



- (12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ
(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN) (11) 
CỤC SỞ HỮU TRÍ TUỆ 1-0050312
(51)^{2020.01} H04N 19/44; H04N 19/33; H04N 19/136; H04N 19/30 (13) B

-
- (21) 1-2022-02313 (22) 15/09/2020
(86) PCT/US2020/050905 15/09/2020 (87) WO 2021/061454 01/04/2021
(30) 62/905,136 24/09/2019 US
(45) 25/08/2025 449 (43) 25/07/2022 412A
(73) HUAWEI TECHNOLOGIES CO., LTD. (CN)
Huawei Administration Building, Bantian, Longgang District, Shenzhen, Guangdong
518129, P. R. China
(72) WANG, Ye-Kui (US).
(74) Công ty Luật TNHH WINCO (WINCO LAW FIRM)
-

- (54) THIẾT BỊ GIẢI MÃ, PHƯƠNG PHÁP ĐỀ GIẢI MÃ ĐƯỢC THỰC HIỆN BỞI BỘ
GIẢI MÃ VIDEO, THIẾT BỊ LẬP MÃ, PHƯƠNG PHÁP ĐỂ LẬP MÃ ĐƯỢC
THỰC HIỆN BỞI BỘ LẬP MÃ VIDEO, THIẾT BỊ MÃ HÓA, HỆ THỐNG MÃ
HÓA VÀ PHƯƠNG TIỆN ĐỀ MÃ HÓA

(21) 1-2022-02313

(57) Sáng chế đề cập đến thiết bị giải mã, phương pháp để giải mã được thực hiện bởi bộ giải mã video, thiết bị lập mã, phương pháp để lập mã được thực hiện bởi bộ lập mã video, thiết bị mã hóa, hệ thống mã hóa và phương tiện để mã hóa. Phương pháp để giải mã được thực hiện bởi bộ giải mã video bao gồm các bước: thu, bởi bộ giải mã video, luồng bit video bao gồm tập thông số video (video parameter set, VPS) và nhiều lớp, trong đó không có lớp nào không phải là lớp đầu ra của ít nhất một tập lớp đầu ra (output layer set, OLS) mà cũng không phải là lớp tham chiếu trực tiếp của lớp bất kỳ khác, trong đó VPS bao gồm lớp được sử dụng làm cờ tham chiếu (LayerUsedAsRefLayerFlag) và lớp được sử dụng làm cờ lớp đầu ra (LayerUsedAsOutputLayerFlag), và trong đó giá trị của LayerUsedAsRefLayerFlag và giá trị của LayerUsedAsOutputLayerFlag đều không phải là không; và giải mã, bởi bộ giải mã video, hình từ một trong số nhiều lớp.

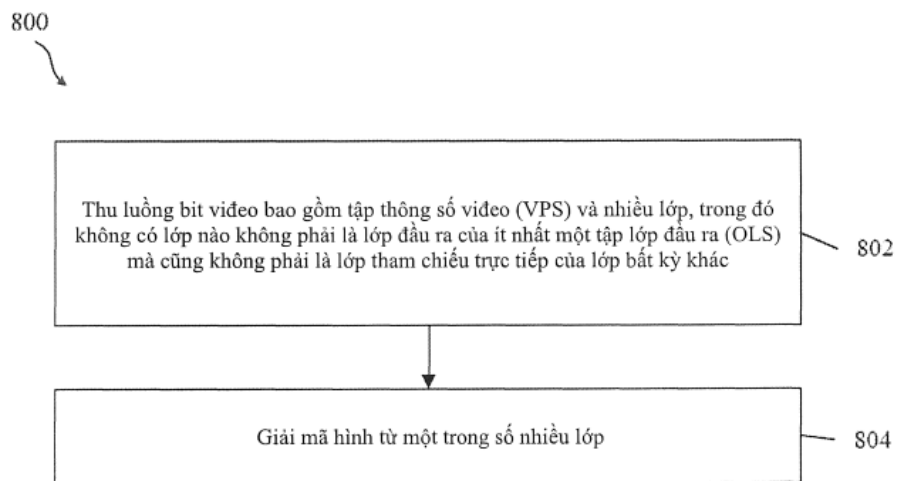


Fig.8

Lĩnh vực kỹ thuật được đề cập

Nói chung, sáng chế đề cập đến các kỹ thuật dành cho các luồng bit video nhiều lớp khi mã hóa video. Cụ thể hơn là, sáng chế đề cập đến thiết bị giải mã, phương pháp để giải mã được thực hiện bởi bộ giải mã video, thiết bị lập mã, phương pháp để lập mã được thực hiện bởi bộ lập mã video, thiết bị mã hóa, hệ thống mã hóa và phương tiện để mã hóa để đảm bảo rằng các lớp không cần thiết và/hoặc không được sử dụng không được cho phép trong luồng bit nhiều lớp khi mã hóa video.

Tình trạng kỹ thuật của sáng chế

Lượng dữ liệu video cần để thể hiện ngay cả với video tương đối ngắn có thể là đáng kể, việc này có thể dẫn đến các khó khăn khi dữ liệu này được truyền theo luồng hoặc truyền thông qua mạng truyền thông với dung lượng băng thông hạn chế. Do đó, dữ liệu video thường được nén trước khi được truyền thông qua các mạng viễn thông hiện đại ngày nay. Kích thước của video cũng có thể là vấn đề khi video được lưu trên thiết bị lưu trữ vì các tài nguyên bộ nhớ có thể bị hạn chế. Các thiết bị nén video thường sử dụng phần mềm và/hoặc phần cứng tại nguồn để mã hóa dữ liệu video trước khi truyền hoặc lưu trữ, do đó làm giảm lượng dữ liệu cần để thể hiện các ảnh video kỹ thuật số. Sau đó dữ liệu được nén được thu tại đích đến bởi thiết bị giải nén video mà giải mã dữ liệu video này. Với các tài nguyên mạng hạn chế và nhu cầu về chất lượng video cao hơn ngày càng tăng, thì mong muốn các kỹ thuật nén và giải nén được cải tiến cải thiện tỷ số nén mà hầu như không hề có sự hy sinh về chất lượng ảnh.

Bản chất kỹ thuật của sáng chế

Khía cạnh thứ nhất liên quan đến phương pháp để giải mã được thực hiện bởi bộ giải mã video, bao gồm các bước: thu, bởi bộ giải mã video, luồng bit video bao gồm tập thông số video (video parameter set, VPS) và nhiều lớp, trong đó không có lớp nào không phải là lớp đầu ra của ít nhất một tập lớp đầu ra (output layer set, OLS) mà cũng không phải là lớp tham

chiếu trực tiếp của lớp bất kỳ khác; và giải mã, bởi bộ giải mã video, hình từ một trong số nhiều lớp.

Phương pháp này cung cấp các kỹ thuật mà không cho phép các lớp không cần thiết trong các luồng bit video nhiều lớp. Tức là, mọi lớp được chứa trong tập lớp đầu ra (OLS) hoặc là lớp đầu ra hoặc được sử dụng như lớp tham chiếu trực tiếp hoặc gián tiếp của lớp đầu ra. Việc này tránh có thông tin ngoại lai trong quy trình mã hóa và làm tăng hiệu quả mã hóa. Do đó, bộ mã hóa/bộ giải mã (còn được gọi là, “bộ mã hóa/giải mã”) khi mã hóa video được cải thiện so với các bộ mã hóa/giải mã hiện thời. Một vấn đề thực tế, quy trình mã hóa video được cải thiện đem lại cho người dùng trải nghiệm người dùng tốt hơn khi các video được gửi, được thu, và/hoặc được xem.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh được đề cập trước đó, dạng thực hiện khác của khía cạnh này đề xuất rằng hình được chứa trong lớp đầu ra của ít nhất một OLS.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh được đề cập trước đó, dạng thực hiện khác của khía cạnh này đề xuất việc chọn lựa lớp đầu ra từ ít nhất một OLS trước khi giải mã.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh được đề cập trước đó, dạng thực hiện khác của khía cạnh này đề xuất việc chọn lựa hình từ lớp đầu ra sau khi lớp đầu ra được chọn lựa.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh được đề cập trước đó, dạng thực hiện khác của khía cạnh này đề xuất rằng mỗi lớp trong nhiều lớp bao gồm tập các đơn vị lớp trừu tượng mạng (network abstraction layer, NAL) lớp mã hóa video (video coding layer, VCL) mà đều có giá trị cụ thể của bộ nhận dạng (identifier, ID) lớp và các đơn vị NAL không phải VCL được liên kết

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh được đề cập trước đó, dạng thực hiện khác của khía cạnh này đề xuất rằng mỗi lớp của ít nhất một OLS bao gồm lớp đầu ra hoặc lớp tham chiếu trực tiếp của lớp bất kỳ khác.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh được đề cập trước đó, dạng thực hiện khác của khía cạnh này đề xuất rằng ít nhất một OLS chứa một hoặc nhiều lớp đầu ra.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh được đề cập trước đó, dạng thực hiện khác của khía cạnh này đề xuất rằng VPS bao gồm lớp được sử dụng làm cờ tham chiếu và lớp được sử dụng làm cờ lớp đầu ra, và trong đó giá trị của lớp được sử dụng làm cờ tham chiếu và giá trị của lớp được sử dụng làm cờ lớp đầu ra đều không phải là không.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh được đề cập trước đó, dạng thực hiện khác của khía cạnh này đề xuất việc hiển thị hình được giải mã trên bộ hiển thị của thiết bị điện tử.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh được đề cập trước đó, dạng thực hiện khác của khía cạnh này đề xuất việc thu, bởi bộ giải mã video, luồng bit video thứ hai bao gồm tập thông số video (VPS) thứ hai và nhiều lớp thứ hai, trong đó ít nhất một lớp không phải là lớp đầu ra của ít nhất một tập lớp đầu ra (OLS) mà cũng không phải là lớp tham chiếu trực tiếp của lớp bất kỳ khác; và nhằm đáp lại việc thu này, thì thực hiện một số biện pháp hiệu chỉnh khác để đảm bảo rằng luồng bit phù hợp tương ứng với luồng bit video thứ hai được thu trước khi giải mã hình từ một trong số nhiều lớp thứ hai.

Khía cạnh thứ hai liên quan đến phương pháp để lập mã được thực hiện bởi bộ lập mã video, phương pháp này bao gồm các bước: tạo ra, bởi bộ lập mã video, nhiều lớp và tập thông số video (VPS) quy định ít nhất một tập lớp đầu ra (OLS), trong đó bộ lập mã video bị ràng buộc sao cho không có lớp là không phải là lớp đầu ra của ít nhất một OLS mà cũng không phải là lớp tham chiếu trực tiếp của lớp bất kỳ khác; lập mã, bởi bộ lập mã video, nhiều lớp và VPS thành luồng bit video; và lưu, bởi bộ lập mã video, luồng bit video để truyền thông với bộ giải mã video.

Phương pháp này cung cấp các kỹ thuật mà không cho phép các lớp không cần thiết trong các luồng bit video nhiều lớp. Tức là, mọi lớp được chứa trong tập lớp đầu ra (OLS) hoặc là lớp đầu ra hoặc được sử dụng như lớp tham chiếu trực tiếp hoặc gián tiếp của lớp đầu ra. Việc này tránh có thông tin ngoại lai trong quy trình mã hóa và làm tăng hiệu quả mã hóa. Do đó, bộ mã hóa/bộ giải mã (còn được gọi là, “bộ mã hóa/giải mã”) khi mã hóa video được cải thiện so với các bộ mã hóa/giải mã hiện thời. Một vấn đề thực tế, quy trình mã hóa video được cải thiện đem lại cho người dùng trải nghiệm người dùng tốt hơn khi các video được gửi, được thu, và/hoặc được xem.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh được đề cập trước đó, dạng

thực hiện khác của khía cạnh này đề xuất rằng mỗi OLS trong số một hoặc nhiều OLS chứa một hoặc nhiều lớp đầu ra, và trong đó mỗi lớp đầu ra trong số các lớp đầu ra bao gồm một hoặc nhiều hình.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh được đề cập trước đó, dạng thực hiện khác của khía cạnh này đề xuất rằng mỗi lớp trong nhiều lớp bao gồm tập các đơn vị lớp trừu tượng mạng (NAL) lớp mã hóa video (VCL) mà đều có giá trị cụ thể của bộ nhận dạng (ID) lớp và các đơn vị NAL không phải VCL được liên kết

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh được đề cập trước đó, dạng thực hiện khác của khía cạnh này đề xuất rằng một trong số các OLS bao gồm hai lớp đầu ra, và trong đó một trong số hai lớp đầu ra tham chiếu tới lớp khác trong số hai lớp đầu ra.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh được đề cập trước đó, dạng thực hiện khác của khía cạnh này đề xuất rằng VPS bao gồm lớp được sử dụng làm cờ tham chiếu và lớp được sử dụng làm cờ lớp đầu ra, và trong đó giá trị của lớp được sử dụng làm cờ tham chiếu và giá trị của lớp được sử dụng làm cờ lớp đầu ra đều không phải là không.

Khía cạnh thứ ba liên quan đến thiết bị giải mã, bao gồm bộ thu được tạo cấu hình để thu luồng bit video bao gồm tập thông số video (VPS) và nhiều lớp, trong đó không có lớp nào không phải là lớp đầu ra của ít nhất một OLS mà cũng không phải là lớp tham chiếu trực tiếp của lớp bất kỳ khác; bộ nhớ được ghép nối với bộ thu, bộ nhớ lưu 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 làm cho thiết bị giải mã giải mã hình từ một trong số nhiều lớp để nhận hình được giải mã.

Thiết bị giải mã cung cấp các kỹ thuật mà không cho phép các lớp không cần thiết trong các luồng bit video nhiều lớp. Tức là, mọi lớp được chứa trong tập lớp đầu ra (OLS) hoặc là lớp đầu ra hoặc được sử dụng như lớp tham chiếu trực tiếp hoặc gián tiếp của lớp đầu ra. Việc này tránh có thông tin ngoại lai trong quy trình mã hóa và làm tăng hiệu quả mã hóa. Do đó, bộ mã hóa/bộ giải mã (còn được gọi là, “bộ mã hóa/giải mã”) khi mã hóa video được cải thiện so với các bộ mã hóa/giải mã hiện thời. Một vấn đề thực tế, quy trình mã hóa video được cải thiện đem lại cho người dùng trải nghiệm người dùng tốt hơn khi các video được gửi, được thu, và/hoặc được xem.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh được đề cập trước đó, dạng thực hiện khác của khía cạnh này đề xuất rằng hình được chứa trong lớp đầu ra của ít nhất một

OLS.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh được đề cập trước đó, dạng thực hiện khác của khía cạnh này đề xuất rằng bộ xử lý còn được tạo cấu hình để chọn lựa lớp đầu ra từ ít nhất một OLS trước khi hình được giải mã.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh được đề cập trước đó, dạng thực hiện khác của khía cạnh này đề xuất rằng bộ xử lý còn được tạo cấu hình để chọn lựa hình từ lớp đầu ra sau khi lớp đầu ra đã được chọn lựa.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh được đề cập trước đó, dạng thực hiện khác của khía cạnh này đề xuất rằng mỗi lớp trong nhiều lớp bao gồm tập các đơn vị lớp trừu tượng mạng (NAL) lớp mã hóa video (VCL) mà đều có giá trị cụ thể của bộ nhận dạng (ID) lớp và các đơn vị NAL không phải VCL được liên kết

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh được đề cập trước đó, dạng thực hiện khác của khía cạnh này đề xuất rằng mỗi lớp của ít nhất một OLS bao gồm lớp đầu ra hoặc lớp tham chiếu trực tiếp của lớp bất kỳ khác.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh được đề cập trước đó, dạng thực hiện khác của khía cạnh này đề xuất rằng ít nhất một OLS chứa một hoặc nhiều lớp đầu ra.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh được đề cập trước đó, dạng thực hiện khác của khía cạnh này đề xuất rằng bộ hiển thị được tạo cấu hình để hiển thị hình được giải mã.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh được đề cập trước đó, dạng thực hiện khác của khía cạnh này đề xuất rằng bộ xử lý được tạo cấu hình để thực thi các lệnh còn làm cho thiết bị giải mã thu luồng bit video thứ hai bao gồm tập thông số video (VPS) thứ hai và nhiều lớp thứ hai, trong đó ít nhất một lớp không phải là lớp đầu ra của ít nhất một tập lớp đầu ra (OLS) mà cũng không phải là lớp tham chiếu trực tiếp của lớp bất kỳ khác; và nhằm đáp lại việc thu luồng bit video thứ hai, thì thực hiện một số biện pháp hiệu chỉnh khác để đảm bảo rằng luồng bit phù hợp tương ứng với luồng bit video thứ hai được thu trước khi giải mã hình từ một trong số nhiều lớp thứ hai.

Khía cạnh thứ tư liên quan đến thiết bị lập mã, 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 làm cho thiết

bị lập mã để tạo ra nhiều lớp và tập thông số video (VPS) quy định một hoặc nhiều tập lớp đầu ra (OLS), trong đó thiết bị lập mã bị ràng buộc sao cho không có lớp nào không phải là lớp đầu ra của ít nhất một OLS mà cũng không phải là lớp tham chiếu trực tiếp của lớp bất kỳ khác; lập mã nhiều lớp và VPS thành luồng bit video; và bộ truyền được ghép nối với bộ xử lý, bộ truyền được tạo cấu hình để truyền luồng bit video về phía bộ giải mã video.

Thiết bị lập mã cung cấp các kỹ thuật mà không cho phép các lớp không cần thiết trong các luồng bit video nhiều lớp. Tức là, mọi lớp được chứa trong tập lớp đầu ra (OLS) hoặc là lớp đầu ra hoặc được sử dụng như lớp tham chiếu trực tiếp hoặc gián tiếp của lớp đầu ra. Việc này tránh có thông tin ngoại lai trong quy trình mã hóa và làm tăng hiệu quả mã hóa. Do đó, bộ mã hóa/bộ giải mã (còn được gọi là, “bộ mã hóa/giải mã”) khi mã hóa video được cải thiện so với các bộ mã hóa/giải mã hiện thời. Một vấn đề thực tế, quy trình mã hóa video được cải thiện đem lại cho người dùng trải nghiệm người dùng tốt hơn khi các video được gửi, được thu, và/hoặc được xem.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh được đề cập trước đó, dạng thực hiện khác của khía cạnh này đề xuất rằng mỗi OLS trong số một hoặc nhiều OLS chứa một hoặc nhiều lớp đầu ra, và trong đó mỗi lớp đầu ra trong số các lớp đầu ra bao gồm một hoặc nhiều hình.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh được đề cập trước đó, dạng thực hiện khác của khía cạnh này đề xuất rằng mỗi lớp trong nhiều lớp bao gồm tập các đơn vị lớp trừu tượng mạng (NAL) lớp mã hóa video (VCL) mà đều có giá trị cụ thể của bộ nhận dạng (ID) lớp và các đơn vị NAL không phải VCL được liên kết

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh được đề cập trước đó, dạng thực hiện khác của khía cạnh này đề xuất rằng mỗi lớp của ít nhất một OLS bao gồm lớp đầu ra hoặc lớp tham chiếu trực tiếp của lớp bất kỳ khác.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh được đề cập trước đó, dạng thực hiện khác của khía cạnh này đề xuất rằng VPS bao gồm lớp được sử dụng làm cờ tham chiếu và lớp được sử dụng làm cờ lớp đầu ra, và trong đó giá trị của lớp được sử dụng làm cờ tham chiếu và giá trị của lớp được sử dụng làm cờ lớp đầu ra đều không phải là không.

Khía cạnh thứ năm liên quan đến thiết bị mã hóa. Thiết bị mã hóa bao gồm bộ thu được tạo cấu hình để thu hình để lập mã hoặc để thu luồng bit để giải mã; bộ truyền được ghép nối

với bộ thu, bộ truyền được tạo cấu hình để truyền luồng bit tới bộ giải mã hoặc để truyền ảnh được giải mã tới bộ hiển thị; bộ nhớ được ghép nối với ít nhất một trong số bộ thu hoặc bộ truyền, bộ nhớ được tạo cấu hình để lưu 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 trong bộ nhớ để thực hiện phương pháp bất kỳ trong số các phương pháp được bộc lộ ở đây.

Thiết bị mã hóa cung cấp các kỹ thuật mà không cho phép các lớp không cần thiết trong các luồng bit video nhiều lớp. Tức là, mọi lớp được chứa trong tập lớp đầu ra (OLS) hoặc là lớp đầu ra hoặc được sử dụng như lớp tham chiếu trực tiếp hoặc gián tiếp của lớp đầu ra. Việc này tránh có thông tin ngoại lai trong quy trình mã hóa và làm tăng hiệu quả mã hóa. Do đó, bộ mã hóa/bộ giải mã (còn được gọi là, “bộ mã hóa/giải mã”) khi mã hóa video được cải thiện so với các bộ mã hóa/giải mã hiện thời. Một vấn đề thực tế, quy trình mã hóa video được cải thiện đem lại cho người dùng trải nghiệm người dùng tốt hơn khi các video được gửi, được thu, và/hoặc được xem.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh được đề cập trước đó, dạng thực hiện khác của khía cạnh này đề xuất bộ hiển thị được tạo cấu hình để hiển thị hình được giải mã.

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

Hệ thống cung cấp các kỹ thuật không cho phép các lớp không cần thiết trong các luồng bit video nhiều lớp. Tức là, mọi lớp được chứa trong tập lớp đầu ra (OLS) hoặc là lớp đầu ra hoặc được sử dụng như lớp tham chiếu trực tiếp hoặc gián tiếp của lớp đầu ra. Việc này tránh có thông tin ngoại lai trong quy trình mã hóa và làm tăng hiệu quả mã hóa. Do đó, bộ mã hóa/bộ giải mã (còn được gọi là, “bộ mã hóa/giải mã”) khi mã hóa video được cải thiện so với các bộ mã hóa/giải mã hiện thời. Một vấn đề thực tế, quy trình mã hóa video được cải thiện đem lại cho người dùng trải nghiệm người dùng tốt hơn khi các video được gửi, được thu, và/hoặc được xem.

Khía cạnh thứ bảy liên quan đến phương tiện để mã hóa. Phương tiện để mã hóa bao gồm phương tiện thu được tạo cấu hình để thu hình để lập mã hoặc để thu luồng bit để giải mã; phương tiện truyền dẫn được ghép nối với phương tiện thu, phương tiện truyền dẫn được

tạo cấu hình để truyền luồng bit tới phương tiện giải mã hoặc để truyền ảnh được giải mã tới phương tiện hiển thị; phương tiện lưu trữ được ghép nối với ít nhất một trong số phương tiện thu hoặc phương tiện truyền dẫn, phương tiện lưu trữ được tạo cấu hình để lưu 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 trong phương tiện lưu trữ để thực hiện phương pháp bất kỳ trong số các phương pháp được bộc lộ ở đây.

Phương tiện để mã hóa cung cấp các kỹ thuật mà không cho phép các lớp không cần thiết trong các luồng bit video nhiều lớp. Tức là, mọi lớp được chứa trong tập lớp đầu ra (OLS) hoặc là lớp đầu ra hoặc được sử dụng như lớp tham chiếu trực tiếp hoặc gián tiếp của lớp đầu ra. Việc này tránh có thông tin ngoại lai trong quy trình mã hóa và làm tăng hiệu quả mã hóa. Do đó, bộ mã hóa/bộ giải mã (còn được gọi là, “bộ mã hóa/giải mã”) khi mã hóa video được cải thiện so với các bộ mã hóa/giải mã hiện thời. Một vấn đề thực tế, quy trình mã hóa video được cải thiện đem lại cho người dùng trải nghiệm người dùng tốt hơn khi các video được gửi, được thu, và/hoặc được xem.

Với mục đích rõ ràng, phương án bất kỳ trong số các phương án được đề cập ở trên có thể được kết hợp với một hoặc nhiều phương án bất kỳ trong số các phương án khác được đề cập ở trên để tạo nên phương án mới, và các phương án như vậy đều nằm trong phạm vi của sáng chế.

Các dấu hiệu này và các dấu hiệu khác sẽ được hiểu rõ ràng hơn từ phần mô tả chi tiết sau đây kết hợp với các hình vẽ và các điểm yêu cầu bảo hộ kèm theo.

Mô tả vắn tắt các hình vẽ

Để hiểu rõ hơn về sáng chế, thì phần tham khảo được thực hiện theo phần mô tả ngắn gọn sau đây kết hợp với các hình vẽ kèm theo và phần mô tả chi tiết, trong đó các số chỉ dẫn giống nhau sử dụng để chỉ các bộ phận giống nhau.

Fig.1 là hình vẽ lưu đồ của phương pháp ví dụ để mã hóa tín hiệu video.

Fig.2 là hình vẽ sơ đồ của hệ thống mã hóa và giải mã (mã hóa/giải mã) ví dụ để mã hóa video.

Fig.3 là hình vẽ sơ đồ thể hiện bộ lập mã video ví dụ.

Fig.4 là hình vẽ sơ đồ thể hiện bộ giải mã video ví dụ.

Fig.5 là hình vẽ thể hiện một ví dụ về việc mã hóa nhiều lớp dành cho khả năng mở rộng

không gian.

Fig.6 là hình vẽ thể hiện việc mã hóa nhiều lớp sử dụng các tập lớp đầu ra (OLS).

Fig.7 là hình vẽ thể hiện một phương án về luồng bit video.

Fig.8 là một phương án của phương pháp để giải mã luồng bit video được mã hóa.

Fig.9 là một phương án của phương pháp để lập mã luồng bit video được mã hóa.

Fig.10 là hình vẽ sơ đồ của thiết bị mã hóa video.

Fig.11 là hình vẽ sơ đồ của một phương án về phương tiện để mã hóa.

Mô tả chi tiết sáng chế

Phải hiểu ngay từ đầu rằng mặc dù dạng thực hiện làm ví dụ của một hoặc nhiều phương án được cung cấp bên dưới, nhưng các hệ thống và/hoặc phương pháp được bộc lộ có thể được thực hiện bằng cách sử dụng số lượng kỹ thuật bất kỳ, cho dù các kỹ thuật đó hiện thời đã được biết đến hoặc đã tồn tại hay không. Sáng chế không bị hạn chế bởi các dạng thực hiện làm ví dụ, hình vẽ, và kỹ thuật được thể hiện bên dưới, bao gồm các phương án và dạng thực hiện làm ví dụ được minh họa và mô tả ở đây, nhưng có thể được cải biến trong phạm vi của các điểm yêu cầu bảo hộ kèm theo cùng với phạm vi đầy đủ của các phương án tương đương.

Các thuật ngữ sau đây được định nghĩa như sau trừ trường hợp được sử dụng trong ngữ cảnh ngược lại ở đây. Cụ thể là, các phần định nghĩa sau đây chỉ nhằm để cung cấp thêm sự rõ ràng cho sáng chế. Tuy nhiên, các thuật ngữ này có thể được mô tả khác nhau trong các ngữ cảnh khác nhau. Do đó, các phần định nghĩa sau đây chỉ được coi là các phần bổ sung và không được coi là để hạn chế các phần định nghĩa bất kỳ khác của các phần mô tả được cung cấp cho các thuật ngữ đó ở đây.

Luồng bit là một trình tự của các bit bao gồm dữ liệu video mà được nén dành cho việc truyền dẫn giữa bộ lập mã và bộ giải mã. Bộ lập mã là thiết bị mà được tạo cấu hình để dùng các quy trình lập mã để nén dữ liệu video thành luồng bit. Bộ giải mã là thiết bị mà được tạo cấu hình để dùng các quy trình giải mã để cấu thành lại dữ liệu video từ luồng bit để hiển thị. Hình là mảng các mẫu độ sáng và/hoặc mảng các mẫu sắc độ mà tạo lập khung hoặc trường của nó. Để rõ ràng, thì hình mà được lập mã hoặc giải mã có thể được gọi là hình hiện thời. Hình tham chiếu là hình mà chứa các mẫu tham chiếu mà có thể được sử dụng khi mã hóa các hình khác bằng cách tham chiếu theo dự báo liên cấu trúc và/hoặc dự báo liên lớp. Danh sách hình tham chiếu là danh sách của các hình tham chiếu để dự báo liên cấu trúc và/hoặc dự báo

liên lớp. Một số hệ thống mã hóa video sử dụng hai danh sách hình tham chiếu, mà có thể được ký hiệu là danh sách hình tham chiếu một và danh sách hình tham chiếu không. Cấu trúc danh sách hình tham chiếu là cấu trúc cú pháp có thể định địa chỉ mà chứa nhiều danh sách hình tham chiếu. Dự báo liên cấu trúc là cơ chế để mã hóa các mẫu của hình hiện thời bằng cách tham chiếu tới các mẫu được chỉ báo trong hình tham chiếu mà khác với hình hiện thời trong đó hình tham chiếu và hình hiện thời ở trong cùng một lớp. Mục nhập cấu trúc danh sách hình tham chiếu là vị trí có thể định địa chỉ trong cấu trúc danh sách hình tham chiếu mà chỉ báo hình tham chiếu được liên kết với danh sách hình tham chiếu. Phần đầu phần chia là một phần của phần chia được mã hóa chứa các phần tử dữ liệu liên quan đến tất cả dữ liệu video trong khung lát được thể hiện trong phần chia. Tập thông số hình (picture parameter set, PPS) là tập thông số mà chứa dữ liệu liên quan đến toàn bộ hình. Cụ thể hơn là, PPS là cấu trúc cú pháp chứa các phần tử cú pháp mà áp dụng cho không hoặc nhiều hình được mã hóa toàn bộ như được xác định bởi phần tử cú pháp được tìm thấy trong mỗi phần đầu hình. Tập thông số trình tự (sequence parameter set, SPS) là tập thông số mà chứa dữ liệu liên quan đến trình tự của các hình. Đơn vị truy nhập (access unit, AU) là tập của một hoặc nhiều hình được mã hóa được liên kết với cùng một thời gian hiển thị (ví dụ, cùng một giá trị đếm thứ bậc hình) để xuất ra từ bộ đệm hình được giải mã (decoded picture buffer, DPB) (ví dụ, để hiển thị cho người dùng). Bộ phân cách đơn vị truy nhập (access unit delimiter, AUD) là bộ chỉ báo hoặc cấu trúc dữ liệu được sử dụng để chỉ báo điểm bắt đầu của AU hoặc biên giữa các AU. Trình tự video được giải mã là trình tự của các hình mà đã được cấu thành lại bởi bộ giải mã để chuẩn bị hiển thị cho người dùng.

Đơn vị lớp trừu tượng mạng (NAL) là cấu trúc cú pháp chứa dữ liệu dưới dạng tải trình tự byte thô (raw byte sequence payload, RBSP), thông tin chỉ báo về kiểu dữ liệu, và được đặt rải rác như mong muốn với các byte ngăn chặn mô phỏng. Đơn vị NAL lớp mã hóa video (VCL) là đơn vị NAL được mã hóa để chứa dữ liệu video, chẳng hạn như phần chia được mã hóa của hình. Đơn vị NAL không phải VCL là đơn vị NAL mà chứa dữ liệu không phải video chẳng hạn như cú pháp và/hoặc các thông số mà hỗ trợ việc giải mã dữ liệu video, hiệu năng để kiểm tra sự phù hợp, hoặc các hoạt động khác. Lớp là tập các đơn vị VCL NAL mà chia sẻ đặc tính được quy định (ví dụ, độ phân giải chung, tỷ lệ khung, kích thước ảnh, v.v.) và các đơn vị NAL không phải VCL được liên kết. Các đơn vị VCL NAL của lớp có thể chia sẻ giá

trị cụ thể của bộ nhận dạng lớp phần đầu đơn vị NAL (`nuh_layer_id`). Hình được mã hóa là dạng biểu diễn được mã hóa của hình bao gồm các đơn vị VCL NAL với giá trị cụ thể của bộ nhận dạng lớp phần đầu đơn vị NAL (`nuh_layer_id`) trong đơn vị truy nhập (AU) và chứa tất cả các đơn vị cây mã hóa (coding tree unit, CTU) của hình. Hình được giải mã là hình được tạo ra bằng cách áp dụng quy trình giải mã cho hình được mã hóa.

Tập lớp đầu ra (OLS) là tập các lớp mà một hoặc nhiều lớp được quy định như (các) lớp đầu ra. Lớp đầu ra là lớp mà được chỉ định dành cho đầu ra (ví dụ, cho bộ hiển thị). OLS thứ 0 là OLS mà chỉ chứa lớp thấp nhất (lớp với bộ nhận dạng lớp thấp nhất) và do đó chỉ chứa lớp đầu ra. Tập thông số video (VPS) là đơn vị dữ liệu mà chứa các thông số liên quan đến toàn bộ video. Dữ liệu liên lớp là cơ chế để mã hóa hình hiện thời trong lớp hiện thời bằng cách tham chiếu tới hình tham chiếu trong lớp tham chiếu, trong đó hình hiện thời và hình tham chiếu được chứa trong cùng một AU và lớp tham chiếu bao gồm `nuh_layer_id` thấp hơn so với lớp hiện thời.

Các từ viết tắt sau đây được sử dụng, khối cây mã hóa (Coding Tree Block, CTB), đơn vị cây mã hóa (Coding Tree Unit, CTU), đơn vị mã hóa (Coding Unit, CU), trình tự video được mã hóa (Coded Video Sequence, CVS), nhóm chuyên gia video chung (Joint Video Experts Team, JVET), tập khung lát bị ràng buộc chuyển động (Motion-Constrained Tile Set, MCTS), đơn vị chuyển tối đa (Maximum Transfer Unit, MTU), lớp trừu tượng mạng (Network Abstraction Layer, NAL), giá trị đếm thứ bậc hình (Picture Order Count, POC), tập thông số hình (Picture Parameter Set, PPS), tải trình tự byte thô (Raw Byte Sequence Payload, RBSP), tập thông số trình tự (Sequence Parameter Set, SPS), mã hóa video đa năng (Versatile Video Coding, VVC), và bản nháp (Working Draft, WD).

Fig.1 là hình vẽ lưu đồ của phương pháp hoạt động ví dụ 100 để mã hóa tín hiệu video. Cụ thể là, tín hiệu video được lập mã tại bộ lập mã. Quy trình lập mã nén tín hiệu video bằng cách dùng các cơ chế khác nhau để làm giảm kích thước tập tin video. Kích thước tập tin nhỏ hơn cho phép tập tin video được nén được truyền tới người dùng, trong khi làm giảm mào đầu băng thông được liên kết. Bộ giải mã sau đó giải mã tập tin video được nén để cấu thành lại tín hiệu video gốc để hiển thị cho người dùng. Quy trình giải mã thường phản chiếu quy trình lập mã để cho phép bộ giải mã để cấu thành lại một cách nhất quán tín hiệu video.

Ở bước 101, tín hiệu video được đưa vào bộ lập mã. Ví dụ, tín hiệu video có thể là tập

tin video không được nén được lưu trong bộ nhớ. Ví dụ khác, tập tin video có thể được quay bằng thiết bị quay video, chẳng hạn như camera video, và được lập mã để hỗ trợ việc truyền theo luồng trực tiếp video. Tập tin video có thể bao gồm cả thành phần âm thanh và thành phần video. Thành phần video chứa một loạt khung ảnh mà, khi được quan sát theo trình tự, thì mang lại cảm giác nhìn thấy được chuyển động. Các khung chứa các điểm ảnh mà được thể hiện dưới dạng ánh sáng, ở đây có thể được gọi là các thành phần độ sáng (hoặc các mẫu độ sáng), và màu sắc, có thể được gọi là các thành phần sắc độ (hoặc các mẫu sắc độ). Theo một số ví dụ, các khung cũng có thể chứa các giá trị độ sâu để hỗ trợ việc quan sát ba chiều.

Ở bước 103, video được chia thành các khối. Việc phân chia bao gồm việc chia nhỏ các điểm ảnh trong mỗi khung thành các khối vuông và/hoặc chữ nhật để nén. Ví dụ, trong mã hóa video hiệu quả cao (High Efficiency Video Coding, HEVC) (còn được biết đến là H.265 và MPEG-H Part 2) khung trước hết có thể được chia thành các đơn vị cây mã hóa (CTU), mà là các khối có kích thước định trước (ví dụ, 64x64 điểm ảnh). Các CTU chứa cả hai mẫu độ sáng và sắc độ. Các cây mã hóa có thể được dùng để phân chia các CTU thành các khối và sau đó chia nhỏ đệ quy các khối này cho đến khi các cấu hình đạt được hỗ trợ hơn nữa việc lập mã. Ví dụ, các thành phần độ sáng của khung có thể được chia nhỏ cho đến khi các khối riêng chứa các giá trị độ sáng tương đối đồng đều. Hơn nữa, các thành phần sắc độ của khung có thể được chia nhỏ cho đến khi các khối riêng chứa các giá trị màu sắc tương đối đồng đều. Do đó, các cơ chế phân chia thay đổi tùy thuộc vào nội dung của các khung video.

Ở bước 105, các cơ chế nén khác nhau được dùng để nén các khối ảnh được phân chia ở bước 103. Ví dụ, dự báo liên cấu trúc và/hoặc dự báo nội cấu trúc có thể được dùng. Dự báo liên cấu trúc được thiết kế để tận dụng việc các đối tượng trong quang cảnh chung có xu hướng xuất hiện trong các khung liên tiếp. Do đó, khối thể hiện đối tượng trong khung tham chiếu không cần phải được mô tả lặp lại trong các khung gần kề. Cụ thể là, đối tượng, chẳng hạn như chiếc bàn, có thể vẫn ở một vị trí không đổi trong nhiều khung. Do đó chiếc bàn này được mô tả một lần và các khung gần kề có thể tham chiếu ngược khung tham chiếu này. Các cơ chế so khớp mẫu có thể được dùng để so khớp các đối tượng trong nhiều khung. Hơn nữa, các đối tượng di chuyển có thể được thể hiện trong nhiều khung, ví dụ do sự di chuyển của đối tượng hoặc sự di chuyển của camera. Ví dụ cụ thể, video có thể thể hiện xe ô tô di chuyển trong màn hình trong nhiều khung. Các vectơ chuyển động có thể được dùng để mô tả sự di

chuyển như vậy. Vector chuyển động là vector hai chiều mà cung cấp độ lệch từ các tọa độ của đối tượng trong khung tới các tọa độ của đối tượng trong khung tham chiếu. Như vậy, dự báo liên cấu trúc có thể lập mã khối ảnh trong khung hiện thời như một tập các vector chuyển động chỉ báo độ lệch từ khối tương ứng trong khung tham chiếu.

Dự báo nội cấu trúc lập mã các khối trong khung chung. Dự báo nội cấu trúc tận dụng việc các thành phần sắc độ và độ sáng có xu hướng gộp lại trong một khung. Ví dụ, mảng màu xanh lá cây trong một phần của cây có xu hướng được bố trí gần kề với các mảng màu xanh lá cây tương tự. Dự báo nội cấu trúc dùng nhiều chế độ dự báo nhiều chiều (ví dụ ba mươi ba trong HEVC), chế độ phẳng, và chế độ một chiều (Direct Current, DC). Các chế độ theo chiều chỉ báo rằng khối hiện thời tương tự/giống với các mẫu của khối lân cận theo chiều tương ứng. Chế độ phẳng chỉ báo rằng một loạt các khối dọc theo hàng/cột (ví dụ mặt phẳng) có thể được nội suy dựa trên các khối lân cận tại các biên của hàng. Chế độ phẳng, có hiệu quả, chỉ báo chuyển tiếp bằng phẳng của ánh sáng/màu sắc trên hàng/cột bằng cách dùng độ nghiêng không đổi một cách tương đối khi thay đổi giá trị. Chế độ DC được dùng để làm phẳng biên và chỉ báo việc khối tương tự/giống như giá trị trung bình được liên kết với các mẫu của các khối lân cận được liên kết với các chiều góc của chế độ dự báo theo chiều. Do đó, các khối dự báo nội cấu trúc có thể thể hiện các khối ảnh như các giá trị chế độ dự báo có quan hệ khác nhau thay vì các giá trị thực sự. Hơn nữa, các khối dự báo liên cấu trúc có thể thể hiện các khối ảnh như các giá trị vector chuyển động thay vì các giá trị thực sự. Trong cả hai trường hợp, các khối dự báo có thể không thể hiện một cách chính xác các khối ảnh trong một số trường hợp. Giá trị chênh lệch bất kỳ được lưu trong các khối dư. Phép biến đổi có thể được áp dụng cho các khối dư để nén hơn nữa tập tin.

Ở bước 107, các kỹ thuật lọc khác nhau có thể được áp dụng. Trong HEVC, các bộ lọc có thể được áp dụng theo sơ đồ lọc vòng trong. Dự báo dựa trên khối được thảo luận ở trên có thể dẫn đến việc tạo nên các ảnh khối tại bộ giải mã. Hơn nữa, sơ đồ dự báo dựa trên khối có thể lập mã khối và sau đó cấu thành lại khối được lập mã để sử dụng sau này làm khối tham chiếu. Sơ đồ lọc vòng trong áp dụng lặp lại các bộ lọc khử nhiễu, các bộ lọc giải khối, các bộ lọc thích nghi, và các bộ lọc độ lệch thích nghi mẫu (Sample Adaptive Offset, SAO) cho các khối/khung. Các bộ lọc này giảm thiểu các thành phần nhiễu khối sao cho tập tin được lập mã có thể được cấu thành lại một cách chính xác. Hơn nữa, các bộ lọc này giảm thiểu các thành

phần nhiều khối trong các khối tham chiếu được cấu thành lại sao cho các thành phần nhiều khối này ít có khả năng tạo nên các thành phần nhiều khối bổ sung trong các khối tiếp theo mà được lập mã dựa trên các khối tham chiếu được cấu thành lại.

Một khi tín hiệu video được phân chia, được nén, và được lọc, thì dữ liệu kết quả được lập mã trong luồng bit ở bước 109. Luồng bit bao gồm dữ liệu được thảo luận ở trên cũng như dữ liệu báo hiệu bất kỳ được đòi hỏi để hỗ trợ việc cấu thành lại tín hiệu video thích hợp tại bộ giải mã. Ví dụ, dữ liệu như vậy có thể bao gồm dữ liệu phân chia, dữ liệu dự báo, các khối dư, và các cờ khác nhau cung cấp các lệnh mã hóa cho bộ giải mã. Luồng bit có thể được lưu trong bộ nhớ để truyền dẫn đến bộ giải mã khi có yêu cầu. Luồng bit này cũng có thể được phát quảng bá và/hoặc đa điểm đến nhiều bộ giải mã. Việc tạo nên luồng bit là quy trình lặp lại. Do đó, các bước 101, 103, 105, 107, và 109 có thể xảy ra liên tục và/hoặc đồng thời trên nhiều khung và khối. Thứ bậc được thể hiện trên Fig.1 được trình bày cho rõ ràng và dễ dàng thảo luận, và không nhằm hạn chế quy trình mã hóa video theo thứ bậc cụ thể.

Bộ giải mã thu luồng bit và bắt đầu quy trình giải mã ở bước 111. Cụ thể là, bộ giải mã dùng sơ đồ giải mã entropy để chuyển đổi luồng bit thành dữ liệu video và cú pháp tương ứng. Bộ giải mã dùng dữ liệu cú pháp từ luồng bit để xác định các phân chia cho các khung ở bước 111. Việc phân chia phải phù hợp với các kết quả của việc phân chia khối ở bước 103. Việc lập mã/giải mã entropy được dùng ở bước 111 được mô tả dưới đây. Bộ lập mã đưa ra nhiều chọn lựa trong quá trình nén, chẳng hạn như chọn lựa các sơ đồ phân chia khối từ một số chọn lựa có thể dựa trên vị trí không gian của các giá trị trong (các) ảnh đầu vào. Báo hiệu các chọn lựa chính xác có thể dùng một số lượng bit nhị phân lớn. Được sử dụng ở đây, bit nhị phân là giá trị nhị phân mà được xử lý như biến (ví dụ giá trị bit mà có thể thay đổi tùy thuộc vào ngữ cảnh). Mã hóa entropy cho phép bộ lập mã loại bỏ tùy chọn bất kỳ mà rõ ràng là không khả thi cho trường hợp cụ thể, để lại tập các tùy chọn có thể được cho phép. Sau đó mỗi tùy chọn có thể được cho phép được gán từ mã. Độ dài của các từ mã dựa trên số lượng các tùy chọn có thể được cho phép (ví dụ một bit nhị phân cho hai tùy chọn, hai bit nhị phân cho ba đến bốn tùy chọn, v.v.). Bộ lập mã sau đó lập mã từ mã cho tùy chọn được chọn lựa. Sơ đồ này làm giảm kích thước của các từ mã này vì các từ mã này lớn như được đòi hỏi để chỉ báo duy nhất sự chọn lựa từ tập con nhỏ của các tùy chọn có thể được cho phép trái ngược với việc chỉ báo duy nhất sự chọn lựa từ tập tiềm năng lớn của tất cả tùy chọn có thể. Bộ giải mã sau đó

giải mã sự chọn lựa bằng cách xác định tập các tùy chọn có thể được cho phép theo cách thức tương tự với bộ lập mã. Bằng cách xác định tập các tùy chọn có thể được cho phép, bộ giải mã có thể đọc từ mã và xác định sự chọn lựa được thực hiện bởi bộ lập mã.

Ở bước 113, bộ giải mã thực hiện giải mã khối. Cụ thể là, bộ giải mã dùng biến đổi ngược để tạo ra các khối dư. Sau đó bộ giải mã dùng các khối dư và các khối dự báo tương ứng để cấu thành lại các khối ảnh theo dạng phân chia. Các khối dự báo có thể bao gồm cả các khối dự báo nội cấu trúc và các khối dự báo liên cấu trúc như được tạo ra tại bộ lập mã ở bước 105. Các khối ảnh được cấu thành lại sau đó được bố trí vào trong các khung của tín hiệu video được cấu thành lại theo dữ liệu phân chia được xác định ở bước 111. Cú pháp dành cho bước 113 cũng có thể được báo hiệu trong luồng bit qua việc mã hóa entropy như được thảo luận ở trên.

Ở bước 115, việc lọc được thực hiện trên các khung của tín hiệu video được cấu thành lại theo cách thức tương tự với bước 107 tại bộ lập mã. Ví dụ, các bộ lọc khử nhiễu, bộ lọc giải khối, và bộ lọc lặp thích nghi, và bộ lọc SAO có thể được áp dụng cho các khung để loại bỏ các thành phần nhiễu khối. Một khi các khung được lọc, thì tín hiệu video có thể được xuất ra bộ hiển thị ở bước 117 để quan sát bởi người dùng.

Fig.2 là hình vẽ sơ đồ thể hiện hệ thống mã hóa và giải mã (codec) ví dụ 200 để mã hóa video. Cụ thể là, hệ thống mã hóa và giải mã 200 cung cấp chức năng để hỗ trợ thực hiện phương pháp hoạt động 100. Hệ thống mã hóa và giải mã 200 được tạo ra để thể hiện các thành phần được dùng trong cả bộ lập mã và bộ giải mã. Hệ thống mã hóa và giải mã 200 thu và phân chia tín hiệu video như được thảo luận liên quan đến các bước 101 và 103 theo phương pháp hoạt động 100, mà dẫn đến kết quả là tín hiệu video được phân chia 201. Hệ thống mã hóa và giải mã 200 sau đó nén tín hiệu video được phân chia 201 thành luồng bit được mã hóa khi hoạt động như bộ lập mã như được thảo luận liên quan đến các bước 105, 107, và 109 theo phương pháp 100. Khi hoạt động như bộ giải mã, thì hệ thống mã hóa và giải mã 200 tạo ra tín hiệu video đầu ra từ luồng bit như được thảo luận liên quan đến các bước 111, 113, 115, và 117 theo phương pháp hoạt động 100. Hệ thống mã hóa và giải mã 200 bao gồm thành phần điều khiển bộ mã hóa chung 211, thành phần lượng tử và chia tỷ lệ biến đổi 213, thành phần ước lượng nội cấu trúc hình 215, thành phần dự báo nội cấu trúc hình 217, thành phần bù chuyển động 219, thành phần ước lượng chuyển động 221, thành phần biến đổi ngược và

chia tỷ lệ 229, thành phần phân tích điều khiển bộ lọc 227, thành phần các bộ lọc vòng trong 225, thành phần lưu đệm hình được giải mã 223, và thành phần mã hóa số học nhị phân thích nghi ngữ cảnh (Context Adaptive Binary Arithmetic Coding, CABAC) và định dạng phần đầu 231. Các thành phần như vậy được ghép nối như được thể hiện. Trên Fig.2, các đường màu đen thể hiện sự di chuyển của dữ liệu được lập mã/giải mã trong khi các đường nét đứt thể hiện sự di chuyển của dữ liệu điều khiển mà điều khiển hoạt động của các thành phần khác. Tất cả các thành phần của hệ thống mã hóa và giải mã 200 có thể có mặt trong bộ lập mã. Bộ giải mã có thể bao gồm một tập con của các thành phần của hệ thống mã hóa và giải mã 200. Ví dụ, bộ giải mã có thể bao gồm thành phần dự báo nội cấu trúc hình 217, thành phần bù chuyển động 219, thành phần biến đổi ngược và chia tỷ lệ 229, thành phần các bộ lọc vòng trong 225, và thành phần lưu đệm hình được giải mã 223. Các thành phần này sẽ được mô tả bên dưới.

Tín hiệu video được phân chia 201 là trình tự video được quay mà đã được phân chia thành các khối của các điểm ảnh bởi cây mã hóa. Cây mã hóa dùng các chế độ chia tách khác nhau để chia nhỏ khối các điểm ảnh thành các khối nhỏ hơn của các điểm ảnh. Sau đó các khối này có thể được chia nhỏ hơn nữa thành các khối nhỏ hơn. Các khối này có thể được gọi là các nút trên cây mã hóa. Các nút cha mẹ lớn hơn được chia tách thành các nút con cái nhỏ hơn. Số lần nút được chia nhỏ được gọi là độ sâu của nút/cây mã hóa. Các khối được chia có thể được chứa trong các đơn vị mã hóa (CU) trong một số trường hợp. Ví dụ, CU có thể là phần con của CTU mà chứa khối độ sáng, (các) khối sắc độ (Cr) khác nhau màu đỏ, và (các) khối sắc độ (Cb) khác nhau màu xanh lam cùng với các lệnh cú pháp tương ứng dành cho CU. Các chế độ chia tách có thể bao gồm cây nhị phân (Binary Tree, BT), cây tam phân (Triple Tree, TT), và cây tứ phân (Quad Tree, QT) lần lượt được dùng để phân chia nút thành hai, ba, hoặc bốn nút con cái, có hình dạng thay đổi tùy thuộc vào các chế độ chia tách được dùng. Tín hiệu video được phân chia 201 được chuyển tới thành phần điều khiển bộ mã hóa chung 211, thành phần lượng tử và chia tỷ lệ biến đổi 213, thành phần ước lượng nội cấu trúc hình 215, thành phần phân tích điều khiển bộ lọc 227, và thành phần ước lượng chuyển động 221 để nén.

Thành phần điều khiển bộ mã hóa chung 211 được tạo cấu hình để đưa ra quyết định liên quan đến việc mã hóa các ảnh của trình tự video thành luồng bit theo các ràng buộc ứng dụng.

Ví dụ, thành phần điều khiển bộ mã hóa chung 211 quản lý việc tối ưu tốc độ bit/kích thước luồng bit so với chất lượng cấu thành lại. Các quyết định như vậy có thể được thực hiện dựa trên không gian lưu trữ/độ khả dụng băng thông và các yêu cầu độ phân giải ảnh. Thành phần điều khiển bộ mã hóa chung 211 cũng quản lý việc sử dụng bộ đệm theo tốc độ truyền dẫn ánh sáng để giảm thiểu các vấn đề thiếu và tràn bộ đệm. Để quản lý các vấn đề này, thì thành phần điều khiển bộ mã hóa chung 211 quản lý việc phân chia, dự báo, và lọc bởi các thành phần khác. Ví dụ, thành phần điều khiển bộ mã hóa chung 211 có thể tăng động độ phức tạp nén để làm tăng độ phân giải và tăng mức sử dụng băng thông hoặc giảm độ phức tạp nén để làm giảm độ phân giải và mức sử dụng băng thông. Do đó, thành phần điều khiển bộ mã hóa chung 211 điều khiển các thành phần khác của hệ thống mã hóa và giải mã 200 để cân bằng chất lượng cấu thành lại tín hiệu video với tốc độ bit liên quan. Thành phần điều khiển bộ mã hóa chung 211 tạo nên dữ liệu điều khiển, mà điều khiển hoạt động của các thành phần khác. Dữ liệu điều khiển cũng được chuyển tới thành phần CABAC và định dạng phần đầu 231 được lập mã trong luồng bit thành các thông số báo hiệu để giải mã tại bộ giải mã.

Tín hiệu video được phân chia 201 cũng được gửi tới thành phần ước lượng chuyển động 221 và thành phần bù chuyển động 219 để dự báo liên cấu trúc. Khung hoặc phân chia của tín hiệu video được phân chia 201 có thể được chia thành nhiều khối video. Thành phần ước lượng chuyển động 221 và thành phần bù chuyển động 219 thực hiện mã hóa dự báo liên cấu trúc của khối video thu được so với một hoặc nhiều khối trong một hoặc nhiều khung tham chiếu để cung cấp sự dự báo thời gian. Hệ thống mã hóa và giải mã 200 có thể thực hiện nhiều lần mã hóa, ví dụ để chọn lựa chế độ mã hóa thích hợp cho mỗi khối dữ liệu video.

Thành phần ước lượng chuyển động 221 và thành phần bù chuyển động 219 có thể được tích hợp ở mức cao, nhưng không được thể hiện riêng biệt vì mục đích khái niệm. Việc ước lượng chuyển động, được thực hiện bởi thành phần ước lượng chuyển động 221, là quy trình để tạo ra các vectơ chuyển động, mà ước lượng chuyển động cho các khối video. Vectơ chuyển động, ví dụ, có thể chỉ báo độ dịch chuyển của đối tượng được mã hóa so với khối dự báo. Khối dự báo là khối mà được tìm thấy trùng khớp với khối được mã hóa, về mặt chênh lệch điểm ảnh. Khối dự báo cũng có thể được gọi là khối tham chiếu. Độ chênh lệch điểm ảnh có thể được xác định bằng tổng của độ chênh lệch tuyệt đối (Absolute Difference, SAD), tổng của độ chênh lệch hình vuông (Square Difference, SSD), hoặc chuẩn đo độ chênh lệch khác.

HEVC dùng một số đối tượng được mã hóa bao gồm CTU, các khối cây mã hóa (CTB), và các CU. Ví dụ, CTU có thể được chia thành các CTB, mà sau đó có thể được chia thành các CB để đưa vào các CU. CU có thể được lập mã làm đơn vị dự báo (Prediction Unit, PU) chứa đơn vị dự báo và/hoặc biến đổi (Transform Unit, TU) chứa dữ liệu dư được biến đổi cho CU. Thành phần ước lượng chuyển động 221 tạo ra các vector chuyển động, các PU, và các TU bằng cách sử dụng phân tích tỷ lệ-độ méo như một phần của quy trình tối ưu tỷ lệ-độ méo. Ví dụ, thành phần ước lượng chuyển động 221 có thể xác định nhiều khối tham chiếu, nhiều vector chuyển động, v.v., cho khối/khung hiện thời, và có thể chọn lựa các khối tham chiếu, vector chuyển động, v.v., có các đặc tính tỷ lệ-độ méo tốt nhất. Các đặc tính tỷ lệ-độ méo tốt nhất cân bằng cả chất lượng cấu thành lại video (ví dụ lượng dữ liệu bị mất bởi quá trình nén) với hiệu quả mã hóa (ví dụ kích thước lập mã cuối cùng).

Theo một số ví dụ, hệ thống mã hóa và giải mã 200 có thể tính các giá trị cho các vị trí điểm ảnh nguyên con của các hình tham chiếu được lưu trong thành phần lưu đệm hình được giải mã 223. Ví dụ, hệ thống mã hóa và giải mã video 200 có thể nội suy các giá trị của các vị trí điểm ảnh một phần tư, vị trí điểm ảnh một phần tám, hoặc vị trí điểm ảnh phân số khác của hình tham chiếu. Do đó, thành phần ước lượng chuyển động 221 có thể thực hiện tìm kiếm chuyển động so với các vị trí điểm ảnh toàn phần và các vị trí điểm ảnh phân số và xuất ra vector chuyển động với độ chính xác điểm ảnh phân số. Thành phần ước lượng chuyển động 221 tính vector chuyển động cho PU của khối video trong phần chia được mã hóa liên cấu trúc bằng các so sánh vị trí của PU với vị trí của khối dự báo của hình tham chiếu. Thành phần ước lượng chuyển động 221 xuất ra vector chuyển động được tính làm dữ liệu chuyển động cho thành phần CABAC và định dạng phần đầu 231 để lập mã và chuyển động cho thành phần bù chuyển động 219.

Việc bù chuyển động, được thực hiện bởi thành phần bù chuyển động 219, có thể bao gồm việc truy xuất và tạo ra khối dự báo dựa trên vector chuyển động được xác định bởi thành phần ước lượng chuyển động 221. Hơn nữa, thành phần ước lượng chuyển động 221 và thành phần bù chuyển động 219 có thể được tích hợp về mặt chức năng, theo một số ví dụ. Khi thu vector chuyển động cho PU của khối video hiện thời, thì thành phần bù chuyển động 219 có thể đặt khối dự báo vào vị trí mà vector chuyển động chỉ. Khối video dư sau đó được tạo thành bằng cách trừ đi các giá trị điểm ảnh của khối dự báo khỏi các giá trị điểm ảnh của khối video

hiện thời được mã hóa, tạo thành các giá trị độ chênh điểm ảnh. Nói chung, thành phần ước lượng chuyển động 221 thực hiện ước lượng chuyển động so với các thành phần độ sáng, và thành phần bù chuyển động 219 sử dụng các vector chuyển động được tính dựa trên các thành phần độ sáng cho cả các thành phần sắc độ và các thành phần độ sáng. Khối dự báo và khối dư được chuyển tới thành phần lượng tử và chia tỷ lệ biến đổi 213.

Tín hiệu video được phân chia 201 cũng được gửi tới thành phần ước lượng nội cấu trúc hình 215 và thành phần dự báo nội cấu trúc hình 217. Với thành phần ước lượng chuyển động 221 và thành phần bù chuyển động 219, thành phần ước lượng nội cấu trúc hình 215 và thành phần dự báo nội cấu trúc hình 217 có thể được tích hợp ở mức độ cao, nhưng không được thể hiện riêng biệt vì mục đích khái niệm. Thành phần ước lượng nội cấu trúc hình 215 và thành phần dự báo nội cấu trúc hình 217 dự báo nội cấu trúc khối hiện thời so với các khối trong khung hiện thời, như là một sự thay thế cho việc dự báo liên cấu trúc được thực hiện bởi thành phần ước lượng chuyển động 221 và thành phần bù chuyển động 219 giữa các khung, như được mô tả ở trên. Cụ thể, thành phần ước lượng nội cấu trúc hình 215 xác định chế độ dự báo nội cấu trúc để sử dụng để lập mã khối hiện thời. Theo một số ví dụ, thành phần ước lượng nội cấu trúc hình 215 chọn lựa chế độ dự báo nội cấu trúc thích hợp để lập mã khối hiện thời từ nhiều chế độ dự báo nội cấu trúc đã được thử nghiệm. Các chế độ dự báo nội cấu trúc được chọn lựa sau đó được chuyển tới thành phần CABAC và định dạng phần đầu 231 để lập mã.

Ví dụ, thành phần ước lượng nội cấu trúc hình 215 tính các giá trị tỷ lệ-độ méo sử dụng phân tích tỷ lệ-độ méo cho các chế độ dự báo nội cấu trúc đã được thử nghiệm, và chọn lựa chế độ dự báo nội cấu trúc có các đặc tính tỷ lệ-độ méo tốt nhất trong số các chế độ đã được thử nghiệm. Phân tích tỷ lệ-độ méo thường xác định lượng méo (hoặc lỗi) giữa khối được lập mã và khối không được lập mã gốc mà đã được lập mã để tạo ra khối được lập mã, cũng như tốc độ bit (ví dụ số lượng bit) được sử dụng để tạo ra khối được lập mã. Thành phần ước lượng nội cấu trúc hình 215 tính tỷ số từ độ méo và tỷ lệ cho các khối được lập mã khác nhau để xác định chế độ dự báo nội cấu trúc nào có giá trị tỷ lệ-độ méo tốt nhất cho khối này. Ngoài ra, thành phần ước lượng nội cấu trúc hình 215 có thể được tạo cấu hình để mã hóa độ sâu các khối của bản đồ độ sâu sử dụng chế độ mô hình độ sâu (DMM) dựa trên việc tối ưu tỷ lệ-độ méo (RDO).

Thành phần dự báo nội cấu trúc hình 217 có thể tạo ra khối dư từ khối dự báo dựa trên

các chế độ dự báo nội cấu trúc được chọn lựa được xác định bởi thành phần ước lượng nội cấu trúc hình 215 khi được thực hiện trên bộ lập mã hoặc đọc khối dư từ luồng bit khi được thực hiện trên bộ giải mã. Khối dư bao gồm độ chênh lệch theo giá trị giữa khối dự báo và khối gốc, được thể hiện dưới dạng ma trận. Khối dư sau đó được chuyển tới thành phần lượng tử và chia tỷ lệ biến đổi 213. Thành phần ước lượng nội cấu trúc hình 215 và thành phần dự báo nội cấu trúc hình 217 có thể vận hành trên cả thành phần độ sáng và sắc độ.

Thành phần lượng tử và chia tỷ lệ biến đổi 213 được tạo cấu hình để nén hơn nữa khối dư. Thành phần lượng tử và chia tỷ lệ biến đổi 213 áp dụng biến đổi, chẳng hạn như biến đổi cosin rời rạc (Discrete Cosine Transform, DCT), biến đổi sin rời rạc (Discrete Sine Transform, DST), hoặc biến đổi tương tự về mặt khái niệm, cho khối dư, tạo ra khối video bao gồm các giá trị hệ số biến đổi dư. Biến đổi wavelet, biến đổi số nguyên, biến đổi băng con hoặc kiểu biến đổi khác cũng có thể được sử dụng. Phép biến đổi này có thể chuyển đổi thông tin dư từ miền giá trị điểm ảnh sang miền biến đổi, chẳng hạn như 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, ví dụ dựa trên tần số. Việc chia tỷ lệ này bao gồm việc áp dụng hệ số tỷ lệ cho thông tin dư sao cho thông tin tần số khác nhau được lượng tử tại các độ chia nhỏ khác nhau, mà có thể ảnh hưởng tới chất lượng trực quan cuối cùng của video được cấu thành lại. 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 làm giảm xuống hơn nữa tốc độ bit. Quy trình lượng tử có thể làm giảm độ sâu bit được liên kết với một số hoặc toàn bộ các hệ số. Mức độ lượng tử có thể được cải biến bằng cách điều chỉnh thông số lượng tử. Theo một số ví dụ, thành phần lượng tử và chia tỷ lệ biến đổi 213 sau đó có thể thực hiện quét ma trận bao gồm các hệ số biến đổi được lượng tử. Các hệ số biến đổi được lượng tử được chuyển tới thành phần CABAC và định dạng phần đầu 231 được lập mã trong luồng bit.

Thành phần biến đổi ngược và chia tỷ lệ 229 áp dụng phép toán nghịch đảo của thành phần lượng tử và chia tỷ lệ biến đổi 213 để hỗ trợ ước lượng chuyển động. Thành phần biến đổi ngược và chia tỷ lệ 229 áp dụng phép chia tỷ lệ, biến đổi, và/hoặc lượng tử ngược để cấu thành lại khối dư trong miền điểm ảnh, ví dụ để sau này sử dụng làm khối tham chiếu mà có thể trở thành khối dự báo cho khối hiện thời khác. Thành phần ước lượng chuyển động 221 và/hoặc thành phần bù chuyển động 219 có thể tính khối tham chiếu bằng cách thêm khối dư

trở lại khối dự báo tương ứng để sử dụng ước lượng chuyển động của khối/khung sau đó. Các bộ lọc được áp dụng cho các khối tham chiếu được cấu thành lại để giảm thiểu các thành phần nhiễu được tạo nên trong khi chia tỷ lệ, lượng tử, và biến đổi. Các thành phần nhiễu khối này có thể làm cho việc dự báo không chính xác (và tạo nên các thành phần nhiễu khối bổ sung) khi khối tiếp theo được dự báo.

Thành phần phân tích điều khiển bộ lọc 227 và thành phần các bộ lọc vòng trong 225 áp dụng các bộ lọc cho các khối dư và/hoặc cho các khối ảnh được cấu thành lại. Ví dụ, 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 nội cấu trúc hình 217 và/hoặc thành phần bù chuyển động 219 để cấu thành lại khối ảnh gốc. Sau đó các bộ lọc có thể được áp dụng cho khối ảnh được cấu thành lại. Theo một số ví dụ, thay vào đó các bộ lọc có thể được áp dụng cho các khối dư. Với các thành phần khác trên Fig.2, thành phần phân tích điều khiển bộ lọc 227 và thành phần các bộ lọc vòng trong 225 được tích hợp ở mức độ cao và có thể được thực hiện cùng nhau, nhưng không được thể hiện riêng biệt vì mục đích khái niệm. Các bộ lọc được áp dụng cho các khối tham chiếu được cấu thành lại được áp dụng cho các vùng không gian cụ thể và bao gồm nhiều thông số để điều chỉnh cách để các bộ lọc này được áp dụng. Thành phần phân tích điều khiển bộ lọc 227 phân tích các khối tham chiếu được cấu thành lại để xác định vị trí các bộ lọc này nên được áp dụng và thiết đặt các thông số tương ứng. Dữ liệu như vậy được chuyển tới thành phần CABAC và định dạng phần đầu 231 làm dữ liệu điều khiển bộ lọc để lập mã. Thành phần các bộ lọc vòng trong 225 áp dụng các bộ lọc này dựa trên dữ liệu điều khiển bộ lọc. Các bộ lọc có thể bao gồm bộ lọc giải khối, bộ lọc khử nhiễu, bộ lọc SAO, và bộ lọc vòng thích nghi. Các bộ lọc này có thể được áp dụng trong miền không gian/điểm ảnh (ví dụ trên khối điểm ảnh được cấu thành lại) hoặc trong miền tần số, tùy thuộc vào ví dụ.

Khi vận hành như bộ lập mã, thì khối ảnh được cấu thành lại được lọc, khối dư, và/hoặc khối dự báo được lưu trong thành phần lưu đệm hình được giải mã 223 để sau này sử dụng ước lượng chuyển động như được thảo luận ở trên. Khi vận hành như bộ giải mã, thì thành phần lưu đệm hình được giải mã 223 lưu và chuyển các khối được cấu thành lại và được lọc tới bộ hiển thị như một phần của tín hiệu video đầu ra. Thành phần lưu đệm hình được giải mã 223 có thể là thiết bị nhớ bất kỳ có khả năng lưu các khối dự báo, các khối dư, và/hoặc các khối ảnh được cấu thành lại.

Thành phần CABAC và định dạng phần đầu 231 thu dữ liệu từ các thành phần khác nhau của hệ thống mã hóa và giải mã 200 và lập mã dữ liệu như vậy thành luồng bit được mã hóa để truyền dẫn tới bộ giải mã. Cụ thể là, thành phần CABAC và định dạng phần đầu 231 tạo ra các phần đầu khác nhau để lập mã dữ liệu điều khiển, chẳng hạn như tạo ra dữ liệu điều khiển và dữ liệu điều khiển bộ lọc. Hơn nữa, dữ liệu dự báo, bao gồm dữ liệu chuyển động và dự báo nội cấu trúc, cũng như dữ liệu dư dưới dạng dữ liệu hệ số biến đổi được lượng tử được lập mã tất cả trong luồng bit. Luồng bit cuối cùng bao gồm tất cả thông tin được đòi hỏi bởi bộ giải mã để cấu thành lại tín hiệu video được phân chia gốc 201. Thông tin như vậy cũng có thể bao gồm các bảng chỉ số chế độ dự báo nội cấu trúc (còn được gọi là các bảng ánh xạ từ mã), các định nghĩa của các ngữ cảnh lập mã cho các khối khác nhau, các chỉ báo của các chế độ dự báo nội cấu trúc có thể xảy ra nhất, chỉ báo của thông tin phân chia, v.v.. Dữ liệu như vậy có thể được lập mã bằng cách dùng mã hóa entropy. Ví dụ, thông tin này có thể được lập mã bằng cách dùng kỹ thuật mã hóa chiều dài biến thiên thích nghi ngữ cảnh (context adaptive variable length coding, CAVLC), CABAC, mã hóa số học nhị phân thích nghi ngữ cảnh dựa trên cú pháp (Syntax-based context-adaptive Binary Arithmetic Coding, SBAC), mã hóa entropy phân chia khoảng xác suất (Probability Interval Partitioning Entropy, PIPE), hoặc kỹ thuật mã hóa entropy khác. Tiếp theo việc mã hóa entropy, thì luồng bit được mã hóa có thể được truyền tới thiết bị khác (ví dụ bộ giải mã video) hoặc được ghi để truyền dẫn sau này hoặc để truy hồi.

Fig.3 là hình vẽ sơ đồ thể hiện bộ lập mã video ví dụ 300. Bộ lập mã video 300 có thể được dùng để thực hiện các chức năng lập mã của hệ thống mã hóa và giải mã 200 và/hoặc thực hiện các bước 101, 103, 105, 107, và/hoặc 109 của phương pháp hoạt động 100. Bộ lập mã 300 phân chia tín hiệu video đầu vào, kết quả là tín hiệu video được phân chia 301, mà gần như tương tự với tín hiệu video được phân chia 201. Tín hiệu video được phân chia 301 sau đó được nén và được lập mã thành luồng bit bởi các thành phần của bộ lập mã 300.

Cụ thể là, tín hiệu video được phân chia 301 được chuyển tới thành phần dự báo nội cấu trúc hình 317 để dự báo nội cấu trúc. Thành phần dự báo nội cấu trúc hình 317 gần như có thể tương tự với thành phần ước lượng nội cấu trúc hình 215 và thành phần dự báo nội cấu trúc hình 217. Tín hiệu video được phân chia 301 cũng được chuyển tới thành phần bù chuyển động 321 để dự báo liên cấu trúc dựa trên các khối tham chiếu trong thành phần lưu đệm hình

được giải mã 323. Thành phần bù chuyển động 321 gần như có thể tương tự với thành phần ước lượng chuyển động 221 và thành phần bù chuyển động 219. Các khối dự báo và khối dư từ thành phần dự báo nội cấu trúc hình 317 và thành phần bù chuyển động 321 được chuyển tới thành phần lượng tử và biến đổi 313 để biến đổi và lượng tử các khối dư. Thành phần lượng tử và biến đổi 313 gần như có thể tương tự với thành phần lượng tử và chia tỷ lệ biến đổi 213. Các khối dư được lượng tử và biến đổi và các khối dự báo tương ứng (cùng với dữ liệu điều khiển được liên kết) được chuyển tới thành phần mã hóa entropy 331 để mã hóa thành luồng bit. Thành phần mã hóa entropy 331 gần như có thể tương tự với thành phần CABAC và định dạng phần đầu 231.

Các khối dư được lượng tử và biến đổi và/hoặc các khối dự báo tương ứng cũng được chuyển từ thành phần lượng tử và biến đổi 313 tới thành phần lượng tử và biến đổi ngược 329 để cấu thành lại thành các khối tham chiếu để sử dụng bởi thành phần bù chuyển động 321. Thành phần lượng tử và biến đổi ngược 329 gần như có thể tương tự với thành phần biến đổi ngược và chia tỷ lệ 229. Các bộ lọc vòng trong thành phần các bộ lọc vòng trong 325 cũng được áp dụng cho các khối dư và/hoặc các khối tham chiếu được cấu thành lại, tùy thuộc vào ví dụ này. Thành phần các bộ lọc vòng trong 325 gần như có thể tương tự với thành phần phân tích điều khiển bộ lọc 227 và thành phần các bộ lọc vòng trong 225. Thành phần các bộ lọc vòng trong 325 có thể bao gồm nhiều bộ lọc như được thảo luận liên quan tới thành phần các bộ lọc vòng trong 225. Các khối được lọc sau đó được lưu trong thành phần lưu đệm hình được giải mã 323 để sử dụng làm các khối tham chiếu bởi thành phần bù chuyển động 321. Thành phần lưu đệm hình được giải mã 323 gần như có thể tương tự với thành phần lưu đệm hình được giải mã 223.

Fig.4 là hình vẽ sơ đồ thể hiện bộ giải mã video ví dụ 400. Bộ giải mã video 400 có thể được dùng để thực hiện các chức năng giải mã của hệ thống mã hóa và giải mã 200 và/hoặc thực hiện các bước 111, 113, 115, và/hoặc 117 của phương pháp hoạt động 100. Bộ giải mã 400 thu luồng bit, ví dụ từ bộ lập mã 300, và tạo ra tín hiệu video đầu ra được cấu thành lại dựa trên luồng bit để hiển thị cho người dùng.

Luồng bit được thu bởi thành phần giải mã entropy 433. Thành phần giải mã entropy 433 được tạo cấu hình để thực hiện sơ đồ giải mã entropy, chẳng hạn như CAVLC, CABAC, SBAC, mã hóa PIPE, hoặc các kỹ thuật mã hóa entropy khác. Ví dụ, thành phần giải mã

entropy 433 có thể dùng thông tin phần đầu để cung cấp ngữ cảnh để diễn dịch dữ liệu bổ sung được lập mã làm các từ mã trong luồng bit. Thông tin được giải mã bao gồm thông tin được đòi hỏi bất kỳ để giải mã tín hiệu video, chẳng hạn như dữ liệu điều khiển chung, dữ liệu điều khiển bộ lọc, thông tin phân chia, 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 tới thành phần lượng tử và biến đổi ngược 429 để cấu thành lại thành các khối dư. Thành phần lượng tử và biến đổi ngược 429 có thể tương tự với thành phần lượng tử và biến đổi ngược 329.

Các khối dư được cấu thành lại và/hoặc các khối dự báo được chuyển tới thành phần dự báo nội cấu trúc hình 417 để cấu thành lại thành các khối ảnh dựa trên các hoạt động dự báo nội cấu trúc. Thành phần dự báo nội cấu trúc hình 417 có thể tương tự với thành phần ước lượng nội cấu trúc hình 215 và thành phần dự báo nội cấu trúc hình 217. Cụ thể là, thành phần dự báo nội cấu trúc hình 417 dùng các chế độ dự báo để đặt khối tham chiếu vào trong khung và áp dụng khối dư cho kết quả để cấu thành lại các khối ảnh được dự báo nội cấu trúc. Các khối ảnh được dự báo nội cấu trúc được cấu thành lại và/hoặc các khối dư và dữ liệu dự báo liên cấu trúc tương ứng được chuyển tới thành phần lưu đệm hình được giải mã 423 qua thành phần các bộ lọc vòng trong 425, mà gần như có thể lần lượt tương tự với thành phần lưu đệm hình được giải mã 223 và thành phần các bộ lọc vòng trong 225. Thành phần các bộ lọc vòng trong 425 lọc các khối ảnh được cấu thành lại, các khối dư và/hoặc các khối dự báo, và thông tin như vậy được lưu trong thành phần lưu đệm hình được giải mã 423. Các khối ảnh được cấu thành lại từ thành phần lưu đệm hình được giải mã 423 được chuyển tới thành phần bù chuyển động 421 để dự báo liên cấu trúc. Thành phần bù chuyển động 421 gần như có thể tương tự với thành phần ước lượng chuyển động 221 và/hoặc thành phần bù chuyển động 219. Cụ thể là, thành phần bù chuyển động 421 dùng các vector chuyển động từ khối tham chiếu để tạo ra khối dự báo và áp dụng khối dư cho kết quả này để cấu thành lại khối ảnh. Các khối được cấu thành lại cũng có thể được chuyển qua thành phần các bộ lọc vòng trong 425 tới thành phần lưu đệm hình được giải mã 423. Thành phần lưu đệm hình được giải mã 423 tiếp tục lưu các khối ảnh được cấu thành lại bổ sung, mà có thể được cấu thành lại thành các khung qua thông tin phân chia. Các khung như vậy cũng có thể được đặt theo trình tự. Trình tự này được xuất ra bộ hiển thị như tín hiệu video đầu ra được cấu thành lại.

Ghi nhớ những phần ở trên, các kỹ thuật nén video thực hiện dự báo (nội cấu trúc hình)

không và/hoặc dự báo (liên cấu trúc hình) thời gian để làm giảm hoặc loại bỏ phần dư vốn có trong các trình tự video. Để mã hóa video dựa trên khối, phần chia video (tức là hình video hoặc một phần của hình video) có thể được phân chia thành các khối video, mà còn có thể được gọi là các khối cây, các khối cây mã hóa (CTB), các đơn vị cây mã hóa (CTU), các đơn vị mã hóa (CU), và/hoặc các nút mã hóa. Các khối video trong phần chia (I) được mã hóa nội cấu trúc của hình được lập mã sử dụng dự báo không gian đối với các mẫu tham chiếu trong các khối lân cận trong cùng một hình. Các khối video trong phần chia (P hoặc B) được mã hóa liên cấu trúc của hình có thể sử dụng dự báo không gian đối với các mẫu tham chiếu trong các khối lân cận trong cùng một hình hoặc dự báo thời gian đối với các mẫu tham chiếu trong các hình tham chiếu khác. Các hình có thể được gọi là các khung, và các hình tham chiếu có thể được gọi là các khung tham chiếu. POC là biến được liên kết với mỗi hình mà nhận dạng duy nhất hình được liên kết trong số tất cả các hình trong trình tự video lớp được mã hóa (coded layer video sequence, CLVS), chỉ báo khi hình được liên kết sẽ được xuất ra từ DPB, và chỉ báo vị trí của hình được liên kết theo thứ bậc xuất ra so với các vị trí thứ bậc xuất ra của các hình khác trong cùng một CLVS mà sẽ được xuất ra từ DPB. Cờ là biến hoặc phần tử cú pháp một bit mà có thể có một trong số hai giá trị có thể: 0 và 1.

Dự báo không gian hoặc thời gian dẫn đến kết quả khối dự báo dành cho khối được mã hóa. Dữ liệu dư thể hiện các độ chênh lệch điểm ảnh giữa khối gốc được mã hóa và khối dự báo. Khối được mã hóa liên cấu trúc được lập mã theo vectơ chuyển động mà trở tới khối của các mẫu tham chiếu tạo thành khối dự báo, và dữ liệu dư chỉ báo độ chênh lệch giữa khối được mã hóa và khối dự báo. Khối được mã hóa nội cấu trúc được lập mã theo chế độ mã hóa nội cấu trúc và dữ liệu dư. Để nén thêm nữa, dữ liệu dư có thể được biến đổi từ miền điểm ảnh thành miền biến đổi, kết quả là các hệ số biến đổi dư, mà sau đó có thể được lượng tử. Các hệ số biến đổi được lượng tử, được sắp xếp khởi tạo trong mảng hai chiều, có thể được quét để tạo ra vectơ một chiều của các hệ số biến đổi, và mã hóa entropy có thể được áp dụng để đạt được hiệu quả nén nhiều hơn nữa.

Nén ảnh và video đã trải qua sự phát triển nhanh chóng, dẫn đến các tiêu chuẩn mã hóa khác nhau. Các tiêu chuẩn mã hóa video như vậy bao gồm ITU-T H.261, tổ chức tiêu chuẩn hóa quốc tế (International Organization for Standardization, ISO)/ủy ban kỹ thuật điện quốc tế (International Electrotechnical Commission, IEC) 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, mã hóa video cao cấp (Advanced Video Coding, AVC), cũng được biết đến như ITU-T H.264 hoặc ISO/IEC MPEG-4 Part 10, và mã hóa video hiệu quả cao (High Efficiency Video Coding, HEVC), cũng được biết đến như ITU-T H.265 hoặc MPEG-H Part 2. AVC bao gồm các phần mở rộng chẳng hạn như mã hóa video có thể mở rộng (Scalable Video Coding, SVC), mã hóa video nhiều cảnh nhìn (Multiview Video Coding, MVC) và mã hóa video nhiều cảnh nhìn cộng độ sâu (Multiview Video Coding plus Depth, MVC+D), và 3D AVC (3D-AVC). HEVC bao gồm các phần mở rộng chẳng hạn như HEVC có thể mở rộng (Scalable HEVC, SHVC), HEVC nhiều cảnh nhìn (Multiview HEVC, MV-HEVC), và 3D HEVC (3D-HEVC).

Cũng có thể có chuẩn mã hóa video mới, có tên là mã hóa video đa năng (VVC), được phát triển bởi nhóm chuyên gia video chung (JVET) của ITU-T và ISO/IEC. Trong khi chuẩn VVC có một số bản nháp, thì một bản nháp (WD) của VVC nói riêng, cụ thể là B. Bross, J. Chen, và S. Liu, “Versatile Video Coding (Draft 5)”, JVET-N1001-v3, 13th JVET Meeting, 27 tháng 3 năm 2019 (VVC Draft 5) được tham khảo ở đây.

Khả năng mở rộng khi mã hóa video thường được hỗ trợ bằng cách sử dụng các kỹ thuật mã hóa nhiều lớp. Luồng bit nhiều lớp bao gồm lớp cơ sở (base layer, BL) và một hoặc nhiều lớp tăng cường (enhancement layer, EL). Một ví dụ về khả năng mở rộng bao gồm khả năng mở rộng không gian, khả năng mở rộng chất lượng/tỷ số tín hiệu trên tạp (signal-to-noise, SNR), khả năng mở rộng nhiều cảnh nhìn, v.v.. Khi kỹ thuật mã hóa nhiều lớp được sử dụng, thì hình hoặc một phần của nó có thể được mã hóa (1) mà không sử dụng hình tham chiếu, tức là, sử dụng dự báo nội cấu trúc; (2) bằng cách tham chiếu tới các hình tham chiếu mà ở trong cùng một lớp, tức là, sử dụng dự báo liên cấu trúc; hoặc (3) bằng cách tham chiếu tới các hình tham chiếu mà ở trong (các) lớp khác, tức là, sử dụng dự báo liên lớp. Hình tham chiếu được sử dụng dành cho việc dự báo liên lớp của hình hiện thời được gọi là hình tham chiếu liên lớp (inter-layer reference picture, ILRP).

Fig.5 là hình vẽ sơ đồ thể hiện một ví dụ về dự báo dựa trên lớp 500, ví dụ như được thực hiện để xác định các MV ở bước nén khối 105, bước giải mã khối 113, thành phần ước lượng chuyển động 221, thành phần bù chuyển động 219, thành phần bù chuyển động 321, và/hoặc thành phần bù chuyển động 421. Dự báo dựa trên lớp 500 tương thích với dự báo liên cấu trúc một chiều và/hoặc dự báo liên cấu trúc hai chiều, nhưng cũng được thực hiện giữa

các hình trong các lớp khác nhau.

Dự báo dựa trên lớp 500 được áp dụng giữa các hình 511, 512, 513, và 514 và các hình 515, 516, 517, và 518 trong các lớp khác nhau. Trong ví dụ được thể hiện, các hình 511, 512, 513, và 514 là một phần của lớp N+1 532 và các hình 515, 516, 517, và 518 là một phần của lớp N 531. Một lớp, chẳng hạn như lớp N 531 và/hoặc lớp N+1 532, là một nhóm của các hình mà tất cả được liên kết với giá trị tương tự của đặc tính, chẳng hạn như kích thước tương tự, chất lượng tương tự, độ phân giải tương tự, tỷ số tín hiệu trên tạp tương tự, khả năng tương tự, v.v.. Trong ví dụ được thể hiện, lớp N+1 532 được liên kết với kích thước ảnh lớn hơn so với lớp N 531. Do đó, các hình 511, 512, 513, và 514 trong lớp N+1 532 có kích thước hình lớn hơn (ví dụ, độ cao và độ rộng lớn hơn và do đó nhiều mẫu hơn) so với các hình 515, 516, 517, và 518 trong lớp N 531 trong ví dụ này. Tuy nhiên, các hình như vậy có thể bị tách rời giữa lớp N+1 532 và lớp N 531 bởi các đặc tính khác. Trong khi chỉ hai lớp, lớp N+1 532 và lớp N 531, được thể hiện, thì tập các hình có thể được tách rời thành số lớp bất kỳ dựa trên các đặc tính được liên kết. Lớp N+1 532 và lớp N 531 cũng có thể được ký hiệu bởi ID lớp. ID lớp là mục dữ liệu mà được liên kết với hình và ký hiệu hình này là một phần của lớp được chỉ báo. Do đó, mỗi hình 511-518 có thể được liên kết với ID lớp tương ứng để chỉ báo lớp nào trong số lớp N+1 532 hoặc lớp N 531 bao gồm hình tương ứng.

Các hình 511-518 trong các lớp khác nhau 531-532 được tạo cấu hình để được hiển thị luân phiên. Như vậy, các hình 511-518 trong các lớp khác nhau 531-532 có thể chia sẻ cùng một bộ nhận dạng (ID) thời gian và có thể được chứa trong cùng một AU. Như được sử dụng ở đây, AU là tập của một hoặc nhiều hình được mã hóa được liên kết với cùng một thời gian hiển thị để xuất ra từ DPB. Ví dụ, bộ giải mã có thể giải mã và hiển thị hình 515 ở thời gian hiển thị hiện thời nếu hình nhỏ hơn được đòi hỏi hoặc bộ giải mã có thể giải mã và hiển thị hình 511 ở thời gian hiển thị hiện thời nếu hình lớn hơn được đòi hỏi. Như vậy, các hình 511-514 ở lớp cao hơn N+1 532 chứa gần như cùng một dữ liệu ảnh như các hình tương ứng 515-518 ở lớp thấp hơn N 531 (bất kể độ chênh lệch về kích thước hình). Cụ thể là, hình 511 chứa gần như cùng một dữ liệu ảnh như hình 515, hình 512 chứa gần như cùng một dữ liệu ảnh như hình 516, v.v..

Các hình 511-518 có thể được mã hóa bằng cách tham chiếu tới các hình khác 511-518 trong cùng một lớp N 531 hoặc N+1 532. Việc mã hóa hình có tham chiếu tới hình khác trong

cùng một lớp dẫn đến kết quả là dự báo liên cấu trúc 523, mà tương thích với dự báo liên cấu trúc một chiều và/hoặc dự báo liên cấu trúc hai chiều. Dự báo liên cấu trúc 523 được thể hiện bởi các mũi tên đường liền nét. Ví dụ, hình 513 có thể được mã hóa bằng cách dùng dự báo liên cấu trúc 523 sử dụng một hoặc hai hình 511, 512, và/hoặc 514 trong lớp $N+1$ 532 làm tham chiếu, trong đó một hình được tham chiếu dành cho dự báo liên cấu trúc một chiều và/hoặc hai hình được tham chiếu dành cho dự báo liên cấu trúc hai chiều. Ngoài ra, hình 517 có thể được mã hóa bằng cách dùng dự báo liên cấu trúc 523 sử dụng một hoặc hai hình 515, 516, và/hoặc 518 trong lớp N 531 làm tham chiếu, trong đó một hình được tham chiếu dành cho dự báo liên cấu trúc một chiều và/hoặc hai hình được tham chiếu dành cho dự báo liên cấu trúc hai chiều. Khi hình được sử dụng làm tham chiếu dành cho hình khác trong cùng một lớp khi thực hiện dự báo liên cấu trúc 523, thì hình này có thể được gọi là hình tham chiếu. Ví dụ, hình 512 có thể là hình tham chiếu được sử dụng để mã hóa hình 513 theo dự báo liên cấu trúc 523. Dự báo liên cấu trúc 523 cũng có thể được gọi là dự báo liên lớp trong ngữ cảnh nhiều lớp. Như vậy, dự báo liên cấu trúc 523 là cơ chế để mã hóa các mẫu của hình hiện thời bằng cách tham chiếu tới các mẫu được chỉ báo trong hình tham chiếu mà khác với hình hiện thời trong đó hình tham chiếu và hình hiện thời ở trong cùng một lớp.

Các hình 511-518 cũng có thể được mã hóa bằng cách tham chiếu tới các hình khác 511-518 trong các lớp khác nhau. Quy trình này được biết đến là dự báo liên lớp 521, và được thể hiện bởi các mũi tên nét đứt. Dự báo liên lớp 521 là cơ chế để mã hóa các mẫu của hình hiện thời bằng cách tham chiếu tới các mẫu được chỉ báo trong hình tham chiếu trong đó hình hiện thời và hình tham chiếu ở trong các lớp khác nhau và do đó có các ID lớp khác nhau. Ví dụ, hình trong lớp thấp hơn N 531 có thể được sử dụng làm hình tham chiếu để mã hóa hình tương ứng ở lớp cao hơn $N+1$ 532. Một ví dụ cụ thể, hình 511 có thể được mã hóa bằng cách tham chiếu tới hình 515 theo dự báo liên lớp 521. Trong trường hợp như vậy, hình 515 được sử dụng làm hình tham chiếu liên lớp. Hình tham chiếu liên lớp là hình tham chiếu được sử dụng dành cho dự báo liên lớp 521. Trong hầu hết các trường hợp, dự báo liên lớp 521 được ràng buộc sao cho hình hiện thời, chẳng hạn như hình 511, có thể chỉ sử dụng (các) hình tham chiếu liên lớp mà được chứa trong cùng một AU và ở tại lớp thấp hơn, chẳng hạn như hình 515. Khi nhiều lớp (ví dụ, nhiều hơn hai lớp) khả dụng, thì dự báo liên lớp 521 có thể lập mã/giải mã hình hiện thời dựa trên nhiều (hình) tham chiếu liên lớp ở các lớp thấp hơn so với hình hiện

thời.

Bộ lập mã video có thể dùng dự báo dựa trên lớp 500 để lập mã các hình 511-518 qua nhiều dạng kết hợp và/hoặc các hoán vị khác nhau của dự báo liên cấu trúc 523 và dự báo liên lớp 521. Ví dụ, hình 515 có thể được mã hóa theo dự báo nội cấu trúc. Các hình 516-518 sau đó có thể được mã hóa theo dự báo liên cấu trúc 523 bằng cách sử dụng hình 515 làm hình tham chiếu. Ngoài ra, hình 511 có thể được mã hóa theo dự báo liên lớp 521 bằng cách sử dụng hình 515 làm hình tham chiếu liên lớp. Các hình 512-514 sau đó có thể được mã hóa theo dự báo liên cấu trúc 523 bằng cách sử dụng hình 511 làm hình tham chiếu. Như vậy, hình tham chiếu có thể phục vụ như cả hình tham chiếu một lớp và hình tham chiếu liên lớp dành cho các cơ chế mã hóa khác nhau. Bằng cách mã hóa các hình ở lớp cao hơn $N+1$ 532 dựa trên các hình ở lớp thấp hơn N 531, lớp cao hơn $N+1$ 532 có thể tránh dùng dự báo nội cấu trúc, mà có hiệu quả mã hóa thấp hơn nhiều so với dự báo liên cấu trúc 523 và dự báo liên lớp 521. Như vậy, hiệu quả mã hóa kém của dự báo nội cấu trúc có thể bị hạn chế với các hình chất lượng nhỏ nhất/thấp nhất, và do đó bị hạn chế với việc mã hóa lượng dữ liệu video nhỏ nhất. Các hình được sử dụng làm các hình tham chiếu và/hoặc các hình tham chiếu liên lớp có thể được chỉ báo trong các mục nhập của (các) danh sách hình tham chiếu được chứa trong cấu trúc danh sách hình tham chiếu.

Mỗi AU 506 trên Fig.5 có thể chứa một số hình. Ví dụ, một AU 506 có thể chứa các hình 511 và 515. AU 506 khác có thể chứa các hình 512 và 516. Thật vậy, mỗi AU 506 là một tập của một hoặc nhiều hình được mã hóa được liên kết với cùng một thời gian hiển thị (ví dụ, cùng một ID thời gian) để xuất ra từ bộ đệm hình được giải mã (DPB) (ví dụ, để hiển thị cho người dùng). Mỗi AUD 508 là một bộ chỉ báo hoặc cấu trúc dữ liệu được sử dụng để chỉ báo điểm bắt đầu của AU (ví dụ, AU 508) hoặc biên giữa các AU.

Các họ mã hóa video H.26x trước đó đã cung cấp sự hỗ trợ dành cho khả năng mở rộng trong (các) đặc tả riêng biệt từ (các) đặc tả để mã hóa một lớp. Mã hóa video có thể mở rộng (SVC) là phần mở rộng có thể mở rộng của AVC/H.264 mà cung cấp sự hỗ trợ dành cho các khả năng mở rộng không gian, thời gian, và chất lượng. Đối với SVC, cờ được báo hiệu trong mỗi khối macro (macroblock, MB) trong các hình EL để chỉ báo việc EL MB có được sử dụng bằng cách sử dụng khối được sắp xếp cùng một chỗ từ lớp thấp hơn hay không. Dự báo từ khối được sắp xếp cùng một chỗ có thể bao gồm bố cục, các vectơ chuyển động, và/hoặc các

chế độ mã hóa. Các dạng thực hiện của SVC không thể trực tiếp sử dụng lại các dạng thực hiện H.264/AVC chưa được cải biến trong thiết kế của chúng. Quy trình giải mã và cú pháp khối macro SVC EL khác với quy trình giải mã và cú pháp H.264/AVC.

HEVC có thể mở rộng (SHVC) là phần mở rộng của chuẩn HEVC/H.265 mà cung cấp sự hỗ trợ dành cho các khả năng mở rộng không gian và chất lượng, HEVC nhiều cảnh nhìn (MV-HEVC) là phần mở rộng của HEVC/H.265 mà cung cấp sự hỗ trợ dành cho khả năng mở rộng nhiều cảnh nhìn, và 3D HEVC (3D-HEVC) là phần mở rộng của HEVC/H.264 mà cung cấp sự hỗ trợ dành cho việc mã hóa video ba chiều (3D) cao cấp và hiệu quả hơn so với MV-HEVC. Lưu ý rằng khả năng mở rộng thời gian được coi là một phần không thể thiếu trong bộ mã hóa/giải mã HEVC một lớp. Thiết kế của phần mở rộng nhiều lớp của HEVC dùng ý tưởng mà các hình được giải mã được sử dụng dành cho dự báo liên lớp chỉ đến từ cùng một đơn vị truy nhập (AU) và được xử lý như các hình tham chiếu dài hạn (long-term reference picture, LTRP), và được gán các chỉ số tham chiếu trong (các) danh sách hình tham chiếu cùng với các hình tham chiếu thời gian khác trong lớp hiện thời. Dự báo liên lớp (ILP) đạt được ở mức đơn vị dự báo (prediction unit, PU) bằng cách thiết đặt giá trị của chỉ số tham chiếu để tham chiếu tới (các) hình tham chiếu liên lớp trong (các) danh sách hình tham chiếu.

Đáng chú ý là, cả hai tính năng lấy mẫu lại hình tham chiếu và mở rộng không gian đều yêu cầu việc lấy mẫu lại của hình tham chiếu hoặc một phần của nó. Lấy mẫu lại hình tham chiếu (reference picture resampling, RPR) có thể được thực hiện ở mức hình hoặc mức khối mã hóa. Tuy nhiên, khi RPR được gọi là tính năng mã hóa, thì nó là tính năng dành cho việc mã hóa một lớp. Ngay cả như vậy, thì có thể hoặc thậm chí phù hợp hơn từ quan điểm thiết kế bộ mã hóa/giải mã để sử dụng cùng một bộ lọc lấy mẫu lại dành cho cả tính năng RPR để mã hóa một lớp và tính năng mở rộng không gian để mã hóa nhiều lớp.

Fig.6 là hình vẽ thể hiện một ví dụ về dự báo dựa trên lớp 600 dùng các tập lớp đầu ra (OLS), ví dụ như được thực hiện để xác định các MV ở bước nén khối 105, bước giải mã khối 113, thành phần ước lượng chuyển động 221, thành phần bù chuyển động 219, thành phần bù chuyển động 321, và/hoặc thành phần bù chuyển động 421. Dự báo dựa trên lớp 500 tương thích với dự báo liên cấu trúc một chiều và/hoặc dự báo liên cấu trúc hai chiều, nhưng cũng được thực hiện giữa các hình trong các lớp khác nhau. Dự báo dựa trên lớp trên Fig.6 tương tự với Fig.5. Do đó, để cho ngắn gọn, thì phần mô tả đầy đủ về dự báo dựa trên lớp không

được lặp lại.

Một số lớp trong số các lớp trong trình tự video được mã hóa (coded video sequence, CVS) 690 trên Fig.6 được chứa trong OLS. OLS là một tập của các lớp mà một hoặc nhiều lớp được quy định là các lớp đầu ra. Lớp đầu ra là một lớp của OLS mà được xuất ra. Fig.6 thể hiện ba OLS khác nhau, cụ thể là OLS 1, OLS 2, và OLS 3. Như được thể hiện, OLS 1 bao gồm lớp N 631 và lớp N+1 632. OLS 2 bao gồm lớp N 631, lớp N+1 632, lớp N+2 633, và lớp N+1 634. OLS 3 bao gồm lớp N 631, lớp N+1 632, và lớp N+2 633. Mặc dù ba OLS được thể hiện, nhưng số lượng khác nhau của các OLS có thể được sử dụng trong các ứng dụng thực tế.

Mỗi OLS trong số các OLS khác nhau có thể chứa số lượng lớp bất kỳ. Các OLS khác nhau được tạo ra với nỗ lực để phù hợp với các khả năng mã hóa của các thiết bị khác nhau có các khả năng mã hóa khác nhau. Ví dụ, OLS 1, mà chỉ chứa hai lớp, có thể được tạo ra để phù hợp với điện thoại di động với các khả năng mã hóa tương đối hạn chế. Mặt khác, OLS 2, mà chứa bốn lớp, có thể được tạo ra để phù hợp với máy thu hình màn hình lớn, mà có thể giải mã các lớp cao hơn so với điện thoại di động. OLS 3, mà chứa ba lớp, có thể được tạo ra để phù hợp với máy tính cá nhân, máy tính xách tay, hoặc máy tính bảng, mà có thể giải mã các lớp cao hơn so với điện thoại di động nhưng không thể giải mã các lớp cao hơn như máy thu hình màn hình lớn.

Các lớp trên Fig.6 có thể độc lập với nhau. Tức là, mỗi lớp có thể được giải mã mà không sử dụng dự báo liên lớp (ILP). Trong trường hợp này, các lớp này được gọi là các lớp đồng thời truyền đi qua đài phát thanh và đài truyền hình. Một hoặc nhiều lớp trên Fig.6 cũng có thể được mã hóa sử dụng ILP. Việc các lớp này có phải là các lớp đồng thời truyền đi qua đài phát thanh và đài truyền hình hay không hoặc việc một số lớp trong số các lớp này có được mã hóa bằng cách sử dụng ILP hay không được báo hiệu bởi cờ trong tập thông số video (VPS), mà sẽ được thảo luận đầy đủ hơn bên dưới. Khi một số lớp sử dụng ILP, thì mối quan hệ phụ thuộc lớp giữa các lớp cũng được báo hiệu trong VPS.

Theo một phương án, khi các lớp này là các lớp đồng thời truyền đi qua đài phát thanh và đài truyền hình, thì chỉ một lớp được chọn lựa để giải mã và xuất ra. Theo một phương án, khi một số lớp sử dụng ILP, tất cả các lớp (ví dụ, toàn bộ luồng bit) được quy định sẽ được giải mã, và các lớp nhất định trong số các lớp được quy định là các lớp đầu ra. Lớp đầu ra

hoặc các lớp có thể là, ví dụ, 1) chỉ duy nhất lớp cao nhất, 2) tất cả các lớp, hoặc 3) lớp cao nhất cộng tập các lớp thấp hơn được chỉ báo. Ví dụ, khi lớp cao nhất cộng tập các lớp thấp hơn được chỉ báo được chỉ định dành cho đầu ra bởi cờ trong VPS, thì lớp N+3 634 (mà là lớp cao nhất) và các lớp N 631 và N+1 632 (mà là các lớp thấp hơn) từ OLS 2 được xuất ra.

Tham khảo Fig.6, một số lớp được chứa trong OLS, nhưng không được xuất ra hoặc được sử dụng để tham chiếu bởi lớp đầu ra. Ví dụ, lớp N + 2 633 được chứa trong OLS 2, nhưng không được xuất ra. Thay vào đó, lớp N+3 634 là lớp cao nhất trong OLS 2 và, do đó, được xuất ra như lớp đầu ra. Ngoài ra, lớp N + 2 633 trong OLS 2 không được sử dụng để tham chiếu bởi lớp đầu ra, lớp N+3 634. Thật vậy, lớp N + 3 634 trong OLS 2 dựa vào lớp N + 2 633, lớp N+1 632, và/hoặc lớp N 631 trong OLS 2. Do đó, lớp N+2 633 trong OLS 2 được gọi là lớp không cần thiết. Không may là, SHVC và MV-HEVC cho phép các lớp không cần thiết như vậy được chứa trong các luồng bit video nhiều lớp, mà tạo gánh nặng không cần thiết cho các tài nguyên mã hóa và làm giảm hiệu quả mã hóa.

Được bộc lộ ở đây là các kỹ thuật mà không cho phép các lớp không cần thiết trong các luồng bit video nhiều lớp. Tức là, mọi lớp được chứa trong tập lớp đầu ra (OLS) hoặc là lớp đầu ra hoặc được sử dụng như lớp tham chiếu trực tiếp hoặc gián tiếp của lớp đầu ra. Việc này tránh có thông tin ngoại lai trong quy trình mã hóa và làm tăng hiệu quả mã hóa. Do đó, bộ mã hóa/bộ giải mã (còn được gọi là, “bộ mã hóa/giải mã”) khi mã hóa video được cải thiện so với các bộ mã hóa/giải mã hiện thời. Một vấn đề thực tế, quy trình mã hóa video được cải thiện đem lại cho người dùng trải nghiệm người dùng tốt hơn khi các video được gửi, được thu, và/hoặc được xem.

Fig.7 là hình vẽ thể hiện một phương án về luồng bit video 700. Như được sử dụng ở đây luồng bit video 700 cũng có thể được gọi là luồng bit video được mã hóa, luồng bit, hoặc các biến thể khác của nó. Như được thể hiện trên Fig.7, luồng bit 700 bao gồm ít nhất một đơn vị hình (PU) 701. Trong khi ba trong số các PU 701 được thể hiện trên Fig.7, thì một số lượng khác nhau của các PU 701 có thể có mặt trong luồng bit 700 trong các ứng dụng thực tế. Mỗi PU 701 là một tập các đơn vị NAL mà được liên kết với nhau theo quy tắc phân loại được quy định, liên tiếp theo thứ bậc giải mã, và chứa chính xác một hình được giải mã (ví dụ, hình 714).

Theo một phương án, mỗi PU 701 chứa một hoặc nhiều thông số sau đây: thông tin khả

năng giải mã (decoding capability information, DCI) 702, tập thông số video (VPS) 704, tập thông số trình tự (SPS) 706, tập thông số hình (PPS) 708, phần đầu hình (picture header, PH) 712, và hình 714. Mỗi thông số trong số DCI 702, VPS 704, SPS 706, và PPS 708 có thể được gọi chung là tập thông số. Theo một phương án, các tập thông số khác không được thể hiện trên Fig.7 cũng có thể được chứa trong luồng bit 700 chẳng hạn như, ví dụ, tập thông số thích nghi (adaption parameter set, APS), mà là cấu trúc cú pháp chứa các phần tử cú pháp mà áp dụng cho không hoặc nhiều phần chia như được xác định bởi không hoặc nhiều phần tử cú pháp được tìm thấy trong các phần đầu phần chia.

DCI 702, mà cũng có thể được gọi là tập thông số giải mã (decoding parameter set, DPS) hoặc tập thông số bộ giải mã, là cấu trúc cú pháp chứa các phần tử cú pháp mà áp dụng cho toàn bộ luồng bit. DCI 702 bao gồm các thông số mà không đổi trong suốt vòng đời của luồng bit video (ví dụ, luồng bit 700), mà có thể chuyển thành vòng đời của một phiên. DCI 702 có thể bao gồm thông tin đặc tả, mức, và đặc tả con để xác định điểm tương tác có độ phức tạp tối đa mà được đảm bảo không được vượt quá, ngay cả nếu sự ghép nối của các trình tự video xảy ra trong một phiên. Còn tùy chọn bao gồm các cờ ràng buộc, mà chỉ báo rằng luồng bit video sẽ bị ràng buộc việc sử dụng các tính năng nhất định như được chỉ báo bởi các giá trị của các cờ. Với điều này, luồng bit có thể được gắn nhãn là không sử dụng các công cụ nhất định, mà cho phép sự phân bổ tài nguyên giữa các thành phần khác theo dạng thực hiện bộ giải mã. Giống như tất cả các tập thông số, DCI 702 có mặt khi được tham chiếu thứ nhất, và được tham chiếu bởi hình thứ nhất trong trình tự video, ngụ ý rằng nó đã được gửi trong số các đơn vị NAL thứ nhất trong luồng bit. Trong khi nhiều DCI 702 có thể ở trong luồng bit, thì giá trị của các phần tử cú pháp trong đó không thể nhất quán khi được tham chiếu.

VPS 704 bao gồm độ phụ thuộc giải mã hoặc thông tin dành cho việc cấu thành tập hình tham chiếu của các lớp tăng cường. VPS 704 cung cấp quan điểm hoặc tầm nhìn tổng thể về trình tự có thể mở rộng, bao gồm kiểu kiểm hoạt động nào được cung cấp, đặc tả, tầng, và mức của các điểm hoạt động, và một số các thuộc tính mức cao khác của luồng bit mà có thể được sử dụng như cơ sở dành cho việc thỏa thuận phiên và chọn lựa nội dung, v.v..

Theo một phương án, khi được chỉ báo rằng một số lớp sử dụng ILP, thì VPS 704 chỉ báo rằng tổng số lượng các OLS được quy định bởi VPS bằng với số lượng các lớp, chỉ báo rằng OLS thứ i bao gồm các lớp với các chỉ số lớp từ 0 đến i , gồm cả 0 và i , và chỉ báo rằng

đối với mỗi OLS chỉ lớp cao nhất trong OLS được xuất ra.

Theo một phương án, VPS 704 bao gồm cú pháp và ngữ nghĩa tương ứng với các OLS trong CLVS và/hoặc luồng bit video. Cú pháp và ngữ nghĩa sau đây tương ứng với VPS 704 có thể được dùng để thực hiện các phương án được bộc lộ ở đây.

Cú pháp dành cho VPS 704 có thể như sau.

video_parameter_set_rbsp() {	Bộ mô tả
vps_video_parameter_set_id	u(4)
vps_max_layers_minus1	u(6)
vps_max_sub_layers_minus1	u(3)
if(vps_max_layers_minus1 > 0)	
vps_all_independent_layers_flag	u(1)
for(i = 0; i <= vps_max_layers_minus1; i++) {	
vps_layer_id[i]	u(6)
if(i > 0 && !vps_all_independent_layers_flag) {	
vps_independent_layer_flag[i]	u(1)
if(!vps_independent_layer_flag[i])	
for(j = 0; j < i; j++)	
vps_direct_dependency_flag[i][j]	u(1)
}	
}	
if(vps_max_layers_minus1 > 0) {	
if(vps_all_independent_layers_flag)	
each_layer_is_an_ols_flag	u(1)
if(!each_layer_is_an_ols_flag) {	
if(!vps_all_independent_layers_flag)	
ols_mode_idc	u(2)
if(ols_mode_idc == 2) {	
num_output_layer_sets_minus1	u(8)

for(i = 1; i < num_output_layer_sets_minus1 + 1;	
i++) {	
for(j = 0; j <= vps_max_layers_minus1; j++	
)	
layer_included_flag[i][j]	u(1)
if(!vps_all_independent_layers_flag)	
for(j = 0; j < NumLayersInOls[i] - 1;	
j++)	
vps_output_layer_flag[i][j]	u(1)
}	
}	
}	
}	
vps_constraint_info_present_flag	u(1)
vps_reserved_zero_7bits	u(7)
if(vps_constraint_info_present_flag)	
general_constraint_info()	
general_hrd_params_present_flag	u(1)
if(general_hrd_params_present_flag) {	
num_units_in_tick	u(32)
time_scale	u(32)
general_hrd_parameters()	
}	
vps_extension_flag	u(1)
if(vps_extension_flag)	
while(more_rbsp_data())	
vps_extension_data_flag	u(1)
rbp_trailing_bits()	
}	

Ngữ nghĩa dành cho VPS 704 có thể như sau. Theo một phương án, VPS 704 bao gồm một hoặc nhiều cờ và thông số được ký hiệu bên dưới.

Tải trình tự byte thô (RBSP) VPS sẽ khả dụng dành cho quy trình giải mã trước khi nó được tham chiếu, được chứa trong ít nhất một đơn vị truy nhập với TemporalId bằng 0 hoặc được cung cấp thông qua các phương tiện bên ngoài, và đơn vị VPS NAL chứa VPS RBSP sẽ có nuh_layer_id bằng vps_layer_id[0]. nuh_layer_id quy định bộ nhận dạng của lớp mà đơn vị VCL NAL thuộc về hoặc bộ nhận dạng của lớp mà đơn vị NAL không phải VCL áp dụng. Giá trị của nuh_layer_id sẽ nằm trong khoảng từ 0 đến 55, bao gồm cả 0 và 55. TemporalId là ID được sử dụng để nhận dạng duy nhất đơn vị NAL cụ thể so với các đơn vị NAL khác. Giá trị của TemporalId sẽ giống nhau đối với tất cả các đơn vị VCL NAL của AU. Giá trị của TemporalId của hình được mã hóa, PU, hoặc AU là giá trị của TemporalId của các đơn vị VCL NAL của hình được mã hóa, PU, hoặc AU.

Tất cả các đơn vị VPS NAL với giá trị cụ thể của vps_video_parameter_set_id trong CVS sẽ có cùng một nội dung. vps_video_parameter_set_id cung cấp bộ nhận dạng dành cho VPS để tham chiếu bởi các phần tử cú pháp khác. vps_max_layers_minus1 cộng 1 quy định số lượng được cho phép tối đa của các lớp trong mỗi CVS tham chiếu tới VPS. vps_max_sub_layers_minus1 cộng 1 quy định số lượng tối đa của các lớp con thời gian mà có thể có mặt trong mỗi CVS tham chiếu tới VPS. Giá trị của vps_max_sub_layers_minus1 sẽ nằm trong khoảng từ 0 đến 6, bao gồm cả 0 và 6.

vps_all_independent_layers_flag bằng 1 quy định rằng tất cả các lớp trong CVS được mã hóa độc lập mà không sử dụng dự báo liên lớp. vps_all_independent_layers_flag bằng 0 quy định rằng một hoặc nhiều lớp trong số các lớp trong CVS có thể sử dụng dự báo liên lớp. Khi không có mặt, giá trị của vps_all_independent_layers_flag được suy ra là bằng 1. Khi vps_all_independent_layers_flag bằng 1, thì giá trị của vps_independent_layer_flag[i] được suy ra là bằng 1. Khi vps_all_independent_layers_flag bằng 0, thì giá trị của vps_independent_layer_flag[0] được suy ra là bằng 1.

vps_layer_id[i] quy định nuh_layer_id giá trị của lớp thứ i. Đối với hai giá trị nguyên không âm bất kỳ của m và n, khi m nhỏ hơn n, thì giá trị của vps_layer_id[m] sẽ nhỏ hơn vps_layer_id[n]. vps_independent_layer_flag[i] bằng 1 quy định rằng lớp với chỉ số i không

sử dụng dự báo liên lớp. $vps_independent_layer_flag[i]$ bằng 0 quy định rằng lớp với chỉ số i có thể sử dụng dự báo liên lớp và $vps_layer_dependency_flag[i]$ có mặt trong VPS. Khi không có mặt, thì giá trị của $vps_independent_layer_flag[i]$ được suy ra là bằng 1.

$vps_direct_dependency_flag[i][j]$ bằng 0 quy định rằng lớp với chỉ số j không phải là lớp tham chiếu trực tiếp dành cho lớp với chỉ số i . $vps_direct_dependency_flag[i][j]$ bằng 1 quy định rằng lớp với chỉ số j là lớp tham chiếu trực tiếp dành cho lớp với chỉ số i . Khi $vps_direct_dependency_flag[i][j]$ không có mặt đối với i và j nằm trong khoảng từ 0 đến $vps_max_layers_minus1$, bao gồm cả 0 và $vps_max_layers_minus1$, thì nó được suy ra là bằng 0.

Biến $DirectDependentLayerIdx[i][j]$, quy định lớp phụ thuộc trực tiếp thứ j của lớp thứ i , được suy ra như sau.

```
for( i = 1; i < vps_max_layers_minus1; i++ )
    if( !vps_independent_layer_flag[i] )
        for( j = i, k = 0; j >= 0; j-- )
            if( vps_direct_dependency_flag[i][j] )
                DirectDependentLayerIdx[i][k++] = j
```

Biến $GeneralLayerIdx[i]$, quy định chỉ số lớp của lớp với nuh_layer_id bằng $vps_layer_id[i]$, được suy ra như sau.

```
for( i = 0; i <= vps_max_layers_minus1; i++ )
    GeneralLayerIdx[vps_layer_id[i]] = i
```

$each_layer_is_an_ols_flag$ bằng 1 quy định rằng mỗi tập lớp đầu ra chỉ chứa một lớp và chính mỗi lớp trong luồng bit là tập lớp đầu ra với một lớp được chứa là lớp đầu ra duy nhất. $each_layer_is_an_ols_flag$ bằng 0 quy định rằng tập lớp đầu ra có thể chứa nhiều hơn một lớp. Nếu $vps_max_layers_minus1$ bằng 0, thì giá trị của $each_layer_is_an_ols_flag$ được suy ra là bằng 1. Nếu không, khi $vps_all_independent_layers_flag$ bằng 0, thì giá trị của $each_layer_is_an_ols_flag$ được suy ra là bằng 0.

ols_mode_idc bằng 0 quy định rằng tổng số lượng của các OLS được quy định bởi VPS bằng $vps_max_layers_minus1 + 1$, OLS thứ i bao gồm các lớp với các chỉ số lớp từ 0 đến i , bao gồm cả 0 và i , và đối với mỗi OLS chỉ lớp cao nhất trong OLS được xuất ra. ols_mode_idc bằng 1 quy định rằng tổng số lượng của các OLS được quy định bởi VPS bằng

$vps_max_layers_minus1 + 1$, OLS thứ i bao gồm các lớp với các chỉ số lớp từ 0 đến i , bao gồm cả 0 và i , và đối với mỗi OLS tất cả các lớp trong OLS được xuất ra. ols_mode_idc bằng 2 quy định rằng tổng số lượng của các OLS được quy định bởi VPS được báo hiệu rõ ràng và đối với mỗi OLS lớp cao nhất và tập được báo hiệu rõ ràng của các lớp thấp hơn trong OLS được xuất ra. Giá trị của ols_mode_idc sẽ nằm trong khoảng từ 0 đến 2, bao gồm cả 0 và 2. Giá trị 3 của ols_mode_idc được dành riêng để sử dụng trong tương lai bởi ITU-T | ISO/IEC. Khi $vps_all_independent_layers_flag$ bằng 1 và $each_layer_is_an_ols_flag$ bằng 0, thì giá trị của ols_mode_idc được suy ra là bằng 2.

$num_output_layer_sets_minus1$ cộng 1 quy định tổng số lượng của các OLS được quy định bởi VPS khi ols_mode_idc bằng 2.

Biến $TotalNumOlss$, quy định tổng số lượng của các OLS được quy định bởi VPS, được suy ra như sau:

```
if( vps_max_layers_minus1 == 0 )
    TotalNumOlss = 1
else if( each_layer_is_an_ols_flag || ols_mode_idc == 0 || ols_mode_idc == 1 )
    TotalNumOlss = vps_max_layers_minus1 + 1
else if( ols_mode_idc == 2 )
    TotalNumOlss = num_output_layer_sets_minus1 + 1
```

$layer_included_flag[i][j]$ quy định việc lớp thứ j (tức là, lớp với nuh_layer_id bằng $vps_layer_id[j]$) được chứa trong OLS thứ i khi ols_mode_idc bằng 2. $layer_included_flag[i][j]$ bằng 1 quy định rằng lớp thứ j được chứa trong OLS thứ i . $layer_included_flag[i][j]$ bằng 0 quy định lớp thứ j không được chứa trong OLS thứ i .

Biến $NumLayersInOls[i]$, quy định số lượng của các lớp trong OLS thứ i , và biến $LayerIdInOls[i][j]$, quy định nuh_layer_id giá trị của lớp thứ j trong OLS thứ i , được suy ra như sau.

```
NumLayersInOls[ 0 ] = 1
LayerIdInOls[ 0 ][ 0 ] = vps_layer_id[ 0 ]
for( i = 1, i < TotalNumOlss; i++ ) {
    if( each_layer_is_an_ols_flag ) {
        NumLayersInOls[ i ] = 1
```

```

    LayerIdInOls[ i ][ 0 ] = vps_layer_id[ i ]
} else if( ols_mode_idc == 0 || ols_mode_idc == 1 ) {
    NumLayersInOls[ i ] = i + 1
    for( j = 0; j < NumLayersInOls[ i ]; j++ )
        LayerIdInOls[ i ][ j ] = vps_layer_id[ j ]
} else if( ols_mode_idc == 2 ) {
    for( k = 0, j = 0; k <= vps_max_layers_minus1; k++ )
        if( layer_included_flag[ i ][ k ] )
            LayerIdInOls[ i ][ j++ ] = vps_layer_id[ k ]
    NumLayersInOls[ i ] = j
}

```

Biến `OlsLayerIdx[ i ][ j ]`, quy định chỉ số lớp OLS của lớp với `nuh_layer_id` bằng `LayerIdInOls[ i ][ j ]`, được suy ra như sau.

```

for( i = 0, i < TotalNumOlss; i++ )
    for j = 0; j < NumLayersInOls[ i ]; j++ )
        OlsLayerIdx[ i ][ LayerIdInOls[ i ][ j ] ] = j

```

Lớp thấp nhất trong mỗi OLS sẽ là lớp độc lập. Nói cách khác, đối với mỗi i trong khoảng từ 0 đến $TotalNumOlss - 1$, bao gồm cả 0 và $TotalNumOlss - 1$, thì giá trị của `vps_independent_layer_flag[ GeneralLayerIdx[ LayerIdInOls[ i ][ 0 ] ] ]` sẽ bằng 1.

Mỗi lớp sẽ được chứa trong ít nhất một OLS được quy định bởi VPS. Tức là, đối với mỗi lớp với giá trị cụ thể của `nuh_layer_id` `nuhLayerId` bằng một giá trị trong số `vps_layer_id[ k ]` đối với k nằm trong khoảng từ 0 đến `vps_max_layers_minus1`, bao gồm cả 0 và `vps_max_layers_minus1`, sẽ có ít nhất một cặp giá trị của i và j , trong đó i nằm trong khoảng từ 0 đến $TotalNumOlss - 1$, bao gồm cả 0 và $TotalNumOlss - 1$, và j nằm trong khoảng $NumLayersInOls[i] - 1$, bao gồm cả $NumLayersInOls[i] - 1$, sao cho giá trị của `LayerIdInOls[ i ][ j ]` bằng `nuhLayerId`.

Lớp bất kỳ trong OLS sẽ là lớp đầu ra của OLS hoặc lớp tham chiếu (trực tiếp hoặc gián tiếp) của lớp đầu ra của OLS.

`vps_output_layer_flag[ i ][ j ]` quy định việc lớp thứ j trong OLS thứ i có được xuất ra hay không khi `ols_mode_idc` bằng 2. `vps_output_layer_flag[ i ]` bằng 1 quy định rằng lớp thứ

j trong OLS thứ i được xuất ra. $vps_output_layer_flag[i]$ bằng 0 quy định rằng lớp thứ j trong OLS thứ i không được xuất ra. Khi $vps_all_independent_layers_flag$ bằng 1 và $each_layer_is_an_ols_flag$ bằng 0, thì giá trị của $vps_output_layer_flag[i]$ được suy ra là bằng 1.

Biến $OutputLayerFlag[i][j]$, mà giá trị 1 quy định rằng lớp thứ j trong OLS thứ i được xuất ra và giá trị 0 quy định rằng lớp thứ j trong OLS thứ i không được xuất ra, được suy ra như sau.

```
for( i = 0, i < TotalNumOlss; i++ ) {
    OutputLayerFlag[ i ][ NumLayersInOls[ i ] - 1 ] = 1
    for( j = 0; j < NumLayersInOls[ i ] - 1; j++ )
        if( ols_mode_idc[ i ] == 0 )
            OutputLayerFlag[ i ][ j ] = 0
        else if( ols_mode_idc[ i ] == 1 )
            OutputLayerFlag[ i ][ j ] = 1
        else if( ols_mode_idc[ i ] == 2 )
            OutputLayerFlag[ i ][ j ] = vps_output_layer_flag[ i ][ j ]
}
```

Lưu ý – OLS thứ 0 chỉ chứa lớp thấp nhất (tức là, lớp với nuh_layer_id bằng $vps_layer_id[0]$) và đối với OLS thứ 0 lớp được chứa duy nhất được xuất ra.

$vps_constraint_info_present_flag$ bằng 1 quy định rằng cấu trúc cú pháp $general_constraint_info()$ có mặt trong VPS. $vps_constraint_info_present_flag$ bằng 0 quy định rằng cấu trúc cú pháp $general_constraint_info()$ không có mặt trong VPS.

$vps_reserved_zero_7bits$ sẽ bằng 0 trong các luồng bit phù hợp với phiên bản của bản nháp VVC này. Các giá trị khác đối với $vps_reserved_zero_7bits$ được dành riêng để sử dụng trong tương lai bởi ITU-T | ISO/IEC. Theo một phương án, các bộ giải mã sẽ loại bỏ giá trị của $vps_reserved_zero_7bits$.

$general_hrd_params_present_flag$ bằng 1 quy định rằng các phần tử cú pháp $num_units_in_tick$ và $time_scale$ và cấu trúc cú pháp $general_hrd_parameters()$ có mặt trong cấu trúc cú pháp SPS RBSP. $general_hrd_params_present_flag$ bằng 0 quy định rằng các phần tử cú pháp $num_units_in_tick$ và $time_scale$ và cấu trúc cú pháp $general_hrd_parameters()$

không có mặt trong cấu trúc cú pháp SPS RBSP.

`num_units_in_tick` là số lượng các đơn vị thời gian của đồng hồ hoạt động ở tần số `time_scale` Hz mà tương ứng với một số gia (được gọi là nhịp đồng hồ) của bộ đếm nhịp đồng hồ. `num_units_in_tick` sẽ lớn hơn 0. Nhịp đồng hồ, theo đơn vị giây, bằng thương của `num_units_in_tick` chia cho `time_scale`. Ví dụ, khi tỷ lệ hình của tín hiệu video là 25Hz, thì `time_scale` có thể bằng 27,000,000 và `num_units_in_tick` có thể bằng 1,080,000, và do đó nhịp đồng hồ có thể bằng 0,04 giây.

`time_scale` là số lượng các đơn vị thời gian trôi qua trong một giây. Ví dụ, hệ thống tọa độ thời gian mà đo thời gian sử dụng đồng hồ 27MHz có `time_scale` là 27,000,000. Giá trị của `time_scale` sẽ lớn hơn 0.

`vps_extension_flag` bằng 0 quy định rằng không có các phần tử cú pháp `vps_extension_data_flag` nào có mặt trong cấu trúc cú pháp VPS RBSP. `vps_extension_flag` bằng 1 quy định rằng có các phần tử cú pháp `vps_extension_data_flag` có mặt trong cấu trúc cú pháp VPS RBSP.

`vps_extension_data_flag` có thể có giá trị bất kỳ. Sự có mặt và giá trị nó không ảnh hưởng tới sự phù hợp bộ giải mã với các đặc tả được quy định trong phiên bản của phần mô tả kỹ thuật này. Các bộ giải mã phù hợp với phiên bản của phần mô tả kỹ thuật mà sẽ loại bỏ tất cả các phần tử cú pháp `vps_extension_data_flag`.

SPS 706 chứa dữ liệu chung cho tất cả các hình theo trình tự của các hình (sequence of picture, SOP). SPS 706 là cấu trúc cú pháp chứa các phần tử cú pháp mà áp dụng cho không hoặc nhiều toàn bộ CLVS như được xác định bởi nội dung của phần tử cú pháp được tìm thấy trong PPS được tham chiếu tới bởi phần tử cú pháp được tìm thấy trong mỗi phần đầu hình. Ngược lại, PPS 708 chứa dữ liệu chung cho toàn bộ hình. PPS 708 là cấu trúc cú pháp chứa các phần tử cú pháp mà áp dụng cho không hoặc nhiều toàn bộ các hình được giải mã như được xác định bởi phần tử cú pháp được tìm thấy trong mỗi phần đầu hình (ví dụ, PH 712).

DCI 702, VPS 704, SPS 706, và PPS 708 được chứa trong các kiểu khác nhau của các đơn vị lớp trừu tượng mạng (NAL). Đơn vị NAL là cấu trúc cú pháp chứa thông tin chỉ báo về kiểu của dữ liệu để tuân theo (ví dụ, dữ liệu video được mã hóa). Các đơn vị NAL được phân loại thành lớp mã hóa video (VCL) và các đơn vị NAL không phải VCL. Các đơn vị VCL NAL chứa dữ liệu mà thể hiện các giá trị của các mẫu trong các hình video, và các đơn

vị NAL không phải VCL chứa thông tin bổ sung được liên kết bất kỳ chẳng hạn như các tập thông số (dữ liệu quan trọng mà có thể áp dụng cho một số lượng các đơn vị VCL NAL) và thông tin tăng cường phụ (thông tin định thời và dữ liệu phụ khác mà có thể tăng cường khả năng sử dụng của tín hiệu video được giải mã nhưng không cần thiết để giải mã các giá trị của các mẫu trong các hình video).

Theo một phương án, DCI 702 được chứa trong đơn vị NAL không phải VCL được chỉ định như đơn vị DCI NAL hoặc đơn vị DPS NAL. Tức là, đơn vị DCI NAL có kiểu đơn vị DCI NAL (NUT) và đơn vị DPS NAL có DPS NUT. Theo một phương án, VPS 704 được chứa trong đơn vị NAL không phải VCL được chỉ định như đơn vị VPS NAL. Do đó, đơn vị VPS NAL có VPS NUT. Theo một phương án, SPS 706 là đơn vị NAL không phải VCL được chỉ định như đơn vị SPS NAL. Do đó, đơn vị SPS NAL có SPS NUT. Theo một phương án, PPS 708 được chứa trong đơn vị NAL không phải VCL được chỉ định như đơn vị PPS NAL. Do đó, đơn vị PPS NAL có PPS NUT.

PH 712 là cấu trúc cú pháp chứa các phần tử cú pháp mà áp dụng cho tất cả các phần chia (ví dụ, các phần chia 718) của hình được giải mã (ví dụ, hình 714). Theo một phương án, PH 712 ở trong kiểu mới của đơn vị NAL không phải VCL được chỉ định như đơn vị PH NAL. Do đó, đơn vị PH NAL có PH NUT (ví dụ, PH_NUT). Theo một phương án, có một và chỉ một PH 712 được chứa trong mỗi PU 701. Tức là, PU 701 chứa một PH 712. Theo một phương án, chính xác một đơn vị PH NAL có mặt trong mỗi hình 701 trong luồng bit 700.

Theo một phương án, đơn vị PH NAL được liên kết với PH 712 có ID thời gian và ID lớp. Bộ nhận dạng ID thời gian chỉ báo vị trí của đơn vị PH NAL, theo thời gian, so với các đơn vị PH NAL khác trong luồng bit (ví dụ, luồng bit 701). ID lớp chỉ báo lớp (ví dụ, lớp 531 hoặc lớp 532) mà chứa đơn vị PH NAL. Theo một phương án, ID thời gian tương tự, nhưng khác với, POC. POC nhận dạng duy nhất mỗi hình theo thứ bậc. Trong luồng bit một lớp, ID thời gian và POC sẽ giống nhau. Trong luồng bit nhiều lớp (ví dụ, xem Fig.5), các hình trong cùng một AU sẽ có các POC khác nhau, nhưng có ID thời gian giống nhau.

Theo một phương án, đơn vị PH NAL ở trước đơn vị VCL NAL chứa phần chia 718 thứ nhất của hình được liên kết 714. Việc này thiết lập sự liên kết giữa PH 712 và các phần chia 718 của hình 714 được liên kết với PH 712 mà không cần phải có ID phần đầu hình được báo hiệu trong PH 712 và được tham chiếu từ phần đầu phần chia 720. Do đó, có thể được suy ra

rằng tất cả các đơn vị VCL NAL giữa hai PH 712 thuộc về cùng một hình 714 và rằng hình 714 được liên kết với PH 712 thứ nhất giữa hai PH 712. Theo một phương án, đơn vị VCL NAL thứ nhất mà theo sau PH 712 chứa phần chia 718 thứ nhất của hình 714 được liên kết với PH 712.

Theo một phương án, đơn vị PH NAL theo sau các tập thông số mức hình (ví dụ, PPS) hoặc các tập thông số mức cao hơn chẳng hạn như DCI (được biết đến là DPS), VPS, SPS, PPS, v.v. có cả ID thời gian và ID lớp lần lượt nhỏ hơn so với ID thời gian và ID lớp của đơn vị PH NAL. Do đó, các tập thông số này không được lặp lại trong hình hoặc đơn vị truy nhập. Vì thứ bậc này, nên PH 712 có thể được phân giải ngay lập tức. Tức là, các tập thông số mà chứa các thông số liên quan đến toàn bộ hình được định vị trong luồng bit trước đơn vị PH NAL. Bất cứ thành phần nào mà chứa các thông số dành cho một phần của hình được định vị sau đơn vị PH NAL.

Theo một phương án thay thế, đơn vị PH NAL theo sau các tập thông số mức hình và các bản tin thông tin tăng cường phụ (supplemental enhancement information, SEI) tiền tố, hoặc các tập thông số mức cao hơn chẳng hạn như DCI (được biết đến là DPS), VPS, SPS, PPS, APS, bản tin SEI, v.v..

Hình 714 là một mảng của các mẫu độ sáng theo định dạng đơn sắc hoặc một mảng của các mẫu độ sáng và hai mảng tương ứng của các mẫu sắc độ theo định dạng màu sắc 4:2:0, 4:2:2, và 4:4:4. Theo một phương án, có một và chỉ một hình 714 được chứa trong mỗi PU 701. Như vậy, chỉ có một PH 712 và chỉ một hình 714 tương ứng với PH 712 trong mỗi PU 701. Tức là, PU 701 chứa một hình 714.

Hình 714 có thể là khung hoặc trường. Tuy nhiên, trong một CVS 716, hoặc là tất cả các hình 714 là các khung hoặc là tất cả các hình 714 là các trường. CVS 716 là trình tự video được mã hóa dành cho mọi trình tự video lớp được mã hóa (CLVS) trong luồng bit video 700. Đáng chú ý là, CVS 716 và CLVS giống nhau khi luồng bit video 700 bao gồm một lớp. CVS 716 và CLVS chỉ khác nhau khi luồng bit video 700 bao gồm nhiều lớp (ví dụ, như được thể hiện trên Fig.5 và Fig.6).

Mỗi hình 714 chứa một hoặc nhiều phần chia 718. Phần chia 718 là số lượng nguyên các khung lát đầy đủ hoặc số lượng nguyên các hàng CTU đầy đủ liên tiếp trong khung lát của hình (ví dụ, hình 714). Mỗi phần chia 718 chỉ được chứa trong một đơn vị NAL (ví dụ, đơn

vị VCL NAL). Khung lát (không được thể hiện trên hình vẽ) là vùng hình chữ nhật của các CTU trong cột khung lát cụ thể và hàng khung lát cụ thể trong hình (ví dụ, hình 714). CTU (không được thể hiện trên hình vẽ) là CTB của các mẫu độ sáng, hai CTB tương ứng của các mẫu sắc độ của hình mà có ba mảng mẫu, hoặc CTB của các mẫu của hình đơn sắc hoặc hình mà được mã hóa sử dụng ba mặt phẳng màu riêng biệt và các cấu trúc cú pháp được sử dụng để mã hóa các mẫu này. CTB (không được thể hiện trên hình vẽ) là khối $N \times N$ của các mẫu dành cho một số giá trị của N sao cho phép chia của một thành phần thành các CTB là dạng phân chia. Khối (không được thể hiện trên hình vẽ) là mảng $M \times N$ (M -cột nhân N -hàng) của các mẫu (ví dụ, các điểm ảnh), hoặc mảng $M \times N$ của các hệ số biến đổi.

Theo một phương án, mỗi phần chia 718 chứa phần đầu phần chia 720. Phần đầu phần chia 720 là một phần của phần chia được mã hóa 718 chứa các phần tử dữ liệu liên quan đến tất cả các khung lát hoặc các hàng CTU trong khung lát được thể hiện trong phần chia 718. Tức là, phần đầu phần chia 720 chứa thông tin về phần chia 718 chẳng hạn như, ví dụ, kiểu phần chia, mà của các hình tham chiếu sẽ được sử dụng, và v.v..

Các hình 714 và các phần chia 718 của chúng bao gồm dữ liệu được liên kết với các ảnh hoặc video được lập mã hoặc giải mã. Do đó, các hình 714 và các phần chia 718 của chúng có thể được gọi một cách đơn giản là tải hoặc dữ liệu được mang trong luồng bit 700.

Người có hiểu biết trung bình về lĩnh vực kỹ thuật này sẽ đánh giá cao rằng luồng bit 700 có thể chứa các thông số và thông tin khác theo các ứng dụng cụ thể.

Fig.8 là một phương án về phương pháp 800 để giải mã được thực hiện bởi bộ giải mã video (ví dụ, bộ giải mã video 400). Phương pháp 800 có thể được thực hiện sau khi luồng bit được thu trực tiếp hoặc gián tiếp từ bộ lập mã video (ví dụ, bộ lập mã video 300). Phương pháp 800 cải thiện quy trình giải mã bằng cách không cho phép các lớp không cần thiết trong các luồng bit video nhiều lớp. Tức là, mọi lớp được chứa trong tập lớp đầu ra (OLS) hoặc là lớp đầu ra hoặc được sử dụng như lớp tham chiếu trực tiếp hoặc gián tiếp của lớp đầu ra. Việc này tránh có thông tin ngoại lai trong quy trình mã hóa và làm tăng hiệu quả mã hóa. Do đó, bộ mã hóa/bộ giải mã (còn được gọi là, “bộ mã hóa/giải mã”) khi mã hóa video được cải thiện so với các bộ mã hóa/giải mã hiện thời. Một vấn đề thực tế, quy trình mã hóa video được cải thiện đem lại cho người dùng trải nghiệm người dùng tốt hơn khi các video được gửi, được thu, và/hoặc được xem.

Ở bước 802, bộ giải mã video thu luồng bit video bao gồm VPS (ví dụ, VPS 704) và nhiều lớp (ví dụ, lớp N 631, lớp N+1 632, v.v.). Theo một phương án, không có lớp nào không phải là lớp đầu ra của ít nhất một OLS (ví dụ, OLS 1, OLS 2, v.v.) mà cũng không phải là lớp tham chiếu trực tiếp của lớp bất kỳ khác. Tức là, mỗi lớp của ít nhất một OLS là lớp đầu ra (ví dụ, lớp N+2 633 trong OLS 3) hoặc lớp tham chiếu trực tiếp của lớp bất kỳ khác (ví dụ, lớp N+1 632 và lớp N 631 trong OLS 3). Lớp bất kỳ mà không phải là lớp đầu ra hoặc lớp mà được tham chiếu trực tiếp bởi lớp khác là không cần thiết và, do đó, bị loại trừ khỏi CVS (ví dụ, 690) và/hoặc luồng bit video.

Theo một phương án, bộ giải mã video mong muốn không có lớp nào sẽ không phải là lớp đầu ra của ít nhất một tập lớp đầu ra (OLS) mà cũng không phải là lớp tham chiếu trực tiếp của lớp bất kỳ khác như được mô tả ở trên dựa trên VVC hoặc một số chuẩn khác. Tuy nhiên, nếu bộ giải mã xác định rằng điều kiện này không đúng, thì bộ giải mã có thể phát hiện lỗi, báo hiệu lỗi, yêu cầu luồng bit được sửa đổi (hoặc một phần của nó) được gửi lại, hoặc thực hiện một số biện pháp hiệu chỉnh khác để đảm bảo rằng luồng bit phù hợp được thu.

Theo một phương án, VPS bao gồm lớp được sử dụng làm cờ tham chiếu (ví dụ, LayerUsedAsRefLayerFlag) và lớp được sử dụng làm cờ lớp đầu ra (ví dụ, LayerUsedAsOutputLayerFlag), và trong đó giá trị (ví dụ, [i]) của lớp này được sử dụng làm cờ tham chiếu và giá trị (ví dụ, [i]) của lớp này được sử dụng làm cờ lớp đầu ra đều không phải là không.

Theo một phương án, mỗi lớp trong nhiều lớp bao gồm tập các đơn vị lớp trừu tượng mạng (NAL) lớp mã hóa video (VCL) mà đều có giá trị cụ thể của bộ nhận dạng (ID) lớp và các đơn vị NAL không phải VCL được liên kết. Theo một phương án, ít nhất một OLS chứa một hoặc nhiều lớp đầu ra. Theo một phương án, đối với mỗi lớp trong số nhiều lớp với giá trị cụ thể của ID lớp được quy định trong VPS, một trong số các lớp trong ít nhất một OLS cũng sẽ có giá trị cụ thể của ID lớp.

Ở bước 804, bộ giải mã video giải mã hình (ví dụ, hình 615) từ một trong số nhiều lớp. Theo một phương án, hình được chứa trong lớp đầu ra của ít nhất một OLS. Theo một phương án, phương pháp 800 còn bao gồm bước chọn lựa lớp đầu ra từ ít nhất một OLS trước khi giải mã. Theo một phương án, phương pháp 800 còn bao gồm bước chọn lựa hình từ lớp đầu ra sau khi lớp đầu ra được chọn lựa.

Một khi hình được giải mã, thì hình này có thể được sử dụng để tạo ra hoặc lập ra ảnh hoặc trình tự video để hiển thị cho người dùng trên bộ hiển thị hoặc màn hình của thiết bị điện tử (ví dụ, điện thoại thông minh, máy tính bảng, máy tính xách tay, máy tính cá nhân, v.v.).

Fig.9 là một phương án về phương pháp 900 để lập mã luồng bit video được thực hiện bởi bộ lập mã video (ví dụ, bộ lập mã video 300). Phương pháp 900 có thể được thực hiện khi hình (ví dụ, từ video) sẽ được lập mã thành luồng bit video và sau đó được truyền về phía bộ giải mã video (ví dụ, bộ giải mã video 400). Phương pháp 900 cải thiện quy trình giải mã bằng cách không cho phép các lớp không cần thiết trong các luồng bit video nhiều lớp. Tức là, mọi lớp được chứa trong tập lớp đầu ra (OLS) hoặc là lớp đầu ra hoặc được sử dụng như lớp tham chiếu trực tiếp hoặc gián tiếp của lớp đầu ra. Việc này tránh có thông tin ngoại lai trong quy trình mã hóa và làm tăng hiệu quả mã hóa. Do đó, bộ mã hóa/bộ giải mã (còn được gọi là, “bộ mã hóa/giải mã”) khi mã hóa video được cải thiện so với các bộ mã hóa/giải mã hiện thời. Một vấn đề thực tế, quy trình mã hóa video được cải thiện đem lại cho người dùng trải nghiệm người dùng tốt hơn khi các video được gửi, được thu, và/hoặc được xem.

Ở bước 902, bộ lập mã video tạo ra nhiều lớp (ví dụ, lớp N 631, lớp N+1 632, v.v.) và VPS (ví dụ, VPS 704) quy định một hoặc nhiều OLS (ví dụ, OLS 1, OLS 2, v.v.). Theo một phương án, mỗi lớp từ nhiều lớp được chứa trong ít nhất một trong số các OLS được quy định bởi VPS. Theo một phương án, không có lớp nào không phải là lớp đầu ra của ít nhất một OLS (ví dụ, OLS 1, OLS 2, v.v.) mà cũng không phải là lớp tham chiếu trực tiếp của lớp bất kỳ khác. Tức là, mỗi lớp của ít nhất một OLS là lớp đầu ra (ví dụ, lớp N+2 633 trong OLS 3) hoặc lớp tham chiếu trực tiếp của lớp bất kỳ khác (ví dụ, lớp N+1 632 và lớp N 631 trong OLS 3). Lớp bất kỳ mà không phải là lớp đầu ra hoặc lớp mà được tham chiếu trực tiếp bởi lớp khác là không cần thiết và, do đó, bị loại trừ khỏi CVS (ví dụ, 690) và/hoặc luồng bit video. Theo một phương án, bộ lập mã video bị ràng buộc sao cho không có lớp nào không phải là lớp đầu ra của ít nhất một tập lớp đầu ra (OLS) mà cũng không phải là lớp tham chiếu trực tiếp của lớp bất kỳ khác. Tức là, bộ lập mã video được yêu cầu để lập mã sao cho không có lớp nào không phải là lớp đầu ra của ít nhất một tập lớp đầu ra (OLS) mà cũng không phải là lớp tham chiếu trực tiếp của lớp bất kỳ khác. Ràng buộc hoặc yêu cầu như vậy đảm bảo rằng luồng bit phù hợp với, ví dụ, VVC hoặc một số chuẩn khác.

Theo một phương án, mỗi OLS trong số một hoặc nhiều OLS chứa một hoặc nhiều lớp

đầu ra, và trong đó mỗi lớp đầu ra trong số các lớp đầu ra bao gồm một hoặc nhiều hình. Theo một phương án, mỗi lớp trong nhiều lớp bao gồm tập các đơn vị lớp trừu tượng mạng (NAL) lớp mã hóa video (VCL) mà đều có giá trị cụ thể của bộ nhận dạng (ID) lớp và các đơn vị NAL không phải VCL được liên kết

Theo một phương án, một trong số các OLS bao gồm hai lớp đầu ra, và trong đó một trong số hai lớp đầu ra tham chiếu tới lớp khác trong số hai lớp đầu ra. Theo một phương án, trong đó VPS bao gồm lớp được sử dụng làm cờ tham chiếu và lớp được sử dụng làm cờ lớp đầu ra, và trong đó giá trị của lớp được sử dụng làm cờ tham chiếu và giá trị của lớp được sử dụng làm cờ lớp đầu ra đều không phải là không.

Theo một phương án, bộ giải mã tham chiếu giả thiết (HRD) được bố trí trong bộ lập mã sẽ kiểm tra để đảm bảo rằng mọi lớp được chứa trong tập lớp đầu ra (OLS) hoặc là lớp đầu ra hoặc được sử dụng như lớp tham chiếu trực tiếp hoặc gián tiếp của lớp đầu ra. Khi HRD tìm thấy lớp mà không cần thiết như được mô tả ở đây, thì lúc đó HRD sẽ trả về lỗi kiểm tra sự phù hợp. Tức là, kiểm tra sự phù hợp HRD đảm bảo rằng không có các lớp không cần thiết. Do đó, bộ lập mã lập mã theo yêu cầu không có lớp không được sử dụng, nhưng HRD thực thi yêu cầu này.

Ở bước 904, bộ lập mã video lập mã nhiều lớp và VPS thành luồng bit video. Ở bước 906, bộ lập mã video lưu luồng bit video để truyền thông với bộ giải mã video. Luồng bit video có thể được lưu trong bộ nhớ cho đến khi luồng bit video được truyền về phía bộ giải mã video. Một khi được thu bởi bộ giải mã video, thì luồng bit video được lập mã có thể được giải mã (ví dụ, như được mô tả ở trên) để tạo ra hoặc lập ra ảnh hoặc trình tự video để hiển thị cho người dùng trên bộ hiển thị hoặc màn hình của thiết bị điện tử (ví dụ, điện thoại thông minh, máy tính bảng, máy tính xách tay, máy tính cá nhân, v.v.).

Fig.10 là hình vẽ sơ đồ của thiết bị mã hóa video 1000 (ví dụ, bộ lập mã video 300 hoặc bộ giải mã video 400) theo một phương án của sáng chế. Thiết bị mã hóa video 1000 thích hợp để thực hiện các phương án được bộc lộ như được mô tả ở đây. Thiết bị mã hóa video 1000 bao gồm các cổng lối vào 1010 và các đơn vị bộ thu (Rx) 1020 để thu dữ liệu; bộ xử lý, đơn vị logic, hoặc bộ xử lý trung tâm (CPU) 1030 để xử lý dữ liệu; các đơn vị bộ truyền (Tx) 1040 và các cổng lối ra 1050 để truyền dữ liệu; và bộ nhớ 1060 để lưu dữ liệu. Thiết bị mã hóa video 1000 cũng có thể bao gồm các bộ phận quang-điện (optical-to-electrical, OE) và

các bộ phận điện-quang (electrical-to-optical, EO) được ghép nối với các cổng lối vào 1010, các đơn vị bộ thu 1020, các đơn vị bộ truyền 1040, và các cổng lối ra 1050 để cho ra hoặc cho vào các tín hiệu quang hoặc điện.

Bộ xử lý 1030 được thực hiện bằng phần cứng và phần mềm. Bộ xử lý 1030 có thể được thực hiện như một hoặc nhiều chip CPU, các lõi (ví dụ bộ xử lý đa lõi), mảng cổng khả lập trình bằng trường (Field-Programmable Gate Array, FPGA), mạch tích hợp chuyên dụng (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 lối vào 1010, các đơn vị bộ thu 1020, các đơn vị bộ truyền 1040, các cổng lối ra 1050, và bộ nhớ 1060. Bộ xử lý 1030 bao gồm môđun mã hóa 1070. Môđun mã hóa 1070 thực hiện các phương án được bộc lộ được mô tả ở trên. Ví dụ, môđun mã hóa 1070 thực hiện, xử lý, chuẩn bị, hoặc cung cấp các chức năng mã hóa/giải mã khác nhau. Do đó, việc bao gồm môđun mã hóa 1070 cung cấp sự cải tiến đáng kể với chức năng của thiết bị mã hóa video 1000 và tác động đến sự biến đổi của thiết bị mã hóa video 1000 với trạng thái khác nhau. Ngoài ra, môđun mã hóa 1070 được thực hiện như các lệnh mà được lưu trong bộ nhớ 1060 và được thực thi bởi bộ xử lý 1030.

Thiết bị mã hóa video 1000 cũng có thể bao gồm các thiết bị nhập và/hoặc xuất (input/output, I/O) 1080 để truyền thông dữ liệu tới và từ người dùng. Thiết bị I/O 1080 có thể bao gồm các thiết bị xuất chẳng hạn như bộ hiển thị để hiển thị dữ liệu video, loa để xuất ra dữ liệu âm thanh, v.v.. Thiết bị I/O 1080 cũng có thể bao gồm thiết bị nhập, chẳng hạn như bàn phím, chuột, bi xoay, v.v. và/hoặc các giao diện tương ứng để tương tác với các thiết bị xuất.

Bộ nhớ 1060 bao gồm một hoặc nhiều đĩa, ổ băng, và ổ đĩa ở trạng thái rắn và có thể được sử dụng như thiết bị lưu trữ dữ liệu, để lưu các chương trình khi các chương trình như vậy được chọn lựa để thực thi, và để lưu lệnh và dữ liệu mà được đọc trong khi thực thi chương trình. Bộ nhớ 1060 có thể là bộ nhớ khả biến và/hoặc bất khả biến và có thể là bộ nhớ chỉ đọc (ROM), bộ nhớ truy nhập ngẫu nhiên (RAM), bộ nhớ địa chỉ-nội dung tam phân (ternary content-addressable memory, TCAM), và/hoặc bộ nhớ truy nhập ngẫu nhiên tĩnh (SRAM).

Fig.11 là hình vẽ sơ đồ của một phương án về phương tiện để mã hóa 1100. Theo một phương án, phương tiện để mã hóa 1100 được thực hiện trong thiết bị mã hóa video 1102 (ví dụ, bộ lập mã video 300 hoặc bộ giải mã video 400). Thiết bị mã hóa video 1102 bao gồm

phương tiện thu 1101. Phương tiện thu 1101 được tạo cấu hình để thu hình để lập mã hoặc để thu luồng bit để giải mã. Thiết bị mã hóa video 1102 bao gồm phương tiện truyền dẫn 1107 được ghép nối với phương tiện thu 1101. Phương tiện truyền dẫn 1107 được tạo cấu hình để truyền luồng bit tới bộ giải mã hoặc để truyền ảnh được giải mã tới phương tiện hiển thị (ví dụ, một trong số các thiết bị I/O 1080).

Thiết bị mã hóa video 1102 bao gồm phương tiện lưu trữ 1103. Phương tiện lưu trữ 1103 được ghép nối với ít nhất một trong số phương tiện thu 1101 hoặc phương tiện truyền dẫn 1107. Phương tiện lưu trữ 1103 được tạo cấu hình để lưu các lệnh. Thiết bị mã hóa video 1102 cũng 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 trong phương tiện lưu trữ 1103 để thực hiện các phương pháp được bộc lộ ở đây.

Phải hiểu rằng các bước của các phương pháp ví dụ được nêu ở đây không nhất thiết phải được thực hiện theo thứ bậc được mô tả, và thứ bậc của các bước của các phương pháp như vậy phải được hiểu chỉ đơn thuần là ví dụ. Tương tự như vậy, các bước bổ sung có thể được chứa trong các phương pháp như vậy, và các bước nhất định có thể được bỏ qua hoặc được kết hợp, trong các phương pháp nhất quán với các phương pháp khác nhau của sáng chế.

Mặc dù một số phương án đã được thể hiện trong phần mô tả sáng chế, nhưng nên hiểu rằng các hệ thống và phương pháp được bộc lộ có thể được thể hiện dưới nhiều hình thức cụ thể khác mà không nằm ngoài phạm vi của sáng chế. Các ví dụ được thể hiện trong phần mô tả sáng chế được coi là các ví dụ minh họa và không nhằm hạn chế, và sáng chế không bị hạn chế bởi các phần mô tả chi tiết được nêu ở đây. Ví dụ, các phần tử hoặc thành phần khác nhau có thể được kết hợp hoặc tích hợp trong hệ thống khác hoặc các dấu hiệu nhất định có thể bị bỏ qua hoặc không được thực hiện.

Ngoài ra, các kỹ thuật, hệ thống, hệ thống con và phương pháp được mô tả và thể hiện trong các phương án khác nhau như là các kỹ thuật, hệ thống, hệ thống con và phương pháp rời rạc hoặc riêng biệt có thể được kết hợp hoặc tích hợp với các hệ thống, môđun, kỹ thuật hoặc phương pháp khác mà không nằm ngoài phạm vi của sáng chế. Các mục khác được thể hiện hoặc thảo luận như việc được ghép nối hoặc ghép nối trực tiếp hoặc truyền thông với nhau có thể là việc được ghép nối gián tiếp hoặc truyền thông qua một số giao diện, thiết bị, hoặc thành phần trung gian cho dù bằng điện, cơ khí hoặc cách khác. Người có hiểu biết trung

bình về lĩnh vực kỹ thuật này có thể thực hiện các phương án thay đổi, phương án thay thế, và phương án bổ sung và các phương án này có thể được thực hiện mà không nằm ngoài phạm vi được bộc lộ ở đây.

YÊU CẦU BẢO HỘ

1. Phương pháp để giải mã được thực hiện bởi bộ giải mã video, bao gồm các bước:

thu, bởi bộ giải mã video, luồng bit video bao gồm tập thông số video (video parameter set, VPS) và nhiều lớp, trong đó không có lớp nào không phải là lớp đầu ra của ít nhất một tập lớp đầu ra (output layer set, OLS) mà cũng không phải là lớp tham chiếu trực tiếp của lớp bất kỳ khác, trong đó VPS bao gồm lớp được sử dụng làm cờ tham chiếu (LayerUsedAsRefLayerFlag) và lớp được sử dụng làm cờ lớp đầu ra (LayerUsedAsOutputLayerFlag), và trong đó giá trị của LayerUsedAsRefLayerFlag và giá trị của LayerUsedAsOutputLayerFlag đều không phải là không; và

giải mã, bởi bộ giải mã video, hình từ một trong số nhiều lớp.

2. Phương pháp theo điểm 1, trong đó hình được chứa trong lớp đầu ra của ít nhất một OLS.

3. Phương pháp theo điểm 1 hoặc điểm 2, trong đó phương pháp này còn bao gồm bước chọn lựa lớp đầu ra từ ít nhất một OLS trước khi giải mã.

4. Phương pháp theo điểm 1 hoặc điểm 2, trong đó phương pháp này còn bao gồm bước chọn lựa hình từ lớp đầu ra sau khi lớp đầu ra được chọn lựa.

5. Phương pháp theo điểm 1 hoặc điểm 2, trong đó mỗi lớp trong nhiều lớp bao gồm tập các đơn vị lớp trừu tượng mạng (network abstraction layer, NAL) lớp mã hóa video (video coding layer, VCL) mà đều có giá trị cụ thể của bộ nhận dạng (identifier, ID) lớp và các đơn vị NAL không phải VCL được liên kết.

6. Phương pháp theo điểm 1 hoặc điểm 2, trong đó mỗi lớp của ít nhất một OLS bao gồm lớp đầu ra hoặc lớp tham chiếu trực tiếp của lớp bất kỳ khác.

7. Phương pháp theo điểm 6, trong đó ít nhất một OLS chứa một hoặc nhiều lớp đầu ra.

8. Phương pháp theo điểm 1 hoặc điểm 2, trong đó VPS bao gồm lớp được sử dụng làm cờ tham chiếu và lớp được sử dụng làm cờ lớp đầu ra, và trong đó giá trị của lớp được sử dụng làm cờ tham chiếu và giá trị của lớp được sử dụng làm cờ lớp đầu ra đều không phải là không.

9. Phương pháp theo điểm 1 hoặc điểm 2, trong đó phương pháp này còn bao gồm bước hiển thị hình được giải mã trên bộ hiển thị của thiết bị điện tử.

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

thu, bởi bộ giải mã video, luồng bit video thứ hai bao gồm tập thông số video (VPS) thứ hai và nhiều lớp thứ hai, trong đó ít nhất một lớp không phải là lớp đầu ra của ít nhất một tập

lớp đầu ra (OLS) mà cũng không phải là lớp tham chiếu trực tiếp của lớp bất kỳ khác; và

nhằm đáp lại việc thu, thì thực hiện một số biện pháp hiệu chỉnh khác để đảm bảo rằng luồng bit phù hợp tương ứng với luồng bit video thứ hai được thu trước khi giải mã hình từ một trong số nhiều lớp thứ hai.

11. Phương pháp để lập mã được thực hiện bởi bộ lập mã video, phương pháp này bao gồm các bước:

tạo ra, bởi bộ lập mã video, nhiều lớp và tập thông số video (VPS) quy định ít nhất một tập lớp đầu ra (OLS), trong đó bộ lập mã video bị ràng buộc sao cho không có lớp nào không phải là lớp đầu ra của ít nhất một OLS mà cũng không phải là lớp tham chiếu trực tiếp của lớp bất kỳ khác, trong đó VPS bao gồm lớp được sử dụng làm cờ tham chiếu (LayerUsedAsRefLayerFlag) và lớp được sử dụng làm cờ lớp đầu ra (LayerUsedAsOutputLayerFlag), và trong đó giá trị của LayerUsedAsRefLayerFlag và giá trị của LayerUsedAsOutputLayerFlag đều không phải là không;

lập mã, bởi bộ lập mã video, nhiều lớp và VPS thành luồng bit video; và

lưu, bởi bộ lập mã video, luồng bit video để truyền thông với bộ giải mã video.

12. Phương pháp theo điểm 11, trong đó mỗi OLS trong số một hoặc nhiều OLS chứa một hoặc nhiều lớp đầu ra, và trong đó mỗi lớp đầu ra trong số các lớp đầu ra bao gồm một hoặc nhiều hình.

13. Phương pháp theo điểm 11 hoặc điểm 12, trong đó mỗi lớp trong nhiều lớp bao gồm tập các đơn vị lớp trừu tượng mạng (NAL) lớp mã hóa video (VCL) mà đều có giá trị cụ thể của bộ nhận dạng (ID) lớp và các đơn vị NAL không phải VCL được liên kết.

14. Phương pháp theo điểm 11 hoặc điểm 12, trong đó một trong số các OLS bao gồm hai lớp đầu ra, và trong đó một trong số hai lớp đầu ra tham chiếu tới lớp khác trong số hai lớp đầu ra.

15. Phương pháp theo điểm 11 hoặc điểm 12, trong đó VPS bao gồm lớp được sử dụng làm cờ tham chiếu và lớp được sử dụng làm cờ lớp đầu ra, và trong đó giá trị của lớp được sử dụng làm cờ tham chiếu và giá trị của lớp được sử dụng làm cờ lớp đầu ra đều không phải là không.

16. Thiết bị giải mã bao gồm:

bộ thu được tạo cấu hình để thu luồng bit video bao gồm tập thông số video (VPS) và nhiều lớp, trong đó không có lớp nào không phải là lớp đầu ra của ít nhất một OLS mà cũng

không phải là lớp tham chiếu trực tiếp của lớp bất kỳ khác;

bộ nhớ được ghép nối với bộ thu, bộ nhớ lưu các lệnh, trong đó VPS bao gồm lớp được sử dụng làm cờ tham chiếu (LayerUsedAsRefLayerFlag) và lớp được sử dụng làm cờ lớp đầu ra (LayerUsedAsOutputLayerFlag), và trong đó giá trị của LayerUsedAsRefLayerFlag và giá trị của LayerUsedAsOutputLayerFlag đều không phải là không; và

bộ xử lý được ghép nối với bộ nhớ, bộ xử lý được tạo cấu hình để thực thi các lệnh làm cho thiết bị giải mã giải mã hình từ một trong số nhiều lớp để nhận hình được giải mã.

17. Thiết bị giải mã theo điểm 16, trong đó hình được chứa trong lớp đầu ra của ít nhất một OLS.

18. Thiết bị giải mã theo điểm 16 hoặc điểm 17, trong đó bộ xử lý còn được tạo cấu hình để chọn lựa lớp đầu ra từ ít nhất một OLS trước khi hình được giải mã.

19. Thiết bị giải mã theo điểm 16 hoặc điểm 17, trong đó bộ xử lý còn được tạo cấu hình để chọn lựa hình từ lớp đầu ra sau khi lớp đầu ra đã được chọn lựa.

20. Thiết bị giải mã theo điểm 16 hoặc điểm 17, trong đó mỗi lớp trong nhiều lớp bao gồm tập các đơn vị lớp trừu tượng mạng (NAL) lớp mã hóa video (VCL) mà đều có giá trị cụ thể của bộ nhận dạng (ID) lớp và các đơn vị NAL không phải VCL được liên kết.

21. Thiết bị giải mã theo điểm 16 hoặc điểm 17, trong đó mỗi lớp của ít nhất một OLS bao gồm lớp đầu ra hoặc lớp tham chiếu trực tiếp của lớp bất kỳ khác.

22. Thiết bị giải mã theo điểm 16 hoặc điểm 17, trong đó ít nhất một OLS chứa một hoặc nhiều lớp đầu ra.

23. Thiết bị giải mã theo điểm 16 hoặc điểm 17, trong đó thiết bị này còn bao gồm bộ hiển thị được tạo cấu hình để hiển thị hình được giải mã.

24. Thiết bị giải mã theo điểm 16 hoặc điểm 17, trong đó bộ xử lý được tạo cấu hình để thực thi các lệnh còn làm cho thiết bị giải mã:

thu luồng bit video thứ hai bao gồm tập thông số video (VPS) thứ hai và nhiều lớp thứ hai, trong đó ít nhất một lớp không phải là lớp đầu ra của ít nhất một tập lớp đầu ra (OLS) mà cũng không phải là lớp tham chiếu trực tiếp của lớp bất kỳ khác; và

nhằm đáp lại việc thu luồng bit video thứ hai, thì thực hiện một số biện pháp hiệu chỉnh khác để đảm bảo rằng luồng bit phù hợp tương ứng với luồng bit video thứ hai được thu trước khi giải mã hình từ một trong số nhiều lớp thứ hai.

25. Thiết bị lập mã 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 làm cho thiết bị lập mã:

tạo ra nhiều lớp và tập thông số video (VPS) quy định một hoặc nhiều tập lớp đầu ra (OLS), trong đó thiết bị lập mã bị ràng buộc sao cho không có lớp nào không phải là lớp đầu ra của ít nhất một OLS mà cũng không phải là lớp tham chiếu trực tiếp của lớp bất kỳ khác, trong đó VPS bao gồm lớp được sử dụng làm cờ tham chiếu (LayerUsedAsRefLayerFlag) và lớp được sử dụng làm cờ lớp đầu ra (LayerUsedAsOutputLayerFlag), và trong đó giá trị của LayerUsedAsRefLayerFlag và giá trị của LayerUsedAsOutputLayerFlag đều không phải là không;

lập mã nhiều lớp và VPS thành luồng bit video; và

bộ truyền được ghép nối với bộ xử lý, bộ truyền được tạo cấu hình để truyền luồng bit video về phía bộ giải mã video.

26. Thiết bị lập mã theo điểm 25, trong đó mỗi OLS trong số một hoặc nhiều OLS chứa một hoặc nhiều lớp đầu ra, và trong đó mỗi lớp đầu ra trong số các lớp đầu ra bao gồm một hoặc nhiều hình.

27. Thiết bị lập mã theo điểm 25 hoặc điểm 26, trong đó mỗi lớp trong nhiều lớp bao gồm tập các đơn vị lớp trừu tượng mạng (NAL) lớp mã hóa video (VCL) mà đều có giá trị cụ thể của bộ nhận dạng (ID) lớp và các đơn vị NAL không phải VCL được liên kết.

28. Thiết bị lập mã theo điểm 25 hoặc điểm 26, trong đó mỗi lớp của ít nhất một OLS bao gồm lớp đầu ra hoặc lớp tham chiếu trực tiếp của lớp bất kỳ khác.

29. Thiết bị lập mã theo điểm 25 hoặc điểm 26, trong đó VPS bao gồm lớp được sử dụng làm cờ tham chiếu và lớp được sử dụng làm cờ lớp đầu ra, và trong đó giá trị của lớp được sử dụng làm cờ tham chiếu và giá trị của lớp được sử dụng làm cờ lớp đầu ra đều không phải là không.

30. Thiết bị mã hóa bao gồm:

bộ thu được tạo cấu hình để thu hình để lập mã hoặc để thu luồng bit để giải mã;

bộ truyền được ghép nối với bộ thu, bộ truyền được tạo cấu hình để truyền luồng bit tới bộ giải mã hoặc để truyền ảnh được giải mã tới bộ hiển thị;

bộ nhớ được ghép nối với ít nhất một trong số bộ thu hoặc bộ truyền, bộ nhớ được tạo

cấu hình để lưu 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 trong bộ nhớ để thực hiện phương pháp theo điểm bất kỳ trong số các điểm từ điểm 1 đến 10 và theo điểm bất kỳ trong số các điểm từ điểm 11 đến 15.

31. Thiết bị mã hóa theo điểm 30, trong đó thiết bị này còn bao gồm bộ hiển thị được tạo cấu hình để hiển thị hình được giải mã.

32. Hệ thống mã hóa bao gồm:

bộ lập mã; và

bộ giải mã truyền thông với bộ lập mã hoặc bộ giải mã, trong đó bộ giải mã bao gồm thiết bị giải mã, thiết bị giải mã theo điểm bất kỳ trong số các điểm từ điểm 16 đến 24, bộ lập mã bao gồm thiết bị lập mã điểm bất kỳ trong số các điểm từ điểm 25 đến 29, hoặc thiết bị mã hóa theo điểm 30 hoặc điểm 31.

33. Phương tiện để mã hóa bao gồm:

phương tiện thu được tạo cấu hình để thu hình để lập mã hoặc để thu luồng bit để giải mã;

phương tiện truyền dẫn được ghép nối với phương tiện thu, phương tiện truyền dẫn được tạo cấu hình để truyền luồng bit tới phương tiện giải mã hoặc để truyền ảnh được giải mã tới phương tiện hiển thị;

phương tiện lưu trữ được ghép nối với ít nhất một trong số phương tiện thu hoặc phương tiện truyền dẫn, phương tiện lưu trữ được tạo cấu hình để lưu các lệnh; và

phương tiện xử lý được ghép nối với phương tiện lưu trữ, các phương tiện xử lý được tạo cấu hình để thực thi các lệnh được lưu trong phương tiện lưu trữ để thực hiện phương pháp theo điểm bất kỳ trong số các điểm từ điểm 1 đến 10 và theo điểm bất kỳ trong số các điểm từ điểm 11 đến 15.

1/11

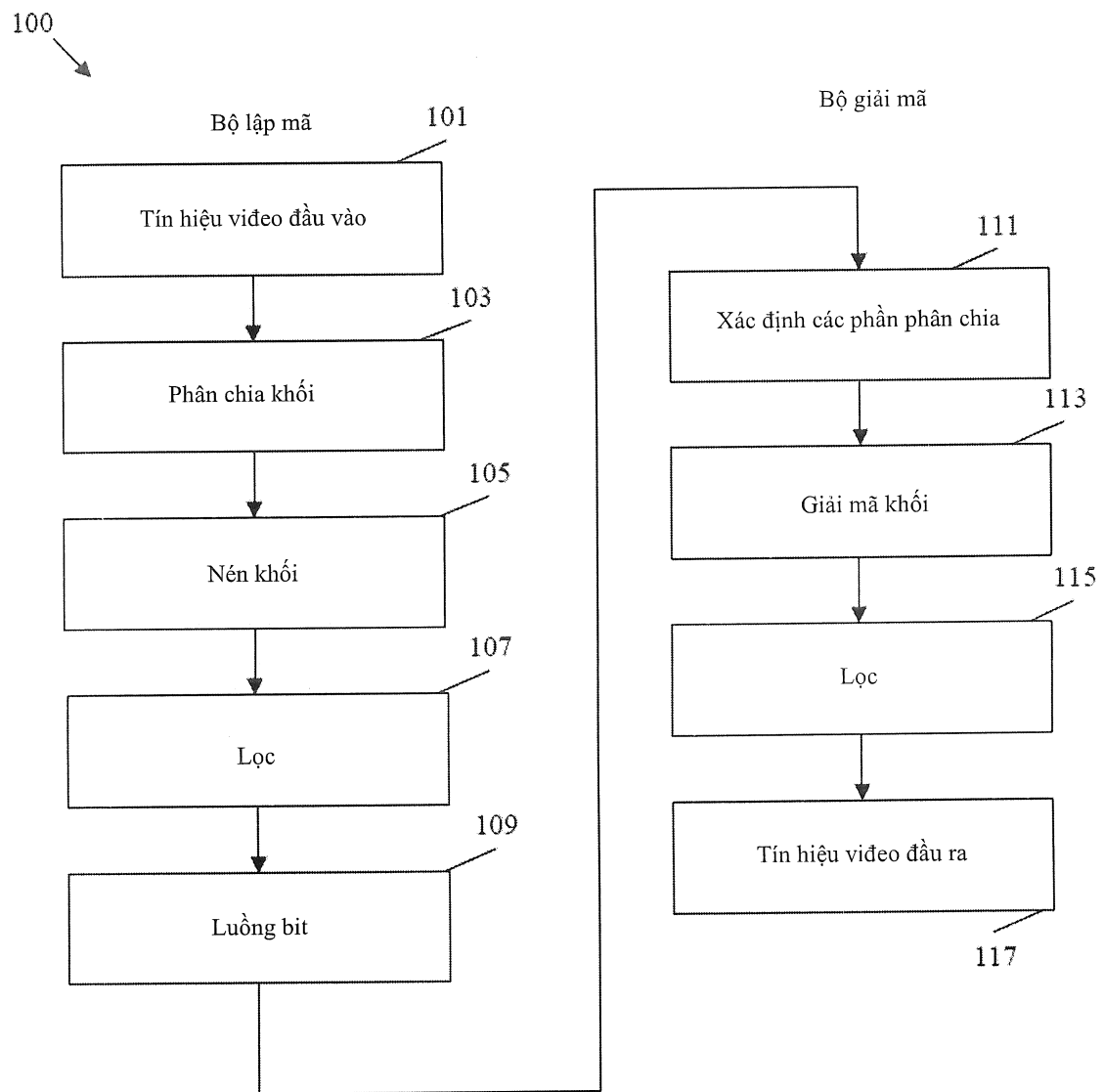


Fig.1

2/11

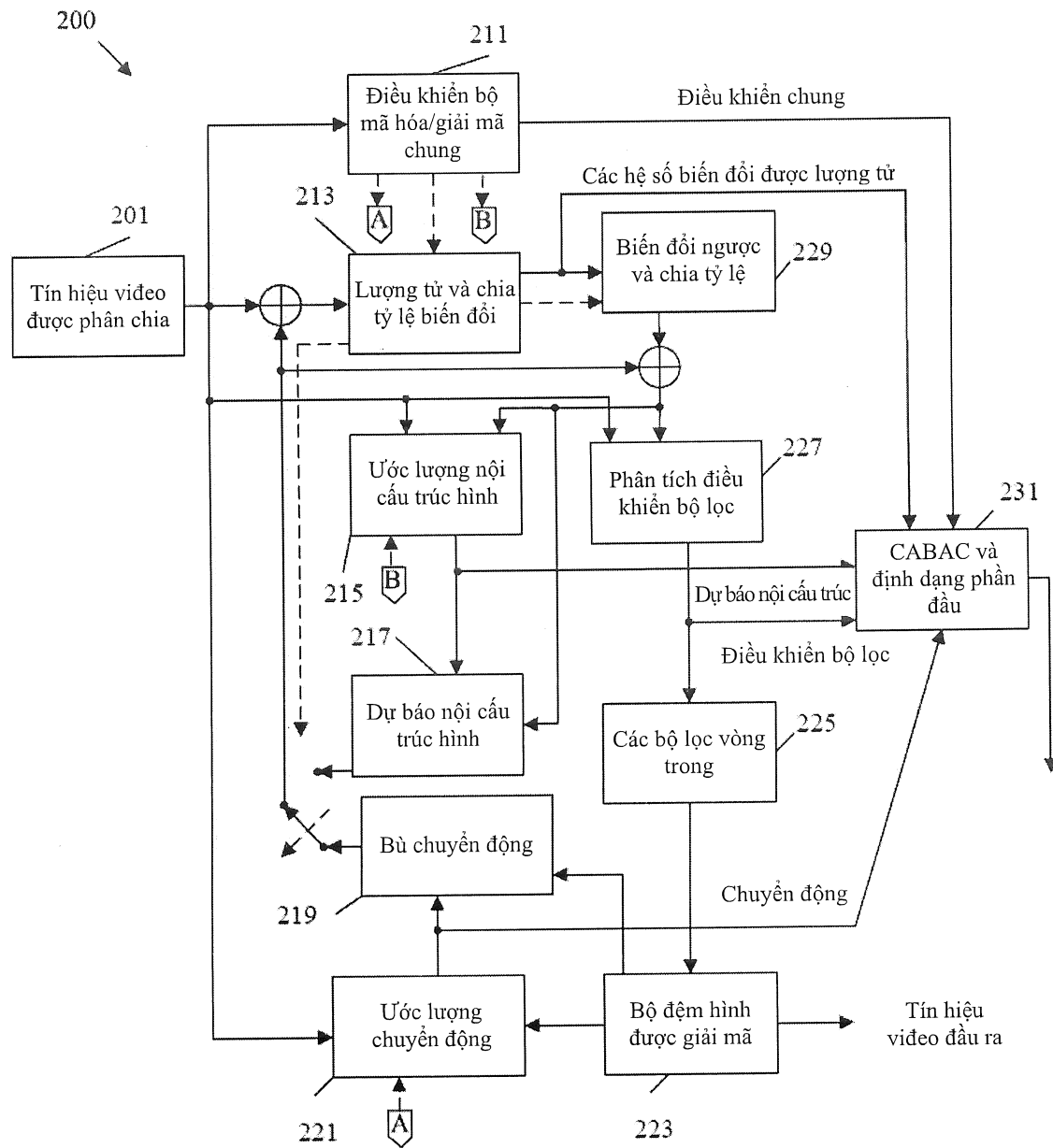


Fig.2

3/11

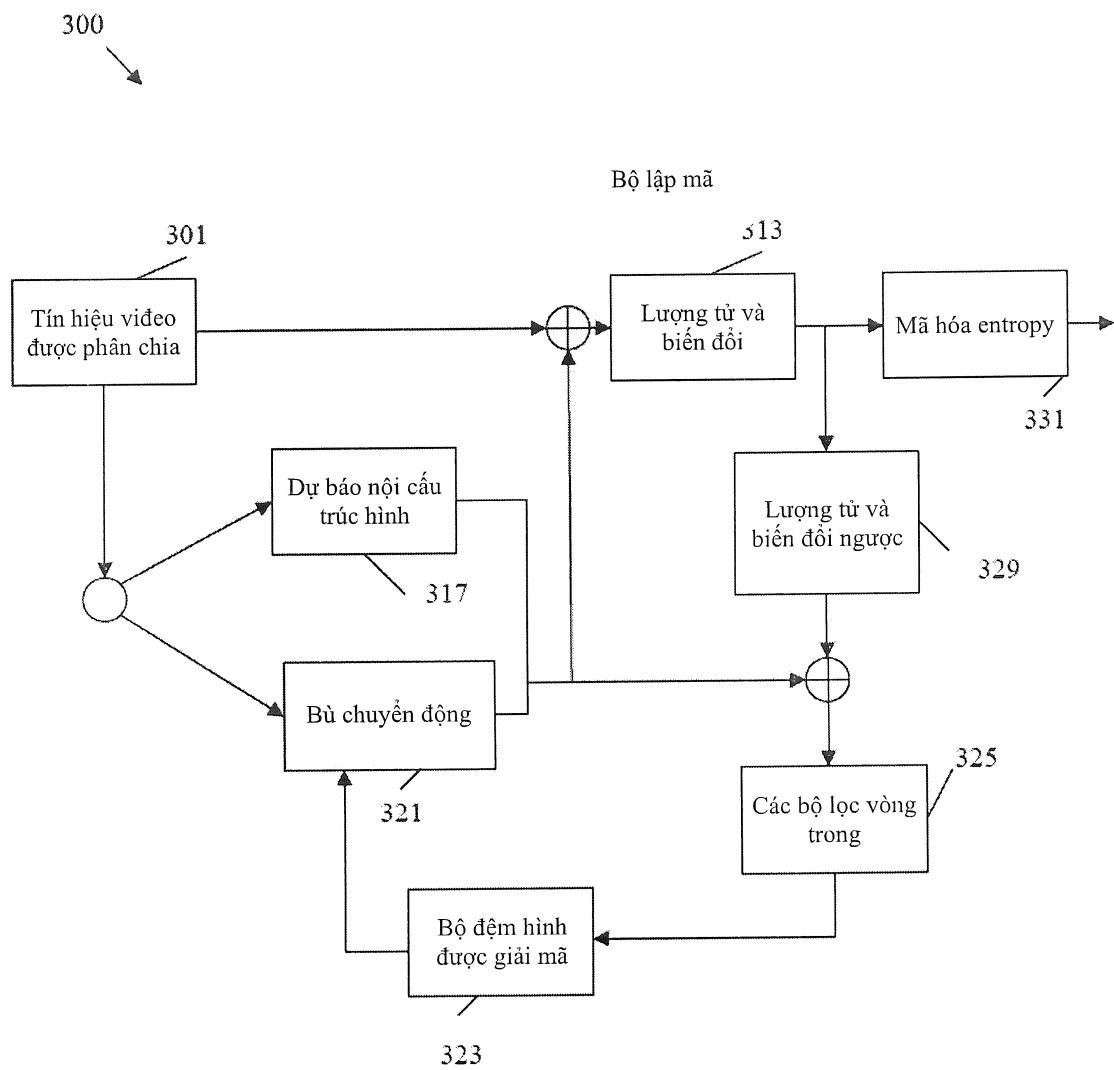


Fig.3

4/11

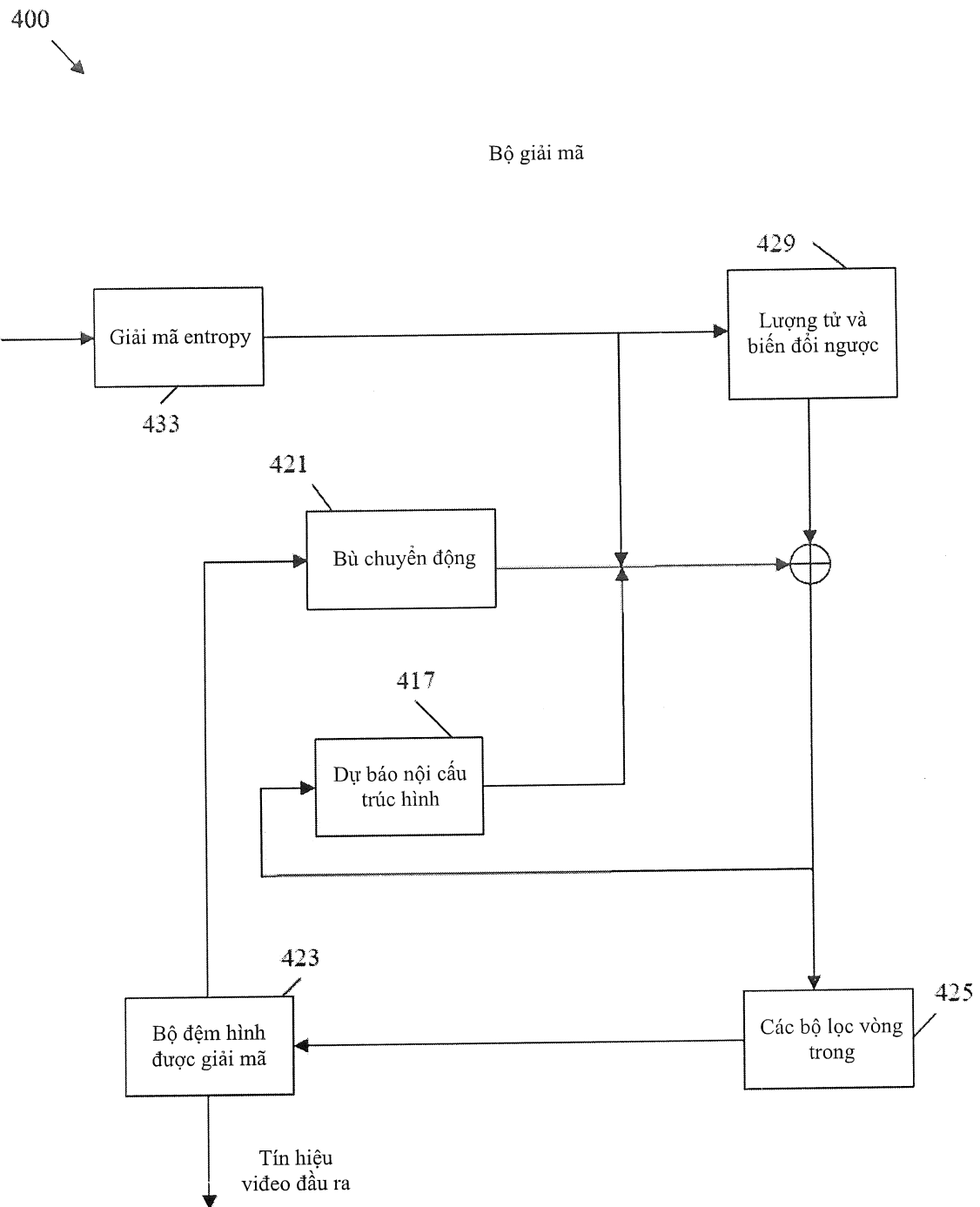


Fig.4

5/11

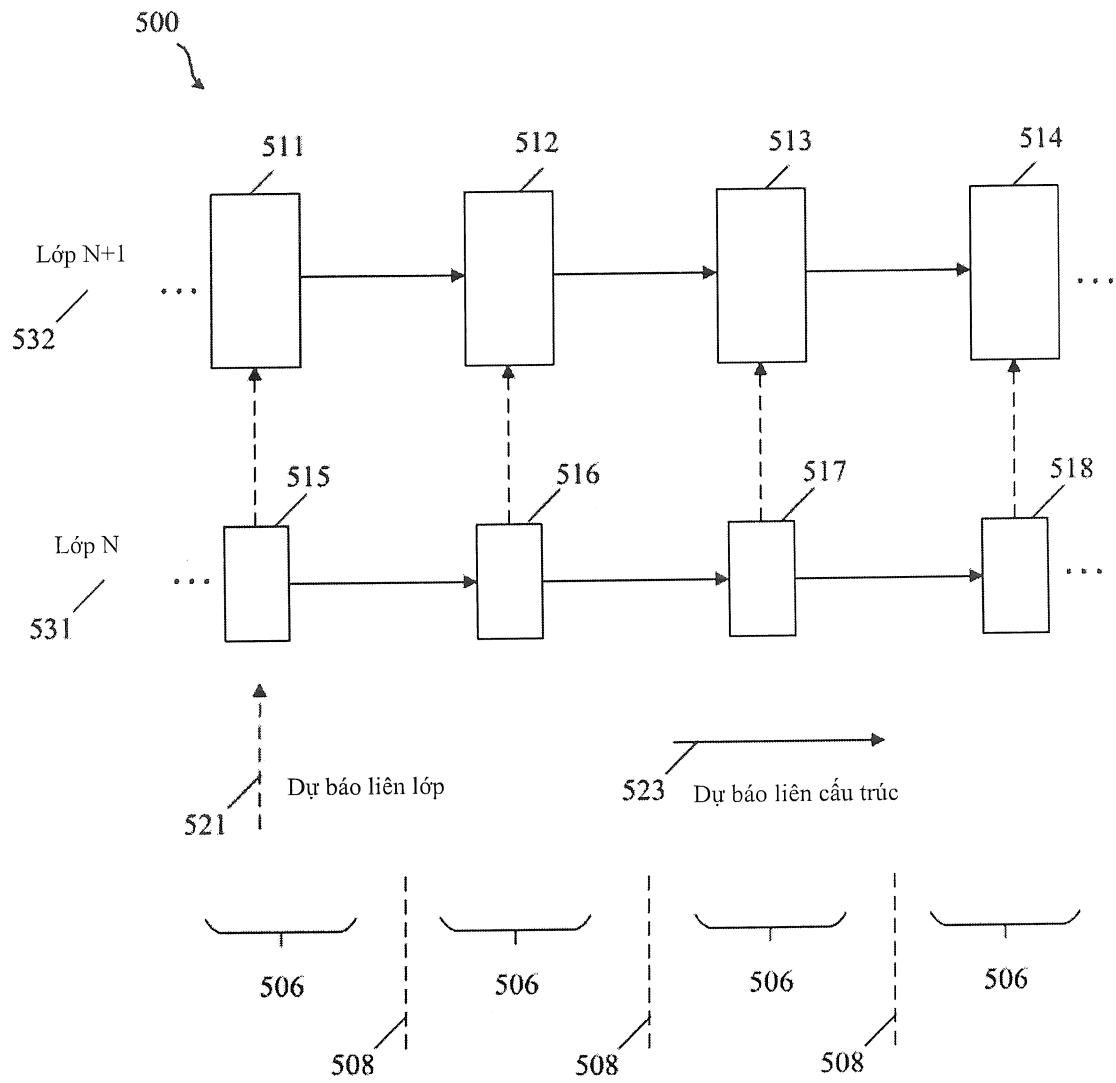
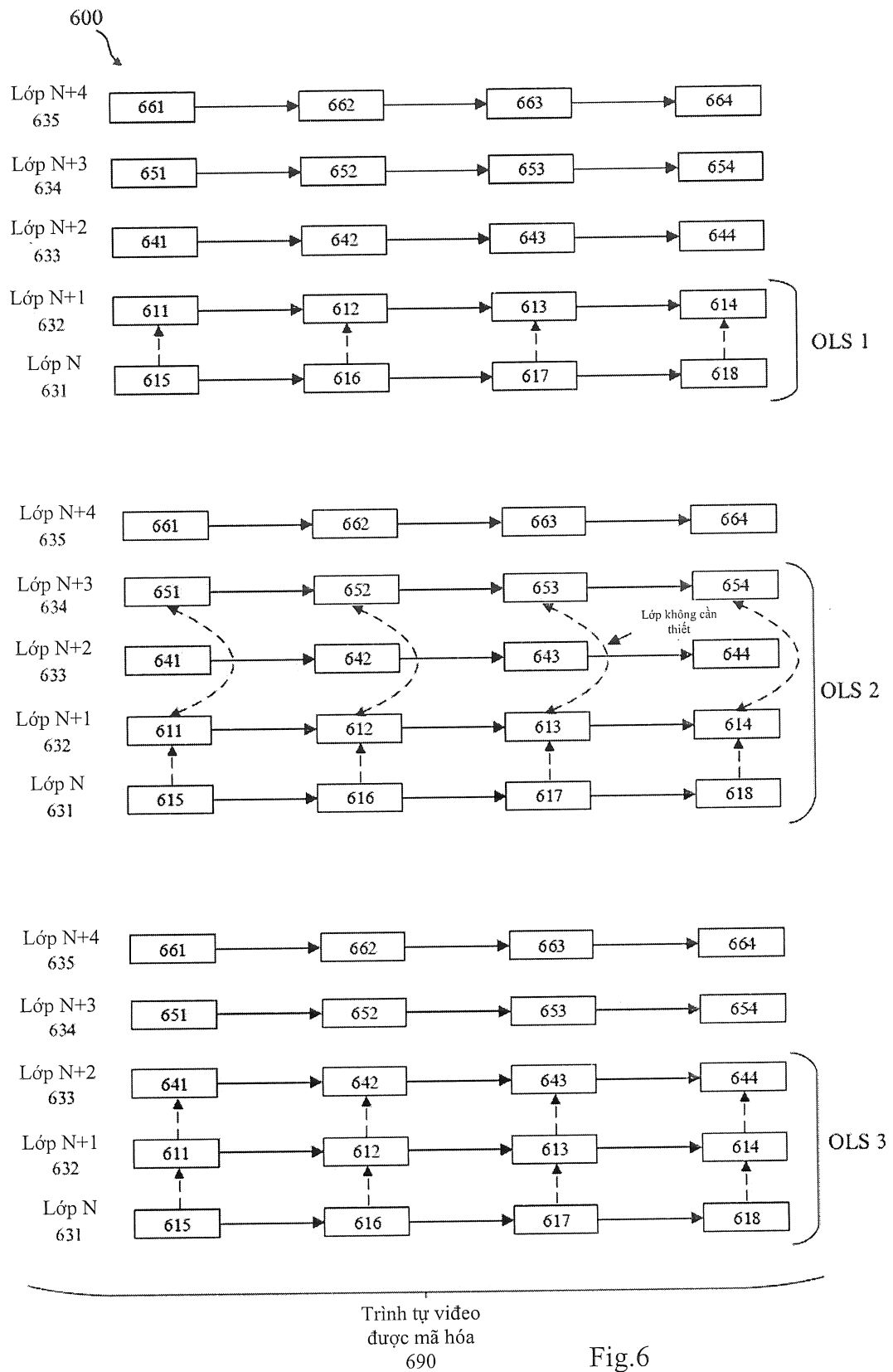


Fig.5

6/11



7/11

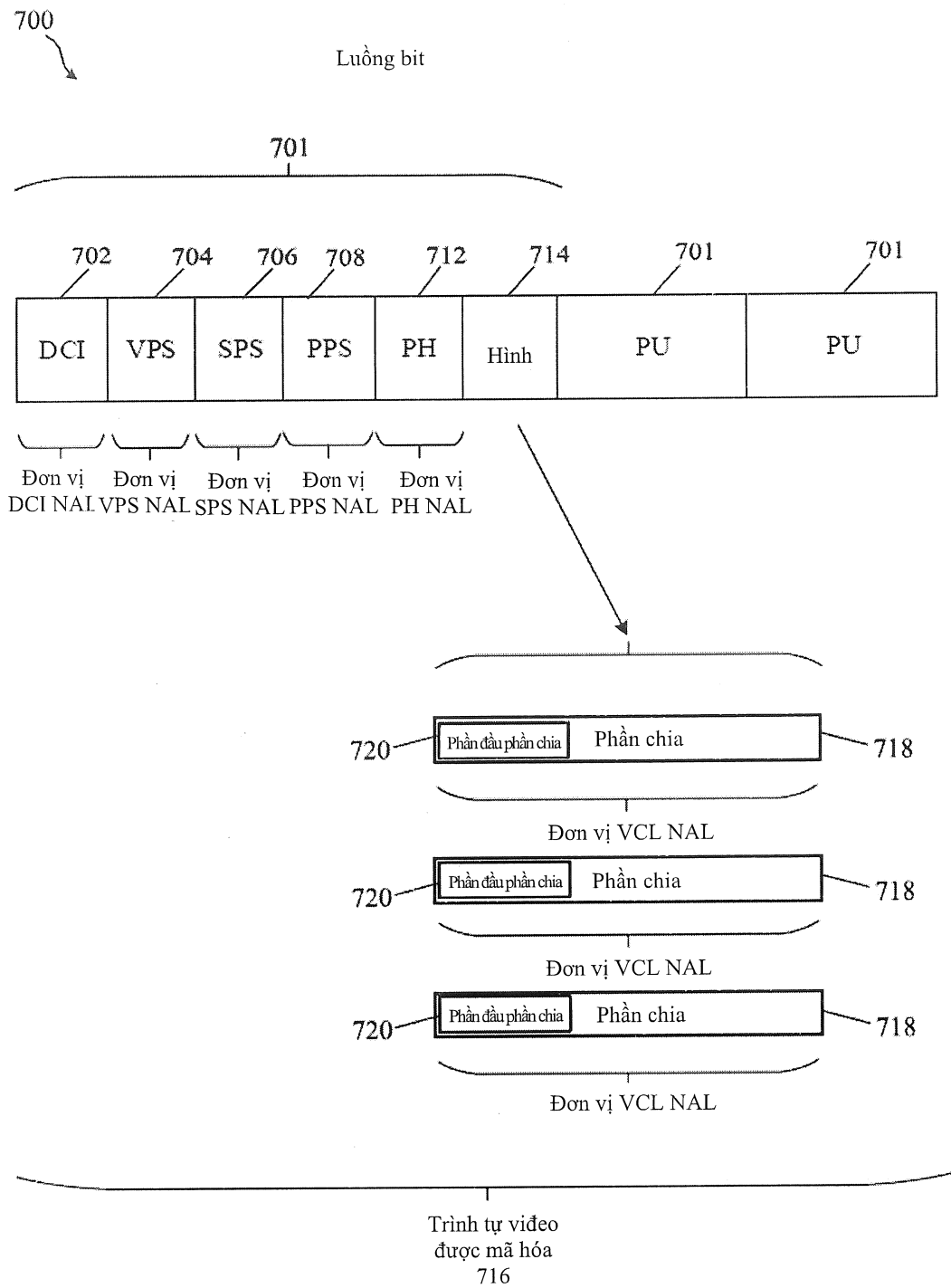


Fig.7

8/11

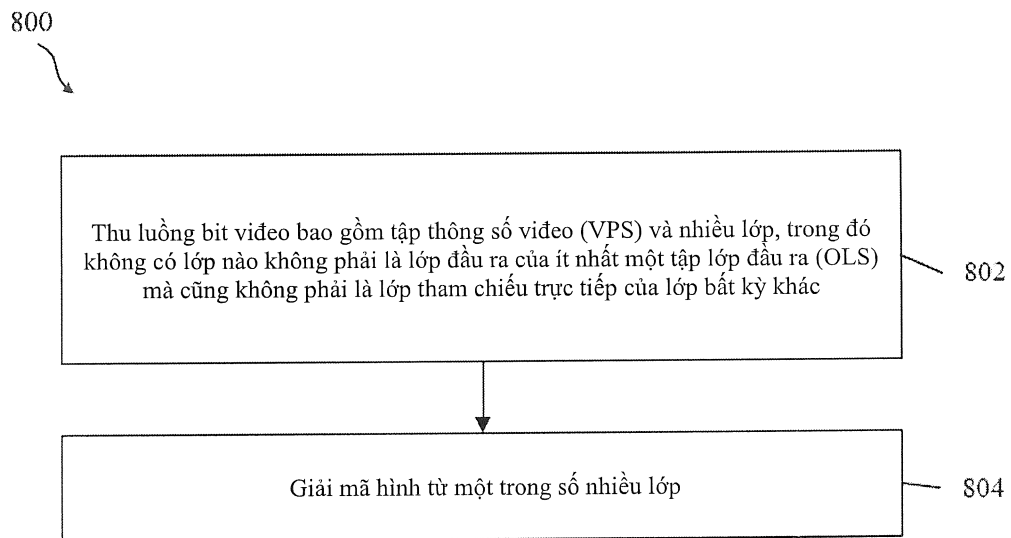


Fig.8

9/11

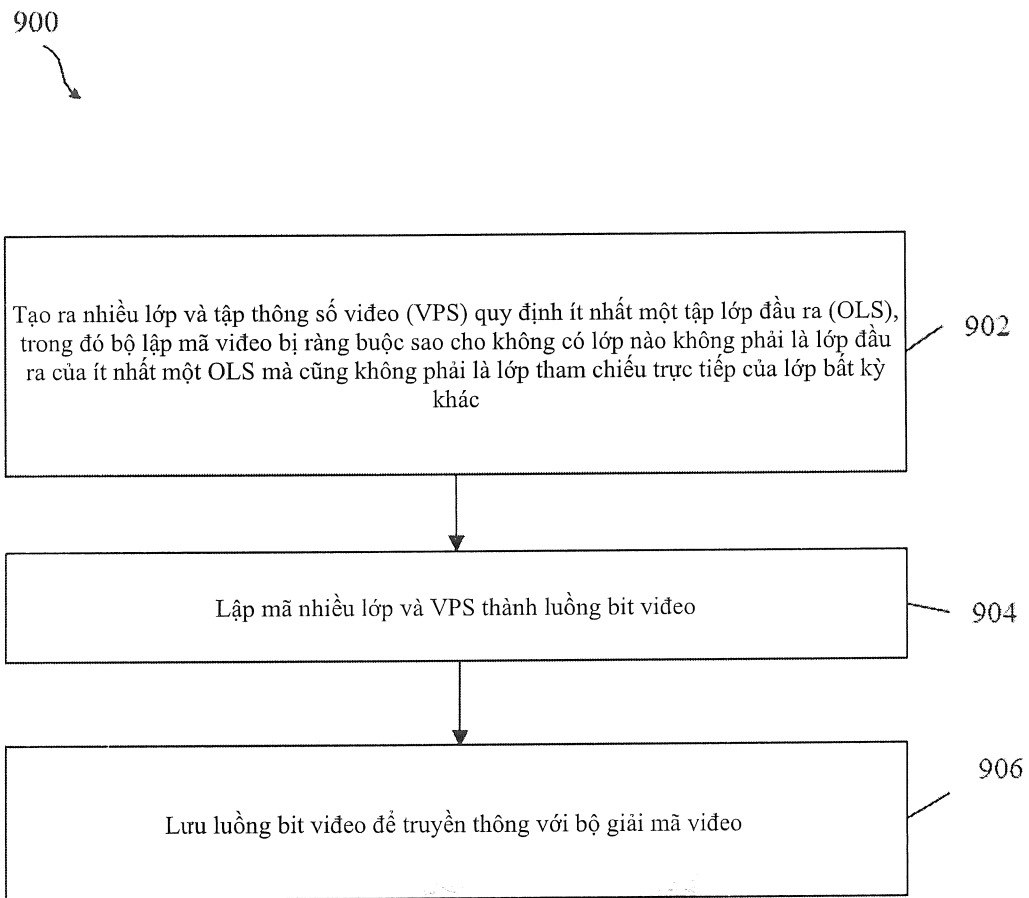


Fig.9

10/11

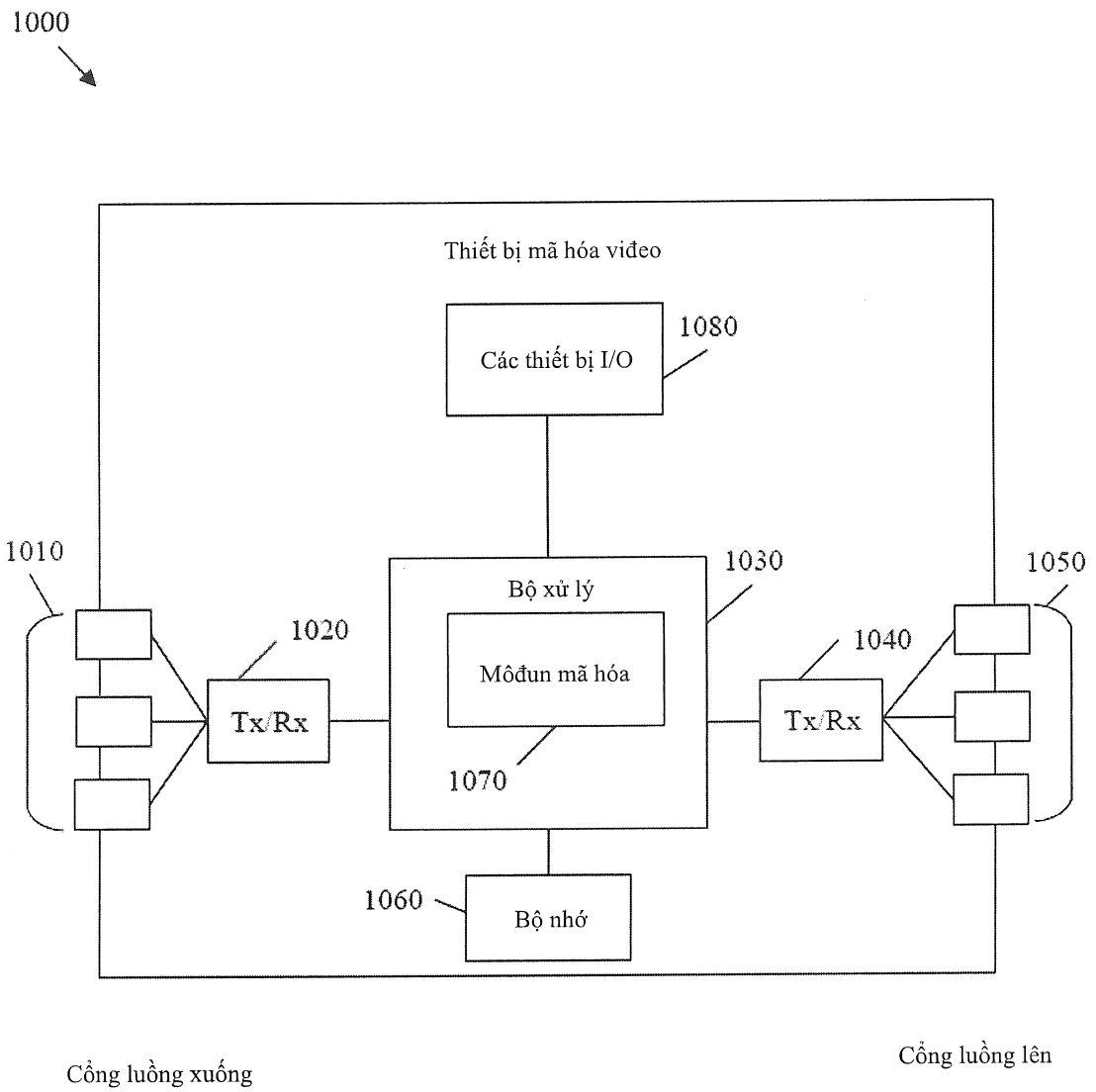


Fig.10

11/11

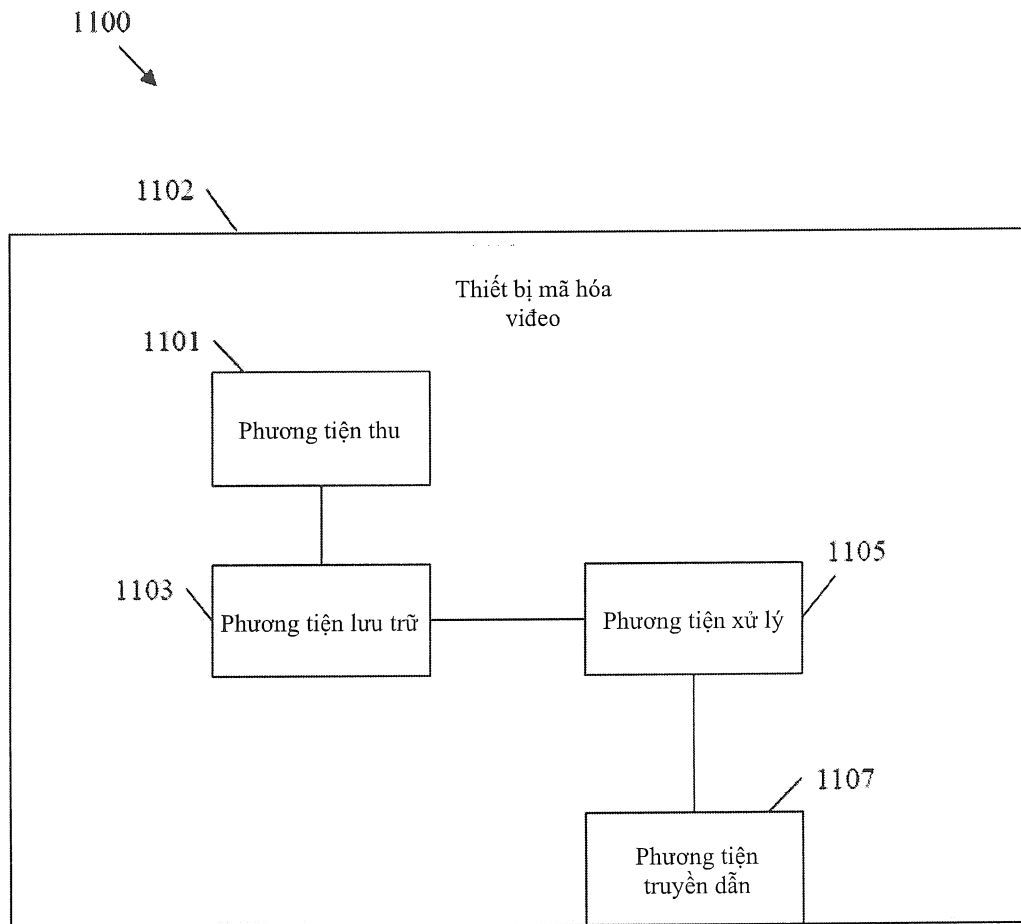


Fig.11