



(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)



1-0036598

(51)⁷ H04N 19/56; H04N 19/513; H04N
19/537; G06T 7/246; H04N 19/52

(13) B

(21) 1-2019-00678

(22) 11/08/2017

(86) PCT/FI2017/050570 11/08/2017

(87) WO 2018/033661 22/02/2018

(30) 20165613 15/08/2016 FI

(45) 25/08/2023 425

(43) 25/04/2019 373A

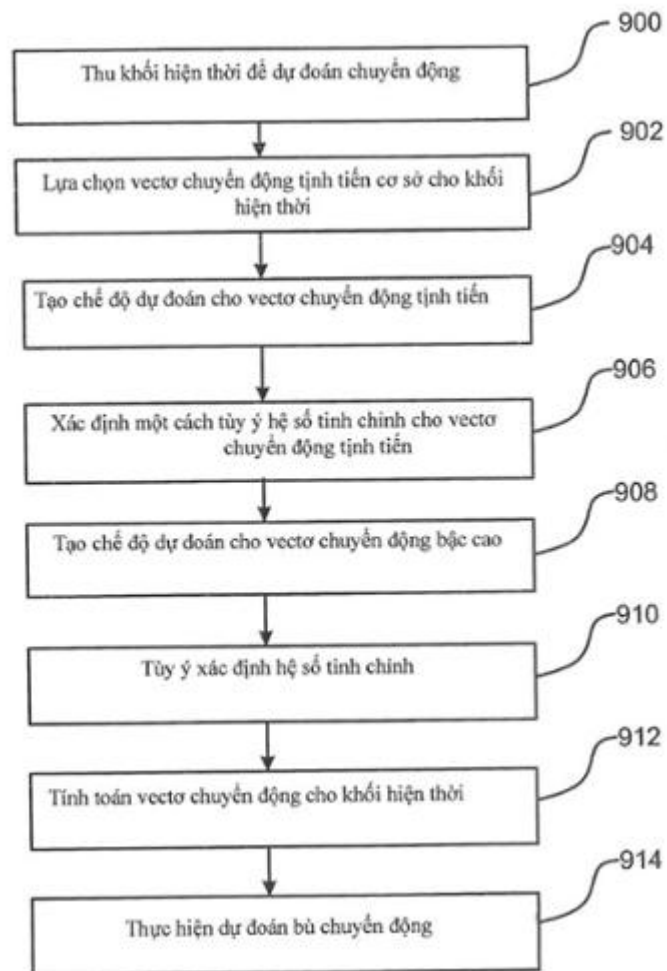
(73) Nokia Technologies Oy (FI)
Karaportti 3, 02610 Espoo, Finland

(72) LAINEMA, Jani (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 cập đến phương pháp, thiết bị và sản phẩm chương trình máy tính dùng để mã hóa và giải mã video. Theo một số phương án mã hóa, khối mẫu được thu để dự đoán chuyển động. Ít nhất một thành phần vector chuyển động tịnh tiến cơ sở cho khối mẫu được tính toán bằng cách xác định chế độ dự đoán cho thành phần vector chuyển động tịnh tiến cơ sở này. Thành phần vector chuyển động vi sai thứ nhất được xác định bằng cách sử dụng độ chính xác thứ nhất. Thành phần vector chuyển động vi sai thứ nhất này được bổ sung vào chế độ dự đoán cho thành phần vector chuyển động tịnh tiến cơ sở. Ít nhất một thành phần vector chuyển động bậc cao được tính toán bằng cách xác định chế độ dự đoán cho thành phần vector chuyển động bậc cao này và xác định thành phần vector chuyển động vi sai thứ hai bằng cách sử dụng độ chính xác thứ hai khác với độ chính xác thứ nhất. Thành phần vector chuyển động vi sai thứ hai này được bổ sung vào chế độ dự đoán cho thành phần vector chuyển động bậc cao. Bước bù chuyển động được thực hiện cho khối mẫu bằng cách sử dụng ít nhất một thành phần vector chuyển động tịnh tiến cơ sở và ít nhất một thành phần vector chuyển động vi sai.



Lĩnh vực kỹ thuật được đề cập

Nói chung, sáng chế đề cập đến phương pháp và thiết bị mã hóa và giải mã các dạng biểu diễn đa phương tiện.

Tình trạng kỹ thuật của sáng chế

Phần này nhằm mục đích cung cấp tình trạng kỹ thuật hoặc bối cảnh cho sáng chế được yêu cầu bảo hộ. Phần mô tả ở đây có thể bao gồm các khái niệm mà có thể được sử dụng theo, nhưng không nhất thiết phải là các khái niệm đã được nghĩ ra hoặc sử dụng theo từ trước đó. Do đó, trừ khi có quy định khác trong bản mô tả này, các nội dung được mô tả trong phần này không phải là tình trạng kỹ thuật đối với phần mô tả và yêu cầu bảo hộ của đơn sáng chế này, và cũng không được thừa nhận là tình trạng kỹ thuật bởi việc đưa vào trong phần này.

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

Bản chất kỹ thuật của sáng chế

Một số phương án của sáng chế đề xuất phương pháp mã hóa và giải mã thông tin video. Theo một số phương án, sáng chế đề xuất thiết bị, sản phẩm chương trình máy tính, vật ghi đọc được bằng máy tính dùng để thực hiện phương pháp này.

Các khía cạnh khác của các phương án của sáng chế được thể hiện trong phần mô tả chi tiết sáng chế.

Theo khía cạnh thứ nhất, sáng chế đề xuất phương pháp bao gồm các bước nêu dưới đây:

thu khối mẫu để dự đoán chuyển động;

tính toán ít nhất một thành phần vector chuyển động tịnh tiến cơ sở cho khối mẫu bằng cách xác định chế độ dự đoán cho thành phần vector chuyển động tịnh tiến cơ sở này;

xác định thành phần vector chuyển động vi sai thứ nhất bằng cách sử dụng độ chính xác thứ nhất; và

bổ sung thành phần vector chuyển động vi sai thứ nhất này vào chế độ dự đoán cho thành phần vector chuyển động tịnh tiến cơ sở;

tính toán ít nhất một thành phần vector chuyển động bậc cao bằng cách

xác định chế độ dự đoán cho thành phần vector chuyển động bậc cao này;

xác định thành phần vector chuyển động vi sai thứ hai bằng cách sử dụng độ chính xác thứ hai khác với độ chính xác thứ nhất;

bổ sung thành phần vector chuyển động vi sai thứ hai này vào chế độ dự đoán cho thành phần vector chuyển động bậc cao; và

thực hiện bước bù chuyển động cho khối mẫu bằng cách sử dụng ít nhất một thành phần vector chuyển động tịnh tiến cơ sở và ít nhất một thành phần vector chuyển động vi sai.

Theo khía cạnh thứ hai, sáng chế đề xuất thiết bị bao gồm ít nhất một bộ xử lý; và ít nhất một bộ nhớ bao gồm mã chương trình máy tính, trong đó ít nhất một bộ nhớ và mã chương trình máy tính này được tạo cấu hình để, với ít nhất một bộ xử lý, giúp thiết bị thực hiện ít nhất là bước nêu dưới đây:

thu khối mẫu để dự đoán chuyển động;

tính toán ít nhất một thành phần vector chuyển động tịnh tiến cơ sở cho khối mẫu bằng cách xác định chế độ dự đoán cho thành phần vector chuyển động tịnh tiến cơ sở này;

xác định thành phần vector chuyển động vi sai thứ nhất bằng cách sử dụng độ chính xác thứ nhất; và

bổ sung thành phần vector chuyển động vi sai thứ nhất này vào chế độ dự đoán cho thành phần vector chuyển động tịnh tiến cơ sở;

tính toán ít nhất một thành phần vector chuyển động bậc cao bằng cách

xác định chế độ dự đoán cho thành phần vector chuyển động bậc cao này;

xác định thành phần vector chuyển động vi sai thứ hai bằng cách sử dụng độ chính xác thứ hai khác với độ chính xác thứ nhất;

bổ sung thành phần vector chuyển động vi sai thứ hai này vào chế độ dự đoán cho thành phần vector chuyển động bậc cao; và

thực hiện bước bù chuyển động cho khối mẫu bằng cách sử dụng ít nhất một thành phần vector chuyển động tịnh tiến cơ sở và ít nhất một thành phần vector chuyển động vi sai.

Theo khía cạnh thứ ba, sáng chế đề xuất vật ghi đọc được bằng máy tính có lưu trữ mã trên đó để sử dụng bằng thiết bị, mà khi được thực thi bởi bộ xử lý, sẽ giúp thiết bị thực hiện các bước nêu dưới đây:

thu khối mẫu để dự đoán chuyển động;

tính toán ít nhất một thành phần vector chuyển động tịnh tiến cơ sở cho khối mẫu bằng cách xác định chế độ dự đoán cho thành phần vector chuyển động tịnh tiến cơ sở này;

xác định thành phần vector chuyển động vi sai thứ nhất bằng cách sử dụng độ chính xác thứ nhất; và

bổ sung thành phần vector chuyển động vi sai thứ nhất này vào chế độ dự đoán cho thành phần vector chuyển động tịnh tiến cơ sở;

tính toán ít nhất một thành phần vector chuyển động bậc cao bằng cách

xác định chế độ dự đoán cho thành phần vector chuyển động bậc cao này;

xác định thành phần vector chuyển động vi sai thứ hai bằng cách sử dụng độ chính xác thứ hai khác với độ chính xác thứ nhất;

bổ sung thành phần vector chuyển động vi sai thứ hai này vào chế độ dự đoán cho thành phần vector chuyển động bậc cao; và

thực hiện bước bù chuyển động cho khối mẫu bằng cách sử dụng ít nhất một thành phần vectơ chuyển động tịnh tiến cơ sở và ít nhất một thành phần vectơ chuyển động vi sai.

Theo khía cạnh thứ tư, sáng chế đề xuất thiết bị bao gồm bộ mã hóa video được tạo cấu hình để mã hóa luồng bit chứa chuỗi hình ảnh, bộ mã hóa video này bao gồm:

phương tiện để thu khối mẫu để dự đoán chuyển động;

phương tiện để tính toán ít nhất một thành phần vectơ chuyển động tịnh tiến cơ sở cho khối mẫu bằng cách xác định chế độ dự đoán cho thành phần vectơ chuyển động tịnh tiến cơ sở này;

phương tiện để xác định thành phần vectơ chuyển động vi sai thứ nhất bằng cách sử dụng độ chính xác thứ nhất;

phương tiện để bổ sung thành phần vectơ chuyển động vi sai thứ nhất này vào chế độ dự đoán cho thành phần vectơ chuyển động tịnh tiến cơ sở;

phương tiện để tính toán ít nhất một thành phần vectơ chuyển động bậc cao bằng cách

xác định chế độ dự đoán cho thành phần vectơ chuyển động bậc cao này;

xác định thành phần vectơ chuyển động vi sai thứ hai bằng cách sử dụng độ chính xác thứ hai khác với độ chính xác thứ nhất;

phương tiện để bổ sung thành phần vectơ chuyển động vi sai thứ hai này vào chế độ dự đoán cho thành phần vectơ chuyển động bậc cao; và

phương tiện để thực hiện bước bù chuyển động cho khối mẫu bằng cách sử dụng ít nhất một thành phần vectơ chuyển động tịnh tiến cơ sở và ít nhất một thành phần vectơ chuyển động vi sai.

Mô tả vắn tắt các hình vẽ

Để hiểu đầy đủ hơn về các phương án ví dụ của sáng chế, cần tham khảo phần mô tả dưới đây cùng với các hình vẽ đi kèm, trong đó:

Fig.1 minh họa sơ đồ khối của hệ thống mã hóa video theo một phương án;

Fig.2 minh họa cấu trúc của thiết bị theo một phương án;

Fig.3 minh họa hệ thống mã hóa video bao gồm các thiết bị, mạng và phân tử mạng theo một phương án;

Fig.4 minh họa sơ đồ khối của bộ mã hóa video theo một phương án;

Fig.5 minh họa sơ đồ khối của bộ giải mã video theo một phương án;

Fig.6 thể hiện ví dụ về ảnh gồm hai ô;

Fig.7 minh họa phần tử xác định vector chuyển động dưới dạng sơ đồ khối đơn giản theo một phương án của sáng chế;

Fig.8a minh họa các vector chuyển động cơ sở b_0 và b_1 bên trong đơn vị mã hóa;

Fig.8b minh họa các vector chuyển động cơ sở b_0 và b_1 bên ngoài đơn vị mã hóa;

Fig.8c minh họa việc sử dụng mô hình afin với giá trị dương cho tham số bậc nhất a_0 ;

Fig.9 minh họa phương pháp mã hóa theo một phương án của sáng chế dưới dạng lưu đồ;

Fig.10 minh họa phương pháp giải mã theo một phương án của sáng chế dưới dạng lưu đồ; và

Fig.11 thể hiện sơ đồ khối của hệ thống truyền thông đa phương tiện ví dụ trong đó các phương án của sáng chế có thể được thực hiện.

Mô tả chi tiết sáng chế

Fig.1 thể hiện hệ thống mã hóa video dưới dạng sơ đồ khối của phương tiện hoặc thiết bị điện tử 50 theo một phương án của sáng chế. Thiết bị điện tử 50 có thể kết hợp bộ mã hóa - giải mã theo một phương án của sáng chế. Fig.2 thể hiện cấu trúc của thiết bị theo một phương án của sáng chế. Các thành phần được thể hiện trên Fig.1 và Fig.2 sẽ được mô tả dưới đây.

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 thực hiện trong phương tiện hoặc thiết bị điện tử bất kỳ mà có thể thực hiện việc mã hóa và giải mã, hoặc mã hóa hoặc giải mã hình ảnh video.

Thiết bị 50 có thể bao gồm vỏ 30 để chứa và bảo vệ thiết bị. Thiết bị 50 còn có thể bao gồm màn hiển thị 32 ở dạng màn hiển thị tinh thể lỏng. Theo các phương án khác, màn hiển thị có thể là công nghệ màn hiển thị thích hợp bất kỳ thích hợp để hiển thị hình ảnh hoặc video. Thiết bị 50 còn có thể bao gồm bàn phím 34. Theo một phương án, cơ cấu giao diện người dùng hoặc dữ liệu thích hợp bất kỳ có thể được sử dụng. Ví dụ, giao diện người dùng có thể được thực hiện dưới dạng bàn phím ảo hoặc hệ thống nhập dữ liệu dưới dạng một phần của màn hiển thị nhạy tiếp xúc. Thiết bị này có thể bao gồm micro 36 hoặc thiết bị đầu vào audio thích hợp bất kỳ mà có thể là thiết bị đầu vào tín hiệu dạng số hoặc dạng tương tự. Thiết bị 50 còn có thể bao gồm thiết bị đầu ra audio, mà theo một phương án có thể là thiết bị bất kỳ trong số các thiết bị sau: 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 một phương án của sáng chế, thiết bị này có thể được cấp năng lượng bởi thiết bị năng lượng di động thích hợp bất kỳ, như pin mặt trời, pin nhiên liệu hoặc máy phát đều đặn). Thiết bị này còn có thể bao gồm camera 42 có khả năng ghi lại hoặc chụp hình ảnh và/hoặc video. Theo một phương án, thiết bị 50 còn có thể bao gồm cổng hồng ngoại để truyền thông tầm gần đến các thiết bị khác. Theo một phương án, thiết bị 50 còn có thể bao gồm giải pháp truyền thông tầm gần thích hợp bất kỳ như, ví dụ, kết nối không dây Bluetooth® hoặc kết nối có dây USB/Firewire.

Thiết bị 50 có thể bao gồm bộ điều khiển 56 hoặc bộ xử lý để điều khiển thiết bị 50. Bộ điều khiển 56 có thể được kết nối với bộ nhớ 58 mà theo một phương án có thể lưu trữ cả dữ liệu ở dạng hình ảnh và dữ liệu audio và/hoặc cũng có thể lưu trữ các lệnh thực hiện trên bộ điều khiển 56. Bộ điều khiển 56 còn có thể được kết nối với mạch mã hóa - giải mã 54 thích hợp để thực hiện việc mã hóa và/hoặc giải mã dữ liệu audio và/hoặc video hoặc hỗ trợ việc mã hóa và/hoặc giải mã được thực hiện bởi bộ điều khiển 56.

Thiết bị 56 còn có thể bao gồm đầu đọc thẻ 48 và thẻ thông minh 46, ví dụ, UICC (Universal Integrated Circuit Card: thẻ mạch tích hợp toàn cầu) và đầu đọc UICC để

cung cấp thông tin người dùng và thích hợp để cung cấp thông tin xác thực nhằm xác thực và trao quyền cho người dùng trên mạng.

Thiết bị 50 còn có thể bao gồm 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 di động, hệ thống truyền thông không dây hoặc mạng cục bộ không dây. Thiết bị 50 còn có thể bao gồmăng ten 44 được kết nối với mạch giao diện vô tuyến 52 để truyền các tín hiệu tần số vô tuyến được tạo ra tại 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.

Theo một phương án, thiết bị 50 bao gồm camera có khả năng ghi lại hoặc phát hiện các khung riêng lẻ mà sau đó được chuyển tiếp đến bộ mã hóa - giải mã 54 hoặc bộ điều khiển để xử lý. Theo một phương án, thiết bị này có thể thu dữ liệu hình ảnh video để xử lý từ thiết bị khác trước khi truyền và/hoặc lưu trữ. Theo một phương án, thiết bị 50 có thể thu hình ảnh để mã hóa/giải mã bằng kết nối có dây hoặc không dây.

Fig.3 thể hiện hệ thống mã hóa video bao gồm các thiết bị, mạng và phần tử mạng theo một phương án của sáng chế. Fig.3 thể hiện ví dụ về hệ thống trong đó các phương án của sáng chế có thể được sử dụng. Hệ thống 10 bao gồm nhiều thiết bị truyền thông mà có thể truyền thông qua một hoặc nhiều mạng. Hệ thống 10 có thể bao gồm tổ hợp mạng có dây hoặc không dây bất kỳ, 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: wireless local area network) như được định nghĩa theo chuẩn bất kỳ trong số các chuẩn IEEE 802.x, mạng vùng cá nhân Bluetooth, mạng cục bộ Ethernet, mạng cục bộ Token Ring, mạng diện rộng và Internet.

Hệ thống 10 có thể bao gồm cả thiết bị truyền thông có dây và không dây 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 dạng biểu diễn của Internet 28. Sự kết nối với Internet 28 có thể bao gồm, nhưng không giới hạn ở, kết nối không dây tầm xa, kết nối không dây tầm gần và các kết nối có dây khác bao gồm, nhưng không giới hạn ở, đường điện thoại, đường cáp, đường điện và các đường truyền thông tương tự.

Các thiết bị truyền thông ví dụ được thể hiện trong hệ thống 10 có thể bao gồm, nhưng không giới hạn ở, phương tiện hoặc thiết bị điện tử 50, tổ hợp bất kỳ của thiết bị hỗ trợ kỹ thuật số cá nhân (PDA: personal digital assistant) và điện thoại di động 14, PDA 16, thiết bị nhắn tin tích hợp (IMD: integrated messaging device) 18, máy tính để bàn 20, máy tính xách tay 22. Thiết bị 50 có thể là tĩnh hoặc động khi được mang bởi cá thể đang chuyển động. Thiết bị 50 cũng có thể được đặt trên phương tiện vận tải bao gồm, nhưng không giới hạn ở, ô tô, xe tải, xe taxi, xe buýt, tàu, thuyền, máy bay, xe đạp, xe máy hoặc phương tiện vận tải tương tự thích hợp bất kỳ.

Các phương pháp này cũng có thể được thực hiện trong hộp thu giải mã tín hiệu truyền hình (set-top box); nghĩa là đầu thu TV kỹ thuật số, mà có thể có/có thể không có màn hiển thị hoặc khả năng kết nối không dây, trong máy tính bảng hoặc máy tính cá nhân (PC: personal computer) (máy tính xách tay), có phần cứng hoặc phần mềm hoặc tổ hợp của bộ mã hóa/bộ giải mã, trong các hệ điều hành và trong bộ chip, bộ xử lý, DSP và/hoặc hệ thống nhúng thực hiện chức năng mã hóa dựa trên phần cứng/phần mềm.

Một số hoặc nhiều thiết bị khác có thể gửi và nhận cuộc gọi và tin nhắn và truyền thông với nhà cung cấp dịch vụ thông qua kết nối không dây 25 với trạm cơ sở 24. Trạm cơ sở 24 có thể được kết nối với máy chủ 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 các loại thiết bị truyền thông khác nhau.

Các thiết bị truyền thông này có thể truyền thông bằng cách sử dụng các công nghệ truyền bao gồm, nhưng không giới hạn ở, đa truy cập phân chia theo mã (CDMA: code division multiple access), hệ thống truyền thông di động toàn cầu (GSM: global system for mobile communication), hệ thống viễn thông di động toàn cầu (UMTS: universal mobile telecommunications system), đa truy cập phân chia theo thời gian (TDMA: time divisional multiple access), đa truy cập phân theo tần số (FDMA: frequency division multiple access), giao thức điều khiển truyền-giao thức internet (TCP-IP: transmission control protocol-internet protocol), dịch vụ tin nhắn ngắn (SMS: short messaging service), dịch vụ tin nhắn đa phương tiện (MMS: multimedia messaging service), thư điện tử, dịch vụ tin nhắn tức thì (IMS: instant messaging service),

Bluetooth, IEEE 802.11 và công nghệ truyền thông không dây tương tự bất kỳ. Thiết bị truyền thông tham gia vào việc thực hiện các phương án của sáng chế có thể truyền thông bằng cách sử dụng các phương tiện khác nhau bao gồm, nhưng không giới hạn ở, kết nối vô tuyến, hồng ngoại, laze, cáp và kết nối thích hợp bất kỳ.

Bộ mã hóa video có thể bao gồm bộ mã hóa mà biến đổi video đầu vào thành dạng biểu diễn được nén thích hợp để lưu trữ/truyền, và bộ giải mã có khả năng giải nén dạng biểu diễn video đã nén thành dạng xem được. Bộ mã hóa video và/hoặc bộ giải mã video cũng có thể là tách biệt khỏi nhau, nghĩa là không cần thiết phải tạo thành bộ mã hóa - giải mã. Bộ mã hóa có thể loại bỏ một số thông tin nào đó trong chuỗi video ban đầu để biểu diễn video này ở dạng nén hơn (nghĩa là tại tốc độ bit thấp hơn).

Các bộ mã hóa - giải mã video lai, ví dụ ITU-T H.263 và H.264, có thể mã hóa thông tin video trong hai pha. Trước tiên, các giá trị điểm ảnh trong ảnh (hoặc “khối”) nhất định được dự đoán, ví dụ, bằng phương tiện bù chuyển động (tìm và chỉ ra vùng trong một trong số các khung video đã mã hóa trước đó mà tương ứng chặt chẽ với khối đang được mã hóa) hoặc bằng phương tiện không gian (bằng cách sử dụng các giá trị điểm ảnh xung quanh khối cần mã hóa theo cách cụ thể). Thứ hai là, lỗi dự đoán, nghĩa là sự khác biệt giữa khối điểm ảnh được dự đoán và khối điểm ảnh ban đầu, được mã hóa. Việc này có thể được thực hiện bằng cách biến đổi sự khác biệt về các giá trị điểm ảnh bằng cách sử dụng phương pháp biến đổi cụ thể (ví dụ, biến đổi cosin rời rạc (DCT: discrete cosine transform) hoặc biến thể của nó), lượng tử hóa các hệ số và mã hóa entropy các hệ số đã lượng tử hóa. Bằng cách thay đổi tính chính xác của quy trình lượng tử hóa, bộ mã hóa có thể điều khiển sự cân bằng giữa độ chính xác của dạng biểu diễn điểm ảnh (chất lượng ảnh) và kích thước của dạng biểu diễn video đã mã hóa thu được (kích thước tệp của tốc độ bit truyền).

Dự đoán liên ảnh, còn có thể được gọi là dự đoán theo thời gian, bù chuyển động, hoặc dự đoán được bù chuyển động, làm giảm dữ liệu thừa theo thời gian. Trong dự đoán liên ảnh, các nguồn để dự đoán là các ảnh đã giải mã trước đó. Thực tế là, chế độ dự đoán nội ảnh sử dụng các điểm ảnh liền kề trong cùng ảnh có khả năng có tương quan. Chế độ dự đoán nội ảnh có thể được thực hiện trong miền không gian hoặc miền biến đổi, nghĩa là, các giá trị mẫu hoặc các hệ số biến đổi có thể được dự đoán. Chế độ

dự đoán nội ảnh thường được sử dụng trong mã hóa nội ảnh, trong đó không áp dụng sự dự đoán liên ảnh.

Một kết quả của quy trình mã hóa là tập tham số mã hóa, như các vector chuyển động và các hệ số biến đổi đã lượng tử hóa. Nhiều tham số có thể được mã hóa entropy hiệu quả hơn nếu chúng được dự đoán trước tiên từ các tham số lân cận theo không gian hoặc theo thời gian. Ví dụ, vector chuyển động có thể được dự đoán từ các vector chuyển động liền kề theo không gian và chỉ có sự khác biệt so với bộ dự đoán vector chuyển động mới có thể được mã hóa. Dự đoán tham số mã hóa và chế độ dự đoán nội ảnh có thể được gọi chung là dự đoán trong ảnh.

Chuẩn mã hóa video cải tiến (H.264/AVC, còn được gọi là AVC - Advanced Video Coding) được phát triển bởi đội ngũ hợp tác video (JVT - Joint Video Team) thuộc nhóm chuyên gia về mã hóa video (VCEG – Video Coding Experts Group) thuộc Ban tiêu chuẩn viễn thông của Liên hiệp viễn thông quốc tế (ITU-T - International Telecommunication Union) và nhóm chuyên gia về ảnh động (MPEG - Moving Picture Experts Group) thuộc Tổ chức xây dựng tiêu chuẩn quốc tế (ISO - International Organisation for Standardization) / Ủy ban kỹ thuật điện quốc tế (IEC - International Electrotechnical Commission). Chuẩn H.264/AVC được công bố bởi cả hai tổ chức xây dựng tiêu chuẩn mẹ, và được gọi là Tiêu chuẩn khuyến nghị ITU-T H.264 (ITU-T Recommendation H.264) và Tiêu chuẩn quốc tế ISO/IEC 14496-10 (ISO/IEC International Standard 14496-10), còn được gọi là chuẩn mã hóa video cải tiến (AVC - Advanced Video Coding) MPEG-4 Part 10. Có nhiều phiên bản của chuẩn H.264/AVC, tích hợp các nội dung mở rộng hoặc các đặc điểm mới vào bản đặc tả. Các nội dung mở rộng này bao gồm chuẩn mã hóa video có thể mở rộng (SVC - Scalable Video Coding) và chuẩn mã hóa video đa điểm (MVC - Multiview Video Coding).

Phiên bản 1 của chuẩn mã hóa video hiệu suất cao (H.265/HEVC, còn được gọi là HEVC - High Efficiency Video Coding) được phát triển bởi đội ngũ hợp tác về mã hóa video (JCT-VC - Joint Collaborative Team - Video Coding) thuộc VCEG và MPEG. Chuẩn này được công bố bởi cả hai tổ chức xây dựng tiêu chuẩn mẹ, và được gọi là Tiêu chuẩn khuyến nghị ITU-T H.265 (ITU-T Recommendation H.265) và Tiêu chuẩn quốc tế ISO/IEC 23008-2 (ISO/IEC International Standard 23008-2), còn được gọi là chuẩn

mã hóa video hiệu suất cao (HEVC - High Efficiency Video Coding) MPEG-H Part 2. Phiên bản 2 của H.265/HEVC bao gồm các nội dung mở rộng về phạm vi có thể mở rộng, đa điểm và tính chính xác, có thể lần lượt được ký hiệu là SHVC, MV-HEVC và REXT. Phiên bản 2 của H.265/HEVC được công bố trước đó với tên Tiêu chuẩn khuyến nghị ITU-T H.265 (10/2014) và có thể được công bố với tên Phiên bản 2 của ISO/IEC 23008-2 vào 2015. Hiện có nhiều dự án xây dựng tiêu chuẩn đang diễn ra nhằm phát triển hơn nữa các nội dung mở rộng cho H.265/HEVC, bao gồm các nội dung mở rộng mã hóa ba chiều và mã hóa nội dung màn hình, có thể lần lượt được ký hiệu là 3D-HEVC và SCC.

SHVC, MV-HEVC và 3D-HEVC sử dụng bản đặc tả có cơ sở chung, được quy định trong Phụ lục F của phiên bản 2 của chuẩn HEVC. Cơ sở chung này bao gồm, ví dụ, cú pháp và ngữ nghĩa mức cao, ví dụ, quy định rõ một số đặc điểm của các lớp luồng bit, như các ràng buộc liên lớp, cũng như các quy trình giải mã, như xây dựng danh mục ảnh tham chiếu bao gồm các ảnh tham chiếu liên lớp và suy ra số thứ tự hình ảnh cho luồng bit đa lớp. Phụ lục F cũng có thể được sử dụng trong các nội dung mở rộng đa lớp có thể có sau đó của HEVC. Cần hiểu rằng mặc dù bộ mã hóa video, bộ giải mã video, phương pháp mã hóa, phương pháp giải mã, cấu trúc luồng bit và/hoặc các phương án có thể được mô tả trong phần dưới đây với viện dẫn đến các nội dung mở rộng cụ thể, như SHVC và/hoặc MV-HEVC, chúng có thể được áp dụng chung cho các nội dung mở rộng đa lớp bất kỳ của HEVC, và thậm chí chung hơn là cho bất kỳ kỹ thuật mã hóa video đa lớp nào.

Một số định nghĩa chính, luồng bit và cấu trúc mã hóa và các khái niệm về H.264/AVC và HEVC được mô tả trong phần này làm ví dụ về bộ mã hóa, bộ giải mã, phương pháp mã hóa, phương pháp giải mã video, và cấu trúc luồng bit, trong đó các phương án này có thể được thực hiện. Một số định nghĩa chính, cấu trúc luồng bit và cấu trúc mã hóa, và các khái niệm về H.264/AVC là giống như trong HEVC, do đó, chúng được mô tả dưới đây theo cách kết hợp. Các khía cạnh của sáng chế không bị giới hạn ở H.264/AVC hoặc HEVC, mà đúng hơn là phần mô tả này đưa ra một cơ sở có thể mà dựa vào đó sáng chế có thể được nhận biết một phần hoặc hoàn toàn.

Tương tự với nhiều chuẩn mã hóa video trước đó, cú pháp và ngữ nghĩa luồng bit cũng như quy trình giải mã đối với luồng bit không chứa lỗi được quy định cụ thể trong H.264/AVC và HEVC. Quy trình mã hóa không được quy định cụ thể, nhưng các bộ mã hóa phải tạo ra các luồng bit phù hợp. Sự phù hợp giữa luồng bit và bộ giải mã có thể được xác nhận bằng bộ giải mã tham chiếu giả thuyết (HRD – Hypothetical Reference Decoder). Các chuẩn này chứa các công cụ mã hóa giúp đối phó với các tổn hao và lỗi truyền, nhưng việc sử dụng các công cụ này trong việc mã hóa chỉ là tùy chọn và không có quy trình giải mã nào được quy định đối với các luồng bit chứa lỗi.

Trong phần mô tả các chuẩn hiện có cũng như trong phần mô tả các phương án ví dụ, phần tử cú pháp có thể được định nghĩa là phần tử dữ liệu được biểu diễn trong luồng bit. Cấu trúc cú pháp có thể được định nghĩa là không hoặc nhiều phần tử cú pháp cùng có mặt trong luồng bit theo trình tự xác định. Trong phần mô tả về các chuẩn hiện có cũng như trong phần mô tả các phương án ví dụ, cụm từ “bảng thành phần ngoại vi” hoặc “thông qua thành phần ngoại vi” có thể được sử dụng. Ví dụ, một thực thể, như cấu trúc cú pháp hoặc giá trị biến số được sử dụng trong quy trình giải mã, có thể được cung cấp “bảng thành phần ngoại vi” cho quy trình giải mã. Cụm từ “bảng thành phần ngoại vi” có thể thể hiện rằng thực thể này không được bao gồm trong luồng bit được tạo ra bởi bộ mã hóa, mà được truyền từ bên ngoài từ luồng bit, ví dụ, bằng cách sử dụng giao thức điều khiển. Ngoài ra hoặc theo cách khác, cụm từ này có thể có nghĩa là thực thể này không được tạo ra bởi bộ mã hóa, nhưng có thể được tạo ra, ví dụ, trong trình phát hoặc logic điều khiển giải mã hoặc tương tự mà có sử dụng bộ giải mã. Bộ giải mã có thể có giao diện để nhập thành phần ngoại vi, như các giá trị biến số.

Profile có thể được định nghĩa là tập con của toàn bộ cú pháp luồng bit được quy định bởi tiêu chuẩn mã hóa/giải mã hoặc bản đặc tả. Trong giới hạn được quy định bởi cú pháp của profile nhất định, vẫn có thể yêu cầu thay đổi rất lớn về hiệu suất của các bộ mã hóa và bộ giải mã tùy thuộc vào các giá trị được sử dụng bởi các phần tử cú pháp trong luồng bit như kích thước được quy định của các ảnh đã giải mã. Trong nhiều ứng dụng, có thể là không khả thi cũng như không kinh tế khi áp dụng các bộ giải mã có khả năng xử lý tất cả các trường hợp lý thuyết của cú pháp trong một profile cụ thể. Để giải quyết vấn đề này, các cấp độ có thể được sử dụng. Cấp độ có thể được định nghĩa là tập

hợp các ràng buộc xác định đối với các giá trị của các phần tử cú pháp trong luồng bit và các biến số được quy định trong tiêu chuẩn mã hóa/giải mã hoặc bản đặc tả. Các ràng buộc này có thể là các giới hạn đơn giản đối với các giá trị. Ngoài ra hoặc theo cách khác, chúng có thể có dạng các ràng buộc đối với các tổ hợp số học của các giá trị (ví dụ, chiều rộng ảnh nhân với chiều cao ảnh nhân với số lượng ảnh được giải mã mỗi giây). Các thành phần khác để quy định các ràng buộc đối với các cấp độ cũng có thể được sử dụng. Ví dụ, một số trong số các ràng buộc được quy định ở một cấp độ có thể có liên quan đến kích thước ảnh tối đa, tốc độ bit tối đa và tốc độ dữ liệu tối đa xét theo các đơn vị mã hóa, như các khối macro, theo đơn vị thời gian, như giây. Tập hợp cấp độ giống như vậy có thể được xác định cho tất cả các profile. Ưu tiên nếu, ví dụ, làm tăng khả năng tương tác của các thiết bị đầu cuối thực hiện các profile khác nhau mà hầu hết hoặc tất cả các khía cạnh của định nghĩa về mỗi cấp độ có thể là chung giữa các profile khác nhau. Lớp có thể được định nghĩa là hạng mục xác định gồm các ràng buộc cấp độ đối với các giá trị của các phần tử cú pháp trong luồng bit, trong đó các ràng buộc cấp độ được lồng trong lớp và bộ giải mã phù hợp với một lớp và cấp độ nhất định sẽ có khả năng giải mã tất cả các luồng bit mà phù hợp với cùng lớp đó hoặc lớp thấp hơn của cấp độ đó hoặc cấp độ bất kỳ dưới nó.

Trong một số trường hợp, điểm tương thích có thể được định nghĩa là tổ hợp của profile cụ thể và cấp độ cụ thể hoặc tổ hợp của profile cụ thể, lớp cụ thể và cấp độ cụ thể. Cần hiểu rằng điểm tương thích có thể được định nghĩa theo nhiều cách thay thế, trong khi ý định của nó nhằm quy định các đặc điểm và các giới hạn của luồng bit và/hoặc các đặc điểm và các tài nguyên (tối đa) của các bộ giải mã có thể được giữ không đổi.

Đơn vị cơ sở cho đầu vào của 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 ảnh. Ảnh được cung cấp làm đầu vào cho bộ mã hóa còn có thể được gọi là ảnh nguồn, và ảnh được giải mã bởi bộ giải mã có thể được gọi là ảnh đã giải mã.

Mỗi ảnh nguồn và ảnh đã giải mã đều bao gồm một hoặc nhiều mảng mẫu, như một trong số các tập hợp mảng mẫu nêu dưới đây:

Chỉ độ chói (Y) (đơn sắc).

Độ chói và hai sắc độ (YCbCr hoặc YCgCo).

Xanh lục, xanh lam và đỏ (GBR, còn được gọi là RGB).

Các mảng biểu diễn các mẫu màu đơn sắc hoặc màu bộ ba không cụ thể khác (ví dụ, YZX, còn được gọi là XYZ).

Trong phần dưới đây, các mảng này có thể được gọi là độ chói (L hoặc Y) và sắc độ, trong đó hai mảng sắc độ này có thể được gọi là Cb và Cr; bất kể phương pháp biểu diễn màu thực đang được sử dụng. Phương pháp biểu diễn màu thực đang được sử dụng có thể được chỉ ra, ví dụ, trong luồng bit đã mã hóa ví dụ bằng cách sử dụng cú pháp Thông tin khả dụng của video (VUI – Video Usability Information) của H.264/AVC và/hoặc HEVC. Thành phần có thể được định nghĩa là mảng hoặc mẫu đơn lẻ từ một trong ba mảng mẫu (độ chói và hai sắc độ) hoặc mảng hoặc mẫu đơn lẻ của mảng gồm hình ảnh có 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 độ chói và có thể bao gồm các mẫu sắc độ tương ứng. Trường là một tập hợp các hàng mẫu luân phiên của khung và có thể được sử dụng làm đầu vào của bộ mã hoá, khi tín hiệu nguồn được trộn lẫn. Các mảng mẫu sắc độ có thể vắng mặt (và do đó có thể sử dụng việc lấy mẫu đơn sắc) hoặc các mảng mẫu sắc độ có thể được chọn làm mẫu con khi so sánh với các mảng mẫu độ chói. Các định dạng sắc độ có thể được tóm tắt như sau:

Trong việc lấy mẫu đơn sắc, chỉ có một mảng mẫu mà có thể được coi trên danh nghĩa là mảng độ chói.

Trong việc lấy mẫu 4:2:0, mỗi mảng trong số hai mảng sắc độ này có chiều cao bằng một nửa và chiều rộng bằng một nửa của mảng độ chói.

Trong việc lấy mẫu 4:2:2, mỗi mảng trong số hai mảng sắc độ này có cùng chiều cao và chiều dài bằng một nửa chiều dài của mảng độ chói.

Trong việc lấy mẫu 4:4:4, khi không mặt phẳng màu riêng biệt nào được sử dụng, mỗi mảng trong số hai mảng sắc độ này có cùng chiều cao và chiều rộng như mảng độ chói.

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

Khi sử dụng việc lấy mẫu con sắc độ (ví dụ, lấy mẫu sắc độ 4:2:0 hoặc 4:2:2), vị trí của các mẫu sắc độ so với các mẫu độ chói có thể được xác định ở phía bộ mã hóa (ví dụ, dưới dạng bước tiền xử lý hoặc một phần của quy trình mã hóa). Vị trí của các mẫu sắc độ so với vị trí của các mẫu độ chói có thể được xác định trước, ví dụ theo chuẩn mã hóa, như H.264/AVC hoặc HEVC, hoặc có thể được chỉ ra trong luồng bit, ví dụ như dưới dạng một phần thông tin thuộc VUI của H.264/AVC hoặc HEVC.

Quy trình phân vùng có thể được định nghĩa là sự phân chia một tập hợp thành các tập hợp con sao cho mỗi phần tử của tập hợp chỉ thuộc đúng một trong các tập hợp con.

Trong H.264/AVC, khối macro là khối 16x16 của các mẫu độ chói và các khối tương ứng của các mẫu sắc độ. Ví dụ, trong kiểu lấy mẫu 4:2:0, khối macro chứa một khối 8x8 của các mẫu sắc độ trong mỗi thành phần sắc độ. Trong H.264/AVC, hình ảnh được phân vùng thành một hoặc nhiều nhóm lát, và mỗi nhóm lát gồm một hoặc nhiều lát. Trong H.264/AVC, lát gồm số nguyên các khối macro được xếp một cách tuần tự khi quét mảnh trong một 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 mẫu NxN đối với một số giá trị N sao cho việc chia khối cây mã hóa thành các khối mã hóa là việc phân vùng. Khối cây mã hóa (CTB – Coding Tree Block) có thể được định nghĩa là khối mẫu NxN đối với một số giá trị N sao cho việc chia một thành phần thành các khối cây mã hóa là việc phân vùng. Đơn vị cây mã hóa (CTU - Coding Tree Unit) có thể được định nghĩa là khối cây mã hóa của các mẫu độ chói, hai khối cây mã hóa tương ứng của các mẫu sắc độ của hình ảnh có ba mảng mẫu, hoặc khối cây mã hóa của mẫu hình ảnh đơn sắc hoặc hình ảnh được mã hóa bằng cách 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 (CU - Coding Unit)

có thể được định nghĩa là khối mã hóa của các mẫu độ chói, hai khối mã hóa tương ứng của các mẫu sắc độ của hình ảnh có ba mảng mẫu, hoặc khối mã hoá của các mẫu hình ảnh đơn sắc hoặc hình ảnh được mã hóa bằng cách 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 suất cao (HEVC), các hình ảnh video được chia thành các đơn vị mã hóa (CU: coding unit) che phủ diện tích ảnh này. CU bao gồm một hoặc nhiều đơn vị dự đoán (PU: prediction unit) xác định quy trình dự đoán đối với các mẫu trong CU và một hoặc nhiều đơn vị biến đổi (TU – transform unit) xác định quy trình mã hoá lỗi dự đoán đối với các mẫu trong CU này. Thông thường, CU bao gồm khối mẫu hình vuông với kích cỡ có thể lựa chọn từ một tập hợp định trước gồm các kích cỡ CU khả thi. Một CU với kích cỡ tối đa cho phép có thể được gọi là đơn vị mã hóa lớn nhất (LCU) hoặc đơn vị cây mã hóa (CTU) và hình ảnh video được chia thành các LCU không xếp chồng. LCU có thể được chia nhỏ tiếp thành tổ hợp của các CU nhỏ hơn, ví dụ, bằng cách phân tách đệ quy LCU này và các CU kết quả. Mỗi CU kết quả thường có ít nhất một PU và ít nhất một TU liên kết với nó. Mỗi PU và TU có thể được chia nhỏ tiếp thành các PU và TU nhỏ hơn để lần lượt làm tăng độ chi tiết của sự dự đoán và các quy trình mã hoá lỗi dự đoán. Mỗi PU có thông tin dự đoán liên kết với nó để xác định loại dự đoán sẽ được áp dụng cho các hình ảnh điểm trong PU đó (ví dụ, thông tin vectơ chuyển động đối với các PU đã dự đoán liên ảnh và thông tin hướng dự đoán nội ảnh đối với các PU được dự đoán nội ảnh).

Mỗi TU có thể được liên kết với thông tin mô tả quy trình giải mã lỗi dự đoán đối với các mẫu trong TU này (bao gồm ví dụ, thông tin hệ số DCT). Thường được báo hiệu ở mức CU xem liệu việc mã hoá lỗi dự đoán có được áp dụng hay không cho mỗi CU. Trong trường hợp không có lỗi dự đoán dự liên kết với CU này, có thể coi là không có TU nào đối với CU này. Sự phân chia hình ảnh thành các CU, và sự phân chia các CU thành các PU và TU có thể được báo hiệu trong luồng bit, cho phép bộ giải mã tái dựng cấu trúc có định hướng của các đơn vị này.

Trong HEVC, hình ảnh có thể được phân vùng thành các ô, có hình chữ nhật và gồm số nguyên các LCU. Trong HEVC, sự phân vùng thành các ô tạo thành lưới có quy

tắc, trong đó chiều cao và chiều rộng của các ô khác nhau tối đa là một LCU. Trong HEVC, lát được định nghĩa là số nguyên đơn vị các cây mã hóa chứa trong một đoạn lát độc lập và tất cả các đoạn lát phụ thuộc sau đó (nếu có) ở trước đoạn lát độc lập tiếp theo (nếu có) trong cùng đơn vị truy cập. Trong HEVC, đoạn lát được định nghĩa là số nguyên các đơn vị cây mã hóa được xếp một cách tuần tự theo trình tự quét ô và chứa trong một đơn vị NAL. Sự phân chia mỗi hình ảnh thành các đoạn lát là việc phân vùng. Trong HEVC, đoạn lát độc lập được xác định là đoạn lát mà đối với các đoạn lát này, giá trị của các phần tử cú pháp của phần đầu đoạn lát không suy ra được từ các giá trị của đoạn lát trước đó, và đoạn lát phụ thuộc được định nghĩa là đoạn lát mà đối với các đoạn lát này, các giá trị của một số phần tử cú pháp của phần đầu đoạn lát suy ra được từ các giá trị của đoạn lát độc lập trước đó theo trình tự giải mã. Trong HEVC, phần đầu lát được định nghĩa là phần đầu đoạn lát của đoạn lát độc lập mà là đoạn lát hiện thời hoặc là đoạn lát độc lập ở trước đoạn lát phụ thuộc hiện thời, và phần đầu đoạn lát được định nghĩa là một phần của đoạn lát đã mã hoá chứa các phần tử dữ liệu gắn với 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 trình tự quét mảnh của các LCU trong các ô hoặc trong hình ảnh, nếu các ô không được sử dụng. Trong một LCU, các CU có một trình tự quét cụ thể. Fig.6 thể hiện ví dụ về hình ảnh gồm hai ô được phân vùng thành các đơn vị mã hóa hình vuông (các đường nét liền), các đơn vị này còn được phân vùng thành các đơn vị mã hóa hình chữ nhật (các đường nét đứt).

Bộ giải mã có thể tái dựng video đầu ra bằng cách áp dụng phương tiện dự đoán tương tự với bộ mã hóa để tạo thành dạng biểu diễn đã dự đoán của các khối ảnh điểm (sử dụng thông tin chuyển động hoặc không gian được tạo ra bởi bộ mã hóa và được lưu trữ dưới dạng biểu diễn nén) và giải mã lỗi dự đoán (sự vận hành ngược quy trình mã hoá lỗi dự đoán để khôi phục tín hiệu lỗi dự đoán được lượng tử hóa trong miền điểm ảnh không gian). Sau khi áp dụng sự dự đoán và phương tiện giải mã lỗi dự đoán, bộ giải mã có thể tổng hợp sự dự đoán và các tín hiệu lỗi dự đoán (các giá trị điểm ảnh) để tạo thành khung video đầu ra. Bộ giải mã (và bộ mã hóa) cũng có thể áp dụng phương tiện lọc bổ sung để cải thiện chất lượng của video đầu ra trước khi chuyển tiếp nó để hiển thị và/hoặc lưu trữ làm tham chiếu dự đoán cho các khung tiếp đến trong chuỗi video.

Bước lọc có thể bao gồm thêm một trong các bước sau, ví dụ: tách khối, độ lệch tương thích mẫu (SAO – Sample Adaptive Offset), và/hoặc lọc vòng lặp tương thích (ALF – Adaptive Loop Filtering). H.264/AVC bao gồm bước tách khối, trong khi HEVC bao gồm cả bước tách khối và SAO.

Thay vì, hoặc ngoài các phương pháp tiếp cận sử dụng sự dự đoán giá trị mẫu và mã hoá biến đổi để chỉ ra các giá trị mẫu đã mã hoá, thì mã hoá dựa trên bảng màu có thể được sử dụng. Mã hoá dựa trên bảng màu đề cập đến tập hợp các phương pháp tiếp cận mà bảng màu, nghĩa là bộ màu và các chỉ số liên quan, được xác định và giá trị đối với mỗi mẫu trong đơn vị mã hóa được biểu hiện bằng cách chỉ ra chỉ số của mẫu trong bảng màu này. Mã hoá dựa trên bảng màu có thể đạt được hiệu quả mã hoá tốt trong các đơn vị mã hóa với số lượng màu tương đối nhỏ (như các diện tích hình ảnh biểu diễn nội dung màn hình máy tính, như văn bản hoặc đồ họa đơn giản). Để cải thiện hiệu quả mã hóa của mã hoá bảng màu, nhiều loại phương pháp tiếp cận dự đoán chỉ số bảng màu có thể được sử dụng, hoặc các chỉ số bảng màu có thể được mã hoá loạt dài để có thể biểu diễn diện tích hình ảnh đồng nhất lớn hơn một cách hiệu quả. Hơn nữa, trong trường hợp CU gồm các giá trị mẫu không lặp lại trong CU, mã hoá thoát (escape) có thể được sử dụng. Các mẫu đã được mã hoá thoát được truyền đi mà không đề cập đến các chỉ số bảng màu bất kỳ. Thay vào đó, các giá trị của chúng được chỉ ra riêng lẻ đối với mỗi mẫu đã mã hoá thoát.

Bộ đệm hình ảnh đã giải mã (DPB - Decoded Picture Buffer) có thể được sử dụng trong bộ mã hóa và/hoặc trong bộ giải mã. Có hai lý do để đệm các hình ảnh đã giải mã, để tham chiếu khi dự đoán liên ảnh và để sắp xếp lại các hình ảnh đã giải mã thành trình tự đầu ra. Nhờ H.264/AVC và HEVC cung cấp độ linh hoạt tốt để vừa đánh dấu hình ảnh tham chiếu vừa sắp xếp lại đầu ra, nên các bộ đệm riêng biệt để đệm hình ảnh tham chiếu và đệm hình ảnh đầu ra có thể gây lãng phí tài nguyên bộ nhớ. Do đó, DPB có thể bao gồm quy trình đệm hình ảnh đã giải mã hợp nhất để tham chiếu hình ảnh và sắp xếp lại đầu ra. Hình ảnh đã giải mã có thể được loại bỏ khỏi DPB khi nó không còn được sử dụng làm tham chiếu và không cần cho ra.

Trong nhiều chế độ mã hoá của H.264/AVC và HEVC, hình ảnh tham chiếu để dự đoán liên ảnh được chỉ báo bởi chỉ số của danh mục hình ảnh tham chiếu. Chỉ số này

có thể được mã hoá với việc mã hoá có độ dài thay đổi, thường làm cho chỉ số nhỏ hơn có giá trị ngắn hơn đối với phần tử cú pháp tương ứng. Trong H.264/AVC và HEVC, hai danh mục hình ảnh tham chiếu (danh mục hình ảnh tham chiếu 0 và danh mục hình ảnh tham chiếu 1) được tạo ra cho mỗi lát dự đoán đôi (B), và một danh mục hình ảnh tham chiếu (danh mục hình ảnh tham chiếu 0) được tạo ra đối với mỗi lát đã hóa liên ảnh (P).

Danh mục hình ảnh tham chiếu, như danh mục hình ảnh tham chiếu 0 và danh mục hình ảnh tham chiếu 1, thường được cấu tạo theo hai bước: Thứ nhất, danh mục hình ảnh tham chiếu ban đầu được tạo ra. Danh mục hình ảnh tham chiếu ban đầu này có thể được tạo ra ví dụ dựa trên `frame_num`, `POC`, `temporal_id` (hoặc `TemporalId` hoặc dạng tương tự), hoặc thông tin về hệ phân cấp dự đoán như cấu trúc GOP, hoặc tổ hợp bất kỳ của chúng. Thứ hai, danh mục hình ảnh tham chiếu ban đầu này có thể được sắp xếp lại bởi các lệnh sắp xếp lại danh mục hình ảnh tham chiếu (RPLR - reference picture list reordering), còn được gọi là cấu trúc cú pháp chuyển đổi danh mục hình ảnh tham chiếu, mà có thể được chứa trong các phần đầu lát. Trong H.264/AVC, các câu lệnh RPLR chỉ ra các hình ảnh được sắp xếp về phần mở đầu của danh mục hình ảnh tham chiếu tương ứng. Bước thứ hai này cũng có thể được gọi là quy trình chuyển đổi danh mục hình ảnh tham chiếu và các câu lệnh RPLR có thể được bao gồm trong cấu trúc cú pháp chuyển đổi danh mục hình ảnh tham chiếu. Nếu các tập hợp hình ảnh tham chiếu được sử dụng, danh mục hình ảnh tham chiếu 0 có thể được mở đầu để đầu tiên gồm `RefPicSetStCurr0`, sau đó là `RefPicSetStCurr1`, sau đó là `RefPicSetLtCurr`. Danh mục hình ảnh tham chiếu 1 có thể được mở đầu để đầu tiên gồm `RefPicSetStCurr1`, sau đó là `RefPicSetStCurr0`. Trong HEVC, các danh mục hình ảnh tham chiếu ban đầu này có thể được chuyển đổi thông qua cấu trúc cú pháp chuyển đổi danh mục hình ảnh tham chiếu, trong đó các hình ảnh trong các danh mục hình ảnh tham chiếu ban đầu này có thể được nhận dạng thông qua một chỉ số mục nhập vào danh mục này. Nói cách khác, trong HEVC, chuyển đổi danh mục hình ảnh tham chiếu được mã hóa thành cấu trúc cú pháp bao gồm vòng lặp đối với mỗi mục nhập trong danh mục hình ảnh tham chiếu cuối cùng, trong đó mỗi mục nhập vòng lặp là chỉ số được mã hoá có độ dài không đổi vào danh mục hình ảnh tham chiếu ban đầu này và chỉ ra hình ảnh theo trình tự vị trí từ dưới lên trong danh mục hình ảnh tham chiếu cuối cùng.

Nhiều tiêu chuẩn mã hóa, bao gồm H.264/AVC và HEVC, có thể có quy trình giải mã để suy ra chỉ số hình ảnh tham chiếu vào danh mục hình ảnh tham chiếu, các tiêu chuẩn này có thể được sử dụng để chỉ ra một trong số nhiều hình ảnh tham chiếu mà 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 luồng bit theo một số chế độ mã hóa liên ảnh hoặc có thể được suy ra (bởi bộ mã hóa và bộ giải mã) ví dụ, bằng cách sử dụng các khối lân cận theo một số chế độ mã hóa liên ảnh khác.

Thông tin chuyển động có thể được chỉ ra trong các bộ mã hóa - giải mã video có các vectơ chuyển động được liên kết với mỗi khối ảnh được bù chuyển động. Mỗi vectơ trong số các vectơ chuyển động này có thể biểu diễn sự dịch chuyển của khối ảnh trong hình ảnh cần được mã hóa (ở phía bộ mã hoá) hoặc được giải mã (ở phía bộ giải mã) và khối nguồn dự đoán trong một trong số các hình ảnh đã mã hoá hoặc đã giải mã trước đó. Để biểu diễn các vectơ chuyển động một cách hiệu quả, các vectơ đó có thể được mã hóa khác so với các vectơ chuyển động đã dự đoán cụ thể về khối. Trong các bộ mã hóa - giải mã video, các vectơ chuyển động đã được dự đoán có thể được tạo ra theo cách định trước, ví dụ, bằng cách tính toán trung vị của các vectơ chuyển động được mã hoá hoặc được giải mã của các khối liền kề. Cách khác để tạo ra các dự đoán vectơ chuyển động là tạo ra danh mục các dự đoán ứng viên từ các khối liền kề và/hoặc các khối đồng vị trong các hình ảnh tham chiếu về mặt thời gian và báo hiệu ứng viên được chọn làm dự đoán vectơ chuyển động. Ngoài việc dự đoán các giá trị vectơ chuyển động, chỉ số tham chiếu của hình ảnh đã mã hoá/giải mã trước đó có thể được dự đoán. Chỉ số tham chiếu này có thể được dự đoán, ví dụ, từ các khối liền kề và/hoặc các khối đồng vị trong hình ảnh tham chiếu về mặt thời gian. Hơn nữa, các bộ mã hóa - giải mã video hiệu suất cao có thể sử dụng cơ chế mã hóa/giải mã thông tin chuyển động bổ sung, được gọi là “chế độ hợp nhất”, trong đó tất cả trường thông tin chuyển động, các trường này có thể bao gồm vectơ chuyển động và chỉ số hình ảnh tham chiếu tương ứng đối với mỗi danh mục hình ảnh tham chiếu khả dụng, được dự đoán và sử dụng mà không có chuyển đổi/sửa đổi nào. Tương tự, dự đoán trường thông tin chuyển động có thể được thực hiện bằng cách sử dụng trường thông tin chuyển động của các khối liền kề và/hoặc các khối đồng vị trong các hình ảnh tham chiếu về mặt thời gian và trường thông tin chuyển động đã sử dụng được báo hiệu trong số danh mục của danh mục ứng

viên trường chuyển động có trường thông tin chuyển động của các khối liền kề/đồng vị khả dụng.

Một số bộ mã hóa - giải mã video có thể sử dụng sự dự đoán đơn, trong đó một khối dự đoán được sử dụng để một khối được mã hoá (giải mã), và dự đoán đôi, trong đó hai khối dự đoán được tạo tổ hợp để tạo thành dự đoán để một khối được mã hoá (giải mã). Một số bộ mã hóa - giải mã video có thể dự đoán có trọng số, trong đó các giá trị mẫu của khối dự đoán được xác định trọng số trước khi bổ sung thông tin dư. Ví dụ, có thể áp dụng hệ số trọng số nhân và độ lệch bổ sung. Theo các dự đoán có trọng số rõ ràng được cho phép bởi một số bộ mã hóa - giải mã video, hệ số có trọng số và độ lệch có thể được mã hóa, ví dụ, trong phần đầu lát đối với mỗi chỉ số hình ảnh tham chiếu được phép. Theo sự dự đoán có trọng số ẩn, được cho phép bởi một số bộ mã hóa - giải mã video, thì các hệ số có trọng số và/hoặc độ lệch không được mã hoá nhưng được suy ra, ví dụ, dựa trên khoảng cách đếm trình tự hình ảnh tương đối (POC) của các hình ảnh tham chiếu.

Ngoài việc áp dụng bù chuyển động để dự đoán liên ảnh, phương pháp tiếp cận tương tự có thể được áp dụng vào chế độ dự đoán nội ảnh. Trong trường hợp này, vector dịch chuyển chỉ ra vị trí khối mẫu có thể được sao chép từ cùng một hình ảnh để tạo thành sự dự đoán về khối cần mã hóa hoặc giải mã. Loại phương pháp sao chép khối nội ảnh này có thể cải thiện hiệu quả mã hóa căn bản là khi có mặt các cấu trúc lặp lại trong khung – như văn bản hoặc các hình hoạ khác.

Trong các bộ mã hóa - giải mã video, dự đoán dư sau khi bù chuyển động có thể đầu tiên được biến đổi bằng nhân biến đổi (ví dụ, DCT) và sau đó được mã hóa. Lý do của dự đoán này là có thể vẫn tồn tại một số tương quan giữa phần dư và có thể biến đổi trong nhiều trường hợp giúp làm giảm mối tương quan này và cung cấp phương pháp mã hoá hiệu quả hơn.

Các bộ mã hóa video có thể sử dụng các hàm giá trị Lagrangian để tìm ra chế độ mã hoá tối ưu, ví dụ, chế độ khối macro mong muốn và các vector chuyển động liên quan. Loại hàm giá trị này sử dụng hệ số có trọng số λ (lamda) để liên kết biến dạng hình ảnh (chính xác hoặc ước lượng) do các phương pháp mã hoá gây hao tổn và lượng

thông tin (chính xác hoặc ước lượng) cần để biểu diễn các giá trị điểm ảnh với nhau trên diện tích ảnh là:

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

Trong đó C là giá trị Lagrangian sẽ được giảm thiểu, D là biến dạng hình ảnh (ví dụ, sai số bình phương trung bình) với chế độ và các vector chuyển động đã được xem xét, và R là số chỉ số lượng bit cần để biểu diễn dữ liệu cần thiết để tái dựng khối ảnh trong bộ giải mã (bao gồm lượng dữ liệu để biểu diễn các vector chuyển động ứng viên).

Các tiêu chuẩn mã hóa video và bản đặc tả có thể cho phép các bộ mã hóa phân chia hình ảnh đã mã hóa thành các lát đã mã hoá hoặc dạng tương tự. Dự đoán trong ảnh thường mất tác dụng qua các giới hạn lát. Do đó, các lát có thể được coi là một cách để phân tách hình ảnh đã mã hóa thành các phần có thể giải mã được một cách độc lập. Trong H.264/AVC và HEVC, dự đoán trong ảnh có thể mất tác dụng qua các giới hạn lát. Bởi vậy, các lát có thể được coi là một cách để phân tách hình ảnh đã mã hóa thành các phần có thể giải mã được một cách độc lập, và do đó các lát thường được coi là các đơn vị cơ sở để truyền phát. Trong nhiều trường hợp, các bộ mã hóa có thể chỉ ra trong luồng bit các loại dự đoán trong ảnh mà mất tác dụng qua các giới hạn lát, và sự vận hành bộ giải mã xem xét thông tin này, ví dụ, khi kết luận các nguồn dự đoán nào là khả dụng. Ví dụ, các mẫu từ khối macro hoặc CU lân cận có thể được coi là không khả dụng để dự đoán nội ảnh, nếu khối macro hoặc CU lân cận này nằm trong lát khác.

Đơn vị cơ sở 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 lần lượt là đơn vị lớp trừu tượng mạng (NAL – Network Abstraction Layer). Để chuyển tải qua các mạng theo gói hoặc lưu trữ thành các tệp có cấu trúc, các đơn vị NAL có thể được bao thành các gói hoặc các cấu trúc tương tự. Định dạng luồng bit đã được định rõ trong H.264/AVC và HEVC để tạo các môi trường truyền và lưu trữ không cung cấp các cấu trúc khung. Định dạng luồng bit tách biệt các đơn vị NAL với nhau bằng cách gắn mã khởi đầu ở trước mỗi đơn vị NAL. Để tránh xác định sai về các giới hạn đơn vị NAL, các bộ mã hóa chạy thuật toán ngăn cản mô phỏng mã khởi đầu theo bit, thêm bit ngăn cản mô phỏng vào trọng tải đơn vị NAL nếu mã khởi đầu xuất hiện theo cách khác. Để cho phép vận hành công truyền thẳng giữa hệ thống theo gói và theo dòng, việc ngăn mô phỏng mã khởi đầu có thể luôn

được thực hiện bất kể là dù định dạng luồng bit có đang được sử dụng hay không. Đơn vị NAL có thể được định nghĩa là cấu trúc cú pháp chứa chỉ báo về loại dữ liệu để theo dõi và các bit chứa dữ liệu đó ở dạng RBSP được đặt rải rác khi cần thiết giữa các bit ngăn cản mô phỏng. Tải chuỗi bit thô (RBSP) có thể được định nghĩa là cấu trúc cú pháp chứa số nguyên các bit được bao trong đơn vị NAL. RBSP hoặc là rỗng hoặc có dạng chuỗi dữ liệu bit chứa các phần tử cú pháp được theo sau bởi một bit dừng RBSP và được theo sau bởi không hoặc nhiều bit sau đó bằng 0.

Các đơn vị NAL gồm phần đầu và phần tải. Trong H.264/AVC và HEVC, phần đầu đơn vị NAL chỉ ra loại đơn vị NAL.

Phần đầu đơn vị NAL trong H.264/AVC bao gồm phần tử cú pháp `nal_ref_idc` gồm 2 bit, mà khi bằng 0 thể hiện rằng lát đã mã hoá chứa trong đơn vị NAL là một phần của hình ảnh không tham chiếu và khi lớn hơn 0 chỉ báo rằng lát đã mã hoá chứa trong đơn vị NAL là một phần của hình ảnh tham chiếu. Phần mở đầu cho các đơn vị NAL SVC và MVC có thể chứa thêm nhiều chỉ báo liên quan đến khả năng mở rộng và phân cấp đa chiều.

Trong HEVC, phần đầu đơn vị NAL gồm hai bit được sử dụng cho tất cả các loại đơn vị NAL đã định rõ. Phần đầu đơn vị NAL chứa một bit dự trữ, chỉ báo loại đơn vị NAL gồm sáu bit, chỉ báo `nuh_temporal_id_plus1` gồm ba bit đối với các mức thời gian (có thể cần phải lớn hơn hoặc bằng 1) và phần tử cú pháp `nuh_layer_id` gồm sáu bit. Phần tử cú pháp `temporal_id_plus1` có thể được coi là ký tự nhận dạng thời gian đối với đơn vị NAL, và biến số `TemporalId` dựa trên không có thể được suy ra như sau: $\text{TemporalId} = \text{temporal_id_plus1} - 1$. `TemporalId` bằng 0 tương ứng với mức thời gian thấp nhất. Giá trị của `temporal_id_plus1` cần khác không để tránh mô phỏng mã khởi đầu bao gồm hai bit đầu đơn vị NAL. Luồng bit được tạo ra bằng cách loại trừ tất cả các đơn vị VCL NAL có `TemporalId` lớn hơn hoặc bằng giá trị được lựa chọn và bao gồm tất cả các đơn vị VCL NAL khác vẫn thích hợp. Do đó, hình ảnh có `TemporalId` bằng TID không sử dụng hình ảnh bất kỳ có `TemporalId` lớn hơn TID làm tham chiếu dự đoán liên ảnh. Lớp con hoặc lớp con theo thời gian có thể được xác định là lớp có thể mở rộng theo thời gian của luồng bit có thể mở rộng về mặt thời gian, gồm các đơn vị VCL NAL với giá trị cụ thể của biến số `TemporalId` và các đơn vị NAL khác VCL liên quan.

Phần tử cú pháp `nuh_layer_id` của HEVC có thể mang thông tin về phân cấp khả năng mở rộng.

Các đơn vị NAL có thể được phân loại thành các đơn vị NAL lớp mã hoá video (VCL) và các đơn vị NAL khác VCL. Các đơn vị VCL NAL thường là các đơn vị NAL lát đã mã hoá. Trong H.264/AVC, các đơn vị NAL lát đã mã hoá chứa các phần tử cú pháp biểu diễn một hoặc nhiều khối macro đã mã hóa, mỗi phần tử cú pháp này tương ứng với khối mẫu trong hình ảnh đã giải nén. Trong HEVC, các đơn vị VCL NAL chứa các phần tử cú pháp biểu diễn một hoặc nhiều CU.

Trong H.264/AVC, đơn vị NAL lát đã mã hoá có thể được chỉ báo là lát đã mã hoá trong hình ảnh làm mới quy trình giải mã tức thời (IDR - Instantaneous Decoding Refresh) hoặc lát đã mã hoá trong hình ảnh khác IDR.

Trong HEVC, đơn vị NAL lát đã mã hoá có thể được chỉ ra là một trong số các dạng nêu dưới 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 đã mã hoá của hình ảnh tạo vệt khác TSA, khác STSA <code>slice_segment_layer_rbsp()</code>
2, 3	TSA_N, TSA_R	Đoạn lát đã mã hoá của hình ảnh TSA <code>slice_segment_layer_rbsp()</code>
4, 5	STSA_N, STSA_R	Đoạn lát đã mã hoá của hình ảnh STSA <code>slice_layer_rbsp()</code>
6, 7	RADL_N, RADL_R	Đoạn lát đã mã hoá của hình ảnh RADL <code>slice_layer_rbsp()</code>
8, 9	RASL_N, RASL_R,	Đoạn lát đã mã hoá của hình ảnh RASL <code>slice_layer_rbsp()</code>
10, 12, 14	RSV_VCL_N10 RSV_VCL_N12 RSV_VCL_N14	Các loại đơn vị VCL NAL không tham chiếu dự trữ // khác RAP dự trữ
11, 13, 15	RSV_VCL_R11 RSV_VCL_R13 RSV_VCL_R15	Các loại đơn vị VCL NAL tham chiếu dự trữ // khác RAP dự trữ

16, 17, 18	BLA_W_LP BLA_W_DLP (còn được gọi là IDR_W_RADL) BLA_N_LP	Đoạn lát đã mã hoá của hình ảnh BLA slice_segment_layer_rbsp()
19, 20	IDR_W_DLP (còn được gọi là IDR_W_RADL) IDR_N_LP	Đoạn lát đã mã hoá của hình ảnh IDR slice_segment_layer_rbsp()
21	CRA_NUT	Đoạn lát đã mã hoá của hình ảnh CRA slice_segment_layer_rbsp()
22, 23	RSV_IRAP_VCL22.. RSV_IRAP_VCL23	Các loại đơn vị VCL NAL dự trữ // RAP dự trữ
24..31	RSV_VCL24.. RSV_VCL31	Các loại đơn vị VCL NAL dự trữ //khác RAP dự trữ

Trong HEVC, các ký hiệu viết tắt cho các loại hình ảnh có thể được xác định như sau: hình ảnh tạo vệt (TRAIL), Truy cập lớp con theo thời gian (TSA), Truy cập lớp con theo thời gian theo bước (STSA), Hình ảnh dẫn đầu có thể giải mã truy cập ngẫu nhiên (RADL - Random Access Decodable Leading), hình ảnh dẫn đầu bỏ qua truy cập ngẫu nhiên (RASL - Random Access Skipped Leading), hình ảnh truy cập link gãy (BLA - Broken Link Access), hình ảnh làm mới quy trình giải mã tức thời (IDR), hình ảnh truy cập ngẫu nhiên không lỗi (CRA – Clean Random Access).

Hình ảnh điểm truy cập ngẫu nhiên (RAP: Random Access Point), còn có thể được gọi là hình ảnh điểm truy cập ngẫu nhiên nội ảnh (IRAP: intra random access point), là hình ảnh trong đó mỗi lát hoặc đoạn lát có nal_unit_type trong khoảng từ 16 đến 23. Hình ảnh IRAP trong một lớp độc lập chỉ gồm các lát đã mã hoá nội ảnh. Hình ảnh IRAP thuộc lớp đã dự đoán với giá trị nuh_layer_id là currLayerId có thể gồm các lát P, B, và I, không thể sử dụng sự dự đoán liên ảnh từ các hình ảnh khác với nuh_layer_id bằng currLayerId, và có thể sử dụng sự dự đoán liên lớp từ các lớp tham chiếu trực tiếp của nó. Theo phiên bản HEVC này, hình ảnh IRAP có thể là hình ảnh BLA, hình ảnh CRA hoặc hình ảnh IDR. Hình ảnh thứ nhất trong luồng bit chứa một lớp cơ sở là hình ảnh IRAP. Miễn là các tập tham số cần thiết là khả dụng khi cần được kích hoạt, hình ảnh IRAP ở lớp độc lập và tất cả các hình ảnh khác RASL tiếp theo ở

lớp độc lập này theo trình tự giải mã có thể được giải mã chính xác mà không thực hiện quy trình giải mã bất kỳ hình ảnh nào ở trước hình ảnh IRAP theo trình tự giải mã. Có thể có các hình ảnh trong luồng bit chỉ gồm các lát đã mã hoá nội ảnh mà không phải là các hình ảnh RAP. Hình ảnh IRAP thuộc lớp đã dự đoán với giá trị `nuh_layer_id` là `currLayerId` và tất cả các hình ảnh khác RASL tiếp theo với `nuh_layer_id` bằng `currLayerId` theo trình tự giải mã có thể được giải mã chính xác mà không thực hiện quy trình giải mã bất kỳ hình ảnh nào với `nuh_layer_id` bằng `currLayerId` ở trước hình ảnh IRAP theo trình tự giải mã, khi các tập tham số cần thiết là khả dụng khi cần được kích hoạt và khi quy trình giải mã mỗi lớp tham chiếu trực tiếp của lớp có `nuh_layer_id` bằng `currLayerId` đã được bắt đầu (nghĩa là khi `LayerInitializedFlag[ refLayerId ]` bằng 1 với `refLayerId` bằng tất cả các giá trị `nuh_layer_id` của các lớp tham chiếu trực tiếp của lớp có `nuh_layer_id` bằng `currLayerId`).

Trong HEVC, hình ảnh CRA có thể là hình ảnh thứ nhất trong luồng bit theo trình tự giải mã, hoặc có thể xuất hiện muộn hơn trong luồng bit. Các hình ảnh CRA trong HEVC cho phép cái gọi là các hình ảnh dẫn đầu mà theo sau hình ảnh CRA theo trình tự giải mã nhưng ở trước nó theo trình tự đầu ra. Một số hình ảnh dẫn đầu, được gọi là các hình ảnh RASL, có thể sử dụng các hình ảnh được giải mã trước hình ảnh CRA làm tham chiếu. Các hình ảnh theo sau hình ảnh CRA trong cả quy trình giải mã và trình tự đầu ra có thể giải mã được nếu truy cập ngẫu nhiên được thực hiện ở hình ảnh CRA, và do đó có thể đạt được truy cập ngẫu nhiên không lỗi tương tự với chức năng truy cập ngẫu nhiên không lỗi của hình ảnh IDR.

Hình ảnh CRA có thể có các hình ảnh RADL hoặc RASL liên quan. Khi hình ảnh CRA là hình ảnh thứ nhất trong luồng bit theo trình tự giải mã, hình ảnh CRA này là hình ảnh thứ nhất của chuỗi video đã mã hoá theo trình tự giải mã, và bất kỳ các hình ảnh RASL liên quan nào không được đưa ra bởi bộ giải mã và có thể không giải mã được, do chúng có thể gồm các tham chiếu đến các hình ảnh không có mặt trong luồng bit.

Hình ảnh dẫn đầu là hình ảnh ở trước hình ảnh RAP liên quan theo trình tự đầu ra. Hình ảnh RAP liên quan này là hình ảnh RAP trước đó theo trình tự giải mã (nếu có mặt). Hình ảnh dẫn đầu có thể hoặc là hình ảnh RADL hoặc là hình ảnh RASL.

Tất cả các hình ảnh RASL là các hình ảnh dẫn đầu của hình ảnh BLA hoặc CRA liên quan. Khi hình ảnh RAP liên quan là hình ảnh BLA hoặc là hình ảnh đã mã hóa thứ nhất trong luồng bit, hình ảnh RASL này không được đưa ra và có thể không giải mã chính xác được, do hình ảnh RASL này có thể gồm các tham chiếu đến các hình ảnh không có mặt trong luồng bit. Tuy nhiên, hình ảnh RASL có thể được giải mã chính xác nếu quy trình giải mã này đã bắt đầu từ hình ảnh RAP trước hình ảnh RAP liên quan của hình ảnh RASL. Các hình ảnh RASL không được sử dụng làm các hình ảnh tham chiếu cho quy trình giải mã các hình ảnh khác RASL. Khi có mặt, tất cả các hình ảnh RASL ở trước, theo trình tự giải mã, tất cả các hình ảnh tạo vệt của cùng hình ảnh RAP liên quan. Theo một số dự thảo của chuẩn HEVC, hình ảnh RASL được đề cập đến hình ảnh được gắn thẻ để loại bỏ (TFD - Tagged for Discard).

Tất cả các hình ảnh RADL là các hình ảnh dẫn đầu. Các hình ảnh RADL không được sử dụng làm các hình ảnh tham chiếu cho quy trình giải mã các hình ảnh tạo vệt của cùng hình ảnh RAP liên quan. Khi có mặt, tất cả các hình ảnh RADL ở trước, theo trình tự giải mã, tất cả các hình ảnh tạo vệt của cùng hình ảnh RAP liên quan. Các hình ảnh RADL không đề cập đến hình ảnh nào ở trước hình ảnh RAP liên quan theo trình tự giải mã và do đó có thể được giải mã chính xác khi quy trình giải mã bắt đầu từ hình ảnh RAP liên quan này. Theo một số dự thảo của chuẩn HEVC, hình ảnh RADL được đề cập đến hình ảnh dẫn đầu có thể giải mã được (DLP – Decodable Leading Picture).

Khi một phần của luồng bit bắt đầu từ hình ảnh CRA được bao gồm trong một luồng bit khác, các hình ảnh RASL liên kết với hình ảnh CRA này có thể không giải mã chính xác được, do một số hình ảnh tham chiếu của chúng có thể không có mặt trong luồng bit đã tạo tổ hợp. Để giúp quy trình vận hành ghép nối truyền thẳng, loại đơn vị NAL của hình ảnh CRA có thể được thay đổi để chỉ báo rằng đây là hình ảnh BLA. Các hình ảnh RASL liên kết với hình ảnh BLA có thể không giải mã chính xác được do đó không được đưa ra/hiển thị. Hơn nữa, các hình ảnh RASL liên kết với hình ảnh BLA có thể được bỏ ra khỏi quy trình giải mã.

Hình ảnh BLA có thể là hình ảnh thứ nhất trong luồng bit theo trình tự giải mã, hoặc có thể xuất hiện muộn hơn trong luồng bit. Mỗi hình ảnh BLA bắt đầu một chuỗi video đã mã hoá mới và có hiệu quả tương tự trên quy trình giải mã như hình ảnh IDR.

Tuy nhiên, hình ảnh BLA có thể chứa các phân tử cú pháp định rõ tập hợp hình ảnh tham chiếu không rỗng. Khi hình ảnh BLA có `nal_unit_type` bằng `BLA_W_LP`, có thể có các hình ảnh RASL liên quan, không được đưa ra bởi bộ giải mã và có thể không giải mã được, do chúng có thể gồm các tham chiếu đến các hình ảnh không có mặt trong luồng bit. Khi hình ảnh BLA có `nal_unit_type` bằng `BLA_W_LP`, nó cũng có thể có các hình ảnh RADL liên quan, được nêu cụ thể để được giải mã. Khi hình ảnh BLA có `nal_unit_type` bằng `BLA_W_DLP`, nó không có các hình ảnh RASL liên quan nhưng có thể có các hình ảnh RADL liên quan, được nêu cụ thể để được giải mã. Khi hình ảnh BLA có `nal_unit_type` bằng `BLA_N_LP`, nó không có các hình ảnh dẫn đầu liên quan nào.

Hình ảnh IDR có `nal_unit_type` bằng `IDR_N_LP` không có các hình ảnh dẫn đầu liên quan có mặt trong luồng bit. Hình ảnh IDR có `nal_unit_type` bằng `IDR_W_LP` không có các hình ảnh RASL liên quan có mặt trong luồng bit, nhưng có thể có các hình ảnh RADL liên quan trong luồng 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 đã giải mã không được sử dụng làm tham chiếu cho bất kỳ hình ảnh nào khác của cùng lớp con theo thời gian. Nghĩa là, trong HEVC, khi giá trị của `nal_unit_type` bằng `TRAIL_N`, `TSA_N`, `STSA_N`, `RADL_N`, `RASL_N`, `RSV_VCL_N10`, `RSV_VCL_N12`, hoặc `RSV_VCL_N14`, hình ảnh đã giải mã không được bao gồm trong bất kỳ giá trị `RefPicSetStCurrBefore`, `RefPicSetStCurrAfter` và `RefPicSetLtCurr` của hình ảnh bất kỳ có cùng giá trị `TemporalId`. Hình ảnh đã 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 được loại bỏ mà không gây ảnh hưởng đối với khả năng giải mã của các hình ảnh có cùng giá trị `TemporalId`.

Hình ảnh tạo vệt có thể được định nghĩa là các hình ảnh theo sau hình ảnh RAP liên quan theo trình tự đầu ra. Hình ảnh nào mà là hình ảnh tạo vệt không có `nal_unit_type` bằng `RADL_N`, `RADL_R`, `RASL_N` hoặc `RASL_R`. Hình ảnh nào mà là hình ảnh dẫn đầu có thể bị buộc phải ở trước, theo trình tự giải mã, tất cả các hình ảnh tạo vệt liên kết với cùng hình ảnh RAP. Không hình ảnh RASL nào có mặt trong luồng

bit mà liên kết với hình ảnh BLA có `nal_unit_type` bằng `BLA_W_DLP` hoặc `BLA_N_LP`. Không có hình ảnh RADL nào trong luồng bit mà liên kết với hình ảnh BLA có `nal_unit_type` bằng `BLA_N_LP` hoặc liên kết với hình ảnh IDR có `nal_unit_type` bằng `IDR_N_LP`. Hình ảnh RASL bất kỳ liên kết với hình ảnh CRA hoặc BLA có thể bị ràng buộc ở trước hình ảnh RADL bất kỳ liên kết với hình ảnh CRA hoặc BLA theo trình tự đầu ra. Hình ảnh RASL bất kỳ liên kết với hình ảnh CRA có thể bị ràng buộc theo sau, theo trình tự đầu ra, hình ảnh RAP bất kỳ khác ở trước hình ảnh CRA này theo trình tự giải mã.

Trong HEVC có hai loại hình ảnh, loại hình ảnh TSA và STSA mà có thể được sử dụng để chỉ ra điểm hoán đổi lớp con theo thời gian. Nếu các lớp con theo thời gian có `TemporalId` lên tới N đã 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 trình tự giải mã) có `TemporalId` bằng $N+1$. Loại hình ảnh TSA có thể tạo giới hạn lên chính hình ảnh TSA và tất cả các hình ảnh trong cùng lớp con theo sau hình ảnh TSA theo trình tự giải mã. Không có hình ảnh nào được cho phép sử dụng dự đoán liên ảnh từ hình ảnh bất kỳ trong cùng lớp con ở trước hình ảnh TSA theo trình tự giải mã. Định nghĩa TSA còn có thể tạo giới hạn lên các hình ảnh trong các lớp con bậc cao hơn theo sau hình ảnh TSA theo trình tự giải mã. Không hình ảnh nào được cho phép đề cập hình ảnh ở trước hình ảnh TSA theo trình tự giải mã nếu hình ảnh đó thuộc cùng lớp con hoặc lớp con bậc cao hơn như hình ảnh TSA. Các hình ảnh TSA có `TemporalId` lớn hơn 0. STSA này tương tự với hình ảnh TSA nhưng không tạo giới hạn lên các hình ảnh trong các lớp con bậc cao hơn theo sau hình ảnh STSA này theo trình tự giải mã và do đó chỉ cho phép hoán đổi lên trên lớp con chứa hình ảnh STSA này.

Ví dụ, đơn vị NAL khác VCL có thể là một trong các loại sau: tập tham số chuỗi, tập tham số hình ảnh, đơn vị NAL thông tin tăng cường bổ sung (SEI), phần phân định đơn vị truy cập, đầu của đơn vị NAL chuỗi, đầu của đơn vị NAL luồng bit, hoặc đơn vị NAL dữ liệu bộ lọc. Các tập tham số có thể cần cho sự tái dựng các hình ảnh đã giải mã, trong khi nhiều đơn vị NAL khác VCL khác không cần thiết cho sự tái dựng của các giá trị mẫu đã giải mã.

Các tham số vẫn không thay đổi qua chuỗi video đã mã hoá có thể được bao gồm trong tập tham số chuỗi. Ngoài các tham số mà có thể là cần thiết cho quy trình giải mã, tập tham số chuỗi này có thể tùy chọn gồm thông tin khả dụng video (VUI), bao gồm các tham số có thể là quan trọng cho quy trình đệm, tính thời gian đưa ra hình ảnh, kết xuất, và dự trữ tài nguyên. Có ba đơn vị NAL được định rõ trong H.264/AVC để mang các tập tham số chuỗi: đơn vị NAL tập tham số chuỗi này chứa tất cả dữ liệu cho các đơn vị VCL NAL H.264/AVC trong chuỗi này, đơn vị NAL mở rộng tập tham số chuỗi chứa dữ liệu cho các hình ảnh phụ đã mã hóa, và tập tham số chuỗi tập hợp con cho các đơn vị VCL NAL MVC và SVC. Trong HEVC, RBSP tập tham số chuỗi bao gồm các tham số mà có thể được đề cập bởi một hoặc nhiều các RBSP tập tham số hình ảnh hoặc một hoặc nhiều các đơn vị SEI NAL chứa thông báo SEI khoảng đệm. Tập tham số hình ảnh gồm các tham số như vậy có thể không thay đổi trong một số hình ảnh đã mã hoá. RBSP tập tham số hình ảnh có thể bao gồm các tham số mà có thể được đề cập bởi các đơn vị NAL lát đã mã hoá của một hoặc nhiều hình ảnh đã mã hoá.

Trong HEVC, tập tham số video (VPS) có thể được định nghĩa là cấu trúc cú pháp chứa các phần tử cú pháp áp dụng vào không hoặc các chuỗi video đã mã hoá hoàn thiện hơn như được xác định bởi nội dung của phần tử cú pháp được tìm thấy ở SPS được đề cập bởi phần tử cú pháp được tìm thấy ở PPS được đề cập bởi phần tử cú pháp được tìm thấy ở mỗi phần đầu đoạn lát.

RBSP tập tham số video có thể bao gồm các tham số mà có thể được đề cập bởi một hoặc nhiều RBSP tập tham số chuỗi.

Mối liên hệ và phân cấp giữa tập tham số video (VPS), tập tham số chuỗi (SPS), và tập tham số hình ảnh (PPS) có thể được mô tả như sau. VPS ở một mức trên SPS theo phân cấp tập tham số và theo phạm vi khả năng mở rộng và/hoặc video 3D. VPS có thể bao gồm các tham số chung cho tất cả các lát qua tất cả các lớp (khả năng mở rộng hoặc hiển thị) trong toàn bộ chuỗi video đã mã hoá. SPS bao gồm các tham số chung cho tất cả các lát trong lớp (khả năng mở rộng hoặc hiển thị) cụ thể trong toàn bộ chuỗi video đã mã hoá, và có thể được dùng chung bởi nhiều lớp (khả năng mở rộng hoặc hiển thị). PPS bao gồm các tham số chung cho tất cả các lát theo dạng biểu diễn lớp cụ thể (dạng

biểu diễn của một khả năng mở rộng hoặc lớp hiển thị trong một đơn vị truy cập) và có thể được dùng chung bởi tất cả các lát theo nhiều dạng biểu diễn lớp.

VPS có thể cung cấp thông tin về các mối liên hệ phụ thuộc của các lớp trong luồng bit, cũng như nhiều thông tin khác mà có thể ứng dụng cho tất cả các lát qua tất cả các lớp (khả năng mở rộng hoặc hiển thị) trong toàn bộ chuỗi video đã mã hoá. VPS có thể được coi là bao gồm ba phần, VPS cơ sở, nội dung mở rộng VPS, và VUI VPS, trong đó nội dung mở rộng VPS và VUI VPS có thể là có mặt một cách tùy ý. Trong HEVC, VPS cơ sở có thể được coi là bao gồm cấu trúc cú pháp `video_parameter_set_rbsp()` mà không có cấu trúc cú pháp `vps_extension()`. Cấu trúc cú pháp `video_parameter_set_rbsp()` đã được định rõ chủ yếu cho HEVC phiên bản 1 và bao gồm các phần tử cú pháp mà có thể là sử dụng để giải mã lớp cơ sở. Trong HEVC, nội dung mở rộng VPS có thể được coi là bao gồm cấu trúc cú pháp `vps_extension()`. Cấu trúc cú pháp `vps_extension()` được định rõ trong HEVC phiên bản 2 chủ yếu cho các nội dung mở rộng đa lớp và bao gồm các phần tử cú pháp mà có thể sử dụng để giải mã một hoặc nhiều các lớp không cơ sở, như các phần tử cú pháp chỉ ra các mối liên hệ phụ thuộc lớp. VUI VPS bao gồm các phần tử cú pháp có thể là hữu ích để giải mã hoặc dùng cho các mục đích khác nhưng không cần phải được sử dụng theo quy trình giải mã HEVC.

Cú pháp H.264/AVC và HEVC cho phép nhiều ví dụ về các tập tham số, và mỗi ví dụ được nhận dạng với bộ nhận dạng duy nhất. Để giới hạn việc sử dụng bộ nhớ cần cho các tập tham số, khoảng giá trị cho các bộ nhận dạng tập tham số đã được giới hạn. Trong H.264/AVC và HEVC, mỗi phần đầu lát bao gồm bộ nhận dạng của tập tham số hình ảnh là hoạt động đối với quy trình giải mã hình ảnh gồm lát này, và mỗi tập tham số hình ảnh gồm bộ nhận dạng của tập tham số chuỗi hoạt động. Do đó, sự truyền hình ảnh và các tập tham số chuỗi không phải được đồng bộ một cách chính xác với sự truyền các lát. Thay vào đó, các tập tham số hình ảnh và chuỗi hoạt động là đủ để thu được tại thời điểm bất kỳ trước khi chúng được tham chiếu, cho phép truyền các tập tham số “ngoài dải tần” bằng cách sử dụng cơ chế truyền đáng tin hơn so với các giao thức được sử dụng cho dữ liệu lát. Ví dụ, các tập tham số có thể được bao gồm dưới dạng tham số trong phần mô tả phiên về các phiên Giao thức chuyển tải thời gian thực (RTP). Nếu các

tập tham số được truyền đi trong dải tần, chúng có thể được lặp lại để cải thiện độ mạnh sai số.

Truyền ngoài dải tần, báo hiệu hoặc thiết bị lưu trữ còn có thể hoặc theo cách khác được sử dụng cho các mục đích khác ngoài việc chống lại các lỗi truyền, như dễ dàng khi truy cập hoặc giao phiên. Ví dụ, mục nhập mẫu của một rãnh ghi trong tệp phù hợp với định dạng tệp đa phương tiện cơ sở ISO có thể bao gồm các tập tham số, trong khi dữ liệu đã mã hoá trong luồng bit được lưu trữ đầu đó trong tệp này hoặc trong một tệp khác. Cụm từ theo luồng bit (ví dụ, chỉ ra theo luồng bit) có thể được sử dụng trong các điểm yêu cầu bảo hộ và các phương án được mô tả để đề cập đến việc truyền, báo hiệu, hoặc thiết bị lưu trữ ngoài dải tần theo cách mà dữ liệu ngoài dải tần được liên kết với luồng bit. Cụm từ giải mã theo luồng bit hoặc tương tự có thể đề cập đến việc giải mã dữ liệu ngoài dải tần đã đề cập (mà có thể thu được từ việc truyền, báo hiệu, hoặc thiết bị lưu trữ ngoài dải tần) mà được liên kết với luồng bit.

Tập tham số có thể được kích hoạt bằng tham chiếu từ một lát hoặc từ một tập tham số hoạt động khác hoặc trong một số trường hợp từ một cấu trúc cú pháp khác như thông báo SEI khoảng đệm.

Đơn vị SEI NAL có thể gồm một hoặc nhiều thông báo SEI, không cần thiết cho quy trình giải mã các hình ảnh đầu ra nhưng có thể hỗ trợ trong các quy trình liên quan, như tính thời gian đưa ra hình ảnh, kết xuất, xác định lỗi, che lỗi, và dự trữ tài nguyên. Một số thông báo SEI được nêu cụ thể trong H.264/AVC và HEVC, và thông báo SEI dữ liệu người dùng cho phép các tổ chức và công ty định rõ thông báo SEI tùy sử dụng. H.264/AVC và HEVC gồm cú pháp và ngữ nghĩa cho thông báo SEI đã định rõ nhưng không xác định được quy trình nào để xử lý các thông báo này ở bên nhận. Do đó, các bộ mã hóa cần phải theo chuẩn H.264/AVC hoặc chuẩn HEVC khi tạo thông báo SEI, và các bộ giải mã phù hợp với chuẩn H.264/AVC hoặc chuẩn HEVC tương ứng, không cần phải xử lý thông báo SEI để hợp cách trình tự đầu ra. Một trong các lý do để bao gồm cú pháp và ngữ nghĩa của thông báo SEI trong H.264/AVC và HEVC là cho phép các bản đặc tả hệ thống khác nhau giải thích thông tin bổ sung y hệt nhau và do đó vận hành với nhau. Được dự định là các bản đặc tả hệ thống có thể yêu cầu việc sử dụng

thông báo SEI cụ thể cả trong đầu mã hoá và trong đầu giải mã, và còn cả quy trình để xử lý thông báo SEI cụ thể ở bên nhận có thể được định rõ.

Một số thông báo SEI lồng đã được nêu trong các chuẩn AVC và HEVC hoặc được đề xuất khác. Ý tưởng về thông báo SEI lồng là để gồm một hoặc nhiều thông báo SEI trong một thông báo SEI lồng và cung cấp cơ chế để liên kết thông báo SEI đã bao gồm với tập con của luồng bit và/hoặc tập con của dữ liệu đã giải mã. Thông báo SEI lồng có thể được yêu cầu là gồm một hoặc nhiều thông báo SEI mà chính các thông báo này không phải là thông báo SEI lồng. Thông báo SEI chứa trong một thông báo SEI lồng có thể được gọi là thông báo SEI đã lồng. Thông báo SEI không được bao gồm trong một thông báo SEI lồng có thể được gọi là thông báo SEI không được lồng. Thông báo SEI lồng có thể mở rộng của HEVC cho phép nhận dạng tập con luồng bit (tạo ra từ quy trình trích xuất luồng bit con) hoặc tập hợp của các lớp mà áp dụng thông báo SEI đã lồng. Tập con luồng bit cũng có thể được gọi là luồng bit con.

Hình ảnh đã mã hóa là dạng biểu diễn đã mã hoá của hình ảnh. Hình ảnh đã mã hóa trong H.264/AVC bao gồm các đơn vị VCL NAL cần cho quy trình giải mã hình ảnh. Trong H.264/AVC, hình ảnh đã mã hóa có thể hình ảnh đã mã hoá sơ cấp hoặc hình ảnh đã mã hóa dư. Các hình ảnh đã mã hóa sơ cấp được sử dụng trong quy trình giải mã luồng bit có hiệu lực, trong khi hình ảnh đã mã hóa dư là dạng biểu diễn dư chỉ nên được giải mã khi các hình ảnh đã mã hóa sơ cấp này không được giả mã thành công. Trong HEVC, không có hình ảnh đã mã hóa dư nào được định rõ.

Trong H.264/AVC, đơn vị truy cập (AU – access unit) bao gồm các hình ảnh đã mã hóa sơ cấp và các đơn vị NAL liên kết với nó. Trong H.264/AVC, trình tự xuất hiện của các đơn vị NAL trong đơn vị truy cập được giới hạn như sau. Đơn vị NAL phân định đơn vị truy cập tùy chọn có thể chỉ ra khởi đầu của đơn vị truy cập. Theo sau đó là không hoặc nhiều đơn vị SEI NAL. Các lát đã mã hoá của các hình ảnh đã mã hóa sơ cấp xuất hiện sau đó. Trong H.264/AVC, lát đã mã hoá của các hình ảnh đã mã hóa sơ cấp có thể được theo sau bởi các lát đã mã hoá cho không hoặc nhiều hình ảnh đã mã hóa dư. Hình ảnh đã mã hóa dư là dạng biểu diễn đã mã hoá của hình ảnh hoặc một phần của hình ảnh. Hình ảnh đã mã hóa dư có thể được giải mã nếu các hình ảnh đã mã hóa

sơ cấp này không thu được bởi bộ giải mã, ví dụ do mất đường truyền hoặc vật ghi vật lý bị hỏng.

Trong H.264/AVC, đơn vị truy cập cũng có thể bao gồm hình ảnh phụ đã mã hóa, hình ảnh này là hình ảnh bổ sung các hình ảnh đã mã hóa sơ cấp và có thể được sử dụng, ví dụ trong quy trình hiển thị. Ví dụ, hình ảnh phụ đã mã hóa có thể được sử dụng làm kênh alpha hoặc mặt phẳng alpha định rõ mức trong suốt của các mẫu trong các hình ảnh đã giải mã. Kênh hoặc mặt phẳng alpha có thể được sử dụng theo kết cấu tạo lớp hoặc hệ thống kết xuất, trong đó hình ảnh đầu ra được tạo ra bằng cách xếp chồng các hình ảnh mà ít nhất là trong suốt một phần lên nhau. Hình ảnh phụ đã mã hóa có cùng các giới hạn cú pháp và ngữ nghĩa như hình ảnh đơn sắc đã mã hóa dư. Trong H.264/AVC, hình ảnh phụ đã mã hóa chứa cùng số lượng khối macro như các hình ảnh đã mã hóa sơ cấp này.

Trong HEVC, hình ảnh đã mã hóa có thể được định nghĩa là dạng biểu diễn đã mã hoá của hình ảnh chứa tất cả các đơn vị cây mã hóa của hình ảnh này. Trong HEVC, đơn vị truy cập (AU – access unit) có thể được định nghĩa là tập hợp của các đơn vị NAL liên kết với nhau theo quy định phân cấp đã định rõ, là tuần tự theo trình tự giải mã, và gồm nhiều nhất là một hình ảnh với giá trị `nuh_layer_id` bất kỳ. Ngoài việc chứa các đơn vị VCL NAL của hình ảnh đã mã hóa, đơn vị truy cập cũng có thể gồm các đơn vị NAL khác VCL.

Luồng bit có thể được định nghĩa là chuỗi bit, ở dạng luồng đơn vị NAL hoặc luồng bit, tạo thành dạng biểu diễn của các hình ảnh đã mã hóa và dữ liệu liên quan tạo thành một hoặc nhiều chuỗi video đã mã hoá. Luồng bit thứ nhất có thể được theo sau bởi luồng bit thứ hai trên cùng một kênh logic, như trong cùng một tệp hoặc có cùng kết nối về giao thức truyền thông. Luồng cơ bản (trong trường hợp mã hóa video) có thể được định nghĩa là chuỗi của một hoặc nhiều luồng bit. Đầu của luồng bit thứ nhất có thể được chỉ ra bởi đơn vị NAL cụ thể, có thể được gọi là đơn vị NAL đầu luồng bit (EOB) và là đơn vị NAL cuối cùng của luồng bit. Trong HEVC và các nội dung mở rộng dự thảo hiện tại của nó, đơn vị NAL EOB cần phải có `nuh_layer_id` bằng 0.

Trong H.264/AVC, chuỗi video đã mã hoá được xác định là chuỗi của các đơn vị truy cập liên tiếp theo trình tự giải mã từ một đơn vị truy cập IDR, bao gồm cả đơn vị

này, đến đơn vị truy cập IDR tiếp theo, loại trừ đơn vị này, hoặc đến đầu của luồng bit, bất kể cái nào xuất hiện trước.

Trong HEVC, chuỗi video đã mã hoá (CVS – Coded Video Sequence) có thể được xác định, ví dụ, làm chuỗi của các đơn vị truy cập gồm, theo trình tự giải mã, đơn vị truy cập IRAP với NoRaslOutputFlag bằng 1, sau đó là 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 sau đó đến tận, nhưng không bao gồm, bất kỳ đơn vị truy cập nào sau đó mà là đơn vị truy cập IRAP có NoRaslOutputFlag bằng 1. Đơn vị truy cập IRAP có thể được định nghĩa là đơn vị truy cập có hình ảnh lớp cơ sở là hình ảnh IRAP. Giá trị của NoRaslOutputFlag bằng 1 đối với mỗi hình ảnh IDR, mỗi hình ảnh BLA, và mỗi hình ảnh IRAP là hình ảnh thứ nhất trong lớp cụ thể đó trong luồng bit theo trình tự giải mã, là hình ảnh IRAP thứ nhất theo sau đầu của đơn vị NAL chuỗi có cùng giá trị `nuh_layer_id` theo trình tự giải mã. Trong HEVC đa lớp, giá trị của NoRaslOutputFlag bằng 1 đối với mỗi hình ảnh IRAP khi `nuh_layer_id` của nó là giá trị sao cho `LayerInitializedFlag[ nuh_layer_id ]` bằng 0 và `LayerInitializedFlag[ refLayerId ]` bằng 1 với tất cả các giá trị `refLayerId` bằng `IdDirectRefLayer[ nuh_layer_id ][ j ]`, trong đó `j` nằm trong khoảng từ 0 đến `NumDirectRefLayers[ nuh_layer_id ] - 1`, khoảng này bao gồm cả hai giá trị đầu khoảng. Mặt khác, giá trị của NoRaslOutputFlag bằng `HandleCraAsBlaFlag`. NoRaslOutputFlag bằng 1 có tác động rằng các hình ảnh RASL liên kết với hình ảnh IRAP mà NoRaslOutputFlag là tập hợp không được đưa ra bởi bộ giải mã. Có thể có các phương tiện để cung cấp giá trị của `HandleCraAsBlaFlag` cho bộ giải mã từ thực thể bên ngoài, như bộ trình phát hoặc bộ thu, các thực thể này có thể điều khiển bộ giải mã. `HandleCraAsBlaFlag` có thể được thiết đặt bằng 1, ví dụ bởi bộ trình phát để tìm kiếm vị trí mới trong luồng bit hoặc điều chỉnh thành bộ phát 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 này được xử lý và được giải mã nếu nó là hình ảnh BLA.

Trong HEVC, chuỗi video đã mã hoá còn có thể hoặc theo cách khác (với bản đặc tả ở trên) được định rõ để kết thúc, khi một đơn vị NAL cụ thể, có thể được gọi là đơn vị NAL đầu chuỗi (EOS), xuất hiện trong luồng bit và có `nuh_layer_id` bằng 0.

Trong HEVC, nhóm chuỗi video đã mã hoá (CVSG – Coded Video Sequence Group) có thể được xác định, ví dụ, làm một hoặc nhiều CVS liên tiếp theo trình tự giải mã nói chung gồm đơn vị truy cập IRAP kích hoạt RBSP VPS firstVpsRbsp chưa hoạt động được theo sau bởi tất cả các đơn vị truy cập sau đó, theo trình tự giải mã, trong đó firstVpsRbsp là RBSP VPS hoạt động đến đầu của luồng bit hoặc đến, nhưng loại trừ đơn vị truy cập kích hoạt một RBSP VPS khác với firstVpsRbsp, dùng bất kỳ trường hợp nào xuất hiện trước theo trình tự giải mã.

Nhóm các hình ảnh (GOP – Group of Pictures) và các đặc điểm của nhóm này có thể được xác định như sau. GOP có thể được giải mã bất kể nếu các hình ảnh bất kỳ trước đó được giải mã hay không. GOP mở là nhóm các hình ảnh như trên trong đó các hình ảnh ở trước hình ảnh bên trong ban đầu theo trình tự đầu ra có thể không giải mã chính xác được khi quy trình giải mã bắt đầu từ hình ảnh bên trong ban đầu của GOP mở. Nói cách khác, các hình ảnh của GOP mở có thể đề cập (trong dự đoán liên ảnh) đến các hình ảnh thuộc một GOP trước đó. Bộ giải mã H.264/AVC có thể nhận ra hình ảnh bên trong bắt đầu GOP mở từ thông báo SEI điểm khôi phục trong luồng bit H.264/AVC. Bộ giải mã HEVC có thể nhận ra hình ảnh bên trong bắt đầu GOP mở, do một loại đơn vị NAL cụ thể, loại đơn vị NAL CRA, có thể được sử dụng cho các lát đã mã hoá của nó. GOP kín là nhóm các hình ảnh như vậy trong đó tất cả các hình ảnh có thể được giải mã chính xác khi quy trình giải mã bắt đầu từ hình ảnh bên trong ban đầu của GOP kín. Nói cách khác, không hình ảnh nào trong GOP kín đề cập đến bất kỳ các hình ảnh nào trong các GOP trước đó. Trong H.264/AVC và HEVC, GOP kín có thể bắt đầu từ hình ảnh IDR. Trong HEVC, GOP kín cũng có thể bắt đầu từ hình ảnh BLA_W_RADL hoặc BLA_N_LP. Cấu trúc mã hóa GOP mở có khả năng nén hiệu quả hơn so với cấu trúc mã hóa GOP kín, do có sự linh hoạt hơn khi lựa chọn các hình ảnh tham chiếu.

Cấu trúc của các hình ảnh (SOP – Structure of Pictures) có thể được định nghĩa là một hoặc nhiều hình ảnh đã mã hóa liên tiếp theo trình tự giải mã, trong đó hình ảnh đã mã hóa thứ nhất theo trình tự giải mã là hình ảnh tham chiếu thuộc lớp con bậc thấp nhất theo thời gian và không có hình ảnh đã mã hóa nào ngoài khả năng là hình ảnh đã mã hóa thứ nhất theo trình tự giải mã là hình ảnh RAP. Tất cả các hình ảnh trong SOP

trước đó theo trình tự giải mã ở trước là tất cả các hình ảnh trong SOP hiện tại và tất cả các hình ảnh trong SOP tiếp đó theo trình tự giải mã nối tiếp tất cả các hình ảnh trong SOP hiện tại. SOP có thể biểu diễn cấu trúc dự đoán liên ảnh phân cấp và lặp lại. Thuật ngữ nhóm các hình ảnh (GOP – group of pictures) đôi khi có thể được sử dụng thay thế bằng thuật ngữ SOP và có cùng ngữ nghĩa như ngữ nghĩa của SOP.

Cú pháp luồng bit của H.264/AVC và HEVC chỉ ra xem liệu hình ảnh cụ thể có là hình ảnh tham chiếu để dự đoán liên ảnh của hình ảnh bất kỳ khác hay không. Các hình ảnh của loại mã hoá bất kỳ (I, P, B) có thể là các hình ảnh tham chiếu hoặc hình ảnh không tham chiếu trong H.264/AVC và HEVC.

H.264/AVC định rõ quy trình để đánh dấu hình ảnh tham chiếu đã giải mã để điều khiển sự tiêu thụ bộ nhớ trong bộ giải mã. Số lượng các hình ảnh tham chiếu tối đa được sử dụng để dự đoán liên ảnh, được gọi là M, được xác định thuộc tập tham số chuỗi này. Khi hình ảnh tham chiếu được giải mã, hình ảnh này được đánh dấu là “được sử dụng để tham chiếu”. Nếu quy trình giải mã hình ảnh tham chiếu khiến nhiều hơn M hình ảnh được đánh dấu là “được sử dụng để tham chiếu”, thì ít nhất một hình ảnh được đánh dấu là “không được sử dụng để tham chiếu”. Có hai loại vận hành để đánh dấu hình ảnh tham chiếu đã giải mã: điều khiển bộ nhớ tương thích và cửa sổ trượt. Chế độ vận hành để đánh dấu hình ảnh tham chiếu đã giải mã được lựa chọn dựa trên cơ sở hình ảnh. Điều khiển bộ nhớ tương thích cho phép báo hiệu rõ ràng rằng các hình ảnh nào được đánh dấu là “không được sử dụng để tham chiếu” và cũng có thể ấn định các chỉ số dài hạn cho các hình ảnh tham chiếu ngắn hạn. Điều khiển bộ nhớ tương thích có thể yêu cầu sự có mặt của các tham số vận hành điều khiển quản lý bộ nhớ (MMCO – memory management control operation) trong luồng bit. Các tham số MMCO có thể được bao gồm trong hình ảnh tham chiếu đã giải mã để đánh dấu cấu trúc cú pháp. Nếu chế độ vận hành cửa sổ trượt đang được sử dụng và có M hình ảnh được đánh dấu là “được sử dụng để tham chiếu”, thì hình ảnh tham chiếu ngắn hạn mà là hình ảnh đã giải mã thứ nhất giữa các hình ảnh tham chiếu ngắn hạn mà được đánh dấu là “được sử dụng để tham chiếu”, được đánh dấu là “không được sử dụng để tham chiếu”. Nói cách khác, chế độ vận hành cửa sổ trượt dẫn đến vận hành đệm vào trước ra trước giữa các hình ảnh tham chiếu ngắn hạn.

Một trong các vận hành điều khiển quản lý bộ nhớ trong H.264/AVC làm cho tất cả các hình ảnh tham chiếu, ngoại trừ hình ảnh hiện thời, được đánh dấu là “không được sử dụng để tham chiếu”. Hình ảnh làm mới quy trình giải mã tức thời (IDR – instantaneous decoding refresh) chỉ gồm các lát đã mã hoá nội ảnh và làm cho “thiết đặt lại” tương tự lên các hình ảnh tham chiếu.

Trong HEVC, các cấu trúc cú pháp đánh dấu hình ảnh tham chiếu và các quy trình giải mã liên quan không được sử dụng, mà thay vào đó cấu trúc cú pháp tập hình ảnh tham chiếu (RPS – reference picture set) và quy trình giải mã được sử dụng thay thế với mục đích tương tự. Tập hình ảnh tham chiếu có hiệu lực hoặc hoạt động đối với hình ảnh bao gồm tất cả các hình ảnh tham chiếu được sử dụng làm tham chiếu cho hình ảnh và tất cả các hình ảnh tham chiếu vẫn được đánh dấu là “được sử dụng để tham chiếu” cho các hình ảnh bất kỳ tiếp đó trong trình tự giải mã. Có sáu tập hợp con của tập hình ảnh tham chiếu, được gọi cụ thể là RefPicSetStCurr0 (còn được gọi là RefPicSetStCurrBefore), RefPicSetStCurr1 (còn được gọi là RefPicSetStCurrAfter), RefPicSetStFoll0, RefPicSetStFoll1, RefPicSetLtCurr, và RefPicSetLtFoll. RefPicSetStFoll0 và RefPicSetStFoll1 cũng có thể được coi là để cùng nhau tạo thành một tập con RefPicSetStFoll. Chú giải của sáu tập hợp con là như sau. “Curr” đề cập đến các hình ảnh tham chiếu được bao gồm trong danh mục các hình ảnh tham chiếu của hình ảnh hiện thời và do đó có thể được sử dụng làm 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 danh mục hình ảnh tham chiếu của hình ảnh hiện thời nhưng có thể được sử dụng trong các hình ảnh tiếp đó theo trình tự giải mã làm các hình ảnh tham chiếu. “St” đề cập đến các hình ảnh tham chiếu ngắn hạn, “St” đề cập đến các hình ảnh tham chiếu ngắn hạn, thường có thể được xác định thông qua một số bit ít quan trọng nhất định của giá trị POC của chúng. “Lt” đề cập đến các hình ảnh tham chiếu dài hạn, được nhận dạng cụ thể và thường có khác biệt lớn về các giá trị POC về hình ảnh hiện thời so với giá trị có thể được biểu diễn bởi số lượng bit ít quan trọng nhất định đã đề cập. “0” đề cập đến các hình ảnh tham chiếu mà có giá trị POC nhỏ hơn giá trị POC của hình ảnh hiện thời. “1” đề cập đến các hình ảnh tham chiếu mà có giá trị POC lớn hơn giá trị POC của hình ảnh hiện thời. RefPicSetStCurr0, RefPicSetStCurr1, RefPicSetStFoll0 và RefPicSetStFoll1 được gọi chung là tập hợp con ngắn hạn của tập hình ảnh tham chiếu.

RefPicSetLtCurr và RefPicSetLtFoll được gọi chung là tập hợp con dài hạn của tập hình ảnh tham chiếu.

Trong HEVC, tập ảnh tham chiếu có thể được định rõ trong tập tham số chuỗi và được đưa vào sử dụng trong phần đầu lát thông qua một chỉ số đến tập ảnh tham chiếu. Tập ảnh tham chiếu cũng có thể được định rõ trong phần đầu lát. Tập ảnh tham chiếu có thể được mã hóa một cách độc lập hoặc có thể được dự đoán từ tập ảnh tham chiếu khác (được gọi là dự đoán liên RPS). Trong cả hai loại mã hóa tập ảnh tham chiếu, cờ (used_by_curr_pic_X_flag) được gửi bổ sung cho mỗi ảnh tham chiếu, cho biết liệu ảnh tham chiếu có được dùng để tham chiếu bởi ảnh hiện thời (được bao gồm trong danh mục *Curr) hay không (được bao gồm trong danh mục *Foll). Các ảnh được bao gồm trong tập ảnh tham chiếu được sử dụng bởi lát hiện thời được đánh dấu là “dùng để tham chiếu”, và các ảnh không nằm trong tập ảnh tham chiếu được sử dụng bởi lát hiện thời được đánh dấu là “không dùng để tham chiếu”. Nếu ảnh hiện thời là ảnh IDR, thì tất cả RefPicSetStCurr0, RefPicSetStCurr1, RefPicSetStFoll0, RefPicSetStFoll1, RefPicSetLtCurr và RefPicSetLtFoll đều được thiết đặt là rỗng.

Mã hóa video có thể mở rộng có thể chỉ cấu trúc mã hóa trong đó một luồng bit có thể chứa nhiều dạng biểu diễn của nội dung tại các tốc độ bit, độ phân giải hoặc tốc độ khung khác nhau. Trong các trường hợp này, bộ thu có thể trích xuất dạng biểu diễn mong muốn tùy thuộc vào các đặc điểm của nó (ví dụ, độ phân giải mà phù hợp nhất với thiết bị hiển thị). Theo cách khác, máy chủ hoặc phần tử mạng có thể trích xuất các phần của luồng bit cần truyền đến bộ thu tùy thuộc vào, ví dụ, các đặc điểm mạng hoặc khả năng xử lý của bộ thu. Luồng bit có thể mở rộng có thể bao gồm “lớp cơ sở” cung cấp video chất lượng thấp nhất khả dụng và một hoặc nhiều lớp tăng cường mà tăng cường chất lượng video khi được thu và được giải mã cùng với các lớp thấp hơn. Để cải thiện hiệu quả mã hóa cho các lớp tăng cường, dạng biểu diễn đã mã hóa của lớp đó có thể phụ thuộc vào 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 sự dự đoán cho lớp tăng cường.

Bộ mã hóa - giải mã video có thể mở rộng để cung cấp khả năng mở rộng chất lượng (còn được gọi là Signal-to-Noise hoặc SNR) và/hoặc khả năng mở rộng không

gian có thể được thực thi như sau. Đối với lớp cơ sở, bộ mã hóa và bộ giải mã video không thể mở rộng thông thường được sử dụng. Các ảnh đã tái dựng/đã giải mã của lớp cơ sở được bao gồm trong bộ đệm ảnh tham chiếu cho lớp tăng cường. Trong H.264/AVC, HEVC và các bộ mã hóa - giải mã tương tự có sử dụng (các) danh mục ảnh tham chiếu để dự đoán liên ảnh, các ảnh đã giải mã lớp cơ sở có thể được đưa vào (các) danh mục ảnh tham chiếu để mã hóa/giải mã ảnh lớp tăng cường theo cách tương tự với các ảnh tham chiếu đã giải mã của lớp tăng cường. Sau đó, bộ mã hóa có thể lựa chọn ảnh tham chiếu lớp cơ sở làm tham chiếu dự đoán liên ảnh và chỉ định việc sử dụng nó với chỉ số ảnh tham chiếu trong luồng bit đã mã hóa. Bộ giải mã giải mã từ luồng bit, ví dụ, từ chỉ số ảnh tham chiếu, rằng ảnh lớp cơ sở được sử dụng làm tham chiếu dự đoán liên ảnh cho lớp tăng cường. Khi ảnh lớp cơ sở đã giải mã được sử dụng làm tham chiếu dự đoán cho lớp tăng cường, nó được gọi là ảnh tham chiếu liên lớp.

Ngoài khả năng mở rộng chất lượng, cũng có thể có các chế độ khả năng mở rộng hoặc chiều mở rộng khác, có thể bao gồm, nhưng không giới hạn ở, các chế độ hoặc chiều mở rộng nêu dưới đây:

Khả năng mở rộng chất lượng: Các ảnh lớp cơ sở được mã hóa với chất lượng thấp hơn so với các ảnh lớp tăng cường, có thể đạt được, ví dụ, bằng cách sử dụng giá trị tham số lượng tử hóa lớn hơn (nghĩa là, kích thước bước lượng tử hóa lớn hơn đối với bước lượng tử hóa hệ số biến đổi) trong lớp cơ sở so với trong lớp tăng cường. Khả năng mở rộng chất lượng còn có thể được phân loại thành khả năng mở rộng hạt mịn hoặc độ hạt mịn (FGS), khả năng mở rộng hạt trung bình hoặc độ hạt trung bình (MGS), và/hoặc khả năng mở rộng hạt thô hoặc độ hạt thô (CGS), như được mô tả dưới đây.

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

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

Khả năng mở rộng định dạng sắc độ: Các ảnh lớp cơ sở cung cấp độ phân giải không gian thấp hơn trong các mảng mẫu sắc độ (ví dụ, được mã hóa ở định dạng sắc độ 4:2:0) so với các ảnh lớp tăng cường (ví dụ, định dạng 4:4:4).

Khả năng mở rộng gam màu: các ảnh lớp tăng cường có sự biểu diễn màu tốt hơn/rộng hơn so với sự biểu diễn màu của các ảnh lớp cơ sở - ví dụ, lớp tăng cường có thể có gam màu UHDTV (ITU-R BT.2020) và lớp cơ sở có thể có gam màu ITU-R BT.709.

Khả năng mở rộng hiển thị, còn có thể được gọi là mã hóa đa hiển thị. Lớp cơ sở biểu diễn hiển thị thứ nhất, trong khi lớp tăng cường biểu diễn hiển thị thứ hai.

Khả năng mở rộng độ sâu, còn có thể được gọi là mã hóa tăng cường độ sâu. Một lớp hoặc một số lớp của luồng bit có thể biểu diễn (các) hiển thị kết cấu, trong khi lớp hoặc các lớp khác có thể biểu diễn (các) hiển thị độ sâu.

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

Khả năng mở rộng từ đơn xen đến tuần tự (còn được gọi là khả năng mở rộng từ trường đến khung): phần nội dung nguồn đơn xen đã mã hóa của lớp cơ sở được tăng cường bằng lớp tăng cường để biểu diễn nội dung nguồn tuần tự. Nội dung nguồn đơn xen đã mã hóa trong lớp cơ sở có thể bao gồm các trường đã mã hóa, các khung đã mã hóa biểu diễn các cặp trường, hoặc tổ hợp của chúng. Trong khả năng mở rộng từ đơn xen đến tuần tự, ảnh lớp cơ sở có thể được lấy mẫu lại sao cho nó trở thành ảnh tham chiếu thích hợp đối với một hoặc nhiều ảnh lớp tăng cường.

Khả năng mở rộng bộ mã hóa - giải mã lai (còn được gọi là khả năng mở rộng chuẩn mã hóa): Trong khả năng mở rộng bộ mã hóa - giải mã lai, cú pháp luồng bit, ngữ nghĩa và quy trình giải mã của lớp cơ sở và lớp tăng cường được định rõ trong các chuẩn mã hóa video khác nhau. Do đó, các ảnh lớp cơ sở được mã hóa theo định dạng hoặc chuẩn mã hóa khác so với các ảnh lớp tăng cường. Ví dụ, lớp cơ sở có thể được mã hóa bằng H.264/AVC và lớp tăng cường có thể được mã hóa bằng nội dung mở rộng đa lớp HEVC. Thông thường hơn, trong khả năng mở rộng bộ mã hóa - giải mã lai, một hoặc nhiều lớp có thể được mã hóa theo một chuẩn mã hóa hoặc bản đặc tả và một hoặc nhiều lớp khác có thể được mã hóa theo chuẩn mã hóa hoặc bản đặc tả khác. Ví dụ, có thể có

hai lớp được mã hóa theo nội dung mở rộng MVC của H.264/AVC (một trong số đó là lớp cơ sở được mã hóa theo H.264/AVC), và một hoặc nhiều lớp bổ sung được mã hóa theo MV-HEVC. Ngoài ra, số lượng chuẩn mã hóa hoặc bản đặc tả mà các lớp khác nhau của cùng một luồng bit được mã hóa theo đó có thể không bị giới hạn ở hai trong khả năng mở rộng bộ mã hóa - giải mã lai.

Cần hiểu rằng nhiều loại khả năng mở rộng có thể được kết hợp và áp dụng cùng nhau. Ví dụ, khả năng mở rộng gam màu và khả năng mở rộng độ sâu bit có thể được kết hợp.

Trong các trường hợp khả năng mở rộng nêu trên, thông tin về lớp cơ sở có thể được sử dụng để mã hóa lớp tăng cường nhằm giảm thiểu phí tổn tốc độ bit bổ sung.

Khả năng mở rộng có thể được kích hoạt theo hai cách. Hoặc bằng cách đưa các chế độ mã hóa mới vào để thực hiện việc dự đoán các giá trị điểm ảnh hoặc cú pháp từ các lớp thấp hơn của dạng biểu diễn có thể mở rộng, hoặc bằng cách đặt các ảnh lớp thấp hơn vào bộ đệm ảnh tham chiếu (bộ đệm ảnh đã giải mã, DPB) của lớp cao hơn. Cách thứ nhất linh hoạt hơn và do đó có thể cung cấp hiệu quả mã hóa tốt hơn trong hầu hết các trường hợp. Tuy nhiên, cách thứ hai, là kích hoạt khả năng mở rộng dựa trên khung tham chiếu, có thể được thực thi rất hiệu quả với các thay đổi tối thiểu đối với các bộ mã hóa - giải mã đơn lớp trong khi vẫn đạt được phần lớn các lợi ích hiệu quả mã hóa khả dụng. Về cơ bản, bộ mã hóa - giải mã khả năng mở rộng dựa trên khung tham chiếu có thể được thực thi bằng cách sử dụng cùng chế độ thực thi phần cứng hoặc phần mềm cho tất cả các lớp, chỉ quan tâm đến việc quản lý DPB bằng thành phần ngoại vi.

Để có thể sử dụng chế độ xử lý song song, các hình ảnh có thể được chia thành các đoạn hình ảnh (các lát hoặc ô) có thể mã hóa và có thể giải mã một cách độc lập. Các lát có thể chỉ các đoạn hình ảnh được xây dựng bởi một số lượng nhất định các đơn vị mã hóa cơ bản mà được xử lý theo thứ tự mã hóa hoặc giải mã mặc định, trong khi các ô có thể chỉ các đoạn hình ảnh được định nghĩa là các vùng hình ảnh hình chữ nhật mà được xử lý ít nhất là đến một mức độ nào đó như các khung riêng lẻ.

Trong phần dưới đây, phương pháp theo một phương án của sáng chế sẽ được mô tả chi tiết hơn dựa vào sơ đồ khối được thể hiện trên các Fig.7, Fig.8a - 8c và lưu đồ

được thể hiện trên Fig.9. Fig.7 minh họa phần tử xác định vector chuyển động 700 dưới dạng sơ đồ khối đơn giản, theo một phương án của sáng chế. Chức năng của phần tử xác định vector chuyển động 700 có thể được bao gồm trong bộ dự đoán liên ảnh 306, 406 (Fig.4) hoặc nó có thể là phần tử riêng biệt của bộ mã hóa 500. Đơn vị mã hóa mà hiện tại cần được mã hóa 900 có thể được gọi là đơn vị mã hóa hiện thời. Vector chuyển động tịnh tiến cơ sở b0 có thể được lựa chọn 902 cho đơn vị mã hóa hiện thời 820 bằng bộ lựa chọn vector chuyển động cơ sở 702. Vector chuyển động tịnh tiến cơ sở b0 có thể được chọn trong số các vector chuyển động của một hoặc nhiều đơn vị mã hóa đã mã hóa trước đó. Nói cách khác, bộ lựa chọn vector chuyển động cơ sở 702 có thể có quyền truy cập vào thông tin của tập vị trí ứng viên mà vector chuyển động tịnh tiến cơ sở b0 có thể được lựa chọn từ đó. Vector chuyển động tịnh tiến cơ sở được lựa chọn còn có thể được gọi là vector chuyển động nguồn. Vị trí của vector chuyển động nguồn để dự đoán có thể được thể hiện trong luồng bit và có thể được lựa chọn bởi bộ lựa chọn vector chuyển động cơ sở 702 từ tập vị trí ứng viên. Ví dụ, tập vị trí ứng viên có thể bao gồm vị trí ngay trên đơn vị mã hóa hiện thời (hoặc đơn vị dự đoán) và vị trí bên trái đơn vị mã hóa.

Thông tin về tập vị trí ứng viên và có thể là các vector chuyển động tương ứng có thể đã được lưu trữ vào bộ nhớ trong quá trình mã hóa trước đó đối với một hoặc nhiều đơn vị mã hóa khác.

Khối tạo trường chuyển động 704 có thể tạo ra trường vector chuyển động mượt liên tục 830 bao gồm đơn vị mã hóa hiện thời, ví dụ, như sau. Vector chuyển động cơ sở được lựa chọn b0 có các thành phần b_x , b_y có thể được sử dụng bởi khối tạo trường chuyển động 704 để xác định 904 chế độ dự đoán cho ít nhất là thành phần vector chuyển động tịnh tiến cơ sở thứ nhất b_x , b_y . Khối tạo trường chuyển động 704 còn có thể xác định 906 hệ số tinh chỉnh r_{bx} , r_{by} cho thành phần vector chuyển động tịnh tiến thứ nhất b_x , b_y bằng cách sử dụng độ chính xác thứ nhất. Khối tạo trường chuyển động 704 cũng có thể thực hiện dự đoán vector chuyển động cho một hoặc nhiều điểm ảnh (mẫu) khác của đơn vị mã hóa hiện thời (ví dụ, khối mẫu). Theo một phương án của sáng chế, khối tạo trường chuyển động 704 thực hiện dự đoán vector chuyển động cho mỗi điểm ảnh của đơn vị mã hóa hiện thời để thu các dự đoán vector chuyển động cho mỗi điểm ảnh.

Khi sự dự đoán vector chuyển động cho vector chuyển động cơ sở được lựa chọn b_0 và sự dự đoán vector chuyển động cho một hoặc nhiều điểm ảnh khác đã được xác định, khối tạo trường chuyển động 704 cũng có thể xác định 908 sự dự đoán cho ít nhất một thành phần vector chuyển động bậc cao a_0, a_1 dựa trên các vector chuyển động này. Khối tạo trường chuyển động 704 còn có thể xác định 910 hệ số tinh chỉnh r_{a0}, r_{a1} cho (các) thành phần vector chuyển động bậc cao a_0, a_1 bằng cách sử dụng độ chính xác thứ hai khác với độ chính xác thứ nhất. Khối tạo trường chuyển động 704 có thể sử dụng 912 thành phần vector chuyển động tịnh tiến cơ sở thứ nhất b_x, b_y , hệ số tinh chỉnh r_{bx}, r_{by} cho thành phần vector chuyển động tịnh tiến thứ nhất b_x, b_y , sự dự đoán cho ít nhất một thành phần vector chuyển động bậc cao a_0, a_1 , và hệ số tinh chỉnh r_{a0}, r_{a1} cho thành phần vector chuyển động bậc cao a_0, a_1 để tính toán vector chuyển động thứ nhất. Tuy nhiên, cần lưu ý rằng có thể không phải lúc nào cũng cần xác định và sử dụng hệ số tinh chỉnh r_{bx}, r_{by} cho thành phần vector chuyển động tịnh tiến thứ nhất b_x, b_y và/hoặc hệ số tinh chỉnh r_{a0}, r_{a1} cho thành phần vector chuyển động bậc cao a_0, a_1 để tính toán vector chuyển động thứ nhất.

Trường vector chuyển động mượt liên tục có thể được tạo ra, ví dụ, bằng cách sử dụng dữ liệu đã xác định ở trên để tạo ra trường (tập) vector chuyển động dưới dạng hàm của vị trí mẫu hoặc khối mẫu trong đơn vị mã hóa hiện thời. Ví dụ, các thành phần vector chuyển động tịnh tiến cơ sở (b_x, b_y) có thể được tính toán bằng cách bổ sung hệ số tinh chỉnh đã mở rộng (r_{bx}, r_{by}) vào các thành phần vector chuyển động đã dự đoán (p_{bx}, p_{by}) bằng cách sử dụng thừa số s_b :

$$b_x = p_{bx} + s_b * r_{bx} \quad (2)$$

$$b_y = p_{by} + s_b * r_{by}$$

Tuy nhiên, trong trường hợp độ chính xác mong muốn của vector chuyển động cơ sở và thông tin của hệ số tinh chỉnh là giống nhau (ví dụ, nếu cả vector cơ sở và vector đã dự đoán đều được đánh giá với độ chính xác mẫu 1/4 và thông tin của hệ số tinh chỉnh cũng được báo hiệu với độ chính xác mẫu 1/4), tham số mở rộng s_b có thể được thiết đặt đến 1 và bước mở rộng có thể được bỏ qua.

Các hệ số chuyển động bậc cao (a_0, a_1) cũng có thể được dự đoán từ các giá trị đã dự đoán tương ứng của chúng (p_{a0}, p_{a1}). Trong trường hợp này, các hệ số chuyển

động a_0 và a_1 có thể được tính toán dựa trên thừa số độ chính xác bậc cao s_a và các giá trị hệ số tinh chỉnh đã báo hiệu (r_{a0} , r_{a1}) dưới dạng:

$$a_0 = p_{a0} + s_a * r_{a0} \quad (3)$$

$$a_1 = p_{a1} + s_a * r_{a1}$$

Ngoài ra, trong trường hợp độ chính xác của hệ số chuyển động bậc cao giống với độ chính xác của hệ số tinh chỉnh, bước mở rộng có thể được bỏ qua.

Theo một phương án của sáng chế, mô hình chuyển động afin giới hạn có khả năng xử lý xoay và thu phóng ngoài chuyển động tịnh tiến có thể được sử dụng. Trong trường hợp này, vectơ chuyển động cho mẫu hoặc nhóm mẫu nhất định có thể được tính toán như sau:

$$v_x = b_x + a_0(x - x_0) + a_1(y - y_0) \quad (4)$$

$$v_y = b_y - a_1(x - x_0) + a_0(y - y_0)$$

trong đó (v_x, v_y) là vectơ chuyển động thu được cho vị trí (x, y) dựa trên vectơ chuyển động cơ sở (b_x, b_y) ở vị trí (x_0, y_0) và hai tham số a_0 và a_1 xác định sự thích ứng trường vectơ chuyển động dưới dạng hàm của vị trí mẫu.

Fig.8a và Fig.8b mô tả một số lựa chọn có thể có đối với các vectơ chuyển động cơ sở để xác định trường vectơ chuyển động của đơn vị mã hóa. Trong trường hợp được thể hiện trên Fig.8a, các vectơ chuyển động cho các khối con (hoặc các mẫu) phía trên bên trái 822 (b_0) và phía trên bên phải 824 (b_1) được tính toán và sự chênh lệch ngang giữa các vectơ chuyển động này được sử dụng để suy ra giá trị bước d_x mà mô tả cách trường vectơ chuyển động 830 thích ứng theo chiều ngang (liên quan tới tham số a_0 trong các phương trình nêu trên). Fig.8b mô tả trường hợp trong đó vectơ chuyển động cơ sở nằm ngoài đơn vị mã hóa hiện thời. Trong ví dụ này, các vectơ chuyển động cho các khối con (hoặc các mẫu) phía trên bên trái 822 (b_0) và phía trên bên phải 824 (b_1) được tính toán bằng cách sử dụng khối đã mã hóa trước đó 826 ở trên và ở phía trái của khối con phía trên bên trái 822 và khối đã mã hóa trước đó 828 ở trên khối con phía trên bên phải 824. Sự chênh lệch ngang giữa các vectơ chuyển động này được sử dụng để suy ra giá trị bước d_x mà mô tả cách trường vectơ chuyển động 830 thích ứng theo chiều ngang (liên quan tới tham số a_0 trong các phương trình nêu trên). Cách này có thể có hai

ưu điểm so với trường hợp được thể hiện trên Fig.8a. Thứ nhất, trong trường hợp độ rộng của khối trong các điểm ảnh là lũy thừa của hai, khoảng cách ngang giữa khối con liên kết với vector cơ sở b_0 và khối con liên kết với vector cơ sở b_1 cũng là lũy thừa của hai và độ chênh lệch dx có thể được tính toán bằng các bước dịch chuyển theo bit. Thứ hai, các vector chuyển động cơ sở b_0 và/hoặc b_1 giờ đây có thể là các vector chuyển động thực đã được mã hóa/giải mã của các khối trước đó và không nhất thiết phải được tinh chỉnh để đạt được các trường vector chuyển động thay đổi liên tục trên ranh giới của các đơn vị mã hóa (hoặc các đơn vị dự đoán).

Khi trường vector chuyển động đã được xác định, bước dự đoán bù chuyển động có thể được thực hiện 914, ví dụ bằng khối dự đoán chuyển động 706, sử dụng vector chuyển động đã nêu cho mẫu hoặc nhóm mẫu. Nói cách khác, nếu các vector chuyển động được xác định theo mẫu, một mẫu của đơn vị mã hóa hiện thời được lựa chọn để dự đoán và vector chuyển động được xác định cho đơn vị mã hóa này được sử dụng trong chế độ dự đoán này. Sau đó, ví dụ, mẫu tiếp theo trong mẫu hiện thời được lựa chọn và vector chuyển động được xác định cho mẫu này trong trường các vector chuyển động được sử dụng trong chế độ dự đoán chuyển động này. Quy trình này có thể được lặp lại cho đến khi việc dự đoán chuyển động cho tất cả các mẫu trong đơn vị mã hóa hiện thời đã được thực hiện.

Trong trường hợp các vector chuyển động đã được xác định theo khối, thì cùng một vector chuyển động có thể được sử dụng cho mỗi mẫu của khối hiện thời để dự đoán chuyển động. Quy trình này có thể được lặp lại cho đến khi việc dự đoán chuyển động cho tất cả các khối trong đơn vị mã hóa hiện thời đều được thực hiện. Sau đó, đơn vị mã hóa khác có thể được lựa chọn làm đơn vị mã hóa hiện thời, nếu có, và các quy trình nêu trên có thể được lặp lại nếu chế độ dự đoán chuyển động được sử dụng cho đơn vị mã hóa này.

Theo một phương án của sáng chế, vector chuyển động tịnh tiến cơ sở (b_x , b_y) có thể được xác định dựa trên vector chuyển động đã giải mã của đơn vị mã hóa khác. Vị trí của vector chuyển động nguồn để dự đoán này có thể được thể hiện trong luồng bit và có thể được lựa chọn bởi bộ mã hóa từ tập vị trí ứng viên. (Các) thành phần vector chuyển động bậc cao có thể được chỉ báo bằng cách báo hiệu các tham số chuyển động còn lại

(a_0, a_1) theo các cách khác nhau. Những thành phần đó có thể được thể hiện dưới dạng vector chuyển động bổ sung cho một vị trí nhất định bên trong hoặc bên ngoài đơn vị mã hóa, hoặc có thể được thể hiện dưới dạng các hệ số chuyển động riêng lẻ.

Để báo hiệu các tham số chuyển động hoặc các hệ số tinh chỉnh cho các tham số chuyển động, các phương pháp khác nhau có thể được sử dụng. Ví dụ, các vector chuyển động ở các vị trí khác nhau có thể được tính toán và các tham số này có thể được suy ra từ các vector chuyển động này.

Theo một cách thực hiện thay thế, mô hình afin đầy đủ với sáu tham số; b_x , b_y và a_0 đến a_3 ; có thể được sử dụng:

$$v_x = b_x + a_0(x - x_0) + a_1(y - y_0) \quad (5)$$

$$v_y = b_y + a_2(x - x_0) + a_3(y - y_0)$$

Trong trường hợp tổng quát hơn, mô hình chuyển động bất kỳ với độ phụ thuộc không gian có thể được ứng dụng, ví dụ mô hình tuyến tính của các hàm cơ bản bậc cao $N+M$:

$$v_x = b_x + a_1 f_1(x, y) + a_2 f_2(x, y) + \dots + a_N f_N(x, y) \quad (6)$$

$$v_y = b_y + b_1 g_1(x, y) + b_2 g_2(x, y) + \dots + b_M g_M(x, y)$$

Theo một phương án, trường vector chuyển động có thể được báo hiệu bằng cách sử dụng các vector chuyển động góc của đơn vị mã hóa. Trong trường hợp này, vector chuyển động delta được sử dụng để tính toán vector chuyển động phía trên bên trái cho đơn vị mã hóa được mã hóa với độ chính xác thứ nhất (ví dụ, độ chính xác mẫu 1/4) và vector chuyển động delta được sử dụng để tính toán vector chuyển động cho mẫu khác (ví dụ, vector chuyển động phía trên bên phải) trong đơn vị mã hóa được mã hóa với độ chính xác thứ hai (ví dụ, độ chính xác mẫu 1/16).

Theo một phương án, vector chuyển động cơ sở cho trường vector chuyển động của đơn vị mã hóa được xác định dưới dạng vector chuyển động cho khối mẫu 4x4 ngay bên ngoài đơn vị mã hóa hiện thời trên chiều phía trên bên trái của góc phía trên bên trái của đơn vị mã hóa.

Theo một phương án, vector chuyển động cơ sở cho đơn vị mã hóa được điều chỉnh bằng cách bổ sung thành phần vector chuyển động bậc cao từ vị trí được sử dụng làm nguồn để dự đoán vector chuyển động cơ sở cho đơn vị mã hóa hiện thời.

Theo một phương án, ảnh hoặc các ảnh tham chiếu để dự đoán vector chuyển động được lựa chọn dựa trên ảnh hoặc các ảnh tham chiếu được sử dụng cho vị trí mà được sử dụng làm nguồn để dự đoán vector chuyển động cơ sở cho đơn vị mã hóa hiện thời.

Theo một phương án, độ chính xác của các hệ số tinh chỉnh cho các hệ số chuyển động bậc cao phụ thuộc vào kích thước của đơn vị mã hóa mà các tham số chuyển động được lựa chọn này được áp dụng lên đó.

Theo một phương án, không có sự báo hiệu về hệ số tinh chỉnh cho vector chuyển động tịnh tiến cơ sở nhưng vector chuyển động đã giải mã của đơn vị mã hóa lân cận hoặc đơn vị dự đoán có thể được sử dụng làm vector chuyển động cơ sở cho đơn vị hiện thời.

Theo một phương án, các hệ số mở rộng cho hệ số tinh chỉnh vector chuyển động cơ sở, hệ số tinh chỉnh hệ số chuyển động bậc cao, hoặc cả hai, có thể được báo hiệu trong tập tham số hoặc phần đầu ảnh/lát/ô hoặc ở vị trí khác trong luồng bit.

Theo một phương án, các hệ số mở rộng cho vector chuyển động cơ sở, hệ số chuyển động bậc cao, hoặc cả hai, có thể được báo hiệu trong tập tham số hoặc phần đầu ảnh/lát/ô hoặc ở vị trí khác trong luồng bit.

Fig.8a minh họa các vector chuyển động cơ sở b_0 và b_1 bên trong đơn vị mã hóa. Sự chênh lệch vector chuyển động ngang giữa các vector chuyển động có thể được tính toán dưới dạng sự chênh lệch giữa các thành phần ngang của các vector chuyển động cơ sở và chia cho khoảng cách của các vector chuyển động cơ sở.

Fig.8b minh họa các vector chuyển động cơ sở b_0 và b_1 bên ngoài đơn vị mã hóa. Sự chênh lệch vector chuyển động ngang giữa các vector chuyển động có thể được tính toán dưới dạng sự chênh lệch giữa các thành phần ngang của các vector chuyển động cơ sở và chia cho khoảng cách của các vector chuyển động cơ sở. Bây giờ, các vector chuyển

động cơ sở có thể là các vector chuyển động thực dùng để bù chuyển động cho các khối trước đó.

Fig.8c minh họa việc sử dụng mô hình afin với giá trị dương cho tham số bậc nhất a_0 . Như có thể thấy trên Fig.8c, tham số đó có thể tác động đến lượng thay đổi của thành phần vector chuyển động ngang khi đi qua đơn vị mã hóa theo chiều ngang.

Fig.8a, Fig.8b và Fig.8c chỉ minh họa sự thích ứng theo chiều ngang của các vector chuyển động, nhưng các cách tương tự cũng có thể được sử dụng theo chiều dọc.

Trong phần dưới đây, một số ví dụ thực hiện sáng chế sẽ được mô tả ngắn gọn. Các bước xử lý trên cả phía bộ mã hóa và bộ giải mã có thể được thực hiện theo thứ tự khác nhau hoặc đan xen nhau, và các bước khác nhau có thể được thực hiện song song với nhau.

Các độ chính xác khác nhau có thể được sử dụng để biểu diễn các vector chuyển động cuối cùng và các thành phần vector chuyển động. Ví dụ, độ chính xác mẫu 1/4 có thể được sử dụng cho các thành phần vector chuyển động, nhưng khi đánh giá các vector chuyển động cuối cùng cho các vị trí khác nhau trong đơn vị mã hóa, thì ví dụ, độ chính xác mẫu 1/16 có thể được sử dụng.

Các vector chuyển động có thể được đánh giá cho các điểm ảnh riêng lẻ trong đơn vị mã hóa hoặc tập mẫu. Ví dụ, vector chuyển động có thể được tính toán cho mỗi tập mẫu 4x4.

Trường vector chuyển động có thể được biểu diễn theo các cách khác nhau. Ví dụ, nó có thể được biểu diễn dưới dạng hàm đa thức của các tọa độ điểm ảnh hoặc có thể được biểu diễn dưới dạng các vector chuyển động ở các góc của các đơn vị mã hóa hoặc có thể được biểu diễn dưới dạng các vector chuyển động cho một số vị trí khác bên trong, bên ngoài hoặc trên ranh giới của đơn vị mã hóa.

Các tham số khác nhau có thể được báo hiệu trong luồng bit hoặc được xác định bên ngoài luồng bit. Ví dụ, hệ số tinh chỉnh cho vector chuyển động tịnh tiến cơ sở có thể được báo hiệu trong luồng bit hoặc có thể được thiết đặt đến một giá trị nhất định (ví dụ, không) bằng cách khác.

Phương pháp này có thể sử dụng chế độ dự đoán đơn với một tham chiếu đơn lẻ, chế độ dự đoán kép với hai tham chiếu hoặc chế độ dự đoán đa giả thuyết với nhiều hơn hai khối tham chiếu để dự đoán bù chuyển động.

Trong bộ giải mã, các bước tương tự có thể được thực hiện để tái dựng các mẫu của khối, ngoại trừ việc bộ giải mã có thể không cần phải lựa chọn các vector chuyển động tịnh tiến cơ sở và tính toán các tham số vì bộ giải mã có thể thu thông tin này trong luồng bit từ bộ mã hóa. Bộ giải mã có thể thu 1000 thông tin liên quan đến đơn vị mã hóa hiện thời và các tham số liên quan đến chế độ dự đoán chuyển động và kiểm tra 1002 thông tin thu được để tìm ra thông tin về vector chuyển động tịnh tiến cơ sở b_0 mà bộ mã hóa đã sử dụng khi mã hóa. Bộ giải mã cũng có thể kiểm tra 1004 loại nào trong số các hệ số tinh chỉnh r_{bx} , r_{by} cho thành phần vector chuyển động tịnh tiến thứ nhất b_x , b_y và thành phần vector chuyển động bậc cao a_0 , a_1 , và các hệ số tinh chỉnh r_{a0} , r_{a1} cho thành phần vector chuyển động bậc cao a_0 , a_1 được xác định bởi bộ mã hóa. Bộ giải mã cũng có thể sử dụng thông tin thu được để tìm ra giá trị của độ chính xác thứ nhất và độ chính xác thứ hai mà bộ mã hóa đã sử dụng khi mã hóa. Sau đó, bộ giải mã có thể sử dụng thông tin này để tính toán 1006 vector chuyển động thứ nhất cơ sở. Tương tự, bộ giải mã có thể tính toán 1008 các vector chuyển động còn lại để xác định trường của các vector chuyển động cho đơn vị mã hóa hiện thời. Theo một phương án của sáng chế, trường của các vector chuyển động bao gồm vector chuyển động cho mỗi mẫu của đơn vị mã hóa hiện thời. Sau đó, chế độ dự đoán chuyển động có thể được thực hiện 1010 bằng bộ giải mã bằng cách sử dụng trường của các vector chuyển động để tái dựng các mẫu của đơn vị mã hóa hiện thời.

Fig.4 thể hiện sơ đồ khối của bộ mã hóa video theo một 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 có thể được mở rộng một cách tương tự để mã hóa nhiều hơn hai lớp, hoặc chỉ một lớp. Fig.4 minh họa một phương án về bộ mã hóa video bao gồm phần mã hóa thứ nhất 500 cho lớp cơ sở và phần mã hóa thứ hai 502 cho lớp tăng cường. Mỗi trong số phần mã hóa thứ nhất 500 và phần mã hóa thứ hai 502 có thể bao gồm các phần tử tương tự để mã hóa các ảnh đầu vào. Các phần 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 về bộ dự đoán điểm ảnh 302, 402 trong đó bao gồm bộ dự đoán liên ảnh 306, 406, bộ dự đoán nội ảnh 308, 408, bộ lựa 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 mã hóa thứ nhất 500 thu 300 các hình ảnh lớp cơ sở của luồng video cần mã hóa tại cả bộ dự đoán liên ảnh 306 (xác định sự khác biệt giữa hình ảnh và khung tham chiếu bù chuyển động 318) và bộ dự đoán nội ảnh 308 (xác định chế độ dự đoán cho khối hình ảnh chỉ dựa trên các phần đã xử lý của ảnh hoặc khung hiện thời). Đầu ra của cả bộ dự đoán liên ảnh và bộ dự đoán nội ảnh được đưa vào bộ lựa 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 đã dự đoán cho bộ lựa chọn chế độ 310. Bộ lựa chọn chế độ 310 cũng thu bản sao của ảnh lớp cơ sở 300. Một cách tương ứng, bộ dự đoán điểm ảnh 402 của phần mã hóa thứ hai 502 thu 400 các hình ảnh lớp tăng cường của luồng video cần mã hóa tại cả bộ dự đoán liên ảnh 406 (xác định sự khác biệt giữa hình ảnh và khung tham chiếu bù chuyển động 418) và bộ dự đoán nội ảnh 408 (xác định chế độ dự đoán cho khối hình ảnh chỉ dựa trên các phần đã xử lý của ảnh hoặc khung hiện thời). Đầu ra của cả bộ dự đoán liên ảnh và bộ dự đoán nội ảnh được đưa vào bộ lựa 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 đã dự đoán cho bộ lựa chọn chế độ 410. Bộ lựa chọn chế độ 410 cũng thu bản sao của ảnh lớp tăng cường 400.

Phụ thuộc vào chế độ mã hóa được lựa chọn để mã hóa đơn vị mã hóa hiện thời, đầu ra của bộ dự đoán liên ảnh 306, 406 hoặc đầu ra của một trong số các chế độ dự đoán nội ảnh tùy chọn hoặc đầu ra của bộ mã hóa bề mặt trong bộ lựa chọn chế độ được đưa vào đầu ra của bộ lựa chọn chế độ 310, 410. Đầu ra của bộ lựa chọn chế độ được đưa vào phương tiện tính tổng thứ nhất 321, 421. Phương tiện tính tổng thứ nhất này có thể trừ đầu ra của bộ dự đoán điểm ảnh 302, 402 từ ảnh lớp cơ sở 300/ảnh lớp tăng cường 400 để tạo ra tín hiệu lỗi dự đoán thứ nhất 320, 420 mà được nhập vào bộ mã hóa lỗi dự đoán 303, 403.

Bộ dự đoán điểm ảnh 302, 402 cũng thu từ bộ tái dựng sơ bộ 339, 439 tổ hợp của dạng biểu diễn dự đoán của khối hình ảnh 312, 412 và đầu ra 338, 438 của bộ giải mã

lỗi dự đoán 304, 404. Hình ảnh đã tái dựng sơ bộ 314, 414 có thể được đưa vào bộ dự đoán nội ảnh 308, 408 và vào bộ lọc 316, 416. Bộ lọc 316, 416 thu dạng biểu diễn sơ bộ có thể lọc dạng biểu diễn sơ bộ này và xuất ra hình ảnh đã tái dựng cuối cùng 340, 440 mà có thể được lưu lại 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 để được sử dụng làm hình ảnh tham chiếu mà ảnh lớp cơ sở tương lai 300 được so sánh dựa vào đó trong các bước dự đoán liên ảnh. Tùy vào lớp cơ sở được lựa chọn và được chỉ báo là nguồn để dự đoán mẫu liên lớp và/hoặc dự đoán thông tin chuyển động liên lớp cho 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 để được sử dụng làm hình ảnh tham chiếu mà các ảnh lớp tăng cường tương lai 400 được so sánh dựa vào đó trong các bước dự đoán liên ảnh. Ngoài ra, bộ nhớ khung tham chiếu 418 có thể được kết nối với bộ dự đoán liên ảnh 406 để được sử dụng làm hình ảnh tham chiếu mà ảnh lớp tăng cường tương lai 400 được so sánh dựa vào đó trong các bước dự đoán liên ảnh.

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

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 thành miền biến đổi. 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 đã lượng tử hóa.

Bộ giải mã lỗi dự đoán 304, 404 thu đầu ra của bộ mã hóa lỗi dự đoán 303, 403 và thực hiện các quy trình đối lập với bộ mã hóa lỗi dự đoán 303, 403 để tạo ra tín hiệu lỗi dự đoán đã giải mã 338, 438 mà, khi được kết hợp với dạng biểu diễn dự đoán của khối hình ảnh 312, 412 tại phương tiện tính tổng thứ hai 339, 439, tạo ra hình ảnh đã tái dựng sơ bộ 314, 414. Bộ giải mã lỗi dự đoán có thể được coi là bao gồm bộ giải lượng tử hóa 361, 461, mà giải lượng tử hóa các giá trị hệ số đã lượng tử hóa, ví dụ, các hệ số DCT, để tái dựng tín hiệu biến đổi và đơn vị biến đổi ngược 363, 463, mà thực hiện việc

biến đổi ngược cho tín hiệu biến đổi đã tái dựng trong đó đầu ra của đơn vị biến đổi ngược 363, 463 chứa (các) khối đã tái dựng. Bộ giải mã lỗi dự đoán cũng có thể bao gồm bộ lọc khối mà có thể lọc (các) khối đã tái dựng theo thông tin đã giải mã và các tham số lọc bổ sung.

Bộ mã hóa entropy 330, 430 thu đầu ra của bộ mã hóa lỗi dự đoán 303, 403 và có thể thực hiện bước mã hóa entropy/mã hóa chiều dài biến số thích hợp đối với tín hiệu để cung cấp sự phát hiện lỗi và khả năng hiệu chỉnh. Đầu ra của các bộ mã hóa entropy 330, 430 có thể được đưa vào luồng bit, ví dụ, bằng bộ dồn kênh 508.

Fig.5 thể hiện sơ đồ khối của bộ giải mã video theo một phương án của sáng chế. Fig.5 mô tả cấu trúc của bộ giải mã hai lớp, nhưng cần hiểu rằng các bước 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 giải mã thứ nhất 552 cho các thành phần hiển thị cơ sở và phần giải mã thứ hai 554 cho các thành phần hiển thị không cơ sở. Khối 556 minh họa bộ giải dồn kênh để phân phối thông tin liên quan đến các thành phần hiển thị cơ sở đến phần giải mã thứ nhất 552 và để phân phối thông tin liên quan đến các thành phần hiển thị không cơ sở đến phần giải mã thứ hai 554. Tham chiếu P'n đại diện cho dạng biểu diễn đã dự đoán của khối hình ảnh. Tham chiếu D'n đại diện cho tín hiệu lỗi dự đoán đã tái dựng. Các khối 704, 804 minh họa các hình ảnh đã tái dựng sơ bộ (I'n). Tham chiếu R'n đại diện cho hình ảnh đã tái dựng cuối cùng. Các khối 703, 803 minh họa bước biến đổi ngược (T^{-1}). Các khối 702, 802 minh họa bước lượng tử hóa ngược (Q^{-1}). Các khối 701, 801 minh họa bước giải mã entropy (E^{-1}). Các khối 706, 806 minh họa bộ nhớ khung tham chiếu (RFM). Các khối 707, 807 minh họa bước dự đoán (P) (dự đoán liên ảnh hoặc chế độ dự đoán nội ảnh). Các khối 708, 808 minh họa bước lọc (F). Các khối 709, 809 có thể được sử dụng để kết hợp thông tin lỗi dự đoán đã giải mã với các thành phần hiển thị cơ sở/hiển thị không cơ sở đã dự đoán để thu các hình ảnh đã tái dựng sơ bộ 705, 805 (I'n). Các hình ảnh hiển thị cơ sở đã tái dựng sơ bộ và đã lọc có thể là đầu ra 710 từ phần giải mã thứ nhất 552 và các hình ảnh hiển thị cơ sở đã tái dựng sơ bộ và đã lọc có thể là đầu ra 810 từ phần giải mã thứ hai 554.

Mô hình kiểm tra nội dung màn hình H.265/HEVC mô tả quá trình thực hiện sao chép khối nội ảnh đơn giản. Phương pháp sao chép khối nội ảnh tạo ra chế độ dự đoán

cho khối mẫu bằng cách sao chép một vùng nhất định của cùng hình ảnh đến vùng đã dự đoán. Vị trí của vùng nguồn dùng cho bước sao chép có thể được thể hiện bằng vector khối mà có thể được mã hóa khác với vector khối đã dự đoán được lựa chọn từ tập vector khối ứng viên được tạo ra bởi các vector khối của một số khối đã mã hóa trước đó.

Fig.11 là đồ thị biểu diễn hệ thống truyền thông đa phương tiện ví dụ, trong đó các phương án của sáng chế có thể được thực hiện. Nguồn dữ liệu 1110 cung cấp tín hiệu nguồn ở định dạng tương tự, định dạng số đã giải nén hoặc định dạng số đã nén, hoặc tổ hợp bất kỳ của các định dạng này. Bộ mã hóa 1120 có thể bao gồm hoặc được kết nối với bộ phận xử lý sơ bộ, như chuyển đổi định dạng dữ liệu và/hoặc lọc tín hiệu nguồn. Bộ mã hóa 1120 mã hóa tín hiệu nguồn thành luồng bit đa phương tiện đã mã hóa. Cần lưu ý rằng luồng bit cần giải mã có thể được thu trực tiếp hoặc gián tiếp từ thiết bị từ xa được đặt trong loại mạng thực bất kỳ. Ngoài ra, luồng bit có thể được thu từ phần mềm hoặc phần cứng cục bộ. Bộ mã hóa 1120 có thể có khả năng mã hóa nhiều hơn một loại nội dung đa phương tiện, như audio và video, hoặc có thể cần đến nhiều hơn một bộ mã hóa 1120 để mã hóa các loại nội dung đa phương tiện khác nhau của tín hiệu nguồn. Bộ mã hóa 1120 cũng có thể nhận đầu vào được tạo ra một cách tổng hợp, như đồ họa và văn bản, hoặc có thể có khả năng tạo ra các luồng bit đã mã hóa của nội dung đa phương tiện tổng hợp. Trong phần dưới đây, chỉ có quá trình xử lý một luồng bit đa phương tiện đã mã hóa của một loại nội dung đa phương tiện được xem xét để làm đơn giản việc mô tả. Tuy nhiên, cần lưu ý rằng thông thường các dịch vụ phát rộng thời gian thực bao gồm một số luồng (thường là ít nhất một luồng audio, video và phụ đề bằng 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 1120 được thể hiện để đơn giản hóa việc mô tả mà không làm mất tính tổng quát. Cần hiểu thêm rằng, mặc dù văn cảnh và các ví dụ nêu trong đó có thể mô tả một cách cụ thể một quy trình mã hóa, nhưng người có hiểu biết trung bình trong lĩnh vực này cần hiểu rằng các khái niệm và nguyên tắc tương tự cũng áp dụng cho quy trình giải mã tương ứng và ngược lại.

Luồng bit đa phương tiện đã mã hóa có thể được truyền đến thiết bị lưu trữ 1130. Thiết bị lưu trữ 1130 có thể bao gồm loại bộ nhớ khối bất kỳ để lưu trữ luồng bit đa phương tiện đã mã hóa. Định dạng của luồng bit đa phương tiện đã mã hóa trong thiết

bị lưu trữ 1130 có thể là định dạng luồng bit tự chứa cơ bản, hoặc một hoặc nhiều luồng bit đa phương tiện đã mã hóa có thể được bao gói trong tệp chứa. Khi một hoặc nhiều luồng bit đa phương tiện được bao gói trong tệp chứa, thì bộ tạo tệp (không được thể hiện trên hình vẽ) có thể được sử dụng để lưu trữ một hoặc nhiều luồng bit đa phương tiện này trong tệp và tạo siêu dữ liệu định dạng tệp, mà cũng có thể được lưu trữ trong tệp. Bộ mã hóa 1120 hoặc thiết bị lưu trữ 1130 có thể bao gồm bộ tạo tệp, hoặc bộ tạo tệp được gắn theo cách có thể hoạt động được vào bộ mã hóa 1120 hoặc thiết bị lưu trữ 1130. Một số hệ thống vận hành “trực tiếp”, nghĩa là bỏ qua việc lưu trữ và truyền luồng bit đa phương tiện đã mã hóa từ bộ mã hóa 1120 trực tiếp đến thiết bị gửi 1140. Sau đó, luồng bit đa phương tiện đã mã hóa có thể được truyền đến thiết bị gửi 1140, còn được gọi là máy chủ, khi cần. Định dạng được sử dụng trong quá trình truyền có thể là định dạng luồng bit tự chứa cơ bản, định dạng luồng gói, hoặc một hoặc nhiều luồng bit đa phương tiện đã mã hóa có thể được bao gói trong tệp chứa. Bộ mã hóa 1120, thiết bị lưu trữ 1130 và máy chủ 1140 có thể nằm trên cùng một thiết bị vật lý hoặc chúng có thể được chứa trong các thiết bị riêng biệt. Bộ mã hóa 1120 và máy chủ 1140 có thể vận hành với nội dung thời gian thực trực tiếp, trong trường hợp này, luồng bit đa phương tiện đã mã hóa thường không được lưu trữ vĩnh viễn, mà được đệm trong các khoảng thời gian ngắn trong bộ mã hóa nội dung 1120 và/hoặc trong máy chủ 1140 để làm mượt các thay đổi về độ trễ xử lý, độ trễ truyền và tốc độ bit đa phương tiện đã mã hóa.

Máy chủ 1140 có thể gửi luồng bit đa phương tiện đã mã hóa bằng cách sử dụng chồng giao thức truyền thông. Chồng này có thể bao gồm, nhưng không giới hạn ở, một hoặc nhiều trong số giao thức truyền tải thời gian thực (RTP: Real-Time Transport Protocol), giao thức gói dữ liệu người dùng (UDP: User Datagram Protocol), giao thức truyền dữ liệu siêu văn bản (HTTP: Hypertext Transfer Protocol), giao thức điều khiển truyền (TCP: Transmission Control Protocol) và giao thức Internet (IP: Internet Protocol). Khi chồng giao thức truyền thông được định hướng gói, thì máy chủ 1140 bao gói luồng bit đa phương tiện đã mã hóa thành các gói. Ví dụ, khi RTP được sử dụng, máy chủ 1140 bao gói luồng bit đa phương tiện đã mã hóa thành các gói RTP theo định dạng tải RTP. Thông thường, mỗi loại nội dung đa phương tiện có một định dạng tải RTP riêng. Cần lưu ý lại rằng hệ thống có thể bao gồm nhiều hơn một máy chủ 1140, nhưng để đơn giản, phần mô tả dưới đây chỉ xem xét một máy chủ 1140.

Khi nội dung đa phương tiện được bao gói trong tệp chứa cho thiết bị lưu trữ 1130 hoặc để nhập dữ liệu vào thiết bị gửi 1140, thiết bị gửi 1140 có thể bao gồm hoặc được gắn theo cách có thể hoạt động được vào "bộ phân tách tệp gửi" (không được thể hiện trên hình vẽ). Cụ thể, khi tệp chứa không được truyền như vậy nhưng ít nhất một trong số các luồng bit đa phương tiện đã mã hóa chứa trong đó được bao gói để truyền thông qua giao thức truyền thông, thì bộ phân tách tệp gửi định vị các phần thích hợp của luồng bit đa phương tiện đã mã hóa cần truyền thông qua giao thức truyền thông này. Bộ phân tách tệp gửi cũng có thể hỗ trợ việc tạo ra định dạng chính xác cho giao thức truyền thông, như các tải và phần đầu gói. Tệp chứa đa phương tiện có thể chứa các lệnh bao gói, như các dữ liệu gợi ý ở định dạng tệp đa phương tiện cơ sở (Base Media File Format) ISO, để bao gói ít nhất một trong số các luồng bit đa phương tiện chứa trong đó trên giao thức truyền thông.

Máy chủ 1140 có thể hoặc có thể không được kết nối với cổng nối 1150 thông qua mạng truyền thông. Cổng nối còn có thể hoặc theo cách khác được gọi là hộp trung gian. Cần lưu ý rằng hệ thống này thường có thể bao gồm số cổng nối bất kỳ hoặc tương tự, nhưng để đơn giản, phần mô tả dưới đây chỉ xem xét một cổng nối 1150. Cổng nối 1150 có thể thực hiện các dạng chức năng khác nhau, như lưu trữ đệm các gói, luồng hoặc tài nguyên, tìm nạp trước dữ liệu đa phương tiện một cách chủ động, tịnh tiến luồng gói theo một chồng giao thức truyền thông sang chồng giao thức truyền thông khác, sáp nhập và chia nhánh các luồng dữ liệu, và thao tác luồng dữ liệu theo khả năng của đường xuống và/hoặc bộ thu, như điều khiển tốc độ bit của luồng chuyển tiếp theo điều kiện mạng đường xuống hiện hành. Các ví dụ về cổng nối 1150 bao gồm các đơn vị điều khiển hội nghị đa điểm (MCU), các cổng nối giữa hệ thống điện thoại video chuyển mạch và chuyển gói, các máy chủ ấn để đàm qua mạng di động (PoC), các bộ bao gói IP trong các hệ thống phát rộng video số cầm tay (DVB-H), hoặc các hộp thu giải mã tín hiệu truyền hình hoặc các thiết bị khác mà chuyển tiếp các cuộc truyền phát rộng cục bộ đến các mạng không dây gia đình. Khi RTP được sử dụng, cổng nối 1150 có thể được gọi là bộ trộn RTP hoặc bộ dịch RTP và có thể hoạt động như một điểm đầu của kết nối RTP. Thay cho hoặc ngoài cổng nối 1150, hệ thống này có thể bao gồm bộ nối để nối chuỗi video hoặc các luồng bit.

Hệ thống này bao gồm một hoặc nhiều bộ thu 1160, thường có khả năng thu, giải điều biến, và mở gói tín hiệu đã truyền thành luồng bit đa phương tiện đã mã hóa. Luồng bit đa phương tiện đã mã hóa có thể được truyền đến thiết bị ghi 1170. Thiết bị ghi 1170 có thể bao gồm loại bộ nhớ khối bất kỳ để lưu trữ luồng bit đa phương tiện đã mã hóa. Ngoài ra hoặc theo cách khác, thiết bị ghi 1170 có thể còn bao gồm bộ nhớ điện toán, như bộ nhớ truy cập ngẫu nhiên. Định dạng của luồng bit đa phương tiện đã mã hóa trong thiết bị ghi 1170 có thể là định dạng luồng bit tự chứa cơ bản, hoặc một hoặc nhiều luồng bit đa phương tiện đã mã hóa có thể được bao gói trong tệp chứa. Nếu có nhiều luồng bit đa phương tiện đã mã hóa, như luồng audio và luồng video, liên kết với nhau, tệp chứa thường được sử dụng và bộ thu 1160 bao gồm hoặc được gắn vào bộ tạo tệp chứa mà tạo ra tệp chứa từ các luồng đầu vào. Một số hệ thống hoạt động “trực tiếp”, nghĩa là bỏ qua thiết bị ghi 1170 và chuyển luồng bit đa phương tiện đã mã hóa từ bộ thu 1160 một cách trực tiếp sang bộ giải mã 1180. Trong một số hệ thống, chỉ có phần gần đây nhất của luồng đã ghi lại, ví dụ, trích đoạn 10 phút gần nhất của luồng đã ghi lại, là được giữ lại trong thiết bị ghi 1170, trong khi dữ liệu đã ghi lại trước đó bất kỳ đều bị loại ra khỏi thiết bị ghi 1170.

Luồng bit đa phương tiện đã mã hóa có thể được truyền từ thiết bị ghi 1170 sang bộ giải mã 1180. Khi có nhiều luồng bit đa phương tiện đã mã hóa, như luồng audio và luồng video, liên kết với nhau và được bao gói trong tệp chứa hoặc một luồng bit đa phương tiện được bao gói trong tệp chứa, ví dụ, để dễ truy cập, thì bộ phân tách tệp (không được thể hiện trên hình vẽ) được sử dụng để mở gói từng luồng bit đa phương tiện đã mã hóa ra khỏi tệp chứa. Thiết bị ghi 1170 hoặc bộ giải mã 1180 có thể bao gồm bộ phân tách tệp, hoặc bộ phân tách tệp được gắn vào thiết bị ghi 1170 hoặc bộ giải mã 1180. Cũng cần lưu ý rằng hệ thống này có thể bao gồm nhiều bộ giải mã, nhưng ở đây chỉ có một bộ giải mã 1170 được xem xét để đơn giản hóa việc mô tả mà không làm mất tính tổng quát.

Luồng bit đa phương tiện đã mã hóa có thể được xử lý tiếp bởi bộ giải mã 1170, mà đầu ra của nó là một hoặc nhiều luồng đa phương tiện đã giải nén. Cuối cùng, bộ kết xuất 1190 có thể tái dựng các luồng đa phương tiện đã giải nén bằng, ví dụ, loa hoặc màn hiển thị. Bộ thu 1160, thiết bị ghi 1170, bộ giải mã 1170, và bộ kết xuất 1190 có

thể nằm trên cùng một thiết bị vật lý hoặc chúng có thể được chứa trong các thiết bị riêng biệt.

Trong phần nêu trên, một số phương án được mô tả theo các thuật ngữ dạng biểu diễn và/hoặc luồng bit. Cần hiểu rằng các phương án này có thể được mô tả một cách tương tự theo các thuật ngữ tương tự, ví dụ, dạng biểu diễn thay vì luồng bit, luồng bit thay vì dạng biểu diễn, hoặc luồng cơ bản thay vì luồng bit hoặc dạng biểu diễn.

Trong phần nêu trên, khi các phương án ví dụ được mô tả với viện dẫn đến bộ mã hóa, cần hiểu rằng luồng bit thu được và bộ giải mã có thể có các phần tử tương ứng trong đó. Tương tự, khi các phương án ví dụ được mô tả với viện dẫn đến bộ giải mã, cần hiểu rằng bộ mã hóa có thể có cấu trúc và/hoặc chương trình máy tính để tạo ra luồng bit cần giải mã bởi bộ giải mã.

Các phương án của sáng chế nêu trên mô tả bộ mã hóa - giải mã theo khía cạnh bộ mã hóa và bộ giải mã riêng biệt để hỗ trợ cho việc hiểu các quy trình có liên quan. Tuy nhiên, cần hiểu rằng các thiết bị, các cấu trúc và các bước có thể được thực hiện dưới dạng một bộ mã hóa - giải mã/cấu trúc/bước riêng lẻ. Ngoài ra, có thể là bộ mã hóa và bộ giải mã có thể có chung một số hoặc tất cả các thành phần chung.

Mặc dù các ví dụ nêu trên mô tả các phương án của sáng chế hoạt động trong bộ mã hóa - giải mã trong thiết bị điện tử, nhưng cần hiểu rằng sáng chế được xác định theo các yêu cầu bảo hộ có thể được thực hiện dưới dạng một phần của bộ mã hóa - giải mã video bất kỳ. Do đó, ví dụ, các phương án của sáng chế có thể được thực hiện trong bộ mã hóa - giải mã video mà có thể thực thi việc mã hóa video trên các đường truyền thông có dây hoặc cố định.

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

Ngoài ra, các thành phần của mạng di động mặt đất công cộng (PLMN) cũng có thể bao gồm các bộ mã hóa - giải mã video nêu trên.

Nói chung, các phương án khác nhau của sáng chế có thể được thực hiện trong phần cứng hoặc mạch chuyên dụng, phần mềm, logic hoặc tổ hợp bất kỳ của chúng. Ví dụ, một số khía cạnh có thể được thực hiện trong phần cứng, trong khi các khía cạnh khác có thể được thực hiện trong phần sụn hoặc phần mềm mà có thể được thực thi bằng bộ điều khiển, bộ vi xử lý hoặc thiết bị điện toán khác, mặc dù sáng chế không bị giới hạn ở đó. Mặc dù các khía cạnh khác nhau của sáng chế có thể được minh họa và được mô tả dưới dạng sơ đồ khối, lưu đồ, hoặc sử dụng một số dạng biểu diễn bằng hình ảnh khác, nhưng có thể hiểu rõ rằng các khối, thiết bị, hệ thống, kỹ thuật hoặc phương pháp được nêu trong bản mô tả này có thể được thực hiện trong, là các ví dụ không giới hạn, phần cứng, phần mềm, phần sụn, mạch chuyên dụng hoặc logic, phần cứng đa năng hoặc bộ điều khiển hoặc các thiết bị điện toán khác, hoặc tổ hợp bất kỳ của chúng.

Các phương án của sáng chế có thể được thực hiện bằng phần mềm máy tính có thể thực thi bằng bộ xử lý dữ liệu của thiết bị điện tử, như trong thực thể bộ xử lý, hoặc bằng phần cứng, hoặc bằng tổ hợp của phần mềm và phần cứng. Phần mềm máy tính có thể giúp các thiết bị liên quan, như bộ mã hóa hoặc bộ giải mã, thực hiện sáng chế. Ví dụ, thiết bị có thể bao gồm mạch và các thiết bị điện tử để xử lý, thu và truyền dữ liệu, mã chương trình máy tính trong bộ nhớ, và bộ xử lý mà, khi chạy mã chương trình máy tính, giúp thiết bị thực hiện các dấu hiệu của phương án của sáng chế. Ngoài ra, thiết bị mạng như máy chủ có thể bao gồm mạch và các thiết bị điện tử để xử lý, thu và truyền dữ liệu, mã chương trình máy tính trong bộ nhớ, và bộ xử lý mà, khi chạy mã chương trình máy tính, giúp thiết bị mạng thực hiện các dấu hiệu của phương án của sáng chế.

Các phương án khác nhau có thể được thực hiện với sự trợ giúp của vật ghi đọc được bằng máy tính không khả biến đã được mã hóa với các lệnh mà, khi được thực thi bởi máy tính, sẽ thực hiện các phương án khác nhau này.

Nếu muốn, các chức năng khác nhau được thảo luận trong bản mô tả này có thể được thực hiện theo thứ tự khác nhau và/hoặc cùng lúc với nhau. Ngoài ra, nếu muốn, một hoặc nhiều chức năng trong số các chức năng nêu trên có thể là tùy chọn hoặc có thể được kết hợp. Ngoài ra, các phương án này được bộc lộ liên quan đến phương pháp giải mã và bộ giải mã. Tuy nhiên, phần bộc lộ này có thể được ứng dụng trong bộ mã

hóa được tạo cấu hình để thực hiện việc mã hóa các đơn vị mã hóa và mã hóa chỉ báo sự tồn tại của mã hóa thoát trong đơn vị mã hóa.

Ngoài ra, về vấn đề này, cần lưu ý rằng các khối bất kỳ của dòng logic như được thể hiện trên các hình vẽ có thể biểu diễn các bước chương trình, hoặc các mạch logic liên kết với nhau, các khối và các chức năng, hoặc tổ hợp của các bước chương trình và các mạch logic, các khối và các chức năng. Phần mềm có thể được lưu trữ trên vật ghi vật lý như chip bộ nhớ, hoặc các khối bộ nhớ được thực thi bên trong bộ xử lý, vật ghi từ tính như đĩa cứng hoặc đĩa mềm, và vật ghi quang học như ví dụ DVD và các biến thể dữ liệu của chúng, CD.

Bộ nhớ có thể là loại bất kỳ thích hợp đối với môi trường kỹ thuật cục bộ và có thể được thực thi bằng cách sử dụng công nghệ lưu trữ dữ liệu thích hợp bất kỳ, như thiết bị nhớ trên cơ sở bán dẫn, hệ thống và thiết bị nhớ từ tính, hệ thống và thiết bị nhớ quang học, bộ nhớ cố định và bộ nhớ có thể tháo được. Các bộ xử lý dữ liệu có thể là loại bất kỳ thích hợp đối với môi trường kỹ thuật cục bộ, và có thể bao gồm một hoặc nhiều thiết bị trong số máy tính đa năng, máy tính chuyên dụng, bộ vi xử lý, bộ xử lý tín hiệu số (DSP) và bộ xử lý dựa trên kiến trúc bộ xử lý đa lõi, là các ví dụ không giới hạn.

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

Các chương trình, như được cung cấp bởi Synopsys, Inc. tại Mountain View, California và Cadence Design tại San Jose, California tự động định tuyến các đường dẫn và định vị các thành phần trên chip bán dẫn bằng cách sử dụng các quy tắc đã thiết lập rõ của thiết kế cũng như các thư viện mô đun thiết kế đã lưu trữ trước đó. Khi thiết kế mạch bán dẫn đã được hoàn thành, thiết kế tạo thành, ở định dạng điện tử đã chuẩn hóa (ví dụ, Opus, GDSII, hoặc dạng tương tự) có thể được chuyển đến cơ sở chế tạo bán dẫn hoặc "fab" để chế tạo.

Phần mô tả nêu trên cung cấp mô tả đầy đủ và nhiều thông tin về phương án ví dụ của sáng chế làm ví dụ minh họa và không giới hạn. Tuy nhiên, các cải biến và sửa

đôi khác nhau có thể trở nên rõ ràng với người có hiểu biết trung bình trong lĩnh vực kỹ thuật liên quan khi xem phần mô tả nêu trên, khi đọc cùng với các hình vẽ kèm theo và yêu cầu bảo hộ đi kèm. Tuy nhiên, tất cả các cải biến như vậy và tương tự của sáng chế vẫn sẽ nằm trong phạm vi của sáng chế.

Các phương án này có thể tạo ra một số ưu điểm. Ví dụ, độ chính xác của chế độ dự đoán bù chuyển động có thể được cải thiện bằng cách đề xuất cách có hiệu quả về tốc độ bit để báo hiệu thông tin bổ sung cần có để lập mô hình các trường vector chuyển động bậc cao.

Theo phương án ví dụ thứ nhất, sáng chế đề xuất phương pháp bao gồm các bước:

thu khối mẫu để dự đoán chuyển động;

lựa chọn vector chuyển động tịnh tiến cơ sở cho khối mẫu;

xác định chế độ dự đoán cho ít nhất thành phần vector chuyển động tịnh tiến cơ sở;

xác định hệ số tinh chỉnh thứ nhất cho thành phần vector chuyển động tịnh tiến cơ sở bằng cách sử dụng độ chính xác thứ nhất;

xác định chế độ dự đoán cho ít nhất một thành phần vector chuyển động bậc cao;

xác định hệ số tinh chỉnh thứ hai cho thành phần vector chuyển động bậc cao bằng cách sử dụng độ chính xác thứ hai khác với độ chính xác thứ nhất; và

mã hóa thông tin về thành phần vector chuyển động tịnh tiến cơ sở, hệ số tinh chỉnh thứ nhất và hệ số tinh chỉnh thứ hai thành một hoặc nhiều luồng bit.

Theo phương án ví dụ thứ hai, sáng chế đề xuất phương pháp bao gồm các bước:

thu khối mẫu để dự đoán chuyển động;

lựa chọn vector chuyển động tịnh tiến cơ sở có độ chính xác thứ nhất cho khối mẫu;

xác định chế độ dự đoán cho ít nhất một thành phần vector chuyển động bậc cao;

xác định hệ số tinh chỉnh cho thành phần vector chuyển động bậc cao bằng cách sử dụng độ chính xác thứ hai khác với độ chính xác thứ nhất; và

mã hóa thông tin về thành phần vectơ chuyển động tịnh tiến cơ sở, và hệ số tinh chỉnh thành một hoặc nhiều luồng bit.

Theo một số phương án, phương pháp theo phương án ví dụ thứ nhất và/hoặc thứ hai bao gồm các bước nêu dưới đây:

tính toán vectơ chuyển động dựa trên thành phần vectơ chuyển động cơ sở, hệ số tinh chỉnh cho thành phần vectơ chuyển động cơ sở, thành phần vectơ chuyển động bậc cao, và hệ số tinh chỉnh cho thành phần vectơ chuyển động bậc cao.

Theo một số phương án, phương pháp theo phương án ví dụ thứ nhất và/hoặc thứ hai bao gồm các bước:

nhân hệ số tinh chỉnh thứ nhất với thừa số thứ nhất để thu hệ số tinh chỉnh thứ nhất đã mở rộng, nếu độ chính xác của vectơ chuyển động tịnh tiến cơ sở và hệ số tinh chỉnh thứ nhất không giống nhau; và

sử dụng hệ số tinh chỉnh thứ nhất đã mở rộng trong bước tính toán thành phần vectơ chuyển động tịnh tiến.

Theo một số phương án, phương pháp theo phương án ví dụ thứ nhất và/hoặc thứ hai bao gồm các bước:

nhân hệ số tinh chỉnh thứ hai với thừa số thứ hai để thu hệ số tinh chỉnh thứ hai đã mở rộng, nếu độ chính xác của vectơ chuyển động bậc cao và hệ số tinh chỉnh thứ hai không giống nhau; và

sử dụng hệ số tinh chỉnh thứ hai đã mở rộng trong bước tính toán thành phần vectơ chuyển động bậc cao.

Theo một số phương án, phương pháp theo phương án ví dụ thứ nhất và/hoặc thứ hai bao gồm bước:

sử dụng mô hình chuyển động afin hạn chế có khả năng xử lý xoay và thu phóng để xác định chế độ dự đoán cho thành phần vectơ chuyển động bậc cao.

Theo một số phương án, phương pháp theo phương án ví dụ thứ nhất và/hoặc thứ hai bao gồm bước:

sử dụng mô hình chuyển động afin đầy đủ có khả năng xử lý xoay và thu phóng để xác định chế độ dự đoán cho thành phần vector chuyển động bậc cao.

Theo một số phương án, phương pháp theo phương án ví dụ thứ nhất và/hoặc thứ hai bao gồm bước:

báo hiệu trường vector chuyển động bằng cách sử dụng các vector chuyển động góc của đơn vị mã hóa.

Theo một số phương án của phương pháp theo phương án ví dụ thứ nhất và/hoặc thứ hai, bước mã hóa bao gồm:

báo hiệu ít nhất một phần thông tin trong tập tham số hoặc trong phần đầu ảnh, trong phần đầu lát hoặc trong phần đầu ô.

Theo một số phương án của phương pháp theo phương án ví dụ thứ nhất và/hoặc thứ hai, độ chính xác thứ nhất thấp hơn độ chính xác thứ hai.

Theo một số phương án của phương pháp này, độ chính xác thứ nhất cao hơn độ chính xác thứ hai.

Theo một số phương án, phương pháp theo phương án ví dụ thứ nhất và/hoặc thứ hai còn bao gồm bước mã hóa thông tin của ít nhất một trong số độ chính xác thứ nhất và độ chính xác thứ hai thành một hoặc nhiều luồng bit.

Theo phương án ví dụ thứ ba, sáng chế đề xuất thiết bị bao gồm ít nhất một bộ xử lý; và ít nhất một bộ nhớ bao gồm mã chương trình máy tính, trong đó ít nhất một bộ nhớ và mã chương trình máy tính này được tạo cấu hình để, với ít nhất một bộ xử lý, giúp thiết bị có thể thực hiện ít nhất các bước:

thu khối mẫu để dự đoán chuyển động;

lựa chọn vector chuyển động tịnh tiến cơ sở cho khối mẫu;

xác định chế độ dự đoán cho ít nhất thành phần vector chuyển động tịnh tiến cơ sở;

xác định hệ số tinh chỉnh thứ nhất cho thành phần vector chuyển động tịnh tiến cơ sở bằng cách sử dụng độ chính xác thứ nhất;

xác định chế độ dự đoán cho ít nhất một thành phần vector chuyển động bậc cao;

xác định hệ số tinh chỉnh thứ hai cho thành phần vector chuyển động bậc cao bằng cách sử dụng độ chính xác thứ hai khác với độ chính xác thứ nhất; và

mã hóa thông tin về thành phần vector chuyển động tịnh tiến cơ sở, hệ số tinh chỉnh thứ nhất và hệ số tinh chỉnh thứ hai thành một hoặc nhiều luồng bit.

Theo phương án ví dụ thứ tư, sáng chế đề xuất thiết bị bao gồm ít nhất một bộ xử lý; và ít nhất một bộ nhớ bao gồm mã chương trình máy tính, trong đó ít nhất một bộ nhớ và mã chương trình máy tính này được tạo cấu hình để, với ít nhất một bộ xử lý, giúp thiết bị thực hiện ít nhất các bước:

thu khối mẫu để dự đoán chuyển động;

lựa chọn vector chuyển động tịnh tiến cơ sở có độ chính xác thứ nhất cho khối mẫu;

xác định chế độ dự đoán cho ít nhất một thành phần vector chuyển động bậc cao;

xác định hệ số tinh chỉnh cho thành phần vector chuyển động bậc cao bằng cách sử dụng độ chính xác thứ hai khác với độ chính xác thứ nhất; và

mã hóa thông tin về thành phần vector chuyển động tịnh tiến cơ sở, và hệ số tinh chỉnh thành một hoặc nhiều luồng bit.

Theo một số phương án, thiết bị theo phương án ví dụ thứ ba và/hoặc thứ tư còn được giúp để thực hiện bước :

tính toán vector chuyển động dựa trên thành phần vector chuyển động cơ sở, hệ số tinh chỉnh cho thành phần vector chuyển động cơ sở, thành phần vector chuyển động bậc cao, và hệ số tinh chỉnh cho thành phần vector chuyển động bậc cao.

Theo một số phương án, thiết bị theo phương án ví dụ thứ ba và/hoặc thứ tư còn được giúp để thực hiện các bước:

nhân hệ số tinh chỉnh thứ nhất với thừa số thứ nhất để thu hệ số tinh chỉnh thứ nhất đã mở rộng, nếu độ chính xác của vector chuyển động tịnh tiến cơ sở và hệ số tinh chỉnh thứ nhất không giống nhau; và

sử dụng hệ số tinh chỉnh thứ nhất đã mở rộng trong bước tính toán thành phần vector chuyển động tịnh tiến.

Theo một số phương án, thiết bị theo phương án ví dụ thứ ba và/hoặc thứ tư còn được giúp để thực hiện các bước:

nhân hệ số tinh chỉnh thứ hai với thừa số thứ hai để thu hệ số tinh chỉnh thứ hai đã mở rộng, nếu độ chính xác của vector chuyển động bậc cao và hệ số tinh chỉnh thứ hai không giống nhau; và

sử dụng hệ số tinh chỉnh thứ hai đã mở rộng này trong bước tính toán thành phần vector chuyển động bậc cao.

Theo một số phương án, thiết bị theo phương án ví dụ thứ ba và/hoặc thứ tư còn được giúp để thực hiện bước:

sử dụng một trong số mô hình chuyển động afin hạn chế và mô hình chuyển động afin đầy đủ có khả năng xử lý xoay và thu phóng để xác định chế độ dự đoán cho thành phần vector chuyển động bậc cao.

Theo một số phương án, thiết bị theo phương án ví dụ thứ ba và/hoặc thứ tư còn được giúp để thực hiện bước:

báo hiệu trường vector chuyển động bằng cách sử dụng các vector chuyển động góc của đơn vị mã hóa.

Theo một số phương án, thiết bị theo phương án ví dụ thứ ba và/hoặc thứ tư còn được giúp để thực hiện bước:

báo hiệu ít nhất một phần thông tin trong tập tham số hoặc trong phần đầu ảnh, trong phần đầu lát hoặc trong phần đầu ô.

Theo phương án ví dụ thứ năm, sáng chế đề xuất vật ghi đọc được bằng máy tính có lưu trữ mã trên đó để sử dụng bằng thiết bị, mà khi được thực thi bởi bộ xử lý, sẽ giúp thiết bị có thể thực hiện các bước:

thu khối mẫu để dự đoán chuyển động;

lựa chọn vector chuyển động tịnh tiến cơ sở cho khối mẫu;

xác định chế độ dự đoán cho ít nhất thành phần vector chuyển động tịnh tiến cơ sở;

xác định hệ số tinh chỉnh thứ nhất cho thành phần vector chuyển động tịnh tiến cơ sở bằng cách sử dụng độ chính xác thứ nhất;

xác định chế độ dự đoán cho ít nhất một thành phần vector chuyển động bậc cao;

xác định hệ số tinh chỉnh thứ hai cho thành phần vector chuyển động bậc cao bằng cách sử dụng độ chính xác thứ hai khác với độ chính xác thứ nhất; và

mã hóa thông tin về thành phần vector chuyển động tịnh tiến cơ sở, hệ số tinh chỉnh thứ nhất và hệ số tinh chỉnh thứ hai thành một hoặc nhiều luồng bit.

Theo phương án ví dụ thứ sáu, sáng chế đề xuất vật ghi đọc được bằng máy tính có lưu trữ mã trên đó để sử dụng bằng thiết bị, mà khi được thực thi bởi bộ xử lý, sẽ giúp thiết bị thực hiện các bước:

thu khối mẫu để dự đoán chuyển động;

lựa chọn vector chuyển động tịnh tiến cơ sở có độ chính xác thứ nhất cho khối mẫu;

xác định chế độ dự đoán cho ít nhất một thành phần vector chuyển động bậc cao;

xác định hệ số tinh chỉnh cho thành phần vector chuyển động bậc cao bằng cách sử dụng độ chính xác thứ hai khác với độ chính xác thứ nhất; và

mã hóa thông tin về thành phần vector chuyển động tịnh tiến cơ sở, và hệ số tinh chỉnh thành một hoặc nhiều luồng bit.

Theo phương án thứ bảy, sáng chế đề xuất thiết bị bao gồm bộ mã hóa video được tạo cấu hình để mã hóa luồng bit chứa chuỗi hình ảnh, bộ mã hóa video này bao gồm:

phương tiện để thu khối mẫu để dự đoán chuyển động;

phương tiện để lựa chọn vector chuyển động tịnh tiến cơ sở cho khối mẫu;

phương tiện để xác định chế độ dự đoán cho ít nhất thành phần vector chuyển động tịnh tiến cơ sở;

phương tiện để xác định hệ số tinh chỉnh thứ nhất cho thành phần vector chuyển động tịnh tiến cơ sở bằng cách sử dụng độ chính xác thứ nhất;

phương tiện để xác định chế độ dự đoán cho ít nhất một thành phần vector chuyển động bậc cao;

phương tiện để xác định hệ số tinh chỉnh thứ hai cho thành phần vector chuyển động bậc cao bằng cách sử dụng độ chính xác thứ hai khác với độ chính xác thứ nhất; và

phương tiện để mã hóa thông tin về thành phần vector chuyển động tịnh tiến cơ sở, hệ số tinh chỉnh thứ nhất và hệ số tinh chỉnh thứ hai thành một hoặc nhiều luồng bit.

Theo phương án thứ tám, sáng chế đề xuất thiết bị bao gồm bộ mã hóa video được tạo cấu hình để mã hóa luồng bit chứa chuỗi hình ảnh, bộ mã hóa video này bao gồm:

phương tiện để thu khối mẫu để dự đoán chuyển động;

phương tiện để lựa chọn vector chuyển động tịnh tiến cơ sở có độ chính xác thứ nhất cho khối mẫu;

phương tiện để xác định chế độ dự đoán cho ít nhất một thành phần vector chuyển động bậc cao;

phương tiện để xác định hệ số tinh chỉnh cho thành phần vector chuyển động bậc cao bằng cách sử dụng độ chính xác thứ hai khác với độ chính xác thứ nhất; và

phương tiện để mã hóa thông tin về thành phần vector chuyển động tịnh tiến cơ sở, và hệ số tinh chỉnh thành một hoặc nhiều luồng bit.

Theo phương án ví dụ thứ chín, sáng chế đề xuất phương pháp bao gồm các bước:

thu khối mẫu để dự đoán chuyển động;

thu thông tin của vectơ chuyển động tịnh tiến cơ sở có độ chính xác thứ nhất được lựa chọn cho khối mẫu;

thu thông tin về chế độ dự đoán cho ít nhất một thành phần vectơ chuyển động bậc cao;

thu thông tin về hệ số tinh chỉnh cho thành phần vectơ chuyển động bậc cao có độ chính xác thứ hai khác với độ chính xác thứ nhất; và

tính toán vectơ chuyển động bằng cách sử dụng thông tin thu được.

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

thu thông tin về hệ số tinh chỉnh cho thành phần vectơ chuyển động tịnh tiến cơ sở có độ chính xác thứ nhất.

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

tính toán vectơ chuyển động dựa trên thành phần vectơ chuyển động cơ sở, hệ số tinh chỉnh cho thành phần vectơ chuyển động cơ sở, thành phần vectơ chuyển động bậc cao, và hệ số tinh chỉnh cho thành phần vectơ chuyển động bậc cao.

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

sử dụng một trong số mô hình chuyển động afin hạn chế và mô hình chuyển động afin đầy đủ có khả năng xử lý xoay và thu phóng để xác định chế độ dự đoán cho thành phần vectơ chuyển động bậc cao.

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

thu trường vectơ chuyển động dưới dạng các vectơ chuyển động góc của đơn vị mã hóa.

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

thu ít nhất một phần thông tin trong tập tham số hoặc trong phần đầu ảnh, trong phần đầu lát hoặc trong phần đầu ô.

Theo phương án ví dụ thứ mười, sáng chế đề xuất thiết bị bao gồm ít nhất một bộ xử lý; và ít nhất một bộ nhớ bao gồm mã chương trình máy tính, trong đó ít nhất một bộ nhớ và mã chương trình máy tính này được tạo cấu hình để, với ít nhất một bộ xử lý, giúp thiết kế thực hiện ít nhất các bước:

thu khối mẫu để dự đoán chuyển động;

thu thông tin của vector chuyển động tịnh tiến cơ sở có độ chính xác thứ nhất được lựa chọn cho khối mẫu;

thu thông tin về chế độ dự đoán cho ít nhất một thành phần vector chuyển động bậc cao;

thu thông tin về hệ số tinh chỉnh cho thành phần vector chuyển động bậc cao có độ chính xác thứ hai khác với độ chính xác thứ nhất; và

tính toán vector chuyển động bằng cách sử dụng thông tin thu được.

Theo một số phương án, thiết bị còn được giúp để thực hiện bước:

thu thông tin về hệ số tinh chỉnh cho thành phần vector chuyển động tịnh tiến cơ sở có độ chính xác thứ nhất.

Theo một số phương án, thiết bị còn được giúp để thực hiện bước:

tính toán vectơ chuyển động dựa trên thành phần vectơ chuyển động cơ sở, hệ số tinh chỉnh cho thành phần vectơ chuyển động cơ sở, thành phần vectơ chuyển động bậc cao, và hệ số tinh chỉnh cho thành phần vectơ chuyển động bậc cao.

Theo một số phương án, thiết bị còn được giúp để thực hiện bước:

sử dụng một trong số mô hình chuyển động afin hạn chế và mô hình chuyển động afin đầy đủ có khả năng xử lý xoay và thu phóng để xác định chế độ dự đoán cho thành phần vectơ chuyển động bậc cao.

Theo một số phương án, thiết bị còn được giúp để thực hiện bước:

thu trường vectơ chuyển động dưới dạng các vectơ chuyển động góc của đơn vị mã hóa.

Theo một số phương án, thiết bị còn được giúp để thực hiện bước:

thu ít nhất một phần thông tin trong tập tham số hoặc trong phần đầu ảnh, trong phần đầu lát hoặc trong phần đầu ô.

Theo phương án ví dụ thứ mười một, sáng chế đề xuất vật ghi đọc được bằng máy tính có lưu trữ mã trên đó để sử dụng bằng thiết bị, mà khi được thực thi bởi bộ xử lý, giúp thiết bị thực hiện các bước:

thu khối mẫu để dự đoán chuyển động;

thu thông tin của vectơ chuyển động tịnh tiến cơ sở có độ chính xác thứ nhất được lựa chọn cho khối mẫu;

thu thông tin về chế độ dự đoán cho ít nhất một thành phần vectơ chuyển động bậc cao;

thu thông tin về hệ số tinh chỉnh cho thành phần vectơ chuyển động bậc cao có độ chính xác thứ hai khác với độ chính xác thứ nhất; và

tính toán vectơ chuyển động bằng cách sử dụng thông tin thu được.

Theo phương án thứ mười hai, sáng chế đề xuất thiết bị bao gồm bộ giải mã video được tạo cấu hình để giải mã luồng bit chứa chuỗi hình ảnh, bộ mã hóa video này bao gồm:

phương tiện để thu khối mẫu để dự đoán chuyển động;

phương tiện để thu thông tin của vector chuyển động tịnh tiến cơ sở có độ chính xác thứ nhất được lựa chọn cho khối mẫu;

phương tiện để thu thông tin về chế độ dự đoán cho ít nhất một thành phần vector chuyển động bậc cao;

phương tiện để thu thông tin về hệ số tinh chỉnh cho thành phần vector chuyển động bậc cao có độ chính xác thứ hai khác với độ chính xác thứ nhất; và

phương tiện để tính toán vector chuyển động bằng cách sử dụng thông tin thu được.

YÊU CẦU BẢO HỘ

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

thu khối mẫu để dự đoán chuyển động;

tính toán ít nhất một thành phần vector chuyển động tịnh tiến cơ sở cho khối mẫu bằng cách xác định chế độ dự đoán cho thành phần vector chuyển động tịnh tiến cơ sở này;

xác định thành phần vector chuyển động vi sai thứ nhất bằng cách sử dụng độ chính xác thứ nhất;

bổ sung thành phần vector chuyển động vi sai thứ nhất này vào chế độ dự đoán cho thành phần vector chuyển động tịnh tiến cơ sở; và

tính toán ít nhất một thành phần vector chuyển động bậc cao bằng cách:

xác định chế độ dự đoán cho thành phần vector chuyển động bậc cao;

xác định thành phần vector chuyển động vi sai thứ hai bằng cách sử dụng độ chính xác thứ hai khác với độ chính xác thứ nhất;

bổ sung thành phần vector chuyển động vi sai thứ hai này vào chế độ dự đoán cho thành phần vector chuyển động bậc cao; và

thực hiện bước bù chuyển động cho khối mẫu bằng cách sử dụng ít nhất một thành phần vector chuyển động tịnh tiến cơ sở và ít nhất một thành phần vector chuyển động vi sai.

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

mã hóa thành hoặc thu từ một hoặc nhiều luồng bit thông tin về hệ số tinh chỉnh cho thành phần vector chuyển động tịnh tiến cơ sở.

3. Phương pháp theo điểm 1, trong đó phương pháp này còn bao gồm bước:

tính toán vector chuyển động dựa trên ít nhất một thành phần vector chuyển động tịnh tiến cơ sở, thành phần vector chuyển động vi sai thứ nhất, thành phần vector chuyển động bậc cao, hoặc thành phần vector chuyển động vi sai thứ hai.

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

nhân thành phần vector chuyển động vi sai thứ nhất với thừa số thứ nhất để thu thành phần vector chuyển động vi sai thứ nhất đã mở rộng, nếu độ chính xác của thành phần vector chuyển động tịnh tiến cơ sở và thành phần vector chuyển động vi sai thứ nhất không giống nhau; và

sử dụng thành phần vector chuyển động vi sai thứ nhất đã mở rộng trong bước tính toán thành phần vector chuyển động tịnh tiến cơ sở.

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

nhân thành phần vector chuyển động vi sai thứ hai với thừa số thứ hai để thu thành phần vector chuyển động vi sai thứ hai đã mở rộng, nếu độ chính xác của thành phần vector chuyển động bậc cao và thành phần vector chuyển động vi sai thứ hai không giống nhau; và

sử dụng thành phần vector chuyển động vi sai thứ hai đã mở rộng này trong bước tính toán thành phần vector chuyển động bậc cao.

6. Phương pháp theo điểm 1, trong đó phương pháp này còn bao gồm bước:

mã hóa thành hoặc thu từ một hoặc nhiều luồng bit, thông tin của ít nhất một trong số độ chính xác thứ nhất hoặc độ chính xác thứ hai.

7. Phương pháp theo điểm 1, trong đó phương pháp này còn bao gồm ít nhất một trong số các bước:

sử dụng mô hình chuyển động afin hạn chế có khả năng xử lý xoay và thu phóng để xác định chế độ dự đoán cho thành phần vector chuyển động bậc cao; hoặc

sử dụng mô hình chuyển động afin đầy đủ có khả năng xử lý xoay và thu phóng để xác định chế độ dự đoán cho thành phần vector chuyển động bậc cao.

8. Phương pháp theo điểm 1, trong đó phương pháp này còn bao gồm bước:

báo hiệu hoặc thu trường vector chuyển động bằng cách sử dụng các vector chuyển động góc của đơn vị mã hóa.

9. Phương pháp theo điểm 1, trong đó phương pháp này còn bao gồm bước:

báo hiệu hoặc thu ít nhất một phần thông tin trong tập tham số, trong phần đầu ảnh, trong phần đầu lát hoặc trong phần đầu ô.

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

thu khối mẫu để dự đoán chuyển động;

tính toán ít nhất một thành phần vector chuyển động tịnh tiến cơ sở cho khối mẫu bằng cách xác định chế độ dự đoán cho thành phần vector chuyển động tịnh tiến cơ sở này;

xác định thành phần vector chuyển động vi sai thứ nhất bằng cách sử dụng độ chính xác thứ nhất;

bổ sung thành phần vector chuyển động vi sai thứ nhất này vào chế độ dự đoán cho thành phần vector chuyển động tịnh tiến cơ sở; và

tính toán ít nhất một thành phần vector chuyển động bậc cao bằng cách

xác định chế độ dự đoán cho thành phần vector chuyển động bậc cao này;

xác định thành phần vector chuyển động vi sai thứ hai bằng cách sử dụng độ chính xác thứ hai khác với độ chính xác thứ nhất;

bổ sung thành phần vector chuyển động vi sai thứ hai này vào chế độ dự đoán cho thành phần vector chuyển động bậc cao; và

thực hiện bước bù chuyển động cho khối mẫu bằng cách sử dụng ít nhất một thành phần vector chuyển động tịnh tiến cơ sở và ít nhất một thành phần vector chuyển động vi sai.

11. Thiết bị theo điểm 10, trong đó thiết bị này còn được giúp để thực hiện bước:

mã hóa thành hoặc thu từ một hoặc nhiều luồng bit, thông tin về hệ số tinh chỉnh cho thành phần vector chuyển động tịnh tiến cơ sở.

12. Thiết bị theo điểm 10, trong đó thiết bị này còn được giúp để thực hiện bước:

tính toán vectơ chuyển động dựa trên ít nhất một thành phần vectơ chuyển động tịnh tiến cơ sở, thành phần vectơ chuyển động vi sai thứ nhất, thành phần vectơ chuyển động bậc cao, hoặc thành phần vectơ chuyển động vi sai thứ hai.

13. Thiết bị theo điểm 10, trong đó thiết bị này còn được giúp để thực hiện các bước:

nhân thành phần vectơ chuyển động vi sai thứ nhất với thừa số thứ nhất để thu thành phần vectơ chuyển động vi sai thứ nhất đã mở rộng, nếu độ chính xác của thành phần vectơ chuyển động tịnh tiến cơ sở và thành phần vectơ chuyển động vi sai thứ nhất không giống nhau; và

sử dụng thành phần vectơ chuyển động vi sai thứ nhất đã mở rộng trong bước tính toán thành phần vectơ chuyển động tịnh tiến cơ sở.

14. Thiết bị theo điểm 10, trong đó thiết bị này còn được giúp để thực hiện các bước:

nhân thành phần vectơ chuyển động vi sai thứ hai với thừa số thứ hai để thu thành phần vectơ chuyển động vi sai thứ hai đã mở rộng, nếu độ chính xác của thành phần vectơ chuyển động bậc cao và thành phần vectơ chuyển động vi sai thứ hai không giống nhau; và

sử dụng thành phần vectơ chuyển động vi sai thứ hai đã mở rộng này trong bước tính toán thành phần vectơ chuyển động bậc cao.

15. Thiết bị theo điểm 10, trong đó thiết bị này còn được giúp để thực hiện bước:

mã hóa thành hoặc thu từ một hoặc nhiều luồng bit, thông tin của ít nhất một trong số độ chính xác thứ nhất hoặc độ chính xác thứ hai.

16. Thiết bị theo điểm 10, trong đó thiết bị này còn được giúp để thực hiện một trong số các bước sau:

sử dụng mô hình chuyển động afin hạn chế có khả năng xử lý xoay và thu phóng để xác định chế độ dự đoán cho thành phần vectơ chuyển động bậc cao; hoặc

sử dụng mô hình chuyển động afin đầy đủ có khả năng xử lý xoay và thu phóng để xác định chế độ dự đoán cho thành phần vectơ chuyển động bậc cao.

17. Thiết bị theo điểm 10, trong đó thiết bị này còn được giúp để thực hiện bước:

báo hiệu hoặc thu trường vector chuyển động bằng cách sử dụng các vector chuyển động góc của đơn vị mã hóa.

18. Thiết bị theo điểm 10, trong đó thiết bị này còn được giúp để thực hiện bước:

báo hiệu hoặc thu ít nhất một phần thông tin trong tập tham số hoặc trong phần đầu ảnh, trong phần đầu lát hoặc trong phần đầu ô.

19. Vật ghi đọc được bằng máy tính, trong đó vật ghi này có lưu trữ mã trên đó để sử dụng bằng thiết bị, mà khi được thực thi bởi bộ xử lý, giúp thiết bị thực hiện các bước:

thu khối mẫu để dự đoán chuyển động;

tính toán ít nhất một thành phần vector chuyển động tịnh tiến cơ sở cho khối mẫu bằng cách xác định chế độ dự đoán cho thành phần vector chuyển động tịnh tiến cơ sở;

xác định thành phần vector chuyển động vi sai thứ nhất bằng cách sử dụng độ chính xác thứ nhất;

bổ sung thành phần vector chuyển động vi sai thứ nhất này vào chế độ dự đoán cho thành phần vector chuyển động tịnh tiến cơ sở; và

tính toán ít nhất một thành phần vector chuyển động bậc cao bằng cách

xác định chế độ dự đoán cho thành phần vector chuyển động bậc cao này;

xác định thành phần vector chuyển động vi sai thứ hai bằng cách sử dụng độ chính xác thứ hai khác với độ chính xác thứ nhất;

bổ sung thành phần vector chuyển động vi sai thứ hai này vào chế độ dự đoán cho thành phần vector chuyển động bậc cao; và

thực hiện bước bù chuyển động cho khối mẫu bằng cách sử dụng ít nhất một thành phần vector chuyển động tịnh tiến cơ sở và ít nhất một thành phần vector chuyển động vi sai.

20. Vật ghi đọc được bằng máy tính theo điểm 19, trong đó vật ghi này còn giúp thiết bị thực hiện bước:

mã hóa thành hoặc thu từ một hoặc nhiều luồng bit, thông tin về hệ số tinh chỉnh cho thành phần vectơ chuyển động tịnh tiến cơ sở.

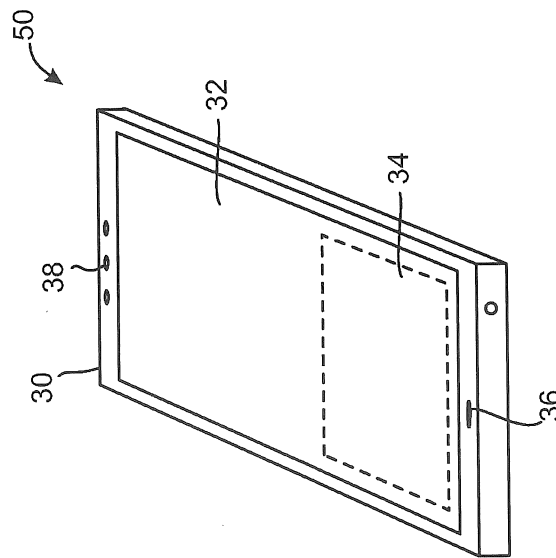


Fig. 2

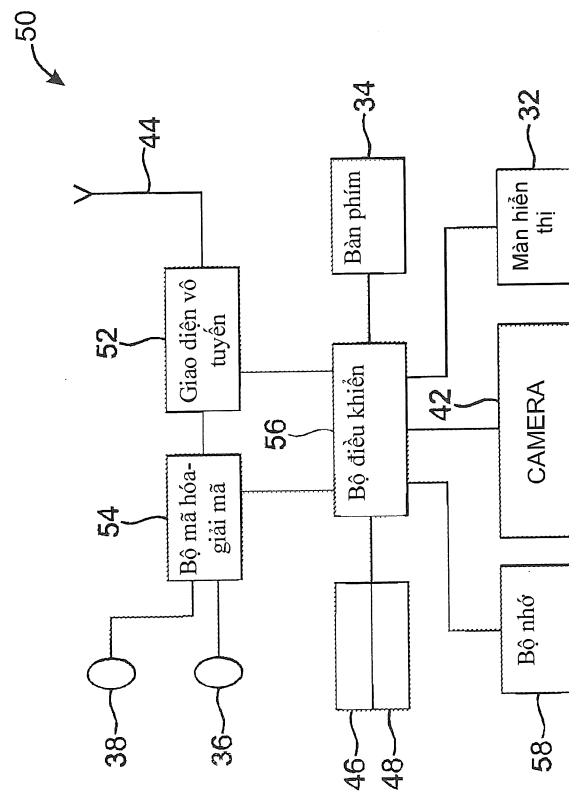


Fig. 1

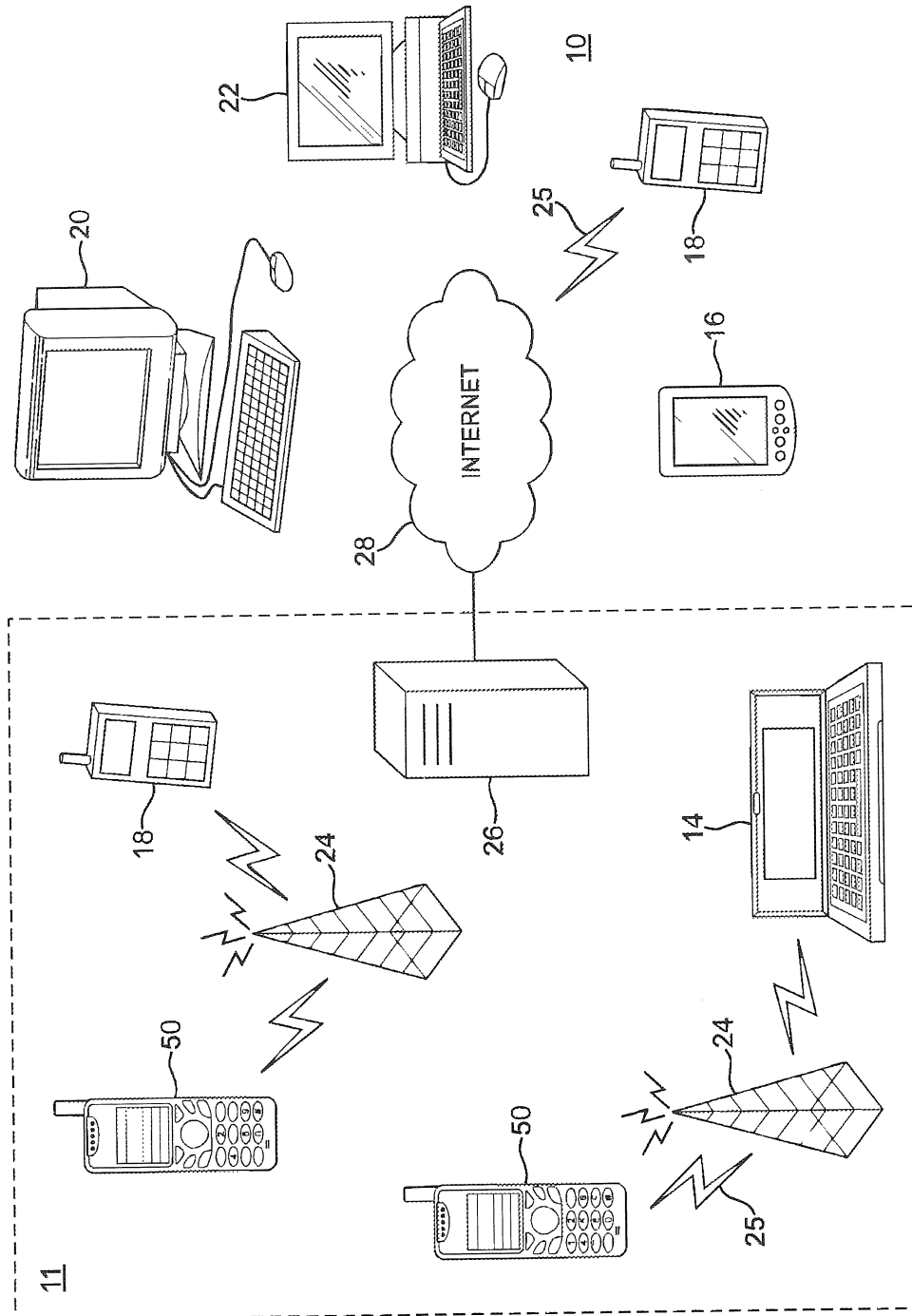


Fig. 3

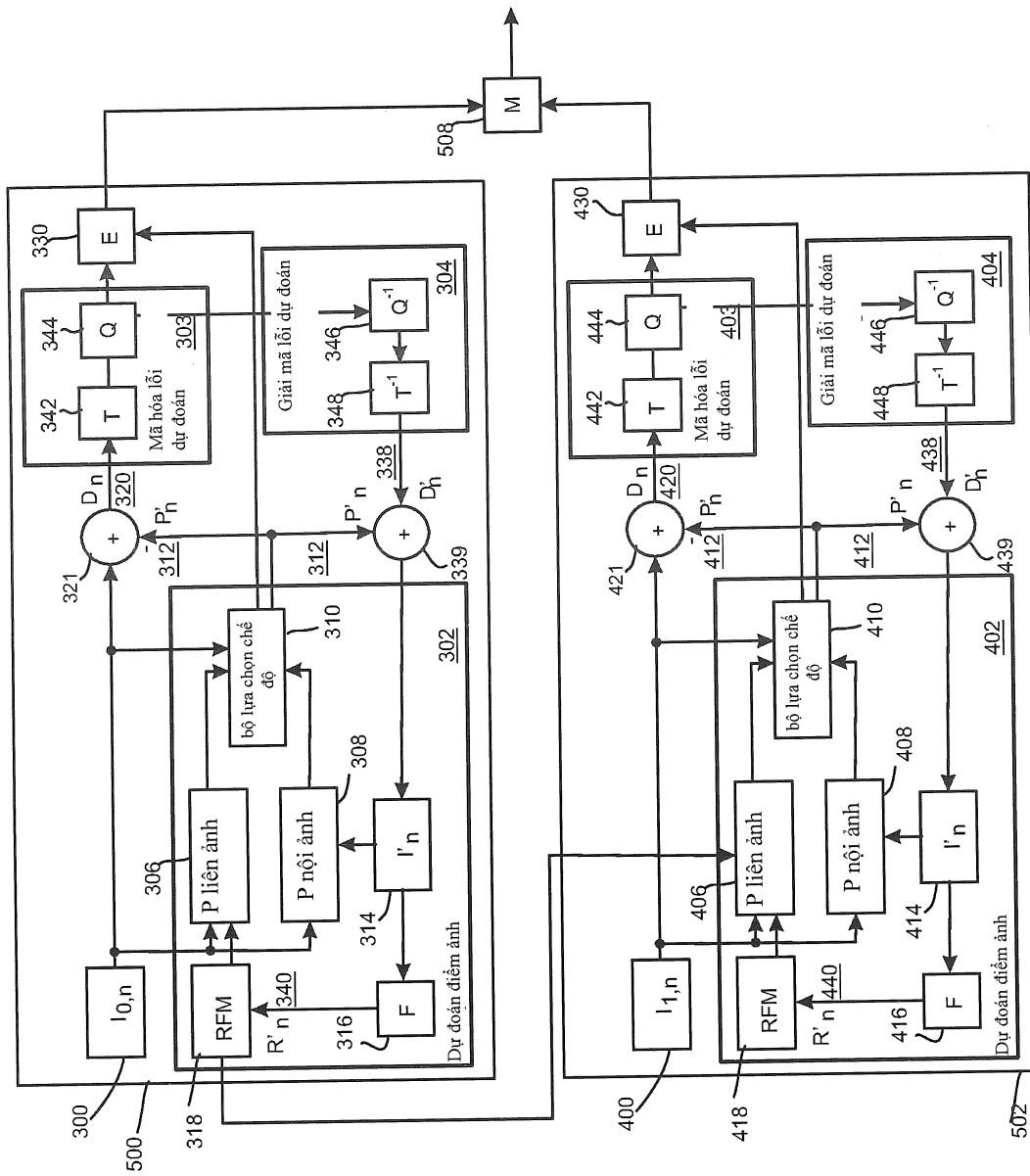


Fig. 4

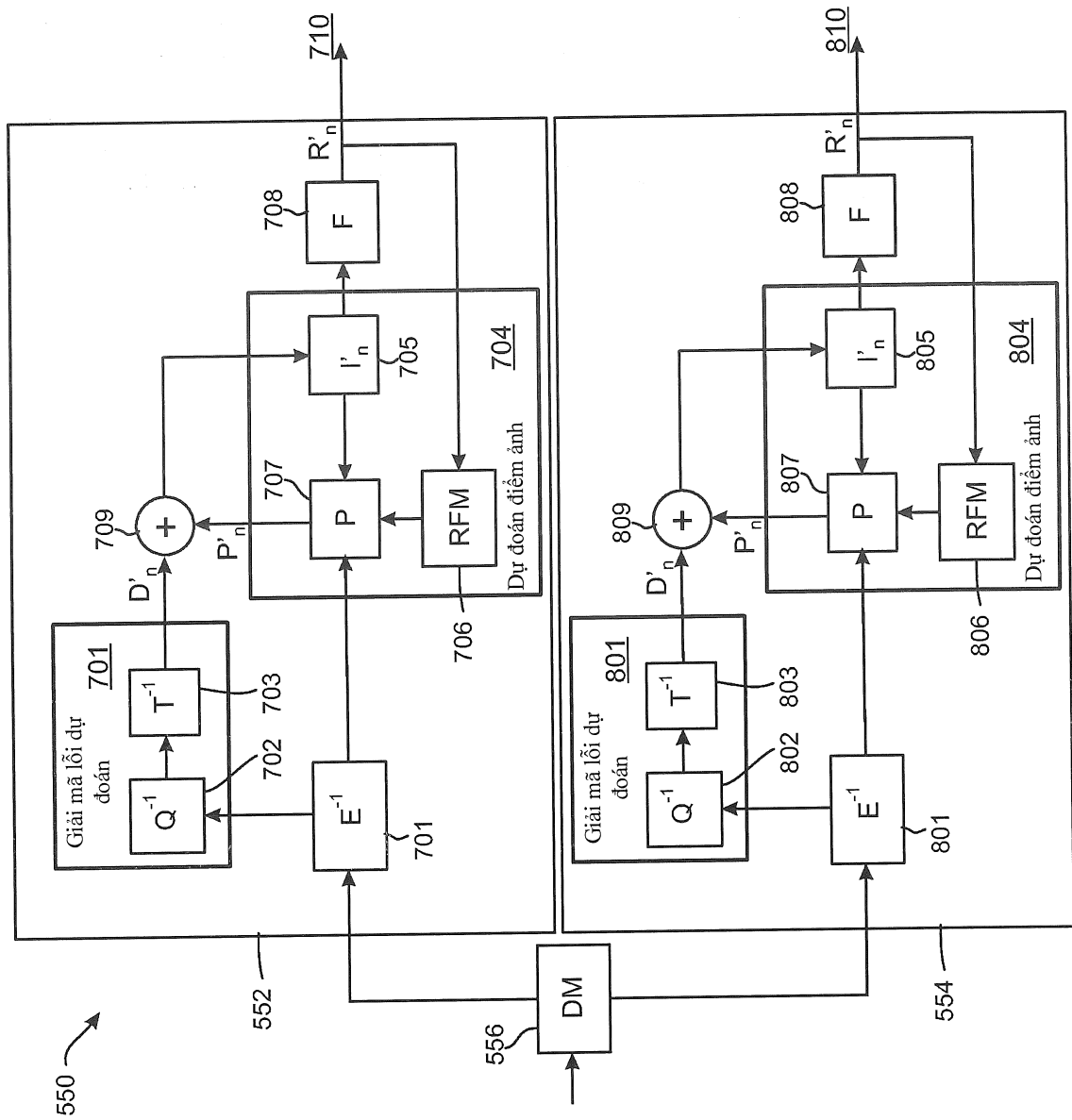


Fig. 5

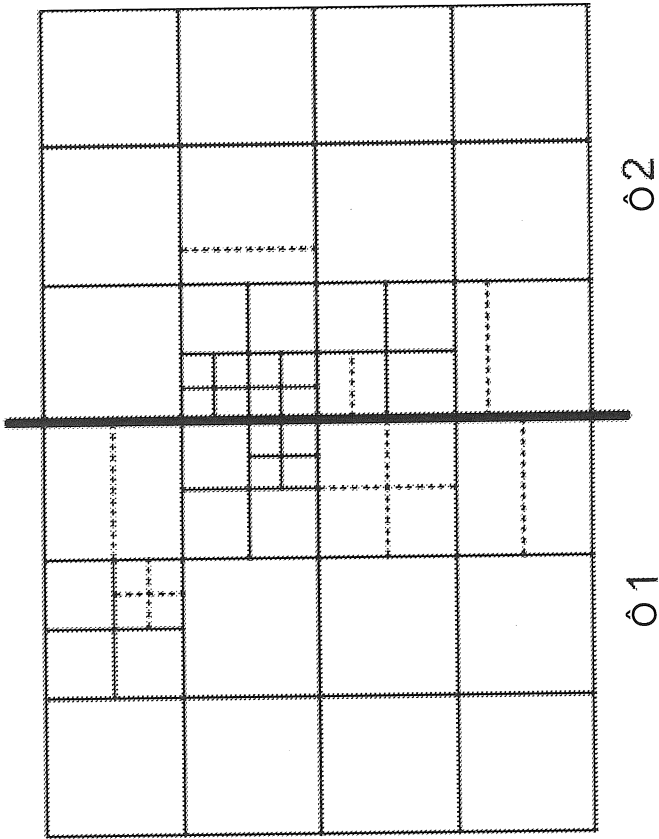


Fig. 6

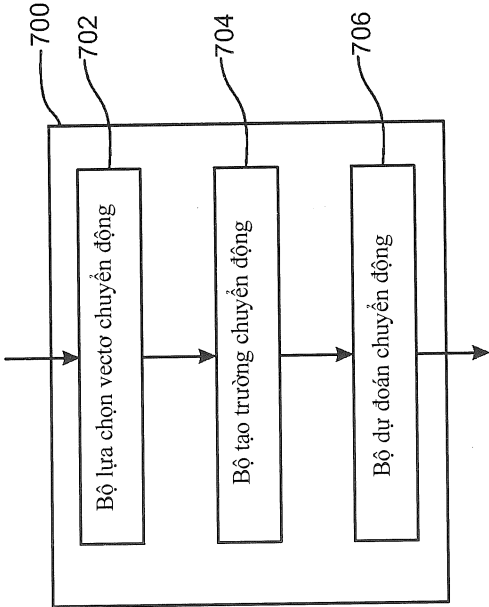


Fig. 7

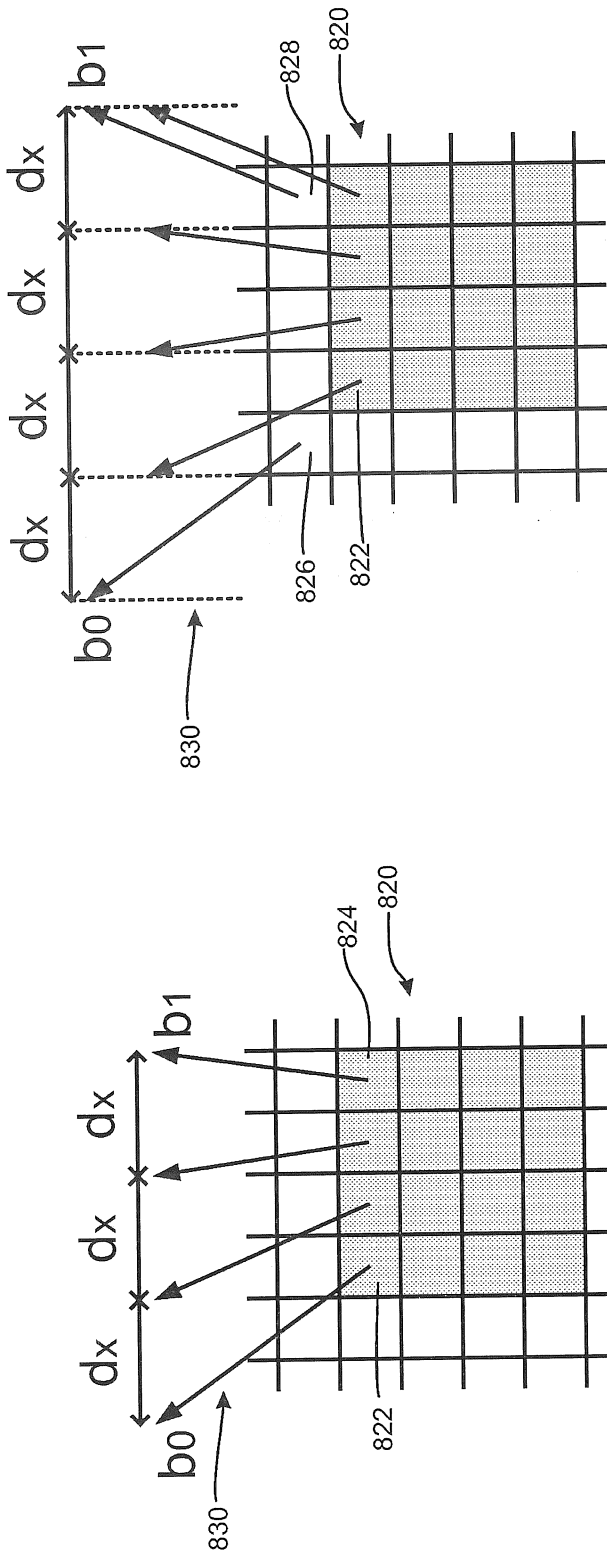


Fig. 8a

Fig. 8b

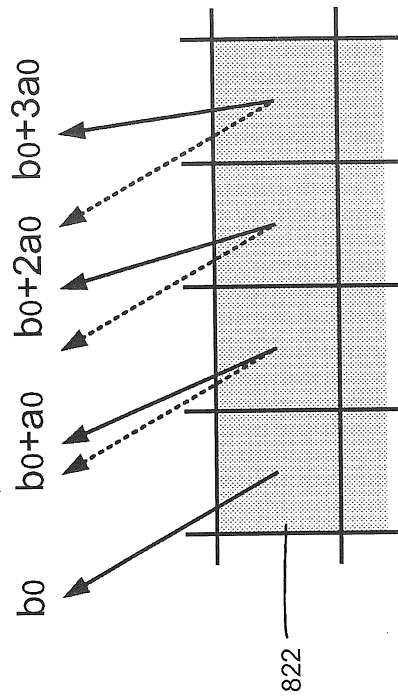


Fig. 8c

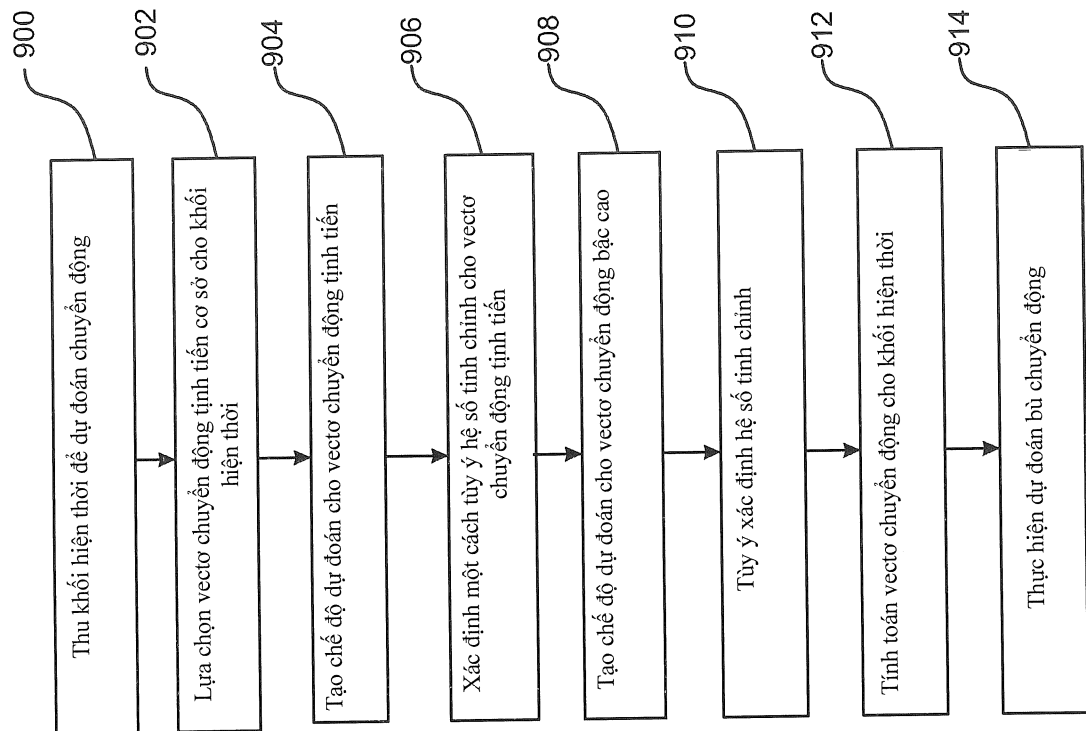


Fig. 9

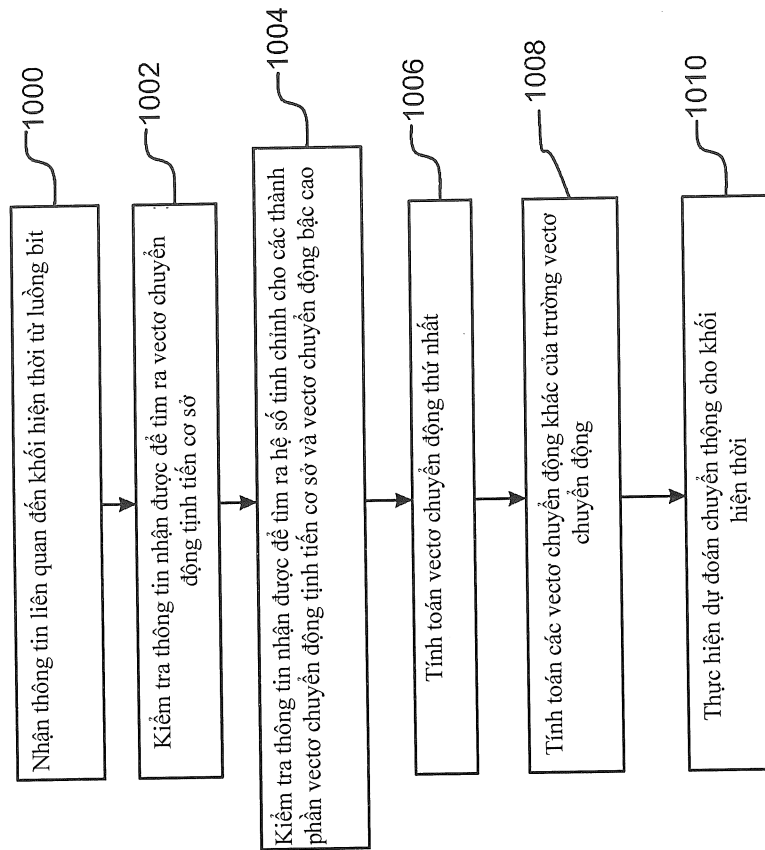


Fig. 10

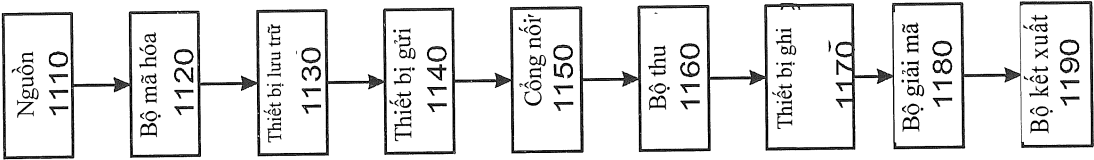


Fig. 11