



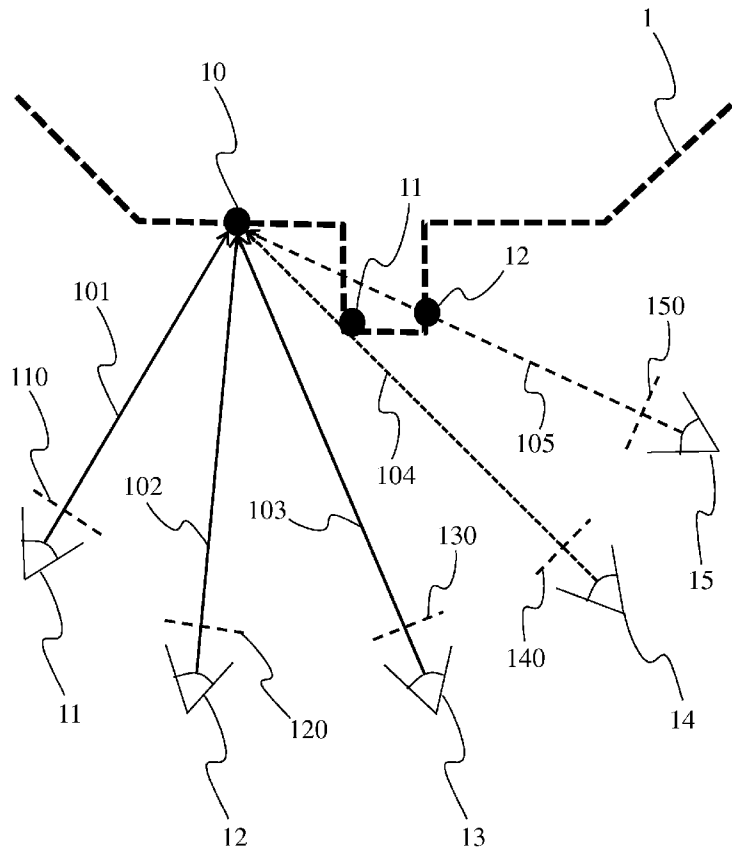
US 20140354632A1

(19) **United States**(12) **Patent Application Publication**
Alj et al.(10) **Pub. No.: US 2014/0354632 A1**(43) **Pub. Date: Dec. 4, 2014**(54) **METHOD FOR MULTI-VIEW MESH
TEXTURING AND CORRESPONDING
DEVICE****Publication Classification**

(51) **Int. Cl.**
G06T 15/04 (2006.01)
H04N 13/02 (2006.01)
(52) **U.S. Cl.**
CPC **G06T 15/04** (2013.01); **H04N 13/0282**
(2013.01); **H04N 13/0275** (2013.01)
USPC **345/419**

(71) Applicant: **THOMSON LICENSING**, Issy de
Moulineaux (FR)(72) Inventors: **Youssef Alj**, Rennes (FR); **Guillaume
Boisson**, Pleumeleuc (FR); **Philippe
Bordes**, Laille (FR); **Luce Morin**,
Rennes (FR); **Muriel Pressigout**,
Rennes (FR)(73) Assignee: **Thomson Licensing**, Issy de
Moulineaux (FR)(21) Appl. No.: **14/372,018**(22) PCT Filed: **Oct. 8, 2012**(86) PCT No.: **PCT/EP2012/069862**§ 371 (c)(1),
(2), (4) Date: **Jul. 13, 2014**(30) **Foreign Application Priority Data**Jan. 13, 2012 (EP) 12305047.8
Oct. 8, 2012 (WO) PCTEP2012/069862(57) **ABSTRACT**

A method for texturing a mesh associated with a surface representative of a scene captured in a plurality of images, wherein at least a mesh element of the mesh is at least partially visible from at least two first images of the plurality of images. As to reduce the amount of texture information associated with multi-views data representative of the scene, the method comprises for each first image (110 to 140), association of a first texture information with the mesh element, projection of the first texture information in the at least two first images, for each first texture information, estimation of an error information according to a comparison result between the projected first texture information and the texture information of said at least two first images, and selection of one of the at least two first texture information according to the error information.



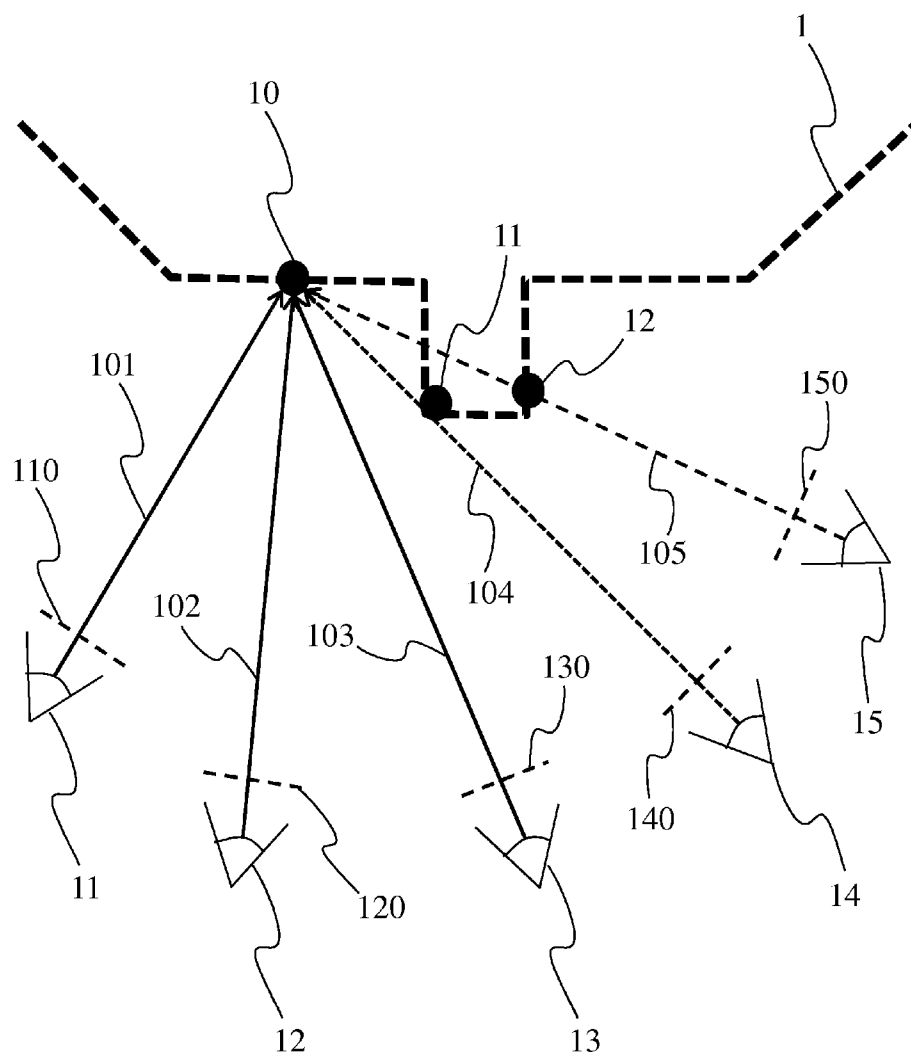


Fig 1

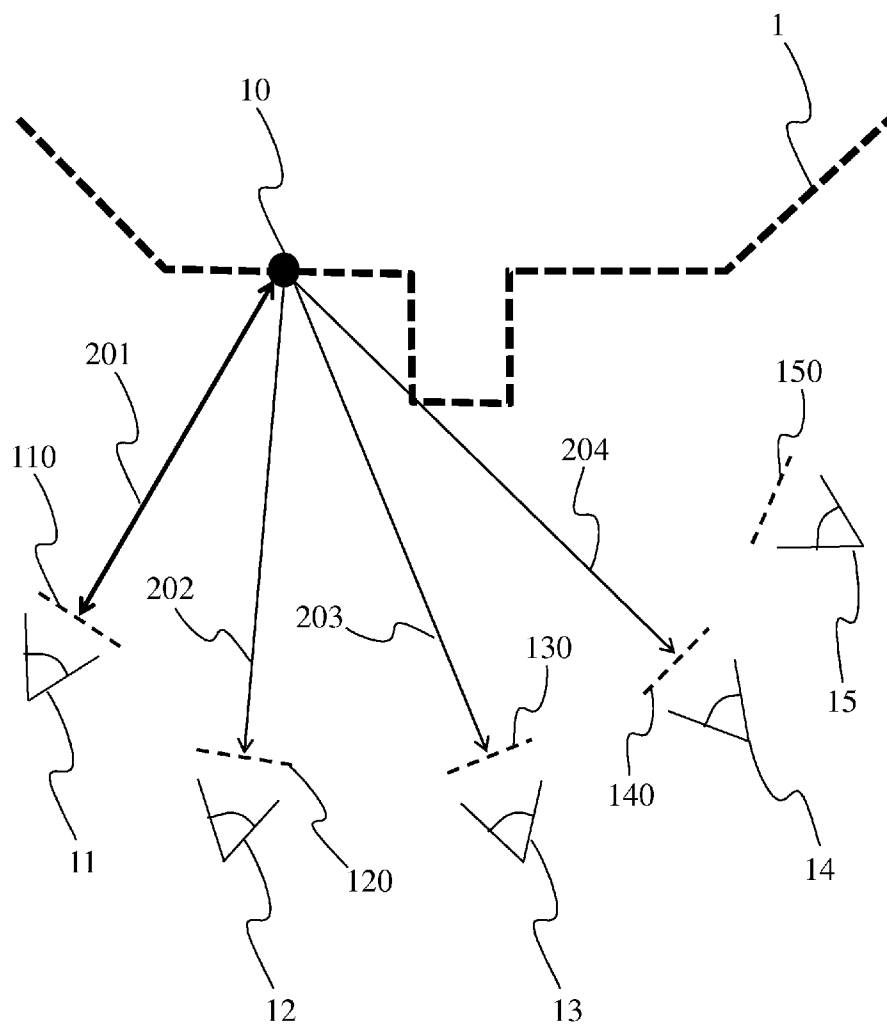


Fig 2

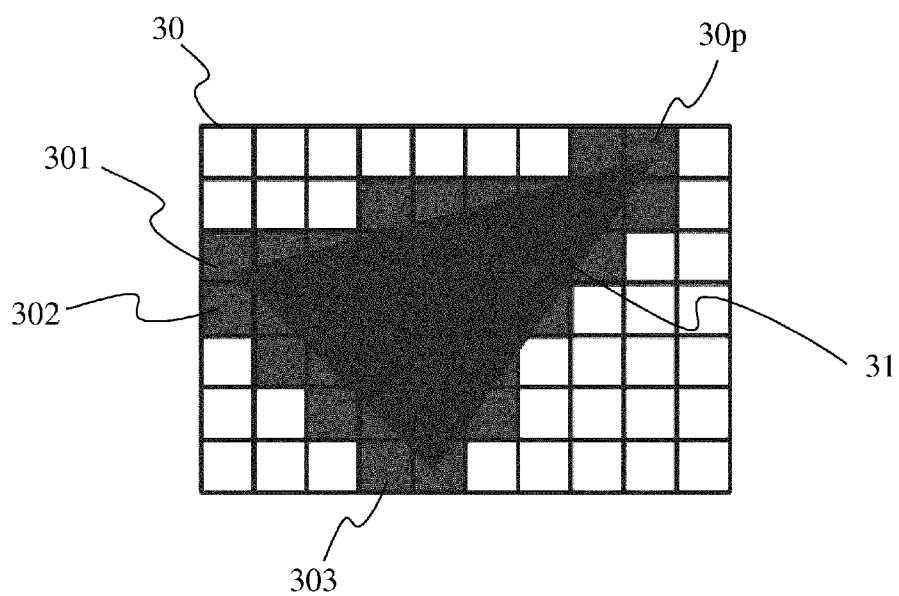


Fig 3

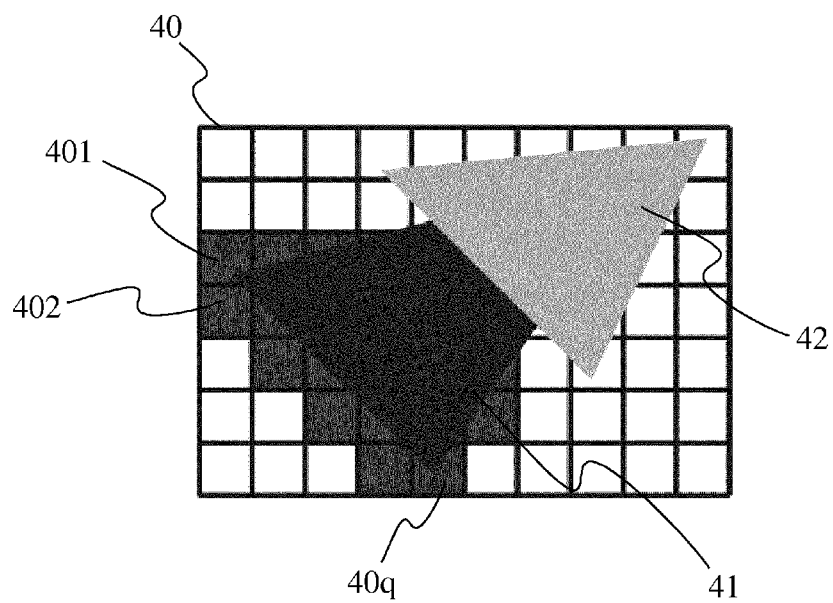


Fig 4

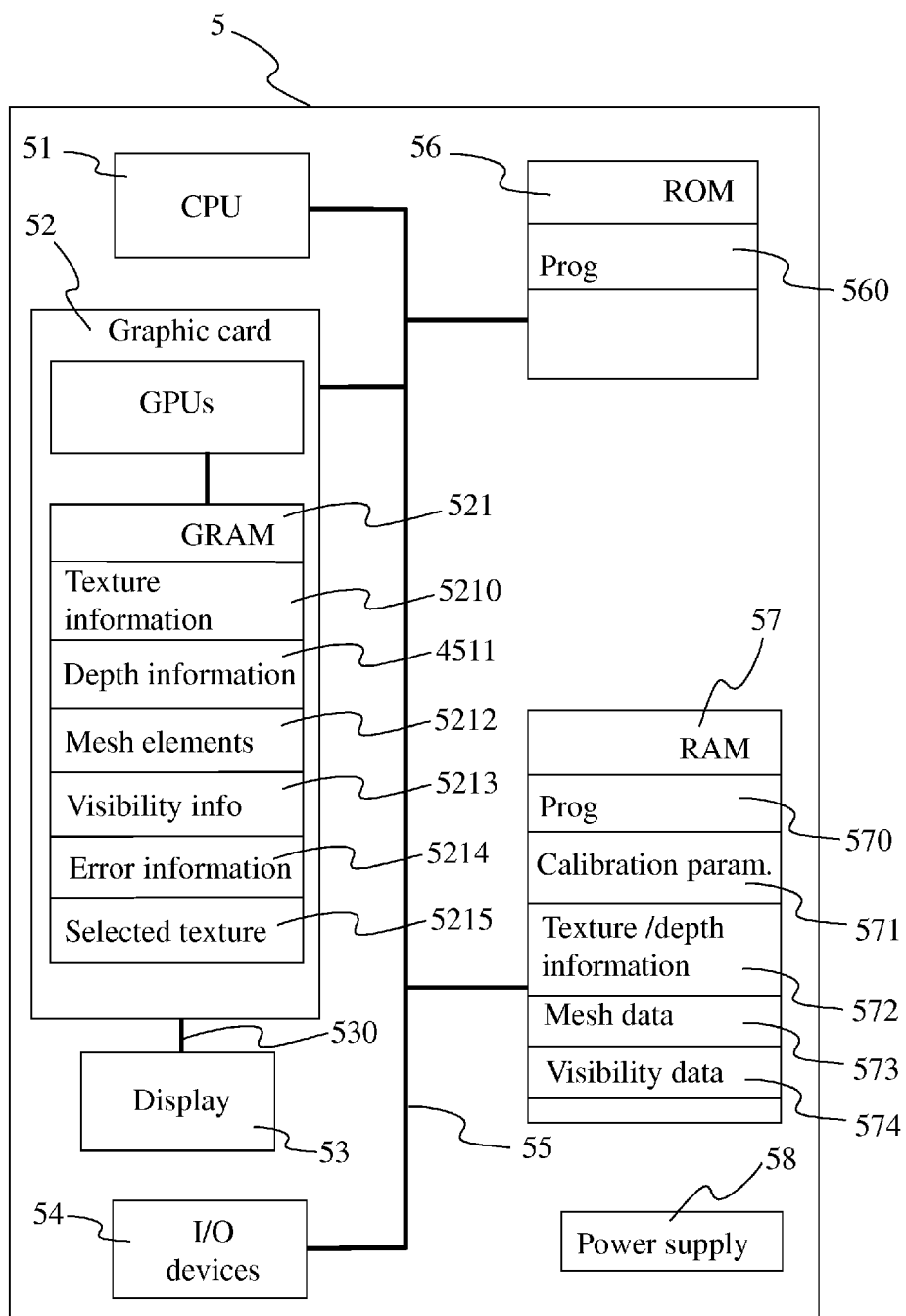
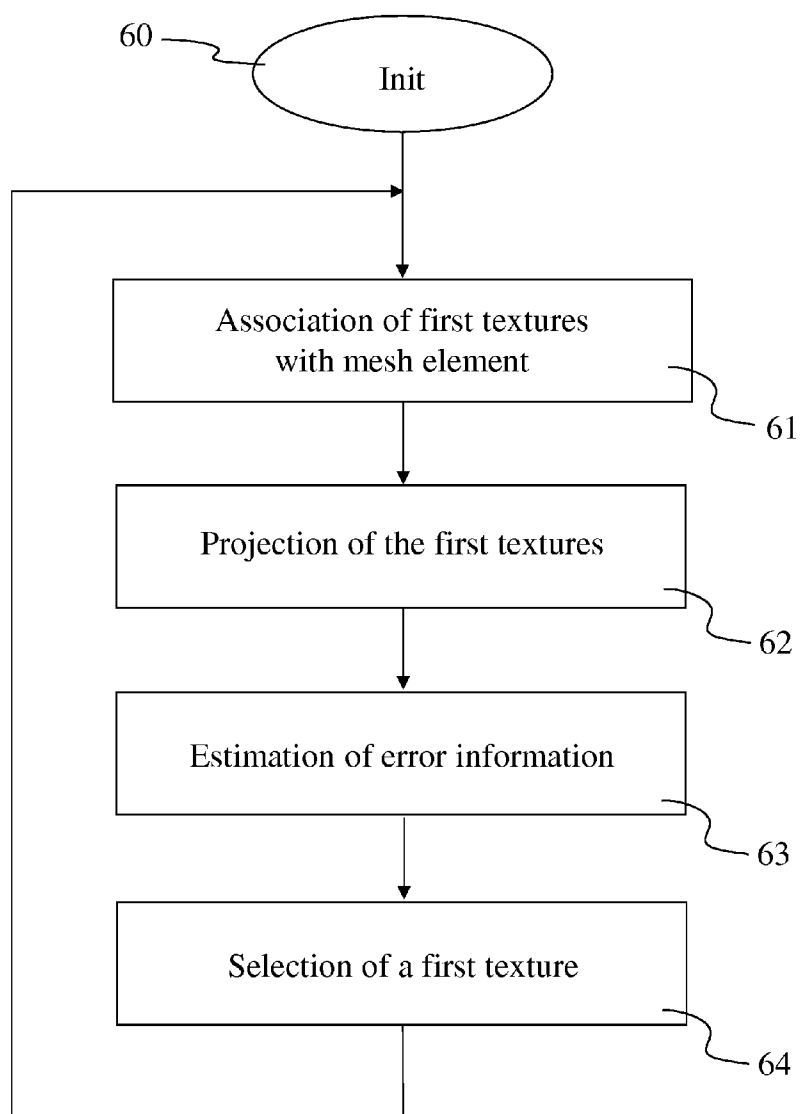


Fig 5

**Fig 6**

METHOD FOR MULTI-VIEW MESH TEXTURING AND CORRESPONDING DEVICE

1. SCOPE OF THE INVENTION

[0001] The invention relates to the domain of image or video processing and more particularly in the processing of data representative of texture associated to images of a multi-views video stream. The invention also relates to the domain of modelling the texture of real scene by using texture data and depth data associated to images representative of the scene according to several points of view.

2. PRIOR ART

[0002] Combined with the use of stereoscopic displays, Multi-View Imaging (MVI) provides a realistic depth perception to the user and allows a virtual navigation around the scene. It also offers the possibility to synthesize virtual views, opening therefore a broad spectrum of applications such as 3DTV and Free-viewpoint TV (FTV). An end-to-end MVI chain presents many challenges due to the inherent problems related to scene capture, data representation and transmission. Indeed, during scene acquisition, multiple cameras are used; hence several photometric errors may occur due to the changing illumination across different views, which increase the signal ambiguity during depth estimation. Scene representation is also a challenging task. One should strive to build a dense yet non-redundant scene model from the set of overlapping views, each view comprising information related to the texture and to the depth. The compromise between model simplicity and easy rendering is usually hard to satisfy. The representation is then to be transmitted through a communication channel. Determining the proper rate-distortion trade-off is another key point. At decoder side, high fidelity to original views is required. Finally to assess the quality of virtual views, conventional image based objective metrics such as Peak of Signal to Noise Ratio (PSNR) or Structural Similarity (SSIM) are not applicable since the reference images for such views are not available.

3. SUMMARY OF THE INVENTION

[0003] The purpose of the invention is to overcome at least one of these disadvantages of the prior art.

[0004] More particularly, the invention has the notable purpose of reducing the amount of texture information associated to multi-views data representative of a scene.

[0005] The invention relates to a method for texturing a mesh, the mesh being associated with a surface representative of a scene captured according to a plurality of points of view, an image of the scene comprising texture information being associated with each point of view, wherein at least a mesh element of the mesh is at least partially visible from at least two first images of the plurality of images. The method comprises the following steps:

[0006] for each first image, association of a first texture information with the mesh element,

[0007] projection of the first texture information in the at least two first images,

[0008] for each first texture information, estimation of an error information according to a comparison result between the projected first texture information and the texture information of the at least two first images, and

[0009] selection of one of the at least two first texture information according to said error information.

[0010] Advantageously, the error information associated to a first texture information corresponds to a sum of the comparison results between the projected first texture information and each of the at least two first images.

[0011] According to a particular characteristic, the selected first texture information corresponds to the first texture information having the least error information.

[0012] Advantageously, the method further comprises the steps of:

[0013] projection of the mesh element in at least a part of the images of the scene,

[0014] for each image of the at least a part of images, comparison between a depth information associated with the projected mesh element and a depth information associated with the image,

[0015] for each image of said at least a part of images, determination of a visibility information associated to said mesh element (t) according to the result of the comparison of previous step, a mesh element being labeled as at least partially visible or not visible from said image according to said visibility information.

[0016] Advantageously, the projection of the mesh element corresponds to a conversion of the coordinates of the mesh element from world space into the image space of the image.

[0017] According to another characteristic, each image comprises a plurality of pixels, video information being associated to the plurality of pixels.

[0018] According to a specific characteristic, each image is represented by using Multi-view Video plus Depth (MVD) representation.

[0019] The invention also relates to a computation unit configured to texture a mesh, the mesh being associated with a surface (1) representative of a scene captured according to a plurality of images, each image comprising texture information, wherein at least a mesh element of the mesh is at least partially visible from at least two first images of the plurality of images, the computation unit comprising:

[0020] means for associating, for each first image, a first texture information with the mesh element,

[0021] means for projecting the first texture information in the at least two first images,

[0022] means for estimating, for each first texture information, an error information according to a comparison result between the projected first texture information and the texture information of the at least two first images, and

[0023] means for selecting one of the at least two first texture information according to the error information.

[0024] Advantageously, the error information associated to a first texture information corresponds to a sum of the comparison results between the projected first texture information and each of the at least two first images.

[0025] According to another characteristic, the selected first texture information corresponds to the first texture information having the least error information.

[0026] According to a particular characteristic, the computation unit further comprises:

[0027] means for projecting the mesh element in at least a part of the images of the scene,

[0028] means for comparing, for each image of the at least a part of images, between a depth information

associated with the projected mesh element and a depth information associated with the image,

[0029] means for determining, for each image at said least a part of the images, a visibility information associated to the mesh element according to the result of the comparison, a mesh element being labeled as at least partially visible or not visible in the image according to the visibility information.

[0030] According to a specific characteristic, the means for projecting said mesh element comprise means for converting the coordinates of the mesh element from world space into the image space of the image.

4. LIST OF FIGURES

[0031] The invention will be better understood, and other specific features and advantages will emerge upon reading the following description, the description making reference to the annexed drawings wherein:

[0032] FIG. 1 illustrates a meshed surface **1** representative of a scene viewed and captured according to several points of views, according to a particular embodiment of the invention;

[0033] FIG. 2 illustrates the association of a first texture information with a mesh element of the surface **1** of FIG. 1 and the projection of the first texture information on at least a part of the views, according to a particular embodiment of the invention;

[0034] FIG. 3 illustrates a mesh element of the meshed surface **1** of FIG. 1 entirely visible according to at least one of the points of view, according to a particular embodiment of the invention;

[0035] FIG. 4 illustrates a mesh element of the meshed surface of FIG. 1 partially visible according to at least one of the points of view, according to a particular embodiment of the invention;

[0036] FIG. 5 illustrates a device implementing a method for texturing a mesh of surface **1** of FIG. 1, according to a particular implementation of the invention;

[0037] FIG. 6 illustrates a method for texturing a mesh of a surface **1** of FIG. 1, according to a particular implementation of the invention.

5. DETAILED DESCRIPTION OF EMBODIMENTS OF THE INVENTION

[0038] The invention will be described in reference to a particular embodiment of a method for texturing a mesh of surface, which is representative of the scene. The scene has been acquired from several points of view, an image or view of the scene being associated with a point of view. A mesh is associated with the surface, the mesh comprising a plurality of mesh elements (for example triangles or quadrilaterals). Each image of the scene comprises a plurality of pixels, the number of pixels depending on the resolution of the image of the scene associated with the view. A texture information is associated with each image of the scene, the texture information corresponding for example to a video information (for example RGB data representative of the grey level associated with each colour (RGB) information) associated with the pixels of the image corresponding to a view of the scene. For a mesh element which is at least partially visible from two or more views or, equivalently, in two or more images (called first images in the following description), a first texture information extracted from one of the first images is associated to the mesh element and projected on all or part of the first

image, i.e. on the image where the mesh element is at least partially visible. The projected first texture information is compared with each texture information associated to the first images on which the first texture information has been projected. From the comparison results, an error information is estimated which is for example representative of the difference between the projected first texture information and the texture information associated with the first images on which the first texture information has been projected. The step of associating a first texture information with the mesh element is reproduced for another first image and the steps of projecting and comparing this new first texture information are reproduced. These steps are advantageously reiterated for each first image, forming a set of several candidate first texture information. When an error information has been estimated for all (candidate) first texture information, one of them is selected according to the error information assigned with each of them. The selection of one single texture information (to be assigned with a mesh element) among the plurality of texture information associated with the plurality of views for which the mesh element is at least partially visible enables to reduce the amount of data to be transmitted to a display or any receiver over a network. According to the invention, minimal input data needed to associate a texture with a mesh element is a meshed surface representative of a scene and texture information associated with a plurality of images representing the scenes according to different points of view. It also enables to select the best texture information candidate to map a given mesh element, reducing photometric errors, which occur for example when the texture information associated with a mesh element correspond to the average of all texture information available in the different views where the mesh element is visible or partially visible.

[0039] FIG. 1 illustrates a surface **1** representative of a scene viewed and captured according to several points of views. The surface **1** is representative of a scene acquired according to a plurality of points of view **11, 12, 13, 14, 15**. An image **110, 120, 130, 140, 150** is associated with each point of view **11** to **15**, corresponding to a representation of the scene according to a given point of view. A mesh is associated with the surface **1**, the mesh comprising a plurality of mesh elements (for example thousands of mesh elements), for example triangles or quadrilaterals. The surface **1** corresponds advantageously to a 3D representation of the scene.

[0040] The surface **1** and the associated mesh are obtained by combining methods well-known by the skilled person in the art, for example by combining the “shape from silhouette” method with the marching-cubes algorithm or by combining the space carving method with the marching-cubes algorithm. The “shape from silhouette” method (described for example in “*Spatio-Temporal Shape from Silhouette using Four-Dimensional Delaunay Meshing*” by Ehsan Aganj, Jean-Philippe Pons, Florent Ségonne and Renaud Keriven) and the space carving method (described in “*A Theory of Shape by Space Carving*” published in *International Journal of Computer Vision*, 38(3), 199-219 (2000), by Kiriakos N. Kutulakos and Steven M. Seitz) are used for obtaining a 3D representation of a scene from a plurality of views of the scene. The marching cubes algorithm (published in the 1987 *SIGGRAPH proceedings* by Lorensen and Cline) is then applied to the 3D representation of the scene as to obtain a meshed surface as a 3D representation of the scene. French Patent application FR1161788, filed on Dec. 16, 2011, describes a method for obtaining a meshed surface corre-

sponding to a 3D representation of a scene. FR1161788 recites a 3D modeling system designed for Multi-view Video plus Depth (MVD) sequences. The aim is to remove redundancy in depth information present in the MVD data. To this end, a volumetric framework is employed in order to merge the input depth maps. Hereby a variant of the space carving algorithm is proposed: voxels are iteratively carved by ray-casting from each view, until the 3D model be geometrically consistent with every input depth map. A surface mesh is then extracted from this volumetric representation thanks to the marching cubes algorithm.

[0041] At least one of the mesh elements is visible or partially visible in different images associated to the views of the scene. According to the non limitative example of FIG. 1, the mesh element 10 is entirely visible from the points of view 11, 12, 13, i.e. the mesh element 10 is entirely visible in the first images 110, 120, 130 associated with the points of view 11, 12, 13. The mesh element 10 is partially visible from the point of view 14, i.e. the mesh element 10 is partially visible in the first image 140. Indeed, as illustrated on FIG. 1, the viewing direction 104 having as origin the point of view 14 and as extremity the mesh element 10 is tangent to a point 11 (or mesh element 11) of the surface 1 on the path separating the point of view 14 and the mesh element 10. The mesh element is not visible from the point of view 15 and is not visible on the associated image 150. The image 150 is thus not considered as being a first image. Indeed, as illustrated on FIG. 1, the viewing direction 105 having as origin the point of view 15 and as extremity the mesh element 10 has an intersection with the surface 1, the intersection corresponding to the point or mesh element 12 of surface 1, the intersection 12 being positioned between the point of view 15 and the mesh element 10 on the path corresponding to the viewing direction 105. For each image 110 to 150, a visibility label is advantageously associated with the mesh element 10. The visibility label takes for example two values, i.e. visible and non visible, or three value, i.e. entirely visible, partially visible and non visible.

[0042] According to a non limitative embodiment, the visibility information associated to the mesh element 10 is determined by comparing a depth information associated with the mesh element 10 and a depth information associated to pixels of a first image 110 to 150 on which the mesh element 10 is projected. The mesh element visibility with respect to each image 110 to 150 is determined for example using OpenGL z-buffer. In a first pass, the mesh (i.e. all mesh elements) is projected onto a current image (for example the image 110), and the z-buffer is extracted, the z-buffer comprising a depth information associated to each mesh element as seen in the current image (for example, a depth information is associated with each pixel of the current image). Depth information is determined by projecting mesh elements into an image, i.e. by converting the coordinates of the mesh element from world space into the image space. A second pass is dedicated to visibility determination using the computed z-buffer. Each vertex of the mesh element 10 is projected onto the current image 110 and the depth component, denoted as $z_{projected}$, is checked against the pixel depth z-buffer of the current image 110. If the projected vertex is behind the pixel in the z-buffer (i.e. the value $z_{projected}$ is greater than the depth value associated to the pixel in the z-buffer of the current image), then this vertex is hidden, and thus the set of mesh triangles lying on this vertex are not visible. In contrast, if the projected vertex is ahead the pixel in the z-buffer (i.e. the value $z_{projected}$ is less

than the depth value associated to the pixel in the z-buffer of the current image), then this vertex is visible, and thus the set of mesh triangles lying on this vertex are at least partially visible. The visibility information associated with the mesh element is for example determined according to the following algorithm:

```
// Views traversal
for each view  $V_j$  do
    Initialize all mesh elements to visible.
    Project M onto  $V_j$ .
    Read  $z_{buffer}$ .
    // Mesh vertices traversal
    for each vertex v do
        Determine (q, l,  $z_{projected}$ ) the projection of the vertex v(X, Y, Z)
        onto  $V_j$ .
        if  $z_{buffer}[q, l] > z_{projected}$  then
            v is a hidden vertex. Mark all the triangles lying on v as
            hidden.
        end
    end
end
end
```

[0043] According to a variant, the mesh is not projected into the current image as to determine the z-buffer associated with the current image. The depth of the mesh element is compared to a depth information comprised in a depth map associated with the current image, the depth map being received with the image (for example in a stream of the type MVD).

[0044] According to a variant, the visibility information associated with the mesh element 10 (and with each mesh element of the surface 1) is received with the data representatives of the images 110 to 150 and is not to be determined. These data comprises for example RGB information for each pixel and visibility information associated with the mesh elements of the mesh of FIG. 1 and comprising information about the visibility of mesh elements in each and every images representing the scene (the visibility information being for example stored in a visibility map).

[0045] Images 110 to 150 belong to a multi-view video stream comprising a sequence of several images representative of a same scene. According to a non-limitative example, the multi-view video stream is MVD type (Multi-view Video plus Depth). In case of MVD data, depth information is associated with each image representing the scene. A depth information is thus associated with each image 110 to 150, the depth information correspond to data representative of depth, for example a depth map (for example of the type of z-buffer) or a disparity map. A depth information is a generic name and corresponds to any data structure representative of a depth or of a disparity. The depth information associated with the pixels of the images 110 to 150 are for example captured with the use of appropriate sensors, for example by using infrared sensors associated with an infrared emitter (for example system of the type Kinect®) or by using a system comprising a laser emitter and an associated sensor configured for determining the crossing time of a laser ray emitted toward the captured scene and reflected by the scene, the determined time being representative of the path crossed and thus of the depth associated with the point of the scene having reflected the laser ray. According to a variant, video information (corresponding for example to colour information) and depth information associated with the pixels of images 110 to 150 are stored in memories under the form of RGB α (Red, Green, Blue, α channels) information, channels RGB being used for

storing video information associated with each pixel (for example 8, 10 or 12 bits per channel) and the α channel being used for storing the depth information (or disparity information), for example on 8, 10 or 12 bits.

[0046] Advantageously, intrinsic and extrinsic parameters of acquisition devices used for acquiring the images **110** to **150** are known and for example stored in a memory. Intrinsic parameters comprise for example focal length, enlargement factor of the image, coordinates of the projection of the optical centre of the acquisition device on image plane and/or a parameter representative of potential non-orthogonality of lines and columns of photosensitive cells forming the acquisition device. Extrinsic parameters comprise for example the orientation matrix for passing from world space to image space (and inversely) and/or components of translation vector for passing from world space to image space (and inversely).

[0047] Naturally, the number of images is not limited to 5 images but extends to any number greater than 2, for example 2, 3, 4, 5, 10, 20, 50, 100 images, data representative of depth being associated with each image. In a same way, the number of first images is not limited to 4 but extends to any number greater than 2. Advantageously, each image is acquired by a particular acquisition device according to a particular point of view, acquisition devices being spatially staggered. According to a variant, only a pair of left and right images is acquired by means of two acquisition devices, the other images of the plurality of images representative of the acquired scene being estimated from the left and right images by disparity compensated interpolation. According to this variant, each image is also representative of the same scene according to a particular point of view and depth information is associated with each image, independently of the fact that the image be acquired or interpolated.

[0048] FIG. 2 illustrates the association of a first texture information with a mesh element of the surface **1** and the projection of a first texture information on some views, according to a particular and non limitative embodiment of the invention. The mesh element **10** is visible or partially visible from the points of view **11**, **12**, **13** and **14**, i.e. the mesh element **10** is visible or partially visible in the images **110**, **120**, **130** and **140** respectively associated with the points of view **11**, **12**, **13** and **14**. The images **110** to **140** in which the mesh element is partially visible are called first images. One first image **110** is selected among the plurality of first images **110** to **140** and a first texture information extracted from thus first image **110** is extracted and associated with the mesh elements. The first texture information corresponds advantageously to the video information (or RGB colour information) associated with the pixels of the first image **110** corresponding to the projection of the mesh element **10** onto the first image **110**. The pixels of the first image on which is projected the mesh element **10** are illustrated on FIG. 3. According to a variant, the texture information associated with the images corresponds to YUV data.

[0049] FIG. 3 illustrates the projection **31** of the mesh element **10** onto an image **30** (corresponding for example to the image **110**) on which the mesh element is entirely visible. The first texture information associated with the mesh element **10** corresponds to the video information associated with the pixels on which the mesh element **10** is projected, these pixels being "covered" or partially "covered" by the projection **31** of the mesh elements. These pixels **301** to **30i** are illustrated in gray on FIG. 3.

[0050] As illustrated on FIG. 2 by the double arrow **201**, the mesh element **10** is first textured with a first texture information of image **110**. This first texture information is then projected onto all first images **110** to **140**, which is illustrated by the arrows **201**, **202**, **203** and **204**. The projection of the mesh element **10** (textured with the first texture information of corresponding pixels of first image **110**) onto first images **110**, **120**, **130** for which the mesh element **10** is entirely visible is illustrated by FIG. 3 and the projection of the mesh element **10** (textured with the first texture information of corresponding pixels of first image **110**) onto the first image **140** for which the mesh element **10** is only partially visible is illustrated by FIG. 4. The projection of the mesh element **10** onto one first image **110** to **140** corresponds to a conversion of the coordinates of the mesh element **10** from the world space into the image space associated to the first image onto which the mesh element is projected.

[0051] On FIG. 3, the image **30** corresponds to a first image (first images **110**, **120** or **130**) on which the mesh element **10** is projected. The projection of the mesh element **10** corresponds to the mesh element referenced **31**. Pixels in gray **301** to **30p** represents the p pixels (p being an integer) of the projection image covered at least partially by the projection **31** of the mesh element **10**. On FIG. 4, the image **40** corresponds to the first image **140** of FIG. 2 on which the mesh element is only partially visible, a part of the mesh element **10** being hidden by another mesh element **11** illustrated on FIG. 1. The projection of the mesh element **10** on image **40** is referenced **41** and the projection of the mesh element **11** on image **40** is referenced **42**. According to this example, the number q (q being an integer) of pixels **401** to **40q** of image **40** at least partially covered by the projection **41** of the mesh element **10** are less than the number p of the pixels of image **30** (p>q) (with a same resolution for all first images **110** to **140**).

[0052] After the projection of the first texture information of first image **110** associated with the mesh element **10** on the first images **110** to **140**, the first texture information is compared with the texture information associated with the pixels of each first image **110** to **140** corresponding to the projection of the mesh elements. The pixels of a first image **110** to **140** corresponding to the projection of the mesh element **10** also correspond to the pixels of this first image for which the mesh element is visible according to the point of view associated with this image. The result of the comparison advantageously corresponds to the distortion between the projected first texture information and the texture information of the pixels of the image onto which the mesh element is projected, and is:

$$D_{i,j} = \|M_{I_i \rightarrow V_j} - I_j\|^2 \quad \text{equation 1}$$

[0053] Where:

[0054] $M_{I_i \rightarrow V_j}$ corresponds to the projection onto the image V_j (for example one of the image **120**, **130** or **140**) of the mesh element **10** textured with the first texture information of image I_i (for example first image **110**)

[0055] $D_{i,j}$ corresponds to the distortion between the texture information of images i and j.

[0056] The first texture information of image **110** associated with the mesh element **10** is compared with each texture information of each first image **110** to **140**. According to a variant, the first texture information is compared with each texture information of each first image except with the first image **110**, the distortion between the projected first texture

information (of first image **110**) and the texture information of image **110** being equal to 0 or close to 0.

[0057] The same process is repeated by associating a first texture information extracted from the first image **120** with the mesh element **10**, by projecting this new first texture information onto the first images **110** to **140** and by comparing it with the texture information associated with the pixels corresponding to the projection of the mesh element **10** onto the first images **110** to **140** as to determine the distortion between the projected first texture information and each first image. This process is reiterated by associating first texture information extracted from each first image **130** and **140** as to determine the distortion between each projected first texture information and each first image. According to a variant, the distortion is determined between each first texture information (from each first image **110** to **140**) with only a part of the first images (by excepting the first image providing the first texture information to be projected onto the first images for comparison). A non limitative example of an algorithm used for determining the result of comparison between the first texture information and the first images is as follow:

```
// Textures traversal
for each texture Ii do
  // Views traversal
  for each view Vj do
    Compute MTi→Vj: projection of the mesh textured with texture Ii
    onto view Vj.
    Di,j = ||MTi→Vj - Ij||2
  end
end
end
```

[0058] For each first texture information associated with the mesh element **10** (extracted from each first image **110** to **140**), an error is determined and associated with respectively each first texture information. The determination of the error associated with a first texture information is based on the comparison results between the projected first texture information and the texture information associated with the pixels of the first images onto which the textured mesh element is projected. The calculation of an error information associated with the mesh element (T) **10** textured with a first texture information originating from a first image I_i (i corresponding to a first image, i being equal to 4 according to the non limitative example of FIGS. **1** and **2**) is performed for example by using the following equation:

$$Error_{T,i} = \sum_{j=1}^n \left(\sum_{(q,l) \in M_{T,V_j}} D_{i,j}[u, v] \right) \quad \text{equation 2}$$

[0059] u, v corresponding to the coordinates of pixel in an image I having a resolution of q×l pixels.

[0060] One first texture information is then selected among the plurality of first texture information according to the errors associated with each first texture information. The selected first texture information corresponds advantageously to the first texture information having the least error value. A first texture information is for example selected according to the following equation:

$$\hat{I}_T = \text{Error}_{T,i} (I_i \in \mathcal{F})$$

[0061] \mathcal{F} corresponding to the set of first texture information (available in the first images **110** to **140**).

[0062] A non limitative example of an algorithm used for selecting the first texture information which minimizes the error is as follow:

```
// Triangles traversal
for each triangle T do
  // Views traversal
  for each view Vj do
    Determine MTi→Vj the projection of the triangle T onto Vj where
    T is visible.
  end
  // Textures traversal
  for each texture Ii do
    Compute ErrorT,i.
  end
  Determine  $\hat{I}_T$ .
end
```

[0063] The selection of a first texture information among the set of candidate first texture information enables to assign one single texture information to a mesh element for every image in which the mesh element is partially visible, which reduces the amount of data to be coded and transmitted to any display device displaying the images or to any device decoding the data representative of the images to be displayed. The selection of the first texture information minimizing photometric error also enables to reduce artefacts which could appear if the texture associated to a mesh element would correspond to an arbitrary selection of one candidate texture information or to a mix of the candidate texture information.

[0064] Advantageously, a first texture information is assigned to each mesh element of the mesh or to a part of the mesh elements of the mesh.

[0065] FIG. **5** diagrammatically shows a hardware embodiment of a device **5** adapted and configured for texturing a mesh of a surface **1** and for the generation of display signals of one or several images, according to a particular and non limitative embodiment of the invention. The device **5** corresponds for example to a personal computer PC, a laptop, a tablet or a mobile phone.

[0066] The device **5** comprises the following elements, connected to each other by a bus **55** of addresses and data that also transports a clock signal:

[0067] a microprocessor **51** (or CPU),

[0068] a graphics card **52** comprising:

[0069] several Graphical Processor Units (or GPUs) **520**,

[0070] a Graphical Random Access Memory (GRAM) **521**,

[0071] a non-volatile memory of ROM (Read Only Memory) type **56**,

[0072] a Random Access Memory or RAM **57**,

[0073] one or several I/O (Input/Output) devices **54** such as for example a keyboard, a mouse, a webcam, and

[0074] a power source **58**.

[0075] The device **5** also comprises a display device **53** of display screen type directly connected to the graphics card **52** to display notably the displaying of synthesized images calculated and composed in the graphics card, for example live. The use of a dedicated bus to connect the display device **53** to the graphics card **52** offers the advantage of having much greater data transmission bitrates and thus reducing the latency time for the displaying of images composed by the

graphics card. According to a variant, a display device is external to the device **5** and is connected to the device **5** by a cable transmitting the display signals. The device **5**, for example the graphics card **52**, comprises a means for transmission or connection (not shown in FIG. **5**) adapted to transmit a display signal to an external display means such as for example an LCD or plasma screen or a video-projector.

[0076] It is noted that the word “register” used in the description of memories **52**, **56** and **57** designates in each of the memories mentioned, both a memory zone of low capacity (some binary data) as well as a memory zone of large capacity (enabling a whole program to be stored or all or part of the data representative of data calculated or to be displayed).

[0077] When switched-on, the microprocessor **51** loads and executes the instructions of the program contained in the RAM **57**.

[0078] The random access memory **57** notably comprises:

[0079] in a register **530**, the operating program of the microprocessor **51** responsible for switching on the device **5**,

[0080] calibration parameters **571** representative of intrinsic and extrinsic parameters of the acquisition devices used for acquiring the images **110** to **150** representative of the scene to be modelled;

[0081] data **572** representative of the texture associated with the images **110** to **150**, for example RGB information associated with the pixels of the images **110** to **150**;

[0082] data **573** representative of the mesh (for example the coordinates of the vertex of the mesh elements of the mesh) of a surface representative of the scene represented in the images **110** to **150**;

[0083] data **574** representative of the visibility of the mesh elements in the image **110** to **150** (for example a label visible and a label non visible associated with the pixels of the image and with the mesh elements of the mesh) when these data are available.

[0084] The algorithms implementing the steps of the method specific to the invention and described hereafter are stored in the memory GRAM **57** of the graphics card **52** associated with the device **5** implementing these steps. When switched on and once the parameters **570** representative of the environment are loaded into the RAM **57**, the graphic processors **520** of the graphics card **52** load these parameters into the GRAM **521** and execute the instructions of these algorithms in the form of microprograms of “shader” type using HLSL (High Level Shader Language) language or GLSL (OpenGL Shading Language) for example.

[0085] The random access memory GRAM **521** notably comprises:

[0086] in a register **5210**, the data representative of the texture associated with the images **110** to **150**,

[0087] data **5211** representative of a depth information (which is estimated or received with the texture information) associated with the pixels of the images **110** to **150**,

[0088] data **5212** representative of the mesh elements (for example coordinates of the vertices),

[0089] parameters **5213** representative of the visibility information associated with the mesh elements of the mesh, and

[0090] data **5214** representative of an error information associated with the candidate first texture information to be assigned to a given mesh element,

[0091] parameter **5215** representative of the selected first texture information among the plurality of candidate first texture information.

[0092] According to a variant, a part of the RAM **57** is assigned by the CPU **51** for storage of the values **5211** to **5214** and the parameters **5215** if the memory storage space available in GRAM **521** is insufficient. This variant however causes greater latency time in the composition of an image comprising a representation of the environment **1** composed from microprograms contained in the GPUs as the data must be transmitted from the graphics card to the random access memory **57** passing by the bus **55** for which the transmission capacities are generally inferior to those available in the graphics card for transmission of data from the GPUs to the GRAM and vice-versa.

[0093] According to another variant, the power supply **58** is external to the device **5**.

[0094] FIG. **6** illustrates a method for texturing a mesh of a surface **1** implemented for example in a computation unit illustrated on FIG. **5**, according to a particular and non limitative embodiment of the invention.

[0095] During an initialisation step **60**, the different parameters of the device **5** are updated. In particular, the parameters representative of the images and associated depth maps are initialised in any way.

[0096] Then, during a step **61**, for each first image of a set of first images, a first texture information is associated with the mesh element. A first image corresponds to an image of a plurality of images representing a scene according to several points of view, in which image a given mesh element is visible or at least partially visible. The visibility of the mesh element is determined by projecting it in the first images and comparing the depth of the mesh element (determined by converting the coordinates of the mesh element from the world space to the image space) with the depth information associated with the pixels of the images onto which the mesh element is projected. According to a variant, the visibility information associated with the mesh element and each image is received in addition to the data representative of the images (texture information, for example RGB information associated with the pixels of the images). As the mesh element is visible in at least two first images, there is several first texture information (called candidate first texture information) which is associated with the mesh element.

[0097] Then, during a step **62**, the candidate first texture information is each projected onto the first images. According to a variant, the candidate first texture information are each projected onto only a part of the first images, for example on all first images excepted the first image providing the candidate first texture information which is projected on the first images. A projection of a mesh element onto a first image corresponds to a conversion of the coordinates of a mesh element from the world space to the image space of the first image.

[0098] Then, during a step **63**, an error information is estimated for each candidate first texture information. The error information is computed according to the results of a comparison between the projected first texture information and the texture information associated with the first image on which is projected the candidate first texture information. The comparison is performed for each first image onto which the candidate first texture information is projected. According to a non limitative example, the error information associated with a candidate first texture information corresponds to the

sum of all comparison results between the candidate first texture information and the texture information associated with the first images onto which the candidate first texture information is projected. According to a variant, the error information associated with a candidate first texture information corresponds to the average of all comparison results between the candidate first texture information and the texture information associated with the first images onto which the candidate first texture information is projected.

[0099] Lastly, during a step **64**, one of the candidate first texture information is selected among the plurality of candidate first texture information according to the error information estimated for and associated with each candidate first texture information. Advantageously, the selected first texture information corresponds to the candidate first texture information having the least value of error information.

[0100] Naturally, the invention is not limited to the embodiments previously described.

[0101] In particular, the invention is not restricted to a method for texturing a mesh but extends to the computation unit implementing such a method and to the display device or mobile device comprising a computation unit implementing the texturing of a mesh or the display of the images resulting from the texturing process. The invention also concerns a method for selecting a texture information to be assigned to a mesh element among a plurality of candidate texture information and also to a method for coding and transmitting the selected texture information associated with a mesh of a 2D or 3D representation of a scene in a multi-view system.

[0102] The implementation of calculations necessary to the texturing of the mesh and to the selection of a texture information to be assigned to a mesh element is not limited either to an implementation in shader type microprograms but also extends to an implementation in any program type, for example programs that can be executed by a CPU type microprocessor.

[0103] The use of the invention is not limited to a live utilisation but also extends to any other utilisation, for example for processing known as postproduction processing in a recording studio for the display of synthesis images for example. The implementation of the invention in postproduction offers the advantage of providing an excellent visual display in terms of realism notably while reducing the required calculation time.

[0104] The implementations described herein may be implemented in, for example, a method or a process, an apparatus, a software program, a data stream, or a signal. Even if only discussed in the context of a single form of implementation (for example, discussed only as a method), the implementation of features discussed may also be implemented in other forms (for example, an apparatus or program). An apparatus may be implemented in, for example, appropriate hardware, software, and firmware. The methods may be implemented in, for example, an apparatus such as, for example, a processor, which refers to processing devices in general, including, for example, a computer, a microprocessor, an integrated circuit, or a programmable logic device. Processors also include communication devices, such as, for example, computers, cell phones, portable/personal digital assistants ("PDAs"), and other devices that facilitate communication of information between end-users.

[0105] Implementations of the various processes and features described herein may be embodied in a variety of different equipment or applications, particularly, for example,

equipment or applications associated with data encoding, data decoding, view generation, texture processing, and other processing of images and related texture information and/or depth information. Examples of such equipment include an encoder, a decoder, a post-processor processing output from a decoder, a pre-processor providing input to an encoder, a video coder, a video decoder, a video codec, a web server, a set-top box, a laptop, a personal computer, a cell phone, a PDA, and other communication devices. As should be clear, the equipment may be mobile and even installed in a mobile vehicle.

[0106] Additionally, the methods may be implemented by instructions being performed by a processor, and such instructions (and/or data values produced by an implementation) may be stored on a processor-readable medium such as, for example, an integrated circuit, a software carrier or other storage device such as, for example, a hard disk, a compact diskette ("CD"), an optical disc (such as, for example, a DVD, often referred to as a digital versatile disc or a digital video disc), a random access memory ("RAM"), or a read-only memory ("ROM"). The instructions may form an application program tangibly embodied on a processor-readable medium. Instructions may be, for example, in hardware, firmware, software, or a combination. Instructions may be found in, for example, an operating system, a separate application, or a combination of the two. A processor may be characterized, therefore, as, for example, both a device configured to carry out a process and a device that includes a processor-readable medium (such as a storage device) having instructions for carrying out a process. Further, a processor-readable medium may store, in addition to or in lieu of instructions, data values produced by an implementation.

[0107] As will be evident to one of skill in the art, implementations may produce a variety of signals formatted to carry information that may be, for example, stored or transmitted. The information may include, for example, instructions for performing a method, or data produced by one of the described implementations. For example, a signal may be formatted to carry as data the rules for writing or reading the syntax of a described embodiment, or to carry as data the actual syntax-values written by a described embodiment. Such a signal may be formatted, for example, as an electromagnetic wave (for example, using a radio frequency portion of spectrum) or as a baseband signal. The formatting may include, for example, encoding a data stream and modulating a carrier with the encoded data stream. The information that the signal carries may be, for example, analog or digital information. The signal may be transmitted over a variety of different wired or wireless links, as is known. The signal may be stored on a processor-readable medium.

[0108] A number of implementations have been described. Nevertheless, it will be understood that various modifications may be made. For example, elements of different implementations may be combined, supplemented, modified, or removed to produce other implementations. Additionally, one of ordinary skill will understand that other structures and processes may be substituted for those disclosed and the resulting implementations will perform at least substantially the same function(s), in at least substantially the same way(s), to achieve at least substantially the same result(s) as the implementations disclosed. Accordingly, these and other implementations are contemplated by this application.

1. A method for texturing a mesh, the mesh being associated with a surface representative of a scene captured accord-

ing to a plurality of points of view, an image of the scene comprising a texture information being associated with each point of view, wherein at least a mesh element of said mesh is at least partially visible from at least two first images of the plurality of images, the method comprising:

- for each first image, association of a first texture information with said mesh element,
- projection of the first texture information in said at least two first images,
- for each first texture information, estimation of an error information according to a comparison result between the projected first texture information and the texture information of said at least two first images, the error information being representative of a difference between the projected first texture information and the texture information of said at least two first images, and
- selection of one of the at least two first texture information according to said error information.

2. The method according to claim 1, wherein the error information associated with a first texture information corresponds to a sum of the comparison results between the projected first texture information and each said at least two first images.

3. The method according to claim 1, wherein the selected first texture information corresponds to the first texture information having the least error information.

4. The method according to claim 1, wherein the method further comprises:

- projecting said mesh element in at least a part of said images of the scene,
- for each image of the at least a part of images, comparing a depth information associated with the projected mesh element and a depth information associated with the image,
- for each image of said at least a part of images, determining a visibility information associated with said mesh element according to the result of the comparison of previous step, a mesh element being labeled as at least partially visible or not visible in said image according to said visibility information.

5. The method according to claim 4, wherein said projection of said mesh element corresponds to a conversion of the coordinates of the mesh element from world space into the image space of the image.

6. The method according to claim 1, wherein each image comprises a plurality of pixels, video information being associated with the plurality of pixels.

7. The method according to claim 1, wherein each image is represented by using Multi-view Video plus Depth (MVD) representation.

8. A computation unit configured to texture a mesh, the mesh being associated with a surface representative of a scene captured according to a plurality of images, each image comprising texture information, wherein at least a mesh element of said mesh is at least partially visible from at least two first images of the plurality of images, the computation unit comprising at least one processor configured for:

- associating, for each first image, a first texture information with said mesh element,
- projecting the first texture information in said at least two first images,
- estimating, for each first texture information, an error information according to a comparison result between the projected first texture information and the texture information of said at least two first images, the error information being representative of a difference between the projected first texture information and the texture information of said at least two first images, and
- selecting one of the at least two first texture information according to said error information.

9. The computation unit according to claim 8, wherein the error information associated with a first texture information corresponds to a sum of the comparison results between the projected first texture information and each said at least two first images.

10. The computation unit according to claim 8, wherein the selected first texture information corresponds to the first texture information having the least error information.

11. The computation unit according to claim 8, wherein the at least one processor is further configured for:

- projecting said mesh element in at least a part of said images of the scene,
- comparing, for each image of the at least a part of images, between a depth information associated with the projected mesh element and a depth information associated with the image,
- determining, for each image at said least a part of said images, a visibility information associated with said mesh element according to the result of the comparison, a mesh element being labeled as at least partially visible or not visible in said image according to said visibility information.

12. The computation unit according to claim 11, wherein said means for projecting said mesh element comprise means for converting the coordinates of the mesh element from world space into the image space of the image.

* * * * *