



(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ

(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN)
CỤC SỞ HỮU TRÍ TUỆ

(11)



1-0049142

(51)^{2020.01} H04N 19/107; H04N 19/70; H04N
19/577; H04N 19/105; H04N 19/176

(13) B

(21) 1-2021-04194

(22) 09/01/2020

(86) PCT/US2020/012826 09/01/2020

(87) WO2020/146562 16/07/2020

(30) 62/790,421 09/01/2019 US

(45) 25/07/2025 448

(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 100085, China

(72) XIU, Xiaoyu (CN); CHEN, Yi-Wen (CN); WANG, Xianglin (US).

(74) Công ty TNHH Đại Tín và Liên Danh (DAITIN AND ASSOCIATES CO.,LTD)

(54) PHƯƠNG PHÁP MÃ HÓA VIDEO, THIẾT BỊ MÃ HÓA VIDEO VÀ THIẾT BỊ
LƯU TRỮ

(21) 1-2021-04194

(57) Sáng chế đề cập đến phương pháp, thiết bị mã hóa video và thiết bị lưu trữ nâng cao hiệu quả dự đoán kết hợp liên và nội bộ (CIIP) trong mã hóa video. Phương pháp bao gồm bước: thu hình ảnh tham chiếu thứ nhất và hình ảnh tham chiếu thứ hai được liên kết với khối mã hiện thời, thu dự đoán thứ nhất L0 dựa trên vector chuyển động thứ nhất MV0 từ khối mã hiện thời đến khối tham chiếu trong hình ảnh tham chiếu thứ nhất, thu dự đoán thứ hai L1 dựa trên vector chuyển động thứ hai MV1 từ khối mã hiện thời đến khối tham chiếu trong hình ảnh tham chiếu thứ hai, xác định xem hoạt động của dòng quang học hai chiều (BDOF) có được áp dụng hay không và tính toán dự đoán kép của khối mã hiện thời dựa trên dự đoán thứ nhất L0 và dự đoán thứ hai L1, và các giá trị gradient thứ nhất và giá trị gradient thứ hai.

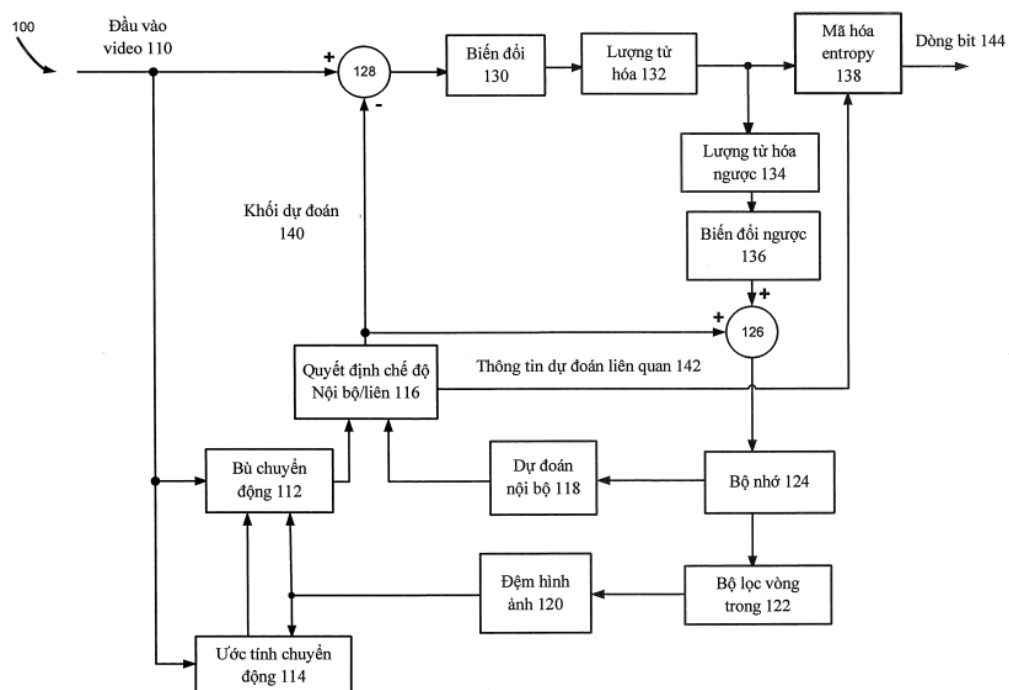


FIG. 1

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến lĩnh vực mã hóa và nén video. Cụ thể hơn là, sáng chế đề cập đến phương pháp mã hóa video, thiết bị mã hóa video và thiết bị lưu trữ nâng cao hiệu quả dự đoán kết hợp liên và nội bộ (CIIP) trong mã hóa video.

Tình trạng kỹ thuật của sáng chế

Để nén video thì có thể dùng rất nhiều kỹ thuật mã hóa. Mã hóa video được thực hiện dựa theo một hoặc nhiều tiêu chuẩn mã hóa video. Ví dụ, các tiêu chuẩn mã hóa video bao gồm mã hóa video đa năng (versatile video coding, VVC), thử nghiệm thăm dò chung (joint exploration test model, JEM), mã hóa video hiệu năng cao (H.265/HEVC), mã hóa video tiên tiến (H.264/AVC), nhóm chuyên gia hình ảnh động (moving picture experts group, MPEG), hoặc tương tự. Mã hóa video thường sử dụng các phương pháp dự đoán (ví dụ: dự đoán liên, dự đoán nội bộ hoặc tương tự) tận dụng phần dư thừa có trong các chuỗi hình ảnh hoặc video. Mục tiêu quan trọng của kỹ thuật mã hóa video là nén dữ liệu video thành dạng tốc độ bit thấp hơn trong khi vẫn đảm bảo chất lượng video.

Bản chất kỹ thuật của sáng chế

Các ví dụ thực hiện sáng chế minh họa giải pháp để cải thiện hiệu năng báo hiệu cú pháp của các chế độ hợp nhất liên quan.

Theo khía cạnh thứ nhất, sáng chế đề xuất phương pháp mã hóa video, phương pháp gồm các bước: thu hình ảnh tham chiếu thứ nhất và hình ảnh tham chiếu thứ hai được liên kết với khối mã hiện thời, trong đó hình ảnh tham chiếu thứ nhất đặt trước hình ảnh hiện thời và hình ảnh tham chiếu thứ hai đặt sau hình ảnh hiện thời theo thứ tự hiển thị, thu dự đoán thứ nhất L0 dựa trên vector chuyển động thứ nhất MV0 từ khối mã hiện thời đến khối tham chiếu trong hình ảnh tham chiếu thứ nhất, thu dự đoán thứ hai L1 dựa trên vector chuyển động thứ hai MV1 từ khối mã hiện thời đến khối tham chiếu trong hình ảnh tham chiếu thứ hai, xác định xem hoạt động của luồng quang học hai chiều

(BDOF) có được áp dụng hay không, trong đó BDOF tính toán các giá trị gradient ngang và dọc thứ nhất $\frac{\partial I^{(0)}}{\partial x}(i, j)$'s và $\frac{\partial I^{(0)}}{\partial y}(i, j)$'s cho các mẫu dự đoán được liên kết với dự đoán thứ nhất L0 và các giá trị gradient ngang và dọc thứ hai $\frac{\partial I^{(1)}}{\partial x}(i, j)$'s và $\frac{\partial I^{(1)}}{\partial y}(i, j)$'s được kết hợp với dự đoán thứ hai L1, và tính toán dự đoán kép của khối mã hiện thời dựa trên dự đoán thứ nhất L0 và dự đoán thứ hai L1 và giá trị gradient thứ nhất $\frac{\partial I^{(0)}}{\partial x}(i, j)$'s và $\frac{\partial I^{(0)}}{\partial y}(i, j)$'s và giá trị gradient thứ hai $\frac{\partial I^{(1)}}{\partial x}(i, j)$'s và $\frac{\partial I^{(1)}}{\partial y}(i, j)$'s.

Theo khía cạnh thứ hai, sáng chế đề xuất phương pháp mã hóa video, trong đó gồm các bước: thu hình ảnh tham chiếu trong danh sách hình ảnh tham chiếu được liên kết với khối mã hiện thời, tạo dự đoán liên dựa trên vector chuyển động thứ nhất từ hình ảnh hiện thời đến hình ảnh tham chiếu thứ nhất, thu chế độ dự đoán nội bộ được liên kết với khối mã hiện thời, tạo dự đoán nội bộ của khối mã hiện thời dựa trên dự đoán nội bộ, tạo dự đoán cuối cùng của khối mã hiện thời bằng cách lấy trung bình của dự đoán liên và dự đoán nội bộ, xác định xem khối mã hiện thời được coi là chế độ liên hoặc chế độ nội bộ đối với chế độ đúng nhất (MPM) dựa trên chế độ dự đoán nội bộ.

Theo khía cạnh thứ ba, sáng chế đề xuất phương tiện lưu trữ lâu có thể đọc được bằng máy tính trong đó các chỉ thị. Khi thực thi bởi một hoặc nhiều bộ xử lý, làm cho thiết bị tính toán thực hiện các hoạt động gồm thu hình ảnh tham chiếu thứ nhất và hình ảnh tham chiếu thứ hai được liên kết với khối mã hiện thời, trong đó hình ảnh tham chiếu thứ nhất đặt trước hình ảnh hiện thời và hình ảnh tham chiếu thứ hai đặt sau hình ảnh hiện thời theo thứ tự hiển thị, thu dự đoán thứ nhất L0 dựa trên vector chuyển động thứ nhất MV0 từ khối mã hiện thời đến khối tham chiếu trong hình ảnh tham chiếu thứ nhất, thu dự đoán thứ hai L1 dựa trên vector chuyển động thứ hai MV1 từ khối mã hiện thời tới khối tham chiếu trong hình ảnh tham chiếu thứ hai, xác định xem hoạt động của luồng quang học hai chiều (BDOF) có được áp dụng hay không, trong đó BDOF tính toán giá trị gradient ngang và dọc thứ nhất $\frac{\partial I^{(0)}}{\partial x}(i, j)$'s và $\frac{\partial I^{(0)}}{\partial y}(i, j)$'s cho các mẫu dự đoán được liên kết với dự đoán thứ nhất L0 và các giá trị gradient ngang và dọc thứ hai $\frac{\partial I^{(1)}}{\partial x}(i, j)$'s và $\frac{\partial I^{(1)}}{\partial y}(i, j)$'s được kết hợp với dự đoán thứ hai L1 và tính toán dự đoán kép của khối mã hiện thời.

Theo khía cạnh thứ tư, sáng chế đề xuất phương tiện lưu trữ lâu có thể đọc được bằng máy tính trong đó lưu trữ các chỉ thị. Khi thực thi bởi một hoặc nhiều bộ xử lý, làm cho thiết bị tính toán thực hiện các hoạt động gồm thu hình ảnh tham chiếu trong danh sách hình ảnh tham chiếu được liên kết với khối mã hiện thời, tạo dự đoán liên dựa trên vectơ chuyển động thứ nhất từ hình ảnh hiện thời tới hình ảnh tham chiếu thứ nhất, thu chế độ dự đoán nội bộ được liên kết với khối mã hiện thời, tạo dự đoán nội bộ của khối mã hiện thời dựa trên dự đoán nội bộ, tạo dự đoán cuối cùng của khối mã hiện thời bằng cách lấy trung bình giữa dự đoán liên và dự đoán nội bộ, và xác định xem khối mã hiện thời có được coi là chế độ liên hoặc chế độ nội bộ đối với chế độ đúng nhất (MPM) dựa trên chế độ dự đoán nội bộ.

Cần hiểu rằng phần mô tả trên đây và mô tả chi tiết sau đây chỉ là các ví dụ minh họa và không giới hạn phạm vi của sáng chế.

Mô tả vắn tắt các hình vẽ.

Sau đây, sáng chế sẽ được bộc lộ rõ ràng hơn thông qua phần mô tả các phương án thực hiện cùng với việc tham khảo các hình vẽ kèm theo. Tuy nhiên, phần mô tả các phương án thực hiện cùng với các hình vẽ kèm theo chỉ nhằm mục đích minh họa mà không giới hạn phạm vi của sáng chế. Trong đó:

Fig.1 là sơ đồ khối minh họa bộ mã hóa theo một phương án thực hiện mẫu của sáng chế.

Fig.2 là sơ đồ khối minh họa bộ giải mã theo một phương án thực hiện mẫu của sáng chế.

Fig.3 là lưu đồ minh họa phương pháp tạo ra dự đoán kết hợp liên và nội bộ (combined inter and intra prediction, CIIP), theo một phương án thực hiện mẫu của sáng chế.

Fig.4 là lưu đồ minh họa phương pháp tạo ra CIIP theo một phương án thực hiện mẫu của sáng chế.

Fig.5A là sơ đồ minh họa phân vùng khối trong cấu trúc cây đa loại theo một phương án thực hiện mẫu của sáng chế.

Fig.5B là sơ đồ minh họa phân vùng khối trong cấu trúc cây đa loại theo một phương án thực hiện mẫu của sáng chế.

Fig.5C là sơ đồ minh họa phân vùng khối trong cấu trúc cây đa loại theo một phương án thực hiện mẫu của sáng chế.

Fig.5D là sơ đồ minh họa phân vùng khối trong cấu trúc cây đa loại theo một phương án thực hiện mẫu của sáng chế.

Fig.5E là sơ đồ minh họa các phân vùng khối trong cấu trúc cây đa loại theo một phương án thực hiện mẫu của sáng chế.

Fig.6A là sơ đồ minh họa dự đoán kết hợp liên và nội bộ theo một phương án thực hiện mẫu của sáng chế.

Fig.6B là sơ đồ minh họa dự đoán kết hợp liên và nội bộ theo một phương án thực hiện mẫu của sáng chế.

Fig.6C là sơ đồ minh họa dự đoán kết hợp liên và nội bộ theo một phương án thực hiện mẫu của sáng chế.

Fig.7A là sơ đồ quy trình tạo ra danh sách đối tượng MPM theo một phương án thực hiện mẫu của sáng chế.

Fig. 7B là sơ đồ quy trình tạo ra danh sách đối tượng MPM theo một phương án thực hiện mẫu của sáng chế.

Fig.8 là sơ đồ minh họa tiến trình của thiết kế CIIP hiện thời trong VVC theo một phương án thực hiện mẫu của sáng chế.

Fig.9 là sơ đồ minh họa tiến trình của phương pháp CIIP được đề xuất loại bỏ BDOF theo một phương án thực hiện mẫu của sáng chế.

Fig.10 là sơ đồ minh họa tiến trình của dự đoán đơn nhất dựa trên CIIP bằng cách lựa chọn danh sách dự đoán dựa trên khoảng cách POC theo một phương án thực hiện mẫu của sáng chế.

Fig.11A là sơ đồ các phương pháp khi kích hoạt các khối CIIP để tạo danh sách đối tượng MPM theo một phương án thực hiện mẫu của sáng chế.

Fig.11B là sơ đồ các phương pháp khi vô hiệu hóa các khối CIIP để tạo danh sách đối tượng MPM theo một phương án thực hiện mẫu của sáng chế.

Fig.12 là sơ đồ minh họa môi trường tính toán kết hợp với giao diện người dùng theo một phương án thực hiện mẫu của sáng chế.

Mô tả chi tiết sáng chế

Sau đây là phần mô tả chi tiết sáng chế thông qua các ví dụ thực hiện các phương án của sáng chế cùng với việc tham chiếu tới các hình vẽ kèm theo. Trong phần mô tả này, các số tham chiếu giống nhau sẽ được sử dụng trong toàn bộ bản vẽ để chỉ các bộ phận giống nhau hoặc tương tự trừ khi được biểu diễn khác. Các ví dụ thực hiện được mô tả không phải là toàn bộ các phương án thực hiện sáng chế. Thay vào đó, đây chỉ là những ví dụ về thiết bị và phương pháp theo phương án thực hiện ưu tiên theo một vài trong số các khía cạnh của sáng chế nằm trong phạm vi yêu cầu bảo hộ kèm theo.

Thuật ngữ được sử dụng trong sáng chế chỉ nhằm mục đích mô tả các phương án cụ thể và không làm giới hạn phạm vi của sáng chế. Như được sử dụng trong các phương án thực hiện sáng chế và các yêu cầu bổ sung, các dạng số ít “một” “duy nhất” và “đó” cũng bao gồm các dạng số nhiều, trừ khi ngữ cảnh chỉ ra rõ ràng khác. Cũng cần hiểu rằng thuật ngữ “và/hoặc” được sử dụng ở đây là để chỉ và bao gồm bất kỳ hoặc tất cả các sự kết hợp có thể có của một hoặc nhiều đối tượng liên quan được liệt kê.

Cần hiểu rằng, mặc dù các thuật ngữ “thứ nhất”, “thứ hai”, “thứ ba”, v.v. có thể được sử dụng để mô tả các nội dung thông tin khác nhau, nhưng nội dung thông tin của sáng chế không bị giới hạn bởi các thuật ngữ này. Các thuật ngữ này chỉ được sử dụng để phân biệt loại thông tin này với loại thông tin khác. Ví dụ, nhưng không nằm ngoài phạm vi của sáng chế, thông tin thứ nhất có thể được gọi là thông tin thứ hai; và tương tự, thông tin thứ hai cũng có thể được gọi là thông tin thứ nhất. Như được sử dụng ở đây, thuật ngữ “nếu” có thể được hiểu có nghĩa là “khi nào” hoặc “ngay khi” hoặc “đáp lại một tranh luận” tùy thuộc vào ngữ cảnh.

Phiên bản đầu tiên của tiêu chuẩn HEVC đã được hoàn thiện vào tháng 10 năm 2013, cho khả năng tiết kiệm tốc độ bit gần 50% hoặc chất lượng cảm nhận tương đương với tiêu chuẩn mã hóa video thế hệ trước H.264/MPEG AVC. Mặc dù tiêu chuẩn HEVC cung cấp những cải tiến đáng kể về mã hóa so với tiền thân của nó, nhưng có bằng chứng cho thấy các công cụ mã hóa bổ sung có thể đạt được hiệu quả mã hóa vượt trội hơn cả HEVC. Trên cơ sở đó, cả VCEG và MPEG đều bắt đầu khai phá các công nghệ mã hóa mới để tiêu chuẩn hóa mã hóa video trong tương lai. Một đội thăm dò video chung (Joint Video Exploration Team, JVET) được thành lập vào tháng 10 năm 2015 bởi ITU-T VCEG và ISO/IEC MPEG bắt đầu nghiên cứu quan trọng về các công nghệ tiên tiến có thể cho phép nâng cao đáng kể hiệu quả mã hóa. Một phần mềm tham chiếu được gọi là mô hình thăm dò chung (joint exploration model, JEM) đã được JVET duy trì bằng cách tích hợp một số công cụ mã hóa bổ sung trên phần đầu mô hình kiểm tra HEVC (HM).

Vào tháng 10 năm 2017, lời kêu gọi chung cho các đề xuất (call for proposals, CfP) về nén video với khả năng vượt trội hơn HEVC đã được ITU-T và ISO/IEC đưa ra. Vào tháng 4 năm 2018, 23 phản hồi về CfP đã được tiếp nhận và đánh giá tại cuộc họp JVET lần thứ 10, chứng tỏ hiệu suất nén đạt được vượt trội hơn so với HEVC khoảng 40%. Dựa trên kết quả đánh giá đó, JVET đã khởi động một dự án mới nhằm phát triển tiêu chuẩn mã hóa video thế hệ mới được đặt tên là mã hóa video đa năng (Versatile Video Coding, VVC). Cùng tháng đó, một cơ sở mã phần mềm tham chiếu, được gọi là mô hình kiểm tra VVC (VVC test model, VTM), cũng đã được thiết lập để chứng minh việc triển khai tham chiếu của tiêu chuẩn VVC.

Giống như HEVC, VVC được xây dựng dựa trên khung mã hóa video kết hợp dựa trên khối. Fig.1 (mô tả chi tiết dưới đây) cho thấy sơ đồ khối của hệ thống mã hóa video kết hợp dựa trên khối chung. Tín hiệu video đầu vào được xử lý từng khối một (gọi là các đơn vị mã hóa (CU)). Trong VTM-1.0, CU có thể lên đến 128x128 pixel. Tuy nhiên, khác với HEVC phân vùng khối chỉ dựa trên cây tứ phân, trong VVC, một đơn vị cây mã hóa (coding tree unit, CTU) được chia thành các CU để thích ứng với các đặc điểm cục bộ khác nhau dựa trên cây tứ phân/nhị phân/bậc ba. Ngoài ra, khái niệm loại đơn vị đa phân vùng trong HEVC bị loại bỏ, tức là, sự tách biệt của CU, đơn vị dự đoán (prediction unit, PU) và đơn vị biến đổi (transform unit, TU) không còn tồn tại trong VVC nữa; thay vào đó, mỗi CU luôn được sử dụng làm đơn vị cơ bản cho cả dự đoán và

biến đổi mà không cần phân vùng bổ sung. Trong cấu trúc cây đa loại, một CTU trước hết được phân vùng theo cấu trúc cây tứ phân. Sau đó, mỗi nút lá cây tứ phân có thể được phân chia thêm bằng cấu trúc cây nhị phân và cây bậc ba. Như được minh họa trên Fig.5A, Fig.5B, Fig.5C, Fig.5D và Fig.5E (mô tả chi tiết dưới đây), có năm kiểu chia, phân vùng bậc bốn, phân vùng nhị phân ngang, phân vùng nhị phân dọc, phân vùng bậc ba ngang và phân vùng bậc ba dọc, và tương tự.

Tham khảo Fig.1 (mô tả chi tiết dưới đây), dự đoán không gian và/hoặc dự đoán thời gian có thể được thực hiện. Dự đoán không gian (hoặc "dự đoán nội bộ") sử dụng pixel từ các mẫu của các khối lân cận đã được mã hóa (được gọi là mẫu tham chiếu) trong cùng một hình ảnh/lớp video để dự đoán khối video hiện thời. Dự đoán không gian làm giảm dư thừa về không gian vốn có trong tín hiệu video. Dự đoán thời gian (còn được gọi là "dự đoán liên" hoặc "dự đoán bù chuyển động") sử dụng các pixel được tái tạo từ các hình ảnh video đã được mã hóa để dự đoán khối video hiện thời. Dự đoán thời gian làm giảm dư thừa về thời gian vốn có trong tín hiệu video. Tín hiệu dự đoán thời gian cho CU nhất định thường được báo hiệu bởi một hoặc nhiều vector chuyển động (MV) cho biết số lượng và hướng chuyển động giữa CU hiện thời và tham chiếu thời gian của nó. Ngoài ra, khi đa hình ảnh tham chiếu được hỗ trợ, một chỉ số hình ảnh tham chiếu sẽ được gửi thêm, sử dụng để xác định từ hình ảnh tham chiếu nào trong hình ảnh tham chiếu lưu trữ tín hiệu dự đoán thời gian đến. Sau khi dự đoán theo không gian và/hoặc thời gian, khối quyết định chế độ trong bộ mã hóa sẽ chọn chế độ dự đoán tốt nhất, ví dụ dựa trên phương pháp tối ưu hóa tỉ lệ lệch. Khối dự đoán sau đó được loại khỏi khối video hiện thời; và phần dự đoán dư thừa được ngắt liên kết bằng cách sử dụng phép biến đổi và lượng tử hóa.

Các hệ số dư lượng tử hóa được lượng tử hóa ngược và biến đổi ngược để tạo thành phần dư thừa được tái tạo, sau đó được thêm trở lại khối dự đoán để tạo thành tín hiệu được tái tạo của CU. Thêm lọc vòng trong, chẳng hạn như bộ lọc gỡ lỗi, độ lệch tương thích mẫu (sample adaptive offset, SAO) và bộ lọc vòng trong thích ứng (adaptive in-loop filter, ALF) có thể được áp dụng trên CU tái tạo trước khi nó được đưa vào kho ảnh tham chiếu và được sử dụng để mã hóa các khối video trong tương lai. Để hình thành dòng bit video đầu ra, chế độ mã hóa (liên hoặc nội bộ), thông tin chế độ dự đoán, thông

tin chuyển động và hệ số dư lượng tử hóa đều được gửi đến đơn vị mã hóa entropy để được nén và gói để tạo thành dòng bit.

Fig.2 (mô tả chi tiết dưới đây) là sơ đồ khối chung của bộ giải mã video dựa trên khối. Dòng bit video thứ nhất được giải mã entropy tại đơn vị giải mã entropy. Chế độ mã hóa và thông tin dự đoán được gửi đến cả hai đơn vị dự đoán không gian (khi được mã hóa nội bộ) hoặc đơn vị dự đoán thời gian (khi được mã hóa liên) để tạo thành khối dự đoán. Các hệ số biến đổi dư thừa được gửi đến đơn vị lượng tử hóa ngược và đơn vị biến đổi ngược để tái tạo lại khối dư. Khối dự đoán và khối dư thừa sau đó được thêm vào nhau. Khối được tái tạo có thể tiếp tục trải qua quá trình lọc vòng trong trước khi lưu trữ trong kho ảnh tham chiếu. Sau đó, video được tái tạo trong kho ảnh tham chiếu sẽ được gửi đến để điều khiển thiết bị hiển thị, cũng như được sử dụng để dự đoán các khối video trong tương lai.

Fig.1 minh họa một bộ mã hóa điển hình 100. Bộ mã hóa 100 có đầu vào video 110, bù chuyển động 112, ước tính chuyển động 114, quyết định chế độ nội bộ/liên 116, bộ dự đoán khối 140, bộ cộng 128, biến đổi 130, lượng tử hóa 132, thông tin dự đoán liên quan 142, dự đoán nội bộ 118, bộ đệm hình ảnh 120, lượng tử hóa ngược 134, biến đổi ngược 136, bộ cộng 126, bộ nhớ 124, bộ lọc vòng trong 122, mã hóa entropy 138 và dòng bit 144.

Fig.2 là sơ đồ khối 200 bộ giải mã điển hình. Bộ giải mã 200 có dòng bit 210, giải mã entropy 212, lượng tử hóa ngược 214, biến đổi ngược 216, bộ cộng 218, lựa chọn chế độ nội bộ/liên 220, dự đoán nội bộ 222, bộ nhớ 230, bộ lọc vòng trong 228, bù chuyển động 224, bộ đệm hình ảnh 226, dự đoán thông tin liên quan 234 và đầu ra video 232.

Fig.3 minh họa phương pháp mẫu 300 để tạo dự đoán kết hợp liên và nội bộ (CIIP) theo phương án thực hiện của sáng chế.

Trong bước 310, thu nhận hình ảnh tham chiếu thứ nhất và hình ảnh tham chiếu thứ hai được liên kết với khối mã hiện thời, trong đó hình ảnh tham chiếu thứ nhất đặt trước hình ảnh hiện thời và hình ảnh tham chiếu thứ hai đặt sau hình ảnh hiện thời theo thứ tự hiển thị.

Trong bước 312, thu nhận dự đoán thứ nhất L0 dựa trên vectơ chuyển động thứ nhất MV0 từ khối mã hiện thời đến khối tham chiếu trong hình ảnh tham chiếu thứ nhất.

Trong bước 314, thu nhận dự đoán thứ hai L1 dựa trên vectơ chuyển động thứ hai MV1 từ khối mã hiện thời đến khối tham chiếu trong hình ảnh tham chiếu thứ hai.

Trong bước 316, xác định quy trình hoạt động dòng quang học hai chiều (BDOF) có được áp dụng hay không, trong đó BDOF tính toán các giá trị gradient ngang và dọc thứ nhất cho các mẫu dự đoán được liên kết với dự đoán thứ nhất L0 và các giá trị gradient ngang và dọc thứ hai được liên kết với dự đoán thứ hai L1. Ví dụ, BDOF tính toán các giá trị gradient ngang và dọc thứ nhất $\frac{\partial I^{(0)}}{\partial x}(i,j)$'s và $\frac{\partial I^{(0)}}{\partial y}(i,j)$'s đối với các mẫu dự đoán được liên kết với dự đoán thứ nhất L0 và các giá trị gradient ngang và dọc thứ hai $\frac{\partial I^{(1)}}{\partial x}(i,j)$'s và $\frac{\partial I^{(1)}}{\partial y}(i,j)$'s được liên kết với dự đoán thứ hai L1.

Trong bước 318, tính toán dự đoán kép của khối mã hiện thời dựa trên dự đoán thứ nhất L0 và dự đoán thứ hai L1 cũng như các giá trị gradient thứ nhất và giá trị gradient thứ hai. Ví dụ: các giá trị gradient thứ nhất $\frac{\partial I^{(0)}}{\partial x}(i,j)$'s và $\frac{\partial I^{(0)}}{\partial y}(i,j)$'s và giá trị gradient thứ hai $\frac{\partial I^{(1)}}{\partial x}(i,j)$'s và $\frac{\partial I^{(1)}}{\partial y}(i,j)$'s.

Fig.4 minh họa phương pháp mẫu 400 để tạo CIIP theo phương án thực hiện của sáng chế. Cụ thể là, phương pháp bao gồm bước dự đoán đơn nhất dựa trên dự đoán liên và dự đoán nội bộ dựa trên MPM để tạo CIIP.

Trong bước 410, thu nhận hình ảnh tham chiếu trong danh sách hình ảnh tham chiếu được liên kết với khối mã hiện thời.

Trong bước 412, tạo dự đoán liên dựa trên vectơ chuyển động thứ nhất từ hình ảnh hiện thời đến hình ảnh tham chiếu thứ nhất.

Trong bước 414, thu nhận chế độ dự đoán nội bộ được liên kết với khối mã hiện thời.

Trong bước 416, tạo dự đoán nội bộ của khối mã hiện thời dựa trên dự đoán nội bộ.

Trong bước 418, tạo dự đoán cuối cùng của khối mã hiện thời bằng cách lấy trung bình dự đoán liên và dự đoán nội bộ.

Trong bước 420, xác định xem khối mã hiện thời được xét là chế độ liên hay chế độ nội bộ đối với chế độ đúng nhất (MPM) dựa trên chế độ dự đoán nội bộ.

Fig.5A là sơ đồ minh họa các phân vùng bậc bốn khối trong cấu trúc cây đa loại theo phương án thực hiện mẫu của sáng chế.

Fig.5B là sơ đồ minh họa khối phân vùng nhị phân dọc trong cấu trúc cây đa loại theo phương án thực hiện mẫu của sáng chế.

Fig.5C là sơ đồ minh họa khối phân vùng nhị phân theo chiều ngang trong cấu trúc cây đa loại theo phương án thực hiện mẫu của sáng chế.

Fig.5D cho thấy sơ đồ minh họa khối phân vùng bậc ba theo chiều dọc trong cấu trúc cây đa loại theo phương án thực hiện mẫu của sáng chế.

Fig.5E cho thấy sơ đồ minh họa khối phân vùng bậc ba theo chiều ngang trong cấu trúc cây đa loại theo phương án thực hiện mẫu của sáng chế.

Kết hợp dự đoán liên và nội bộ

Tham khảo Fig.1 và Fig.2, các phương pháp dự đoán nội bộ và liên được sử dụng trong sơ đồ mã hóa video kết hợp, trong đó mỗi PU chỉ được phép chọn dự đoán nội bộ hoặc dự đoán liên để khai thác mối tương quan trong miền thời gian hoặc miền không gian và không được chọn cả hai cùng lúc. Tuy nhiên, như đã được đề cập trong các giải pháp kỹ thuật trước đây, tín hiệu dư thừa được tạo ra bởi các khối dự đoán liên và các khối dự đoán nội bộ có các đặc điểm rất khác nhau. Do đó, khi kết hợp hai loại dự đoán một cách hiệu quả, một dự đoán chính xác hơn có thể được kì vọng giảm năng lượng của phần dự đoán dư thừa và từ đó cải thiện hiệu quả mã hóa. Ngoài ra, trong tính chất nội dung video, chuyển động của các đối tượng chuyển động có thể phức tạp. Cụ thể hơn là, có thể tồn tại các khu vực chứa các nội dung cũ (ví dụ: các đối tượng có trong ảnh được mã hóa trước đó) và nội dung mới xuất hiện (ví dụ: các đối tượng bị loại trừ trong ảnh được mã hóa trước đó). Trong trường hợp như vậy, cả dự đoán liên lẫn dự đoán nội bộ đều không thể cung cấp một dự đoán chính xác về khối hiện thời.

Để nâng cao hơn nữa hiệu quả dự đoán, kết hợp dự đoán nội bộ và dự đoán liên (CIIP), kết hợp dự đoán nội bộ và dự đoán liên của một CU được mã hóa bằng chế độ hợp nhất, được áp dụng trong tiêu chuẩn VVC. Cụ thể là, đối với mỗi CU hợp nhất, một cờ bỏ sung được báo hiệu để cho biết liệu CIIP có được kích hoạt cho CU hiện thời hay không. Đối với thành phần độ sáng, CIIP hỗ trợ bốn chế độ nội bộ thường được sử dụng, bao gồm chế độ phẳng, DC, chế độ ngang và dọc. Đối với các thành phần sắc độ, DM (nghĩa là, sắc độ tái sử dụng cùng một chế độ nội bộ của thành phần độ sáng) luôn được áp dụng mà không cần thêm tín hiệu. Ngoài ra, trong thiết kế CIIP hiện có, trung bình có trọng số được áp dụng để kết hợp các mẫu dự đoán liên các mẫu và các mẫu dự đoán nội bộ của một CIIP CU. Cụ thể, khi chế độ phẳng hoặc DC được chọn, trọng lượng bằng nhau (tức là bằng 0,5) sẽ được áp dụng. Nếu không (nghĩa là áp dụng chế độ ngang hoặc dọc), CU hiện thời trước tiên được chia theo chiều ngang (đối với chế độ ngang) hoặc theo chiều dọc (đối với chế độ dọc) thành bốn vùng có kích thước bằng nhau.

Bốn bộ trọng số, được ký hiệu là (w_intra_i, w_inter_i) , sẽ được áp dụng để kết hợp các mẫu dự đoán liên và nội bộ trong các khu vực khác nhau, trong đó $i = 0$ and $i = 3$ đại diện cho các khu vực gần nhất và xa nhất với các mẫu lân cận được tái tạo được sử dụng cho dự đoán nội bộ. Trong thiết kế CIIP hiện thời, các giá trị của bộ trọng số được đặt là $(w_intra_0, w_inter_0) = (0.75, 0.25)$, $(w_intra_1, w_inter_1) = (0.625, 0.375)$, $(w_intra_2, w_inter_2) = (0.375, 0.625)$ và $(w_intra_3, w_inter_3) = (0.25, 0.75)$. Fig.6A, Fig.6B và Fig.6C (mô tả chi tiết dưới đây) minh họa các ví dụ để minh họa chế độ CIIP.

Ngoài ra, trong hoạt động của VVC hiện thời, chế độ nội bộ của một CIIP CU có thể được sử dụng làm công cụ dự đoán để dự đoán chế độ nội bộ của các CIIP CU lân cận của nó thông qua cơ quan chế độ đúng nhất (MPM). Cụ thể, đối với mỗi CIIP CU, khi các khối lân cận của nó cũng là CIIP CU, các chế độ nội bộ của các khối lân cận đó trước tiên được làm tròn thành chế độ gần nhất trong chế độ phẳng, DC, ngang và dọc và sau đó được thêm vào danh sách đối tượng MPM của CU hiện thời. Tuy nhiên, khi xây dựng danh sách MPM cho mỗi CU nội bộ, khi một trong các khối lân cận của nó được mã hóa bằng chế độ CIIP, nó được coi là không khả dụng, tức là chế độ nội bộ của một CIIP CU không được phép dự đoán các chế độ nội bộ của các CU nội bộ lân cận nó. Fig.7A và Fig.7B (mô tả chi tiết dưới đây) so sánh các quy trình tạo danh sách MPM của CU trong và các CIIP CU.

Luồng quang học hai chiều

Dự đoán kép tiêu chuẩn trong mã hóa video là sự kết hợp đơn giản của hai khối dự đoán thời gian thu nhận từ các hình ảnh tham chiếu đã được tái tạo lại. Tuy nhiên, do giới hạn của bù chuyển động dựa trên khối, có thể có những chuyển động nhỏ còn lại có thể quan sát được giữa các mẫu của hai khối dự đoán, do đó làm giảm hiệu quả của dự đoán bù chuyển động. Để giải quyết vấn đề này, luồng quang học hai chiều (BDOF) được áp dụng trong VVC để giảm tác động của chuyển động như vậy đối với mọi mẫu bên trong một khối.

Cụ thể, như trong Fig.6A, Fig.6B và Fig.6C (được mô tả chi tiết dưới đây), BDOF là sự tinh chỉnh chuyển động theo mẫu được thực hiện ở phần trên của các dự đoán bù chuyển động dựa trên khối khi sử dụng dự đoán kép. Tinh chỉnh chuyển động (v_x, v_y) của mỗi khối con 4×4 được tính toán bằng cách giảm sự chênh lệch giữa các mẫu dự đoán L0 và L1 sau khi áp dụng BDOF bên trong một cửa sổ 6×6 xung quanh khối phụ. Cụ thể, giá trị của (v_x, v_y) được suy ra là

$$v_x = S_1 > 0 ? \text{clip3} \left(-th_{BDOF}, th_{BDOF}, -((S_3 \cdot 2^3) \gg \lfloor \log_2 S_1 \rfloor) \right) : 0$$

$$v_y = S_5 > 0 ? \text{clip3} \left(-th_{BDOF}, th_{BDOF}, -\left((S_6 \cdot 2^3 - ((v_x S_{2,m}) \ll n_{S_2} + v_x S_{2,s}) / 2) \gg \lfloor \log_2 S_5 \rfloor \right) \right) :$$

trong đó $\lfloor \cdot \rfloor$ là hàm tăng; $\text{clip3}(\min, \max, x)$ là một hàm kẹp một giá trị x cho trước bên trong phạm vi $[\min, \max]$; biểu tượng $\gg$ đại diện cho hoạt động dịch chuyển sang phải theo đảo bit; biểu tượng $\ll$ đại diện cho hoạt động phân tích trái đảo bit; th_{BDOF} là ngưỡng lọc chuyển động để ngăn chặn các lỗi lan truyền do chuyển động cục bộ không đều, tính bằng 2^{13-BD} ; trong đó BD là độ sâu bit của video đầu vào. Trong (1), $S_{2,m} = S_2 \gg n_{S_2}$, $S_{2,s} = S_2 \& (2^{n_{S_2}} - 1)$.

Giá trị của S_1, S_2, S_3, S_5 and S_6 được tính bằng

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

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

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

$$\psi_y(i,j)$$

Trong đó

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

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

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

Trong đó $I^{(k)}(i,j)$ là giá trị mẫu tại tọa độ (i,j) của tín hiệu dự đoán trong danh sách k , $k = 0,1$, được tạo ra ở độ chính xác trung gian cao (tức là 16-bit); $\frac{\partial I^{(k)}}{\partial x}(i,j)$ và $\frac{\partial I^{(k)}}{\partial y}(i,j)$ là các gradient ngang và dọc của mẫu thu được bằng cách tính toán trực tiếp sự khác biệt giữa hai mẫu lân cận của nó, tức là

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

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

Dựa trên sự tinh chỉnh chuyển động thu được ở (1), các mẫu dự đoán kép cuối cùng của CU được tính bằng cách nội suy các mẫu dự đoán L0/L1 dọc theo quỹ đạo chuyển động dựa trên mô hình luồng quang học, xác định bởi

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

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

Trong đó *shift* và *offset* là giá trị thay đổi bên phải và giá trị bù được áp dụng để kết hợp các tín hiệu dự đoán L0 và L1 cho dự đoán kép, tương ứng bằng $15 - BD$ và $1 \ll (14 - BD) + 2 \cdot (1 \ll 13)$, tương tự.

Fig.6A cho thấy sơ đồ minh họa dự đoán kết hợp liên và nội bộ cho chế độ ngang theo phương án thực hiện mẫu của sáng chế.

Fig.6B cho thấy sơ đồ minh họa dự đoán kết hợp liên và nội bộ cho chế độ dọc theo phương án thực hiện mẫu của sáng chế.

Fig.6C cho thấy sơ đồ minh họa dự đoán kết hợp liên và nội bộ cho các chế độ phẳng và DC theo phương án thực hiện mẫu của sáng chế.

Fig.7A cho thấy sơ đồ các quy trình tạo danh sách đối tượng MPM của CUS nội bộ theo phương án thực hiện mẫu của sáng chế.

Fig.7B cho thấy sơ đồ các quy trình tạo danh sách đối tượng MPM của các CIIP CU theo phương án thực hiện mẫu của sáng chế.

Các cải tiến đối với CIIP

Mặc dù CIIP có thể nâng cao hiệu quả của dự đoán bù chuyển động tiêu chuẩn, thiết kế của nó vẫn có thể được cải thiện hơn nữa. Cụ thể, các vấn đề sau đây trong thiết kế CIIP hiện thời trong VVC được xác định theo phương án này.

Theo khía cạnh thứ nhất, như đã đề cập trong phần “dự đoán kết hợp liên và nội bộ”, vì CIIP kết hợp các mẫu dự đoán liên và nội bộ, mỗi CIIP CU cần sử dụng các mẫu lân cận của nó được tái tạo lại để tạo tín hiệu dự đoán. Điều này có nghĩa là việc giải mã một CIIP CU phụ thuộc vào việc xây dựng lại toàn bộ các khối lân cận của nó. Do sự phụ thuộc lẫn nhau như vậy, để triển khai phần cứng theo lý thuyết, CIIP cần được thực

hiện trong giai đoạn tái tạo, nơi các mẫu được tái tạo lân cận sẵn sàng để dự đoán nội bộ. Vì việc giải mã các CU trong giai đoạn tái tạo phải được thực hiện tuần tự (tức là từng cái một), số lượng các phép tính (ví dụ: nhân, cộng và thay đổi bit) liên quan đến quá trình CIIP không được quá cao để đảm bảo đủ công suất giải mã thời gian thực.

Như đã đề cập trong phần “luồng quang học hai chiều”, BDOF được kích hoạt để nâng cao chất lượng dự đoán khi một CU liên mã hóa được dự đoán từ hai khối tham chiếu từ cả hai hướng thời gian tiến và lùi. Như được minh họa trên Fig.8 (được mô tả chi tiết dưới đây), trong VVC hiện thời, BDOF cũng tham gia để tạo các mẫu dự đoán liên cho chế độ CIIP. Do sự phức tạp bổ sung được đưa vào bởi BDOF, thiết kế như vậy có thể làm giảm đáng kể công suất mã hóa/giải mã của codec phần cứng khi CIIP được kích hoạt.

Theo khía cạnh thứ hai, trong thiết kế CIIP hiện thời, khi một CIIP CU tham chiếu đến một đối tượng hợp nhất được dự đoán kép, cả hai tín hiệu dự đoán bù chuyển động trong danh sách L0 và L1 đều cần được tạo. Khi một hoặc nhiều MV không có độ chính xác số nguyên, các quy trình nội suy bổ sung phải được dẫn để nội suy các mẫu tại các vị trí mẫu phân số. Quá trình như vậy không chỉ làm tăng độ tính toán phức tạp mà còn tăng băng thông bộ nhớ do cần phải truy cập nhiều mẫu tham chiếu hơn từ bộ nhớ ngoài.

Theo khía cạnh thứ ba, như đã đề cập trong phần “kết hợp dự đoán nội bộ và liên”, trong thiết kế CIIP hiện thời, các chế độ nội bộ của các CIIP CU và chế độ nội bộ của các CU nội bộ được xử lý khác nhau khi xây dựng danh sách MPM của các khối lân cận của chúng. Cụ thể, khi một CU hiện thời được mã hóa bằng chế độ CIIP, các CIIP CU lân cận của nó được coi là nội bộ, tức là, các chế độ nội bộ của các CU CIIP lân cận có thể được thêm vào danh sách đối tượng MPM. Tuy nhiên, khi CU hiện thời được mã hóa theo chế độ nội bộ, các CIIP CU lân cận của nó được coi là liên, tức là, các chế độ nội bộ của các CU CIIP lân cận bị loại khỏi danh sách đối tượng MPM. Thiết kế không thống nhất như vậy có thể không tối ưu cho phiên bản cuối cùng của chuẩn VVC.

Fig.8 là sơ đồ minh họa tiến trình của thiết kế CIIP hiện có trong VVC theo phương án thực hiện mẫu của sáng chế.

Đơn giản hóa CIIP

Theo một phương án thực hiện sáng chế, các phương pháp được đề xuất để đơn giản hóa thiết kế CIIP hiện có để tạo điều kiện triển khai codec phần cứng. Nhìn chung, các khía cạnh chính của các kỹ thuật đề xuất trong sáng chế này được tóm tắt như sau:

Thứ nhất, để cải thiện thông lượng mã hóa/giải mã CIIP, sáng chế đề xuất lược bỏ BDOF khỏi việc tạo các mẫu dự đoán liên trong chế độ CIIP.

Thứ hai, để giảm độ phức tạp tính toán và mức tiêu thụ băng thông bộ nhớ, khi một CIIP CU được dự đoán kép (tức là có cả MV L0 và L1), các phương pháp được đề xuất để chuyển khối từ dự đoán kép thành dự đoán đơn nhất để tạo ra các mẫu dự đoán liên.

Thứ ba, hai phương pháp được đề xuất để cân bằng giữa các chế độ nội bộ của CIIP và CU khi hình thành các đối tượng MPM từ các khối lân cận của chúng.

CIIP không có BDOF

Như đã đề cập trong phần “tình trạng kỹ thuật”, BDOF luôn được kích hoạt để tạo các mẫu dự đoán liên cho chế độ CIIP khi CU hiện thời được dự đoán kép. Do sự phức tạp bổ sung của BDOF, thiết kế CIIP hiện thời có thể làm giảm đáng kể thông lượng mã hóa/giải mã, đặc biệt là làm cho việc giải mã thời gian thực trở nên khó khăn đối với bộ giải mã VVC. Mặt khác, đối với các CIIP CU, các mẫu dự đoán cuối cùng của chúng được tạo ra bằng cách lấy trung bình các mẫu dự đoán liên và các mẫu dự đoán nội bộ. Nói cách khác, các mẫu dự đoán được tinh chỉnh bởi BDOF sẽ không sử dụng trực tiếp làm tín hiệu dự đoán cho các CIIP CU. Do đó, so với các CU dự đoán kép thông thường (nơi BDOF áp dụng trực tiếp để tạo ra các mẫu dự đoán), cải tiến tương ứng thu được từ BDOF kém hiệu quả hơn đối với các CIIP CU. Do đó, dựa trên các cân nhắc ở trên, Do đó, dựa trên các cân nhắc ở trên, sáng chế đề xuất vô hiệu hóa BDOF khi tạo các mẫu dự đoán liên của chế độ CIIP. Fig.9 (mô tả chi tiết dưới đây) minh họa tiến trình tương ứng của quy trình CIIP được đề xuất sau khi loại bỏ BDOF.

Fig.9 là sơ đồ minh họa tiến trình của phương pháp CIIP được đề xuất bằng cách lược bỏ BDOF theo phương án thực hiện mẫu của sáng chế.

CIIP dựa trên dự đoán đơn nhất

Như đã đề cập ở phần trên, khi đối tượng hợp nhất được tham chiếu bởi một CIIP CU được dự đoán kép, cả tín hiệu dự đoán L0 và L1 đều được tạo ra để dự đoán các mẫu bên trong CU. Để giảm băng thông bộ nhớ và độ phức tạp của phép nội suy theo phương án thực hiện của sáng chế, đề xuất chỉ sử dụng các mẫu dự đoán liên được tạo ra bằng cách sử dụng dự đoán đơn nhất (ngay cả khi CU hiện thời được dự đoán kép) được kết hợp với các mẫu dự đoán nội bộ ở chế độ CIIP. Cụ thể hơn là, khi CIIP CU hiện thời được dự đoán đơn nhất, các mẫu dự đoán liên sẽ được kết hợp trực tiếp với các mẫu dự đoán nội bộ. Nếu không (nghĩa là CU hiện thời được dự đoán kép), các mẫu dự đoán liên được CIIP sử dụng được tạo dựa trên dự đoán đơn nhất từ một danh sách dự đoán (L0 hoặc L1). Để chọn danh sách dự đoán, có thể áp dụng các phương pháp khác nhau. Trong phương pháp thứ nhất, sáng chế đề xuất luôn chọn dự đoán thứ nhất (tức là danh sách L0) cho bất kỳ khối CIIP nào được dự đoán bằng hai hình ảnh tham chiếu.

Trong phương pháp thứ hai, sáng chế đề xuất luôn chọn dự đoán thứ hai (tức là danh sách L1) cho bất kỳ khối CIIP nào được dự đoán bởi hai hình ảnh tham chiếu. Trong phương pháp thứ ba, một phương pháp thích ứng được áp dụng trong đó danh sách dự đoán được liên kết với một ảnh tham chiếu có khoảng cách số đếm thứ tự ảnh (picture order count, POC) nhỏ hơn so với ảnh hiện thời được chọn. Fig.10 (mô tả chi tiết dưới đây) minh họa tiến trình của CIIP dựa trên dự đoán đơn nhất với việc chọn danh sách dự đoán dựa trên khoảng cách POC.

Cuối cùng, phương pháp cuối, sáng chế đề xuất chỉ kích hoạt chế độ CIIP khi CU hiện thời không được dự đoán đơn nhất. Hơn nữa, để giảm chi phí, tín hiệu của cờ bật/tắt CIIP phụ thuộc vào hướng dự đoán của CIIP CU hiện thời. Khi CU hiện thời không được dự đoán, cờ CIIP sẽ được báo hiệu trong luồng bit để cho biết CIIP được bật hay tắt. Nếu không (nghĩa là CU hiện thời được dự đoán kép), tín hiệu của cờ CIIP sẽ bị bỏ qua và luôn được suy ra là sai, tức là CIIP luôn bị vô hiệu.

Fig.10 là sơ đồ minh họa quy trình làm việc của CIIP dựa trên dự đoán đơn nhất với việc lựa chọn danh sách dự đoán dựa trên khoảng cách POC theo phương án thực hiện mẫu của sáng chế.

Sự cân bằng của các phương thức nội bộ của CIIP và các CU nội bộ để xây dựng danh sách đối tượng MPM

Như đã đề cập ở phần trên, thiết kế CIIP hiện thời không thống nhất về cách sử dụng các chế độ nội bộ của các CIIP CU và các CU nội bộ để tạo thành danh sách đối tượng MPM của các khối lân cận tương ứng. Cụ thể hơn là, cả chế độ nội bộ của các CIIP CU và các CU nội bộ có thể dự đoán chế độ nội bộ của các khối lân cận tương ứng mã hóa trong chế độ CIIP. Tuy nhiên, chỉ các chế độ nội bộ của các CU nội bộ mới có thể dự đoán các chế độ nội bộ của các CU nội bộ. Để đạt được một thiết kế thống nhất hơn, hai phương pháp được đề xuất trong phần này để cân bằng việc sử dụng các chế độ nội bộ của CIIP và các CU nội bộ để xây dựng danh sách MPM.

Trong phương án thực hiện thứ nhất, sáng chế đề xuất coi chế độ CIIP là chế độ liên để xây dựng danh sách MPM. Cụ thể, khi tạo danh sách MPM của một CIIP CU hoặc một CU nội bộ, khi khối lân cận tương ứng được mã hóa ở chế độ CIIP, thì chế độ nội bộ của khối lân cận được đánh dấu là không khả dụng. Bằng phương pháp này, không có chế độ nội bộ nào của khối CIIP có thể được sử dụng để xây dựng danh sách MPM. Ngược lại, trong phương pháp thứ hai, sáng chế đề xuất coi chế độ CIIP là chế độ nội bộ để xây dựng danh sách MPM. Cụ thể, trong phương pháp này, các chế độ nội bộ của các CIIP CU có thể dự đoán các chế độ nội bộ của cả các khối CIIP lân cận và các khối nội bộ tương ứng. Fig.11A và Fig.11B (mô tả chi tiết dưới đây) minh họa quy trình tạo danh sách đối tượng MPM khi hai phương pháp trên được áp dụng.

Fig.11A là sơ đồ phương pháp khi kích hoạt các khối CIIP để tạo danh sách đối tượng MPM theo phương án thực hiện mẫu của sáng chế.

Fig.11B là sơ đồ các phương pháp khi vô hiệu các khối CIIP để tạo danh sách đối tượng MPM theo phương án thực hiện mẫu của sáng chế.

Fig.12 cho thấy môi trường tính toán 1210 cùng với giao diện người dùng 1260. Môi trường tính toán 1210 có thể là một phần của máy chủ xử lý dữ liệu. Môi trường tính toán 1210 bao gồm bộ xử lý 1220, bộ nhớ 1240 và giao diện I/O 1250.

Bộ xử lý 1220 thường kiểm soát các hoạt động tổng thể của môi trường tính toán 1210, chẳng hạn như các hoạt động liên quan đến trình chiếu, thu thập dữ liệu, truyền dữ liệu và xử lý hình ảnh. Bộ xử lý 1220 có thể bao gồm một hoặc nhiều bộ xử lý để thực thi các lệnh để thực hiện tất cả hoặc một số bước trong các phương pháp được mô tả ở trên. Hơn nữa, bộ xử lý 1220 có thể bao gồm một hoặc nhiều mạch tạo điều kiện cho sự

tương tác giữa bộ xử lý 1220 và các thành phần khác. Bộ xử lý có thể là bộ xử lý trung tâm (Central Processing Unit, CPU), bộ vi xử lý, máy chip đơn, GPU hoặc tương tự.

Bộ nhớ 1240 được định cấu hình để lưu trữ nhiều loại dữ liệu khác nhau nhằm hỗ trợ hoạt động của môi trường tính toán 1210. Ví dụ về dữ liệu đó bao gồm chỉ thị cho bất kỳ ứng dụng hoặc phương pháp nào hoạt động trên môi trường tính toán 1210, dữ liệu video, dữ liệu hình ảnh, v.v. Bộ nhớ 1240 có thể được triển khai bằng cách sử dụng bất kỳ loại bộ nhớ khả biến hoặc bất biến nào, hoặc kết hợp chúng, chẳng hạn như bộ nhớ truy cập ngẫu nhiên tĩnh (static random access memory, SRAM), bộ nhớ chỉ đọc lập trình có thể xóa bằng điện (electrically erasable programmable read-only memory, EEPROM), bộ nhớ chỉ đọc lập trình có thể xóa (erasable PROM, EPROM), bộ nhớ chỉ đọc có thể lập trình (programmable ROM, PROM), bộ nhớ chỉ đọc (read-only memory, ROM), bộ nhớ từ tính, bộ nhớ flash, đĩa từ tính hoặc đĩa quang.

Giao diện I/O 1250 cung cấp giao diện giữa bộ xử lý 1220 và các mô-đun giao diện ngoại vi, chẳng hạn như bàn phím, con lăn chuột, các nút, v.v. Các nút có thể bao gồm nhưng không giới hạn ở, nút trang chủ, nút bắt đầu quét và nút dừng quét. Giao diện I/O 1250 có thể được kết hợp với một bộ mã hóa và bộ giải mã.

Theo một khía cạnh khác, sáng chế cũng đề xuất phương tiện lưu trữ lâu có thể đọc được bằng máy tính bao gồm nhiều chương trình 1242, cụ thể là bao gồm trong bộ nhớ 1240, được thực thi bởi bộ xử lý 1220 trong môi trường tính toán 1210, để thực thi các phương pháp đã được mô tả trên đây. Cụ thể hơn là: phương tiện lưu trữ lâu có thể đọc được bằng máy tính có thể là ROM, RAM, CD-ROM, băng từ, đĩa mềm, thiết bị lưu trữ dữ liệu quang học hoặc những thiết bị tương tự.

Phương tiện lưu trữ lâu có thể đọc được bằng máy tính đã lưu trữ trong đó nhiều chương trình để thiết bị tính toán có một hoặc nhiều bộ xử lý để thực thi, trong đó nhiều chương trình, khi được thực thi bởi một hoặc nhiều bộ xử lý, khiến thiết bị tính toán thực hiện phương pháp được mô tả ở trên để dự đoán chuyển động.

Theo một phương án thực hiện của sáng chế, môi trường tính toán 1210 có thể được triển khai với một hoặc nhiều mạch tích hợp dành riêng cho ứng dụng (Application Specific Integrated Circuit, ASIC), bộ xử lý tín hiệu kỹ thuật số (Digital Signal Processing, DSP), thiết bị xử lý tín hiệu kỹ thuật số (DSP Device, DSPD), thiết bị logic

lập trình được (Programmable Logic Device, PLD), mảng cổng lập trình trường (Field-Programmable Gate Array, FPGA), đơn vị xử lý đồ họa (GPU), bộ điều khiển, bộ vi điều khiển, bộ vi xử lý hoặc các thành phần điện tử khác, để thực hiện các phương pháp được mô tả trên đây.

Các phương án thực hiện khác của sáng chế sẽ rõ ràng đối với những người có trình độ trong cùng lĩnh vực kỹ thuật khi xem xét đặc điểm kỹ thuật và ví dụ thực hiện sáng chế được bộc lộ ở đây. Đơn sáng chế này bao gồm mọi cải biến, cách sử dụng hoặc sự điều chỉnh dựa trên các nguyên tắc chung được bộc lộ của sáng chế và bao gồm cả những điểm khác biệt so với sáng chế như đã biết hoặc thông lệ trong lĩnh vực này. Bản mô tả và các ví dụ chỉ nhằm mục đích minh họa, với phạm vi và tinh thần của sáng chế được xác định bởi các yêu cầu bảo hộ sau đây.

Cần hiểu rõ rằng sáng chế không giới hạn ở các ví dụ chính xác được mô tả ở trên và được minh họa trong các bản vẽ kèm theo, và có thể thực hiện các sửa đổi và thay đổi khác nhau mà không vượt khỏi phạm vi của sáng chế. Dự kiến phạm vi của sáng chế chỉ bị giới hạn bởi các yêu cầu bổ sung.

YÊU CẦU BẢO HỘ

1. Phương pháp mã hóa video, trong đó bao gồm các bước:

thu nhận hình ảnh tham chiếu thứ nhất và hình ảnh tham chiếu thứ hai được liên kết với khối mã hiện thời của hình ảnh hiện thời, trong đó hình ảnh tham chiếu thứ nhất đặt trước hình ảnh hiện thời và hình ảnh tham chiếu thứ hai đặt sau hình ảnh hiện thời theo thứ tự hiển thị;

thu nhận dự đoán thứ nhất dựa trên vector chuyển động thứ nhất từ khối mã hiện thời đến khối tham chiếu trong hình tham chiếu thứ nhất;

thu nhận dự đoán thứ hai dựa trên vector chuyển động thứ hai từ khối mã hiện thời đến khối tham chiếu trong hình tham chiếu thứ hai; và

xác định rằng hoạt động luồng quang học hai chiều (BDOF) không được áp dụng dựa trên dự đoán kết hợp liên và nội bộ được áp dụng để tạo ra dự đoán cuối cùng của khối mã hiện thời, trong đó hoạt động BDOF bao gồm các bước:

tính toán giá trị gradient ngang thứ nhất $\frac{\partial I^{(0)}}{\partial x}(i, j)$'s và giá trị gradient dọc thứ nhất $\frac{\partial I^{(0)}}{\partial y}(i, j)$'s cho các mẫu dự đoán được liên kết với dự đoán thứ nhất, và tính toán giá trị gradient ngang thứ hai $\frac{\partial I^{(1)}}{\partial x}(i, j)$'s và giá trị gradient dọc thứ hai $\frac{\partial I^{(1)}}{\partial y}(i, j)$'s cho các mẫu dự đoán được kết hợp với dự đoán thứ hai, trong đó $I^{(0)}(i, j)$ tượng trưng cho mẫu dự đoán thứ nhất tại vị trí mẫu (i, j) , và $I^{(1)}(i, j)$ tượng trưng cho mẫu dự đoán thứ hai tại vị trí mẫu (i, j) ; và

thu được dự đoán kép của khối mã hiện thời dựa trên dự đoán thứ nhất, dự đoán thứ hai, các giá trị gradient ngang thứ nhất, các giá trị gradient dọc thứ nhất, các giá trị gradient ngang thứ hai, và các giá trị gradient dọc thứ hai.

2. Phương pháp theo điểm 1, bao gồm thêm việc xác định hoạt động BDOF có được áp dụng với điều kiện dự đoán kết hợp liên và nội bộ (CIIP) không được áp dụng để tạo ra dự đoán cuối cùng của khối mã hiện thời.

3. Phương pháp theo điểm 1, trong đó còn bao gồm bước tính toán một dự đoán kép của khối mã hiện thời dựa trên giá trị trung bình của dự đoán thứ nhất và dự đoán thứ hai.

4. Phương pháp theo điểm 2, tính toán dự đoán khối mã hiện thời gradient

trong đó bước tính toán dự đoán kép của khối mã hiện thời còn bao gồm bước:

tính toán giá trị BDOF dựa trên giá trị gradient ngang thứ nhất, giá trị gradient dọc thứ nhất, giá trị gradient ngang thứ hai, và giá trị gradient dọc thứ hai, và

tính toán dự đoán kép của khối mã hiện thời dựa trên giá trị dự đoán thứ nhất, giá trị dự đoán thứ hai và giá trị BDOF.

5. Thiết bị mã hóa video, trong đó bao gồm một hoặc nhiều bộ xử lý và một hay nhiều bộ lưu trữ được gắn với một hay nhiều bộ xử lý, thiết bị mã hóa video được cấu hình để thực hiện các hoạt động bao gồm:

thu nhận một hình ảnh tham chiếu thứ nhất và một hình ảnh tham chiếu thứ hai liên kết với khối mã hiện thời của hình ảnh hiện thời, trong đó hình ảnh tham chiếu thứ nhất đặt trước hình ảnh hiện thời và hình ảnh tham chiếu thứ hai đặt sau hình ảnh hiện thời theo thứ tự trình chiếu;

thu nhận dự đoán thứ nhất dựa trên vector chuyển động thứ nhất từ khối mã hiện thời đến khối tham chiếu trong hình tham chiếu thứ nhất;

thu nhận dự đoán thứ hai dựa trên vector chuyển động thứ hai từ khối mã hiện thời đến khối tham chiếu trong hình tham chiếu thứ hai; và

xác định rằng hoạt động luồng quang học hai chiều (BDOF) không được áp dụng dựa trên dự đoán kết hợp liên và nội bộ được áp dụng để tạo ra dự đoán cuối cùng của khối mã hóa hiện tại, trong đó hoạt động BDOF bao gồm các bước:

tính toán giá trị gradient ngang thứ nhất $\frac{\partial I^{(0)}}{\partial x}(i, j)$'s và giá trị gradient dọc thứ nhất $\frac{\partial I^{(0)}}{\partial y}(i, j)$'s cho các mẫu dự đoán được liên kết với dự đoán thứ nhất, và

tính toán giá trị gradient ngang thứ hai $\frac{\partial I^{(1)}}{\partial x}(i, j)$'s và giá trị gradient dọc thứ hai $\frac{\partial I^{(1)}}{\partial y}(i, j)$'s cho các mẫu dự đoán được kết hợp với dự đoán thứ hai, trong đó $I^{(0)}(i, j)$ tượng trưng cho mẫu dự đoán thứ nhất tại vị trí mẫu (i, j) , và $I^{(1)}(i, j)$ tượng trưng cho mẫu dự đoán thứ hai tại vị trí mẫu (i, j) ; và

thu được dự đoán kép của khối mã hiện thời dựa trên dự đoán thứ nhất, dự đoán thứ hai, các giá trị gradient ngang thứ nhất, các giá trị gradient dọc thứ nhất, các giá trị gradient ngang thứ hai, và các giá trị gradient dọc thứ hai.

6. Thiết bị theo điểm 5, trong đó các hoạt động còn bao gồm: xác định hoạt động BDOF được áp dụng với điều kiện dự đoán kết hợp liên và nội bộ (CIIP) không được áp dụng để tạo dự đoán cuối cùng của khối mã hiện thời.

7. Thiết bị theo điểm 5, trong đó các hoạt động còn bao gồm tính toán dự đoán kép của khối mã hiện thời dựa trên tính trung bình dự đoán thứ nhất và dự đoán thứ hai.

8. Thiết bị theo điểm 6, trong đó tính toán dự đoán kép của khối mã hiện thời còn bao gồm:

tính toán giá trị BDOF dựa trên giá trị gradient ngang thứ nhất, giá trị gradient dọc thứ nhất, giá trị gradient ngang thứ hai, giá trị gradient dọc thứ hai, và

tính toán dự đoán kép của khối mã hiện thời dựa trên dự đoán thứ nhất, dự đoán thứ hai và giá trị BDOF.

9. Phương tiện lưu trữ lâu có thể đọc được bằng máy tính lưu trữ nhiều chương trình thực thi bởi một thiết bị tính toán có một hoặc nhiều bộ xử lý, trong đó nhiều chương trình, khi thực thi bởi một hoặc nhiều bộ xử lý, làm cho thiết bị tính toán thực hiện phương pháp theo một trong các điểm 1-4, và lưu trữ một dòng bit được giải mã bằng phương pháp này.

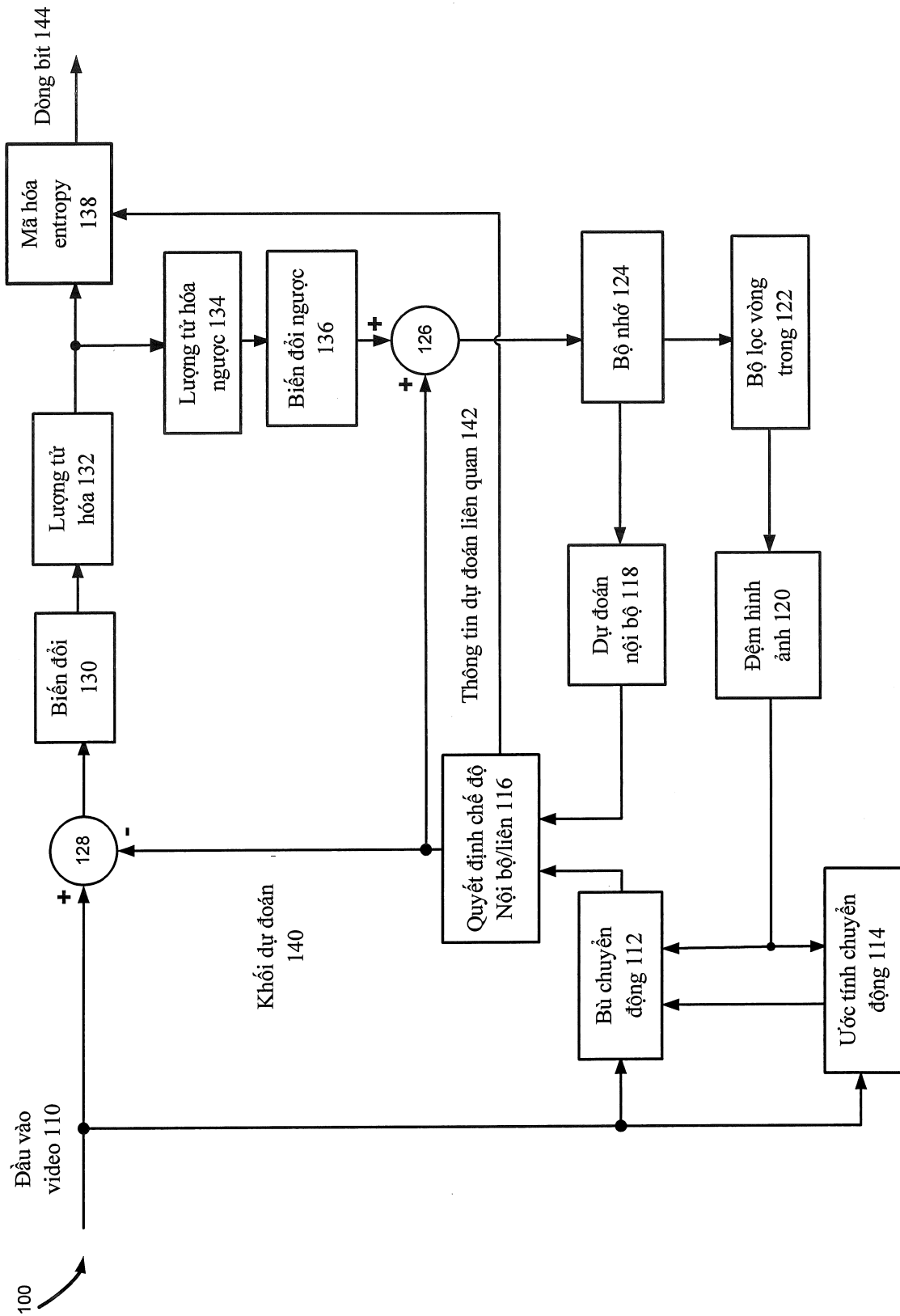


FIG. 1
24

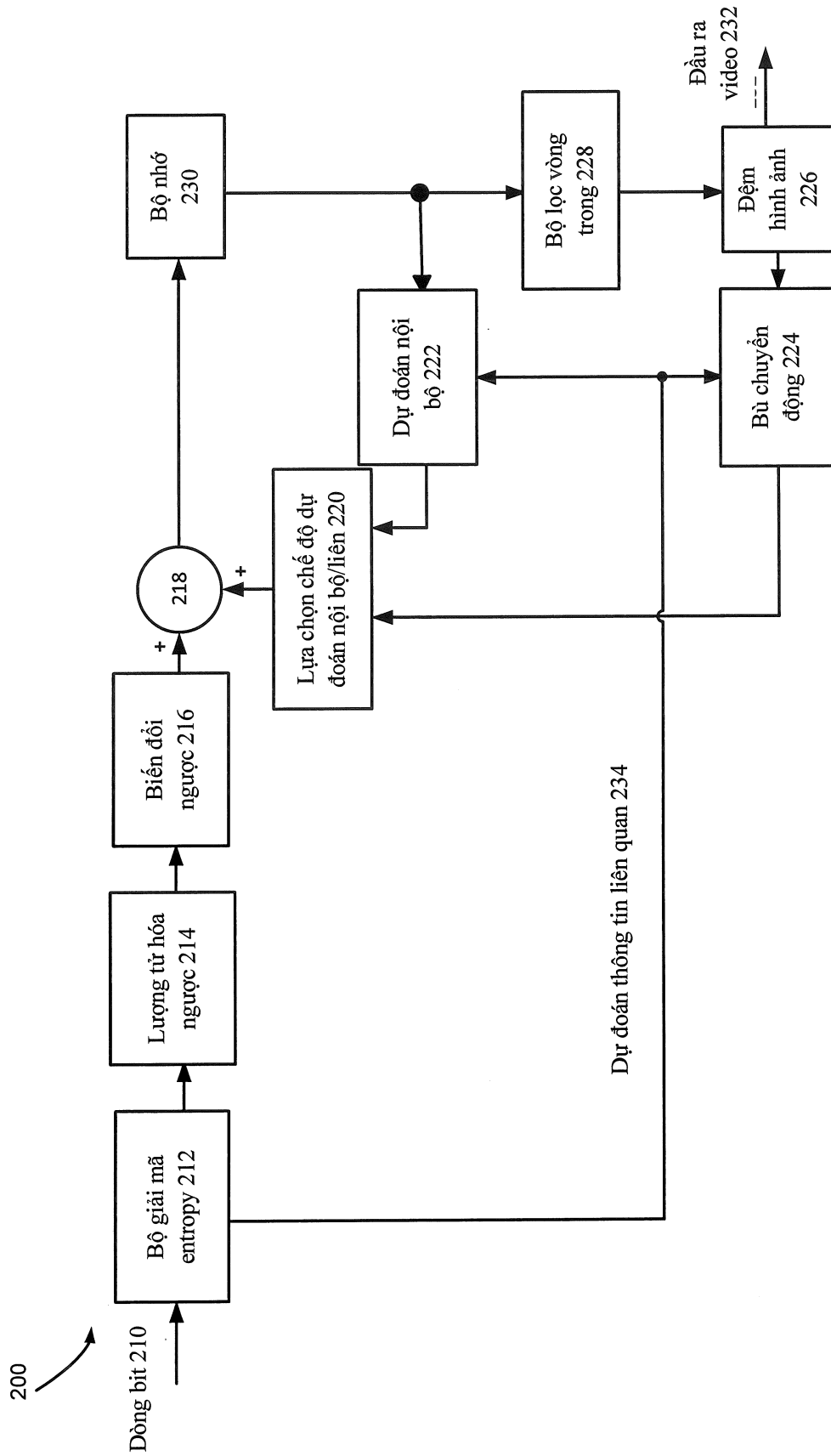


FIG. 2

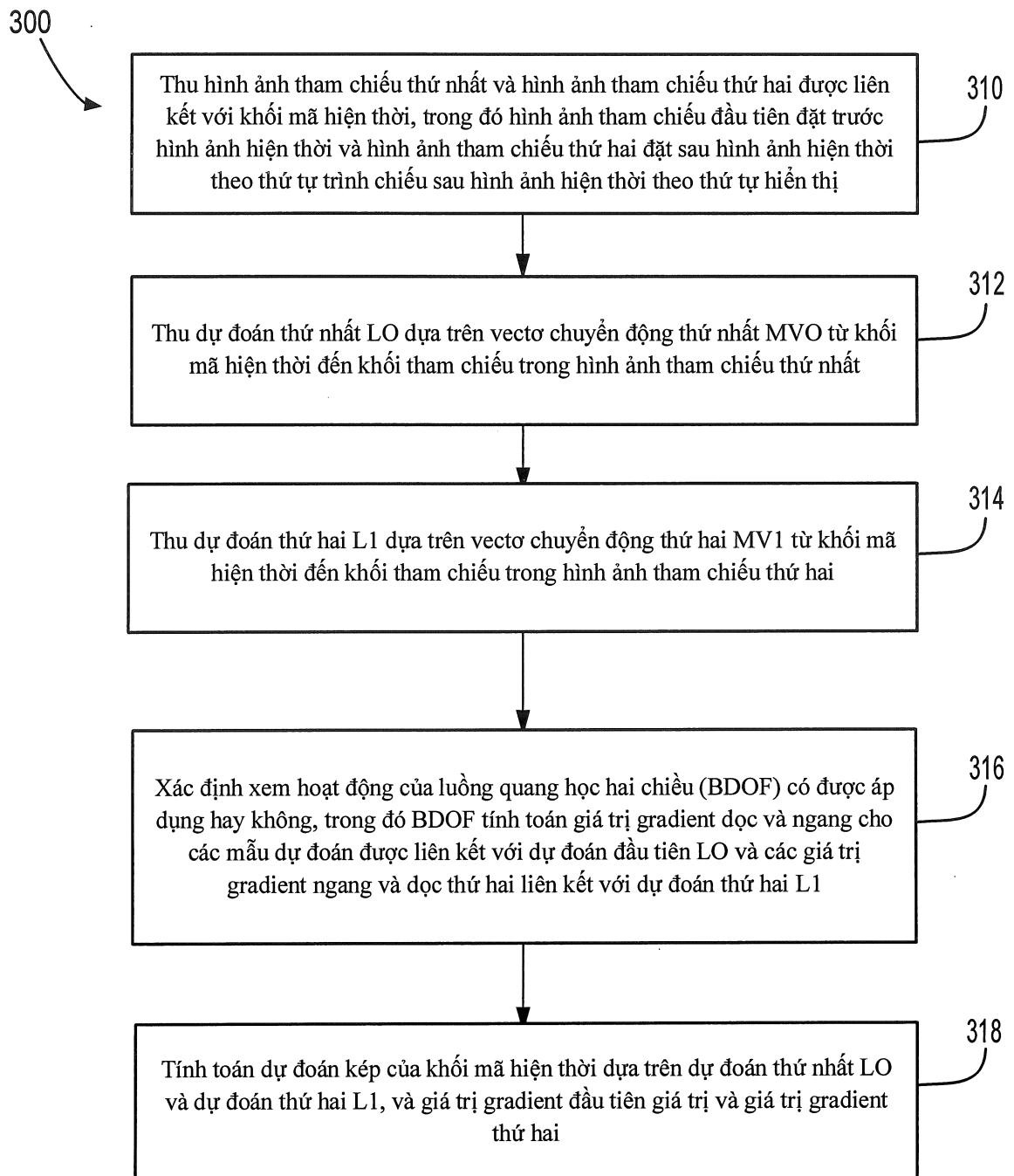


FIG. 3

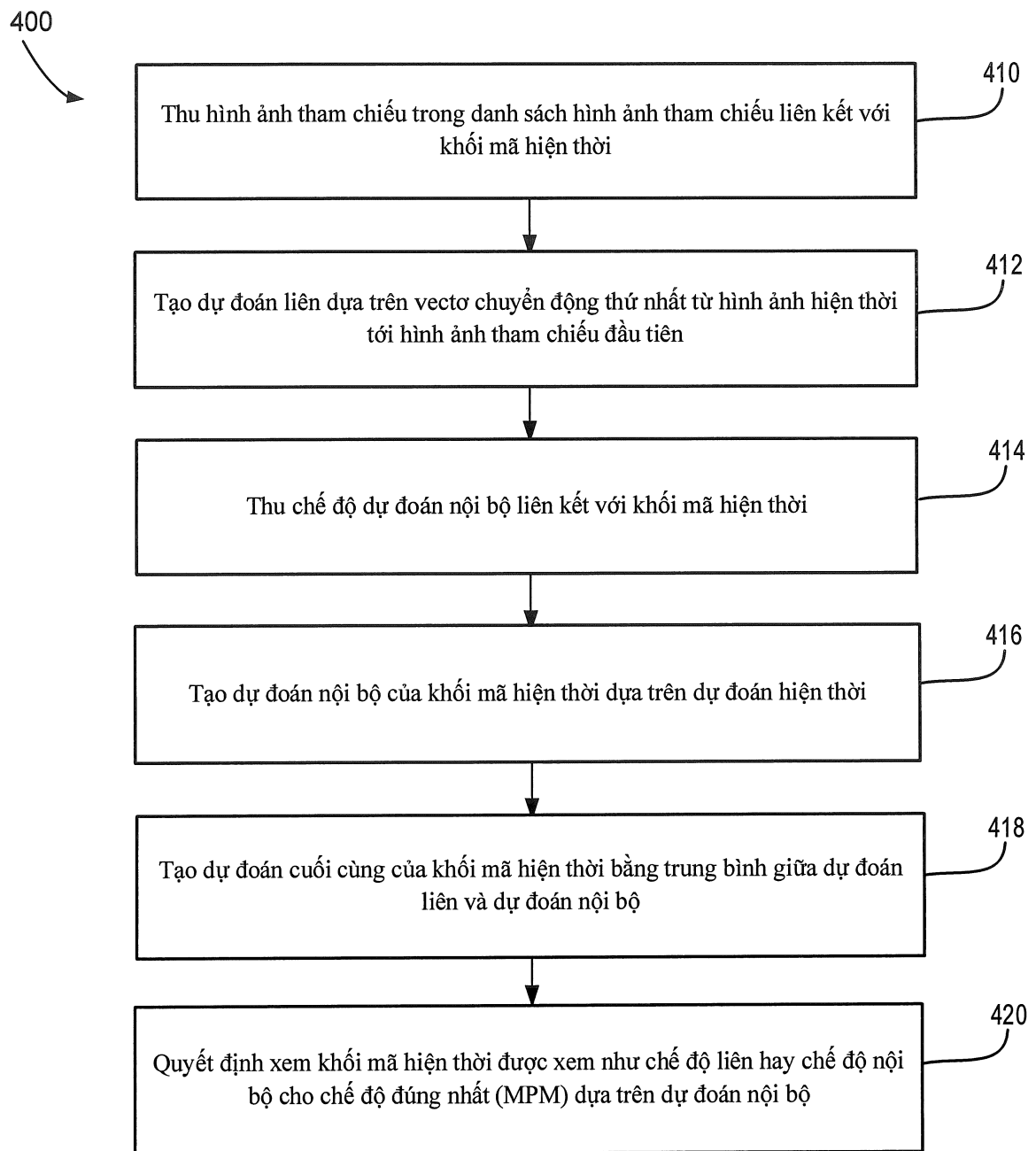
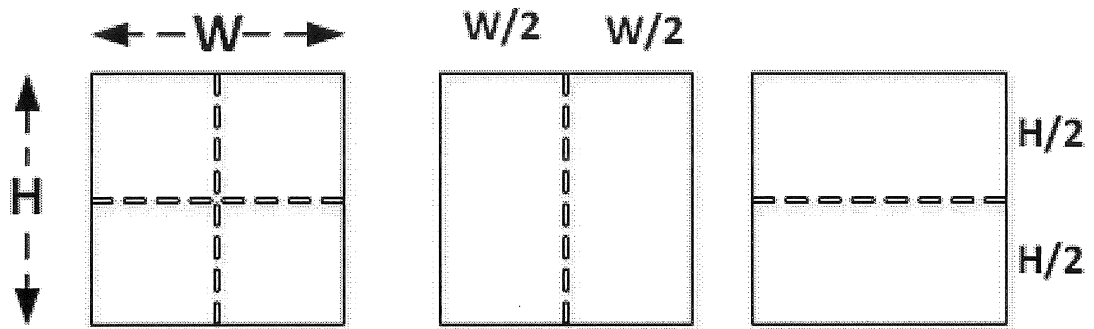
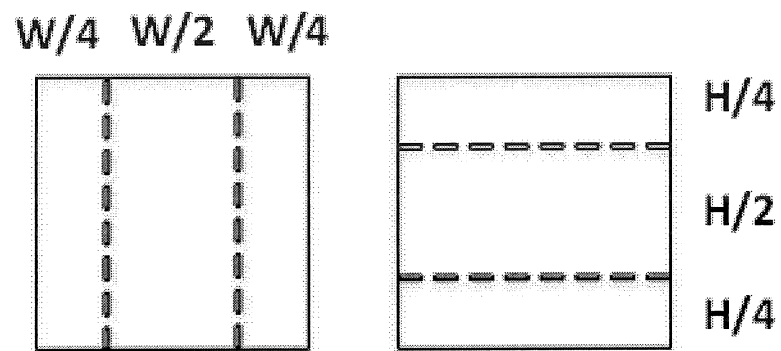


FIG. 4

**FIG. 5A****FIG. 5B****FIG. 5C****FIG. 5D****FIG. 5E**

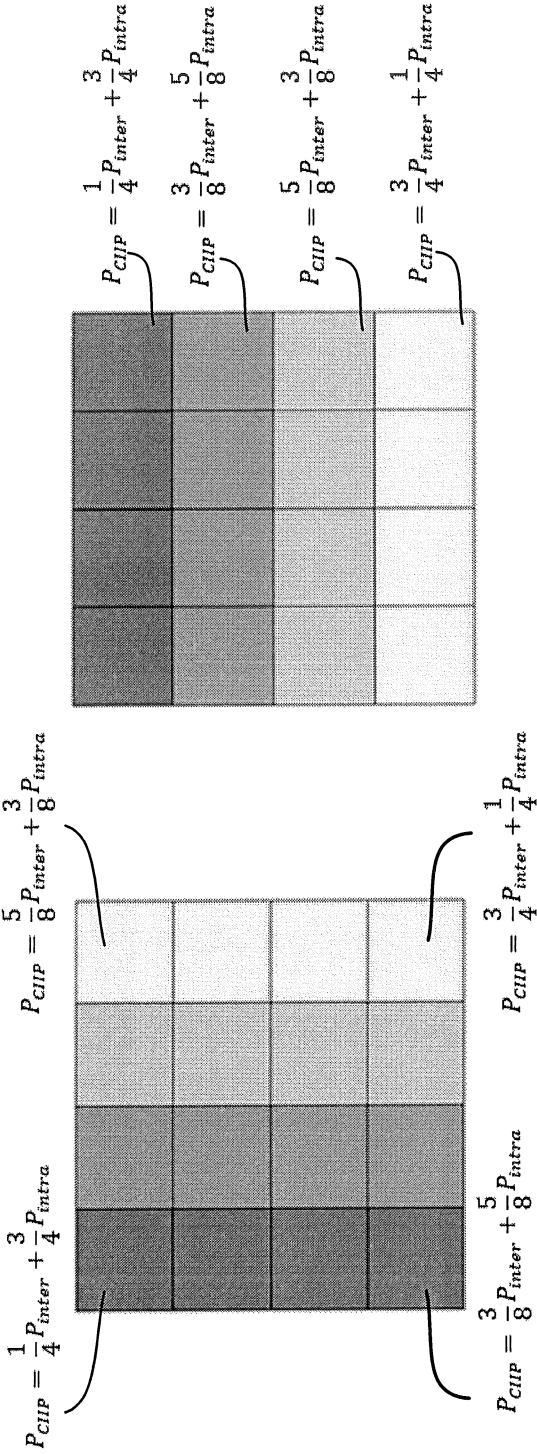


FIG. 6A

FIG. 6B

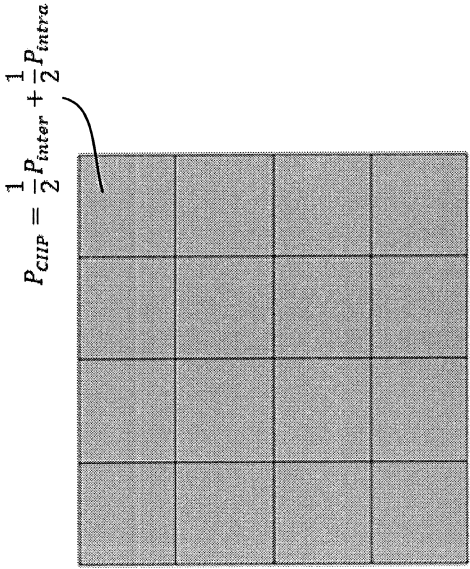


FIG. 6C

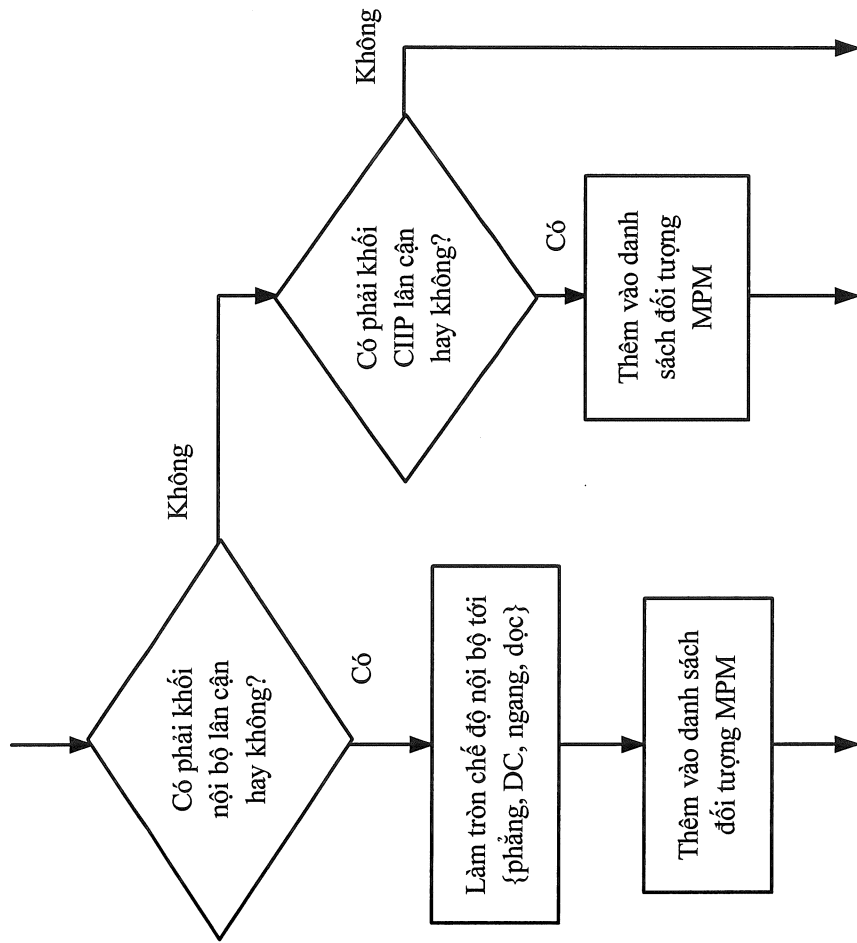


FIG. 7B

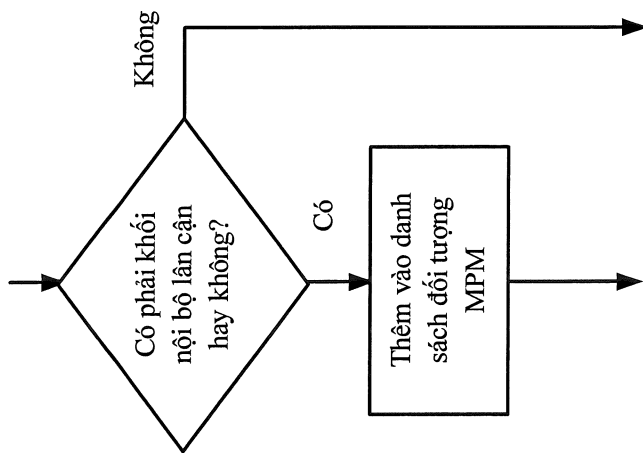


FIG. 7A

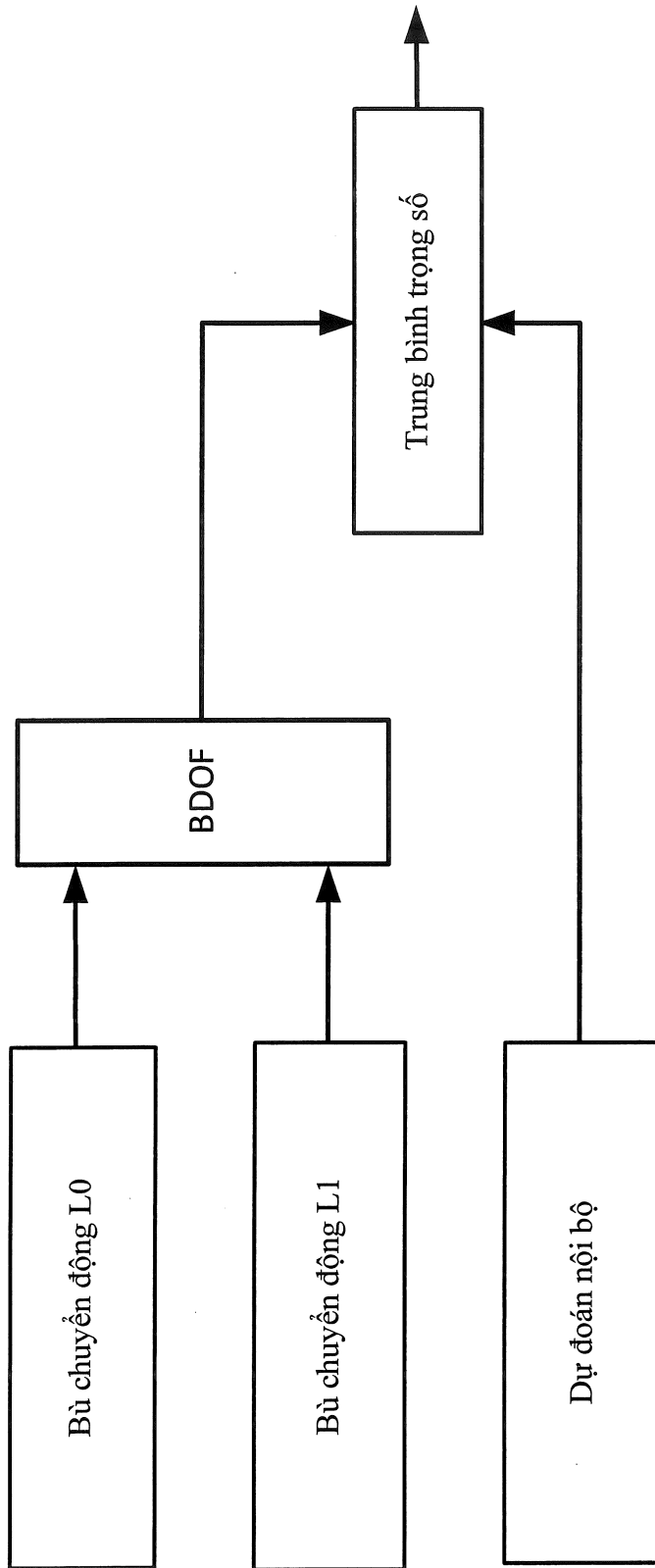


FIG. 8

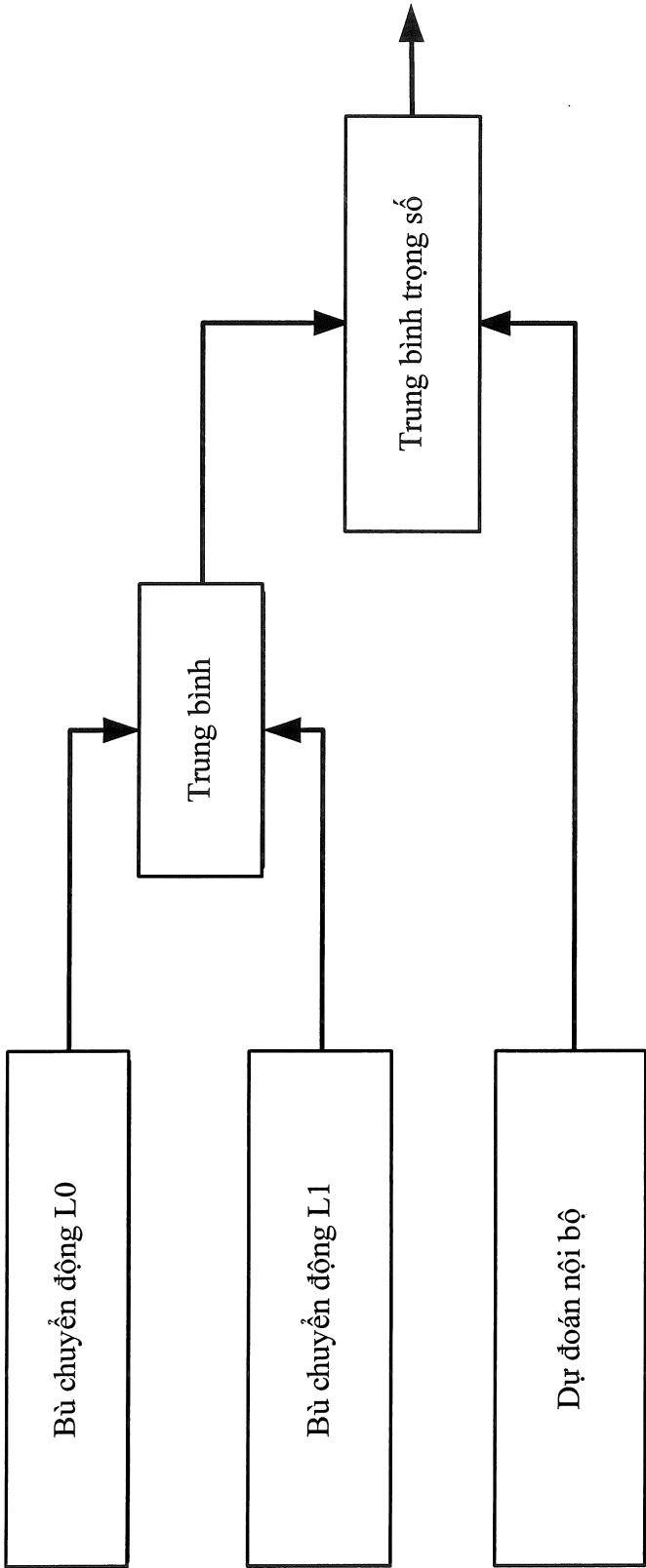
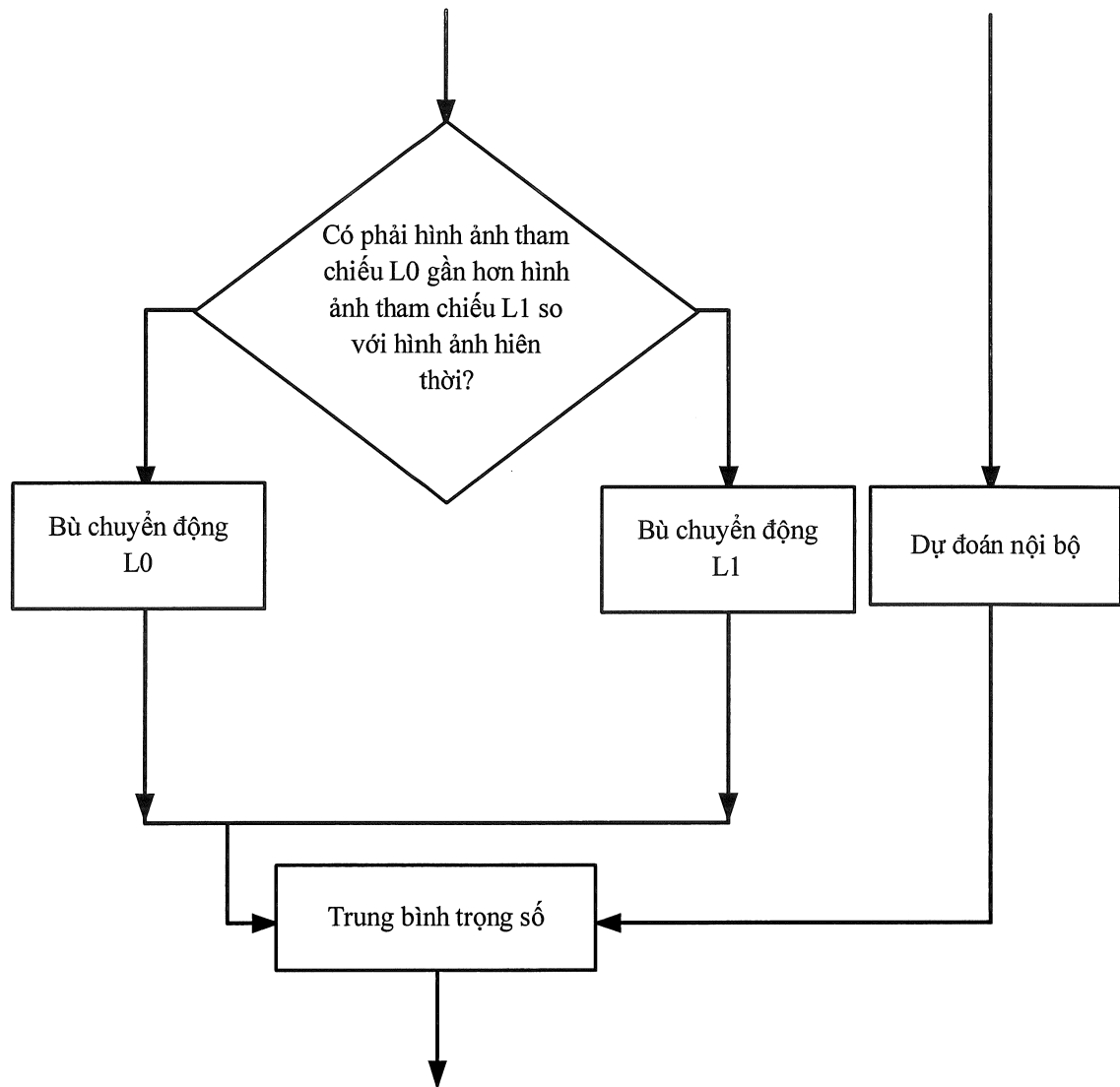
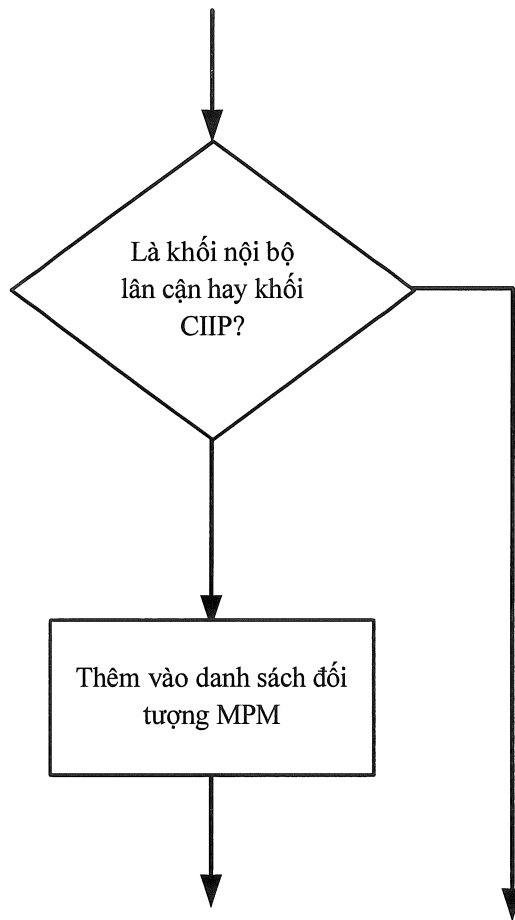
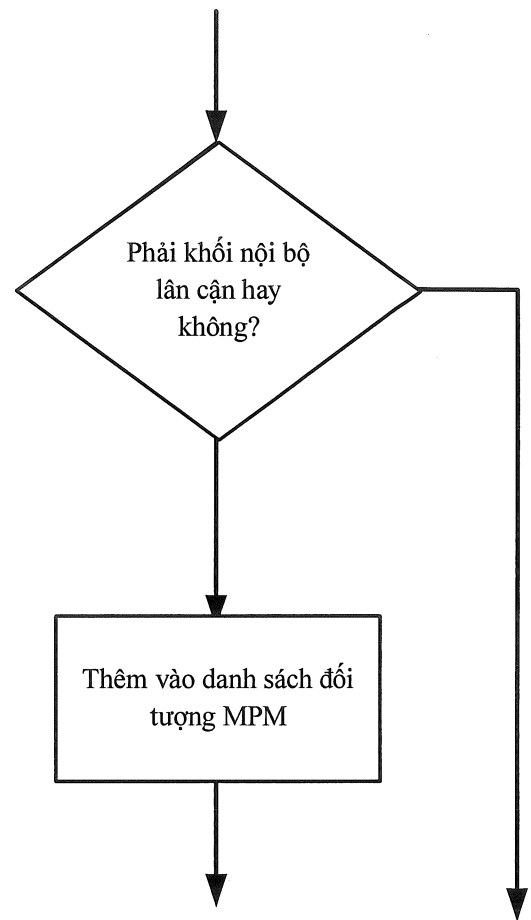
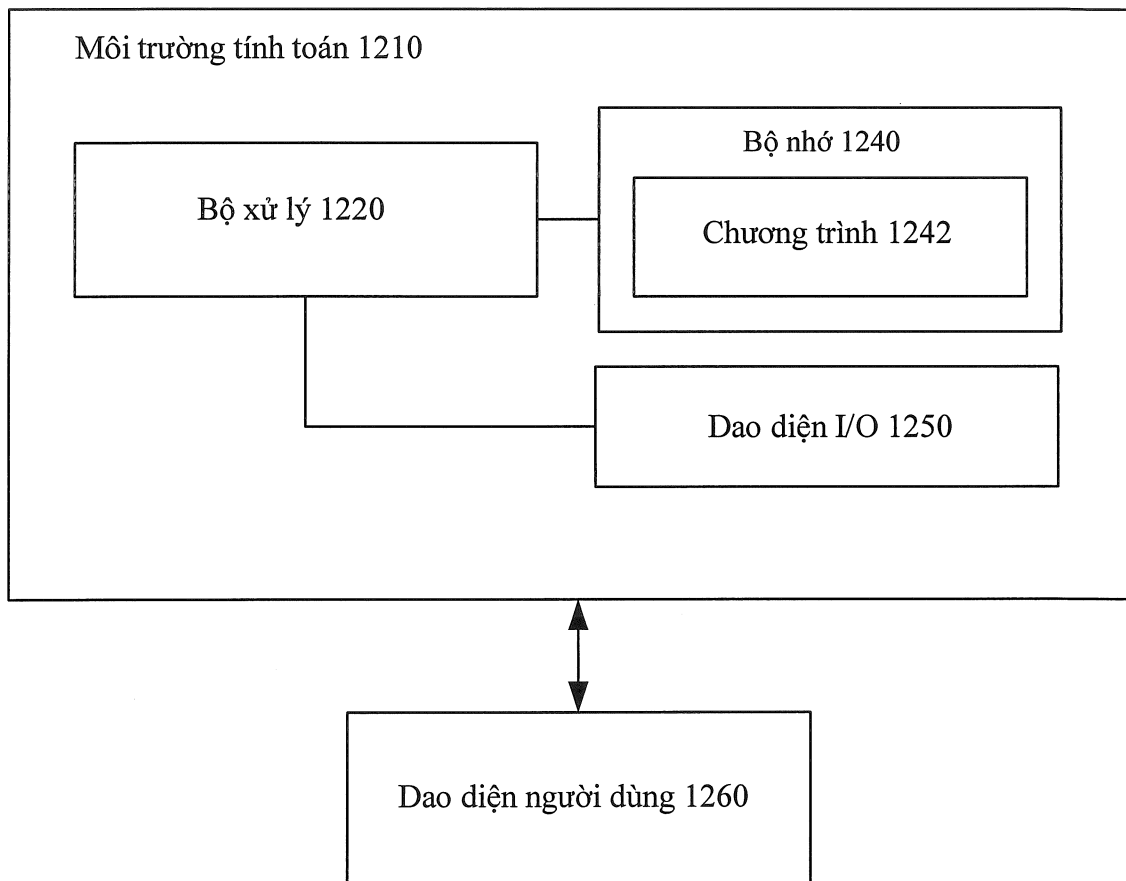


FIG. 9

**FIG. 10**

**FIG. 11A****FIG. 11B**

**FIG. 12**