

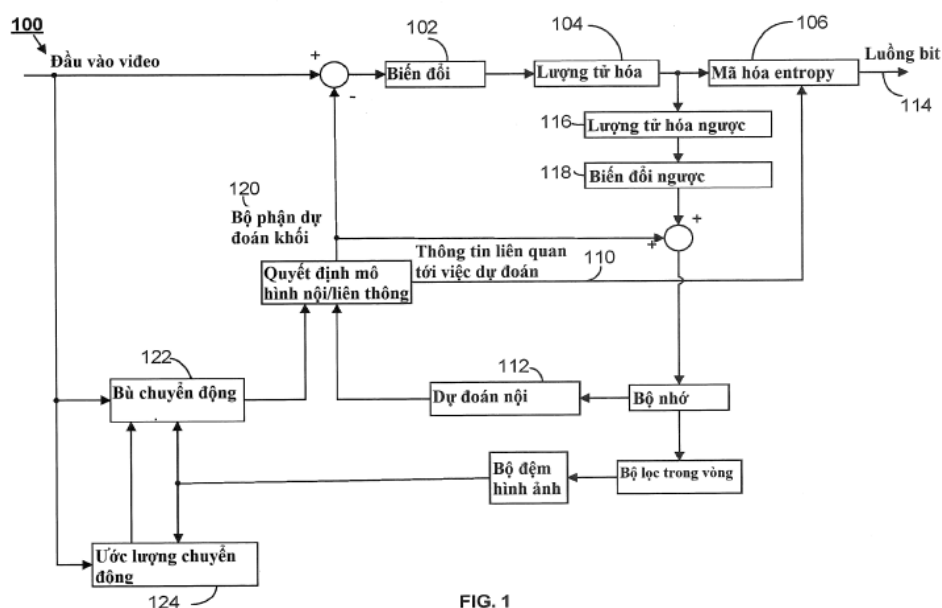


- (12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ
(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN) (11) 
CỤC SỞ HỮU TRÍ TUỆ
1-0050900
(51)^{2020.01} H04N 19/577; H04N 19/105; H04N 19/513 (13) B

-
- (21) 1-2021-04090 (22) 06/01/2020
(86) PCT/US2020/012338 06/01/2020 (87) WO2020/142762 09/07/2020
(30) 62/788,898 06/01/2019 US
(45) 25/08/2025 449 (43) 25/10/2021 403A
(73) BEIJING DAJIA INTERNET INFORMATION TECHNOLOGY CO., LTD. (CN)
Room 101D1-7, 1st Floor, Building 1, No. 6, Shangdi West Road, Haidian District,
Beijing, Beijing 100085, China
(72) XIU, Xiaoyu (CN); CHEN, Yi-Wen (TW); WANG, Xianglin (US).
(74) Công ty Luật TNHH T&G (TGVN)
-

- (54) PHƯƠNG PHÁP GIẢI MÃ VIDEO, THIẾT BỊ TÍNH TOÁN VÀ VẬT GHI ĐỌC
ĐƯỢC BẰNG MÁY TÍNH

(57) Sáng chế đề cập tới phương pháp giải mã video, thiết bị tính toán và vật ghi phi chuyển tiếp đọc được bằng máy tính. Phương pháp này bao gồm bước giải mã hình ảnh tham chiếu thứ nhất và hình ảnh tham chiếu thứ hai, hình ảnh tham chiếu thứ hai đang khác với hình ảnh tham chiếu thứ nhất, trong đó hình ảnh tham chiếu thứ nhất là được hiển thị trước hình ảnh hiện tại và hình ảnh tham chiếu thứ hai được hiển thị sau hình ảnh hiện tại. Phần tinh chỉnh chuyển động của đơn vị mã hóa (coding unit - CU) được tính toán bằng cách tối thiểu hóa sự chênh lệch giữa dự đoán thứ nhất L_0 và dự đoán thứ hai L_1 . Các trị số gradient thứ nhất cho dự đoán thứ nhất L_0 và các trị số gradient thứ hai cho dự đoán thứ hai L_1 được tính toán. Việc dự đoán đôi cuối cùng của CU được tính toán. Sáng chế cũng đề xuất phương pháp liên quan và vật ghi phi chuyển tiếp đọc được bằng máy tính liên quan.



Lĩnh vực kỹ thuật của sáng chế

Sáng chế này đề cập chung tới việc mã hóa (ví dụ, ghi mã và giải mã) dữ liệu video. Cụ thể hơn, phần bộc lộ này đề cập tới các phương pháp và các thiết bị để cải thiện các phương pháp xử lý luồng quang học hai chiều (bi-directional optical flow - BDOF) cho việc mã hóa video. Theo các phương án thực hiện làm ví dụ, các phương pháp được mô tả nhờ đó khôi phục dự đoán hai chiều từ các hình ảnh tham chiếu xuất hiện trong các hướng đối diện so với hình ảnh hiện tại (ví dụ, hình ảnh tham chiếu thứ nhất có thứ tự hiển thị trước hình ảnh hiện tại và hình ảnh tham chiếu thứ hai đang có thứ tự hiển thị sau hình ảnh hiện tại) sử dụng BDOF.

Tình trạng kỹ thuật của sáng chế

Phần này cung cấp thông tin tình trạng kỹ thuật của sáng chế này. Thông tin được chứa trong phần này không nhất thiết phải được coi là tình trạng kỹ thuật của sáng chế.

Các kỹ thuật mã hóa video khác nhau có thể được sử dụng để nén dữ liệu video. Việc mã hóa video được thực hiện theo một hoặc nhiều tiêu chuẩn mã hóa video. Ví dụ, các tiêu chuẩn mã hóa video chứa việc mã hóa video vạn năng (Versatile Video Coding - VVC), mã hóa mô hình thử nghiệm thăm dò chung (joint exploration test model coding - JEM), mã hóa video hiệu suất cao (high-efficiency video coding - H.265/HEVC), mã hóa video cải tiến (advanced video coding - H.264/AVC), mã hóa nhóm chuyên gia hình ảnh chuyển động (moving picture experts group coding - MPEG), hoặc dạng tương tự. Việc mã hóa video thường tận dụng các phương pháp dự đoán (ví dụ, việc dự đoán liên thông, việc dự đoán nội, hoặc dạng tương tự) tận dụng ưu điểm của sự dư thừa có mặt trong các hình ảnh hoặc các chuỗi video. Mục tiêu quan trọng của các kỹ thuật mã hóa video là để nén dữ liệu video thành dạng sử dụng tỉ lệ bit thấp hơn, trong khi loại bỏ hoặc tối thiểu hóa các suy giảm chất lượng video.

Phiên bản thứ nhất của tiêu chuẩn HEVC được hoàn thành trong tháng 10 năm, đề xuất việc tiết kiệm xấp xỉ 50% tỉ lệ bit hoặc chất lượng nhận thức tương tự khi so với

tiêu chuẩn mã hóa video thế hệ trước đó H.264/MPEG AVC. Mặc dù tiêu chuẩn HEVC cung cấp các cải tiến mã hóa đáng kể hơn so với các tiêu chuẩn tiền nhiệm của nó, nhưng có bằng chứng là có thể đạt được hiệu suất mã hóa vượt trội hơn HEVC, với các công cụ mã hóa bổ sung. Dựa trên đó, cả VCEG và MPEG đã bắt đầu công việc khám phá các kỹ thuật mã hóa mới cho việc chuẩn hóa mã hóa video trong tương lai. Một nhóm khám phá video hợp tác (Joint Video Exploration Team - JVET) được hình thành trong tháng 10 năm 2015 bởi ITU-T VCEG và ISO/IEC MPEG để bắt đầu nghiên cứu một cách đáng kể các công nghệ cải tiến có thể cho phép thực hiện việc tăng cường một cách đáng kể cho hiệu suất mã hóa. JVET đã phát triển phần mềm tham chiếu được gọi là mô hình khám phá hợp tác (joint exploration model - JEM) bằng cách tích hợp nhiều công cụ mã hóa bổ sung trên đỉnh của mô hình thử nghiệm HEVC (HEVC test model - HM).

Trong tháng 10 năm 2017, ITU-T và ISO/IEC đã phát hành cuộc gọi hợp tác cho các đề xuất (call for proposals - CfP) trên việc nén video với khả năng vượt quá HEVC. Trong tháng 4 năm 2018, 23 phản hồi CfP được nhận và được đánh giá tạo phiên họp JVET lần thứ 10, đã thể hiện rằng hiệu suất nén đã vượt quá HEVC khoảng quanh 40%. Dựa trên các kết quả đánh giá này, JVET đã triển khai dự án mới để giúp phát triển tiêu chuẩn mã hóa video thế hệ mới có tên là mã hóa video vạn năng (Versatile Video Coding - VVC).. Trong cùng tháng đó, một cơ sở mã phần mềm tham chiếu, được gọi là mô hình kiểm tra VVC (VVC test model - VTM), được thiết lập để thể hiện việc áp dụng tham chiếu của tiêu chuẩn VVC.

Bản chất kỹ thuật của sáng chế

Phần này cung cấp tóm tắt chung của sáng chế, và là không phải là bộ lọc toàn diện của phạm vi đầy đủ của nó hoặc tất cả các dấu hiệu của nó.

Theo khía cạnh thứ nhất, sáng chế đề xuất phương pháp điều khiển độ rộng bit cho luồng quang học hai chiều (bi-directional optical flow - BDOF) cho việc mã hóa video, được thực hiện tại thiết bị tính toán đang có một hoặc nhiều bộ xử lý và bộ nhớ đang lưu trữ nhiều chương trình cần được thực thi bởi một hoặc nhiều bộ xử lý, phương pháp chứa bước giải mã hình ảnh tham chiếu thứ nhất và hình ảnh tham chiếu thứ hai. Hình ảnh tham chiếu thứ hai đang khác với hình ảnh tham chiếu thứ nhất, trong đó hình ảnh tham chiếu thứ nhất là được hiển thị trước hình ảnh hiện tại và hình ảnh tham chiếu thứ

hai được hiển thị sau hình ảnh hiện tại. Vector chuyển động thứ nhất (MV_0) được xác định từ khối hiện tại của hình ảnh hiện tại tới dự đoán thứ nhất L_0 của hình ảnh tham chiếu thứ nhất. Vector chuyển động thứ hai (MV_1) được xác định từ khối hiện tại tới dự đoán thứ hai L_1 của hình ảnh tham chiếu thứ hai. Khoảng cách theo thời gian thứ nhất (τ_0) được xác định từ hình ảnh hiện tại tới hình ảnh tham chiếu thứ nhất và khoảng cách theo thời gian thứ hai (τ_1) được xác định từ hình ảnh hiện tại tới hình ảnh tham chiếu thứ hai. Phần tinh chỉnh chuyển động được tính toán bằng cách tối thiểu hóa sự chênh lệch giữa dự đoán thứ nhất L_0 và dự đoán thứ hai L_1 . Trị số gradient thứ nhất và trị số gradient thứ hai được tính toán, trong đó dịch chuyển phải bổ sung được đưa vào để làm giảm độ rộng bit bên trong của trị số gradient thứ nhất và trị số gradient thứ hai. Việc lấy tổng của trị số gradient thứ nhất và trị số gradient thứ hai được tính toán, trong đó dịch chuyển phải, bổ sung được đưa vào việc lấy tổng để làm giảm độ rộng bit của nó. Khác biệt giữa mẫu dự đoán thứ nhất L_0 và mẫu dự đoán thứ hai L_1 được tính toán, trong đó dịch chuyển phải, bổ sung được đưa vào mẫu dự đoán thứ nhất L_0 và mẫu dự đoán thứ hai L_1 để làm giảm độ rộng bit bên trong của khác biệt.

Theo khía cạnh thứ hai, sáng chế đề xuất thiết bị tính toán chứa một hoặc nhiều bộ xử lý, vật ghi phi chuyển tiếp được ghép nối tới một hoặc nhiều bộ xử lý, và nhiều chương trình được lưu trữ trong vật ghi phi chuyển tiếp. Khi nhiều chương trình được thực thi bởi một hoặc nhiều bộ xử lý, thiết bị tính toán giải mã hình ảnh tham chiếu thứ nhất và hình ảnh tham chiếu thứ hai. Hình ảnh tham chiếu thứ hai là khác với hình ảnh tham chiếu thứ nhất, trong đó hình ảnh tham chiếu thứ nhất là được hiển thị trước hình ảnh hiện tại và hình ảnh tham chiếu thứ hai được hiển thị sau hình ảnh hiện tại; Vector chuyển động thứ nhất (MV_0) được xác định từ khối hiện tại của hình ảnh hiện tại tới dự đoán thứ nhất L_0 của hình ảnh tham chiếu thứ nhất. Vector chuyển động thứ hai (MV_1) được xác định từ khối hiện tại tới dự đoán thứ hai L_1 của hình ảnh tham chiếu thứ hai. Khoảng cách theo thời gian thứ nhất (τ_0) được xác định từ hình ảnh hiện tại tới hình ảnh tham chiếu thứ nhất và khoảng cách theo thời gian thứ hai (τ_1) được xác định từ hình ảnh hiện tại tới hình ảnh tham chiếu thứ hai. Phần tinh chỉnh chuyển động được tính toán bằng cách tối thiểu hóa sự chênh lệch giữa dự đoán thứ nhất L_0 và dự đoán thứ hai L_1 . Trị số gradient thứ nhất và trị số gradient thứ hai được tính toán, trong đó dịch chuyển phải, bổ sung được đưa vào để làm giảm độ rộng bit bên trong của trị số gradient thứ

nhất và trị số gradient thứ hai. Việc lấy tổng của trị số gradient thứ nhất và trị số gradient thứ hai được tính toán, trong đó dịch chuyển phải, bổ sung được đưa vào việc lấy tổng để làm giảm độ rộng bit của nó. Khác biệt giữa mẫu dự đoán thứ nhất L_0 và mẫu dự đoán thứ hai L_1 được tính toán, trong đó dịch chuyển phải, bổ sung được đưa vào mẫu dự đoán thứ nhất L_0 và mẫu dự đoán thứ hai L_1 để làm giảm độ rộng bit bên trong của khác biệt.

Theo khía cạnh thứ ba, sáng chế đề xuất vật ghi phi chuyển tiếp đọc được bằng máy tính lưu trữ nhiều chương trình để thực thi bởi thiết bị tính toán đang có một hoặc nhiều bộ xử lý. Nhiều chương trình, khi được thực thi bởi một hoặc nhiều bộ xử lý, sẽ làm cho thiết bị tính toán giải mã hình ảnh tham chiếu thứ nhất và hình ảnh tham chiếu thứ hai. Hình ảnh tham chiếu thứ hai là khác với hình ảnh tham chiếu thứ nhất, trong đó hình ảnh tham chiếu thứ nhất là được hiển thị trước hình ảnh hiện tại và hình ảnh tham chiếu thứ hai được hiển thị sau hình ảnh hiện tại; Vector chuyển động thứ nhất (MV_0) được xác định từ khối hiện tại của hình ảnh hiện tại tới dự đoán thứ nhất L_0 của hình ảnh tham chiếu thứ nhất. Vector chuyển động thứ hai (MV_1) được xác định từ khối hiện tại tới dự đoán thứ hai L_1 của hình ảnh tham chiếu thứ hai. Khoảng cách theo thời gian thứ nhất (τ_0) được xác định từ hình ảnh hiện tại tới hình ảnh tham chiếu thứ nhất và khoảng cách theo thời gian thứ hai (τ_1) được xác định từ hình ảnh hiện tại tới hình ảnh tham chiếu thứ hai. Phần tinh chỉnh chuyển động được tính toán bằng cách tối thiểu hóa sự chênh lệch giữa dự đoán thứ nhất L_0 và dự đoán thứ hai L_1 . Trị số gradient thứ nhất và trị số gradient thứ hai được tính toán, trong đó dịch chuyển phải, bổ sung được đưa vào để làm giảm độ rộng bit bên trong của trị số gradient thứ nhất và trị số gradient thứ hai. Việc lấy tổng của trị số gradient thứ nhất và trị số gradient thứ hai được tính toán, trong đó dịch chuyển phải, bổ sung được đưa vào việc lấy tổng để làm giảm độ rộng bit của nó. Khác biệt giữa mẫu dự đoán thứ nhất L_0 và mẫu dự đoán thứ hai L_1 được tính toán, trong đó dịch chuyển phải, bổ sung được đưa vào mẫu dự đoán thứ nhất L_0 và mẫu dự đoán thứ hai L_1 để làm giảm độ rộng bit bên trong của khác biệt.

Mô tả vắn tắt các hình vẽ

Dưới đây, các tập hợp của các phương án thực hiện minh họa, và không làm hạn chế của sáng chế này sẽ được mô tả cùng với các hình vẽ kèm theo. Các biến thể của

cấu trúc, phương pháp, hoặc chức năng có thể được áp dụng bởi những người có hiểu biết trung bình trong lĩnh vực kỹ thuật có liên quan, dựa trên các ví dụ được thể hiện ở đây, và các biến thể đều được chứa trong phạm vi bảo hộ của sáng chế này. Trong các trường hợp mà trong đó không có mặt xung đột, thì các chỉ dẫn của các phương án thực hiện khác nhau này có thể, nhưng không nhất thiết, được kết hợp với chỉ dẫn khác.

Fig.1 là giản đồ khối của hệ thống ghi mã video lai, dựa trên khối theo các phương án thực hiện làm ví dụ chứa bộ ghi mã làm ví dụ thích hợp cho việc sử dụng với các phương pháp luồng quang học hai chiều (bi-directional optical flow - BDOF);

Fig.2 là giản đồ khối của bộ giải mã làm ví dụ theo các phương án thực hiện làm ví dụ thích hợp cho việc sử dụng với các phương pháp luồng quang học hai chiều (bi-directional optical flow - BDOF);

Các hình vẽ từ Fig. 3A đến Fig.3E thể hiện các loại chia tách ví dụ, cụ thể là phân chia tứ phân (Fig.3A), phân chia nhị phân nằm ngang (Fig.3B), phân chia nhị phân thẳng đứng (Fig.3C), phân chia tam phân nằm ngang (Fig.3D), và phân chia tam phân thẳng đứng (Fig.3E), theo các phương án thực hiện làm ví dụ.

Fig.4 thể hiện mô hình luồng quang học hai chiều (bi-directional optical flow - BDOF) làm ví dụ, và

Fig.5 là lưu đồ của phương pháp làm ví dụ để dự đoán khối của dữ liệu video theo các phương án thực hiện làm ví dụ.

Mô tả chi tiết sáng chế

Các thuật ngữ được sử dụng ở đây theo phần bộc lộ này đề cập tới các ví dụ cụ thể làm minh họa, hơn là để hạn chế phần bộc lộ này. Dang số ít “một”, “cái” và “chiếc” như được sử dụng trong sáng chế này cũng như trong các yêu cầu bảo hộ kèm theo cũng đề cập tới các dạng số nhiều, trừ khi các ý nghĩa khác được bao hàm một cách rõ rệt trong văn cảnh. Cần hiểu rằng thuật ngữ “và/hoặc” như được sử dụng ở đây đề cập tới tổ hợp bất kỳ hoặc tất cả các tổ hợp khả thi của một hoặc nhiều mục được liệt kê, được kết hợp.

Cần hiểu rằng, mặc dù các thuật ngữ “thứ nhất”, “thứ hai”, “thứ ba”, v.v. có thể được sử dụng ở đây để mô tả các thông tin khác nhau, nhưng không tin không bị hạn

chế vào các thuật ngữ này. Các thuật ngữ này chỉ được sử dụng để phân biệt một phân loại của thông tin so với thông tin khác. Ví dụ, không tách khỏi phạm vi của sáng chế này, thông tin thứ nhất có thể được đặt là thông tin thứ hai, và tương tự, thông tin thứ hai cũng có thể được đặt là thông tin thứ nhất. Như được sử dụng ở đây, thuật ngữ “nếu” có thể được hiểu nghĩa là “khi” hoặc “theo” hoặc “đáp lại”, phụ thuộc vào văn cảnh.

Đề cập trong suốt bản mô tả này tới “một phương án thực hiện”, “phương án thực hiện”, “phương án thực hiện khác”, hoặc dạng tương tự ở dạng số ít hoặc số nhiều có nghĩa là một hoặc nhiều dấu hiệu, cấu trúc, hoặc đặc tính cụ thể được mô tả cùng với phương án thực hiện được chứa trong ít nhất một phương án thực hiện của sáng chế này. Do đó, các biểu diễn của các cụm “theo một phương án thực hiện” hoặc “theo phương án thực hiện”, “theo phương án thực hiện khác”, hoặc dạng tương tự ở dạng số ít hoặc số nhiều trong các vị trí khác nhau trong suốt bản mô tả này không nhất thiết đều phải đề cập tới cùng một phương án thực hiện. Hơn nữa, các dấu hiệu, cấu trúc, hoặc các đặc tính cụ thể theo một hoặc nhiều phương án thực hiện có thể được kết hợp theo cách thích hợp bất kỳ.

Sáng chế này đề cập chung tới việc mã hóa (ví dụ, ghi mã và giải mã) dữ liệu video. Cụ thể hơn, phần bộc lộ này đề cập tới các phương pháp và các thiết bị để cải thiện các phương pháp xử lý luồng quang học hai chiều (bi-directional optical flow - BDOF) cho việc mã hóa video. Theo các phương án thực hiện làm ví dụ, các phương pháp được mô tả nhờ đó, bằng cách sử dụng BDOF, khối được dự đoán hai chiều từ các hình ảnh tham chiếu xuất hiện trong các hướng đối diện so với hình ảnh hiện tại (ví dụ, hình ảnh tham chiếu thứ nhất có thứ tự hiển thị trước hình ảnh hiện tại và hình ảnh tham chiếu thứ hai đang có thứ tự hiển thị sau hình ảnh hiện tại).

Các phương pháp như được mô tả ở đây liên quan tới việc cải thiện hoặc tăng cường BDOF. BDOF có thể được áp dụng trong suốt việc bù chuyển động và được sử dụng để biến đổi vector chuyển động trên cơ sở mỗi điểm ảnh cho khối hiện tại, sao cho các điểm ảnh của khối hiện tại được dự đoán sử dụng các trị số dịch vị tương ứng được áp dụng cho vector chuyển động. Theo các phương án thực hiện làm ví dụ, BDOF được sử dụng khi các vector chuyển động được sử dụng để dự đoán liên thông khối của dữ liệu video của hình ảnh hiện tại so với các khối tham chiếu trong các hình ảnh tham chiếu

theo các hướng dự đoán đối diện so với khối hiện tại. Các phương pháp mã hóa video chứa việc dự đoán theo không gian (nội hình ảnh) và/hoặc việc dự đoán theo thời gian (liên hình ảnh) để làm giảm hoặc loại bỏ sự dư thừa cố hữu của các trình tự video. Với việc mã hóa video dựa trên khối, lát video (ví dụ, khung video hoặc một phần của khung video) có thể được phân chia thành các khối video, vốn có thể còn được đề cập tới như là các khối cây, các đơn vị mã hóa (coding unit - CU), và/hoặc các nút mã hóa. Các khối video trong lát được mã hóa nội (I) của hình ảnh có thể được ghi mã sử dụng việc dự đoán theo không gian liên quan tới các mẫu tham chiếu trong các khối lân cận trong cùng một hình ảnh. Các khối video trong lát được mã hóa liên thông (P hoặc B) của hình ảnh có thể được sử dụng trong việc dự đoán theo không gian liên quan tới các mẫu trong các khối lân cận trong cùng một hình ảnh hoặc việc dự đoán theo thời gian liên quan tới các mẫu tham chiếu trong các hình ảnh tham chiếu khác. Các hình ảnh có thể được đề cập tới như là các khung, và các hình ảnh tham chiếu có thể được đề cập tới như là các khung tham chiếu.

Việc dự đoán theo không gian hoặc việc dự đoán theo thời gian tạo thành khối dự đoán hoặc khối cần được mã hóa. Dữ liệu còn dư biểu diễn các khác biệt hình ảnh giữa khối gốc cần được giải mã và khối dự đoán. Khối được giải mã liên thông được giải mã theo vectơ chuyển động chỉ tới khối của các mẫu tham chiếu đang tạo thành khối dự đoán, và hình ảnh còn dư đang chỉ thị khác biệt giữa khối được mã hóa và khối dự đoán. Khối được mã hóa nội được mã hóa theo chế độ mã hóa nội và dữ liệu còn dư. Với việc nén thêm, dữ liệu còn dư có thể được biến đổi từ miền điểm ảnh tới miền biến đổi, tạo thành các hệ số biến đổi còn dư, vốn sau đó có thể được lượng tử hóa. Các hệ số biến đổi được lượng tử hóa, được sắp xếp từ đầu vào mảng hai chiều, có thể được quét để tạo ra vectơ một chiều của các hệ số biến đổi, và việc mã hóa entropy có thể được áp dụng để đạt được việc nén hơn nữa.

Với từng khối được dự đoán liên thông, tập hợp của thông tin chuyển động có thể có sẵn. Tập hợp của thông tin chuyển động có thể chứa thông tin chuyển động cho các hướng dự đoán về phía trước và về phía sau. Như được sử dụng ở đây, hướng về phía trước hoặc hướng dự đoán thứ nhất và hướng về phía sau hoặc hướng dự đoán thứ hai là hai hướng dự đoán của chế độ dự đoán hai hướng, và các thuật ngữ “về phía trước”,

“thứ nhất”, “về phía sau”, và “thứ hai” không nhất thiết phải có ý nghĩa về mặt hình học. Thay vào đó, các thuật ngữ này nói chung là tương ứng với việc xem liệu các hình ảnh tham chiếu cần được hiển thị trước (“về phía sau”) hay sau (“về phía trước”) hình ảnh hiện tại. Theo các phương án thực hiện làm ví dụ, hướng dự đoán về phía trước có thể tương ứng với danh sách hình ảnh tham chiếu 0 (RefPicList0) và hướng dự đoán về phía sau có thể tương ứng với danh sách hình ảnh tham chiếu 1 (RefPicList1) của hình ảnh hiện tại.

Với từng hướng dự đoán, thông tin chuyển động chứa chỉ số tham chiếu và vector chuyển động. Theo các phương án thực hiện làm ví dụ, chỉ số tham chiếu có thể được sử dụng để nhận diện hình ảnh tham chiếu trong danh sách hình ảnh tham chiếu hiện tại (RefPicList0 hoặc RefPicList1). Vector chuyển động có thành phần nằm ngang (x) và thành phần thẳng đứng (y). Nói chung, thành phần nằm ngang chỉ thị dịch chuyển nằm ngang nằm trong hình ảnh tham chiếu, so với vị trí của khối hiện tại trong hình ảnh hiện tại, cho việc định vị tọa độ x của khối tham chiếu, trong khi thành phần thẳng đứng chỉ thị dịch chuyển thẳng đứng nằm trong hình ảnh tham chiếu, so với vị trí của khối hiện tại, để định vị tọa độ y của khối tham chiếu.

Các trị số số đếm thứ tự hình ảnh (Picture order count - POC) được sử dụng rộng rãi trong các tiêu chuẩn mã hóa video để nhận diện thứ tự hiển thị của hình ảnh. Mặc dù trong một số tình huống hai hình ảnh nằm trong một trình tự video được mã hóa có thể có cùng một trị số POC, nhưng nó thường không xảy ra trong một trình tự video được mã hóa. Do đó, các trị số POC của các hình ảnh nói chung là duy nhất, và do đó có thể nhận diện một cách duy nhất các hình ảnh tương ứng. Khi nhiều trình tự video được mã hóa có mặt trong dòng bit, thì các hình ảnh đang có cùng một trị số POC có thể gần với nhau hơn, theo nghĩa của thứ tự giải mã. Các trị số POC của các hình ảnh thường được sử dụng cho việc xây dựng danh sách hình ảnh tham chiếu, việc dẫn ra của các tập hợp hình ảnh tham chiếu như trong HEVC, và việc định cỡ vector chuyển động.

Theo khái niệm, nhiều tiêu chuẩn mã hóa video là tương tự, chứa các phần được đề cập tới ở trên trong phần tình trạng kỹ thuật của sáng chế. Ví dụ, gần như tất cả các tiêu chuẩn mã hóa video sử dụng việc xử lý dựa trên khối và chia sẻ cùng các giản đồ khối video tương tự để đạt được việc nén video.

Giống như HEVC, VVC được xây dựng trên khung mã hóa video lai, dựa trên khối. Fig.1 là giản đồ khối của hệ thống ghi mã video lai, dựa trên khối chứa bộ ghi mã làm ví dụ 100 có thể sử dụng các phương pháp cho luồng quang học hai chiều theo các phương án thực hiện làm ví dụ. Tín hiệu video đầu vào được xử lý theo từng khối (các CU). Trong VTM-1.0, CU có thể lên tới 128x128 điểm ảnh. Tuy nhiên, khác biệt từ HEVC vốn phân chia các khối chỉ dựa trên các cây tứ phân, trong VVC, một đơn vị cây mã hóa (coding tree unit - CTU) được chia tách thành các CU để thích ứng để thay đổi các đặc tính cục bộ dựa trên cây tứ phân/nhi phân/tam phân. Thêm vào đó, khác niệm của = loại đơn vị nhiều phân chia trong HEVC được loại bỏ, tức là, việc tách ra của CU, đơn vị dự đoán (prediction unit - PU) và đơn vị biến đổi (transform unit - TU) không tồn tại trong VVC nữa; thay vào đó, mỗi CU luôn được sử dụng như là đơn vị cơ bản cho cả việc dự đoán và biến đổi mà không có thêm các phần chia khác nữa. Trong cấu trúc cây nhiều loại, một CTU được phân chia bởi cấu trúc cây tứ phân và sau đó nút lá cây tứ phân có thể còn được phân chia nữa bởi cấu trúc cây nhị phân hoặc tam phân.

Đề cập tới Fig.1, bộ ghi mã 100, khung video được phân chia thành nhiều khối video để xử lý. Với khối video đã cho, phần dự đoán được tạo thành dựa trên hoặc là cách tiếp cận dự đoán liên thông hoặc là cách tiếp cận dự đoán nội. Theo việc dự đoán liên thông, một hoặc nhiều bộ dự đoán được tạo thành qua ước lượng chuyển động và việc bù chuyển động, dựa trên các điểm ảnh từ các khung được tái tạo trước đó. Trong việc dự đoán nội, các bộ phận dự đoán được tạo thành dựa trên các điểm ảnh được tái tạo trong khung hiện tại. Qua việc quyết định chế độ, bộ dự đoán tốt nhất có thể được chọn để dự đoán khối hiện tại.

Phần còn dư của dự đoán, biểu diễn khác biệt giữa khối video hiện tại và bộ phận dự đoán của nó, được gửi tới mạch biến đổi 102. Thuật ngữ “mạch” như được sử dụng ở đây chứa phần cứng và phần mềm để vận hành phần cứng. Mạch biến đổi 102 áp dụng việc biến đổi, như là biến đổi cosin rời rạc (discrete cosine transform - DCT) hoặc biến đổi tương tự theo khái niệm (conceptually similar transform), cho khối còn dư, tạo ra khối video chứa các trị số hệ số biến đổi. Các biến đổi sóng nhỏ, các biến đổi băng phụ, các biến đổi sin rời rạc (discrete sine transform - DST), hoặc các loại biến đổi khác có thể được sử dụng thay cho DCT. Mạch biến đổi 102 áp dụng biến đổi cho khối còn dư,

tạo ra khối của các hệ số biến đổi. Việc biến đổi có thể biến đổi thông tin còn dư từ miền điểm ảnh tới miền biến đổi, như miền tần số.

Các hệ số biến đổi được tạo thành sau đó được gửi từ mạch biến đổi 102 tới mạch lượng tử hóa 104 để làm giảm entropy. Mạch lượng tử hóa 104 lượng tử hóa các hệ số để tiếp tục làm giảm tỉ lệ bit. Quy trình lượng tử hóa có thể làm giảm độ sâu bit được kết hợp với một số hoặc tất cả các hệ số trong số các hệ số. Mức độ lượng tử hóa có thể được biến đổi bằng cách điều chỉnh thông số lượng tử hóa. Các hệ số được lượng tử hóa sau đó được nạp vào mạch mã hóa entropy 106 để sinh ra dòng bit video đã được nén. Mạch mã hóa entropy 106 mã hóa entropy các hệ số biến đổi được lượng tử hóa. Ví dụ, mạch mã hóa entropy 106 có thể thực hiện việc mã hóa chiều dài thay đổi thích ứng văn cảnh (context adaptive variable length coding - CAVLC), mã hóa số học nhị phân thích ứng văn cảnh (context adaptive binary arithmetic coding - CABAC), mã hóa số học nhị phân thích ứng văn cảnh dựa trên cú pháp (syntax-based context-adaptive binary arithmetic coding - SBAC), mã hóa entropy phân chia khoảng gián cách khả năng (probability interval partitioning entropy - PIPE) hoặc kỹ thuật mã hóa entropy khác. Với việc mã hóa entropy dựa trên văn cảnh, văn cảnh có thể là dựa trên các khối lân cận. Việc tuân theo việc mã hóa entropy bởi mạch mã hóa entropy 106, dòng bit được ghi mã có thể được truyền tới thiết bị khác (ví dụ, video bộ giải mã) hoặc đạt được cho việc truyền hoặc truy hồi sau này. Như được thể hiện trên Fig.1, thông tin liên quan tới việc dự đoán 110 từ mạch dự đoán liên thông và/hoặc mạch dự đoán nội 112, như thông tin phân chia khối video, các vectơ chuyển động, chỉ số hình ảnh tham chiếu, và chế độ dự đoán nội, cũng được nạp qua mạch mã hóa entropy 106 và được lưu vào trong dòng bit video được nén 114.

Trong bộ ghi mã 100, mạch liên quan tới bộ giải mã cũng là cần thiết để tái tạo các điểm ảnh cho mục đích dự đoán. Đầu tiên, phần còn dư của việc dự đoán được tái tạo qua mạch lượng tử hóa ngược 116 và mạch biến đổi ngược 118. Phần còn dư của việc dự đoán được tái tạo này được kết hợp với bộ phận dự đoán khối 120 để sinh ra các điểm ảnh được tái tạo chưa được lọc cho khối video hiện tại. Mạch lượng tử hóa ngược 116 và mạch biến đổi ngược 118 áp dụng việc lượng tử hóa ngược và biến đổi ngược, theo cách tương ứng, để tái tạo khối còn dư trong miền điểm ảnh. Theo các phương án thực

hiện cụ thể, bộ phận lấy tổng cộng khối còn dư được tái tạo vào khối dự đoán được bù chuyển động được tạo ra sớm hơn bởi mạch bù chuyển động 122 hoặc mạch dự đoán nội 112 để tạo ra khối video được tái tạo để lưu trữ trong bộ nhớ hình ảnh tham chiếu. Khối video được tái tạo có thể được sử dụng bởi mạch ước lượng chuyển động 124 và mạch bù chuyển động 122 như là khối tham chiếu để mã hóa liên thông khối trong khung video tiếp theo.

Để cải thiện hiệu suất mã hóa và chất lượng trực quan, bộ lọc trong vòng thường được sử dụng. Ví dụ, bộ lọc khử khối là có sẵn trong AVC và HEVC, cũng như phiên bản hiện tại của VVC. Trong HEVC, bộ lọc trong vòng bổ sung còn được gọi là SAO (sample adaptive offset - dịch vị thích ứng mẫu) được định nghĩa để còn tiếp tục cải thiện hiệu suất mã hóa. Trong phiên bản hiện tại nhất của tiêu chuẩn VVC, bộ lọc trong vòng khác còn được gọi là ALF (adaptive loop filter - bộ lọc vòng lặp thích ứng) đang được khảo sát một cách chủ động để tích hợp vào sản phẩm cuối cùng. Các hoạt động của bộ lọc trong vòng này là tùy chọn. Việc thực hiện các hoạt động này giúp cải thiện hiệu suất mã hóa và chất lượng trực quan. Các hoạt động cũng có thể được tắt như là quyết định được phục hồi bởi bộ ghi mã 100 để tiết kiệm độ phức tạp tính toán. Cần chú ý rằng việc dự đoán nội thường dựa trên các điểm ảnh được tái tạo không được lọc, trong khi việc dự đoán liên thông là dựa trên các điểm ảnh được tái tạo được lọc nếu các lựa chọn bộ lọc này được bật lên bởi bộ ghi mã 100.

Fig.2 là giản đồ khối của bộ giải mã làm ví dụ 200 có thể sử dụng các phương pháp cho luồng quang học hai chiều. Theo các phương án thực hiện làm ví dụ, bộ giải mã 200 là tương tự như phần liên quan tới việc xây dựng lại nằm trong bộ ghi mã 100 của Fig.1. Đề cập tới Fig.2, trong bộ giải mã 200 luồng video đang đi tới 201 đầu tiên được giải mã qua mạch giải mã entropy 202 để dẫn ra các mức hệ số được lượng tử hóa và thông tin liên quan tới việc dự đoán. Các mức hệ số được lượng tử hóa sau đó được xử lý qua mạch lượng tử hóa ngược 204 và mạch biến đổi ngược 206 để thu được phần còn dư, dự đoán, được tái tạo. Mạch lượng tử hóa ngược 204 đảo ngược việc lượng tử hóa, tức là, giải lượng tử, các hệ số biến đổi được lượng tử hóa được tạo ra trong dòng bit và được giải mã bởi mạch giải mã entropy 202. Quy trình lượng tử hóa ngược có thể chứa việc sử dụng của thông số lượng tử hóa QPy được tính toán bởi bộ giải mã 200 cho từng

khối video trong lát video để xác định mức độ lượng tử hóa, và theo cách tương tự, là mức độ lượng tử hóa ngược cần được áp dụng. Mạch biến đổi ngược 206 áp dụng việc biến đổi ngược, ví dụ, DCT ngược, biến đổi số nguyên ngược, hoặc quy trình biến đổi hóa ngược tương tự về mặt ý tưởng, cho các hệ số biến đổi để tạo ra các khối phần còn dư trong miền điểm ảnh.

Cơ chế bộ dự đoán khối, vốn có thể được áp dụng trong mạch chế độ nội/liên thông 208, chứa mạch dự đoán nội 210 được đặt cấu hình để thực hiện quy trình dự đoán nội và/hoặc mạch bù chuyển động 212 được đặt cấu hình để thực hiện quy trình bù chuyển động dựa trên thông tin dự đoán được giải mã. Tập hợp của các điểm ảnh được tái tạo chưa được lọc bằng cách lấy tổng phần dư dự đoán được tái tạo từ mạch biến đổi ngược 206 và đầu ra dự đoán được sinh ra bởi cơ chế bộ phận dự đoán khối, sử dụng bộ lấy tổng 214. Trong các tình huống mà trong đó bộ lọc trong vòng 216 được bật lên, thao tác lọc được thực hiện trên các điểm ảnh được tái tạo này để dẫn ra video được tái tạo cuối cùng để đưa ra.

Khi lát video được mã hóa như là lát được mã hóa nội (I), thì mạch dự đoán nội 210 có thể sinh ra dữ liệu dự đoán cho khối video của lát video hiện tại dựa trên mô hình dự đoán nội được tạo tín hiệu và dữ liệu từ các khối được giải mã trước đó của khung hoặc hình ảnh hiện tại. Khi khung video được mã hóa như là lát được mã hóa liên thông (tức là, B, P hoặc GPB), thì mạch bù chuyển động 212 tạo ra các khối dự đoán cho khối video của lát video hiện tại dựa trên các vector chuyển động và các thành phần cú pháp khác nhận được từ mạch giải mã entropy 202. Các khối dự đoán có thể được tạo ra từ một trong số các hình ảnh tham chiếu nằm trong một trong số các danh sách hình ảnh tham chiếu. Bộ giải mã 200 có thể tạo nên các danh sách khung tham chiếu, List 0 và List 1, sử dụng các kỹ thuật tạo mặc định dựa trên các hình ảnh tham chiếu được lưu trữ trong bộ nhớ hình ảnh tham chiếu.

Mạch bù chuyển động 212 xác định thông tin dự đoán cho khối video của lát video hiện tại bằng cách phân tích cú pháp các vector chuyển động và các thành phần cú pháp khác và sử dụng thông tin dự đoán để tạo ra các khối dự đoán cho khối video hiện tại đang được giải mã. Ví dụ, mạch bù chuyển động 212 sử dụng một số thành phần trong số các thành phần cú pháp nhận được để xác định mô hình dự đoán (ví dụ, dự đoán nội

hoặc dự đoán liên thông) được sử dụng để mã hóa các khối video của lát video, loại lát dự đoán liên thông (ví dụ, lát B, lát P, hoặc lát GPB), thông tin kết cấu cho một hoặc nhiều danh sách trong số các danh sách hình ảnh tham chiếu cho lát, các vector chuyển động cho từng khối video được mã hóa liên thông của lát, trạng thái dự đoán liên thông cho từng khối video được mã hóa liên thông của lát, và thông tin khác để giải mã các khối video trong lát video hiện tại.

Mạch bù chuyển động 212 cũng có thể thực hiện việc ngoại suy dựa trên các bộ lọc ngoại suy cho độ chính xác dưới điểm ảnh. Mạch bù chuyển động 212 có thể sử dụng các bộ lọc ngoại suy như được sử dụng bởi bộ ghi mã 100 trong suốt quá trình giải mã của các khối video để tính toán các trị số được ngoại suy cho các điểm ảnh dưới số nguyên của các khối tham chiếu. Theo các phương án thực hiện làm ví dụ, mạch bù chuyển động 212 xác định các bộ lọc ngoại suy được sử dụng bởi bộ ghi mã 100 từ các thành phần cú pháp nhận được và sử dụng các bộ lọc ngoại suy để tạo ra các khối dự đoán. Mạch bù chuyển động 212 có thể được đặt cấu hình để thực hiện phương pháp bất kỳ hoặc tất cả các phương pháp trong số các phương pháp được mô tả trong phần bổ cục này (riêng rẽ hoặc theo tổ hợp bất kỳ). Ví dụ, mạch bù chuyển động 212 có thể được đặt cấu hình để thực hiện các phương pháp BDOF làm ví dụ được thảo luận ở đây.

Theo các phương án thực hiện làm ví dụ, mạch bù chuyển động 212 được đặt cấu hình để áp dụng BDOF cho các khối có hình ảnh tham chiếu thứ nhất tạm thời trước hình ảnh hiện tại (ví dụ, trong danh sách hình ảnh tham chiếu 0) và hình ảnh tham chiếu thứ hai tạm thời sau hình ảnh hiện tại (ví dụ, trong danh sách hình ảnh tham chiếu 1). Đáp lại để xác định để áp dụng BDOF cho khối, mạch bù chuyển động 212 có thể áp dụng các phương pháp được thảo luận ở đây (ví dụ, liên quan tới các công thức được mô tả ở đây) để dự đoán khối theo BDOF.

Các hình vẽ từ Fig. 3A đến Fig.3E thể hiện năm loại chia tách ví dụ, cụ thể là phân chia tứ phân (Fig.3A), phân chia nhị phân nằm ngang (Fig.3B), phân chia nhị phân thẳng đứng (Fig.3C), phân chia tam phân nằm ngang (Fig.3D), và phân chia tam phân thẳng đứng (Fig.3E).

Đề cập lại tới Fig.1, việc dự đoán theo không gian và/hoặc việc dự đoán theo thời gian có thể được thực hiện. Việc dự đoán theo không gian (hoặc “dự đoán nội”) sử dụng

các điểm ảnh từ các mẫu của các khối lân cận hiện đã được mã hóa (vốn được gọi là các mẫu tham chiếu) trong cùng một hình ảnh hoặc lát video để dự đoán khối video hiện tại. Việc dự đoán theo không gian làm giảm tính dư thừa về mặt không gian cố hữu trong tín hiệu video. Việc dự đoán theo thời gian (còn được đề cập tới như là “dự đoán liên thông” hoặc “dự đoán bù chuyển động”) sử dụng các điểm ảnh được tái tạo từ các hình ảnh video hiện đã được mã hóa để dự đoán khối video hiện tại. Việc dự đoán theo thời gian làm giảm tính dư thừa về mặt thời gian cố hữu trong tín hiệu video. Tín hiệu dự đoán theo thời gian cho CU đã cho thường được tạo tín hiệu bởi một hoặc nhiều vector chuyển động (motion vector - MV) vốn chỉ thị số lượng và hướng chuyển động giữa CU hiện tại và tham chiếu theo thời gian của CU hiện tại. Cũng vậy, nếu nhiều hình ảnh tham chiếu được hỗ trợ, thì một chỉ số hình ảnh tham chiếu sẽ được gửi bổ sung, vốn được sử dụng để nhận diện nơi mà hình ảnh tham chiếu trong hình ảnh tham chiếu lưu tín hiệu dự đoán theo thời gian, đi ra từ đó. Sau việc dự đoán theo không gian và/hoặc theo thời gian, khối quyết định chế độ trong bộ ghi mã chọn chế độ dự đoán tốt nhất, ví dụ dựa trên phương pháp tối ưu hóa nhiễu loạn tỉ lệ. Khối dự đoán sau đó được trừ từ khối video hiện tại và phần còn dư dự đoán được giải tương quan sử dụng việc biến đổi và lượng tử hóa. Các hệ số phần còn dư được lượng tử hóa được lượng tử hóa ngược và biến đổi ngược để tạo thành phần còn dư được tái tạo, vốn sau đó được thêm lại vào khối dự đoán để tạo thành tín hiệu được tái tạo của CU. Việc lọc trong vòng tiếp theo, sử dụng bộ lọc khử khối, như dịch vị tương thích mẫu (sample adaptive offset - SAO) và bộ lọc trong vòng tương thích (adaptive in-loop filter - ALF) ví dụ, có thể được áp dụng cho CU được tái tạo trước khi CU được tái tạo được đặt trong hình ảnh tham chiếu lưu và được sử dụng để mã hóa các khối video tương lai. Để tạo thành dòng bit video đầu ra, chế độ mã hóa (liên thông hoặc nội), thông tin chế độ dự đoán, thông tin chuyển động, và các hệ số phần còn dư được lượng tử hóa tất cả được gửi tới mạch mã hóa entropy 106 để còn được nén và đóng gói để tạo thành dòng bit.

Fig.4 thể hiện mô hình luồng quang học hai chiều (bi-directional optical flow - BDOF) làm ví dụ. Như được thể hiện trên Fig.4, dự đoán thứ nhất, L_0 , trong khối tham chiếu thứ nhất, I_0 , được định vị tại khoảng cách theo thời gian thứ nhất, τ_0 , từ khối hiện tại, CurBlk, mà việc dự đoán được thực hiện cho nó. Vector chuyển động thứ nhất, A , đang có các thành phần vector chuyển động, (V_x, V_y) , nằm trong khối tham chiếu thứ nhất

I_0 . Theo cách tương tự, dự đoán thứ hai, L_1 , trong khối tham chiếu thứ hai, I_1 , được định vị tại khoảng cách theo thời gian thứ hai, τ_1 , từ khối hiện tại. Vector chuyển động thứ hai B , đang có các thành phần vector chuyển động, $(-V_x, -V_y)$, nằm trong khối tham chiếu thứ hai I_1 .

Luồng quang học hai chiều (Bi-Directional Optical Flow): Việc dự đoán đôi truyền thông trong việc mã hóa video là tổ hợp đơn giản của hai khối dự đoán theo thời gian thu được từ các hình ảnh tham chiếu hiện được tái tạo. Tuy nhiên, do hạn chế của việc bù chuyển động dựa trên khối, nên có thể có chuyển động nhỏ còn lại có thể được quan sát giữa các mẫu của hai khối dự đoán, do đó làm giảm hiệu suất của việc dự đoán bù chuyển động. Để giải quyết vấn đề này, luồng quang học hai chiều (bi-directional optical flow - BDOF) được áp dụng trong VVC để làm giảm tác động của chuyển động này cho mọi mẫu nằm bên trong khối. Cụ thể hơn, như được thể hiện trên Fig.4, BDOF là phần tinh chỉnh chuyển động đảo mẫu được thực hiện trên đỉnh của các dự đoán được bù chuyển động dựa trên khối khi việc dự đoán đôi được sử dụng. Phần tinh chỉnh chuyển động (v_x, v_y) của từng khối con 4×4 được tính toán bằng cách tối thiểu hóa sự chênh lệch giữa các mẫu dự đoán L_0 và L_1 sau khi BDOF được áp dụng bên trong một cửa sổ 6×6 Ω quanh khối con. Cụ thể hơn, trị số của (v_x, v_y) được dẫn ra theo:

$$v_x = S_1 > 0 ? \text{clip3} \left(-th_{BDOF}, th_{BDOF}, -((S_3 \cdot 2^{n_c}) \gg \lfloor \log_2 S_1 \rfloor) \right) : 0$$

$$v_y = S_5 > 0 ? \text{clip3} \left(-th_{BDOF}, th_{BDOF}, - \left((S_6 \cdot 2^{n_c} - \left((v_x S_{2,m}) \ll n_{S_2} + v_x S_{2,s} \right) / 2 \right) \gg \lfloor \log_2 S_5 \rfloor \right) : 0,$$

Phương trình(1)

trong đó $\lfloor \cdot \rfloor$ là hàm làm tròn xuống (floor); $\text{clip3}(\min, \max, x)$ là hàm xén trị số định trước x nằm trong khoảng giới hạn $[\min, \max]$; ký hiệu $\gg$ biểu diễn thao tác dịch sang phải của đảo bit; ký hiệu $\ll$ biểu diễn thao tác dịch sang trái của đảo bit; th_{BDOF} là ngưỡng tinh chỉnh chuyển động để ngăn chặn các lỗi được lan truyền do chuyển động cục bộ không đều đặn, vốn bằng với 2^{13-BD} , trong đó BD là độ sâu bit của video nhập vào. Trong Phương trình (1), $S_{2,m} = S_2 \gg n_{S_2}$, $S_{2,s} = S_2 \& (2^{n_{S_2}} - 1)$.

Các trị số của S_1 , S_2 , S_3 , S_5 , và S_6 được tính toán theo:

$$S_1 = \sum_{(i,j) \in \Omega} \psi_x(i,j) \cdot \psi_x(i,j), S_3 = \sum_{(i,j) \in \Omega} \theta(i,j) \cdot \psi_x(i,j)$$

$$S_2 = \sum_{(i,j) \in \Omega} \psi_x(i,j) \cdot \psi_y(i,j)$$

*Phương
trình(2)*

$$S_5 = \sum_{(i,j) \in \Omega} \psi_y(i,j) \cdot \psi_y(i,j) S_6 = \sum_{(i,j) \in \Omega} \theta(i,j) \cdot \psi_y(i,j),$$

trong đó Ω là tập hợp của các mẫu dự đoán lân cận bao quanh một khối con của phần dự đoán L_0 và L_1 và trong đó

$$\psi_x(i,j) = \left(\frac{\partial I^{(1)}}{\partial x}(i,j) + \frac{\partial I^{(0)}}{\partial x}(i,j) \right) \gg 3$$

$$\psi_y(i,j) = \left(\frac{\partial I^{(1)}}{\partial y}(i,j) + \frac{\partial I^{(0)}}{\partial y}(i,j) \right) \gg 3$$

*Phương
trình(3)*

$$\theta(i,j) = (I^{(1)}(i,j) \gg 6) - (I^{(0)}(i,j) \gg 6)$$

trong đó $I^{(k)}(i,j)$ là các mẫu dự đoán tại vị trí mẫu (i,j) của tín hiệu dự đoán trong danh sách k , $k = 0, 1$, vốn được sinh ra tại độ chính xác cao ở giữa (tức là, 16-bit); $\frac{\partial I^{(k)}}{\partial x}(i,j)$ và $\frac{\partial I^{(k)}}{\partial y}(i,j)$ là gradien nằm ngang và gradien thẳng đứng, theo cách tương ứng, của mẫu thu được bằng cách tính toán một cách trực tiếp khác biệt giữa hai mẫu lân cận của nó, tức là,

$$\frac{\partial I^{(k)}}{\partial x}(i,j) = (I^{(k)}(i+1,j) - I^{(k)}(i-1,j)) \gg 4$$

$$\frac{\partial I^{(k)}}{\partial y}(i,j) = (I^{(k)}(i,j+1) - I^{(k)}(i,j-1)) \gg 4$$

*Phương
trình(4)*

Dựa trên phần tinh chỉnh chuyển động được dẫn ra trong *Phương trình (1)*, các mẫu dự đoán đôi cuối cùng của CU được tính toán bằng ngoại suy phần dự đoán L_0 và phần dự đoán L_1 dọc theo quỹ đạo chuyển động dựa trên mô hình luồng quang học, như được chỉ thị bởi:

$$pred_{BDOF}(x,y) = (I^{(0)}(x,y) + I^{(1)}(x,y) + b + o_{offset}) \gg shift$$

*Phương
trình(5)*

$$b = rnd \left(\left(v_x \left(\frac{\partial I^{(1)}(x, y)}{\partial x} - \frac{\partial I^{(0)}(x, y)}{\partial x} \right) \right) / 2 \right) \\ + rnd \left(\left(v_y \left(\frac{\partial I^{(1)}(x, y)}{\partial y} - \frac{\partial I^{(0)}(x, y)}{\partial y} \right) \right) / 2 \right)$$

trong đó $shift$ và o_{offset} là trị số dịch chuyển phải và trị số dịch vị được áp dụng để kết hợp tín hiệu dự đoán L_0 và tín hiệu dự đoán L_1 cho việc dự đoán đôi, vốn bằng với $15 - BD$ và $1 \ll (14 - BD) + 2 \cdot (1 \ll 13)$, theo cách tương ứng. Bảng 1 minh họa các độ rộng bit cụ thể của các thông số trung gian tham gia vào trong quy trình BDOF. Như được thể hiện trên bảng 1, độ rộng bit bên trong của toàn bộ quy trình BDOF không vượt quá 32-bit. Thêm vào đó, việc nhân với đầu vào khả thi xấu nhất xảy ra tại tích của $v_x S_{2,m}$ trong *Phương trình (1)* với độ rộng bit đầu vào là 15-bit và 4-bit, theo cách tương ứng, for $S_{2,m}$ và v_x . Do đó, số nhân 15-bit là đủ cho BDOF.

Bảng 1: Các độ rộng bit của các thông số trung gian của BDOF trong VVC

Thao tác	Thông số	Độ rộng bit
Việc dự đoán L_0/L_1	$I^{(0)}(i, j), I^{(1)}(i, j)$	16
Độ lệch gradient	$\frac{\partial I^{(0)}}{\partial x}(i, j), \frac{\partial I^{(0)}}{\partial y}(i, j)$	13
	$\frac{\partial I^{(1)}}{\partial x}(i, j), \frac{\partial I^{(1)}}{\partial y}(i, j)$	13
Tính toán số tương quan	$\theta(i, j) = (I^{(1)}(i, j) \gg 6) - (I^{(0)}(i, j) \gg 6)$	11
	$\psi_x(i, j) = \left(\frac{\partial I^{(1)}}{\partial x}(i, j) + \frac{\partial I^{(0)}}{\partial x}(i, j) \right) \gg 3$	11
	$\psi_y(i, j) = \left(\frac{\partial I^{(1)}}{\partial y}(i, j) + \frac{\partial I^{(0)}}{\partial y}(i, j) \right) \gg 3$	11
	$\psi_x(i, j) \cdot \psi_x(i, j)$	21
	$\psi_x(i, j) \cdot \psi_y(i, j)$	21
	$\theta(i, j) \cdot \psi_x(i, j)$	21
	$\psi_y(i, j) \cdot \psi_y(i, j)$	21
	$\theta(i, j) \cdot \psi_y(i, j)$	21
	$S_1 = \sum_{(i, j) \in \Omega} \psi_x(i, j) \cdot \psi_x(i, j)$	27

	$S_2 = \sum_{(i,j) \in \Omega} \psi_x(i,j) \cdot \psi_y(i,j)$	27
	$S_3 = \sum_{(i,j) \in \Omega} \theta(i,j) \cdot \psi_x(i,j)$	27
	$S_5 = \sum_{(i,j) \in \Omega} \psi_y(i,j) \cdot \psi_y(i,j)$	27
	$S_6 = \sum_{(i,j) \in \Omega} \theta(i,j) \cdot \psi_y(i,j)$	27
Độ lệch phân tinh chỉnh chuyển động	$v_x = S_1 > 0? \text{clip3} \left(-th'_{BIO}, th'_{BIO}, -((S_3 \cdot 2^3) \gg \lfloor \log_2 S_1 \rfloor) \right) : 0$	4
	$S_{2,m} = S_2 \gg 12$	15
	$S_{2,s} = S_2 \& (2^{12} - 1)$	12
	$((v_x S_{2,m}) \ll 12 + v_x S_{2,s}) / 2$	30
	$S_6 \cdot 2^3 - ((v_x S_{2,m}) \ll 12 + v_x S_{2,s}) / 2$	31
	$v_y = S_5 > 0? \text{clip3} \left(-th'_{BIO}, th'_{BIO}, -((S_6 \cdot 2^3 - ((v_x S_{2,m}) \ll 12 + v_x S_{2,s}) / 2) \gg \lfloor \log_2 S_5 \rfloor) \right) : 0$	4
Việc sinh ra việc dự đoán đôi cuối cùng	$v_x \cdot \left(\frac{\partial I^{(1)}}{\partial x}(i,j) - \frac{\partial I^{(0)}}{\partial x}(i,j) \right)$	17
	$v_y \cdot \left(\frac{\partial I^{(1)}}{\partial y}(i,j) - \frac{\partial I^{(0)}}{\partial y}(i,j) \right)$	17

Mặc dù BDOF có thể tăng cường hiệu suất của việc dự đoán hai hướng, nhưng có cơ hội để tiếp tục cải thiện thiết kế của BDOF. Cụ thể hơn, các vấn đề sau đây trong thiết kế BDOF hiện đang tồn tại trong VVC để điều khiển các độ rộng bit của các thông số trung gian được nhận biết trong phần mô tả này.

Đầu tiên, như được thể hiện trên Bảng 1, thông số $\theta(i,j)$ (tức là, khác biệt giữa các phần dự đoán L0 và L1), và thông số $\psi_x(i,j)$ and $\psi_y(i,j)$ (tức là, tổng của trị số gradient nằm ngang và trị số gradient thẳng đứng cho L0 và L1) được biểu diễn trong cùng một độ rộng bit là 11-bit. Mặc dù phương pháp này có thể tạo thuận tiện cho việc

điều khiển tổng thể của độ rộng bit bên trong cho BDOF, nhưng phương pháp này có thể chưa được tối ưu liên quan tới độ chính xác của các phần tinh chỉnh chuyển động được dẫn ra. Đó là do như được thể hiện trên Phương trình (4), các trị số gradien được tính toán theo khác biệt giữa các phần dự đoán lân cận. Do bản chất thông cao của quy trình này, nên các gradien được dẫn ra là ít tin cậy hơn với sự có mặt của nhiễu, ví dụ nhiễu chụp được trong video gốc và nhiễu mã hóa được sinh ra trong suốt quá trình xử lý mã hóa. Nó có nghĩa là nó không phải lúc nào cũng cỡ lợi để biểu diễn các trị số gradien trong độ rộng bit cao.

Thêm vào đó, như được thể hiện trên Bảng 1, việc sử dụng độ rộng bit tối đa của toàn bộ quy trình BDOF xuất hiện với việc tính toán của phần tinh chỉnh chuyển động thẳng đứng vậy trong đó S6 (27-bit) được dịch chuyển trái đầu tiên bởi 3-bit và sau đó được trừ đi bởi $((vxS2,m) \ll 12 + vxS2,s)/2$ (30-bit). Do đó, độ rộng bit tối đa của thiết kế hiện tại là bằng với 31-bit. Trong áp dụng phần cứng thực tế, quy trình mã hóa với độ rộng bit bên trong tối đa lớn hơn 16-bit thường được áp dụng bởi việc áp dụng 32-bit. Do đó, thiết kế đang tồn tại không tận dụng được hoàn toàn khoảng giới hạn động hợp lệ của việc áp dụng 32-bit. Nó có thể dẫn tới tổn thất độ chính xác không cần thiết của các phần tinh chỉnh chuyển động được dẫn ra bởi BDOF.

Theo các phương án thực hiện làm ví dụ, như phương pháp làm ví dụ như được thể hiện trên Fig.5, phương pháp điều khiển độ rộng bit làm ví dụ giải quyết các vấn đề nêu trên của phương pháp điều khiển độ rộng bit cho thiết kế BDOF truyền thống.

Đầu tiên, để khắc phục các tác động có hại của các lỗi ước lượng gradien, dịch chuyển phải, bổ sung ngrad được đưa vào trong phương án thực hiện làm ví dụ khi tính toán các trị số gradien $\frac{\partial I^{(k)}}{\partial x}(i,j)$ và $\frac{\partial I^{(k)}}{\partial y}(i,j)$ trong Phương trình (4), tức là, để làm giảm độ rộng bit bên trong của các trị số gradien. Cụ thể hơn, các trị số gradien thứ nhất, tức là, gradien nằm ngang và gradien thẳng đứng, tại vị trí dự đoán thứ nhất L0 ($\frac{\partial I^{(0)}}{\partial x}(i,j)$ và $\frac{\partial I^{(0)}}{\partial y}(i,j)$) và các trị số gradien thứ hai, tức là, gradien nằm ngang và gradien thẳng đứng, tại vị trí dự đoán thứ hai L1 ($\frac{\partial I^{(1)}}{\partial x}(i,j)$ và $\frac{\partial I^{(1)}}{\partial y}(i,j)$) được tính toán theo:

$$\begin{aligned}\frac{\partial I^{(k)}}{\partial x}(i, j) &= \left(I^{(k)}(i+1, j) - I^{(k)}(i-1, j) \right) \gg (4 + n_{grad}), k = 0, 1 \\ \frac{\partial I^{(k)}}{\partial y}(i, j) &= \left(I^{(k)}(i, j+1) - I^{(k)}(i, j-1) \right) \gg (4 + n_{grad}), k = 0, 1\end{aligned}\quad \begin{array}{l} \text{Phương} \\ \text{trình(6)} \end{array}$$

Hơn nữa, dịch bit bổ sung n_{adj} được đưa vào tính toán của các biến $\psi_x(i, j)$, $\psi_y(i, j)$ và $\theta(i, j)$ để điều khiển toàn bộ quy trình BDOF sao cho quy trình BDOF được vận hành tại các độ rộng bit bên trong thích hợp, như được mô tả theo:

$$\begin{aligned}\psi_x(i, j) &= \left(\frac{\partial I^{(1)}}{\partial x}(i, j) + \frac{\partial I^{(0)}}{\partial x}(i, j) \right) \gg (3 - n_{adj}) \\ \psi_y(i, j) &= \left(\frac{\partial I^{(1)}}{\partial y}(i, j) + \frac{\partial I^{(0)}}{\partial y}(i, j) \right) \gg (3 - n_{adj}) \\ \theta(i, j) &= \left(I^{(1)}(i, j) \gg (6 - n_{adj}) \right) - \left(I^{(0)}(i, j) \gg (6 - n_{adj}) \right).\end{aligned}\quad \begin{array}{l} \text{Phương} \\ \text{trình(7)} \end{array}$$

Bảng 2 bên dưới thể hiện độ rộng bit tương ứng của từng thông số trung gian khi phương pháp điều khiển độ rộng bit làm ví dụ được áp dụng cho BDOF. Bảng 2 thể hiện các thay đổi được áp dụng trong phương pháp điều khiển độ rộng bit làm ví dụ khi so sánh với thiết kế BDOF truyền thống trong VVC. Như có thể thấy trên bảng 2, với phương pháp điều khiển độ rộng bit làm ví dụ, độ rộng bit bên trong của toàn bộ quy trình BDOF không vượt quá 32-bit. Thêm nữa, nhờ phương pháp làm ví dụ, độ rộng bit tối đa chỉ là 32-bit, vốn có thể tận dụng một cách đầy đủ khoảng giới hạn động sẵn có là 32-bit của ứng dụng phần cứng. Thêm nữa, như được thể hiện trên bảng 2, việc nhân với đầu vào khả thi kém nhất xuất hiện tại tích của $vxS2,m$ trong đó đầu vào $S2,m$ là 14-bit và đầu vào vx là 6-bit. Do đó, giống như các phương pháp BDOF truyền thống, một số nhân 16-bit là cũng đủ lớn khi phương pháp làm ví dụ được áp dụng.

Đề cập tới bảng 2, do việc biến đổi tới cho số các bit được dịch chuyển phải được áp dụng cho Phương trình (6) và Phương trình (7), nên các khoảng giới hạn động của các thông số $\psi_x(i, j)$, $\psi_y(i, j)$ và $\theta(i, j)$ là khác nhau, so với các phương pháp BDOF truyền thống được thể hiện trên Bảng 1, trong đó ba thông số được biểu diễn theo cùng

một khoảng giới hạn động (tức là, 21-bit). Thay đổi này có thể tăng các độ rộng bit của các thông số bên trong S1, S2, S3, S5, và S6, vốn có thể có tiềm năng làm tăng độ rộng bit tối đa của quy trình BDOF bên trong tới mức lớn hơn 32-bit. Do đó, để đảm bảo việc áp dụng 32-bit, hai thao tác cắt xén bổ sung được đưa vào trong tính toán các trị số của S2 và S6. Cụ thể hơn, theo phương pháp làm ví dụ, các trị số của hai thông số được tính toán theo:

$$S_2 = \text{Clip3} \left(-2^{B_2}, 2^{B_2} - 1, \sum_{(i,j) \in \Omega} \psi_x(i,j) \cdot \psi_y(i,j) \right) \quad \text{Phương}$$

$$S_6 = \text{Clip3} \left(-2^{B_6}, 2^{B_6} - 1, \sum_{(i,j) \in \Omega} \theta(i,j) \cdot \psi_y(i,j) \right), \quad \text{trình(8)}$$

trong đó B2 và B6 là các thông số để điều khiển các khoảng giới hạn động của đầu ra của S2 và S6, theo cách tương ứng. Cần chú ý rằng khác biệt từ tính toán gradient của các thao tác cắt xén trong Phương trình (8) chỉ được áp dụng một lần để tính toán phần tinh chỉnh chuyển động của từng khối con 4x4 nằm bên trong một BDOF CU, tức là, đang được gọi dựa trên đơn vị 4x4. Do đó, độ phức tạp tương ứng tăng do các thao tác cắt xén được đưa vào trong phương pháp được đề xuất sẽ là rất nhỏ.

Theo các phương án thực hiện làm ví dụ, việc dự đoán đôi cuối cùng của CU được tính toán dựa trên dự đoán thứ nhất L₀ và dự đoán thứ hai L₁, phần tinh chỉnh chuyển động (v_x, v_y) và các trị số gradient thứ nhất $\frac{\partial I^{(0)}}{\partial x}(i,j)$ và $\frac{\partial I^{(0)}}{\partial y}(i,j)$ và các trị số gradient thứ hai $\frac{\partial I^{(1)}}{\partial x}(i,j)$ và $\frac{\partial I^{(1)}}{\partial y}(i,j)$. Trong thực tế, các trị số khác biệt của n_{grad} , n_{adj} , B_2 , và B_6 có thể được áp dụng để đạt được các việc dàn xếp khác nhau giữa độ rộng bit trung gian và độ chính xác của các độ lệch BDOF bên trong. Theo một phương án thực hiện làm ví dụ, n_{grad} và n_{adj} được đặt là 2, B_2 được đặt là 25 và B_6 được đặt là 27. Theo các phương án thực hiện làm ví dụ cụ thể, trị số của n_c được đặt bằng với 2 và/hoặc th_{BDOF} được đặt bằng với 2 (xem *Phương trình (1)*), ví dụ.

Bảng 2: Các độ rộng bit của các thông số trực tiếp của phương pháp làm ví dụ

Việc vận hành	Thông số	Độ rộng bit
Việc dự đoán L ₀ /L ₁	$I^{(0)}(i,j), I^{(1)}(i,j)$	16
Độ lệch gradient	$\frac{\partial I^{(0)}}{\partial x}(i,j), \frac{\partial I^{(0)}}{\partial y}(i,j)$	11

	$\frac{\partial I^{(1)}}{\partial x}(i, j), \frac{\partial I^{(1)}}{\partial y}(i, j)$	11
Tính toán thông số tương quan	$\theta(i, j) = (I^{(1)}(i, j) \gg 4) - (I^{(0)}(i, j) \gg 4)$	13
	$\psi_x(i, j) = \left(\frac{\partial I^{(1)}}{\partial x}(i, j) + \frac{\partial I^{(0)}}{\partial x}(i, j) \right) \gg 1$	11
	$\psi_y(i, j) = \left(\frac{\partial I^{(1)}}{\partial y}(i, j) + \frac{\partial I^{(0)}}{\partial y}(i, j) \right) \gg 1$	11
	$\psi_x(i, j) \cdot \psi_x(i, j)$	21
	$\psi_x(i, j) \cdot \psi_y(i, j)$	21
	$\theta(i, j) \cdot \psi_x(i, j)$	23
	$\psi_y(i, j) \cdot \psi_y(i, j)$	21
	$\theta(i, j) \cdot \psi_y(i, j)$	23
Tổng	$S_1 = \sum_{(i,j) \in \Omega} \psi_x(i, j) \cdot \psi_x(i, j)$	27
	$S_2 = Clip3 \left(-2^{25}, 2^{25} - 1, \sum_{(i,j) \in \Omega} \psi_x(i, j) \cdot \psi_y(i, j) \right)$	26
	$S_3 = \sum_{(i,j) \in \Omega} \theta(i, j) \cdot \psi_x(i, j)$	29
	$S_5 = \sum_{(i,j) \in \Omega} \psi_y(i, j) \cdot \psi_y(i, j)$	27
	$S_6 = Clip3 \left(-2^{27}, 2^{27} - 1, \sum_{(i,j) \in \Omega} \theta(i, j) \cdot \psi_y(i, j) \right)$	28
Độ lệch phản tinh chỉnh chuyển động	$v_x = S_1 > 0 ? clip3 \left(-th'_{BDOF}, th'_{BDOF}, -((S_3 \cdot 2^3) \gg \lfloor \log_2 S_1 \rfloor) \right) : 0$	6
	$S_{2,m} = S_2 \gg 12$	14
	$S_{2,s} = S_2 \& (2^{12} - 1)$	12
	$((v_x S_{2,m}) \ll 12 + v_x S_{2,s}) / 2$	31
	$S_6 \cdot 2^3 - ((v_x S_{2,m}) \ll 12 + v_x S_{2,s}) / 2$	32
	$v_y = S_5 > 0 ? clip3 \left(-th'_{BDOF}, th'_{BDOF}, - \left((S_6 \cdot 2^3 - ((v_x S_{2,m}) \ll 12 + v_x S_{2,s}) / 2) \gg \lfloor \log_2 S_5 \rfloor \right) \right) : 0$	6

Việc sinh ra việc dự đoán đôi cuối cùng	$v_x \cdot \left(\frac{\partial I^{(1)}}{\partial x}(i, j) - \frac{\partial I^{(0)}}{\partial x}(i, j) \right)$	17
	$v_y \cdot \left(\frac{\partial I^{(1)}}{\partial y}(i, j) - \frac{\partial I^{(0)}}{\partial y}(i, j) \right)$	17

Fig.5 là lưu đồ của phương pháp làm ví dụ 300 để dự đoán khối của dữ liệu video theo các phương án thực hiện làm ví dụ. Phương pháp được thể hiện trên Fig.5 được giải thích liên quan tới bộ ghi mã 100 trên Fig.1 và/hoặc bộ giải mã 200 trên Fig.2. Tuy nhiên, theo các ví dụ khác, các thiết bị khác có thể được đặt cấu hình để thực hiện phương pháp được thể hiện trên Fig.5.

Theo các phương án thực hiện làm ví dụ, bộ ghi mã 100 ban đầu ghi mã, và sau đó giải mã, các hình ảnh để sử dụng như là các hình ảnh tham chiếu, được lưu trữ trong bộ nhớ hình ảnh tham chiếu, mặc dù các bước này không được minh họa trên phương pháp làm ví dụ của Fig.5. Bộ ghi mã 100 sau đó nhận khối của dữ liệu video. Khối có thể, ví dụ, là khối dự đoán (prediction unit - PU) của đơn vị mã hóa (coding unit - CU) của hình ảnh hiện tại (hoặc lát của hình ảnh) của dữ liệu video. Bộ ghi mã 100 có thể xác định chế độ nào trong số nhiều chế độ dự đoán để sử dụng để dự đoán khối, như dự đoán nội hoặc dự đoán liên thông, và nếu việc dự đoán liên thông được xác định, thì xem liệu sử dụng việc dự đoán liên thông một hướng hay hai hướng. Theo phương pháp làm ví dụ của Fig.5, bộ ghi mã 100 xác định để dự đoán khối sử dụng việc dự đoán đôi. Ví dụ, bộ ghi mã 100 có thể tính toán các số đo nhiễu loạn tỉ lệ cho các chế độ dự đoán khả thi khác nhau (ví dụ, một hoặc nhiều mô hình dự đoán nội và dự đoán một hướng hoặc hai hướng từ các hình ảnh tham chiếu được giải mã được lưu trữ trong bộ nhớ hình ảnh tham chiếu) và xác định rằng việc dự đoán đôi tạo ra các số đo nhiễu loạn tỉ lệ tốt nhất cho khối.

Bộ ghi mã 100, ví dụ, mạch ước lượng chuyển động 124, có thể còn tính toán các vector chuyển động cho khối theo việc dự đoán đôi. Các vector chuyển động này có thể đề cập tới hình ảnh tham chiếu thứ nhất cần được hiển thị trước hình ảnh hiện tại và hình ảnh tham chiếu thứ hai để được hiển thị sau hình ảnh hiện tại. Ví dụ, hình ảnh tham chiếu thứ nhất có thể xuất hiện trong danh sách hình ảnh tham chiếu 0 (RefPicList0) và

hình ảnh tham chiếu thứ hai có thể xuất hiện trong danh sách hình ảnh tham chiếu 1 (RefPicList1).

Theo một phương án thực hiện, bộ ghi mã 100 xác định khác biệt POC thứ nhất giữa trị số POC cho hình ảnh hiện tại và trị số POC cho hình ảnh tham chiếu thứ nhất, mà vectơ chuyển động thứ nhất (MV_0) đề cập tới đó, tại khoảng cách theo thời gian thứ nhất τ_0 , và khác biệt POC thứ hai giữa trị số POC cho hình ảnh hiện tại và trị số POC cho hình ảnh tham chiếu thứ hai, mà vectơ chuyển động thứ hai (MV_1) đề cập tới đó, tại khoảng cách theo thời gian thứ hai τ_1 . MV_0 và MV_1 có thể chứa thành phần x và thành phần y tương ứng. Ví dụ, MV_0 có thể chứa các thành phần MV_{x0} và MV_{y0} , trong đó MV_{x0} mô tả dịch vị nằm ngang và MV_{y0} mô tả dịch vị thẳng đứng liên quan tới vị trí của khối trong hình ảnh hiện tại để xác định vị trí mà khối tham chiếu trong hình ảnh tham chiếu thứ nhất nằm tại đó. Theo cách tương tự, MV_1 có thể chứa các thành phần MV_{x1} và MV_{y1} , trong đó MV_{x1} mô tả dịch vị nằm ngang và MV_{y1} mô tả dịch vị thẳng đứng liên quan tới vị trí của khối trong hình ảnh hiện tại để xác định vị trí mà khối tham chiếu trong hình ảnh tham chiếu thứ hai nằm tại đó.

Theo các phương án thực hiện làm ví dụ, hình ảnh tham chiếu thứ nhất và hình ảnh tham chiếu thứ hai được giải mã 302. Hình ảnh tham chiếu thứ hai là khác với hình ảnh tham chiếu thứ nhất, trong đó hình ảnh tham chiếu thứ nhất là được hiển thị trước hình ảnh hiện tại và hình ảnh tham chiếu thứ hai được hiển thị sau hình ảnh hiện tại; Theo các phương án thực hiện làm ví dụ, vectơ chuyển động thứ nhất (MV_0) từ khối hiện tại của hình ảnh hiện tại tới dự đoán thứ nhất L_0 của hình ảnh tham chiếu thứ nhất được xác định 304 và vectơ chuyển động thứ hai (MV_1) từ khối hiện tại tới dự đoán thứ hai L_1 của hình ảnh tham chiếu thứ hai được xác định 306. Theo các phương án thực hiện làm ví dụ cụ thể, vectơ chuyển động thứ nhất (MV_0) từ khối hiện tại của hình ảnh hiện tại tới dự đoán thứ nhất L_0 của hình ảnh tham chiếu thứ nhất tại khoảng cách theo thời gian thứ nhất (τ_0) từ hình ảnh hiện tại tới hình ảnh tham chiếu thứ nhất được xác định và vectơ chuyển động thứ hai (MV_1) từ khối hiện tại tới dự đoán thứ hai L_1 của hình ảnh tham chiếu thứ hai tại khoảng cách theo thời gian thứ hai (τ_1) từ hình ảnh hiện tại tới hình ảnh tham chiếu thứ hai được xác định.

Bộ ghi mã 100 sau đó sử dụng các trị số để tính toán 308 phần tinh chỉnh chuyển động (v_x, v_y) của đơn vị mã hóa (coding unit - CU) bằng cách tối thiểu hóa sự chênh lệch giữa dự đoán thứ nhất L_0 và dự đoán thứ hai L_1 . Theo các phương án thực hiện làm ví dụ, các trị số gradient thứ nhất $\frac{\partial I^{(0)}}{\partial x}(i, j)$ và $\frac{\partial I^{(0)}}{\partial y}(i, j)$ cho dự đoán thứ nhất L_0 và các trị số gradient thứ hai $\frac{\partial I^{(1)}}{\partial x}(i, j)$ và $\frac{\partial I^{(1)}}{\partial y}(i, j)$ cho dự đoán thứ hai L_1 được tính toán 310. Theo các phương án thực hiện làm ví dụ, các trị số gradient thứ nhất và các trị số gradient thứ hai sau đó có thể được tính toán bằng cách đưa dịch chuyển phải, bổ sung n_{grad} để làm giảm độ rộng bit bên trong của các trị số gradient thứ nhất và các trị số gradient thứ hai, như được thể hiện trên *Phương trình (6)*. Theo các phương án thực hiện làm ví dụ, tính toán 310 trị số gradient thứ nhất và trị số gradient thứ hai chứa tính toán gradient nằm ngang và gradient thẳng đứng tại từng vị trí trong số vị trí dự đoán thứ nhất của dự đoán thứ nhất và vị trí dự đoán thứ hai của dự đoán thứ hai. Đề cập tới *Phương trình (6)* gradient nằm ngang được tính toán tại từng vị trí trong số vị trí dự đoán thứ nhất và vị trí dự đoán thứ hai, và gradient thẳng đứng được tính toán tại từng vị trí trong số vị trí dự đoán thứ nhất và vị trí dự đoán thứ hai.

Theo các phương án thực hiện làm ví dụ, trị số của (v_x, v_y) được tính toán sử dụng *Phương trình (1)*, trong đó dịch bit bổ sung n_{adj} được đưa vào tính toán của nhiều biến $\psi_x(i, j)$, $\psi_y(i, j)$, và $\theta(i, j)$, như được thể hiện trên *Phương trình (7)*, để tính toán trị số của S_1, S_2, S_3, S_5 , và S_6 . Ví dụ, trị số của S_2 và S_6 được tính toán sử dụng *Phương trình (8)*. Theo các phương án thực hiện làm ví dụ, trị số của S_2 và S_6 được tính toán bằng cách đưa vào thao tác cắt xén thứ nhất và thao tác cắt xén thứ hai như được thể hiện trên *Phương trình (8)*, trong đó B_2 là thông số để điều khiển khoảng giới hạn động của đầu ra của S_2 và B_6 là thông số để điều khiển khoảng giới hạn động của đầu ra của S_6 . Việc dự đoán đôi cuối cùng của CU sau đó có thể được tính toán 312 dựa trên dự đoán thứ nhất L_0 và dự đoán thứ hai L_1 , phần tinh chỉnh chuyển động (v_x, v_y), các trị số gradient thứ nhất $\frac{\partial I^{(0)}}{\partial x}(i, j)$ và $\frac{\partial I^{(0)}}{\partial y}(i, j)$, và các trị số gradient thứ hai $\frac{\partial I^{(1)}}{\partial x}(i, j)$ và $\frac{\partial I^{(1)}}{\partial y}(i, j)$. Theo các phương án thực hiện làm ví dụ cụ thể, các mẫu việc dự đoán đôi cuối cùng của CU sau đó có thể được tính toán dựa trên phần tinh chỉnh chuyển động bằng cách nội suy dự đoán thứ nhất L_0 và/hoặc dự đoán thứ hai L_1 dọc theo quỹ đạo chuyển động, như

được chỉ thị trong *Phương trình (5)*. Theo các phương án thực hiện làm ví dụ, tổng của trị số gradien thứ nhất và trị số gradien thứ hai được tính toán, trong đó dịch chuyển phải, bổ sung được đưa vào tổng để làm giảm độ rộng bit của nó. Khác biệt giữa mẫu dự đoán thứ nhất L_0 và mẫu dự đoán thứ hai L_1 được tính toán, trong đó dịch chuyển phải, bổ sung được đưa vào mẫu dự đoán thứ nhất L_0 và mẫu dự đoán thứ hai L_1 để làm giảm độ rộng bit bên trong của khác biệt.

Phương pháp làm ví dụ được thể hiện trên Fig.5 biểu diễn ví dụ của phương pháp ghi mã dữ liệu video chứa bước giải mã hình ảnh tham chiếu thứ nhất và hình ảnh tham chiếu thứ hai, hình ảnh tham chiếu thứ hai đang khác với hình ảnh tham chiếu thứ nhất, trong đó hình ảnh tham chiếu thứ nhất được hiển thị trước hình ảnh hiện tại và hình ảnh tham chiếu thứ hai được hiển thị sau hình ảnh hiện tại, xác định vector chuyển động thứ nhất (MV_0) từ khối hiện tại của hình ảnh hiện tại tới khối tham chiếu thứ nhất của hình ảnh tham chiếu thứ nhất, xác định vector chuyển động thứ hai (MV_1) từ khối hiện tại tới khối tham chiếu thứ hai của hình ảnh tham chiếu thứ hai, và ghi mã khối hiện tại theo luồng quang học hai chiều (bi-directional optical flow - BDOF).

Theo một hoặc nhiều ví dụ, các chức năng được mô tả có thể được áp dụng trong phần cứng, phần mềm, phần sụn, hoặc tổ hợp bất kỳ của chúng. Nếu được áp dụng trong phần mềm, các chức năng có thể được lưu trữ hoặc được truyền qua, như là một hoặc nhiều lệnh hoặc mã, môi trường đọc được bởi máy tính và được thực thi bởi đơn vị xử lý dựa trên phần cứng. Phương tiện đọc được bởi máy tính có thể chứa vật ghi đọc được bởi máy tính, vốn tương ứng với vật ghi hữu hình như vật ghi lưu trữ dữ liệu, hoặc môi trường liên lạc chứa môi trường bất kỳ tạo thuận tiện cho việc truyền chương trình máy tính từ một vị trí này tới vị trí khác, ví dụ, theo giao thức truyền thông. Theo cách này, môi trường đọc được bởi máy tính nói chung có thể tương ứng với (1) vật ghi hữu hình đọc được bởi máy tính vốn là phi chuyển tiếp hoặc (2) môi trường liên lạc như tín hiệu hoặc sóng mang. Môi trường lưu trữ dữ liệu có thể là môi trường sẵn có bất kỳ có thể được truy cập bởi một hoặc nhiều máy tính hoặc một hoặc nhiều bộ xử lý để truy hồi các lệnh, các mã và/hoặc các cấu trúc dữ liệu để áp dụng của các việc áp dụng được mô tả trong sáng chế này. Sản phẩm chương trình máy tính có thể chứa môi trường đọc được bởi máy tính.

Hơn nữa, các phương pháp nêu trên có thể được áp dụng sử dụng thiết bị chứa một hoặc nhiều mạch điện, vốn chứa các mạch tích hợp ứng dụng cụ thể (application specific integrated circuit - ASIC), các bộ xử lý tín hiệu số (digital signal processor - DSP), các thiết bị xử lý tín hiệu số (digital signal processing device - DSPD), các thiết bị logic lập trình được (programmable logic device - PLD), các mảng cổng lập trình được dạng trường (field programmable gate array - FPGA), các bộ điều khiển, các bộ vi điều khiển, các bộ vi xử lý, hoặc các thành phần điện tử khác. Thiết bị có thể sử dụng các mạch điện trong kết hợp với các thành phần phần cứng hoặc phần mềm để thực hiện các phương pháp được mô tả ở trên. Mỗi môđun, môđun con, hoặc đơn vị con được bộc lộ ở trên có thể được áp dụng ít nhất một phần sử dụng một hoặc nhiều mạch điện.

Các phương án thực hiện khác của sáng chế sẽ là rõ ràng đối với người có hiểu biết trung bình trong lĩnh vực kỹ thuật từ việc xem xét của phần mô tả này và việc thực hiện sáng chế được bộc lộ ở đây. Sáng chế này nhằm mục đích che phủ các biến thể, các sử dụng, hoặc các thích ứng bất kỳ của sáng chế tuân theo các nguyên lý chung của chúng và chứa các biến thể này, từ phần bộc lộ kèm theo khi nằm trong việc thực hiện đã biết hoặc được tùy chỉnh trong tình trạng kỹ thuật của sáng chế. Mục đích là bản mô tả và các ví dụ chỉ được xem xét như là ví dụ, với phạm vi bảo hộ thực tế và nguyên lý của sáng chế được chỉ ra trong các yêu cầu bảo hộ dưới đây.

Sẽ hiểu rằng sáng chế này không bị giới hạn ở các ví dụ chính xác được mô tả ở trên và được minh họa trong các hình vẽ kèm theo, và các sự cải biến hoặc các sự thay đổi khác nhau có thể được tạo ra mà không lệch khỏi phạm vi của chúng. Điều này nhằm mục đích để phạm vi bảo hộ của sáng chế chỉ bị hạn chế bởi các yêu cầu bảo hộ kèm theo.

YÊU CẦU BẢO HỘ

1. Phương pháp giải mã video, phương pháp này bao gồm các bước:

giải mã hình ảnh tham chiếu thứ nhất và hình ảnh tham chiếu thứ hai, hình ảnh tham chiếu thứ hai khác với hình ảnh tham chiếu thứ nhất, trong đó hình ảnh tham chiếu thứ nhất được hiển thị trước hình ảnh hiện tại và hình ảnh tham chiếu thứ hai được hiển thị sau hình ảnh hiện tại;

xác định, đối với khối hiện tại của hình ảnh hiện tại, dự đoán thứ nhất L_0 tương ứng với hình ảnh tham chiếu thứ nhất;

xác định, đối với khối hiện tại, dự đoán thứ hai L_1 tương ứng với hình ảnh tham chiếu thứ hai;

tính toán phần tinh chỉnh chuyển động (v_x, v_y) tương ứng của mỗi khối con của khối hiện tại bằng cách tối thiểu hóa sự chênh lệch giữa dự đoán thứ nhất L_0 và dự đoán thứ hai L_1 dựa trên ngưỡng tinh chỉnh chuyển động và nhiều thông số bên trong đại diện cho các ký tự tính tổng khác nhau của dự đoán thứ nhất L_0 và dự đoán thứ hai L_1 ;

tính toán các trị số gradient thứ nhất $\frac{\partial I^{(0)}}{\partial x}(i, j)$ và $\frac{\partial I^{(0)}}{\partial y}(i, j)$ đối với dự đoán thứ nhất L_0 và các trị số gradient thứ hai $\frac{\partial I^{(1)}}{\partial x}(i, j)$ và $\frac{\partial I^{(1)}}{\partial y}(i, j)$ đối với dự đoán thứ hai L_1 , trong đó $I^{(0)}(i, j)$ là mẫu dự đoán tại vị trí mẫu (i, j) của dự đoán thứ nhất L_0 , và $I^{(1)}(i, j)$ là mẫu dự đoán tại vị trí mẫu (i, j) của dự đoán thứ hai L_1 ; và

tính toán việc dự đoán đôi cuối cùng của khối hiện tại dựa trên dự đoán thứ nhất L_0 , dự đoán thứ hai L_1 , các phần tinh chỉnh chuyển động (v_x, v_y) tương ứng, các trị số gradient thứ nhất $\frac{\partial I^{(0)}}{\partial x}(i, j)$ và $\frac{\partial I^{(0)}}{\partial y}(i, j)$, và các trị số gradient thứ hai $\frac{\partial I^{(1)}}{\partial x}(i, j)$ và $\frac{\partial I^{(1)}}{\partial y}(i, j)$.

2. Phương pháp theo điểm 1, trong đó trị số của (v_x, v_y) được tính toán dựa trên các phương trình sau đây:

$$v_x = S_1 > 0? \text{clip3}(-th_{BDOF}, th_{BDOF}, -((S_3 \cdot 2^{n_c}) \gg \lfloor \log_2 S_1 \rfloor)) : 0$$

$$v_y = S_5 > 0? \text{clip3} \left(-th_{BDOF}, th_{BDOF}, - \left((S_6 \cdot 2^{n_c} - (v_x S_{2,m}) \ll n_{S_2} + v_x S_{2,s}) / 2 \right) \gg \lfloor \log_2 S_5 \rfloor \right) : 0,$$

trong đó th_{BDOF} là ngưỡng tình chỉnh chuyển động, $S_{2,m} = S_2 \gg n_{S_2}$, $S_{2,s} = S_2 \& (2^{n_{S_2}} - 1)$, S_1, S_2, S_3, S_5 và S_6 là các thông số bên trong đại diện cho các ký tự tính tổng khác nhau của dự đoán thứ nhất L_0 và dự đoán thứ hai L_1 , và n_{S_2} và n_c là các số nguyên.

3. Phương pháp theo điểm 2, trong đó trị số của S_1 được tính toán dựa trên $\psi_x(i, j)$ đối với tập hợp của các mẫu dự đoán xung quanh khối con tương ứng của mỗi dự đoán trong số dự đoán thứ nhất L_0 và dự đoán thứ hai L_1 , trong đó trị số của S_2 được tính toán dựa trên $\psi_x(i, j)$ và $\psi_y(i, j)$ đối với tập hợp của các mẫu dự đoán, trong đó trị số của S_3 được tính toán dựa trên $\theta(i, j)$ và $\psi_x(i, j)$ đối với tập hợp của các mẫu dự đoán, trong đó trị số của S_5 được tính toán dựa trên $\psi_y(i, j)$ đối với tập hợp của các mẫu dự đoán, và trong đó trị số của S_6 được tính toán dựa trên $\theta(i, j)$ và $\psi_y(i, j)$ đối với tập hợp của các mẫu dự đoán,

trong đó $\psi_x(i, j)$, $\psi_y(i, j)$ và $\theta(i, j)$ là các thông số tương quan, trong đó $\psi_x(i, j)$ được tính toán dựa trên trị số gradient nằm ngang thứ nhất của dự đoán thứ nhất L_0 và trị số gradient nằm ngang thứ hai của dự đoán thứ hai L_1 , $\psi_y(i, j)$ được tính toán dựa trên trị số gradient thẳng đứng thứ nhất của dự đoán thứ nhất L_0 và trị số gradient thẳng đứng thứ hai của dự đoán thứ hai L_1 , và $\theta(i, j)$ được tính toán dựa trên dự đoán thứ nhất L_0 và dự đoán thứ hai L_1 .

4. Phương pháp theo điểm 2, trong đó trị số của n_c được đặt bằng 2.

5. Phương pháp theo điểm 3, phương pháp này còn bao gồm bước đưa dịch bit bổ sung n_{adj} vào việc tính toán của $\psi_x(i, j)$, $\psi_y(i, j)$, và $\theta(i, j)$, trong đó:

$$\psi_x(i, j) = \left(\frac{\partial I^{(1)}}{\partial x}(i, j) + \frac{\partial I^{(0)}}{\partial x}(i, j) \right) \gg (3 - n_{adj});$$

$$\psi_y(i, j) = \left(\frac{\partial I^{(1)}}{\partial y}(i, j) + \frac{\partial I^{(0)}}{\partial y}(i, j) \right) \gg (3 - n_{adj}); \text{ và}$$

$$\theta(i, j) = \left(I^{(1)}(i, j) \gg (6 - n_{adj}) \right) - \left(I^{(0)}(i, j) \gg (6 - n_{adj}) \right).$$

6. Phương pháp theo điểm 5, trong đó trị số của n_{adj} được đặt bằng 2.

7. Phương pháp theo điểm 5, trong đó trị số của S_2 được tính toán dựa trên $\psi_x(i, j) \cdot \psi_y(i, j)$ đối với tập hợp của các mẫu dự đoán, và trong đó trị số của S_6 được tính toán dựa trên $\theta(i, j) \cdot \psi_y(i, j)$ đối với tập hợp của các mẫu dự đoán.

8. Phương pháp theo điểm 1, trong đó bước tính toán trị số gradient $\frac{\partial I^{(k)}}{\partial x}(i, j), k = 0, 1$ và trị số gradient $\frac{\partial I^{(k)}}{\partial y}(i, j), k = 0, 1$ còn bao gồm tính toán gradient nằm ngang và gradient thẳng đứng tại từng vị trí trong số vị trí dự đoán thứ nhất của dự đoán thứ nhất L_0 và vị trí dự đoán thứ hai của dự đoán thứ hai L_1 .

9. Phương pháp theo điểm 8, trong đó gradient nằm ngang được tính toán tại từng vị trí trong số vị trí dự đoán thứ nhất và vị trí dự đoán thứ hai theo:

$$\frac{\partial I^{(k)}}{\partial x}(i, j) = \left(I^{(k)}(i + 1, j) - I^{(k)}(i - 1, j) \right) \gg (4 + n_{grad}), k = 0, 1;$$

và gradient thẳng đứng được tính toán tại từng vị trí trong số vị trí dự đoán thứ nhất và vị trí dự đoán thứ hai theo:

$$\frac{\partial I^{(k)}}{\partial y}(i, j) = \left(I^{(k)}(i, j + 1) - I^{(k)}(i, j - 1) \right) \gg (4 + n_{grad}), k = 0, 1,$$

trong đó n_{grad} là dịch chuyển phải bổ sung.

10. Phương pháp theo điểm 1, trong đó bước tính toán việc dự đoán đôi cuối cùng của khối hiện tại bao gồm:

tính toán việc dự đoán đôi cuối cùng của khối hiện tại còn dựa trên trị số dịch vị O_{offset} để kết hợp dự đoán thứ nhất L_0 và dự đoán thứ hai L_1 .

11. Phương pháp theo điểm 1, trong đó dự đoán thứ nhất L_0 được xác định dựa trên vector chuyển động thứ nhất (MV0) từ khối hiện tại đến hình ảnh tham chiếu thứ nhất,

và trong đó dự đoán thứ hai L_1 được xác định dựa trên vector chuyển động thứ hai (MV1) từ khối hiện tại đến hình ảnh tham chiếu thứ hai.

12. Thiết bị tính toán, thiết bị này bao gồm:

một hoặc nhiều bộ xử lý;

bộ lưu trữ phi chuyển tiếp được ghép nối với một hoặc nhiều bộ xử lý; và

nhiều chương trình được lưu trữ trong bộ lưu trữ phi chuyển tiếp mà, khi được thực thi bởi một hoặc nhiều bộ xử lý, làm cho thiết bị tính toán thực hiện các hoạt động bao gồm:

giải mã hình ảnh tham chiếu thứ nhất và hình ảnh tham chiếu thứ hai, hình ảnh tham chiếu thứ hai khác với hình ảnh tham chiếu thứ nhất, trong đó hình ảnh tham chiếu thứ nhất được hiển thị trước hình ảnh hiện tại và hình ảnh tham chiếu thứ hai được hiển thị sau hình ảnh hiện tại;

xác định, đối với khối hiện tại của hình ảnh hiện tại, dự đoán thứ nhất L_0 tương ứng với hình ảnh tham chiếu thứ nhất;

xác định, đối với khối hiện tại, dự đoán thứ hai L_1 tương ứng với hình ảnh tham chiếu thứ hai;

tính toán phần tinh chỉnh chuyển động (v_x, v_y) tương ứng của mỗi khối con của khối hiện tại bằng cách tối thiểu hóa sự chênh lệch giữa dự đoán thứ nhất L_0 và dự đoán thứ hai L_1 dựa trên ngưỡng tinh chỉnh chuyển động và nhiều các thông số bên trong đại diện cho các ký tự tính tổng khác nhau của dự đoán thứ nhất L_0 và dự đoán thứ hai L_1 ;

tính toán các trị số gradient thứ nhất $\frac{\partial I^{(0)}}{\partial x}(i, j)$ và $\frac{\partial I^{(0)}}{\partial y}(i, j)$ đối với dự đoán thứ nhất L_0 và các trị số gradient thứ hai $\frac{\partial I^{(1)}}{\partial x}(i, j)$ và $\frac{\partial I^{(1)}}{\partial y}(i, j)$ đối với dự đoán thứ hai L_1 , trong đó $I^{(0)}(i, j)$ là mẫu dự đoán tại vị trí mẫu (i, j) của dự đoán thứ nhất L_0 , và $I^{(1)}(i, j)$ là mẫu dự đoán tại vị trí mẫu (i, j) của dự đoán thứ hai L_1 ; và

tính toán việc dự đoán đôi cuối cùng của khối hiện tại dựa trên dự đoán thứ nhất L_0 , dự đoán thứ hai L_1 , các phần tinh chỉnh chuyển động (v_x, v_y) tương ứng, các trị số

gradien thứ nhất $\frac{\partial I^{(0)}}{\partial x}(i, j)$ và $\frac{\partial I^{(0)}}{\partial y}(i, j)$, và các trị số gradien thứ hai $\frac{\partial I^{(1)}}{\partial x}(i, j)$ và $\frac{\partial I^{(1)}}{\partial y}(i, j)$.

13. Thiết bị tính toán theo điểm 12, trong đó trị số của (v_x, v_y) được tính toán dựa trên các phương trình sau đây:

$$v_x = S_1 > 0? \text{clip3} \left(-th_{BDOF}, th_{BDOF}, -((S_3 \cdot 2^{n_c}) \gg \lfloor \log_2 S_1 \rfloor) \right) : 0$$

$$v_y = S_5 > 0? \text{clip3} \left(-th_{BDOF}, th_{BDOF}, - \left(\left(S_6 \cdot 2^{n_c} - ((v_x S_{2,m}) \ll n_{S_2} + v_x S_{2,s}) / 2 \right) \gg \lfloor \log_2 S_5 \rfloor \right) \right) : 0,$$

trong đó th_{BDOF} là ngưỡng tinh chỉnh chuyển động, $S_{2,m} = S_2 \gg n_{S_2}$, $S_{2,s} = S_2 \& (2^{n_{S_2}} - 1)$, S_1, S_2, S_3, S_5 và S_6 là các thông số bên trong đại diện cho các ký tự tính tổng khác nhau của dự đoán thứ nhất L_0 và dự đoán thứ hai L_1 , và n_{S_2} và n_c là các số nguyên.

14. Thiết bị tính toán theo điểm 13, trong đó trị số của S_1 được tính toán dựa trên $\psi_x(i, j)$ đối với tập hợp của các mẫu dự đoán xung quanh khối con tương ứng của mỗi dự đoán trong số dự đoán thứ nhất L_0 và dự đoán thứ hai L_1 , trong đó trị số của S_2 được tính toán dựa trên $\psi_x(i, j)$ và $\psi_y(i, j)$ đối với tập hợp của các mẫu dự đoán, trong đó trị số của S_3 được tính toán dựa trên $\theta(i, j)$ và $\psi_x(i, j)$ đối với tập hợp của các mẫu dự đoán, trong đó trị số của S_5 được tính toán dựa trên $\psi_y(i, j)$ đối với tập hợp của các mẫu dự đoán, và trong đó trị số của S_6 được tính toán dựa trên $\theta(i, j)$ và $\psi_y(i, j)$ đối với tập hợp của các mẫu dự đoán,

trong đó $\psi_x(i, j)$, $\psi_y(i, j)$ và $\theta(i, j)$ là các thông số tương quan, trong đó $\psi_x(i, j)$ được tính toán dựa trên trị số gradien nằm ngang thứ nhất của dự đoán thứ nhất L_0 và trị số gradien nằm ngang thứ hai của dự đoán thứ hai L_1 , $\psi_y(i, j)$ được tính toán dựa trên trị số gradien thẳng đứng thứ nhất của dự đoán thứ nhất L_0 và trị số gradien thẳng

đứng thứ hai của dự đoán thứ hai L_1 , và $\theta(i, j)$ được tính toán dựa trên dự đoán thứ nhất L_0 và dự đoán thứ hai L_1 .

15. Thiết bị tính toán theo điểm 13, trong đó trị số của n_c được đặt bằng 2.

16. Thiết bị tính toán theo điểm 14, còn bao gồm việc đưa dịch bit bổ sung n_{adj} vào việc tính toán của $\psi_x(i, j)$, $\psi_y(i, j)$, và $\theta(i, j)$, trong đó:

$$\psi_x(i, j) = \left(\frac{\partial I^{(1)}}{\partial x}(i, j) + \frac{\partial I^{(0)}}{\partial x}(i, j) \right) \gg (3 - n_{adj});$$

$$\psi_y(i, j) = \left(\frac{\partial I^{(1)}}{\partial y}(i, j) + \frac{\partial I^{(0)}}{\partial y}(i, j) \right) \gg (3 - n_{adj}); \text{ và}$$

$$\theta(i, j) = \left(I^{(1)}(i, j) \gg (6 - n_{adj}) \right) - \left(I^{(0)}(i, j) \gg (6 - n_{adj}) \right).$$

17. Thiết bị tính toán theo điểm 16, trong đó trị số của n_{adj} được đặt bằng 2.

18. Thiết bị tính toán theo điểm 16, trong đó trị số của S_2 được tính toán dựa trên $\psi_x(i, j) \cdot \psi_y(i, j)$ đối với tập hợp của các mẫu dự đoán, và trong đó trị số của S_6 được tính toán dựa trên $\theta(i, j) \cdot \psi_y(i, j)$ đối với tập hợp của các mẫu dự đoán.

19. Thiết bị tính toán theo điểm 12, trong đó việc tính toán trị số gradien $\frac{\partial I^{(k)}}{\partial x}(i, j)$, $k = 0, 1$ và trị số gradien $\frac{\partial I^{(k)}}{\partial y}(i, j)$, $k = 0, 1$ còn bao gồm tính toán gradien nằm ngang và gradien thẳng đứng tại từng vị trí trong số vị trí dự đoán thứ nhất của dự đoán thứ nhất L_0 và vị trí dự đoán thứ hai của dự đoán thứ hai L_1 .

20. Thiết bị tính toán theo điểm 19, trong đó gradien nằm ngang được tính toán tại từng vị trí trong số vị trí dự đoán thứ nhất và vị trí dự đoán thứ hai theo:

$$\frac{\partial I^{(k)}}{\partial x}(i, j) = \left(I^{(k)}(i + 1, j) - I^{(k)}(i - 1, j) \right) \gg (4 + n_{grad}), k = 0, 1;$$

và gradien thẳng đứng được tính toán tại từng vị trí trong số vị trí dự đoán thứ nhất và vị trí dự đoán thứ hai theo:

$$\frac{\partial I^{(k)}}{\partial y}(i, j) = \left(I^{(k)}(i, j+1) - I^{(k)}(i, j-1) \right) \gg (4 + n_{grad}), k = 0, 1,$$

trong đó n_{grad} là dịch chuyển phải bỏ sung.

21. Vật ghi phi chuyển tiếp đọc được bằng máy tính lưu trữ nhiều chương trình để thực thi bởi thiết bị tính toán có một hoặc nhiều bộ xử lý, trong đó các chương trình, khi được thực thi bởi một hoặc nhiều bộ xử lý, làm cho thiết bị tính toán thực hiện các hoạt động bao gồm:

giải mã hình ảnh tham chiếu thứ nhất và hình ảnh tham chiếu thứ hai, hình ảnh tham chiếu thứ hai khác với hình ảnh tham chiếu thứ nhất, trong đó hình ảnh tham chiếu thứ nhất được hiển thị trước hình ảnh hiện tại và hình ảnh tham chiếu thứ hai được hiển thị sau hình ảnh hiện tại;

xác định, đối với khối hiện tại của hình ảnh hiện tại, dự đoán thứ nhất L_0 tương ứng với hình ảnh tham chiếu thứ nhất;

xác định, đối với khối hiện tại, dự đoán thứ hai L_1 tương ứng với hình ảnh tham chiếu thứ hai;

tính toán phân tích chỉnh chuyển động (v_x, v_y) tương ứng của mỗi khối con của khối hiện tại bằng cách tối thiểu hóa sự chênh lệch giữa dự đoán thứ nhất L_0 và dự đoán thứ hai L_1 dựa trên ngưỡng phân tích chỉnh chuyển động và nhiều thông số bên trong đại diện cho các ký tự tính tổng khác nhau của dự đoán thứ nhất L_0 và dự đoán thứ hai L_1 ;

tính toán các trị số gradient thứ nhất $\frac{\partial I^{(0)}}{\partial x}(i, j)$ và $\frac{\partial I^{(0)}}{\partial y}(i, j)$ đối với dự đoán thứ nhất L_0 và các trị số gradient thứ hai $\frac{\partial I^{(1)}}{\partial x}(i, j)$ và $\frac{\partial I^{(1)}}{\partial y}(i, j)$ đối với dự đoán thứ hai L_1 , trong đó $I^{(0)}(i, j)$ là mẫu dự đoán tại vị trí mẫu (i, j) của dự đoán thứ nhất L_0 , và $I^{(1)}(i, j)$ là mẫu dự đoán tại vị trí mẫu (i, j) của dự đoán thứ hai L_1 ; và

tính toán việc dự đoán đôi cuối cùng của khối hiện tại dựa trên dự đoán thứ nhất L_0 , dự đoán thứ hai L_1 , các phân tích chỉnh chuyển động (v_x, v_y) tương ứng, các trị số

gradient thứ nhất $\frac{\partial I^{(0)}}{\partial x}(i, j)$ và $\frac{\partial I^{(0)}}{\partial y}(i, j)$, và các trị số gradient thứ hai $\frac{\partial I^{(1)}}{\partial x}(i, j)$ và $\frac{\partial I^{(1)}}{\partial y}(i, j)$.

2 / 4

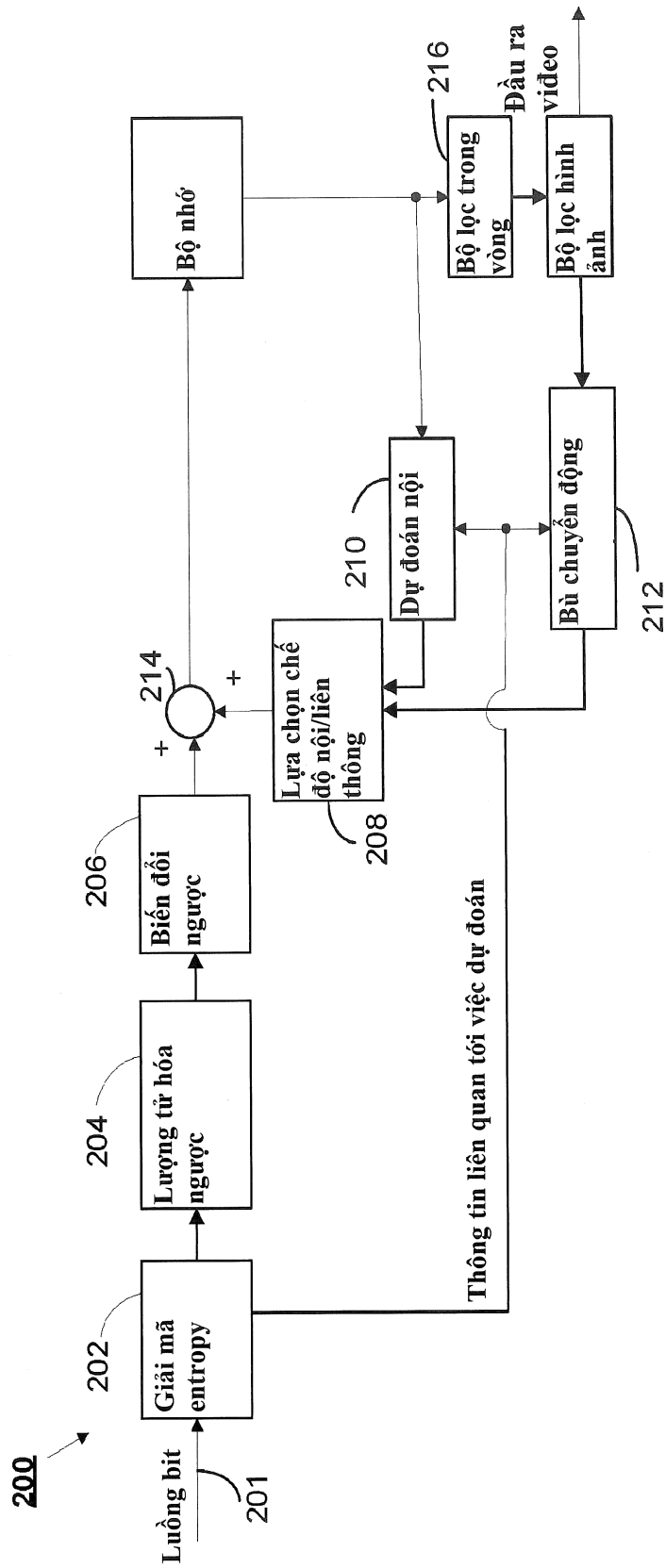


FIG. 2

3 / 4

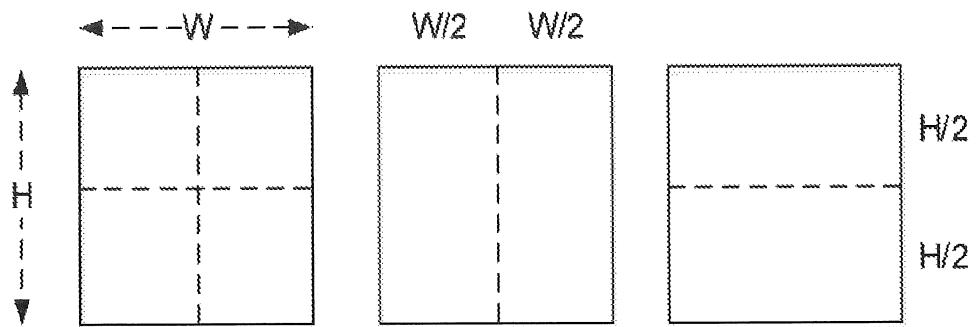


FIG. 3A

FIG. 3B

FIG. 3C

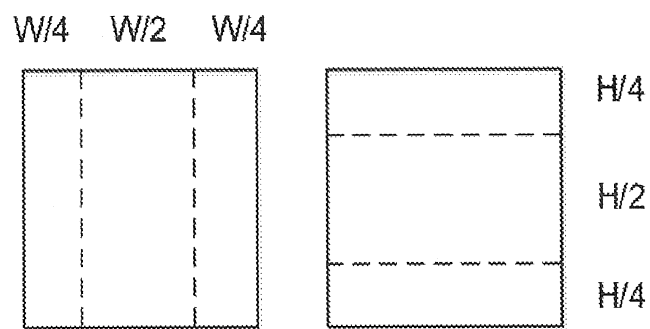


FIG. 3D

FIG. 3E

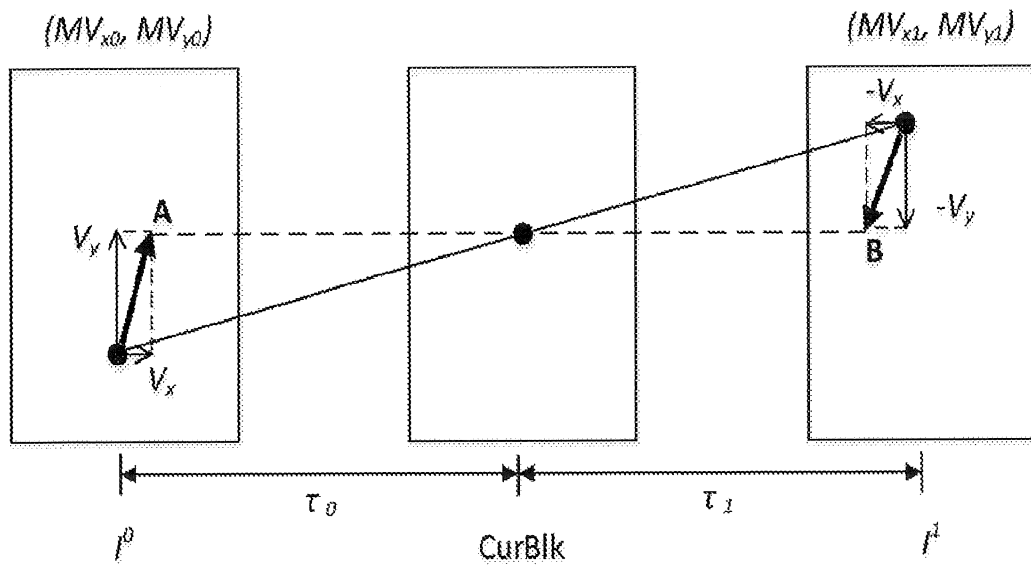


FIG. 4

4 / 4

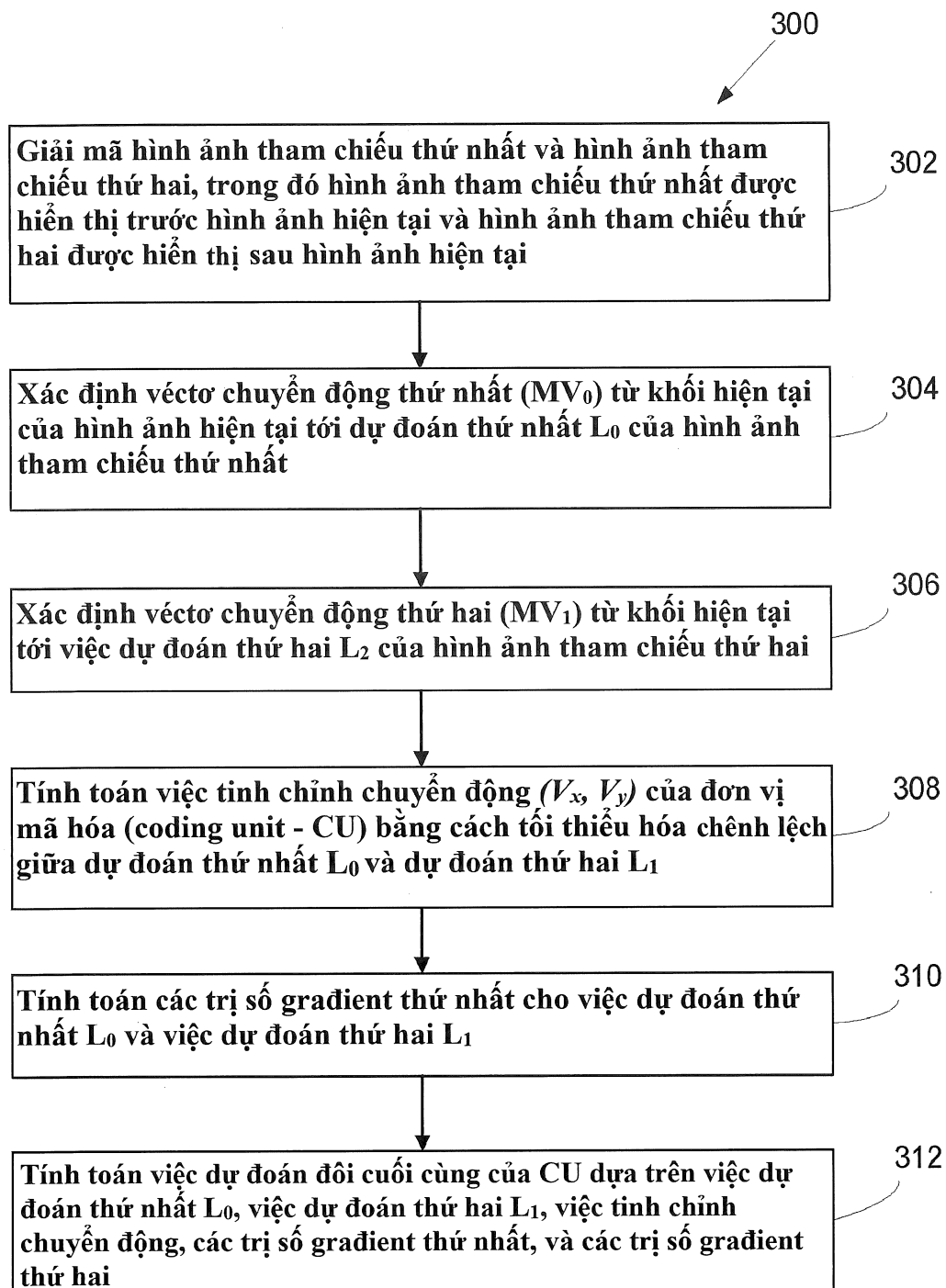


FIG. 5