



(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



1-0049537

(51)^{2021.01} G06T 15/08; H04N 19/597; G06T 17/00 (13) B

-
- | | |
|--|--------------------------------|
| (21) 1-2022-04133 | (22) 17/12/2020 |
| (86) PCT/EP2020/086623 17/12/2020 | (87) WO/2021/122881 24/06/2021 |
| (30) 19306697.4 19/12/2019 EP | |
| (45) 25/08/2025 449 | (43) 26/09/2022 414A |
| (73) InterDigital CE Patent Holdings, SAS (FR)
3 rue du Colonel Moll, Paris, 75017 France | |
| (72) Renaud DORE (FR); Bertrand CHUPEAU (FR); Franck THUDOR (FR); Julien FLEUREAU (FR). | |
| (74) Văn phòng Luật sư Ân Nam (ANNAM IP & LAW) | |
-

(54) PHƯƠNG PHÁP VÀ THIẾT BỊ MÃ HÓA VÀ GIẢI MÃ CẢNH BA CHIỀU TRONG LUỒNG DỮ LIỆU

(21) 1-2022-04133

(57) Sáng chế đề xuất các phương pháp và thiết bị để mã hóa và giải mã dữ liệu mô tả cảnh 3D. Tập hợp các miếng vá thứ nhất được tạo từ nội dung MVD thứ nhất thu được từ vùng thứ nhất của cảnh 3D. Miếng vá là một phần của một trong các góc nhìn của nội dung MVD. Tập hợp các miếng vá thứ hai được tạo từ nội dung MVD thứ hai thu được từ vùng thứ hai của cảnh 3D. Tập bản ảnh đóng gói các miếng vá thứ nhất và thứ hai được tạo và liên kết với siêu dữ liệu biểu thị, đối với miếng vá của tập bản ảnh, liệu miếng vá đó là miếng vá thứ nhất hay thứ hai. Ở phía giải mã, các miếng vá thứ nhất được sử dụng để kết xuất hình ảnh vùng nhìn và các miếng vá thứ hai được sử dụng để xử lý trước hoặc xử lý sau hình ảnh vùng nhìn.

Lĩnh vực kỹ thuật được đề cập

Sáng chế nhìn chung đề cập đến miền cảnh ba chiều (3D) và nội dung video ba chiều. Tài liệu bằng sáng chế cũng được hiểu theo ngữ cảnh mã hóa, định dạng và giải mã dữ liệu biểu thị kết cấu và hình học của cảnh 3D để kết xuất nội dung ba chiều trên thiết bị của người dùng cuối như thiết bị di động hoặc Màn hình gắn trên đầu (HMD).

Tình trạng kỹ thuật của sáng chế

Phần này có chủ đích giới thiệu cho người đọc các phương án khác nhau của lĩnh vực kỹ thuật, có thể liên quan đến các phương án khác nhau của sáng chế được mô tả và/hoặc yêu cầu bảo hộ bên dưới. Phần trình bày này được cho là hữu dụng khi cung cấp cho người đọc thông tin về tình trạng kỹ thuật của sáng chế nhằm tạo điều kiện hiểu rõ hơn về các phương án khác nhau của sáng chế. Theo đó, cần hiểu rằng các tuyên bố này phải đọc theo hướng vừa được trình bày mà không phải là sự thừa nhận lĩnh vực kỹ thuật trước đó.

Gần đây, đã có sự phát triển nội dung trường nhìn rộng lớn có sẵn (lên đến 360°). Nội dung này có thể sẽ không được hiển thị đầy đủ khi người dùng xem nội dung đó trên các thiết bị hiển thị sống động như màn hình gắn trên đầu, kính thông minh, màn hình PC, máy tính bảng, điện thoại thông minh, và các thiết bị tương tự. Điều đó có nghĩa là tại một thời điểm nhất định, người dùng có thể chỉ đang xem một phần nội dung. Tuy nhiên, người dùng thường có thể điều hướng bên trong nội dung bằng nhiều cách khác nhau như di chuyển đầu, di chuyển chuột, màn hình cảm ứng, giọng nói và những thứ tương tự. Người ta thường mong muốn mã hóa và giải mã nội dung này.

Video nhập vai, còn được gọi là video phẳng 360° , cho phép người dùng xem toàn cảnh xung quanh mình thông qua các thao tác xoay đầu xung quanh một điểm tĩnh. Các thao tác xoay chỉ cho phép trải nghiệm 3 bậc tự do (3DoF). Ngay cả khi video 3DoF đã đủ cho trải nghiệm video đa hướng đầu tiên, chẳng hạn như sử dụng thiết bị Màn hình gắn trên đầu (HMD), thì video 3DoF có thể nhanh chóng trở nên khó chịu đối với người xem yêu cầu cao hơn về sự thoải mái, chẳng hạn như bằng trải nghiệm thị sai. Ngoài ra, 3DoF cũng có thể gây chóng mặt vì người dùng không bao giờ chỉ xoay đầu mà còn phải dịch chuyển đầu theo ba hướng, các thao tác dịch chuyển không được tái tạo trong trải nghiệm video 3DoF.

Trong số những nội dung khác, trường nhìn rộng lớn có thể là cảnh hình ảnh đồ họa máy tính ba chiều (cảnh 3D CGI), đám mây điểm hoặc video sống động. Nhiều thuật ngữ có thể được sử dụng để thiết kế các video sống động này: Thực tế ảo (VR), 360, toàn cảnh, 4π steradian, trường nhìn sống động, đa hướng hoặc rộng lớn.

Video ba chiều (còn được gọi là video 6 Bậc tự do (6DoF)) là một giải pháp thay thế cho video 3DoF. Khi xem video 6DoF, ngoài các thao tác xoay, người dùng cũng có thể dịch chuyển đầu và thậm chí cả cơ thể của mình trong nội dung được xem, đồng thời trải nghiệm thị sai và thậm chí cả không gian ba chiều. Những video này giúp làm tăng đáng kể cảm giác sống động và cảm nhận về độ sâu của cảnh, đồng thời cũng ngăn ngừa hiện tượng chóng mặt bằng việc cung cấp phản hồi trực quan nhất quán trong quá trình dịch chuyển đầu. Nội dung được tạo bằng các cảm biến chuyên dụng cho phép ghi lại đồng thời màu sắc và độ sâu của cảnh đề cập. Việc sử dụng hệ thống máy ảnh màu kết hợp với kỹ thuật quang trắc là một cách để thực hiện việc ghi hình này, ngay cả khi vẫn có những khó khăn về mặt kỹ thuật.

Trong khi video 3DoF bao gồm chuỗi hình ảnh thu được từ việc không kết nối các hình ảnh kết cấu (ví dụ, hình ảnh hình cầu được mã hóa theo công nghệ hình ảnh 2 chiều được kết nối vĩ độ/kinh độ hoặc công nghệ hình ảnh 2 chiều được kết nối đẳng hình chữ nhật), khung video 6DoF nhúng thông tin từ một số điểm nhìn. Chúng có thể được xem như là một chuỗi các đám mây điểm theo thời gian thu được từ bước chụp ba chiều. Hai loại video ba chiều có thể được xem xét tùy thuộc vào điều kiện xem. Loại thứ nhất (nghĩa là 6DoF hoàn chỉnh) cho phép điều hướng hoàn toàn tự do bên trong nội dung video trong khi loại thứ hai (hay còn gọi là 3DoF+) giới hạn không gian xem của người dùng đến không gian giới hạn được gọi là hộp giới hạn xem, cho phép dịch chuyển giới hạn của trải nghiệm đầu và thị sai. Ngữ cảnh thứ hai này là sự đánh đổi có giá trị giữa điều hướng tự do và điều kiện xem/nhìn thụ động của khán giả đang ngồi.

Trong trường hợp 3DoF+, phương pháp chỉ bao gồm việc gửi thông tin cần thiết để xem cảnh 3D từ điểm bất kỳ trong hộp giới hạn xem. Một phương pháp khác xem xét việc gửi thêm thông tin về hình dạng và/hoặc màu sắc, không thấy được từ hộp giới hạn xem nhưng hữu ích để thực hiện quy trình khác ở phía bộ giải mã như chiếu sáng lại, phát hiện va chạm hoặc tương tác xúc giác. Thông tin bổ sung này có thể được truyền tải theo cùng một định dạng so với các điểm nhìn thấy được. Tuy nhiên, cần có định dạng và phương pháp để chỉ báo cho bộ giải mã biết rằng một phần thông tin sẽ được

sử dụng để kết xuất và một phần thông tin khác sẽ được sử dụng cho quy trình xử lý khác.

Bản chất kỹ thuật của sáng chế

Phần sau đây trình bày ngắn gọn bản chất kỹ thuật của sáng chế để tạo sự hiểu biết cơ bản về một số phương án của sáng chế. Phần bản chất kỹ thuật này không phải là sự miêu tả chung bao quát về sáng chế. Phần này không nhằm mục đích xác định các yếu tố chính hoặc trọng yếu của sáng chế. Phần tóm tắt sau chỉ trình bày một số phương án của sáng chế ở dạng giản lược dưới dạng phần mở đầu cho mô tả chi tiết hơn được đề xuất bên dưới.

Sáng chế đề cập đến phương pháp mã hóa dữ liệu mô tả cảnh 3D trong luồng dữ liệu. Phương pháp này bao gồm các bước:

- tạo tập hợp các miếng vá thứ nhất từ nội dung đa góc nhìn cộng với độ sâu (MVD) thứ nhất thu được để kết xuất cảnh 3D. MVD thứ nhất được lấy từ vùng thứ nhất của cảnh 3D. Miếng vá là một phần của một trong các góc nhìn của nội dung MVD.
- tạo tập hợp các miếng vá thứ hai từ nội dung MVD thứ hai thu được để áp dụng bước xử lý trước hoặc xử lý sau; MVD thứ hai thu được từ vùng thứ hai của cảnh 3D. Vùng thứ hai có thể chồng chéo lên vùng thứ nhất hoặc được tách riêng.
- tạo tập bản ảnh với các miếng vá thứ nhất và thứ hai. Tập bản ảnh là hình ảnh đóng gói các miếng vá theo bố cục tập bản ảnh và được liên kết với siêu dữ liệu biểu thị, đối với miếng vá của tập bản ảnh, liệu miếng vá đó là miếng vá thứ nhất hay thứ hai; và
- mã hóa tập bản ảnh nêu trên trong luồng dữ liệu nêu trên.

Sáng chế cũng đề cập đến phương pháp giải mã dữ liệu mô tả cảnh 3D từ luồng dữ liệu. Phương pháp này bao gồm các bước:

- giải mã luồng dữ liệu để truy xuất tập bản ảnh và siêu dữ liệu được liên kết. Tập bản ảnh là hình ảnh đóng gói các miếng vá theo bố cục tập bản ảnh. Miếng vá là một phần của một khung hình trong nội dung MVD thu được từ vùng của cảnh 3D. Đối với miếng vá của tập bản ảnh, siêu dữ liệu bao gồm dữ liệu biểu thị liệu miếng vá đó là miếng vá thứ nhất hay thứ hai; miếng vá thứ nhất là một phần của nội dung MVD thu được từ vùng thứ nhất của cảnh

3D và miếng vá thứ hai là một phần của MVD thu được từ vùng thứ hai của cảnh 3D. Vùng thứ nhất và vùng thứ hai có thể chồng chéo lên nhau hoặc được tách riêng.

- kết xuất hình ảnh vùng nhìn từ điểm quan sát trong cảnh 3D bằng cách sử dụng các miếng vá được biểu thị là miếng vá thứ nhất trong siêu dữ liệu; và
- sử dụng các miếng vá được biểu thị là miếng vá thứ hai trong siêu dữ liệu để xử lý trước và/hoặc xử lý sau hình ảnh vùng nhìn nêu trên.

Sáng chế cũng đề cập đến thiết bị bao gồm bộ xử lý được tạo cấu hình để thực hiện phương pháp mã hóa ở trên và đề cập đến thiết bị bao gồm bộ xử lý được tạo cấu hình để thực hiện phương pháp giải mã ở trên.

Sáng chế cũng đề cập đến luồng dữ liệu và/hoặc phương tiện không chuyển tiếp mang dữ liệu mô tả cảnh 3D. Luồng dữ liệu hoặc phương tiện không chuyển tiếp bao gồm:

- hình ảnh tập bản ảnh đóng gói các miếng vá thứ nhất và thứ hai theo bố cục tập bản ảnh, miếng vá thứ nhất là một phần của một khung hình trong nội dung MVD thu được để kết xuất cảnh 3D, miếng vá thứ hai là một phần của một khung hình trong nội dung MVD thu được để áp dụng bước xử lý trước hoặc xử lý sau, và
- siêu dữ liệu được liên kết với tập bản ảnh nêu trên, siêu dữ liệu này bao gồm, đối với miếng vá của tập bản ảnh, dữ liệu biểu thị liệu miếng vá là miếng vá thứ nhất hay thứ hai.

Mô tả vắn tắt các hình vẽ

Sáng chế sẽ được hiểu rõ hơn và các đặc điểm và ưu điểm cụ thể khác sẽ xuất hiện khi đọc phần mô tả sau, phần mô tả tham chiếu đến các bản vẽ phụ lục trong đó:

Hình 1 minh họa mô hình ba chiều (3D) của vật thể và các điểm của đám mây điểm theo mô hình 3D, theo các phương án không giới hạn của sáng chế;

Hình 2 minh họa ví dụ không giới hạn về bước mã hóa, phát và giải mã dữ liệu biểu thị chuỗi cảnh 3D, theo các phương án không giới hạn của sáng chế;

Hình 3 minh họa ví dụ về cấu trúc của thiết bị mà có thể được tạo cấu hình để thực hiện phương pháp được mô tả theo hình 8 và 9, theo phương án không giới hạn của sáng chế;

Hình 4 minh họa ví dụ về phương án cú pháp của luồng khi dữ liệu được phát qua giao thức phát dựa trên gói, theo các phương án không giới hạn của sáng chế;

Hình 5 minh họa phép chiếu hình cầu từ điểm nhìn trung tâm, theo một phương án không giới hạn của sáng chế;

Hình 6 minh họa ví dụ về việc bộ mã hóa tạo tập bản ảnh 60 và 61, theo phương án không giới hạn của sáng chế

Hình 7 minh họa việc thu khung hình để kết xuất 3DoF+ cộng với các khung hình bổ sung dành cho miếng vá bổ trợ, theo phương án không giới hạn của sáng chế

Hình 8 minh họa phương pháp 80 giải mã nội dung video ba chiều bao gồm thông tin bổ trợ, theo phương án không giới hạn của sáng chế.

Hình 9 minh họa phương pháp 90 giải mã nội dung video ba chiều bao gồm thông tin bổ trợ, theo phương án không giới hạn của sáng chế.

Mô tả chi tiết về phương án minh họa

Sáng chế sẽ được mô tả đầy đủ hơn trong tài liệu này theo các hình kèm theo, trong đó trình bày các ví dụ về sáng chế. Tuy nhiên, sáng chế có thể được thể hiện dưới nhiều dạng thay thế và không nên được hiểu là bị giới hạn ở các ví dụ nêu trong tài liệu này. Do đó, mặc dù sáng chế dễ bị cải biến theo nhiều cách và các hình thức thay thế khác nhau, các ví dụ cụ thể của sáng chế được minh họa bằng các ví dụ trong hình vẽ và sẽ được mô tả chi tiết trong tài liệu này. Tuy nhiên cần phải hiểu rằng không có dự định giới hạn sáng chế theo các dạng thức cụ thể được đề xuất, nhưng ngược lại, sáng chế phải bao gồm tất cả sự cải biến, phương thức tương đương và thay thế nằm trong tinh thần và phạm vi của sáng chế được định nghĩa bởi các điểm yêu cầu bảo hộ.

Thuật ngữ được sử dụng trong bản mô tả này chỉ nhằm mục đích mô tả các ví dụ cụ thể, và không nhằm mục đích giới hạn sáng chế. Như được sử dụng trong tài liệu này, các dạng số ít “một”, và mạo từ được dùng để bao hàm các dạng số nhiều, trừ khi được chỉ rõ khác theo ngữ cảnh. Cần phải hiểu thêm rằng các thuật ngữ “gồm có “đang gồm có”, “gồm có” và/hoặc “bao gồm” khi được sử dụng trong phần đặc tả này, có thể chỉ rõ sự hiện diện của các đặc điểm nêu trên, ví dụ như các tính năng, số nguyên, các bước, hoạt động, phần tử và/hoặc các thành phần nhưng không loại trừ sự hiện diện hoặc bổ sung của một hoặc nhiều tính năng, số nguyên, bước, hoạt động, phần tử, thành phần và/hoặc nhóm của chúng. Hơn nữa, khi một phần tử được gọi là “đáp ứng” hoặc “được kết nối” với một phần tử khác, thì phần tử đó có thể đáp ứng trực tiếp hoặc được kết nối

với phần tử kia, hoặc có thể có sự hiện diện của các phần tử xen kẽ. Ngược lại, khi một phần tử được gọi là “phản hồi trực tiếp” hoặc “được kết nối trực tiếp” với phần tử khác, thì sẽ không có sự hiện diện của bất kỳ phần tử xen kẽ nào. Như được sử dụng trong tài liệu này, “và/hoặc” bao gồm bất kỳ và tất cả sự kết hợp của một hoặc nhiều mục được liệt kê liên quan và có thể được viết tắt dưới dạng “/”.

Điều này sẽ được hiểu là, mặc dù các thuật ngữ thứ nhất, thứ hai, v.v. có thể được sử dụng trong tài liệu này để mô tả các yếu tố khác nhau, nhưng các yếu tố này không nên bị giới hạn bởi các thuật ngữ này. Các thuật ngữ này chỉ được sử dụng để phân biệt một phần tử này với một phần tử khác. Ví dụ, phần tử thứ nhất có thể được gọi là phần tử thứ hai, và tương tự, phần tử thứ hai có thể được gọi là phần tử thứ nhất mà không tách rời khỏi những đề xuất trong đơn xin cấp bằng sáng chế này.

Mặc dù một số sơ đồ gồm có các mũi tên trên các đường dẫn giao tiếp để minh họa hướng giao tiếp chính, nhưng cần hiểu rằng giao tiếp vẫn có thể xảy ra theo hướng ngược lại với các mũi tên được mô tả.

Một số ví dụ được mô tả liên quan đến sơ đồ khối và lưu đồ hoạt động trong đó mỗi khối đại diện cho cho thành phần mạch, mô-đun, một phần của mã, bao gồm một hoặc nhiều lệnh thực thi để triển khai (các) hàm logic được chỉ định. Cũng cần lưu ý rằng, trong các phương án triển khai khác, (các) hàm được đề cập trong các khối có thể xảy ra không theo thứ tự đã lưu ý. Ví dụ, trên thực tế, hai khối được minh họa liên tiếp, về cơ bản, có thể được thực thi đồng thời hoặc, đôi khi, các khối có thể được thực thi theo thứ tự ngược lại, tùy thuộc vào chức năng liên quan.

Tham chiếu trong tài liệu này “theo ví dụ” hoặc “trong ví dụ” nghĩa là tính năng, cấu trúc, hoặc đặc điểm được mô tả liên quan đến ví dụ có thể được bao gồm trong ít nhất một sự thực hiện sáng chế. Sự xuất hiện của cụm từ “theo ví dụ” hoặc “trong ví dụ” ở các vị trí khác nhau trong phần đặc tả không nhất thiết phải đề cập đến cùng một ví dụ, các ví dụ riêng biệt hoặc thay thế cũng không nhất thiết phải loại trừ lẫn nhau các ví dụ khác.

Các chữ số tham chiếu xuất hiện trong các điểm yêu cầu bảo hộ chỉ mang tính chất minh họa và sẽ không có ảnh hưởng hạn chế đến phạm vi của các yêu cầu bảo hộ. Mặc dù không được mô tả rõ ràng, nhưng các ví dụ và biến thể hiện tại có thể được sử dụng trong bất kỳ phương án kết hợp hoặc phương án kết hợp phụ nào.

Hình 1 minh họa mô hình (3D) 10 của vật thể và các điểm của đám mây điểm 11 tương ứng với mô hình 3D 10. Ví dụ, mô hình 3D 10 và đám mây điểm 11 có thể tương ứng với sự biểu diễn 3D có thể về vật thể của cảnh 3D bao gồm các vật thể khác. Mô hình 10 có thể là biểu diễn lưới 3D và các điểm của đám mây điểm 11 có thể là các đỉnh của lưới. Các điểm của đám mây điểm 11 cũng có thể là các điểm trải dài trên bề mặt các mặt của lưới. Mô hình 10 cũng có thể được biểu diễn dưới dạng phiên bản rải rác của đám mây điểm 11, bề mặt của mô hình 10 được tạo ra bằng cách rải rác các điểm của đám mây điểm 11. Mô hình 10 có thể được biểu diễn bằng rất nhiều cách biểu diễn khác nhau như voxel (đại diện cho một giá trị trên một lưới thông thường trong không gian ba chiều) hoặc spline. Hình 1 minh họa thực tế rằng một đám mây điểm có thể được xác định với biểu diễn bề mặt của vật thể D và sự biểu diễn bề mặt của vật thể 3D có thể được tạo từ điểm đám mây. Như được sử dụng trong tài liệu này, việc chiếu các điểm của vật thể 3D (bằng các điểm mở rộng của cảnh 3D) lên hình ảnh tương đương với việc chiếu mọi sự biểu diễn bất kỳ của đối tượng 3D này, ví dụ như đám mây điểm, lưới, mô hình spline hoặc mô hình voxel.

Ví dụ, đám mây điểm có thể được biểu diễn trong bộ nhớ dưới dạng cấu trúc dựa trên vector, trong đó mỗi điểm có tọa độ riêng trong hệ quy chiếu của điểm nhìn (ví dụ, tọa độ ba chiều XYZ, hoặc một góc khối và một khoảng cách (còn được gọi là chiều sâu) từ/đến điểm nhìn) và một hoặc nhiều thuộc tính, còn được gọi là thành phần. Ví dụ về thành phần là thành phần màu có thể được thể hiện trong các không gian màu khác nhau, ví dụ RGB (Đỏ, Xanh lục và Xanh lam) hoặc YUV (Y là thành phần luma và hai thành phần sắc độ UV). Đám mây điểm là sự biểu diễn cảnh 3D bao gồm các vật thể. Cảnh 3D có thể được nhìn thấy từ một góc nhìn nhất định hoặc phạm vi điểm nhìn. Có thể thu đám mây điểm theo nhiều cách, ví dụ:

- từ việc chụp hình ảnh thực được chụp bởi hệ thống camera, được bổ sung tùy chọn bởi thiết bị cảm biến độ sâu chủ động;
- từ việc chụp vật thể ảo/tổng hợp được chụp bởi hệ thống camera ảo trong công cụ mô hình hóa;
- từ sự kết hợp của cả vật thể thực và vật thể ảo.

Hình 2 minh họa ví dụ không giới hạn về bước mã hóa, phát và giải mã dữ liệu biểu thị chuỗi cảnh 3D. Ví dụ, định dạng mã hóa có thể tương thích đồng thời với bước giải mã 3DoF, 3DoF+ và 6DoF.

Thu được chuỗi cảnh 3D 20. Vì chuỗi ảnh là video 2D, chuỗi cảnh 3D là video 3D (còn gọi là ba chiều). Chuỗi cảnh 3D có thể được cung cấp cho thiết bị kết xuất video ba chiều để kết xuất và hiển thị 3DoF, 3DoF+ hoặc 6DoF.

Chuỗi cảnh 3D 20 được cung cấp cho bộ mã hóa 21. Bộ mã hóa 21 lấy cảnh 3D hoặc chuỗi cảnh 3D làm đầu vào và cung cấp chuỗi bit biểu thị đầu vào. Có thể lưu trữ chuỗi bit trong bộ nhớ 22 và/hoặc trên phương tiện dữ liệu điện tử và có thể được phát qua mạng 22. Chuỗi bit biểu thị chuỗi cảnh 3D có thể được đọc từ bộ nhớ 22 và/hoặc thu được từ mạng 22 bởi bộ giải mã 23. Bộ giải mã 23 được nhập bởi chuỗi bit này và cung cấp chuỗi các cảnh 3D, ví dụ tại định dạng đám mây điểm.

Bộ mã hóa 21 có thể bao gồm một số mạch thực hiện một số bước. Ở bước thứ nhất, bộ mã hóa 21 chiếu mỗi cảnh 3D lên ít nhất một hình ảnh 2D. Chiếu 3D là phương pháp chiếu các điểm ba chiều với mặt phẳng hai chiều. Vì hầu hết các phương pháp hiện tại để hiển thị dữ liệu đồ họa dựa trên phương tiện hai chiều phẳng (thông tin pixel từ một số mặt phẳng bit), việc sử dụng kiểu chiếu này rất phổ biến, đặc biệt là trong đồ họa máy tính, kỹ thuật và vẽ thiết kế. Mạch chiếu 211 cung cấp ít nhất khung hai chiều 2111 cho cảnh 3D của chuỗi 20. Khung 2111 gồm có thông tin màu và thông tin độ sâu biểu diễn cảnh 3D được chiếu lên khung 2111. Theo một biến thể, thông tin màu và thông tin độ sâu được mã hóa trong hai khung hình riêng biệt 2111 và 2112.

Siêu dữ liệu 212 được sử dụng và cập nhật bởi mạch chiếu 211. Siêu dữ liệu 212 bao gồm thông tin về hoạt động chiếu (ví dụ, tham số chiếu) và về cách tổ chức thông tin về độ sâu và màu trong các khung hình 2111 và 2112 như được mô tả liên quan đến các hình từ 5 đến 7.

Mạch mã hóa video 213 bao gồm chuỗi khung hình 2111 và 2112 dưới dạng video. Hình ảnh của cảnh 3D 2111 và 2112 (hoặc chuỗi hình ảnh của cảnh 3D) được mã hóa trong luồng bằng bộ mã hóa video 213. Sau đó, dữ liệu video và siêu dữ liệu 212 được đóng gói trong luồng dữ liệu bởi mạch đóng gói dữ liệu 214.

Ví dụ, bộ mã hóa 213 tương thích với bộ mã hóa như là:

- JPEG, tiêu chuẩn ISO/CEI 10918-1 UIT-T khuyến nghị T.81, <https://www.itu.int/rec/T-REC-T.81/en>;
- AVC, còn được gọi là MPEG-4 AVC hoặc h264. Được định rõ trong cả phần UIT-T H.264 và ISO/CEI MPEG-4 10 (ISO/CEI 14496-10), <http://www.itu.int/rec/T-REC-H.264/en>, HEVC (có thể tìm đặc điểm kỹ

thuật tại trang web ITU, khuyến nghị T, sê-ri H, h265, <http://www.itu.int/rec/T-REC-H.265-201612-I/en>);

- 3D-HEVC (phần mở rộng của HEVC có đặc điểm kỹ thuật được tìm thấy tại trang web ITU, khuyến nghị T, sê-ri H, h265, <http://www.itu.int/rec/T-REC-H.265-201612-I/en> phụ lục G và I);
- VP9 được phát triển bởi Google; hoặc
- AV1 (AOMedia Video 1) được phát triển bởi Alliance for Open Media.

Luồng dữ liệu được lưu trữ trong bộ nhớ có thể truy cập được, ví dụ qua mạng 22, bởi bộ giải mã 23. Bộ giải mã 23 gồm có các mạch khác nhau thực hiện các bước giải mã khác nhau. Bộ giải mã 23 lấy luồng dữ liệu do bộ mã hóa 21 tạo làm đầu vào và cung cấp chuỗi cảnh 3D 24 sẽ được kết xuất và hiển thị bằng thiết bị hiển thị video ba chiều, như Thiết bị gắn trên đầu (HMD). Bộ giải mã 23 lấy luồng từ nguồn 22. Ví dụ, nguồn 22 thuộc về tập hợp gồm có:

- bộ nhớ cục bộ, ví dụ, bộ nhớ video hoặc RAM (hoặc Bộ nhớ truy cập ngẫu nhiên), bộ nhớ flash, ROM (hoặc Bộ nhớ chỉ đọc), đĩa cứng;
- giao diện lưu trữ, ví dụ, giao diện với bộ nhớ chung, RAM, bộ nhớ flash, ROM, đĩa quang hoặc hỗ trợ từ tính;
- giao diện truyền thông, ví dụ, giao diện có dây (ví dụ, giao diện bus, giao diện mạng diện rộng, giao diện mạng cục bộ) hoặc giao diện không dây (như là giao diện IEEE 802.11 hoặc giao diện Bluetooth®); và
- giao diện người dùng như là Graphical User Interface (Giao diện đồ họa người dùng) cho phép người dùng nhập dữ liệu.

Bộ giải mã 23 gồm có mạch 234 để trích xuất dữ liệu được mã hóa trong luồng dữ liệu. Mạch 234 lấy luồng dữ liệu làm đầu vào và cung cấp siêu dữ liệu 232 tương ứng với siêu dữ liệu 212 được mã hóa trong luồng và video hai chiều. Video được giải mã bởi bộ giải mã video 233 cung cấp một chuỗi các khung hình. Khung hình được giải mã bao gồm thông tin về màu và độ sâu. Trong một biến thể, bộ giải mã video 233 cung cấp hai chuỗi khung hình, một chuỗi bao gồm thông tin về màu, chuỗi còn lại bao gồm thông tin về độ sâu. Mạch 231 sử dụng siêu dữ liệu 232 để không chiếu (un-project) thông tin về độ sâu và màu từ các khung hình được giải mã để cung cấp chuỗi các cảnh 3D 24. Chuỗi cảnh 3D 24 tương ứng với chuỗi cảnh 3D 20, có thể mất độ chính xác liên quan đến việc mã hóa dưới dạng video 2D và nén video.

Các mạch và chức năng khác có thể được thêm vào, chẳng hạn như trước bước bỏ chiếu theo mạch 231 hoặc trong bước xử lý sau sau khi bỏ chiếu. Ví dụ, các mạch có thể được thêm vào để chiếu sáng lại cảnh từ một ánh sáng khác đặt ở bất kỳ đâu trong cảnh. Có thể thực hiện phát hiện va chạm để tổng hợp chiều sâu, như thêm đối tượng mới vào cảnh 3DoF+ theo cách thực tế nhất quán hoặc để lập kế hoạch đường dẫn. Mạch này có thể cần thông tin hình dạng và/hoặc màu sắc về cảnh 3D mà không được sử dụng để tự kết xuất 3DoF+. Ngữ nghĩa của các loại thông tin khác nhau phải được biểu thị chuỗi bit biểu diễn cảnh 3DoF+.

Hình 3 minh họa ví dụ về cấu trúc của thiết bị 30 mà có thể được tạo cấu hình để thực hiện phương pháp được mô tả theo hình 8 và 9. Bộ mã hóa 21 và/hoặc bộ giải mã 23 của hình 2 có thể thực hiện cấu trúc này. Ngoài ra, mỗi mạch của bộ mã hóa 21 và/hoặc bộ giải mã 23 có thể là thiết bị theo cấu trúc theo Hình 3, được liên kết với nhau, ví dụ, thông qua bus 31 và/hoặc qua giao diện I/O 36 của chúng.

Thiết bị 30 bao gồm các thành phần sau được liên kết với nhau bằng bus dữ liệu và địa chỉ 31:

- bộ vi xử lý 32 (hoặc CPU), ví dụ, là DSP (hoặc Bộ xử lý tín hiệu số);
- ROM (hoặc là Bộ nhớ chỉ đọc) 33;
- RAM (hoặc là Bộ nhớ truy cập ngẫu nhiên) 34;
- giao diện lưu trữ 35;
- giao diện I/O 36 để thu dữ liệu để phát, từ ứng dụng; và
- nguồn điện, ví dụ ắc quy.

Theo ví dụ, nguồn điện là bên ngoài thiết bị. Trong mỗi bộ nhớ được đề cập, từ « thanh ghi » được sử dụng trong đặc điểm kỹ thuật có thể tương ứng với vùng có dung lượng nhỏ (một số bit) hoặc vùng rất lớn (ví dụ, toàn bộ chương trình hoặc lượng lớn dữ liệu đã thu hoặc được giải mã). ROM 33 bao gồm ít nhất chương trình và tham số. ROM 33 có thể lưu trữ các thuật toán và hướng dẫn để thực hiện các kỹ thuật theo sáng chế. Khi được bật, CPU 32 tải chương trình lên RAM và thực hiện các lệnh tương ứng.

RAM 34 bao gồm, trong thanh ghi, chương trình được thực thi bởi CPU 32 và được tải lên sau khi bật thiết bị 30, dữ liệu đầu vào trong thanh ghi, dữ liệu trung gian ở các trạng thái khác nhau của phương pháp trong thanh ghi và các biến khác được sử dụng để thực hiện phương pháp trong thanh ghi.

Các phương án thực hiện được mô tả trong bản mô tả này có thể được thực hiện, ví dụ phương pháp hoặc quy trình, thiết bị, sản phẩm chương trình máy tính, luồng dữ liệu, hoặc tín hiệu. Ngay cả khi chỉ được thảo luận trong ngữ cảnh về hình thức của phương án thực hiện duy nhất (ví dụ, được thảo luận chỉ là phương pháp hoặc thiết bị), thì phương án thực hiện có các tính năng được thảo luận này cũng có thể được thực hiện dưới các hình thức khác (ví dụ, chương trình). Thiết bị có thể được thực hiện trong, ví dụ như phần cứng, phần mềm và phần mềm điều khiển thích hợp. Các phương pháp có thể được thực hiện trong, ví dụ, thiết bị, chẳng hạn như bộ xử lý, dùng để chỉ các thiết bị xử lý nói chung, bao gồm, ví dụ, máy tính, bộ vi xử lý, mạch tích hợp hoặc thiết bị logic khả lập. Bộ xử lý cũng bao gồm các thiết bị truyền thông, chẳng hạn như máy tính, điện thoại di động, thiết bị kỹ thuật số hỗ trợ cá nhân/di động (“PDA”) và các thiết bị khác hỗ trợ truyền thông tin giữa những người dùng cuối.

Theo các ví dụ, thiết bị 30 được tạo cấu hình để thực hiện phương pháp được mô tả liên quan đến hình 8 và 9, và thuộc về tập hợp bao gồm:

- thiết bị di động;
- thiết bị giao tiếp;
- thiết bị game;
- máy tính bảng (hoặc máy tính bảng);
- máy tính xách tay;
- máy chụp ảnh tĩnh;
- camera video;
- chip mã hóa;
- máy chủ (ví dụ, máy chủ broadcast, máy chủ video theo yêu cầu hoặc máy chủ web).

Hình 4 minh họa ví dụ về phương án cú pháp của luồng khi dữ liệu được phát qua giao thức phát dựa trên gói. Hình 4 minh họa cấu trúc ví dụ 4 của luồng video ba chiều. Cấu trúc bao gồm trong vùng chứa tổ chức luồng trong các phần tử cú pháp độc lập. Cấu trúc có thể bao gồm phần đầu 41 là tập hợp dữ liệu chung cho mọi phần tử cú pháp của luồng. Ví dụ, phần đầu bao gồm một số siêu dữ liệu về các phần tử cú pháp, mô tả bản chất và vai trò của mỗi phần tử cú pháp đó. Phần đầu cũng có thể bao gồm một phần siêu dữ liệu 212 theo hình 2, ví dụ tọa độ của điểm nhìn trung tâm được sử dụng để chiếu các điểm của cảnh 3D lên khung 2111 và 2112. Cấu trúc bao gồm tải

trọng bao gồm phần tử của cú pháp 42 và ít nhất một phần tử cú pháp 43. Phần tử cú pháp 42 bao gồm dữ liệu biểu thị khung màu và độ sâu. Hình ảnh có thể đã được nén theo phương pháp nén video.

Phần tử cú pháp 43 là một phần tải trọng của luồng dữ liệu và có thể bao gồm siêu dữ liệu về cách các khung của phần tử cú pháp 42 được mã hóa, ví dụ các tham số được sử dụng để chiếu và đóng gói các điểm của cảnh 3D vào khung hình. Siêu dữ liệu này có thể được liên kết với từng khung hình của video hoặc với nhóm khung hình (còn được gọi là Nhóm hình ảnh (GoP) trong tiêu chuẩn nén video).

Hình 5 minh họa cách tiếp cận tập bản ảnh của miếng vá với ví dụ về 4 tâm chiếu. Cảnh 3D 50 bao gồm nhân vật. Ví dụ, tâm của hình chiếu 51 là camera phối cảnh và camera 53 là camera chiếu trực giao. Camera cũng có thể là camera đẳng hướng, như phép chiếu hình cầu (ví dụ, phép chiếu đẳng hình chữ nhật) hoặc phép chiếu hình khối. Các điểm 3D của cảnh 3D được chiếu lên mặt phẳng 2D được liên kết với các camera ảo đặt tại các tâm chiếu, theo hoạt động chiếu được mô tả trong dữ liệu chiếu của siêu dữ liệu. Theo ví dụ của hình 5, hình chiếu của các điểm được camera 51 chụp được chiếu lên miếng vá 52 theo phép chiếu phối cảnh và phép chiếu của các điểm được camera 53 chụp được ánh xạ lên miếng vá 54 theo phép chiếu trực giao.

Việc tạo nhóm các điểm ảnh được chiếu tạo ra miếng vá 2D, được đóng gói trong tập bản ảnh hình chữ nhật 55. Việc sắp xếp các miếng vá trong tập bản ảnh xác định bố cục tập bản ảnh. Theo một phương án, hai tập bản ảnh có bố cục giống nhau: một dành cho thông tin về kết cấu (nghĩa là màu) và một dành cho thông tin về độ sâu. Hai miếng vá được chụp bởi cùng một camera hoặc bởi hai camera khác nhau có thể bao gồm thông tin biểu diễn cùng một phần của cảnh 3D, ví dụ các miếng vá 54 và 56.

Hoạt động đóng gói tạo ra dữ liệu miếng vá cho mỗi miếng vá được tạo. Dữ liệu miếng vá bao gồm tham chiếu đến dữ liệu chiếu (ví dụ, chỉ mục trong bảng dữ liệu chiếu hoặc con trỏ (nghĩa là địa chỉ trong bộ nhớ hoặc trong luồng dữ liệu) tới dữ liệu chiếu) và thông tin mô tả vị trí và kích thước của miếng vá trong tập bản ảnh (ví dụ, tọa độ góc trên cùng bên trái, kích thước và chiều rộng tính bằng điểm ảnh). Các mục dữ liệu miếng vá được thêm vào siêu dữ liệu để được đóng gói trong luồng dữ liệu cùng với dữ liệu nén của một hoặc hai tập bản ảnh.

Hình 6 minh họa ví dụ về việc tạo ra tập bản ảnh 60 và 61 bởi bộ mã hóa. Tập bản ảnh 60 và 61 bao gồm thông tin kết cấu (ví dụ, dữ liệu RGB hoặc dữ liệu YUV) về

các điểm của cảnh 3D, theo phương án không giới hạn của sáng chế. Như đã giải thích liên quan đến hình 5, tập bản ảnh là các miếng vá nén hình ảnh. Ví dụ, bộ mã hóa lấy đầu vào là video nhiều khung hình+độ sâu bao gồm ba khung hình 62, 63 và 64 theo ví dụ trong hình 6. Bộ mã hóa loại bỏ phần dư thừa giữa các khung hình (bước cắt bớt) và đóng gói các miếng vá kết cấu và chiều sâu đã chọn vào một hoặc nhiều tập bản ảnh. Do đó, chuỗi bit bao gồm nhiều chuỗi video (ví dụ, chuỗi video HEVC) mang tập bản ảnh kết cấu và độ sâu, đi kèm với siêu dữ liệu mô tả thông số camera của khung hình đầu vào và bố cục tập bản ảnh.

Tập bản ảnh miếng vá bao gồm các cặp thành phần tập bản ảnh kết cấu và chiều sâu, có cùng kích thước hình ảnh và bố cục (cùng cách nén) cho kết cấu và chiều sâu. Theo một phương pháp, tập bản ảnh chỉ mang thông tin cần thiết để kết xuất 3DoF+ của cảnh từ bất kỳ điểm nào trong hộp giới hạn xem. Theo một phương pháp khác, tập bản ảnh có thể mang thêm thông tin về hình dạng và/hoặc màu sắc hữu ích cho quá trình xử lý khác như chiếu sáng lại cảnh hoặc phát hiện va chạm. Ví dụ, thông tin bổ sung này có thể là hình dạng mặt sau của đối tượng trong cảnh 3D. Các miếng vá này được gọi là miếng vá bổ trợ. Những miếng vá này không có nghĩa được kết xuất bởi bộ giải mã, mà được sử dụng bởi mạch xử lý trước hoặc xử lý sau của bộ giải mã.

Hình 7 minh họa việc thu khung hình để kết xuất 3DoF+ cộng với các khung hình bổ sung cho các miếng vá bổ trợ. Ở phía bộ mã hóa, việc tạo miếng vá bổ trợ có thể được thực hiện bằng các phương tiện khác nhau. Ví dụ, để thu được cảnh 70, có thể đặt nhóm 71 gồm các camera thực hoặc ảo thứ nhất hướng về phía trước của cảnh 70. Nhóm thứ hai 72 gồm các camera thực hoặc ảo được đặt để quan sát mặt sau và các mặt của cảnh 3 chiều. Theo phương án, camera 72 chụp khung hình ở độ phân giải thấp hơn camera 71. Camera 72 lấy hình dạng và/hoặc màu sắc của phần đối tượng ẩn. Những miếng vá thu được từ khung hình được chụp bằng camera 71 là miếng vá để kết xuất 3DoF+, còn miếng vá thu được từ khung hình do camera 72 chụp là miếng vá bổ trợ để hoàn thành mô tả về thông tin hình dạng và/hoặc màu sắc để xử lý trước hoặc xử lý sau. Siêu dữ liệu liên quan đến các miếng vá của tập bản ảnh có thể được định dạng để báo hiệu ngữ nghĩa của mỗi miếng vá. Ở phía bộ giải mã, bộ kết xuất vùng nhìn phải bỏ qua các miếng vá không hợp lệ cho việc kết xuất. Siêu dữ liệu cũng biểu thị mô-đun nào của bộ giải mã có thể sử dụng những miếng vá không hợp lệ để kết xuất này. Ví dụ, mạch tái tạo ánh sáng sẽ sử dụng thông tin bổ trợ này để cập nhật bản đồ hình dạng của nó từ

điểm nhìn của ánh sáng và thay đổi kết cấu chiếu sáng của toàn bộ cảnh tương ứng để tạo ra vùng bóng thích hợp.

Có nghĩa là để tạo các miếng vá bổ trợ liên kết với camera 72 chụp từ phía sau và các bên mà mô tả hình dạng của phần phía sau của đối tượng, thường ở độ phân giải thấp hơn, được thêm vào bộ mã hóa. Đầu tiên phải thu được khung hình độ sâu bổ sung từ phía sau và bên cạnh, có thể được thực hiện theo nhiều cách khác nhau. Trong trường hợp các đối tượng được tạo tổng hợp, hình ảnh độ sâu kết hợp với camera ảo đặt ở bất kỳ đâu sẽ trực tiếp thu được từ mô hình 3D. Để chụp 3D tự nhiên, có thể thêm camera màu và/hoặc camera độ sâu hoạt động ở giai đoạn chụp: camera độ sâu trực tiếp cung cấp khung hình độ sâu, còn thuật toán quang trắc ước tính độ sâu từ khung hình màu. Khi cả mô hình 3D và chụp bổ sung đều không dùng được, thuật toán hoàn thiện hình dạng lỗi có thể được sử dụng để tạo hình dạng khép kín hợp lý từ hình dạng mở được khôi phục từ camera trước. Sau đó, phần dư giữa các khung hình được loại bỏ bằng cách cắt bớt, theo cách tương tự như được thực hiện trên khung hình từ camera 71. Theo phương án, bước cắt bớt được thực hiện độc lập trên hai nhóm khung hình. Do đó, phần dư có thể có giữa các miếng vá thông thường và bổ sung sẽ không bị loại bỏ. Các miếng vá độ sâu bổ trợ kết quả được nén cùng với các miếng vá thông thường trong tập bản ảnh miếng vá độ sâu.

Theo một phương án khác, nếu các miếng vá bổ trợ chỉ được xác định cho độ sâu thì phần kết cấu trên tập bản ảnh sẽ vẫn trống nếu bố cục đồng nhất được sử dụng cho tập bản ảnh kết cấu và tập bản ảnh độ sâu. Ngay cả khi các miếng vá bổ trợ này được cho là được xác định ở độ phân giải thấp hơn, điều này cũng sẽ khiến mất không gian trong tập bản ảnh kết cấu. Theo phương án này, các bố cục khác nhau có thể được sử dụng cho tập bản ảnh có độ sâu và kết cấu, sự khác biệt này được biểu thị trong siêu dữ liệu liên quan đến tập bản ảnh.

Cú pháp khả thi đối với siêu dữ liệu mô tả tập bản ảnh có thể bao gồm khái niệm cấp cao, được gọi là 'entity_id': entity_id này cho phép kết nối nhóm các miếng vá vào một chỉ mục để xử lý ngữ nghĩa cấp cao, chẳng hạn như lọc hoặc ghép đối tượng. Cú pháp khả dĩ dành cho siêu dữ liệu thông số tập bản ảnh được trình bày trong bảng sau.

atlas_params(a) {	Bộ mô tả
num_patches_minus1[a]	u(16)

atlas_width_minus1[a]	u(16)
atlas_height_minus1[a]	u(16)
depth_occ_params_present_flag[a]	u(1)
for(p = 0; p <= num_patches_minus1; p++) {	
view_id[a][p]	u(v)
if(max_entities_minus1 > 0)	
entity_id[a][p]	u(v)
patch_width_in_view[a][p]	u(v)
patch_height_in_view[a][p]	u(v)
patch_pos_in_atlas_x[a][p]	u(v)
patch_pos_in_atlas_y[a][p]	u(v)
patch_pos_in_view_x[a][p]	u(v)
patch_pos_in_view_y[a][p]	u(v)
patch_rotation[a][p]	u(3)
if(depth_occ_params_present_flag[a])	
depth_occupancy(a, p)	
}	
}	

Theo phương án của sáng chế, các miềng và bổ trợ được xác định là các thực thể cụ thể, được gọi là thực thể bổ trợ. Sau đó, số lượng thực thể và chức năng của chúng (nghĩa là, có phải là thực thể bổ trợ hay không) được mô tả trong siêu dữ liệu như trình bày trong bảng sau:

iv_sequence_params() {	Bộ mô tả
ivs_profile_tier_level()	
depth_params_num_bits_minus8	u(4)
view_params_list()	
depth_low_quality_flag	u(1)
num_groups_minus1	ue(v)
max_entities_minus1	ue(v)
auxiliary_flag	u(1)
if(auxiliary_flag và max_entities_minus1 != 0) {	

for (e==0; e <= max_entities_minus1; e++) {	
auxiliary_entity_flag [e]	u(1)
}	
viewing_space_present_flag	u(1)
if(viewing_space_present_flag)	u(1)
viewing_space()	
ivs_sp_extension_present_flag	u(1)
if(ivs_sp_extension_present_flag) {	
while(more_data_in_payload())	
ivs_sp_extension_data_flag	u(1)
rbsp_trailing_bits()	
}	

auxiliary_flag bằng 1 biểu thị rằng phần mô tả hỗ trợ có sẵn cho mỗi cấu trúc thực thể.

auxiliary_entity_flag[e] bằng 1 biểu thị rằng các miếng vá liên quan đến thực thể e không dành cho việc kết xuất vùng nhìn.

Theo một phương án khác của sáng chế, các miếng vá hỗ trợ được báo hiệu ở cấp độ miếng vá bằng cách sửa đổi cú pháp thông số tập bản ảnh như được trình bày trong bảng sau:

atlas_params(a) {	Bộ mô tả
num_patches_minus1 [a]	u(16)
atlas_width_minus1 [a]	u(16)
atlas_height_minus1 [a]	u(16)
depth_occ_params_present_flag [a]	u(1)
auxiliary_flag [a]	u(1)
for(p = 0; p <= num_patches_minus1; p++) {	
view_id [a][p]	u(v)
if(max_entities_minus1 > 0)	
entity_id [a][p]	u(v)
if(auxiliary_flag[a])	
auxiliary_patch_flag [a][p]	u(1)
patch_width_in_view [a][p]	u(v)

patch_height_in_view [a][p]	u(v)
patch_pos_in_atlas_x [a][p]	u(v)
patch_pos_in_atlas_y [a][p]	u(v)
patch_pos_in_view_x [a][p]	u(v)
patch_pos_in_view_y [a][p]	u(v)
patch_rotation [a][p]	u(3)
if(depth_occ_params_present_flag[a])	
depth_occupancy(a, p)	
}	
}	

auxiliary_flag bằng 1 biểu thị rằng phần mô tả hỗ trợ có sẵn cho mỗi cấu trúc miếng vá.

auxiliary_patch_flag[a][p] bằng 1 biểu thị rằng miếng vá p của tập bản ảnh a không dành cho bước kết xuất vùng nhìn.

Theo một phương án khác, cú pháp dữ liệu thông tin miếng vá xác định cờ miếng vá hỗ trợ như được trình bày trong bảng sau:

patch_information_data (patchIndex, patch_mode) () {	Bộ mô tả
auxiliary_patch_flag	u(1)
if(patch_mode == P_SKIP)	
skip_patch_data_unit(patchIndex)	
else if(patch_mode == P_MERGE)	
merge_patch_data_unit(patchIndex)	
else if(patch_mode == I_INTRA patch_mode == P_INTRA)	
patch_data_unit(patchIndex)	
else if(patch_mode == P_INTER)	
inter_patch_data_unit(patchIndex)	
else if(patch_mode == I_RAW patch_mode == P_RAW)	
raw_patch_data_unit(patchIndex)	
}	

Ở phía giải mã, **auxiliary_patch_flag** được sử dụng để xác định xem miếng vá có bao gồm thông tin để kết xuất và/hoặc cho mô-đun khác hay không.

Hình 8 minh họa phương pháp 80 để mã hóa nội dung video ba chiều bao gồm thông tin hỗ trợ, theo phương án không giới hạn của sáng chế. Trong bước 81, các miếng vá được sử dụng để tạo kết xuất 3DoF+, chẳng hạn như bằng cách cắt bớt thông tin dư thừa từ nội dung đa góc nhìn cộng với độ sâu (Multiview-plus-depth) mà nhóm camera thứ nhất chụp được. Trong bước 82, các miếng vá hỗ trợ được tạo ra từ khung hình do camera thu được khi chụp các phần của cảnh mà không được kết xuất. Bước 81 và 82 có thể được thực hiện song song hoặc nối tiếp nhau. Khung hình được sử dụng để tạo miếng vá hỗ trợ được chụp bởi nhóm camera thứ hai, ví dụ, nằm ở phía sau và bên cạnh cảnh 3D. Miếng vá hỗ trợ được tạo ra, ví dụ, bằng cách cắt bớt thông tin thừa có trong khung hình được camera của nhóm thứ nhất và thứ hai chụp lại. Theo một phương án khác, miếng vá hỗ trợ được tạo ra, ví dụ, bằng cách cắt bớt thông tin thừa có trong khung hình chỉ được chụp bằng camera của nhóm thứ hai. Theo phương án này, có thể tồn tại phần dư thừa giữa các miếng vá 3DoF+ và miếng vá hỗ trợ. Trong bước 83, tập bản ảnh được tạo ra bằng cách nén 3DoF+ và các miếng vá hỗ trợ trong cùng một hình ảnh. Theo một phương án, bố cục nén đối với các thành phần độ sâu và màu sắc của tập bản ảnh là khác nhau. Siêu dữ liệu mô tả thông số tập bản ảnh và thông số miếng vá được tạo theo cú pháp như mô tả trong bảng trên. Siêu dữ liệu bao gồm thông tin biểu thị, đối với mỗi miếng vá, miếng vá đó là miếng vá 3DoF+ được kết xuất, hay miếng vá hỗ trợ được sử dụng để xử lý trước và/hoặc xử lý sau. Trong bước 84, tập bản ảnh đã tạo và siêu dữ liệu liên quan được mã hóa trong luồng dữ liệu.

Hình 9 minh họa phương pháp 90 để giải mã nội dung video ba chiều bao gồm thông tin hỗ trợ, theo phương án không giới hạn của sáng chế. Trong bước 91, luồng dữ liệu đại diện cho nội dung ba chiều thu được từ luồng. Luồng dữ liệu được giải mã để truy xuất tập bản ảnh và siêu dữ liệu liên quan. Tập bản ảnh là hình ảnh nén ít nhất miếng vá theo bố cục nén. Miếng vá là hình ảnh bao gồm độ sâu và/hoặc thông tin màu đại diện cho một phần của cảnh 3D. Siêu dữ liệu bao gồm thông tin để bỏ chiếu miếng vá và truy xuất cảnh 3D. Trong bước 92, các miếng vá được giải nén khỏi tập bản ảnh và bản chất được ghi nhận cho mỗi miếng vá theo thông tin có trong siêu dữ liệu. Miếng vá có thể là miếng vá 3DoF+, được sử dụng để kết xuất hình ảnh vùng nhìn ở bước 93

hoặc miếng vá bổ trợ, được sử dụng cho hoạt động xử lý trước hoặc xử lý sau ở bước 94. Các bước 93 và 94 có thể được thực hiện song song hoặc nối tiếp nhau.

Các phương án thực hiện được mô tả trong bản mô tả này có thể được thực hiện, ví dụ phương pháp hoặc quy trình, thiết bị, sản phẩm chương trình máy tính, luồng dữ liệu, hoặc tín hiệu. Ngay cả khi chỉ được thảo luận trong ngữ cảnh về hình thức của phương án thực hiện duy nhất (ví dụ, được thảo luận chỉ là phương pháp hoặc thiết bị), thì phương án thực hiện có các tính năng được thảo luận này cũng có thể được thực hiện dưới các hình thức khác (ví dụ, chương trình). Thiết bị có thể được thực hiện trong, ví dụ như phần cứng, phần mềm và phần mềm điều khiển thích hợp. Các phương pháp có thể được thực hiện trong, ví dụ, thiết bị, chẳng hạn như bộ xử lý, dùng để chỉ các thiết bị xử lý nói chung, bao gồm, ví dụ, máy tính, bộ vi xử lý, mạch tích hợp hoặc thiết bị logic khả lập. Bộ xử lý cũng bao gồm các thiết bị truyền thông, chẳng hạn như điện thoại thông minh, máy tính bảng, máy tính, điện thoại di động, thiết bị kỹ thuật số hỗ trợ cá nhân/di động ("PDA") và các thiết bị khác hỗ trợ truyền thông tin giữa những người dùng cuối.

Việc triển khai các quy trình và tính năng khác nhau được mô tả trong tài liệu này có thể được thể hiện trong nhiều thiết bị hoặc ứng dụng khác nhau, đặc biệt, ví dụ, thiết bị hoặc ứng dụng liên quan đến mã hóa dữ liệu, giải mã dữ liệu, tạo khung hình, xử lý kết cấu và xử lý hình ảnh khác và các ứng dụng liên quan thông tin kết cấu và/hoặc thông tin độ sâu. Ví dụ về thiết bị này bao gồm bộ mã hóa, bộ giải mã, đầu ra xử lý của bộ xử lý sau từ bộ giải mã, bộ xử lý trước cung cấp đầu vào cho bộ mã hóa, bộ mã hóa video, bộ giải mã video, codec video, máy chủ web, đầu thu giải mã tín hiệu cho tivi, máy tính xách tay, máy tính cá nhân, điện thoại di động, PDA và các thiết bị liên lạc khác. Cần phải rõ ràng rằng thiết bị có thể di động và thậm chí được lắp đặt trên phương tiện di động.

Ngoài ra, các phương pháp có thể được thực hiện bằng các lệnh do bộ xử lý thực hiện và các lệnh đó (và/hoặc các giá trị dữ liệu do phương án triển khai tạo ra) có thể được lưu trữ trên phương tiện mà bộ xử lý có thể đọc được, chẳng hạn như mạch tích hợp, sóng mang phần mềm hoặc thiết bị lưu trữ khác, chẳng hạn như đĩa cứng, đĩa compact ("CD"), đĩa quang (như là DVD, thường được gọi là đĩa đa năng kỹ thuật số hoặc đĩa video kỹ thuật số), bộ nhớ truy cập ngẫu nhiên ("RAM") hoặc bộ nhớ chỉ đọc ("ROM"). Các lệnh có thể tạo thành một chương trình ứng dụng được đưa vào một

phương tiện mà bộ xử lý có thể đọc được. Ví dụ, lệnh có thể là phần cứng, phần mềm điều khiển, phần mềm hoặc giải pháp kết hợp. Ví dụ, lệnh có thể được tìm thấy trong một hệ điều hành, một ứng dụng riêng biệt hoặc kết hợp cả hai. Do đó, đặc điểm của bộ xử lý có thể, ví dụ như cả thiết bị được định cấu hình để thực hiện một quy trình và một thiết bị gồm một phương tiện có thể đọc được bằng bộ xử lý (chẳng hạn như một thiết bị lưu trữ) có các lệnh thực thi một quy trình. Hơn nữa, phương tiện có thể đọc được bằng bộ xử lý có thể lưu trữ, ngoài hoặc thay cho các lệnh, các giá trị dữ liệu do một phương án triển khai tạo ra.

Như một người có hiểu biết trong lĩnh vực kỹ thuật này sẽ hiểu rõ, các phương án thực hiện có thể tạo ra nhiều loại tín hiệu được định dạng để mang thông tin có thể, ví dụ như được lưu trữ hoặc phát. Thông tin có thể bao gồm, ví dụ như lệnh để thực hiện phương pháp hoặc dữ liệu được tạo ra bởi một trong các phương án thực hiện được mô tả. Ví dụ, tín hiệu có thể được định dạng để mang dưới dạng dữ liệu các quy tắc để viết hoặc đọc cú pháp của phương án được mô tả, hoặc để mang dưới dạng dữ liệu các giá trị cú pháp thực tế được viết bởi phương án được mô tả. Tín hiệu như vậy có thể được định dạng, ví dụ như thành sóng điện từ (ví dụ, sử dụng phần tần số vô tuyến của quang phổ) hoặc thành một tín hiệu dải tần cơ sở. Ví dụ, định dạng có thể bao gồm mã hóa luồng dữ liệu và điều biến sóng mang với luồng dữ liệu được mã hóa. Thông tin mà tín hiệu mang theo có thể là, ví dụ như thông tin kỹ thuật số hoặc tương tự. Như đã biết, tín hiệu có thể được truyền qua nhiều dạng kết nối có dây hoặc không dây. Tín hiệu có thể được lưu trữ trên phương tiện mà bộ xử lý có thể đọc được.

Một số phương án triển khai đã được mô tả. Tuy nhiên, cần hiểu rằng có thể thực hiện nhiều sửa đổi khác nhau. Ví dụ, các phần tử của các phương án triển khai khác nhau có thể được kết hợp, bổ sung, sửa đổi hoặc loại bỏ để tạo ra các phương án triển khai khác. Ngoài ra, người có hiểu biết trung bình trong lĩnh vực sẽ hiểu được rằng các cấu trúc và quy trình khác có thể được thay thế cho những cấu trúc và quy trình được đề xuất và các phương án triển khai sau cùng, về cơ bản, sẽ thực hiện ít nhất cùng (các) chức năng, ít nhất, về cơ bản, theo cùng một cách, để đạt được, về cơ bản, ít nhất là (các) kết quả giống như các phương án triển khai được đề xuất. Theo đó, những phương án triển khai này và các phương án triển khai khác đều được tính đến trong đơn xin cấp bằng sáng chế này.

YÊU CẦU BẢO HỘ

1. Phương pháp mã hóa cảnh 3D trong luồng dữ liệu, phương pháp này bao gồm các bước:

tạo tập hợp các miếng vá thứ nhất từ nội dung đa góc nhìn cộng với độ sâu (MVD) thứ nhất được thu thập cho việc kết xuất cảnh 3D, miếng vá thứ nhất là một phần của một trong các góc nhìn của nội dung MVD thứ nhất;

tạo tập hợp các miếng vá thứ hai từ nội dung MVD thứ hai, trong đó các miếng vá thứ hai không được dự định sử dụng cho việc kết xuất góc nhìn và được thu thập cho việc xử lý trước hoặc xử lý sau của việc kết xuất cảnh 3D, miếng vá thứ hai là một phần của một trong các góc nhìn của nội dung MVD thứ hai;

tạo tập bản ảnh thứ nhất đóng gói các miếng vá thứ nhất và tập bản ảnh thứ hai đóng gói các miếng vá thứ hai, tập bản ảnh là hình ảnh chứa các miếng vá theo bố cục của tập bản ảnh; và

mã hóa các tập bản ảnh thứ nhất và thứ hai trong luồng dữ liệu.

2. Phương pháp theo điểm 1, trong đó nội dung MVD thứ hai được thu thập với độ phân giải thấp hơn độ phân giải của nội dung MVD thứ nhất.

3. Phương pháp theo điểm 1, trong đó miếng vá là một phần của một góc nhìn của nội dung MVD được thu được bằng cách loại bỏ sự dư thừa thông tin giữa các góc nhìn của nội dung MVD.

4. Phương pháp theo điểm 1, trong đó các miếng vá thứ nhất và thứ hai được đóng gói trong tập bản ảnh chung, tập bản ảnh chung được liên kết với siêu dữ liệu biểu thị, đối với mỗi miếng vá, miếng vá đó là miếng vá thứ nhất hay miếng vá thứ hai.

5. Phương pháp theo điểm 1, trong đó tập bản ảnh thứ hai được liên kết với siêu dữ liệu biểu thị rằng các miếng vá thứ hai không được dự định sử dụng cho việc kết xuất góc nhìn và bao gồm danh sách các loại xử lý trước hoặc xử lý sau của việc kết xuất cảnh 3D.

6. Phương pháp theo điểm 5, trong đó loại xử lý trước hoặc xử lý sau của việc kết xuất cảnh 3D thuộc nhóm các loại xử lý trước hoặc xử lý sau của việc kết xuất cảnh 3D bao gồm phát hiện va chạm, tương tác và định sáng lại cảnh.

7. Phương pháp giải mã cảnh 3D từ luồng dữ liệu, phương pháp này bao gồm các bước:

giải mã luồng dữ liệu để truy xuất tập bản ảnh thứ nhất và tập bản ảnh thứ hai, trong đó tập bản ảnh thứ nhất và tập bản ảnh thứ hai là hình ảnh chứa các miếng vá thứ nhất hoặc thứ hai tương ứng theo bố cục của tập bản ảnh, miếng vá thứ nhất là một phần của một góc nhìn của nội dung MVD thứ nhất được thu thập để kết xuất cảnh 3D, miếng vá thứ hai là một phần của một góc nhìn của nội dung MVD thứ hai, không được dự định sử dụng cho việc kết xuất góc nhìn và được thu thập cho việc xử lý trước hoặc xử lý sau của việc kết xuất cảnh 3D;

sử dụng các miếng vá thứ hai cho việc xử lý trước và/hoặc xử lý sau của việc kết xuất cảnh 3D; và

kết xuất hình ảnh vùng nhìn từ điểm quan sát trong cảnh 3D bằng cách sử dụng các miếng vá thứ nhất.

8. Phương pháp theo điểm 7, trong đó nội dung MVD thứ hai có độ phân giải thấp hơn độ phân giải của nội dung MVD thứ nhất.

9. Phương pháp theo điểm 7, trong đó tập bản ảnh thứ nhất và tập bản ảnh thứ hai là tập bản ảnh duy nhất, tập bản ảnh duy nhất được liên kết với siêu dữ liệu biểu thị, đối với mỗi miếng vá, miếng vá đó là miếng vá thứ nhất hay miếng vá thứ hai.

10. Phương pháp theo điểm 7, trong đó tập bản ảnh thứ hai được liên kết với siêu dữ liệu biểu thị rằng các miếng vá thứ hai không được dự định sử dụng cho việc kết xuất góc nhìn và bao gồm danh sách các loại xử lý trước hoặc xử lý sau của việc kết xuất cảnh 3D.

11. Phương pháp theo điểm 10, trong đó loại xử lý trước hoặc xử lý sau của việc kết xuất cảnh 3D thuộc nhóm các loại xử lý trước hoặc xử lý sau của việc kết xuất cảnh 3D bao gồm phát hiện va chạm, tương tác và định sáng lại cảnh.

12. Thiết bị mã hóa cảnh 3D trong luồng dữ liệu, thiết bị bao gồm bộ nhớ liên kết với bộ xử lý được tạo cấu hình để:

tạo tập hợp các miếng vá thứ nhất từ nội dung đa góc nhìn cộng với độ sâu (MVD) thứ nhất được thu thập cho việc kết xuất cảnh 3D, miếng vá thứ nhất là một phần của một trong các góc nhìn của nội dung MVD thứ nhất;

tạo tập hợp các miếng vá thứ hai từ nội dung MVD thứ hai, trong đó các miếng vá thứ hai không được dự định sử dụng cho việc kết xuất góc nhìn và được thu thập cho

việc xử lý trước hoặc xử lý sau của việc kết xuất cảnh 3D, miếng vá thứ hai là một phần của một trong các góc nhìn của nội dung MVD thứ hai;

tạo tập bản ảnh thứ nhất đóng gói các miếng vá thứ nhất và tập bản ảnh thứ hai đóng gói các miếng vá thứ hai, tập bản ảnh là hình ảnh chứa các miếng vá theo bố cục của tập bản ảnh; và

mã hóa các tập bản ảnh thứ nhất và thứ hai trong luồng dữ liệu.

13. Thiết bị theo điểm 12, trong đó nội dung MVD thứ hai được thu thập với độ phân giải thấp hơn độ phân giải của nội dung MVD thứ nhất.

14. Thiết bị theo điểm 12, trong đó miếng vá là một phần của một góc nhìn của nội dung MVD được thu được bằng cách loại bỏ sự dư thừa thông tin giữa các góc nhìn của nội dung MVD.

15. Thiết bị theo điểm 12, trong đó các miếng vá thứ nhất và thứ hai được đóng gói trong tập bản ảnh duy nhất, tập bản ảnh duy nhất được liên kết với siêu dữ liệu biểu thị, đối với mỗi miếng vá, miếng vá đó là miếng vá thứ nhất hay miếng vá thứ hai.

16. Thiết bị theo điểm 12, trong đó tập bản ảnh thứ hai được liên kết với siêu dữ liệu biểu thị rằng các miếng vá thứ hai không được dự định sử dụng cho việc kết xuất góc nhìn và bao gồm danh sách các loại xử lý trước hoặc xử lý sau của việc kết xuất cảnh 3D.

17. Thiết bị theo điểm 16, trong đó loại xử lý trước hoặc xử lý sau của việc kết xuất cảnh 3D thuộc nhóm các loại xử lý trước hoặc xử lý sau của việc kết xuất cảnh 3D bao gồm phát hiện va chạm, tương tác và định sáng lại cảnh.

18. Thiết bị giải mã cảnh 3D từ luồng dữ liệu, thiết bị bao gồm bộ xử lý được tạo cấu hình để:

giải mã luồng dữ liệu để truy xuất tập bản ảnh thứ nhất và tập bản ảnh thứ hai, trong đó tập bản ảnh thứ nhất và tập bản ảnh thứ hai là hình ảnh chứa các miếng vá thứ nhất hoặc thứ hai tương ứng theo bố cục của tập bản ảnh, miếng vá thứ nhất là một phần của một góc nhìn của nội dung MVD thứ nhất được thu thập để kết xuất cảnh 3D, miếng vá thứ hai là một phần của một góc nhìn của nội dung MVD thứ hai, không được dự định sử dụng cho việc kết xuất góc nhìn và được thu thập cho việc xử lý trước hoặc xử lý sau của việc kết xuất cảnh 3D;

sử dụng các miếng vá thứ hai cho việc xử lý trước và/hoặc xử lý sau của việc kết xuất cảnh 3D; và

kết xuất hình ảnh vùng nhìn từ điểm quan sát trong cảnh 3D bằng cách sử dụng các miếng vá thứ nhất.

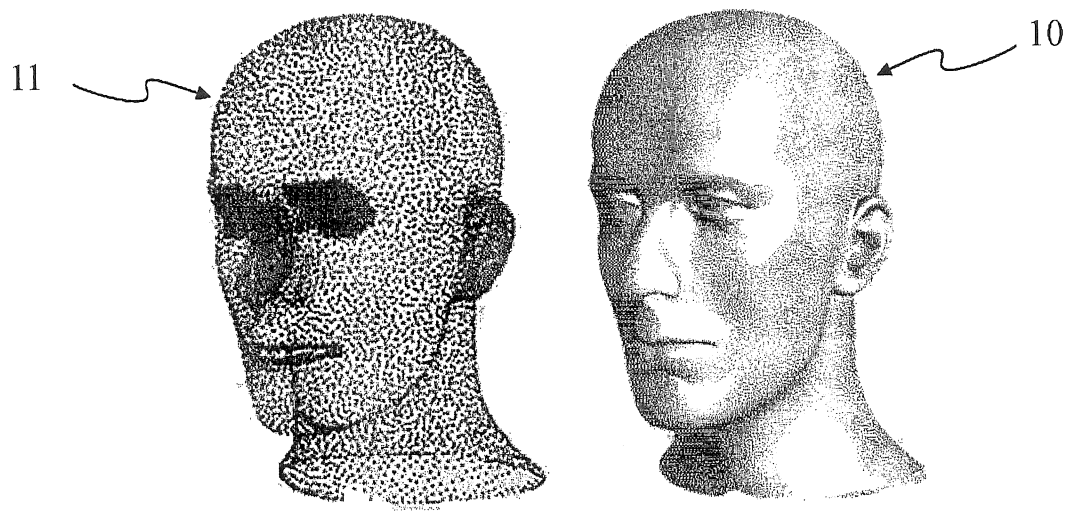
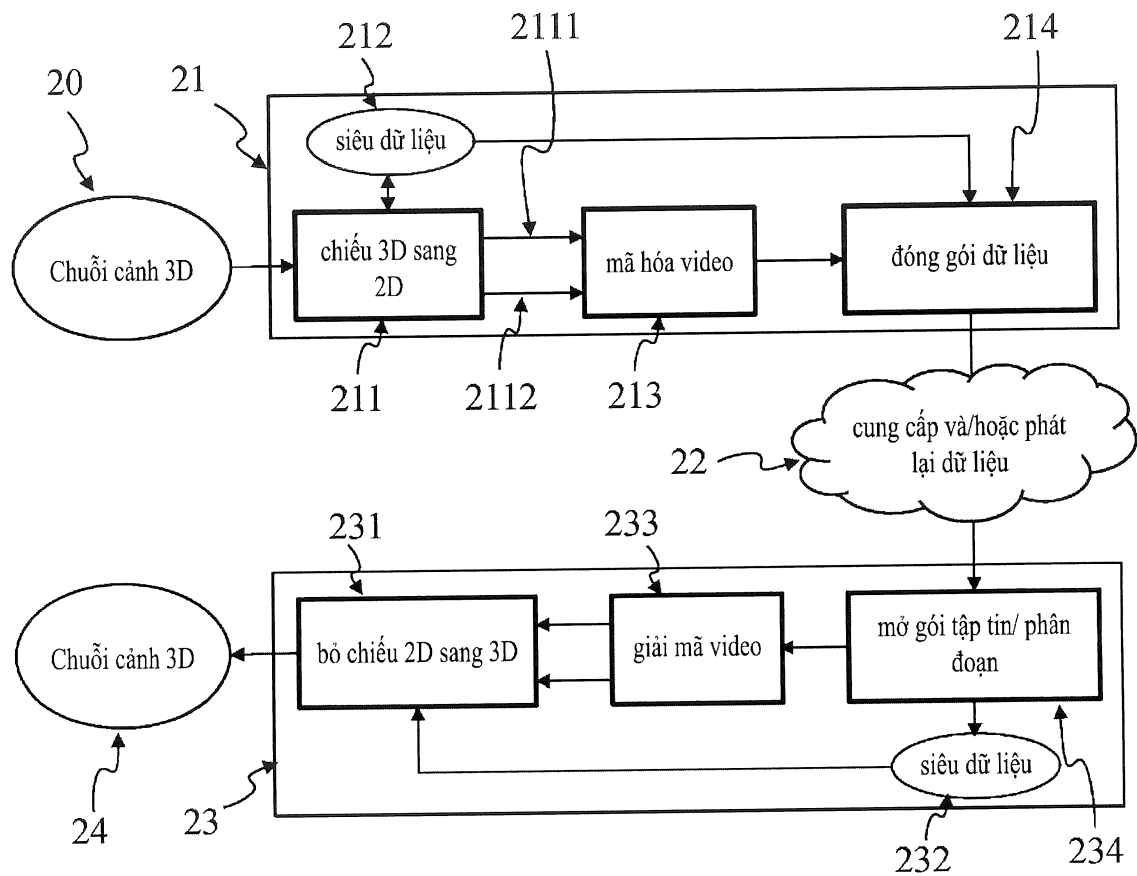
19. Thiết bị theo điểm 18, trong đó nội dung MVD thứ hai có độ phân giải thấp hơn độ phân giải của nội dung MVD thứ nhất.

20. Thiết bị theo điểm 18, trong đó tập bản ảnh thứ nhất và tập bản ảnh thứ hai là tập bản ảnh duy nhất, tập bản ảnh duy nhất được liên kết với siêu dữ liệu biểu thị, đối với mỗi miếng vá, miếng vá đó là miếng vá thứ nhất hay miếng vá thứ hai.

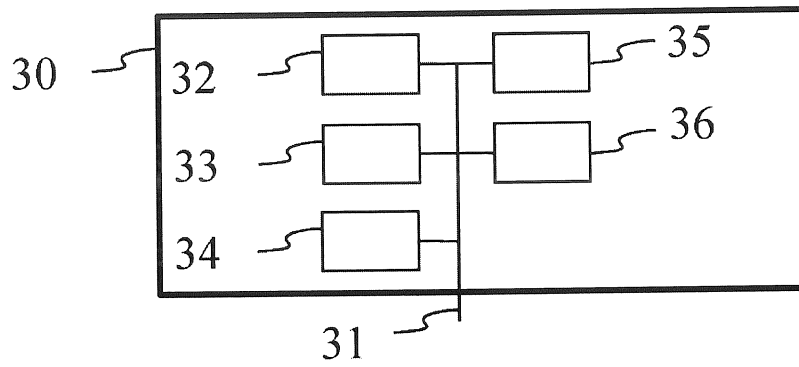
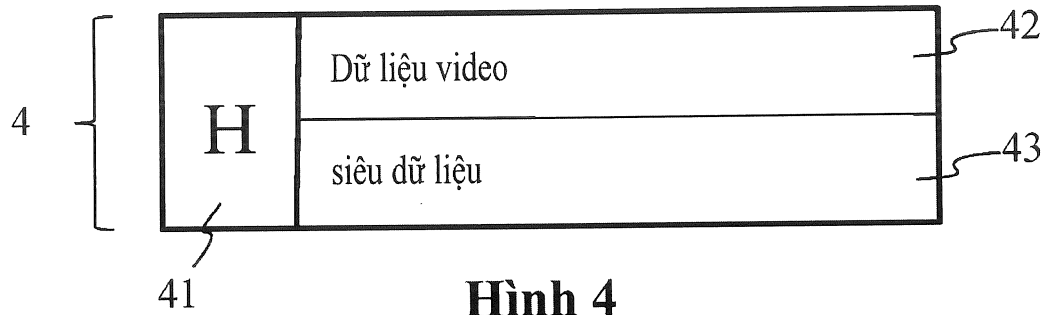
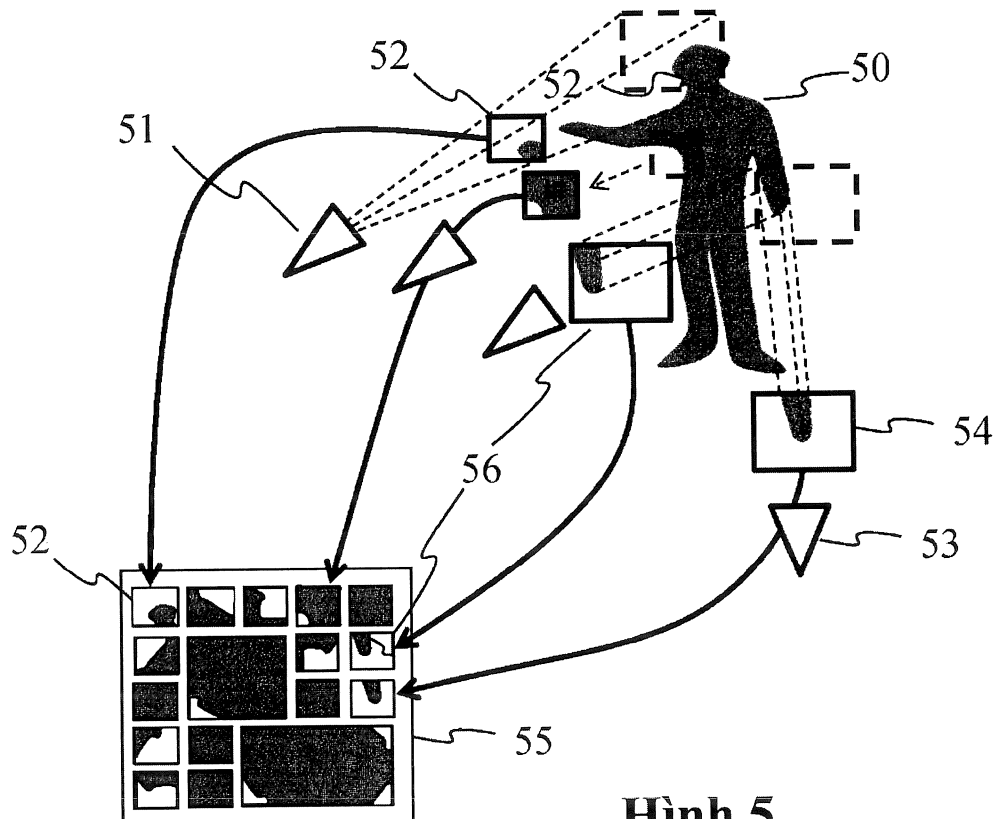
21. Thiết bị theo điểm 18, trong đó tập bản ảnh thứ hai được liên kết với siêu dữ liệu biểu thị rằng các miếng vá thứ hai không được dự định sử dụng cho việc kết xuất góc nhìn và bao gồm danh sách các loại xử lý trước hoặc xử lý sau của việc kết xuất cảnh 3D.

22. Thiết bị theo điểm 21, trong đó loại xử lý trước hoặc xử lý sau của việc kết xuất cảnh 3D thuộc nhóm các loại xử lý trước hoặc xử lý sau của việc kết xuất cảnh 3D bao gồm phát hiện va chạm, tương tác và định sáng lại cảnh.

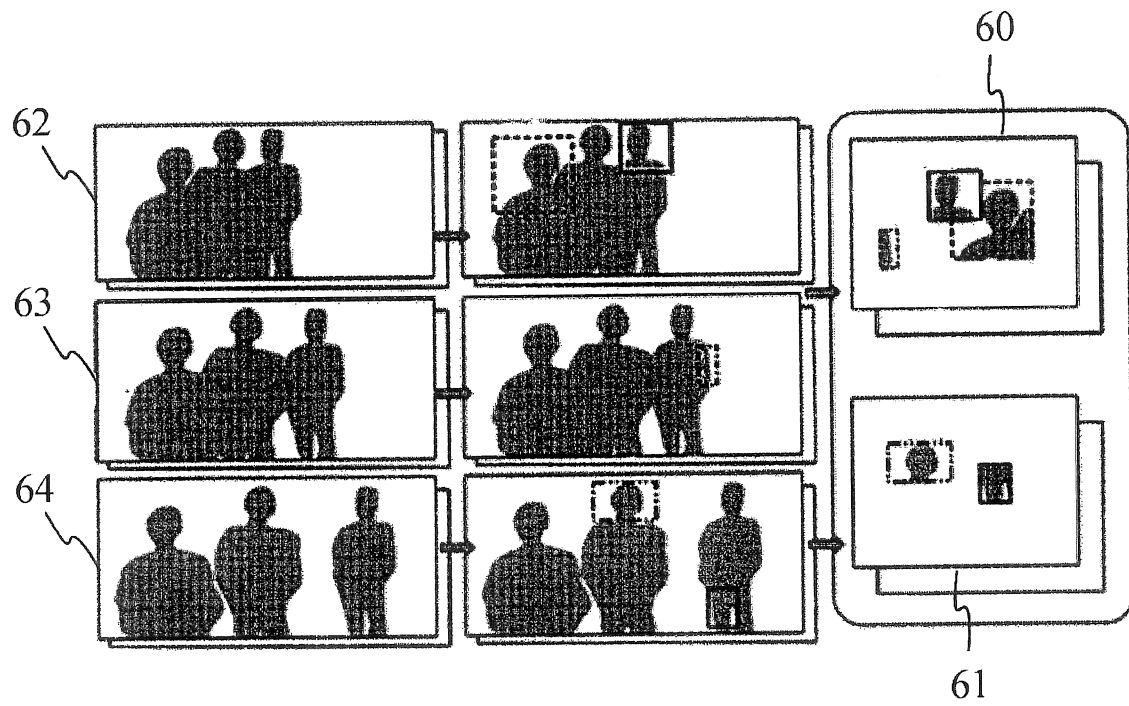
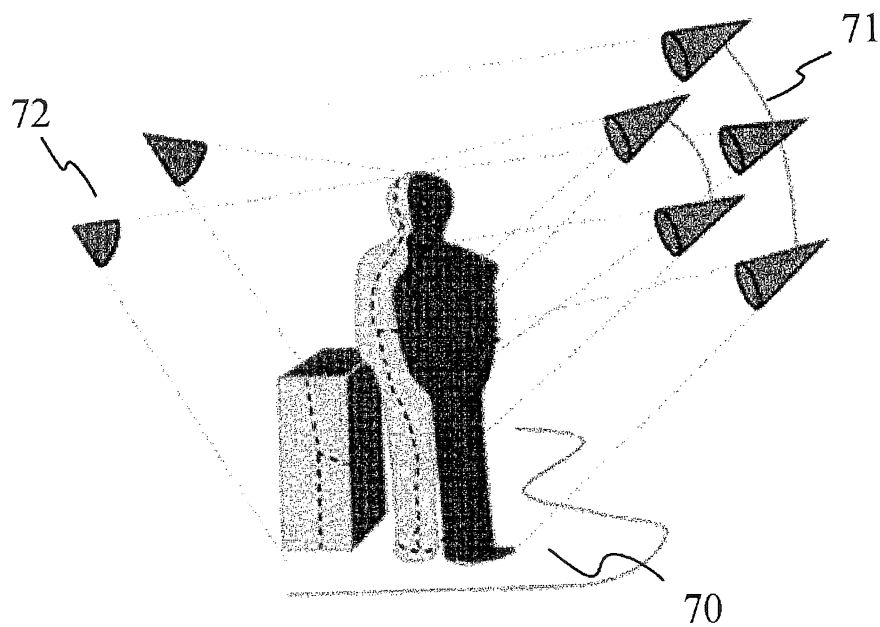
1/4

**Hình 1****Hình 2**

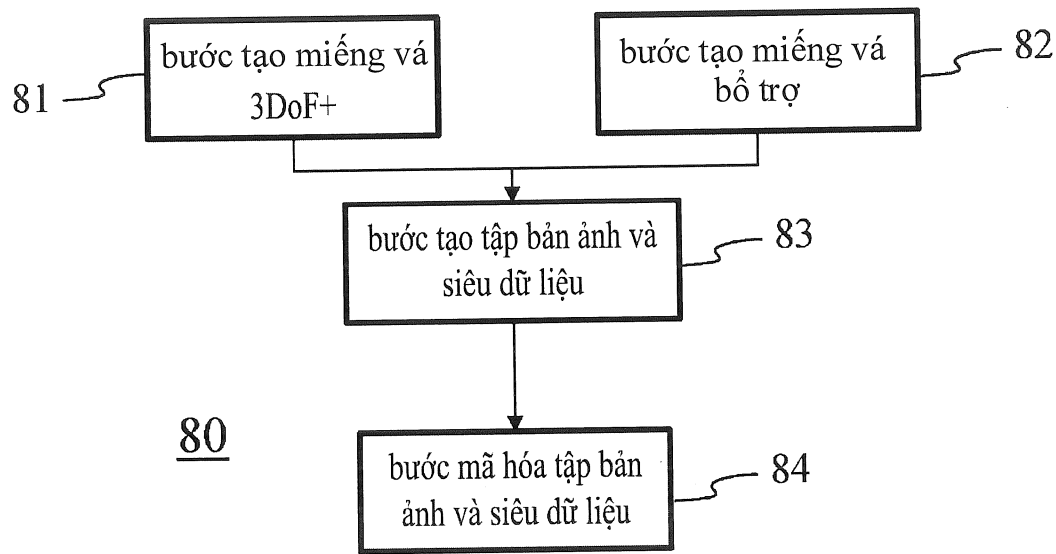
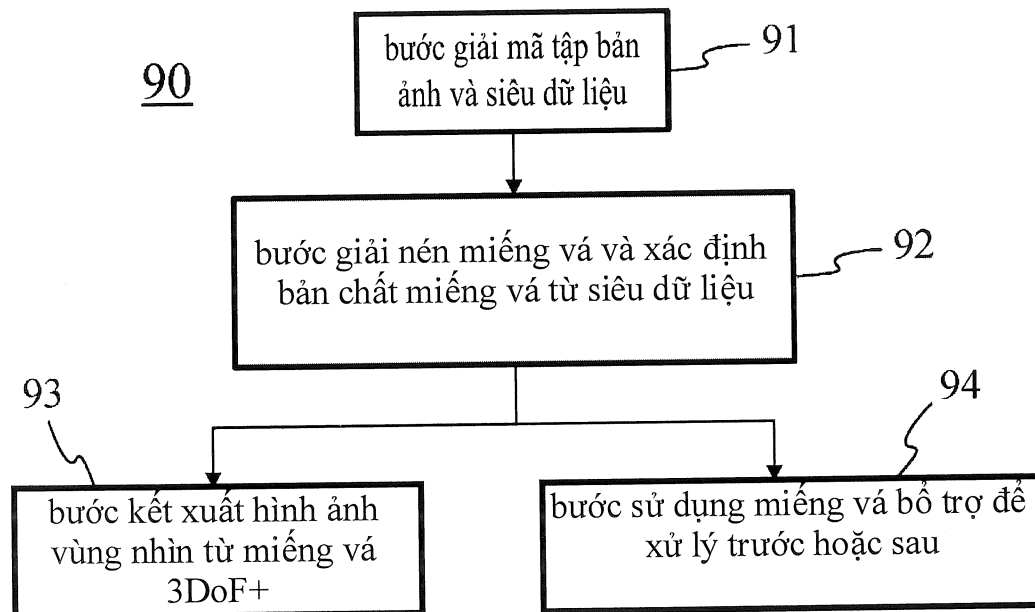
2/4

**Hình 3****Hình 4****Hình 5**

3/4

**Hình 6****Hình 7**

4/4

**Hình 8****Hình 9**