



- (12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ
(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN) (11) 
CỤC SỞ HỮU TRÍ TUỆ
1-0047627
(51)^{2020.01} H04N 19/46; H04N 19/169; H04N 19/176; H04N 19/159; H04N 19/172 (13) B
-

- (21) 1-2021-07782 (22) 01/05/2020
(86) PCT/US2020/030951 01/05/2020 (87) WO2020/227064 12/11/2020
(30) 62/843,991 06/05/2019 US
(45) 25/06/2025 447 (43) 25/02/2022 407A
(73) HUAWEI TECHNOLOGIES CO., LTD. (CN)
Huawei Administration Building, Bantian, Longgang District, Shenzhen, Guangdong
518129, China
(72) WANG, Ye-Kui (US).
(74) Công ty Luật TNHH T&G (TGVN)
-

- (54) PHƯƠNG PHÁP GIẢI MÃ, PHƯƠNG PHÁP MÃ HÓA, THIẾT BỊ GIẢI MÃ, VÀ
THIẾT BỊ MÃ HÓA

(21) 1-2021-07782

(57) Sáng chế đề cập đến phương pháp giải mã, phương pháp mã hóa, thiết bị giải mã, thiết bị mã hóa, máy tạo mã, hệ thống tạo mã, phương tiện tạo mã, phương tiện lưu trữ và bộ tạo mã. Phương pháp giải mã dòng bit video được tạo mã bao gồm bước nhận dòng bit video được tạo mã, trong đó dòng bit video được tạo mã chứa hình ảnh làm mới giải mã từng bước (GDR) và cờ thứ nhất có giá trị thứ nhất; thiết đặt giá trị thứ hai của cờ thứ hai bằng giá trị thứ nhất của cờ thứ nhất; làm trống các hình ảnh được giải mã trước đó bất kỳ từ bộ đệm hình ảnh được giải mã (DPB) dựa trên cờ thứ hai có giá trị thứ hai; và giải mã hình ảnh hiện hành sau khi DPB đã được làm trống.

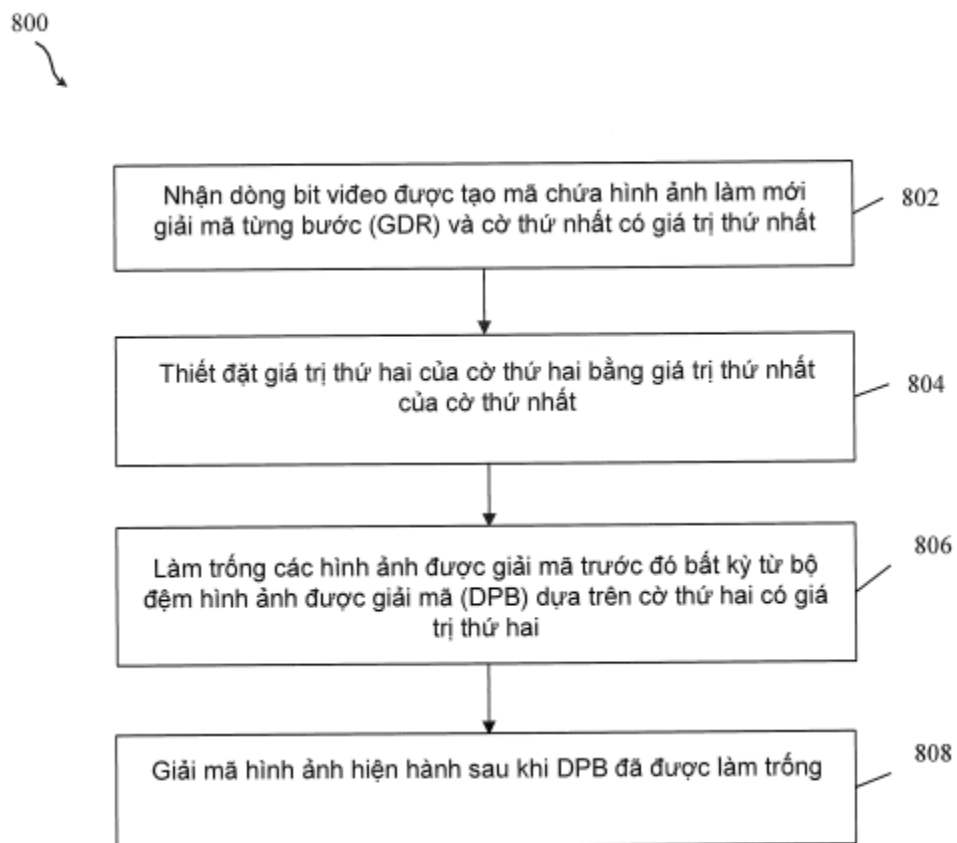


FIG. 8

Lĩnh vực kỹ thuật được đề cập

Nói chung, sáng chế đề cập đến các kỹ thuật hỗ trợ việc kết xuất các hình ảnh được giải mã trước đó khi tạo mã video. Cụ thể hơn, sáng chế cho phép các hình ảnh được giải mã trước đó tương ứng với hình ảnh điểm truy cập ngẫu nhiên bắt đầu chuỗi video được tạo mã (coded video sequence, CVS) được kết xuất từ bộ đệm hình ảnh được giải mã (decoded picture buffer, DPB).

Tình trạng kỹ thuật của sáng chế

Lượng dữ liệu video cần để mô tả thậm chí video tương đối ngắn có thể là quan trọng, mà có thể dẫn đến những khó khăn khi dữ liệu cần được tạo luồng hoặc nếu không thì được truyền thông qua mạng truyền thông với dung lượng dải thông hạn chế. Vì vậy, dữ liệu video thường được nén trước khi được truyền thông qua các mạng viễn thông đương thời. Kích thước của video cũng có thể là vấn đề khi video được lưu trữ trên thiết bị lưu trữ do các tài nguyên bộ nhớ có thể bị giới hạn. Các thiết bị nén video thường sử dụng phần mềm và/hoặc phần cứng ở nguồn để tạo mã dữ liệu video trước khi truyền hoặc lưu trữ, nhờ đó làm giảm lượng dữ liệu cần để đại diện cho các ảnh video số. Dữ liệu đã nén sau đó được nhận tại điểm đến bởi thiết bị giải nén video để giải mã dữ liệu video. Với các tài nguyên mạng hạn chế và các nhu cầu luôn tăng lên về chất lượng video cao hơn, các kỹ thuật nén và giải nén cải tiến để nâng cao tỷ lệ nén với việc đánh đổi một ít cho đến không phải đánh đổi về chất lượng ảnh luôn được mong đợi.

Bản chất kỹ thuật của sáng chế

Khía cạnh thứ nhất đề cập đến phương pháp giải mã được thực hiện bởi bộ giải mã video. Phương pháp này bao gồm bước nhận, bởi bộ giải mã video, dòng bit video được tạo mã, trong đó dòng bit video được tạo mã chứa hình ảnh làm mới giải mã từng bước (gradual decoding refresh, GDR) và cờ thứ nhất có giá trị thứ nhất; thiết đặt, bởi bộ giải mã video, giá trị thứ hai của cờ thứ hai bằng giá trị thứ nhất của cờ thứ nhất; làm

trống, bởi bộ giải mã video, các hình ảnh được giải mã trước đó bất kỳ từ bộ đệm hình ảnh được giải mã (DPB) dựa trên cờ thứ hai có giá trị thứ hai sau khi hình ảnh GDR đã được giải mã; và giải mã, bởi bộ giải mã video, hình ảnh hiện hành sau khi DPB đã được làm trống.

Phương pháp này cung cấp các kỹ thuật để kết xuất các hình ảnh trước (ví dụ, các hình ảnh được giải mã trước đó) trong bộ đệm hình ảnh được giải mã (DPB) khi hình ảnh điểm truy cập ngẫu nhiên (ví dụ, hình ảnh truy cập ngẫu nhiên sạch (clean random access, CRA), hình ảnh truy cập ngẫu nhiên từng bước (gradual random access, GRA), hoặc hình ảnh làm mới giải mã từng bước (GDR), hình ảnh CVSS, v.v.) hơn là là hình ảnh làm mới bộ giải mã tức thời (instantaneous decoder refresh, IDR) được gặp phải theo thứ tự giải mã. Việc làm trống các hình ảnh được giải mã trước đó từ DPB khi hình ảnh điểm truy cập ngẫu nhiên được đạt đến ngăn không cho DPB bị tràn và xúc tiến phát lại liên tục hơn. Vì vậy, bộ tạo mã / bộ giải mã (cũng được gọi là “codec”) trong tạo mã video được nâng cao so với các codec hiện hành. Như một vấn đề thực tế, quy trình tạo mã video được nâng cao đem lại cho người dùng trải nghiệm người dùng tốt hơn khi các video được gửi, được nhận, và/hoặc được xem.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh trên đây, một cách thực hiện khác của khía cạnh này đề xuất việc hình ảnh GDR không phải là hình ảnh thứ nhất của dòng bit video được tạo mã.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh trên đây, một cách thực hiện khác của khía cạnh này đề xuất việc hình ảnh GDR được bố trí trong đơn vị lớp trừu tượng mạng (network abstraction layer, NAL) của lớp tạo mã video (video coding layer, VCL) có loại đơn vị lớp trừu tượng mạng (NAL) làm mới giải mã từng bước (GDR) (GDR_NUT).

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh trên đây, một cách thực hiện khác của khía cạnh này đề xuất việc cờ thứ nhất được chỉ định là `no_output_of_prior_pics_flag` và cờ thứ hai được chỉ định là `NoOutputOfPriorPicsFlag`.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh trên đây, một cách thực hiện khác của khía cạnh này đề xuất việc thiết đặt thông số đầy DPB về không (zero) khi cờ thứ nhất được thiết đặt ở giá trị thứ nhất.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh trên đây, một cách thực hiện khác của khía cạnh này đề xuất việc DPB được làm trống sau khi hình ảnh GDR đã được giải mã.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh trên đây, một cách thực hiện khác của khía cạnh này đề xuất việc hiển thị hình ảnh được tạo ra dựa trên hình ảnh hiện hành.

Khía cạnh thứ hai đề cập đến phương pháp mã hóa được thực hiện bởi bộ mã hóa video. Phương pháp này bao gồm bước xác định, bởi bộ mã hóa video, điểm truy cập ngẫu nhiên cho chuỗi video; mã hóa, bởi bộ mã hóa video, hình ảnh làm mới giải mã từng bước (GDR) thành chuỗi video tại điểm truy cập ngẫu nhiên; thiết đặt, bởi bộ mã hóa video, cờ ở giá trị thứ nhất để ra lệnh cho bộ giải mã video làm trống các hình ảnh được giải mã trước đó bất kỳ từ bộ đệm hình ảnh được giải mã (DPB); tạo, bởi bộ mã hóa video, dòng bit video chứa chuỗi video có hình ảnh GDR tại điểm truy cập ngẫu nhiên và cờ; và lưu trữ, bởi bộ mã hóa video, dòng bit video để truyền đến bộ giải mã video.

Phương pháp này cung cấp các kỹ thuật để kết xuất các hình ảnh trước (ví dụ, các hình ảnh được giải mã trước đó) trong bộ đệm hình ảnh được giải mã (DPB) khi hình ảnh điểm truy cập ngẫu nhiên (ví dụ, hình ảnh truy cập ngẫu nhiên sạch (clean random access, CRA), hình ảnh truy cập ngẫu nhiên từng bước (gradual random access, GRA), hoặc hình ảnh làm mới giải mã từng bước (GDR), hình ảnh CVSS, v.v.) hơn là là hình ảnh làm mới bộ giải mã tức thời (instantaneous decoder refresh, IDR) được gặp phải theo thứ tự giải mã. Việc làm trống các hình ảnh được giải mã trước đó từ DPB khi hình ảnh điểm truy cập ngẫu nhiên được đạt đến ngăn không cho DPB bị tràn và xúc tiến phát lại liên tục hơn. Vì vậy, bộ tạo mã / bộ giải mã (cũng được gọi là “codec”) trong tạo mã video được nâng cao so với các codec hiện hành. Như một vấn đề thực tế, quy trình tạo mã video được nâng cao đem lại cho người dùng trải nghiệm người dùng tốt hơn khi các video được gửi, được nhận, và/hoặc được xem.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh trên đây, một cách thực hiện khác của khía cạnh này đề xuất việc hình ảnh GDR không phải là hình ảnh thứ nhất của dòng bit video, và trong đó bộ giải mã video được ra lệnh để làm trống DPB sau khi

hình ảnh GDR đã được giải mã.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh trên đây, một cách thực hiện khác của khía cạnh này đề xuất việc hình ảnh GDR được bố trí trong đơn vị lớp trừu tượng mạng (network abstraction layer, NAL) của lớp tạo mã video (video coding layer, VCL) có loại đơn vị lớp trừu tượng mạng (NAL) làm mới giải mã từng bước (GDR) (GDR_NUT).

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh trên đây, một cách thực hiện khác của khía cạnh này đề xuất việc ra lệnh cho bộ giải mã video thiết đặt thông số đầy DPB ở không (zero) khi cờ này được thiết đặt ở giá trị thứ nhất.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh trên đây, một cách thực hiện khác của khía cạnh này đề xuất việc cờ này được chỉ định là `no_output_of_prior_pics_flag`.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh trên đây, một cách thực hiện khác của khía cạnh này đề xuất việc giá trị thứ nhất của cờ này là một.

Khía cạnh thứ ba đề cập đến thiết bị giải mã. Thiết bị giải mã bao gồm bộ nhận được tạo cấu hình để nhận dòng bit video được tạo mã; bộ nhớ được ghép nối với bộ nhận, bộ nhớ lưu trữ các lệnh; và bộ xử lý được ghép nối với bộ nhớ, bộ xử lý được tạo cấu hình để thực hiện các lệnh để làm cho thiết bị giải mã: nhận dòng bit video được tạo mã, trong đó dòng bit video được tạo mã chứa hình ảnh làm mới giải mã từng bước (GDR) và cờ thứ nhất có giá trị thứ nhất; thiết đặt giá trị thứ hai của cờ thứ hai bằng giá trị thứ nhất của cờ thứ nhất; làm trống các hình ảnh được giải mã trước đó bất kỳ từ bộ đệm hình ảnh được giải mã (DPB) dựa trên cờ thứ hai có giá trị thứ hai; và giải mã hình ảnh hiện hành sau khi DPB đã được làm trống.

Thiết bị giải mã này cung cấp các kỹ thuật để kết xuất các hình ảnh trước (ví dụ, các hình ảnh được giải mã trước đó) trong bộ đệm hình ảnh được giải mã (DPB) khi hình ảnh điểm truy cập ngẫu nhiên (ví dụ, hình ảnh truy cập ngẫu nhiên sạch (clean random access, CRA), hình ảnh truy cập ngẫu nhiên từng bước (gradual random access, GRA), hoặc hình ảnh làm mới giải mã từng bước (GDR), hình ảnh CVSS, v.v.) hơn là là hình ảnh làm mới bộ giải mã tức thời (instantaneous decoder refresh, IDR) được gặp

phải theo thứ tự giải mã. Việc làm trống các hình ảnh được giải mã trước đó từ DPB khi hình ảnh điểm truy cập ngẫu nhiên được đặt đến ngăn không cho DPB bị tràn và xúc tiến phát lại liên tục hơn. Vì vậy, bộ tạo mã / bộ giải mã (cũng được gọi là “codec”) trong tạo mã video được nâng cao so với các codec hiện hành. Như một vấn đề thực tế, quy trình tạo mã video được nâng cao đem lại cho người dùng trải nghiệm người dùng tốt hơn khi các video được gửi, được nhận, và/hoặc được xem.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh trên đây, một cách thực hiện khác của khía cạnh này đề xuất việc hình ảnh GDR không phải là hình ảnh thứ nhất của dòng bit video được tạo mã.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh trên đây, một cách thực hiện khác của khía cạnh này đề xuất việc cờ thứ nhất được chỉ định là `no_output_of_prior_pics_flag`, và trong đó cờ thứ hai được chỉ định là `NoOutputOfPriorPicsFlag`.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh trên đây, một cách thực hiện khác của khía cạnh này đề xuất bộ hiển thị được tạo cấu hình để hiển thị ảnh như được tạo ra dựa trên hình ảnh hiện hành.

Khía cạnh thứ tư đề cập đến thiết bị mã hóa. Thiết bị mã hóa bao gồm bộ nhớ chứa các lệnh; bộ xử lý được ghép nối với bộ nhớ, bộ xử lý được tạo cấu hình để thực hiện các lệnh để làm cho thiết bị mã hóa: xác định điểm truy cập ngẫu nhiên cho chuỗi video; mã hóa hình ảnh làm mới giải mã từng bước (GDR) thành chuỗi video tại điểm truy cập ngẫu nhiên; thiết đặt cờ ở giá trị thứ nhất để ra lệnh cho bộ giải mã video làm trống các hình ảnh được giải mã trước đó bất kỳ từ bộ đệm hình ảnh được giải mã (DPB); và tạo dòng bit video chứa chuỗi video có hình ảnh GDR tại điểm truy cập ngẫu nhiên và cờ; và bộ truyền được ghép nối với bộ xử lý, bộ truyền được tạo cấu hình để truyền dòng bit video đến bộ giải mã video.

Thiết bị mã hóa này cung cấp các kỹ thuật để kết xuất các hình ảnh trước (ví dụ, các hình ảnh được giải mã trước đó) trong bộ đệm hình ảnh được giải mã (DPB) khi hình ảnh điểm truy cập ngẫu nhiên (ví dụ, hình ảnh truy cập ngẫu nhiên sạch (clean random access, CRA), hình ảnh truy cập ngẫu nhiên từng bước (gradual random access, GRA), hoặc hình ảnh làm mới giải mã từng bước (GDR), hình ảnh CVSS, v.v.) hơn là là

hình ảnh làm mới bộ giải mã tức thời (instantaneous decoder refresh, IDR) được gặp phải theo thứ tự giải mã. Việc làm trống các hình ảnh được giải mã trước đó từ DPB khi hình ảnh điểm truy cập ngẫu nhiên được đặt đến ngăn không cho DPB bị tràn và xúc tiến phát lại liên tục hơn. Vì vậy, bộ tạo mã / bộ giải mã (cũng được gọi là “codec”) trong tạo mã video được nâng cao so với các codec hiện hành. Như một vấn đề thực tế, quy trình tạo mã video được nâng cao đem lại cho người dùng trải nghiệm người dùng tốt hơn khi các video được gửi, được nhận, và/hoặc được xem.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh trên đây, một cách thực hiện khác của khía cạnh này đề xuất việc hình ảnh GDR không phải là hình ảnh thứ nhất của dòng bit video.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh trên đây, một cách thực hiện khác của khía cạnh này đề xuất việc cờ này được chỉ định là `no_output_of_prior_pics_flag`.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh trên đây, một cách thực hiện khác của khía cạnh này đề xuất việc bộ nhớ lưu trữ dòng bit trước khi bộ truyền truyền dòng bit đến bộ giải mã video.

Khía cạnh thứ năm đề cập đến máy tạo mã. Máy tạo mã bao gồm bộ nhận được tạo cấu hình để nhận hình ảnh để mã hóa hoặc để nhận dòng bit để giải mã; bộ truyền được ghép nối với bộ nhận, bộ truyền được tạo cấu hình để truyền dòng bit đến bộ giải mã hoặc để truyền ảnh đã giải mã đến bộ hiển thị; bộ nhớ được ghép nối với ít nhất một trong số bộ nhận hoặc bộ truyền, bộ nhớ được tạo cấu hình để lưu trữ các lệnh; và bộ xử lý được ghép nối với bộ nhớ, bộ xử lý được tạo cấu hình để thực hiện các lệnh được lưu trữ trong bộ nhớ để thực hiện phương pháp bất kỳ trong số các phương pháp được bộc lộ ở đây.

Máy tạo mã này cung cấp các kỹ thuật để kết xuất các hình ảnh trước (ví dụ, các hình ảnh được giải mã trước đó) trong bộ đệm hình ảnh được giải mã (DPB) khi hình ảnh điểm truy cập ngẫu nhiên (ví dụ, hình ảnh truy cập ngẫu nhiên sạch (clean random access, CRA), hình ảnh truy cập ngẫu nhiên từng bước (gradual random access, GRA), hoặc hình ảnh làm mới giải mã từng bước (GDR), hình ảnh CVSS, v.v.) hơn là là hình ảnh làm mới bộ giải mã tức thời (instantaneous decoder refresh, IDR) được gặp phải

theo thứ tự giải mã. Việc làm trống các hình ảnh được giải mã trước đó từ DPB khi hình ảnh điểm truy cập ngẫu nhiên được đạt đến ngăn không cho DPB bị tràn và xúc tiến phát lại liên tục hơn. Vì vậy, bộ tạo mã / bộ giải mã (cũng được gọi là “codec”) trong tạo mã video được nâng cao so với các codec hiện hành. Như một vấn đề thực tế, quy trình tạo mã video được nâng cao đem lại cho người dùng trải nghiệm người dùng tốt hơn khi các video được gửi, được nhận, và/hoặc được xem.

Tùy chọn là, theo khía cạnh bất kỳ trong số các khía cạnh trên đây, một cách thực hiện khác của khía cạnh này đề xuất bộ hiển thị được tạo cấu hình để hiển thị ảnh.

Khía cạnh thứ sáu đề cập đến hệ thống. Hệ thống này bao gồm bộ mã hóa; và bộ giải mã truyền thông với bộ mã hóa, trong đó bộ mã hóa hoặc bộ giải mã bao gồm thiết bị giải mã, thiết bị mã hóa, hoặc máy tạo mã được bộc lộ ở đây.

Hệ thống này cung cấp các kỹ thuật để kết xuất các hình ảnh trước (ví dụ, các hình ảnh được giải mã trước đó) trong bộ đệm hình ảnh được giải mã (DPB) khi hình ảnh điểm truy cập ngẫu nhiên (ví dụ, hình ảnh truy cập ngẫu nhiên sạch (clean random access, CRA), hình ảnh truy cập ngẫu nhiên từng bước (gradual random access, GRA), hoặc hình ảnh làm mới giải mã từng bước (GDR), hình ảnh CVSS, v.v.) hơn là là hình ảnh làm mới bộ giải mã tức thời (instantaneous decoder refresh, IDR) được gặp phải theo thứ tự giải mã. Việc làm trống các hình ảnh được giải mã trước đó từ DPB khi hình ảnh điểm truy cập ngẫu nhiên được đạt đến ngăn không cho DPB bị tràn và xúc tiến phát lại liên tục hơn. Vì vậy, bộ tạo mã / bộ giải mã (cũng được gọi là “codec”) trong tạo mã video được nâng cao so với các codec hiện hành. Như một vấn đề thực tế, quy trình tạo mã video được nâng cao đem lại cho người dùng trải nghiệm người dùng tốt hơn khi các video được gửi, được nhận, và/hoặc được xem.

Khía cạnh thứ bảy đề cập đến phương tiện tạo mã. Phương tiện tạo mã bao gồm phương tiện nhận được tạo cấu hình để nhận hình ảnh để mã hóa hoặc để nhận dòng bit để giải mã; phương tiện truyền được ghép nối với phương tiện nhận, phương tiện truyền được tạo cấu hình để truyền dòng bit đến phương tiện giải mã hoặc để truyền ảnh đã giải mã đến phương tiện hiển thị; phương tiện lưu trữ được ghép nối với ít nhất một trong số phương tiện nhận hoặc phương tiện truyền, phương tiện lưu trữ được tạo cấu hình để lưu trữ các lệnh; và phương tiện xử lý được ghép nối với phương tiện lưu trữ, phương tiện

xử lý được tạo cấu hình để thực hiện các lệnh được lưu trữ trong phương tiện lưu trữ để thực hiện phương pháp bất kỳ trong số các phương pháp được bộc lộ ở đây.

Phương tiện tạo mã này cung cấp các kỹ thuật để kết xuất các hình ảnh trước (ví dụ, các hình ảnh được giải mã trước đó) trong bộ đệm hình ảnh được giải mã (DPB) khi hình ảnh điểm truy cập ngẫu nhiên (ví dụ, hình ảnh truy cập ngẫu nhiên sạch (clean random access, CRA), hình ảnh truy cập ngẫu nhiên từng bước (gradual random access, GRA), hoặc hình ảnh làm mới giải mã từng bước (GDR), hình ảnh CVSS, v.v.) hơn là là hình ảnh làm mới bộ giải mã tức thời (instantaneous decoder refresh, IDR) được gặp phải theo thứ tự giải mã. Việc làm trống các hình ảnh được giải mã trước đó từ DPB khi hình ảnh điểm truy cập ngẫu nhiên được đạt đến ngăn không cho DPB bị tràn và xúc tiến phát lại liên tục hơn. Vì vậy, bộ tạo mã / bộ giải mã (cũng được gọi là “codec”) trong tạo mã video được nâng cao so với các codec hiện hành. Như một vấn đề thực tế, quy trình tạo mã video được nâng cao đem lại cho người dùng trải nghiệm người dùng tốt hơn khi các video được gửi, được nhận, và/hoặc được xem.

Nhằm mục đích làm rõ, phương án bất kỳ trong số các phương án trên đây có thể được kết hợp với một hoặc nhiều phương án bất kỳ trong số các phương án khác trên đây để tạo phương án mới nằm trong phạm vi của sáng chế.

Các dấu hiệu này và các dấu hiệu khác sẽ được hiểu rõ ràng hơn từ phần mô tả chi tiết dưới đây kết hợp với các hình vẽ kèm theo và yêu cầu bảo hộ.

Mô tả vắn tắt các hình vẽ

Để hiểu đầy đủ hơn về sáng chế, có thể tham khảo phần mô tả vắn tắt sau đây, kết hợp với các hình vẽ kèm theo và phần mô tả chi tiết, trong đó các số chỉ dẫn giống nhau đại diện cho các phần giống nhau.

Fig.1 là sơ đồ khối minh họa hệ thống tạo mã làm ví dụ mà có thể sử dụng các kỹ thuật GDR.

Fig.2 là sơ đồ khối minh họa bộ mã hóa video làm ví dụ mà có thể thực hiện các kỹ thuật GDR.

Fig.3 là sơ đồ khối minh họa một ví dụ về bộ giải mã video mà có thể thực hiện các kỹ thuật GDR.

Fig.4 là hình vẽ thể hiện mối quan hệ giữa hình ảnh IRAP cân xứng với các hình ảnh chủ đạo và các hình ảnh kế tiếp theo thứ tự giải mã và thứ tự trình diễn.

Fig.5 minh họa kỹ thuật làm mới giải mã từng bước.

Fig.6 là hình vẽ giản lược minh họa sự tìm kiếm chuyển động không mong muốn.

Fig.7 minh họa dòng bit video được tạo cấu hình để thực hiện kỹ thuật truy cập ngẫu nhiên sạch (CRA).

Fig.8 thể hiện một phương án của phương pháp giải mã dòng bit video được tạo mã.

Fig.9 thể hiện một phương án của phương pháp mã hóa dòng bit video được tạo mã.

Fig.10 là hình vẽ giản lược của thiết bị tạo mã video.

Fig.11 là hình vẽ giản lược của một phương án của phương tiện tạo mã.

Mô tả chi tiết sáng chế

Nên được hiểu từ đầu rằng mặc dù cách thực hiện minh họa của một hoặc nhiều phương án được đưa ra dưới đây, nhưng các hệ thống và/hoặc các phương pháp được bộc lộ có thể được thực hiện nhờ sử dụng số lượng kỹ thuật bất kỳ, liệu hiện đã được biết đến hay đang tồn tại. Sáng chế sẽ hoàn toàn không bị giới hạn ở các cách thực hiện minh họa, các hình vẽ, và các kỹ thuật được minh họa dưới đây, bao gồm các thiết kế và các cách thực hiện làm ví dụ được minh họa và được mô tả ở đây, mà có thể được cải biến nằm trong phạm vi của yêu cầu bảo hộ kèm theo cùng với phạm vi tương đương đầy đủ của chúng.

Fig.1 là sơ đồ khối minh họa hệ thống tạo mã 10 làm ví dụ mà có thể sử dụng các kỹ thuật tạo mã video như được mô tả ở đây. Như được thể hiện trên Fig.1, hệ thống tạo mã 10 bao gồm thiết bị nguồn 12 để cung cấp dữ liệu video được mã hóa để được giải mã vào thời điểm muộn hơn bởi thiết bị đích 14. Cụ thể, thiết bị nguồn 12 có thể cung cấp dữ liệu video cho thiết bị đích 14 qua phương tiện đọc được bằng máy tính 16. Thiết bị nguồn 12 và thiết bị đích 14 có thể bao gồm thiết bị bất kỳ trong phạm vi thiết bị rộng, bao gồm các máy tính để bàn, các máy tính xách tay loại notebook (ví dụ,

máy tính xách tay loại laptop), các máy tính bảng, các thiết bị giải mã tín hiệu và truyền dữ liệu (set-top box), các máy điện thoại như các điện thoại "thông minh", các bàn điều khiển (pad) "thông minh", các máy thu hình, các camera, các thiết bị hiển thị, các máy phát phương tiện số, các bàn giao tiếp trò chơi video, thiết bị tạo luồng video, hoặc tương tự. Trong một số trường hợp, thiết bị nguồn 12 và thiết bị đích 14 có thể được trang bị để truyền thông vô tuyến.

Thiết bị đích 14 có thể nhận dữ liệu video được mã hóa để được giải mã qua phương tiện đọc được bằng máy tính 16. Phương tiện đọc được bằng máy tính 16 có thể bao gồm loại phương tiện hoặc thiết bị bất kỳ có khả năng chuyển dữ liệu video được mã hóa từ thiết bị nguồn 12 đến thiết bị đích 14. Theo một ví dụ, phương tiện đọc được bằng máy tính 16 có thể bao gồm phương tiện truyền thông để cho phép thiết bị nguồn 12 truyền dữ liệu video được mã hóa một cách trực tiếp đến thiết bị đích 14 theo thời gian thực. Dữ liệu video được mã hóa có thể được điều biến theo tiêu chuẩn truyền thông, như giao thức truyền thông vô tuyến, và được truyền đến thiết bị đích 14. Phương tiện truyền thông có thể bao gồm phương tiện truyền thông vô tuyến hoặc hữu tuyến bất kỳ, như phổ tần số radiô (RF) hoặc một hoặc nhiều đường truyền vật lý. Phương tiện truyền thông có thể tạo thành một phần của mạng dựa trên gói, như mạng cục bộ, mạng diện rộng, hoặc mạng toàn cầu như Internet. Phương tiện truyền thông có thể bao gồm các bộ định tuyến, các bộ chuyển mạch, các trạm gốc, hoặc thiết bị khác bất kỳ mà có thể là hữu ích để hỗ trợ truyền thông từ thiết bị nguồn 12 đến thiết bị đích 14.

Trong một số ví dụ, dữ liệu được mã hóa có thể được kết xuất từ giao diện đầu ra 22 đến thiết bị lưu trữ. Tương tự, dữ liệu được mã hóa có thể được truy cập từ thiết bị lưu trữ bởi giao diện đầu vào. Thiết bị lưu trữ có thể bao gồm loại phương tiện bất kỳ trong số các loại phương tiện lưu trữ dữ liệu truy cập cục bộ hoặc phân tán như ổ cứng, các đĩa Blu-ray, các đĩa video số (digital video disk, DVD), các bộ nhớ chỉ đọc dạng đĩa compac (CD-ROM), bộ nhớ tác động nhanh (flash memory), bộ nhớ khả biến hoặc bất khả biến, hoặc phương tiện lưu trữ kỹ thuật số thích hợp khác bất kỳ để lưu trữ dữ liệu video được mã hóa. Trong ví dụ thêm nữa, thiết bị lưu trữ có thể tương ứng với máy chủ tệp hoặc thiết bị lưu trữ trung gian khác mà có thể lưu trữ video đã mã hóa được tạo ra bởi thiết bị nguồn 12. Thiết bị đích 14 có thể truy cập dữ liệu video được lưu trữ từ thiết bị lưu trữ qua tạo luồng hoặc tải xuống. Máy chủ tệp có thể là loại máy chủ bất kỳ có

khả năng lưu trữ dữ liệu video được mã hóa và truyền dữ liệu video được mã hóa đó đến thiết bị đích 14. Các máy chủ tệp làm ví dụ bao gồm máy chủ web (ví dụ, đối với vị trí web), máy chủ giao thức truyền tệp (file transfer protocol, FTP), các thiết bị lưu trữ gắn vào mạng (network attached storage, NAS), hoặc ổ đĩa cục bộ. Thiết bị đích 14 có thể truy cập dữ liệu video được mã hóa qua kết nối dữ liệu chuẩn bất kỳ, bao gồm kết nối Internet. Điều này có thể bao gồm kênh vô tuyến (ví dụ, kết nối Wi-Fi), kết nối hữu tuyến (ví dụ, đường thuê bao số (digital subscriber line, DSL), modem cáp, v.v.), hoặc sự kết hợp của cả hai loại mà là thích hợp để truy cập dữ liệu video được mã hóa được lưu trữ trên máy chủ tệp. Hoạt động truyền dữ liệu video được mã hóa từ thiết bị lưu trữ có thể là hoạt động truyền tạo luồng, hoạt động truyền tải xuống, hoặc sự kết hợp của chúng.

Các kỹ thuật của sáng chế không nhất thiết phải bị giới hạn ở các ứng dụng hoặc các thiết đặt vô tuyến. Các kỹ thuật này có thể được áp dụng vào tạo mã video với sự hỗ trợ của ứng dụng đa phương tiện bất kỳ trong số nhiều loại ứng dụng đa phương tiện, như phát rộng vô tuyến truyền hình qua không trung, truyền vô tuyến truyền hình cáp, truyền vô tuyến truyền hình vệ tinh, truyền video tạo luồng Internet, như tạo luồng thích ứng động qua HTTP (dynamic adaptive streaming over HTTP, DASH), video số mà được mã hóa trên phương tiện lưu trữ dữ liệu, giải mã video số được lưu trữ trên phương tiện lưu trữ dữ liệu, hoặc các ứng dụng khác. Trong một số ví dụ, hệ thống tạo mã 10 có thể được tạo cấu hình để hỗ trợ truyền video một-hướng hoặc hai-hướng để hỗ trợ các ứng dụng như tạo luồng video, phát lại video, phát rộng video, và/hoặc điện thoại video.

Trong ví dụ trên Fig.1, thiết bị nguồn 12 bao gồm nguồn video 18, bộ mã hóa video 20, và giao diện đầu ra 22. Thiết bị đích 14 bao gồm giao diện đầu vào 28, bộ giải mã video 30, và thiết bị hiển thị 32. Theo sáng chế, bộ mã hóa video 20 của thiết bị nguồn 12 và/hoặc bộ giải mã video 30 của thiết bị đích 14 có thể được tạo cấu hình để áp dụng các kỹ thuật để tạo mã video. Trong các ví dụ khác, thiết bị nguồn và thiết bị đích có thể bao gồm các thành phần hoặc các cách bố trí khác. Ví dụ, thiết bị nguồn 12 có thể nhận dữ liệu video từ nguồn video ngoài, như camera ngoài. Tương tự, thiết bị đích 14 có thể giao tiếp với thiết bị hiển thị ngoài, ngoài việc bao gồm thiết bị hiển thị được tích hợp.

Hệ thống tạo mã được minh họa 10 trên Fig.1 chỉ là một ví dụ. Các kỹ thuật để tạo mã video có thể được thực hiện bởi thiết bị mã hóa và/hoặc giải mã video số bất kỳ. Mặc dù các kỹ thuật của sáng chế thường được thực hiện bởi thiết bị tạo mã video, nhưng các kỹ thuật này cũng có thể được thực hiện bởi bộ mã hóa/bộ giải mã video, thường được gọi là “CODEC.” Hơn nữa, các kỹ thuật của sáng chế cũng có thể được thực hiện bởi bộ tiền xử lý video. Bộ mã hóa video và/hoặc bộ giải mã có thể là bộ phận xử lý đồ họa (GPU) hoặc thiết bị tương tự.

Thiết bị nguồn 12 và thiết bị đích 14 chỉ là các ví dụ của các thiết bị tạo mã như vậy trong đó thiết bị nguồn 12 tạo ra dữ liệu video được tạo mã để truyền đến thiết bị đích 14. Trong một số ví dụ, thiết bị nguồn 12 và thiết bị đích 14 có thể hoạt động theo cách gần như đối xứng sao cho mỗi thiết bị trong số thiết bị nguồn 12 và thiết bị đích 14 bao gồm các thành phần mã hóa và giải mã video. Vì vậy, hệ thống tạo mã 10 có thể hỗ trợ truyền video một-hướng hoặc hai-hướng giữa các thiết bị video 12, 14, ví dụ, để tạo luồng video, phát lại video, phát rộng video, hoặc điện thoại video.

Nguồn video 18 của thiết bị nguồn 12 có thể bao gồm thiết bị thu video, như camera video, kho trữ video chứa video thu được trước đó, và/hoặc giao diện cung cấp video để nhận video từ bên cung cấp nội dung video. Theo một cách khác nữa, nguồn video 18 có thể tạo dữ liệu dựa trên đồ họa máy tính làm video nguồn, hoặc sự kết hợp của video trước tiếp, video lưu trữ, và video được tạo ra bởi máy tính.

Trong một số trường hợp, khi nguồn video 18 là camera video, thiết bị nguồn 12 và thiết bị đích 14 có thể tạo thành các điện thoại camera hoặc các điện thoại video. Tuy nhiên, như đã nêu trên, các kỹ thuật được mô tả trong sáng chế có thể áp dụng được vào tạo mã video nói chung, và có thể được áp dụng vào các ứng dụng vô tuyến và/hoặc hữu tuyến. Trong từng trường hợp, video thu được, video thu được từ trước, hoặc video được tạo ra bởi máy tính có thể được mã hóa bởi bộ mã hóa video 20. Thông tin video đã mã hóa sau đó có thể được kết xuất bởi giao diện đầu ra 22 vào phương tiện đọc được bằng máy tính 16.

Phương tiện đọc được bằng máy tính 16 có thể bao gồm phương tiện nhất thời, như phát rộng vô tuyến hoặc truyền mạng hữu tuyến, hoặc phương tiện lưu trữ (tức là, phương tiện lưu trữ lâu dài), như đĩa cứng, ổ tác động nhanh, đĩa compac, đĩa video số,

đĩa Blu-ray, hoặc phương tiện đọc được bằng máy tính khác. Trong một số ví dụ, mạng máy chủ (không được thể hiện trên hình vẽ) có thể nhận dữ liệu video được mã hóa từ thiết bị nguồn 12 và cung cấp dữ liệu video được mã hóa cho thiết bị đích 14, ví dụ, qua hoạt động truyền mạng. Tương tự, thiết bị tính toán của nhà máy sản xuất phương tiện, như nhà máy dập khuôn đĩa, có thể nhận dữ liệu video được mã hóa từ thiết bị nguồn 12 và tạo ra đĩa chứa dữ liệu video được mã hóa. Do đó, phương tiện đọc được bằng máy tính 16 có thể được hiểu là bao gồm một hoặc nhiều phương tiện đọc được bằng máy tính ở các dạng khác nhau, trong các ví dụ khác nhau.

Giao diện đầu vào 28 của thiết bị đích 14 nhận thông tin từ phương tiện đọc được bằng máy tính 16. Thông tin của phương tiện đọc được bằng máy tính 16 có thể bao gồm thông tin cú pháp được xác định bởi bộ mã hóa video 20, mà cũng được sử dụng bởi bộ giải mã video 30, mà bao gồm các phần tử cú pháp mô tả các đặc tính và/hoặc các hoạt động xử lý các khối và các đơn vị được tạo mã khác, ví dụ, nhóm các hình ảnh (group of pictures, GOP). Thiết bị hiển thị 32 hiển thị dữ liệu video được giải mã cho người dùng, và có thể bao gồm thiết bị hiển thị bất kỳ trong số nhiều loại thiết bị hiển thị như ống tia catốt (CRT), bộ hiển thị tinh thể lỏng (LCD), bộ hiển thị plasma, bộ hiển thị diot phát quang hữu cơ (OLED), hoặc loại thiết bị hiển thị khác.

Bộ mã hóa video 20 và bộ giải mã video 30 có thể hoạt động theo tiêu chuẩn tạo mã video, như tiêu chuẩn tạo mã video hiệu suất cao (High Efficiency Video Coding, HEVC) hiện đang sự phát triển, và có thể phù hợp với mô hình thử nghiệm HEVC (HM). Theo cách khác, bộ mã hóa video 20 và bộ giải mã video 30 có thể hoạt động theo các tiêu chuẩn độc quyền hoặc công nghiệp khác, như tiêu chuẩn H.264 của lĩnh vực chuẩn hóa viễn thông của hiệp hội viễn thông quốc tế (ITU-T), theo cách khác được gọi là nhóm chuyên gia hình ảnh động (Moving Picture Expert Group, MPEG)-4, phần 10, tạo mã video tiên tiến (AVC), H.265/HEVC, hoặc các sự mở rộng của các tiêu chuẩn này. Tuy nhiên, các kỹ thuật của sáng chế không bị giới hạn ở tiêu chuẩn tạo mã cụ thể bất kỳ. Các ví dụ khác về các tiêu chuẩn tạo mã video bao gồm MPEG-2 và ITU-T H.263. Mặc dù không được thể hiện trên Fig.1, nhưng theo một số khía cạnh, bộ mã hóa video 20 và bộ giải mã video 30, mỗi bộ có thể mỗi được tích hợp với bộ mã hóa và bộ giải mã audio, và có thể bao gồm các bộ phận dồn kênh-bộ phận phân kênh (MUX-DEMUX) thích hợp, hoặc phần cứng và phần mềm khác, để xử lý việc mã hóa cả audio

lẫn video trong một luồng dữ liệu chung hoặc các luồng dữ liệu riêng rẽ. Nếu áp dụng được, các bộ phận MUX-DEMUX có thể phù hợp với giao thức bộ dồn kênh ITU H.223, hoặc các giao thức khác như giao thức gói dữ liệu người dùng (user datagram protocol, UDP).

Bộ mã hóa video 20 và bộ giải mã video 30, mỗi bộ có thể được thực thi như là loại bất kỳ trong số nhiều loại mạch bộ mã hóa thích hợp, như một hoặc nhiều bộ vi xử lý, bộ xử lý tín hiệu số (digital signal processor, DSP), mạch tích hợp chuyên dụng (application specific integrated circuit, ASIC), mảng cổng lập trình được dạng trường (field programmable gate array, FPGA), bộ logic rời rạc, phần mềm, phần cứng, phần sụn hoặc các sự kết hợp bất kỳ của chúng. Khi các kỹ thuật được thực hiện một phần trong phần mềm, thì thiết bị có thể lưu trữ các lệnh đối với phần mềm trong phương tiện lâu dài đọc được bằng máy tính thích hợp và thực hiện các lệnh trong phần cứng nhờ sử dụng một hoặc nhiều bộ xử lý để thực hiện các kỹ thuật của sáng chế. Mỗi bộ trong số bộ mã hóa video 20 và bộ giải mã video 30 có thể có trong một hoặc nhiều bộ mã hóa hoặc bộ giải mã, mỗi loại có thể được tích hợp như là một phần của bộ mã hóa/bộ giải mã kết hợp (CODEC - CODEC) trong thiết bị tương ứng. Thiết bị bao gồm bộ mã hóa video 20 và/hoặc bộ giải mã video 30 có thể bao gồm mạch tích hợp, bộ vi xử lý, và/hoặc thiết bị truyền thông vô tuyến, như điện thoại mạng dạng ô.

Fig.2 là sơ đồ khối minh họa một ví dụ về bộ mã hóa video 20 mà có thể thực hiện các kỹ thuật tạo mã video. Bộ mã hóa video 20 có thể thực hiện tạo mã trong ảnh và liên ảnh của các khối video trong các lát video. Tạo mã trong ảnh dựa trên dự đoán không gian để làm giảm hoặc loại bỏ dư thừa không gian trong video trong khung video hoặc hình ảnh đã cho. Tạo mã liên ảnh dựa trên dự đoán thời gian để làm giảm hoặc loại bỏ dư thừa thời gian trong video trong các khung liên kề hoặc các hình ảnh của chuỗi video. Chế độ trong ảnh (chế độ I) có thể liên quan đến chế độ bất kỳ trong số một số chế độ tạo mã dựa trên không gian. Các chế độ liên ảnh, như dự đoán đơn hướng (cũng được gọi là dự đoán đơn) (chế độ P) hoặc song dự đoán (cũng được gọi là dự đoán song) (chế độ B), có thể liên quan đến chế độ bất kỳ trong số một số chế độ tạo mã dựa trên thời gian.

Như được thể hiện trên Fig.2, bộ mã hóa video 20 nhận khối video hiện hành

trong khung video cần được mã hóa. Trong ví dụ trên Fig.2, bộ mã hóa video 20 bao gồm bộ phận chọn chế độ 40, bộ nhớ khung tham chiếu 64, bộ cộng 50, bộ phận xử lý biến đổi 52, bộ phận lượng tử hóa 54, và bộ phận tạo mã entropi 56. Bộ phận chọn chế độ 40 bao gồm bộ phận bù chuyển động 44, bộ phận ước lượng chuyển động 42, bộ phận dự đoán trong ảnh (cũng được gọi là dự đoán-trong ảnh) 46, và bộ phận phân chia 48. Đối với việc tái tạo khối video, bộ mã hóa video 20 cũng bao gồm bộ phận lượng tử hóa ngược 58, bộ phận biến đổi ngược 60, và bộ cộng 62. Bộ lọc giải khối (không được thể hiện trên Fig.2) cũng có thể được bao gồm để lọc các biên khối để loại bỏ các vật tạo tác dạng khối từ video được tái tạo. Nếu muốn, bộ lọc giải khối sẽ thường lọc đầu ra của bộ cộng 62. Các bộ lọc bổ sung (nội vòng hoặc sau vòng) cũng có thể được sử dụng ngoài bộ lọc giải khối. Các bộ lọc này không được thể hiện trên hình vẽ cho ngắn gọn, nhưng nếu muốn, có thể lọc đầu ra của bộ cộng 50 (như là bộ lọc nội vòng).

Trong suốt quá trình mã hóa, bộ mã hóa video 20 nhận khung video hoặc lát cần được tạo mã. Khung hoặc lát có thể được chia thành nhiều khối video. Bộ phận ước lượng chuyển động 42 và bộ phận bù chuyển động 44 thực hiện việc tạo mã dự đoán liên ảnh đối với khối video nhận được so với một hoặc nhiều khối trong một hoặc nhiều khung tham chiếu để cung cấp dự đoán thời gian. Bộ phận dự đoán trong ảnh 46 theo cách khác có thể thực hiện việc tạo mã dự đoán trong ảnh đối với khối video nhận được so với một hoặc nhiều khối lân cận trong cùng khung hoặc lát như là khối cần được tạo mã để cung cấp dự đoán không gian. Bộ mã hóa video 20 có thể thực hiện nhiều sự cho phép tạo mã, ví dụ, để chọn chế độ tạo mã thích hợp cho mỗi khối của dữ liệu video.

Hơn nữa, bộ phận phân chia 48 có thể phân chia các khối của dữ liệu video thành các khối con, dựa trên việc đánh giá các sơ đồ phân chia trước đó trong các sự cho phép tạo mã trước đó. Ví dụ, bộ phận phân chia 48 có thể phân chia ban đầu khung hoặc lát thành các đơn vị tạo mã lớn nhất (largest coding unit, LCU), và phân chia mỗi trong số các LCU này thành các đơn vị tạo mã con (các CU con) dựa trên phân tích tỷ lệ-méo (ví dụ, tối ưu hóa tỷ lệ-méo). Bộ phận chọn chế độ 40 còn có thể tạo ra cấu trúc dữ liệu cây tứ phân chỉ báo sự phân chia LCU thành các CU con. Các CU nút lá của cây tứ phân có thể bao gồm một hoặc nhiều đơn vị dự đoán (prediction unit, PU) và một hoặc nhiều đơn vị biến đổi (transform unit, TU).

Sáng chế sử dụng thuật ngữ “khối” để đề cập đến loại bất kỳ trong số CU, PU, hoặc TU, trong ngữ cảnh của HEVC, hoặc các cấu trúc dữ liệu tương tự trong ngữ cảnh của các tiêu chuẩn khác (ví dụ, các khối macro và các khối con của nó trong H.264/AVC). CU bao gồm nút tạo mã, các PU, và các TU được liên kết với nút tạo mã. Kích thước của CU tương ứng với kích thước của nút tạo mã và có hình vuông. Kích thước của CU có thể nằm trong khoảng từ 8×8 điểm ảnh đến kích thước của khối cây với tối đa là 64×64 điểm ảnh hoặc lớn hơn. Mỗi CU có thể chứa một hoặc nhiều PU và một hoặc nhiều TU. Dữ liệu cú pháp được liên kết với CU có thể mô tả, ví dụ, việc phân chia CU thành một hoặc nhiều PU. Các chế độ phân chia có thể khác nhau giữa việc liệu CU là chế độ bỏ qua hay chế độ trực tiếp được mã hóa, chế độ dự đoán trong ảnh được mã hóa, hoặc chế độ dự đoán liên ảnh (cũng được gọi là dự đoán-liên ảnh) được mã hóa. Các PU có thể được phân chia để có hình dạng không phải hình vuông. Dữ liệu cú pháp được liên kết với CU cũng có thể mô tả, ví dụ, việc phân chia CU thành một hoặc nhiều TU theo cây tứ phân. TU có thể là hình vuông hoặc hình không vuông (ví dụ, hình chữ nhật).

Bộ phận chọn chế độ 40 có thể chọn một trong số các chế độ tạo mã, trong ảnh hoặc liên ảnh, ví dụ, dựa trên các kết quả sai số, và cung cấp khối được tạo mã liên ảnh hoặc trong ảnh tạo thành cho bộ cộng 50 để tạo dữ liệu khối dư và cho bộ cộng 62 để tái tạo khối được mã hóa để sử dụng làm khung tham chiếu. Bộ phận chọn chế độ 40 cũng cung cấp các phần tử cú pháp, như các vector chuyển động, các bộ chỉ báo chế độ trong ảnh, thông tin phân chia, và thông tin cú pháp khác, cho bộ phận tạo mã entropi 56.

Bộ phận ước lượng chuyển động 42 và bộ phận bù chuyển động 44 có thể được tích hợp ở mức cao, nhưng được minh họa riêng rẽ nhằm các mục đích liên quan đến khái niệm. Ước lượng chuyển động, được thực hiện bởi bộ phận ước lượng chuyển động 42, là quy trình tạo các vector chuyển động, mà ước lượng chuyển động đối với các khối video. Vector chuyển động, ví dụ, có thể chỉ báo sự dịch chuyển của PU của khối video trong khung video hoặc hình ảnh hiện hành so với khối dự đoán trong khung tham chiếu (hoặc đơn vị tạo mã khác) so với khối hiện hành đang được tạo mã trong khung hiện hành (hoặc đơn vị tạo mã khác). Khối dự đoán là khối mà được thấy là khớp gần với khối cần được tạo mã, trên phương diện chênh lệch điểm ảnh, mà có thể được xác định bởi tổng sai phân tuyệt đối (SAD), tổng sai phân bình phương (SSD), hoặc các chuẩn đo

chênh lệch khác. Trong một số ví dụ, bộ mã hóa video 20 có thể tính toán các giá trị đối với các vị trí điểm ảnh nguyên con của các hình ảnh tham chiếu được lưu trữ trong bộ nhớ khung tham chiếu 64. Ví dụ, bộ mã hóa video 20 có thể nội suy các giá trị của các vị trí điểm ảnh một phần tư, các vị trí điểm ảnh một phần tám, hoặc các vị trí điểm ảnh phân số khác của hình ảnh tham chiếu. Do đó, bộ phận ước lượng chuyển động 42 có thể thực hiện việc tìm kiếm chuyển động so với các vị trí điểm ảnh đầy đủ và các vị trí điểm ảnh phân số và kết xuất vector chuyển động với độ chính xác điểm ảnh thập phân.

Bộ phận ước lượng chuyển động 42 tính toán vector chuyển động đối với PU của khối video trong lát được tạo mã liên ảnh bằng cách so sánh vị trí của PU với vị trí của khối dự đoán của hình ảnh tham chiếu. Hình ảnh tham chiếu có thể được chọn từ danh sách hình ảnh tham chiếu thứ nhất (List 0) hoặc danh sách hình ảnh tham chiếu thứ hai (List 1), mỗi danh sách trong số đó nhận dạng một hoặc nhiều hình ảnh tham chiếu được lưu trữ trong bộ nhớ khung tham chiếu 64. Bộ phận ước lượng chuyển động 42 gửi vector chuyển động tính toán được đến bộ phận mã hóa entropi 56 và bộ phận bù chuyển động 44.

Bù chuyển động, được thực hiện bởi bộ phận bù chuyển động 44, có thể bao gồm việc tìm hoặc tạo khối dự đoán dựa trên vector chuyển động được xác định bởi bộ phận ước lượng chuyển động 42. Mặt khác, bộ phận ước lượng chuyển động 42 và bộ phận bù chuyển động 44 có thể được tích hợp về mặt chức năng trong một số ví dụ. Khi nhận vector chuyển động đối với PU của khối video hiện hành, bộ phận bù chuyển động 44 có thể định vị khối dự đoán mà vector chuyển động trỏ đến trong một trong số các danh sách hình ảnh tham chiếu. Bộ cộng 50 tạo ra khối video dư bằng cách trừ các giá trị điểm ảnh của khối dự đoán từ các giá trị điểm ảnh của khối video hiện hành được tạo mã, tạo thành các giá trị chênh lệch điểm ảnh, như được thảo luận dưới đây. Nói chung, bộ phận ước lượng chuyển động 42 thực hiện ước lượng chuyển động so với các thành phần độ chói, và bộ phận bù chuyển động 44 sử dụng các vector chuyển động được tính toán dựa trên các thành phần độ chói đối với cả các thành phần sắc độ lẫn các thành phần độ chói. Bộ phận chọn chế độ 40 cũng có thể tạo các phần tử cú pháp được liên kết với các khối video và lát video để sử dụng bởi bộ giải mã video 30 khi giải mã các khối video của lát video.

Bộ phận dự đoán trong ảnh 46 có thể dự đoán nội bộ khối hiện hành, như là một sự lựa chọn đối với việc dự đoán liên ảnh được thực hiện bởi bộ phận ước lượng chuyển động 42 và bộ phận bù chuyển động 44, như được mô tả trên đây. Cụ thể, bộ phận dự đoán trong ảnh 46 có thể xác định chế độ dự đoán trong ảnh để sử dụng để mã hóa khối hiện hành. Trong một số ví dụ, bộ phận dự đoán trong ảnh 46 có thể mã hóa khối hiện hành sử dụng các chế độ dự đoán trong ảnh khác nhau, ví dụ, trong suốt quá trình cho phép mã hóa riêng rẽ, và bộ phận dự đoán trong ảnh 46 (hoặc bộ phận chọn chế độ 40, trong một số ví dụ) có thể chọn chế độ dự đoán trong ảnh thích hợp để sử dụng từ các chế độ đã kiểm tra.

Ví dụ, bộ phận dự đoán trong ảnh 46 có thể tính toán các giá trị tỷ lệ méo sử dụng phân tích tỷ lệ méo đối với các chế độ dự đoán trong ảnh đã kiểm tra khác nhau, và chọn chế độ dự đoán trong ảnh có các đặc tính tỷ lệ méo tốt nhất trong số các chế độ đã kiểm tra. Việc phân tích tỷ lệ méo xác định chung lượng méo (hoặc sai số) giữa khối được mã hóa và khối chưa được mã hóa gốc mà đã được mã hóa để tạo ra khối được mã hóa, cũng như tốc độ bit (tức là, số lượng bit) được sử dụng để tạo khối được mã hóa. Bộ phận dự đoán trong ảnh 46 có thể tính toán các tỷ lệ từ các phần méo và các tỷ lệ đối với các khối được mã hóa khác nhau để xác định chế độ dự đoán trong ảnh nào thể hiện giá trị tỷ lệ méo tốt nhất đối với khối.

Ngoài ra, bộ phận dự đoán trong ảnh 46 có thể được tạo cấu hình để tạo mã các khối độ sâu của bản đồ độ sâu sử dụng chế độ tạo mô hình độ sâu (depth modeling mode, DMM). Bộ phận chọn chế độ 40 có thể xác định liệu chế độ DMM khả dụng tạo ra các kết quả tạo mã tốt hơn so với chế độ dự đoán trong ảnh và các chế độ DMM khác, ví dụ, sử dụng tối ưu hóa tỷ lệ-méo (rate-distortion optimization, RDO). Dữ liệu cho ảnh bố cục tương ứng với bản đồ độ sâu có thể được lưu trữ trong bộ nhớ khung tham chiếu 64. Bộ phận ước lượng chuyển động 42 và bộ phận bù chuyển động 44 cũng có thể được tạo cấu hình để dự đoán liên ảnh các khối độ sâu của bản đồ độ sâu.

Sau khi chọn chế độ dự đoán trong ảnh đối với khối (ví dụ, chế độ dự đoán trong ảnh thông thường hoặc một trong số các chế độ DMM), bộ phận dự đoán trong ảnh 46 có thể cung cấp thông tin chỉ báo chế độ dự đoán trong ảnh được chọn đối với khối cho bộ phận tạo mã entropi 56. Bộ phận tạo mã entropi 56 có thể mã hóa thông tin chỉ báo

chế độ dự đoán trong ảnh được chọn. Bộ mã hóa video 20 có thể bao gồm trong dữ liệu cấu hình dòng bit được truyền, mà có thể bao gồm các bảng chỉ số chế độ dự đoán trong ảnh và các bảng chỉ số chế độ dự đoán trong ảnh được thay đổi (cũng được gọi là các bảng ánh xạ từ mã), các định nghĩa về các ngữ cảnh mã hóa đối với các khối khác nhau, và các chỉ báo của chế độ dự đoán trong ảnh có khả năng nhất, bảng chỉ số chế độ dự đoán trong ảnh, và bảng chỉ số chế độ dự đoán trong ảnh thay đổi để sử dụng cho mỗi ngữ cảnh trong số các ngữ cảnh này.

Bộ mã hóa video 20 tạo thành khối video dư bằng cách trừ dữ liệu dự đoán từ bộ phận chọn chế độ 40 từ khối video gốc đang được tạo mã. Bộ cộng 50 đại diện cho thành phần hoặc các thành phần thực hiện phép toán trừ này.

Bộ phận xử lý biến đổi 52 áp dụng biến đổi, như biến đổi cosin rời rạc (DCT) hoặc biến đổi tương tự về mặt khái niệm, vào khối dư, tạo ra khối video bao gồm các giá trị hệ số biến đổi dư. Bộ phận xử lý biến đổi 52 có thể thực hiện các sự biến đổi khác mà tương tự về mặt khái niệm với DCT. Biến đổi sóng nhỏ (Wavelet), biến đổi nguyên, biến đổi dải con hoặc các loại biến đổi khác cũng có thể được sử dụng.

Bộ phận xử lý biến đổi 52 áp dụng biến đổi vào khối dư, tạo ra khối của các hệ số biến đổi dư. Việc biến đổi có thể chuyển đổi thông tin còn lại từ miền giá trị điểm ảnh sang miền biến đổi, như miền tần số. Bộ phận xử lý biến đổi 52 có thể gửi các hệ số biến đổi tạo thành đến bộ phận lượng tử hóa 54. Bộ phận lượng tử hóa 54 lượng tử hóa các hệ số biến đổi để làm giảm tốc độ bit thêm nữa. Quy trình lượng tử hóa có thể làm giảm độ sâu bit được liên kết với một số hoặc tất cả hệ số. Mức độ lượng tử hóa có thể được thay đổi bằng cách điều chỉnh thông số lượng tử hóa. Trong một số ví dụ, bộ phận lượng tử hóa 54 sau đó có thể thực hiện việc quét ma trận bao gồm các hệ số biến đổi được lượng tử hóa. Theo cách khác, bộ phận mã hóa entropi 56 có thể thực hiện việc quét.

Tiếp theo việc lượng tử hóa, bộ phận tạo mã entropi 56 tạo mã entropi các hệ số biến đổi được lượng tử hóa. Ví dụ, bộ phận tạo mã entropi 56 có thể thực hiện tạo mã độ dài biến số thích ứng ngữ cảnh (context adaptive variable length coding, CAVLC), tạo mã số học nhị phân thích ứng ngữ cảnh (context adaptive binary arithmetic coding, CABAC), tạo mã số học nhị phân thích ứng ngữ cảnh dựa trên cú pháp (syntax-based context-adaptive binary arithmetic coding, SBAC), tạo mã entropi phân chia khoảng xác

suất (probability interval partitioning entropy, PIPE) hoặc kỹ thuật tạo mã entropi khác. Trong trường hợp tạo mã entropi dựa trên ngữ cảnh, ngữ cảnh có thể dựa trên các khối lân cận. Tiếp theo tạo mã entropi bởi bộ phận tạo mã entropi 56, dòng bit được mã hóa có thể được truyền đến thiết bị khác (ví dụ, bộ giải mã video 30) hoặc thu được để truyền hoặc truy hồi sau này.

Bộ phận lượng tử hóa ngược 58 và bộ phận biến đổi ngược 60 áp dụng tương ứng lượng tử hóa ngược và biến đổi ngược để tái tạo khối dư trong miền điểm ảnh, ví dụ, để sử dụng sau này làm khối tham chiếu. Bộ phận bù chuyển động 44 có thể tính toán khối tham chiếu bằng cách cộng khối dư vào khối dự đoán của một trong số các khung của bộ nhớ khung tham chiếu 64. Bộ phận bù chuyển động 44 cũng có thể áp dụng một hoặc nhiều bộ lọc nội suy vào khối dư được tái tạo để tính toán các giá trị điểm ảnh nguyên con để sử dụng khi ước lượng chuyển động. Bộ cộng 62 cộng khối dư được tái tạo vào khối dự đoán bù chuyển động được tạo ra bởi bộ phận bù chuyển động 44 để tạo ra khối video được tái tạo để lưu trữ trong bộ nhớ khung tham chiếu 64. Khối video được tái tạo có thể được sử dụng bởi bộ phận ước lượng chuyển động 42 và bộ phận bù chuyển động 44 làm khối tham chiếu để tạo mã liên ảnh một khối trong khung video sau đó.

Fig.3 là sơ đồ khối minh họa một ví dụ về bộ giải mã video 30 mà có thể thực hiện các kỹ thuật tạo mã video. Trong ví dụ trên Fig.3, bộ giải mã video 30 bao gồm bộ phận giải mã entropi 70, bộ phận bù chuyển động 72, bộ phận dự đoán trong ảnh 74, bộ phận lượng tử hóa ngược 76, bộ phận biến đổi ngược 78, bộ nhớ khung tham chiếu 82, và bộ cộng 80. Bộ giải mã video 30 có thể, trong một số ví dụ, thực hiện cách giải mã thường ngược với cách mã hóa được mô tả liên quan đến bộ mã hóa video 20 (Fig.2). Bộ phận bù chuyển động 72 có thể tạo dữ liệu dự đoán dựa trên các vector chuyển động được nhận từ bộ phận giải mã entropi 70, trong khi bộ phận dự đoán trong ảnh 74 có thể tạo dữ liệu dự đoán dựa trên các bộ chỉ báo chế độ dự đoán trong ảnh được nhận từ bộ phận giải mã entropi 70.

Trong suốt quá trình giải mã, bộ giải mã video 30 nhận dòng bit video được mã hóa mà đại diện cho các khối video của lát video được mã hóa và các phần tử cú pháp liên kết từ bộ mã hóa video 20. Bộ phận giải mã entropi 70 của bộ giải mã video 30 giải mã entropi dòng bit để tạo các hệ số được lượng tử hóa, các vector chuyển động hoặc các

bộ chỉ báo chế độ dự đoán trong ảnh, và các phần tử cú pháp khác. Bộ phận giải mã entropi 70 chuyển tiếp các vector chuyển động và các phần tử cú pháp khác đến bộ phận bù chuyển động 72. Bộ giải mã video 30 có thể nhận các phần tử cú pháp ở mức lát video và/hoặc mức khối video.

Khi lát video được tạo mã dưới dạng lát được tạo mã trong ảnh (I), bộ phận dự đoán trong ảnh 74 có thể tạo dữ liệu dự đoán đối với khối video của lát video hiện hành dựa trên chế độ dự đoán trong ảnh và dữ liệu được báo hiệu từ các khối được giải mã trước đó của khung hoặc hình ảnh hiện hành. Khi khung video được tạo mã như là lát được tạo mã liên ảnh (ví dụ, B, P, hoặc GPB), bộ phận bù chuyển động 72 tạo ra các khối dự đoán đối với khối video của lát video hiện hành dựa trên các vector chuyển động và các phần tử cú pháp khác được nhận từ bộ phận giải mã entropi 70. Các khối dự đoán có thể được tạo ra từ một trong số các hình ảnh tham chiếu trong một trong số các danh sách hình ảnh tham chiếu. Bộ giải mã video 30 có thể xây dựng các danh sách khung tham chiếu, List (danh sách) 0 và List (danh sách) 1, sử dụng các kỹ thuật xây dựng mặc định dựa trên các hình ảnh tham chiếu được lưu trữ trong bộ nhớ khung tham chiếu 82.

Bộ phận bù chuyển động 72 xác định thông tin dự đoán đối với khối video của lát video hiện hành bằng cách phân tách các vector chuyển động và các phần tử cú pháp khác, và sử dụng thông tin dự đoán để tạo ra các khối dự đoán đối với khối video hiện hành đang được giải mã. Ví dụ, bộ phận bù chuyển động 72 sử dụng một số trong số các phần tử cú pháp được nhận để xác định chế độ dự đoán (ví dụ, dự đoán trong ảnh hoặc liên ảnh) được sử dụng để tạo mã các khối video của lát video, loại lát dự đoán liên ảnh (ví dụ, lát B, lát P, hoặc lát GPB), thông tin xây dựng đối với một hoặc nhiều danh sách trong số các danh sách hình ảnh tham chiếu đối với lát, các vector chuyển động đối với mỗi khối video được mã hóa liên ảnh của lát, trạng thái dự đoán liên ảnh đối với mỗi khối video được mã hóa liên ảnh của lát, và thông tin khác để giải mã các khối video trong lát video hiện hành.

Bộ phận bù chuyển động 72 cũng có thể thực hiện nội suy dựa trên các bộ lọc nội suy. Bộ phận bù chuyển động 72 có thể sử dụng các bộ lọc nội suy như được sử dụng bởi bộ mã hóa video 20 trong suốt quá trình mã hóa các khối video để tính toán các giá trị được nội suy đối với các điểm ảnh nguyên con của các khối tham chiếu. Trong

trường hợp này, bộ phận bù chuyển động 72 có thể xác định các bộ lọc nội suy được sử dụng bởi bộ mã hóa video 20 từ các phần tử cú pháp được nhận và sử dụng các bộ lọc nội suy để tạo ra các khối dự đoán.

Dữ liệu cho ảnh bố cục tương ứng với bản đồ độ sâu có thể được lưu trữ trong bộ nhớ khung tham chiếu 82. Bộ phận bù chuyển động 72 cũng có thể được tạo cấu hình để dự đoán liên ảnh các khối độ sâu của bản đồ độ sâu.

Theo một phương án, bộ giải mã video 30 bao gồm giao diện người dùng (UI) 84. Giao diện người dùng 84 được tạo cấu hình để nhận đầu vào từ người dùng của bộ giải mã video 30 (ví dụ, người quản trị mạng). Qua giao diện người dùng 84, người dùng có thể quản lý hoặc thay đổi các thiết đặt trên bộ giải mã video 30. Ví dụ, người dùng có thể nhập hoặc nếu không thì cung cấp giá trị đối với thông số (ví dụ, cờ) để điều khiển cấu hình và/hoặc hoạt động của bộ giải mã video 30 theo quyền ưu tiên của người dùng. Giao diện người dùng 84 có thể, ví dụ, là giao diện người dùng đồ họa (GUI) mà cho phép người dùng tương tác với bộ giải mã video 30 qua các biểu tượng đồ họa, các bảng chọn thả xuống, các hộp kiểm tra, v.v.. Trong một số trường hợp, giao diện người dùng 84 có thể nhận thông tin từ người dùng qua bàn phím, chuột, hoặc thiết bị ngoại vi khác. Theo một phương án, người dùng có thể truy cập giao diện người dùng 84 qua điện thoại thông minh, thiết bị máy tính bảng, máy tính cá nhân được đặt ở xa từ bộ giải mã video 30, v.v.. Như được sử dụng ở đây, giao diện người dùng 84 có thể được gọi là đầu vào bên ngoài hoặc phương tiện bên ngoài.

Lưu ý, các kỹ thuật nén video thực hiện dự đoán không gian (trong hình ảnh) và/hoặc dự đoán thời gian (liên hình ảnh) để làm giảm hoặc loại bỏ phần dư thừa vốn có trong các chuỗi video. Đối với việc tạo mã video dựa trên khối, lát video (tức là, hình ảnh video hoặc một phần của hình ảnh video) có thể được phân chia thành các khối video, mà cũng có thể được gọi là các khối cây (treeblocks), các khối cây tạo mã (coding tree block, CTB), các đơn vị cây tạo mã (coding tree unit, CTU), các đơn vị tạo mã (coding unit, CU) và/hoặc các nút tạo mã. Các khối video trong lát được tạo mã trong ảnh (I) của hình ảnh được mã hóa sử dụng dự đoán không gian đối với các mẫu tham chiếu trong các khối lân cận trong cùng một hình ảnh. Các khối video trong lát được tạo mã liên ảnh (P hoặc B) của hình ảnh có thể sử dụng dự đoán không gian đối

với các mẫu tham chiếu trong các khối lân cận trong cùng một hình ảnh hoặc dự đoán thời gian đối với các mẫu tham chiếu trong các hình ảnh tham chiếu khác. Các hình ảnh có thể được gọi là các khung, và các hình ảnh tham chiếu có thể được gọi các khung tham chiếu.

Dự đoán không gian hoặc thời gian dẫn đến khối dự đoán cho khối cần được tạo mã. Dữ liệu dư thể hiện các sự chênh lệch điểm ảnh giữa khối gốc cần được tạo mã và khối dự đoán. Khối được tạo mã liên ảnh được mã hóa theo vector chuyển động mà trở đến khối của các mẫu tham chiếu tạo thành khối dự đoán, và dữ liệu dư chỉ báo sự chênh lệch giữa khối được tạo mã và khối dự đoán. Khối được tạo mã trong ảnh được mã hóa theo chế độ tạo mã trong ảnh và dữ liệu dư. Để nén thêm, dữ liệu dư có thể được biến đổi từ miền điểm ảnh sang miền biến đổi, tạo ra các hệ số biến đổi dư, mà sau đó có thể được lượng tử hóa. Các hệ số biến đổi được lượng tử hóa, ban đầu được bố trí trong mảng hai chiều, có thể được quét để tạo ra vector một chiều của các hệ số biến đổi, và mã hóa entropi có thể được áp dụng để thu được kết quả nén thậm chí là tốt hơn.

Ảnh và video nén đã trải qua tăng trưởng nhanh chóng, dẫn đến các tiêu chuẩn tạo mã khác nhau. Các tiêu chuẩn tạo mã video này bao gồm ITU-T H.261, MPEG-1 phần 2 của tổ chức chuẩn hóa quốc tế/Ủy ban kỹ thuật điện quốc tế (ISO/IEC), ITU-T H.262 hoặc ISO/IEC MPEG-2 phần 2, ITU-T H.263, ISO/IEC MPEG-4 phần 2, tạo mã video tiên tiến (Advanced Video Coding, AVC), cũng được gọi là ITU-T H.264 hoặc ISO/IEC MPEG-4 phần 10, và tạo mã video hiệu suất cao (HEVC), cũng được gọi là ITU-T H.265 hoặc MPEG-H phần 2. AVC bao gồm các loại mở rộng như tạo mã video thay đổi được (Scalable Video Coding, SVC), tạo mã video đa cảnh nhìn (Multiview Video Coding, MVC) và tạo mã video đa cảnh nhìn cộng độ sâu (Multiview Video Coding plus Depth, MVC+D), và 3D AVC (3D-AVC). HEVC bao gồm các loại mở rộng như HEVC thay đổi được (SHVC), HEVC đa cảnh nhìn (MV-HEVC), và 3D HEVC (3D-HEVC).

Cũng có tiêu chuẩn tạo mã video mới, được đặt tên là tạo mã video đa năng (Versatile Video Coding, VVC), được phát triển bởi nhóm chuyên gia video chung (joint video experts team, JVET) của ITU-T và ISO/IEC. Mặc dù tiêu chuẩn VVC có một số bản phác thảo công việc, nhưng một bản phác thảo công việc (WD) của VVC cụ

thể, cụ thể B. Bross, J. Chen, và S. Liu, “Versatile Video Coding (Draft 5),” JVET-N1001-v3, 13th JVET Meeting, March 27, 2019 (phác thảo VVC 5) được tham chiếu ở đây.

Phần mô tả về các kỹ thuật được bộc lộ ở đây dựa trên tiêu chuẩn tạo mã video đang phát triển Tạo mã video đa năng (Versatile Video Coding, VVC) bởi nhóm chuyên gia video chung (joint video experts team, JVET) của ITU-T và ISO/IEC. Tuy nhiên, các kỹ thuật này cũng áp dụng vào các đặc tả codec video khác.

Fig.4 là hình vẽ thể hiện sự đại diện 400 của mối quan hệ giữa hình ảnh điểm truy cập ngẫu nhiên trong ảnh (intra random access picture, IRAP) 402 cân xứng với các hình ảnh chủ đạo 404 và các hình ảnh kế tiếp 406 theo thứ tự giải mã 408 và thứ tự trình diễn 410. Theo một phương án, hình ảnh IRAP 402 được gọi là hình ảnh truy cập ngẫu nhiên sạch (CRA) hoặc hình ảnh làm mới bộ giải mã tức thời (IDR) với hình ảnh truy cập ngẫu nhiên giải mã được (random access decodable, RADL). Trong HEVC, các hình ảnh IDR, các hình ảnh CRA, và các hình ảnh truy cập liên kết gãy (Broken Link Access, BLA) đều được coi là các hình ảnh IRAP 402. Đối với VVC, hội nghị JVET thứ 12 vào tháng 10 năm 2018, điều được đồng ý là có cả các hình ảnh IDR và CRA là các hình ảnh IRAP. Theo một phương án, các hình ảnh truy cập liên kết gãy (BLA) và làm mới giải mã từng bước (GDR) cũng có thể được coi là các hình ảnh IRAP. Quy trình giải mã đối với chuỗi video được tạo mã luôn bắt đầu tại IRAP.

Hình ảnh CRA là hình ảnh IRAP mà mỗi đơn vị lớp trừu tượng mạng (NAL) của lớp tạo mã video (VCL) có nal_unit_type bằng CRA_NUT. Hình ảnh CRA không liên quan đến các hình ảnh bất kỳ không phải là chính nó đối với dự đoán liên ảnh trong quy trình giải mã của nó, và có thể là hình ảnh thứ nhất trong dòng bit theo thứ tự giải mã, hoặc có thể xuất hiện sau đó trong dòng bit. Hình ảnh CRA có thể có các hình ảnh chủ đạo bỏ qua truy cập ngẫu nhiên (random access skipped leading, RASL) hoặc RADL liên kết. Khi hình ảnh CRA có NoOutputBeforeRecoveryFlag bằng 1, các hình ảnh RASL liên kết không được kết xuất bởi bộ giải mã, do chúng có thể không giải mã được, do chúng có thể chứa các tham chiếu đến các hình ảnh mà không có trong dòng bit.

Như được thể hiện trên Fig.4, các hình ảnh chủ đạo 404 (ví dụ, các hình ảnh 2 và

3) theo sau hình ảnh IRAP 402 theo thứ tự giải mã 408, nhưng trước hình ảnh IRAP 402 theo thứ tự trình diễn 410. Hình ảnh kế tiếp 406 theo sau hình ảnh IRAP 402 cả theo thứ tự giải mã 408 lẫn theo thứ tự trình diễn 410. Mặc dù hai hình ảnh chủ đạo 404 và một hình ảnh kế tiếp 406 được thể hiện trên Fig.4, nhưng người có hiểu biết trung bình trong lĩnh vực tương ứng sẽ hiểu rằng có thể có nhiều hoặc ít hình ảnh chủ đạo 404 và/hoặc hình ảnh kế tiếp 406 hơn theo thứ tự giải mã 408 và thứ tự trình diễn 410 trong các ứng dụng thực tế.

Các hình ảnh chủ đạo 404 trên Fig.4 đã được chia thành hai loại, cụ thể chủ đạo bỏ qua truy cập ngẫu nhiên (random access skipped leading, RASL) và RADL. Khi việc giải mã bắt đầu với hình ảnh IRAP 402 (ví dụ, hình ảnh 1), hình ảnh RADL (ví dụ, hình ảnh 3) có thể được giải mã chính xác; tuy nhiên, hình ảnh RASL (ví dụ, hình ảnh 2) không thể được giải mã chính xác. Vì vậy, hình ảnh RASL được loại bỏ. Xét đến sự phân biệt giữa các hình ảnh RADL và RASL, loại của hình ảnh chủ đạo 404 được liên kết với hình ảnh IRAP 402 sẽ được nhận dạng là RADL hoặc RASL để tạo mã hiệu quả và chính xác. Trong HEVC, khi các hình ảnh RASL và RADL có mặt, điều bị ràng buộc là đối với các hình ảnh RASL và RADL mà được liên kết với cùng một hình ảnh IRAP 402, các hình ảnh RASL sẽ ở trước các hình ảnh RADL theo thứ tự trình diễn 410.

Hình ảnh IRAP 402 cung cấp hai chức năng / lợi ích quan trọng sau đây. Trước tiên, sự có mặt của hình ảnh IRAP 402 chỉ báo rằng quy trình giải mã có thể bắt đầu từ hình ảnh đó. Chức năng này cho phép dấu hiệu truy cập ngẫu nhiên trong đó quy trình giải mã bắt đầu ở vị trí đó trong dòng bit, không nhất thiết là đầu của dòng bit, miễn là hình ảnh IRAP 402 có mặt ở vị trí đó. Thứ hai, sự có mặt của hình ảnh IRAP 402 làm mới quy trình giải mã sao cho hình ảnh được tạo mã bắt đầu tại hình ảnh IRAP 402, trừ các hình ảnh RASL, được tạo mã mà không cần tham chiếu bất kỳ đến các hình ảnh trước đó. Việc có hình ảnh IRAP 402 trong dòng bit do đó sẽ dùng lỗi bất kỳ mà có thể xảy ra trong suốt quá trình giải mã các hình ảnh được tạo mã trước hình ảnh IRAP 402 để lan truyền đến hình ảnh IRAP 402 và các hình ảnh đó mà cho phép hình ảnh IRAP 402 theo thứ tự giải mã 408.

Mặc dù các hình ảnh IRAP 402 cung cấp các chức năng quan trọng, nhưng chúng sẽ mang lại bất lợi về hiệu suất nén. Sự có mặt của hình ảnh IRAP 402 gây ra tăng đột

biến về tốc độ bit. Bất lợi này đối với hiệu suất nén là bởi hai lý do. Trước tiên, do một hình ảnh IRAP 402 là hình ảnh được dự đoán trong ảnh, nên chính hình ảnh sẽ đòi hỏi khá nhiều bit hơn để đại diện khi so với các hình ảnh khác (ví dụ, các hình ảnh chủ đạo 404, các hình ảnh kế tiếp 406) mà là các hình ảnh được dự đoán liên ảnh. Thứ hai, do sự có mặt của hình ảnh IRAP 402 phá vỡ dự đoán thời gian (điều này là do bộ giải mã sẽ làm mới quy trình giải mã, trong đó một trong số các hoạt động của quy trình giải mã dùng cho việc này là để loại bỏ các hình ảnh tham chiếu trước đó trong bộ đệm hình ảnh được giải mã (DPB)), hình ảnh IRAP 402 làm cho việc tạo mã các hình ảnh mà theo sau hình ảnh IRAP 402 theo thứ tự giải mã 408 có hiệu quả ít hơn (tức là, cần nhiều bit hơn để đại diện) do chúng có ít hình ảnh tham chiếu hơn cho việc tạo mã dự đoán liên ảnh của chúng.

Trong số các loại hình ảnh mà được coi là các hình ảnh IRAP 402, hình ảnh IDR trong HEVC có báo hiệu và dẫn xuất khác khi so với các loại hình ảnh khác. Một số sự khác biệt là như sau.

Để báo hiệu và dẫn xuất giá trị đếm thứ tự hình ảnh (POC) của hình ảnh IDR, phần bit có trọng số lớn nhất (MSB) của POC không được dẫn xuất từ hình ảnh khóa trước đó nhưng được thiết đặt đơn giản để bằng 0.

Để báo hiệu thông tin cần để quản lý hình ảnh tham chiếu, phần đầu lát của hình ảnh IDR không chứa thông tin cần để được báo hiệu để hỗ trợ quản lý hình ảnh tham chiếu. Đối với các loại hình ảnh khác (tức là, CRA, kế tiếp, truy cập lớp con thời gian (temporal sub-layer access, TSA), v.v.), thông tin như tập hợp hình ảnh tham chiếu (RPS) được mô tả dưới đây hoặc các dạng thông tin tương tự khác (ví dụ, các danh sách hình ảnh tham chiếu) là cần cho quy trình đánh dấu các hình ảnh tham chiếu (tức là, quy trình để xác định trạng thái của các hình ảnh tham chiếu trong bộ đệm hình ảnh được giải mã (decoded picture buffer, DPB), một loại trong số được sử dụng để tham chiếu và không được sử dụng để tham chiếu). Tuy nhiên, đối với hình ảnh IDR, thông tin này không cần được báo hiệu do sự có mặt của IDR chỉ báo rằng quy trình giải mã sẽ đánh dấu một cách đơn giản tất cả các hình ảnh tham chiếu trong DPB là không được sử dụng để tham chiếu.

Trong HEVC và VVC, các hình ảnh IRAP 402 và các hình ảnh chủ đạo 404 mỗi

hình ảnh có thể được có trong đơn vị lớp trừu tượng mạng (NAL) đơn lẻ. Tập hợp của các đơn vị NAL có thể được gọi là đơn vị truy cập. Các hình ảnh IRAP 402 và các hình ảnh chủ đạo 404 là các loại đơn vị NAL khác nhau đã cho sao cho chúng có thể được nhận dạng dễ dàng bởi các ứng dụng mức hệ thống. Ví dụ, bộ nối video cần hiểu các loại hình ảnh được tạo mã mà không phải hiểu quá nhiều chi tiết về phần tử cú pháp trong dòng bit đã tạo mã, cụ thể để nhận dạng các hình ảnh IRAP 402 từ các hình ảnh phi IRAP và để nhận dạng các hình ảnh chủ đạo 404, bao gồm việc xác định các hình ảnh RASL và RADL, từ các hình ảnh kế tiếp 406. Các hình ảnh kế tiếp 406 là các hình ảnh mà được liên kết với hình ảnh IRAP 402 và theo sau hình ảnh IRAP 402 theo thứ tự trình diễn 410. Hình ảnh có thể theo sau hình ảnh IRAP 402 cụ thể theo thứ tự giải mã 408 và ở trước hình ảnh IRAP 402 khác bất kỳ theo thứ tự giải mã 408. Đối với điều này, việc cho các hình ảnh IRAP 402 đã cho và các hình ảnh chủ đạo 404 loại đơn vị NAL của chính chúng giúp các ứng dụng này.

Đối với HEVC, các loại đơn vị NAL đối với các hình ảnh IRAP bao gồm các loại sau đây:

BLA với hình ảnh chủ đạo (BLA_W_LP): Đơn vị NAL của hình ảnh truy cập liên kết gây (BLA) mà có thể được theo sau bởi một hoặc nhiều hình ảnh chủ đạo theo thứ tự giải mã.

BLA với RADL (BLA_W_RADL): Đơn vị NAL của hình ảnh BLA mà có thể được theo sau bởi một hoặc nhiều hình ảnh RADL nhưng không có hình ảnh RASL theo thứ tự giải mã.

BLA không có hình ảnh chủ đạo (BLA_N_LP): Đơn vị NAL của hình ảnh BLA mà không được theo sau bởi hình ảnh chủ đạo theo thứ tự giải mã.

IDR với RADL (IDR_W_RADL): Đơn vị NAL của hình ảnh IDR mà có thể được theo sau bởi một hoặc nhiều hình ảnh RADL nhưng không có hình ảnh RASL theo thứ tự giải mã.

IDR không có hình ảnh chủ đạo (IDR_N_LP): Đơn vị NAL của hình ảnh IDR mà không được theo sau bởi hình ảnh chủ đạo theo thứ tự giải mã.

CRA: Đơn vị NAL của hình ảnh truy cập ngẫu nhiên sạch (CRA) mà có thể được

theo sau bởi các hình ảnh chủ đạo (tức là, các hình ảnh RASL hoặc các hình ảnh RADL hoặc cả hai).

RADL: Đơn vị NAL của hình ảnh RADL.

RASL: Đơn vị NAL của hình ảnh RASL.

Đối với VVC, loại đơn vị NAL đối với các hình ảnh IRAP 402 và các hình ảnh chủ đạo 404 như sau:

IDR với RADL (IDR_W_RADL): Đơn vị NAL của hình ảnh IDR mà có thể được theo sau bởi một hoặc nhiều hình ảnh RADL nhưng không có hình ảnh RASL theo thứ tự giải mã.

IDR không có hình ảnh chủ đạo (IDR_N_LP): Đơn vị NAL của hình ảnh IDR mà không được theo sau bởi hình ảnh chủ đạo theo thứ tự giải mã.

CRA: Đơn vị NAL của hình ảnh truy cập ngẫu nhiên sạch (CRA) mà có thể được theo sau bởi các hình ảnh chủ đạo (tức là, các hình ảnh RASL hoặc các hình ảnh RADL hoặc cả hai).

RADL: Đơn vị NAL của hình ảnh RADL.

RASL: Đơn vị NAL của hình ảnh RASL.

Fig.5 minh họa dòng bit video 550 được tạo cấu hình để thực hiện kỹ thuật làm mới giải mã từng bước (GDR) 500. Như được sử dụng ở đây, dòng bit video 550 cũng có thể được gọi là dòng bit video được tạo mã, dòng bit, hoặc các biến thể của chúng. Như được thể hiện trên Fig.5, dòng bit 550 bao gồm tập hợp thông số chuỗi (SPS) 552, tập hợp thông số hình ảnh (PPS) 554, phần đầu lát 556, và dữ liệu ảnh 558.

SPS 552 chứa dữ liệu mà là chung cho tất cả các hình ảnh trong chuỗi các hình ảnh (SOP). Ngược lại, PPS 554 chứa dữ liệu mà là chung cho toàn bộ hình ảnh. Phần đầu lát 556 chứa thông tin về lát hiện hành như, ví dụ, loại lát, mà là của các hình ảnh tham chiếu sẽ được sử dụng, v.v.. SPS 552 và PPS 554 thường có thể được gọi là tập hợp thông số. SPS 552, PPS 554, và phần đầu lát 556 là các loại đơn vị lớp trừu tượng mạng (NAL). Đơn vị NAL là cấu trúc cú pháp chứa chỉ báo về loại dữ liệu cần tuân theo (ví dụ, dữ liệu video đã tạo mã). Các đơn vị NAL được phân loại thành các đơn vị NAL

lớp tạo mã video (VCL) và các đơn vị NAL phi VCL. Các đơn vị NAL VCL chứa dữ liệu mà đại diện cho các giá trị của các mẫu trong các hình ảnh video, và các đơn vị NAL phi VCL chứa thông tin bổ sung liên kết bất kỳ như các tập hợp thông số (dữ liệu phần đầu quan trọng mà có thể áp dụng vào số lượng đơn vị NAL VCL lớn) và thông tin nâng cao phụ (thông tin định thời và dữ liệu bổ sung khác mà có thể nâng cao khả năng sử dụng của tín hiệu video được giải mã nhưng không cần giải mã các giá trị của các mẫu trong các hình ảnh video). Người có hiểu biết trung bình trong lĩnh vực tương ứng sẽ hiểu rằng dòng bit 550 có thể chứa các thông số khác và thông tin trong các ứng dụng thực tế.

Dữ liệu ảnh 558 trên Fig.5 bao gồm dữ liệu được liên kết với các ảnh hoặc video đang được mã hóa hoặc được giải mã. Dữ liệu ảnh 558 có thể được gọi đơn giản là phần tải dữ liệu hoặc dữ liệu đang được mang trong dòng bit 550. Theo một phương án, dữ liệu ảnh 558 bao gồm CVS 508 (hoặc CLVS) chứa hình ảnh GDR 502, một hoặc nhiều hình ảnh kế tiếp 504, và hình ảnh điểm phục hồi 506. Theo một phương án, hình ảnh GDR 502 được gọi là hình ảnh bắt đầu CVS (CVSS). CVS 508 là chuỗi video được tạo mã cho mọi chuỗi video lớp được tạo mã (coded layer video sequence, CLVS) trong dòng bit video 550. Đáng chú ý là, CVS và CLVS là giống nhau khi dòng bit video 550 bao gồm một lớp đơn lẻ. CVS và CLVS chỉ khác nhau khi dòng bit video 550 bao gồm nhiều lớp. Theo một phương án, các hình ảnh kế tiếp 504 có thể được xem xét là một dạng của hình ảnh GDR do chúng ở trước hình ảnh điểm phục hồi 506 trong chu kỳ GDR.

Theo một phương án, hình ảnh GDR 502, các hình ảnh kế tiếp 504, và hình ảnh điểm phục hồi 506 có thể xác định chu kỳ GDR trong CVS 508. Theo một phương án, thứ tự giải mã bắt đầu với hình ảnh GDR 502, tiếp tục với các hình ảnh kế tiếp 504, và sau đó chuyển sang hình ảnh điểm phục hồi 506.

CVS 508 là một chuỗi các hình ảnh (hoặc các phần của nó) bắt đầu với hình ảnh GDR 502 và bao gồm tất cả các hình ảnh (hoặc các phần của nó) cho đến, nhưng không bao gồm, hình ảnh GDR tiếp theo hoặc cho đến khi kết thúc dòng bit. Chu kỳ GDR là một chuỗi các hình ảnh bắt đầu với hình ảnh GDR 502 và bao gồm tất cả các hình ảnh cho đến và bao gồm hình ảnh điểm phục hồi 506. Quy trình giải mã đối với CVS 508

luôn bắt đầu tại hình ảnh GDR 502.

Như được thể hiện trên Fig.5, kỹ thuật GDR 500 hoặc các công việc nguyên tắc qua một chuỗi các hình ảnh bắt đầu với hình ảnh GDR 502 và kết thúc với hình ảnh điểm phục hồi 506. Hình ảnh GDR 502 chứa vùng được làm mới / sạch 510 chứa các khối mà đều đã được tạo mã sử dụng dự đoán trong ảnh (tức là, các khối được dự đoán trong ảnh) và vùng không được làm mới / sạch 512 chứa các khối mà đều đã được tạo mã sử dụng dự đoán liên ảnh (tức là, các khối được dự đoán liên ảnh).

Hình ảnh kế tiếp 504 liên sát với hình ảnh GDR 502 chứa vùng được làm mới / sạch 510 có phần thứ nhất 510A được tạo mã nhờ sử dụng dự đoán trong ảnh và phần thứ hai 510B được tạo mã nhờ sử dụng dự đoán liên ảnh. Phần thứ hai 510B được tạo mã bằng cách tham chiếu vùng được làm mới / sạch 510 của, ví dụ, hình ảnh trước trong chu kỳ GDR của CVS 508. Như được thể hiện, vùng được làm mới / sạch 510 của các hình ảnh kế tiếp 504 mở rộng do quy trình tạo mã di chuyển hoặc tiến triển theo chiều nhất quán (ví dụ, từ trái sang phải), mà làm co lại tương ứng vùng không được làm mới / sạch 512. Cuối cùng, hình ảnh điểm phục hồi 506, mà chứa chỉ vùng được làm mới / sạch 510, được thu nhận từ quy trình tạo mã. Đáng chú ý là, và như sẽ được thảo luận thêm dưới đây, phần thứ hai 510B của vùng được làm mới / sạch 510, mà được tạo mã như là các khối được dự đoán liên ảnh, có thể chỉ đề cập đến vùng được làm mới / sạch 510 trong hình ảnh tham chiếu.

Như được thể hiện trên Fig.5, hình ảnh GDR 502, các hình ảnh kế tiếp 504, và hình ảnh điểm phục hồi 506 trong CVS 508, mỗi hình ảnh được chứa trong đơn vị NAL VCL 530 của chính chúng. Tập hợp của các đơn vị NAL VCL 530 trong CVS 508 có thể được gọi là đơn vị truy cập.

Theo một phương án, đơn vị NAL VCL 530 chứa hình ảnh GDR 502 trong CVS 508 có loại đơn vị NAL GDR (GDR_NUT). Điều này có nghĩa là, theo một phương án, đơn vị NAL VCL 530 chứa hình ảnh GDR 502 trong CVS 508 có loại đơn vị NAL duy nhất của chính nó cân xứng với các hình ảnh kế tiếp 504 và hình ảnh điểm phục hồi 506. Theo một phương án, GDR_NUT cho phép dòng bit 550 bắt đầu với hình ảnh GDR 502 thay vì dòng bit 550 phải bắt đầu với hình ảnh IRAP. Việc chỉ định đơn vị NAL VCL 530 của hình ảnh GDR 502 làm GDR_NUT có thể chỉ báo cho, ví dụ, bộ giải mã rằng

đơn vị NAL VCL 530 ban đầu trong CVS 508 chứa hình ảnh GDR 502. Theo một phương án, hình ảnh GDR 502 là hình ảnh ban đầu trong CVS 508. Theo một phương án, hình ảnh GDR 502 là hình ảnh ban đầu trong chu kỳ GDR.

Fig.6 là hình vẽ giản lược minh họa sự tìm kiếm chuyển động không mong muốn 600 khi sử dụng sự hạn chế của bộ mã hóa để hỗ trợ GDR. Như được thể hiện, sự tìm kiếm chuyển động 600 thể hiện hình ảnh hiện hành 602 và hình ảnh tham chiếu 604. Hình ảnh hiện hành 602 và hình ảnh tham chiếu 604, mỗi hình ảnh bao gồm vùng được làm mới 606 được tạo mã với dự đoán trong ảnh, vùng được làm mới 608 được tạo mã với dự đoán liên ảnh, và vùng không được làm mới 608. Vùng được làm mới 604, vùng được làm mới 606, và vùng không được làm mới 608 tương tự như phần thứ nhất 510A của vùng được làm mới / sạch 510, phần thứ hai 510B của vùng được làm mới / sạch 510, và vùng không được làm mới / bẩn 512 trên Fig.5.

Trong suốt quá trình tìm kiếm chuyển động, bộ mã hóa bị ràng buộc hoặc bị ngăn không cho chọn vector chuyển động 610 bất kỳ mà dẫn đến một số trong số các mẫu của khối tham chiếu 612 được định vị bên ngoài vùng được làm mới 606. Điều này xảy ra thậm chí khi khối tham chiếu 612 cung cấp tiêu chuẩn chi phí tỷ lệ méo tốt nhất khi dự đoán khối hiện hành 614 trong hình ảnh hiện hành 602. Vì vậy, Fig.6 minh họa lý do cho tính không tối ưu trong việc tìm kiếm chuyển động 600 khi sử dụng sự hạn chế của bộ mã hóa để hỗ trợ GDR.

Fig.7 minh họa dòng bit video 750 được tạo cấu hình để thực hiện kỹ thuật làm mới giải mã từng bước (GDR) 700. Như được sử dụng ở đây, dòng bit video 750 cũng có thể được gọi là dòng bit video được tạo mã, dòng bit, hoặc các biến thể của chúng. Như được thể hiện trên Fig.7, dòng bit 750 bao gồm tập hợp thông số chuỗi (SPS) 752, tập hợp thông số hình ảnh (PPS) 754, phần đầu lát 756, và dữ liệu ảnh 758. Dòng bit 750, SPS 752, PPS 754, và phần đầu lát 756 trên Fig.7 tương tự như dòng bit 550, SPS 552, PPS 554, và phần đầu lát 556 trên Fig.5. Do đó, để ngắn gọn, phần mô tả của các phần tử này sẽ không được lặp lại.

Dữ liệu ảnh 758 trên Fig.7 bao gồm dữ liệu được liên kết với các ảnh hoặc video đang được mã hóa hoặc được giải mã. Dữ liệu ảnh 758 có thể được gọi đơn giản là phần tải dữ liệu hoặc dữ liệu đang được mang trong dòng bit 750. Theo một phương án, dữ

liệu ảnh 758 bao gồm CVS 708 (hoặc CLVS) chứa hình ảnh GDR 702, một hoặc nhiều hình ảnh kế tiếp 704, và cuối của hình ảnh sau đó 706. Theo một phương án, hình ảnh GDR 702 được gọi là hình ảnh CVSS. Quy trình giải mã đối với CVS 708 luôn bắt đầu tại hình ảnh GDR 702.

Như được thể hiện trên Fig.7, hình ảnh GDR 702, các hình ảnh kế tiếp 704, và cuối của hình ảnh sau đó 706 trong CVS 708, mỗi hình ảnh có trong chính đơn vị NAL VCL 730 của chúng. Tập hợp của các đơn vị NAL VCL 730 trong CVS 708 có thể được gọi là đơn vị truy cập.

Trong đặc tả phác thảo mới nhất của VVC, việc kết xuất các hình ảnh trước đối với các hình ảnh IRAP được quy định như sau. Các hình ảnh trước (ví dụ, các hình ảnh được giải mã trước đó) đối với hình ảnh IRAP liên quan đến các hình ảnh mà 1) được giải mã sớm hơn hình ảnh IRAP, 2) được chỉ báo để kết xuất, 3) có trong bộ đệm hình ảnh được giải mã (DPB) khi bắt đầu giải mã hình ảnh IRAP, và 4) đã không được kết xuất khi bắt đầu giải mã hình ảnh IRAP. Như được sử dụng ở đây, các hình ảnh trước có thể được gọi là các hình ảnh được giải mã trước đó.

Cú pháp phần đầu lát bao gồm phần tử cú pháp `no_output_of_prior_pics_flag` đối với các hình ảnh IDR và CRA. Các ngữ nghĩa là như sau:

`no_output_of_prior_pics_flag` tác động đến việc kết xuất các hình ảnh được giải mã trước đó trong bộ đệm hình ảnh được giải mã sau khi giải mã hình ảnh IDR mà không phải là hình ảnh thứ nhất trong dòng bit như được quy định trong Phụ lục C của Phác thảo VVC 5.

Điều C.3.2 (Loại bỏ các hình ảnh từ DPB trước khi giải mã hình ảnh hiện hành) trong Phác thảo VVC 5 bao gồm nội dung sau đây:

– Khi hình ảnh hiện hành là hình ảnh IRAP với `NoIncorrectPicOutputFlag` bằng 1 mà không phải là hình ảnh 0, các bước có thứ tự sau đây được áp dụng:

1. Biến số `NoOutputOfPriorPicsFlag` được dẫn xuất đối với bộ giải mã theo thử nghiệm như sau:

– Nếu hình ảnh hiện hành là hình ảnh CRA, `NoOutputOfPriorPicsFlag` được thiết đặt bằng 1 (mà không quan tâm đến giá trị của `no_output_of_prior_pics_flag`).

– Nếu không, nếu giá trị của `pic_width_in_luma_samples`, `pic_height_in_luma_samples`, `chroma_format_idc`, `separate_colour_plane_flag`, `bit_depth_luma_minus8`, `bit_depth_chroma_minus8` hoặc `sps_max_dec_pic_buffering_minus1[ HighestTid ]` được dẫn xuất từ SPS hoạt động khác với giá trị của `pic_width_in_luma_samples`, `pic_height_in_luma_samples`, `chroma_format_idc`, `separate_colour_plane_flag`, `bit_depth_luma_minus8`, `bit_depth_chroma_minus8` hoặc `sps_max_dec_pic_buffering_minus1[ HighestTid ]`, một cách tương ứng, được dẫn xuất từ SPS hoạt động đối với hình ảnh trước, `NoOutputOfPriorPicsFlag` có thể (nhưng không nên) được thiết đặt ở 1 bởi bộ giải mã theo thử nghiệm, mà không quan tâm đến giá trị của `no_output_of_prior_pics_flag`.

Lưu ý – Mặc dù việc thiết đặt `NoOutputOfPriorPicsFlag` bằng `no_output_of_prior_pics_flag` được ưu tiên theo các điều kiện này, nhưng bộ giải mã theo thử nghiệm được phép thiết đặt `NoOutputOfPriorPicsFlag` ở 1 trong trường hợp này.

– Nếu không, `NoOutputOfPriorPicsFlag` được thiết đặt bằng `no_output_of_prior_pics_flag`.

2. Giá trị của `NoOutputOfPriorPicsFlag` được dẫn xuất đối với bộ giải mã theo thử nghiệm được áp dụng cho bộ giải mã tham chiếu giả định (hypothetical reference decoder, HRD), sao cho khi giá trị của `NoOutputOfPriorPicsFlag` bằng 1, tất cả các bộ đệm lưu trữ hình ảnh trong DPB được làm trống mà không kết xuất các hình ảnh mà chúng chứa, và giá trị đầy DPB được thiết đặt bằng 0.

Điều C.5.2.2 (Kết xuất và loại bỏ các hình ảnh từ DPB) của Phác thảo VVC 5 bao gồm nội dung sau đây:

– Nếu hình ảnh hiện hành là hình ảnh IRAP với `NoIncorrectPicOutputFlag` bằng 1 mà không phải là hình ảnh 0, các bước có thứ tự sau đây được áp dụng:

1. Biến số `NoOutputOfPriorPicsFlag` được dẫn xuất đối với bộ giải mã theo thử nghiệm như sau:

– Nếu hình ảnh hiện hành là hình ảnh CRA, `NoOutputOfPriorPicsFlag` được thiết đặt bằng 1 (mà không quan tâm đến giá trị của `no_output_of_prior_pics_flag`).

– Nếu không, nếu giá trị của `pic_width_in_luma_samples`, `pic_height_in_luma_samples`, `chroma_format_idc`, `separate_colour_plane_flag`, `bit_depth_luma_minus8`, `bit_depth_chroma_minus8` hoặc `sps_max_dec_pic_buffering_minus1[ HighestTid ]` được dẫn xuất từ SPS hoạt động khác với giá trị của `pic_width_in_luma_samples`, `pic_height_in_luma_samples`, `chroma_format_idc`, `separate_colour_plane_flag`, `bit_depth_luma_minus8`, `bit_depth_chroma_minus8` hoặc `sps_max_dec_pic_buffering_minus1[ HighestTid ]`, một cách tương ứng, được dẫn xuất từ SPS hoạt động đối với hình ảnh trước, `NoOutputOfPriorPicsFlag` có thể (nhưng không nên) được thiết đặt ở 1 bởi bộ giải mã theo thử nghiệm, mà không quan tâm đến giá trị của `no_output_of_prior_pics_flag`.

Lưu ý – Mặc dù việc thiết đặt `NoOutputOfPriorPicsFlag` bằng `no_output_of_prior_pics_flag` được ưu tiên theo các điều kiện này, nhưng bộ giải mã theo thử nghiệm được phép thiết đặt `NoOutputOfPriorPicsFlag` ở 1 trong trường hợp này.

– Nếu không, `NoOutputOfPriorPicsFlag` được thiết đặt bằng `no_output_of_prior_pics_flag`.

2. Giá trị của `NoOutputOfPriorPicsFlag` được dẫn xuất đối với bộ giải mã theo thử nghiệm được áp dụng cho HRD như sau:

– Nếu `NoOutputOfPriorPicsFlag` bằng 1, tất cả các bộ đệm lưu trữ hình ảnh trong DPB được làm trống mà không kết xuất các hình ảnh mà chúng chứa, và giá trị đầy DPB được thiết đặt bằng 0.

– Nếu không (`NoOutputOfPriorPicsFlag` bằng 0), tất cả các bộ đệm lưu trữ hình ảnh chứa hình ảnh mà được đánh dấu là “không cần để kết xuất” và “không được sử dụng để tham chiếu” được làm trống (mà không kết xuất) và tất cả các bộ đệm lưu trữ hình ảnh không làm trống trong DPB được làm trống bằng cách gọi lặp lại quy trình “bó” (“bumping”) được quy định ở điều C.5.2.4 và giá trị đầy DPB được thiết đặt bằng 0.

Các vấn đề về thiết kế tồn tại được thảo luận.

Trong đặc tả phác thảo mới nhất của VVC, đối với hình ảnh CRA với

NoIncorrectPicOutputFlag bằng 1 (tức là, hình ảnh CRA mà bắt đầu CVS mới), giá trị của `no_output_of_prior_pics_flag` không được sử dụng, do giá trị của NoOutputOfPriorPicsFlag được thiết đặt bằng 1 mà không quan tâm đến giá trị của `no_output_of_prior_pics_flag`. Điều đó có nghĩa là, các hình ảnh trước của mỗi hình ảnh CRA bắt đầu CVS không được kết xuất. Tuy nhiên, tương tự đối với hình ảnh IDR, việc kết xuất/hiển thị các hình ảnh trước có thể mang lại việc phát lại liên tục hơn và vì vậy dẫn đến trải nghiệm tốt hơn của người dùng miễn là DPB không bị tràn khi giải mã hình ảnh bắt đầu CVS mới và các hình ảnh sau đó theo thứ tự giải mã.

Để giải quyết các vấn đề được thảo luận trên đây, sáng chế đề xuất khía cạnh sáng tạo sau đây. Giá trị của `no_output_of_prior_pics_flag` được sử dụng trong đặc tả của việc kết xuất các hình ảnh trước đối với mỗi hình ảnh CRA mà bắt đầu CVS mới và không phải là hình ảnh thứ nhất trong dòng bit. Điều này cho phép phát lại liên tục hơn và vì vậy dẫn đến trải nghiệm tốt hơn của người dùng.

Sáng chế cũng áp dụng các loại hình ảnh khác mà bắt đầu CVS mới, ví dụ, hình ảnh truy cập ngẫu nhiên từng bước (GRA) như hiện được quy định trong đặc tả phác thảo VVC mới nhất. Theo một phương án, hình ảnh GRA có thể được tham chiếu đến hoặc đồng nghĩa với hình ảnh GDR.

Ví dụ, khi giải mã dòng bit video, cờ tương ứng với hình ảnh truy cập ngẫu nhiên sạch (CRA) được báo hiệu trong dòng bit. Cờ này quy định liệu các hình ảnh được giải mã trong bộ đệm hình ảnh được giải mã mà được giải mã sớm hơn hình ảnh CRA được kết xuất khi hình ảnh CRA bắt đầu chuỗi video được tạo mã mới. Điều này có nghĩa là, khi giá trị của cờ chỉ báo rằng các hình ảnh trước được kết xuất (ví dụ, khi giá trị bằng 0), thì kết xuất các hình ảnh trước. Theo một phương án, cờ này được chỉ định là `no_output_of_prior_pics_flag`.

Theo một ví dụ khác, khi giải mã dòng bit video, cờ tương ứng với hình ảnh truy cập ngẫu nhiên từng bước (GRA) được báo hiệu trong dòng bit. Cờ này quy định liệu các hình ảnh được giải mã trong bộ đệm hình ảnh được giải mã mà được giải mã sớm hơn hình ảnh GRA được kết xuất khi hình ảnh GRA bắt đầu chuỗi video được tạo mã mới. Điều này có nghĩa là, khi giá trị của cờ chỉ báo rằng các hình ảnh trước được kết xuất (ví dụ, khi giá trị bằng 0), thì kết xuất các hình ảnh trước. Theo một phương án, cờ

này được chỉ định là `no_output_of_prior_pics_flag`.

Được bộc lộ ở đây là các kỹ thuật để kết xuất các hình ảnh trước (ví dụ, các hình ảnh được giải mã trước đó) trong bộ đệm hình ảnh được giải mã (DPB) khi hình ảnh điểm truy cập ngẫu nhiên (ví dụ, hình ảnh truy cập ngẫu nhiên sạch (clean random access, CRA), hình ảnh truy cập ngẫu nhiên từng bước (gradual random access, GRA), hoặc hình ảnh làm mới giải mã từng bước (GDR), hình ảnh CVSS, v.v.) hơn là là hình ảnh làm mới bộ giải mã tức thời (instantaneous decoder refresh, IDR) được gặp phải theo thứ tự giải mã. Việc làm trống các hình ảnh được giải mã trước đó từ DPB khi hình ảnh điểm truy cập ngẫu nhiên được đạt đến ngăn không cho DPB bị tràn và xúc tiến phát lại liên tục hơn. Vì vậy, bộ tạo mã / bộ giải mã (cũng được gọi là “codec”) trong tạo mã video được nâng cao so với các codec hiện hành. Như một vấn đề thực tế, quy trình tạo mã video được nâng cao đem lại cho người dùng trải nghiệm người dùng tốt hơn khi các video được gửi, được nhận, và/hoặc được xem.

Fig.8 thể hiện một phương án của phương pháp 800 để giải mã dòng bit video được tạo mã được thực hiện bởi bộ giải mã video (ví dụ, bộ giải mã video 30). Phương pháp 800 có thể được thực hiện sau khi dòng bit được giải mã đã được nhận trực tiếp hoặc gián tiếp từ bộ mã hóa video (ví dụ, bộ mã hóa video 20). Phương pháp 800 cải tiến quy trình giải mã bằng cách làm trống DPB trước khi hình ảnh hiện hành được giải mã khi hình ảnh điểm truy cập ngẫu nhiên được gặp phải. Phương pháp 800 ngăn không cho DPB bị tràn và xúc tiến phát lại liên tục hơn. Do đó, như một vấn đề thực tế, hiệu năng của codec được cải thiện, mà dẫn đến trải nghiệm tốt hơn của người dùng.

Trong khối 802, bộ giải mã video nhận dòng bit video được tạo mã (ví dụ, dòng bit 750). Dòng bit video được tạo mã chứa hình ảnh làm mới giải mã từng bước (GDR) và cờ thứ nhất có giá trị thứ nhất. Theo một phương án, hình ảnh GDR không phải là hình ảnh thứ nhất của dòng bit video được tạo mã. Theo một phương án, cờ thứ nhất được chỉ định là `no_output_of_prior_pics_flag`. Theo một phương án, hình ảnh GDR được bố trí trong đơn vị lớp trừu tượng mạng (NAL) của lớp tạo mã video (VCL) có loại đơn vị lớp trừu tượng mạng (NAL) làm mới giải mã từng bước (GDR) (GDR_NUT).

Ở khối 804, bộ giải mã video thiết đặt giá trị thứ hai của cờ thứ hai bằng giá trị thứ nhất của cờ thứ nhất. Theo một phương án, cờ thứ hai được chỉ định là

NoOutputOfPriorPicsFlag. Theo một phương án, cờ thứ hai ở bên trong bộ giải mã.

Ở khối 806, bộ giải mã video làm trống các hình ảnh được giải mã trước đó bất kỳ tương ứng với hình ảnh GDR từ DPB dựa trên cờ thứ hai có giá trị thứ hai. Theo một phương án, các hình ảnh được giải mã trước đó được làm trống từ DPB sau khi hình ảnh GDR đã được giải mã. Điều này có nghĩa là, bộ giải mã video loại bỏ các hình ảnh được giải mã trước đó từ các bộ đệm lưu trữ hình ảnh trong DPB. Theo một phương án, các hình ảnh được giải mã trước đó không được kết xuất hoặc được hiển thị khi các hình ảnh được giải mã trước đó được loại bỏ khỏi DPB. Theo một phương án, thông số đầy DPB được thiết đặt ở zero khi cờ thứ nhất được thiết đặt ở giá trị thứ nhất. Thông số đầy DPB chỉ báo có bao nhiêu hình ảnh được giữ trong DPB. Việc thiết đặt thông số đầy DPB ở không (zero) biểu thị rằng DPB là trống.

Ở khối 808, bộ giải mã video giải mã hình ảnh hiện hành sau khi DPB đã được làm trống. Theo một phương án, hình ảnh hiện hành là từ cùng một CVS với hình ảnh CRA và được thu nhận hoặc được gộp phải sau CRA theo thứ tự giải mã. Theo một phương án, ảnh được tạo ra dựa trên hình ảnh hiện hành được hiển thị cho người dùng thiết bị điện tử (ví dụ, điện thoại thông minh, máy tính bảng, máy tính xách tay loại laptop, máy tính cá nhân, v.v.).

Fig.9 thể hiện một phương án của phương pháp 900 để mã hóa dòng bit video được thực hiện bởi bộ mã hóa video (ví dụ, bộ mã hóa video 20). Phương pháp 900 có thể được thực hiện khi hình ảnh (ví dụ, từ video) cần được mã hóa thành dòng bit video và sau đó được truyền đến bộ giải mã video (ví dụ, bộ giải mã video 30). Phương pháp 900 cải tiến quy trình mã hóa bằng cách ra lệnh cho bộ giải mã video làm trống DPB trước khi hình ảnh hiện hành được giải mã khi hình ảnh điểm truy cập ngẫu nhiên được gộp phải. Phương pháp 900 ngăn không cho DPB bị tràn và xúc tiến phát lại liên tục hơn. Do đó, như một vấn đề thực tế, hiệu năng của codec được cải thiện, mà dẫn đến trải nghiệm tốt hơn của người dùng.

Ở khối 902, bộ mã hóa video xác định điểm truy cập ngẫu nhiên cho chuỗi video. Ở khối 904, bộ mã hóa video mã hóa hình ảnh làm mới giải mã từng bước (GDR) thành chuỗi video tại điểm truy cập ngẫu nhiên. Theo một phương án, hình ảnh GDR không phải là hình ảnh thứ nhất của dòng bit video. Theo một phương án, hình ảnh GDR được

bố trí trong đơn vị lớp trừu tượng mạng (NAL) của lớp tạo mã video (VCL) có loại đơn vị lớp trừu tượng mạng (NAL) làm mới giải mã từng bước (GDR) (GDR_NUT).

Ở khối 906, bộ mã hóa video thiết đặt cờ ở giá trị thứ nhất để ra lệnh cho bộ giải mã video làm trống các hình ảnh được giải mã trước đó bất kỳ từ bộ đệm hình ảnh được giải mã (DPB). Theo một phương án, bộ giải mã video được ra lệnh để làm trống các hình ảnh được giải mã trước đó bất kỳ từ DPB sau khi hình ảnh GDR đã được giải mã. Theo một phương án, cờ này được chỉ định là `no_output_of_prior_pics_flag`. Theo một phương án, bộ mã hóa video ra lệnh cho bộ giải mã video thiết đặt thông số đầy DPB ở không (zero) khi cờ này được thiết đặt ở giá trị thứ nhất. Theo một phương án, giá trị thứ nhất của cờ là một.

Ở khối 908, bộ mã hóa video tạo ra dòng bit video chứa chuỗi video có hình ảnh GDR tại điểm truy cập ngẫu nhiên và cờ. Ở khối 910, bộ mã hóa video lưu trữ dòng bit video để truyền đến bộ giải mã video.

Cú pháp và các ngữ nghĩa sau đây có thể được sử dụng để thực hiện các phương án được bộc lộ ở đây. Phần mô tả sau đây có liên quan đến văn bản cơ bản, mà là đặc tả phác thảo VVC mới nhất. Nói cách khác, chỉ delta được mô tả, trong khi văn bản trong văn bản cơ bản mà không được nêu dưới đây áp dụng đúng như chính nó. Phần văn bản bổ sung vào văn bản cơ bản được thể hiện đậm, và phần văn bản được loại bỏ được thể hiện nghiêng.

Cú pháp phần đầu lát chung (7.3.5.1 trong VVC).

<code>slice_header()</code> {	Bộ mô tả
...	
<code>if(NalUnitType == IDR_W_RADL NalUnitType == IDR_N_LP NalUnitType == CRA_NUT NalUnitType == GRA_NUT)</code>	
<code>no_output_of_prior_pics_flag</code>	u(1)
...	

Các ngữ nghĩa phần đầu lát chung (7.4.6.1 trong VVC).

Khi có mặt, giá trị của mỗi trong số các phần tử cú pháp phần đầu lát

`slice_pic_parameter_set_id`, `slice_pic_order_cnt_lsb`, **`no_output_of_prior_pics_flag`**, và `slice_temporal_mvp_enabled_flag` sẽ là giống nhau trong tất cả các phần đầu lát của hình ảnh được tạo mã.

...

`no_output_of_prior_pics_flag` tác động đến đầu ra của các hình ảnh được giải mã trước đó trong bộ đệm hình ảnh được giải mã sau khi giải mã hình ảnh IDR, hình ảnh CVSS mà không phải là hình ảnh thứ nhất trong dòng bit như được quy định trong Phụ lục C.

...

Loại bỏ các hình ảnh từ DPB trước khi giải mã hình ảnh hiện hành (C.3.2 trong VVC).

...

– Khi hình ảnh hiện hành là *hình ảnh IRAP* với *NoIncorrectPicOutputFlag* bằng 1, hình ảnh CVSS mà không phải là hình ảnh 0, các bước có thứ tự sau đây được áp dụng:

1. Biến số `NoOutputOfPriorPicsFlag` được dẫn xuất đối với bộ giải mã theo thử nghiệm như sau:

– Nếu hình ảnh hiện hành là hình ảnh CRA, *NoOutputOfPriorPicsFlag* được thiết đặt bằng 1 (mà không quan tâm đến giá trị của `no_output_of_prior_pics_flag`).

– Nếu không, nếu giá trị của `pic_width_in_luma_samples`, `pic_height_in_luma_samples`, `chroma_format_idc`, `separate_colour_plane_flag`, `bit_depth_luma_minus8`, `bit_depth_chroma_minus8` hoặc `sps_max_dec_pic_buffering_minus1[ HighestTid ]` được dẫn xuất từ SPS hoạt động khác với giá trị của `pic_width_in_luma_samples`, `pic_height_in_luma_samples`, `chroma_format_idc`, `separate_colour_plane_cờ`, `bit_depth_luma_minus8`, `bit_depth_chroma_minus8` hoặc `sps_max_dec_pic_buffering_minus1[ HighestTid ]`, một cách tương ứng, được dẫn xuất từ SPS hoạt động đối với hình ảnh trước, `NoOutputOfPriorPicsFlag` có thể (nhưng không nên) được thiết đặt ở 1 bởi bộ giải mã theo thử nghiệm, mà không quan tâm đến giá trị của `no_output_of_prior_pics_flag`.

Lưu ý – Mặc dù việc thiết đặt NoOutputOfPriorPicsFlag bằng no_output_of_prior_pics_flag được ưu tiên theo các điều kiện này, nhưng bộ giải mã theo thử nghiệm được phép thiết đặt NoOutputOfPriorPicsFlag ở 1 trong trường hợp này.

– Nếu không, NoOutputOfPriorPicsFlag được thiết đặt bằng no_output_of_prior_pics_flag.

2. Giá trị của NoOutputOfPriorPicsFlag được dẫn xuất đối với bộ giải mã theo thử nghiệm được áp dụng cho bộ giải mã tham chiếu giả định (HRD), sao cho khi giá trị của NoOutputOfPriorPicsFlag bằng 1, tất cả các bộ đệm lưu trữ hình ảnh trong DPB được làm trống mà không kết xuất các hình ảnh mà chúng chứa, và giá trị đầy DPB được thiết đặt bằng 0.

Fig.10 là hình vẽ giản lược của thiết bị tạo mã video 1000 (ví dụ, bộ mã hóa video 20 hoặc bộ giải mã video 30) theo một phương án của sáng chế. Thiết bị tạo mã video 1000 thích hợp để thực hiện các phương án được bộc lộ như được mô tả ở đây. Thiết bị tạo mã video 1000 bao gồm các cổng vào 1010 và các bộ phận nhận (Rx) 1020 để nhận dữ liệu; bộ xử lý, bộ phận logic, hoặc bộ xử lý trung tâm (CPU) 1030 để xử lý dữ liệu; các bộ phận truyền (Tx) 1040 và các cổng ra 1050 để truyền dữ liệu; và bộ nhớ 1060 để lưu trữ dữ liệu. Thiết bị tạo mã video 1000 cũng có thể bao gồm các thành phần quang sang điện (optical-to-electrical, OE) và các thành phần điện sang quang (electrical-to-optical, EO) được ghép nối với các cổng vào 1010, các bộ phận nhận 1020, các bộ phận truyền 1040, và các cổng ra 1050 để lấy ra hoặc lấy vào các tín hiệu điện hoặc quang.

Bộ xử lý 1030 được thực thi bởi phần cứng và phần mềm. Bộ xử lý 1030 có thể được thực thi như là một hoặc nhiều chip CPU, nhân (ví dụ, bộ xử lý nhiều nhân), mảng cổng lập trình được dạng trường (field-programmable gate array, FPGA), mạch tích hợp chuyên dụng (application specific integrated circuit, ASIC), và bộ xử lý tín hiệu số (digital signal processor, DSP). Bộ xử lý 1030 truyền thông với các cổng vào 1010, các bộ phận nhận 1020, các bộ phận truyền 1040, các cổng ra 1050, và bộ nhớ 1060. Bộ xử lý 1030 bao gồm môđun tạo mã 1070. Môđun tạo mã 1070 thực thi các phương án được bộc lộ được mô tả trên đây. Ví dụ, môđun tạo mã 1070 thực hiện, xử lý, chuẩn bị, hoặc

cung cấp các chức năng codec khác nhau. Việc bao gồm môđun tạo mã 1070 do đó sẽ mang lại sự cải tiến quan trọng cho chức năng của thiết bị tạo mã video 1000 và đem lại sự biến đổi thiết bị tạo mã video 1000 sang trạng thái khác. Theo cách khác, môđun tạo mã 1070 được thực thi dưới dạng các lệnh được lưu trữ trong bộ nhớ 1060 và được thực hiện bởi bộ xử lý 1030.

Thiết bị tạo mã video 1000 cũng có thể bao gồm các thiết bị đầu vào và/hoặc đầu ra (I/O) 1080 để truyền thông dữ liệu đến và từ người dùng. Các thiết bị I/O 1080 có thể bao gồm thiết bị đầu ra như bộ hiển thị để hiển thị dữ liệu video, loa để kết xuất dữ liệu audio, v.v.. Các thiết bị I/O 1080 cũng có thể bao gồm các thiết bị đầu vào, như bàn phím, chuột, bi xoay, v.v., và/hoặc các giao diện tương ứng để tương tác với thiết bị đầu ra đó.

Bộ nhớ 1060 bao gồm một hoặc nhiều đĩa, ổ băng, và ổ trạng thái rắn và có thể được sử dụng làm thiết bị lưu trữ dữ liệu tràn, để lưu trữ các chương trình khi các chương trình này được chọn để thực hiện, và để lưu trữ các lệnh và dữ liệu được đọc trong suốt quá trình thực hiện chương trình. Bộ nhớ 1060 có thể là khả biến và/hoặc bất khả biến và có thể là bộ nhớ chỉ đọc (ROM), bộ nhớ truy cập ngẫu nhiên (RAM), bộ nhớ khả lập địa chỉ nội dung bậc ba (ternary content-addressable memory, TCAM), và/hoặc bộ nhớ truy cập ngẫu nhiên tĩnh (SRAM).

Fig.11 là hình vẽ giản lược của một phương án của phương tiện tạo mã 1100. Theo một phương án, phương tiện tạo mã 1100 được thực thi trong thiết bị tạo mã video 1102 (ví dụ, bộ mã hóa video 20 hoặc bộ giải mã video 30). Thiết bị tạo mã video 1102 bao gồm phương tiện nhận 1101. Phương tiện nhận 1101 được tạo cấu hình để nhận hình ảnh để mã hóa hoặc để nhận dòng bit để giải mã. Thiết bị tạo mã video 1102 bao gồm phương tiện truyền 1107 được ghép nối với phương tiện nhận 1101. Phương tiện truyền 1107 được tạo cấu hình để truyền dòng bit đến bộ giải mã hoặc để truyền ảnh đã giải mã đến phương tiện hiển thị (ví dụ, một trong số các thiết bị I/O 1080).

Thiết bị tạo mã video 1102 bao gồm phương tiện lưu trữ 1103. Phương tiện lưu trữ 1103 được ghép nối với ít nhất một phương tiện trong số phương tiện nhận 1101 hoặc phương tiện truyền 1107. Phương tiện lưu trữ 1103 được tạo cấu hình để lưu trữ các lệnh. Thiết bị tạo mã video 1102 cũng bao gồm phương tiện xử lý 1105. Phương tiện

xử lý 1105 được ghép nối với phương tiện lưu trữ 1103. Phương tiện xử lý 1105 được tạo cấu hình để thực hiện các lệnh được lưu trữ trong phương tiện lưu trữ 1103 để thực hiện các phương pháp được bộc lộ ở đây.

Cũng cần hiểu rằng các bước của các phương pháp làm ví dụ được nêu ở đây không nhất thiết phải được thực hiện theo thứ tự được mô tả, và thứ tự của các bước của các phương pháp này nên được hiểu là chỉ làm ví dụ. Tương tự, các bước bổ sung có thể có trong các phương pháp này, và các bước nhất định có thể được bỏ qua hoặc được kết hợp, trong các phương pháp phù hợp với các phương án khác nhau của sáng chế.

Mặc dù một số phương án đã được đề xuất trong sáng chế, nhưng cần hiểu rằng các hệ thống và các phương pháp được bộc lộ có thể có trong nhiều dạng cụ thể khác mà không nằm ngoài mục đích hoặc phạm vi của sáng chế. Các ví dụ này cần được coi là minh họa và không phải giới hạn, và sáng chế không bị giới hạn ở các nội dung chi tiết được đưa ra ở đây. Ví dụ, các phần tử hoặc các thành phần khác nhau có thể được kết hợp hoặc được tích hợp trong hệ thống khác hoặc các dấu hiệu nhất định có thể được bỏ qua, hoặc không được thực hiện.

Ngoài ra, các kỹ thuật, các hệ thống, các hệ thống con, và các phương pháp được mô tả và được minh họa trong các phương án khác nhau rời rạc hoặc riêng rẽ có thể được kết hợp hoặc được tích hợp với các hệ thống, các môđun, các kỹ thuật, hoặc các phương pháp khác mà không nằm ngoài phạm vi của sáng chế. Các mục được thể hiện hoặc thảo luận là được ghép nối hoặc được ghép nối trực tiếp hoặc truyền thông với nhau có thể được ghép nối gián tiếp hoặc truyền thông qua giao diện, thiết bị, hoặc thành phần trung gian nào đó bằng điện, cơ học, hoặc loại khác. Các ví dụ khác về các sự thay đổi, thay thế, và cải biến có thể biết được bởi người có hiểu biết trung bình trong lĩnh vực tương ứng và có thể được tạo ra mà không nằm ngoài mục đích và phạm vi được bộc lộ ở đây.

YÊU CẦU BẢO HỘ

1. Phương pháp giải mã được thực hiện bởi bộ giải mã video, phương pháp này bao gồm các bước:

nhận, bởi bộ giải mã video, dòng bit video được tạo mã, trong đó dòng bit video được tạo mã chứa hình ảnh làm mới giải mã từng bước (gradual decoding refresh, GDR) và cờ thứ nhất có giá trị thứ nhất;

thiết đặt, bởi bộ giải mã video, giá trị thứ hai của cờ thứ hai bằng giá trị thứ nhất của cờ thứ nhất;

làm trống, bởi bộ giải mã video, các hình ảnh được giải mã trước đó bất kỳ từ bộ đệm hình ảnh được giải mã (decoded picture buffer, DPB) dựa trên cờ thứ hai có giá trị thứ hai sau khi hình ảnh GDR đã được giải mã; và

giải mã, bởi bộ giải mã video, hình ảnh hiện hành sau khi DPB đã được làm trống.

2. Phương pháp theo điểm 1, trong đó hình ảnh GDR không phải là hình ảnh thứ nhất của dòng bit video được tạo mã, và trong đó giá trị thứ nhất của cờ là một.

3. Phương pháp theo điểm 1, trong đó hình ảnh GDR được bố trí trong đơn vị lớp trừu tượng mạng (network abstraction layer, NAL) của lớp tạo mã video (video coding layer, VCL) có loại đơn vị lớp trừu tượng mạng (NAL) làm mới giải mã từng bước (GDR) (GDR_NUT).

4. Phương pháp theo điểm 1, phương pháp này còn bao gồm bước thiết đặt thông số đầy DPB bằng không (zero) khi cờ thứ nhất được thiết đặt ở giá trị thứ nhất.

5. Phương pháp theo điểm 1, trong đó cờ thứ nhất được chỉ định là `no_output_of_prior_pics_flag` và cờ thứ hai được chỉ định là `NoOutputOfPriorPicsFlag`.

6. Phương pháp theo điểm 1, trong đó DPB được làm trống sau khi hình ảnh GDR đã được giải mã.

7. Phương pháp theo điểm 1, phương pháp này còn bao gồm bước hiển thị ảnh được tạo ra dựa trên hình ảnh hiện hành.

8. Phương pháp mã hóa được thực hiện bởi bộ mã hóa video, phương pháp này bao gồm các bước:

xác định, bởi bộ mã hóa video, điểm truy cập ngẫu nhiên cho chuỗi video;

mã hóa, bởi bộ mã hóa video, hình ảnh làm mới giải mã từng bước (GDR) thành chuỗi video tại điểm truy cập ngẫu nhiên;

thiết đặt, bởi bộ mã hóa video, cờ ở giá trị thứ nhất để ra lệnh cho bộ giải mã video làm trống các hình ảnh được giải mã trước đó bất kỳ từ bộ đệm hình ảnh được giải mã (decoded picture buffer, DPB);

tạo, bởi bộ mã hóa video, dòng bit video chứa chuỗi video có hình ảnh GDR tại điểm truy cập ngẫu nhiên và cờ; và

lưu trữ, bởi bộ mã hóa video, dòng bit video để truyền đến bộ giải mã video.

9. Phương pháp theo điểm 8, trong đó hình ảnh GDR không phải là hình ảnh thứ nhất của dòng bit video, và trong đó bộ giải mã video được ra lệnh để làm trống DPB sau khi hình ảnh GDR đã được giải mã.

10. Phương pháp theo điểm 8, trong đó hình ảnh GDR được bố trí trong đơn vị lớp trừu tượng mạng (network abstraction layer, NAL) của lớp tạo mã video (video coding layer, VCL) có loại đơn vị lớp trừu tượng mạng (NAL) làm mới giải mã từng bước (GDR) (GDR_NUT).

11. Phương pháp theo điểm 8, phương pháp này còn bao gồm bước ra lệnh cho bộ giải mã video thiết đặt thông số đầy DPB bằng không (zero) khi cờ này được thiết đặt ở giá trị thứ nhất.

12. Phương pháp theo điểm 8, trong đó cờ được chỉ định là `no_output_of_prior_pics_flag`.

13. Phương pháp theo điểm 8, trong đó giá trị thứ nhất của cờ này là một.

14. Thiết bị giải mã, thiết bị này bao gồm:

bộ nhận được tạo cấu hình để nhận dòng bit video được tạo mã;

bộ nhớ được ghép nối với bộ nhận, bộ nhớ lưu trữ các lệnh; và

bộ xử lý được ghép nối với bộ nhớ, bộ xử lý được tạo cấu hình để thực hiện các lệnh để làm cho thiết bị giải mã:

nhận dòng bit video được tạo mã, trong đó dòng bit video được tạo mã chứa hình ảnh làm mới giải mã từng bước (GDR) và cờ thứ nhất có giá trị thứ nhất;

thiết đặt giá trị thứ hai của cờ thứ hai bằng giá trị thứ nhất của cờ thứ nhất;

làm trống các hình ảnh được giải mã trước đó bất kỳ từ bộ đệm hình ảnh được giải mã (DPB) dựa trên cờ thứ hai có giá trị thứ hai; và

giải mã hình ảnh hiện hành sau khi DPB đã được làm trống.

15. Thiết bị giải mã theo điểm 14, trong đó hình ảnh GDR không phải là hình ảnh thứ nhất của dòng bit video được tạo mã.

16. Thiết bị giải mã theo điểm 14, trong đó cờ thứ nhất được chỉ định là `no_output_of_prior_pics_flag`, và trong đó cờ thứ hai được chỉ định là `NoOutputOfPriorPicsFlag`.

17. Thiết bị giải mã theo điểm 14, thiết bị này còn bao gồm bộ hiển thị được tạo cấu hình để hiển thị ảnh như được tạo ra dựa trên hình ảnh hiện hành.

18. Thiết bị mã hóa, thiết bị này bao gồm:

bộ nhớ chứa các lệnh;

bộ xử lý được ghép nối với bộ nhớ, bộ xử lý được tạo cấu hình để thực hiện các lệnh để làm cho thiết bị mã hóa:

xác định điểm truy cập ngẫu nhiên cho chuỗi video;

mã hóa hình ảnh làm mới giải mã từng bước (GDR) thành chuỗi video tại điểm truy cập ngẫu nhiên;

thiết đặt cờ ở giá trị thứ nhất để ra lệnh cho bộ giải mã video làm trống các hình ảnh được giải mã trước đó bất kỳ từ bộ đệm hình ảnh được giải mã (DPB); và

tạo dòng bit video chứa chuỗi video có hình ảnh GDR tại điểm truy cập ngẫu nhiên và cờ; và

bộ truyền được ghép nối với bộ xử lý, bộ truyền được tạo cấu hình để truyền dòng bit video đến bộ giải mã video.

19. Thiết bị mã hóa theo điểm 18, trong đó hình ảnh GDR không phải là hình ảnh thứ nhất của dòng bit video.

20. Thiết bị mã hóa theo điểm 18, trong đó cờ được chỉ định là `no_output_of_prior_pics_flag`.

21. Thiết bị mã hóa theo điểm 18, trong đó bộ nhớ lưu trữ dòng bit video trước khi bộ truyền truyền dòng bit video đến bộ giải mã video.

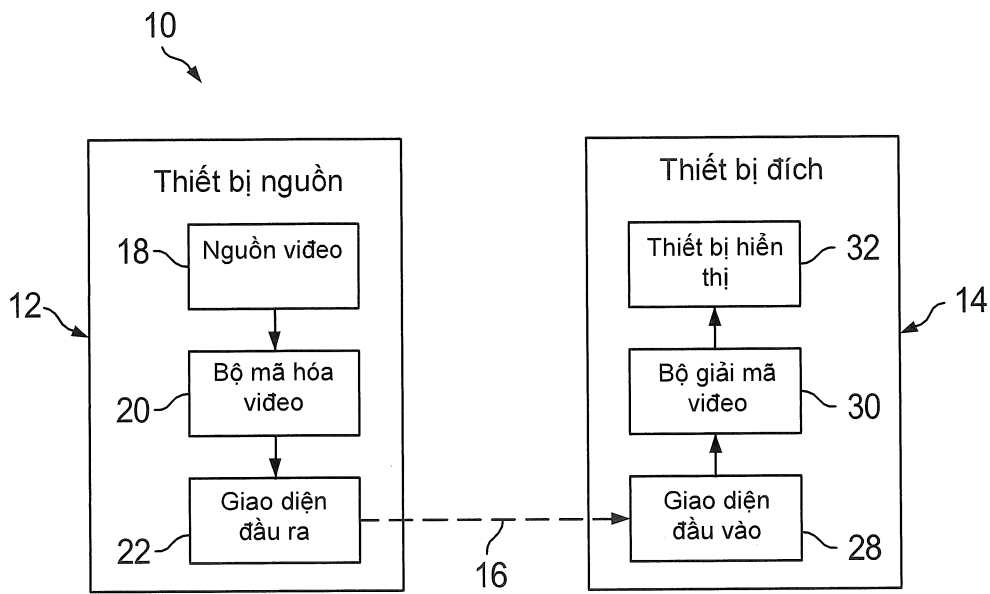


FIG. 1

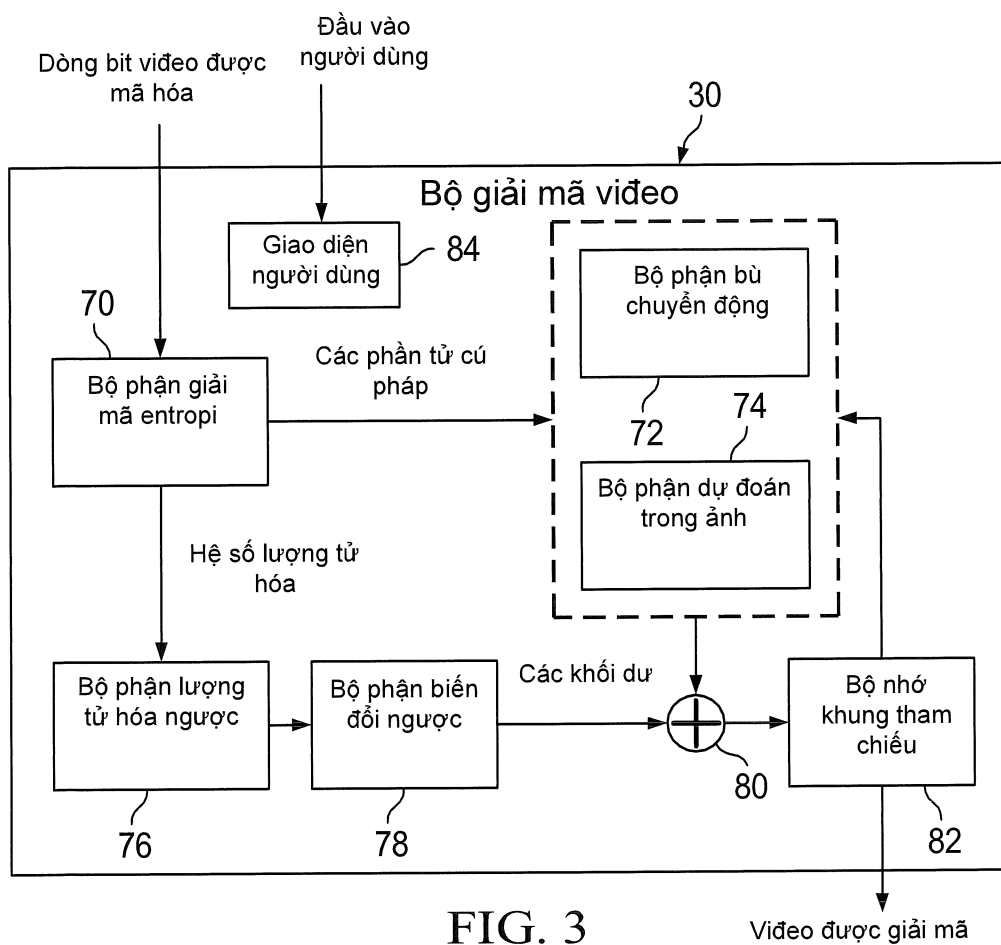
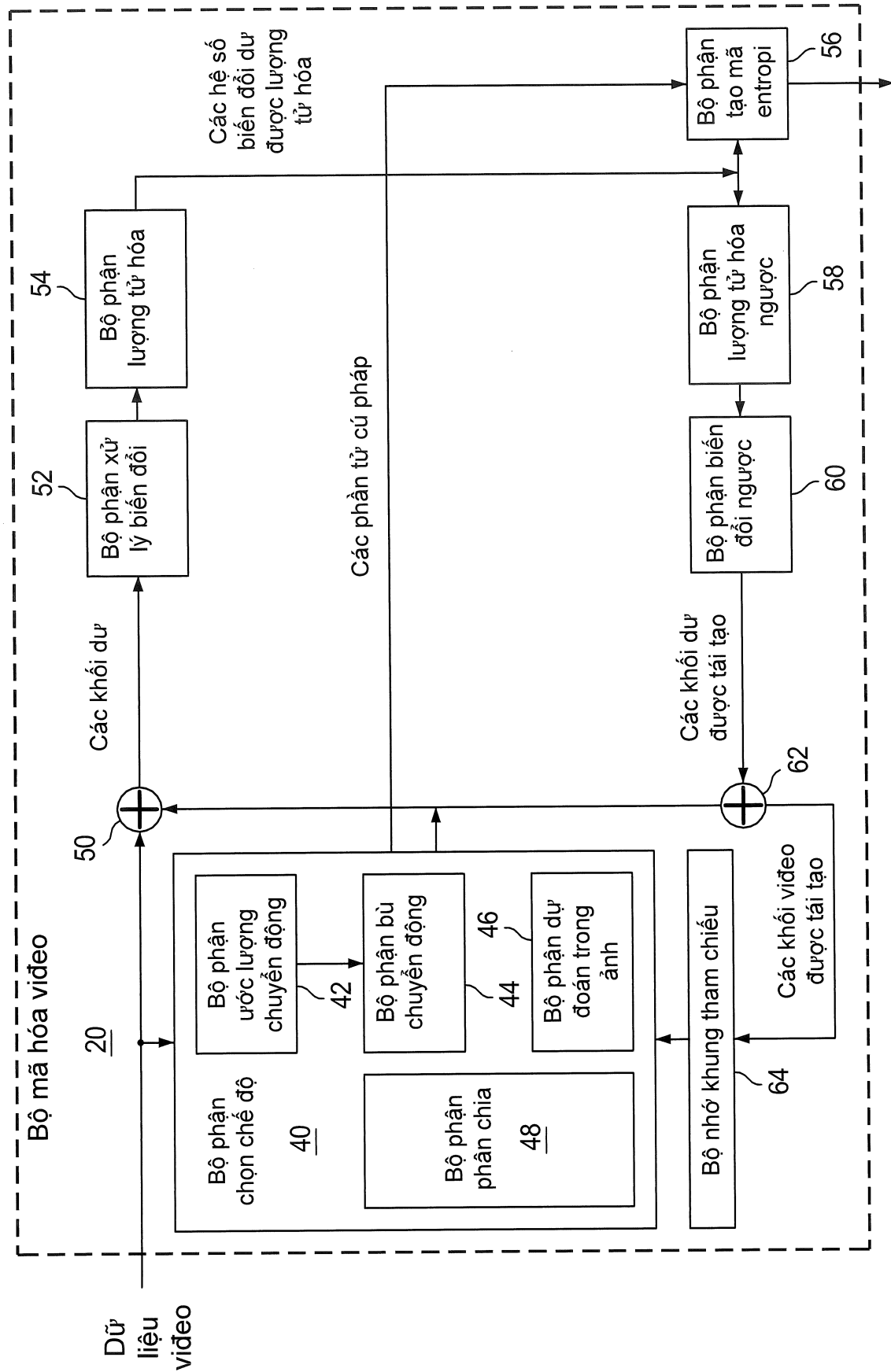


FIG. 3



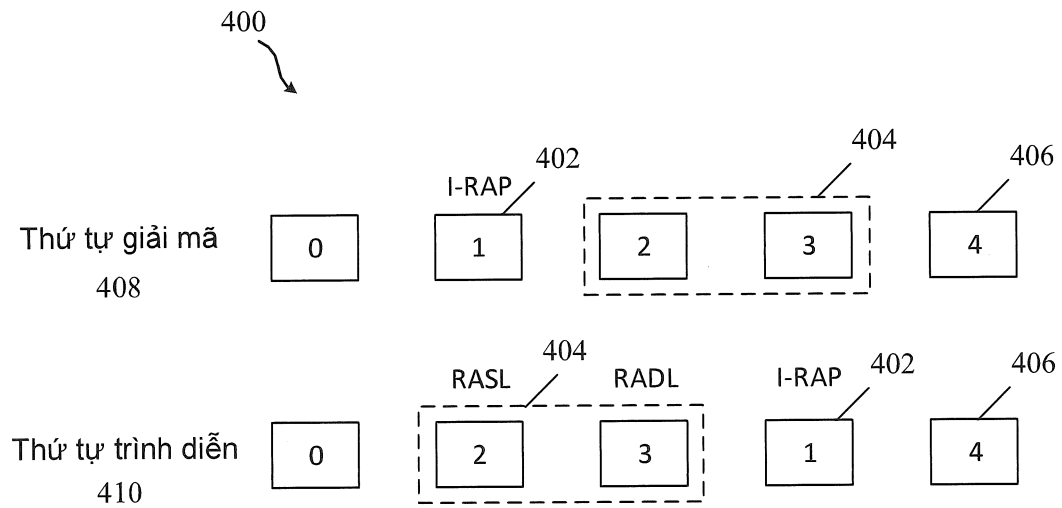


FIG. 4

4/10

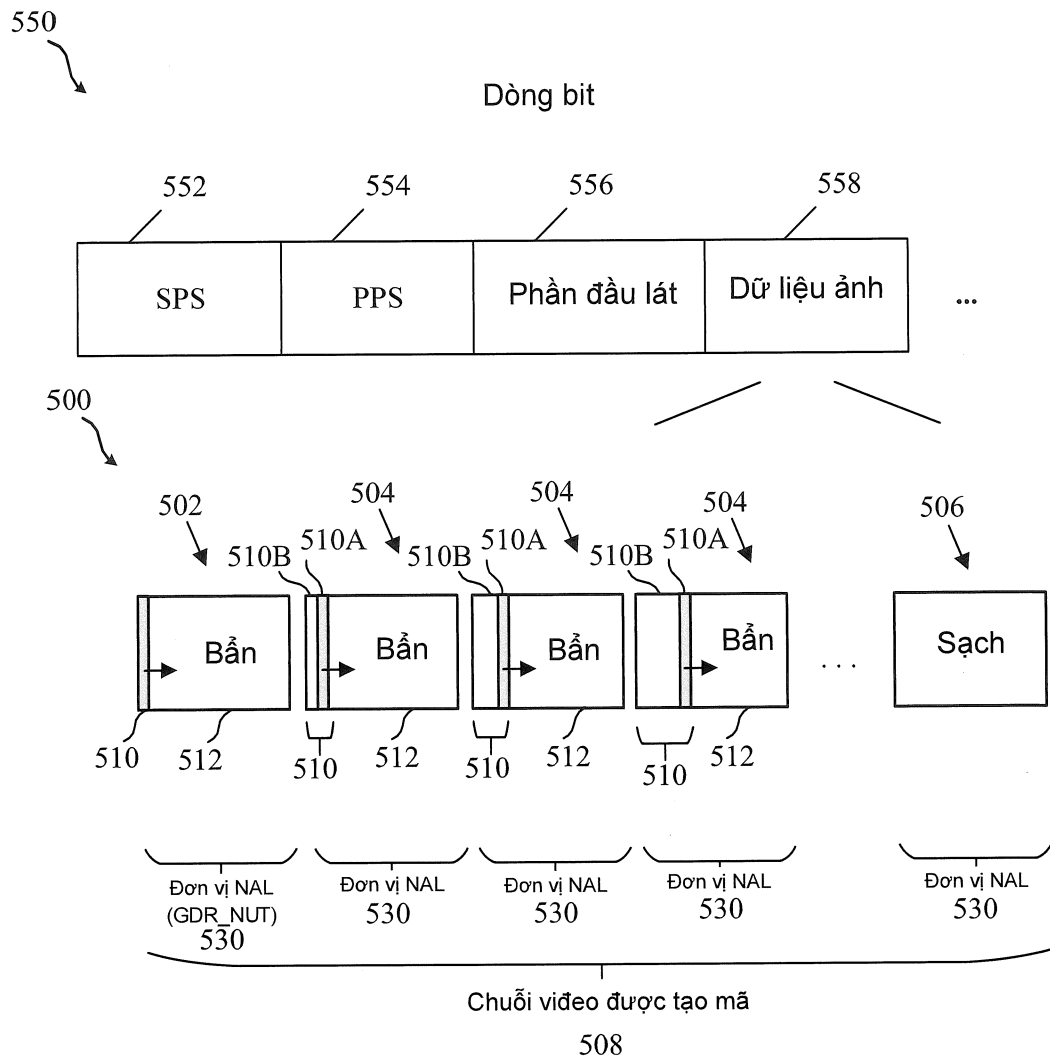


FIG. 5

5/10

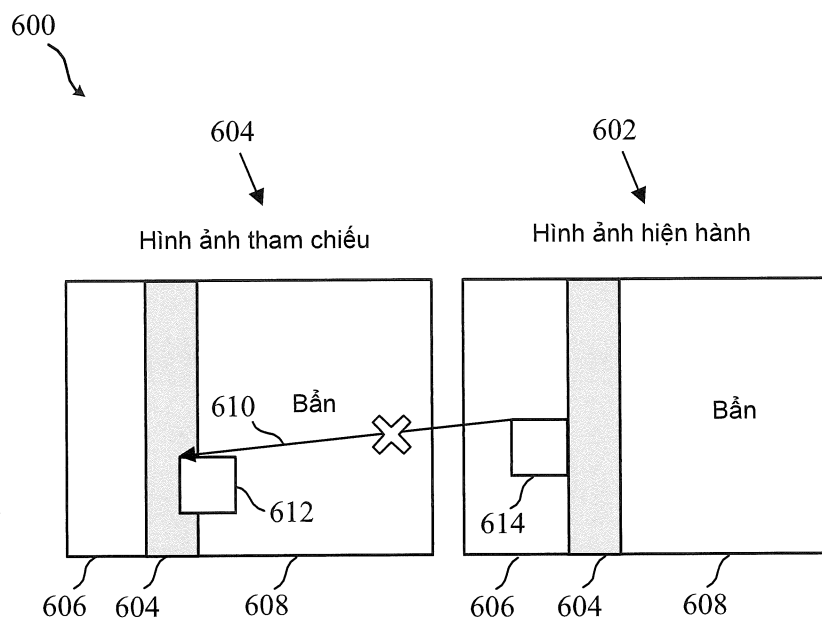


FIG. 6

6/10

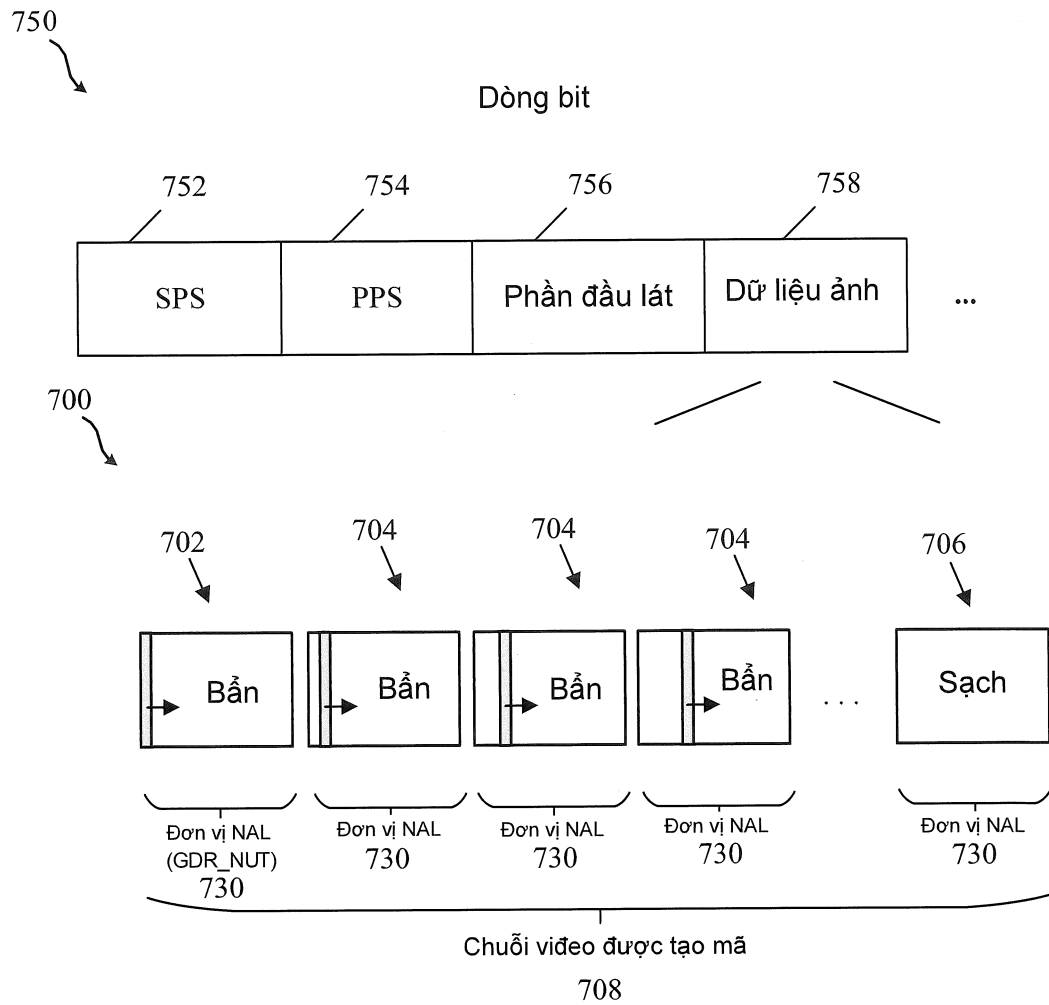


FIG. 7

7/10

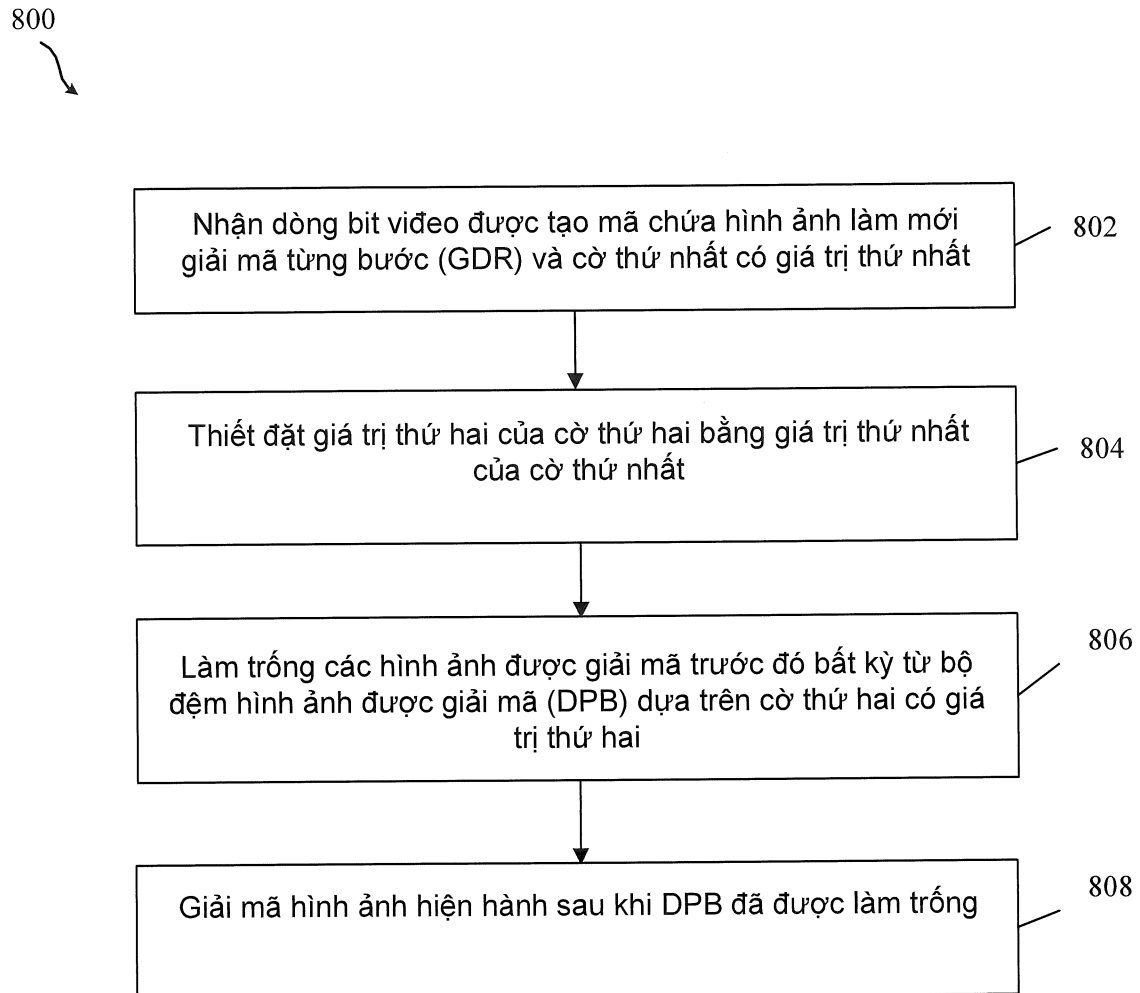


FIG. 8

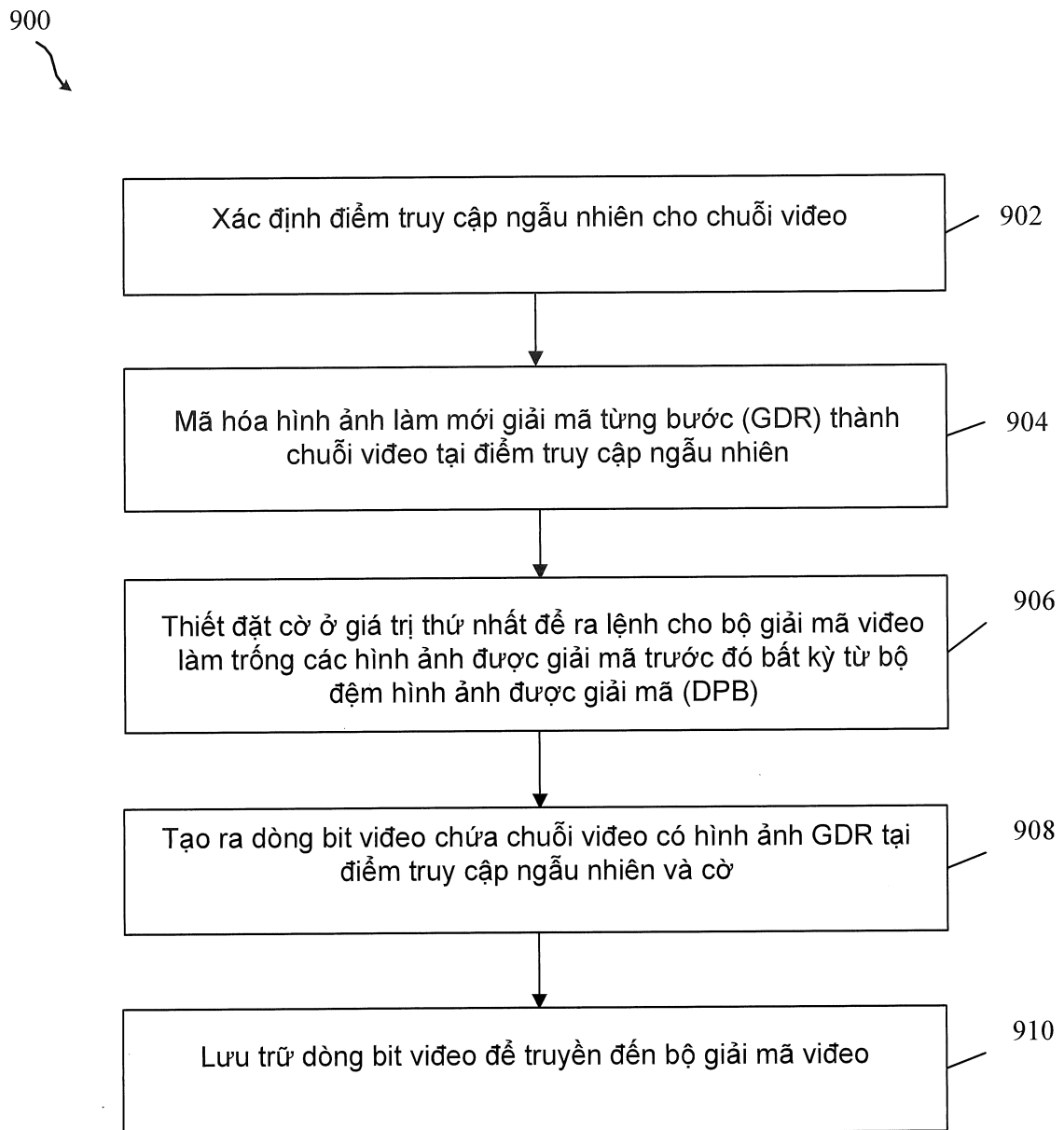


FIG. 9

9/10

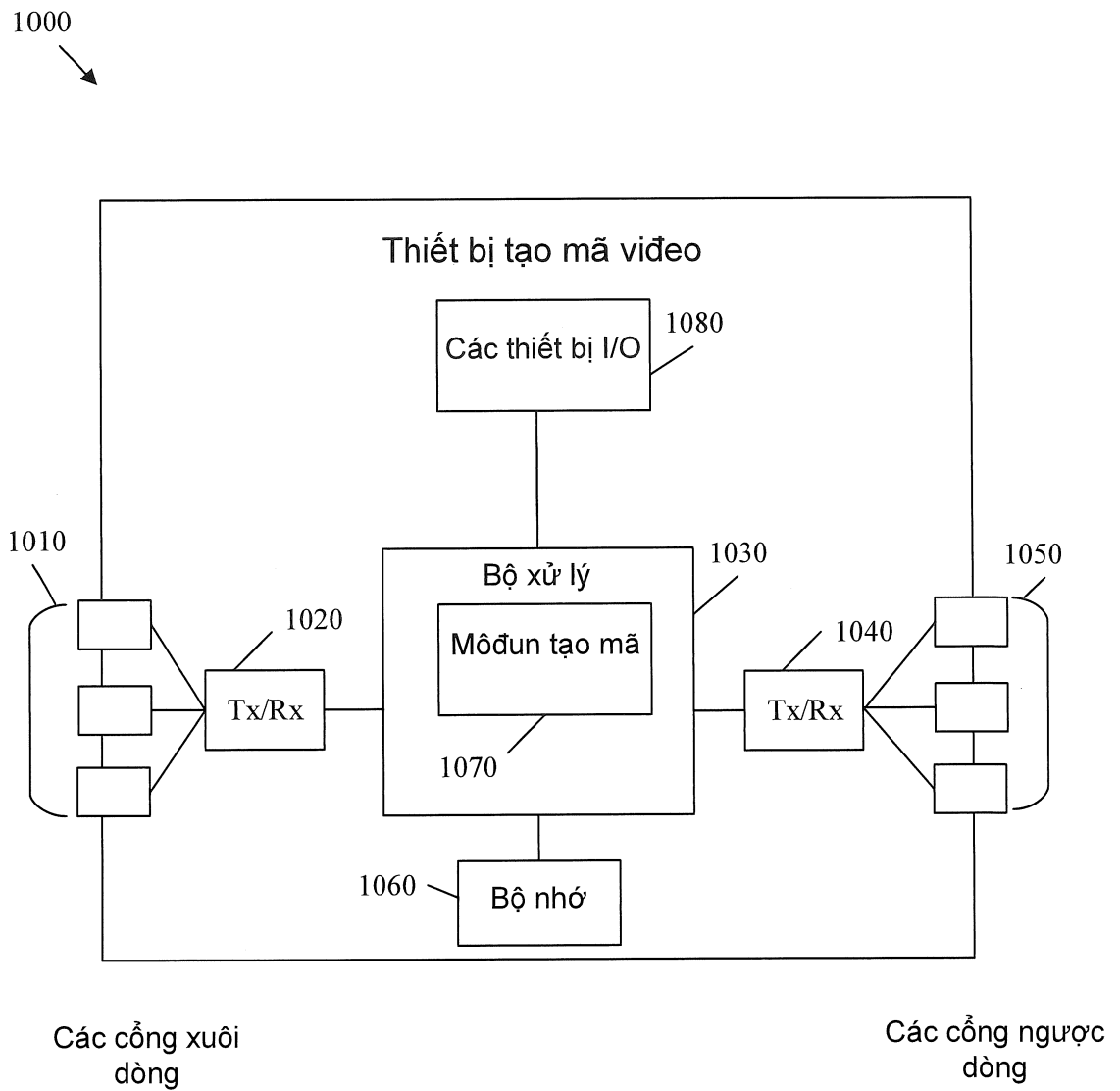


FIG. 10

10/10

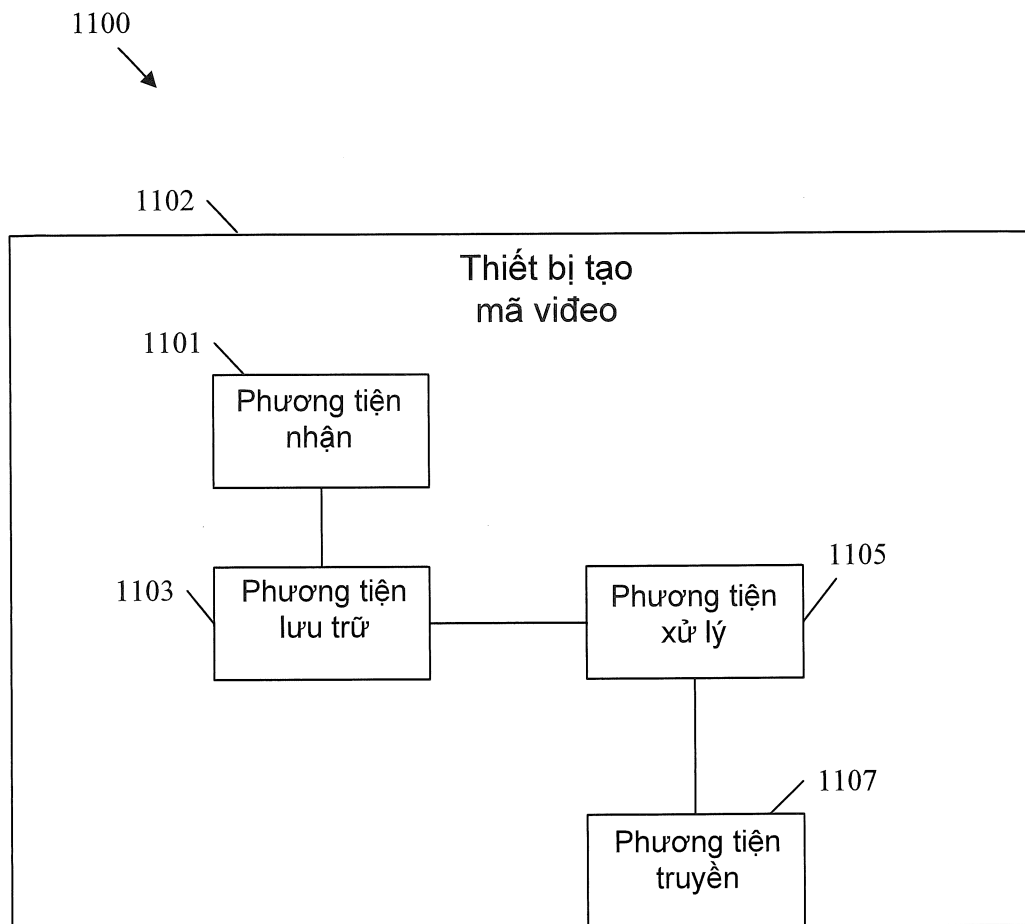


FIG. 11