



(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)



1-0050053

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

(21) 1-2022-04069

(22) 17/12/2020

(86) PCT/EP2020/086756 17/12/2020

(87) WO/2021/122983 24/06/2021

(30) 19306717.0 20/12/2019 EP

(45) 25/08/2025 449

(43) 25/08/2022 413A

(73) InterDigital CE Patent Holdings, SAS (FR)

3 rue du Colonel Moll, Paris, 75017 France

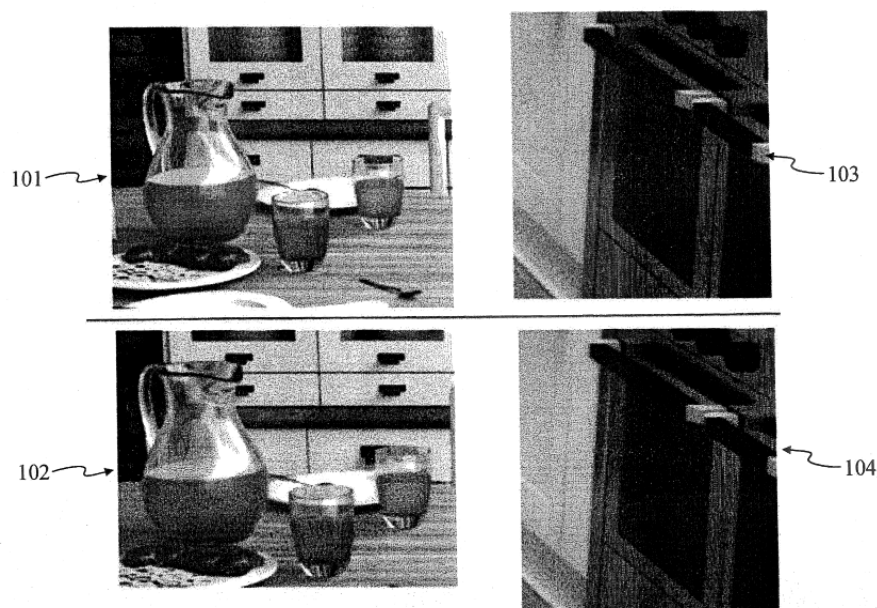
(72) Julien FLEUREAU (FR); Franck THUDOR (FR); Thierry TAPIE (FR); Renaud DORE (FR).

(74) Văn phòng Luật sư Ân Nam (ANNAM IP & LAW)

(54) PHƯƠNG PHÁP VÀ THIẾT BỊ MÃ HÓA CẢNH BA CHIỀU, PHƯƠNG PHÁP VÀ THIẾT BỊ KẾT XUẤT HÌNH ẢNH KHUNG NHÌN BIỂU DIỄN CẢNH BA CHIỀU VÀ PHƯƠNG TIỆN LƯU TRỮ CÓ THỂ ĐỌC ĐƯỢC BẰNG MÁY TÍNH KHÔNG TẠM THỜI

(21) 1-2022-04069

(57) Các phương pháp và thiết bị được đề xuất để mã hóa, truyền và giải mã video ba chiều 3DoF+. Ở giai đoạn mã hóa, một khung hình đầu vào (trong số tất cả các khung hình đầu vào) được chọn để truyền tải hiệu ứng ánh sáng phụ thuộc vào khung nhìn và mã nhận biết của nó được truyền đến bộ giải mã dưới dạng siêu dữ liệu bổ sung. Về phía bộ giải mã, khi các phân mảnh có nguồn gốc từ khung hình được chọn này khả dụng để kết xuất khung nhìn, chúng được ưu tiên sử dụng so với các ứng viên khác bất kể đối với khung hình nào để tổng hợp vị trí.



Hình 10

Lĩnh vực kỹ thuật được đề cập

Sáng chế nhìn chung đề cập đến lĩnh vực cảnh ba chiều (Three Demension - 3D) và nội dung video ba chiều. Sáng chế cũng được hiểu trong bối cảnh mã hóa, định dạng và giải mã dữ liệu đại diện cho kết cấu và hình học của cảnh 3D để hiển thị nội dung ba chiều trên thiết bị của người dùng cuối như thiết bị di động hoặc màn hình hiển thị đội đầu (Head- Mounted Displays - HMD).

Tình trạng kỹ thuật của sáng chế

Phần này nhằm mục đích giới thiệu cho người đọc về các khía cạnh khác nhau của lĩnh vực kỹ thuật tương ứng, có thể liên quan đến các khía cạnh 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 khía cạnh khác nhau của sáng chế. Theo đó, cần hiểu rằng các tuyên bố này phải đọc theo hướng như được trình bày chứ không phải là sự thừa nhận về các giải pháp 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 hiển thị đầy đủ khi người dùng xem nội dung đó trên các thiết bị hiển thị nhập vai như màn hình hiển thị đội 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, chạm vào màn hình, 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 quay đầu quanh điểm nhìn tĩnh. Các thao tác quay 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 hiển thị đội 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 chẳng bao giờ chỉ quay 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, nội dung trường nhìn rộng có thể là cảnh hình ảnh đồ họa máy tính ba chiều (cảnh 3D CGI), điểm đám mây hoặc video nhập vai. Nhiều thuật ngữ có thể được sử dụng để chỉ định các video nhập vai này, ví dụ, Thực tế ảo (Virtual Reality - VR), 360, toàn cảnh, 4π steradian, nhập vai, đa hướng hoặc trường nhìn rộng.

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 quay, người dùng cũng có thể dịch chuyển đầu và thậm chí cả cơ thể của mình trong nội dung đã 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 nhập vai và cảm nhận về độ sâu của phối cảnh, đồng thời cũng ngăn ngừa hiện tượng chóng mặt bằng cách cung cấp phản hồi trực quan nhất quán trong quá trình dịch chuyển đầu. Nội dung được tạo ra bằng các phương tiện cảm biến chuyên dụng cho phép ghi lại đồng thời màu sắc và độ sâu của phối cảnh quan tâm. Việc sử dụng hệ thống máy thu hình 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 ánh xạ các hình ảnh kết cấu (ví dụ, hình ảnh hình cầu được mã hóa theo kỹ thuật ánh xạ chiếu hình ngang/dọc hoặc kỹ thuật ánh xạ chiếu hình chữ nhật), các khung video 6DoF nhúng thông tin từ một số điểm nhìn. Chúng có thể được xem như là chuỗi tạm thời của các đám mây điểm thu được từ sự quay chụp ba chiều. Hai loại video ba chiều có thể được xem xét tùy thuộc vào các điều kiện xem. Loại thứ nhất (tức là 6DoF hoàn chỉnh) cho phép điều hướng tự do hoàn toàn trong nội dung video trong khi loại thứ hai (còn gọi là 3DoF+) giới hạn không gian xem của người dùng trong không gian ba chiều hạn chế được gọi là hộp giới hạn xem, cho phép sự dịch chuyển đầu hạn chế và trải nghiệm thị sai. Loại thứ hai này là sự đánh đổi có giá trị giữa điều hướng tự do và điều kiện xem thụ động của khán giả đang ngồi.

Nội dung 3DoF+ có thể được cung cấp dưới dạng tập hợp các khung nhiều khung hình cộng độ sâu (MVD). Những nội dung đó có thể xuất phát từ các ảnh chụp thực

hoặc có thể được tạo ra từ nội dung CG hiện có bằng cách kết xuất chuyên dụng (có thể là ảnh hiện thực). Khung MVD được quay chụp bởi bộ camera (có cấu trúc như hệ thống camera) đặt ở các điểm nhìn khác nhau và quay chụp cùng một cảnh 3D. Ví dụ, hệ thống có thể tạo cấu trúc bộ camera thành dãy camera hoặc như thiết bị thu nhận 360° với các camera hướng theo các hướng khác nhau. Tuy nhiên, kết xuất 3DoF+ của cảnh 3D được tạo ra dưới dạng khung MVD có thể tạo ra các ảnh giả trực quan khi khung bao gồm nhiều thành phần phản xạ / phản chiếu hoặc trong suốt. Khôi phục độ sáng thực về mặt vật lý từ khung MVD là một thách thức rất phức tạp. Quả thực, tập hợp khung MVD không trực tiếp quay chụp đủ thông tin để khôi phục trường ánh sáng phù hợp. Các thuộc tính vật chất (Chức năng phân bố phản xạ hai chiều (Bidirectional Reflectance Distribution Function) hay còn gọi là BRDF) của các vật thể của phối cảnh cũng như mô tả về các nguồn sáng có thể chỉ được tạm ước tính từ tập hợp các khung MVD và quá trình này rất phức tạp.

Trong trường hợp 3DoF+, người dùng cuối được cho là chỉ trải nghiệm lượng thị sai hạn chế với các chuyển động “nhỏ” trong không gian xem hạn chế. Trong trường hợp đó, theo kinh nghiệm có thể cho thấy người xem thông thường rất ít nhạy cảm với các thay đổi về độ phản xạ / phản chiếu / trong suốt mà họ có thể gặp phải khi dịch chuyển vị trí đầu của mình. Tuy nhiên, người xem rất nhạy cảm với độ nhất quán của những hiệu ứng ánh sáng này qua khung hình tổng hợp. Hiện thiếu giải pháp mã hóa, truyền và giải mã nội dung video ba chiều được cung cấp dưới dạng chuỗi khung MVD với kết xuất những hiệu ứng ánh sáng này một cách nhất quán.

Bản chất kỹ thuật của sáng chế

Phần sau đây trình bày ngắn gọn bản chất kỹ thuật của sáng chế để cung cấp hiểu biết cơ bản về một số khía cạnh theo sáng chế. Phần bản chất kỹ thuật này không phải là tổng quan về sáng chế. Phần này không nhằm mục đích xác định các yếu tố chính hoặc trọng yếu của sáng chế. Phần bản chất kỹ thuật sau đây chỉ trình bày một số khía cạnh của sáng chế ở dạng giản lược dưới dạng phần mở đầu cho phần mô tả chi tiết hơn được đưa ra bên dưới.

Sáng chế đề cập đến phương pháp mã hóa cảnh 3D được tạo ra dưới dạng khung nhiều khung hình cộng độ sâu (MVD) trong luồng dữ liệu. Phương pháp bao gồm các bước:

- chọn một khung hình trong khung MVD làm khung hình tham chiếu trên cơ sở phạm vi bao quát trường nhìn của cảnh 3D từ vị trí trung bình trong không gian xem;

- tạo hình ảnh atlas đóng gói ít nhất một phân mảnh. Phân mảnh là một phần hình ảnh của khung hình trong khung MVD. Hình ảnh atlas bao gồm thông tin đặc trưng của cảnh 3D có thể nhìn thấy từ không gian xem.

- tạo siêu dữ liệu bao gồm các thông số thu nhận của mỗi khung hình trong khung MVD, thông tin liên kết từng phân mảnh với khung hình, và ký hiệu nhận biết khung hình tham chiếu đã chọn; và

- mã hóa atlas và siêu dữ liệu trong luồng dữ liệu.

Sáng chế cũng đề cập đến phương pháp giải mã luồng dữ liệu biểu diễn cảnh 3D có thể nhìn thấy từ không gian xem. Phương pháp bao gồm các bước:

- giải mã hình ảnh atlas và siêu dữ liệu được liên kết từ luồng dữ liệu,

hình ảnh atlas đang đóng gói ít nhất một phân mảnh. Phân mảnh là một phần hình ảnh của khung hình trong khung nhiều khung hình cộng độ sâu (MVD). Hình ảnh atlas bao gồm thông tin đặc trưng của cảnh 3D có thể nhìn thấy từ không gian xem.

Siêu dữ liệu bao gồm các thông số thu nhận của mỗi khung hình trong khung MVD, thông tin liên kết mỗi phân mảnh với khung hình và ký hiệu nhận biết khung hình tham chiếu, phân mảnh được liên kết với khung hình tham chiếu là phân mảnh tham chiếu;

- kết xuất hình ảnh khung nhìn cho điểm nhìn hiện tại trong không gian xem bằng cách:

khi hai phân mảnh góp phần vào hình ảnh khung nhìn, nếu một trong hai phân mảnh là phân mảnh tham chiếu, sử dụng phần màu sắc của phân mảnh tham chiếu cho hình ảnh khung nhìn.

Sáng chế cũng đề cập đến thiết bị bao gồm bộ xử lý thực hiện các bước của phương pháp mã hóa được mô tả ở trên và đề cập đến thiết bị bao gồm bộ xử lý thực hiện các bước của phương pháp giải mã được mô tả ở trên.

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 đây, phần mô tả này dựa trên các hình vẽ kèm theo trong đó:

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

- Hình 2 thể hiện ví dụ không giới hạn về bước mã hóa, truyền và giải mã dữ liệu biểu thị chuỗi phối cảnh 3D, theo phương án không giới hạn của sáng chế; - Hình 3 thể hiện kiến trúc ví dụ của thiết bị mà có thể được tạo cấu hình để thực hiện phương pháp được mô tả theo Hình 11 và 12, theo phương án không giới hạn của sáng chế;

- Hình 4 thể hiện ví dụ về phương án cú pháp của luồng khi dữ liệu được truyền qua giao thức truyền theo gói, theo phương án không giới hạn của sáng chế;

- Hình 5 thể hiện phép chiếu hình cầu từ điểm nhìn trung tâm, theo phương án không giới hạn của sáng chế;

- Hình 6 thể hiện ví dụ về atlas bao gồm thông tin kết cấu của các điểm của phối cảnh 3D, theo phương án không giới hạn của sáng chế;

- Hình 7 thể hiện ví dụ về atlas bao gồm thông tin độ sâu của các điểm của phối cảnh 3D trên Hình 6, theo phương án không giới hạn của sáng chế;

- Hình 8 là sơ đồ minh họa việc quay chụp khung nhiều khung hình cộng độ sâu (MVD) bởi một dãy hai mươi camera phối cảnh, theo phương án không giới hạn của sáng chế;

- Hình 9 thể hiện hai trong số hai mươi khung hình của cảnh 3D được quay chụp bằng dãy máy quay trên Hình 8, theo phương án không giới hạn của sáng chế;

- Hình 10 thể hiện các chi tiết của hình ảnh khung nhìn của phối cảnh trên Hình 9 từ điểm nhìn nhất định trong không gian xem 3DoF+ được liên kết, theo phương án không giới hạn của sáng chế;

- Hình 11 minh họa phương pháp mã hóa nội dung video ba chiều, theo phương án không giới hạn của sáng chế;

- Hình 12 minh họa phương pháp 120 để giải mã nội dung video ba chiều, theo phương án không giới hạn của sáng chế.

Mô tả chi tiết sáng chế

Sáng chế sẽ được mô tả đầy đủ hơn trong phần này tham chiếu đến các hình vẽ kèm theo, trong đó các ví dụ theo sáng chế được thể hiện. 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 bản mô tả 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 thể hiện bằng các ví dụ trong các hình vẽ và sẽ được mô tả chi tiết trong bản mô tả này. Tuy nhiên cần phải hiểu rằng không có dự định giới hạn sáng chế theo các dạng thức cụ thể được bộc lộ, nhưng ngược lại, sáng chế phải bao gồm tất cả sự cải biến, phương thức tương đương và thay thế nằm trong bản chất và phạm vi của sáng chế như được xác định 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 bản mô tả này, các dạng mạo từ số ít được dự định bao hàm cả các dạng số nhiều, trừ khi được chỉ rõ khác theo ngữ cảnh. Cần phải hiểu thêm rằng các thuật ngữ “bao gồm” và/hoặc “bao gồm cả” khi được sử dụng trong bản mô tả này chỉ rõ sự hiện diện của các đặc điểm, các số nguyên, các bước, hoạt động, phần tử và/hoặc các thành phần đã nêu nhưng không loại trừ sự hiện diện hoặc bổ sung của một hoặc nhiều đặc điểm, số nguyên, bước, hoạt động, phần tử, thành phần khác, và/hoặc các 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 “đáp ứng 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 bản mô tả này, thuật ngữ “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 bản mô tả này để mô tả các phần tử khác nhau, nhưng các phần tử 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 truyền thông để minh họa hướng truyền thông chính, nhưng cần hiểu rằng truyền thông 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 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 thể hiện liên tiếp, về cơ bản, có thể được thực thi về cơ bản là 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 bản mô tả này “theo ví dụ” hoặc “trong ví dụ” nghĩa là tính năng, cấu trúc, hoặc đặc điểm cụ thể 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 bản mô 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 số tham chiếu xuất hiện trong các điểm yêu cầu bảo hộ chỉ mang tính chất minh họa và sẽ không có tác dụng giới hạn phạm vi của các điểm 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 thể hiện mô hình ba chiều (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 trình diễn 3D có thể về vật thể của phối cảnh 3D bao gồm các vật thể khác. Mô hình 10 có thể là trình 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 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 điểm ảnh ba chiều (voxel) hoặc đường cong bậc ba (spline). Hình 1 minh họa thực tế rằng đám mây điểm có thể được xác định với sự biểu diễn bề mặt của vật thể 3D và sự biểu diễn bề mặt của vật thể 3D có thể được tạo từ điểm đám mây. Như được sử dụng trong bản mô tả này, việc chiếu các điểm của vật thể 3D (bằng các điểm mở rộng của cảnh 3D) lên hình ảnh là tương đương với việc chiếu sự biểu diễn bất kỳ của vật thể 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 góc khối và 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 điểm nhìn nhất định hoặc một loạt các điểm nhìn. Đám mây điểm có thể thu được theo nhiều cách, ví dụ:

- từ ảnh chụp vật thể thực được chụp bởi hệ thống camera, được bổ sung tùy chọn bởi thiết bị cảm biến độ sâu chủ động;
- từ ảnh 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 thể hiện ví dụ không giới hạn về bước mã hóa, truyền và giải mã dữ liệu biểu thị chuỗi phối cảnh 3D. Định dạng mã hóa có thể, ví dụ và đồng thời, tương thích cho việc giải mã 3DoF, 3DoF+ và 6DoF.

Chuỗi phối cảnh 3D 20 là có thể thu được. Vì chuỗi hình 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 một cảnh 3D hoặc chuỗi cảnh 3D làm đầu vào và tạo ra chuỗi bit biểu thị đầu vào. Chuỗi bit có thể được lưu trữ trong bộ nhớ 22 và/hoặc trên phương tiện dữ liệu điện tử và có thể được truyền qua mạng 22. Chuỗi bit biểu thị chuỗi cảnh 3D có thể được đọc từ bộ nhớ 22 và/hoặc được thu 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 ở định dạng đám mây điểm. Bộ giải mã 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 bất kỳ để ánh xạ các điểm ba chiều đến mặt phẳng hai chiều. Vì hầu hết các phương pháp hiện tại để hiển thị dữ liệu đồ họa là dựa trên phương tiện truyền thông 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, thiết kế và phác thảo. Mạch chiếu 211 cung cấp ít nhất một khung hai

chiều 2111 cho chuỗi cảnh 3D 20. Khung 2111 bao gồm thông tin màu và thông tin độ sâu biểu diễn cảnh 3D được chiếu lên khung 2111. Theo biến thể, thông tin màu và thông tin độ sâu được mã hóa trong hai khung riêng biệt 2111 và 2112. Siêu dữ liệu 212 được sử dụng và được cập nhật bởi mạch chiếu 211. Siêu dữ liệu 212 bao gồm thông tin về hoạt động chiếu (ví dụ các tham số chiếu) và về cách mà thông tin về màu và độ sâu được tổ chức hóa trong các khung 2111 và 2112 như được mô tả liên quan đến các hình từ 5 đến 7.

Mạch mã hóa video 213 mã hóa chuỗi các khung 2111 và 2112 dưới dạng video. Các hình ảnh của các cảnh 3D 2111 và 2112 (hoặc chuỗi các hình ảnh của cảnh 3D) được mã hóa thành 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, mô tả kỹ thuật ISO/CEI 10918-1 UIT-T Recommendation 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 UIT-T H.264 và ISO/CEI MPEG-4 Part 10 (ISO/CEI 14496-10), <http://www.itu.int/rec/T-REC-H.264/en>, HEVC (mô tả kỹ thuật của nó 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>);
- 3D-HEVC (phần mở rộng của HEVC có mô tả 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> annex G and 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 bao gồm 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 tạo ra chuỗi cảnh 3D 24 để được kết xuất và hiển thị bằng thiết bị hiển thị video ba chiều, như thiết bị hiển thị đội đầu (HMD). Bộ giải mã 23 lấy luồng từ nguồn 22. Ví dụ, nguồn 22 thuộc tập hợp bao gồm:

- 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 lưu trữ khối, 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 bao gồm 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 chuỗi các khung. Các khung được giải mã bao gồm thông tin về màu và độ sâu. Theo 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 tạo ra 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.

Hình 3 minh họa cấu trúc ví dụ của thiết bị 30 mà có thể được tạo cấu hình để thực hiện phương pháp được mô tả theo hình 11 và 12 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 Bộ nhớ chỉ đọc) 33;
- RAM (hoặc 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 trong, 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 sụn. 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, ví dụ, 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 định cấu hình để thực hiện phương pháp được mô tả liên quan đến hình 11 và 12, và thuộc về tập hợp bao gồm:

- thiết bị di động;
- thiết bị truyền thông;
- 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 của siêu dữ liệu 212 trên Hình 2, ví dụ, các tọa độ của điểm nhìn trung tâm được sử dụng để chiếu các điểm của cảnh 3D lên các 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 trên các khung. Siêu dữ liệu này có thể được liên kết với mỗi khung của video hoặc với nhóm các khung (còn được gọi là Nhóm hình ảnh (GoP) trong các tiêu chuẩn nén video).

Hình 5 minh họa phương pháp tạo atlas phân mảnh 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 các 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 ánh xạ lên phân mảnh 52 theo phép ánh xạ phối cảnh và hình chiếu của các điểm được camera 53 chụp được ánh xạ lên phân mảnh 54 theo phép ánh xạ trực giao.

Việc tạo nhóm các pixel được chiếu tạo ra phân mảnh 2D, được đóng gói trong Atlas hình chữ nhật 55. Việc sắp xếp các phân mảnh trong Atlas xác định bố cục Atlas. Theo một phương án, hai Atlas 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 phân mảnh 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 phân mảnh 54 và 56. Một phân mảnh cũng có thể là tổng hợp của khung hình trung tâm 4pi steradian (ví dụ như đẳng chữ nhật như trên hình 6

và 7 hoặc ánh xạ hình khối) mà trong hầu hết các trường hợp, sẽ bao phủ chín mươi phần trăm phần nhìn thấy được của cảnh 3D.

Hoạt động đóng gói tạo ra dữ liệu phân mảnh cho mỗi phân mảnh được tạo. Dữ liệu phân mảnh 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 phân mảnh trong Atlas (ví dụ: tọa độ góc trên cùng bên trái, kích thước và chiều rộng tính bằng pixel). Các mục dữ liệu phân mảnh được thêm vào siêu dữ liệu để được đóng gói trong luồng dữ liệu cùng với dữ liệu nén của một hoặc hai Atlas.

Hình 6 minh họa ví dụ về atlas 60 bao gồm thông tin kết cấu (ví dụ dữ liệu RGB hoặc dữ liệu YUV) của các điểm cảnh 3D, theo phương án không giới hạn của sáng chế. Như đã giải thích liên quan đến hình 5, Atlas là phân mảnh đóng gói hình ảnh, phân mảnh là hình ảnh thu được bằng cách chiếu một phần của các điểm của cảnh 3D. Theo ví dụ trên Hình 6, atlas 60 bao gồm phần thứ nhất 61 bao gồm thông tin kết cấu của các điểm của cảnh 3D có thể nhìn thấy từ điểm nhìn và một hoặc nhiều phần thứ hai 62. Ví dụ, thông tin về kết cấu của phần thứ nhất 61 có thể thu được theo công nghệ hình ảnh 2 chiều được kết nối đẳng hình chữ nhật, công nghệ hình ảnh 2 chiều được kết nối đẳng hình chữ nhật là ví dụ về công nghệ hình ảnh 2 chiều được kết nối đẳng hình cầu. Theo ví dụ trong hình 6, phần thứ hai 62 được sắp xếp ở các đường viền bên trái và bên phải của phần thứ nhất 61 nhưng phần thứ hai có thể được sắp xếp khác nhau. Các phần thứ hai 62 bao gồm thông tin kết cấu của các phần của cảnh 3D bổ sung cho phần có thể nhìn thấy từ điểm nhìn. Các phần thứ hai có thể thu được bằng cách xóa khỏi cảnh 3D các điểm có thể nhìn thấy từ điểm nhìn thứ nhất (kết cấu được lưu trữ trong phần thứ nhất) và bằng cách chiếu các điểm còn lại theo cùng một điểm nhìn. Quá trình thứ hai có thể được lặp đi lặp lại để thu được các phần bị ẩn của cảnh 3D mỗi lần. Theo một biến thể, các phần thứ hai có thể thu được bằng cách xóa khỏi cảnh 3D những điểm có thể nhìn thấy từ điểm nhìn, ví dụ như điểm nhìn trung tâm, (kết cấu được lưu trữ trong phần thứ nhất) và bằng cách chiếu các điểm còn lại theo điểm nhìn khác với điểm nhìn thứ nhất, ví dụ từ một hoặc nhiều điểm nhìn thứ hai của không gian nhìn tập trung vào điểm nhìn trung tâm (ví dụ: không gian nhìn của kết xuất 3Dof).

Phần thứ nhất 61 có thể được xem như phân mảnh kết cấu lớn thứ nhất (tương ứng với phần thứ nhất của cảnh 3D) và các phần thứ hai 62 bao gồm các phân mảnh kết

cầu nhỏ hơn (tương ứng với các phần thứ hai của cảnh 3D bổ sung cho phần thứ nhất). Atlas như vậy có lợi thế là tương thích đồng thời với sự kết xuất 3DoF (khi chỉ kết xuất phần thứ nhất 61) và với sự kết xuất 3DoF + / 6DoF.

Hình 7 minh họa ví dụ về atlas 70 bao gồm thông tin độ sâu của của các điểm cảnh 3D theo hình 6, theo phương án không giới hạn của sáng chế. Atlas 70 có thể được xem là hình ảnh chiều sâu tương ứng với hình ảnh kết cấu 60 của Hình 6.

Atlas 70 bao gồm phần thứ nhất 71 bao gồm thông tin độ sâu của các điểm của cảnh 3D có thể nhìn thấy từ điểm nhìn trung tâm và một hoặc nhiều phần thứ hai 72. Atlas 70 có thể thu được theo cách tương tự như atlas 60 nhưng chứa thông tin về độ sâu được liên kết với các điểm của cảnh 3D thay vì thông tin về kết cấu. Đối với sự kết xuất 3DoF của cảnh 3D, chỉ một điểm nhìn, thường là điểm nhìn trung tâm, được xem xét. Người dùng có thể xoay đầu theo ba bậc tự do xung quanh điểm nhìn thứ nhất để xem các phần khác nhau của cảnh 3D, nhưng người dùng không thể di chuyển điểm nhìn độc nhất này. Các điểm của cảnh được mã hóa là các điểm có thể nhìn thấy từ điểm nhìn độc nhất này và chỉ có thông tin kết cấu là cần thiết để được mã hóa/giải mã cho sự kết xuất 3DoF. Không cần mã hóa các điểm của cảnh không thể nhìn thấy từ điểm nhìn duy nhất này để kết xuất 3DoF vì người dùng không thể truy cập chúng.

Đối với kết xuất 6DoF, người dùng có thể di chuyển điểm nhìn ở khắp mọi nơi trong cảnh. Trong trường hợp này, cần phải mã hóa mọi điểm (độ sâu và kết cấu) của cảnh trong chuỗi bit vì mọi điểm đều được tiếp cận bởi người dùng có thể di chuyển điểm nhìn của mình. Ở giai đoạn mã hóa, không có phương tiện nào để biết tiên nghiệm, từ điểm nhìn nào người dùng sẽ quan sát cảnh 3D.

Đối với sự kết xuất 3DoF+, người dùng có thể di chuyển điểm nhìn trong một không gian giới hạn xung quanh điểm nhìn trung tâm. Điều này cho phép trải nghiệm thị sai. Dữ liệu biểu diễn phần của phối cảnh có thể nhìn thấy từ điểm bất kỳ trong không gian nhìn sẽ được mã hóa vào luồng, bao gồm dữ liệu biểu diễn cảnh 3D có thể nhìn thấy theo điểm nhìn trung tâm (tức là các phần thứ nhất 61 và 71). Ví dụ, kích thước và hình dạng của không gian nhìn có thể được quyết định và xác định ở bước mã hóa và được mã hóa trong chuỗi bit. Bộ giải mã có thể thu được thông tin này từ chuỗi bit và trình kết xuất giới hạn không gian nhìn trong không gian được xác định bởi thông tin thu được. Theo ví dụ khác, trình kết xuất xác định không gian nhìn theo các giới hạn phần cứng, ví dụ liên quan đến khả năng của (các) cảm biến phát hiện chuyển động của

người dùng. Trong trường hợp này, nếu ở giai đoạn mã hóa, một điểm có thể nhìn thấy từ một điểm trong không gian nhìn của trình kết xuất chưa được mã hóa trong chuỗi bit, thì điểm này sẽ không được kết xuất. Theo ví dụ khác, dữ liệu (ví dụ, kết cấu và/hoặc hình học) biểu diễn mọi điểm của cảnh 3D được mã hóa trong luồng mà không tính đến không gian nhìn kết xuất. Để tối ưu hóa kích thước của luồng, chỉ một tập hợp con các điểm của cảnh có thể được mã hóa, như là tập hợp con các điểm có thể được nhìn thấy theo không gian nhìn kết xuất.

Hình 8 là sơ đồ minh họa ảnh chụp khung nhiều khung hình cộng độ sâu (MVD) bởi một dãy hai mươi camera phối cảnh. Trong ví dụ ở hình 8, hai mươi camera (thực hoặc ảo) 811 đến 845 được tổ chức như một dãy và nhắm đến cùng một hướng xem. Điểm nhìn (tức là tâm chiếu của hình ảnh do camera chụp) được dịch chuyển theo chiều ngang và/hoặc theo chiều dọc so với nhau. Hệ thống camera đó cung cấp hai mươi hình ảnh của cùng một cảnh, mỗi khung hình bao gồm thông tin về màu sắc và độ sâu tiềm năng. Khi các khung hình không bao gồm thành phần độ sâu, thông tin độ sâu có thể được truy xuất bằng cách xử lý các khung hình theo các thông số của camera.

Khi nội dung 3D được tạo ra dưới dạng khung MVD, thông tin ba chiều cũng được truyền tải dưới dạng kết hợp các phân mảnh màu sắc và độ sâu được lưu trữ trong các atlas màu sắc và độ sâu tương ứng, như trong hình 5 đến 7. Nội dung video được mã hóa bằng cách sử dụng codec thông thường (ví dụ: HEVC). Mỗi sự kết hợp các phân mảnh màu và độ sâu biểu thị một phần nhỏ của các khung hình đầu vào của MVD và tập hợp tất cả các phân mảnh được thiết kế ở giai đoạn mã hóa để “bao phủ” toàn bộ cảnh trong khi duy trì tình trạng càng ít dư thừa càng tốt. Ở giai đoạn giải mã, các atlas được giải mã video đầu tiên và các phân mảnh được kết xuất trong quy trình tổng hợp khung hình để khôi phục khung nhìn được liên kết với vị trí nhìn mong muốn. Cách mà các phân mảnh được tạo ra vừa đủ, không dư thừa và bổ sung cho nhau là yếu tố chính. Một cách cần làm là cắt xén các khung hình nguồn đầu vào (hoặc khung hình tổng hợp khác bất kỳ thu được từ các khung hình nguồn đầu vào chẳng hạn như khung hình trung tâm trên hình 6 và 7) để loại bỏ mọi thông tin dư thừa. Để thực hiện như vậy, mỗi khung hình đầu vào (màu sắc + độ sâu), ban đầu hoặc tổng hợp, lần lượt được cắt xén lặp đi lặp lại liên tiếp. Chính xác hơn, tập hợp các khung hình “cơ bản” chưa được cắt xén được chọn đầu tiên trong số các khung hình nguồn và được phát hoàn toàn. Tập hợp các khung hình “bổ sung” còn lại sau đó được xử lý lặp đi lặp lại để loại bỏ thông tin dư

thừa (về độ tương đồng màu sắc và độ sâu) với “các khung hình cơ bản” và với khung hình “bổ sung” đã được cắt xén khác bất kỳ. Tuy nhiên, một số tình trạng gổ chồng giữa các vùng được cắt giảm có thể được cố ý bảo toàn để tránh bất kỳ ảnh giả ghép nối nào.

Hình 9 thể hiện hai trong số hai mươi khung hình của cảnh 3D được chụp bởi dây camera trên hình 8. Để làm rõ hơn, chỉ có khung hình 811 và khung hình 835 được thể hiện. Hình ảnh 811 là khung hình của cảnh từ phía trên và lệch sang trái nhiều hơn so với hình ảnh 835. Cảnh 3D bao gồm các vật thể trong suốt 92 và vật thể phản chiếu 91 (cửa lò phản chiếu con nhện khổng lồ trên sàn). Các khung hình 811 và 835 bao gồm thông tin tương ứng với các điểm giống nhau của cảnh 3D. Tuy nhiên, do việc chiếu sáng của cảnh và các vị trí thu nhận khác nhau, thông tin màu sắc liên quan đến những điểm này có thể khác nhau giữa các khung hình. Các khung hình 811 cũng bao gồm thông tin về các điểm của cảnh 3D mà không thể nhìn thấy từ điểm nhìn của khung hình 835 và ngược lại.

Như đã mô tả ở trên, ít nhất một atlas được tạo để mã hóa cảnh 3D từ khung MVD bằng cách cắt giảm thông tin dư thừa và bảo toàn một số tình trạng gổ chồng giữa các vùng được cắt giảm của không gian 3D để giúp ghép nối khi kết xuất. Atlas được cho là đủ để tái tạo / tổng hợp bất kỳ hình ảnh khung nhìn nào từ bất kỳ điểm nhìn nào trong không gian xem 3DoF+ mà người dùng có thể di chuyển trong đó. Để làm như vậy, quy trình tổng hợp được thực hiện để ghép nối tất cả các phân mảnh từ atlas để khôi phục hình ảnh khung nhìn mong muốn. Tuy nhiên, bước ghép nối này có thể phụ thuộc vào ảnh giả mạnh khi cảnh được biểu diễn trong atlas bao gồm các thành phần phản xạ / phản chiếu hoặc trong suốt như được minh họa trong hình 9. Những hiệu ứng ánh sáng này phụ thuộc vào vị trí quan sát, do đó, màu sắc cảm nhận của các phần không gian liên quan có thể thay đổi từ điểm nhìn này đến điểm nhìn khác. Do đó, tập hợp các khung nguồn MVD được sử dụng để chụp cảnh ba chiều sẽ bị ảnh hưởng bởi những thay đổi đó nên các phân mảnh của atlas (đã được trích xuất hoặc tổng hợp từ tập hợp các khung MVD này). Các phân mảnh khác nhau liên quan đến các vùng gổ chồng của không gian 3D có thể có các màu sắc khác nhau. Tuy nhiên, việc trộn các phân mảnh xuất phát từ các khung hình khác nhau có thể dẫn đến chấp phân mảnh không đồng nhất và không tương thích với cách hình dung xxphù hợp về cảnh ba chiều.

Khi xảy ra sự gộp chồng giữa các phân mảnh xảy ra, các chiến lược khác nhau có thể được hình dung để giảm thiểu vấn đề này. Cách tiếp cận đầu tiên bao gồm việc chọn phân mảnh đến từ khung hình gần với khung hình hiện tại nhất để tổng hợp. Mặc dù cách tiếp cận này đơn giản nhưng có thể gây ra các ảnh giả tạm thời quan trọng khi tổng hợp các khung nhìn từ một camera ảo đang chuyển động. Quả thực, trong trường hợp ví dụ về hai phân mảnh được trích xuất từ hai khung hình gắn liền với vị trí quan sát P1 và P2 và biểu diễn cùng một phần của không gian, khi tâm camera ảo O đang di chuyển và vượt qua đường biên đồng khoảng cách ($OP1 = OP2$), bộ tổng hợp đột ngột chuyển từ phân mảnh này sang phân mảnh khác gây ra hiện tượng gọi là "ảnh giả nhấp nháy". Những ảnh giả này thậm chí còn đáng lo ngại hơn nếu camera đang chuyển động thường xuyên vượt qua đường biên đồng khoảng cách. Cách tiếp cận thứ hai bao gồm tính trọng số phần góp của mỗi phân mảnh thay vì chọn một trong số chúng. Trọng số có thể, ví dụ, được suy luận từ các khoảng cách $OP1$ và $OP2$. Cách tiếp cận này tránh "vệch nhấp nháy" ảnh giả, nhưng hình ảnh tổng hợp cuối cùng bị mờ vì phần góp của mỗi phân mảnh được tính trung bình. Tuy nhiên, hai cách tiếp cận này không phù hợp khi các phân mảnh không chồng lên nhau. Hình 10 minh họa chi tiết hình ảnh khung nhìn của cảnh trong hình 9 từ một điểm nhìn nhất định trong không gian xem 3DoF+ liên kết. Các hình ảnh 101 và 103 là các chi tiết của hình ảnh khung nhìn khi được tạo bằng cách sử dụng sáng chế. Các hình ảnh 102 và 104 là các chi tiết giống nhau của cùng một hình ảnh khung nhìn khi được tạo bằng cách sử dụng phương pháp tiếp cận trọng số (được gọi là phương pháp tiếp cận thứ hai ở trên). Trên hình ảnh 102, bình và ly ít trong suốt hơn trên hình 101 và độ phản chiếu con nhện trong kính lò trên hình 104 kém sắc nét hơn hình 103 do tình trạng kết hợp và/hoặc trung bình của các phân mảnh khác nhau xuất phát từ các điểm nhìn khác nhau đáng kể.

Theo sáng chế, ở giai đoạn mã hóa, một khung hình nguồn (hoặc một tập hợp con khung hình nguồn hạn chế) được chọn làm khung hình tham chiếu và được coi ở giai đoạn giải mã là truyền tải thông tin chiếu sáng. Các phân mảnh liên quan đến khung hình tham chiếu này được chọn ở giai đoạn giải mã làm màu sắc kết xuất (ưu tiên) ngay sau khi phân mảnh được sử dụng để tạo hình ảnh khung nhìn từ vị trí quan sát hiện tại.

Khung hình tham chiếu được chọn trong số các khung hình làm khung hình bao phủ tốt hơn cảnh ba chiều về trường nhìn từ vị trí trung bình của người dùng cuối trong không gian xem 3DoF+. Đảm bảo điều kiện sau này làm giảm các hiệu ứng nhấp vào /

chấp phân mảnh có thể xảy ra khi khung hình đã chọn không bao phủ cổng nhìn của người dùng nữa (do đổ bóng khe hở hoặc trường nhìn quá nhỏ). Ví dụ: khung hình tham chiếu được chọn trong số các khung hình trong khung MVD nguồn là khung hình có vị trí trung tâm nhất (gần nhất tính từ khối tâm của tất cả tâm quang học chế độ đầu vào). Trên ví dụ của hình 8, khung hình 823 hoặc khung hình 833 (hoặc cả hai hoặc khung hình được tổng hợp từ cả hai) được chọn làm khung hình tham chiếu. Trong một biến thể, tổng hợp của khung hình trung tâm 4π steradian (như trong hình 6 và 7) được thực thi dưới dạng giai đoạn mã hóa và khung hình trung tâm này được chọn làm khung hình tham chiếu (chỉ với một phân mảnh liên quan).

Khung hình tham chiếu phải được phát tín hiệu đến bộ giải mã để bộ giải mã này có thể sử dụng đúng cách ở giai đoạn kết xuất. Thông tin này có thể được chỉ báo trong siêu dữ liệu ở cấp độ mô tả thông số camera với dãy Boole tùy chọn bổ sung `specular_confidence` (khả dụng khi và chỉ khi `specular_confidence_params_equal_flag` được đặt thành true (đúng)) trong đó mỗi thành phần là đúng nếu khung hình / camera đầu vào liên quan được sử dụng làm chuẩn để tạo bóng, như được đề xuất trong bảng sau:

<code>camera_params_list() {</code>	Mô tả
<code>num_cameras_minus1</code>	u(16)
<code>for (i = 0; i <= num_cameras_minus1; i++) {</code>	
<code>cam_pos_x[i]</code>	u(32)
<code>cam_pos_y[i]</code>	u(32)
<code>cam_pos_z[i]</code>	u(32)
<code>cam_yaw[i]</code>	u(32)
<code>cam_pitch[i]</code>	u(32)
<code>cam_roll[i]</code>	u(32)
<code>}</code>	
<code>intrinsic_params_equal_flag</code>	u(1)
<code>for (i = 0; i <= intrinsic_params_equal_flag ? 0 : num_cameras_minus1; i++)</code>	
<code>camera_intrinsics([i])</code>	

depth_quantization_params_equal_flag	u(1)
for (i = 0; i <= depth_quantization_equal_flag ? 0 : num_cameras_minus1; i++)	
depth_quantization([i])	
specular_confidence_params_equal_flag	u(1)
for (i = 0; i <= specular_confidence_params_equal_flag ? 0 : num_cameras_minus1; i++)	
specular_confidence[i]	u(1)

Trong trường hợp sử dụng một khung hình duy nhất, sử dụng một trường số nguyên không dấu 32 bit duy nhất specular_confidence_id chỉ định khung hình tham chiếu có thể được xem xét theo cách khác. Một giá trị cụ thể (ví dụ: 0xffffffff hoặc num_cameras_minus1 + 1) cũng có thể được chọn để tắt tính năng này, như được minh họa trong bảng sau:

camera_params_list() {	Mô tả
num_cameras_minus1	u(16)
for (i= 0; i <= num_cameras_minus1; i++) {	
cam_pos_x[i]	u(32)
cam_pos_y[i]	u(32)
cam_pos_z[i]	u(32)
cam_yaw[i]	u(32)
cam_pitch[i]	u(32)
cam_roll[i]	u(32)
}	
intrinsic_params_equal_flag	u(1)
for (i = 0; i <= intrinsic_params_equal_flag ? 0 : num_ cameras_minus1; i++)	
camera_intrinsics([i])	
depth_quantization_params_equal_flag	u(1)

for (i = 0; i <= depth_quantization_equal_flag ? 0 : num_cameras_minus1; i++)	
depth_quantization([i])	
specular_confidence_id	u(32)

Ở giai đoạn giải mã, (tập hợp) mã nhận biết camera / khung hình tham chiếu cho (hiệu ứng ánh sáng) tạo bóng được khôi phục. Mỗi phân mảnh liên kết với (những) camera tham chiếu này được sử dụng ưu tiên mỗi khi camera bao phủ khung nhìn để tổng hợp. Khi làm như vậy, các hiệu ứng ánh sáng được bắt sáng và không làm thay đổi bất kỳ vị trí nhìn nào như mong đợi.

Hình 11 minh họa phương pháp 110 mã hóa nội dung video ba chiều, theo một phương án không giới hạn của sáng chế. Tại bước 111, cảnh 3D cần mã hóa được thu từ nguồn. Cảnh 3D được cung cấp dưới dạng nội dung nhiều khung hình cộng với độ sâu, ví dụ như khung MVD, được chụp bởi các camera khác nhau, mỗi camera có các thông số bên ngoài và bên trong riêng. Theo sáng chế, một khung hình (hoặc một tập hợp con khung hình nguồn hạn chế) được chọn làm khung hình tham chiếu. khung hình tham chiếu được chọn trong số các khung hình làm khung hình bao phủ tốt hơn cảnh ba chiều về trường nhìn từ vị trí trung bình của người dùng cuối trong không gian xem 3DoF+. Các khung hình được sử dụng để tạo một tập hợp phân mảnh có thể được đóng gói trong một hình ảnh atlas, một phân mảnh là một hình ảnh được trích xuất từ một khung hình. Ở bước 112, siêu dữ liệu được tạo và liên kết với các phân mảnh và/hoặc với atlas. Siêu dữ liệu đã tạo bao gồm thông tin mô tả các thông số bên ngoài và bên trong của camera (tức là các thông số liên quan đến từng khung hình) và liên kết từng phân mảnh với khung hình mà nó được trích xuất. Theo sáng chế, siêu dữ liệu bao gồm thông tin chỉ báo khung hình tham chiếu được chọn ở bước 111. Ở bước 113, cảnh ba chiều được mã hóa trong một luồng dữ liệu, luồng dữ liệu bao gồm dữ liệu biểu diễn atlas phân mảnh và siêu dữ liệu liên quan. Phương pháp này áp dụng trên chuỗi nội dung MVD để mã hóa video ba chiều. Hình 12 minh họa phương pháp 120 giải mã nội dung video ba chiều, theo một phương án không giới hạn của sáng chế. Ở bước 121, luồng dữ liệu biểu diễn cảnh ba chiều được thu từ nguồn. Luồng dữ liệu được giải mã để truy xuất hình ảnh atlas bao gồm ít nhất một hình ảnh phân mảnh, phân mảnh là hình ảnh được trích xuất từ khung hình cảnh 3D. Hình ảnh atlas được liên kết trong luồng dữ liệu với

siêu dữ liệu đã tạo bao gồm thông tin mô tả các thông số bên ngoài và bên trong của camera (tức là các thông số liên quan đến từng khung hình) và liên kết từng phân mảnh với khung hình mà nó được trích xuất. Ở bước 122, mã nhận biết của khung hình tham chiếu được truy xuất từ siêu dữ liệu. Phân mảnh được liên kết với khung hình tham chiếu là phân mảnh tham chiếu. Ở bước 123, hình ảnh khung nhìn được tạo cho điểm nhìn hiện tại của người dùng trong không gian xem 3DoF+ của cảnh 3D. Thành phần màu sắc của pixel trong hình ảnh khung nhìn được chọn theo pixel trong phân mảnh làm chức năng của điểm nhìn hiện tại và các thông số của từng khung hình và phân mảnh. Khi hai phân mảnh được liên kết với các khung hình khác nhau có thể góp phần vào thành phần màu sắc của pixel trong hình ảnh khung nhìn, nếu một trong hai phân mảnh này được liên kết với khung hình tham chiếu, màu sắc do phân mảnh tham chiếu này cung cấp được quy cho pixel trong khung nhìn. Nếu không, kết hợp hai màu sắc (hoặc lựa chọn một trong số chúng) được sử dụng cho pixel trong khung nhìn.

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. Các 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 trình độ 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 thông thường 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 đăng ký sáng chế này.

YÊU CẦU BẢO HỘ

1. Phương pháp mã hóa cảnh ba chiều (Three dimensions - 3D) trong luồng dữ liệu, cảnh 3D được tạo ra dưới dạng khung nhiều khung hình cộng độ sâu (Multi-view plus Depth - MVD), phương pháp này bao gồm các bước:

chọn khung hình từ khung nhiều khung hình cộng độ sâu (Multi-view plus Depth - MVD) làm khung hình tham chiếu dựa trên phạm vi bao quát trường nhìn của cảnh 3D từ vị trí trung bình trong hộp giới hạn xem trong cảnh 3D;

tạo hình ảnh atlas, hình ảnh atlas đóng gói các phân mảnh, phân mảnh là một phần hình ảnh của khung hình trong khung nhiều khung hình cộng độ sâu (Multi-view plus Depth - MVD), hình ảnh atlas bao gồm thông tin đặc trưng của cảnh 3D có thể nhìn thấy từ hộp giới hạn xem;

tạo siêu dữ liệu bao gồm thông tin liên kết từng phân mảnh với khung hình, và ký hiệu nhận biết khung hình tham chiếu đã chọn; và

mã hóa atlas và siêu dữ liệu trong luồng dữ liệu.

2. Phương pháp theo điểm 1, trong đó các phân mảnh được thu bằng cách cắt xén thông tin dư thừa giữa các khung hình và bảo toàn các điểm ảnh gối chồng giữa các phân mảnh.

3. Phương pháp theo điểm 1, trong đó khung hình tham chiếu được đóng gói trong atlas như là phân mảnh chưa được cắt xén.

4. Phương tiện lưu trữ có thể đọc được bằng máy tính không tạm thời trên đó có lưu trữ các lệnh để khiến cho một hoặc nhiều bộ xử lý thực hiện phương pháp theo điểm 1.

5. Phương pháp kết xuất hình ảnh khung nhìn đại diện của cảnh 3D có thể nhìn thấy từ hộp giới hạn xem trong cảnh 3D, phương pháp này bao gồm các bước:

giải mã hình ảnh atlas và siêu dữ liệu liên quan từ luồng dữ liệu, hình ảnh atlas đóng gói các phân mảnh, phân mảnh là một phần hình ảnh của khung hình trong khung nhiều khung hình cộng độ sâu (Multi-view plus Depth - MVD), hình ảnh atlas bao gồm thông tin đặc trưng của cảnh 3D có thể nhìn thấy từ hộp giới hạn xem, và siêu dữ liệu

bao gồm thông tin liên kết từng phân mảnh với khung hình, và ký hiệu nhận biết khung hình tham chiếu, phân mảnh được liên kết với khung hình tham chiếu là phân mảnh tham chiếu; và

khi hai phân mảnh góp phần vào hình ảnh khung nhìn, kết xuất hình ảnh khung nhìn cho điểm nhìn hiện tại trong hộp giới hạn xem nếu một trong hai phân mảnh là phân mảnh tham chiếu, chỉ sử dụng phần góp của phân mảnh tham chiếu cho hình ảnh khung nhìn.

6. Phương pháp theo điểm 5, trong đó khung hình tham chiếu là một trong số các phân mảnh của atlas.

7. Thiết bị mã hóa cảnh 3D trong luồng dữ liệu, cảnh 3D được tạo ra dưới dạng khung nhiều khung hình cộng độ sâu (Multi-view plus Depth - MVD), thiết bị này bao gồm bộ xử lý được định cấu hình để:

chọn khung hình từ khung nhiều khung hình cộng độ sâu (Multi-view plus Depth - MVD) làm khung hình tham chiếu dựa trên phạm vi bao quát trường nhìn của cảnh 3D từ vị trí trung bình trong hộp giới hạn xem trong cảnh 3D;

tạo hình ảnh atlas, hình ảnh atlas đóng gói các phân mảnh, phân mảnh là một phần hình ảnh của khung hình trong khung nhiều khung hình cộng độ sâu (Multi-view plus Depth - MVD), hình ảnh atlas bao gồm thông tin đặc trưng của cảnh 3D có thể nhìn thấy từ hộp giới hạn xem;

tạo siêu dữ liệu bao gồm thông tin liên kết từng phân mảnh với khung hình, và ký hiệu nhận biết khung hình tham chiếu đã chọn; và

mã hóa atlas và siêu dữ liệu trong luồng dữ liệu.

8. Thiết bị theo điểm 7, trong đó các phân mảnh được thu bằng cách cắt xén thông tin dư thừa giữa các khung hình và bảo toàn các điểm ảnh gối chồng giữa các phân mảnh.

9. Thiết bị theo điểm 7, trong đó khung hình tham chiếu được đóng gói trong atlas như là phân mảnh chưa được cắt xén.

10. Thiết bị kết xuất hình ảnh khung nhìn đại diện của cảnh 3D có thể nhìn thấy từ hộp giới hạn xem trong cảnh 3D, thiết bị này bao gồm bộ xử lý được định cấu hình để:

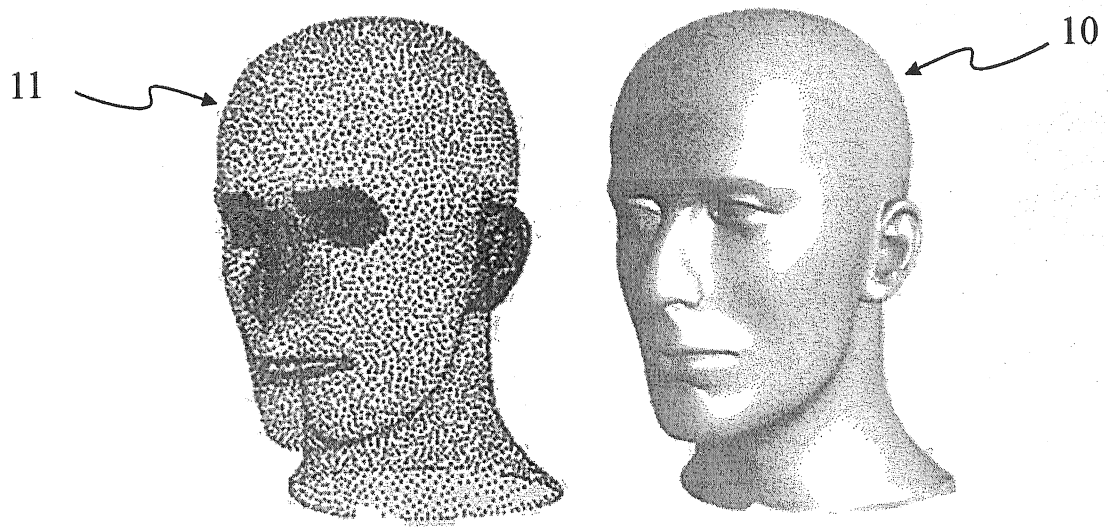
giải mã hình ảnh atlas và siêu dữ liệu liên quan từ luồng dữ liệu, hình ảnh atlas đóng gói các phân mảnh, phân mảnh là một phần hình ảnh của khung hình trong khung nhiều khung hình cộng độ sâu (Multi-view plus Depth - MVD), hình ảnh atlas bao gồm thông tin đặc trưng của cảnh 3D có thể nhìn thấy từ hộp giới hạn xem, và siêu dữ liệu bao gồm thông tin liên kết từng phân mảnh với khung hình, và ký hiệu nhận biết khung hình tham chiếu, phân mảnh được liên kết với khung hình tham chiếu là phân mảnh tham chiếu; và

kết xuất hình ảnh khung nhìn cho điểm nhìn hiện tại trong hộp giới hạn xem bằng cách:

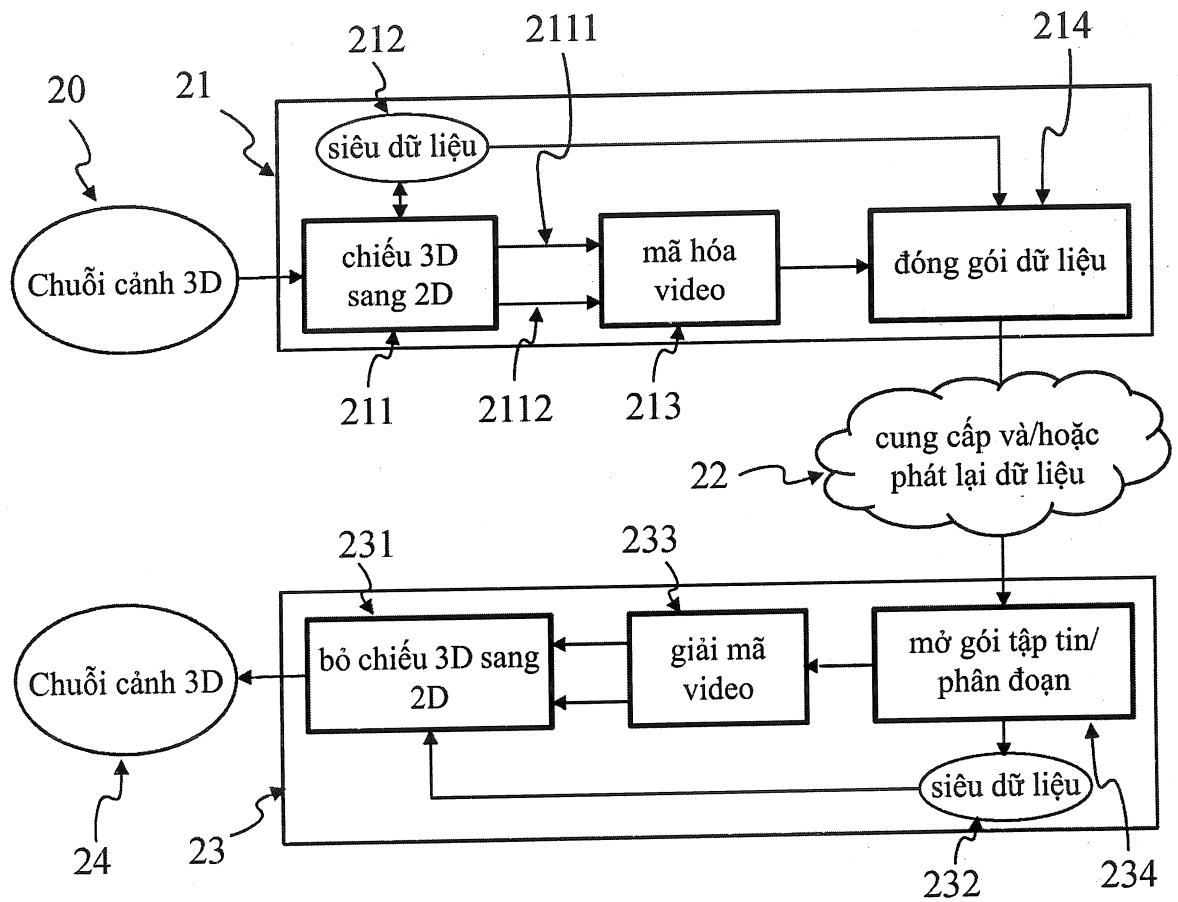
khi hai phân mảnh góp phần vào hình ảnh khung nhìn, nếu một trong hai phân mảnh là phân mảnh tham chiếu, chỉ sử dụng phần góp của phân mảnh tham chiếu cho hình ảnh khung nhìn.

11. Thiết bị theo điểm 10, trong đó khung hình tham chiếu là một trong số các phân mảnh của atlas.

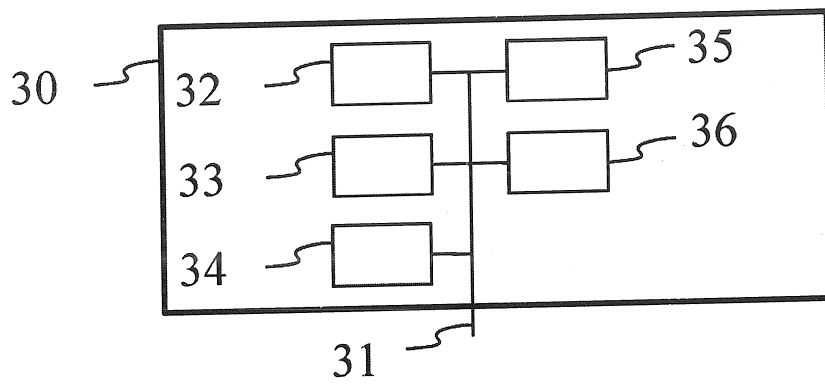
12. Phương tiện lưu trữ có thể đọc được bằng máy tính không tạm thời trên đó có lưu trữ các lệnh để khiến cho một hoặc nhiều bộ xử lý thực hiện phương pháp theo điểm 5.



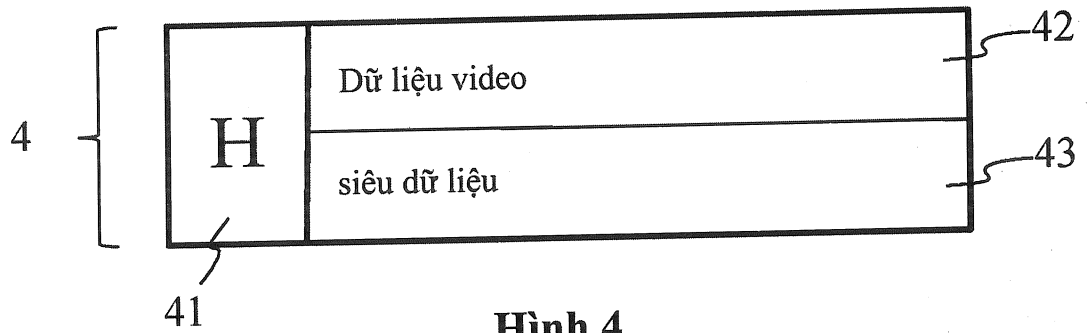
Hình 1



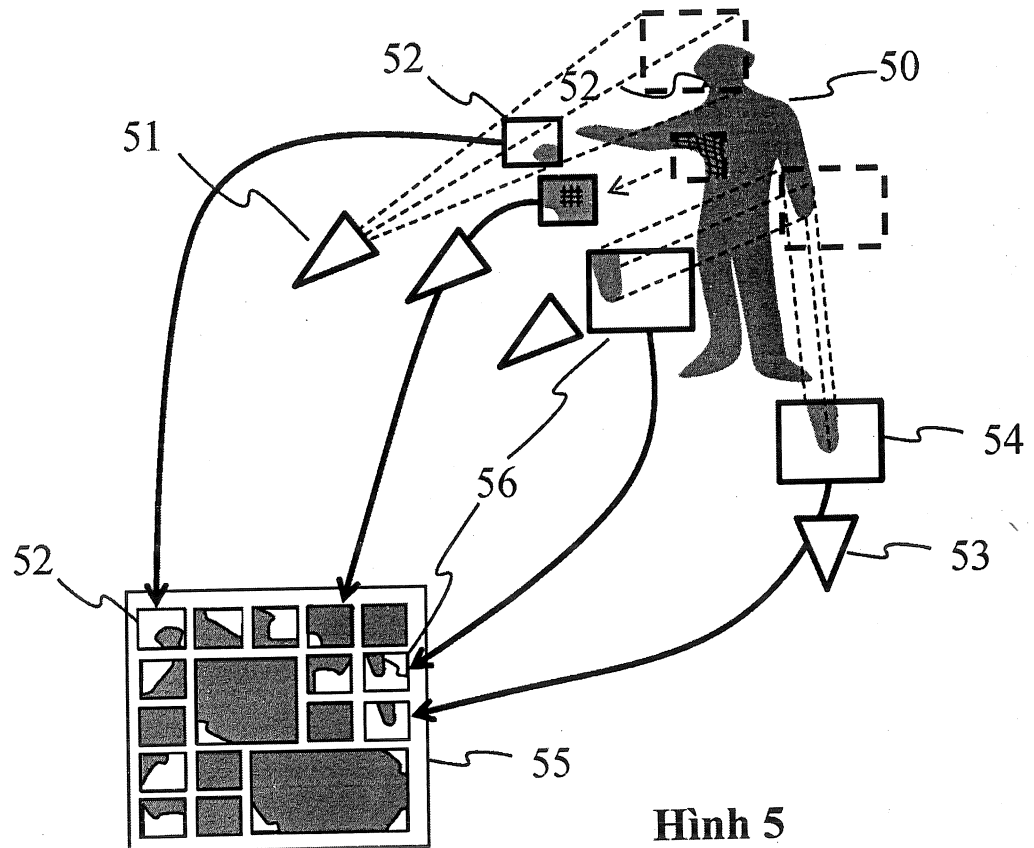
Hình 2



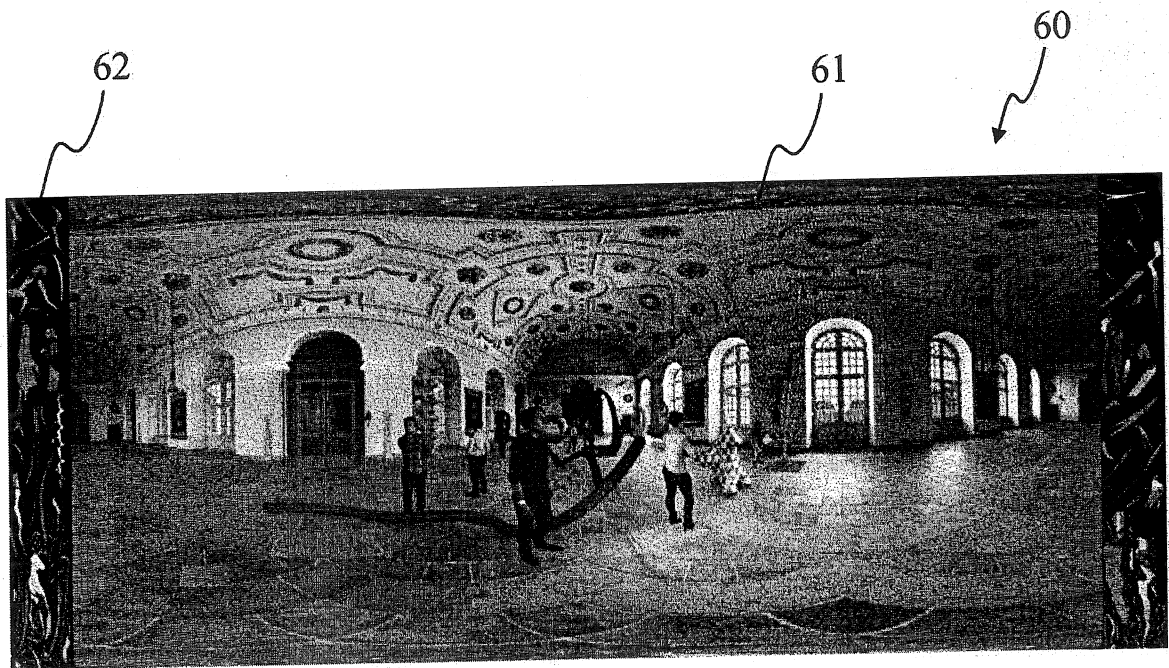
Hình 3



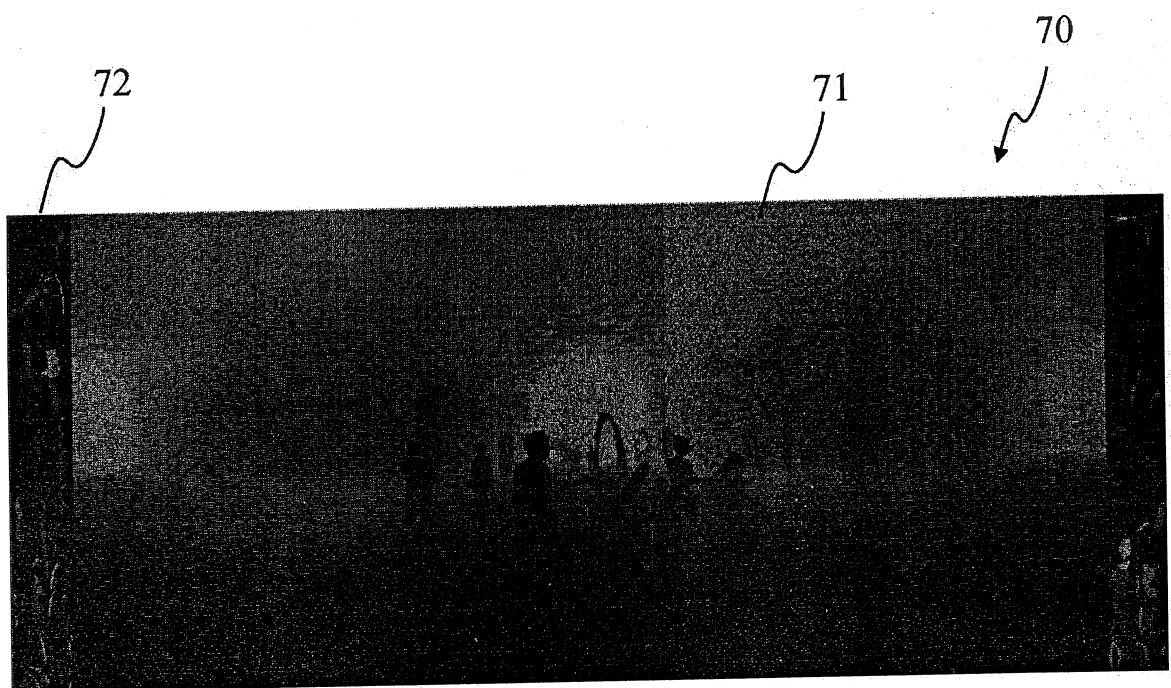
Hình 4



Hình 5



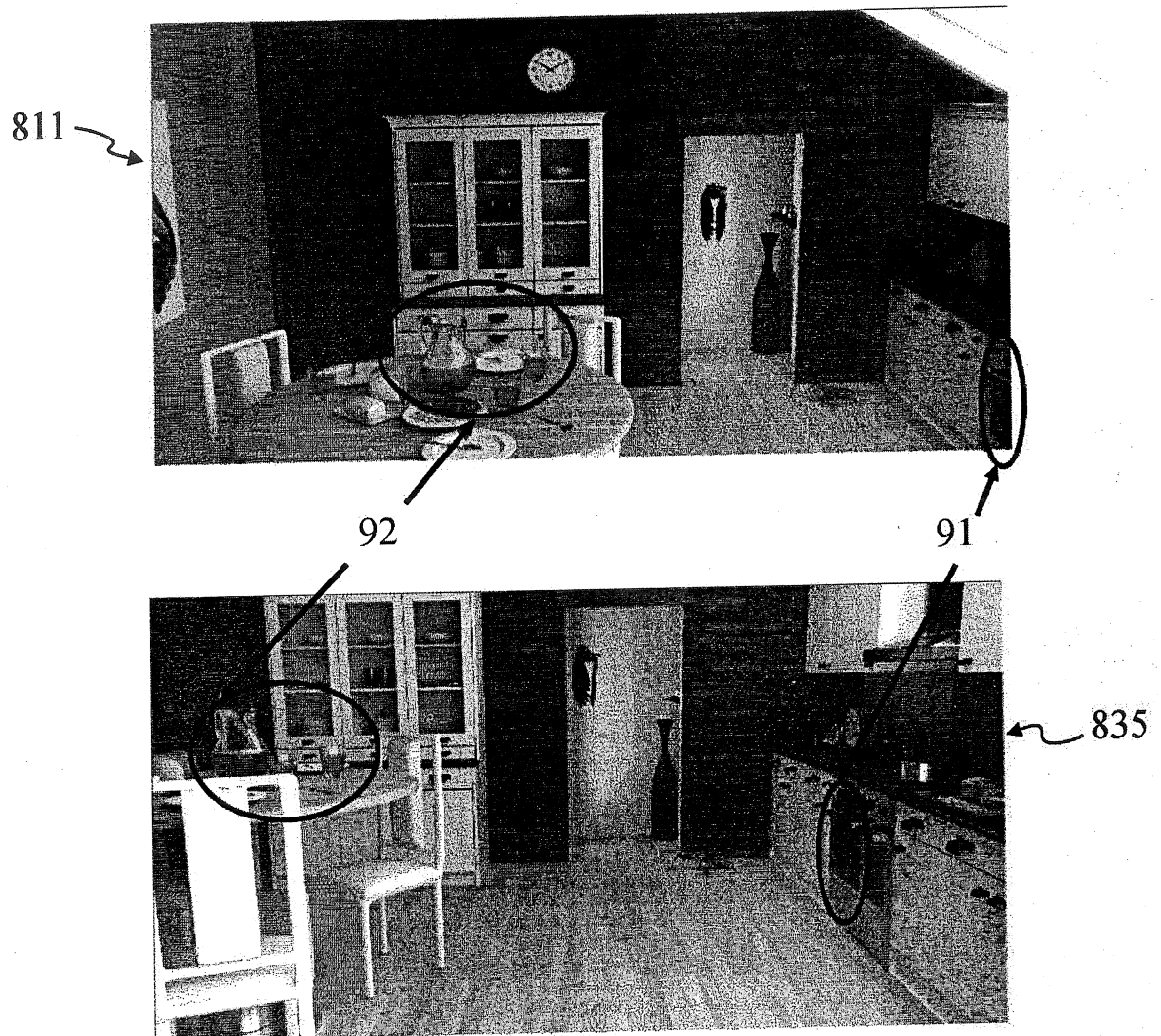
Hình 6



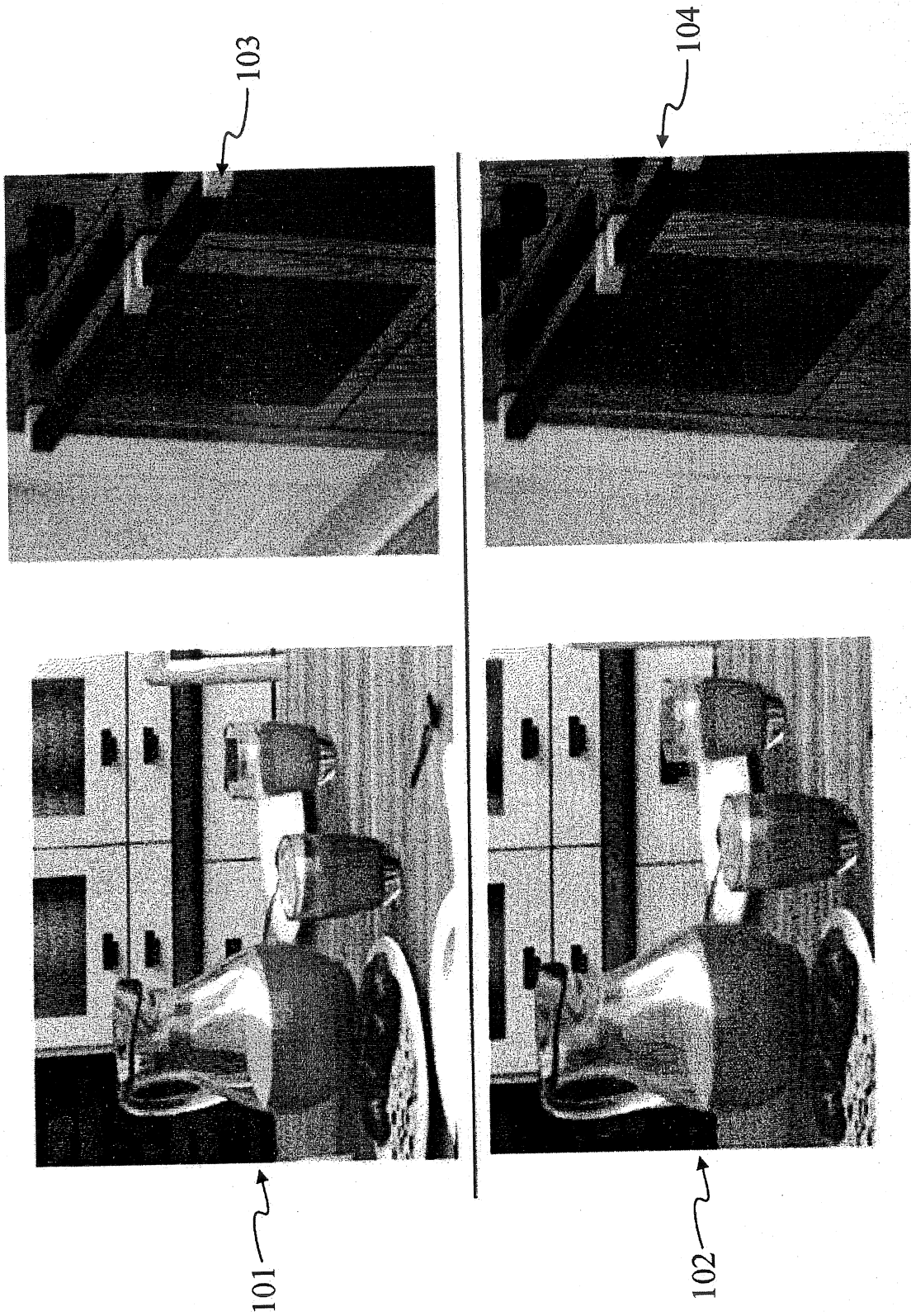
Hình 7

811	812	813	814	815
821	822	823	824	825
831	832	833	834	835
841	842	843	844	845

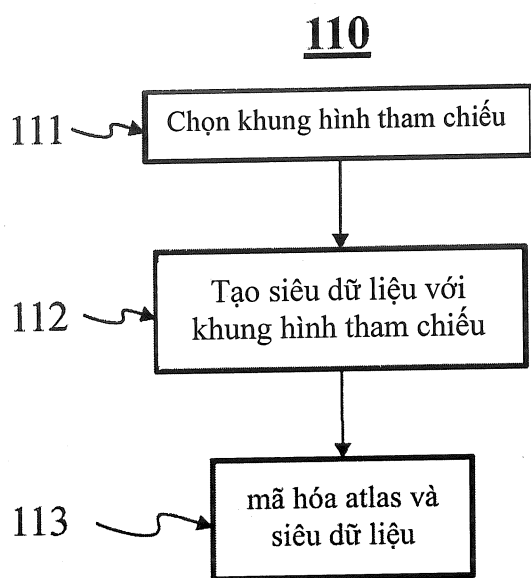
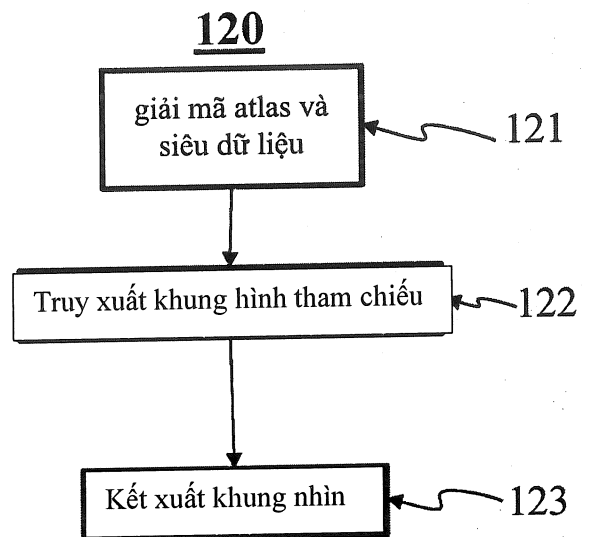
Hình 8



Hình 9



Hình 10

**Hình 11****Hình 12**