



(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)



1-0034918

(51)⁷ H04N 19/55; H04N 19/70; H04N
19/563; H04N 13/00

(13) B

(21) 1-2018-04652

(22) 17/03/2017

(86) PCT/FI2017/050180 17/03/2017

(87) WO 2017/162912 28/09/2017

(30) 20165256 24/03/2016 FI

(45) 27/03/2023 420

(43) 25/01/2019 370A

(73) NOKIA TECHNOLOGIES OY (FI)

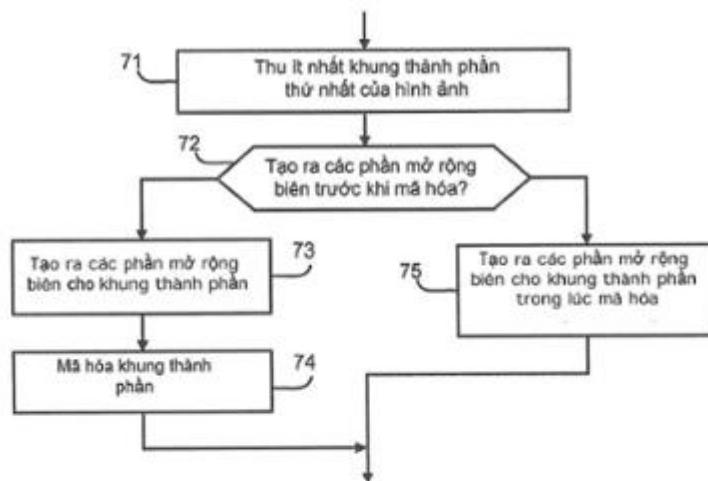
Karaportti 3, 02610 Espoo, Finland

(72) HANNUKSELA, Miska (FI).

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

(54) PHƯƠNG PHÁP VÀ THIẾT BỊ MÃ HÓA VÀ GIẢI MÃ VIDEO

(57) Sáng chế đề xuất phương pháp và thiết bị mã hóa và giải mã video. Phương pháp này bao gồm bước mã hóa khung thành phần không nén thành hình ảnh thứ nhất được mã hóa, việc mã hóa nêu trên cũng dẫn đến hình ảnh khôi phục thứ nhất, và khung thành phần nêu trên có vùng hình ảnh hiệu dụng trong hình ảnh khôi phục thứ nhất, thực hiện một trong các bước sau như là một phần của bước mã hóa nêu trên: chèn ít nhất một giá trị mẫu bên ngoài vùng hình ảnh hiệu dụng để tạo ra phần mở rộng biên cho khung thành phần trong hình ảnh khôi phục thứ nhất; hoặc làm bão hòa hoặc gói các vị trí mẫu bên ngoài vùng hình ảnh hiệu dụng ở trong vùng hình ảnh hiệu dụng. Sáng chế đề xuất phương pháp bao gồm bước nhận hình ảnh được mã hóa, giải mã hình ảnh được mã hóa để tạo ra khung thành phần được khôi phục của hình ảnh có vùng hình ảnh hiệu dụng; và thực hiện một trong các bước sau: làm đầy vùng bên ngoài vùng hình ảnh hiệu dụng để tạo ra hình ảnh tham chiếu được đệm, trong đó vùng được làm đầy tạo ra phần mở rộng biên; hoặc xác định rằng khi tham chiếu đến các vị trí mẫu bên ngoài vùng hình ảnh hiệu dụng trong bước giải mã, các vị trí mẫu nêu trên được làm bão hòa hoặc được gói ở trong vùng hình ảnh hiệu dụng.



Lĩnh vực kỹ thuật được đề cập

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

Tình trạng kỹ thuật của sáng chế

Phần này nhằm mục đích cung cấp tình trạng kỹ thuật hoặc ngữ cảnh của sáng chế được trích dẫn trong phần yêu cầu bảo hộ. Phần mô tả dưới đây có thể gồm các khái niệm cần được tuân theo, nhưng không nhất thiết là các khái niệm đã được biết hoặc tuân theo từ trước. Vì vậy, trừ khi được nêu rõ trong bản đặc tả này, nội dung được mô tả trong phần này không phải là tình trạng kỹ thuật cho phần mô tả và yêu cầu bảo hộ trong đơn sáng chế này và không được thừa nhận như là tình trạng kỹ thuật do có mặt trong phần này.

Hệ thống mã hóa video có thể bao gồm bộ mã hóa trong đó biến đổi video đầu vào thành biểu diễn nén thích hợp để lưu trữ/truyền và bộ giải mã có thể giải nén dạng biểu diễn video nén trở lại thành dạng xem được. Bộ mã hóa có thể loại bỏ một số thông tin trong chuỗi video ban đầu để biểu diễn video ở dạng nén hơn, ví dụ, để cho phép việc lưu trữ/truyền thông tin video với tỷ lệ bit thấp hơn cần thiết.

Bản chất kỹ thuật của sáng chế

Một số phương án đề xuất phương pháp mã hóa và giải mã thông tin video. Theo một số phương án, sáng chế đề xuất phương pháp, thiết bị và sản phẩm chương trình máy tính để mã hóa video.

Phương pháp theo khía cạnh thứ nhất bao gồm các bước:

nhận hình ảnh được mã hóa;

giải mã hình ảnh được mã hóa để tạo ra khung thành phần được khôi phục của hình ảnh có vùng hình ảnh hiệu dụng; và

thực hiện một trong các bước sau:

làm đầy vùng bên ngoài vùng hình ảnh hiệu dụng để tạo thành hình ảnh tham chiếu được đệm, trong đó vùng được làm đầy này tạo thành phần mở rộng biên; hoặc

xác định rằng khi tham chiếu đến các vị trí mẫu bên ngoài vùng hình ảnh hiệu dụng trong bước giải mã, các vị trí mẫu nêu trên được làm bão hòa hoặc được gói ở trong vùng hình ảnh hiệu dụng.

Thiết bị theo khía cạnh thứ hai bao gồm ít nhất một bộ xử lý và ít nhất một bộ nhớ, ít nhất một bộ nhớ nêu trên được lưu trữ với mã trên đó, khi được chạy bởi ít nhất một bộ xử lý nêu trên, khiến thiết bị thực hiện ít nhất các bước:

nhận hình ảnh được mã hóa;

giải mã hình ảnh được mã hóa để tạo ra khung thành phần được khôi phục của hình ảnh có vùng hình ảnh hiệu dụng; và

thực hiện một trong các bước sau:

làm đầy vùng bên ngoài vùng hình ảnh hiệu dụng để tạo ra hình ảnh tham chiếu được đệm, trong đó vùng được làm đầy tạo thành phần mở rộng biên; hoặc

xác định rằng khi tham chiếu đến các vị trí mẫu bên ngoài vùng hình ảnh hiệu dụng trong bước giải mã, các vị trí mẫu nêu trên được làm bão hòa hoặc được gói trong vùng hình ảnh hiệu dụng.

Vật ghi đọc được bằng máy tính theo khía cạnh thứ ba bao gồm mã để sử dụng bởi thiết bị, khi được chạy bởi bộ xử lý, khiến thiết bị thực hiện:

nhận hình ảnh được mã hóa;

giải mã hình ảnh được mã hóa để tạo ra khung thành phần được khôi phục của hình ảnh có vùng hình ảnh hiệu dụng; và

thực hiện một trong các bước sau:

làm đầy vùng bên ngoài vùng hình ảnh hiệu dụng để tạo ra hình ảnh tham chiếu được đệm, trong đó vùng được làm đầy tạo ra phần mở rộng biên; hoặc

xác định rằng khi tham chiếu đến các vị trí mẫu bên ngoài vùng hình ảnh hiệu dụng trong bước giải mã, các vị trí mẫu nêu trên được làm bão hòa hoặc được gói trong vùng hình ảnh hiệu dụng.

Thiết bị theo khía cạnh thứ tư bao gồm:

các phương tiện để nhận hình ảnh được mã hóa;

các phương tiện để giải mã hình ảnh được mã hóa để tạo ra khung thành phần được khôi phục của hình ảnh có vùng hình ảnh hiệu dụng; và

các phương tiện để thực hiện một trong các bước sau:

làm đầy vùng bên ngoài vùng hình ảnh hiệu dụng để tạo ra hình ảnh tham chiếu được đệm, trong đó vùng được làm đầy tạo thành phần mở rộng biên; hoặc

xác định rằng khi tham chiếu đến các vị trí mẫu bên ngoài vùng hình ảnh hiệu dụng trong bước giải mã, các vị trí mẫu nêu trên được làm bão hòa hoặc được gói trong vùng hình ảnh hiệu dụng.

Phương pháp theo khía cạnh thứ năm bao gồm bước:

thu khung thành phần của hình ảnh thứ nhất có vùng hình ảnh hiệu dụng để mã hóa;

thực hiện một trong các bước sau:

chèn ít nhất một giá trị mẫu bên ngoài vùng hình ảnh hiệu dụng của khung thành phần để tạo ra phần mở rộng biên đối với khung thành phần; hoặc

làm bão hòa hoặc gói các vị trí mẫu bên ngoài vùng hình ảnh hiệu dụng ở trong vùng hình ảnh hiệu dụng.

Thiết bị theo khía cạnh thứ sáu bao gồm ít nhất một bộ xử lý và ít nhất một bộ nhớ, ít nhất một bộ nhớ nêu trên được lưu trữ với mã trên đó, khi được chạy với ít nhất một bộ xử lý nêu trên, khiến thiết bị thực hiện ít nhất:

mã hóa khung thành phần không nén thành hình ảnh thứ nhất được mã hóa, việc mã hóa nêu trên cũng dẫn đến hình ảnh khôi phục thứ nhất và khung thành phần nêu trên có vùng hình ảnh hiệu dụng trong hình ảnh khôi phục thứ nhất;

thực hiện một trong các bước sau dưới dạng một phần của bước mã hóa nêu trên:

chèn ít nhất một giá trị mẫu bên ngoài vùng hình ảnh hiệu dụng để tạo ra phần mở rộng biên đối với khung thành phần trong hình ảnh khôi phục thứ nhất; hoặc

làm bão hòa hoặc gói các vị trí mẫu bên ngoài vùng hình ảnh hiệu dụng trong vùng hình ảnh hiệu dụng.

Vật ghi đọc được bằng máy tính theo khía cạnh thứ bảy bao gồm mã để sử dụng bởi thiết bị, khi được chạy bởi bộ xử lý, khiến thiết bị thực hiện:

mã hóa khung thành phần không nén thành hình ảnh thứ nhất được mã hóa, việc mã hóa nêu trên cũng dẫn đến hình ảnh khôi phục thứ nhất và khung thành phần nêu trên có vùng hình ảnh hiệu dụng trong hình ảnh khôi phục thứ nhất;

thực hiện một trong các bước sau dưới dạng một phần của bước mã hóa nêu trên:

chèn ít nhất một giá trị mẫu bên ngoài vùng hình ảnh hiệu dụng để tạo ra phần mở rộng biên đối với khung thành phần trong hình ảnh khôi phục thứ nhất; hoặc

làm bão hòa hoặc gói các vị trí mẫu bên ngoài vùng hình ảnh hiệu dụng ở trong vùng hình ảnh hiệu dụng.

Thiết bị theo khía cạnh thứ tám bao gồm:

các phương tiện mã hóa khung thành phần không nén thành hình ảnh thứ nhất được mã hóa, việc mã hóa nêu trên cũng mang đến hình ảnh khôi

phục thứ nhất và khung thành phần nêu trên có vùng hình ảnh hiệu dụng trong hình ảnh khôi phục thứ nhất;

các phương tiện để thực hiện một trong các bước sau dưới dạng một phần của bước mã hóa nêu trên:

chèn ít nhất một giá trị mẫu bên ngoài vùng hình ảnh hiệu dụng để tạo ra phần mở rộng biên đối với khung thành phần trong hình ảnh khôi phục thứ nhất; hoặc

làm bão hòa hoặc gói các vị trí mẫu bên ngoài vùng hình ảnh hiệu dụng ở trong vùng hình ảnh hiệu dụng.

Các khía cạnh khác bao gồm ít nhất các thiết bị và các sản phẩm/mã chương trình máy tính được lưu trữ trên vật ghi không khả biến được bố trí để thực hiện các phương pháp nêu trên.

Mô tả vắn tắt các hình vẽ

Để hiểu rõ hơn về sáng chế, giờ đây tham chiếu sẽ được tạo ra theo cách làm ví dụ với các hình vẽ kèm theo trong đó:

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

Fig.2 thể hiện sơ đồ của thiết bị người dùng thích hợp để thực hiện các phương án của sáng chế;

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

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

Fig.5a thể hiện các nguồn ứng viên theo không gian của bộ dự đoán vector chuyển động ứng viên, theo một phương án;

Fig.5b thể hiện các nguồn ứng viên theo thời gian của bộ dự đoán vector chuyển động ứng viên, theo một phương án;

Fig.6a minh họa vấn đề liên quan đến các vectơ chuyển động phụ được xử lý một cách tối ưu;

Fig.6b thể hiện ví dụ về phần mở rộng biên, theo một phương án;

Fig.6c thể hiện các nguồn ứng viên theo không gian của bộ dự đoán vectơ chuyển động ứng viên, theo một phương án;

Fig.6d thể hiện các nguồn ứng viên theo thời gian của bộ dự đoán vectơ chuyển động ứng viên, theo một phương án;

Fig.7a thể hiện lưu đồ mã hóa hình ảnh, theo một phương án;

Fig.7b thể hiện lưu đồ giải mã hình ảnh, theo một phương án;

Fig.8 thể hiện sơ đồ của bộ giải mã thích hợp để thực hiện các phương án của sáng chế; và

Fig.9 thể hiện sơ đồ của hệ thống truyền thông đa phương tiện làm ví dụ trong đó các phương án khác nhau có thể được triển khai; và

Các hình vẽ từ Fig.10a đến Fig.10c minh họa ánh xạ hình chóp, theo một phương án.

Mô tả chi tiết sáng chế

Sau đây, các phương án của sáng chế sẽ được mô tả trong ngữ cảnh của một bố trí mã hóa video. Tuy nhiên, lưu ý rằng sáng chế không bị giới hạn với bố trí cụ thể này. Thực tế, các phương án khác có ứng dụng rộng rãi trong bất kỳ môi trường nào trong đó việc tăng cường mã hóa video không mở rộng được, mở rộng được và/hoặc đa góc nhìn được yêu cầu. Ví dụ, sáng chế có thể được áp dụng với các hệ thống mã hóa video như các hệ thống tạo dòng, các bộ phát DVD, các bộ thu truyền hình số, máy quay video cá nhân, các hệ thống và các chương trình máy tính trên các máy tính cá nhân, các máy tính cầm tay và các thiết bị truyền thông, cũng như các thành phần mạng như các bộ chuyển mã và các bố trí điện toán đám mây trong đó dữ liệu video được xử lý.

Sau đây mô tả chi tiết hơn thiết bị tương thích và các cơ chế khả dĩ để triển khai một số phương án. Về vấn đề này trước hết tham chiếu được thực hiện với

Fig.1 và Fig.2, trong đó Fig.1 thể hiện sơ đồ khối của hệ thống mã hóa video theo một phương án làm ví dụ dưới dạng sơ đồ khối của thiết bị hoặc thiết bị điện tử 50 làm ví dụ, có thể hợp nhất với bộ mã hóa-giải mã theo một phương án của sáng chế. Fig.2 thể hiện sơ đồ của thiết bị theo một phương án làm ví dụ. Các thành phần của Fig.1 và Fig.2 sẽ được giải thích tiếp theo.

Ví dụ, thiết bị điện tử 50 có thể là thiết bị đầu cuối di động hoặc thiết bị người dùng của hệ thống truyền thông không dây. Tuy nhiên, cần hiểu rằng các phương án của sáng chế có thể được triển khai trong bất kỳ thiết bị điện tử hoặc thiết bị nào có thể yêu cầu mã hóa và giải mã hoặc mã hóa hoặc giải mã các ảnh video.

Thiết bị 50 có thể bao gồm vỏ 30 để chứa và bảo vệ thiết bị này. Thiết bị 50 có thể còn bao gồm màn hình 32 ở dạng của màn hình tinh thể lỏng. Theo các phương án khác của sáng chế màn hình này có thể là bất kỳ công nghệ hiển thị thích hợp nào thích hợp để hiển thị ảnh hoặc video. Thiết bị 50 cũng có thể bao gồm bàn phím 34. Theo các phương án khác của sáng chế bất kỳ cơ chế dữ liệu hoặc giao diện người dùng thích hợp nào có thể được sử dụng. Ví dụ giao diện người dùng này có thể được triển khai dưới dạng bàn phím ảo hoặc hệ thống đầu vào dữ liệu dưới dạng một phần của màn hình cảm ứng.

Thiết bị có thể bao gồm micrô 36 hoặc bất kỳ đầu vào audio thích hợp nào có thể là đầu vào tín hiệu số hoặc đầu vào tín hiệu tương tự. Thiết bị 50 cũng có thể bao gồm thiết bị đầu ra audio theo các phương án của sáng chế có thể là bất kỳ một trong số: lỗ cắm tai nghe 38, loa, hoặc kết nối đầu ra audio tương tự hoặc audio số. Thiết bị 50 cũng có thể bao gồm pin 40 (hoặc theo các phương án khác của sáng chế thiết bị này có thể được cấp nguồn bởi bất kỳ thiết bị năng lượng di động thích hợp nào như pin mặt trời, pin nhiên liệu hoặc máy phát điện đồng hồ). Thiết bị này cũng có thể bao gồm camera 42 có thể quay hoặc chụp các ảnh và/hoặc video. Thiết bị 50 cũng có thể bao gồm cổng hồng ngoại đối với đường nhìn truyền thông tầm ngắn đến các thiết bị khác. Theo các phương án khác thiết bị 50 cũng có thể bao gồm bất kỳ giải pháp truyền thông tầm gần thích hợp nào như ví dụ kết nối không dây Bluetooth hoặc kết nối có dây cổng 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 theo các phương án của sáng chế có thể lưu trữ cả dữ liệu trong dạng dữ liệu ảnh và audio và/hoặc cũng có thể lưu trữ các lệnh để triển khai trên bộ điều khiển 56. Bộ điều khiển 56 còn có thể được kết nối để hệ mạch mã hóa-giải mã 54 thích hợp để tiến hành mã hóa và giải mã của dữ liệu audio và/hoặc video hoặc hỗ trợ việc mã hóa và giải mã được thực thi bởi bộ điều khiển.

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

Thiết bị 50 có thể bao gồm hệ mạch giao diện vô tuyến 52 được kết nối với bộ điều khiển và thích hợp để tạo ra các tín hiệu truyền thông không dây ví dụ để truyền thông với mạng truyền thông tế bào, hệ thống truyền thông không dây hoặc mạng cục bộ không dây. Thiết bị 50 cũng có thể bao gồm anten 44 được kết nối với hệ mạch giao diện vô tuyến 52 để phát các tín hiệu tần số vô tuyến được tạo ra ở hệ mạch giao diện vô tuyến 52 đến (các) thiết bị khác và để thu các tín hiệu tần số vô tuyến từ (các) thiết bị khác.

Thiết bị 50 có thể bao gồm camera có thể quay hoặc phát hiện các khung riêng biệt mà sau đó được truyền đến bộ mã hóa-giải mã 54 hoặc bộ điều khiển để xử lý. Thiết bị này có thể thu dữ liệu ảnh video để xử lý từ thiết bị khác trước khi truyền và/hoặc lưu trữ. Thiết bị 50 cũng có thể nhận ảnh bằng kết nối không dây hoặc bằng kết nối có dây để mã hóa/giải mã.

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

Hệ thống 10 có thể bao gồm cả 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 được thể hiện trên Fig.3 thể hiện mạng điện thoại di động 11 và biểu diễn của internet 28. Kết nối với internet 28 có thể bao gồm, nhưng không giới hạn ở, các kết nối không dây tầm xa, các kết nối không dây tầm gần, và các kết nối có dây khác nhau bao gồm, nhưng không giới hạn ở, các đường dây điện thoại, các đường dây cáp, các đường dây điện, và các đường truyền thông tương tự.

Các thiết bị truyền thông làm ví dụ được thể hiện trong hệ thống 10 có thể bao gồm, nhưng không giới hạn ở, thiết bị điện tử hoặc thiết bị 50, kết hợp của các thiết bị hỗ trợ kỹ thuật số cá nhân (personal digital assistant - PDA) và điện thoại di động 14, PDA 16, thiết bị nhắn tin tích hợp (integrated messaging device - IMD) 18, máy tính để bàn 20, máy tính notebook 22. Thiết bị 50 có thể là cố định hoặc di động được mang bởi người đang chuyển động. Thiết bị 50 cũng có thể được định vị ở chế độ vận chuyển bao gồm, nhưng không giới hạn ở, xe hơi, xe tải, taxi, xe buýt, tàu hỏa, thuyền, máy bay, xe đạp, xe máy hoặc bất kỳ chế độ vận chuyển thích hợp tương tự nào.

Các phương án này cũng có thể được thực hiện trong hộp đổi tín hiệu truyền hình (set-top box); tức là bộ thu TV kỹ thuật số, có thể/không thể có màn hình hoặc các khả năng không dây, trong các máy tính bảng hoặc (laptop) các máy tính cá nhân (PC), có phần cứng hoặc phần mềm hoặc kết hợp của các lắp đặt 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 cài sẵn cung cấp mã hóa dựa vào phần cứng/phần mềm.

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

Các thiết bị truyền thông có thể truyền thông sử dụng các công nghệ truyền khác nhau bao gồm, nhưng không giới hạn ở, truy cập đa chia mã (code division multiple access - CDMA), các hệ thống truyền thông di động (global system for mobile communications - GSM), hệ thống viễn thông di động toàn cầu (universal mobile telecommunications system - UMTS), truy cập đa chia thời gian (time divisional multiple access - TDMA), truy cập đa chia tần số (frequency division multiple access - FDMA), giao thức điều khiển truyền – giao thức internet (transmission control protocol-internet protocol - TCP-IP), dịch vụ nhắn tin ngắn (short messaging service - SMS), dịch vụ nhắn tin đa phương tiện (multimedia messaging service - MMS), email, dịch vụ nhắn tin tức thời (instant messaging service - IMS), Bluetooth, IEEE 802.11 và bất kỳ công nghệ truyền thông không dây tương tự nào. Thiết bị truyền thông liên quan đến việc thực hiện các phương án khác nhau của sáng chế có thể truyền thông sử dụng các phương tiện khác nhau bao gồm, nhưng không giới hạn ở, các kết nối vô tuyến, hồng ngoại, laze, kết nối cáp, và bất kỳ kết nối thích hợp nào.

Bộ mã hóa-giải mã video bao gồm bộ mã hóa biến đổi video đầu vào thành biểu diễn nén tương thích để lưu trữ/truyền và bộ giải mã có thể giải nén dạng biểu diễn video nén trở lại thành dạng xem được. Bộ mã hóa video và/hoặc bộ giải mã video cũng có thể riêng biệt với nhau, tức là không cần tạo ra bộ mã hóa-giải mã. Bộ mã hóa điển hình loại bỏ một số thông tin trong chuỗi video ban đầu để biểu diễn video ở dạng nén hơn (tức là, ở tỷ lệ bit thấp hơn). Bộ mã hóa video có thể được sử dụng để mã hóa chuỗi ảnh, như được mô tả dưới đây, và bộ giải mã video có thể được sử dụng để giải mã chuỗi ảnh được mã hóa. Bộ mã hóa video hoặc phần mã hóa nội ảnh của bộ mã hóa video hoặc bộ mã hóa ảnh có thể được sử dụng để mã hóa ảnh, và bộ giải mã video hoặc phần giải mã liên ảnh của bộ giải mã video hoặc bộ giải mã ảnh có thể được sử dụng để giải mã ảnh được mã hóa.

Một số bộ mã hóa video lai hóa, ví dụ các dạng triển khai bộ mã hóa của ITU-T H.263 và H.264, mã hóa thông tin video trong hai pha. Trước hết, các giá trị điểm ảnh trong vùng hình ảnh xác định (hoặc “khối”) được dự đoán ví dụ bằng

các phương tiện bù chuyển động (tìm và chỉ báo trong một trong số các khung video được mã hóa trước đó liên quan chặt chẽ đến khối được mã hóa) hoặc bởi các phương tiện không gian (sử dụng các giá trị điểm ảnh xung quanh khối được mã hóa theo cách xác định). Thứ hai, lỗi dự đoán, tức là sự khác nhau giữa khối các điểm ảnh được dự đoán và khối các điểm ảnh ban đầu, được mã hóa. Điều này được thực hiện thông thường bằng cách biến đổi sự khác nhau trong các giá trị điểm ảnh sử dụng biến đổi được xác định (ví dụ biến đổi cosin rời rạc (Discrete Cosine Transform - DCT) hoặc biến thể của nó), lượng tử hóa các hệ số và mã hóa entropy các hệ số lượng tử hóa. Bằng cách thay đổi độ chính xác của quy trình lượng tử hóa, bộ mã hóa có thể điều khiển sự cân bằng giữa sự chính xác của biểu diễn điểm ảnh (chất lượng hình ảnh) và kích cỡ của dạng biểu diễn video được mã hóa cuối cùng (kích cỡ tệp tin hoặc tốc độ bit truyền).

Trong dự đoán theo thời gian, các nguồn dự đoán là các hình ảnh được giải mã trước đó (còn được biến đến là các hình ảnh tham chiếu). Trong sao chép nội khối (còn được biến đến là dự đoán sao chép nội khối), dự đoán được áp dụng tương tự với dự đoán theo thời gian nhưng hình ảnh tham chiếu là hình ảnh hiện thời và chỉ có các mẫu được giải mã trước đó có thể được tham chiếu trong quy trình dự đoán. Dự đoán liên lớp hoặc liên góc nhìn có thể được áp dụng tương tự với dự đoán theo thời gian, nhưng hình ảnh tham chiếu là hình ảnh được giải mã từ lớp mở rộng được khác hoặc từ góc nhìn khác, một cách tương ứng. Trong một số trường hợp, dự đoán liên ảnh có thể chỉ đề cập đến dự đoán theo thời gian, trong khi trong các trường hợp khác dự đoán liên ảnh có thể gọi chung là dự đoán theo thời gian và bất kỳ trong số sao chép nội khối, dự đoán liên lớp, và dự đoán liên góc nhìn giả định rằng chúng được thực hiện với quy trình giống hoặc tương tự với dự đoán theo thời gian. Dự đoán liên ảnh hoặc dự đoán theo thời gian đôi khi có thể được gọi là sự bù chuyển động hoặc dự đoán bù chuyển động.

Dự đoán nội ảnh sử dụng thực tế là các điểm ảnh liền kề trong cùng hình ảnh có khả năng tương quan với nhau. Dự đoán nội ảnh có thể được thực hiện trong miền không gian hoặc miền biến đổi, tức là, các giá trị mẫu hoặc các hệ số

biến đổi có thể được dự đoán. Dự đoán nội ảnh thường được sử dụng trong dự đoán nội ảnh, trong đó dự đoán liên ảnh được áp dụng.

Có các kiểu chế độ dự đoán nội ảnh khác nhau khả dụng trong mô hình mã hóa, trong đó bộ mã hóa có thể chọn và chỉ báo ra kiểu được sử dụng, ví dụ trên cơ sở khối hoặc đơn vị mã hóa. Bộ giải mã có thể giải mã chế độ dự đoán nội ảnh và khôi phục khối dự đoán được chỉ báo theo đó. Ví dụ, các chế độ dự đoán nội ảnh theo góc, mỗi chế độ cho hướng góc khác, có thể khả dụng. Dự đoán nội ảnh theo góc có thể được xem xét để ngoại suy các mẫu biên của các khối liên kề theo hướng dự đoán tuyến tính. Ngoài ra hoặc thay vào đó, chế độ dự đoán phẳng có thể khả dụng. Dự đoán phẳng có thể được xem xét để về cơ bản tạo ra khối dự đoán, trong đó mỗi mẫu của khối dự đoán có thể được xác định là trung bình của khối được đóng thẳng trong cột mẫu liền kề ở bên trái của khối hiện thời và mẫu được đóng chiều ngang trong đường mẫu liền kề ở trên khối hiện thời. Ngoài ra hoặc theo cách khác, chế độ dự đoán DC có thể khả dụng, trong đó khối dự đoán nhất thiết là giá trị mẫu trung bình của khối hoặc các khối lân cận.

Kết quả của thủ tục mã hóa là tập hợp các tham số mã hóa, như các vector chuyển động và các hệ số biến đổi lượng tử hóa. Các tham số có thể được mã hóa entropy hiệu quả hơn nếu chúng được dự đoán trước hết từ các tham số lân cận theo không gian hoặc thời gian. Ví dụ, vector chuyển động có thể được dự đoán từ các vector chuyển động liền kề theo không gian và chỉ sự khác nhau tương đối với bộ dự đoán vector chuyển động có thể được mã hóa. Dự đoán về các tham số mã hóa và dự đoán nội ảnh có thể được gọi chung là dự đoán trong hình ảnh.

Tiêu chuẩn H.264/AVC được phát triển bởi Nhóm Video Kết hợp (Joint Video Team - JVT) của nhóm chuyên gia mã hóa video (Video Coding Experts Group - VCEG) của bộ phận tiêu chuẩn hóa viễn thông của Hiệp hội Viễn Thông Quốc tế (International Telecommunication Union - ITU-T) và Nhóm Chuyên gia Hình ảnh Động (Moving Picture Experts Group - MPEG) của tổ chức tiêu chuẩn hóa Quốc tế (International Organisation for Standardization - ISO) / Ủy ban kỹ thuật điện Quốc tế (International Electrotechnical Commission - IEC). Tiêu chuẩn H.264/AVC được công bố bởi cả hai tổ chức tiêu chuẩn mẹ, và thường được gọi

là Khuyến cáo ITU-T H.264 và tiêu chuẩn quốc tế ISO/IEC 14496-10, cũng được biết là Mã hóa Video Tiên tiến MPEG-4 Phần 10. Có nhiều phiên bản của tiêu chuẩn H.264/AVC, có nhiều phiên bản của tiêu chuẩn H.264/AVC, tích hợp phần mở rộng hoặc đặc điểm mới vào bản đặc tả. Nội dung mở rộng này bao gồm Mã hóa video mở rộng được (Scalable Video Coding - SVC) và Mã hóa Video đa góc nhìn (Multiview Video Coding - MVC).

Phiên bản 1 của tiêu chuẩn mã hóa video hiệu suất cao (H.265/HEVC còn được biến đến là HEVC) được phát triển bởi nhóm cộng tác kết hợp - Mã hóa Video (Joint Collaborative Team – Video Coding - JCT-VC) của VCEG và MPEG. Tiêu chuẩn này đã công bố bởi các tổ chức tiêu chuẩn mẹ, và thường được gọi là khuyến cáo ITU-T H.265 và tiêu chuẩn quốc tế ISO/IEC 23008-2, cũng được biết là mã hóa video hiệu quả cao (High Efficiency Video Coding - HEVC) MPEG-H Phần 2. Phiên bản 2 của H.265/HEVC được bao gồm các mở rộng phạm vi mở rộng được, đa góc nhìn, và phạm vi chính xác, có thể được viết tắt là SHVC, MV-HEVC, và REXT, tương ứng. Phiên bản 2 của H.265/HEVC đã được công bố của khuyến nghị ITU-T H.265 (10/2014) và dưới dạng Ấn bản 2 của ISO/IEC 23008-2. Hiện tại có các dự án chuẩn hóa để phát triển các nội dung mở rộng khác là H.265/HEVC, bao gồm các nội dung mở rộng mã hóa nội dung ba chiều và màn hình, có thể được viết tắt là 3D-HEVC và SCC, một cách tương ứng.

SHVC, MV-HEVC, và 3D-HEVC sử dụng đặc tả cơ sở chung, được xác định trong Phụ lục F của phiên bản 2 của tiêu chuẩn HEVC. Cơ sở chung này bao gồm ví dụ cú pháp cấp độ cao và các ngữ nghĩa ví dụ xác định một vài trong số các đặc tính của các lớp của dòng bit, như các phụ thuộc liên lớp, cũng như các quy trình giải mã, như cấu trúc danh sách hình ảnh tham chiếu bao gồm việc thu nhận các hình ảnh tham chiếu liên lớp và số lượng thứ tự hình ảnh cho dòng bit đa lớp. Phụ lục F cũng có thể được sử dụng trong các phần mở rộng đa lớp nối tiếp tiềm năng của HEVC. Cần 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 với tham chiếu đến các phần

mở rộng cụ thể, như SHVC và/hoặc MV-HEVC, chúng thường áp dụng được cho bất kỳ các phần mở rộng đa lớp nào của HEVC, và thậm chí tổng quát hơn cho mô hình mã hóa video đa lớp.

Một số định nghĩa chính, các cấu trúc dòng bit và mã hóa, và các khái niệm về 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 triển khai. Một số định nghĩa chính, các cấu trúc dòng bit và mã hóa, và các khái niệm của H.264/AVC giống như Theo HEVC – do đó, chúng được mô tả sau đây một cách đồng thời. Các khía cạnh của sáng chế không bị giới hạn với H.264/AVC hoặc HEVC, nhưng tốt hơn là phần mô tả này được đưa ra đối với cơ sở khả dĩ mà sáng chế có thể được nhận biết một phần hoặc đầy đủ.

Tương tự với các tiêu chuẩn mã hóa video trước đây, cú pháp dòng bit và các ngữ nghĩa cũng như quy trình giải mã cho các dòng bit không lỗi được xác 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 thích hợp. Sự thích ứng dòng bit và bộ giải mã có thể được xác minh bằng bộ giải mã tham chiếu giả định (Hypothetical Reference Decoder - HRD). Các tiêu chuẩn này chứa các công cụ mã hóa giúp giải quyết lỗi truyền và mất mát dữ liệu khi truyền, nhưng việc sử dụng các công cụ này trong bước mã hóa là tùy chọn và không có quy trình giải mã được xác định đối với các dòng bit sai sót.

Trong mô tả các tiêu chuẩn có sẵn cũng như trong việc mô tả các phương án làm ví dụ, thành phần cú pháp có thể được xác định là yếu tố dữ liệu được biểu diễn trong dòng bit. Cấu trúc cú pháp có thể được quy định là không có yếu cú pháp hoặc nhiều thành phần cú pháp có mặt cùng nhau trong dòng bit theo thứ tự xác định. Trong mô tả các tiêu chuẩn có sẵn cũng như trong mô tả các phương án làm ví dụ, cụm từ “bằng các phương tiện bên ngoài” hoặc “thông qua các 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 đưa ra “bằng các phương tiện bên ngoài” cho quy trình giải mã. Cụm từ “bằng các phương tiện

bên ngoài” có thể chỉ báo rằng thực thể này không có trong dòng bit được tạo ra bởi bộ mã hóa, nhưng tốt hơn là được truyền ra bên ngoài từ dòng bit ví dụ sử dụng giao thức điều khiển. Nó có thể có nghĩa bao gồm hoặc thay thế là thực thể này không được tạo ra bởi bộ mã hóa, nhưng có thể được tạo ra ví dụ trong bộ phát hoặc logic điều khiển giải mã hoặc tương tự sử dụng bộ giải mã. Bộ giải mã có thể có giao diện để nhập vào các phương tiện bên ngoài, như các giá trị biến.

Đơn vị sơ cấp đối với đầu vào cho bộ mã hóa H.264/AVC hoặc HEVC và đầu ra của bộ giải mã H.264/AVC hoặc HEVC, một cách tương ứng, là hình ảnh. Hình ảnh đã cho dưới dạng đầu vào đến bộ mã hóa cũng có thể được gọi là hình ảnh nguồn, và hình ảnh được giải mã bởi giải mã có thể được gọi là hình ảnh được giải mã.

Mỗi trong số các hình ảnh nguồn và các hình ảnh được giải mã đều bao gồm một hoặc nhiều mảng mẫu, như một trong số các tập hợp của các mảng mẫu sau:

- chỉ có luma (Y) (đơn sắc).
- luma và hai chroma (YCbCr hoặc YCgCo).
- lục, lam và đỏ (GBR, cũng được biết là RGB).
- các mảng biểu diễn các mẫu đơn sắc không xác định hoặc các các mẫu ba màu cơ bản khác (ví dụ, YZX, cũng được biết là XYZ).

Sau đây, các mảng này có thể được gọi là luma (hoặc L hoặc Y) và chroma, trong đó các mảng hai chroma có thể được gọi là Cb và Cr; bất kể phương pháp biểu diễn màu thực trong việc sử dụng. Phương pháp biểu diễn màu thực trong việc sử dụng có thể được chỉ báo ví dụ trong dòng bit được mã hóa ví dụ 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 một mảng hoặc một mẫu từ một trong số ba mảng mẫu (luma và hai chroma) hoặc mảng này hoặc một mẫu đơn của mảng bao gồm hình ảnh trong định dạng đơn sắc.

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

- Trong lấy mẫu đơn sắc chỉ có một mảng mẫu, có thể được coi như định là mảng luma.
- Trong lấy mẫu 4:2:0, mỗi trong số hai mảng chroma có nửa chiều cao và nửa chiều rộng của mảng luma.
- Trong lấy mẫu 4:2:2, mỗi trong số hai mảng chroma có cùng chiều cao và nửa chiều rộng của mảng luma.
- Trong lấy mẫu 4:4:4 khi không có các mặt phẳng màu riêng biệt được sử dụng, mỗi trong số hai mảng chroma có cùng chiều cao và chiều rộng với mảng luma.

Trong H.264/AVC và HEVC, có thể mã hóa các mảng mẫu dưới dạng các mặt phẳng màu riêng biệt thành dòng bit và giải mã tương ứng các mặt phẳng màu được mã hóa riêng biệt từ dòng bit. Khi các mặt phẳng màu riêng biệt được sử dụng, mỗi một trong số chúng được xử lý riêng biệt (bằng bộ mã hóa và/hoặc bộ giải mã) dưới dạng hình ảnh nhờ lấy mẫu đơn sắc.

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

Trong H.264/AVC, khối macro là khối 16x16 của các mẫu luma và các khối tương ứng của các mẫu chroma. Ví dụ, trong mô hình mẫu 4:2:0, khối macro chứa một khối 8x8 trong số các mẫu chroma trên mỗi thành phần chroma. Trong H.264/AVC, hình ảnh được phân vùng thành một hoặc nhiều nhóm lát, và nhóm

lát chứa một hoặc nhiều lát. Trong H.264/AVC, lát bao gồm số nguyên của các khối macro theo thứ tự liên tiếp theo sự quét mảnh trong nhóm lát cụ thể.

Khi mô tả hoạt động mã hóa và/hoặc giải mã HEVC, các thuật ngữ sau đây có thể được sử dụng. Khối mã hóa có thể được định nghĩa là khối các mẫu $N \times N$ đối với một số giá trị của N sao cho việc chia khối cây mã hóa thành các khối mã hóa là phân vùng. Khối cây mã hóa (coding tree block - CTB) có thể được định nghĩa là các mẫu khối các mẫu $N \times N$ đối với một số giá trị của N sao cho việc phân chia một thành phần thành các khối cây mã hóa là phân vùng. Đơn vị cây mã hóa (coding tree unit - CTU) có thể được định nghĩa là khối cây mã hóa của các mẫu luma, hai khối cây mã hóa tương ứng của các mẫu chroma của hình ảnh mà có ba mảng mẫu, hoặc khối cây mã hóa của các mẫu có hình ảnh đơn sắc hoặc hình ảnh được mã hóa sử dụng ba mặt phẳng màu riêng biệt và các cấu trúc cú pháp được sử dụng để mã hóa các mẫu này. Đơn vị mã hóa (coding unit - CU) có thể được định nghĩa là khối mã hóa của các mẫu luma, hai khối mã hóa tương ứng của các mẫu chroma của hình ảnh có ba mảng mẫu, hoặc các mẫu khối mã hóa có hình ảnh đơn sắc hoặc hình ảnh được mã hóa sử dụng ba mặt phẳng màu riêng biệt và các cấu trúc cú pháp được sử dụng để mã hóa các mẫu này.

Trong một số bộ mã hóa-giải mã video, như bộ mã hóa-giải mã mã hóa video hiệu quả cao (High Efficiency Video Coding - HEVC), các hình ảnh video được chia thành các đơn vị mã hóa (CU) bao phủ vùng hình ảnh. CU bao gồm một hoặc nhiều đơn vị dự đoán (PU) quy định quy trình dự đoán đối với các mẫu trong CU và một hoặc nhiều đơn vị 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 các mẫu có khối vuông có kích thước có thể được chọn từ tập hợp các kích thước CU khả dĩ định trước. CU có kích thước cho phép tối đa có thể được gọi là LCU (đơn vị mã hóa lớn nhất – largest coding unit) hoặc đơn vị cây mã hóa (CTU) và hình ảnh video được chia thành các LCU không xếp chồng. LCU còn có thể được chia thành kết hợp của các CU nhỏ hơn, ví dụ bằng cách chia LCU và các CU thu được theo cách đệ quy. Mỗi CU thu được thường có ít nhất một PU và ít nhất một TU được kết hợp với nó. Mỗi PU và TU còn có thể được chia thành các PU và

TU nhỏ hơn để làm tăng độ chi tiết của các quy trình dự đoán và quy trình mã hóa lỗi dự đoán, một cách tương ứng. Mỗi PU có thông tin dự đoán được kết hợp với nó mà quy định kiểu dự đoán nào được áp dụng đối với các điểm ảnh trong PU đó (ví dụ thông tin vectơ chuyển động đối với các PU được dự đoán liên ảnh và thông tin định hướng dự đoán nội ảnh đối với các PU được dự đoán nội ảnh).

Mỗi TU có thể được kết hợp với thông tin mô tả quy trình giải mã lỗi dự đoán đối với các mẫu trong TU nêu trên (bao gồm ví dụ thông tin hệ số DCT). Có thể được báo hiệu ở cấp độ CU liệu việc mã hóa lỗi dự đoán có được áp dụng hay không cho mỗi CU. Trong trường hợp không có phân dư lỗi dự đoán được kết hợp với CU, có thể xem là không có TU đối với CU nêu trên. Việc chia ảnh thành các CU, và việc chia các CU thành các PU và các TU có thể được báo hiệu trong dòng bit cho phép bộ giải mã khôi phục cấu trúc được dự định của các đơn vị này.

Theo HEVC, hình ảnh có thể được phân vùng thành các tấm (tiles), là hình chữ nhật và chứa số nguyên các LCU. Theo HEVC, việc phân vùng thành các tấm tạo thành lưới bao gồm một hoặc nhiều cột tấm và một hoặc nhiều hàng tấm. Tấm được mã hóa được căn thẳng theo byte, có thể đạt được bằng cách thêm các bit được căn thẳng theo byte ở đầu cuối của tấm được mã hóa.

Theo HEVC, lát được xác định là số nguyên của các đơn vị cây mã hóa được chứa trong một đoạn lát độc lập và tất cả các đoạn lát phụ thuộc tiếp theo (nếu có) liền trước đoạn lát độc lập tiếp theo (nếu có) trong cùng đơn vị truy cập. Theo HEVC, đoạn lát được định nghĩa là số nguyên của các đơn vị cây mã hóa theo thứ tự liên tiếp trong việc quét tấm và được chứa trong một đơn vị NAL. Việc chia mỗi hình ảnh thành các đoạn lát là việc phân vùng. Theo HEVC, đoạn lát độc lập được định nghĩa là đoạn lát mà trong đó các giá trị của các thành phần cú pháp của phần đầu đoạn lát không được suy ra từ các giá trị đối với đoạn lát liền trước, và đoạn lát phụ thuộc được định nghĩa là đoạn lát mà trong đó các giá trị của một số thành phần cú pháp của phần đầu đoạn lát được suy ra từ các giá trị cho đoạn lát độc lập liền trước theo thứ tự giải mã. Theo HEVC, phần đầu của lát được định nghĩa là phần đầu đoạn của lát của đoạn lát độc lập là đoạn lát hiện thời hoặc là đoạn lát độc lập liền trước đoạn lát phụ thuộc hiện thời, và phần đầu đoạn

của lát được định nghĩa là một phần của đoạn lát được mã hóa chứa các thành phần dữ liệu thuộc về các đơn vị cây mã hóa thứ nhất hoặc tất cả các đơn vị cây mã hóa được biểu diễn trong đoạn lát. Các CU được quét theo thứ tự quét kiểu mảnh của các LCU trong các tấm hoặc trong hình ảnh, nếu các tấm không được sử dụng. Trong LCU, các CU có thứ tự quét xác định.

Theo HEVC, tấm chứa số nguyên của các đơn vị cây mã hóa, và có thể bao gồm các đơn vị cây mã hóa được chứa trong nhiều hơn một lát. Tương tự, lát có thể bao gồm các đơn vị cây mã hóa được chứa trong nhiều hơn một tấm. Theo HEVC, tất cả các đơn vị cây mã hóa trong lát thuộc về cùng tấm và/hoặc tất cả các đơn vị cây mã hóa trong tấm thuộc về cùng lát. Hơn nữa, Theo HEVC, tất cả các đơn vị cây mã hóa trong đoạn lát thuộc về cùng tấm và/hoặc tất cả các đơn vị cây mã hóa trong tấm thuộc về cùng đoạn lát.

Tập hợp tấm bị ràng buộc chuyển động thỏa mã quy trình dự đoán liên ảnh bị ràng buộc trong bước mã hóa sao cho không có giá trị mẫu bên ngoài tập hợp tấm bị ràng buộc chuyển động, và không có giá trị mẫu ở vị trí mẫu phân đoạn thu được sử dụng một hoặc nhiều giá trị mẫu bên ngoài tập hợp tấm bị ràng buộc chuyển động, được sử dụng cho dự đoán liên ảnh của bất kỳ mẫu nào trong tập hợp tấm bị ràng buộc chuyển động.

Lưu ý rằng các vị trí mẫu được sử dụng trong dự đoán liên ảnh được làm bão hòa để vị trí ở bên ngoài hình ảnh mặc khác được làm bão hòa để trở về mẫu biên tương ứng của hình ảnh. Do đó, nếu biên của tấm cũng là biên của hình ảnh, các vectơ chuyển động có thể đi qua biên đó một cách hiệu quả hoặc vectơ chuyển động có thể thực hiện hiệu quả phép nội suy mẫu phân đoạn sẽ tham chiếu đến vị trí bên ngoài biên đó, vì các vị trí mẫu được làm bão hòa trên biên này.

Tin nhắn SEI của các tập hợp tấm bị giới hạn chuyển động theo thời gian của HEVC có thể được sử dụng để chỉ báo sự có mặt của các tập hợp tấm bị ràng buộc chuyển động trong dòng bit.

Tập hợp tấm bị ràng buộc liên lớp sao cho quy trình dự đoán liên lớp bị giới hạn trong bước mã hóa sao cho không có giá trị mẫu bên ngoài mỗi tập hợp tấm tham chiếu được kết hợp, và không có giá trị mẫu ở vị trí mẫu phân đoạn thu

được sử dụng một hoặc nhiều giá trị mẫu bên ngoài mỗi tập hợp tầm tham chiếu được kết hợp, được sử dụng cho dự đoán liên lớp của bất kỳ mẫu nào trong tập hợp tầm bị ràng buộc liên lớp.

Tin nhắn SEI của tập hợp tầm bị ràng buộc liên lớp HEVC có thể được sử dụng để chỉ báo sự có mặt của các tập hợp tầm bị ràng buộc liên lớp trong dòng bit.

Bộ giải mã khôi phục video đầu ra bằng cách áp dụng các phương tiện dự đoán tương tự với bộ mã hóa để tạo ra biểu diễn được dự đoán của các khối điểm ảnh (sử dụng thông tin chuyển động hoặc không gian được tạo ra với bộ mã hóa và được lưu trữ trong dạng biểu diễn nén) và giải mã lỗi dự đoán (phép nghịch đảo của việc mã hóa lỗi dự đoán phục hồi tín hiệu lỗi dự đoán được lượng tử hóa trong miền điểm ảnh theo không gian). Sau khi áp dụng các phương tiện dự đoán và giải mã lỗi dự đoán giải mã, bộ giải mã gộp các tín hiệu dự đoán và tín hiệu lỗi dự đoán (các giá trị điểm ảnh) để tạo ra khung video đầu ra. Bộ giải mã (và bộ mã hóa) cũng có thể áp dụng các phương tiện lọc bổ sung để cải thiện chất lượng của video đầu ra trước khi truyền nó để hiển thị và/hoặc lưu trữ nó dưới dạng tham chiếu dự đoán đối với các khung sắp tới trong chuỗi video.

Ví dụ, việc lọc có thể bao gồm một hoặc nhiều thành phần sau đây: giải khối, khoảng dịch thích ứng mẫu (sample adaptive offset - SAO), và/hoặc bộ lọc lặp thích ứng (adaptive loop filtering - ALF). H.264/AVC bao gồm giải khối, trong đó HEVC bao gồm cả giải khối và SAO.

Trong các bộ mã hóa-giải mã video thông thường, thông tin chuyển động được chỉ báo với các vectơ chuyển động được kết hợp với mỗi khối ảnh bù chuyển động, như đơn vị dự đoán. Mỗi trong số các vectơ chuyển động này biểu diễn sự dịch chuyển của khối 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 số các hình ảnh được mã hóa hoặc được giải mã trước đó. Để biểu diễn các vectơ chuyển động hiệu quả chúng thường được mã hóa khác biệt tương ứng với các vectơ chuyển động được dự đoán cụ thể của khối. Trong các bộ mã hóa-giải mã video thường các vectơ chuyển động được dự đoán được tạo ra theo cách định

trước, ví dụ tính số trung bình của các vectơ chuyển động được mã hóa hoặc giải mã của các khối liên kề. Cách khác để tạo ra các dự đoán về vectơ chuyển động là tạo ra danh sách các dự đoán ứng viên từ các khối liên kề và/hoặc các khối đồng vị trí ở các hình ảnh tham chiếu tạm thời và báo hiệu ứng viên được chọn dưới dạng bộ dự đoán vectơ chuyển động. Ngoài việc dự đoán các giá trị vectơ chuyển động ra, có thể được dự đoán (các) hình ảnh tham chiếu nào được sử dụng cho dự đoán bù chuyển động và thông tin dự đoán này có thể được biểu diễn ví dụ bởi chỉ số tham chiếu của hình ảnh được mã hóa/giải mã trước đó. Chỉ số tham chiếu thường được dự đoán từ các khối liên kề và/hoặc các khối đồng vị trí trong hình ảnh tham chiếu theo thời gian. Ngoài ra, các bộ mã hóa-giải mã video hiệu quả cao thường sử dụng cơ chế mã hóa/giải mã thông tin chuyển động bổ sung, thường được gọi là chế độ kết hợp (merging/merge mode), trong đó tất cả thông tin trường chuyển động, bao gồm vectơ chuyển động và chỉ số hình ảnh tham chiếu tương ứng cho mỗi danh sách hình ảnh tham chiếu khả dụng, được dự đoán và được sử dụng mà không có bất kỳ biến đổi/sửa đổi nào. Tương tự, dự đoán thông tin trường chuyển động được thực hiện sử dụng thông tin trường chuyển động của các khối liên kề và/hoặc các khối đồng vị trí trong các hình ảnh tham chiếu tạm thời và thông tin trường chuyển động được sử dụng được báo hiệu trong số danh sách của danh sách ứng viên trường chuyển động được làm đầy với thông tin trường chuyển động của các khối liên kề/đồng vị trí khả dụng.

Các bộ mã hóa-giải mã video thường cho phép việc sử dụng dự đoán đơn, trong đó một khối dự đoán được sử dụng đối với khối được mã hóa (giải mã), và dự đoán kép, trong đó hai khối dự đoán được kết hợp để tạo ra dự đoán đối với khối được mã hóa (giải mã). Một số bộ mã hóa-giải mã video cho phép dự đoán có trọng số, trong đó các giá trị mẫu của các khối dự đoán được lấy trọng số trước khi bổ sung thông tin phần dư. Ví dụ, hệ số trọng số nhân và khoảng dịch bổ sung có thể được áp dụng. Trong dự đoán có trọng số tường minh, được cho phép bởi một số bộ mã hóa-giải mã video, hệ số trọng số và khoảng dịch có thể được mã hóa ví dụ trong phần đầu lát đối với mỗi chỉ số hình ảnh tham chiếu được cho phép. Trong dự đoán có trọng số ngầm định, được cho phép bởi một số bộ mã hóa-giải mã video, các hệ số trọng số và/hoặc các khoảng dịch không được mã

hóa nhưng thu được ví dụ dựa trên các khoảng cách số đếm thứ tự hình ảnh (picture order count - POC) tương đối của các hình ảnh tham chiếu.

Trong các bộ mã hóa-giải mã video thông thường, phần dư dự đoán sau khi bù chuyển động trước hết được biến đổi với nhân biến đổi (giống DCT) và sau đó được mã hóa. Lý do cho điều này là thường tồn tại một số tương quan trong số phần dư và biến đổi trong nhiều trường hợp có thể giúp giảm sự tương quan này và cung cấp việc mã hóa hiệu quả hơn.

Các bộ mã hóa video thường sử dụng các hàm giá trị Lagrangian để tìm các chế độ mã hóa tối ưu, ví dụ chế độ khối macro mong muốn và các vector chuyển động được kết hợp. Kiểu hàm giá trị này sử dụng hệ số trọng số λ để kết hợp với độ biến dạng ảnh (chính xác hoặc được ước lượng) do các phương pháp mã hóa mất mát và lượng thông tin (chính xác hoặc được ước lượng) được yêu cầu để biểu diễn các giá trị điểm ảnh trong vùng ảnh:

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

trong đó C là giá trị Lagrangian được tối thiểu hóa, D là độ biến dạng ảnh (ví dụ sai số toàn phương trung bình) với chế độ và các vector chuyển động được xem xét, và R số lượng các bit cần thiết để biểu diễn dữ liệu được yêu cầu để khôi phục khối ảnh trong bộ giải mã (bao gồm lượng dữ liệu để biểu diễn các vector chuyển động ứng viên).

Các tiêu chuẩn và đặc tả mã hóa video có thể cho phép các bộ mã hóa chia hình ảnh mã hóa thành các lát mã hóa hoặc tương tự. Dự đoán trong hình ảnh thường không khả dụng đối với các biên lát. Do đó, các lát có thể được xem là cách để tách hình ảnh được mã hóa thành các phần giải mã được độc lập. Trong H.264/AVC và HEVC, dự đoán trong hình ảnh có thể không khả dụng đối với các biên lát. Do đó, các lát có thể được xem là cách để tách hình ảnh được mã hóa thành các phần giải mã được độc lập, và do đó các lát thường được xem là các đơn vị sơ cấp để truyền. Trong nhiều trường hợp, các bộ mã hóa có thể chỉ báo trong dòng bit kiểu dự đoán trong hình ảnh nào bị bất hoạt đối với các biên lát, và hoạt động bộ giải mã sử dụng thông tin này làm ví dụ khi kết luận các nguồn dự đoán nào là khả dụng. Ví dụ, các mẫu từ khối macro lân cận hoặc CU có thể được

xem là không khả dụng cho dự đoán nội ảnh, nếu khối macro lân cận hoặc CU ở tại lát khác.

Đơn vị sơ cấp đối với đầu ra của bộ mã hóa H.264/AVC hoặc HEVC và đầu vào của bộ giải mã H.264/AVC hoặc HEVC, một cách tương ứng, là đơn vị lớp trừu tượng mạng (Network Abstraction Layer - NAL). Để vận chuyển qua các mạng định hướng gói hoặc lưu trữ trong các tệp có cấu trúc, các đơn vị NAL có thể được gói vào các gói hoặc các cấu trúc tương tự. Đơn vị NAL có thể được định nghĩa là cấu trúc cú pháp chứa chỉ báo về kiểu dữ liệu để theo dõi và các byte chứa dữ liệu trong dạng RBSP được phân tán khi cần thiết với các byte ngăn ngừa mô phỏng mã hóa bắt đầu. Phần tải trình tự byte thô (raw byte sequence payload - RBSP) có thể được định nghĩa là cấu trúc cú pháp chứa số nguyên các byte được gói trong đơn vị NAL. RBSP là trống hoặc có dạng chuỗi của các bit dữ liệu chứa các thành phần cú pháp được theo sau bởi bit dừng RBSP và được theo sau bởi không (0) hoặc nhiều hơn bit liên tiếp bằng 0. Các đơn vị NAL bao gồm phần đầu và phần tải.

Theo HEVC, phần đầu đơn vị NAL hai byte được sử dụng cho tất cả các kiểu đơn vị NAL xác định. Phần đầu đơn vị NAL chứa một bit dành riêng, chỉ báo kiểu đơn vị NAL sáu bit, chỉ báo `nuh_temporal_id_plus1` ba bit cho cấp độ thời gian (có thể được yêu cầu lớn hơn hoặc bằng 1) và thành phần cú pháp `nuh_layer_id` sáu bit. Thành phần cú pháp `temporal_id_plus1` có thể được xem là mã nhận dạng theo thời gian cho đơn vị NAL, và biến `TemporalId` trên cơ sở không có thể thu được như sau: $\text{TemporalId} = \text{temporal_id_plus1} - 1$. `TemporalId` bằng 0 tương ứng với mức độ thời gian thấp nhất. Giá trị của `temporal_id_plus1` được yêu cầu không phải là không để tránh mô phỏng mã bắt đầu liên quan đến hai byte phần đầu đơn vị NAL. Dòng bit được tạo ra bằng cách loại trừ tất cả các đơn vị NAL VCL có `TemporalId` lớn hơn hoặc bằng giá trị được chọn và bao gồm tất cả các đơn vị NAL VCL khác duy trì thích hợp. Do đó, hình ảnh có `TemporalId` bằng TID không sử dụng bất kỳ hình ảnh nào có `TemporalId` lớn hơn TID dưới dạng tham chiếu dự đoán liên ảnh. Lớp con hoặc lớp con theo thời gian có thể được định nghĩa là lớp mở rộng được theo thời gian của dòng bit mở rộng

được theo thời gian, bao gồm các đơn vị NAL VCL với giá trị cụ thể của biến TemporalId và các đơn vị NAL phi VCL được kết hợp. nuh_layer_id có thể được hiểu là mã nhận dạng lớp khả năng mở rộng được.

Các đơn vị NAL có thể được phân loại thành các đơn vị NAL lớp mã hóa video (VCL) và các đơn vị NAL phi VCL. Trong H.264/AVC, các đơn vị NAL lát mã hóa chứa các thành phần cú pháp biểu diễn một hoặc nhiều các khối macro được mã hóa, mỗi trong số đó tương ứng với khối có các mẫu trong hình ảnh không nén. Theo HEVC, các đơn vị NAL VCL chứa các thành phần cú pháp biểu diễn một hoặc nhiều CU.

Theo HEVC, đơn vị NAL lát mã hóa có thể được chỉ báo là một trong số các kiểu sau đây:

nal_unit_type	Tên của nal_unit_type	Nội dung của đơn vị NAL và cấu trúc cú pháp RBSP
0, 1	TRAIL_N, TRAIL_R	Đoạn lát mã hóa của hình ảnh theo sau phi TSA, phi STSA slice_segment_layer_rbsp()
2, 3	TSA_N, TSA_R	Đoạn lát mã hóa của hình ảnh TSA slice_segment_layer_rbsp()
4, 5	STSA_N, STSA_R	Đoạn lát mã hóa của hình ảnh STSA slice_layer_rbsp()
6, 7	RADL_N, RADL_R	Đoạn lát mã hóa của hình ảnh RADL slice_layer_rbsp()
8, 9	RASL_N, RASL_R,	Đoạn lát mã hóa của hình ảnh RASL slice_layer_rbsp()
10, 12, 14	RSV_VCL_N10 RSV_VCL_N12 RSV_VCL_N14	Được dự phòng // các kiểu đơn vị NAL VCL không tham chiếu phi RAP được dự phòng
11, 13, 15	RSV_VCL_R11 RSV_VCL_R13 RSV_VCL_R15	Được dự phòng // các kiểu đơn vị NAL VCL tham chiếu phi RAP được dự phòng
16, 17, 18	BLA_W_LP IDR_W_RADL BLA_N_LP	Đoạn lát mã hóa của hình ảnh BLA slice_segment_layer_rbsp()

19, 20	IDR_W_RADL IDR_N_LP	Đoạn lát mã hóa của hình ảnh IDR slice_segment_layer_rbsp()
21	CRA_NUT	Đoạn lát mã hóa của hình ảnh CRA slice_segment_layer_rbsp()
22, 23	RSV_IRAP_VCL22.. RSV_IRAP_VCL23	Được dự phòng // các kiểu đơn vị RAP VCL NAL được dự phòng
24..31	RSV_VCL24.. RSV_VCL31	Được dự phòng // các kiểu đơn vị VCL NAL phi RAP được dự phòng

Theo HEVC, các dạng viết tắt của các kiểu hình ảnh có thể được định nghĩa như sau: hình ảnh theo sau (TRAIL), truy cập lớp con tạm thời (TSA), truy cập lớp con tạm thời theo bước (STSA), hình ảnh dẫn có thể giải mã truy cập ngẫu nhiên (RADL), hình ảnh dẫn bỏ qua truy cập ngẫu nhiên (Random Access Skipped Leading - RASL), hình ảnh truy cập kết nối đứt (Broken Link Access - BLA), hình ảnh làm tươi giải mã tức thời (Instantaneous Decoding Refresh - IDR), hình ảnh truy cập ngẫu nhiên rõ ràng (Clean Random Access - CRA).

Hình ảnh điểm truy cập điểm ngẫu nhiên (Random Access Point - RAP), cũng có thể được gọi là hình ảnh điểm truy cập ngẫu nhiên nội ảnh (Intra Random Access Point - IRAP), là hình ảnh trong đó mỗi lát hoặc đoạn lát có nal_unit_type nằm trong khoảng từ 16 đến 23 bao gồm đầu mút. Hình ảnh IRAP trong lớp độc lập không tham chiếu đến bất kỳ hình ảnh nào ngoài chính nó để dự đoán liên ảnh trong quy trình giải mã của nó. Khi không sử dụng sao chép nội khối, hình ảnh IRAP trong lớp độc lập chỉ chứa các lát được mã hóa nội ảnh. Hình ảnh IRAP thuộc về lớp dự đoán với giá trị nuh_layer_id currLayerId có thể chứa các lát P, B, và I, không thể sử dụng dự đoán liên ảnh 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 thời 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 tại lớp cơ sở này. Với điều kiện là các tập hợp các tham số cần thiết khả dụng khi chúng cần được kích hoạt, hình ảnh IRAP tại lớp độc lập và tất cả các hình ảnh không phải RASL liên tiếp ở lớp

độc lập theo thứ tự giải mã có thể được giải mã chính xác mà không thực hiện quy trình giải mã của bất kỳ các hình ảnh nào liền trước hình ảnh IRAP theo thứ tự giải mã. Hình ảnh IRAP thuộc về lớp dự đoán với giá trị `nuh_layer_id` bằng `currLayerId` và tất cả các hình ảnh không phải RASL nối tiếp với `nuh_layer_id` bằng `currLayerId` theo thứ tự giải mã có thể được giải mã chính xác mà không thực hiện quy trình giải mã của bất kỳ các hình ảnh nào với `nuh_layer_id` bằng `currLayerId` liền trước hình ảnh IRAP theo thứ tự giải mã, khi các tập hợp các tham số cần thiết 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 bắt đầu (tức là khi `LayerInitializedFlag[ refLayerId ]` bằng 1 đối với `refLayerId` bằng tất cả giá trị `nuh_layer_id` của các lớp tham chiếu trực tiếp của lớp với `nuh_layer_id` bằng `currLayerId`). Có thể là có các hình ảnh trong dòng bit chỉ chứa các lát được nội ảnh không phải là các hình ảnh IRAP.

Theo HEVC hình ảnh CRA có thể là hình ảnh thứ nhất trong dòng bit theo thứ tự giải mã, hoặc có thể xuất hiện sau trong dòng bit. Các hình ảnh CRA Theo HEVC cho phép cái gọi là các hình ảnh dẫn theo sau hình ảnh CRA theo thứ tự giải mã nhưng liền trước nó theo thứ tự kết xuất. Một vài trong số các hình ảnh dẫn, được gọi là các hình ảnh RASL, có thể sử dụng các hình ảnh được giải mã trước hình ảnh CRA dưới dạng tham chiếu. Các hình ảnh theo sau hình ảnh CRA theo cả thứ tự giải mã và kết xuất giải mã được nếu truy cập ngẫu nhiên được thực hiện tại hình ảnh CRA, và do đó truy cập ngẫu nhiên rõ ràng đạt được tương tự với chức năng truy cập ngẫu nhiên rõ ràng của hình ảnh IDR.

Hình ảnh CRA có thể có các hình ảnh RADL hoặc RASL được kết hợp. Khi hình ảnh CRA là hình ảnh thứ nhất trong dòng bit theo thứ tự giải mã, hình ảnh CRA là hình ảnh thứ nhất của chuỗi video mã hóa theo thứ tự giải mã, và bất kỳ các hình ảnh RASL được kết hợp nào không được kết xuất bởi bộ giải mã và không thể giải mã được, vì chúng có thể chứa các tham chiếu đến các hình ảnh không có mặt trong dòng bit.

Hình ảnh dẫn là hình ảnh liền trước hình ảnh RAP được kết hợp theo thứ tự kết xuất. Hình ảnh RAP được kết hợp là hình ảnh RAP trước đó theo thứ tự giải mã (nếu có mặt). Hình ảnh dẫn là hình ảnh RADL hoặc hình ảnh RASL.

Tất cả các hình ảnh RASL là các hình ảnh dẫn của hình ảnh BLA hoặc CRA được kết hợp. Khi hình ảnh RAP được kết hợp là hình ảnh BLA hoặc là hình ảnh thứ nhất được mã hóa trong dòng bit, hình ảnh RASL không được kết xuất và không thể giải mã được chính xác, dưới dạng hình ảnh RASL có thể chứa các tham chiếu đến các hình ảnh không có mặt trong dòng bit. Tuy nhiên, hình ảnh RASL có thể được giải mã chính xác nếu việc giải mã được bắt đầu từ hình ảnh RAP trước hình ảnh RAP được kết hợp của hình ảnh RASL. Các hình ảnh RASL không được sử dụng dưới dạng các hình ảnh tham chiếu đối với quy trình giải mã của các hình ảnh không phải RASL. Khi có mặt, tất cả các hình ảnh RASL liền trước, theo thứ tự giải mã, tất cả các hình ảnh theo sau của cùng một hình ảnh RAP được kết hợp. Trong một số bản nháp của tiêu chuẩn HEVC, hình ảnh RASL thường được gọi là hình ảnh được gắn thẻ để loại bỏ (Tagged for Discard - TFD).

Tất cả các hình ảnh RADL là các hình ảnh dẫn. Các hình ảnh RADL không được sử dụng dưới dạng các hình ảnh tham chiếu cho quy trình giải mã các hình ảnh theo sau của cùng một hình ảnh RAP được kết hợp. Khi có mặt, tất cả các hình ảnh RADL liền trước, theo thứ tự giải mã, tất cả các hình ảnh theo sau của cùng một hình ảnh RAP được kết hợp. Các hình ảnh RADL không tham chiếu đến bất kỳ hình ảnh nào liền trước hình ảnh RAP được kết hợp theo thứ tự giải mã và do đó có thể được giải mã chính xác khi việc giải mã mở đầu từ hình ảnh RAP được kết hợp.

Khi một phần của dòng bit bắt đầu từ hình ảnh CRA được bao gồm trong dòng bit khác, các hình ảnh RASL được kết hợp với hình ảnh CRA không thể giải mã chính xác được, vì một vài trong số các hình ảnh tham chiếu của chúng không thể có mặt trong dòng bit kết hợp. Để thực hiện hoạt động ghép nối một cách dễ dàng, kiểu đơn vị NAL của hình ảnh CRA có thể được thay đổi để chỉ báo rằng đó là hình ảnh BLA. Các hình ảnh RASL được kết hợp với hình ảnh BLA không

thể giải mã được chính xác do đó không được kết xuất/hiển thị. Hơn nữa, các hình ảnh RASL được kết hợp với hình ảnh BLA có thể được bỏ qua khỏi việc giải mã.

Hình ảnh BLA có thể là hình ảnh thứ nhất trong dòng bit theo thứ tự giải mã, hoặc có thể xuất hiện sau đó trong dòng bit. Mỗi hình ảnh BLA bắt đầu chuỗi video mã hóa mới, và có hiệu quả tương tự trong quy trình giải mã dưới dạng hình ảnh IDR. Tuy nhiên, hình ảnh BLA chứa các thành phần cú pháp xác định tập hợp hình ảnh tham chiếu không rỗng. Khi hình ảnh BLA có `nal_unit_type` bằng `BLA_W_LP`, nó có thể có các hình ảnh RASL được kết hợp, vốn không được kết xuất bởi bộ giải mã và không thể giải mã được, do 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`, thì cũng có thể có các hình ảnh RADL được kết hợp, mà được xác định là cần được giải mã. Khi hình ảnh BLA có `nal_unit_type` bằng `BLA_W_RADL`, nó không có các hình ảnh RASL được kết hợp nhưng có thể có các hình ảnh RADL được kết hợp, mà xác định là cần được giải mã. Khi hình ảnh BLA có `nal_unit_type` bằng `BLA_N_LP`, thì không có bất kỳ hình ảnh dẫn được kết hợp nào.

Hình ảnh IDR có `nal_unit_type` bằng `IDR_N_LP` không có các hình ảnh dẫn được kết hợp có mặt trong dòng bit. Hình ảnh IDR có `nal_unit_type` bằng `IDR_W_LP` không có các hình ảnh RASL được kết hợp có mặt trong dòng bit, nhưng có thể có các hình ảnh RADL được kết hợp 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 dưới dạng tham chiếu đối với bất kỳ hình ảnh khác nào của cùng một lớp con theo thời gian. Tức là, theo 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 được bao gồm trong bất kỳ trong số `RefPicSetStCurrBefore`, `RefPicSetStCurrAfter` và `RefPicSetLtCurr` của bất kỳ hình ảnh nào với cùng một giá trị của `TemporalId`. Hình ảnh được mã hóa có `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 bị loại bỏ không ảnh hưởng đến khả năng giải mã của các hình ảnh khác với cùng một giá trị của TemporalId.

Hình ảnh theo sau có thể được xác định là hình ảnh theo sau hình ảnh RAP được kết hợp theo thứ tự kết xuất. Bất kỳ hình ảnh nào là hình ảnh theo sau không có nal_unit_type bằng RADL_N, RADL_R, RASL_N hoặc RASL_R. Bất kỳ hình ảnh nào là hình ảnh dẫn có thể bị ràng buộc là liền trước, theo thứ tự giải mã, tất cả các hình ảnh theo sau được kết hợp với cùng một hình ảnh RAP. Không có các hình ảnh RASL có mặt trong dòng bit được kết hợp với hình ảnh BLA có nal_unit_type bằng BLA_W_RADL hoặc BLA_N_LP. Không có hình ảnh RADL có mặt trong dòng bit được kết hợp với hình ảnh BLA có nal_unit_type bằng BLA_N_LP hoặc được kết hợp với hình ảnh IDR có nal_unit_type bằng IDR_N_LP. Bất kỳ hình ảnh nào RASL được kết hợp với hình ảnh CRA hoặc BLA có thể bị ràng buộc liền trước bất kỳ hình ảnh RADL nào được kết hợp với hình ảnh CRA hoặc BLA theo thứ tự kết xuất. Bất kỳ hình ảnh RASL nào được kết hợp với hình ảnh CRA có thể bị ràng buộc là theo sau, theo thứ tự kết xuất, bất kỳ hình ảnh RAP khác nào liền trước hình ảnh CRA theo thứ tự giải mã.

Theo HEVC có hai kiểu hình ảnh, các kiểu hình ảnh TSA và STSA có thể được sử dụng để chỉ báo các điểm chuyển lớp con theo thời gian. Nếu các lớp con theo thời gian với TemporalId lên đến N được giải mã cho đến khi hình ảnh TSA hoặc STSA (loại trừ) 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 (theo thứ tự giải mã) có TemporalId bằng N+1. Kiểu hình ảnh TSA có thể chịu các giới hạn đối với chính hình ảnh TSA và tất cả các hình ảnh trong cùng lớp con theo sau hình ảnh TSA theo thứ tự giải mã. Không có hình ảnh nào trong số các hình ảnh này được phép sử dụng dự đoán liên ảnh từ bất kỳ hình ảnh nào trong cùng lớp con liền trước hình ảnh TSA theo thứ tự giải mã. Định nghĩa TSA còn có thể áp đặt các giới hạn đối với các hình ảnh ở các lớp con cao hơn theo sau hình ảnh TSA theo thứ tự giải mã. Không có hình ảnh nào trong số các hình ảnh này được cho phép tham chiếu đến hình ảnh liền trước hình ảnh TSA theo thứ tự giải mã nếu hình ảnh đó thuộc về cùng lớp con hoặc lớp con cao hơn với hình ảnh TSA.

Các hình ảnh TSA có TemporalId lớn hơn 0. STSA tương tự với hình ảnh TSA nhưng áp đặt các giới hạn đối với các hình ảnh ở các lớp con cao hơn theo sau hình ảnh STSA theo thứ tự giải mã và do đó chỉ cho phép việc chuyển lên lớp con mà hình ảnh STSA có mặt.

Ví dụ, đơn vị NAL phi VCL có thể là một trong số các kiểu sau đây: tập hợp các tham số trình tự, tập hợp các tham số hình ảnh, đơn vị NAL thông tin cải tiến bổ sung (SEI), bộ phân tách đơn vị truy cập, đầu của đơn vị NAL chuỗi, một đầu của đơn vị NAL dòng bit, hoặc đơn vị NAL dữ liệu lọc. Các tập hợp các tham số có thể cần thiết cho việc khôi phục các hình ảnh được giải mã, trong đó nhiều trong số các đơn vị NAL phi VCL khác không cần thiết để khôi phục các giá trị mẫu được giải mã.

Các tham số giữ không đổi thông qua chuỗi video mã hóa có thể được bao gồm trong tập hợp các tham số trình tự. Ngoài các tham số này có thể cần thiết với quy trình giải mã, tập hợp các tham số trình tự này có thể chứa thông tin sử dụng video (VUI) một cách tùy chọn, bao gồm các tham số có thể quan trọng cho bước đệm, bước định thời tạo hình ảnh, bước tạo, và dự trữ tài nguyên. Theo HEVC, tập hợp các tham số trình tự RBSP bao gồm các tham số có thể được tham chiếu với một hoặc nhiều tập hợp các tham số hình ảnh RBSP hoặc một hoặc nhiều đơn vị NAL SEI chứa tin nhắn SEI chu kỳ đệm. Tập hợp các tham số hình ảnh chứa các tham số này có thể không đổi trong các hình ảnh được mã hóa. Tập hợp các tham số hình ảnh RBSP có thể bao gồm các tham số mà có thể được tham chiếu với các đơn vị NAL lát mã hóa của một hoặc nhiều hình ảnh được mã hóa.

Theo HEVC, tập hợp các tham số video (video parameter set - VPS) có thể được xác định là cấu trúc cú pháp chứa các thành phần cú pháp áp dụng cho không hoặc nhiều của toàn bộ chuỗi video mã hóa như được xác định bởi nội dung của thành phần cú pháp thấy được trong SPS được tham chiếu đến bởi thành phần cú pháp thấy được trong PPS được tham chiếu đến bởi thành phần cú pháp thấy được trong mỗi phần đầu đoạn của lát. Tập hợp các tham số video RBSP có

thể bao gồm các tham số có thể được tham chiếu đến bởi một hoặc nhiều tập hợp các tham số trình tự RBSP.

Mối tương quan và hệ thống thứ bậc giữa tập hợp các tham số video (VPS), tập hợp các tham số trình tự (SPS), và tập hợp các tham số hình ảnh (PPS) có thể được mô tả như sau. VPS tập trung vào một mức trên SPS trong hệ thống thứ bậc của tập hợp các tham số và trong ngữ cảnh khả năng mở rộng được và/hoặc video 3D. VPS có thể bao gồm các tham số phổ biến cho tất cả các lát theo tất cả các lớp (khả năng mở rộng được hoặc góc nhìn) trong toàn bộ chuỗi video mã hóa. SPS bao gồm các tham số phổ biến cho tất cả các lát trong lớp cụ thể (khả năng mở rộng được hoặc góc nhìn) trong toàn bộ chuỗi video mã hóa, và có thể được chia sẻ với nhiều lớp (khả năng mở rộng được hoặc góc nhìn). PPS bao gồm các tham số phổ biến cho tất cả các lát trong biểu diễn lớp cụ thể (biểu diễn của một lớp khả năng mở rộng được hoặc góc nhìn trong đơn vị truy cập) và có thể được chia sẻ bởi tất cả các lát trong nhiều biểu diễn lớp.

VPS có thể đưa ra thông tin về mối quan hệ phụ thuộc của các các lớp trong dòng bit, cũng như nhiều thông tin khác mà có thể áp dụng cho tất cả các lát theo tất cả các lớp (khả năng mở rộng được hoặc góc nhìn) trong toàn bộ chuỗi video mã hóa. VPS có thể được xem xét bao gồm hai phần, VPS cơ sở và phần mở rộng VPS, trong đó nội dung mở rộng VPS có thể có mặt theo cách tùy chọn. Theo HEVC, VPS cơ sở có thể được xem xét bao gồm cấu trúc cú pháp `video_parameter_set_rbsp()` không có cấu trúc cú pháp `vps_extension()`. Cấu trúc cú pháp `video_parameter_set_rbsp()` đã xác định ban đầu cho HEVC phiên bản 1 và bao gồm các thành phần cú pháp có thể sử dụng để giải mã lớp cơ sở. Theo HEVC, phần mở rộng VPS có thể được xem xét để bao gồm cấu trúc cú pháp `vps_extension()`. Cấu trúc cú pháp `vps_extension()` này được xác định Theo HEVC phiên bản 2 chủ yếu cho các phần mở rộng đa lớp và bao gồm các thành phần cú pháp có thể sử dụng để giải mã một hoặc nhiều lớp không phải lớp cơ sở, như các thành phần cú pháp chỉ báo các tương quan phụ thuộc vào lớp.

Thành phần cú pháp `max_tid_il_ref_pics_plus1` ở nội dung mở rộng VPS có thể được sử dụng để chỉ báo rằng các hình ảnh không phải IRAP không được

sử dụng làm tham chiếu để dự đoán liên lớp và, nếu không phải vậy, các lớp con tạm thời không được sử dụng dưới dạng tham chiếu để dự đoán liên lớp:

$\text{max_tid_il_ref_pics_plus1}[i][j]$ bằng 0 quy định rằng các hình ảnh không phải IRAP có nuh_layer_id bằng $\text{layer_id_in_nuh}[i]$ không được sử dụng làm các hình ảnh nguồn để dự đoán liên lớp đối với 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 rằng các hình ảnh có 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 các hình ảnh nguồn 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 trường hợp của các tập hợp các tham số, và mỗi trường hợp được nhận dạng với mã nhận dạng duy nhất. Để làm giới hạn việc sử dụng bộ nhớ cần thiết cho các tập hợp các tham số, miền giá trị cho các mã nhận dạng của tập hợp tham số bị giới hạn. Theo H.264/AVC và HEVC, mỗi phần đầu của lát bao gồm mã nhận dạng của tập hợp các tham số hình ảnh kích hoạt cho việc giải mã hình ảnh mà chứa lát, và mỗi tập hợp các tham số hình ảnh chứa mã nhận dạng của tập hợp các tham số trình tự kích hoạt. Do đó, việc truyền các tập hợp các tham số hình ảnh và tập hợp các tham số trình tự không cần được đồng bộ hóa chính xác với việc truyền các lát. Thay vào đó, đủ để các tập hợp các tham số trình tự và tập hợp tham số hình ảnh kích hoạt được thu được ở bất kỳ thời điểm nào trước khi chúng được tham chiếu, cho phép việc truyền các tập hợp các tham số “ngoài dải” sử dụng cơ chế truyền tin cậy so với các giao thức được sử dụng cho dữ liệu của lát. Ví dụ, các tập hợp các tham số có thể được bao gồm dưới dạng tham số trong mô tả phiên cho các phiên giao thức vận chuyển thời gian thực (RTP). Nếu tập hợp các tham số được truyền trong dải, chúng có thể được lặp lại để cải tiến sự vượt trội của lỗi.

Ngoài ra hoặc theo cách khác, việc truyền ngoài dải, việc báo hiệu hoặc lưu trữ có thể được sử dụng cho các mục đích khác ngoài dung sai đối với các lỗi truyền, như việc giám truy cập hoặc đàm phán phiên. Ví dụ, đầu vào mẫu của vết

(track) trong tệp thích hợp với ISOBMFF có thể bao gồm các tập hợp các tham số, trong khi dữ liệu được mã hóa trong dòng bit được lưu trữ ở nơi nào đó trong tệp tin này hoặc trong tệp tin khác. Cụm từ theo dòng bit (ví dụ chỉ báo theo dòng bit) có thể được sử dụng trong bộ yêu cầu bảo hộ và được mô tả các phương án để tham chiếu đến việc truyền ngoài dải, báo hiệu, hoặc lưu trữ theo cách trong đó dữ liệu ngoài dải được kết hợp với dòng bit. Cụm từ giải mã theo dòng bit hoặc tương tự có thể tham chiếu đến việc giải mã dữ liệu ngoài dải được tham chiếu (có thể thu được từ việc truyền ngoài dải, báo hiệu, hoặc lưu trữ) mà được kết hợp với dòng bit. Hình ảnh được mã hóa là biểu diễn hình ảnh được mã hóa.

Theo HEVC, hình ảnh được mã hóa có thể được xác định là biểu diễn hình ảnh được mã hóa chứa tất cả các đơn vị cây mã hóa của hình ảnh. Theo HEVC, đơn vị truy cập (AU) có thể được xác định là tập hợp các đơn vị NAL liên quan đến nhau theo quy tắc phân loại xác định, là liên tiếp theo thứ tự giải mã, và chứa nhiều nhất một hình ảnh với bất kỳ giá trị cụ thể nào của `nuh_layer_id`. Ngoài việc chứa các đơn vị VCL NAL của hình ảnh được mã hóa, đơn vị truy cập cũng có thể chứa các đơn vị NAL phi VCL.

Có thể được yêu cầu các hình ảnh mã hóa xuất hiện theo thứ tự nhất định trong đơn vị truy cập. Ví dụ hình ảnh được mã hóa với `nuh_layer_id` bằng `NuhLayerIdA` có thể được yêu cầu liên trước, theo thứ tự giải mã, tất cả các hình ảnh mã hóa với `nuh_layer_id` lớn hơn `nuhLayerIdA` trong cùng một đơn vị truy cập. AU thường chứa tất cả các hình ảnh mã hóa mà biểu diễn cùng thời gian kết xuất và/hoặc cùng thời gian chụp.

Dòng bit có thể được xác định là chuỗi các bit, ở dạng của dòng đơn vị NAL hoặc dòng byte, tạo ra biểu diễn của các hình ảnh mã hóa và dữ liệu được kết hợp tạo ra một hoặc nhiều chuỗi video mã hóa. Dòng bit thứ nhất có thể được theo sau bởi dòng bit thứ hai trong cùng kênh logic, như trong cùng tệp tin hoặc trong cùng kết nối của giao thức truyền thông. Dòng sơ cấp (trong ngữ cảnh mã hóa video) có thể được xác định là chuỗi của một hoặc nhiều dòng bit. Đầu của dòng bit thứ nhất có thể được chỉ báo bởi đơn vị NAL xác định, có thể được gọi là đơn vị NAL của đầu của dòng bit (EOB - end of bitstream) và là đơn vị NAL

cuối cùng của dòng bit. Theo HEVC và các bản nháp mở rộng hiện thời của nó, đơn vị NAL của EOB được yêu cầu để có `nuh_layer_id` bằng 0.

Định dạng dòng byte được xác định trong H.264/AVC và HEVC để truyền hoặc các môi trường lưu trữ không cung cấp các cấu trúc khung. Định dạng dòng byte phân tách các đơn vị NAL với nhau bằng cách gắn mã mở đầu phía trước mỗi đơn vị NAL. Để tránh phát hiện lỗi sai của các biên đơn vị NAL, các bộ mã hóa chạy thuật toán ngăn cản mô phỏng mã bắt đầu định hướng byte, thêm byte ngăn cản mô phỏng vào phần tải của đơn vị NAL nếu mã bắt đầu đã xuất hiện theo cách khác. Để, ví dụ, cho phép hoạt động công không phức tạp giữa hệ thống hướng gói và hướng dòng, sự ngăn cản mô phỏng mã bắt đầu có thể luôn được thực hiện bất kể định dạng dòng byte được sử dụng hay không. Thứ tự bit cho định dạng dòng byte có thể được xác định để bắt đầu với bit quan trọng nhất (most significant bit - MSB) của byte thứ nhất, tiếp theo đến bit ít quan trọng nhất (least significant bit - LSB) của byte thứ nhất, được theo sau bởi MSB của byte thứ hai, v.v.. Định dạng dòng byte có thể được xem xét bao gồm chuỗi các cấu trúc cú pháp đơn vị NAL dòng byte. Mỗi cấu trúc cú pháp đơn vị NAL dòng byte có thể được xem xét để chứa một tiền tố mã mở đầu được theo sau bởi một cấu trúc cú pháp đơn vị NAL, tức là cấu trúc cú pháp `nal_unit( NumBytesInNalUnit)` nếu các tên thành phần cú pháp được tham chiếu đến. Đơn vị NAL dòng byte cũng có thể chứa thành phần cú pháp `zero_byte` bổ sung. Cũng có thể chứa một hoặc nhiều thành phần cú pháp bổ sung `trailing_zero_8bits`. Khi đơn vị NAL dòng byte là đơn vị NAL dòng byte thứ nhất trong dòng bit, nó cũng có thể chứa một hoặc nhiều thành phần cú pháp `leading_zero_8bits` bổ sung. Cú pháp của đơn vị NAL dòng byte có thể được định nghĩa như sau:

<code>byte_stream_nal_unit(NumBytesInNalUnit) {</code>	Thông tin mô tả
<code>while(next_bits(24) != 0x000001 && next_bits(32) != 0x00000001)</code>	
<code>leading_zero_8bits /* equal to 0x00 */</code>	f(8)
<code>if(next_bits(24) != 0x000001)</code>	
<code>zero_byte /* equal to 0x00 */</code>	f(8)
<code>start_code_prefix_one_3bytes /* equal to 0x000001 */</code>	f(24)
<code>nal_unit(NumBytesInNalUnit)</code>	
<code>while(more_data_in_byte_stream() && next_bits(24) != 0x000001 && next_bits(32) != 0x00000001)</code>	
<code>trailing_zero_8bits /* equal to 0x00 */</code>	f(8)
<code>}</code>	

Thứ tự của các đơn vị NAL dòng byte trong dòng byte có thể được yêu cầu để tuân theo thứ tự giải mã của các đơn vị NAL được chứa trong các đơn vị NAL dòng byte. Các ngữ nghĩa của các thành phần cú pháp có thể được xác định như sau. `leading_zero_8bits` là byte bằng 0x00. Thành phần cú pháp `leading_zero_8bits` có thể chỉ có mặt trong đơn vị NAL dòng byte thứ nhất của dòng bit, vì bất kỳ các byte nào bằng 0x00 tuân theo cấu trúc cú pháp đơn vị NAL và liền trước chuỗi bốn byte 0x00000001 (được hiểu là `zero_byte` được theo sau bởi `start_code_prefix_one_3bytes`) sẽ được xem xét là các thành phần cú pháp `trailing_zero_8bits` là một phần của đơn vị NAL dòng byte liền trước. `zero_byte` là một byte bằng 0x00. `start_code_prefix_one_3bytes` là chuỗi giá trị không đổi của 3 byte bằng 0x000001. Thành phần cú pháp này có thể gọi là tiền tố mã mở đầu (hoặc đơn giản là mã mở đầu). `trailing_zero_8bits` là một byte bằng 0x00.

Đơn vị NAL có thể được xác định là cấu trúc cú pháp chứa chỉ báo của kiểu dữ liệu để tuân theo và các byte chứa dữ liệu trong dạng của RBSP được đặt khi cần thiết với các byte ngăn ngừa mô phỏng. Phần tải trình tự byte thô (RBSP) có thể được định nghĩa là cấu trúc cú pháp chứa số nguyên các byte được gói trong đơn vị NAL. RBSP là trống hoặc có dạng chuỗi dữ liệu bit chứa các thành phần cú pháp được theo sau bởi bit dừng RBSP và được theo sau bởi không hoặc nhiều hơn bit liên tiếp bằng 0.

Các đơn vị NAL bao gồm phần đầu và phần tải. Trong H.264/AVC và HEVC, phần đầu đơn vị NAL chỉ báo kiểu đơn vị NAL.

Cú pháp HEVC của cấu trúc cú pháp `nal_unit( NumBytesInNalUnit )` được đưa ra tiếp theo dưới dạng ví dụ của cú pháp của đơn vị NAL.

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

Theo HEVC, chuỗi video mã hóa (CVS) có thể được định nghĩa, ví dụ, là chuỗi của các đơn vị truy cập bao gồm, theo thứ tự giải mã, của đơn vị truy cập IRAP với `NoRaslOutputFlag` bằng 1, được theo sau bởi không hoặc nhiều đơn vị truy cập không phải là các đơn vị truy cập IRAP với `NoRaslOutputFlag` bằng 1, bao gồm tất cả các đơn vị truy cập tiếp theo lên đến nhưng không bao gồm bất kỳ đơn vị truy cập tiếp theo nào là đơn vị truy cập IRAP với `NoRaslOutputFlag` bằng 1. Đơn vị truy cập IRAP có thể được xác định là đơn vị truy cập trong đó hình ảnh lớp cơ sở là hình ảnh IRAP. Giá trị của `NoRaslOutputFlag` bằng 1 cho mỗi hình ảnh IDR, mỗi hình ảnh BLA, và mỗi hình ảnh IRAP trong đó là hình ảnh thứ nhất trong đó lớp cụ thể trong dòng bit theo thứ tự giải mã, là hình ảnh IRAP thứ nhất theo sau đầu của đơn vị NAL chuỗi có cùng giá trị của `nuh_layer_id` theo thứ tự giải mã. Theo HEVC đa lớp, giá trị của `NoRaslOutputFlag` bằng 1 cho mỗi hình ảnh IRAP khi `nuh_layer_id` của nó sao cho `LayerInitializedFlag[ nuh_layer_id ]` bằng 0 và `LayerInitializedFlag[ refLayerId ]` bằng 1 cho tất cả các giá trị của `refLayerId` bằng `IdDirectRefLayer[ nuh_layer_id ][ j ]`, trong đó `j` nằm trong miền từ 0 đến `NumDirectRefCac lop[ nuh_layer_id ] - 1`. Mặc khác, giá trị của `NoRaslOutputFlag` bằng `HandleCraAsBlaFlag`. `NoRaslOutputFlag` bằng 1 có ảnh hưởng trong đó các hình ảnh RASL được kết hợp với hình ảnh IRAP mà

trong đó NoRaslOutputFlag được thiết lập không được kết xuất bởi bộ giải mã. Có thể có phương tiện đưa ra giá trị của HandleCraAsBlaFlag cho bộ giải mã từ thực thể bên ngoài, như máy phát hoặc bộ thu, có thể được điều khiển bộ giải mã.

HandleCraAsBlaFlag có thể được thiết lập là 1 ví dụ bằng máy phát tìm kiếm vị trí mới trong dòng bit hoặc điều chỉnh thành phát quảng bá 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 được xử lý và được giải mã như thể nó là hình ảnh BLA.

Theo HEVC, chuỗi video mã hóa có thể ngoài ra hoặc theo cách khác (với mô tả ở trên) được quy định đối với đầu, khi đơn vị NAL xác định, có thể được gọi là đầu của đơn vị NAL chuỗi (end of sequence - EOS), xuất hiện trong dòng bit và có nuh_layer_id bằng 0.

Nhóm hình ảnh (group of pictures - GOP) và các đặc tính của nó có thể được định nghĩa như sau. GOP có thể được giải mã giải mã bất kể hình ảnh bất kỳ trước đó được giải mã hay không. GOP mở là nhóm hình ảnh trong đó các hình ảnh trước hình ảnh nội bộ ban đầu trong thứ tự đầu ra có thể không được giải mã một cách chính xác khi bước giải mã bắt đầu từ hình ảnh nội bộ ban đầu của GOP mở. Nói cách khác, các hình ảnh của GOP mở có thể tham chiếu (theo dự đoán liên ảnh) đến các hình ảnh thuộc về GOP trước. Bộ giải mã HEVC có thể nhận dạng hình ảnh nội bộ bắt đầu GOP mở, vì kiểu đơn vị NAL xác định, kiểu đơn vị NAL CRA, có thể được sử dụng cho các lát được mã hóa của nó. GOP đóng là nhóm các hình ảnh trong đó tất cả các hình ảnh có thể được giải mã chính xác khi giải mã mở đầu từ hình ảnh nội bộ ban đầu của GOP đóng. Nói cách khác, không có hình ảnh nào trong GOP đóng tham chiếu đến bất kỳ các hình ảnh nào trong các GOP trước. Trong H.264/AVC và HEVC, GOP đóng có thể bắt đầu từ hình ảnh IDR. Theo HEVC, GOP đóng cũng có thể bắt đầu từ hình ảnh BLA_W_RADL hoặc BLA_N_LP. Cấu trúc mã hóa GOP mở là tiềm năng hiệu quả hơn khi nén so với cấu trúc mã hóa GOP đóng, vì độ linh hoạt lớn hơn khi chọn hình ảnh tham chiếu.

Cấu trúc hình ảnh (SOP) có thể được xác định là một hoặc nhiều hình ảnh mã hóa liên tiếp theo thứ tự giải mã, trong đó hình ảnh được mã hóa thứ nhất theo thứ tự giải mã là hình ảnh tham chiếu tại lớp con theo thời gian thấp nhất và không có hình ảnh được mã hóa ngoại trừ hình ảnh được mã hóa thứ nhất theo thứ tự giải mã là hình ảnh RAP. Tất cả các hình ảnh trong SOP trước liền trước theo thứ tự giải mã tất cả các hình ảnh trong SOP hiện thời và tất cả các hình ảnh trong SOP kế tiếp nối tiếp theo thứ tự giải mã tất cả các hình ảnh trong SOP hiện thời. SOP có thể biểu diễn cấu trúc phân tầng và cấu trúc dự đoán liên ảnh lặp lại. Thuật ngữ nhóm hình ảnh (GOP) đôi khi có thể được sử dụng hoán đổi với thuật ngữ SOP và có cùng nghĩa với nghĩa của SOP.

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

Theo 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 là hợp lệ hoặc kích hoạt cho hình ảnh bao gồm tất cả các hình ảnh tham chiếu được sử dụng dưới dạng tham chiếu cho hình ảnh này và tất cả các hình ảnh tham chiếu được giữ đánh dấu là “được sử dụng để tham chiếu” cho bất kỳ các hình ảnh tiếp theo nào theo thứ tự giải mã. Có sáu tập con của tập hợp hình ảnh tham chiếu, được gọi là RefPicSetStCurr0 (còn được biến đến là RefPicSetStCurrBefore), RefPicSetStCurr1 (còn được biến đến là RefPicSetStCurrAfter), RefPicSetStFoll0, RefPicSetStFoll1, RefPicSetLtCurr, và RefPicSetLtFoll. RefPicSetStFoll0 và RefPicSetStFoll1 cũng có thể được xem xét để tạo ra đồng thời một tập con RefPicSetStFoll. Ký hiệu của sáu tập con như sau. “Curr” đề cập đến các hình ảnh tham chiếu được bao gồm trong các danh sách hình ảnh tham chiếu của hình ảnh hiện thời và do đó có thể được sử dụng dưới dạng tham chiếu dự đoán liên ảnh cho hình ảnh hiện thời. “Foll” đề cập đến các hình ảnh tham chiếu không được bao gồm trong các danh sách hình ảnh tham chiếu của hình ảnh hiện thời nhưng có thể được sử dụng trong các hình ảnh tiếp theo theo thứ tự giải

mã dưới dạng các hình ảnh tham chiếu. “St” đề cập đến các hình ảnh tham chiếu ngắn hạn, thường có thể được nhận dạng thông qua số lượng nhất định của các bit ít quan trọng nhất của giá trị POC của chúng. “Lt” đề cập đến các hình ảnh tham chiếu dài hạn, mà được xác định cụ thể và thường có sự khác biệt lớn hơn của các giá trị POC tương đối so với các hình ảnh hiện tại hơn so với cái được biểu diễn bằng cách đề cập đến số lượng nhất định của các bit ít quan trọng nhất. “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 giá trị của hình ảnh hiện tại. RefPicSetStCurr0, RefPicSetStCurr1, RefPicSetStFoll0 và RefPicSetStFoll1 cùng nhau được tham chiếu đến như là tập con ngắn hạn của tập hình ảnh tham chiếu. RefPicSetLtCurr và RefPicSetLtFoll cùng nhau được tham chiếu đến như là tập con dài hạn của tập hình ảnh tham chiếu.

Theo HEVC, tập hợp hình ảnh tham chiếu có thể được xác định trong tập hợp các tham số trình tự và được sử dụng trong phần đầu của lát thông qua chỉ số của tập hợp hình ảnh tham chiếu. Tập hợp hình ảnh tham chiếu cũng có thể được xác định trong phần đầu của lát. 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 đến là dự đoán liên RPS). Trong cả hai kiểu của việc mã hóa tập hợp hình ảnh tham chiếu, cờ (used_by_curr_pic_X_flag) còn được gửi cho mỗi hình ảnh tham chiếu chỉ ra liệu hình ảnh tham chiếu được sử dụng để tham chiếu bởi hình ảnh hiện tại (được bao gồm trong *Curr danh sách) hay không (chứa trong danh sách *Foll). Các hình ảnh chứa trong tập hình ảnh tham chiếu được sử dụng bởi lát hiện tại được tham chiếu như là “được sử dụng để tham chiếu”, và các hình ảnh không trong tập hình ảnh tham chiếu được sử dụng bởi lát hiện tại được đánh dấu như là “chưa được sử dụng để tham chiếu”. Nếu hình ảnh hiện tại là hình ảnh IDR, RefPicSetStCurr0, RefPicSetStCurr1, RefPicSetStFoll0, RefPicSetStFoll1, RefPicSetLtCurr, và RefPicSetLtFoll tất cả đều trống.

Đệm hình ảnh 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, để tham chiếu trong dự đoán liên ảnh và để tái thứ tự các hình ảnh được giải mã

thành thứ tự đầu ra. Khi H.264/AVC và HEVC đề xuất tính linh hoạt cho cả bước đánh dấu hình ảnh tham chiếu và tái thứ tự đầu ra, các đệm riêng biệt để đệm hình ảnh tham chiếu và đệm hình ảnh xuất ra có thể làm lãng phí tài nguyên bộ nhớ. Vì vậy, DPB có thể bao gồm quy trình đệm hình ảnh được giải mã thống nhất đối với các hình ảnh tham chiếu và tái thứ tự đầu ra. Hình ảnh được giải mã có thể được loại bỏ khỏi DPB khi không còn được sử dụng để tham chiếu và không cần thiết cho đầu ra.

Trong nhiều chế độ mã hóa theo H.264/AVC và HEVC, hình ảnh tham chiếu để dự đoán liên ảnh được chỉ ra cùng với chỉ số của danh sách hình ảnh tham chiếu. Chỉ số có thể được mã hóa bởi bước mã hóa chiều dài khả biến, thường khiến chỉ số nhỏ hơn có giá trị ngắn hơn với thành phần cú pháp tương ứng. Trong H.264/AVC và HEVC, hai danh sách hình ảnh tham chiếu (danh sách hình ảnh tham chiếu 0 và danh sách hình ảnh tham chiếu 1) được tạo ra cho mỗi lát dự đoán đối (B), và một danh sách hình ảnh tham chiếu (danh sách hình ảnh tham chiếu 0) được tạo ra đối với mỗi lát (P) được mã hóa liên ảnh.

Danh sách hình ảnh tham chiếu, như danh sách hình ảnh tham chiếu 0 và danh sách hình ảnh tham chiếu 1, có được xây dựng theo hai bước: Đầu tiên, danh sách hình ảnh tham chiếu ban đầu được tạo ra. Danh sách 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 hệ thống thứ bậc dự đoán như cấu trúc GOP, hoặc bất kỳ kết hợp nào của chúng. Thứ hai, danh sách hình ảnh tham chiếu ban đầu có thể được tái thứ tự với việc tái thứ tự bằng các lệnh tái thứ tự danh sách hình ảnh tham chiếu (RPLR), cũng được biết là cấu trúc cú pháp biến đổi của danh sách hình ảnh tham chiếu, có thể được chứa trong các phần đầu của lát. Nếu các tập hợp hình ảnh tham chiếu được sử dụng, đầu tiên danh sách hình ảnh tham chiếu 0 có thể được khởi tạo để chứa danh sách `RefPicSetStCurr0`, được theo sau bởi `RefPicSetStCurr1`, được theo sau bởi `RefPicSetLtCurr`. Danh sách hình ảnh tham chiếu 1 có thể được khởi tạo để chứa `RefPicSetStCurr1` trước tiên, được theo sau bởi `RefPicSetStCurr0`. Theo HEVC, các danh sách hình ảnh tham chiếu ban đầu có thể được biến đổi thông qua cấu trúc cú pháp biến đổi của

danh sách hình ảnh tham chiếu, trong đó các hình ảnh trong các danh sách hình ảnh tham chiếu ban đầu có thể được nhận dạng thông qua chỉ số điểm vào của danh sách. Nói cách khác, theo HEVC, biến đổi danh sách hình ảnh tham chiếu được mã hóa thành cấu trúc cú pháp bao gồm vòng lặp trên mỗi điểm vào trong danh sách hình ảnh tham chiếu cuối cùng, trong đó mỗi điểm vào vòng lặp là chỉ số mã hóa độ dài cố định vào danh sách hình ảnh tham chiếu ban đầu và chỉ báo hình ảnh theo thứ tự vị trí tăng dần trong danh sách 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 thu chỉ số hình ảnh tham chiếu vào danh sách hình ảnh tham chiếu, có thể được sử dụng để chỉ báo hình ảnh nào trong số các hình ảnh tham chiếu được sử dụng để dự đoán liên ảnh cho khối cụ thể. Chỉ số hình ảnh tham chiếu có thể được mã hóa bởi bộ mã hóa thành dòng bit là một số chế độ mã hóa liên ảnh hoặc nó có thể thu được (bằng bộ mã hóa và bộ giải mã) ví dụ sử dụng các khối lân cận trong một số chế độ mã hóa liên kết khác.

Để biểu diễn các vector chuyển động hiệu quả trong các dòng bit, các vector chuyển động có thể được mã hóa khác nhau tương ứng với vector chuyển động được dự đoán dành riêng cho khối. Trong nhiều bộ mã hóa-giải mã video, các vector chuyển động được dự đoán được tạo ra theo cách định trước, ví dụ bằng việc tính số trung bình của các vector chuyển động được mã hóa hoặc giải mã của các khối liền kề. Cách khác để tạo ra các dự đoán vector chuyển động, đôi khi được gọi là dự đoán vector chuyển động cải tiến (AMVP), để tạo ra danh sách các dự đoán ứng viên từ các khối liền kề và/hoặc các khối đồng vị trí trong các hình ảnh tham chiếu tạm thời và báo hiệu ứng viên được chọn dưới dạng bộ dự đoán vector chuyển động. Ngoài việc dự đoán các giá trị vector chuyển động, chỉ số tham chiếu của hình ảnh được mã hóa/giải mã trước đó có thể được dự đoán. Chỉ số tham chiếu thường được dự đoán từ các khối liền kề và/hoặc các khối đồng vị trí trong hình ảnh tham chiếu theo thời gian. Việc mã hóa khác nhau của các vector chuyển động thường không hoạt động qua các biên lát.

Mã hóa video mở rộng được có thể đề cập đến cấu trúc mã hóa trong đó một dòng bit có thể chứa nhiều biểu diễn nội dung, ví dụ, ở các tỷ lệ bit, các độ phân giải hoặc các tốc độ khung khác nhau. Trong trường hợp này bộ thu có thể trích xuất biểu diễn mong muốn phụ thuộc và các đặc tính của nó (ví dụ độ phân giải phù hợp nhất với độ phân giải của màn hình hiển thị của thiết bị). Theo cách khác, máy chủ hoặc thành phần mạng có thể trích xuất các phần của dòng bit được truyền đến bộ thu phụ thuộc vào đặc điểm mạng hoặc khả năng giải mã chỉ có các phần nhất định của dòng bit mở rộng được. Dòng bit mở rộng được thường bao gồm “lớp cơ sở” tạo ra video chất lượng thấp nhất khả dụng và một hoặc nhiều lớp tăng cường để tăng cường chất lượng video khi thu được và được giải mã cùng với các lớp thấp hơn. Để cải tiến hiệu quả mã hóa cho các lớp tăng cường, biểu diễn được mã hóa của lớp đó cơ thể phụ thuộc và các lớp thấp hơn. Ví dụ, thông tin chuyển động và chế độ của lớp tăng cường có thể được dự đoán từ các lớp thấp hơn. Tương tự, dữ liệu điểm ảnh của các lớp thấp hơn có thể được sử dụng để tạo ra dự đoán cho các lớp tăng cường.

Trong một số mô hình mã hóa video mở rộng được, tín hiệu video có thể được mã hóa thành lớp cơ sở và một hoặc nhiều lớp tăng cường. Lớp tăng cường có thể tăng cường, ví dụ, độ phân giải tạm thời (tức là tốc độ khung), độ phân giải không gian, hoặc đơn giản là chất lượng của nội dung video được biểu diễn bởi lớp khác hoặc một phần của nó. Mỗi lớp cùng với tất cả các lớp phụ thuộc của nó là một biểu diễn của tín hiệu video, ví dụ, tại độ phân giải không gian nhất định, độ phân giải thời gian, và mức chất lượng. Theo tài liệu này, lớp mở rộng được cùng với tất cả các lớp phụ thuộc của nó được tham chiếu đến như là “biểu diễn lớp mở rộng được”. Một phần của dòng bit khả năng mở rộng được tương ứng với biểu diễn của lớp mở rộng được có thể được trích xuất và được giải mã để tạo ra biểu diễn của tín hiệu gốc với độ chính xác nhất định.

Các chế độ khả năng mở rộng được hoặc các kích thước khả năng mở rộng được có thể bao gồm nhưng không giới hạn ở nội dung sau đây:

- Khả năng mở rộng được chất lượng: Các hình ảnh lớp cơ sở được mã hóa với chất lượng thấp hơn các hình ảnh lớp tăng cường, có thể đạt

được ví dụ sử dụng giá trị tham số lượng tử hóa lớn hơn (tức là, kích thước bước lượng tử hóa lớn hơn cho việc lượng tử hóa hệ số biến đổi) trong lớp cơ sở so với trong lớp tăng cường.

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

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

- Khả năng mở rộng được vùng động học: Các lớp mở rộng được biểu diễn các vùng và/hoặc hình ảnh động học khác thu được sử dụng hàm ánh xạ tông màu khác và/hoặc hàm chuyển quang học khác.

- Khả năng mở rộng được định dạng chroma: Các hình ảnh lớp cơ sở tạo ra độ phân giải không gian thấp hơn trong các mảng mẫu màu (ví dụ, được mã hóa với định dạng chroma 4:2:0) so với các hình ảnh lớp tăng cường (ví dụ định dạng 4:4:4).

- Khả năng mở rộng được dải màu: các hình ảnh lớp tăng cường có vùng biểu diễn màu đa dạng hơn/rộng hơn các hình ảnh lớp cơ sở - ví dụ lớp tăng cường có thể có dải màu UHDTV (ITU-R BT.2020) và lớp cơ sở có thể có dải màu ITU-R BT.709.

- Khả năng mở rộng được góc nhìn, cũng có thể được gọi là mã hóa đa góc nhìn. Lớp cơ sở biểu diễn góc nhìn thứ nhất, trong đó lớp tăng cường biểu diễn góc nhìn thứ hai.

- Khả năng mở rộng được độ sâu, cũng có thể được gọi là mã hóa tăng cường độ sâu. Một lớp hoặc một số lớp của dòng bit có thể biểu diễn (các) góc nhìn kết cấu, trong khi lớp hoặc các lớp khác có thể biểu diễn (các) góc nhìn độ sâu.

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

- Khả năng mở rộng được xen kẽ đến lũy tiến (cũng được biết là khả năng mở rộng được trường đến khung): nội dung nguồn được mã hóa xen kẽ của lớp cơ sở được tăng cường với lớp tăng cường để biểu diễn nội dung nguồn lũy tiến.

- Khả năng mở rộng được bộ mã hóa-giải mã lai hóa (cũng được biết là khả năng mở rộng được tiêu chuẩn mã hóa): Trong khả năng mở rộng được bộ mã hóa-giải mã lai hóa, cú pháp dòng bit, các ngữ nghĩa và quy trình giải mã của lớp cơ sở và lớp tăng cường được xác định trong các tiêu chuẩn mã hóa video khác. Do đó, các hình ảnh lớp cơ sở được mã hóa theo tiêu chuẩn hoặc định dạng mã hóa khác 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 phần mở rộng đa lớp HEVC.

Cần hiểu rằng các kiểu khả năng mở rộng được có thể được kết hợp và được áp dụng cùng nhau. Ví dụ khả năng mở rộng được dải màu và khả năng mở rộng được độ sâu bit có thể được kết hợp.

Thuật ngữ lớp có thể được sử dụng trong ngữ cảnh của bất kỳ kiểu khả năng mở rộng được nào, bao gồm khả năng mở rộng được góc nhìn và tăng cường độ sâu. Lớp tăng cường có thể đề cập đến bất kỳ kiểu tăng cường nào, như tăng cường SNR, không gian, đa góc nhìn, độ sâu, độ sâu bit, chroma, và/hoặc dải màu. Lớp cơ sở có thể đề cập đến bất kỳ kiểu chuỗi video cơ sở nào, như góc nhìn cơ sở, lớp cơ sở cho khả năng mở rộng được SNR/không gian, hoặc góc nhìn cơ sở kết cấu cho việc mã hóa video được tăng cường theo độ sâu.

Các công nghệ khác nhau để cung cấp nội dung video ba chiều (3D đang được nghiên cứu và được phát triển. Có thể xem xét rằng trong video lập thể hoặc video hai góc nhìn, một chuỗi video hoặc góc nhìn được biểu diễn cho mắt trái trong khi góc nhìn song song được biểu diễn cho mắt phải. Nhiều hơn hai góc nhìn song song có thể cần thiết cho các ứng dụng cho phép chuyển điểm nhìn hoặc các màn hình lập thể tự động có thể thể hiện số lượng lớn các hình ảnh một cách đồng thời và để người xem quan sát nội dung từ các góc nhìn khác.

Góc nhìn có thể được xác định là chuỗi các hình ảnh biểu diễn một camera hoặc điểm nhìn. Các hình ảnh biểu diễn góc nhìn cũng có thể được gọi là các thành phần góc nhìn. Nói cách khác, thành phần góc nhìn có thể được xác định là biểu diễn góc nhìn được mã hóa trong một đơn vị truy cập. Trong mã hóa video đa góc nhìn, nhiều hơn một góc nhìn được mã hóa trong dòng bit. Vì các góc nhìn thường được dự định được hiển thị trên màn hình lập thể hoặc màn hình lập thể tự động đa góc nhìn hoặc được sử dụng cho các thiết bị 3D khác, chúng thường biểu diễn trên cùng cảnh và chồng lấp một phần theo nội dung mặc dù biểu diễn các điểm nhìn khác đối với nội dung đó. Do đó, dự đoán liên góc nhìn có thể được sử dụng trong bước mã hóa video đa góc nhìn để tận dụng tương quan liên góc nhìn và tăng cường hiệu quả nén. Một cách để nhận biết dự đoán liên góc nhìn là đưa vào một hoặc nhiều hình ảnh được giải mã của một hoặc nhiều góc nhìn khác trong (các) danh sách hình ảnh tham chiếu của hình ảnh được mã hóa hoặc giải mã nằm trong góc nhìn thứ nhất. Khả năng mở rộng được góc nhìn có thể đề cập đến mã hóa video đa góc nhìn hoặc các dòng bit video đa góc nhìn, mà cho phép việc loại bỏ hoặc bỏ qua một hoặc nhiều các góc nhìn được mã hóa, trong khi dòng bit thu được duy trì tương thích với và biểu diễn video với số lượng các góc nhìn nhỏ hơn ban đầu. Mã hóa vùng quan tâm (Region of Interest - ROI) có thể được định nghĩa là đề cập việc mã hóa miền cụ thể trong video với độ chính xác cao hơn.

Khả năng mở rộng được ROI có thể được xác định là kiểu khả năng mở rộng được trong đó lớp tăng cường tăng cường chỉ một phần của hình ảnh lớp tham chiếu ví dụ theo không gian, chất lượng, độ sâu bit, và/hoặc theo các kích thước khả năng mở rộng được khác. Khi khả năng mở rộng được ROI có thể được sử dụng cùng với các kiểu khả năng mở rộng được khác, có thể xem xét để tạo ra hạng mục khác của các kiểu khả năng mở rộng được. Tồn tại các ứng dụng khác đối với việc mã hóa ROI với các yêu cầu khác, có thể được nhận biết bằng cách sử dụng khả năng mở rộng được ROI. Ví dụ, lớp tăng cường có thể được truyền để tăng cường chất lượng và/hoặc độ phân giải của vùng trong lớp cơ sở. Bộ giải mã mà nhận dòng bit tăng cường và dòng bit lớp cơ sở có thể giải mã các lớp và phủ các hình ảnh được giải mã ở trên nhau và hiển thị hình ảnh cuối cùng.

Sự tương ứng không gian của hình ảnh lớp tham chiếu và hình ảnh lớp tăng cường có thể được suy ra hoặc có thể được chỉ báo với một hoặc nhiều kiểu của các khoảng dịch vị trí lớp tham chiếu. Theo HEVC, các khoảng dịch vị trí lớp tham chiếu có thể được bao gồm trong PPS bởi bộ mã hóa và được giải mã từ PPS bởi bộ giải mã. Các khoảng dịch vị trí lớp tham chiếu có thể được sử dụng cho nhưng không giới hạn ở khả năng mở rộng được ROI thu được. Các khoảng dịch vị trí lớp tham chiếu có thể được chỉ báo giữa hai lớp hoặc các hình ảnh của hai lớp ngay cả khi các lớp này không có tương quan dự đoán liên lớp với nhau. Các khoảng dịch vị trí lớp tham chiếu có thể bao gồm một hoặc nhiều khoảng dịch lớp tham chiếu định tỉ lệ, các khoảng dịch vùng tham chiếu, và các tập hợp pha lấy mẫu lại. Các khoảng dịch lớp tham chiếu định tỉ lệ có thể được xem xét để xác định các khoảng dịch chiều ngang và chiều dọc giữa mẫu trong hình ảnh hiện thời được xếp cùng vị trí với mẫu luma bên trái phía trên của vùng tham chiếu trong hình ảnh giải mã trong lớp tham chiếu và các khoảng dịch chiều ngang và chiều dọc giữa mẫu trong hình ảnh hiện thời được xếp cùng vị trí với mẫu luma bên phải phía dưới của vùng tham chiếu trong hình ảnh được giải mã trong lớp tham chiếu. Một cách khác là để xem xét các khoảng dịch lớp tham chiếu định tỉ lệ để xác định các vị trí của các mẫu góc của vùng tham chiếu được lấy mẫu lên liên quan đến các mẫu góc tương ứng của hình ảnh lớp tăng cường. Các giá trị khoảng dịch lớp tham chiếu định tỉ lệ có thể được đánh dấu và thường được cho phép bằng 0. Các khoảng dịch vùng tham chiếu có thể được xem xét để xác định các khoảng dịch chiều ngang và chiều dọc giữa mẫu luma bên trái phía trên của vùng tham chiếu trong hình ảnh được giải mã trong lớp tham chiếu và mẫu luma bên trái phía trên của cùng một hình ảnh được giải mã cũng như các khoảng dịch chiều ngang và chiều dọc giữa mẫu luma bên phải phía dưới của vùng tham chiếu trong hình ảnh được giải mã trong lớp tham chiếu và mẫu luma bên phải phía dưới của cùng một hình ảnh giải mã. Các giá trị khoảng dịch vùng tham chiếu có thể được đánh dấu và thường được cho phép bằng 0. Tập hợp pha lấy mẫu lại có thể được xem xét để xác định các khoảng dịch pha được sử dụng trong quy trình lấy mẫu lại của hình ảnh nguồn để dự đoán liên lớp. Các khoảng dịch pha khác có thể được cung cấp đối với các thành phần luma và chroma.

Sự đóng gói khung có thể được định nghĩa để bao gồm bước bố trí nhiều hơn một hình ảnh đầu vào, có thể được gọi là các khung thành phần (đầu vào), trong hình ảnh đầu ra. Nói chung, đóng gói khung không bị giới hạn với bất kỳ kiểu cụ thể nào của các khung thành phần hoặc các khung thành phần không cần tương quan cụ thể nào với nhau. Trong nhiều trường hợp, đóng gói khung được sử dụng cho việc bố trí các khung thành phần của video clip lập thể trong một chuỗi hình ảnh, như được giải thích chi tiết hơn trong đoạn tiếp theo. Việc bố trí có thể bao gồm việc đặt các hình ảnh đầu vào trong không gian các vùng không xếp chồng trong hình ảnh đầu ra. Ví dụ, trong việc bố trí cạnh nhau, hai hình ảnh đầu vào được đặt trong hình ảnh đầu ra liền kề với nhau theo hướng chiều ngang. Việc bố trí này cũng có thể bao gồm phân vùng một hoặc nhiều hình ảnh đầu vào trong hai hoặc nhiều hơn hai phân vùng khung thành phần và đặt các phần khung thành phần trong các vùng không xếp chồng theo không gian trong hình ảnh đầu ra. Hình ảnh đầu ra hoặc chuỗi hình ảnh đầu ra được gói trong khung có thể được mã hóa thành dòng bit ví dụ bởi bộ mã hóa video. Dòng bit có thể được giải mã ví dụ bởi bộ giải mã video. Bộ giải mã hoặc hoạt động sau xử lý sau khi giải mã có thể trích xuất các khung thành phần được giải mã từ (các) hình ảnh được giải mã ví dụ để hiển thị.

Trong video lập thể tương thích với khung (còn được biến đến là đóng gói khung của video lập thể), việc đóng gói không gian của cặp hình ảnh lập thể thành một khung được thực hiện ở phía bộ mã hóa dưới dạng bước xử lý trước để mã hóa và sau đó các khung đóng gói khung được mã hóa với mô hình video mã hóa 2D thông thường. Các khung đầu ra được tạo ra bởi bộ giải mã chứa các khung thành phần của cặp hình ảnh lập thể.

Trong chế độ hoạt động thông thường, độ phân giải không gian của các khung ban đầu của mỗi góc nhìn và khung được gói có cùng độ phân giải. Trong trường hợp này bộ mã hóa lấy mẫu xuống hai góc nhìn của video lập thể trước hoạt động đóng gói. Việc đóng gói trong không gian có thể sử dụng ví dụ định dạng cạnh-cạnh hoặc trên-dưới, và việc lấy mẫu xuống cần được thực hiện theo đó.

Đóng gói khung có thể được ưu tiên so với mã hóa video đa góc nhìn (ví dụ phần mở rộng H.264/AVC hoặc MV-HEVC của H.265/HEVC) ví dụ vì các lý do sau:

- Quy trình hậu sản xuất có thể được điều chỉnh cho một tín hiệu video. Một số công cụ sau sản xuất không thể xử lý hai chuỗi hình ảnh riêng biệt và/hoặc không thể giữ các chuỗi hình ảnh riêng biệt đồng bộ với nhau.
- Hệ thống phân phối, như các giao thức truyền, có thể hỗ trợ chỉ một chuỗi mã hóa và/hoặc không thể giữ các chuỗi mã hóa tách biệt đồng bộ với nhau và/hoặc có thể yêu cầu nhiều đệm hoặc độ trễ để giữ các chuỗi mã hóa riêng biệt đồng bộ với nhau.
- Việc giải mã các dòng bit với các công cụ mã hóa video đa góc nhìn có thể yêu cầu sự hỗ trợ của các chế độ mã hóa riêng biệt, mà có thể không thể khả dụng trong các máy phát. Ví dụ, nhiều điện thoại thông minh hỗ trợ profile chính H.265/HEVC nhưng không thể xử lý giải mã profile chính H.265/HEVC Multiview mặc dù nó chỉ yêu cầu các bổ sung cấp độ cao so với profile chính.

Phép chiếu cụ thể để ánh xạ ảnh toàn cảnh bao phủ trường nhìn 360 độ theo hướng chiều ngang và trường nhìn 180 độ theo hướng chiều dọc (do đó biểu diễn hình cầu) đến mặt phẳng ảnh hai chiều hình chữ nhật được biết đến là phép chiếu đẳng chữ nhật (equirectangular). Trong trường hợp này, tọa độ ngang có thể được xem xét tương đương với kinh độ, và tọa độ dọc có thể được xem xét tương đương với vĩ độ, không có biến đổi hoặc định tỷ lệ được áp dụng. Trong một số trường hợp, nội dung toàn cảnh với trường nhìn chiều ngang 360 độ nhưng với trường nhìn chiều dọc nhỏ hơn 180 độ có thể được xem xét các trường hợp đặc biệt của phép chiếu hình cầu, trong đó các vùng cực của hình cầu không được ánh xạ trên mặt phẳng ảnh hai chiều. Trong một số trường hợp tính toàn cảnh có thể có trường nhìn chiều ngang nhỏ hơn 360 độ và lên đến trường nhìn chiều dọc 180 độ, trong khi nếu không thì có các đặc tính của định dạng phép đẳng chữ nhật.

Trong định dạng phép chiếu ánh xạ lập phương, video cần được chiếu trên sáu mặt (còn được biến đến là các mặt) của hình lập phương. Ánh xạ lập phương có thể được tạo ra ví dụ trước tiên bằng cách biểu diễn cảnh mặt cầu sáu lần từ một điểm nhìn, với các góc nhìn được xác định bởi khối cắt góc nhìn 90 độ biểu diễn mỗi mặt lập phương. Các mặt lập phương có thể được gói khung trong cùng một khung hoặc mỗi mặt lập phương có thể được xử lý riêng biệt (ví dụ trong bước mã hóa). Có các thứ tự khả dĩ để đặt các mặt lập phương trên khung và/hoặc các mặt lập phương có thể được quay hoặc được phản chiếu. Khung chiều rộng và chiều cao để đóng gói khung có thể được chọn để khớp các mặt lập phương “chặt” ví dụ tại lưới mặt lập phương 3x2, hoặc có thể bao gồm các khung thành phần không được sử dụng ví dụ với lưới mặt lập phương 4x3.

Công cụ hoặc chế độ mã hóa sao chép nội khối (IBC) tương tự với dự đoán liên ảnh nhưng sử dụng hình ảnh hiện thời được mã hóa hoặc giải mã làm hình ảnh tham chiếu. Hiển nhiên, chỉ có các khối được mã hóa hoặc giải mã trước khi khối hiện thời được mã hóa hoặc giải mã có thể được sử dụng dưới dạng các tham chiếu để dự đoán. Phần mở rộng mã hóa nội dung màn hình (SCC) của HEVC được dự định bao gồm IBC.

IBC Theo HEVC SCC nghĩa là chỉ cho mẫu dự đoán. Rõ ràng là không được phép trong bản đặc tả HEVC SCC để sử dụng hình ảnh hiện thời dưới dạng hình ảnh đồng vị trí cho TMVP.

Quy trình dự đoán liên ảnh có thể liên quan đến việc tham chiếu đến các vị trí mẫu bên ngoài các biên hình ảnh ít nhất cho (nhưng không nhất thiết bị giới hạn với) hai lý do sau đây. Thứ nhất, các vectơ chuyển động có thể chỉ đến các khối dự đoán bên ngoài các biên hình ảnh, và thứ hai, các vectơ chuyển động có thể trỏ đến vị trí mẫu không phải số nguyên mà trong đó giá trị mẫu được nội suy sử dụng việc lọc lấy các mẫu đầu vào từ các vị trí bên ngoài các biên hình ảnh.

Trong các tiêu chuẩn mã hóa video, như H.263, H.264/AVC, và H.265/HEVC, các vectơ chuyển động được cho phép để trỏ đến vùng bên ngoài các biên hình ảnh để thu khối dự đoán và nội suy mẫu phân đoạn có thể sử dụng các vị trí mẫu bên ngoài các biên hình ảnh. Để thu các mẫu bên ngoài các biên

hình ảnh là một phần quy trình dự đoán, các mẫu tương ứng của các biên hình ảnh có thể được sao chép hiệu quả. Điều này được minh họa trên Fig.5a. Trên Fig.5a và Fig.5b hình chữ nhật nét đứt minh họa mảng mẫu hình ảnh tham chiếu, và các hình vuông với các đường liền nét minh họa các mẫu của mảng mẫu. Trong các hình vuông, ký hiệu phía trên chỉ báo tọa độ x của vị trí của mẫu quan tâm và ký hiệu phía dưới chỉ báo tọa độ y của vị trí mẫu quan tâm.

Fig.5a minh họa ví dụ xử lý tham chiếu đến các mẫu bên ngoài các biên hình ảnh trong quy trình dự đoán liên ảnh trong đó các mẫu biên được sao chép tới vùng bên ngoài biên của hình ảnh.

Cơ chế để hỗ trợ các vị trí mẫu bên ngoài các biên hình ảnh trong quy trình dự đoán liên ảnh có thể được thực hiện theo nhiều cách. Một cách là cấp phát mảng mẫu mà lớn hơn kích thước hình ảnh giải mã, tức là có các mép ở bên trên của, bên dưới, ở bên phải, và ở bên trái của ảnh này. Ngoài ra hoặc thay vì sử dụng các mép này, vị trí của mẫu được sử dụng để dự đoán (dưới dạng đầu vào để phép nội suy mẫu phân đoạn cho khối dự đoán hoặc dưới dạng mẫu trong chính khối dự đoán) có thể được làm bão hòa để vị trí không vượt quá các biên hình ảnh (với các mép, nếu được sử dụng). Một vài trong số các tiêu chuẩn mã hóa video mô tả việc hỗ trợ các vector chuyển động trên các biên hình ảnh theo cách này.

Ví dụ, phương trình sau đây được sử dụng theo HEVC để thu các vị trí ($x_{Ai,j}$, $y_{Ai,j}$) bên trong mảng đã cho $refPicLXL$ của các mẫu luma (của hình ảnh tham chiếu) cho phép nội suy mẫu phân đoạn để tạo ra khối dự đoán (trung gian):

$$x_{Ai,j} = \text{Clip3}(0, \text{pic_width_in_luma_samples} - 1, x_{IntL} + i)$$

$$y_{Ai,j} = \text{Clip3}(0, \text{pic_height_in_luma_samples} - 1, y_{IntL} + j)$$

trong đó (x_{IntL} , y_{IntL}) là vị trí luma trong các đơn vị mẫu toàn phần; i và j các vị trí tương đối trong cửa sổ lọc nội suy; $\text{pic_width_in_luma_samples}$ là chiều rộng của hình ảnh trong các mẫu luma; $\text{pic_height_in_luma_samples}$ là chiều cao của hình ảnh trong các mẫu luma; và $\text{Clip3}()$ được xác định như sau:

$$\text{Clip3}(x, y, z) = \begin{cases} x & ; \quad z < x \\ y & ; \quad z > y \\ z & ; \quad \text{khác} \end{cases}$$

Có thể thấy rằng nếu vị trí mẫu sẽ có tọa độ âm (trước hoạt động Clip3), nó được làm bão hòa đến 0 bằng hoạt động Clip3. Tương tự, nếu vị trí mẫu sẽ có tọa độ lớn hơn chiều rộng (đối với vị trí chiều ngang) hoặc chiều cao (đối với vị trí chiều dọc), chiều rộng được làm bão hòa là -1 hoặc chiều cao -1 , tương ứng.

Các phương trình tương tự cũng có sẵn để làm bão hòa các vị trí mẫu chroma ở trong các biên mảng mẫu chroma trong phép nội suy mẫu phân đoạn để dự đoán liên ảnh của các thành phần chroma.

Trong video toàn cảnh 360 độ các giá trị mẫu từ phía đối diện của hình ảnh có thể được sử dụng thay vì sử dụng mẫu biên khi mẫu theo hướng chiều ngang bên ngoài biên của hình ảnh cần thiết trong quy trình dự đoán. Điều này được minh họa trên Fig.5b thể hiện việc xử lý các vectơ chuyển động trên các biên hình ảnh trong bước mã hóa video toàn cảnh, theo một phương án.

Việc xử lý các điểm ảnh bên ngoài các biên hình ảnh có thể được xử lý bằng cách mở rộng hình ảnh tham chiếu lớn hơn (theo chiều rộng và/hoặc chiều cao) so với hình ảnh được mã hóa. Các triển khai khác như được mô tả sau đây cũng khả dĩ.

Theo ví dụ của Fig.5b có thể thấy rằng vị trí mẫu chiều ngang có thể được gộp đối với các mẫu được tham chiếu trong quy trình dự đoán liên ảnh. Ví dụ, cột liền kề với bên trái của biên hình ảnh tham chiếu ($x=-1$) có thể tham chiếu đến các giá trị mẫu của cột ngoài cùng bên phải của hình ảnh tham chiếu ($x=w-1$), cột tiếp theo với bên trái ($x=-2$) có thể tham chiếu đến các giá trị mẫu của cột ngoài cùng bên phải thứ hai của hình ảnh tham chiếu ($x=w-2$), v.v.. Nói cách khác, thay vì làm bão hòa các vị trí chiều ngang âm đến 0 và các vị trí mẫu chiều ngang lớn hơn chiều rộng -1 đến chiều rộng -1 (tức là vị trí mẫu chiều ngang của cột mẫu ngoài cùng bên phải của hình ảnh), các vị trí mẫu chiều ngang được tham chiếu bên ngoài các biên hình ảnh có thể được gộp, tức là, các vị trí mẫu chiều ngang lớn hơn chiều rộng -1 được gộp để chúng để tham chiếu các cột mẫu ở bên trái

của hình ảnh. Ngược lại, các vị trí mẫu chiều ngang nhỏ hơn 0 được gói để chúng đề cập đến cột mẫu ở bên phải của hình ảnh.

Do đó, hàm $\text{Wrap}()$ có thể được xác định như sau:

$$\text{Wrap}(x, y, z) = \begin{cases} y - (x - z) + 1 & ; \quad z < x \\ x + (z - y) - 1 & ; \quad z > y \\ z & ; \quad \text{khác} \end{cases}$$

Khi tham chiếu đến các phương trình trên cho nội suy mẫu luma phân đoạn theo HEVC, phương trình sau đây có thể được sử dụng

$$xA_{i,j} = \text{Wrap}(0, \text{pic_width_in_luma_samples} - 1, xInt_L + i)$$

thay cho phương trình:

$$xA_{i,j} = \text{Clip3}(0, \text{pic_width_in_luma_samples} - 1, xInt_L + i)$$

Vùng hình ảnh hiệu dụng có thể được xác định là vùng hình ảnh hoặc tập hợp các mẫu hoặc các vị trí mẫu trong hình ảnh được giải mã chứa được giải mã các mẫu hữu ích để hiển thị. Các vùng bên ngoài vùng hình ảnh hiệu dụng không thể hữu ích để hiển thị nhưng có thể có mặt ví dụ để tạo chiều rộng của hình ảnh hoặc chiều cao để phù hợp với lưới khối được sử dụng trong mã hóa.

Đóng gói khung có thể kém hơn để mã hóa video đa góc nhìn về mặt hiệu quả nén (còn được biến đến là hiệu quả biến dạng tỉ lệ) do, ví dụ, các lý do sau đây. Trong việc đóng gói khung, dự đoán mẫu liên góc nhìn và dự đoán chuyển động liên góc nhìn không được cho phép giữa các góc nhìn. Hơn nữa, trong việc đóng gói khung, các vector chuyển động trở ra bên ngoài các biên của khung thành phần (đến khung thành phần khác) hoặc khiến phép nội suy điểm ảnh phụ sử dụng các mẫu bên ngoài các biên của khung thành phần (trong khung thành phần khác) có thể được xử lý ở dưới điều kiện tốt nhất. Trong bước mã hóa video đa góc nhìn thông thường, các vị trí mẫu được sử dụng trong dự đoán liên ảnh và nội suy điểm ảnh phụ có thể được làm bão hòa trong các biên hình ảnh hoặc các vùng tương đương bên ngoài biên của hình ảnh trong các hình ảnh khôi phục có thể được đệm với các giá trị mẫu biên.

Fig.6a minh họa vấn đề liên quan đến các vectơ chuyển động được xử lý tối ưu. Theo ví dụ này, khối ngoài cùng bên phải 604 của hàng khối thứ hai trong khung thành phần #1 601 của khung #N+1 được dịch chuyển từ vùng một phần bên ngoài vùng của khung thành phần #1 trong khung #N, được minh họa với khối 603 trên Fig.6a. Mũi tên 605 minh họa ví dụ của dự đoán liên ảnh có thể được tối ưu hoặc, trong trường hợp của các dòng bit có khả năng giải mã song song, không thể được cho phép. Trong bước mã hóa không được gói khung, các vùng bên ngoài các biên hình ảnh có thể được xử lý trong dự đoán liên ảnh bằng cách đệm, như được mô tả ở trên, tương đương với làm bão hòa các vị trí mẫu được sử dụng trong dự đoán liên ảnh hoặc phép nội suy mẫu phân đoạn ở trong các biên hình ảnh. Trong bước mã hóa video được gói khung, trong việc đệm hoặc việc bão hòa vị trí mẫu không khả dụng và do đó hiệu quả nén gần các biên giữa các khung thành phần có thể kém hơn hiệu quả nén trong bước mã hóa video đa góc nhìn không được gói khung.

Các phương án được mô tả sau đây để mở rộng biên của một hoặc nhiều khung thành phần hoặc để làm bão hòa các vị trí mẫu được sử dụng trong quy trình dự đoán liên ảnh trong các biên của khung thành phần. Các cụm từ sau đây như mở rộng biên của khung thành phần hoặc đệm có thể đề cập đến ngoại suy tín hiệu của vùng hình ảnh hiệu dụng của khung thành phần, ví dụ bằng cách sao chép điểm ảnh biên, hoặc bằng cách sao chép tín hiệu của vùng hình ảnh hiệu dụng biểu diễn nội dung toàn cảnh 360 độ từ phía đối diện của hình ảnh hiệu dụng, như được mô tả chi tiết hơn sau đây.

Fig.6b thể hiện ví dụ của phần mở rộng biên, theo một phương án. Trên Fig.6b, khung thành phần #1 601 được mở rộng với một cột khối 606 sang bên phải, và, tương ứng, khung thành phần #2 602 được mở rộng với một cột khối 607 sang trái (ví dụ cột CTU theo HEVC), mặc dù trong trường hợp tổng quát hơn chỉ có tập con của các khung thành phần có thể có phần mở rộng biên. Các mẫu trong phần mở rộng biên 606, 607 có thể ví dụ chứa giá trị mẫu biên của vùng hình ảnh hiệu dụng của khung thành phần tương ứng. Mũi tên 608 minh họa

ví dụ về dự đoán liên ảnh với vectơ chuyển động trên vùng hình ảnh hiệu dụng của khung thành phần #1 601.

Các phương án được thể hiện sau đây với một số phương pháp để mã hóa và/hoặc giải mã vùng mở rộng biên.

Bộ mã hóa có thể thu một hoặc nhiều hình ảnh để mã hóa (khối 71 trong lưu đồ của Fig.7a). Theo một phương án, bộ mã hóa có thể quyết định (khối 72) để bố trí (khối 73) các hình ảnh không nén trước khi mã hóa (khối 74) để đưa vào (các) hàng khối và/hoặc (các) cột khối được dự định cho các phần mở rộng biên hoặc bao gồm (các) hàng khối và/hoặc (các) cột khối cho các phần mở rộng biên trong các hình ảnh mã hóa trong lúc mã hóa (khối 75). Theo một phương án, phần mở rộng biên được tạo ra trong khối 75 bằng dự đoán nội ảnh, dự đoán sao chép nội khối, hoặc dự đoán liên ảnh, như được giải thích chi tiết hơn sau đây. Theo phương án khác, phần mở rộng biên được tạo ra trong khối 75 bằng cách đệm khung tham chiếu, như được giải thích chi tiết hơn sau đây.

Một cách tương ứng, bộ giải mã có thể nhận một hoặc nhiều hình ảnh được mã hóa để giải mã (khối 81 trong lưu đồ của Fig.7b). Theo một phương án, bộ giải mã giải mã các hình ảnh để thu (các) khung được khôi phục (khối 82) và thực hiện đệm khung để làm đầy trong các vùng bên ngoài vùng hình ảnh hiệu dụng để tạo ra hình ảnh tham chiếu được đệm, nếu cần (khối 83).

Theo một phương án, cột khối mở rộng khung thành phần sang bên phải (ví dụ phần mở rộng biên 606 của khung thành phần #1 601 trên Fig.6b) được mã hóa với dự đoán nội ảnh theo chiều ngang. Theo một phương án, bộ mã hóa có thể chọn để mã hóa phần mở rộng biên không mã hóa lỗi dự đoán. Theo phương án khác, bộ mã hóa luôn mã hóa phần mở rộng biên mà không mã hóa lỗi dự đoán.

Theo một phương án, hàng khối mở rộng khung thành phần xuống dưới (ví dụ phần mở rộng biên của khung thành phần #1 theo cách bố trí đóng gói khung trên-dưới) được mã hóa với dự đoán nội ảnh theo chiều dọc. Bộ mã hóa có thể chọn để mã hóa phần mở rộng biên không mã hóa lỗi dự đoán. Theo phương án khác, bộ mã hóa luôn mã hóa phần mở rộng biên mà không mã hóa lỗi dự đoán.

Theo một phương án, bộ mã hóa chỉ báo trong dòng bit các khối mã hóa lỗi dự đoán không có mặt hoặc chứa lỗi dự đoán có giá trị không. Theo một phương án, bộ mã hóa suy luận rằng không có lỗi dự đoán được bao gồm cho phần mở rộng biên và bỏ qua mã hóa các thành phần cú pháp liên quan đến lỗi dự đoán bất kỳ cho phần mở rộng biên trong dòng bit. Theo một phương án, bộ giải mã kết luận các khối nào thuộc về phần mở rộng biên và bỏ qua việc giải mã các thành phần cú pháp liên quan đến lỗi dự đoán bất kỳ cho phần mở rộng biên từ dòng bit.

Theo một phương án, có thể được áp dụng cùng với hoặc độc lập với các phương án khác, khung thành phần thứ nhất biểu diễn góc nhìn thứ nhất và khung thành phần thứ hai biểu diễn góc nhìn thứ hai của cùng nội dung đa góc nhìn (ví dụ lập thể) được gói khung. Dự đoán sao chép nội khối từ khung thành phần thứ nhất đến khung thành phần thứ hai được cho phép trong bước mã hóa, nhưng dự đoán sao chép nội khối từ khung thành phần thứ hai đến khung thành phần thứ nhất không được cho phép trong bước mã hóa. Do đó, dự đoán hình ảnh liên kết về cơ bản đạt được bằng cách áp dụng dự đoán sao chép nội khối giữa các khung thành phần của các góc nhìn khác. Hiệu quả nén do đó được tăng cường so với việc mã hóa nội dung đa góc nhìn được gói khung mà không dự đoán sao chép nội khối.

Theo một phương án, được chỉ báo (ví dụ bởi bộ mã hóa) dự đoán sao chép nội khối được áp dụng chỉ trong khung thành phần thứ hai theo cách bị giới hạn như sau. Đối với mỗi khối của khung thành phần thứ hai, ở đây được gọi là khối được mã hóa/giải mã, việc mã hóa bị giới hạn như sau. Sau đây, cho khối đồng vị trí là khối trong khung thành phần thứ nhất được xếp cùng vị trí với khối được mã hóa/giải mã, trong đó đồng vị trí có thể được xác định là đặc điểm có cùng tọa độ ngang và tọa độ dọc tương đối tương ứng với góc bên trái phía trên của khung thành phần chứa, hoặc có cùng các tọa độ kết hợp với khoảng dịch chiều dọc và/hoặc chiều ngang ví dụ trong các đơn vị của các đơn vị cây mã hóa. (Các) khoảng dịch có thể được chỉ báo bởi bộ mã hóa trong dòng bit và/hoặc được giải mã từ dòng bit bởi bộ giải mã. Không có các giá trị mẫu của các khối nối tiếp,

theo thứ tự quét khối, khối đồng vị trí trong khung thành phần thứ nhất, được sử dụng để dự đoán sao chép nội khối.

Theo một phương án, bộ mã hóa biết về, ví dụ thông qua giao diện lập trình ứng dụng (API) hoặc tương tự, sự chênh lệch chiều dọc và/hoặc chiều ngang tối đa giữa các khung thành phần hoặc bộ mã hóa phân tích nội dung nguồn để phát hiện sự chênh lệch chiều dọc và/hoặc chiều ngang tối đa giữa các khung thành phần hoặc bộ mã hóa tạo giả định về sự chênh lệch chiều dọc và/hoặc chiều ngang tối đa giữa các khung thành phần này. Bộ mã hóa giới hạn miền tìm kiếm dự đoán sao chép khối liên kết theo sự chênh lệch chiều dọc và/hoặc chiều ngang tối đa. Ví dụ, sự chênh lệch theo chiều chiều dọc tối đa có thể được giả sử là 0, và miền tìm kiếm do đó bị giới hạn là cùng hàng với khối hiện thời.

Theo một phương án, có thể được sử dụng cùng với hoặc độc lập với các phương án trên mô tả việc sử dụng dự đoán sao chép nội khối để đạt được dự đoán hình ảnh liên kết về nội dung đa góc nhìn được gói khung, cột khối mở rộng khung thành phần sang bên phải (ví dụ phần mở rộng biên 606 của khung thành phần #1 601 trên Fig.6b) được mã hóa với dự đoán sao chép nội khối từ cột khối ngoài cùng bên trái của khung thành phần. Bộ mã hóa có thể xác định để mã hóa phần mở rộng biên cột khối không mã hóa lỗi dự đoán. Chế độ hoạt động này thích hợp với khung thành phần là toàn cảnh 360 độ (ví dụ toàn cảnh đẳng chữ nhật), vì nhờ áp dụng với phần mở rộng biên bằng dự đoán sao chép nội khối, tín hiệu này giữ liên tục vượt biên hiệu dụng của hình ảnh. Phương án này có thể được sử dụng cũng trong trường hợp trong đó có một khung thành phần trong hình ảnh, tức là với mã hóa toàn cảnh với việc mã hóa toàn cảnh monoscopic.

Theo một phương án, cột khối mở rộng khung thành phần sang trái (ví dụ phần mở rộng biên 607 của khung thành phần #2 602 trên Fig.6b) được mã hóa với dự đoán sao chép nội khối từ khung thành phần khác trong hình ảnh. Theo một phương án, hàng khối mở rộng khung thành phần lên trên (ví dụ phần mở rộng biên 607 của khung thành phần #2 602 theo cách bố trí đóng gói khung trên-dưới) được mã hóa với dự đoán sao chép nội khối từ một khung thành phần khác trong hình ảnh. Hai phương án này có thể được xem là áp dụng dự đoán liên góc

nhìn để làm đầy trong các vùng biên, có thể hiệu quả khi sự chênh lệch hoặc khoảng cách của các mẫu được sao chép bởi dự đoán IBC thành phần mở rộng biên là ổn định và tương ứng với sự chênh lệch hoặc khoảng cách các mẫu biên của vùng hình ảnh hiệu dụng.

Theo một phương án, khối trong phần mở rộng biên được dự đoán liên ảnh. Điều này có thể hiệu quả khi sự dịch chuyển sao cho (các) hình ảnh trước đó, theo thứ tự giải mã, chứa các mẫu thích hợp cho phần mở rộng biên, ví dụ khi camera chuyển động quay theo cách thích hợp hoặc khi camera và nội dung hình ảnh là tĩnh và do đó phần mở rộng biên của hình ảnh trước đó cũng đóng vai trò là phần mở rộng biên cho hình ảnh hiện thời.

Theo một phương án, bộ mã hóa chỉ báo, ví dụ trong tin nhắn PPS, SPS, VPS, SEI hoặc VUI, vùng hình ảnh hiệu dụng của khung thành phần. Theo một phương án, bộ giải mã giải mã chỉ báo về vùng hình ảnh hiệu dụng của khung thành phần. Theo một phương án, bộ giải mã xén đi các hình ảnh giải mã để loại trừ các vùng bên ngoài vùng hình ảnh hiệu dụng dưới dạng một phần của quy trình xén đầu ra được áp dụng cho các hình ảnh được giải mã trước khi chúng được xuất ra bởi bộ giải mã. Theo một phương án, bộ giải mã xuất ra chỉ báo được giải mã của vùng hình ảnh hiệu dụng của khung thành phần. Theo một phương án, bộ giải mã hoặc quy trình hiển thị chỉ sử dụng vùng hình ảnh hiệu dụng của khung thành phần trong quy trình hiển thị, loại trừ (các) phần mở rộng biên từ quy trình hiển thị.

Theo một phương án, bộ tạo tệp tin hoặc tương tự chỉ báo vùng hình ảnh hiệu dụng của khung thành phần trong tệp tin chứa. Theo một phương án, bộ phân tích tệp tin hoặc tương tự giải mã chỉ báo vùng hình ảnh hiệu dụng của khung thành phần từ tệp tin chứa. Theo một phương án, bộ phân tích tệp tin hoặc tương tự xén các hình ảnh được giải mã để loại trừ các vùng bên ngoài vùng hình ảnh hiệu dụng trước khi hiển thị các hình ảnh được giải mã. Theo một phương án, quy trình hiển thị chỉ sử dụng vùng hình ảnh hiệu dụng của khung thành phần trong quy trình hiển thị, loại trừ (các) phần mở rộng biên từ quy trình hiển thị.

Theo phương án làm ví dụ khác, các khối mở rộng biên được mã hóa và khôi phục trước tiên thành khung tham chiếu trung gian thông thường nhưng có thể chứa bất kỳ các giá trị mẫu nào. Tương tự, các khối mở rộng biên được giải mã theo cách thông thường thành khung tham chiếu trung gian trước tiên. Sau đó phần mở rộng biên của khung tham chiếu trung gian được biến đổi. Ví dụ, giá trị mẫu được khôi phục của mẫu biên của vùng hình ảnh hiệu dụng có thể được sử dụng để làm đầy trong các mẫu trong cùng hàng mẫu trong phần mở rộng biên. Nếu khung thành phần biểu diễn hình ảnh toàn cảnh 360 độ, các mẫu được khôi phục ở phía đối diện của vùng hình ảnh hiệu dụng có thể được sử dụng để làm đầy trong các mẫu trong phần mở rộng biên. Ví dụ, các mẫu trong phần mở rộng biên có thể được làm đầy bằng cách sao chép các mẫu được giải mã từ phía đối diện của vùng hình ảnh hiệu dụng trên cơ sở hàng mẫu: mẫu thứ nhất bên ngoài vùng hình ảnh hiệu dụng (trong phần mở rộng biên) được sao chép từ mẫu thứ nhất bên trong vùng hình ảnh hiệu dụng ở phía đối diện, mẫu thứ hai bên ngoài vùng hình ảnh hiệu dụng được sao chép từ mẫu thứ hai bên trong vùng hình ảnh hiệu dụng ở phía đối diện, và tương tự. Vùng nguồn để sao chép nêu trên có thể được gọi là vùng biên. Ví dụ, khi phần mở rộng biên ở bên phải của vùng hình ảnh hiệu dụng, cột mẫu ngoài cùng bên trái của phần mở rộng biên có thể được làm đầy bằng cách sao chép các mẫu được giải mã từ cột mẫu ngoài cùng bên trái của vùng hình ảnh hiệu dụng, cột mẫu thứ hai từ bên trái trong phần mở rộng biên có thể được làm đầy bằng cách sao chép các mẫu được giải mã từ cột mẫu thứ hai từ bên trái trong vùng hình ảnh hiệu dụng, và tương tự.

Theo một phương án liên quan đến việc đệm khung tham chiếu trong vòng lặp, bộ mã hóa chỉ báo, ví dụ trong tập hợp các tham số như PPS hoặc SPS, vùng hình ảnh hiệu dụng của khung thành phần. Theo một phương án, bộ giải mã giải mã chỉ báo và kết luận các phần mở rộng biên theo đó.

Theo một phương án, bộ mã hóa chỉ báo, ví dụ trên cơ sở phần mở rộng biên hoặc khung thành phần ví dụ trong PPS hoặc SPS, phương pháp nào được sử dụng để làm đầy trong phần mở rộng biên của khung tham chiếu trung gian. Theo một phương án, bộ giải mã giải mã chỉ báo và sử dụng phương pháp được chỉ báo

để làm đầy trong phần mở rộng biên của khung tham chiếu trung gian. Theo một phương án, bộ giải mã kết luận phương pháp làm đầy phần mở rộng biên của khung tham chiếu trung gian, ví dụ sử dụng bước sao chép từ phía đối diện của vùng hình ảnh hiệu dụng khi khung thành phần này được chỉ báo là hình ảnh toàn cảnh 360 độ, và nếu không thì sử dụng việc sao chép mẫu biên.

Theo một phương án áp dụng được khi (các) hình ảnh được mã hóa bao gồm ít nhất khung thành phần thứ nhất và khung thành phần thứ hai, bộ mã hóa chọn hoặc được tạo cấu hình để sử dụng tám phân vùng sao cho tập hợp tám thứ nhất bao phủ khung thành phần thứ nhất và toàn bộ (các) phần mở rộng biên của nó và không bao gồm dữ liệu từ khung thành phần thứ hai, và tập hợp tám thứ hai bao phủ khung thành phần thứ hai và toàn bộ (các) phần mở rộng biên của nó và không bao gồm dữ liệu từ khung thành phần thứ nhất.

Theo một phương án, bộ mã hóa mã hóa tập hợp tám thứ nhất dưới dạng tập hợp tám bị ràng buộc chuyển động và tập hợp tám thứ hai dưới dạng tập hợp tám khác bị giới hạn chuyển động. Bộ mã hóa có thể chỉ báo rằng việc mã hóa các tập hợp tám thứ nhất và thứ hai được thực hiện theo cách chuyển động bị giới hạn ví dụ với tin nhắn SEI, như tin nhắn SEI của các tập hợp tám bị giới hạn chuyển động theo thời gian của HEVC.

Theo một phương án, bộ giải mã, bộ phát, hoặc tương tự giải mã chỉ báo ví dụ từ tin nhắn SEI, như tin nhắn SEI của các tập hợp tám bị giới hạn chuyển động theo thời gian của HEVC, việc mã hóa các tập hợp tám thứ nhất và thứ hai được thực hiện theo cách chuyển động bị giới hạn. Theo một phương án, bộ giải mã, bộ phát, hoặc tương tự xác định trên cơ sở của chỉ báo được giải mã để giải mã tập hợp tám thứ nhất với thực thể bộ giải mã thứ nhất (ví dụ quy trình giải mã thứ nhất) và tập hợp tám thứ hai với thực thể bộ giải mã thứ hai (ví dụ quy trình giải mã thứ hai).

Theo một phương án, bộ mã hóa mã hóa tập hợp tám thứ nhất dưới dạng tập hợp tám bị ràng buộc chuyển động. Theo một phương án, bộ mã hóa mã hóa tập hợp tám thứ hai dưới dạng tập hợp tám bị ràng buộc chuyển động ngoại trừ khi tham chiếu đến hình ảnh hiện thời trong việc dự đoán sao chép nội khối hoặc

tương tự (như được mô tả theo một phương án khác). Theo một phương án, bộ mã hóa chỉ báo trong dòng bit, ví dụ trong tin nhắn SEI, các tập hợp tằm bị ràng buộc chuyển động được sử dụng trong bước mã hóa như được mô tả trong đoạn này.

Theo một phương án, bộ giải mã, bộ phát, hoặc tương tự giải mã chỉ báo ví dụ từ tin nhắn SEI, như tin nhắn SEI của các tập hợp tằm tạm thời bị giới hạn chuyển động của HEVC, việc mã hóa các tập hợp tằm thứ nhất và thứ hai được thực hiện theo cách chuyển động bị giới hạn ngoại trừ khi tham chiếu đến hình ảnh hiện thời trong dự đoán sao chép nội khối hoặc tương tự (như được mô tả theo một phương án khác). Theo một phương án, bộ giải mã, bộ phát, hoặc tương tự xác định trên cơ sở chỉ báo được giải mã để giải mã tập hợp tằm thứ nhất với thực thể bộ giải mã thứ nhất (ví dụ quy trình giải mã thứ nhất) và tập hợp tằm thứ hai với thực thể bộ giải mã thứ hai (ví dụ quy trình giải mã thứ hai). Thực thể bộ giải mã thứ nhất được tạo cấu hình để giải mã tập hợp tằm thứ nhất trước khi thực thể bộ giải mã thứ hai giải mã tập hợp tằm thứ hai của cùng một hình ảnh. Tuy nhiên, thực thể bộ giải mã thứ nhất có thể giải mã tập hợp tằm thứ nhất của hình ảnh kế tiếp (theo sau hình ảnh hiện thời) theo thứ tự giải mã, trong khi thực thể bộ giải mã thứ hai giải mã tập hợp tằm thứ hai của hình ảnh hiện thời.

Theo một phương án làm ví dụ liên quan đến làm bão hòa hoặc gói các vị trí mẫu trên biên của khung thành phần, các khung thành phần có thể được gói “nối tiếp nhau”, tức là không có các vùng mở rộng biên ở giữa chúng. Quy trình này để thu nhận các mẫu để dự đoán liên ảnh có thể được biến đổi theo cách trong đó các vị trí mẫu bên ngoài vùng khung thành phần không được sử dụng.

Bộ mã hóa chỉ báo, ví dụ trong tập hợp các tham số như PPS hoặc SPS, vùng hình ảnh của khung thành phần. Theo một phương án, bộ giải mã giải mã chỉ báo. Sau đây, giả sử rằng góc bên trái phía trên của khung thành phần có các tọa độ mẫu luma leftCF, topCF và góc bên phải phía dưới có các tọa độ mẫu luma rightCF, bottomCF.

Theo một phương án, bộ mã hóa chỉ báo, ví dụ, trên cơ sở của khung thành phần ví dụ trong PPS hoặc SPS, liệu các vị trí mẫu bên ngoài các biên khung thành phần có được làm bão hòa hay được gói. Theo một phương án, bộ giải mã

giải mã chỉ báo và sử dụng phương pháp được chỉ báo trong dự đoán liên ảnh. Theo một phương án, bộ giải mã kết luận phương pháp, ví dụ sử dụng gói khi khung thành phần được chỉ báo là hình ảnh toàn cảnh 360 độ, và nếu không thì sử dụng việc làm bão hòa.

Bộ mã hóa và/hoặc bộ giải mã có thể sử dụng phương trình sau đây để gói các vị trí mẫu:

$$xA_{i,j} = \text{Wrap}(\text{leftCF}, \text{rightCF}, xInt_L + i)$$

Bộ mã hóa và/hoặc bộ giải mã có thể sử dụng phương trình sau đây để làm bão hòa các vị trí mẫu ở trong các biên khung thành phần:

$$xA_{i,j} = \text{Clip3}(\text{leftCF}, \text{rightCF}, xInt_L + i)$$

Các phương trình hoặc việc xử lý tương tự có thể áp dụng cho hướng chiều dọc.

Sau đây, chi tiết hơn về việc đệm khung tham chiếu trong vòng lặp trong bộ mã hóa và bộ giải mã sẽ được mô tả, theo một phương án.

Fig.4 thể hiện sơ đồ khối của bộ mã hóa video thích hợp để sử dụng các phương án của sáng chế. Fig.4 thể hiện bộ mã hóa cho hai lớp, nhưng cần hiểu rằng bộ mã hóa được thể hiện cũng được làm 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 phần bộ mã hóa thứ nhất 500 đối với lớp cơ sở và phần bộ mã hóa thứ hai 502 cho lớp tăng cường. Mỗi trong số phần bộ mã hóa thứ nhất 500 và phần bộ mã hóa thứ hai 502 có thể bao gồm các thành phần tương tự để mã hóa các hình ảnh mới đến. Các phần bộ mã hóa 500, 502 có thể bao gồm bộ 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 dự đoán điểm ảnh 302, 402 là bao gồm bộ dự đoán liên ảnh 306, 406, bộ dự đoán nội ảnh 308, 408, bộ chọn chế độ 310, 410, bộ lọc 316, 416, và bộ nhớ khung tham chiếu 318, 418. Bộ dự đoán điểm ảnh 302 của phần bộ mã hóa thứ nhất 500 thu 300 ảnh lớp cơ sở của dòng video cần mã hóa ở cả bộ dự đoán liên ảnh 306 (mà xác định sự khác nhau giữa ảnh và khung tham chiếu bù chuyển

động 318) và bộ dự đoán nội ảnh 308 (mà xác định dự đoán cho khối ảnh chỉ dựa trên các phần được xử lý sẵn của khung hoặc hình ảnh hiện thời). Đầu ra của cả bộ dự đoán liên ảnh và bộ dự đoán nội ảnh được truyền đến bộ chọn chế độ 310. Bộ dự đoán nội ảnh 308 có thể có nhiều hơn một chế độ dự đoán nội ảnh. Do đó, mỗi chế độ có thể thực hiện dự đoán nội ảnh và cung cấp tín hiệu được dự đoán cho bộ chọn chế độ 310. Bộ chọn chế độ 310 cũng thu bản sao của hình ảnh lớp cơ sở 300. Tương ứng, bộ dự đoán điểm ảnh 402 của phần bộ mã hóa thứ hai 502 thu 400 ảnh lớp tăng cường của dòng video được mã hóa ở cả bộ dự đoán liên ảnh 406 (mà xác định sự khác nhau giữa ảnh và khung tham chiếu bù chuyển động 418) và bộ dự đoán nội ảnh 408 (xác định dự đoán cho khối ảnh chỉ dựa trên các phần được xử lý sẵn của khung hoặc hình ảnh hiện thời). Đầu ra của cả bộ dự đoán liên ảnh và bộ dự đoán nội ảnh được truyền đến bộ chọn chế độ 410. Bộ dự đoán nội ảnh 408 có thể có nhiều hơn một chế độ dự đoán nội ảnh. Do đó, mỗi chế độ có thể thực hiện dự đoán nội ảnh và cung cấp tín hiệu được dự đoán đến bộ chọn chế độ 410. Bộ chọn chế độ 410 cũng thu nhận bản sao của hình ảnh lớp tăng cường 400.

Phụ thuộc vào chế độ mã hóa nào được chọn để mã hóa khối hiện thời, đầu ra của bộ dự đoán liên ảnh 306, 406 hoặc đầu ra của một trong số các chế độ bộ dự đoán nội ảnh tùy chọn hoặc đầu ra của bộ mã hóa bề mặt trong bộ chọn chế độ được truyền đến đầu ra của bộ chọn chế độ 310, 410. Đầu ra của bộ chọn chế độ được truyền đến bộ cộng thứ nhất 321, 421. Bộ cộng thứ nhất có thể trừ đi đầu ra của bộ 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 được nhập vào bộ mã hóa lỗi dự đoán 303, 403.

Bộ dự đoán điểm ảnh 302, 402 còn nhận từ bộ khôi phục sơ bộ 339, 439 kết hợp của biểu diễn dự đoán của khối ảnh 312, 412 và đầu ra 338, 438 của bộ giải mã lỗi dự đoán 304, 404. Ảnh được khôi phục sơ bộ 314, 414 có thể được truyền đến bộ dự đoán nội ảnh 308, 408 và đến đơn vị đệm khung tham chiếu 315, 415. Đơn vị đệm khung tham chiếu 315, 415 có thể làm đầy trong các vùng bên ngoài vùng hình ảnh hiệu dụng theo một hoặc nhiều phương án khác để tạo ra

ảnh tham chiếu được đệm. Ảnh tham chiếu được đệm này có thể được truyền đến bộ lọc 316, 416. Bộ lọc 316, 416 nhận ảnh tham chiếu được đệm có thể lọc ảnh tham chiếu được đệm và xuất ra ảnh được khôi phục cuối cùng 340, 440 có thể được lưu trong bộ nhớ khung tham chiếu 318, 418. Bộ nhớ khung tham chiếu 318 có thể được kết nối với bộ dự đoán liên ảnh 306 được sử dụng dưới dạng ảnh tham chiếu trong đó hình ảnh lớp cơ sở tương lai 300 được so sánh trong các hoạt động dự đoán liên ảnh. Tùy theo lớp cơ sở được chọn và được chỉ báo là nguồn đối với dự đoán mẫu liên lớp và/hoặc dự đoán thông tin chuyển động liên lớp của lớp tăng cường theo một số phương án, bộ nhớ khung tham chiếu 318 cũng có thể được kết nối với bộ dự đoán liên ảnh 406 được sử dụng dưới dạng ảnh tham chiếu trong đó các hình ảnh lớp tăng cường trong tương lai 400 được so sánh trong các hoạt động dự đoán liên ảnh. Ngoài ra, bộ nhớ khung tham chiếu 418 có thể được kết nối với bộ dự đoán liên ảnh 406 được sử dụng làm ảnh tham chiếu trong đó hình ảnh lớp tăng cường trong tương lai 400 được so sánh trong các hoạt động dự đoán liên ảnh.

Cần hiểu rằng đơn vị đệm khung tham chiếu 315, 415 có thể hoạt động trên các đơn vị khác nhau, có thể được xác định trước ví dụ trong tiêu chuẩn mã hóa hoặc có thể được chỉ báo bởi bộ mã hóa trong dòng bit. Theo một phương án, đơn vị đệm khung tham chiếu 315, 415 sử dụng toàn bộ ảnh được khôi phục sơ bộ 314, 414 làm đầu vào. Theo một phương án khác, đơn vị đệm khung tham chiếu 315, 415 sử dụng khung thành phần được khôi phục sơ bộ làm đầu vào, và do đó có thể được viện dẫn nhiều hơn một lần trên mỗi ảnh được khôi phục sơ bộ (phụ thuộc vào số lượng các khung thành phần mà nó bao quanh). Khi gọi đơn vị đệm khung tham chiếu 315, 415 với khung thành phần được khôi phục sơ bộ làm đầu vào, biên mở rộng của khung thành phần có thể có lợi khi sử dụng dự đoán sao chép nội khối từ khung thành phần này đến khung thành phần khác của cùng một hình ảnh. Vẫn theo phương án khác, đơn vị đệm khung tham chiếu 315, 415 sử dụng khối được khôi phục sơ bộ, hàng mẫu biên của khối được khôi phục sơ bộ (khi mở rộng biên theo hướng chiều dọc), hoặc cột mẫu biên của khối được khôi phục sơ bộ (khi mở rộng biên theo hướng chiều ngang) làm đầu vào. Phương án

này có thể có lợi ví dụ khi sử dụng dự đoán sao chép nội khối từ khung thành phần này đến một khung thành phần khác của cùng một hình ảnh.

Cần hiểu rằng theo các phương án khác bộ dự đoán điểm ảnh có thể được bố trí để thực hiện một chuỗi bước lọc và đệm theo các cách khác nhau. Ví dụ, bộ dự đoán điểm ảnh có thể thực hiện đệm và lọc khung tham chiếu theo thứ tự ngược lại, tức là lọc trước khi đệm khung tham chiếu. Theo một ví dụ khác, bộ dự đoán điểm ảnh có thể thực hiện lọc trong hai giai đoạn, như bước lọc khoảng dịch thích ứng mẫu và bước lọc giải khối, trong đó giai đoạn thứ nhất diễn ra trước khi đệm khung tham chiếu và giai đoạn thứ hai diễn ra sau khi đệm khung tham chiếu. Cần lưu ý rằng đó là các ví dụ không giới hạn và bộ dự đoán điểm ảnh có thể thực hiện lọc theo cách khác.

Bước lọc các tham số từ bộ lọc 316 của phần bộ mã hóa thứ nhất 500 có thể được cấp đến phần bộ mã hóa thứ hai 502 tùy theo lớp cơ sở được chọn và được chỉ báo là nguồn để dự đoán các tham số lọc của lớp tăng cường theo một số phương án.

Bộ mã hóa lỗi dự đoán 303, 403 bao gồm đơn vị biến đổi 342, 442 và bộ lượng tử hóa 344, 444. Đơn vị biến đổi 342, 442 biến đổi tín hiệu lỗi dự đoán thứ nhất 320, 420 sang miền biến đổi. Sự biến đổi là, ví dụ, 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 lượng tử hóa.

Bộ giải mã lỗi dự đoán 304, 404 nhận đầu ra từ bộ mã hóa lỗi dự đoán 303, 403 và thực hiện các quy trình ngược của bộ mã hóa lỗi dự đoán 303, 403 để tạo ra lỗi dự đoán tín hiệu được giải mã 338, 438, khi được kết hợp với biểu diễn dự đoán của khối ảnh 312, 412 ở bộ cộng thứ hai 339, 439, tạo ra ảnh được khôi phục sơ bộ 314, 414. Bộ giải mã lỗi dự đoán có thể được xem xét để bao gồm bộ giải lượng tử hóa 361, 461, giải lượng tử hóa các giá trị hệ số được lượng tử hóa, ví dụ các hệ số DCT, để khôi phục tín hiệu biến đổi và đơn vị biến đổi ngược 363, 463, thực hiện biến đổi ngược với tín hiệu biến đổi được khôi phục trong đó đầu ra của đơn vị biến đổi ngược 363, 463 chứa (các) khối được khôi phục. Bộ giải

mã lỗi dự đoán cũng có thể bao gồm bộ lọc khối có thể lọc (các) khối được khôi phục theo thông tin được giải mã khác và các tham số lọc.

Bộ mã hóa entropy 330, 430 nhận đầu ra của bộ mã hóa lỗi dự đoán 303, 403 và có thể thực hiện mã hóa entropy thích hợp/mã hóa độ dài biến đổi đối với tín hiệu để tạo ra khả năng phát hiện lỗi và khả năng hiệu chỉnh. Các đầu ra của các bộ mã hóa entropy 330, 430 có thể được đưa vào dòng bit ví dụ với bộ trộn kênh 508.

Fig.8 thể hiện sơ đồ khối của bộ giải mã video thích hợp để sử dụng các phương án của sáng chế. Fig.8 minh họa cấu trúc của bộ giải mã hai lớp, nhưng cần hiểu rằng các hoạt động giải mã có thể được sử dụng một cách tương tự trong bộ giải mã một lớp.

Bộ giải mã video 550 bao gồm phần bộ giải mã thứ nhất 552 cho các thành phần góc nhìn cơ sở và phần bộ giải mã thứ hai 554 đối với các thành phần không phải góc nhìn cơ sở. Khối 556 minh họa bộ phân kênh để truyền thông tin liên quan đến các thành phần góc nhìn cơ sở đến phần bộ giải mã thứ nhất 552 và để truyền thông tin liên quan đến các thành phần không phải góc nhìn cơ sở đến phần bộ giải mã thứ hai 554. Tham chiếu P'n viết tắt bởi biểu diễn được dự đoán của khối hình ảnh. Tham chiếu D'n đại diện cho tín hiệu lỗi dự đoán được khôi phục. Các khối 704, 804 minh họa các ảnh được khôi phục sơ bộ ($I'n$). Tham chiếu R'n đại diện cho ảnh được khôi phục cuối cùng. Các khối 703, 803 minh họa biến đổi ngược (T^{-1}). Các khối 702, 802 minh họa lượng tử hóa ngược (Q^{-1}). Các khối 700, 800 minh họa giải mã entropy (E^{-1}). Các khối 706, 806 minh họa bộ nhớ khung tham chiếu (RFM). Các khối 707, 807 minh họa dự đoán (P) (dự đoán liên ảnh hoặc dự đoán nội ảnh). Các khối 708, 808 minh họa việc lọc (F). Các khối 709, 809 có thể được sử dụng để kết hợp thông tin lỗi dự đoán được giải mã với các thành phần góc nhìn cơ sở/không phải góc nhìn cơ sở được dự đoán để thu các ảnh được khôi phục sơ bộ ($I'n$). Các ảnh góc nhìn cơ sở được khôi phục sơ bộ và được lọc có thể được xuất ra 710 từ phần bộ giải mã thứ nhất 552 và các ảnh góc nhìn cơ sở được khôi phục và được lọc sơ bộ có thể được xuất ra 810 từ phần bộ giải mã thứ hai 554.

Ở đây, bộ giải mã cần được hiểu bao gồm bất kỳ đơn vị hoạt động nào có thể thực thi các hoạt động giải mã, như bộ đọc, bộ thu, công, bộ phân kênh và/hoặc bộ giải mã.

Bộ giải mã entropy 700, 800 thực hiện giải mã entropy đối với tín hiệu thu được. Bộ giải mã entropy 700, 800 do đó thực hiện phép nghịch đảo với bộ mã hóa entropy của bộ mã hóa 330, 430 được mô tả ở trên. Bộ giải mã entropy 700, 800 xuất ra các kết quả của giải mã entropy đến bộ giải mã lỗi dự đoán 701, 801 và bộ dự đoán điểm ảnh 704, 804.

Bộ dự đoán điểm ảnh 704, 804 nhận đầu ra của bộ giải mã entropy 700, 800. Đầu ra của bộ giải mã entropy 700, 800 có thể bao gồm chỉ báo về chế độ dự đoán được sử dụng trong bước mã hóa khôi hiện thời. Bộ dự đoán 707, 807 có thể thực hiện dự đoán nội ảnh hoặc liên ảnh như được xác định bởi chỉ báo và xuất ra biểu diễn được dự đoán của khối ảnh đến bộ tổ hợp thứ nhất 709, 809. Biểu diễn được dự đoán của khối ảnh được sử dụng kết hợp với tín hiệu lỗi dự đoán được khôi phục để tạo ra ảnh được khôi phục sơ bộ. Ảnh được khôi phục sơ bộ có thể được sử dụng trong bộ dự đoán hoặc có thể được truyền đến đơn vị đệm khung tham chiếu 711, 811. Đơn vị đệm khung tham chiếu 711, 811 có thể làm đầy trong các vùng bên ngoài vùng hình ảnh hiệu dụng theo các phương án khác để tạo ra ảnh tham chiếu được đệm. Hình ảnh tham chiếu được đệm có thể được truyền đến bộ lọc 708, 808. Bộ lọc 708, 808 có thể áp dụng việc lọc mà kết xuất tín hiệu được khôi phục cuối cùng. Tín hiệu được khôi phục cuối cùng có thể được lưu trữ trong bộ nhớ khung tham chiếu 706, 806. Bộ nhớ khung tham chiếu 706, 806 còn có thể được kết nối với bộ dự đoán 707, 807 cho các hoạt động dự đoán.

Cần hiểu rằng đơn vị đệm khung tham chiếu 711, 811 có thể hoạt động trong các đơn vị khác nhau, có thể được xác định trước ví dụ trong tiêu chuẩn mã hóa hoặc có thể được giải mã bởi bộ giải mã từ dòng bit. Theo một phương án, đơn vị đệm khung tham chiếu 711, 811 sử dụng toàn bộ hình ảnh được khôi phục sơ bộ làm đầu vào. Theo phương án khác, đơn vị đệm khung tham chiếu 711, 811 sử dụng khung thành phần được khôi phục sơ bộ làm đầu vào, và do đó có thể

được gọi nhiều hơn một lần trên mỗi ảnh được khôi phục sơ bộ (phụ thuộc vào số lượng các khung thành phần mà nó bao quanh). Vẫn theo phương án khác, đơn vị đệm khung tham chiếu 711, 811 sử dụng khối được khôi phục sơ bộ, hàng mẫu biên của khối được khôi phục sơ bộ (khi mở rộng biên theo hướng chiều dọc), hoặc cột mẫu biên của khối được khôi phục sơ bộ (khi mở rộng biên theo hướng chiều ngang) làm đầu vào.

Cần hiểu rằng theo các phương án khác bộ dự đoán điểm ảnh 704, 804 có thể được bố trí để thực hiện chuỗi bước lọc và đệm theo các cách khác nhau. Ví dụ, bộ dự đoán điểm ảnh 704, 804 có thể thực hiện đệm khung tham chiếu và lọc theo thứ tự ngược lại, tức là lọc trước khi đệm khung tham chiếu. Theo một ví dụ khác, bộ dự đoán điểm ảnh 704, 804 có thể thực hiện lọc trong hai giai đoạn, như bước lọc khoảng dịch thích ứng mẫu và bước lọc giải khối, trong đó giai đoạn thứ nhất diễn ra trước khi đệm khung tham chiếu và giai đoạn thứ hai diễn ra sau khi đệm khung tham chiếu. Cần lưu ý rằng đó là các ví dụ không giới hạn và bộ dự đoán điểm ảnh 704, 804 có thể thực hiện lọc theo cách khác.

Tín hiệu được khôi phục cuối cùng cũng có thể được truyền đến bộ xén đầu ra 712, 812, có thể thực hiện một trong số hoặc cả hai (nhưng không giới hạn ở) bước sau. Bộ xén đầu ra 712, 812 có thể xén ảnh được khôi phục cuối cùng theo vùng hình ảnh hiệu dụng được chỉ báo, hoặc cửa sổ xén, hoặc tương tự. Bộ xén đầu ra 712, 812 có thể thiết lập các giá trị mẫu bên ngoài vùng hình ảnh hiệu dụng (trong phần mở rộng biên) ví dụ với giá trị định trước ví dụ biểu diễn màu đen.

Bộ xén đầu ra 712, 812 có thể tạo ra hình ảnh xuất ra dưới dạng đầu ra. Bộ xén đầu ra 712, 812 cũng có thể chỉ báo (các) vùng hình ảnh hiệu dụng cùng với ảnh đầu ra.

Bộ giải mã lỗi dự đoán 701, 801 nhận đầu ra của bộ giải mã entropy 700, 800. Bộ giải lượng tử hóa 702, 802 của bộ giải mã lỗi dự đoán có thể giải lượng tử hóa đầu ra của bộ giải mã entropy và khối biến đổi ngược 703, 803 có thể thực hiện hoạt động biến đổi ngược với đầu ra tín hiệu được giải lượng tử hóa bởi bộ giải lượng tử hóa 702, 802. Nói cách khác hoặc ngoài ra, bộ giải mã lỗi dự đoán

701, 801 có thể bao gồm bộ giải mã mạch rẽ biến đổi, trong đó các giá trị mẫu được giải mã không liên quan đến hoạt động biến đổi ngược, nhưng tốt hơn là ví dụ với việc giải mã giá trị mẫu khác nhau. Đầu ra của bộ giải mã entropy 700, 800 cũng có thể chỉ báo rằng lỗi dự đoán tín hiệu không được áp dụng và trong trường hợp này bộ giải mã lỗi dự đoán 701, 801 có thể tạo ra tất cả tín hiệu đầu ra bằng không.

Sau đây, một số ví dụ về việc sửa đổi khả dĩ đối với bộ lọc trong bộ mã hóa sẽ được đề xuất.

Khi xác định các giá trị tham số lọc, bộ mã hóa có thể bao gồm các mẫu bên trong vùng hình ảnh hiệu dụng để xem xét và loại trừ các mẫu bên ngoài vùng hình ảnh hiệu dụng khỏi việc xem xét. Ví dụ, các giá trị tham số SAO có thể thu được bằng cách chỉ xét các giá trị mẫu bên trong vùng hình ảnh hiệu dụng.

Để xử lý khối trong vùng mở rộng biên, bộ mã hóa có thể thực hiện một hoặc nhiều bước sau. Bộ mã hóa có thể bỏ qua việc mã hóa các khối ở bên ngoài vùng hình ảnh hiệu dụng. Ở đây, các khối có thể là ví dụ các đơn vị cây mã hóa và/hoặc các đơn vị mã hóa. Dự đoán nội ảnh và việc lọc sử dụng các mẫu của các khối bên ngoài vùng hình ảnh hiệu dụng có thể không có hiệu lực. Bộ mã hóa có thể xác định khối phân vùng trên cơ sở của vùng hình ảnh hiệu dụng. Có thể xác định được rằng các khối hoàn toàn ở bên ngoài vùng hình ảnh hiệu dụng không được phân vùng. Bộ mã hóa có thể chọn chế độ trên cơ sở của vùng hình ảnh hiệu dụng. Có thể được xác định rằng các khối hoàn toàn ở bên ngoài vùng hình ảnh hiệu dụng được mã hóa với chế độ cụ thể, có thể được xác định trước trên cơ sở mà nó cung cấp tốc độ bit nhỏ. Ví dụ, bộ mã hóa có thể chọn chế độ nội ảnh khả dĩ nhất (như được xác định bằng bộ mã hóa entropy). Bộ mã hóa cũng có thể bỏ qua việc mã hóa lỗi dự đoán của các khối hoàn toàn ở bên ngoài vùng hình ảnh hiệu dụng. Bộ mã hóa có thể chỉ báo rằng không có lỗi dự đoán có mặt trong dòng bit đối với khối hoàn toàn ở bên ngoài vùng hình ảnh hiệu dụng.

Tương ứng, bộ giải mã có thể thực hiện một hoặc nhiều bước sau đây. Bộ giải mã có thể bỏ qua việc giải mã các khối ở bên ngoài vùng hình ảnh hiệu dụng. Ở đây, ví dụ các khối có thể là các đơn vị cây mã hóa và/hoặc các đơn vị mã hóa.

Trong một số trường hợp, không có các thành phần cú pháp có mặt trong dòng bit đối với các khối ở bên ngoài vùng hình ảnh hiệu dụng và bộ giải mã theo dõi vị trí khối nào mà dữ liệu mã hóa tiếp theo tương ứng với. Trong một số trường hợp, các khối ở bên ngoài vùng hình ảnh hiệu dụng sử dụng cùng cú pháp với các khối hoàn toàn hoặc một phần ở bên trong vùng hình ảnh hiệu dụng, và bộ giải mã có thể giải mã entropy các khối này nhưng bỏ qua việc giải mã lỗi dự đoán và dự đoán điểm ảnh đối với các khối này. Bộ giải mã có thể vô hiệu hóa dự đoán nội ảnh và việc lọc sử dụng các mẫu của các khối bên ngoài vùng hình ảnh hiệu dụng. Bộ giải mã có thể xác định khối phân vùng trên cơ sở của vùng hình ảnh hiệu dụng. Có thể được xác định rằng các khối hoàn toàn ở bên ngoài vùng hình ảnh hiệu dụng không được phân vùng. Trong một số trường hợp, không có các thành phần cú pháp có mặt trong dòng bit đối với việc phân vùng khối của các khối nằm hoàn toàn ở bên ngoài vùng hình ảnh hiệu dụng. Bộ giải mã có thể chọn chế độ trên cơ sở của vùng hình ảnh hiệu dụng. Có thể xác định được rằng các khối hoàn toàn ở bên ngoài vùng hình ảnh hiệu dụng được giải mã với chế độ cụ thể, có thể được xác định trước ví dụ trong tiêu chuẩn mã hóa. Ví dụ, bộ giải mã có thể chọn chế độ nội ảnh khả dĩ nhất (như được xác định bởi bộ giải mã entropy). Bộ giải mã có thể bỏ qua việc giải mã lỗi dự đoán của các khối hoàn toàn ở bên ngoài vùng hình ảnh hiệu dụng. Trong một số trường hợp, không có các thành phần cú pháp có mặt trong dòng bit đối với các khối lỗi dự đoán hoàn toàn ở bên ngoài vùng hình ảnh hiệu dụng.

Việc dự đoán vector chuyển động của H.265/HEVC được mô tả sau đây dưới dạng ví dụ về hệ thống hoặc phương pháp trong đó các phương án có thể được áp dụng.

H.265/HEVC bao gồm hai mô hình dự đoán vector chuyển động, cụ thể là dự đoán vector chuyển động cải tiến (advanced motion vector prediction - AMVP) và chế độ hợp nhất. Trong AMVP hoặc chế độ hợp nhất, danh sách của các ứng viên của vector chuyển động thu được đối với PU. Có hai kiểu ứng viên: các ứng viên không gian và các ứng viên thời gian, trong đó các ứng viên thời gian cũng có thể được gọi là các ứng viên TMVP. Các nguồn của các bộ dự đoán vector

chuyển động ứng viên được thể hiện trên Fig.6c và Fig.6d. X viết tắt của đơn vị dự đoán hiện thời. A_0, A_1, B_0, B_1, B_2 trên Fig.6c là các ứng viên không gian trong khi C_0, C_1 trên Fig.6d là các ứng viên thời gian. Khối bao gồm hoặc tương ứng với ứng viên C_0 hoặc C_1 trên Fig.6d, bất kỳ thứ gì là nguồn cho ứng viên thời gian, có thể được gọi là khối đồng vị trí.

Việc thu nhận danh sách ứng viên có thể được thực hiện ví dụ như sau, trong khi cần hiểu rằng các khả năng khác có thể có để thu nhận danh sách ứng viên. Nếu việc chiếm chỗ danh sách ứng viên không ở mức tối đa, các ứng viên không gian được đưa vào danh sách ứng viên trước tiên nếu chúng khả dụng và không có mặt trong danh sách ứng viên. Sau đó, nếu việc chiếm chỗ của danh sách ứng viên không ở mức tối đa, ứng viên thời gian được đưa vào trong danh sách ứng viên. Nếu số lượng các ứng viên không chạm mức tối đa số lượng cho phép, các ứng viên dự đoán kép được kết hợp (đối với các lát B) và vectơ chuyển động không được thêm vào. Sau khi danh sách ứng viên này được tạo ra, bộ mã hóa quyết định thông tin chuyển động cuối cùng từ các ứng viên ví dụ dựa trên quyết định tối ưu hóa biến dạng tỷ lệ (rate-distortion optimization - RDO) và mã hóa chỉ số của ứng viên được chọn trong dòng bit. Tương tự, bộ giải mã giải mã chỉ số của ứng viên được chọn từ dòng bit, tạo thành danh sách ứng viên, và sử dụng chỉ số được giải mã để chọn bộ dự đoán vectơ chuyển động từ danh sách ứng viên.

Trong H.265/HEVC, AMVP và chế độ hợp nhất có thể khác biệt như sau. Trong AMVP, bộ mã hóa chỉ báo liệu dự đoán đơn hay dự đoán kép được sử dụng và các hình ảnh tham chiếu nào được sử dụng cũng như mã hóa sự chênh lệch vectơ chuyển động. Trong chế độ hợp nhất, chỉ có ứng viên được chọn từ danh sách ứng viên được mã hóa thành dòng bit chỉ báo đơn vị dự đoán hiện thời có cùng thông tin chuyển động như với bộ dự đoán được chỉ báo. Do đó, chế độ hợp nhất tạo ra các vùng bao gồm các khối dự đoán lân cận chia sẻ thông tin chuyển động giống nhau, chỉ được báo hiệu một lần cho mỗi vùng. Sự khác biệt khác giữa AMVP và chế độ hợp nhất trong H.265/HEVC là số lượng các ứng

viên tối đa của AMVP là 2 trong khi số lượng các ứng viên tối đa của chế độ hợp nhất là 5.

Dự đoán vector chuyển động cải tiến có thể hoạt động ví dụ như sau, trong khi các nhận biết tương tự khác của dự đoán vector chuyển động cải tiến cũng là khả dĩ ví dụ với các tập hợp vị trí ứng viên khác và các vị trí ứng viên với các tập hợp vị trí ứng viên. Hai bộ dự đoán vector chuyển động theo không gian (MVP) có thể thu được và bộ dự đoán vector chuyển động theo thời gian (TMVP) có thể thu được. Chúng có thể được chọn trong số các vị trí: các vị trí ba bộ dự đoán vector chuyển động ứng viên theo không gian được định vị bên trên khối dự đoán hiện thời (B_0, B_1, B_2) và hai tập ở bên trái (A_0, A_1). Bộ dự đoán vector chuyển động thứ nhất là khả dụng (ví dụ nằm trong cùng lát, được mã hóa liên ảnh, v.v..) theo thứ tự định trước của mỗi tập hợp vị trí ứng viên, (B_0, B_1, B_2) hoặc (A_0, A_1), có thể được chọn biểu diễn hướng dự đoán (lên hoặc sang trái) trong cạnh tranh vector chuyển động. Chỉ số tham chiếu đối với bộ dự đoán vector chuyển động theo thời gian có thể được chỉ báo bởi bộ mã hóa trong phần đầu của lát (ví dụ dưới dạng thành phần cú pháp `collocated_ref_idx`). Bộ dự đoán vector chuyển động thứ nhất là khả dụng (ví dụ được mã hóa liên kết) theo thứ tự định trước của các vị trí ứng viên thời gian tiềm năng, ví dụ theo thứ tự (C_0, C_1), có thể được chọn làm nguồn cho bộ dự đoán vector chuyển động theo thời gian. Vector chuyển động thu được từ vị trí ứng viên khả dụng thứ nhất trong hình ảnh đồng vị trí có thể được định tỉ lệ theo các tỷ lệ của các chênh lệch của số đếm thứ tự hình ảnh của hình ảnh tham chiếu của bộ dự đoán vector chuyển động theo thời gian, hình ảnh đồng vị trí, và hình ảnh hiện thời. Ngoài ra, việc kiểm tra thừa có thể được thực hiện trong số các ứng viên để loại bỏ các ứng viên giống nhau, có thể dẫn đến việc đưa vector chuyển động không vào danh sách ứng viên. Bộ dự đoán vector chuyển động có thể được chỉ báo trong dòng bit ví dụ bằng chỉ báo hướng của bộ dự đoán vector chuyển động theo không gian (lên hoặc sang trái) hoặc việc chọn ứng viên bộ dự đoán vector chuyển động theo thời gian. Hình ảnh đồng vị trí (co-located) cũng có thể được gọi là hình ảnh đồng vị trí (collocated), nguồn cho dự đoán vector chuyển động, hoặc hình ảnh nguồn cho dự đoán vector chuyển động.

Chế độ/quy trình/cơ chế hợp nhất (merging/merge) có thể hoạt động ví dụ như sau, trong khi các nhận biết tương tự khác của chế độ hợp nhất cũng khả dĩ ví dụ với các tập hợp vị trí ứng viên khác và các vị trí ứng viên với các tập hợp vị trí ứng viên.

Trong chế độ/quy trình/cơ chế hợp nhất (merging/merge), trong đó tất cả thông tin chuyển động của khối/PU được dự đoán và được sử dụng mà không có bất kỳ sửa đổi/hiệu chỉnh nào. Thông tin chuyển động nêu trên đối với PU có thể bao gồm một hoặc nhiều thông tin sau: 1) Thông tin liệu ‘PU được dự đoán đơn chỉ sử dụng hình ảnh tham chiếu list0’ hoặc ‘PU được dự đoán đơn chỉ sử dụng hình ảnh tham chiếu list1’ hoặc ‘PU được dự đoán kép sử dụng cả hình ảnh tham chiếu list0 và list1’; 2) Giá trị vector chuyển động tương ứng với hình ảnh tham chiếu list0, mà có thể bao gồm thành phần vector chuyển động chiều ngang và chiều dọc; 3) Chỉ số hình ảnh tham chiếu trong hình ảnh tham chiếu list0 và/hoặc mã nhận dạng của hình ảnh tham chiếu được trỏ đến bởi vector chuyển động tương ứng với danh sách hình ảnh tham chiếu 0, trong đó mã nhận dạng của hình ảnh tham chiếu có thể là ví dụ giá trị số đếm thứ tự hình ảnh, giá trị nhận dạng lớp (để dự đoán liên lớp), hoặc cặp giá trị số đếm thứ tự hình ảnh và giá trị nhận dạng lớp; 4) Thông tin về hình ảnh tham chiếu đánh dấu hình ảnh tham chiếu, ví dụ thông tin liệu hình ảnh tham chiếu được đánh dấu là “được sử dụng cho tham chiếu ngắn hạn” hoặc “được sử dụng cho tham chiếu dài hạn”; 5) - 7) tương ứng giống với 2) - 4), nhưng cho hình ảnh tham chiếu list1.

Tương tự, dự đoán thông tin chuyển động được thực thi sử dụng thông tin chuyển động của các khối liền kề và/hoặc các khối đồng vị trí trong các hình ảnh tham chiếu theo thời gian. Danh sách, thường gọi là danh sách kết hợp, có thể được tạo thành bằng cách đưa vào các ứng viên dự đoán chuyển động được kết hợp với các khối liền kề/đồng vị trí khả dụng và chỉ số của ứng viên dự đoán chuyển động được chọn trong danh sách được báo hiệu và thông tin chuyển động của ứng viên được chọn được sao chép vào thông tin chuyển động của PU hiện thời. Khi cơ chế hợp nhất được sử dụng cho toàn bộ CU và tín hiệu dự đoán cho CU được sử dụng dưới dạng tín hiệu khôi phục, tức là phần dư dự đoán không

được xử lý, kiểu mã hóa/giải mã này CU thường gọi là chế độ bỏ qua hoặc chế độ bỏ qua dựa trên việc hợp nhất. Ngoài chế độ bỏ qua, cơ chế hợp nhất cũng có thể được sử dụng cho các PU riêng biệt (không nhất thiết là toàn bộ CU ở chế độ bỏ qua) và trong trường hợp này, phần dư dự đoán có thể được sử dụng để tăng cường chất lượng dự đoán. Kiểu chế độ dự đoán này được gọi chế độ hợp nhất liên ảnh.

Một trong số các ứng viên trong danh sách kết hợp và/hoặc danh sách ứng viên đối với AMVP hoặc bất kỳ danh sách ứng viên vector chuyển động tương tự nào có thể là ứng viên TMVP hoặc tương tự, mà có thể thu được từ khối đồng vị trí trong hình ảnh tham chiếu được chỉ báo hoặc được suy ra, như hình ảnh tham chiếu được chỉ báo ví dụ trong phần đầu của lát. Theo HEVC, danh sách hình ảnh tham chiếu được sử dụng để thu phân vùng đồng vị trí được chọn theo thành phần cú pháp `collocated_from_10_flag` trong phần đầu của lát. Khi cờ bằng 1, xác định rằng hình ảnh mà có phân vùng đồng vị trí được thu nhận từ danh sách 0, nếu không thì hình ảnh này được thu nhận từ danh sách 1. Khi `collocated_from_10_flag` không có mặt, nó được suy ra bằng 1. `collocated_ref_idx` trong phần đầu của lát quy định chỉ số tham chiếu của hình ảnh có phân vùng đồng vị trí. Khi lát hiện thời là lát P, `collocated_ref_idx` đề cập đến hình ảnh trong danh sách 0. Khi lát hiện thời là lát B, `collocated_ref_idx` đề cập đến hình ảnh trong danh sách 0 nếu `collocated_from_10` là 1, nếu không thì nó đề cập đến hình ảnh trong danh sách 1. `collocated_ref_idx` luôn đề cập đến lối vào danh sách hợp lệ, và hình ảnh thu được là giống nhau cho tất cả các lát của hình ảnh được mã hóa. Khi `collocated_ref_idx` không có mặt, nó được suy ra bằng 0.

Theo HEVC chỉ số tham chiếu đích cho dự đoán vector chuyển động theo thời gian trong danh sách kết hợp được thiết lập là 0 khi chế độ mã hóa chuyển động là chế độ hợp nhất. Khi chế độ mã hóa chuyển động theo HEVC sử dụng dự đoán vector chuyển động theo thời gian gian là chế độ dự đoán vector chuyển động cải tiến, các giá trị chỉ số tham chiếu đích được chỉ báo rõ ràng (ví dụ trên mỗi PU).

Theo HEVC, độ khả dụng của vector chuyển động được dự đoán ứng viên (PMV) có thể được xác định như sau (cả các ứng viên theo không gian và thời gian) (SRTP = hình ảnh tham chiếu ngắn hạn, LTRP = hình ảnh tham chiếu dài hạn):

Hình ảnh tham chiếu cho chỉ số tham chiếu đích	Hình ảnh tham chiếu cho ứng viên PMV	Độ khả dụng ứng viên PMV
STRP	STRP	“khả dụng” (và được định tỉ lệ)
STRP	LTRP	“không khả dụng”
LTRP	STRP	“không khả dụng”
LTRP	LTRP	“khả dụng” (nhưng không được định tỉ lệ)

Theo HEVC, khi giá trị chỉ số tham chiếu đích được xác định, giá trị vector chuyển động của dự đoán vector chuyển động theo không gian có thể thu được như sau: Vector chuyển động PMV ở khối được xếp cùng vị trí với vị trí liền kề bên phải phía dưới (vị trí C0 trên Fig.6d) của đơn vị dự đoán hiện thời được theo. Hình ảnh trong đó khối đồng vị trí được bố trí có thể ví dụ được xác định theo chỉ số tham chiếu được báo hiệu trong phần đầu của lát như được mô tả ở trên. Nếu PMV ở vị trí C0 không khả dụng, vector chuyển động PMV ở vị trí C1 (xem Fig.6d) của hình ảnh đồng vị trí thu được. Vector chuyển động khả dụng PMV được xác định tại khối đồng vị trí được định tỷ lệ tương ứng với tỷ lệ của sự chênh lệch số đếm thứ tự hình ảnh thứ nhất và chênh lệch số đếm thứ tự hình ảnh thứ hai. Sự chênh lệch số đếm thứ tự hình ảnh thứ nhất được thu nhận giữa hình ảnh chứa khối đồng vị trí và hình ảnh tham chiếu của vector chuyển động của khối đồng vị trí. Sự chênh lệch số đếm thứ tự hình ảnh thứ hai được thu nhận giữa hình ảnh hiện thời và hình ảnh tham chiếu đích. Nếu một mà không phải cả hai hình ảnh tham chiếu đích và hình ảnh tham chiếu của vector chuyển động của khối đồng vị trí là hình ảnh tham chiếu dài hạn (trong khi hình ảnh khác là hình ảnh

tham chiếu ngắn hạn), ứng viên TMVP có thể được xem xét không khả dụng. Nếu cả hình ảnh tham chiếu đích và hình ảnh tham chiếu của vector chuyển động của khối đồng vị trí là các hình ảnh tham chiếu dài hạn, không có việc định tỉ lệ vector chuyển động dựa vào POC có thể được áp dụng.

Các kiểu tham số chuyển động hoặc thông tin chuyển động có thể bao gồm nhưng không giới hạn ở một hoặc nhiều kiểu sau đây:

- chỉ báo kiểu dự đoán (ví dụ dự đoán nội ảnh, dự đoán đơn, dự đoán kép) và/hoặc số lượng các hình ảnh tham chiếu;
 - chỉ báo hướng dự đoán, như dự đoán liên ảnh (còn được biến đến là dự đoán theo thời gian), dự đoán liên lớp, dự đoán liên góc nhìn, dự đoán tổng hợp góc nhìn (VSP), và dự đoán liên thành phần (có thể được chỉ báo trên từng hình ảnh tham chiếu và/hoặc mỗi kiểu dự đoán và trong đó theo một số phương án dự đoán liên góc nhìn và tổng hợp góc nhìn có thể được xem xét cùng nhau dưới dạng một hướng dự đoán) và/hoặc
 - chỉ báo về kiểu hình ảnh tham chiếu, như hình ảnh tham chiếu ngắn hạn và/hoặc hình ảnh tham chiếu dài hạn và/hoặc hình ảnh tham chiếu liên lớp (có thể được chỉ báo ví dụ trên từng hình ảnh tham chiếu)
 - chỉ số tham chiếu về danh sách hình ảnh tham chiếu và/hoặc bất kỳ mã nhận dạng nào khác của hình ảnh tham chiếu (có thể được chỉ báo ví dụ trên từng hình ảnh tham chiếu và kiểu của hình ảnh tham chiếu có thể phụ thuộc vào hướng dự đoán và/hoặc kiểu hình ảnh tham chiếu và có thể được đi kèm với các phần thông tin liên quan khác, như danh sách hình ảnh tham chiếu hoặc tương tự mà chỉ số tham chiếu áp dụng);
 - thành phần vector chuyển động ngang (có thể được chỉ báo ví dụ trên từng khối dự đoán hoặc trên từng chỉ số tham chiếu hoặc tương tự);
 - thành phần vector chuyển động chiều dọc (có thể được chỉ báo ví dụ trên từng khối dự đoán hoặc trên từng chỉ số tham chiếu hoặc tương tự);
 - một hoặc nhiều tham số, như chênh lệch số đếm thứ tự hình ảnh và/hoặc sự tách biệt camera tương đối giữa hình ảnh chứa hoặc được kết hợp

với các tham số chuyển động và hình ảnh tham chiếu của nó, mà có thể được sử dụng để định tỉ lệ của thành phần vector chuyển động ngang và/hoặc thành phần vector chuyển động chiều dọc trong một hoặc nhiều quy trình dự đoán vector chuyển động (trong đó một hoặc nhiều tham số nêu trên có thể được chỉ báo ví dụ trên mỗi hình ảnh tham chiếu hoặc mỗi chỉ số tham chiếu hoặc tương tự);

- các tọa độ của khối trong đó các tham số chuyển động và/hoặc thông tin chuyển động áp dụng, ví dụ các tọa độ của mẫu bên trái phía trên của khối trong các đơn vị mẫu luma;
- mở rộng (ví dụ chiều rộng và chiều cao) của khối trong đó các tham số chuyển động và/hoặc thông tin chuyển động áp dụng.

Nói chung, các cơ chế dự đoán vector chuyển động, như các cơ chế dự đoán vector chuyển động được mô tả trên dưới dạng các ví dụ, có thể bao gồm dự đoán hoặc kế thừa của các tham số chuyển động được định trước hoặc được chỉ báo nhất định.

Trường chuyển động được kết hợp hình ảnh có thể được xem xét để bao gồm tập hợp thông tin chuyển động được tạo ra cho mọi khối mã hóa của hình ảnh. Trường chuyển động có thể truy cập được với các tọa độ của khối, chẳng hạn. Trường chuyển động có thể được sử dụng ví dụ trong TMVP hoặc bất kỳ cơ chế dự đoán chuyển động khác nào trong đó nguồn hoặc tham chiếu để dự đoán ngoài hình ảnh được mã hóa (giải mã) được sử dụng.

Độ mịn không gian khác hoặc các đơn vị có thể được áp dụng để biểu diễn và/hoặc lưu trữ trường chuyển động. Ví dụ, lưới thông thường của các đơn vị không gian có thể được sử dụng. Ví dụ, hình ảnh có thể được chia thành các khối hình chữ nhật có kích thước nhất định (với sự loại trừ khả dĩ của các khối ở các mép của hình ảnh, như trên mép phải và mép đáy). Ví dụ, kích thước của đơn vị không gian có thể bằng kích thước nhỏ nhất mà chuyển động riêng biệt có thể được chỉ báo bằng bộ mã hóa trong dòng bit, như khối 4x4 trong các đơn vị mẫu luma. Ví dụ, trường chuyển động được nén có thể được sử dụng, trong đó đơn vị không gian có thể bằng kích thước được định trước hoặc được chỉ báo, như khối

16x16 trong các đơn vị mẫu luma, kích thước nào có thể lớn hơn kích thước nhỏ nhất để chỉ báo chuyển động riêng biệt. Ví dụ, bộ mã hóa và/hoặc bộ giải mã HEVC có thể được chỉ báo theo cách mà việc giảm lưu trữ dữ liệu chuyển động (MDSR) hoặc việc nén trường chuyển động được thực hiện cho mỗi trường chuyển động được giải mã (trước khi sử dụng trường chuyển động đối với bất kỳ dự đoán nào giữa các hình ảnh). Theo một phương án triển khai HEVC, MDSR có thể làm giảm độ mịn của dữ liệu chuyển động đến các khối 16x16 trong các đơn vị mẫu luma bằng cách duy trì chuyển động áp dụng được cho mẫu bên trái phía trên của khối 16x16 trong trường chuyển động nén. Bộ mã hóa có thể mã hóa (các) chỉ báo liên quan đến đơn vị không gian của trường chuyển động nén dưới dạng một hoặc nhiều thành phần cú pháp và/hoặc các giá trị thành phần cú pháp ví dụ trong cấu trúc cú pháp cấp độ chuỗi, như tập hợp các tham số video hoặc tập hợp các tham số trình tự. Trong một số phương pháp và/hoặc thiết bị mã hóa (giải mã), trường chuyển động có thể được biểu diễn và/hoặc được lưu trữ theo sự phân vùng khối của dự đoán chuyển động (ví dụ theo các đơn vị dự đoán của tiêu chuẩn HEVC). Theo một số phương pháp và/hoặc thiết bị mã hóa (giải mã), kết hợp của lưới thông thường và sự phân vùng khối có thể được áp dụng sao cho chuyển động được kết hợp với các phân vùng lớn hơn kích thước đơn vị không gian định trước hoặc được chỉ báo được biểu diễn và/hoặc được lưu trữ được kết hợp với các phân vùng này, trong đó chuyển động được kết hợp với các phân vùng nhỏ hơn hoặc không được căn thẳng với kích thước hoặc lưới đơn vị không gian định trước hoặc được chỉ báo được biểu diễn và/hoặc được lưu trữ đối với các đơn vị định trước hoặc được chỉ báo.

Phân mở rộng mã hóa video khả năng mở rộng được (SHVC) của HEVC cung cấp cơ chế để tạo ra khả năng mở rộng được không gian, độ sâu bit, dải màu, và chất lượng trong khi tận dụng phần dư liên lớp. Phân mở rộng đa góc nhìn (MV-HEVC) của HEVC cho phép việc mã hóa dữ liệu video đa góc nhìn thích hợp ví dụ cho các màn hình lập thể. Đối với MV-HEVC, các chuỗi video đa góc nhìn đầu vào để mã hóa thường được chụp bằng nhiều camera được bố trí theo hàng. Các trung tâm chiếu camera thường cùng nằm trên cùng một đường thẳng và cách đều nhau và các camera thường hướng theo cùng chiều. SHVC và

MV-HEVC chia sẻ cùng một cú pháp cấp độ cao và hầu hết các phần của quy trình giải mã cũng giống nhau, khiến nó cần hỗ trợ cả SHVC và MV-HEVC với cùng cách triển khai bộ mã hóa-giải mã.

SHVC và MV-HEVC cho phép hai kiểu dự đoán liên lớp, cụ thể là dự đoán mẫu và dự đoán chuyển động. Trong dự đoán mẫu liên lớp, hình ảnh tham chiếu liên lớp (ILR) được sử dụng để thu các giá trị mẫu của khối dự đoán. Trong MV-HEVC, hình ảnh lớp cơ sở được giải mã đóng vai trò là hình ảnh ILR mà không cần biến đổi. Trong khả năng mở rộng được không gian và dải màu của SHVC, xử lý liên lớp, như việc lấy mẫu lại, được áp dụng cho hình ảnh lớp cơ sở được giải mã để thu hình ảnh ILR. Trong quy trình lấy mẫu lại của SHVC, hình ảnh lớp cơ sở có thể được xén, được lấy mẫu lên và/hoặc được đệm để thu hình ảnh ILR. Vị trí tương đối của hình ảnh lớp cơ sở được lấy mẫu lên đến hình ảnh lớp tăng cường được chỉ báo thông qua các khoảng dịch vị trí lớp tham chiếu. Dấu hiệu này cho phép khả năng mở rộng được vùng quan tâm (ROI), trong đó chỉ có tập con của vùng hình ảnh của lớp cơ sở được tăng cường trong hình ảnh lớp tăng cường.

Dự đoán chuyển động liên lớp tận dụng cơ chế dự đoán vectơ chuyển động theo không gian (TMVP) của HEVC. Hình ảnh ILR được chỉ định là hình ảnh đồng vị trí đối với TMVP, và do đó là nguồn của các ứng viên TMVP trong quy trình dự đoán vectơ chuyển động. Trong khả năng mở rộng được không gian của SHVC, ánh xạ trường chuyển động (MFM) được sử dụng để thu thông tin chuyển động cho hình ảnh ILR từ thông tin chuyển động của hình ảnh lớp cơ sở. Trong MFM, sự phụ thuộc dự đoán trong các hình ảnh lớp cơ sở có thể được xem xét được lặp lại để tạo ra (các) danh sách hình ảnh tham chiếu cho các hình ảnh ILR. Đối với mỗi khối của trường chuyển động được ánh xạ lại với lưới của khối cụ thể (ví dụ lưới 16x16 theo HEVC), vị trí mẫu đồng vị trí trong hình ảnh nguồn để dự đoán liên lớp có thể thu được. Vị trí mẫu tham chiếu có thể ví dụ được thu được cho mẫu chính giữa của khối. Trong việc thu vị trí mẫu tham chiếu, tỷ lệ lấy mẫu lại, vùng tham chiếu, và vùng tham chiếu được lấy mẫu lại có thể được tính đến – xem thêm dưới đây về việc định nghĩa tỷ lệ lấy mẫu lại, vùng tham chiếu,

và vùng tham chiếu được lấy mẫu lại. Ngoài ra, vị trí mẫu tham chiếu có thể được khoanh tròn hoặc được làm vát để căn thẳng với lưới của khối của trường chuyển động của vùng tham chiếu hoặc hình ảnh nguồn để dự đoán liên lớp. Vector chuyển động tương ứng với vị trí mẫu tham chiếu thu được từ trường chuyển động của vùng tham chiếu hoặc hình ảnh nguồn để dự đoán liên lớp. Vector chuyển động này được định tỷ lệ lại theo tỷ lệ lấy mẫu lại và sau đó được bao gồm trong trường chuyển động được ánh xạ lại. Trường chuyển động được ánh xạ lại có thể được sử dụng tiếp theo dưới dạng nguồn cho TMVP hoặc tương tự. Ngược lại, MFM không được áp dụng trong MV-HEVC cho các hình ảnh lớp cơ sở được tham chiếu trong quy trình dự đoán chuyển động liên lớp.

Sự tương ứng không gian của hình ảnh lớp tham chiếu và hình ảnh lớp tăng cường có thể được suy ra hoặc có thể được chỉ báo với một hoặc nhiều kiểu của các khoảng dịch vị trí lớp tham chiếu. Theo HEVC, các khoảng dịch vị trí lớp tham chiếu có thể được bao gồm trong PPS bởi bộ mã hóa và được giải mã từ PPS bởi bộ giải mã. Các khoảng dịch vị trí lớp tham chiếu có thể được sử dụng nhưng không giới hạn ở việc đạt được khả năng mở rộng được vùng quan tâm (ROI). Các khoảng dịch vị trí lớp tham chiếu có thể bao gồm một hoặc nhiều khoảng dịch lớp tham chiếu định tỉ lệ, các khoảng dịch vùng tham chiếu, và các tập hợp pha lấy mẫu lại. Các khoảng dịch lớp tham chiếu định tỉ lệ có thể được xem xét để xác định các khoảng dịch chiều ngang và chiều dọc giữa mẫu trong hình ảnh hiện thời được xếp cùng vị trí với mẫu luma bên trái phía trên của vùng tham chiếu trong hình ảnh được giải mã trong lớp tham chiếu và các khoảng dịch chiều ngang và chiều dọc giữa mẫu trong hình ảnh hiện thời được xếp cùng vị trí với mẫu luma bên phải phía dưới của vùng tham chiếu trong hình ảnh được giải mã trong lớp tham chiếu. Cách khác để xem xét các khoảng dịch lớp tham chiếu định tỉ lệ để xác định các vị trí của các mẫu góc của vùng tham chiếu được lấy mẫu lên (hoặc tổng quát hơn, vùng tham chiếu được lấy mẫu lại) tương đối với các mẫu góc tương ứng của hình ảnh lớp tăng cường. Các khoảng dịch lớp tham chiếu định tỉ lệ có thể được xem xét để xác định sự tương ứng không gian của hình ảnh lớp hiện thời (trong đó các khoảng dịch vị trí lớp tham chiếu được chỉ báo) tương đối với vùng tham chiếu được định tỷ lệ của hình ảnh lớp tham chiếu

được định tỷ lệ. Các giá trị khoảng dịch lớp tham chiếu định tỉ lệ có thể được đánh dấu. Khi các khoảng dịch lớp tham chiếu định tỉ lệ là âm, hình ảnh trong đó các khoảng dịch vị trí lớp tham chiếu được chỉ báo tương ứng với vùng được xén của hình ảnh lớp tham chiếu. Các khoảng dịch vùng tham chiếu có thể được xem xét để xác định các khoảng dịch chiều ngang và chiều dọc giữa mẫu luma bên trái phía trên của vùng tham chiếu trong hình ảnh được giải mã trong lớp tham chiếu và mẫu luma bên trái phía trên của cùng một hình ảnh giải mã cũng như các khoảng dịch chiều ngang và chiều dọc giữa mẫu luma bên phải phía dưới của vùng tham chiếu trong hình ảnh được giải mã trong lớp tham chiếu và mẫu luma bên phải phía dưới của cùng một hình ảnh giải mã. Các khoảng dịch vùng tham chiếu có thể được xem xét để xác định sự tương ứng không gian của vùng tham chiếu trong hình ảnh lớp tham chiếu tương đối với hình ảnh lớp tham chiếu được giải mã. Các giá trị khoảng dịch vùng tham chiếu có thể được đánh dấu. Khi các khoảng dịch vùng tham chiếu là âm, hình ảnh lớp tham chiếu tương ứng với vùng hình ảnh được xén mà các khoảng dịch vị trí lớp tham chiếu được chỉ báo. Tập hợp pha lấy mẫu lại có thể được xem xét để xác định các khoảng dịch pha được sử dụng trong quy trình lấy mẫu lại của hình ảnh nguồn để dự đoán liên lớp. Các khoảng dịch pha khác có thể được cung cấp cho các thành phần luma và chroma.

Một số mô hình mã hóa video mở rộng được có thể yêu cầu các hình ảnh IRAP được căn thẳng qua các lớp theo cách mà tất cả các hình ảnh trong đơn vị truy cập là các hình ảnh IRAP hoặc không có hình ảnh nào trong đơn vị truy cập là hình ảnh IRAP. Các mô hình mã hóa video mở rộng được khác, như các phần mở rộng đa lớp của HEVC, có thể cho phép các hình ảnh IRAP không được căn thẳng, tức là một hoặc nhiều hình ảnh trong đơn vị truy cập là các hình ảnh IRAP, trong khi một hoặc nhiều hình ảnh khác trong đơn vị truy cập không phải là các hình ảnh IRAP. Các dòng bit mở rộng được với các hình ảnh IRAP hoặc tương tự không được căn thẳng qua các lớp có thể được sử dụng ví dụ để cung cấp các hình ảnh IRAP thường xuyên hơn trong lớp cơ sở, trong đó chúng có thể có kích thước mã hóa nhỏ hơn ví dụ do độ phân giải không gian nhỏ hơn. Quy trình hoặc cơ chế để khởi động theo lớp của việc giải mã có thể được bao gồm trong mô hình 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 theo lớp các lớp khác khi chúng chứa các hình ảnh IRAP. Nói cách khác, trong việc khởi động theo lớp cơ chế hoặc quy trình giải mã, các bộ giải mã tăng lên dần dần số lượng các lớp được giải mã (trong đó các lớp có thể biểu diễn sự tăng cường trong độ phân giải không gian, cấp độ chất lượng, các góc nhìn, các thành phần bổ sung như độ sâu, hoặc kết hợp của chúng) dưới dạng các hình ảnh tiếp theo từ các lớp tăng cường bổ sung được giải mã trong quy trình giải mã. Sự tăng dần của số lượng các lớp được giải mã có thể được nhận biết ví dụ dưới dạng sự tăng dần về chất lượng hình ảnh (trong trường hợp khả năng mở rộng được chất lượng và không gian).

Cơ chế khởi động theo lớp có thể tạo ra các hình ảnh không khả dụng cho các hình ảnh tham chiếu của hình ảnh thứ nhất theo thứ tự giải mã trong lớp tăng cường cụ thể. Nói cách khác, bộ giải mã có thể bỏ qua việc giải mã các hình ảnh liền trước, theo thứ tự giải mã, hình ảnh IRAP mà từ đó việc giải mã của lớp có thể được bắt đầu. Các hình ảnh này có thể được bỏ qua có thể được dán nhãn cụ thể bởi bộ mã hóa hoặc thực thể khác trong dòng bit. Ví dụ, một hoặc nhiều kiểu đơn vị NAL xác định có thể được sử dụng từ chúng. Các hình ảnh này, không kể việc chúng có được dán nhãn cụ thể với kiểu đơn vị NAL hay được suy ra ví dụ bởi bộ giải mã, có thể được gọi là các hình ảnh truy cập ngẫu nhiên chéo giữa các lớp (CL-RAS). Bộ giải mã có thể bỏ qua đầu ra của các hình ảnh không khả dụng được tạo ra và các hình ảnh CL-RAS được giải mã.

Khả năng mở rộng được có thể được cho phép theo hai cách cơ bản. Một trong số đó là bằng cách đưa vào các chế độ mã hóa mới để thực hiện dự đoán các giá trị điểm ảnh hoặc cú pháp từ các lớp thấp hơn của biểu diễn mở rộng được hoặc bằng cách đặt các hình ảnh lớp thấp hơn vào đệm hình ảnh tham chiếu (ví dụ đệm hình ảnh giải mã, DPB) của lớp cao hơn. Cách thứ nhất có thể linh động hơn và do đó có thể cung cấp hiệu suất mã hóa tốt hơn trong hầu hết các trường hợp. Tuy nhiên, khả năng mở rộng được dựa vào khung tham chiếu thứ hai, phương pháp này có thể được triển khai hiệu quả với các thay đổi tối thiểu cho các bộ mã hóa-giải mã một lớp trong khi vẫn đạt được phần lớn hiệu suất mã hóa đạt sự khả dụng. Về cơ bản bộ mã hóa-giải mã khả năng mở rộng được dựa vào tham chiếu

khung có thể được triển khai bằng cách sử dụng cùng phần cứng hoặc phần mềm cho tất cả các lớp, mà chỉ quan tâm đến việc quản lý DPB bởi các phương tiện bên ngoài.

Bộ mã hóa video mở rộng được (cũng được biết là tín hiệu trên tạp âm hoặc SNR) và/hoặc khả năng mở rộng được không gian có thể được triển khai dưới dạng sau đây. Đối với lớp cơ sở, bộ mã hóa và bộ giải mã video không mở rộng được thông thường có thể được sử dụng. Các hình ảnh được khôi phục/giải mã của lớp cơ sở được bao gồm trong đệm hình ảnh tham chiếu và/hoặc các danh sách hình ảnh tham chiếu cho lớp tăng cường. Trong trường hợp khả năng mở rộng được không gian, hình ảnh lớp cơ sở được khôi phục/giải mã có thể được lấy mẫu lên trước khi nó được đưa vào các danh sách hình ảnh tham chiếu cho hình ảnh lớp tăng cường. Các hình ảnh lớp cơ sở được giải mã có thể được đưa vào (các) danh sách hình ảnh tham chiếu để mã hóa/giải mã của hình ảnh lớp tăng cường tương tự với các hình ảnh tham chiếu được giải mã của lớp tăng cường. Do đó, bộ mã hóa có thể chọn hình ảnh lớp cơ sở tham chiếu dưới dạng tham chiếu dự đoán liên ảnh và chỉ báo việc sử dụng của nó với chỉ số hình ảnh tham chiếu trong dòng bit được mã hóa. Bộ giải mã giải mã từ dòng bit, ví dụ từ chỉ số hình ảnh tham chiếu, trong đó hình ảnh lớp cơ sở được sử dụng làm tham chiếu dự đoán liên ảnh cho lớp tăng cường. Khi hình ảnh lớp cơ sở được giải mã được sử dụng dưới dạng tham chiếu dự đoán cho lớp tăng cường, thường được gọi là hình ảnh tham chiếu liên lớp.

Trong khi đoạn trước mô tả bộ mã hóa-giải mã video mở rộng được với hai lớp có khả năng mở rộng được với lớp tăng cường và lớp cơ sở, cần hiểu rằng mô tả này có thể được tổng quát hóa tới bất kỳ hai lớp nào trong hệ thống thứ bậc có khả năng mở rộng được 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 xem là lớp cơ sở để mã hóa và/hoặc giải mã cho lớp tăng cường thứ hai. Hơn nữa, cần hiểu rằng có thể là các hình ảnh tham chiếu liên lớp từ nhiều hơn một lớp trong đệm hình ảnh tham chiếu hoặc các danh sách hình ảnh tham chiếu của lớp tăng cường,

và mỗi trong số các hình ảnh tham chiếu liên lớp này có thể được xem xét để được bố trí trong lớp cơ sở hoặc lớp tham chiếu đối với lớp tăng cường được mã hóa và/hoặc được giải mã. Hơn nữa, thay vào đó hoặc ngoài ra, cần hiểu rằng các kiểu khác nào của việc xử lý liên lớp so với việc lấy mẫu lên hình ảnh lớp tham chiếu có thể diễn ra. Ví dụ, độ sâu bit của các mẫu của hình ảnh lớp tham chiếu có thể được chuyển đổi thành độ sâu bit của lớp tăng cường và/hoặc các giá trị mẫu có thể trải qua ánh xạ từ không gian màu của lớp tham chiếu đến không gian màu của lớp tăng cường.

Mô hình mã hóa và/hoặc giải mã video mở rộng được có thể sử dụng việc mã hóa và/hoặc giải mã đa vòng lặp, có thể khác biệt như sau. Trong bước mã hóa/giải mã, hình ảnh lớp cơ sở có thể được khôi phục/giải để được sử dụng dưới dạng hình ảnh tham chiếu bù chuyển động cho các hình ảnh tiếp theo, theo thứ tự mã hóa/giải mã, trong cùng lớp hoặc dưới dạng tham chiếu cho dự đoán liên lớp (hoặc liên góc nhìn hoặc liên thành phần). Hình ảnh lớp cơ sở được khôi phục/giải mã có thể được lưu trữ trong DPB. Tương tự, hình ảnh lớp cơ sở được khôi phục/giải mã có thể được sử dụng làm hình ảnh tham chiếu bù chuyển động cho các hình ảnh tiếp theo, theo thứ tự mã hóa/giải mã, trong cùng lớp hoặc dưới dạng tham chiếu cho dự đoán liên lớp (hoặc liên góc nhìn hoặc liên thành phần) cho các lớp tăng cường cao hơn, nếu có. Ngoài các giá trị mẫu được khôi phục/giải mã, các giá trị thành phần cú pháp của lớp cơ sở/tham chiếu hoặc các biến thu được từ các giá trị thành phần cú pháp của lớp cơ sở/tham chiếu có thể được sử dụng trong dự đoán liên lớp/liên thành phần/liên góc nhìn.

Dự đoán liên lớp có thể được xác định là dự đoán theo cách mà độc lập về các thành phần dữ liệu (ví dụ, các giá trị mẫu hoặc các vector chuyển động) của các hình ảnh tham chiếu từ lớp khác với lớp của hình ảnh hiện thời (được mã hóa hoặc giải mã). Các kiểu dự đoán liên lớp tồn tại và có thể được áp dụng trong bộ mã hóa video/bộ giải mã mở rộng được. Các kiểu dự đoán liên lớp khả dụng có thể ví dụ phụ thuộc vào profile mã hóa theo đó dòng bit hoặc lớp cụ thể trong dòng bit được mã hóa hoặc, khi giải mã, profile mã hóa mà dòng bit hoặc lớp cụ thể trong dòng bit được chỉ báo để thích ứng với. Nói cách khác hoặc ngoài ra,

các kiểu dự đoán liên lớp khả dụng có thể phụ thuộc vào các kiểu khả năng mở rộng được hoặc kiểu bộ mã hóa-giải mã mở rộng được hoặc kiểu sửa đổi tiêu chuẩn mã hóa video (ví dụ SHVC, MV-HEVC, hoặc 3D-HEVC) được sử dụng.

Các kiểu dự đoán liên lớp có thể bao gồm, nhưng không giới hạn ở, một hoặc nhiều dự đoán sau: dự đoán mẫu liên lớp, dự đoán chuyển động liên lớp, dự đoán phần dư liên lớp. Trong dự đoán mẫu liên lớp, ít nhất tập con của các giá trị mẫu khôi phục của hình ảnh nguồn để dự đoán liên lớp được sử dụng dưới dạng tham chiếu để dự đoán các giá trị mẫu của hình ảnh hiện thời. Trong dự đoán chuyển động liên lớp, ít nhất tập con của các vectơ chuyển động của hình ảnh nguồn để dự đoán liên lớp được sử dụng dưới dạng tham chiếu để dự đoán các vectơ chuyển động của hình ảnh hiện thời. Thông thường, việc dự đoán thông tin về các hình ảnh tham chiếu nào được kết hợp với các vectơ chuyển động cũng được bao gồm trong dự đoán chuyển động liên lớp. Ví dụ, các chỉ số tham chiếu của các hình ảnh tham chiếu cho các vectơ chuyển động có thể được dự đoán liên lớp và/hoặc số đếm thứ tự hình ảnh hoặc bất kỳ nhận dạng nào khác của hình ảnh tham chiếu có thể được dự đoán liên lớp. Trong một số trường hợp, dự đoán chuyển động liên lớp cũng có thể bao gồm dự đoán về chế độ mã hóa khối, thông tin phần đầu, phân vùng khối, và/hoặc các tham số tương tự khác. Trong một số trường hợp, dự đoán tham số mã hóa, như dự đoán liên lớp của bước phân vùng khối, có thể được xem là một kiểu dự đoán liên lớp khác. Trong dự đoán phần dư liên lớp, lỗi dự đoán hoặc phần dư của các khối được chọn của hình ảnh nguồn để dự đoán liên lớp được sử dụng để dự đoán hình ảnh hiện thời. Trong bước mã hóa độ sâu đa góc nhìn, như 3D-HEVC, dự đoán liên lớp chéo thành phần có thể được áp dụng, trong đó hình ảnh của kiểu thứ nhất, như hình ảnh độ sâu, có thể ảnh hưởng đến dự đoán liên lớp của hình ảnh thuộc kiểu thứ hai, như hình ảnh bố cục thông thường. Ví dụ, giá trị mẫu liên lớp bù chênh lệch và/hoặc dự đoán chuyển động có thể được áp dụng, trong đó sự chênh lệch có thể là ít nhất một phần được thu nhận từ hình ảnh độ sâu.

Lớp tham chiếu trực tiếp có thể được xác định là lớp có thể được sử dụng để dự đoán liên lớp của lớp khác mà lớp này là lớp tham chiếu trực tiếp. Lớp dự

đoán trực tiếp có thể được định nghĩa là lớp trong đó lớp khác là lớp tham chiếu trực tiếp. Lớp tham chiếu gián tiếp có thể được định nghĩa là lớp không phải lớp tham chiếu trực tiếp của lớp thứ hai nhưng là lớp tham chiếu trực tiếp của lớp thứ ba mà là lớp tham chiếu trực tiếp hoặc lớp tham chiếu gián tiếp của lớp tham chiếu trực tiếp của lớp thứ hai trong đó lớp này là lớp tham chiếu gián tiếp. Lớp dự đoán gián tiếp có thể được định nghĩa là lớp trong đó lớp khác là lớp tham chiếu gián tiếp. Lớp độc lập có thể được xác định là lớp mà không có các lớp tham chiếu trực tiếp. Nói cách khác, lớp độc lập không được dự đoán sử dụng dự đoán liên lớp. Lớp phi cơ sở có thể được định nghĩa là bất kỳ lớp nào khác ngoài lớp cơ sở, và lớp cơ sở có thể được xác định là lớp thấp nhất trong dòng bit. Lớp phi cơ sở độc lập có thể được xác định là lớp vừa là lớp độc lập và lớp phi cơ sở.

Hình ảnh nguồn để dự đoán liên lớp có thể được định nghĩa là hình ảnh được giải mã mà hoặc được sử dụng trong việc thu nhận, hình ảnh tham chiếu liên lớp có thể được sử dụng làm hình ảnh tham chiếu để dự đoán hình ảnh hiện thời. Trong các phần mở rộng HEVC đa lớp, hình ảnh tham chiếu liên lớp được bao gồm trong tập hợp hình ảnh tham chiếu liên lớp của hình ảnh hiện thời. Hình ảnh tham chiếu liên lớp có thể được định nghĩa là hình ảnh tham chiếu có thể được sử dụng để dự đoán liên lớp của hình ảnh hiện thời. Trong quy trình mã hóa và/hoặc giải mã, các hình ảnh tham chiếu liên lớp có thể được xử lý dưới dạng các hình ảnh tham chiếu dài hạn.

Hình ảnh nguồn để dự đoán liên lớp có thể được yêu cầu là trong cùng đơn vị truy cập dưới dạng hình ảnh hiện thời. Trong một số trường hợp, ví dụ khi không có việc 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 để thích ứng với lưới lấy mẫu của lớp tham chiếu đến lưới lấy mẫu của lớp của hình ảnh hiện thời (được mã hóa hoặc giải mã), việc xử lý liên lớp được áp dụng để thu nhận hình ảnh tham chiếu liên lớp từ hình ảnh nguồn để dự đoán liên lớp. Các ví dụ về việc xử lý liên lớp này được mô tả trong các đoạn sau.

Dự đoán mẫu liên lớp có thể bao gồm việc lấy mẫu lại (các) mảng mẫu của hình ảnh nguồn để dự đoán liên lớp. Bộ mã hóa và/hoặc bộ giải mã có thể thu nhận hệ số tỉ lệ theo hướng chiều ngang (ví dụ được lưu trữ trong biến ScaleFactorX) và hệ số tỉ lệ theo hướng chiều dọc (ví dụ được lưu trữ trong biến ScaleFactorY) đối với cặp của lớp tăng cường và lớp tham chiếu của nó ví dụ dựa trên các khoảng dịch vị trí lớp tham chiếu đối với cặp này. Nếu một trong số hoặc cả hai hệ số tỉ 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 chỉ báo bởi bộ mã hóa trong dòng bit (ví dụ dưới dạng chỉ số trong số các quy trình hoặc bộ lọc lấy mẫu lại định trước) và/hoặc được giải mã bởi bộ giải mã từ dòng bit. Quy trình lấy mẫu lại khác có thể được chỉ báo bởi bộ mã hóa và/hoặc được giải mã bởi bộ giải mã và/hoặc được suy ra bằng bộ mã hóa và/hoặc bộ giải mã phụ thuộc vào các giá trị của hệ số tỉ lệ. Ví dụ, khi cả hai hệ số tỉ lệ nhỏ hơn 1, quy trình giảm tốc độ lấy mẫu định trước có thể được suy ra; và khi cả hai hệ số tỉ lệ đều lớn hơn 1, quy trình lấy mẫu lên định trước có thể được suy ra. Ngoài ra hoặc theo cách khác, quy trình lấy mẫu lại khác có thể được chỉ báo bởi bộ mã hóa và/hoặc được giải mã bởi bộ giải mã và/hoặc được suy ra bằng bộ mã hóa và/hoặc bộ giải mã phụ thuộc vào mảng mẫu nào được xử lý. Ví dụ, quy trình lấy mẫu lại thứ nhất có thể được suy ra được sử dụng cho các mảng mẫu luma và quy trình lấy mẫu lại thứ hai có thể được suy ra được sử dụng cho các mảng chroma.

SHVC cho phép việc sử dụng dự đoán có trọng số hoặc quy trình ánh xạ màu dựa trên bảng dò tìm 3D (LUT) đối với (nhưng không giới hạn ở) khả năng mở rộng được dải màu. Phương pháp 3D LUT có thể được mô tả như sau. Vùng giá trị mẫu của mỗi thành phần màu có thể được chia thành hai dải trước tiên, tạo thành các cung phần tám $2 \times 2 \times 2$, và sau đó các dải luma cũng có thể chia thành bốn phần, dẫn đến các cung phần tám $8 \times 2 \times 2$. Trong mỗi cung phần tám, mô hình tuyến tính của thành phần màu chéo được áp dụng để thực hiện ánh xạ màu. Đối với mỗi cung phần tám, bốn đỉnh được mã hóa thành và/hoặc được giải mã từ dòng bit để biểu diễn mô hình tuyến tính trong cung phần tám. Bảng ánh xạ màu

được mã hóa thành và/hoặc được giải mã từ dòng bit một cách riêng biệt cho mỗi thành phần màu. Ánh xạ màu có thể được xem xét để liên quan đến ba bước: trước hết, cung phần tám trong đó bộ ba lớp tham chiếu mẫu đã cho (Y, Cb, Cr) thuộc về đó được xác định. Thứ hai, các vị trí mẫu luma và chroma có thể được căn thẳng thông qua áp dụng quy trình điều chỉnh thành phần màu. Thứ ba, ánh xạ tuyến tính được xác định cho cung phần tám xác định được áp dụng. Ánh xạ này có thể có bản chất thành phần chéo, tức là giá trị đầu vào của một thành phần màu có thể ảnh hưởng đến giá trị được ánh xạ của thành phần màu khác. Ngoài ra, nếu việc lấy mẫu lại liên lớp cũng được yêu cầu, đầu vào với quy trình lấy mẫu lại là hình ảnh được ánh xạ màu. Ánh xạ màu có thể (nhưng không cần thiết để) 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 nhận biết như sau. Quy trình dự đoán vector chuyển động theo thời gian, như TMVP của H.265/HEVC, có thể được sử dụng để tận dụng phần dư 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 lấy mẫu lên, 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 sử dụng dự đoán vector chuyển động từ hình ảnh lớp cơ sở ví dụ với cơ chế dự đoán vector chuyển động theo không gian như TMVP của H.265/HEVC, bộ dự đoán vector chuyển động tương ứng được bắt nguồn từ trường chuyển động lớp cơ sở được ánh xạ. Cách này là tương quan giữa dữ liệu chuyển động của các lớp khác có thể được tận dụng để tăng cường hiệu suất 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 bằng cách thiết lập hình ảnh tham chiếu liên lớp dưới dạng hình ảnh tham chiếu đồng vị trí cho việc thu nhận TMVP.

Tương tự với MVC, trong MV-HEVC, hình ảnh tham chiếu liên góc nhìn cá có thể được bao gồm trong (các) danh sách hình ảnh tham chiếu của hình ảnh hiện thời được mã hóa hoặc giải mã. SHVC sử dụng hoạt động giải mã đa vòng lặp (không giống phần mở rộng SVC của H.264/AVC). SHVC có thể được xem xét để sử dụng phương pháp dựa vào chỉ số tham chiếu, tức là hình ảnh tham

chiếu liên lớp có thể được bao gồm trong một hoặc nhiều danh sách hình ảnh tham chiếu của hình ảnh hiện thời được mã hóa hoặc giải mã (như được mô tả ở trên).

Đối với việc 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, sử dụng dữ liệu được mã hóa sẵn (bao gồm các mẫu hình ảnh khôi phục và các tham số chuyển động còn được gọi là thông tin chuyển động) trong lớp tham chiếu để mã hóa hiệu quả lớp tăng cường, có thể được tích hợp với SHVC, MV-HEVC, và/hoặc bộ mã hóa-giải mã tương tự.

Tiêu chuẩn hoặc hệ thống mã hóa có thể đề cập đến thuật ngữ điểm hoạt động hoặc tương tự, có thể chỉ báo các lớp và/hoặc lớp con mở rộng được trong đó việc giải mã hoạt động và/hoặc có thể được kết hợp với dòng bit con bao gồm các lớp và/hoặc lớp con mở rộng được được giải mã. Theo HEVC, điểm hoạt động được xác định làm dòng bit được tạo ra từ dòng bit khác bằng hoạt động của quy trình trích xuất dòng bit phụ bằng dòng bit khác, TemporalId đích cao nhất, và danh sách mã nhận dạng lớp đích làm các đầu vào.

VPS của HEVC quy định các tập hợp lớp và các tham số HRD đối với các tập hợp lớp này. Tập hợp lớp có thể được sử dụng làm danh sách mã nhận dạng lớp đích trong quy trình trích xuất dòng bit. Theo HEVC, tập hợp lớp có thể được xác định là tập hợp các lớp được biểu diễn trong dòng bit được tạo ra từ dòng bit khác bằng hoạt động của quy trình trích xuất dòng bit với dòng bit khác, TemporalId đích cao nhất bằng 6, và danh sách mã nhận dạng lớp đích bằng danh sách mã nhận dạng lớp được kết hợp với tập hợp lớp làm các đầu vào.

Lớp đầu ra có thể được định nghĩa là lớp mà các hình ảnh của nó được giải mã được kết xuất với quy trình giải mã. Các lớp đầu ra có thể phụ thuộc vào tập con nào của dòng bit đa lớp được giải mã. Các hình ảnh được xuất ra bởi quy trình giải mã cũng có thể được xử lý, ví dụ sự chuyển đổi không gian màu từ không gian màu YUV đến RGB có thể được thực hiện, và chúng có thể được hiển

thị. Tuy nhiên, việc xử lý và/hoặc hiển thị khác có thể được xem xét là các quy trình bên ngoài bộ giải mã và/hoặc quy trình giải mã và có thể không diễn ra.

Trong các dòng bit video đa lớp, việc xác định điểm hoạt động có thể bao gồm việc xem xét tập hợp lớp đầu ra đích. Ví dụ, điểm hoạt động có thể được định nghĩa là dòng bit được tạo ra từ dòng bit khác với hoạt động của quy trình trích xuất dòng bit với dòng bit khác, lớp con tạm theo thời gian cao nhất (ví dụ TemporalId đích cao nhất), và danh sách mã nhận dạng lớp đích dưới dạng các đầu vào, và được kết hợp với tập hợp các lớp đầu ra. Nói cách khác, thuật ngữ khác, như điểm hoạt động đầu ra, có thể được sử dụng khi tham chiếu đến điểm hoạt động và tập hợp các lớp đầu ra được kết hợp. Ví dụ, trong MV-HEVC/SHVC, điểm hoạt động đầu ra có thể được định nghĩa là dòng bit được tạo ra từ dòng bit đầu vào bằng hoạt động của quy trình trích xuất dòng bit phụ với dòng bit đầu vào, TemporalId đích cao nhất, và danh sách mã nhận dạng lớp đích dưới dạng các đầu vào, và được kết hợp với tập hợp các lớp đầu ra.

Tập hợp lớp đầu ra (OLS) có thể được xác định là tập hợp của các lớp bao gồm các lớp của một trong số các tập hợp lớp được xác định, trong đó một hoặc nhiều lớp trong tập hợp các lớp được chỉ báo là các lớp đầu ra. Lớp đầu ra có thể được định nghĩa là lớp của tập hợp lớp đầu ra mà là đầu ra khi bộ giải mã và/hoặc HRD hoạt động sử dụng tập hợp lớp đầu ra dưới dạng tập hợp lớp đầu ra đích. Trong MV-HEVC/SHVC, biến TargetOlsIdx có thể quy định tập hợp lớp đầu ra nào là tập hợp lớp đầu ra đích bằng cách thiết lập TargetOlsIdx bằng chỉ số của tập hợp lớp đầu ra mà là tập hợp lớp đầu ra đích. Tập hợp lớp đầu ra đích có thể được định nghĩa là tập hợp lớp đầu ra trong đó chỉ số bằng TargetOlsIdx. TargetOlsIdx có thể được thiết lập ví dụ bằng HRD và/hoặc có thể được thiết lập bằng các phương tiện bên ngoài, ví dụ bằng bộ phát hoặc tương tự thông qua giao diện được cấp bởi bộ giải mã. Trong MV-HEVC/SHVC, lớp đầu ra có thể được định nghĩa là lớp của tập hợp lớp đầu ra được xuất ra khi TargetOlsIdx bằng chỉ số của tập hợp lớp đầu ra.

Bộ gửi, công, máy khách, hoặc tương tự có thể chọn các lớp và/hoặc các lớp con được truyền của dòng bit video mở rộng được. Các thuật ngữ trích xuất

lớp, trích xuất của các lớp, hoặc chuyển lớp xuống có thể đề cập đến việc truyền số lớp ít hơn số lớp khả dụng trong dòng bit được nhận bằng bộ gửi, công, máy khách, hoặc tương tự. Việc chuyển lên có thể đề cập đến việc truyền (các) lớp bổ sung so với chúng được truyền trước khi chuyển lớp lên bằng bộ gửi, công, máy khách, hoặc tương tự, tức là bắt đầu lại việc truyền một hoặc nhiều lớp mà việc truyền đó bị ngừng trước đó trong việc chuyển lớp xuống. Tương tự với việc chuyển lớp xuống và/hoặc việc chuyển lớp lên, bộ gửi, công, máy khách, hoặc tương tự có thể thực hiện việc chuyển lớp xuống và/hoặc chuyển lớp lên đối với các lớp con theo thời gian. Bộ gửi, công, máy khách, hoặc tương tự cũng có thể thực hiện việc chuyển lớp xuống và/hoặc chuyển đối lên đối với cả lớp và lớp con. Việc chuyển lớp xuống và/hoặc chuyển đối lên đối với lớp và lớp con có thể được thực hiện trong cùng đơn vị truy cập hoặc tương tự (tức là ảo một cách đồng thời) hoặc có thể được thực hiện trong các đơn vị truy cập khác hoặc tương tự (tức là ảo ở các thời điểm cụ thể).

Trong khi tập hợp các lớp đầu ra không đổi thích hợp các trường hợp sử dụng và các dòng bit trong đó lớp cao nhất giữ không đổi trong mỗi đơn vị truy cập, chúng không thể hỗ trợ các trường hợp sử dụng trong đó lớp cao nhất thay đổi từ một đơn vị truy cập này thành một đơn vị truy cập khác. Do đó sáng chế đề xuất các bộ mã hóa có thể quy định việc sử dụng các lớp đầu ra thay thế trong dòng bit và đáp lại việc sử dụng được xác định của các lớp đầu ra thay thế, bộ giải mã xuất ra hình ảnh được giải mã từ lớp đầu ra thay thế khi không có hình ảnh trong lớp đầu ra trong cùng đơn vị truy cập. Nhiều khả năng chỉ ra cách để chỉ báo các lớp đầu ra thay thế. Ví dụ, như được xác định theo HEVC, cơ chế tập hợp lớp đầu ra thay thế có thể bị ràng buộc chỉ được sử dụng cho các tập hợp lớp đầu ra chứa chỉ một lớp đầu ra, và cờ theo thiết lập lớp đầu ra (`alt_output_layer_flag[olsIdx]` theo HEVC) có thể được sử dụng để quy định rằng bất kỳ lớp tham chiếu trực tiếp hoặc gián tiếp nào của lớp đầu ra có thể dùng làm lớp đầu ra thay thế cho lớp đầu ra của tập hợp lớp đầu ra. Khi nhiều hơn một lớp đầu ra thay thế được cho phép sử dụng, có thể xác định được rằng hình ảnh tham chiếu liên lớp trực tiếp hoặc gián tiếp thứ nhất có mặt trong đơn vị truy cập

theo thứ tự mã nhận dạng lớp đi xuống xuống đến lớp đầu ra thay thế tối thiểu được chỉ báo được xuất ra.

Fig.9 là biểu diễn sơ đồ của hệ thống truyền thông đa phương tiện làm ví dụ trong đó các phương án khác nhau có thể được triển khai. Nguồn dữ liệu 1510 cung cấp tín hiệu nguồn trong định dạng tương tự, định dạng số không nén, hoặc định dạng số nén, hoặc bất kỳ kết hợp nào của các định dạng này. Bộ mã hóa 1520 có thể bao gồm hoặc được kết nối với việc xử lý trước, như chuyển đổi định dạng dữ liệu và/hoặc lọc tín hiệu nguồn. Bộ mã hóa 1520 mã hóa tín hiệu nguồn thành dòng bit đa phương tiện được mã hóa. Cần lưu ý rằng dòng bit cần được giải mã có thể được nhận trực tiếp hoặc gián tiếp từ thiết bị từ xa được định vị ảo bên trong bất kỳ kiểu mạng nào. Ngoài ra, dòng bit này có thể được nhận từ phần cứng hoặc phần mềm cục bộ. Bộ mã hóa 1520 có thể mã hóa nhiều hơn một kiểu đa phương tiện, như audio và video, hoặc nhiều hơn một bộ mã hóa 1520 có thể được yêu cầu để mã hóa các kiểu phương tiện khác của tín hiệu nguồn. Bộ mã hóa 1520 cũng có thể nhận đầu vào tạo ra một cách tổng hợp, như đồ họa và văn bản, hoặc nó có thể có khả năng tạo các dòng bit được mã hóa của các đa phương tiện tổng hợp. Sau đây, chỉ xử lý một dòng bit đa phương tiện được mã hóa của một kiểu đa phương tiện được xem xét để làm đơn giản hóa mô tả. Lưu ý rằng, tuy nhiên, các dịch vụ quảng bá thời gian thực thường bao gồm các dòng (thường là ít nhất một dòng trong số audio, video và phụ đề văn bản). Cũng cần lưu ý rằng hệ thống này có thể bao gồm nhiều bộ mã hóa, nhưng trong hình vẽ chỉ có một bộ mã hóa 1520 được biểu diễn để làm đơn giản phần mô tả mà không làm mất tính tổng quát. Cũng cần hiểu rằng, mặc dù văn bản và các ví dụ có trong bản mô tả có thể mô tả cụ thể quy trình mã hóa, người có hiểu biết trung bình trong lĩnh vực kỹ thuật này hiểu được rằng cùng các khái niệm và nguyên lý cũng áp dụng cho quy trình giải mã tương tự và ngược lại.

Dòng bit đa phương tiện được mã hóa có thể được truyền đến bộ lưu trữ 1530. Bộ lưu trữ 1530 có thể bao gồm bất kỳ kiểu bộ nhớ lớn nào để lưu trữ dòng bit đa phương tiện được mã hóa. Định dạng của dòng bit đa phương tiện được mã hóa trong bộ lưu trữ 1530 có thể là định dạng dòng bit độc lập sơ cấp, hoặc một

hoặc nhiều dòng bit đa phương tiện được mã hóa có thể được bao gói trong tệp tin chứa, hoặc dòng bit đa phương tiện được mã hóa có thể được gói trong định dạng đoạn thích hợp cho DASH (hoặc hệ thống tạo dòng tương tự) và được lưu trữ dưới dạng chuỗi của các đoạn. Nếu một hoặc nhiều dòng bit đa phương tiện được gói trong tệp tin chứa, bộ tạo tệp tin (không được thể hiện trên hình vẽ) có thể được sử dụng để lưu trữ nhiều hơn một dòng bit đa phương tiện trong tệp tin và tạo ra siêu dữ liệu định dạng tệp tin, cũng có thể được lưu trữ trong tệp tin. Bộ mã hóa 1520 hoặc bộ lưu trữ 1530 có thể bao gồm bộ tạo tệp tin, hoặc bộ tạo tệp tin được gắn theo cách phối hợp bộ mã hóa 1520 hoặc bộ lưu trữ 1530. Một số hệ thống vận hành “trực tiếp”, tức là bỏ qua lưu trữ và truyền dòng bit đa phương tiện được mã hóa từ bộ mã hóa 1520 một cách trực tiếp đến bộ gửi 1540. Sau đó, dòng bit đa phương tiện được mã hóa có thể được truyền đến bộ gửi 1540, cũng được gọi là máy chủ, trên cơ sở cần thiết. Định dạng được sử dụng trong việc truyền có thể là định dạng dòng bit độc lập sơ cấp, định dạng dòng gói, định dạng đoạn thích hợp cho DASH (hoặc hệ thống tạo dòng tương tự), hoặc một hoặc nhiều dòng bit đa phương tiện được mã hóa có thể được gói trong tệp tin chứa. Bộ mã hóa 1520, bộ lưu trữ 1530, và máy chủ 1540 có thể được bố trí trong cùng thiết bị vật lý hoặc chúng có thể được bao gồm trong các thiết bị riêng biệt. Bộ mã hóa 1520 và máy chủ 1540 có thể hoạt động với nội dung trực tiếp trong thời gian thực, trong đó trường hợp dòng bit đa phương tiện được mã hóa thường không được lưu trữ vĩnh viễn, nhưng tốt hơn là được đệm trong các khoảng thời gian nhỏ trong bộ mã hóa nội dung 1520 và/hoặc trong máy chủ 1540 để tạo thuận lợi cho các biến đổi trong làm chậm việc xử lý, làm chậm việc truyền, và tốc độ bit đa phương tiện được mã hóa.

Máy chủ 1540 gửi dòng bit đa phương tiện được mã hóa sử dụng chồng giao thức liên lạc. Chồng giao thức này có thể nhưng không giới hạn bởi Giao thức vận chuyển thời gian thực (RPT), Giao thức dữ liệu người dùng internet (UDP), và giao thức truyền siêu văn bản (HTTP), Giao thức điều khiển việc truyền (TCP), và Giao thức Internet (IP). Khi chồng giao thức truyền thông được định hướng gói, máy chủ 1540 bao gói dòng bit đa phương tiện được mã hóa thành các gói. Ví dụ, khi RTP được sử dụng, máy chủ 1540 bao gói dòng bit đa

phương tiện được mã hóa thành các gói RTP theo định dạng tải RTP. Thông thường, mỗi loại đa phương tiện có định dạng tải RTP chuyên dụng. Cần lưu ý thêm rằng hệ thống có thể chứa nhiều hơn một máy chủ 1540, nhưng để đơn giản, phần mô tả sau chỉ đề cập đến một máy chủ 1540.

Nếu nội dung đa phương tiện được bao gói trong tệp tin chứa cho bộ lưu trữ 1530 hoặc làm đầu vào dữ liệu cho bộ gửi 1540, bộ gửi 1540 có thể bao gồm hoặc có thể chứa hoặc được gắn một cách vận hành và “bộ phân tích tệp tin gửi” (không được thể hiện trên hình vẽ). Cụ thể, nếu tệp tin chứa không được truyền như vậy mà ít nhất một trong số các dòng bit đa phương tiện được mã hóa được chứa được bao gói để vận chuyển qua giao thức truyền thông, thì bộ phân tích tệp tin gửi bố trí các phần phù hợp của dòng bit đa phương tiện được mã hóa để vận chuyển qua giao thức truyền thông. Bộ phân tích tệp tin gửi cũng có thể giúp tạo định dạng chính xác đối với giao thức truyền thông, như là các phần đầu và các phần tải của gói. Tệp tin chứa đa phương tiện có thể chứa các lệnh đóng gói, như là rãnh gợi ý (hint track) trong ISOBMFF, để đóng gói ít nhất một trong số dòng bit đa phương tiện được chứa theo giao thức truyền thông.

Máy chủ 1540 có thể hoặc không thể được kết nối với cổng 1550 thông qua mạng truyền thông, có thể ví dụ là kết hợp của CDN, Internet và/hoặc một hoặc nhiều mạng truy cập. Cổng vốn cũng có thể hoặc thay thế tham chiếu đến như là hộp trung gian. Đối với DASH, cổng có thể là máy chủ cạnh (của CDN) hoặc ủy nhiệm web. Lưu ý rằng hệ thống này có thể thường bao gồm số lượng cổng bất kỳ hoặc tương tự, nhưng để đơn giản, mô tả sau đây chỉ xét một cổng 1550. Cổng 1550 có thể thực hiện các kiểu chức năng khác nhau, như dịch dòng theo gói từ chồng giao thức truyền thông này đến chồng giao thức truyền thông khác, kết hợp và chia nhánh các dòng dữ liệu, và thao tác dòng dữ liệu theo liên kết xuống và/hoặc các khả năng của bộ thu, như điều khiển tốc độ bit của dòng được chuyển tiếp theo các điều kiện mạng liên kết xuống đang thịnh hành.

Hệ thống gồm một hoặc nhiều bộ thu 1560, thường có khả năng nhận, giải điều biết, và/hoặc khử đóng gói tín hiệu được truyền thành dòng bit đa phương tiện được mã hóa. Dòng bit đa phương tiện được mã hóa có thể được truyền đến

vật ghi 1570. Vật ghi 1570 có thể bao gồm bất kỳ kiểu bộ nhớ lớn nào để lưu trữ dòng bit đa phương tiện được mã hóa. Vật ghi 1570 có thể bao gồm ngoài ra hoặc theo cách khác bộ nhớ điện toán, như bộ nhớ truy cập ngẫu nhiên. Định dạng của dòng bit đa phương tiện được mã hóa trong vật ghi 1570 có thể là định dạng dòng bit độc lập sơ cấp, hoặc một hoặc nhiều dòng bit đa phương tiện được mã hóa có thể được gói trong tệp tin chứa. Nếu có các dòng bit đa phương tiện được mã hóa, như dòng audio và dòng video, được kết hợp với nhau, tệp tin chứa thường được sử dụng và bộ thu 1560 bao gồm hoặc được gắn với bộ tạo tệp tin chứa mà tạo ra tệp tin chứa từ các dòng đầu vào. Một số hệ thống vận hành “trực tiếp”, tức là bỏ qua vật ghi 1570 và truyền dòng bit đa phương tiện được mã hóa từ bộ thu 1560 một cách trực tiếp đến bộ giải mã 1580. Trong một số hệ thống, chỉ một phần của dòng được mã hóa gần nhất ví dụ việc trích dòng được ghi 10 phút gần đây, được duy trì trong vật ghi 1570, trong khi bất kỳ dữ liệu được ghi sớm hơn đều bị loại bỏ khỏi vật ghi 1570.

Dòng bit đa phương tiện được mã hóa được truyền từ vật ghi 1570 đến bộ giải mã 1580. Nếu có nhiều dòng bit đa phương tiện, như là dòng âm thanh và dòng video, được kết hợp với nhau và được bao gói trong tệp tin chứa hoặc một dòng bit đa phương tiện được bao gói trong tệp tin chứa ví dụ để truy cập dễ dàng hơn, bộ phân tích tệp tin (không được thể hiện trên hình vẽ) được sử dụng để khử bao gói dòng bit đa phương tiện được mã hóa từ tệp tin chứa. Vật ghi 1570 hoặc bộ giải mã 1580 có thể bao gồm bộ phân tích tệp tin, hoặc bộ phân tích tệp tin được gắn vào vật ghi 1570 hoặc bộ giải mã 1580. Cũng cần lưu ý rằng hệ thống này có thể bao gồm nhiều bộ giải mã, nhưng ở đây chỉ có bộ giải mã 1570 được thảo luận để làm đơn giản phần mô tả mà không làm mất tính tổng quát.

Dòng bit đa phương tiện được mã hóa thường được tiếp tục xử lý bởi bộ giải mã 1570, đầu ra của chúng là một hoặc nhiều dòng đa phương tiện không nén. Cuối cùng, bộ phận tạo 1590 có thể mô phỏng các dòng đa phương tiện không nén bằng, ví dụ, loa hoặc màn hình hiển thị, chẳng hạn. Bộ thu 1560, vật ghi 1570, bộ giải mã 1570, và bộ phận tạo 1590 có thể nằm trong cùng một thiết bị vật lý hoặc chúng có thể được bao gồm trong các thiết bị riêng biệt.

Bộ gửi 1540 và/hoặc cổng 1550 có thể được tạo cấu hình để thực hiện việc truyền giữa các biểu diễn khác ví dụ để chuyển đổi góc nhìn, làm thích ứng tốc độ bit và/hoặc khởi động nhanh, và/hoặc bộ gửi 1540 và/hoặc cổng 1550 có thể được tạo cấu hình để chọn (các) dạng biểu diễn được truyền. Việc truyền giữa các biểu diễn khác có thể diễn ra vì nhiều lý do, như đáp lại các yêu cầu của bộ thu 1560 hoặc các điều kiện chiếm ưu thế, như thông lượng, của mạng trên đó dòng bit này được truyền. Yêu cầu từ bộ thu có thể là, ví dụ, yêu cầu cho đoạn hoặc đoạn con từ biểu diễn khác với biểu diễn trước đó, yêu cầu đối với sự thay đổi của các lớp và/hoặc các lớp con có khả năng mở rộng được được truyền, hoặc sự thay đổi của bộ tạo có các khả năng khác nhau so với thành phần trước đó. Yêu cầu đối với đoạn có thể là yêu cầu HTTP GET. Yêu cầu đối với đoạn con là yêu cầu HTTP GET với miền byte. Ngoài ra hoặc theo cách khác, điều chỉnh tốc độ bit hoặc thích ứng tốc độ bit có thể được sử dụng ví dụ để mang lại khởi động nhanh trong các dịch vụ tạo dòng, trong đó tốc độ bit của dòng bit được truyền thấp hơn tốc độ bit của kênh sau khi khởi động hoặc truy cập ngẫu nhiên dòng này để bắt đầu phát lại ngay lập tức và đạt được mức chiếm đệm để cho phép việc làm chậm gói ngẫu nhiên và/hoặc các lần truyền lại. Sự thích ứng tốc độ bit có thể bao gồm việc biểu diễn hoặc việc chuyển lớp lên và việc biểu diễn hoặc các hoạt động chuyển lớp xuống diễn ra theo các thứ tự khác nhau.

Bộ giải mã 1580 có thể được tạo cấu hình để thực hiện việc chuyển đổi giữa các biểu diễn khác nhau ví dụ để chuyển đổi góc nhìn, thích ứng tốc độ bit và/hoặc khởi động nhanh, và/hoặc bộ giải mã 1580 có thể được tạo cấu hình để chọn (các) biểu diễn được truyền. Việc chuyển đổi giữa các biểu diễn khác có thể diễn ra với các lý do, như đạt được hoạt động giải mã nhanh hơn hoặc để thích ứng dòng bit được truyền, ví dụ liên quan đến tốc độ bit, cho các điều kiện ưu thế, như thông lượng, của mạng trên đó dòng bit này được truyền. Hoạt động giải mã nhanh hơn có thể cần thiết ví dụ nếu thiết bị này bao gồm bộ giải mã 580 là đa nhiệm và sử dụng các tài nguyên tính toán cho các mục đích khác ngoài việc giải mã mở rộng được video dòng bit. Theo ví dụ khác, hoạt động giải mã nhanh hơn có thể cần thiết khi nội dung được phát lại với tốc độ nhanh hơn tốc độ phát lại thông thường, ví dụ nhanh hơn hai hoặc ba lần tốc độ phát lại thông thường trong

thời gian thực. Tốc độ hoạt động của bộ giải mã có thể được thay đổi trong khi giải mã hoặc phát lại ví dụ đáp lại lại sự thay đổi từ việc phát chuyển tệp nhanh từ tốc độ phát lại thông thường hoặc ngược lại, và vì vậy các hoạt động chuyển lớp lên và chuyển lớp xuống có thể diễn ra theo nhiều thứ tự khác nhau.

Hình ảnh khả năng mở rộng được không gian và/hoặc mã hóa video có thể được nhận biết bằng cách lấy mẫu lại hình ảnh lớp tham chiếu được khôi phục để dùng làm hình ảnh tham chiếu để mã hóa/giải mã (sau đây được gọi là mã hóa (giải mã)) hình ảnh lớp được dự đoán. Nói cách khác, mã hóa video mở rộng được theo không gian có thể được nhận biết bằng cách lấy mẫu lại các phần của hình ảnh lớp tham chiếu, như khối được mã hóa nội ảnh, để dùng làm khối dự đoán để mã hóa (giải mã) một phần của hình ảnh lớp dự đoán. Việc lấy mẫu lại có thể bao gồm hoạt động lọc trong đó số lượng các mẫu lớp tham chiếu được lọc để thu nhận mẫu được sử dụng dưới dạng tham chiếu để dự đoán liên lớp (ví dụ, mẫu của hình ảnh tham chiếu liên lớp trong SHVC). Do đó, việc lấy mẫu lại có thể truy cập ác vị trí mẫu bên ngoài các biên hình ảnh. Tương tự, nếu hình ảnh lớp dự đoán được gói khung, tức là, bao quanh các khung thành phần, việc lấy mẫu lại này có thể truy cập các vị trí mẫu bên ngoài các biên của khung thành phần. Hơn nữa, các khoảng dịch vị trí lớp tham chiếu hoặc tương tự có thể được sử dụng để chỉ báo sự tương ứng không gian của hình ảnh lớp dự đoán tương đối với hình ảnh lớp tham chiếu và có thể thực hiện quy trình lấy mẫu lại liên lớp để đề cập đến các mẫu bên ngoài các biên hình ảnh của hình ảnh lớp tham chiếu và/hoặc bên ngoài các biên của khung thành phần trong hình ảnh lớp tham chiếu. Cần hiểu rằng các phương án được mô tả ở trên áp dụng ngoài ra hoặc bổ sung dự đoán liên lớp. Ví dụ, phần mở rộng biên có thể được sử dụng cho quy trình lọc được sử dụng trong dự đoán liên lớp. Theo ví dụ khác, các vị trí mẫu được sử dụng trong quy trình lọc để dự đoán liên lớp được gói lại (ví dụ, trong trường hợp của hướng chiều ngang của khung thành phần 360 độ) hoặc (nếu không) được làm bão hòa ở trong các biên của khung thành phần.

Cần hiểu rằng các phương án được mô tả ở trên áp dụng ngoài ra hoặc bổ sung cho dự đoán hình ảnh liên kết. Theo một ví dụ, hình ảnh thứ nhất của lớp

tham chiếu hoặc góc nhìn bao quanh nhiều hơn một khung thành phần, ví dụ biểu diễn hai góc nhìn của nội dung. Hình ảnh thứ hai của lớp hoặc góc nhìn được dự đoán bao quanh ít nhất một khung thành phần, ví dụ biểu diễn một góc nhìn khác của nội dung (ngoài các góc nhìn được biểu diễn bởi hình ảnh thứ nhất). Dự đoán liên góc nhìn diễn ra từ một hoặc nhiều khung thành phần của hình ảnh thứ nhất đến hình ảnh thứ hai. Vector riêng biệt (tương tự hoặc giống với vector chuyển động) được sử dụng trong dự đoán hình ảnh liên kết có thể đề cập đến khối dự đoán một phần ở bên ngoài khung thành phần hoặc có thể đề cập đến vị trí mẫu phân đoạn để yêu cầu các mẫu đầu vào từ nhiều hơn một khung thành phần. Các phương án có thể được sử dụng để bố trí phần mở rộng biên, hoặc làm bão hòa hoặc bao gói các vị trí mẫu trong khung thành phần, sao cho khối dự đoán liên góc nhìn bị giới hạn để khởi đầu từ các mẫu của chỉ một khung thành phần.

Ở phần trên, một số phương án được mô tả có tham chiếu đến thuật ngữ khung thành phần ở dạng số ít của nó. Cần hiểu rằng các phương án có thể được áp dụng cho nhiều khung thành phần, và số lượng các khung thành phần đó thường không bị giới hạn với các phương án này.

Ở phần trên, một số phương án được mô tả với tham chiếu đến thuật ngữ khung thành phần. Cần hiểu rằng khung thành phần có thể biểu diễn khung đầu vào hoặc trường đầu vào, trong đó trường là tập hợp của các hàng mẫu xen kẽ của khung và có thể được sử dụng dưới dạng đầu vào bộ mã hóa, khi tín hiệu nguồn được xen kẽ.

Ở phần trên, một số phương án được mô tả với tham chiếu đến thuật ngữ khung thành phần. Cần hiểu rằng các phương án thường áp dụng được cho bất kỳ các phần liên kề trong không gian nào của hình ảnh được mã hóa hoặc được giải mã ngay cả khi phần này không được gọi rõ ràng là khung thành phần. Cần hiểu rằng các phương án áp dụng tương tự với các đơn vị nhỏ hơn khung thành phần miễn là đơn vị này chiếm vùng không gian liên kề trong hình ảnh được mã hóa hoặc giải mã. Theo một phương án làm ví dụ, khung thành phần được phân vùng thành hai hoặc nhiều hơn hai phần tách khung thành phần rời nhau trong không gian trong bước trước xử lý trước khi mã hóa hoặc là một phần của bước mã hóa.

Một hoặc nhiều phân vùng của khung thành phần có phần mở rộng biên như được mô tả theo các phương án khác. Cùng một hình ảnh có thể bao gồm các biên của khung thành phần từ nhiều hơn một khung thành phần. Mỗi biên của khung thành phần, bao gồm hoặc loại trừ phần mở rộng biên, nếu có, có thể được mã hóa dưới dạng tập hợp tám bị ràng buộc chuyển động. Các phần mở rộng biên có thể được loại bỏ và các phân vùng khung thành phần có thể được bố trí theo cách liên tục trong không gian bởi bộ giải mã, ví dụ là một phần của quy trình xén đầu ra, hoặc bởi bước hậu xử lý sau khi giải mã, để các khung thành phần liên tục toàn bộ. Ví dụ, khung thành phần toàn cảnh 360 độ có thể được phân vùng ví dụ sử dụng lưới 45, 90, và 45 độ theo hướng chiều dọc và các khoảng lưới 90 độ theo hướng chiều ngang, và được xử lý như được mô tả theo phương án làm ví dụ. Theo một ví dụ khác, mỗi mặt lập phương của hình ảnh cầu được chiếu với ánh xạ lập phương có thể được xem là phân vùng khung thành phần và được xử lý như được mô tả theo phương án làm ví dụ.

Theo một phương án, các hình ảnh đầu vào để mã hóa video bắt nguồn từ ánh xạ hình chóp như được minh họa trên các hình vẽ từ Fig.10a đến Fig.10c. Mỗi hình ảnh đầu vào của chuỗi video cầu, chuỗi video biểu diễn phép chiếu video cầu, như toàn cảnh hình cầu, video đa góc nhìn của việc thiết lập camera hướng ra ngoài, hoặc tương tự được chiếu trên hình chóp 1000 được minh họa trên Fig.10a. Phép chiếu hình chóp cũng có thể được bỏ qua và ánh xạ trực tiếp từ nội dung nguồn đến phép chiếu hình chóp không gấp 1002, ví dụ như trên Fig.10b, hoặc đến phép chiếu hình chóp đóng gói khung 1004, ví dụ như trên Fig.10c, có thể được thực hiện. Mặt hình chữ nhật 1006 của hình chóp có thể được gọi là mặt trước, giả sử hướng quan sát ban đầu hướng chính xác hoặc gần hướng về phía mặt trước. Trục chiều ngang đối với mặt trước được chọn theo định hướng quan sát ban đầu (thích ứng với định hướng đầu giả định). Vị trí quan sát ảo trong hình chóp có thể được chọn theo cách mà mặt trước có trường quan sát mong muốn. Ví dụ, trường quan sát chiều ngang và chiều dọc 90 độ có thể được chọn cho mặt trước. Các mặt chóp hình tam giác có thể được gọi là các mặt trái 1008, mặt trên 1010, mặt phải 1012, và mặt đáy 1014 theo các vị trí của chúng so với hướng quan sát ban đầu được giả định và hướng quan sát ban đầu.

Các mặt chóp hình tam giác có thể được xem xét là không bị gập trên mặt phẳng của mặt trước 1006, như được minh họa trên Fig.10b, mặc dù bước trung gian này có thể được bỏ qua và ánh xạ trực tiếp từ hình chóp 1000 của Fig.10a đến phép chiếu hình chóp đóng gói khung 1006, ví dụ như trên Fig.10c, có thể được thực hiện. Dưới dạng một phần không bị gập, các mặt chóp hình tam giác có thể được lấy mẫu lại theo cách mà khi chúng được bố trí giống nhau hoặc tương tự với bố trí được minh họa trên Fig.10c, tổng chiều rộng và chiều cao của chúng khớp với chiều rộng và chiều cao của mặt trước. Các mặt hình tam giác có thể được bố trí trong hình chữ nhật biên theo các cách khác nhau, một trong số đó được minh họa trên Fig.10c, trong đó các mặt bên phải không gập 1012 và các mặt bên trái không gập 1008 được giữ không đổi, trong đó các mặt trên không gập 1010 và đáy không gập 1014 được quay 180 độ. Theo một phương án, mặt trước 1006 được xem là khung thành phần thứ nhất và hình chữ nhật biên bao gồm các mặt hình tam giác có thể được xem là khung thành phần thứ hai theo các phương án khác. Theo một phương án, mặt trước 1006 có phần mở rộng biên ở phía liền kề với hình chữ nhật biên của các mặt hình tam giác. Ví dụ, phần mở rộng biên có thể tập trung ở bên phải của mặt trước 1006, khi hình chữ nhật biên của các mặt hình tam giác cũng tập trung ở bên phải của mặt trước 1006, như được minh họa trên Fig.10c. Phần mở rộng biên được xử lý tương tự với phần mở rộng biên được mô tả theo các phương án khác. Trong khi Fig.10c minh họa phần mở rộng biên, cần hiểu rằng các phương án khác trong đó phần mở rộng biên không có mặt (và, ví dụ, các tham chiếu bên ngoài khung thành phần được xử lý bằng cách làm bão hòa các vị trí mẫu ở trong khung thành phần) cũng áp dụng được cho sự bố trí tương tự với bố trí trên Fig.10c nhưng không có phần mở rộng biên. Theo một phương án, mặt trước 1006, bao gồm phần mở rộng biên của nó (nếu có), được bao quanh trong tập hợp tám thứ nhất trong bước mã hóa và tập hợp tám thứ nhất được mã hóa dưới dạng tập hợp tám bị ràng buộc chuyển động. Theo một phương án, hình chữ nhật biên bao gồm các mặt hình tam giác 1008—1014 được bao quanh trong tập hợp tám thứ hai trong bước mã hóa và tập hợp tám thứ hai được mã hóa dưới dạng tập hợp tám bị ràng buộc chuyển động. (Các) tập hợp tám bị ràng buộc chuyển động được xử lý tương tự với tập hợp tám được mô tả theo các

phương án khác. Trong khi đoạn này được mô tả các phương án mã hóa, cần hiểu rằng các phương án giải mã có thể được nhận biết tương tự bằng cách xử lý mặt trước 1006 dưới dạng khung thành phần theo các phương án khác nhau để giải mã.

Ở phần trên, một số phương án được mô tả với tham chiếu đến thuật ngữ khối. Cần hiểu rằng các hệ thống mã hóa có thể có thuật ngữ khác đối với cùng khái niệm hoặc khái niệm tương tự mà các phương án đề cập đến. Ví dụ, theo HEVC thuật ngữ khối có thể được hiểu ví dụ là đơn vị cây mã hóa, khối cây mã hóa, đơn vị mã hóa, hoặc khối mã hóa theo các phương án khác.

Ở phần trên, một số phương án được mô tả liên quan đến HEVC và/hoặc các thuật ngữ được sử dụng trong đặc tả HEVC. Cần hiểu rằng các phương án áp dụng tương tự với các bộ mã hóa-giải mã khác và các định dạng mã hóa và thuật ngữ khác với sự tương đương hoặc tương tự với các thuật ngữ được sử dụng trong các phương án được mô tả ở trên.

Ở phần trên, trong đó các phương án làm ví dụ được mô tả với tham chiếu đến bộ mã hóa, cần hiểu rằng dòng bit thu được và bộ giải mã có thể có các thành phần tương ứng trong đó. Tương tự, trong đó các phương án làm ví dụ được mô tả với tham chiếu đến bộ giải mã, cần hiểu rằng bộ mã hóa có thể có cấu trúc và/hoặc chương trình máy tính để tạo ra dòng bit được giải mã bởi bộ giải mã.

Các phương án của sáng chế được mô tả ở trên mô tả bộ mã hóa-giải mã trong các thuật ngữ thiết bị mã hóa và giải mã riêng biệt để hỗ trợ việc hiểu các quy trình liên quan. Tuy nhiên, cần hiểu rằng thiết bị, các cấu trúc và các hoạt động có thể được triển khai dưới dạng một thiết bị/cấu trúc/hoạt động mã hóa-giải mã đơn. Hơn nữa, có thể bộ mã hóa và bộ giải mã có thể chia sẻ một vài hoặc tất cả các thành phần chung.

Mặc dù, các ví dụ trên mô tả các phương án của sáng chế sử dụng trong bộ mã hóa-giải mã trong thiết bị điện tử, cần hiểu rằng sáng chế như được xác định trong các yêu cầu bảo hộ có thể được triển khai dưới dạng một phần của bất kỳ bộ mã hóa-giải mã video nào. Do đó, ví dụ, các phương án của sáng chế có thể được

triển khai trong bộ mã hóa-giải mã video có thể thực hiện mã hóa video trên các đường truyền thông cố định hoặc có dây.

Do đó, thiết bị người dùng có thể bao gồm bộ mã hóa-giải mã video như được mô tả theo các phương án của sáng chế ở trên. Cần hiểu rằng thuật ngữ thiết bị người dùng được dự định bao gồm bất kỳ kiểu thiết bị người dùng không dây thích hợp nào, như các điện thoại di động, các thiết bị xử lý dữ liệu di động hoặc các trình duyệt web di động.

Ngoài ra các thành phần của mạng di động đất liền công cộng (public land mobile network - PLMN) cũng có thể 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 triển khai trong phần cứng hoặc các mạch chuyên dụng, phần mềm, logic hoặc bất kỳ kết hợp nào của chúng. Ví dụ, một số khía cạnh có thể được triển khai trong phần cứng, trong khi các khía cạnh khác có thể được triển khai trong phần sụn hoặc phần mềm có thể được chạy bởi bộ điều khiển, bộ vi xử lý hoặc thiết bị điện toán 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 dưới dạng các sơ đồ khối, các lưu đồ, hoặc sử dụng một số biểu diễn hình ảnh khác, cũng hiểu rằng các khối, thiết bị, các hệ thống, các kỹ thuật hoặc phương pháp này được mô tả trong bản mô tả này có thể được triển khai trong, dưới dạng các ví dụ không giới hạn, phần cứng, phần mềm, phần sụn, các mạch chuyên dụng hoặc logic, phần cứng đa năng hoặc bộ điều khiển hoặc các thiết bị điện toán khác, hoặc một số kết hợp của chúng.

Các phương án của sáng chế có thể được triển khai bằng phần mềm máy tính chạy được bởi bộ xử lý dữ liệu của thiết bị di động, như trong thực thể bộ xử lý, hoặc bằng phần cứng, hoặc bằng kết hợp của phần mềm và phần cứng. Ngoài ra cần lưu ý rằng bất kỳ khối nào của lưu đồ logic như trên các hình vẽ có thể biểu diễn các bước chương trình, hoặc các mạch logic kết nối, các khối và các chức năng, hoặc kết hợp của các bước chương trình và các mạch logic, các khối và các chức năng. Phần mềm có thể được lưu trữ trên các vật ghi vật lý dưới dạng các

chip bộ nhớ, hoặc các khối bộ nhớ được lắp đặt trong bộ xử lý, các vật ghi từ như đĩa cứng hoặc các đĩa mềm, và các vật ghi quang học như ví dụ DVD và các biến thể dữ liệu của chúng, CD.

Bộ nhớ có thể là bất kỳ kiểu nào thích hợp cho môi trường kỹ thuật cục bộ và có thể được triển khai sử dụng bất kỳ công nghệ lưu trữ dữ liệu thích hợp nào, như các thiết bị bộ nhớ dựa vào bán dẫn, các thiết bị và hệ thống bộ nhớ từ, các thiết bị và hệ thống bộ nhớ quang học, bộ nhớ cố định và bộ nhớ tháo rời được. Các bộ xử lý dữ liệu có thể là bất kỳ kiểu nào thích hợp cho môi trường kỹ thuật cục bộ, và có thể bao gồm một hoặc nhiều máy tính đa năng, máy tính chuyên dụng, bộ vi xử lý, bộ xử lý tín hiệu số (DSP) và các bộ xử lý dựa trên kiến trúc bộ xử lý đa lõi, dưới dạng các ví dụ không giới hạn.

Các phương án của sáng chế có thể được thực hiện trong các thành phần như các môđun mạch tích hợp. Thiết kế của các mạch tích hợp nói chung là quy trình tự động hóa cao. Các công cụ phần mềm phức tạp và mạnh mẽ là khả dụng để chuyển đổi thiết kế cấp độ logic thành thiết kế mạch bán dẫn được khắc và tạo ra trên đế bán dẫn.

Các chương trình, như được cung cấp bởi Synopsys, Inc. của Mountain View, California và Cadence Design, của San Jose, California được đính tuyến tự động đến các vật dẫn điện và đặt các thành phần trên chip bán dẫn sử dụng các quy tắc thiết kế được thiết lập cũng như các thư viện của các môđun thiết kế được lưu trữ trước. Ngay khi thiết kế đối với mạch bán dẫn được hoàn thành, thiết kế thu được, trong định dạng điện tử chuẩn hóa (ví dụ, Opus, GDSII, hoặc tương tự) có thể được truyền đến phương tiện sản xuất bán dẫn hoặc “fab” cho việc chế tạo.

Mô tả trước được đề xuất với cách làm ví dụ và các ví dụ không giới hạn mô tả đầy đủ và mang thông tin của các phương án làm ví dụ của sáng chế. Tuy nhiên, người có hiểu biết trung bình trong lĩnh vực kỹ thuật liên quan có thể hiểu được các biến đổi và tương thích khác nhau khi xem xét mô tả ở trên, khi đọc kết hợp với các hình vẽ kèm theo và các yêu cầu bảo hộ kèm theo. Tuy nhiên, tất cả các biến đổi này và biến đổi tương tự của việc hiểu sáng chế vẫn nằm trong phạm vi bảo hộ của các yêu cầu bảo hộ.

Sau đây một số ví dụ sẽ được đề xuất.

Theo ví dụ thứ nhất phương pháp được đề xuất bao gồm:

nhận hình ảnh được mã hóa;

giải mã hình ảnh được mã hóa để tạo ra khung thành phần được khôi phục của hình ảnh có vùng hình ảnh hiệu dụng; và

thực hiện một trong các bước sau:

làm đầy vùng bên ngoài vùng hình ảnh hiệu dụng để tạo ra hình ảnh tham chiếu được đệm, trong đó vùng được làm đầy tạo ra phần mở rộng biên; hoặc

xác định rằng khi tham chiếu đến các vị trí mẫu bên ngoài vùng hình ảnh hiệu dụng trong bước giải mã, các vị trí mẫu nêu trên được làm bão hòa hoặc được gói trong vùng hình ảnh hiệu dụng.

Theo một phương án của phương pháp bước làm đầy nêu trên bao gồm ít nhất một trong số bước dưới đây đối với ít nhất một mẫu bên ngoài vùng hình ảnh hiệu dụng:

giải mã ít nhất một giá trị mẫu bằng dự đoán nội ảnh theo chiều ngang;

giải mã ít nhất một giá trị mẫu bằng dự đoán nội ảnh theo chiều dọc;

giải mã ít nhất một giá trị mẫu bằng dự đoán sao chép nội khối từ vùng biên của vùng hình ảnh hiệu dụng đối diện với phần mở rộng biên;

giải mã ít nhất một giá trị mẫu bằng dự đoán sao chép nội khối từ khung thành phần khác trong hình ảnh;

giải mã ít nhất một giá trị mẫu bằng dự đoán liên ảnh từ hình ảnh trước đó theo thứ tự giải mã.

Theo một phương án phương pháp này còn bao gồm các bước:

giải mã chỉ báo liệu các vị trí mẫu bên ngoài các biên khung thành phần được làm bão hòa hoặc được gói; và

xác định theo chỉ báo được giải mã liệu các vị trí mẫu nêu trên được làm bão hòa hay được gói trong vùng hình ảnh hiệu dụng.

Theo một phương án phương pháp này còn bao gồm bước:

giải mã chỉ báo của vùng hình ảnh hiệu dụng của khung thành phần.

Theo một phương án phương pháp này còn bao gồm bước:

chèn ít nhất một giá trị mẫu bằng cách sao chép từ phía đối diện của vùng hình ảnh hiệu dụng khi khung thành phần được chỉ báo là hình ảnh toàn cảnh 360 độ.

Theo một phương án, phương pháp này bao gồm các bước:

giải mã hình ảnh được mã hóa để tạo ra khung thành phần được khôi phục thứ hai của hình ảnh.

Theo một phương án, phương pháp này bao gồm các bước:

nhận hình ảnh mã hóa thứ hai;

giải mã vectơ chuyển động từ hình ảnh mã hóa thứ hai, vectơ chuyển động khiến việc tham chiếu đến vị trí mẫu bên ngoài vùng hình ảnh hiệu dụng của khung thành phần được khôi phục trong bước giải mã; và

làm bão hòa hoặc gói vị trí mẫu ở trong vùng hình ảnh hiệu dụng.

Theo ví dụ thứ hai thiết bị được đề xuất bao gồm ít nhất một bộ xử lý và ít nhất một bộ nhớ, ít nhất một bộ nhớ nêu trên được lưu trữ với mã trên đó, khi được chạy với ít nhất một bộ xử lý nêu trên, khiến thiết bị thực hiện ít nhất:

nhận hình ảnh được mã hóa;

giải mã hình ảnh được mã hóa để tạo ra khung thành phần được khôi phục của hình ảnh có vùng hình ảnh hiệu dụng; và

thực hiện một trong các bước sau:

làm đầy vùng bên ngoài vùng hình ảnh hiệu dụng để tạo ra hình ảnh tham chiếu được đệm, trong đó vùng được làm đầy tạo ra phần mở rộng biên; hoặc

xác định rằng khi tham chiếu đến các vị trí mẫu bên ngoài vùng hình ảnh hiệu dụng trong bước giải mã, các vị trí mẫu nêu trên được làm bão hòa hoặc được gói trong vùng hình ảnh hiệu dụng.

Theo một phương án của thiết bị làm đầy nêu trên bao gồm mã làm cho thiết bị thực hiện ít nhất một trong số các bước:

giải mã ít nhất một giá trị mẫu bằng dự đoán nội ảnh theo chiều ngang;

giải mã ít nhất một giá trị mẫu bằng dự đoán nội ảnh theo chiều dọc;

giải mã ít nhất một giá trị mẫu bằng dự đoán sao chép nội khối từ vùng biên của vùng hình ảnh hiệu dụng đối diện với phần mở rộng biên;

giải mã ít nhất một giá trị mẫu bằng dự đoán sao chép nội khối từ khung thành phần khác trong hình ảnh;

giải mã ít nhất một giá trị mẫu bằng dự đoán liên ảnh từ hình ảnh trước đó theo thứ tự giải mã.

Theo một phương án thiết bị còn bao gồm mã làm cho thiết bị thực hiện ít nhất các bước:

giải mã chỉ báo liệu các vị trí mẫu bên ngoài các biên khung thành phần có được làm bão hòa hoặc được gói hay không; và

xác định theo chỉ báo được giải mã liệu các vị trí mẫu nêu trên có được làm bão hòa hoặc được gói trong vùng hình ảnh hiệu dụng hay không.

Theo một phương án thiết bị còn bao gồm mã làm cho thiết bị thực hiện ít nhất bước:

giải mã chỉ báo của vùng hình ảnh hiệu dụng của khung thành phần.

Theo một phương án thiết bị còn bao gồm mã làm cho thiết bị thực hiện ít nhất bước:

chèn ít nhất một giá trị mẫu bằng cách sao chép từ phía đối diện của vùng hình ảnh hiệu dụng khi khung thành phần được chỉ báo là hình ảnh toàn cảnh 360 độ.

Theo một phương án thiết bị còn bao gồm mã làm cho thiết bị thực hiện ít nhất bước:

giải mã vùng bên ngoài vùng hình ảnh hiệu dụng để thu khung tham chiếu trung gian; và

trong đó còn việc làm đầy bao gồm bước biến đổi ít nhất một giá trị mẫu của phần mở rộng biên của khung tham chiếu trung gian.

Theo một phương án thiết bị còn bao gồm mã làm cho thiết bị thực hiện ít nhất bước:

giải mã hình ảnh được mã hóa để tạo ra khung thành phần được khôi phục thứ hai của hình ảnh.

Theo một phương án thiết bị còn bao gồm mã làm cho thiết bị thực hiện ít nhất bước:

nhận hình ảnh mã hóa thứ hai;

giải mã vectơ chuyển động từ hình ảnh mã hóa thứ hai, vectơ chuyển động thực hiện tham chiếu đến vị trí mẫu bên ngoài vùng hình ảnh hiệu dụng của khung thành phần được khôi phục trong bước giải mã; và

làm bão hòa hoặc gói vị trí mẫu trong vùng hình ảnh hiệu dụng.

Theo ví dụ thứ ba, sáng chế đề xuất vật ghi đọc được bằng máy tính được lưu trữ với mã trên đó để sử dụng với thiết bị, khi được chạy bởi bộ xử lý, khiến thiết bị thực hiện:

nhận hình ảnh được mã hóa;

giải mã hình ảnh được mã hóa để tạo ra khung thành phần được khôi phục của hình ảnh có vùng hình ảnh hiệu dụng; và

thực hiện một trong các bước sau:

làm đầy vùng bên ngoài vùng hình ảnh hiệu dụng để tạo ra hình ảnh tham chiếu được đệm, trong đó vùng được làm đầy tạo ra phần mở rộng biên; hoặc

xác định rằng khi tham chiếu đến các vị trí mẫu bên ngoài vùng hình ảnh hiệu dụng trong bước giải mã, các vị trí mẫu nêu trên được làm bão hòa hoặc được gói trong vùng hình ảnh hiệu dụng.

Theo ví dụ thứ tư sáng chế đề xuất thiết bị bao gồm:

các phương tiện để nhận hình ảnh được mã hóa;

các phương tiện để giải mã hình ảnh được mã hóa để tạo ra khung thành phần được khôi phục của hình ảnh có vùng hình ảnh hiệu dụng; và

các phương tiện để thực hiện một trong các bước sau:

làm đầy vùng bên ngoài vùng hình ảnh hiệu dụng để tạo ra hình ảnh tham chiếu được đệm, trong đó vùng được làm đầy tạo ra phần mở rộng biên; hoặc

xác định rằng khi tham chiếu đến các vị trí mẫu bên ngoài vùng hình ảnh hiệu dụng trong bước giải mã, các vị trí mẫu nêu trên được làm bão hòa hoặc được gói ở trong vùng hình ảnh hiệu dụng.

Theo ví dụ thứ năm, sáng chế đề xuất phương pháp bao gồm:

mã hóa khung thành phần không nén thành hình ảnh thứ nhất được mã hóa, việc mã hóa nêu trên cũng dẫn đến hình ảnh khôi phục thứ nhất và khung thành phần nêu trên có vùng hình ảnh hiệu dụng trong hình ảnh khôi phục thứ nhất;

thực hiện một trong các bước sau dưới dạng một phần của bước mã hóa nêu trên:

chèn ít nhất một giá trị mẫu bên ngoài vùng hình ảnh hiệu dụng để tạo ra phần mở rộng biên đối với khung thành phần trong hình ảnh khôi phục thứ nhất; hoặc

làm bão hòa hoặc gói các vị trí mẫu bên ngoài vùng hình ảnh hiệu dụng ở trong vùng hình ảnh hiệu dụng.

Theo một phương án, phương pháp này còn bao gồm các bước:

thu ít nhất một giá trị mẫu bằng dự đoán nội ảnh theo chiều ngang;

thu ít nhất một giá trị mẫu bằng dự đoán nội ảnh theo chiều dọc;

thu ít nhất một giá trị mẫu bằng dự đoán sao chép nội khối từ biên của vùng hình ảnh hiệu dụng đối diện với phần mở rộng biên;

thu ít nhất một giá trị mẫu bằng dự đoán sao chép nội khối từ khung thành phần khác trong hình ảnh khôi phục;

thu ít nhất một giá trị mẫu bằng dự đoán liên ảnh từ hình ảnh trước đó theo thứ tự giải mã.

Theo một phương án phương pháp này còn bao gồm bước:

mã hóa chỉ báo liệu các vị trí mẫu bên ngoài các biên khung thành phần có được làm bão hòa hoặc được gói hay không.

Theo một phương án, phương pháp này còn bao gồm bước:

mã hóa chỉ báo về vùng hình ảnh hiệu dụng của khung thành phần.

Theo một phương án, phương pháp còn bao gồm bước:

chèn ít nhất một giá trị mẫu bằng cách sao chép từ phía đối diện của vùng hình ảnh hiệu dụng khi khung thành phần được chỉ báo là hình ảnh toàn cảnh 360 độ.

Theo một phương án của phương pháp mã hóa nêu trên bao gồm:

khôi phục vùng bên ngoài vùng hình ảnh hiệu dụng để thu khung tham chiếu trung gian; và

ngoài ra trong đó việc làm đầy bao gồm biến đổi ít nhất một giá trị mẫu của phần mở rộng biên của khung tham chiếu trung gian.

Theo một phương án phương pháp này bao gồm:

mã hóa khung thành phần thứ hai không nén thành hình ảnh thứ nhất được mã hóa.

Theo một phương án phương pháp còn bao gồm:

mã hóa hình ảnh mã hóa thứ hai;

bước mã hóa nêu trên của hình ảnh mã hóa thứ hai bao gồm mã hóa vector chuyển động thực hiện tham chiếu với vị trí mẫu bên ngoài vùng hình ảnh hiệu dụng của khung thành phần được khôi phục trong bước giải mã; và

làm bão hòa hoặc gói vị trí mẫu ở trong vùng hình ảnh hiệu dụng.

Theo ví dụ thứ sáu, sáng chế đề xuất thiết bị bao gồm ít nhất một bộ xử lý và ít nhất một bộ nhớ, ít nhất một bộ nhớ nêu trên được lưu trữ với mã trên đó, mã này khi được chạy với ít nhất một bộ xử lý nêu trên, khiến thiết bị thực hiện ít nhất:

mã hóa khung thành phần không nén thành hình ảnh thứ nhất được mã hóa, việc mã hóa nêu trên cũng dẫn đến hình ảnh khôi phục thứ nhất và khung thành phần nêu trên có vùng hình ảnh hiệu dụng trong hình ảnh khôi phục thứ nhất;

thực hiện một trong các bước sau như là một phần của bước mã hóa nêu trên:

chèn ít nhất một giá trị mẫu bên ngoài vùng hình ảnh hiệu dụng để tạo ra phần mở rộng biên đối với khung thành phần trong hình ảnh khôi phục thứ nhất; hoặc

làm bão hòa hoặc gói các vị trí mẫu bên ngoài vùng hình ảnh hiệu dụng ở trong vùng hình ảnh hiệu dụng.

Theo một phương án thiết bị bao gồm:

thu ít nhất một giá trị mẫu bằng dự đoán nội ảnh theo chiều ngang;

thu ít nhất một giá trị mẫu bằng dự đoán nội ảnh theo chiều dọc;

thu ít nhất một giá trị mẫu bằng dự đoán sao chép nội khối từ biên của vùng hình ảnh hiệu dụng đối diện với phần mở rộng biên;

thu ít nhất một giá trị mẫu bằng dự đoán sao chép nội khối từ khung thành phần khác trong hình ảnh khôi phục;

thu ít nhất một giá trị mẫu bằng dự đoán liên ảnh từ hình ảnh trước đó theo thứ tự giải mã.

Theo một phương án thiết bị còn bao gồm mã làm cho thiết bị thực hiện ít nhất một trong số:

mã hóa chỉ báo liệu các vị trí mẫu bên ngoài các biên khung thành phần có được làm bão hòa hoặc được gói hay không.

Theo một phương án thiết bị còn bao gồm mã làm cho thiết bị thực hiện ít nhất bước:

mã hóa chỉ báo về vùng hình ảnh hiệu dụng của khung thành phần.

Theo một phương án thiết bị còn bao gồm mã làm cho thiết bị thực hiện ít nhất:

chèn ít nhất một giá trị mẫu bằng cách sao chép từ phía đối diện của vùng hình ảnh hiệu dụng khi khung thành phần được chỉ báo là hình ảnh toàn cảnh 360 độ.

Theo một phương án, thiết bị mã hóa nêu trên bao gồm mã làm cho thiết bị thực hiện ít nhất:

tái cấu trúc vùng bên ngoài vùng hình ảnh hiệu dụng để thu khung tham chiếu trung gian; và

ngoài ra trong đó việc làm đầy bao gồm biến đổi ít nhất một giá trị mẫu của phần mở rộng biên của khung tham chiếu trung gian.

Theo một phương án thiết bị còn bao gồm mã làm cho thiết bị thực hiện ít nhất:

mã hóa khung thành phần thứ hai không nén thành hình ảnh thứ nhất được mã hóa.

Theo một phương án thiết bị còn bao gồm mã làm cho thiết bị thực hiện ít nhất một trong số bước:

mã hóa hình ảnh mã hóa thứ hai;

bước mã hóa hình ảnh được mã hóa thứ hai nêu trên bao gồm mã hóa vectơ chuyển động thực hiện tham chiếu đến vị trí mẫu bên ngoài vùng hình ảnh hiệu dụng của khung thành phần được khôi phục trong bước giải mã; và

làm bão hòa hoặc gói vị trí mẫu ở trong vùng hình ảnh hiệu dụng.

Theo ví dụ thứ bảy, sáng chế đề xuất vật ghi đọc được bằng máy tính được lưu trữ với mã trên đó để sử dụng bởi thiết bị, khi được chạy bởi bộ xử lý, khiến thiết bị thực hiện:

mã hóa khung thành phần không nén thành hình ảnh thứ nhất được mã hóa, việc mã hóa nêu trên cũng dẫn đến hình ảnh khôi phục thứ nhất và khung thành phần nêu trên có vùng hình ảnh hiệu dụng trong hình ảnh khôi phục thứ nhất;

thực hiện một trong các bước sau dưới dạng một phần của bước mã hóa nêu trên:

chèn ít nhất một giá trị mẫu bên ngoài vùng hình ảnh hiệu dụng để tạo ra phần mở rộng biên đối với khung thành phần trong hình ảnh khôi phục thứ nhất; hoặc

làm bão hòa hoặc gói các vị trí mẫu bên ngoài vùng hình ảnh hiệu dụng ở trong vùng hình ảnh hiệu dụng.

Theo ví dụ thứ tám sáng chế đề xuất thiết bị bao gồm:

các phương tiện mã hóa khung thành phần không nén thành hình ảnh thứ nhất được mã hóa, việc mã hóa nêu trên cũng dẫn đến hình ảnh khôi phục thứ nhất và khung thành phần nêu trên mà có vùng hình ảnh hiệu dụng trong hình ảnh khôi phục thứ nhất;

các phương tiện để thực hiện một trong các bước sau dưới dạng một phần của bước mã hóa nêu trên:

chèn ít nhất một giá trị mẫu bên ngoài vùng hình ảnh hiệu dụng để tạo ra phần mở rộng biên đối với khung thành phần trong hình ảnh khôi phục thứ nhất; hoặc

làm bão hòa hoặc gói các vị trí mẫu bên ngoài vùng hình ảnh hiệu dụng ở trong vùng hình ảnh hiệu dụng.

YÊU CẦU BẢO HỘ

1. Phương pháp mã hóa và giải mã video bao gồm các bước:

nhận hình ảnh monoscopic được mã hóa;

giải mã hình ảnh monoscopic được mã hóa để tạo ra phần liên kết theo không gian được khôi phục của hình ảnh monoscopic được mã hóa bao gồm vùng hình ảnh hiệu dụng;

giải mã hình ảnh monoscopic được mã hóa để tạo ra phần liên kết theo không gian được khôi phục thứ hai của hình ảnh monoscopic được mã hóa;

giải mã chỉ báo để chỉ báo, trên cơ sở phần liên kết theo không gian, liệu các vị trí mẫu ở ngoài các biên phần liên kết theo không gian có được làm bão hòa hay không; và

xác định, khi tham chiếu đến các vị trí mẫu ở ngoài vùng hình ảnh hiệu dụng trong bước giải mã, liệu các vị trí mẫu nêu trên có được làm bão hòa hay không.

2. Thiết bị mã hóa và giải mã video bao gồm ít nhất một bộ xử lý; và ít nhất một bộ nhớ bao gồm mã chương trình máy tính, ít nhất một bộ nhớ và mã chương trình máy tính được tạo cấu hình để, với ít nhất một bộ xử lý, khiến thiết bị này thực hiện:

nhận hình ảnh monoscopic được mã hóa;

giải mã hình ảnh monoscopic được mã hóa để tạo ra phần liên kết theo không gian được khôi phục của hình ảnh monoscopic được mã hóa có vùng hình ảnh hiệu dụng;

giải mã hình ảnh monoscopic được mã hóa để tạo ra phần liên kết theo không gian được khôi phục thứ hai của hình ảnh monoscopic được mã hóa;

giải mã chỉ báo để chỉ báo, trên cơ sở phần liên kết theo không gian, liệu các vị trí mẫu ở ngoài các biên phần liên kết theo không gian có được làm bão hòa hay không; và

xác định, khi tham chiếu đến các vị trí mẫu ở ngoài vùng hình ảnh hiệu dụng trong việc giải mã, liệu các vị trí mẫu nêu trên có được làm bão hòa hay không.

3. Thiết bị theo điểm 2, trong đó để làm đầy vùng ở ngoài vùng hình ảnh hiệu dụng để tạo ra hình ảnh tham chiếu được đệm, trong đó vùng được làm đầy tạo ra phần mở rộng biên, thiết bị này còn thực hiện ít nhất một trong số:

giải mã ít nhất một giá trị mẫu bằng dự đoán trong ảnh theo chiều ngang;

giải mã ít nhất một giá trị mẫu bằng dự đoán trong ảnh theo chiều dọc;

giải mã ít nhất một giá trị mẫu bằng dự đoán bản sao nội khối từ vùng biên của vùng hình ảnh hiệu dụng đối diện với phần mở rộng biên;

giải mã ít nhất một giá trị mẫu bằng dự đoán bản sao nội khối từ phần liên kề theo không gian khác trong hình ảnh monoscopic được mã hóa; hoặc

giải mã ít nhất một giá trị mẫu bằng cách dự đoán liên ảnh từ hình ảnh monoscopic được mã hóa trước theo thứ tự giải mã.

4. Thiết bị theo điểm 2, trong đó thiết bị này còn thực hiện: giải mã chỉ báo liệu các vị trí mẫu ở ngoài các biên phần liên kề theo không gian được làm bão hòa hoặc được gói; và

xác định theo chỉ báo được giải mã liệu các vị trí mẫu nêu trên được làm bão hòa hay được gói bên trong vùng hình ảnh hiệu dụng.

5. Thiết bị theo điểm 2, trong đó thiết bị này còn thực hiện: giải mã chỉ báo của vùng hình ảnh hiệu dụng của phần liên kề theo không gian.

6. Thiết bị theo điểm 2, trong đó thiết bị này còn thực hiện: chèn ít nhất một giá trị mẫu bằng cách sao chép từ phía đối diện của vùng hình ảnh hiệu dụng khi phần liên kề theo không gian được chỉ báo là hình ảnh toàn cảnh 360 độ.

7. Thiết bị theo điểm 2, trong đó thiết bị này còn thực hiện trong đó việc giải mã hình ảnh monoscopic được mã hóa bao gồm:

giải mã vùng ở ngoài vùng hình ảnh hiệu dụng để thu khung tham chiếu trung gian; và

hơn nữa trong đó, để làm đầy vùng ở ngoài vùng hình ảnh hiệu dụng, thiết bị này thực hiện sửa đổi ít nhất một giá trị mẫu của phần mở rộng biên của khung tham chiếu trung gian.

8. Thiết bị theo điểm 2, trong đó thiết bị này còn thực hiện:

nhận hình ảnh monoscopic được mã hóa thứ hai;

giải mã vectơ chuyển động từ hình ảnh monoscopic được mã hóa thứ hai, vectơ chuyển động làm cho tham chiếu đến vị trí mẫu ở ngoài vùng hình ảnh hiệu dụng của phần liên kề theo không gian được khôi phục trong bước giải mã; và

làm bão hòa hoặc gói vị trí mẫu ở bên trong vùng hình ảnh hiệu dụng.

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

mã hóa phần liên kề theo không gian không nén thành hình ảnh monoscopic được mã hóa thứ nhất, bước mã hóa nêu trên cũng dẫn đến hình ảnh thứ nhất được khôi phục và phần liên kề theo không gian nêu trên bao gồm vùng hình ảnh hiệu dụng bên trong hình ảnh thứ nhất được khôi phục, trong đó bước mã hóa bao gồm thực hiện một trong số;

chèn ít nhất một giá trị mẫu ở ngoài vùng hình ảnh hiệu dụng để tạo ra phần mở rộng biên đối với phần liên kề theo không gian không nén trong hình ảnh thứ nhất được khôi phục; hoặc

làm bão hòa các vị trí mẫu ở ngoài vùng hình ảnh hiệu dụng ở bên trong vùng hình ảnh hiệu dụng;

mã hóa phần liên kề theo không gian không nén thứ hai thành hình ảnh monoscopic được mã hóa thứ nhất; và

mã hóa chỉ báo để chỉ báo, trên cơ sở phần liên kề theo không gian, liệu các vị trí mẫu ở ngoài các biên phần liên kề theo không gian có được làm bão hòa hay không.

10. Phương pháp theo điểm 9, trong đó phương pháp này còn bao gồm ít nhất một trong số bước sau đây:

thu ít nhất một giá trị mẫu bằng dự đoán trong ảnh theo chiều ngang;

thu ít nhất một giá trị mẫu bằng dự đoán trong ảnh theo chiều dọc;

thu ít nhất một giá trị mẫu bằng dự đoán bản sao nội khối từ biên của vùng hình ảnh hiệu dụng đối diện với phần mở rộng biên;

thu ít nhất một giá trị mẫu bằng dự đoán bản sao nội khối từ phần liên kề theo không gian khác trong hình ảnh thứ nhất được khôi phục; hoặc

thu ít nhất một giá trị mẫu bằng cách dự đoán liên ảnh từ hình ảnh trước theo thứ tự giải mã.

11. Phương pháp theo điểm 9, trong đó phương pháp này còn bao gồm bước:

mã hóa chỉ báo liệu các vị trí mẫu ở ngoài các biên phần liên kề theo không gian được làm bão hòa hay được gói.

12. Thiết bị mã hóa và giải mã video bao gồm ít nhất một bộ xử lý; và ít nhất một bộ nhớ bao gồm mã chương trình máy tính, ít nhất một bộ nhớ và mã chương trình máy tính được tạo cấu hình để, với ít nhất một bộ xử lý, khiến thiết bị này thực hiện:

mã hóa phần liên kề theo không gian không nén thành hình ảnh monoscopic được mã hóa thứ nhất, bước mã hóa nêu trên còn dẫn đến hình ảnh thứ nhất được khôi phục và phần liên kề theo không gian nêu trên có vùng hình ảnh hiệu dụng bên trong hình ảnh được khôi phục thứ nhất;

thực hiện một trong số các hoạt động sau là một phần của việc mã hóa nêu trên:

chèn ít nhất một giá trị mẫu ở ngoài vùng hình ảnh hiệu dụng để tạo ra phần mở rộng biên đối với phần liên kề theo không gian trong hình ảnh thứ nhất được khôi phục; hoặc

làm bão hòa các vị trí mẫu ở ngoài vùng hình ảnh hiệu dụng ở bên trong vùng hình ảnh hiệu dụng;

mã hóa phần liên kề theo không gian không nén thứ hai thành hình ảnh monoscopic được mã hóa thứ nhất; và

mã hóa chỉ báo để chỉ báo, trên cơ sở phần liên kề theo không gian, liệu các vị trí mẫu ở ngoài các biên phần liên kề theo không gian có được làm bão hòa hay không.

13. Thiết bị theo điểm 12, trong đó thiết bị này còn thực hiện ít nhất một trong số hoạt động sau:

thu ít nhất một giá trị mẫu bằng dự đoán trong ảnh theo chiều ngang;

thu ít nhất một giá trị mẫu bằng dự đoán trong ảnh theo chiều dọc;

thu ít nhất một giá trị mẫu bằng dự đoán bản sao nội khối từ biên của vùng hình ảnh hiệu dụng đối diện với phần mở rộng biên;

thu ít nhất một giá trị mẫu bằng dự đoán bản sao nội khối từ phần liên kề theo không gian khác trong hình ảnh được khôi phục; hoặc

thu ít nhất một giá trị mẫu bằng cách dự đoán liên ảnh từ hình ảnh trước theo thứ tự giải mã.

14. Thiết bị theo điểm 12, trong đó thiết bị này còn thực hiện: mã hóa chỉ báo liệu các vị trí mẫu ở ngoài các biên phần liên kề theo không gian được làm bão hòa hoay được gói.

15. Thiết bị theo điểm 12, trong đó thiết bị này còn thực hiện: mã hóa chỉ báo của vùng hình ảnh hiệu dụng của phần liên kề theo không gian.

16. Thiết bị theo điểm 12, trong đó thiết bị này còn thực hiện: chèn ít nhất một giá trị mẫu bằng cách sao chép từ phía đối diện của vùng hình ảnh hiệu dụng khi phần liên kề theo không gian được chỉ báo là hình ảnh toàn cảnh 360 độ.

17. Thiết bị theo điểm 12, trong đó thiết bị này còn thực hiện: khôi phục vùng ở ngoài vùng hình ảnh hiệu dụng để thu khung tham chiếu trung gian.

18. Thiết bị theo điểm 12, trong đó thiết bị này còn thực hiện:

mã hóa hình ảnh monoscopic được mã hóa thứ hai;

mã hóa vectơ chuyển động khiến cho tham chiếu đến vị trí mẫu ở ngoài vùng hình ảnh hiệu dụng của phần liên kề theo không gian được khôi phục trong việc giải mã; và

làm bão hòa hoặc gói vị trí mẫu ở bên trong vùng hình ảnh hiệu dụng.

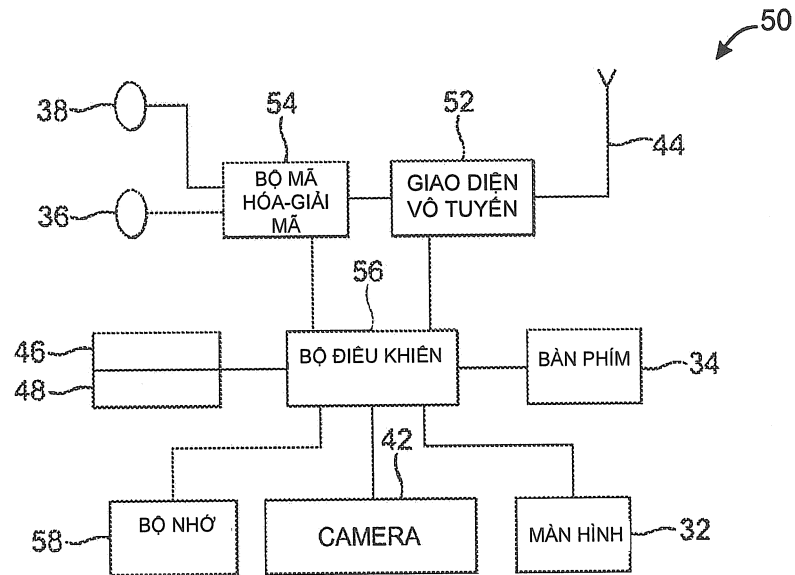


Fig. 1

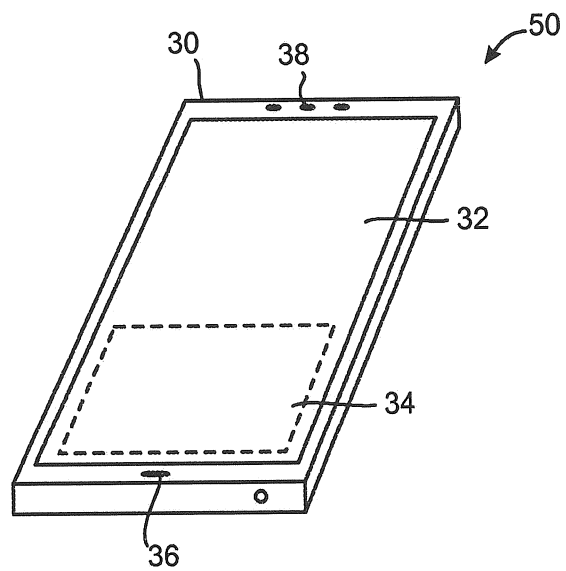


Fig. 2

2/10

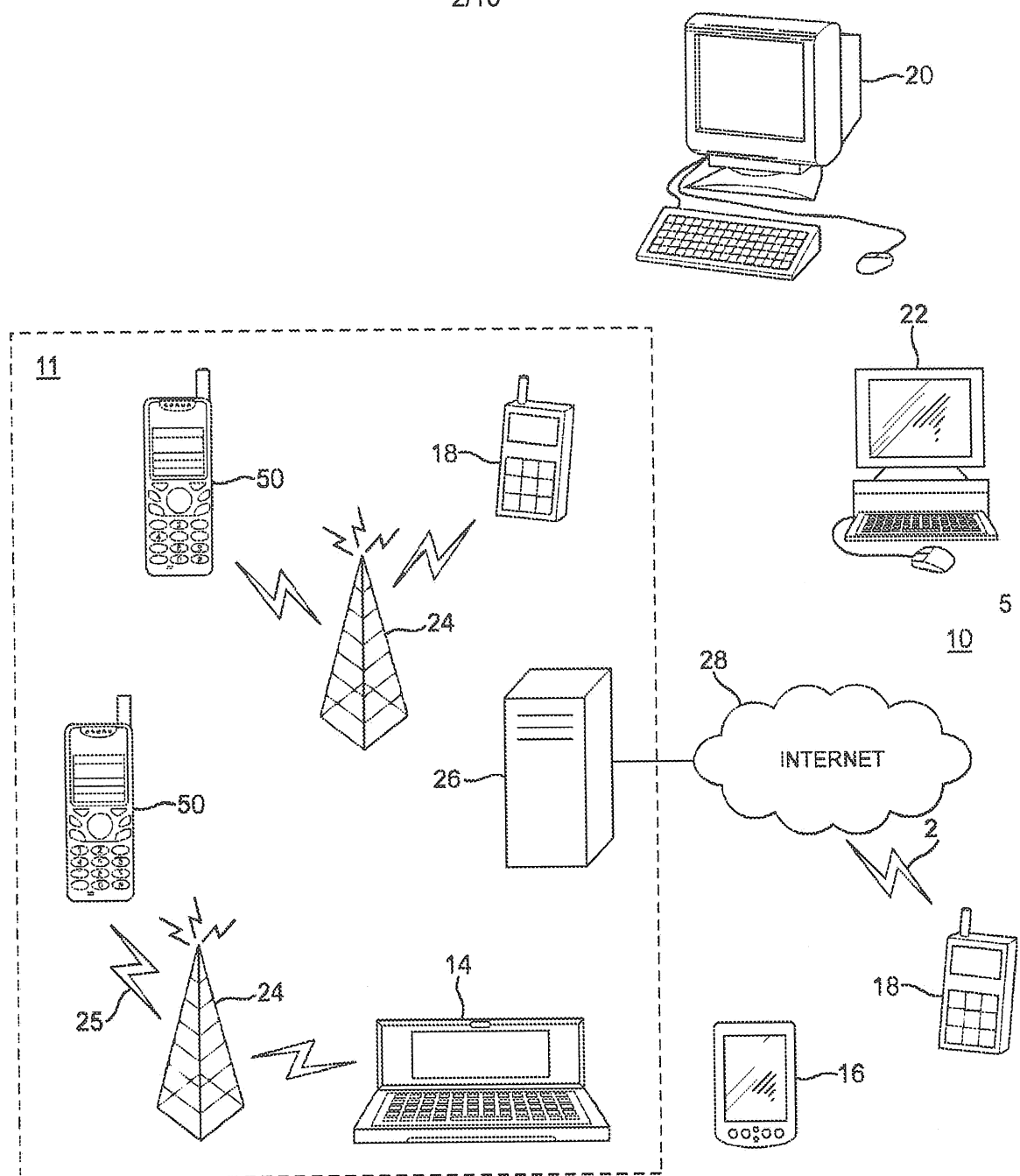


Fig. 3

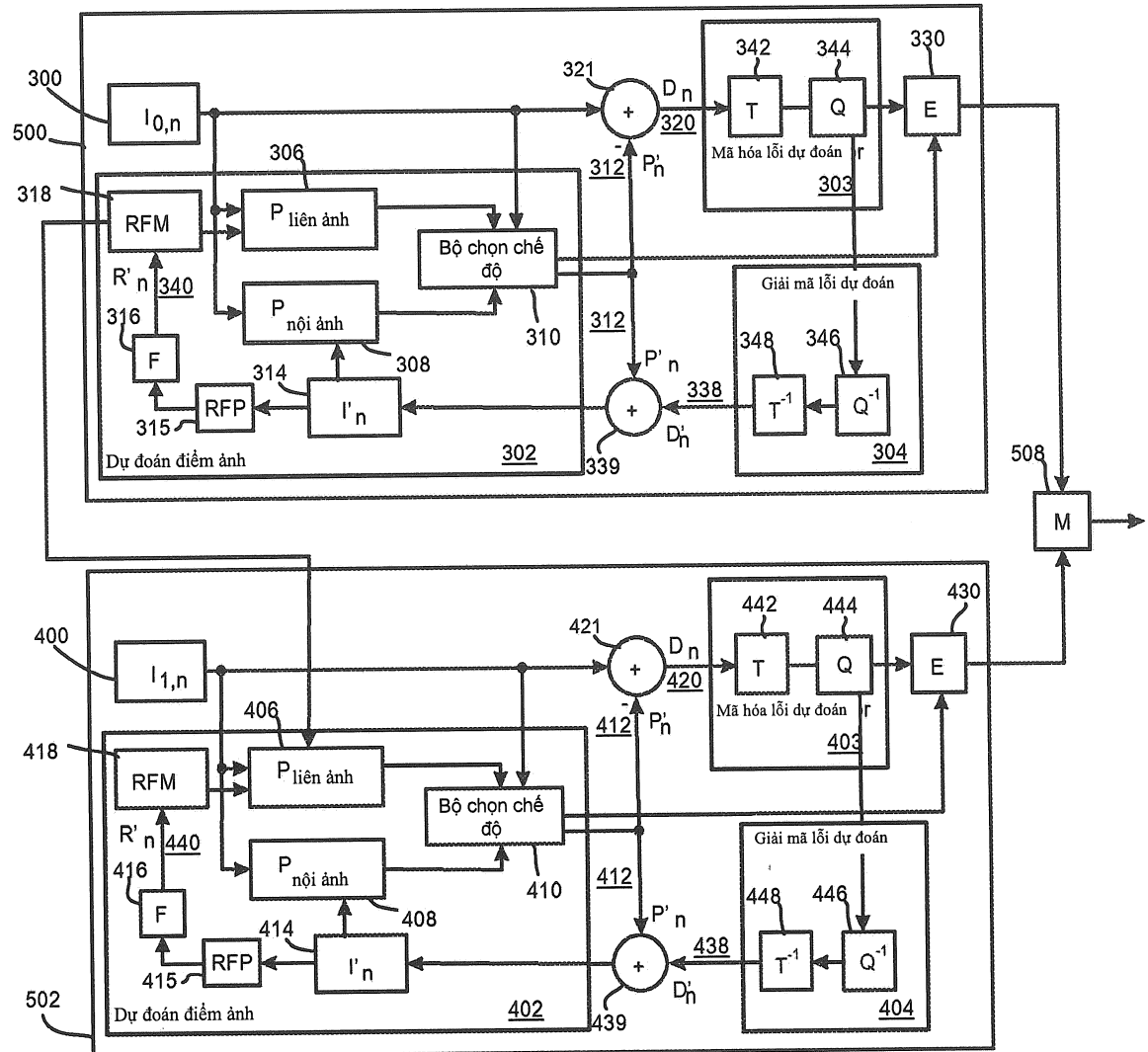


Fig. 4

4/10

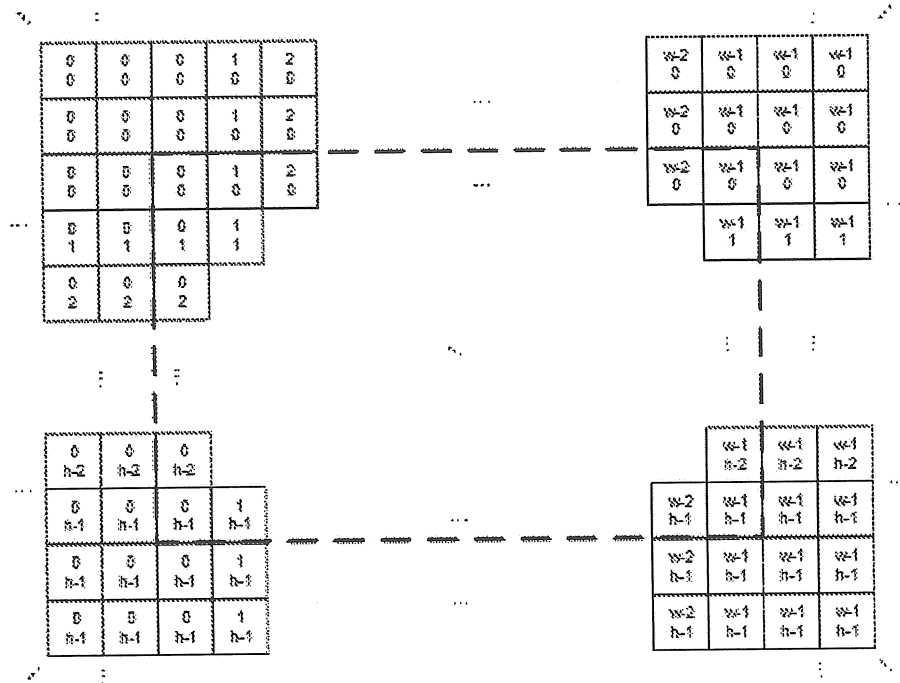


Fig.5a

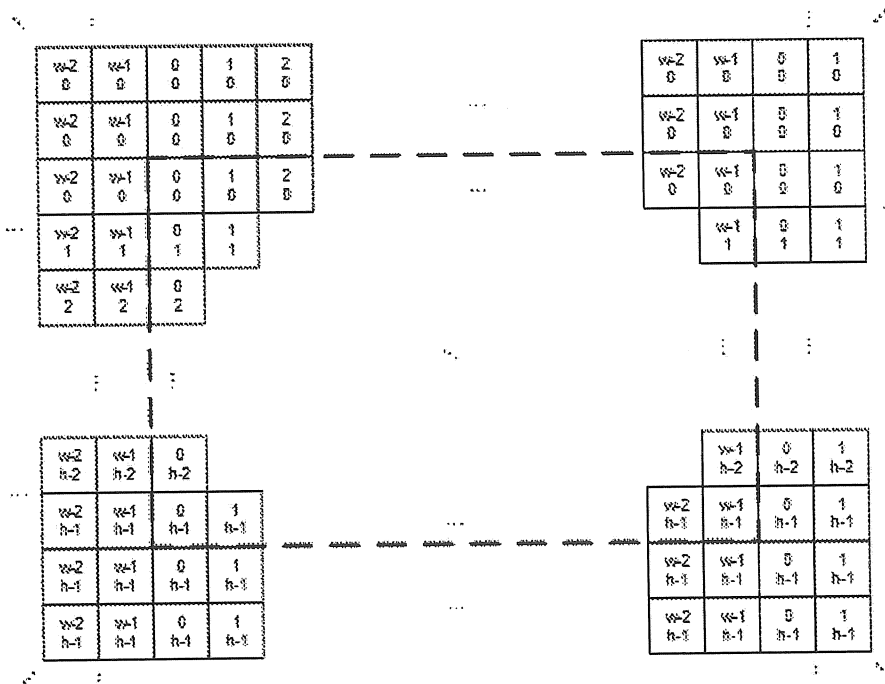


Fig.5b

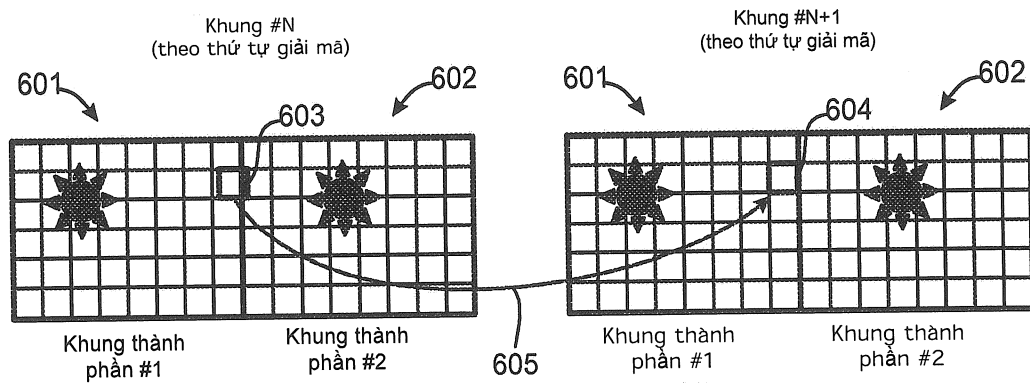


Fig. 6a

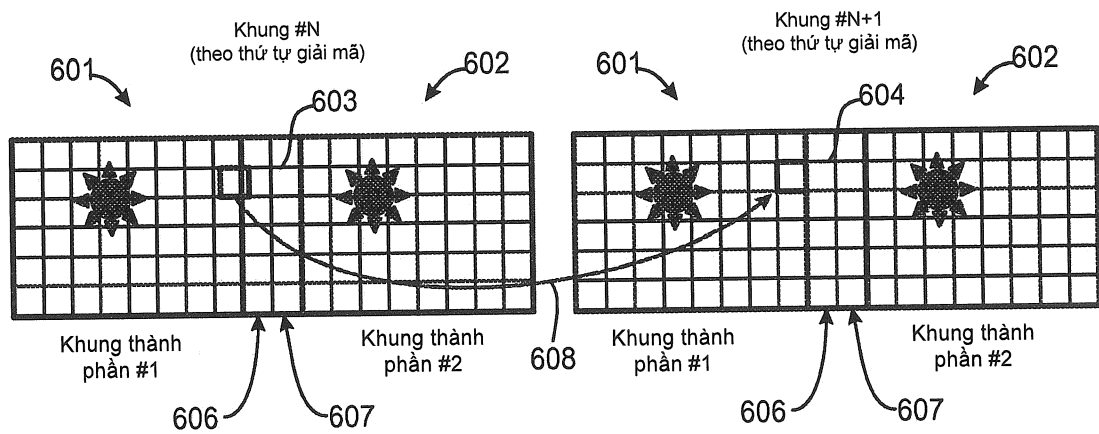


Fig. 6b

6/10

50

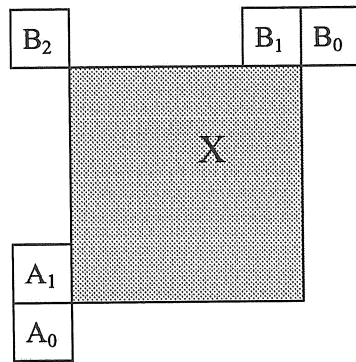


Fig. 6c

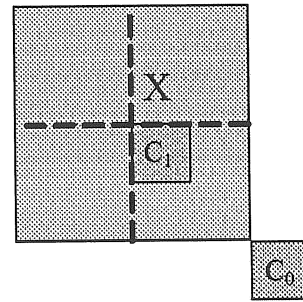


Fig. 6d

7/10

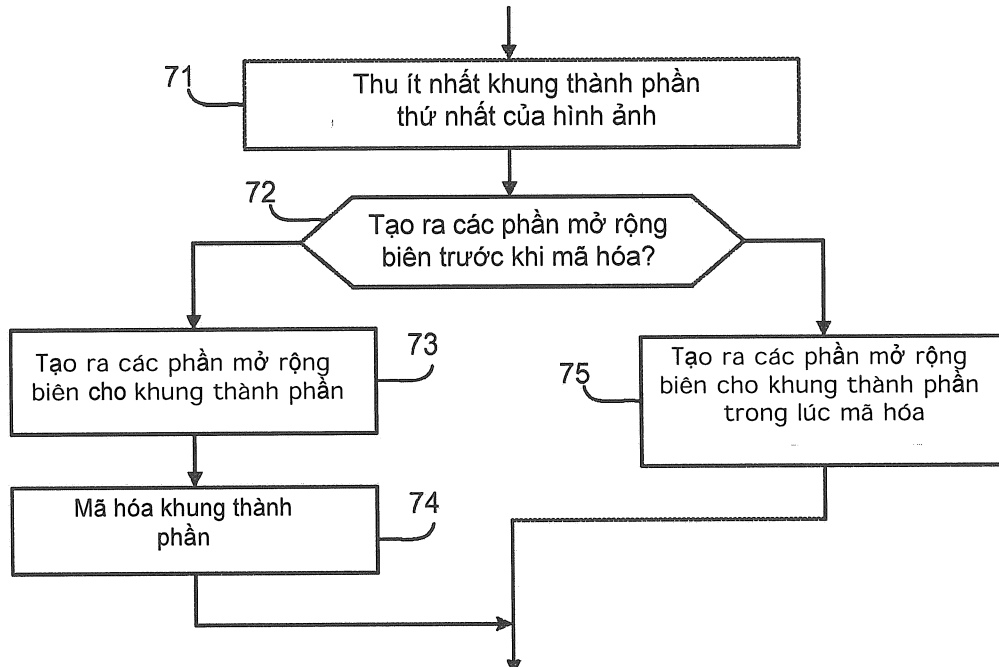


Fig. 7a

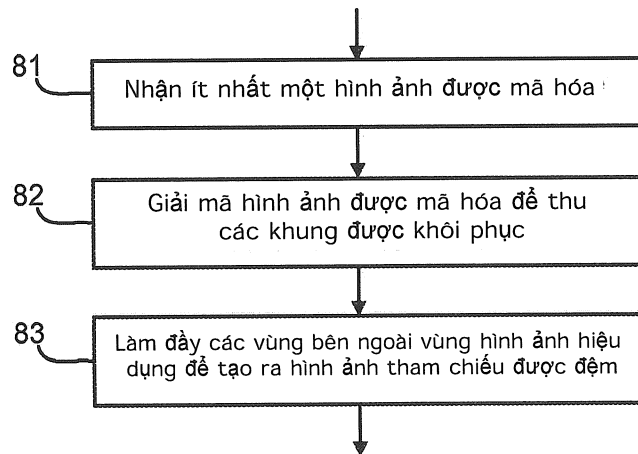


Fig. 7b

8/10

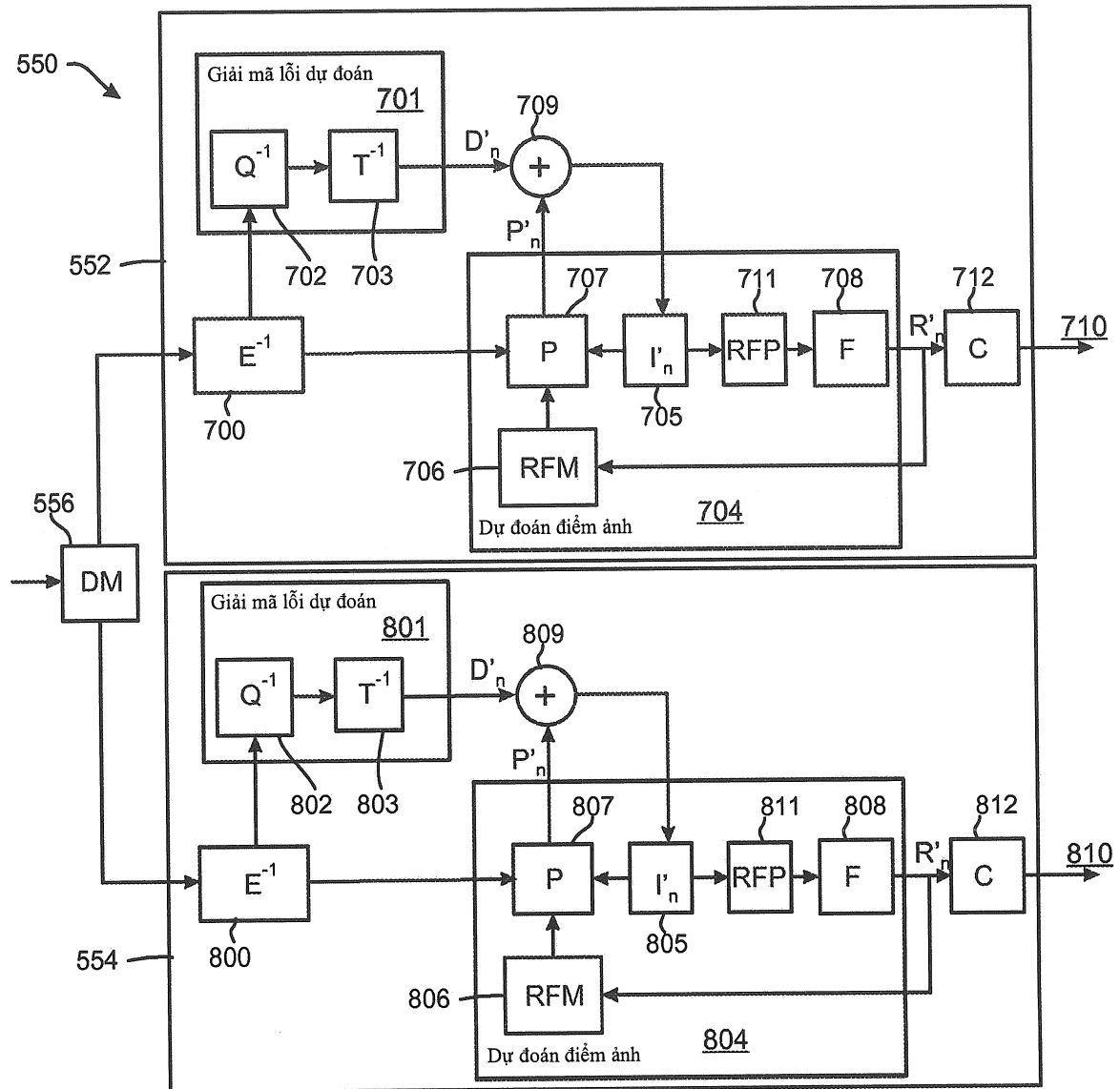


Fig. 8

9/10

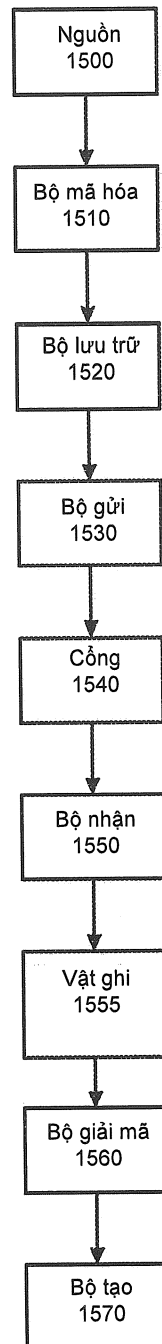


Fig. 9

10/10

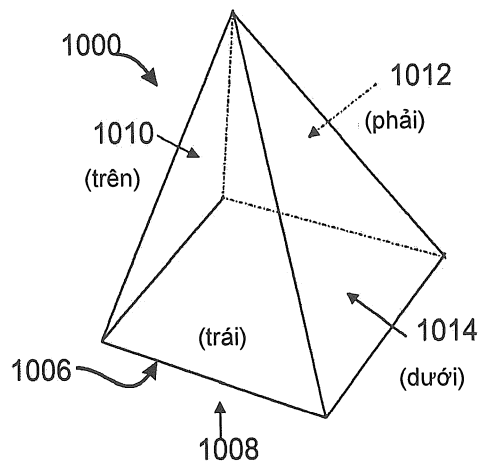


Fig. 10a

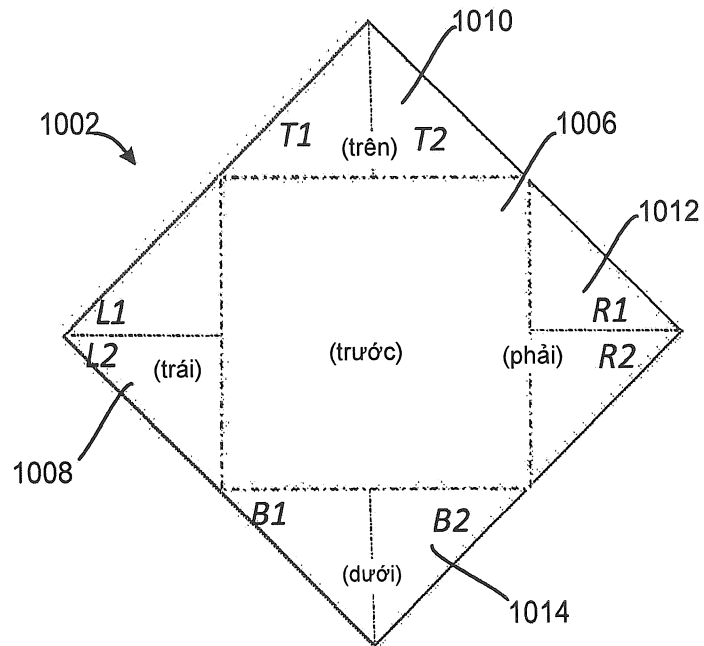


Fig. 10b

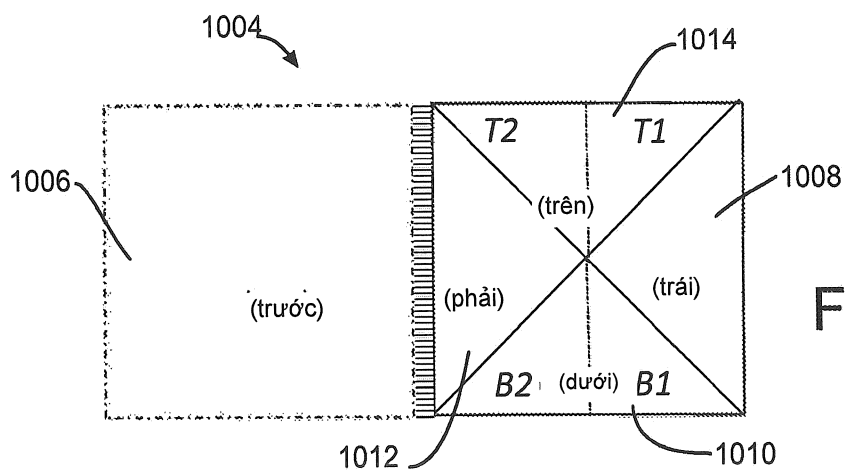


Fig. 10c