CLAIMS

1. A method for deriving a 3D data from image data, comprising:

receiving, from at least one camera, image data representing an environment, wherein the image data comprises a time-ordered sequence of frames, the time-ordered sequence of frames comprising frames collected from each camera of the at least one cameras;

detecting, from the image data, at least one object within the environment, comprising detecting said at least one object in each frame of the time-ordered sequence of frames;

classifying the at least one detected object, comprising classifying the at least one detected object in each frame of the sequence of frames, wherein the method comprises, for each classified object of the classified at least one objects:

determining a 2D skeleton of the classified object by implementing a neural network to identify features of the classified object in the image data corresponding to the classified object; and

constructing a 3D skeleton for the classified object, comprising mapping the determined 2D skeleton to 3D, the mapping comprising:

determining a plurality of 3D skeletons to form a time-varying 3D skeleton, and

applying a recursive estimator across the time-ordered sequence of frames to determine a time-varying 3D skeleton for the classified object, wherein a derived 3D skeleton at a first time instance is used as a further input to the derivation of a second 3D skeleton at a later second time instance,

wherein each classified object of the classified at least one objects is tracked across the time-ordered sequence of frames by implementing a recursive estimator.

2. The method of claim 1, wherein each of said at least one object comprises a plurality of associated sub-objects, and wherein, for each of said at least one object:

detecting at least one object comprises detecting each of the plurality of associated sub-objects;

classifying at least one object comprises classifying each of the plurality of associated sub-objects;

determining the 2D skeleton of the classified object comprises identifying features of each of the plurality of classified sub-objects in the image data corresponding to the classified sub-objects.

3. The method of claim 1 or claim 2, wherein, for each classified object, mapping the determined 2D skeleton to 3D comprises:

implementation of a neural network; and/or

applying statistical and/or probabilistic methods to the determined 2D skeleton and applying holonomic constraints to the statistical and/or probabilistic methods.

4. The method of claim 3, wherein classifying the at least one detected object comprises classifying a first detected object of the at least one detected object as a human object, and, when statistical and/or probabilistic methods are applied to the determined 2D skeleton, the holonomic constraints comprise human anatomical holonomic constraints.

5. The method of claim 3 or claim 4, wherein the 3D skeleton comprises an anchor point and a plurality of child points, each child point is defined by applying a transformation to the anchor point or another child point of the plurality of child points, wherein the holonomic constraints define the degrees of freedom of the transformation defining each child point, and/or the range of possible values for each degree of freedom of the transformation.

6. The method of any preceding claim, wherein classifying the at least one detected object comprises classifying a second detected object of the at least one detected object as a non-human object.

7. The method of any preceding claim, wherein the at least one camera comprises a plurality of cameras;

wherein detecting at least one object comprises detecting said at least one object in image data collected from each camera;

classifying the at least one object comprises classifying said at least one object detected in the image data from each camera;

determining the 2D skeleton for each classified object comprises determining a plurality of 2D skeletons, each of said plurality of 2D skeletons determined from identifying features of the classified object in image data collected from a different camera of the plurality of cameras; and

wherein constructing the 3D skeleton comprises combining the determined plurality of determined 2D skeletons.

8. The method of claim 1, wherein the application of recursive estimators comprises applying time-dependent holonomic constraints.

9. The method of claim 8, wherein:

the 3D skeleton comprises an anchor point and a plurality of child points, each child point is defined by applying a transformation to the anchor point or another child point of the plurality of child points;

the time-varying 3D skeleton is defined by time-variable transformations for each child point, and the time-dependent holonomic constraints define, for each point in time, the degrees of freedom of transformation defining each child point and/or the range of possible values for each degree of freedom of the transformation.

10. The method of any one of claims 1 – 9, wherein the at least one camera comprises at least one first type of camera and at least one second type of camera, wherein each of the first type of camera captures image data at a first framerate, and each second type of camera captures image data at a second framerate that is different to the first framerate.

11. The method of any one of claims 1 – 10, further comprising constructing, from the image data, a 3D model of the environment.

12. The method of claim 11, wherein the 3D environment model is constructed by

applying Simultaneous Localisation and Mapping techniques to the image data to estimate a position and an orientation of each camera of the at least one camera and to construct an environment map; and

constructing the 3D model by mapping the environment map to a pre-determined 3D model of the environment.

13. The method of claim 11 or claim 12 further comprising integrating the time-varying 3D skeleton with the 3D environment model to construct a time-varying integrated environment.

14. The method of any preceding claim, the method further comprising, for each classified object of the classified at least one objects:

constructing a 3D avatar for the classified object comprising integrating the constructed 3D skeleton with a 3D model corresponding to the classified object.

15. The method of claim 14, wherein the method further comprises capturing the 3D model corresponding to the classified object at a higher resolution than the resolution of the image data.

16. The method of claim 11 or 12, further comprising:

for each classified object of the classified at least one objects: constructing a 3D avatar for the classified object comprising integrating the constructed 3D skeleton with a 3D model corresponding to the classified object, and

integrating the determined 3D environment model with the constructed 3D avatar for each of the classified objects to construct a 3D model of an integrated environment, said integrated environment comprising said environment and said objects within the environment.

17. The method of claim 16, further comprising refining the construction of the integrated environment by applying time-dependent smoothing to the constructed 3D avatar, using filtering methods.

18. The method of claim 16 or claim 17, further comprising refining the construction of the integrated environment by constraining interactions between each constructed 3D avatar and the 3D environment model.

19. The method of any one of claims 16 – 18, wherein constructing a 3D avatar for each of the classified objects comprises constructing a plurality of 3D avatars, and the method further comprises refining the construction of the integrated environment by constraining interactions between the plurality of 3D avatars.

20. A method for deriving a 3D data from image data of a sporting environment for use in determining compliance with an offside rule in a sporting match, said method comprising:

receiving, from at least one camera, image data representing the environment wherein the image data comprises a time-ordered sequence of frames, the time-ordered sequence of frames comprising frames collected from each camera of the at least one cameras;

detecting, from the image data, a plurality of objects within the environment comprising detecting the plurality of objects in each frame of the time-ordered sequence of frames;

classifying a plurality of the detected objects as players, comprising classifying the plurality of the detected objects as players in each frame of the sequence of frames, wherein the method comprises, for each player:

determining a 2D skeleton of the player by implementing a neural network to identify features of the payer in the image data corresponding to the player;

constructing a 3D skeleton for the player, comprising mapping the determined 2D skeleton to 3D, the mapping comprising:

determining a plurality of 3D skeletons to form a time-varying 3D skeleton, and

applying a recursive estimator across the time-ordered sequence of frames to determine a time-varying 3D skeleton for the classified objects, wherein a derived 3D skeleton at a first time instance is used as a further input to the derivation of a second 3D skeleton at a later second time instance,

wherein each classified object of the classified plurality of objects is tracked across the time-ordered sequence of frames by implementing a recursive estimator; and

integrating the constructed 3D skeleton with a 3D model corresponding to the classified object to construct a 3D avatar for the player,

the method further comprising in response to a user-detected or automatically-detected event:

defining a two-dimensional plane within the environment;

determining on which side of the two-dimensional plane predefined portions of the avatars of the players may be found.

21. A method for deriving a 3D data from image data of a sporting environment for use in determining compliance with an offside rule in a sporting match, said method comprising:

receiving, from at least one camera, image data representing the environment, wherein the image data comprises a time-ordered sequence of frames, the time-ordered sequence of frames comprising frames collected from each camera of the at least one cameras;

detecting, from the image data, a plurality of objects within the environment, comprising detecting the plurality of objects in each frame of the time-ordered sequence of frames;

classifying a plurality of the detected objects as players, comprising classifying the plurality of the detected objects as players in each frame of the sequence of frames, wherein the method comprises, for each player:

determining a 2D skeleton of the player by implementing a neural network to identify features of the payer in the image data corresponding to the player;

constructing a 3D skeleton for the player, comprising mapping the determined 2D skeleton to 3D, the mapping comprising:

determining a plurality of 3D skeletons to form a time-varying 3D skeleton, and

applying a recursive estimator across the time-ordered sequence of frames to determine a time-varying 3D skeleton for the classified object, wherein a derived 3D skeleton at a first time instance is used as a further input to the derivation of a second 3D skeleton at a later second time instance,

wherein each classified object of the classified plurality of objects is tracked across the time-ordered sequence of frames by implementing a recursive estimator; and

the method further comprising in response to a user-detected or automatically-detected event:

defining a two-dimensional plane within the environment;

determining on which side of the two-dimensional plane predefined portions of the skeletons of the players may be found.

22. A computer comprising a processor configured to carry out the method of any preceding claim.

23. A computer program comprising instructions that, when executed by a processor, cause the processor to carry out the method of any of claims 1 – 21.