

Article

# Structured-Light-Based System for Shape Measurement of the Human Body in Motion

Paweł Liberadzki \*, Marcin Adamczyk, Marcin Witkowski and Robert Sitnik 

Institute of Micromechanics and Photonics, Faculty of Mechatronics, Warsaw University of Technology, ul. Św. Andrzeja Boboli 8, 02-525 Warsaw, Poland; m.adamczyk@mchtr.pw.edu.pl (M.A.); m.witkowski@mchtr.pw.edu.pl (M.W.); r.sitnik@mchtr.pw.edu.pl (R.S.)

\* Correspondence: p.liberadzki@mchtr.pw.edu.pl

Received: 20 July 2018; Accepted: 23 August 2018; Published: 27 August 2018



**Abstract:** The existing methods for measuring the shape of the human body in motion are limited in their practical application owing to immaturity, complexity, and/or high price. Therefore, we propose a method based on structured light supported by multispectral separation to achieve multidirectional and parallel acquisition. Single-frame fringe projection is employed in this method for detailed geometry reconstruction. An extended phase unwrapping method adapted for measurement of the human body is also proposed. This method utilizes local fringe parameter information to identify the optimal unwrapping path for reconstruction. Subsequently, we present a prototype 4DBODY system with a working volume of  $2.0 \times 1.5 \times 1.5 \text{ m}^3$ , a measurement uncertainty less than 0.5 mm and an average spatial resolution of 1.0 mm for three-dimensional (3D) points. The system consists of eight directional 3D scanners functioning synchronously with an acquisition frequency of 120 Hz. The efficacy of the proposed system is demonstrated by presenting the measurement results obtained for known geometrical objects moving at various speeds as well actual human movements.

**Keywords:** 4D imaging; whole-body scanning; structured light; single frame

## 1. Introduction

Over the last few years, substantial improvements in three-dimensional (3D) scanning technology have increased the popularity of this method for fast and accurate surface reconstruction. Three-dimensional scanning technology is employed in many science and technology fields for a variety of purposes, including law enforcement work, cultural heritage investigations, medicine and entertainment. In law enforcement, 3D scanners are used for crime scene reconstruction [1,2] and enable further investigation, such as bloodstain pattern analysis [3], to be conducted. In cultural heritage studies, this technology can be used to characterize cultural heritage objects completely, including their shapes and colours [4,5]. In some cases, the reconstruction is performed with the use of additional information from multispectral analysis [6,7]. In the medical field, 3D scanning technology is employed for many different purposes [8], including body surface analysis for anatomical structure detection [9,10] and internal body analysis [11]. Three-dimensional scanners are also utilized in entertainment to capture accurate and realistic 3D textured models for computer graphics [12,13].

In this paper, we focus on the acquisition of comprehensive measurements of a full human body. Currently, the most popular full-field measurement techniques used for 3D human surface acquisition are the laser triangulation (LT) [14,15], time of flight (TOF) [16,17], structured light (SL) [18,19] and structure from motion (SfM) [20,21] methods. Most of the available scanning techniques are only intended to measure the human figure in static poses.

When considering 3D scanning of the human body, integrating the fourth dimension, time, presents an opportunity for improvement. Data captured over an extended period of time could

easily be incorporated to enhance practical applications such as medical diagnostics and computer graphics. The lack of four-dimensional (4D) information, that is, 3D data obtained over an extended period of time, limits the applications of many existing 3D scanning techniques. Although these techniques are adequate for 3D data and each technique can be used to obtain dynamic measurements in specific conditions, the existing approaches all encounter difficulties when capturing 4D information. For example, motion capture (mocap) systems are commonly used to gather similar types of data (i.e., 4D data) but instead of the whole surface geometry, mocap systems only track the positions of certain points [22,23].

Next, we review the existing full-field measurement methods considering the objective of delivering surface information that extends beyond fiducial landmarks. Table 1 presents the advantages and disadvantages of the existing full-field techniques in 4D measurement applications. The SL technique is predominantly based on the projection of a sequence of certain pattern images. Thus, the temporal consistency of an object within a single measurement is required to obtain accurate results. Several reports have suggested that the optimal 4D measurement approach should be based on the projection of a single frame pattern [24,25]. This variant of the SL technique encounters two problems. First, when phase reconstruction relies on data from a single frame, the resulting high error rate diminishes the quality of the generated models. Second, phase unwrapping is impossible without additional information, which creates the problem of appropriate modification of the projected pattern [26,27].

**Table 1.** Comparison of 4D full-field acquisition techniques.

Technique	Advantages	Disadvantages
<b>Structured Light</b>	<ul style="list-style-type: none"> <li>• High capture frequency</li> <li>• High resolution</li> </ul>	<ul style="list-style-type: none"> <li>• Prone to noise</li> </ul>
<b>Structure from Motion</b>	<ul style="list-style-type: none"> <li>• High capture frequency</li> <li>• Not affected by ambient light</li> </ul>	<ul style="list-style-type: none"> <li>• Time-consuming processing</li> <li>• Low resolution in some areas</li> </ul>
<b>Time of Flight</b>	<ul style="list-style-type: none"> <li>• Not affected by ambient light</li> </ul>	<ul style="list-style-type: none"> <li>• Low capture frequency</li> </ul>
<b>Laser Triangulation</b>	<ul style="list-style-type: none"> <li>• Not affected by ambient light</li> </ul>	<ul style="list-style-type: none"> <li>• Low resolution</li> </ul>

Alternatively, SfM algorithms consider multiple images of the different sides of the measured object. Real-time scanners based on this technique must include numerous devices to provide complete sets of synchronously captured images. This approach is encumbered by extremely time-consuming data post-processing, and, depending on the density of unique surface features in the images, the local surface reconstruction quality may fluctuate [28,29].

TOF cameras are commonly used to gather dynamic (4D) data rather than static data. Although TOF cameras are fast, the captured human body surface measurements are often inaccurate [30,31].

LT can be applied to dynamic shape measurements under the condition that multiple laser stripes are used. This approach is problematic because the resulting body surface sampling rate varies with the projected pattern direction [32,33]. For example, if the lines are projected horizontally, then spatial resolution in the vertical direction is much lower than that in the horizontal direction.

In this paper, we present a 4D measurement system (called 4DBODY) that was developed for imaging the human body surface while in motion. The SL method is employed in the proposed system, because this approach provides high-resolution 3D geometry measurements together with relatively high measurement speed and accuracy. We defined three goals that should be achieved to enable the use of measurement data in practical applications. The first goal was full-body measurement, with minimal uncovered areas. The second one was a sampling frequency around 100 Hz when capturing the human body in motion. This value is acceptable for most human movements. We adjusted the sampling

frequency up to 120 Hz in this study due to the technical parameters of the available digital projectors. The third goal was the definition of the output data as a cloud of points in terms of spatial resolution and measurement uncertainty. The spatial resolution was defined as the maximum distance between neighbouring points and the measurement uncertainty was defined as the acceptable error of the location of each point ( $x, y, z$ ). We assumed that the spatial resolution and measurement uncertainty should be equal to or less than 1.0 mm and 0.5 mm, respectively. These values enable rendering of the resultant data without interpolation in standard, full high definition (HD).

The remainder of this paper is organized as follows. Section 2 summarizes the previous studies and outlines their main points and limitations. Section 3 introduces the system concept as well as the system structure and its components and Section 4 describes the developed method and explains the system calibration process. To confirm the reliability of the proposed system, Section 5 discusses the system validation and shows exemplary results. Finally, Section 6 summarizes the study.

## 2. Previous Works

The methods mentioned in Section 1 are used in some existing systems to obtain dynamic shape measurements. For instance, Lenar et al. [27] used a measurement system based on SL called OGX|4DSCANNER for real-time evaluation of lower body kinematics. A single frame pattern and a fringe projection method are employed in this system and the absolute phase distribution is reconstructed based on a single fringe with a known absolute phase value. Two-directional measurements of the lower human body can be obtained using four detectors and two pattern projectors. This system utilizes only two measurement directions, with no overlap of the projected patterns. In their research, Lenar et al. successfully demonstrated the suitability of 4D surface data for musculoskeletal analysis. However, the utilization of only two measurement directions leads to large areas with no data, which is unacceptable for most applications. Imaging only the lower part of the body is a significant limitation as well.

Using 3dMD technology for research on soft-tissue deformations, Pons-Moll et al. [34] created a model for predicting these deformations. Zhang et al. [35] then used that system in their research on estimating body shapes under clothing. The system utilized by Zhang et al. consists of 22 pairs of stereo cameras, 22 colour cameras, 34 speckle projectors and arrays of white-light light-emitting diode (LED) panels. It can capture 3D full-body scans at 60 Hz but the projectors and LEDs flash at 120 Hz to alternate between stereo capture and colour capture. However, the effectiveness of the 3dMD system comes with the trade-off of expensiveness.

In another approach, a Microsoft Kinect v2.0 depth sensor was employed [36]. The acquired 4D data were used for augmented reality to visualize the fit of clothing to the human body in images that followed the movements of customers. As this system is single-directional, this approach cannot be used to capture complete 360° scans. Moreover, the spatial accuracy of the Kinect sensor and its acquisition rate of 30 Hz [37] may be insufficient for dynamic human body measurements.

The biomedical technology manufacturer DIERS International GmbH proposed a system [38] for dynamic spine and posture measurements, which consists of a single detector–projector pair and primarily relies upon SL and LT to perform back reconstruction, in which the position of the spine of the patient is adjusted according to the surface topography generated based on a mathematical model. Gipsman et al. [32] and Betsch et al. [33] used DIERS International GmbH technology in their research to verify the applicability of the system to spine curvature analysis. However, complete 360° scans cannot be captured using this system.

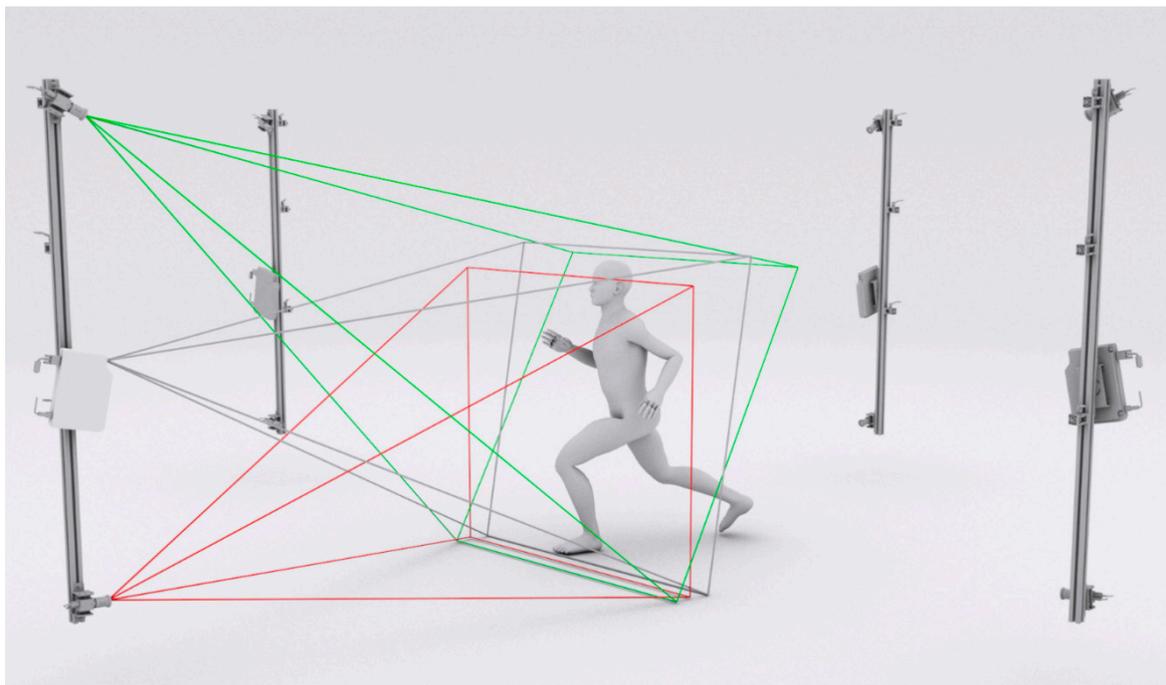
Brahme et al. [39] presented a system for 4D patient pose estimation during diagnostic and therapeutic procedures. This system is based on LT and involves the use of multiple laser stripes with a measured volume of approximately  $40 \times 40 \times 20 \text{ cm}^3$ . Although multiple laser stripes can be employed to capture simultaneous scans of the entire measurement area, the spatial resolutions of these data are uneven because, while high resolution can be achieved in the direction of the laser stripes, the areas between these stripes exhibit low resolutions.

Collet et al. [40] proposed an integrated approach that combines three different techniques for 4D reconstruction: shape from silhouette, SfM and infrared SL. The final model is represented by an animated 3D mesh textured by its natural red, green and blue (RGB) colour. This solution is complete and mature from technical and practical perspectives but it is very expensive and requires the use of a green-box.

These systems have numerous limitations such as small measurement volume, limited number of measurement directions, low acquisition frequency, or high cost. In this paper, we address these limitations with an approach extended to 4D body scanning, which is partially based on the method previously proposed by Lenar et al. [27].

### 3. Acquisition System Design

We decided to use four directional measurement heads (DMHs) evenly distributed around the measurement volume, as a compromise between the completeness of the final model and the equipment cost. This setup provides almost complete reconstruction, excluding some areas around the armpits and groin. Increasing the number of DMHs would not significantly affect the areas in these regions that are not measured but would increase the equipment cost. The resultant measurement volume is  $2.0 \times 1.5 \times 1.5 \text{ m}^3$ . To reconstruct as much of the human body surface area properly as possible without skipping portions such as the shoulders or perineum region, two detectors and one projector are used in each DMH, as illustrated in Figure 1.



**Figure 1.** Architecture of the proposed acquisition system.

The detectors are Grasshopper 3.0 cameras with 2.3 megapixel Sony sensors, 163 Hz capture frequencies in free-run mode and 120 Hz capture frequencies in synchronous mode [41]. Each DMH includes one Casio XJ-A242 projector (1280 × 800 pixels, 2500 ANSI LUMENS) with two light sources, an LED for the R channel and a laser for the B and G channels [42]. To avoid crosstalk between the cameras and projectors of the neighbouring DMHs, spectral separation is applied. Each DMH uses one R, G, or B channel for projection. The corresponding spectral filters are mounted on the camera lenses. Figure 2 depicts the two personal computers (PCs) and the single custom-designed hardware synchronization unit (HSU) that controls the proposed system. Each PC is responsible

for the management and acquisition of two DMHs. The HSU is responsible for the synchronous projection-acquisition of all of the DMHs. This task is realized by a wired synchronization connection between the projector, cameras and HSU. A photograph of one side of the developed measurement system is presented in Figure 3. Aside from the projectors, the room with the system does not contain any other light sources. We used additional blackout curtains to provide the highest possible fringes modulation and remove influence of external lighting. The curtains are outside of the measurement volume; thus, they do not appear as part of the measurements.

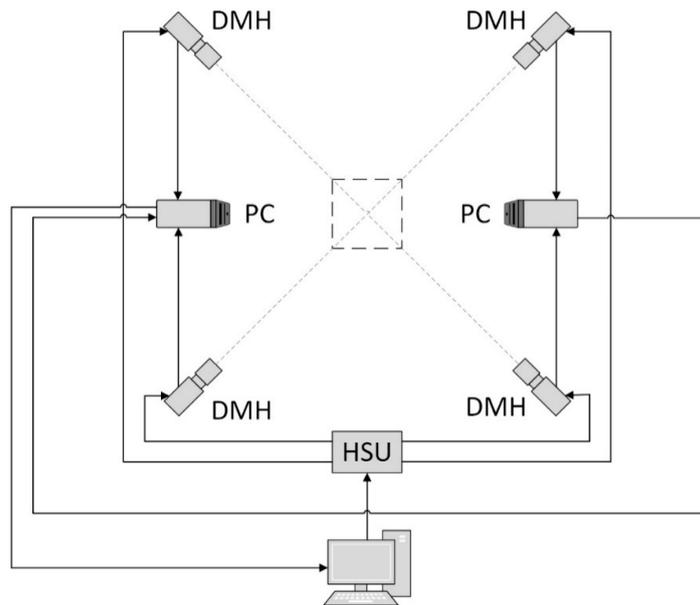


Figure 2. Communication model used in the proposed system.

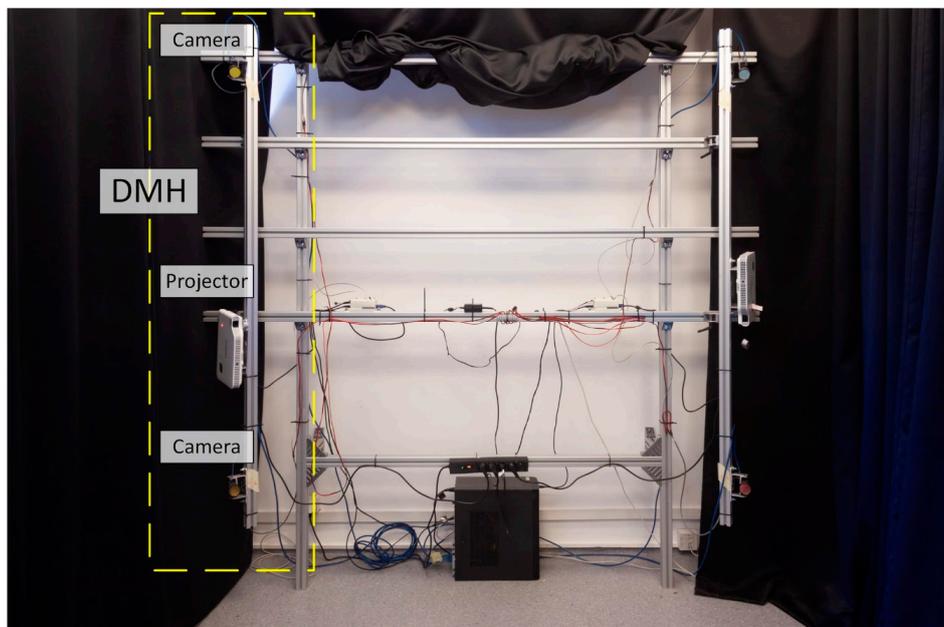


Figure 3. Photograph of the developed 4DBODY system (two directional measurement heads (DMHs) are visible). A single DMH with its constituent devices is enclosed within the dashed rectangle.

To summarize, the 4DBODY system was designed to realize synchronous projection and acquisition of images with a frequency of 120 Hz, which is twice as high as the frequencies of the

systems presented in Section 2. Compared to the other systems that can provide full human body measurements, this system can capture a similar amount of data. However, due to the use of only four measurement directions, the cost of the equipment in our system is significantly lower than it is for the other described systems. A comparison of the proposed acquisition system with the previous systems is presented in Table 2.

**Table 2.** Comparison of the 4DBODY system with the previously proposed systems.

	Amount of Data	Acquisition Frequency	Equipment Cost	Marker-Less System
<b>4DBODY</b>	+	+	+	+
<b>3dMD system</b>	+	+/-	-	+
<b>DIERS International GmbH system</b>	-	+/-	+/-	+
<b>System proposed by Collet et al.</b>	+	+/-	-	+
<b>Microsoft Kinect 2.0</b>	-	-	+	+
<b>VICON</b>	-	+	+/-	-

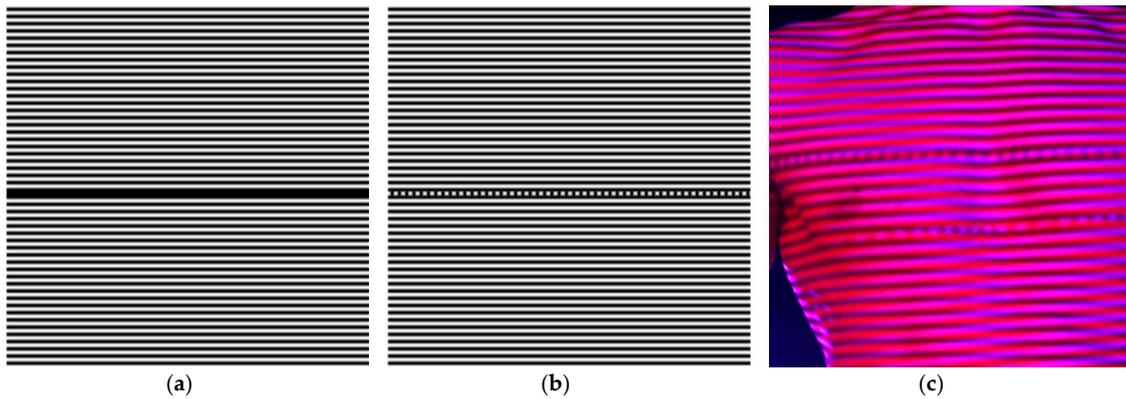
#### 4. Measurement Process

The entire 4DBODY system must be calibrated before measurement. The same single-frame processing method is used in the calibration and measurement procedures. In the following subsections, we will describe the single-frame processing method followed by the calibration of the whole system.

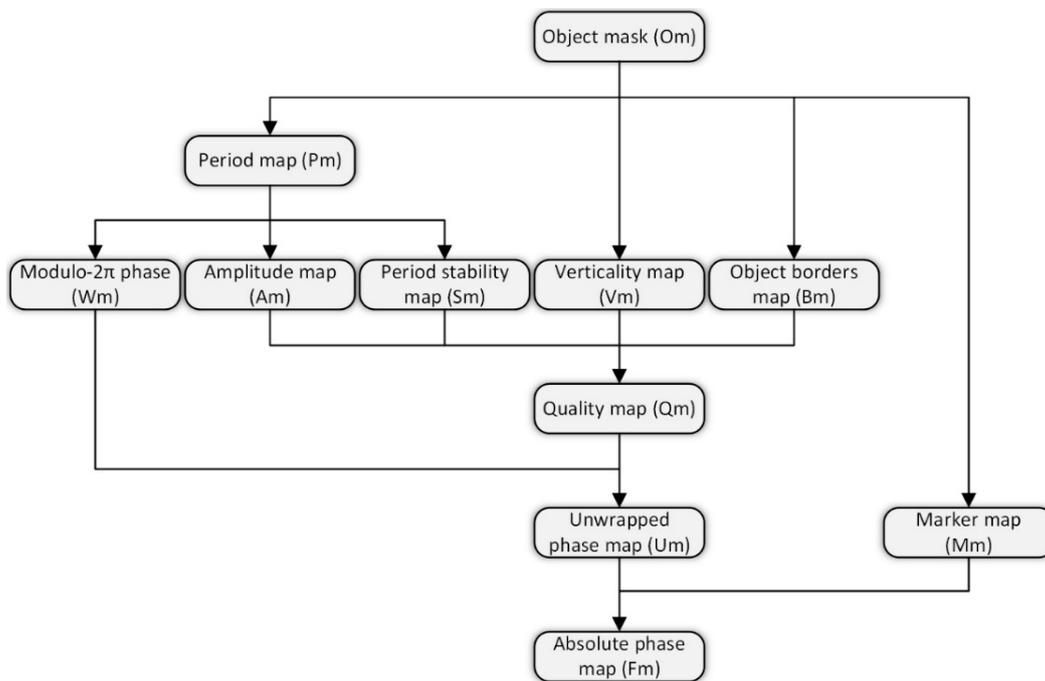
##### 4.1. Single-Frame Processing

To avoid motion artefacts, a single frame method is used in the 4DBODY system for shape measurement. For this purpose, we modified the single sine pattern method proposed by Sitnik [43]. The modification involves using a different design for a single distinguished fringe, which is modulated transversely for fringe orientation in the proposed technique, as displayed in Figure 4. This distinguished fringe is called the marker and is used to determine the absolute phase value. During image analysis, the marker is localized in the image space by performing one-dimensional (1D) fast Fourier transform (FFT) frequency filtering [44]. The spatial-carrier phase-shifting (SCPS) method proposed by Larkin [45] is employed to calculate a modulo- $2\pi$  phase. The seven-point method was selected as a compromise between phase quality and minimal analysed fringe period. Furthermore, the selected SCPS method is resistant to inaccurate intensity sampling within the area of a single fringe period.

One of the most difficult aspects of the SCPS method is determining a phase unwrapping technique that is reliable, especially in areas in which the local shape derivative is discontinuous. When applying phase unwrapping to the human body, problems usually arise in several regions, specifically, the cervical (neck), mammary (breast), axillary (armpit), antecubital (elbow), inguinal (groin) and popliteal (knee pit) regions. Thus, the proposed method is focused on identifying areas to perform phase unwrapping reliably and accurately. This objective is realized by creating a quality map ( $Q_m$ ) enabling proper phase unwrapping, based on a reliable spanning tree [46]. The general processing path of a single frame is presented in Figure 5. A brief explanation of Figure 5 is given next, followed by a more detailed description of the algorithms used in the proposed approach.



**Figure 4.** Projected patterns: (a) original Sitnik pattern from [43]; (b) modified pattern presented in this paper; and (c) two spectrally separated patterns (red and blue) projected onto a human body surface.



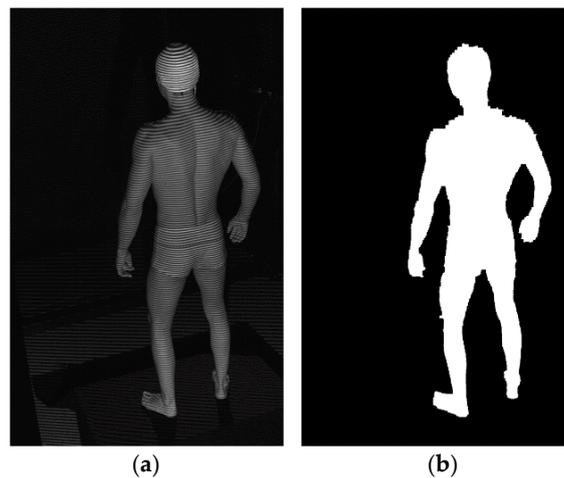
**Figure 5.** Single-frame processing scheme.

The first processing step involves separating the analysed surface from the background using an object mask (Om). The Om limits the subsequent calculations to only the object area, eliminating erroneous off-object areas and dramatically increasing the processing speed. In the next step, the fringe period per pixel map (Pm) and modulo- $2\pi$  phase map (Wm) are calculated. To achieve the final Qm values, the following two-dimensional maps are calculated:

- Fringe amplitude map (Am)—favours areas with high fringe contrast, eliminating errors due to incorrect fringe period estimation;
- Period stability map (Sm)—favours areas with stable fringe periods, avoiding areas with local period discontinuities;
- Fringe verticality map (Vm)—favours areas consisting of fringes with locally constant orientations, according to the projected orientation, thus avoiding high curvature areas;
- Border areas map (Bm)—favours areas with the greatest distances to the edges of the object, thus eliminating errors due to surface discontinuities.

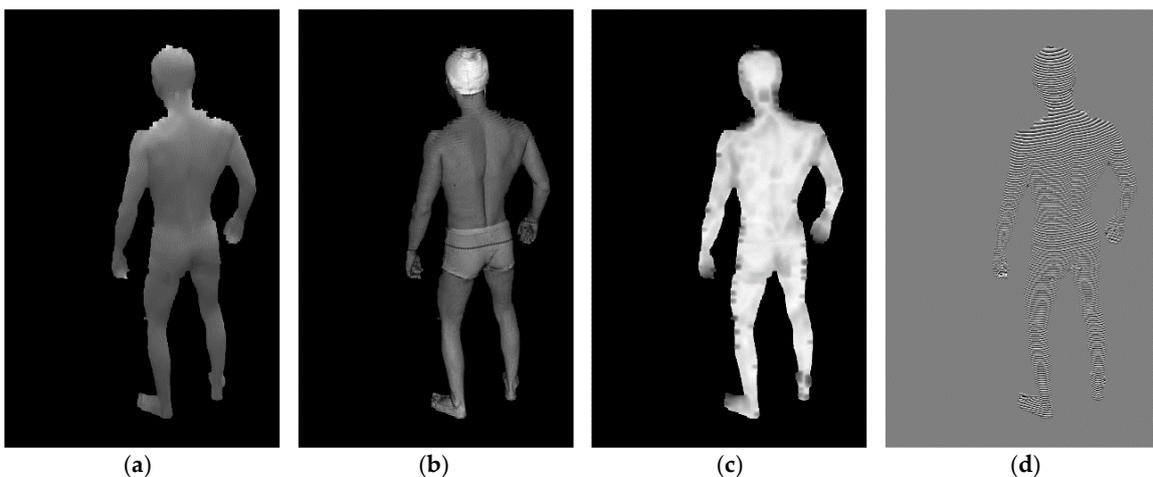
The  $Q_m$  calculation method is heuristic and is adjusted to achieve proper human body measurements. The calculated spanning tree is applied to the  $W_m$ , with branch weights based on the  $Q_m$ , to derive the  $U_m$ . Then, the  $U_m$  and  $M_m$ , which provide information about the marker location, are used to generate the final absolute phase map ( $F_m$ ).

As shown in Figure 6, the  $O_m$  calculation begins with the masking of overexposed pixels and Otsu thresholding [47]. Next, dilation and erosion are applied to produce smooth contours. Finally, the largest segment in the image is selected and everything else is masked.



**Figure 6.** Example images depicting the  $O_m$  calculation procedure: (a) input image; (b) output image.

The corresponding  $P_m$  and  $A_m$  are calculated for each object pixel in the  $O_m$  and are derived from the median, maximum and minimum values of the local intensity in the neighbourhood. To calculate the  $P_m$ , such as that shown in Figure 7a, the median intensity is used for the thresholding of the local intensities as well as for counting the period values in the directions perpendicular to the fringes. The  $A_m$  values in Figure 7b represent the differences between the local maximum and minimum intensities. Next, the  $S_m$ , such as that presented in Figure 7c, is calculated as the variance of the  $P_m$  in the direction of estimation. The  $W_m$ , such as that depicted in Figure 7d, is calculated for each object pixel in the  $O_m$  using the  $P_m$  values to select samples based on the local fringe period. Linear interpolation is used for non-integer coordinate sampling.



**Figure 7.** Example images depicting certain maps calculated in this study: (a)  $P_m$ ; (b)  $A_m$ ; (c)  $S_m$ ; (d)  $W_m$ .

Next, the  $V_m$ , such as that in Figure 8a, is derived based on the calculated intensity gradients in both the vertical and horizontal directions. The quotient of the horizontal gradient over the vertical gradient can be taken as a measure of fringe verticality, as detailed in Equations (1)–(3). The window size in pixel  $(r, c)$  is equal to local fringe period taken from the  $P_m$ .

$$grad_V(r, c) = \sum_{i=r+1}^{r+w} |I(i, c) - I(i-1, c)| + \sum_{i=r-w}^{r-1} |I(i, c) - I(i+1, c)| \quad (1)$$

$$grad_H(r, c) = \sum_{j=c+1}^{c+w} |I(r, j) - I(r, j-1)| + \sum_{j=c-w}^{c-1} |I(r, j) - I(r, j+1)| \quad (2)$$

$$grad_H(r, c) = \sum_{j=c+1}^{c+w} |I(r, j) - I(r, j-1)| + \sum_{j=c-w}^{c-1} |I(r, j) - I(r, j+1)| \quad (3)$$

where

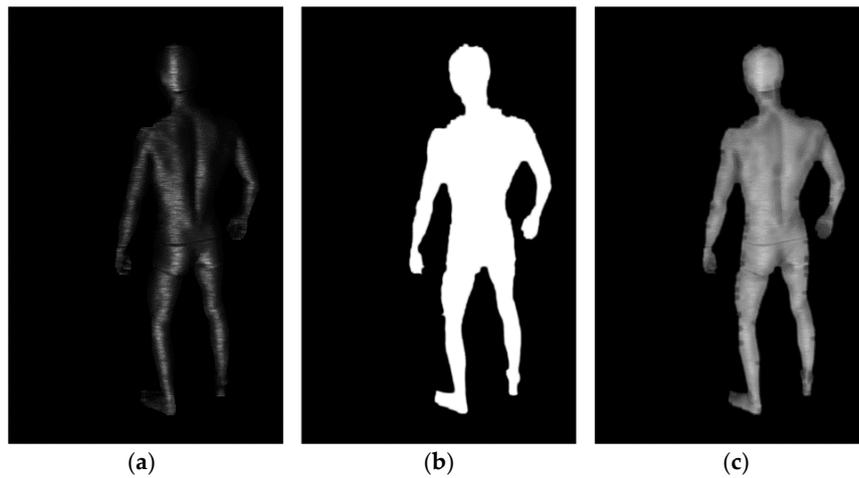
- $r$  and  $c$ : row and column number, respectively, of the central pixel;
- $i$  and  $j$ : row and column number, respectively, of the current pixel;
- $w$ : window size;
- $I(r, c)$ : intensity of pixel  $(r, c)$  in the image;
- $grad_V(r, c)$  and  $grad_H(r, c)$ : vertical and horizontal gradients, respectively, of pixel  $(r, c)$ ;
- $V_m(r, c)$ : verticality of pixel  $(r, c)$  in the  $V_m$ .

As shown in Figure 8b, the  $B_m$  calculation simply blurs the  $O_m$ . The  $B_m$  is used to reduce the weights of the pixels near the border between object and background, which is a highly erroneous area. Together with the  $A_m$  and  $S_m$ , the  $V_m$  and  $B_m$  are normalized and used to construct the  $Q_m$ , such as that depicted in Figure 8c. Equation (4) is used to evaluate the quality as the weighted arithmetic mean of the powers of the pixel values from the four contributing maps. The weights that we established experimentally, enabling us to achieve reliable results, were as follows:  $b_w = 1$ ,  $a_w = 5$ ,  $v_w = 3$ ,  $s_w = 1$ ,  $b_e = 1$ ,  $a_e = 2$ ,  $v_e = 1$ ,  $s_e = 1$ , where these variables are as defined after Equation (4). According to our experience, the  $A_m$  map has the greatest influence on the proper unwrapping procedure. Additional power weights enable the gradients of values in a particular map to be increased, leading to less error-prone unwrapping of the human body geometry. The weight values were established in this study for human body measurement and should be adopted for other subjects.

$$Q_m(r, c) = \frac{b_w * B_m(r, c)^{b_e} + a_w * A_m(r, c)^{a_e} + v_w * V_m(r, c)^{v_e} + s_w * S_m(r, c)^{s_e}}{b_w + a_w + v_w + s_w} \quad (4)$$

where

- $r$  and  $c$ : row and column, respectively, of a pixel;
- $b_w$ ,  $a_w$ ,  $v_w$  and  $s_w$ : weights of the  $B_m$ ,  $A_m$ ,  $V_m$  and  $S_m$  components, respectively;
- $b_e$ ,  $a_e$ ,  $v_e$  and  $s_e$ : exponents of the  $B_m$ ,  $A_m$ ,  $V_m$  and  $S_m$  components, respectively;
- $B_m(r, c)$ ,  $A_m(r, c)$ ,  $V_m(r, c)$  and  $S_m(r, c)$ : pixel values in the  $B_m$ ,  $A_m$ ,  $V_m$  and  $S_m$ , respectively;
- $Q_m(r, c)$ : quality value of pixel  $(r, c)$ .



**Figure 8.** Example images showing the (a)  $V_m$  and (b)  $B_m$  calculation results and (c)  $Q_m$  calculated based on the  $V_m$  and  $B_m$ , together with the  $A_m$  and  $S_m$ .

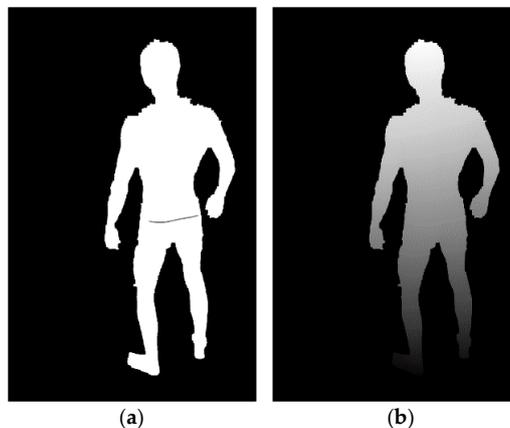
The  $U_m$  calculation begins by selecting a random object pixel from the top 10% of the  $Q_m$  as the initial point for the minimum spanning tree algorithm [46]. Therefore, the calculated  $U_m$  must be shifted by a certain value to obtain the absolute phase distribution. The  $M_m$ , which contains information about the distinguished fringe location, is used to determine the value of the shift and thus to construct the final output, the  $F_m$ . An example  $F_m$  obtained in this study is depicted in Figure 9b. The  $M_m$  calculation begins by performing one-dimensional (1D) FFT [44] to filter out frequencies with orientations similar to those of the calculated fringes. Subsequently, after conducting an adequate inverse FFT [44], thresholding and segmentation are performed to identify the largest segment representing the marker, as shown in Figure 9a. Marker pixels are used to calculate the marker phase value, thereby providing the marker fringe number. Then, the median of the phase values under the marker pixels is obtained and the necessary phase shift is calculated by applying Equation (5):

$$\Phi_x = 2\pi * \left[ N - \text{floor} \left( \frac{\text{median} + \pi}{2\pi} \right) \right], \quad (5)$$

where

$N$ : projected marker index;

$\Phi_x$ : phase shift.



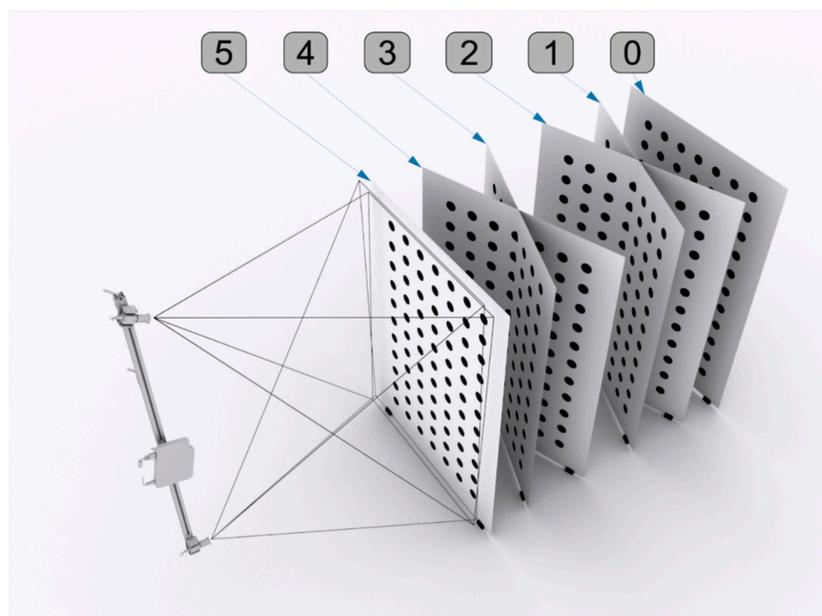
**Figure 9.** Visualization of the last steps of single-frame processing: (a)  $M_m$ ; (b)  $F_m$ .

Later, the  $F_m$  is scaled to a point cloud in real-world coordinates  $(x, y, z)$  based on phase/geometry calibration data from the DMHs. The next section describes the calibration procedure for a single DMH and the global relative calibration of all of the measurement modules.

#### 4.2. Calibration Procedure

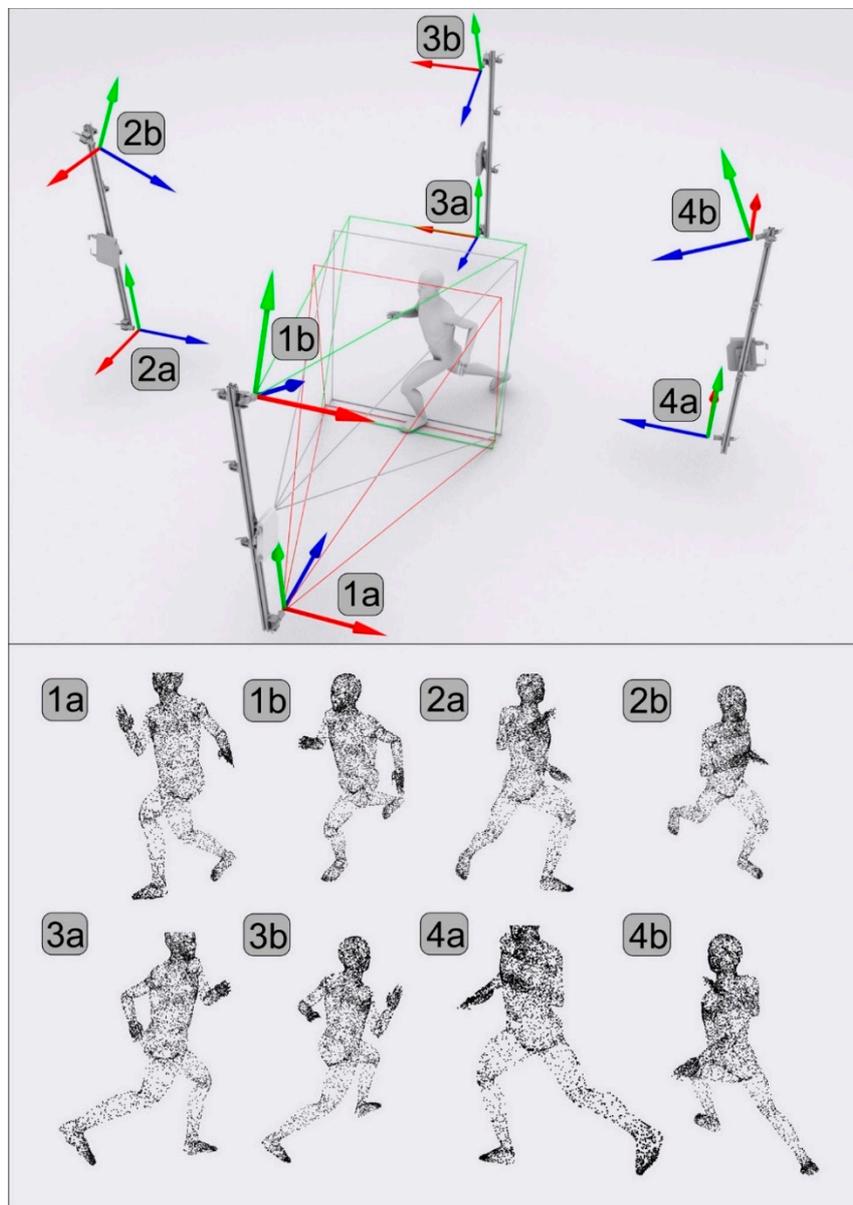
The procedure for calibrating the proposed 4DBODY system has three stages. In the first stage, both detectors of each DMH are calibrated together, with each DMH being calibrated independently. In the next stage, phase calibration is conducted for each projector–detector set that constitutes a single DMH. The camera and phase calibration of the DMHs enables every DMH to collect measurements independently. The final calibration stage, called global calibration, involves calculating transformations between individual module measurement volumes, enabling  $(x, y, z)$  data to be received in the common, global coordinate system.

The first two calibration stages pertain to the local calibration of individual DHMs, as conveyed in Figure 10. These two stages are executed according to the procedure described by Sitnik [48]. A dedicated calibration artefact is employed, which is validated using a coordinate measuring machine (CMM) and has the form of a white board with black, circular markers aligned in rows and columns. The dimensions of this artefact are  $2.0 \times 1.5 \text{ m}^2$ . During camera calibration, a line in the 3D coordinate system is assigned to each detector pixel. Subsequently, the real  $x$  and  $y$  coordinates are determined for each pixel. The phase calibration is based on triangulation with the distinction that, in the proposed method, the projector acts as the second device in the triangulation pair. This alteration produces a phase-to-depth distribution for each detector pixel. This distribution, together with the 3D lines from camera calibration, constitutes a mapping from pixel coordinates and pixel absolute phase values to a specific 3D point in the real coordinate system. Thus, the  $F_m$  calculated from a single image is processed into a 3D reconstruction of a measured object. The output data are presented in the form of a cloud of points. In addition to the  $(x, y, z)$  coordinates, the output contains two more data buffers in the individual points: intensities and normal vectors. The normal vectors are calculated based on the  $(x, y, z)$  coordinates of the neighbouring detector pixels for each point in the cloud. This form of the output data was selected owing to its suitability for further processing and conversion into other forms, such as that used in the triangle mesh model.



**Figure 10.** Single DMH calibration process: calibration artefact positions 0–5 used during camera calibration.

Two directional point clouds generated by a single DMH are spatially aligned with each other. However, as the clouds generated by different DMHs are defined by different, local coordinate systems, global calibration is performed to realize proper, integrated multiview measurement. The same calibration artefact employed for local calibration is used for global calibration. The calibration artefact is placed in the measurement volume in a position that is visible to the detectors of two neighbouring DMHs. The calibration pattern positions expressed in the local coordinate system of each DMH are analysed and used to estimate the mutual transformations between the two devices. This process is executed on each possible pair of DMHs, thereby capturing all of the relative transformations. Thus, the four sets of individual point clouds, as exhibited in Figure 11, can be merged into a single multidirectional cloud by applying the calculated transformations only. The final measurement results are represented by all of the points from the directional clouds of points with the corresponding 3D transformations.



**Figure 11.** Integration of the four DMH coordinate systems using the camera coordinate systems depicted in 1a–4b and respective point clouds. Note that the number of points was reduced to improve the clarity of the figure.

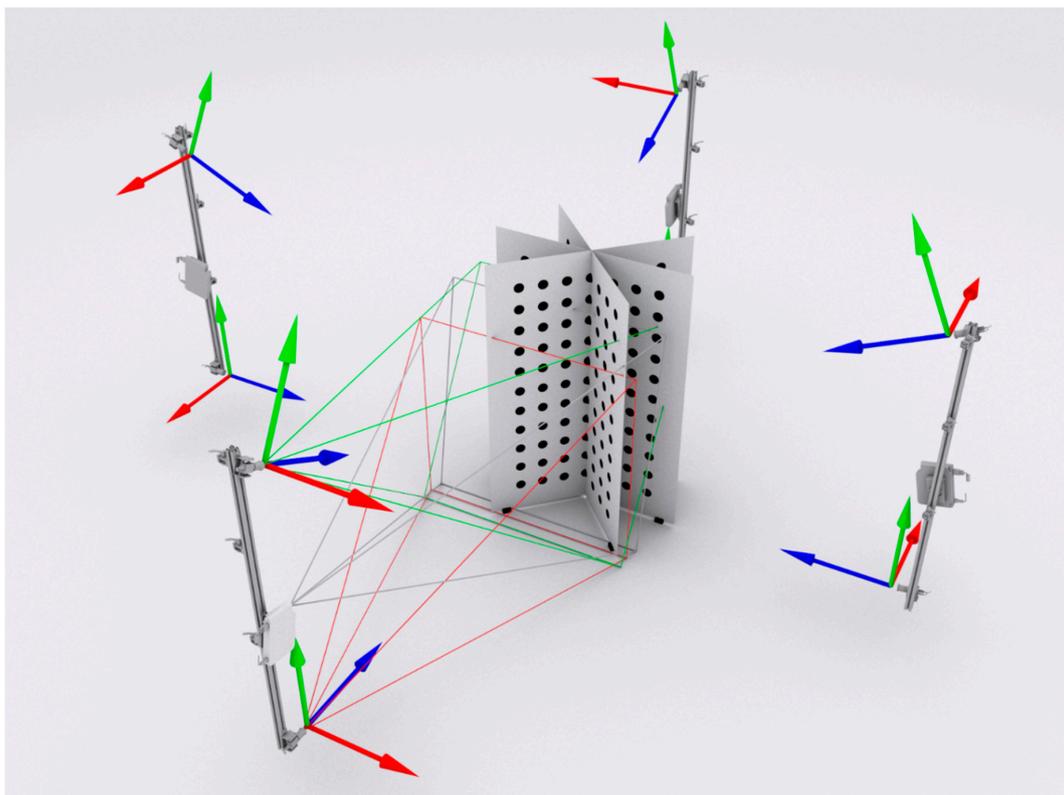
## 5. Validation of the Proposed System

### 5.1. Initial Validation

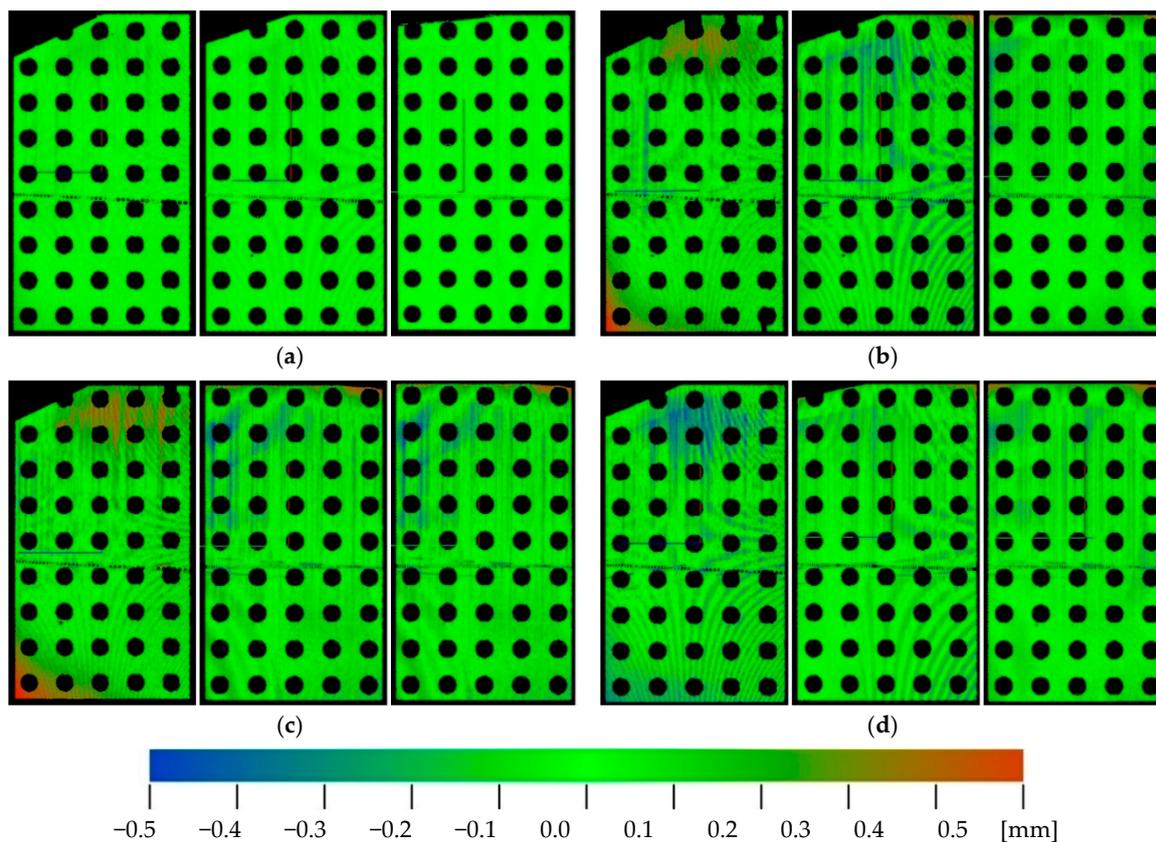
After completing the calibration process, the accuracy of the overall 4DBODY system was evaluated. Generally, SL scanners are validated according to the recommendations in VDI/VDE 2617-6 [49]. Considering the dynamic aspects of the measurements performed using the proposed system, we developed a simplified and adapted approach based on the measurement of an object with a known geometry. The calibration artefact was also used in this initial validation. The artefact was attached to a turntable, which allowed us to examine both static and dynamic cases using stable and rotating models, respectively. We tested model rotation speeds of 4.0 rpm, that is, approximately 0.20 m/s on the side of the model used; 9.0 rpm, or approximately 0.45 m/s; and 14.0 rpm, or approximately 0.70 m/s. For each case, validation was executed separately for each DMH. The validation process consisted of two error analysis steps.

- Step 1: A virtual plane was fit to the received cloud of points, that is, the captured model.
- Step 2: The distances between the outermost marker centres on both model diagonals were measured. This distance was determined as the mean of the distances between the points on the outer/inner edges of the outermost markers.

For the validation, we used data originating from two frames in which the calibration artefact was oriented parallel to the diagonals of the measurement volume and one frame in which the calibration artefact was perpendicular to the DMH, as depicted in Figure 12. Such placement provides reliable estimates because, in most cases, extreme positions produce the highest measurement errors. Examples of the root mean square (RMS) errors obtained from plane fitting are presented in Figure 13. Table 3 lists the errors calculated for the analysed dynamic scenarios.



**Figure 12.** Calibration artefact positions employed in the validation analysis.



**Figure 13.** Examples of RMS errors obtained by plane fitting for a single DMH, for stable and rotating calibration artefact: (a) 0.0 rpm; (b) 4.0 rpm; (c) 9.0 rpm; (d) 14.0 rpm. In each panel, the outermost left and right images correspond to when the calibration artefact was oriented diagonally in the measurement volume. The middle image was taken when the calibration artefact was oriented perpendicular to the DMH. A portion of the upper left corner of the calibration artefact is missing from some of the images because part of the artefact was outside the calibrated volume.

**Table 3.** Root mean square (RMS) errors for the analysed dynamic scenarios.

	Rotation Speed [rpm]			
	0.0	4.0	9.0	14.0
Average RMS error of plane fitting [mm]	0.17	0.22	0.25	0.23
Average RMS error of the distance between the outermost marker centres [mm]	0.21	0.27	0.23	0.23

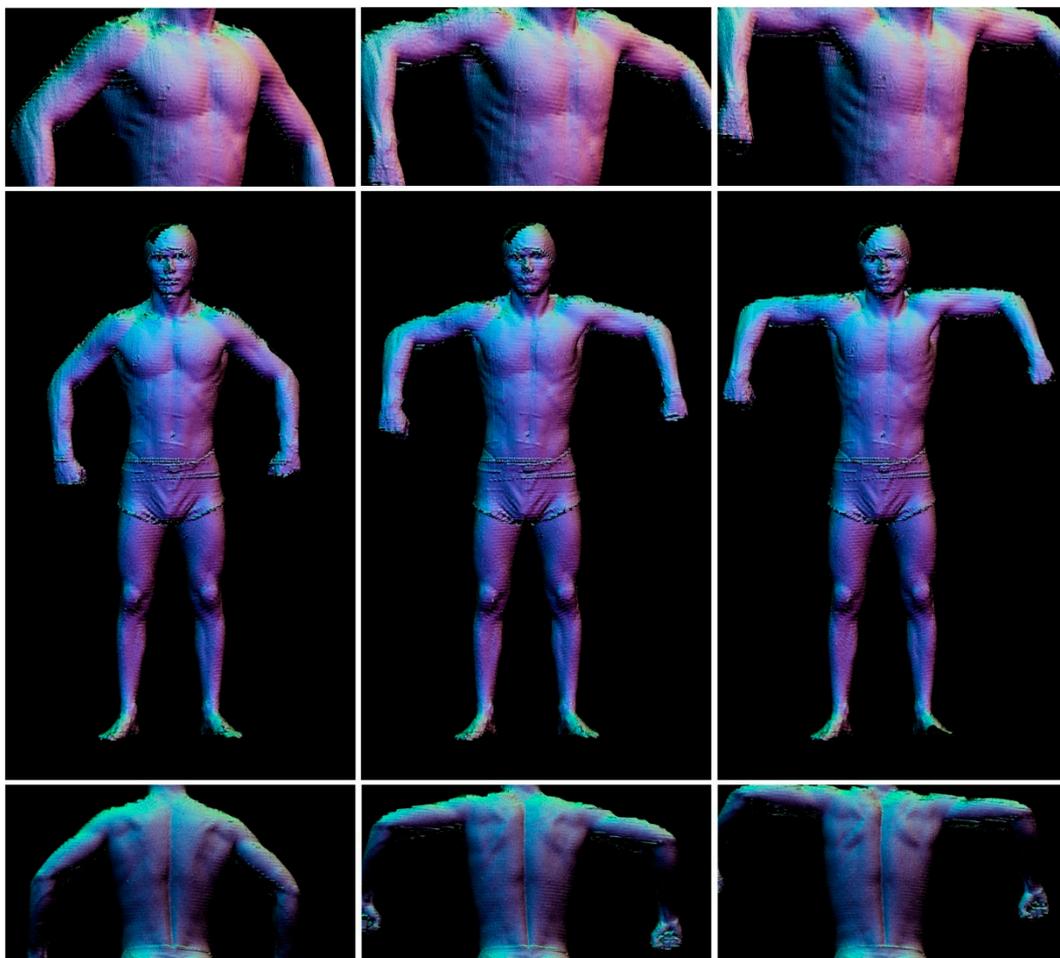
The calculated RMS errors are lower in the static case (0.0 rpm) than in the dynamic cases. In the dynamic cases, no correlation between the artefact rotation speed and RMS error is evident due to the relatively short shutter speed employed for camera image acquisition. However, we suggest that the influence of the rotation speed on the RMS error would be observable at higher rotation speeds.

## 5.2. Validation Using Human Subjects

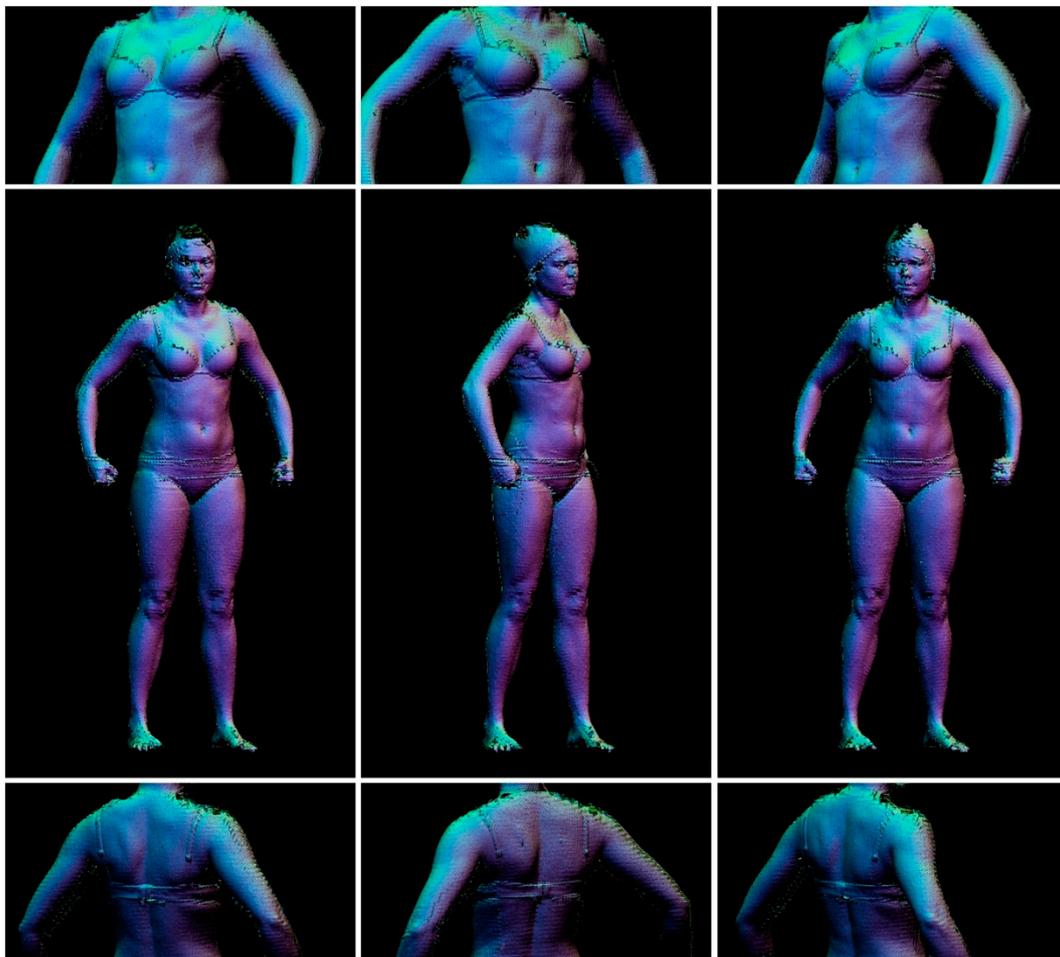
To perform quantitative validation using human subjects, we employed a 3D body scanning system (OGX|MMS) and the body measurement algorithms developed by Markiewicz et al. [50] for reference. The OGX|MMS system was developed for whole-body static measurements as well as body dimension calculations. We performed parallel measurements of four subjects using our 4DBODY scanner and the reference OGX|MMS scanner [50]. Next, we calculated three body dimensions (namely, the waist, hip and chest girths) by applying OGX|MMS-validated algorithms to both

measurements for each individual and compared the results. The average difference between the girths was 0.74 mm, while the maximum difference was 3.21 mm. After careful analysis, we ascertained that these differences originated from two main sources: the different postures adopted during measurement by the 4DBODY and OGX|MMS systems and the measurement uncertainties of both systems. When we performed the same comparison on a mannequin, the average and maximum differences were 0.27 mm and 0.38 mm, respectively, leading us to conclude that the measurement uncertainty is less than 0.5 mm. However, it is very difficult to obtain this level of uncertainty when using actual human subjects.

The 4DBODY system was then tested on several individuals while they performed various basic movements. Measurements were taken at a frequency of 120 Hz, although the geometry reconstruction process was approximately 100 times longer than the acquisition process. The measurement data for each frame consisted of eight directional clouds integrated into one coordinate system. Each single cloud consisted of about 500,000 points. In addition to its  $(x, y, z)$  coordinates, each point contained information about the intensity and normal vector. With a measurement frequency of 30–120 Hz, we collected approximately 4–16 GB of data per second. Each constituent cloud had an average point-to-point distance of 1.5 mm, whereas the average point-to-point distance of the multidirectional cloud was approximately 1.0 mm. These temporal and spatial cloud densities enabled highly effective observation and representation of anatomical structures. Still images demonstrating the exemplary results of the proposed system are presented in Figures 14 and 15. Example animations are also attached to this paper.



**Figure 14.** Three frames depicting a subject raising his shoulders. **Middle row:** full images; **top and bottom rows:** zoomed-in portions of the images in the middle row.



**Figure 15.** Three frames depicting a subject while turning her hips. **Middle row:** full images; **top and bottom rows:** zoomed-in portions of the images in the middle row.

## 6. Conclusions and Future Work

This paper presented a 4DBODY system developed at the Virtual Reality Techniques Division of the Faculty of Mechatronics, Warsaw University of Technology, Poland. This system can realize full human body measurements at frequencies up to 120 Hz and the output point clouds can provide up to four million points per frame. The spatial resolution is approximately 1.0 mm and the uncertainty is less than 0.5 mm in both static and dynamic cases. Each point contains information about its intensity ( $i$ ) and normal vector ( $n_x, n_y, n_z$ ), in addition to the standard ( $x, y, z$ ) coordinates.

These features make the 4DBODY system potentially usable for supporting medical diagnostics and medical rehabilitation monitoring process. The primary advantage of the proposed system is that it does not require physical markers to be attached to the measured human body. Thus, the proposed system is entirely non-invasive, which is crucial for numerous applications. Marker-less measurement also provides information about the geometry of the entire body, which cannot be deduced based on measurements performed only at certain points. The current system performance supports medical diagnosis and rehabilitation monitoring from the point of view of input data for analysis. However, to enable the 4DBODY system to reach its full potential, new analytical algorithms need to be developed. These algorithms should use the full information provided by the extensive surface details during motion.

The 4DBODY system could also be applied in computer graphics and entertainment because the output, which is a cloud of points, can easily be converted into a mesh data type. Furthermore, our system provides a low-cost, or at most medium-cost, solution for dynamic full body measurements.

Although the system performance enables full HD rendering of measurement data, there is a lack of colour mapped on the measured geometry.

In future work, we plan to focus on improving the system accuracy and reducing the processing time. We are considering increasing the number of processing units and re-implementing the graphics processing unit-based algorithms to achieve pseudo real-time reconstruction. We are also contemplating using RGB cameras to add colour information to each measurement point. Furthermore, new methods for 4D data analysis in various applications should be developed and validated.

**Author Contributions:** P.L. and R.S. wrote the article and developed the system. P.L. and M.W. conducted the experiments. P.L. and M.A. performed the validation. M.W. and M.A. also participated in the development of the system. R.S. oversaw the research.

**Funding:** The work described in this article was part of the project PBS3/B9/43/2015, funded by the National Centre for Research and Development with public money for science and statutory work at Warsaw University of Technology.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Buck, U.; Naether, S.; Räss, B.; Jackowski, C.; Tahli, M.J. Accident or homicide—Virtual crime scene reconstruction using 3D methods. *Forensic Sci. Int.* **2013**, *225*, 75–84. [[CrossRef](#)] [[PubMed](#)]
- Se, P.; Jasiobedzki, P. Instant scene modeler for crime scene reconstruction. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, CA, USA, 21–23 September 2005.
- Adamczyk, M.; Sienilo, M.; Sitnik, R.; Woźniak, A. Hierarchical, three-dimensional measurement system for crime scene documentation. *J. Forensic Sci.* **2017**, *2017*, 889–899.
- Yastikli, N. Documentation of cultural heritage using digital photogrammetry and laser scanning. *J. Cult. Herit.* **2007**, *8*, 423–427. [[CrossRef](#)]
- Zlot, R.; Bosse, M.; Greenop, K.; Jarzab, Z.; Juckes, E.; Roberts, J. Efficiently capturing large, complex cultural heritage sites with a handheld mobile 3D laser mapping system. *J. Cult. Herit.* **2014**, *15*, 670–678. [[CrossRef](#)]
- Sitnik, R.; Krzesłowski, J.; Mączkowski, G. Archiving shape and appearance of cultural heritage objects using structured light projection and multispectral imaging. *Opt. Eng.* **2012**, *51*, 021115. [[CrossRef](#)]
- Sitnik, R.; Mączkowski, G.; Krzesłowski, J. Calculation methods for digital model creation based on integrated shape, color and angular reflectivity measurement. In Proceedings of the 2010 Euro-Mediterranean Conference: Digital Heritage, Lemessos, Cyprus, 8–13 November 2010.
- Treleaven, P.; Wells, J. 3D body scanning and healthcare applications. *Computer* **2007**, *40*, 28–34. [[CrossRef](#)]
- Michoński, J.; Glinkowski, W.; Witkowski, M.; Sitnik, R. Automatic recognition of surface landmarks of anatomical structures of back and posture. *J. Biomed. Opt.* **2012**, *17*, 056015. [[CrossRef](#)] [[PubMed](#)]
- Glinkowski, W.; Michoński, J.; Sitnik, R.; Witkowski, M. 3D diagnostic system for anatomical structures detection based on a parameterized method of body surface analysis. In *Information Technologies in Biomedicine*; Piętka, E., Kawa, J., Eds.; Springer: Berlin, Germany, 2010; Volume 2, pp. 153–164. ISBN 9783642131042.
- Schmalz, C.; Forster, F.; Schick, A.; Angelopoulou, E. An endoscopic 3D scanner based on structured light. *Med. Image Anal.* **2012**, *16*, 1063–1072. [[CrossRef](#)] [[PubMed](#)]
- Kontogianni, G.; Georgopoulos, A. Developing and exploiting textured 3D models for a serious game application. In Proceedings of the 2016 8th International Conference on Virtual Worlds and Games for Serious Applications (VS-GAMES), Barcelona, Spain, 7–9 September 2016.
- Szabó, C.; Korečko, Š.; Sobota, B. Processing 3D scanner data for virtual reality. In Proceedings of the 2010 10th International Conference on Intelligent Systems Design and Applications, Cairo, Egypt, 29 November–1 December 2010.
- Ebrahim, M.A.B. 3D laser scanners' techniques overview. *Int. J. Sci. Res.* **2015**, *4*, 323–331.
- Human Solution Informational Material. Available online: [http://www.human-solutions.com/fashion/front\\_content.php?idcat=813&lang=7](http://www.human-solutions.com/fashion/front_content.php?idcat=813&lang=7) (accessed on 27 June 2018).
- Marshall, G.F.; Stutz, G.E. *Handbook of Optical and Laser Scanning*, 2nd ed.; CRC Press: Boca Raton, FL, USA, 2011; ISBN 9781439808795.

17. Schaller, C.; Penne, J.; Hornegger, J. Time-of-Flight sensor for respiratory motion gating. *Int. J. Med. Phys. Res. Pract.* **2008**, *35*, 3090–3093. [[CrossRef](#)] [[PubMed](#)]
18. Geng, J. Structured-light 3D surface imaging: A tutorial. *Adv. Opt. Photonics* **2011**, *3*, 128–160. [[CrossRef](#)]
19. Dunn, S.M.; Keizer, R.L.; Yu, J. Measuring the area and volume of the human body with structured light. *IEEE Trans. Syst. Man Cybern.* **1989**, *19*, 1350–1364. [[CrossRef](#)]
20. Bregler, C.; Hertzmann, A.; Biermann, H. Recovering non-rigid 3D shape from image streams. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Hilton Head Island, SC, USA, 13–15 June 2000.
21. Dellaert, F.; Seitz, S.; Thorpe, C.; Thrun, S. Structure from motion without correspondence. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Hilton Head Island, SC, USA, 13–15 June 2000.
22. Wei, Q.; Shan, J.; Cheng, H.; Yu, Z.; Lijuan, B.; Haimei, Z. A method of 3D human-motion capture and reconstruction based on depth information. In Proceedings of the 2016 IEEE International Conference on Mechatronics and Automation, Harbin, China, 7–10 August 2016.
23. Ceseracciu, E.; Sawacha, Z.; Cobelli, C. Comparison of markerless and marker-based motion capture technologies through simultaneous data collection during gait: Proof of concept. *PLoS ONE* **2014**, *9*, e87640. [[CrossRef](#)] [[PubMed](#)]
24. Sagawa, R.; Ota, Y.; Yagi, Y.; Furukawa, R.; Asada, N. Dense 3D reconstruction method using a single pattern for fast moving object. In Proceedings of the IEEE 12th International Conference on Computer Vision, Kyoto, Japan, 27 September–4 October 2009.
25. Zhang, Z.H. Review of single-shot 3D shape measurement by phase calculation-based fringe projection techniques. *Opt. Lasers Eng.* **2012**, *50*, 1097–1106. [[CrossRef](#)]
26. Griesser, A.; Koninckx, T.P.; Van Gool, L. Adaptive real-time 3D acquisition and contour tracking within a multiple structure light system. In Proceedings of the 12th Pacific Conference on Computer Graphics and Applications, Seoul, Korea, 6–8 October 2004.
27. Lenar, J.; Witkowski, M.; Carbone, V.; Kolk, S.; Adamczyk, M.; Sitnik, R.; van der Krogt, M.; Verdonschot, N. Lower body kinematics based on a multidirectional four-dimensional structured light measurement. *J. Biomed. Opt.* **2013**, *18*, 56014. [[CrossRef](#)] [[PubMed](#)]
28. Wang, T.Y.; Kohli, P.; Mitra, N.J. Dynamic SfM: Detecting scene changes from image pairs. *Comput. Graph. Forum* **2015**, *34*, 177–189. [[CrossRef](#)]
29. Mouragnon, E.; Lhuillier, M.; Dhome, M.; Dekeyser, F.; Sayd, P. Generic and real-time structure from motion using local bundle adjustment. *Image Vis. Comput.* **2009**, *27*, 1178–1193. [[CrossRef](#)]
30. Schwarz, L.A.; Mkhitarayan, A.; Mateus, D.; Navab, N. Estimating human 3D pose from Time-of-Flight images based on geodesic distances and optical flow. In Proceedings of the Face and Gesture 2011, Santa Barbara, CA, USA, 21–25 March 2011.
31. Zhang, L.; Sturm, J.; Cremers, D.; Lee, D. Real-time human motion tracking using multiple depth cameras. In Proceedings of the 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, Vilamoura, Portugal, 7–12 October 2012.
32. Gipsman, A.; Rauschert, L.; Daneshvar, M.; Knott, P. Evaluating the reproducibility of motion analysis scanning of the spine during walking. *Adv. Med.* **2014**, *2014*, 721829. [[CrossRef](#)] [[PubMed](#)]
33. Betsch, M.; Wild, M.; Johnstone, B.; Jungbluth, P.; Hakimi, M.; Kühlmann, B.; Rapp, W. Evaluation of a novel spine and surface topography system for dynamic spinal curvature analysis during gait. *PLoS ONE* **2013**, *8*, e70581. [[CrossRef](#)] [[PubMed](#)]
34. Pons-Moll, G.; Romero, J.; Mahmood, N.; Black, M.J. Dyna: A model of dynamic human shape in motion. *ACM Trans. Graph. TOG* **2015**, *34*, 120. [[CrossRef](#)]
35. Zhang, C.; Pujades, S.; Black, M.; Pons-Moll, G. Detailed, accurate, human shape estimation from clothed 3D scan sequences. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017.
36. Pachoulakis, I.; Kapetanakis, K. Augmented reality platforms for virtual fitting rooms. *Int. J. Multimed. Appl.* **2012**, *4*, 35–46. [[CrossRef](#)]
37. Microsoft Informational Material. Available online: <https://developer.microsoft.com/en-us/windows/kinect/hardware> (accessed on 27 June 2018).

38. DIERS International GmbH Informational Material. Available online: <http://diers.eu/en/products/spine-posture-analysis/diers-formetric-4d/> (accessed on 27 June 2018).
39. Brahme, A.; Nyman, P.; Skatt, B. 4D laser camera for accurate patient positioning collision avoidance, image fusion and adaptive approaches during diagnostic and therapeutic procedure. *Med. Phys.* **2008**, *35*, 1670–1681. [[CrossRef](#)] [[PubMed](#)]
40. Collet, A.; Chuang, M.; Sweeney, P.; Gillett, D.; Evseev, D.; Calabrese, D.; Hoppe, H.; Kirk, A.; Sullivan, S. High-quality streamable free-viewpoint video. *ACM Trans. Graph.* **2015**, *34*, 69. [[CrossRef](#)]
41. Point Grey Informational Material. Available online: <https://eu.ptgrey.com/grasshopper3-23-mp-mono-usb3-vision-sony-pregius-imx174> (accessed on 27 June 2018).
42. Casio Information Material. Available online: <https://www.casio.com/products/projectors/slim-projectors/xj-a242> (accessed on 27 June 2018).
43. Sitnik, R. Four-dimensional measurement by a single-frame structure light method. *Appl. Opt.* **2009**, *48*, 3344–3354. [[CrossRef](#)] [[PubMed](#)]
44. Bergland, G.D. A guided tour of the fast Fourier transform. *IEEE Spectr.* **1969**, *6*, 41–52. [[CrossRef](#)]
45. Takeda, M. Spatial-carrier fringe-pattern analysis and its applications to precision interferometry and profilometry: An overview. *Ind. Metrol.* **1990**, *1*, 79–99. [[CrossRef](#)]
46. Ghiglia, D.C.; Pritt, M.D. *Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software*, 1st ed.; Wiley-Interscience: Hoboken, NJ, USA, 1998; ISBN 9780471249351.
47. Otsu, N. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* **1979**, *9*, 62–66. [[CrossRef](#)]
48. Sitnik, R. New method of structure light measurement system calibration based on adaptive and effective evaluation of 3-D phase distribution. *Opt. Meas. Syst. Ind. Insp. IV* **2005**, *5856*, 109–118.
49. VDI/VDE 2617-6: Accuracy of CMMs—Guideline for the Application of ISO 10360 to CMMs with Optical Distance Sensors. Available online: [https://www.vdi.de/uploads/tx\\_vdirili/pdf/9778569.pdf](https://www.vdi.de/uploads/tx_vdirili/pdf/9778569.pdf) (accessed on 27 June 2018).
50. Markiewicz, Ł.; Witkowski, M.; Sitnik, R.; Mielicka, E. 3D anthropometric algorithms for the estimation of measurements required for specialized garment design. *Expert Syst. Appl.* **2017**, *85*, 366–385. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).