



- (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
(51)^{2020.01} H04N 21/218; H04N 21/2387; H04N 21/81; H04N 21/234 (13) B



1-0044693

-
- (21) 1-2021-00602 (22) 21/06/2019
(86) PCT/US2019/038407 21/06/2019 (87) WO2020/013975 16/01/2020
(30) 18305943.5 12/07/2018 EP
(45) 25/04/2025 445 (43) 26/04/2021 397A
(71) INTERDIGITAL VC HOLDINGS, INC. (US)
200 Bellevue Parkway, Suite 300, Wilmington, Delaware 19809, United States of America
(72) FLEUREAU, Julien (FR); DORE, Renaud (FR); SALMON-LEGAGNEUR, Charles (FR); HOUDAILLE, Remi (FR).
(74) Công ty Luật TNHH T&G (TGVN)
-

- (54) PHƯƠNG PHÁP VÀ THIẾT BỊ ĐỂ XỬ LÝ NỘI DUNG VIDEO, VÀ PHƯƠNG TIỆN ĐỌC ĐƯỢC BỞI BỘ XỬ LÝ KHÔNG CHUYỂN TIẾP

(21) 1-2021-00602

(57) Phương pháp và thiết bị được đề xuất để truyền thông tin biểu diễn điểm nhìn trong cảnh ba chiều (Three-Dimensional, 3D) được biểu diễn với tập hợp các nội dung video thể tích; nhận nội dung video thể tích thứ nhất của tập hợp, nội dung video thể tích thứ nhất đang theo phạm vi của các điểm nhìn bao gồm điểm nhìn, nội dung video thể tích thứ nhất được biểu diễn với tập hợp các bản vá thứ nhất, mỗi bản vá trong số các bản vá thứ nhất tương ứng với sự tham số hóa hai chiều (Two-Dimensional, 2D) của nhóm thứ nhất của các điểm trong phần 3D của cảnh 3D được kết hợp với nội dung video thể tích thứ nhất, và ít nhất một bản vá thứ nhất tham chiếu đến khu vực của bản vá thứ hai tương ứng với sự tham số hóa 2D của nhóm thứ hai của các điểm trong một phần 3D khác của cảnh 3D được kết hợp với nội dung video thể tích thứ hai của tập hợp các nội dung video thể tích.

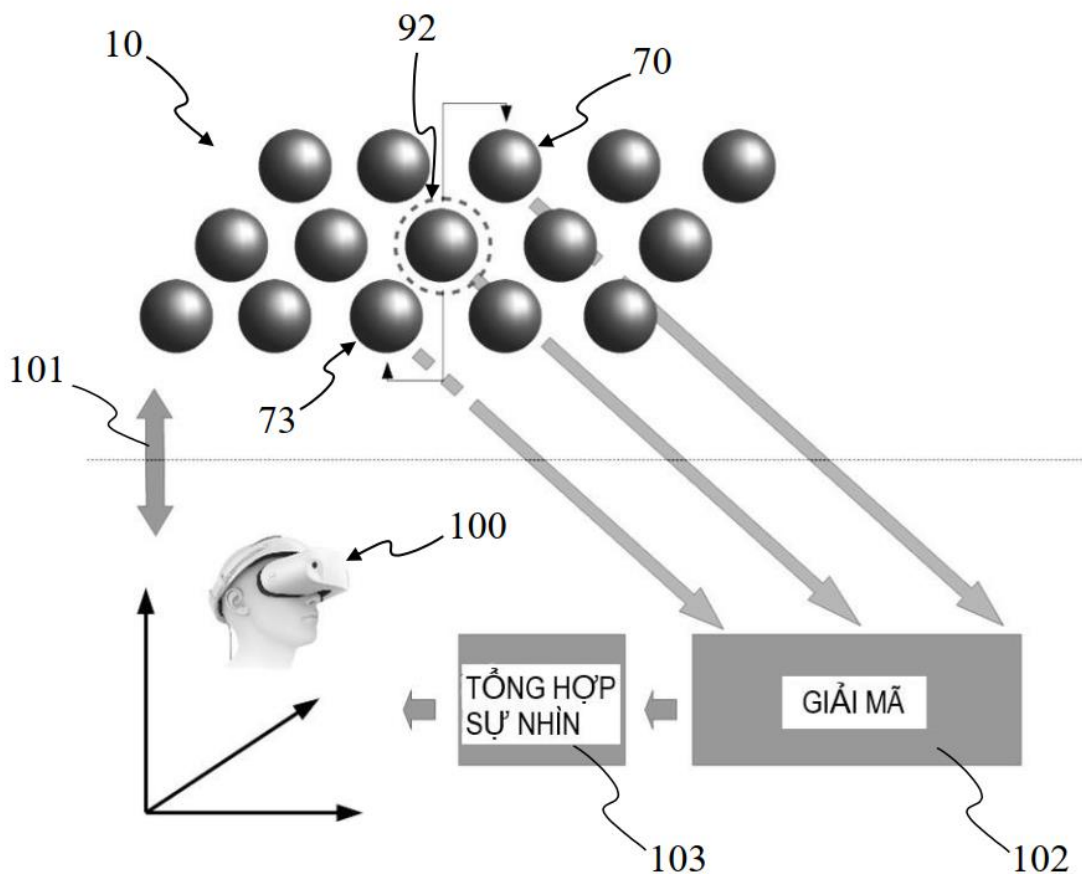


Fig.10

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến lĩnh vực cảnh ba chiều (Three-Dimensional, 3D) và nội dung video thể tích. Sáng chế cũng được hiểu theo ngữ cảnh của việc mã hóa và/hoặc việc tạo định dạng của dữ liệu biểu diễn cho độ sâu của cảnh 3D, ví dụ, đối với việc kết xuất của nội dung thể tích trên các thiết bị người dùng cuối như là các thiết bị di động hoặc các bộ hiển thị gắn trên đầu.

Tình trạng kỹ thuật của sáng chế

Phần này nhằm để giới thiệu cho người đọc các khía cạnh khác nhau của lĩnh vực kỹ thuật, vốn có thể liên quan đến các khía cạnh khác nhau mà được mô tả và/hoặc được yêu cầu bảo hộ dưới đây. Phần thảo luận này sẽ hữu ích trong việc cung cấp cho người đọc thông tin cơ bản để tạo điều kiện cho việc hiểu rõ hơn về sáng chế. Theo đó, phần này cần được hiểu như đã nêu trên, chứ không phải là sự thừa nhận giải pháp đã biết.

Gần đây đã có sự gia tăng của nội dung trường nhìn lớn sẵn có (lên đến 360°). Người dùng có khả năng không thể xem hết toàn bộ nội dung như vậy khi xem các nội dung đó trên các thiết bị đa dạng như là như là các bộ hiển thị gắn trên đầu, kính thông minh, màn hình PC, máy tính bảng, điện thoại thông minh và loại tương tự. Điều đó có nghĩa là tại một thời điểm nào đó, người dùng có thể chỉ xem một phần của nội dung. Tuy nhiên, người dùng thường có thể điều hướng trong nội dung theo các cách thức khác nhau như là chuyển động đầu, chuyển động chuột, màn hình chạm, giọng nói và

loại tương tự. Người ta cũng thường mong muốn thường mã hóa và giải mã nội dung này.

Video nhập vai, cũng được gọi là video phẳng 360°, cho phép người dùng xem toàn bộ xung quanh mình thông qua các thao tác quay đầu xung quanh điểm nhìn tĩnh. Các thao tác quay chỉ cho phép việc trải nghiệm 3 cấp độ tự do (3 Degrees of Freedom, 3DoF). Ngay cả nếu video 3DoF đủ cho việc trải nghiệm video toàn hướng thứ nhất, ví dụ, sử dụng bộ hiển thị gắn trên đầu thiết bị (Head-Mounted Display, HMD), thì video 3DoF có thể nhanh chóng trở nên gây bức dọc cho người xem vốn mong đợi nhiều tự do hơn, ví dụ, bằng trải nghiệm thị sai. Bên cạnh đó, 3DoF cũng có thể gây ra sự chóng mặt bởi vì người dùng không bao giờ chỉ quay đầu mà còn dịch chuyển đầu theo ba hướng, các dịch chuyển không được tái lập trong các trải nghiệm video 3DoF.

Video thể tích (cũng được biết đến như video 6 cấp độ tự do (6 Degrees of Freedom, 6DoF)) là sự thay thế cho video 3DoF. Khi xem video 6DoF, cùng với việc quay, thì người dùng cũng có thể dịch chuyển đầu mình, và thậm chí thân mình, theo nội dung được xem và trải nghiệm thị sai và thậm chí các thể tích. Video như vậy tăng đáng kể cảm nhận của sự nhúng và thụ cảm độ sâu của cảnh và cũng ngăn sự hoa mắt bằng cách cấp sự hồi đáp thị giác nhất quán trong quá trình dịch chuyển đầu. Nội dung được tạo nên bằng các cảm biến dành riêng cho phép sự ghi màu sắc và độ sâu của cảnh quan tâm đồng thời. Việc sử dụng trang thiết bị gồm các camera màu được kết hợp với các công nghệ trắc lượng chụp ảnh là cách phổ biến để thực hiện việc ghi này.

Lượng dữ liệu nội dung video thể tích như vậy là quan trọng và đòi hỏi các dung lượng lưu trữ lớn và tốc độ bit cao cần để vận chuyển dữ liệu như vậy.

Bản chất kỹ thuật của sáng chế

Các tham chiếu trong bản mô tả đến “một phương án”, “phương án”, “phương án ví dụ”, “phương án cụ thể” chỉ ra rằng phương án được mô tả có thể gồm dấu hiệu, cấu trúc, hoặc đặc tính cụ thể, nhưng mọi phương án có thể không nhất thiết phải gồm dấu hiệu, cấu trúc, hoặc đặc tính cụ thể này. Ngoài ra, các cụm này không nhất thiết là đang đề cập đến cùng một phương án. Hơn nữa, khi dấu hiệu, cấu trúc, hoặc đặc tính cụ thể được mô tả trong sự kết nối với phương án, thì cần được xem xét rằng trong phạm vi hiểu biết của người có hiểu biết trung bình trong lĩnh vực kỹ thuật tương ứng để ảnh hưởng tới dấu hiệu, cấu trúc, hoặc đặc tính như vậy trong sự kết nối với các phương án khác sẽ được mô tả một cách rõ ràng hoặc không.

Theo một phương án, phương pháp bao gồm các bước:

- nhận thông tin biểu diễn cho điểm nhìn trong cảnh 3D, cảnh 3D được biểu diễn với tập hợp các nội dung video thể tích;
- truyền nội dung video thể tích thứ nhất của tập hợp, nội dung video thể tích thứ nhất đang theo phạm vi của các điểm nhìn bao gồm điểm nhìn, nội dung video thể tích thứ nhất được biểu diễn với tập hợp các bản vá thứ nhất, mỗi bản vá thứ nhất của tập hợp tương ứng với sự tham số hóa hai chiều của nhóm thứ nhất của các điểm được bao gồm trong phần 3D của cảnh 3D được kết hợp với nội dung video thể tích thứ nhất, ít nhất bản vá thứ nhất của tập hợp tham chiếu đến ít nhất khu vực của ít nhất bản vá thứ hai tương ứng với sự tham số hóa hai chiều của nhóm thứ hai của các điểm được bao gồm trong một phần 3D khác của cảnh 3D được kết hợp với nội dung video thể tích thứ hai của tập hợp các nội dung video thể tích.

Theo một phương án khác, thiết bị được tạo cấu hình để:

- nhận thông tin biểu diễn cho điểm nhìn trong cảnh 3D, cảnh 3D được biểu diễn với tập hợp các nội dung video thể tích;
- truyền nội dung video thể tích thứ nhất của tập hợp, nội dung video thể tích thứ nhất đang theo phạm vi của các điểm nhìn bao gồm điểm nhìn, nội dung video thể tích thứ nhất được biểu diễn với tập hợp các bản vá thứ nhất, mỗi bản vá thứ nhất của tập hợp tương ứng với sự tham số hóa hai chiều của nhóm thứ nhất của các điểm được bao gồm trong phần 3D của cảnh 3D được kết hợp với nội dung video thể tích thứ nhất, ít nhất bản vá thứ nhất của tập hợp tham chiếu đến ít nhất khu vực của ít nhất bản vá thứ hai tương ứng với sự tham số hóa hai chiều của nhóm thứ hai của các điểm được bao gồm trong một phần 3D khác của cảnh 3D được kết hợp với nội dung video thể tích thứ hai của tập hợp các nội dung video thể tích.

Theo phương án khác, phương pháp bao gồm các bước:

- truyền thông tin biểu diễn cho điểm nhìn trong cảnh 3D, cảnh 3D được biểu diễn với tập hợp các nội dung video thể tích;
- nhận nội dung video thể tích thứ nhất của tập hợp, nội dung video thể tích thứ nhất đang theo phạm vi của các điểm nhìn bao gồm điểm nhìn, nội dung video thể tích thứ nhất được biểu diễn với tập hợp các bản vá thứ nhất, mỗi bản vá thứ nhất của tập hợp tương ứng với sự tham số hóa hai chiều của nhóm thứ nhất của các điểm được bao gồm trong phần 3D của cảnh 3D được kết hợp với nội dung video thể tích thứ nhất, ít nhất bản vá thứ nhất của tập hợp tham chiếu đến ít nhất khu vực của ít nhất bản vá thứ hai tương ứng với sự tham số hóa hai chiều của nhóm thứ hai của các điểm được bao gồm trong một phần 3D khác

của cảnh 3D được kết hợp với nội dung video thể tích thứ hai của tập hợp các nội dung video thể tích.

Theo một phương án khác, thiết bị được tạo cấu hình để:

- truyền thông tin biểu diễn cho điểm nhìn trong cảnh 3D, cảnh 3D được biểu diễn với tập hợp các nội dung video thể tích;
- nhận nội dung video thể tích thứ nhất của tập hợp, nội dung video thể tích thứ nhất đang theo phạm vi của các điểm nhìn bao gồm điểm nhìn, nội dung video thể tích thứ nhất được biểu diễn với tập hợp các bản vá thứ nhất, mỗi bản vá thứ nhất của tập hợp tương ứng với sự tham số hóa hai chiều của nhóm thứ nhất của các điểm được bao gồm trong phần 3D của cảnh 3D được kết hợp với nội dung video thể tích thứ nhất, ít nhất bản vá thứ nhất của tập hợp tham chiếu đến ít nhất khu vực của ít nhất bản vá thứ hai tương ứng với sự tham số hóa hai chiều của nhóm thứ hai của các điểm được bao gồm trong một phần 3D khác của cảnh 3D được kết hợp với nội dung video thể tích thứ hai của tập hợp các nội dung video thể tích.

Theo phương án khác, sản phẩm chương trình máy tính bao gồm mã chương trình các lệnh để thực thi các bước của phương pháp (các phương pháp), khi chương trình này được thực thi trên máy tính.

Theo một phương án khác, phương tiện đọc được bởi bộ xử lý (ví dụ, không chuyển tiếp) mà đã được lưu trữ các lệnh trong đó để làm cho bộ xử lý thực hiện các bước của phương pháp (các phương pháp).

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 nổi bật lên khi đọc phần mô tả sau đây, phần mô tả có tham chiếu đến các hình vẽ kèm theo trong đó:

Fig.1 thể hiện hình ảnh biểu diễn cảnh ba chiều (3D), theo phương án không hạn chế;

Fig.2 thể hiện mô hình ba chiều (3D) của vật thể của cảnh 3D trên Fig.1 và các điểm của đám mây điểm tương ứng với mô hình 3D, theo phương án không hạn chế;

Fig.3 thể hiện ảnh bao gồm thông tin kết cấu của các điểm của cảnh 3D trên Fig.1, theo phương án không hạn chế;

Fig.4 thể hiện ảnh bao gồm thông tin độ sâu của các điểm của cảnh 3D trên Fig.1, theo phương án không hạn chế;

Fig.5 thể hiện sự tham số hóa hai chiều (Two-Dimensional, 2D) của ít nhất một phần của cảnh 3D trên Fig.1, theo phương án không hạn chế;

Fig.6 thể hiện ví dụ về việc mã hóa, việc truyền, sự nhận và việc giải mã của dữ liệu biểu diễn cho cảnh 3D, theo phương án không hạn chế;

Fig.7 và Fig.8 thể hiện việc lưu trữ và việc phát luồng của nội dung video thể tích tương ứng với cảnh 3D trên Fig.1 đối với việc kết xuất 6DoF của nội dung thể tích, theo phương án không hạn chế;

Fig.9 và Fig.10 thể hiện phương án không làm giới hạn khác của việc lưu trữ và việc phát luồng của nội dung thể tích tương ứng với cảnh 3D trên Fig.1 đối với việc kết xuất 6DoF của nội dung thể tích;

Fig.11 thể hiện ví dụ về cú pháp của luồng bit mang dữ liệu biểu diễn cho ít nhất một phần của nội dung video thể tích trên Fig.7 hoặc Fig.9, theo phương án không hạn chế;

Fig.12 thể hiện kiến trúc ví dụ của thiết bị vốn có thể được tạo cấu hình để triển khai phương pháp hoặc quy trình được mô tả liên quan tới các hình Fig.7, Fig.8, Fig.9, Fig.10, Fig.13 và/hoặc và/hoặc Fig.14, theo phương án không hạn chế;

Fig.13 thể hiện ví dụ về phương pháp truyền dữ liệu biểu diễn cho nội dung video thể tích biểu diễn cho ít nhất một phần của cảnh 3D trên Fig.1, được triển khai, ví dụ, trong thiết bị trên Fig.12, theo phương án không hạn chế;

Fig.14 thể hiện ví dụ về phương pháp nhận dữ liệu biểu diễn cho nội dung video thể tích biểu diễn cho ít nhất một phần của cảnh 3D trên Fig.1, được triển khai, ví dụ, trong thiết bị trên Fig.12, theo phương án không làm giới hạn.

Mô tả chi tiết sáng chế

Đối tượng được yêu cầu bảo hộ hiện được mô tả với tham khảo tới các hình vẽ, trong đó, các số chỉ dẫn giống nhau được sử dụng để đề cập tới các phần tử giống nhau. Trong phần mô tả sau, chỉ nhằm mục đích minh họa, nhiều chi tiết cụ thể được đặt ra nhằm tạo nên việc hiểu đầy đủ vật thể được yêu cầu bảo hộ. Tuy nhiên, có thể thấy rằng, vật thể được yêu cầu bảo hộ theo các phương án của sáng chế có thể được thực hiện mà không có các chi tiết cụ thể này.

Theo các phương án không hạn chế, thì các phương pháp và các thiết bị để truyền và/hoặc nhận dữ liệu của nội dung video thể tích sẽ được bộc lộ. Rõ ràng là những người có hiểu biết trung bình trong lĩnh vực kỹ thuật tương ứng sẽ có thể đặt ra các sắp xếp

khác nhau, mặc dù không được mô tả hoặc thể hiện rõ ràng, gồm các nguyên lý của sáng chế.

Một hoặc nhiều phương án của phương pháp (và thiết bị được tạo cấu hình để) truyền nội dung video thể tích thứ nhất của tập hợp các nội dung video thể tích biểu diễn cảnh 3D sẽ được mô tả. Thông tin biểu diễn cho điểm nhìn đã cho trong cảnh 3D sẽ được nhận. Để đáp lại sự nhận thông tin biểu diễn cho điểm nhìn đã cho, thì nội dung video thể tích thứ nhất được truyền, nội dung thể tích thứ nhất đang theo phạm vi của các điểm nhìn (tức là, nội dung video thể tích thứ nhất biểu diễn phần 3D của cảnh 3D như được nhìn thấy từ phạm vi của các điểm nhìn), phạm vi của các điểm nhìn bao gồm điểm nhìn đã cho. Nội dung video thể tích thứ nhất được biểu diễn với tập hợp các bản vá thứ nhất, mỗi bản vá thứ nhất của tập hợp tương ứng với sự tham số hóa hai chiều của nhóm thứ nhất của các điểm được bao gồm trong phần 3D của cảnh 3D được kết hợp với nội dung video thể tích thứ nhất. Ít nhất bản vá thứ nhất của tập hợp tham chiếu đến ít nhất khu vực của ít nhất bản vá thứ hai, vốn tương ứng với sự tham số hóa hai chiều của nhóm thứ hai của các điểm được bao gồm trong một phần 3D khác của cảnh 3D được kết hợp với nội dung video thể tích thứ hai của tập hợp các nội dung video thể tích.

Một hoặc nhiều phương án của phương pháp (và thiết bị được tạo cấu hình để) nhận nội dung video thể tích thứ nhất của tập hợp các nội dung video thể tích biểu diễn cảnh 3D sẽ được mô tả. Thông tin biểu diễn cho điểm nhìn đã cho trong cảnh 3D sẽ được truyền. Để đáp lại việc truyền của thông tin biểu diễn cho điểm nhìn đã cho, thì nội dung video thể tích thứ nhất được nhận, nội dung thể tích thứ nhất đang theo phạm vi của các điểm nhìn (tức là, nội dung video thể tích thứ nhất biểu diễn phần 3D của cảnh 3D như được nhìn thấy từ phạm vi của các điểm nhìn), phạm vi của các điểm nhìn

bao gồm điểm nhìn đã cho. Nội dung video thể tích thứ nhất được biểu diễn với tập hợp các bản vá thứ nhất, mỗi bản vá thứ nhất của tập hợp tương ứng với sự tham số hóa hai chiều của nhóm thứ nhất của các điểm được bao gồm trong phần 3D của cảnh 3D sẽ được kết hợp với nội dung video thể tích thứ nhất. Ít nhất bản vá thứ nhất của tập hợp tham chiếu đến ít nhất khu vực của ít nhất bản vá thứ hai, vốn tương ứng với sự tham số hóa hai chiều của nhóm thứ hai của các điểm được bao gồm trong một phần 3D khác của cảnh 3D sẽ được kết hợp với nội dung video thể tích thứ hai của tập hợp các nội dung video thể tích.

Fig.1 thể hiện hình ảnh biểu diễn cảnh ba chiều (3D) 10 bao gồm sự biểu diễn bề mặt của vài vật thể. Cảnh có thể đã được thu nhận nhờ sử dụng công nghệ phù hợp bất kỳ. Ví dụ, nó có thể đã được tạo nên nhờ sử dụng các công cụ giao diện đồ họa máy tính (Computer Graphics Interface, CGI). Nó có thể đã được thu nhận với các thiết bị thu nhận hình ảnh màu sắc và độ sâu. Trong trường hợp như vậy, có thể là một hoặc nhiều phần của các vật thể mà không thể nhìn thấy được từ các thiết bị thu nhận (ví dụ, các camera) có thể không được biểu diễn trong cảnh như được mô tả liên quan tới Fig.1. Dữ liệu được kết hợp với các phần tử (ví dụ, các điểm) của cảnh 3D được thu nhận, dữ liệu tương ứng với các thuộc tính được kết hợp với các phần tử của cảnh, tức là, các thuộc tính kết cấu (màu sắc) và/hoặc các thuộc tính hình học. Ví dụ, chuỗi các hình ảnh liên tiếp theo thời gian có thể được thu nhận. Các thuộc tính kết cấu có thể được thu nhận với một hoặc nhiều cảm biến quang và các thuộc tính hình học có thể, ví dụ, được thu nhận với một hoặc nhiều cảm biến độ sâu. Theo một biến thể, cảnh 3D dành được với công nghệ tạo hình ảnh bằng máy tính (Computer-Generated Imagery, CGI). Ít nhất một phần của cảnh 3D có thể nhìn thấy được theo nhiều điểm nhìn, ví dụ, theo phạm vi của các điểm nhìn gồm điểm nhìn trung tâm thứ nhất. Theo một biến thể, cảnh 3D không được thu nhận cũng không được tạo ra qua CGI nhưng được truy xuất từ đám mây, thu

viện của các nội dung thể tích hoặc bộ lưu trữ hoặc thiết bị lưu trữ bất kỳ. Cảnh ví dụ được minh họa trên Fig.1 bao gồm các nhân vật và các vật thể trong phòng. Cảnh 3D 10 được biểu diễn theo điểm nhìn được xác định trên Fig.1. Điểm nhìn này có thể, ví dụ, là một phần của không gian nhìn mà người dùng có thể quan sát cảnh 3D từ đó. Theo một biến thể, nội dung của cảnh 3D (thông tin độ sâu và/hoặc kết cấu) khả dụng chỉ tương ứng với các phần tử của cảnh (ví dụ, các điểm) mà có thể nhìn thấy được từ điểm nhìn được xác định trên Fig.1.

Fig.2 thể hiện mô hình ba chiều (3D) của vật thể 20 và các điểm của đám mây điểm 21 tương ứng với mô hình 3D 20. Mô hình 3D 20 và đám mây điểm 21 có thể, ví dụ, tương ứng với sự biểu diễn 3D khả thi của vật thể của cảnh 3D 10, ví dụ, đầu của nhân vật. Mô hình 20 có thể là sự biểu diễn lưới 3D và các điểm của đám mây điểm 21 có thể là các đỉnh của lưới. Các điểm của đám mây điểm 21 cũng có thể là các điểm trải trên bề mặt của các mặt của lưới. Mô hình 20 cũng có thể được biểu diễn dưới dạng phiên bản được làm phẳng của đám mây điểm 21, bề mặt của mô hình 20 được tạo nên bằng cách làm phẳng các điểm của đám mây điểm 21. Mô hình 20 có thể được biểu diễn bởi nhiều sự biểu diễn khác nhau như là các điểm ảnh ba chiều hoặc các hàm nối trục. Fig.2 minh họa thực tế rằng đám mây điểm có thể được định rõ với sự biểu diễn bề mặt của vật thể 3D và sự biểu diễn bề mặt của vật thể 3D có thể được tạo ra từ điểm của đám mây. Như được sử dụng ở đây, việc chiếu các điểm của vật thể 3D (bởi các điểm mở rộng của cảnh 3D) lên trên hình ảnh là tương ứng với việc chiếu sự biểu diễn hình ảnh bất kỳ của vật thể 3D này để tạo nên vật thể.

Đám mây điểm có thể được nhìn thấy dưới dạng cấu trúc dựa trên véctor, trong đó mỗi điểm có các tọa độ của nó (ví dụ, các tọa độ ba chiều XYZ, hoặc độ sâu/khoảng cách từ điểm nhìn đã cho) và một hoặc nhiều thuộc tính, cũng được gọi là thành phần.

Ví dụ về thành phần là thành phần màu sắc mà có thể được trình bày trong các không gian màu sắc khác nhau, ví dụ, đỏ, xanh lá cây và xanh da trời RGB (Red, Green và Blue) hoặc YUV (Y là thành phần độ chói và UV là hai thành phần sắc độ). Đám mây điểm là sự biểu diễn của vật thể như được nhìn thấy từ điểm nhìn đã cho, hoặc phạm vi của các điểm nhìn. Đám mây điểm có thể dành được bằng nhiều cách, ví dụ:

- từ việc chụp của vật thể thực được chụp bởi một giàn các camera, được bổ sung theo cách tùy chọn bởi thiết bị cảm ứng độ sâu chủ động;
- từ việc chụp của vật thể ảo/nhân tạo được chụp bởi một giàn các camera ảo trong công cụ mô hình hóa;
- từ sự hợp tác của cả vật thể thực và ảo.

Các phần thể tích của cảnh 3D có thể, ví dụ, được biểu diễn với một hoặc vài đám mây điểm như là đám mây điểm 21.

Fig.3 thể hiện ví dụ về ảnh 30 bao gồm thông tin kết cấu (ví dụ, dữ liệu RGB hoặc dữ liệu YUV) của các điểm của một phần của cảnh 3D 10, ví dụ, một phần của cảnh 3D như được nhìn thấy từ không gian nhìn (tương ứng với phạm vi của các điểm nhìn liên tục), theo phương án không làm giới hạn.

Ảnh 30 bao gồm phần thứ nhất 301 bao gồm thông tin kết cấu của các phần tử (các điểm) của một phần của cảnh 3D mà có thể nhìn thấy được từ điểm nhìn thứ nhất và một hoặc nhiều phần thứ hai 302. Thông tin kết cấu của phần thứ nhất 301 có thể, ví dụ, dành được theo việc ánh xạ chiếu hình bình hành, việc ánh xạ chiếu hình bình hành là ví dụ về việc ánh xạ chiếu hình cầu và ví dụ về sự tham số hóa hai chiều (Two-Dimensional, 2D) của một phần của cảnh 3D. Trong ví dụ trên Fig.3, các phần thứ hai được sắp xếp tại các biên trái và phải của phần thứ nhất 301 nhưng các phần thứ hai có thể được sắp

xếp theo cách khác. Các phần thứ hai 302 bao gồm thông tin kết cấu của các phần của cảnh 3D bù cho phần có thể nhìn thấy được từ điểm nhìn thứ nhất (vốn có thể tương ứng với các phần của cảnh 3D có thể nhìn thấy được theo các điểm nhìn khác của không gian nhìn nhưng từ điểm nhìn thứ nhất). Các phần thứ hai có thể dành được bằng cách loại bỏ khỏi cảnh 3D các điểm mà có thể nhìn thấy được từ điểm nhìn thứ nhất (kết cấu của chúng được lưu trữ trong phần thứ nhất) và bằng cách chiếu các điểm còn lại theo cùng điểm nhìn thứ nhất. Quy trình sau có thể được lặp đi lặp lại để dành được, tại mỗi thời điểm, các phần ẩn của cảnh 3D. Theo một biến thể, các phần thứ hai có thể dành được bằng cách loại bỏ khỏi cảnh 3D các điểm mà có thể nhìn thấy được từ điểm nhìn thứ nhất (kết cấu của chúng 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 được định tâm ở trên điểm nhìn thứ nhất. Mỗi phần thứ hai có thể được nhìn thấy dưới dạng bản vá dành được bởi sự tham số hóa 2D của một phần của cảnh tương ứng với phần thứ hai nói trên.

Phần thứ nhất 301 có thể được nhìn thấy dưới dạng bản vá kết cấu lớn thứ nhất (tương ứng với phần thứ nhất của cảnh 3D) và các phần thứ hai 302 như bao gồm tập hợp của các bản vá kết cấu nhỏ hơn (tương ứng với các phần thứ hai của cảnh 3D mà bù cho phần thứ nhất).

Ảnh 30 có thể tương ứng với thông tin kết cấu/dữ liệu của nội dung video thể tích được kết hợp với một phần của cảnh 3D 10 được nhìn thấy từ không gian nhìn. Không gian nhìn, có thể tương ứng với phạm vi của các điểm nhìn liên tục (ví dụ, được định tâm xung quanh điểm nhìn thứ nhất) được kết hợp với nội dung video thể tích này.

Fig.4 thể hiện ví dụ về ảnh 40 bao gồm thông tin độ sâu của các điểm của một phần của cảnh 3D 10, theo phương án không làm giới hạn. Ảnh 40 có thể được nhìn thấy dưới dạng độ sâu ảnh tương ứng với kết cấu ảnh 30.

Ảnh 40 bao gồm phần thứ nhất 401 bao gồm thông tin độ sâu của các phần tử (các điểm) của cảnh 3D mà có thể nhìn thấy được từ điểm nhìn thứ nhất và một hoặc nhiều phần thứ hai 402. Ảnh 40 có thể dành được theo cùng một cách với ảnh 30 nhưng chứa thông tin độ sâu được kết hợp với các điểm của cảnh 3D thay cho thông tin kết cấu như trong ảnh 30.

Phần thứ nhất 401 có thể được nhìn thấy dưới dạng bản vá độ sâu lớn thứ nhất (tương ứng với phần thứ nhất của cảnh 3D) và các phần thứ hai 402 bao gồm các bản vá kết cấu nhỏ hơn (tương ứng với các phần thứ hai của cảnh 3D mà bù cho phần thứ nhất).

Đối với việc kết xuất 3DoF của một phần của cảnh 3D, thì chỉ một điểm nhìn, ví dụ, điểm nhìn thứ nhất, sẽ được xem xét. Người dùng có thể quay đầu của họ theo ba cấp độ 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 thứ nhất. Các điểm của cảnh để được mã hóa là các điểm mà có thể nhìn thấy được từ điểm nhìn thứ nhất này, và chỉ thông tin kết cấu được cần để được mã hóa/được giải mã đối với việc kết xuất 3DoF. Không cần phải mã hóa các điểm của cảnh mà không thể nhìn thấy được từ điểm nhìn thứ nhất này vì người dùng không thể truy cập vào chúng bằng cách di chuyển điểm nhìn thứ nhất.

Về việc kết xuất 3DoF+, thì người dùng có thể di chuyển điểm nhìn trong không gian bị giới hạn xung quanh điểm nhìn, ví dụ, xung quanh điểm nhìn thứ nhất, không gian bị giới hạn xung quanh điểm nhìn tương ứng với không gian nhìn được kết hợp với các ảnh 30 và 40. Ví dụ, người dùng có thể di chuyển điểm nhìn của họ trong không

gian nhìn được xác định được định tâm trên điểm nhìn thứ nhất. Việc này cho phép trải nghiệm thị sai. Dữ liệu biểu diễn cho một phần của cảnh có thể nhìn thấy được từ điểm bất kỳ của không gian nhìn để được mã hóa thành luồng, gồm dữ liệu biểu diễn cho cảnh 3D có thể nhìn thấy được theo điểm nhìn thứ nhất (tức là, các phần thứ nhất 301 và 401). Kích thước và hình dạng của không gian nhìn có thể, ví dụ, được quyết định và được xác định tại bước mã hóa và được mã hóa trong luồng bit. Bộ giải mã có thể dành được thông tin này từ luồng bit và bộ kết xuất giới hạn không gian nhìn ở không gian được xác định bởi thông tin dành được. Theo một ví dụ khác, bộ kết xuất xác định không gian nhìn theo các ràng buộc về phần cứng, ví dụ, liên quan tới các khả năng của cảm biến (các cảm biến) mà phát hiện các chuyển động của người dùng. Trong trường hợp như vậy, tại pha mã hóa, nếu điểm có thể nhìn thấy được từ điểm trong không gian nhìn của bộ kết xuất đã không được mã hóa trong luồng 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 cho mọi điểm của cảnh 3D được mã hóa trong luồng mà không cần xem xét tới không gian nhìn kết xuất. Để tối ưu hóa kích thước của luồng, thì chỉ tập hợp con của các điểm của cảnh có thể được mã hóa, ví dụ, tập hợp con của các điểm mà có thể được nhìn thấy theo không gian nhìn kết xuất.

Về việc kết xuất 6DoF, người dùng có thể di chuyển điểm nhìn mọi nơi trong cảnh (tức là, theo 3 cách quay (xung quanh trục xoay ngang, trục chúc góc và trục nghiêng) và 3 cách dịch chuyển (dọc theo trục xoay ngang, trục chúc góc và trục nghiêng)). Đối với việc kết xuất 6DoF, mọi điểm (độ sâu và kết cấu) của cảnh phải được mã hóa trong luồng bit vì mọi điểm đều có thể truy cập được bởi người dùng mà có thể di chuyển điểm nhìn của họ. Tại giai đoạn mã hóa, không có cách thức nào để biết trước là từ điểm nhìn nào thì người dùng sẽ quan sát cảnh 3D 10.

Fig.5 thể hiện việc dành được các bản vá của một phần của cảnh 3D (ví dụ, một nhân vật trong số các nhân vật của cảnh 3D 10) mà có thể được đóng gói trên một hoặc nhiều tập bản đồ bản vá 54, 55, hoặc trên các ảnh 30 và 40, theo phương án không làm giới hạn. Một phần của cảnh 3D được nhìn thấy từ không gian nhìn đã cho. Đám mây điểm biểu diễn vật thể 3D 5 được phân chia thành nhiều phần 3D, ví dụ, 5, 10, 50, 100, 1000 hoặc nhiều phần 3D, 3 phần trong số các phần này được minh họa trên Fig.5, tức là, các phần 3D 52, 53 và 54, phần 3D 54 bao gồm các điểm của đám mây điểm biểu diễn một phần của đầu người, phần 3D 52 bao gồm các điểm của đám mây điểm biểu diễn nách người và phần 3D 53 bao gồm các điểm của đám mây điểm biểu diễn tay người. Một hoặc nhiều bản vá của mỗi phần 3D hoặc của một phần của các phần 3D được tạo ra để biểu diễn mỗi phần 3D theo hai chiều, tức là, theo sự tham số hóa 2D. Ví dụ, sự tham số hóa 2D 501 được dành cho phần 3D 51, sự tham số hóa 2D 502 được dành cho phần 3D 52 và 2 sự tham số hóa 2D khác nhau 503 và 504 có thể được dành cho phần 3D 53. Sự tham số hóa 2D có thể biến đổi từ phần 3D thành một phần khác. Ví dụ, sự tham số hóa 2D 501 được kết hợp với phần 3D là sự chiếu phối cảnh tuyến tính trong khi sự tham số hóa 2D 502 được kết hợp với phần 3D 52 là LLE và các sự tham số hóa 2D 503 và 504 được kết hợp với phần 3D 53 đều là các sự chiếu trực giao theo các điểm nhìn khác nhau. Theo một biến thể, toàn bộ các sự tham số hóa 2D được kết hợp với toàn bộ các phần 3D là cùng loại, ví dụ, sự chiếu phối cảnh tuyến tính hoặc sự chiếu trực giao. Theo một biến thể, các sự tham số hóa 2D khác nhau có thể được sử dụng cho cùng phần 3D.

Sự tham số hóa 2D được kết hợp với một phần 3D đã cho của đám mây điểm tương ứng với sự duyệt theo 2 chiều của phần 3D đã cho của đám mây điểm cho phép lấy mẫu phần 3D đã cho, tức là, sự biểu diễn 2D của nội dung (tức là, điểm (các điểm)) của phần 3D đã cho này bao gồm nhiều mẫu (mà có thể tương ứng với các điểm ảnh của hình ảnh

thứ nhất), số lượng của chúng vốn phụ thuộc vào bước lấy mẫu được áp dụng. Sự tham số hóa 2D có thể dành được theo nhiều cách khác nhau, ví dụ, bằng triển khai một phương pháp bất kỳ trong số các phương pháp sau đây:

- sự chiếu phối cảnh tuyến tính của các điểm của phần 3D của đám mây điểm lên trên mặt phẳng được kết hợp với điểm nhìn, các tham số biểu diễn cho sự chiếu phối cảnh tuyến tính bao gồm địa điểm của camera ảo, bước lấy mẫu theo không gian và trường nhìn theo 2 chiều;
- sự chiếu trực giao của các điểm của phần 3D của đám mây điểm lên trên bề mặt, các tham số biểu diễn cho sự chiếu trực giao bao gồm hình học (hình dạng, kích thước và hướng) của bề mặt chiếu và bước lấy mẫu theo không gian;
- Việc nhúng cục bộ tuyến tính (Locally-Linear Embedding, LLE) mà tương ứng với phép tính toán học giảm kích thước, được áp dụng ở đây để chuyển đổi/biến đổi từ 3D thành 2D, các tham số biểu diễn cho LLE bao gồm các hệ số biến đổi.

Mỗi bản vá có ưu điểm là hình dạng hình chữ nhật để làm dễ dàng quá trình đóng gói trên tập bản đồ bản vá 54, 55. Tập bản đồ bản vá 54 có thể là tập bản đồ bản vá hình học, tức là, ảnh của các điểm ảnh bao gồm các bản vá khác nhau 541, 542, 543 và 544 (mà có thể được nhìn thấy dưới dạng các mảng của các điểm ảnh, ví dụ), thông tin hình học dành được bởi sự chiếu/sự tham số hóa 2D của các điểm của phần 3D được kết hợp được kết hợp với mỗi điểm ảnh. Thông tin hình học có thể tương ứng với thông tin độ sâu hoặc thông tin về vị trí của các đỉnh của phần tử lưới. Tập bản đồ bản vá kết cấu tương ứng 55 bao gồm thông tin kết cấu được kết hợp với các phần 3D có thể dành được theo cùng một cách.

Thông tin ánh xạ, mà liên kết mỗi sự tham số hóa 2D với bản vá được kết hợp của nó trong tập bản đồ bản vá hình học và trong tập bản đồ bản vá kết cấu, sẽ có thể được

tạo ra. Thông tin ánh xạ có thể được tạo ra để giữ kết nối giữa sự tham số hóa 2D và bản vá hình học và bản vá kết cấu được kết hợp trong tập bản đồ bản vá hình học và tập bản đồ bản vá kết cấu theo cách tương ứng. Thông tin ánh xạ có thể, ví dụ, ở trong dạng:

{các tham số của sự tham số hóa 2D; bản vá hình học ID; bản vá kết cấu ID}

trong đó bản vá hình học ID có thể là giá trị kiểu số nguyên hoặc cặp giá trị bao gồm chỉ số cột U và chỉ số hàng V bản vá hình học thuộc về trong ma trận của các bản vá của tập bản đồ bản vá hình học; bản vá kết cấu ID có thể là giá trị kiểu số nguyên hoặc cặp giá trị bao gồm chỉ số cột U' và chỉ số hàng V' bản vá kết cấu thuộc về trong ma trận của các bản vá của tập bản đồ bản vá kết cấu.

Khi các bản vá hình học và các bản vá kết cấu được sắp xếp theo cùng sự sắp xếp trong tập bản đồ bản vá hình học và tập bản đồ bản vá kết cấu, thì bản vá hình học ID và bản vá kết cấu ID là giống nhau và thông tin ánh xạ có thể ở trong dạng, ví dụ:

{các tham số của sự tham số hóa 2D; bản vá hình học và kết cấu ID}

trong đó 'bản vá hình học và kết cấu ID' nhận dạng cả bản vá hình học trong tập bản đồ bản vá hình học và bản vá kết cấu trong tập bản đồ bản vá kết cấu, hoặc qua cùng giá trị kiểu số nguyên được kết hợp với cả bản vá hình học và bản vá kết cấu hoặc qua cặp giá trị chỉ số cột U và chỉ số hàng V bản vá hình học và các bản vá kết cấu thuộc về trong tập bản đồ bản vá hình học và tập bản đồ bản vá kết cấu theo cách tương ứng.

Thông tin ánh xạ giống nhau được tạo ra cho mỗi sự tham số hóa 2D và bản vá hình học và bản vá kết cấu được kết hợp. Thông tin ánh xạ như vậy cho phép tái tạo các phần tương ứng của cảnh 3D bằng cách thiết lập sự kết hợp của sự tham số hóa 2D với bản vá hình học và bản vá kết cấu tương ứng. Nếu sự tham số hóa 2D là sự chiếu, thì phần tương ứng của cảnh 3D có thể được tái tạo bằng cách bỏ chiếu (thực hiện sự chiếu

ngược) thông tin hình học được bao gồm trong bản vá hình học được kết hợp và thông tin kết cấu trong bản vá kết cấu được kết hợp. Sau đó, thông tin ánh xạ tương ứng với danh sách thông tin ánh xạ:

$\{\text{các tham số của sự tham số hóa 2D; bản vá hình học và kết cấu ID}\}_i$,

Đối với $i = 1$ đến n , với n là số lượng của các sự tham số hóa 2D.

Thông tin được bao gồm trong tập bản đồ bản vá 54, 55 cho phép việc kết xuất 3DoF+ của phần 3D của cảnh như được nhìn thấy từ không gian nhìn đã cho được kết hợp với một phần của cảnh 3D, tức là, người dùng xem tại một phần của cảnh 3D được kết xuất từ tập bản đồ bản vá có thể duyệt vào trong cảnh theo các điểm nhìn khác nhau được bao gồm trong không gian nhìn.

Fig.7 thể hiện tập hợp các nội dung video thể tích 70, 71, 72 và 73, theo phương án không làm giới hạn.

Mỗi nội dung video thể tích 70 đến 73 biểu diễn phần 3D của cảnh 3D theo phạm vi đã cho của các điểm nhìn, mỗi nội dung video thể tích 70 đến 73 cho phép việc kết xuất 3DoF+ của một phần của cảnh 3D được biểu diễn với mỗi nội dung video thể tích nói trên. Tập hợp các nội dung video thể tích 70 đến 73 cho phép, ví dụ, sự biểu diễn hoàn toàn của cảnh 3D và cho phép người dùng duyệt vào trong toàn bộ cảnh 3D. Theo một biến thể, chỉ một phần của cảnh 3D được biểu diễn với tập hợp các nội dung video thể tích. Ví dụ, đường dẫn 3D có thể được xác định trong cảnh và tập hợp các nội dung video thể tích cho phép người dùng duyệt vào trong cảnh 3D dọc theo đường dẫn 3D này.

Không gian 3D của cảnh 3D có thể được lấy mẫu thành nhiều điểm nhìn được xác định (ví dụ, các điểm nhìn $X=0$, $X=1$, $X=2$ và $X=3$ của đường dẫn 3D được minh họa

trên Fig.7) và nội dung video thể tích được tạo ra từ mỗi điểm nhìn trong số các điểm nhìn được xác định để dành được sự biểu diễn đầy đủ của cảnh 3D. Không gian nhìn hoặc phạm vi của các điểm nhìn được kết hợp với mỗi nội dung video thể tích, không gian nhìn của nội dung video thể tích đã cho được định tâm, ví dụ, xung quanh điểm nhìn mà nội dung video thể tích được tạo ra từ đó. Có việc chồng lấn về nội dung giữa hai nội dung video thể tích lân cận để đảm bảo việc duyệt mượt mà cho người dùng cuối. Ví dụ, các nội dung video thể tích 70 và 71 có thể chồng lấn một phần, các nội dung video thể tích 71 và 72 có thể chồng lấn một phần và các nội dung video thể tích 72 và 73 có thể chồng lấn một phần. Việc chồng lấn về nội dung của hai nội dung video thể tích có nghĩa là một số điểm nhìn được bao gồm trong các không gian nhìn được kết hợp với hai nội dung video thể tích lân cận là chung cho hai không gian nhìn.

Mỗi nội dung video thể tích 70 đến 73 được biểu diễn với tập hợp các bản vá 700, 701, 702 và 703, vốn bao gồm thông tin kết cấu và hình học. Tập hợp các bản vá có thể được sắp xếp trong các ảnh hoặc tập bản đồ bản vá như được mô tả liên quan tới Fig.3, Fig.4 và/hoặc Fig.5.

Fig.8 thể hiện quy trình cho việc kết xuất 6DoF của cảnh 3D hoặc một phần của nó từ tập hợp 18 của các nội dung video thể tích trên Fig.7, theo phương án không hạn chế.

Theo phương án trên Fig.8, người dùng cuối đang đeo thiết bị kết xuất nhập vai 80 (ví dụ, HMD) để duyệt vào trong sự biểu diễn video của cảnh 3D. Vị trí của người dùng cuối được xác định trong cảnh 3D (tức là, trong không gian của cảnh 3D), ví dụ, sử dụng bộ đo quán tính IMU (Inertial Measurement Unit, IMU) của HMD. HMD có thể bao gồm một hoặc nhiều màn hình hiển thị (ví dụ, màn hình tinh thể lỏng (Liquid Crystal Display, LCD), điốt phát sáng hữu cơ (Organic Light-Emitting Diode, OLED) hoặc silic

trên màn hình tinh thể lỏng (Liquid Crystal On Silicon, LCOS)) và cảm biến (các cảm biến) được tạo cấu hình để đo sự thay đổi (các sự thay đổi) của vị trí của HMD, ví dụ, các con quay hồi chuyển hoặc IMU (Inertial Measurement Unit), theo một, hai hoặc ba trục của thế giới thực (trục chúc góc, trục xoay ngang, và/hoặc trục nghiêng). Một phần của nội dung thể tích tương ứng với vị trí đo được của HMD có thể được xác định với chức năng cụ thể là thiết lập sự liên hệ giữa điểm nhìn được kết hợp với HMD trong thế giới thực và điểm nhìn của camera ảo được kết hợp với nội dung thể tích. Việc điều khiển một phần của nội dung thể tích để được hiển thị trên màn hình hiển thị (các màn hình hiển thị) của HMD theo vị trí đo được của HMD cho phép người dùng đeo HMD duyệt vào trong nội dung thể tích (và cảnh 3D), vốn rộng hơn so với trường nhìn được kết hợp với màn hình hiển thị (các màn hình hiển thị) của HMD. Theo một biến thể, thiết bị kết xuất nhập vai 80 là máy tính bảng có màn hình hiển thị chạm, người dùng duyệt vào trong nội dung bằng cách cuộn nội dung với một hoặc nhiều ngón tay trượt lên trên màn hình hiển thị chạm. Theo một biến thể khác, bộ kết xuất nhập vai 80 là bộ điều khiển được kết hợp với thiết bị hiển thị như là HMD hoặc máy tính bảng.

Thông tin biểu diễn cho điểm nhìn hiện tại của người dùng cuối sẽ được truyền, ví dụ, đến máy chủ hoặc thiết bị lưu trữ từ xa. Thông tin biểu diễn cho điểm nhìn tương ứng, ví dụ, với địa điểm của thiết bị kết xuất nhập vai trong thế giới thực hoặc điểm nhìn của thiết bị kết xuất nhập vai trong không gian của cảnh 3D.

Dựa trên thông tin nhận được biểu diễn cho điểm nhìn hiện tại, thì nội dung video thể tích 71 được lựa chọn trong số tập hợp 8 của các nội dung video thể tích 70 đến 73 được lưu trữ trên máy chủ hoặc thiết bị lưu trữ từ xa. Nội dung video thể tích 71 được lựa chọn, ví dụ, bằng cách so sánh điểm nhìn hiện tại nhận được với các không gian nhìn (hoặc các phạm vi của các điểm nhìn) được kết hợp với mỗi nội dung video thể tích của

tập hợp 8. Điểm nhìn hiện tại nhận được có thể, ví dụ, được so sánh với điểm nhìn tham chiếu (ví dụ, điểm nhìn trung tâm) của mỗi không gian trong số các không gian nhìn của các nội dung video thể tích và điểm nhìn tham chiếu mà gần nhất về mặt không gian so với điểm nhìn hiện tại nhận được sẽ được lựa chọn, với nội dung video thể tích tương ứng (ví dụ, nội dung video thể tích 71 và điểm nhìn tham chiếu được kết hợp $X=1$ theo phương án trên Fig.7 và Fig.8).

Nội dung video thể tích được lựa chọn 71 được truyền đến thiết bị kết xuất nhập vai 80. Dữ liệu (kết cấu và hình học) biểu diễn cho nội dung video thể tích 71, ví dụ, được mã hóa dưới dạng tập hợp các bản vá 701 được sắp xếp trong các ảnh của tập bản đồ bản vá. Dữ liệu được truyền và được nhận bởi thiết bị kết xuất nhập vai sẽ được giải mã bởi bộ giải mã 802. Sự nhìn của cảnh 3D tương ứng với điểm nhìn hiện tại sẽ được tổng hợp, ví dụ, bởi bộ tổng hợp 803, từ dữ liệu được giải mã (bằng cách bỏ chiếu các bản vá) và được kết xuất trong công nhìn người dùng cuối. Sau đó, sự nhìn được tổng hợp sẽ được truyền đến màn hình hiển thị của thiết bị kết xuất nhập vai (hoặc đến màn hình hiển thị được kết nối với thiết bị kết xuất nhập vai) để hiển thị.

Khi sự thay đổi của điểm nhìn của người dùng cuối xảy ra, thì thông tin biểu diễn cho điểm nhìn mới sẽ được truyền, và nội dung video thể tích tương ứng sẽ được lựa chọn (ví dụ, nội dung video thể tích 70). Tập hợp các bản vá 700 được sử dụng để biểu diễn nội dung video thể tích 70 được truyền đến thiết bị kết xuất nhập vai để giải mã 802, tổng hợp sự nhìn 803 và hiển thị.

Việc kết xuất của nội dung video thể tích đã cho được nhận từ máy chủ bởi thiết bị kết xuất nhập vai tương ứng với việc kết xuất 3DoF+ của một phần của cảnh được kết hợp với nội dung video thể tích đã cho. Việc kết xuất liên nhau của vài nội dung video thể tích của tập hợp 8 cho phép việc kết xuất 6DoF của cảnh 3D (hoặc một phần

của nó) trong khi giới hạn tốc độ bit mà được cần để phát luồng dữ liệu biểu diễn cho nội dung video thể tích đối với việc kết xuất 6DoF (chỉ một nội dung video thể tích biểu diễn một phần của cảnh 3D được cần để được phát luồng tại một thời điểm và không phải toàn bộ cảnh 3D).

Fig.9 thể hiện tập hợp các nội dung video thể tích 70, 91, 92 và 73, theo phương án không làm giới hạn khác.

Mỗi nội dung video thể tích 70, 91, 92 và 73 biểu diễn phần 3D của cảnh 3D theo phạm vi đã cho của các điểm nhìn, mỗi nội dung video thể tích 70, 91, 92 và 73 cho phép việc kết xuất 3DoF+ của một phần của cảnh 3D được biểu diễn với mỗi nội dung video thể tích nói trên 70, 91, 92 và 73. Tập hợp các nội dung video thể tích 70, 91, 92 và 73 cho phép, ví dụ, sự biểu diễn hoàn toàn của cảnh 3D và cho phép người dùng duyệt vào trong toàn bộ cảnh 3D. Theo một biến thể, chỉ một phần của cảnh 3D được biểu diễn với tập hợp các nội dung video thể tích. Ví dụ, đường dẫn 3D có thể được xác định trong cảnh và tập hợp các nội dung video thể tích cho phép người dùng duyệt vào trong cảnh 3D dọc theo đường dẫn 3D này.

Không gian 3D của cảnh 3D có thể được lấy mẫu thành nhiều điểm nhìn được xác định (ví dụ, các điểm nhìn $X=0$, $X=1$, $X=2$ và $X=3$ của đường dẫn 3D được minh họa trên Fig.9) và nội dung video thể tích được tạo ra từ mỗi điểm nhìn trong số các điểm nhìn được xác định để dành được sự biểu diễn đầy đủ của cảnh 3D. Không gian nhìn hoặc phạm vi của các điểm nhìn được kết hợp với mỗi nội dung video thể tích 70, 91, 92 và 73, không gian nhìn của nội dung video thể tích đã cho được định tâm, ví dụ, xung quanh điểm nhìn ($X=0$, $X=1$, $X=2$ và $X=3$) mà nội dung video thể tích được tạo ra từ đó. Có việc chồng lấn về nội dung giữa hai nội dung video thể tích lân cận để đảm bảo việc duyệt mượt mà cho người dùng cuối. Ví dụ, các nội dung video thể tích 70 và 91

có thể chồng lẫn một phần, các nội dung video thể tích 91 và 92 có thể chồng lẫn một phần và các nội dung video thể tích 92 và 73 có thể chồng lẫn một phần. Việc chồng lẫn về nội dung của hai nội dung video thể tích có nghĩa là một số điểm nhìn được bao gồm trong các không gian nhìn được kết hợp với hai nội dung video thể tích lân cận là chung cho hai không gian nhìn.

Mỗi nội dung video thể tích 70, 91, 92 và 73 được biểu diễn với tập hợp các bản vá, vốn bao gồm thông tin kết cấu và hình học. Tập hợp các bản vá có thể được sắp xếp trong các ảnh hoặc tập bản đồ bản vá như được mô tả liên quan tới Fig.3, Fig.4 và/hoặc Fig.5.

Để giảm các sự đòi hỏi bộ nhớ để lưu trữ các nội dung video thể tích khác nhau 70, 91, 92 và 73, thì các bản vá của tập hợp của nội dung video thể tích đã cho được so sánh với các bản vá của các nội dung video thể tích khác, ví dụ, các bản vá của tập hợp 901 được so sánh với các bản vá của các tập hợp 700, 902 và 703 của các nội dung video thể tích theo cách tương ứng 70, 92 và 73 để phát hiện phần dư nội dung giữa các bản vá. Thật vậy, vì các nội dung video thể tích 70, 91, 92 và 73 chồng lẫn một phần, nên một số bản vá của các nội dung video thể tích nói trên có thể có cùng nội dung, ít nhất là một phần. Theo một biến thể, các bản vá của tập hợp 901 được so sánh với chỉ một phần của các nội dung video thể tích khác, ví dụ, chỉ với các tập hợp 700 và 703 của các nội dung video thể tích theo cách tương ứng 70 và 73. Theo biến thể này, một số nội dung trong số các nội dung video thể tích có thể được xác định là các tham chiếu, tức là, các bản vá của nội dung video thể tích khác có thể tham chiếu đến các bản vá của các nội dung video thể tích tham chiếu nhưng các bản vá của nội dung video thể tích tham chiếu do không tham chiếu đến các bản bất kỳ của các nội dung video thể tích khác. Theo phương án cụ thể trên Fig.9, các nội dung video thể tích 70 và 73 có thể được xem

xét như là nội dung video thể tích tham chiếu trong khi các nội dung video thể tích khác 91 và 92 không được xem xét như vậy (mỗi nội dung trong số các nội dung video thể tích 91 và 92 bao gồm ít nhất bản vá mà tham chiếu đến một hoặc nhiều bản vá của các nội dung video thể tích tham chiếu 70 và 73).

Để giảm dung lượng bộ nhớ, thì các bản vá mà đã được nhận dạng là bao gồm nội dung (kết cấu và/hoặc hình học) mà giống hệt với bản vá của một nội dung video thể tích khác không chứa thông tin kết cấu và/hoặc hình học mà chỉ đơn thuần là thông tin nhận dạng bản vá khác mà gồm cùng nội dung (kết cấu và/hoặc hình học). Ví dụ, bản vá 9021 của tập hợp 902 đã được nhận dạng là bao gồm cùng nội dung với bản vá 7001 của tập hợp 700. Bản vá 9022 của tập hợp 902 đã được nhận dạng là bao gồm cùng nội dung với bản vá 7032 của tập hợp 703. Bản vá 9021 bao gồm thông tin nhận dạng bản vá 7001 mà tham chiếu đến (thay cho dữ liệu biểu diễn cho các thuộc tính, kết cấu hoặc hình học, của một phần của cảnh 3D mà nó biểu diễn), thông tin như vậy được lập mã chỉ với một ít bit trong khi việc lập mã dữ liệu biểu diễn cho các thuộc tính sẽ đòi hỏi số lượng bit lớn hơn nhiều. Theo cùng một cách, bản vá 9022 bao gồm thông tin nhận dạng bản vá 7032 mà nó tham chiếu đến.

Theo một phương án biến thể, bản vá 9021 và/hoặc 9022 có thể tham chiếu đến chỉ khu vực (phần phụ) của bản vá 7001 và/hoặc 7032, theo cách tương ứng, khi chỉ khu vực nói trên của bản vá 7001 và/hoặc 7032 bao gồm nội dung của bản vá 9021 và/hoặc 9022. Theo phương án biến thể này, bản vá 9021 bao gồm thông tin nhận dạng bản vá 7001 mà nó tham chiếu đến với thông tin nhận dạng khu vực cụ thể của bản vá 7001 mà bao gồm cùng nội dung với bản vá 9021. Khu vực cụ thể có thể, ví dụ, được nhận dạng với các tọa độ của điểm ảnh tham chiếu trong bản vá 7001 (ví dụ, các số hàng và cột)

với chiều cao và chiều rộng (ví dụ, mỗi khu vực được trình bày với một số lượng các điểm ảnh).

Theo phương án khác, bản vá tham chiếu đến một bản vá khác sẽ không bao gồm thông tin nhận dạng bản vá mà nó tham chiếu đến. Theo phương án khác này, danh sách thiết lập việc ánh xạ giữa các bản vá sẽ được kết hợp với mỗi nội dung video thể tích, danh sách nói trên còn bao gồm thông tin chỉ ra liệu bản vá có tham chiếu đến một bản vá hay không (ví dụ, với cờ 1-bit, cờ lấy giá trị '0' khi bản vá không tham chiếu đến một bản vá khác và giá trị '1' khi bản vá tham chiếu đến một bản vá khác, hoặc theo cách ngược lại). Danh sách có thể, ví dụ, có dạng bảng tra cứu (Look-Up-table, LUT), ví dụ:

ID bản vá	Tham chiếu (Có/Không)	ID bản vá tham chiếu
0000	0	
0001	1	XXXXXXX
0002	0	
0003	0	
...		
0111	1	YYYYYYY

Bảng 1: danh sách ánh xạ

Fig.10 thể hiện quy trình cho việc kết xuất 6DoF của cảnh 3D hoặc một phần của nó từ tập hợp 10 của các nội dung video thể tích trên Fig.9, theo phương án không hạn chế.

Theo phương án trên Fig.10, người dùng cuối đang đeo thiết bị kết xuất nhập vai 100 (ví dụ, HMD) để duyệt vào trong sự biểu diễn video của cảnh 3D. Vị trí của người dùng cuối được xác định trong cảnh 3D (tức là, trong không gian của cảnh 3D), ví dụ,

sử dụng IMU (Inertial Measurement Unit) của HMD. HMD có thể bao gồm một hoặc nhiều màn hình hiển thị (ví dụ, LCD (Liquid Crystal Display), OLED (Organic Light-Emitting Diode) hoặc LCOS (Liquid Crystal On Silicon)) và cảm biến (các cảm biến) được tạo cấu hình để đo sự thay đổi (các sự thay đổi) của vị trí của HMD, ví dụ, các con quay hồi chuyển hoặc IMU (Inertial Measurement Unit), theo một, hai hoặc ba trục của thế giới thực (trục chúc góc, trục xoay ngang, và/hoặc trục nghiêng). Một phần của nội dung thể tích tương ứng với vị trí đo được của HMD có thể được xác định với chức năng cụ thể là thiết lập sự liên hệ giữa điểm nhìn được kết hợp với HMD trong thế giới thực và điểm nhìn của camera ảo được kết hợp với nội dung thể tích. Việc điều khiển một phần của nội dung thể tích để được hiển thị trên màn hình hiển thị (các màn hình hiển thị) của HMD theo vị trí đo được của HMD cho phép người dùng đeo HMD duyệt vào trong nội dung thể tích (và cảnh 3D), vốn rộng hơn so với trường nhìn được kết hợp với màn hình hiển thị (các màn hình hiển thị) của HMD. Theo một biến thể, thiết bị kết xuất nhập vai 100 là máy tính bảng có màn hình hiển thị chạm, người dùng duyệt vào trong nội dung bằng cách cuộn nội dung với một hoặc nhiều ngón tay trượt lên trên màn hình hiển thị chạm. Theo một biến thể khác, bộ kết xuất nhập vai 100 là bộ điều khiển được kết hợp với thiết bị hiển thị như là HMD hoặc máy tính bảng.

Thông tin biểu diễn cho điểm nhìn hiện tại của người dùng cuối được truyền, ví dụ, đến máy chủ hoặc thiết bị lưu trữ từ xa. Thông tin biểu diễn cho điểm nhìn tương ứng, ví dụ, với địa điểm của thiết bị kết xuất nhập vai trong thế giới thực hoặc điểm nhìn của thiết bị kết xuất nhập vai trong không gian của cảnh 3D.

Dựa trên thông tin nhận được biểu diễn cho điểm nhìn hiện tại, thì nội dung video thể tích thứ nhất 92 được lựa chọn trong số tập hợp 10 của các nội dung video thể tích 70, 91, 92 và 73 được lưu trữ trên máy chủ hoặc thiết bị lưu trữ từ xa. Nội dung video

thể tích thứ nhất 92 được xác định, ví dụ, bằng cách so sánh điểm nhìn hiện tại nhận được với các không gian nhìn (hoặc các phạm vi của các điểm nhìn) được kết hợp với mỗi nội dung video thể tích của tập hợp 10. Điểm nhìn hiện tại nhận được có thể, ví dụ, được so sánh với điểm nhìn tham chiếu (ví dụ, điểm nhìn trung tâm) của mỗi không gian trong số các không gian nhìn của các nội dung video thể tích và điểm nhìn tham chiếu mà gần nhất về mặt không gian so với điểm nhìn hiện tại nhận được sẽ được lựa chọn, với nội dung video thể tích tương ứng (ví dụ, nội dung video thể tích 92 và điểm nhìn tham chiếu được kết hợp $X=2$ theo phương án trên Fig.9 và Fig.10).

Nội dung video thể tích thứ nhất 92 được truyền đến thiết bị kết xuất nhập vai 100. Dữ liệu (kết cấu và hình học) biểu diễn cho nội dung video thể tích thứ nhất 92, ví dụ, được mã hóa dưới dạng tập hợp các bản vá thứ nhất 92 được sắp xếp trong các ảnh của tập bản đồ bản vá. Ít nhất bản vá thứ nhất 9021, 9022 tham chiếu đến bản vá thứ hai 7001, 7032 của nội dung video thể tích thứ hai 70, 73. Dữ liệu được truyền và được nhận bởi thiết bị kết xuất nhập vai sẽ được giải mã bởi bộ giải mã 102. Dữ liệu dành được bằng cách giải mã thông tin được bao gồm trong các bản vá thứ nhất của tập hợp 902 của các bản vá. Một số bản vá trong số các bản vá thứ nhất của tập hợp 902 bao gồm các thuộc tính được kết hợp với các điểm của phần 3D của cảnh 3D được biểu diễn với nội dung video thể tích thứ nhất 92. Một hoặc nhiều bản vá thứ nhất 9021, 9022 của tập hợp 902 không bao gồm các thuộc tính mà chỉ đơn thuần là thông tin chỉ ra các bản vá thứ hai nào hoặc khu vực cụ thể nào của các bản vá thứ hai 7001, 7032 mà các bản vá thứ nhất 9021, 9022 tham chiếu đến.

Nội dung video thể tích thứ hai 70 và 73, vốn bao gồm các bản vá thứ hai 7001, 7032 các bản vá thứ nhất 9021, 9022 có thể được truyền với nội dung video thể tích thứ nhất 92. Theo một biến thể, chỉ các bản vá thứ hai của các ô dạng lợp ngói mà bao gồm

bản vá thứ hai 7001 và 7032 sẽ được truyền, khi các nội dung video thể tích của tập hợp 10 được chia nhỏ thành các ô dạng lợp ngói ((ô dạng lợp ngói tương ứng với khu vực phụ của khung và, ví dụ, được định rõ trong tiêu chuẩn mã hóa video hiệu quả cao (High Efficiency Video Coding, HEVC) hoặc trong VP9 hoặc AV1 của Google với từ viết tắt (các từ viết tắt) khác nhau)), thì một phần của cảnh 3D được kết hợp với mỗi ô dạng lợp ngói.

Theo một biến thể, nội dung video thể tích thứ hai 70 và 73 (hoặc các ô dạng lợp ngói bao gồm các bản vá thứ hai mà bản vá thứ nhất tham chiếu đến) được truyền đến thiết bị kết xuất nhập vai 100 sau việc truyền (bởi thiết bị kết xuất nhập vai) của yêu cầu mà yêu cầu việc truyền của các bản vá thứ hai hoặc các nội dung video thể tích thứ hai được kết hợp 70, 73.

Theo biến thể khác, các bản vá thứ nhất tham chiếu đến các bản vá thứ hai của nội dung video thể tích thứ hai 70, 73 sẽ không được truyền bởi máy chủ đến thiết bị kết xuất nhập vai 100. Theo biến thể này, máy chủ có thể phân tích nội dung video thể tích thứ nhất 92 hoặc danh sách ánh xạ được kết hợp (như bảng 1, ví dụ) để phát hiện liệu nội dung video thể tích thứ nhất 92 có bao gồm bản vá (các bản vá) thứ nhất tham chiếu đến bản vá (các bản vá) thứ hai của nội dung (các nội dung) video thể tích khác hay không (được gọi là nội dung (các nội dung) video thể tích thứ hai). Khi nó được xác định rằng nội dung video thể tích thứ nhất 92 bao gồm một hoặc nhiều bản vá (các bản vá) thứ nhất tham chiếu đến một hoặc nhiều bản vá (các bản vá) thứ hai, thì máy chủ có thể lập lịch một cách tự động việc truyền của nội dung video thể tích thứ nhất 92 (hoặc ô dạng lợp ngói (các ô dạng lợp ngói) của nội dung video thể tích thứ nhất 92) và nội dung video thể tích thứ hai 70, 73 (hoặc ô dạng lợp ngói (các ô dạng lợp ngói) của nội dung video thể tích thứ hai 70, 73) đến thiết bị kết xuất nhập vai 100.

Theo biến thể khác, khi nội dung video thể tích thứ hai đã được nhận bởi thiết bị kết xuất nhập vai (ví dụ, khi tổng hợp các sự nhìn trước đó của cảnh tương ứng với các điểm nhìn của nội dung video thể tích thứ hai 70, 73), thì nội dung video thể tích thứ hai không được truyền lại và thiết bị kết xuất nhập vai 100 có thể truy xuất dữ liệu từ các bản vá thứ hai mà bản vá thứ nhất của nội dung video thể tích thứ nhất 92 tham chiếu đến.

Theo một biến thể khác, để tránh các vấn đề về độ trễ và sự chậm trễ trong việc tổng hợp các sự nhìn, thì điểm nhìn tương lai theo sau điểm nhìn hiện tại có thể được xác định bởi thiết bị kết xuất nhập vai 100 hoặc máy chủ bao gồm tập hợp các nội dung video thể tích (ví dụ, từ các điểm nhìn trước đó bởi quỹ đạo dịch chuyển được xác định của các điểm nhìn trong cảnh 3D). Nội dung video thể tích tiềm năng thứ nhất và thứ hai có thể được truyền trước một cách tự động đến thiết bị kết xuất nhập vai mà có thể quyết định tổng hợp sự nhìn nếu điểm nhìn của người dùng cuối thực sự tương ứng với một điểm nhìn được xác định trước.

Sự nhìn của cảnh 3D tương ứng với điểm nhìn hiện tại sẽ được tổng hợp, ví dụ, bởi bộ tổng hợp 103, từ dữ liệu được giải mã dành được từ các bản vá thứ nhất và thứ hai (bằng cách bỏ chiếu các bản vá hoặc bằng cách áp dụng hoạt động ngược với hoạt động được thực hiện để dành được sự tham số hóa 2D) và được kết xuất trong công nhìn người dùng cuối. Sự nhìn được tổng hợp sẽ được truyền đến màn hình hiển thị của thiết bị kết xuất nhập vai (hoặc đến màn hình hiển thị được kết nối với thiết bị kết xuất nhập vai) để hiển thị.

Khi sự thay đổi của điểm nhìn của người dùng cuối xảy ra, thì thông tin biểu diễn cho điểm nhìn mới sẽ được truyền, và các hoạt động được mô tả ở trên đây được lặp lại.

Việc kết xuất của nội dung video thể tích đã cho được nhận từ máy chủ bởi thiết bị kết xuất nhập vai tương ứng với việc kết xuất 3DoF+ của một phần của cảnh được kết hợp với nội dung video thể tích đã cho. Việc kết xuất liên nhau của vài nội dung video thể tích của tập hợp 8 cho phép việc kết xuất 6DoF của cảnh 3D (hoặc một phần của nó) trong khi giới hạn tốc độ bit mà được cần để phát luồng dữ liệu biểu diễn cho nội dung video thể tích đối với việc kết xuất 6DoF (chỉ một nội dung video thể tích biểu diễn một phần của cảnh 3D được cần để được phát luồng tại một thời điểm và không phải toàn bộ cảnh 3D).

Fig.6 thể hiện ví dụ không giới hạn về việc mã hóa, việc truyền, việc giải mã và việc kết xuất của dữ liệu biểu diễn cho cảnh 3D trong định dạng mà có thể, đồng thời, tương thích cho việc kết xuất 3DoF, 3DoF+ và 6DoF.

Ảnh hoặc tập bản đồ bản vá của cảnh 3D 60 (hoặc chuỗi của các ảnh của cảnh 3D) được mã hóa trong luồng 62 bởi bộ mã hóa 61. Luồng 62 bao gồm phần tử thứ nhất của cú pháp mang dữ liệu biểu diễn cho cảnh 3D đối với việc kết xuất 3DoF (dữ liệu của phần tử thứ nhất của ảnh 30, ví dụ) và ít nhất phần tử thứ hai của cú pháp mang dữ liệu biểu diễn cho cảnh 3D đối với việc kết xuất 3DoF+ (ví dụ, dữ liệu của các phần tử thứ hai của ảnh 30 và ảnh 40 hoặc tập bản đồ bản vá 54, 55). Bộ mã hóa có thể mã hóa vài ảnh hoặc tập bản đồ bản vá tương ứng với các điểm nhìn khác nhau (và không gian nhìn được kết hợp) của cảnh (mỗi điểm nhìn tương ứng với nội dung video thể tích như được mô tả liên quan tới các hình từ Fig.7 đến Fig.10).

Bộ mã hóa 61, ví dụ, phù hợp với bộ mã hóa như là:

- JPEG, specification ISO/CEI 10918-1 UIT-T Recommendation T.81, <https://www.itu.int/rec/T-REC-T.81/en>;

- AVC, cũng được gọi là MPEG-4 AVC hoặc h264. Được đặc tả 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 (bản mô tả của nó được tìm thấy tại trang web ITU, phần khuyến nghị T, se-ri H, h265, <http://www.itu.int/rec/T-REC-H.265-201612-I/en>);
- 3D-HEVC (phần mở rộng của HEVC mà bản mô tả của nó được tìm thấy tại trang web ITU, phần khuyến nghị T, se-ri H, h265, <http://www.itu.int/rec/T-REC-H.265-201612-I/en> annex G và I);
- VP9 được phát triển bởi Google; hoặc
- AV1 (AOMedia Video 1) được phát triển bởi Alliance for Open Media.

Bộ giải mã 63 dành được luồng 62 từ nguồn. Ví dụ, nguồn thuộc về tập hợp bao gồm:

- bộ nhớ cục bộ, ví dụ, bộ nhớ video hoặc RAM (hoặc bộ nhớ truy cập ngẫu nhiên (Random-Access Memory)), bộ nhớ tốc độ nhanh, ROM (hoặc bộ nhớ chỉ đọc (Read Only Memory)), đĩa cứng;
- giao diện lưu trữ, ví dụ, giao diện với bộ lưu trữ khối, RAM, bộ nhớ tốc độ nhanh, ROM, đĩa quang hoặc bộ hỗ trợ từ;
- 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 giao diện đồ họa người - máy cho phép người dùng nhập dữ liệu.

Bộ giải mã 63 giải mã phần tử thứ nhất của cú pháp của luồng 62 đối với việc kết xuất 3DoF 64. Đối với việc kết xuất 3DoF+ 65, thì bộ giải mã sẽ giải mã cả phần tử thứ nhất của cú pháp và phần tử thứ hai của cú pháp của luồng 62. Đối với việc kết xuất 6DoF, thì bộ giải mã sẽ giải mã phần tử thứ nhất của cú pháp và phần tử thứ hai của cú pháp tương ứng với vài nội dung video thể tích.

Bộ giải mã 63 phù hợp với bộ mã hóa 61, ví dụ, phù hợp với bộ giải mã như là:

- JPEG;
- AVC;
- HEVC;
- 3D-HEVC (phần mở rộng của HEVC);
- VP9; hoặc
- AV1.

Fig.11 thể hiện ví dụ không làm giới hạn về phương án của cú pháp của luồng mang dữ liệu biểu diễn cho cảnh 3D khi dữ liệu được truyền trên giao thức truyền dựa trên gói. Fig.11 thể hiện cấu trúc ví dụ 11 của luồng video. Cấu trúc gồm trong bộ chứa mà tổ chức luồng trong các phần tử độc lập của cú pháp. Cấu trúc có thể bao gồm phần đầu 111, vốn là tập hợp của dữ liệu chung cho mọi phần tử cú pháp của luồng. Ví dụ, phần đầu bao gồm siêu dữ liệu về các phần tử cú pháp, mô tả bản chất và vai trò của mỗi phần tử cú pháp trong số chúng. Phần đầu cũng có thể bao gồm các tọa độ của điểm nhìn được sử dụng cho việc mã hóa của ảnh 30, 40 (hoặc tập bản đồ bản vá 54, 55) và thông tin về kích thước và độ phân giải của ảnh. Cấu trúc bao gồm phần tải bao gồm phần tử thứ nhất của cú pháp 112 và ít nhất một phần tử thứ hai của cú pháp 113. Phần tử cú pháp thứ nhất 112 có thể bao gồm dữ liệu biểu diễn cho nội dung video thể tích

thứ nhất và một hoặc nhiều phần tử cú pháp thứ hai 113 có thể bao gồm dữ liệu biểu diễn cho một hoặc nhiều nội dung video thể tích thứ hai.

Nhằm mục đích minh họa, trong ngữ cảnh của tiêu chuẩn định dạng tệp ISOBMF, thì các bản vá kết cấu, các bản vá hình học và siêu dữ liệu sẽ thường được tham chiếu trong các rãnh từ ISOBMFF hộp loại moov, với dữ liệu kết cấu và dữ liệu hình học tự nhúng trong hộp phương tiện-dữ liệu loại mdat.

Fig.12 minh họa sơ đồ khối của ví dụ về hệ thống mà các khía cạnh và các phương án khác nhau được triển khai trong đó. Hệ thống 1200 có thể được biểu hiện dưới dạng thiết bị gồm các thành phần khác nhau được mô tả dưới đây và được tạo cấu hình để thực hiện một hoặc nhiều khía cạnh trong số các khía cạnh được mô tả trong sáng chế này. Các ví dụ về các thiết bị như vậy, gồm, nhưng không bị giới hạn ở, các thiết bị điện tử như là các máy tính cá nhân, các máy tính xách tay, các điện thoại thông minh, các máy tính bảng, các đầu thu giải mã tín hiệu truyền thông đa phương tiện kỹ thuật số, các bộ nhận tivi kỹ thuật số, các hệ thống quay video cá nhân, các thiết bị gia dụng được kết nối, và các máy chủ. Các phần tử của hệ thống 1200, theo kiểu đơn lẻ hoặc tổ hợp, có thể được biểu hiện trong mạch tích hợp đơn, nhiều IC, và/hoặc các thành phần rời rạc. Ví dụ, trong ít nhất một phương án, các phần tử xử lý và bộ mã hóa/bộ giải mã của hệ thống 1200 được phân bổ qua nhiều IC và/hoặc các thành phần rời rạc. Trong các phương án khác nhau, hệ thống 1200 được ghép nối theo kiểu truyền thông với các hệ thống tương tự khác, hoặc với các thiết bị điện tử khác, qua, ví dụ, bus truyền thông hoặc thông qua các cổng đầu vào và/hoặc đầu ra chuyên dụng. Trong các phương án khác nhau, hệ thống 1200 được tạo cấu hình để triển khai một hoặc nhiều khía cạnh trong số các khía cạnh được mô tả trong sáng chế này.

Hệ thống 1200 gồm ít nhất một bộ xử lý 1210 được tạo cấu hình để thực thi các lệnh được tải trong đó để triển khai, ví dụ, các khía cạnh khác nhau được mô tả trong sáng chế này. Bộ xử lý 1210 có thể gồm bộ nhớ nhúng, giao diện đầu vào đầu ra, và các hệ mạch khác nhau như đã biết trong tình trạng kỹ thuật. Hệ thống 1200 gồm ít nhất một bộ nhớ 1220 (ví dụ, thiết bị nhớ khả biến, và/hoặc thiết bị nhớ bất khả biến). Hệ thống 1200 gồm thiết bị lưu trữ 1240, vốn có thể gồm bộ nhớ bất khả biến và/hoặc bộ nhớ khả biến, gồm, nhưng không bị giới hạn ở, EEPROM, ROM, PROM, RAM, DRAM, SRAM, tác động nhanh, ổ đĩa từ, và/hoặc ổ đĩa quang. Thiết bị lưu trữ 1240 có thể gồm thiết bị lưu trữ trong, thiết bị lưu trữ được đính kèm, và/hoặc thiết bị lưu trữ có thể truy cập được vào mạng, như các ví dụ không làm giới hạn.

Hệ thống 1200 gồm môđun bộ mã hóa/bộ giải mã 1230 được tạo cấu hình, ví dụ, để xử lý dữ liệu để cung cấp video được mã hóa hoặc video được giải mã, và môđun bộ mã hóa/bộ giải mã 1230 có thể gồm bộ xử lý và bộ nhớ riêng của nó. Môđun bộ mã hóa/bộ giải mã 1230 biểu diễn môđun (các môđun) mà có thể được gồm trong thiết bị để thực hiện các chức năng mã hóa và/hoặc giải mã. Như đã biết, thiết bị có thể gồm một hoặc cả hai môđun trong số các môđun mã hóa và giải mã. Theo cách tùy chọn, môđun bộ mã hóa/bộ giải mã 1230 có thể được triển khai dưới dạng phần tử tách rời của hệ thống 1200 hoặc có thể được hợp nhất trong bộ xử lý 1210 như là tổ hợp của phần cứng và phần mềm như đã biết đến bởi những người có hiểu biết trung bình trong lĩnh vực kỹ thuật tương ứng.

Mã chương trình để được tải lên bộ xử lý 1210 hoặc bộ mã hóa/bộ giải mã 1230 để thực hiện các khía cạnh khác nhau được mô tả trong sáng chế này có thể được lưu trữ trong thiết bị lưu trữ 1240 và tiếp đó được tải lên bộ nhớ 1220 cho việc thực thi bởi bộ xử lý 1210. Theo các phương án khác nhau, một hoặc nhiều thành phần trong số bộ

xử lý 1210, bộ nhớ 1220, thiết bị lưu trữ 1240, và môđun bộ mã hóa/bộ giải mã 1230 có thể lưu trữ một hoặc nhiều mục trong số các mục khác nhau trong quá trình thực hiện các quy trình được mô tả trong sáng chế này. Các mục được lưu trữ như vậy có thể gồm, nhưng không bị giới hạn ở, video đầu vào, video được giải mã hoặc các bộ phận của video được giải mã, luồng bit, các ma trận, các biến, và các kết quả trung gian hoặc cuối cùng từ việc xử lý các phương trình, các công thức, các phép toán, và logic tính toán.

Trong vài phương án, bộ nhớ nằm trong bộ xử lý 1210 và/hoặc môđun bộ mã hóa/bộ giải mã 1230 được sử dụng để lưu trữ các lệnh và để cung cấp bộ nhớ làm việc để xử lý mà được cần trong quá trình mã hóa hoặc giải mã.

Tuy nhiên, trong các phương án khác, bộ nhớ nằm ngoài thiết bị xử lý (ví dụ, thiết bị xử lý có thể hoặc là bộ xử lý 1210 hoặc là môđun bộ mã hóa/bộ giải mã 1230) được sử dụng cho một hoặc nhiều chức năng trong số các chức năng này. Bộ nhớ ngoài có thể là bộ nhớ 1220 và/hoặc thiết bị lưu trữ 1240, ví dụ, bộ nhớ khả biến động và/hoặc bộ nhớ tốc độ nhanh bất khả biến. Trong vài phương án, bộ nhớ tốc độ nhanh bất khả biến ngoài được sử dụng để lưu trữ hệ thống hoạt động của tivi. Trong ít nhất một phương án, bộ nhớ khả biến động ngoài nhanh như là RAM được sử dụng làm bộ nhớ làm việc cho các hoạt động lập mã và giải mã video, như là đối với MPEG-2, HEVC, hoặc tạo mã video đa năng (Versatile Video Coding, VVC).

Đầu vào của các phần tử của hệ thống 1200 có thể được cung cấp thông qua các thiết bị đầu vào khác nhau được chỉ ra trong khối 1130. Các thiết bị đầu vào như vậy gồm, nhưng không bị giới hạn ở, (i) bộ phận RF mà nhận tín hiệu RF được truyền, ví dụ, qua không gian bởi bộ phát rộng, (ii) thiết bị đầu cuối đầu vào hỗn hợp, (iii) thiết bị đầu cuối đầu vào USB, và/hoặc (iv) thiết bị đầu cuối đầu vào HDMI.

Trong các phương án khác nhau, các thiết bị đầu vào của khối 1130 có các phần tử xử lý đầu vào tương ứng được kết hợp như đã biết trong tình trạng kỹ thuật. Ví dụ, bộ phận RF có thể được kết hợp với các phần tử cần thiết để (i) lựa chọn tần số mong muốn (cũng được gọi là việc lựa chọn tín hiệu, hoặc giới hạn dải tín hiệu ở dải tần số), (ii) chuyển đổi xuống tín hiệu được lựa chọn, (iii) giới hạn dải một lần nữa ở dải tần số hẹp hơn, để lựa chọn (ví dụ) dải tần số tín hiệu vốn có thể được gọi là kênh trong các phương án nhất định, (iv) giải điều biến tín hiệu được chuyển đổi xuống và được giới hạn dải, (v) thực hiện việc sửa lỗi, và (vi) phân kênh để lựa chọn luồng mong muốn của các gói dữ liệu. Bộ phận RF của các phương án khác nhau gồm một hoặc nhiều phần tử để thực hiện các chức năng này, ví dụ, các bộ lựa chọn tần số, các bộ lựa chọn tín hiệu, các bộ giới hạn dải, các bộ lựa chọn kênh, các bộ lọc, các bộ chuyển đổi xuống, các bộ giải điều biến, các bộ sửa lỗi, và các bộ phân kênh. Bộ phận RF có thể gồm bộ điều hướng mà thực hiện các chức năng khác nhau trong số các chức năng, gồm, ví dụ, chuyển đổi xuống tín hiệu nhận được xuống tần số thấp hơn (ví dụ, tần số trung gian hoặc tần số gần dải tần cơ sở) hoặc xuống dải tần cơ sở.

Trong một phương án bộ giải mã tín hiệu truyền hình, thì bộ phận RF và phần tử xử lý đầu vào được kết hợp của nó nhận tín hiệu RF được truyền trên phương tiện có dây (ví dụ, cáp), và thực hiện sự lựa chọn tần số bằng cách lọc, chuyển đổi xuống, và lọc một lần nữa ở dải tần số mong muốn.

Các phương án khác nhau sắp xếp lại thứ tự của các phần tử được mô tả ở trên (và các phần tử khác), loại bỏ một số phần tử trong số các phần tử này, và/hoặc bổ sung các phần tử khác thực hiện các chức năng tương tự hoặc khác nhau.

Việc bổ sung các phần tử có thể gồm việc chèn các phần tử vào giữa các phần tử sẵn có, như là, ví dụ, chèn các bộ khuếch đại và bộ chuyển đổi từ kỹ thuật tương tự sang kỹ thuật số. Trong các phương án khác nhau, bộ phận RF gồm ăngten.

Theo cách tùy chọn, các thiết bị đầu cuối USB và/hoặc HDMI có thể gồm các bộ xử lý giao diện tương ứng để kết nối hệ thống 1200 với các thiết bị điện tử khác qua các kết nối USB và/hoặc HDMI. Cần hiểu rằng các khía cạnh khác nhau của việc xử lý đầu vào, ví dụ, việc sửa lỗi Reed-Solomon, có thể được triển khai, ví dụ, trong IC xử lý đầu vào tách rời hoặc trong bộ xử lý 1210 nếu cần thiết. Theo cách tương tự, các khía cạnh của việc xử lý giao diện USB hoặc HDMI có thể được triển khai trong các IC giao diện tách rời hoặc trong bộ xử lý 1210 nếu cần thiết. Luồng được giải điều biến, được sửa lỗi và được phân kênh sẽ được cung cấp cho các phần tử xử lý khác nhau, gồm, ví dụ, bộ xử lý 1210, và bộ mã hóa/bộ giải mã 1230 hoạt động trong sự tổ hợp với bộ nhớ và các phần tử lưu trữ để xử lý luồng dữ liệu nếu cần thiết cho việc giới thiệu trên thiết bị đầu ra.

Các phần tử khác nhau của hệ thống 1200 có thể được cung cấp trong phần vỏ được tích hợp. Trong phần vỏ được tích hợp, các phần tử khác nhau có thể được nối kết và truyền dữ liệu giữa chúng nhờ sử dụng sự sắp xếp kết nối phù hợp, ví dụ, bus trong như đã biết trong tình trạng kỹ thuật, gồm bus I2C, nối dây, và các bo mạch in.

Hệ thống 1200 gồm giao diện truyền thông 1250 mà cho phép truyền thông với các thiết bị khác qua kênh truyền thông 1260. Giao diện truyền thông 1250 có thể gồm, nhưng không bị giới hạn ở, bộ thu phát được tạo cấu hình để truyền và nhận dữ liệu trên kênh truyền thông 1260. Giao diện truyền thông 1250 có thể gồm, nhưng không bị giới hạn ở, modem hoặc thẻ mạng và kênh truyền thông 1260 có thể được triển khai, ví dụ, trong phương tiện có dây và/hoặc không dây.

Dữ liệu được phát luồng đến hệ thống 1200, trong các phương án khác nhau, nhờ sử dụng mạng Wi-Fi như là IEEE 802.11. Tín hiệu Wi-Fi của các phương án này được nhận trên kênh truyền thông 1260 và giao diện truyền thông 1250 vốn được làm thích ứng cho truyền thông Wi-Fi. Kênh truyền thông 1260 của các phương án này thường được kết nối với điểm truy cập hoặc bộ định tuyến mà cung cấp truy cập vào các mạng bên ngoài gồm Internet cho phép các ứng dụng phát luồng và truyền thông tốc độ cao qua Internet khác.

Các phương án khác cung cấp dữ liệu được phát luồng cho hệ thống 1200 nhờ sử dụng bộ giải mã tín hiệu truyền hình mà truyền tải dữ liệu trên kết nối HDMI của khối đầu vào 1130.

Các phương án khác nữa cung cấp dữ liệu được phát luồng cho hệ thống 1200 nhờ sử dụng kết nối RF của khối đầu vào 1130.

Hệ thống 1200 có thể cung cấp tín hiệu đầu ra cho các thiết bị đầu ra khác nhau, gồm bộ hiển thị 1100, các loa 1110, và các thiết bị ngoại vi khác 1120. Các thiết bị ngoại vi khác 1120 gồm, trong các ví dụ khác nhau của các phương án, một hoặc nhiều thành phần trong số DVR độc lập, máy chơi nhạc, hệ thống stereo, hệ thống ánh sáng, và các thiết bị khác mà cung cấp chức năng dựa trên đầu ra của của hệ thống 1200.

Trong các phương án khác nhau, các điều khiển tín hiệu được truyền thông giữa hệ thống 1200 và bộ hiển thị 1100, các loa 1110, hoặc các thiết bị ngoại vi khác 1120 nhờ sử dụng việc phát tín hiệu như là AV.Link, CEC, hoặc các giao thức truyền thông khác mà cho phép việc điều khiển từ thiết bị đến thiết bị mà có hoặc không có sự can thiệp của người dùng.

Các thiết bị đầu ra có thể được ghép nối theo kiểu truyền thông với hệ thống 1200 qua chuyên dụng các kết nối thông qua các giao diện tương ứng 1270, 1280, và 1290.

Theo cách thay thế, các thiết bị đầu ra có thể được kết nối với hệ thống 1200 nhờ sử dụng kênh truyền thông 1260 qua giao diện truyền thông 1250. Bộ hiển thị 1100 và các loa 1110 có thể được tích hợp trong một bộ đơn với các thành phần khác của hệ thống 1200 trong thiết bị điện tử như là, ví dụ, tivi.

Trong các phương án khác nhau, bộ hiển thị giao diện 1270 gồm trình điều khiển hiển thị, như là, ví dụ, chip điều khiển định thời (Timing controller, T Con).

Bộ hiển thị 1100 và loa 1110 có thể, theo cách thay thế, được tách rời với một hoặc nhiều thành phần trong số các thành phần khác, ví dụ, nếu bộ phận RF của đầu vào 1130 là một phần của bộ giải mã tín hiệu truyền hình tách rời. Trong các phương án khác nhau mà trong số bộ hiển thị 1100 và các loa 1110 là các thành phần ngoài, thì tín hiệu đầu vào có thể được cung cấp qua các kết nối đầu ra chuyên dụng, gồm, ví dụ, các cổng HDMI, các cổng USB, hoặc các đầu ra COMP.

Fig.13 minh họa phương pháp truyền dữ liệu biểu diễn cho cảnh 3D, ví dụ, cảnh 3D 10, theo phương án không hạn chế. Phương pháp có thể, ví dụ, được triển khai trong bộ mã hóa 61 và/hoặc trong thiết bị 1200. Các tham số khác nhau của thiết bị 1200 có thể được cung cấp. Cảnh 3D có thể, ví dụ, dành được từ nguồn, một hoặc nhiều điểm nhìn có thể được xác định trong không gian của cảnh 3D, các tham số được kết hợp với ánh xạ (các ánh xạ) chiếu có thể được khởi tạo.

Trong hoạt động thứ nhất 131, thông tin biểu diễn cho điểm nhìn trong cảnh 3D sẽ được nhận. Cảnh 3D được biểu diễn với tập hợp các nội dung video thể tích, không gian nhìn hoặc phạm vi của các điểm nhìn được kết hợp với mỗi nội dung video thể tích.

Trong hoạt động thứ hai 132, dữ liệu biểu diễn cho nội dung video thể tích thứ nhất của tập hợp các nội dung video thể tích sẽ được truyền. Nội dung video thể tích thứ nhất đang theo phạm vi của các điểm nhìn mà bao gồm điểm nhìn. Nội dung video thể

tích thứ nhất được biểu diễn với tập hợp các bản vá thứ nhất, mỗi bản vá thứ nhất của tập hợp tương ứng với sự tham số hóa hai chiều của nhóm thứ nhất của các điểm được bao gồm trong phần 3D của cảnh 3D được kết hợp với nội dung video thể tích thứ nhất. Ít nhất bản vá thứ nhất của tập hợp các bản vá thứ nhất tham chiếu đến bản vá thứ hai (hoặc khu vực của bản vá thứ hai), bản vá thứ hai tương ứng với sự tham số hóa hai chiều của nhóm thứ hai của các điểm được bao gồm trong một phần 3D khác của cảnh 3D được kết hợp với nội dung video thể tích thứ hai của tập hợp các nội dung video thể tích. Nội dung video thể tích thứ hai là khác với nội dung video thể tích thứ nhất, phạm vi của các điểm nhìn của nội dung video thể tích thứ nhất là khác với phạm vi của các điểm nhìn của nội dung video thể tích thứ hai, ngay cả nếu chúng có thể chồng lấn một phần với nhau.

Theo phương án cụ thể, bản vá (các bản vá) thứ hai mà một hoặc nhiều bản vá thứ nhất tham chiếu đến sẽ được truyền với các bản vá thứ nhất. Theo một phương án biến thể, các bản vá thứ hai đã được truyền trong hoạt động trước đó và được lưu trữ trong bộ nhớ của thiết bị 1200.

Theo phương án cụ thể khác, bản vá (các bản vá) thứ hai được truyền lúc nhận được yêu cầu mà yêu cầu việc truyền của bản vá (các bản vá) thứ hai.

Theo một phương án cụ thể khác, các thuộc tính (ví dụ, kết cấu và/hoặc hình học/độ sâu) được kết hợp với các điểm của ít nhất nhóm thứ nhất của các điểm được kết hợp với ít nhất bản vá thứ nhất tham chiếu đến ít nhất bản vá thứ hai sẽ được bao gồm trong ít nhất bản vá thứ hai. Ít nhất bản vá thứ nhất bao gồm thông tin nhận dạng ít nhất bản vá thứ hai (hoặc khu vực của ít nhất bản vá thứ hai).

Theo phương án khác, các thuộc tính được kết hợp với các điểm của ít nhất nhóm thứ nhất của các điểm được kết hợp với ít nhất bản vá thứ nhất tham chiếu đến ít nhất

bản và thứ hai sẽ được bao gồm trong ít nhất bản và thứ hai. Thông tin ánh xạ biểu diễn cho việc tham chiếu của ít nhất bản và thứ nhất đến ít nhất bản và thứ hai sẽ được kết hợp với nội dung video thể tích thứ nhất. Thông tin ánh xạ có thể, ví dụ, được bao gồm trong bảng được kết hợp với nội dung thể tích thứ nhất.

Fig.14 minh họa phương pháp nhận dữ liệu biểu diễn cho cảnh 3D, ví dụ, cảnh 3D 10, theo phương án không hạn chế. Phương pháp có thể, ví dụ, được triển khai trong bộ giải mã 63 và/hoặc trong thiết bị kết xuất nhập vai 100 và/hoặc trong thiết bị 1200. Các tham số khác nhau của thiết bị 1200 có thể được cung cấp. Dữ liệu có thể, ví dụ, dành được từ nguồn, ví dụ, bộ nhớ của luồng bit.

Trong hoạt động thứ nhất 141, thông tin biểu diễn cho điểm nhìn trong cảnh 3D sẽ được truyền. Cảnh 3D được biểu diễn với tập hợp các nội dung video thể tích, không gian nhìn hoặc phạm vi của các điểm nhìn được kết hợp với mỗi nội dung video thể tích.

Trong hoạt động thứ hai 142, dữ liệu biểu diễn cho nội dung video thể tích thứ nhất của tập hợp các nội dung video thể tích được nhận, ví dụ, từ luồng bit hoặc bộ nhớ cục bộ hoặc thiết bị nhớ từ xa. Nội dung video thể tích thứ nhất đang theo phạm vi của các điểm nhìn mà Nội dung video thể tích thứ nhất được biểu diễn với tập hợp các bản và thứ nhất, mỗi bản và thứ nhất của tập hợp tương ứng với sự tham số hóa hai chiều của nhóm thứ nhất của các điểm được bao gồm trong phần 3D của cảnh 3D được kết hợp với nội dung video thể tích thứ nhất. Ít nhất bản và thứ nhất của tập hợp các bản và thứ nhất tham chiếu đến bản và thứ hai (hoặc khu vực của bản và thứ hai), bản và thứ hai tương ứng với sự tham số hóa hai chiều của nhóm thứ hai của các điểm được bao gồm trong một phần 3D khác của cảnh 3D được kết hợp với nội dung video thể tích thứ hai của tập hợp các nội dung video thể tích. Nội dung video thể tích thứ hai là khác với nội dung video thể tích thứ nhất, phạm vi của các điểm nhìn của nội dung video thể tích

thứ nhất là khác với phạm vi của các điểm nhìn của nội dung video thể tích thứ hai, ngay cả nếu chúng có thể chồng lấn một phần với nhau.

Theo phương án cụ thể, bản vá (các bản vá) thứ hai mà một hoặc nhiều bản vá thứ nhất tham chiếu đến sẽ được nhận với các bản vá thứ nhất. Theo một phương án biến thể, các bản vá thứ hai đã được nhận trong hoạt động trước đó và được truy xuất từ bộ lưu trữ (ví dụ, bộ nhớ của thiết bị 1200).

Theo phương án cụ thể khác, yêu cầu mà yêu cầu việc truyền của bản vá (các bản vá) thứ hai sẽ được truyền.

Theo một phương án cụ thể khác, các thuộc tính (ví dụ, kết cấu và/hoặc hình học/độ sâu) được kết hợp với các điểm của ít nhất nhóm thứ nhất của các điểm được kết hợp với ít nhất bản vá thứ nhất tham chiếu đến ít nhất bản vá thứ hai sẽ được bao gồm trong ít nhất bản vá thứ hai. Ít nhất bản vá thứ nhất bao gồm thông tin nhận dạng ít nhất bản vá thứ hai (hoặc khu vực của ít nhất bản vá thứ hai).

Theo phương án khác, các thuộc tính được kết hợp với các điểm của ít nhất nhóm thứ nhất của các điểm được kết hợp với ít nhất bản vá thứ nhất tham chiếu đến ít nhất bản vá thứ hai sẽ được bao gồm trong ít nhất bản vá thứ hai. Thông tin ánh xạ biểu diễn cho việc tham chiếu của ít nhất bản vá thứ nhất đến ít nhất bản vá thứ hai sẽ được kết hợp với nội dung video thể tích thứ nhất. Thông tin ánh xạ có thể, ví dụ, được bao gồm trong bảng được kết hợp với nội dung thể tích thứ nhất.

Theo phương án tùy chọn, một hoặc nhiều sự nhìn của cảnh được kết xuất từ dữ liệu nhận được. Sự nhìn (các sự nhìn) được kết xuất có thể được truyền để hiển thị.

Đương nhiên, sáng chế không bị giới hạn ở các phương án được mô tả trước đó.

Cụ thể, sáng chế không bị giới hạn ở phương pháp và thiết bị truyền/nhận dữ liệu biểu diễn cho cảnh 3D mà còn mở rộng thành phương pháp mã hóa/giải mã dữ liệu và/hoặc phương pháp tạo ra luồng bit bao gồm dữ liệu được mã hóa và thành thiết bị triển khai phương pháp này và nhất là các thiết bị bao gồm ít nhất một CPU và/hoặc ít nhất một GPU.

Sáng chế cũng đề cập đến phương pháp (và thiết bị được tạo cấu hình) để hiển thị các hình ảnh được kết xuất từ dữ liệu được giải mã của luồng bit.

Sáng chế cũng đề cập đến phương pháp (và thiết bị được tạo cấu hình) để truyền và/hoặc nhận luồng bit.

Các cách triển khai được mô tả ở đây có thể được triển khai trong, ví dụ, phương pháp hoặc quy trình, thiết bị, sản phẩm chương trình máy tính, luồng dữ liệu, hoặc tín hiệu. Ngay cả nếu chỉ được thảo luận trong ngữ cảnh của dạng triển khai đơn (ví dụ, chỉ được thảo luận như là phương pháp hoặc thiết bị), cách triển khai của các dấu hiệu được thảo luận cũng có thể được triển khai trong các dạng khác (ví dụ, chương trình). Thiết bị có thể được triển khai trong, ví dụ, phần cứng, phần mềm, và phần sụn thích hợp. Các phương pháp có thể được triển khai trong, ví dụ, thiết bị như là, ví dụ, bộ xử lý, mà tham chiếu đến các thiết bị xử lý nói chung, gồm, ví dụ, máy tính, bộ vi xử lý, mạch tích hợp, hoặc thiết bị logic lập trình được. Các bộ xử lý cũng gồm các thiết bị truyền thông, như là, ví dụ, các điện thoại thông minh, máy tính bảng, máy tính, điện thoại di động, thiết bị trợ lý kỹ thuật số cầm tay/cá nhân (Portable/Personal Digital Assistant, PDA), và các thiết bị khác mà tạo điều kiện cho truyền thông của thông tin giữa các người dùng cuối.

Các cách triển khai của các quy trình và các dấu hiệu khác nhau được mô tả ở đây có thể được biểu hiện trong nhiều thiết bị hoặc ứng dụng khác nhau, cụ thể là, ví dụ, thiết bị hoặc ứng dụng được kết hợp với việc mã hóa dữ liệu, việc giải mã dữ liệu, tạo

chế độ xem, việc xử lý kết cấu, và các việc xử lý hình ảnh khác và thông tin kết cấu và/hoặc thông tin độ sâu liên quan. Các ví dụ về thiết bị như vậy gồm bộ mã hóa, bộ giải mã, bộ xử lý sau sẽ xử lý đầu ra từ bộ giải mã, bộ xử lý trước sẽ cung cấp đầu vào cho bộ mã hóa, bộ lập mã, bộ giải mã video, bộ mã hóa-giải mã (codec) video, máy chủ web, bộ giải mã tín hiệu truyền hình, máy tính xách tay, máy tính cá nhân, điện thoại di động, PDA, và các thiết bị truyền thông khác. Rõ ràng là, thiết bị có thể là di động và thậm chí là được cài đặt trong xe di động.

Theo cách tùy chọn, các phương pháp có thể được triển khai bởi các lệnh được thực hiện bởi bộ xử lý, và các lệnh như vậy (và/hoặc các giá trị dữ liệu được tạo thành bởi cách triển khai) có thể được lưu trữ trên phương tiện đọc được bởi bộ xử lý như là, ví dụ, mạch tích hợp, bộ mang phần mềm hoặc thiết bị lưu trữ khác như là, ví dụ, đĩa cứng, đĩa compact (Compact Diskette, CD), đĩa quang (như là, ví dụ, DVD, thường được gọi là đĩa đa năng kỹ thuật số, hoặc đĩa video kỹ thuật số), bộ nhớ truy cập ngẫu nhiên (Random Access Memory, RAM), hoặc bộ nhớ chỉ đọc (Read-Only Memory, ROM). Các lệnh có thể tạo ra chương trình ứng dụng được biểu hiện một cách hữu hình trên phương tiện đọc được bởi bộ xử lý. Các lệnh có thể là, ví dụ, ở dạng phần cứng, phần mềm, hoặc tổ hợp của chúng. Các lệnh có thể được tìm thấy trong, ví dụ, hệ thống hoạt động, ứng dụng tách rời, hoặc tổ hợp của chúng. Do đó, bộ xử lý có thể được đặc trưng bởi, ví dụ, cả thiết bị được tạo cấu hình để tiến hành quy trình và thiết bị mà gồm phương tiện đọc được bởi bộ xử lý (như là thiết bị lưu trữ) có các lệnh để tiến hành quy trình. Hơn nữa, phương tiện đọc được bởi bộ xử lý có thể lưu trữ, bên cạnh hoặc thay cho các lệnh, các giá trị dữ liệu được sản xuất bởi cách triển khai.

Người có hiểu biết trung bình trong lĩnh vực kỹ thuật tương ứng có thể nhận ra một cách hiển nhiên rằng, các cách triển khai có thể được sản xuất nhiều tín hiệu được

tạo định dạng để mang thông tin mà có thể, ví dụ, được lưu trữ hoặc được truyền. Thông tin có thể gồm, ví dụ, các lệnh để thực hiện phương pháp, hoặc dữ liệu được sản xuất bởi một cách triển khai trong số các cách triển khai được mô tả. Ví dụ, tín hiệu có thể được tạo định dạng để mang dưới dạng dữ liệu các quy tắc để ghi hoặc đọc cú pháp của phương án được mô tả, hoặc để mang dưới dạng dữ liệu các cú pháp-giá trị được ghi bởi phương án được mô tả. Tín hiệu như vậy có thể được tạo định dạng, ví dụ, là sóng điện từ (ví dụ, sử dụng bộ phận tần số radio của phổ) hoặc là tín hiệu dải tần cơ sở. Việc tạo định dạng có thể gồm, ví dụ, việc mã hoá luồng dữ liệu và việc điều biến bộ mang với luồng dữ liệu được mã hóa. Thông tin mà tín hiệu mang có thể là, ví dụ, thông tin kỹ thuật tương tự hoặc kỹ thuật số. Tín hiệu có thể được truyền trên nhiều liên kết có dây hoặc không dây khác nhau, như đã biết. Tín hiệu có thể được lưu trữ trên phương tiện đọc được bởi bộ xử lý.

Một số cách triển khai đã được mô tả. Tuy nhiên, cần hiểu rằng có thể thực hiện nhiều sự sửa đổi khác nhau. Ví dụ, các phần tử của cách triển khai khác nhau có thể được tổ hợp, bổ sung, sửa đổi, hoặc loại bỏ để sản xuất các cách triển khai khác. Theo cách tùy chọn, một người có hiểu biết trung bình trong lĩnh vực kỹ thuật tương ứng sẽ hiểu rằng các cấu trúc và các quy trình khác có thể được thay thế cho các cấu trúc và các quy trình được bộc lộ và các cách triển khai kết quả sẽ thực hiện ít nhất về cơ bản là cùng chức năng (các chức năng), ít nhất về cơ bản là cùng cách (các cách), để đạt được ít nhất về cơ bản là kết quả (các kết quả) giống như các cách triển khai được bộc lộ. Theo đó, các cách triển khai này và các cách triển khai khác đều nằm trong phạm vi của sáng chế.

YÊU CẦU BẢO HỘ

1. Phương pháp để xử lý nội dung video, phương pháp bao gồm các bước:

nhận thông tin biểu diễn cho điểm nhìn thứ nhất trong cảnh ba chiều (Three-Dimensional, 3D), cảnh 3D nói trên được biểu diễn với tập hợp các nội dung video thể tích;

truyền nội dung video thể tích thứ nhất của tập hợp nói trên được lựa chọn theo điểm nhìn thứ nhất, nội dung video thể tích thứ nhất nói trên được xác định theo phạm vi của các điểm nhìn bao gồm điểm nhìn nói trên, nội dung video thể tích thứ nhất nói trên được biểu diễn với tập hợp các bản vá thứ nhất, bản vá thứ nhất của tập hợp tương ứng với sự tham số hóa hai chiều của nhóm thứ nhất của các điểm được bao gồm trong phần 3D của cảnh 3D được kết hợp với nội dung video thể tích thứ nhất nói trên.

2. Phương pháp theo điểm 1, trong đó nội dung video thể tích thứ hai được truyền lúc nhận được điểm nhìn thứ hai khác với điểm nhìn thứ nhất nói trên, nội dung video thể tích thứ hai được xác định theo phạm vi của các điểm nhìn bao gồm điểm nhìn thứ hai nói trên, nội dung video thể tích thứ hai nói trên được biểu diễn với tập hợp các bản vá thứ hai, mỗi bản vá thứ hai của tập hợp các bản vá thứ hai tương ứng với sự tham số hóa hai chiều của nhóm thứ hai của các điểm được bao gồm trong phần 3D của cảnh 3D được kết hợp với nội dung video thể tích thứ hai.

3. Phương pháp theo điểm 2, trong đó phạm vi của các điểm nhìn được kết hợp với nội dung video thể tích thứ nhất và phạm vi của các điểm nhìn được kết hợp với nội dung video thể tích thứ hai chồng lấn một phần.

4. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 3, trong đó ít nhất bản vá thứ nhất của tập hợp nói trên của các bản vá thứ nhất tham chiếu đến ít nhất khu vực của

ít nhất bản vá thứ ba tương ứng với sự tham số hóa hai chiều của nhóm thứ ba của các điểm được bao gồm trong một phần 3D khác của cảnh 3D được kết hợp với nội dung video thể tích thứ ba của tập hợp các nội dung video thể tích.

5. Phương pháp theo điểm 4, trong đó các thuộc tính được kết hợp với các điểm của ít nhất nhóm thứ nhất của các điểm được kết hợp với ít nhất bản vá thứ nhất nói trên tham chiếu đến ít nhất khu vực của ít nhất bản vá thứ ba nói trên được bao gồm trong ít nhất khu vực nói trên của ít nhất bản vá thứ ba nói trên, ít nhất bản vá thứ nhất nói trên bao gồm thông tin nhận dạng ít nhất khu vực nói trên của ít nhất bản vá thứ ba nói trên.

6. Phương pháp theo điểm 4, trong đó các thuộc tính được kết hợp với các điểm của ít nhất nhóm thứ nhất của các điểm được kết hợp với ít nhất bản vá thứ nhất nói trên tham chiếu đến ít nhất khu vực của ít nhất bản vá thứ ba nói trên được bao gồm trong ít nhất bản vá thứ ba nói trên, thông tin ánh xạ biểu diễn cho việc tham chiếu của ít nhất bản vá thứ nhất nói trên đến ít nhất khu vực nói trên của ít nhất bản vá thứ ba nói trên được kết hợp với nội dung video thể tích thứ nhất nói trên.

7. Thiết bị để xử lý nội dung video, thiết bị bao gồm: 2

bộ nhận được tạo cấu hình để nhận thông tin biểu diễn cho điểm nhìn thứ nhất trong cảnh 3D, cảnh 3D nói trên được biểu diễn với tập hợp các nội dung video thể tích;

bộ truyền được tạo cấu hình để truyền nội dung video thể tích thứ nhất của tập hợp nói trên được lựa chọn theo điểm nhìn thứ nhất, nội dung video thể tích thứ nhất nói trên được xác định theo phạm vi của các điểm nhìn bao gồm điểm nhìn nói trên, nội dung video thể tích thứ nhất nói trên được biểu diễn với tập hợp các bản vá thứ nhất, bản vá thứ nhất của tập hợp tương ứng với sự tham số hóa hai chiều của nhóm thứ nhất của các điểm được bao gồm trong phần 3D của cảnh 3D được kết hợp với nội dung video thể tích thứ nhất nói trên.

8. Thiết bị theo điểm 7, trong đó nội dung video thể tích thứ hai được truyền lúc nhận được điểm nhìn thứ hai khác với điểm nhìn thứ nhất nói trên, nội dung video thể tích thứ hai được xác định theo phạm vi của các điểm nhìn bao gồm điểm nhìn thứ hai nói trên, nội dung video thể tích thứ hai nói trên được biểu diễn với tập hợp các bản vá thứ hai, mỗi bản vá thứ hai của tập hợp các bản vá thứ hai tương ứng với sự tham số hóa hai chiều của nhóm thứ hai của các điểm được bao gồm trong phần 3D của cảnh 3D được kết hợp với nội dung video thể tích thứ hai.

9. Thiết bị theo điểm 8, trong đó phạm vi của các điểm nhìn được kết hợp với nội dung video thể tích thứ nhất và phạm vi của các điểm nhìn được kết hợp với nội dung video thể tích thứ hai chồng lấn một phần.

10. Thiết bị theo điểm bất kỳ trong số các điểm từ 7 đến 9, trong đó ít nhất bản vá thứ nhất của tập hợp nói trên của các bản vá thứ nhất tham chiếu đến ít nhất khu vực của ít nhất bản vá thứ ba tương ứng với sự tham số hóa hai chiều của nhóm thứ ba của các điểm được bao gồm trong một phần 3D khác của cảnh 3D được kết hợp với nội dung video thể tích thứ ba của tập hợp các nội dung video thể tích.

11. Thiết bị theo điểm 10, trong đó các thuộc tính được kết hợp với các điểm của ít nhất nhóm thứ nhất của các điểm được kết hợp với ít nhất bản vá thứ nhất nói trên tham chiếu đến ít nhất khu vực của ít nhất bản vá thứ ba nói trên được bao gồm trong ít nhất khu vực nói trên của ít nhất bản vá thứ ba nói trên, ít nhất bản vá thứ nhất nói trên bao gồm thông tin nhận dạng ít nhất khu vực nói trên của ít nhất bản vá thứ ba nói trên.

12. Thiết bị theo điểm 10, trong đó các thuộc tính được kết hợp với các điểm của ít nhất nhóm thứ nhất của các điểm được kết hợp với ít nhất bản vá thứ nhất nói trên tham chiếu đến ít nhất khu vực của ít nhất bản vá thứ ba nói trên được bao gồm trong ít nhất bản vá thứ ba nói trên, thông tin ánh xạ biểu diễn cho việc tham chiếu của ít nhất bản vá thứ

nhất nói trên đến ít nhất khu vực nói trên của ít nhất bản vá thứ ba nói trên được kết hợp với nội dung video thể tích thứ nhất nói trên.

13. Phương pháp để xử lý nội dung video, phương pháp bao gồm các bước:

truyền thông tin biểu diễn cho điểm nhìn thứ nhất trong cảnh 3D, cảnh 3D nói trên được biểu diễn với tập hợp các nội dung video thể tích;

nhận nội dung video thể tích thứ nhất của tập hợp nói trên được lựa chọn theo điểm nhìn thứ nhất, nội dung video thể tích thứ nhất nói trên đang theo phạm vi của các điểm nhìn bao gồm điểm nhìn thứ nhất, nội dung video thể tích thứ nhất nói trên được biểu diễn với tập hợp các bản vá thứ nhất, bản vá thứ nhất của tập hợp tương ứng với sự tham số hóa hai chiều của nhóm thứ nhất của các điểm được bao gồm trong phần 3D của cảnh 3D được kết hợp với nội dung video thể tích thứ nhất nói trên.

14. Phương pháp theo điểm 13, trong đó nội dung video thể tích thứ hai được nhận sau khi việc truyền của điểm nhìn thứ hai khác với điểm nhìn thứ nhất nói trên, nội dung video thể tích thứ hai được xác định theo phạm vi của các điểm nhìn bao gồm điểm nhìn thứ hai nói trên, nội dung video thể tích thứ hai nói trên được biểu diễn với tập hợp các bản vá thứ hai, bản vá thứ hai của tập hợp tương ứng với sự tham số hóa hai chiều của nhóm thứ hai của các điểm được bao gồm trong phần 3D của cảnh 3D được kết hợp với nội dung video thể tích thứ hai nói trên.

15. Phương pháp theo điểm 14, trong đó ít nhất bản vá thứ nhất của tập hợp nói trên của các bản vá thứ nhất tham chiếu đến ít nhất khu vực của ít nhất bản vá thứ ba tương ứng với sự tham số hóa hai chiều của nhóm thứ ba của các điểm được bao gồm trong một phần 3D khác của cảnh 3D được kết hợp với nội dung video thể tích thứ ba của tập hợp nói trên của các nội dung video thể tích.

16. Phương pháp theo điểm 15, trong đó các thuộc tính được kết hợp với các điểm của ít nhất nhóm thứ nhất của các điểm được kết hợp với ít nhất bản vá thứ nhất nói trên tham chiếu đến ít nhất khu vực của ít nhất bản vá thứ ba nói trên được bao gồm trong ít nhất khu vực nói trên của ít nhất bản vá thứ ba nói trên, ít nhất bản vá thứ nhất nói trên bao gồm thông tin nhận dạng ít nhất khu vực nói trên của ít nhất bản vá thứ ba nói trên.

17. Phương pháp theo điểm 15, trong đó các thuộc tính được kết hợp với các điểm của ít nhất nhóm thứ nhất của các điểm được kết hợp với ít nhất bản vá thứ nhất nói trên tham chiếu đến ít nhất khu vực của ít nhất bản vá thứ ba nói trên được bao gồm trong ít nhất bản vá thứ ba nói trên, thông tin ánh xạ biểu diễn cho việc tham chiếu của ít nhất bản vá thứ nhất nói trên đến ít nhất khu vực nói trên của ít nhất bản vá thứ ba nói trên được kết hợp với nội dung video thể tích thứ nhất nói trên.

18. Thiết bị để xử lý nội dung video, thiết bị bao gồm:

bộ truyền được tạo cấu hình để truyền thông tin biểu diễn cho điểm nhìn thứ nhất trong cảnh 3D, cảnh 3D nói trên được biểu diễn với tập hợp các nội dung video thể tích;

bộ nhận được tạo cấu hình để nhận nội dung video thể tích thứ nhất của tập hợp nói trên được lựa chọn theo điểm nhìn thứ nhất, nội dung video thể tích thứ nhất nói trên đang theo phạm vi của các điểm nhìn bao gồm điểm nhìn thứ nhất, nội dung video thể tích thứ nhất nói trên được biểu diễn với tập hợp các bản vá thứ nhất, bản vá thứ nhất của tập hợp tương ứng với sự tham số hóa hai chiều của nhóm thứ nhất của các điểm được bao gồm trong phần 3D của cảnh 3D được kết hợp với nội dung video thể tích thứ nhất nói trên.

19. Thiết bị theo điểm 18, trong đó nội dung video thể tích thứ hai được nhận sau khi việc truyền của điểm nhìn thứ hai khác với điểm nhìn thứ nhất nói trên, nội dung video thể tích thứ hai được xác định theo phạm vi của các điểm nhìn bao gồm điểm nhìn thứ

hai nói trên, nội dung video thể tích thứ hai nói trên được biểu diễn với tập hợp các bản vá thứ hai, bản vá thứ hai của tập hợp tương ứng với sự tham số hóa hai chiều của nhóm thứ hai của các điểm được bao gồm trong phần 3D của cảnh 3D được kết hợp với nội dung video thể tích thứ hai nói trên.

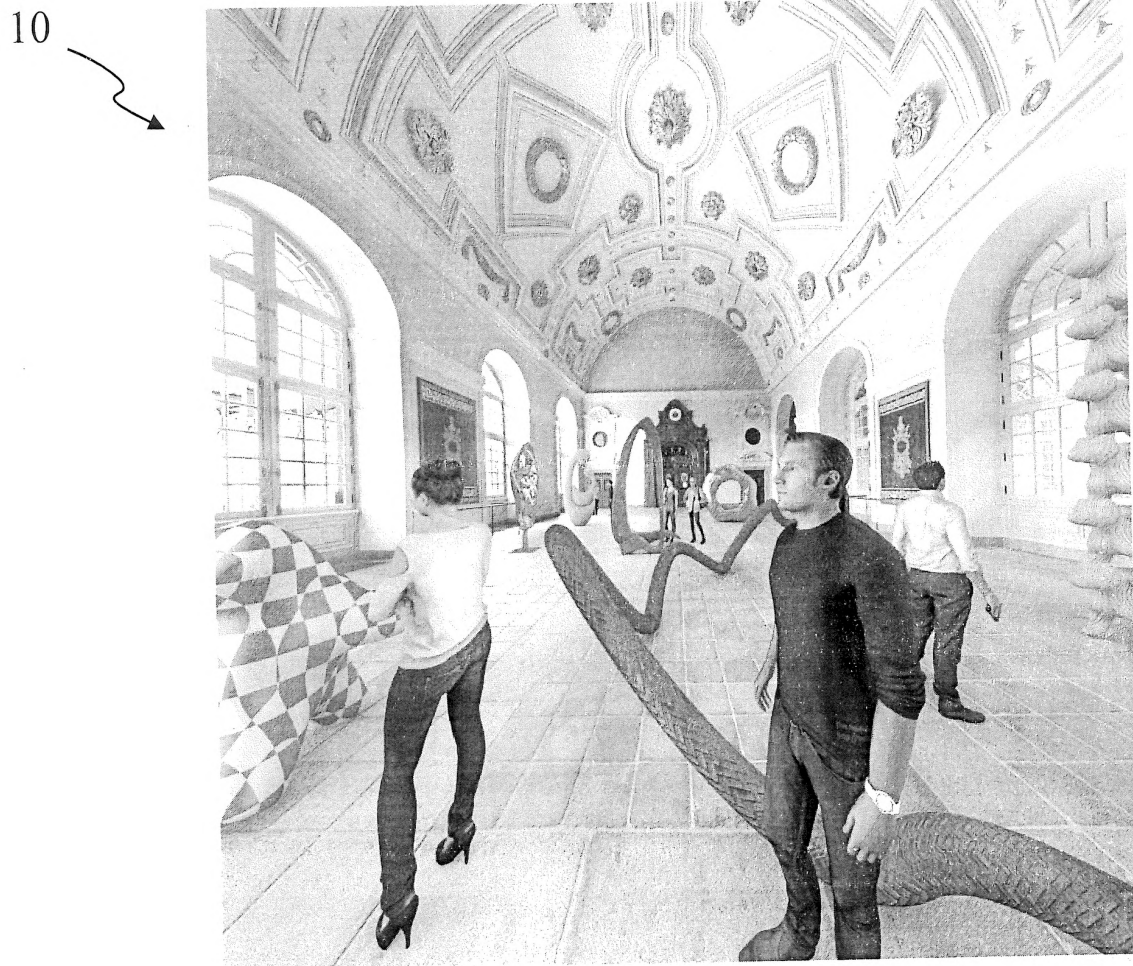
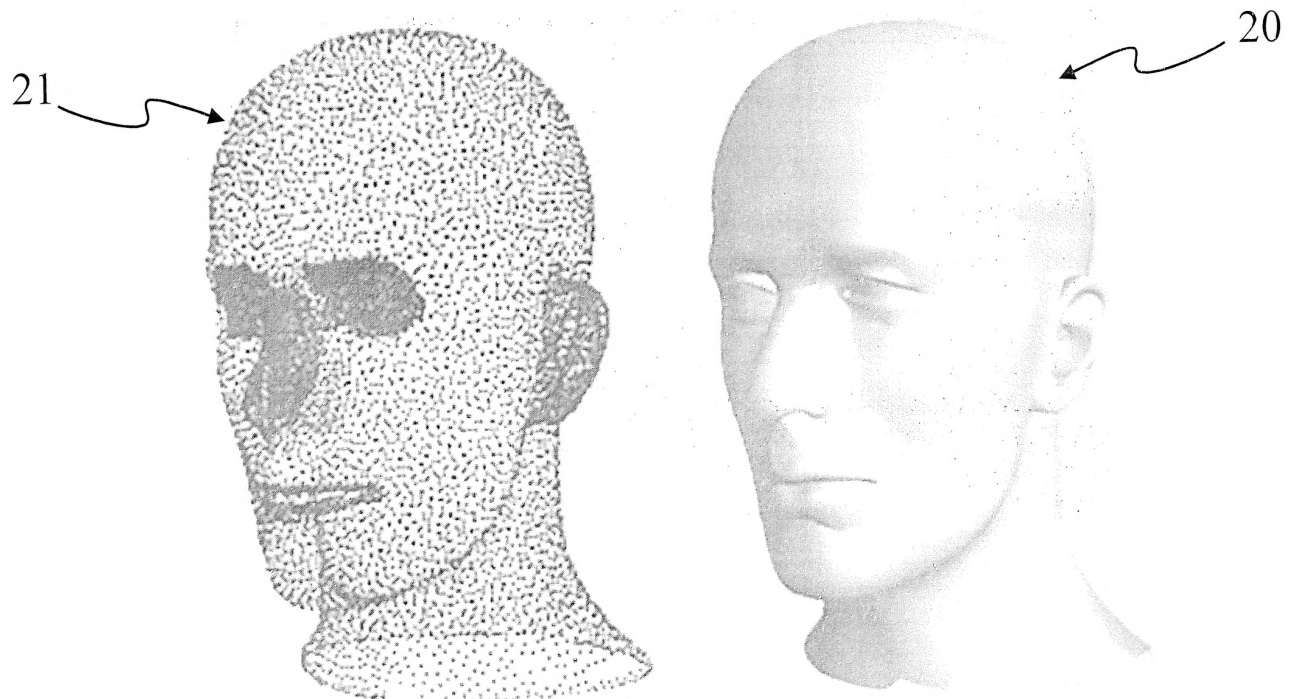
20. Thiết bị theo điểm 18, trong đó ít nhất bản vá thứ nhất của tập hợp nói trên của các bản vá thứ nhất tham chiếu đến ít nhất khu vực của ít nhất bản vá thứ ba tương ứng với sự tham số hóa hai chiều của nhóm thứ ba của các điểm được bao gồm trong một phần 3D khác của cảnh 3D được kết hợp với nội dung video thể tích thứ ba của tập hợp nói trên của các nội dung video thể tích.

21. Thiết bị theo điểm 20, trong đó các thuộc tính được kết hợp với các điểm của ít nhất nhóm thứ nhất của các điểm được kết hợp với ít nhất bản vá thứ nhất nói trên tham chiếu đến ít nhất khu vực của ít nhất bản vá thứ ba nói trên được bao gồm trong ít nhất khu vực nói trên của ít nhất bản vá thứ ba nói trên, ít nhất bản vá thứ nhất nói trên bao gồm thông tin nhận dạng ít nhất khu vực nói trên của ít nhất bản vá thứ ba nói trên.

22. Thiết bị theo điểm 20, trong đó các thuộc tính được kết hợp với các điểm của ít nhất nhóm thứ nhất của các điểm được kết hợp với ít nhất bản vá thứ nhất nói trên tham chiếu đến ít nhất khu vực của ít nhất bản vá thứ ba nói trên được bao gồm trong ít nhất bản vá thứ ba nói trên, thông tin ánh xạ biểu diễn cho việc tham chiếu của ít nhất bản vá thứ nhất nói trên đến ít nhất khu vực nói trên của ít nhất bản vá thứ ba nói trên được kết hợp với nội dung video thể tích thứ nhất nói trên.

23. Phương tiện đọc được bởi bộ xử lý không chuyên tiếp có lưu trữ trên đó các lệnh để làm cho bộ xử lý thực hiện ít nhất các bước của phương pháp theo một điểm bất kỳ trong số các điểm từ 1 đến 6.

24. Phương tiện đọc được bởi bộ xử lý không chuyển tiếp có lưu trữ trên đó các lệnh để làm cho bộ xử lý thực hiện ít nhất các bước của phương pháp theo một điểm bất kỳ trong số các điểm từ 13 đến 17.

**Fig.1****Fig.2**

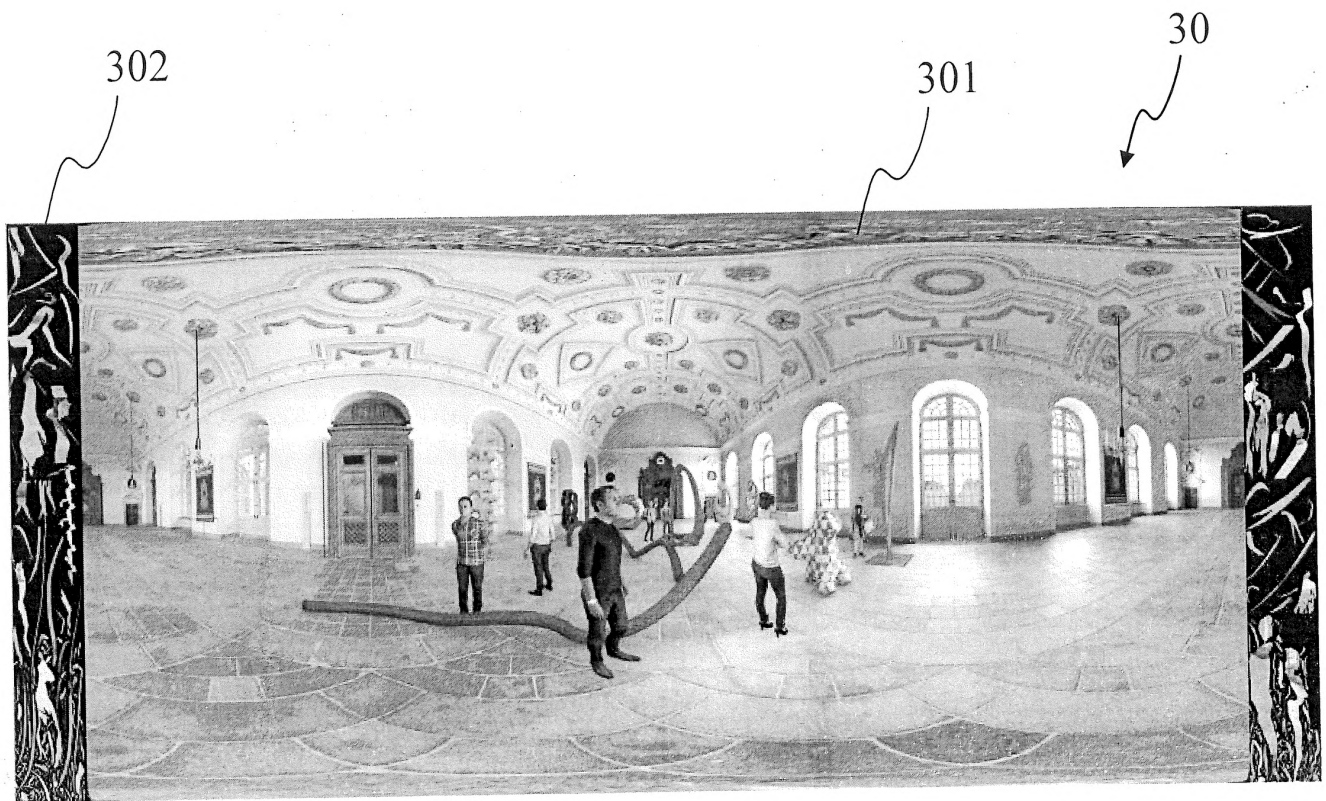
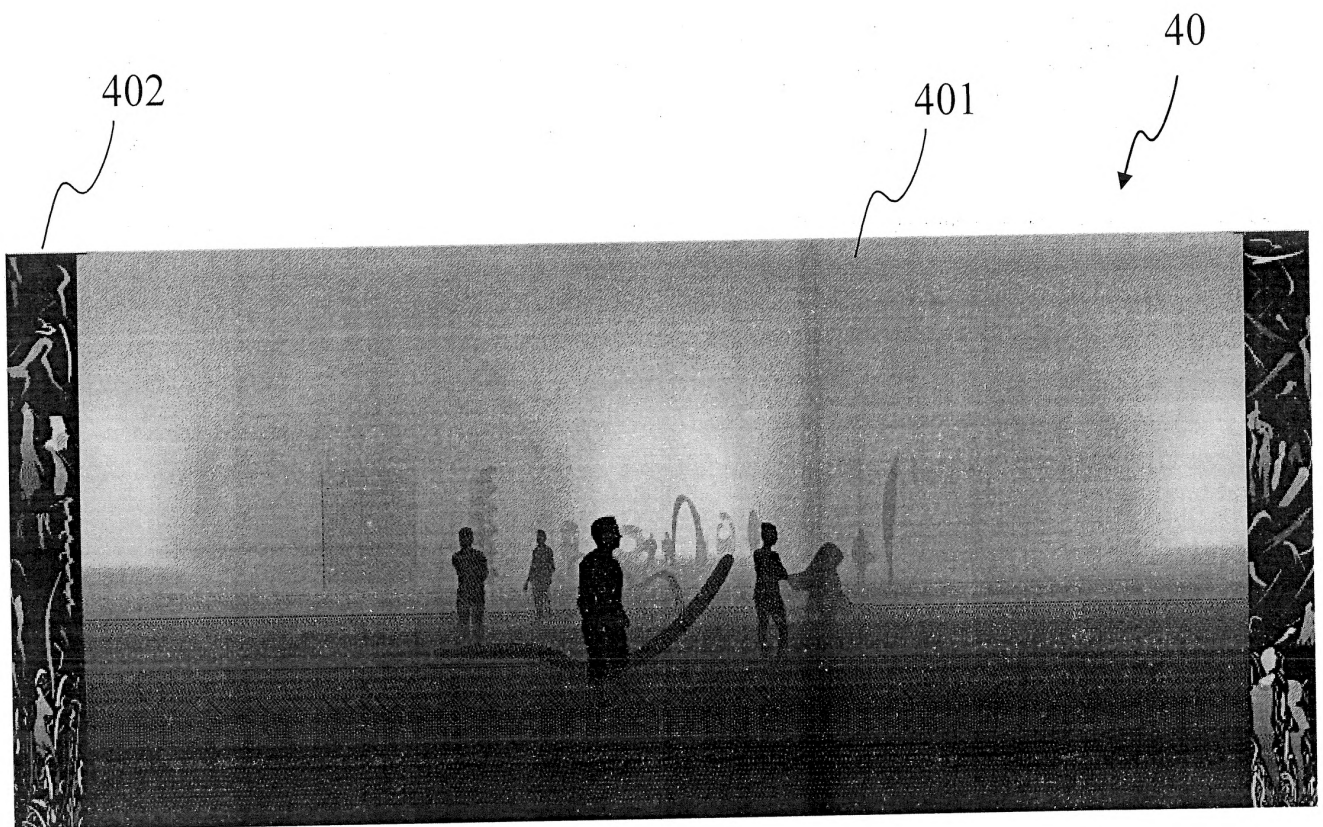
**Fig.3****Fig.4**

Fig.5

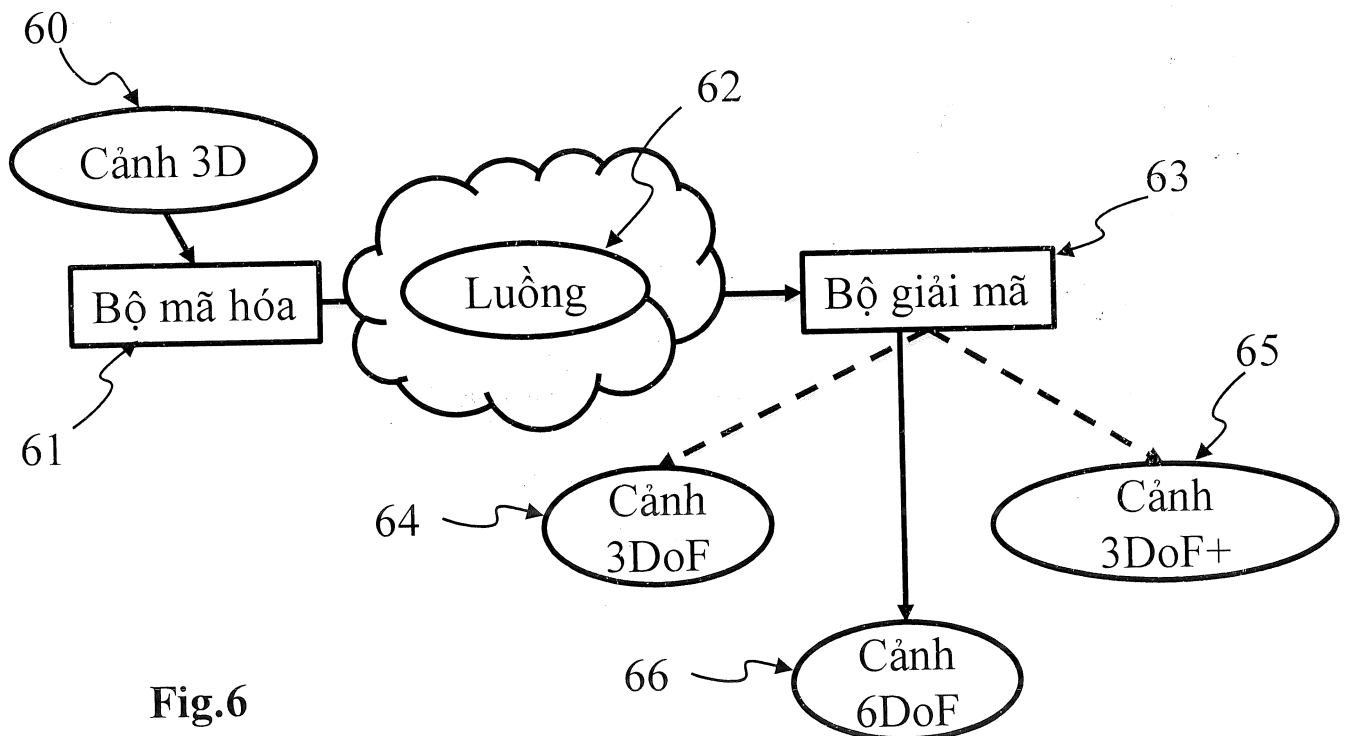
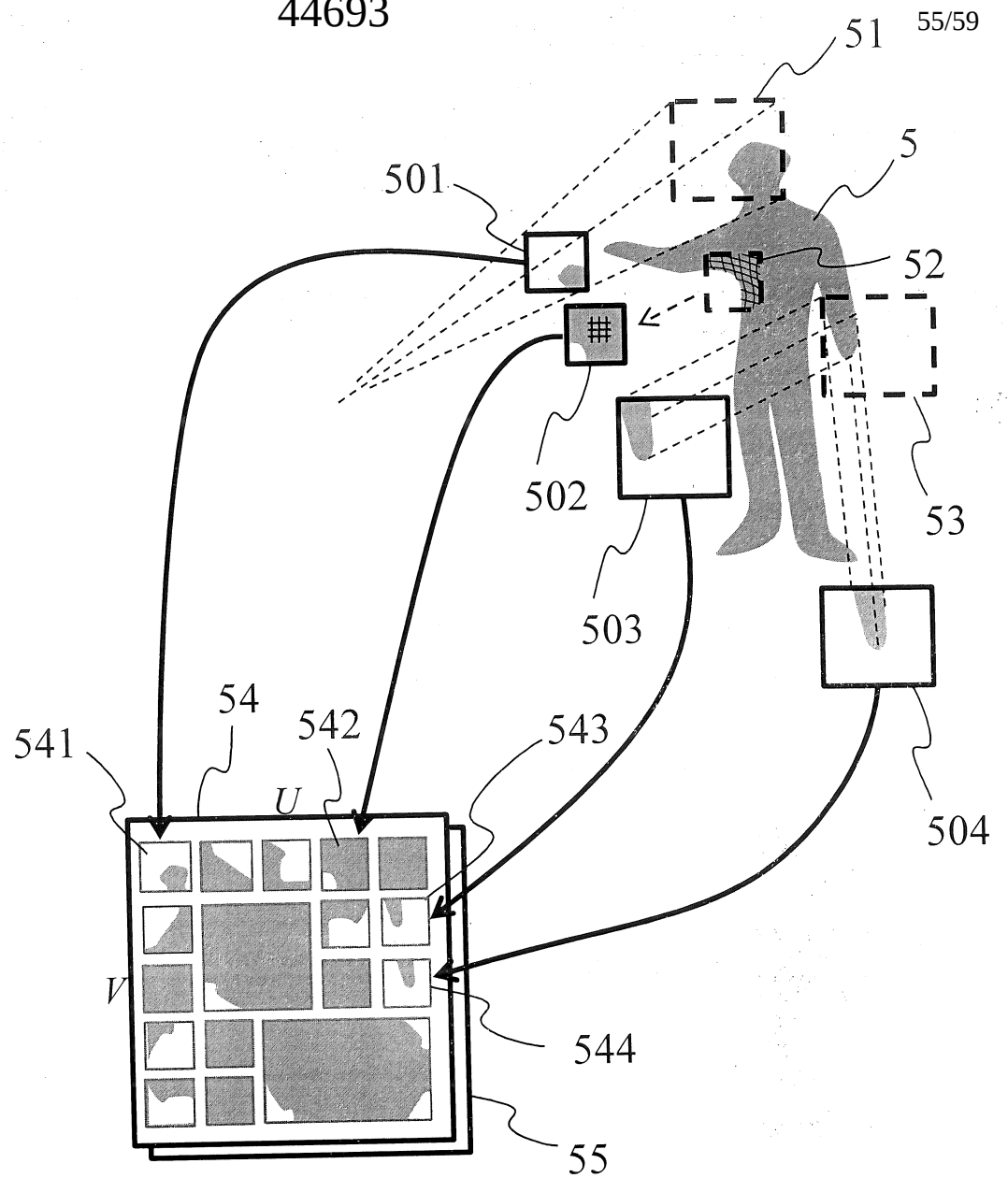


Fig.6

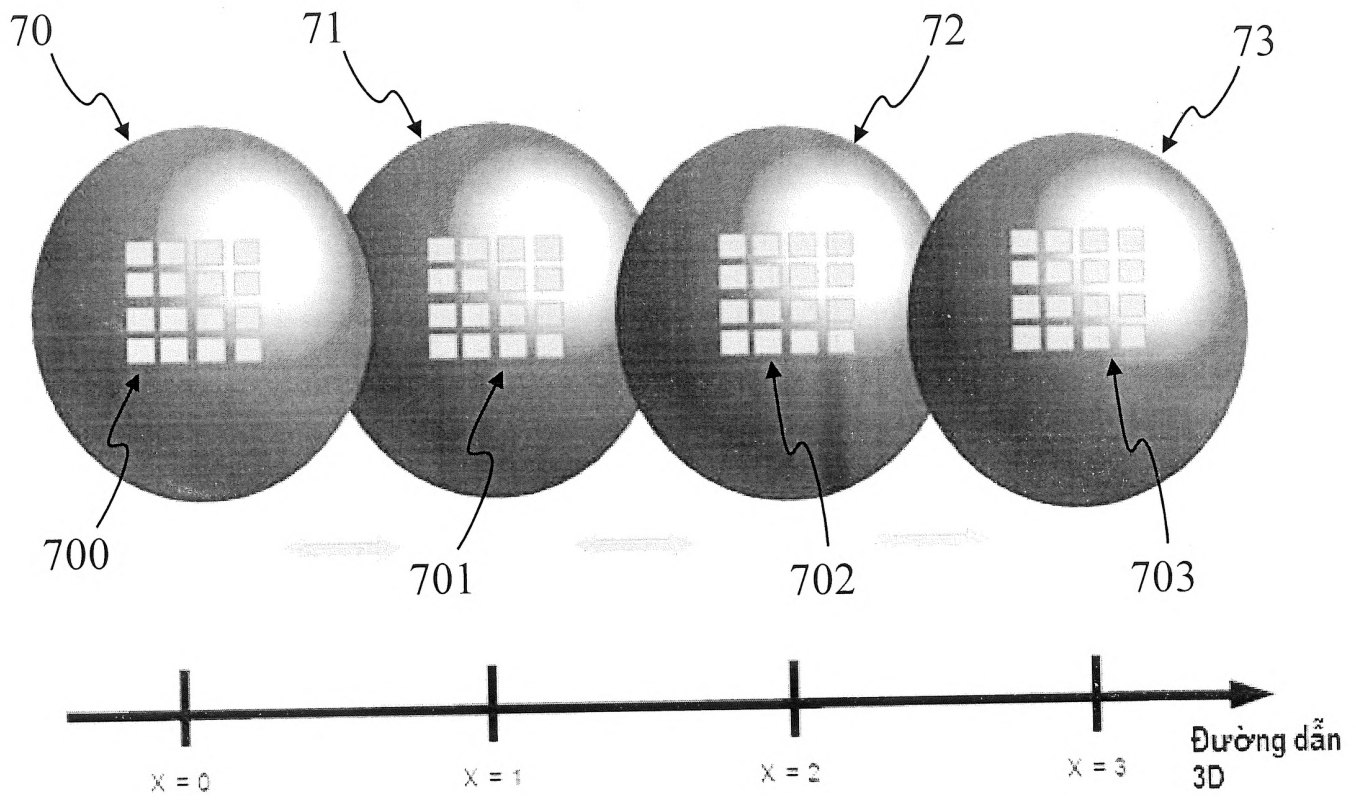


Fig. 7

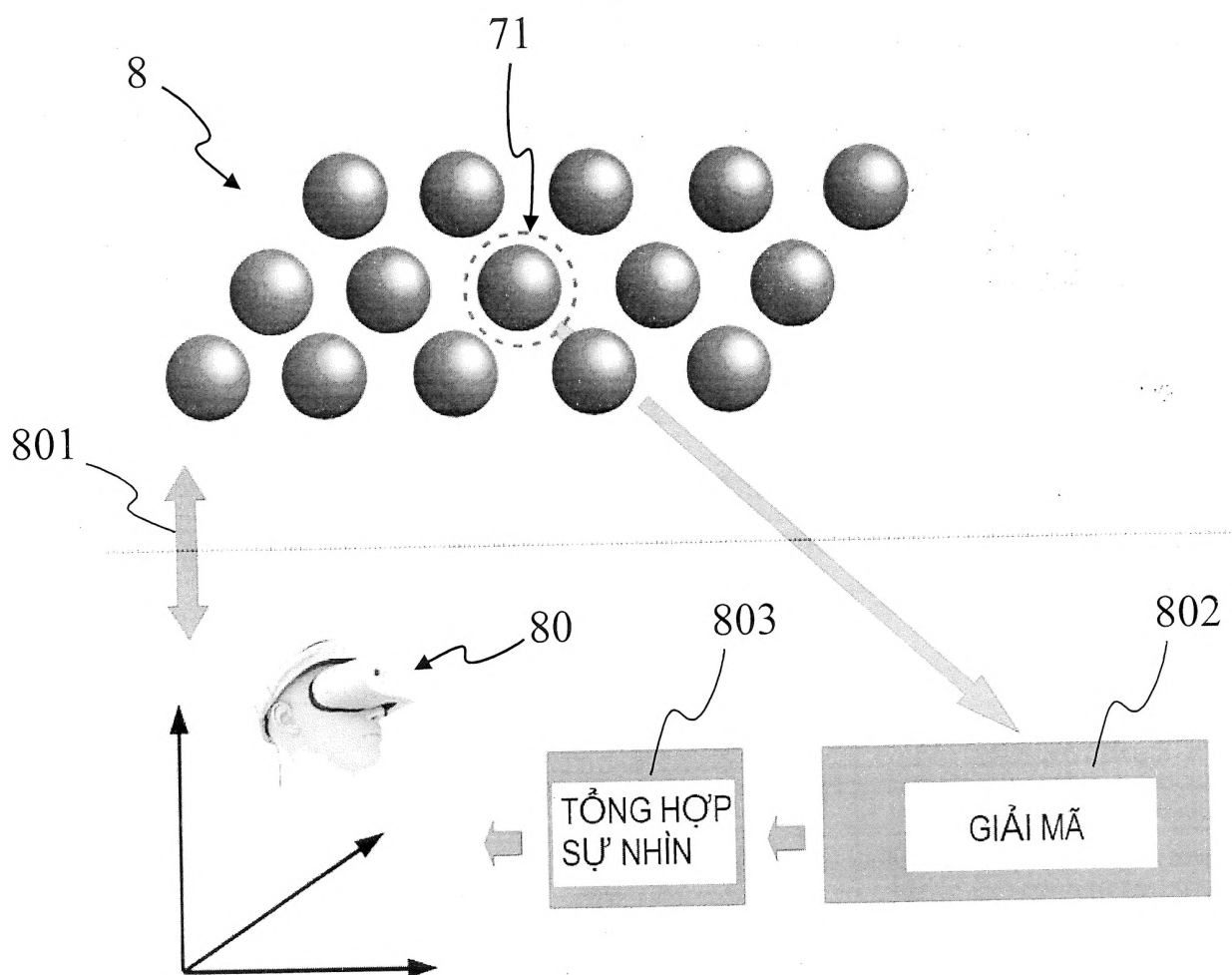


Fig. 8

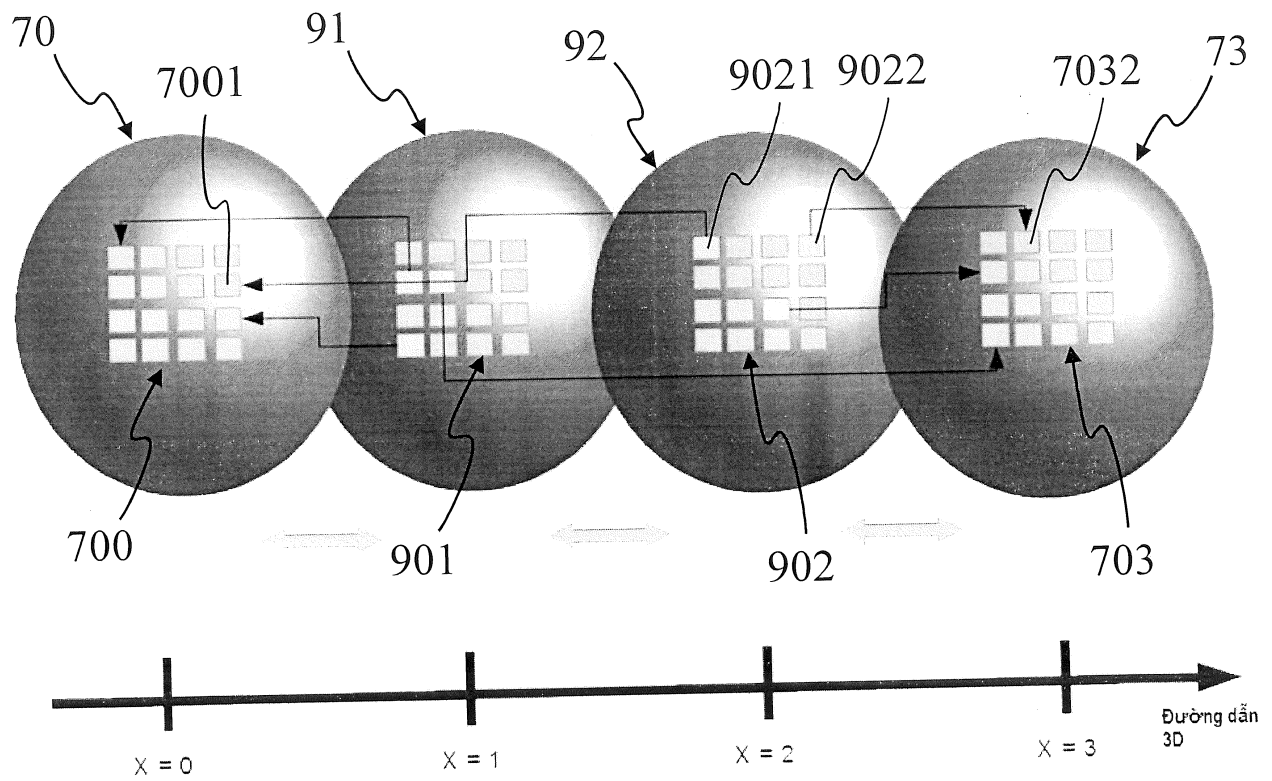


Fig.9

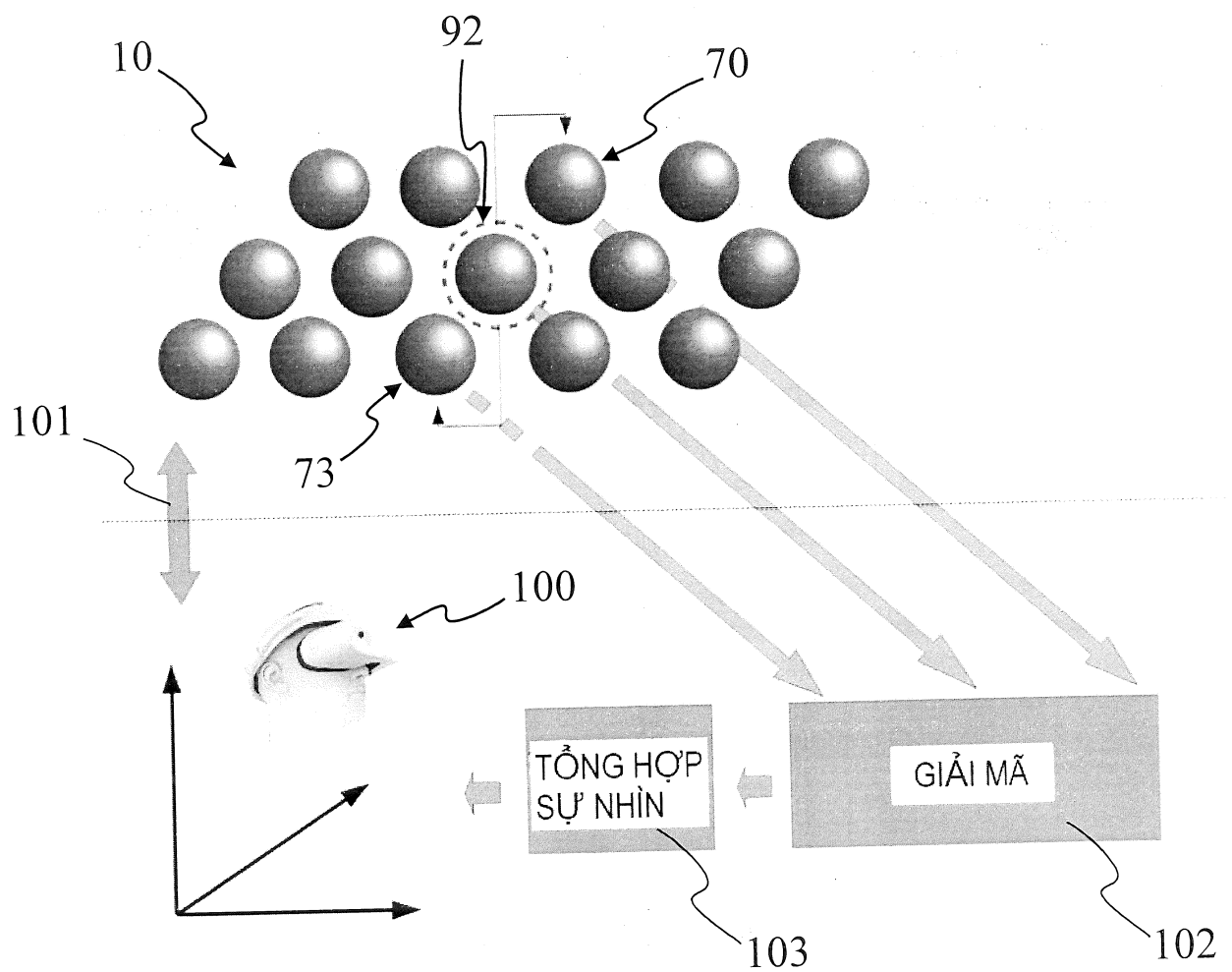


Fig.10

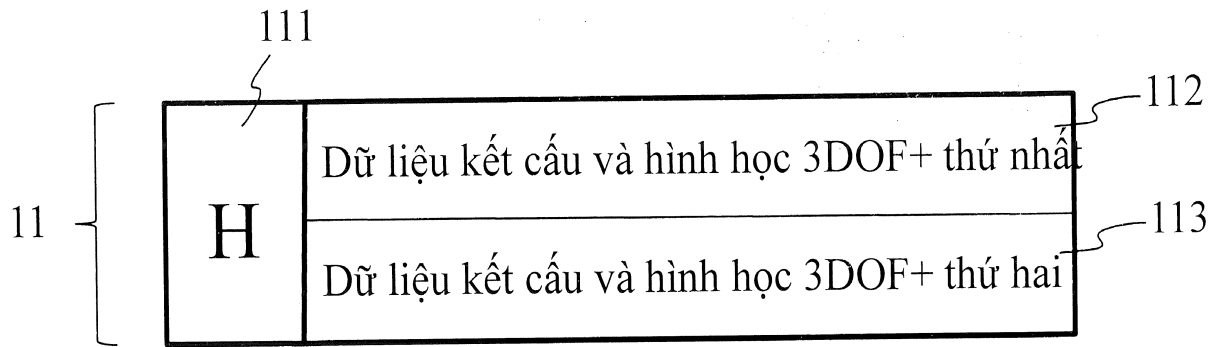


Fig.11

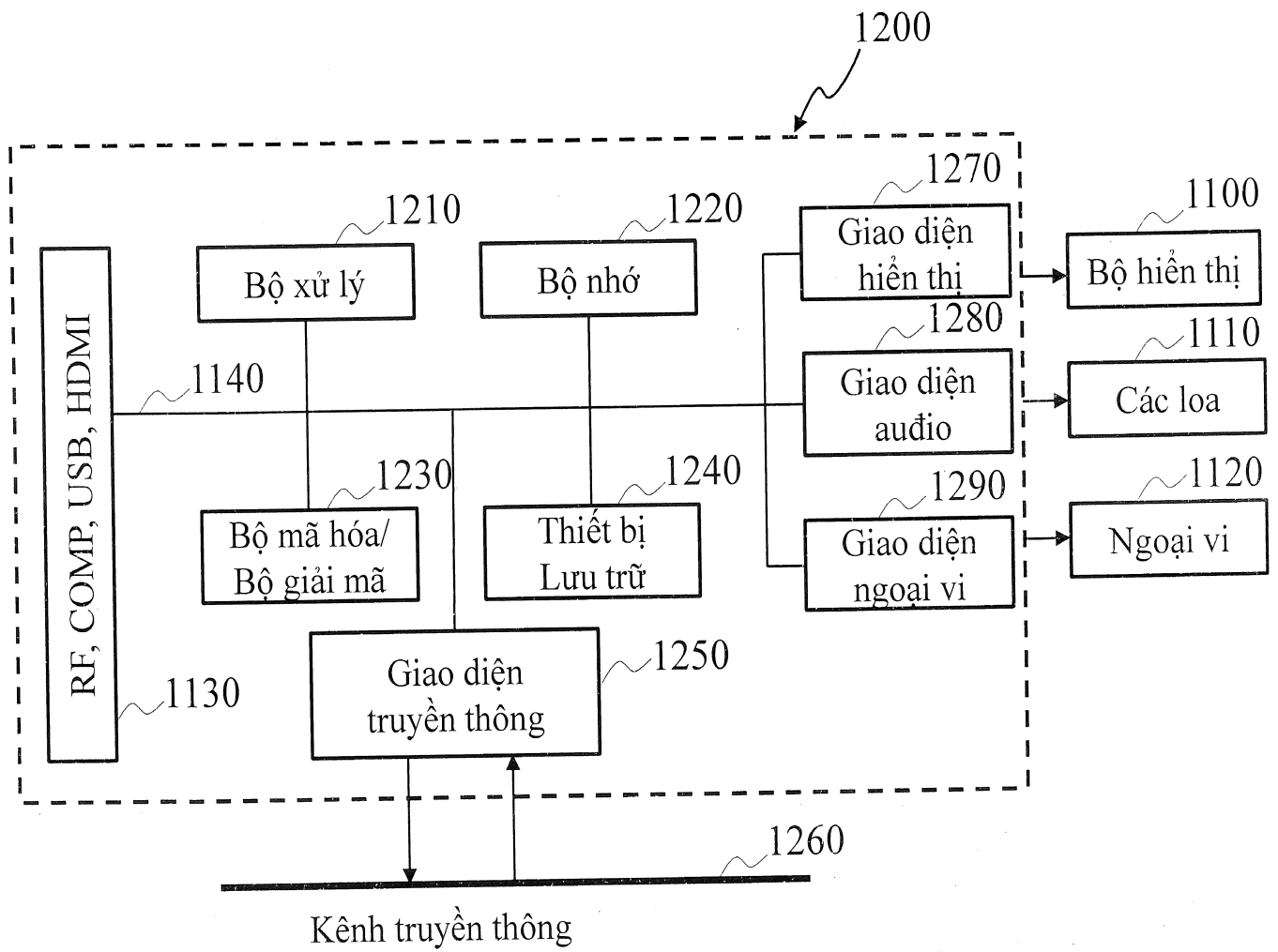
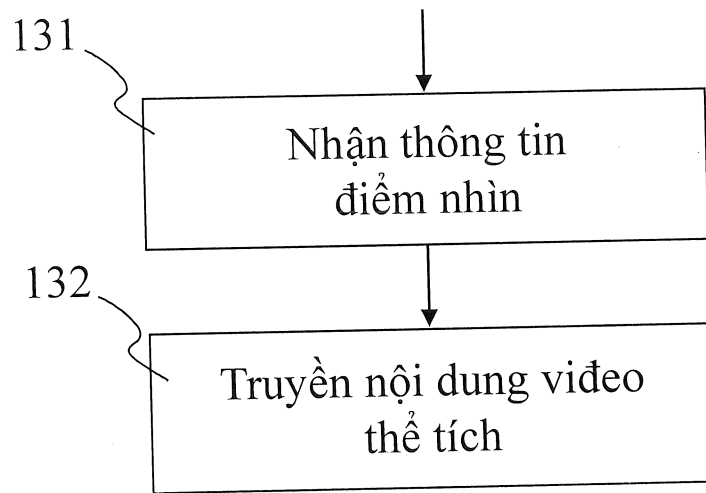
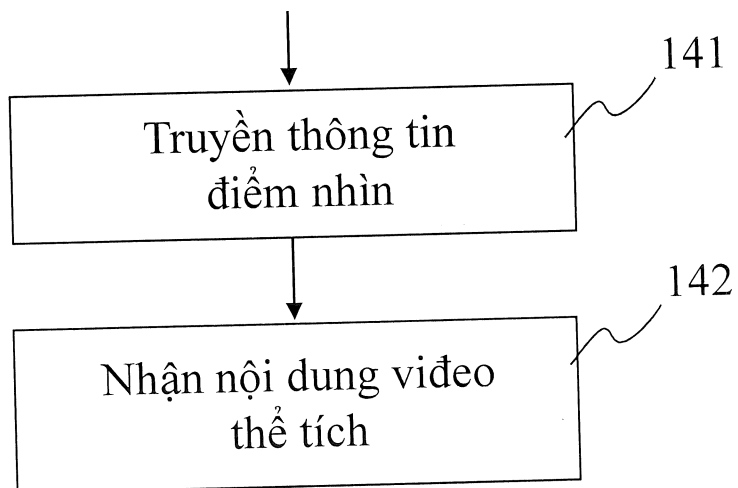


Fig.12

**Fig.13****Fig.14**