



(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)



1-0036736

(51)⁷ H04N 19/105; H04N 19/137; H04N 19/167; H04N 21/4402; H04N 19/42; H04N 19/46; H04N 19/517; H04N 21/2343; G02B 27/01; H04N 19/174 (13) B

(21) 1-2018-02843

(22) 14/02/2017

(86) PCT/FI2017/050086 14/02/2017

(87) WO 2017/140945 24/08/2017

(30) 20165114 17/02/2016 FI

(45) 25/08/2023 425

(43) 26/11/2018 368A

(73) Nokia Technologies Oy (FI)

Karaportti 3, 02610 Espoo, Finland

(72) HANNUKSELA, Miska (FI); AMINLOU, Alireza (IR); ZARE, Alireza (IR).

(74) Công ty TNHH Tầm nhìn và Liên danh (VISION & ASSOCIATES CO.LTD.)

(54) THIẾT BỊ, PHƯƠNG PHÁP VÀ VẬT GHI ĐỀ MÃ HÓA VÀ GIẢI MÃ VIDEO

(57) Sáng chế đề cập đến phương pháp bao gồm bước thu được đường ghi hoặc dòng bit của hình ảnh đầy đủ bao gồm tập hợp ô bị hạn chế chuyển động; và xây dựng đường ghi chứa tập hợp ô hoặc dòng bit tuân theo hình ảnh đầy đủ trên cơ sở tập hợp ô bị hạn chế chuyển động hoặc tạo ra các lệnh để xây dựng đường ghi chứa tập hợp ô hoặc dòng bit tuân theo hình ảnh đầy đủ trên cơ sở tập hợp ô bị hạn chế chuyển động.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến thiết bị, phương pháp và chương trình máy tính dùng để mã hóa và giải mã video.

Tình trạng kỹ thuật của sáng chế

Trong nhiều tình huống, chỉ một phần hình ảnh cần được hiển thị trong khi phần còn lại đòi hỏi phải được giải mã nhưng không được hiển thị. Những tình huống hiển thị này gồm có, ví dụ, các màn hiển thị đeo đầu (head-mounted display, HMD) hiển thị với phạm vi quan sát khoảng 100 độ, trong khi đó video đầu vào cho nhu cầu của HMD thường bao quát toàn bộ 360 độ. Các môi trường quan sát màn hình phẳng thông thường hiển thị phạm vi quan sát lên đến 40 độ. Khi hiển thị nội dung phạm vi quan sát rộng (ví dụ, mắt cá) trên màn hiển thị này, có thể ưu tiên hiển thị tập hợp con không gian thay vì toàn bộ hình ảnh.

Thông thường, các hình ảnh được mã hóa cần được giải mã hoàn toàn, do chúng có các phần phụ thuộc mạnh về không gian (ví dụ, dự đoán nội bộ), các phần phụ thuộc mã hóa entropy, các phần phụ thuộc lọc theo vòng lặp, và các phần phụ thuộc không - thời gian (các vector chuyển động chỉ vào phần khác biệt của hình ảnh). Tập hợp ô (tile set) bị hạn chế chuyển động cho phép giải mã tập hợp ô này độc lập với các tập hợp ô khác; tuy nhiên, do tập hợp ô được mã hóa có thể được chèn dữ liệu được mã hóa khác trong dòng bit và do các bộ thông số và các phần đầu (ví dụ, phần đầu phân đoạn mảng) dùng cho toàn bộ dòng bit, nên quy trình giải mã chuyên dụng là cần thiết để giải mã tập hợp ô bị hạn chế chuyển động cụ thể, trong khi bỏ qua việc giải mã các ô (tile) khác.

Trong các cách bố trí phát trực tuyến thông thường đối với video 360 độ, video 360 độ được truyền toàn phần. Ví dụ, toàn bộ hình ảnh toàn cảnh hình chữ nhật được truyền. Tuy nhiên, trong nhiều môi trường hiển thị, như nhu cầu trên HMD, chỉ một tập hợp con của video 360-độ là được hiển thị tại một thời điểm. Tốc độ bit phát trực tuyến vì thế được sử dụng cho các vùng ảnh vốn không được hiển thị. Điều này có

thể gây ra nguy cơ cho mạng và/hoặc nghẽn đường truy cập và tạo đệm lại và/hoặc việc sử dụng chất lượng tương đối thấp cho các phần của vùng ảnh không hiển thị.

Bản chất kỹ thuật của sáng chế

Để khắc phục các vấn đề nêu trên, sáng chế đề xuất phương pháp cải tiến để tạo ra dòng bit video.

Phương pháp theo khía cạnh thứ nhất bao gồm bước thu được đường ghi hoặc dòng bit của hình ảnh đầy đủ bao gồm tập hợp ô bị hạn chế chuyển động; và bước xây dựng đường ghi chứa tập hợp ô hoặc dòng bit tuân theo hình ảnh đầy đủ trên cơ sở của tập hợp ô bị hạn chế chuyển động.

Phương pháp theo khía cạnh thứ hai bao gồm bước thu được đường ghi hoặc dòng bit của hình ảnh đầy đủ bao gồm tập hợp ô bị hạn chế chuyển động; và bước xây dựng đường ghi chứa tập hợp ô hoặc dòng bit tuân theo hình ảnh đầy đủ trên cơ sở của tập hợp ô bị hạn chế chuyển động.

Theo một phương án, tập hợp ô bị hạn chế chuyển động đặc trưng ở chỗ được chứa trong một hay nhiều phân đoạn mảng vốn cũng chứa các ô bên ngoài tập hợp ô bị hạn chế chuyển động, và bước xây dựng hoặc tạo ra các lệnh để xây dựng bao gồm: bổ sung, vào đường ghi chứa tập hợp ô hoặc dòng bit tuân theo hình ảnh đầy đủ hoặc vào các lệnh để xây dựng đường ghi chứa tập hợp ô hoặc dòng bit tuân theo hình ảnh đầy đủ, cột ô ở bên phải của tập hợp ô bị hạn chế chuyển động.

Theo một phương án, phương pháp này còn bao gồm bước điều khiển sự kết thúc của phân đoạn mảng với các ô của cột ô được bổ sung.

Theo một phương án, phương pháp này còn bao gồm bước chỉ ra trong ô cuối cùng của cột ô được bổ sung sự kết thúc của phân đoạn mảng.

Theo một phương án, phương pháp này còn bao gồm bước bổ sung, vào đường ghi chứa tập hợp ô hoặc dòng bit tuân theo hình ảnh đầy đủ hoặc vào các lệnh để xây dựng đường ghi chứa tập hợp ô hoặc dòng bit tuân theo hình ảnh đầy đủ, thông tin trên cửa sổ cắt ảnh phù hợp khiến cho việc giải mã loại trừ cột ô được bổ sung từ việc giải mã đầu ra.

Theo một phương án, tập hợp ô bị hạn chế chuyển động đặc trưng ở chỗ

được chứa trong một hay nhiều phân đoạn mảng vốn cũng chứa các ô bên ngoài tập hợp ô bị hạn chế chuyển động, và xây dựng hoặc tạo ra các lệnh để xây dựng bao gồm: nhận diện ô ngoài cùng bên phải trong tập hợp ô bị hạn chế chuyển động mà không phải là ô cuối cùng của phân đoạn mảng thứ nhất chứa ô ngoài cùng bên phải; và sửa đổi ô ngoài cùng bên phải sao cho nó gây ra sự kết thúc của mảng thứ nhất.

Phương pháp theo khía cạnh thứ ba bao gồm bước thu được đường ghi chứa tập hợp ô hoặc dòng bit tuân theo hình ảnh đầy đủ được xây dựng, hoặc các lệnh để xây dựng đường ghi chứa tập hợp ô hoặc dòng bit tuân theo hình ảnh đầy đủ; và bước tạo ra thông tin, trong phần mô tả trình bày phương tiện, của hiển thị thứ nhất được tạo thành từ đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ.

Theo một phương án, phương pháp này còn bao gồm việc tạo ra thông tin, trong phần mô tả trình bày phương tiện, của hiển thị thứ hai được tạo thành từ đường ghi hoặc dòng bit của hình ảnh đầy đủ; kết hợp hiển thị thứ nhất và hiển thị thứ hai vào cùng một không gian tham chiếu theo phần mô tả trình bày phương tiện; và biểu thị vị trí và kích thước của hiển thị thứ nhất liên quan đến hiển thị thứ hai theo phần mô tả trình bày phương tiện.

Theo một phương án, các lệnh để xây dựng đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ bao gồm bộ trích xuất mà trích xuất một phần của bộ phận cơ bản từ đường ghi hoặc dòng bit của hình ảnh đầy đủ.

Theo một phương án, bộ phận cơ bản là gói lớp trừu tượng hóa mạng (Bộ NAL).

Theo một phương án, việc xây dựng hoặc tạo ra các lệnh để xây dựng đường ghi chứa tập hợp ô hoặc dòng bit tuân theo hình ảnh đầy đủ bao gồm việc ghi lại các bộ thông số và/hoặc toàn bộ hoặc một phần các phần đầu của phân đoạn mảng nếu so với những việc đó trong đường ghi hoặc dòng bit của hình ảnh đầy đủ.

Theo một phương án, các lệnh để xây dựng đường ghi chứa tập hợp ô hoặc dòng bit tuân theo hình ảnh đầy đủ đưa đến kết quả thành chuỗi của các phân đoạn hoặc các phân đoạn con.

Phương pháp theo khía cạnh thứ tư bao gồm bước thu được các đặc tính thứ

nhất của đường ghi tuân theo hình ảnh đầy đủ thứ nhất từ phần mô tả trình bày phương tiện và/hoặc từ một tập đa phương tiện; việc xác định dựa trên các đặc tính thứ nhất để thu được một phần hoặc toàn bộ đường ghi tuân theo hình ảnh đầy đủ thứ nhất; và thu được một phần hoặc toàn bộ đường ghi tuân theo hình ảnh đầy đủ thứ nhất.

Theo một phương án, phương pháp này còn bao gồm bước giải mã và phát một phần hoặc toàn bộ đường ghi tuân theo hình ảnh đầy đủ thứ nhất.

Theo một phương án, phương pháp này còn bao gồm bước thu được, theo các đặc tính thứ nhất, thông tin của vùng không gian được bao phủ hoặc cả phạm vi quan sát và hướng nhìn được bao phủ, trong đó việc xác định này dựa trên vùng không gian được bao phủ hoặc cả phạm vi quan sát và hướng nhìn được bao phủ.

Theo một phương án, phương pháp này còn bao gồm bước thu được các đặc tính thứ hai của đường ghi tuân theo hình ảnh đầy đủ thứ hai từ phần mô tả trình bày phương tiện và/hoặc từ tập đa phương tiện, kết thúc từ các đặc tính thứ nhất và thứ hai mà đường ghi tuân theo hình ảnh đầy đủ thứ hai có độ phân giải không gian, tốc độ bit, chất lượng hình ảnh, độ sâu bit cho mỗi thành phần màu, dải tần nhạy sáng, và/hoặc độ rộng của dải màu thấp hơn so với các đặc tính này của đường ghi tuân theo hình ảnh đầy đủ thứ nhất; kết thúc từ các đặc tính thứ nhất và thứ hai mà các đường ghi tuân theo hình ảnh đầy đủ thứ nhất và thứ hai đại diện cho các phần không gian khác nhau của nội dung; xác định dựa trên các kết luận trên để thu được một phần hoặc toàn bộ đường ghi tuân theo hình ảnh đầy đủ thứ hai; và thu được một phần hoặc toàn bộ đường ghi tuân theo hình ảnh đầy đủ thứ hai. Các khía cạnh khác bao gồm ít nhất là các thiết bị và các sản phẩm chương trình máy tính/mã được lưu giữ trên phương tiện nhớ không chuyển tiếp được sắp xếp để thực hiện các phương pháp trên.

Mô tả vắn tắt các hình vẽ

Để hiểu rõ ràng hơn về sáng chế, cần phải tham khảo các hình vẽ kèm theo, trong đó:

Fig.1 thể hiện bằng sơ đồ thiết bị điện tử ứng dụng các phương án của sáng chế;

Fig.2 thể hiện bằng sơ đồ thiết bị của người dùng thích hợp để ứng dụng các

phương án của sáng chế;

Fig.3 thể hiện thêm bằng sơ đồ các thiết bị điện tử ứng dụng các phương án của sáng chế được kết nối nhờ sử dụng các kết nối mạng không dây và có dây;

Fig.4 thể hiện bằng sơ đồ bộ mã hóa thích hợp để thực hiện các phương án của sáng chế;

Fig.5 thể hiện ví dụ về mô hình dữ liệu phân cấp sử dụng trong DASH;

Fig.6a và Fig.6b thể hiện các lưu đồ tạo ra các lệnh dừng cho việc xây dựng dòng bit và xây dựng dòng bit theo các phương án của sáng chế;

Fig.7 thể hiện khái niệm cơ bản làm cơ sở cho các phương án khác nhau của sáng chế;

Fig.8 thể hiện ví dụ về sự phân chia các tập hợp ô bị hạn chế chuyển động;

Fig.9 thể hiện quy trình tạo ra các lệnh dừng cho việc xây dựng đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ theo một phương án;

Fig.10 thể hiện ví dụ về cấu trúc tệp theo một phương án;

Fig.11a và Fig.11b thể hiện ví dụ của việc sắp xếp các khúc dữ liệu của các đường ghi khác nhau trong tệp theo một phương án, và ví dụ về toán tử hàm dừng theo một phương án;

Fig.12 thể hiện ví dụ về việc tạo ra MPD theo một phương án;

Fig.13 thể hiện ví dụ về việc áp dụng các phương án đối với các phiên bản khác nhau của cùng một nội dung với độ phân giải không gian khác nhau;

Fig.14 thể hiện ví dụ về việc áp dụng các phương án đối với các lớp có thể mở rộng khác nhau của cùng một nội dung với độ phân giải không gian khác nhau;

Fig.15 thể hiện ví dụ về phương án được áp dụng để mã hóa video lập thể;

Fig.16 thể hiện ví dụ về phương án được áp dụng để mã hóa video lập thể với dự đoán liên khung hình ;

Fig.17 thể hiện ví dụ tạo ra gợi ý phân đoạn theo một phương án;

Fig.18 thể hiện ví dụ về đường ghi gợi ý phân đoạn hoặc tệp phân đoạn theo một phương án;

Fig.19 thể hiện hoạt động của máy chủ hoặc bộ phân đoạn theo một phương án;

Fig.20 thể hiện ví dụ về một phương án liên quan đến việc tạo ra các đường ghi chứa tập hợp ô đã được xây dựng;

Fig.21 thể hiện ví dụ về việc sắp xếp các khúc dữ liệu của các đường ghi chứa tập hợp ô đã được xây dựng và các đường ghi hình ảnh đầy đủ trong tệp theo một phương án;

Fig.22 thể hiện sơ đồ của bộ giải mã thích hợp để thực hiện các phương án của sáng chế; và

Fig.23 thể hiện sơ đồ ví dụ về hệ thống truyền thông đa phương tiện trong đó các phương án khác nhau có thể được sử dụng.

Mô tả chi tiết sáng chế

Phần mô tả sau đây sẽ mô tả chi tiết hơn về thiết bị thích hợp và các cơ cấu khả thi để làm giảm tốc độ bit nhận dữ liệu (bitrate) của việc phát trực tuyến dòng bit video mở rộng được. Trong đó, việc tham chiếu được thực hiện đối với Fig.1 và Fig.2, trong đó Fig.1 thể hiện sơ đồ khối của hệ thống mã hóa video theo phương án ví dụ như là sơ đồ khối của thiết bị hoặc thiết bị điện tử mẫu 50, vốn có thể kết hợp bộ mã hóa-giải mã theo một phương án của sáng chế. Fig.2 thể hiện bố trí của thiết bị theo phương án ví dụ. Các chi tiết của Fig.1 và Fig.2 sẽ được giải thích tiếp theo.

Thiết bị điện tử 50 có thể, ví dụ là thiết bị đầu cuối di động hoặc thiết bị người dùng của hệ thống truyền thông vô tuyến. Tuy nhiên, có thể là tốt nếu các phương án của sáng chế có thể được sử dụng trong thiết bị điện tử hoặc thiết bị bất kỳ vốn có thể đòi hỏi việc mã hóa và giải mã hoặc mã hóa hoặc giải mã các hình ảnh video.

Thiết bị 50 có thể bao gồm vỏ 30 để kết hợp và bảo vệ thiết bị. Thiết bị 50 còn có thể bao gồm màn hình 32 dưới dạng màn hình tinh thể lỏng. Theo các phương án khác của sáng chế, màn hình này có thể là công nghệ màn hình bất kỳ thích hợp để hiển thị hình ảnh hoặc video. Thiết bị 50 có thể bao gồm thêm bàn phím 34. Theo các phương án khác của sáng chế, bất kỳ dữ liệu hoặc cơ cấu giao diện người dùng nào

cũng có thể được sử dụng. Ví dụ, giao diện người dùng có thể được sử dụng là bàn phím ảo hoặc hệ thống nhập dữ liệu như là một phần của màn hình cảm ứng.

Thiết bị có thể bao gồm micrô 36 hoặc đầu vào âm thanh thích hợp bất kỳ có thể là đầu vào tín hiệu số hoặc tương tự. Thiết bị 50 có thể bao gồm thêm thiết bị đầu ra âm thanh mà trong các phương án của sáng chế, có thể là thiết bị bất kỳ trong số: tai nghe 38, loa, hoặc kết nối đầu ra âm thanh số hoặc âm thanh tương tự. Thiết bị 50 có thể còn bao gồm pin 40 (hoặc theo các phương án khác của sáng chế, thiết bị có thể được cấp nguồn bởi thiết bị cung cấp năng lượng di động thích hợp bất kỳ chẳng hạn như pin mặt trời, pin nhiên liệu hoặc máy phát kiểu bộ máy đồng hồ). Thiết bị có thể bao gồm thêm camera 42 có thể ghi hoặc chụp hình ảnh và/hoặc video. Thiết bị 50 có thể bao gồm thêm cổng hồng ngoại dùng để giao tiếp đường nhìn thẳng phạm vi ngắn đến các thiết bị khác. Theo các phương án khác, thiết bị 50 có thể bao gồm thêm giải pháp giao tiếp phạm vi ngắn thích hợp bất kỳ như, ví dụ kết nối không dây Bluetooth hoặc kết nối có dây USB/firewire.

Thiết bị 50 có thể bao gồm bộ điều khiển 56 hoặc bộ xử lý để điều khiển thiết bị 50. Bộ điều khiển 56 có thể được kết nối với bộ nhớ 58 vốn trong các phương án của sáng chế có thể lưu cả dữ liệu dưới dạng hình ảnh và dữ liệu âm thanh và/hoặc còn có thể lưu các lệnh để thực hiện trên bộ điều khiển 56. Bộ điều khiển 56 có thể được kết nối thêm với mạch mã hóa-giải mã 54 thích hợp để thực hiện việc mã hóa và giải mã dữ liệu audio và/hoặc video hoặc hỗ trợ trong việc mã hóa và giải mã được thực hiện bởi bộ điều khiển này.

Thiết bị 50 có thể bao gồm thêm bộ đọc thẻ 48 và thẻ thông minh 46, ví dụ bộ đọc UICC và UICC để cung cấp thông tin người dùng và là thích hợp để cung cấp thông tin xác thực để xác thực và ủy quyền của người dùng trên mạng.

Thiết bị 50 có thể bao gồm mạch giao diện radiô 52 được kết nối với bộ điều khiển và thích hợp để tạo ra các tín hiệu truyền thông không dây, ví dụ để kết nối với mạng truyền thông di động, hệ thống truyền thông không dây hoặc mạng nội bộ không dây. Thiết bị 50 có thể bao gồm thêm ăng-ten 44 được kết nối với mạch giao diện radiô 52 để truyền các tín hiệu tần số radiô phát ra tại mạch giao diện radiô 52 đến (các) thiết bị khác và để nhận các tín hiệu tần số radiô từ (các) thiết bị khác.

Thiết bị 50 có thể bao gồm camera có thể ghi lại hoặc phát hiện các khung hình riêng lẻ vốn sau đó được chuyển đến bộ mã hóa-giải mã 54 hoặc bộ điều khiển để xử lý. Thiết bị này có thể nhận dữ liệu hình ảnh video để xử lý từ thiết bị khác trước khi truyền và/hoặc lưu. Thiết bị 50 cũng có thể nhận theo cách không dây cũng như bằng kết nối có dây hình ảnh để mã hóa /giải mã.

Theo Fig.3, ví dụ về hệ thống trong đó thể hiện các phương án của sáng chế có thể được ứng dụng. Hệ thống 10 bao gồm các thiết bị đa truyền thông có thể kết nối thông qua một hay nhiều mạng. Hệ thống 10 có thể bao gồm sự kết hợp bất kỳ của mạng có dây và không dây gồm có, nhưng không giới hạn ở mạng điện thoại di động không dây (như mạng GSM, UMTS, CDMA, v.v.), mạng nội bộ không dây (WLAN) như được xác định bởi một trong số các tiêu chuẩn IEEE 802.x, mạng cá nhân Bluetooth, mạng nội bộ Ethernet, mạng nội bộ token ring, mạng diện rộng (wide area network), và Internet.

Hệ thống 10 có thể bao gồm cả các thiết bị truyền thông có dây và không dây và/hoặc thiết bị 50 thích hợp để thực hiện các phương án của sáng chế.

Ví dụ, hệ thống trên Fig.3 thể hiện mạng điện thoại di động 11 và hiển thị của internet 28. Sự kết nối với internet 28 có thể gồm có, nhưng không giới hạn ở, các kết nối không dây phạm vi dài, các kết nối không dây phạm vi ngắn, và các kết nối có dây khác nhau gồm có, nhưng không giới hạn ở, dây điện thoại, dây cáp, dây điện, và các đường dẫn truyền thông cùng loại.

Các thiết bị truyền thông làm ví dụ thể hiện trong hệ thống 10 có thể gồm có, nhưng không bị giới hạn ở, thiết bị điện tử hoặc thiết bị 50, sự kết hợp của thiết bị kỹ thuật số hỗ trợ cá nhân (PDA) và điện thoại di động 14, PDA 16, được tích hợp thiết bị nhắn tin (IMD) 18, máy tính bàn 20, máy tính xách tay 22. Thiết bị 50 có thể là cố định hoặc di động khi được thực hiện bởi cá nhân đang chuyển động. Thiết bị 50 cũng có thể được đặt ở chế độ thay đổi gồm có, nhưng không giới hạn ở, xe ô tô, xe tải, taxi, xe buýt, tàu hỏa, tàu thuyền, máy bay, xe đạp, mô tô hoặc bất kỳ loại vận tải thích hợp cùng loại nào.

Các phương án còn có thể được sử dụng trong thiết bị giải mã thiết bị truyền hình (set-top box); ví dụ bộ thu TV kỹ thuật số, mà có thể/không thể có màn hình hoặc

các năng lực không dây, trong các máy tính xách tay (laptop) hoặc các máy tính cá nhân (PC), vốn có phần cứng hoặc phần mềm hoặc sự kết hợp của các ứng dụng bộ mã hóa /bộ giải mã, trong các hệ điều hành khác nhau, và trong các bộ chip, các bộ xử lý, các DSP và/hoặc các hệ thống nhúng cung cấp việc mã hóa dựa trên phần cứng/phần mềm.

Một số hoặc thêm thiết bị có thể gửi và nhận các cuộc gọi và các tin nhắn và kết nối với các nhà cung cấp dịch vụ thông qua kết nối không dây 25 với trạm cơ sở 24. Trạm cơ sở 24 có thể được kết nối với máy chủ của mạng 26 mà cho phép kết nối giữa mạng điện thoại di động 11 và internet 28. Hệ thống có thể bao gồm các thiết bị truyền thông bổ sung và các loại thiết bị truyền thông khác nhau.

Các thiết bị truyền thông có thể kết nối nhờ sử dụng các công nghệ truyền dẫn khác nhau gồm có, nhưng không giới hạn ở, công nghệ đa truy nhập phân chia theo mã (CDMA), các hệ thống truyền thông di động toàn cầu (GSM), hệ thống viễn thông di động toàn cầu (UMTS), công nghệ đa truy nhập phân chia theo thời gian (TDMA), công nghệ đa truy nhập phân chia theo tần số (FDMA), giao thức TCP-IP, dịch vụ tin nhắn ngắn (SMS), dịch vụ tin nhắn đa phương tiện (MMS), email, dịch vụ tin nhắn nhanh (IMS), Bluetooth, IEEE 802.11 và công nghệ truyền thông không dây cùng loại bất kỳ. Thiết bị truyền thông liên quan đến thực hiện các phương án khác nhau của sáng chế có thể kết nối nhờ sử dụng các phương tiện khác nhau gồm có, nhưng không giới hạn ở, các kết nối radiô, hồng ngoại, laze, cáp, và kết nối thích hợp bất kỳ.

Giao thức truyền tải thời gian thực (RTP) được sử dụng rộng rãi để truyền tải thời gian thực của phương tiện theo thời gian như audio và video. RTP có thể vận hành trên đỉnh của Giao thức UDP, vốn lần lượt có thể vận hành trên đỉnh của Giao thức Internet (IP). RTP được xác định trong Lực lượng Quản lý Kỹ thuật Internet (IETF) Đề nghị duyệt thảo và bình luận (RFC) 3550, có trên www.ietf.org/rfc/rfc3550.txt. Trong Giao thức RTP, dữ liệu đa phương tiện được gói thành các gói RTP. Thông thường, mỗi loại phương tiện hoặc định dạng mã hóa phương tiện có định dạng Giao thức RTP chuyên dụng.

Giao thức RTP là sự kết hợp giữa nhóm thành viên liên lạc với RTP. Đây là

kênh liên lạc của nhóm có thể có khả năng mang một số lượng các dòng RTP. Dòng RTP là dòng các gói RTP bao gồm dữ liệu đa phương tiện. Dòng RTP được xác định bởi SSRC thuộc Giao thức RTP cụ thể. SSRC hoặc là nguồn đồng bộ hóa hoặc bộ định danh nguồn đồng bộ hóa là trường SSRC 32-bit trong phần đầu gói RTP. Nguồn đồng bộ hóa đặc trưng ở chỗ tất cả các gói từ phần hình thành nguồn đồng bộ hóa của cùng không gian tính giờ và số lượng dãy, sao cho bộ phận tiếp nhận có thể nhóm các gói bởi nguồn đồng bộ hóa để phát lại. Các ví dụ về các nguồn đồng bộ hóa bao gồm bộ phận chuyển dòng các gói xuất phát từ tín hiệu nguồn như micro hoặc camera, hoặc bộ trộn RTP. Mỗi dòng RTP được xác định bởi SSRC vốn là duy nhất trong Giao thức RTP.

Dòng truyền tải MPEG-2 (TS), được xác định trong ISO/IEC 13818-1 hoặc tương đương trong Khuyến nghị ITU-T H.222.0, là định dạng để mang audio, video, và các phương tiện khác cũng như siêu dữ liệu chương trình hoặc các siêu dữ liệu khác, trong dòng ghép kênh. Bộ định danh gói (PID) được sử dụng để nhận dạng luồng cơ bản (cũng được gọi là luồng cơ bản được gói hóa) trong TS.

Các tiêu chuẩn định dạng tệp đa phương tiện khả dụng bao gồm định dạng tệp đa phương tiện cơ bản ISO (ISO/IEC 14496-12, mà có thể được viết tắt là ISOBMFF), định dạng tệp MPEG-4 (ISO/IEC 14496-14, còn được biết là định dạng MP4), định dạng tệp dùng cho video cấu trúc khối NAL (ISO/IEC 14496-15) và định dạng tệp 3GPP (3GPP TS 26.244, còn được biết là định dạng 3GP). ISOBMFF là cơ sở cho nguồn gốc của tất cả các định dạng tệp nêu trên (ngoại trừ chính ISOBMFF).

Một số các khái niệm, các cấu trúc, và các thông số kỹ thuật của ISOBMFF được mô tả dưới đây làm ví dụ về định dạng tệp của vật chứa, dựa vào đó các phương án có thể được sử dụng. Các khía cạnh của sáng chế không bị giới hạn ở ISOBMFF, mà đúng hơn là việc mô tả được nêu ra làm cơ sở mà dựa trên đó nó sáng chế có thể được thực hiện một phần hoặc toàn bộ.

Một khối cơ bản trong ISOBMFF được gọi là một hộp. Mỗi hộp có thể có phần đầu và tải. Phần đầu của hộp biểu thị loại hộp và kích thước của hộp về mặt byte. Hộp có thể kèm theo các hộp khác, và định dạng tệp ISO xác định loại hộp nào được phép có trong loại hộp nhất định. Ngoài ra, sự có mặt của một số các hộp có thể là bắt

buộc trong mỗi tệp, trong khi sự có mặt của các hộp khác có thể là tùy ý. Ngoài ra, đối với một số loại hộp, có thể cho phép có nhiều hơn một hộp có mặt trong tệp. Do đó, ISOBMFF có thể được coi là xác định cấu trúc phân cấp của các hộp. Mỗi hộp của tệp đa phương tiện cơ bản ISO có thể được xác định bởi mã bốn ký tự (4CC, fourCC). Mã bốn ký tự có thể được đại diện thay đổi được bởi số nguyên không dấu 32-bit (bằng cách giả định một sự chuyển đổi nhất định các ký tự thành các giá trị 8-bit, có tính endian của bit nhất định, và tính endian của byte nhất định). Phần đầu có thể cung cấp thông tin về loại và kích thước của hộp.

Theo ISOBMFF, một tệp có thể bao gồm dữ liệu đa phương tiện và siêu dữ liệu mà có thể được kèm theo trong các hộp riêng biệt. Trong phương án ví dụ, dữ liệu đa phương tiện có thể được cung cấp trong hộp dữ liệu đa phương tiện (siêu dữ liệu-metadata) và hộp phim (cũng được gọi là MovieBox) có thể được sử dụng kèm theo siêu dữ liệu. Trong một số trường hợp, đối với tệp cần phải hoạt động được, cả siêu dữ liệu và các hộp phim phải có mặt. Hộp phim (moov) có thể bao gồm một hay nhiều đường ghi (trak), và mỗi đường ghi có thể cư trú tại một hộp đường ghi tương ứng. Mỗi đường ghi được kèm theo bộ xử lý, được xác định bởi mã bốn ký tự, xác định loại đường ghi. Các đường ghi chuỗi video, audio, và hình ảnh có thể được gọi chung là các đường ghi phương tiện, và chúng chứa dòng phương tiện cơ bản. Các loại đường ghi khác bao gồm đường ghi gợi ý và đường ghi siêu dữ liệu hẹn giờ. Các đường ghi bao gồm các mẫu, như các khung hình audio hoặc video. Đường ghi phương tiện đề cập đến các mẫu (còn có thể được gọi là các mẫu phương tiện) được định dạng theo định dạng nén dữ liệu đa phương tiện (và gói nó vào ISOBMFF). Đường ghi gợi ý đề cập đến các mẫu gợi ý, chứa các lệnh để xây dựng các gói để biến đổi đối với giao thức truyền thông được chỉ ra. Các lệnh có thể bao gồm hướng dẫn cấu trúc phân đầu của gói và có thể bao gồm cấu trúc gói tải. Trong cấu trúc gói tải, dữ liệu cư trú trong đường ghi hoặc các mục khác có thể được tham chiếu. Như vậy, ví dụ, dữ liệu cư trú trong các đường ghi hoặc các mục khác có thể được biểu thị bởi sự tham chiếu giống như là một mẫu dữ liệu của nó trong đường ghi hoặc mục cụ thể được hướng dẫn để được sao chép vào gói trong quy trình xây dựng gói. Đường ghi siêu dữ liệu hẹn giờ có thể là các mẫu mô tả các mẫu phương tiện và/hoặc gợi ý được đề cập. Để trình bày

một loại phương tiện, một đường ghi của phương tiện có thể được chọn.

Hộp 'trak' chứa hộp Bảng Mẫu. Hộp Bảng Mẫu bao gồm ví dụ toàn bộ việc lập bảng mục lục thời gian và dữ liệu của các mẫu phương tiện trong đường ghi. Hộp Bảng Mẫu được yêu cầu phải chứa Hộp Mô tả Mẫu. Hộp Mô tả Mẫu bao gồm trường đếm mục nhập, xác định số lượng các mục nhập mẫu có trong hộp. Hộp Mô tả Mẫu được yêu cầu phải chứa ít nhất một mục nhập mẫu. Định dạng mục nhập mẫu phụ thuộc vào loại bộ xử lý đối với đường ghi. Các mục nhập mẫu cung cấp thông tin chi tiết về loại mã hóa được sử dụng và thông tin khởi tạo bất kỳ cần thiết để mã hóa.

Tính chất của đoạn phim có thể cho phép chia nhỏ siêu dữ liệu không thì có thể cư trú trong hộp phim thành nhiều mảnh. Mỗi mảnh có thể tương ứng với khoảng thời gian nhất định của đường ghi. Nói cách khác, tính chất của đoạn phim có thể cho phép xen kẽ siêu dữ liệu tệp và dữ liệu đa phương tiện. Do đó, kích thước của hộp phim có thể bị giới hạn và các trường hợp sử dụng trình bày ở trên được thực hiện.

Trong một số ví dụ, các mẫu phương tiện dùng cho các đoạn phim có thể cư trú trong hộp mdat. Tuy nhiên, hộp moof có thể được tạo ra cho siêu dữ liệu của các đoạn phim. Hộp moof có thể bao gồm thông tin dùng cho thời lượng phát lại nhất định mà có thể có trước trong hộp moov. Hộp moov có thể tự nó vẫn hiển thị phim hợp lệ, nhưng ngoài ra, nó có thể bao gồm hộp mvex biểu thị các đoạn phim sẽ nối tiếp nhau trong cùng một tệp. Các đoạn phim có thể mở rộng việc hiển thị mà cùng lúc được kèm theo hộp phim.

Trong đoạn phim có thể có tập hợp các đoạn ghi, bao gồm bất cứ đâu từ không đến nhiều đoạn ghi cho mỗi đường ghi. Các đoạn ghi có thể lần lượt bao gồm bất cứ đâu từ không đến nhiều tuyến đường ghi, mỗi tài liệu trong số đó là sự hoạt động liên tục của các mẫu cho đường ghi đó (và do đó là giống với các khối). Trong các cấu trúc này, nhiều trường là tùy chọn và có thể là mặc định. Siêu dữ liệu mà có thể chứa trong hộp phim có thể bị giới hạn ở tập hợp con của siêu dữ liệu mà có thể chứa trong hộp phim và có thể được mã hóa khác nhau trong một số trường hợp. Các chi tiết liên quan đến các hộp mà có thể chứa trong hộp phim có thể tìm thấy từ thông số kỹ thuật ISOBMFF. Đoạn phim khép kín có thể được xác định để bao gồm hộp moof và hộp mdat vốn là liên tục trong trình tự tệp và ở đó hộp mdat chứa các mẫu

của đoạn phim (mà hộp moof cung cấp siêu dữ liệu cho nó) và không chứa các mẫu của bất kỳ đoạn phim nào khác (tức là hộp moof khác bất kỳ).

Việc nhóm mẫu trong ISOBMFF và các phát sinh của nó, như định dạng tệp cho video cấu trúc bộ NAL (ISO/IEC 14496-15), có thể được xác định như là việc chuyển mỗi mẫu trong đường ghi trở thành thành phần của một nhóm mẫu, dựa trên tiêu chí tạo nhóm. Nhóm mẫu trong khi tạo nhóm mẫu không bị giới hạn ở các mẫu gần kề và có thể chứa các mẫu không liên kề. Như có thể có nhiều hơn một nhóm mẫu đối với các mẫu trong đường ghi, mỗi nhóm mẫu có thể có loại trường để thể hiện loại nhóm. Các nhóm mẫu có thể được biểu diễn bởi hai cấu trúc dữ liệu được liên kết: (1) hộp SampleToGroup (hộp sbgp) thể hiện việc chuyển giao các mẫu cho các nhóm mẫu; và (2) hộp SampleGroupDescription (hộp sgpd) chứa mục nhập nhóm mẫu cho mỗi nhóm mẫu mô tả các đặc tính của nhóm. Ở đây có thể có nhiều trường hợp các hộp SampleToGroup và SampleGroupDescription dựa trên các tiêu chí tạo nhóm khác nhau. Điều này có thể được phân biệt bởi một loại trường được sử dụng để biểu thị loại nhóm. Các hộp 'sbgp' và 'sgpd' có thể được kết nối nhờ sử dụng giá trị `grouping_type` và, trong một số phiên bản của các hộp, cũng như giá trị của `grouping_type_parameter`. Hộp 'sbgp' biểu thị chỉ mục của mục nhập mô tả nhóm mẫu mà mẫu cụ thể thuộc về nó.

Định dạng tệp Matroska có thể (nhưng không giới hạn ở) việc lưu đường ghi bất kỳ trong số video, audio, hình ảnh, hoặc phụ đề trong một tệp. các phần mở rộng tệp Matroska bao gồm .mkv cho video (với các phụ đề và audio), .mk3d cho video lập thể, .mka chỉ dành cho các tệp audio, và .mks chỉ dành cho các phụ đề. Matroska có thể được sử dụng như là định dạng cơ sở cho các định dạng tệp xuất phát, như WebM.

Matroska sử dụng Ngôn ngữ Meta Nhị phân Mở rộng (EBML) làm nền tảng. EBML xác định định dạng nhị phân và bát tử (byte) sắp hàng lấy cảm hứng bởi nguyên tắc XML. Chính EBML là sự mô tả tổng quát hóa của kỹ thuật đánh dấu nhị phân. File Matroska bao gồm các phần tử mà bù đắp cho "tài liệu" EBML. Các phần tử kết hợp phần tử ID, bộ mô tả dùng cho kích thước của phần tử, và chính dữ liệu nhị phân. Các phần tử có thể lồng nhau.

Phần tử Phân đoạn của Matroska là vật chứa cho các phần tử cấp cao nhất

khác (cấp 1). Một tệp Matroska có thể bao gồm (nhưng không giới hạn việc bao gồm) một phân đoạn. Dữ liệu đa phương tiện trong các tệp Matroska được tổ chức trong các Cụm (hoặc các phần tử của Cụm), mỗi Cụm thường chứa vài giây dữ liệu đa phương tiện. Một Cụm bao gồm các phần tử BlockGroup, phần tử này lại bao gồm các phần tử Khối. Phần tử Tín hiệu bao gồm siêu dữ liệu mà có thể hỗ trợ khi truy cập ngẫu nhiên hoặc tìm kiếm và có thể bao gồm các con trỏ tệp hoặc các nhãn thời gian tương ứng để tìm kiếm các điểm.

Bộ mã hóa-giải mã video bao gồm bộ mã hóa mà biến đổi video đầu vào thành sự trình diễn được nén thích hợp để lưu/truyền dẫn và bộ giải mã mà có thể giải nén sự biểu diễn video được nén trở lại dạng xem được. Bộ mã hóa video và/hoặc bộ giải mã video cũng có thể tách rời khỏi nhau, tức là không cần tạo ra bộ mã hóa-bộ giải mã. Thông thường, bộ mã hóa loại bỏ một số thông tin trong chuỗi video gốc để biểu diễn video ở dạng nhỏ gọn hơn (nghĩa là, ở tốc độ bit thấp hơn). Bộ mã hóa video có thể được sử dụng để mã hóa chuỗi hình ảnh, như được xác định sau đó, và bộ giải mã video có thể được sử dụng để giải mã chuỗi hình ảnh được mã hóa. Bộ mã hóa video hoặc bộ phận mã hóa trong của bộ mã hóa video hoặc bộ mã hóa hình ảnh có thể được sử dụng để mã hóa hình ảnh, và bộ giải mã video hoặc bộ phận giải mã nội bộ của bộ giải mã video hoặc bộ giải mã hình ảnh có thể được sử dụng để giải mã hình ảnh được mã hóa.

Các bộ mã hóa video ghép diễn hình, ví dụ nhiều công cụ mã hóa ITU-T H.263 và H.264, mã hóa thông tin video theo hai giai đoạn. Thứ nhất, các giá trị điểm ảnh trong vùng hình ảnh nhất định (hoặc “khối”) được dự đoán, ví dụ bởi phương tiện bù chuyển động (tìm kiếm và biểu thị vùng trong một khung hình video được mã hóa từ trước mà tương ứng chặt chẽ với khối đang được mã hóa) hoặc bởi phương tiện không gian (nhờ sử dụng các giá trị điểm ảnh xung quanh khối cần được mã hóa theo cách xác định). Thứ hai, lỗi dự đoán, tức là sự chênh lệch giữa khối điểm ảnh được dự đoán và khối điểm ảnh ban đầu, được mã hóa. Điều này thường được thực hiện bằng cách biến đổi sự chênh lệch về các giá trị điểm ảnh nhờ sử dụng biến đổi được xác định (ví dụ, Biến đổi Cosin Gián đoạn (DCT) hoặc biến thể của nó), lượng tử hóa các hệ số và mã hóa entropy các hệ số đã lượng tử hóa. Bằng cách thay đổi độ tin cậy của

quá trình lượng tử hóa, bộ mã hóa có thể điều khiển sự cân bằng giữa độ chính xác của sự biểu diễn điểm ảnh (chất lượng hình ảnh) và kích thước của sự biểu diễn video được mã hóa thu được (kích thước tệp hoặc tốc độ bit biến đổi).

Dự đoán liên khung, vốn còn có thể được gọi là dự đoán tạm thời, bù chuyển động, hoặc dự đoán bù chuyển động, làm giảm phần dư tạm thời. Trong dự đoán liên khung, nguồn dự đoán là các hình ảnh được mã hóa trước. Dự đoán liên khung vận dụng thực tế là các điểm ảnh liên kề trong cùng hình ảnh có thể cần phải có tương quan. Dự đoán liên khung có thể được thực hiện trong miền không gian hoặc biến đổi, tức là, hoặc các giá trị của mẫu hoặc các hệ số biến đổi có thể được dự đoán. Dự đoán liên khung thường được khai thác khi mã hóa trong, là nơi không áp dụng dự đoán liên khung.

Một kết quả của trình tự mã hóa là một tập hợp các thông số mã hóa, như các vector chuyển động và các hệ số biến đổi đã lượng tử hóa. Nhiều thông số có thể được mã hóa entropy hiệu quả hơn nếu chúng được dự đoán trước tiên từ các thông số lân cận về không gian hoặc tạm thời. Ví dụ, một vector chuyển động có thể được dự đoán từ các vector chuyển động liên kề về mặt không gian và chỉ ra sự chênh lệch tương đối với bộ dự đoán vector chuyển động có thể được mã hóa. Việc dự đoán các thông số mã hóa và dự đoán liên khung có thể được gọi chung là dự đoán trong-ảnh.

Fig.4 thể hiện sơ đồ khối của bộ mã hóa video thích hợp để ứng dụng các phương án của sáng chế. Fig.4 trình bày bộ mã hóa cho hai lớp, nhưng có thể tốt hơn nếu bộ mã hóa trình bày có thể được đơn giản hóa một cách tương tự để mã hóa chỉ một lớp hoặc được mở rộng để mã hóa nhiều hơn hai lớp. Fig.4 minh họa một phương án của bộ mã hóa video bao gồm bộ mã hóa thứ nhất 500 dùng cho lớp cơ sở và bộ mã hóa thứ hai 502 dùng cho lớp tăng cường. Mỗi bộ mã hóa thứ nhất 500 và bộ mã hóa thứ hai 502 có thể bao gồm các phần tử cùng loại dùng để mã hóa các hình ảnh đến. Các bộ mã hóa 500, 502 có thể bao gồm bộ phận dự đoán điểm ảnh 302, 402, bộ mã hóa lỗi dự đoán 303, 403 và bộ giải mã lỗi dự đoán 304, 404. Fig.4 cũng thể hiện một phương án của bộ phận dự đoán điểm ảnh 302, 402 bao gồm bộ phận dự đoán liên khung 306, 406, bộ phận dự đoán trong 308, 408, bộ chọn chế độ 310, 410, bộ lọc 316, 416, và bộ nhớ khung hình tham chiếu 318, 418. Bộ phận dự đoán điểm ảnh 302 của

bộ mã hóa thứ nhất 500 nhận hình ảnh lớp cơ sở 300 của dòng video cần được mã hóa tại cả bộ phận dự đoán liên khung 306 (vốn xác định sự khác biệt giữa hình ảnh và khung hình tham chiếu bù chuyển động 318) và bộ phận dự đoán trong 308 (vốn xác định dự đoán cho khối hình ảnh chỉ dựa trên các phần đã được xử lý của khung hoặc hình ảnh hiện tại). Đầu ra của cả bộ phận dự đoán liên khung và bộ phận dự đoán trong được chuyển sang bộ chọn chế độ 310. Bộ phận dự đoán trong 308 có thể có nhiều hơn một chế độ dự đoán trong. Do đó, mỗi chế độ có thể thực hiện dự đoán trong và tạo ra tín hiệu được dự đoán cho bộ chọn chế độ 310. Bộ chọn chế độ 310 cũng nhận bản sao của hình ảnh lớp cơ sở 300. Một cách tương ứng, bộ phận dự đoán điểm ảnh 402 của bộ mã hóa thứ hai 502 nhận 400 hình ảnh lớp tăng cường của dòng video cần được mã hóa tại cả bộ phận dự đoán liên khung 406 (vốn xác định sự khác biệt giữa hình ảnh và khung hình tham chiếu bù chuyển động 418) và bộ phận dự đoán trong 408 (vốn xác định dự đoán cho khối hình ảnh chỉ dựa trên các phần đã được xử lý của khung hình hoặc hình ảnh hiện tại). Đầu ra của cả bộ phận dự đoán liên khung và bộ phận dự đoán trong được chuyển đến bộ chọn chế độ 410. Bộ phận dự đoán trong 408 có thể có nhiều hơn một chế độ dự đoán trong. Do đó, mỗi chế độ có thể thực hiện dự đoán trong và tạo ra tín hiệu được dự đoán đến bộ chọn chế độ 410. Bộ chọn chế độ 410 cũng nhận bản sao của hình ảnh lớp tăng cường 400.

Tùy vào việc chế độ mã hóa được chọn để mã hóa khối hiện hành, đầu ra của bộ phận dự đoán liên khung 306, 406 hoặc đầu ra của một trong số các chế độ dự đoán trong tùy ý hoặc đầu ra của bộ mã hóa bề mặt trong bộ chọn chế độ được đưa sang đầu ra của bộ chọn chế độ 310, 410. Đầu ra của bộ chọn chế độ được đưa sang thiết bị cộng thứ nhất 321, 421. Thiết bị cộng thứ nhất có thể trừ đi đầu ra của bộ phận dự đoán điểm ảnh 302, 402 từ hình ảnh lớp cơ sở 300/hình ảnh lớp tăng cường 400 để tạo ra tín hiệu lỗi dự đoán thứ nhất 320, 420 vốn là đầu vào đến bộ mã hóa lỗi dự đoán 303, 403.

Bộ phận dự đoán điểm ảnh 302, 402 nhận thêm từ bộ phận phục hồi sơ bộ 339, 439 sự kết hợp của sự biểu diễn dự đoán của khối hình ảnh 312, 412 và đầu ra 338, 438 của bộ giả mã lỗi dự đoán 304, 404. Hình ảnh được phục hồi sơ bộ 314, 414 có thể được đưa đến bộ phận dự đoán trong 308, 408 và đến bộ lọc 316, 416. Bộ lọc

316, 416 nhận sự biểu diễn sơ bộ có thể lọc sự biểu diễn sơ bộ và tạo ra hình ảnh được phục hồi lần cuối 340, 440 vốn có thể được lưu trong bộ nhớ khung hình tham chiếu 318, 418. Bộ nhớ khung hình tham chiếu 318 có thể được kết nối với bộ phận dự đoán liên khung 306 cần được sử dụng làm hình ảnh tham chiếu, mà hình ảnh lớp cơ sở tương lai 300 được so sánh với nó trong các hoạt động dự đoán liên khung. Tùy thuộc vào lớp cơ sở đang được chọn và được chỉ định làm nguồn dự đoán mẫu liên lớp và/hoặc thông tin dự đoán chuyển động liên lớp của lớp tăng cường theo một số phương án, bộ nhớ khung hình tham chiếu 318 có thể cũng được kết nối với bộ phận dự đoán liên khung 406 để được sử dụng làm hình ảnh tham chiếu mà các hình ảnh lớp tăng cường tương lai 400 được so sánh với nó trong các hoạt động dự đoán liên khung. Hơn nữa, bộ nhớ khung hình tham chiếu 418 có thể được kết nối với bộ phận dự đoán liên khung 406 để được sử dụng làm hình ảnh tham chiếu mà hình ảnh lớp tăng cường tương lai 400 được so sánh với nó trong các hoạt động dự đoán liên khung.

Các thông số lọc từ bộ lọc 316 của bộ mã hóa thứ nhất 500 có thể được đề xuất cho bộ mã hóa thứ hai 502 tùy thuộc vào lớp cơ sở đang được chọn và được chỉ định để làm nguồn để dự đoán các thông số lọc của lớp tăng cường theo một số phương án.

Bộ mã hóa lỗi dự đoán 303, 403 bao gồm bộ biến đổi 342, 442 và bộ lượng tử hóa 344, 444. Bộ biến đổi 342, 442 biến đổi tín hiệu lỗi dự đoán thứ nhất 320, 420 thành miền biến đổi. Bộ biến đổi là, ví dụ, phép biến đổi DCT. Bộ lượng tử hóa 344, 444 lượng tử hóa tín hiệu miền biến đổi, ví dụ các hệ số DCT, để tạo ra các hệ số đã lượng tử hóa.

Bộ giải mã lỗi dự đoán 304, 404 nhận tín hiệu đầu ra từ bộ mã hóa lỗi dự đoán 303, 403 và thực hiện các quá trình ngược với bộ mã hóa lỗi dự đoán 303, 403 để tạo ra tín hiệu lỗi dự đoán được giải mã 338, 438 mà khi được kết hợp với hiển thị dự đoán của khối hình ảnh 312, 412 tại thiết bị cộng thứ hai 339, 439, tạo ra hình ảnh được phục hồi sơ bộ 314, 414. Bộ giải mã lỗi dự đoán có thể được coi là bao gồm bộ khử lượng tử 361, 461, mà khử lượng tử các giá trị hệ số đã lượng tử hóa, ví dụ các hệ số DCT, để phục hồi tín hiệu biến đổi và bộ biến đổi tín hiệu nghịch đảo 363, 463, vốn thực hiện biến đổi nghịch đảo thành tín hiệu biến đổi được phục hồi trong đó tín hiệu

đầu ra của bộ biến đổi nghịch đảo 363, 463 chứa (các) khối được phục hồi. Bộ giải mã lỗi dự đoán có thể cũng bao gồm bộ lọc khối mà có thể lọc (các) khối được phục hồi theo thông tin được giải mã thêm và các thông số của bộ lọc.

Bộ mã hóa entropy 330, 430 nhận tín hiệu đầu ra của bộ mã hóa lỗi dự đoán 303, 403 và có thể thực hiện giải mã entropy/giải mã độ dài khác nhau thích hợp đối với tín hiệu để xác định việc dò ra lỗi và khả năng sửa lỗi. Các tín hiệu đầu ra của các bộ mã hóa entropy 330, 430 có thể được chèn vào dòng bit ví dụ bởi bộ ghép kênh 508.

Tiêu chuẩn H.264/AVC đã được phát triển bởi Joint Video Team (JVT) của Nhóm chuyên gia mã hóa video (VCEG) của Ngành Tiêu chuẩn Viễn thông của Liên minh Viễn thông quốc tế (ITU-T) và Nhóm chuyên gia hình ảnh động (MPEG) của Tổ chức tiêu chuẩn hóa quốc tế (ISO)/Ủy ban kỹ thuật điện quốc tế (IEC). Tiêu chuẩn H.264/AVC được công bố bởi cả hai tổ chức tiêu chuẩn hóa gốc, và nó được gọi là ITU-T khuyến nghị sử dụng H.264 và tiêu chuẩn quốc tế ISO/IEC 14496-10, còn được biết là Mã hóa video tiên tiến (AVC) MPEG-4 Phần 10. Có nhiều phiên bản của tiêu chuẩn H.264/AVC, kết hợp các phần mở rộng hoặc các tính năng mới vào thông số kỹ thuật. Các phần mở rộng này bao gồm Mã hóa video mở rộng được (SVC) và Mã hóa video hiển thị nhiều khung hình (MVC).

Phiên bản 1 của chuẩn Mã hóa video hiệu quả cao (H.265/HEVC cũng được gọi là HEVC) đã được phát triển bởi Joint Collaborative Team – Video Coding (JCT-VC) của VCEG và MPEG. Tiêu chuẩn này đã được công bố bởi cả hai tổ chức tiêu chuẩn hóa gốc, và nó được gọi là ITU-T khuyến nghị sử dụng H.265 và tiêu chuẩn quốc tế ISO/IEC 23008-2, còn được biết là Mã hóa video hiệu quả cao (HEVC) MPEG-H Phần 2. Phiên bản 2 của H.265/HEVC chứa các phần mở rộng, hiển thị nhiều khung hình, và phạm vi tín hiệu mở rộng được, mà có thể được viết tắt là SHVC, MV-HEVC, và REXT, tương ứng. Phiên bản 2 của H.265/HEVC đã được công bố là ITU-T khuyến nghị sử dụng H.265 (10/2014) và như Phiên bản 2 của ISO/IEC 23008-2. Hiện nay đang tiến hành các dự án tiêu chuẩn hóa để phát triển thêm các phần mở rộng cho H.265/HEVC, bao gồm các mở rộng việc mã hóa nội dung ba chiều và kịch bản, mà có thể được viết tắt là 3D-HEVC và SCC, tương ứng.

SHVC, MV-HEVC, và 3D-HEVC sử dụng thông số kỹ thuật cơ bản chung, được quy định trong Phụ lục F của Phiên bản 2 của Tiêu chuẩn HEVC. Thông số kỹ thuật cơ bản chung này bao gồm, ví dụ cú pháp và ngữ nghĩa cao cấp, ví dụ xác định một số đặc tính của lớp dòng bit, như các phần phụ thuộc liên lớp, cũng như các quy trình giải mã, như xây dựng danh mục hình ảnh bao gồm các hình ảnh tham chiếu liên lớp và hàm đếm thứ tự hình ảnh cho dòng bit đa lớp. Phụ lục F có thể cũng được sử dụng trong các phần mở rộng đa lớp tiếp theo tiềm năng của HEVC. Cần phải hiểu rằng mặc dù bộ mã hóa video, bộ giải mã video, các phương pháp mã hóa, các phương pháp giải mã, các cấu trúc dòng bit, và/hoặc các phương án có thể được mô tả sau đây có tham chiếu đến các phần mở rộng cụ thể, như SHVC và/hoặc MV-HEVC, chúng có thể áp dụng chung các phần mở rộng nhiều lớp bất kỳ của HEVC, và nói chung ngay cả nhiều hơn đối với hệ thống mã hóa video nhiều lớp bất kỳ.

Một số định nghĩa cơ bản, dòng bit và các cấu trúc mã hóa, và các khái niệm của H.264/AVC và HEVC được mô tả trong phần này làm ví dụ của bộ mã hóa video, bộ giải mã, phương pháp mã hóa, phương pháp giải mã, và cấu trúc dòng bit, trong đó các phương án có thể được sử dụng. Một số định nghĩa cơ bản, dòng bit và các cấu trúc mã hóa, và các khái niệm của H.264/AVC giống như trong HEVC - do đó, chúng được mô tả dưới đây cùng nhau. Các khía cạnh của sáng chế không bị giới hạn ở H.264/AVC hoặc HEVC, mà đúng hơn là phần mô tả đưa ra một căn cứ thực hiện, dựa trên đó sáng chế có thể được thực hiện một phần hoặc toàn bộ.

Tương tự như nhiều tiêu chuẩn mã hóa video trước đây, cú pháp và ngữ nghĩa dòng bit cũng như quy trình giải mã các dòng bit không-lỗi được quy định trong H.264/AVC và HEVC. Quy trình mã hóa không được quy định, nhưng các bộ mã hóa phải tạo ra các dòng bit phù hợp. Sự phù hợp của dòng bit và bộ giải mã có thể được xác nhận bằng Bộ giải mã Tham chiếu giả định (Hypothetical Reference Decoder - HRD). Các tiêu chuẩn này chứa các công cụ mã hóa mà giúp sao chép với các lỗi biến đổi và các hao tổn, nhưng việc sử dụng các công cụ này để mã hóa là tùy ý và không có quy trình giải mã nào được quy định cho các dòng bit lỗi.

Trong phần mô tả các tiêu chuẩn hiện hành cũng như trong phần mô tả các phương án ví dụ, phần tử cú pháp có thể được xác định như là phần tử của dữ liệu

được có trong dòng bit. Cấu trúc cú pháp có thể được xác định là không hoặc nhiều phần tử cú pháp cùng có mặt trong dòng bit theo trình tự quy định. Trong phần mô tả các tiêu chuẩn hiện hành cũng như trong phần mô tả các phương án ví dụ, đoạn “bởi phương tiện bên ngoài” hoặc “nhờ phương tiện bên ngoài” có thể được sử dụng. Ví dụ, thực thể, như cấu trúc cú pháp hoặc giá trị của biến được sử dụng trong quy trình giải mã, có thể được cung cấp “bởi phương tiện bên ngoài” cho quy trình giải mã. Đoạn “bởi phương tiện bên ngoài” có thể biểu thị là thực thể không được chứa trong dòng bit được tạo ra bởi bộ mã hóa, mà đúng hơn là từ dòng bit được chuyển đến từ bên ngoài. Ví dụ, nhờ sử dụng giao thức điều khiển. Có thể là phương tiện thay thế được hoặc bổ sung mà thực thể không được tạo ra bởi bộ mã hóa, nhưng có thể được tạo ra, ví dụ trong logic lớp hoặc điều khiển giải mã hoặc tương tự nhờ sử dụng bộ giải mã. Bộ giải mã có thể có giao diện để nhập phương tiện bên ngoài vào, như các giá trị biến.

Bộ phận cơ bản cho tín hiệu đầu vào bộ mã hóa H.264/AVC hoặc HEVC và tín hiệu đầu ra của bộ giải mã H.264/AVC hoặc HEVC, tương ứng, là hình ảnh. Hình ảnh được cho là tín hiệu đầu vào bộ mã hóa cũng có thể được gọi là hình ảnh nguồn, và hình ảnh được giải mã bởi bộ giải mã có thể được gọi là hình ảnh được giải mã.

Nguồn và các hình ảnh được giải mã mỗi thứ bao gồm một hay nhiều mảng mẫu, như một bộ mảng mẫu sau:

- Chỉ Luma (Y) (đơn sắc).
- Luma và hai sắc độ (YCbCr hoặc YCgCo).
- Xanh lục (Green), Xanh da trời (Blue) và Đỏ (Red) (GBR), còn được biết là RGB).
- Các mảng hiển thị đơn sắc không xác định khác hoặc lấy mẫu màu bộ ba (Ví dụ, YZX, còn được biết là XYZ).

Tiếp theo, các mảng này có thể được gọi là luma (hoặc L hoặc Y) và sắc độ, nơi hai mảng sắc độ có thể được gọi là Cb và Cr; bất kể phương pháp biểu diễn màu thực tế khi sử dụng. Phương pháp biểu diễn màu thực tế khi sử dụng có thể được chỉ định ví dụ trong dòng bit được mã hóa ví dụ nhờ sử dụng cú pháp thông tin khả năng sử dụng

video (Video Usability Information, VUI) của H.264/AVC và/hoặc HEVC. Thành phần có thể được xác định là mảng hoặc mẫu đơn từ một trong ba mảng mẫu (luma và hai sắc độ) hoặc mảng hoặc mẫu đơn của mảng mà hợp thành hình ảnh ở định dạng đơn sắc.

Trong H.264/AVC và HEVC, hình ảnh có thể là khung hình hoặc trường. Khung hình bao gồm ma trận của các mẫu luma và có thể các mẫu sắc độ tương ứng. Trường là tập hợp của các hàng mẫu xen kẽ khung hình và có thể được sử dụng làm đầu vào của bộ mã hóa, nếu tín hiệu nguồn được trộn lẫn. Các mảng mẫu sắc độ có thể vắng mặt (và do đó có thể lấy mẫu đơn sắc khi sử dụng) hoặc mảng mẫu sắc độ có thể được lấy mẫu khi được so sánh với mảng mẫu luma. Các định dạng sắc độ có thể được tóm tắt như sau:

- Khi lấy mẫu đơn sắc chỉ có một mảng mẫu, vốn về danh nghĩa có thể được coi là mảng luma.
- Khi lấy mẫu 4:2:0, mỗi trong số hai mảng sắc độ có một nửa chiều cao và một nửa chiều rộng của mảng luma.
- Khi lấy mẫu 4:2:2, mỗi trong số hai mảng sắc độ có cùng chiều cao và một nửa chiều rộng của mảng luma.
- Khi lấy mẫu 4:4:4, nếu không sử dụng các các mặt màu riêng biệt, mỗi trong số hai mảng sắc độ có cùng chiều cao và chiều rộng là mảng luma.

Trong H.264/AVC và HEVC, có thể mã hóa mảng mẫu như các mặt màu riêng biệt thành dòng bit và giải mã riêng biệt các mặt màu được mã hóa tương ứng từ dòng bit. Khi các mặt màu riêng biệt được sử dụng, mỗi trong số đó được xử lý riêng biệt (bởi bộ mã hóa và/hoặc bộ giải mã) như là hình ảnh với việc lấy mẫu đơn sắc.

Việc phân vùng có thể được xác định như là sự phân chia tập hợp thành tập hợp con sao cho mỗi phần tử của tập hợp chính xác là một trong số các tập hợp con.

Trong H.264/AVC, khối macro là khối 16x16 của các mẫu luma và các khối tương ứng của các mẫu sắc độ. Ví dụ, trong mẫu lấy mẫu 4:2:0, khối macro chứa một khối 8x8 của các mẫu sắc độ cho mỗi thành phần sắc độ. Trong H.264/AVC, hình ảnh được phân chia thành một hay nhiều nhóm mảng, và nhóm mảng chứa một hay nhiều

mảng. Trong H.264/AVC, mảng bao gồm số nguyên các khối macro được sắp xếp liên tục trong quét mảnh trong nhóm mảng cụ thể.

Khi mô tả hoạt động mã hóa và/hoặc giải mã HEVC, có thể sử dụng các điều khoản sau đây. Khối mã hóa có thể được xác định như là khối $N \times N$ của các mẫu đối với một số giá trị của N để cho sự phân chia khối cây mã hóa thành các khối mã hóa là việc phân vùng. Khối cây mã hóa (CTB) có thể được xác định như là khối $N \times N$ của các mẫu đối với một số giá trị của N để cho sự phân chia của thành phần thành các khối cây mã hóa là việc phân vùng. Bộ cây mã hóa (CTU) có thể được xác định như là khối cây mã hóa của các mẫu luma, hai khối cây mã hóa tương ứng của các mẫu sắc độ của hình ảnh mà có ba mảng mẫu, hoặc khối cây mã hóa của các mẫu của hình ảnh đơn sắc hoặc hình ảnh mà được mã hóa nhờ sử dụng ba mặt màu riêng biệt và các cấu trúc cú pháp được sử dụng để mã hóa các mẫu. Bộ mã hóa (CU) có thể được xác định như là khối mã hóa của các mẫu luma, hai khối mã hóa tương ứng của các mẫu sắc độ của hình ảnh mà có ba mảng mẫu, hoặc khối mã hóa của các mẫu của hình ảnh đơn sắc hoặc hình ảnh mà được mã hóa nhờ sử dụng ba mặt màu riêng biệt và các cấu trúc cú pháp được sử dụng để mã hóa các mẫu.

Trong một số bộ mã hóa-giải mã video, như bộ mã hóa-giải mã mã hóa video hiệu suất cao (High Efficiency Video Coding, HEVC), các hình ảnh video được chia thành các bộ mã hóa (CU) bao phủ vùng hình ảnh. CU bao gồm một hay nhiều bộ dự đoán (PU) xác định quy trình dự đoán đối với các mẫu trong CU và một hay nhiều bộ biến đổi (TU) xác định quy trình mã hóa lỗi dự đoán đối với các mẫu trong CU nêu trên. Thông thường, CU bao gồm khối vuông của các mẫu với kích thước có thể chọn từ tập hợp được xác định trước của các kích thước CU thích hợp. CU với kích thước cực đại cho phép có thể được gọi là LCU (bộ mã hóa rộng nhất) hoặc bộ cây mã hóa (CTU) và hình ảnh video được chia thành các LCU không chồng lên nhau. LCU có thể chia thêm ra thành sự kết hợp của các CU nhỏ hơn, ví dụ bằng cách chia đệ quy LCU và các CU thu được. Mỗi CU thu được thường có ít nhất một PU và ít nhất một TU đi kèm với nó. Mỗi PU và TU có thể chia thêm ra thành các PU và các TU nhỏ hơn để làm tăng mức độ chi tiết của dự đoán các quy trình mã hóa lỗi dự đoán, tương ứng. Mỗi PU có thông tin dự đoán đi kèm với nó xác định loại dự đoán nào cần được áp

dụng cho các điểm ảnh trong PU đó (ví dụ thông tin vector chuyển động đối với các PU được dự đoán liên khung và thông tin định hướng dự đoán liên khung đối với các PU được dự đoán trong).

Mỗi TU có thể đi kèm với thông tin mô tả quy trình giải mã lỗi dự đoán đối với các mẫu trong TU nêu trên (gồm có, ví dụ thông tin hệ số DCT). Thông thường nó được báo hiệu tại cấp CU xem liệu việc mã hóa lỗi dự đoán có được áp dụng hay không đối với mỗi CU. Trong trường hợp không còn lại lỗi dự đoán đi kèm với CU, nó có thể được coi là không có các TU cho CU nêu trên. Sự phân chia hình ảnh thành các CU, và sự phân chia các CU thành các PU và các TU thường được báo hiệu trong dòng bit cho phép bộ giải mã sao chép lại cấu trúc được dự kiến của các bộ phận này.

Trong HEVC, hình ảnh có thể được phân chia trong các ô, chúng vốn là hình chữ nhật và chứa một số nguyên của các LCU. Trong HEVC, việc phân chia thành các ô tạo thành lưới ô bao gồm một hay nhiều cột ô và một hay nhiều hàng ô. Ô được mã hóa được liên kết byte, mà có thể đạt được nhờ việc bổ sung các bit liên kết byte tại sự kết thúc của ô được mã hóa.

Trong HEVC, mảng được xác định là số nguyên các bộ cây mã hóa được chứa trong một phân đoạn mảng độc lập và tất cả các phân đoạn mảng phụ thuộc tiếp theo (nếu có) mà đứng trước phân đoạn mảng độc lập tiếp theo (nếu có) trong cùng một bộ truy nhập. Trong HEVC, phân đoạn mảng được xác định là một số nguyên các bộ cây mã hóa được sắp xếp liên tục trong phần quét ô và được chứa trong bộ NAL đơn. Việc phân chia mỗi hình ảnh thành các phân đoạn mảng là việc phân vùng. Trong HEVC, phân đoạn mảng độc lập được xác định là phân đoạn mảng mà đối với nó, các giá trị của các phần tử cú pháp của phần đầu phân đoạn mảng không được suy ra từ các giá trị đối với phân đoạn mảng đứng trước, và phân đoạn mảng phụ thuộc được xác định là phân đoạn mảng mà đối với nó, các giá trị của một số phần tử cú pháp của phần đầu phân đoạn mảng được suy ra từ các giá trị đối với phân đoạn mảng độc lập đứng trước trong trình tự giải mã. Trong HEVC, phần đầu mảng được xác định là phần đầu phân đoạn mảng của phân đoạn mảng độc lập mà là phân đoạn mảng hiện hành hoặc là phân đoạn mảng độc lập đứng trước phân đoạn mảng phụ thuộc hiện hành, và phần đầu phân đoạn mảng được xác định là một phần của phân đoạn mảng được mã hóa chứa

các phần tử dữ liệu liên quan đến bộ thứ nhất hoặc tất cả các bộ cây mã hóa có mặt trong phân đoạn mảng. Các CU được quét theo trình tự quét mảnh của các LCU trong các ô hoặc trong hình ảnh, nếu các ô đang không được sử dụng. Trong LCU, các CU có trình tự quét riêng.

Trong HEVC, ô chứa một số nguyên các bộ cây mã hóa, và có thể bao gồm các bộ cây mã hóa chứa trong hơn một mảng. Tương tự, mảng có thể bao gồm các bộ cây mã hóa chứa trong nhiều hơn một ô. Trong HEVC, tất cả các bộ cây mã hóa trong mảng thuộc về cùng một ô và/hoặc tất cả các bộ cây mã hóa trong ô thuộc về cùng một mảng. Ngoài ra, trong HEVC, tất cả các bộ cây mã hóa trong phân đoạn mảng thuộc về cùng một ô và/hoặc tất cả các bộ cây mã hóa trong ô thuộc về cùng một phân đoạn mảng.

tập hợp ô bị hạn chế chuyển động là tập hợp mà quy trình dự đoán liên khung bị hạn chế khi mã hóa sao cho không có giá trị mẫu bên ngoài tập hợp ô bị hạn chế chuyển động, và không có giá trị mẫu tại vị trí mẫu phân đoạn mà xuất phát nhờ sử dụng một hay nhiều giá trị của mẫu bên ngoài tập hợp ô bị hạn chế chuyển động, được sử dụng để dự đoán liên khung của mẫu bất kỳ trong tập hợp ô bị hạn chế chuyển động.

Lưu ý là các vị trí mẫu được sử dụng trong dự đoán liên khung được bảo hòa sao cho vị trí mà có thể ở bên ngoài hình ảnh nếu không thì được bảo hòa để trở đến mẫu ranh giới tương ứng của hình ảnh. Do đó, nếu ranh giới ô cũng là ranh giới hình ảnh, các vector chuyển động có thể cắt hiệu quả ranh giới đó hoặc vector chuyển động có thể khiến cho việc nội suy mẫu phân đoạn hiệu quả mà có thể là vị trí bên ngoài ranh giới đó, vì các vị trí mẫu được bảo hòa lên trên ranh giới này.

Tin nhắn SEI của các tập hợp ô tạm thời bị hạn chế chuyển động của HEVC có thể được sử dụng để biểu thị sự có mặt của các tập hợp ô bị hạn chế chuyển động trong dòng bit.

Tập hợp ô bị hạn chế liên lớp là quy trình dự đoán liên lớp bị hạn chế khi mã hóa như vậy mà không có giá trị mẫu ngoài mỗi tập hợp ô tham chiếu kèm theo, và không có giá trị mẫu tại vị trí mẫu phân đoạn vốn xuất phát nhờ sử dụng một hay nhiều giá trị của mẫu bên ngoài mỗi tập hợp ô tham chiếu kèm theo, được sử dụng để

dự đoán liên lớp của mẫu bất kỳ trong tập hợp ô bị hạn chế liên lớp.

Tin nhắn SEI của tập hợp ô bị hạn chế liên lớp SEI của HEVC có thể được sử dụng để chỉ sự có mặt của các tập hợp ô bị hạn chế liên lớp trong dòng bit.

Bộ giải mã phục hồi tín hiệu video đầu ra nhờ áp dụng các phương tiện dự đoán tương tự như bộ mã hóa để tạo ra hiển thị được dự đoán của các khối điểm ảnh (sử dụng thông tin chuyển động hoặc không gian được tạo ra bởi bộ mã hóa và được lưu trong hiển thị được nén) và giải mã lỗi dự đoán (thao tác nghịch đảo của việc mã hóa lỗi dự đoán phục hồi tín hiệu lỗi dự đoán đã lượng tử hóa trong miền điểm ảnh không gian). Sau khi áp dụng dự đoán và phương tiện giải mã lỗi dự đoán, bộ giải mã tóm tắt dự đoán và các tín hiệu lỗi dự đoán (các giá trị điểm ảnh) để tạo ra khung hình video đầu ra. Bộ giải mã (và bộ mã hóa) có thể cũng áp dụng phương tiện lọc bổ sung để nâng cao chất lượng của video đầu ra trước khi chuyển sang màn hình và/hoặc lưu nó như tham chiếu dự đoán cho các khung hình sắp tới trong chuỗi video.

Việc lọc có thể, ví dụ, bao gồm thêm một bước sau: giải khối, độ lệch tương thích mẫu (SAO), và/hoặc lọc vòng tương thích (ALF). H.264/AVC bao gồm giải khối, trong khi HEVC bao gồm cả giải khối và SAO.

Trong các bộ mã hóa-giải mã video điển hình, thông tin chuyển động được biểu thị với các vector chuyển động đi kèm với mỗi khối hình ảnh bù chuyển động, như bộ dự đoán. Mỗi véc-tơ chuyển động biểu thị cho sự dịch chuyển của khối hình ảnh trong hình ảnh cần được mã hóa (ở phía bộ mã hóa) hoặc được giải mã (ở phía bộ giải mã) và khối nguồn dự đoán trong một trong các hình ảnh được mã hóa hoặc được giải mã trước. Để biểu thị các vector chuyển động hiệu quả chúng thường được mã hóa theo cách khác theo cách khóa các véc-tơ chuyển động được dự đoán riêng. Trong các bộ mã hóa-giải mã video điển hình, các véc-tơ chuyển động được dự đoán được tạo ra theo cách đã định, ví dụ tính toán giá trị trung bình của các vector chuyển động được mã hóa hoặc được giải mã của các khối liền kề. Còn cách khác để tạo ra các dự đoán vector chuyển động là tạo ra danh mục của các dự đoán dự kiến từ các khối liền kề và/hoặc các khối cùng vị trí trong các hình ảnh tham chiếu tạm thời và báo hiệu về các dự đoán dự kiến đã chọn làm bộ dự đoán vector chuyển động. Ngoài việc dự đoán các giá trị vector chuyển động, nó có thể dự đoán (các) hình ảnh tham chiếu nào được sử

dụng để dự đoán bù chuyển động và thông tin dự đoán này có thể được thể hiện, ví dụ bởi chỉ mục tham chiếu của hình ảnh được mã hóa/được giải mã từ trước. Chỉ mục tham chiếu thường được dự đoán từ các khối liền kề và/hoặc các khối cùng vị trí trong hình ảnh tham chiếu tạm thời. Hơn nữa, các bộ mã hóa-giải mã video hiệu quả cao thông thường sử dụng cơ chế mã hóa/giải mã thông tin chuyển động bổ sung, thường được gọi là chế độ hợp nhất/hợp nhất, nơi tất cả thông tin trường chuyển động, vốn bao gồm vector chuyển động và chỉ mục hình ảnh tham chiếu tương ứng cho mỗi danh mục hình ảnh tham chiếu khả dụng, được dự đoán và được sử dụng không cần bất kỳ sự thay đổi/sửa chữa nào. Tương tự, việc dự đoán thông tin trường chuyển động được thực hiện nhờ sử dụng thông tin trường chuyển động của các khối liền kề và/hoặc các khối cùng vị trí trong các hình ảnh tham chiếu tạm thời và thông tin trường chuyển động được sử dụng được báo hiệu trong danh mục của danh mục dự kiến trường chuyển động được điền đầy với thông tin trường chuyển động của các khối liền kề/cùng-vị trí khả dụng.

Các bộ mã hóa-giải mã video thông thường cho phép sử dụng một-dự đoán, nơi một khối dự đoán được sử dụng cho khối đang được mã hóa (giải mã), và hai-dự đoán, nơi các khối hai-dự đoán được kết hợp để tạo ra dự đoán cho khối đang được mã hóa (giải mã). Một số bộ mã hóa-giải mã video cho phép dự đoán có trọng số, nơi các giá trị mẫu của các khối dự đoán có trọng số trước khi bổ sung thông tin còn lại. Ví dụ, yếu tố nhân ảnh và độ lệch bổ sung có thể được áp dụng. Trong dự đoán trọng số rõ ràng, được phép bởi một số bộ mã hóa-giải mã video, yếu tố ảnh và độ lệch có thể được mã hóa ví dụ trong phần đầu mảng đối với mỗi chỉ mục hình ảnh tham chiếu cho phép. Trong dự đoán trọng số tiềm ẩn, được phép bởi một số bộ mã hóa-giải mã video, các yếu tố ảnh và/hoặc độ lệch không được mã hóa nhưng xuất phát, ví dụ, dựa trên các khoảng cách đếm trình tự hình ảnh tương đối (POC) của các hình ảnh tham chiếu.

Trong các bộ mã hóa-giải mã video thông thường, phần dự đoán còn lại sau khi bù chuyển động được biến đổi trước tiên với hạt nhân biến đổi (giống như DCT) và sau đó được mã hóa. Nguyên nhân của việc này là do vẫn thường tồn tại một số liên quan giữa phần còn lại và sự biến đổi có thể trong nhiều trường hợp giúp làm giảm sự liên quan này và làm cho việc mã hóa hiệu quả hơn.

Các bộ mã hóa-giải mã video thông thường sử dụng các hàm chi phí Lagrangian để tìm ra các chế độ mã hóa tối ưu, ví dụ chế độ khối macro mong muốn và các véc-tơ chuyển động đi kèm. Loại hàm chi phí này sử dụng yếu tố ảnh hoặc λ để ràng buộc với sự biến dạng hình ảnh (chính xác hoặc ước lượng) do các phương pháp mã hóa tổn hao và (chính xác hoặc ước lượng) lượng thông tin đòi hỏi hiển thị các giá trị điểm ảnh trong vùng hình ảnh :

$$C = D + \lambda R, (1)$$

Trong đó C là chi phí Lagrangian cần được giảm tối thiểu, D là sự biến dạng hình ảnh (ví dụ, sai số toàn phương trung bình MSE) với chế độ và các véc-tơ chuyển động được xem xét, và R là số lượng các bit cần để biểu diễn dữ liệu cần để phục hồi khối hình ảnh trong bộ giải mã (bao gồm lượng dữ liệu để biểu diễn các véc-tơ chuyển động dự kiến).

Các tiêu chuẩn mã hóa video và các thông số kỹ thuật có thể cho phép các bộ mã hóa phân chia hình ảnh được mã hóa thành các mảng được mã hóa hoặc tương tự và/hoặc các ô hoặc tương tự. Thông thường, dự đoán trong-hình ảnh là các ranh giới mảng cắt ngang và các ranh giới ô bị vô hiệu. Do đó, các mảng và các ô có thể được coi như cách chia hình ảnh được mã hóa thành các mảnh giải mã độc lập được. Trong H.264/AVC và HEVC, dự đoán trong-hình ảnh có thể là các ranh giới mảng cắt ngang bị khuyết, và trong HEVC dự đoán trong-hình ảnh có thể là các ranh giới ô cắt ngang bị khuyết. Do đó, các mảng có thể được coi là cách chia hình ảnh được mã hóa thành các mảnh giải mã độc lập được, và do đó các mảng thường được coi như các bộ phận cơ bản để biến đổi và có thể cũng được sử dụng như các bộ phận cơ bản để song song hóa. Các ô có thể được coi như các bộ phận cơ bản để song song hóa khi mã hóa và/hoặc giải mã. Trong nhiều trường hợp, các bộ mã hóa có thể biểu thị trong dòng bit loại dự đoán trong-hình ảnh nào đã tắt các ranh giới mảng cắt ngang hoặc các ranh giới ô (riêng biệt hoặc cùng nhau đối với các ranh giới mảng và ô), và thao tác giải mã đưa thông tin này vào tài khoản, ví dụ khi kết luận nguồn dự đoán nào là khả dụng. Ví dụ, các mẫu từ khối macro hoặc CU lân cận có thể được coi là không khả dụng để dự đoán liên khung, nếu khối macro hoặc CU lân cận cư trú trong mảng khác.

Bộ phận cơ bản cho đầu ra của bộ mã hóa H.264/AVC hoặc HEVC và đầu

vào của bộ giải mã H.264/AVC hoặc HEVC, tương ứng, là bộ phận lớp trừu tượng hóa mạng (Network Abstraction Layer-NAL). Để vận chuyển qua các mạng định hướng gói hoặc lưu vào các tệp cấu trúc, các bộ NAL có thể được gói thành các gói hoặc các cấu trúc tương tự. Bộ NAL có thể được xác định như là cấu trúc cú pháp chứa dấu hiệu của loại dữ liệu tiếp theo và các byte chứa dữ liệu đó dưới dạng RBSP được xen kẽ như là sự cần thiết với các byte ngăn chặn mô phỏng mã khởi đầu. Tái chuỗi byte thô (RBSP) có thể được xác định là cấu trúc cú pháp chứa một số nguyên các byte được gói trong bộ NAL. RBSP hoặc là trống hoặc có dạng chuỗi các bit dữ liệu chứa các phân tử cú pháp được tiếp nối bởi bit dừng RBSP và được tiếp nối bởi các bit không hoặc tiếp theo nữa là bằng 0. Các bộ NAL bao gồm phần đầu và tải.

Trong HEVC, phần đầu của bộ NAL hai-byte được sử dụng cho tất cả các loại bộ NAL được quy định. Phần đầu của bộ NAL chứa một bit riêng, dấu hiệu loại bộ NAL sáu-bit, dấu hiệu cho mức tạm thời ba-bit `nuh_temporal_id_plus1` (có thể được yêu cầu lớn hơn hoặc bằng 1) và phân tử cú pháp sáu-bit `nuh_layer_id`. phân tử cú pháp `temporal_id_plus1` có thể được coi như là bộ định danh tạm thời cho bộ NAL, và biến `TemporalId` nền tảng không có thể xuất phát như sau: $\text{TemporalId} = \text{temporal_id_plus1} - 1$. `TemporalId` bằng 0 tương ứng với mức tạm thời thấp nhất. Giá trị của `temporal_id_plus1` được yêu cầu không bằng không để tránh mô phỏng mã khởi động liên quan đến hai byte phần đầu của bộ NAL. Dòng bit được tạo ra bằng cách loại trừ tất cả các bộ NAL VCL có `TemporalId` lớn hơn hoặc bằng giá trị được chọn và bao gồm tất cả các bộ NAL VCL khác duy trì tính phù hợp. Do đó, hình ảnh có `TemporalId` bằng `TID` không sử dụng bất kỳ hình ảnh nào có `TemporalId` lớn hơn `TID` như tham khảo dự đoán liên khung. Lớp-phụ hoặc lớp-phụ tạm thời có thể được quy định là lớp mở rộng tạm thời được của dòng bit mở rộng tạm thời được, bao gồm các bộ NAL VCL với giá trị cụ thể của biến `TemporalId` và các bộ NAL không có VCL đi kèm. `nuh_layer_id` có thể được hiểu như là bộ định danh lớp mở rộng được.

Các bộ NAL có thể được phân loại thành các bộ NAL của lớp mã hóa video (VCL) và các bộ NAL không có-VCL. Trong H.264/AVC, các bộ NAL mảng được mã hóa chứa các phân tử cú pháp biểu diễn một hay nhiều khối macro được mã hóa, mỗi trong số đó tương ứng với khối các mẫu trong hình ảnh không nén. Trong HEVC,

các bộ NAL VCL chứa các phân tử cú pháp biểu diễn một hay nhiều CU.

Trong HEVC, bộ NAL mảng được mã hóa có thể được chỉ ra là một trong các loại sau:

nal_unit_type	Tên của nal_unit_type	Nội dung của bộ NAL và cấu trúc cú pháp RBSP
0, 1	TRAIL_N, TRAIL_R	Phân đoạn mảng được mã hóa của non-TSA, non-STSA sau hình ảnh slice_segment_layer_rbsp()
2, 3	TSA_N, TSA_R	Phân đoạn mảng được mã hóa của hình ảnh TSA slice_segment_layer_rbsp()
4, 5	STSA_N, STSA_R	Phân đoạn mảng được mã hóa của hình ảnh TSA slice_layer_rbsp()
6, 7	RADL_N, RADL_R	Phân đoạn mảng được mã hóa của hình ảnh RADL slice_layer_rbsp()
8, 9	RASL_N, RASL_R,	Phân đoạn mảng được mã hóa của hình ảnh RASL slice_layer_rbsp()
10, 12, 14	RSV_VCL_N10 RSV_VCL_N12 RSV_VCL_N14	Các loại bộ NAL VCL Reserved// reserved non-RAP non-reference
11, 13, 15	RSV_VCL_R11 RSV_VCL_R13 RSV_VCL_R15	Các loại bộ NAL VCL Reserved // reserved non-RAP reference
16, 17, 18	BLA_W_LP IDR_W_RADL BLA_N_LP	Phân đoạn mảng được mã hóa của hình ảnh BLA slice_segment_layer_rbsp()
19, 20	IDR_W_RADL IDR_N_LP	Phân đoạn mảng được mã hóa của hình ảnh IDR slice_segment_layer_rbsp()
21	CRA_NUT	Phân đoạn mảng được mã hóa của hình ảnh CRA slice_segment_layer_rbsp()
22, 23	RSV_IRAP_VCL22.. RSV_IRAP_VCL23	Các loại bộ NAL VCL Reserved // reserved RAP
24..31	RSV_VCL24.. RSV_VCL31	Các loại bộ NAL VCL Reserved // reserved non-RAP

Trong HEVC, các từ viết tắt cho các loại hình ảnh có thể được xác định như sau: Hình ảnh sau (TRAIL), Truy cập lớp-phụ tạm thời (TSA), Truy cập lớp-phụ tạm thời từng bước (STSA), Hình ảnh dẫn đầu có thể giải mã truy nhập ngẫu nhiên (RADL), Hình ảnh dẫn đầu bỏ qua truy cập ngẫu nhiên (RASL), Hình ảnh truy nhập liên kết hỏng (BLA), Hình ảnh làm mới giải mã tức thời (IDR), Hình ảnh truy cập ngẫu nhiên sạch (CRA).

Hình ảnh điểm truy cập ngẫu nhiên (RAP), vốn có thể cũng được gọi là hình ảnh điểm truy cập ngẫu nhiên nội bộ (IRAP), là hình ảnh ở đó mỗi mảng hoặc phân đoạn mảng có nal_unit_type trong phạm vi từ 16 đến 23, gộp cả. Hình ảnh IRAP trong lớp độc lập chỉ chứa các mảng được mã hóa-trong. Hình ảnh IRAP thuộc về lớp được dự đoán với giá trị nuh_layer_id currLayerId có thể chứa các mảng P, B, và I, không thể sử dụng dự đoán hướng trong từ các hình ảnh khác với nuh_layer_id bằng currLayerId, và có thể sử dụng dự đoán liên lớp từ các lớp tham chiếu trực tiếp của nó. Trong phiên bản hiện tại của HEVC, hình ảnh IRAP có thể là hình ảnh BLA, hình ảnh CRA hoặc hình ảnh IDR. Hình ảnh thứ nhất trong dòng bit chứa lớp cơ sở là hình ảnh IRAP lớp cơ sở. Các bộ thông số cần thiết được cung cấp là khả dụng khi chúng được kích hoạt, hình ảnh IRAP ở lớp độc lập và tất cả các hình ảnh non-RASL tiếp theo ở lớp độc lập trong trình tự giải mã có thể được giải mã đúng mà không cần thực hiện quy trình giải mã của các hình ảnh bất kỳ đứng trước hình ảnh IRAP trong trình tự giải mã. Hình ảnh IRAP thuộc về lớp được dự đoán với giá trị nuh_layer_id currLayerId và tất cả các hình ảnh non-RASL tiếp theo với nuh_layer_id bằng currLayerId trong trình tự giải mã có thể được giải mã đúng mà không cần thực hiện quy trình giải mã của các hình ảnh bất kỳ với nuh_layer_id bằng currLayerId mà đứng trước hình ảnh IRAP trong trình tự giải mã, khi các bộ thông số cần thiết là khả dụng khi chúng cần được kích hoạt và khi việc giải mã của mỗi lớp tham chiếu trực tiếp của lớp với nuh_layer_id bằng currLayerId được khởi tạo (tức là khi LayerInitializedFlag[refLayerId] bằng 1 đối với refLayerId bằng tất cả các giá trị nuh_layer_id của lớp tham chiếu trực tiếp của lớp với nuh_layer_id bằng currLayerId). Có thể các hình ảnh trong dòng bit mà chỉ chứa các mảng được mã hóa-nội bộ không phải các hình ảnh IRAP.

Trong HEVC, hình ảnh CRA có thể là hình ảnh đầu tiên trong dòng bit trong trình tự giải mã, hoặc có thể xuất hiện sau đó trong dòng bit. Các hình ảnh CRA trong HEVC cho phép được gọi là các hình ảnh dẫn đầu mà tiếp theo hình ảnh CRA trong trình tự giải mã nhưng đứng trước nó theo trình tự đầu ra. Một số hình ảnh dẫn đầu, được gọi là các hình ảnh RASL, có thể sử dụng các hình ảnh được giải mã trước hình ảnh CRA như là sự tham khảo. Các hình ảnh mà tiếp theo hình ảnh CRA trong cả trình tự giải mã và đầu ra là giải mã được nếu việc truy cập ngẫu nhiên được thực hiện ở hình ảnh CRA, và do đó việc truy cập ngẫu nhiên không có lỗi đạt được tương tự như đối với chức năng truy cập ngẫu nhiên không có lỗi của hình ảnh IDR.

Hình ảnh CRA có thể đi kèm các hình ảnh RADL hoặc RASL. Khi hình ảnh CRA là hình ảnh đầu tiên trong dòng bit trong trình tự giải mã, hình ảnh CRA là hình ảnh đầu tiên của chuỗi video được mã hóa trong trình tự giải mã, và các hình ảnh RASL đi kèm bất kỳ không phải là tín hiệu đầu ra bởi bộ giải mã và có thể không giải mã được, cũng như chúng có thể chứa các tham chiếu đối với các hình ảnh mà không có mặt trong dòng bit.

Hình ảnh dẫn đầu là hình ảnh đứng trước hình ảnh RAP đi kèm trong trình tự đầu ra. Hình ảnh RAP đi kèm là hình ảnh RAP trước đó trong trình tự giải mã (nếu có). Hình ảnh dẫn đầu hoặc là hình ảnh RADL hoặc hình ảnh RASL.

Tất cả các hình ảnh RASL là các hình ảnh dẫn đầu của hình ảnh BLA hoặc hình ảnh CRA đi kèm. Khi hình ảnh RAP đi kèm là hình ảnh BLA hoặc là hình ảnh đầu tiên được mã hóa trong dòng bit, hình ảnh RASL không phải đầu ra và có thể không được giải mã đúng, cũng như hình ảnh RASL có thể chứa các tham chiếu đến các hình ảnh mà không có mặt trong dòng bit. Tuy nhiên, hình ảnh RASL có thể được giải mã đúng nếu việc giải mã được bắt đầu từ hình ảnh RAP trước hình ảnh RAP đi kèm của hình ảnh RASL. Các hình ảnh RASL không được sử dụng làm các hình ảnh tham chiếu cho quy trình giải mã của các hình ảnh non-RASL. Khi có mặt, tất cả các hình ảnh RASL đứng trước, trong trình tự giải mã, tất cả các hình ảnh đứng sau của cùng hình ảnh RAP đi kèm. Trong một số bản thảo của Tiêu chuẩn HEVC, hình ảnh RASL đã được gọi là hình ảnh được gắn thẻ để loại bỏ (TFD).

Tất cả các hình ảnh RADL là các hình ảnh dẫn đầu. Các hình ảnh RADL

không được sử dụng làm các hình ảnh tham chiếu đối với quy trình giải mã các hình ảnh đứng sau của cùng một hình ảnh RAP đi kèm. Khi có mặt, tất cả các hình ảnh RADL đứng trước, trong trình tự giải mã, tất cả các hình ảnh đứng sau của cùng hình ảnh RAP đi kèm. Các hình ảnh RADL không viện dẫn đến hình ảnh bất kỳ đứng trước hình ảnh RAP đi kèm trong trình tự giải mã và do đó có thể được giải mã đúng khi việc giải mã bắt đầu từ hình ảnh RAP đi kèm.

Khi một phân dòng bit bắt đầu từ hình ảnh CRA chứa trong dòng bit khác, các hình ảnh RASL đi kèm với hình ảnh CRA có thể không giải mã đúng được, vì một số hình ảnh tham chiếu của chúng có thể không có mặt trong dòng bit được kết hợp. Để làm cho thao tác ghép nối đơn giản, loại bộ NAL của hình ảnh CRA có thể thay đổi để biểu thị đó là hình ảnh BLA. Các hình ảnh RASL đi kèm với hình ảnh BLA có thể không giải mã đúng được do đó không được xuất/hiển thị. Ngoài ra, các hình ảnh RASL đi kèm với hình ảnh BLA có thể bỏ qua việc giải mã.

Hình ảnh BLA có thể là hình ảnh đầu tiên trong dòng bit trong trình tự giải mã, hoặc có thể xuất hiện sau đó trong dòng bit. Mỗi hình ảnh BLA bắt đầu chuỗi video được mã hóa mới, và tương tự như hiệu ứng lên quy trình giải mã như là hình ảnh IDR. Tuy nhiên, hình ảnh BLA chứa các phần tử cú pháp mà xác định bộ hình ảnh tham chiếu không-trống. Khi hình ảnh BLA có `nal_unit_type` bằng `BLA_W_LP`, có thể có đi kèm các hình ảnh RASL, mà không phải là đầu ra bởi bộ giải mã và có thể giải mã được, cũng như chúng có thể chứa các tham chiếu đến các hình ảnh không có mặt trong dòng bit. Khi hình ảnh BLA có `nal_unit_type` bằng `BLA_W_LP`, có thể cũng có đi kèm các hình ảnh RADL, vốn được chỉ định để được giải mã. Khi hình ảnh BLA có `nal_unit_type` bằng `BLA_W_RADL`, nó không có các hình ảnh RASL đi kèm nhưng có thể có đi kèm các hình ảnh RADL, vốn được chỉ định để được giải mã. Khi hình ảnh BLA có `nal_unit_type` bằng `BLA_N_LP`, không có bất kỳ hình ảnh dẫn đầu đi kèm nào.

Hình ảnh IDR có `nal_unit_type` bằng `IDR_N_LP` không có các hình ảnh dẫn đầu đi kèm có mặt trong dòng bit. Hình ảnh IDR có `nal_unit_type` bằng `IDR_W_LP` không có các hình ảnh RASL đi kèm có mặt trong dòng bit, nhưng có thể có các hình ảnh RADL đi kèm trong dòng bit.

Khi giá trị của `nal_unit_type` bằng `TRAIL_N`, `TSA_N`, `STSA_N`, `RADL_N`, `RASL_N`, `RSV_VCL_N10`, `RSV_VCL_N12`, hoặc `RSV_VCL_N14`, hình ảnh được giải mã không được sử dụng như là tham chiếu đối với hình ảnh bất kỳ khác của cùng một lớp phụ tạm thời. Tức là, trong HEVC, khi giá trị của `nal_unit_type` bằng `TRAIL_N`, `TSA_N`, `STSA_N`, `RADL_N`, `RASL_N`, `RSV_VCL_N10`, `RSV_VCL_N12`, hoặc `RSV_VCL_N14`, hình ảnh được giải mã không chứa trong của `RefPicSetStCurrBefore`, `RefPicSetStCurrAfter` và `RefPicSetLtCurr` của hình ảnh bất kỳ với cùng giá trị của `TemporalId`. Hình ảnh được mã hóa với `nal_unit_type` bằng `TRAIL_N`, `TSA_N`, `STSA_N`, `RADL_N`, `RASL_N`, `RSV_VCL_N10`, `RSV_VCL_N12`, hoặc `RSV_VCL_N14` có thể được loại bỏ mà không làm ảnh hưởng khả năng giải mã của các hình ảnh khác có cùng giá trị của `TemporalId`.

Hình ảnh đứng sau có thể được xác định như là hình ảnh mà tiếp theo hình ảnh RAP đi kèm theo trình tự đầu ra. Hình ảnh bất kỳ là hình ảnh đứng sau không có `nal_unit_type` bằng `RADL_N`, `RADL_R`, `RASL_N` hoặc `RASL_R`. Hình ảnh bất kỳ mà là hình ảnh dẫn đầu có thể bị hạn chế để đứng trước, trong trình tự giải mã, tất cả các hình ảnh đứng sau mà đi kèm với cùng hình ảnh RAP. Không có các hình ảnh `RASL` có mặt trong dòng bit mà đi kèm với cùng hình ảnh `BLA` có `nal_unit_type` bằng `BLA_W_RADL` hoặc `BLA_N_LP`. Không có các hình ảnh `RADL` có mặt trong dòng bit mà đi kèm với hình ảnh `BLA` có `nal_unit_type` bằng `BLA_N_LP` hoặc đi kèm với hình ảnh `IDR` có `nal_unit_type` bằng `IDR_N_LP`. Hình ảnh `RASL` bất kỳ đi kèm với hình ảnh `CRA` hoặc `BLA` có thể bị hạn chế để đứng trước hình ảnh `RADL` bất kỳ đi kèm với hình ảnh `CRA` hoặc `BLA` theo trình tự đầu ra. Hình ảnh `RASL` bất kỳ đi kèm với hình ảnh `CRA` có thể bị hạn chế tiếp theo, theo trình tự đầu ra, bất kỳ hình ảnh RAP nào khác mà đứng trước hình ảnh `CRA` trong trình tự giải mã.

Trong HEVC có hai loại hình ảnh, loại hình ảnh `TSA` và loại hình ảnh `STSA` mà có thể được sử dụng để biểu thị các điểm chuyển lớp-phụ tạm thời. Nếu lớp-phụ tạm thời với `TemporalId` lên đến `N` được giải mã cho đến khi hình ảnh `TSA` hoặc `STSA` (độc quyền) và hình ảnh `TSA` hoặc `STSA` có `TemporalId` bằng `N+1`, hình ảnh `TSA` hoặc `STSA` cho phép giải mã tất cả các hình ảnh tiếp theo (trong trình tự giải mã) có `TemporalId` bằng `N+1`. Loại hình ảnh `TSA` có thể áp đặt các hạn chế lên chính hình

ảnh TSA và tất cả các hình ảnh trong cùng lớp-phụ tiếp theo hình ảnh TSA trong trình tự giải mã. Không hình ảnh nào được phép sử dụng dự đoán hướng trong từ hình ảnh bất kỳ trong cùng lớp-phụ mà đứng trước hình ảnh TSA trong trình tự giải mã. Định nghĩa TSA có thể áp đặt thêm các hạn chế lên các hình ảnh trong các lớp-phụ cao hơn mà tiếp theo hình ảnh TSA trong trình tự giải mã. Không hình ảnh nào được phép hình ảnh tham chiếu đứng trước hình ảnh TSA trong trình tự giải mã nếu hình ảnh đó thuộc cùng lớp-phụ hoặc lớp-phụ cao hơn như hình ảnh TSA. Các hình ảnh TSA có TemporalId lớn hơn 0. STSA tương tự như hình ảnh TSA nhưng không áp đặt các hạn chế lên các hình ảnh trong lớp-phụ cao hơn tiếp theo hình ảnh STSA trong trình tự giải mã và do đó cho phép chuyển đổi lên trên lớp-phụ nơi các hình ảnh STSA khu trú.

Bộ NAL non-VCL có thể là, ví dụ, một trong số các loại sau: bộ thông số chuỗi, bộ thông số hình ảnh, bộ NAL của thông tin tăng cường bổ sung (SEI), dấu phân cách bộ truy nhập, sự kết thúc của bộ NAL chuỗi, sự kết thúc của bộ NAL dòng bit, hoặc bộ NAL dữ liệu đệm. Các bộ thông số có thể cần để phục hồi các hình ảnh được giải mã, trong khi nhiều bộ NAL non-VCL khác không cần thiết để phục hồi các giá trị của mẫu được giải mã.

Các thông số mà duy trì không thay đổi thông qua chuỗi video được mã hóa có thể chứa trong bộ thông số chuỗi. Ngoài các thông số mà có thể cần thiết bởi quy trình giải mã, bộ thông số chuỗi có thể tùy ý chứa thông tin khả năng sử dụng video (VUI), vốn bao gồm các thông số có thể quan trọng để đệm, định giờ xuất hình ảnh, dựng hình, và lưu giữ tài nguyên. Trong HEVC bộ thông số chuỗi RBSP bao gồm các thông số có thể được tham chiếu đến bởi một hay nhiều tập hợp thông số hình ảnh RBSP hoặc một hay nhiều bộ NAL SEI chứa tin nhắn giai đoạn đệm SEI. Bộ thông số hình ảnh chứa các thông số mà có thể không thay đổi trong nhiều hình ảnh được mã hóa. Bộ thông số hình ảnh RBSP có thể bao gồm các thông số mà có thể được tham chiếu đến bởi các bộ NAL mảng được mã hóa của một hay nhiều hình ảnh được mã hóa.

Trong HEVC, bộ thông số video (VPS) có thể được xác định như là cấu trúc cú pháp chứa các phần tử cú pháp mà áp dụng cho các chuỗi video được mã hóa số không hoặc toàn bộ như được xác định bởi nội dung của phần tử cú pháp tìm thấy

trong SPS được tham chiếu bởi phần tử cú pháp tìm thấy trong PPS được tham chiếu bởi phần tử cú pháp tìm thấy trong mỗi phần đầu phân đoạn mảng. Bộ thông số video RBSP có thể bao gồm các thông số mà có thể được tham chiếu bởi một hay nhiều bộ thông số chuỗi RBSP.

Mối liên hệ và sự phân cấp giữa bộ thông số video (VPS), bộ thông số chuỗi (SPS), và bộ thông số hình ảnh (PPS) có thể được mô tả như sau. VPS cư trú một cấp trên SPS trong phân cấp bộ thông số và trong bối cảnh khả năng mở rộng và/hoặc video 3D. VPS có thể bao gồm các thông số là thông thường đối với tất cả các mảng giao với tất cả các lớp (có khả năng mở rộng hoặc nhìn) trong toàn bộ chuỗi video được mã hóa. SPS bao gồm các thông số là thông thường đối với tất cả các mảng trong lớp cụ thể (có khả năng mở rộng hoặc nhìn) trong toàn bộ chuỗi video được mã hóa, và có thể được chia sẻ bởi nhiều lớp (có khả năng mở rộng hoặc nhìn). PPS bao gồm các thông số mà là thông thường đối với tất cả các mảng trong hiển thị lớp cụ thể (hiển thị khả năng mở rộng hoặc lớp nhìn trong một bộ truy nhập) và có thể được chia sẻ bởi tất cả các mảng trong các lớp hiển thị.

VPS có thể cung cấp thông tin về các mối quan hệ phụ thuộc của các lớp trong dòng bit, cũng như nhiều thông tin khác mà có thể áp dụng với tất cả các mảng trên tất cả các lớp (có khả năng mở rộng hoặc nhìn) trong toàn bộ chuỗi video được mã hóa. VPS có thể được coi là bao gồm hai phần, VPS cơ bản và VPS mở rộng, tại đó VPS mở rộng có thể tùy ý có mặt. Trong HEVC, VPS cơ bản có thể được coi là bao gồm cấu trúc cú pháp của bộ thông số video (rbps) mà không có cấu trúc cú pháp mở rộng (VPS). Cấu trúc cú pháp của bộ thông số video (RBSP) đã được xác định chủ yếu đối với Phiên bản HEVC 1 và bao gồm các phần tử cú pháp vốn có thể sử dụng cho lớp giải mã cơ bản. Trong HEVC, VPS mở rộng có thể được coi là bao gồm cấu trúc cú pháp mở rộng (VPS). Cấu trúc cú pháp mở rộng (VPS) đã được xác định trong Phiên bản HEVC 2 chủ yếu cho các phần mở rộng nhiều-lớp và bao gồm các phần tử cú pháp vốn có thể sử dụng để giải mã một hay nhiều lớp không-cơ sở, như các phần tử cú pháp biểu thị các quan hệ phụ thuộc về lớp.

Phần tử cú pháp `max_tid_il_ref_pics_plus1` trong VPS mở rộng có thể được sử dụng để biểu thị là các hình ảnh không-IRAP không sử dụng tham chiếu để dự đoán

liên lớp và, nếu không như vậy, các lớp-phụ tạm thời không được sử dụng làm tham chiếu cho dự đoán liên lớp:

$\text{max_tid_il_ref_pics_plus1}[i][j]$ bằng 0 xác định mà các hình ảnh non-IRAP với nuh_layer_id bằng $\text{layer_id_in_nuh}[i]$ không được sử dụng làm nguồn các hình ảnh để dự đoán liên lớp cho các hình ảnh với nuh_layer_id bằng $\text{layer_id_in_nuh}[j]$. $\text{max_tid_il_ref_pics_plus1}[i][j]$ lớn hơn 0 xác định là các hình ảnh với nuh_layer_id bằng $\text{layer_id_in_nuh}[i]$ và TemporalId lớn hơn $\text{max_tid_il_ref_pics_plus1}[i][j] - 1$ không được sử dụng làm nguồn các hình ảnh để dự đoán liên lớp cho các hình ảnh với nuh_layer_id bằng $\text{layer_id_in_nuh}[j]$. Khi không có mặt, giá trị của $\text{max_tid_il_ref_pics_plus1}[i][j]$ được suy ra bằng 7.

Cú pháp H.264/AVC và HEVC cho phép nhiều ví dụ về các bộ thông số, và mỗi ví dụ được xác định với bộ định danh duy nhất. Để giới hạn việc sử dụng bộ nhớ cần cho các bộ thông số, giá trị phạm vi đối với các bộ định danh bộ thông số được giới hạn. Trong H.264/AVC và HEVC, mỗi phần đầu mảng bao gồm bộ định danh của bộ thông số hình ảnh mà hoạt động để giải mã hình ảnh chứa mảng, và mỗi bộ thông số hình ảnh chứa bộ định danh của bộ thông số chuỗi đang hoạt động. Do đó, việc truyền hình ảnh và các bộ thông số chuỗi không được đồng bộ hóa chính xác với việc truyền dẫn các mảng. Thay vào đó, cũng đủ là chuỗi hoạt động và các bộ thông số hình ảnh được nhận tại thời điểm bất kỳ trước khi chúng được tham chiếu, mà cho phép truyền các bộ thông số “out-of-band” nhờ sử dụng cơ chế truyền tin cậy hơn được so sánh với các giao thức được sử dụng dữ liệu mảng. Ví dụ, các bộ thông số có thể chứa thông số trong phiên mô tả đối với các phiên giao thức truyền tải thời gian thực (RTP). Nếu các bộ thông số được truyền trong băng, chúng có thể được lặp lại để nâng cao độ bền lỗi.

Việc biến đổi out-of-band, gửi tín hiệu hoặc lưu có thể được sử dụng bổ sung hoặc thay thế vì các mục đích khác với dung sai chống lại các lỗi biến đổi, như sự giảm truy nhập hoặc đàm phán phiên. Ví dụ, phần nhập mẫu của đường ghi trong tệp phù hợp với ISOBMFF có thể bao gồm các bộ thông số, trong khi dữ liệu được mã hóa trong dòng bit được lưu tại nơi khác trong tệp hoặc trong tệp khác. Cụm từ dọc theo dòng bit (ví dụ việc biểu thị dọc theo dòng bit) có thể được sử dụng trong yêu cầu bảo

hộ và các phương án tham chiếu được mô tả để tham chiếu sự biến đổi out-of-band, gửi tín hiệu, hoặc lưu theo cách là dữ liệu out-of-band được kèm với dòng bit. Việc giải mã đoạn văn đọc theo dòng bit hoặc tương tự có thể tham chiếu để giải mã dữ liệu được gọi là out-of-band (mà có thể thu được từ việc biến đổi out-of-band, truyền tín hiệu, hoặc lưu) được kèm theo với dòng bit. Hình ảnh được mã hóa là sự hiển thị được mã hóa của hình ảnh.

Trong HEVC, hình ảnh được mã hóa có thể được xác định như là sự hiển thị được mã hóa của hình ảnh chứa tất cả các bộ cây mã hóa của hình ảnh. Trong HEVC, bộ truy nhập (AU) có thể được xác định như là tập hợp của các bộ NAL đi kèm với nhau theo quy tắc phân loại được quy định, là liên tục trong trình tự giải mã, và chứa nhiều nhất một hình ảnh với giá trị riêng bất kỳ của `nuh_layer_id`. Ngoài việc chứa các bộ NAL VCL của hình ảnh được mã hóa, bộ truy nhập có thể cũng chứa các bộ NAL non-VCL.

Có thể yêu cầu là các hình ảnh được mã hóa xuất hiện theo trình tự nhất định trong bộ truy nhập. Ví dụ, hình ảnh được mã hóa với `nuh_layer_id` bằng `nuhLayerIdA` có thể được yêu cầu đứng trước, trong trình tự giải mã, tất cả các hình ảnh được mã hóa với `nuh_layer_id` lớn hơn `nuhLayerIdA` trong cùng một bộ truy nhập. Thông thường AU chứa tất cả các hình ảnh được mã hóa mà biểu diễn cho cùng thời gian đầu ra và/hoặc thời gian chụp.

Dòng bit có thể được xác định như là chuỗi của các bit, dưới dạng dòng của bộ NAL hoặc dòng byte, mà tạo thành sự biểu diễn các hình ảnh được mã hóa và dữ liệu đi kèm tạo thành một hay nhiều chuỗi video được mã hóa. Dòng bit thứ nhất có thể được tiếp nối bởi dòng bit thứ hai trong cùng một kênh hợp lý, như trong cùng tệp hoặc trong cùng kết nối của giao thức truyền thông. Dòng cơ bản (trong bối cảnh mã hóa video) có thể được xác định như là một chuỗi của một hay nhiều dòng bit. Sự kết thúc của dòng bit thứ nhất có thể được biểu thị bởi bộ NAL riêng, mà có thể được tham chiếu như là sự kết thúc của bộ NAL dòng bit (EOB) và nó là bộ NAL cuối cùng của dòng bit. Trong HEVC và các phần mở rộng dự thảo hiện tại của nó, bộ NAL EOB yêu cầu có `nuh_layer_id` bằng 0.

Định dạng của dòng byte đã được quy định trong H.264/AVC và

HEVC để biến đổi hoặc lưu các môi trường mà không tạo ra các cấu trúc tạo khung hình. Định dạng dòng byte phân chia các bộ NAL ra khỏi nhau nhờ đánh kèm mã khởi động ở phía trước mỗi bộ NAL. Để tránh phát hiện sai các ranh giới của bộ NAL, các bộ mã hóa vận hành thao tác ngăn ngừa mô phỏng mã khởi động định hướng-byte, mà bổ sung byte ngăn ngừa mô phỏng vào tải của bộ NAL nếu mã khởi động có thể xuất hiện theo cách khác. Để, ví dụ, cho phép vận hành cổng vào đơn giản giữa các hệ thống định hướng-gói và định hướng-dòng, mã khởi động ngăn ngừa mô phỏng có thể luôn luôn thực hiện bất kể liệu có sử dụng định dạng dòng byte hay không. Trình tự bit dùng cho định dạng dòng byte có thể được xác định bắt đầu với bit quan trọng nhất (MSB) của byte thứ nhất, tiếp theo bit ít quan trọng (LSB) của byte thứ nhất, được tiếp nối bởi MSB của byte thứ hai, v. v.. Định dạng dòng byte có thể được coi là để bao gồm chuỗi các cấu trúc cú pháp bộ NAL của dòng byte. Mỗi cấu trúc cú pháp bộ NAL của dòng byte có thể được coi là chứa một tiền tố mã khởi động được tiếp nối bởi một cấu trúc cú pháp bộ NAL, tức là cấu trúc cú pháp nal_unit(Num bytesInNalUnit) nếu các tên phần tử cú pháp được tham chiếu đến. Bộ NAL của dòng byte còn có thể chứa phần tử cú pháp zero_byte bổ sung. Nó có thể còn chứa một hay nhiều phần tử cú pháp trailing_zero_8bits bổ sung. Khi bộ NAL của dòng byte là bộ NAL của dòng byte thứ nhất trong dòng bit, nó có thể cũng chứa một hay nhiều phần tử cú pháp leading_zero_8bits bổ sung. Cú pháp của bộ NAL của dòng byte có thể được quy định như sau:

byte_stream_nal_unit(NumBytesInNalUnit) {	Bộ mô tả
while(next_bits(24) != 0x000001 && next_bits(32) != 0x00000001)	
leading_zero_8bits /* bằng 0x00 */	f(8)
if(next_bits(24) != 0x000001)	
zero_byte /* bằng 0x00 */	f(8)
start_code_prefix_one_3bytes /* bằng 0x000001 */	f(24)
nal_unit(NumBytesInNalUnit)	
while(more_data_in_byte_stream() && next_bits(24) != 0x000001 && next_bits(32) != 0x00000001)	
trailing_zero_8bits /* bằng 0x00 */	f(8)
}	

Trình tự các bộ NAL của dòng byte trong dòng byte có thể được yêu cầu để tiếp trong trình tự giải mã của các bộ NAL chứa trong các bộ NAL của dòng byte. Các nghĩa của các phần tử cú pháp có thể được quy định như sau. `leading_zero_8bits` là byte bằng 0x00. Phần tử cú pháp `leading_zero_8bits` có thể chỉ có mặt trong bộ NAL của dòng byte thứ nhất của dòng bit, vì các byte bất kỳ bằng 0x00 mà tiếp theo cấu trúc cú pháp bộ NAL và đứng trước chuỗi four-byte 0x00000001 (mà cần được dẫn giải như là `zero_byte` được tiếp nối bởi `start_code_prefix_one_3byte`) sẽ được coi là các phần tử cú pháp `trailing_zero_8bits` là một phần của bộ NAL của dòng byte đứng trước. `zero_byte` là một byte đơn bằng 0x00. `start_code_prefix_one_3byte` là chuỗi giá trị cố định của 3 byte bằng 0x000001. Phần tử cú pháp này có thể được gọi là tiền tố mã khởi động (hoặc đơn giản là mã khởi động). `trailing_zero_8bits` là byte bằng 0x00.

Bộ NAL có thể được xác định như là cấu trúc cú pháp chứa dấu hiệu của loại dữ liệu tiếp theo và các byte chứa dữ liệu đó dưới dạng RBSP được xen kẽ như là sự cần thiết với các byte ngăn ngừa mô phỏng. Tái chuỗi byte thô (RBSP) có thể được xác định như là cấu trúc cú pháp chứa một số nguyên các byte mà được gói trong bộ NAL. RBSP hoặc là trống hoặc có dạng chuỗi các bit dữ liệu chứa các phần tử cú pháp được tiếp nối bởi bit dừng RBSP và được tiếp nối bởi các bit tiếp theo bit không hoặc nhiều hơn bằng 0.

Các bộ NAL bao gồm phần đầu và tải. Trong H.264/AVC và HEVC, phần đầu của bộ NAL biểu thị thể loại của bộ NAL.

Cú pháp HEVC của cấu trúc cú pháp `nal_unit(NumbytesInNalUnit)` được cung cấp tiếp theo làm ví dụ của cú pháp của bộ NAL.

nal_unit(NumBytesInNalUnit) {	Bộ mô tả
nal_unit_header()	
NumBytesInRbsp = 0	
for(i = 2; i < NumBytesInNalUnit; i++)	
if(i + 2 < NumBytesInNalUnit && next_bits(24) == 0x000003) {	
 rbsp_byte[NumBytesInRbsp++]	b(8)
 rbsp_byte[NumBytesInRbsp++]	b(8)
 i += 2	
 emulation_prevention_three_byte /* bằng 0x03 */	f(8)
 } else	
 rbsp_byte[NumBytesInRbsp++]	b(8)
}	

Trong HEVC, chuỗi video được mã hóa (CVS) có thể được xác định, ví dụ, như là chuỗi các bộ truy nhập mà bao gồm, trong trình tự giải mã, bộ truy nhập IRAP với NoRaslOutputFlag bằng 1, được tiếp nối bởi các bộ truy nhập 0 hoặc nhiều hơn mà không phải là các bộ truy nhập IRAP với NoRaslOutputFlag bằng 1, bao gồm tất cả các bộ truy nhập tiếp theo trở lên nhưng không bao gồm bộ truy nhập tiếp theo bất kỳ là bộ truy nhập IRAP với NoRaslOutputFlag bằng 1. Bộ truy nhập IRAP có thể được xác định như là bộ truy nhập trong đó hình ảnh lớp cơ sở là hình ảnh IRAP. Giá trị của NoRaslOutputFlag là bằng 1 đối với mỗi hình ảnh IDR, mỗi hình ảnh BLA, và mỗi hình ảnh IRAP vốn là hình ảnh đầu tiên trong đó lớp cụ thể trong dòng bit trong trình tự giải mã, là hình ảnh IRAP đầu tiên tiếp theo sự kết thúc của chuỗi bộ NAL có cùng giá trị trình tự giải mã nuh_layer_idin. Trong HEVC nhiều-lớp, giá trị NoRaslOutputFlag là bằng 1 đối với mỗi hình ảnh IRAP khi nuh_layer_id của nó là LayerInitializedFlag[nuh_layer_id] bằng 0 và LayerInitializedFlag[refLayerId] bằng 1 đối với tất cả các giá trị refLayerId bằng IdDirectRefLayer[nuh_layer_id][j], trong đó j ở trong phạm vi từ 0 đến NumDirectRefLayers [nuh_layer_id] – 1, gộp vào. Mặt khác, giá trị NoRaslOutputFlag là bằng HandleCraAsBlaFlag. NoRaslOutputFlag bằng 1 có tác động mà các hình ảnh RASL đi kèm với hình ảnh IRAP mà NoRaslOutputFlag là tập hợp không được xuất ra bởi bộ giải mã. Đây có thể là cách tạo ra giá trị HandleCraAsBlaFlag cho bộ giải mã từ thực thể bên ngoài, như bộ phát hoặc bộ thu, vốn có thể điều khiển bộ giải mã. HandleCraAsBlaFlag có thể là tập hợp

cho 1, ví dụ bởi bộ phát mà tìm kiếm vị trí mới trong dòng bit hoặc các âm điệu vào phát sóng và bắt đầu giải mã và sau đó bắt đầu giải mã từ hình ảnh CRA. Khi HandleCraAsBlaFlag bằng 1 đối với hình ảnh CRA, hình ảnh CRA xử lý và được giải mã nếu nó là hình ảnh BLA.

Trong HEVC, ngoài ra hoặc theo cách khác, chuỗi video được mã hóa có thể (đối với thông số kỹ thuật ở trên) được xác định đến cuối, khi bộ NAL riêng biệt, vốn có thể được gọi là sự kết thúc của bộ NAL của chuỗi (EOS), xuất hiện trong dòng bit và có nuh_layer_id bằng 0.

Nhóm các hình ảnh (GOP) và các đặc tính của chúng có thể được xác định như sau. GOP có thể được giải mã bất kể các hình ảnh bất kỳ trước đó có được giải mã hay không. GOP mở là loại mà nhóm các hình ảnh trong đó các hình ảnh đứng trước hình ảnh bên trong ban đầu theo trình tự đầu ra có thể không được giải mã đúng khi việc giải mã bắt đầu từ hình ảnh bên trong ban đầu của GOP mở. Nói cách khác, các hình ảnh của GOP mở có thể tham chiếu (trong dự đoán liên khung) đến các hình ảnh thuộc về GOP trước đó. Bộ giải mã HEVC có thể xác nhận hình ảnh bên trong bắt đầu GOP mở, vì loại bộ NAL riêng biệt, loại bộ NAL CRA, có thể được sử dụng cho các mảng được mã hóa mảng của nó. GOP kín là nhóm các hình ảnh trong đó tất cả các hình ảnh có thể được giải mã đúng khi việc giải mã bắt đầu từ hình ảnh bên trong ban đầu của GOP kín. Nói cách khác, không có hình ảnh trong GOP kín tham chiếu đến các hình ảnh bất kỳ trong các GOP trước đó. Trong H.264/AVC và HEVC, GOP kín có thể bắt đầu từ hình ảnh IDR. Trong HEVC, GOP kín cũng có thể bắt đầu từ hình ảnh BLA_W_RADL hoặc BLA_N_LP. Cấu trúc mã hóa GOP mở có khả năng hiệu quả hơn khi nén so với cấu trúc mã hóa GOP kín, do tính linh hoạt rộng hơn khi lựa chọn các hình ảnh tham chiếu.

Cấu trúc của các hình ảnh (SOP) có thể được xác định như là một hay nhiều hình ảnh được mã hóa liên tục trong trình tự giải mã, trong đó hình ảnh được mã hóa đầu tiên trong trình tự giải mã là hình ảnh tham chiếu tại lớp-phụ tạm thời thấp nhất và không có hình ảnh được mã hóa có khả năng ngoại trừ hình ảnh được mã hóa đầu tiên trong trình tự giải mã là hình ảnh RAP. Tất cả các hình ảnh trong SOP trước đứng trước trong trình tự giải mã, tất cả các hình ảnh trong SOP hiện tại và tất cả các hình

ảnh trong SOP tiếp theo thành công trong trình tự giải mã tất cả các hình ảnh trong SOP hiện tại. SOP có thể đại diện cho cấu trúc dự đoán liên khung phân cấp và lặp lại. Nhóm thời hạn các hình ảnh (GOP) đôi khi có thể được sử dụng thay đổi được với thuật ngữ SOP và có cùng nghĩa như nghĩa của SOP.

Cú pháp dòng bit của H.264/AVC và HEVC cho biết liệu hình ảnh cụ thể có phải là hình ảnh tham chiếu để dự đoán liên khung hình ảnh bất kỳ khác hay không. Các hình ảnh của loại mã hóa (I, P, B) bất kỳ có thể là các hình ảnh tham chiếu hoặc các hình ảnh không-tham chiếu trong H.264/AVC và HEVC.

Trong HEVC, cấu trúc cú pháp của tập hợp hình ảnh tham chiếu (RPS) và quy trình giải mã được sử dụng. Tập hợp hình ảnh tham chiếu có hiệu lực hoặc hoạt động đối với hình ảnh bao gồm tất cả các hình ảnh tham chiếu được sử dụng làm tham chiếu đối với hình ảnh và tất cả các hình ảnh tham chiếu mà được đánh dấu như “được sử dụng để tham chiếu” các hình ảnh tiếp theo bất kỳ trong trình tự giải mã. Có sáu tập hợp con của bộ hình ảnh tham chiếu, mà được gọi cụ thể là RefPicSetStCurr0 (cũng được gọi là RefPicSetStCurrBefore), RefPicSetStCurr1 (cũng được gọi là RefPicSetStCurrAfter), RefPicSetStFoll0, RefPicSetStFoll1, RefPicSetLtCurr, và RefPicSetLtFoll. RefPicSetStFoll0 và RefPicSetStFoll1 có thể cũng được coi là để cùng nhau tạo ra một tập hợp con RefPicSetStFoll. Ký hiệu của sáu tập hợp con như sau. “Curr” đề cập đến các hình ảnh tham chiếu mà chứa trong các danh mục hình ảnh tham chiếu của hình ảnh hiện tại và do đó có thể được sử dụng làm tham chiếu dự đoán liên khung đối với hình ảnh hiện tại. “Foll” đề cập đến các hình ảnh tham chiếu mà không chứa trong các danh mục hình ảnh tham chiếu của hình ảnh hiện tại nhưng có thể được sử dụng trong các hình ảnh tiếp theo trong trình tự giải mã như là các hình ảnh tham chiếu. “St” đề cập đến các hình ảnh tham chiếu ngắn hạn, nói chung vốn có thể được xác định thông qua số lượng nhất định của các bit ít quan trọng của giá trị POC của chúng. “Lt” đề cập đến các hình ảnh tham chiếu dài hạn, vốn được xác định cụ thể và nói chung có sự khác biệt lớn hơn của các giá trị POC liên quan tới hình ảnh hiện tại so với cái có thể được thể hiện bởi số lượng nhất định được đề cập của các bit ít quan trọng. “0” đề cập đến các hình ảnh tham chiếu mà có giá trị POC nhỏ hơn của hình ảnh hiện tại. “1” đề cập đến các hình ảnh tham chiếu mà có giá trị POC lớn hơn

của hình ảnh hiện tại. RefPicSetStCurr0, RefPicSetStCurr1, RefPicSetStFoll0 và RefPicSetStFoll1 được gọi chung là tập hợp con ngắn hạn của bộ hình ảnh tham chiếu. RefPicSetLtCurr và RefPicSetLtFoll được gọi chung là tập hợp con dài hạn của bộ hình ảnh tham chiếu.

Trong HEVC, tập hợp hình ảnh tham chiếu có thể được xác định trong bộ thông số chuỗi và được đưa vào sử dụng sử dụng trong phần đầu mảng thông qua chỉ mục đến bộ hình ảnh tham chiếu. Tập hợp hình ảnh tham chiếu có thể cũng được xác định trong phần đầu mảng. Tập hợp hình ảnh tham chiếu có thể được mã hóa độc lập hoặc có thể được dự đoán từ tập hợp hình ảnh tham chiếu khác (được biết là dự đoán liên khung RPS). Trong cả hai loại mã hóa tập hợp hình ảnh tham chiếu, cờ (được sử dụng bởi `_curr_pic_X_flag`) được gửi bổ sung cho mỗi hình ảnh tham chiếu chỉ rõ liệu hình ảnh tham chiếu được sử dụng để tham chiếu bởi hình ảnh hiện tại (chứa trong `*Curr list`) hoặc không (chứa trong `*Foll list`). Các hình ảnh mà chứa trong tập hợp hình ảnh tham chiếu được sử dụng bởi mảng hiện tại được đánh dấu như “được sử dụng để tham chiếu”, và các hình ảnh mà không có trong tập hợp hình ảnh tham chiếu được sử dụng bởi mảng hiện tại được đánh dấu như “không được sử dụng để tham chiếu”. Nếu hình ảnh hiện tại là hình ảnh IDR, RefPicSetStCurr0, RefPicSetStCurr1, RefPicSetStFoll0, RefPicSetStFoll1, RefPicSetLtCurr, và RefPicSetLtFoll tất cả tập hợp là trống.

Bộ đệm hình ảnh được giải mã (DPB) có thể được sử dụng trong bộ mã hóa và/hoặc trong bộ giải mã. Có hai lý do để đệm các hình ảnh được giải mã, cho các tham chiếu trong dự đoán liên khung và để sắp xếp lại các hình ảnh được giải mã vào trình tự đầu ra. H.264/AVC và HEVC tạo ra rất nhiều tính linh hoạt cho cả việc tạo ra hình ảnh tham chiếu và sắp xếp lại đầu ra, các bộ đệm riêng biệt để đệm hình ảnh tham chiếu và đệm hình ảnh đầu ra có thể làm lãng phí các nguồn bộ nhớ. Do đó, DPB có thể bao gồm quy trình đệm hình ảnh thống nhất cho các hình ảnh tham chiếu và sắp xếp lại đầu ra. Hình ảnh được giải mã có thể lấy ra từ DPB khi nó không được sử dụng lâu hơn làm tham chiếu và không cần thiết cho đầu ra.

Trong nhiều chế độ mã hóa của H.264/AVC và HEVC, hình ảnh tham chiếu để dự đoán liên khung được biểu thị với chỉ mục cho danh mục hình ảnh tham chiếu.

Chỉ mục có thể được mã hóa với việc mã hóa độ dài khác nhau, vốn thường sinh ra chỉ mục nhỏ hơn để có giá trị ngắn hơn đối với phần tử cú pháp tương ứng. Trong H.264/AVC và HEVC, hai danh mục hình ảnh tham chiếu (danh mục hình ảnh tham chiếu 0 và danh mục hình ảnh tham chiếu 1) được sinh ra đối với mỗi mảng hai-dự đoán (B), và một danh mục hình ảnh tham chiếu (danh mục hình ảnh tham chiếu 0) được tạo thành cho mỗi mảng được mã hóa-trong (P).

Danh mục hình ảnh tham chiếu, như danh mục hình ảnh tham chiếu 0 và danh mục hình ảnh tham chiếu 1, thường được xây dựng theo hai bước: Thứ nhất, danh mục hình ảnh tham chiếu ban đầu được tạo ra. Danh mục hình ảnh tham chiếu ban đầu có thể được tạo ra, ví dụ trên cơ sở của `frame_num`, `POC`, `temporal_id` (hoặc `TemporalId` hoặc tương tự), hoặc thông tin về sự phân cấp dự đoán như cấu trúc GOP, hoặc sự kết hợp bất kỳ của chúng. Thứ hai, danh mục hình ảnh tham chiếu ban đầu có thể được sắp xếp lại nhờ các lệnh sắp xếp lại danh mục hình ảnh tham chiếu (RPLR), còn được biết là cấu trúc cú pháp sửa đổi danh mục hình ảnh tham chiếu, vốn có thể được chứa trong các phần đầu mảng. Nếu bộ hình ảnh tham chiếu được sử dụng, danh mục hình ảnh tham chiếu 0 có thể được khởi tạo để chứa `RefPicSetStCurr0` thứ nhất, được tiếp nối bởi `RefPicSetStCurr1`, được tiếp nối bởi `RefPicSetLtCurr`. Danh mục hình ảnh tham chiếu 1 có thể được khởi tạo để chứa `RefPicSetStCurr1` thứ nhất, được tiếp nối bởi `RefPicSetStCurr0`. Trong HEVC, các danh mục hình ảnh tham chiếu ban đầu có thể được thay đổi thông qua cấu trúc cú pháp thay đổi danh mục hình ảnh tham chiếu, tại đó các hình ảnh trong các danh mục hình ảnh tham chiếu ban đầu có thể được xác định thông qua chỉ mục truy nhập đến danh mục này. Nói cách khác, trong HEVC, việc thay đổi danh mục hình ảnh tham chiếu được mã hóa thành cấu trúc cú pháp bao gồm vòng lặp đối với mỗi cửa vào trong danh mục hình ảnh tham chiếu cuối cùng, nơi mỗi cửa vào vòng lặp là chỉ mục được mã hóa độ dài cố định đối với danh mục hình ảnh tham chiếu ban đầu và biểu diễn hình ảnh theo trình tự vị trí tăng dần trong danh mục hình ảnh tham chiếu cuối cùng.

Nhiều tiêu chuẩn mã hóa, bao gồm H.264/AVC và HEVC, có thể có quy trình giải mã để chuyển hóa chỉ mục hình ảnh tham chiếu thành danh mục hình ảnh tham chiếu, vốn có thể được sử dụng để biểu thị cái nào trong số các hình ảnh tham

chiều được sử dụng để dự đoán liên khung đối với một khối cụ thể. Chỉ mục hình ảnh tham chiếu có thể được mã hóa bởi bộ mã hóa thành dòng bit là một số chế độ mã hóa nội bộ hoặc nó có thể xuất phát (bởi bộ mã hóa và bộ giải mã) ví dụ, nhờ sử dụng các khối lân cận trong một số chế độ mã hóa nội bộ khác.

Để biểu diễn các véc-tơ chuyển động hiệu quả trong các dòng bit, các véc-tơ chuyển động có thể được mã hóa khác biệt về mặt vector chuyển động được dự đoán riêng biệt theo khối. Trong nhiều bộ mã hóa-giải mã video, các véc-tơ chuyển động được dự đoán được tạo ra theo cách đã định, ví dụ nhờ tính toán giá trị trung bình của các véc-tơ chuyển động được mã hóa hoặc được giải mã của các khối liên kề. Cách khác để tạo ra các dự đoán vector chuyển động, đôi khi được gọi là dự đoán vector chuyển động tiên tiến (AMVP), để tạo ra danh mục của các dự đoán dự kiến từ các khối liên kề và/hoặc các khối cùng vị trí trong các hình ảnh tham chiếu tạm thời và báo hiệu dự kiến được chọn như là bộ phận dự đoán vector chuyển động. Ngoài việc dự đoán các giá trị vector chuyển động, chỉ mục tham chiếu của hình ảnh được mã hóa/được giải mã trước có thể được dự đoán. Thông thường, chỉ mục tham chiếu được dự đoán từ các khối liên kề và/hoặc các khối cùng vị trí trong hình ảnh tham chiếu tạm thời. Việc mã hóa khác biệt các véc-tơ chuyển động thường là các ranh giới mảng cắt ngang bị hỏng.

Việc mã hóa video mở rộng được có thể tham chiếu đến cấu trúc mã hóa nơi một dòng bit có thể chứa nhiều cách thể hiện của nội dung, ví dụ, ở các tốc độ bit, độ phân giải hoặc các tỷ lệ khung hình khác nhau. Trong trường hợp này, bộ thu có thể trích xuất sự biểu diễn mong muốn tùy thuộc vào các đặc tính của nó (ví dụ, độ phân giải mà phù hợp nhất với thiết bị hiển thị). Theo cách khác, máy chủ hoặc bộ phận mạng có thể trích xuất các phần của dòng bit cần được truyền đến bộ thu tùy thuộc vào, ví dụ các đặc tính của mạng hoặc khả năng xử lý của bộ thu. Việc hiển thị được giải mã có ý nghĩa có thể được tạo ra bởi chỉ các phần được giải mã nhất định của dòng bit mở rộng được. Thông thường, dòng bit mở rộng được bao gồm “lớp cơ sở” tạo ra video chất lượng thấp nhất khả dụng và một hay nhiều lớp tăng cường mà làm tăng chất lượng video khi được nhận và được giải mã cùng với các lớp thấp hơn. Để nâng cao hiệu quả mã hóa cho lớp tăng cường, sự hiển thị được mã hóa của lớp này

thường phụ thuộc vào lớp thấp hơn. Ví dụ chuyển động và chế độ thông tin của lớp tăng cường có thể được dự đoán từ lớp thấp hơn. Tương tự, dữ liệu điểm ảnh của lớp thấp hơn có thể được sử dụng để tạo ra dự đoán cho lớp tăng cường.

Trong một số hệ thống mã hóa video mở rộng được, tín hiệu video có thể được mã hóa thành lớp cơ sở và một hay nhiều lớp tăng cường. Lớp tăng cường có thể nâng cao, ví dụ, độ phân giải tạm thời (tức là, tỷ lệ khung hình), độ phân giải không gian, hoặc đơn giản là chất lượng của nội dung video được thể hiện bởi lớp khác hoặc một phần của nó. Mỗi lớp cùng với tất cả các lớp phụ thuộc của nó là một hiển thị của tín hiệu video, ví dụ, ở độ phân giải không gian, độ phân giải tạm thời và mức chất lượng nhất định. Trong tài liệu này, chúng ta tham chiếu lớp mở rộng được cùng với tất cả các lớp phụ thuộc của nó như là “hiển thị lớp mở rộng được”. Một phần của dòng bit mở rộng được tương ứng việc hiển thị lớp mở rộng được có thể được trích xuất và được giải mã để tạo ra sự hiển thị của tín hiệu gốc với sự tin cậy nhất định.

Các chế độ mở rộng được hoặc các kích thước mở rộng được có thể bao gồm nhưng không bị giới hạn ở các tiêu chí dưới đây:

- Khả năng mở rộng chất lượng: Các hình ảnh lớp cơ sở được mã hóa ở chất lượng thấp hơn các hình ảnh lớp tăng cường, vốn có thể đạt được, ví dụ nhờ sử dụng giá trị thông số lượng tử hóa lớn hơn (tức là, kích thước bước lượng tử hóa lớn hơn để lượng tử hóa hệ số biến đổi) trong lớp cơ sở so với trong lớp tăng cường.

- Khả năng mở rộng không gian: Các hình ảnh lớp cơ sở được mã hóa ở độ phân giải thấp hơn (tức là có các mẫu ít hơn) so với các hình ảnh lớp tăng cường. Khả năng mở rộng không gian và khả năng mở rộng chất lượng, cụ thể là loại khả năng mở rộng hạt thô của nó, đôi khi có thể được coi là cùng loại của khả năng mở rộng.

- Khả năng mở rộng độ sâu-bit: Các hình ảnh lớp cơ sở được mã hóa ở độ sâu bit (ví dụ 8 bit) so với các hình ảnh lớp tăng cường (ví dụ 10 hoặc 12 bit).

- Khả năng mở rộng dải tần nhạy sáng: Các lớp mở rộng được thể hiện dải tần nhạy sáng khác nhau và/hoặc các hình ảnh thu được nhờ sử dụng chức năng chỉnh sửa ánh sáng khác nhau và/hoặc chức năng truyền quang khác nhau.

- Khả năng mở rộng định dạng sắc độ: Các hình ảnh lớp cơ sở tạo ra độ phân

giải không gian thấp hơn trong mảng mẫu sắc độ (ví dụ, được mã hóa trong định dạng sắc độ 4:2:0) so với các hình ảnh lớp tăng cường (ví dụ định dạng 4:4:4).

- Khả năng mở rộng độ rộng của giải màu: Các hình ảnh lớp tăng cường có phạm vi hiển thị màu giàu hơn/rộng hơn so với của các hình ảnh lớp cơ sở ví dụ lớp tăng cường có thể có độ rộng của giải màu UHDTV (ITU-R BT.2020) và lớp cơ sở có thể có độ rộng của giải màu ITU-R BT.709.

- Khả năng mở rộng khung hình, vốn có thể cũng được gọi là việc mã hóa hiển thị nhiều khung hình. Lớp cơ sở hiển thị khung hình thứ nhất, trong khi lớp tăng cường hiển thị khung hình thứ hai.

- Khả năng mở rộng độ sâu, vốn có thể cũng được gọi là mã hóa tăng cường độ sâu. Lớp hoặc một số lớp của dòng bit có thể hiển thị (các) khung hình dạng vãi dẹt, trong khi lớp hoặc các lớp khác có thể hiển thị (các) khung hình sâu.

- Khả năng mở rộng vùng quan tâm (như được mô tả dưới đây).

- Khả năng mở rộng Xen kẽ tuần tự (còn được biết là khả năng mở rộng trường đến khung): Tài liệu nội dung nguồn được mã hóa xen kẽ của lớp cơ sở được nâng cao nhờ lớp tăng cường để hiển thị nội dung nguồn lũy tiến.

- Khả năng mở rộng bộ mã hóa-giải mã hỗn hợp (còn được biết là khả năng mở rộng tiêu chuẩn mã hóa): Khi mở rộng mã hỗn hợp, cú pháp dòng bit, các ngữ nghĩa và quy trình giải mã của lớp cơ sở và lớp tăng cường được quy định trong các tiêu chuẩn mã hóa video khác nhau. Do đó, các hình ảnh lớp cơ sở được mã hóa theo tiêu chuẩn hoặc định dạng mã hóa video khác nhau so với các hình ảnh lớp tăng cường. Ví dụ, lớp cơ sở có thể được mã hóa với H.264/AVC và lớp tăng cường có thể được mã hóa với sự mở rộng HEVC nhiều-lớp.

Cần phải hiểu là nhiều loại khả năng mở rộng có thể được kết hợp và được áp dụng với nhau. Ví dụ, có thể kết hợp khả năng mở rộng độ rộng của giải màu và khả năng mở rộng độ sâu bit.

Lớp giới hạn có thể được sử dụng trong bối cảnh của loại khả năng mở rộng bất kỳ, bao gồm khả năng mở rộng khung hình và tăng cường độ sâu. Lớp tăng cường có thể tham chiếu loại tăng cường bất kỳ, như SNR, về không gian, hiển thị nhiều

khung hình, độ sâu, độ sâu bit, định dạng sắc độ, và/hoặc tăng cường độ rộng của giải màu. Lớp cơ sở có thể là loại chuỗi video cơ sở bất kỳ, như hình cơ sở, lớp cơ sở đối với khả năng mở rộng SNR/không gian, hoặc hình cơ sở dạng dẹt để mã hóa video độ sâu nâng cao.

Các công nghệ khác nhau để tạo ra nội dung video ba chiều (3D) đang được nghiên cứu và phát triển. Điều có thể được xem xét là trong video hai-khung hình hoặc lập thể, một chuỗi hoặc khung hình video được hiển thị cho mắt trái trong khi khung hình song song được hiển thị cho mắt phải. Nhiều hơn hai khung hình song song có thể cần cho các ứng dụng mà cho phép chuyển điểm nhìn hoặc để hiển thị lập thể tự động vốn có thể hiển thị số lượng lớn các khung hình đồng thời và cho phép người xem quan sát nội dung từ các điểm nhìn khác nhau.

Khung hình có thể được xác định như là chuỗi các hình ảnh hiển thị một camera hoặc điểm nhìn. Các hình ảnh hiển thị khung hình có thể cũng được gọi là các thành phần của khung hình. Nói cách khác, thành phần của khung hình có thể được xác định như là sự hiển thị được mã hóa khung hình trong bộ phận truy nhập đơn. Trong mã hóa video hiển thị nhiều khung hình, nhiều hơn một khung hình được mã hóa trong dòng bit. Thông thường, vì các khung hình được dự kiến để được hiển thị trên màn hình lập thể hoặc hiển thị nhiều khung hình lập thể tự động hoặc được sử dụng cho các bố trí 3D khác, chúng thường hiển thị cùng một cảnh và là sự trùng lặp một phần về nội dung mặc dù hiển thị các điểm nhìn khác nhau đối với nội dung. Do đó, dự đoán liên khung hình có thể được ứng dụng khi mã hóa video hiển thị nhiều khung hình để có lợi về sự tương quan hình xen kẽ và nâng cao hiệu quả nén. Một cách để thực hiện dự đoán liên khung hình là bao gồm một hay nhiều các hình ảnh được giải mã của một hay nhiều khung hình khác trong (các) danh mục hình ảnh tham chiếu của hình ảnh đang được mã hóa hoặc được giải mã cư trú trong hình ảnh đầu tiên. Khả năng mở rộng khung hình có thể tham chiếu việc mã hóa video nhiều khung hình hoặc các dòng bit video nhiều khung hình, mà cho phép loại bỏ hoặc bỏ sót một hay nhiều khung hình được mã hóa trong khi dòng bit thu được duy trì sự phù hợp và hiển thị video với số lượng các khung hình nhỏ hơn so với ban đầu. Việc mã hóa vùng quan tâm (ROI) có thể được xác định tham chiếu việc mã hóa một vùng cụ thể trong

video với độ tin cậy cao hơn.

Khả năng mở rộng ROI có thể được xác định như là loại khả năng mở rộng trong đó các lớp tăng cường nâng cao chỉ là một phần của hình ảnh lớp-tham chiếu, ví dụ về không gian, về chất lượng, trong độ sâu bit, và/hoặc cùng với các kích thước mở rộng được khác. Khả năng mở rộng ROI có thể được sử dụng cùng với các loại khả năng mở rộng khác, nó có thể được cho là để tạo ra sự phân loại khác của các loại khả năng mở rộng. Có nhiều ứng dụng khác nhau đối với việc mã hóa ROI với các điều kiện khác nhau, vốn có thể được thực hiện nhờ sử dụng khả năng mở rộng ROI. Ví dụ, lớp tăng cường có thể được biến đổi để nâng cao chất lượng và/hoặc độ phân giải của vùng trong lớp cơ sở. Bộ giải mã nhận cả sự tăng cường và dòng bit lớp cơ sở có thể giải mã cả hai lớp và bao phủ các hình ảnh được giải mã trên đỉnh của nhau và hiển thị hình ảnh cuối cùng.

Sự tương ứng về không gian hình ảnh lớp tham chiếu và hình ảnh lớp tăng cường có thể được suy ra hoặc có thể được biểu thị với một hay nhiều loại được gọi là các độ lệch vị trí lớp tham chiếu. Trong HEVC, các độ lệch vị trí lớp tham chiếu có thể chứa trong PPS bởi bộ mã hóa và được giải mã từ PPS bởi bộ giải mã. Các độ lệch vị trí lớp tham chiếu có thể được sử dụng cho nhưng không bị giới hạn ở việc đạt được khả năng mở rộng ROI. Các độ lệch vị trí lớp tham chiếu có thể bao gồm một hay nhiều độ lệch vị trí lớp tham chiếu được thu nhỏ, các độ lệch vùng tham chiếu, và các bộ giai đoạn lấy mẫu lại. Các độ lệch lớp tham chiếu thu nhỏ có thể được coi là xác định các độ lệch theo chiều ngang và chiều đứng giữa mẫu trong hình ảnh hiện tại mà được đặt cùng chỗ với mẫu luma ở trên cùng-bên trái của vùng tham chiếu trong hình ảnh được giải mã trong lớp tham chiếu và các độ lệch theo chiều ngang và chiều đứng giữa mẫu trong hình ảnh hiện tại mà được đặt cùng chỗ với mẫu luma ở dưới cùng-bên phải của vùng tham chiếu trong hình ảnh được giải mã trong lớp tham chiếu. Cách khác là xem xét các độ lệch lớp tham chiếu được thu nhỏ để xác định các vị trí của các mẫu góc của vùng tham chiếu không được lấy mẫu liên quan đến các mẫu góc tương ứng của hình ảnh lớp tăng cường. Các giá trị độ lệch lớp tham chiếu thu nhỏ có thể được đánh dấu. Các độ lệch vùng tham chiếu có thể được coi là xác định các độ lệch theo chiều ngang và chiều đứng giữa mẫu luma ở trên cùng-bên trái của vùng tham

chiều trong hình ảnh được giải mã trong lớp tham chiếu và mẫu luma ở trên cùng-bên trái của cùng hình ảnh được giải mã cũng như các độ lệch theo chiều ngang và chiều đứng giữa mẫu luma ở dưới cùng-bên phải của vùng tham chiếu trong hình ảnh được giải mã trong lớp tham chiếu và mẫu luma ở dưới cùng-bên phải của cùng hình ảnh được giải mã. Các giá trị độ lệch vùng tham chiếu có thể được đánh dấu. Tập hợp giai đoạn lấy mẫu lại có thể được coi là xác định các độ lệch giai đoạn được sử dụng trong quy trình lấy mẫu lại của ảnh nguồn cho dự đoán liên lớp. Các độ lệch của giai đoạn khác nhau có thể được tạo ra cho các thành phần luma và sắc độ.

Trong video lập thể khung hình tương thích (cũng được gọi là gói khung hình của video lập thể), việc gói về không gian của cặp video thành khung hình đơn được thực hiện ở phía bộ mã hóa như là bước xử lý trước để mã hóa và sau đó các khung hình được gói khung được mã hóa bởi hệ thống mã hóa video 2D thông thường. Các khung hình đầu ra được tạo ra bởi bộ giải mã chứa các khung hình cấu thành của cặp stereo.

Trong chế độ vận hành thông thường, độ phân giải không gian của các khung hình gốc của mỗi hình và khung hình đơn được gói có cùng độ phân giải. Trong trường hợp này bộ mã hóa hạ tần số lấy mẫu hai hình của video lập thể trước thao tác gói. Việc gói về không gian có thể sử dụng ví dụ định dạng cạnh nhau (side-by-side) hoặc trên-dưới (top-bottom), và việc hạ tần số lấy mẫu cần được thực hiện một cách phù hợp.

Việc gói khung hình có thể được ưu tiên đối với việc mã hóa video nhiều khung hình (ví dụ mở rộng MVC của H.264/AVC hoặc sự mở rộng MV-HEVC của H.265/HEVC) ví dụ do những nguyên nhân sau:

- Các quy trình làm việc sau sản xuất có thể được thiết kế cho tín hiệu video đơn. Một số công cụ sau sản xuất có thể không có khả năng xử lý hai chuỗi hình ảnh riêng biệt và/hoặc có thể không có khả năng giữ các chuỗi hình ảnh riêng biệt đồng bộ với nhau.
- Hệ thống phân phối, như các giao thức biến đổi, có thể chỉ hỗ trợ chuỗi được mã hóa đơn và/hoặc có thể không có khả năng giữ các chuỗi được mã hóa riêng biệt đồng bộ với nhau và/hoặc có thể yêu cầu đệm nhiều hoặc độ trễ giữ lại các chuỗi được

mã hóa riêng biệt đồng bộ với nhau.

- Việc giải mã các dòng bit với các công cụ mã hóa video nhiều khung hình có thể đòi hỏi sự hỗ trợ của các chế độ mã hóa riêng biệt, mà có thể không khả dụng với bộ phát. Ví dụ, nhiều điện thoại thông minh hỗ trợ giải mã hồ sơ chính H.265/HEVC nhưng không thể xử lý giải mã hồ sơ chính nhiều khung hình H.265/HEVC ngay cả khi nó chỉ yêu cầu các bổ sung mức độ cao so với hồ sơ chính.

- Một số phương án mã hóa video mở rộng được có thể yêu cầu các hình ảnh IRAP cần được sắp thẳng hàng đi qua các lớp theo cách mà hoặc tất cả các hình ảnh trong bộ truy nhập là các hình ảnh IRAP hoặc không có hình ảnh nào trong bộ truy nhập là hình ảnh IRAP. Các phương án mã hóa video mở rộng được khác, như các phần mở rộng nhiều-lớp của HEVC, có thể cho phép các hình ảnh IRAP mà không sắp thẳng hàng, tức là một hay nhiều hình ảnh trong bộ truy nhập là các hình ảnh IRAP, trong khi một hay nhiều hình ảnh khác trong bộ truy nhập không phải là các hình ảnh IRAP. Các dòng bit mở rộng được với các hình ảnh IRAP hoặc tương tự mà không được sắp thẳng hàng ngang qua các lớp có thể được sử dụng, ví dụ để tạo ra nhiều hình ảnh IRAP thường xuyên trong lớp cơ sở, nơi chúng có thể có kích thước được mã hóa nhỏ hơn do ví dụ độ phân giải không gian nhỏ hơn. Quy trình hoặc cơ cấu để khởi động từng lớp giải mã có thể chứa trong hệ thống giải mã video. Do đó các bộ giải mã có thể bắt đầu giải mã dòng bit khi lớp cơ sở chứa hình ảnh IRAP và bắt đầu giải mã từng bước lớp khác khi chúng chứa các hình ảnh IRAP. Nói cách khác, khi khởi động từng lớp của cơ cấu hoặc quy trình giải mã, các bộ giải mã làm tăng dần dần số lượng các lớp được giải mã lớp (nơi các lớp có thể hiển thị sự tăng cường về độ phân giải không gian, mức chất lượng, các khung hình, các thành phần bổ sung như độ sâu, hoặc sự kết hợp của chúng) như là các hình ảnh tiếp theo từ các lớp tăng cường bổ sung được giải mã trong quy trình giải mã. Việc làm tăng theo lũy tiến số lượng các lớp được giải mã có thể nhận ra ví dụ sự nâng cao theo lũy tiến của chất lượng hình ảnh (trong trường hợp khả năng mở rộng chất lượng và không gian).

Cơ cấu khởi động từng lớp có thể tạo ra các hình ảnh không khả dụng đối với các hình ảnh tham chiếu của hình ảnh đầu tiên trong trình tự giải mã trong lớp tăng cường cụ thể. Theo cách khác, bộ giải mã có thể bỏ sót giải mã các hình ảnh đứng

trước, trong trình tự giải mã, hình ảnh IRAP mà từ đó việc giải mã của lớp có thể bắt đầu. Các hình ảnh mà có thể bị bỏ sót có thể được đánh dấu riêng bởi bộ mã hóa hoặc thực thể khác trong dòng bit. Ví dụ, một hay nhiều loại bộ NAL riêng biệt có thể được sử dụng. Các hình ảnh này, bất kể liệu chúng có được đánh dấu riêng với loại bộ NAL hoặc được suy ra ví dụ bởi bộ giải mã hay không, có thể được gọi là các hình ảnh bỏ qua truy nhập ngẫu nhiên lớp cắt ngang (CL-RAS). Bộ giải mã có thể bỏ sót đầu ra của các hình ảnh không khả dụng được tạo ra và các hình ảnh được giải mã CL-RAS.

Khả năng mở rộng có thể được phép theo hai cách cơ bản. Hoặc bằng cách đưa vào các chế độ mã hóa mới để thực hiện dự đoán các giá trị điểm ảnh hoặc cú pháp từ lớp thấp hơn của hiển thị mở rộng được hoặc bằng cách đặt các hình ảnh lớp thấp hơn vào phần đệm hình ảnh tham chiếu (ví dụ phần đệm hình ảnh được giải mã, DPB) của lớp cao hơn. Phương pháp tiếp cận đầu tiên có thể linh hoạt hơn và do đó có thể tạo ra hiệu quả mã hóa tốt hơn trong phần lớn trường hợp. Tuy nhiên, thứ hai, khả năng mở rộng dựa trên khung hình tham chiếu, phương pháp tiếp cận có thể được sử dụng hiệu quả với những thay đổi tối thiểu đối với các bộ mã hóa-giải mã lớp đơn trong khi vẫn đạt được phần lớn những lợi ích của hiệu quả mã hóa có ích. Về cơ bản, bộ mã hóa-giải mã có thể mở rộng dựa trên nền tảng khung hình tham chiếu có thể được sử dụng bằng cách ứng dụng cùng một phần cứng hoặc phần mềm thực hiện cho tất cả các lớp, chỉ quan tâm tới việc quản lý DPB bởi phương tiện bên ngoài.

Bộ mã hóa video mở rộng được dùng để mở rộng chất lượng (còn được biết là tín hiệu/nhiều hoặc SNR) và/hoặc khả năng mở rộng không gian có thể được sử dụng như sau. Đối với lớp cơ sở, bộ mã hóa video thông thường không mở rộng được và bộ giải mã có thể được sử dụng. Các hình ảnh được phục hồi/ được giải mã của lớp cơ sở được chứa trong bộ đệm hình ảnh tham chiếu và/hoặc các danh mục hình ảnh tham chiếu dùng cho lớp tăng cường. Trong trường hợp khả năng mở rộng về không gian, hình ảnh lớp cơ sở được phục hồi/được giải mã có thể không được lấy mẫu trước khi chèn nó vào các danh mục hình ảnh tham chiếu dùng cho hình ảnh lớp tăng cường. Các hình ảnh được giải mã lớp cơ sở có thể được chèn vào (các) danh mục hình ảnh tham chiếu để mã hóa/giải mã hình ảnh lớp tăng cường tương tự như các hình ảnh tham chiếu được giải mã của lớp tăng cường. Do đó, bộ mã hóa có thể chọn hình ảnh

tham chiếu lớp cơ sở như là sự tham chiếu dự đoán liên khung và chỉ định việc sử dụng của nó với chỉ mục hình ảnh tham chiếu trong dòng bit được mã hóa. Bộ giải mã giải mã từ dòng bit, ví dụ từ chỉ mục hình ảnh tham chiếu, mà hình ảnh lớp cơ sở được sử dụng để tham chiếu dự đoán liên khung cho lớp tăng cường. Khi ấy hình ảnh lớp cơ sở được giải mã được sử dụng để tham chiếu dự đoán cho lớp tăng cường, ưu tiên là hình ảnh tham chiếu liên lớp.

Trong khi đoạn trên được mô tả, mã video mở rộng được với hai lớp mở rộng được với lớp tăng cường và lớp cơ sở, cần được hiểu là sự mô tả có thể được tổng quát hóa thành hai lớp theo sự phân cấp khả năng mở rộng với nhiều hơn hai lớp. Trong trường hợp này, lớp tăng cường thứ hai có thể phụ thuộc vào lớp tăng cường thứ nhất trong các quy trình mã hóa và/hoặc giải mã, và lớp tăng cường thứ nhất có thể do đó được coi như lớp cơ sở cho việc mã hóa và/hoặc giải mã lớp tăng cường thứ hai. Ngoài ra, cần phải hiểu là có thể các hình ảnh tham chiếu liên lớp từ nhiều hơn một lớp trong bộ đệm hình ảnh tham chiếu hoặc các danh mục hình ảnh tham chiếu của lớp tăng cường, và mỗi trong số các hình ảnh tham chiếu liên lớp có thể được coi là cư trú trong lớp cơ sở hoặc lớp tham chiếu đối với lớp tăng cường đang được mã hóa và/hoặc được giải mã. Ngoài ra, cần phải hiểu là các loại liên lớp khác xử lý tăng tốc độ lấy mẫu hình ảnh lớp tham chiếu có thể được thực hiện thay vào đó hoặc bổ sung. Ví dụ, độ sâu bit của các mẫu của hình ảnh lớp tham chiếu có thể được chuyển hóa thành độ sâu bit của lớp tăng cường và/hoặc các giá trị của mẫu có thể trải qua ánh xạ từ không gian màu của lớp tham chiếu đến không gian màu của lớp tăng cường.

Việc mã hóa video mở rộng được và/hoặc hệ thống giải mã có thể sử dụng mã hóa và/hoặc giải mã nhiều vòng lặp, có thể đặc trưng như sau. Trong mã hóa/giải mã, hình ảnh lớp cơ sở có thể được phục hồi/được giải mã để được sử dụng như là hình ảnh tham chiếu bù chuyển động cho các hình ảnh tiếp theo, theo trình tự mã hóa/giải mã, trong cùng lớp hoặc như là tham chiếu đối với dự đoán liên lớp (hoặc hình xen kẽ hoặc thành phần bên trong). Hình ảnh lớp cơ sở được phục hồi/được giải mã có thể được lưu trong DPB. Hình ảnh lớp tăng cường có thể được phục hồi/được giải mã tương tự như vậy để được sử dụng như là hình ảnh tham chiếu bù chuyển động cho các hình ảnh tiếp theo, theo trình tự mã hóa/giải mã, trong cùng lớp hoặc như tham chiếu

cho dự đoán liên lớp (hoặc hình xen kẽ hoặc thành phần xen kẽ) đối với lớp tăng cường cao hơn, nếu có. Ngoài các giá trị của mẫu được phục hồi/được giải mã, các giá trị phần tử cú pháp của lớp cơ sở/tham chiếu hoặc các biến xuất phát từ các giá trị phần tử cú pháp của lớp cơ sở/tham chiếu có thể được sử dụng trong dự đoán liên lớp/thành phần xen kẽ/hình xen kẽ.

Dự đoán liên lớp có thể được xác định như là dự đoán theo cách mà phụ thuộc vào các phần tử dữ liệu (ví dụ, các giá trị của mẫu hoặc các véc-tơ chuyển động) của các hình ảnh tham chiếu từ lớp khác với lớp của hình ảnh hiện tại (đang được mã hóa hoặc được giải mã). Nhiều loại dự đoán liên lớp tồn tại và có thể được áp dụng trong bộ mã hóa/bộ giải mã video mở rộng được. Các loại dự đoán liên lớp khả dụng có thể ví dụ phụ thuộc vào hồ sơ mã hóa theo đó dòng bit hoặc lớp cụ thể trong dòng bit được mã hóa hoặc, khi giải mã, hồ sơ mã hóa dòng bit đó hoặc lớp cụ thể trong dòng bit được biểu thị phù hợp. Theo cách khác hoặc ngoài ra, các loại dự đoán liên lớp khả dụng có thể phụ thuộc vào loại khả năng mở rộng hoặc loại mã mở rộng được hoặc việc sửa đổi tiêu chuẩn mã hóa video (ví dụ SHVC, MV-HEVC, hoặc 3D-HEVC) được sử dụng.

Các loại dự đoán liên lớp có thể bao gồm, nhưng không bị giới hạn ở, một hay nhiều loại sau: dự đoán mẫu liên lớp, dự đoán chuyển động liên lớp, dự đoán phân dư liên lớp. Trong dự đoán mẫu liên lớp, ít nhất một tập hợp con của các giá trị của mẫu được phục hồi của hình ảnh nguồn để dự đoán liên lớp được sử dụng như tham chiếu để dự đoán các giá trị của mẫu của hình ảnh hiện tại. Trong dự đoán chuyển động liên lớp, ít nhất một tập hợp con của các véc-tơ chuyển động của hình ảnh nguồn để dự đoán liên lớp được sử dụng như là tham chiếu để dự đoán các vectơ chuyển động của hình ảnh hiện tại. Thông thường, thông tin dự đoán mà nhờ nó các hình ảnh tham chiếu được đi kèm với các vectơ chuyển động cũng chứa trong dự đoán chuyển động liên lớp. Ví dụ, các chỉ số tham chiếu của các hình ảnh tham chiếu đối với các vectơ chuyển động có thể được dự đoán xen kẽ và/hoặc việc đếm trình tự hình ảnh hoặc sự nhận biết bất kỳ nào khác của của hình ảnh tham chiếu có thể được dự đoán liên lớp. Trong một số trường hợp, dự đoán chuyển động liên lớp còn có thể bao gồm dự đoán chế độ mã hóa khối, thông tin phân đầu, phân chia khối, và/hoặc các thông số

tương tự khác. Trong một số trường hợp, dự đoán thông số mã hóa, như dự đoán liên lớp của việc phân chia khối, có thể được coi là loại dự đoán liên lớp khác. Trong dự đoán phần dư liên lớp, lỗi dự đoán hoặc phần dư của các khối được chọn hình ảnh nguồn cho dự đoán liên lớp được sử dụng để dự đoán hình ảnh hiện tại. Khi mã hóa multiview-plus-depth, như 3D-HEVC, dự đoán liên lớp thành phần-cắt ngang có thể được áp dụng, trong đó hình ảnh loại thứ nhất, như hình ảnh sâu, có thể ảnh hưởng đến dự đoán liên lớp của hình ảnh loại thứ hai, như hình ảnh kiểu vải dệt thông thường. Ví dụ, giá trị mẫu liên lớp được bù sự khác biệt và/hoặc dự đoán chuyển động có thể được áp dụng, nơi sự khác biệt có thể ít nhất một phần xuất phát từ hình ảnh sâu.

Lớp tham chiếu trực tiếp có thể được xác định như lớp mà có thể được sử dụng để dự đoán liên lớp của lớp khác mà lớp này là lớp tham chiếu trực tiếp đối với nó. Lớp được dự đoán trực tiếp có thể được xác định như là lớp mà lớp khác là lớp tham chiếu trực tiếp đối với nó. Lớp tham chiếu gián tiếp có thể được xác định như là lớp mà không phải lớp tham chiếu trực tiếp của lớp thứ hai nhưng là lớp tham chiếu trực tiếp của lớp thứ ba là lớp tham chiếu trực tiếp hoặc lớp tham chiếu gián tiếp của lớp tham chiếu trực tiếp của lớp thứ hai, lớp này là lớp tham chiếu gián tiếp đối với thứ hai đó. Lớp được dự đoán gián tiếp có thể được xác định như là lớp mà đối với nó lớp khác là lớp tham chiếu gián tiếp. Lớp độc lập có thể được xác định như là lớp mà không có các lớp tham chiếu trực tiếp. Nói cách khác, lớp độc lập không được dự đoán nhờ sử dụng dự đoán liên lớp. Lớp không-cơ sở có thể được xác định như là lớp bất kỳ khác với lớp cơ sở, và lớp cơ sở có thể được xác định như là lớp thấp nhất trong dòng bit. Lớp độc lập không-cơ sở có thể được xác định như là lớp mà là cả hai lớp độc lập và lớp không-cơ sở.

Hình ảnh nguồn để dự đoán liên lớp có thể được xác định như hình ảnh được giải mã mà hoặc ở, hoặc được sử dụng để chuyển hóa, hình ảnh tham chiếu liên lớp mà có thể được sử dụng làm hình ảnh tham chiếu để dự đoán hình ảnh hiện tại. Trong các phần mở rộng HEVC nhiều-lớp, hình ảnh tham chiếu liên lớp chứa trong tập hợp hình ảnh tham chiếu liên lớp của hình ảnh hiện tại. Hình ảnh tham chiếu liên lớp có thể được xác định như là hình ảnh tham chiếu mà có thể được sử dụng để dự đoán liên lớp của hình ảnh hiện tại. Trong quy trình mã hóa và/hoặc giải mã, các hình ảnh tham

chiều liên lớp có thể được xử lý như là các hình ảnh tham chiếu dài hạn.

Hình ảnh nguồn để dự đoán liên lớp có thể được yêu cầu để ở trong cùng bộ truy nhập như là hình ảnh hiện tại. Trong một số trường hợp, ví dụ khi không lấy mẫu lại, ánh xạ trường chuyển động hoặc việc xử lý liên lớp khác là cần thiết, hình ảnh nguồn để dự đoán liên lớp và hình ảnh tham chiếu liên lớp tương ứng có thể giống nhau. Trong một số trường hợp, ví dụ khi việc lấy mẫu lại là cần thiết phù hợp mạng lưới lấy mẫu của lớp tham chiếu với mạng lưới lấy mẫu của lớp hình ảnh hiện tại (đang được mã hóa hoặc giải mã), việc xử lý liên lớp được áp dụng để chuyển hóa hình ảnh tham chiếu liên lớp từ hình ảnh nguồn để dự đoán liên lớp. Các ví dụ về xử lý liên lớp được mô tả trong các đoạn tiếp theo.

Dự đoán mẫu liên lớp có thể bao gồm việc lấy mẫu lại của (các) mảng mẫu của hình ảnh nguồn để dự đoán liên lớp. Bộ mã hóa và/hoặc bộ giải mã có thể chuyển hóa hệ số tỷ lệ theo chiều ngang hoặc (ví dụ được lưu trong biến `ScaleFactorX`) và hệ số tỷ lệ theo chiều đứng hoặc (ví dụ được lưu trong biến `ScaleFactorY`) đối với cặp lớp tăng cường và lớp tham chiếu của nó, ví dụ dựa trên các độ lệch vị trí lớp tham chiếu cho cặp này. Nếu một hoặc cả hai hệ số tỷ lệ không bằng 1, hình ảnh nguồn để dự đoán liên lớp có thể được lấy mẫu lại để tạo ra hình ảnh tham chiếu liên lớp để dự đoán hình ảnh lớp tăng cường. Quy trình và/hoặc bộ lọc được sử dụng để lấy mẫu lại có thể được xác định trước, ví dụ trong tiêu chuẩn mã hóa và/hoặc được biểu thị bởi bộ mã hóa trong dòng bit (ví dụ như là chỉ mục trong các quy trình lấy mẫu lại được xác định trước hoặc các bộ lọc) và/hoặc được giải mã bởi bộ giải mã từ dòng bit. Quy trình lấy mẫu lại khác có thể được biểu thị bởi bộ mã hóa và/hoặc được giải mã bởi bộ giải mã và/hoặc được suy ra bởi bộ mã hóa và/hoặc bộ giải mã tùy thuộc vào các giá trị của hệ số tỷ lệ. Ví dụ, khi cả hai hệ số tỷ lệ nhỏ hơn 1, quy trình hạ tần số lấy mẫu được xác định trước có thể được suy ra; và khi cả hai hệ số tỷ lệ là lớn hơn 1, quy trình tăng tần số lấy mẫu được xác định trước có thể được suy ra. Ngoài ra hoặc theo cách khác, quy trình lấy mẫu lại khác có thể được biểu thị bởi bộ mã hóa và/hoặc được giải mã bởi bộ giải mã và/hoặc được suy ra bởi bộ mã hóa và/hoặc bộ giải mã tùy thuộc vào mảng mẫu nào được xử lý. Ví dụ, quy trình lấy mẫu lại thứ nhất có thể được suy ra để được sử dụng cho các mảng mẫu luma và quy trình lấy mẫu lại thứ hai có thể được suy

ra để được sử dụng cho các mảng mẫu sắc độ.

SHVC cho phép sử dụng dự đoán trọng số hoặc quy trình ánh xạ màu dựa trên bảng tra cứu 3D (LUT) đối với (nhưng không giới hạn ở) khả năng mở rộng độ rộng của giải màu. Phương pháp tiếp cận 3D LUT có thể được mô tả như sau. Phạm vi giá trị mẫu của mỗi thành phần màu có thể trước tiên được chia ra thành hai phạm vi, tạo thành các octant $2 \times 2 \times 2$, và sau đó các phạm vi luma có thể chia thêm ra thành bốn phần, dẫn đến thành các octant $8 \times 2 \times 2$. Trong mỗi octant, mô hình tuyến tính thành phần màu cắt ngang được áp dụng để thực hiện ánh xạ màu. Đối với mỗi octant, bốn đỉnh được mã hóa thành và/hoặc được giải mã từ dòng bit để hiển thị mô hình tuyến tính trong octant. Bảng ánh xạ màu được mã hóa thành và/hoặc được giải mã từ dòng bit riêng biệt đối với mỗi thành phần màu. Ánh xạ màu có thể được coi là bao gồm ba bước: Bước thứ nhất, xác định octant mà bộ ba mẫu lớp tham chiếu (Y, Cb, Cr) thuộc về nó. Bước thứ hai, các vị trí mẫu của luma và sắc độ có thể được sắp hàng nhờ áp dụng quy trình điều chỉnh thành phần màu. Bước thứ ba, ánh xạ tuyến tính được xác định đối với octant xác định được áp dụng. Ánh xạ có thể có đặc tính thành phần-cắt ngang, tức là giá trị đầu vào của một thành phần màu có thể ảnh hưởng tới giá trị được ánh xạ của thành phần màu khác. Ngoài ra, nếu lấy mẫu lại liên lớp cũng được yêu cầu, đầu vào cho quy trình lấy mẫu lại là hình ảnh đã được ánh xạ màu. Ánh xạ có thể (nhưng không cần phải) ánh xạ các mẫu của độ sâu bit thứ nhất đến các mẫu của độ sâu bit khác.

Dự đoán chuyển động liên lớp có thể được thực hiện như sau. Quy trình dự đoán vector chuyển động tạm thời, như TMVP của H.265/HEVC, có thể được sử dụng để khai thác phần dư của dữ liệu chuyển động giữa các lớp khác nhau. Điều này có thể được thực hiện như sau: khi hình ảnh lớp cơ sở được giải mã được tăng tần số lấy mẫu, dữ liệu chuyển động của hình ảnh lớp cơ sở cũng được ánh xạ đến độ phân giải của lớp tăng cường. Nếu hình ảnh lớp tăng cường vận dụng dự đoán vector chuyển động từ hình ảnh lớp cơ sở, ví dụ với cơ chế dự đoán vector chuyển động tạm thời như TMVP của H.265/HEVC, bộ phận dự đoán vector chuyển động tương ứng có nguồn gốc từ trường chuyển động lớp cơ sở đã được ánh xạ. Theo cách này, sự liên quan giữa dữ liệu chuyển động của các lớp khác nhau có thể được khai thác để nâng cao hiệu quả

mã hóa của bộ mã hóa video mở rộng được. Trong SHVC và/hoặc tương tự, dự đoán chuyển động liên lớp có thể được thực hiện nhờ cài đặt hình ảnh tham chiếu liên lớp như là hình ảnh tham chiếu được đặt cùng chỗ đối với nguồn gốc TMVP.

Tương tự đối với MVC, trong MV-HEVC, các hình ảnh tham chiếu hình xen kẽ có thể chứa trong (các) danh mục hình ảnh tham chiếu của hình ảnh hiện tại đang được mã hóa hoặc được giải mã. SHVC sử dụng thao tác giải mã nhiều vòng lặp (không giống sự mở rộng SVC của H.264/AVC). SHVC có thể được coi là sử dụng chỉ mục tham chiếu dựa trên phương pháp tiếp cận, tức là hình ảnh tham chiếu liên lớp có thể chứa trong một hay nhiều danh mục hình ảnh tham chiếu của hình ảnh hiện tại đang được mã hóa hoặc được giải mã (như được mô tả ở trên).

Để mã hóa lớp tăng cường, các khái niệm và các công cụ mã hóa của lớp cơ sở HEVC có thể được sử dụng trong SHVC, MV-HEVC, và/hoặc tương tự. Tuy nhiên, các công cụ dự đoán liên lớp bổ sung, vốn sử dụng dữ liệu đã được mã hóa (bao gồm các mẫu hình ảnh được phục hồi và các thông tin chuyển động a.k.a thông số chuyển động) trong lớp tham chiếu để mã hóa hiệu quả lớp tăng cường, có thể được tích hợp vào SHVC, MV-HEVC, và/hoặc mã tương tự.

Tiêu chuẩn hoặc hệ thống mã hóa có thể là điểm thao tác tạm thời hoặc tương tự, nó có thể biểu thị các lớp và/hoặc các lớp-phụ mở rộng được nhờ đó việc giải mã hoạt động và/hoặc có thể được đi kèm với dòng bit phụ mà bao gồm các lớp mở rộng được và/hoặc các lớp phụ đang được giải mã. Trong HEVC, điểm thao tác được xác định là dòng bit được tạo ra từ dòng bit khác nhờ thao tác của quy trình trích xuất dòng bit phụ với dòng bit khác, TemporalId đích cao nhất, và danh mục bộ định danh lớp đích là các đầu vào.

VPS của HEVC xác định các bộ lớp và các thông số HRD cho các bộ lớp này. Bộ lớp có thể được sử dụng làm danh mục bộ định danh lớp đích trong quy trình trích xuất dòng bit phụ. Trong HEVC, bộ lớp có thể được xác định như là bộ của các lớp được hiển thị trong dòng bit được tạo ra từ dòng bit khác bởi thao tác của quy trình trích xuất dòng bit phụ với dòng bit khác, TemporalId đích cao nhất bằng 6, và danh mục bộ định danh lớp đích bằng danh mục bộ định danh lớp đi kèm với bộ lớp như là các đầu vào.

Lớp đầu ra có thể được xác định như là lớp mà các hình ảnh được giải mã của nó được xuất ra bởi quy trình giải mã. Các lớp đầu ra có thể phụ thuộc vào bộ phụ của dòng bit nhiều lớp được giải mã. Các hình ảnh xuất ra nhờ quy trình giải mã có thể được xử lý tiếp, ví dụ sự chuyển đổi không gian màu từ không gian màu YUV thành RGB có thể được thực hiện, và chúng có thể được hiển thị. Tuy nhiên, việc xử lý và/hoặc hiển thị thêm có thể được coi là các quy trình bên ngoài bộ giải mã và/hoặc quy trình giải mã và có thể không thực hiện được.

Trong các dòng bit video đa lớp, định nghĩa điểm thao tác có thể bao gồm việc xem xét bộ lớp đầu ra đích. Ví dụ, điểm thao tác có thể được xác định như dòng bit mà được tạo ra từ dòng bit khác nhờ thao tác quy trình trích xuất dòng bit phụ với dòng bit khác, lớp-phụ tạm thời đích cao nhất (ví dụ TemporalId đích cao nhất), và danh mục bộ định danh lớp đích là các đầu vào, và nó được đi kèm với bộ của các lớp đầu ra. Theo cách khác, mục khác, như là điểm thao tác đầu ra, có thể được sử dụng khi tham chiếu điểm thao tác và bộ các lớp đầu ra đi kèm. Ví dụ, trong MV-HEVC/SHVC, điểm thao tác đầu ra có thể được xác định như là dòng bit được tạo ra từ dòng bit đầu vào nhờ thao tác của quy trình trích xuất dòng bit phụ với dòng bit đầu vào, TemporalId đích cao nhất, và danh mục bộ định danh lớp đích là các đầu vào, và được đi kèm với bộ các lớp đầu ra.

Bộ lớp đầu ra (OLS) có thể được xác định như là một bộ của các lớp bao gồm các lớp của một trong số các bộ lớp được xác định, nơi một hay nhiều lớp trong bộ các lớp được biểu thị để là các lớp đầu ra. Lớp đầu ra có thể được xác định là lớp của bộ lớp đầu ra mà là đầu ra khi bộ giải mã và/hoặc HRD hoạt động nhờ sử dụng bộ lớp đầu ra như là bộ lớp đầu ra đích. Trong MV-HEVC/SHVC, biến TargetOlsIdx có thể xác định bộ lớp đầu ra nào là bộ lớp đầu ra đích nhờ thiết lập TargetOlsIdx bằng chỉ mục của bộ lớp đầu ra là bộ lớp đầu ra đích. Bộ lớp đầu ra đích có thể được xác định là bộ lớp đầu ra mà đối với nó chỉ mục bằng TargetOlsIdx. TargetOlsIdx có thể được thiết lập, ví dụ bởi HRD và/hoặc có thể được thiết lập bởi phương tiện bên ngoài, ví dụ bởi bộ phát hoặc tương tự thông qua giao diện được tạo ra bởi bộ giải mã. Trong MV-HEVC/SHVC, lớp đầu ra có thể được xác định là lớp của bộ lớp đầu ra là đầu ra khi TargetOlsIdx bằng chỉ mục của bộ lớp đầu ra.

Bộ phát, cổng vào, máy khách, hoặc tương tự có thể chọn các lớp được biến đổi và/hoặc các lớp phụ của dòng bit video mở rộng được. Việc trích xuất lớp tạm thời, trích xuất các lớp, hoặc chuyển lớp xuống có thể là các lớp ít biến đổi hơn so với lớp khả dụng trong dòng bit được nhận bởi bộ phát, cổng vào, máy khách, hoặc tương tự. Việc chuyển lớp lên có thể là (các) lớp biến đổi bổ sung so với các lớp được biến đổi trước khi chuyển lớp lên bởi bộ phát, cổng vào, máy khách, hoặc tương tự, tức là bắt đầu lại sự biến đổi của một hay nhiều lớp mà việc biến đổi của nó đã ngừng sớm hơn khi chuyển lớp xuống. Tương tự việc chuyển lớp xuống và/hoặc chuyển lớp lên, bộ phát, cổng vào, máy khách, hoặc tương tự có thể thực hiện chuyển-xuống và/hoặc chuyển-lên của lớp-phụ tạm thời. Bộ phát, cổng vào, máy khách, hoặc tương tự cũng có thể thực hiện cả lớp và lớp-phụ chuyển-xuống và/hoặc chuyển-lên. Lớp và lớp-phụ chuyển-xuống và/hoặc chuyển-lên có thể được thực hiện trong cùng một bộ truy nhập hoặc tương tự (tức là hầu như cùng lúc) hoặc có thể được thực hiện trong các bộ truy nhập khác nhau hoặc tương tự (tức là hầu như tại các thời gian khác biệt).

Trong khi bộ không đổi các lớp đầu ra phù hợp đối với các trường hợp sử dụng và các dòng bit nơi lớp cao nhất giữ nguyên không thay đổi trong mỗi bộ truy nhập, chúng có thể không hỗ trợ các trường hợp sử dụng nơi mà lớp cao nhất thay đổi từ một bộ truy nhập sang bộ khác. Do đó, đã được đề xuất là các bộ mã hóa có thể xác định việc sử dụng các lớp đầu ra thay thế trong dòng bit và để đáp ứng cho việc sử dụng riêng các lớp đầu ra thay thế, các bộ giải mã xuất ra hình ảnh được giải mã từ lớp đầu ra thay thế khi vắng mặt hình ảnh trong lớp đầu ra trong cùng bộ truy nhập. Tồn tại một vài khả năng làm cách nào để biểu thị các lớp đầu ra thay thế. Ví dụ, như được xác định trong HEVC, cơ chế bộ lớp đầu ra thay thế có thể bị hạn chế để được sử dụng chỉ dành cho các bộ lớp đầu ra chứa chỉ một lớp đầu ra, và cờ output-layer-set-wise (`alt_output_layer_flag [ olsIdx ]` trong HEVC) có thể được sử dụng để xác định là bất kỳ lớp tham chiếu trực tiếp hoặc gián tiếp của lớp đầu ra có thể phục vụ như là lớp đầu ra thay thế đối với lớp đầu ra của bộ lớp đầu ra. Khi có nhiều hơn một lớp đầu ra thay thế được phép sử dụng, nó có thể được xác định là hình ảnh tham chiếu liên lớp trực tiếp hoặc gián tiếp thứ nhất có mặt trong bộ truy nhập theo trình tự của bộ định danh lớp giảm dần xuống đến lớp đầu ra thay thế cực tiểu được biểu thị là đầu ra.

Mẫu theo các mã nhập mẫu được xác định có mặt trong ISO/IEC 14496-15 (ví dụ 'hvc1', 'hev1') bao gồm chuỗi các bộ NAL không cần mã khởi đầu, mỗi bộ được tiếp đầu ngữ bởi cách tính byte của chúng. Độ dài của trường đếm byte có thể được biểu thị trong mục nhập mẫu.

Các loại phương tiện internet, còn được biết là các loại MIME (Multipurpose Internet Mail Extension), được sử dụng bởi các ứng dụng khác nhau để nhận dạng loại nguồn hoặc tệp. Các loại MIME bao gồm loại phương tiện, phân loại, và 0 hoặc nhiều thông số tùy ý.

Như được mô tả, MIME là sự mở rộng đối với giao thức email vốn làm cho việc biến đổi và nhận các loại tệp dữ liệu khác nhau trên internet có thể thực hiện được, ví dụ video và audio, các hình ảnh, phần mềm, v. v.. Loại phương tiện internet là bộ định danh được sử dụng trên internet để biểu thị loại dữ liệu mà tệp chứa. Các loại phương tiện internet như vậy có thể cũng được gọi các loại nội dung. Một số kết hợp loại/phân loại MIME tồn tại mà có thể biểu thị các định dạng truyền thông khác nhau. Thông tin loại nội dung có thể được bao gồm nhờ thực thể biến đổi trong phần đầu MIME tại lúc bắt đầu biến đổi phương tiện. Do đó, thực thể nhận có thể cần thẩm định chi tiết của nội dung phương tiện để xác định nếu các phần tử cụ thể có thể kết xuất cho bộ mã khả dụng. Đặc biệt, khi hệ thống đầu cuối có nguồn giới hạn, hoặc sự kết nối vào các hệ thống đầu cuối có chiều rộng dải bị giới hạn, nó có thể hữu ích để biết từ một mình loại nội dung loại nếu nội dung có thể được kết xuất.

Hai thông số, 'các bộ mã hóa-giải mã' và 'các hồ sơ', được xác định để được sử dụng với các loại MIME khác nhau hoặc các kết hợp loại/phân loại để cho phép thông số kỹ thuật rõ ràng của các bộ mã hóa-giải mã được sử dụng bởi các định dạng phương tiện chứa bên trong, hoặc (các) hồ sơ của toàn bộ định dạng vật chứa.

Bằng cách ghi nhãn nội dung bằng các mã riêng được biểu thị để kết xuất phương tiện chứa trong, các hệ thống nhận có thể xác định nếu các mã được hỗ trợ bởi hệ thống đầu cuối, và nếu không, có thể ra hành động thích hợp (như từ chối nội dung, gửi thông báo về tình trạng, chuyển mã nội dung sang loại được hỗ trợ, tìm nạp và cài đặt các mã được yêu cầu, kiểm tra thêm để xác định nếu nó là đủ để hỗ trợ bộ con của các mã được chỉ ra, v.v.). Đối với các định dạng tệp xuất phát từ ISO/BMFF, các thông

số mã có thể được coi là bao gồm danh mục được tách bằng dấu phẩy của một hay nhiều mục danh mục. Khi mục của danh mục của các thông số mã hiển thị đường ghi của tệp tương thích ISOBMFF, mục của danh mục có thể bao gồm mã bốn ký tự của mục nhập mẫu của đường ghi. Mục của danh mục có thể biểu thị thêm hồ sơ mã hóa (ví dụ hồ sơ chính HEVC) mà dòng bit của nó được chứa bởi các mẫu của đường ghi phù hợp.

Các thông số hồ sơ MIME có thể tạo ra dấu hiệu tổng thể, đối với bộ thu, của các thông số kỹ thuật mà nội dung phù hợp với nó. Đó là dấu hiệu khả năng tương thích của định dạng vật chứa và các nội dung của nó đối với một số thông số kỹ thuật. Bộ thu có thể hoạt động bên ngoài mức độ mà nó có thể xử lý và kết xuất vào đó những nội dung bằng cách thăm định để nhìn thấy hồ sơ nào trong số các hồ sơ được khai báo được hỗ trợ, và chúng là gì. Thông số của các hồ sơ đối với tệp ISOBMFF có thể được xác định để bao gồm danh mục của các nhãn tương thích chứa trong tệp này.

Bộ định danh nguồn đồng nhất (URI) có thể được xác định như là chuỗi của các ký tự được sử dụng để nhận dạng tên nguồn. Việc nhận dạng như vậy cho phép tương tác với các hiển thị của nguồn đối với mạng, nhờ sử dụng các giao thức riêng biệt. URI được xác định thông qua hệ thống xác định cú pháp cụ thể và giao thức kèm theo đối với URI. Bộ định vị tài nguyên đồng nhất (URL) và tên nguồn đồng nhất (URN) là các dạng URI. URL có thể được xác định như là URI mà nhận dạng nguồn web và xác định phương tiện hành động theo hoặc thu được hiển thị của nguồn, xác định cả cơ chế truy nhập cơ bản của nó và vị trí mạng. URN có thể được xác định như là URI mà nhận dạng nguồn bởi tên trong không gian tên cụ thể. URN có thể được sử dụng để nhận diện nguồn mà không cần bao hàm vị trí của nó hoặc cách truy nhập nó.

Giao thức truyền siêu văn bản (Hypertext Transfer Protocol - HTTP) đã được sử dụng rộng rãi cho việc chuyển nội dung đa truyền thông thời gian thực đối với internet, như trong các ứng dụng phát video.

Một số giải pháp thương mại để phát trực tuyến đối với HTTP, như Microsoft® Smooth Streaming, Apple® Adaptive HTTP Live Streaming và Adobe® Dynamic Streaming, đã được giới thiệu cũng như các dự án tiêu chuẩn hóa đã được thực hiện. Adaptive HTTP streaming (AHS) đã được tiêu chuẩn hóa lần đầu tiên trong

Phiên bản 9 của dịch vụ phát trực tuyến chuyên-gói (PSS) của Dự án Hợp tác Thế hệ thứ 3 (3GPP) (3GPP TS 26.234 Phiên bản 9: “Transparent end-to-end packet-switched streaming service (PSS); các giao thức và bộ mã hóa-giải mã”). MPEG lấy 3GPP AHS Phiên bản 9 như là điểm khởi đầu đối với tiêu chuẩn MPEG DASH (ISO/IEC 23009-1: “Dong thích ứng động trên HTTP (DASH)-phần 1: mô tả trình diễn phương tiện và các phân đoạn”, tiêu chuẩn quốc tế, phiên bản thứ 2, 2014). MPEG DASH và 3GP-DASH gần gũi với nhau về mặt kỹ thuật và do đó có thể gọi chung là DASH. Một số các khái niệm, các định dạng, và các thao tác của DASH được mô tả dưới đây làm ví dụ về hệ thống phát trực tuyến video, trong đó các phương án có thể được sử dụng. Các khía cạnh của sáng chế không bị giới hạn ở DASH, mà đúng hơn là sự mô tả được đưa ra với một căn cứ khả thi mà dựa trên đó sáng chế có thể được thực hiện một phần hoặc toàn bộ.

Trong DASH, nội dung đa phương tiện có thể được lưu trên máy chủ HTTP và có thể được phân phối nhờ sử dụng HTTP. Nội dung có thể được lưu trên máy chủ theo hai phần: Phần mô tả trình bày truyền thông (MPD), vốn mô tả sự biểu thị của nội dung khả dụng, các nội dung thay thế khác nhau của nó, các địa chỉ URL của chúng, và các đặc tính khác; và các phân đoạn, vốn chứa các dòng bit đa phương tiện thực tế dưới dạng các mảnh, trong các tệp đơn hoặc nhiều tệp. MPD đề xuất thông tin cần thiết cho các máy khách để thiết lập dòng thích ứng động đối với HTTP. MPD chứa thông tin mô tả trình bày phương tiện, như HTTP- bộ định vị tài nguyên đồng nhất (URL) của mỗi phân đoạn để tạo ra phân đoạn yêu cầu GET. Để phát nội dung, máy khách DASH có thể thu được MPD ví dụ nhờ sử dụng HTTP, email, ổ USB, phát sóng, hoặc các phương pháp chuyển khác. Bằng cách phân tích cú pháp MPD, máy khách DASH có thể trở nên nhận thức được việc định giờ chương trình, tính khả dụng của nội dung phương tiện, các loại phương tiện, độ phân giải, các băng thông cực tiểu và cực đại, và sự tồn tại của các thay thế được mã hóa khác nhau của các thành phần đa phương tiện, các đặc tính về khả năng truy cập và yêu cầu về quản lý bản quyền kỹ thuật số (DRM), các vị trí thành phần phương tiện trên mạng, và các đặc tính về nội dung khác. Nhờ sử dụng thông tin này, máy khách DASH có thể lựa chọn các thay thế được mã hóa thích hợp và bắt đầu phát trực tuyến nội dung bằng cách tìm nạp các

phân đoạn nhờ sử dụng ví dụ các yêu cầu HTTP GET. Sau khi đệm thích hợp để cho phép các biến thể thông lượng mạng, máy khách có thể tiếp tục tìm nạp các phân đoạn tiếp theo và cũng giám sát các biến động băng thông mạng. Máy khách có thể quyết định cách thích nghi với băng thông khả dụng bằng cách tìm nạp các phân đoạn của các lựa chọn thay thế khác nhau (với các tốc độ bit thấp hơn hoặc cao hơn) để duy trì bộ đệm đầy đủ.

Trong DASH, mô hình dữ liệu phân cấp được sử dụng để cấu trúc trình bày phương tiện như được thể hiện trong Fig.5. Sự trình diễn phương tiện bao gồm chuỗi của một hay nhiều giai đoạn, mỗi giai đoạn chứa một hay nhiều nhóm, mỗi nhóm chứa một hay nhiều bộ thích ứng, mỗi bộ thích ứng chứa một hay nhiều sự trình diễn, mỗi sự trình diễn bao gồm một hay nhiều phân đoạn. Sự trình diễn một trong các lựa chọn thay thế của nội dung phương tiện hoặc tập hợp con của nó thường khác nhau bởi lựa chọn mã hóa, ví dụ bởi tốc độ bit, độ phân giải, ngôn ngữ, mã, v. v.. Phân đoạn chứa thời lượng nhất định dữ liệu đa phương tiện, và siêu dữ liệu để mã hóa và trình bày nội dung phương tiện được chứa. Phân đoạn được xác định bởi URI và có thể thường được yêu cầu bởi HTTP GET yêu cầu. Phân đoạn có thể được xác định như là bộ dữ liệu đi kèm với HTTP-URL và tùy ý phạm vi byte mà được xác định bởi MPD.

DASH MPD tuân theo Ngôn ngữ đánh dấu mở rộng (Extensible Markup Language XML) và do đó được xác định thông qua các phần tử và thuộc tính như được xác định trong XML. MPD có thể được xác định nhờ sử dụng các quy ước sau: Các phần tử trong tài liệu XML có thể được xác định bởi chữ các thứ nhất viết hoa và có thể được in đậm như **Element**. Để biểu thị là phần tử **Element 1** được chứa trong phần tử khác, **Element 2**, một cách có thể viết là **Element 2.Element 1**. Nếu tên phần tử bao gồm hai hoặc nhiều từ được kết hợp, camel-casing (dạng lạc đà) có thể được sử dụng, ví dụ **ImportantElement**. Các phần tử có thể có mặt hoặc chính xác một, hoặc sự xuất hiện tối thiểu và tối đa có thể được xác định bởi <minOccurs> ... <maxOccurs>. Các thuộc tính trong tài liệu XML có thể được xác định bởi chữ cái đầu viết hoa cũng như chúng có thể được đứng trước bởi ký hiệu '@', ví dụ thuộc tính @. Để trỏ đến thuộc tính riêng thuộc tính @ được chứa trong phần tử **Element**, một có thể viết thuộc tính @**Element**. Nếu tên của thuộc tính bao gồm hai hoặc nhiều từ được kết

hợp, camel-casing có thể được sử dụng sau từ thứ nhất, ví dụ `@veryImportantAttribute`. Các thuộc tính có thể có được chuyển giao trạng thái trong XML như là (M) bắt buộc, (O) tùy ý, tùy ý với giá trị mặc định (OD) và bắt buộc có điều kiện (CM).

Trong DASH, sự biểu diễn độc lập có thể được xác định như là sự biểu diễn mà có thể được xử lý độc lập của bất kỳ các trình diễn khác. Sự biểu diễn độc lập có thể được hiểu là bao gồm dòng bit độc lập hoặc lớp độc lập của dòng bit. Sự biểu diễn phụ thuộc có thể được xác định như là sự biểu diễn mà đối với nó các phân đoạn từ các sự biểu diễn bổ sung của nó cần thiết cho việc trình diễn và/hoặc giải mã các thành phần nội dung phương tiện được chứa. Sự biểu diễn phụ thuộc có thể được hiểu là bao gồm ví dụ lớp được dự đoán của dòng bit mở rộng được. Sự biểu diễn bổ sung có thể được xác định như là sự biểu diễn mà bổ sung ít nhất một sự biểu diễn phụ thuộc. Sự biểu diễn bổ sung có thể là sự biểu diễn độc lập hoặc sự biểu diễn phụ thuộc. Các sự biểu diễn phụ thuộc có thể được mô tả bởi phần tử **Representation** mà chứa thuộc tính `@dependencyId`. Các sự biểu diễn phụ thuộc có thể được coi như là các sự biểu diễn thường xuyên ngoại trừ là chúng phụ thuộc vào tập hợp của các sự biểu diễn bổ sung để giải mã và/hoặc trình diễn. `@dependencyId` chứa các giá trị của `@id` attribute của tất cả các sự biểu diễn bổ sung, tức là các sự biểu diễn mà cần thiết để trình diễn và/hoặc giải mã các thành phần nội dung phương tiện được chứa trong các sự biểu diễn phụ thuộc này.

Trong bối cảnh DASH, các định nghĩa sau có thể được sử dụng: thành phần nội dung phương tiện hoặc thành phần phương tiện có thể được xác định như là một thành phần liên tiếp của nội dung phương tiện với loại thành phần phương tiện được giao mà có thể được mã hóa riêng lẻ thành dòng phương tiện. Nội dung phương tiện có thể được xác định như là một giai đoạn nội dung phương tiện hoặc chuỗi liên tục của các giai đoạn nội dung phương tiện. Loại thành phần nội dung phương tiện có thể được xác định như là loại đơn của nội dung phương tiện như audio, video, hoặc văn bản. Dòng phương tiện có thể được xác định như là phiên bản được mã hóa của thành phần nội dung phương tiện.

Phân đoạn khởi tạo có thể được xác định như là phân đoạn chứa siêu dữ liệu

mà cần thiết hiển thị các dòng phương tiện được gói trong các phân đoạn phương tiện. Trong các định dạng phân đoạn dựa trên ISOBMFF, Phân đoạn khởi tạo có thể bao gồm Hộp phim ('moov') mà có thể không bao gồm siêu dữ liệu cho các mẫu bất kỳ, tức là siêu dữ liệu bất kỳ cho các mẫu được cung cấp trong các hộp 'moof'.

Phân đoạn phương tiện chứa thời lượng nhất định của dữ liệu đa phương tiện để phát lại ở tốc độ thường, thời lượng này được gọi là thời lượng phân đoạn phương tiện hoặc thời lượng phân đoạn. Nhà sản xuất nội dung hoặc nhà cung cấp dịch vụ có thể chọn thời lượng phân đoạn theo các đặc tính mong muốn của dịch vụ. Ví dụ, thời lượng phân đoạn tương đối ngắn có thể được sử dụng trong dịch vụ trực tiếp (live service) để đạt được độ trễ end-to-end ngắn. Nguyên nhân là thời lượng phân đoạn thường là ranh giới thấp hơn đối với độ trễ end-to-end được nhận ra bởi máy khách DASH vì phân đoạn là bộ rời rạc của dữ liệu đa phương tiện tạo ra cho DASH. Việc tạo ra nội dung thường được thực hiện theo cách mà toàn bộ phân đoạn của dữ liệu đa phương tiện được làm cho khả dụng đối với máy chủ. Ngoài ra, nhiều ứng dụng của máy khách sử dụng phân đoạn như là một bộ cho Các yêu cầu GET. Do đó, trong các bố trí thông thường cho các dịch vụ trực tiếp (live services) phân đoạn có thể được yêu cầu bởi máy khách DASH chỉ khi toàn bộ thời lượng của phân đoạn phương tiện là khả dụng cũng như được mã hóa và được gói thành phân đoạn. Đối với dịch vụ theo yêu cầu, có thể sử dụng các chính sách khác nhau để lựa chọn thời lượng phân đoạn.

Phân đoạn có thể được phân chia thêm thành các phân đoạn con ví dụ để cho phép tải các phân đoạn trong các phần. Các phân đoạn con có thể được yêu cầu để chứa các bộ truy nhập đầy đủ. Các phân đoạn con có thể được lập chỉ mục bởi hộp Chỉ mục phân đoạn, mà chứa thông tin để ánh xạ phạm vi thời gian trình bày và phạm vi byte cho mỗi phân đoạn con. Hộp Chỉ mục phân đoạn cũng có thể mô tả các phân đoạn con và các điểm truy nhập dòng trong phân đoạn nhờ báo hiệu các thời lượng của chúng và bù đắp byte. Máy khách DASH có thể sử dụng thông tin thu được từ các Hộp Chỉ mục phân đoạn để tạo ra HTTP GET yêu cầu cho phân đoạn con riêng nhờ sử dụng yêu cầu HTTP phạm vi byte. Nếu sử dụng thời lượng phân đoạn tương đối dài, sau đó các phân đoạn con có thể được sử dụng để giữ kích thước của các đáp ứng HTTP hợp lý và linh hoạt cho sự thích nghi tốc độ bit. Thông tin chỉ mục của phân

đoạn có thể đặt trong hộp đơn tại lúc bắt đầu phân đoạn đó, hoặc trải ra trong nhiều hộp chỉ mục trong phân đoạn. Các phương pháp trải khác nhau đều có thể, như phân cấp, chuỗi daisy, và hỗn hợp. Kỹ thuật này có thể tránh bỏ sung hộp lớn tại lúc bắt đầu phân đoạn và do đó có thể ngăn ngừa việc tải xuống chậm có thể lúc đầu.

Các hiển thị phụ được nhúng trong các hiển thị thường xuyên và được mô tả bởi phần tử **SubRepresentation**. Các phần tử **SubRepresentation** chứa trong phần tử **Representation**. Phần tử **SubRepresentation** mô tả các đặc tính của một hoặc vài thành phần nội dung phương tiện mà được nhúng trong hiển thị. Nó có thể ví dụ mô tả chính xác các đặc tính của thành phần audio được nhúng (ví dụ, mã, tỷ lệ lấy mẫu, v.v.), tiêu đề phụ được nhúng (ví dụ, mã) hoặc nó có thể mô tả một số lớp video chất lượng thấp hơn được nhúng (ví dụ một số tỷ lệ khung hình thấp hơn, v.v.). Các hiển thị phụ và hiển thị chia sẻ một số thuộc tính chung và các phần tử. Trong trường hợp thuộc tính mức @ có mặt trong phần tử **SubRepresentation**, áp dụng các bước sau:

- Các hiển thị phụ đề xuất khả năng truy cập phiên bản chất lượng thấp hơn hiển thị mà chúng được bao hàm trong đó. Trong trường hợp này, các hiển thị phụ ví dụ cho phép trích xuất đường ghi audio trong hiển thị ghép kênh hoặc có thể cho phép đối với các thao tác tua nhanh hoặc tua lại hiệu quả nếu được cung cấp với tỷ lệ khung hình thấp hơn;

- Phân đoạn khởi tạo và/hoặc các phân đoạn phương tiện và/hoặc các phân đoạn chỉ mục sẽ đề xuất đủ thông tin như vậy mà dữ liệu có thể được truy cập dễ dàng thông qua các yêu cầu GET một phần HTTP. Các chi tiết về đề xuất thông tin này được xác định bởi định dạng phương tiện khi sử dụng.

- Khi sử dụng các phân đoạn ISOBMFF, các bước sau được áp dụng:

- Phân đoạn khởi tạo chứa hộp chuyển cấp.
- Hộp Chỉ mục phân đoạn con ('ssix') có mặt cho mỗi phân đoạn con.
- Cấp thuộc tính @ xác định cấp mà hiển thị phụ được mô tả được kèm theo vào Chỉ mục phân đoạn con. Thông tin trong hiển thị, hiển thị phụ và trong hộp Chuyển cấp ('leva') chứa thông tin về việc chuyển giao dữ liệu đa phương tiện đến các cấp.

- Dữ liệu đa phương tiện phải có trình tự như vậy mà mỗi cấp đưa ra sự tăng cường so với các cấp thấp hơn.

Nếu thuộc tính cấp @ vắng mặt, sau đó phần tử hiển thị phụ chỉ được sử dụng để đưa ra sự mô tả chi tiết hơn đối với các dòng phương tiện vốn được nhúng trong hiển thị.

ISOBMFF bao gồm cơ cấu được gọi là cấp xác định tập hợp con của tệp. Các cấp tiếp theo phân cấp phụ thuộc sao cho các mẫu được ánh xạ đến mức n có thể phụ thuộc vào các mẫu bất kỳ của các mức m , trong đó $m \leq n$, và không phụ thuộc vào các mẫu bất kỳ của các mức p , trong đó $p > n$. Ví dụ, các cấp có thể được xác định theo lớp-phụ tạm thời (ví dụ, `temporal_id` của SVC hoặc MVC hoặc `TemporalId` của HEVC). Các cấp có thể được công bố trong hộp Chuyển cấp ('leva') chứa trong hộp Phim mở rộng ('mvex'). Các cấp không thể xác định được đối với phim nguyên bản. Khi hộp Chuyển cấp có mặt, nó áp dụng cho tất cả các đoạn phim tiếp theo phim nguyên bản. Đối với hộp Chuyển cấp, phân đoạn được xác định để bao gồm một hay nhiều hộp đoạn phim và các hộp dữ liệu đa phương tiện đi kèm, có thể chỉ bao gồm phần ban đầu phần của hộp dữ liệu đa phương tiện cuối cùng. Trong phân đoạn, dữ liệu cho mỗi cấp xuất hiện liên tục. Dữ liệu cho các cấp trong phân đoạn xuất hiện theo trình tự tăng dần của giá trị cấp. Tất cả dữ liệu trong phân đoạn sẽ được đánh dấu vào các cấp. Hộp Chuyển cấp tạo ra ánh xạ từ các tính năng, như các lớp có thể mở rộng hoặc các lớp phụ tạm thời, đến các cấp. Tính năng có thể được xác định nhờ đường ghi, đường ghi phụ trong đường ghi, hoặc việc nhóm mẫu của đường ghi. Ví dụ, việc nhóm mẫu cấp tạm thời có thể được sử dụng để biểu thị ánh xạ của các hình ảnh đến các cấp tạm thời, vốn tương đương với lớp-phụ tạm thời trong HEVC. Tức là, các hình ảnh HEVC của giá trị `TemporalId` nhất định có thể được ánh xạ đến cấp tạm thời cụ thể nhờ sử dụng việc nhóm mẫu Cấp tạm thời (và cũng có thể lặp lại đối với tất cả các giá trị `TemporalId`). Sau đó hộp Chuyển cấp có thể tham chiếu nhóm mẫu Cấp tạm thời trong ánh xạ được biểu thị đến các cấp. Hộp Chuyển cấp bao gồm phần tử cú pháp `padding_flag`. `padding_flag` bằng 1 biểu thị phân đoạn phù hợp mà có thể được tạo thành bởi việc nối lại một số nguyên dương bất kỳ của các cấp trong phân đoạn và thêm vào hộp dữ liệu đa phương tiện cuối cùng bởi các byte không đến kích thước đầy

đủ mà được biểu thị trong phần khởi đầu của hộp Dữ liệu đa phương tiện cuối cùng. Ví dụ, `padding_flag` có thể là tập hợp bằng 1 khi mỗi phân đoạn chứa hai hoặc nhiều đường ghi AVC, SVC, hoặc MVC của cùng dòng bit video, các mẫu đối với mỗi đường ghi của phân đoạn là gần kề và theo trình tự giải mã trong hộp dữ liệu đa phương tiện, và các mẫu của cấp AVC, SVC, hoặc MVC đầu tiên chứa các bộ trích xuất NAL để bao gồm các bộ NAL mã hóa video từ các cấp khác của cùng phân đoạn.

Hộp Chỉ mục phân đoạn con ('ssix') tạo ra ánh xạ từ các cấp (như được xác định bởi Hộp Chuyển cấp) cho các phạm vi byte của phân đoạn con được tạo mục. Nói cách khác, hộp này tạo ra chỉ mục nhỏ gọn để dữ liệu trong phân đoạn con được sắp xếp theo các cấp thành các phân đoạn con một phần. Nó cho phép máy khách truy cập dữ liệu dễ dàng đối với các phân đoạn con một phần nhờ việc tải các phạm vi của dữ liệu trong phân đoạn con. Khi hộp Chỉ mục phân đoạn có mặt, mỗi byte trong phân đoạn con được chuyển cấp. Nếu phạm vi này không đi kèm với thông tin nào đó khi chuyển cấp, sau đó có thể sử dụng cấp bất kỳ mà không trong chuyển cấp. Đó là các hộp Chỉ mục phân đoạn con 0 hoặc 1 có mặt cho mỗi hộp Chỉ mục phân đoạn mà chỉ lập chỉ mục các phân đoạn con, tức là chỉ lập chỉ mục các phân đoạn con nhưng không lập chỉ mục phân đoạn. Hộp Chỉ mục phân đoạn, nếu có, là hộp tiếp theo sau hộp Chỉ mục phân đoạn đi kèm. Hộp Chỉ mục phân đoạn dẫn chứng phân đoạn con mà được biểu thị trong hộp Chỉ mục phân đoạn đứng ngay trước. Mỗi cấp có thể được chuyển đến chính xác một phần phân đoạn con, tức là các phạm vi byte đối với một cấp là kề nhau. Các cấp của các phần phân đoạn con được chuyển bởi các số tăng dần trong phân đoạn con, tức là, các mẫu của phần phân đoạn con có thể phụ thuộc vào các mẫu bất kỳ của các phần phân đoạn con đứng trước trong cùng phân đoạn con, nhưng không phải là cách khác xung quanh. Ví dụ, mỗi phần phân đoạn con chứa các mẫu có lớp-phụ tạm thời giống nhau và các phần phân đoạn con xuất hiện trong trình tự lớp-phụ tạm thời tăng lên trong phân đoạn con đó. Khi một phần phân đoạn con được truy cập theo cách này, hộp dữ liệu đa phương tiện cuối cùng có thể không hoàn thành, tức là, dữ liệu được truy cập ít hơn chỉ thị độ dài của các chỉ thị hộp dữ liệu đa phương tiện biểu thị có mặt. Độ dài của hộp dữ liệu đa phương tiện có thể cần điều chỉnh, hoặc có thể sử dụng phần đệm. `padding_flag` trong hộp Chuyển cấp biểu thị xem việc mất dữ liệu này

có thể được thay thế bởi các bit 0 hay không. Nếu không, dữ liệu mẫu cho các mẫu được chuyển tới các cấp mà không được truy cập là không có mặt, và nên chăm sóc.

Có thể được yêu cầu là đối với các hiển thị phụ thuộc X bất kỳ mà phụ thuộc vào hiển thị bổ sung Y, phân đoạn con thứ - m của X và phân đoạn con thứ- n của Y sẽ không chồng chéo ngay cả khi m không bằng n . Có thể được yêu cầu là đối với các hiển thị phụ thuộc việc ghép nối của phân đoạn khởi tạo với chuỗi các phân đoạn con của các hiển thị phụ thuộc, mỗi cái được đứng trước bởi phân đoạn con tương ứng của mỗi hiển thị bổ sung như được cung cấp về thuộc tính @dependencyId sẽ đại diện cho chuỗi phân đoạn con phù hợp với định dạng phương tiện như được xác định trong thuộc tính loại @mime đối với các hiển thị phụ thuộc này.

MPEG-DASH xác định các định dạng chứa phân đoạn cho cả hai dòng chuyển ISO BMFF và MPEG-2. Các thông số kỹ thuật khác có thể xác định các định dạng của phân đoạn dựa trên các định dạng của vật chứa khác. Ví dụ, định dạng phân đoạn dựa trên tệp định dạng vật chứa Matroska đã được đề xuất và có thể được tóm tắt như sau. Khi các tệp Matroska được mang theo như là các phân đoạn DASH hoặc tương tự, sự kết hợp của các bộ DASH và các bộ Matroska có thể được xác định như sau. Phân đoạn con (của DASH) có thể được xác định như là một hay nhiều cụm liên tục của nội dung Matroska-được gói. Phân đoạn khởi tạo của DASH có thể được yêu cầu để bao gồm phần đầu EBML, phần đầu phân đoạn (của Matroska), thông tin phân đoạn (của Matroska) và đường ghi, và có thể tùy ý bao gồm các phần tử cấp 1 khác và đệm. Chỉ mục của phân đoạn DASH có thể bao gồm phần tử tín hiệu của Matroska.

SRD được xác định trong Phụ lục quy tắc H của MPEG-DASH. Phần sau đây chứa một số trích đoạn của thông số kỹ thuật SRD.

SRD hệ thống cho phép các tác giả bản mô tả trình bày phương tiện biểu diễn các mối quan hệ về không gian giữa các đối tượng không gian. Đối tượng không gian được thể hiện bởi hoặc là Bộ thích ứng hoặc hiển thị phụ. Ví dụ, mối quan hệ về không gian có thể được biểu diễn là video trình bày một phần không gian của video toàn khung khác (ví dụ vùng quan tâm, hoặc ô).

Các bộ mô tả SupplementalProperty và/hoặc EssentialProperty với @schemeIdUri bằng "urn:mpeg:dash:srd:2014" được sử dụng để tạo ra thông tin mỗi

quan hệ không gian đi kèm với việc chứa đối tượng không gian. SRD sẽ được chứa riêng trong hai phần tử MPD này (Bộ thích ứng và hiển thị phụ).

Để giữ gìn khả năng tương thích với các máy khách cũ, MPD đòi hỏi sử dụng SupplementalProperty và EssentialProperty theo cách là ít nhất một hiển thị có thể được dẫn giải bởi các máy khách cũ sau khi loại bỏ phần tử chứa EssentialProperty.

Các SRD cấp hiển thị phụ có thể được sử dụng để trình bày các đối tượng không gian trong một hiển thị như các dòng xếp HEVC. Trong trường hợp này, các bộ mô tả SRD có thể có mặt ở Bộ thích ứng cũng như các cấp hiển thị phụ.

Giá trị @ của các phần tử SupplementalProperty hoặc EssentialProperty nhờ sử dụng SRD hệ thống là danh mục được tách bằng dấu phẩy của các giá trị đối với các thông số SRD. Các thông số SRD source_id, object_x, object_y, object_width, và object_height được yêu cầu có mặt và các thông số SRD total_width, total_height, và spatial_set_id có mặt có điều kiện hoặc tùy ý.

source_id là số nguyên không âm trong hiển thị thập phân tạo ra bộ định danh cho nguồn nội dung. Thông số source_id tạo ra bộ định danh duy nhất, trong giai đoạn này, cho nguồn nội dung. Nó hàm ý định nghĩa hệ tọa độ đi kèm với nguồn này. Hệ tọa độ này sử dụng gốc tùy ý (0; 0); trục-x hướng từ trái sang phải và trục-y từ trên xuống dưới. Tất cả SRD chia sẻ cùng giá trị source_id có cùng gốc và các hướng trục. Các mối quan hệ về không gian đối với các đối tượng không gian nhờ sử dụng SRD với các giá trị source_id khác nhau không được xác định.

Đối với các giá trị source_id đã cho, không gian tham chiếu được xác định, tương ứng với vùng hình chữ nhật bao gồm toàn bộ nguồn nội dung, mà góc trên cùng-bên trái của nó là gốc của hệ tọa độ. Các giá trị total_width và total_height trong SRD tạo ra kích thước của không gian tham chiếu này được biểu diễn theo các đơn vị tùy ý. total_width là số nguyên không âm trong hiển thị thập phân biểu diễn chiều rộng của không gian tham chiếu theo các đơn vị tùy ý. total_height là số nguyên không âm trong hiển thị thập phân biểu diễn chiều cao của không gian tham chiếu theo các đơn vị tùy ý. Được phép không có đối tượng không gian trong MPD mà bao gồm toàn bộ nguồn nội dung, ví dụ khi toàn bộ nguồn nội dung được thể hiện bởi hai video tách rời.

`object_x` là số nguyên không âm trong hiển thị thập phân biểu diễn vị trí theo chiều ngang của góc trên cùng-bên trái của đối tượng không gian theo các đơn vị tùy ý. `object_y` là số nguyên không âm trong hiển thị thập phân biểu diễn vị trí theo chiều đứng của góc trên cùng-bên trái của đối tượng không gian theo các đơn vị tùy ý. `object_width` là số nguyên không âm trong hiển thị thập phân biểu diễn chiều rộng của đối tượng không gian theo các đơn vị tùy ý. `object_height` là số nguyên không âm trong hiển thị thập phân biểu diễn chiều cao của đối tượng không gian theo các đơn vị tùy ý. Các thông số `object_x` và `object_y` (tương ứng `object_width` và `object_height`) biểu diễn các vị trí 2D (tương ứng các kích thước 2D) của đối tượng không gian đi kèm trong hệ tọa độ đi kèm với nguồn. Các giá trị của các thông số `object_x`, `object_y`, `object_width`, và `object_height` là tương ứng với các giá trị của các thông số `total_width` và `total_height`, như được xác định ở trên. Các vị trí (`object_x`, `object_y`) và các kích thước (`object_width`, `object_height`) của các SRD chia sẻ giá trị `source_id` có thể được so sánh sau khi tính kích thước của không gian tham chiếu, tức là sau khi các giá trị `object_x` và `object_width` được chia bởi giá trị `total_width` và các giá trị `object_y` và `object_height` được chia bởi giá trị `total_height` của các bộ mô tả tương ứng của chúng. Các giá trị `total_width` và `total_height` khác nhau có thể được sử dụng trong các bộ mô tả khác nhau để tạo ra thông tin về các vị trí và các kích thước theo các đơn vị khác nhau đối với cùng một không gian tham chiếu.

`spatial_set_id` là số nguyên không âm trong hiển thị thập phân tạo ra bộ định danh cho các nhóm đối tượng không gian. Khi không có mặt, đối tượng không gian đi kèm với bộ mô tả này không thuộc về một bộ không gian nào và thông tin bộ không gian không được cho. Các tác giả của MPD có thể biểu diễn, nhờ sử dụng thông số `spatial_set_id`, mà một số đối tượng không gian, trong `source_id` đã cho, có mối quan hệ cụ thể về không gian. Lấy ví dụ, tác giả MPD có thể nhóm tất cả các bộ thích ứng tương ứng với các ô ở cùng cấp độ phân giải. Theo cách này, thông số `spatial_set_id` có thể được sử dụng bởi máy khách DASH để chọn nhanh các đối tượng liên quan về không gian.

Việc chiếu để ánh xạ hình ảnh toàn cảnh bao phủ góc-nhìn 360-độ theo chiều ngang và góc-nhìn 180-độ theo chiều đứng (tức là hiển thị hình cầu) vào mặt phẳng

hình ảnh hai-chiều hình chữ nhật đã được biết đến là chiếu hình chữ nhật đều. Trong trường hợp này, tọa độ theo chiều ngang có thể được coi là tương đương với kinh độ, và tọa độ theo chiều đứng có thể được coi là tương đương với vĩ độ, với việc không áp dụng sự biến đổi hoặc thay đổi kích thước. Trong một số trường hợp, nội dung toàn cảnh với góc-nhìn ngang 360-độ nhưng với góc-nhìn đứng ít hơn 180-độ có thể được coi là trường hợp đặc biệt của việc trình chiếu hình chữ nhật đều, nơi mà các vùng cực của hình cầu không được ánh xạ lên mặt phẳng hình ảnh hai-chiều. Trong một số trường hợp toàn cảnh có thể có trường-nhìn ngang ít hơn 360-độ và góc-nhìn đứng đến 180-độ, trong khi nếu không thì có các đặc tính của định dạng chiếu hình chữ nhật đều.

Toàn cảnh, như hình ảnh toàn cảnh hình chữ nhật, có thể là lập thể. Trong định dạng toàn cảnh lập thể, một hình ảnh toàn cảnh có thể trình bày hình ảnh bên trái và hình ảnh toàn cảnh khác (của cùng thời gian tức thì hoặc bộ truy cập) có thể trình bày hình ảnh bên phải. Khi trình bày hình ảnh toàn cảnh lập thể trên màn hình lập thể màn hình, như tai nghe thực tại ảo, hình ảnh toàn cảnh bên trái có thể được trình bày theo phạm vi quan sát thích hợp và là phạm vi quan sát cho mắt trái, và tương tự, hình ảnh toàn cảnh bên phải có thể được trình bày cho mắt phải. Trong hình ảnh toàn cảnh lập thể, việc xem lập thể có thể được giả định xảy ra theo hướng đường xích đạo (tức là hàng điểm ảnh ở giữa nhất theo chiều đứng) của hình ảnh toàn cảnh, gây ra độ nghiêng tuyệt đối của phạm vi quan sát càng lớn, độ chính xác của trình bày lập thể ba-chiều càng kém.

Trong định dạng chiếu khối cắt, video hình cầu được chiếu lên sáu mặt (cũng được gọi là các mặt bên) của khối lập phương. Khối cắt có thể được tạo ra, ví dụ bằng cách trước tiên dựng cảnh hình cầu sáu lần từ điểm nhìn, với các khung hình được xác định bởi hình chóp cắt quan sát 90 độ trình bày từng mặt của khối lập phương. Các mặt bên của khối lập phương có thể là khung hình-được gói thành cùng khung hình hoặc mỗi mặt bên của khối lập phương có thể được xử lý riêng lẻ (ví dụ khi mã hóa). Có nhiều trình tự khả thi để đặt các mặt bên của khối lập phương lên khung hình và/hoặc các mặt bên của khối lập phương có thể quay hoặc được phản chiếu. Chiều rộng và chiều cao của khung hình để gói khung hình có thể được chọn để lấp vừa

“khít” các mặt bên của khối lập phương ví dụ ở lưới mặt bên của khối lập phương 3x2, hoặc có thể bao gồm các khung hình cấu thành không được sử dụng ví dụ ở lưới mặt bên của khối lập phương 4x3.

Khối cắt có thể là lập thể. Khối cắt lập thể có thể ví dụ đạt được bằng cách chiếu lại mỗi hình ảnh của hình ảnh toàn cảnh lập thể vào định dạng khối cắt.

Trong nhiều tình huống hiển thị, chỉ phần hình ảnh là cần thiết để được hiển thị trong khi hình ảnh còn lại đòi hỏi cần được giải mã nhưng không được hiển thị. Những tình huống hiển thị này gồm có:

- Các màn hình lắp đầu thông thường (HMD) hiển thị phạm vi quan sát ~100 độ, trong khi thường video đầu vào cho HMD tiêu dùng cá nhân bao phủ toàn bộ 360 độ.

- Các môi trường xem màn hình phẳng thông thường hiển thị góc-nhìn đến 40-độ. Khi hiển thị nội dung FOV-rộng (ví dụ mắt cá) trên màn hình này, có thể ưu tiên hiển thị tập hợp con không gian thay vì toàn bộ hình ảnh.

Bây giờ, để làm giảm thiểu các nhược điểm trên, phương pháp giải mã theo tiêu chuẩn đối với các tập hợp ô bị hạn chế chuyển động được thể hiện dưới đây.

Trong phương pháp này, mà được mô tả trong Fig.6a, bước thu được đường ghi hoặc dòng bit của hình ảnh đầy đủ bao gồm tập hợp ô bị hạn chế chuyển động (600), và trên cơ sở của tập hợp ô bị hạn chế chuyển động, các lệnh để xây dựng đường ghi chứa tập hợp ô hoặc dòng bit tuân theo hình ảnh đầy đủ được tạo ra (602). Phương pháp này có thể được thực hiện bởi bộ tạo tệp (ví dụ 1530 hoặc 1570 của Fig.23).

Trong phương pháp này, như được bộc lộ trong Fig.6b, thu được đường ghi hoặc dòng bit của hình ảnh đầy đủ bao gồm tập hợp ô bị hạn chế chuyển động (610), và trên cơ sở của tập hợp ô bị hạn chế chuyển động, xây dựng đường ghi chứa tập hợp ô hoặc dòng bit tuân theo hình ảnh đầy đủ (612). Phương pháp này có thể được thực hiện bởi tệp chương trình phân tích cú pháp (ví dụ 1580 của Fig.23) hoặc bộ phát trước khi giải mã, hoặc bởi máy chủ phát trực tuyến (ví dụ 1540 của Fig.23).

Fig.7 thể hiện ví dụ về bối cảnh làm cơ sở cho các phương án. Như là đầu

vào cho phương pháp này, dòng bit 700 thu được mà chứa một hay nhiều tập hợp ô bị hạn chế chuyển động. Ví dụ, hoặc chuỗi video hình ảnh toàn cảnh hình chữ nhật đều (700, trên cùng) hoặc chuỗi multiview mất cá gói khung hình (700, dưới cùng) có thể được sử dụng làm đầu vào. Tập hợp ô bị hạn chế chuyển động là tập hợp hình chữ nhật của các ô mà có thể chứa một hay nhiều ô. Trong Fig.7, các hình chữ nhật màu trắng minh họa các ranh giới ô. Việc mã hóa tập hợp ô bị hạn chế chuyển động được thực hiện theo cách là quy trình dự đoán liên khung bị hạn chế sao cho không có giá trị mẫu bên ngoài tập hợp ô bị hạn chế chuyển động, và không có giá trị mẫu tại vị trí mẫu phân đoạn mà xuất phát nhờ sử dụng một hay nhiều giá trị của mẫu bên ngoài tập hợp ô bị hạn chế chuyển động, được sử dụng để dự đoán liên khung của mẫu bất kỳ trong tập hợp ô bị hạn chế chuyển động.

Tiếp theo, tại 702, tập hợp con của các ô bị hạn chế chuyển động được chọn để xây dựng hoặc tạo ra các lệnh để xây dựng đường ghi chứa tập hợp ô hoặc dòng bit tuân theo hình ảnh đầy đủ từ tập hợp ô bị hạn chế chuyển động của dòng bit đầu vào. Khi xây dựng đường ghi chứa tập hợp ô hoặc dòng bit tuân theo hình ảnh đầy đủ, bước 702 có thể được thực hiện ví dụ ở máy chủ phát trực tuyến (ví dụ 1540 của Fig.23), trong khi hoặc trước khi phát trực tuyến, hoặc trong tệp chương trình phân tích cú pháp (ví dụ 1580 của Fig.23) hoặc bộ phát, trước khi giải mã. Khi tạo ra các lệnh để xây dựng đường ghi chứa tập hợp ô hoặc dòng bit tuân theo hình ảnh đầy đủ, bước 702 có thể được thực hiện ví dụ bởi bộ tạo tệp (ví dụ 1530 hoặc 1570 của Fig.23). Đầu ra của phương pháp này là đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ 704 hoặc dòng bit được xây dựng trên cơ sở tập hợp ô bị hạn chế chuyển động, trong đó đường ghi chứa tập hợp ô hoặc dòng bit tuân theo hình ảnh đầy đủ là giải mã được, như vậy, nhờ bộ mã hóa video tiêu chuẩn 706, như bộ giải mã H.265/ HEVC. Theo cách khác, đầu ra có thể là các lệnh để xây dựng đường ghi chứa tập hợp ô hoặc dòng bit tuân theo hình ảnh đầy đủ được xây dựng trên cơ sở của tập hợp ô bị hạn chế chuyển động.

Một số ví dụ được minh họa với (nhưng không giới hạn ở) sự phân chia các tập hợp ô bị hạn chế chuyển động thể hiện trong Fig.8. Trong Fig.8, giá trị 1..8 biểu thị vị trí của tập hợp ô. Giá trị trong ngoặc vuông biểu thị hình ảnh chứa tập hợp ô. Trong một số các phương án, phân đoạn hoặc phân đoạn con bao gồm các hình ảnh từ N đến

$N + M$, bao gồm, và ngoại trừ các hình ảnh khác.

Fig.9 minh họa quy trình bao gồm bước tạo ra các lệnh để xây dựng đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ. Trong "Gói tệp/Phân đoạn", dòng bit với các tập hợp ô bị hạn chế chuyển động được gói thành tệp hoặc các phân đoạn của vật chứa (ví dụ để chuyển DASH). Trong "Tạo ra đường ghi chứa tập hợp ô được xây dựng", đường ghi chứa tập hợp ô được xây dựng được tạo ra với tham chiếu tới đường ghi chứa tập hợp ô bị hạn chế chuyển động trong tệp chứa hoặc các phân đoạn.

Kết quả của quy trình trên có thể được minh họa (với sự tham chiếu nhưng không giới hạn ở ISO BMFF) với cấu trúc tệp hợp lý thể hiện trong Fig.10 chứa đường ghi hình ảnh đầy đủ và một hay nhiều đường ghi chứa tập hợp ô được xây dựng. Một hay nhiều đường ghi chứa tập hợp ô được xây dựng được tạo ra bao gồm các mẫu nhờ tham chiếu theo các toán tử hàm dựng từ đường ghi của hình ảnh đầy đủ.

Dữ liệu mẫu dùng cho các đường ghi có thể được sắp xếp cơ học trong các mảnh hoặc đường chạy trong tệp này. Ví dụ, các mẫu $(M+1)$ bắt đầu từ mẫu N trong mỗi đường ghi có thể là mảnh nhỏ hoặc đường chạy mà không được xen kẽ bởi dữ liệu của mảnh hoặc đường chạy bất kỳ khác, tương ứng. Các mảnh hoặc đường chạy, tương ứng của các đường ghi khác nhau có thể ví dụ được sắp xếp liên tục trong tệp này. Điều này được minh họa (với việc tham chiếu nhưng không giới hạn ở ISO BMFF) trong Fig.11a.

Toán tử "Constr" viết tắt của các lệnh để xây dựng mẫu định dạng tệp. "Constr" có thể ví dụ chứa một hay nhiều cấu trúc giống-bộ-NAL, mỗi trong số đưa đến các kết quả thành một bộ NAL thành mẫu định dạng tệp. Fig.11b thể hiện ví dụ của toán tử hàm dựng đối với bộ NAL VCL.

Trong phần mô tả này, khi giải thích các phương án khác nhau, có thể sử dụng các khái niệm và các định nghĩa của chúng sau đây.

Đường ghi chứa tập hợp ô được xây dựng là đường ghi chứa tập hợp ô, ví dụ đường ghi theo ISO BMFF, chứa các hàm dựng mà, khi thực hiện, dẫn đến kết quả là dòng bit của tập hợp ô.

Hàm dựng là tập hợp của các kết cấu mà, khi thực hiện, dẫn đến kết quả là

mảnh hợp lệ dữ liệu mẫu theo định dạng mẫu cơ bản.

Bộ trích xuất là hàm dựng mà, khi thực hiện, sao chép dữ liệu mẫu của phạm vi byte được biểu thị của mẫu được biểu thị của đường ghi được chỉ ra. Kết luận nhờ tham chiếu có thể được xác định như là bộ trích xuất hoặc tương tự mà, khi thực hiện, sao chép dữ liệu mẫu của phạm vi byte được biểu thị của mẫu được biểu thị của đường ghi được chỉ ra.

Tập hợp ô tuân theo hình ảnh đầy đủ {đường ghi | dòng bit } là tập hợp ô {đường ghi | dòng bit } mà phù hợp với định dạng của hình ảnh đầy đủ {đường ghi | dòng bit }. Ở đây, ký hiệu {optionA | optionB} minh họa các thay thế, tức là hoặc optionA hoặc optionB, mà được chọn một cách nhất quán trong tất cả các lựa chọn. Đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ có thể chạy như với đường ghi hình ảnh đầy đủ bất kỳ nhờ sử dụng phép phân tích cú pháp và quy trình giải mã của các đường ghi hình ảnh đầy đủ. Dòng bit tuân theo hình ảnh đầy đủ có thể được giải mã như với dòng bit hình ảnh đầy đủ bất kỳ nhờ sử dụng quy trình giải mã của các dòng bit hình ảnh đầy đủ.

Đường ghi hình ảnh đầy đủ là đường ghi thể hiện dòng bit gốc (bao gồm tất cả các ô của nó).

Hàm dựng nội tuyến là hàm dựng mà, khi thực hiện, trả lại dữ liệu mẫu mà nó chứa. Ví dụ, hàm dựng nội tuyến có thể bao gồm tập hợp của các lệnh để ghi lại phần đầu mảng mới. Đoạn nội tuyến có thể được sử dụng để biểu thị dữ liệu được mã hóa mà chứa trong mẫu của đường ghi.

Cấu trúc giống-bộ-NAL là cấu trúc với các đặc tính của bộ NAL ngoại trừ là việc chống mô phỏng mã khởi động không được thực hiện.

Đường ghi chứa tập hợp ô được xây dựng trước là đường ghi chứa tập hợp ô chứa dữ liệu mẫu nội tuyến.

Dòng bit chứa tập hợp ô là dòng bit mà chứa tập hợp ô của dòng bit gốc nhưng không trình bày toàn bộ dòng bit gốc.

Đường ghi chứa tập hợp ô là đường ghi trình bày tập hợp ô của dòng bit gốc nhưng không trình bày toàn bộ dòng bit gốc.

Các phương án khác nhau liên quan đến việc xây dựng dòng bit chứa tập hợp ô hoặc sự tạo thành các đường ghi chứa tập hợp ô được xây dựng bao gồm các lệnh để xây dựng dòng bit chứa tập hợp ô được mô tả dưới đây. Các phương án có thể áp dụng riêng biệt, nhưng không bị giới hạn, ở HEVC.

Theo một phương án, dữ liệu được mã hóa dùng cho tập hợp ô bị hạn chế chuyển động được trích xuất từ dòng bit hình ảnh đầy đủ, và cột ô được bổ sung ở bên phải của tập hợp ô bị hạn chế chuyển động được trích xuất. Cột ô được bổ sung này có thể được gọi là cột ô giả.

Theo một phương án, dữ liệu được mã hóa cho tập hợp ô bị hạn chế chuyển động được trích xuất từ dòng bit hình ảnh đầy đủ, và dòng ô được bổ sung bên dưới tập hợp ô bị hạn chế chuyển động được trích xuất. Hàng ô được bổ sung này có thể được gọi là hàng ô giả. Phương án này có thể được sử dụng ví dụ khi chiều rộng của tập hợp ô bị hạn chế chuyển động bằng chiều rộng của dòng bit hình ảnh đầy đủ hoặc khi tất cả các ô của tập hợp ô bị hạn chế chuyển động ở trong cùng phân đoạn mảng.

Theo một phương án, cột ô giả hoặc hàng ô giả được biểu thị để được loại trừ hình ảnh đầu ra của việc giải mã. Ví dụ, cửa sổ cắt ảnh không phù hợp (như được quy định Trong HEVC) hoặc tương tự được chứa ví dụ trong dữ liệu cấp chuỗi, như bộ thông số chuỗi, có thể được sử dụng loại trừ cột ô giả hoặc hàng ô giả, và bao gồm tập hợp ô bị hạn chế chuyển động được trích xuất.

Theo một phương án, các ô trong cột ô giả hoặc hàng ô giả được sử dụng khi biểu thị liệu các ô giả có kết thúc phân đoạn mảng hay không.

Theo một phương án bao gồm việc bổ sung cột ô giả, mỗi ô giả được tạo ra theo cách là nó kết thúc phân đoạn mảng. Phần đầu của phân đoạn mảng được bổ sung trước khi ô thứ nhất (tức là cực trái) của hàng ô tiếp theo trong tập hợp ô bị hạn chế chuyển động.

Theo một phương án bao gồm việc bổ sung cột ô giả, ô giả được tạo ra theo cách là nó kết thúc phân đoạn mảng, khi ô giả là ô cuối cùng, trong trình tự giải mã, của hình ảnh được xây dựng, hoặc khi ô thứ nhất (tức là cực-trái) của hàng ô tiếp theo trong tập hợp ô bị hạn chế chuyển động chứa trong phân đoạn mảng khác nhau so với ô ngoài cùng bên phải của hàng ô có mặt trong tập hợp ô bị hạn chế chuyển động. Mặt

khác, ô giả được tạo ra theo cách là nó không kết thúc phân đoạn mảng.

Theo một phương án bao gồm việc bổ sung cột ô giả, các phần đầu phân đoạn mảng của các phân đoạn mảng bao gồm các ô của tập hợp ô bị hạn chế chuyển động có thể được nghiên cứu. Nếu các phần đầu của phân đoạn mảng là, thông qua nghiên cứu hoặc thông qua các cách khác (ví dụ các hạn chế được xác định trước), được tìm thấy hầu như giống nhau, hình ảnh được xây dựng được tạo ra theo cách là nó bao gồm phần đầu phân đoạn mảng chỉ khi bắt đầu, ô giả mà là cuối cùng, trong trình tự giải mã, trong hình ảnh được xây dựng được tạo ra theo cách là nó kết thúc phân đoạn mảng, và các ô giả khác được tạo ra theo cách là chúng không kết thúc phân đoạn mảng.

Theo một phương án bao gồm việc bổ sung hàng ô giả, ô giả mà là cuối cùng, trong trình tự giải mã, trong hình ảnh được xây dựng được tạo ra theo cách là nó kết thúc phân đoạn mảng, và các ô giả khác được tạo ra theo cách là chúng không kết thúc phân đoạn mảng.

Cấu trúc cú pháp của `slice_segment_data ( )` của HEVC được xác định như sau, với `ae(v)` biểu thị phần tử cú pháp được mã hóa với việc mã hóa số học, tức là trong HEVC, việc mã hóa số học bậc hai hợp bối cảnh (CABAC):

<code>slice_segment_data() {</code>	Bộ mô tả
<code>do {</code>	
<code>coding_tree_unit()</code>	
<code>end_of_slice_segment_flag</code>	ae(v)
<code>CtbAddrInTs++</code>	
<code>CtbAddrInRs = CtbAddrTsToRs[CtbAddrInTs]</code>	
<code>if(!end_of_slice_segment_flag && ((tiles_enabled_flag && TileId[CtbAddrInTs] != TileId[CtbAddrInTs - 1]) (entropy_coding_sync_enabled_flag && (CtbAddrInTs % PicWidthInCtbsY == 0 TileId[CtbAddrInTs] != TileId[CtbAddrRsToTs[CtbAddrInRs - 1]]))))) {</code>	
<code>end_of_subset_one_bit</code> /* equal 1 */	ae(v)
<code>byte_alignment()</code>	
<code>}</code>	
<code>} while(!end_of_slice_segment_flag)</code>	
<code>}</code>	

Theo một phương án có thể áp dụng cho HEVC hoặc cú pháp tương tự, ô giả mà kết thúc phân đoạn mảng được tạo ra bởi việc cài đặt `end_of_slice_segment_flag` bằng 1 sau CTU cuối cùng, trong trình tự giải mã, của ô giả. Ô giả mà không kết thúc phân đoạn mảng được tạo ra bằng cách cài đặt `end_of_slice_segment_flag` bằng 0 (và `end_of_subset_one_bit` bằng 1, như được biểu thị trong cú pháp ở trên).

Theo một phương án, chiều rộng của cột ô giả là tập hợp bằng một CTU. Theo một phương án, chiều cao của hàng ô giả là tập hợp bằng một CTU.

Theo một phương án, ô giả được mã hóa theo cách hiệu quả khi nén. Ví dụ, các bộ mã hóa của ô giả có thể được mã hóa mà không có mặt dữ liệu lỗi dự đoán. Có thể sử dụng CU kích thước lớn nhất có thể; ví dụ trong HEVC `split_cu_flag []` có thể là tập hợp bằng 0 trong các ô giả. Đối với các ô giả HEVC được mã hóa nội bộ, `rqt_root_cbf` có thể là tập hợp bằng 0 (giả định là chế độ PCM không được sử dụng). Đối với các ô giả HEVC được mã hóa nội bộ, `cu_skip_flag []` có thể là tập hợp bằng 1.

Các phương án khác nhau liên quan đến việc xây dựng dòng bit chứa tập hợp

ô từ dòng bit hình ảnh đầy đủ với việc ghi lại phần đuôi của ô (tức là số lượng của các byte bắt đầu từ sự kết thúc của ô về phía bắt đầu của ô) được mô tả dưới đây.

Theo một phương án, tập hợp của các ô bao gồm các ô ngoài cùng bên phải của tập hợp ô bị hạn chế chuyển động mà có chiều rộng nhỏ hơn chiều rộng hình ảnh của dòng bit hình ảnh đầy đủ. Mỗi ô ngoài cùng bên phải nêu trên là tập hợp để kết thúc phân đoạn mảng trong dòng bit chứa tập hợp ô. Phần đầu phân đoạn mảng được bổ sung trước khi ô thứ nhất (tức là cực-trái) của hàng ô tiếp theo trong tập hợp ô bị hạn chế chuyển động.

Theo một phương án, tập hợp của các ô bao gồm một ô cho mỗi mỗi phân đoạn mảng chứa các ô của tập hợp ô bị hạn chế chuyển động, nơi mà một ô là ô cuối cùng, trong trình tự giải mã trong tập hợp ô bị hạn chế chuyển động, của phân đoạn mảng tương ứng. Mỗi ô trong tập hợp của các ô là tập hợp kết thúc phân đoạn mảng trong dòng bit chứa tập hợp ô, và nếu ô tiếp theo trong tập hợp ô bị hạn chế chuyển động theo sau, phần đầu phân đoạn mảng được bổ sung.

Theo một phương án, các phần đầu phân đoạn mảng của (các) phân đoạn mảng bao gồm các ô của tập hợp ô bị hạn chế chuyển động có thể được nghiên cứu. nếu các phần đầu phân đoạn mảng là, thông qua nghiên cứu hoặc thông qua các cách khác (ví dụ các hạn chế được xác định trước), được tìm thấy hầu như giống nhau, hình ảnh được xây dựng được tạo ra theo cách là nó bao gồm phần đầu phân đoạn mảng chỉ ở lúc bắt đầu. Đối với các ô ngoại trừ ô cuối cùng của tập hợp ô bị hạn chế chuyển động, được biểu thị là phân đoạn mảng không phải là kết thúc, và đối với ô cuối cùng của tập hợp ô bị hạn chế chuyển động, được biểu thị là phân đoạn mảng kết thúc.

Các phương án khác nhau liên quan đến việc xây dựng đường ghi chứa tập hợp ô với việc ghi lại phần đuôi của ô được mô tả dưới đây.

Theo một phương án, hai phiên bản của ô được mã hóa, mà nếu không thì giống nhau nhưng phiên bản thứ nhất kết thúc phân đoạn mảng (hoặc mảng, trong một số phương án) và phiên bản thứ hai không kết thúc phân đoạn mảng (hoặc mảng, trong một số các phương án). Phiên bản mà phù hợp với mục đích dòng bit hình ảnh đầy đủ được chứa trong dòng bit hình ảnh đầy đủ. Điều này được kết luận liệu phiên bản giống hoặc khác có phù hợp với mục đích của dòng bit chứa tập hợp ô hay không, vì

dụ như được mô tả ở trên. Trong trường hợp mà phiên bản giống phù hợp mục đích của dòng bit chứa tập hợp ô, ô được chứa bằng cách tham chiếu trong đường ghi chứa tập hợp ô. Trong trường hợp mà phiên bản khác phù hợp với mục đích của dòng bit chứa tập hợp ô, ô của dòng bit hình ảnh đầy đủ được chứa bằng cách tham chiếu trong đường ghi chứa tập hợp ô, ngoại trừ phần đuôi của ô (ví dụ byte cuối cùng) nơi sự kết thúc phân đoạn mảng được biểu thị riêng biệt. Phần đuôi của ô được chứa nội tuyến trong đường ghi chứa tập hợp ô.

Theo một phương án, dòng bit hình ảnh đầy đủ được giải mã-entropy một phần hoặc hoàn toàn (ví dụ như được mô tả ở trên) và nó được kết luận nếu dấu hiệu của sự kết thúc của phân đoạn mảng có giá trị đúng đối với mục đích của dòng bit chứa tập hợp ô (ví dụ như được mô tả ở trên). Trong trường hợp mà sự kết thúc của phân đoạn mảng có giá trị đúng đối với mục đích của dòng bit chứa tập hợp ô, ô đó được chứa bởi tham chiếu trong đường ghi chứa tập hợp ô. Trong trường hợp mà sự kết thúc của phân đoạn mảng không có giá trị đúng đối với mục đích của dòng bit chứa tập hợp ô, ô của dòng bit hình ảnh đầy đủ được chứa bởi tham chiếu trong đường ghi chứa tập hợp ô, ngoại trừ phần đuôi của ô (ví dụ byte cuối cùng). Ô này được mã hóa-entropy hoàn toàn hoặc một phần với dấu hiệu được thay đổi của sự kết thúc phân đoạn mảng (ví dụ như được mô tả ở trên) và phần đuôi của ô được mã hóa-entropy được chứa nội tuyến trong đường ghi chứa tập hợp ô.

Như được trình bày ở trên, khi sử dụng định dạng dòng byte để chuyển giao hoặc lưu giữ dữ liệu video (như được quy định, ví dụ trong H.264/AVC hoặc trong HEVC), việc mô phỏng mã khởi động và xử lý ngăn ngừa mô phỏng mã khởi động cần phải được tính toán.

Khi xây dựng dòng bit chứa tập hợp ô từ các ô mà là không-liên tục trong dòng bit hình ảnh đầy đủ hoặc khi bổ sung dữ liệu vào dòng bit chứa tập hợp ô không có mặt trong dòng bit hình ảnh đầy đủ (ví dụ, bổ sung các phần đầu phân đoạn mảng), việc mô phỏng mã khởi động có thể xuất hiện. Do đó, cần phải quan tâm để bao gồm phương tiện ngăn ngừa mô phỏng mã khởi động trong các trường hợp này.

Khi xây dựng dòng bit chứa tập hợp ô từ các ô mà không-liên tục trong dòng bit hình ảnh đầy đủ hoặc khi bổ sung dữ liệu vào dòng bit chứa tập hợp ô mà không có

mặt trong dòng bit hình ảnh đầy đủ (ví dụ bổ sung các phần đầu phân đoạn mảng), phần tử cú pháp ngăn ngừa mô phỏng mã khởi động giả có thể xuất hiện. Khi chuỗi của ba byte trong dòng bit chứa tập hợp ô chứa các byte từ nguồn khác (ví dụ các ô không-liên tục) và bằng 0x000003 trong trường hợp HEVC hoặc, tổng quát hơn, bằng với mẫu giá trị byte ngăn ngừa mô phỏng mã khởi động, bộ giải mã kết luận sai lầm là byte cuối cùng của ba byte là byte ngăn ngừa mô phỏng. Do đó, cần phải quan tâm để tránh phương tiện ngăn ngừa mô phỏng mã khởi động giả trong các trường hợp này.

Theo một phương án, các phần tử cú pháp ngăn ngừa mô phỏng mã khởi động (ví dụ các phần tử cú pháp emulation_prevention_tree_byte trong HEVC) được lấy ra từ dữ liệu video được mã hóa (ví dụ các bộ NAL VCL trong HEVC) của dòng bit hình ảnh đầy đủ, dẫn đến kết quả dòng bit hình ảnh đầy đủ được xử lý. Dòng bit chứa tập hợp ô trung gian sau đó được xây dựng từ dòng bit hình ảnh đầy đủ được xử lý, ví dụ như được mô tả trong các phương án khác. Các phần tử cú pháp ngăn ngừa mô phỏng mã khởi động được bổ sung vào dòng bit chứa tập hợp ô trung gian để tạo ra dòng bit chứa tập hợp ô.

Theo một phương án, đã được nghiên cứu liệu hai hoặc nhiều mảnh bất kỳ của dữ liệu liên tục từ nguồn khác trong dòng bit chứa tập hợp ô cần-được-xây dựng tạo ra mã khởi động hoặc ngăn ngừa mô phỏng mã khởi động giả, trong trường hợp đó việc ngăn ngừa mô phỏng mã khởi động được bổ sung vào dòng bit chứa tập hợp ô.

Theo một phương án, cú pháp ngăn ngừa mô phỏng mã khởi động mà là cần được bổ sung trong dòng bit chứa tập hợp ô, như được kết luận bởi (các) phương án thực hiện ở trên, được chứa nội tuyến trong đường ghi chứa tập hợp ô.

Phương án thực hiện được minh họa trên Fig.9 có thể được kèm theo bởi thể hệ MPD như sau và như được minh họa trên Fig.12. So sánh với toàn bộ cấu hình thể hiện trên Fig.9, MPD được tạo ra mà bao gồm thông tin trên các đường ghi chứa tập hợp ô và các phân đoạn hoặc các phân đoạn con của chúng, ví dụ trên cơ sở cấu hình để tạo ra các phân đoạn (phụ) trong gói tệp/phân đoạn và/hoặc cấu hình tạo ra các tập hợp ô bị hạn chế chuyển động. Theo cách khác hoặc ngoài ra, thể hệ MPD có thể nhận được như là thông tin đầu vào trên tập hợp các ô được xây dựng được tạo ra, ví dụ vị trí không gian của chúng và kích thước tương ứng với hình ảnh đầy đủ.

Theo một phương án, các đường ghi hình ảnh đầy đủ tương ứng với Hiện thị thứ nhất và đường ghi chứa tập hợp ô được xây dựng (có các tham chiếu đến đường ghi hình ảnh đầy đủ) tương ứng với hiện thị thứ hai. Trong thể hệ MPD, hiện thị thứ nhất được biểu thị để có trong bộ thích ứng thứ nhất và hiện thị thứ hai được biểu thị để có trong bộ thích ứng thứ hai.

Theo một phương án, SRD thứ nhất được chứa trong bộ thích ứng thứ nhất và SRD thứ hai được chứa trong bộ thích ứng thứ hai. SRD thứ nhất và SRD thứ hai biểu thị các bộ thích ứng để có cùng một không gian tham chiếu, ví dụ có cùng giá trị của `source_id` như được giải thích trước đó. Vị trí và kích thước của tập hợp ô bị hạn chế chuyển động được thể hiện bởi đường ghi chứa tập hợp ô được xây dựng được biểu thị trong SRD thứ hai liên quan đến các kích thước hình ảnh đầy đủ mà có thể được biểu thị trong SRD thứ nhất.

Theo một phương án, bộ phát kết luận là bộ thích ứng thứ nhất và bộ thích ứng thứ hai thể hiện cùng một không gian tham chiếu, ví dụ dựa trên chúng có cùng giá trị của `source_id` như đã được giải thích trước đó. Bộ phát còn kết luận kích thước và vị trí tương đối của bộ thích ứng thứ nhất đối với bộ thích ứng thứ hai. Dựa trên, ví dụ, hướng nhìn thịnh hành và trường hiển thị thịnh hành của hình ảnh, bộ phát chọn bộ thích ứng mà phù hợp với các điều kiện nhìn nhất.

Trong ISO/IEC 14496-15 ("mang video cấu trúc bộ phận lớp trừu tượng hóa mạng (NAL) trong định dạng tệp cơ sở ISO"), cấu trúc của bộ trích xuất có nghĩa là để trích xuất một hay nhiều bộ NAL trọn vẹn. Đối với mục đích của các phương án được mô tả ở đây, các bộ NAL trích xuất một phần (không trọn vẹn) là cần thiết, mà có thể đạt được bởi hàm dựng của bộ trích xuất, như được xác định ở trên. Cấu trúc giống-bộ-NAL có thể được xác định để bao gồm các hàm dựng mà việc thi hành chúng dẫn đến kết quả một hay nhiều bộ NAL trọn vẹn. Tính chất này có thể đạt được với những thay đổi sau trong cấu trúc của bộ trích xuất theo ISO/IEC 14496-15, nhưng nó cần được hiểu là các phương án này không bị giới hạn ở cú pháp hoặc ngữ nghĩa này.

Cấu trúc `ExtractorPayload`

```
class aligned(8) Extractor () {
    NALUnitHeader();
```

```

unsignedint(8) track_ref_index;
signedint(8) sample_offset;
unsignedint((lengthSizeMinusOne+1)*8)
    data_offset;
unsignedint((lengthSizeMinusOne+1)*8)
    data_length;
}

```

được đổi thành:

```

class aligned(8) ExtractorPayload () {
    unsignedint(8) track_ref_index;
    signedint(8) sample_offset;
    unsignedint((lengthSizeMinusOne+1)*8)
        data_offset;
    unsignedint((lengthSizeMinusOne+1)*8)
        data_length;
}

class aligned(8) Extractor () {
    NALUnitHeader();
    ExtractorPayload();
}

```

Ở đây, `track_ref_index` biểu thị đường ghi từ nơi dữ liệu được trích xuất, `sample_offset` biểu thị mẫu từ nơi dữ liệu được trích xuất, `data_offset` biểu thị độ lệch trong mẫu bắt đầu từ đó dữ liệu được trích xuất, và `data_length` biểu thị số lượng của các byte cần được trích xuất. Cấu trúc `NALUnitHeader` chứa phần đầu của bộ NAL.

Cấu trúc `InlineNALUnitData` chứa dữ liệu của các byte cần chứa trong bộ NAL được xây dựng.

```
class aligned(8) InlineNALUnitData () {
    unsignedint(8) length;
    unsignedint(8) inline_data[length];
}
```

Trong đó, độ dài là số lượng của các byte cần được chứa trong bộ NAL được xây dựng, và inline_data là dữ liệu các byte cần được chứa trong bộ NAL được xây dựng.

Để xây dựng bộ NAL, ít nhất hai lựa chọn khác nhau có thể được đề xuất. Trong lựa chọn thứ nhất, các bộ trích xuất của HEVC được xác định lại sao cho chúng có thể được xây dựng từ nhiều hơn một mức độ. Mức độ có thể hoặc là ExtractorPayload hoặc InlineNALUnitData.

Cú pháp có thể như sau:

```
class aligned(8) Extractor () {
    NALUnitHeader()
    do {
        unsignedint(8)    nal_unit_extnt_type;
        if(nal_unit_extnt_type == 0 )
            ExtractorPayload();
        else // if( nal_unit_extnt_type == 1 )
            InlineNALUnitData();
    } while( !EndOfNALUnit() )
}
```

Trong lựa chọn thứ hai, cấu trúc NALUnitConstructor được xác định. Việc giải quyết NALUnitConstructor dẫn đến kết quả một hay nhiều bộ NAL trọn vẹn.

```
class aligned(8) NALUnitConstructor () {
    NALUnitHeader();
    do {
        unsignedint(8)    nal_unit_extnt_type;
```

```

        if(nal_unit_extent_type == 0 )
            ExtractorPayload();
        else // if( nal_unit_extent_type == 1 )
            InlineNALUnitData();
    } while( !EndOfNALUnit() )
}

```

Ngữ nghĩa của NALUnitHeader là giống với bộ trích xuất (Extractor) ngoại trừ là giá trị `nal_unit_type` khác được sử dụng, ví dụ `nal_unit_type` bằng 50 được sử dụng.

Chi phí của việc thay đổi sinh ra bởi lựa chọn thứ nhất được so với việc sử dụng thông thường của các bộ trích xuất mà trích xuất một hay nhiều bộ NAL trọn vẹn là một byte bổ sung. Vì chi phí này là vừa phải, nên lựa chọn thứ nhất có thể được ưu tiên đối với lựa chọn thứ hai của cấu trúc NALUnitConstructor riêng biệt.

Trong cả hai lựa chọn, `nal_unit_extent_type` bằng 0 xác định là mức tiếp theo là `ExtractorPayload`. `nal_unit_extent_type` bằng 1 xác định mức này là `InlineNALUnitData`. Các giá trị lớn hơn 1 được bảo toàn.

Việc giải quyết bộ trích xuất được đề xuất hoặc cấu trúc hàm dựng của bộ NAL có thể được yêu cầu để được hỗ trợ ở bộ phát hỗ trợ (các) mã bốn-ký tự nhập mẫu nhất định, như 'hvc2' và 'hev2'.

Theo một phương án, đường ghi chứa tập hợp ô được xây dựng được tạo ra bao gồm bộ trích xuất được đề xuất hoặc (các) cấu trúc hàm dựng của bộ NAL, mà trích xuất tập hợp ô bị hạn chế chuyển động từ đường ghi hình ảnh đầy đủ (ví dụ với mã nhập mẫu 'hvc1' hoặc 'hev1'). Đường ghi chứa tập hợp ô được xây dựng có thể có mã nhập mẫu, như 'hvc2' hoặc 'hev2', mà biểu thị định dạng mã hóa cơ sở, như HEVC, và việc sử dụng bộ trích xuất hoặc cấu trúc hàm dựng của bộ NAL.

Theo một phương án, bộ phát xác định phân tích cú pháp, giải mã và chạy đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ, ví dụ được đánh dấu bằng mã nhập mẫu 'hvc2' hoặc 'hev2'. Bộ phát giải quyết hoặc hoàn thành các hàm dựng chứa trong các mẫu của đường ghi để thu được dòng bit chứa tập hợp ô tuân theo hình ảnh

đầy đủ. Sau đó, bộ phát giải mã dòng bit chứa tập hợp ô tuân theo hình ảnh đầy đủ với quy trình giải mã thông thường, như quy trình giải mã HEVC.

Các phương án khác nhau liên quan đến việc thực hiện với HEVC hoặc bộ giải mã bất kỳ khác với các tính năng tương tự được mô tả dưới đây.

Theo một phương án, các bộ thông số của đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ được sao chép từ dòng bit gốc với một hay nhiều biến thể sau:

- Cấp được biểu thị có thể được thiết lập đến giá trị đủ để giải mã tập hợp ô thay vì giải mã hình ảnh đầy đủ. Tin nhắn SEI của tập hợp ô bị hạn chế chuyển động tạm thời có thể chứa giá trị của mức (`mcts_level_idc[ i ]`) mà thích hợp để được sử dụng trong bộ thông số được cải biến.

- Chiều rộng và chiều cao hình ảnh có thể được thiết lập đến các giá trị phù hợp với tập hợp ô thay vì hình ảnh đầy đủ.

- Cửa sổ cắt ảnh phù hợp có thể được thiết lập đến các giá trị phù hợp với tập hợp ô, hoặc có thể được biểu thị là vắng mặt trong trường hợp đó vùng hình ảnh đầy đủ của dòng bit chứa tập hợp ô là đầu ra.

- Lưới ô có thể được biểu thị đối với tập hợp ô thay vì đối với hình ảnh đầy đủ.

Theo một phương án, các bộ thông số được chứa trong phần mô tả nhập mẫu của (các) đường ghi chứa tập hợp ô. Theo một phương án, các bộ thông số được chứa trong các mẫu của (các) đường ghi chứa tập hợp ô (theo cách khác, hoặc ngoài ra bao gồm các bộ thông số trong phần mô tả nhập mẫu của (các) đường ghi chứa tập hợp ô). Theo một phương án, các bộ thông số được chứa nội tuyến trong các mẫu. Theo một phương án, các bộ thông số, theo cách khác hoặc ngoài ra bao gồm các bộ thông số nội tuyến trong các mẫu, được xây dựng trong các mẫu. Ví dụ, một phần của bộ thông số có thể được trích xuất từ mẫu của đường ghi khác, và phần khác của bộ thông số có thể được chứa nhờ sử dụng hàm dựng nội tuyến.

Theo một phương án, (các) phần đầu của bộ NAL của đường ghi chứa tập hợp ô được chứa nhờ sử dụng hàm dựng nội tuyến.

Theo một phương án, các phần đầu của phân đoạn mảng chứa trong các mẫu của đường ghi chứa tập hợp ô nhờ sử dụng hàm dựng nội tuyến. Theo một phương án,

ngoài ra hoặc theo cách khác, bao gồm các phần đầu của phân đoạn mảng với hàm dựng nội tuyến, các phần đầu của phân đoạn mảng có thể được trích xuất từ đường ghi khác. Ví dụ, phần của phần đầu phân đoạn mảng có thể được trích xuất từ mẫu của đường ghi khác, và phần khác của phần đầu phân đoạn mảng có thể được chứa nhờ sử dụng hàm dựng nội tuyến. Một hay nhiều nội dung sau có thể cần để được cải biến so với phần đầu phân đoạn mảng tương ứng của dòng bit hình ảnh đầy đủ gốc:

- dấu hiệu của mảng thứ nhất của hình ảnh (trong HEVC, `first_slice_segment_in_pic_flag`)
- địa chỉ mảng hoặc vị trí của khối thứ nhất (ví dụ CTU) của mảng (trong HEVC, `slice_segment_address`)
- các độ lệch điểm truy nhập (mỗi bù đắp nhắm vào điểm bắt đầu của ô), khi có mặt và khi số lượng của các ô trong phân đoạn mảng thay đổi.

Theo một phương án, phần đầu của phân đoạn mảng được bổ sung vào đường ghi chứa tập hợp ô khi hàng ô trong tập hợp ô bị hạn chế chuyển động không bắt đầu với phần đầu phân đoạn mảng và chiều rộng của tập hợp ô bị hạn chế chuyển động nhỏ hơn chiều rộng hình ảnh của dòng bit hình ảnh đầy đủ gốc.

Như được hiểu hiển nhiên từ những nội dung trên, phương pháp này có thể được áp dụng theo cách phối hợp đối với một số phiên bản được mã hóa của cùng một nội dung. Fig.13 minh họa cách các đường ghi/dòng bit chứa tập hợp ô bị hạn chế chuyển động có thể được xây dựng từ hai dòng bit H.265/HEVC có cùng một nội dung nhưng với độ phân giải khác nhau.(97)

Các phiên bản được mã hóa có thể khác nhau, ví dụ trong một hay nhiều tiêu chí sau:

- Tốc độ bit
- Chất lượng hình ảnh, mà có thể có thu được nhờ sử dụng giá trị thông số lượng tử hóa khác nhau
- Độ phân giải không gian
- Độ sâu bit cho mỗi thành phần màu
- Dải tần nhạy sáng

- Độ rộng của giải màu

Việc xem xét, ví dụ, hai phiên bản của cùng một nội dung được mã hóa với độ phân giải không gian khác nhau như thể hiện trong Fig.13, sau đó theo một phương án, các ranh giới của các tập hợp ô bị hạn chế chuyển động trong hai phiên bản này được đặt cùng chỗ về không gian. Điều này cho phép tạo ra các đường ghi chứa tập hợp ô hoặc các dòng bit không chồng chéo về không gian có độ phân giải khác nhau. Theo một phương án, các ranh giới của các tập hợp ô bị hạn chế chuyển động trong hai phiên bản này không cần được đặt cùng chỗ về không gian. Các tập hợp ô bị hạn chế chuyển động ở độ phân giải khác nhau có thể được chọn ví dụ để phát trực tuyến và/hoặc để giải mã theo cách chồng chéo giữa các đường ghi chứa tập hợp ô hoặc các dòng bit của các độ phân giải khác nhau không tồn tại hoặc vừa phải.

Một hay nhiều đường ghi chứa tập hợp ô hoặc các dòng bit độ phân giải cao được xây dựng từ phiên bản độ phân giải cao. Một hay nhiều đường ghi chứa tập hợp ô hoặc các dòng bit độ phân giải thấp được xây dựng từ phiên bản độ phân giải thấp. Theo một phương án, các đường ghi chứa tập hợp ô hoặc các dòng bit được xây dựng theo cách là toàn bộ vùng hình ảnh (hoặc tương đương toàn bộ phạm vi quan sát) của nội dung được bao hàm. Trong phương án khác, các đường ghi chứa tập hợp ô hoặc các dòng bit được xây dựng theo cách là phạm vi quan sát đủ để việc hiển thị được đảm bảo, tính đến các thay đổi có thể tiếp theo trong hướng nhìn. Theo một phương án, (các) đường ghi chứa tập hợp ô độ phân giải cao có thể được xây dựng từ (các) tập hợp ô bị hạn chế chuyển động mà đảm bảo có mặt hướng nhìn và trường nhìn.

Như điều trở nên hiển nhiên từ phần trên, phương pháp này có thể được áp dụng theo cách phối hợp đối với một số lớp được mã hóa của cùng một nội dung. Fig.14 minh họa cách đường ghi/các dòng bit tập hợp ô bị hạn chế chuyển động có thể được xây dựng từ hai lớp có thể mở rộng H.265/HEVC có cùng một nội dung nhưng với độ phân giải khác nhau.

Lớp được mã hóa khác nhau, ví dụ trong một hay nhiều tiêu chí sau:

- Tốc độ bit
- Chất lượng hình ảnh, mà có thể có thu được nhờ sử dụng giá trị thông số lượng tử hóa

- Độ phân giải không gian
- Độ sâu bit cho mỗi thành phần màu
- Dải tần nhạy sáng
- Độ rộng của giải màu

Việc xem xét, ví dụ, hai lớp của cùng một nội dung được mã hóa với độ phân giải không gian khác nhau, sau đó lớp cơ sở không cần được mã hóa với các ô hoặc các ô bị hạn chế chuyển động. Lớp cơ sở có thể được mã hóa với mã khác, như H.264/AVC. Không có đường ghi chứa tập hợp ô hoặc việc xây dựng dòng bit được thực hiện đối với lớp cơ sở.

Lớp được dự đoán sử dụng dự đoán liên lớp từ lớp cơ sở.

Một hay nhiều đường ghi chứa tập hợp ô hoặc lớp có độ phân giải cao được xây dựng từ lớp được dự đoán. Theo một phương án, (các) đường ghi chứa tập hợp ô hoặc (các) lớp độ phân giải cao có thể được xây dựng từ (các) tập hợp ô bị hạn chế chuyển động mà đảm bảo có mặt hướng nhìn và phạm vi quan sát.

Theo một phương án, các bù đắp vị trí lớp tham chiếu, như bù đắp lớp tham chiếu được thu nhỏ, hoặc tương tự được chèn vào đường ghi chứa tập hợp ô hoặc dòng bit (ví dụ trong bộ thông số hình ảnh) để biểu thị phần nào của vùng hình ảnh lớp cơ sở mà lớp được dự đoán tương ứng với. Theo một phương án, bộ giải mã giải mã bù đắp vị trí lớp tham chiếu, như các bù đắp lớp tham chiếu được thu nhỏ, hoặc tương tự từ đường ghi chứa tập hợp ô hoặc dòng bit (ví dụ từ bộ thông số hình ảnh) để kết luận phần nào của vùng hình ảnh lớp cơ sở mà lớp được dự đoán tương ứng với. Bộ giải mã hoặc quy trình hiển thị có thể sử dụng thông tin để chọn các mẫu từ lớp cơ sở được tham chiếu và các hình ảnh lớp tăng cường cần được hiển thị. Ví dụ, nếu phạm vi quan sát được yêu cầu để hiển thị không được đảm bảo đầy đủ bởi (các) hình ảnh lớp tăng cường, các phần của hình ảnh lớp cơ sở không được đảm bảo bởi lớp tăng cường có thể được chọn để hiển thị.

Theo một phương án, lớp cơ sở được mã hóa với các ô bị hạn chế chuyển động, mà đặt chung với các ô của lớp được dự đoán. Ngoài ra, dự đoán liên lớp bị hạn chế khi mã hóa theo cách là không có các giá trị của mẫu hoặc thông tin vectơ chuyển

động bên ngoài ô được đặt cùng chỗ trong lớp cơ sở cần cho dự đoán liên lớp của ô trong lớp được dự đoán. Lớp cơ sở hoặc đường ghi của lớp cơ sở là tùy thuộc vào lớp hoặc đường ghi chứa kết cấu tập hợp ô theo cách đặt cùng chỗ các tập hợp ô bị hạn chế chuyển động được xây dựng cho lớp cơ sở và lớp được dự đoán.

Fig.15 minh họa một phương án được áp dụng để mã hóa video lập thể. Cần lưu ý là trong khi hình vẽ làm ví dụ minh họa hình ảnh toàn cảnh hình chữ nhật lập thể, phương pháp này không bị giới hạn ở định dạng này, nhưng áp dụng như nhau cho các định dạng lập thể khác, như định dạng khối cắt lập thể. Nó cũng được đánh dấu là trong khi hình vẽ dưới đây minh họa việc mã hóa hai- hình, phương pháp này không bị giới hạn ở hai hình chỉ áp dụng để hiển thị việc mã hóa nhiều hình với số lượng xem lớn hơn hai.

Trong Fig.15, một khung hình (ví dụ hình trái) được mã hóa với hoặc không có các ô bị hạn chế chuyển động và có thể được mã hóa với mã bất kỳ, như HEVC (ví dụ hồ sơ chính) hoặc H.264/AVC. Hình khác được mã hóa nhờ sử dụng các ô bị hạn chế chuyển động, và một hay nhiều đường ghi chứa tập hợp ô hoặc các dòng bit được xây dựng từ hình khác này. Các hình không cần có cùng độ phân giải không gian hoặc các đặc tính khác (ví dụ tốc độ bit), chất lượng hình ảnh, độ sâu bit cho mỗi thành phần màu, dải tần nhạy sáng, và/hoặc độ rộng của giải màu).

Bộ phát có thể chọn đường ghi chứa tập hợp ô hoặc dòng bit cần truyền trực tuyến và/hoặc được giải mã, ví dụ dựa trên hướng nhìn hiện tại. Khi một hình được truyền toàn bộ, bộ phát có thể xử lý sự thay đổi bất chợt của hướng nhìn nhờ hiển thị nội dung chỉ từ một hình (khi vắng mặt nội dung tương ứng từ hình khác).

Theo một phương án, cả hai hình được mã hóa nhờ sử dụng các ô bị hạn chế chuyển động, và một hay nhiều đường ghi chứa tập hợp ô hoặc các dòng bit được xây dựng từ mỗi hình. Các hình này không cần phải có cùng độ phân giải không gian hoặc các đặc tính khác (ví dụ bitrate), chất lượng hình ảnh, độ sâu bit cho mỗi thành phần màu, dải tần nhạy sáng, và/hoặc độ rộng của giải màu).

Bộ phát có thể chọn đường ghi chứa tập hợp ô hoặc dòng bit truyền dẫn trực tuyến và/hoặc được giải mã từ mỗi hình ví dụ dựa trên hướng nhìn hiện tại. Phương án này có thể phù hợp ví dụ các môi trường nhìn nơi các thay đổi nhanh của hướng nhìn

không thể hoặc sẽ không xảy ra.

Theo một phương án, cả hai hình được mã hóa với nhiều độ phân giải, như được mô tả ở trên trong các phương án của "Giải mã một phần hình ảnh nhiều-độ phân giải dựa trên các tập hợp ô bị hạn chế chuyển động". Đối với một hình hoặc cả hai hình, bộ phát chọn một hay nhiều tập hợp ô bị hạn chế chuyển động của hình cần truyền trực tuyến và/hoặc được giải mã ở độ phân giải cao, trong khi một số hoặc tất cả các tập hợp ô còn lại (tức là không-được đặt cùng chỗ) được truyền trực tuyến và/hoặc được giải mã ở độ phân giải thấp.

Các phương án có thể được áp dụng để mã hóa lập thể và video nhiều hình với dự đoán liên khung hình giữa các hình. Phương pháp này không bị giới hạn ở định dạng cụ thể bất kỳ nhưng áp dụng như nhau cho định dạng nhiều hình hoặc lập thể bất kỳ, như định dạng hình ảnh toàn cảnh hình chữ nhật lập thể và định dạng khối cắt lập thể.

Theo một phương án, một hình (ví dụ hình trái) được mã hóa với hoặc không cần các ô bị hạn chế chuyển động và có thể được mã hóa với mã bất kỳ, như HEVC (ví dụ hồ sơ chính) hoặc H.264/AVC. Hình khác được mã hóa nhờ sử dụng các ô bị hạn chế chuyển động, và một hay nhiều đường ghi chứa tập hợp ô hoặc các dòng bit được xây dựng từ hình khác đó. Dự đoán liên khung hình được áp dụng giữa các hình.

Bộ phát có thể chọn đường ghi chứa tập hợp ô hoặc dòng bit cần được truyền trực tuyến và/hoặc được giải mã ví dụ dựa trên hướng nhìn hiện tại. Khi một hình được truyền hoàn toàn, bộ phát có thể xử lý sự thay đổi đột ngột của hướng nhìn bằng cách hiển thị nội dung chỉ từ một hình (khi vắng mặt nội dung tương ứng nội dung từ hình khác).

Phương án này đòi hỏi định dạng mã hóa và khả năng giải mã để hỗ trợ các hình có các độ phân giải không gian khác nhau. Lưu ý là ví dụ MV-HEVC không cho phép như vậy. Nói chung, dự đoán liên khung hình có thể bao gồm việc xử lý liên lớp, như chia tỷ lệ đối với các độ sâu bit khác nhau hoặc lấy mẫu lại các mảng mẫu và/hoặc các trường chuyển động, và do đó các hình không cần có các đặc tính (ví dụ độ sâu bit cho mỗi thành phần màu, dải tần nhạy sáng, và/hoặc độ rộng của giải màu).

Phương án khác được minh họa trong Fig.16. Hình cơ sở (ví dụ hình trái)

được mã hóa với các ô bị hạn chế chuyển động. Một hay nhiều đường ghi chứa tập hợp ô hoặc các dòng bit được xây dựng từ hình cơ sở. Hình được dự đoán được mã hóa với dự đoán liên khung hình từ hình cơ sở. Hình được dự đoán được mã hóa nhờ sử dụng các ô bị hạn chế chuyển động mà là các ô bị hạn chế liên lớp bổ sung. Các ô bị hạn chế liên lớp bị hạn chế theo cách là chỉ các mẫu và dữ liệu chuyển động của các ô đặt cùng chỗ trong hình cơ sở được sử dụng làm tham chiếu cho dự đoán liên khung hình. Một hay nhiều đường ghi chứa tập hợp ô hoặc các dòng bit được xây dựng từ hình được dự đoán.

Khi các đường ghi chứa tập hợp ô hoặc các lớp đặt cùng chỗ được xây dựng từ các hình, các đường ghi chứa tập hợp ô và dòng bit thu được tuân theo hình ảnh đầy đủ. Ví dụ, khi hình cơ sở tuân theo hồ sơ chính HEVC và hình được dự đoán tuân theo hồ sơ chính nhiều hình HEVC, các hình cơ sở và được dự đoán trong các đường ghi chứa tập hợp ô và dòng bit thu được cũng tuân theo hồ sơ chính HEVC và hồ sơ chính nhiều khung hình HEVC, tương ứng.

Bộ phát có thể chọn đường ghi chứa tập hợp ô hoặc dòng bit cần truyền trực tuyến và/hoặc được giải mã từ cả hai hình ví dụ dựa trên hướng nhìn hiện tại. Phương án này có thể phù hợp ví dụ các môi trường nhìn nơi các thay đổi nhanh của hướng nhìn là không thể hoặc sẽ không xảy ra.

Theo một phương án, có nhiều hơn một dòng bit nhiều hình được mã hóa, mỗi hình với độ phân giải khác nhau và/hoặc đặc tính khác, như tốc độ bit, chất lượng hình ảnh, độ sâu bit cho mỗi thành phần màu, dải tần nhạy sáng, và/hoặc độ rộng của giải màu. Đối với một hình hoặc cả hai hình, bộ phát chọn một hay nhiều tập hợp ô bị hạn chế chuyển động của hình cần được truyền trực tuyến và/hoặc được giải mã ở độ phân giải cao, trong khi một số hoặc tất cả các tập hợp ô còn lại (tức là không-đặt cùng chỗ) được truyền trực tuyến và/hoặc được giải mã ở độ phân giải thấp. Ví dụ, bộ phát có thể chọn một hay nhiều tập hợp ô bị hạn chế chuyển động để được truyền trực tuyến và/được giải mã ở độ phân giải cao dựa trên hướng nhìn hiện tại.

Một phương án để tạo gợi ý phân đoạn được minh họa trong Fig.17. Các bước "Việc mã hóa với các tập hợp ô bị hạn chế chuyển động" và "Gói tập/Phân đoạn" hoạt động giống như được mô tả theo các phương án khác. Bước "Tạo gợi ý phân

đoạn" tạo ra các lệnh (mà có thể cũng được gọi các gợi ý) theo đó phân đoạn hoặc phân đoạn con có thể được tạo ra cho các đường ghi chứa tập hợp ô. Các lệnh có thể ví dụ bao gồm danh mục các hàm dựng, mà được dự kiến để thực hiện tuần tự. Có thể có các loại khác nhau của các hàm dựng khả dựng, như hàm dựng nội tuyến chứa dữ liệu để chứa trong phân đoạn hoặc phân đoạn con được tạo ra theo các lệnh và bộ trích xuất sao chép dữ liệu bằng cách tham chiếu tệp hoặc đường ghi khác, tại vị trí byte được biểu thị (mà có thể được biểu thị nhờ nhận diện mẫu và vị trí byte trong mẫu), và phạm vi byte được biểu thị để được sao chép bằng cách tham khảo.

Theo một phương án, việc gói tệp/phân đoạn tạo ra các đường ghi chứa tập hợp ô đã được xây dựng, mà có thể tuân theo hình ảnh đầy đủ. Ngoài ra, việc gói tệp/phân đoạn tạo ra (các) đường ghi hình ảnh đầy đủ được xây dựng mà sử dụng các đường ghi chứa tập hợp ô đã được xây dựng như là tham chiếu cho kết cấu. Các lệnh có thể được lưu trong cùng tệp với (các) phân đoạn hoặc (các) tệp đa phương tiện, hoặc chúng có thể được lưu trong (các) tệp gợi ý phân đoạn riêng biệt. Định dạng của các lệnh có thể nhưng không cần phải tuân theo ISO BMFF (hoặc chung hơn, định dạng được sử dụng cho (các) phân đoạn hoặc (các) tệp đa phương tiện). Ví dụ, các lệnh có thể tạo ra đường ghi (mà có thể được gọi ví dụ đường ghi gợi ý phân đoạn MPEG-DASH) theo ISO BMFF, và mỗi mẫu của đường ghi có thể đề xuất các lệnh để xây dựng phân đoạn hoặc phân đoạn con.

Theo một phương án, MPD được tạo ra mà bao gồm thông tin trên các đường ghi chứa tập hợp ô và các phân đoạn hoặc các phân đoạn con của chúng, được xây dựng theo các gợi ý phân đoạn. Việc tạo ra MPD có thể tương tự như điều được mô tả sớm hơn trong phần "Tạo ra MPD".

Theo một phương án, đường ghi hình ảnh đầy đủ, nếu có, tương ứng với Hiện thị thứ nhất, đường ghi chứa tập hợp ô tương ứng với Hiện thị thứ hai. Khi tạo ra MPD, Hiện thị thứ nhất được biểu thị để ở trong bộ thích ứng thứ nhất và Hiện thị thứ hai được biểu thị để ở trong bộ thích ứng thứ hai.

Theo một phương án, SRD thứ nhất được chứa trong bộ thích ứng thứ nhất và SRD thứ hai được chứa trong bộ thích ứng thứ hai. SRD thứ nhất và SRD thứ hai biểu thị các bộ thích ứng để có cùng một không gian tham chiếu, ví dụ có cùng giá trị

của `source_id` như đã được giải thích trước đó. Vị trí và kích thước của tập hợp ô bị hạn chế chuyển động được thể hiện bởi đường ghi chứa tập hợp ô được biểu thị trong SRD thứ hai liên quan đến các kích thước của hình ảnh đầy đủ mà có thể được biểu thị trong SRD thứ nhất.

Theo một phương án, bộ phát kết luận là bộ thích ứng thứ nhất và bộ thích ứng thứ hai đại diện cho cùng một không gian tham chiếu, ví dụ dựa trên đó có cùng giá trị `source_id` như đã được giải thích trước đó. Bộ phát còn kết luận kích thước và vị trí tương đối của bộ thích ứng thứ nhất đối với bộ thích ứng thứ hai. Dựa trên ví dụ hướng nhìn thịnh hành và phạm vi quan sát được hiển thị thịnh hành, bộ phát chọn bộ thích ứng mà phù hợp nhất với các điều kiện xem của nó.

Fig.18 minh họa đường ghi gợi ý phân đoạn hoặc tệp phân đoạn với các hàm dựng trích xuất các tập hợp ô bị hạn chế chuyển động từ đường ghi hình ảnh đầy đủ.

Theo một phương án, thông tin URL có thể được bổ sung cho đường ghi gợi ý phân đoạn hoặc tệp phân đoạn, như vậy cho phép làm phù hợp phần đúng của đường ghi gợi ý phân đoạn (ví dụ mẫu gợi ý) hoặc tệp phân đoạn với URL được sử dụng trong yêu cầu và được kết luận từ thông tin MPD và/hoặc các phân đoạn (phụ) sớm hơn (ví dụ phạm vi byte thông tin). Thông tin URL có thể được chứa ví dụ nhờ sử dụng hệ thống mẫu URL tương tự như hoặc giống với thông tin được sử dụng trong hệ thống MPD hoặc bao gồm URL cho mỗi mỗi phân đoạn (phụ) mà có thể được tạo ra nhờ giải mã đường ghi gợi ý phân đoạn hoặc tệp phân đoạn.

Fig.19 minh họa việc vận hành máy chủ hoặc bộ phân đoạn, trong đó theo một phương án, máy chủ hoặc bộ phân đoạn mà được kết nối hoạt động được với máy chủ phân tích cú pháp các gợi ý của bộ phân đoạn. Bộ phân đoạn có thể ví dụ là tập lệnh phía máy chủ mà được giới thiệu bởi máy chủ HTTP. Máy chủ hoặc bộ phân đoạn phân tích cú pháp các gợi ý của bộ phân đoạn và tạo ra các phân đoạn hoặc các phân đoạn con theo các gợi ý được phân tích cú pháp. Phân đoạn hoặc các phân đoạn con là một phần của đường ghi chứa tập hợp ô được xây dựng trước dạng có hiệu lực.

Theo một phương án, máy chủ và/hoặc bộ phân đoạn nhận URL của phân đoạn (phụ) được yêu cầu, ví dụ thông qua yêu cầu HTTP GET. Máy chủ hoặc bộ phân đoạn phân tích cú pháp thông tin URL chứa trong đường ghi gợi ý phân đoạn hoặc tệp

phân đoạn để tìm ra các lệnh hoặc các gợi ý mà phù hợp với URL của phân đoạn (phụ) được yêu cầu. Máy chủ hoặc bộ phân đoạn sau đó hoàn thành các lệnh được tìm thấy hoặc các gợi ý để tạo ra phân đoạn (phụ) được yêu cầu. Máy chủ hoặc bộ phân đoạn sau đó đáp ứng yêu cầu bằng cách truyền phân đoạn (phụ) được tạo ra, ví dụ khi trả lời yêu cầu HTTP GET.

Theo một phương án, hướng dẫn hoặc gợi ý được đánh dấu cụ thể được tạo ra như là một phần của "Việc tạo ra gợi ý phân đoạn" đối với phân đoạn khởi tạo. Khi được giải mã, hướng dẫn hoặc gợi ý này tạo ra phân đoạn khởi tạo cho đường ghi chứa tập hợp ô được xây dựng trước. Phân đoạn khởi tạo có thể ví dụ bao gồm FileTypeBox và MovieBox.

Theo một phương án, việc gói tệp/phân đoạn tạo ra các đường ghi chứa tập hợp ô đã được xây dựng, mà có thể tuân theo hình ảnh đầy đủ. Ngoài ra, việc gói tệp/phân đoạn tạo ra (các) đường ghi hình ảnh đầy đủ được xây dựng mà sử dụng các đường ghi chứa tập hợp ô đã được xây dựng làm tham chiếu cho việc xây dựng.

Theo một phương án, MPD được tạo ra mà bao gồm thông tin trên các đường ghi chứa tập hợp ô đã được xây dựng, và (các) đường ghi hình ảnh đầy đủ được xây dựng mà sử dụng các đường ghi chứa tập hợp ô đã được xây dựng làm tham chiếu để xây dựng. Theo một phương án, MPD được tạo ra tương tự với cái được mô tả ở trên đối với (các) đường ghi chứa tập hợp ô được xây dựng theo các gợi ý của bộ phân đoạn và (các) đường ghi hình ảnh đầy đủ.

Theo một phương án, các mẫu của đoạn phim của các đường ghi chứa tập hợp ô đã được xây dựng được sắp xếp liên tục theo cách là các mẫu của đường ghi chứa tập hợp ô thứ nhất được xây dựng trước đứng trước các mẫu của đường ghi chứa tập hợp ô thứ hai được xây dựng trước và v. v.. Các mẫu của (các) đường ghi hình ảnh đầy đủ được xây dựng tiếp theo các mẫu của các đường ghi chứa tập hợp ô đã được xây dựng mà được sử dụng làm tham chiếu để xây dựng. Các ranh giới giữa các mẫu của các đường ghi khác nhau có thể được biểu thị ví dụ nhờ sử dụng SubsegmentIndexBox của ISOBMFF. LevelAssignmentBox hoặc tương tự có thể được chứa bởi tệp writer để biểu thị các cấp theo đường ghi được chỉ thị bởi SubsegmentIndexBox khi sử dụng.

Theo một phương án, mà được minh họa trong Fig.20, việc tạo ra MPD bao gồm thông tin SubRepresentations thành MPD, dựa trên các cấp được tạo ra cho (các) phân đoạn. Ví dụ, loại MIME và/hoặc thông tin quan hệ về không gian (ví dụ SRD) có thể được chứa trong MPD. Loại MIME được nêu ra đối với SubRepresentation tương ứng với đường ghi chứa tập hợp ô được xây dựng trước có thể biểu thị hồ sơ và mức độ đủ để giải mã tập hợp ô bị hạn chế chuyển động có liên quan. Thông tin có quan hệ về không gian đối với mỗi đường ghi chứa tập hợp ô được xây dựng trước có thể biểu thị vị trí và kích thước của nó liên quan tới kích thước hình ảnh đầy đủ được biểu thị. Thông tin có thể thực hiện nhờ sử dụng mức hiển thị phụ SRD, như được quy định trong MPEG-DASH.

Kết quả của quy trình trên được minh họa (với việc tham chiếu, nhưng không giới hạn ở, ISO/BMFF) trong Fig.21.

Theo một phương án, bộ phát nhận (các) phân đoạn như được mô tả ở trên hoạt động như sau: Bộ phát phân tích cú pháp thông tin mối quan hệ về không gian đối với các đường ghi chứa tập hợp ô đã được xây dựng, ví dụ từ mức hiển thị phụ SRD, như được quy định trong MPEG-DASH. Dựa trên thông tin mối quan hệ về không gian được phân tích cú pháp, bộ phát tự động xác định, ví dụ trên cơ sở phân đoạn hoặc phân đoạn con, mà yêu cầu các đường ghi chứa tập hợp ô đã được xây dựng. Bộ phát thu được sự bắt đầu của phân đoạn hoặc phân đoạn con, chứa SubsegmentIndexBox hoặc tương tự mà biểu thị các ranh giới byte giữa các mức, tức là giữa các đường ghi chứa tập hợp ô đã được xây dựng. Bộ phát kết luận các phạm vi byte cần được yêu cầu một cách phù hợp, và đưa ra các yêu cầu, như các yêu cầu HTTP GET, một cách phù hợp. Sự đáp ứng từ phía-máy chủ, bộ phát có được dữ liệu được yêu cầu đối với các đường ghi chứa tập hợp ô đã được xây dựng, mà sau đó được giải mã và có thể chạy.

Theo một phương án, hình cơ sở (ví dụ, hình trái) được mã hóa với các tập hợp ô bị hạn chế chuyển động. Dự đoán được mã hóa tương tự như vậy với các tập hợp ô bị hạn chế chuyển động. Các ranh giới của các tập hợp ô bị hạn chế chuyển động có thể được sắp hàng giữa các hình (nhưng nói chung không cần phải sắp hàng). Các tập hợp ô trong hình được dự đoán cũng bị hạn chế liên lớp sao cho dự đoán liên

lớp sử dụng các giá trị của mẫu chỉ trong tập hợp ô tương ứng của hình cơ sở. Dự đoán liên khung hình bị hạn chế đến mức chỉ tập hợp con của các hình ảnh của hình cơ sở có thể được sử dụng làm tham chiếu để dự đoán liên khung hình. Tập hợp con này có thể được gọi là tập hợp con tham chiếu hình xen kẽ. Ví dụ tập hợp con tham chiếu hình xen kẽ có thể bao gồm các hình ảnh IRAP của hình cơ sở, và các hình ảnh IRAP có thể được sắp hàng cắt ngang các hình, tức là khi có hình ảnh IRAP trong hình cơ sở, thì cũng có hình ảnh IRAP trong hình được dự đoán của cùng bộ truy nhập, và ngược lại.

Trong giai đoạn xây dựng đường ghi chứa tập hợp ô/dòng bit, một hay nhiều đường ghi chứa tập hợp ô hoặc các dòng bit được xây dựng từ hình cơ sở (ví dụ như được mô tả theo các phương án khác, ví dụ liên quan đến Fig.16). Một hay nhiều đường ghi chứa tập hợp ô hoặc các dòng bit được xây dựng từ hình được dự đoán (ví dụ như được mô tả theo các phương án khác, ví dụ liên quan đến Fig.16). Khi các đường ghi chứa tập hợp ô hoặc lớp đặt cùng chỗ được xây dựng từ các hình, các đường ghi chứa tập hợp ô và dòng bit thu được được tuân theo hình ảnh đầy đủ. Ví dụ, khi hình cơ sở tuân theo hồ sơ chính HEVC và hình được dự đoán tuân theo hồ sơ chính nhiều khung hình HEVC, các hình cơ sở và được dự đoán trong các đường ghi chứa tập hợp ô và dòng bit thu được cũng tuân theo hồ sơ chính HEVC và hồ sơ chính nhiều khung hình HEVC, tương ứng.

Phần hình cơ sở của đường ghi chứa tập hợp ô hoặc dòng bit có thể được trích xuất trừ khi đã được thực hiện trong giai đoạn trước.

Trong giai đoạn được gọi là "ghi lại lớp được dự đoán thành dòng bit hồ sơ đơn-hình", đường ghi chứa tập hợp ô hoặc dòng bit được xây dựng của hình được dự đoán được xử lý để tuân thủ hồ sơ đơn-hình. Đường ghi hình thứ hai được tạo ra bao gồm dữ liệu hình cơ sở bởi tham chiếu vào đường ghi hình thứ hai. Dữ liệu được chứa bởi tham chiếu này bao gồm toàn bộ hoặc một phần tập hợp con tham chiếu hình xen kẽ của các hình cơ sở từ đó các hình ảnh được mã hóa của hình được dự đoán có thể được dự đoán. Đường ghi hình thứ hai cũng bao gồm dữ liệu được mã hóa, nội tuyến (trong một số phương án) hoặc bởi tham chiếu (trong các phương án khác), của hình được dự đoán. Các mẫu của đường ghi hình thứ hai có thể bao gồm các hàm dựng, như

các bộ trích xuất và các hàm dựng nội tuyến. Đường ghi hình thứ hai được biểu thị để tuân thủ quy trình giải mã đơn-hình, như việc giải mã profin chính HEVC.

Các đầu ra của phương án là đường ghi chứa tập hợp ô hoặc dòng bit tuân theo hình ảnh đầy đủ của hình cơ sở và đường ghi thứ hai chứa tập hợp ô hoặc dòng bit tuân theo hình ảnh đầy đủ của hình thứ hai. Cả hai đường ghi chứa tập hợp ô hoặc các dòng bit tuân theo hình ảnh đầy đủ là có thể giải mã được, như vậy, nhờ bộ mã hóa video tiêu chuẩn, như bộ giải mã H.265/ HEVC. Theo cách khác, đầu ra có thể là các lệnh để xây dựng hình cơ sở và các đường ghi chứa tập hợp ô hoặc các dòng bit tuân theo hình ảnh đầy đủ của hình thứ hai.

Các giai đoạn của phương án có thể được thực hiện theo trình tự ngược lại, nghĩa là trình tự xây dựng đường ghi chứa tập hợp ô/dòng bit và việc ghi lại lớp được dự đoán thành dòng bit hồ sơ đơn-hình có thể đảo ngược. Nói cách khác, dòng bit của đơn-hình thứ nhất (của các hình ảnh đầy đủ) có thể được tạo ra và sau đó dòng bit chứa tập hợp ô của đơn-hình có thể được tạo ra.

Các giai đoạn của phương án có thể được kết hợp chủ yếu vào một thao tác. Ví dụ, cùng một đường ghi có thể được tạo ra để trích xuất các tập hợp ô của hình được dự đoán và bao gồm tập hợp con tham chiếu hình xen kẽ bởi tham chiếu. Nói cách khác, cùng đường ghi hoặc các mẫu có thể thực hiện cả hai việc xây dựng tập hợp ô và ghi lại dòng bit nhiều khung hình thành dòng bit đơn-hình.

Ở trên, một số các phương án được mô tả liên quan đến định dạng video hình ảnh toàn cảnh hình chữ nhật, trong phiên bản hoặc một chiều hoặc lập thể của nó. Cần phải hiểu là các phương án không bị giới hạn ở định dạng hình ảnh toàn cảnh hình chữ nhật. Các phương án tương tự có thể được áp dụng đối với hình ảnh hoặc định dạng video bất kỳ, bao gồm nhưng không giới hạn ở, ví dụ, các loại sau:

- Định dạng hình ảnh hai-chiều thông thường, ví dụ được chụp bởi các ống kính và cảm biến thông thường. Định dạng này có thể là trực tuyến.
- Định dạng hình ảnh kiểu mắt cá, ví dụ được chụp với ống kính kiểu mắt cá, đảm bảo phạm vi quan sát rộng, khoảng 180 độ
- Định dạng khối cắt, trong phiên bản một chiều hoặc lập thể của nó

- Định dạng gói-khung hình nhiều hình, trong đó các hình ảnh của các camera khác nhau được gói (như các khung hình cấu thành) vào cùng khung hình trước khi mã hóa và có thể được trích xuất từ các khung hình được giải mã sau khi giải mã. Các khung hình cấu thành có thể đại diện cho ví dụ định dạng bất kỳ trong các định dạng ở trên, ví dụ các hình ảnh hai-chiều thông thường, các hình ảnh kiểu mắt cá, các mặt bên của khối cắt từ các khối cắt một chiều hoặc lập thể, hoặc các hình của định dạng hình ảnh toàn cảnh lập thể.

Theo một phương án, phương án bất kỳ trong số các phương án được mô tả ở trên được áp dụng cho các hình ảnh đơn-hình hoặc kiểu mắt cá gói-khung hình theo cách là tập hợp ô bị hạn chế chuyển động được mã hóa đối với tập hợp con không gian của vùng hình ảnh kiểu mắt cá. Tương tự đối với (các) hình ảnh kiểu mắt cá, phương án có thể áp dụng đối với (các) loại hình ảnh có góc-nhìn rộng bất kỳ khác. Tập hợp con không gian có thể ví dụ đại diện cho vùng hình ảnh phù hợp với việc hiển thị trên màn hình tấm phẳng thông thường. Các ô khác bên ngoài tập hợp ô bị hạn chế chuyển động không cần phải bị hạn chế chuyển động, tức là có thể sử dụng dự đoán liên khung với tham chiếu đến các mẫu trong tập hợp ô bị hạn chế chuyển động trong hình ảnh tham chiếu. Đường ghi chứa tập hợp ô và/hoặc dòng bit đối với tập hợp ô bị hạn chế chuyển động được xây dựng, như được mô tả theo các phương án khác.

Bộ phát có thể chọn để truyền trực tuyến và/hoặc giải mã đường ghi hoặc dòng bit chứa tập hợp ô bị hạn chế chuyển động được xây dựng, ví dụ trên cơ sở của hiểu biết của thiết bị nhìn hoặc hiển thị có khả năng hiển thị phạm vi quan sát phù hợp tương đối hoặc chính xác với của tập hợp ô bị hạn chế chuyển động.

Theo một phương án mà có thể được áp dụng với các phương án được mô tả ở trên đối với các phương án giải mã một phần hình ảnh nhiều-độ phân giải có hoặc không có việc mã hóa video mở rộng được, định dạng hiển thị hình ảnh khác nhau có thể được sử dụng theo các độ phân giải hoặc các lớp khác nhau. Ví dụ, phiên bản độ phân giải thấp hoặc lớp cơ sở của nội dung có thể sử dụng định dạng hình ảnh toàn cảnh hình chữ nhật, trong khi phiên bản độ phân giải cao hoặc lớp được dự đoán của nội dung có thể sử dụng định dạng khối cắt.

Theo một phương án, các tập hợp ô bị hạn chế chuyển động được chọn khi

mã hóa để phù hợp với các ranh giới khung hình cấu thành của định dạng hiển thị gói-khung hình. Ví dụ, nếu khối cắt được thể hiện trong định dạng gói-khung hình, mỗi mặt bên của khối cắt là khung hình cấu thành riêng của nó, và mỗi khung hình cấu thành có thể được mã hóa như là tập hợp ô bị hạn chế chuyển động.

Theo một phương án, nội dung được mã hóa trong định dạng khối cắt theo ít nhất hai độ phân giải. Theo một phương án, bộ phát chọn các tập hợp ô bị hạn chế chuyển động của tập hợp con của các mặt bên của khối cắt có độ phân giải cao theo hướng nhìn chính. Giả định là việc nhìn được thực hiện với phạm vi quan sát của màn hình được lắp đầu thu khoảng 100 độ, bộ phát có thể ví dụ chọn đến ba mặt bên của khối cắt để được truyền trực tuyến và/hoặc được giải mã ở độ phân giải cao. Các mặt bên còn lại của khối cắt để được truyền trực tuyến và/hoặc được giải mã có thể được chọn từ phiên bản độ phân giải thấp hoặc có thể bỏ qua việc phát trực tuyến và/hoặc giải mã.

Theo một phương án, kích thước của các tập hợp ô bị hạn chế chuyển động được chọn để mã hóa trên cơ sở phạm vi quan sát mong đợi hoặc đã biết của thiết bị nhìn và/hoặc thói quen nhìn. Ví dụ, có thể giả định là việc định hướng đầu theo chiều đứng thông thường khi nhìn với màn hình lắp đầu thu là khoảng trên đường xích đạo của video hình cầu.

Theo một phương án, video toàn cảnh 360-độ được phân chia thành 12 ô bị hạn chế chuyển động, mỗi ô bao phủ phạm vi quan sát ngang 90-độ (tức là bốn cột ô có chiều rộng bằng nhau), hàng ô cao nhất bao phủ phạm vi quan sát đứng 45 độ, hàng ô trung tâm bao phủ phạm vi quan sát đứng 90 độ, và hàng ô dưới cùng bao phủ phạm vi quan sát đứng 45 độ. Giả định là phạm vi quan sát trên màn hình lắp đầu thu là gần 100 độ, việc phát trực tuyến và/hoặc giải mã bốn tập hợp ô thường là đủ để bao phủ phạm vi quan sát cần thiết để hiển thị, đặc biệt khi phiên bản hoặc lớp cơ sở độ phân giải thấp có thể được sử dụng để bao phủ các vùng khuyết có thể có bất kỳ trên các mặt bên của các hình ảnh được hiển thị.

Theo một phương án, video toàn cảnh 360-độ được phân chia thành 8 ô bị hạn chế chuyển động, mỗi ô bao phủ 90 độ ngang và đứng. Giả định là phạm vi quan sát trên màn hình lắp đầu thu là gần 100 độ, việc phát trực tuyến và/hoặc giải mã bốn tập

hợp ô thường là đủ để bao phủ phạm vi quan sát cần thiết để hiển thị, đặc biệt khi phiên bản hoặc lớp cơ sở độ phân giải thấp có thể được sử dụng để bao phủ các vùng khuyết có thể có bất kỳ trên các mặt bên của các hình ảnh được hiển thị.

Theo một phương án, video toàn cảnh 360-độ được phân chia thành 12 ô bị hạn chế chuyển động, mỗi ô bao phủ phạm vi quan sát ngang 90-độ (tức là bốn cột ô có chiều rộng bằng nhau), hàng ô cao nhất bao phủ phạm vi quan sát đứng 30 độ, hàng ô trung tâm bao phủ phạm vi quan sát đứng 120 độ, và hàng ô dưới cùng bao phủ phạm vi quan sát đứng 30 độ. Giả định là phạm vi quan sát trên màn hình laptop đầu thu là gần 100 độ và nó nhìn gần như thẳng về trước (thay vì nhìn lên hoặc nhìn xuống), việc phát trực tuyến và/hoặc giải mã hai tập hợp ô thường là đủ để bao phủ phạm vi quan sát cần để hiển thị, đặc biệt khi phiên bản hoặc lớp cơ sở độ phân giải thấp có thể được sử dụng để bao phủ các vùng khuyết có thể có bất kỳ trên các mặt bên của các hình ảnh được hiển thị.

Cần phải hiểu là các giá trị của phạm vi quan sát trong các phương án ở trên của các phân vùng ô bị hạn chế chuyển động tiềm năng là các ví dụ và là các giá trị chính xác có thể được chọn ví dụ theo cách mà các hàng và/hoặc các cột CTU phù hợp các ranh giới ô bị hạn chế chuyển động. Các phương án không bị giới hạn ở số lượng của các ô hoặc các giá trị của phạm vi quan sát được trình bày.

Trong một số phương án, tập hợp của các ô cùng tạo thành tập hợp ô bị hạn chế chuyển động. Theo một phương án có thể áp dụng ví dụ các phân vùng 12-ô như được mô tả ở trên, hàng ô trên cùng tạo thành một tập hợp ô bị hạn chế chuyển động (trong khi các ô của hàng ô trên cùng không cần bị hạn chế chuyển động riêng biệt), và/hoặc hàng ô dưới cùng tạo thành một tập hợp ô bị hạn chế chuyển động (trong khi các ô của hàng ô dưới cùng không cần bị hạn chế chuyển động riêng biệt).

Lưu ý là một số tập hợp ô có thể được kết hợp thành cùng đường ghi chứa tập hợp ô hoặc dòng bit được xây dựng. Theo một phương án, tập hợp ô thứ nhất bị hạn chế chuyển động và tập hợp ô thứ hai bị hạn chế chuyển động mà không liên kết với mỗi tập hợp ô khác được kết hợp thành cùng đường ghi chứa tập hợp ô hoặc dòng bit được xây dựng. Theo một phương án, ô cực-trái hoặc tập hợp ô của dòng bit hình ảnh đầy đủ được kết hợp vào cùng đường ghi chứa tập hợp ô hoặc dòng bit được xây

dựng với ô ngoài cùng bên phải hoặc tập hợp ô của dòng bit hình ảnh đầy đủ. Sự kết hợp như vậy có thể đặc biệt phù hợp video toàn cảnh 360-độ, trong trường hợp đó ô ngoài cùng bên phải có thể được đặt trong đường ghi chứa tập hợp ô hoặc dòng bit được xây dựng bên trái của ô cực-trái của cùng hàng ô, để duy trì tính liên tục về không gian.

Fig.22 thể hiện sơ đồ khối của bộ mã hóa video thích hợp để ứng dụng các phương án của sáng chế. Fig.22 mô tả cấu trúc của bộ giải mã hai-lớp, nhưng nó có thể được ưa thích là các thao tác giải mã có thể giống như được sử dụng trong bộ giải mã đơn-lớp.

Bộ mã hóa video 550 bao gồm bộ giải mã thứ nhất 552 cho các thành phần hình cơ sở và bộ giải mã thứ hai 554 cho các thành phần không có hình cơ sở. Khối 556 minh họa bộ tách kênh để cung cấp thông tin về các thành phần hình cơ sở cho bộ giải mã thứ nhất 552 và để cung cấp thông tin về các thành phần không có hình cơ sở cho bộ giải mã thứ hai 554. Tham chiếu P'n là viết tắt của hiển thị được dự đoán của khối hình ảnh. Tham chiếu D'n là viết tắt của tín hiệu lỗi dự đoán được phục hồi. Các khối 704, 804 minh họa hình ảnh được phục hồi sơ bộ ($I'n$). Tham chiếu R'n là viết tắt của hình ảnh được phục hồi cuối cùng. Các khối 703, 803 minh họa biến đổi nghịch đảo (T^{-1}). Các khối 702, 802 minh họa sự lượng tử hóa nghịch đảo (Q^{-1}). Các khối 701, 801 minh họa giải mã entropy (E^{-1}). Các khối 705, 805 minh họa bộ nhớ khung hình tham chiếu (RFM). Các khối 706, 806 minh họa dự đoán (P) (hoặc dự đoán liên khung hoặc dự đoán xen kẽ). Các khối 707, 807 minh họa việc lọc (F). Các khối 708, 808 có thể được sử dụng để kết hợp thông tin lỗi dự đoán được giải mã với các thành phần hình cơ sở/không-hình cơ sở được dự đoán để thu được các hình ảnh được phục hồi sơ bộ ($I'n$). Các hình ảnh của hình cơ sở được phục hồi sơ bộ và được lọc có thể là đầu ra 709 từ bộ phận giải mã thứ nhất 552 và được phục hồi sơ bộ và được lọc và các hình ảnh của hình cơ sở được lọc có thể là đầu ra 809 từ bộ phận giải mã thứ nhất 554.

Ở đây, bộ giải mã cần được diễn giải để đảm bảo phạm vi hành bất kỳ có thể thực hiện các thao tác giải mã, như bộ phát, bộ thu, công vào, bộ tách kênh và/hoặc bộ giải mã.

Fig.23 là biểu diễn đồ họa của ví dụ hệ thống truyền thông đa phương tiện

trong đó các phương án khác nhau có thể được sử dụng. Nguồn dữ liệu 1510 tạo ra tín hiệu nguồn theo analog, số không nén, hoặc định dạng số được nén, hoặc sự kết hợp bất kỳ của các các định dạng này. Bộ mã hóa 1520 có thể bao gồm hoặc được kết nối với bộ xử lý trước, như sự chuyển đổi định dạng dữ liệu và/hoặc lọc tín hiệu nguồn. Bộ mã hóa 1520 mã hóa tín hiệu nguồn thành dòng bit phương tiện được mã hóa. Cần lưu ý là dòng bit cần được giải mã có thể được nhận trực tiếp hoặc gián tiếp từ thiết bị điều khiển từ xa đặt trong hầu như loại mạng bất kỳ. Ngoài ra, dòng bit có thể được nhận từ phần cứng hoặc phần mềm cục bộ. Bộ mã hóa 1520 có thể có khả năng mã hóa nhiều hơn một loại phương tiện, như audio và video, hoặc nhiều hơn một bộ mã hóa 1520 có thể được yêu cầu để mã hóa các loại phương tiện khác nhau của tín hiệu nguồn. Bộ mã hóa 1520 có thể cũng là đầu vào được tạo ra theo cách tổng hợp, như đồ họa và văn bản, hoặc nó có thể có khả năng tạo ra các dòng bit được mã hóa của phương tiện tổng hợp. Tiếp theo, chỉ xử lý một dòng bit phương tiện được mã hóa của một loại phương tiện được coi là để đơn giản hóa việc mô tả. Tuy nhiên, cần phải lưu ý là các dịch vụ phát sóng thời gian thực thông thường bao gồm một số dòng (thường ít nhất một dòng audio, video và phụ đề văn bản). Cũng cần lưu ý là hệ thống này có thể bao gồm nhiều bộ mã hóa, nhưng trong hình vẽ chỉ một bộ mã hóa 1520 được thể hiện để đơn giản hóa bản mô tả mà không làm mất tính tổng quát. Cần phải hiểu thêm là, mặc dù văn bản và các ví dụ được bao hàm ở đây có thể mô tả cụ thể quy trình mã hóa, người có kỹ năng trong lĩnh vực kỹ thuật có thể hiểu là cùng các khái niệm và các nguyên tắc giống nhau cũng áp dụng cho quy trình giải mã tương ứng và ngược lại.

Dòng bit phương tiện được mã hóa có thể được chuyển sang bộ lưu trữ 1530. Bộ lưu trữ 1530 có thể bao gồm loại bộ nhớ thứ cấp bất kỳ để lưu trữ dòng bit phương tiện được mã hóa. Định dạng của dòng bit phương tiện được mã hóa trong bộ lưu trữ 1530 có thể là định dạng dòng bit khép kín cơ bản, hoặc một hay nhiều dòng bit phương tiện được mã hóa có thể được gói vào tệp chứa, hoặc dòng bit phương tiện được mã hóa có thể được gói vào định dạng phân đoạn thích hợp đối với DASH (hoặc hệ thống tương tự như phát trực tuyến) và được lưu như là chuỗi của các phân đoạn. Nếu một hay nhiều dòng bit phương tiện được gói trong tệp chứa, bộ tạo tệp (không thể hiện trong hình vẽ) có thể được sử dụng để lưu trữ thêm một dòng bit phương tiện

trong tệp và tạo ra siêu dữ liệu định dạng tệp, mà có thể cũng được lưu trong tệp. Bộ mã hóa 1520 hoặc bộ lưu trữ 1530 có thể bao gồm bộ tạo tệp, hoặc bộ tạo tệp được gắn thao tác được vào hoặc bộ mã hóa 1520 hoặc bộ lưu trữ 1530. Một số hệ thống vận hành “live”, tức là bỏ qua việc lưu và chuyển dòng bit phương tiện được mã hóa từ bộ mã hóa 1520 trực tiếp đến bộ phận chuyển 1540. Sau đó, dòng bit phương tiện được mã hóa có thể được chuyển đến bộ phận chuyển 1540, cũng được gọi là máy chủ, trên nền tảng cần thiết. Định dạng được sử dụng khi biến đổi có thể là định dạng dòng bit dòng khép kín cơ bản, định dạng dòng gói, định dạng phân đoạn thích hợp đối với DASH (hoặc hệ thống phát trực tuyến tương tự), hoặc một hay nhiều dòng bit phương tiện được mã hóa có thể được gói vào tệp chứa. Bộ mã hóa 1520, bộ lưu trữ 1530, và máy chủ 1540 có thể lưu trữ trong cùng một thiết bị vật lý hoặc chúng có thể được chứa trong các thiết bị riêng biệt. Bộ mã hóa 1520 và máy chủ 1540 có thể vận hành với nội dung trực tiếp thời gian thực, trong trường hợp đó dòng bit phương tiện được mã hóa thường không được lưu vĩnh viễn, mà đúng hơn là được đệm cho các giai đoạn thời gian ngắn trong bộ mã hóa nội dung 1520 và/hoặc trong máy chủ 1540 để làm thông suốt các biến thể khi chậm xử lý, chậm chuyển, và tốc độ bit phương tiện được mã hóa.

Máy chủ 1540 gửi dòng bit phương tiện được mã hóa nhờ sử dụng chùm giao thức truyền thông. Chùm này có thể bao gồm, nhưng không giới hạn ở, một hay nhiều Real-Time Transport Protocol (RTP), User Datagram Protocol (UDP), Hypertext Transfer Protocol (HTTP), Transmission Control Protocol (TCP), và Internet Protocol (IP). Khi chùm giao thức truyền thông được định hướng gói, máy chủ 1540 gói dòng bit phương tiện được mã hóa thành các gói. Ví dụ, khi RTP được sử dụng, máy chủ 1540 gói dòng bit phương tiện được mã hóa vào các gói RTP theo định dạng tải RTP. Thông thường, mỗi loại phương tiện có định dạng tải RTP chuyên dụng. Có thể lưu ý lần nữa là hệ thống có thể chứa nhiều hơn một máy chủ 1540, nhưng để đơn giản, phần mô tả sau chỉ xem xét một máy chủ 1540.

Nếu nội dung phương tiện được gói trong tệp chứa cho bộ lưu trữ 1530 hoặc để nhập dữ liệu vào bộ phận chuyển 1540, bộ phận chuyển 1540 có thể bao gồm hoặc được gắn theo cách thao tác được vào "bước gửi tệp chương trình phân tích cú pháp"

(không thể hiện trong hình vẽ). Cụ thể, nếu tệp chứa không được truyền như vậy nhưng ít nhất một trong các dòng bit phương tiện được mã hóa được gói để chuyển qua giao thức truyền thông, bước gửi tệp chương trình phân tích cú pháp tệp gửi đặt các phần thích hợp của dòng bit phương tiện được mã hóa cần được chuyển qua giao thức truyền thông. bước gửi tệp chương trình phân tích cú pháp tệp gửi có thể cũng giúp trong việc tạo ra định dạng đúng đối với giao thức truyền thông, như gói các phần đầu và tải. tệp chứa đa truyền thông có thể chứa các lệnh cách gói, như đường ghi đánh dấu trong ISOBMFF, để gói ít nhất một dòng bit phương tiện được chứa trên giao thức truyền thông.

Máy chủ 1540 có thể hoặc không được kết nối vào cổng vào 1550 thông qua mạng truyền thông, mà có thể ví dụ là sự kết hợp của CDN, internet và/hoặc một hay nhiều mạng truy cập. Cổng vào có thể hoặc theo cách khác còn được gọi là hộp trung gian. Đối với DASH, cổng vào có thể là máy chủ cạnh (của CDN) hoặc web proxy. Nói chung, cần lưu ý là hệ thống có thể bao gồm số lượng cổng vào bất kỳ hoặc tương tự, nhưng để đơn giản, phần mô tả sau chỉ xem xét một cổng vào 1550. Cổng vào 1550 có thể thực hiện các loại khác nhau của các chức năng, như phiên dịch dòng gói theo một chùm giao thức truyền thông sang chùm giao thức truyền thông khác, hợp nhất và chia nhánh các dòng dữ liệu, và xử lý dòng dữ liệu theo tuyến xuống và/hoặc các khả năng tiếp nhận, như điều khiển tốc độ bit của dòng chuyển tiếp theo các điều kiện tuyến xuống thịnh hành.

Hệ thống bao gồm một hay nhiều bộ thu 1560, thường có thể nhận, giải điều biến, và mở gói tín hiệu được truyền thành dòng bit phương tiện được mã hóa. Dòng bit phương tiện được mã hóa có thể được chuyển cho bộ lưu trữ ghi 1570. Bộ lưu trữ ghi 1570 có thể bao gồm loại bộ nhớ thứ cấp bất kỳ để lưu dòng bit phương tiện được mã hóa. Bộ lưu trữ ghi 1570 có thể theo cách khác hoặc ngoài ra bao gồm bộ nhớ tính toán, như bộ nhớ truy cập ngẫu nhiên. Định dạng của dòng bit phương tiện được mã hóa trong bộ lưu trữ ghi 1570 có thể là định dạng dòng bit dòng khép kín cơ bản, hoặc một hay nhiều dòng bit phương tiện được mã hóa có thể được gói vào tệp chứa. Nếu có nhiều dòng bit phương tiện được mã hóa, như dòng audio và dòng video, đi kèm với nhau, tệp chứa thường được sử dụng và bộ thu 1560 bao gồm hoặc được gắn vào

bộ tạo tệp chứa tạo ra tệp chứa từ các dòng đầu vào. Một số hệ thống vận hành “live,” tức là bỏ sót bộ lưu trữ ghi 1570 và chuyển dòng bit phương tiện được mã hóa từ bộ thu 1560 trực tiếp đến bộ giải mã 1580. Trong một số hệ thống, chỉ phần đầu nhất của dòng được ghi, ví dụ, trích đoạn 10 phút gần đây nhất của dòng được ghi, được giữ trong bộ lưu trữ ghi 1570, trong khi dữ liệu được ghi sớm hơn bất kỳ được loại bỏ từ bộ lưu trữ ghi 1570.

Dòng bit phương tiện được mã hóa có thể được chuyển từ bộ lưu trữ ghi 1570 đến bộ giải mã 1580. Nếu có nhiều dòng bit phương tiện được mã hóa, như dòng audio và dòng video, đi kèm với nhau và được gói vào tệp chứa hoặc một dòng bit phương tiện được gói trong tệp chứa ví dụ để dễ truy cập, tệp chương trình phân tích cú pháp (không thể hiện trong hình vẽ) được sử dụng để mở gói mỗi dòng bit phương tiện được mã hóa từ tệp chứa. Bộ lưu trữ ghi 1570 hoặc bộ giải mã 1580 có thể bao gồm tệp chương trình phân tích cú pháp, hoặc tệp chương trình phân tích cú pháp được gắn kèm vào hoặc bộ lưu trữ ghi 1570 hoặc bộ giải mã 1580. Cũng cần lưu ý là hệ thống có thể bao gồm nhiều bộ giải mã, nhưng ở đây chỉ một bộ giải mã 1580 được thảo luận để làm đơn giản phần mô tả mà không làm mất đi tính tổng quát.

Dòng bit phương tiện được mã hóa có thể được xử lý tiếp nhờ bộ giải mã 1580, đầu ra của nó là một hay nhiều dòng phương tiện không nén. Cuối cùng, bộ kết xuất 1590 có thể sao chép lại dòng phương tiện không nén với loa âm thanh hoặc màn hình, ví dụ. Bộ thu 1560, bộ lưu trữ ghi 1570, bộ giải mã 1580, và bộ kết xuất 1590 có thể cư trú trong cùng thiết bị vật lý hoặc chúng có thể chứa trong các thiết bị riêng biệt.

Bộ phận chuyển 1540 và/hoặc cổng vào 1550 có thể được cấu hình để thực hiện chuyển giữa các sự biểu diễn khác nhau ví dụ để chuyển hình, sự thích ứng tốc độ bit và/hoặc khởi động nhanh, và/hoặc bộ phận chuyển 1540 và/hoặc cổng vào 1550 có thể được cấu hình để chọn (các) sự biểu diễn được truyền. Việc chuyển giữa các sự biểu diễn khác nhau có thể thực hiện vì nhiều lý do, như đáp ứng các yêu cầu của bộ thu 1560 hoặc các điều kiện thường xuyên, như thông lượng, của mạng đối với nó dòng bit được chuyển. Yêu cầu từ bộ thu có thể, ví dụ, là yêu cầu đối với phân đoạn hoặc phân đoạn con từ hiển thị khác với cái trước, yêu cầu về sự thay đổi của các lớp

có thể mở rộng được truyền và/hoặc lớp-phụ, hoặc sự thay đổi của thiết bị dựng hình có khả năng khác nhau được so với cái trước. Yêu cầu đối với phân đoạn có thể là yêu cầu HTTP GET. Yêu cầu đối với phân đoạn con có thể là yêu cầu HTTP GET với phạm vi byte. Ngoài ra hoặc theo cách khác, việc điều chỉnh tốc độ bit hoặc sự thích ứng tốc độ bit có thể được sử dụng ví dụ để tạo ra cái được gọi là khởi động-nhanh trong các thiết bị phát trực tuyến, nơi tốc độ bit của dòng được truyền thấp hơn tốc độ bit kênh sau khi bắt đầu hoặc truy cập ngẫu nhiên việc phát trực tuyến để bắt đầu phát lại ngay và đạt được mức chiếm dụng bộ đệm mà thỉnh thoảng chịu trì hoãn gói tin và/hoặc truyền lại. Sự thích ứng tốc độ bit có thể bao gồm nhiều Hiển thị hoặc chuyển lớp lên và Hiển thị hoặc các thao tác chuyển lớp xuống thực hiện theo các trình tự khác nhau.

Bộ giải mã 1580 có thể được cấu hình để thực hiện chuyển giữa các Hiển thị khác nhau ví dụ để chuyển hình, sự thích ứng tốc độ bit và/hoặc khởi động-nhanh, và/hoặc bộ giải mã 1580 có thể được cấu hình để chọn (các) Hiển thị được truyền. Việc chuyển giữa (các) Hiển thị khác nhau có thể thực hiện do nhiều nguyên nhân, như đạt được thao tác giải mã nhanh hơn hoặc làm thích nghi dòng bit được truyền, ví dụ về mặt tốc độ bit, với các điều kiện hiện hành, như thông lượng, của mạng trên đó dòng bit được chuyển. Thao tác giải mã nhanh hơn có thể là cần thiết ví dụ nếu thiết bị bao gồm bộ giải mã 1580 là đa-nhiệm và sử dụng nguồn tính toán cho các mục đích khác với việc giải mã dòng bit video mở rộng được. Trong ví dụ khác, thao tác giải mã nhanh hơn có thể cần nếu nội dung được phát ngược tại tốc độ nhanh hơn tốc độ phát lại bình thường, ví dụ hai hoặc ba lần nhanh hơn so với tốc độ phát lại thời gian thực thông thường. Tốc độ của bộ thao tác giải mã có thể thay đổi trong quá trình giải mã hoặc phát lại ví dụ như đáp ứng sự thay đổi từ phát nhanh về trước từ tốc độ phát lại bình thường hoặc ngược lại, và do đó các thao tác chuyển đổi nhiều lớp lên và chuyển lớp xuống có thể thực hiện theo các trình tự khác nhau.

Như đã nêu ở trên, một số phương án được mô tả với tham chiếu tới phân đoạn mảng tạm thời. Cần phải được hiểu là các phương án áp dụng tương tự cho các bộ phận phân chia hình ảnh cùng loại khác. Ví dụ, một số hệ thống mã hóa có thể không bao gồm bối cảnh của các phân đoạn mảng nhưng có thể có bối cảnh của các

mảng, như được xác định trong các tiêu chuẩn mã hóa video bao gồm H.264/AVC và HEVC, trong trường hợp đó, các phương án áp dụng cho các mảng.

Như đã nêu ở trên, một số phương án được mô tả liên quan đến các đường ghi chứa tập hợp ô, các dòng bit, và/hoặc lớp có độ phân giải cao và/hoặc độ phân giải thấp. Cần phải hiểu rằng thay vì hoặc ngoài ra để phân biệt về độ phân giải không gian, độ phân giải cao thời hạn có thể được sử dụng trong các phương án để biểu thị kết quả thành nhận thức về chất lượng cao hơn so với độ phân giải thấp thời hạn, ví dụ về một hay nhiều mục sau:

- Chất lượng hình ảnh, mà có thể thu được nhờ sử dụng các giá trị thông số lượng tử hóa khác nhau
- Độ sâu bit cho mỗi thành phần màu
- Dải tần nhạy sáng
- Độ rộng của giải màu

Như đã nêu ở trên, một số phương án được mô tả liên quan đến hai lớp. Cần phải hiểu là các phương án có thể áp dụng tương tự cho nhiều hơn hai lớp.

Như đã nêu ở trên, một số phương án được mô tả liên quan đến HEVC và/hoặc các điều khoản được sử dụng trong thông số kỹ thuật HEVC. Cần phải được hiểu là các phương án áp dụng tương tự các bộ mã hóa-giải mã khác và việc mã hóa các định dạng và thuật ngữ khác với nghĩa tương đương hoặc giống với các điều khoản được sử dụng trong các phương án được mô tả ở trên.

Như đã nêu ở trên, một số phương án được mô tả liên quan đến ISO/BMFF và/hoặc các định dạng xuất phát từ ISO/BMFF. Cần phải hiểu là các phương án áp dụng tương tự đối với tệp khác và các định dạng của phân đoạn, như định dạng tệp Matroska.

Như đã nêu ở trên, một số phương án được mô tả liên quan đến MPEG-DASH hoặc DASH. Cần phải hiểu là các phương án áp dụng tương tự cho các dạng phát trực tuyến khác đối với HTTP, như phát sóng trực tiếp HTTP của Apple (Apple HTTP Live Streaming - HLS).

Như đã nêu ở trên, một số phương án được mô tả bằng cách viện dẫn đến

việc phát trực tuyến có giới hạn. Cần phải hiểu là các phương án áp dụng tương tự cho các dạng biến đổi video khác, như tải lũy tiến, chuyển tệp, và các giao tiếp video đàm thoại, như điện thoại video.

Như đã nêu ở trên, nơi các phương án ví dụ được mô tả với tham chiếu bộ mã hóa, cần phải hiểu là dòng bit thu được và bộ giải mã có thể có các phần tử tương ứng trong đó. Tương tự như vậy, các phương án ví dụ được mô tả với việc tham chiếu bộ giải mã, cần phải hiểu là bộ mã hóa có thể có cấu trúc và/hoặc chương trình máy tính để tạo ra dòng bit cần được giải mã bởi bộ giải mã.

Các phương án của sáng chế được mô tả ở trên mô tả bộ giải mã về thiết bị mã hóa và thiết bị giải mã riêng biệt để hỗ trợ việc hiểu các quy trình liên quan. Tuy nhiên, sẽ thích hợp là thiết bị, các cấu trúc và các thao tác có thể được sử dụng như là thiết bị mã hóa - giải mã/cấu trúc/thao tác đơn lẻ. Ngoài ra, có khả năng là bộ mã hóa và bộ giải mã có thể chia sẻ một số hoặc tất cả các phần tử chung.

Mặc dù các ví dụ trên mô tả các phương án của sáng chế vận hành bên trong bộ giải mã trong thiết bị điện tử, có thể thích hợp là sáng chế như được xác định trong các điểm yêu cầu bảo hộ có thể được sử dụng như một phần của mã video bất kỳ. Do đó, ví dụ, các phương án của sáng chế có thể được sử dụng trong bộ mã hóa-giải mã video mà có thể sử dụng để mã hóa video trên các kênh truyền thông cố định hoặc hữu tuyến.

Do đó, thiết bị người dùng có thể bao gồm bộ mã hóa-giải mã video như được mô tả trong các phương án của sáng chế ở trên. Sẽ tốt hơn là thiết bị người dùng kỳ hạn được dự kiến để bao gồm loại thiết bị người dùng không dây thích hợp bất kỳ, như các loại điện thoại di động, các thiết bị xử lý dữ liệu cầm tay hoặc các trình duyệt web cầm tay.

Ngoài ra, các phần tử của mạng di động đất liền công cộng (PLMN) có thể cũng bao gồm các bộ mã hóa-giải mã video như được mô tả ở trên.

Nói chung, các phương án khác nhau của sáng chế có thể được sử dụng trong phần cứng hoặc các mạch mục đích đặc thù, phần mềm, logic hoặc sự kết hợp bất kỳ của chúng. Ví dụ, một số các khía cạnh có thể được sử dụng trong phần cứng, trong khi các khía cạnh khác có thể được sử dụng trong phần sụn hoặc phần mềm mà có thể

thực hiện bởi bộ điều khiển, bộ vi xử lý hoặc máy tính khác, mặc dù sáng chế không bị giới hạn ở đó. Trong khi các khía cạnh khác nhau của sáng chế có thể được minh họa và được mô tả như các sơ đồ khối, các lưu đồ, hoặc nhờ sử dụng một số hình ảnh khác, nó được hiểu rõ là các khối, thiết bị, các hệ thống, các kỹ thuật hoặc các phương pháp được mô tả ở đây có thể được sử dụng trong, như là các ví dụ không nhằm mục đích hạn chế, phần cứng, phần mềm, phần sụn, các mạch mục đích đặc thù hoặc logic, phần cứng đa năng hoặc bộ điều khiển hoặc các thiết bị tính toán khác, hoặc một số kết hợp của chúng.

Các phương án của sáng chế này có thể được sử dụng bởi phần mềm máy tính có thể thực thi bởi bộ xử lý dữ liệu của thiết bị di động, như trong bộ xử lý, hoặc bởi phần cứng, hoặc bởi sự kết hợp của phần mềm và phần cứng. Ngoài ra, về vấn đề này cần phải hiểu là các khối bất kỳ của dòng logic như trong các hình vẽ có thể đại diện cho các bước của chương trình, hoặc các mạch logic được kết nối với nhau, các khối và các chức năng, hoặc sự kết hợp của các bước của chương trình và các mạch logic, các khối và các chức năng. Phần mềm có thể được lưu trên các phương tiện vật lý như các chip nhớ, hoặc các khối bộ nhớ được sử dụng trong bộ xử lý, các phương tiện từ tính như đĩa cứng hoặc đĩa mềm, và phương tiện quang học như ví dụ DVD và các biến thể dữ liệu của chúng, CD.

Bộ nhớ có thể là loại bất kỳ thích hợp với môi trường kỹ thuật địa phương và có thể được sử dụng nhờ sử dụng công nghệ lưu trữ dữ liệu thích hợp bất kỳ, như các thiết bị nhớ bán dẫn, các thiết bị và các hệ thống nhớ từ tính, các thiết bị và hệ thống nhớ quang học, các bộ nhớ cố định và các bộ nhớ tháo rời được. Các bộ xử lý dữ liệu có thể là loại bất kỳ thích hợp đối với môi trường kỹ thuật địa phương, và có thể bao gồm một hay nhiều máy tính đa năng, máy tính chuyên dụng, các bộ xử lý, các bộ xử lý tín hiệu số (DSPs) và các bộ xử lý dựa trên kiến trúc bộ xử lý đa lõi, như các ví dụ không nhằm hạn chế.

Các phương án của sáng chế có thể được thực hành trong các thành phần khác nhau như các môđun mạch tích hợp. Thiết kế của các mạch tích hợp là bởi quy trình được tự động hóa cao. Các công cụ phần mềm phức tạp và mạnh là khả dụng để biến thiết kế mức logic thành thiết kế mạch bán dẫn sẵn sàng để được khắc và được

tạo thành trên nền bán dẫn.

Các chương trình, như được cung cấp bởi Synopsys, Inc. của Mountain View, California và Cadence Design, của San Jose, California tự động dẫn dây dẫn và đặt các thành phần lên chip bán dẫn nhờ sử dụng các quy tắc được thiết lập chuẩn của bản thiết kế cũng như các thư viện của các môđun thiết kế được lưu trước. Khi thiết kế cho mạch bán dẫn được hoàn thành, thiết kế thu được, theo định dạng điện tử được tiêu chuẩn hóa (ví dụ, Opus, GDSII, hoặc tương tự) có thể được chuyển đến nhà máy sản xuất bán dẫn hoặc "fab" để sản xuất.

Phần mô tả được cung cấp nhờ các ví dụ để dẫn chứng và không nhằm giới hạn phần mô tả đầy đủ và có nhiều thông tin của phương án để dẫn chứng của sáng chế. Tuy nhiên, các phương án cải biến và các thích ứng có thể trở nên rõ ràng đối với người có hiểu biết trung bình về lĩnh vực kỹ thuật liên quan theo quan điểm của bản mô tả trên, khi đọc kết hợp với các hình vẽ kèm theo và các điểm yêu cầu bảo hộ. Tuy nhiên, tất cả các cải biến như vậy và tương tự với ý tưởng của sáng chế này vẫn thuộc phạm vi của yêu cầu bảo hộ.

YÊU CẦU BẢO HỘ

1. Thiết bị mã hóa và giải mã video, thiết bị này bao gồm:

ít nhất một bộ xử lý; và

ít nhất một bộ nhớ lâu dài bao gồm mã chương trình máy tính, mà khi thực hiện bởi ít nhất một bộ xử lý, khiến cho thiết bị:

thu được dòng bit của hình ảnh đầy đủ bao gồm:

ảnh được mã hóa, mỗi ảnh được mã hóa này bao gồm lưới của bộ cây mã hóa và được phân chia thành một hoặc nhiều ô chứa số nguyên của bộ cây mã hóa tạo thành hình chữ nhật, và

tập hợp ô bị hạn chế chuyển động bao gồm một hoặc nhiều ô trên mỗi ảnh được mã hóa theo cách mà không có giá trị mẫu bên ngoài tập hợp ô bị hạn chế chuyển động được đề cập trong dự đoán mẫu bất kỳ trong tập hợp ô bị hạn chế chuyển động;

tạo ra các lệnh để xây dựng đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ trên cơ sở tập hợp ô bị hạn chế chuyển động, trong đó đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ nêu trên có thể phát được dưới dạng đường ghi hình ảnh đầy đủ bất kỳ nhờ sử dụng quy trình phân tích cú pháp và giải mã các đường ghi hình ảnh đầy đủ, các lệnh nêu trên bao gồm cấu trúc bộ trích xuất để trích xuất một phần của bộ phận cơ bản từ đường ghi hình ảnh đầy đủ và trong đó bộ phận cơ bản là bộ lớp trừu tượng hóa mạng;

bao gồm, trong các lệnh nêu trên, cấu trúc bộ trích xuất bao gồm cấu trúc hàm dựng nội tuyến mà bao gồm một hoặc cả hai phần đầu lớp trừu tượng mạng và phần đầu phân đoạn mảng, và cấu trúc hàm dựng mà trích xuất tập hợp ô bị hạn chế chuyển động từ đường ghi hình ảnh đầy đủ.

2. Thiết bị theo điểm 1, trong đó tập hợp ô bị hạn chế chuyển động được chứa trong một hoặc nhiều phân đoạn mảng bao gồm các ô bên ngoài tập hợp ô bị hạn chế chuyển động, trong đó để xây dựng tập hợp ô bị hạn chế chuyển động nêu trên, thiết bị này còn khiến cho:

việc bổ sung cột ô ở bên phải của tập hợp ô bị hạn chế chuyển động vào một trong số các đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ, dòng bit tuân theo

hình ảnh đầy đủ, các lệnh để xây dựng đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ, hoặc dòng bit.

3. Thiết bị theo điểm 2, trong đó thiết bị này còn khiến cho: việc điều khiển sự kết thúc của phân đoạn mảng với các ô của cột ô được bổ sung.

4. Thiết bị theo điểm 2, trong đó thiết bị này còn khiến cho: việc chỉ định trong ô cuối cùng của cột ô được bổ sung sự kết thúc của phân đoạn mảng.

5. Thiết bị theo điểm 2, trong đó thiết bị này còn khiến cho:

việc bổ sung, vào đường ghi chứa tập hợp ô hoặc dòng bit tuân theo hình ảnh đầy đủ hoặc vào các lệnh để xây dựng đường ghi chứa tập hợp ô hoặc dòng bit tuân theo hình ảnh đầy đủ, thông tin về cửa sổ cắt ảnh phù hợp được tạo cấu hình để khiến bước giải mã loại trừ cột ô được bổ sung từ bước giải mã đầu ra.

6. Thiết bị theo điểm 1, trong đó tập hợp ô bị hạn chế chuyển động có trong một hoặc nhiều phân đoạn mảng bao gồm các ô bên ngoài tập hợp ô bị hạn chế chuyển động, trong đó để xây dựng tập hợp ô bị hạn chế chuyển động, thiết bị này còn khiến cho:

việc nhận diện ô ngoài cùng bên phải trong tập hợp ô bị hạn chế chuyển động mà không phải ô cuối cùng của phân đoạn mảng thứ nhất chứa ô ngoài cùng bên phải; và

việc sửa đổi ô ngoài cùng bên phải sao cho nó tạo ra sự kết thúc của mảng thứ nhất.

7. Thiết bị theo điểm 1, trong đó để xây dựng đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ bao gồm việc ghi lại ít nhất một bộ thông số hoặc toàn bộ hoặc một phần phần đầu phân đoạn mảng khi được so sánh với chúng trong đường ghi hình ảnh đầy đủ.

8. Thiết bị theo điểm 1, trong đó các lệnh để xây dựng đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ dẫn đến kết quả một chuỗi các phân đoạn hoặc các phân đoạn con và trong đó bộ định vị tài nguyên đồng nhất của phân đoạn được chỉ ra với tập hợp con của các lệnh dẫn đến phân đoạn đó.

9. Phương pháp mã hóa và giải mã video, phương pháp này bao gồm các bước:

thu được dòng bit của hình ảnh đầy đủ bao gồm:

ảnh được mã hóa, mỗi ảnh được mã hóa này bao gồm lưới của bộ cây mã hóa và được phân chia thành một hoặc nhiều ô chứa số nguyên của bộ cây mã hóa tạo thành hình chữ nhật, và

tập hợp ô bị hạn chế chuyển động bao gồm một hoặc nhiều ô trên mỗi ảnh được mã hóa theo cách mà không có giá trị mẫu bên ngoài tập hợp ô bị hạn chế chuyển động được đề cập trong dự đoán mẫu bất kỳ trong tập hợp ô bị hạn chế chuyển động;

tạo ra các lệnh để xây dựng đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ trên cơ sở tập hợp ô bị hạn chế chuyển động, trong đó đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ nêu trên có thể phát được dưới dạng đường ghi hình ảnh đầy đủ bất kỳ nhờ sử dụng quy trình phân tích cú pháp và giải mã các đường ghi hình ảnh đầy đủ, các lệnh nêu trên bao gồm cấu trúc bộ trích xuất để trích xuất một phần của bộ phận cơ bản từ đường ghi hình ảnh đầy đủ và trong đó bộ phận cơ bản là bộ lớp trừu tượng hóa mạng; và

bao gồm, trong các lệnh nêu trên, cấu trúc bộ trích xuất bao gồm cấu trúc hàm dựng nội tuyến mà bao gồm một hoặc cả hai phần đầu lớp trừu tượng mạng và phần đầu phân đoạn mảng, và cấu trúc hàm dựng mà trích xuất tập hợp ô bị hạn chế chuyển động từ đường ghi hình ảnh đầy đủ.

10. Phương pháp theo điểm 9, trong đó tập hợp ô bị hạn chế chuyển động khác biệt ở chỗ được chứa trong một hoặc nhiều phân đoạn mảng vốn cũng chứa các ô bên ngoài tập hợp ô bị hạn chế chuyển động, trong đó phương pháp này còn bao gồm bước xây dựng hoặc tạo ra các lệnh để xây dựng với:

bước bổ sung, vào đường ghi chứa tập hợp ô hoặc dòng bit tuân theo hình ảnh đầy đủ hoặc vào các lệnh để xây dựng đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ, cột ô ở bên phải của tập hợp ô bị hạn chế chuyển động.

11. Phương pháp theo điểm 10, trong đó phương pháp này còn bao gồm bước:

điều khiển sự kết thúc của phân đoạn mảng với các ô của cột ô được bổ sung.

12. Phương pháp theo điểm 10, trong đó phương pháp này còn bao gồm bước:

chỉ định trong ô cuối cùng của cột ô được bổ sung sự kết thúc của phân đoạn mảng.

13. Phương pháp theo điểm 10, trong đó phương pháp này còn bao gồm bước:

bổ sung, vào đường ghi chứa tập hợp ô hoặc dòng bit tuân theo hình ảnh đầy đủ hoặc vào các lệnh để xây dựng đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ, thông tin về cửa sổ cắt ảnh phù hợp được tạo cấu hình để khiến bước giải mã loại trừ cột ô được bổ sung từ việc giải mã đầu ra.

14. Vật ghi lâu dài đọc được bằng máy tính lưu trữ sản phẩm chương trình máy tính, bao gồm mã chương trình máy tính có thể thực thi được bằng máy tính mà khi được thực thi bằng thiết bị được tạo cấu hình để khiến cho thiết bị thực hiện phương pháp theo điểm 9.

15. Thiết bị mã hóa và giải mã video, thiết bị này bao gồm:

ít nhất một bộ xử lý; và

ít nhất một bộ nhớ lâu dài bao gồm mã chương trình máy tính, mà khi được thực thi bởi ít nhất một bộ xử lý, khiến cho thiết bị:

thu được các đặc tính thứ nhất của đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ thứ nhất từ ít nhất một phần mô tả trình bày phương tiện hoặc từ tệp đa phương tiện, trong đó đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ nêu trên có thể phát được dưới dạng đường ghi hình ảnh đầy đủ bằng quy trình phân tích cú pháp và giải mã của các đường ghi hình ảnh đầy đủ;

xác định dựa trên các đặc tính thứ nhất để thu được một phần hoặc toàn bộ đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ thứ nhất;

thu được một phần hoặc toàn bộ đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ thứ nhất bằng cách trích xuất cấu trúc bộ trích xuất mà trích xuất một phần của bộ phận cơ bản từ đường ghi hình ảnh đầy đủ và trong đó bộ phận cơ bản là bộ lớp trừu tượng hóa mạng; và

thu được một phần hoặc toàn bộ đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ thứ nhất bằng cách trích xuất cấu trúc bộ trích xuất bao gồm cấu trúc hàm dựng nội tuyến mà bao gồm một hoặc cả hai phần đầu lớp trừu tượng mạng và phần đầu phân đoạn mảng, và cấu trúc hàm dựng để trích xuất tập hợp ô bị hạn chế chuyển động từ đường ghi hình ảnh đầy đủ;

trong đó tập hợp ô bị hạn chế chuyển động có trong dòng bit hình ảnh đầy đủ bao gồm ảnh được mã hóa, mỗi ảnh được mã hóa bao gồm lưới của bộ cây mã hóa và được phân chia thành một hoặc nhiều ô chứa số nguyên của bộ cây mã hóa tạo thành hình chữ nhật;

trong đó tập hợp ô bị hạn chế chuyển động bao gồm một hoặc nhiều ô trên mỗi ảnh được mã hóa theo cách mà không có giá trị mẫu bên ngoài tập hợp ô bị hạn chế chuyển động được đề cập trong dự đoán mẫu bất kỳ trong tập hợp ô bị hạn chế chuyển động.

16. Thiết bị theo điểm 15, trong đó thiết bị này còn khiến cho:

việc giải mã và phát một phần hoặc toàn bộ đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ thứ nhất.

17. Thiết bị theo điểm 15, trong đó thiết bị này còn khiến cho:

việc thu được, trong các đặc tính thứ nhất, thông tin của vùng không gian được bao phủ hoặc cả phạm vi quan sát được bao phủ và hướng nhìn, trong đó việc xác định nêu trên dựa trên vùng không gian được bao phủ hoặc cả phạm vi quan sát được bao phủ và hướng nhìn.

18. Thiết bị theo điểm 15, trong đó thiết bị này còn khiến cho:

việc thu được các đặc tính thứ hai của đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ thứ hai từ ít nhất một phần mô tả trình bày phương tiện hoặc từ tệp đa phương tiện;

kết luận từ các đặc tính thứ nhất và thứ hai mà đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ thứ hai có ít nhất một trong số độ phân giải không gian, tốc độ bit, chất lượng hình ảnh, độ sâu bit cho mỗi thành phần màu, dải tần nhạy sáng hoặc độ rộng của dải màu thấp hơn so với đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ thứ nhất;

kết luận từ các đặc tính thứ nhất và thứ hai mà đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ thứ nhất và thứ hai đại diện cho các phần không gian khác nhau của nội dung;

xác định dựa trên các kết luận nêu trên để thu được một phần hoặc toàn bộ đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ thứ hai; và

thu được một phần hoặc toàn bộ đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ thứ hai.

19. Phương pháp mã hóa và giải mã video, phương pháp này bao gồm các bước:

thu được đặc tính thứ nhất của đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ thứ nhất từ phần mô tả trình bày phương tiện và/hoặc từ tệp đa phương tiện, trong đó đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ nêu trên có thể phát được dưới dạng đường ghi hình ảnh đầy đủ bất kỳ nhờ sử dụng quy trình phân tích cú pháp và giải mã các đường ghi hình ảnh đầy đủ;

xác định dựa trên các đặc tính thứ nhất để thu được một phần hoặc toàn bộ đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ thứ nhất;

thu được một phần hoặc toàn bộ đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ thứ nhất bằng cách trích xuất cấu trúc bộ trích xuất mà trích xuất một phần của bộ phận cơ bản từ đường ghi hình ảnh đầy đủ và trong đó bộ phận cơ bản là bộ lớp trừu tượng hóa mạng; và

thu được một phần hoặc toàn bộ đường ghi chứa tập hợp ô tuân theo hình ảnh đầy đủ thứ nhất bằng cách trích xuất cấu trúc bộ trích xuất bao gồm cấu trúc hàm dựng nội tuyến mà bao gồm một hoặc cả hai phần đầu lớp trừu tượng mạng và phần đầu phân đoạn mảng, và cấu trúc hàm dựng để trích xuất tập hợp ô bị hạn chế chuyển động từ đường ghi hình ảnh đầy đủ;

trong đó tập hợp ô bị hạn chế chuyển động được bao gồm trong dòng bit hình ảnh đầy đủ bao gồm ảnh được mã hóa, mỗi ảnh được mã hóa này bao gồm lưới của bộ cây mã hóa và được phân chia thành một hoặc nhiều ô chứa số nguyên của bộ cây mã hóa tạo thành hình chữ nhật;

trong đó tập hợp ô bị hạn chế chuyển động bao gồm một hoặc nhiều ô trên mỗi ảnh được mã hóa theo cách mà không có giá trị mẫu bên ngoài tập hợp ô bị hạn chế chuyển động được đề cập trong dự đoán mẫu bất kỳ trong tập hợp ô bị hạn chế chuyển động.

20. Vật ghi lâu dài đọc được bằng máy tính lưu trữ sản phẩm chương trình máy tính, bao gồm mã chương trình máy tính có thể thực thi được bằng máy tính mà khi được

thực thi bằng thiết bị được tạo cấu hình để khiến cho thiết bị thực hiện phương pháp theo điểm 19.

1/15

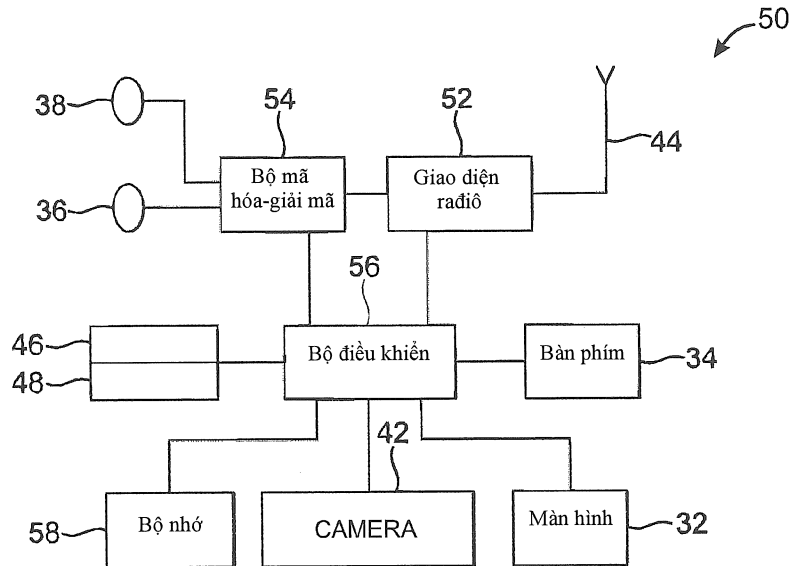


Fig. 1

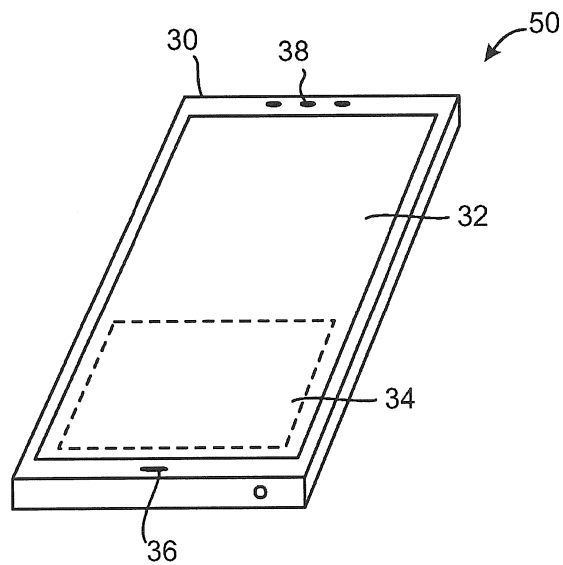


Fig. 2

2/15

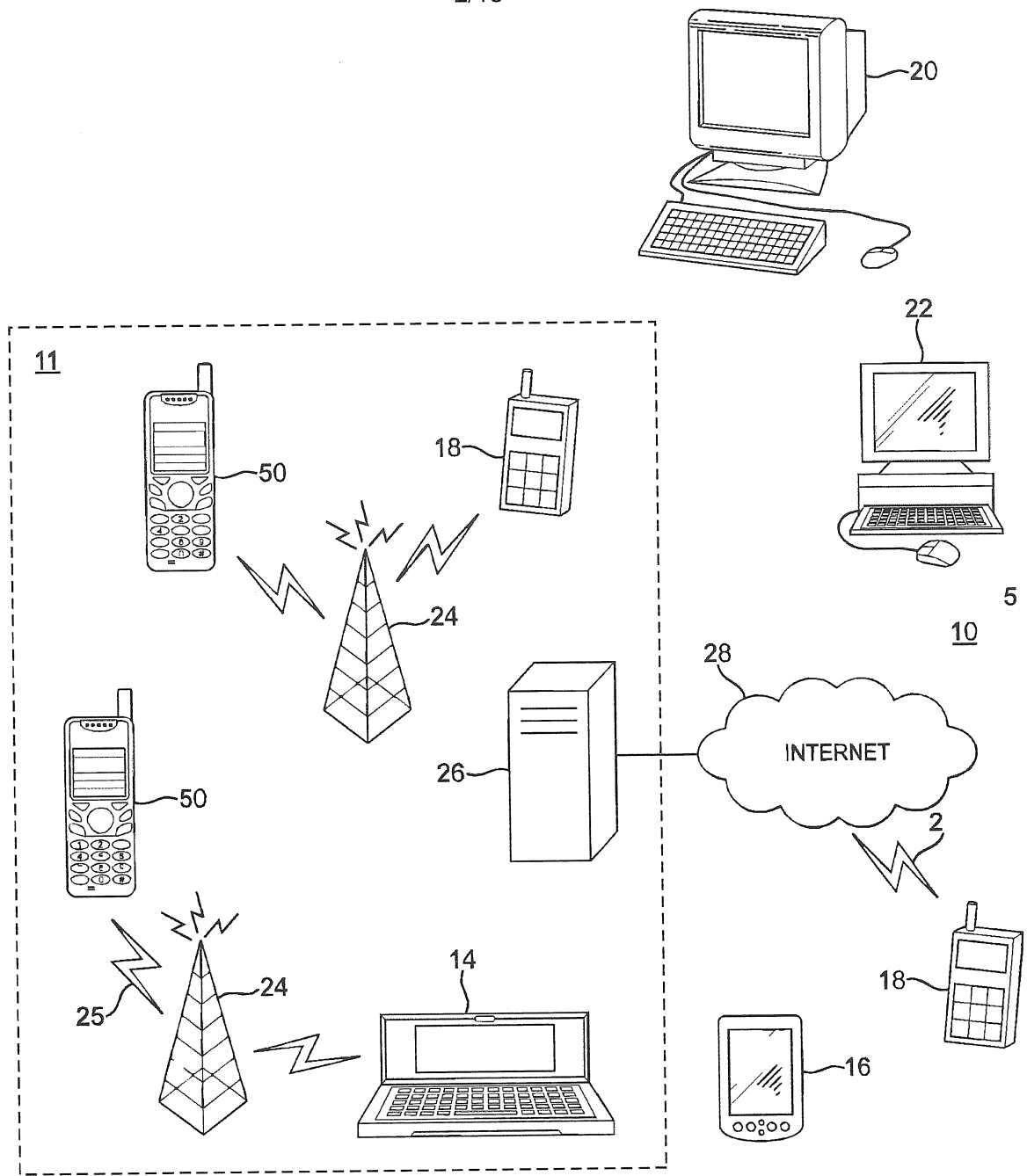


Fig. 3



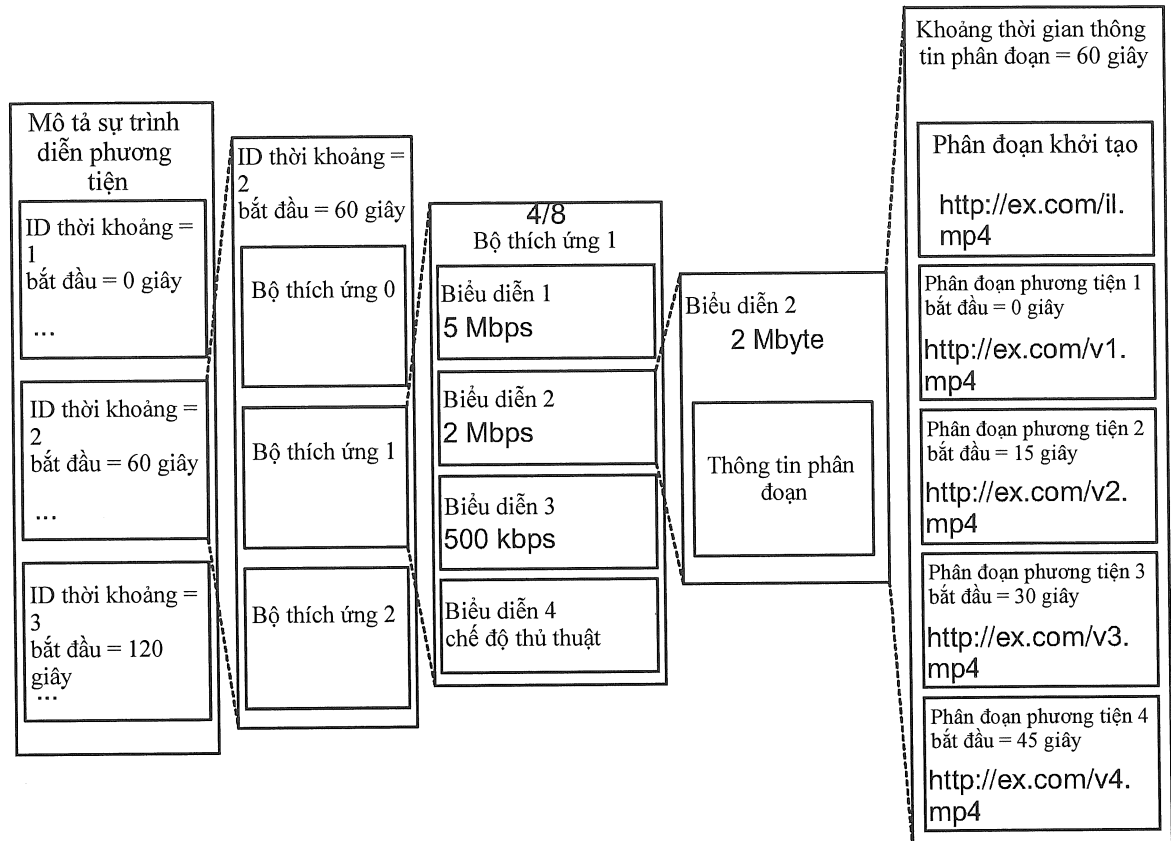


Fig. 5

5/15

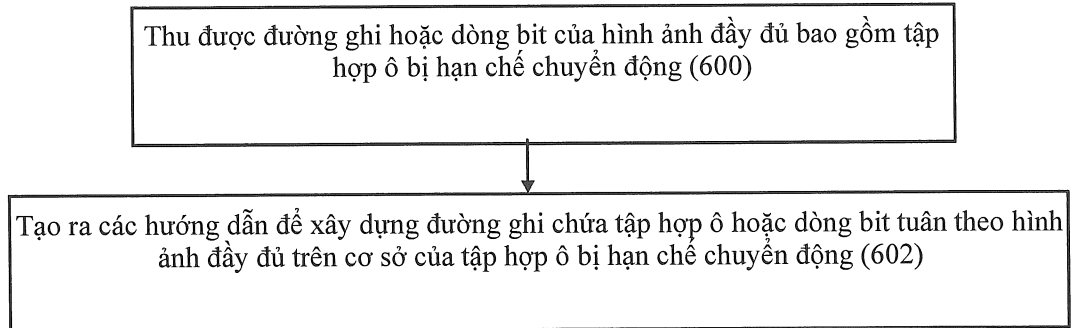


Fig. 6a

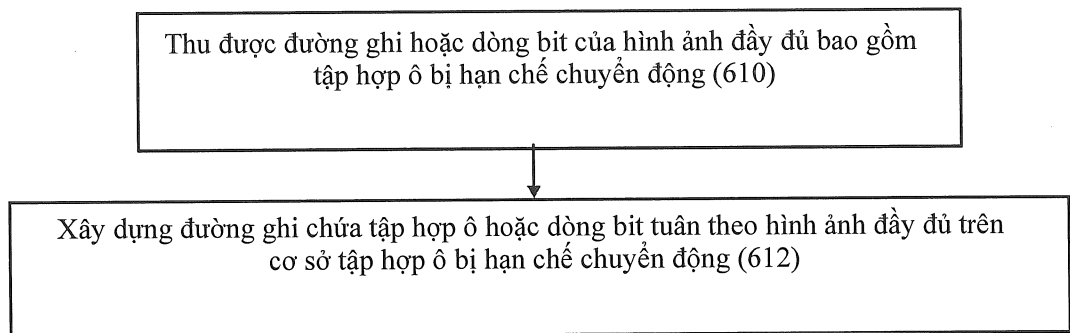


Fig. 6b

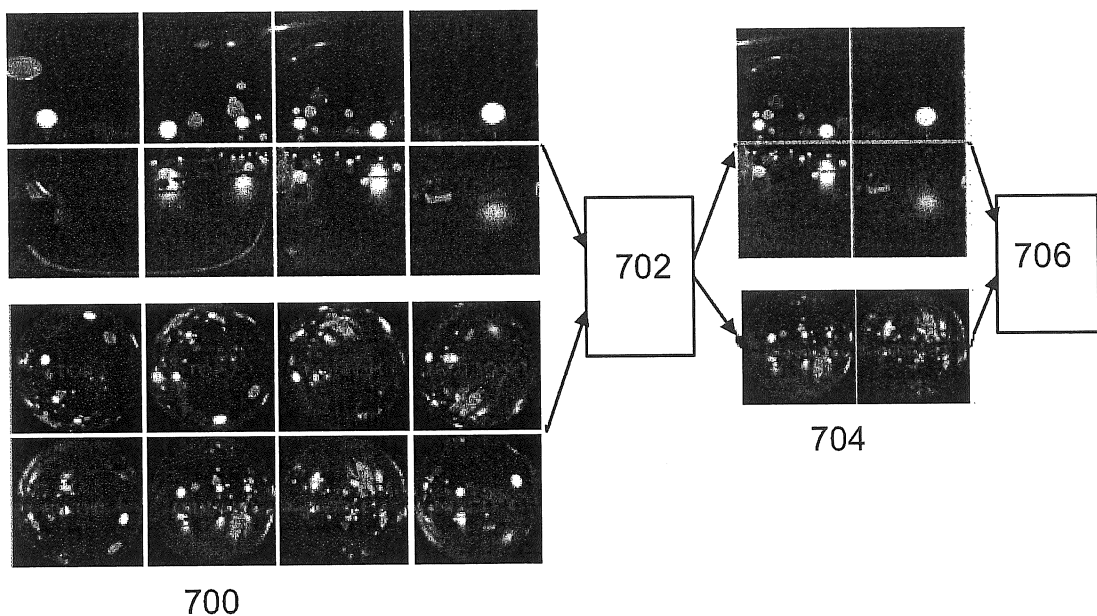


Fig. 7

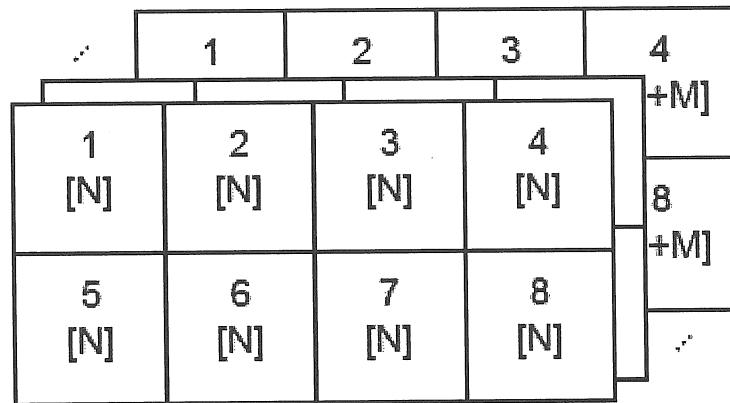


Fig. 8

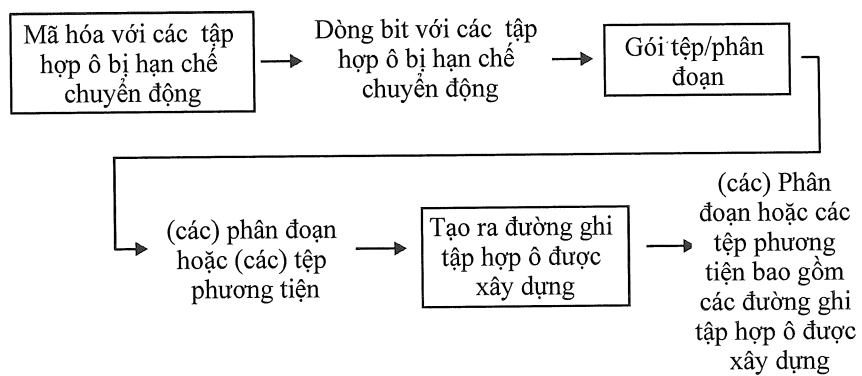


Fig. 9

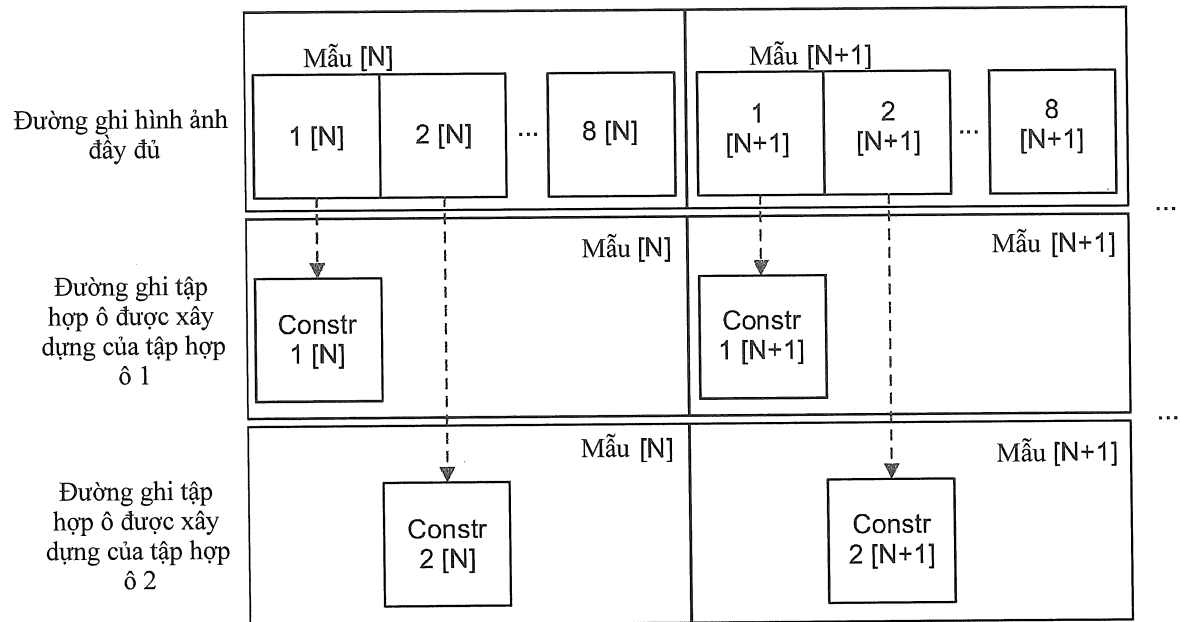


Fig. 10

8/15

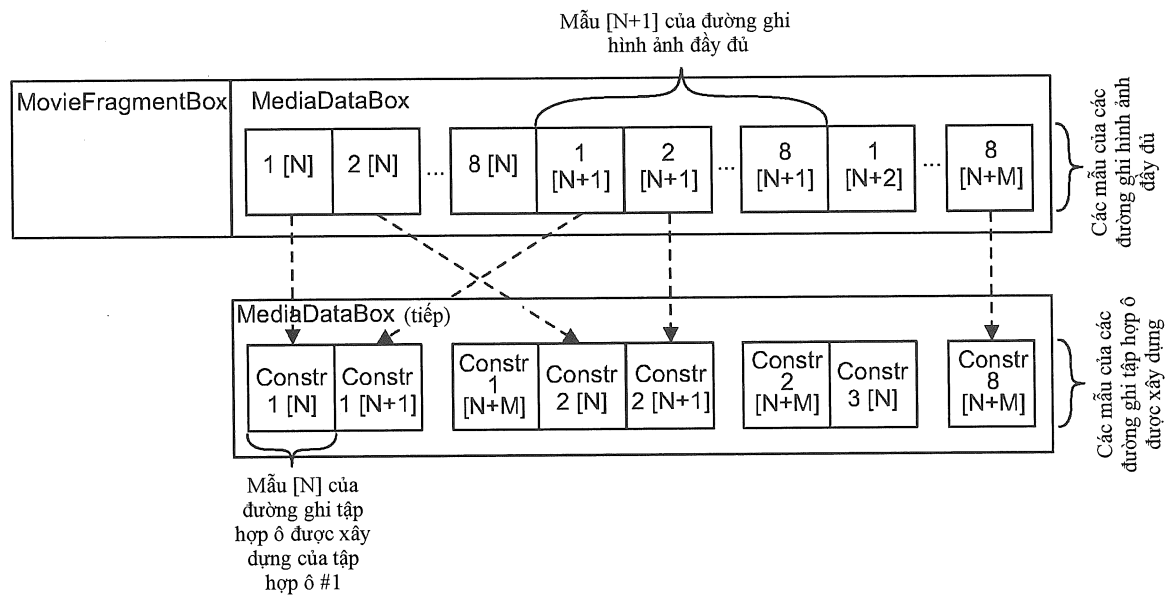


Fig. 11a

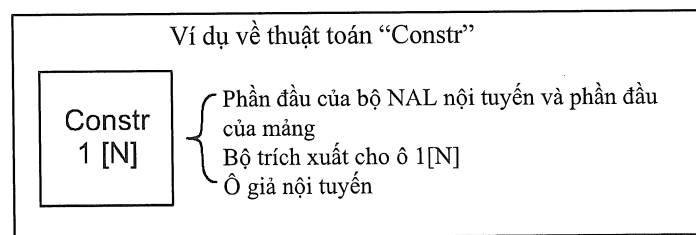


Fig. 11b

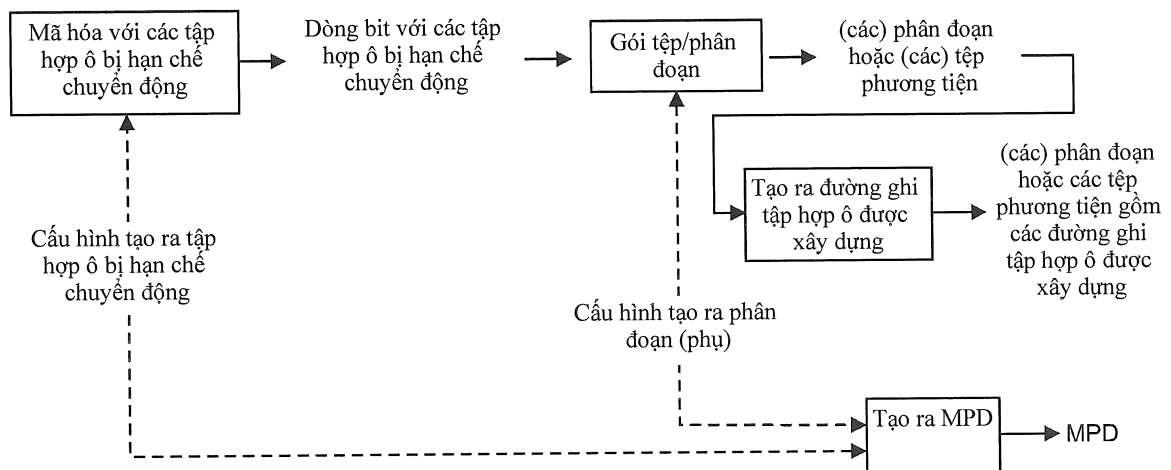
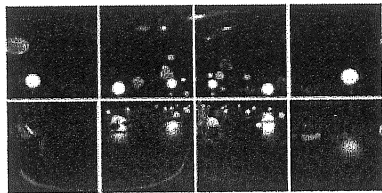


Fig. 12

9/15

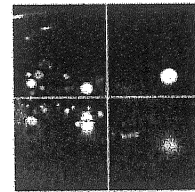
Dòng bit H.265/HEVC với các ô bị hạn chế chuyển động (độ phân giải cao)



Lưu ý: Hoạt động như nhau đối với loại nội dung bất kỳ, ví dụ nội dung nhiều khung hình được gói khung

Xây dựng đường ghi tập hợp ô/dòng bit trong máy chủ trước khi phát trực tuyến hoặc ở máy phát, trước khi giải mã

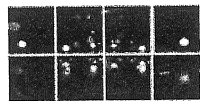
Đầu ra: (các) dòng bit H.265/HEVC phù hợp
- Có thể được chọn trên cơ sở hướng nhìn hiện tại
- Ví dụ:



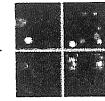
Ví dụ, 2x2 ô ngoài cùng bên phải

Giải mã H.265/HEVC

Dòng bit H.265/HEVC với các ô bị hạn chế chuyển động (độ phân giải thấp)



Xây dựng đường ghi tập hợp ô/dòng bit trong máy chủ trước khi phát trực tuyến hoặc ở máy phát, trước khi giải mã



Ví dụ, 2x2 ô ngoài cùng bên phải

Giải mã H.265/HEVC

Fig. 13

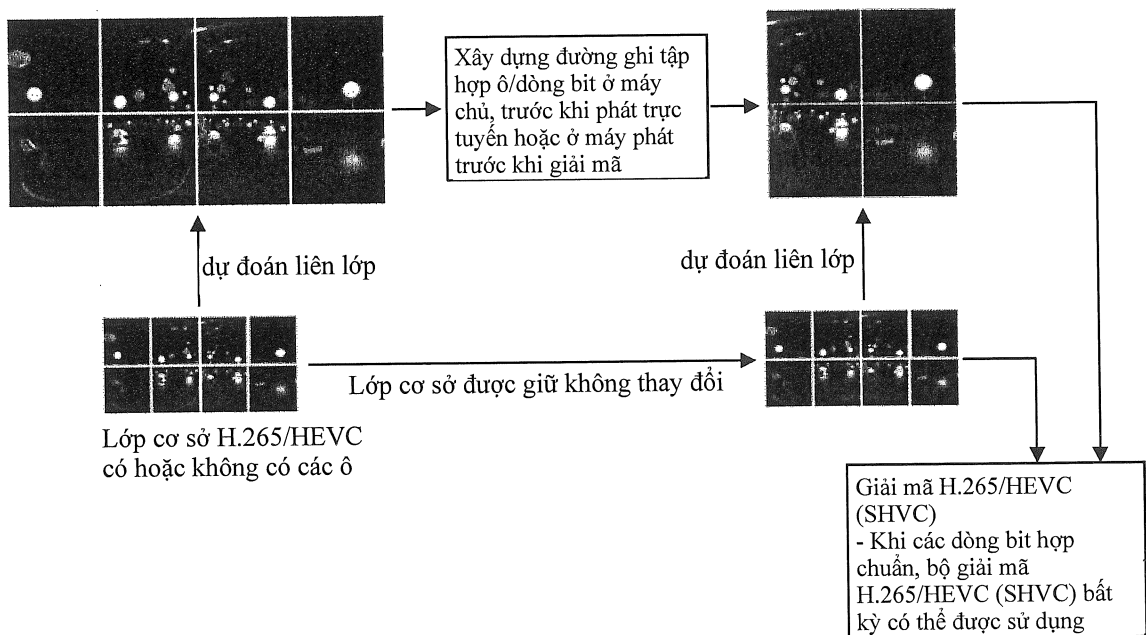
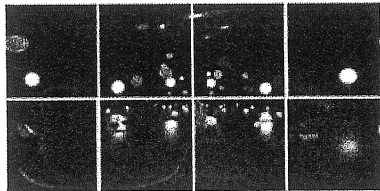


Fig. 14

10/15

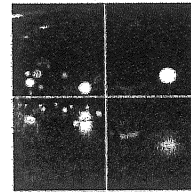
Dòng bit H.265/HEVC với các ô bị hạn chế chuyển động



Ví dụ, khung hình bên phải

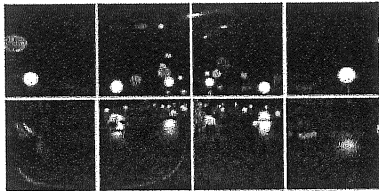
Xây dựng đường ghi tập hợp ô/dòng bit trong máy chủ trước khi phát trực tuyến hoặc ở máy phát, trước khi giải mã

Đầu ra: (các) dòng bit H.265/HEVC phù hợp
- Ví dụ:



Giải mã H.265/HEVC

Dòng bit H.265/HEVC có hoặc không có các ô bị hạn chế chuyển động



Ví dụ, khung hình bên trái

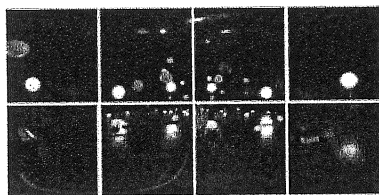
Lưu ý: Các độ phân giải về không gian của các dòng bit không cần giống nhau

Một khung hình được giữ không thay đổi

Giải mã H.265/HEVC

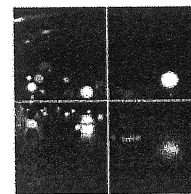
Fig. 15

Khung hình được dự báo H.265/HEVC có các ô bị hạn chế chuyển động và liên lớp



Ví dụ, khung hình bên phải

Xây dựng đường ghi tập hợp ô/dòng bit ở máy chủ, trước khi phát trực tuyến hoặc ở máy phát, trước khi giải mã



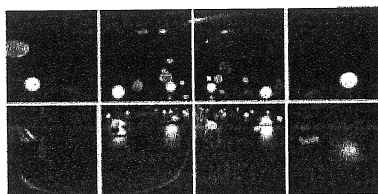
Giải mã khung hình được dự báo H.265/HEVC (MV-HEVC)

Dự đoán liên khung hình

Đầu ra: Làm phù hợp dòng bit H.265/HEVC với khung hình cơ bản và khung hình được dự đoán

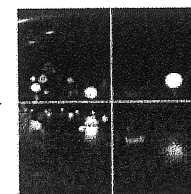
Dự đoán liên khung hình

Dòng bit H.265/HEVC có các ô bị hạn chế chuyển động



Ví dụ, khung hình bên trái

Xây dựng đường ghi tập hợp ô/dòng bit ở máy chủ, trước khi phát trực tuyến hoặc ở máy phát, trước khi giải mã



Giải mã H.265/HEVC

Fig. 16

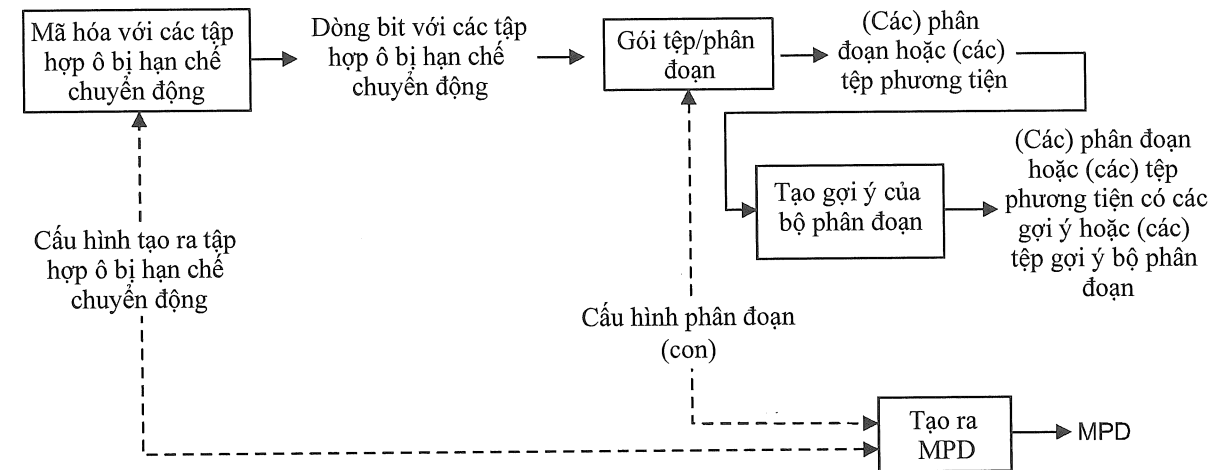


Fig. 17

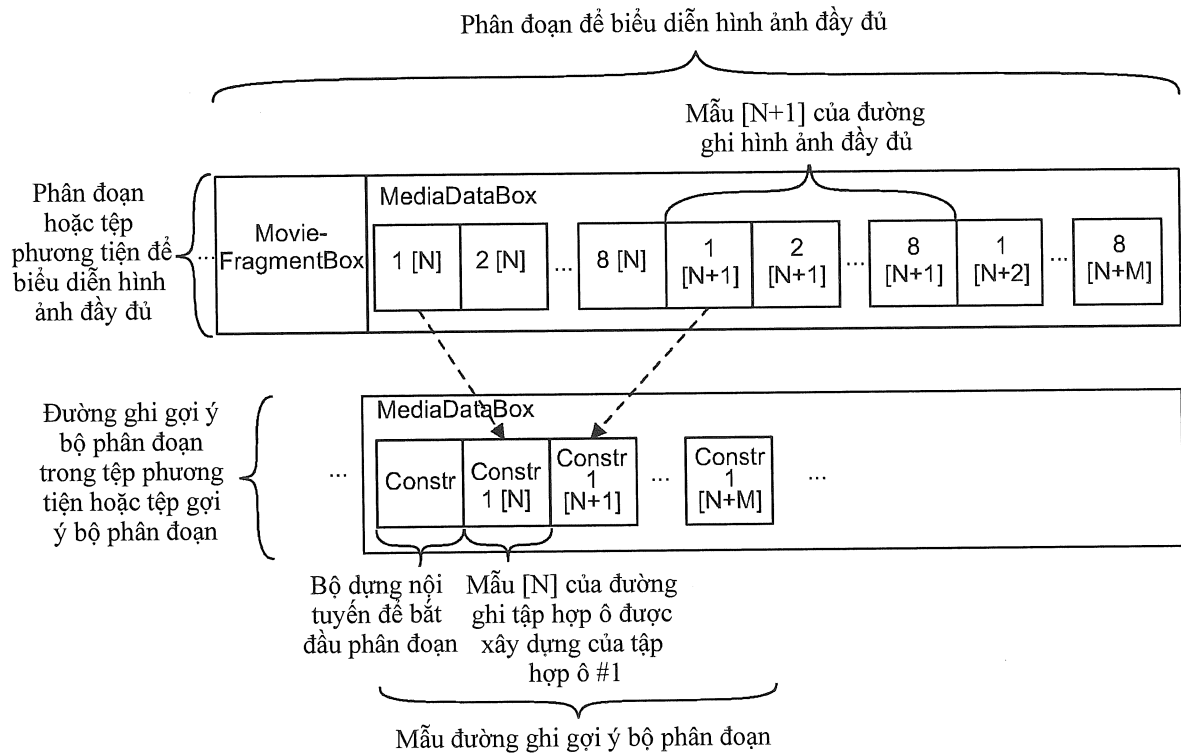


Fig. 18

12/15

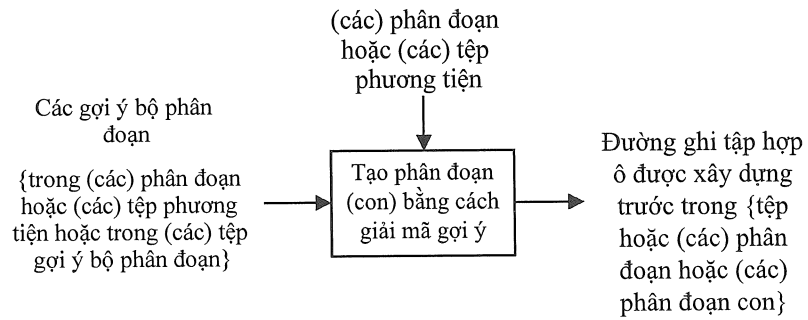


Fig. 19

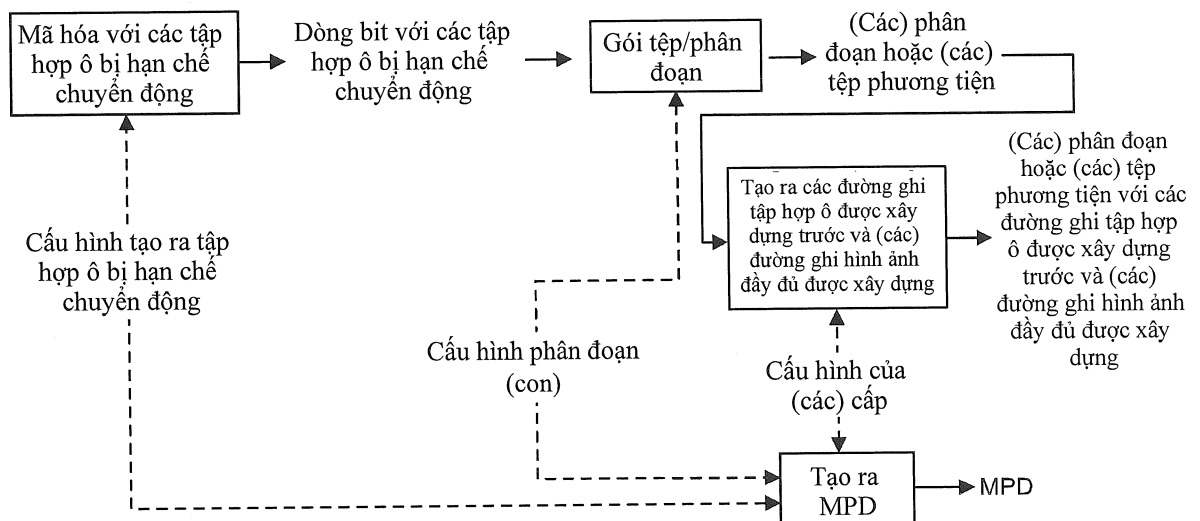


Fig. 20

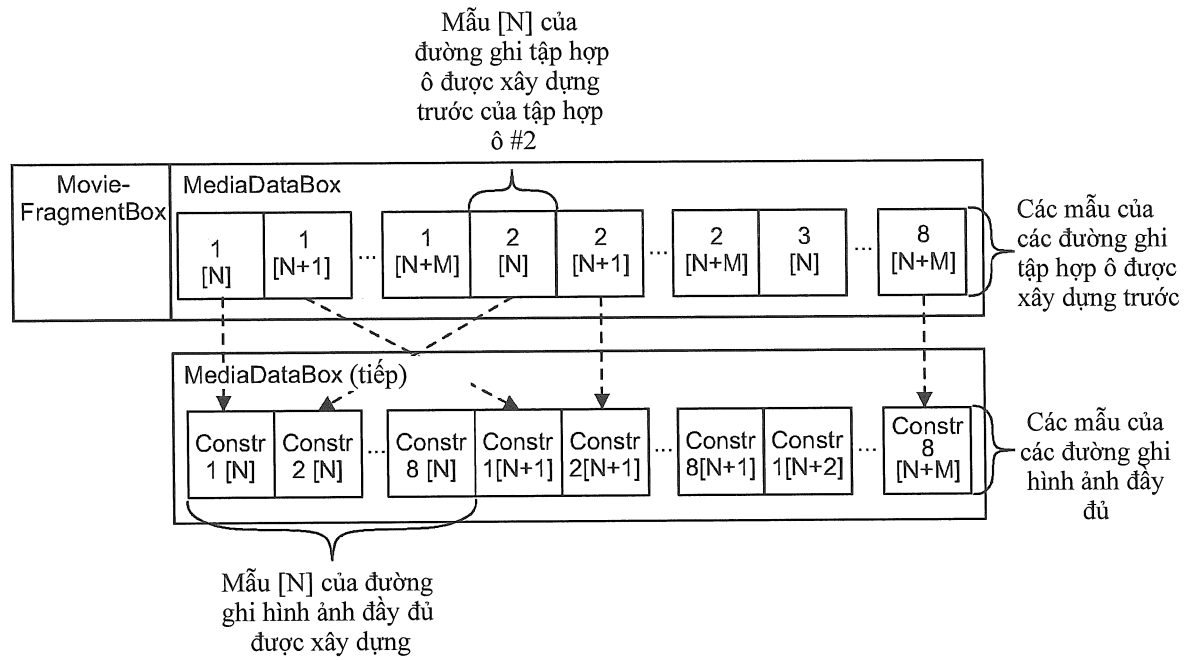


Fig. 21

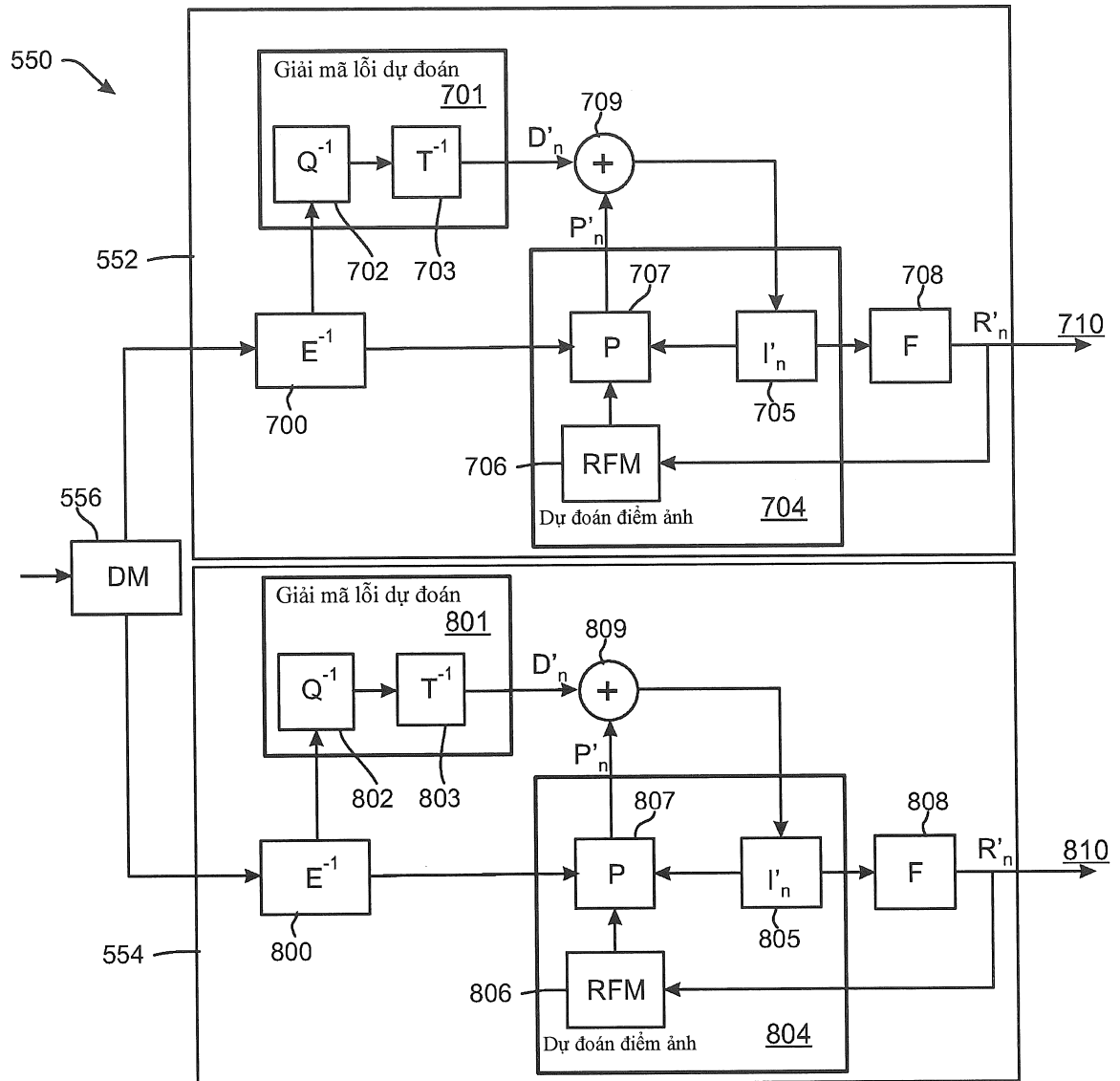


Fig. 22

15/15

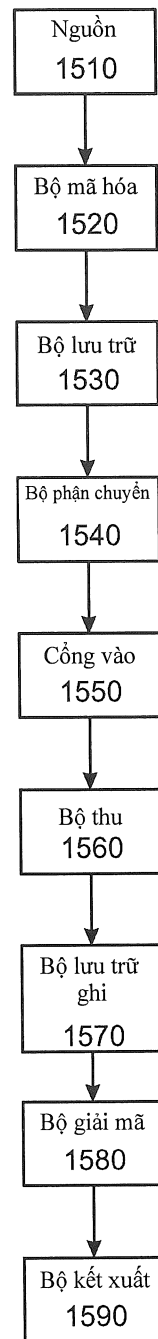


Fig. 23