Object Recognition Method

Technical Field

The present invention relates to the field of object recognition in 2-dimensional and 3-dimensional space, as may be used as part of an artificial intelligence system.

Background

Associated technologies required for 3-dimensional object recognition have already entered the commercial market, such as 3-dimensional laser scanners and other non-laser reticle projection systems which enable a 3-dimensional object to be scanned and a 3-dimensional digital model to be generated. For example such systems enable the human face to be scanned for computer games and for the generation of 3-dimensional laser images in crystal.

Some current 3-dimensional object recognition systems rely upon environment-dependant conditions such as the distinct colouring of objects.

Other current 3-dimensional object recognition systems rely upon complete 3-dimensional imaging of an object in order to create a normalised set of data pertaining to the object, for example normalisation methods employing an inferred centre of mass of an object. This method is demonstrated in the paper “A Geometric Approach to 3D object Comparison”, Novotni et al, smi, pp.0167, International Conference on Shape Modelling and Applications, 2001.

Other current 3-dimensional object recognition systems rely upon the extraction of affine invariant surface patches, where these patches are normalised for object recognition. This method is demonstrated in the paper “3D Object Modelling and Recognition Using Affine-Invariant Patches and Multi-View Spatial Constraints”, Rothganger et al, Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 2, 2003, pp. II-272-7. This technology enables both the use of appearance and 3-dimensional structure in object recognition.

It is an object of the present invention to provide a 2-dimensional (2D) and/or 3-dimensional (3D) object recognition algorithm which will operate independently of the object position and orientation, and the viewpoint position and orientation, operate in a large range of lighting conditions and, operate efficiently. The object recognition method will also operate without the need for artificial colouring to be applied to the objects, will not require an object to be viewed from all angles during either the characterisation or recognition phase, will make use of the appearance in object recognition, and for 3D specific recognition, will make use of the 3-dimensional structure, will operate on objects without planar surfaces or sharp edges to aid identification, and will not require a complete 3D image of the object to be obtained.

Geometric hashing has been well documented as a means of recognising objects at arbitrary angles and positions within a scene, and the method according to the present invention is an extension of this method. See “Geometric Hashing: An Overview”, Wolfson et al, Computing in Science and Engineering, Vol. 4, 1997, pp 10-21.

It should be realized that the method according to the present invention will reduce in performance where object features (points of interest on the object) cannot be extracted. This can occur in the scenario where an object being recognised has both no distinct areas of high curvature in shape, for example corners, and no distinct areas of high curvature in surface texture. However there are other well known methods to identify object features. Also some "degenerate" objects may look the same regardless of the viewing angle and may still be recognisable regardless of the accuracy and repeatability of the object feature extraction method.

Summary of Invention

The present invention consists of a method for generating one or more sets of characterising data for one or more objects residing in one or more scenes in an N-dimensional space defined by an axis system passing through an origin, comprising the steps of, for each scene:

1. deriving a set of N-dimensional object data points for an object, each object data point defined by N-dimensional coordinates and one or more light intensity values, based on a first data-processing of the coordinates and the one or more light intensity values of the object data points, and also deriving a set of feature data points corresponding to object features of the object, each feature data point also defined by N-dimensional coordinates;
2. grouping combinations of the coordinates of the feature data points as the coordinates of apexes of one or more object triangles, each object triangle lying in a plane;
3. for each of one or more of the sides of each of the object triangles, applying a transformation function to the coordinates of one or more of the data points, including at least one object data point or one feature data point, to generate respective transformed data points with respective new coordinates in a new axis system, wherein the transformation function is a function of the coordinates of the apexes of the respective object triangle; and
4. for each of one or more of the sides of each of the object triangles, carrying out a second data-processing of the transformed data points, and generating a set of characterising data.

It is preferred that the one or more light intensity values correspond to one or more light frequencies, or frequency bands, or a function of these frequencies or frequency bands.

It is preferred that the object data points are located on the surface of the respective object and the one or more light intensity values characterize the light radiating from the object data points.

It is preferred that a unique object index is assigned to each of the objects.

It is preferred that the first data-processing in Step 1 comprises subdividing the object data points into at least one region of contiguous object data points with similar one or more light intensity values or a similar first function of their one or more light intensity values, deriving a base point for the region as a second function of the coordinates of the object data points in the region and, based on a third function of the base point, deriving feature data points.

It is preferred that the first function is a luminosity contrast function of the one or more light intensity values.

It is preferred that the second function is a geometric centroid of the coordinates of the object data points in the region.

It is preferred that the third function relates to the distance of one or more of the object data points in the region from the base point.

It is preferred that the third function is a relative and/or absolute minima and/or maxima of the distance of one or more of the object data points in the region from the base point.

It is preferred that one more of the object data points in the region are on a boundary of the region.

It is preferred that the base point is also a feature data point.

It is preferred that the third function is an equality function and, correspondingly, the base point is the only feature data point for the region.

It is preferred that the first data-processing in Step 1 comprises determining at least one point of high local curvature of the object data points, and designating the point as a feature data point.

It is preferred that the first data-processing in Step 1 comprises determining at least one point of high local curvature of the one or more light intensity values, and designating the point as a feature data point.

It is preferred that the object data points are derived by N-dimensionally imaging the scene from one or more viewpoints, each viewpoint characterised by a position and an orientation in the space and predefined viewing properties and, for each of the viewpoints, generating a 2-dimensional image of the object, each image comprising an array of pixels, each pixel corresponding to a viewable point in the respective scene and being characterised by one or more light intensity values and coordinates in the space and, either during the calculating of the coordinates for each viewpoint or after the calculating of the coordinates for all viewpoints, isolating the object data points for the object.

It is preferred that the object data points are isolated by determining boundaries of the object in the image based on a boolean depth contrast map derived by applying an arbitrary threshold to a depth contrast map of the image, or a boolean depth gradient contrast map derived by applying an arbitrary threshold to a depth gradient contrast map of the image, or a boolean luminosity (overall light intensity) contrast map derived by applying an arbitrary threshold to a luminosity contrast map of the image, or linear or non-linear functional combination of these maps.

It is preferred that the position and the orientation of each viewpoint in the space are predefined.

It is preferred that at least one of the viewable points reside on the surface of the object.

It is preferred that the object data points are derived by 3-dimensional imaging of the scene, the space is a 3-dimensional space, and the coordinates in the space of the viewable point are 3-dimensional coordinates.

It is preferred that, for each viewpoint, the N-dimensional imaging comprises creating at least two 2-dimensional sub-images from two alternative viewpoints, slightly offset by a predetermined distance either side of the viewpoint, either sequentially or using at least two corresponding imaging systems, and a resulting parallax offset between the positions of the corresponding viewable point in the resulting pixels arrays of the sub-images, and/or the difference between the corresponding one or more light intensity values, is used to calculate the distance of the viewable point from the viewpoint.

It is preferred that the coordinates in the space of the viewable point is derived based on the location of the corresponding pixel in the array, the distance of the viewable point from the viewpoint, and the position, the orientation, and the viewing properties of the viewpoint.

It is preferred that the viewing properties comprise a view width angle and a view height angle.

It is preferred that the first data processing in Step 1 is performed using hardware acceleration hardware, such as a PC graphics card.

It is preferred that the space is a 3-dimensional space and the coordinates are correspondingly 3-dimensional coordinates, and the transformation function in Step 3 comprises shifting the position of the origin to a new origin position and reorientating the axis system to a new axis orientation such that, in the new axis system (X', Y', Z'), defined by a first new axis (X'), a second new axis (Y'), and a third new axis (Z'), the new axis orientation and new origin position are a function of the coordinates of the apexes of the respective object triangle.

It is preferred that the new axis system (X', Y', Z') is positioned and aligned such that the third new axis (Z') is aligned perpendicular to the plane of the respective object triangle, and passes through a mid-point between the two apexes at the extremities of the respective side of the object triangle, the first new axis (X') is aligned parallel to the side, and the second new axis (Y') is directed through the mid-point in the direction of the remaining apex of the object triangle.

It is preferred that the new axis orientation is a function of the 3-dimensional orientation of the respective side and the 3-dimensional orientation of the plane of the respective object triangle.

It is preferred that the space is a 2-dimensional space and the coordinates are correspondingly 2-dimensional coordinates, and the transformation function in Step 3 comprises transforming the coordinates of the data points such that, in the new axis system (X', Y'), defined by a first new axis (X') and a second new axis (Y'), the three apexes of the respective object triangle are coincident with the three apexes of a predefined triangle in the new axis system.

It is preferred that the predefined triangle is an equilateral triangle.

It is preferred that the three apexes of the respective object triangle are made coincident with the three apexes of the predefined triangle by scaling the coordinates of the data points in a first direction such that the respective side of the object triangle is of same length as a predefined side of the predefined triangle, scaling the coordinates of the data points in a second direction such that the perpendicular distance between the side of the object triangle and the corresponding remote apex of the object triangle and perpendicular distance between the predefined side of the predefined triangle and the corresponding remote apex of the predefined triangle are equal, then shearing the coordinates of the data points along an axis defined by the orientation of the side of the object triangle, and then translating the coordinates of the data points such that the three apexes of the object triangle are coincident with the three apexes of the predefined triangle.

It is preferred that a predefined side of the predefined triangle is parallel to the first new axis (X'), and the three apexes of the respective object triangle are made coincident with the three apexes of the predefined triangle by firstly scaling the coordinates of the data points such that the respective side of the object triangle is of same length as the predefined side of the predefined triangle, then rotating the coordinates of the data points such that the side of the object triangle is parallel with the first new axis (X'), then scaling the coordinates of the data points in the direction of second new axis (Y') such that the perpendicular distance between the side of the object triangle and the corresponding remote apex of the object triangle and perpendicular distance between the predefined side of the predefined triangle and the corresponding remote apex of the predefined triangle are equal, then shearing the coordinates of the data points in the direction of the first new axis (X'), and then translating the coordinates of the data points such that the three apexes of the object triangle are coincident with the three apexes of the predefined triangle.

It is preferred that the new axis system has an orientation and an origin position equivalent to that of the axis system.

It is preferred that, in Step 3, a unique transformation index is assigned to each of the one or more sides of each of the object triangles of each of the objects.

It is preferred that the transformation function in Step 3 is performed using hardware acceleration hardware, such as a PC graphics card.

It is preferred that the second data-processing in Step 4 comprises creating an M-dimensional image of the transformed data points.

It is preferred that the image is created by forming an interpolated or non-interpolated N-dimensional surface mesh defined by surface mesh points, the coordinates of each surface mesh point corresponding to a coordinate of, or an interpolation of one or more of the coordinates of, the transformed data points, each polygon of the surface mesh formed by proximate surface mesh points and having associated with it one or more light intensity values derived as a function of the one or more light intensity values of the nearby transformed data points.

It is preferred that the image comprises a depth map or a function of the depth map calculated based on the surface mesh.

It is preferred that the image comprises a luminosity map or a function of the luminosity map calculated based on the surface mesh.

It is preferred that the image is a 2-dimensional image.

It is preferred that the set of characterising data comprises one or more subsets of characterising data, each subset comprising coordinates of one or more of the transformed data points corresponding to feature data points, or a fourth function of the coordinates of one or more of the transformed data points corresponding to feature data points, or the M-dimensional image of the transformed data points, or a seventh function of the M-dimensional image of the transformed data points.

It is preferred that the set of characterising data comprises one or more subsets of characterising data, each subset comprising the coordinates of one or more of the transformed data points corresponding to feature data points, or a sixth function of the coordinates of one or more of the transformed data points corresponding to feature data points.

It is preferred that the seventh function of the M-dimensional image of the transformed data points is a spacial convolution function, and this is applied to the image to generate one or more convolution coefficients, or a fifth function of the one or more convolution coefficients.

It is preferred that the fourth or fifth functions are a binning function.

It is preferred that the sixth function is a binning function.

It is preferred that the binning function is also a function of a subset index number uniquely identifying a subset of characterising data such that, for one or more of the transformed data points corresponding to feature data points, or for one or more of the convolution coefficients, an adjacent bin is filled instead of an optimum bin.

It is preferred that the binning function is also a function of a subset index number uniquely identifying a subset of characterising data such that, for one or more of the transformed data points corresponding to feature data points, an adjacent bin is filled instead of an optimum bin.

It is preferred that the set of characterising data is represented as a 1-dimensional binary string.

It is preferred that the one or more subsets of characterising data are each represented as a 1-dimensional binary string.

It is preferred that the one or more subsets of characterising data are each represented as a 1-dimensional binary string.

It is preferred that the image is generated by generating a virtual image of the surface mesh points from a particular virtual viewpoint.

It is preferred that the image comprises mesh surface data points generated by interpolating the surface mesh at particular coordinate intervals along the first new axis (X'), or a function of the first new axis (X'), and along the second new axis (Y'), or a function of the second new axis (Y'), each mesh surface data point corresponding to a point on or off of the mesh surface and having one or more light intensity values derived as a function of the light intensity values of the transformed data points proximate to the point, and a coordinate derived as a function of the coordinates of the transformed data points proximate to the point.

It is preferred that the mesh surface data points of the image have their one or more light intensity values set to an arbitrary one or more light intensity values if, based in the new axis system, their coordinates lie outside of the triangle formed by the coordinates of the transformed data points corresponding to feature data points, in turn corresponding to the apexes of the object triangle.

It is preferred that the second data processing in Step 4 is performed using hardware acceleration hardware, such as in a PC graphics card.

It is preferred that the set of characterising data in Step 4 also includes data relating to the object index.

It is preferred that the set of characterising data in Step 4 also includes data relating to the transformation index.

It is preferred that the set of characterising data in Step 4 also includes data relating to the new axis system.

It is preferred that, in Step 4, the set of characterising data is generated for each combination of scene, object, object triangle, and side.

It is preferred that the one or more sets of characterising data are initially generated as training data, and another one or more sets of characterising data are also later generated as test data, and the training data and test data is then compared using an algorithm, thereby enabling the recognition of the one or more objects in the one or more scenes.

It is preferred that the one or more sets of characterising data is generated for multiple scenes as training data, and another one or more sets of characterising data is later generated for a different scene as test data, and the training data and test data is then compared using an algorithm, thereby enabling the recognition of the one or more objects in the different scene.

It is preferred that the algorithm comprises determining if one of the one or more sets of characterising data comprising the training data, and one of one or more sets of characterising data comprising the test data, are substantially equal.

It is preferred that the algorithm comprises using a neural network to determine the degree of equality of the one or more sets of characterising data comprising the training data, and the one or more sets of characterising data comprising the test data.

It is preferred that the one or more sets of characterising data each comprise subsets of characterising data, and the algorithm comprises determining if one or more of the subsets of characterising data comprising the training data, and the one or more subsets of characterising data comprising the test data, are substantially equal.

It is preferred that the one or more sets of characterising data each comprise subsets of characterising data, and the algorithm comprises using a neural network to determine the degree of equality of the one or more subsets of characterising data comprising the training data, and the one or more subsets of characterising data comprising the test data.

It is preferred that the seventh function of the M-dimensional image of the transformed data points is a neural network, or the fourth function of the coordinates of one or more of the transformed data points corresponding to feature data points is a neural network.

It is preferred that the fifth function of the one or more convolution coefficients is a neural network.

It is preferred that the seventh function of the M-dimensional image of the transformed data points is a decision tree, or the fourth function of the coordinates of one or more of the transformed data points corresponding to feature data points is a decision tree.

It is preferred that the fifth function of the one or more convolution coefficients is a decision tree.

It is preferred that the seventh function of the M-dimensional image of the transformed data points is an image processing function (such as a Gaussian filter) resulting in significant quantization (spatial resolution or intensity resolution reduction) of the output, and where this quantization is insensitive to slight (spatial or intensity) variations in the input data.

It is preferred that all three features points of the combination of the coordinates of the feature points grouped as the coordinates of apexes of an object triangle in Step 3 are derived from a single region.

It is preferred the operations in Step 3 and Step 4 performed for each of one or more of the sides of each of the object triangles are instead performed for each of one or more of the apexes of each of the object triangles, and where these operations are equivalent because the transformation function is a function of the coordinates of the apexes of the respective object triangle.

Brief Description of Drawings

Fig 1. shows a cubic object residing a 3-dimensional space and viewed from two viewpoints,

Fig 2. shows a flow diagram of the data generation method according to the present invention,

Fig 3. shows a 2D image produced from one of the viewpoints in Fig. 1,

Fig 4. shows diagrammatically a method of calculating the distance of a viewable point from a viewpoint using a parallax technique,

Fig 5. shows a depth map based on the image in Fig. 3 and other depth data,

Fig 6. shows the object data points produced from the image in Fig. 3,

Fig 7. shows an RGB map of the image in Fig. 3,

Fig 8. shows a luminosity map based on the RGB map in Fig. 7,

Fig 9. shows a luminosity contrast map based on the luminosity map in Fig. 8,

Fig 10. shows a depth contrast map based on the depth map in Fig. 5,

Fig 11. shows a depth gradient map based on the depth map in Fig. 5,

Fig 12. shows a depth gradient contrast map based on the depth gradient map in Fig. 11 and the depth contrast map in Fig. 10,

Fig 13. shows a region based upon the luminosity contrast map in Fig. 9,

Fig 14. shows a combination of relative minimas and maximas from the centroidal base point of the region in Fig. 13,

Fig 15. shows the corners map in Fig. 14, also with relevant object triangles,

Fig 16. shows the start of a 3D transformation function with the object triangle in the original coordinates system,

Fig 17. shows step A of the transformation function in which a third new axis (Z') is aligned perpendicular to the plane of the object triangle,

Fig 18. shows step B of the transformation function in which the third new axis (Z') is positioned such that it passes through the mid-point between the two apexes at the extremities of the side,

Fig 19. shows step C of the transformation function in which a first new axis (X') is aligned parallel to the side, and a second new axis (Y') is directed through the mid-point in the direction of the third apex of the object triangle,

Fig 20. shows step D of the transformation function in which the coordinates (x, y, z) of the data points are transformed into transformed data points, with new coordinates (x', y', z') in the new axis system,

Fig 21. shows the start of an alternative 2D transformation function with the object triangle and the predefined equilateral triangle shown with respect to both the original coordinates system and the new coordinates system, where a predefined side of the predefined triangle is parallel to the new X axis.

Fig 22. shows step A of the alternative transformation function in which the object data is scaled such that the side of the object triangle is of same length as a predefined side of the predefined triangle.

Fig 23. shows step B of the alternative transformation function in which the object data is rotated such that the side of the object triangle is parallel with the new X axis.

Fig 24. shows step C of the alternative transformation function in which the object data is scaled in the new Y axis direction such that the perpendicular distance between the side of the object triangle and the third apex of the object triangle and perpendicular distance between the side of the predefined triangle and the third apex of the predefined triangle are the same.

Fig 25. shows step D of the alternative transformation function in which the object data is sheared along the new X-axis.

Fig 26. shows step E of the alternative transformation function in which the object data is translated such that the object triangle is centred about a predefined point, the centre of the predefined triangle.

Fig 27. shows an interpolated 3D surface mesh with surface mesh points,

Fig 28. shows an interpolated 3D surface mesh with light intensity values of the mesh surface polygons,

Fig 29. shows the interpolated 3D surface mesh in Fig. 27 with the object triangles included, and

Fig 30. shows a 2D image of the interpolated 3D surface mesh in Fig. 27.

Best Mode for Carrying Out the Invention

Referring to Fig. 1, the data generation method according to the present invention will be described based, for simplicity, on a single cubic object 1 residing in a single scene 2 in a 3-dimensional cartesian space defined by an X, Y, Z axis system passing through an origin 3. Fig. 1 also shows two alternative viewpoints 4 and 5 of the object 1, each with respective viewpoint positions 6 and 7 and respective viewpoint orientations 8 and 9 in the space. It will be recognised by those skilled in the art of object recognition that the same method could also be used to generate data pertinent to an object where multiple objects reside in one or more scenes.

Referring to the flow diagram for the data generation method in Fig. 2, it can be seen that the method described can therefore be "looped through" for more than one scene, for more than one viewpoint and for more than one object. For example, it might be necessary in a particular application for a vision system of an industrial robot to be able to recognise a set of three objects, say objects O1, O2 and O3 in relatively complex scenes in which one or more of the objects O1, O2 and O3 are resident. The vision system could, for example, be trained by arranging the vision system to successively view O1 in a first scene from a number of viewpoint positions and orientations in that scene, then view O2 in a second scene also from a number of viewpoint positions and orientations, and then view O3 in the third scene also from a number of viewpoint positions and orientations. The characterising data generated for each object according to the present invention would then constitute "training data".

After this training process, the capability of the object recognition system of the industrial robot could be measured by arranging the (or another) vision system to view a scene comprising one or more of the objects O1, O2 and O3, plus also potentially previous "untrained" objects introduced in an attempt to contaminate the scene. The method according to the present invention allows the segmentation of the characterising data pertinent to a particular object (say O2), and then the comparison of this "test data" to the "training data" for that object (O2), in order to quantifiably measure the "recognition level' of that object.

However, for reason described above, the method will herein be described by passing through the data generation procedure once, based on a single object 1 residing in a single scene 2 and viewed from a single viewpoint 4. For clarity, repeatable procedures will be identified in the description by the designation "Perform the following procedure X" and "End of procedure X".

Perform the following procedure A for each scene.

Procedure A comprises of 4 steps.

<u>Step 1</u>:

Set up a digital camera (or other digital imaging system) at an arbitrary original viewpoint 4, with a viewpoint position 6 and a viewpoint orientation 8 in space, to image the scene 2, with known optics parameters, for example a view width angle 10 and a view height angle 11. Also set a 2-dimensional pixel resolution for the camera.

Perform the following Procedure B for each new viewpoint.

Each new viewpoint, during each pass through this loop, must be at a known relative position and orientation from the original viewpoint 4. This loop firstly involves generation of a 2D image 12 of object 1 as shown in Fig. 3, comprising an array of pixels, each pixel 14 (for example) corresponding to a viewable point 13 (for example) in the scene 2. Viewable points are shown here in Fig. 3 as all being positioned on the surface of the object 1. However, in the general sense, viewable points can also be located somewhere else in the scene 2, not on the object 1. In this embodiment each pixel has associated with it Red-Green-Blue (RBG) light intensity values, however other well known colour or monochrome light intensity value parameter sets could also be employed.

A distance 15 of each viewable point 13 from the viewpoint 4 is now calculated using a parallax technique. This is facilitated by creating at least two 2D sub-images from two alternative viewpoints, each alternative viewpoint itself of known position and orientation and, by implication, offset by a predetermined separation distance. Then, either sequentially or using different cameras, the distance 15 can be calculated by techniques well known in the art, using the parallax offset between the different pixel positions 14 of the corresponding viewable point 13 in the two images 12.

Now referring in more detail to one particular embodiment of this technique in Figs. 4a - 4d, the calculation of the distance 15 involves creation of a contrast map of the 2D image (refer to Fig. 4a), and then performing sub-pixel edge detection on this contrast map (refer to Fig. 4b). This is repeated for a number of pixel value offset positions, adjusting the position in the direction of the sub-viewpoint separation distance, of pixel values in one sub-image (refer to Fig. 4c), testing the sub-image associated with each sub-viewpoints, and measuring how similar they are (refer to Fig. 4d). Upon locating the most similar sub-image, the distance 15 between the viewable point 13 of the surface of the object 1 and the viewpoint 4 can be calculated based upon their separation pixel value offset in the sub-image.

A depth map as shown in Fig. 5 can now be generated based on the distances calculated for all the viewable points 13 as visible from the viewpoint 4.

The coordinates (x, y, z) in the space 2 of each viewable point 13 can now be calculated based on the location of the corresponding pixel 14 in the array, its corresponding distance 15, viewpoint position 6 and viewpoint orientation 8 in the space 2, and the viewing properties of the viewpoint 4 such as the view width angle 10 and the view height angle 11. Object data points 17 are generated for object 1 based on the “view” from the viewpoint 4 as shown in Fig. 6.

Based on the light intensity values (RBG values in this case) and corresponding coordinates of each pixel 14 in the image 12, a set of “object features” can be now identified. In the embodiment described herein in respect to the present invention one type of object feature, specifically “corners” of the object, or regions of high curvature on the object 1, are derived as follows.

Using the RGB map shown in Fig. 7 (the light intensity RGB values of every pixel in the image 12), a corresponding luminosity map shown in Fig. 8 is generated, in which each luminosity map pixel value is equal to the red plus green plus blue component values of the corresponding RGB map pixel.

Using the luminosity map, a luminosity contrast map shown in Fig. 9 is generated, in which each pixel value is equal to a function of the difference in luminosity between neighbouring pixels.

Also, using the depth map already described in reference to Fig. 5 (the distance values of every pixel in the image 12), a depth contrast map is generated shown in Fig. 10, in which each pixel value is equal to a function of the difference in depth between neighbouring pixels.

Also, using the depth map in Fig. 5, a depth gradient map is generated shown in Fig. 11, in which each pixel value has a vector associated with it which is equal to the change in depth between neighbouring pixels.

Using this depth gradient map, a depth gradient contrast map is generated shown in Fig. 12, in which each pixel value is equal to a function of the difference in depth gradient (in all directions) between neighbouring pixels.

At this point in the procedure, object data points corresponding to unique objects may be isolated based upon, for example, contiguous regions of zero or near-zero depth contrast or contiguous regions of zero or near-zero luminosity contrast. For every object isolated, a unique object index is assigned. It should be noted that if procedure B is executed for more than one viewpoint, the indices of the objects isolated with the indices of the objects isolated for previous viewpoints are mapped based upon the coordinates of the viewable points in the current image and the coordinates of the viewable points in previous images and their corresponding object indices.

A function of the original object data, such as the luminosity contrast map, the depth contrast map, or depth gradient contrast map, is then used to identify a region 40 of contiguous object data points shown in Fig. 13, corresponding in this case to a side of the cubic object 1. The luminosity contrast map in Fig. 9 is chosen to be used in this case to generate the single contiguous region of zero or near-zero luminosity contrast.

A base point is then calculated based upon one or more points contained within this region 40. The base point calculated is chosen in the case of this embodiment to be the centroid 41 of the region, the average position (x, y, z) of all pixels in the region 40 - or all pixels in the contiguous region's outer boundary, which happens to provide the same centroid position in this particular case.

The coordinates (x,y,z) of the centroid 41 provides an initial feature data point for object 1. Additional object features may be calculated based upon a function of the relative position of the centroid 41 and the coordinates of all other points in the region 40. Sequentially, taking adjacent points, the position (x,y,z) of every point 42, 43, 44 on the outer boundary of the region 40 is then compared to the position of the centroid 41 (x,y,z) measuring a distance 45, and it is noted where relative minimas and maximas occur in this distance 45. These relative minimas and maximas are used to generate coordinates (x,y,z) of additional feature data points 18 as shown in Fig. 14. In this particular case, the identified feature data points 18 are object features corresponding to virtual corners of the object 1.

For each of the objects 1 isolated from this particular viewpoint 4, a set of 3D “data points” is created. This set of data points is made up by (a) a set of 3D coordinates (x, y, z) of the feature data points, the virtual corners 18 plus the centroids 41 in this case for each of the identified regions 40, and (b) a set of object data points where each object data point contains the 3D coordinates (x, y, z) of the respective viewable points 13 in the image 12 pertaining to the particular object 1 when viewed from this particular viewpoint 4, plus its associated RBG values.

This is now the end of Procedure B.

<u>Step 2</u>:

Referring to Fig. 15, for each of the objects, combinations of the coordinates of the virtual corners 18 of the set of feature data points pertaining to object 1 are now grouped as apexes of object triangles, for example the three apexes 19, 20 and 21 of the object triangle 22 bounded by the three sides 23, 24 and 25. The triangles, four in the case of Fig. 15, are differently 3-dimensionally orientated in space 2. The specific orientation of plane 26 which is coplanar with object triangle 22 is fully defined by the coordinates of the three respective apexes 19, 20 and 21.

To minimise the number of feature point combinations (and hence the number of object triangles defined), it may be asserted that all three features points of each combination of feature points must originate from the feature point calculations for single contiguous region (rather than originating from more than one contiguous region's feature point calculation process).

<u>Step 3</u>:

For each of the one or more of the sides 23, 24 and 25 of each of the object triangles 22 of each of the objects 1 for each of the viewpoints 4, a unique transformation index assigned, and the coordinates of the set of the data points are transformed into transformed coordinates in a new axis system, where the transformation function is a function of the coordinates of the apexes of the particular object triangle.

This transformation procedure is as follows.

The transformation function comprises, shifting the position of the origin to a new origin position and reorientating the axis system to a new axis orientation such that, in the new axis system (X', Y', Z'), defined by a first new axis (X'), a second new axis (Y'), and a third new axis (Z'), the new axis orientation, and new origin position, are a function of the coordinates of the apexes of the respective object triangle.

Referring to the axis transformation sequence shown in Fig. 16, 17, 18 and 19, the origin 3 is shifted to a new origin position 27 and reorientated to a new axis orientation 28 such that, in the new X', Y', Z' axis system, a third new axis (Z') is aligned perpendicular to the plane 26 of the object triangle 22 (Fig. 17), this third new axis (Z') passes through a mid-point 29 between the apexes 19 and 21 at the extremities of side 25 (Fig. 18), a first new axis (X') of the new axis system is aligned parallel to side 25, and a second new axis (Y') of the new axis system is directed through mid-point 29 in the direction 35 of the third apex 20 of the triangle (Fig. 19).

The coordinates (x, y, z) of the data points (object data points and feature data points) are therefore appropriately transformed into new coordinates (x', y', z') in the new X', Y', Z' axis system, hence forming a set of transformed data points as shown in Fig. 20.

Alternatively, in an embodiment of the present invention where only 2-dimensional object data is available (e.g. a 2D image of the object from a viewpoint), the transformation function may involve the following.

The new axis system (X', Y') is defined by a first new axis (X') and a second new axis (Y'), and the coordinates (x, y) of the set of data points are transformed such that, in the new axis system, the three apexes of the object triangle are coincident with the three apexes of a predefined triangle in the new axis system.

Referring to the coordinate transformation sequence shown in Fig. 21, 22, 23, 24, 25, and 26, the three apexes of the object triangle 22 are made coincident with the three apexes of a predefined equilateral triangle 36, where a predefined side 39 of the predefined triangle 36 is parallel to the new X' axis (Fig. 21). This is achieved by, firstly, scaling the coordinates of the data points such that the side 25 of the object triangle is of same length 37 as the side 39 of the predefined triangle 36 (Fig. 22), then rotating the coordinates of the data points such that the side 25 of the object triangle 22 is parallel with the new X' axis (Fig. 23), then scaling the coordinates of the data points in the new Y' axis direction such that the perpendicular distance 38 between the side 25 of the object triangle and the third apex of the object triangle 22 and perpendicular distance between the side 39 of the predefined triangle 36 and the third apex of the predefined triangle are the same (Fig. 24), then shearing the coordinates of the data points along the new X' axis (Fig. 25), and translating the coordinates of the data points such that the object triangle is centred about a predefined point, the centre of the predefined triangle 36 (Fig. 26).

Step 4:

Referring to Figs. 27 and 28, for each of the axis transformations, an interpolated 3D surface mesh 30 is formed defined by surface mesh points 31, the coordinates of each surface mesh point 31 corresponding to the coordinates of a transformed data point 33 associated with the object 1 and the relevant axis transformation or, as in the case of surface mesh point 31B, corresponding to an interpolation of one or more of the coordinates of the transformed data points 33. Each polygon 32 of the surface mesh 30 is formed by proximate surface mesh points 31 and 31B and has associated with it the light intensity values (or “RGB values” in this embodiment) derived as a function of the RGB values associated with nearby transformed data points.

Referring to Figs. 29 and 30, for each of the axis transformations, a 2D image of the surface mesh 30 is generated, where the image data points 34 in the 2D image are calculated as the coordinates (or array positions) and RGB values of pixels in the 2D image generated when, using computer graphics, a virtual mesh viewpoint is orientated along the Z’ axis passing through the new origin position 27, each pixel corresponding to a viewable surface mesh point 31 on the surface mesh 30 and having RGB values derived as a function of the RGB values of transformed data points proximate to the viewable surface mesh point 31, and coordinates derived as a function of the coordinates of the data points proximate to the viewable surface mesh point 31. A viewable mesh depth map or depth map image (a 2D array of Z’ values of each of the viewable surface mesh points 31) is also created along with the viewable mesh RGB map or RGB map image (a 2D array of RGB values of each of the viewable surface mesh points 31).

For each of the axis transformations, a set of characterising data is generated. In the case of this embodiment, this set of characterising data also comprises the respective object index, and the respective transformation index.

The set of characterising data generated according to the present invention can be used for purposes of object recognition. As mentioned at the start of the description of this embodiment, the characterising data generating method according to the present invention can be executed sequentially for multiple scenes, each with multiple objects and viewpoints. Training data can be initially generated and this training data later compared with test data, also generated using the same method, and algorithms then employed to compare one or more objects in the test data with one or more objects in the training data, hence enabling the recognition of the one or more objects. For each of the one or more sides of each of the object triangles of each of the objects of the test data, and for each of the one or more sides of each of the object triangles of each of the objects of the training data, the one or more sets of characterising data pertaining to the test data are compared with the one or more sets of characterising data pertaining to the training data.

Two embodiments of the algorithm used for this comparison will be described in this specification: a “database comparison algorithm” and a “network comparison algorithm”. Importantly, the method of data processing of the transformed data points, and hence the format of the set of characterising data, depends on which comparison algorithm is eventually used.

When the “database comparison algorithm” will be used, the set of charactersing data may be represented as a 1-dimensional binary string.

For each of the axis transformations, binned coefficient values of a spatial convolution of the viewable mesh RGB map(s), binned coefficient values of a spatial convolution of the the viewable mesh depth map(s), and the binned coordinate values of the transformed data points corresponding to the object features (for example the apexes of the object triangles), may be recorded as the set of characterising data. The set of charactersing data represented as a 1-dimensional binary string, may for example contain 4 bits corresponding to each of the binned spatial convolution coefficient values, and 4 bits corresponding to each of the binned coordinate values of the transformed data points corresponding to the object features.

Alternatively, one or more bits of the 1-dimensional binary string may be generated as a function of the spatial convolution coefficient values, coordinate values of the transformed data points corresponding to the object features, viewable mesh RGB map(s), or viewable mesh depth (s). This function could be a neural network, or decision tree, where these are input data to the neural network(s) or decision tree(s). Alternatively, where the viewable mesh RGB map(s), or viewable mesh depth (s) are used, this function could be an image processing function (such as a Gaussian filter) where it has been configured to result in significant quantization (spatial resolution and intensity resolution reduction) of the output, and where this quantization is insensitive to slight (spatial or intensity) variations in the input data. The quantization should ideally be independent of aliasing (or the pixel map structure), and produce repeatable (the same) output with slight variations to the input maps (to ensure train versus test data will match).

The spatial convolution utilised may be a low resolution DCT (discreet cosine transformation) such is implemented in image compression algorithms well known to the art (e.g. JPEG). When the one or more set of characterising data currently being generated is training data, two or more subsets of each set of characterising data may be formed, each with slight variations to the binning of the coordinates of the transformed data points corresponding to the object features, and/or the binning of the DCT coefficients. When the one or more set of characterising data currently being generated is test data, each set of characterising data will only include a single subset of the set of characterising data. In the case of the “database comparison algorithm”, the generation of such multiple subsets of each set of characterising data enables instantaneous database lookup since, for example, direct comparisons are possible between binary strings of the one or more sets of characterising data pertaining to test data and binary strings of the one or more sets (or subsets) of characterising data pertaining to training data.

When the “network comparison algorithm” will be used, for each of the axis transformations, the viewable mesh RGB map(s), the viewable mesh depth map(s), and the coordinates of the transformed data points corresponding to the object features (for example the apexes of the object triangles), are recorded as the set of characterising data. When the one or more sets of characterising data currently being generated is training data, two or more subsets of the set of characterising data may be formed, each with slight positional or orientation variations of the viewpoint, in order to improve the performance of the neural network. Again, when the one or more set of characterising data currently being generated is test data, the set of characterising data will only include a single subset of the set of characterising data.

This is now the end of Procedure A.

As referred to above, the set of characterising data generated according to the present invention can be used for purposes of object recognition, and this involves a comparison of one or more sets (or subsets) of characterising data comprising the training data with one or more sets of characterising data comprising the test data, using an algorithm.

When the “database comparison algorithm” is used, every set of characterising data in the test data is compared with every set (or subset) of characterising data in the training data. The object indicies corresponding to the set of characterising data in the training data which result in matches are recorded, and tallied. An object is recognised based upon a function of the number of sets (or subsets) of characterising data associated with that object that produced matches.

When the “network comparison algorithm” is used, the comparison may be performed by first comparing the coordinates of the transformed data points corresponding to the object features (for example the apexes of the object triangles) of the test data and training data - a method sometimes referred to as a “geometric hashing algorithm”. In practice the actual comparison is performed by comparing the distance each (transformed) object triangle apex in the test data is away from the corresponding (transformed) object triangle apex in the training data. For each of the one or more sides of each of the object triangles of each of the objects of the test data, the maximum comparison accuracy value experienced across all training data sets is calculated, and the transformation indices of the training data is recorded for each case where their maximum “comparison accuracy” is above a certain arbitrary value. This creates a subset of the test data termed in this specification the “geometry matched test data” (GMTD), and for each transformation index of the GMTD, the transformation indices of the object triangle side(s) of the training data that give “comparison accuracies” above the certain arbitrary value, called in this specification the “geometry matched test data training data matches” (GMTDTDM), are also recorded.

Light intensity recognition is now performed on those test data object triangle sides and their corresponding training object triangle sides which are found to have a high geometric “comparison accuracy” i.e. the GMTD and their corresponding GMTDTDM, by performing the following procedure.

As described below, a network based artificial intelligence algorithm (such as a neural network or a decision tree) is used to train a network with the training data. The input values into the network are the values of a viewable mesh RGB map and/or the values of a colour saturation contrast map of the viewable mesh RGB map and/or the values of a viewable mesh depth map, and the output values of the network are a function of an object index or an object transformation index (for example the object index x 3 + the transformation index).

For each of the one or more sides of each of the object triangles of each of the objects of the training data, the network is trained with an experience, where the input values are the values of the viewable mesh RGB map(s) and/or the values of the colour saturation contrast map(s) of the viewable mesh RGB map(s) and/or the values of the viewable mesh depth map(s) from the training data, and the output value is the training output value value (for example the object index x 3 + the transformation index).

The network based artificial intelligence algorithm can now be used to test the network against the test data object(s), thereby determining whether or not the testing object(s) is/are recognised and which of the trained objects this test data object(s) corresponds to.

For each of the one or more sides of each of the object triangles of the object of the GMTD, and for every possible training output value, the test data experience is tested against the trained network, where the input values are the values of the viewable mesh RGB map and/or the values of the colour saturation contrast map of the viewable mesh RGB map and/or the values of the viewable mesh depth map from the test data, and the output is the training output value, where the test data object (index) equates to the training data object (index).

It is also verified that the object triangle side (transformation index) recognised in the light intensity recognition phase is also recognised as the same object triangle side (transformation index) in the geometric recognition phase. The transformation indices of the object triangle sides of the GMTD and training data matched by use of the network in the light intensity recognition phase are compared with the transformation indices of GMTD and each of their GMTDTDM matched in the geometric recognition phase. An object is then recognised based upon a function of the number of object triangle sides associated with that object that were recognised.

The term “comprising” as used herein is used in the inclusive sense of “including” or “having” and not in the exclusive sense of “consisting only of”.

Claims

The claims defining the invention are as follows:

Claim 1

A method for generating one or more sets of characterising data for one or more objects residing in one or more scenes in an N-dimensional space defined by an axis system passing through an origin, comprising the steps of, for each scene:

1. deriving a set of N-dimensional object data points for an object, each object data point defined by N-dimensional coordinates and one or more light intensity values, based on a first data-processing of the coordinates and the one or more light intensity values of the object data points, and also deriving a set of feature data points corresponding to object features of the object, each feature data point also defined by N-dimensional coordinates;
2. grouping combinations of the coordinates of the feature data points as the coordinates of apexes of one or more object triangles, each object triangle lying in a plane;
3. for each of one or more of the sides of each of the object triangles, applying a transformation function to the coordinates of one or more of the data points, including at least one object data point or one feature data point, to generate respective transformed data points with respective new coordinates in a new axis system, wherein the transformation function is a function of the coordinates of the apexes of the respective object triangle; and
4. for each of one or more of the sides of each of the object triangles, carrying out a second data-processing of the transformed data points, and generating a set of characterising data.

Claim 2

A method as claimed in Claim 1 wherein the one or more light intensity values correspond to one or more light frequencies, or frequency bands, or a function of these frequencies or frequency bands.

Claim 3

A method as claimed in Claim 1 wherein the object data points are located on the surface of the respective object and the one or more light intensity values characterize the light radiating from the object data points.

Claim 4

A method as claimed in Claim 1 wherein a unique object index is assigned to each of the objects.

Claim 5

A method as claimed in Claim 1 wherein the first data-processing in Step 1 comprises subdividing the object data points into at least one region of contiguous object data points with similar one or more light intensity values or a similar first function of their one or more light intensity values, deriving a base point for the region as a second function of the coordinates of the object data points in the region and, based on a third function of the base point, deriving feature data points.

Claim 6

A method as claimed in Claim 5, wherein the first function is a luminosity contrast function of the one or more light intensity values.

Claim 7

A method as claimed in Claim 5, wherein the second function is a geometric centroid of the coordinates of the object data points in the region.

Claim 8

A method as Claimed in Claim 5, wherein the third function relates to the distance of one or more of the object data points in the region from the base point.

Claim 9

A method as claimed in Claim 5, wherein the third function is a relative and/or absolute minima and/or maxima of the distance of one or more of the object data points in the region from the base point.

Claim 10

A method as claimed in Claim 5, wherein one more of the object data points in the region are on a boundary of the region.

Claim 11

A method as claimed in Claim 5, wherein the base point is also a feature data point.

Claim 12

A method as claimed in Claim 5, wherein the third function is an equality function and, correspondingly, the base point is the only feature data point for the region.

Claim 13

A method as claimed in Claim 1 wherein the first data-processing in Step 1 comprises determining at least one point of high local curvature of the object data points, and designating the point as a feature data point.

Claim 14

A method as claimed in Claim 1 wherein the first data-processing in Step 1 comprises determining at least one point of high local curvature of the one or more light intensity values, and designating the point as a feature data point.

Claim 15

A method as claimed in Claim 1, wherein the object data points are derived by N-dimensionally imaging the scene from one or more viewpoints, each viewpoint characterised by a position and an orientation in the space and predefined viewing properties and, for each of the viewpoints, generating a 2-dimensional image of the object, each image comprising an array of pixels, each pixel corresponding to a viewable point in the respective scene and being characterised by one or more light intensity values and coordinates in the space and, either during the calculating of the coordinates for each viewpoint or after the calculating of the coordinates for all viewpoints, isolating the object data points for the object.

Claim 16

A method as claimed in Claim 15, wherein the object data points are isolated by determining boundaries of the object in the image based on a boolean depth contrast map derived by applying an arbitrary threshold to a depth contrast map of the image, or a boolean depth gradient contrast map derived by applying an arbitrary threshold to a depth gradient contrast map of the image, or a boolean luminosity (overall light intensity) contrast map derived by applying an arbitrary threshold to a luminosity contrast map of the image, or linear or non-linear functional combination of these maps.

Claim 17

A method as claimed in Claim 15, wherein the position and the orientation of each viewpoint in the space are predefined.

Claim 18

A method as claimed in Claim 15, wherein at least one of the viewable points reside on a surface of the object.

Claim 19

A method as claimed in Claim 15, wherein the object data points are derived by 3-dimensional imaging of the scene, the space is a 3-dimensional space, and the coordinates in the space of the viewable point are 3-dimensional coordinates.

Claim 20

A method as claimed in Claim 15 wherein, for each viewpoint, the N-dimensional imaging comprises creating at least two 2-dimensional sub-images from two alternative viewpoints, slightly offset by a predetermined distance either side of the viewpoint, either sequentially or using at least two corresponding imaging systems, and a resulting parallax offset between the positions of the corresponding viewable point in the resulting pixels arrays of the sub-images, and/or the difference between the corresponding one or more light intensity values, is used to calculate the distance of the viewable point from the viewpoint.

Claim 21

A method as claimed in Claim 20, wherein the coordinates in the space of the viewable point is derived based on the location of the corresponding pixel in the array, the distance of the viewable point from the viewpoint, and the position, the orientation, and the viewing properties of the viewpoint.

Claim 22

A method as claimed in Claim 15, wherein the viewing properties comprise a view width angle and a view height angle.

Claim 23

A method as claimed in Claim 1 wherein the first data processing in Step 1 is performed using hardware acceleration hardware, such as a PC graphics card.

Claim 24

A method as claimed in Claim 1, wherein the space is a 3-dimensional space and the coordinates are correspondingly 3-dimensional coordinates, and the transformation function in Step 3 comprises shifting the position of the origin to a new origin position and reorientating the axis system to a new axis orientation such that, in the new axis system (X', Y', Z'), defined by a first new axis (X'), a second new axis (Y'), and a third new axis (Z'), the new axis orientation and new origin position are a function of the coordinates of the apexes of the respective object triangle.

Claim 25

A method as claimed in Claim 24, wherein the new axis system (X', Y', Z') is positioned and aligned such that the third new axis (Z') is aligned perpendicular to the plane of the respective object triangle, and passes through a mid-point between the two apexes at the extremities of the respective side of the object triangle, the first new axis (X') is aligned parallel to the side, and the second new axis (Y') is directed through the mid-point in the direction of the remaining apex of the object triangle.

Claim 26

A method as claimed in Claim 24, wherein the new axis orientation is a function of the 3-dimensional orientation of the respective side and the 3-dimensional orientation of the plane of the respective object triangle.

Claim 27

A method as claimed in Claim 1, wherein the space is a 2-dimensional space and the coordinates are correspondingly 2-dimensional coordinates, and the transformation function in Step 3 comprises transforming the coordinates of the data points such that, in the new axis system (X', Y'), defined by a first new axis (X') and a second new axis (Y'), the three apexes of the respective object triangle are coincident with the three apexes of a predefined triangle in the new axis system.

Claim 28

A method as claimed in Claim 27, wherein the predefined triangle is an equilateral triangle.

Claim 29

A method as claimed in Claim 27, wherein the three apexes of the respective object triangle are made coincident with the three apexes of the predefined triangle by scaling the coordinates of the data points in a first direction such that the respective side of the object triangle is of same length as a predefined side of the predefined triangle, scaling the coordinates of the data points in a second direction such that the perpendicular distance between the side of the object triangle and the corresponding remote apex of the object triangle and perpendicular distance between the predefined side of the predefined triangle and the corresponding remote apex of the predefined triangle are equal, then shearing the coordinates of the data points along an axis defined by the orientation of the side of the object triangle, and then translating the coordinates of the data points such that the three apexes of the object triangle are coincident with the three apexes of the predefined triangle.

Claim 30

A method as claimed in Claim 27, wherein a predefined side of the predefined triangle is parallel to the first new axis (X'), and the three apexes of the respective object triangle are made coincident with the three apexes of the predefined triangle by firstly scaling the coordinates of the data points such that the respective side of the object triangle is of same length as the predefined side of the predefined triangle, then rotating the coordinates of the data points such that the side of the object triangle is parallel with the first new axis (X'), then scaling the coordinates of the data points in the direction of second new axis (Y') such that the perpendicular distance between the side of the object triangle and the corresponding remote apex of the object triangle and perpendicular distance between the predefined side of the predefined triangle and the corresponding remote apex of the predefined triangle are equal, then shearing the coordinates of the data points in the direction of the first new axis (X'), and then translating the coordinates of the data points such that the three apexes of the object triangle are coincident with the three apexes of the predefined triangle.

Claim 31

A method as claimed in Claim 1, wherein the new axis system has an orientation and an origin position equivalent to that of the axis system.

Claim 32

A method as claimed in Claim 1 wherein in Step 3 a unique transformation index is assigned to each of the one or more sides of each of the object triangles of each of the objects.

Claim 33

A method as claimed in Claim 1 wherein the transformation function in Step 3 is performed using hardware acceleration hardware, such as a PC graphics card.

Claim 34

A method as claimed in Claim 1 wherein the second data-processing in Step 4 comprises creating an M-dimensional image of the transformed data points.

Claim 35

A method as claimed in Claim 34 wherein the image is created by forming an interpolated or non-interpolated N-dimensional surface mesh defined by surface mesh points, the coordinates of each surface mesh point corresponding to a coordinate of, or an interpolation of one or more of the coordinates of, the transformed data points, each polygon of the surface mesh formed by proximate surface mesh points and having associated with it one or more light intensity values derived as a function of the one or more light intensity values of the nearby transformed data points.

Claim 36

A method as claimed in Claim 35 wherein the image comprises a depth map or a function of the depth map calculated based on the surface mesh.

Claim 37

A method as claimed in Claim 35 wherein the image comprises a luminosity map or a function of the luminosity map calculated based on the surface mesh.

Claim 38

A method as claimed in Claim 34 wherein the image is a 2-dimensional image.

Claim 39

A method as claimed in Claim 34 wherein the set of characterising data comprises one or more subsets of characterising data, each subset comprising coordinates of one or more of the transformed data points corresponding to feature data points, or a fourth function of the coordinates of one or more of the transformed data points corresponding to feature data points, or the M-dimensional image of the transformed data points, or a seventh function of the M-dimensional image of the transformed data points.

Claim 40

A method as claimed in Claim 1 wherein the set of characterising data comprises one or more subsets of characterising data, each subset comprising the coordinates of one or more of the transformed data points corresponding to feature data points, or a sixth function of the coordinates of one or more of the transformed data points corresponding to feature data points.

Claim 41

A method as claimed in Claim 39 wherein the seventh function of the M-dimensional image of the transformed data points is a spacial convolution function, and this is applied to the image to generate one or more convolution coefficients, or a fifth function of the one or more convolution coefficients.

Claim 42

A method as claimed in Claim 41 wherein the fourth or fifth functions are a binning function.

Claim 43

A method as claimed in Claim 40 wherein the sixth function is a binning function.

Claim 44

A method as claimed in Claim 42 wherein the binning function is also a function of a subset index number uniquely identifying a subset of characterising data such that, for one or more of the transformed data points corresponding to feature data points, or for one or more of the convolution coefficients, an adjacent bin is filled instead of an optimum bin.

Claim 45

A method as claimed in Claim 43 wherein the binning function is also a function of a subset index number uniquely identifying a subset of characterising data such that, for one or more of the transformed data points corresponding to feature data points, an adjacent bin is filled instead of an optimum bin.

Claim 46

A method as claimed in Claim 1 wherein the set of characterising data is represented as a 1-dimensional binary string.

Claim 47

A method as claimed in Claim 39 wherein the one or more subsets of characterising data are each represented as a 1-dimensional binary string.

Claim 48

A method as claimed in Claim 40 wherein the one or more subsets of characterising data are each represented as a 1-dimensional binary string.

Claim 49

A method as claimed in Claim 35 wherein the image is generated by generating a virtual image of the surface mesh points from a particular virtual viewpoint.

Claim 50

A method as claimed in Claim 35 wherein the image comprises mesh surface data points generated by interpolating the surface mesh at particular coordinate intervals along the first new axis (X'), or a function of the first new axis (X'), and along the second new axis (Y'), or a function of the second new axis (Y'), each mesh surface data point corresponding to a point on or off of the mesh surface and having one or more light intensity values derived as a function of the light intensity values of the transformed data points proximate to the point, and a coordinate derived as a function of the coordinates of the transformed data points proximate to the point.

Claim 51

A method as claimed in Claim 50 wherein the mesh surface data points of the image have their one or more light intensity values set to an arbitrary one or more light intensity values if, based in the new axis system, their coordinates lie outside of the triangle formed by the coordinates of the transformed data points corresponding to feature data points, in turn corresponding to the apexes of the object triangle.

Claim 52

A method as claimed in Claim 1 wherein the second data processing in Step 4 is performed using hardware acceleration hardware, such as in a PC graphics card.

Claim 53

A method as claimed in Claim 4 wherein the set of characterising data in Step 4 also includes data relating to the object index.

Claim 54

A method as claimed in Claim 32 wherein the set of characterising data in Step 4 also includes data relating to the transformation index.

Claim 55

A method as claimed in Claim 1 wherein the set of characterising data in Step 4 also includes data relating to the new axis system.

Claim 56

A method as claimed in Claim 1 wherein, in Step 4, the set of characterising data is generated for each combination of scene, object, object triangle, and side.

Claim 57

A method as claimed in Claim 1 wherein the one or more sets of characterising data are initially generated as training data, and another one or more sets of characterising data are also later generated as test data, and the training data and test data is then compared using an algorithm, thereby enabling the recognition of the one or more objects in the one or more scenes.

Claim 58

A method as claimed in Claim 1 wherein the one or more sets of characterising data is generated for multiple scenes as training data, and another one or more sets of characterising data is later generated for a different scene as test data, and the training data and test data is then compared using an algorithm, thereby enabling the recognition of the one or more objects in the different scene.

Claim 59

A method as claimed in Claim 57 or Claim 58 wherein the algorithm comprises determining if one of the one or more sets of characterising data comprising the training data, and one of one or more sets of characterising data comprising the test data, are substantially equal.

Claim 60

A method as claimed in Claim 57 or Claim 58 wherein the algorithm comprises using a neural network to determine the degree of equality of the one or more sets of characterising data comprising the training data, and the one or more sets of characterising data comprising the test data.

Claim 61

A method as claimed in Claim 57 or Claim 58 wherein the one or more sets of characterising data each comprise subsets of characterising data, and the algorithm comprises determining if one or more of the subsets of characterising data comprising the training data, and the one or more subsets of characterising data comprising the test data, are substantially equal.

Claim 62

A method as claimed in Claim 57 or Claim 58 wherein the one or more sets of characterising data each comprise subsets of characterising data, and the algorithm comprises using a neural network to determine the degree of equality of the one or more subsets of characterising data comprising the training data, and the one or more subsets of characterising data comprising the test data.

Claim 63

A method as claimed in Claim 39 wherein the seventh function of the M-dimensional image of the transformed data points is a neural network, or the fourth function of the coordinates of one or more of the transformed data points corresponding to feature data points is a neural network.

Claim 64

A method as claimed in Claim 41 wherein the fifth function of the one or more convolution coefficients is a neural network.

Claim 65

A method as claimed in Claim 39 wherein the seventh function of the M-dimensional image of the transformed data points is a decision tree, or the fourth function of the coordinates of one or more of the transformed data points corresponding to feature data points is a decision tree.

Claim 66

A method as claimed in Claim 41 wherein the fifth function of the one or more convolution coefficients is a decision tree.

Claim 67

A method as claimed in Claim 39 wherein the seventh function of the M-dimensional image of the transformed data points is an image processing function (such as a Gaussian filter) resulting in significant quantization (spatial resolution or intensity resolution reduction) of the output, and where this quantization is insensitive to slight (spatial or intensity) variations in the input data.

Claim 68

A method as claimed in Claim 5 wherein in all three features points of the combination of the coordinates of the feature points grouped as the coordinates of apexes of an object triangle in Step 3 are derived from a single region.

Claim 69

A method as claimed in Claim 1 wherein the operations in Step 3 and Step 4 performed for each of one or more of the sides of each of the object triangles are instead performed for each of one or more of the apexes of each of the object triangles, and where these operations are equivalent because the transformation function is a function of the coordinates of the apexes of the respective object triangle.

Abstract

A method is described for generating characterising data for one or more objects residing in one or more scenes in a space defined by an axis system passing through an origin. The steps involve deriving a set of data points for the object comprising coordinates, one or more of the data points corresponding to feature data points for object features. Object triangles are then defined based on the apexes formed from combinations of the coordinates of the feature data points. The coordinates of the data points are then transformed based on the apexes of each object triangle, thereby forming a set of transformed data points. Characterising data is then generated based on the set of transformed data points for one or more side of each object triangle of each object. This method enables a similar object, with an arbitrary position and orientation, to be recognised by an observer at an arbitrary position and orientation.

Figures

1/30

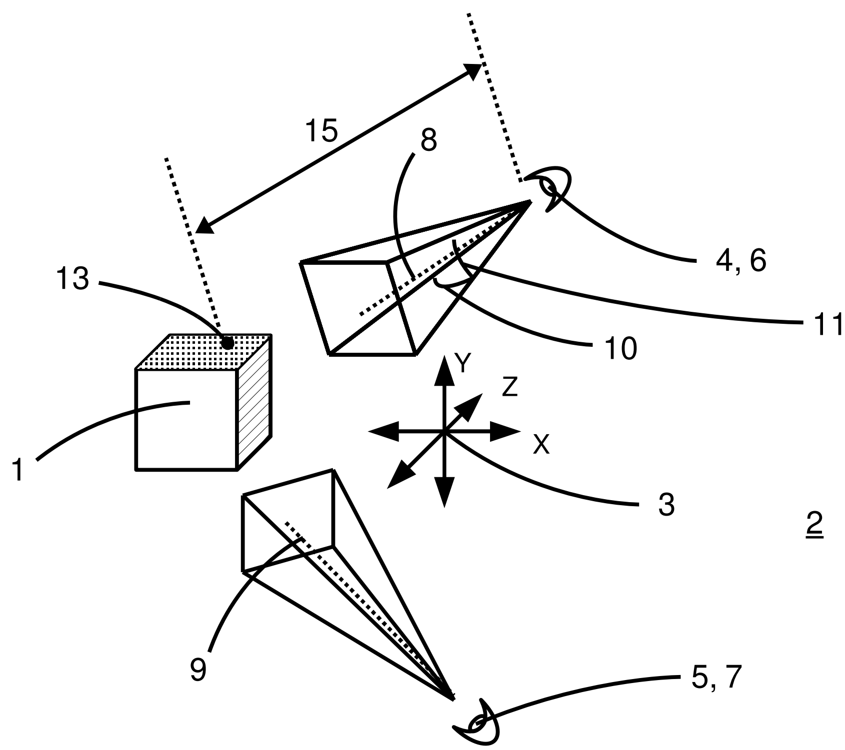

Fig. 1

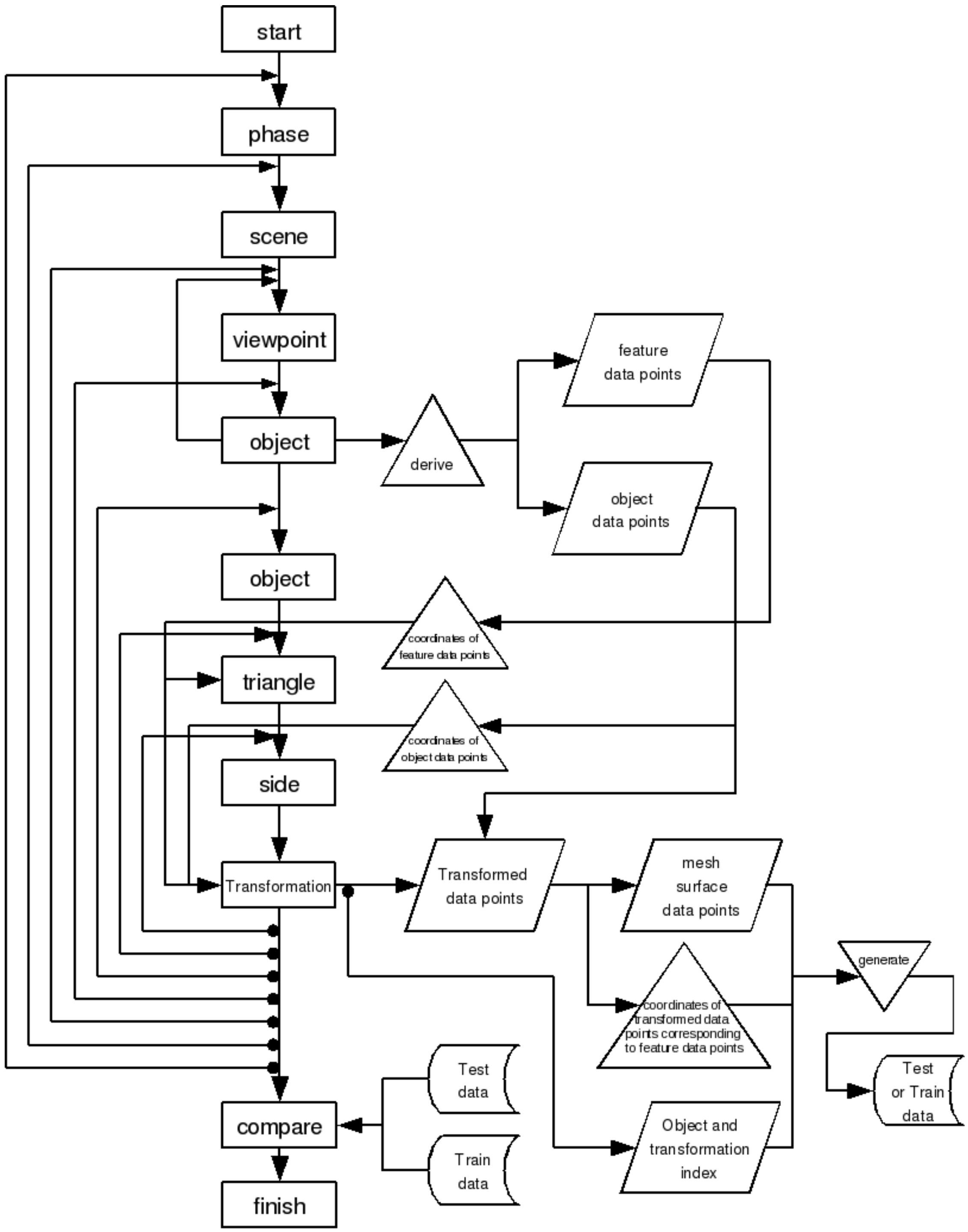
start
phase
scene
viewpoint
object
derive
feature data points
object data points
object
coordinates of feature data points
triangle
coordinates of object data points
side
Transformation
Transformed data points
mesh surface data points
coordinates of transformed data points corresponding to feature data points
generate
Test or Train data
Object and transformation index
Test data
Train data
compare
finish

Fig. 2

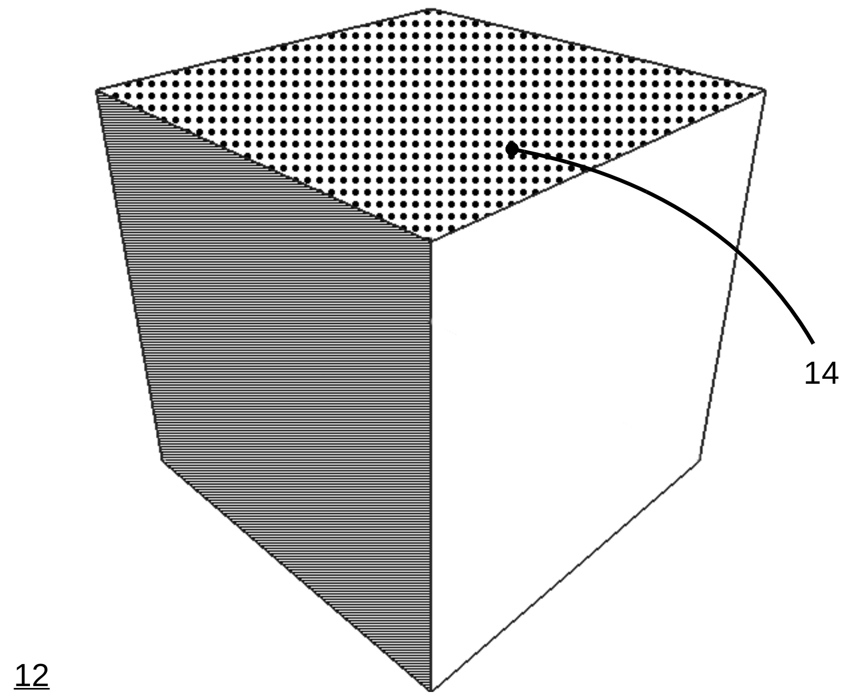

Fig. 3

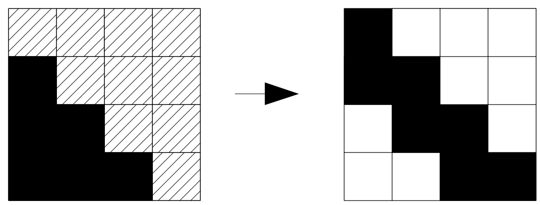

Fig. 4a

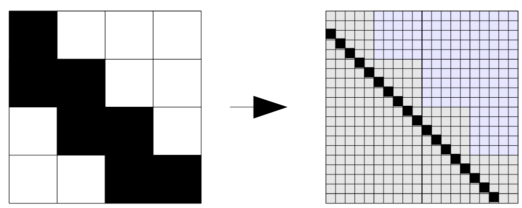

Fig. 4b

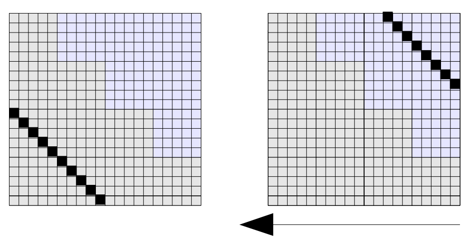

Fig. 4c

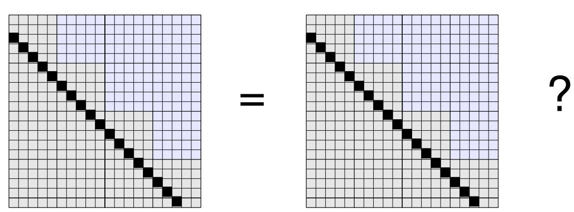

Fig. 4d

Fig. 4

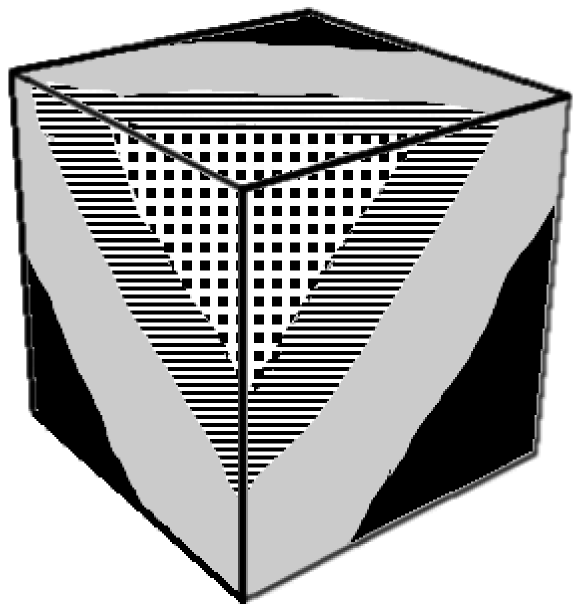

Fig. 5

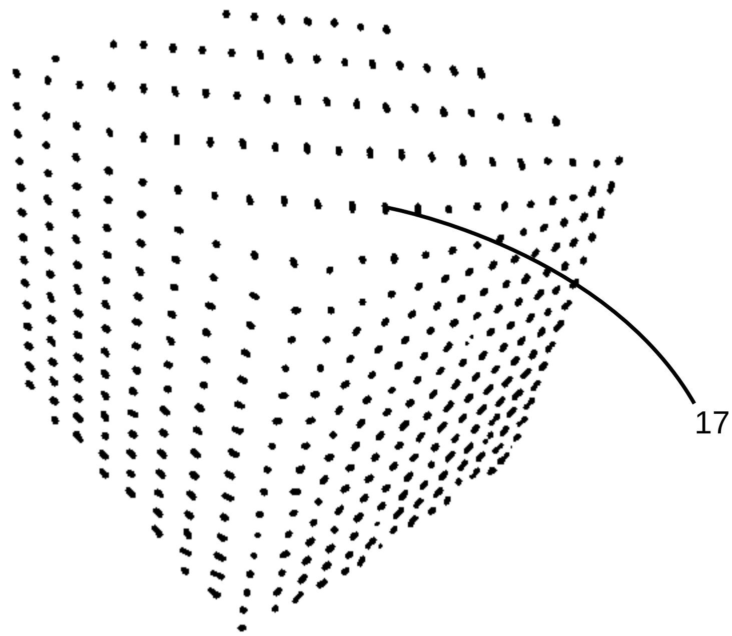

Fig. 6

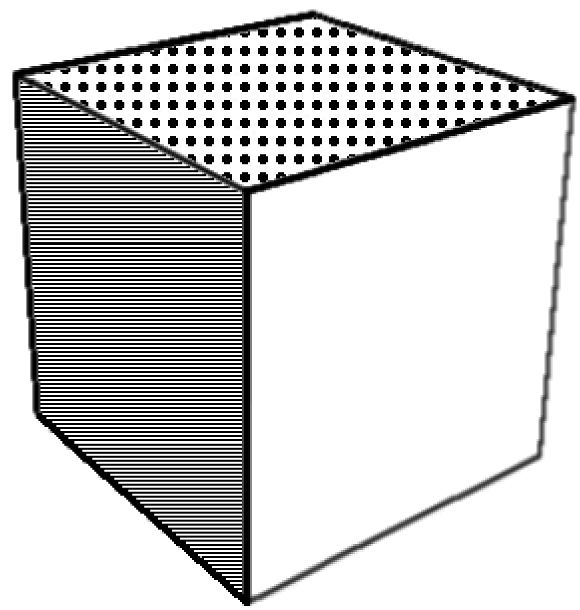

Fig. 7

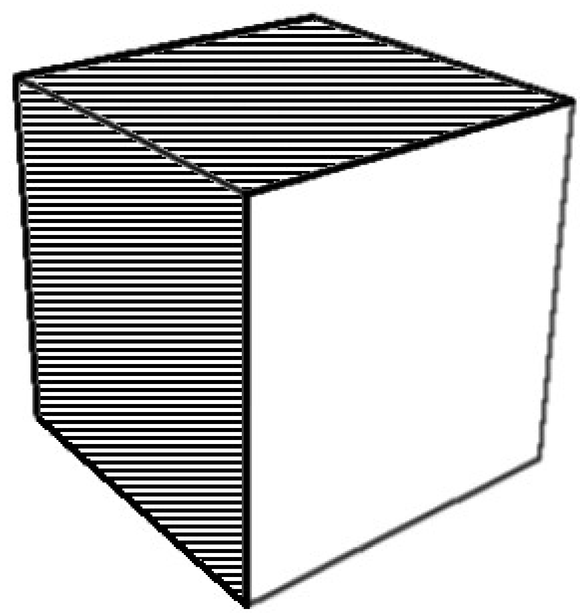

Fig. 8

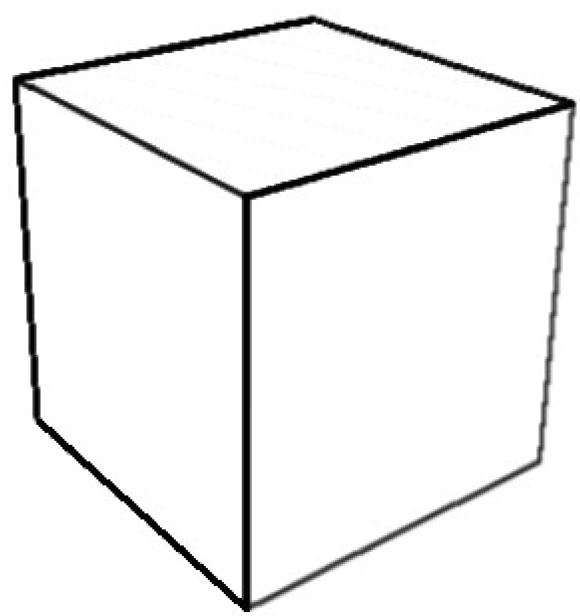

Fig. 9

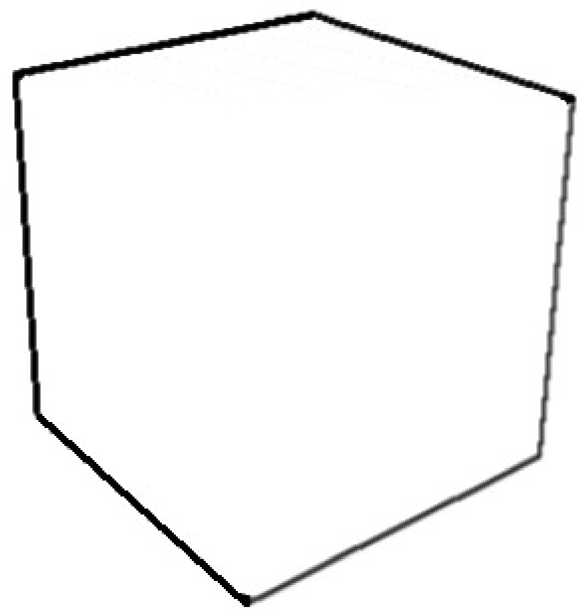

Fig. 10

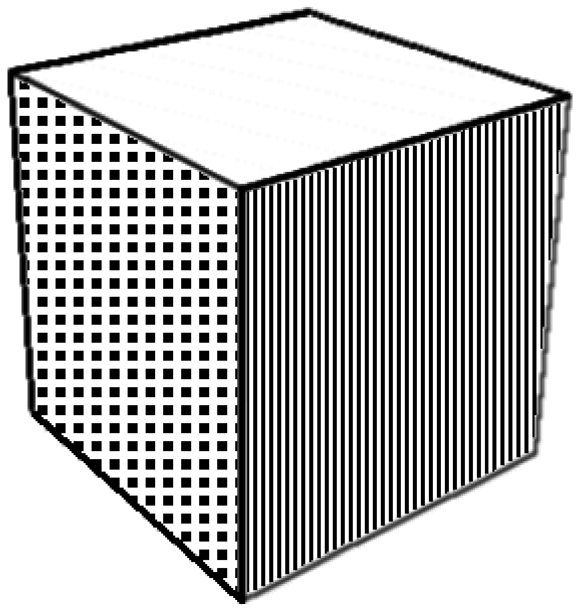

Fig. 11

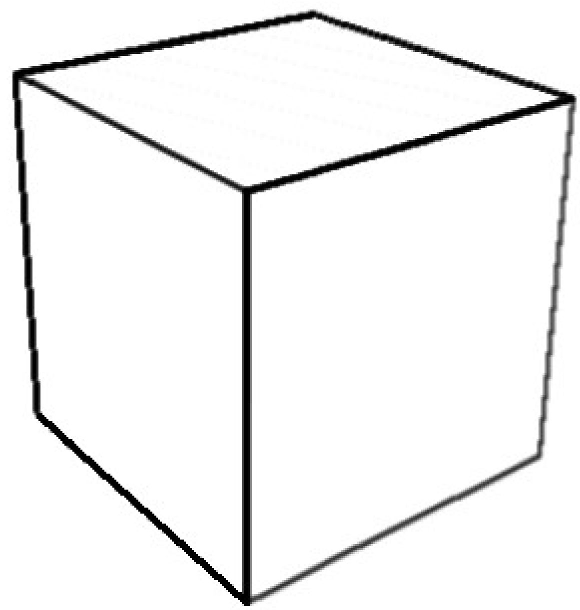

Fig. 12

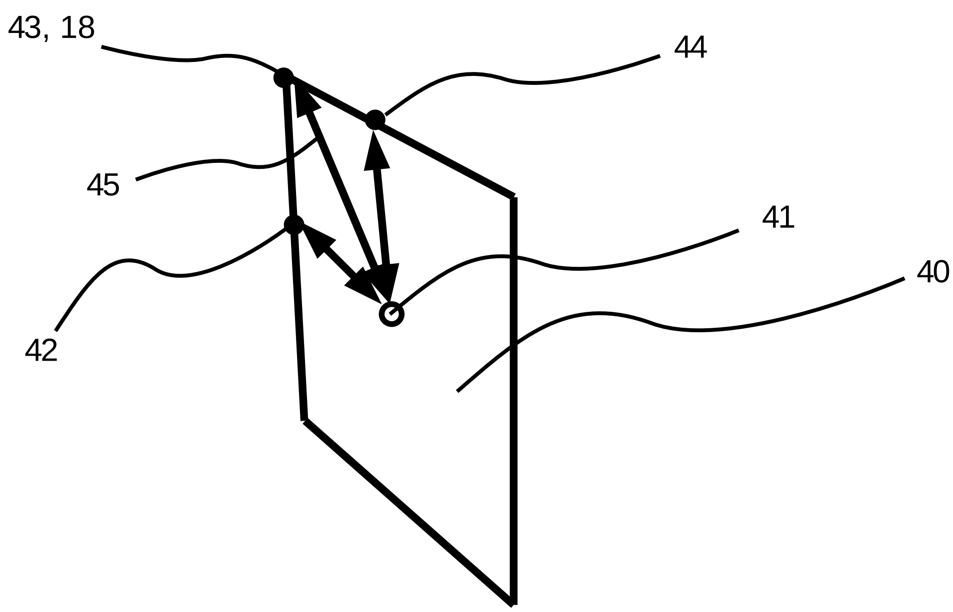

Fig. 13

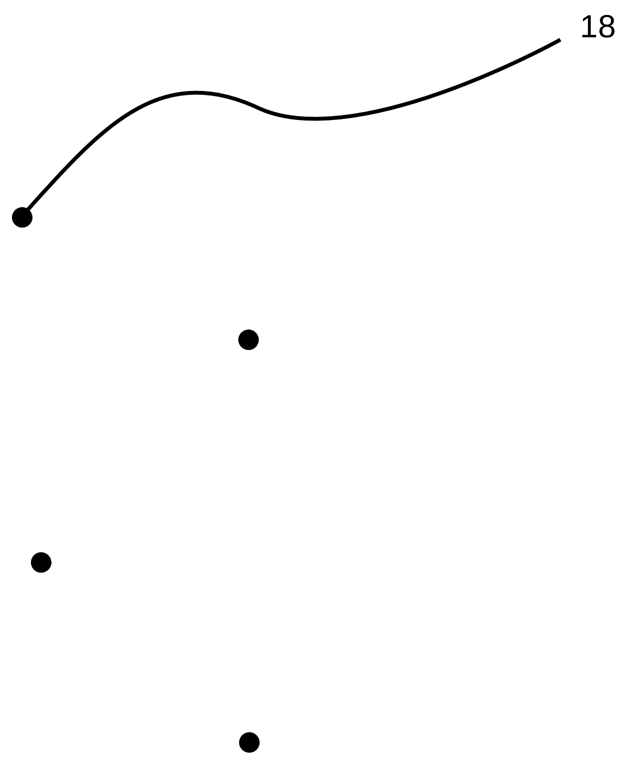

Fig. 14

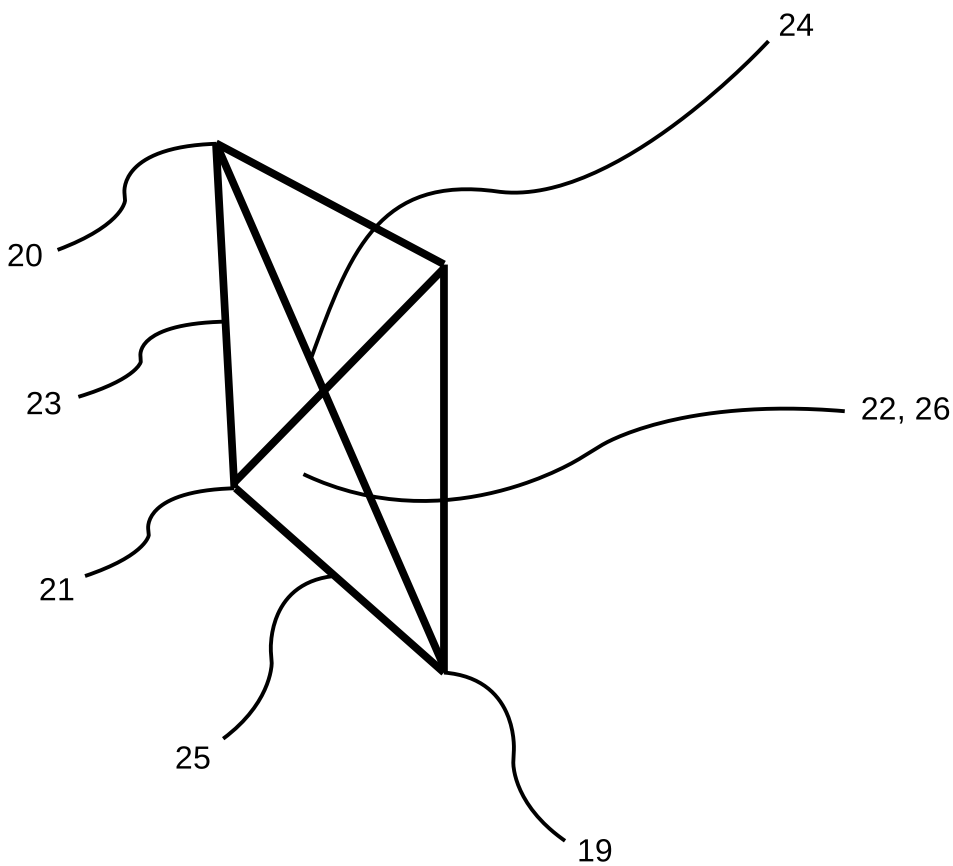

Fig. 15

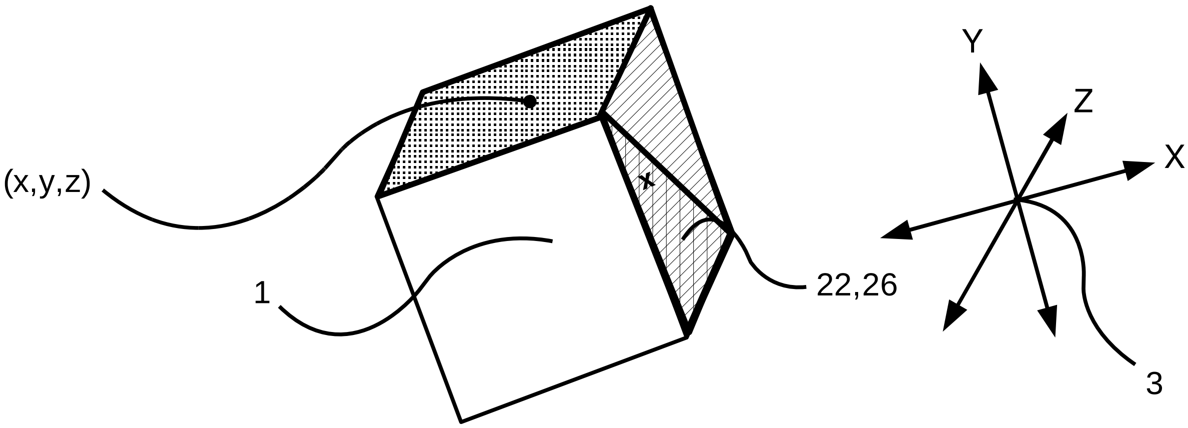

Fig. 16

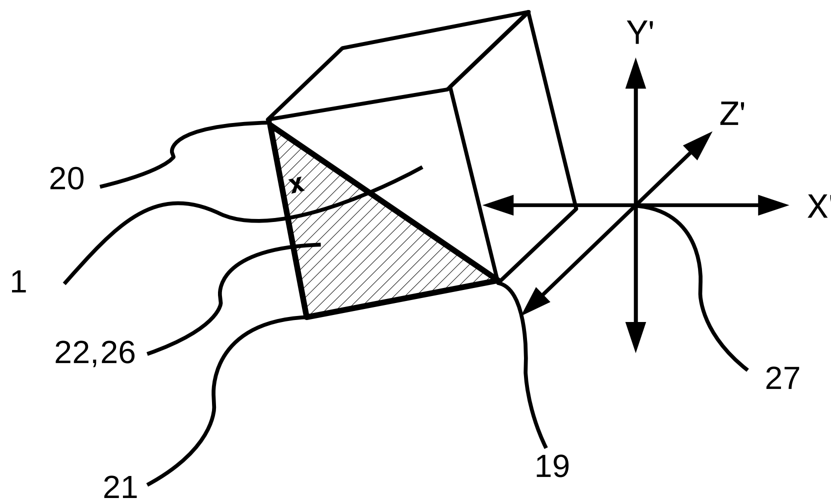

Fig. 17

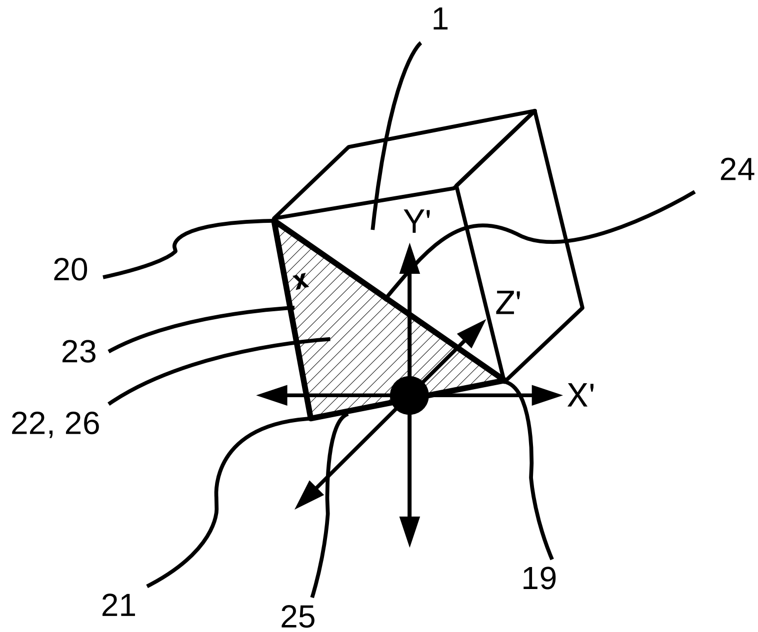

Fig. 18

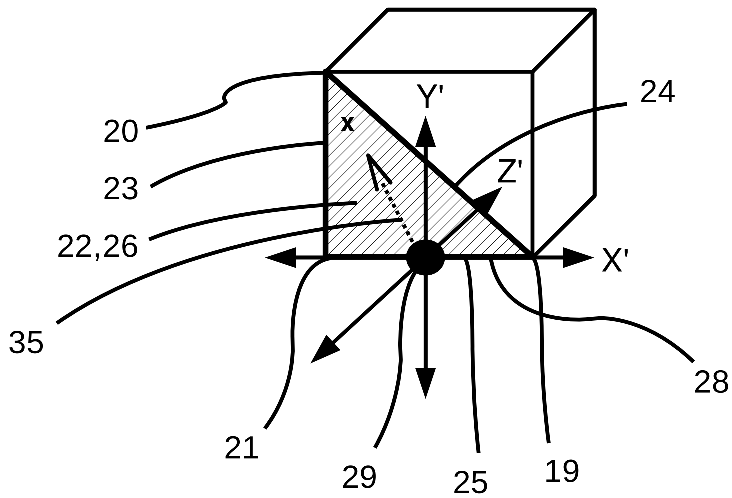

Fig. 19

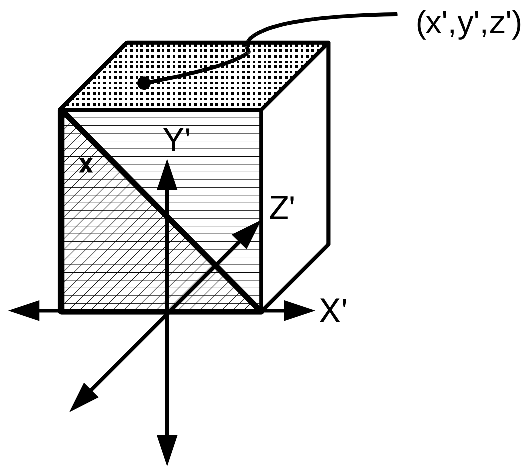

Fig. 20

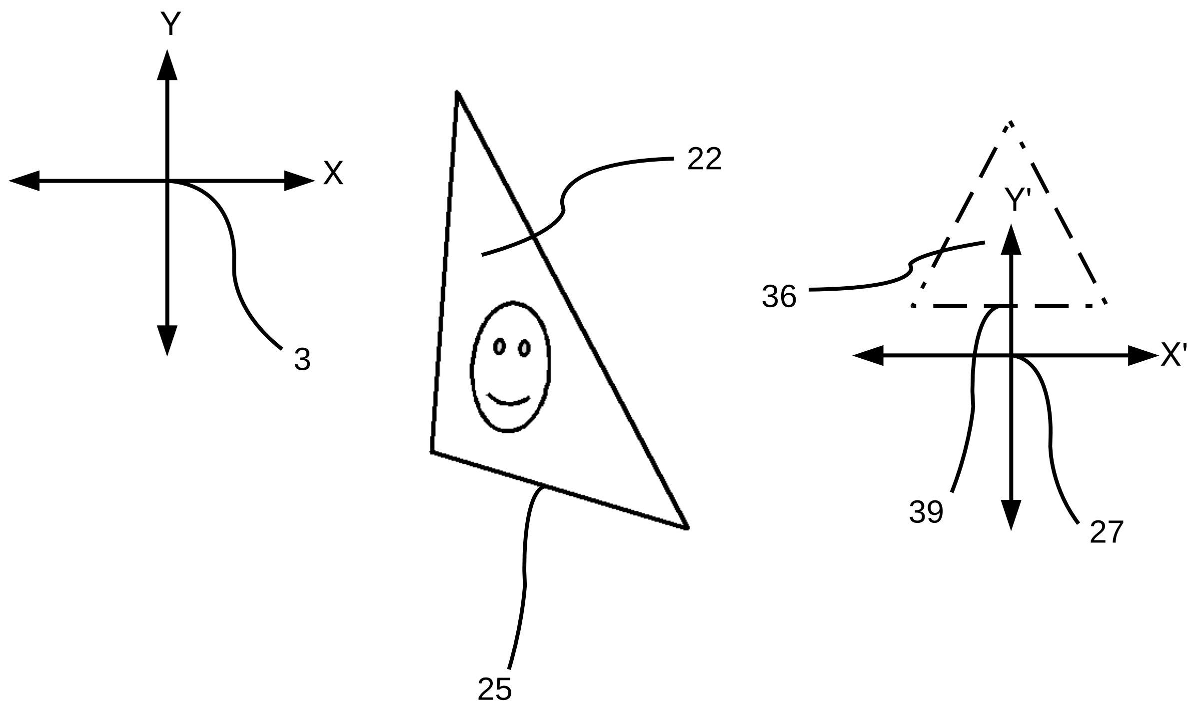

Fig. 21

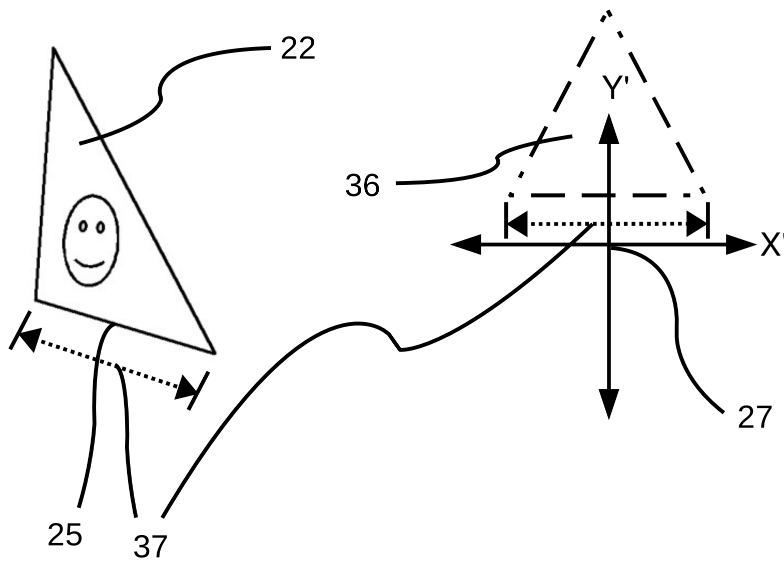

Fig. 22

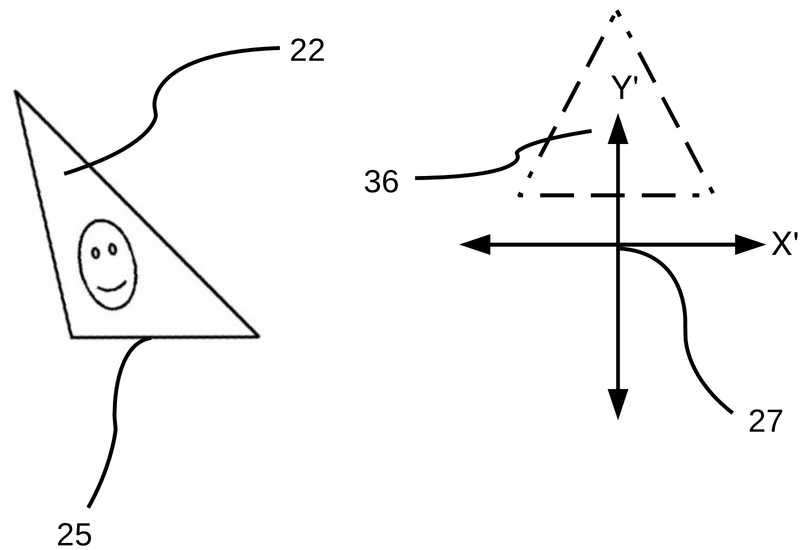

Fig. 23

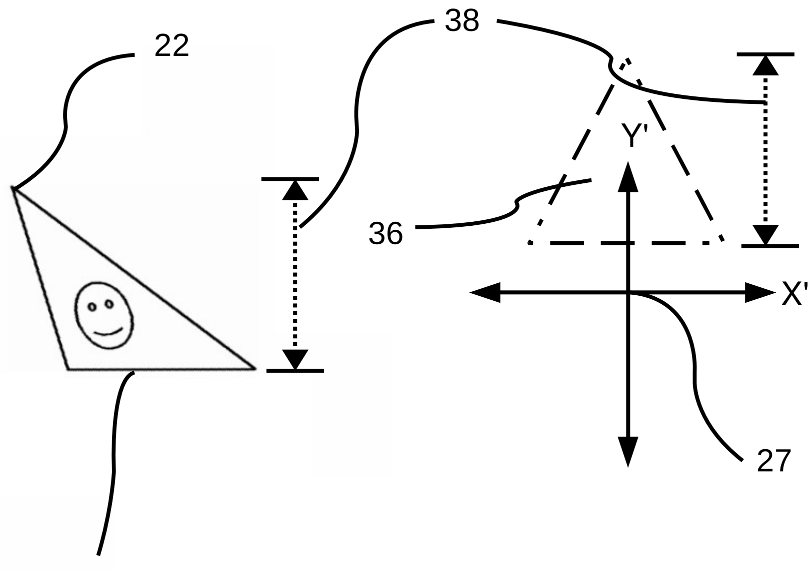

Fig. 24

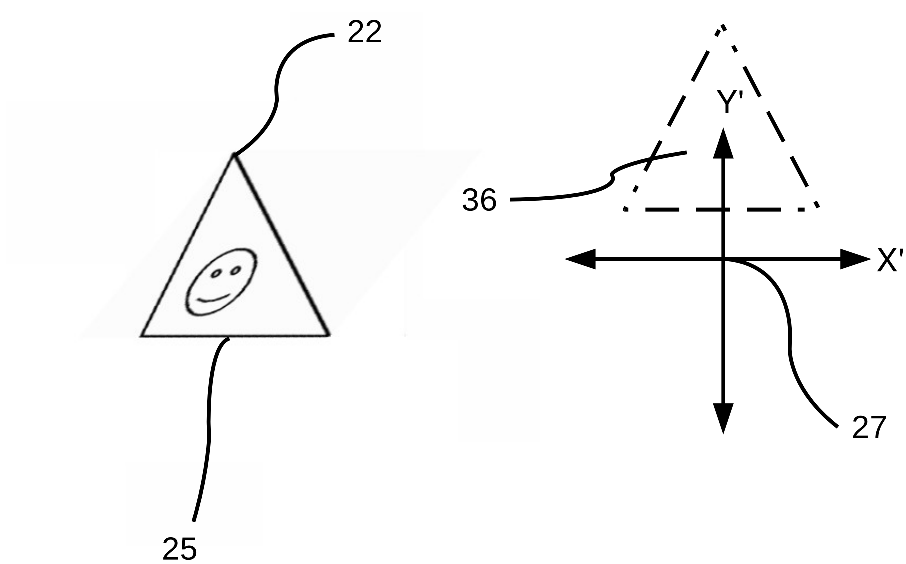

Fig. 25

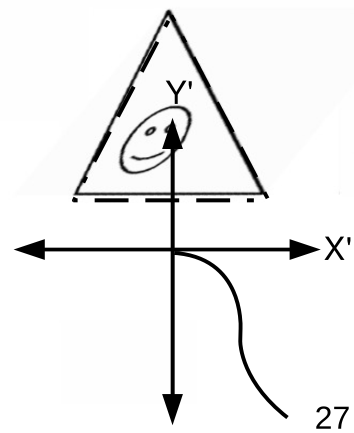

Fig. 26

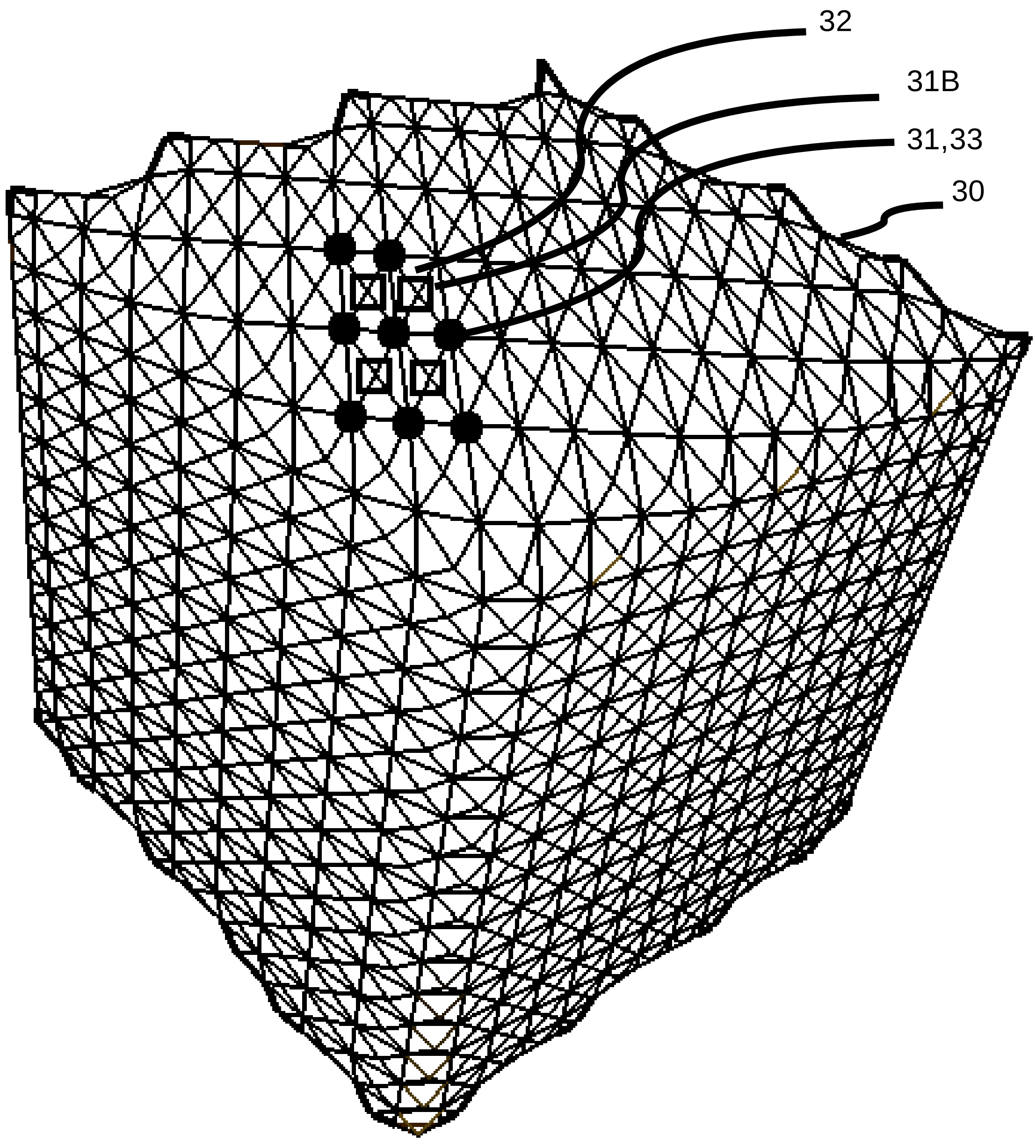

Fig. 27

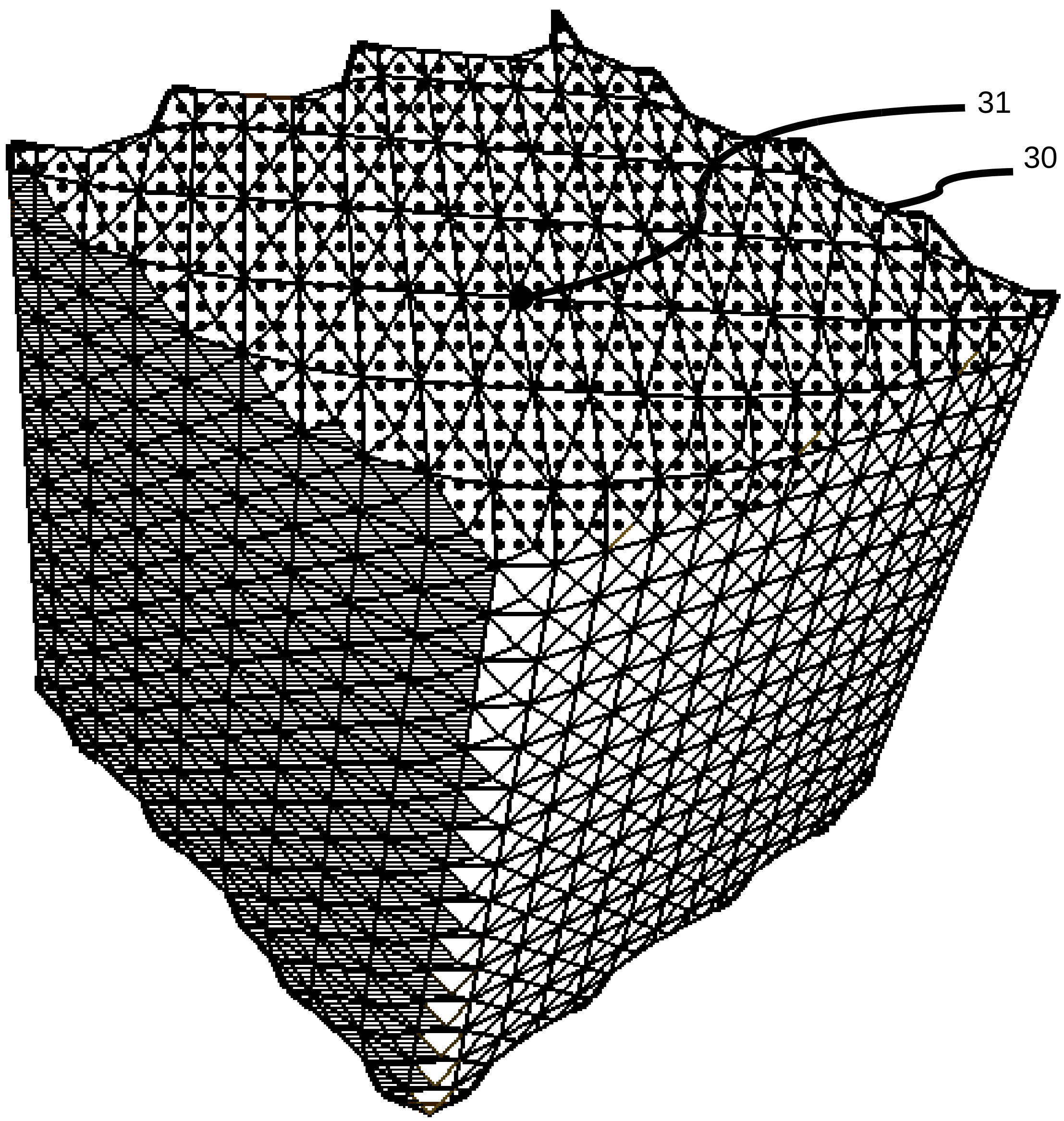

Fig. 28

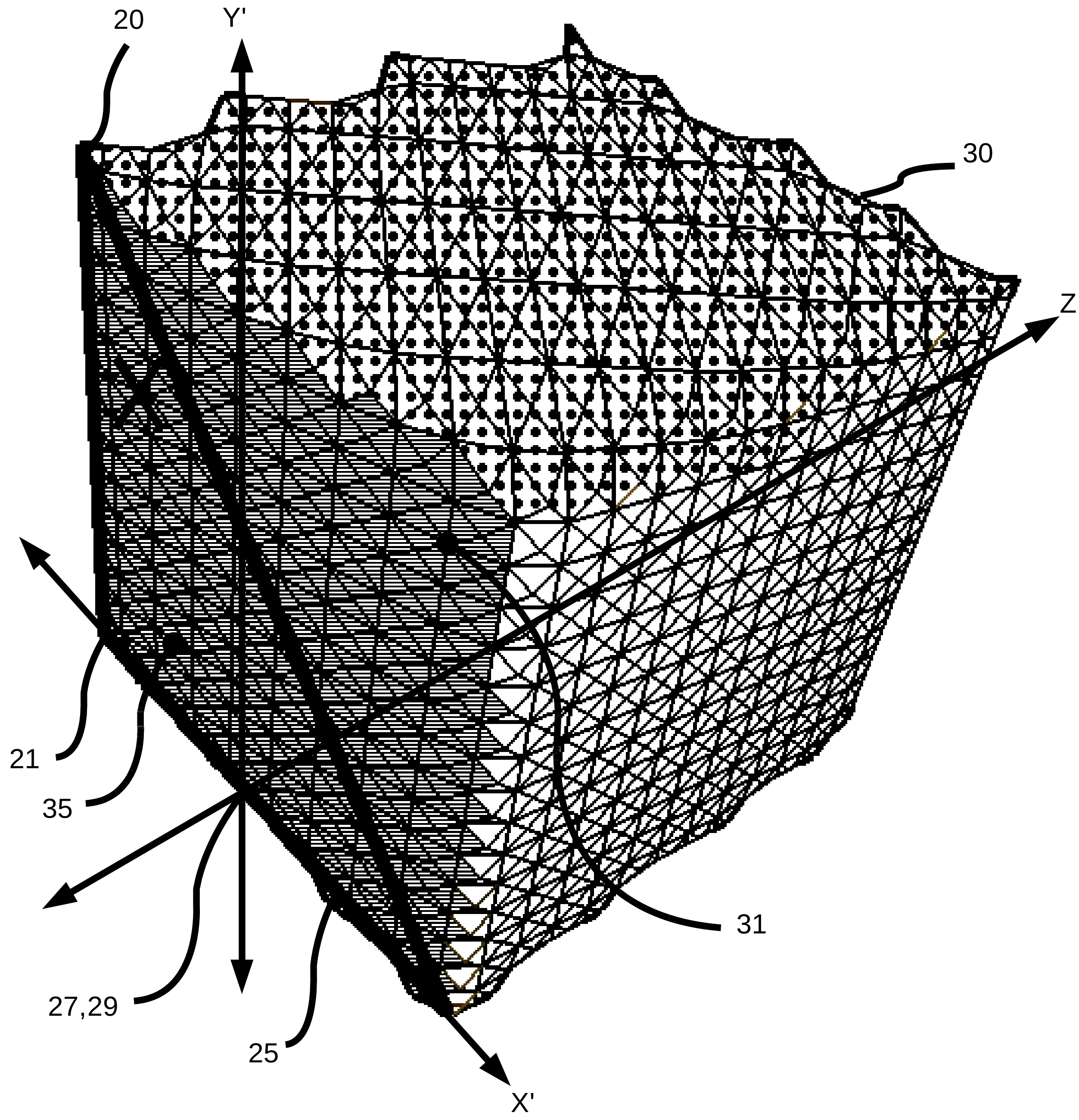

Fig. 29

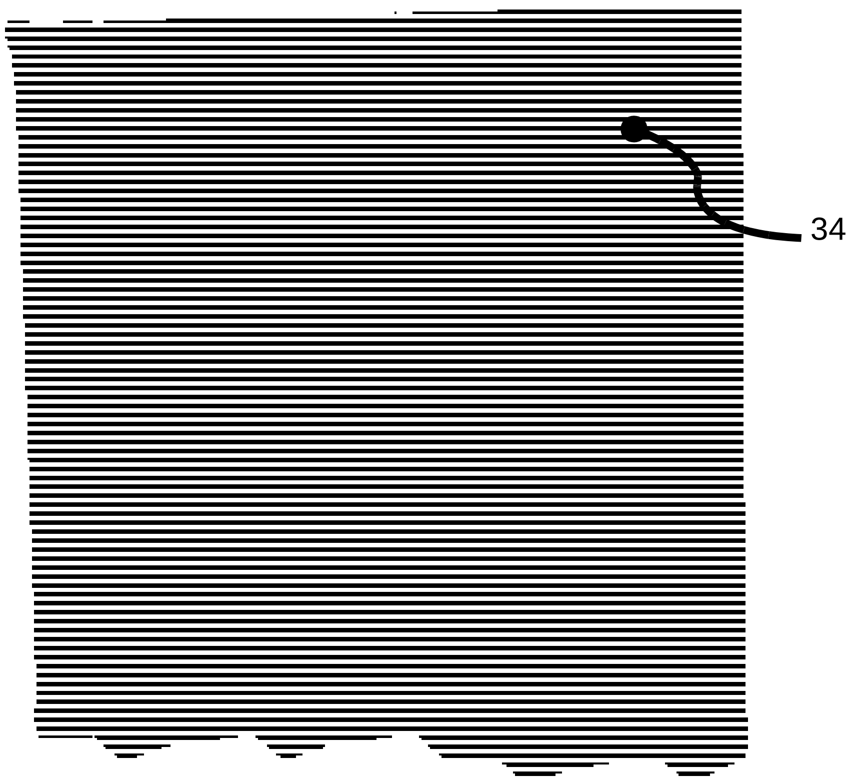

Fig. 30