(12) UK Patent Application (19) GB (11) 2 358 980 (13) A

(43) Date of A Publication 08.08.2001

(21) Application No **0002771.4**

(22) Date of Filing **07.02.2000**

(71) Applicant(s)
British Broadcasting Corporation
(Incorporated in the United Kingdom)
Broadcasting House, LONDON, W1A 1AA, United Kingdom

(72) Inventor(s)
Graham Alexander Thomas
Richard Fawconer Stevens

(74) Agent and/or Address for Service
Mathys & Squire
100 Grays Inn Road, LONDON, WC1X 8AL, United Kingdom

(51) INT CL[7]
H04N 13/00

(52) UK CL (Edition S)
H4F FDD
H4T TBAS T131

(56) Documents Cited
GB 2325836 A **GB 2316256 A** **GB 2296152 A**
GB 2273577 A **GB 2272597 A** **GB 2206763 A**
WO 91/16664 A1

(58) Field of Search
UK CL (Edition R) **H4F FDD , H4T**
INT CL[7] **H04N 13/00**
ONLINE: EPODOC, WPI, JAPIO.

(54) Abstract Title
Processing of images for 3D display.

(57) A method and display are disclosed in which a representation of a 3D model is provided for presentation as a 3D image. The image may be presented under an array of spherical or lenticular microlenses 1-5 so that different images are presented at different viewing angles A-D. The images are rendered using a set of orthographic projections; this can avoid the need for multiple cameras or the highly computer intensive processing associated with generation of simulated camera images and can also give improved results as compared to prior art multiview images.

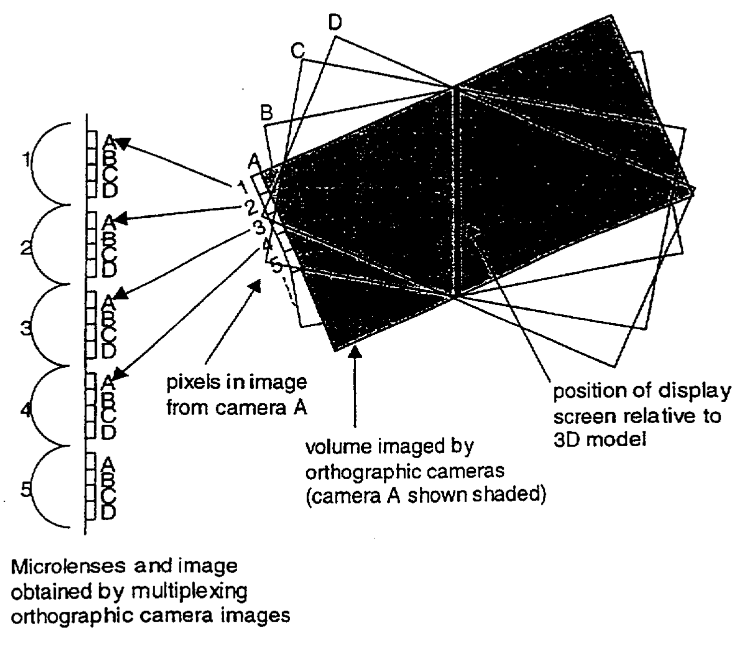

Figure 3

GB 2 358 980 A

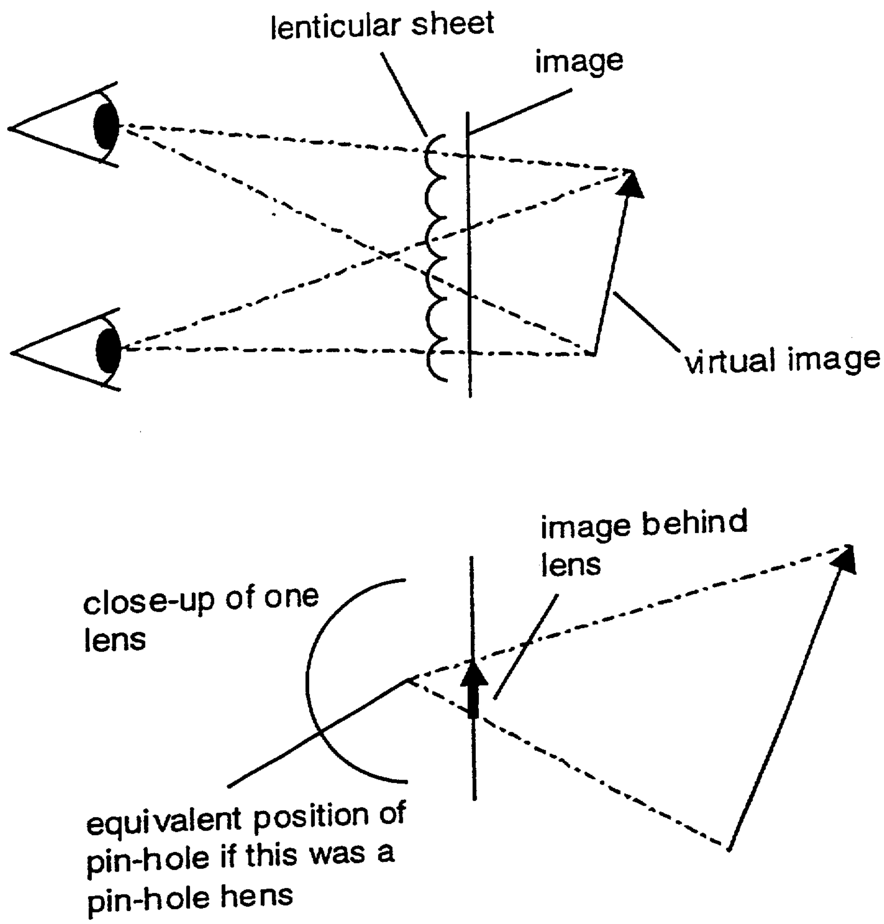
lenticular sheet
image
virtual image
image behind
lens
close-up of one
lens
equivalent position of
pin-hole if this was a
pin-hole hens

Figure 1

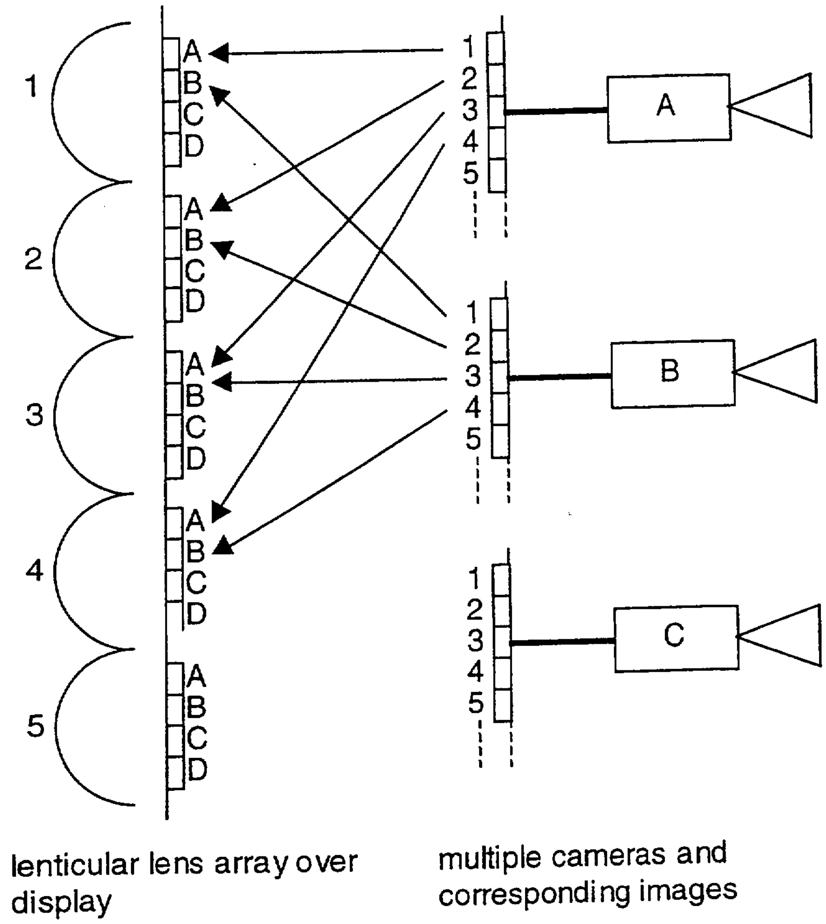
1
2
3
4
5
A
B
C
D
A
B
C
D
A
B
C
D
A
B
C
D
A
B
C
D
1
2
3
4
5
A
1
2
3
4
5
B
1
2
3
4
5
C
lenticular lens array over display
multiple cameras and corresponding images

Figure 2

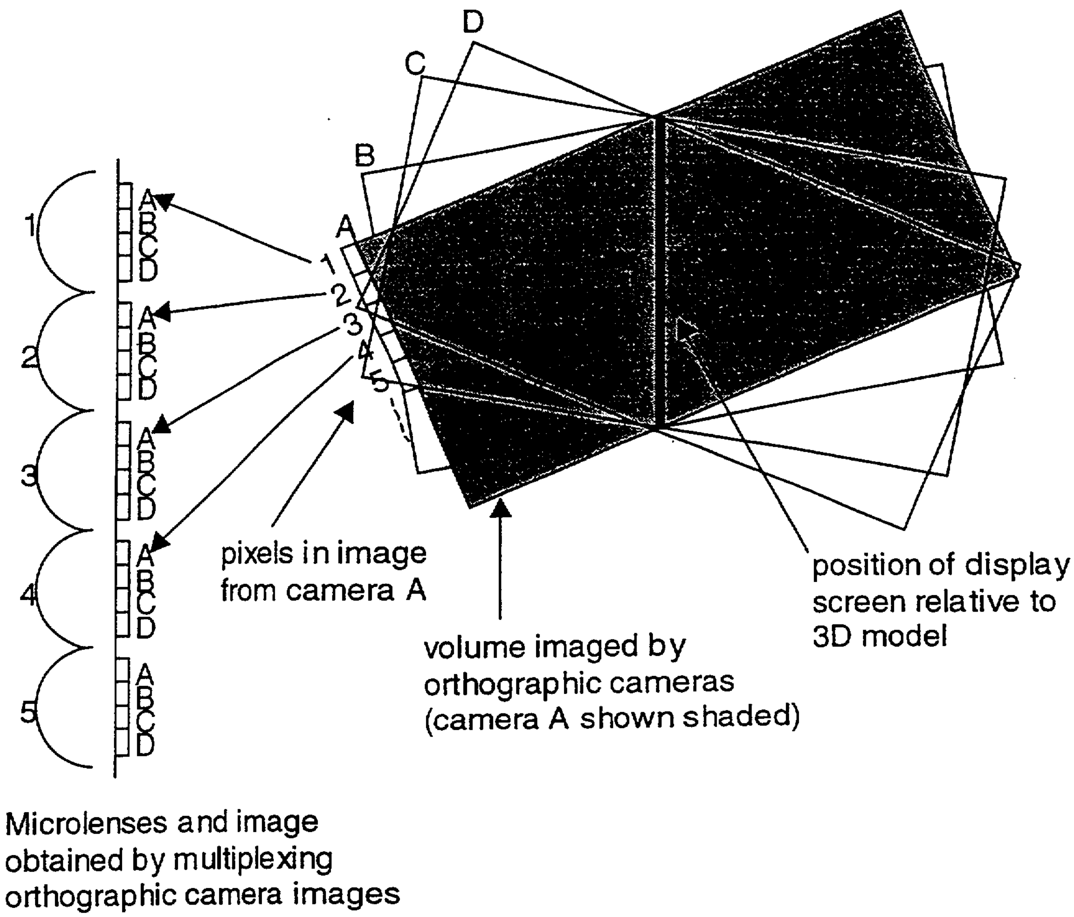
D
C
B
A
1
2
3
4
5
A
B
C
D
pixels in image
from camera A
volume imaged by
orthographic cameras
(camera A shown shaded)
position of display
screen relative to
3D model
Microlenses and image
obtained by multiplexing
orthographic camera images

Figure 3

PROCESSING OF IMAGES FOR 3D DISPLAY

This invention relates to the field of 3D displays.

One known method of producing a 3D image that exhibits full parallax is integral imaging, an example of which is described by McCormick et al., "Examination of the requirements for autostereoscopic, full parallax, 3D TV", IBC '94, IEE Conference Publication 397, September 1994, pp.477-482, and WO-A-9534018. In this method, a special kind of camera captures an image which is 'replayed' by placing a sheet of microlenses over the image. When the image is viewed in this way, each microlens appears to have a colour and brightness that is a function of the angle at which it is viewed. The form of the image underneath the microlens array is such that the overall appearance is of a 3D image. The microlens array should ideally be composed of spherical microlenses, although lenticular lenses can also be used, in which case the image only exhibits parallax in the direction normal to the lenticular elements.

It is known that the image underneath the microlens array can be generated by a computer, rather than by a special kind of camera. This allows a 3D image of a computer model to be generated. However, techniques for doing this have involved ray tracing and simulation of the behaviour of the integral image camera, and are therefore complex and processor-intensive. For example, the image underneath each microlens can be computed by rendering an image of the scene from the viewpoint of that microlens, with an appropriate modification to allow objects that are on both sides of the image plane to be visible. For a typical microlens array consisting of, say 400 by 300 microlenses, this would require 120,000 separate rendering processes.

Another approach to computing the image to place underneath the microlens array is to use a number of discrete images of the scene taken with a conventional camera, or generated using conventional 3D rendering software. The image underneath each microlens is formed by taking one pixel from each image in turn, so that the display presents the viewer with each image in sequence as the viewer's head moves parallel to the image. An example of this "multiview" approach is described in "Design and applications of multiview 3D-LCD" by C. van Berkel et al, 1996 EuroDisplay Conference, Birmingham, October 1996, and "Characterisation and optimisation of 3D-LCD module design" by C. van Berkel and J. Clarke, SPIE International Conference on Electronic Imaging, San Jose, 11-14 February 1997. This method has the advantage that conventional cameras or image rendering systems may be used, since the number of separate images required is relatively small, being equal to the number of pixels underneath a microlens. However, this method does not produce a fully accurate 3D representation of the scene; it merely presents the viewer with a number of discrete views. Often, the viewer will be sufficiently close to the screen that one of his eyes is presented with parts of one image from microlenses on the left of the screen and parts of another image from microlenses on the right, due to the changing angle between the screen and the direction of view. These combined views will not represent an accurate portrayal of the scene.

The present invention seeks to provide - in one of its forms - a method of allowing conventional image rendering hardware to be used efficiently to produce the image to place underneath the microlens array, whilst maintaining the accurate representation of the scene of which the integral imaging method is capable.

Accordingly, the present invention consists, in one aspect, in a method of providing a viewable representation of a 3D model, comprising the steps of identifying a set of different viewing angles and rendering from said model a set of orthographic projections corresponding respectively with said viewing angles.

Advantageously, wherein the set of orthographic projections are spatially multiplexed to provide a set of micro images, each micro image containing one pixel from each of the projections.

In another aspect, the present invention consists in representation of a 3D scene or model for display utilising 2D display means, comprising a set of 2D images representing respective orthographic projections of the 3D scene or model at respective different angles.

In still another aspect, the present invention consists in a method of processing a viewable representation of a 3D model or scene, comprising the steps of demultiplexing the representation to form a set of orthographic projections corresponding respectively with different viewing angles, and processing each projection.

Advantageously, the processed projections are remultiplexed.

The present invention recognises that each pixel underneath a spherical microlens in a true integral image represents the brightness and colour of the scene when viewed from the position of that microlens, at an angle that is a function of the position of the pixel with respect to the focal point of the microlens. We will assume for now that the arrangement of pixels underneath each microlens are the same for all microlenses. Therefore, for example, the top left-hand pixel underneath each microlens corresponds to the light ray leaving the scene at a given angle, and this angle is the same for all top left-hand pixels under all microlenses. The position at which the light ray intersects the display surface varies according to the position of the microlens, but the angle does not vary. Thus, if we were to take the top left-hand pixel from under each microlens and assemble them into an image, we would obtain a kind of image of the scene, similar to that which would be obtained by a camera a very long way away, having a very narrow field-of-view, such that the angle of view was essentially unchanging across the image.

There exists a well-known method for generating an image of a 3D model held in a computer for which the angle of view is constant for all pixels in the image. The method is called orthographic projection. It is supported, for example, by the "OpenGL" 3D graphics programming interface, the "Open Inventor" programming API, and many 3D graphics hardware systems. Therefore, the present invention provides through its use of orthographic projection, a convenient and efficient method of generating the image to place behind a microlens sheet in order to provide an accurate 3D view of a 3D scene.

The invention will now be described by way of example with reference to the accompanying drawings, in which:

Figure 1 is a diagram illustrating the known principle of integral image display;

Figure 2 is a diagram illustrating the known use of multiple cameras to generate a multi-view image viewed through a lenticular screen; and

Figure 3 is a diagram illustrating an embodiment of the present invention.

Referring briefly to Figure 1, this illustrates the known technique mentioned above, in which a special, integral image camera can provide a 2D representation, which when viewed through the matching microlens array, creates a 3D display with full parallax.

Figure 2 illustrates the other known technique mentioned above in which multiple cameras are used to generate a multi-view image to be viewed through a lenticular screen.

A 3D display according to this invention may be realised using an arrangement shown in Fig. 3. The figure shows a cross-section view looking down onto the display; it is assumed that the lenses are spherical, and that the view from the side would look similar. The volume imaged by each orthographic camera is represented as a rectangle. The orientation of each viewing volume corresponds to the angle of view of a particular pixel under each microlens. For example, the orientation of the shaded rectangle, corresponding to 'camera' A, corresponds to the angle at which the pixels labelled 'A' are visible through the microlenses. The width of each view corresponds to the width of the display screen when viewed at that angle. The extent of each viewing volume to the front and behind of the display corresponds to the range of depths between which objects are to be visible. Information describing the orientation and size of each viewing volume is given to the computer rendering system in order to allow it to compute each orthographic view. The image computed from the view corresponding to area A in Fig. 3 is shown as a row of pixels on the left-hand side of the viewing volume. These pixels are multiplexed with pixels from the other views and displayed on the appropriate pixels underneath each microlens.

The arrows in Fig. 3 show how the pixel values from view A are used to form the first pixel underneath each microlens; the pixels from view B are similarly used to form the second pixel underneath each lens, and so on. For example, with a regular array of spherical microlenses with a spacing of 8 pixels horizontally and 8 pixels vertically, it would be necessary to render a total of 64 orthographic images. The resolution of each orthographic image would be equal to the number of microlenses in the display.

The operation of the orthographic camera may be represented as a 4x4 matrix, known as the projection matrix in openGL. This is described, for example, in the "OpenGL Reference Manual", published by Addison-Wesley. The standard form of the matrix used to create an orthographic view is:

$$\begin{matrix} 2/(r-l) & 0 & 0 & Tx \\ 0 & 2/(t-b) & 0 & Ty \\ 0 & 0 & 2/(f-n) & Tz \\ 0 & 0 & 0 & 1 \end{matrix}$$

where

t, b define the positions of the planes at the top and bottom of the viewing volume, being equal to the x values of the top and bottom planes respectively, and the planes being normal to the x (vertical) axis;

l,r define the planes on the left and right of the viewing volume, being equal to the y values of the left and right-hand planes respectively, and the planes being normal to the y axis;

f,n define the planes on the far and near ends of the viewing volume, being equal to the z values of the far and near planes respectively, and the planes being normal to the z axis. The camera is positioned at the origin, looking along the z axis in the direction of -ve z. in accordance with the usual convention in OpenGL.

Tx = - (r+l) / (r-l)

Ty = - (t+b) / (t-b)

Tz = - (f+n) / (f-n)

This matrix premultiplies the coordinates of points in the 3D model being viewed, these points being represented in homogeneous coordinates as a vector of the form $[x, y, z, 1]^T$. So a point (x, y, z) is transformed to the position

$$((2x - r - l) / (r - l), (2y - t - b) / (t - b))$$

in the rendered image, with x coordinates ranging from -1 on the left of the image to +1 on the right, and y ranging from -1 at the bottom to +1 at the top. Note that the coordinates of the point in the image are independent of z; this is the fundamental characteristic of an orthographic projection.

Although the best quality 3D images are likely to be obtained by using spherical microlenses, an alternative arrangement is to use cylindrical microlenses, for example as used in the multiview LCD system shown in Fig. 2. This provides parallax only in the horizontal direction (if the lenticular lenses run vertically). However, lenticular lenses are cheaper and easier to manufacture than spherical microlens arrays, and may be preferable in some applications.

The invention may also be used to generate images for such lenticular lens displays, by using a virtual camera which exhibits the properties of an orthographic camera in the horizontal direction, and behaves like a perspective camera in the vertical direction. This may be achieved by using the expression for the y display coordinate given by the conventional perspective camera transformation (see the OpenGL Reference manual), i.e.

$$y = (-2ny/z + t + b) / (b-t)$$

Note that such a combined orthographic/perspective transformation cannot be implemented by multiplication with a viewing matrix, since the perspective transformation requires that the w coordinate of the transformed point (the 4th component of the position, in homogeneous coordinates) be equal to the depth value of the point (which is -z, since the camera looks in the direction of the negative z axis). OpenGL divides the transformed x, y, z coordinates by w to obtain screen coordinates; this allows the size of objects to appear inversely proportional to their distance from the camera. In this application, we require this division to happen for the y coordinate, but <u>not</u> the x coordinate. Therefore, the transformation we require in this case would have to be implemented by special-purpose hardware or software, since OpenGL divides <u>all</u> coordinates by w.

In some applications, the arrangement of pixels underneath the microlenses may vary across the display, for example if the spacing between lenses is not an integer multiple of the pixel spacing. In this case, the orthographic cameras may be used to generate images on a regular pixel lattice under each lens, and the resulting image data, after multiplexing, may be resampled using conventional filtering and sampling techniques. If the microlenses themselves are arranged in some way other than on a rectangular lattice, such as a hexagonal pattern, then the orthographically-rendered images may be resampled to provide pixel data at locations corresponding to the positions of the lenses, before the data is multiplexed to place under the microlenses. Alternatively, the lattice of microlens positions may be broken down into several subsets of points on a square lattice, and several sets of orthographic images may be rendered, to correspond to the lens positions in each lattice.

Filtering and resampling of the image data before or after multiplexing may also be used to apply 'anti-aliasing'. For example, the orthographic images may be rendered to a higher resolution than that corresponding to the number of microlenses in the image, and each orthographic image may be low-pass filtered and resampled to the required resolution, before the pixels are multiplexed. More orthographic images than there are pixels under each lens may be rendered, and the images under each microlens (after the pixels are multiplexed), may be low-pass filtered and resampled. Either or both of these approaches may be used to obtain higher quality images.

Similar filtering and resampling techniques may be applied to allow the use of a smaller number of orthographically-rendered images than there are pixels under each lens, by interpolating additional pixel values in the images under each microlens. Equally, the resolution of each rendered orthographic image may be lower than the number of microlenses in the image, and additional pixels interpolated in each orthographic image to match the number of microlenses.

The described techniques of filtering and resampling the othographically-rendered images can be extended to the representations provided by the known "real" integral image camera. Thus, such a representation formed with an N x M array of microlenses can be de-multiplexed to provide a set of orthographic projections. The number of such projections will be determined by the number of pixels under each microlens. The resolution of each projection will be N x M pixels. Each of these demultiplexed projections can then be subjected to a variety of conventional video processing techniques, as if they were true images. It would be possible, for example, by resampling to provide a set of images having different resolution, say P x Q. Remultiplexing would then provide a 2D representation which can be viewed under a P x Q microlens array. In other senses, the availability of demultiplexed othographic projections will enable types of processing which would otherwise be impossible or impracticable, or will improve significantly the performance of a specific technique. Thus, an orthographic projection will be expected to share the spatial correlation of a real image. Compression techniques such as JPEG and the intra-coding within MPEG should therefore be highly efficient. A compression technique which first demultiplexed the representation into a set of orthographic projections, should therefore be more efficient than one operating on spatial blocks of the representation itself. Related advantages will be found with other spatial processing techniques such as noise reduction. Similar effects should be seen in the temporal correlation of successive pictures. Thus, a succession of orthographic projections can be processed as if it were a video signal.

Although the invention has been described in terms of its application to a display using microlenses, it can be applied to any type of 3D display where each small area of the screen appears to have a brightness and colour that varies as a function of viewing angle. Such displays include those using multiple displays or high frame-rate displays with an associated mechanism to project particular images at particular angles.

CLAIMS

1. A method of providing a viewable representation of a 3D model, comprising the steps of identifying a set of different viewing angles and rendering from said model a set of orthographic projections corresponding respectively with said viewing angles.

2. A method according to Claim 1, wherein the set of orthograpic projections are spatially multiplexed to provide a set of micro images, each micro image containing one pixel from each of the projections.

3. A representation of a 3D scene or model for display utilising 2D display means, comprising a set of 2D images representing respective orthographic projections of the 3D scene or model at respective different angles.

4. A method of processing a viewable representation of a 3D model or scene, comprising the steps of demultiplexing the representation to form a set of orthographic projections corresponding respectively with different viewing angles, and processing each projection.

5. A method according to Claim 4, wherein the processed projections are remultiplexed.

11.

Application No: GB 0002771.4
Claims searched: 1-5
Examiner: Matthew Males
Date of search: 7 September 2000

Patents Act 1977
Search Report under Section 17

Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

UK Cl (Ed.R): H4F (FDD); H4T

Int Cl (Ed.7): H04N 13/00

Other: Online: EPODOC, WPI, JAPIO

Documents considered to be relevant:

Category	Identity of document and relevant passage		Relevant to claims
X	GB 2,325,836 A	(Hewlett-Packard) - whole document but see Abstract.	1 at least
X	GB 2,316,256 A	(Fujitsu) - whole document.	1, 3 & 4 at least
X	GB 2,296,152 A	(GEC-Marconi) - whole document but see Abstract and Figure 1.	1 at least
X	GB 2,273,577 A	(Sharp) - whole document but see page 1, second paragraph and page 2, paragraphs 4-6.	1, 3 & 4 at least
X	GB 2,272,597 A	(Sharp) - whole document but see Abstract and Figure 3.	1, 3 & 4 at least
X	GB 2,206,763 A	(Travis) - whole document but see Abstract, Figures 2 & 3, and page 3, lines 10-30.	1, 3 & 4 at least
X	WO 91/16664 A1	(Dominguez) - see Abstract, Figures.	1, 3 & 4 at least

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
		E	Patent document published on or after, but with priority date earlier than, the filing date of this application.
&	Member of the same patent family		