp SEI liên quan MCTS thời gian, thông điệp SEI của tập hợp thông tin trích xuất MCTS, và thông điệp SEI lồng thông tin trích xuất các MCTS.

Thông điệp SEI liên quan MCTS thời gian chỉ báo sự tồn tại của các MCTS trong dòng bit và báo hiệu các MCTS. Đối với mỗi MCTS, các vector chuyển động bị giới hạn để trở đến các vị trí mẫu nguyên bên trong MCTS và đến các vị trí mẫu phân số đòi hỏi chỉ các vị trí mẫu nguyên bên trong MCTS để nội suy, và việc sử dụng các ứng viên vector chuyển động để dự báo MV thời gian được dẫn xuất từ các khối bên ngoài MCTS không được phép. Theo cách này, mỗi MCTS có thể được giải mã độc lập mà không có các ô không được bao gồm trong MCTS.

Thông điệp SEI của các tập hợp thông tin trích xuất các MCTS cung cấp thông tin bổ sung có thể được sử dụng trong trích xuất dòng bit thành phần MCTS (được xác định như là một phần ngữ nghĩa của thông điệp SEI) để tạo dòng bit tương hợp cho tập hợp MCTS. Thông tin bao gồm số lượng tập hợp thông tin trích xuất, mà mỗi tập hợp định nghĩa số lượng tập hợp MCTS và chứa các byte RBSP của các byte của các VPS, các SPS, và các PPS thay thế cần được sử dụng trong quá trình trích xuất dòng bit thành phần MCTS. Khi trích xuất dòng bit thành phần theo quá trình trích xuất dòng bit thành phần MCTS, các tập hợp tham số (VPS, SPS, và PPS) cần được viết lại hoặc thay thế, các tiêu đề lát cần được cập nhật một chút do một hoặc tất cả các phần tử cú pháp liên quan địa chỉ lát (bao gồm `first_slice_segment_in_pic_flag` và `slice_segment_address`) thường sẽ cần có các giá trị khác nhau.

Trong đặc tả phác thảo VVC mới nhất, ảnh có thể được phân vùng thành nhiều ảnh con, mỗi ảnh con bao phủ vùng chữ nhật và chứa số lượng nguyên các lát hoàn chỉnh. Quá trình phân vùng ảnh con tiếp tục trên tất cả các ảnh trong CVS, và thông tin phân vùng (tức là, các vị trí và các kích thước ảnh con) được báo hiệu trong SPS. Ảnh con có thể được chỉ báo như được tạo mã mà không sử dụng các giá trị mẫu từ ảnh con khác bất kỳ để bù chuyển động.

Trong thành phần JVET JVET-O0141, công khai tại địa chỉ: http://phenix.it-sudparis.eu/jvet/doc_end_user/documents/13_Marrakech/wg11/JVET-M0261-v1.zip, thiết kế ảnh con giống như trong đặc tả phác thảo VVC mới nhất với một số khác biệt, một trong số chúng là báo hiệu tường minh của ID ảnh con đối với mỗi ảnh con trong SPS cũng như báo hiệu ID ảnh con trong mỗi tiêu đề lát để kích hoạt trích xuất dòng bit thành phần dựa trên ảnh con mà không cần thay đổi các đơn vị NAL lát được tạo mã. Theo cách này, ID ảnh con đối với mỗi ảnh con luôn không đổi giữa tất cả các ảnh trong CVS.

Trong các thảo luận trước đây, đề cập đến việc cách báo hiệu khác về các ID ảnh con cần được sử dụng cũng để kích hoạt hợp nhất dòng bit thành phần dựa trên ảnh con mà không cần thay đổi các đơn vị NAL lát được tạo mã.

Chẳng hạn, và như sẽ được giải thích đầy đủ hơn dưới đây, khi phát trực tuyến video 360° phụ thuộc viewport (cổng nhìn), máy khách chọn ảnh con nào

được nhận và được hợp nhất vào dòng bit cần được giải mã. Khi hướng xem thay đổi, một số giá trị ID ảnh con có thể vẫn y nguyên trong khi các giá trị ID ảnh con khác được thay đổi so với các giá trị ID ảnh con được nhận trước khi thay đổi hướng nhìn. Do vậy, các ID ảnh con mà bộ giải mã nhận trong dòng bit được hợp nhất thay đổi khi hướng nhìn thay đổi.

Một cách đã được xem xét như sau.

Các vị trí và kích thước ảnh con được chỉ báo trong SPS nhờ sử dụng, chẳng hạn, cú pháp trong bản mô tả phác thảo VVC mới nhất. Các chỉ số ảnh con được gán (từ 0 đến $N-1$, trong đó N là số lượng ảnh con). Việc ánh xạ ID ảnh con đến chỉ số ảnh con nằm trong PPS do việc ánh xạ đó có thể cần được thay đổi ở giữa CVS. Đây có thể là vòng lặp từ $i = 0$ đến $N-1$, giả sử `subpicture_id[i]` ánh xạ đến chỉ số ảnh con i . Ánh xạ của ID ảnh con này cần được viết lại bởi máy khách khi máy khách lựa chọn tập hợp ảnh con mới cần được giải mã. Tiêu đề lát chứa ID ảnh con (chẳng hạn, `subpicture_id`), bộ mã hóa nào được chọn (và điều này không cần được viết lại trong hợp nhất dòng bit thành phần).

Không có thể miễn, có các vấn đề với cách thức báo hiệu ID ảnh con hiện tại. Nhiều kịch bản ứng dụng bằng cách sử dụng tạo mã video dựa trên ảnh con bao gồm trích xuất dòng bit thành phần mà không hợp nhất dòng bit thành phần. Chẳng hạn, mỗi dòng bit thành phần được trích xuất có thể được giải mã bằng thực thể bộ giải mã của riêng chúng. Thực tế, không cần hợp nhất các dòng bit thành phần được trích xuất thành một dòng bit cần được giải mã bởi chỉ một thực thể bộ giải mã. Trong các kịch bản này, ID ảnh con cho mỗi ảnh con không thay đổi trong CVS. Do vậy, các ID ảnh con có thể được báo hiệu trong SPS. Việc báo hiệu của các ID ảnh con này trong SPS thay vì PPS có lợi theo quan điểm tiết kiệm bit lẫn quan điểm thương lượng phiên.

Trong các kịch bản ứng dụng bao gồm cả trích xuất dòng bit thành phần và hợp nhất dòng bit thành phần, các ID ảnh con không thay đổi trong CVS trong các dòng bit ban đầu. Do vậy, việc báo hiệu các ID ảnh con trong SPS kích hoạt sử dụng các ID ảnh con trong cả SPS và trong các tiêu đề lát, vốn hữu ích cho phần trích xuất dòng bit thành phần của quá trình này. Tuy nhiên, cách này chỉ linh hoạt khi các ID ảnh con không thay đổi trong CVS trong dòng bit ban đầu.

Các kỹ thuật được bộc lộ ở đây đảm bảo bảo hiệu hiệu quả của các ID ảnh con thậm chí khi các ID ảnh con thay đổi trong CVS trong các kịch bản ứng dụng bao gồm cả trích xuất dòng bit thành phần và hợp nhất dòng bit thành phần. Bảo hiệu hiệu quả được hoàn thành bằng cách thiết lập các cờ trong SPS hoặc SPS để chỉ báo liệu các ID ảnh con trong CVS có thể thay đổi và, nếu vậy, trong đó đặt các ID ảnh con. Điều này giảm độ dư thừa và tăng hiệu suất tạo mã. Do vậy, bộ mã hóa/bộ giải mã (cũng được biết đến là “codec”) trong khi tạo mã video được cải thiện liên quan đến các codec hiện tại. Thực tế là, quá trình tạo mã video được cải tiến cho người dùng trải nghiệm sử dụng tốt hơn khi các video được gửi, được nhận, và/hoặc được xem.

Fig.1 là lưu đồ của phương pháp hoạt động 100 lấy làm ví dụ tạo mã tín hiệu video. Cụ thể là, tín hiệu video được mã hóa ở bộ mã hóa. Quá trình mã hóa nén tín hiệu video nhờ sử dụng các cơ chế khác nhau để giảm kích thước tệp tin video. Kích thước tệp tin nhỏ hơn cho phép tệp tin video được nén được truyền về phía người dùng, trong khi giảm phụ tải băng thông liên kết. Sau đó, bộ giải mã giải mã tệp tin video được nén để tái tạo tín hiệu video ban đầu để hiển thị đến người dùng cuối. Quá trình giải mã thường phản chiếu quá trình mã hóa để cho phép bộ giải mã tái tạo nhất quán tín hiệu video.

Ở bước 101, tín hiệu video được đưa vào bộ mã hóa. Chẳng hạn, tín hiệu video có thể là tệp tin video không được nén được lưu trữ trong bộ nhớ. Lấy ví dụ khác, tệp tin video có thể được ghi lại bởi thiết bị ghi lại video, chẳng hạn video camera, và được mã hóa để hỗ trợ live streaming (phát trực tuyến trực tiếp) video. Tệp tin video có thể bao gồm cả thành phần audio lẫn thành phần video. Thành phần video chứa chuỗi khung ảnh, khi được xem theo trình tự, tạo ấn tượng chuyển động bằng hình ảnh. Các khung chứa các pixel được biểu diễn liên quan ánh sáng, ở đây được gọi là các thành phần độ sáng (hoặc các mẫu độ sáng), và màu sắc, được gọi là các thành phần sắc độ (hoặc các mẫu màu sắc). Trong một số ví dụ, các khung cũng có thể chứa các giá trị độ sâu để hỗ trợ xem 3D.

Ở bước 103, video được phân vùng thành các khối. Việc phân vùng bao gồm bước phân chia phụ các pixel trong mỗi khung thành các khối hình vuông

và/hoặc chữ nhật để nén. Chẳng hạn, trong HEVC (cũng được biết là H.265 và MPEG-H Part 2), khung có thể trước hết được phân chia thành các CTU, vốn là các khối có kích thước định trước (chẳng hạn, 64 pixel x 64 pixel). Các CTU chứa cả các mẫu độ sáng và sắc độ. Các cây mã có thể được sử dụng để phân chia các CTU thành các khối và sau đó phân chia phụ đệ quy các khối cho đến khi đạt được các cấu hình hỗ trợ mã hóa tiếp. Chẳng hạn, các thành phần độ sáng của khung có thể được phân chia phụ cho đến khi các khối riêng rẽ chứa các giá trị độ sáng tương đối đồng nhất. Ngoài ra, các thành phần sắc độ của khung có thể được phân chia phụ cho đến khi các khối riêng rẽ chứa các giá trị màu tương đối đồng nhất. Do đó, các cơ chế phân vùng thay đổi phụ thuộc vào nội dung của các khung video.

Ở bước 105, các cơ chế nén khác nhau được sử dụng để nén các khối ảnh được phân vùng ở bước 103. Chẳng hạn, dự báo ngoài và/hoặc dự báo trong có thể được sử dụng. Dự báo ngoài được thiết kế để tận dụng thực tế rằng các đối tượng trong khung cảnh thông thường có xu hướng xuất hiện trong các khung liên tiếp. Do đó, khối mô tả đối tượng trong khung tham chiếu cần được mô tả lặp lại trong các khung liên kế. Cụ thể là, đối tượng, chẳng hạn bàn, có thể vẫn ở vị trí không đổi trên nhiều khung. Do vậy, bảng được mô tả một lần và các khung liên kế có thể quay ngược lại khung tham chiếu. Các cơ chế so khớp khuôn mẫu có thể được sử dụng để so khớp các đối tượng trên nhiều khung. Ngoài ra, các đối tượng di chuyển có thể được biểu diễn trên nhiều khung, chẳng hạn do di chuyển đối tượng hoặc di chuyển camera. Lấy ví dụ cụ thể, video có thể thể hiện xe ô tô di chuyển dọc màn hình trên nhiều khung. Các vector chuyển động có thể được sử dụng để mô tả chuyển động này. Vector chuyển động là vector hai chiều tạo độ lệch từ các tọa độ của đối tượng trong khung đến các tọa độ của đối tượng trong khung tham chiếu. Thực tế, dự báo ngoài có thể mã hóa khối ảnh trong khung hiện tại như là tập hợp MV chỉ báo độ lệch từ khối tương ứng trong khung tham chiếu.

Dự báo trong mã hóa các khối trong khung thường. Dự báo trong tận dụng thực tế rằng các thành phần độ sáng và sắc độ có xu hướng tạo cụm trong khung. Chẳng hạn, mảng lục trong phần cây có xu hướng được đặc lần các mảng lục

giống nhau. Dự báo trong sử dụng nhiều chế độ dự báo định hướng (chẳng hạn, 33 trong HEVC), chế độ phẳng, và chế độ dòng một chiều (direct current, DC). Các chế độ định hướng chỉ báo rằng khối hiện tại là tương tự/giống với các mẫu của khối lân cận theo hướng tương ứng. Chế độ phẳng chỉ báo rằng chuỗi khối dọc theo hàng/cột (chẳng hạn, mặt phẳng) có thể được nội suy dựa trên các khối lân cận ở các mép của hàng. Chế độ phẳng, thực tế, chỉ báo dịch chuyển tron tru ánh sáng/màu sắc giữa hàng/cột nhờ sử dụng đường cong tương đối không đổi trong các giá trị thay đổi. Chế độ DC được sử dụng để làm mượt đường biên và chỉ báo rằng khối tương tự/giống như giá trị trung bình được liên kết với các mẫu của tất cả các khối lân cận được liên kết với các hướng góc của các chế độ dự báo định hướng. Do đó, các khối dự báo trong có thể biểu diễn các khối ảnh dưới dạng các giá trị chế độ dự báo quan hệ khác nhau thay cho các giá trị thực. Ngoài ra, các khối dự báo ngoài có thể biểu diễn các khối ảnh dưới dạng các giá trị MV thay cho các giá trị thực. Trong trường hợp bất kỳ, các khối dự báo không thể biểu diễn chính xác các khối ảnh trong một số trường hợp. Bất kỳ khác biệt nào được lưu trữ trong các khối dư. Các phép biến đổi có thể được áp dụng cho các khối dư để nén tiếp tệp tin này.

Ở bước 107, các kỹ thuật lọc khác nhau có thể được áp dụng. Trong chuẩn HEVC, các bộ lọc được áp dụng theo phương tiện lọc trong vòng. Dự báo dựa trên khối nêu trên có thể dẫn đến việc tạo ra các ảnh có khối ở bộ giải mã. Ngoài ra, phương tiện dự báo dựa trên khối có thể mã hóa khối và sau đó tái tạo khối được mã hóa để sử dụng sau này làm khối tham chiếu. Phương tiện lọc trong vòng áp dụng lặp lại các bộ lọc chặn tạp âm, các bộ lọc tách khối, các ALF, và các bộ lọc SAO cho các khối/khung. Các bộ lọc này di trú các giả tượng khối này sao cho tệp tin được mã hóa có thể được tái tạo chính xác. Ngoài ra, các bộ lọc này di trú các giả tượng trong các khối tham chiếu được tái tạo sao cho các giả tượng ít khả năng hơn trong việc tạo các giả tượng bổ sung trong các khối tiếp theo mà được mã hóa dựa trên các khối tham chiếu được tái tạo.

Khi tín hiệu video đã được phân vùng, được nén, và được lọc, dữ liệu thu được sẽ được mã hóa trong dòng bit ở bước 109. Dòng bit bao gồm dữ liệu được nêu trên cũng như bất kỳ dữ liệu báo hiệu nào cần hỗ trợ tái tạo tín hiệu video

đúng ở bộ giải mã. Chẳng hạn, dữ liệu này có thể bao gồm dữ liệu phân vùng, dữ liệu dự báo, các khối dư, và các cờ khác nhau cung cấp các lệnh mã cho bộ giải mã. Dòng bit có thể được lưu trữ trong bộ nhớ để truyền đến bộ giải mã theo yêu cầu. Dòng bit cũng có thể được phát quảng bá và/hoặc phát đa hướng về phía các bộ giải mã. Việc tạo dòng bit là quá trình lặp. Do đó, các bước 101, 103, 105, 107, và 109 có thể xuất hiện liên tục và/hoặc đồng thời trên nhiều khung và khối. Thứ tự được thể hiện trên Fig.1 được nêu cho rõ ràng và dễ thảo luận, và không nhằm giới hạn quá trình tạo mã video ở thứ tự cụ thể.

Bộ giải mã nhận dòng bit và bắt đầu quá trình giải mã ở bước 111. Cụ thể là, bộ giải mã sử dụng phương tiện giải mã entropy để biến đổi dòng bit thành cú pháp và dữ liệu video tương ứng. Bộ giải mã sử dụng dữ liệu cú pháp từ dòng bit để xác định các phân vùng cho các khung ở bước 111. Hoạt động phân vùng nên so khớp các kết quả phân vùng khối ở bước 103. Kỹ thuật mã hóa/giải mã entropy như được sử dụng ở bước 111 đang được mô tả. Bộ mã hóa thực hiện nhiều lựa chọn trong quá trình nén, chẳng hạn lựa chọn các phương tiện phân vùng khối từ vài lựa chọn khả thi dựa trên định vị không gian các giá trị trong (các) ảnh đầu vào. Việc báo hiệu các lựa chọn chính xác có thể sử dụng số lượng bin lớn. Như được sử dụng ở đây, bin là giá trị nhị phân được xem như là biến (chẳng hạn, giá trị bit có thể thay đổi tùy thuộc ngữ cảnh). Tạo mã entropy cho phép bộ mã hóa loại bỏ bất kỳ tùy chọn nào mà rõ ràng không quan trọng đối với trường hợp cụ thể, bỏ lại tập hợp tùy chọn được phép. Sau đó, mỗi tùy chọn được phép được gán từ mã. Chiều dài của các từ mã dựa trên số lượng tùy chọn được phép (chẳng hạn, một bin cho hai tùy chọn, hai bin cho ba đến bốn tùy chọn, v.v.). Sau đó, bộ mã hóa mã hóa từ mã đối với tùy chọn được chọn. Phương thức này giảm kích thước của các từ mã do các từ mã lớn như mong muốn để chỉ báo duy nhất lựa chọn từ tập con nhỏ của các tùy chọn được phép ngược với chỉ báo duy nhất lựa chọn từ tập hợp có thể lớn của tất cả các tùy chọn khả thi. Sau đó, bộ giải mã giải mã lựa chọn bằng cách xác định tập hợp của các tùy chọn được phép theo cách tương tự với bộ mã hóa. Bằng cách xác định tập hợp của các tùy chọn được phép, bộ giải mã có thể đọc từ mã và xác định lựa chọn được thực hiện bởi bộ mã hóa.

Ở bước 113, bộ giải mã thực hiện giải mã khối. Cụ thể là, bộ giải mã sử dụng các phép biến đổi ngược để tạo các khối dư. Sau đó, bộ giải mã sử dụng các khối dư và các khối dự báo tương ứng để tái tạo các khối ảnh theo phân vùng này. Các khối dự báo có thể bao gồm cả hai khối dự báo trong lẫn các khối dự báo ngoài như được tạo ở bộ mã hóa ở bước 105. Sau đó, các khối ảnh được tái tạo được định vị vào các khung của tín hiệu video được tái tạo theo dữ liệu phân vùng được xác định ở bước 111. Cú pháp cho bước 113 cũng có thể được báo hiệu trong dòng bit qua tạo mã entropy như nêu trên.

Ở bước 115, thực hiện lọc trên các khung của tín hiệu video được tái tạo theo cách giống như bước 107 ở bộ mã hóa. Chẳng hạn, các bộ lọc chặn nhiễu, các DBF, các ALF, và các bộ lọc SAO có thể được áp dụng cho các khung để loại bỏ các giả tượng khối. Khi các khung được lọc, tín hiệu video có thể được xuất ra bộ hiển thị ở bước 117 để xem bởi người dùng cuối.

Fig.2 là sơ đồ của hệ thống tạo mã và giải mã (codec) 200 ví dụ để tạo mã video. Cụ thể là, hệ thống codec 200 cung cấp chức năng để hỗ trợ thực hiện phương pháp hoạt động 100. Hệ thống codec 200 được tổng quát để mô tả các thành phần được sử dụng trong cả bộ mã hóa và bộ giải mã. Hệ thống codec 200 nhận và phân vùng tín hiệu video như được nêu với các bước 101 và 103 trong phương pháp hoạt động 100, dẫn đến tín hiệu video được phân vùng 201. Sau đó, hệ thống codec 200 nén tín hiệu video được phân vùng 201 thành dòng bit được tạo mã khi hoạt động như là bộ mã hóa như được nêu dựa vào các bước 105, 107, và 109 trong phương pháp 100. Khi hoạt động như là bộ giải mã hệ thống codec 200 tạo tín hiệu video đầu ra từ dòng bit như được nêu dựa vào các bước 111, 113, 115, và 117 trong phương pháp hoạt động 100. Hệ thống codec 200 bao gồm thành phần điều khiển bộ tạo mã chung 211, thành phần lượng tử và chia tỷ lệ biến đổi 213, thành phần ước tính trong ảnh 215, thành phần dự báo trong ảnh 217, thành phần bù chuyển động 219, thành phần ước tính chuyển động 221, thành phần biến đổi ngược và chia tỷ lệ 229, thành phần phân tích điều khiển lọc 227, thành phần lọc trong vòng 225, thành phần DPB 223, và thành phần định dạng tiêu đề và tạo mã số học nhị phân thích ứng (context adaptive binary arithmetic coding, CABAC) 231. Các thành phần này được ghép

nổi như được thể hiện trên hình vẽ. Trên Fig.2, các đường màu đen chỉ báo di chuyển của dữ liệu cần được mã hóa/giải mã trong khi các đường nét đứt chỉ báo di chuyển của dữ liệu điều khiển mà điều khiển hoạt động của các thành phần khác. Các thành phần của hệ thống codec 200 có thể đều xuất hiện trong bộ mã hóa. Bộ giải mã có thể bao gồm tập con của các thành phần của hệ thống codec 200. Chẳng hạn, bộ giải mã có thể bao gồm thành phần dự báo trong ảnh 217, thành phần bù chuyển động 219, thành phần biến đổi ngược và chia tỷ lệ 229, thành phần lọc trong vòng 225, và thành phần DPB 223. Các thành phần này đang được mô tả.

Tín hiệu video được phân vùng 201 là chuỗi video được chụp đã được phân vùng thành các khối pixel bằng cây mã. Cây mã sử dụng các chế độ tách khác nhau để phân chia phụ khối pixel theo các khối pixel nhỏ hơn. Sau đó, các khối này có thể còn được phân chia phụ thành các khối nhỏ hơn. Các khối có thể được gọi là các nút trên cây mã. Các nút cha mẹ lớn hơn được tách thành các nút con nhỏ hơn. Số lần nút được phân chia phụ được gọi là chiều sâu của nút/cây mã. Các khối được phân chia có thể được bao gồm trong các CU trong một số trường hợp. Chẳng hạn, CU có thể là phần phụ của CTU chứa khối độ sáng, (các) khối sắc độ lệch đỏ (Cr), và (các) khối sắc độ lệch lam (Cb) cùng với các lệnh cú pháp tương ứng cho CU. Các nút tách có thể bao gồm cây nhị phân (binary tree, BT), cây tam phân (triple tree, TT), và cây tứ phân (quad tree, QT) được sử dụng để phân vùng nút thành hai, ba, hoặc bốn nút con, lần lượt, có các hình dạng thay đổi tùy thuộc vào các chế độ tách được sử dụng. Tín hiệu video được phân vùng 201 được chuyển tiếp đến thành phần điều khiển bộ tạo mã chung 211, thành phần lượng tử và chia tỷ lệ biến đổi 213, thành phần ước tính trong ảnh 215, thành phần phân tích điều khiển lọc 227, và thành phần ước tính chuyển động 221 để nén.

Thành phần điều khiển bộ tạo mã chung 211 được tạo cấu hình để ra quyết định liên quan đến tạo mã các ảnh của chuỗi video thành dòng bit theo các ràng buộc ứng dụng. Chẳng hạn, thành phần điều khiển bộ tạo mã chung 211 quản lý tối ưu hóa tốc độ bit/kích thước dòng bit với chất lượng tái tạo. Các quyết định này có thể được thực hiện dựa trên không gian lưu trữ/tính khả dụng băng thông

và các yêu cầu độ phân giải ảnh. Thành phần điều khiển bộ tạo mã chung 211 cũng quản lý sử dụng bộ đệm theo tốc độ truyền để giảm bớt các vấn đề chạy quá mức và chạy quá thấp của bộ đệm. Để quản lý các vấn đề này, thành phần điều khiển bộ tạo mã chung 211 quản lý phân vùng, dự báo và lọc bằng các thành phần khác. Chẳng hạn, thành phần điều khiển bộ tạo mã chung 211 có thể tăng động độ phức tạp nén để tăng độ phân giải và tăng sử dụng băng thông hoặc giảm độ phức tạp nén để giảm độ phân giải và sử dụng băng thông. Do vậy, thành phần điều khiển bộ tạo mã chung 211 điều khiển các thành phần khác của hệ thống codec 200 để cân bằng chất lượng tái tạo tín hiệu video với các quan ngại tốc độ bit. Thành phần điều khiển bộ tạo mã chung 211 tạo dữ liệu điều khiển, điều khiển hoạt động của các thành phần khác. Dữ liệu điều khiển cũng được chuyển tiếp đến thành phần định dạng tiêu đề và CABAC 231 cần được mã hóa trong dòng bit để báo hiệu các tham số để giải mã ở bộ giải mã.

Tín hiệu video được phân vùng 201 cũng được gửi đến thành phần ước tính chuyển động 221 và thành phần bù chuyển động 219 để dự báo ngoài. Khung hoặc lát của tín hiệu video được phân vùng 201 có thể được phân chia thành nhiều khối video. Thành phần ước tính chuyển động 221 và thành phần bù chuyển động 219 thực hiện tạo mã dự báo ngoài của khối video được nhận so với một hoặc nhiều khối trong một hoặc nhiều khung tham chiếu để dự báo thời gian. Hệ thống codec 200 có thể thực hiện nhiều bước tạo mã, chẳng hạn, để chọn chế độ tạo mã thích hợp cho mỗi khối của dữ liệu video.

Thành phần ước tính chuyển động 221 và thành phần bù chuyển động 219 có thể được tích hợp mạnh, nhưng được minh họa riêng rẽ theo các khái niệm. Ước tính chuyển động, được thực hiện bằng thành phần ước tính chuyển động 221, là quá trình tạo các vector chuyển động, mà ước tính chuyển động cho các khối video. Vector chuyển động, chẳng hạn, có thể chỉ báo dịch chuyển của đối tượng được tạo mã so với khối dự báo. Khối dự báo là khối được thấy so khớp gần nhất với khối cần được tạo mã, liên quan đến độ lệch pixel. Khối dự báo cũng có thể được gọi là khối tham chiếu. Độ lệch pixel này có thể được xác định bằng tổng hiệu số tuyệt đối (sum of absolute difference, SAD), tổng hiệu số bình phương (sum of square difference, SSD), hoặc các hệ đo hiệu số khác. HEVC sử dụng

vài đối tượng được tạo mã bao gồm CTU, các CTB, và các CU. Chẳng hạn, CTU có thể được phân chia thành các CTB, mà sau đó có thể được phân chia thành các CB để bao gồm trong các CU. CU có thể được mã hóa dưới dạng đơn vị dự báo (prediction unit, PU) chứa dữ liệu dự báo và/hoặc đơn vị biến đổi (transform unit, TU) chứa dữ liệu dư được biến đổi cho CU. Thành phần ước tính chuyển động 221 tạo các vector chuyển động, các PU, và các TU nhờ sử dụng phân tích méo dạng tốc độ làm một phần của quá trình tối ưu hóa méo dạng tốc độ. Chẳng hạn, thành phần ước tính chuyển động 221 có thể xác định nhiều khối tham chiếu, nhiều vector chuyển động, v.v. đối với khối/khung hiện tại, và có thể chọn các khối tham chiếu, các vector chuyển động, v.v. có các đặc tính méo dạng tốc độ tốt nhất. Các đặc tính méo dạng tốc độ tốt nhất cân bằng cả hai chất lượng tái tạo video (chẳng hạn, lượng tổn hao dữ liệu bằng cách nén) với hiệu suất tạo mã (chẳng hạn, kích thước mã hóa cuối cùng).

Trong một số ví dụ, hệ thống codec 200 có thể tính toán các giá trị đối với các vị trí pixel nguyên con của các ảnh tham chiếu được lưu trữ trong thành phần DPB 223. Chẳng hạn, hệ thống codec video 200 có thể nội suy các giá trị của các vị trí pixel $\frac{1}{4}$, các vị trí pixel $\frac{1}{8}$, hoặc các vị trí pixel phân số khác của ảnh tham chiếu. Do vậy, thành phần ước tính chuyển động 221 có thể thực hiện tìm kiếm chuyển động so với các vị trí pixel nguyên và các vị trí pixel phân số và xuất ra vector chuyển động có độ chính xác pixel phân số. Thành phần ước tính chuyển động 221 tính toán vector chuyển động cho PU của khối video trong lát được tạo mã ngoài bằng cách so sánh vị trí của PU với vị trí của khối dự báo của ảnh tham chiếu. Thành phần ước tính chuyển động 221 xuất ra vector chuyển động được tính toán dưới dạng dữ liệu chuyển động với thành phần định dạng tiêu đề và CABAC 231 để mã hóa và chuyển động đến thành phần bù chuyển động 219.

Bù chuyển động, được thực hiện bởi thành phần bù chuyển động 219, có thể bao gồm bước nạp hoặc tạo khối dự báo dựa trên vector chuyển động được xác định bởi thành phần ước tính chuyển động 221. Thành phần ước tính chuyển động 221 và thành phần bù chuyển động 219 lại có thể được tích hợp về mặt chức năng, trong một số ví dụ. Khi nhận vector chuyển động cho PU của khối

video hiện tại, thành phần bù chuyển động 219 có thể định vị khối dự báo mà vector chuyển động trỏ đến. Sau đó, khối video dư được tạo bằng cách lấy các giá trị pixel của khối video hiện tại được tạo mã trừ đi các giá trị pixel của khối dự báo, tạo thành các giá trị độ lệch pixel. Nói chung, thành phần ước tính chuyển động 221 thực hiện ước tính chuyển động so với các thành phần độ sáng, và thành phần bù chuyển động 219 sử dụng các vector chuyển động được tính toán dựa trên các thành phần độ sáng cho cả các thành phần sắc độ lẫn các thành phần độ sáng. Khối dự báo và khối dư được chuyển tiếp về thành phần lượng tử và chia tỷ lệ biến đổi 213.

Tín hiệu video được phân vùng 201 cũng được gửi đến thành phần ước tính trong ảnh 215 và thành phần dự báo trong ảnh 217. Như với thành phần ước tính chuyển động 221 và thành phần bù chuyển động 219, thành phần ước tính trong ảnh 215 và thành phần dự báo trong ảnh 217 có thể được tích hợp chặt chẽ, nhưng được minh họa riêng rẽ cho các khái niệm. Thành phần ước tính trong ảnh 215 và thành phần dự báo trong ảnh 217 dự báo trong khối hiện tại so với các khối trong khung hiện tại, như là tùy chọn với dự báo ngoài được thực hiện bởi thành phần ước tính chuyển động 221 và thành phần bù chuyển động 219 giữa các khung, như nêu trên. Cụ thể là, thành phần ước tính trong ảnh 215 xác định chế độ dự báo trong sử dụng để mã hóa khối hiện tại. Trong một số ví dụ, thành phần ước tính trong ảnh 215 chọn chế độ dự báo trong thích hợp để mã hóa khối hiện tại từ nhiều chế độ dự báo trong được kiểm tra. Sau đó, các chế độ dự báo trong được chọn được chuyển tiếp về thành phần định dạng tiêu đề và CABAC 231 để mã hóa.

Chẳng hạn, thành phần ước tính trong ảnh 215 tính toán các giá trị méo dạng tốc độ nhờ sử dụng phân tích méo dạng tốc độ đối với các chế độ dự báo trong được kiểm tra khác nhau, và lựa chọn chế độ dự báo trong có các đặc tính méo dạng tốc độ tốt nhất trong số các chế độ được kiểm tra. Phân tích méo dạng tốc độ thường xác định lượng méo dạng (hoặc sai số) giữa khối được mã hóa và khối không được mã hóa ban đầu mà được mã hóa để tạo khối được mã hóa, cũng như tốc độ bit (chẳng hạn, số lượng bit) được sử dụng để tạo khối được mã hóa. Thành phần ước tính trong ảnh 215 tính toán các tỷ lệ từ các méo dạng và các tốc

độ cho các khối được mã hóa khác nhau để xác định chế độ dự báo trong nào biểu diễn giá trị méo dạng tốc độ tốt nhất cho khối. Ngoài ra, thành phần ước tính trong ảnh 215 có thể được tạo cấu hình để tạo mã các khối chiều sâu của bản đồ chiều sâu nhờ sử dụng chế độ tạo mô hình chiều sâu (depth modeling mode, DMM) dựa trên tối ưu hóa méo dạng tốc độ (rate-distortion optimization, RDO).

Thành phần dự báo trong ảnh 217 có thể tạo khối dư từ khối dự báo dựa trên các chế độ dự báo trong được chọn được xác định bằng thành phần ước tính trong ảnh 215 khi được triển khai trên bộ mã hóa hoặc đọc khối dư từ dòng bit khi được triển khai trên bộ giải mã. Khối dư bao gồm hiệu số ở các giá trị giữa khối dự báo và khối ban đầu, được biểu diễn dưới dạng ma trận. Sau đó, khối dư được chuyển tiếp đến thành phần lượng tử và chia tỷ lệ biến đổi 213. Thành phần ước tính trong ảnh 215 và thành phần dự báo trong ảnh 217 có thể hoạt động trên cả các thành phần độ sáng lẫn sắc độ.

Thành phần lượng tử và chia tỷ lệ biến đổi 213 được tạo cấu hình để nén tiếp khối dư. Thành phần lượng tử và chia tỷ lệ biến đổi 213 áp dụng phép biến đổi, chẳng hạn, biến đổi côsin rời rạc (discrete cosine transform, DCT), biến đổi sin rời rạc (discrete sine transform, DST), hoặc phép biến đổi tương tự về khái niệm, với khối dư, tạo khối video bao gồm các giá trị hệ số biến đổi dư. Các phép biến đổi wavelet, các phép biến đổi số nguyên, các phép biến đổi băng thành phần hoặc các loại biến đổi khác cũng có thể được sử dụng. Phép biến đổi có thể biến đổi thông tin dư từ miền giá trị pixel sang miền biến đổi, chẳng hạn miền tần số. Thành phần lượng tử và chia tỷ lệ biến đổi 213 cũng được tạo cấu hình để chia tỷ lệ thông tin dư được biến đổi, chẳng hạn dựa trên tần số. Phép chia tỷ lệ này bao gồm áp dụng hệ số tỷ lệ cho thông tin dư sao cho thông tin tần số khác nhau được lượng tử ở các độ hạt khác nhau, mà có thể ảnh hưởng chất lượng hình ảnh cuối cùng của video được tái tạo. Thành phần lượng tử và chia tỷ lệ biến đổi 213 cũng được tạo cấu hình để lượng tử các hệ số biến đổi để giảm tiếp tốc độ bit. Quá trình lượng tử có thể giảm chiều sâu bit được liên kết với một số hoặc tất cả các hệ số. Mức độ lượng tử có thể được chỉnh sửa bằng cách điều chỉnh tham số Mức độ lượng tử có thể được chỉnh sửa bằng cách điều chỉnh tham số lượng tử (quantization parameter, QP). Sau đó, trong một số ví dụ, thành phần lượng tử và

chia tỷ lệ biến đổi 213 có thể thực hiện phép quét ma trận bao gồm các hệ số biến đổi được lượng tử. Các hệ số biến đổi được lượng tử được chuyển tiếp đến thành phần định dạng tiêu đề và CABAC 231 cần được mã hóa trong dòng bit.

Thành phần biến đổi ngược và chia tỷ lệ 229 áp dụng hoạt động ngược lại của thành phần lượng tử và chia tỷ lệ biến đổi 213 để hỗ trợ ước tính chuyển động. Thành phần biến đổi ngược và chia tỷ lệ 229 áp dụng phép chia tỷ lệ ngược, biến đổi, và/hoặc lượng tử để tái tạo khối dư trong miền pixel, chẳng hạn, để sử dụng sau này làm khối tham chiếu mà có thể trở thành khối dự báo cho khối hiện tại khác. Thành phần ước tính chuyển động 221 và/hoặc thành phần bù chuyển động 219 có thể tính toán khối tham chiếu bằng cách thêm khối dư lại vào khối dự báo tương ứng để sử dụng khi ước tính chuyển động của khối/khung sau này. Các bộ lọc được áp dụng cho các khối tham chiếu được tái tạo để di trú các giả tượng được tạo trong khi chia tỷ lệ, lượng tử, và phép biến đổi. Ngược lại, các giả tượng này có thể gây ra dự báo không chính xác (và tạo các giả tượng bổ sung) khi các khối tiếp theo được dự báo.

Thành phần phân tích điều khiển lọc 227 và thành phần lọc trong vòng 225 áp dụng các bộ lọc cho các khối dư và/hoặc cho các khối ảnh được tái tạo. Chẳng hạn, khối dư được biến đổi từ thành phần biến đổi ngược và chia tỷ lệ 229 có thể được kết hợp với khối dự báo tương ứng từ thành phần dự báo trong ảnh 217 và/hoặc thành phần bù chuyển động 219 để tái tạo khối ảnh ban đầu. Sau đó, các bộ lọc có thể được áp dụng cho khối ảnh được tái tạo. Trong một số ví dụ, thay vào đó, các bộ lọc có thể được áp dụng cho các khối dư. Như với các thành phần khác trên Fig.2, thành phần phân tích điều khiển lọc 227 và thành phần lọc trong vòng 225 được tích hợp chặt chẽ và có thể được thực hiện với nhau, nhưng được mô tả riêng rẽ cho các khái niệm. Các bộ lọc được áp dụng cho các khối tham chiếu được tái tạo được áp dụng cho các vùng không gian cụ thể và bao gồm nhiều tham số để điều chỉnh cách thức các bộ lọc này được áp dụng. Thành phần phân tích điều khiển lọc 227 phân tích các khối tham chiếu được tái tạo để xác định trong đó các bộ lọc này cần được áp dụng và thiết lập các tham số tương ứng. Dữ liệu này được chuyển tiếp đến thành phần định dạng tiêu đề và CABAC 231 dưới dạng dữ liệu điều khiển lọc để mã hóa. Thành phần lọc trong vòng 225

áp dụng các bộ lọc này dựa trên dữ liệu điều khiển lọc. Các bộ lọc có thể bao gồm DBF, bộ lọc chặn nhiễu, bộ lọc SAO, và ALF. Các bộ lọc này có thể được áp dụng trong miền không gian/pixel (chẳng hạn, trên khối pixel được tái tạo) hoặc trong miền tần số, tùy thuộc vào ví dụ.

Khi hoạt động như là bộ mã hóa, khối ảnh được tái tạo được lọc, khối dư, và/hoặc khối dự báo được lưu trữ trong thành phần DPB 223 để sử dụng sau này khi ước tính chuyển động như nêu trên. Khi hoạt động như là bộ giải mã, thành phần DPB 223 lưu trữ và chuyển tiếp các khối được tái tạo và được lọc về phía bộ phận hiển thị như là một phần của tín hiệu video đầu ra. Thành phần DPB 223 có thể là bộ nhớ bất kỳ có thể lưu trữ các khối dự báo, các khối dư, và/hoặc các khối ảnh được tái tạo.

Thành phần định dạng tiêu đề và CABAC 231 nhận dữ liệu từ các thành phần khác nhau của hệ thống codec 200 và mã hóa dữ liệu này vào dòng bit được tạo mã để truyền về phía bộ giải mã. Cụ thể là, thành phần định dạng tiêu đề và CABAC 231 tạo các tiêu đề khác nhau để mã hóa dữ liệu điều khiển, chẳng hạn dữ liệu điều khiển tổng quát và dữ liệu điều khiển lọc. Ngoài ra, dữ liệu dự báo, bao gồm dữ liệu chuyển độ và dự báo trong, cũng như dữ liệu dư ở dạng dữ liệu hệ số biến đổi được lượng tử đều được mã hóa trong dòng bit. Dòng bit cuối cùng bao gồm tất cả thông tin mong muốn bởi bộ giải mã để tái tạo tín hiệu video được phân vùng ban đầu 201. Thông tin này cũng có thể bao gồm các bảng chỉ số chế độ dự báo trong (cũng được gọi là các bảng ánh xạ từ mã), các định nghĩa về các ngữ cảnh mã hóa cho các khối khác nhau, các chỉ báo về các chế độ dự báo trong có xác suất lớn nhất, chỉ báo thông tin phân vùng, v.v.. Dữ liệu này có thể được mã hóa bằng cách sử dụng tạo mã entropy. Chẳng hạn, thông tin này có thể được mã hóa bằng cách sử dụng mã hóa có chiều dài biến thiên thích ứng ngữ cảnh (context adaptive variable length coding, CAVLC), CABAC, mã hóa số học nhị phân ngữ cảnh thích ứng dựa trên cú pháp (syntax-based context-adaptive binary arithmetic coding, SBAC), mã hóa entropy phân vùng khoảng xác suất (probability interval partitioning entropy, PIPE), hoặc kỹ thuật mã hóa entropy khác. Tiếp sau mã hóa entropy, dòng bit được tạo mã có thể được truyền

đến thiết bị khác (chẳng hạn, bộ giải mã video) hoặc được lưu trữ để truyền hoặc truy xuất sau này.

Fig.3 là sơ đồ khối minh họa bộ mã hóa video 300 ví dụ. Bộ mã hóa video 300 có thể được sử dụng để thực hiện các chức năng mã hóa của hệ thống codec 200 và/hoặc thực hiện các bước 101, 103, 105, 107, và/hoặc 109 của phương pháp hoạt động 100. Bộ mã hóa 300 phân vùng tín hiệu video đầu vào, dẫn đến tín hiệu video được phân vùng 301, gần như tương tự với tín hiệu video được phân vùng 201. Sau đó, tín hiệu video được phân vùng 301 được nén và được mã hóa thành dòng bit bởi các thành phần của bộ mã hóa 300.

Cụ thể là, tín hiệu video được phân vùng 301 được chuyển tiếp đến thành phần dự báo trong ảnh 317 để dự báo trong. Thành phần dự báo trong ảnh 317 có thể gần như tương tự với thành phần ước tính trong ảnh 215 và thành phần dự báo trong ảnh 217. Tín hiệu video được phân vùng 301 cũng được chuyển tiếp đến thành phần bù chuyển động 321 để dự báo ngoài dựa trên các khối tham chiếu trong thành phần DPB 323. Thành phần bù chuyển động 321 có thể gần như tương tự với thành phần ước tính chuyển động 221 và thành phần bù chuyển động 219. Các khối dự báo và các khối dư từ thành phần dự báo trong ảnh 317 và thành phần bù chuyển động 321 được chuyển tiếp đến thành phần biến đổi và lượng tử 313 để biến đổi và lượng tử các khối dư. Thành phần biến đổi và lượng tử 313 có thể gần như tương tự với thành phần lượng tử và chia tỷ lệ biến đổi 213. Các khối dư được biến đổi và được lượng tử và các khối dự báo tương ứng (cùng với dữ liệu điều khiển liên kết) được chuyển tiếp đến thành phần tạo mã entropy 331 để tạo mã vào dòng bit. Thành phần tạo mã entropy 331 có thể gần như tương tự với thành phần định dạng tiêu đề và CABAC 231.

Các khối dư được biến đổi và được lượng tử và/hoặc các khối dự báo tương ứng cũng được chuyển tiếp từ thành phần biến đổi và lượng tử 313 đến thành phần biến đổi ngược và lượng tử 329 để tái tạo thành các khối tham chiếu để sử dụng bởi thành phần bù chuyển động 321. Thành phần biến đổi ngược và lượng tử 329 có thể gần như tương tự với thành phần biến đổi ngược và chia tỷ lệ 229. Các bộ lọc trong vòng trong thành phần lọc trong vòng 325 cũng được áp dụng cho các khối dư và/hoặc các khối tham chiếu được tái tạo, phụ thuộc vào ví dụ.

Thành phần lọc trong vòng 325 có thể gần như tương tự với thành phần phân tích điều khiển lọc 227 và thành phần lọc trong vòng 225. Thành phần lọc trong vòng 325 có thể bao gồm nhiều bộ lọc như được đề cập dựa vào thành phần lọc trong vòng 225. Sau đó, các khối được lọc được lưu trữ trong thành phần DPB 323 để sử dụng làm các khối tham chiếu bằng thành phần bù chuyển động 321. Thành phần DPB 323 có thể gần như tương tự với thành phần DPB 223.

Fig.4 là sơ đồ khối minh họa bộ giải mã video 400 ví dụ. Bộ giải mã video 400 có thể được sử dụng để thực hiện các chức năng giải mã của hệ thống codec 200 và/hoặc thực hiện các bước 111, 113, 115, và/hoặc 117 của phương pháp hoạt động 100. Bộ giải mã 400 nhận dòng bit, chẳng hạn từ bộ mã hóa 300, và tạo tín hiệu video đầu ra được tái tạo dựa trên dòng bit để hiển thị đến người dùng cuối.

Dòng bit được nhận bởi thành phần giải mã entropy 433. Thành phần giải mã entropy 433 được tạo cấu hình để thực hiện phương tiện giải mã entropy, chẳng hạn CAVLC, CABAC, SBAC, tạo mã PIPE, hoặc các kỹ thuật tạo mã entropy khác. Chẳng hạn, thành phần giải mã entropy 433 có thể sử dụng thông tin tiêu đề để tạo ngữ cảnh để hiểu dữ liệu bổ sung được mã hóa dưới dạng các từ mã trong dòng bit. Thông tin được giải mã bao gồm thông tin mong muốn bất kỳ để giải mã tín hiệu video, chẳng hạn dữ liệu điều khiển chung, dữ liệu điều khiển lọc, thông tin phân vùng, dữ liệu chuyển động, dữ liệu dự báo, và các hệ số biến đổi được lượng tử từ các khối dư. Các hệ số biến đổi được lượng tử được chuyển tiếp đến thành phần biến đổi ngược và lượng tử 429 để tái tạo thành các khối dư. Thành phần biến đổi ngược và lượng tử 429 có thể giống như thành phần biến đổi ngược và lượng tử 329.

Các khối dư được tái tạo và/hoặc các khối dự báo được chuyển tiếp đến thành phần dự báo trong ảnh 417 để tái tạo thành các khối ảnh dựa trên các hoạt động dự báo trong. Thành phần dự báo trong ảnh 417 có thể giống như thành phần ước tính trong ảnh 215 và thành phần dự báo trong ảnh 217. Cụ thể là, thành phần dự báo trong ảnh 417 sử dụng các chế độ dự báo để định vị khối tham chiếu trong khung và áp dụng khối dư cho kết quả để tái tạo các khối ảnh được dự báo trong. Các khối ảnh được dự báo trong được tái tạo và/hoặc các khối dư

và dữ liệu dự báo ngoài tương ứng được chuyển tiếp đến thành phần DPB 423 qua thành phần lọc trong vòng 425, có thể lần lượt gần như tương tự với thành phần DPB 223 và thành phần lọc trong vòng 225. Thành phần lọc trong vòng 425 lọc các khối ảnh được tái tạo, các khối dư và/hoặc các khối dự báo, và thông tin này được lưu trữ trong thành phần DPB 423. Các khối ảnh được tái tạo từ thành phần DPB 423 được chuyển tiếp đến thành phần bù chuyển động 421 để dự báo ngoài. Thành phần bù chuyển động 421 có thể gần như tương tự với thành phần ước tính chuyển động 221 và/hoặc thành phần bù chuyển động 219. Cụ thể là, thành phần bù chuyển động 421 sử dụng các vector chuyển động từ khối tham chiếu để tạo khối dự báo và áp dụng khối dư cho kết quả để tái tạo khối ảnh. Các khối được tái tạo thu được cũng có thể được chuyển tiếp qua thành phần lọc trong vòng 425 đến thành phần DPB 423. Thành phần DPB 423 tiếp tục lưu trữ các khối ảnh được tái tạo bổ sung, có thể được tái tạo thành các khung qua thông tin phân vùng. Các khung này cũng có thể được đặt vào chuỗi. Chuỗi này được xuất về phía bộ hiển thị dưới dạng tín hiệu video đầu ra được tái tạo.

Fig.5 là sơ đồ minh họa dòng bit ví dụ 500 chứa chuỗi video được mã hóa. Chẳng hạn, dòng bit 500 có thể được tạo bởi hệ thống codec 200 và/hoặc bộ mã hóa 300 để giải mã bởi hệ thống codec 200 và/hoặc bộ giải mã 400. Lấy ví dụ khác, dòng bit 500 có thể được tạo bởi bộ mã hóa ở bước 109 của phương pháp 100 để sử dụng bởi bộ giải mã ở bước 111.

Dòng bit 500 bao gồm SPS 510, các PPS 512, các tiêu đề nhóm ô 514, và dữ liệu ảnh 520. SPS 510 chứa dữ liệu chuỗi chung cho tất cả các ảnh trong chuỗi video được chứa trong dòng bit 500. Dữ liệu này có thể bao gồm kích thước ảnh, chiều sâu bit, các tham số công cụ mã, các giới hạn tốc độ bit, v.v.. Theo phương án thực hiện, SPS 510 bao gồm cờ SPS 565. Theo phương án thực hiện, cờ SPS 565 có giá trị thứ nhất (chẳng hạn, 1) hoặc giá trị thứ hai (chẳng hạn, 0). Cờ SPS 565 có giá trị thứ nhất xác định rằng ánh xạ của ID ảnh con 575 được báo hiệu trong SPS 510, và cờ SPS 565 có giá trị thứ hai xác định rằng ánh xạ của ID ảnh con 575 được báo hiệu trong PPS 512.

PPS 512 chứa các tham số riêng cho một hoặc nhiều ảnh tương ứng. Do vậy, mỗi ảnh trong chuỗi video có thể đề cập đến một PPS 512. PPS 512 có thể chỉ

báo các công cụ mã khả dụng cho các ô trong các ảnh tương ứng, các QP, các độ lệch, các tham số công cụ mã riêng cho ảnh (chẳng hạn, các điều khiển lọc), v.v.. Theo phương án thực hiện, PPS 512 bao gồm cờ PPS 567. Theo phương án thực hiện, cờ PPS 567 có giá trị thứ nhất (chẳng hạn, 1) hoặc giá trị thứ hai (chẳng hạn, 0). Cờ PPS 567 có giá trị thứ nhất xác định rằng ánh xạ của ID ảnh con 575 được báo hiệu trong PPS 512, và cờ PPS 567 có giá trị thứ hai xác định rằng ánh xạ của ID ảnh con 575 không được báo hiệu trong PPS 512.

Tiêu đề nhóm ô 514 chứa các tham số riêng cho mỗi nhóm ô trong ảnh. Do vậy, có thể có một tiêu đề nhóm ô 514 trên nhóm ô trong chuỗi video. Tiêu đề nhóm ô 514 có thể chứa thông tin nhóm ô, các POC, các danh sách ảnh tham chiếu, các trọng số dự báo, các điểm đầu vào ô, các tham số tách khối, v.v.. Cần lưu ý rằng một số hệ thống đề cập đến tiêu đề nhóm ô 514 dưới dạng tiêu đề lát, và sử dụng thông tin này để hỗ trợ các lát thay vì các nhóm ô.

Dữ liệu ảnh 520 chứa dữ liệu video được mã hóa theo dự báo ngoài và/hoặc dự báo trong cũng như dữ liệu dư được biến đổi và lượng tử tương ứng. Dữ liệu ảnh 520 này được sắp xếp theo phân vùng được sử dụng để phân vùng ảnh trước khi mã hóa. Chẳng hạn, ảnh trong dữ liệu ảnh 520 bao gồm một hoặc nhiều ảnh 521. Chuỗi ảnh 527 có thể được gọi là CVS 527. Như được sử dụng ở đây, CVS 527 là chuỗi AU bao gồm, theo thứ tự giải mã, AU CVSS, tiếp theo bởi 0 hoặc nhiều AU không phải là các AU CVSS, bao gồm tất cả các AU tiếp theo mà không bao gồm bất kỳ AU tiếp theo vốn là AU CVSS. AU CVSS là AU trong đó có PU cho mỗi lớp được xác định bởi VPS và ảnh được tạo mã trong mỗi PU là ảnh CLVSS. Theo phương án thực hiện, mỗi ảnh 521 trong AU. PU là tập hợp đơn vị NAL được liên kết với nhau theo quy tắc phân loại cụ thể, là liên tục trong thứ tự giải mã, và chứa chính xác một ảnh được mã hóa.

CVS là CVS cho mỗi chuỗi vi đeo được tạo mã theo lớp (coded layer video sequence, CLVS) trong dòng bit video. Đáng lưu ý, CVS và CLVS là giống nhau khi dòng bit video bao gồm một lớp. CVS và CLVS chỉ khác nhau khi dòng bit video bao gồm nhiều lớp.

Mỗi ảnh trong các ảnh 521 có thể được phân chia thành các ô 523. Ô 523 là phần được phân vùng của ảnh được tạo bởi các đường biên ngang và dọc. Các ô

523 có thể là hình chữ nhật và/hoặc hình vuông. Cụ thể là, ô 523 bao gồm bốn cạnh được nối ở các góc vuông. Bốn cạnh bao gồm hai cặp cạnh song song. Ngoài ra, các cạnh theo cặp cạnh song song có chiều dài bằng nhau. Thực tế, ô 523 có thể có hình chữ nhật, trong đó hình vuông là trường hợp đặc biệt của hình chữ nhật trong đó bốn cạnh có chiều dài bằng nhau. Ảnh có thể chứa một hoặc nhiều ô 523.

Ảnh (chẳng hạn, ảnh 521) có thể được phân vùng thành các hàng và các cột của các ô 523. Hàng ô là tập hợp của các ô 523 được định vị theo cách liền kề ngang để tạo đường liên tục từ đường biên trái sang đường biên phải của ảnh (hoặc ngược lại). Cột ô là tập hợp ô 523 được định vị theo cách liền kề dọc để tạo đường liên tục từ đường biên trên đến đường biên dưới của ảnh (hoặc ngược lại).

Các ô 523 có thể hoặc có thể không cho phép dự báo dựa trên các ô 523 khác, tùy thuộc vào ví dụ. Mỗi ô 523 có thể có chỉ số ô duy nhất trong ảnh. Chỉ số ô là ID số được chọn theo thủ tục có thể được sử dụng để phân biệt một ô 523 với ô khác. Chẳng hạn, các chỉ số ô có thể tăng theo số theo thứ tự quét mảnh. Thứ tự quét mảnh là từ trái sang phải và từ trên xuống dưới.

Cần lưu ý rằng, trong một số ví dụ, các ô 523 cũng có thể được gán các ID ô. ID ô là ID được gán có thể được sử dụng để phân biệt một ô 523 với ô khác. Các tính toán có thể sử dụng các ID ô thay cho các chỉ số ô trong một số ví dụ. Ngoài ra, các ID ô có thể được gán để có các giá trị giống như các chỉ số ô trong một số ví dụ. Các chỉ số ô và/hoặc các ID có thể được báo hiệu để chỉ báo các nhóm ô chứa các ô 523. Chẳng hạn, các chỉ số ô và/hoặc các ID có thể được sử dụng để ánh xạ dữ liệu ảnh được liên kết với ô 523 đến vị trí thích hợp để hiển thị. Nhóm ô là tập hợp ô liên quan 523 có thể được trích xuất và tạo mã riêng rẽ, chẳng hạn để hỗ trợ hiển thị vùng liên quan và/hoặc để hỗ trợ xử lý song song. Các ô 523 trong nhóm ô có thể được tạo mã không dựa vào các ô 523 nằm ngoài nhóm ô. Mỗi ô 523 có thể được gán cho nhóm ô tương ứng, và do vậy ảnh có thể chứa các nhóm ô.

Các ô 523 còn được phân chia thành các CTU. Các CTU còn được phân chia thành các CB dựa trên các cây mã. Sau đó, các CB có thể được mã hóa/giải mã theo các cơ chế dự báo.

Fig.6A đến Fig.6E là các hình vẽ minh họa cơ chế 600 lấy làm ví dụ để tạo rãnh trích xuất 610 (cũng được biết là dòng bit được hợp nhất) để kết hợp các ảnh con có nhiều độ phân giải từ các dòng bit khác nhau thành một ảnh để sử dụng trong các ứng dụng thực tế ảo (virtual reality, VR). Cơ chế 600 có thể được sử dụng để hỗ trợ trường hợp sử dụng ví dụ của phương pháp 100. Chẳng hạn, cơ chế 600 có thể được sử dụng để tạo dòng bit 500 để truyền từ hệ thống codec 200 và/hoặc bộ mã hóa 300 về phía hệ thống codec 200 và/hoặc bộ giải mã 400. Lấy ví dụ cụ thể, cơ chế 600 có thể được sử dụng cùng với VR, định dạng phương tiện đa hướng (Ominidirectional Media Format, OMAF), video 360°, v.v..

Trong VR, chỉ một phần video được hiển thị đến người dùng. Chẳng hạn, video VR có thể được tạo phim để bao gồm hình cầu bao quanh người dùng. Người dùng có thể sử dụng phần hiển thị đội đầu (head mounted display, HMD) để xem video VR. Người dùng có thể trở HMD về phía vùng liên quan. Vùng liên quan được hiển thị đến người dùng và dữ liệu video khác bị bỏ qua. Theo cách này, người dùng xem chỉ phần được người dùng chọn của video VR ở thời điểm bất kỳ. Cách này bắt chước các nhận thức của người dùng, và do vậy khiến người dùng trải nghiệm môi trường ảo theo cách bắt chước môi trường thực. Một trong các vấn đề với cách tiếp cận này chính là toàn bộ video VR có thể được truyền đến người dùng, nhưng chỉ cổng nhìn hiện tại của video thực sự được sử dụng và phần còn lại bị bỏ qua. Để tăng hiệu suất báo hiệu cho các ứng dụng streaming, cổng nhìn hiện tại của người dùng có thể được truyền ở độ phân giải thứ nhất cao hơn và các cổng nhìn khác có thể được truyền ở độ phân giải thứ hai thấp hơn. Theo cách này, các cổng nhìn có thể bị loại bỏ chiếm băng thông nhỏ hơn (các) cổng nhìn có thể được xem bởi người dùng. Nếu người dùng chọn cổng nhìn mới, nội dung có độ phân giải thấp hơn có thể được thể hiện cho đến khi bộ giải mã có thể yêu cầu rằng cổng nhìn hiện tại khác được truyền ở độ phân giải thứ nhất cao hơn. Cơ chế 600 có thể được sử dụng để tạo

rãnh trích xuất 610, như được thể hiện trên Fig.6E, để hỗ trợ chức năng này. Rãnh trích xuất 610 là rãnh dữ liệu ảnh đóng gói ảnh (chẳng hạn, ảnh 521) ở nhiều độ phân giải để sử dụng như nêu trên.

Cơ chế 600 mã hóa cùng nội dung video ở độ phân giải thứ nhất 611 và độ phân giải thứ hai 612, lần lượt được thể hiện trên Fig.6A và Fig.6B. Lấy ví dụ cụ thể, độ phân giải thứ nhất 611 có thể là các mẫu độ sáng 5120×2560 và độ phân giải thứ hai 612 có thể là các mẫu độ sáng 2560×1280 . Các ảnh của video có thể lần lượt được phân vùng thành các ô 601 ở độ phân giải thứ nhất 611 và các ô 603 ở độ phân giải thứ hai 612. Như được sử dụng ở đây, các ô có thể được gọi là các ảnh con. Trong ví dụ được thể hiện, mỗi ô trong các ô 601 và 603 được phân vùng thành lưới 4×2 . Ngoài ra, MCTS có thể được tạo mã cho vị trí mỗi ô 601 và 603. Các ảnh ở độ phân giải thứ nhất 611 và độ phân giải thứ hai 612 mà mỗi ảnh dẫn đến chuỗi MCTS mô tả video theo thời gian ở độ phân giải tương ứng. Mỗi chuỗi MCTS được tạo mã được lưu trữ dưới dạng rãnh ảnh con hoặc rãnh ô. Sau đó, cơ chế 600 có thể sử dụng các ảnh để tạo các phân đoạn để hỗ trợ lựa chọn MCTS công nhìn thích ứng. Chẳng hạn, mỗi khoảng hướng nhìn tạo lựa chọn khác của các MCTS độ phân giải cao và độ phân giải thấp được xem xét. Trong ví dụ được minh họa, bốn ô 601 chứa các MCTS ở độ phân giải thứ nhất 611 và bốn ô 603 chứa các MCTS ở độ phân giải thứ hai 612 thu được.

Sau đó, cơ chế 600 có thể tạo rãnh trích xuất 610 cho mỗi lựa chọn MCTS công nhìn thích ứng khả thi. Fig.6C và Fig.6D minh họa lựa chọn MCTS công nhìn thích ứng ví dụ. Cụ thể là, tập hợp ô được chọn 605 và 607 lần lượt được chọn ở độ phân giải thứ nhất 611 và độ phân giải thứ hai 612. Các ô được chọn 605 và 607 được minh họa trong tô xám. Trong ví dụ được thể hiện, các ô 605 được chọn là các ô 601 ở độ phân giải thứ nhất 611 được thể hiện với người dùng và các ô được chọn 607 là các ô 603 ở độ phân giải thứ hai 612 có thể bị loại bỏ nhưng được duy trì để hỗ trợ hiển thị nếu người dùng chọn công nhìn mới. Sau đó, các ô được chọn 605 và 607 được kết hợp thành một ảnh chứa dữ liệu ảnh ở cả độ phân giải thứ nhất 611 lẫn độ phân giải thứ hai 612. Các ảnh này được kết hợp để tạo rãnh trích xuất 610. Fig.6E minh họa một ảnh từ rãnh trích xuất 610 tương ứng để minh họa. Như được thể hiện trên hình vẽ, ảnh trong rãnh

trích xuất 610 chứa các ô được chọn 605 và 607 từ độ phân giải thứ nhất 611 và độ phân giải thứ hai 612. Như lưu ý trên đây, các hình vẽ Fig.6C đến Fig.6E minh họa một lựa chọn MCTS công nhìn thích ứng. Để cho phép người dùng lựa chọn công nhìn bất kỳ, rãnh trích xuất 610 cần được tạo cho mỗi tổ hợp khả thi của các ô được chọn 605 và 607.

Trong ví dụ được thể hiện, mỗi lựa chọn ô 603 đóng gói nội dung từ dòng bit độ phân giải thứ hai 612 chứa hai lát. RegionWisePackingBox có thể được bao gồm trong rãnh trích xuất 610 để tạo ánh xạ giữa ảnh được đóng gói và ảnh được chiếu theo định dạng chiếu gần như hình chữ nhật (equirectangular projection, ERP). Trong ví dụ hiện tại, các dòng bit thu được từ các rãnh trích xuất 610 có độ phân giải 3200×2560 . Kết quả là, bộ giải mã có thể có bốn nghìn mẫu (4K) có thể giải mã nội dung trong đó công nhìn được trích xuất từ dòng bit được tạo mã có độ phân giải năm nghìn mẫu 5K (5120×2560).

Như được thể hiện trên Fig.6C, các ô được chọn 605 (được thể hiện là màu xám) ở độ phân giải thứ nhất 611 có các ID ô sau: 10, 11, 14 và 15. Như được sử dụng ở đây, các ID ô cũng có thể được gọi là các ID ảnh con. Các ô được chọn 607 ở độ phân giải thứ hai 612 có các ID sau: 1, 5, 4, và 8. Do vậy, rãnh trích xuất 610 chứa các ID ô sau: 10, 11, 14, 15, 1, 5, 4, và 8. Các ID ô được sử dụng để nhận diện các ảnh con cụ thể nhờ sử dụng chỉ số ảnh con, có thể được gọi ở đây là ánh xạ của ID ảnh con.

Fig.7A đến Fig.7E minh họa cơ chế 700 ví dụ để tạo rãnh trích xuất 710 để kết hợp các ảnh con có nhiều độ phân giải từ các dòng bit khác nhau thành một ảnh để sử dụng trong các ứng dụng VR khi người dùng đã thay đổi các công nhìn so với công nhìn được chọn cho các hình vẽ từ Fig.6A đến Fig.6E. Tức là, Fig.7A đến Fig.7E là các hình vẽ minh họa cách thức rãnh trích xuất 710 mới được tạo khi có thay đổi trong hướng nhìn trong CVS, bao gồm rãnh trích xuất 610 và rãnh trích xuất 710.

Như được thể hiện trên Fig.7A đến Fig.7B, các ảnh của video lần lượt đã được phân vùng thành các ô 701 ở độ phân giải thứ nhất 711 và các ô 703 ở độ phân giải thứ hai 712. Tuy nhiên, đã có thay đổi trong hướng nhìn trong cơ chế 700 so với cơ chế 600. Do vậy, như được thể hiện trên Fig.7C và Fig.7D, các ô

được chọn 705 (được thể hiện là màu xám) ở độ phân giải thứ nhất 711 đang có các ID ô sau: 9, 10, 13, và 14, và các ô được chọn 707 ở độ phân giải thứ hai 712 hiện có các ID sau: 3, 4, 7, và 8. Do vậy, rãnh trích xuất 710 chứa các ID ô sau: 3, 4, 7, 8, 9, 10, 13, và 14 do thay đổi trong hướng nhìn.

Khi CVS được truyền bao gồm tập hợp ảnh con không đổi, sau đó các ID ảnh con liên quan bị bỏ lại trong SPS (tất cả các ID khác bị loại bỏ). Khi xuất hiện hợp nhất (chẳng hạn, để tạo một trong các rãnh trích xuất 610 hoặc 710), các ID ảnh con được di chuyển đến PPS. Trong trường hợp khác, cờ được thiết lập trong dòng bit để chỉ báo trong đó các ID ảnh con đang được đặt.

Thông thường, khi xuất hiện thay đổi trong hướng nhìn, các ảnh IRAP mới phải được gửi. Ảnh IRAP là ảnh được tạo mã mà tất cả các đơn VCL NAL có cùng giá trị của loại đơn vị NAL. Ảnh IRAP cung cấp hai chức năng/lợi ích quan trọng. Trước hết, sự xuất hiện của ảnh IRAP chỉ báo rằng quá trình giải mã có thể bắt đầu từ ảnh đó. Chức năng này cho phép đặc tính truy nhập ngẫu nhiên trong đó quá trình giải mã bắt đầu ở vị trí đó trong dòng bit, không nhất thiết là bắt đầu của dòng bit, miễn là ảnh IRAP xuất hiện ở vị trí đó. Thứ hai, sự xuất hiện của ảnh IRAP làm mới quá trình giải mã sao cho ảnh được tạo mã bắt đầu ở ảnh IRAP, không bao gồm ảnh dẫn đầu bỏ qua truy nhập ngẫu nhiên (random access skipped leading, RASL), được tạo mã không dựa vào các ảnh trước đó. Kết quả là, việc có ảnh IRAP xuất hiện trong dòng bit sẽ dừng bất kỳ lỗi có thể xuất hiện trong quá trình giải mã các ảnh được tạo mã trước ảnh IRAP để lan truyền đến ảnh IRAP và các ảnh đó tuân theo ảnh IRAP theo thứ tự giải mã.

Trong khi các ảnh IRAP cung cấp các chức năng quan trọng, chúng gây ảnh hưởng đến hiệu suất nén. Sự có mặt của ảnh IRAP làm tăng đột ngột tốc độ bit. Việc ảnh hưởng đến hiệu suất nén là do hai nguyên nhân. Trước hết, do ảnh IRAP là ảnh được dự báo trong, chính ảnh này sẽ yêu cầu nhiều bit hơn để biểu diễn khi so sánh với các ảnh khác (chẳng hạn, ảnh đứng đầu, ảnh đứng cuối) vốn là các ảnh được dự báo ngoài. Thứ hai, do sự có mặt của ảnh IRAP phá vỡ dự báo thời gian (đây là do bộ giải mã làm mới quá trình giải mã, trong đó một trong các hành động của quá trình giải mã chính là loại bỏ các ảnh tham chiếu trước đó trong DPB), ảnh IRAP khiến việc tạo mã các ảnh đi sau ảnh IRAP theo

thứ tự giải mã kém hiệu quả (tức là, cần nhiều bit hơn để biểu diễn) do chúng có ít ảnh tham chiếu hơn để mã hóa dự báo ngoài..

Theo phương án thực hiện, ảnh IRAP được gọi là ảnh truy nhập ngẫu nhiên sạch (clean random access, CRA) hoặc ảnh làm mới bộ giải mã tức thời (instantaneous decoder refresh, IDR) với ảnh truy nhập ngẫu nhiên có thể giải mã (random access decodable, RADL). Trong chuẩn HEVC, các ảnh IDR, các ảnh CRA, và các ảnh truy nhập liên kết đứt (Broken Link Access, BLA) đều được xem là ảnh IRAP. Đối với VVC, trong cuộc họp JVET thứ 12 vào tháng 10 2018, đồng thuận xem cả ảnh IDR và CRA làm ảnh IRAP. Theo phương án thực hiện, các ảnh BLA và làm mới bộ giải mã dần dần (Gradual Decoder Refresh, GDR) cũng có thể được xem là ảnh IRAP. Quá trình giải mã cho CVS luôn bắt đầu ở ảnh IRAP.

Ngược với việc truyền các ảnh IRAP mới như nêu trên, cách tốt hơn là tiếp tục truyền bất kỳ ô nào (tức là, ảnh con) được chia sẻ giữa rãnh trích xuất 610 và rãnh trích xuất 710. Tức là, tiếp tục truyền các ô có các ID ô sau: 4, 8, 10, và 14 do các ô đó đều nằm trong rãnh trích xuất 610 và rãnh trích xuất 710. Khi làm vậy, các ảnh IRAP mới chỉ cần được truyền cho các ô trong rãnh trích xuất 710 cũng không nằm tổng rãnh trích xuất 610. Tức là, các ảnh IRAP mới chỉ cần được truyền cho các ô có các ID ô sau: 1, 5, 9, và 13 khi hướng nhìn thay đổi. Tuy nhiên, thay đổi trong các ID ảnh con trong CVS có thể gây ra các vấn đề về báo hiệu.

Để giải quyết ít nhất vấn đề báo hiệu, chỉ báo (chẳng hạn, cờ) được báo hiệu trong dòng bit để chỉ báo liệu ID ô (tức là, ID ảnh con) cho mỗi ô (tức là, ảnh con) có thể thay đổi trong CVS. Cờ này có thể được báo hiệu trong SPS hoặc PPS. Bên cạnh báo hiệu liệu ID ô có thể thay đổi trong CVS, cờ cũng có thể có chức năng khác.

Theo một cách, các ID ảnh con được báo hiệu hoặc trong cú pháp SPS (khi chỉ báo rằng ID ảnh con cho mỗi ảnh con không thay đổi trong CVS) hoặc cú pháp PPS (khi chỉ báo rằng ID ảnh con cho mỗi ảnh con có thể thay đổi trong CVS). Theo phương án thực hiện, các ID ảnh con không bao giờ được báo hiệu trong cả cú pháp SPS và cú pháp PPS.

Theo cách khác, các ID ảnh con luôn được báo hiệu trong cú pháp SPS, và khi cờ chỉ báo rằng ID ảnh con cho mỗi ảnh con có thể thay đổi trong CVS, các giá trị ID ảnh con được báo hiệu trong SPS có thể bị chồng bởi các ID ảnh con được báo hiệu trong cú pháp PPS.

Theo cách khác nữa, các ID ảnh con không được báo hiệu trong cú pháp SPS nhưng chỉ trong cú pháp PPS, và chỉ báo liệu ID ảnh con cho mỗi ảnh con có thể thay đổi trong CVS cũng chỉ được báo hiệu trong cú pháp PPS.

Theo cách này, thông tin ảnh con khác, chẳng hạn vị trí và kích thước của mỗi ảnh con, chiều dài của các ID ảnh con trong các bit, cũng như các cờ `subpic_treated_as_pic_flag[i]` và `loop_filter_across_subpic_enabled_flag[i]` như trong đặc tả phác thảo VVC mới nhất, cũng có thể được báo hiệu trong PPS thay vì trong SPS, nhưng tất cả là giống nhau đối với tất cả các PPS được tham chiếu bởi các ảnh được tạo mã trong CVS.

Khi các vị trí và các kích thước ảnh con được báo hiệu trong PPS, thay vì được báo hiệu tương tự như trong đặc tả phác thảo VVC mới nhất, chúng có thể được báo hiệu dựa trên các lát được bao gồm trong mỗi ảnh con. Chẳng hạn, cho mỗi ảnh con, chỉ số lát hoặc ID của lát được đặt ở góc trên bên trái ảnh con và chỉ số lát hoặc ID hoặc lát được đặt ở góc dưới bên phải của ảnh con có thể được báo hiệu, để dẫn xuất vị trí và kích thước ảnh con, trong đó việc báo hiệu này có thể dựa trên delta, và trong một số trường hợp cụ thể, việc báo hiệu chỉ số lát hoặc ID hoặc delta của nó có thể được tránh và giá trị được suy ra, chẳng hạn, tương tự như cách thức các chỉ số khối viên trên bên trái và dưới bên phải được báo hiệu cho các lát chữ nhật trong đặc tả phác thảo VVC mới nhất.

Theo phương án thực hiện, các ID ảnh con được báo hiệu như sau.

Trong cú pháp SPS, khi chỉ báo rằng số lượng ảnh con trong mỗi ảnh trong CVS lớn hơn 1, phương trình sau áp dụng.

Cờ (chẳng hạn, `designated_subpicture_ids_signalled_in_sps_flag` hoặc `sps_subpic_id_mapping_present_flag`) được báo hiệu trong SPS. Theo phương án thực hiện, cờ có ngữ nghĩa sau: `subpicture_ids_signalled_in_sps_flag` bằng 1 xác định rằng các ID ảnh con được báo hiệu trong SPS, 1 cho mỗi ảnh con, và giá trị ID ảnh con cho mỗi ảnh con cụ thể không thay đổi trong CVS.

Subpicture_ids_signalled_in_sps_flag bằng 0 xác định rằng các ID ảnh con không được báo hiệu trong SPS mà thay vào đó được báo hiệu trong PPS, và giá trị ID ảnh con cho mỗi ảnh con cụ thể có thể thay đổi trong CVS. Khi subpicture_ids_signalled_in_sps_flag bằng 1, ID ảnh con được báo hiệu cho mỗi ảnh con trong SPS.

Trong cú pháp PPS, cờ (chẳng hạn, subpicture_ids_signalled_in_pps_flag hoặc pps_subpic_id_mapping_present_flag được chỉ định) được báo hiệu trong PPS. Cờ có các ngữ nghĩa sau: subpicture_ids_signalled_in_pps_flag bằng 1 xác định rằng các ID ảnh con được báo hiệu trong PPS, 1 cho mỗi ảnh con, và giá trị ID ảnh con cho mỗi ảnh con cụ thể có thể thay đổi trong CVS. Subpicture_ids_signalled_in_pps_flag bằng 0 xác định rằng các ID ảnh con không được báo hiệu trong PPS mà thay vào đó được báo hiệu trong SPS, và giá trị ID ảnh con cho mỗi ảnh con cụ thể không thay đổi trong CVS.

Theo phương án thực hiện, giá trị của subpicture_ids_signalled_in_pps_flag sẽ bằng 1 – subpicture_ids_signalled_in_sps_flag. Khi subpicture_ids_signalled_in_pps_flag bằng 1, ID ảnh con được báo hiệu cho mỗi ảnh con trong PPS.

Trong cú pháp tiêu đề lát, ID ảnh con được báo hiệu bất kể số lượng ảnh con được xác định bởi SPS được tham chiếu.

Theo cách khác, khi số lượng ảnh con trong mỗi ảnh lớn hơn 1, cú pháp SPS luôn bao gồm các ID ảnh con, 1 cho mỗi ảnh con, và cú pháp SPS cũng bao gồm cờ xác định liệu các ID ảnh con có thể bị chồng bởi các ID ảnh con được báo hiệu trong cú pháp PPS và liệu giá trị ID ảnh con cho mỗi ảnh con cụ thể có thể thay đổi trong CVS. Việc ghi chồng các ID ảnh con có thể được thực hiện hoặc luôn luôn cho tất cả các ID ảnh con, hoặc chỉ cho tập con được chọn của tất cả các ID ảnh con.

Theo phương án thực hiện, Fig.8 là phương pháp 800 giải mã được thực hiện bởi bộ giải mã video (chẳng hạn, bộ giải mã video 400). Phương pháp 800 có thể được thực hiện sau khi dòng bit được giải mã đã được nhận trực tiếp hoặc gián tiếp từ bộ mã hóa video (chẳng hạn, bộ mã hóa video 300). Phương pháp 800 cải thiện quá trình giải mã bằng cách đảm bảo báo hiệu hiệu quả các ID ảnh

con thậm chí khi các ID ảnh con thay đổi trong CVS trong các kịch bản ứng dụng bao gồm cả trích xuất dòng bit thành phần và hợp nhất dòng bit thành phần. Điều này giảm độ dư thừa và tăng hiệu suất tạo mã. Do vậy, bộ mã hóa/bộ giải mã (tức là, “codec”) trong khi tạo mã video được cải thiện liên quan đến các codec hiện tại. Thực tế là, quá trình tạo mã video được cải tiến cho người dùng trải nghiệm sử dụng tốt hơn khi các video được gửi, được nhận, và/hoặc được xem.

Trong khối 802, bộ giải mã video nhận dòng bit video (chẳng hạn, dòng bit 500) bao gồm SPS (chẳng hạn, SPS 510), PPS (chẳng hạn, PPS 512), và các ảnh con (chẳng hạn, các ô 605 và 607) được liên kết với ánh xạ ID ảnh con 575. Lưu ý trên đây, ánh xạ của ID ảnh con 575 là ánh xạ của các ID ảnh con đến các ảnh con cụ thể theo chỉ số ảnh con (chẳng hạn, ID ảnh con 8 tương ứng với chỉ số ảnh con 8, nhận diện ảnh con cụ thể từ các ảnh con). Theo phương án thực hiện, SPS 510 chứa cờ SPS 565. Theo phương án thực hiện, PPS 512 chứa cờ PPS 567.

Trong khối 804, bộ giải mã video xác định liệu cờ SPS 565 có giá trị thứ nhất hoặc giá trị thứ hai. Cờ SPS 565 có giá trị thứ nhất xác định rằng ánh xạ của ID ảnh con 575 được bao gồm trong SPS 510, và cờ SPS 565 có giá trị thứ hai xác định rằng ánh xạ của ID ảnh con 575 được báo hiệu trong PPS 512.

Khi cờ SPS 565 có giá trị thứ hai, trong khối 806, bộ giải mã video xác định liệu cờ PPS 567 có giá trị thứ nhất hoặc giá trị thứ hai. Cờ PPS 567 có giá trị thứ nhất xác định rằng ánh xạ của ID ảnh con 575 được bao gồm trong PPS 512, và cờ PPS 567 có giá trị thứ hai xác định rằng ánh xạ của ID ảnh con 575 không được báo hiệu trong PPS 512.

Theo phương án thực hiện, cờ SPS 656 có giá trị thứ nhất khi cờ PPS 567 có giá trị thứ hai. Theo phương án thực hiện, cờ SPS 565 có giá trị thứ hai khi cờ PPS 567 có giá trị thứ nhất. Theo phương án thực hiện, giá trị thứ nhất bằng 1 và giá trị thứ hai bằng 0.

Theo phương án thực hiện, SPS 510 chứa cờ SPS thứ hai 569. Cờ SPS thứ hai 569 xác định liệu việc ánh xạ của ID ảnh con 575 được báo hiệu tường minh trong SPS 510 hoặc PPS 512. Theo phương án thực hiện, dòng bit 500 còn bao

gồm cờ thay đổi CVS 571. Cờ thay đổi CVS 571 chỉ báo liệu ánh xạ của ID ảnh con 575 có thể thay đổi trong CVS 590 of dòng bit 500. Theo phương án thực hiện, cờ thay đổi CVS 571 được bao gồm trong SPS 510, PPS 512, hoặc tập hợp tham số khác hoặc tiêu đề trong dòng bit 500.

Trong khối 808, bộ giải mã video thu được ánh xạ của ID ảnh con 575 từ SPS 510 khi cờ SPS 565 có giá trị thứ nhất và từ PPS 512 khi SPS có giá trị thứ hai và/hoặc khi cờ PPS 567 có giá trị thứ nhất. Theo phương án thực hiện, dòng bit bao gồm dòng bit được hợp nhất. Theo phương án thực hiện, ánh xạ của ID ảnh con 575 đã thay đổi trong CVS 590 của dòng bit 500.

Trong khối 810, bộ giải mã video giải mã các ảnh con nhờ sử dụng ánh xạ của ID ảnh con 575. Khi được giải mã, các ảnh con có thể được sử dụng để tạo ảnh hoặc chuỗi video để hiển thị đến người dùng trên màn hình hoặc bộ phận hiển thị của thiết bị điện tử (chẳng hạn, điện thoại thông minh, tablet, máy tính xách tay, máy tính cá nhân, v.v.).

Theo phương án thực hiện, Fig.9 là phương pháp 900 mã hóa dòng bit video được thực hiện bởi bộ mã hóa video (chẳng hạn, bộ mã hóa video 300). Phương pháp 900 có thể được thực hiện khi ảnh (chẳng hạn, từ video) cần được mã hóa thành dòng bit video và sau đó được truyền về phía bộ giải mã video (chẳng hạn, bộ giải mã video 400). Phương pháp 900 cải thiện quá trình mã hóa bằng cách đảm bảo báo hiệu hiệu quả các ID ảnh con thậm chí khi các ID ảnh con thay đổi trong CVS trong các kịch bản ứng dụng bao gồm cả trích xuất dòng bit thành phần và hợp nhất dòng bit thành phần. Điều này giảm độ dư thừa và tăng hiệu suất tạo mã. Do vậy, bộ mã hóa/bộ giải mã (tức là, “codec”) trong khi tạo mã video được cải thiện liên quan đến các codec hiện tại. Thực tế là, quá trình tạo mã video được cải tiến cho người dùng trải nghiệm sử dụng tốt hơn khi các video được gửi, được nhận, và/hoặc được xem.

Trong khối 902, bộ mã hóa video mã hóa dòng bit bao gồm SPS (chẳng hạn, SPS 510), PPS (chẳng hạn, PPS 512), và các ảnh con (chẳng hạn, các ô 605 và 607) được liên kết với ánh xạ ID ảnh con. Như lưu ý trên đây, ánh xạ của ID ảnh con 575 là ánh xạ của các ID ảnh con đến các ảnh con cụ thể thông qua chỉ số ảnh con (chẳng hạn, ID ảnh con 8 tương ứng với chỉ số ảnh con 8, nhận diện ảnh

con cụ thể từ các ảnh con). Theo phương án thực hiện, SPS 510 chứa cờ SPS 565 và PPS 512 chứa cờ PPS 567.

Trong khối 904, bộ mã hóa video gán cờ SPS 565 bằng giá trị thứ nhất khi ánh xạ của ID ảnh con 575 được bao gồm trong SPS 510 và bằng giá trị thứ hai khi ánh xạ của ID ảnh con 575 được báo hiệu trong PPS 512.

Trong khối 906, bộ mã hóa video gán cờ PPS 567 bằng giá trị thứ nhất khi ánh xạ của ID ảnh con 575 được bao gồm trong PPS 512 và bằng giá trị thứ hai khi ánh xạ của ID ảnh con 575 không được báo hiệu trong PPS 512.

Theo phương án thực hiện, cờ SPS 565 có giá trị thứ nhất khi cờ PPS 567 có giá trị thứ hai. Theo phương án thực hiện, cờ SPS 565 có giá trị thứ hai khi cờ PPS 567 có giá trị thứ nhất. Theo phương án thực hiện, giá trị thứ nhất bằng 1 và giá trị thứ hai bằng 0.

Theo phương án thực hiện, SPS 510 chứa cờ SPS thứ hai 569. Cờ SPS thứ hai 569 xác định liệu ánh xạ của ID ảnh con 575 được báo hiệu tường minh trong SPS 510 hoặc PPS 512. Theo phương án thực hiện, dòng bit 500 còn bao gồm cờ thay đổi CVS 571. Cờ thay đổi CVS 571 chỉ báo liệu ánh xạ của ID ảnh con 575 có thể thay đổi trong CVS 590 của dòng bit 500. Theo phương án thực hiện, cờ thay đổi CVS 571 được bao gồm trong SPS 510, PPS 512, hoặc tập hợp tham số khác hoặc tiêu đề trong dòng bit 500.

Trong khối 908, bộ mã hóa video lưu trữ dòng bit để truyền thông về phía bộ giải mã. Dòng bit có thể được lưu trữ trong bộ nhớ cho đến khi dòng bit video được truyền về phía bộ giải mã video. Khi được nhận bởi bộ giải mã video, video được mã hóa dòng bit có thể được giải mã (chẳng hạn, như nêu trên) để tạo ảnh hoặc chuỗi video để hiển thị đến người dùng trên bộ phận hiển thị hoặc màn hình của thiết bị điện tử (chẳng hạn, điện thoại thông minh, tablet, máy tính xách tay, máy tính cá nhân, v.v.).

Theo phương án thực hiện, các ID ảnh con được báo hiệu như sau.

cú pháp SPS không bao gồm báo hiệu các ID ảnh con.

Trong cú pháp PPS, giá trị chỉ báo số lượng ảnh con được báo hiệu, bắt buộc là giống nhau cho tất cả các PPS được tạo tham chiếu bởi các ảnh được tạo mã

trong CVS, và khi số lượng ảnh con được chỉ báo lớn hơn 1, điều kiện sau áp dụng:

Cờ, chẳng hạn, tên là `subpicture_id_unchanging_in_cvs_flag`, được báo hiệu trong PPS, với các ngữ nghĩa sau: `subpicture_id_unchanging_in_cvs_flag` bằng 1 xác định rằng ID ảnh con cho mỗi ảnh con cụ thể được báo hiệu trong PPS không thay đổi trong CVS. `subpicture_id_unchanging_in_cvs_flag` bằng 0 xác định rằng ID ảnh con cho mỗi ảnh con cụ thể được báo hiệu trong PPS có thể thay đổi trong CVS.

Giá trị chỉ báo số lượng ảnh con được báo hiệu trong PPS. Số lượng ảnh con được chỉ báo sẽ giống nhau cho tất cả các PPS được tạo tham chiếu bởi các ảnh được tạo mã trong CVS.

ID ảnh con được báo hiệu cho mỗi ảnh con trong PPS. Khi `subpicture_id_unchanging_in_cvs_flag` bằng 1, ID ảnh con cho mỗi ảnh con cụ thể sẽ giống nhau cho tất cả các PPS được tạo tham chiếu bởi các ảnh được tạo mã trong CVS.

Trong cú pháp tiêu đề lát, ID ảnh con được báo hiệu, bất kể số lượng ảnh con được xác định bởi SPS được tạo tham chiếu.

Cũng cần hiểu rằng các bước của các phương pháp lấy làm ví dụ được nêu ở đây không bắt buộc phải được thực hiện theo thứ tự được mô tả, và thứ tự của các bước của các phương pháp này nên được hiểu là chỉ lấy làm ví dụ. Tương tự, các bước bổ sung có thể được bao gồm trong các phương pháp này, và các bước cụ thể có thể bị bỏ qua hoặc kết hợp, trong các phương pháp nhất quán với các phương án thực hiện khác nhau của sáng chế.

Fig.10 là sơ đồ của thiết bị tạo mã video 1000 (chẳng hạn, bộ mã hóa video 300 hoặc bộ giải mã video 400) theo phương án thực hiện sáng chế. Thiết bị tạo mã video 1000 thích hợp để thực hiện các phương án thực hiện được bộc lộ như được nêu ở đây. Thiết bị tạo mã video 1000 bao gồm các cổng vào 1010 và các bộ thu (Rx) 1020 để nhận dữ liệu; bộ xử lý, đơn vị logic, hoặc khối xử lý trung tâm (central processing unit, CPU) 1030 để xử lý dữ liệu; các bộ phát (Tx) 1040 và các cổng ra 1050 để truyền dữ liệu; và bộ nhớ 1060 để lưu trữ dữ liệu. Thiết bị tạo mã video 1000 cũng có thể bao gồm các thành phần quang sang điện

(optical-to-electrical, OE) và các thành phần điện sang quang (electrical-to-optical, EO) được ghép nối với các cổng vào 1010, các bộ thu 1020, các bộ phát 1040, và các cổng ra 1050 để xuất hoặc nhập các tín hiệu quang hoặc điện.

Bộ xử lý 1030 được thực hiện bởi phần cứng và phần mềm. Bộ xử lý 1030 có thể được thực hiện dưới dạng một hoặc nhiều chip CPU, lõi (chẳng hạn, dưới dạng bộ xử lý nhiều lõi), mảng cổng dạng trường lập trình được (field-programmable gate-array, FPGA), mạch tích hợp ứng dụng cụ thể (application specific integrated circuit, ASIC), và bộ xử lý tín hiệu số (digital signal processor, DSP). Bộ xử lý 1030 truyền thông với các cổng vào 1010, bộ thu 1020, bộ phát 1040, cổng ra 1050, và bộ nhớ 1060. Bộ xử lý 1030 bao gồm mô-đun tạo mã 1070. Mô-đun tạo mã 1070 thực hiện các phương án thực hiện được bộc lộ nêu trên. Chẳng hạn, mô-đun tạo mã 1070 thực hiện, xử lý, chuẩn bị, hoặc cấp các chức năng codec khác nhau. Do vậy, việc bao gồm mô-đun tạo mã 1070 cải thiện đáng kể chức năng của thiết bị tạo mã video 1000 và biến đổi thiết bị tạo mã video 1000 sang trạng thái khác. Theo cách khác, mô-đun tạo mã 1070 được triển khai dưới dạng các lệnh được lưu trữ trong bộ nhớ 1060 và được thực thi bởi bộ xử lý 1030.

Thiết bị tạo mã video 1000 cũng có thể bao gồm các bộ phận nhập và/hoặc xuất (I/O) 1080 để truyền thông dữ liệu đến và từ người dùng. Các bộ phận I/O 1080 có thể bao gồm các bộ phận xuất, chẳng hạn bộ phận hiển thị để hiển thị dữ liệu video, loa để xuất dữ liệu audio v.v.. Các bộ phận I/O 1080 có thể cũng bao gồm các bộ phận nhập, chẳng hạn, bàn phím, chuột, bi xoay, v.v., và/hoặc các giao diện tương ứng để tương tác với các bộ phận xuất này.

Bộ nhớ 1060 bao gồm một hoặc nhiều đĩa, ổ băng, và ổ trạng thái rắn và có thể được sử dụng làm bộ lưu trữ dữ liệu tràn, để lưu trữ chương trình khi các chương trình này được chọn để thực thi, và để lưu trữ các lệnh và dữ liệu được đọc trong khi thực thi chương trình. Bộ nhớ 1060 có thể là khả biến và/hoặc bất biến và có thể là bộ nhớ chỉ đọc (read-only memory, ROM), bộ nhớ truy cập ngẫu nhiên (random access memory, RAM), bộ nhớ có nội dung đánh địa chỉ được tam phân (ternary content-addressable memory, TCAM), và/hoặc RAM tĩnh (static random-access memory, SRAM).

Theo phương án thực hiện, Fig.11 là sơ đồ của phương tiện tạo mã 1100. Theo phương án thực hiện, phương tiện tạo mã 1100 được triển khai trong thiết bị tạo mã video 1102 (chẳng hạn, bộ mã hóa video 300 hoặc bộ giải mã video 400). Thiết bị tạo mã video 1102 bao gồm phương tiện nhận 1101. Phương tiện nhận 1101 được tạo cấu hình để nhận ảnh để mã hóa hoặc để nhận dòng bit để giải mã. Thiết bị tạo mã video 1102 bao gồm phương tiện truyền 1107 được ghép nối với phương tiện nhận 1101. Phương tiện truyền 1107 được tạo cấu hình để truyền dòng bit đến bộ giải mã hoặc để truyền ảnh được giải mã đến phương tiện hiển thị (chẳng hạn, một trong các bộ phận I/O 1080).

Thiết bị tạo mã video 1102 bao gồm phương tiện lưu trữ 1103. Phương tiện lưu trữ 1103 được ghép nối với ít nhất một trong phương tiện nhận 1101 hoặc phương tiện truyền 1107. Phương tiện lưu trữ 1103 được tạo cấu hình để lưu trữ các lệnh. Thiết bị tạo mã video 1102 also bao gồm phương tiện xử lý 1105. Phương tiện xử lý 1105 được ghép nối với phương tiện lưu trữ 1103. Phương tiện xử lý 1105 được tạo cấu hình để thực thi các lệnh được lưu trữ trong phương tiện lưu trữ 1103 để thực hiện các phương pháp được bộc lộ ở đây.

Trong khi vài phương án thực hiện được nêu theo sáng chế, có thể hiểu rằng các hệ thống và các phương pháp được bộc lộ có thể được triển khai ở nhiều dạng cụ thể khác mà không xa rời phạm vi sáng chế. Các ví dụ này được xem là minh họa và không giới hạn, và ý tưởng sẽ không bị giới hạn ở các chi tiết được nêu ở đây. Chẳng hạn, các phần tử hoặc các thành phần khác nhau có thể được kết hợp hoặc tích hợp trong hệ thống khác hoặc các dấu hiệu khác có thể bị bỏ qua, hoặc không được triển khai.

Ngoài ra, các kỹ thuật, các hệ thống, các hệ thống phụ, và các phương pháp được mô tả và minh họa theo các phương án thực hiện khác nhau dưới dạng rời rạc hoặc riêng rẽ có thể được kết hợp hoặc tích hợp vào các hệ thống, các thành phần, các kỹ thuật, hoặc các phương pháp khác mà không xa rời phạm vi sáng chế. Các ví dụ khác về các thay đổi, thay thế, và biến đổi được người có hiểu biết trung bình trong lĩnh vực hiểu và có thể được thực hiện mà không xa rời phạm vi được bộc lộ ở đây.

YÊU CẦU BẢO HỘ

1. Phương pháp giải mã được triển khai bởi bộ giải mã bao gồm các bước:

nhận (802), bởi bộ giải mã, dòng bit bao gồm tập hợp tham số chuỗi (sequence parameter set, SPS), tập hợp tham số ảnh (picture parameter set, PPS), và các ảnh con được liên kết với ánh xạ bộ nhận dạng (identifier, ID) ảnh con, trong đó SPS bao gồm cờ SPS;

xác định (804), bởi bộ giải mã, liệu cờ SPS có giá trị thứ nhất hoặc giá trị thứ hai, trong đó cờ SPS có giá trị thứ nhất xác định rằng ánh xạ ID ảnh con được báo hiệu trong SPS, và trong đó cờ SPS có giá trị thứ hai xác định rằng ánh xạ ID ảnh con được báo hiệu trong PPS;

thu được (808), bởi bộ giải mã, ánh xạ ID ảnh con từ SPS khi cờ SPS có giá trị thứ nhất và từ PPS khi cờ SPS có giá trị thứ hai; và

giải mã (810), bởi bộ giải mã, các ảnh con sử dụng ánh xạ ID ảnh con.

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

xác định (806), bởi bộ giải mã, liệu cờ PPS có giá trị thứ nhất hoặc giá trị thứ hai, trong đó cờ PPS có giá trị thứ nhất xác định rằng ánh xạ ID ảnh con được báo hiệu trong PPS, và trong đó cờ PPS có giá trị thứ hai xác định rằng ánh xạ ID ảnh con không được báo hiệu trong PPS; và

thu được (808), bởi bộ giải mã, ánh xạ ID ảnh con từ PPS khi cờ PPS có giá trị thứ nhất.

3. Phương pháp theo điểm 2, trong đó cờ SPS có giá trị thứ nhất khi cờ PPS có giá trị thứ hai, và trong đó cờ SPS có giá trị thứ hai khi cờ PPS có giá trị thứ nhất.

4. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 3, trong đó giá trị thứ nhất bằng 1 và giá trị thứ hai bằng 0.

5. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 4, trong đó SPS bao gồm cờ SPS thứ hai, và trong đó cờ SPS thứ hai xác định liệu ánh xạ ID ảnh con được báo hiệu tường minh trong SPS hoặc PPS.

6. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 5, trong đó dòng bit còn bao gồm cờ thay đổi chuỗi video được tạo mã (coded video sequence, CVS), và trong đó cờ thay đổi CVS chỉ báo liệu ánh xạ ID ảnh con được phép thay đổi trong CVS của dòng bit.

7. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 6, trong đó dòng bit bao gồm dòng bit được hợp nhất, và trong đó ánh xạ ID ảnh con đã thay đổi trong CVS của dòng bit.

8. Phương pháp mã hóa được triển khai bởi bộ mã hóa, trong đó phương pháp bao gồm các bước:

mã hóa (902), bởi bộ mã hóa, dòng bit bao gồm SPS, PPS, và các ảnh con được liên kết với ánh xạ ID ảnh con, trong đó SPS bao gồm cờ SPS;

gán (904), bởi bộ mã hóa, cờ SPS bằng giá trị thứ nhất khi ánh xạ ID ảnh con được báo hiệu trong SPS và bằng giá trị thứ hai khi ánh xạ ID ảnh con được báo hiệu trong PPS; và

lưu trữ (908), bởi bộ giải mã, dòng bit để truyền về phía bộ giải mã.

9. Phương pháp theo điểm 8, trong đó phương pháp còn bao gồm bước thiết lập tiếp (906), bởi bộ mã hóa, cờ PPS trong PPS bằng giá trị thứ nhất khi ánh xạ ID ảnh con được báo hiệu trong PPS và bằng giá trị thứ hai khi ánh xạ ID ảnh con không được báo hiệu trong PPS.

10. Phương pháp theo điểm 9, trong đó cờ SPS có giá trị thứ nhất khi cờ PPS có giá trị thứ hai, và trong đó cờ SPS có giá trị thứ hai khi cờ PPS có giá trị thứ nhất.

11. Phương pháp theo điểm bất kỳ trong số các điểm từ 8 đến 10, trong đó giá trị thứ nhất bằng 1 và giá trị thứ hai bằng 0.

12. Phương pháp theo điểm bất kỳ trong số các điểm từ 8 đến 11, trong đó SPS bao gồm cờ SPS thứ hai, và trong đó cờ SPS thứ hai xác định liệu ánh xạ ảnh con được báo hiệu tường minh trong SPS hoặc PPS.

13. Phương pháp theo điểm bất kỳ trong số các điểm từ 8 đến 12, trong đó dòng bit còn bao gồm cờ thay đổi CVS, và trong đó cờ thay đổi CVS chỉ báo liệu ánh xạ ID ảnh con có thể thay đổi trong CVS của dòng bit.

14. Thiết bị giải mã bao gồm:

bộ nhận (1020) được tạo cấu hình để nhận dòng bit video bao gồm SPS, PPS, và các ảnh con được liên kết với ánh xạ ID ảnh con, trong đó SPS bao gồm cờ SPS;

bộ nhớ (1032) được ghép nối với bộ nhận, bộ nhớ lưu trữ các lệnh; và

bộ xử lý (1030) được ghép nối với bộ nhớ (1032), bộ xử lý (1030) được tạo cấu hình để thực thi các lệnh để khiến thiết bị giải mã để:

xác định liệu cờ SPS có giá trị thứ nhất hoặc giá trị thứ hai, trong đó cờ SPS có giá trị thứ nhất xác định rằng ánh xạ ID ảnh con được báo hiệu trong SPS, và trong đó cờ SPS có giá trị thứ hai xác định rằng ánh xạ ID ảnh con được báo hiệu trong PPS;

thu được ánh xạ ID ảnh con từ SPS khi cờ SPS có giá trị thứ nhất và từ PPS khi cờ SPS có giá trị thứ hai; và

giải mã các ảnh con sử dụng ánh xạ ID ảnh con.

1/10

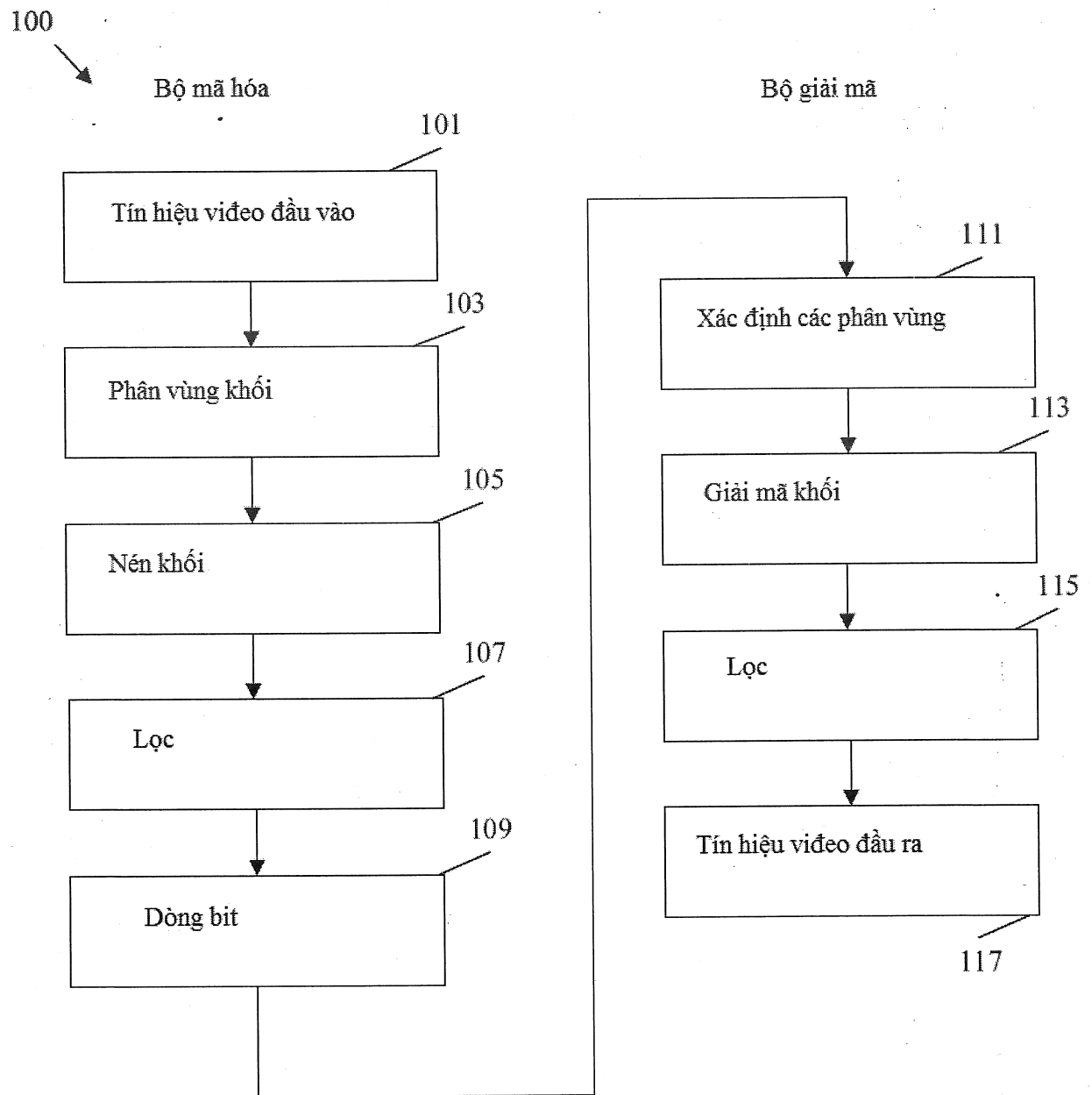


Fig.1

2/10

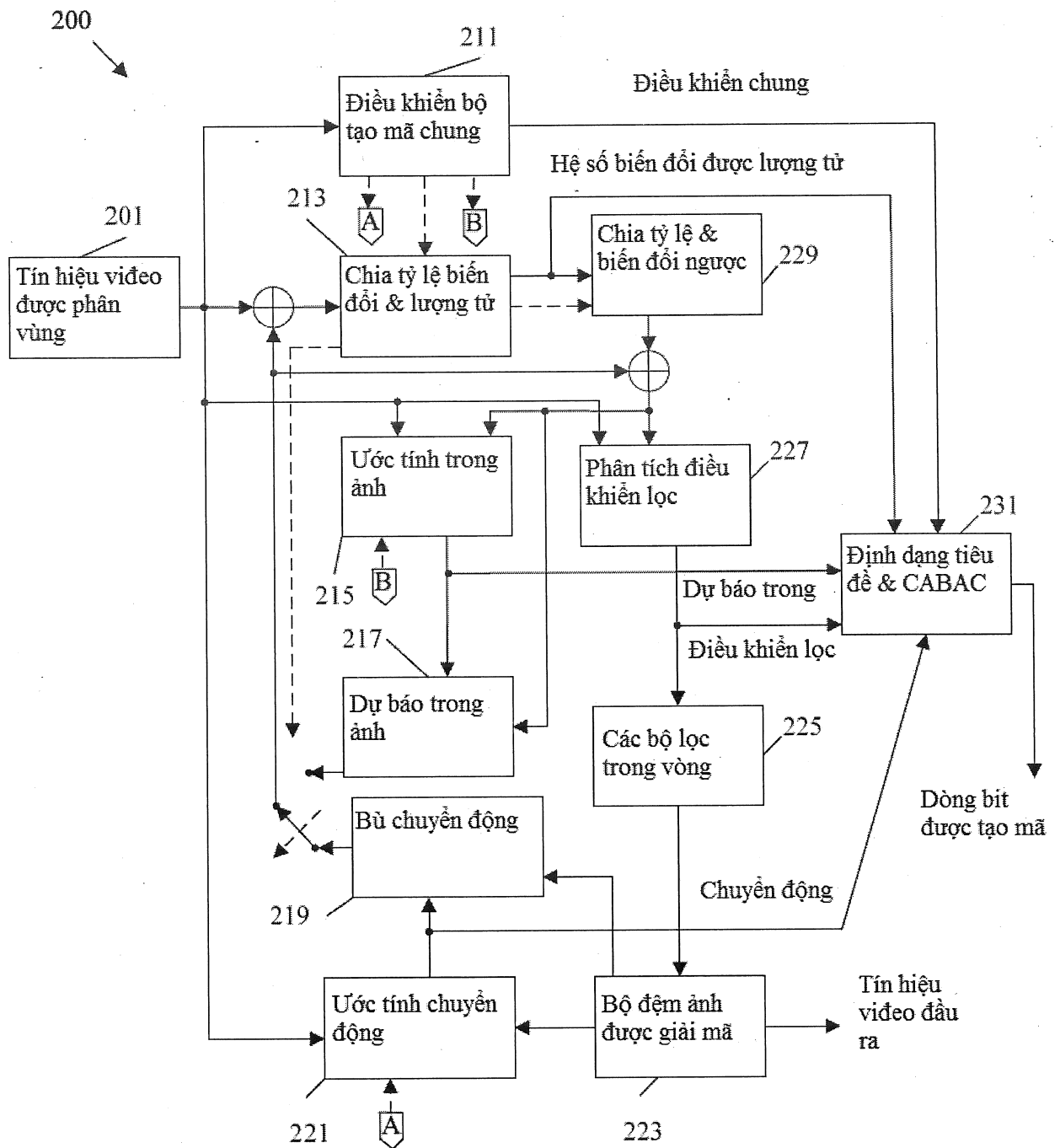


Fig.2

3/10

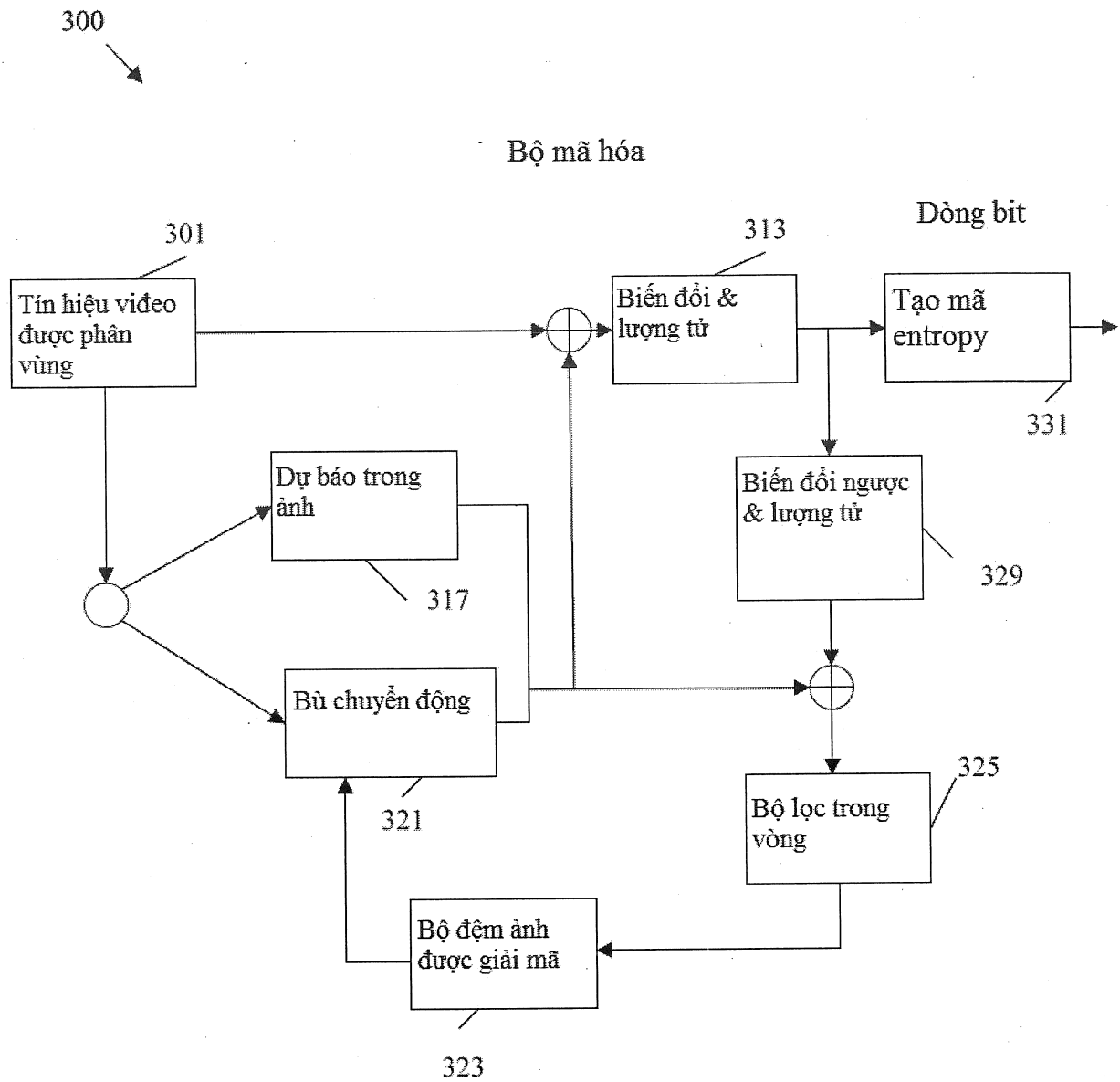


Fig.3

4/10

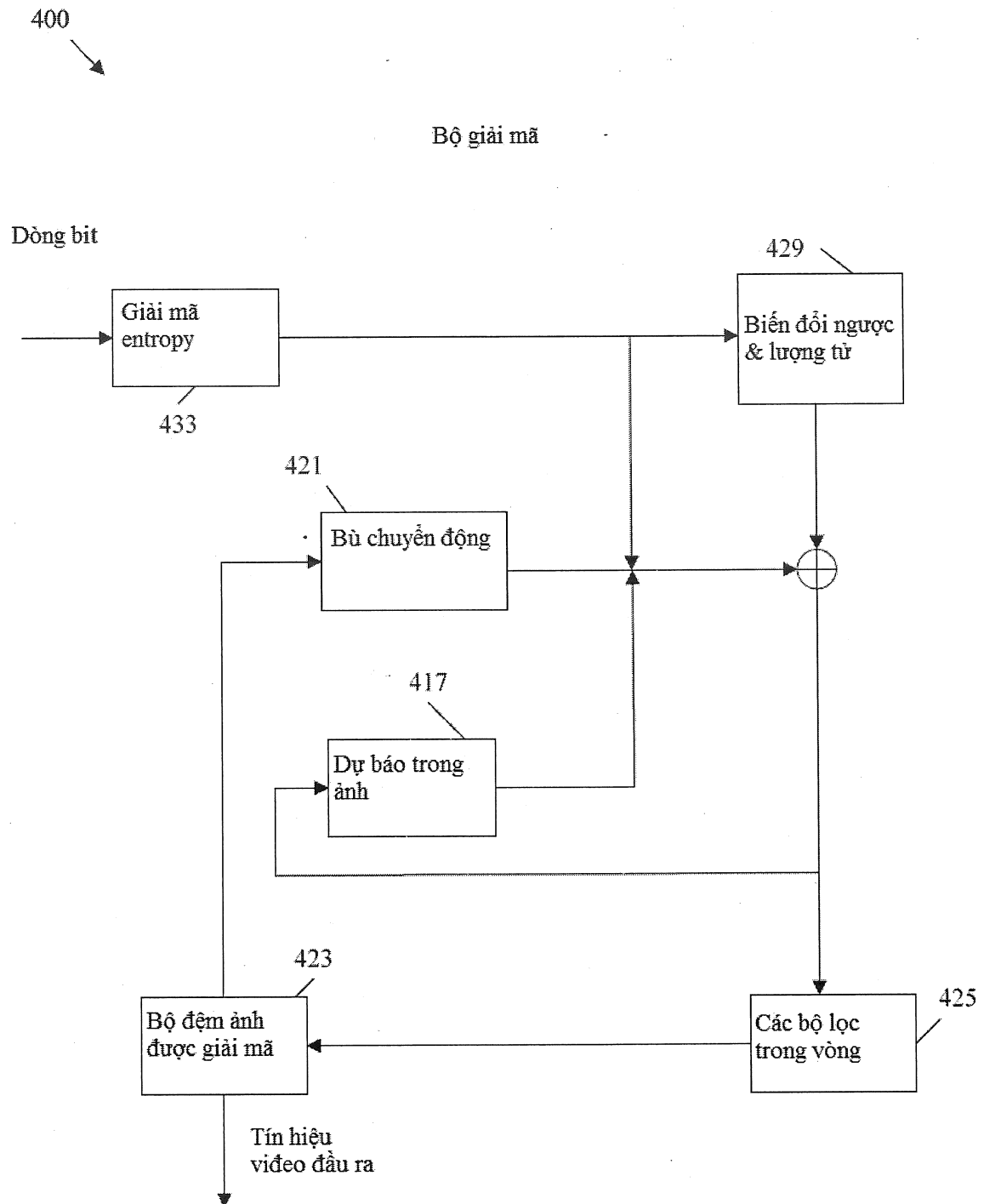


Fig.4

5/10

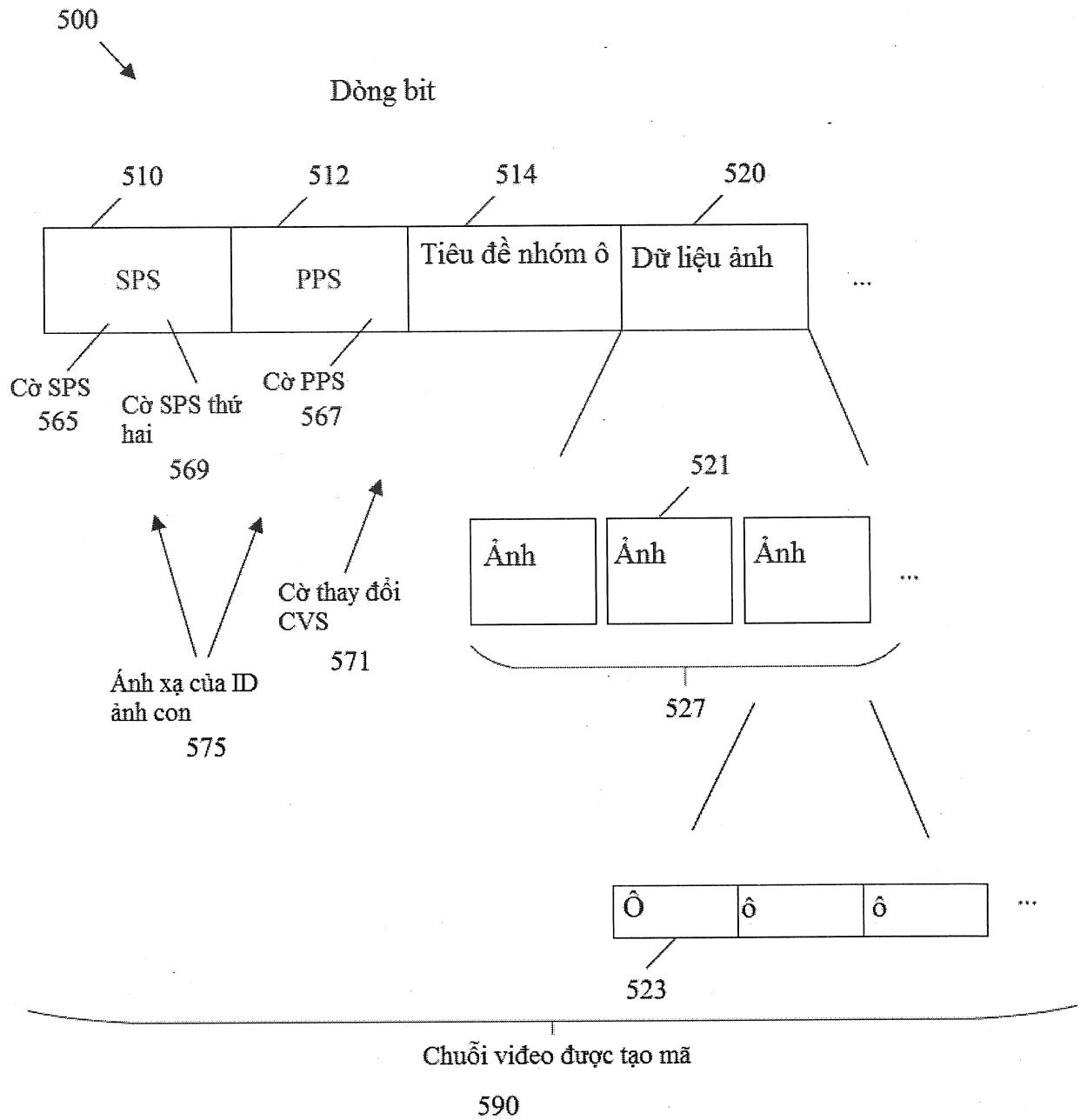


Fig.5

6/10

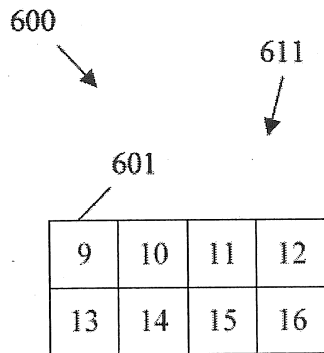


Fig. 6A

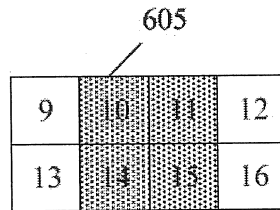


Fig. 6C

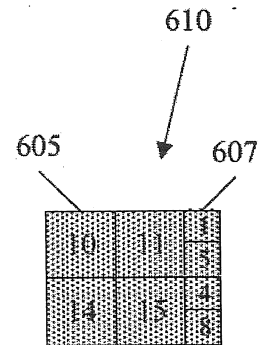


Fig. 6E

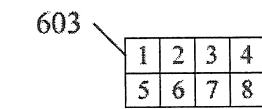


Fig. 6B

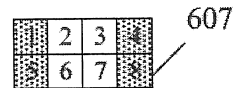


Fig. 6D

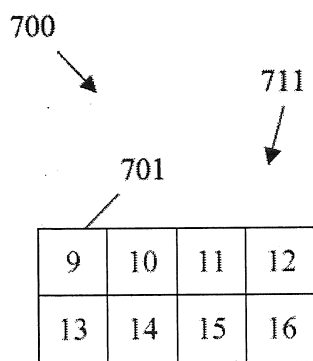


Fig. 7A

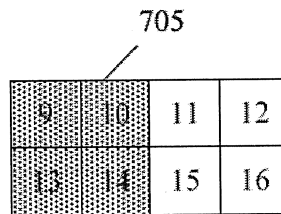


Fig. 7C

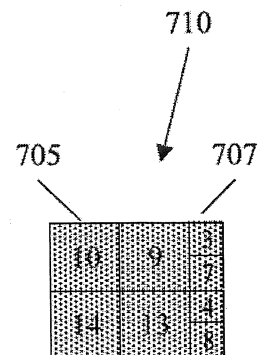


Fig. 7E

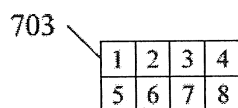


Fig. 7B

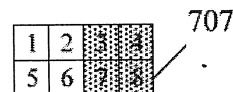


Fig. 7D

7/10

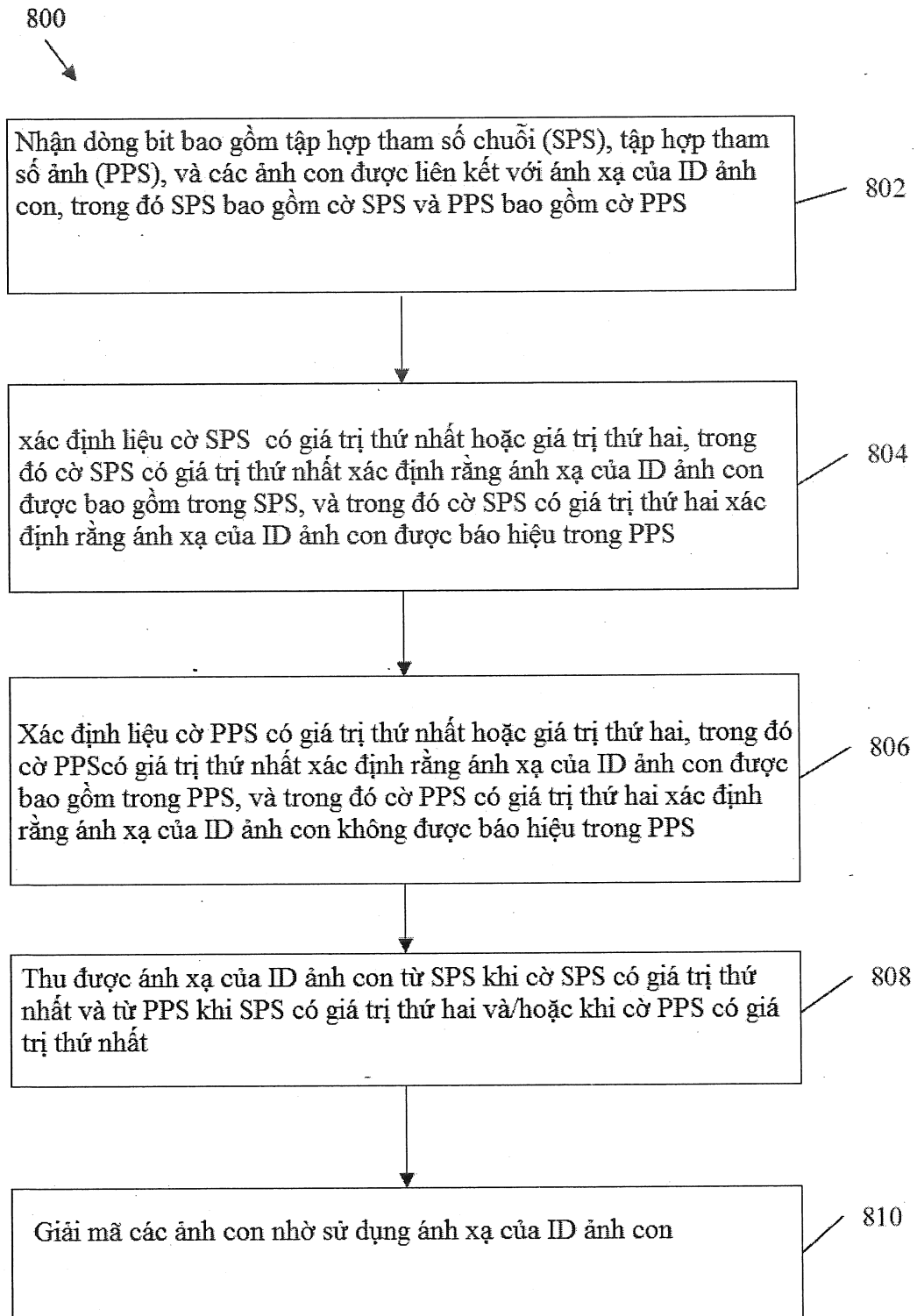


Fig.8

8/10

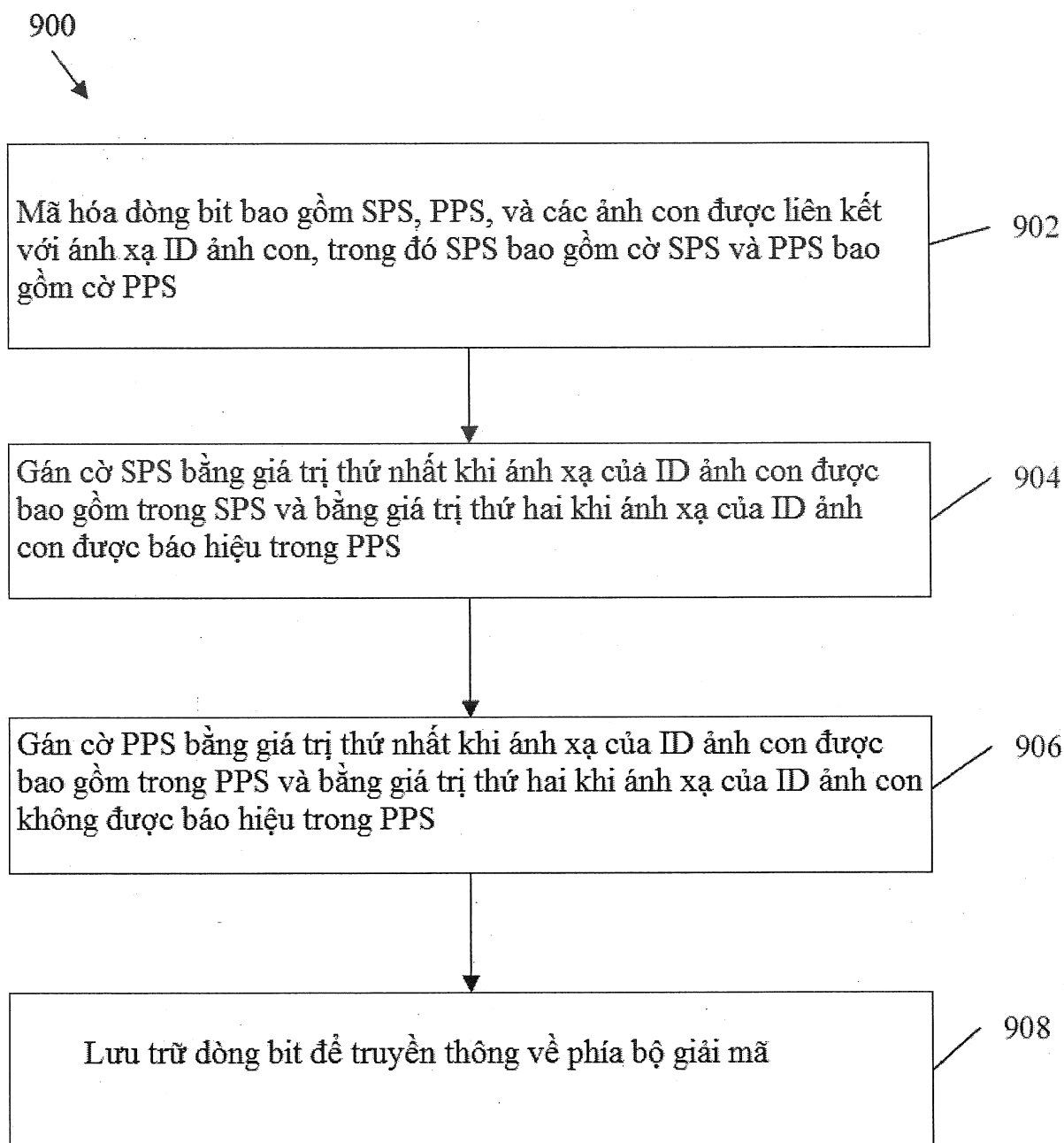


Fig.9

9/10

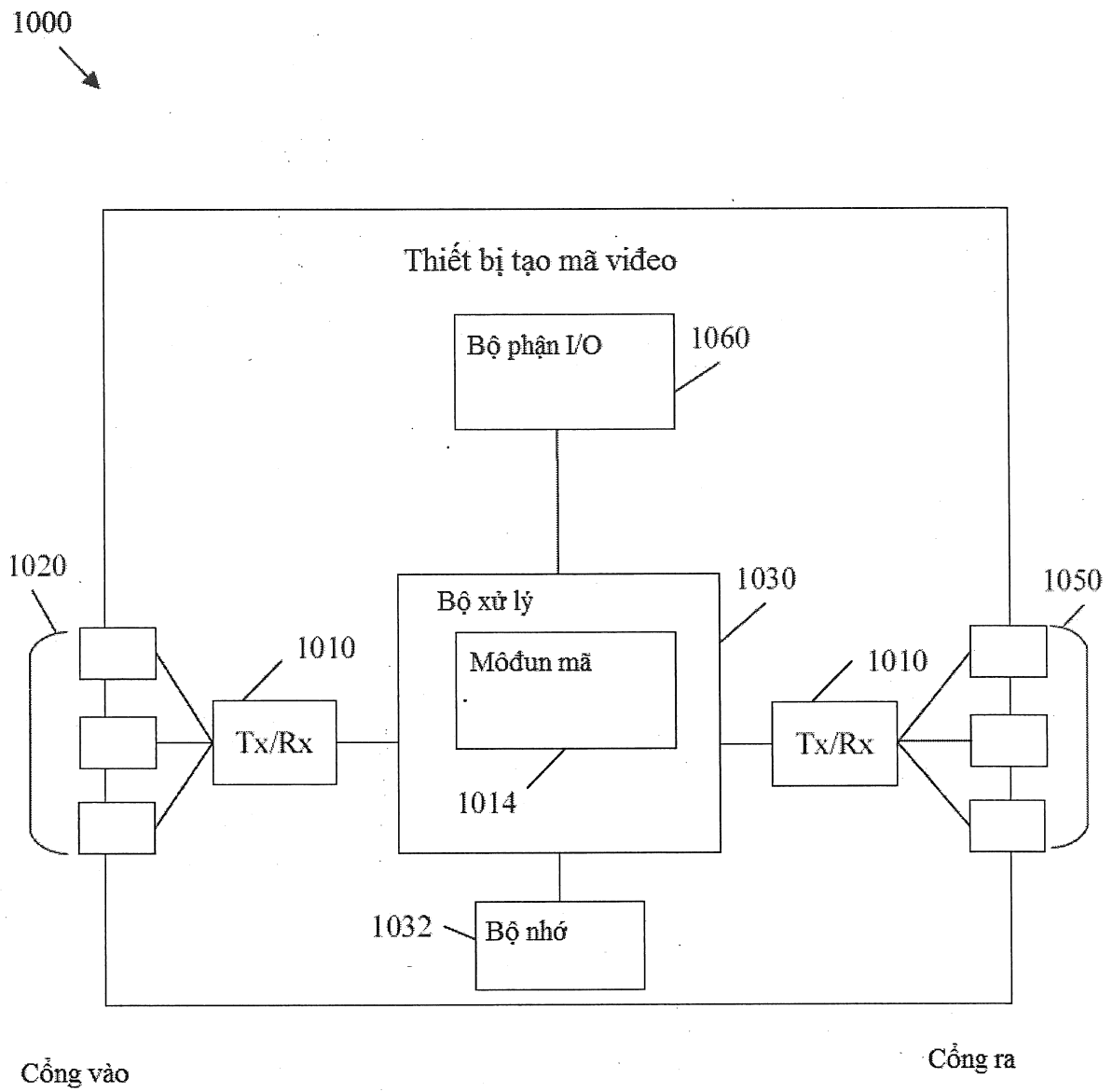


Fig.10

10/10

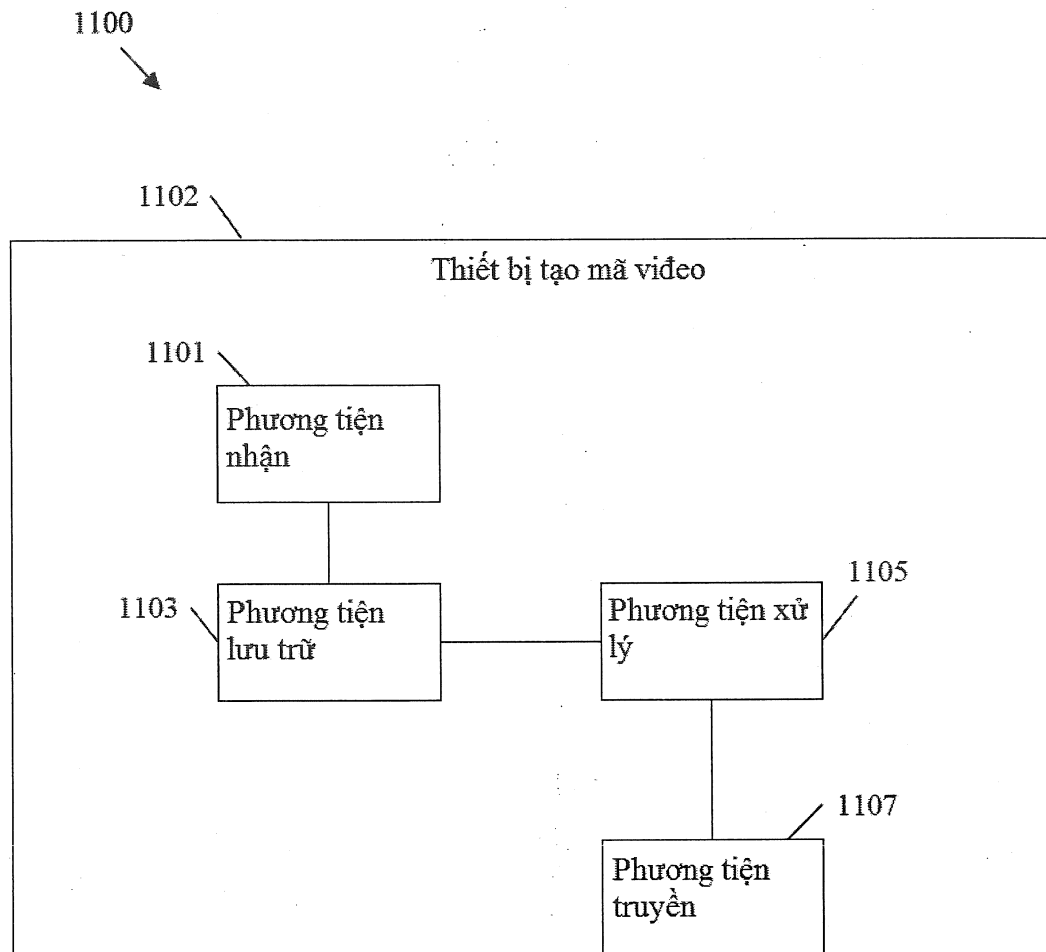


Fig.11