

CLAIMS

The invention claimed is:

1. A device for imaging a body of a user, the device comprising:
a controller;
storage storing electronic program instructions for controlling the controller;
a display for displaying a user interface; and
an image capturing device configured to capture visual images;
wherein the controller is operable, under control of the electronic program instructions, to:
 - receive input at least in part via the image capturing device, the input comprising a first representation of the body, and the first representation comprising one or more visual representations of the body;
 - display the first representation via the display;
 - generate, using (i) a model of skeleton joint positions, and (ii) at least a weight of the user, a user-specific skeleton that will appear on the display once the input is received;
 - enable the user to align the body in the first representation with the user-specific skeleton, at least in part by (i) displaying the user-specific skeleton along with one or more real time captured images of the body and (ii) instructing the user to move in such a manner that the displayed body is aligned to the displayed user-specific skeleton;
 - process the first representation, when the displayed body has been aligned with the displayed user-specific skeleton, by segmenting the first representation of the body to obtain a plurality of silhouettes which correspond to projected shadows of a substantially true three dimensional scan of the body;
 - generate a second representation of the body on the basis of the plurality of silhouettes; and
 - display the generated second representation via the display.
2. A device according to claim 1, wherein the controller is operable, under control of the electronic program instructions, to:
 - generate the second representation of the body on the basis of the plurality of silhouettes and thousands of known human shapes learned offline using intelligent machine learning techniques.

3. A device according to claim 1, wherein the controller is also operable, under control of the electronic program instructions, to:

instruct the user via audible sounds, words, or speech to align parts of the displayed body to the displayed user-specific skeleton, wherein the electronic program instructions are operable to control the alignment process by errors calculated between characteristics including shape features, pose features, and spatiotemporal features that are extracted from the generated skeleton and the one or more real time captured images of the body.

4. A device according to claim 3, wherein the controller is also operable, under control of the electronic program instructions, to:

calculate, on the basis of user height information submitted and image height and width in pixels, and using blob analysis of binary images, projection theories and camera models, the following:

initial estimates of intrinsic and extrinsic parameters of a capturing camera which includes camera position and orientation in each image, defined as pose P ; and

initial estimates of joint kinematics of a skeletal model representing a skeleton of the body, defined as JK , including 3D position and 3D orientation of each joint of the skeletal model.

5. A device according to claim 3, wherein the controller is also operable, under control of the electronic program instructions, to:

predict, on the basis of user height and weight information submitted, or the user height information only, an initial on-average avatar, defined as Av , which varies with the user's entered height, weight or other body measurements if known; and

rig the on-average avatar Av to a reference skeleton of size N -joints with known skeletal model JK in a reference pose, and a bone weight/height matrix defined as W .

6. A device according to claim 5, wherein the matrix W is calculated offline just once during a learning process of the prediction process, then saved together with the reference skeletal model JK to be used for prediction or generation of other avatars, wherein W is used to constrain, control and model the relationship between joints, bones and the actual 3D avatar surface represented by its vertices V , edges E and faces F .

7. A device according to claim 6, wherein the process of predicting the initial on-average avatar Av follows a multivariate-based machine learning approach.
8. A device according to claim 1, wherein the input further comprises a classification of the body, and the controller is operable, under control of the electronic program instructions, to:
- on the basis of the classification of the body, obtain data corresponding to the body classification;
 - process the first representation further by comparing the first representation and the obtained data; and
 - generate the second representation of the body further on the basis of the comparison.
9. A device according to claim 8, wherein the obtained data comprises at least one of:
- a template;
 - an earlier representation of the body; and
 - an integration of, or an integration of data associated with, one or more earlier representations of one or both of (i) the body, and (ii) other bodies.
10. A device according to claim 1, wherein the input is in part received via one or more of:
- a motion sensor;
 - an infra-red sensor;
 - a depth sensor;
 - a three dimensional imaging sensor;
 - an inertial sensor;
 - a Micro-Electromechanical (MEMS) sensor;
 - an acceleration sensor;
 - an orientation sensor;
 - a direction sensor; and
 - a position sensor.
11. A device according to claim 10, wherein the input is in part received via an orientation sensor operable to provide orientation data for use during capture of the one or more visual representations of the body to facilitate alignment thereof to a plane for increased accuracy.

12. A device according to claim 11, wherein the one or more visual representations of the body include at least one photograph of a front view of the body and at least one photograph of a side view of the body, and wherein the photographs comprise at least one of:

- standard two dimensional (2D) binary, gray or color images;
- depth images with or without one or both of colors and texture;
- a complete three dimensional (3D) point cloud or a number of incomplete point clouds of the body with or without one or both of colors and texture; or
- a three dimensional (3D) mesh of the body with or without one or both of colors and texture.

13. A device according to claim 12, wherein the controller is further operable, under control of the electronic program instructions, to:

- instruct the user via audible sounds, words, or speech to align parts of the body to the displayed user-specific skeleton, wherein the electronic program instructions are operable to control the alignment process by errors calculated between characteristics and various data including shape features, pose features, and spatiotemporal features that are extracted from the generated skeleton and the one or more real time captured images of the body;

- segment one or more foreground regions or features of the one or more visual representations of the body that are definitely, highly likely, or likely/probably the body;

- convert the one or more segmented foreground regions or features of the one or more visual representations into respective ones of the plurality of silhouettes;

- use the one or more segmented foreground regions or features and their respective silhouettes to one or more of (i) construct a hull of a shape of the body, (ii) extract features, or (iii) extract measurements of key points; and

- use one or more of the hull, features, or key point measurements to one or more of modify, rig, and morph a 3D model of an average body of the selected template to create a modified subject-specific 3D model image being the second representation.

14. A device according to claim 13, wherein in the case of depth images, point clouds and meshes, any of which are with or without one or both of colors and textures, the controller is operable, under control of the electronic program instructions, to reconstruct a three dimensional subject-specific shape of the body.

15. A method for imaging a body of a user, the method comprising:
storing electronic program instructions for controlling a controller; and
controlling the controller via the electronic program instructions, to:
 receive an input via an image capturing device configured to capture visual
images, the input comprising a first representation of the body, and the first
representation comprising one or more visual representations of the body;
 display the first representation on a user display;
 generate, using (i) a model of skeleton joint positions, and (ii) at least a weight
of the user, a user-specific skeleton that will appear on the display once the input is
received;
 enable the user to align the body in the first representation with the user-
specific skeleton, at least in part by (i) displaying the user-specific skeleton along with
one or more real time captured images of the body and (ii) instructing the user to
move in such a manner that the displayed body is aligned to the displayed user-
specific skeleton;
 process the first representation, when the displayed body has been aligned
with the displayed user-specific skeleton, by segmenting the first representation of
the body to obtain a plurality of silhouettes which correspond to projected shadows of
a substantially true three dimensional scan of the body;
 generate a second representation of the body on the basis of the plurality of
silhouettes; and
 display the generated second representation.

16. A method according to claim 15, wherein enabling the user to align the body with the
user-specific skeleton includes instructing the user via audible sounds, words, or speech to
align parts of the body to the displayed user-specific skeleton, and wherein the electronic
program instructions are operable to control the alignment process by errors calculated
between characteristics including shape features, pose features, and spatiotemporal
features that are extracted from the generated skeleton and the one or more real time
captured-images of the body.

17. A method according to claim 16, further comprising controlling the controller via the
electronic program instructions, to:
 calculate, on the basis of user height information submitted and image height and
width in pixels, and using blob analysis of binary images, projection theories and camera
models, the following:

initial estimates of intrinsic and extrinsic parameters of the capturing camera which includes camera position and orientation in each image, defined as pose P;
and

initial estimates of joint kinematics of a skeletal model representing a skeleton of the body, defined as JK, including 3D position and 3D orientation of each joint of the skeletal model.

18. A method according to claim 17, further comprising controlling the controller via the electronic program instructions, to:

predict, on the basis of user height and weight information submitted, or the user height information only, an initial on-average avatar, defined as Av, which varies with the user's entered height, weight or other body measurements if known; and

rig the on-average avatar Av to a reference skeleton of size N-joints with known skeletal model JK in a reference pose, and

a bone weight/height matrix defined as W.

19. A method according to claim 18, wherein the matrix W is calculated offline just once during a learning process of the prediction process, then saved together with the reference skeletal model JK to be used for prediction or generation of other avatars, wherein W is used to constrain, control and model the relationship between joints, bones and the actual 3D avatar surface represented by its vertices V, edges E and faces F.

20. A method according to claim 19, wherein the process of predicting the initial on-average avatar Av follows a multivariate-based machine learning approach.

21. A method according to claim 20, wherein the multivariate-based machine learning approach comprises offline machine intelligence learning of human shapes using 3D features extracted from a plurality of rigged and rendered three dimensional scans of real humans of different ages and poses.

22. A method according to claim 21, wherein the multivariate-based machine learning approach further comprises machine intelligence learning of various statistical relationships between different body measurements defined as vector $M=(m_1, m_2, \dots, m_L)$ with L number of different measurements wherein, in use, one or more measurements can be predicted given one or more different measurements and an on-average avatar Av can be predicted given one or more of these measurements.

23. A method according to claim 22, wherein in order to deform or simply animate an avatar to a new avatar defined as Av1 of the body as represented in a new first representation, the reference or on-average avatar data (V, E, F, JK, W) and a known, or an estimate of, user joint kinematics defined as JK1 of the new first representation are fed to a cost function that optimizes and deforms Av to Av1 subject to a number of physical constraints known or learned from natural human motion wherein, in use, the new animated avatar Av1 with same body measurement as the on-average avatar Av is a function of the reference or on-average data, such that $Av1=f(Av, W, JK, JK1)$.

24. A method according to claim 23, wherein a function is derived that combines two weighted energy minimization functions:

- a surface smoothness function utilizing a Laplacian cotangent matrix which uses V, F and E; and,

- a bone attachment function which uses V, F and W to ensure that correspondence is constrained between the avatar vertices and its skeletal structure.

25. A method according to claim 15, wherein the input further comprises a classification of the body, and further comprising controlling the controller via the electronic program instructions, to:

- on the basis of the classification of the body, obtain data corresponding to the body classification;

- process the first representation further by comparing the first representation and the obtained data; and

- generate the second representation of the body further on the basis of the comparison.

26. A non-transitory computer readable storage medium on which is stored instructions that, when executed by one or more processors causes the one or more processors to:

- receive an input via an image capturing device configured to capture visual images, the input comprising a first representation of the body, and the first representation comprising one or more visual representation of the body;

- display the first representation on a user display;

- generate, using (i) a model of skeleton joint positions, and (ii) at least a weight of the user, a user-specific skeleton that will appear on the display once the input is received;

- enable the user to align the body in the first representation with the user-specific skeleton, at least in part by (i) displaying the user-specific skeleton along with one or more

real time captured images of the body and (ii) instructing the user to move in such a manner that the displayed body is aligned to the displayed user-specific skeleton;

process the first representation, when the displayed body has been aligned with the displayed user-specific skeleton, by segmenting the first representation of the body to obtain a plurality of silhouettes which correspond to projected shadow of a substantially true three dimensional scan of the body;

generate a second representation of the body on the basis of the plurality of silhouettes; and

display the generated second representation.