



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

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

(11)



1-0049110

(51)^{2021.01} H04N 19/46; H04N 19/172; H04N
19/96; H04N 19/70; H04N 19/105;
H04N 19/187

(13) B

(21) 1-2022-06112

(22) 22/03/2021

(86) PCT/CN2021/082029 22/03/2021

(87) WO 2021/190438 30/09/2021

(30) PCT/CN2020/080533 21/03/2020 CN

(45) 25/07/2025 448

(43) 26/12/2022 417A

(73) 1. BEIJING BYTEDANCE NETWORK TECHNOLOGY CO., LTD. (CN)
Room B-0035, 2/F, No. 3 Building No. 30, Shixing Road, Shijingshan District
Beijing 100041, China

2. BYTEDANCE INC. (US)

12655 West Jefferson Boulevard Sixth Floor, Suite No. 137 Los Angeles, California
90066, USA

(72) WANG, Ye-Kui (US); ZHANG, Li (CN); ZHANG, Kai (CN); DENG, Zhipin (CN).

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

(54) PHƯƠNG PHÁP VÀ THIẾT BỊ XỬ LÝ VIDEO, VÀ VẬT LƯU TRỮ MÁY TÍNH
ĐƯỢC ĐƯỢC

(21)1-2022-06112

(57) Sáng chế đề xuất một số kỹ thuật để mã hóa video và giải mã video. Một phương pháp ví dụ bao gồm bước thực hiện sự biến đổi giữa ảnh con trong ảnh video của video và dòng bit của video theo quy tắc. Quy tắc này chỉ ra rằng, trong trường hợp ảnh con được xử lý như ảnh video để biến đổi, sự hạn chế căn chỉnh ngang lớp được áp dụng cho ít hơn tất cả nhiều lớp này bao gồm lớp hiện tại mà bao gồm ảnh con và tập con của các lớp được liên kết với lớp hiện tại.

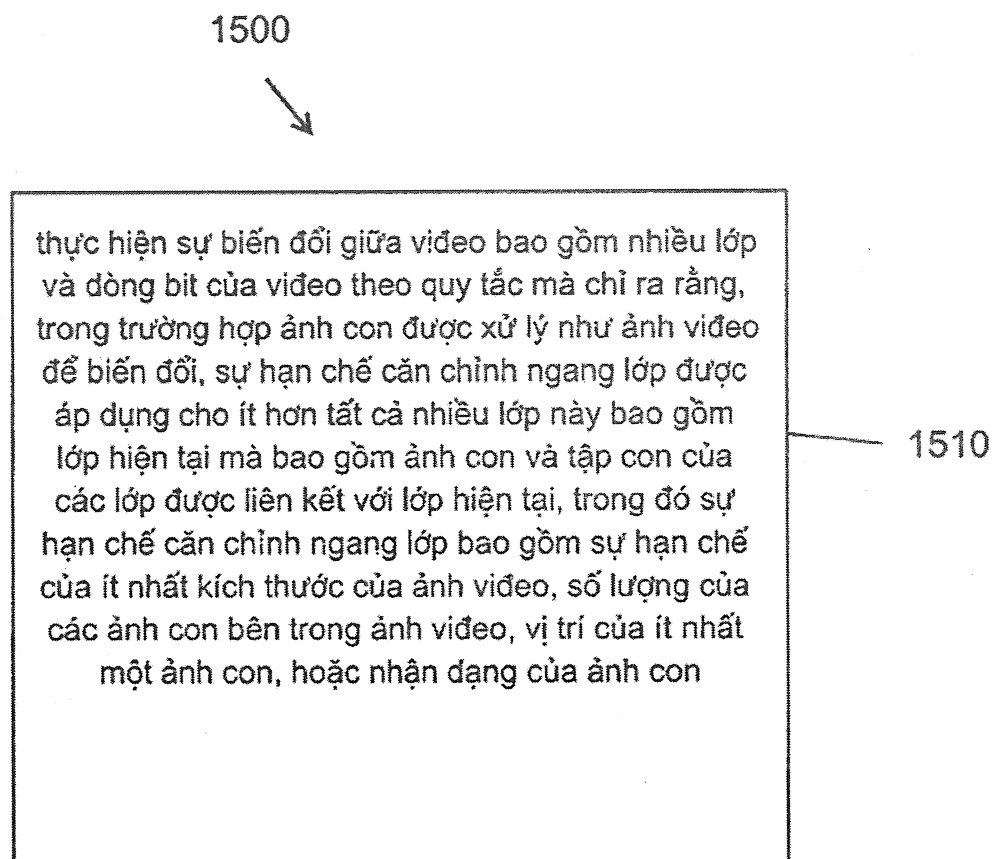


FIG. 15

Lĩnh vực kỹ thuật được đề cập

Sáng chế liên quan tới việc tạo mã và giải mã hình ảnh và video.

Tình trạng kỹ thuật của sáng chế

Video kỹ thuật số sử dụng băng thông lớn nhất trên internet và các mạng truyền thông kỹ thuật số khác. Do số lượng các thiết bị người dùng được kết nối có khả năng nhận và hiển thị video tăng lên, dự kiến rằng yêu cầu về băng thông dành cho việc sử dụng video kỹ thuật số sẽ tiếp tục tăng trưởng.

Bản chất kỹ thuật của sáng chế

Sáng chế đề xuất các kỹ thuật mà có thể được sử dụng bởi các bộ mã hóa và các bộ giải mã video để xử lý sự biểu diễn được tạo mã của video bằng cách sử dụng thông tin điều khiển hữu ích để giải mã sự biểu diễn được tạo mã.

Trong một khía cạnh ví dụ, phương pháp xử lý video được đề xuất. Phương pháp bao gồm bước thực hiện sự biến đổi giữa ảnh hiện tại của video và dòng bit của video theo quy tắc. Quy tắc này chỉ ra rằng nhiều phần tử cú pháp được sử dụng để chỉ ra việc sử dụng của công cụ lấy mẫu lại ảnh tham chiếu mà trong đó ảnh tham chiếu mà có độ phân giải khác với ảnh hiện tại được lấy mẫu lại để biến đổi.

Trong một khía cạnh ví dụ khác, phương pháp xử lý video được đề xuất. Phương pháp bao gồm bước thực hiện sự biến đổi giữa ảnh hiện tại của video và dòng bit của video theo quy tắc. Quy tắc này chỉ ra rằng phần tử cú pháp mà có giá trị không phải nhị phân được sử dụng để chỉ ra việc sử dụng của (1) công cụ lấy mẫu lại ảnh tham chiếu mà trong đó ảnh tham chiếu mà có độ phân giải khác với ảnh hiện tại được lấy mẫu lại và (2) sự thay đổi của độ phân giải ảnh bên trong dãy video lớp được tạo mã (CLVS).

Trong một khía cạnh ví dụ khác, phương pháp xử lý video được đề xuất. Phương pháp bao gồm bước thực hiện sự biến đổi giữa video bao gồm nhiều lớp và dòng bit của video theo quy tắc. Quy tắc này chỉ ra rằng, trong trường hợp ảnh con được xử lý như ảnh video để biến đổi, sự hạn chế căn chỉnh ngang lớp được áp dụng cho ít hơn tất cả nhiều lớp này bao gồm lớp hiện tại mà bao gồm ảnh con và tập con của các lớp được liên kết với lớp hiện tại, trong đó sự hạn chế căn chỉnh ngang lớp bao gồm sự hạn chế của ít nhất kích thước của ảnh video, số lượng của các ảnh con bên trong ảnh video, vị trí của ít nhất một ảnh con, hoặc nhận dạng của ảnh con.

Trong một khía cạnh ví dụ khác, phương pháp xử lý video được đề xuất. Phương pháp bao gồm bước thực hiện sự biến đổi giữa lớp hiện tại của video và dòng bit của video theo quy tắc. Quy tắc này chỉ ra rằng sự hạn chế căn chỉnh ngang lớp được áp dụng cho tất cả các lớp trong cây phụ thuộc được liên kết với lớp hiện tại không phụ thuộc vào việc liệu lớp bất kỳ trong số tất cả các lớp có là lớp đầu ra trong tập lớp đầu ra. Sự hạn chế căn chỉnh ngang lớp bao gồm sự hạn chế của ít nhất kích thước của ảnh video, số lượng của các ảnh con bên trong ảnh video, vị trí của ít nhất một ảnh con, hoặc nhận dạng của ảnh con. Tất cả các lớp trong cây phụ thuộc bao gồm lớp hiện tại, tất cả các lớp mà có lớp hiện tại làm lớp tham chiếu, và tất cả các lớp tham chiếu của lớp hiện tại.

Trong một khía cạnh ví dụ khác, phương pháp xử lý video được đề xuất. Phương pháp bao gồm bước thực hiện sự biến đổi giữa ảnh hiện tại của video và dòng bit của video theo quy tắc. Quy tắc này chỉ ra rằng việc lấy mẫu lại của ảnh tham chiếu mà ở trong cùng lớp như ảnh hiện tại là được phép không phụ thuộc vào giá trị của phần tử cú pháp mà chỉ ra việc liệu sự thay đổi của độ phân giải ảnh có được cho phép trong dãy video lớp được tạo mã (CLVS).

Trong một khía cạnh ví dụ khác, phương pháp xử lý video được đề xuất. Phương pháp bao gồm bước thực hiện sự biến đổi giữa ảnh hiện tại của video bao gồm nhiều lớp và dòng bit của video theo quy tắc. Quy tắc này chỉ ra một trong số: (1) ảnh tham chiếu của ảnh hiện tại là không được phép để được sắp xếp, hoặc (2) trong trường hợp mà ảnh tham chiếu của ảnh hiện tại được sắp xếp,

thì các vector chuyển động mà trở tới ảnh tham chiếu được sử dụng trong khi biến đổi của ảnh hiện tại mà không định tỷ lệ.

Trong một khía cạnh ví dụ khác, phương pháp xử lý video được đề xuất. Phương pháp bao gồm bước thực hiện sự biến đổi giữa video và dòng bit của video theo quy tắc. Quy tắc này chỉ ra rằng thông số dịch vị của số định tỷ lệ là giống nhau đối với hai ảnh video bất kỳ trong cùng dãy video lớp được tạo mã (CLVS) hoặc dãy video được tạo mã (CVS) mà có kích thước giống nhau được biểu diễn trong số lượng các mẫu độ sáng.

Trong một khía cạnh ví dụ khác, phương pháp xử lý video được đề xuất. Phương pháp bao gồm bước thực hiện sự biến đổi giữa ảnh hiện tại của video và dòng bit của video theo quy tắc. Quy tắc này chỉ ra rằng, đáp lại độ phân giải ảnh của ảnh hiện tại là khác với ít nhất một ảnh khác trong cùng đơn vị truy nhập của ảnh hiện tại, công cụ tạo mã dự báo ngoài lớp là được phép chỉ trong trường hợp mà ảnh hiện tại là ảnh điểm truy nhập ngẫu nhiên trong (IRAP).

Trong một khía cạnh ví dụ khác, phương pháp xử lý video được đề xuất. Phương pháp bao gồm bước thực hiện sự biến đổi giữa video bao gồm một hoặc nhiều ảnh video, trong đó sự biểu diễn được tạo mã tương thích với quy tắc định dạng; trong đó quy tắc định dạng chỉ ra rằng hai hoặc nhiều trường cú pháp trong tập thông số dãy điều khiển sự thay đổi độ phân giải ảnh tham chiếu (RPR) trong video.

Trong một khía cạnh ví dụ khác, một phương pháp xử lý video khác được đề xuất. Phương pháp bao gồm bước thực hiện sự biến đổi giữa video bao gồm một hoặc nhiều ảnh video, trong đó sự biểu diễn được tạo mã tương thích với quy tắc định dạng; trong đó quy tắc định dạng chỉ ra rằng một trường cú pháp trong tập thông số dãy điều khiển sự thay đổi độ phân giải ảnh tham chiếu (RPR) trong video; và trong đó, quy tắc định dạng chỉ ra rằng, không phụ thuộc vào giá trị của một trường cú pháp này, việc lấy mẫu lại của ảnh tham chiếu ngoài lớp là được phép để biến đổi.

Trong một khía cạnh ví dụ khác, một phương pháp xử lý video khác được đề xuất. Phương pháp bao gồm bước thực hiện sự biến đổi giữa video bao gồm một hoặc nhiều lớp bao gồm một hoặc nhiều ảnh video bao gồm một hoặc nhiều

ảnh con, trong đó sự biểu diễn được tạo mã tương thích với quy tắc định dạng; trong đó quy tắc định dạng chỉ ra điều kiện ràng buộc thứ nhất đối với căn chỉnh ngang lớp hoặc điều kiện ràng buộc thứ hai đối với sự kết hợp của các ảnh con và khả năng mở rộng của các ảnh ngoài lớp.

Trong một khía cạnh ví dụ khác, một phương pháp xử lý video khác được đề xuất. Phương pháp bao gồm bước thực hiện sự biến đổi giữa video bao gồm một hoặc nhiều lớp bao gồm một hoặc nhiều ảnh video bao gồm một hoặc nhiều ảnh con, trong đó sự biến đổi tương thích với quy tắc định dạng mà chỉ ra rằng ảnh tham chiếu ngoài lớp hoặc ảnh tham chiếu dài hạn là không được phép làm ảnh được sắp xếp của ảnh hiện tại để biến đổi.

Trong một khía cạnh ví dụ khác, một phương pháp xử lý video khác được đề xuất. Phương pháp bao gồm bước thực hiện sự biến đổi giữa video bao gồm nhiều ảnh và sự biểu diễn được tạo mã của video, trong đó sự biến đổi tương thích với quy tắc mà chỉ ra rằng các giá trị của mỗi trong số `scaling_win_left_offset`, `scaling_win_right_offset`, `scaling_win_top_offset`, và `scaling_win_bottom_offset` là giống nhau đối với hai ảnh bất kỳ bên trong cùng dãy video lớp được tạo mã hoặc dãy video được tạo mã có các giá trị của `pic_width_in_luma_samples` và `pic_height_in_luma_samples` giống nhau.

Trong một khía cạnh ví dụ khác, một phương pháp xử lý video khác được đề xuất. Phương pháp bao gồm bước thực hiện sự biến đổi giữa video bao gồm nhiều ảnh và sự biểu diễn được tạo mã của video, trong đó sự biến đổi tương thích với quy tắc mà chỉ ra rằng trong trường hợp mà độ phân giải ảnh hoặc của số định tỷ lệ là khác nhau đối với ảnh hiện tại và ảnh khác trong cùng đơn vị truy nhập, thì dự báo ngoài lớp chỉ được cho phép khi ảnh hiện tại là ảnh điểm truy nhập ngẫu nhiên trong.

Trong một khía cạnh ví dụ khác nữa, thiết bị mã hóa video được đề xuất. Bộ mã hóa video bao gồm bộ xử lý được tạo cấu hình để thực hiện các phương pháp được mô tả trên đây.

Trong một khía cạnh ví dụ khác nữa, thiết bị giải mã video được đề xuất. Bộ giải mã video bao gồm bộ xử lý được tạo cấu hình để thực hiện các phương pháp được mô tả trên đây.

Trong một khía cạnh ví dụ khác nữa, môi trường máy tính đọc được có mã được lưu trên đó được đề xuất. Mã này thực hiện một trong số các phương pháp được mô tả ở đây dưới dạng mã bộ xử lý thực thi được.

Các dấu hiệu này và khác nữa được mô tả trong toàn bộ bản mô tả này.

Mô tả vắn tắt các hình vẽ

FIG. 1 thể hiện ví dụ về phân vùng lát quét màn hình của ảnh, trong đó ảnh được chia thành 12 ô ảnh và 3 lát quét màn hình.

FIG. 2 thể hiện ví dụ về phân vùng lát hình chữ nhật của ảnh, trong đó ảnh được chia thành 24 ô ảnh (6 cột ô ảnh và 4 hàng ô ảnh) và 9 lát hình chữ nhật.

FIG. 3 thể hiện ví dụ của ảnh được phân vùng thành các ô ảnh và các lát hình chữ nhật, trong đó ảnh được chia thành 4 ô ảnh (2 cột ô ảnh và 2 hàng ô ảnh) và 4 lát hình chữ nhật.

FIG. 4 thể hiện ảnh mà được phân vùng thành 15 ô ảnh, 24 lát và 24 ảnh con.

FIG. 5 là sơ đồ khối của hệ thống xử lý video ví dụ.

FIG. 6 là sơ đồ khối của thiết bị xử lý video.

FIG. 7 là lưu đồ của phương pháp xử lý video ví dụ.

FIG. 8 là sơ đồ khối mà minh họa hệ thống tạo mã video theo một số phương án của sáng chế.

FIG. 9 là sơ đồ khối mà minh họa bộ mã hóa theo một số phương án của sáng chế.

FIG. 10 là sơ đồ khối mà minh họa bộ giải mã theo một số phương án của sáng chế.

FIG. 11 thể hiện ví dụ của sơ đồ tạo mã video 360° lệ thuộc công nhìn dựa trên ảnh con điển hình.

FIG. 12 thể hiện sơ đồ tạo mã video 360° lệ thuộc công nhìn dựa trên các ảnh con và khả năng mở rộng không gian.

FIG. 13 là sự biểu diễn dưới dạng lưu đồ của phương pháp xử lý video theo công nghệ theo sáng chế.

FIG. 14 là sự biểu diễn dưới dạng lưu đồ của phương pháp xử lý video khác theo công nghệ theo sáng chế.

FIG. 15 là sự biểu diễn dưới dạng lưu đồ của phương pháp xử lý video khác theo công nghệ theo sáng chế.

FIG. 16 là sự biểu diễn dưới dạng lưu đồ của phương pháp xử lý video khác theo công nghệ theo sáng chế.

FIG. 17 là sự biểu diễn dưới dạng lưu đồ của phương pháp xử lý video khác theo công nghệ theo sáng chế.

FIG. 18 là sự biểu diễn dưới dạng lưu đồ của phương pháp xử lý video khác theo công nghệ theo sáng chế.

FIG. 19 là sự biểu diễn dưới dạng lưu đồ của phương pháp xử lý video khác theo công nghệ theo sáng chế.

FIG. 20 là sự biểu diễn dưới dạng lưu đồ của phương pháp xử lý video khác nữa theo công nghệ theo sáng chế.

Mô tả chi tiết sáng chế

Các tiêu đề của các phần được sử dụng trong bản mô tả để làm cho dễ hiểu và không giới hạn khả năng ứng dụng của các kỹ thuật và các phương án được mô tả trong mỗi phần chỉ trong phần đó. Hơn nữa, thuật ngữ của chuẩn H.266 được sử dụng trong một số phần mô tả chỉ để làm cho dễ hiểu và không nhằm giới hạn phạm vi của các kỹ thuật được mô tả. Như vậy, các kỹ thuật được mô tả ở đây cũng có thể được áp dụng cho các giao thức tạo mã-giải mã video và các thiết kế khác. Trong bản mô tả này, các thay đổi biên tập được thể hiện trong văn bản bằng cách gạch ngang để chỉ báo đoạn văn bản bị loại bỏ và làm nổi bật để chỉ báo đoạn văn bản được thêm vào (bao gồm in đậm nghiêng), liên quan tới dự thảo hiện tại của đặc tả VVC.

1. Tổng quan

Bản mô tả này đề cập đến các công nghệ tạo mã video. Cụ thể là, về 1) sự kết hợp của hai hoặc nhiều trong số việc lấy mẫu lại ảnh tham chiếu (RPR), các ảnh con, và khả năng mở rộng trong tạo mã video, 2) việc sử dụng RPR giữa ảnh hiện tại và các ảnh tham chiếu mà có cùng độ phân giải không gian, và 3) sự kết

hợp của ảnh tham chiếu dài hạn và ảnh được sắp xếp. Các ý đồ này có thể được áp dụng một cách riêng rẽ hoặc trong sự kết hợp khác nhau, vào bộ mã hóa-giải mã video tiêu chuẩn hoặc phi tiêu chuẩn tạo mã video bất kỳ mà hỗ trợ tạo mã video nhiều lớp, ví dụ, chuẩn Tạo mã video đa năng (Versatile Video Coding - VVC) đang được phát triển.

2. Các chữ viết tắt

APS	Tập thông số thích ứng
AU	Đơn vị truy nhập
AUD	Dấu cách đơn vị truy nhập
AVC	Tạo mã video nâng cao
CLVS	Dãy video lớp được tạo mã
CPB	Bộ đệm ảnh được tạo mã
CRA	Truy nhập ngẫu nhiên sạch
CTU	Đơn vị cây tạo mã
CVS	Dãy video được tạo mã
DCI	Thông tin khả năng giải mã
DPB	Bộ đệm ảnh được giải mã
EOB	Cuối dòng bit
EOS	Cuối dãy
GDR	Làm mới giải mã từng bậc
HEVC	Tạo mã video hiệu suất cao
HRD	Bộ giải mã tham chiếu giả định
IDR	Làm mới giải mã tức thời
ILP	Dự báo ngoài lớp
ILRP	Ảnh tham chiếu ngoài lớp
IRAP	Ảnh truy nhập ngẫu nhiên trong
JEM	Mô hình nghiên cứu chung
LTRP	Ảnh tham chiếu dài hạn
MCTS	Các tập ô ảnh hạn chế chuyển động
NAL	Lớp trừu tượng mạng
OLS	Tập lớp đầu ra

PH Tiêu đề ảnh
 PPSTập thông số ảnh
 PTLTiêu sử, Cấp và Bậc
 PU Đơn vị ảnh
 RAP Điểm truy nhập ngẫu nhiên
 RBSP Tải đầy byte thô
 SEI Thông tin tăng cường bổ sung
 SPSTập thông số dãy
 STRP Ảnh tham chiếu ngắn hạn
 SVC Tạo mã video khả mở
 VCL Lớp tạo mã video
 VPS Tập thông số video
 VTM Mô hình thử VVC
 VUI Thông tin khả năng sử dụng video
 VVC Tạo mã video đa năng

3. Nhập đề

Các chuẩn tạo mã video đã tiến hóa chủ yếu thông qua sự phát triển của các chuẩn ITU-T và ISO/IEC đã biết. ITU-T đã đề xuất H.261 và H.263, ISO/IEC đã đề xuất MPEG-1 và MPEG-4 Visual, và hai tổ chức này cùng đề xuất các chuẩn H.262/MPEG-2 Video và H.264/MPEG-4 Advanced Video Coding (AVC) và H.265/HEVC. Kể từ chuẩn H.262, các chuẩn tạo mã video là dựa trên cấu trúc tạo mã video lai trong đó việc tạo mã dự báo theo thời gian cộng chuyển đổi được sử dụng. Để tìm kiếm các công nghệ tạo mã video tương lai sau HEVC, Joint Video Exploration Team (JVET) đã được đồng thành lập bởi VCEG và MPEG trong năm 2015. Kể từ đó, nhiều phương pháp mới đã được chấp nhận bởi JVET và được đưa vào phần mềm tham chiếu được gọi là Joint Exploration Model (JEM). Hội nghị JVET được đồng thời tổ chức mỗi lần một quý, và chuẩn tạo mã mới nhằm giảm 50% tốc độ bit so với HEVC. Chuẩn tạo mã video mới được đặt tên chính thức là chuẩn Tạo mã video vạn năng (VVC) tại hội nghị JVET vào tháng tư năm 2018, và phiên bản thứ nhất của mô hình thử VVC (VTM) đã được phát hành vào lúc đó. Do có các nỗ lực không ngừng nghỉ đóng góp vào

chuẩn hóa VVC, các kỹ thuật tạo mã mới được chấp nhận cho chuẩn VVC trong mỗi hội nghị JVET. Dự thảo làm việc VVC và mô hình thử VTM được cập nhật sau mỗi hội nghị. Dự án VVC nhằm hoàn thiện kỹ thuật (FDIS) tại hội nghị tháng 7 năm 2020.

3.1. Các sơ đồ phân vùng ảnh trong HEVC

HEVC bao gồm bốn sơ đồ phân vùng ảnh khác nhau, mà chính là các lát chữ nhật, các lát phụ thuộc, các ô ảnh, và xử lý song song đầu sóng (Wavefront Parallel Processing - WPP), mà có thể được áp dụng cho việc so khớp kích thước đơn vị truyền cực đại (Maximum Transfer Unit - MTU), xử lý song song, và độ trễ từ đầu này tới đầu kia được giảm thấp.

Các lát chữ nhật là giống như trong H.264/AVC. Mỗi lát chữ nhật được đóng gói trong bản thân đơn vị NAL của mình, và dự báo trong ảnh (dự báo trong mẫu, dự báo thông tin chuyển động, dự báo chế độ tạo mã) và sự phụ thuộc tạo mã entropi ngang qua các đường biên lát là không được phép. Do đó lát chữ nhật có thể được tái tạo không phụ thuộc vào các lát chữ nhật khác bên trong cùng một ảnh (mặc dù vẫn có thể có các sự phụ thuộc lẫn nhau do các hoạt động lọc vòng).

Lát chữ nhật là công cụ duy nhất có thể được sử dụng để song song hóa mà cũng khả dụng, dưới dạng gần như giống hệt, trong H.264/AVC. Việc song song hóa dựa trên các lát chữ nhật không yêu cầu nhiều truyền thông liên bộ xử lý hoặc liên lõi (ngoại trừ chia sẻ dữ liệu liên bộ xử lý hoặc liên lõi để bù chuyển động khi giải mã ảnh được tạo mã theo cách dự báo, mà thường nặng hơn nhiều chia sẻ dữ liệu liên bộ xử lý hoặc liên lõi do dự báo trong ảnh). Tuy nhiên, cũng vì lý do này, việc sử dụng các lát chữ nhật có thể phải chịu chi phí phụ tạo mã đáng kể do chi phí bit của tiêu đề lát và do không có dự báo ngang qua các đường biên lát. Tiếp theo, các lát chữ nhật (trái ngược với các công cụ khác được nêu dưới đây) cũng làm nhiệm vụ của cơ cấu khóa mã đối với việc phân vùng dòng bit để đáp ứng các yêu cầu về kích thước MTU, do sự phụ thuộc trong ảnh của các lát chữ nhật và mỗi lát chữ nhật được đóng gói trong bản thân đơn vị NAL của mình. Trong nhiều trường hợp, mục đích của song song hóa và mục đích của so khớp kích thước MTU đặt ra các đòi hỏi trái ngược đối với sự bố trí lát trong

ảnh. Việc hiện thực hóa tình huống này dẫn đến sự phát triển của các công cụ song song hóa được nêu dưới đây.

Các lát phụ thuộc có các tiêu đề lát ngắn và cho phép phân vùng dòng bit ở các biên khối cây mà không phá vỡ dự báo trong ảnh bất kỳ. Về cơ bản, các lát phụ thuộc tạo ra sự phân đoạn của các lát chữ nhật thành nhiều đơn vị NAL, để giảm thấp độ trễ từ đầu này tới đầu kia bằng cách cho phép một phần của lát chữ nhật được gửi ra trước khi việc mã hóa của toàn bộ lát chữ nhật được kết thúc.

Trong WPP, ảnh được phân vùng thành các hàng đơn của các khối cây tạo mã (coding tree block - CTB). Giải mã entropi và dự báo được cho phép sử dụng dữ liệu từ các CTB trong các phân vùng khác. Việc xử lý song song là có khả năng thông qua giải mã song song của các hàng CTB, trong đó sự bắt đầu của việc giải mã hàng CTB bị trễ bởi hai CTB, để bảo đảm rằng dữ liệu liên quan tới CTB bên trên và bên phải CTB được xử lý là khả dụng trước khi CTB được xử lý được giải mã. Nhờ sử dụng sự bắt đầu so le này (mà trông giống như đầu sóng khi được biểu diễn bằng đồ họa) mà việc song song hóa là có khả năng với số bộ xử lý/lỗi nhiều bằng số lượng mà ảnh chứa các hàng CTB. Do dự báo trong ảnh giữa các hàng khối cây lân cận bên trong ảnh là được phép, truyền thông liên bộ xử lý/liên lỗi được yêu cầu để cho phép dự báo trong ảnh có thể là đáng kể. Việc phân vùng WPP không dẫn đến việc tạo ra các đơn vị NAL bổ sung so với khi nó không được áp dụng, nên WPP không phải là công cụ để so khớp kích thước MTU. Tuy nhiên, nếu cần phải so khớp kích thước MTU, thì các lát chữ nhật có thể được sử dụng với WPP, với một số chi phí phụ tạo mã đã biết.

Các ô ảnh xác định các biên ngang và dọc mà phân vùng ảnh thành các cột và các hàng ô ảnh. Cột ô ảnh chạy từ đỉnh của ảnh tới đáy của ảnh. Tương tự, hàng ô ảnh chạy từ bên trái của ảnh sang bên phải của ảnh. Số lượng của các ô ảnh trong ảnh có thể được suy ra một cách đơn giản là số lượng của các cột ô ảnh nhân với số lượng của các hàng ô ảnh.

Thứ tự quét của các CTB được thay đổi thành nội bộ bên trong ô ảnh (theo thứ tự quét mảnh CTB của ô ảnh), trước khi giải mã CTB trên cùng bên trái của ô ảnh kế tiếp theo thứ tự quét mảnh ô ảnh của ảnh. Tương tự như các lát chữ nhật, các ô ảnh làm gãy các phụ thuộc dự báo trong ảnh cũng như các phụ thuộc giải

mã entropi. Tuy nhiên, chúng không cần phải được bao gồm trong các đơn vị NAL riêng lẻ (về mặt này giống như WPP); do các ô ảnh không thể được sử dụng để so khớp kích thước MTU. Mỗi ô ảnh có thể được xử lý bởi một bộ xử lý/lỗi, và truyền thông liên bộ xử lý/liên lỗi được yêu cầu để dự báo trong ảnh giữa các đơn vị xử lý giải mã các ô ảnh lân cận được giới hạn ở vận chuyển tiêu đề lát được chia sẻ trong các trường hợp mà lát mở rộng hơn một ô ảnh, và sự chia sẻ liên quan đến lọc vòng của các mẫu được tái tạo và siêu dữ liệu. Khi nhiều hơn một ô ảnh hoặc đoạn WPP được bao gồm trong lát, dịch vị byte điểm nhập đối với mỗi ô ảnh hoặc đoạn WPP khác với dịch vị đầu tiên trong lát được báo tín hiệu trong tiêu đề lát.

Để đơn giản hóa, các hạn chế đối với việc áp dụng bốn sơ đồ phân vùng ảnh khác nhau đã được chỉ ra trong HEVC. Dãy video được tạo mã đã cho không thể bao gồm cả các ô ảnh lẫn các đầu sóng đối với phần lớn trong số các tiểu sử được chỉ ra trong HEVC. Đối với mỗi lát và ô ảnh, từng điều kiện hoặc cả hai điều kiện sau đây phải được thỏa mãn: 1) tất cả các khối cây được tạo mã trong lát thuộc về cùng một ô ảnh; 2) tất cả các khối cây được tạo mã trong ô ảnh thuộc về cùng một lát. Cuối cùng, đoạn đầu sóng chứa chính xác một hàng CTB, và khi WPP được sử dụng, nếu lát bắt đầu bên trong hàng CTB, thì nó phải kết thúc trong cùng hàng CTB.

Với sự sửa đổi gần đây đối với HEVC, HEVC chỉ ra ba thông điệp SEI liên quan đến MCTS, mà chính là thông điệp SEI MCTS theo thời gian, thông điệp SEI tập thông tin trích xuất MCTS, và thông điệp SEI lồng thông tin trích xuất MCTS.

Thông điệp SEI MCTS theo thời gian chỉ báo sự tồn tại của các MCTS trong dòng bit và báo tín hiệu các MCTS. Đối với mỗi MCTS, các vector chuyển động bị giới hạn để trở tới các vị trí mẫu toàn phần bên trong MCTS và tới các vị trí mẫu phân số mà chỉ yêu cầu các vị trí mẫu toàn phần bên trong MCTS để nội suy, và việc sử dụng các ứng viên vector chuyển động để dự báo vector chuyển động theo thời gian được suy ra từ các khối bên ngoài MCTS không được cho phép. Theo cách này, mỗi MCTS có thể được giải mã theo cách không phụ thuộc mà không có sự tồn tại của các ô ảnh không được bao gồm trong MCTS.

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

3.2. Phân vùng các ảnh trong VVC

Trong VVC, ảnh được chia thành một hoặc nhiều hàng ô ảnh và một hoặc nhiều cột ô ảnh. Ô ảnh là dãy của các CTU mà bao phủ vùng hình chữ nhật của ảnh. Các CTU trong ô ảnh được quét theo thứ tự quét mảnh bên trong ô ảnh đó.

Lát chứa số nguyên các ô ảnh hoàn chỉnh hoặc số nguyên các hàng CTU hoàn chỉnh liên tiếp bên trong ô ảnh của ảnh.

Hai chế độ của các lát được hỗ trợ, mà chính là chế độ lát quét mảnh và chế độ lát hình chữ nhật. Trong chế độ lát quét mảnh, lát chứa dãy của các ô ảnh hoàn chỉnh trong quét mảnh ô ảnh của ảnh. Trong chế độ lát hình chữ nhật, lát chứa hoặc số lượng của các ô ảnh hoàn chỉnh mà cùng nhau tạo thành vùng hình chữ nhật của ảnh hoặc số lượng của các hàng CTU hoàn chỉnh liên tiếp của một ô ảnh mà cùng nhau tạo thành vùng hình chữ nhật của ảnh. Các ô ảnh bên trong lát hình chữ nhật được quét theo thứ tự quét mảnh ô ảnh bên trong vùng hình chữ nhật tương ứng với lát đó.

Ảnh con chứa một hoặc nhiều lát mà cùng nhau bao phủ vùng hình chữ nhật của ảnh.

FIG. 1 thể hiện ví dụ về phân vùng lát quét mảnh của ảnh, trong đó ảnh được chia thành 12 ô ảnh và 3 lát quét mảnh.

FIG. 2 thể hiện ví dụ về phân vùng lát hình chữ nhật của ảnh, trong đó ảnh được chia thành 24 ô ảnh (6 cột ô ảnh và 4 hàng ô ảnh) và 9 lát hình chữ nhật.

FIG. 3 thể hiện ví dụ về ảnh được phân vùng thành các ô ảnh và các lát hình chữ nhật, trong đó ảnh được chia thành 4 ô ảnh (2 cột ô ảnh và 2 hàng ô ảnh) và 4 lát hình chữ nhật.

FIG. 4 thể hiện ví dụ về phân vùng ảnh con của ảnh, trong đó ảnh được phân vùng thành 18 ô ảnh, 12 ở phía tay trái mỗi ô che phủ một lát gồm 4 nhân 4 CTU và 6 ô ảnh ở phía tay phải mỗi ô che phủ 2 lát được xếp chồng thẳng đứng gồm 2 nhân 2 CTU, tất cả cùng tạo thành 24 lát và 24 ảnh con có các kích thước thay đổi (mỗi lát là một ảnh con).

3.3. Sự thay đổi độ phân giải ảnh bên trong dãy

Trong AVC và HEVC, độ phân giải không gian của các ảnh không thể thay đổi trừ phi dãy mới mà sử dụng SPS mới bắt đầu, với ảnh IRAP. VVC cho phép thay đổi độ phân giải ảnh bên trong dãy ở một vị trí mà không cần mã hóa ảnh IRAP, mà luôn luôn được tạo mã trong. Dấu hiệu này đôi khi được gọi là lấy mẫu lại ảnh tham chiếu (RPR), do dấu hiệu này cần phải lấy mẫu lại ảnh tham chiếu được sử dụng để dự báo ngoài khi ảnh tham chiếu đó có độ phân giải khác với ảnh hiện tại mà đang được giải mã.

Hệ số định tỷ lệ bị giới hạn là lớn hơn hoặc bằng $1/2$ (2 lần giảm kích thước mẫu từ ảnh tham chiếu tới ảnh hiện tại), và nhỏ hơn hoặc bằng 8 (8 lần tăng kích thước mẫu). Ba nhóm của các bộ lọc lấy mẫu lại với các tần suất cắt khác nhau được chỉ ra để xử lý các hệ số định tỷ lệ khác nhau giữa ảnh tham chiếu và ảnh hiện tại. Ba nhóm của các bộ lọc lấy mẫu lại này lần lượt được áp dụng cho hệ số định tỷ lệ nằm trong khoảng từ $1/2$ tới $1/1,75$, từ $1/1,75$ tới $1/1,25$, và từ $1/1,25$ tới 8. Mỗi nhóm của các bộ lọc lấy mẫu lại có 16 pha đối với độ sáng và 32 pha đối với sắc độ mà giống như trường hợp của các bộ lọc nội suy bù chuyển động. Thực tế quá trình nội suy MC thông thường là trường hợp cụ thể của quá trình lấy mẫu lại với hệ số định tỷ lệ nằm trong khoảng từ $1/1,25$ tới 8. Các hệ số định tỷ lệ ngang và dọc là được rút ra dựa trên chiều rộng ảnh và chiều cao ảnh, và các dịch vị định tỷ lệ trái, phải, đỉnh và đáy được chỉ ra đối với ảnh tham chiếu và ảnh hiện tại.

Các khía cạnh khác của thiết kế VVC để hỗ trợ dấu hiệu này mà khác với HEVC bao gồm: i) Độ phân giải ảnh và cửa sổ tương thích tương ứng được báo

tín hiệu trong PPS thay vì trong SPS, trong khi trong SPS độ phân giải ảnh tối đa được báo tín hiệu. ii) Đối với dòng bit một lớp, mỗi ảnh lưu trữ (khe trong DPB để lưu trữ của một ảnh được giải mã) chiếm kích thước bộ đệm theo yêu cầu để lưu trữ ảnh được giải mã có độ phân giải ảnh tối đa.

3.4. Tạo mã video khả mở (SVC) nói chung và trong chuẩn VVC

Tạo mã video khả mở (SVC, đôi khi còn được gọi là khả năng mở rộng trong tạo mã video) dùng để chỉ việc tạo mã video mà trong đó lớp đề (BL), đôi khi được gọi là lớp tham chiếu (RL), và một hoặc nhiều lớp tăng cường khả mở (EL) được sử dụng. Trong SVC, lớp đề có thể mang dữ liệu video với bậc cơ sở của chất lượng. Một hoặc nhiều lớp tăng cường có thể mang dữ liệu video bổ sung để hỗ trợ, ví dụ, các bậc không gian, thời gian, và/hoặc tín hiệu trên nhiều (SNR) cao hơn. Các lớp tăng cường có thể được xác định tương đối với lớp được mã hóa trước đó. Ví dụ, lớp đáy có thể dùng làm BL, trong khi lớp đỉnh có thể dùng làm EL. Các lớp giữa có thể dùng làm hoặc các EL hoặc các RL, hoặc cả hai. Ví dụ, lớp giữa (ví dụ, lớp mà không phải lớp thấp nhất cũng không phải lớp cao nhất) có thể là EL đối với các lớp bên dưới lớp giữa, như lớp đề hoặc các lớp tăng cường xen kẽ bất kỳ, và đồng thời dùng làm RL đối với một hoặc nhiều lớp tăng cường bên trên lớp giữa. Tương tự, trong phân mở rộng đa khung nhìn hoặc 3D của chuẩn HEVC, có thể có nhiều khung nhìn, và thông tin về một khung nhìn có thể được sử dụng để tạo mã (ví dụ, mã hóa hoặc giải mã) thông tin về khung nhìn khác (ví dụ, ước tính chuyển động, dự báo vector chuyển động và/hoặc các độ dư khác).

Trong SVC, các thông số được sử dụng bởi bộ mã hóa hoặc bộ giải mã được nhóm lại thành các tập thông số dựa trên bậc tạo mã (ví dụ, bậc video, bậc dãy, bậc ảnh, bậc lát, v.v.) mà trong đó chúng có thể được sử dụng. Ví dụ, các thông số mà có thể được sử dụng bởi một hoặc nhiều dãy video được tạo mã của các lớp khác nhau trong dòng bit có thể được bao gồm trong tập thông số video (VPS), và các thông số mà được sử dụng bởi một hoặc nhiều ảnh trong dãy video được tạo mã có thể được bao gồm trong tập thông số dãy (SPS). Tương tự, các thông số mà được sử dụng bởi một hoặc nhiều lát trong ảnh có thể được bao gồm trong tập thông số ảnh (PPS), và các thông số khác mà là cụ thể với một lát có

thể được bao gồm trong tiêu đề lát. Tương tự, sự chỉ báo về việc (các) tập thông số nào mà lớp cụ thể đang sử dụng ở thời điểm đã cho có thể được cấp ở các bậc tạo mã khác nhau.

Nhờ có sự hỗ trợ của việc lấy mẫu lại ảnh tham chiếu (RPR) trong VVC, sự hỗ trợ của dòng bit chứa nhiều lớp, ví dụ, hai lớp với độ phân giải SD và HD trong VVC có thể được thiết kế mà không cần công cụ tạo mã bậc xử lý tín hiệu bổ sung bất kỳ, do việc tăng kích thước mẫu cần thiết để hỗ trợ khả năng mở rộng không gian có thể chỉ sử dụng bộ lọc tăng kích thước mẫu RPR. Tuy nhiên, các thay đổi cú pháp bậc cao (so với việc không hỗ trợ khả năng mở rộng) là cần thiết để hỗ trợ khả năng mở rộng. Sự hỗ trợ khả năng mở rộng được chỉ ra trong phiên bản 1 VVC. Khác với các hỗ trợ khả năng mở rộng trong các chuẩn tạo mã video sớm hơn bất kỳ, bao gồm trong các phần mở rộng của AVC và HEVC, thiết kế của khả năng mở rộng VVC đã được thực hiện một cách thân thiện với các thiết kế bộ giải mã một lớp càng nhiều càng tốt. Khả năng giải mã đối với các dòng bit nhiều lớp được chỉ ra theo cách cứ như chỉ có một lớp trong dòng bit. Ví dụ, khả năng giải mã, như kích thước DPB, được chỉ ra theo cách mà không phụ thuộc vào số lượng của các lớp trong dòng bit cần được giải mã. Về cơ bản, bộ giải mã được thiết kế cho các dòng bit một lớp không cần thay đổi nhiều để có thể giải mã các dòng bit nhiều lớp. So với các thiết kế của các phần mở rộng nhiều lớp của AVC và HEVC, các khía cạnh HLS đã được đơn giản hóa đáng kể khi hy sinh một số mức độ linh hoạt. Ví dụ, IRAP ATI cần phải chứa ảnh đối với mỗi lớp trong số các lớp hiện diện trong CVS.

3.5. Phát trực tiếp video 360° lệ thuộc công nhìn dựa trên các ảnh con

Khi phát trực tiếp video 360°, hay còn gọi là video đẳng hướng, vào thời điểm cụ thể bất kỳ chỉ tập con (ví dụ, công nhìn hiện tại) của toàn bộ hình cầu video đẳng hướng được cấp cho người dùng, trong khi người dùng có thể quay đầu của anh/chị ta bất kỳ lúc nào để thay đổi hướng nhìn và do đó công nhìn hiện tại. Trong khi mong muốn là có ít nhất sự biểu diễn chất lượng thấp nào đó của vùng không được bao phủ bởi công nhìn hiện tại khả dụng đối với khách hàng và sẵn sàng để được cấp cho người dùng chỉ trong trường hợp người dùng bất ngờ thay đổi hướng nhìn của anh/chị ta tới vị trí bất kỳ trên hình cầu này, sự biểu

diễn chất lượng cao của video đẳng hướng chỉ cần thiết đối với công nhìn hiện tại mà đang được cấp để sử dụng ngay bây giờ. Việc tách sự biểu diễn chất lượng cao của toàn bộ video đẳng hướng thành các ảnh con ở độ chi tiết thích hợp cho phép sự tối ưu hóa này. Bằng cách sử dụng VVC, hai sự biểu diễn có thể được mã hóa như hai lớp mà là độc lập với nhau.

Sơ đồ chuyển video 360° lệ thuộc công nhìn dựa trên ảnh con diễn hình được thể hiện trên FIG. 11 trong đó sự biểu diễn độ phân giải cao hơn của video đầy đủ bao gồm các ảnh con, trong khi sự biểu diễn độ phân giải thấp hơn của video đầy đủ không sử dụng các ảnh con và có thể được tạo mã với các điểm truy nhập ngẫu nhiên ít thường xuyên hơn sự biểu diễn độ phân giải cao hơn. Máy khách nhận video đầy đủ trong độ phân giải thấp hơn và đối với độ phân giải cao hơn video nó chỉ nhận và giải mã các ảnh con mà bao phủ công nhìn hiện tại.

3.6. Các tập thông số

AVC, HEVC, và VVC chỉ ra các tập thông số. Các kiểu của các tập thông số bao gồm SPS, PPS, APS, và VPS. SPS và PPS được hỗ trợ trong tất cả AVC, HEVC, và VVC. VPS được đưa vào từ chuẩn HEVC và được bao gồm trong cả HEVC lẫn VVC. APS đã không được bao gồm trong AVC hoặc HEVC mà được bao gồm trong văn bản dự thảo VVC mới nhất.

SPS đã được thiết kế để mang thông tin tiêu đề bậc dãy, và PPS đã được thiết kế để mang thông tin tiêu đề bậc ảnh không thường xuyên thay đổi. Với SPS và PPS, thông tin không thường xuyên thay đổi không cần được nhắc lại đối với mỗi dãy hoặc ảnh, nên có thể tránh được việc báo tín hiệu dư thừa về thông tin này. Hơn nữa, việc sử dụng SPS và PPS cho phép truyền ngoài dải băng thông tin tiêu đề quan trọng, nên không chỉ tránh được việc phải truyền dư thừa mà còn nâng cao độ bền chịu lỗi.

VPS được đưa vào để mang thông tin tiêu đề bậc dãy mà là thông dụng đối với tất cả các lớp trong các dòng bit nhiều lớp.

APS được đưa vào để mang thông tin bậc ảnh hoặc bậc lát mà cần một số bit đáng kể để tạo mã, có thể được chia sẻ bởi nhiều ảnh, và trong dãy có thể có khá nhiều biến thể khác nhau.

4. Các vấn đề kỹ thuật được giải quyết bởi các giải pháp kỹ thuật theo sáng chế

Thiết kế hiện có trong văn bản VVC mới nhất có các vấn đề sau đây:

1) Thiết kế VVC hiện tại hỗ trợ sơ đồ tạo mã điển hình đối với video 360° như được thể hiện trên FIG. 11. Tuy nhiên, mặc dù khả năng mở rộng được hỗ trợ trong thiết kế VVC hiện tại, sơ đồ tạo mã video 360° nâng cao như được thể hiện trên FIG. 12 không được hỗ trợ. Điều khác biệt duy nhất so với cách tiếp cận được thể hiện trên FIG. 11 là dự báo ngoài lớp (ILP) được áp dụng cho cách tiếp cận được thể hiện trên FIG. 12.

Hai chỗ sau đây trong dự thảo VVC không cho phép sự sử dụng kết hợp của các ảnh con và khả năng mở rộng không gian:

a. Thiết kế khả năng mở rộng không gian trong VVC dựa trên dấu hiệu RPR. Tuy nhiên, sự kết hợp của RPR và các ảnh con hiện không được phép bởi điều kiện ràng buộc ngữ nghĩa sau đây:

Khi `res_change_in_clvs_allowed_flag` là bằng 1, giá trị của `subpic_info_present_flag` sẽ bằng 0.

Do đó, sơ đồ tạo mã nâng cao là không được phép, bởi vì đối với SPS mà được tham chiếu bởi lớp cao hơn, điều kiện ràng buộc nêu trên không cho phép thiết lập `subpic_info_present_flag` bằng 1 (để sử dụng nhiều ảnh con trên ảnh) và đồng thời thiết lập `res_change_in_clvs_allowed_flag` bằng 1 (để cho phép RPR, mà là cần thiết đối với khả năng mở rộng không gian với ILP).

b. Dự thảo VVC hiện tại có điều kiện ràng buộc sau đây đối với sự kết hợp của các ảnh con và khả năng mở rộng:

*Khi `subpic_treated_as_pic_flag[i]` là bằng 1, thì yêu cầu về sự tương thích dòng bit là tất cả các điều kiện sau đây là đúng đối với mỗi lớp đầu ra và các lớp tham chiếu của nó trong OLS mà bao gồm lớp chứa ảnh con thứ *i* làm lớp đầu ra:*

– *Tất cả các ảnh trong lớp đầu ra và các lớp tham chiếu của nó sẽ có giá trị giống nhau của*

pic_width_in_luma_samples và giá trị giống nhau của *pic_height_in_luma_samples*.

– Tất cả các SPS được tham chiếu bởi lớp đầu ra và các lớp tham chiếu của nó sẽ có giá trị giống nhau của *sps_num_subpics_minus1* và sẽ lần lượt có các giá trị giống nhau của *subpic_ctu_top_left_x[j]*, *subpic_ctu_top_left_y[j]*, *subpic_width_minus1[j]*, *subpic_height_minus1[j]*, và *loop_filter_across_subpic_enabled_flag[j]*, đối với mỗi giá trị của *j* nằm trong khoảng từ 0 tới *sps_num_subpics_minus1*, bao gồm cả các giá trị biên.

– Tất cả các ảnh trong mỗi đơn vị truy nhập trong lớp đầu ra và các lớp tham chiếu của nó sẽ có giá trị giống nhau của *SubpicIdVal[j]* đối với mỗi giá trị của *j* nằm trong khoảng từ 0 tới *sps_num_subpics_minus1*, bao gồm cả các giá trị biên.

Điều kiện ràng buộc trên đây chủ yếu không cho phép sự kết hợp bất kỳ của các ảnh con và khả năng mở rộng với ILP khác với sự kết hợp bị giới hạn của các ảnh con với khả năng mở rộng SNR, trong đó các lớp bên trong mỗi cây phụ thuộc phải có độ phân giải không gian giống nhau cũng như sơ đồ bố trí ảnh con giống nhau.

2) Khi *subpic_treated_as_pic_flag[i]* là bằng 1, các biên ảnh con của ảnh con thứ *i* sẽ được xử lý như các biên ảnh trong bù chuyển động. Các xử lý này được thực hiện trong văn bản dự thảo VVC bằng cách áp dụng một số hoạt động xén đã biết trong các quá trình giải mã liên quan tới bù chuyển động. Tuy nhiên, đối với sơ đồ tạo mã nâng cao được thể hiện trên FIG. 12 do lớp thấp hơn là hoàn toàn khả dụng đối với bộ giải mã, chứ không chỉ khu vực tương ứng với ảnh con thứ *i*, việc xén này không cần phải được áp dụng trong trường hợp này, để tránh tổn thất hiệu quả tạo mã không cần thiết.

3) Không xét đến sự hỗ trợ của sơ đồ tạo mã nâng cao được thể hiện trên FIG. 12, điều kiện ràng buộc hiện có nêu trên đối với sự kết hợp của các ảnh con và khả năng mở rộng với ILP (phân mô tả này được bao gồm trong phân mô tả của vấn đề 1b) có các vấn đề sau đây:

a. Điều kiện ràng buộc cũng phải được áp dụng khi lớp chứa ảnh con thứ i không phải là lớp đầu ra của OLS. Toàn bộ điều kiện ràng buộc phải được chỉ ra theo cách mà không tính đến việc liệu lớp có là lớp đầu ra của OLS.

b. Yêu cầu đối với giá trị của `subpic_treated_as_pic_flag[i]` cần được căn chỉnh ngang qua các lớp phải được bao gồm, nếu không thì việc trích ra của dãy ảnh con với chỉ số giống nhau ngang qua các lớp sẽ không khả thi.

c. Yêu cầu đối với giá trị của `loop_filter_across_subpic_enabled_flag[i]` cần được căn chỉnh ngang qua các lớp phải được loại trừ, vì không phụ thuộc vào giá trị của cờ này, miễn là `subpic_treated_as_pic_flag[i]` là bằng 1 dãy ảnh con là trích ra được. Việc thiết lập giá trị của `loop_filter_across_subpic_enabled_flag[i]` phải được dành lại cho bộ mã hóa quyết định để cân bằng chất lượng của một dãy ảnh con trích ra được với chất lượng của các nhóm dãy ảnh con trích ra được, chính là lý do mà hai cờ được báo tín hiệu một cách độc lập với nhau.

d. Toàn bộ điều kiện ràng buộc chỉ được áp dụng khi `sps_num_subpics_minus1` là lớn hơn 0, để tránh tất cả các trường hợp của một ảnh con trên ảnh con được bao trùm một cách vô tình bởi điều kiện ràng buộc này.

e. Phạm vi thời gian, ví dụ, tập của các AU, mà trong đó điều kiện ràng buộc được áp dụng cần phải được chỉ ra một cách rõ ràng.

f. Yêu cầu đối với giá trị của mỗi trong số các thông số của số định tỷ lệ `scaling_win_left_offset`, `scaling_win_right_offset`, `scaling_win_top_offset`, và `scaling_win_bottom_offset` cần được

căn chỉnh ngang qua các lớp phải được bao gồm, để bảo đảm RPR của các ITRP là không cần thiết khi có nhiều ảnh con trên ảnh.

4) Hiện nay, ảnh được sắp xếp đối với ảnh hiện tại có thể là ảnh tham chiếu dài hạn (LTRP) trong cùng lớp như ảnh hiện tại và cũng có thể là ảnh tham chiếu ngoài lớp (ILRP), ví dụ, ảnh tham chiếu trong lớp khác với ảnh hiện tại. Tuy nhiên, trong cả hai trường hợp, việc định tỷ lệ dựa trên POC của các vector chuyển động sẽ không được áp dụng, do đó hiệu suất tạo mã do cho phép điều này được dự kiến là rất thấp. Do đó, sẽ tốt hơn khi không cho phép ảnh được sắp xếp đối với ảnh hiện tại là LTRP hoặc ILRP.

5) Hiện nay, cho phép đối với các ảnh với độ phân giải không gian giống nhau trong CLVS có các cửa sổ định tỷ lệ khác nhau. Tuy nhiên, điều đó phải không được phép, bởi nếu không thì cờ SPS đối với RPR và cờ điều kiện ràng buộc chung đối với RPR sẽ không thể được sử dụng để vô hiệu hóa hoàn toàn các công cụ RPR.

5. Liệt kê các giải pháp kỹ thuật và các phương án

Để giải quyết các vấn đề nêu trên, và khác nữa, các phương pháp như được tổng kết dưới đây được đề xuất. Các điều khoản này phải được coi như các ví dụ để giải thích các ý đồ chung và không được diễn giải theo cách hẹp. Hơn nữa, các điều khoản này có thể được áp dụng một cách riêng rẽ hoặc được kết hợp theo cách bất kỳ.

1) Để giải quyết vấn đề 1a, thay vì chỉ có một cờ SPS (ví dụ, `res_change_in_clvs_allowed_flag` như trong dự thảo VVC hiện tại) để điều khiển RPR, nhiều (ví dụ hai) cờ SPS có thể được chỉ ra và/hoặc được báo tín hiệu cho mục đích này.

a. Ví dụ, cờ thứ nhất (ví dụ, `ref_pic_resampling_enabled_flag`) chỉ ra liệu việc sử dụng RPR có thể là cần thiết để giải mã một hoặc nhiều ảnh, trong khi cờ thứ hai (ví dụ, `res_change_in_clvs_allowed_flag`) chỉ ra việc liệu độ phân giải ảnh có được cho phép để thay đổi bên trong CLVS.

b. Theo cách khác, hơn nữa, cờ thứ hai chỉ được báo tín hiệu khi cờ thứ nhất chỉ ra rằng việc sử dụng RPR có thể là cần thiết để giải mã một hoặc nhiều ảnh. Hơn nữa, khi không được báo tín hiệu, giá trị của cờ thứ hai được suy ra bằng giá trị mà chỉ ra rằng độ phân giải ảnh không được cho phép để thay đổi bên trong CLVS.

i. Theo cách khác, hai cờ này được báo tín hiệu độc lập với nhau.

c. Theo cách khác, hơn nữa, thêm một cờ điều kiện ràng buộc chung được bổ sung, sao cho có một cờ điều kiện ràng buộc chung đối với mỗi trong số các cờ thứ nhất và cờ thứ hai.

d. Hơn nữa, sự kết hợp của nhiều ảnh con trên ảnh với `res_change_in_clvs_allowed_flag` bằng 1 là không được phép, nhưng sự kết hợp của nhiều ảnh con trên ảnh với `ref_pic_resampling_enabled_flag` bằng 1 là được cho phép.

e. Hơn nữa, điều kiện ràng buộc của giá trị của `scaling_window_explicit_signalling_flag` dựa trên giá trị của `res_change_in_clvs_allowed_flag` là được thay đổi để dựa trên giá trị của `ref_pic_resampling_enabled_flag`, như sau: Khi ***ref_pic_resampling_enabled_flag*** `res_change_in_clvs_allowed_flag` là bằng 0, giá trị của `scaling_window_explicit_signalling_flag` sẽ bằng 0.

f. Theo cách khác, hoặc một hoặc tất cả trong số nhiều (ví dụ hai) cờ có thể được báo tín hiệu trong VPS thay vì trong SPS.

i. Trong một ví dụ, hoặc một hoặc tất cả trong số nhiều (ví dụ hai) cờ trong VPS được áp dụng cho tất cả các lớp được chỉ ra bởi VPS.

ii. Trong một ví dụ khác, hoặc một hoặc tất cả trong số nhiều (ví dụ hai) cờ trong VPS có thể mỗi cờ có nhiều phiên bản được báo tín hiệu trong VPS và mỗi phiên bản áp dụng cho tất cả các lớp trong một cây phụ thuộc.

g. Trong một ví dụ, mỗi cờ trong số nhiều cờ được tạo mã dưới dạng số nguyên không dấu bằng cách sử dụng 1 bit $u(1)$.

h. Theo cách khác, một phần tử cú pháp có thể là được báo tín hiệu với giá trị không phải nhị phân, ví dụ, trong SPS/VPS để chỉ ra việc sử dụng RPR trong quá trình giải mã và sự cho phép thay đổi độ phân giải ảnh bên trong CLVS.

i. Trong một ví dụ, khi giá trị của phần tử cú pháp là bằng 0, nó chỉ ra rằng việc sử dụng RPR là không cần thiết để giải mã một hoặc nhiều ảnh.

ii. Trong một ví dụ, khi giá trị của phần tử cú pháp là bằng 1, nó chỉ ra rằng việc sử dụng RPR có thể là cần thiết để giải mã một hoặc nhiều ảnh trong khi độ phân giải ảnh không được phép thay đổi bên trong CLVS.

iii. Trong một ví dụ, khi giá trị của phần tử cú pháp là bằng 2, nó chỉ ra rằng việc sử dụng RPR có thể là cần thiết để giải mã một hoặc nhiều ảnh trong khi độ phân giải ảnh được cho phép để thay đổi bên trong CLVS.

iv. Theo cách khác, hơn nữa, cách thức báo tín hiệu phần tử cú pháp có thể phụ thuộc vào việc liệu dự báo ngoài lớp có được cho phép.

v. Trong một ví dụ, phần tử cú pháp được tạo mã với $ue(v)$, mà chỉ báo phần tử cú pháp được tạo mã Exp-Golomb bậc 0 số nguyên không dấu với bit trái đầu tiên.

vi. Trong một ví dụ khác, phần tử cú pháp được tạo mã dưới dạng số nguyên không dấu bằng cách sử dụng N bit $u(N)$, ví dụ, với N bằng 2.

2) Theo cách khác, hoặc bổ sung cho điều khoản 1 để giải quyết vấn đề 1a, vẫn chỉ có một cờ, ví dụ, `res_change_in_clvs_allowed_flag`, nhưng ngữ nghĩa có thể được thay đổi sao cho không phụ thuộc vào giá trị của cờ này, việc lấy mẫu lại của ảnh tham chiếu ngoài lớp là được cho phép.

a. Trong một ví dụ, ngữ nghĩa có thể được thay đổi như sau: `res_change_in_clvs_allowed_flag` bằng 1 chỉ ra rằng độ phân giải không gian ảnh có thể thay đổi bên trong CLVS mà tham chiếu tới SPS và *khi giải mã ảnh hiện tại trong CLVS, việc lấy mẫu lại của ảnh tham chiếu trong cùng lớp như ảnh hiện tại có thể là cần thiết*. `res_change_in_clvs_allowed_flag` bằng 0 chỉ ra rằng độ phân giải không gian ảnh không thay đổi bên trong CLVS bất kỳ mà tham chiếu tới SPS và *khi giải mã ảnh hiện tại bất kỳ trong CLVS, việc lấy mẫu lại của ảnh tham chiếu trong cùng lớp như ảnh hiện tại là không cần thiết*.

b. Với sự thay đổi này, thì ngay cả khi `res_change_in_clvs_allowed_flag` là bằng 0, việc giải mã ảnh con/ảnh vẫn có thể sử dụng RPR cho các ảnh tham chiếu ngoài lớp (ILRP).

3) Để giải quyết vấn đề 1b, điều kiện ràng buộc đối với sự kết hợp của các ảnh con và khả năng mở rộng với ILP được cập nhật sao cho điều kiện ràng buộc này chỉ áp đặt các hạn chế căn chỉnh ngang lớp lên lớp hiện tại và tất cả các lớp cao hơn mà phụ thuộc vào lớp hiện tại, trong khi không áp đặt lên các lớp cao hơn mà không phụ thuộc vào lớp hiện tại hoặc lên các lớp thấp hơn.

a. Theo cách khác, điều kiện ràng buộc được cập nhật để chỉ áp đặt các hạn chế căn chỉnh ngang lớp lên lớp hiện tại và tất cả các lớp mà là cao hơn lớp hiện tại.

b. Theo cách khác, điều kiện ràng buộc được cập nhật để chỉ áp đặt các hạn chế căn chỉnh ngang lớp lên lớp hiện tại và tất cả các lớp cao hơn trong mỗi OLS chứa lớp hiện tại.

c. Theo cách khác, điều kiện ràng buộc được cập nhật để chỉ áp đặt các hạn chế căn chỉnh ngang lớp lên lớp hiện tại và tất cả các lớp thấp hơn mà là lớp tham chiếu của lớp hiện tại.

d. Theo cách khác, điều kiện ràng buộc được cập nhật để chỉ áp đặt các hạn chế căn chỉnh ngang lớp lên lớp hiện tại và tất cả các lớp mà là thấp hơn lớp hiện tại.

e. Theo cách khác, điều kiện ràng buộc được cập nhật để chỉ áp đặt các hạn chế căn chỉnh ngang lớp lên lớp hiện tại và tất cả các lớp thấp hơn trong mỗi OLS chứa lớp hiện tại.

f. Theo cách khác, điều kiện ràng buộc được cập nhật để chỉ áp đặt các hạn chế căn chỉnh ngang lớp lên tất cả các lớp mà là thấp hơn lớp cao nhất.

g. Theo cách khác, điều kiện ràng buộc được cập nhật để chỉ áp đặt các hạn chế căn chỉnh ngang lớp lên tất cả các lớp mà là cao hơn lớp thấp nhất.

4) Để giải quyết vấn đề 2, trong một hoặc nhiều trong số các quá trình giải mã mà bao hàm các hoạt động xen trong các quá trình liên quan tới dự báo ngoài để xử lý các biên ảnh con trong bù chuyển động/dự báo chuyển động như các biên ảnh (ví dụ, trong các mục 8.5.2.11 Quá trình rút ra để dự báo vector chuyển động độ sáng theo thời gian, 8.5.3.2.2 Quá trình nội suy song tuyến tính mẫu độ sáng, 8.5.5.3 Quá trình rút ra đối với các ứng viên hợp nhất theo thời gian dựa trên khối con, 8.5.5.4 Quá trình rút ra đối với dữ liệu chuyển động cơ sở hợp nhất theo thời gian dựa trên khối con, 8.5.5.6 Quá trình rút ra đối với các ứng viên hợp nhất vector chuyển động điểm điều khiển afin được xây dựng, 8.5.6.3.2 Quá trình lọc nội suy mẫu độ sáng, 8.5.6.3.3 Quá trình tìm nạp mẫu nguyên độ sáng, và 8.5.6.3.4 Quá trình nội suy mẫu sắc độ), các thay đổi sau đây được áp dụng:

a. Trong một ví dụ, các quá trình được thay đổi sao cho các hoạt động xen được áp dụng nếu `subpic_treated_as_pic_flag[CurrSubpicIdx]` là bằng 1 và `sps_num_subpics_minus1` đối với ảnh tham chiếu `refPicLX` là lớn hơn 0, và không được áp dụng nếu khác đi.

i. Theo cách khác, khi ảnh được sắp xếp đối với ảnh là không được phép là ILRP, thì chỉ các quá trình mà trong đó ảnh tham chiếu refPicLX không phải ảnh được sắp xếp là được thay đổi như được mô tả trên đây, và các quá trình mà trong đó ảnh tham chiếu refPicLX là ảnh được sắp xếp thì không được thay đổi.

b. Trong một ví dụ, các quá trình được thay đổi sao cho các hoạt động xén được áp dụng nếu `subpic_treated_as_pic_flag[CurrSubpicIdx]` là bằng 1 và giá trị của `nal_unit_type` của lát hiện tại không bằng `IDR_W_RADL`, `IDR_N_LP`, hoặc `CRA_NUT`, và không được áp dụng nếu ngược lại. Đồng thời, ILP được cho phép chỉ để tạo mã các ảnh IRAP.

c. Trong một ví dụ, không có sự thay đổi nào đối với các quá trình giải mã này được thực hiện, ví dụ, các hoạt động xén được áp dụng nếu `subpic_treated_as_pic_flag[CurrSubpicIdx]` là bằng 1 và không được áp dụng nếu ngược lại, giống như trong văn bản VVC hiện tại.

5) Để giải quyết vấn đề 3a, điều kiện ràng buộc đối với sự kết hợp của các ảnh con và khả năng mở rộng với ILP được cập nhật sao cho điều kiện ràng buộc này áp đặt các hạn chế căn chỉnh ngang lớp lên tất cả các lớp trong mỗi cây phụ thuộc. Cây phụ thuộc chứa lớp cụ thể, tất cả các lớp mà có lớp cụ thể là lớp tham chiếu, và tất cả các lớp tham chiếu của lớp cụ thể), không phụ thuộc vào việc liệu lớp bất kỳ trong số các lớp này có là lớp đầu ra của OLS.

6) Để giải quyết vấn đề 3b, điều kiện ràng buộc đối với sự kết hợp của các ảnh con và khả năng mở rộng với ILP được cập nhật sao cho điều kiện ràng buộc này áp đặt sự hạn chế căn chỉnh ngang lớp lên giá trị của `subpic_treated_as_pic_flag[i]`.

7) Để giải quyết vấn đề 3c, điều kiện ràng buộc đối với sự kết hợp của các ảnh con và khả năng mở rộng với ILP được cập nhật sao cho điều

kiện ràng buộc này không áp đặt sự hạn chế căn chỉnh ngang lớp lên giá trị của `loop_filter_across_subpic_enabled_flag[i]`.

8) Để giải quyết vấn đề 3d, điều kiện ràng buộc đối với sự kết hợp của các ảnh con và khả năng mở rộng với ILP được cập nhật sao cho điều kiện ràng buộc này không được áp dụng khi `sps_num_subpics_minus1` là bằng 0.

a. Theo cách khác, điều kiện ràng buộc được cập nhật sao cho điều kiện ràng buộc này không được áp dụng khi `subpic_info_present_flag` là bằng 0.

9) Để giải quyết vấn đề 3e, điều kiện ràng buộc đối với sự kết hợp của các ảnh con và khả năng mở rộng với ILP được cập nhật sao cho điều kiện ràng buộc này áp đặt các hạn chế căn chỉnh ngang lớp lên các ảnh trong tập đích đã biết của các AU.

a. Trong một ví dụ, đối với mỗi CLVS của lớp hiện tại mà tham chiếu tới SPS, giả sử tập đích của các AU `targetAuSet` là tất cả các AU bắt đầu từ AU chứa ảnh thứ nhất của CLVS theo thứ tự giải mã, tới AU chứa ảnh cuối của CLVS theo thứ tự giải mã, bao gồm cả các giá trị biên.

10) Để giải quyết vấn đề 3f, điều kiện ràng buộc đối với sự kết hợp của các ảnh con và khả năng mở rộng với ILP được cập nhật sao cho điều kiện ràng buộc này áp đặt sự hạn chế căn chỉnh ngang lớp đối với giá trị của mỗi thông số trong số các thông số cửa sổ định tỷ lệ `scaling_win_left_offset`, `scaling_win_right_offset`, `scaling_win_top_offset`, và `scaling_win_bottom_offset`.

11) Để giải quyết vấn đề 4, điều kiện ràng buộc là ảnh được sắp xếp đối với ảnh hiện tại sẽ không phải ảnh tham chiếu dài hạn (LTRP).

a. Theo cách khác, điều kiện ràng buộc là ảnh được sắp xếp đối với ảnh hiện tại sẽ không phải ảnh tham chiếu ngoài lớp (ILRP).

b. Theo cách khác, điều kiện ràng buộc là ảnh được sắp xếp đối với ảnh hiện tại sẽ không phải LTRP hoặc ILRP.

c. Theo cách khác, thay vào đó, nếu ảnh được sắp xếp đối với ảnh hiện tại là LTRP hoặc ILRP, không có sự định tỷ lệ nào được áp dụng để lấy vector chuyển động trở tới ảnh được sắp xếp.

12) Để giải quyết vấn đề 5, điều kiện ràng buộc là giá trị của mỗi trong số `scaling_win_left_offset`, `scaling_win_right_offset`, `scaling_win_top_offset`, và `scaling_win_bottom_offset` sẽ là giống nhau đối với hai ảnh bất kỳ bên trong cùng một CLVS lần lượt có các giá trị của `pic_width_in_luma_samples` và `pic_height_in_luma_samples` giống nhau.

a. Theo cách khác, thay thế "bên trong cùng một CLVS" trên đây bằng "bên trong cùng một CVS".

b. Theo cách khác, điều kiện ràng buộc được chỉ ra như sau:

Giả sử `ppsA` và `ppsB` là hai PPS bất kỳ mà tham chiếu tới cùng một SPS. Yêu cầu về sự tương thích dòng bit là, khi `ppsA` và `ppsB` lần lượt có các giá trị của `pic_width_in_luma_samples` và `pic_height_in_luma_samples` giống nhau, thì `ppsA` và `ppsB` sẽ lần lượt có các giá trị giống nhau của `scaling_win_left_offset`, `scaling_win_right_offset`, `scaling_win_top_offset`, và `scaling_win_bottom_offset`.

c. Theo cách khác, điều kiện ràng buộc được chỉ ra như sau:

Giá trị của mỗi trong số `scaling_win_left_offset`, `scaling_win_right_offset`, `scaling_win_top_offset`, và `scaling_win_bottom_offset` sẽ là giống nhau đối với hai ảnh bất kỳ bên trong cùng một CVS và thỏa mãn tất cả các điều kiện sau đây:

i. Hai ảnh lần lượt có các giá trị của `pic_width_in_luma_samples` và `pic_height_in_luma_samples` giống nhau.

ii. Hai ảnh thuộc về cùng lớp hoặc hai lớp mà một lớp này là lớp tham chiếu của lớp khác.

13) Được đề xuất rằng khi độ phân giải ảnh/cửa sổ định tỷ lệ là khác nhau đối với ảnh hiện tại và ảnh khác trong cùng đơn vị truy nhập, thì ILP chỉ được cho phép khi ảnh hiện tại là ảnh IRAP.

14) Trong bản mô tả này, độ phân giải ảnh có thể dùng để chỉ chiều rộng và/hoặc chiều cao của ảnh, hoặc có thể dùng để chỉ chiều rộng và/hoặc chiều cao và/hoặc vị trí góc trên cùng bên trái của cửa sổ định tỷ lệ/hoặc cửa sổ tương thích của ảnh.

15) Trong bản mô tả này, RPR không được sử dụng có thể có nghĩa là độ phân giải của các ảnh tham chiếu bất kỳ của ảnh hiện tại là giống như của ảnh hiện tại.

6. Các phương án

Dưới đây là một số các phương án ví dụ đối với một số trong số các khía cạnh của sáng chế được tổng kết trên đây trong Phần 5, mà có thể được áp dụng cho đặc tả VVC. Các phần liên quan nhất mà đã được bổ sung hoặc cải biến được gạch chân in **ngiên đậm**, và một số trong số các phần bị xóa được chỉ báo bằng cách sử dụng [[]].

6.1. Phương án thứ nhất

Phương án này là đối với các điều khoản 1, 1.a, 1.b, 1.c, 1.d, 3, 4.a.i, 5, 6, 7, 8, 9, 9.a, 10, 11, và 12b.

7.3.2.3 Cú pháp tập thông số dãy

seq_parameter_set_rbsp() {	Mô tả
...	
<u>ref_pic_resampling_enabled_flag</u>	<u>u(1)</u>
<u>if(ref_pic_resampling_enabled_flag)</u>	
<u>res_change_in_clvs_allowed_flag</u>	u(1)
<u>pic_width_max_in_luma_samples</u>	ue(v)
<u>pic_height_max_in_luma_samples</u>	ue(v)
...	
}	

7.4.3.3 Ngữ nghĩa RBSP tập thông số dãy

...

ref_pic_resampling_enabled_flag bằng 1 chỉ ra rằng một hoặc nhiều lát của các ảnh trong CLVS có thể tham chiếu tới ảnh tham chiếu với độ phân giải không gian khác nhau trong mục nhập hoạt động của danh sách ảnh

tham chiếu. *ref_pic_resampling_enabled* flag bằng 0 chỉ ra rằng không có lát nào của các ảnh trong CLVS tham chiếu tới ảnh tham chiếu với độ phân giải không gian khác nhau trong mục nhập hoạt động của danh sách ảnh tham chiếu.

NOTE 2 – Khi *ref_pic_resampling_enabled* flag là bằng 1, đối với ảnh hiện tại ảnh tham chiếu với độ phân giải không gian khác nhau có thể hoặc thuộc về cùng lớp hoặc lớp khác với lớp chứa ảnh hiện tại.

res_change_in_clvs_allowed_flag bằng 1 chỉ ra rằng độ phân giải không gian ảnh có thể thay đổi bên trong CLVS mà tham chiếu tới SPS. *res_change_in_clvs_allowed_flag* bằng 0 chỉ ra rằng độ phân giải không gian ảnh không thay đổi bên trong CLVS bất kỳ mà tham chiếu tới SPS. Khi không hiện diện, giá trị của *res_change_in_clvs_allowed_flag* được suy ra bằng 0.

...

subpic_treated_as_pic_flag[i] bằng 1 chỉ ra rằng ảnh con thứ i của mỗi ảnh được tạo mã trong CLVS được xử lý như ảnh trong quá trình giải mã mà loại trừ các hoạt động lọc trong vòng. *subpic_treated_as_pic_flag[i]* bằng 0 chỉ ra rằng ảnh con thứ i của mỗi ảnh được tạo mã trong CLVS không được xử lý như ảnh trong quá trình giải mã mà loại trừ các hoạt động lọc trong vòng. Khi không hiện diện, giá trị của *subpic_treated_as_pic_flag[i]* được suy ra bằng *sps_independent_subpics_flag*.

Khi *sps_num_subpics_minus1* là lớn hơn 0 và *subpic_treated_as_pic_flag[i]* là bằng 1, đối với mỗi CLVS của lớp hiện tại mà tham chiếu tới SPS, giả sử *targetAuSet* là tất cả các AU bắt đầu từ AU chứa ảnh thứ nhất của CLVS theo thứ tự giải mã, tới AU chứa ảnh cuối của CLVS theo thứ tự giải mã, bao gồm cả các giá trị biên, yêu cầu về sự tương thích dòng bit là tất cả các điều kiện sau đây là đúng đối với *targetLayerSet* mà bao gồm lớp hiện tại và tất cả các lớp mà có lớp hiện tại làm lớp tham chiếu:

- Tất cả các ảnh của các AU trong *targetAuSet* và các lớp trong *targetLayerSet* sẽ có giá trị giống nhau của *pic_width_in_luma_samples* và giá trị giống nhau của *pic_height_in_luma_samples*.

– Tất cả các ảnh của các AU trong targetAuSet và các lớp trong targetLayerSet sẽ có giá trị giống nhau của mỗi trong số scaling win left offset, scaling win right offset, scaling win top offset, và scaling win bottom offset.

– Tất cả các SPS được tham chiếu bởi các lớp trong targetLayerSet sẽ có giá trị giống nhau của sps_num_subpics_minus1 và sẽ có các giá trị giống nhau lần lượt của subpic_ctu_top_left_x[j], subpic_ctu_top_left_y[j], subpic_width_minus1[j], subpic_height_minus1[j], và subpic treated as pic flag[j] [[loop_filter_across_subpic_enabled_flag[j]],], đối với mỗi giá trị của j nằm trong khoảng từ 0 tới sps_num_subpics_minus1, bao gồm cả các giá trị biên.

– Tất cả các ảnh của các AU trong targetAuSet và các lớp trong targetLayerSet sẽ có giá trị giống nhau của SubpicIdVal[j] đối với mỗi giá trị của j nằm trong khoảng từ 0 tới sps_num_subpics_minus1, bao gồm cả các giá trị biên.

...

7.4.3.4 Ngữ nghĩa RBSP tập thông số ảnh

...

scaling_window_explicit_signalling_flag bằng 1 chỉ ra rằng các thông số dịch vị cửa sổ định tỷ lệ là hiện diện trong PPS. scaling_window_explicit_signalling_flag bằng 0 chỉ ra rằng các thông số dịch vị cửa sổ định tỷ lệ không hiện diện trong PPS. Khi ref pic resampling enabled flag [[res_change_in_clvs_allowed_flag]] là bằng 0, giá trị của scaling_window_explicit_signalling_flag sẽ bằng 0.

scaling_win_left_offset, scaling_win_right_offset, scaling_win_top_offset, và scaling_win_bottom_offset chỉ ra các dịch vị mà được áp dụng cho kích thước ảnh để tính toán hệ số định tỷ lệ. Khi không hiện diện, các giá trị của scaling_win_left_offset, scaling_win_right_offset, scaling_win_top_offset, và scaling_win_bottom_offset được suy ra lần lượt bằng pps_conf_win_left_offset, pps_conf_win_right_offset, pps_conf_win_top_offset, và pps_conf_win_bottom_offset.

Giá trị của $\text{SubWidthC} * (\text{scaling_win_left_offset} + \text{scaling_win_right_offset})$ sẽ nhỏ hơn $\text{pic_width_in_luma_samples}$, và giá trị của $\text{SubHeightC} * (\text{scaling_win_top_offset} + \text{scaling_win_bottom_offset})$ sẽ nhỏ hơn $\text{pic_height_in_luma_samples}$.

Giả sử ppsA và ppsB là hai PPS bất kỳ mà tham chiếu tới cùng một SPS. Yêu cầu về sự tương thích dòng bit là, khi ppsA và ppsB lần lượt có các giá trị của $\text{pic_width_in_luma_samples}$ và $\text{pic_height_in_luma_samples}$ giống nhau, ppsA và ppsB sẽ lần lượt có các giá trị giống nhau của $\text{scaling_win_left_offset}$, $\text{scaling_win_right_offset}$, $\text{scaling_win_top_offset}$, và $\text{scaling_win_bottom_offset}$.

Các biến số PicOutputWidthL và PicOutputHeightL được rút ra như sau:

$$\text{PicOutputWidthL} = \text{pic_width_in_luma_samples} - (78)$$

$$\text{SubWidthC} * (\text{scaling_win_right_offset} + \text{scaling_win_left_offset})$$

$$\text{PicOutputHeightL} = \text{pic_height_in_luma_samples} - (79)$$

$$\text{SubWidthC} * (\text{scaling_win_bottom_offset} + \text{scaling_win_top_offset})$$

Giả sử $\text{refPicOutputWidthL}$ và $\text{refPicOutputHeightL}$ lần lượt là PicOutputWidthL và PicOutputHeightL , của ảnh tham chiếu của ảnh hiện tại mà tham chiếu tới PPS này. Yêu cầu về sự tương thích dòng bit là tất cả các điều kiện sau đây được thỏa mãn:

– $\text{PicOutputWidthL} * 2$ sẽ lớn hơn hoặc bằng $\text{refPicWidthInLumaSamples}$.

– $\text{PicOutputHeightL} * 2$ sẽ lớn hơn hoặc bằng $\text{refPicHeightInLumaSamples}$.

– PicOutputWidthL sẽ nhỏ hơn hoặc bằng $\text{refPicWidthInLumaSamples} * 8$.

– PicOutputHeightL sẽ nhỏ hơn hoặc bằng $\text{refPicHeightInLumaSamples} * 8$.

– $\text{PicOutputWidthL} * \text{pic_width_max_in_luma_samples}$ sẽ lớn hơn hoặc bằng

$\text{refPicOutputWidthL} * (\text{pic_width_in_luma_samples} - \text{Max}(8, \text{MinCbSizeY}))$.

– $\text{PicOutputHeightL} * \text{pic_height_max_in_luma_samples}$ sẽ lớn hơn hoặc bằng $\text{refPicOutputHeightL} * (\text{pic_height_in_luma_samples} - \text{Max}(8, \text{MinCbSizeY}))$.

...

7.3.3.2 Cú pháp thông tin điều kiện ràng buộc chung

general_constraint_info() {	Mô tả
...	
<u><i>no_ref_pic_resampling_constraint_flag</i></u>	<u><i>u(1)</i></u>
<u><i>no_res_change_in_clvs_constraint_flag</i></u>	<u><i>u(1)</i></u>
...	
}	

7.4.4.2 Ngữ nghĩa thông tin điều kiện ràng buộc chung

...

no_ref_pic_resampling_constraint_flag bằng 1 chỉ ra rằng *ref_pic_resampling_enabled_flag* sẽ bằng 0. *no_ref_pic_resampling_constraint_flag* bằng 0 không áp đặt điều kiện ràng buộc này.

no_res_change_in_clvs_constraint_flag bằng 1 chỉ ra rằng *res_change_in_clvs_allowed_flag* sẽ bằng 0. *no_res_change_in_clvs_constraint_flag* bằng 0 không áp đặt điều kiện ràng buộc này.

...

7.4.8.1 Ngữ nghĩa tiêu đề lát chung

...

slice_collocated_from_10_flag bằng 1 chỉ ra rằng ảnh được sắp xếp được sử dụng để dự báo vector chuyển động theo thời gian được rút ra từ danh sách ảnh tham chiếu 0. *slice_collocated_from_10_flag* bằng 0 chỉ ra rằng ảnh được sắp xếp được sử dụng để dự báo vector chuyển động theo thời gian được rút ra từ danh sách ảnh tham chiếu 1.

Khi slice_type là bằng B hoặc P, ph_temporal_mvp_enabled_flag là bằng 1, và slice_collocated_from_10_flag không hiện diện, điều sau đây được áp dụng:

- Nếu rpl_info_in_ph_flag là bằng 1, slice_collocated_from_10_flag được suy ra bằng ph_collocated_from_10_flag.
- Ngược lại (rpl_info_in_ph_flag là bằng 0 và slice_type là bằng P), giá trị của slice_collocated_from_10_flag được suy ra bằng 1.

slice_collocated_ref_idx chỉ ra chỉ số tham chiếu của ảnh được sắp xếp được sử dụng để dự báo vector chuyển động theo thời gian.

Khi slice_type là bằng P hoặc khi slice_type là bằng B và slice_collocated_from_10_flag là bằng 1, slice_collocated_ref_idx tham chiếu tới mục nhập trong danh sách ảnh tham chiếu 0, và giá trị của slice_collocated_ref_idx sẽ nằm trong khoảng từ 0 tới NumRefIdxActive[0] – 1, bao gồm cả các giá trị biên.

Khi slice_type là bằng B và slice_collocated_from_10_flag là bằng 0, slice_collocated_ref_idx tham chiếu tới mục nhập trong danh sách ảnh tham chiếu 1, và giá trị của slice_collocated_ref_idx sẽ nằm trong khoảng từ 0 tới NumRefIdxActive[1] – 1, bao gồm cả các giá trị biên.

Khi slice_collocated_ref_idx không hiện diện, điều sau đây được áp dụng:

- Nếu rpl_info_in_ph_flag là bằng 1, giá trị của slice_collocated_ref_idx được suy ra bằng ph_collocated_ref_idx.
- Ngược lại (rpl_info_in_ph_flag là bằng 0), giá trị của slice_collocated_ref_idx được suy ra bằng 0.

Yêu cầu về sự tương thích dòng bit là ảnh được tham chiếu bởi slice_collocated_ref_idx sẽ là giống nhau đối với tất cả các lát của ảnh được tạo mã và **sẽ không phải ảnh tham chiếu dài hạn.**

Yêu cầu về sự tương thích dòng bit là các giá trị của pic_width_in_luma_samples và pic_height_in_luma_samples của ảnh tham chiếu mà được tham chiếu bởi slice_collocated_ref_idx sẽ lần lượt bằng các giá trị của pic_width_in_luma_samples và pic_height_in_luma_samples, của ảnh hiện tại, và RprConstraintsActive[slice_collocated_from_10_flag ? 0 : 1][slice_collocated_ref_idx] sẽ bằng 0.

...

8.5.3.2.2 Quá trình nội suy song tuyến tính mẫu độ sáng

...

Các vị trí độ sáng trong các đơn vị mẫu toàn phần ($xInt_i$, $yInt_i$) là được rút ra như sau đối với $i = 0..1$:

– Nếu `subpic_treated_as_pic_flag[CurrSubpicIdx]` là bằng 1 và *sps_num_subpics_minus1* đối với ảnh tham chiếu *refPicLX* là lớn hơn 0, điều sau đây được áp dụng:

$$xInt_i = Clip3(SubpicLeftBoundaryPos, SubpicRightBoundaryPos, xInt_L + i) \quad (640)$$

$$yInt_i = Clip3(SubpicTopBoundaryPos, SubpicBotBoundaryPos, yInt_L + i) \quad (641)$$

– Ngược lại (`subpic_treated_as_pic_flag[CurrSubpicIdx]` là bằng 0 hoặc *sps_num_subpics_minus1* đối với ảnh tham chiếu *refPicLX* là bằng 0), điều sau đây được áp dụng:

$$\begin{aligned} xInt_i &= Clip3(0, picW - 1, refWraparoundEnabledFlag ? \\ &ClipH((PpsRefWraparoundOffset) * MinCbSizeY, picW, (xInt_L + i)) : \\ &xInt_L + i) \end{aligned} \quad (642)$$

$$yInt_i = Clip3(0, picH - 1, yInt_L + i) \quad (643)$$

...

8.5.6.3.2 Quá trình lọc nội suy mẫu độ sáng

...

– Nếu `subpic_treated_as_pic_flag[CurrSubpicIdx]` là bằng 1 và *sps_num_subpics_minus1* đối với ảnh tham chiếu *refPicLX* là lớn hơn 0, điều sau đây được áp dụng:

$$xInt_i = Clip3(SubpicLeftBoundaryPos, SubpicRightBoundaryPos, xInt_i) \quad (959)$$

$$yInt_i = Clip3(SubpicTopBoundaryPos, SubpicBotBoundaryPos, yInt_i) \quad (960)$$

– Ngược lại (*subpic_treated_as_pic_flag*[CurrSubpicIdx] là bằng 0 hoặc *sps_num_subpics_minus1* đối với ảnh tham chiếu refPicLX là bằng 0), điều sau đây được áp dụng:

$xInt_i = \text{Clip3}(0, \text{picW} - 1, \text{refWraparoundEnabledFlag} ?$

$\text{ClipH}((\text{PpsRefWraparoundOffset}) * \text{MinCbSizeY}, \text{picW}, xInt_i) : xInt_i)$
(961)

$yInt_i = \text{Clip3}(0, \text{picH} - 1, yInt_i)$ (962)

...

8.5.6.3.3 Quá trình tìm nạp mẫu nguyên độ sáng

...

Các vị trí độ sáng trong các đơn vị mẫu toàn phần ($xInt, yInt$) là được rút ra như sau:

– Nếu *subpic_treated_as_pic_flag*[CurrSubpicIdx] là bằng 1 và *sps_num_subpics_minus1* đối với ảnh tham chiếu refPicLX là lớn hơn 0, điều sau đây được áp dụng:

$xInt = \text{Clip3}(\text{SubpicLeftBoundaryPos}, \text{SubpicRightBoundaryPos}, xInt_L)$ (968)

$yInt = \text{Clip3}(\text{SubpicTopBoundaryPos}, \text{SubpicBotBoundaryPos}, yInt_L)$ (969)

– Ngược lại (*subpic_treated_as_pic_flag*[CurrSubpicIdx] là bằng 0 hoặc *sps_num_subpics_minus1* đối với ảnh tham chiếu refPicLX là bằng 0), điều sau đây được áp dụng:

$xInt = \text{Clip3}(0, \text{picW} - 1, \text{refWraparoundEnabledFlag} ?$ (970)

$\text{ClipH}((\text{PpsRefWraparoundOffset}) * \text{MinCbSizeY}, \text{picW}, xInt_L) : xInt_L)$

$yInt = \text{Clip3}(0, \text{picH} - 1, yInt_L)$ (971)

...

8.5.6.3.4 Quá trình nội suy mẫu sắc độ

...

– Nếu *subpic_treated_as_pic_flag*[CurrSubpicIdx] là bằng 1 và *sps_num_subpics_minus1* đối với ảnh tham chiếu refPicLX là lớn hơn 0, điều sau đây được áp dụng:

$xInt_i = \text{Clip3}(\text{SubpicLeftBoundaryPos} / \text{SubWidthC}, \text{SubpicRightBoundaryPos} / \text{SubWidthC}, xInt_i)$ (977)

$$yInt_i = \text{Clip3}(\text{SubpicTopBoundaryPos} / \text{SubHeightC}, \text{SubpicBotBoundaryPos} / \text{SubHeightC}, yInt_i) \quad (978)$$

– Ngược lại (`subpic_treated_as_pic_flag[CurrSubpicIdx]` là bằng 0 hoặc *sps_num_subpics_minus1* đối với ảnh tham chiếu *refPicLX* là bằng 0), điều sau đây được áp dụng:

$$xInt_i = \text{Clip3}(\quad \quad \quad 0, \quad \quad \quad \text{picWc} - 1, \text{refWraparoundEnabledFlag} ? \text{ClipH}(xOffset, \text{picWc}, xInt_i) : \quad \quad \quad (979)$$

$$xIntc + i - 1)$$

$$yInt_i = \text{Clip3}(0, \text{picHc} - 1, yInt_i) \quad (980)$$

...

Theo cách khác, phần được làm nổi bật "và `sps_num_subpics_minus1` đối với ảnh tham chiếu `refPicLX` là lớn hơn 0" có thể được thay thế bằng "và nếu ảnh tham chiếu `refPicLX` là ILRP có cùng độ phân giải không gian với ảnh hiện tại".

Theo cách khác, phần được làm nổi bật "hoặc `sps_num_subpics_minus1` đối với ảnh tham chiếu `refPicLX` là bằng 0" có thể là được thay thế bằng "hoặc nếu ảnh tham chiếu `refPicLX` là ILRP có độ phân giải không gian khác với ảnh hiện tại".

Theo cách khác, yêu cầu đối với ảnh được sắp xếp, ví dụ, "Yêu cầu về sự tương thích dòng bit là ảnh được tham chiếu bởi `slice_collocated_ref_idx` sẽ là giống nhau đối với tất cả các lát của ảnh được tạo mã và *sẽ không phải ảnh tham chiếu dài hạn.*" có thể được thay thế bằng "Yêu cầu về sự tương thích dòng bit là ảnh được tham chiếu bởi `slice_collocated_ref_idx` sẽ là giống nhau đối với tất cả các lát của ảnh được tạo mã và *sẽ không phải ảnh tham chiếu ngoài lớp.*".

Theo cách khác, yêu cầu đối với ảnh được sắp xếp, ví dụ, "Yêu cầu về sự tương thích dòng bit là ảnh được tham chiếu bởi `slice_collocated_ref_idx` sẽ là giống nhau đối với tất cả các lát của ảnh được tạo mã và *sẽ không phải ảnh tham chiếu dài hạn.*" có thể được thay thế bằng "Yêu cầu về sự tương thích dòng bit là ảnh được tham chiếu bởi `slice_collocated_ref_idx` sẽ là giống nhau đối với tất cả các lát của ảnh được tạo mã và *sẽ không phải ảnh tham chiếu dài hạn hoặc ảnh tham chiếu ngoài lớp.*"

6.2. Các phương án thay thế

Trong một số phương án thay thế, điều kiện ràng buộc sau đây trong phương án thứ nhất :

Khi $sps_num_subpics_minus1$ là lớn hơn 0 và $subpic_treated_as_pic_flag[i]$ là bằng 1, đối với mỗi CLVS của lớp hiện tại mà tham chiếu tới SPS, giả sử $targetAuSet$ là tất cả các AU bắt đầu từ AU chứa ảnh thứ nhất của CLVS theo thứ tự giải mã, tới AU chứa ảnh cuối của CLVS theo thứ tự giải mã, bao gồm cả các giá trị biên, yêu cầu về sự tương thích dòng bit là tất cả các điều kiện sau đây là đúng đối với $targetLayerSet$ mà bao gồm lớp hiện tại và tất cả các lớp mà có lớp hiện tại làm lớp tham chiếu:

- Tất cả các ảnh của các AU trong $targetAuSet$ và các lớp trong $targetLayerSet$ sẽ có giá trị giống nhau của $pic_width_in_luma_samples$ và giá trị giống nhau của $pic_height_in_luma_samples$.
- Tất cả các ảnh của các AU trong $targetAuSet$ và các lớp trong $targetLayerSet$ sẽ có giá trị giống nhau của mỗi trong số $scaling_win_left_offset$, $scaling_win_right_offset$, $scaling_win_top_offset$, và $scaling_win_bottom_offset$.
- Tất cả các SPS được tham chiếu bởi các lớp trong $targetLayerSet$ sẽ có giá trị giống nhau của $sps_num_subpics_minus1$ và sẽ có các giá trị giống nhau lần lượt của $subpic_ctu_top_left_x[j]$, $subpic_ctu_top_left_y[j]$, $subpic_width_minus1[j]$, $subpic_height_minus1[j]$, và $subpic_treated_as_pic_flag[j]$ $[[loop_filter_across_subpic_enabled_flag[j],]]$, đối với mỗi giá trị của j nằm trong khoảng từ 0 tới $sps_num_subpics_minus1$, bao gồm cả các giá trị biên.
- Tất cả các ảnh của các AU trong $targetAuSet$ và các lớp trong $targetLayerSet$ sẽ có giá trị giống nhau của $SubpicIdVal[j]$ đối

với mỗi giá trị của j nằm trong khoảng từ 0 tới $\text{sps_num_subpics_minus1}$, bao gồm cả các giá trị biên.

được thay thế bằng một trong số các điều sau đây:

1) Khi $\text{sps_num_subpics_minus1}$ là lớn hơn 0 và $\text{subpic_treated_as_pic_flag}[i]$ là bằng 1, đối với mỗi CLVS của lớp hiện tại mà tham chiếu tới SPS, giả sử targetAuSet là tất cả các AU bắt đầu từ AU chứa ảnh thứ nhất của CLVS theo thứ tự giải mã, tới AU chứa ảnh cuối của CLVS theo thứ tự giải mã, bao gồm cả các giá trị biên, yêu cầu về sự tương thích dòng bit là tất cả của các điều kiện sau đây là đúng đối với targetLayerSet mà bao gồm lớp hiện tại và tất cả các lớp mà có lớp hiện tại làm lớp tham chiếu:

- Tất cả các ảnh của các AU trong targetAuSet và các lớp trong targetLayerSet sẽ có giá trị giống nhau của $\text{pic_width_in_luma_samples}$ và giá trị giống nhau của $\text{pic_height_in_luma_samples}$.

- Tất cả các ảnh của các AU trong targetAuSet và các lớp trong targetLayerSet sẽ có giá trị giống nhau của mỗi trong số $\text{scaling_win_left_offset}$, $\text{scaling_win_right_offset}$, $\text{scaling_win_top_offset}$, và $\text{scaling_win_bottom_offset}$.

- Tất cả các SPS được tham chiếu bởi các lớp trong targetLayerSet sẽ có giá trị giống nhau của $\text{sps_num_subpics_minus1}$ và sẽ có các giá trị giống nhau lần lượt của $\text{subpic_ctu_top_left_x}[j]$, $\text{subpic_ctu_top_left_y}[j]$, $\text{subpic_width_minus1}[j]$, và $\text{subpic_height_minus1}[j]$, [[and $\text{loop_filter_across_subpic_enabled_flag}[j]$]], đối với mỗi giá trị của j nằm trong khoảng từ 0 tới $\text{sps_num_subpics_minus1}$, bao gồm cả các giá trị biên.

- Tất cả các ảnh của các AU trong targetAuSet và các lớp trong targetLayerSet sẽ có giá trị giống nhau của $\text{SubpicIdVal}[j]$ đối với mỗi giá trị của j nằm trong khoảng từ 0 tới $\text{sps_num_subpics_minus1}$, bao gồm cả các giá trị biên.

2) Khi $sps_num_subpics_minus1$ là lớn hơn 0 và $subpic_treated_as_pic_flag[i]$ là bằng 1, đối với mỗi CLVS của lớp hiện tại mà tham chiếu tới SPS, giả sử $targetAuSet$ là tất cả các AU bắt đầu từ AU chứa ảnh thứ nhất của CLVS theo thứ tự giải mã, tới AU chứa ảnh cuối của CLVS theo thứ tự giải mã, bao gồm cả các giá trị biên, yêu cầu về sự tương thích dòng bit là tất cả các điều kiện sau đây là đúng đối với $targetLayerSet$ mà bao gồm lớp hiện tại và tất cả các lớp mà có lớp hiện tại làm lớp tham chiếu:

- Tất cả các ảnh của các AU trong $targetAuSet$ và các lớp trong $targetLayerSet$ sẽ có giá trị giống nhau của $pic_width_in_luma_samples$ và giá trị giống nhau của $pic_height_in_luma_samples$.

- Tất cả các ảnh của các AU trong $targetAuSet$ và các lớp trong $targetLayerSet$ sẽ có giá trị giống nhau của mỗi trong số $scaling_win_left_offset$, $scaling_win_right_offset$, $scaling_win_top_offset$, và $scaling_win_bottom_offset$.

- Tất cả các SPS được tham chiếu bởi các lớp trong $targetLayerSet$ sẽ có giá trị giống nhau của $sps_num_subpics_minus1$ và sẽ có các giá trị giống nhau lần lượt của $subpic_ctu_top_left_x[j]$, $subpic_ctu_top_left_y[j]$, $subpic_width_minus1[j]$, $subpic_height_minus1[j]$, và $loop_filter_across_subpic_enabled_flag[j]$, đối với mỗi giá trị của j nằm trong khoảng từ 0 tới $sps_num_subpics_minus1$, bao gồm cả các giá trị biên.

- Tất cả các ảnh của các AU trong $targetAuSet$ và các lớp trong $targetLayerSet$ sẽ có giá trị giống nhau của $SubpicIdVal[j]$ đối với mỗi giá trị của j nằm trong khoảng từ 0 tới $sps_num_subpics_minus1$, bao gồm cả các giá trị biên.

3) Khi $subpic_treated_as_pic_flag[i]$ là bằng 1, đối với mỗi CLVS của lớp hiện tại mà tham chiếu tới SPS, giả sử $targetAuSet$ là tất cả các AU bắt đầu từ AU chứa ảnh thứ nhất của CLVS theo thứ tự giải mã,

tới AU chứa ảnh cuối của CLVS theo thứ tự giải mã, bao gồm cả các giá trị biên, yêu cầu về sự tương thích dòng bit là tất cả của các điều kiện sau đây là đúng đối với targetLayerSet mà bao gồm lớp hiện tại và tất cả các lớp mà có lớp hiện tại làm lớp tham chiếu:

- Tất cả các ảnh của các AU trong targetAuSet và các lớp trong targetLayerSet sẽ có giá trị giống nhau của pic_width_in_luma_samples và giá trị giống nhau của pic_height_in_luma_samples.

- Tất cả các ảnh của các AU trong targetAuSet và các lớp trong targetLayerSet sẽ có giá trị giống nhau của mỗi trong số scaling win left offset, scaling win right offset, scaling win top offset, và scaling win bottom offset.

- Tất cả các SPS được tham chiếu bởi các lớp trong targetLayerSet sẽ có giá trị giống nhau của sps_num_subpics_minus1 và sẽ có các giá trị giống nhau lần lượt của subpic_ctu_top_left_x[j], subpic_ctu_top_left_y[j], subpic_width_minus1[j], subpic_height_minus1[j], và subpic treated as pic flag[j]

[[loop_filter_across_subpic_enabled_flag[j],]], đối với mỗi giá trị của j nằm trong khoảng từ 0 tới sps_num_subpics_minus1, bao gồm cả các giá trị biên.

- Tất cả các ảnh của các AU trong targetAuSet và các lớp trong targetLayerSet sẽ có giá trị giống nhau của SubpicIdVal[j] đối với mỗi giá trị của j nằm trong khoảng từ 0 tới sps_num_subpics_minus1, bao gồm cả các giá trị biên.

4) Khi sps num subpics minus1 là lớn hơn 0 và subpic_treated_as_pic_flag[i] là bằng 1, đối với mỗi CLVS của lớp hiện tại mà tham chiếu tới SPS, giả sử targetAuSet là tất cả các AU bắt đầu từ AU chứa ảnh thứ nhất của CLVS theo thứ tự giải mã, tới AU chứa ảnh cuối của CLVS theo thứ tự giải mã, bao gồm cả các giá trị biên, yêu cầu về sự tương thích dòng bit là tất cả của các điều kiện sau đây là

đúng đối với *targetLayerSet* mà bao gồm lớp hiện tại và tất cả các lớp mà có lớp hiện tại làm lớp tham chiếu:

- Tất cả các ảnh của *các AU trong targetAuSet* và các lớp trong *targetLayerSet* sẽ có giá trị giống nhau của *pic_width_in_luma_samples* và giá trị giống nhau của *pic_height_in_luma_samples*.
- Tất cả các ảnh của các AU trong *targetAuSet* và các lớp trong *targetLayerSet* sẽ có giá trị giống nhau của mỗi trong số *scaling win left offset*, *scaling win right offset*, *scaling win top offset*, và *scaling win bottom offset*.
- Tất cả các SPS được tham chiếu bởi *các lớp trong targetLayerSet* sẽ có giá trị giống nhau của *sps_num_subpics_minus1* và sẽ có các giá trị giống nhau lần lượt của *subpic_ctu_top_left_x[j]*, *subpic_ctu_top_left_y[j]*, *subpic_width_minus1[j]*, *subpic_height_minus1[j]*, *subpic treated as pic flag[j]*, và *loop_filter_across_subpic_enabled_flag[j]*, đối với mỗi giá trị của *j* nằm trong khoảng từ 0 tới *sps_num_subpics_minus1*, bao gồm cả các giá trị biên.
- Tất cả các ảnh của *các AU trong targetAuSet* và các lớp trong *targetLayerSet* sẽ có giá trị giống nhau của *SubpicIdVal[j]* đối với mỗi giá trị của *j* nằm trong khoảng từ 0 tới *sps_num_subpics_minus1*, bao gồm cả các giá trị biên.

FIG. 5 là sơ đồ khối thể hiện hệ thống xử lý video ví dụ 1900 mà trong đó các kỹ thuật khác nhau được mô tả ở đây có thể được thực hiện. Các phương án thực hiện khác nhau có thể bao gồm một số hoặc tất cả các thành phần của hệ thống 1900. Hệ thống 1900 có thể bao gồm đầu vào 1902 để nhận nội dung video. Nội dung video có thể được nhận trong định dạng thô hoặc không nén, ví dụ, các giá trị điểm ảnh đa thành phần 8 hoặc 10 bit, hoặc có thể trong định dạng được nén hoặc được mã hóa. Đầu vào 1902 có thể biểu diễn giao diện mạng, giao diện buýt ngoại vi, hoặc giao diện bộ lưu trữ. Các ví dụ về giao diện mạng bao gồm

các giao diện dây dẫn như Ethernet, mạng quang học thụ động (PON – passive optical network), v.v. và các giao diện không dây như Wi-Fi hoặc các giao diện dạng ô.

Hệ thống 1900 có thể bao gồm thành phần tạo mã 1904 mà có thể thực hiện các phương pháp tạo mã hoặc mã hóa khác nhau mà được mô tả trong bản mô tả này. Thành phần tạo mã 1904 có thể giảm tốc độ bit trung bình của video từ đầu vào 1902 tới đầu ra của thành phần tạo mã 1904 để tạo ra sự biểu diễn được tạo mã của video. Do đó, các kỹ thuật tạo mã này đôi khi được gọi là nén video hoặc các kỹ thuật chuyển mã video. Đầu ra của thành phần tạo mã 1904 có thể hoặc được lưu, hoặc được truyền qua đường truyền thông được kết nối, như được biểu diễn bởi thành phần 1906. Sự biểu diễn dòng bit được lưu hoặc được truyền thông (hoặc được tạo mã) của video được nhận ở đầu vào 1902 có thể được sử dụng bởi thành phần 1908 để tạo ra các giá trị điểm ảnh hoặc video có thể hiển thị mà được gửi tới giao diện hiển thị 1910. Quá trình tạo ra video người dùng xem được từ sự biểu diễn dòng bit đôi khi được gọi là giải nén video. Hơn nữa, trong khi những hoạt động xử lý video đã biết được gọi là các hoạt động hoặc các công cụ “tạo mã”, thì cần hiểu rằng các công cụ hoặc các hoạt động tạo mã được sử dụng ở bộ mã hóa và các công cụ hoặc các hoạt động giải mã tương ứng mà đảo ngược các kết quả của việc tạo mã sẽ được thực hiện bởi bộ giải mã.

Các ví dụ về giao diện buýt ngoại vi hoặc giao diện hiển thị có thể bao gồm buýt nối tiếp vạn năng (USB - universal serial bus) hoặc giao diện đa phương tiện độ phân giải cao (HDMI - high definition multimedia interface) hoặc Displayport, và v.v. Các ví dụ về các giao diện bộ lưu trữ bao gồm giao diện SATA (serial advanced technology attachment), PCI, IDE, và tương tự. Các kỹ thuật được mô tả trong bản mô tả có thể được thực hiện trong các thiết bị điện tử khác nhau như các điện thoại di động, các máy tính xách tay, các điện thoại thông minh hoặc các thiết bị khác mà có khả năng thực hiện xử lý dữ liệu kỹ thuật số và/hoặc hiển thị video.

FIG. 6 là sơ đồ khối của thiết bị xử lý video 3600. Thiết bị 3600 có thể được sử dụng để thực hiện một hoặc nhiều trong số các phương pháp được mô tả ở đây. Thiết bị 3600 có thể được thực hiện trong điện thoại thông minh, máy tính

bảng, máy tính, bộ thu internet vạn vật (IoT - Internet of Things), và v.v.. Thiết bị 3600 có thể bao gồm một hoặc nhiều bộ xử lý 3602, một hoặc nhiều bộ nhớ 3604 và phần cứng xử lý video 3606. (Các) bộ xử lý 3602 có thể được tạo cấu hình để thực hiện một hoặc nhiều phương pháp được mô tả trong bản mô tả. Bộ nhớ (các bộ nhớ) 3604 có thể được sử dụng để lưu trữ dữ liệu và mã được sử dụng để thực hiện các phương pháp và các kỹ thuật được mô tả ở đây. Phần cứng xử lý video 3606 có thể được sử dụng để thực hiện, trong hệ mạch phần cứng, một số kỹ thuật được mô tả trong bản mô tả này.

FIG. 8 là sơ đồ khối mà minh họa hệ thống tạo mã video 100 ví dụ mà có thể sử dụng các kỹ thuật theo sáng chế.

Như được thể hiện trên FIG. 8, hệ thống tạo mã video 100 có thể bao gồm thiết bị nguồn 110 và thiết bị đích 120. Thiết bị nguồn 110 tạo ra dữ liệu video được mã hóa mà có thể được gọi là thiết bị mã hóa video. Thiết bị đích 120 có thể giải mã dữ liệu video được mã hóa được tạo ra bởi thiết bị nguồn 110 mà có thể được gọi là thiết bị giải mã video.

Thiết bị nguồn 110 có thể bao gồm nguồn video 112, bộ mã hóa video 114, và giao diện đầu vào/đầu ra (I/O) 116.

Nguồn video 112 có thể bao gồm nguồn như thiết bị quay video, giao diện để nhận dữ liệu video từ nhà cung cấp nội dung video, và/hoặc hệ thống đồ họa máy tính để tạo ra dữ liệu video, hoặc sự kết hợp của các nguồn này. Dữ liệu video có thể bao gồm một hoặc nhiều ảnh. Bộ mã hóa video 114 mã hóa dữ liệu video từ nguồn video 112 để tạo ra dòng bit. Dòng bit có thể bao gồm dãy của các bit mà tạo thành sự biểu diễn được tạo mã của dữ liệu video. Dòng bit có thể bao gồm các ảnh được tạo mã và dữ liệu liên kết. Ảnh được tạo mã là sự biểu diễn được tạo mã của ảnh. Dữ liệu liên kết có thể bao gồm các tập thông số dãy, các tập thông số ảnh, và các cấu trúc cú pháp khác. Giao diện I/O 116 có thể bao gồm bộ điều biến/giải điều biến (môdem) và/hoặc bộ truyền. Dữ liệu video được mã hóa có thể được truyền trực tiếp tới thiết bị đích 120 qua giao diện I/O 116 thông qua mạng 130a. Dữ liệu video được mã hóa cũng có thể được lưu trong vật lưu trữ/máy chủ 130b để truy nhập bởi thiết bị đích 120.

Thiết bị đích 120 có thể bao gồm giao diện I/O 126, bộ giải mã video 124, và thiết bị hiển thị 122.

Giao diện I/O 126 có thể bao gồm bộ thu và/hoặc môđem. Giao diện I/O 126 có thể thu thập dữ liệu video được mã hóa từ thiết bị nguồn 110 hoặc vật lưu trữ/máy chủ 130b. Bộ giải mã video 124 có thể giải mã dữ liệu video được mã hóa. Thiết bị hiển thị 122 có thể hiển thị dữ liệu video được giải mã tới người dùng. Thiết bị hiển thị 122 có thể được tích hợp với thiết bị đích 120, hoặc có thể là ngoại vi đối với thiết bị đích 120 mà được tạo cấu hình để giao diện với thiết bị hiển thị ngoại vi.

Bộ mã hóa video 114 và bộ giải mã video 124 có thể hoạt động theo chuẩn nén video, như chuẩn tạo mã video hiệu quả cao (HEVC - High Efficiency Video Coding), chuẩn tạo mã video đa năng (VVC – Versatile Video Coding) và các chuẩn hiện tại và/hoặc tiếp theo khác.

FIG. 9 là sơ đồ khối minh họa ví dụ về bộ mã hóa video 200, mà có thể là bộ mã hóa video 114 trong hệ thống 100 được minh họa trên FIG. 8.

Bộ mã hóa video 200 có thể được tạo cấu hình để thực hiện kỹ thuật bất kỳ hoặc tất cả các kỹ thuật theo sáng chế. Trong ví dụ trên FIG. 9, bộ mã hóa video 200 bao gồm nhiều thành phần chức năng. Các kỹ thuật được mô tả trong bản mô tả này có thể được chia sẻ giữa các thành phần khác nhau của bộ mã hóa video 200. Trong một số ví dụ, bộ xử lý có thể được tạo cấu hình để thực hiện kỹ thuật bất kỳ hoặc tất cả các kỹ thuật được mô tả trong bản mô tả này.

Các thành phần chức năng của bộ mã hóa video 200 có thể bao gồm khối phân vùng 201, khối xác nhận 202 mà có thể bao gồm khối chọn chế độ 203, khối ước tính chuyển động 204, khối bù chuyển động 205 và khối dự báo trong 206, khối tạo dư 207, khối chuyển đổi 208, khối lượng tử hóa 209, khối lượng tử hóa ngược 210, khối chuyển đổi ngược 211, khối tái tạo 212, bộ đệm 213, và khối mã hóa entropi 214.

Trong các ví dụ khác, bộ mã hóa video 200 có thể bao gồm nhiều hơn, ít hơn, hoặc các thành phần chức năng khác nhau. Trong một ví dụ, khối xác nhận 202 có thể bao gồm khối sao chép khối trong (IBC). Khối IBC có thể thực hiện

việc xác nhận trong chế độ IBC mà trong đó ít nhất một ảnh tham chiếu là ảnh mà khối video hiện tại nằm trong đó.

Hơn nữa, một số thành phần, như khối ước tính chuyển động 204 và khối bù chuyển động 205 có thể được tích hợp ở mức độ cao, nhưng được biểu diễn trong ví dụ trên FIG. 9 một cách riêng biệt cho mục đích giải thích.

Khối phân vùng 201 có thể phân vùng ảnh thành một hoặc nhiều khối video. Bộ mã hóa video 200 và bộ giải mã video 300 có thể hỗ trợ các kích thước khối video khác nhau.

Khối chọn chế độ 203 có thể chọn một trong số các chế độ tạo mã, trong hoặc ngoài, ví dụ, dựa trên các kết quả lỗi, và cấp khối được tạo mã trong hoặc ngoài thu được tới khối tạo dư 207 để tạo ra khối dữ liệu dư và tới khối tái tạo 212 để tái tạo khối được mã hóa để sử dụng làm ảnh tham chiếu. Trong một số ví dụ, khối chọn chế độ 203 có thể chọn chế độ sự kết hợp của xác nhận trong và ngoài (CIIP – combination of intra and inter predication) mà trong đó sự xác nhận là dựa trên tín hiệu xác nhận ngoài và tín hiệu xác nhận trong. Khối chọn chế độ 203 cũng có thể chọn độ phân giải đối với vector chuyển động (ví dụ, độ chính xác điểm ảnh con hoặc điểm ảnh nguyên) đối với khối này trong trường hợp xác nhận ngoài.

Để thực hiện dự báo ngoài trên khối video hiện tại, khối ước tính chuyển động 204 có thể tạo ra thông tin chuyển động đối với khối video hiện tại bằng cách so sánh một hoặc nhiều khung tham chiếu từ bộ đệm 213 với khối video hiện tại. Khối bù chuyển động 205 có thể xác định khối video được dự báo đối với khối video hiện tại dựa trên thông tin chuyển động và các mẫu được giải mã của các ảnh từ bộ đệm 213 khác với ảnh được liên kết với khối video hiện tại.

Khối ước tính chuyển động 204 và khối bù chuyển động 205 có thể thực hiện các hoạt động khác nhau đối với khối video hiện tại, ví dụ, phụ thuộc vào việc liệu khối video hiện tại đang ở trong lát I, lát P, hoặc lát B.

Trong một số ví dụ, khối ước tính chuyển động 204 có thể thực hiện dự báo đơn hướng đối với khối video hiện tại, và khối ước tính chuyển động 204 có thể tìm kiếm các ảnh tham chiếu của danh sách 0 hoặc danh sách 1 đối với khối video tham chiếu đối với khối video hiện tại. Sau đó khối ước tính chuyển động

204 có thể tạo ra chỉ số tham chiếu mà chỉ báo ảnh tham chiếu trong danh sách 0 hoặc danh sách 1 mà chứa khối video tham chiếu và vector chuyển động mà chỉ báo sự dịch chuyển không gian giữa khối video hiện tại và khối video tham chiếu. Khối ước tính chuyển động 204 có thể xuất ra chỉ số tham chiếu, bộ chỉ báo hướng dự báo, và vector chuyển động làm thông tin chuyển động của khối video hiện tại. Khối bù chuyển động 205 có thể tạo ra khối video được dự báo của khối hiện tại dựa trên khối video tham chiếu được chỉ báo bởi thông tin chuyển động của khối video hiện tại.

Trong các ví dụ khác, khối ước tính chuyển động 204 có thể thực hiện dự báo hai hướng đối với khối video hiện tại, khối ước tính chuyển động 204 có thể tìm kiếm các ảnh tham chiếu trong danh sách 0 đối với khối video tham chiếu đối với khối video hiện tại và cũng có thể tìm kiếm các ảnh tham chiếu trong danh sách 1 đối với khối video tham chiếu khác đối với khối video hiện tại. Sau đó khối ước tính chuyển động 204 có thể tạo ra các chỉ số tham chiếu mà chỉ báo các ảnh tham chiếu trong danh sách 0 và danh sách 1 chứa các khối video tham chiếu và các vector chuyển động mà chỉ báo các dịch chuyển không gian giữa các khối video tham chiếu và khối video hiện tại. Khối ước tính chuyển động 204 có thể xuất ra các chỉ số tham chiếu và các vector chuyển động của khối video hiện tại làm thông tin chuyển động của khối video hiện tại. Khối bù chuyển động 205 có thể tạo ra khối video được dự báo của khối video hiện tại dựa trên các khối video tham chiếu được chỉ báo bởi thông tin chuyển động của khối video hiện tại.

Trong một số ví dụ, khối ước tính chuyển động 204 có thể xuất ra tập đầy đủ của thông tin chuyển động cho quá trình xử lý giải mã của bộ giải mã.

Trong một số ví dụ, khối ước tính chuyển động 204 có thể không xuất ra tập đầy đủ của thông tin chuyển động đối với video hiện tại. Thay vào đó, khối ước tính chuyển động 204 có thể báo tín hiệu thông tin chuyển động của khối video hiện tại nhờ tham chiếu tới thông tin chuyển động của khối video khác. Ví dụ, khối ước tính chuyển động 204 có thể xác định rằng thông tin chuyển động của khối video hiện tại là đủ giống với thông tin chuyển động của khối video lân cận.

Trong một ví dụ, khối ước tính chuyển động 204 có thể chỉ báo, trong cấu trúc cú pháp liên kết với khối video hiện tại, giá trị mà chỉ báo tới bộ giải mã video 300 rằng khối video hiện tại có thông tin chuyển động giống như khối video khác.

Trong một ví dụ khác, khối ước tính chuyển động 204 có thể nhận dạng, trong cấu trúc cú pháp liên kết với khối video hiện tại, khối video khác và hiệu số vector chuyển động (MVD). Hiệu số vector chuyển động chỉ báo hiệu số giữa vector chuyển động của khối video hiện tại và vector chuyển động của khối video được chỉ báo. Bộ giải mã video 300 có thể sử dụng vector chuyển động của khối video được chỉ báo và hiệu số vector chuyển động để xác định vector chuyển động của khối video hiện tại.

Như đã đề cập trên đây, bộ mã hóa video 200 có thể báo tín hiệu theo cách dự báo vector chuyển động. Hai ví dụ về các kỹ thuật báo tín hiệu theo cách dự báo mà có thể được thực hiện bởi bộ mã hóa video 200 bao gồm xác nhận vector chuyển động cao cấp (AMVP – advanced motion vector predication) và báo tín hiệu chế độ kết hợp.

Khối dự báo trong 206 có thể thực hiện dự báo trong trên khối video hiện tại. Khi khối dự báo trong 206 thực hiện dự báo trong trên khối video hiện tại, khối dự báo trong 206 có thể tạo ra dữ liệu dự báo đối với khối video hiện tại dựa trên các mẫu được giải mã của các khối video khác trong cùng một ảnh. Dữ liệu dự báo đối với khối video hiện tại có thể bao gồm khối video được dự báo và các phần tử cú pháp khác nhau.

Khối tạo dư 207 có thể tạo ra dữ liệu dư đối với khối video hiện tại bằng cách trừ (ví dụ, được chỉ báo bởi dấu trừ) (các) khối video được dự báo của khối video hiện tại từ khối video hiện tại. Dữ liệu dư của khối video hiện tại có thể bao gồm các khối video dư mà tương ứng với các thành phần mẫu khác nhau của các mẫu trong khối video hiện tại.

Trong các ví dụ khác, có thể không có dữ liệu dư đối với khối video hiện tại đối với khối video hiện tại, ví dụ trong chế độ bỏ qua, và khối tạo dư 207 có thể không thực hiện hoạt động trừ.

Khối xử lý chuyển đổi 208 có thể tạo ra một hoặc nhiều khối video hệ số chuyển đổi đối với khối video hiện tại bằng cách áp dụng một hoặc nhiều sự chuyển đổi vào khối video dư liên kết với khối video hiện tại.

Sau khi khối xử lý chuyển đổi 208 tạo ra khối video hệ số chuyển đổi liên kết với khối video hiện tại, khối lượng tử hóa 209 có thể lượng tử hóa khối video hệ số chuyển đổi liên kết với khối video hiện tại dựa trên một hoặc nhiều giá trị thông số lượng tử hóa (QP) liên kết với khối video hiện tại.

Khối lượng tử hóa ngược 210 và khối chuyển đổi ngược 211 có thể lần lượt áp dụng lượng tử hóa ngược và các chuyển đổi ngược vào khối video hệ số chuyển đổi, để tái tạo khối video dư từ khối video hệ số chuyển đổi. Khối tái tạo 212 có thể bổ sung khối video dư được tái lập vào các mẫu tương ứng từ một hoặc nhiều khối video được dự báo được tạo ra bởi khối xác nhận 202 để tạo ra khối video được tái lập liên kết với khối hiện tại để lưu trữ trong bộ đệm 213.

Sau khi khối tái tạo 212 tái tạo khối video, hoạt động lọc vòng có thể được thực hiện để giảm các thành phần lạ tạo khối video trong khối video.

Khối mã hóa entropi 214 có thể nhận dữ liệu từ các thành phần chức năng khác của bộ mã hóa video 200. Khi khối mã hóa entropi 214 nhận dữ liệu, khối mã hóa entropi 214 có thể thực hiện một hoặc nhiều hoạt động mã hóa entropi để tạo ra dữ liệu được mã hóa entropi và xuất ra dòng bit mà bao gồm dữ liệu được mã hóa entropi.

FIG. 10 là sơ đồ khối minh họa ví dụ của bộ giải mã video 300 mà có thể là bộ giải mã video 114 trong hệ thống 100 được minh họa trên FIG. 8.

Bộ giải mã video 300 có thể được tạo cấu hình để thực hiện kỹ thuật bất kỳ hoặc tất cả các kỹ thuật theo sáng chế. Trong ví dụ trên FIG. 10, bộ giải mã video 300 bao gồm nhiều thành phần chức năng. Các kỹ thuật được mô tả trong bản mô tả này có thể được chia sẻ giữa các thành phần khác nhau của bộ giải mã video 300. Trong một số ví dụ, bộ xử lý có thể được tạo cấu hình để thực hiện kỹ thuật bất kỳ hoặc tất cả các kỹ thuật được mô tả trong bản mô tả này.

Trong ví dụ trên FIG. 10, bộ giải mã video 300 bao gồm khối giải mã entropi 301, khối bù chuyển động 302, khối dự báo trong 303, khối lượng tử hóa ngược 304, khối chuyển đổi ngược 305, và khối tái tạo 306 và bộ đệm 307. Bộ giải mã

video 300 có thể, trong một số ví dụ, thực hiện chuyển tác giải mã mà thường là thuận nghịch với chuyển tác mã hóa được mô tả đối với bộ mã hóa video 200 (FIG. 9).

Khối giải mã entropi 301 có thể truy xuất dòng bit được mã hóa. Dòng bit được mã hóa có thể bao gồm dữ liệu video được tạo mã entropi (ví dụ, các khối được mã hóa của dữ liệu video). Khối giải mã entropi 301 có thể giải mã dữ liệu video được tạo mã entropi, và từ dữ liệu video được giải mã entropi, khối bù chuyển động 302 có thể xác định thông tin chuyển động bao gồm các vector chuyển động, độ chính xác vector chuyển động, các chỉ số danh sách ảnh tham chiếu, và thông tin chuyển động khác. Khối bù chuyển động 302 có thể, ví dụ, xác định thông tin này bằng cách thực hiện AMVP và chế độ kết hợp.

Khối bù chuyển động 302 có thể tạo ra các khối bù chuyển động, có thể bằng cách thực hiện nội suy dựa trên các bộ lọc nội suy. Các bộ nhận dạng đối với các bộ lọc nội suy cần được sử dụng với độ chính xác điểm ảnh con có thể được bao gồm trong các phần tử cú pháp.

Khối bù chuyển động 302 có thể sử dụng các bộ lọc nội suy như được sử dụng bởi bộ mã hóa video 200 trong khi mã hóa của khối video để tính toán các giá trị nội suy đối với các điểm ảnh con số nguyên của khối tham chiếu. Khối bù chuyển động 302 có thể xác định các bộ lọc nội suy được sử dụng bởi bộ mã hóa video 200 theo thông tin cú pháp nhận được và sử dụng các bộ lọc nội suy để tạo ra các khối dự báo.

Khối bù chuyển động 302 có thể sử dụng một số thông tin cú pháp để xác định các kích thước của các khối được sử dụng để mã hóa (các) khung và/hoặc (các) lát của dãy video được mã hóa, thông tin phân vùng mà mô tả cách thức mỗi khối macro của ảnh của dãy video được mã hóa được phân vùng, các chế độ mà chỉ báo cách thức mỗi phân vùng được mã hóa, một hoặc nhiều khung tham chiếu (và các danh sách khung tham chiếu) đối với mỗi khối được mã hóa ngoài, và thông tin khác để giải mã dãy video được mã hóa.

Khối dự báo trong 303 có thể sử dụng các chế độ dự báo trong ví dụ được nhận trong dòng bit để tạo thành khối dự báo từ các khối liền kề trong không gian. Khối lượng tử hóa ngược 303 lượng tử hóa ngược, tức là, giải lượng tử hóa,

các hệ số khối video được lượng tử hóa được tạo ra trong dòng bit và được giải mã bởi khối giải mã entropi 301. Khối chuyển đổi ngược 303 áp dụng chuyển đổi ngược.

Khối tái tạo 306 có thể lấy tổng các khối dư với các khối dự báo tương ứng được tạo ra bởi khối bù chuyển động 202 hoặc khối dự báo trong 303 để tạo thành các khối được giải mã. Nếu mong muốn, bộ lọc giải khối cũng có thể được áp dụng để lọc các khối được giải mã nhằm loại bỏ các thành phần lạ tạo khối. Sau đó các khối video được giải mã được lưu trong bộ đệm 307, mà cấp các khối tham chiếu để bù chuyển động/xác nhận trong tiếp sau và còn tạo ra video được giải mã để trình diễn trên thiết bị hiển thị.

Việc liệt kê các giải pháp ưu tiên bởi một số phương án được thực hiện tiếp theo đây.

Các giải pháp sau đây thể hiện các phương án ví dụ của các kỹ thuật được đề cập trong phần trước (ví dụ, điều khoản 1).

1. Phương pháp xử lý video (ví dụ, phương pháp 700 được thể hiện trên FIG. 7), bao gồm các bước: thực hiện (702) sự biến đổi giữa video bao gồm một hoặc nhiều ảnh video, trong đó sự biểu diễn được tạo mã tương thích với quy tắc định dạng; trong đó quy tắc định dạng chỉ ra rằng hai hoặc nhiều trường cú pháp trong tập thông số dãy điều khiển sự thay đổi độ phân giải ảnh tham chiếu (RPR) trong video.

2. Phương pháp theo giải pháp 1, trong đó trường cú pháp thứ nhất trong số hai hoặc nhiều trường cú pháp chỉ báo việc liệu RPR có được sử dụng đối với một hoặc nhiều ảnh và trường cú pháp thứ hai trong số hai hoặc nhiều trường cú pháp chỉ báo việc liệu độ phân giải ảnh có được cho phép để thay đổi trong sự biểu diễn được tạo mã ở bậc dãy.

Các giải pháp sau đây thể hiện các phương án ví dụ của các kỹ thuật được đề cập trong phần trước (ví dụ, điều khoản 2).

3. Phương pháp xử lý video, phương pháp này bao gồm các bước: thực hiện sự biến đổi giữa video bao gồm một hoặc nhiều ảnh video, trong đó sự biểu diễn được tạo mã tương thích với quy tắc định dạng; trong đó quy tắc định dạng chỉ ra rằng một trường cú pháp trong tập thông số dãy điều khiển sự thay đổi độ phân

giải ảnh tham chiếu (RPR) trong video; và trong đó, quy tắc định dạng chỉ ra rằng, không phụ thuộc vào giá trị của một trường cú pháp này, việc lấy mẫu lại của ảnh tham chiếu ngoài lớp là được phép để biến đổi.

Các giải pháp sau đây thể hiện các phương án ví dụ của các kỹ thuật được đề cập trong phần trước (ví dụ, các điều khoản 3, 5, 6, 7, 9, 10).

4. Phương pháp xử lý video, phương pháp này bao gồm các bước: thực hiện sự biến đổi giữa video bao gồm một hoặc nhiều lớp bao gồm một hoặc nhiều ảnh video bao gồm một hoặc nhiều ảnh con, trong đó sự biểu diễn được tạo mã tương thích với quy tắc định dạng; trong đó quy tắc định dạng chỉ ra điều kiện ràng buộc thứ nhất đối với căn chỉnh ngang lớp hoặc điều kiện ràng buộc thứ hai đối với sự kết hợp của các ảnh con và khả năng mở rộng của các ảnh ngoài lớp.

5. Phương pháp theo giải pháp 4, trong đó điều kiện ràng buộc thứ nhất xác định sự hạn chế căn chỉnh ngang lớp đối với lớp hiện tại và tất cả các lớp cao hơn mà phụ thuộc vào lớp hiện tại mà không áp đặt sự hạn chế căn chỉnh trên các lớp thấp hơn của lớp hiện tại và tất cả các lớp cao hơn mà không phụ thuộc vào lớp hiện tại.

6. Phương pháp theo giải pháp 4, trong đó điều kiện ràng buộc thứ hai áp đặt sự hạn chế căn chỉnh ngang lớp trên tất cả các lớp trong mỗi cây phụ thuộc của lớp cụ thể.

7. Phương pháp theo giải pháp 4, trong đó các điều kiện ràng buộc thứ hai giới hạn giá trị của `subpic_treated_as_pic_flag[i]` theo sự hạn chế căn chỉnh ngang lớp.

8. Phương pháp theo giải pháp 4, trong đó các điều kiện ràng buộc thứ hai giới hạn giá trị của `loop_filter_across_subpic_enabled_flag[i]` theo sự hạn chế căn chỉnh ngang lớp.

9. Phương pháp theo giải pháp bất kỳ trong số các giải pháp từ 4 tới 8, trong đó điều kiện ràng buộc thứ nhất và/hoặc điều kiện ràng buộc thứ hai được chỉ ra đối với tập đích của các đơn vị truy nhập.

10. Phương pháp theo giải pháp 4, trong đó các điều kiện ràng buộc thứ hai giới hạn giá trị của mỗi thông số trong số các thông số của sơ đồ định tỷ lệ

`scaling_win_left_offset`, `scaling_win_right_offset`, `scaling_win_top_offset`, và `scaling_win_bottom_offset` theo sự hạn chế căn chỉnh ngang lớp.

Các giải pháp sau đây thể hiện các phương án ví dụ của các kỹ thuật được đề cập trong phần trước (ví dụ, điều khoản 11).

11. Phương pháp xử lý video, phương pháp này bao gồm các bước: thực hiện sự biến đổi giữa video bao gồm một hoặc nhiều lớp bao gồm một hoặc nhiều ảnh video bao gồm một hoặc nhiều ảnh con, trong đó sự biến đổi tương thích với quy tắc định dạng mà chỉ ra rằng ảnh tham chiếu ngoài lớp hoặc ảnh tham chiếu dài hạn là không được phép làm ảnh được sắp xếp của ảnh hiện tại để biến đổi.

Các giải pháp sau đây thể hiện các phương án ví dụ của các kỹ thuật được đề cập trong phần trước (ví dụ, điều khoản 12).

12. Phương pháp xử lý video, phương pháp này bao gồm các bước: thực hiện sự biến đổi giữa video bao gồm nhiều ảnh và sự biểu diễn được tạo mã của video, trong đó sự biến đổi tương thích với quy tắc mà chỉ ra rằng các giá trị của mỗi trong số `scaling_win_left_offset`, `scaling_win_right_offset`, `scaling_win_top_offset`, và `scaling_win_bottom_offset` là giống nhau đối với hai ảnh bất kỳ bên trong cùng dãy video lớp được tạo mã hoặc dãy video được tạo mã có các giá trị của `pic_width_in_luma_samples` và `pic_height_in_luma_samples` giống nhau.

Các giải pháp sau đây thể hiện các phương án ví dụ của các kỹ thuật được đề cập trong phần trước (ví dụ, điều khoản 13).

13. Phương pháp xử lý video, phương pháp này bao gồm các bước: thực hiện sự biến đổi giữa video bao gồm nhiều ảnh và sự biểu diễn được tạo mã của video, trong đó sự biến đổi tương thích với quy tắc mà chỉ ra rằng trong trường hợp mà độ phân giải ảnh hoặc cửa sổ định tỷ lệ là khác nhau đối với ảnh hiện tại và ảnh khác trong cùng đơn vị truy nhập, thì việc dự báo ngoài lớp chỉ được cho phép khi ảnh hiện tại là ảnh điểm truy nhập ngẫu nhiên trong.

14. Phương pháp theo giải pháp bất kỳ trong số các giải pháp từ 1 tới 13, trong đó sự biến đổi bao gồm việc mã hóa video thành sự biểu diễn được tạo mã.

15. Phương pháp theo giải pháp bất kỳ trong số các giải pháp từ 1 tới 13, trong đó sự biến đổi bao gồm việc giải mã sự biểu diễn được tạo mã để tạo ra các giá trị điểm ảnh của video.

16. Thiết bị giải mã video bao gồm bộ xử lý được tạo cấu hình để thực hiện phương pháp được nêu trong một hoặc nhiều trong số các giải pháp từ 1 tới 15.

17. Thiết bị mã hóa video bao gồm bộ xử lý được tạo cấu hình để thực hiện phương pháp được nêu trong một hoặc nhiều trong số các giải pháp từ 1 tới 15.

18. Sản phẩm chương trình máy tính có mã máy tính được lưu trên đó, mã này, khi được thực thi bởi bộ xử lý, khiến bộ xử lý thực hiện phương pháp được nêu trong giải pháp bất kỳ trong số các giải pháp từ 1 tới 15.

19. Phương pháp, thiết bị hoặc hệ thống được mô tả trong bản mô tả này.

FIG. 13 là sự biểu diễn dưới dạng lưu đồ của phương pháp 1300 để xử lý video theo công nghệ theo sáng chế. Phương pháp 1300 bao gồm, ở hoạt động 1310, bước thực hiện sự biến đổi giữa ảnh hiện tại của video và dòng bit của video theo quy tắc. Quy tắc này chỉ ra rằng nhiều phần tử cú pháp được sử dụng để chỉ ra việc sử dụng công cụ lấy mẫu lại ảnh tham chiếu. Công cụ lấy mẫu lại ảnh tham chiếu là công cụ tạo mã mà trong đó ảnh tham chiếu mà có độ phân giải khác với ảnh hiện tại được lấy mẫu lại để biến đổi.

Trong một số phương án, nhiều phần tử cú pháp được bao gồm trong tập thông số dãy (SPS). Trong một số phương án, nhiều phần tử cú pháp bao gồm phần tử cú pháp thứ nhất mà chỉ ra việc liệu công cụ lấy mẫu lại ảnh tham chiếu có được phép để biến đổi, và phần tử cú pháp thứ hai mà chỉ ra việc liệu độ phân giải của ảnh có được cho phép để thay đổi bên trong dãy video lớp được tạo mã (CLVS). Trong một số phương án, phần tử cú pháp thứ hai được bao gồm trong SPS trong trường hợp phần tử cú pháp thứ nhất chỉ ra rằng công cụ lấy mẫu lại ảnh tham chiếu là được phép để biến đổi. Trong một số phương án, trong trường hợp phần tử cú pháp thứ hai không được bao gồm trong SPS, phần tử cú pháp thứ hai được suy ra là tương đương với giá trị mà chỉ báo rằng độ phân giải của ảnh là không được phép để thay đổi bên trong CLVS. Trong một số phương án, phần tử cú pháp thứ nhất và phần tử cú pháp thứ hai được chỉ báo một cách độc lập với nhau để biến đổi.

Trong một số phương án, một hoặc nhiều cờ điều kiện ràng buộc chung tương ứng với mỗi trong số nhiều phần tử cú pháp được chỉ báo để biến đổi. Trong một số phương án, quy tắc chỉ ra rằng độ phân giải của ảnh là không được phép để thay đổi bên trong dãy video lớp được tạo mã (CLVS) trong trường hợp ảnh hiện tại bao gồm nhiều ảnh con. Trong một số phương án, quy tắc chỉ ra rằng công cụ lấy mẫu lại ảnh tham chiếu là được phép để biến đổi trong trường hợp ảnh hiện tại bao gồm nhiều ảnh con. Trong một số phương án, cờ điều kiện ràng buộc mà chỉ ra việc liệu thông số dịch vị của sổ định tỷ lệ có hiện diện trong tập thông số ảnh được xác định dựa trên phần tử cú pháp thứ nhất. Trong một số phương án, trong trường hợp phần tử cú pháp thứ nhất chỉ báo rằng công cụ lấy mẫu lại ảnh tham chiếu là không được phép để biến đổi, cờ điều kiện ràng buộc chỉ ra rằng thông số dịch vị của sổ định tỷ lệ được bỏ qua trong tập thông số ảnh.

Trong một số phương án, nhiều phần tử cú pháp được bao gồm trong tập thông số video. Trong một số phương án, ít nhất một trong số nhiều phần tử cú pháp có thể áp dụng cho tất cả các lớp được chỉ ra bởi tập thông số video. Trong một số phương án, ít nhất một trong số nhiều phần tử cú pháp có nhiều phiên bản được chỉ báo trong tập thông số video, và mỗi trong số nhiều phiên bản có thể áp dụng cho tất cả các lớp trong cây phụ thuộc. Trong một số phương án, mỗi trong số nhiều phần tử cú pháp được tạo mã bằng cách sử dụng phương pháp tạo mã đơn phân với một bit.

FIG. 14 là sự biểu diễn dưới dạng lưu đồ của phương pháp 1400 để xử lý video theo công nghệ theo sáng chế. Phương pháp 1400 bao gồm, ở hoạt động 1410, bước thực hiện sự biến đổi giữa ảnh hiện tại của video và dòng bit của video theo quy tắc. Quy tắc này chỉ ra rằng phần tử cú pháp mà có giá trị không phải nhị phân được sử dụng để chỉ ra việc sử dụng của (1) công cụ lấy mẫu lại ảnh tham chiếu mà trong đó ảnh tham chiếu mà có độ phân giải khác với ảnh hiện tại được lấy mẫu lại và (2) sự thay đổi của độ phân giải ảnh bên trong dãy video lớp được tạo mã (CLVS).

Trong một số phương án, giá trị của phần tử cú pháp bằng 0 chỉ báo rằng công cụ lấy mẫu lại ảnh tham chiếu là không được phép để biến đổi. Trong một số phương án, giá trị của phần tử cú pháp bằng 1 chỉ báo rằng công cụ lấy mẫu

lại ảnh tham chiếu là được phép để biến đổi và sự thay đổi của độ phân giải ảnh là không được phép bên trong dãy video lớp được tạo mã (CLVS). Trong một số phương án, giá trị của phần tử cú pháp bằng 2 chỉ báo rằng công cụ lấy mẫu lại ảnh tham chiếu là được phép để biến đổi và sự thay đổi của độ phân giải ảnh là được cho phép bên trong dãy video lớp được tạo mã (CLVS). Trong một số phương án, sự chỉ báo của phần tử cú pháp là dựa trên việc liệu công cụ tạo mã dự báo ngoài lớp có được cho phép để biến đổi. Trong một số phương án, phần tử cú pháp là phần tử cú pháp được tạo mã Exp-Golobm bậc 0 số nguyên không dấu. Trong một số phương án, phần tử cú pháp được tạo mã dưới dạng số nguyên không dấu bằng cách sử dụng N bit, N là số nguyên. Trong một số phương án, N là bằng 2.

FIG. 15 là sự biểu diễn dưới dạng lưu đồ của phương pháp 1500 để xử lý video theo công nghệ theo sáng chế. Phương pháp 1500 bao gồm, ở hoạt động 1510, bước thực hiện sự biến đổi giữa video bao gồm nhiều lớp và dòng bit của video theo quy tắc. Quy tắc này chỉ ra rằng, trong trường hợp ảnh con được xử lý như ảnh video để biến đổi, sự hạn chế căn chỉnh ngang lớp được áp dụng cho ít hơn tất cả nhiều lớp này bao gồm lớp hiện tại mà bao gồm ảnh con và tập con của các lớp được liên kết với lớp hiện tại. Sự hạn chế căn chỉnh ngang lớp bao gồm sự hạn chế của ít nhất kích thước của ảnh video, số lượng của các ảnh con bên trong ảnh video, vị trí của ít nhất một ảnh con, hoặc nhận dạng của ảnh con.

Trong một số phương án, tập con của các lớp bao gồm tất cả các lớp cao hơn mà phụ thuộc vào lớp hiện tại. Trong một số phương án, tập con của các lớp loại trừ tất cả các lớp thấp hơn của lớp hiện tại và các lớp cao hơn mà không phụ thuộc vào lớp hiện tại. Trong một số phương án, tập con của các lớp bao gồm tất cả các lớp mà là cao hơn lớp hiện tại. Trong một số phương án, tập con của các lớp bao gồm tất cả các lớp cao hơn mà là trong cùng tập lớp đầu ra như lớp hiện tại. Trong một số phương án, tập con của các lớp bao gồm tất cả các lớp thấp hơn của lớp hiện tại. Trong một số phương án, tập con của các lớp bao gồm tất cả các lớp mà là thấp hơn lớp hiện tại. Trong một số phương án, tập con của các lớp bao gồm tất cả các lớp thấp hơn mà là trong cùng tập lớp đầu ra như lớp hiện tại. Trong một số phương án, tập con của các lớp bao

gồm tất cả các lớp mà là thấp hơn lớp cao nhất. Trong một số phương án, tập con của các lớp bao gồm tất cả các lớp mà là cao hơn lớp thấp nhất.

FIG. 16 là sự biểu diễn dưới dạng lưu đồ của phương pháp 1600 để xử lý video theo công nghệ theo sáng chế. Phương pháp 1600 bao gồm, ở hoạt động 1610, bước thực hiện sự biến đổi giữa lớp hiện tại của video và dòng bit của video theo quy tắc. Quy tắc này chỉ ra rằng sự hạn chế căn chỉnh ngang lớp được áp dụng cho tất cả các lớp trong cây phụ thuộc được liên kết với lớp hiện tại không phụ thuộc vào việc liệu lớp bất kỳ trong số tất cả các lớp có là lớp đầu ra trong tập lớp đầu ra. Sự hạn chế căn chỉnh ngang lớp bao gồm sự hạn chế của ít nhất kích thước của ảnh video, số lượng của các ảnh con bên trong ảnh video, vị trí của ít nhất một ảnh con, hoặc nhận dạng của ảnh con. Tất cả các lớp trong cây phụ thuộc bao gồm lớp hiện tại, tất cả các lớp mà có lớp hiện tại làm lớp tham chiếu, và tất cả các lớp tham chiếu của lớp hiện tại.

Trong một số phương án, sự hạn chế căn chỉnh ngang lớp còn bao gồm sự hạn chế liên quan đến việc liệu ảnh con trong ảnh video có được xử lý dưới dạng ảnh ngang qua tất cả các lớp. Trong một số phương án, sự hạn chế căn chỉnh ngang lớp không giới hạn việc liệu hoạt động lọc vòng có được áp dụng trên các biên của các ảnh con bên trong ảnh video ngang qua tất cả các lớp. Trong một số phương án, quy tắc còn chỉ ra rằng sự hạn chế căn chỉnh ngang lớp không được áp dụng trong trường hợp phân tử cú pháp trong tập thông số dãy chỉ báo rằng ảnh video bao gồm một ảnh con. Trong một số phương án, quy tắc còn chỉ ra rằng sự hạn chế căn chỉnh ngang lớp không được áp dụng trong trường hợp phân tử cú pháp chỉ báo rằng thông tin ảnh con được bỏ qua trong tập thông số dãy. Trong một số phương án, quy tắc còn chỉ ra rằng sự hạn chế căn chỉnh ngang lớp được áp dụng cho các ảnh trong tập đích của các đơn vị truy nhập. Trong một số phương án, đối với mỗi CLVS của lớp hiện tại mà tham chiếu tới tập thông số dãy, tập đích của các đơn vị truy nhập bao gồm tất cả các đơn vị truy nhập bắt đầu từ đơn vị truy nhập thứ nhất mà bao gồm ảnh thứ nhất của CLVS tới đơn vị truy nhập thứ hai mà bao gồm ảnh cuối của CLVS theo thứ tự giải mã.

Trong một số phương án, sự hạn chế căn chỉnh ngang lớp còn bao gồm thông số dịch vị của sổ định tỷ lệ, thông số dịch vị của sổ định tỷ lệ bao gồm ít nhất

một trong số: (1) dịch vị trái của cửa sổ định tỷ lệ, (2) dịch vị phải của cửa sổ định tỷ lệ, (3) dịch vị đỉnh của cửa sổ định tỷ lệ, hoặc (4) dịch vị đáy của cửa sổ định tỷ lệ.

FIG. 17 là sự biểu diễn dưới dạng lưu đồ của phương pháp 1700 để xử lý video theo công nghệ theo sáng chế. Phương pháp 1700 bao gồm, ở hoạt động 1710, bước thực hiện sự biến đổi giữa ảnh hiện tại của video và dòng bit của video theo quy tắc. Quy tắc này chỉ ra rằng việc lấy mẫu lại của ảnh tham chiếu mà ở trong cùng lớp như ảnh hiện tại là được phép không phụ thuộc vào giá trị của phần tử cú pháp mà chỉ ra việc liệu sự thay đổi của độ phân giải ảnh có được cho phép trong dãy video lớp được tạo mã (CLVS).

Trong một số phương án, trong trường hợp phần tử cú pháp chỉ báo rằng sự thay đổi của độ phân giải ảnh là không được phép trong CLVS, công cụ lấy mẫu lại ảnh tham chiếu mà trong đó ảnh tham chiếu mà có độ phân giải ảnh khác với ảnh hiện tại được lấy mẫu lại là được phép đối với ảnh tham chiếu.

FIG. 18 là sự biểu diễn dưới dạng lưu đồ của phương pháp 1800 để xử lý video theo công nghệ theo sáng chế. Phương pháp 1800 bao gồm, ở hoạt động 1810, bước thực hiện sự biến đổi giữa ảnh hiện tại của video bao gồm nhiều lớp và dòng bit của video theo quy tắc. Quy tắc này chỉ ra một trong số: (1) ảnh tham chiếu của ảnh hiện tại là không được phép để được sắp xếp, hoặc (2) trong trường hợp mà ảnh tham chiếu của ảnh hiện tại được sắp xếp, các vectơ chuyển động mà trỏ tới ảnh tham chiếu được sử dụng trong khi biến đổi của ảnh hiện tại mà không định tỷ lệ. Trong một số phương án, ảnh tham chiếu là ảnh tham chiếu dài hạn. Trong một số phương án, ảnh tham chiếu là ảnh tham chiếu ngoài lớp.

FIG. 19 là sự biểu diễn dưới dạng lưu đồ của phương pháp 1900 để xử lý video theo công nghệ theo sáng chế. Phương pháp 1900 bao gồm, ở hoạt động 1910, bước thực hiện sự biến đổi giữa video và dòng bit của video theo quy tắc. Quy tắc này chỉ ra rằng thông số dịch vị của cửa sổ định tỷ lệ là giống nhau đối với hai ảnh video bất kỳ trong cùng dãy video lớp được tạo mã (CLVS) hoặc dãy video được tạo mã (CVS) mà có kích thước giống nhau được biểu diễn trong số lượng các mẫu độ sáng.

Trong một số phương án, thông số dịch vị của cửa sổ định tỷ lệ bao gồm ít nhất một trong số: (1) dịch vị trái của cửa sổ định tỷ lệ, (2) dịch vị phải của cửa sổ định tỷ lệ, (3) dịch vị đỉnh của cửa sổ định tỷ lệ, hoặc (4) dịch vị đáy của cửa sổ định tỷ lệ. Trong một số phương án, kích thước bao gồm chiều rộng hoặc chiều cao của ảnh. Trong một số phương án, quy tắc còn chỉ ra rằng hai ảnh video bất kỳ thuộc về cùng lớp. Trong một số phương án, quy tắc còn chỉ ra rằng hai ảnh video bất kỳ thuộc về hai lớp bao gồm lớp thứ nhất và lớp thứ hai. lớp thứ nhất là lớp tham chiếu của lớp thứ hai.

FIG. 20 là sự biểu diễn dưới dạng lưu đồ của phương pháp 2000 để xử lý video theo công nghệ theo sáng chế. Phương pháp 2000 bao gồm, ở hoạt động 2010, bước thực hiện sự biến đổi giữa ảnh hiện tại của video và dòng bit của video theo quy tắc. Quy tắc này chỉ ra rằng, đáp lại độ phân giải ảnh của ảnh hiện tại là khác với ít nhất một ảnh khác trong cùng đơn vị truy nhập của ảnh hiện tại, công cụ tạo mã dự báo ngoài lớp là được phép chỉ trong trường hợp mà ảnh hiện tại là ảnh điểm truy nhập ngẫu nhiên trong (IRAP) mà đối với nó tất cả các đơn vị lớp trừu tượng mạng (NAL) lớp tạo mã video (VCL) có giá trị giống nhau đối với kiểu đơn vị NAL.

Trong một số phương án, sự thay đổi của độ phân giải ảnh của ảnh tham chiếu mà không được phép chỉ báo rằng độ phân giải ảnh của ảnh tham chiếu là giống như độ phân giải ảnh của ảnh hiện tại. Trong một số phương án, độ phân giải ảnh của ảnh video bao gồm chiều rộng hoặc chiều cao của ảnh video. Trong một số phương án, độ phân giải ảnh của ảnh video bao gồm chiều rộng, chiều cao, và/hoặc vị trí trên cùng bên trái của cửa sổ định tỷ lệ hoặc cửa sổ tương thích của ảnh video.

Trong một số phương án, sự biến đổi bao gồm việc mã hóa video thành dòng bit. Trong một số phương án, sự biến đổi bao gồm việc giải mã dòng bit để tạo ra video.

Trong các giải pháp được mô tả ở đây, bộ mã hóa có thể tương thích với quy tắc định dạng bằng cách tạo ra sự biểu diễn được tạo mã theo quy tắc định dạng. Trong các giải pháp được mô tả ở đây, bộ giải mã có thể sử dụng quy tắc định dạng để phân tách các phần tử cú pháp trong sự biểu diễn được tạo mã trong khi

biết về sự hiện diện và không hiện diện của các phần tử cú pháp theo quy tắc định dạng để tạo video được giải mã.

Trong bản mô tả, thuật ngữ “xử lý video” có thể được dùng để chỉ mã hóa video, giải mã video, nén video hoặc giải nén video. Ví dụ, các thuật toán nén video có thể được áp dụng trong khi biến đổi từ biểu diễn điểm ảnh của video sang sự biểu diễn dòng bit tương ứng hoặc ngược lại. Sự biểu diễn dòng bit của khối video hiện tại có thể, ví dụ, tương ứng với các bit mà hoặc cùng nằm ở hoặc trải dài trong các vị trí khác nhau bên trong dòng bit, như được xác định bởi cú pháp. Ví dụ, khối macro có thể được mã hóa theo các giá trị dư lỗi được chuyển đổi và được tạo mã và còn bằng cách sử dụng các bit trong các tiêu đề và các trường khác trong dòng bit. Hơn nữa, trong khi biến đổi, bộ giải mã có thể phân tích dòng bit khi biết rằng một số trường có thể hiện diện, hoặc không hiện diện, dựa trên sự xác định, như được mô tả trong các giải pháp trên đây. Tương tự, bộ mã hóa có thể xác định rằng các trường cú pháp đã biết là có hoặc không được bao gồm và tạo ra sự biểu diễn được tạo mã một cách tương ứng bằng cách bao gồm hoặc loại trừ các trường cú pháp từ sự biểu diễn được tạo mã.

Các giải pháp được mô tả và các giải pháp khác, các ví dụ, các phương án, các mô đun và các hoạt động chức năng được mô tả trong bản mô tả này có thể được thực hiện trong hệ mạch điện tử kỹ thuật số, hoặc trong phần mềm, phần sụn, hoặc phần cứng máy tính, bao gồm các cấu trúc được mô tả trong bản mô tả này và các phương án tương đương về mặt cấu trúc của chúng, hoặc trong các sự kết hợp của một hoặc nhiều trong số chúng. Các phương án được mô tả và khác nữa có thể được thực hiện dưới dạng một hoặc nhiều sản phẩm chương trình máy tính, tức là, một hoặc nhiều mô đun của các lệnh chương trình máy tính được mã hóa trong môi trường máy tính đọc được để thực thi bởi, hoặc để điều khiển hoạt động của, thiết bị xử lý dữ liệu. Môi trường máy tính đọc được có thể là thiết bị lưu trữ máy đọc được, nền lưu trữ máy đọc được, thiết bị nhớ, sản phẩm thực hiện tín hiệu được lan truyền máy đọc được, hoặc sự kết hợp của một hoặc nhiều đối tượng này. Thuật ngữ “thiết bị xử lý dữ liệu” bao hàm tất cả thiết bị, các bộ phận, và các máy để xử lý dữ liệu, bao gồm ví dụ bộ xử lý lập trình được, máy tính, hoặc nhiều bộ xử lý hoặc máy tính. Thiết bị có thể bao gồm,

ngoài phần cứng, mã mà tạo ra môi trường thực thi đối với chương trình máy tính đang được xem xét, ví dụ, mã mà cấu thành phần sụn bộ xử lý, ngăn xếp giao thức, hệ thống quản lý cơ sở dữ liệu, hệ điều hành, hoặc sự kết hợp của một hoặc nhiều trong số chúng. Tín hiệu được lan truyền là tín hiệu được tạo ra một cách nhân tạo, ví dụ, tín hiệu điện, quang học, hoặc điện từ được tạo ra bằng máy, mà được tạo ra để mã hóa thông tin để truyền tới thiết bị thu thích hợp.

Chương trình máy tính (còn được biết như chương trình, phần mềm, ứng dụng phần mềm, tập lệnh, hoặc mã) có thể được viết dưới dạng bất kỳ của ngôn ngữ lập trình, bao gồm các ngôn ngữ biên dịch hoặc diễn dịch, và có thể được triển khai dưới dạng bất kỳ, bao gồm dưới dạng chương trình độc lập hoặc dưới dạng môđun, thành phần, thủ tục con, hoặc đơn vị khác thích hợp để sử dụng trong môi trường máy tính. Chương trình máy tính không nhất thiết phải tương ứng với tập tin trong hệ thống tập tin. Chương trình có thể được lưu trong phần của tập tin mà giữ các chương trình khác hoặc dữ liệu (ví dụ, một hoặc nhiều tập lệnh được lưu trong tài liệu ngôn ngữ đánh dấu), trong một tập tin được dành riêng cho chương trình đang được xem xét, hoặc trong nhiều tập tin được kết hợp (ví dụ, các tập tin mà lưu trữ một hoặc nhiều môđun, các chương trình con, hoặc các phần của mã). Chương trình máy tính có thể được triển khai để được thực thi trên một máy tính hoặc trên nhiều máy tính mà nằm ở một địa điểm hoặc được phân bố trên nhiều địa điểm và được liên kết bởi mạng truyền thông.

Các quá trình và các luồng logic được mô tả trong bản mô tả này có thể được thực hiện bởi một hoặc nhiều bộ xử lý lập trình được thực thi một hoặc nhiều chương trình máy tính để thực hiện các chức năng bằng cách điều hành trên dữ liệu đầu vào và tạo ra đầu ra. Các quá trình và các luồng logic này cũng có thể được thực hiện bởi, và thiết bị cũng có thể được thực hiện dưới dạng, hệ mạch logic chuyên dụng, ví dụ, mảng cổng trường lập trình được FPGA (field programmable gate array) hoặc mạch tích hợp chuyên dụng ASIC (application specific integrated circuit).

Các bộ xử lý thích hợp để thực thi chương trình máy tính bao gồm, ví dụ, cả bộ vi xử lý vạn năng lẫn bộ vi xử lý chuyên dụng, và một hoặc nhiều bộ xử lý bất kỳ của loại máy tính kỹ thuật số bất kỳ. Thông thường, bộ xử lý sẽ nhận các

lệnh và dữ liệu từ bộ nhớ chỉ đọc hoặc bộ nhớ truy cập ngẫu nhiên hoặc từ cả hai. Các phần tử cơ bản của máy tính là bộ xử lý để thực hiện các lệnh và một hoặc nhiều thiết bị nhớ để lưu trữ các lệnh và dữ liệu. Thông thường, máy tính sẽ còn bao gồm, hoặc được kết nối theo cách hoạt động được để nhận dữ liệu từ hoặc chuyển dữ liệu tới, hoặc cả hai, một hoặc nhiều thiết bị lưu trữ dung lượng lớn để lưu trữ dữ liệu, ví dụ, các đĩa từ, quang từ, hoặc các đĩa quang học. Tuy nhiên, máy tính không cần phải có các thiết bị này. Môi trường máy tính đọc được thích hợp để lưu trữ các lệnh chương trình máy tính và dữ liệu bao gồm tất cả các dạng của bộ nhớ bất biến, phương tiện và các thiết bị nhớ, bao gồm ví dụ các thiết bị nhớ bán dẫn, ví dụ, EPROM, EEPROM, và các thiết bị nhớ nhanh; các đĩa từ, ví dụ, các ổ cứng trong hoặc ổ cứng lưu động; các đĩa quang từ; và các đĩa CD ROM và DVD-ROM. Bộ xử lý và bộ nhớ có thể được bổ sung bởi, hoặc được tích hợp trong, hệ mạch logic chuyên dụng.

Trong khi bản mô tả này chứa nhiều dấu hiệu cụ thể, chúng không được coi là các dấu hiệu giới hạn phạm vi của đối tượng bất kỳ hoặc của những gì có thể được yêu cầu bảo hộ, mà là các phần mô tả của các dấu hiệu mà có thể là cụ thể đối với các phương án cụ thể của các kỹ thuật cụ thể. Các dấu hiệu đã biết mà được mô tả trong bản mô tả này trong ngữ cảnh của các phương án riêng biệt cũng có thể được thực hiện trong sự kết hợp trong một phương án. Ngược lại, các dấu hiệu khác nhau mà được mô tả trong ngữ cảnh của một phương án cũng có thể được thực hiện trong nhiều phương án một cách riêng biệt hoặc trong sự kết hợp phụ thích hợp bất kỳ. Hơn nữa, mặc dù các dấu hiệu có thể được mô tả trên đây là hoạt động trong các sự kết hợp đã biết và thậm chí ban đầu được yêu cầu bảo hộ như vậy, một hoặc nhiều dấu hiệu từ sự kết hợp được yêu cầu bảo hộ có thể trong một số trường hợp được loại bỏ khỏi sự kết hợp, và sự kết hợp được yêu cầu bảo hộ có thể hướng đến sự kết hợp nhỏ hoặc biến thể của sự kết hợp nhỏ.

Tương tự, trong khi các hoạt động được mô tả trên các hình vẽ theo thứ tự cụ thể, điều này không được hiểu là yêu cầu các hoạt động này được thực hiện theo thứ tự cụ thể được thể hiện hoặc theo thứ tự liên tiếp, hoặc rằng tất cả các hoạt động được minh họa được thực hiện, để đạt được các kết quả mong muốn.

Hơn nữa, sự tách biệt của các thành phần hệ thống khác nhau trong các phương án được mô tả trong bản mô tả này không được hiểu là yêu cầu rằng sự tách biệt này trong tất cả các phương án.

Chỉ một vài phương án thực hiện và các phương án ví dụ được mô tả và các phương án thực hiện, nâng cao và các biến thể khác nữa có thể được thực hiện dựa trên những gì được mô tả và được minh họa trong bản mô tả này.

YÊU CẦU BẢO HỘ

1. Phương pháp xử lý video, phương pháp này bao gồm các bước:

thực hiện sự biến đổi giữa video bao gồm nhiều lớp và dòng bit của video theo quy tắc,

trong đó quy tắc chỉ ra rằng, trong trường hợp rằng phần tử cú pháp thứ nhất được bao gồm trong tập hợp tham số chuỗi chỉ báo số lượng ảnh con trong ảnh video lớn hơn 1, và rằng ảnh con với chỉ số ảnh con thứ nhất được xử lý như một ảnh video để biến đổi, thì sự hạn chế căn chỉnh ngang lớp được áp dụng cho lớp hiện tại mà bao gồm ảnh con và tập con của các lớp được liên kết với lớp hiện tại,

trong đó sự hạn chế căn chỉnh ngang lớp bao gồm sự hạn chế của ảnh hiện tại bao gồm ảnh con và ảnh thứ nhất trong tập hợp con của các lớp có cùng giá trị đối với ít nhất một trong các phần tử cú pháp sau:

giá trị của phần tử cú pháp thứ nhất:

giá trị của phần tử cú pháp thứ hai mà chỉ ra kích thước của ảnh video,

giá trị của phần tử cú pháp thứ ba mà chỉ báo kích thước của ảnh con thứ i,

giá trị của phần tử cú pháp thứ tư mà chỉ báo vị trí của ảnh con thứ i, hoặc

giá trị của chỉ số ảnh phụ thứ nhất.

2. Phương pháp theo điểm 1, trong đó tập con của các lớp được liên kết với lớp hiện tại bao gồm một hoặc nhiều lớp cao hơn mà phụ thuộc vào lớp hiện tại.

3. Phương pháp theo điểm 1 hoặc 2, trong đó tập con của các lớp được liên kết với lớp hiện tại loại trừ tất cả các lớp cao hơn mà không phụ thuộc vào lớp hiện tại.

4. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 tới 3, trong đó tập con của các lớp được liên kết với lớp hiện tại loại trừ tất cả các lớp thấp hơn của lớp hiện tại.

5. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 tới 4, trong đó tập con của các lớp được liên kết với lớp hiện tại là tập con của cây phụ thuộc được liên kết với lớp hiện tại, và

trong đó cây phụ thuộc được liên kết với lớp hiện tại bao gồm lớp hiện tại, tất cả các lớp mà có lớp hiện tại là lớp tham chiếu, và tất cả các lớp tham chiếu của lớp hiện tại.

6. Phương pháp theo điểm 5, trong đó tập con của các lớp được liên kết với lớp hiện tại là tập con của cây phụ thuộc không phụ thuộc vào việc liệu tập con bất kỳ trong số tập con của cây phụ thuộc có là lớp đầu ra trong tập lớp đầu ra.

7. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 tới 6, trong đó sự hạn chế căn chỉnh ngang lớp còn bao gồm sự hạn chế liên quan tới giá trị của phần tử cú pháp thứ năm, và trong đó lớp hiện tại và tập con của các lớp được liên kết với lớp hiện tại có cùng giá trị của phần tử cú pháp thứ năm,

trong đó phần tử cú pháp thứ năm chỉ ra việc liệu ảnh con của mỗi ảnh được tạo mã trong dãy video lớp được tạo mã có được xử lý như ảnh trong quá trình giải mã mà loại trừ các hoạt động lọc trong vòng.

8. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 tới 7, trong đó sự hạn chế căn chỉnh ngang lớp loại trừ giá trị của phần tử cú pháp thứ sáu, và

trong đó phần tử cú pháp thứ sáu chỉ ra việc liệu hoạt động lọc trong vòng ngang qua các đường biên ảnh con có được cho phép.

9. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 tới 8, trong đó quy tắc còn chỉ ra rằng sự hạn chế căn chỉnh ngang lớp không được áp dụng trong trường hợp phần tử cú pháp thứ nhất chỉ báo rằng ảnh video bao gồm một ảnh con.

10. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 tới 9, trong đó quy tắc còn chỉ ra rằng sự hạn chế căn chỉnh ngang lớp không được áp dụng trong trường hợp phần tử cú pháp thứ bảy được bao gồm trong tập thông số dãy chỉ báo rằng thông tin ảnh con là không hiện diện.

11. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 tới 10, trong đó quy tắc còn chỉ ra rằng sự hạn chế căn chỉnh ngang lớp được áp dụng cho các ảnh trong tập đích của các đơn vị truy nhập.

12. Phương pháp theo điểm 11, trong đó, đối với mỗi chuỗi video lớp được tạo mã (CLVS) của lớp hiện tại mà tham chiếu tới tập thông số dãy, tập đích của các đơn vị truy nhập bao gồm tất cả các đơn vị truy nhập bắt đầu từ đơn vị truy nhập

thứ nhất mà bao gồm ảnh thứ nhất của CLVS tới đơn vị truy nhập thứ hai mà bao gồm ảnh cuối cùng của CLVS theo thứ tự giải mã.

13. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 tới 12, trong đó sự biến đổi bao gồm mã hóa video thành dòng bit.

14. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 tới 12, trong đó sự biến đổi bao gồm giải mã dòng bit để tạo ra video.

15. Thiết bị xử lý dữ liệu video bao gồm bộ xử lý và bộ nhớ bất biến với các lệnh trên đó, trong đó các lệnh này khi được thực thi bởi bộ xử lý, khiến bộ xử lý:

thực hiện sự biến đổi giữa video bao gồm nhiều lớp và dòng bit của video theo quy tắc,

trong đó quy tắc chỉ ra rằng, trong trường hợp rằng phần tử cú pháp thứ nhất được bao gồm trong tập hợp tham số chuỗi chỉ báo số lượng ảnh con trong ảnh video lớn hơn 1, và rằng ảnh con với chỉ số ảnh con thứ nhất được xử lý như một ảnh video để biến đổi, thì sự hạn chế căn chỉnh ngang lớp được áp dụng cho lớp hiện tại mà bao gồm ảnh con và tập con của các lớp được liên kết với lớp hiện tại,

trong đó sự hạn chế căn chỉnh ngang lớp bao gồm sự hạn chế của of ảnh hiện tại bao gồm ảnh con và ảnh thứ nhất trong tập hợp con của các lớp có cùng giá trị đối với ít nhất một trong các phần tử cú pháp sau:

giá trị của phần tử cú pháp thứ nhất,

giá trị của phần tử cú pháp thứ hai mà chỉ ra kích thước của ảnh video,

giá trị của phần tử cú pháp thứ ba mà chỉ báo kích thước của ảnh con thứ i,

giá trị của phần tử cú pháp thứ tư mà chỉ báo vị trí của ảnh con thứ i, hoặc

giá trị của chỉ số ảnh con thứ nhất.

16. Thiết bị theo điểm 15, trong đó tập con của các lớp được liên kết với lớp hiện tại bao gồm một hoặc nhiều lớp cao hơn mà phụ thuộc vào lớp hiện tại.

17. Thiết bị theo điểm 15 hoặc 16, trong đó tập con của các lớp được liên kết với lớp hiện tại loại trừ tất cả các lớp cao hơn mà không phụ thuộc vào lớp hiện tại.

18. Thiết bị theo điểm bất kỳ trong số các điểm từ 15 tới 17, trong đó tập con của các lớp được liên kết với lớp hiện tại loại trừ tất cả các lớp thấp hơn của lớp hiện tại.

19. Vật lưu trữ máy tính đọc được bất biến để lưu trữ các lệnh mà khiến bộ xử lý:

thực hiện sự biến đổi giữa video bao gồm nhiều lớp và dòng bit của video theo quy tắc,

trong đó quy tắc chỉ ra rằng, trong trường hợp rằng phần tử cú pháp thứ nhất được bao gồm trong tập hợp tham số chuỗi chỉ báo số lượng ảnh con trong ảnh video lớn hơn 1, và rằng ảnh con với chỉ số ảnh con thứ nhất được xử lý như một ảnh video để biến đổi, thì sự hạn chế căn chỉnh ngang lớp được áp dụng cho lớp hiện tại mà bao gồm ảnh con và tập con của các lớp được liên kết với lớp hiện tại,

trong đó sự hạn chế căn chỉnh ngang lớp bao gồm sự hạn chế của ảnh hiện tại bao gồm ảnh con và ảnh thứ nhất trong tập hợp con của các lớp có cùng giá trị đối với ít nhất một trong các phần tử cú pháp sau:

giá trị của phần tử cú pháp thứ nhất,

giá trị của phần tử cú pháp thứ hai mà chỉ ra kích thước của ảnh video,

giá trị của phần tử cú pháp thứ ba mà chỉ báo kích thước của ảnh con thứ i ,

giá trị của phần tử cú pháp thứ tư mà chỉ báo vị trí của ảnh con thứ i , hoặc

giá trị của chỉ số ảnh con thứ nhất.

20. Phương pháp theo điểm 1, trong đó ảnh thứ nhất là ảnh tham chiếu của ảnh hiện tại, hoặc ảnh hiện tại là ảnh tham chiếu của ảnh thứ nhất.

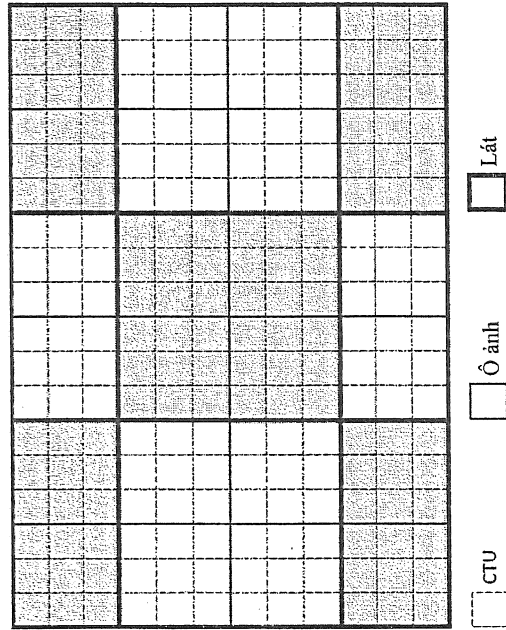


FIG. 1

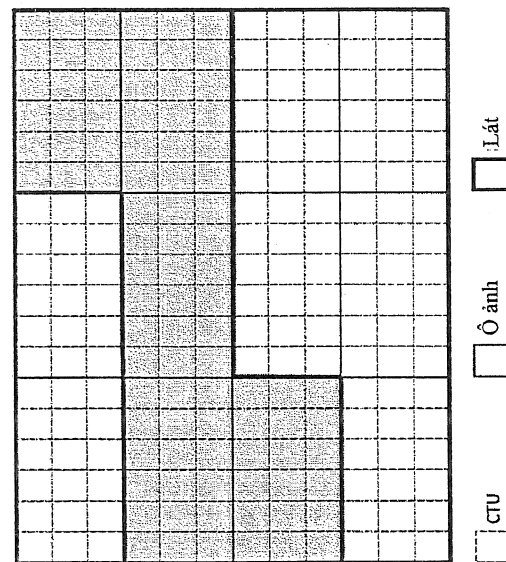


FIG. 2

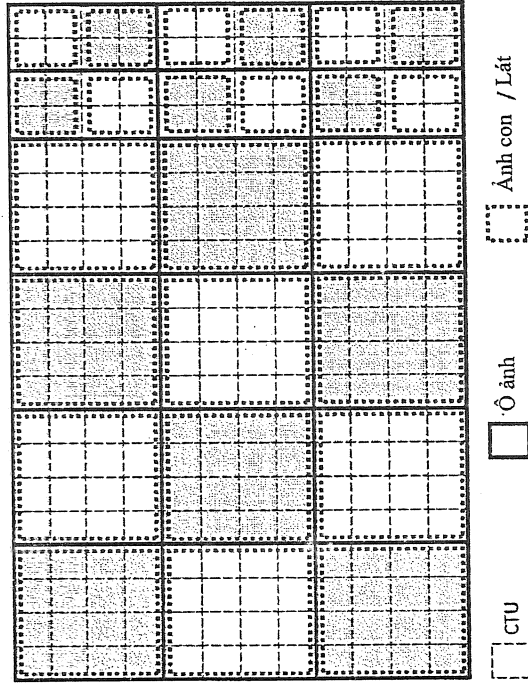


FIG. 3

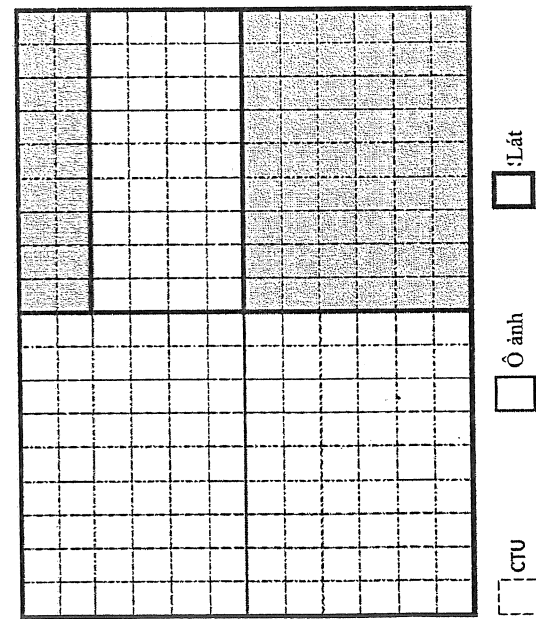
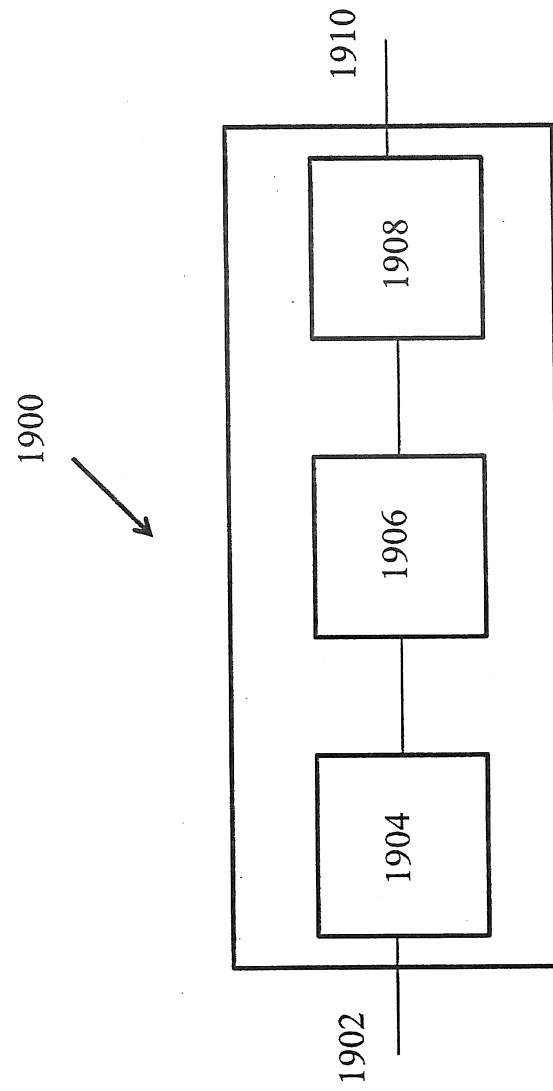


FIG. 4

**FIG. 5**

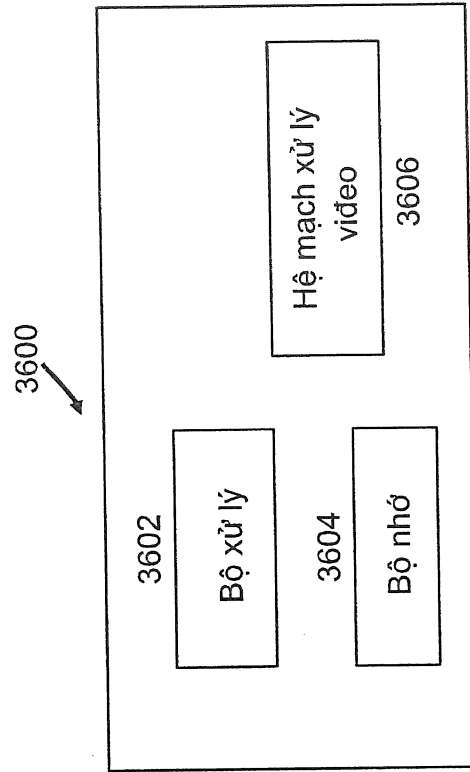
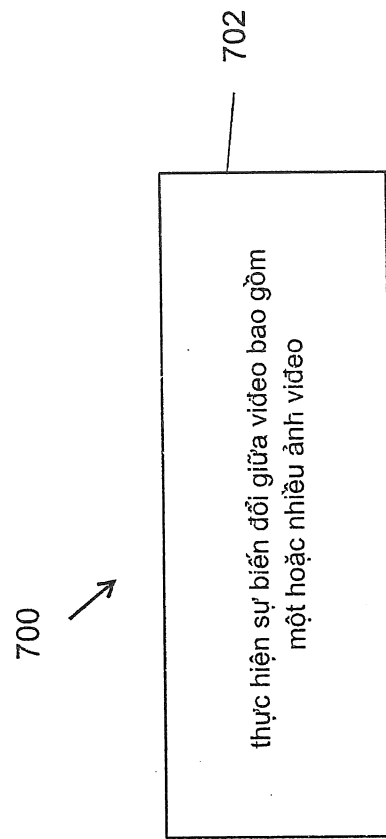


FIG. 6

**FIG. 7**

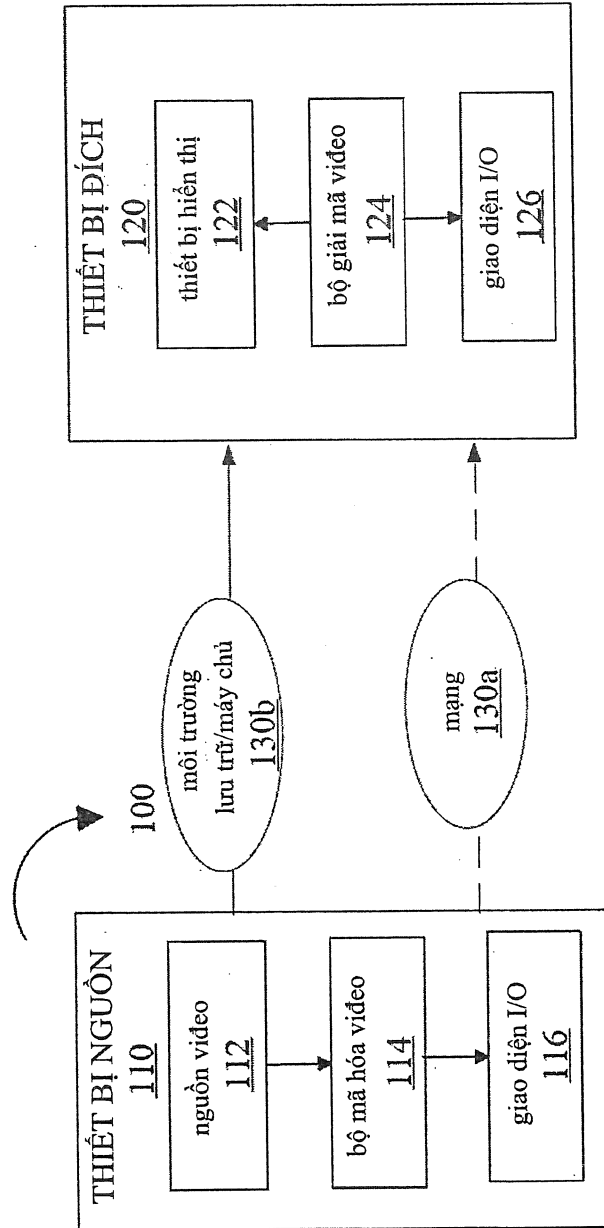


FIG. 8

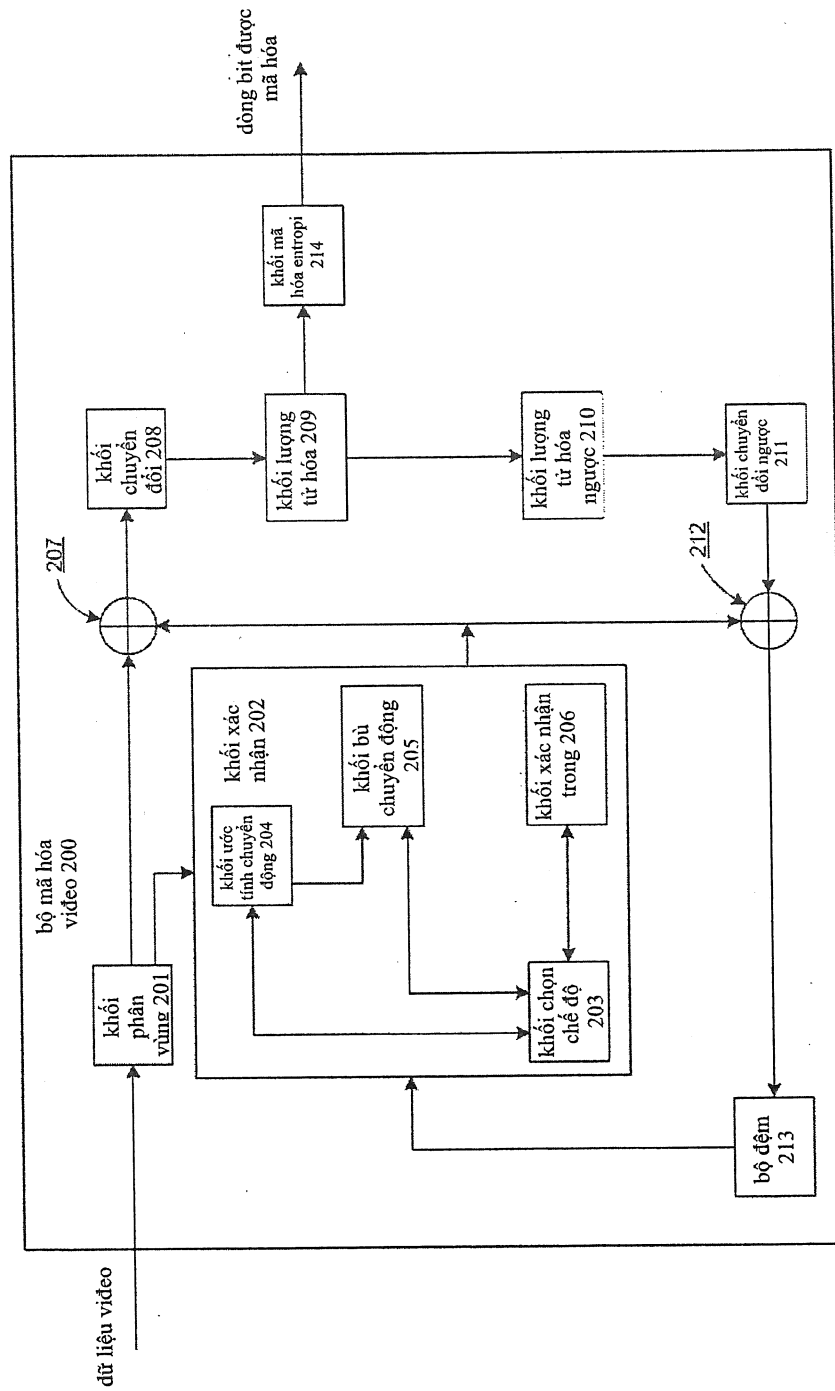


FIG. 9

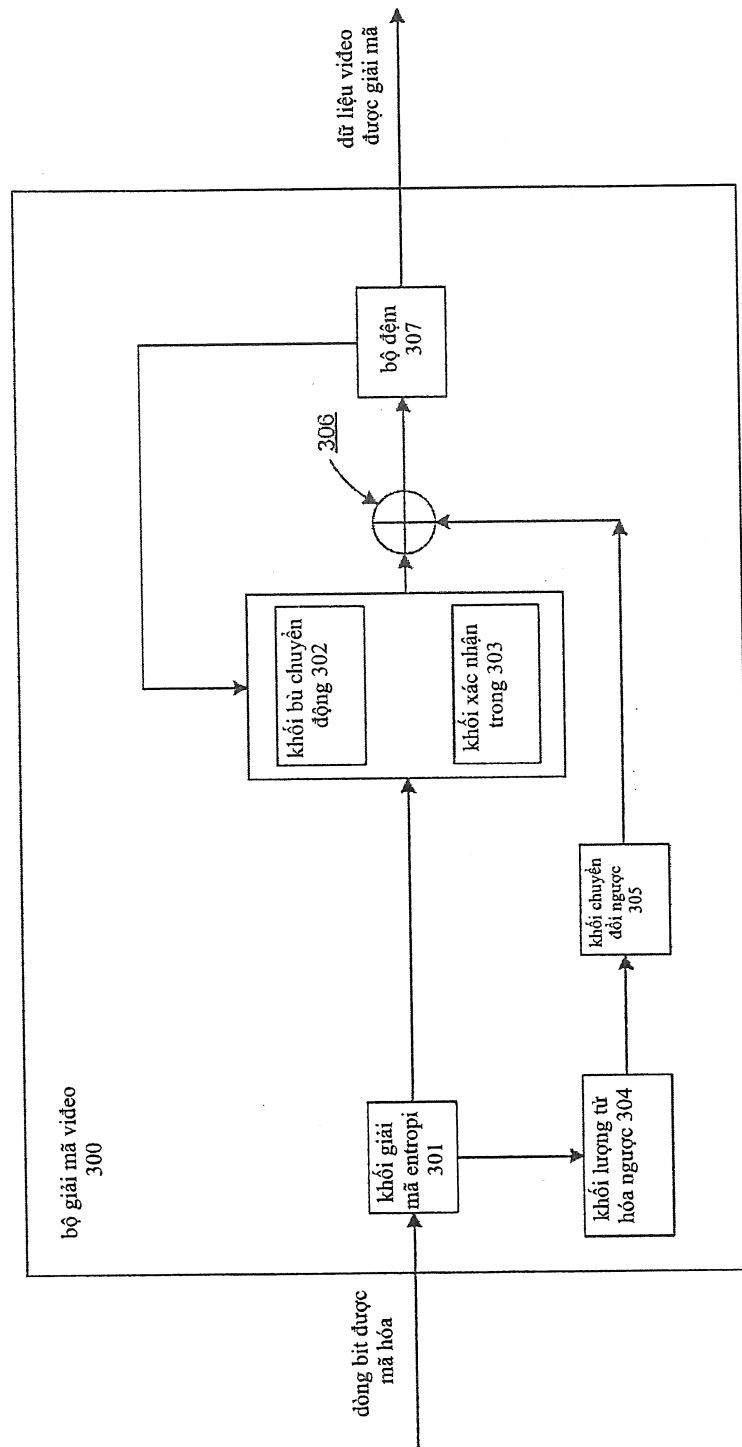


FIG. 10

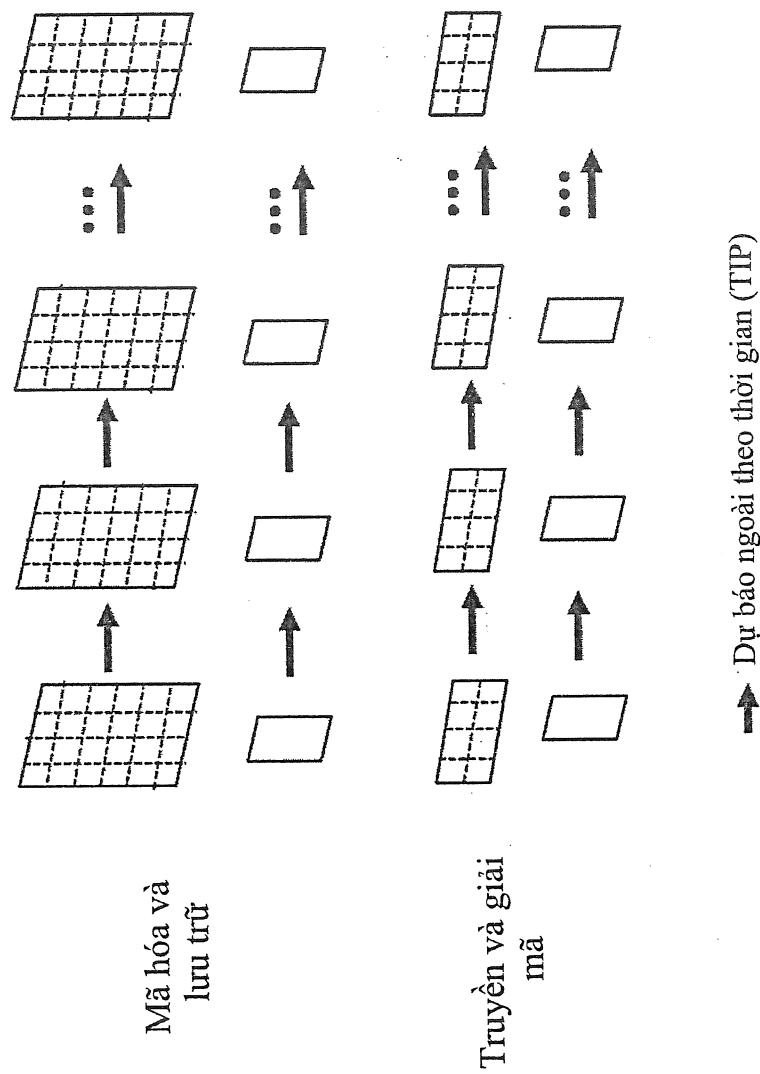


FIG. 11

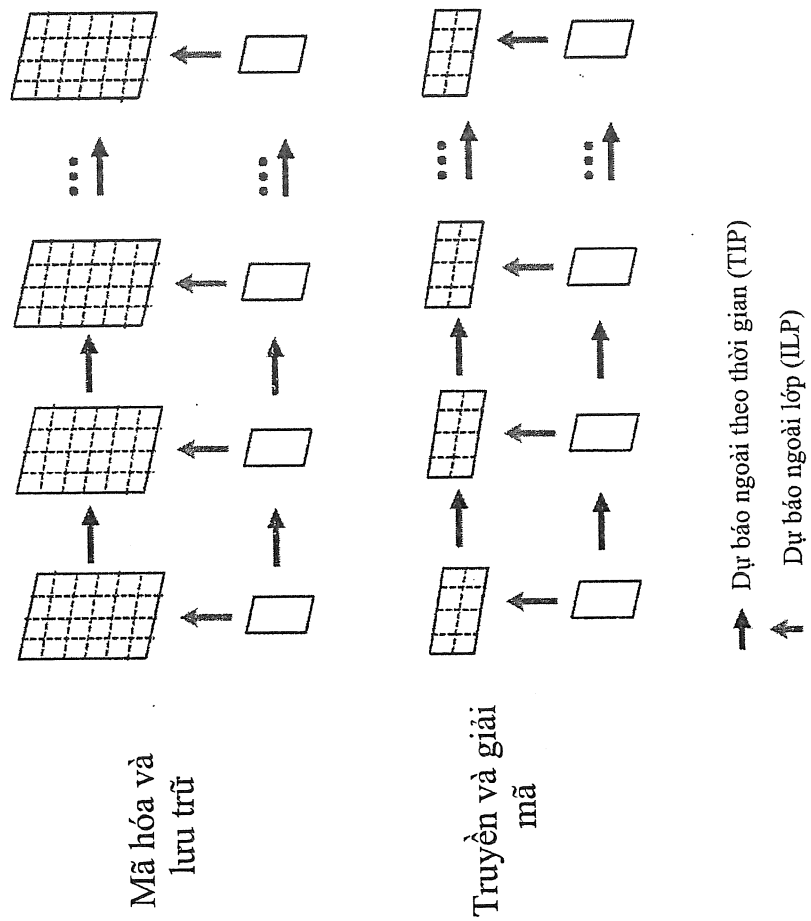
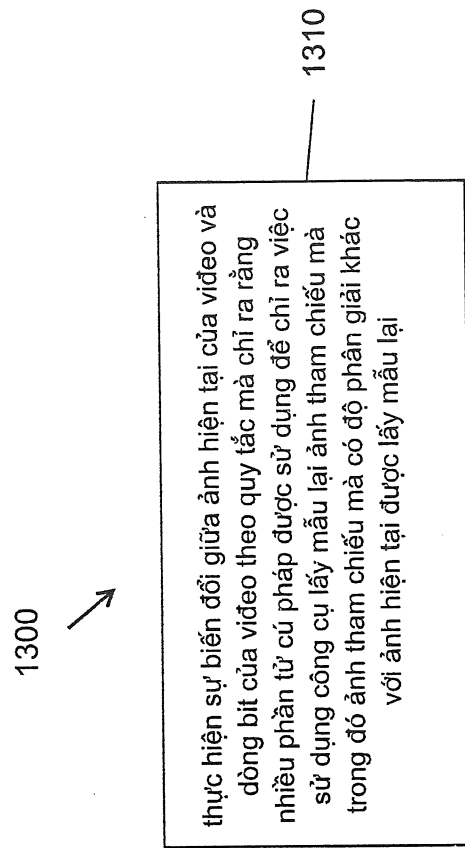
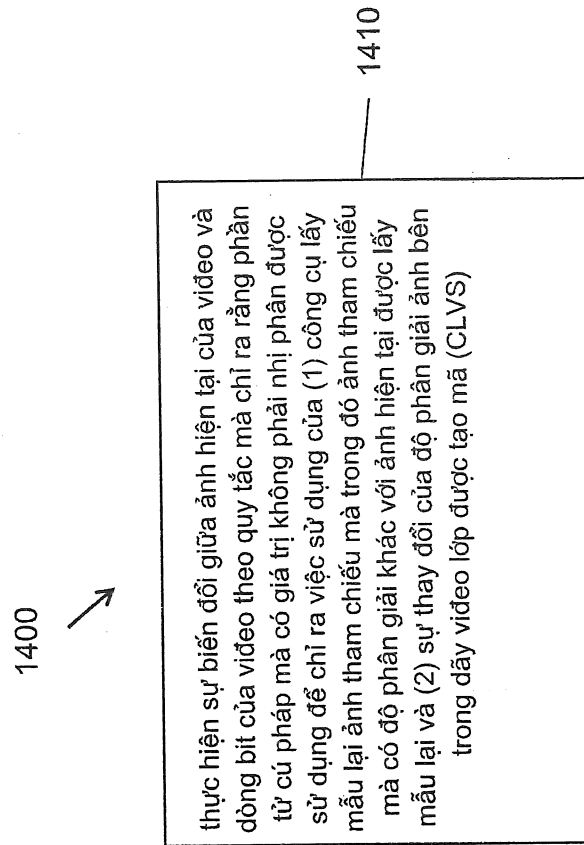


FIG. 12

**FIG. 13**

**FIG. 14**

1500



thực hiện sự biến đổi giữa video bao gồm nhiều lớp và dòng bit của video theo quy tắc mà chỉ ra rằng, trong trường hợp ảnh con được xử lý như ảnh video để biến đổi, sự hạn chế căn chỉnh ngang lớp được áp dụng cho ít hơn tất cả nhiều lớp này bao gồm lớp hiện tại mà bao gồm ảnh con và tập con của các lớp được liên kết với lớp hiện tại, trong đó sự hạn chế căn chỉnh ngang lớp bao gồm sự hạn chế của ít nhất kích thước của ảnh video, số lượng của các ảnh con bên trong ảnh video, vị trí của ít nhất một ảnh con, hoặc nhận dạng của ảnh con

1510

FIG. 15

1600



thực hiện sự biến đổi giữa lớp hiện tại của video và dòng bit của video theo quy tắc mà chỉ ra rằng sự hạn chế căn chỉnh ngang lớp được áp dụng cho tất cả các lớp trong cây phụ thuộc được liên kết với lớp hiện tại không phụ thuộc vào việc liệu lớp bất kỳ trong số tất cả các lớp có là lớp đầu ra trong tập lớp đầu ra, trong đó tất cả các lớp trong cây phụ thuộc bao gồm lớp hiện tại, tất cả các lớp mà có lớp hiện tại làm lớp tham chiếu, và tất cả các lớp tham chiếu của lớp hiện tại

1610

FIG. 16

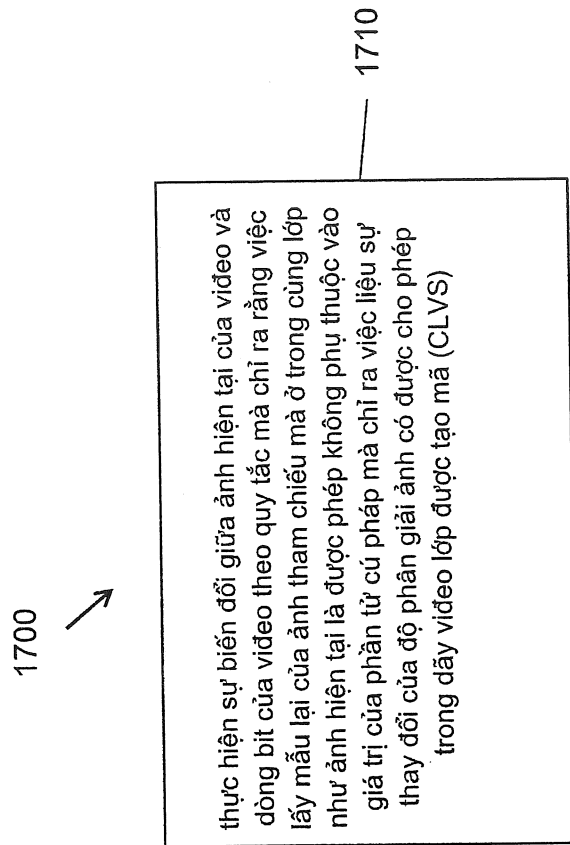
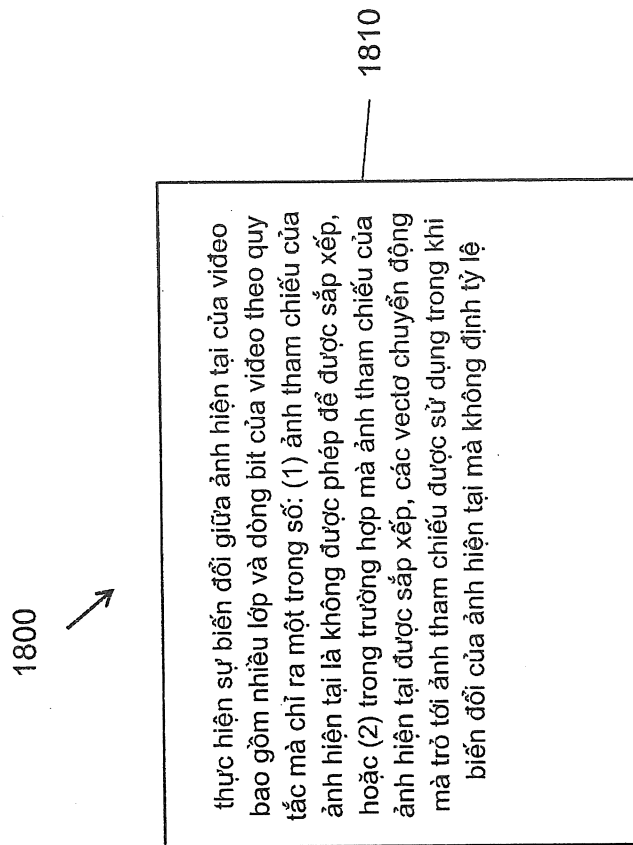
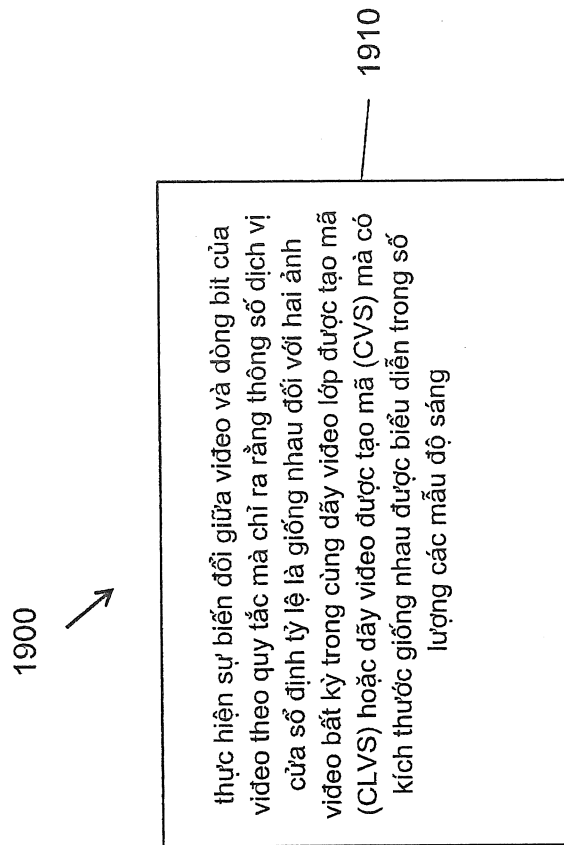


FIG. 17

**FIG. 18**

**FIG. 19**

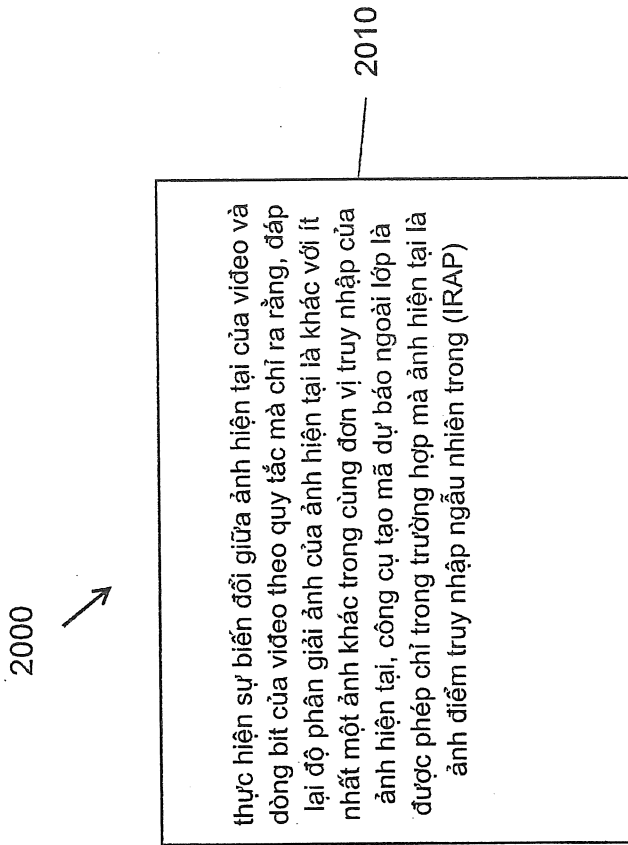


FIG. 20