



(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)



1-0024797

(51)<sup>7</sup> H04N 19/70; H04N 19/30

(13) B

(21) 1-2016-01803

(22) 23/10/2014

(86) PCT/US2014/061955 23/10/2014

(87) WO2015/061561A1 30/04/2015

(30) 61/894,886 23/10/2013 US; 14/521,099 22/10/2014 US

(45) 25/08/2020 389

(43) 25/07/2016 340A

(73) QUALCOMM INCORPORATED (US)

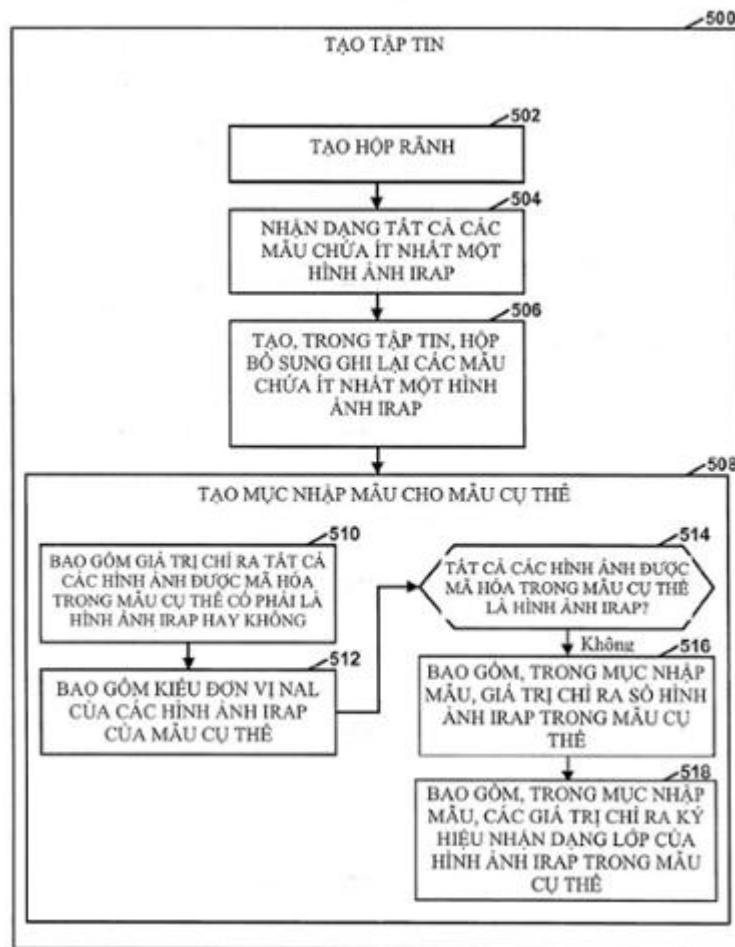
ATTN: International IP Administration, 5775 Morehouse Drive, San Diego, CA  
92121-1714, United States of America

(72) WANG, Ye-Kui (CN); CHEN, Ying (CN); RAMASUBRAMONIAN, Adarsh  
Krishnan (IN); HENDRY, Fnu (ID).

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

(54) PHƯƠNG PHÁP XỬ LÝ DỮ LIỆU VIDEO ĐA LỚP, THIẾT BỊ VIDEO VÀ VẬT  
GHI LƯU TRỮ DỮ LIỆU BẤT BIẾN ĐỌC ĐƯỢC BỞI MÁY TÍNH

(57) Sáng chế đề cập đến phương pháp xử lý dữ liệu video đa lớp, thiết bị video và vật ghi  
lưu trữ dữ liệu bất biến đọc được bởi máy tính có các lệnh được lưu trữ trên đó. Thiết bị  
điện toán tạo tập tin bao gồm hộp rãnh chứa siêu dữ liệu cho một rãnh trong tập tin. Dữ  
liệu phương tiện cho rãnh này bao gồm chuỗi các mẫu. Mỗi mẫu là một đơn vị truy cập  
video của dữ liệu video đa lớp. Một phần của quá trình tạo tập tin là thiết bị điện toán tạo,  
trong tập tin này, hộp bổ sung ghi lại tất cả các mẫu chứa ít nhất một hình ảnh điểm truy  
cập ngẫu nhiên trong ảnh (Intra Random Access Point-IRAP).



## **Lĩnh vực kỹ thuật được đề cập**

Sáng chế đề cập đến việc mã hóa video.

## **Tình trạng kỹ thuật của sáng chế**

Các khả năng xử lý video số có thể được kết hợp vào một loạt các thiết bị bao gồm ti-vi kỹ thuật số, các hệ thống truyền hình trực tiếp kỹ thuật số, các hệ thống truyền hình không dây, các thiết bị hỗ trợ số cá nhân (personal digital assistant - PDA), các máy tính xách tay hoặc máy tính để bàn, các máy tính bảng, các thiết bị đọc sách điện tử, các máy ảnh số, các thiết bị ghi âm số, các thiết bị phát phương tiện số, các thiết bị chơi trò chơi điện tử, bàn giao tiếp trò chơi điện tử, điện thoại di động hoặc điện thoại vô tuyến vệ tinh, được gọi là “điện thoại thông minh”, các thiết bị hội nghị truyền hình, các thiết bị truyền liên tục video, và tương tự. Các thiết bị video số thực hiện các kỹ thuật nén video, như các kỹ thuật được mô tả trong các chuẩn được xác định bởi MPEG-2, MPEG-4, ITU-T H.263, ITU-T H.264/MPEG-4, phần 10, mã hóa video tiên tiến (Advanced Video Coding - AVC), chuẩn mã hóa video hiệu suất cao (High Efficiency Video Coding - HEVC) hiện đang triển khai, và các phiên bản mở rộng của các chuẩn này. Các thiết bị video có thể truyền, nhận, mã hóa, giải mã, và/hoặc lưu trữ thông tin video số một cách có hiệu quả hơn bằng cách thực hiện các kỹ thuật nén video như vậy.

Các kỹ thuật nén video thực hiện dự đoán không gian (trong ảnh) và/hoặc dự đoán thời gian (liên ảnh) để giảm bớt hoặc loại bỏ phần dư vốn có trong các chuỗi video. Đối với việc mã hóa video dựa trên khối, lát video (tức là, khung video hoặc một phần của khung video) có thể được phân chia thành các khối video, còn có thể được gọi là khối cây, đơn vị mã hóa (coding unit-CU) và/hoặc nút mã hóa. Các khối video trong lát được mã hóa trong ảnh (I) của một hình ảnh được mã hóa bằng cách sử dụng kỹ thuật dự đoán không gian liên quan đến các mẫu tham chiếu trong khối lân cận trong cùng một hình ảnh. Khối video trong lát video được mã hóa liên ảnh (P hoặc B) của một hình ảnh có thể sử dụng kỹ thuật dự đoán không gian liên quan đến mẫu tham chiếu trong khối lân cận trong cùng một hình ảnh hoặc dự đoán thời gian liên

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

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

### **Bản chất kỹ thuật của sáng chế**

Nói chung, sáng chế đề cập đến việc lưu trữ nội dung video trong tập tin dựa trên định dạng tập tin phương tiện cơ sở của tổ chức tiêu chuẩn quốc tế (International Standards Organization base media file format-ISOBMFF). Một số ví dụ của bản mô tả này đề cập đến phương pháp lưu trữ dòng video chứa nhiều lớp mã hóa, trong đó mỗi lớp có thể là lớp mở rộng được, khung hình kết cấu, khung hình chiều sâu, v.v. và phương pháp này có thể áp dụng để lưu trữ dữ liệu mã hóa video hiệu suất cao đa khung hình (Multi-View High Efficiency Video Coding -MV-HEVC), HEVC mở rộng được (Scalable HEVC-SHVC), HEVC ba chiều (3D-HEVC), và các loại dữ liệu video khác.

Theo một khía cạnh, sáng chế đề cập đến phương pháp xử lý dữ liệu video đa lớp, phương pháp này bao gồm các bước: tạo tập tin bao gồm hộp rãnh chứa siêu dữ liệu cho rãnh trong tập tin, trong đó: dữ liệu phương tiện cho rãnh bao gồm chuỗi các mẫu, mỗi mẫu là một đơn vị truy cập video của dữ liệu video đa lớp, và việc tạo tập tin bao gồm: tạo, trong tập tin, hộp bổ sung ghi lại tất cả các mẫu chứa ít nhất một hình ảnh điểm truy cập ngẫu nhiên trong ảnh (Intra Random Access Point-IRAP).

Theo khía cạnh khác, sáng chế đề cập đến phương pháp xử lý dữ liệu video đa lớp, phương pháp này bao gồm các bước: thu nhận, từ tập tin, hộp rãnh chứa siêu dữ

liệu cho rãnh trong tập tin, trong đó dữ liệu phương tiện cho rãnh bao gồm chuỗi các mẫu, mỗi mẫu là một đơn vị truy cập video của dữ liệu video đa lớp, và thu nhận từ tập tin, hộp bổ sung ghi lại tất cả các mẫu chứa ít nhất một hình ảnh IRAP; và xác định, dựa trên thông tin trong hộp bổ sung, một hoặc nhiều mẫu chứa ít nhất một hình ảnh IRAP.

Theo khía cạnh khác, sáng chế đề cập đến thiết bị video bao gồm: vật ghi lưu trữ dữ liệu được tạo cấu hình để lưu trữ dữ liệu video đa lớp; và một hoặc nhiều bộ xử lý được tạo cấu hình để: tạo tập tin bao gồm hộp rãnh chứa siêu dữ liệu cho rãnh trong tập tin, trong đó: dữ liệu phương tiện cho rãnh bao gồm chuỗi các mẫu, mỗi mẫu là một đơn vị truy cập video của dữ liệu video đa lớp, và việc tạo tập tin bao gồm: tạo, trong tập tin, hộp bổ sung ghi lại tất cả các mẫu chứa ít nhất một hình ảnh IRAP.

Theo khía cạnh khác, sáng chế đề cập đến thiết bị video bao gồm: vật ghi lưu trữ dữ liệu được tạo cấu hình để lưu trữ dữ liệu video đa lớp; và một hoặc nhiều bộ xử lý được tạo cấu hình để: thu nhận, từ tập tin, hộp rãnh chứa siêu dữ liệu cho rãnh trong tập tin, trong đó dữ liệu phương tiện cho rãnh bao gồm chuỗi các mẫu, mỗi mẫu là một đơn vị truy cập video của dữ liệu video đa lớp, và thu nhận, từ tập tin, hộp bổ sung ghi lại tất cả các mẫu chứa ít nhất một hình ảnh IRAP; và xác định, dựa trên thông tin trong hộp bổ sung, một hoặc nhiều mẫu chứa ít nhất một hình ảnh IRAP.

Theo khía cạnh khác, sáng chế đề cập đến thiết bị video bao gồm: phương tiện tạo tập tin bao gồm hộp rãnh chứa siêu dữ liệu cho rãnh trong tập tin, trong đó: dữ liệu phương tiện cho rãnh bao gồm chuỗi các mẫu, mỗi mẫu là một đơn vị truy cập video của dữ liệu video đa lớp, và việc tạo tập tin bao gồm tạo, trong tập tin, hộp bổ sung ghi lại tất cả các mẫu chứa ít nhất một hình ảnh IRAP.

Theo khía cạnh khác, sáng chế đề cập đến thiết bị video bao gồm: phương tiện thu nhận, từ tập tin bao gồm hộp rãnh chứa siêu dữ liệu cho rãnh trong tập tin, trong đó dữ liệu phương tiện cho rãnh bao gồm chuỗi các mẫu, mỗi mẫu là một đơn vị truy cập video của dữ liệu video đa lớp, và phương tiện thu nhận, từ tập tin, hộp bổ sung ghi lại tất cả các mẫu chứa ít nhất một hình ảnh IRAP; và phương tiện xác định, dựa trên thông tin trong hộp bổ sung, một hoặc nhiều mẫu chứa ít nhất một hình ảnh IRAP.

Theo khía cạnh khác, sáng chế đề cập đến vật ghi lưu trữ dữ liệu được đọc được bằng máy tính có các lệnh được lưu trữ trên đó mà khi được thực hiện khiến cho một hoặc nhiều bộ xử lý: tạo tập tin bao gồm hộp rãnh chứa siêu dữ liệu cho rãnh trong tập tin,

trong đó: dữ liệu phương tiện cho rãnh bao gồm chuỗi các mẫu, mỗi mẫu là một đơn vị truy cập video của dữ liệu video đa lớp, và để tạo tập tin, một hoặc nhiều bộ xử lý tạo, trong tập tin, hộp bổ sung ghi lại tất cả các mẫu chứa ít nhất một hình ảnh IRAP.

Theo khía cạnh khác, sáng chế đề cập đến vật ghi lưu trữ dữ liệu đọc được bằng máy tính có các lệnh được lưu trữ trên đó mà khi được thực hiện khiến cho một hoặc nhiều bộ xử lý: thu nhận, từ tập tin, hộp rãnh chứa siêu dữ liệu cho rãnh trong tập tin, trong đó dữ liệu phương tiện cho rãnh bao gồm chuỗi các mẫu, mỗi mẫu là một đơn vị truy cập video của dữ liệu video đa lớp, và thu nhận, từ tập tin, hộp bổ sung ghi lại tất cả các mẫu chứa ít nhất một hình ảnh IRAP; và xác định, dựa trên thông tin trong hộp bổ sung, một hoặc nhiều mẫu chứa ít nhất một hình ảnh IRAP.

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

### **Mô tả vắn tắt các hình vẽ**

Fig.1 là sơ đồ khối minh họa hệ thống mã hóa và giải mã video làm ví dụ có thể sử dụng các kỹ thuật được mô tả trong bản mô tả này.

Fig.2 là sơ đồ khối minh họa bộ mã hóa video làm ví dụ mà có thể thực hiện các kỹ thuật được mô tả trong bản mô tả này.

Fig.3 là sơ đồ khối minh họa bộ giải mã video làm ví dụ mà có thể thực hiện các kỹ thuật được mô tả trong bản mô tả này.

Fig.4 là sơ đồ khối minh họa tập thiết bị làm ví dụ tạo thành một phần của mạng.

Fig.5 là sơ đồ khái niệm minh họa cấu trúc làm ví dụ của tập tin, theo một hoặc nhiều kỹ thuật của bản mô tả này.

Fig.6 là sơ đồ khái niệm minh họa cấu trúc làm ví dụ của tập tin, theo một hoặc nhiều kỹ thuật của bản mô tả này.

Fig.7 là lưu đồ minh họa thao tác làm ví dụ của thiết bị tạo tập tin, theo một hoặc nhiều kỹ thuật của bản mô tả này.

Fig.8 là lưu đồ minh họa thao tác làm ví dụ trong đó thiết bị máy tính thực hiện việc truy cập ngẫu nhiên và/hoặc chuyển mạch mức, theo một hoặc nhiều kỹ thuật của bản mô tả này.

Fig.9 là lưu đồ minh họa thao tác làm ví dụ của thiết bị tạo tập tin, theo một hoặc nhiều kỹ thuật của bản mô tả này.

Fig.10 là lưu đồ minh họa thao tác làm ví dụ của thiết bị điện toán, theo một hoặc nhiều kỹ thuật của bản mô tả này.

Fig.11 là lưu đồ minh họa thao tác làm ví dụ của thiết bị tạo tập tin, theo một hoặc nhiều kỹ thuật của bản mô tả này.

Fig.12 là lưu đồ minh họa thao tác làm ví dụ của thiết bị đích, theo một hoặc nhiều kỹ thuật của bản mô tả này.

### **Mô tả chi tiết sáng chế**

Định dạng tập tin phương tiện cơ sở (ISO base media file format-ISOBMFF) là định dạng tập tin để lưu trữ dữ liệu phương tiện. ISOBMFF có thể mở rộng để hỗ trợ lưu trữ dữ liệu video theo các chuẩn mã hóa video cụ thể. Ví dụ, ISOBMFF đã được mở rộng trước đó để hỗ trợ việc lưu trữ dữ liệu video theo các chuẩn mã hóa video H.264/AVC và chuẩn mã hóa video hiệu suất cao (High Efficiency Video Coding-HEVC). Hơn nữa, ISOBMFF đã được mở rộng để hỗ trợ việc lưu trữ dữ liệu video theo các phiên bản mở rộng mã hóa đa khung hình (Multi-view coding-MVC) và mã hóa video mở rộng được (Scalable video coding-SVC) của H.264/AVC. MV-HEVC, 3D-HEVC, và SHVC là các phiên bản mở rộng của chuẩn mã hóa video hỗ trợ dữ liệu video đa lớp. Các đặc điểm được bổ sung vào ISOBMFF để lưu trữ dữ liệu video theo các phiên bản mở rộng MVC và SVC của chuẩn H.264/AVC là chưa đủ để lưu trữ hiệu quả dữ liệu video theo MV-HEVC, 3D-HEVC, và SHVC. Nói cách khác, các vấn đề khác nhau có thể xuất hiện nếu sử dụng phiên bản mở rộng của ISOBMFF để lưu trữ dữ liệu video theo các phiên bản mở rộng MVC và SVC của chuẩn H.264/AVC để lưu trữ dữ liệu video theo MV-HEVC, 3D-HEVC, và SHVC.

Chẳng hạn, không giống như dòng bit theo các phiên bản mở rộng MVC hoặc SVC của H.264/AVC, dòng bit theo MV-HEVC, 3D-HEVC, hoặc SHVC có thể bao gồm đơn vị truy cập chứa các hình ảnh điểm truy cập ngẫu nhiên trong ảnh (Intra Random Access Point -IRAP) và các hình ảnh không IRAP. Đơn vị truy cập chứa các hình ảnh IRAP và hình ảnh không IRAP có thể được sử dụng để truy cập ngẫu nhiên theo MV-HEVC, 3D-HEVC, và SHVC. Tuy nhiên, ISOBMFF và các phiên bản mở rộng hiện có của nó không đưa ra cách thức để nhận dạng các đơn vị truy cập này.

Điều này có thể cản trở khả năng thiết bị điện toán thực hiện việc truy cập ngẫu nhiên và chuyển lớp.

Do vậy, theo một ví dụ của bản mô tả này, thiết bị điện toán có thể tạo tập tin bao gồm hộp rãnh chứa siêu dữ liệu cho rãnh trong tập tin. Dữ liệu phương tiện cho rãnh này bao gồm chuỗi các mẫu. Mỗi mẫu này có thể là đơn vị truy cập video của dữ liệu video đa lớp (ví dụ., dữ liệu video MV-HEVC, 3D-HEVC, or SHVC). Là một phần của quá trình tạo tập tin, thiết bị điện toán có thể tạo, trong tập tin này, hộp bổ sung ghi lại tất cả các mẫu chứa ít nhất một hình ảnh IRAP. Việc có thể xác định được các mẫu chứa các hình ảnh IRAP dựa trên thông tin có trong hộp bổ sung này có thể giúp thiết bị điện toán nhận tập tin để thực hiện truy cập ngẫu nhiên và chuyển lớp mà không cần phân tích và diễn dịch các đơn vị NAL. Điều này có thể làm giảm độ phức tạp và thời gian xử lý.

Hơn nữa, dữ liệu video đa lớp, như dữ liệu video MV-HEVC, 3D-HEVC, và SHVC, có thể bao gồm nhiều hình ảnh được mã hóa cho mỗi đơn vị truy cập. Tuy nhiên, ISOBMFF và các phiên bản mở rộng hiện có của nó không cung cấp thông tin về các hình ảnh được mã hóa riêng biệt trong đơn vị truy cập khi có nhiều hình ảnh được mã hóa trong đơn vị truy cập. Do vậy, trong các ví dụ trong đó thiết bị điện toán (ví dụ, máy chủ truyền liên tục) đang xác định xem có chuyển tiếp đơn vị NAL trong một tập tin hay không, thiết bị điện toán có thể cần phải phân tích và diễn dịch thông tin được lưu trữ trong các đơn vị NAL để xác định xem có chuyển tiếp đơn vị NAL hay không. Việc phân tích và diễn dịch thông tin được lưu trữ trong các đơn vị NAL có thể làm tăng tính phức tạp của thiết bị điện toán và có thể làm tăng độ trễ truyền liên tục.

Do vậy, theo một ví dụ của bản mô tả này, thiết bị điện toán có thể tạo tập tin bao gồm hộp rãnh chứa siêu dữ liệu cho rãnh trong tập. Dữ liệu phương tiện cho rãnh này bao gồm chuỗi các mẫu. Mỗi mẫu là đơn vị truy cập video của dữ liệu video đa lớp. Là một phần của quá trình tạo tập tin, thiết bị điện toán tạo ra, trong tập tin này, hộp thông tin mẫu con chứa các cờ chỉ rõ loại thông tin mẫu con có trong hộp thông tin mẫu con. Khi cờ có giá trị cụ thể, mẫu con tương ứng với hộp thông tin mẫu con chứa chính xác một hình ảnh được mã hóa và không hoặc nhiều đơn vị NAL không phải là lớp mã hóa video (Video Coding Layer - VCL) liên quan đến hình ảnh được mã hóa. Theo cách này, thiết bị điện toán nhận tập tin có thể sử dụng thông tin mẫu

con có trong hộp thông tin mẫu con để thực hiện việc xác định liên quan đến các hình ảnh được mã hóa riêng biệt trong một mẫu của tập tin. Đơn vị NAL không VCL liên quan đến hình ảnh được mã hóa bao gồm các đơn vị NAL cho tập tham số (ví dụ, PPS, SPS, VPS) và SEI có thể áp dụng cho hình ảnh được mã hóa.

Trong dữ liệu video đa lớp, đơn vị truy cập bao gồm các hình ảnh được mã hóa mà được đánh dấu là được kết xuất và hình ảnh được mã hóa mà được đánh dấu là không được kết xuất. Bộ giải mã video có thể sử dụng các hình ảnh được mã hóa mà được đánh dấu là không được kết xuất làm hình ảnh tham chiếu để giải mã hình ảnh được đánh dấu là được kết xuất. Tiêu đề của đơn vị NAL đối với đơn vị NAL của một lát hình ảnh có thể bao gồm cờ kết xuất hình ảnh (ví dụ, `pic_output_flag` theo HEVC) chỉ báo hình ảnh này có được đánh dấu là được kết xuất hay không. Trong tập tin ISOBMFF, mỗi mẫu cần được kết hợp với thời gian kết xuất (ví dụ, thời gian hợp thành) chỉ ra khi nào mẫu được kết xuất. Tuy nhiên, các hình ảnh được đánh dấu là không được kết xuất không có thời gian kết xuất. Do vậy, sự có mặt của các hình ảnh được đánh dấu là không được kết xuất có thể ảnh hưởng đến yêu cầu này của ISOBMFF hoặc có thể cần các kỹ thuật không theo chuẩn khác.

Do vậy, theo một hoặc nhiều kỹ thuật của bản mô tả này, thiết bị điện toán có thể tạo tập tin bao gồm hộp dữ liệu phương tiện chứa nội dung phương tiện. Nội dung phương tiện bao gồm chuỗi các mẫu. Mỗi mẫu bao gồm đơn vị truy cập của dữ liệu video đa lớp. Là một phần của quá trình tạo tập tin, thiết bị điện toán có thể, đáp lại việc xác định rằng ít nhất một đơn vị truy cập của dòng bit của dữ liệu video đa lớp bao gồm hình ảnh được mã hóa có cờ kết xuất hình ảnh bằng giá trị thứ nhất (ví dụ, 1) và hình ảnh được mã hóa có cờ kết xuất hình ảnh bằng giá trị thứ hai (ví dụ, 0), sử dụng ít nhất hai rãnh để lưu trữ dòng bit trong tập tin. Đối với mỗi rãnh tương ứng từ ít nhất hai rãnh này, tất cả các hình ảnh được mã hóa trong mỗi mẫu của rãnh tương ứng có cùng giá trị cờ kết xuất hình ảnh. Các hình ảnh có cờ kết xuất hình ảnh bằng giá trị thứ nhất (ví dụ, 1) được phép kết xuất và các hình ảnh có cờ kết xuất hình ảnh bằng giá trị thứ hai (ví dụ, 0) được phép được sử dụng làm hình ảnh tham chiếu nhưng không được phép kết xuất. Việc sử dụng ít nhất hai rãnh có thể giải quyết được vấn đề nêu trên, do mỗi mẫu trong mỗi rãnh có thể được gán một thời gian kết xuất thích hợp và bộ giải mã video không thể kết xuất các hình ảnh trong rãnh chứa các mẫu không được phép kết xuất.

Trong khi phần lớn phần mô tả về kỹ thuật theo sáng chế mô tả MV-HEVC, 3D-HEVC, và SHVC, người đọc sẽ nhận ra rằng các kỹ thuật theo sáng chế có thể được áp dụng cho các chuẩn mã hóa video khác và/hoặc phiên bản mở rộng của chúng.

Fig.1 là sơ đồ khối minh họa hệ thống mã hóa và giải mã video 10 làm ví dụ có thể sử dụng các kỹ thuật được mô tả trong bản mô tả này. Như được thể hiện trên Fig.1, hệ thống 10 bao gồm thiết bị nguồn 12 tạo ra dữ liệu video mã hóa cần được giải mã sau đó bởi thiết bị đích 14. Thiết bị nguồn 12 và thiết bị đích 14 có thể bao gồm thiết bị bất kỳ trong rất nhiều thiết bị, bao gồm máy tính để bàn, máy tính notebook (tức là, máy tính xách tay), máy tính bảng, bộ giải mã tín hiệu truyền hình, các máy điện thoại cầm tay như điện thoại "thông minh", bảng "thông minh", ti-vi, máy ảnh, thiết bị hiển thị, thiết bị phát phương tiện kỹ thuật số, bàn giao tiếp trò chơi video, thiết bị truyền liên tục video hoặc thiết bị tương tự. Theo một số trường hợp, thiết bị nguồn 12 và thiết bị đích 14 có thể được trang bị để truyền thông không dây. Thiết bị nguồn 12 và thiết bị đích 14 có thể được xem là các thiết bị video.

Theo ví dụ trên Fig.1, thiết bị nguồn 12 bao gồm nguồn video 18, bộ mã hóa video 20, và giao diện đầu ra 22. Theo một số trường hợp, giao diện đầu ra 22 có thể bao gồm bộ điều biến/bộ giải điều (môdem) và/hoặc bộ phát. Trong thiết bị nguồn 12, nguồn video 18 có thể bao gồm nguồn như thiết bị thu video, ví dụ, máy quay video, kho lưu trữ video chứa video được thu trước đó, giao diện cấp video để nhận video từ nhà cung cấp nội dung video, và/hoặc hệ thống đồ họa máy tính để tạo dữ liệu đồ họa máy tính như video nguồn, hoặc tổ hợp của các nguồn này. Tuy nhiên, kỹ thuật được mô tả trong bản mô tả này có thể được áp dụng để mã hóa video nói chung và có thể được áp dụng cho các ứng dụng không dây và/hoặc có dây.

Bộ mã hóa video 20 có thể mã hóa video được thu, được thu trước hoặc được tạo ra bởi máy tính. Thiết bị nguồn 12 có thể truyền dữ liệu video mã hóa trực tiếp đến thiết bị đích 14 qua giao diện đầu ra 22 của thiết bị nguồn 12. Dữ liệu video được mã hóa cũng có thể (hoặc theo cách khác) được lưu trữ trên thiết bị lưu trữ 33 để truy cập sau đó bởi thiết bị đích 14 hoặc thiết bị khác để giải mã và/hoặc phát lại.

Thiết bị đích 14 bao gồm giao diện đầu vào 28, bộ giải mã video 30, và thiết bị hiển thị 32. Theo một số trường hợp, giao diện đầu vào 28 có thể bao gồm bộ thu và/hoặc môdem. Giao diện đầu vào 28 của thiết bị đích 14 nhận dữ liệu video mã hóa từ liên kết 16. Dữ liệu video mã hóa được truyền trên liên kết 16, hoặc được lưu trên

thiết bị lưu trữ 33, có thể bao gồm nhiều phần tử cú pháp được tạo ra bởi bộ mã hóa video 20 để sử dụng bởi bộ giải mã video, như bộ giải mã video 30, trong việc giải mã dữ liệu video. Các phần tử cú pháp này có thể được bao gồm với dữ liệu video mã hóa được truyền trên phương tiện truyền thông, được lưu trữ trên phương tiện lưu trữ hoặc lưu trữ trên máy chủ tập tin.

Thiết bị hiển thị 32 có thể được tích hợp vào hoặc có thể ở bên ngoài thiết bị đích 14. Trong một số ví dụ, thiết bị đích 14 có thể bao gồm thiết bị hiển thị được tích hợp và còn có thể được tạo cấu hình để tương tác với thiết bị hiển thị bên ngoài. Trong các ví dụ khác, thiết bị đích 14 có thể là thiết bị hiển thị. Nhìn chung, thiết bị hiển thị 32 hiển thị dữ liệu video được giải mã cho người dùng, và có thể bao gồm thiết bị bất kỳ trong số nhiều thiết bị hiển thị như màn hình tinh thể lỏng (liquid crystal display - LCD), màn hình plasma, màn hình điốt phát quang hữu cơ (organic light emitting diode - OLED), hoặc các loại thiết bị hiển thị khác.

Mỗi bộ mã hóa video 20 và bộ giải mã video 30 có thể được thực thi dưới dạng bất kỳ trong nhiều sơ đồ mạch mã hóa thích hợp, như một hoặc nhiều bộ vi xử lý, bộ xử lý tín hiệu số (digital signal processor - DSP), mạch tích hợp chuyên dụng (Application specific integrated circuits-ASIC), mảng cổng lập trình được dạng trường (Field programmable gate array - FPGA), logic rời rạc, phần mềm, phần cứng, phần sụn hoặc tổ hợp bất kỳ của chúng. Khi các kỹ thuật được thực hiện một phần trong phần mềm, thiết bị có thể lưu trữ các lệnh cho phần mềm trong vật ghi bất biến đọc được bởi máy tính thích hợp và thực thi các lệnh trong phần cứng bằng cách sử dụng một hoặc nhiều bộ xử lý để thực hiện các kỹ thuật theo sáng chế. Mỗi bộ mã hóa video 20 và bộ giải mã video 30 có thể được bao gồm trong một hoặc nhiều bộ mã hóa hoặc giải mã, cả hai có thể được tích hợp làm một bộ phận của bộ mã hóa/bộ giải mã kết hợp (CODEC) trong thiết bị tương ứng.

Thiết bị đích 14 có thể nhận dữ liệu video mã hóa cần được giải mã qua liên kết 16. Liên kết 16 có thể bao gồm loại phương tiện hoặc thiết bị bất kỳ có khả năng di chuyển dữ liệu video mã hóa từ thiết bị nguồn 12 đến thiết bị đích 14. Theo một ví dụ, liên kết 16 có thể bao gồm phương tiện truyền thông làm cho thiết bị nguồn 12 có thể truyền dữ liệu video mã hóa trực tiếp đến thiết bị đích 14 theo thời gian thực. Dữ liệu video mã hóa có thể được điều biến theo chuẩn truyền thông, như giao thức truyền thông không dây và được truyền đến thiết bị đích 14. Phương tiện truyền thông có thể

bao gồm phương tiện truyền thông không dây hoặc có dây bất kỳ, như phổ tần số vô tuyến (radio frequency-RF) hoặc một hoặc nhiều dây truyền vật lý. Phương tiện truyền thông có thể tạo thành một phần của mạng dựa trên gói, như mạng diện rộng, hoặc mạng toàn cầu như Internet. Phương tiện truyền thông có thể bao gồm các bộ định tuyến, bộ chuyển mạch, trạm gốc, hoặc thiết bị bất kỳ khác có thể là hữu ích để tạo điều kiện thuận lợi cho việc truyền thông từ thiết bị nguồn 12 đến thiết bị đích 14.

Theo cách khác, giao diện đầu ra 22 có thể kết xuất dữ liệu mã hóa đến thiết bị lưu trữ 33. Tương tự, giao diện đầu vào 28 có thể truy cập vào thiết bị lưu trữ dữ liệu mã hóa 33. Thiết bị lưu trữ 33 có thể bao gồm thiết bị bất kỳ trong số các phương tiện lưu trữ dữ liệu truy cập phân tán hoặc cục bộ như đĩa cứng, đĩa Blu-ray, DVD, CD-ROM, bộ nhớ nhanh, bộ nhớ khả biến hoặc không khả biến, hoặc phương tiện lưu trữ số thích hợp bất kỳ khác để lưu trữ dữ liệu video mã hóa. Theo ví dụ khác, thiết bị lưu trữ 33 có thể tương đương với máy chủ tập tin hoặc thiết bị lưu trữ trung gian khác có thể lưu trữ video mã hóa được tạo ra bởi thiết bị nguồn 12. Thiết bị đích 14 có thể truy cập dữ liệu video được lưu trữ từ thiết bị lưu trữ 33 thông qua truyền liên tục hoặc tải về. Máy chủ tập tin có thể là loại máy chủ bất kỳ có khả năng lưu trữ dữ liệu video mã hóa và truyền dữ liệu video mã hóa đó đến thiết bị đích 14. Máy chủ tập tin làm ví dụ bao gồm máy chủ web (ví dụ, đối với website), máy chủ FTP, thiết bị lưu trữ gắn mạng (Network attached storage -NAS), hoặc ổ đĩa cục bộ. Thiết bị đích 14 có thể truy cập dữ liệu video mã hóa thông qua kết nối dữ liệu chuẩn bất kỳ, kể cả kết nối Internet. Kết nối này có thể bao gồm kênh không dây (ví dụ, kết nối Wi-Fi), kết nối có dây (ví dụ, DSL, modem cáp, v.v.), hoặc tổ hợp của cả hai mà phù hợp để truy cập dữ liệu video mã hóa được lưu trữ trên máy chủ tập tin. Việc truyền dữ liệu video mã hóa từ thiết bị lưu trữ 33 có thể là truyền liên tục, truyền tải xuống, hoặc kết hợp cả hai.

Kỹ thuật theo sáng chế không nhất thiết giới hạn ở các ứng dụng hoặc thiết lập không dây. Các kỹ thuật này có thể được áp dụng cho mã hóa video để hỗ trợ cho ứng dụng bất kỳ trong nhiều ứng dụng đa phương tiện, chẳng hạn như phát sóng truyền hình qua không gian, phát sóng truyền hình cáp, phát sóng truyền hình vệ tinh, truyền liên tục video, chẳng hạn qua mạng Internet, mã hóa dữ liệu video để lưu trữ trên vật ghi lưu trữ dữ liệu, giải mã video số được lưu trên vật ghi lưu trữ dữ liệu, hoặc các ứng dụng khác. Trong một số ví dụ, hệ thống 10 có thể được tạo cấu hình để hỗ trợ truyền

video một chiều hoặc hai chiều để hỗ trợ các ứng dụng như truyền liên tục video, phát lại video, phát rộng video, và/hoặc điện thoại video.

Hơn nữa, theo ví dụ trên Fig.1, hệ thống mã hóa video 10 bao gồm thiết bị tạo tập tin 34. Thiết bị tạo tập tin 34 có thể nhận dữ liệu video mã hóa được tạo ra bởi thiết bị nguồn 12. Thiết bị tạo tập tin 34 có thể tạo tập tin bao gồm dữ liệu video mã hóa. Thiết bị đích 14 có thể nhận tập tin được tạo ra bởi thiết bị tạo tập tin 34. Theo các ví dụ khác nhau, thiết bị tạo tập tin 34 có thể bao gồm nhiều thiết bị điện toán khác nhau. Chẳng hạn, thiết bị tạo tập tin 34 có thể bao gồm phần tử mạng nhận biết truyền thông (Media Aware Network Element-MANE), thiết bị điện toán máy chủ, thiết bị điện toán cá nhân, thiết bị điện toán chuyên dụng, thiết bị điện toán thương mại hoặc loại thiết bị điện toán khác. Theo một số ví dụ, thiết bị tạo tập tin 34 là một phần của mạng cung cấp nội dung. Thiết bị tạo tập tin 34 có thể nhận được dữ liệu video mã hóa từ thiết bị nguồn 12 thông qua kênh như liên kết 16. Hơn nữa, thiết bị đích 14 có thể nhận tập tin từ thiết bị tạo tập tin 34 thông qua kênh như liên kết 16. Thiết bị tạo tập tin 34 có thể được xem là thiết bị video.

Theo các ví dụ khác, thiết bị nguồn 12 hoặc thiết bị điện toán khác có thể tạo ra tập tin bao gồm dữ liệu video mã hóa. Tuy nhiên, để dễ giải thích, bản mô tả này mô tả thiết bị tạo tập tin 34 khi tạo tập tin. Tuy nhiên, cần hiểu rằng phần mô tả này không được áp dụng cho các thiết bị điện toán nói chung.

Bộ mã hóa video 20 và bộ giải mã video 30 có thể hoạt động theo chuẩn nén video, như chuẩn mã hóa video hiệu suất cao (HEVC) hoặc phiên bản mở rộng của nó. Chuẩn HEVC cũng có thể được gọi là ISO/IEC 23008-2. Gần đây, thiết kế của HEVC đã được hoàn chỉnh bởi Nhóm Cộng tác Chung về Mã hóa Video (Joint Collaboration Team on Video Coding - JCT-VC) của Nhóm Chuyên gia Mã hóa Video ITU-T (Video Coding Experts Group - VCEG) và Nhóm Chuyên Gia Hình Ảnh Động (Motion Picture Experts Group - MPEG) ISO/IEC. Dự thảo HEVC mới nhất và dưới đây được gọi là HEVC WD có sẵn từ [http://phenix.int-evry.fr/jct/doc\\_end\\_user/documents/14\\_Vienna/wg11/JCTVC-N1003-v1.zip](http://phenix.int-evry.fr/jct/doc_end_user/documents/14_Vienna/wg11/JCTVC-N1003-v1.zip). Phiên bản mở rộng đa khung hình đối với HEVC, tức là MV-HEVC, cũng đang được phát triển bởi JCT-3V. Dự thảo làm việc gần đây (Working Draft-WD) của MV-HEVC, có tựa đề “MV-HEVC Draft Text 5” và dưới đây được gọi là MV-HEVC WD5, có sẵn từ [http://phenix.it-sudparis.eu/jct2/doc\\_end\\_user/documents/5\\_Vienna/wg11/JCT3V-](http://phenix.it-sudparis.eu/jct2/doc_end_user/documents/5_Vienna/wg11/JCT3V-)

E1004-v6.zip. Phiên bản mở rộng có thể thay đổi được đối với HEVC, tức là SHVC, cũng đang được phát triển bởi JCT-VC. Dự thảo làm việc gần đây (WD) của SHVC, có tựa đề “High efficiency video coding (HEVC) scalable extension draft 3” và dưới đây được gọi là SHVC WD3, có sẵn từ [http://phenix.it-sudparis.eu/jct/doc\\_end\\_user/documents/14\\_Vienna/wg11/JCTVC-N1008-v3.zip](http://phenix.it-sudparis.eu/jct/doc_end_user/documents/14_Vienna/wg11/JCTVC-N1008-v3.zip). Dự thảo làm việc gần đây (WD) của phiên bản mở rộng của HEVC, có sẵn từ [http://phenix.int-evry.fr/jct/doc\\_end\\_user/documents/14\\_Vienna/wg11/JCTVC-N1005-v3.zip](http://phenix.int-evry.fr/jct/doc_end_user/documents/14_Vienna/wg11/JCTVC-N1005-v3.zip). Dự thảo làm việc gần đây (WD) của phiên bản mở rộng 3D của HEVC, tức là 3D-HEVC, có tựa đề “3D-HEVC Draft Text 1” có sẵn từ [http://phenix.int-evry.fr/jct2/doc\\_end\\_user/documents/5\\_Vienna/wg11/JCT3V-E1001-v3.zip](http://phenix.int-evry.fr/jct2/doc_end_user/documents/5_Vienna/wg11/JCT3V-E1001-v3.zip). Bộ hóa video 20 và bộ giải mã 30 có thể hoạt động theo một hoặc nhiều chuẩn này.

Theo cách khác, bộ mã hóa video 20 và bộ giải mã video 30 có thể hoạt động theo chuẩn độc quyền hoặc công nghiệp khác, như chuẩn ITU-T H.264, theo cách khác được gọi là MPEG-4, phần 10, mã hóa video tiên tiến (AVC), hoặc các phiên bản mở rộng của các chuẩn này. Tuy nhiên, các kỹ thuật trong bản mô tả này không chỉ giới hạn ở chuẩn mã hóa cụ thể bất kỳ. Các ví dụ khác về chuẩn nén video bao gồm ITU-T H.261, ISO/IEC MPEG-1 Visual, ITU-T H.262 or ISO/IEC MPEG-2 Visual, ITU-T H.263, ISO/IEC MPEG-4 Visual và ITU-T H.264 (còn được biết đến là ISO/IEC MPEG-4 AVC), bao gồm các phiên bản mở rộng SVC và MVC của nó.

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

JCT-VC đang cố gắng phát triển chuẩn HEVC. Nỗ lực chuẩn hóa HEVC là dựa trên mẫu tiến hóa của thiết bị mã hóa video được gọi là Mẫu thử nghiệm HEVC (HM). HM giả định các khả năng phụ trợ của các thiết bị mã hóa video liên quan đến các thiết bị sẵn có theo, ví dụ, ITU-T H.264/AVC. Ví dụ, trong khi H.264/AVC cung cấp 9 chế độ mã hóa dự đoán trong ảnh, thì HEVC HM có thể cung cấp 33 chế độ mã hóa dự đoán trong ảnh.

Nói chung, mô hình làm việc của HM mô tả rằng khung video hoặc hình ảnh video có thể được chia thành chuỗi các khối cây hoặc đơn vị mã hóa lớn nhất (Largest coding unit-LCU) bao gồm cả mẫu độ sáng lẫn mẫu độ màu. Khối cây cũng có thể được gọi là đơn vị cây mã hóa (Coding Tree Unit-CTU). Khối cây có mục đích giống như khối marco của chuẩn H.264/AVC. Lát bao gồm một số khối cây liên tiếp theo thứ tự mã hóa. Khung hoặc hình ảnh video có thể được chia thành một hoặc nhiều lát. Mỗi khối cây có thể được chia thành các đơn vị mã hóa (CU) theo cây tứ phân. Chẳng hạn, khối cây, là nút rễ của cây tứ phân, có thể được chia thành bốn nút con và mỗi nút con có thể là nút cha và có thể được chia thành bốn nút con khác. Nút con không được chia cuối cùng, là nút lá của cây tứ phân, bao gồm nút mã hóa, tức là khối video mã hóa. Dữ liệu cú pháp liên quan đến dòng bit mã hóa có thể xác định số lần tối đa một khối cây có thể được phân chia và có thể xác định kích thước tối đa của nút mã hóa.

CU bao gồm nút mã hóa và các đơn vị dự đoán (prediction unit - PU) và đơn vị biến đổi (transform unit - TU) liên quan đến nút mã hóa này. Kích thước của CU tương ứng với kích thước của nút mã hóa và phải là hình vuông. Kích thước của CU có thể nằm trong khoảng từ 8x8 điểm ảnh đến kích thước của khối cây có tối đa 64x64 điểm ảnh hoặc lớn hơn. Mỗi CU có thể bao gồm một hoặc nhiều PU và một hoặc nhiều TU. Dữ liệu cú pháp liên quan đến CU có thể mô tả, ví dụ, sự phân chia CU thành một hoặc nhiều PU. Các chế độ phân chia có thể khác nhau giữa việc CU được mã hóa theo chế độ bỏ qua hoặc trực tiếp, được mã hóa theo chế độ dự đoán trong ảnh hoặc được mã hóa theo chế độ dự đoán liên ảnh. PU có thể được phân chia thành dạng không phải hình vuông. Dữ liệu cú pháp liên quan đến CU cũng có thể mô tả, ví dụ, sự phân chia của CU thành một hoặc nhiều TU theo cây tứ phân. TU có thể có dạng hình vuông hoặc không phải hình vuông.

Chuẩn HEVC cho phép biến đổi theo các đơn vị biến đổi (transform unit - TU), các đơn vị này có thể khác nhau đối với các CU khác nhau. Các TU thường được xác định kích thước dựa trên kích thước của các PU trong CU cho trước được xác định cho LCU được chia, mặc dù có thể không phải luôn gặp trường hợp này. Các TU thường có kích thước bằng hoặc nhỏ hơn các PU. Trong một số ví dụ, các mẫu dư tương ứng với CU có thể được chia nhỏ thành các đơn vị nhỏ hơn bằng cách sử dụng cấu trúc cây tứ phân gọi là “cây tứ phân dư” (residual quad tree - RQT). Các nút lá của RQT có thể

được gọi là đơn vị biến đổi (TU). Các giá trị độ lệch điểm ảnh liên quan đến các TU có thể được biến đổi để tạo ra các hệ số biến đổi, các hệ số này có thể được lượng tử hóa.

Nhìn chung, PU bao gồm dữ liệu liên quan đến quy trình dự đoán. Ví dụ, khi PU được mã hóa theo chế độ trong ảnh, PU có thể bao gồm dữ liệu mà mô tả chế độ dự đoán trong ảnh đối với PU. Theo ví dụ khác, khi PU được mã hóa theo chế độ trong ảnh, PU có thể bao gồm dữ liệu mà xác định vector chuyển động cho PU. Dữ liệu xác định vector chuyển động cho PU có thể mô tả, ví dụ, thành phần ngang của vector chuyển động, thành phần dọc của vector chuyển động, độ phân giải cho vector chuyển động (ví dụ, độ chính xác một phần tư điểm ảnh hoặc độ chính xác một phần tám điểm ảnh), hình ảnh tham chiếu mà vector chuyển động trỏ đến, và/hoặc danh sách hình ảnh tham chiếu (ví dụ, Danh sách 0, Danh sách 1, hoặc Danh sách C) đối với vector chuyển động.

Nhìn chung, TU được sử dụng cho các quy trình biến đổi và lượng tử hóa. CU cho trước có một hoặc nhiều PU có thể còn bao gồm một hoặc nhiều đơn vị biến đổi (TU). Sau khi dự đoán, bộ mã hóa video 20 có thể tính toán giá trị dư tương ứng với PU. Các giá trị dư bao gồm các giá trị độ lệch điểm ảnh mà có thể được biến đổi thành các hệ số biến đổi, được lượng tử hóa, và được quét bằng cách sử dụng các TU để tạo ra các hệ số biến đổi được xếp theo thứ tự để mã hóa entropy. Bản mô tả này thường sử dụng thuật ngữ “khối video” để chỉ nút mã hóa (tức là, khối mã hóa) của CU. Trong một số trường hợp cụ thể, bản mô tả này cũng có thể sử dụng thuật ngữ “khối video” để chỉ khối cây, tức là, LCU, hoặc CU, bao gồm nút mã hóa và các PU và TU.

Chuỗi video thường bao gồm dãy các khung video hoặc hình ảnh video. Nhóm hình ảnh (group of picture - GOP) thường bao gồm một dãy gồm một hoặc nhiều hình ảnh video. GOP có thể bao gồm dữ liệu cú pháp trong tiêu đề của GOP, tiêu đề của một hoặc nhiều hình ảnh, hoặc ở vị trí khác, mà mô tả một số hình ảnh được bao gồm trong GOP. Mỗi lát của một hình ảnh có thể bao gồm dữ liệu cú pháp lát mà mô tả chế độ mã hóa cho lát tương ứng. Bộ mã hóa video 20 thường hoạt động trên các khối video trong các lát video riêng lẻ nhằm mã hóa dữ liệu video. Khối video có thể tương ứng với nút mã hóa trong CU. Các khối video có thể có kích thước cố định hoặc biến thiên, và có thể khác nhau về kích thước theo chuẩn mã hóa riêng.

Ví dụ, HM hỗ trợ việc dự đoán ở các kích thước PU khác nhau. Giả định rằng kích thước của CU cụ thể là  $2N \times 2N$ , HM hỗ trợ dự đoán trong ảnh ở kích thước PU là

$2N \times 2N$  hoặc  $N \times N$ , và dự đoán liên ảnh ở các kích thước PU không đối xứng  $2N \times 2N$ ,  $2N \times N$ ,  $N \times 2N$ , hoặc  $N \times N$ . HM cũng hỗ trợ phân chia không đối xứng để dự đoán liên ảnh ở các kích thước PU  $2N \times nU$ ,  $2N \times nD$ ,  $nL \times 2N$ , và  $nR \times 2N$ . Trong quá trình phân chia không đối xứng, một chiều của CU không được phân chia, còn chiều kia được phân chia thành 25% và 75%. Phần của CU tương ứng với phần chia 25% được thể hiện bởi “n” tiếp sau là ký hiệu “trên”, “dưới” “trái”, hoặc “phải”. Do vậy, ví dụ, “ $2N \times nU$ ” chỉ CU  $2N \times 2N$  được chia theo chiều ngang với PU  $2N \times 0,5N$  ở trên và PU  $2N \times 1,5N$  PU ở dưới.

Theo bản mô tả này, “ $N \times N$ ” and “N nhân N” có thể sử dụng hoán đổi cho nhau để chỉ kích thước điểm ảnh của khối video theo kích thước dọc và ngang, ví dụ  $16 \times 16$  điểm ảnh hoặc 16 nhân 16 điểm ảnh. Nói chung, khối  $16 \times 16$  có 16 điểm ảnh theo chiều dọc ( $y = 16$ ) và 16 điểm ảnh theo chiều ngang ( $x = 16$ ). Tương tự, khối  $N \times N$  thường có N điểm ảnh theo chiều dọc và N điểm ảnh theo chiều ngang, trong đó N biểu diễn giá trị nguyên không âm. Các điểm ảnh trong một khối có thể được sắp xếp theo hàng và cột. Hơn nữa, các khối không nhất thiết phải có số điểm theo chiều ngang bằng số điểm ảnh theo chiều dọc. Ví dụ, các khối có thể bao gồm  $N \times M$  điểm ảnh, trong đó M không nhất thiết bằng N.

Tiếp theo quá trình mã hóa dự đoán liên ảnh hoặc trong ảnh sử dụng các PU của CU, bộ mã hóa video 20 có thể tính toán dữ liệu dư cho các TU của CU. PU có thể bao gồm dữ liệu điểm ảnh trong miền không gian (còn được gọi là miền điểm ảnh) và TU có thể bao gồm các hệ số trong miền biến đổi sau khi áp dụng phép biến đổi, ví dụ, phép biến đổi cosin rời rạc (DCT), biến đổi số nguyên, biến đổi sóng con hoặc phép biến đổi tương tự về mặt khái niệm đối với dữ liệu video dư. Dữ liệu dư có thể tương ứng với các chênh lệch điểm ảnh giữa các điểm ảnh của hình ảnh không được mã hóa và giá trị dự đoán tương ứng với các PU. Bộ mã hóa video 20 có thể tạo ra các TU bao gồm dữ liệu dư cho CU, và sau đó biến đổi các TU để tạo ra hệ số biến đổi cho CU.

Sau phép biến đổi bất kỳ để tạo ra các hệ số biến đổi, bộ mã hóa video 20 có thể thực hiện lượng tử hóa các hệ số biến đổi này. Lượng tử hóa thường chỉ quá trình trong đó các hệ số biến đổi được lượng tử hóa để có thể giảm lượng dữ liệu được sử dụng để biểu diễn các hệ số biến đổi, tạo ra sự nén thêm. Quy trình lượng tử hóa có thể làm giảm độ sâu bit được kết hợp với một số hoặc tất cả các hệ số biến đổi. Ví dụ, giá trị n

bit có thể được làm tròn xuống thành giá trị  $m$  bit trong quá trình lượng tử hóa, trong đó  $n$  lớn hơn  $m$ .

Trong một số ví dụ, bộ mã hóa video 20 có thể sử dụng thứ tự quét được định trước để quét các hệ số biến đổi lượng tử hóa nhằm tạo ra vector nối tiếp hóa mà có thể được mã hóa entropy. Trong các ví dụ khác, bộ mã hóa video 20 có thể thực hiện quét thích ứng. Sau khi quét các hệ số biến đổi đã lượng tử hóa để tạo ra vector một chiều, bộ mã hóa video 20 có thể mã hóa entropy vector một chiều này, ví dụ, theo phương pháp mã hóa chiều dài biến đổi thích ứng với ngữ cảnh (context adaptive variable length coding-CAVLC), mã hóa số học nhị phân thích ứng ngữ cảnh (Context adaptive binary arithmetic coding-CABAC), mã hóa thuật toán nhị phân thích ứng với ngữ cảnh dựa trên cú pháp (Syntax-based context-adaptive binary arithmetic coding-SBAC), Mã hóa entropy phân chia khoảng xác suất (Probability Interval Partitioning Entropy-PIPE) hoặc phương pháp mã hóa entropy khác. Bộ mã hóa video 20 cũng có thể mã hóa entropy các phần tử cú pháp liên quan đến dữ liệu video mã hóa để bộ giải mã video 30 sử dụng để giải mã dữ liệu video.

Để thực hiện CABAC, bộ mã hóa video 20 có thể gán ngữ cảnh trong mẫu ngữ cảnh cho ký hiệu được truyền. Ngữ cảnh có thể liên quan đến việc, ví dụ, các giá trị lân cận của ký hiệu này có khác không hay không. Để thực hiện CABAC, bộ mã hóa video 20 có thể chọn mã chiều dài thay đổi cho một ký hiệu được truyền. Từ mã trong phương pháp mã hóa chiều dài thay đổi (Variable length coding-VLC) có thể được tạo ra để các mã tương đối ngắn hơn tương ứng với các ký hiệu xác suất cao hơn, còn mã dài hơn tương ứng với các ký hiệu xác suất thấp hơn. Theo cách này, việc sử dụng VLC có thể đạt được sự tiết kiệm bit so với, ví dụ, sử dụng từ mã chiều dài bằng nhau đối với một ký hiệu được truyền. Việc xác định xác suất có thể dựa trên ngữ cảnh được gán cho ký hiệu này.

Bộ mã hóa video 20 có thể kết xuất dòng bit bao gồm chuỗi các bit tạo ra sự biểu diễn các hình ảnh được mã hóa và dữ liệu liên quan. Thuật ngữ “dòng bit” có thể là thuật ngữ chung được sử dụng để chỉ dòng đơn vị lớp trừu tượng hóa mạng (Network Abstraction Layer-NAL) (ví dụ, chuỗi đơn vị NAL) hoặc dòng byte (ví dụ, đóng gói dòng đơn vị NAL chứa tiền tố mã bắt đầu và đơn vị NAL như được thể hiện bởi phụ lục B của chuẩn HEVC). Đơn vị NAL là cấu trúc cú pháp chứa chỉ báo về kiểu dữ liệu trong đơn vị NAL và byte chứa dữ liệu ở dạng chuỗi byte thô (Raw byte

sequence payload-RBSP) được đặt rải rác khi cần với các bit ngăn chặn sự mô phỏng. Mỗi đơn vị NAL bao gồm tiêu đề đơn vị NAL và có thể đóng gói một RBSP. Tiêu đề đơn vị NAL có thể bao gồm phần tử cú pháp chỉ báo mã kiểu đơn vị NAL. Mã kiểu đơn vị NAL được chỉ rõ bởi tiêu đề đơn vị NAL của đơn vị NAL chỉ báo kiểu đơn vị NAL. RBSP có thể là cấu trúc cú pháp chứa số nguyên byte được đóng gói trong đơn vị NAL. Trong một số trường hợp, RBSP bao gồm các bit 0.

Các kiểu đơn vị NAL khác nhau có thể đóng gói các loại RBSP khác nhau. Ví dụ, kiểu đơn vị NAL thứ nhất có thể đóng gói một RBSP cho một tập tham số ảnh (picture parameter set - PPS), kiểu đơn vị NAL thứ hai có thể đóng gói một RBSP cho đoạn lát, kiểu đơn vị NAL thứ ba có thể đóng gói một RBSP cho thông tin nâng cao bổ sung (supplemental enhancement information - SEI), v.v. Đơn vị NAL đóng gói RBSP cho dữ liệu mã hóa video (như trái ngược với RBSP cho tập tham số và thông báo SEI) có thể được gọi là đơn vị NAL lớp mã hóa video (Video coding layer-VCL). Đơn vị NAL chứa tập tham số (ví dụ, tập tham số video (video parameter set-VPS), tập tham số chuỗi (sequence parameter set-SPS, PPS, v.v..) có thể được gọi là đơn vị NAL tập tham số.

Bản mô tả này đề cập đến đơn vị NAL đóng gói RBSP cho đoạn lát dưới dạng đơn vị NAL lát mã hóa. Như được định nghĩa trong HEVC WD, đoạn lát là một số nguyên các CTU được sắp xếp liên tiếp trong quét khung ô và được chứa trong một đơn vị NAL. Ngược lại, theo HEVC WD, một lát có thể là số nguyên các CTU có trong một đoạn lát độc lập và tất cả các đoạn lát phụ thuộc kế tiếp (nếu có) đứng trước đoạn lát độc lập tiếp theo (nếu có) trong cùng một đơn vị truy cập. Đoạn lát độc lập là đoạn lát mà các giá trị của phần tử cú pháp của tiêu đề đoạn lát không được suy ra từ các giá trị cho đoạn lát đứng trước. Đoạn lát phụ thuộc là đoạn lát mà các giá trị của một số phần tử cú pháp của tiêu đề đoạn lát được suy ra từ các giá trị cho đoạn lát độc lập đứng trước theo thứ tự giải mã. RBSP của đơn vị NAL đã mã hóa có thể bao gồm tiêu đề đoạn lát và dữ liệu lát. Tiêu đề đoạn lát là một phần của lát đã mã hóa chứa các phần tử dữ liệu liên quan đến CTU thứ nhất hoặc tất cả các CTU được biểu diễn trong đoạn lát. Tiêu đề lát là tiêu đề đoạn lát của đoạn lát độc lập là đoạn lát hiện thời hoặc đoạn lát độc lập gần nhất đứng trước đoạn lát phụ thuộc hiện thời theo thứ tự giải mã.

VPS là cấu trúc cú pháp bao gồm các phần tử cú pháp áp dụng cho không hoặc nhiều chuỗi video được mã hóa (Coded video sequence-CVS) toàn vẹn. SPS là cấu

trúc cú pháp bao gồm các phần tử cú pháp áp dụng cho không hoặc nhiều CVS toàn vẹn. SPS có thể bao gồm phần tử cú pháp nhận dạng VPS hoạt động khi SPS hoạt động. Do vậy, các phần tử cú pháp của VPS có thể có khả năng áp dụng một cách rộng rãi hơn các phần tử cú pháp của SPS.

Tập tham số (ví dụ, VPS, SPS, PPS, v.v.) có thể chứa nhận dạng được tham chiếu, trực tiếp hoặc gián tiếp, từ tiêu đề lát của lát. Quy trình tham chiếu được biết đến là "sự kích hoạt". Do vậy, khi bộ giải mã video 30 đang giải mã một lát cụ thể, tập tham số được tham chiếu, trực tiếp hoặc gián tiếp, bằng phần tử cú pháp trong tiêu đề lát của lát cụ thể được cho là "được kích hoạt". Tùy thuộc vào kiểu tập tham số, sự kích hoạt có thể xảy ra trên cơ sở hình ảnh hoặc trên cơ sở chuỗi. Ví dụ, tiêu đề của lát có thể bao gồm phần tử cú pháp nhận dạng PPS. Do vậy, khi bộ mã hóa video mã hóa lát, PPS được hoạt hóa. Hơn nữa, PPS có thể bao gồm phần tử cú pháp nhận dạng SPS. Do vậy, khi PPS nhận dạng SPS được kích hoạt, SPS có thể được kích hoạt. SPS có thể bao gồm một phần tử cú pháp nhận dạng VPS. Do vậy, khi SPS nhận dạng VPS được kích hoạt, VPS được kích hoạt.

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

Theo HEVC WD, CVS có thể bắt đầu từ hình ảnh làm mới giải mã tức thời (Instantaneous Decoding Refresh-IDR), hoặc hình ảnh truy cập liên kết gãy (Broken link access-BLA) hoặc hình ảnh truy cập ngẫu nhiên sạch (clean random access -CRA) là hình ảnh đầu tiên trong dòng bit, bao gồm tất cả các hình ảnh tiếp theo không phải

ảnh IDR hoặc BLA. Hình ảnh IDR chỉ chứa lát I (tức là, các lát trong đó chỉ sử dụng dự đoán trong ảnh). Hình ảnh IDR có thể là hình ảnh đầu tiên trong dòng bit theo thứ tự mã hóa hoặc có thể xuất hiện muộn hơn trong dòng bit. Mỗi hình ảnh IDR là hình ảnh đầu tiên của CVS theo thứ tự giải mã. Theo HEVC WD, hình ảnh IDR có thể là hình ảnh điểm truy cập ngẫu nhiên (Intra random access point (IRAP) trong ảnh mà đối với hình ảnh này, mỗi đơn vị NAL VCL có `nal_unit_type` bằng `IDR_W_RADL` hoặc `IDR_N_LP`.

Các hình ảnh IDR có thể được sử dụng để truy cập ngẫu nhiên. Tuy nhiên, các hình ảnh đứng sau hình ảnh IDR theo thứ tự giải mã không thể sử dụng các hình ảnh được giải mã trước hình ảnh IDR làm tham chiếu. Do đó, dòng bit dựa vào các hình ảnh IDR để truy cập ngẫu nhiên có hiệu suất mã hóa thấp hơn đáng kể so với dòng bit sử dụng các loại hình ảnh truy cập ngẫu nhiên khác. Theo ít nhất một số ví dụ, đơn vị truy cập IDR là đơn vị truy cập chứa hình ảnh IDR.

Khái niệm về hình ảnh CRA được đưa vào HEVC để cho phép các hình ảnh đứng sau hình ảnh CRA theo thứ tự giải mã, nhưng đứng trước hình ảnh CRA theo thứ tự kết xuất, sử dụng các hình ảnh đã được giải mã trước hình ảnh CRA để tham chiếu. Các hình ảnh đứng sau hình ảnh CRA theo thứ tự giải mã, nhưng đứng trước hình ảnh CRA theo thứ tự kết xuất, được gọi là các hình ảnh hướng dẫn đối với hình ảnh CRA (hoặc các hình ảnh hướng dẫn của hình ảnh CRA). Tức là, để cải thiện hiệu quả mã hóa, khái niệm về hình ảnh CRA được đưa vào HEVC để cho phép các hình ảnh đứng sau hình ảnh CRA theo thứ tự giải mã, nhưng đứng trước hình ảnh CRA theo thứ tự kết xuất, sử dụng các hình ảnh đã được giải mã trước hình ảnh CRA để tham chiếu. Điểm truy cập CRA là điểm truy cập trong đó hình ảnh được mã hóa là hình ảnh CRA. Theo HEVC WD, hình ảnh CRA là hình ảnh truy cập ngẫu nhiên trong ảnh mà đối với hình ảnh này, mỗi đơn vị NAL VCL có `nal_unit_type` bằng `CRA_NUT`.

Hình ảnh hướng dẫn của hình ảnh CRA có thể được giải mã chính xác nếu việc giải mã bắt đầu từ hình ảnh IDR hoặc hình ảnh CRA xảy ra trước hình ảnh CRA theo thứ tự giải mã. Tuy nhiên, hình ảnh hướng dẫn của hình ảnh CRA có thể không được giải mã khi việc truy cập ngẫu nhiên từ hình ảnh CRA xảy ra. Do vậy, bộ giải mã video thường giải mã các hình ảnh hướng dẫn của hình ảnh CRA trong quá trình giải mã truy cập ngẫu nhiên. Để ngăn chặn sự lan truyền lỗi từ các hình ảnh tham chiếu mà có thể là không khả dụng tùy thuộc vào quá trình giải mã bắt đầu ở đâu, không có hình

ảnh nào đứng sau hình ảnh CRA theo cả thứ tự giải mã lẫn thứ tự kết xuất có thể sử dụng hình ảnh bất kỳ đứng trước hình ảnh CRA theo thứ tự giải mã hoặc thứ tự kết xuất (bao gồm các hình ảnh hướng dẫn) để tham chiếu.

Khái niệm về hình ảnh BLA được đưa vào HEVC sau khi đưa các hình ảnh CRA và dựa trên khái niệm về hình ảnh CRA. Hình ảnh BLA thường bắt đầu từ dòng bit ghép nối ở vị trí hình ảnh CRA và trong dòng bit đã ghép nối, hình ảnh CRA tại điểm ghép nối được biến đổi thành hình ảnh BLA. Do vậy, hình ảnh BLA có thể là hình ảnh CRA trong dòng bit ban đầu và hình ảnh CRA được biến đổi thành hình ảnh BLA bởi bộ ghép nối dòng bit sau khi ghép nối dòng bit ở vị trí hình ảnh CRA. Theo một số trường hợp, đơn vị truy cập chứa hình ảnh RAP có thể được gọi là đơn vị truy cập RAP. Đơn vị truy cập BLA là đơn vị truy cập chứa hình ảnh BLA. Theo HEVC WD, hình ảnh BLA có thể là hình ảnh truy cập ngẫu nhiên trong ảnh mà đối với hình ảnh này, mỗi đơn vị NAL VCL có `nal_unit_type` bằng `BLA_W_LP`, `BLA_W_RADL`, hoặc `BLA_N_LP`.

Nói chung, hình ảnh IRAP chỉ chứa các lát I và có thể là hình ảnh BLA, hình ảnh CRA hoặc hình ảnh IDR. Chẳng hạn, HEVC WD chỉ ra rằng hình ảnh IRAP có thể là hình ảnh được mã hóa mà đối với hình ảnh này mỗi đơn vị NAL VCL có `nal_unit_type` nằm trong khoảng từ `BLA_W_LP` đến `RSV_IRAP_VCL23`. Hơn nữa, HEVC WD chỉ ra rằng hình ảnh đầu tiên trong dòng bit theo thứ tự giải mã phải là hình ảnh IRAP. Bảng 7-1 của HEVC WD thể hiện các mã kiểu đơn vị NAL và nhóm kiểu đơn vị NAL. Bảng 7-1 của HEVC WD được đưa ra dưới đây.

Bảng 7-1 – Các mã kiểu đơn vị NAL và nhóm kiểu đơn vị NAL

| <code>nal_unit_type</code> | Tên của <code>nal_unit_type</code> | Nội dung của đơn vị NAL và cấu trúc cú pháp RBSP                                                  | Nhóm kiểu đơn vị NAL |
|----------------------------|------------------------------------|---------------------------------------------------------------------------------------------------|----------------------|
| 0<br>1                     | TRAIL_N<br>TRAIL_R                 | Đoạn lát mã hóa của hình ảnh trễ không TSA, không STSA<br><code>slice_segment_layer_rbsp()</code> | VCL                  |
| 2<br>3                     | TSA_N<br>TSA_R                     | Đoạn lát mã hóa của hình ảnh TSA<br><code>slice_segment_layer_rbsp()</code>                       | VCL                  |
| 45.                        | STSA_N<br>STSA_R                   | Đoạn lát mã hóa của hình ảnh STSA                                                                 | VCL                  |

|        |                          |                                                                    |           |
|--------|--------------------------|--------------------------------------------------------------------|-----------|
|        |                          | slice_segment_layer_rbsp( )                                        |           |
| 6      | RADL_N                   | Đoạn lát mã hóa của hình ảnh                                       | VCL       |
| 7      | RADL_R                   | RASL<br>slice_segment_layer_rbsp( )                                |           |
| 8      | RASL_N                   | Đoạn lát mã hóa của hình ảnh                                       | VCL       |
| 9      | RASL_R                   | RASL<br>slice_segment_layer_rbsp( )                                |           |
| 10     | RSV_VCL_N10              | Kiểu đơn vị NAL VCL không                                          | VCL       |
| 12     | RSV_VCL_N12              | tham chiếu của lớp con không                                       |           |
| 14)    | RSV_VCL_N14              | IRAP dự trữ                                                        |           |
| 11     | RSV_VCL_R11              | Kiểu đơn vị NAL VCL tham                                           | VCL       |
| 13     | RSV_VCL_R13              | chiếu của lớp con không IRAP dự                                    |           |
| 15)    | RSV_VCL_R15              | trữ                                                                |           |
| 16     | BLA_W_LP                 | Đoạn lát mã hóa của hình ảnh                                       | VCL       |
| 17     | BLA_W_RADL               | BLA                                                                |           |
| 18)    | BLA_N_LP                 | slice_segment_layer_rbsp( )                                        |           |
| 19     | IDR_W_RADL               | Đoạn lát mã hóa của hình ảnh                                       | VCL       |
| 20     | IDR_N_LP                 | IDR<br>slice_segment_layer_rbsp( )                                 |           |
| 21     | CRA_NUT                  | Đoạn lát mã hóa của hình ảnh<br>CRA<br>slice_segment_layer_rbsp( ) | VCL       |
| 22     | RSV_IRAP_VCL             | Kiểu đơn vị NAL IRAP VCL dự                                        | VCL       |
| 23     | 22<br>RSV_IRAP_VCL<br>23 | trữ                                                                |           |
| 24..31 | RSV_VCL24..<br>RSV_VCL31 | Kiểu đơn vị NAL VCL không<br>IRAP dự trữ                           | VCL       |
| 32     | VPS_NUT                  | Tập tham số video<br>video_parameter_set_rbsp( )                   | không VCL |
| 33     | SPS_NUT                  | Tập tham số chuỗi                                                  | không VCL |

|          |                                          |                                                           |           |
|----------|------------------------------------------|-----------------------------------------------------------|-----------|
|          |                                          | seq_parameter_set_rbsp( )                                 |           |
| 34       | PPS_NUT                                  | Tập tham số ảnh<br>pic_parameter_set_rbsp( )              | không VCL |
| 35       | AUD_NUT                                  | Dấu tách đơn vị truy cập<br>access_unit_delimiter_rbsp( ) | không VCL |
| 36       | EOS_NUT                                  | Điểm cuối của chuỗi<br>end_of_seq_rbsp( )                 | không VCL |
| 37       | EOB_NUT                                  | Đầu cuối của dòng bit<br>end_of_bitstream_rbsp( )         | không VCL |
| 38       | FD_NUT                                   | Dữ liệu đệm<br>filler_data_rbsp( )                        | không VCL |
| 39<br>40 | PREFIX_SEI_N<br>UT<br>SUFFIX_SEI_N<br>UT | Thông tin nâng cao bổ sung<br>sei_rbsp( )                 | không VCL |
| 41..47   | RSV_NVCL41..<br>RSV_NVCL47               | Dự trữ                                                    | không VCL |
| 48..63   | UNSPEC48..<br>UNSPEC63                   | Không xác định                                            | không VCL |

Một sự khác nhau giữa hình ảnh BLA và hình ảnh CRA là như sau. Đối với hình ảnh CRA, hình ảnh hướng dẫn liên quan có thể được giải mã chính xác nếu việc giải mã bắt đầu từ hình ảnh RAP đứng trước hình ảnh CRA theo thứ tự giải mã. Tuy nhiên, hình ảnh hướng dẫn liên quan đến hình ảnh CRA không thể được giải mã chính xác khi xảy ra sự truy cập ngẫu nhiên từ hình ảnh CRA này (tức là, khi việc giải mã bắt đầu từ hình ảnh CRA hoặc nói cách khác, khi hình ảnh CRA là hình ảnh đầu tiên trong dòng bit). Ngược lại, không có trường hợp nào trong đó hình ảnh hướng dẫn liên quan đến hình ảnh BLA có thể giải mã được, ngay cả khi việc giải mã bắt đầu từ hình ảnh RAP đứng trước hình ảnh BLA theo thứ tự giải mã.

Một số hình ảnh hướng dẫn liên quan đến hình ảnh CRA cụ thể hoặc hình ảnh BLA cụ thể có thể được giải mã chính xác ngay cả khi hình ảnh CRA hoặc hình ảnh BLA cụ thể là hình ảnh đầu tiên trong dòng bit. Các hình ảnh hướng dẫn này có thể

được gọi là hình ảnh hướng dẫn giải mã được (decodable leading picture-DLP) hoặc hình ảnh hướng dẫn giải mã truy cập ngẫu nhiên (Random Access Decodable Leading-RADL). Theo HEVC WD, hình ảnh RADL có thể là hình ảnh được mã hóa mà đối với hình ảnh này, mỗi đơn vị NAL VCL có `nal_unit_type` bằng `RADL_R` hoặc `RADL_N`. Hơn nữa, HEVC WD chỉ ra rằng tất cả các hình ảnh RADL là hình ảnh hướng dẫn và rằng các hình ảnh RADL không được sử dụng làm hình ảnh tham chiếu để giải mã các hình ảnh hướng dẫn của cùng một hình ảnh IRAP liên quan. Khi có mặt, tất cả các hình ảnh RADL đứng trước, theo thứ tự giải mã, tất cả các hình ảnh trễ của cùng một hình ảnh IRAP liên quan. HEVC WD chỉ ra rằng đơn vị truy cập RADL có thể là đơn vị truy cập trong đó hình ảnh được mã hóa là hình ảnh RADL. Hình ảnh trễ có thể là hình ảnh đứng sau hình ảnh IRAP liên quan (tức là, hình ảnh IRAP trước đó theo thứ tự giải mã) theo thứ tự kết xuất.

Các hình ảnh hướng dẫn khác có thể được gọi là hình ảnh hướng dẫn không giải mã được (non-decodable leading picture-NLP) hoặc hình ảnh hướng dẫn bỏ qua truy cập ngẫu nhiên (Random Access Skipped Leading-RSDL). Theo HEVC WD, hình ảnh RSDL có thể là hình ảnh được mã hóa mà đối với hình ảnh này, mỗi đơn vị NAL VCL có `nal_unit_type` bằng `RASL_R` hoặc `RASL_N`. Tất cả các hình ảnh RASL là hình ảnh hướng dẫn của hình ảnh BLA hoặc CRA liên quan.

Với điều kiện là tập tham số cần thiết là khả dụng khi chúng cần được kích hoạt, hình ảnh IRAP và tất cả các hình ảnh không RASL tiếp theo theo thứ tự giải mã có thể được giải mã chính xác mà không cần thực hiện quá trình giải mã hình ảnh bất kỳ đứng trước hình ảnh IRAP theo thứ tự giải mã. Có thể có các hình ảnh trong dòng bit chỉ chứa lát I không phải là hình ảnh IRAP.

Theo kỹ thuật mã hóa đa khung hình, có thể có nhiều khung hình của cùng một màn hình theo các góc nhìn khác nhau. Thuật ngữ “đơn vị truy cập” có thể được sử dụng để chỉ tập các hình ảnh tương ứng với cùng một thời điểm thời gian. Do vậy, dữ liệu video có thể khái niệm hóa là một loạt các đơn vị truy cập xảy ra trong một khoảng thời gian. “Thành phần khung hình” có thể là dạng biểu diễn mã hóa của khung hình trong một đơn vị truy cập đơn lẻ. Trong bản mô tả này, “khung hình” có thể là chuỗi hoặc tập thành phần khung hình có cùng một ký hiệu nhận dạng khung hình. Thành phần khung hình có thể chứa thành phần khung hình kết cấu và thành phần khung hình chiều sâu. Trong bản mô tả này, “khung hình” có thể là chuỗi hoặc

tập gồm một hoặc nhiều thành phần khung hình có cùng một ký hiệu nhận dạng khung hình.

Thành phần khung hình kết cấu (tức là, hình ảnh kết cấu) có thể là dạng biểu diễn mã hóa của kết cấu của khung hình trong một đơn vị truy cập. Khung hình kết cấu có thể là chuỗi các thành phần khung hình kết cấu có giá trị chỉ số thứ tự khung hình giống nhau. Chỉ số thứ tự khung hình của một khung hình có thể chỉ báo vị trí camera của khung hình này so với các khung hình khác. Thành phần khung hình chiều sâu (tức là, hình ảnh chiều sâu) có thể là dạng biểu diễn mã hóa chiều sâu của khung hình trong một đơn vị truy cập đơn lẻ. Khung hình chiều sâu có thể là tập hoặc chuỗi gồm một hoặc nhiều các thành phần khung hình chiều sâu có giá trị chỉ số thứ tự khung hình giống nhau.

Theo MV-HEVC, 3D-HEVC và SHVC, bộ mã hóa video có thể tạo ra dòng bit bao gồm một dãy các đơn vị NAL. Các đơn vị NAL khác nhau của dòng bit có thể được kết hợp với các lớp khác nhau của dòng bit. Một lớp có thể được định nghĩa là tập đơn vị NAL VCL và đơn vị NAL không VCL liên quan có cùng ký hiệu nhận dạng lớp. Một lớp có thể tương đương với một khung hình trong mã hóa video đa khung hình. Theo mã hóa video đa khung hình, một lớp có thể chứa tất cả các thành phần khung hình của cùng một lớp với các thời điểm khác nhau. Mỗi thành phần khung hình có thể là hình ảnh được mã hóa của màn hình video thuộc một khung hình cụ thể ở một thời điểm cụ thể. Theo một số ví dụ về mã hóa video 3D, một lớp có thể chứa tất cả các hình ảnh chiều sâu mã hóa của một khung hình cụ thể hoặc các hình ảnh kết cấu mã hóa của một khung hình cụ thể. Theo ví dụ khác về mã hóa video 3D, một lớp có thể chứa cả thành phần khung hình kết cấu và thành phần khung hình chiều sâu của một khung hình cụ thể. Tương tự, trong trường hợp mã hóa video có thể mở rộng được, một lớp thường tương ứng với các hình ảnh được mã hóa có các đặc điểm video khác với hình ảnh được mã hóa trong lớp khác. Các đặc điểm video này thường bao gồm độ phân giải không gian và mức chất lượng (ví dụ, tỷ lệ tín hiệu so với tiếng ồn). Theo HEVC và các phiên bản mở rộng của nó, khả năng mở rộng về thời gian có thể đạt được trong một lớp bằng cách xác định nhóm ảnh có mức thời gian cụ thể làm lớp con.

Đối với mỗi lớp tương ứng của dòng bit, dữ liệu ở lớp thấp có thể được giải mã mà không cần tham chiếu đến dữ liệu ở lớp cao. Theo phương pháp mã hóa video có

thể mở rộng được, ví dụ dữ liệu ở lớp cơ sở có thể được giải mã mà không cần tham chiếu đến dữ liệu ở lớp nâng cao. Nói chung, đơn vị NAL chỉ có thể đóng gói dữ liệu của một lớp đơn. Do vậy, đơn vị NAL đóng gói dữ liệu của lớp còn lại cao nhất của dòng bit có thể được loại bỏ ra khỏi dòng bit mà không làm ảnh hưởng đến khả năng giải mã dữ liệu ở các lớp còn lại của dòng bit. Theo kỹ thuật mã hóa đa khung hình và 3D-HEVC, các lớp cao có thể bao gồm các thành phần khung hình bổ sung. Theo SHVC, các lớp cao có thể bao gồm dữ liệu nâng cao tỷ lệ tín hiệu so với tiếng ồn (signal to noise ratio-SNR), dữ liệu nâng cao về không gian và/hoặc dữ liệu nâng cao về thời gian. Theo MV-HEVC, 3D-HEVC và SHVC, một lớp có thể được gọi là "lớp cơ sở" nếu bộ giải mã video có thể giải mã các hình ảnh trong lớp này mà không cần tham chiếu đến dữ liệu của lớp bất kỳ khác. Lớp cơ sở có thể theo tiêu chuẩn cơ bản của HEVC (ví dụ, HEVC WD).

Theo SVC, các lớp không phải lớp cơ sở có thể được gọi là "lớp nâng cao" và có thể cung cấp thông tin làm tăng chất lượng trực quan của dữ liệu video được giải mã từ dòng bit. SVC có thể làm tăng độ phân giải không gian, tỷ lệ tín hiệu so với tiếng ồn (tức là, chất lượng) hoặc tốc độ thời gian. Theo kỹ thuật mã hóa video có thể mở rộng (ví dụ, SHVC), "dạng biểu diễn lớp" có thể là dạng biểu diễn mã hóa lớp không gian trong một đơn vị truy cập đơn lẻ. Để giải thích, bản mô tả này có thể đề cập đến các thành phần khung hình và/hoặc dạng biểu diễn lớp dưới dạng "thành phần khung hình/dạng biểu diễn lớp" hoặc đơn giản là "các hình ảnh".

Để thực hiện các lớp, tiêu đề của đơn vị NAL có thể bao gồm các phần tử cú pháp `nuh_reserved_zero_6bits`. Theo HEVC WD, phần tử cú pháp `nuh_reserved_zero_6bits` được dự trữ. Tuy nhiên, theo MV-HEVC, 3D-HEVC, và SVC, phần tử cú pháp `nuh_reserved_zero_6bits` được gọi là phần tử cú pháp `nuh_layer_id`. Phần tử cú pháp `nuh_layer_id` mô tả ký hiệu nhận dạng của một lớp. Các đơn vị NAL của một dòng bit có phần tử cú pháp `nuh_layer_id` mô tả các giá trị khác nhau thuộc các lớp khác nhau của dòng bit.

Theo một số ví dụ, phần tử cú pháp `nuh_layer_id` của đơn vị NAL bằng 0 nếu đơn vị NAL liên quan đến lớp cơ sở trong mã hóa đa khung hình (ví dụ, MV-HEVC), mã hóa 3DV (ví dụ, 3D-HEVC), hoặc mã hóa video mở rộng được (ví dụ, SHVC). Dữ liệu ở lớp cơ sở của dòng bit có thể được giải mã mà không cần tham chiếu đến dữ liệu ở lớp bất kỳ khác của dòng bit. Nếu đơn vị NAL không liên quan đến lớp cơ sở trong

mã hóa đa khung hình, 3DV, hoặc mã hóa video mở rộng được, phần tử cú pháp `nuh_layer_id` của đơn vị NAL có thể có giá trị khác không.

Hơn nữa, một số thành phần khung hình/dạng biểu diễn lớp trong một lớp có thể được giải mã mà không cần tham chiếu đến thành phần khung hình/dạng biểu diễn lớp khác trong cùng một lớp. Do vậy, đơn vị NAL đóng gói dữ liệu của một số thành phần khung hình/dạng biểu diễn lớp của một lớp có thể được loại bỏ ra khỏi dòng bit mà không làm hình ảnh hưởng đến khả năng giải mã các thành phần khung hình/dạng biểu diễn lớp khác trong lớp này. Việc loại bỏ các đơn vị NAL đóng gói dữ liệu của các thành phần khung hình/dạng biểu diễn lớp này có thể làm giảm tốc độ khung hình của dòng bit. Tập con của các thành phần khung hình/dạng biểu diễn lớp trong một lớp có thể được giải mã mà không cần tham chiếu đến các thành phần khung hình/dạng biểu diễn lớp khác trong lớp có thể được gọi là "lớp con" hoặc "lớp con thời gian".

Các đơn vị NAL có thể bao gồm các phần tử cú pháp `temporal_id` mô tả các ký hiệu nhận dạng thời gian (tức là, các `TemporalId`) của các đơn vị NAL. Ký hiệu nhận dạng thời gian của đơn vị NAL nhận dạng lớp con mà đơn vị NAL thuộc lớp này. Do vậy, mỗi lớp con của dòng bit có thể có ký hiệu nhận dạng khác nhau. Nói chung, nếu ký hiệu nhận dạng của đơn vị NAL đầu tiên của một lớp nhỏ hơn ký hiệu nhận dạng của đơn vị NAL thứ hai của cùng một lớp, dữ liệu được đóng gói bởi đơn vị NAL thứ nhất có thể được giải mã mà không cần tham chiếu đến dữ liệu được đóng gói bởi đơn vị NAL thứ hai.

Một dòng bit có thể kết hợp với nhiều điểm hoạt động. Mỗi điểm hoạt động của dòng bit được kết hợp với tập ký hiệu nhận dạng lớp (ví dụ, tập các giá trị `nuh_layer_id`) và ký hiệu nhận dạng thời gian. Tập các ký hiệu nhận dạng lớp có thể được thể hiện là `OpLayerIdSet` và ký hiệu nhận dạng thời gian có thể được thể hiện là `TemporalID`. Nếu ký hiệu nhận dạng lớp của một đơn vị nằm trong tập các ký hiệu nhận dạng lớp của điểm hoạt động và ký hiệu nhận dạng thời gian của đơn vị NAL nhỏ hơn hoặc bằng ký hiệu nhận dạng thời gian của điểm hoạt động, đơn vị NAL được kết hợp với điểm hoạt động. Do vậy, điểm hoạt động có thể tương đương với tập con các đơn vị NAL trong dòng bit.

Như được giới thiệu ở trên, bản mô tả này đề cập đến việc lưu trữ nội dung video trong một tập tin dựa trên định dạng tập tin phương tiện cơ sở ISO (ISO/BMFF).

Cụ thể, bản mô tả này mô tả các kỹ thuật để lưu trữ các dòng video chứa nhiều lớp mã hóa, trong đó mỗi lớp có thể là lớp mở rộng được, khung hình kết cấu, khung hình chiều sâu hoặc các loại lớp hoặc khung hình khác. Các kỹ thuật của bản mô tả này có thể được áp dụng để, ví dụ, lưu trữ dữ liệu video MV-HEVC, dữ liệu video SHVC, dữ liệu video 3D-HEVC, và/hoặc các loại dữ liệu video khác.

Các định dạng của tập tin và các chuẩn định dạng tập tin sẽ được mô tả một cách vắn tắt. Các chuẩn định dạng tập tin bao gồm định dạng tập tin phương tiện cơ sở ISO (ISO/BMFF, ISO/IEC 14496-12, dưới đây, “ISO/IEC 14996-12”) và các chuẩn định dạng tập tin khác nhận được từ ISO/BMFF, bao gồm định dạng tập tin MPEG-4 (ISO/IEC 14496-14), định dạng tập tin 3GPP (3GPP TS 26.244) và định dạng tập tin AVC (ISO/IEC 14496-15, dưới đây là “ISO/IEC 14996-15”). Do vậy, ISO/IEC 14496-12 mô tả định dạng tập tin phương tiện cơ sở ISO. Các tài liệu khác mở rộng định dạng tập tin phương tiện cơ sở ISO cho các ứng dụng cụ thể. Ví dụ, ISO/IEC 14496-15 mô tả việc lưu trữ video được tạo cấu trúc bởi đơn vị NAL theo định dạng tập tin phương tiện cơ sở ISO. H.264/AVC và HEVC, cũng như các phiên bản mở rộng của chúng, là các ví dụ về video được tạo cấu trúc bởi đơn vị NAL. ISO/IEC 14496-15 bao gồm các phần mô tả việc mang các đơn vị NAL H.264/AVC. Ngoài ra, phần 8 của ISO/IEC 14496-15 mô tả việc mang các đơn vị NAL HEVC.

ISO/BMFF được sử dụng làm cơ sở cho nhiều định dạng đóng gói của bộ mã hóa-giải mã, như định dạng tập tin AVC, cũng như cho nhiều các định dạng lưu trữ đa phương tiện như định dạng tập tin MPEG-4, định dạng tập tin 3GPP (3GP), và định dạng tập tin DVB. Ngoài việc truyền thông liên tục, như âm thanh và video, truyền thông tĩnh như hình ảnh cũng như siêu dữ liệu, có thể được lưu trữ trong một tập tin theo chuẩn ISO/BMFF. Các tập tin được tạo cấu trúc theo chuẩn ISO/BMFF có thể được sử dụng cho nhiều mục đích, bao gồm phát lại tập tin phương tiện cục bộ, tải liên tục tập tin từ xa, các đoạn để truyền liên tục tương thích động theo HTTP (DASH), bộ lưu trữ cho nội dung được truyền liên tục và các lệnh đóng gói của nó và ghi lại các dòng phương tiện thời gian thực nhận được. Do vậy, mặc dù ban đầu được thiết kế để lưu trữ, ISO/BMFF đã tỏ ra có giá trị đối với việc truyền liên tục, ví dụ, để tải liên tục hoặc DASH. Đối với mục đích truyền liên tục, các đoạn phim được xác định theo ISO/BMFF có thể được sử dụng.

Tập tin theo định dạng tập tin HEVC có thể bao gồm một loạt các đối tượng, được gọi là hộp. Một hộp có thể là khối kiến trúc hướng đến đối tượng được xác định bởi ký hiệu nhận dạng loại duy nhất và chiều dài. Chẳng hạn, một hộp có thể là cấu trúc phần tử cú pháp theo ISOBMFF, bao gồm hộp được mã hóa bởi bốn đặc điểm, tổng số byte của hộp này và tải tin. Nói cách khác, một hộp có thể là cấu trúc cú pháp bao gồm loại hộp được mã hóa, tổng số byte của hộp này và tải tin. Theo một số trường hợp, tất cả các dữ liệu trong một tập tin theo định dạng tập tin HEVC có thể được chứa trong các hộp và có thể không có dữ liệu nào trong tập tin không ở trong hộp. Do vậy, tập tin ISOBMFF có thể bao gồm một chuỗi các hộp và các hộp có thể chứa các hộp khác. Chẳng hạn, tải tin của hộp có thể bao gồm một hoặc nhiều hộp khác. Fig.5 và Fig.6 mô tả chi tiết trong bản mô tả này, thể hiện các hộp ví dụ trong tập tin, theo một hoặc nhiều kỹ thuật của bản mô tả này.

Tập tin theo ISOBMFF có thể bao gồm các loại hộp khác nhau. Ví dụ, tập tin theo ISOBMFF có thể bao gồm hộp kiểu tập tin, hộp dữ liệu phương tiện, hộp phim ảnh, hộp đoạn phim, v.v.. Theo ví dụ này, hộp loại tập tin bao gồm thông tin về kiểu tập tin và thông tin về tính tương thích. Hộp dữ liệu phương tiện có thể chứa các mẫu (ví dụ, các hình ảnh được mã hóa). Hộp phim (“moov”) chứa siêu dữ liệu cho các dòng phương tiện liên tục có trong tập tin. Mỗi dòng phương tiện liên tục này có thể được thể hiện lại trong tập tin dưới dạng rãnh. Chẳng hạn, hộp phim có thể chứa siêu dữ liệu về phim (ví dụ, mối liên hệ logic và định thời giữa các mẫu và cũng về các con trỏ đến vị trí của các mẫu). Hộp phim có thể bao gồm một số kiểu hộp con. Các hộp con trong một hộp phim có thể bao gồm một hoặc nhiều hộp rãnh. Hộp rãnh có thể bao gồm thông tin về một rãnh cụ thể của phim. Hộp rãnh có thể bao gồm tiêu đề của rãnh mô tả thông tin chung của một rãnh đơn lẻ. Ngoài ra, hộp rãnh có thể bao gồm hộp phương tiện chứa hộp thông tin phương tiện. Hộp thông tin phương tiện có thể bao gồm hộp bảng mẫu chứa dữ liệu chỉ ra các mẫu phương tiện trong rãnh. Thông tin trong hộp bảng mẫu có thể được sử dụng để định vị mẫu theo thời gian và cho mỗi mẫu của rãnh, kiểu, kích thước và bộ lưu trữ và độ dịch chuyển vào bộ lưu trữ đó của mẫu. Do vậy, siêu dữ liệu cho một rãnh có thể được chứa trong hộp rãnh (“trak”), còn nội dung phương tiện của rãnh này được chứa trong hộp dữ liệu phương tiện (Media Data box-mdat) hoặc trực tiếp trong một tập tin riêng. Nội dung phương tiện đối với

các rãnh bao gồm (ví dụ, gồm có) chuỗi các mẫu, như đơn vị truy cập âm thanh hoặc video.

ISOBMFF mô tả các kiểu rãnh sau: rãnh phương tiện, chứa một dòng phương tiện cơ bản, rãnh gợi ý, bao gồm các lệnh truyền phương tiện hoặc là dòng gói nhận và rãnh siêu dữ liệu được định thời, bao gồm siêu dữ liệu được đồng bộ hóa theo thời gian. Siêu dữ liệu cho mỗi rãnh bao gồm một danh sách các mục nhập mô tả mẫu, mỗi mục nhập cung cấp định dạng mã hóa hoặc đóng gói được sử dụng ở rãnh này và dữ liệu khởi tạo cần thiết cho việc xử lý định dạng này. Mỗi mẫu có liên quan đến một trong số các mục nhập mô tả mẫu của rãnh.

ISOBMFF có thể mô tả siêu dữ liệu đặc trưng cho mẫu bằng các cơ chế khác nhau. Các hộp đặc trưng trong hộp bảng mẫu (“stbl”) đã được chuẩn hóa để đáp ứng với các nhu cầu chung. Ví dụ, hộp mẫu đồng bộ hóa (“sync sample box-stss”) là một hộp nằm trong hộp bảng mẫu. Hộp mẫu đồng bộ hóa được sử dụng để lập danh sách các mẫu truy cập ngẫu nhiên của rãnh. Bản mô tả này có thể đề cập đến mẫu được lập danh sách bởi hộp mẫu đồng bộ hóa làm mẫu đồng bộ hóa. Theo ví dụ khác, cơ chế tạo nhóm mẫu có thể lập sơ đồ mẫu theo kiểu tạo nhóm 4 đặc điểm thành các nhóm mẫu có chung đặc tính được mô tả là mục nhập mô tả nhóm mẫu trong tập tin. Một số kiểu tạo nhóm đã được mô tả trong ISOBMFF.

Hộp bảng mẫu bao gồm một hoặc nhiều hộp SampleToGroup và một hoặc nhiều hộp mô tả nhóm mẫu (tức là, hộp SampleGroupDescription). Hộp SampleToGroup có thể được sử dụng để xác định nhóm mẫu mà mẫu này thuộc về, cùng với mô tả liên quan về nhóm mẫu này. Nói cách khác, hộp SampleToGroup có thể chỉ ra nhóm mà mẫu này thuộc về. Hộp SampleToGroup có thể là kiểu hộp “sbgp”. Hộp SampleToGroup có thể bao gồm phần tử kiểu tạo nhóm (ví dụ, grouping\_type). Phần tử kiểu tạo nhóm có thể là một số nguyên nhận dạng kiểu (tức là, tiêu chí được sử dụng để tạo nhóm mẫu) của việc tạo nhóm mẫu. Hơn nữa, hộp SampleToGroup có thể bao gồm một hoặc nhiều mục nhập. Mỗi mục nhập trong hộp SampleToGroup có thể liên quan đến một loạt các mẫu liên tiếp khác nhau, không chồng lấp trong rãnh. Mỗi mục nhập có thể chỉ ra phần tử đếm mẫu (ví dụ, sample\_count) và phần tử chỉ số mô tả nhóm (ví dụ, group\_description\_index). Phần tử đếm mẫu của một mục nhập có thể chỉ ra số lượng mẫu kết hợp với mục nhập này. Nói cách khác, phần tử đếm mẫu của mục nhập có thể là số nguyên xác định số mẫu liên tiếp có cùng bộ mô tả nhóm

mẫu. Phần tử chỉ số mô tả nhóm có thể chỉ ra hộp SampleGroupDescription chứa một mô tả về các mẫu kết hợp với mục nhập. Phần tử chỉ số mô tả nhóm của nhiều mục nhập có thể nhận dạng cùng một hộp SampleGroupDescription.

Thiết kế định dạng tập tin hiện nay có một hoặc nhiều vấn đề. Để lưu trữ nội dung video của một bộ mã hóa video cụ thể dựa trên ISOBMFF, việc mô tả về định dạng của tập tin cho bộ mã hóa video đó có thể là cần thiết. Để lưu trữ dòng video chứa nhiều lớp như MV-HEVC và SHVC, có thể sử dụng lại một số khái niệm của định dạng tập tin SVC và MVC. Tuy nhiên, nhiều phần không thể sử dụng trực tiếp cho các dòng video SHVC và MV-HEVC. Việc áp dụng trực tiếp định dạng tập tin HEVC có ít nhất là các nhược điểm sau: Dòng bit SHVC và MV-HEVC có thể bắt đầu bằng đơn vị truy cập chứa hình ảnh IRAP ở lớp cơ sở, nhưng cũng có thể chứa các hình ảnh không IRAP khác ở các lớp khác hoặc ngược lại. Mẫu đồng bộ hóa hiện nay không cho phép chỉ ra điểm như vậy để truy cập ngẫu nhiên.

Bản mô tả này mô tả các giải pháp có thể thực hiện đối với các vấn đề trên, cũng như đưa ra các cải thiện tiềm năng, có thể lưu trữ hiệu quả và linh hoạt dòng video chứa nhiều lớp. Các kỹ thuật được mô tả trong bản mô tả này có thể áp dụng cho định dạng tập tin bất kỳ để lưu trữ nội dung video được mã hóa bởi bộ mã hóa video bất kỳ, mặc dù phần mô tả là dành riêng để lưu trữ các dòng video SHVC và MV-HEVC dựa trên định dạng tập tin HEVC, được mô tả trong mệnh đề 8 của ISO/IEC 14496-15.

Phương án thực hiện chi tiết của các kỹ thuật theo bản mô tả này sẽ được thảo luận dưới đây. Các kỹ thuật của bản mô tả này có thể được nêu vắn tắt trong các ví dụ sau. Các ví dụ sau có thể được sử dụng riêng rẽ. Theo cách khác, sự kết hợp khác nhau của các ví dụ sau có thể cùng được sử dụng.

Theo ví dụ thứ nhất, Compressorname là giá trị được mô tả trong hộp VisualSampleEntry. Như được mô tả trong phần 8.5.2.1 trong ISO/IEC 14496-12, hộp VisualSampleEntry là kiểu hộp bảng mẫu cho các rãnh video lưu trữ thông tin chi tiết về kiểu mã hóa được sử dụng và thông tin khởi tạo bất kỳ cần thiết cho quá trình mã hóa này. Compressorname chỉ ra tên của bộ nén được sử dụng để tạo ra dữ liệu phương tiện. Bộ giải mã video có thể sử dụng giá trị của Compressorname để xác định bằng cách nào và/hoặc có giải mã dữ liệu video trong tập tin hay không. Như được mô tả trong phần 8.5.3 của ISO/IEC 14496-12, Compressorname được định dạng theo

trường 32 byte cố định, với byte đầu tiên được đặt là số byte được hiển thị, tiếp đó là số byte của dữ liệu có thể hiển thị được và sau đó là tạo đệm để hoàn chỉnh tổng số 32 byte (kể cả byte kích thước).

Ví dụ đầu tiên cho phép có hai giá trị Compressorname mới. Giá trị Compressorname mới thứ nhất là “\013SHVC Coding,” đối với tập tin chứa dòng video SHVC. Giá trị Compressorname mới thứ hai là “\016MV-HEVC Coding” đối với tập tin chứa dòng video MV-HEVC. Ví dụ thứ nhất này có thể được thực hiện như được thể hiện trong phần 9.5.3.1.3 và 10.5.3.2, dưới đây.

Như được mô tả một cách vắn tắt trên đây, tập tin có thể bao gồm hộp phim chứa siêu dữ liệu cho các rãnh của tập tin. Hộp phim bao gồm hộp rãnh cho mỗi rãnh của tập tin. Hơn nữa, hộp rãnh có thể bao gồm hộp thông tin phương tiện chứa tất cả các đối tượng thể hiện các thông tin đặc điểm của dữ liệu của rãnh. Hộp thông tin phương tiện có thể bao gồm hộp bảng mẫu. Hộp bảng mẫu có thể mô tả siêu dữ liệu dành riêng cho mẫu. Chẳng hạn, hộp bảng mẫu có thể bao gồm nhiều hộp mô tả mẫu. Mỗi hộp mô tả mẫu là một ví dụ về mục nhập mẫu. ISO/IEC 14496-12, các ví dụ về nhóm VisualSampleEntry có thể được sử dụng làm mục nhập mẫu. Các nhóm mục nhập mẫu dùng riêng cho các chuẩn mã hóa video có thể mở rộng nhóm VisualSampleEntry. Ví dụ, nhóm mục nhập mẫu dùng riêng cho HEVC có thể mở rộng nhóm VisualSampleEntry. Do vậy, bản mô tả này có thể đề cập đến các nhóm khác nhau mở rộng nhóm VisualSampleEntry dưới dạng kiểu mục nhập mẫu khác.

Theo ví dụ thứ hai, hai kiểu mục nhập mẫu mới (tức là, “mẫu”), “hev2” và “hvc2”, được xác định cho các rãnh HEVC. Hai kiểu mục nhập mẫu mới này cho phép sử dụng bộ gộp và bộ trích. Nói chung, bộ gộp gộp nhiều đơn vị NAL ở dạng đơn vị dữ liệu gộp đơn. Chẳng hạn, bộ gộp có thể chứa nhiều đơn vị NAL và/hoặc có thể gắn như là ghép nối nhiều đơn vị NAL. Nói chung, bộ trích chỉ ra kiểu dữ liệu thu được từ các rãnh khác. Chẳng hạn, việc lưu trữ dữ liệu phương tiện (ví dụ, dữ liệu HEVC) trên nhiều rãnh có thể tạo ra các tập tin nén do sự sao chép dữ liệu có thể tránh được bằng cách tham chiếu dữ liệu qua các rãnh phương tiện nhờ sử dụng các đơn vị dữ liệu tương đối nhỏ được gọi là bộ trích được nhúng dưới dạng đơn vị NAL trong dữ liệu phương tiện. Ví dụ thứ hai này có thể được thực hiện như được thể hiện trong phần 9.5.3.1.1, 9.5.3.1.2, 9.5.4, 9.5.6, 10.4.5, 10.5.3.1.1.1, và 10.5.3.2, dưới đây.

Theo ví dụ thứ ba, định nghĩa về mục nhập mẫu liên quan đến một yêu cầu cụ thể đối với việc lưu trữ tập tham số cho dòng bit đa lớp được cải biến đến có thể truy cập ngẫu nhiên một cách thuận lợi vào một lớp cụ thể hoặc điểm hoạt động cụ thể. Chẳng hạn, khi rãnh SHVC, MV-HEVC, hoặc 3D-HEVC có mục nhập mẫu và khi mẫu chứa ít nhất một hình ảnh IRAP, tất cả các tham số cần thiết để giải mã mẫu này sẽ được chứa trong mục nhập mẫu hoặc chính mẫu đó. Theo ví dụ này, khi mẫu không chứa hình ảnh IRAP bất kỳ, tất cả các tập tham số (ví dụ, VPS, SPS, PPS) cần thiết để giải mã mẫu đó sẽ được chứa trong mục nhập mẫu hoặc trong mẫu bất kỳ vì mẫu trước đó chứa ít nhất một hình ảnh IRAP cho chính mẫu đó. Ví dụ thứ ba này có thể được thực hiện như được thể hiện trong phần 9.5.3.1.1 dưới đây.

Theo phiên bản khác của ví dụ thứ ba, khi rãnh SHVC, MV-HEVC, hoặc 3D-HEVC có mục nhập mẫu và khi một hình ảnh trong mẫu là hình ảnh IRAP, tất cả các tham số cần thiết để giải mã mẫu này sẽ được chứa trong mục nhập mẫu hoặc chính mẫu đó. Hơn nữa, theo cách khác, khi mẫu không chứa hình ảnh IRAP bất kỳ, tất cả các tập tham số cần thiết để giải mã hình ảnh này sẽ được chứa trong mục nhập mẫu hoặc trong mẫu bất kỳ đứng sau mẫu trước đó chứa ít nhất một hình ảnh IRAP trong cùng một lớp đối với chính mẫu đó.

Theo ví dụ thứ tư, các trường hợp sau đối với các kiểu mục nhập mẫu đã có được xác định. Theo ví dụ này, các mẫu thuộc kiểu mục nhập mẫu “hev1” và “hvc1” chứa các cấu hình HEVC, SHVC, và MV-HEVC cho các rãnh SHVC và MV-HEVC cùng với đơn vị NAL VCL HEVC. Hơn nữa, kiểu mục nhập mẫu “hev1” và “hvc1” chứa các cấu hình SHVC và MV-HEVC được xác định cho các rãnh SHVC và MV-HEVC mà không có đơn vị NAL HEVC nhưng có đơn vị NAL VCL có `nuh_layer_id` lớn hơn 0, ở đó bộ trích không được chấp nhận. Ví dụ thứ tư này có thể được thực hiện như được thể hiện trong phần 9.5.3.1.1 dưới đây.

Theo ví dụ thứ năm, mẫu đồng bộ hóa trong rãnh SHVC, MV-HEVC, hoặc 3D-HEVC được xác định là mẫu chứa tất cả các hình ảnh IRAP. Ví dụ thứ năm này có thể được thực hiện như được thể hiện trong phần 9.5.5 và 10.4.3, dưới đây. Như được mô tả trong phần 9.5.5, dưới đây, mẫu SHVC được xem là mẫu đồng bộ hóa nếu mỗi hình ảnh được mã hóa trong đơn vị truy cập là hình ảnh IRAP, như được định nghĩa trong HEVC WD. Hơn nữa, như được mô tả trong phần 10.4.3, dưới đây, mẫu MV-HEVC

được xem là mẫu đồng bộ hóa nếu mỗi hình ảnh được mã hóa trong đơn vị truy cập là hình ảnh IRAP không có hình ảnh RASL, như được định nghĩa trong HEVC WD.

Do vậy, theo ví dụ thứ năm này, là một phần của quá trình tạo tập tin, thiết bị tạo tập tin 34 có thể tạo hộp mẫu đồng bộ hóa bao gồm bảng mẫu đồng bộ hóa ghi lại các mẫu đồng bộ hóa của một rãnh dữ liệu video đa lớp. Mỗi mẫu đồng bộ hóa của rãnh này là mẫu truy cập ngẫu nhiên của rãnh. Mẫu mã hóa video mở rộng được là mẫu đồng bộ hóa nếu mỗi hình ảnh được mã hóa trong đơn vị truy cập là hình ảnh IRAP. Mẫu mã hóa video đa khung hình là mẫu đồng bộ hóa nếu mỗi hình ảnh được mã hóa trong đơn vị truy cập là hình ảnh IRAP mà không có hình ảnh RASL.

Theo phiên bản khác của ví dụ thứ năm, mẫu đồng bộ hóa trong rãnh SHVC, MV-HEVC, hoặc 3D-HEVC được xác định là mẫu chứa tất cả các hình ảnh là hình ảnh IRAP không có hình ảnh RASL. Bảng mẫu đồng bộ hóa ghi lại mẫu đồng bộ hóa. Tùy ý, nhóm mẫu đồng bộ hóa ghi lại các mẫu đồng bộ hóa. Nói cách khác, nhóm mẫu đồng bộ hóa bao gồm thông tin nhận dạng mẫu đồng bộ hóa.

Theo ví dụ thứ sáu, nhóm mẫu “rap” được xác định để chứa các mẫu này chứa các hình ảnh là hình ảnh IRAP (có hoặc không có hình ảnh RASL). Ví dụ thứ sáu này có thể được thực hiện như được thể hiện trong phần 9.5.5 dưới đây. Nói cách khác, theo ví dụ thứ sáu, nhóm mẫu “rap” được xác định chứa các mẫu chứa các hình ảnh là hình ảnh IRAP nhưng không chứa các mẫu được chỉ ra là mẫu đồng bộ hóa.

Theo ví dụ thứ bảy, nhóm mẫu mới hoặc hộp mới được xác định là ghi lại tất cả các mẫu chứa ít nhất một hình ảnh IRAP, kiểu đơn vị NAL của các đơn vị NAL VCL trong các hình ảnh IRAP trong mẫu, dù cho tất cả các hình ảnh được mã hóa có là hình ảnh IRAP hay không, và nếu không, số lượng ảnh IRAP trong mẫu và các giá trị ID lớp của các hình ảnh IRAP này trong mẫu.

Do vậy, theo ví dụ thứ bảy, thiết bị tạo tập tin 34 có thể tạo tập tin bao gồm hộp rãnh chứa siêu dữ liệu cho một rãnh trong tập tin. Dữ liệu phương tiện cho rãnh này bao gồm chuỗi các mẫu. Mỗi mẫu có thể là đơn vị truy cập video của dữ liệu video đa lớp. Là một phần của quá trình tạo tập tin, thiết bị tạo tập tin 34 tạo, trong tập tin này, một hộp bổ sung ghi lại tất cả các mẫu chứa ít nhất một hình ảnh IRAP.

Ví dụ thứ bảy này có thể được thực hiện như được thể hiện trong phần 9.5.5.1. dưới đây. Như được thể hiện trong phần 9.5.5.1, dưới đây, nhóm mục nhập mẫu truy cập ngẫu nhiên (Random Accessible Sample Entry) mở rộng nhóm

VisualSampleGroupEntry. Đối tượng của nhóm mục nhập mẫu truy cập ngẫu nhiên (tức là các hộp mục nhập mẫu truy cập ngẫu nhiên) tương ứng với các mẫu chứa ít nhất một hình ảnh IRAP. Hơn nữa, hộp Random Accessible mục nhập mẫu bao gồm một giá trị `all_pics_are_IRAP` mô tả xem tất cả các hình ảnh được mã hóa ở mẫu tương ứng có là hình ảnh IRAP hay không.

Do vậy, theo ví dụ thứ bảy này, thiết bị tạo tập tin 34 có thể tạo ra mục nhập mẫu gồm một giá trị (ví dụ, `all_pics_are_IRAP`). Giá trị này bằng 1 mô tả mỗi hình ảnh được mã hóa trong mẫu là hình ảnh IRAP. Giá trị này bằng 0 mô tả không phải tất cả các hình ảnh được mã hóa trong mẫu là hình ảnh IRAP.

Hơn nữa, theo ví dụ thứ bảy, khi không phải tất cả các hình ảnh được mã hóa của mẫu là hình ảnh IRAP, thiết bị tạo tập tin 34 có thể đưa vào, trong mục nhập mẫu này tương ứng với mẫu, một giá trị chỉ báo số lượng hình ảnh IRAP trong mỗi mẫu của nhóm mẫu. Ngoài ra, khi không phải tất cả các hình ảnh được mã hóa của mẫu là hình ảnh IRAP, thiết bị tạo tập tin 34 có thể đưa vào, trong mục nhập mẫu này tương ứng với mẫu, một giá trị chỉ báo ký hiệu nhận dạng lớp của hình ảnh IRAP trong mẫu.

Theo cách khác, theo ví dụ thứ bảy, nhóm mẫu mới hoặc hộp mới ghi lại các mẫu này nhưng không bao gồm các mẫu được chỉ ra là mẫu hoặc thành viên đồng bộ hóa của nhóm mẫu "rap".

Ví dụ thứ bảy này có thể giải quyết một hoặc nhiều vấn đề có thể nảy sinh khi dữ liệu video đa lớp được lưu trữ bằng cách sử dụng ISOBMFF hoặc các phiên bản mở rộng của nó. Ví dụ, trong mã hóa video một lớp, thường chỉ có một hình ảnh được mã hóa trên một đơn vị truy cập. Tuy nhiên, trong mã hóa video đa lớp, thường có nhiều hơn một hình ảnh được mã hóa trên một đơn vị truy cập. ISOBMFF và các phiên bản mở rộng của nó không cung cấp cách thức chỉ ra mẫu nào chứa một hoặc nhiều hình ảnh IRAP. Điều này có thể cản trở khả năng thiết bị điện toán định vị các điểm truy cập ngẫu nhiên trong tập tin hoặc thực hiện sự chuyển lớp. Chẳng hạn, với việc không có thông tin chỉ ra mẫu nào chứa một hoặc nhiều hình ảnh IRAP, thiết bị điện toán có thể cần phải phân tích và diễn dịch các đơn vị NAL để xác định xem đơn vị truy cập có được sử dụng làm điểm truy cập ngẫu nhiên và/hoặc để chuyển lớp hay không. Việc phân tích và diễn dịch đơn vị NAL có thể là tăng thêm tính phức tạp của thiết bị điện toán và có thể mất nhiều thời gian và các tài nguyên xử lý. Hơn nữa, một

số thiết bị điện toán thực hiện truy cập ngẫu nhiên và/hoặc chuyển lớp, như máy chủ truyền liên tục, không được tạo cấu hình để phân tích hoặc diễn dịch các đơn vị NAL.

Theo ví dụ thứ tám, việc đưa ra một kiểu mẫu con mới được mô tả, trong đó mỗi mẫu con chứa một hình ảnh được mã hóa và các đơn vị NAL không VCL liên quan của nó. Ví dụ thứ tám này có thể được thực hiện như được thể hiện trong phần 9.5.8 dưới đây. Do vậy, theo ví dụ thứ tám, thiết bị tạo tập tin 34 có thể tạo tập tin bao gồm hộp rãnh chứa siêu dữ liệu cho một rãnh trong tập tin. Dữ liệu phương tiện cho rãnh này bao gồm chuỗi các mẫu. Mỗi mẫu là đơn vị truy cập video của dữ liệu video đa lớp. Là một phần của quá trình tạo tập tin, thiết bị tạo tập tin 34 tạo ra, trong tập tin này, hộp thông tin mẫu con chứa các cờ chỉ rõ kiểu thông tin mẫu con có trong hộp thông tin mẫu con. Khi cờ có giá trị cụ thể, mẫu con tương ứng với hộp thông tin mẫu con chứa chính xác một hình ảnh được mã hóa và không hoặc nhiều đơn vị NAL không VCL liên quan đến hình ảnh được mã hóa này.

Ví dụ thứ tám này có thể giải quyết một hoặc nhiều vấn đề có thể nảy sinh khi dữ liệu video đa lớp được lưu trữ bằng cách sử dụng ISOBMFF hoặc các phiên bản mở rộng của nó. Chẳng hạn, trong mã hóa video đa lớp, có thể có nhiều hình ảnh được mã hóa trên một mẫu. Ví dụ, có thể có một hoặc nhiều hình ảnh trong một mẫu đối với mỗi lớp. Tuy nhiên, theo phiên bản mở rộng của ISOBMFF đối với H.264/AVC và HEVC, hộp thông tin mẫu con không cung cấp thông tin về các hình ảnh cụ thể trong một mẫu khi mẫu này chứa nhiều hình ảnh. Kỹ thuật theo ví dụ thứ tám có thể giải quyết vấn đề này bằng cách đưa ra một kiểu hộp thông tin mẫu con mới về mẫu con chỉ chứa một hình ảnh được mã hóa và chứa các đơn vị NAL không VCL liên quan đến hình ảnh được mã hóa này. Việc đưa ra thông tin về hình ảnh được mã hóa riêng biệt trong cấu trúc của tập tin, như được đề xuất chỉ để cung cấp thông tin này trong các đơn vị NAL liên quan đến hình ảnh được mã hóa, có thể cho phép thiết bị điện toán xác định thông tin về hình ảnh được mã hóa mà không cần phải diễn dịch đơn vị NAL. Theo một số ví dụ, để làm giảm tính phức tạp của thiết bị điện toán và/hoặc làm gia tăng công suất của thiết bị điện toán, thiết bị điện toán không được tạo cấu hình để diễn dịch đơn vị NAL. Theo một số ví dụ trong đó thiết bị điện toán là các đơn vị NAL truyền liên tục được lưu trữ trong tập tin, thiết bị điện toán có thể sử dụng thông tin trong hộp thông tin mẫu con để xác định xem có chuyển tiếp đơn vị NAL của mẫu con đến thiết bị khách hay không.

Ví dụ thứ chín đề cập đến việc xử lý các mẫu không kết xuất trong trường hợp đa lớp. Cụ thể là, theo ví dụ thứ chín, khi đơn vị truy cập chứa một số hình ảnh được mã hóa có giá trị `pic_output_flag` bằng 1 và một số hình ảnh được mã hóa khác có `pic_output_flag` bằng 0, phải sử dụng ít nhất hai rãnh để lưu trữ dòng này, để trong mỗi rãnh tất cả các hình ảnh được mã hóa trong mỗi mẫu có cùng giá trị `pic_output_flag`. Ví dụ thứ chín này có thể được thực hiện như được thể hiện trong phần 9.5.9 dưới đây.

Do vậy, theo ví dụ thứ chín, thiết bị tạo tập tin 34 có thể tạo tập tin bao gồm hộp dữ liệu phương tiện chứa nội dung phương tiện. Nội dung phương tiện bao gồm chuỗi các mẫu. Mỗi mẫu là đơn vị truy cập video của dữ liệu video đa lớp. Đáp lại việc xác định rằng ít nhất một đơn vị truy cập của dòng bit của dữ liệu video đa lớp bao gồm hình ảnh được mã hóa có cờ kết xuất hình ảnh (ví dụ, `pic_output_flag`) bằng 1 và hình ảnh được mã hóa có cờ kết xuất hình ảnh bằng 0, thiết bị tạo tập tin 34 có thể sử dụng ít nhất hai rãnh để lưu trữ dòng bit trong tập tin. Đối với mỗi rãnh tương ứng trong số ít nhất hai rãnh này, tất cả các hình ảnh được mã hóa trong mỗi mẫu của rãnh tương ứng có cùng giá trị cờ kết xuất hình ảnh.

Ví dụ thứ chín này có thể giải quyết một hoặc nhiều vấn đề có thể nảy sinh khi dữ liệu video đa lớp được lưu trữ bằng cách sử dụng ISOBMFF hoặc các phiên bản mở rộng của nó. Chẳng hạn, nếu một rãnh đơn được sử dụng để lưu trữ các hình ảnh được mã hóa có cờ kết xuất hình ảnh bằng 0 và cờ kết xuất hình ảnh bằng 1, các quy tắc định dạng tập tin khác nhau có thể bị vi phạm. Ví dụ, các quy tắc định dạng tập tin thường đòi hỏi chỉ có một mẫu trong một rãnh trong một khoảng thời gian. Nếu một rãnh đơn lưu trữ các hình ảnh được mã hóa có cờ kết xuất hình ảnh bằng 0 và cờ kết xuất hình ảnh bằng 1, có thể có nhiều mẫu trong một rãnh trong một khoảng thời gian. Việc bắt các hình ảnh được mã hóa có các giá trị cờ kết xuất hình ảnh khác nhau trong các rãnh khác nhau có thể giải quyết vấn đề này.

Phương án thực hiện ví dụ về một số kỹ thuật của bản mô tả được mô tả dưới đây. Phương án thực hiện ví dụ được mô tả dưới đây là dựa trên phần mô tả tích hợp mới nhất 14496-15 trong tài liệu xuất MPEG W13478. Những sửa đổi đối với Phụ lục A (được thể hiện bằng cách gạch chân) và các phần được bổ sung (phần 9 đối với SHVC và phần 10 đối với MV-HEVC) được mô tả dưới đây. Nói cách khác, các ví dụ cụ thể về bản mô tả này có thể sửa đổi Phụ lục A của ISO/IEC 14496-15 và có thể bỏ

sung phần 9 và/hoặc 10 vào ISO/IEC 14496-15. Phần văn bản được thể hiện bằng cách gạch chân và gạch chân hai lần có thể liên quan cụ thể đến các ví dụ của bản mô tả này. Mặc dù thuật ngữ SHVC được sử dụng ở các chỗ khác nhau trong các ví dụ được mô tả trong bản mô tả này, thiết kế của bản mô tả này thực sự không chỉ để hỗ trợ bộ mã hóa SHVC, mà tất cả các bộ mã hóa đa lớp bao gồm MV-HEVC, 3D-HEVC đều có thể được hỗ trợ, trừ khi có chỉ dẫn rõ ràng khác.

## 9 Các dòng SHVC cơ bản và các định nghĩa về mẫu

### 9.1 Giới thiệu

Mệnh đề này mô tả định dạng lưu trữ dữ liệu SHVC. Mệnh đề này mở rộng các định nghĩa về định dạng lưu trữ của HEVC trong mệnh đề 8.

Định dạng tập tin để lưu trữ nội dung SHVC, như được định nghĩa trong mệnh đề này và các Phụ lục từ A đến D, sử dụng các khả năng đã có của định dạng tập tin phương tiện cơ sở ISO và định dạng tập tin HEVC đơn giản (tức là, định dạng tập tin được mô tả trong mệnh đề 8). Ngoài ra, các cấu trúc hoặc phiên bản mở rộng sau, ngoài những cái khác, hỗ trợ các đặc điểm dành riêng cho SHVC được sử dụng.

Bộ gộp:

cấu trúc có thể thực hiện việc tạo nhóm đơn vị NAL có thể mở rộng được một cách hiệu quả bằng cách biến đổi các mẫu đơn vị NAL không theo quy tắc thành các mẫu đơn vị dữ liệu được gộp theo quy tắc.

Bộ trích:

cấu trúc có thể tách hiệu quả các đơn vị NAL ra khỏi các rãnh khác không phải rãnh chứa dữ liệu phương tiện.

Các câu lệnh siêu dữ liệu thời gian:

cấu trúc để lưu trữ thông tin được sắp hàng theo thời gian của mẫu phương tiện.

Khả năng tương thích với HEVC:

một điều khoản để lưu trữ một dòng bit SHVC theo cách thức tương thích với HEVC, để lớp cơ sở tương thích với HEVC có thể được sử dụng bởi thiết bị đọc tương thích với định dạng tập tin HEVC đơn giản bất kỳ.

### 9.2 Cấu trúc dòng cơ bản

Các dòng SHVC được lưu trữ theo 8.2, với định nghĩa sau về dòng video SHVC cơ bản:

- Dòng video SHVC cơ bản chứa tất cả các đơn vị NAL liên quan đến việc mã hóa video (tức là, các đơn vị NAL chứa dữ liệu video hoặc báo hiệu cấu trúc video) và có thể chứa các đơn vị NAL không liên quan đến việc mã hóa video như thông báo SEI và đơn vị NAL đầu cách đơn vị truy cập. Ngoài ra, bộ gộp (xem A.2) hoặc bộ trích (xem A.3) có thể có mặt. Bộ gộp và bộ trích được xử lý như được mô tả trong chuẩn quốc tế (ví dụ, không được đặt trực tiếp trong bộ đệm kết xuất trong khi truy cập tập tin). Các đơn vị NAL khác không bị cấm tuyệt đối có thể có mặt và nếu chúng không được nhận dạng chúng nên được bỏ qua (ví dụ, không đặt trong bộ đệm kết xuất trong quá trình truy cập tập tin).

Các dòng SHVC sẽ không được lưu trữ bằng cách sử dụng dòng tập tham số liên quan.

Có thể có các đơn vị NAL VCL có `nuh_layer_id` bằng 0, đơn vị NAL VCL có `nuh_layer_id` lớn hơn 0, và đơn vị NAL không VCL có mặt trong dòng video SHVC cơ bản. Ngoài ra, có thể có các đơn vị NAL bộ gộp và đơn vị NAL bộ trích có mặt trong dòng video SHVC cơ bản.

### 9.3 Sử dụng định dạng tập tin HEVC đơn giản

Định dạng tập tin SHVC là phiên bản mở rộng của định dạng tập tin HEVC đơn giản được định nghĩa trong mệnh đề 8.

### 9.4 Định nghĩa mẫu và cấu hình

#### 9.4.1 Giới thiệu

Mẫu SHVC: Mẫu SHVC cũng là một đơn vị truy cập như được định nghĩa trong phụ lục H của ISO/IEC 23008-2.

#### 9.4.2 Thứ tự tiêu chuẩn và các giới hạn

##### 9.4.2.1 Các giới hạn

Các giới hạn sau áp dụng cho dữ liệu SHVC ngoài các yêu cầu trong 8.3.2.

- Đơn vị NAL VCL: Tất cả các đơn vị NAL VCL trong một đơn vị truy cập được chứa trong mẫu mà thời gian hợp thành của nó là thời gian hình ảnh được biểu diễn bởi đơn vị truy cập. Mẫu SHVC chứa ít nhất một đơn vị NAL VCL.
- Bộ gộp/bộ trích: Thứ tự của tất cả các đơn vị NAL trong bộ gộp hoặc được tham chiếu bởi bộ trích chính xác là thứ tự giải mã nếu như các đơn vị NAL có mặt trong mẫu không chứa bộ gộp/bộ trích. Sau khi xử lý bộ gộp hoặc bộ trích,

tất cả các đơn vị NAL phải theo thứ tự giải mã có hiệu lực như được mô tả trong ISO/IEC 23008-2.

#### 9.4.2.2 Bản ghi cấu hình bộ giải mã

Khi bản ghi cấu hình bộ giải mã được xác định trong 8.3.3.1 được sử dụng cho dòng có thể được diễn dịch là dòng SHVC hoặc dòng HEVC, bản ghi cấu hình bộ giải mã HEVC sẽ phản ánh các đặc tính của lớp cơ sở tương thích với HEVC, ví dụ, nó chỉ chứa tập tham số cần thiết để giải mã lớp cơ sở HEVC.

Bản ghi cấu hình bộ giải mã theo SHVC về mặt cấu trúc giống với bản ghi cấu hình bộ giải mã theo HEVC. Cú pháp là như sau:

```
aligned(8) class SHVCDecoderConfigurationRecord {
    // same fields as in HEVCDecoderConfigurationRecord syntax
    structure
}
```

Ngữ nghĩa của các trường trong bản ghi cấu hình bộ giải mã theo SHVC là giống như được xác định đối với bản ghi cấu hình bộ giải mã theo HEVC.

### 9.5 Suy ra từ định dạng tập tin phương tiện cơ sở ISO

#### 9.5.1 Cấu trúc của rãnh theo SHVC

Dòng video có thể mở rộng được được biểu diễn bằng một hoặc nhiều rãnh video trong tập tin. Mỗi rãnh biểu diễn một hoặc nhiều điểm hoạt động của dòng có thể mở rộng. Dòng mở rộng này, có thể được làm mỏng thêm, nếu muốn.

Điểm hoạt động thấp nhất là một điểm trong số tất cả các điểm hoạt động chứa đơn vị NAL có `nuh_layer_id` chỉ bằng 0 và `TemporalID` chỉ bằng 0. Rãnh chứa điểm hoạt động thấp nhất được gọi là "rãnh cơ sở mở rộng được." Tất cả các rãnh khác là một phần của cùng một thông tin mã hóa có thể mở rộng được liên kết với rãnh cơ sở này bằng phương tiện tham chiếu rãnh theo kiểu 'sbas' (cơ sở mở rộng được).

Tất cả các rãnh có cùng chung rãnh cơ sở mở rộng được phải có cùng thang thời gian như rãnh cơ sở mở rộng được.

#### 9.5.2 Chia sẻ dữ liệu và trích dữ liệu

Các rãnh khác nhau có thể chia sẻ dữ liệu theo cách logic. Việc chia sẻ này có thể theo một trong số hai dạng sau:

- a) Dữ liệu mẫu được sao chép từ một rãnh vào rãnh khác (và có thể được nén hoặc được xen kẽ bằng dữ liệu khác, như âm thanh). Việc này tạo ra các tập tin lớn

hơn, dữ liệu tốc độ bit thấp có thể được nén và/hoặc được đan xen kẽ bằng nguyên liệu khác, để dễ dàng trích ra.

- b) Có các hướng dẫn về việc bằng cách nào để thực hiện sự sao chép này ở thời điểm tập tin được đọc.

Đối với trường hợp thứ hai, bộ trích (như được xác định trong A.3) được sử dụng.

### 9.5.3 Định nghĩa về dòng video theo SHVC

#### 9.5.3.1 Tên mục nhập mẫu và định dạng

##### 9.5.3.1.1 Định nghĩa:

Kiểu: 'hvc2', 'hev2', 'shc1', 'shv1', 'shcC'

Bộ lưu trữ: Hộp mô tả mẫu ('std')

Bắt buộc: Mục nhập mẫu 'hvc1', 'hev1', 'hvc2', 'hev2', 'shc1', hoặc 'shv1' là bắt buộc

Số lượng: Một hoặc nhiều mục nhập mẫu có thể có mặt

Khi tên mục nhập mẫu là 'shc1', giá trị mặc định và bắt buộc của array\_completeness là 1 cho các mảng của tất cả các kiểu tập tham số và 0 cho tất cả các mảng khác. Khi tên mục nhập mẫu là 'shv1', giá trị mặc định của array\_completeness là 0 cho tất cả các mảng.

Khi tên mục nhập mẫu là 'shv1', áp dụng các điều kiện sau:

- Nếu mẫu chứa ít nhất một hình ảnh IRAP như được định nghĩa trong ISO/IEC 23008-2, tất cả các tập tham số cần thiết để giải mã mẫu đó được chứa trong mục nhập mẫu hoặc trong chính mẫu đó.
- Nếu không (mẫu không chứa hình ảnh IRAP), tất cả các tập tham số cần thiết để giải mã mẫu đó sẽ được chứa trong mục nhập mẫu hoặc trong mẫu bất kỳ vì mẫu trước đó chứa ít nhất một hình ảnh IRAP cho chính mẫu đó.

Theo cách khác, khi tên mục nhập mẫu là 'shv1', áp dụng như sau:

- Nếu hình ảnh được mã hóa trong mẫu là hình ảnh IRAP như được định nghĩa trong ISO/IEC 23008-2, tất cả các tập tham số cần thiết để giải mã hình ảnh được mã hóa đó được chứa trong mục nhập mẫu hoặc trong chính mẫu đó.
- Nếu không (hình ảnh được mã hóa trong mẫu không phải là hình ảnh IRAP), tất cả các tập tham số cần thiết để giải mã hình ảnh này sẽ được chứa trong mục

nhập mẫu hoặc trong mẫu bất kỳ vì mẫu trước đó chứa một hình ảnh IRAP trong một lớp giống như hình ảnh được mã hóa đó đối với chính mẫu đó.

Nếu dòng cơ bản SHVC chứa lớp cơ sở tương thích với HEVC có thể sử dụng, mục nhập mẫu trực quan HEVC ('hvc1' hoặc 'hev1') sẽ được sử dụng. Do vậy, mục nhập ban đầu chứa hộp cấu hình HEVC, có thể tiếp đến là hộp cấu hình SHVC như được xác định dưới đây. Hộp cấu hình HEVC ghi lại đặc tính, bậc, mức, và có thể là tập tham số liên quan đến lớp cơ sở tương thích với HEVC như được xác định bởi bản ghi cấu hình bộ giải mã HEVC (HEVCDecoderConfigurationRecord). Hộp cấu hình SHVC ghi lại đặc tính, bậc, mức, và có thể là tập tham số liên quan đến toàn bộ dòng chứa các lớp nâng cao tương thích với SHVC như được xác định bởi HEVCDecoderConfigurationRecord, được lưu trữ trong hộp cấu hình SHVC (SHVCConfigurationBox).

Nếu dòng cơ bản SHVC không chứa lớp cơ sở tương thích với HEVC có thể sử dụng, mục nhập mẫu trực quan SHVC ('shc1' hoặc 'shv1') sẽ được sử dụng. Mục nhập mẫu trực quan SHVC chứa một hộp cấu hình SHVC, như được xác định dưới đây. Mục nhập này chứa bản ghi cấu hình bộ giải mã SHVC, như được xác định trong chuẩn quốc tế này.

Trường lengthSizeMinusOne trong cấu hình SHVC và HEVC trong mục nhập mẫu xác định bất kỳ có cùng giá trị.

Bộ trích hoặc bộ gộp có thể được sử dụng cho các đơn vị NAL có nuh\_layer\_id lớn hơn 0 trong các rãnh 'hvc1', 'hev1', 'hvc2', 'hev2', 'shc1', hoặc 'shv1'. 'extra\_boxes' trong mục nhập mẫu 'hvc2' hoặc 'hev2' có thể là SHVCConfigurationBox hoặc các hộp mở rộng khác.

**LƯU Ý** Khi sự tương thích với HEVC được chỉ rõ, có thể cần phải chỉ ra mức độ không thật đối với lớp cơ sở HEVC, để phù hợp với tốc độ bit của toàn bộ dòng, do tất cả các đơn vị NAL được xem là chứa trong lớp cơ sở HEVC và do đó có thể được nạp vào bộ giải mã, được mong đợi là loại bỏ các đơn vị NAL mà nó không nhận ra. Trường hợp này xảy ra khi mục nhập mẫu 'hvc1' or 'hev1' được sử dụng cả hai cấu hình HEVC và SHVC đều có mặt.

SHVCConfigurationBox có thể có mặt trong mục nhập mẫu 'hvc1' hoặc 'hev1'. Trong trường hợp này, áp dụng định nghĩa về mục nhập mẫu HEVCSHVC (HEVCSHVCSampleEntry) dưới đây.

Bảng dưới đây thể hiện đối với một rãnh video thì việc sử dụng các mục nhập mẫu, cấu hình và công cụ SHVC đều có thể (trừ siêu dữ liệu được định thời, luôn được sử dụng trong rãnh khác):

Bảng 10 – Sử dụng các mục nhập mẫu đối với các rãnh HEVC và SHVC

| Tên mục nhập mẫu   | cùng với bản ghi cấu hình | Ý nghĩa                                                                                                                                                                                                                                                                  |
|--------------------|---------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 'hvc1' hoặc 'hev1' | Chỉ có cấu hình HEVC      | Rãnh HEVC đơn giản không có các đơn vị NAL có nuh_layer_id lớn hơn 0; bộ trích và bộ gộp không có mặt.                                                                                                                                                                   |
| 'hvc1' hoặc 'hev1' | Cấu hình HEVC và SHVC     | Rãnh SHVC với cả đơn vị NAL có nuh_layer_id bằng 0 và đơn vị NAL có nuh_layer_id lớn hơn 0; bộ trích và bộ gộp có thể có mặt; bộ trích không tham chiếu các đơn vị NAL có nuh_layer_id bằng 0; Bộ gộp không chứa nhưng tham chiếu các đơn vị NAL có nuh_layer_id bằng 0. |
| 'hvc2' hoặc 'hev2' | Chỉ có cấu hình HEVC      | Rãnh HEVC đơn giản không chứa các đơn vị NAL có giá trị nuh_layer_id lớn hơn 0; bộ trích có thể có mặt và được sử dụng để tham chiếu các đơn vị NAL; bộ gộp có thể có mặt để chứa và tham chiếu các đơn vị NAL.                                                          |
| 'hvc2' hoặc 'hev2' | Cấu hình HEVC và SHVC     | Rãnh SHVC với cả đơn vị NAL có nuh_layer_id bằng 0 và đơn vị NAL có nuh_layer_id lớn hơn 0; bộ trích và bộ gộp có thể có mặt; bộ trích tham chiếu đơn vị NAL bất kỳ; Bộ gộp không chứa và tham chiếu đơn vị NAL bất kỳ.                                                  |
| 'shc1' hoặc 'shv1' | Cấu hình SHVC             | Rãnh SHVC không chứa các đơn vị NAL có giá trị nuh_layer_id bằng 0; bộ trích có thể có mặt và được sử dụng để tham chiếu các đơn vị NAL; bộ gộp có thể có mặt để                                                                                                         |

|  |  |                                    |
|--|--|------------------------------------|
|  |  | chứa và tham chiếu các đơn vị NAL. |
|--|--|------------------------------------|

#### 9.5.3.1.2 Cú pháp

```

class SHVCCConfigurationBox extends Box('shcC') {
    SHVCDecoderConfigurationRecord() SHVCCConfig;
}

class HEVC2SampleEntry() extends VisualSampleEntry ('hev2' or 'hvc2') {
    HEVCConfigurationBox          hevconfig;
    MPEG4BitRateBox ();           // optional
    MPEG4ExtensionDescriptorsBox (); // optional
    extra_boxes          boxes;    // optional
}

class HEVCSHVCSampleEntry() extends HEVCSampleEntry {
    SHVCCConfigurationBox          shvcconfig;
}

class HEVC2SHVCSampleEntry() extends HEVC2SampleEntry() {
    SHVCCConfigurationBox          shvcconfig; // optional
}

// Use this if track is not HEVC compatible
class SHVCSampleEntry() extends VisualSampleEntry ('shv1' or 'shc1') {
    SHVCCConfigurationBox          shvcconfig;
    MPEG4BitRateBox ();           // optional
    MPEG4ExtensionDescriptorsBox (); // optional
    extra_boxes          boxes;    // optional
}

```

#### 9.5.3.1.3 Ngữ nghĩa

Khi dòng mã mục nhập mẫu áp dụng cho dòng này chứa các đơn vị NAL có giá trị `nuh_layer_id` lớn hơn 0, tên bộ nén ở Mục nhập mẫu trực quan (`VisualSampleEntry`) lớp cơ sở chỉ ra tên của bộ nén được sử dụng với giá trị "`\mã hóa 013SHVC`" được gợi ý (013 là 11, chiều dài của vòng "mã hóaSHVC" theo byte).

#### 9.5.4 Chiều rộng và chiều cao hiển thị theo SHVC

Chiều rộng và chiều cao hiển thị được ghi trong VisualSampleEntry của dòng chứa các đơn vị NAL có giá trị nuh\_layer\_id lớn hơn 0 là chiều rộng và chiều cao của lớp cơ sở HEVC, nếu dòng này được mô tả bằng mục nhập mẫu thuộc kiểu 'hvc1', 'hev1', 'hvc2', 'hev2'; theo cách khác, đó là chiều rộng và chiều cao của các hình ảnh được giải mã của lớp cao nhất bằng cách giải mã toàn bộ dòng này.

#### 9.5.5 Mẫu đồng bộ hóa

Mẫu SHVC được xem là mẫu đồng bộ hóa nếu mỗi hình ảnh được mã hóa trong đơn vị truy cập là hình ảnh IRAP, như được định nghĩa trong ISO/IEC 23008-2. Các mẫu đồng bộ hóa được ghi bởi bảng mẫu đồng bộ hóa, và có thể còn được ghi bởi nhóm mẫu đồng bộ hóa và nhóm mẫu 'rap'.

##### 9.5.5.1 Nhóm mẫu truy cập ngẫu nhiên

###### 9.5.5.1.1 Định nghĩa:

Kiểu nhóm: 'ras'

Bộ lưu trữ: Sample Group Description Box ('ras')

Bắt buộc: Không

Số lượng: Không hoặc lớn hơn

Nhóm mẫu truy cập ngẫu nhiên nhận dạng các mẫu chứa ít nhất một hình ảnh IRAP.

###### 9.5.5.1.2 Cú pháp

```
class RandomAccessibleSampleEntry() extends VisualSampleGroupEntry ('ras ') {
    unsigned int(1) reserved = 0;
    unsigned int(1) all_pics_are_IRAP
    unsigned int(6) IRAP_nal_unit_type
    if( !all_pics_are_IRAP ) {
        unsigned int(2) reserved = 0;
        unsigned int(6) num_IRAP_pics;
        for(i=0; i< num_IRAP_pics; i++) {
            unsigned int(2) reserved = 0;
            unsigned int(6) IRAP_pic_layer_id
        }
    }
}
```

### 9.5.5.1.3 Ngữ nghĩa

`all_pics_are_IRAP` bằng 1 mô tả tất cả các hình ảnh được mã hóa trong mỗi mẫu của nhóm là hình ảnh IRAP. Khi giá trị này bằng 0, ràng buộc trên có thể hoặc không thể áp dụng.

`IRAP_nal_unit_type` mô tả kiểu đơn vị NAL của các hình ảnh IRAP ở mỗi mẫu trong nhóm. Giá trị `IRAP_nal_unit_type` sẽ nằm trong khoảng từ 16 đến 23, bao gồm cả hai giá trị nút.

`num_IRAP_pics` mô tả số lượng ảnh IRAP trong mỗi mẫu của nhóm.

`IRAP_pic_layer_id` mô tả giá trị `nuh_layer_id` của hình ảnh IRAP thứ *i* trong mỗi nhóm của nhóm.

### 9.5.6 Nhóm mẫu ở các điểm khôi phục truy cập ngẫu nhiên và các điểm truy cập ngẫu nhiên

Đối với dữ liệu video được mô tả bởi mục nhập mẫu thuộc kiểu ‘hvc1’, ‘hev1’, ‘hvc2’, hoặc ‘hev2,’ nhóm mẫu khôi phục truy cập ngẫu nhiên và nhóm mẫu điểm truy cập ngẫu nhiên lần lượt nhận dạng các điểm khôi phục truy cập ngẫu nhiên và các điểm truy cập ngẫu nhiên, cho cả bộ giải mã HEVC và bộ giải mã SHVC (nếu có) hoạt động trên toàn bộ dòng bit.

Đối với dữ liệu video được mô tả bởi mục nhập mẫu kiểu ‘shc1’ hoặc ‘shv1,’ nhóm mẫu khôi phục truy cập ngẫu nhiên nhận dạng sự khôi phục truy cập ngẫu nhiên trên toàn bộ dòng bit SHVC và nhóm mẫu điểm truy cập ngẫu nhiên nhận dạng các điểm truy cập ngẫu nhiên trên toàn bộ dòng bit SHVC.

Một mẫu SHVC được xem là một điểm truy cập ngẫu nhiên nếu mỗi hình ảnh được mã hóa trong đơn vị truy cập là hình ảnh IRAP (có hoặc không có hình ảnh RASL) như được định nghĩa trong ISO/IEC 23008-2, và các mẫu hướng dẫn trong ISO/IEC 14496-2 là các hình ảnh trong đó tất cả các hình ảnh là hình ảnh RASL như được định nghĩa trong ISO/IEC 23008-2.

### 9.5.7 Hộp mẫu có thể dùng độc lập

Nếu được sử dụng trong một rãnh tương thích với cả HEVC và SHVC, thì điều quan tâm là các câu lệnh là đúng bất kể tập con có hiệu lực của dữ liệu SHVC (có thể chỉ có dữ liệu HEVC) được sử dụng. Giá trị ‘chưa biết’ (giá trị 0 của các trường `sample-depends-on`, `sample-is-depended-on`, và `sample-has-redundancy`) có thể là cần thiết nếu thông tin thay đổi.

### 9.5.8 Định nghĩa về mẫu con đối với SHVC

Mệnh đề phụ này mở rộng định nghĩa về mẫu con của HEVC ở mục 8.4.8.

Đối với việc sử dụng hộp thông tin mẫu con (8.7.7 của ISO/IEC 14496-12) trên dòng SHVC, mẫu con được định nghĩa trên cơ sở giá trị của các cờ của hộp thông tin mẫu con như được mô tả dưới đây. Sự có mặt của hộp này là tùy ý; tuy nhiên, nếu có mặt trong rãnh chứa dữ liệu SHVC, hộp này sẽ có ngữ nghĩa được mô tả trong bản mô tả này.

Các cờ chỉ rõ kiểu thông tin mẫu con có trong hộp này như sau:

- 0: Mẫu con trên cơ sở đơn vị NAL. Mẫu con chứa một hoặc nhiều đơn vị NAL liên kề.
- 1: Mẫu con trên cơ sở đơn vị giải mã. Mẫu con chứa chính xác một đơn vị giải mã.
- 2: Mẫu con trên cơ sở ô. Mẫu con chứa một ô và các đơn vị NAL không VCL liên quan, nếu có, của (các) đơn vị NAL VCL chứa ô hoặc chứa một hoặc nhiều đơn vị NAL không VCL.
- 3: Mẫu con trên cơ sở hàng CTU. Mẫu con chứa một hàng CTU trong một lát và các đơn vị NAL không VCL liên quan, nếu có, của (các) đơn vị NAL VCL chứa hàng CTU hoặc chứa một hoặc nhiều đơn vị NAL không VCL. Kiểu thông tin mẫu con này sẽ không được sử dụng khi giá trị `entropy_coding_sync_enabled_flag` bằng 0.
- 4: Mẫu con trên cơ sở lát. Một mẫu con chứa một lát (khi mỗi lát có thể chứa một hoặc nhiều đoạn lát, mỗi đoạn này là một đơn vị NAL) và các đơn vị NAL không VCL liên quan, nếu có, hoặc chứa một hoặc nhiều đơn vị NAL không VCL.
- 5: Mẫu con trên cơ sở hình ảnh. Mẫu con chứa một hình ảnh được mã hóa và đơn vị NAL không VCL liên quan.

Các giá trị khác của cờ được dự trữ.

Trường `subsample_priority` được thiết lập giá trị theo mô tả về trường này trong ISO/IEC 14496-12.

Trường có thể loại bỏ được thiết lập là 1 chỉ khi mẫu này vẫn có thể được giải mã nếu mẫu con này được loại bỏ (ví dụ, mẫu con này chứa đơn vị NAL SEI).

Khi byte đầu tiên của đơn vị NAL được bao gồm trong một mẫu con, trường độ dài đứng trước cũng phải được bao gồm trong cùng một mẫu con.

```

if(flags == 0) {
    unsigned int(1) SubLayerRefNalUnitFlag;
    unsigned int(1) RapNalUnitFlag;
    unsigned int(1) VclNalUnitFlag;
    unsigned int(1) DiscardableFlag;
    unsigned int(1) NoInterLayerPredFlag;
    unsigned int(6) LayerId;
    unsigned int(3) TempId;
    unsigned int(18) reserved = 0;
} else if(flags == 1)
    unsigned int(32) reserved = 0;
else if(flags == 2) {
    unsigned int(2) vcl_idc;
    unsigned int(2) reserved = 0;
    unsigned int(4) log2_min_luma_ctb;
    unsigned int(12) ctb_x;
    unsigned int(12) ctb_y;
} else if(flags == 3 || flags == 4) {
    unsigned int(2) vcl_idc;
    unsigned int(30) reserved = 0;
} esle if(flags == 5) {
    unsigned int(1) DiscardableFlag;
    unsigned int(6) VclNalUnitType;
    unsigned int(6) LayerId;
    unsigned int(3) TempId;
    unsigned int(1) NoInterLayerPredFlag;
    unsigned int(1) SubLayerRefNalUnitFlag;
    unsigned int(14) reserved = 0;
}

```

SubLayerRefNalUnitFlag bằng 0 chỉ ra rằng tất cả các đơn vị NAL trong mẫu con là đơn vị NAL VCL của một hình ảnh không tham chiếu của lớp con như được mô tả trong ISO/IEC 23008-2. Giá trị 1 chỉ ra rằng tất cả các đơn vị NAL trong mẫu con là đơn vị NAL VCL của một hình ảnh không tham chiếu của lớp con như được mô tả trong ISO/IEC 23008-2.

RapNalUnitFlag bằng 0 chỉ ra rằng không có đơn vị NAL nào trong mẫu con có nal\_unit\_type bằng IDR\_W\_RADL, IDR\_N\_LP, CRA\_NUT, BLA\_W\_LP, BLA\_W\_RADL, BLA\_N\_LP, RSV\_IRAP\_VCL22, hoặc RSV\_IRAP\_VCL23 như được mô tả trong ISO/IEC 23008-2. Giá trị 1 chỉ ra rằng tất cả các đơn vị NAL trong mẫu con có nal\_unit\_type bằng IDR\_W\_RADL, IDR\_N\_LP, CRA\_NUT, BLA\_W\_LP, BLA\_W\_RADL, BLA\_N\_LP, RSV\_IRAP\_VCL22, hoặc RSV\_IRAP\_VCL23 như được mô tả trong ISO/IEC 23008-2.

VclNalUnitFlag bằng 0 chỉ ra rằng tất cả các đơn vị NAL trong mẫu con là các đơn vị NAL không VCL. Giá trị 1 chỉ ra rằng tất cả các đơn vị NAL trong mẫu con là các đơn vị NAL VCL.

DiscardableFlag chỉ ra giá trị discardable\_flag của các đơn vị NAL VCL trong mẫu con. Tất cả các đơn vị NAL VCL trong mẫu con có cùng giá trị discardable\_flag.

LUU Ý rằng điều này không phải là định nghĩa giống như trường có thể loại bỏ trong hộp thông tin mẫu con.

NoInterLayerPredFlag chỉ ra giá trị của inter\_layer\_pred\_enabled\_flag của các đơn vị NAL VCL trong mẫu con. Tất cả các đơn vị NAL VCL trong mẫu con có cùng giá trị inter\_layer\_pred\_enabled\_flag.

LayerId chỉ ra giá trị nuh\_layer\_id của các đơn vị NAL trong mẫu con. Tất cả các đơn vị NAL VCL trong mẫu con có cùng giá trị nuh\_layer\_id.

TempId chỉ ra giá trị TemporalID của các đơn vị NAL trong mẫu con. Tất cả các đơn vị NAL VCL trong mẫu con có cùng giá trị TemporalID.

vcl\_idc chỉ ra mẫu con có chứa dữ liệu lớp mã hóa video (Video Coding Layer-VCL), dữ liệu không VCL hoặc cả hai hay không, như sau:

0: mẫu con chứa dữ liệu VCL và không chứa dữ liệu không VCL

1: mẫu con không chứa dữ liệu VCL và chứa dữ liệu không VCL

2: mẫu con chứa cả dữ liệu VCL và không VCL liên quan tới nhau. Ví dụ, mẫu con có thể chứa thông báo SEI của thông tin đơn vị giải mã tiếp đến là tập các đơn vị NAL liên quan đến thông báo SEI này.

3: được dự trữ

`log2_min_luma_ctb` chỉ ra đơn vị `ctb_x` và `ctb_y`, được mô tả như sau:

0: 8 mẫu độ sáng

1: 16 mẫu độ sáng

2: 32 mẫu độ sáng

3: 64 mẫu độ sáng

`ctb_x` mô tả tọa độ trên cơ sở 0 của các mẫu độ sáng ở ngoài cùng bên phải của ô liên quan đến mẫu con khi các cờ bằng 2 và `vcl_idc` bằng 1 hoặc 2, trong các đơn vị được suy ra từ `log2_min_luma_ctb` như được mô tả ở trên.

`ctb_y` mô tả tọa độ trên cơ sở 0 của các mẫu độ sáng ở dưới cùng của ô liên quan đến mẫu con khi các cờ bằng 2 và `vcl_idc` bằng 1 hoặc 2, trong các đơn vị được suy ra từ `log2_min_luma_ctb` như được mô tả ở trên.

`VclNalUnitType` chỉ ra giá trị `nal_unit_type` của các đơn vị NAL VCL trong mẫu con. Tất cả các đơn vị NAL VCL trong mẫu con có cùng giá trị `nal_unit_type`.

#### 9.5.9 Xử lý các mẫu không kết xuất

Phần mô tả trong 8.4.9 áp dụng với “HEVC” được thay bằng “SHVC”, và mẫu không kết xuất được định nghĩa là mẫu trong đó (các) ảnh của lớp xuất đích có giá trị `pic_output_flag` bằng 0. Khi đơn vị truy cập chứa một số hình ảnh được mã hóa có `pic_output_flag` bằng 1 và một số hình ảnh được mã hóa khác có `pic_output_flag` bằng 0, ít nhất hai rãnh phải được sử dụng để lưu trữ dòng này, để tất cả các hình ảnh được mã hóa trong mỗi mẫu có cùng giá trị `pic_output_flag` trong mỗi rãnh.

## 10 Dòng MV-HEVC cơ bản và các định nghĩa về mẫu

### 10.1 Giới thiệu

Mệnh đề này mô tả định dạng lưu trữ dữ liệu MV-HEVC. Mệnh đề này mở rộng các định nghĩa về định dạng lưu trữ HEVC trong mệnh đề 8.

Định dạng tập tin để lưu trữ nội dung MV-HEVC, như được định nghĩa trong mệnh đề này và các Phụ lục từ A đến D, sử dụng các khả năng đã có của định dạng tập tin phương tiện cơ sở ISO và định dạng tập tin HEVC đơn giản (tức là, định dạng tập

tin được mô tả trong mệnh đề 8). Ngoài ra, các cấu trúc hoặc phiên bản mở rộng sau, ngoài những cái khác, để hỗ trợ các đặc điểm dành riêng cho SHVC được sử dụng.

Bộ gộp:

cấu trúc có thể thực hiện việc tạo nhóm đơn vị NAL có thể mở rộng được một cách hiệu quả bằng cách biến đổi các mẫu đơn vị NAL không theo quy tắc thành các mẫu đơn vị dữ liệu được gộp theo quy tắc.

Bộ trích:

cấu trúc có thể tách hiệu quả các đơn vị NAL ra khỏi các rãnh khác không phải rãnh chứa dữ liệu phương tiện.

Khả năng tương thích với HEVC:

một điều khoản để lưu trữ một dòng bit MV-HEVC theo cách thức tương thích với HEVC, để lớp cơ sở tương thích với HEVC có thể được sử dụng bởi thiết bị đọc tương thích với định dạng tập tin HEVC đơn giản bất kỳ.

Hỗ trợ cho MV-HEVC bao gồm một số công cụ và có các "mô hình" khác nhau về việc sử dụng chúng như thế nào. Cụ thể, dòng MV-HEVC có thể được đặt trong các rãnh theo một số cách, trong số các cách này có cách sau:

1. tất cả các khung hình trong một rãnh, được đánh dấu theo nhóm mẫu;
2. mỗi khung hình trên rãnh của nó, được đánh dấu trong các mục nhập mẫu;
3. một thể lai, một rãnh chứa tất cả các khung hình, và một hoặc nhiều rãnh một khung hình chỉ chứa một khung hình có thể được mã hóa một cách độc lập;
4. các điểm hoạt động được mong đợi mỗi điểm trong một rãnh (ví dụ, cơ sở HEVC, cặp hình nổi, màn hình đa khung hình).

Định dạng tập tin theo MV-HEVC cho phép lưu trữ một hoặc nhiều khung hình vào một rãnh, tương tự với sự hỗ trợ đối với SHVC trong mệnh đề 9. Việc lưu trữ nhiều khung hình trên một rãnh có thể được sử dụng, ví dụ, khi nhà cung cấp nội dung muốn cung cấp dòng bit đa khung hình không dự định để tạo tập con hoặc khi dòng bit được tạo ra cho các tập khung hình kết xuất xác định trước (như 1, 2, 5, hoặc 9 khung hình) trong đó các rãnh có thể được tạo ra. Nếu nhiều hơn một khung hình được lưu trữ trong một rãnh và có một vài rãnh (nhiều hơn một) biểu diễn dòng bit MV-HEVC, việc sử dụng cơ chế tạo nhóm được khuyến cáo.

Khi dòng bit MV-HEVC được biểu diễn bằng nhiều rãnh và bộ đọc sử dụng điểm hoạt động chứa dữ liệu trong nhiều rãnh, bộ đọc phải tái tạo các đơn vị truy cập

MV-HEVC trước khi cho chúng đi qua bộ giải mã MV-HEVC. Điểm hoạt động theo MV-HEVC có thể được biểu diễn một cách tường minh bởi một rãnh, tức là đơn vị truy cập được tái tạo một cách đơn giản bằng cách tách tất cả các đơn vị của bộ trích và bộ gộp của một mẫu. Nếu số lượng các điểm hoạt động lớn, có thể tốn nhiều không gian và không thể thực hiện được việc tạo một rãnh cho mỗi điểm hoạt động. Trong trường hợp này, các đơn vị truy cập MV-HEVC được tái tạo như được mô tả trong mục 10.5.2. Bản ghi cấu hình bộ giải mã theo MV-HEVC chứa trường chỉ ra các mẫu liên quan sử dụng việc tái tạo đơn vị truy cập hiện hoặc ẩn (xem trường `explicit_aud_track`).

## 10.2 Cấu trúc của rãnh theo MV-HEVC

Các dòng MV-HEVC được lưu trữ theo 8.2, với định nghĩa sau về dòng video MV-HEVC cơ bản:

- Dòng video MV-HEVC cơ bản chứa tất cả các đơn vị NAL liên quan đến việc mã hóa video (tức là, các đơn vị NAL chứa dữ liệu video hoặc báo hiệu cấu trúc video) và có thể chứa các đơn vị NAL không liên quan đến việc mã hóa video như thông báo SEI và đơn vị NAL dấu cách đơn vị truy cập. Ngoài ra, bộ gộp (xem A.2) hoặc bộ trích (xem A.3) có thể có mặt. Bộ gộp và bộ trích được xử lý như được mô tả trong chuẩn quốc tế (ví dụ, không được đặt trực tiếp trong bộ đệm kết xuất trong khi truy cập tập tin). Các đơn vị NAL khác không bị cấm tuyệt đối có thể có mặt và nếu chúng không được nhận dạng chúng nên được bỏ qua (ví dụ, không đặt trong bộ đệm kết xuất trong quá trình truy cập tập tin).

Các dòng MV-HEVC sẽ không được lưu trữ bằng cách sử dụng dòng tập tham số liên quan, khi cần.

Có thể có các đơn vị NAL VCL có `nuh_layer_id` bằng 0, đơn vị NAL VCL có `nuh_layer_id` lớn hơn 0, và đơn vị NAL không VCL khác có mặt trong dòng video MV-HEVC cơ bản. Ngoài ra, có thể có các đơn vị NAL bộ gộp và đơn vị NAL bộ trích có mặt trong dòng video MV-HEVC cơ bản.

## 10.3 Sử dụng định dạng tập tin HEVC đơn giản

Định dạng tập tin MV-HEVC là phiên bản mở rộng của định dạng tập tin HEVC đơn giản được định nghĩa trong mệnh đề 8.

## 10.4 Định nghĩa mẫu và cấu hình

### 10.4.1 Giới thiệu

Mẫu MV-HEVC: Mẫu MV-HEVC cũng là một đơn vị truy cập như được định nghĩa trong phụ lục F của ISO/IEC 23008-2.

#### 10.4.2 Thứ tự tiêu chuẩn và các giới hạn

##### 10.4.2.1 Các giới hạn

Các giới hạn sau áp dụng cho dữ liệu MV-HEVC ngoài các yêu cầu trong 8.3.2.

- Đơn vị NAL VCL: Tất cả các đơn vị NAL VCL trong một đơn vị truy cập được chứa trong mẫu mà thời gian hợp thành của nó là thời gian hình ảnh được biểu diễn bởi đơn vị truy cập. Mẫu MV-HEVC chứa ít nhất một đơn vị NAL VCL.
- Bộ gộp/bộ trích: Thứ tự của tất cả các đơn vị NAL trong bộ gộp hoặc được tham chiếu bởi bộ trích chính xác là thứ tự giải mã nếu như các đơn vị NAL có mặt trong mẫu không chứa bộ gộp/bộ trích. Sau khi xử lý bộ gộp hoặc bộ trích, tất cả các đơn vị NAL phải theo thứ tự giải mã có hiệu lực như được mô tả trong ISO/IEC 23008-2.

##### 10.4.2.2 Bản ghi cấu hình bộ giải mã

Khi bản ghi cấu hình bộ giải mã được xác định trong 8.3.3.1 được sử dụng cho dòng có thể được diễn dịch là dòng MV-HEVC hoặc dòng HEVC, bản ghi cấu hình bộ giải mã HEVC sẽ phản ánh các đặc tính của khung hình cơ sở tương thích với HEVC, ví dụ, nó chỉ chứa tập tham số cần thiết để giải mã khung hình cơ sở HEVC.

Bản ghi cấu hình bộ giải mã theo MVHEVC (MVHEVCDecoderConfigurationRecord) về mặt cấu trúc giống với bản ghi cấu hình bộ giải mã theo HEVC (HEVCDecoderConfigurationRecord). Cú pháp là như sau:

```
aligned(8) class MVHEVCDecoderConfigurationRecord {
    // same fields as in HEVCDecoderConfigurationRecord syntax
    structure
}
```

Ngữ nghĩa của các trường trong bản ghi cấu hình bộ giải mã theo MVHEVC là giống như được xác định đối với bản ghi cấu hình bộ giải mã theo HEVC.

##### 10.4.3 Mẫu đồng bộ hóa

Mẫu MV-HEVC được xem là mẫu đồng bộ hóa nếu mỗi hình ảnh được mã hóa trong đơn vị truy cập là hình ảnh IRAP không có hình ảnh RASL, như được định nghĩa trong ISO/IEC 23008-2. Các mẫu đồng bộ hóa được ghi bởi bảng mẫu đồng bộ

hóa, và có thể còn được ghi bởi nhóm mẫu mẫu đồng bộ hóa và nhóm mẫu 'rap' được định nghĩa giống như trong SHVC.

#### 10.4.4 Hộp mẫu có thể dùng và độc lập

Nếu được sử dụng trong một rãnh tương thích với cả HEVC và MV-HEVC, thì điều quan tâm là các câu lệnh là đúng bất kể tập con có hiệu lực của dữ liệu SHVC (có thể chỉ có dữ liệu HEVC) được sử dụng. Giá trị 'chưa biết' (giá trị 0 của các trường sample-depends-on, sample-is-depended-on, và sample-has-redundancy) có thể là cần thiết nếu thông tin thay đổi.

#### 10.4.5 Nhóm mẫu ở các điểm khôi phục truy cập ngẫu nhiên và các điểm truy cập ngẫu nhiên

Đối với dữ liệu video được mô tả bởi mục nhập mẫu thuộc kiểu 'hvc1', 'hev1', 'hvc2', hoặc 'hev2,' nhóm mẫu khôi phục truy cập ngẫu nhiên và nhóm mẫu điểm truy cập ngẫu nhiên lần lượt nhận dạng các điểm khôi phục truy cập ngẫu nhiên và các điểm truy cập ngẫu nhiên, cho cả bộ giải mã HEVC và bộ giải mã MV-HEVC (nếu có) hoạt động trên toàn bộ dòng bit.

Đối với dữ liệu video được mô tả bởi kiểu mục nhập MV-HEVC nhóm mẫu khôi phục truy cập ngẫu nhiên nhận dạng sự khôi phục truy cập ngẫu nhiên trên toàn bộ dòng bit MV-HEVC và nhóm mẫu điểm truy cập ngẫu nhiên nhận dạng các điểm truy cập ngẫu nhiên trên toàn bộ dòng bit MV-HEVC.

### 10.5 Suy ra từ định dạng tập tin phương tiện cơ sở ISO

#### 10.5.1 Cấu trúc của rãnh theo MV-HEVC

Dòng video đa khung hình được biểu diễn bằng một hoặc nhiều rãnh video trong tập tin. Mỗi rãnh biểu diễn một hoặc nhiều khung hình của dòng này.

Có một tập tối thiểu gồm một hoặc nhiều rãnh, khi ghép lại với nhau, chứa tập thông tin mã hóa hoàn chỉnh. Tất cả các rãnh này có cờ "complete\_representation" thiết lập trong tất cả các mục nhập mẫu của chúng. Nhóm các rãnh này tạo ra thông tin mã hóa hoàn chỉnh được gọi là "tập con hoàn chỉnh".

Điểm hoạt động thấp nhất là một điểm trong số tất cả các điểm hoạt động chứa đơn vị NAL có nuh\_layer\_id chỉ bằng 0 và TemporalID chỉ bằng 0. Rãnh chứa điểm hoạt động thấp nhất được gọi là "rãnh khung hình cơ sở." Tất cả các rãnh khác là một phần của cùng một dòng được liên kết với rãnh cơ sở này bằng phương tiện tham chiếu rãnh theo kiểu 'sbas' (cơ sở khung hình).

Tất cả các rãnh có cùng chung rãnh khung hình cơ sở phải có thang thời gian giống như rãnh khung hình cơ sở.

Nếu khung hình được biểu diễn bằng một rãnh sử dụng khung hình khác được biểu diễn bằng rãnh khác làm tham chiếu dự đoán liên khung hình, tham chiếu rãnh thuộc kiểu 'scal' sẽ được bao gồm trong rãnh tham chiếu đến rãnh nguồn để dự đoán liên khung hình.

Nếu áp dụng các soạn thảo cho các rãnh chứa thành phần khung hình của dòng bit MV-HEVC, danh sách soạn thảo phải thống nhất trên tất cả các rãnh chịu ảnh hưởng của các soạn thảo này.

#### 10.5.2 Tái tạo đơn vị truy cập

Để tái tạo đơn vị truy cập từ các mẫu của một hoặc nhiều rãnh MV-HEVC, khung hình kết xuất đích có thể cần được xác định trước tiên.

Hình ảnh cần thiết để giải mã các khung hình kết xuất đích xác định có thể được lấy từ các ký hiệu nhận dạng khung hình tham chiếu có trong hộp ký hiệu nhận dạng khung hình hoặc các tham chiếu rãnh 'scal'.

Nếu một số rãnh chứa dữ liệu cho đơn vị truy cập, việc sắp hàng các mẫu tương ứng trong các rãnh được thực hiện trong thời gian giải mã, tức là, sử dụng bảng thời gian mẫu mà không cần quan tâm đến danh sách soạn thảo.

Đơn vị truy cập được tái tạo từ các mẫu tương ứng trong các rãnh cần thiết để sắp xếp các đơn vị NAL của chúng theo thứ tự phù hợp với chuẩn ISO/IEC 23008-02. Thứ tự sau cùng cấp những chỉ dẫn của thủ tục này để tạo ra một đơn vị truy cập thích hợp:

- Tất cả các đơn vị NAL tập tham số (của các rãnh tập tham số liên quan và của các rãnh dòng cơ sở liên quan).
- Tất cả các đơn vị NAL SEI (của các rãnh tập tham số liên quan và của các rãnh dòng cơ sở liên quan).
- Thành phần khung hình theo thứ tự giá trị chỉ số thứ tự khung hình tăng dần. Các đơn vị NAL trong một thành phần khung hình là theo thứ tự xuất hiện trong mẫu.

#### 10.5.3 Mục nhập mẫu

##### 10.5.3.1 Các hộp cho mục nhập mẫu

##### 10.5.3.1.1 Hộp ký hiệu nhận dạng khung hình

10.5.3.1.1.1 Định nghĩa:

Kiểu hộp: 'vwid'

Bộ lưu trữ: Mục nhập mẫu ('hev1', 'hvc1', 'hev2', 'hvc2', 'mhc1', 'mhv1')  
hoặc MultiviewGroupEntry

Bắt buộc: Có (cho các mục nhập mẫu)

Số lượng: Chính xác một (cho các mục nhập mẫu)

Khi có trong mục nhập mẫu, hộp này chỉ ra khung hình có trong rãnh. Hộp này cũng chỉ ra chỉ số thứ tự khung hình đối với mỗi khung hình được lập danh sách. Ngoài ra, hộp cũng bao gồm các giá trị nhỏ nhất và lớn nhất của temporal\_id có trong rãnh khi hộp ký hiệu nhận dạng khung hình có trong mục nhập mẫu. Hơn nữa, hộp này chỉ ra các khung hình được tham chiếu cần thiết để giải mã các khung hình có trong rãnh.

10.5.3.1.1.2 Cú pháp

class ViewIdentifierBox extends FullBox ('vwid', version=0, flags)

```
{
    unsigned int(2) reserved6 = 0;
    unsigned int(3) min_temporal_id;
    unsigned int(3) max_temporal_id;
    unsigned int(16) num_views;
    for (i=0; i<num_views; i++) {
        unsigned int(6) reserved1 = 0;
        unsigned int(6) layer_id[i];
        unsigned int(10) view_id[i];
        unsigned int(2) base_view_type;
        for (j = 0; j < layer_id[i]; j++) {
            unsigned int(1) depdent_layer[i][j];
        }
    }
}
```

10.5.3.1.1.3 Ngữ nghĩa

min\_temporal\_id, max\_temporal\_id lần lượt là giá trị nhỏ nhất và lớn nhất của phần tử cú pháp temporal\_id có trong phiên bản mở rộng của tiêu đề đơn vị NAL của

các đơn vị NAL được ánh xạ đối với rãnh hoặc lớp khi hộp ký hiệu nhận dạng khung hình có trong mục nhập mẫu. Đối với các dòng AVC, giá trị này có trong hoặc có thể có trong đơn vị NAL tiền tố.

num\_views, khi hộp ký hiệu nhận dạng khung hình có mặt trong mục nhập mẫu, chỉ ra số lượng khung hình có trong rãnh.

layer\_id[i] chỉ ra giá trị của phần tử cú pháp nuh\_layer\_id trong tiêu đề đơn vị NAL của một lớp có trong rãnh khi hộp ký hiệu nhận dạng khung hình có trong mục nhập mẫu.

view\_id chỉ ra ký hiệu nhận dạng khung hình của lớp thứ i có nuh\_layer\_id bằng layer\_id[i], như được mô tả trong phụ lục F của ISO/IEC 23008-2.

base\_view\_type chỉ ra khung hình có phải là khung hình cơ sở không (ảo hay không). Lấy các giá trị sau:

- 0 chỉ ra khung hình không phải là khung hình cơ sở cũng không phải là khung hình cơ sở ảo.
- 1 được sử dụng để đánh dấu khung hình cơ sở không ảo của dòng bit MV-HEVC.
- 2 là giá trị dự trữ và không được sử dụng.
- 3 chỉ ra rằng khung hình có view\_id[i] là khung hình cơ sở ảo. Khung hình không phải khung hình cơ sở được mã hóa độc lập tương ứng có view\_id[i] nằm ở trong rãnh khác. Khi base\_view\_type bằng 3, num\_ref\_views tiếp theo sẽ bằng 0.

depdent\_layer[i][j] chỉ ra lớp thứ j có nuh\_layer\_id bằng j có thể là lớp tham chiếu trực tiếp hoặc gián tiếp của lớp có nuh\_layer\_id bằng layer\_id[i]. Khi hộp ký hiệu nhận dạng khung hình có trong mục nhập mẫu, được khuyến cáo chỉ ra các khung hình được tham chiếu trong cùng một mục nhập mẫu.

#### 10.5.3.2 Định nghĩa về mục nhập mẫu

Kiểu mục nhập mẫu: 'hvc2', 'hev2', 'mhc1', 'mhv1', 'mhcC',

bộ lưu trữ: Hộp mô tả mẫu ('std')

Bắt buộc: Một trong số các hộp 'hvc1', 'hev1', 'hvc2', 'hev2', 'mhc1', hoặc 'mhv1' là bắt buộc

Số lượng: Một hoặc nhiều mục nhập mẫu có thể có mặt

Nếu một dòng MV-HEVC cơ bản chứa khung hình cơ sở tương thích với HEVC có thể sử dụng, mục nhập mẫu trực quan HEVC ('hvc1', 'hev1', 'hvc2', 'hev2') sẽ được sử dụng. Do vậy, mục nhập ban đầu chứa hộp cấu hình HEVC, có thể tiếp đến là hộp cấu hình MV-HEVC như được xác định dưới đây. Hộp cấu hình HEVC ghi lại đặc tính, mức, và có thể là tập tham số liên quan đến khung hình cơ sở tương thích với HEVC như được xác định bởi bản ghi cấu hình bộ giải mã HEVC. Hộp cấu hình MV-HEVC ghi lại đặc tính, mức, và thông tin về tập tham số liên quan đến toàn bộ dòng chứa các khung hình không phải là cơ sở như được xác định bởi bản ghi cấu hình bộ giải mã MVHEVC, được lưu trữ trong hộp cấu hình MVHEVC.

Đối với tất cả các mục nhập mẫu 'hvc1,' 'hev1,' 'hvc2,' 'hev2,' trường chiều rộng và chiều cao trong mục nhập mẫu ghi lại lớp cơ sở HEVC. Đối với một mục nhập mẫu theo MV-HEVC ('mhc1', 'mhv1'), chiều rộng và chiều cao ghi lại độ phân giải thu được bằng cách giải mã khung hình đơn bất kỳ của toàn bộ dòng.

Nếu dòng cơ bản MV-HEVC không chứa khung hình cơ sở HEVC có thể sử dụng, mục nhập mẫu trực quan MV-HEVC ('mhc1', 'mhv1') sẽ được sử dụng. Mục nhập mẫu trực quan MV-HEVC chứa một hộp cấu hình MV-HEVC, như được xác định dưới đây. Mục nhập này chứa bản ghi cấu hình bộ giải mã MV-HEVC, như được xác định trong chuẩn quốc tế này.

Trường lengthSizeMinusOne trong cấu hình MV-HEVC và HEVC trong mục nhập mẫu xác định bất kỳ có cùng giá trị.

Những yêu cầu đối với kiểu mục nhập mẫu 'hvc1' và 'hev1' như được mô tả trong mục 6.5.3.1.1 cũng áp dụng ở đây.

Hộp cấu hình MVHEVC có thể có mặt trong mục nhập mẫu 'hvc1', 'hev1', 'hvc2', 'hev2'. Trong trường hợp này, định nghĩa về HEVCMVHEVCSampleEntry or HEVC2MVHEVCSampleEntry dưới đây lần lượt được áp dụng.

Tên bộ nén trong nhóm mục nhập mẫu trực quan cơ sở chỉ ra tên của bộ nén được sử dụng, có giá trị “mã hóa \014MV-HEVC” được khuyến cáo (\016 là 14, chiều dài của vòng “mã hóa MV-HEVC” theo byte).

Các tập tham số cần thiết để giải mã đơn vị NAL có mặt trong dữ liệu mẫu của dòng video, trực tiếp hoặc bằng cách tham chiếu từ bộ trích, có mặt trong cấu hình bộ giải mã của dòng video đó hoặc trong dòng tập tham số liên quan (nếu được sử dụng).

Bảng dưới đây thể hiện tất cả các khả năng sử dụng có thể có mục nhập mẫu cho một rãnh video khi dòng MV-HEVC cơ bản được lưu trữ trong một hoặc nhiều rãnh, cấu hình và công cụ MV-HEVC.

Bảng 14 – Sử dụng các mục nhập mẫu đối với các rãnh HEVC và MV-HEVC

| Tên mục nhập mẫu   | cùng với bản ghi cấu hình | Ý nghĩa                                                                                                                                                                                                                                                                     |
|--------------------|---------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 'hvc1' hoặc 'hev1' | Chỉ có cấu hình HEVC      | Rãnh HEVC đơn giản không có các đơn vị NAL có nuh_layer_id lớn hơn 0; bộ trích và bộ gộp không có mặt.                                                                                                                                                                      |
| 'hvc1' hoặc 'hev1' | Cấu hình HEVC và MV-HEVC  | Rãnh MV-HEVC với cả đơn vị NAL có nuh_layer_id bằng 0 và đơn vị NAL có nuh_layer_id lớn hơn 0; bộ trích và bộ gộp có thể có mặt; bộ trích không tham chiếu các đơn vị NAL có nuh_layer_id bằng 0; Bộ gộp không chứa nhưng tham chiếu các đơn vị NAL có nuh_layer_id bằng 0. |
| 'hvc2' hoặc 'hev2' | Chỉ có cấu hình HEVC      | Rãnh HEVC đơn giản không chứa các đơn vị NAL có giá trị nuh_layer_id lớn hơn 0; bộ trích có thể có mặt và được sử dụng để tham chiếu các đơn vị NAL; bộ gộp có thể có mặt để chứa và tham chiếu các đơn vị NAL.                                                             |
| 'hvc2' hoặc 'hev2' | Cấu hình HEVC và MV-HEVC  | Rãnh MV-HEVC với cả đơn vị NAL có nuh_layer_id bằng 0 và đơn vị NAL có nuh_layer_id lớn hơn 0; bộ trích và bộ gộp có thể có mặt; bộ trích tham chiếu đơn vị NAL bất kỳ; Bộ gộp không chứa và tham chiếu đơn vị NAL bất kỳ.                                                  |
| 'mhc1' hoặc 'mhv1' | Cấu hình MV-HEVC          | Rãnh MV-HEVC không chứa các đơn vị NAL có giá trị nuh_layer_id bằng 0; bộ trích có thể có mặt và được sử dụng để tham chiếu các đơn vị NAL; bộ gộp có thể                                                                                                                   |

|  |  |                                              |
|--|--|----------------------------------------------|
|  |  | có mặt để chứa và tham chiếu các đơn vị NAL. |
|--|--|----------------------------------------------|

Mục nhập mẫu mvhevc-type trong phần dưới đây là một mục nhập mẫu trong số {mhv1, mhc1}.

#### 10.5.3.3 Cú pháp

```
class MVHEVCConfigurationBox extends Box('mhcC') {
    MVHEVCDecoderConfigurationRecord() MVHEVCConfig;
}

class HEVCMVHEVCSESampleEntry() extends HEVCSESampleEntry() {
    ViewIdentifierBox      view_identifiers;    // optional
    MVHEVCConfigurationBox mvhevconfig;        // optional
}

class HEVC2MVHEVCSESampleEntry() extends HEVC2SESampleEntry() {
    ViewIdentifierBox      view_identifiers;    // optional
    MVHEVCConfigurationBox mvhevconfig;        // optional
}

// Use this if the track is NOT HEVC compatible
class MVHEVCSESampleEntry() extends VisualSampleEntry(mvhevc-type) {
    MVHEVCConfigurationBox mvhevconfig;        // mandatory
    ViewIdentifierBox      view_identifiers;    // mandatory
    MPEG4BitRateBox        bitrate;            // optional
    MPEG4ExtensionDescriptorsBox descr;        // optional
}
```

#### 10.5.4 Định nghĩa về mẫu con dùng cho MV-HEVC

Định nghĩa về mẫu con dùng cho MV-HEVC được định nghĩa giống như mẫu con được định nghĩa cho SHVC.

#### 10.5.5 Xử lý các mẫu không kết xuất

Việc xử lý các mẫu không kết xuất cho MV-HEVC được mô tả giống như việc xử lý được mô tả cho SHVC.

Những sửa đổi đối với Phụ lục A được thể hiện dưới đây.

## Phụ lục A (tiêu chuẩn)

## Các cấu trúc trong dòng

## A.1 Giới thiệu

Bộ gộp và bộ trích là các cấu trúc bên trong của định dạng tập tin có thể phân các đơn vị NAL thành nhóm hiệu quả hoặc tách đơn vị NAL ra khỏi rãnh khác.

Bộ gộp và bộ trích sử dụng cú pháp đơn vị NAL. Các cấu trúc này được xem là các đơn vị NAL trong ngữ cảnh cấu trúc mẫu. Khi truy cập mẫu, bộ gộp phải được loại bỏ (bỏ các đơn vị NAL được chứa hoặc tham chiếu của chúng) và bộ trích phải được thay bằng dữ liệu mà chúng tham chiếu. Bộ gộp và bộ trích không cần có mặt trong dòng ở ngoài định dạng tập tin.

Các cấu trúc này sử dụng kiểu đơn vị NAL được lưu trữ cho lớp ứng dụng/vận chuyển theo ISO/IEC 14496-10 hoặc ISO/IEC 23008-2.

LUU Ý Dưới đây là của chuẩn ISO/IEC 14496-10:

“LUU Ý – kiểu đơn vị NAL 0 và 24..31 có thể được sử dụng như được xác định bởi ứng dụng này. Không có quá trình giải mã đối với các giá trị nal\_unit\_type này được mô tả trong Recommendation | International Standard”.

LUU Ý Dưới đây là của chuẩn ISO/IEC 23008-2:

"LUU Ý 1 – kiểu đơn vị NAL trong phạm vi UNSPEC48..UNSPEC63 có thể được sử dụng như được xác định bởi ứng dụng này. Không có quá trình giải mã đối với các giá trị nal\_unit\_type này được mô tả trong bản mô tả này. Do các ứng dụng khác nhau có thể sử dụng các kiểu đơn vị NAL này cho các mục đích khác nhau, cần thực hiện lưu ý cụ thể trong việc thiết kế các bộ mã hóa tạo ra các đơn vị NAL có các giá trị nal\_unit\_type này và trong việc thiết kế các bộ mã hóa diễn dịch nội dung của các đơn vị NAL có các giá trị nal\_unit\_type này".

## A.2 Bộ gộp

## A.2.1 Định nghĩa:

Mệnh đề này mô tả bộ gộp, làm cho các mục nhập NALU-map-group nhất quán và được lặp lại. (xem Phụ lục B).

Bộ gộp được sử dụng để lập nhóm các đơn vị NAL của cùng một mẫu.

Để lưu trữ video ISO/IEC 14496-10, quy tắc sau được áp dụng:

- Các bộ gộp sử dụng tiêu đề đơn vị NAL giống như các đơn vị NAL VCL SVC hoặc NAL VCL MVC, nhưng có giá trị kiểu đơn vị NAL khác nhau.
- Khi `svc_extension_flag` của cú pháp đơn vị NAL (được mô tả trong mục 7.3.1 của ISO/IEC 14496-10) của bộ gộp bằng 1, tiêu đề đơn vị NAL của các đơn vị NAL VCL SVC được sử dụng cho các bộ gộp này. Theo cách khác, tiêu đề đơn vị NAL của các đơn vị NAL VCL MVC được sử dụng cho bộ gộp.

Để lưu trữ video ISO/IEC 23008-2, bộ gộp sử dụng tiêu đề đơn vị NAL như được mô tả trong ISO/IEC 23008-2, có cú pháp giống như đối với HEVC, SHVC, và MV-HEVC đơn giản.

Bộ gộp có thể gộp, bằng cách cộng, các đơn vị trong các bộ gộp (nằm trong khoảng kích thước được chỉ ra bởi chiều dài của chúng) và cũng có thể gộp, bằng cách tham chiếu, các đơn vị NAL đứng sau chúng (nằm trong khoảng vùng được chỉ ra bởi trường `additional_bytes` nằm trong bộ gộp). Khi dòng này được quét bằng bộ đọc tập tin AVC hoặc HEVC, chỉ có các đơn vị NAL được cộng được xem là "nằm trong" bộ gộp. Điều này cho phép bộ đọc tập tin AVC hoặc HEVC bỏ qua toàn bộ tập các đơn vị NAL không cần thiết khi chúng được gộp bằng cách cộng. Điều này cho phép bộ đọc tập tin AVC hoặc HEVC bỏ qua các đơn vị NAL cần thiết nhưng để các đơn vị này vẫn ở trong dòng khi chúng được gộp bằng cách tham chiếu.

Bộ gộp có thể được sử dụng để lập nhóm các đơn vị NAL của lớp cơ sở hoặc khung hình cơ sở. Nếu các bộ gộp được sử dụng trên rãnh 'avc1', 'hvc1', hoặc 'hev1' thì bộ gộp sẽ không sử dụng phép cộng mà tham chiếu các đơn vị NAL của lớp cơ sở hoặc khung hình cơ sở (chiều dài của bộ gộp chỉ chứa tiêu đề của nó và các đơn vị NAL được tham chiếu bởi bộ gộp được mô tả bằng `additional_bytes`).

Khi bộ gộp được tham chiếu bằng bộ trích có `data_length` bằng 0 hoặc bằng nhóm mẫu Map, bộ gộp được xử lý khi gộp cả các byte được cộng và được tham chiếu.

Bộ gộp có thể cộng hoặc tham chiếu các bộ trích. Bộ trích có thể trích ra từ bộ gộp. Bộ gộp không cộng hoặc tham chiếu bộ gộp khác một cách trực tiếp; tuy nhiên, bộ gộp có thể cộng hoặc tham chiếu bộ trích mà tham chiếu bộ gộp.

Khi quét một dòng:

- Nếu bộ gộp không được nhận dạng (ví dụ, bằng bộ đọc AVC hoặc HEVC) nó được loại bỏ dễ dàng với nội dung được gộp;

- b) nếu bộ gộp là không cần thiết (tức là, nó thuộc lớp không mong muốn) bộ gộp và nội dung của nó cả bằng cách cộng và tham chiếu được loại bỏ dễ dàng (sử dụng trường độ dài và `additional_bytes` của bộ gộp);
- c) nếu bộ gộp là cần thiết, tiêu đề của bộ gộp dễ dàng được loại bỏ và nội dung của nó được giữ lại.

Bộ gộp được lưu trữ trong một mẫu như đơn vị NAL bất kỳ khác.

Tất cả các đơn vị NAL giữ nguyên thứ tự giải mã trong bộ gộp.

#### A.2.2 Cú pháp

```
class aligned(8) Aggregator (AggregatorSize) {
    NALUnitHeader();
    unsigned int i = sizeof(NALUnitHeader());
    unsigned int((lengthSizeMinusOne+1)*8)
        additional_bytes;
    i += lengthSizeMinusOne+1;
    while (i<AggregatorSize) {
        unsigned int((lengthSizeMinusOne+1)*8)
            NALUnitLength;
        unsigned int(NALUnitLength*8) NALUnit;
        i += NALUnitLength+lengthSizeMinusOne+1;
    };
}
```

#### A.2.3 Ngữ nghĩa

Giá trị của kích thước bộ gộp biến đổi bằng kích thước của đơn vị NAL của bộ gộp và hàm kích thước của (X) trả về kích thước của trường X theo byte.

Tiêu đề đơn vị NAL(): bốn byte đầu tiên của các đơn vị NAL VCL SVC và MVC, hoặc hai byte đầu tiên của các đơn vị ISO/IEC 23008-2 NAL.

`nal_unit_type` được thiết lập đối với kiểu đơn vị NAL của bộ gộp (kiểu 30 đối với video ISO/IEC 14496-10 và kiểu 48 đối với video ISO/IEC 23008-2).

Đối với bộ gộp cộng hoặc tham chiếu các đơn vị NAL SVC, áp dụng như sau.

`forbidden_zero_bit` và `reserved_three_2bits` được thiết lập như được mô tả trong ISO/IEC 14496-10.

Các trường khác (`nal_ref_idc`, `idr_flag`, `priority_id`, `no_inter_layer_pred_flag`, `dependency_id`, `quality_id`, `temporal_id`, `use_ref_base_pic_flag`, `discardable_flag`, và `output_flag`) được thiết lập như được mô tả trong A.4.

Đối với bộ gộp cộng hoặc tham chiếu các đơn vị NAL MVC, áp dụng như sau.

`forbidden_zero_bit` và `reserved_one_bit` được thiết lập như được mô tả trong ISO/IEC 14496-10.

Các trường khác (`nal_ref_idc`, `non_idr_flag`, `priority_id`, `view_id`, `temporal_id`, `anchor_pic_flag`, và `inter_view_flag`) được thiết lập như được mô tả trong A.5.

Đối với bộ gộp cộng hoặc tham chiếu các đơn vị NAL ISO/IEC 23008-2, áp dụng như sau.

`forbidden_zero_bit` được thiết lập như được mô tả trong ISO/IEC 23008-2.

Các trường khác (`nuh_layer_id` và `nuh_temporal_id_plus1`) được thiết lập như được mô tả trong A.6.

`additional_bytes`: Số lượng byte đứng sau đơn vị NAL bộ gộp này được xem như được gộp khi bộ gộp này được tham chiếu bởi bộ trích có `data_length` bằng 0 hoặc nhóm mẫu Map.

Chiều dài đơn vị NAL (`NALUnitLength`): mô tả kích thước, theo byte, của các đơn vị tiếp theo. Kích thước của trường này được mô tả cũng với trường `lengthSizeMinusOne`.

Đơn vị NAL (`NALUnit`): đơn vị NAL như được mô tả trong ISO/IEC 14496-10 hoặc ISO/IEC 23008-2, bao gồm tiêu đề đơn vị NAL. Kích thước của đơn vị NAL được mô tả bởi chiều dài đơn vị NAL.

### A.3 Bộ trích:

#### A.3.1 Định nghĩa:

Mệnh đề phụ này mô tả các bộ trích, có thể tạo nên các rãnh tách, bằng cách tham chiếu dữ liệu đơn vị NAL ra khỏi các rãnh khác.

Bộ gộp có thể cộng hoặc tham chiếu các bộ trích. Bộ trích có thể tham chiếu bộ gộp. Khi bộ trích được xử lý bằng bộ đọc tập tin cần đến bộ trích, bộ trích được thay về mặt logic bằng các byte mà nó tham chiếu đến. Các byte này phải không chứa bộ trích; bộ trích phải không tham chiếu, trực tiếp hoặc gián tiếp đến bộ trích khác.

LƯU Ý Rãnh được tham chiếu có thể chứa các bộ trích ngay cả khi không chứa dữ liệu được tham chiếu bằng bộ trích.

Bộ trích chứa lệnh để trích dữ liệu ra khỏi rãnh khác, được liên kết với rãnh này mà bộ trích có trong đó, bằng cách tham chiếu rãnh theo kiểu 'scal'.

Các byte được sao chép là một trong số đơn vị sau:

- a) Một đơn vị NAL nguyên vẹn, lưu ý rằng khi bộ gộp được tham chiếu, cả byte được cộng và tham chiếu đều được sao chép
- b) Nhiều hơn một đơn vị NAL nguyên vẹn

Trong cả hai trường hợp, các byte được trích bắt đầu bằng trường độ dài thực và tiêu đề đơn vị NAL.

Các byte được sao chép chỉ từ mẫu đơn được nhận dạng trong rãnh tham chiếu thông qua việc tham chiếu rãnh "scal" được chỉ báo. Việc sắp hàng trong thời gian giải mã, tức là, chỉ sử dụng bảng thời gian theo mẫu, tiếp đến là độ dịch chuyển được đếm trong một số mẫu. Bộ trích là một khái niệm ở mức phương tiện và do đó áp dụng cho rãnh đích trước khi danh sách soạn thảo được xem xét (Tuy nhiên, thường mong đợi rằng danh sách soạn thảo trong hai rãnh là giống nhau).

### A.3.2 Cú pháp

```
class aligned(8) Extractor () {
    NALUnitHeader();
    unsigned int(8) track_ref_index;
    signed int(8) sample_offset;
    unsigned int((lengthSizeMinusOne+1)*8)
        data_offset;
    unsigned int((lengthSizeMinusOne+1)*8)
        data_length;
}
```

### A.3.3 Ngữ nghĩa

Tiêu đề đơn vị NAL() (NALUnitHeader()): bốn byte đầu tiên của các đơn vị NAL VCL SVC và MVC, hoặc hai byte đầu tiên của các đơn vị ISO/IEC 23008-2 NAL.

nal\_unit\_type sẽ được thiết lập cho kiểu đơn vị NAL của bộ trích (kiểu 31 cho video ISO/IEC 14496-10 và kiểu 49 cho video ISO/IEC 23008-2).

Đối với bộ trích tham chiếu các đơn vị NAL SVC, áp dụng như sau.

forbidden\_zero\_bit và reserved\_three\_2bits được thiết lập như được mô tả trong ISO/IEC 14496-10.

Các trường khác (nal\_ref\_idc, idr\_flag, priority\_id, no\_inter\_layer\_pred\_flag, dependency\_id, quality\_id, temporal\_id, use\_ref\_base\_pic\_flag, discardable\_flag, và output\_flag) được thiết lập như được mô tả trong A.4.

Đối với bộ trích tham chiếu các đơn vị NAL MVC, áp dụng như sau.

forbidden\_zero\_bit và reserved\_one\_bit được thiết lập như được mô tả trong ISO/IEC 14496-10.

Các trường khác (nal\_ref\_idc, non\_idr\_flag, priority\_id, view\_id, temporal\_id, anchor\_pic\_flag, và inter\_view\_flag) được thiết lập như được mô tả trong A.5.

Đối với bộ trích tham chiếu các đơn vị NAL ISO/IEC 23008-2, áp dụng như sau.

forbidden\_zero\_bit được thiết lập như được mô tả trong ISO/IEC 23008-2.

Các trường khác (nuh\_layer\_id và nuh\_temporal\_id\_plus1) được thiết lập như được mô tả trong A.6.

track\_ref\_index mô tả chỉ số của tham chiếu rãnh theo kiểu ‘scal’ được sử dụng để tìm rãnh từ đó tách dữ liệu. Mẫu trong rãnh đó mà dữ liệu được tách từ mẫu này được sắp hàng về mặt thời gian hoặc đứng trước gần nhất theo dòng thời gian giải mã phương tiện, tức là sử dụng bảng thời gian theo mẫu, được điều chỉnh bằng độ lệch được mô tả bởi sample\_offset với mẫu chứa bộ trích. Tham chiếu của rãnh thứ nhất giá trị chỉ số 1; giá trị 0 được dự trữ.

sample\_offset xác định chỉ số tương đối của mẫu trong rãnh liên kết được sử dụng làm nguồn thông tin. Mẫu 0 là mẫu có cùng thời gian giải mã hoặc thời gian giải mã đứng trước gần nhất so với thời gian giải mã của mẫu chứa bộ trích, mẫu 1 (một) là mẫu tiếp theo, mẫu -1 (âm 1) là mẫu trước đó, .v.v..

data\_offset: Độ lệch của byte thứ nhất trong mẫu tham chiếu để sao chép. Nếu việc trích bắt đầu bằng byte dữ liệu thứ nhất trong mẫu đó, độ lệch có giá trị bằng 0. Độ lệch sẽ tham chiếu phần đầu của trường độ dài của đơn vị NAL.

data\_length: Số lượng byte để sao chép. Nếu trường này lấy giá trị 0, sau đó toàn bộ đơn vị NAL tham chiếu được sao chép (tức là, độ dài để sao chép được lấy từ trường độ dài được tham chiếu bởi độ lệch dữ liệu, được bổ sung bằng trường additional\_bytes của bộ gộp).

LƯU Ý Nếu hai rãnh sử dụng các giá trị `lengthSizeMinusOne` khác nhau, dữ liệu trích được sẽ cần phải định dạng lại để phù hợp với kích thước của trường độ dài của rãnh đích.

#### A.4 Các giá trị của tiêu đề đơn vị NAL đối với SVC

Cả bộ gộp và bộ trích sử dụng phiên bản mở rộng SVC của tiêu đề đơn vị NAL. Đơn vị NAL được trích bằng bộ trích hoặc được gộp bằng bộ gộp là tất cả các đơn vị NAL được tham chiếu hoặc cộng bằng cách kiểm tra một cách đệ quy nội dung của các đơn vị NAL của bộ gộp hoặc bộ trích

Các trường `nal_ref_idc`, `idr_flag`, `priority_id`, `temporal_id`, `dependency_id`, `quality_id`, `discardable_flag`, `output_flag`, `use_ref_base_pic_flag`, và `no_inter_layer_pred_flag` có các giá trị sau:

`nal_ref_idc` được thiết lập là giá trị cao nhất của trường này trong tất cả các đơn vị NAL được trích hoặc gộp.

`idr_flag` được thiết lập là giá trị cao nhất của trường này trong tất cả các đơn vị NAL được trích hoặc gộp.

`priority_id`, `temporal_id`, `dependency_id`, và `quality_id` lần lượt được thiết lập là giá trị thấp nhất của trường này trong tất cả các đơn vị NAL được trích hoặc gộp.

`discardable_flag` được thiết lập là 1 nếu và chỉ nếu tất cả các đơn vị NAL được trích hoặc gộp có `discardable_flag` được thiết lập là 1, và 0 theo cách khác.

`output_flag` được thiết lập là 1 nếu ít nhất một đơn vị NAL được trích hoặc gộp có cờ này được thiết lập 1, và theo cách khác được thiết lập là 0.

`use_ref_base_pic_flag` được thiết lập là 1 nếu và chỉ nếu ít nhất một trong số các đơn vị NAL được trích hoặc gộp có `use_ref_base_pic_flag` được thiết lập là 1, và 0 theo cách khác.

`no_inter_layer_pred_flag` được thiết lập là 1 nếu và chỉ nếu tất cả các đơn vị NAL VCL được trích hoặc gộp có `no_inter_layer_pred_flag` được thiết lập là 1, và 0 theo cách khác.

Nếu tập đơn vị NAL được trích hoặc gộp là rỗng, thì mỗi trường này có giá trị phù hợp với phần mô tả về tầng được ánh xạ.

LƯU Ý Các bộ gộp có thể lập nhóm các đơn vị NAL cùng với thông tin có khả năng mở rộng khác.

LUU Ý Bộ gộp có thể được sử dụng để lập nhóm các đơn vị NAL thuộc mức có khả năng mở rộng có thể không được báo hiệu bằng tiêu đề đơn vị NAL (ví dụ, các đơn vị NAL thuộc vùng quan tâm). Việc mô tả về các bộ gộp này có thể được thực hiện cùng với mô tả về tầng và các nhóm sơ đồ đơn vị NAL. Trong trường hợp này, nhiều hơn một bộ gộp với cùng một thông tin có khả năng mở rộng có thể có trong một mẫu.

LUU Ý Nếu nhiều rãnh có thể mở rộng tham chiếu cùng một dữ liệu phương tiện, bộ gộp sẽ lập nhóm các đơn vị NAL chỉ với thông tin có khả năng mở rộng giống nhau. Điều này đảm bảo rằng kiểu thu được có thể được truy cập bởi mỗi rãnh.

LUU Ý Nếu không có đơn vị NAL của một lớp cụ thể tồn tại trong một đơn vị truy cập thì bộ gộp trống (trong đó chiều dài của bộ gộp chỉ bao gồm tiêu đề và additional\_bytes bằng 0) có thể tồn tại.

#### A.5 Các giá trị của tiêu đề đơn vị NAL đối với MVC

Cả bộ gộp và bộ trích sử dụng phiên bản mở rộng MAC của tiêu đề đơn vị NAL. Đơn vị NAL được trích bằng bộ trích hoặc được gộp bằng bộ gộp là tất cả các đơn vị NAL được tham chiếu hoặc cộng bằng cách kiểm tra một cách đệ quy nội dung của các đơn vị NAL của bộ gộp hoặc bộ trích

Các trường nal\_ref\_idc, non\_idr\_flag, priority\_id, view\_id, temporal\_id, anchor\_pic\_flag, and inter\_view\_flag có các giá trị sau:

nal\_ref\_idc được thiết lập là giá trị cao nhất của trường này trong tất cả các đơn vị NAL được gộp hoặc trích.

non\_idr\_flag được thiết lập là giá trị thấp nhất của trường này trong tất cả các đơn vị NAL được gộp hoặc trích.

priority\_id và temporal\_id lần lượt được thiết lập là giá trị thấp nhất của trường này trong tất cả các đơn vị NAL được gộp hoặc trích.

view\_id được thiết lập là giá trị view\_id của đơn vị NAL VCL với chỉ số thứ tự khung hình thấp nhất trong số các đơn vị NAL VCL được gộp hoặc trích.

anchor\_pic\_flag và inter\_view\_flag lần lượt được thiết lập là giá trị cao nhất của trường này trong tất cả các đơn vị NAL được gộp hoặc trích.

Nếu tập đơn vị NAL được trích hoặc gộp là rỗng, thì mỗi trường này có giá trị phù hợp với phần mô tả về tầng được ánh xạ.

#### A.6 Các giá trị tiêu đề đơn vị đối với ISO/IEC 23008-2

Cả bộ gộp và bộ trích sử dụng tiêu đề đơn vị NAL như được mô tả trong ISO/IEC 23008-2. Đơn vị NAL được trích bằng bộ trích hoặc được gộp bằng bộ gộp là tất cả các đơn vị NAL được tham chiếu hoặc cộng bằng cách kiểm tra một cách đệ quy nội dung của các đơn vị NAL của bộ gộp hoặc bộ trích

Các trường `nuh_layer_id` và `nuh_temporal_id_plus1` được thiết lập như sau:

`nuh_layer_id` được thiết lập là giá trị thấp nhất của trường này trong tất cả các đơn vị NAL được gộp hoặc trích.

`nuh_temporal_id_plus1` được thiết lập là giá trị thấp nhất của trường này trong tất cả các đơn vị NAL được gộp hoặc trích.

Theo ví dụ khác, cấu trúc, bảng hoặc nhóm mẫu mới được xác định để ghi tất cả các đơn vị truy cập IRAP như được xác định trong Phụ lục F của MV-HEVC WD5 hoặc SHVC WD3. Theo cách khác, bảng hoặc nhóm mẫu mới được xác định để ghi tất cả các đơn vị truy cập IRAP như được xác định trong Phụ lục F của MV-HEVC WD5 hoặc SHVC WD3 ngoại trừ các đơn vị truy cập mà tất cả các hình ảnh được mã hóa là hình ảnh IRAP. Theo ví dụ khác, mục nhập nhóm mẫu của mẫu đồng bộ hóa `SyncSampleEntry` được xác định lại bao gồm `aligned_sync_flag` trong một trong số các bit bảo lưu mô tả tất cả các hình ảnh trong mẫu thuộc nhóm này là hình ảnh IDR, hình ảnh CRA hoặc hình ảnh BLA. Theo ví dụ khác, định dạng tập tin thông thường đối với SHVC và MV-HEVC được xác định bao gồm tất cả các dấu hiệu thông thường của định dạng tập tin SHVC và MV-HEVC, và chỉ có các định dạng tập tin SHVC và MV-HEVC được xác định lại để chỉ bao gồm các dấu hiệu liên quan đến phiên bản mở rộng này. Theo ví dụ khác, mục nhập mẫu siêu dữ liệu SHVC `SHVCMetadataSampleEntry` và `SHVCMetadataSampleConfigBox` được xác định và kiểu câu lệnh mẫu siêu dữ liệu `scalabilityInfoSHVCStatement` cũng được xác định.

Fig.2 là sơ đồ khối thể hiện bộ mã hóa video 20 làm ví dụ mà có thể thực hiện các kỹ thuật được mô tả trong bản mô tả này. Bộ mã hóa video 20 có thể được tạo cấu hình để kết xuất một khung hình, đa khung hình, dữ liệu mở rộng được, 3D, và các loại dữ liệu video khác. Bộ mã hóa video 20 có thể được tạo cấu hình để kết xuất video cho thực thể sau xử lý 27. Thực thể sau xử lý 27 được dự định là một ví dụ về thực thể video, như MANE hoặc thiết bị nối/soạn thảo, có thể xử lý dữ liệu video mã hóa từ bộ

mã hóa video 20. Trong một số trường hợp, thực thể sau xử lý có thể là một ví dụ về thực thể mạng. Trong một số hệ thống mã hóa video, thực thể sau xử lý 27 và bộ mã hóa video 20 có thể là các bộ phận của các thiết bị riêng biệt, còn theo trường hợp khác, chức năng được mô tả liên quan đến thực thể sau xử lý 27 có thể được thực hiện bởi cùng một thiết bị bao gồm bộ mã hóa 20. Thực thể sau xử lý 27 có thể là thiết bị video. Theo một số ví dụ, thực thể sau xử lý 27 có thể giống như thiết bị tạo tập tin 34 trên Fig.1.

Bộ mã hóa video 20 có thể mã hóa trong ảnh hoặc liên ảnh khối video trong các lát video. Mã hóa trong ảnh dựa vào dự đoán không gian để làm giảm hoặc loại bỏ phần dư không gian trong video trong khung video xác định. Mã hóa liên ảnh dựa vào dự đoán thời gian để giảm hoặc loại bỏ phần dư thời gian trong video trong khung hoặc hình ảnh hiện thời liên kế của chuỗi video. Chế độ trong ảnh (chế độ I) có thể chỉ đến chế độ nén dựa trên không gian bất kỳ. Chế độ liên ảnh, như chế độ dự đoán một chiều (chế độ P) hoặc dự đoán hai chiều (chế độ B), có thể chỉ chế độ nén dựa vào thời gian bất kỳ.

Theo ví dụ trên Fig.2, bộ mã hóa video 20 bao gồm đơn vị phân chia 35, đơn vị xử lý dự đoán 41, đơn vị lọc 63, bộ nhớ hình ảnh tham chiếu 64, bộ cộng 50, đơn vị xử lý biến đổi 52, đơn vị xử lý lượng tử hóa 54 và đơn vị mã hóa entropy 56. Đơn vị xử lý dự đoán 41 bao gồm đơn vị ước lượng chuyển động 42, đơn vị bù chuyển động 44 và đơn vị xử lý dự đoán trong ảnh 46. Đối với việc tái tạo khối video, bộ mã hóa video 20 còn bao gồm đơn vị lượng tử hóa ngược 58, đơn vị xử lý biến đổi ngược 60, và bộ cộng 62. Đơn vị lọc 63 được dự định là một hoặc nhiều bộ lọc vòng lặp như bộ lọc tách khối, bộ lọc vòng lặp tương thích (adaptive loop filter-ALF), và bộ lọc dịch chuyển tương thích mẫu (sample adaptive offset-SAO). Mặc dù đơn vị lọc 63 được thể hiện trên Fig.2 là bộ lọc vòng lặp, theo cấu hình khác, đơn vị lọc 63 có thể được thực hiện như bộ lọc sau vòng lặp.

Bộ nhớ dữ liệu video của bộ mã hóa video 20 có thể lưu trữ dữ liệu video được mã hóa bởi các thành phần của bộ mã hóa video 20. Dữ liệu video được lưu trữ trong bộ nhớ dữ liệu video có thể được thu, ví dụ, từ nguồn video 18. Bộ nhớ hình ảnh tham chiếu 64 có thể là bộ nhớ hình ảnh tham chiếu lưu trữ dữ liệu video tham chiếu để sử dụng trong việc mã hóa dữ liệu video của bộ mã hóa video 20, ví dụ, các chế độ mã hóa trong ảnh hoặc liên ảnh. Bộ nhớ dữ liệu video và bộ nhớ hình ảnh tham chiếu 64

có thể được tạo ra bởi thiết bị nhớ bất kỳ, như bộ nhớ truy cập ngẫu nhiên động (DRAM), bao gồm DRAM đồng bộ (SDRAM), RAM từ trở (MRAM), RAM điện trở (RRAM), hoặc các loại thiết bị nhớ khác. Bộ nhớ dữ liệu video và bộ nhớ hình ảnh tham chiếu 64 có thể được cung cấp bởi cùng một thiết bị nhớ hoặc các thiết bị nhớ riêng. Theo các ví dụ khác nhau, bộ nhớ dữ liệu video có thể nằm trên chip cùng các bộ phận khác của bộ mã hóa video 20, hoặc không nằm trên chip với các bộ phận đó.

Như được thể hiện trên Fig.2, bộ mã hóa video 20 nhận dữ liệu video và đơn vị phân chia 35 phân chia dữ liệu thành các khối video. Việc phân chia này cũng có thể bao gồm phân chia các lát, ô hoặc các đơn vị lớn hơn khác cũng như phân chia khối video, ví dụ, theo cấu trúc cây tứ phân của LCU và CU. Bộ mã hóa video 20 minh họa một cách tổng quát các thành phần mã hóa khối video trong lát video được mã hóa. Lát này có thể được chia thành nhiều khối video (và có thể thành tập các khối video được gọi là ô). Đơn vị xử lý dự đoán 41 có thể chọn một trong số các chế độ mã hóa có thể có như một trong số các chế độ mã hóa trong ảnh, một trong số các chế độ mã hóa liên ảnh, cho khối video hiện thời dựa trên kết quả lỗi (ví dụ, tốc độ mã hóa và độ méo ảnh). Đơn vị xử lý dự đoán 41 có thể cung cấp khối mã hóa trong ảnh hoặc liên ảnh thu được cho bộ cộng 50 để tạo ra khối dữ liệu dư và cho bộ cộng 62 để tái tạo khối hiện thời dùng để sử dụng làm hình ảnh tham chiếu.

Đơn vị xử lý dự đoán trong ảnh 46 nằm trong đơn vị xử lý dự đoán 41 có thể thực hiện mã hóa dự đoán trong ảnh khối video hiện thời so với một hoặc nhiều khối lân cận trong cùng một khung hoặc lát dưới dạng khối hiện thời được mã hóa để tạo ra sự nén không gian. Đơn vị ước lượng chuyển động 42 và đơn vị bù chuyển động 44 nằm trong đơn vị xử lý dự đoán 41 thực hiện mã hóa dự đoán liên ảnh khối video hiện thời so với một hoặc nhiều khối dự đoán trong một hoặc nhiều hình ảnh tham chiếu để tạo ra sự nén về thời gian.

Đơn vị ước lượng chuyển động 42 có thể được tạo cấu hình để xác định chế độ dự đoán liên ảnh cho lát video theo kiểu định trước cho chuỗi video. Kiểu định trước có thể biểu thị các lát video trong chuỗi là lát P, lát B hoặc lát GPB. Đơn vị ước lượng chuyển động 42 và đơn vị bù chuyển động 44 có thể được tích hợp hoàn toàn, nhưng vẫn được minh họa riêng nhằm mục đích giải thích khái niệm. Việc ước lượng chuyển động, được thực hiện bởi đơn vị ước lượng chuyển động 42, là quá trình tạo ra các vector chuyển động, các vector này ước lượng chuyển động cho các khối video. Vector

chuyển động, ví dụ, có thể chỉ báo chuyển vị của PU của khối video nằm trong khung hoặc hình ảnh video hiện thời so với khối dự đoán nằm trong ảnh tham chiếu.

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

Đơn vị ước lượng chuyển động 42 có thể tính toán vector chuyển động cho PU của khối video trong lát được mã hóa trong ảnh bằng cách so sánh vị trí của PU này với vị trí của khối dự đoán của hình ảnh tham chiếu. Hình ảnh tham chiếu có thể được chọn từ danh sách hình ảnh tham chiếu thứ nhất (Danh sách 0) hoặc danh sách hình ảnh tham chiếu thứ hai (Danh sách 1), mỗi danh sách này nhận dạng một hoặc nhiều hình ảnh tham chiếu được lưu trữ trong DPB 64. Đơn vị ước lượng chuyển động 42 gửi vector chuyển động tính toán được đến đơn vị mã hóa entropy 56 và đơn vị bù chuyển động 44.

Bù chuyển động, được thực hiện bởi đơn vị bù chuyển động 44, có thể bao gồm việc tìm nạp hoặc tạo ra khối dự đoán dựa trên vector chuyển động được xác định bằng cách ước lượng chuyển động, có thể thực hiện nội suy độ chính xác điểm ảnh con. Khi nhận vector chuyển động cho PU của khối video hiện thời, đơn vị bù chuyển động 44 có thể định vị khối dự đoán mà vector chuyển động trỏ đến khối này trong một trong số các hình ảnh tham chiếu. Bộ mã hóa video 20 tạo ra khối video dư bằng cách trừ đi giá trị điểm ảnh của khối dự đoán từ giá trị điểm ảnh của khối video hiện thời được mã hóa, tạo ra giá trị độ lệch điểm ảnh. Giá trị độ lệch điểm ảnh này tạo thành khối dữ liệu dư và có thể bao gồm thành phần độ lệch độ sáng và độ sắc. Bộ cộng 50 thể hiện thành phần hoặc các thành phần thực hiện phép trừ này. Đơn vị bù chuyển động 44 cũng có

thể tạo ra phần tử cú pháp liên quan đến các khối video và lát video để bộ giải mã video 30 sử dụng trong việc giải mã khối video của lát video.

Đơn vị xử lý dự đoán trong ảnh 46 có thể dự đoán trong ảnh đối với khối hiện thời, thay cho dự đoán liên ảnh được thực hiện bởi đơn vị ước lượng chuyển động 42 và đơn vị bù chuyển động 44, như nêu trên. Cụ thể là, đơn vị xử lý dự đoán trong ảnh 46 có thể xác định chế độ dự đoán trong ảnh sử dụng để mã hóa khối hiện thời. Theo một số ví dụ, đơn vị xử lý dự đoán trong ảnh 46 có thể mã hóa khối hiện thời bằng cách sử dụng các chế độ dự đoán trong ảnh khác nhau, ví dụ, trong các lần chạy mã hóa riêng biệt và đơn vị xử lý dự đoán trong ảnh 46 (hoặc đơn vị chọn chế độ 40, theo một số ví dụ) có thể chọn chế độ dự đoán trong ảnh thích hợp để sử dụng từ các chế độ đã thử nghiệm. Ví dụ, đơn vị xử lý dự đoán trong ảnh 46 có thể tính toán giá trị tốc độ-méo bằng cách sử dụng phân tích tốc độ-méo đối với các chế độ dự đoán trong ảnh khác nhau và chọn chế độ dự đoán trong ảnh có đặc điểm tốc độ-méo tốt nhất trong số các chế độ đã thử nghiệm. Phân tích tốc độ-méo thường xác định mức độ méo (hoặc lỗi) giữa khối mã hóa và khối gốc, chưa được mã hóa mà khối này được mã hóa để tạo ra khối mã hóa, cũng như tốc độ bit (tức là, số lượng bit) được sử dụng để tạo ra khối mã hóa. Đơn vị xử lý dự đoán trong ảnh 46 có thể tính toán tỷ lệ giữa mức độ méo và tốc độ cho các khối mã hóa khác nhau để xác định chế độ dự đoán trong ảnh nào có giá trị tốc độ-méo tốt nhất cho khối này.

Trong trường hợp bất kỳ, sau khi chọn chế độ dự đoán trong ảnh cho một khối, đơn vị xử lý dự đoán trong ảnh 46 có thể cung cấp thông tin chỉ báo chế độ dự đoán trong ảnh được chọn cho khối này đến đơn vị mã hóa entropy 56. Đơn vị mã hóa entropy 56 có thể mã hóa thông tin chỉ báo chế độ dự đoán trong ảnh được chọn theo kỹ thuật theo sáng chế. Bộ mã hóa video 20 có thể bao gồm trên dữ liệu cấu hiện dòng bit được truyền, bao gồm nhiều bảng chỉ số chế độ dự đoán trong ảnh và nhiều bảng chỉ số chế độ dự đoán trong ảnh được hiệu chỉnh (còn được gọi là bảng ánh xạ từ mã), các định nghĩa về ngữ cảnh mã hóa cho các khối khác nhau và các chỉ báo về chế độ dự đoán trong ảnh có xác suất cao nhất, bảng chỉ số chế độ dự đoán trong ảnh và bảng chỉ số chế độ dự đoán trong ảnh đã hiệu chỉnh để sử dụng cho từng ngữ cảnh.

Sau khi đơn vị xử lý dự đoán 41 tạo ra khối dự đoán cho khối video hiện thời bằng cách dự đoán liên ảnh hoặc dự đoán trong ảnh, bộ mã hóa video 20 tạo ra khối video dư bằng cách lấy khối video hiện thời trừ đi khối dự đoán. Dữ liệu video dư

trong khối dư có thể được chứa trong một hoặc nhiều TU và được sử dụng cho đơn vị xử lý biến đổi 52. Đơn vị xử lý biến đổi 52 biến đổi dữ liệu video dư thành hệ số biến đổi dư bằng cách sử dụng phép biến đổi như phép biến đổi cosin rời rạc (discrete cosine transform-DCT) hoặc phép biến đổi tương tự về mặt khái niệm. Đơn vị xử lý biến đổi 52 có thể biến đổi dữ liệu video dư từ miền điểm ảnh thành miền biến đổi, như miền tần số.

Đơn vị xử lý biến đổi 52 có thể gửi các hệ số biến đổi thu được đến đơn vị lượng tử hóa 54. Đơn vị lượng tử hóa 54 lượng tử hóa các hệ số biến đổi để giảm tiếp tốc độ bit. Quy trình lượng tử hóa có thể làm giảm độ sâu bit liên quan đến một số hoặc tất cả các hệ số biến đổi. Mức độ lượng tử hóa có thể được thay đổi bằng cách điều chỉnh tham số lượng tử hóa. Theo một số ví dụ, đơn vị lượng tử hóa 54 có thể sau đó thực hiện quét ma trận bao gồm các hệ số biến đổi đã lượng tử hóa. Theo cách khác, đơn vị mã hóa entropy 56 có thể thực hiện việc quét này.

Sau khi lượng tử hóa, đơn vị mã hóa entropy 56 mã hóa entropy phần tử cú pháp biểu diễn các hệ số biến đổi được lượng tử hóa. Ví dụ, đơn vị mã hóa entropy 56 có thể thực hiện phương pháp hoặc kỹ thuật mã hóa độ dài biến đổi tương thích ngữ cảnh (CAVLC), mã hóa số học nhị phân tương thích ngữ cảnh (CABAC), mã hóa số học nhị phân tương thích ngữ cảnh dựa trên cú pháp (SBAC), mã hóa entropy phân chia khoảng xác suất (PIPE) hoặc phương pháp hoặc kỹ thuật mã hóa entropy khác. Sau khi mã hóa entropy bằng đơn vị mã hóa entropy 56, dòng bit mã hóa có thể được truyền đến bộ giải mã video 30, hoặc lưu trữ để truyền sau hoặc phục hồi bởi bộ giải mã video 30. Đơn vị mã hóa entropy 56 cũng có thể mã hóa entropy các vector chuyển động và các phần tử cú pháp cho lát video hiện thời được mã hóa.

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

khối tham chiếu để lưu trữ trong bộ nhớ hình ảnh tham chiếu 64. Khối tham chiếu có thể được sử dụng bởi đơn vị ước lượng chuyển động 42 và đơn vị bù chuyển động 44 làm khối tham chiếu để dự đoán liên ảnh khối trong khung hoặc hình ảnh video tiếp theo.

Bộ mã hóa video 20 là một ví dụ về bộ mã hóa video được tạo cấu hình để tạo ra dữ liệu video có thể được lưu trữ bằng cách sử dụng kỹ thuật định dạng tập tin được mô tả trong bản mô tả này.

Fig.3 là sơ đồ khối thể hiện bộ giải mã video 30 ví dụ mà có thể thực hiện các kỹ thuật được mô tả trong bản mô tả này. Bộ giải mã video 30 có thể được tạo cấu hình để giải mã một khung hình, đa khung hình, dữ liệu mở rộng được, 3D, và các loại dữ liệu video khác. Theo ví dụ trên Fig.3, bộ giải mã video 30 bao gồm đơn vị giải mã entropy 80, đơn vị xử lý dự đoán 81, đơn vị xử lý lượng tử hóa ngược 86, đơn vị xử lý biến đổi ngược 88, bộ cộng 90, đơn vị lọc 91 và bộ nhớ hình ảnh tham chiếu 92. Đơn vị xử lý dự đoán 81 bao gồm đơn vị bù chuyển động 82 và đơn vị xử lý dự đoán trong ảnh 84. Bộ giải mã video 30 có thể, trong một số ví dụ, thực hiện một đường chạy giải mã thường ngược với đường chạy mã hóa được mô tả với bộ mã hóa video 20 trên Fig.2.

Bộ nhớ đệm hình ảnh được mã hóa (coded picture buffer-CPB) có thể nhận và lưu trữ dữ liệu video đã mã hóa (ví dụ, đơn vị NAL) của dòng bit. Dữ liệu video được lưu trữ trong bộ CPB 79 có thể được thu, ví dụ, từ liên kết 16, ví dụ, từ nguồn video cục bộ, như máy quay phim, thông qua truyền thông mạng có dây hoặc không dây của dữ liệu video, hoặc bằng cách truy cập vật ghi dữ liệu vật lý. CPB 79 có thể tạo bộ nhớ dữ liệu video mà lưu trữ dữ liệu video được mã hóa từ dòng bit video được mã hóa. CPB 79 có thể là bộ nhớ hình ảnh tham chiếu mà lưu trữ dữ liệu video tham chiếu để dùng để giải mã dữ liệu video bởi bộ giải mã video 30, ví dụ, các chế độ mã hóa trong ảnh hoặc liên ảnh. CPB 79 và bộ nhớ hình ảnh tham chiếu 92 có thể được tạo ra bởi thiết bị nhớ bất kỳ, như bộ nhớ truy cập ngẫu nhiên động (DRAM), bao gồm DRAM đồng bộ (SDRAM), RAM từ trở (MRAM), RAM điện trở (RRAM), hoặc các loại thiết bị nhớ khác. CPB 79 và bộ nhớ hình ảnh tham chiếu 92 có thể được cung cấp bởi cùng một thiết bị nhớ hoặc các thiết bị nhớ riêng. Trong các ví dụ, CPB 79 có thể nằm trên chip cùng các bộ phận khác của bộ giải mã video 30, hoặc không nằm trên chip so với các bộ phận đó.

Trong quá trình giải mã, bộ giải mã video 30 nhận dòng bit video mã hóa biểu diễn khối video của lát video được mã hóa và các phần tử cú pháp liên quan từ bộ mã hóa video 20. Bộ giải mã video 30 có thể nhận dòng bit video mã hóa từ thực thể mạng 29. Thực thể mạng 29 có thể là, ví dụ, máy chủ, MANE, bộ soạn thảo/bộ nối video, hoặc thiết bị như vậy khác được tạo cấu hình để thực thi một hoặc nhiều kỹ thuật được mô tả trên đây. Thực thể mạng 29 có thể bao gồm hoặc không bao gồm bộ mã hóa video, như bộ mã hóa video 20. Một số kỹ thuật được mô tả trong bản mô tả này có thể được thực thi bởi thực thể mạng 29 trước khi thực thể mạng 29 truyền dòng bit video mã hóa đến bộ giải mã video 30. Trong một số hệ thống giải mã hóa video, thực thể mạng 29 và bộ giải mã video 30 có thể là các bộ phận của các thiết bị riêng biệt, còn theo trường hợp khác, chức năng được mô tả liên quan đến thực thể mạng 29 có thể được thực hiện bởi cùng một thiết bị bao gồm bộ mã hóa 30. Thực thể mạng 29 có thể là được xem là thiết bị video. Hơn nữa, thực thể mạng 29 là thiết bị tạo tập tin 34 trên Fig.1.

Đơn vị giải mã entropy 80 của bộ giải mã 30 giải mã entropy phần tử cú pháp cụ thể của dòng bit để tạo ra các hệ số lượng tử hóa, vector chuyển động và các phần tử cú pháp khác. Đơn vị giải mã entropy 80 chuyển tiếp vector chuyển động và các phần tử cú pháp khác đến đơn vị xử lý dự đoán 81. Bộ giải mã video 30 có thể nhận phần tử cú pháp ở mức lát video và/hoặc mức khối video.

Khi lát video được mã hóa dưới dạng lát được mã hóa trong ảnh (I), đơn vị xử lý dự đoán trong ảnh 84 của đơn vị xử lý dự đoán 81 có thể tạo ra dữ liệu dự đoán cho khối video của lát video hiện thời dựa trên chế độ dự đoán trong ảnh được báo hiệu và dữ liệu từ khối được giải mã trước đó của khung hoặc hình ảnh hiện thời. Khi khung video được mã hóa dưới dạng phần được mã hóa liên ảnh (tức là, B, P hoặc GPB) đơn vị bù chuyển động 82 của đơn vị xử lý dự đoán 81 tạo ra các khối dự đoán cho khối video của lát video hiện thời dựa vào các vector chuyển động và các phần tử cú pháp khác nhận được từ đơn vị giải mã entropy 80. Các khối dự đoán có thể được tạo ra từ một trong các hình ảnh tham chiếu trong một trong các danh sách hình ảnh tham chiếu. Bộ giải mã video 30 có thể tạo ra danh sách khung tham chiếu, Danh sách 0 và Danh sách 1, bằng cách sử dụng kỹ thuật tạo mặc định dựa trên các hình ảnh tham chiếu được lưu trữ trong bộ nhớ hình ảnh tham chiếu 92.

Đơn vị bù chuyển động 82 xác định thông tin dự đoán cho khối video của lát video hiện thời bằng cách phân tích các vector chuyển động và phần tử cú pháp và sử dụng thông tin dự đoán để tạo ra khối dự đoán cho khối video hiện thời được giải mã. Ví dụ, đơn vị bù chuyển động 82 sử dụng các phần tử cú pháp nhận được để xác định chế độ dự đoán (ví dụ, dự đoán trong ảnh hoặc liên ảnh) được sử dụng để mã hóa các khối video, loại lát dự đoán liên ảnh (ví dụ, lát B, lát P hoặc lát GPB), thông tin cấu trúc cho một hoặc nhiều danh sách hình ảnh tham chiếu cho lát này, các vector chuyển động cho mỗi khối video được mã hóa liên ảnh của lát này, trạng thái dự đoán liên ảnh đối với mỗi khối video được mã hóa liên ảnh của lát này và thông tin khác để giải mã các khối video trong lát video hiện thời.

Đơn vị bù chuyển động 82 cũng có thể thực hiện nội suy dựa vào bộ lọc nội suy. Đơn vị bù chuyển động 82 có thể sử dụng các bộ lọc nội suy như bộ mã hóa video 20 sử dụng chúng trong việc mã hóa khối video để tính toán các giá trị nội suy cho các điểm ảnh dưới số nguyên của khối tham chiếu. Đơn vị bù chuyển động 82 có thể xác định các bộ lọc nội suy được sử dụng bởi bộ mã hóa video 20 theo thông tin cú pháp nhận được và sử dụng các bộ lọc nội suy để tạo ra các khối dự đoán.

Đơn vị lượng tử hóa ngược 86 lượng tử hóa ngược, tức là, giải lượng tử hóa, các hệ số biến đổi được lượng tử hóa có trong dòng bit và được giải mã bởi bộ giải mã entropy 80. Quy trình lượng tử hóa ngược có thể bao gồm việc sử dụng tham số lượng tử hóa QP được tính bởi bộ mã hóa video 20 cho mỗi khối video trong lát video để xác định mức lượng tử hóa và, tương tự, mức lượng tử hóa ngược cần được áp dụng. Đơn vị xử lý biến đổi ngược 88 áp dụng phép biến đổi ngược, ví dụ, phép biến đổi DCT ngược, phép biến đổi số nguyên ngược hoặc quá trình biến đổi ngược tương tự về mặt khái niệm, đối với các hệ số biến đổi để tạo ra khối dư trong miền điểm ảnh.

Sau khi đơn vị bù chuyển động 82 tạo ra khối dự đoán cho khối video hiện thời dựa vào vector chuyển động và các phần tử cú pháp khác, bộ giải mã video 30 tạo ra khối video giải mã bằng cách cộng khối dư từ đơn vị xử lý biến đổi ngược 88 với các khối dự đoán tương ứng được tạo ra bởi đơn vị bù chuyển động 82. Bộ cộng 90 thể hiện thành phần hoặc các thành phần thực hiện phép cộng này. Nếu muốn, bộ lọc vòng lặp (hoặc trong vòng lặp mã hóa hoặc sau vòng lặp mã hóa) có thể được sử dụng để làm trơn sự chuyển tiếp điểm ảnh hoặc theo cách khác cải thiện chất lượng video. Đơn vị lọc 91 được dự định là một hoặc nhiều bộ lọc vòng lặp như bộ lọc tách khối, bộ lọc

vòng lặp tương thích (adaptive loop filter-ALF), và bộ lọc dịch chuyển tương thích mẫu (sample adaptive offset-SAO). Mặc dù đơn vị lọc 91 được thể hiện tên Fig.3 là bộ lọc vòng lặp, theo cấu hình khác, đơn vị lọc 91 có thể được thực hiện như bộ lọc sau vòng lặp. Các khối video đã giải mã trong một khung hoặc hình ảnh xác định sau đó được lưu trữ trong bộ nhớ hình ảnh tham chiếu 92, lưu trữ các hình ảnh tham chiếu được sử dụng để bù chuyển động tiếp theo. Bộ nhớ hình ảnh tham chiếu 92 cũng lưu trữ video đã giải mã để sau đó biểu diễn trên thiết bị hiển thị, như thiết bị hiển thị 32 trên Fig.1.

Bộ giải mã 30 trên Fig.3 là một ví dụ về bộ giải hóa video được tạo cấu hình để giải mã dữ liệu video có thể được lưu trữ bằng cách sử dụng kỹ thuật định dạng tập tin được mô tả trong bản mô tả này.

Fig.4 là sơ đồ khối minh họa tập thiết bị làm ví dụ tạo thành một phần của mạng 100. Theo ví dụ này, mạng 100 bao gồm các thiết bị định tuyến 104A, 104B (thiết bị định tuyến 104) và thiết bị chuyển mã 106. Thiết bị định tuyến 104 và thiết bị chuyển mã 106 được dự định là một số lượng nhỏ các thiết bị có thể là một phần của mạng 100. Các thiết bị mạng khác như bộ ngắt mạch, ống nối ngoài, cổng nối, tường lửa, cầu nối và thiết bị tương tự khác cũng có thể được bao gồm trong mạng 100. Ngoài ra, các thiết bị mạng khác có thể được cung cấp cùng với đường mạng giữa thiết bị máy chủ 102 và thiết bị khách 108. Thiết bị máy chủ 102 có thể tương ứng với thiết bị nguồn 12 (FIG.1), còn thiết bị khách 108 có thể tương ứng với thiết bị đích 14 (FIG.1), theo một số ví dụ.

Nói chung, thiết bị định tuyến 104 thực hiện một hoặc nhiều giao thức định tuyến để trao đổi dữ liệu mạng qua mạng 100. Theo một số ví dụ, thiết bị định tuyến 104 có thể được tạo cấu hình để thực hiện các thao tác ủy nhiệm hoặc lưu trữ. Do vậy, theo một số ví dụ, thiết bị định tuyến 104 có thể được gọi là thiết bị ủy nhiệm. Nói chung, thiết bị định tuyến 104 chạy các giao thức định tuyến để tìm ra các tuyến qua mạng 100. Bằng cách chạy các giao thức định tuyến này, thiết bị định tuyến 104B có thể phát hiện ra đường mạng từ chính thiết bị này đến thiết bị máy chủ 102 qua thiết bị định tuyến 104A.

Kỹ thuật theo sáng chế có thể được thực hiện bởi thiết bị mạng như thiết bị định tuyến 104 và thiết bị chuyển mã 106, nhưng cũng có thể được thực hiện bởi thiết bị khách 108. Theo cách này, thiết bị định tuyến 104, thiết bị chuyển mã 106, và thiết bị

khách 108 là các ví dụ về thiết bị được tạo cấu hình để thực hiện kỹ thuật của bộ lọc này. Hơn nữa, thiết bị trên Fig.1, và bộ mã hóa 20 được minh họa trên Fig.2, và bộ giải mã 30 được minh họa trên Fig.3 là các ví dụ về thiết bị có thể được tạo cấu hình để thực hiện một hoặc nhiều kỹ thuật theo sáng chế.

Fig.5 là sơ đồ khái niệm minh họa cấu trúc làm ví dụ của tập tin 300, theo một hoặc nhiều kỹ thuật của bản mô tả này. Theo ví dụ trên Fig.5, tập tin 300 bao gồm hộp phim 302 và nhiều hộp dữ liệu phương tiện 304. Mặc dù được minh họa trong ví dụ trên Fig.5 là ở trong cùng một tập tin, theo các ví dụ khác hộp phim 302 và hộp dữ liệu phương tiện 304 có thể trong các tập tin riêng biệt. Như nêu trên, một hộp có thể là khối kiến trúc hướng đến đối tượng được xác định bởi ký hiệu nhận dạng loại duy nhất và chiều dài. Chẳng hạn, một hộp có thể là cấu trúc phần tử cú pháp theo ISOBMFF, bao gồm loại hộp được mã hóa bởi bốn đặc điểm, tổng số byte của hộp này và tải tin.

Hộp phim 302 có thể chứa siêu dữ liệu cho các rãnh của tập tin 300. Mỗi rãnh của tập tin 300 có thể bao gồm dòng dữ liệu phương tiện liên tục. Mỗi hộp dữ liệu phương tiện 304 có thể bao gồm một hoặc nhiều mẫu 305. Mỗi mẫu 305 có thể bao gồm một đơn vị truy cập âm thanh hoặc video. Như được mô tả ở phần khác trong bản mô tả này, mỗi đơn vị truy cập có thể bao gồm nhiều hình ảnh được mã hóa trong mã hóa đa khung hình (ví dụ, MV-HEVC và 3D-HEVC) và mã hóa video mở rộng được (ví dụ, SHVC). Chẳng hạn, đơn vị truy cập có thể bao gồm một hoặc nhiều hình ảnh được mã hóa cho mỗi lớp.

Hơn nữa, theo ví dụ trên Fig.5, hộp phim 302 bao gồm hộp rãnh 306. Hộp rãnh 306 có thể chứa siêu dữ liệu cho một rãnh của tập tin 300. Theo các ví dụ khác, hộp phim 302 có thể bao gồm nhiều hộp rãnh cho các rãnh khác nhau của tập tin 300. Hộp rãnh 306 bao gồm hộp phương tiện 307. Hộp phương tiện 307 có thể chứa tất cả các đối tượng thể hiện thông tin về dữ liệu phương tiện trong rãnh. Hộp phương tiện 307 bao gồm hộp thông tin phương tiện 308. Hộp phương tiện 308 có thể chứa tất cả các đối tượng thể hiện thông tin đặc trưng về phương tiện trong rãnh. Hộp thông tin phương tiện 308 bao gồm hộp bảng mẫu 309. Hộp bảng mẫu 309 có thể mô tả siêu dữ liệu dành riêng cho mẫu.

Theo ví dụ trên Fig.5, hộp bảng mẫu 309 bao gồm hộp mẫu theo nhóm (SampleToGroup) 310 và hộp mô tả nhóm mẫu (SampleGroupDescription) 312. Theo cách khác, hộp bảng mẫu 309 có thể bao gồm các hộp khác ngoài hộp mẫu theo nhóm

310 và hộp mô tả nhóm mẫu 312, và/hoặc có thể bao gồm nhiều hộp mẫu theo nhóm và hộp mô tả nhóm mẫu. Hộp mẫu theo nhóm 310 có thể sắp xếp các mẫu (ví dụ, mẫu cụ thể trong số các mẫu 305) vào nhóm mẫu. Hộp mô tả nhóm mẫu 312 có thể mô tả một đặc tính chung của các mẫu trong nhóm mẫu (tức là, nhóm mẫu). Hơn nữa, hộp bảng mẫu 309 có thể bao gồm các hộp mục nhập mẫu 311. Mỗi hộp mục nhập mẫu 311 có thể tương ứng với một mẫu trong nhóm mẫu. Theo một số ví dụ, hộp mục nhập mẫu 311 là các trường hợp của nhóm mục nhập truy cập ngẫu nhiên mà mở rộng nhóm mô tả nhóm mẫu cơ sở như được định nghĩa trong phần 9.5.5.1.2, nêu trên.

Theo một hoặc nhiều kỹ thuật của bản mô tả này, hộp mô tả nhóm mẫu 312 có thể mô tả mỗi mẫu của nhóm mẫu chứa ít nhất một hình ảnh IRAP. Theo cách này, thiết bị tạo tập tin 34 có thể tạo tập tin bao gồm hộp rãnh 306 chứa siêu dữ liệu cho một rãnh trong tập tin 300. Dữ liệu phương tiện cho rãnh này bao gồm chuỗi các mẫu 305. Mỗi mẫu này có thể là đơn vị truy cập video của dữ liệu video đa lớp (ví dụ, dữ liệu video SHVC, MV-HEVC hoặc 3D-HEVC). Hơn nữa, là một phần của quá trình tạo tập tin 300, thiết bị tạo tập tin 34 có thể tạo ra, trong tập tin 300, một hộp bổ sung (tức là, hộp bảng mẫu 309) ghi lại tất cả các mẫu 305 chứa ít nhất một hình ảnh IRAP. Nói cách khác, hộp bổ sung này nhận dạng tất cả các mẫu 305 chứa ít nhất một hình ảnh IRAP. Theo ví dụ trên Fig.5, hộp bổ sung này xác định nhóm mẫu ghi lại (ví dụ, nhận dạng) tất cả các mẫu 305 chứa ít nhất một hình ảnh IRAP. Nói cách khác, hộp bổ sung này mô tả các mẫu 305 chứa ít nhất một hình ảnh IRAP thuộc nhóm mẫu.

Hơn nữa, theo một hoặc nhiều kỹ thuật của bản mô tả này, mỗi hộp mục nhập mẫu 311 có thể bao gồm giá trị (ví dụ, `all_pics_are_IRAP`) chỉ ra tất cả các hình ảnh được mã hóa trong mẫu tương ứng có phải là hình ảnh IRAP hay không. Theo một số ví dụ, giá trị này bằng 1 chỉ ra rằng không phải tất cả các hình ảnh được mã hóa trong mẫu là hình ảnh IRAP. Giá trị này bằng 0 chỉ ra rằng không đòi hỏi mỗi hình ảnh được mã hóa trong mỗi mẫu của nhóm mẫu là hình ảnh IRAP.

Theo một số ví dụ, khi không phải tất cả các hình ảnh được mã hóa của mẫu là hình ảnh IRAP, thiết bị tạo tập tin 34 có thể bao gồm, trong một trong số các hộp mục nhập mẫu 311 cho một mẫu cụ thể, một giá trị (ví dụ, `num_IRAP_pics`) chỉ ra số hình ảnh IRAP trong mẫu cụ thể. Ngoài ra, thiết bị tạo tập tin 34 có thể đưa vào, trong mục nhập mẫu này cho mẫu cụ thể, các giá trị chỉ báo ký hiệu nhận dạng lớp của các hình ảnh IRAP trong mẫu cụ thể. Thiết bị tạo tập tin 34 cũng có thể đưa vào, trong mục

nhập mẫu này cho mẫu cụ thể, giá trị chỉ báo kiểu đơn vị NAL của các đơn vị NAL VCL trong các hình ảnh IRAP của mẫu cụ thể.

Hơn nữa, theo ví dụ trên Fig.5, hộp bảng mẫu 309 bao gồm hộp thông tin mẫu con 314. Mặc dù ví dụ trên Fig.5 chỉ thể hiện một hộp thông tin mẫu con, hộp bảng mẫu 309 có thể bao gồm nhiều hộp thông tin mẫu con. Nói chung, một hộp thông tin mẫu con được thiết kế để chứa thông tin mẫu con. Mẫu con là một khoảng byte liên tục của một mẫu. ISO/IEC 14496-12 chỉ ra rằng định nghĩa cụ thể về mẫu con được đưa ra đối với một hệ thống mã hóa xác định, như H.264/AVC hoặc HEVC.

Phần 8.4.8 của ISO/IEC 14496-15 chỉ rõ định nghĩa về mẫu con cho HEVC. Cụ thể là, phần 8.4.8 của ISO/IEC 14496-15 chỉ rõ rằng đối với việc sử dụng hộp thông tin mẫu con (8.7.7 trong ISO/IEC 14496-12) trong dòng HEVC, mẫu con được định nghĩa trên cơ sở giá trị của trường cờ của hộp thông tin mẫu con. Theo một hoặc nhiều kỹ thuật của bản mô tả này, nếu trường cờ trong hộp thông tin mẫu con 314 bằng 5, mẫu con tương ứng với hộp thông tin mẫu con 314 chứa một hình ảnh được mã hóa và các đơn vị NAL không VCL liên quan. Các đơn vị NAL không VCL liên quan có thể bao gồm các đơn vị NAL chứa thông báo SEI có thể áp dụng cho hình ảnh được mã hóa và các đơn vị NAL chứa các tập tham số (ví dụ, VPS, SPS, PPS, v.v.) có thể áp dụng cho hình ảnh được mã hóa.

Do vậy, theo một ví dụ, thiết bị tạo tập tin 34 có thể tạo tập tin (ví dụ, tập tin 300) bao gồm hộp rãnh (ví dụ, hộp rãnh 306) chứa siêu dữ liệu cho một rãnh trong tập tin. Theo một ví dụ, dữ liệu phương tiện cho một rãnh có thể bao gồm chuỗi các mẫu, mỗi mẫu là một đơn vị truy cập video của dữ liệu video đa lớp (ví dụ, dữ liệu video SHVC, MV-HEVC, hoặc 3D-HEVC). Hơn nữa, theo ví dụ này, là một phần của thiết bị tạo tập tin 34 tạo tập tin, thiết bị tạo tập tin 34 có thể tạo ra, trong tập tin này, hộp thông tin mẫu con (ví dụ, hộp thông tin mẫu con 314) chứa các cờ chỉ rõ kiểu thông tin mẫu con có trong hộp thông tin mẫu con. Khi cờ có giá trị cụ thể, mẫu con tương ứng với hộp thông tin mẫu con chứa chính xác một hình ảnh được mã hóa và không hoặc nhiều đơn vị NAL không VCL liên quan đến hình ảnh được mã hóa này.

Hơn nữa theo một hoặc nhiều kỹ thuật của bản mô tả này, nếu trường cờ của hộp thông tin mẫu con 314 bằng 0, hộp thông tin mẫu con 314 còn bao gồm giá trị DiscardableFlag, giá trị NoInterLayerPredFlag, giá trị LayerId và giá trị TempId. Nếu trường cờ của hộp thông tin mẫu con 314 bằng 5, hộp thông tin mẫu con 314 có thể

bao gồm giá trị DiscardableFlag, giá trị VclNalUnitType, giá trị LayerId, giá trị TemporalID, giá trị NoInterLayerPredFlag, giá trị SubLayerRefNalUnitFlag và giá trị dự trữ.

SubLayerRefNalUnitFlag bằng 0 chỉ ra rằng tất cả các đơn vị NAL trong mẫu con là đơn vị NAL VCL của một hình ảnh không tham chiếu của lớp con như được mô tả trong ISO/IEC 23008-2 (tức là, HEVC). SubLayerRefNalUnitFlag bằng 1 chỉ ra rằng tất cả các đơn vị NAL trong mẫu con là đơn vị NAL VCL của một hình ảnh tham chiếu của lớp con như được mô tả trong ISO/IEC 23008-2 (tức là, HEVC). Do vậy, khi thiết bị tạo tập tin 34 tạo ra hộp thông tin mẫu con 314 và các cờ có giá trị cụ thể (ví dụ, 5), thiết bị tạo tập tin 34 bao gồm, trong hộp thông tin mẫu con 314, cờ phụ chỉ ra tất cả các đơn vị NAL trong mẫu con có phải là đơn vị NAL VCL của hình ảnh không tham chiếu của lớp con hay không.

DiscardableFlag chỉ ra giá trị discardable\_flag của các đơn vị NAL VCL trong mẫu con. Như được mô tả trong phần A.4 của ISO/IEC 14496-15, discardable\_flag được thiết lập là 1 nếu và chỉ nếu tất cả các đơn vị NAL được trích hoặc gộp có discardable\_flag được thiết lập là 1, và được thiết lập là 0 theo cách khác. Đơn vị NAL có thể có giá trị discardable\_flag được thiết lập là 1 nếu dòng bit chứa đơn vị NAL có thể được giải mã chính xác mà không cần đơn vị NAL. Do vậy, đơn vị NAL "có thể loại bỏ" nếu dòng bit chứa đơn vị NAL có thể được giải mã chính xác mà không cần đơn vị NAL. Tất cả các đơn vị NAL VCL trong mẫu con có cùng giá trị discardable\_flag. Do vậy, khi thiết bị tạo tập tin 34 tạo ra hộp thông tin mẫu con 314 và các cờ có giá trị cụ thể (ví dụ, 5), thiết bị tạo tập tin 34 bao gồm, trong hộp thông tin mẫu con 314, cờ phụ (ví dụ, discardable\_flag) chỉ báo liệu tất cả các đơn vị NAL VCL của mẫu con là có thể loại bỏ hay không.

Giá trị NoInterLayerPredFlag chỉ ra giá trị của inter\_layer\_pred\_enabled\_flag của các đơn vị NAL VCL trong mẫu con. inter\_layer\_pred\_enabled\_flag được thiết lập là 1 nếu và chỉ nếu tất cả các đơn vị NAL VCL được trích hoặc gộp có inter\_layer\_pred\_enabled\_flag được thiết lập là 1, và được thiết lập là 0 theo cách khác. Tất cả các đơn vị NAL VCL trong mẫu con có cùng giá trị inter\_layer\_pred\_enabled\_flag. Do vậy, khi thiết bị tạo tập tin 34 tạo ra hộp thông tin mẫu con 314 và các cờ có giá trị cụ thể (ví dụ, 5), thiết bị tạo tập tin 34 bao gồm, trong hộp thông tin mẫu con 314, cờ phụ (ví dụ, inter\_layer\_pred\_enabled\_flag) chỉ ra việc

dự đoán liên lớp có thể được thực hiện cho tất cả các đơn vị NAL VCL của mẫu con hay không.

LayerId chỉ ra giá trị `nuh_layer_id` của các đơn vị NAL trong mẫu con. Tất cả các đơn vị NAL VCL trong mẫu con có cùng giá trị `nuh_layer_id`. Do đó, khi thiết bị tạo tập tin 34 tạo ra hộp thông tin mẫu con 314 và các cờ có giá trị cụ thể (ví dụ, 5), thiết bị tạo tập tin 34 bao gồm, trong hộp thông tin mẫu con 314, giá trị bổ sung (ví dụ, `LayerId`) chỉ ra ký hiệu nhận dạng lớp của mỗi đơn vị NAL của mẫu con.

`TempId` chỉ ra giá trị `TemporalId` của các đơn vị NAL trong mẫu con. Tất cả các đơn vị NAL VCL trong mẫu con có cùng giá trị `TemporalId`. Do vậy, khi thiết bị tạo tập tin 34 tạo ra hộp thông tin mẫu con 314 và các cờ có giá trị cụ thể (ví dụ, 5), thiết bị tạo tập tin 34 bao gồm, trong hộp thông tin mẫu con 314, cờ phụ (ví dụ, `TempId`) chỉ báo ký hiệu nhận dạng thời gian của mỗi đơn vị NAL của mẫu con.

Kiểu đơn vị NAL VCL (`VclNalUnitType`) chỉ ra phần tử cú pháp `nal_unit_type` của các đơn vị NAL VCL trong mẫu con. Phần tử cú pháp `nal_unit_type` là phần tử cú pháp trong tiêu đề đơn vị NAL của đơn vị NAL. Phần tử cú pháp `nal_unit_type` mô tả kiểu của RBSP có trong đơn vị NAL. Tất cả các đơn vị NAL VCL `nal_unit_type` trong mẫu con có cùng giá trị `nal_unit_type`. Do vậy, khi thiết bị tạo tập tin 34 tạo ra hộp thông tin mẫu con 314 và các cờ có giá trị cụ thể (ví dụ, 5), thiết bị tạo tập tin 34 bao gồm, trong hộp thông tin mẫu con 314, cờ phụ (ví dụ, `VclNalUnitType`) chỉ báo kiểu đơn vị NAL của các đơn vị NAL VCL của mẫu con. Tất cả các đơn vị NAL VCL trong mẫu con có cùng kiểu đơn vị NAL.

Fig.6 là sơ đồ khái niệm minh họa cấu trúc làm ví dụ của tập tin 300, theo một hoặc nhiều kỹ thuật của bản mô tả này. Như được mô tả trong phần 8.4.9 của ISO/IEC 14496-15, HEVC cho phép các mẫu định dạng tập tin chỉ được sử dụng để tham chiếu và không được kết xuất. Ví dụ, HEVC cho phép ảnh tham chiếu không được hiển thị trong video.

Hơn nữa, phần 8.4.9 của ISO/IEC 14496-15 mô tả rằng khi mẫu bất kỳ không kết xuất có mặt trong một rãnh, tập tin sẽ bị ràng buộc như sau.

1. Mẫu không kết xuất được xác định thời gian hợp thành nằm ngoài khoảng thời gian của mẫu được kết xuất.
2. Danh sách soạn thảo được sử dụng để loại bỏ thời gian hợp thành của các mẫu không được kết xuất.

3. Khi rãnh bao gồm hộp độ lệch hợp thành (CompositionOffsetBox) ('ctts'),
    - a. phiên bản 1 của hộp độ lệch hợp thành sẽ được sử dụng
    - b. giá trị của sample\_offset được thiết lập bằng  $-2^{31}$  cho mỗi mẫu không kết xuất,
    - c. Hộp hợp thành để giải mã (CompositionToDecodeBox) ('cslg') sẽ được chứa trong hộp bảng mẫu (SampleTableBox) ('stbl') của rãnh, và
    - d. Khi hộp hợp thành để giải mã có mặt cho rãnh này, giá trị của trường leastDecodeToDisplayDelta là độ lệch hợp thành nhỏ nhất trong hộp độ lệch hợp thành ngoại trừ các giá trị sample\_offset cho các mẫu không kết xuất.
- LƯU Ý: Do vậy, giá trị của trường leastDecodeToDisplayDelta lớn hơn  $-2^{31}$ .

Như được mô tả trong ISO/IEC 14496-12, hộp độ lệch hợp thành cung cấp độ lệch giữa thời gian giải mã và thời gian hợp thành. Hộp độ lệch hợp thành bao gồm tập các giá trị sample\_offset. Mỗi giá trị sample\_offset là một số nguyên không âm đưa ra độ lệch giữa thời gian hợp thành và thời gian giải mã. Thời gian hợp thành chỉ thời gian mà mẫu được kết xuất. Thời gian giải mã chỉ thời gian mà mẫu được giải mã.

Như nêu trên, đơn vị NAL đã mã hóa có thể bao gồm tiêu đề đoạn lát. Tiêu đề đoạn lát là một phần của lát đã mã hóa và có thể chứa các phần tử dữ liệu liên quan đến CTU thứ nhất hoặc tất cả các CTU trong đoạn lát. Theo HEVC, tiêu đề đoạn lát bao gồm phần tử cú pháp pic\_output\_flag. Nói chung phần tử cú pháp pic\_output\_flag có trong tiêu đề đoạn lát thứ nhất của một lát của một hình ảnh. Do vậy, bản mô tả này có thể đề cập đến pic\_output\_flag của tiêu đề đoạn lát thứ nhất của lát của hình ảnh là pic\_output\_flag của hình ảnh.

Như được mô tả trong phần 7.4.7.1 của HEVC WD, phần tử cú pháp pic\_output\_flag tác động đến quá trình kết xuất và loại bỏ ảnh giải mã như được mô tả trong Phụ lục C của HEVC WD. Nói chung, nếu phần tử cú pháp pic\_output\_flag của tiêu đề đoạn lát đối với một đoạn lát là 1, một hình ảnh bao gồm một lát tương ứng với tiêu đề đoạn lát này được kết xuất. Theo cách khác, nếu phần tử cú pháp pic\_output\_flag của tiêu đề đoạn lát đối với một đoạn lát là 0, hình ảnh bao gồm lát tương ứng với tiêu đề đoạn lát này có thể được giải mã để sử dụng làm hình ảnh tham chiếu, nhưng không được kết xuất.

Theo một hoặc nhiều kỹ thuật của bản mô tả này, các tham chiếu trong phần 8.4.9 của ISO/IEC 14496-15 đối với HEVC có thể được thay bằng các tham chiếu tương ứng đối với SHVC, MV-HEVC, hoặc 3D-HEVC. Hơn nữa, theo một hoặc nhiều kỹ thuật của bản mô tả này, khi đơn vị truy cập chứa một số hình ảnh được mã hóa có `pic_output_flag` bằng 1 và một số hình ảnh được mã hóa khác có `pic_output_flag` bằng 0, ít nhất hai rãnh phải được sử dụng để lưu trữ dòng này. Đối với mỗi rãnh tương ứng trong số các rãnh này, tất cả các hình ảnh được mã hóa trong mỗi mẫu của rãnh tương ứng có cùng giá trị `pic_output_flag`. Do vậy, tất cả các hình ảnh được mã hóa trong rãnh thứ nhất trong số các rãnh có `pic_output_flag` bằng 0 và tất cả các hình ảnh được mã hóa trong rãnh thứ hai trong số các rãnh có `pic_output_flag` bằng 1.

Do vậy, theo ví dụ trên Fig.6, thiết bị tạo tập tin 34 có thể tạo tập tin 400. Tương tự với tập tin 300, theo ví dụ trên Fig.5, tập tin 400 bao gồm hộp phim 402 và một hoặc nhiều hộp dữ liệu phương tiện 404. Mỗi hộp dữ liệu phương tiện 404 có thể tương ứng với rãnh khác của tập tin 400. Hộp phim 402 có thể chứa siêu dữ liệu cho các rãnh của tập tin 400. Mỗi rãnh của tập tin 400 có thể bao gồm dòng dữ liệu phương tiện liên tục. Mỗi hộp dữ liệu phương tiện 404 có thể bao gồm một hoặc nhiều mẫu 405. Mỗi mẫu 405 có thể bao gồm một đơn vị truy cập âm thanh hoặc video.

Như nêu trên, theo một số ví dụ, khi đơn vị truy cập chứa một số hình ảnh được mã hóa có `pic_output_flag` bằng 1 và một số hình ảnh được mã hóa khác có `pic_output_flag` bằng 0, ít nhất hai rãnh phải được sử dụng để lưu trữ dòng này. Do vậy, theo ví dụ trên Fig.6, hộp phim 402 bao gồm hộp rãnh 406 và hộp rãnh 408. Mỗi hộp rãnh 406 và 408 chứa siêu dữ liệu cho một rãnh khác của tập tin 400. Chẳng hạn, hộp rãnh 406 có thể chứa siêu dữ liệu cho một rãnh có các hình ảnh được mã hóa với `pic_output_flag` bằng 0, và không có hình ảnh có `pic_output_flag` bằng 1. Hộp rãnh 408 có thể chứa siêu dữ liệu cho một rãnh có các hình ảnh được mã hóa với `pic_output_flag` bằng 1, và không có hình ảnh có `pic_output_flag` bằng 0.

Do vậy, theo một ví dụ, thiết bị tạo tập tin 34 có thể tạo tập tin (ví dụ, tập tin 400) bao gồm hộp dữ liệu phương tiện (ví dụ, hộp dữ liệu phương tiện 404) chứa nội dung phương tiện. Nội dung phương tiện bao gồm chuỗi các mẫu (ví dụ, mẫu 405). Mỗi mẫu có thể là đơn vị truy cập video của dữ liệu video đa lớp. Theo ví dụ này, khi thiết bị tạo tập tin 34 tạo tập tin, đáp lại việc xác định rằng ít nhất một đơn vị truy cập của dòng bit bao gồm hình ảnh được mã hóa có cờ kết xuất hình ảnh bằng 1 và hình

ảnh được mã hóa có cờ kết xuất hình ảnh bằng 0, thiết bị tạo tập tin 34 có thể sử dụng ít nhất hai rãnh để lưu trữ dòng bit trong tập tin. Đối với mỗi rãnh tương ứng từ ít nhất hai rãnh này, tất cả các hình ảnh được mã hóa trong mỗi mẫu của rãnh tương ứng có cùng giá trị cờ kết xuất hình ảnh. Các hình ảnh có cờ kết xuất hình ảnh bằng 1 được phép kết xuất và các hình ảnh có cờ kết xuất hình ảnh bằng 0 được phép được sử dụng làm hình ảnh tham chiếu nhưng không được phép kết xuất.

Fig.7 là lưu đồ minh họa thao tác làm ví dụ của thiết bị tạo tập tin 34, theo một hoặc nhiều kỹ thuật của bản mô tả này. Thao tác trên Fig.7, cùng với các thao tác được minh họa trong các lưu đồ khác của bản mô tả này, là các ví dụ. Các thao tác làm ví dụ khác cùng với kỹ thuật của bản mô tả này có thể bao gồm nhiều hơn, ít hơn hoặc các thao tác khác.

Theo ví dụ trên Fig.7, thiết bị tạo tập tin 34 tạo tập tin 500. Là một phần của việc tạo tập tin, thiết bị tạo tập tin 34 tạo hộp rãnh chứa siêu dữ liệu cho một rãnh trong tập tin (502). Theo cách này, thiết bị tạo tập tin 34 có thể tạo tập tin bao gồm hộp rãnh 306 chứa siêu dữ liệu cho một rãnh trong tập tin. Dữ liệu phương tiện cho rãnh này bao gồm chuỗi các mẫu. Mỗi mẫu là đơn vị truy cập video của dữ liệu video đa lớp. Theo một số ví dụ, thiết bị tạo tập tin 34 mã hóa dữ liệu video đa lớp.

Hơn nữa, là một phần của việc tạo tập tin, thiết bị tạo tập tin 34 nhận dạng tất cả các mẫu chứa ít nhất một hình ảnh IRAP (504). Ngoài ra, thiết bị tạo tập tin 34 có thể tạo ra, trong tập tin này, một hộp bổ sung ghi lại tất cả các mẫu chứa ít nhất một hình ảnh IRAP (506). Theo một số ví dụ, hộp bổ sung này là hộp mới không được định nghĩa trong ISOBMFF hoặc các phiên bản mở rộng hiện có của chuẩn này. Theo một số ví dụ, hộp bổ sung này xác định nhóm mẫu ghi tất cả các mẫu chứa ít nhất một hình ảnh IRAP. Chẳng hạn, hộp bổ sung có thể là hoặc bao gồm hộp bảng mẫu chứa hộp mẫu theo nhóm và hộp mô tả nhóm mẫu. Hộp mẫu theo nhóm nhận dạng các mẫu chứa ít nhất một hình ảnh IRAP. Hộp mô tả nhóm mẫu chỉ ra rằng nhóm mẫu là nhóm các mẫu chứa ít nhất một hình ảnh IRAP.

Hơn nữa, theo ví dụ trên Fig.7, thiết bị tạo tập tin 34 có thể tạo ra mục nhập mẫu cho một mẫu cụ thể trong số các mẫu chứa ít nhất một hình ảnh IRAP (508). Theo một số ví dụ, thiết bị tạo tập tin 34 có thể tạo ra mục nhập mẫu cho một mẫu tương ứng trong số các mẫu chứa ít nhất một hình ảnh IRAP. Mục nhập mẫu này có thể là mục nhập mẫu truy cập ngẫu nhiên như được xác định trong phần 9.5.5.1.2, nêu trên.

Như được minh họa trong ví dụ trên Fig.7, là một phần của quá trình tạo mục nhập mẫu cho mẫu cụ thể, thiết bị tạo tập tin 34 có thể đưa vào, trong mục nhập mẫu này cho mẫu cụ thể, giá trị chỉ ra tất cả các hình ảnh được mã hóa trong mẫu cụ thể có phải là hình ảnh IRAP hay không (510). Theo cách này thiết bị tạo tập tin 34 có thể tạo, trong tập tin, mục nhập mẫu bao gồm giá trị chỉ báo tất cả các hình ảnh được mã hóa trong mẫu cụ thể trong chuỗi mẫu có phải là hình ảnh IRAP hay không. Hơn nữa, thiết bị tạo tập tin 34 có thể đưa vào, trong mục nhập mẫu này cho mẫu cụ thể, giá trị chỉ báo kiểu đơn vị NAL của các đơn vị NAL VCL trong các hình ảnh IRAP của mẫu cụ thể (512).

Ngoài ra, thiết bị tạo tập tin 34 có thể xác định tất cả các hình ảnh được mã hóa trong một mẫu cụ thể có phải là hình ảnh IRAP hay không (514). Khi không phải tất cả các hình ảnh được mã hóa của mẫu là hình ảnh IRAP ("Không" của 514), thiết bị tạo tập tin 34 có thể đưa vào, trong mục nhập mẫu này cho một mẫu cụ thể, một giá trị chỉ ra số hình ảnh IRAP trong mẫu cụ thể (516). Ngoài ra, thiết bị tạo tập tin 34 có thể đưa vào, trong mục nhập mẫu này cho mẫu cụ thể, các giá trị chỉ báo ký hiệu nhận dạng lớp (tức là, `nuh_layer_id`) của các hình ảnh IRAP trong mẫu cụ thể.

Như được chỉ ra trên đây, Fig.7 đưa ra một ví dụ. Các ví dụ khác không bao gồm từng thao tác trên Fig.7. Chẳng hạn, một số ví dụ không bao gồm các bước 502, 504 và 508. Hơn nữa, các ví dụ khác không bao gồm các bước khác nhau trong số các bước 510-518. Hơn nữa, một số ví dụ bao gồm một hoặc nhiều thao tác phụ khác. Chẳng hạn, một số ví dụ bao gồm thao tác phụ đó là tạo ra, là một phần của quá trình tạo tập tin, hộp mẫu đồng bộ hóa bao gồm bảng mẫu đồng bộ hóa ghi các mẫu đồng bộ hóa của một rãnh của dữ liệu video đa lớp. Mỗi mẫu đồng bộ hóa của rãnh này là mẫu truy cập ngẫu nhiên của rãnh. Theo ví dụ này, mẫu mã hóa video mở rộng được là mẫu đồng bộ hóa nếu mỗi hình ảnh được mã hóa trong đơn vị truy cập là hình ảnh IRAP. Hơn nữa, theo ví dụ này, mẫu mã hóa video đa khung hình là mẫu đồng bộ hóa nếu mỗi hình ảnh được mã hóa trong đơn vị truy cập là hình ảnh IRAP mà không có hình ảnh RASL.

Fig.8 là lưu đồ minh họa thao tác làm ví dụ trong đó thiết bị điện toán thực hiện việc truy cập ngẫu nhiên và/hoặc chuyển mạch mức, theo một hoặc nhiều kỹ thuật của bản mô tả này. Theo ví dụ trên Fig.8, thiết bị điện toán nhận tập tin (550). Theo ví dụ trên Fig.8, thiết bị điện toán có thể là thiết bị mạng trung gian (ví dụ, MANE, máy chủ

truyền liên tục), thiết bị giải mã (ví dụ, thiết bị đích 14), hoặc loại thiết bị video khác. Theo một số ví dụ, thiết bị điện toán có thể là một phần của mạng cung cấp nội dung.

Theo ví dụ trên Fig.8, thiết bị điện toán có thể thu nhận, từ tập tin, hộp rãnh chứa siêu dữ liệu cho một rãnh trong tập (552). Dữ liệu phương tiện cho rãnh này bao gồm chuỗi các mẫu. Theo ví dụ trên Fig.8, mỗi mẫu là đơn vị truy cập video của dữ liệu video đa lớp.

Hơn nữa, theo ví dụ trên Fig.8, thiết bị điện toán có thể thu nhận hộp bổ sung từ tập tin (554). Hộp bổ sung này ghi tất cả các mẫu chứa ít nhất một hình ảnh IRAP. Do vậy, thiết bị điện toán có thể xác định, dựa trên thông tin trong hộp bổ sung, tất cả các mẫu chứa ít nhất một hình ảnh IRAP (556).

Hơn nữa, theo một số ví dụ, thiết bị điện toán có thể thu nhận, từ tập tin, mục nhập mẫu bao gồm giá trị chỉ báo tất cả các hình ảnh được mã hóa trong mẫu cụ thể trong chuỗi mẫu có phải là hình ảnh IRAP hay không. Khi không phải tất cả các hình ảnh được mã hóa của mẫu là hình ảnh IRAP, thiết bị điện toán có thể thu nhận, từ mục nhập mẫu, một giá trị chỉ ra số hình ảnh IRAP trong mẫu cụ thể. Ngoài ra, thiết bị điện toán có thể thu nhận, từ mục nhập mẫu, các giá trị chỉ báo ký hiệu nhận dạng lớp của các hình ảnh IRAP trong mẫu cụ thể. Ngoài ra, theo một số ví dụ, thiết bị điện toán có thể ra,, từ mục nhập mẫu, giá trị chỉ báo kiểu đơn vị NAL của các đơn vị NAL VCL trong các hình ảnh IRAP của mẫu cụ thể. Ngoài ra, theo một số ví dụ, thiết bị điện toán có thể thu nhận, từ tập tin, hộp mẫu đồng bộ hóa bao gồm bảng mẫu đồng bộ hóa ghi các mẫu đồng bộ hóa của một rãnh của dữ liệu video. Theo các ví dụ này, mỗi mẫu đồng bộ hóa của rãnh là mẫu truy cập ngẫu nhiên của rãnh, mẫu mã hóa video mở rộng được là mẫu đồng bộ hóa nếu mỗi hình ảnh được mã hóa trong đơn vị truy cập là hình ảnh IRAP, và mẫu mã hóa video đa khung hình là mẫu đồng bộ hóa nếu mỗi hình ảnh được mã hóa trong đơn vị truy cập là hình ảnh IRAP không có hình ảnh RASL.

Hơn nữa, theo ví dụ trên Fig.8, thiết bị điện toán bắt đầu chuyển tiếp và giải mã các đơn vị NAL của mẫu chứa ít nhất một hình ảnh IRAP mà không chuyển tiếp và giải mã các đơn vị NAL của tập tin ở trước theo thứ tự giải mã đến mẫu (558). Theo cách này, thiết bị điện toán có thể thực hiện việc truy cập ngẫu nhiên hoặc chuyển lớp. Chẳng hạn, thiết bị điện toán có thể bắt đầu giải mã dữ liệu video đa lớp ở một trong số một hoặc nhiều mẫu chứa ít nhất một hình ảnh IRAP.

Fig.9 là lưu đồ minh họa thao tác làm ví dụ của thiết bị tạo tập tin 34, theo một hoặc nhiều kỹ thuật của bản mô tả này. Theo ví dụ trên Fig.9, thiết bị tạo tập tin 34 có thể tạo tập tin bao gồm hộp rãnh chứa siêu dữ liệu cho một rãnh trong tập tin 600. Dữ liệu phương tiện cho rãnh này bao gồm chuỗi các mẫu. Theo ví dụ trên Fig.9, mỗi mẫu là đơn vị truy cập video của dữ liệu video đa lớp. Theo một số ví dụ, thiết bị tạo tập tin 34 mã hóa dữ liệu video đa lớp.

Là một phần của việc tạo tập tin, thiết bị tạo tập tin 34 có thể xác định mẫu con có chứa chính xác một hình ảnh được mã hóa và không hoặc nhiều đơn vị NAL không VCL liên quan đến hình ảnh được mã hóa (602) hay không. Đáp lại việc xác định mẫu con chứa chính xác một hình ảnh được mã hóa và không hoặc nhiều đơn vị NAL không VCL liên quan đến hình ảnh được mã hóa (“CÓ” của 602), thiết bị tạo tập tin 34 có thể tạo ra, trong tập tin, hộp thông tin mẫu con chứa các cờ có giá trị (ví dụ, 5) chỉ ra rằng mẫu con chứa chính xác một hình ảnh được mã hóa và không hoặc nhiều đơn vị NAL không VCL liên quan đến hình ảnh được mã hóa này (604). Theo cách khác (“KHÔNG” của 602), thiết bị tạo tập tin 34 có thể tạo ra, trong tập tin, hộp thông tin mẫu con chứa các cờ có giá trị khác (e.g., 0, 1, 2, 3, 4) (606).

Theo cách này, thiết bị tạo tập tin 34 có thể tạo tập tin bao gồm hộp rãnh 306 chứa siêu dữ liệu cho một rãnh trong tập tin. Dữ liệu phương tiện cho rãnh bao gồm chuỗi các mẫu, mỗi mẫu là một đơn vị truy cập video của dữ liệu video đa lớp. Là một phần của quá trình tạo tập tin, thiết bị tạo tập tin 34 tạo ra, trong tập tin này, hộp thông tin mẫu con chứa các cờ chỉ rõ kiểu thông tin mẫu con có trong hộp thông tin mẫu con. Khi cờ có giá trị cụ thể, mẫu con tương ứng với hộp thông tin mẫu con chứa chính xác một hình ảnh được mã hóa và không hoặc nhiều đơn vị NAL không VCL liên quan đến hình ảnh được mã hóa này.

Fig.10 là lưu đồ minh họa thao tác làm ví dụ của thiết bị điện toán, theo một hoặc nhiều kỹ thuật của bản mô tả này. Theo ví dụ trên Fig.10, thiết bị điện toán nhận tập tin (650). Theo ví dụ trên Fig.10, thiết bị điện toán có thể là thiết bị mạng trung gian, như MANE hoặc máy chủ truyền liên tục. Theo một số ví dụ, thiết bị điện toán có thể là một phần của mạng cung cấp nội dung. Hơn nữa, theo ví dụ trên Fig.10, thiết bị điện toán có thể chứa hộp rãnh của tập tin (651). Hộp rãnh chứa siêu dữ liệu cho một rãnh trong tập tin. Dữ liệu phương tiện cho rãnh này bao gồm chuỗi các mẫu. Theo ví dụ trên Fig.10, mỗi mẫu là đơn vị truy cập video của dữ liệu video đa lớp.

Hơn nữa, theo ví dụ trên Fig.10, thiết bị điện toán có thể tạo ra hộp thông tin mẫu con từ tập tin (652). Thiết bị điện toán sử dụng thông tin trong hộp thông tin mẫu con để trích dòng bit phụ (654). Dòng bit phụ có thể bao gồm mỗi đơn vị NAL của điểm thao tác của dòng bit được lưu trữ trong tập tin. Nói cách khác, các đơn vị NAL của dòng bit phụ có thể là tập con các đơn vị NAL được lưu trữ trong tập tin. Thiết bị điện toán có thể thu được hộp thông tin mẫu con từ tập tin và có thể trích dòng bit phụ mà không cần phân tích hoặc diễn dịch các đơn vị NAL trong chuỗi mẫu. Việc không phải phân tích hoặc diễn dịch các đơn vị NAL khi trích dòng bit phụ có thể làm giảm độ phức tạp của thiết bị điện toán và/hoặc thúc đẩy quá trình trích dòng bit phụ.

Hơn nữa, theo một số ví dụ, thiết bị điện toán có thể thu được, khi các cờ có giá trị cụ thể, từ hộp thông tin mẫu con, một hoặc nhiều:

- cờ bổ sung chỉ báo liệu tất cả các đơn vị NAL VCL của mẫu con là có thể loại bỏ hay không,
- giá trị bổ sung chỉ báo kiểu đơn vị NAL của các đơn vị NAL VCL của mẫu con, trong đó tất cả các đơn vị NAL VCL của mẫu con có cùng kiểu đơn vị NAL,
- giá trị bổ sung chỉ báo ký hiệu nhận dạng lớp của mỗi đơn vị của mỗi đơn vị NAL của mẫu con,
- giá trị bổ sung chỉ báo ký hiệu nhận dạng thời gian của mỗi đơn vị của mỗi đơn vị NAL của mẫu con,
- cờ bổ sung chỉ báo liệu dự đoán liên lớp có thể thực hiện được cho tất cả các đơn vị NAL VCL của mẫu con hay không, hoặc
- cờ bổ sung chỉ báo liệu tất cả các đơn vị NAL trong mẫu con có phải là đơn vị NAL VCL của hình ảnh không tham chiếu của lớp con hay không.

Theo ví dụ trên Fig.10, là một phần của việc trích dòng bit phụ, thiết bị điện toán có thể xác định một giá trị “các cờ” của hộp thông tin mẫu con có giá trị cụ thể (ví dụ, 5) chỉ ra hộp thông tin mẫu con có tương ứng với chính xác một hình ảnh được mã hóa và không hoặc nhiều đơn vị NAL không VCL liên quan đến hình ảnh được mã hóa này hay không (656). Khi giá trị “các cờ” của hộp thông tin mẫu con có giá trị cụ thể (“CÓ” của 656), thiết bị điện toán có thể xác định, dựa trên thông tin được mô tả trong hộp thông tin mẫu con, hình ảnh được mã hóa có cần để giải mã điểm hoạt động hay không (658). Ví dụ, thiết bị điện toán có thể xác định, dựa trên cờ có thể loại bỏ, bộ chỉ

báo kiểu đơn vị, ký hiệu nhận dạng lớp, ký hiệu nhận dạng thời gian, cờ không dự đoán liên lớp và/hoặc cờ đơn vị NAL tham chiếu lớp con, hình ảnh được mã hóa có cần để giải mã điểm hoạt động hay không. Khi cần hình ảnh được mã hóa để giải mã điểm hoạt động (“CÓ” của 658), thiết bị điện toán có thể bao gồm các đơn vị NAL của mẫu con trong dòng bit phụ (660). Theo cách khác, theo ví dụ trên Fig.10, khi không cần hình ảnh được mã hóa để giải mã điểm hoạt động (“KHÔNG” của 658), thiết bị điện toán không bao gồm các đơn vị NAL của mẫu con trong dòng bit phụ (662).

Hơn nữa, theo ví dụ trên Fig.10, thiết bị điện toán có thể kết xuất dòng bit phụ (664). Chẳng hạn, thiết bị điện toán có thể lưu trữ dòng bit phụ vào vật ghi đọc được bởi máy tính hoặc truyền dòng bit phụ đến thiết bị điện toán khác.

Như được chỉ ra trên đây, Fig.10 là một ví dụ. Các ví dụ khác có thể bao gồm hoặc bỏ qua các thao tác cụ thể trên Fig.10. Chẳng hạn, một số ví dụ bỏ qua các thao tác 650, 651, 654, và/hoặc 664. Hơn nữa, một số ví dụ bỏ qua một hoặc nhiều thao tác 656-662.

Fig.11 là lưu đồ minh họa thao tác làm ví dụ của thiết bị tạo tập tin 34, theo một hoặc nhiều kỹ thuật của bản mô tả này. Theo ví dụ trên Fig.11, thiết bị tạo tập tin 34 tạo tập tin bao gồm hộp dữ liệu phương tiện chứa nội dung phương tiện (700). Nội dung phương tiện có thể bao gồm chuỗi các mẫu, mỗi mẫu là một đơn vị truy cập video của dữ liệu video đa lớp. Theo các ví dụ khác nhau, dữ liệu video đa lớp có thể là dữ liệu SHVC, dữ liệu MV-HEVC, hoặc dữ liệu 3D-HEVC. Theo một số ví dụ, thiết bị tạo tập tin 34 mã hóa dữ liệu video đa lớp.

Theo ví dụ trên Fig.11, là một phần của việc tạo tập tin, thiết bị tạo tập tin 34 có thể xác định ít nhất một đơn vị truy cập của dòng bit dữ liệu video đa lớp bao gồm hình ảnh được mã hóa có cờ kết xuất hình ảnh bằng giá trị thứ nhất (ví dụ, 1) và hình ảnh được mã hóa có cờ kết xuất hình ảnh bằng giá trị thứ hai (ví dụ, 0) (702). Các hình ảnh có cờ kết xuất hình ảnh bằng giá trị thứ nhất (ví dụ, 1) được phép kết xuất và các hình ảnh có cờ kết xuất hình ảnh bằng giá trị thứ hai (ví dụ, 0) được phép được sử dụng làm hình ảnh tham chiếu nhưng không được phép kết xuất. Theo các ví dụ khác, thiết bị khác có thể thực hiện việc xác định xem ít nhất một đơn vị truy cập của dòng bit dữ liệu video đa lớp có bao gồm hình ảnh được mã hóa có cờ kết xuất hình ảnh bằng giá trị thứ nhất và hình ảnh được mã hóa có cờ kết xuất hình ảnh bằng giá trị thứ hai hay không.

Đáp lại việc xác định rằng ít nhất một đơn vị truy cập của dòng bit của dữ liệu video đa lớp bao gồm hình ảnh được mã hóa có cờ kết xuất hình ảnh bằng giá trị thứ nhất và hình ảnh được mã hóa có cờ kết xuất hình ảnh bằng giá trị thứ hai ("CÓ" của 702), thiết bị tạo tập tin 34 sử dụng ít nhất rãnh thứ nhất và rãnh thứ hai để lưu trữ dòng bit trong tập tin (704). Đối với mỗi rãnh tương ứng từ rãnh thứ nhất và thứ hai, tất cả các hình ảnh được mã hóa trong mỗi mẫu của rãnh tương ứng có cùng giá trị cờ kết xuất hình ảnh.

Hơn nữa, theo ví dụ trên Fig.11, đáp lại việc xác định rằng không có đơn vị truy cập của dòng bit của dữ liệu video đa lớp bao gồm hình ảnh được mã hóa có cờ kết xuất hình ảnh bằng giá trị thứ nhất (ví dụ, 1) và hình ảnh được mã hóa có cờ kết xuất hình ảnh bằng giá trị thứ hai (ví dụ, 0) ("CÓ" của 702), thiết bị tạo tập tin 34 sử dụng một rãnh đơn để lưu trữ dòng bit trong tập tin (706). Theo các ví dụ khác, thiết bị tạo tập tin 34 có thể tạo tập tin có nhiều rãnh ngay cả khi không có đơn vị truy cập của dòng bit bao gồm hình ảnh được mã hóa có cờ kết xuất hình ảnh bằng giá trị thứ nhất (ví dụ, 1) và hình ảnh được mã hóa có cờ kết xuất hình ảnh bằng giá trị thứ hai (ví dụ, 0).

Như được chỉ ra trên đây, Fig.11 là một ví dụ. Các ví dụ khác có thể bao gồm ít thao tác hơn. Chẳng hạn, một số ví dụ bỏ qua các thao tác 702 và 706.

Fig.12 là lưu đồ minh họa thao tác làm ví dụ của thiết bị đích 14, theo một hoặc nhiều kỹ thuật của bản mô tả này. Theo ví dụ trên Fig.12, thiết bị đích 14 nhận tập tin (750). Tập tin này có thể bao gồm hộp dữ liệu phương tiện chứa nội dung phương tiện, nội dung phương tiện bao gồm chuỗi các mẫu. Mỗi mẫu có thể là đơn vị truy cập video của dữ liệu video đa lớp. Theo các ví dụ khác nhau, dữ liệu video đa lớp có thể là dữ liệu SHVC, dữ liệu MV-HEVC, hoặc dữ liệu 3D-HEVC. Hơn nữa, theo ví dụ trên Fig.12, thiết bị đích 14 có thể thu nhận, từ tập tin, hộp rãnh thứ nhất và hộp rãnh thứ hai (751). Hộp rãnh thứ nhất chứa siêu dữ liệu cho rãnh thứ nhất trong tập tin. Hộp rãnh thứ hai chứa siêu dữ liệu cho rãnh thứ hai trong tập tin. Đối với mỗi rãnh tương ứng từ rãnh thứ nhất và thứ hai, tất cả các hình ảnh được mã hóa trong mỗi mẫu của rãnh tương ứng có cùng giá trị cờ kết xuất hình ảnh. Các hình ảnh có cờ kết xuất hình ảnh bằng giá trị thứ nhất (ví dụ, 1) được phép kết xuất và các hình ảnh có cờ kết xuất hình ảnh bằng giá trị thứ hai (ví dụ, 0) được phép được sử dụng làm hình ảnh tham chiếu nhưng không được phép kết xuất.

Bộ giải mã video 30 của thiết bị đích 14 có thể giải mã các hình ảnh trong rãnh đối với các hình ảnh có cờ kết xuất hình ảnh bằng giá trị thứ nhất (ví dụ, 1) và có thể giải mã các hình ảnh trong rãnh đối với các hình ảnh có cờ kết xuất hình ảnh bằng giá trị thứ hai (ví dụ, 0) (752). Trong một số trường hợp, bộ giải mã video 30 có thể sử dụng các hình ảnh có cờ kết xuất hình ảnh bằng 1 để giải mã các hình ảnh có cờ kết xuất hình ảnh bằng 0, và ngược lại. Thiết bị đích 14 có thể kết xuất các hình ảnh có cờ kết xuất hình ảnh bằng giá trị thứ nhất (754). Thiết bị đích 14 không kết xuất các hình ảnh có cờ kết xuất hình ảnh bằng giá trị thứ hai (756). Theo cách này, đối với mỗi rãnh tương ứng của rãnh thứ nhất và thứ hai, thiết bị đích 14 có thể giải mã các hình ảnh được mã hóa trong mỗi mẫu của rãnh tương ứng và kết xuất hình ảnh giải mã có cờ kết xuất hình ảnh bằng giá trị thứ nhất.

Như được chỉ ra trên đây, Fig.12 đưa ra một ví dụ. Các ví dụ khác có thể bỏ qua các thao tác cụ thể trên Fig.12, như các thao tác 752-756.

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

Ví dụ, và không giới hạn, vật ghi đọc được bởi máy tính như vậy có thể bao gồm RAM, ROM, EEPROM, CD-ROM hoặc các thiết bị lưu trữ đĩa quang khác, thiết bị lưu trữ đĩa từ, hoặc các thiết bị lưu trữ từ tính khác, bộ nhớ nhanh, hoặc phương tiện

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

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

Các kỹ thuật của sáng chế có thể được thực hiện trong nhiều loại cơ cấu hoặc thiết bị khác nhau, bao gồm máy cầm tay không dây, mạch tích hợp (integrated circuit - IC) hoặc bộ IC (ví dụ, bộ chip). Các thành phần, môđun hoặc đơn vị khác nhau được mô tả trong bản mô tả này để nhấn mạnh các khía cạnh chức năng của thiết bị được tạo

cấu hình để thực hiện các kỹ thuật được bộc lộ, nhưng không nhất thiết đòi hỏi việc thực hiện bởi các đơn vị phần cứng khác nhau. Thay vào đó, như được mô tả trên đây, các đơn vị khác nhau có thể được kết hợp trong đơn vị phần cứng bộ mã hóa-giải mã hoặc được tạo ra bởi tập hợp các đơn vị phần cứng liên kết hoạt động, bao gồm một hoặc nhiều bộ xử lý như được mô tả trên đây, kết hợp với phần mềm và/hoặc phần sụn thích hợp.

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

## Yêu cầu bảo hộ

1. Phương pháp xử lý dữ liệu video đa lớp, phương pháp này bao gồm các bước:

tạo tập tin bao gồm hộp rãnh chứa siêu dữ liệu cho một rãnh trong tập tin, trong đó:

dữ liệu phương tiện cho rãnh này bao gồm chuỗi các mẫu, mỗi mẫu là một đơn vị truy cập video trong số nhiều đơn vị truy cập video của dữ liệu video đa lớp và tạo tập tin bao gồm:

xác định, trong tập tin, nhóm mẫu ghi lại tất cả các mẫu chứa ít nhất một hình ảnh điểm truy cập ngẫu nhiên trong ảnh (Intra Random Access Point - IRAP), trong đó:

nhóm mẫu này được xác định là bao gồm mục nhập nhóm mẫu mà bao gồm:

giá trị chỉ báo liệu tất cả các hình ảnh được mã hóa trong mẫu cụ thể trong chuỗi mẫu có phải là hình ảnh IRAP hay không,

giá trị chỉ báo số lượng hình ảnh IRAP trong mẫu cụ thể; và

các giá trị chỉ báo các ký hiệu nhận dạng lớp của các hình ảnh IRAP trong mẫu cụ thể, và

trong đó mẫu cụ thể này chứa hình ảnh IRAP trong lớp cơ sở và còn chứa một hoặc nhiều hình ảnh không IRAP trong các lớp khác.

2. Phương pháp theo điểm 1, trong đó bước tạo tập tin bao gồm:

đưa vào, trong mục nhập nhóm mẫu, giá trị chỉ báo kiểu đơn vị lớp trừu tượng hóa mạng (Network Abstraction Layer - NAL) của các đơn vị NAL của lớp mã hóa video (Video Coding Layer - VCL) trong các hình ảnh IRAP của mẫu cụ thể.

3. Phương pháp theo điểm 1, trong đó:

bước tạo tập tin bao gồm tạo, trong tập tin, hộp mẫu đồng bộ hóa bao gồm bảng mẫu đồng bộ hóa ghi lại các mẫu đồng bộ hóa của rãnh,

mỗi mẫu đồng bộ hóa của rãnh là mẫu truy cập ngẫu nhiên của rãnh,

mẫu mã hóa video mở rộng được là mẫu đồng bộ hóa nếu mỗi hình ảnh được mã hóa trong đơn vị truy cập là hình ảnh IRAP, và

mẫu mã hóa video đa khung hình là mẫu đồng bộ hóa nếu mỗi hình ảnh được mã hóa trong đơn vị truy cập là hình ảnh IRAP mà không có hình ảnh hướng dẫn bỏ qua truy cập ngẫu nhiên (Random Access Skipped Leading - RASL).

4. Phương pháp theo điểm 1, trong đó phương pháp này còn bao gồm bước mã hóa dữ liệu video đa lớp.

5. Phương pháp xử lý dữ liệu video đa lớp, phương pháp này bao gồm các bước:

thu nhận, từ tập tin, hộp rãnh chứa siêu dữ liệu cho rãnh trong tập tin, trong đó dữ liệu phương tiện cho rãnh bao gồm chuỗi các mẫu, mỗi mẫu là một đơn vị truy cập video trong số nhiều đơn vị truy cập video của dữ liệu video đa lớp;

thu nhận, từ tập tin, định nghĩa về nhóm mẫu ghi lại tất cả các mẫu chứa ít nhất một hình ảnh điểm truy cập ngẫu nhiên trong ảnh (Intra Random Access Point - IRAP), trong đó:

nhóm mẫu này được xác định là bao gồm mục nhập nhóm mẫu mà bao gồm:

giá trị chỉ báo liệu tất cả các hình ảnh được mã hóa trong mẫu cụ thể trong chuỗi mẫu có phải là hình ảnh IRAP hay không,

giá trị chỉ báo số lượng hình ảnh IRAP trong mẫu cụ thể; và

các giá trị chỉ báo các ký hiệu nhận dạng lớp của các hình ảnh IRAP trong mẫu cụ thể, và

mẫu cụ thể này chứa hình ảnh IRAP trong lớp cơ sở và còn chứa một hoặc nhiều hình ảnh không IRAP trong các lớp khác; và

xác định, dựa trên nhóm mẫu, một hoặc nhiều trong số các mẫu chứa ít nhất một hình ảnh IRAP.

6. Phương pháp theo điểm 5, trong đó phương pháp này còn bao gồm bước:

thu nhận, từ mục nhập nhóm mẫu, giá trị chỉ báo kiểu đơn vị lớp trừu tượng hóa mạng (Network Abstraction Layer-NAL) của các đơn vị NAL của lớp mã hóa video (Video Coding Layer-VCL) trong các hình ảnh IRAP của mẫu cụ thể.

7. Phương pháp theo điểm 5, trong đó phương pháp này còn bao gồm bước:

thu nhận, từ tập tin, hộp mẫu đồng bộ hóa bao gồm bảng mẫu đồng bộ hóa ghi lại các mẫu đồng bộ hóa của rãnh, trong đó:

mỗi mẫu đồng bộ hóa của rãnh này là mẫu truy cập ngẫu nhiên của rãnh, mẫu mã hóa video mở rộng được là mẫu đồng bộ hóa nếu mỗi hình ảnh được mã hóa trong đơn vị truy cập là hình ảnh IRAP, và

mẫu mã hóa video đa khung hình là mẫu đồng bộ hóa nếu mỗi hình ảnh được mã hóa trong đơn vị truy cập là hình ảnh IRAP mà không có hình ảnh hướng dẫn bỏ qua truy cập ngẫu nhiên (Random Access Skipped Leading - RASL).

8. Phương pháp theo điểm 5, trong đó phương pháp này còn bao gồm bước:

bắt đầu giải mã dữ liệu video đa lớp tại một trong số một hoặc nhiều mẫu chứa ít nhất một hình ảnh IRAP.

9. Thiết bị video bao gồm:

phương tiện lưu trữ dữ liệu được tạo cấu hình để lưu trữ dữ liệu video đa lớp; và

một hoặc nhiều bộ xử lý được tạo cấu hình để:

tạo tập tin bao gồm hộp rãnh chứa siêu dữ liệu cho một rãnh trong tập tin, trong đó:

dữ liệu phương tiện cho rãnh này bao gồm chuỗi các mẫu, mỗi mẫu là một đơn vị truy cập video trong số nhiều đơn vị truy cập video của dữ liệu video đa lớp; và

việc tạo tập tin bao gồm xác định, trong tập tin, nhóm mẫu ghi lại tất cả các mẫu chứa ít nhất một hình ảnh điểm truy cập ngẫu nhiên trong ảnh (Intra Random Access Point-IRAP), trong đó:

nhóm mẫu này được xác định là bao gồm mục nhập nhóm mẫu mà bao gồm:

giá trị chỉ báo liệu tất cả các hình ảnh được mã hóa trong mẫu cụ thể trong chuỗi mẫu có phải là hình ảnh IRAP hay không,

giá trị chỉ báo số lượng hình ảnh IRAP trong mẫu cụ thể;

và

các giá trị chỉ báo các ký hiệu nhận dạng lớp của các hình ảnh IRAP trong mẫu cụ thể, và  
mẫu cụ thể này chứa hình ảnh IRAP trong lớp cơ sở và còn chứa một hoặc nhiều hình ảnh không IRAP trong các lớp khác.

10. Thiết bị video theo điểm 9, trong đó một hoặc nhiều bộ xử lý được tạo cấu hình để:  
đưa vào, trong mục nhập nhóm mẫu, giá trị chỉ báo kiểu đơn vị lớp trừu tượng hóa mạng (Network Abstraction Layer-NAL) của các đơn vị NAL của lớp mã hóa video (Video Coding Layer-VCL) trong các hình ảnh IRAP của mẫu cụ thể.

11. Thiết bị video theo điểm 9, trong đó một hoặc nhiều bộ xử lý được tạo cấu hình để tạo hộp mẫu đồng bộ hóa trong tập tin bao gồm bảng mẫu đồng bộ hóa ghi lại các mẫu đồng bộ hóa của rãnh, trong đó:

mỗi mẫu đồng bộ hóa của rãnh này là mẫu truy cập ngẫu nhiên của rãnh,  
mẫu mã hóa video mở rộng được là mẫu đồng bộ hóa nếu mỗi hình ảnh được mã hóa trong đơn vị truy cập là hình ảnh IRAP, và

mẫu mã hóa video đa khung hình là mẫu đồng bộ hóa nếu mỗi hình ảnh được mã hóa trong đơn vị truy cập là hình ảnh IRAP mà không có hình ảnh hướng dẫn bỏ qua truy cập ngẫu nhiên (Random Access Skipped Leading - RASL).

12. Thiết bị video theo điểm 9, trong đó một hoặc nhiều bộ xử lý được tạo cấu hình để mã hóa dữ liệu video đa lớp.

13. Thiết bị video theo điểm 9, trong đó thiết bị video này bao gồm ít nhất một trong số:

mạch tích hợp;  
bộ vi xử lý; hoặc  
máy cầm tay không dây.

14. Thiết bị video theo điểm 9, trong đó thiết bị này còn bao gồm một hoặc nhiều máy quay được tạo cấu hình để ghi lại dữ liệu video đa lớp.

15. Thiết bị video bao gồm:

phương tiện lưu trữ dữ liệu được tạo cấu hình lưu trữ dữ liệu video đa lớp; và một hoặc nhiều bộ xử lý được tạo cấu hình để:

thu nhận, từ tập tin, hộp rãnh chứa siêu dữ liệu cho rãnh trong tập tin, trong đó dữ liệu phương tiện cho rãnh bao gồm chuỗi các mẫu, mỗi mẫu là một đơn vị truy cập video của dữ liệu video đa lớp;

thu nhận, từ tập tin, định nghĩa về nhóm mẫu ghi lại tất cả các mẫu chứa ít nhất một hình ảnh điểm truy cập ngẫu nhiên trong ảnh (Intra Random Access Point-IRAP), trong đó:

nhóm mẫu này được xác định là bao gồm mục nhập nhóm mẫu mà bao gồm:

giá trị chỉ báo liệu tất cả các hình ảnh được mã hóa trong mẫu cụ thể trong chuỗi mẫu có phải là hình ảnh IRAP hay không,

giá trị chỉ báo số lượng hình ảnh IRAP trong mẫu cụ thể;

và

các giá trị chỉ báo các ký hiệu nhận dạng lớp của các hình ảnh IRAP trong mẫu cụ thể, và

mẫu cụ thể này chứa hình ảnh IRAP trong lớp cơ sở và còn chứa một hoặc nhiều hình ảnh không IRAP trong các lớp khác; và

xác định, dựa trên nhóm mẫu này, một hoặc nhiều mẫu chứa ít nhất một hình ảnh IRAP.

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

thu nhận, từ mục nhập nhóm mẫu, giá trị chỉ báo kiểu đơn vị lớp trừu tượng hóa mạng (Network Abstraction Layer-NAL) của các đơn vị NAL của lớp mã hóa video (Video Coding Layer-VCL) trong các hình ảnh IRAP của mẫu cụ thể.

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

thu nhận, từ tập tin, hộp mẫu đồng bộ hóa bao gồm bảng mẫu đồng bộ hóa ghi lại các mẫu đồng bộ hóa của rãnh, trong đó:

mỗi mẫu đồng bộ hóa của rãnh này là mẫu truy cập ngẫu nhiên của rãnh, mẫu mã hóa video mở rộng được là mẫu đồng bộ hóa nếu mỗi hình ảnh được mã hóa trong đơn vị truy cập là hình ảnh IRAP, và

mẫu mã hóa video đa khung hình là mẫu đồng bộ hóa nếu mỗi hình ảnh được mã hóa trong đơn vị truy cập là hình ảnh IRAP mà không có hình ảnh hướng dẫn bỏ qua truy cập ngẫu nhiên (Random Access Skipped Leading - RASL).

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

bắt đầu giải mã dữ liệu video đa lớp tại một trong số một hoặc nhiều mẫu chứa ít nhất một hình ảnh IRAP.

19. Thiết bị video theo điểm 15, trong đó thiết bị video này bao gồm ít nhất một trong số:

mạch tích hợp;

bộ vi xử lý; hoặc

máy cầm tay không dây.

20. Thiết bị video theo điểm 15, trong đó thiết bị này còn bao gồm màn hình hiển thị được tạo cấu hình để hiển thị dữ liệu video đã giải mã.

21. Thiết bị video bao gồm:

phương tiện lưu trữ dữ liệu video đa lớp đã mã hóa; và

phương tiện tạo tập tin bao gồm hộp rãnh chứa siêu dữ liệu cho một rãnh trong tập tin, trong đó:

dữ liệu phương tiện cho rãnh này bao gồm chuỗi các mẫu, mỗi mẫu là một đơn vị truy cập video trong số nhiều đơn vị truy cập video của dữ liệu video đa lớp; và

việc tạo tập tin bao gồm xác định, trong tập tin, nhóm mẫu ghi lại tất cả các mẫu chứa ít nhất một hình ảnh điểm truy cập ngẫu nhiên trong ảnh (Intra Random Access Point-IRAP), trong đó:

nhóm mẫu này được xác định là bao gồm mục nhập nhóm mẫu mà bao gồm:

giá trị chỉ báo liệu tất cả các hình ảnh được mã hóa trong mẫu cụ thể trong chuỗi mẫu có phải là hình ảnh IRAP hay không, giá trị chỉ báo số lượng hình ảnh IRAP trong mẫu cụ thể; và

các giá trị chỉ báo các ký hiệu nhận dạng lớp của các hình ảnh IRAP trong mẫu cụ thể, và mẫu cụ thể này chứa hình ảnh IRAP trong lớp cơ sở và còn chứa một hoặc nhiều hình ảnh không IRAP trong các lớp khác.

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

phương tiện đưa vào, trong mục nhập nhóm mẫu, giá trị chỉ báo kiểu đơn vị lớp trừu tượng hóa mạng (Network Abstraction Layer-NAL) của các đơn vị NAL của lớp mã hóa video (Video Coding Layer-VCL) trong các hình ảnh IRAP của mẫu cụ thể.

23. Thiết bị video theo điểm 21, trong đó:

phương tiện tạo tập tin bao gồm phương tiện tạo, trong tập tin, hộp mẫu đồng bộ hóa bao gồm bảng mẫu đồng bộ hóa ghi lại các mẫu đồng bộ hóa của rãnh, mỗi mẫu đồng bộ hóa của rãnh này là mẫu truy cập ngẫu nhiên của rãnh, mẫu mã hóa video mở rộng được là mẫu đồng bộ hóa nếu mỗi hình ảnh được mã hóa trong đơn vị truy cập là hình ảnh IRAP, và

mẫu mã hóa video đa khung hình là mẫu đồng bộ hóa nếu mỗi hình ảnh được mã hóa trong đơn vị truy cập là hình ảnh IRAP mà không có hình ảnh hướng dẫn bỏ qua truy cập ngẫu nhiên (Random Access Skipped Leading - RASL).

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

phương tiện mã hóa dữ liệu video đa lớp.

25. Thiết bị video bao gồm:

phương tiện thu nhận, từ tập tin, hộp rãnh chứa siêu dữ liệu cho rãnh trong tập tin, trong đó dữ liệu phương tiện cho rãnh bao gồm chuỗi các mẫu, mỗi mẫu là một đơn vị truy cập video trong số nhiều đơn vị truy cập video của dữ liệu video đa lớp;

phương tiện thu nhận, từ tập tin, định nghĩa về nhóm mẫu ghi lại tất cả các mẫu chứa ít nhất một hình ảnh điểm truy cập ngẫu nhiên trong ảnh (Intra Random Access Point-IRAP), trong đó:

nhóm mẫu này được xác định là bao gồm mục nhập nhóm mẫu mà bao gồm:

giá trị chỉ báo liệu tất cả các hình ảnh được mã hóa trong mẫu cụ thể trong chuỗi mẫu có phải là hình ảnh IRAP hay không,

giá trị chỉ báo số lượng hình ảnh IRAP trong mẫu cụ thể; và

các giá trị chỉ báo các ký hiệu nhận dạng lớp của các hình ảnh IRAP trong mẫu cụ thể, và

mẫu cụ thể này chứa hình ảnh IRAP trong lớp cơ sở và còn chứa một hoặc nhiều hình ảnh không IRAP trong các lớp khác; và

phương tiện xác định, dựa trên nhóm mẫu, một hoặc nhiều mẫu chứa ít nhất một hình ảnh IRAP.

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

phương tiện thu nhận, từ mục nhập nhóm mẫu, giá trị chỉ báo kiểu đơn vị lớp trừu tượng hóa mạng (Network Abstraction Layer-NAL) của các đơn vị NAL của lớp mã hóa video (Video Coding Layer-VCL) trong các hình ảnh IRAP của mẫu cụ thể.

27. Thiết bị video theo điểm 25, thiết bị này còn bao gồm:

phương tiện thu nhận, từ tập tin, hộp mẫu đồng bộ hóa bao gồm bảng mẫu đồng bộ hóa ghi lại các mẫu đồng bộ hóa của rãnh, trong đó:

mỗi mẫu đồng bộ hóa của rãnh này là mẫu truy cập ngẫu nhiên của rãnh, mẫu mã hóa video mở rộng được là mẫu đồng bộ hóa nếu mỗi hình ảnh được mã hóa trong đơn vị truy cập là hình ảnh IRAP, và

mẫu mã hóa video đa khung hình là mẫu đồng bộ hóa nếu mỗi hình ảnh được mã hóa trong đơn vị truy cập là hình ảnh IRAP mà không có hình ảnh

hướng dẫn bỏ qua truy cập ngẫu nhiên (Random Access Skipped Leading - RASL).

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

phương tiện bắt đầu giải mã dữ liệu video đa lớp tại một trong số một hoặc nhiều mẫu chứa ít nhất một hình ảnh IRAP.

29. Vật ghi lưu trữ dữ liệu bất biến đọc được bởi máy tính lưu trữ các lệnh trên đó, mà khi được thực thi, khiến cho một hoặc nhiều bộ xử lý:

tạo tập tin bao gồm hộp rãnh chứa siêu dữ liệu cho một rãnh trong tập tin, trong đó:

dữ liệu phương tiện cho rãnh bao gồm chuỗi các mẫu, mỗi mẫu là một đơn vị truy cập video trong số nhiều đơn vị truy cập video của dữ liệu video đa lớp,

để tạo tập tin, một hoặc nhiều bộ xử lý xác định, trong tập tin, nhóm mẫu ghi lại tất cả các mẫu chứa ít nhất một hình ảnh điểm truy cập ngẫu nhiên trong ảnh (Intra Random Access Point-IRAP), trong đó:

nhóm mẫu này được xác định là bao gồm mục nhập nhóm mẫu mà bao gồm:

giá trị chỉ báo liệu tất cả các hình ảnh được mã hóa trong mẫu cụ thể trong chuỗi mẫu có phải là hình ảnh IRAP hay không,

giá trị chỉ báo số lượng hình ảnh IRAP trong mẫu cụ thể; và

các giá trị chỉ báo các ký hiệu nhận dạng lớp của các hình ảnh IRAP trong mẫu cụ thể, và

mẫu cụ thể này chứa hình ảnh IRAP trong lớp cơ sở và còn chứa một hoặc nhiều hình ảnh không IRAP trong các lớp khác.

30. Vật ghi lưu trữ dữ liệu bất biến đọc được bởi máy tính theo điểm 29, trong đó các lệnh khiến cho một hoặc nhiều bộ xử lý:

đưa vào, trong mục nhập nhóm mẫu, giá trị chỉ báo kiểu đơn vị lớp trừu tượng hóa mạng (Network Abstraction Layer-NAL) của các đơn vị NAL của lớp mã hóa video (Video Coding Layer-VCL) trong các hình ảnh IRAP của mẫu cụ thể.

31. Vật ghi lưu trữ dữ liệu bất biến đọc được bởi máy tính theo điểm 29, trong đó các lệnh còn khiến cho một hoặc nhiều bộ xử lý:

tạo, trong tập tin, hộp mẫu đồng bộ hóa bao gồm bảng mẫu đồng bộ hóa ghi lại các mẫu đồng bộ hóa của rãnh, trong đó:

mỗi mẫu đồng bộ hóa của rãnh này là mẫu truy cập ngẫu nhiên của rãnh, mẫu mã hóa video mở rộng được là mẫu đồng bộ hóa nếu mỗi hình ảnh được mã hóa trong đơn vị truy cập là hình ảnh IRAP, và

mẫu mã hóa video đa khung hình là mẫu đồng bộ hóa nếu mỗi hình ảnh được mã hóa trong đơn vị truy cập là hình ảnh IRAP mà không có hình ảnh hướng dẫn bỏ qua truy cập ngẫu nhiên (Random Access Skipped Leading - RASL).

32. Vật ghi lưu trữ dữ liệu bất biến đọc được bởi máy tính có các lệnh trên đó, mà khi được thực thi, khiến cho một hoặc nhiều bộ xử lý:

thu nhận, từ tập tin, hộp rãnh chứa siêu dữ liệu cho rãnh trong tập tin, trong đó dữ liệu phương tiện cho rãnh bao gồm chuỗi các mẫu, mỗi mẫu là một đơn vị truy cập video trong số nhiều đơn vị truy cập video của dữ liệu video đa lớp;

thu nhận, từ tập tin, định nghĩa về nhóm mẫu ghi lại tất cả các mẫu chứa ít nhất một hình ảnh điểm truy cập ngẫu nhiên trong ảnh (Intra Random Access Point-IRAP), trong đó:

nhóm mẫu này được xác định là bao gồm mục nhập nhóm mẫu mà bao gồm:

giá trị chỉ báo liệu tất cả các hình ảnh được mã hóa trong mẫu cụ thể trong chuỗi mẫu có phải là hình ảnh IRAP hay không,

giá trị chỉ báo số lượng hình ảnh IRAP trong mẫu cụ thể; và

các giá trị chỉ báo các ký hiệu nhận dạng lớp của các hình ảnh IRAP trong mẫu cụ thể, và

mẫu cụ thể này chứa hình ảnh IRAP trong lớp cơ sở và còn chứa một hoặc nhiều hình ảnh không IRAP trong các lớp khác; và

xác định, dựa trên nhóm mẫu, một hoặc nhiều mẫu chứa ít nhất một hình ảnh IRAP.

33. Vật ghi lưu trữ dữ liệu bất biến đọc được bởi máy tính theo điểm 32, trong đó các lệnh này khiến cho một hoặc nhiều bộ xử lý:

thu nhận, từ mục nhập nhóm mẫu, giá trị chỉ báo kiểu đơn vị lớp trừu tượng hóa mạng (Network Abstraction Layer-NAL) của các đơn vị NAL của lớp mã hóa video (Video Coding Layer-VCL) trong các hình ảnh IRAP của mẫu cụ thể.

34. Vật ghi lưu trữ dữ liệu bất biến đọc được bởi máy tính theo điểm 32, trong đó các lệnh khiến cho một hoặc nhiều bộ xử lý:

thu nhận, từ tập tin, hộp mẫu đồng bộ hóa bao gồm bảng mẫu đồng bộ hóa ghi lại các mẫu đồng bộ hóa của rãnh, trong đó:

mỗi mẫu đồng bộ hóa của rãnh này là mẫu truy cập ngẫu nhiên của rãnh, mẫu mã hóa video mở rộng được là mẫu đồng bộ hóa nếu mỗi hình ảnh được mã hóa trong đơn vị truy cập là hình ảnh IRAP, và

mẫu mã hóa video đa khung hình là mẫu đồng bộ hóa nếu mỗi hình ảnh được mã hóa trong đơn vị truy cập là hình ảnh IRAP mà không có hình ảnh hướng dẫn bỏ qua truy cập ngẫu nhiên (Random Access Skipped Leading - RASL).

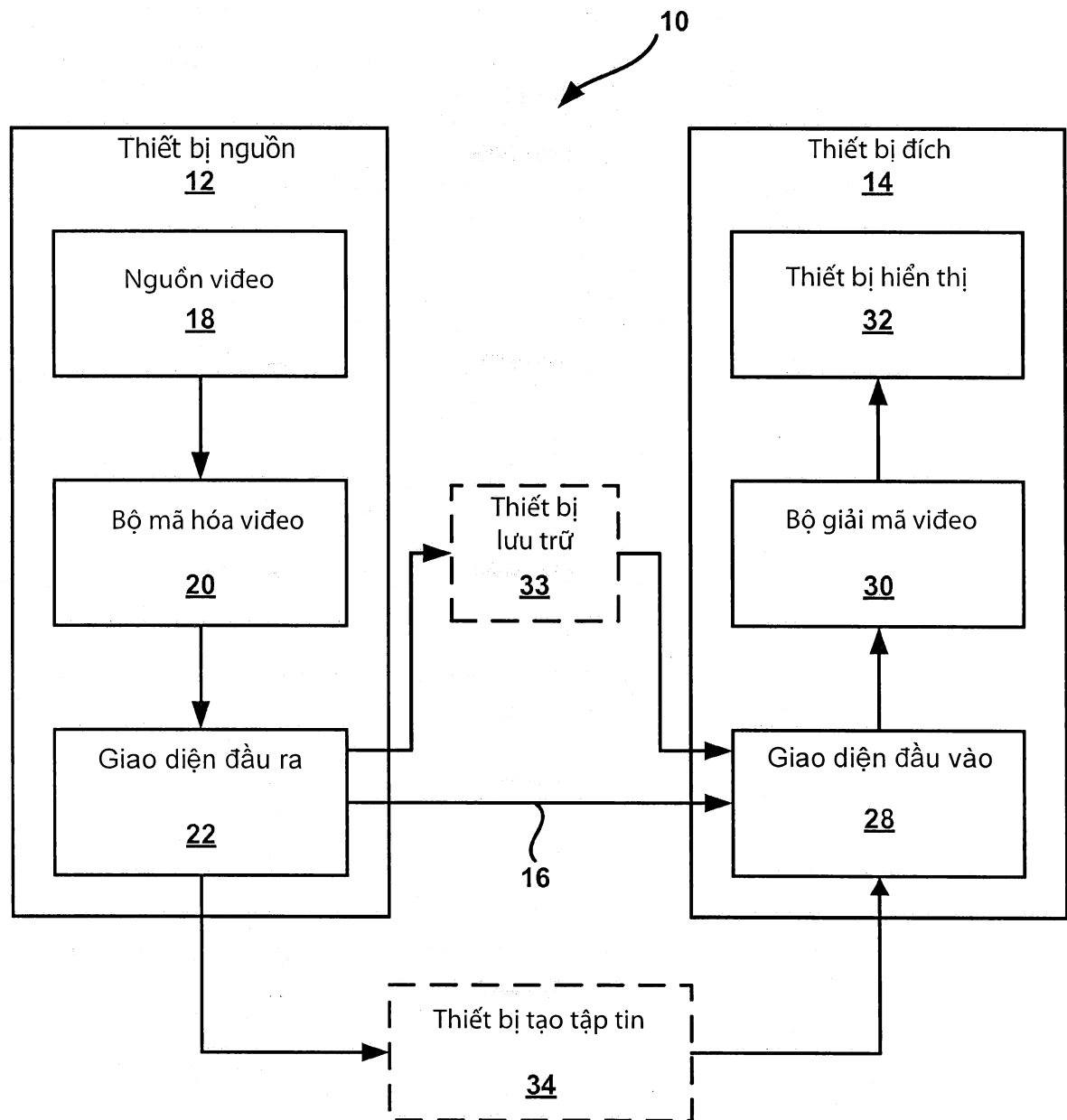
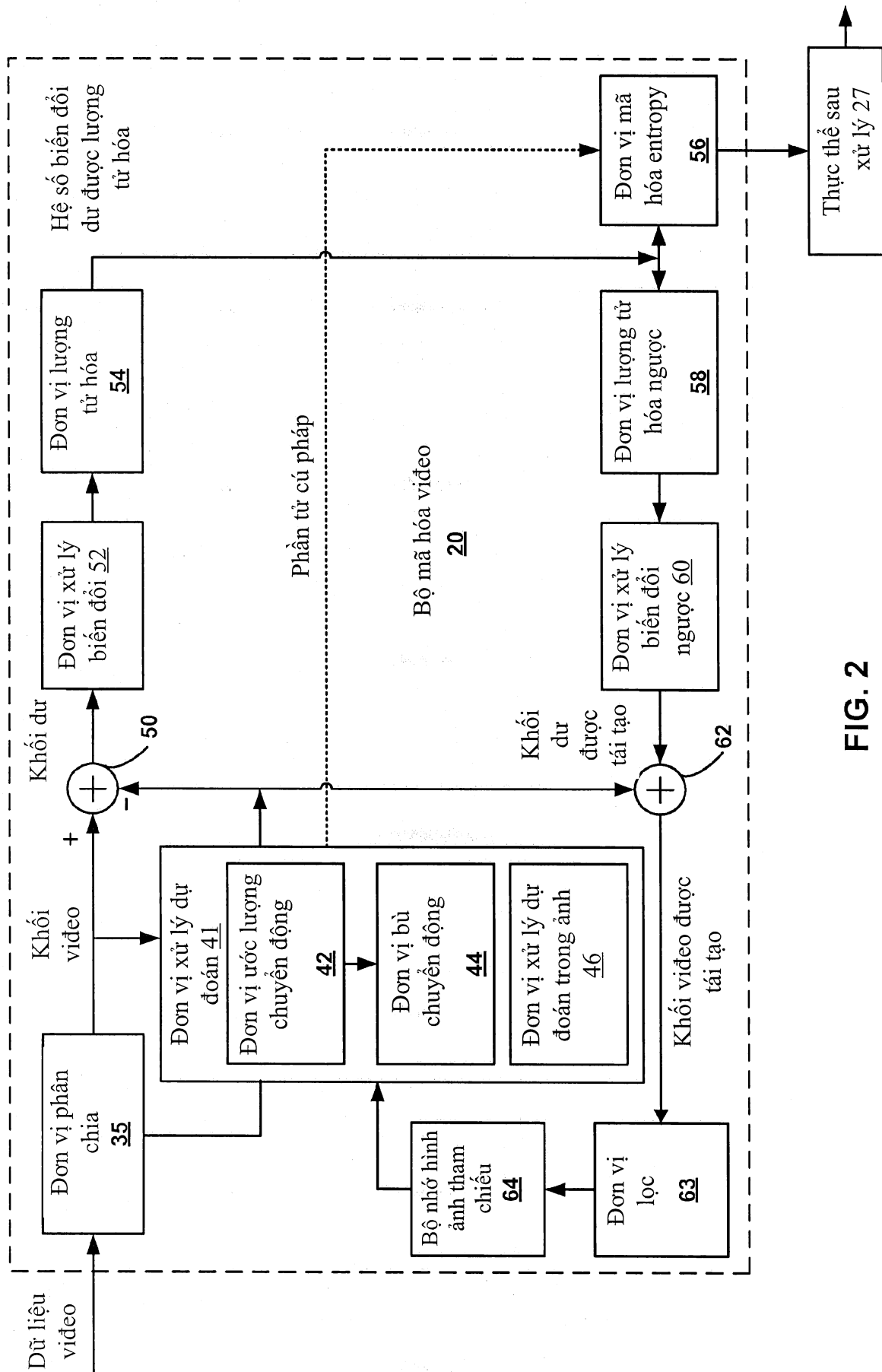


FIG. 1



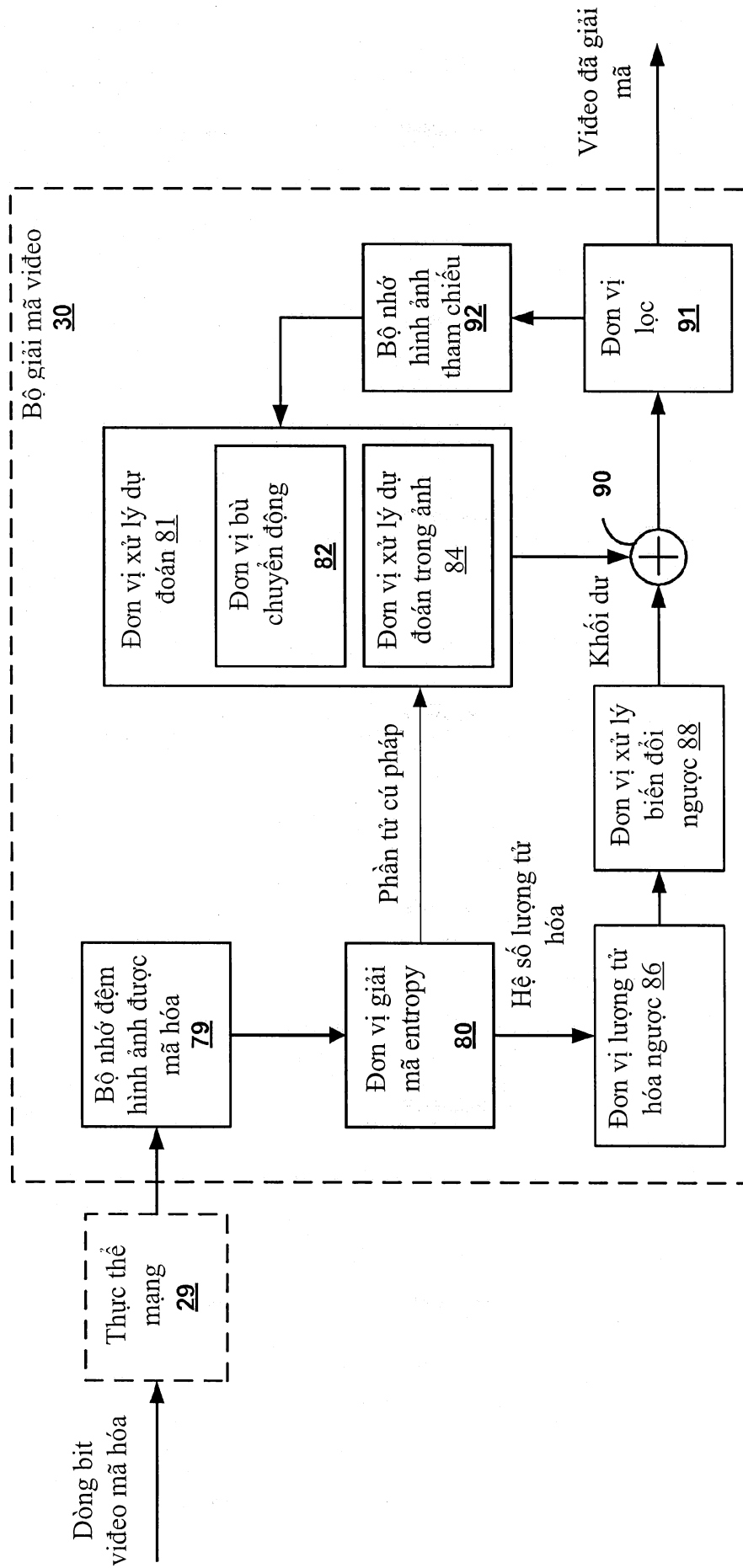


FIG. 3

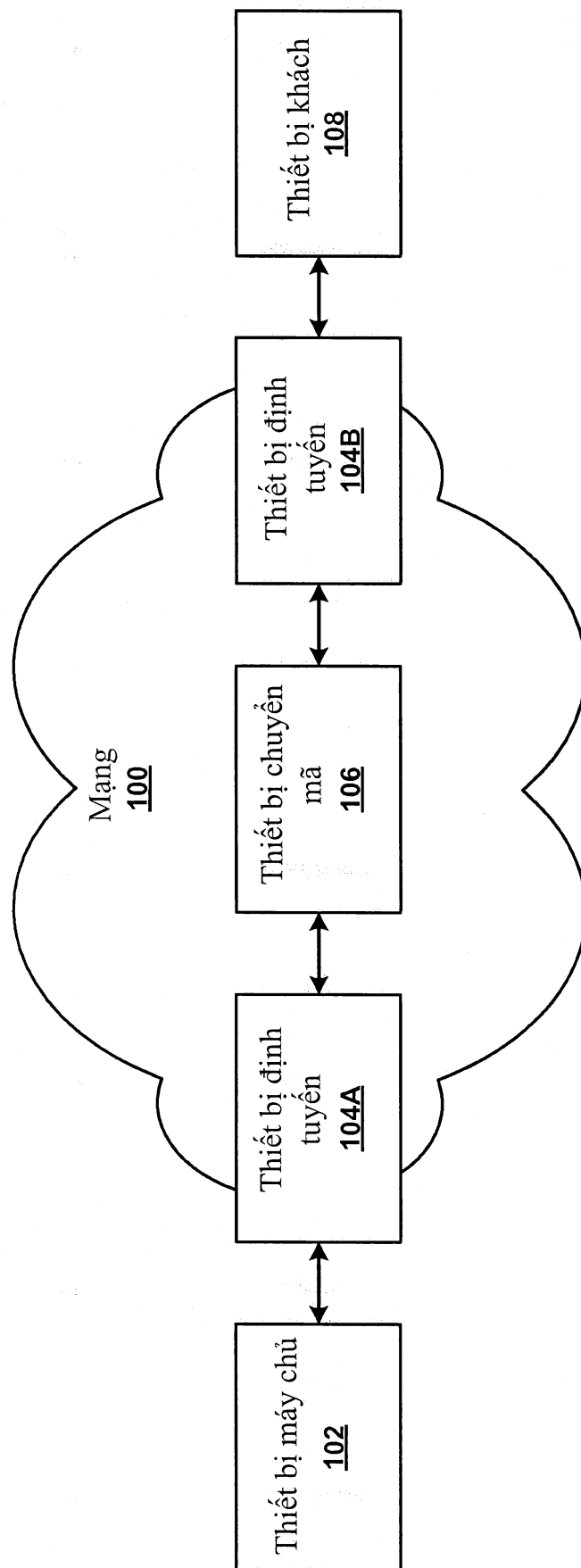


FIG. 4

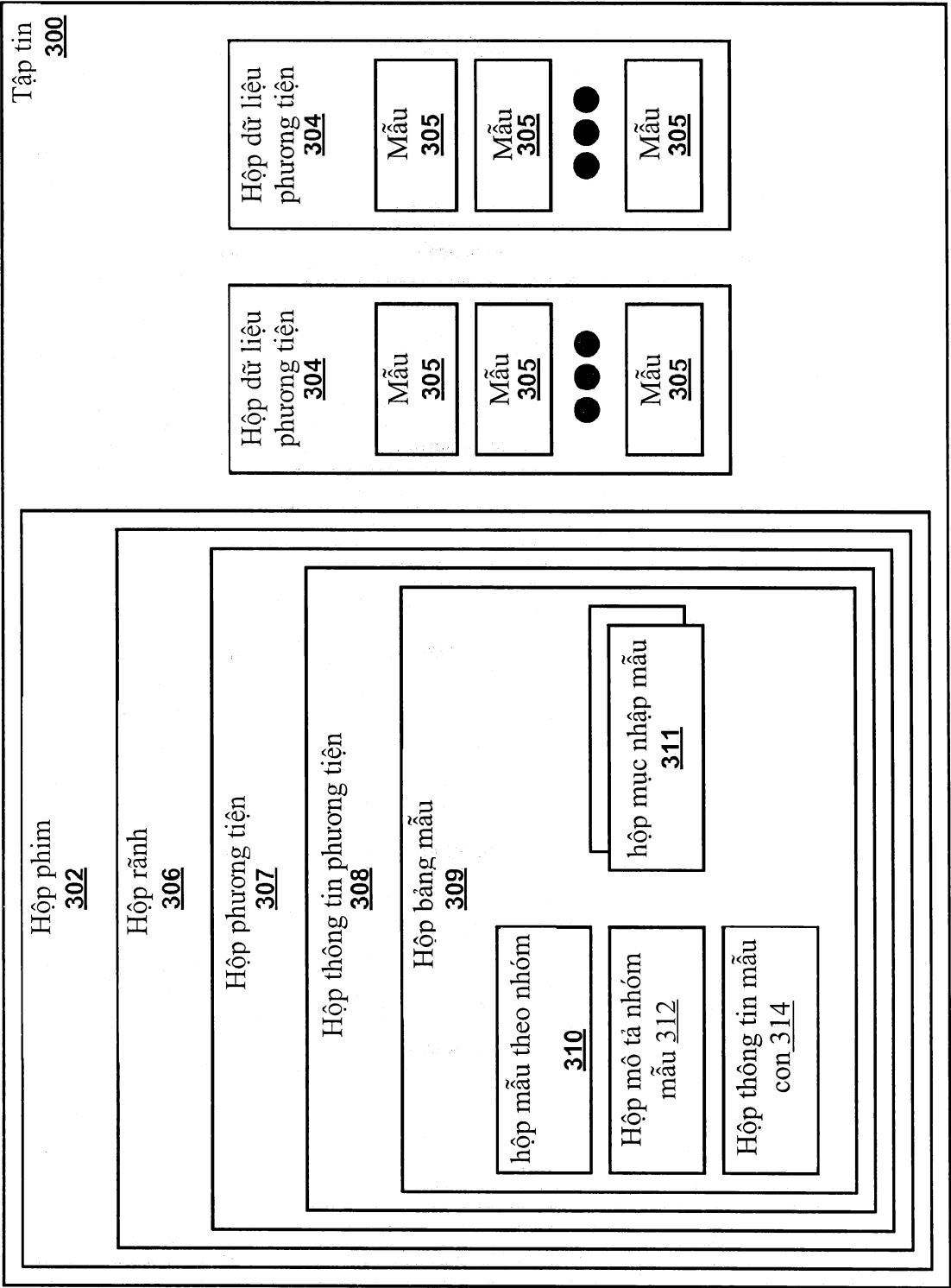


FIG. 5

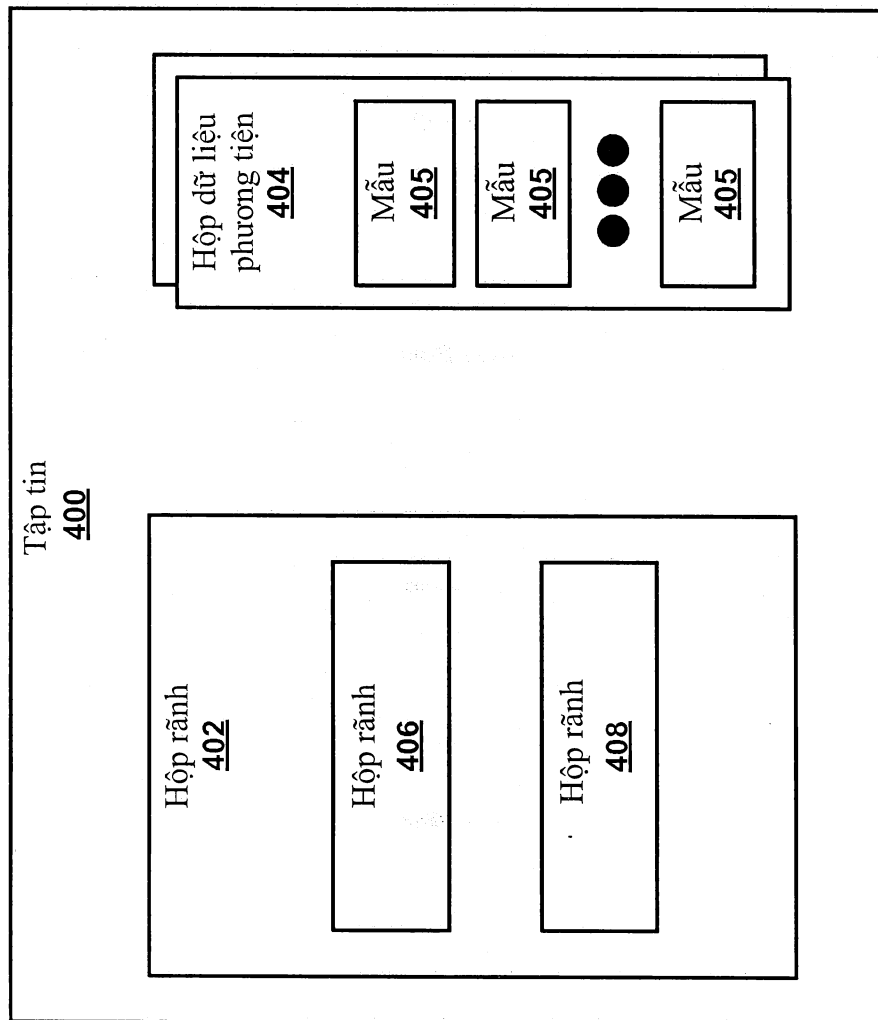


FIG. 6

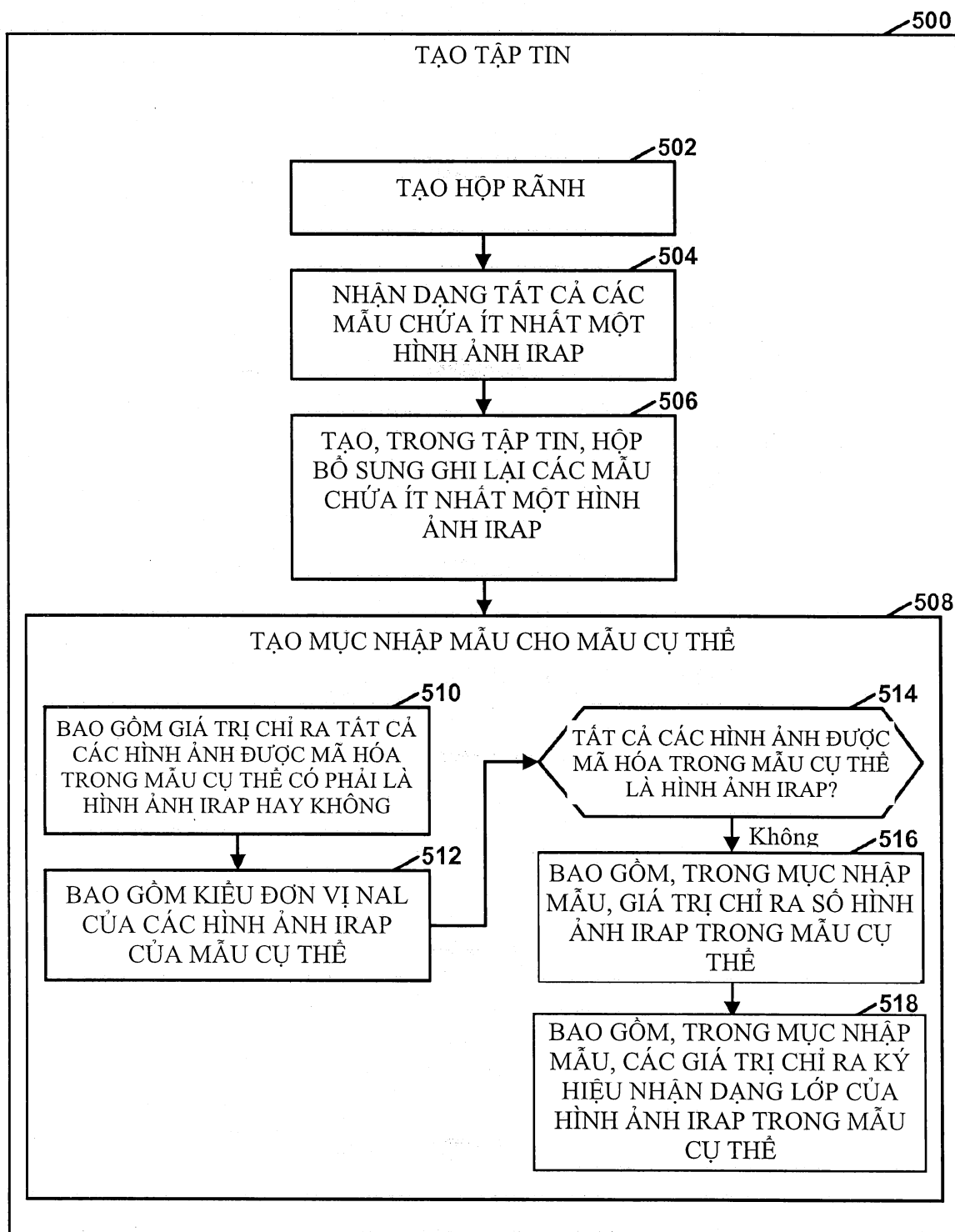


FIG. 7

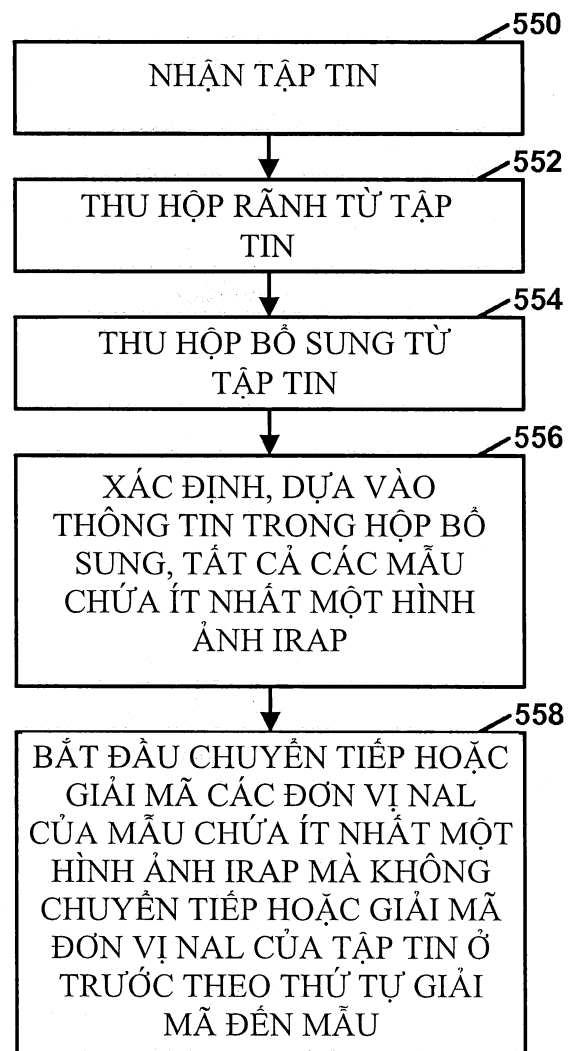


FIG. 8

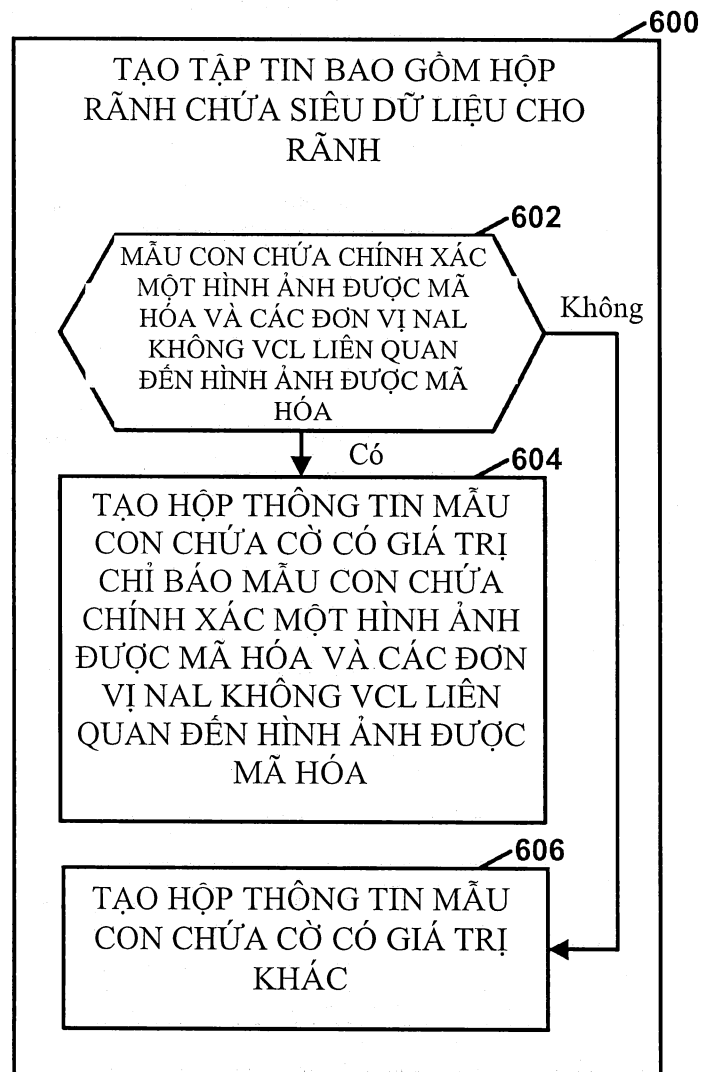


FIG. 9

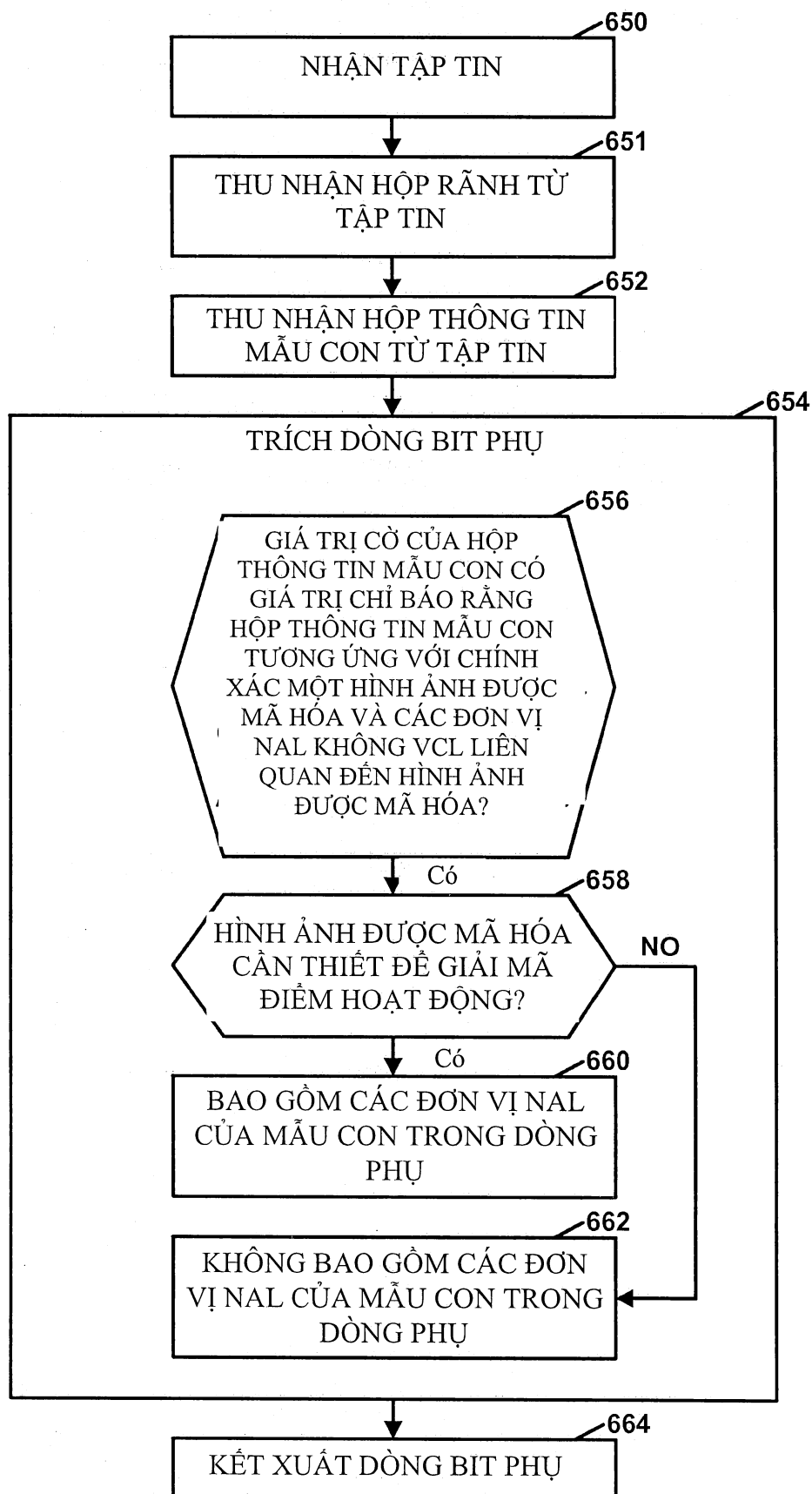


FIG. 10

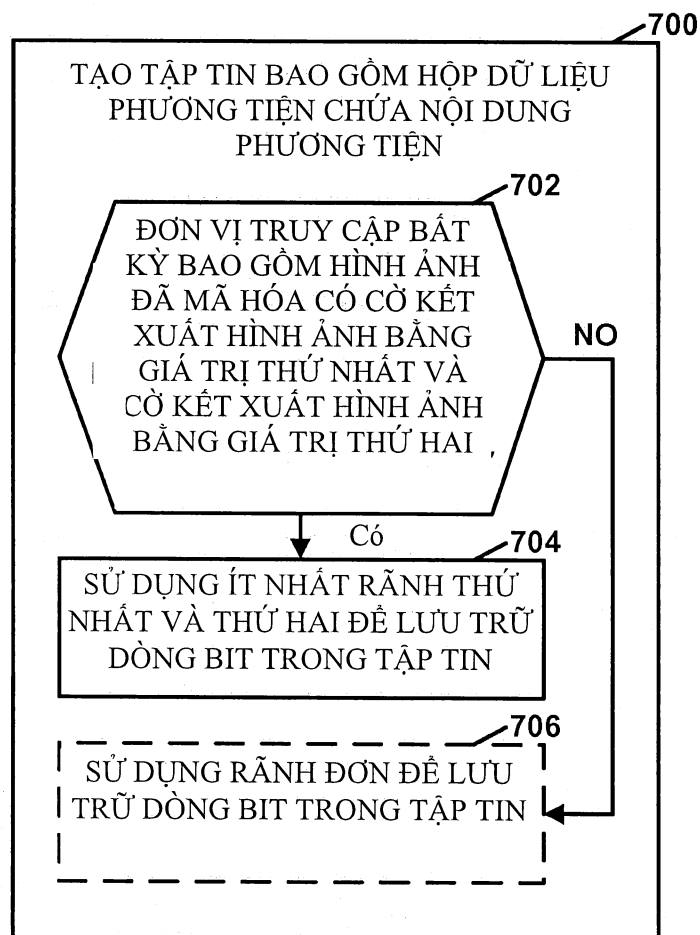


FIG. 11

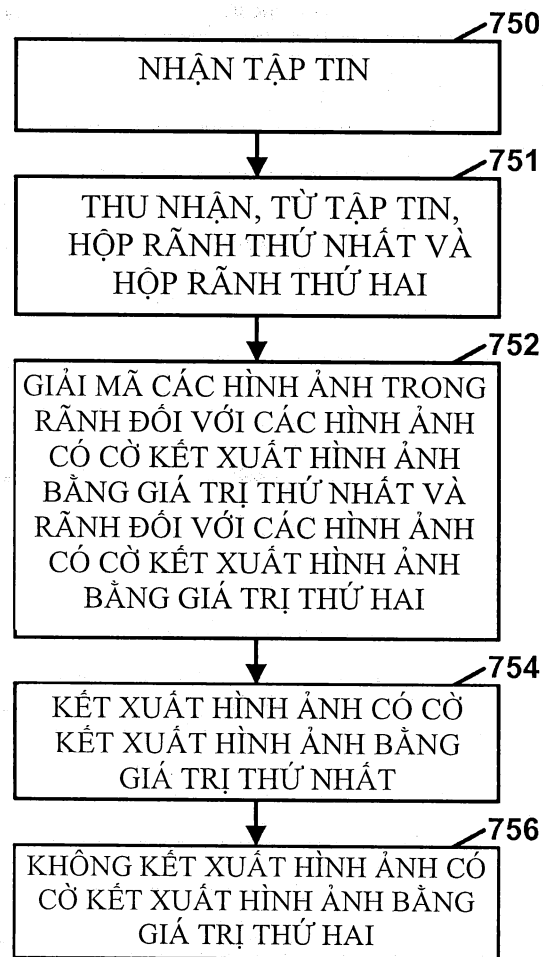


FIG. 12