(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
28 November 2013 (28.11.2013)

WIPO | PCT

(10) International Publication Number
WO 2013/175228 A1

(51) International Patent Classification: *G06T 7/00* (2006.01)

(21) International Application Number: PCT/GB2013/051371

(22) International Filing Date: 24 May 2013 (24.05.2013)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
1209382.9 25 May 2012 (25.05.2012) GB

(71) Applicant: **POIKOS LIMITED** [GB/GB]; Faraday Wharf, Holt Street, Birmingham, West Midlands B7 4BB (GB).

(72) Inventors: **WATSON, Eleanor**; Poikos Limited, Faraday Wharf, Holt Street, Birmingham, West Midlands B7 4BB (GB). **EVANS, David**; Poikos Limited, Faraday Street, Holt Street, Birmingham, West Midlands B7 4BB (GB).

(74) Agent: **BHIMANI, Alan**; Marks & Clerk LLP, Alpha Tower, Suffolk Street Queensway, Birmingham B1 1TT (GB).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV,

[Continued on next page]

(54) Title: BODY MEASUREMENT

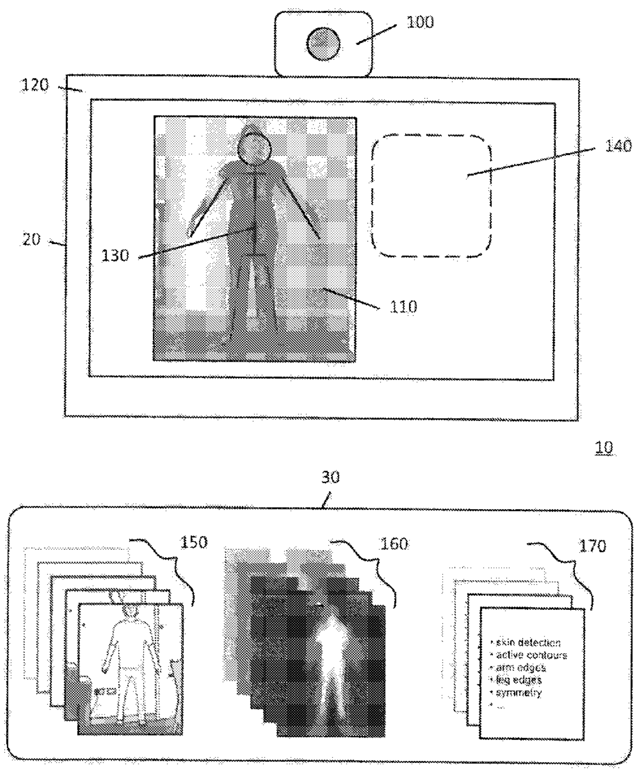

Fig 1

(57) Abstract: This invention relates to a method of generating three dimensional body data of a subject. The method includes the following steps. First or more source images of the subject are captured using a digital imaging device. The one or more source images are partitioned into a plurality of segments or probability distributions using one or more segmentation method, heuristics and/or predetermined mappings. The results of each segmentation method are combined to produce one or more unique probability maps representing the subject and the unique probability maps are then compared with a database of representations of three dimensional bodies to determine a closest match or best mapping between the or each unique probability map and a 'representation determined from the database. Three dimensional body data and'or measurements of the subject is then generated and may be. output based on the best mapping.

WO 2013/175228 A1

MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

Body Measurement

The present invention relates to a method and system for determining body measurements using images of the body. More specifically, the present invention defines a method and system for generating three dimensional body image data of a subject by comparison to a database of existing known three dimensional models to obtain a best mapping. From this mapping, body shape specific data can be determined, including measurements.

Conventional body scanners and body measurement devices rely on high resolution depth sensors, fixed patterns of light or known camera angles to generate three dimensional images of the human body. Such techniques require specialist equipment and present a high burden on the user who is required to travel to specialised centres for such body measurement.

Obtaining accurate three dimensional body measurement data is particularly beneficial in the medical field and may be used, for example, to monitor changes in body volume that may occur as a result of medical conditions, for example renal failure. By assessing the changes in the patient's body images, it is possible to determine factors such as water retention and surgery success. Indeed, with the rise of telemedicine, the need to monitor and provide information about body measurements is likely to grow.

Furthermore, being able to generate accurate body characterisation and measurement is also of particular interest in the clothing industry. It can be appreciated that being able to provide a clothing manufacturer or retailer with accurate body dimensions would be beneficial to both the person seeking clothes and the companies involved in the supply of clothes, this is particularly beneficial for online purchases.

Current methods and devices for providing such measurements are either large and expensive, and/or require the use of complex equipment including detectors capable of determining the per-pixel depth or distance of the subject relative to a reference point. As a result they do not allow subjects, such as private individuals and patients, to monitor their three dimensional body shape easily or at home. Being able to provide a three dimensional body image simply, without using dedicated complex body scanners is therefore desirable. The present invention seeks to address the above problems.

Acording to the present invention there is provided a method of generating three dimensional body data of a subject; said method comprising the steps of:

i capturing one or more source images of the subject using a digital imaging device;

ii partitioning the one or more images into a plurality of segments using one or more segmentation method;

iii combining the results of each segmentation method in step ii) to produce one or more unique probability maps representing the subject;

iv comparing the one or more unique probability maps with a database of representations of three dimensional bodies to determine a closest match or best mapping between the or each unique probability map and a representation determined from the database; and

v generating three dimensional body data and/or measurements of the subject based on the best mapping.

Optionally one or more pre-processing stages may be applied to the one or more source images to enhance the available data prior to segmentation;

Partitioning the one or more images into a plurality of segments may comprise applying one or more image segmentation techniques, heuristics and/or predetermined mappings to partition the one or more enhanced images and/or original images into a plurality of segments and/or probability distributions;

Combining the results of each segmentation method may comprise using mathematical equations and/or heuristics.

Optionally an additional segmentation technique seeded with the results of step iii) may be applied to produce a refined unique probability map for each source image;

Comparing the one or more unique probability maps with a database of representations may comprise determining a best mapping between each or all of the unique probability maps and the database of existing known three dimensional models by applying a Gradient Descent method to minimise the difference between computer-generated silhouettes of the three dimensional data and the probability maps

Generating three dimensional body data and/or measurements may further comprise outputting three dimensional body data of the subject or other personal metrics based on the best mapping from step iv).

The invention allows for data representative of the three dimensional body image of a subject to be generated cheaply and easily based on 2 dimensional images obtainable from a webcam, cameraphone or other image capture device. Additionally, because dedicated imaging equipment is not necessary, a subject can easily update their generated three dimensional body image, allowing changes to their body shape to be easily tracked and monitored.

Generally, one or more probability maps and heuristic methods can be provided which identify the expected shape and pose of the subject in the one or more images. Typically, the probability maps are pre-rendered probability maps stored on the server either generated from earlier images and/or obtained from a pre-existing database of probability maps, and the heuristics are coded methods within the system. For example, if the subject poses in the images to present their left and right side profiles, the probability maps provided and heuristics applied are representative of the respective profile. In this case, an image with a subject posing in a left side profile pose can be compared to a left side profile probability map and processed using segmentation heuristics known to apply to left side profiles.

The one or more probability maps can also be provided based on the determined pose of the subject. The pose of the subject can be determined by feature tracking the subject in a sequence of frames, which may be from a video feed. A typical method of feature tracking utilises Lucas-Kanade Pyramids to calculate the optical flow of keypoints of the subject between images (or frames if applied to video). By taking a frame where the subject's position is known, key points in the image can be located (for example using corner detection techniques). Such points may be fingers, patterns on clothing, tattoos, skin blemishes etc.). A Lucas-Kanade Pyramid technique can then be used to follow these points between frames (or images). The points may be rechecked every few frames to prevent slippage. Such a technique allows the pose of the subject within images to be guessed more accurately and further allows the one or more probability maps to be provided based on the determined pose. Such a pose determination technique increases the flexibility of the system and is particularly useful for determining the pose of subjects from video.

In embodiments of the present invention, the pre-processing stage ii) may include:

- rescaling the image to reduce computational overhead;
- applying noise reduction filters such as Anisotropic Diffusion;
- counteracting known camera characteristics, such as barrel distortion and patterns introduced by the layout of pixels within the CCD (Charge-Coupled Device) sensor;
- applying a per-pixel lighting boost to enhance details in shadow, typically this would follow the formula $\mathbf{I}_{new}(\mathbf{x}) = \mathbf{I}_{raw}(\mathbf{x})^{\alpha}$ where $\mathbf{x}$ is the current pixel, $\mathbf{I}_{new}(\mathbf{x})$ is the enhanced intensity value for the pixel $\mathbf{x}$ (between 0 and 1), $\mathbf{I}_{raw}(\mathbf{x})$ is the original intensity value for the pixel $\mathbf{x}$ (between 0 and 1), and α is a number which may be constant or based on properties of the image. Embodiments may vary α according to the location within the image, such as boosting the bottom of the image more strongly as in this example:

$$\alpha_t = \frac{log(b)}{log\left(\overline{I_{raw}(x \in top\ quarter\ of\ image)}\right)},\ \alpha_b = \frac{log(b)}{log\left(\overline{I_{raw}(x \in bottom\ quarter\ of\ image)}\right)},$$

$$\alpha = \alpha_t + \tfrac{y}{h}\left(\alpha_b - \alpha_t\right)$$

where $\mathbf{y}$ is the current pixel's vertical coordinate and $\mathbf{h}$ is the height of the image.

Additionally the collection of image segmentation methods applied to the one or more images in stage iii) may include:

- invariant methods, such as pre-rendered probability maps for a particular view;
- per-pixel or moving-window comparisons, for example looking for pre-determined ranges of colour which correspond to skin tones or a texture which matches denim;

- learning-based techniques, for example locating a face in a front view to learn the individual's skin colour and/or clothing colours, and/or using the chosen pre-rendered probability map to learn the background colours. The identified colours (in RGB space or otherwise) and/or locations (in image space) can be analysed (for example using a Gaussian Mixture Model) to produce an expected colour distribution for each segment, which may then be used to generate a segmentation of the image;
- iterative optimisations, for example using Active Contours (seeded with a pre-defined human shape according to the current view and/or pose detected by other methods) to maximise the image's edge intensity along the perimeter of the contour, which may also be combined with rigid-body or soft-body simulations for enhanced results;
- comparing regions of pixels in one view against the corresponding region in another view, for example comparing the left and right sides of front-with-arms-out and side-view images to identify arms (assuming the background remains constant between the images). Embodiments may include an additional image of only-background (subject out of frame) for this purpose;
- edge analysis techniques, for example searching for edges with particular properties such as a semi-circle shape (corresponding to feet) or particular angle (corresponding to arms, legs, or torso). Edges may be detected using one or more techniques such as application of the Sobel Operator, the Canny Edge Detector algorithm, and custom methods;
- generalised segmentation methods such as the Graph Cut method or multi-scale methods such as Segmentation by Weighted Aggregation;
- customised heuristics, such as identification of shoulder points in a front view to define a region for the torso, or identification of intersecting lines to locate armpits and the crotch point.

Typically the methods are applied to the image encoded in YCbCr space, with some methods using only the luminosity channel and others potentially assigning extra weight to the Cb and Cr channels, however other colour spaces such as RGB or sRGB can be used.

Segmentation methods which return a plurality of segments (for example: head, shirt, chair, desk) may be further processed by analysing the shape, colour, texture and/or position of each segment. By comparing these metrics against predefined or otherwise calculated expectations, the segments can be reduced into a simpler foreground/background binary, or given a per-segment probability.

If multiple segmentation methods are used, embodiments may combine their results into a single unique probability map per view via various methods. The method used may depend upon the current view and/or segmentation method, or several methods can be attempted, with a rating function deciding the best result. Possible methods for combining results p_1 and p_2 into p_T include:

- averaging: $p_T = \frac{p_1 + p_2}{2}$;
- adding regions known to be the subject: $p_T = \max(p_1, p_2)$;
- adding regions known to be the background: $p_T = \min(p_1, p_2)$;
- combining probabilities: $p_T = \frac{p_0 p_1}{p_0 p_1 + (1 - p_0)(1 - p_1)}$ with special cases for $p_0 = 0, p_1 = 1$ and $p_0 = 1, p_1 = 0$;
- more complex methods such as $p_T = p_0^{2^{n(0.5 - p_1)}}$ for some $n > 0$ or $p_T = 2Ap_1^2 - (A + 2(x^c - x))y + x^c, A = 1 - (1 - x)^c + x^c - 2x$ for some $c > 1$. These particular methods are applicable when p_0 contains some base expectation (such as a pre-rendered probability map) which must be adjusted according to p_1 but maintain hard limits.

It may be appreciated that many of these methods trivially extend to combine more than two results, or can be chained to combine arbitrary numbers of segmentation results. It may also be appreciated that weightings can be applied to each segmentation at this stage, determined manually, empirically, and/or automatically based on properties of the segmentation (such as degree of symmetry).

After producing a single unique probability map for each source image, embodiments may refine the map by the use of further segmentation techniques, such as Segmentation by Weighted Aggregation or Grab Cut. In the case of a Grab Cut, an input of "known foreground", "probably foreground", "probably background" and "known background" is given, which can be generated using trivial thresholding of the probability maps. It can also be appreciated that this 4-state map, or variations thereupon, could be considered as probability maps in their own right, and could therefore be the direct output of previous segmentation stages.

Embodiments of the invention further allow for compensation for variable line sharpness produced by uneven lighting conditions in the source images, with the same method also allowing the final probability maps to be upscaled to equal or greater resolution than the original source images. This is achieved by blurring of the edges of the unique probability map generated by the above method to produce a boundary of uncertain pixels. The blurring of the edges results in a probability gradient in the unique probability map from the background to the subject. The direction and strength of this gradient can then be identified and used to indicate the direction to the subject at each uncertain pixel with a vector line. Typically, the direction of this vector line is substantially perpendicular to the edge. From these gradients a vectored outline probability map can be generated, said map providing a direction vector (vector line) for each uncertain pixel (typically edge pixels) denoting the local expected direction of the subject relative to the pixel.

The vectored outline probability map can then be applied to the original image or images. By following the path indicated by the vectored outline probability map as applied to the original image, it is possible to determine the pixels in the original image that lie on each vector line with the highest contrast. These pixels represent the boundary pixels located on the edge of the subject, which separate the subject from the background of the image. This boundary is optimal in the case of no noise and any blurring of the image being gaussian, which is a good approximation for camera images. Additionally, comparing the colour and/or intensity of the boundary pixel with the colours and/or intensity of neighbouring pixels can enable an estimated true boundary for each boundary pixel to be ascertained with sub-pixel resolution. The estimated true boundary for each boundary pixel may be calculated using: $b_\delta = 1 - \frac{c_B - c_A}{c_C - c_A}$, where b_δ gives the distance along the vector between the pixel boundary and the true sub-pixel boundary, c_B is the intensity of the current pixel, c_A is the intensity of the previous pixel along the vector and c_c is the intensity of the next pixel along the vector. Once the estimated true boundary has been determined, this can be applied to the corresponding location on the unique probability map to provide an estimated true boundary for each boundary pixel of the unique probability map.

The final unique probability maps provide an outline of the subject for each view, where the regions determined to definitely be the subject have a probability close to 100%, the regions determined to be background (definitely not the subject) have a probability close to 0%, and edge regions have probabilities in-between. Typically, said unique probability maps have a resolution of 0.2 to 0.5 pixels depending upon the noise levels in the source images.

Embodiments of the present invention may compare the one or more unique probability maps with an existing database of three dimensional body representations to determine a Closest Match and/or may use an interpolation method to generate new bodies from the existing database to determine a Best Mapping. The existing database is typically populated with empirical data collected from dedicated body scanners. Such representations may include three dimensional body measurements, measurements of the subjects taken manually, and/or scans.

To determine the Closest Match, the test scans can be exhaustively searched, optimising each for pose and camera variations and determining the best result as below, though it may be appreciated that optimisations such as grouping the scans by particular properties (such as gender, weight, age) may be applied to reduce the search space, and an ordering could be applied to enable an approximation of a Binary Search. However, determining the Closest Match cannot offer results which are not already present in the database, and for this reason the Best Mapping is preferred. The following paragraphs will describe the process used in finding a Best Mapping.

Linear interpolation may be used to generate new scans from the existing database, which is advantageous as it allows the use of Linear Optimisation methods to identify the Best Mapping. An average three dimensional body representation and a series of vectors describing variations of the normalised three dimensional body representations within the existing database ("Body Vectors") may be pre-generated via Principle Component Analysis. This stage may include methods to normalise the pose of each subject and/or infer missing data (for example occluded regions of scans from structured-light scanners). The results of this analysis will typically be recorded in advance and retrieved as required, and are referred to as the "Space of Bodies".

Embodiments may also include measurement data in this analysis; producing average measurements and a LOOKUP table denoting how each measurement must be changed as the contribution of each Body Vector is altered (for example, if a Body Vector controls the weight of the scan, it will typically increase the waist circumference as the scan's size increases). These correlations will typically be generated automatically and may have no obvious physical interpretation. To ensure the applicability of Linear methods, the measurements may be pre-processed to produce linearly-related values (for example taking the cube root of any volumetric measurements and decomposing elliptical measurements into major and minor axes and a scaling factor).

In embodiments, only selected body image data is generated, with areas of the subject discounted or discarded. Such exclusions may be undertaken by utilising probability maps that do not include areas of the subject, such as the head area, or by discarding features after segmentation. This can be advantageous as it allows features from the body image data that do not relate to actual body size, such as hair, to be excluded. Excluding these features prevents seemingly arbitrary details such as hair style influencing results.

The Space of Bodies may be searched using a Gradient Descent method, with variables controlling the camera (e.g. position, rotation, field-of-view), pose (e.g. joint angles) and body shape (e.g. the first $\mathbf{N}$ Body Vectors, where typically $\mathbf{N} \simeq 40$) of the scan. This typically involves starting with an average or rough guess for each variable and iteratively refining the value through trial-and-error. The Newton–Raphson method is one possible option, and may be used to minimise the sum of the absolute differences of probabilities for each pixel in a computer-generated silhouette against the calculated unique probability map for the current subject. Mathematically,

$$\operatorname{argmin}\left(\sum_{x \in image} \left| p_{unique}(x) - p_{calc}(x, camera, pose, shape) \right| \right)$$

Where $\mathbf{p_{unique}(x)}$ is the value of the unique probability map at the current pixel and $\mathbf{p_{calc}(x, c, p, s)}$ is 1 if the computer generated silhouette for the current camera **c**, pose **p** and body shape **s** contains the current pixel and 0 otherwise **camera, pose** and **shape** are the arguments to search. It should be noted that the use of the absolute value function is chosen to optimise the Newton–Raphson search, which requires a quadratic function. For other search methods different costs may be appropriate, such as a sum of squares. It may also be beneficial to use a secondary distance function to quickly discard local minima. For example taking the sum of the minimum distances from each of the computer generated silhouette's pixels to a pixel in the unique probability map which has a value above a predetermined threshold c (such as $c = \frac{1}{2}$). Again this is optimised for a Newton-Raphson search. Mathematically,

$$\operatorname{argmin}\left(\sum_{x \in p_{calc}(x,\, camera,\, pose,\, shape)} min\left(\left|y\right| : p_{unique}(x+y) > c\right)\right)$$

Boundaries, such as increasing the cost of the function as arguments vary beyond reasonable limits, are beneficial for preventing unrealistic outcomes. For example including a soft limit as leg separation becomes too wide, with a hard limit after a known impossible separation or separation which does not match the requested pose. Similar principles can be applied to body shape and camera properties.

The body shape is optimised across all views, and the pose and camera properties are per-view. It can be appreciated that in controlled environments multiple pictures can be taken simultaneously, in which case the pose becomes a property which may be optimised across all views as well. Fixed camera positions and properties can further allow the camera variables to be removed from this stage entirely.

In embodiments, this minimisation stage may be achieved by minimising for small numbers of parameters in sequence for improved speed. For example varying the camera and pose variables while locking the body shape variables, then varying the first 5 Body Vectors while locking all others, then varying the first 15 Body Vectors and camera variables, etc. The precise sequence may be determined empirically in advance and/or automatically as a result of analysing the current state (for example giving more time to the optimisation of leg position if it has previously changed significantly).

The end-result of this stage is a unique three dimensional scan for the subject. The same principles apply to finding a Closest Match, but involve iterating through all scans in the database in the place of varying the Body Vectors. In either case the resulting scan (ignoring pose) will now be referred to as the "unique three dimensional scan".

Once the Closest Match or Best Mapping has been obtained, embodiments may calculate measurements using the measurement LOOKUP table and/or measuring paths on the unique three dimensional scan. For a Best Mapping, the determined Body Vector values are used in combination with the LOOKUP table to linearly interpolate the average measurements, producing a unique set of measurements. For a Closest Match, the measurements may be retrieved directly from the LOOKUP table. These measurements are all unscaled, and may be scaled according to several factors including: a known measurement (for example the subject's height); a determined distance between the subject and the camera (for example using a hardware depth sensor or known environment); and/or comparison against a calibrated known object (for example a marker or common object such as a sheet of paper). In the absence of this information, a scale can be approximated by assuming it is similar to that of known subjects with similar body shape (for example a relatively larger head usually implies a smaller scale). This approximation can be calculated using the same LOOKUP methods as for other measurements. At this stage, the measurements will typically be un-normalised; volumes cubed, the circumference of decomposed ellipses recalculated, and so on.

Furthermore, in embodiments the three dimensional body data of the subject can be outputted. The data can be output to one or more of a storage device, online cloud based storage, a mobile phone, a display screen or a personal computing device.

Embodiments of the present invention will now be described, with reference to the following figures, in which:

Figure 1 is a schematic overview of the present invention;

Figure 2 is a flowchart of the present invention relating to data collection;

Figure 3 is a flowchart of the present invention relating to methods used to generate unique probability maps;

Figure 4 is a flowchart of the present invention relating to improving the unique probability map; and

Figure 5 is a flowchart of the present invention relating to generating a three dimensional body image based on the unique probability map and reference three dimensional body representations.

Referring to the figures, system 10 (figure 1) is provided with an input 'dumb' client system 20 and a server unit 30.

The client system 20 comprises a webcam 100 that is used to capture image information from a subject. In use, the webcam 100 is used in conjunction with a screen 120, such as a computer monitor, that displays a representation 110 of the subject. The software controlling the webcam 100, when the webcam is activated, issues a command to the screen 120 such that the screen 120 displays a representation 130 corresponding to a posed position. Typically the posed position required is face-on, with legs and arms spaced apart from the body; the anatomical position. Additionally, side profile positions (along the median plane) may also be requested 225 (figure 2).

When the representation 110 of the subject displayed on the screen aligns with a predetermined outline 220, the subject is deemed to be "in pose" (this may also depend upon remaining stationary for a few seconds). When the subject is ready, and the condition that the subject stand in the requested posed 130 is satisfied, the webcam 100 takes a sequence of images 150, 240 of the subject in the posed position. Once the requisite number of images has been collected, the images are compressed and stored for subsequent upload to the server unit 30, 270.

It can be appreciated that the above steps may be repeated as many times as desired to generate a larger database of posed subject images. Furthermore, as an alternative to a webcam, mobile phone cameras may be used to capture the required images. In such case, the screen 120 and software to run the image capture and issue commands to the subject may be on the phone itself, using a dedicated application. It may also be appreciated that a conventional camera may be used for generating the images as well as other known image capture devices and software. Frames from video sources may also be used.

In addition to the capture of images, the client system 20 can request additional information 140 from the subject to assist the calibration and image identification process. As an example, calibration data to identify relative scales in the two dimensional images may be collected by requesting the height of the subject 200. Other options include asking the subject to hold a calibrated marker to the camera, such as a playing card. Further questions may assist in image processing, such as general amount of and/or colour of worn clothing.

Optionally, the client system 20 may also perform rudimentary checks on the taken and stored images 240. These checks may include analysing the images for background features and requesting that the subject move to a different area if too many background features are detected. Such analysis may also allow a further calibration between the relative height of the subject captured images as, if a common background feature is found, this can be used to ensure all the images of the subject are of a consistent height, to account for variable distance between the subject and the imaging device. Alternatively, or additionally, said analysis can remove common background features (to aid subsequent steps) or may require new images to be obtained if too many background features are detected.

Once images 150 have been captured by the client system 20, they are uploaded to the server unit 30, 270. The server unit 30 may be a subject's computer or smartphone or the like, although typically the server unit 30 is a dedicated off site server. The server comprises a database of known probability maps 160 corresponding to expected locations of the subject based on the posed positions that the subject was instructed to stand in 300, and a collection of segmentation methods 170 which have been determined to be applicable to each pose. Alternatively, the pose of the subject may be determined using a feature tracking technique. A typical method of feature tracking utilises Lucas-Kanade Pyramids to calculate the optical flow of keypoints of the subject between images (or frames if applied to video). By taking a frame where the subject's position is known, key points in the image can be located (for example using corner detection techniques). Such points may be fingers, patterns on clothing, tattoos, skin blemishes etc.). An optical flow technique can then be used to follow these points between frames (or images). The points can be rechecked every few frames to prevent slippage. Such a technique allows the pose of the subject within images to be guessed more accurately and further allows the probability maps to be provided based on the determined pose. Such a pose determination technique increases the flexibility of the system and is particularly useful for determining the pose of subjects from video.

Initially, the server unit 30 analyses the incoming images and pre-processes them 320 to normalise lighting, detect and removing patterns in the images introduced from CCD (Charge coupled device) sensors or the like, such as Active Pixel Sensors (APS) and correct barrel distortion or other known defects. From these filtered images, the server 30 then determines what is likely to be the representation of the subject 110, rather than the background or noise 330-360.

In order to determine the location of the subject within the image, the server 30 applies various image segmentation techniques 330 which have been chosen in advance for the particular pose. These may include manually or automatically tuned methods which produce best results in specific areas of the image, as well as more general methods, and may use both the pre-rendered probability maps 160 and pre-processed source image(s). The segmentation methods may also share results with other segmentation methods (e.g. a hand detector may reduce the search space required by an arm detector). The results of each segmentation method are combined 340, allowing the server to generate a first estimate at the object's (subject) shape. The pre-rendered probability maps 160 can further be used to tune the silhouette identified by aiding in the removal of background objects which may have been incorrectly identified. For example, a colour distribution analysis may identify a colour (e.g. red) as representative of the subject, however if the background contains an identical colour, this will be identified as representative of the subject if based on the colour distribution alone. By combining the colour distribution with the probability map, such background features can be eliminated.

Once a first estimate of the subject's body shape has been determined, and a unique probability map created, the system undertakes a further step to compensate for variable lighting conditions in the source images that affect line sharpness. In order to compensate for this variable line sharpness, the system deliberately blurs the edges present within the probability map 410. This blurring is typically in the order of 1 or 2 pixels (in essence the line is anti-aliased) and may be combined with an upscaling of the probability map for greater accuracy.

By blurring the edges of the generated probability map, the system is aided in determining the direction of pixels in relation to the expected position of the subject 420. As the probability map is a high contrast greyscale image, the system can identify areas (of uncertain pixels) where the change of highest contrast occurs, and assign directionality to that change and those pixels. This direction is normal to the edges of the probability map and identifies the direction of the subject (the white of the probability map) at all pixels within the uncertain (blurred) edge points.

Once the directionality of an uncertain pixel is determined, this can be applied to the original RGB image to find areas (of pixels) with the highest contrast, i.e. where the gradient between neighbouring pixels is greatest 430. This determination allows the boundary of the subject in the RGB image to be determined accurately within a resolution of one pixel. Furthermore, improved, sub-pixel, resolution may be obtained by comparing the boundary pixel to neighbouring pixels to obtain an indication of the true boundary position 440. This may be achieved by determining the colour gradient between the boundary pixel and neighbouring pixel and assigning an estimated true boundary position at the point where the gradient is greatest. An indication of the true estimated boundary position is given by: $\mathbf{b} = 1 - \frac{\mathbf{c_B} - \mathbf{c_A}}{\mathbf{c_C} - \mathbf{c_A}}$, where $\mathbf{b}$ is the true estimated boundary, $\mathbf{c_B}$ is the intensity of the current pixel, $\mathbf{c_A}$ is the intensity of the previous pixel along the vector and $\mathbf{c_C}$ is the intensity of the next pixel along the vector. This step represents an accurate inverse ant-aliasing step for two blocks of solid colour and therefore is especially useful for accounting for anti-aliasing introduced by CCD sensors.

Once the boundary has been determined in the RGB image, this information is then utilised in the probability map to provide an accurate representation of the edge of the generated probability map 450. Noise introduced by the edge identification process can be reduced by blurring the identified edge by ±1 pixel and re-smoothing 460. The end result of this stage is a sequence of accurate, input-specific, probability maps 470, one per source image 150.

With the sequence of accurate, input-specific, probability maps, it is now possible to perform a mapping of these probability maps to pre-calculated average body shapes to attempt to generate the final three-dimensional body shape of the subject. This is achieved by using a database containing a pre-calculated average body shape and extracted principal components "Body Vectors" which are stored on the server as follows. A Body Vector or Eigenbody is similar to the known term Eigenvector or Eigenface and is a representation of how a body may vary (increases in hips leads to corresponding increase in bust etc.). It is calculated in advance using principle component analysis. For example, a particular body image may comprise of an average body shape plus 60% of Body Vector 1, -12% of Body Vector 2 and 18% of Body Vector 3. This allows a mathematical representation of a body scan to be stored without storing the scan itself. Furthermore, the pose of the pre-calculated average body can be selected or calculated to correspond to the pose of the subject as previously determined.

This database of body variations is preloaded by the server 500. When a series of unique probability maps are received 515, the server will perform a search over the Body Space (the possible bodies generated from the average body and Body Vector combinations) to find the Best Mapping 530–575. Initially the Body Vectors are assigned contributions of zero (producing an average body with no variation), and per-pose sensible defaults are assigned to the camera and pose variables (e.g. camera position, arm separation), 530. These defaults may be altered according to features identified by the segmentation methods 330. The corresponding body is rendered for each pose 541, and these renderings are compared against the unique probability maps, 542. The distance function used may be a simple mean-absolute-difference per-pixel, or something more complex, such as a sum of minimum distance between edge pixels. Multiple methods may be used together, for example using minimum distance between edge pixels to determine pose and camera properties, then using mean-absolute-difference to tune the body shape. These methods may benefit from pre-caching of values 520 to improve response times.

By repeatedly varying certain properties (such as a Body Vector's contribution) and re-measuring the distance between rendered scans and the observed unique probability maps 540–575, the Body Vectors, pose variables and camera properties can be iteratively refined. Typically this involves sampling a few variations 550 then fitting a simple mathematical shape (such as a parabola) to the distances to identify a likely minimum-distance. The values at this minimum can then be checked, and the overall minimum from all samples chosen 560, then the search space is reduced and the process repeated. Once the search space is small enough, the process terminates and the final values represent the Best Mapping 580. Typically this optimisation stage will be repeated several times, each time with certain variables "locked" and others "free", for example an initial search may lock all Body Vector contributions, varying only the camera properties and pose variables, and this could be followed by locking the camera while refining the first 5 Body Vectors and rough pose, etc.

The final body shape, pose and camera properties will produce a silhouette which most closely resembles the unique probability maps. It can be appreciated that in systems with known hardware, the camera properties may be known in advance and will not need to be included in this process. Furthermore, in systems with multiple cameras, it may be known that the subject's pose is identical across several images, and a global (rather than per-image) pose may be used.

When the iterative search is complete and a Best Mapping determined, the server can store the three dimensional body data according to its Body Vector contributions. Measurements relating to the three dimensional body shape can be retrieved using a lookup table 580 based upon the average body shape and the Body Vector contributions. Such measurements can include physical measurements such as waist circumference, chest circumference etc., as well as more abstract information, such as gender, age, etc. This lookup table is sourced from known empirical body dimensions generated from precise and accurate full body scanners typically used in hospital environments. The calculated silhouette of best match may then be fed back into the pre-rendered probability maps, potentially increasing the accuracy of the segmentation steps in future uses.

Once a true approximation of the three Dimensional body shape of the subject has been calculated, the body shape and the corresponding measurements can be scaled up to the subject's size by taking into account their entered height or other calibration information 585, though this could also be performed at a later stage by the client 285. Finally, the three dimensional scan and/or measurements can be displayed to the subject 290.

CLAIMS:

1 A method of generating three dimensional body data of a subject; said method comprising the steps of:

i capturing one or more source images of the subject using a digital imaging device;

ii partitioning the one or more source images into a plurality of segments or probability distributions using one or more segmentation method, heuristics and/or predetermined mappings;

iii combining the results of each segmentation method in step ii) to produce one or more unique probability maps representing the subject;

iv comparing the one or more unique probability maps with a database of representations of three dimensional bodies to determine a closest match or best mapping between the or each unique probability map and a representation determined from the database; and

v generating three dimensional body data and/or measurements of the subject based on the best mapping.

2 The method according to claim 1 further comprising, prior to partitioning said one of more images, applying one or more pre-processing stages to the one or more source images to enhance the available data prior to segmentation.

3 The method according to claim 1 or claim 2 wherein combining the results of each segmentation method comprises using mathematical equations and/or heuristics.

4 The method according to any preceding claim wherein comparing the one or more unique probability maps with a database of representations comprises determining a best mapping between each or all of the unique probability maps and a database of existing known three dimensional models by applying a Gradient Descent method to minimise the difference between computer-generated silhouettes of the three dimensional data and the probability maps.

5 The method according to any preceding claim wherein generating three dimensional body data and/or measurements further comprises outputting three dimensional body data of the subject or other personal metrics based on the best mapping from step iv).

6 A method according to any preceding claim , further comprising the step of:

after capturing said one or more images, providing one or more probability map(s) that identify the expected shape and pose of the subject in the one or more images and/or providing a sequence of segmentation methods to apply according to the expected shape and pose of the subject in each image.

7 A method according to claim 6, further comprising the step of:

determining the pose of the subject and providing one or more probability maps based on the determined pose.

8 A method according to any previous claim, wherein comparing the one or more unique probability maps with a database of three dimensional body representations to determine a best mapping comprises the steps of:

igenerating an average three dimensional body image from the database of three dimensional body representations;

igenerating at least one silhouette of the average three dimensional body image; and

icomparing the silhouette with the one or more unique probability maps of the subject and determining the degree of match.

9 A method according to claim 8, further comprising the steps of:

iproviding a predefined series of vectors that define how the average three dimensional body image can be altered;

iapplying the predefined series of vectors with weightings to the average three dimensional body image to generate a modified three dimensional body image;

generating a silhouette of the modified three dimensional body image; and comparing the silhouette with the unique probability map of the subject and determining the degree of match.

10 A method according to claim 9, further comprising the step of:

comparing the degree of match of multiple silhouettes generated by different weightings to determine the best match.

11 A method according to any one of claims 8 to 10, further comprising the step of:

altering the three dimensional body image and/or generated silhouettes to account for imaging factors present within the one or more images of the subject.

12 A method according to claim 11, wherein the imaging factors present are at least one of: camera pitch; camera yaw; camera distance and barrel distortion.

13 A method according to any one of claims 10 to 12, further comprising the steps of:

determining the unscaled measurements of the best match based on a LOOKUP table of known three dimensional body images measurements; and

scaling the best match measurements or scan based on reference data provided by the subject, said reference data being one or more of: subject height; distance between subject and camera; reference to a calibrated known object.

14 A method according to any one of claims 6 to 13, wherein partitioning the one or more images into a plurality of segments comprises the step of:

comparing the location of pixels in the or each image with their corresponding locations on the probability maps to determine the probability that a pixel represents the subject in the image.

15 A method according to any one of claims 6 to 14, wherein partitioning the one or more images into a plurality of segments comprises the step of:

weighting pixels in the or each image according to the corresponding locations on the probability maps to generate a corresponding weighted image, and identifying a colour and/or spatial distribution within these weighted images;

16 A method according to any previous claim, wherein after combining the results of each segmentation method in step ii) to produce one or more unique probability maps representing the subject, the method further comprises the step of:

applying one or more final segmentation method(s) using the combined results of previous segmentation methods as a seed.

17 A method according to any previous claim, wherein after combining the results of each segmentation method in step ii) to produce one or more unique probability maps representing the subject, the method further comprises the step of:

adding a region of uncertainty around the edges of the one or more unique probability maps (for example by blurring) to compensate for variable line sharpness to produce a boundary of uncertain pixels;

identifying the gradient of the edges of the one or more unique probability maps to determine the direction of the subject at each uncertain pixel to produce a corresponding vectored outline probability map; and

applying the vectored outline probability map to the image to determine points with the highest contrast; and

identifying, at said point of highest contrast, a boundary pixel located on the edge of the subject.

18 A method according to claim 17, further comprising the steps of:

comparing the colour of the boundary pixel with the colours of neighbouring pixels to determine from the intensities a true boundary with sub-pixel resolution.

19 A method according to claim 18, wherein the true boundary is calculated using: $\mathbf{b} = 1 - \frac{\mathbf{c_B} - \mathbf{c_A}}{\mathbf{c_C} - \mathbf{c_A}}$, where $\mathbf{b}$ is the true boundary, $\mathbf{c_B}$ is the intensity of the current pixel, $\mathbf{c_A}$ is

the intensity of the previous pixel along the vector and c_C is the intensity of the next pixel along the vector.

20 A method according to any one of claims 15 to 19, wherein a colour distribution is identified within the weighted images and identifying the colour distribution additionally comprises the steps of:

sampling pixels of the image to identify their colour; and

mapping the location of the sampled pixels to the pre-rendered probability map to determine the expected colour of that region of the probability map.

21 A method according to claim 20, wherein identifying the colour distribution further comprises the steps of:

obtaining a probability distribution of the colour of the pixels within all said segments of the image; and

generating a multidimensional (multivariate) probability distribution of the image in intensity, colour and/or spatial space.

22 A method according to claims 20 or 21, further comprising the step of:

comparing the probability distribution with the expected colour of corresponding regions of the probability map to determine the probability that a pixel represents the subject.

23 A method according to any previous claim, further comprising the step of:

converting each of the one or more images into an RGB image for identifying defects in the one or more images.

24 A method according to any previous claim, wherein combining the results of each segmentation method in step ii) to produce one or more unique probability maps representing the subject uses confidence and/or spatial information to weight segmentation methods in areas of disagreement.

25 A method according to any previous claim, further comprising the step of:

outputting the three dimensional body data of the subject.

26 A method according to claim 25, wherein the three dimensional body data is output to one or more of: a storage device; online cloud based storage; a mobile phone; a display screen; or a personal computing device.

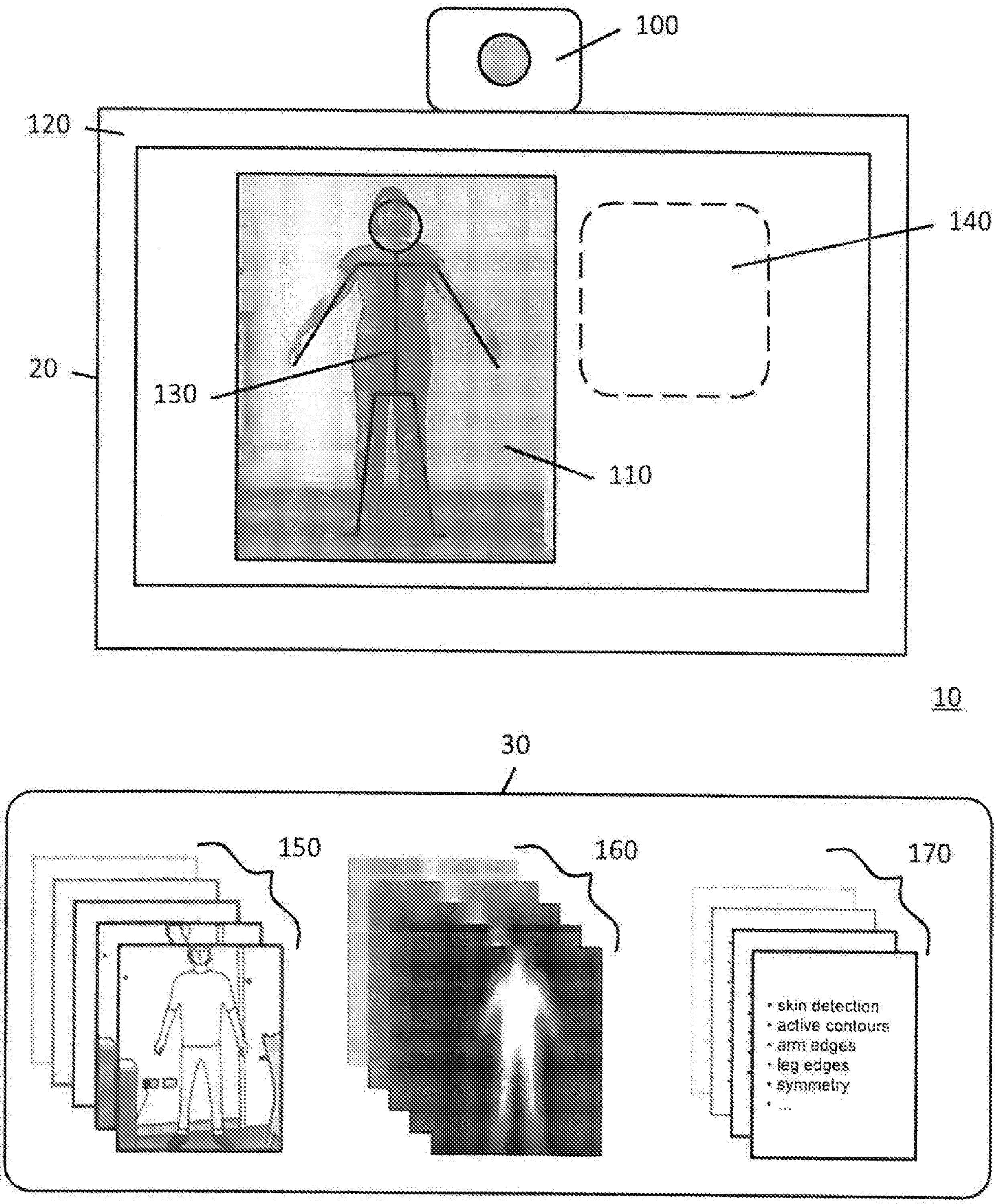
100
120
140
20
130
110
10
30
150
160
170
• skin detection
• active contours
• arm edges
• leg edges
• symmetry
• ...

Fig 1

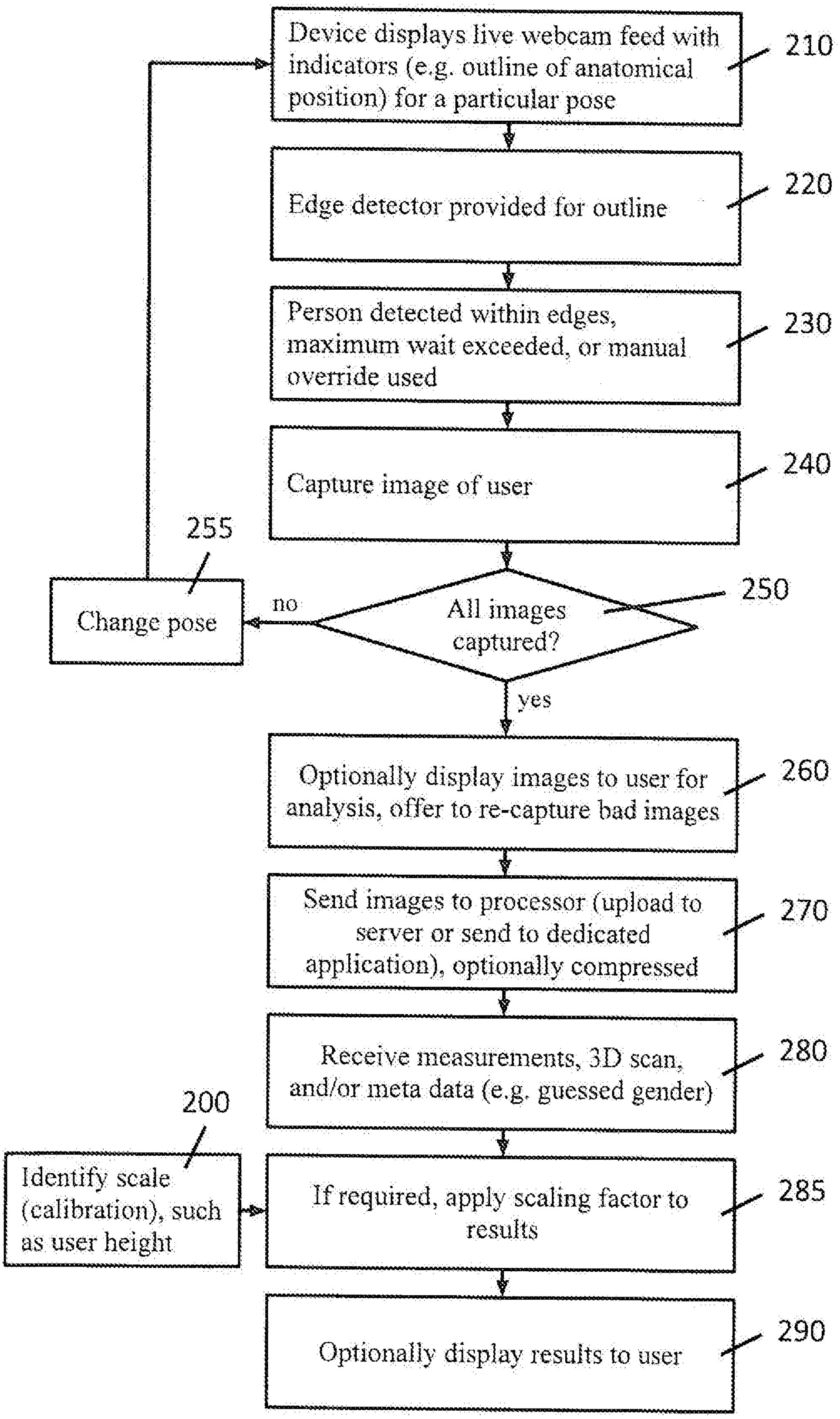
Device displays live webcam feed with indicators (e.g. outline of anatomical position) for a particular pose
210
Edge detector provided for outline
220
Person detected within edges, maximum wait exceeded, or manual override used
230
Capture image of user
240
255
Change pose
no
All images captured?
250
yes
Optionally display images to user for analysis, offer to re-capture bad images
260
Send images to processor (upload to server or send to dedicated application), optionally compressed
270
Receive measurements, 3D scan, and/or meta data (e.g. guessed gender)
280
200
Identify scale (calibration), such as user height
If required, apply scaling factor to results
285
Optionally display results to user
290

Fig 2

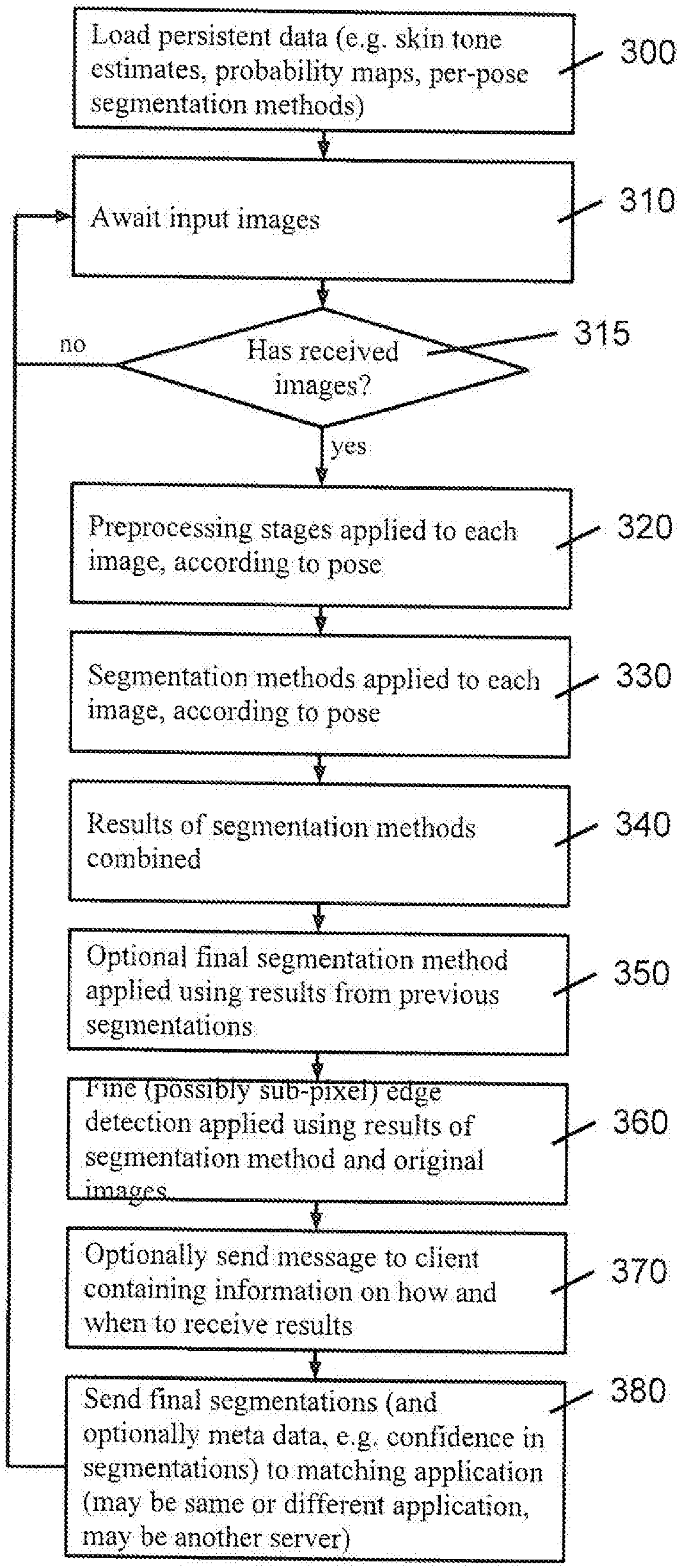
Load persistent data (e.g. skin tone estimates, probability maps, per-pose segmentation methods)
300
Await input images
310
Has received images?
315
no
yes
Preprocessing stages applied to each image, according to pose
320
Segmentation methods applied to each image, according to pose
330
Results of segmentation methods combined
340
Optional final segmentation method applied using results from previous segmentations
350
Fine (possibly sub-pixel) edge detection applied using results of segmentation method and original images
360
Optionally send message to client containing information on how and when to receive results
370
Send final segmentations (and optionally meta data, e.g. confidence in segmentations) to matching application (may be same or different application, may be another server)
380

Fig 3

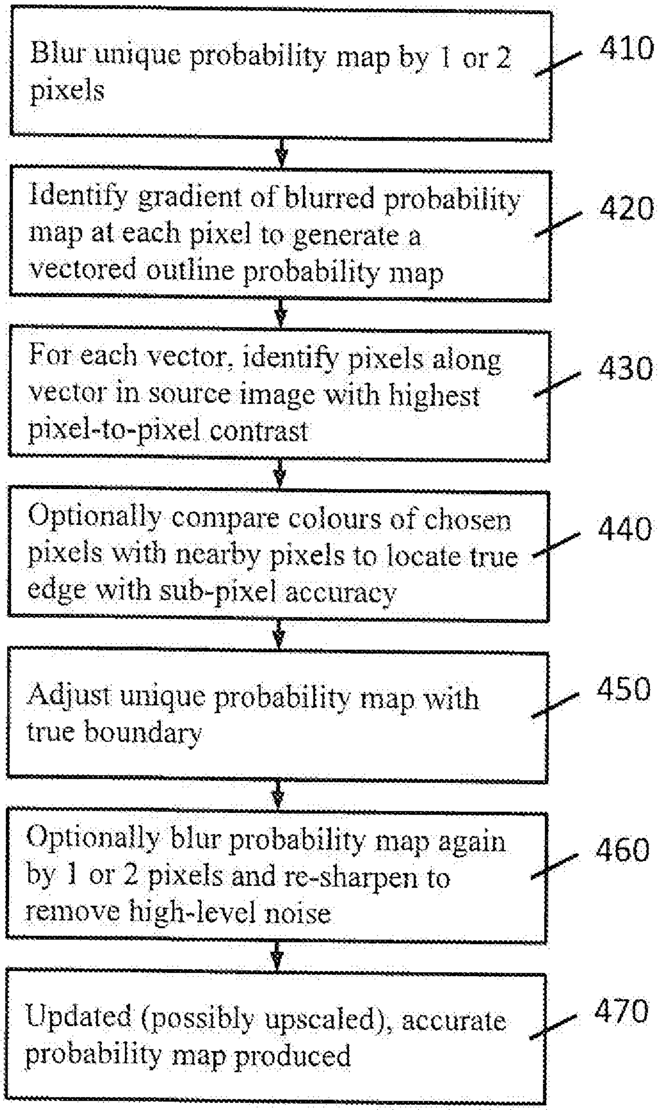
Blur unique probability map by 1 or 2 pixels
410
Identify gradient of blurred probability map at each pixel to generate a vectored outline probability map
420
For each vector, identify pixels along vector in source image with highest pixel-to-pixel contrast
430
Optionally compare colours of chosen pixels with nearby pixels to locate true edge with sub-pixel accuracy
440
Adjust unique probability map with true boundary
450
Optionally blur probability map again by 1 or 2 pixels and re-sharpen to remove high-level noise
460
Updated (possibly upscaled), accurate probability map produced
470

Fig 4

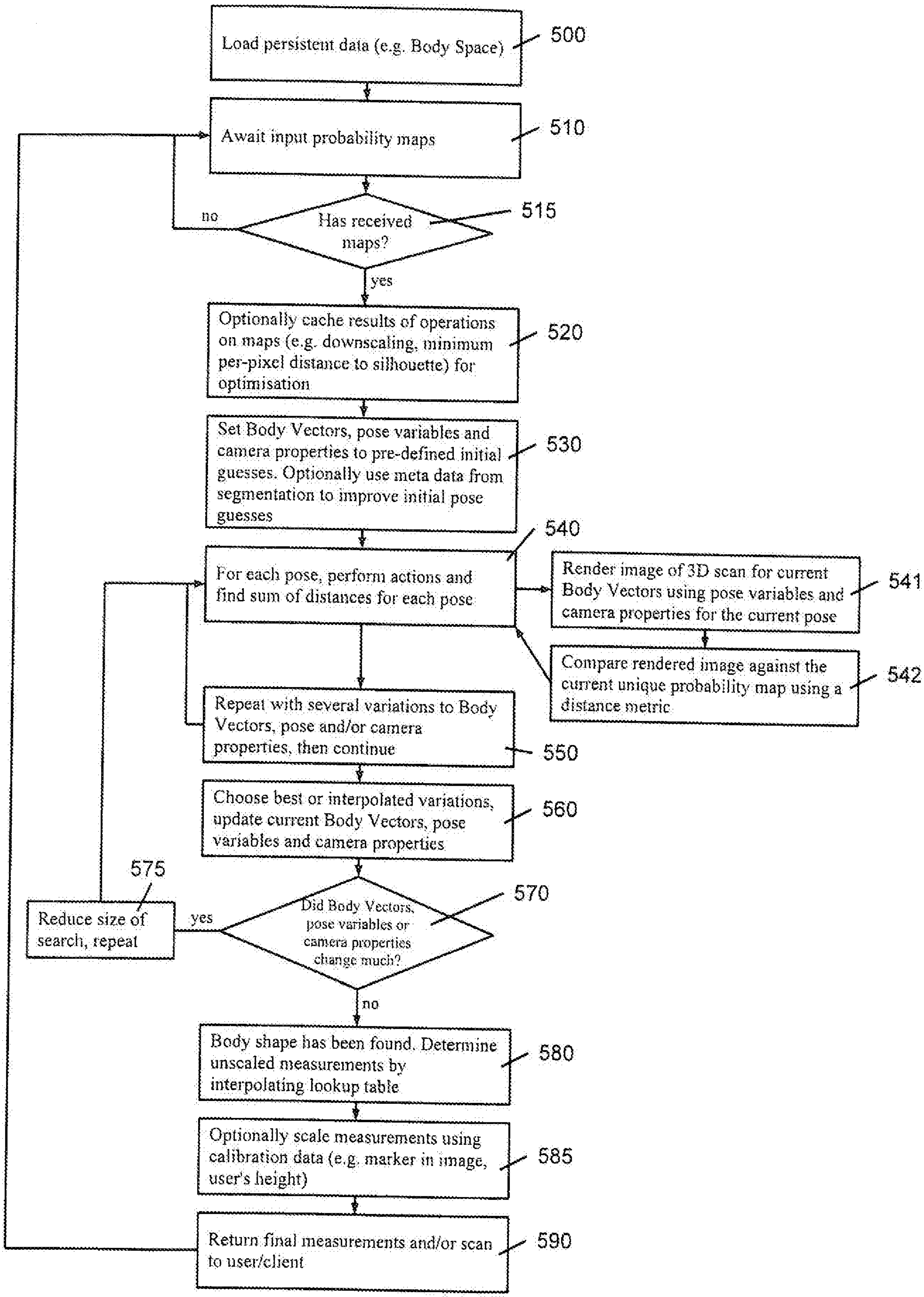
Load persistent data (e.g. Body Space)
500
Await input probability maps
510
Has received maps?
515
no
yes
Optionally cache results of operations on maps (e.g. downscaling, minimum per-pixel distance to silhouette) for optimisation
520
Set Body Vectors, pose variables and camera properties to pre-defined initial guesses. Optionally use meta data from segmentation to improve initial pose guesses
530
540
For each pose, perform actions and find sum of distances for each pose
Render image of 3D scan for current Body Vectors using pose variables and camera properties for the current pose
541
Compare rendered image against the current unique probability map using a distance metric
542
Repeat with several variations to Body Vectors, pose and/or camera properties, then continue
550
Choose best or interpolated variations, update current Body Vectors, pose variables and camera properties
560
575
Reduce size of search, repeat
yes
Did Body Vectors, pose variables or camera properties change much?
570
no
Body shape has been found. Determine unscaled measurements by interpolating lookup table
580
Optionally scale measurements using calibration data (e.g. marker in image, user's height)
585
Return final measurements and/or scan to user/client
590

Fig 5

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2013/051371

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06T7/00
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06T

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, COMPENDEX, INSPEC, IBM-TDB, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Alexandru O Balan: "Detailed Human Shape and Pose from Images", Doctoral Thesis, 1 May 2010 (2010-05-01), XP055069706, Providence, Rhode Island Retrieved from the Internet: URL:http://cs.brown.edu/~black/Papers/BalanThesis2010.pdf [retrieved on 2013-07-04]	1-5, 8-14,16, 24-26
Y	chapters 1, 3, 4	6,7,15, 20-23
A		17-19

	-/--	

[X] Further documents are listed in the continuation of Box C.

[ ] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
11 September 2013	18/09/2013

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Katartzis, Antonios

3

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	MITTAL A ET AL: "Human body pose estimation using silhouette shape analysis", ADVANCED VIDEO AND SIGNAL BASED SURVEILLANCE, 2003. PROCEEDINGS. IEEE CONFERENCE ON 21-22 JULY 2003, PISCATAWAY, NJ, USA,IEEE, 21 July 2003 (2003-07-21), pages 263-270, XP010648393, DOI: 10.1109/AVSS.2003.1217930 ISBN: 978-0-7695-1971-5	6,7,15, 20-23
A	the whole document	1-5, 8-14, 16-19, 24-26

A	HAO JIANG: "Human Pose Estimation Using Consistent Max Covering", TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, IEEE, PISCATAWAY, USA, vol. 33, no. 9, 1 September 2011 (2011-09-01), pages 1911-1918, XP011387229, ISSN: 0162-8828, DOI: 10.1109/TPAMI.2011.92 abstract section 3, equation 1	1-26

3