sensors

MDPI

Article

ECG Electrode Localization: 3D DS Camera System for Use in Diverse Clinical Environments

Jennifer Bayer [1,†], Christoph Hintermüller [1,*,†], Hermann Blessberger [2,3] and Clemens Steinwender [2,3]

[1] Institute for Biomedical Mechatronics, Johannes Kepler University, 4040 Linz, Austria
[2] Department of Cardiology, Kepler University Hospital, 4020 Linz, Austria
[3] Medical Faculty, Johannes Kepler University, 4020 Linz, Austria
* Correspondence: christoph.hintermueller@jku.at
† These authors contributed equally to this work.

Abstract: Models of the human body representing digital twins of patients have attracted increasing interest in clinical research for the delivery of personalized diagnoses and treatments to patients. For example, noninvasive cardiac imaging models are used to localize the origin of cardiac arrhythmias and myocardial infarctions. The precise knowledge of a few hundred electrocardiogram (ECG) electrode positions is essential for their diagnostic value. Smaller positional errors are obtained when extracting the sensor positions, along with the anatomical information, for example, from X-ray Computed Tomography (CT) slices. Alternatively, the amount of ionizing radiation the patient is exposed to can be reduced by manually pointing a magnetic digitizer probe one by one to each sensor. An experienced user requires at least 15 min. to perform a precise measurement. Therefore, a 3D depth-sensing camera system was developed that can be operated under adverse lighting conditions and limited space, as encountered in clinical settings. The camera was used to record the positions of 67 electrodes attached to a patient's chest. These deviate, on average, by 2.0 mm ± 1.5 mm from manually placed markers on the individual 3D views. This demonstrates that the system provides reasonable positional precision even when operated within clinical environments.

Keywords: electrode localization; 3D camera; real-time 3D recording; exposure control; white balancing; system calibration; image processing; surface alignment

check for updates

Citation: Bayer, J.; Hintermüller, C.; Blessberger, H.; Steinwender, C. ECG Electrode Localization: 3D DS Camera System for Use in Diverse Clinical Environments. *Sensors* **2023**, *23*, 5552. https://doi.org/10.3390/s23125552

Academic Editor: Fabio Baselice

Received: 24 March 2023
Revised: 15 May 2023
Accepted: 8 June 2023
Published: 13 June 2023

1. Introduction

Models of the human body have gained increasing interest in clinical research and are essential for delivering personalized diagnoses and treatments to patients. They can be used to build a digital twin of a patient's body that can be used for planning curative interventions, predicting the outcomes of intended treatments, or the likelihood of relapses and complications. For most of these models, apart from the knowledge of the patient's exact anatomy, information about physiological processes is required such as the impedance of fibrous tissue forming an infarction scar, which largely differs from the impedance of intact myocardium.

Electrical impedance tomography (EIT) and electrical capacitance tomography (ECT) are used to measure tissue parameters such as impedance or capacitance [1–3]. The origin of cardiac arrhythmias or myocardial infarctions can be identified by integrating ECG recordings [4–6], and the functions within the human brain [7,8] can be visualized using models, including EEG recordings. All these methods require the positions of between 12 and a few hundred sensors to be exactly known. The larger the positional error, the lower the diagnostic value of the results generated by the model. Consequently, they are less suitable for treatment planning, guidance, outcome stratification, or prevention of complications and relapses.

A commonly used approach is to extract the sensor positions, along with the anatomical details, from Magnet Resonance Image (MRI) stacks or X-ray Computed Tomography

(CT) slices. Both approaches require special markers to be attached to the sensors, which are visible in MRI [9,10] or CT scans [11]. Identifying the sensor positions from MRI or CT scans yields the smallest positional errors compared to the true sensor position. However, this approach significantly hinders the clinical uptake and widespread use of electrical impedance tomography (EIT), electrical capacitance tomography (ECT), noninvasive imaging of cardiac electrophysiology (NICE), and other model-based approaches. Patients either have to be exposed to large amounts of ionizing radiation when using CT scans, limiting the use of the aforementioned methods to three applications per year. Although MRI is not bound by this limitation, it is only covered by insurance companies if it is required for obtaining a proper diagnosis and evaluating outcomes.

Given these limitations, alternative approaches that decouple the generation of the underlying anatomical models from the localization of the sensors have been tested [12–14]. Alternatives such as magnetic digitizer systems, e.g., the Polhemus Fastrak [12], tracked gaming controllers [13], or motion capture systems, have been used to identify the positions of electrodes relative to the patient's body. The use of photogrammetry, visual odometry, and stereoscopic approaches was already considered more than 15 years years ago [15,16]. The Microsoft Kinect 3D depth-sensing camera (3D DS) was one of the first compact and affordable devices. Nowadays, modern coded light and stereo vision-based models are portable and lightweight enough to be easily attached to or even integrated within a standard tablet computer.

In the past few decades, 3D DS cameras have mainly been used in EEG-based studies to locate EEG sensors on the patient's skull [12,14,17]. All of them use the recorded EEG signals to localize brain activity or identify the focus of a seizure within the cortex. In contrast, very few studies report the use of 3D DS cameras to locate ECG sensors on the chest or even the whole torso [18–20]. One reason for this may be that the skull is a rigid structure that does not change its shape when the subject moves during the recording. In contrast, when recording the sensor position on the torso, the patient needs to maintain a specific posture. The instructions provided to the patient on how to achieve and maintain this posture are integral to the entire recording procedure.

In the present work, the positions of 64 ECG electrodes mounted on the torso are recorded using 3D DS camera readings only. Section 2 encompasses descriptions of the overall structure of the developed 3D DS camera-based system, method, and algorithm used for the real-time recording of the individual 3D views of the torso (Section 2.2); the postprocessing steps necessary for extracting the electrode positions (Section 2.3); and the recording protocol used and the instructions provided to each subject participating in the clinical testing (Section 2.4). In Section 3, the results obtained from the five subjects are presented, and in Section 4, these results are discussed.

2. Materials and Methods

The 3D depth-sensing (3D DS) camera-based measurement of electrode positions can be divided into four main steps: (i) selecting the appropriate 3D DS camera, (ii) defining an appropriate measurement protocol, (iii) recording the 3D surfaces in real time, and (iv) extracting the electrode center points.

The most important component for recording the electrode positions is the 3D camera. It can be characterized by various parameters such as the closest distance d_{near} and the vertical ψ_V and horizontal ψ_H fields of view (FOV). These parameters define the volume in front of the camera in which objects must be placed to be accurately captured by the depth sensor. Based on these considerations, the Intel Realsense SR300 3D DS camera [21] was selected. Descriptions of the exact selection criteria that led to this decision can be found in Section 2.1.

The human torso represents a flexible object that offers several degrees of freedom for movement and deformation in contrast to the rather rigid skull. The position of each ECG electrode perceived by the 3D DS camera and its relative position to the other electrodes is directly affected by the movements of the patient's body. Therefore, it is essential to

define an appropriate recording protocol before the first 3D data set is recorded. As large displacements may prevent the successful extraction of the electrode positions, the patient is required to actively maintain the same posture throughout the recording procedure. Details on how this active engagement of the patient can be achieved are described in Section 2.4. For the remaining steps, (iii) real-time recording and (iv) offline processing, Figure 1 provides an overview of the necessary sub-steps and their interdependence:

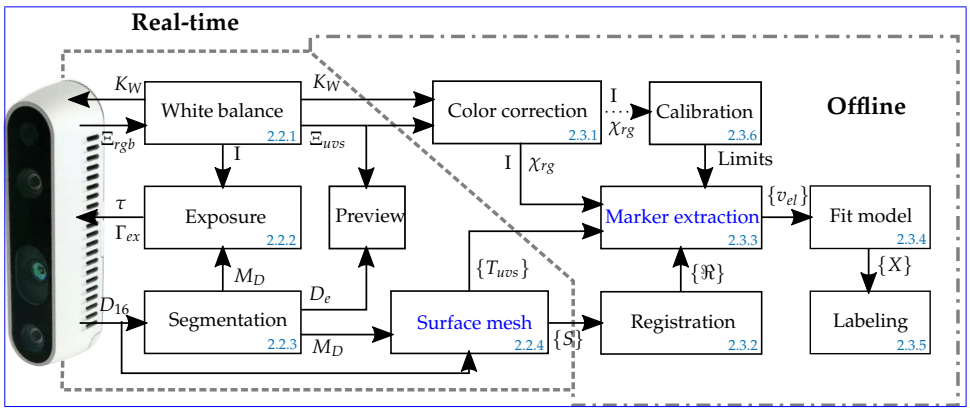

Figure 1. Schematic presentation of the overall approach used to extract ECG electrode positions from 3D depth-sensing camera data. The first step involves methods to control the camera's white balance and exposure settings and generate textured 3D surface meshes from the recorded depth data. During the offline processing step, these surfaces are aligned to extract the electrode positions within clusters of marker vertices found using texture images in the RGB chromacity color space and 3D surfaces.

A 3D DS camera combines a depth sensor and an RGB color sensor in a single device. These two sensors simultaneously record an RGB color image Ξ_{rgb} and a 16-bit depth image D_{16}. The latter image encodes the distance d between the camera and the objects located in front of the camera.

The developed real-time recording system is intended for use in diverse clinical settings such as examination rooms in outpatient clinics or local cardiology practitioner clinics. The lighting conditions encountered depend on the pointing direction of the camera and the number of light sources, as well as their brightness and color hue. In order to properly handle these conditions, the white-balancing settings, exposure time τ, and overall gain γ_{ex} of the color sensor are continuously adjusted in real time. Automatic white balancing (AWB), which is described in detail in Section 2.2.1, uses Ξ_{rgb} to estimate the color temperature K_W of the dominant light source.

At the same time, a binary mask M_D is generated from the depth image D_{16}. This M_D splits D_{16} into foreground pixels representing the torso surface and objects in the background (Section 2.2.3). M_D is used to generate a 3D mesh S of the imaged torso surface (Section 2.2.4) and tune the exposure time τ and global gain setting Γ_{ex} of the color sensor. This is achieved by combining M_D with the brightness information I of the color image Ξ_{rgb} obtained during the AWB step (Section 2.2.2). The mask M_D is also used to outline the patient's contours on the real-time preview screen, along with various system parameters.

When the trigger is pressed, the triangulation component (Figure 1) generates a 3D surface mesh S, which is stored along with the corresponding texture information Ξ_{uvs} of the torso created from the RGB color image Ξ.

In the offline processing step, a pairwise iterative closest-point (ICP) algorithm is used to align the recorded surfaces S with each other. The resulting transformation matrices $\mathfrak{R}$ are used to extract the 3D positions from the color-corrected texture images Ξ_{uvs}, which have been stored alongside each S (Section 2.3.2). In order to facilitate the steps necessary to identify the color markers attached to the electrodes, an additional color-correction step, which is described in Section 2.3.1, is conducted. The aim of this step is to ensure that the

patient's skin color and marker colors are accurately represented across all the recorded texture images Ξ_{uvs}. To achieve this, the Ξ_{uvs} are split into a chromacity image χ_{rg} and the corresponding intensity image I. Both are used to identify the red and blue pixels and related 3D points corresponding to each electrode marker. Details on how this is achieved can be found in Section 2.3.3.

The centers of these markers are coaligned with the centers of the electrode clips and patches. Their positions on the surface are computed by fitting a planar model (Section 2.3.4) to the extracted red and blue points. In the final labeling step (Section 2.3.5), the electrode positions are assigned to the corresponding ECG signals recorded from the patient's torso.

The colors of the markers vary depending on the position and orientation of the electrode clip relative to the torso and 3D DS camera. Therefore, a dedicated calibration procedure is utilized, which is outlined in Section 2.3.6, to determine the ranges of the red and blue color values that represent the electrode markers.

2.1. Selecting the Camera

The selected Intel Realsense SR300 3D DS camera [21] is used in narrow or crowded places such as examination rooms in outpatient clinics and cardiology practitioner clinics. In these places, the patient is typically seated on an examination bed or chair placed close to the wall. Consequently, the closest distance d_{near} relative to the depth sensor at which objects may be placed has to be shorter than the shortest horizontal distance $d_{H,min}$ of the patient's torso to any surrounding obstacles such as walls or furniture. The horizontal ψ_H and vertical ψ_V FOVs determine how tall or wide the closest object can be to be fully captured in its height and width. The minimum required values for ψ_H and ψ_V can be approximated based on the patient's approximated h_{torso} and the 3D DS camera's d_{near} using the following relationships:

$$\begin{gathered} d_{near} < d_{H,min} \\ \psi_{min} = 2\arctan\left(\frac{h_{torso}}{d_{near} + d_{H,min}}\right) \\ \max(\psi_H, \psi_V) >= \psi_{min} \end{gathered} \tag{1}$$

According to the datasheet [21] the depth sensor can capture objects located at distances between 20 cm and 150 cm from the camera. This range is more than sufficient to record the surface of the torso. The depth information of each object is captured using an infrared sensor in combination with a near-infrared projector [21,22].

The depth images D_{16} are recorded in 4:3 image format, covering a horizontal FOV of 69 degrees and a vertical FOV of 54 degrees at a depth resolution of less than 1 mm. The color sensor of the camera generates the RGB images Ξ_{rgb} in 16:9 format. Its horizontal FOV of 68 degrees is sufficiently well-paired with the horizontal FOV of the depth sensor. With a ψ_V of 41 degrees, it covers only 2/3 of the depth sensor in height. This results in a lack of color information for the pixels close to the top and bottom edges of the depth image D_{16}, which was considered when outlining the measurement protocol in Section 2.4.

2.2. Real-Time Recording

2.2.1. Automatic White Balancing

The color sensor used by the Intel Realsense 3D DS camera offers the possibility to manually tune the color gains $\Gamma_{\hat{R}}$, $\Gamma_{\hat{G}}$, $\Gamma_{\hat{B}}$ indirectly by adjusting their color temperature parameter K_W. This was used to implement a custom AWB component (Figure 1), along with the algorithm proposed in [23], which can handle these varying conditions. After applying a lookup table $\hat{v} = \lfloor v \rfloor$ based on linearization (gamma decompression) and normalization to the interval $[0, 1]$ of the red $\hat{R} = \lfloor R \rfloor$, green $\hat{G} = \lfloor G \rfloor$, and blue $\hat{B} = \lfloor B \rfloor$

color channels, the resulting linear RGB image Ξ is converted into an RGB chromacity χ_{rg} image and a linear grayscale image I that encodes the brightness of each pixel.

$$I_i = \hat{R}_i + \hat{G}_i + \hat{B}_i \tag{2}$$
$$r_i = \hat{R}_i / I_i,\ g_i = \hat{G}_i / I_i,\ b_i = \hat{G}_i / I_i$$

From χ_{rg}, all pixels (r_i, g_i, b_i) are selected that encode shades of gray. The red r, green g, and blue b chromacity values of these pixels are located within a small area around the neutral color gamut point, which has a color temperature of 5500 K, as shown in Figure 2. The basic assumption is that these pixels most likely correspond to object surfaces of a neutral gray color. Consequently, a reddish-colored taint in these pixels must be caused by a low K value of the predominant illumination, and a bluish cast most likely results from a light source with a large K. Overexposed pixels are excluded, as their color most likely results from the saturation of at least one of the three color channels and thus does not properly represent the skin color of the patient or the color of the illuminant. Likewise, underexposed pixels are not considered, as their color is most likely caused by camera noise rather than the light reflected by the imaged object.

For adjusting the color temperature setting K_W of the 3D DS camera, only pixels $(r_{i,W}, g_{i,W}, b_{i,W}, I_{i,W})$ that are located within a small area surrounding the neutral color gamut point are selected, which is, according to Cohen [23], defined by the chromacity values $\overline{r} = 0.363$, $\overline{g} = 0.338$, and $\overline{b} = 0.299$. This area encloses all pixels that are located within the following two ellipses centered at the color gamut point:

$$\frac{(r_{i,W} - \overline{r})^2}{\sigma_r^2} + \frac{(g_{i,W} - \overline{g})^2}{\sigma_g^2} <= 1 \text{ and } \frac{(b_{i,W} - \overline{b})^2}{\sigma_b^2} + \frac{(g_{i,W} - \overline{g})^2}{\sigma_g^2} <= 1 \tag{3}$$
$$0.00155 I_{max} < I_i < 0.955 I_{max} \tag{4}$$

Their primary and secondary axes are defined by the standard deviations for the red $\sigma_r = 0.0723$, green $\sigma_g = 0.0097$, and blue $\sigma_b = 0.0749$ chromacity values with respect to the neutral color gamut point, which was determined in [23]. The maximum intensity encountered is $\max I = 3$.

The lower $I_{min} = 0.02$ and upper $I_{max} = 0.98$ exposure limits, as defined for each channel in [23], are linearized to $I_{min} = 0.02/12.92$ and $I_{max} = ((0.98 + 0.055)/1.055)^{2.4}$ before applying them to the overall linear intensity values I.

To match K_W with the color temperature K of the light source, the overall color gain Γ_K of the camera is estimated. The following model is used to simulate how the camera adjusts the gain $\Gamma_{\hat{R}}$ of its red and blue $\Gamma_{\hat{B}}$ channels when K_W is updated.

$$\Gamma_R = \Gamma_K,\ \Gamma_{\hat{B}} = \frac{1}{\Gamma_K},\ \Gamma_K = \frac{2(K_W - K_{W,min})}{K_{W,max} - K_{W,min}},\ K_W = K_{W,n+1} = \gamma K_{W,n} \tag{5}$$

Neither the lower and upper limits for $\Gamma_{\hat{R}}$ and $\Gamma_{\hat{B}}$, nor the color temperature corresponding to equal gain values $\Gamma_{\hat{R}} = \Gamma_{\hat{B}} = 1$, are documented for up-to-date 3D DS cameras. It is assumed that $\Gamma_{\hat{R}} = \Gamma_{\hat{B}} = 1$ corresponds to the center color temperature $\overline{K}_W = (K_{W,min} + K_{W,max})/2$ between the minimum $K_{W,min}$ and maximum $K_{W,max}$ values of the color sensor. On startup, K_W is initialized to $K_{W,0} = \overline{K}_W$. For the recorded color images Ξ_{n+1}, the corresponding $K_{W,n+1} = \gamma K_{W,n}$ is estimated from the previous value of $K_W = K_{W,n}$ and a scaler γ reflecting the relative change of K_W between two consecutive Ξ_n. The color sensor of the used camera has a rolling shutter. Therefore, color images are only considered for estimating the scaler γ and K_W after the next exposure time interval has elapsed.

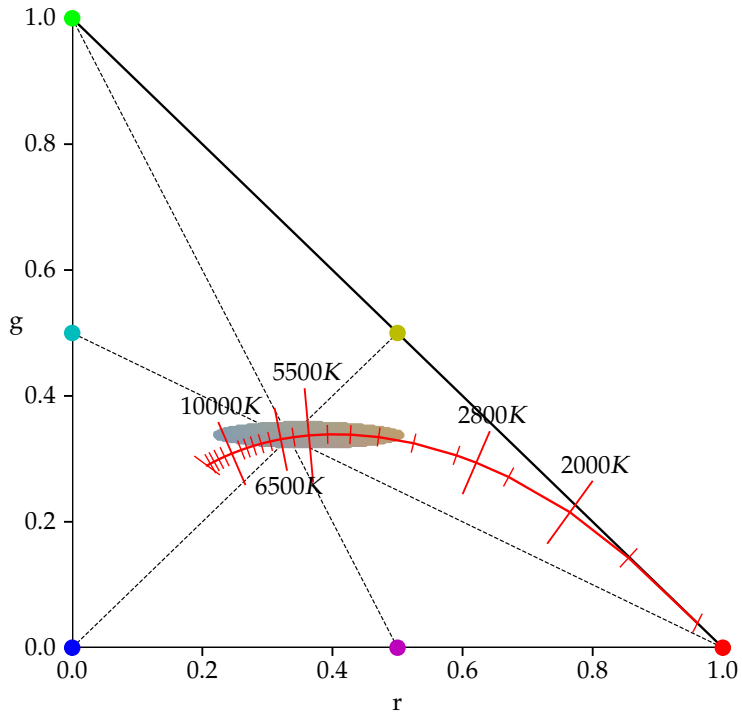

Figure 2. The elliptic area in the RGB chromacity space corresponds to the pixels encoding shades of gray [23]. The red r and green g chromacity values of the natural illumination color gamut of 5500 K define a point that is shifted slightly off the mean RGB chromacity toward yellowish colors. Standard 3D DS cameras designed for indoor use such as the Intel Realsense™ typically allow adjusting the gains for the red and blue channels to illumination color gamuts between, for example, 2800 K and 6500 K, as indicated on the color gamut curve. However, in real clinical settings, gamuts from 2000 K up to 10,000 K can be expected, depending on the number of light sources and shades cast by objects and people.

The goal is to minimize the distance between the average $\overline{r}_W$, green $\overline{g}_W$, and blue $\overline{b}_W$ chromacities of the selected pixels and the $\overline{r}, \overline{g}, \overline{b}$ of the color gamut point that corresponds to a color temperature of $K = 5500$ Kelvin. To achieve this, $\overline{r}_W$, $\overline{g}_W$, and $\overline{b}_W$ are multiplied by the unknown intensity $\dot{y} = \overline{R}_W + \overline{G}_W + \overline{B}_W$ to obtain the corresponding mean red $\overline{R}_W$, green $\overline{G}_W$, and blue $\overline{B}_W$ color values. These values are scaled by γ using (5). After scaling, the updated $\overline{r}_W$, $\overline{g}_W$, and $\overline{b}_W$ are computed using (2).

$$\overline{r} = \frac{\overline{r}_W \Gamma_K \dot{I}}{\overline{r}_W \Gamma_K \dot{I} + \overline{g}_W \dot{I} + \frac{\overline{b}_W \dot{I}}{\Gamma_K}}, \; \overline{g} = \frac{\overline{g}_W \dot{I}}{\overline{r}_W \Gamma_K \dot{I} + \overline{g}_W \dot{I} + \frac{\overline{b}_W \dot{I}}{\Gamma_K}}, \; \overline{b} = \frac{\frac{\overline{b}_W \dot{I}}{\Gamma_K}}{\overline{r}_W \Gamma_K \dot{I} + \overline{g}_W \dot{I} + \frac{\overline{b}_W \dot{I}}{\Gamma_K}} \tag{6}$$

It is obvious that the unknown intensity $\dot{I}$ does not have any impact on the result. It can be omitted from (6) and γ. Consequently, Γ_K can be computed from $\overline{r}, \overline{g}, \overline{b}, \overline{r}_W, \overline{g}_W$, and $\overline{b}_W$ directly.

In Figure 2, it can be observed that the curve along which the color gamut point moves can be approximated for color temperatures $K \leq 5000$ K by the line connecting the red corner $(r = 1, g = 0, b = 0)$ and the midpoint $(r = 0, g = 0.5, b = 0.5)$ between the blue $(r = 0, g = 0, b = 1)$ and green corners $(r = 0, g = 1, b = 0)$ of the chromacity space. For color temperatures $K > 5000$ K, the curve can be approximated by the line connecting the blue corner $(r = 0, g = 0, b = 1)$ with the midpoint $(r = 0.5, g = 0.5, b = 0)$ between the red $(r = 1, g = 0, b = 0)$ and green $(r = 0, g = 1, b = 0)$ corners, respectively. The two midpoints $(r = 0.5, g = 0.5, b = 0)$ and $(r = 0, g = 0.5, b = 0.5)$ correspond to the yellow $\overline{y} = (\overline{r} + \overline{g})/2$ and cyan $\overline{c} = (\overline{g} + \overline{b})/2$ chromacities, respectively. Based on the ratio $\overline{y}/\overline{b}$,

the average chromacity value $\overline{g}_W\gamma$ of the green channel scaled by γ can be expressed. The resulting expression is inserted into the quadratic Equation (8) obtained from the ratio $\overline{r}/\overline{c}$:

$$\frac{\overline{r}}{\overline{c}} = \frac{2\overline{r}}{\overline{g}+\overline{b}} = \frac{2\overline{r}_W\gamma}{\overline{g}_W + \frac{\overline{b}_W}{\gamma}}, \ \frac{\overline{y}}{\overline{b}} = \frac{\overline{r}+\overline{g}}{2\overline{b}} = \frac{\overline{r}_W\gamma + \overline{g}_W}{2\frac{\overline{b}_W}{\gamma}} \tag{7}$$

$$\gamma^2\overline{r}_W - \frac{\overline{r}\overline{b}_W}{\overline{b}} = 0 \tag{8}$$

Solving (8) with respect to γ yields

$$\gamma = \sqrt{\frac{\overline{r}\overline{b}_W}{\overline{r}_W\overline{b}}} \tag{9}$$

Along with γ, the actual error E between the neutral illumination color gamuts $\overline{r}$, $\overline{g}$, $\overline{b}$ and $\overline{r}_W$, $\overline{g}_W$, and $\overline{b}_W$; the expected error E^* after scaling $\overline{r}_W$ and $\overline{b}_W$ by γ; and the updated value $K_W^+ = K_W\Gamma_K$ are computed using (5):

$$E = (\overline{r} - \overline{r}_W)^2 + (\overline{g} - \overline{g}_W)^2 + (\overline{b} - \overline{b}_W)^2 \ \text{ and} \tag{10}$$

$$E^* = (\overline{r} - \overline{r}_W\gamma)^2 + (\overline{g} - \overline{g}_W)^2 + (\overline{b} - \frac{\overline{b}_W}{\gamma})^2 \tag{11}$$

Based on these equations, the K_W setting of the 3D DS camera is updated $K_W = K_W\Gamma_K$ if $E^* < E$. During testing, it was found that numerical inaccuracies can prevent the computation of the appropriate estimates for the color temperature K of the predominant illuminant. Therefore, a numerically stable test is used instead to determine whether K_W has to be updated or its current value can be kept.

$$(E > 1.758 \times 10^{-8}) \wedge (E - E^* > 10^{-15}) \tag{12}$$

2.2.2. Patient-Locked Auto-Exposure

In addition to the overall color appearance, the light sources that are present also affect the overall light intensity I, which among others, can vary depending on the viewing direction of the 3D DS camera. For example, in the case shown in Figure 3a, the camera is pointing toward a window. In Figure 3b, the camera is pointing in the opposite direction toward the door.

In order to maintain a constant illumination intensity I of the patient's torso, independent of the viewing direction and the overall brightness of all present light sources, the histogram-based auto-exposure AE algorithm proposed in [24] was adopted.

This algorithm is implemented in the exposure component (Figure 1). It considers only the pixels in Ξ that correspond to the patient's torso. These pixels are selected by segmenting the depth image D recorded by the 3D DS camera into a foreground object (the patient) and the remaining background using the approach outlined in Section 2.2.3. The binary mask M_D obtained in this segmentation step is mapped to the color image Ξ using the texture coordinates $\boldsymbol{v}_{uvs}$ computed from the depth image D_{16} by the camera control library. All brightness values I_i of all pixels covered by the mapped mask M'_D are considered for adjusting τ. Any other pixels and pixels that are over- or underexposed according to Equation (4) are discarded.

The algorithm proposed by Chen and Li [24] uses the histogram of the gamma-compressed grayscale image $\overline{\Xi}$ computed from Ξ. In order to avoid the computational burden required by an explicit conversion between the linear illumination image I and $\overline{\Xi}$, the histogram $H(V)$ is directly computed from the linearized illumination values I_i of the selected pixels. This is accomplished by maintaining a lookup table that lists the linearized bin boundary values $\hat{h}_v$ corresponding to the uniform boundaries h_v of the grayscale histogram $H(V)$. The histogram $H(V)$ can then be generated for all considered I_i using a

left bisection search to scan this lookup table, which is far less computationally demanding. A further reduction can be achieved by precomputing the differences $\Delta_H^2 = (I_i - 128)^2$ and $\Delta_H^3 = (I_i - 128)^3$ for each bin, which are used to calculate the skewness $S(V)$ of $H(V)$.

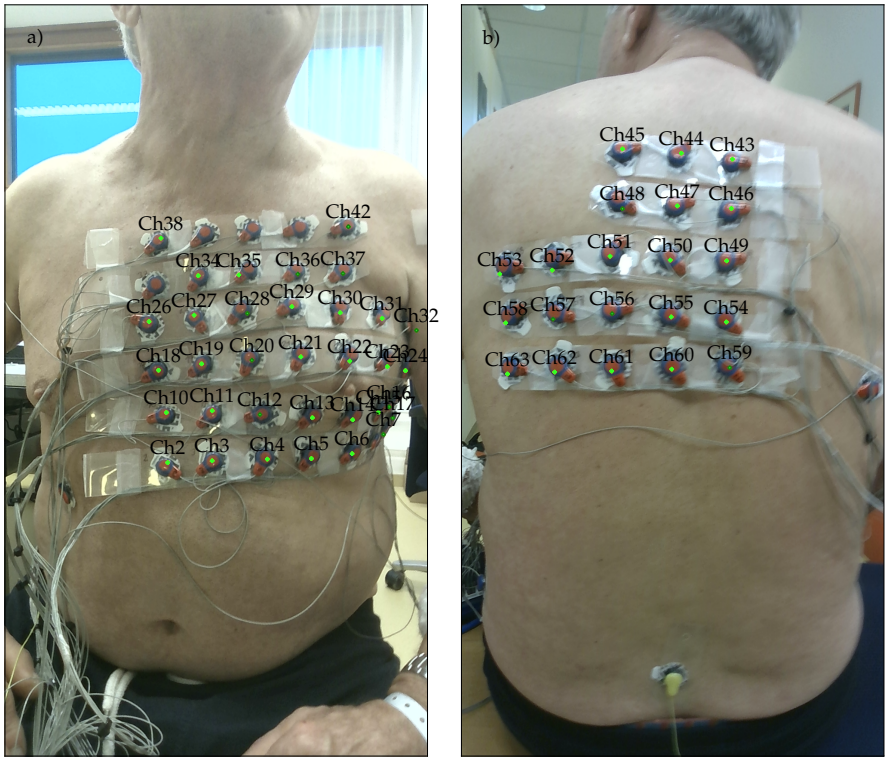

Figure 3. The histogram-based auto-exposure algorithm considers only the pixels that most likely correspond to the patient and ignores any others. This ensures that the brightness of the patient's skin remains as constant as possible, regardless of whether the camera points toward a window (**a**) or the darkest corner of the room (**b**). Each visible electrode is labeled with the corresponding channel number and the '+' markers indicate the projected location of the computed electrode position.

To compute the values of the exposure time τ and overall gain Γ_{ex} to be set on the camera, the overall exposure parameter $\overset{*}{\tau}$ is used.

$$\overset{*}{\tau}_{n+1} = \overset{*}{\tau}_n - \frac{S(V)}{\tau_\Delta} N_\tau \Gamma_{ex,n}$$

$$\Gamma_{ex,n+1} = \max(\min(\frac{\overset{*}{\tau}_{n+1}}{\overset{**}{\tau}}, \Gamma_{ex,max}), \Gamma_{ex,min}) \tag{13}$$

$$\tau = \max(\min(\frac{\overset{*}{\tau}}{\Gamma_{ex,n+1}}, \tau_{max}), \tau_{min})$$

$$\overset{**}{\tau} = \max(\min(\frac{\tau_{frame}}{\tau_\Delta}, \tau_{max}), \tau_{min}) \tag{14}$$

The parameter τ_Δ represents the size of one τ step in milliseconds, $N_\tau = 5$ represents the number of steps to take if $S(V) = 1$, and $\tau_{frame} = 100\,\text{ms}$ represents the optimal exposure time for each frame. The value of τ_Δ depends on the actual step size in ms offered by the 3D DS camera.

2.2.3. Depth Segmentation

The binary mask M_D is created from the 16-bit depth images D_{16} recorded by the 3D DS camera. It splits the image into the patient and any surrounding objects, obstacles, and

relevant edges. This implementation was inspired by the Canny edge detection algorithm proposed in [25]. The algorithm uses two thresholds to find the edges in an image Ξ based on the gradient $\Delta\overline{\Xi}$ of its corresponding grayscale image $\overline{\Xi}$. Pixels that have a gradient value Δx_i that exceeds the upper limit are considered to be part of an edge. Pixels with a value of Δx_i between the two limits are only included in an edge if they are adjacent to an already identified edge pixel. To improve the obtained set of edges and reduce the number of edges caused by noise, the grayscale $\overline{\Xi}$ is smoothed using a Gaussian filter.

This approach was adopted for processing depth images D_{16} that contain pixels for which no valid depth value $D_i = 0$ is available. The computation of the depth value gradient ΔD_i and one of the corresponding Gaussian filter weights w_i are computationally too demanding to be computed in real time. Therefore, the depth gradient values $\Delta D_i = \sqrt{\Delta d_{x,i}^2 + \Delta D_{y,i}^2}$ of D are rounded to the closest 16-bit integer value $\Delta_{16}D$. The resulting reduced number of $\Delta_{16}D$ and corresponding distinct weights w_{16} are stored in a precomputed weights table instead of directly computing w_i on every iteration for each pixel. This avoids the computationally demanding operations of computing e^x and $\sqrt{x}$ in real time. A companion table with squared boundary values $\Delta_{16}^2 D = {}^{(\Delta_{j,16}^2 D + \Delta_{j+1,16}^2 D)}/_4$ between the individual $\Delta_{16}D$ ensures that the w_i for a pixel D_i of D can be generated through a fast left bisection search. Pixels $D_i = 0$ represent objects without a defined depth, and their values are copied to the smoothed depth image $\tilde{D}$ image without any changes.

The smoothed $\tilde{D}$ is filtered using an octagonal Laplace kernel to find the initial set of edge pixels d_e,

$$\begin{bmatrix} -\frac{1}{\sqrt{2}} & -1 & -\frac{1}{\sqrt{2}} \\ -1 & 4+2\sqrt{2} & -1 \\ -\frac{1}{\sqrt{2}} & -1 & -\frac{1}{\sqrt{2}} \end{bmatrix} \tag{15}$$

An octagonal kernel has the advantage that all distances between the eight-connected neighbor pixels d_8 and the central pixel d_c are of equal length.

All pixels d that exhibit a sign change between opposing neighbor pixels ∇d_8 on the Laplacian image $\nabla\tilde{D}$ are included in the initial set of edge points d_e. Pixels d_e that have at least one neighbor $d_{k,8} = 0$ with an undefined depth are considered primary edge pixels e_P. Their actual $\Delta\tilde{D}(d_e)$ values are computed using the following approach:

$$\Delta\tilde{D}(d_e) = \max\left(\begin{cases} \tilde{D}_{+k,8} - \tilde{D}_{-k,8} & \text{if } \tilde{D}_{-k,8} > 0 \text{ and } \tilde{D}_{+k,8} > 0 \\ 2 * \tilde{D}_{-k,8} - \tilde{D}_i & \text{if } \tilde{D}_{-k,8} > 0 \\ 2 * \tilde{D}_{+k,8} - \tilde{D}_i & \text{if } \tilde{D}_{+k,8} > 0 \end{cases}\right) \tag{16}$$

All d_e where $\Delta\tilde{D}(d_e) > \Delta\tilde{D}_P$ are marked d_P, whereas any other d_e are only considered if the Canny rule for minor edge pixels d_M holds. This rule has been modified for use on depth images D as follows:

$$(\Delta\tilde{D}(d_e) > \Delta\tilde{D}_M) \wedge ((\Delta\tilde{D}(d_e) > \Delta\tilde{D}_P) \vee (\frac{\Delta\tilde{D}(d_e) - \Delta\tilde{D}_M}{\Delta\tilde{D}_P - \Delta\tilde{D}(d_e)} > 1)) \tag{17}$$

The upper Canny limit $\Delta\tilde{D}_P$ is set to 1.2 cm and the minor limit $\Delta\tilde{D}_M$ is set to 0.35 cm.

A binary depth mask M_D is created from all pixels $d_i > 0$ in D of a known depth. Pixels d_e located at any of the edges are excluded from M_D. The resulting M_D is split into 9 segments $M_{D,9}$. The pixels M_i within the central $M_{D,9}$ are labeled with respect to the different objects and components they represent. The labeled four-connected components L_4 are sorted by size. The largest L_4 that touches the segment boundary is extended to all other $M_{D,9}$ segments using the flood-fill method, starting from the center of mass of L_4. At the end of this step, all adjacent edge pixels d_e are appended to the extended L_4^+ representation of L_4.

As the depth values at the boundaries of L_4^+ can largely vary, the following approach is used to remove any unrelated outliers. This approach is based on the observation that

the boundaries of the patient's torso are well-separated from the background along the vertical direction and above the head.

$$D_{close} < D_i(M_i) < D_{far} \tag{18}$$

$$D_{far} = \min\left(\max(D_{r,far}) + \frac{\max(D_{r,far}) - \min(D_{r,close})}{3}, \overline{D_{r,far}} + 3\sigma(D_{r,far})\right) \tag{19}$$

$$D_{close} = \max\left(\min(D_{r,close}) - \frac{\max(D_{r,close}) - \min(D_{r,close})}{3}, \overline{D_{r,close}} - 3\sigma(D_{r,far}), D_{near}\right) \tag{20}$$

The values $D_{r,close}$ and $D_{r,far}$ correspond to the smallest and largest depth values encountered for the mask pixels M_i within each row of L_4^+, and $\overline{D_{r,close}}$, $\overline{D_{r,far}}$, $\sigma D_{r,close}$, and $\sigma(D_{r,far})$ represent their mean and standard deviations. Any pixel M_i for which the condition in (18) does not hold is removed from L_4^+. In the case that either the number of pixels of L_4^+ is less than 200 or no appropriate values for D_{far} or D_{close} could be found, the current L_4 is discarded and the search for a suitable L_4^+ representing the patient is attempted with the next larger L_4. If no suitable L_4 is left, segmentation is aborted and real-time processing continues with the next set of depth and color image frames recorded by the 3D DS camera.

2.2.4. Surface Mesh Generation

The final surface mesh is generated by converting the depth image D into a corresponding point cloud P. Therein, each point $\boldsymbol{v}_i$ corresponds to a specific pixel d_i in D. In the case of pixels $d_i = 0$ without a defined depth value, the origin point $\boldsymbol{v}_i = O = (0, 0, 0)$ is assigned. The unique correspondence between any d_i and its corresponding $\boldsymbol{v}_i$ allows creating S by mapping a pre-triangulated grid G to P. Any triangle T that includes at least one $\boldsymbol{v}_i$ for which $d_i = 0$ is dropped from G.

Before S is stored on disk using the **.obj** format, along with Ξ_{uvs} and the color temperature setting K_W it was recorded with, degenerated $T_{A=0}$ and occluded triangles T_{-1} that do not correspond to a valid surface patch are removed. The filtering of T_{-1} is facilitated by the fact that 3D DS cameras, especially those that can capture objects located a short distance from the camera, use a dedicated RGB color sensor to record Ξ_{uvs}. This sensor is typically attached to the left or right side of the depth sensor system and thus views the imaged object from a slightly different angle. This difference in viewing angle and FOV between the depth and the color sensor is sufficiently large to identify triangles that do not represent a part of the object's real surface. This small difference in viewing angle causes the surface normal $\boldsymbol{n}_{-1}$ to flip its direction between the representation of T_{-1} in the depth image D and in Ξ. This flip is not plausible as it would mean that the color sensor is capturing the back side of T_{-1}, whereas the depth sensor captures its front side. This is prevented by the fact that both sensors are mounted on the same support. The following approach exploits this fact by identifying triangles where the sign, and thus the direction, of the surface normal vector appears flipped in Ξ compared to D.

The pre-triangulated grid G is initialized such that the normal vector $\boldsymbol{n}_T$ of each triangle T on S points toward the camera and is oriented in the negative $\langle \boldsymbol{m}, \boldsymbol{Z} \rangle < 0$ viewing direction $\boldsymbol{Z}$ of the camera. For every valid T of initial surface mesh S, the normal vector $\boldsymbol{n}_{uvs}$ of its representation in Ξuvs T_{uvs} must also point in the $-\boldsymbol{Z}_{uvs}$ direction. Triangles T_{-1} where the signs of $\boldsymbol{n}$ and $\boldsymbol{n}_{uvs}$ are opposite, indicated by $\langle \boldsymbol{n}_{uvs}, \boldsymbol{Z}_{uvs} \rangle >= 0$, suggest that triangle T_{-1} likely does not represent a valid part of S and should be removed.

In addition, triangles $T_{A=0}$ with a degenerated representation T_{uvs} in Ξ_{uvs} are removed. This includes triangles with an area $A_{uvs} < 0.25$ pixels, as well as cases where T_{uvs} has a shortest edge of less than half a pixel and triangles that extend beyond the top and bottom corners of Ξ_{uvs}.

Further, skinny triangles $T_{\varphi<13}$ are discarded if they enclose at least one angle φ between any two edges $\boldsymbol{e}_a$, $\boldsymbol{e}_b$, and $\boldsymbol{e}_c$ that is smaller than 13 degrees, and if the lengths $|\boldsymbol{e}_c|$ and $|\boldsymbol{e}_b|$ of its longest two edges $\boldsymbol{e}_c$ and $\boldsymbol{e}_b$ conform to the following conditions:

$$(|\boldsymbol{e}_c| > \overline{|\boldsymbol{e}_{KNN}|} + 4\sigma|\boldsymbol{e}_{KNN}|) \wedge (\boldsymbol{e}_b| > \overline{|\boldsymbol{e}_{KNN}|} + 4\sigma|\boldsymbol{e}_{KNN}|) \tag{21}$$

$$(|\boldsymbol{e}_c| > \overline{|\boldsymbol{e}_{KNN}|} + 4\sigma|\boldsymbol{e}_{KNN}|) \wedge (\text{count}(\varphi < 13) == 2) \tag{22}$$

To compute the average length $\overline{|\boldsymbol{e}_{KNN}|}$ and standard deviation $\sigma|\boldsymbol{e}_{KNN}|$, only triangles T_{KNN} are considered that are formed by any three K-nearest neighbors $\boldsymbol{v}_{KNN}$ located within a radius of $\max(|\boldsymbol{e}_b| * 0.9, |\boldsymbol{e}_a|)$ around the tip vertex $\boldsymbol{v}_{\varphi<13}$ of $T_{\varphi<13}$ and the midpoint of its shortest edge $\boldsymbol{e}_a$. Additionally, any $T_{\varphi<13}$ that has to be discarded according to (21) will result in the deletion of all adjacent $T_{\varphi<13}$ connected to its $\boldsymbol{e}_b$ or $\boldsymbol{e}_c$. In addition, in the case of any $T_{\varphi<13}$ satisfying (22), only the $T_{\varphi<13}$ adjacent to $\boldsymbol{e}_c$ is removed. Finally, duplicate $\boldsymbol{vi}_i \equiv \boldsymbol{v}_j$ encoding the same point and $\boldsymbol{v}$ not referenced by any triangle are removed from the surface S, along with all small disconnected surface patches S_{dis}.

The surface S is stored on disk in **.obj** format, along with the corresponding texture information Ξ_{uvs}. Its triangle $\boldsymbol{n}_T$ and vertex normals $\boldsymbol{n}_v$ are recomputed, and a transformation $\mathfrak{R}$ is applied to all vertices and normals. The latter ensures that the $\boldsymbol{z}$-axis points in the direction of the patient's head and the positive $\boldsymbol{x}$-axis extends from the left to the right side of the torso. The origin point is selected such that it is located on the central viewing axis of the camera. To compute its y-component, the point cloud is divided into 3 sections along the vertical direction, roughly representing the chest, belly, and hips of the patient from top to bottom. The points within the top third are further split into 5 subgroups from right to left along the x-axis. For the rightmost and leftmost groups, the median coordinates $\hat{y}_r$ and $\hat{y}_l$ are computed. Based on these values, the final y-coordinate of the origin point is computed as $y = \hat{y}_r + \hat{y}_l/2$.

This ensures that all surfaces are located close to each other and that they partially overlap. At the same time, the actual relative shift between the surfaces and the angle at which the camera views the surface is retained as much as possible. This is crucial for the registration process described in Section 2.3.2.

2.3. Offline Processing

The electrode positions are computed using a set of at least 14 recordings of the torso surface, covering a minimum angle of approximately ≈270 degrees in the horizontal plane. The necessary steps, depicted in Figure 1, are presented in the following subsections. These steps include the pairwise alignment and registration of the recorded surfaces S, as described in Section 2.3.2; the extraction of the points $\boldsymbol{v}$ representing the colored electrode markers, as described in Section 2.3.3; and the fitting of a model of the marker to identify its central point, as described in Section 2.3.4. In the final step, a unique label is attached to each position, which uniquely links the individual ECG signals and the 3D position of the corresponding electrode.

2.3.1. Color Correction

The color sensor of the Intel Realsense SR300 camera (Intel corporation, Santa Clara, CA, USA) offers only a limited range (between $K_{W,min} = 2500$ and $K_{W,max} = 6500$) within which the color temperature parameter K_W can be tuned using the algorithm discussed in Section 2.2.1. This range is optimized for indoor use [21,22], where typical light sources include incandescent tungsten lamps ($K = 2500$), fluorescent lights ($K = 3800$), and standardized CIE sources such as CIE55 ($K = 5000$) or CIE65 ($K = 6500$).

The space limitations encountered in clinical settings, for example, outpatient and cardiology practitioner clinics, result in more challenging illumination conditions that can vary significantly depending on factors such as the patient's seating position or the camera's direction. Specifically, individual objects and parts of the room may be shaded by other objects, for example, the electrodes on the patient's back. Shaded areas are characterized by

color temperature values $K > 7000$, which are significantly larger than the $K_{W,max} = 6500$ upper limit assumed by the color sensor. Examples of this situation are shown in Figure 4a,c.

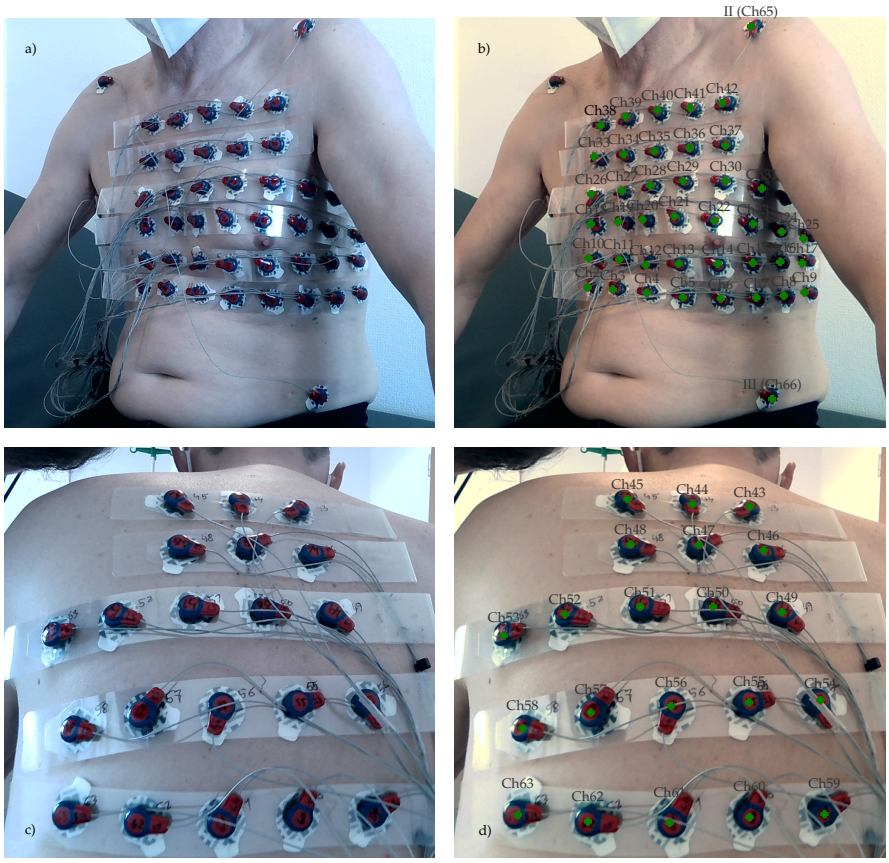

Figure 4. Examples for texture images recorded from the front (**a**) and back (**c**) of the torso, and the corresponding color-corrected versions (**b**,**d**). In (**b**,**d**), each visible electrode is labeled with the corresponding channel number. The '+' markers indicate the projected location of the computed electrode position.

An additional color-correction process is applied to the recorded texture images Ξ and the 3D surfaces. A virtual camera is used to simulate the recording of Ξ with a different K_W setting than the actual one. This virtual camera offers an AWB range between $K_{W,min} = 2000$ and $K_{W,max} = 9000$. It uses the model introduced in Section 2.2.1 to adjust the gain of its red $\Gamma_{\hat{R}} = \Gamma_K$ and blue $\Gamma_{\hat{B}} = 1/\Gamma_K$ color channels.

The virtual camera internally stores a linearized and normalized representation $\hat{\Xi}_=$ of Ξ_{uvs}. This representation corresponds to an image recorded with an equal gain $\Gamma_K = 1$ and $K_W = 5500$.

$$\Gamma_{K,=} = \frac{2(K_{W,uvs} - K_{W,min})}{K_{W,max} - K_{W,min}},\ \hat{R}_{i,=} = \frac{\hat{R}_i}{\Gamma_{K,=}},\ \hat{G}_{i,=} = \hat{G}_i,\ \hat{B}_{i,=} = \hat{B}_i\Gamma_{K,=},\ \Gamma_K = \Gamma_{K,=} \quad (23)$$

Its white-balancing parameter K_W is initialized to the color temperature K_{uvs} at which $\hat{\Xi}_{uvs}$ was recorded by the color sensor of the 3D DS camera.

After initialization, the color-correction approach described in Section 2.2.1 is used to adjust the K_W of the virtual camera until a suitable value for K_W^+ is found. If K_W^+ jitters around its ideal value for at least 20 repetitions, the color correction stops when the following condition is met:

$$-1 <= K_{n+1,W}^+ - K_{n,W}^+ <= 1 \quad (24)$$

In this case, K_W^+ is set to the mean value $\overline{K_W^+}$ of the last 3 minimum updates for which the difference between consecutive K_W^+ values is less than 10. With each update of K_W, a new version of Ξ_{uvs} is created by multiplying the red color values $\hat{R}_=$ of $\hat{\Xi}_=$ by the updated Γ_K^+, multiplying the blue values by $1/\Gamma_K^+$, and performing a left bisection search on the lookup table $\hat{V} = \lfloor V \rfloor$ established in Section 2.2.1. Pixels that are overexposed according to (4) are not modified. Pixels that appear overexposed after scaling and exceed a maximum value of 1 in at least one channel are assumed to be fully saturated in all three channels, which are each set to the maximum value. Pixels that appear underexposed, with at least one channel having a value less than 10^{-8}, are assumed to be unexposed in all channels. Therefore, in such cases, all three channels of the pixel are set to 1 when fully saturated and 0 when unexposed. Additionally, all channels are clipped to the maximum possible value of 1 if necessary. The color-optimized version of Ξ_{uvs} (Figure 4b,d) is then used to extract the 3D points of the electrode markers, as described in Section 2.3.3.

2.3.2. Surface Registration

To align the surfaces, a point-to-plane algorithm was chosen. This kind of ICP algorithm minimizes the distances $l = |\mathfrak{R}\boldsymbol{v}_S - \boldsymbol{v}_T|$ between corresponding $\boldsymbol{v}_S$ and $\boldsymbol{v}_T$ along the direction of the surface normals $\boldsymbol{n}_T$ of S_T.

$$E(\mathfrak{R}) = \sum_{(\boldsymbol{v}_S, \boldsymbol{v}_T) \in K} \|(\boldsymbol{v}_T - \mathfrak{R}\boldsymbol{v}_S)\boldsymbol{n}_T\|^2 = min \tag{25}$$

A precise alignment between S_T and S_S across all surface pairs is achieved when Equation (25) is also minimal in the reverse case with S_T and S_S swapped. The following simple symmetric point-to-plane approach is used by the registration component (Figure 1) to align the surfaces. It was chosen in favor of other symmetric point-to-plane algorithms such as [26], as it can be directly implemented using unidirectional ICP functions from open3D library [27]. In the first step, the forward transformation matrix $\mathfrak{R}_f$ is computed for the set of corresponding points $(\boldsymbol{v}_T, \boldsymbol{v}_S) \in C_f$ by applying (25). In the second step, the reverse transformation $\mathfrak{R}_R$ is computed for the points $(\boldsymbol{v}_S, \boldsymbol{v}_T) \in C_r$ corresponding to the reversed setup. The initial $\mathfrak{R}_{0,r}$ is initialized as $\mathfrak{R}^{-}1_f$. The set $(\boldsymbol{v}_T, \boldsymbol{v}_S) \in C_f$ is selected from a subset of $\boldsymbol{v}_S$ that is located within the maximum correspondence distance l_c of $\boldsymbol{v}_T$. The same selection criterion is used for the reverse set $(\boldsymbol{v}_S, \boldsymbol{v}_T) \in C_r$ with respect to any $\boldsymbol{v}_S$. In the final step, the optimal transformation $\mathfrak{R}$ and the new correspondence distance $l_{c,+1}$ are selected from $\mathfrak{R}_f$, $\mathfrak{R}_r$, $l_{c+1,f}$, and $l_{c+1,r}$ using the following criteria:

$$\begin{gathered}
\mathfrak{R},\, l_{c+1} = \begin{cases} \mathfrak{R}_r,\, l_{c+1,r} \text{ if } E(\mathfrak{R}_r) < \{E(\mathfrak{R}),\, E(\mathfrak{R}_f)\} \wedge l_{c+1,r} < \{l_c,\, l_{c+1,f}\} \\ \mathfrak{R}_f,\, l_{c+1,f} \text{ if } E(\mathfrak{R}_f) < \{E(\mathfrak{R}),\, E(\boldsymbol{R}\boldsymbol{e}_r)\} \wedge l_{c+1,f} < \{l_c,\, l_{c+1,r}\} \end{cases} \\
l_{c+1,f} = \bar{l}_f + 2\sigma l_f,\; l_{c+1,r} = \bar{l}_r + 2\sigma l_r \\
\bar{l}_f = \underset{(\boldsymbol{v}_T, \boldsymbol{v}_S) \in K_f}{\text{mean}} (|\boldsymbol{v}_S - \boldsymbol{v}_T|),\; \sigma l_f = \underset{(\boldsymbol{v}_T, \boldsymbol{v}_S) \in C_f}{\text{std}} (|\boldsymbol{v}_S - \boldsymbol{v}_T|) \\
\bar{l}_r = \underset{(\boldsymbol{v}_S, \boldsymbol{v}_T) \in C_f}{\text{mean}} (|\boldsymbol{v}_T - \boldsymbol{v}_S|),\; \sigma l_r = \underset{(\boldsymbol{v}_S, \boldsymbol{v}_T) \in C_r}{\text{std}} (|\boldsymbol{v}_T - \boldsymbol{v}_S|)
\end{gathered} \tag{26}$$

The surfaces S recorded using the approach described in Section 2.2 are aligned such that they more or less share the same space, apart from the small rotation $\Delta\varphi$ along the horizontal direction and the relative vertical movement Δz between the cameras. No information about their orientation in space or how much each pair overlaps is recorded. For obtaining sufficiently precise positions of the electrodes, the optimal correspondence distance l_o between any $(\boldsymbol{v}_T, \boldsymbol{v}_S)$ should be $l_o \lesssim 1\,\text{mm}$. Therefore, the symmetric ICP registration is repeated for each pair in multiple runs. The results obtained for $\mathfrak{R}$ and l_{c+1} in the previous run are used to initialize $\mathfrak{R}_0$ and l_c in the next run. If the condition in (26) for updating $\mathfrak{R}$ and l_c fails, one last run is attempted with $l_c = l_{min} \approx 1\,\text{mm}$ if $l_c < l_{min}$ and $l_{c-1} - l_c > \sigma l_0$ holds. For the first optimization run, $\mathfrak{R}_0$ is initialized to roughly reflect the

relative rotation about the z-axis between two recorded surfaces S_T and S_S and its relative shift Δz along the z-axis. The following approach is used to estimate the relative rotation angle $\Delta\varphi$ between S_T and S_S:

$$\Delta\varphi = \arccos\left(\frac{\langle \hat{\boldsymbol{v}}_{l,S} - \hat{\boldsymbol{v}}_{r,S} | \hat{\boldsymbol{v}}_{l,T} - \hat{\boldsymbol{v}}_{r,T} \rangle}{|\boldsymbol{h}_S||\boldsymbol{h}_T|}\right) \tag{27}$$

The right $\hat{\boldsymbol{v}}_{r,S}$, $\hat{\boldsymbol{v}}_{r,T}$ and left $\hat{\boldsymbol{v}}_{l,T}$, $\hat{\boldsymbol{v}}_{r,T}$ median points define the horizontal directions of the sagittal planes with respect to S_S and S_T. They are computed using the same approach described in Section 2.2.4 to define the final position of the origin along the y-coordinate.

Suitable estimates for $l_{c,max}$, $l_{c,min}$, and σl_0 are essential for achieving a sufficiently precise alignment of S_T and S_S. When testing the implementation of the symmetric ICP, it was empirically found that the values for $l_{c,max}$, in particular, varied significantly depending on the relative distance and angle between two consecutive surfaces. Initially, constant values were assigned to $l_{c,max}$ and $l_{c,min}$. However, these values resulted in an insufficient alignment between the surfaces on average. Specifically, the alignment of the surfaces at the left side where the front and back sides of the torso meet was rather challenging, and in some cases, not possible at all.

In order to improve the results and ensure a proper alignment between the surfaces, the following approach is used to determine suitable estimates of $l_{c,max}$, $l_{c,min}$, and σl_0 for each pair of S_T and S_S. These estimates are computed based on the distances between the vertices $\boldsymbol{v}_T$ and $\boldsymbol{v}'_S$ within the volume $V_{T\cap S'} = V_T \cap V'_S$, which represents the common region of the axis-aligned bounding boxes V_T and V'_S encompassing the target surface S_T and the source surface S'_S. The latter S'_S is obtained by applying an initial transformation $\mathfrak{R}_0$ to the source surface S_S. The transformation $\mathfrak{R}_0$ shifts all $\boldsymbol{v}'_S \in V_{T\cap S'}$ such that their center of mass $\hat{\boldsymbol{v}'_S}$ aligns with the center of mass $\hat{\boldsymbol{v}_T}$ of all $\boldsymbol{v}_S \in V_{T\cap S'}$. The value for l_{max} is obtained by applying (26) to the distances between the points in the forward correspondence set $(\boldsymbol{v}_T, \boldsymbol{v}'_S) \in C_{f,0}$ and the backward correspondence set $(\boldsymbol{v}_T, \boldsymbol{v}'_S) \in C_{r,0}$. Both sets are found through a KNN search [28,29], which also considers the surface normals $\boldsymbol{n}_T$ and $\boldsymbol{n}'_S$ in each $\boldsymbol{v}_T$ and $\boldsymbol{v}'_S$. This approach has the advantage of considering only $\boldsymbol{v}_T$ and $\boldsymbol{v}'_S$ as corresponding when their surface normals $\boldsymbol{n}_T$ and $\boldsymbol{n}'_S$ are closely aligned. From the resulting $C_{f,0}$ and $C_{r,0}$, any $\boldsymbol{v}_T$ and $\boldsymbol{v}'_S$ are removed if the deviation between their surface normals $\boldsymbol{n}'_S$ and $\boldsymbol{n}_T$ exceeds 30 degrees, ensuring that $\langle \boldsymbol{n}'_S | \boldsymbol{n}_T \rangle < 0.98$.

The estimate for l_{min} is based on the overall mean$(|\boldsymbol{v}_T|, |\boldsymbol{v}'_S|)$ of the shortest neighbor distances within all $\boldsymbol{v}_T \in V_{T\cap S'}$ and $\boldsymbol{v}'_S \in V_{T\cap S'}$.

From the final $\mathfrak{R}$ of all consecutive pairs of S_T and S_S, the global alignment of each S_i is determined by the cumulative transformation $\mathfrak{R} = \prod_j^i \mathfrak{R}_j$, starting with identity $\boldsymbol{Re} = \boldsymbol{I}$ for the first surface S_1. Alternatively, the transformation of the first surface can be initialized by the horizontal camera inclination angle φ about the z-axis using (27). From this, the relative angle between S_1 and the x-axis of the patient's frontal plane is computed. This already provides a rough alignment of the resulting point cloud of the torso with its frontal plane.

2.3.3. Electrode Marker Extraction

In the current setup, the electrodes are attached to g.LADYbird™ active electrode clips from g.tec medical engineering GmbH, Schiedlberg, Austria. These clips have a circular head, with its center aligned with the center of the electrode. The clip itself is covered with red-colored epoxy to protect the integrated electronics from water and other liquids. The circumference of the head is painted blue to model a circular electrode marker with a blue boundary and a red central disk. Figure 5 shows an example of this basic setup.

The blue boundary color (see Figure 5b) is selected such that the electrode marker easily can be detected within the RGB chromacity space representations $\hat{\chi}_{rg}$ of the surface texture images Ξ_{uvs}. The $\hat{\chi}_{rg}$ values are obtained as a byproduct of the white-balancing and light color–temperature correction approaches described in Section 2.3.1.

Each $\hat{\chi}_{rg}$ is scanned for red $\boldsymbol{x}_r = r_r, g_r, b_r$ and blue $\boldsymbol{x}_b = r_b, g_b, b_b$ pixels that are fully described by one of the following two ellipses within the RGB chromacity space.

$$\frac{(\delta r_r \cos(\phi_r) - \delta g_r \sin(\phi_r))^2}{\sigma(r_r)} + \frac{(\delta r_r \sin(\phi_r) + \delta g_r \cos(\phi_r))^2}{\sigma(g_r)} <= 1 \quad \text{with } \delta r_r = r_r - \overline{r_r},\ \delta g_r = g_r - \overline{g_r} \tag{28}$$

$$\frac{(\delta r_b \cos(\phi_b) - \delta g_b \sin(\phi_b))^2}{\sigma(r_b)} + \frac{(\delta r_b \sin(\phi_b) + \delta g_b \cos(\phi_b))^2}{\sigma(g_b)} <= 1 \quad \text{with } \delta r_b = r_b - \overline{r_b},\ \delta g_b = g_b - \overline{g_b} \tag{29}$$

The values $\overline{r_r}$, $\overline{g_r}$, $\overline{r_b}$, ϕ_r, $\overline{g_b}$, and ϕ_b define the red and green coordinates of the center point of the ellipsis and the rotation angle by which each of them is rotated with respect to the red axis of the RGB chromacity space. Their values are determined through the calibration procedure described in Section 2.3.6. All matching $\boldsymbol{x}_r$ and $\boldsymbol{x}_b$ pixels are mapped to their corresponding 3D vertices $\boldsymbol{v}_r$ and $\boldsymbol{v}_b$ on the torso surface S. This mapping is accomplished by computing the barycentric coordinates of each $\boldsymbol{x}_r$ and $\boldsymbol{x}_b$ within the representation of the surface triangle T in Ξ_{uvs}.

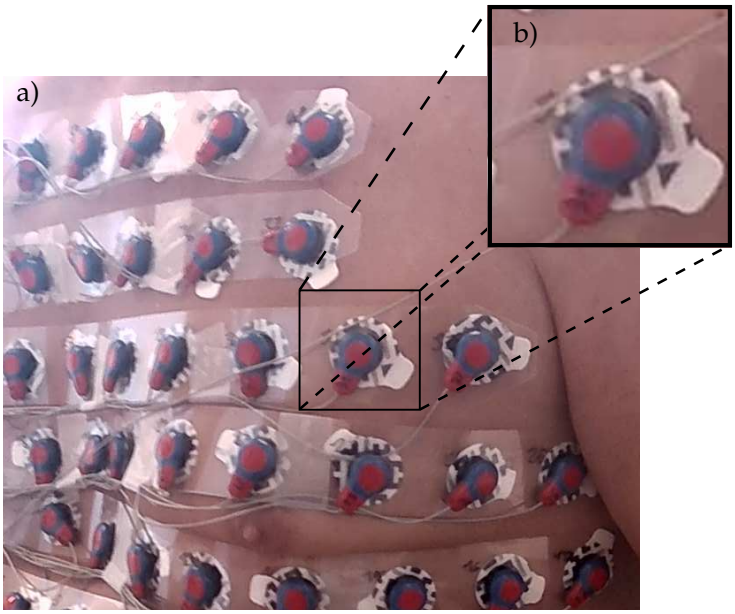

Figure 5. The electrodes mounted on the patient's torso (**a**) are attached to the red electrode clips. The blue boundary of each clip head (**b**) forms a circular marker with the red electrode clip. Based on the red and blue colors, each marker can be recognized from the recorded texture images along with 3D surface information.

The resulting marker point cloud P_M formed by all $\boldsymbol{v}_r$ and $\boldsymbol{v}_b$ is filtered with respect to $\boldsymbol{v}_r$ and $\boldsymbol{v}_b$, which likely correspond to a valid electrode marker, as defined by the color of the clip head. This is achieved by a radius-based KNN search for at least one neighbor of the opposite color. The radius is set to the radius of the clip head for all $\boldsymbol{v}_r$ and the width of the blue boundary ring for all $\boldsymbol{v}_b$. If the neighborhood of radius ρ does not contain any points of the opposite color, $\boldsymbol{v}$ is removed from P_M.

The filtered P_M is split into individual clusters of $\boldsymbol{v}_{el} \in \boldsymbol{v}_r \cup \boldsymbol{v}_b$, representing the individual electrode clips. This is accomplished by applying the HDBSCAN algorithm [30]. The results are more robust compared to the basic DBSCAN algorithm [31], especially in the presence of groups of outliers, for example, generated by a bluish shadow cast on the cables and electrode clips. In addition, a minimum distance ϵ_{split} can be defined, and clusters are not split any further. In contrast to the basic DBSCAN [31] algorithm, ϵ_{split} defines a lower

boundary limit rather than a strict cutting distance. In other words, less dense clusters with an average density exceeding ϵ_{split} are not necessarily forced to split into distinct leaf clusters. The parameters of the minimum cluster size $N_{C,min}$ and the minimum samples $X_{C,min} = 20$ are used to fine-tune and control the extraction of clusters that represent the individual electrode markers, considering the actual number of electrodes N_{el}.

$$\epsilon_{split} = r_{nh} \tag{30}$$

$$N_{C,min} = \max(\frac{\text{count}(\boldsymbol{v}_r)}{4N_{clip}}, X_{C,min}) \tag{31}$$

In order to simplify the subsequent processing steps, the overall point cloud P_S, as well as PM, is realigned such that the frontal plane of the torso is in line with the x-z plane of the coordinate system. This is achieved by once again splitting P_S into chest, belly, and hip sections. The points of the chest section are further split along the x-axis into three parts, representing the right shoulder, neck, and left shoulder. The final transformation $\mathfrak{R}$ is computed by aligning the vector between the median points of the left and right shoulders to the x-axis of the frontal plane.

2.3.4. Fitting Marker Model

The red points $\boldsymbol{v}_r$ and blue points $\boldsymbol{v}_b$ within each cluster are fitted to a planar marker model consisting of a red disk enclosed within a blue ring. Before fitting, all $\boldsymbol{v}_r$ and $\boldsymbol{v}_b$ are projected onto the plane Q_{cl}, which is parallel to all $\boldsymbol{v}_{el}$.

$$\boldsymbol{v}'_e l = \boldsymbol{v}_{el} - \langle \boldsymbol{v}_{el} - \overline{\boldsymbol{v}}_{el},\ \boldsymbol{n}_{mj} \rangle \boldsymbol{n}_{mj} \tag{32}$$

This ensures that all $\boldsymbol{v}_{el}$ are located on Q_{cl}, which is defined by the predominant surface normal vector direction $\boldsymbol{n}_{mj}$ within all surface normal vectors $\boldsymbol{n}_{el}$ $\boldsymbol{v}_{el}$ and their center of mass $\overline{\boldsymbol{v}}_{el}$.

$$\boldsymbol{n}_{mj} = \sum_i^r \boldsymbol{U}_i \sigma_i \boldsymbol{V}_i \text{ with } \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V} = \text{svd}(\boldsymbol{n}_G(\boldsymbol{v}_{el})) \tag{33}$$

$$r = i_{max} \underset{1 \le i \le K}{\Longrightarrow} -10 \log_{10} \left(\frac{\sum_j^i \sigma_j}{\sum_k^K \sigma_k} \right) > 3$$

The shifted $\boldsymbol{v}'_{el}$ are then fitted to the following model, which is based on the distances $\rho(\boldsymbol{v}'_{el})$ between the individual $\boldsymbol{v}'_{el})$ and the electrode center $\boldsymbol{X}$ on Q_{el}.

$$\rho(\boldsymbol{v}'_{el}) = |\boldsymbol{v}'_{el} - \boldsymbol{X}|,\ \delta_\rho(\boldsymbol{b}_b) = \rho(\boldsymbol{v}_b) - \rho_{disc})$$
$$\epsilon = \sum \delta\rho^2(\boldsymbol{v}_b) + \sum \rho^2(\{\boldsymbol{v}'_r | \rho(\boldsymbol{v}'_r) < \rho_{disc}\}) + \sum \rho^2(\boldsymbol{v}'_b) + \langle \boldsymbol{X} - \overline{\boldsymbol{v}}_{el}, \boldsymbol{n}_{mj} \rangle^2 \tag{34}$$

$\boldsymbol{\delta}_\rho$ represents the relative distances of the blue points $\boldsymbol{v}'_b$ from the boundary of the enclosed red disk with a radius ρ_{disc}. From all the red points $\boldsymbol{v}_r$ within a cluster, the model selects those that are within a radius $\rho_r < \rho_{disc}$ from the current $\boldsymbol{X}$. The model in (34) is optimized with respect to $\boldsymbol{X}$ using the L-BFGS-B algorithm provided by the SciPy minimize function. This numerically robust algorithm was selected because it can achieve satisfactory optimization results for least-squares optimization problems. Its implementation details can be found in the SciPy manual and [32,33]. For all clusters for which an $\epsilon_{min}(X)$ could be found, $\boldsymbol{X}_{min}$ is stored, along with $\boldsymbol{n}_{mj}$. Any remaining clusters for which no appropriate $\boldsymbol{X}_{min}$ could be found are not further considered.

In some cases, it is possible that a clip is split into two smaller clusters. For example, if an electrode array is carelessly attached to the torso, electrode leads can shadow the relevant parts of the clip head. This might be the case when the following condition holds with respect to the counts of $\boldsymbol{v}_r$ or $\boldsymbol{v}_b$:

$$\frac{1}{3} < \frac{\text{count}(\boldsymbol{v}_r)}{\text{count}(\boldsymbol{v}_b)} < 3 \tag{35}$$

Two neighboring clusters are considered pieces of the same marker only if at least 10 closest neighbors of any $\boldsymbol{v}_{el}$ in the first cluster are closest to at least 85 distinct $\boldsymbol{v}_{el}$ in the other cluster. The cylindrical model is fitted to the largest piece of the marker only. This prevents nearby image artifacts in the Ξ_{uvs} from causing misalignment of the affected electrode marker and distracting the center point from its true location.

The identified cluster centers $\boldsymbol{X}$ are triangulated using the ball-pivoting method [34,35] implemented in the open3D library. The radii $\rho_1 = \overline{x}/\sqrt{2}$ and $\rho_2 = \overline{x}$ for two distinct balls are derived from the average distance $\overline{x} = \text{mean}(|\boldsymbol{X}_9 - \boldsymbol{X}|)$ between each $\boldsymbol{X}$ and its 9 closest neighbors. Outliers are removed if $|\boldsymbol{X}_9 - \boldsymbol{X}| > \overline{x} + 2 * \sigma(|\boldsymbol{X}_9 - \boldsymbol{X}|)$ before computing $\overline{x}$. For a final check to determine if the $\boldsymbol{X}$ of neighboring clusters resemble two pieces of the same marker, the surface connectivity between individual $\boldsymbol{X}$ is computed. The marker attached to the largest group, where two $\boldsymbol{X}$ for which $|\boldsymbol{X}_1 - \boldsymbol{X}_2| < {}^2/_3\rho_0$ holds, is retained, whereas the other is removed. Ball-pivoting triangulation and the removal of small clip pieces are repeated until no more nearby groups, represented by distinct $\boldsymbol{X}$, are found. The remaining $\boldsymbol{X}$ that are included in the resulting triangular surfaces represent the frontal and dorsal patches of the electrode grid layout proposed in [36]. Clusters that are too far away to be included in the mesh by the ball-pivoting process are considered single electrodes, similar to those used, for example, in Einthoven I, II, and III.

In the final step, the triangular meshes of the frontal and dorsal electrode patches are normalized. In this process, any vertical edge that intersects the horizontal line between two common neighbors of its endpoints is swapped with the edge that connects the common neighbors.

2.3.5. Label Assignment

Starting from the point with the smallest y-coordinate, the triangulation of the frontal patch is scanned line by line. All electrodes that can be connected along consecutive horizontal edges are joined into one row of the frontal patch [36] and stored in right-to-left order. The rows are ordered from bottom to top. After all rows of the frontal patch have been collected, the same approach is applied to collect the electrodes of the dorsal patch. Again, the electrodes are stored in right-to-left and bottom-to-top order.

On the frontal patch, the number labels for each channel are assigned in ascending order from bottom right to top left. The dorsal assignment starts at the top right and ends at the bottom left. The remaining electrode points $\boldsymbol{X}$ that have not been included within the triangulation of the frontal and dorsal patches either correspond to the three Einthoven leads I, II, and III if they are located on the arms close to the front of the left and right shoulders and on the left hip. The electrode array includes two additional electrodes that are placed frontal and dorsal close to the right side of the torso.

2.3.6. Calibration

The proposed method to identify the color electrode markers requires proper calibration of the mean values $\overline{r_r}$, $\overline{g_r}$, $\overline{r_b}$, and $\overline{g_b}$; the standard deviations $\sigma(r_r)$, $\sigma(g_r)$, $\sigma(r_b)$, and $\sigma(g_b)$; and the rotation angles ϕ_r and ϕ_b of the ellipses in Equations (28) and (29). In the first step, the color-corrected chromacity representation $\hat{\chi}_{rg}$ of the texture images Ξ_{uvs} obtained as a byproduct in Section 2.3.1) is roughly segmented. The pixels representing a blue or red pixel of the clips are initialized with the following values: $\overline{r_r} = 0.75$, $\sigma(r_r) = 0.1$, $\overline{g_r} = 0.08$, $\sigma(g_r) = 0.06$, $\overline{r_b} = 0.05$, $\sigma(r_b) = 0.02$, $\overline{g_b} = 0.13$, $\sigma(g_b) = 0.06$, and $\phi_r = \phi_b = 0$.

These values were empirically identified from the chromacity space triangle of the 3D DS camera's color sensor, generated from the pixels of all $\hat{\chi}_{rg}$. The resulting raw pixel masks $M_{\hat{\chi},raw}$ are stored along with the corresponding $\hat{\chi}_{rg}$ obtained from the data sets of at least three patients. In addition, a binary mask M_I selects pixels of Ξ_{uvs} that are properly exposed according to (4). For storing the $\hat{chi}_{rg}$ on disk, the 16-bit PNG format is used. They

are loaded along with the corresponding $M_{\hat{\chi},raw}$ in an image processing program such as Gimp™ or Adobe Photoshop™ for manual segmentation of the clips.

The resulting $M_{\hat{\chi}}$, created by manually removing any pixel that does not represent a clip or electrode marker from $M_{\hat{\chi},raw}$, is used in combination with M_I to extract the pixels that are part of the electrode clips and markers visible on each 16-bit χ_{rg} image. Any pixel that does not correspond to a clip, is over- or underexposed, or meets the condition in (3) is not further considered in the following calibration steps. From all other pixel values, a 2D heat map N_H with 256 bins for red r and green g chromacity values each is generated and median-filtered using a 7 by 7 neighborhood.

The red and blue color shades of the electrode markers appear as distinct, Gaussian-shaped peaks P_H on N_H. They (1, 2) are clearly visible as bright spots on the heat map, as shown in Figure 6. A Gaussian mixture model [37,38] is used to extract the individual clusters C_H that represent each peak. Each peak is described as a 2D Gaussian distribution, which can be characterized by its center point or centroid and the standard deviations along each direction with respect to this center. By fitting the individual Gaussian models to the heat map $N - H$, the actual position, orientation, and area covered by each peak can be found. To compute the initial positions of the cluster centroids, the heat map is binarized and labeled. In this process, any 4-connected set of at least 5 bins is considered a peak if all bin counts n_H conform to the following condition:

$$n_H > |\boldsymbol{n_H}| + 1.9\sigma(\boldsymbol{n_H}) \text{ with } \boldsymbol{n_H} = \{n_H | n_H > 0\} \tag{36}$$

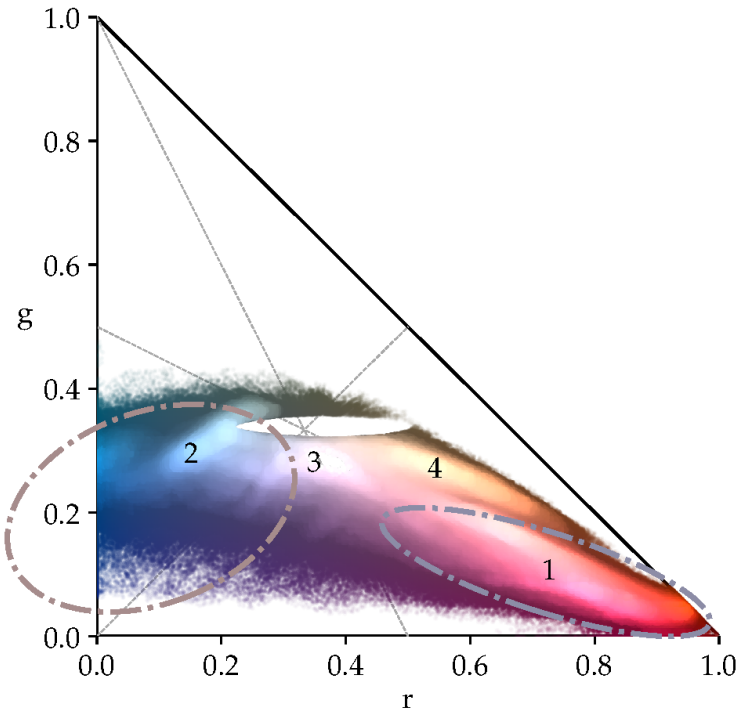

Figure 6. Chromacity heat map of the pixels representing the electrode markers created from the texture images of three patients. The brighter the color, the higher the pixel count for the corresponding point in the RGB chromacity space, represented by its red r and green g chromacity values. For better readability, the RGB chromacity values are displayed in RGB gamma-compressed form. The Gaussian peaks (dash doted ellipses) representing the red (1) and blue (2) pixels of the electrode markers are clearly visible. They can easily be distinguished from the peak (3) representing the color highlights and reflections. Peak (4) is caused by inappropriately chosen values for the parameters required to convert raw color sensor data to the RGB color space.

The cluster $C_{H,r}$ with the highest mean $\overline{r_r}$ red component is used to compute $\sigma(r_r)$, $\sigma(g_r)$, and ϕ_r. The values for $\sigma(r_b)$, $\sigma(g_b)$, and ϕ_b are derived from $C_{H,b}$ for which $1 - \overline{r_b} - \overline{g_b} = max$ holds.

$$\begin{array}{cc} \Sigma_r\, \boldsymbol{U}_r = \mathrm{eig}(\mathrm{cov}(C_{H,r})) & \Sigma_b\, \boldsymbol{U}_b = \mathrm{eig}(\mathrm{cov}(C_{H,b})) \\ \sigma(r_r) = \sqrt{\sigma_{1,r}},\ \sigma(g_r) = \sqrt{\sigma_{2,r}},\ \phi_r = \dfrac{\langle \boldsymbol{u}_{1,r} | \boldsymbol{r} \rangle}{|\boldsymbol{u}_{1,r}|} & \sigma(r_b) = \sqrt{\sigma_{1,b}},\ \sigma(g_b) = \sqrt{\sigma_{2,b}},\ \phi_b = \dfrac{\langle \boldsymbol{u}_{1,b} | \boldsymbol{r} \rangle}{|\boldsymbol{u}_{1,r}|} \end{array} \tag{37}$$

$\sigma_{1,r}$, $\sigma_{2,r}$, $\sigma_{1,b}$, and $\sigma_{2,b}$ represent the first and second eigenvalues Σ of the covariance matrices $\mathrm{cov}(C_{H,r})$, $\mathrm{cov}(C_{H,b})$ of $C_{H,r}$ and $C_{H,b}$, and $\boldsymbol{u}_{1,r}$ and $\boldsymbol{u}_{1,b}$ are the corresponding initial eigenvectors. These values are stored along with the centroids of $C_{H,r}$ and C_H, which define the mean values $\overline{r_r}$, $\overline{g_r}$, $\overline{r_b}$, and $\overline{g_b}$ on disk that are to be used in the extraction step described in Section 2.3.3.

The remaining clusters 3 and 4 are not further considered as they correspond to the color highlights on the clips (3) or are caused by inappropriately chosen parameters affecting the conversion of the raw sensor signals to the RGB color space (4).

2.4. Recording Protocol

The technical approach outlined in Sections 2.2 and 2.3, requires that the patient maintains the same posture throughout the recording. This is only possible if the patient is directly engaged and actively participating in the measurement.

Therefore, prior to the application of the electrodes, the patient is instructed to sit down on a chair. The height of the chair is then adjusted so the patient can comfortably sit upright throughout the recording process. The feet of the patient should rest flat on the floor and the knees should be bent by no more than 90 degrees. If the chair cannot be adjusted in height, an alternative solution is to stack multiple chairs to increase the patient's comfort and encourage them to straighten their back. To ensure optimal recordings without any obstacles, the chair should not have armrests or a backrest and be placed at least 1 meter from any furniture or other objects that can cause shadows. This ensures that the FOV of the 3D DS camera can be optimally used and the operator is able to capture a surface at least every 20 degrees.

After the electrodes have been attached to the torso, the patient is instructed to place the hands on the thighs. The fingers should point inward and the thumbs should point straight toward the hips. The optimal position of the hands is a thumb length before the hips. While the electrode positions are recorded, the patient is instructed to maintain a straight and upright back. Most patients are able to easily maintain this position by slightly straightening their elbows (about 120 degrees between the upper and lower arm). This helps them to move their chest and shoulders into a position that is as upright as possible. This has the effect that the patient is forced into an isometric posture, which can easily be maintained while the electrode positions are recorded. In addition, this position facilitates the recording of electrodes placed under the left axle, for example, the Wilson electrodes V_5 and V_6.

3. Results

In the following section, the results are presented.

The narrow vertical field of view of the color sensor is one of the main reasons why the 3D images of the torso are recorded in portrait mode. In a typical clinical setting, where space is limited, it is likely that the patient is seated close to furniture or walls. For proper recording of the 3D images, a space of at least 2.5 m by 2.5 m is required. This includes a standard chair without armrests or a backrest, with a diameter of 50 cm, that can provide at least 1 m of space on all four sides of the patient for the operator to move around while recording the images. The remaining space between the patient, the operator, and any surrounding furniture or walls should be 50 cm or less. Both sensors of the camera must be able to properly capture the dorsal part of the patient's torso at distances between 20 cm

and 50 cm. This can only be achieved by cameras with FOV angles conforming to (1) such as the Intel Realsense cameras, which have wide viewing angles of ≈70 degrees for both the depth and color sensors when used in portrait recording mode. This is especially important for capturing the dorsal views of the torso.

The color sensor has a viewing ratio of 16:9 between the horizontal and vertical FOVs. This results in a vertical viewing angle of about 40 degrees, which is a lot smaller than the ≈60 degrees of the depth sensor. This can lead to a situation where, for example, around ≈60 columns on the top and bottom of the depth image lack texture information. However, this is acceptable given that consecutive 3D images are recorded in portrait mode with an overlap of about two-thirds, ensuring that the texture images sufficiently overlap.

Thanks to the vertical nature of the patient's torso, in portrait mode, it is easy to keep the patient centered in the image while moving the camera to the next recording position. As the patient's torso covers most of the image space, only very few objects and obstacles located behind the patient are captured by the cameras, which can easily be removed before storing the 3D surface images.

Scanning always starts with the right frontal view of the torso and ends at the right dorsal side. If possible, the right lateral side of the torso can be recorded. This is not essential for extracting the electrode positions and can be omitted in standard recording procedures. It is recommended to explicitly record the right lateral torso surface when there is sufficient space to the right of the patient.

The preview image of the torso, shown in the main area (1) of the user interface shown in Figure 7, is split into a 3-by-3 grid. The center segment of this grid is used as the focus area, representing the central part of the patient's torso. The contours of the largest object containing the focus segment are highlighted in orange. As the camera points at the patient's torso, the contours highlight the boundaries of the patient's torso. The recording of a torso surface segment is initiated by pressing the trigger of the camera. The color of the contour line switches to green and the live preview freezes to indicate that the captured depth and color image have been processed and the 3D surface has been generated and stored. Once the underlying point cloud has been triangulated, occluded and degenerated triangles, as well as detached surface patches, are removed. Then, the contour is updated to mark the parts that will be stored on disk. After the 3D surface information, corresponding texture image, and meta information have been stored, the live preview is started again and the color of the contour reverts to orange. The live preview is updated at a maximum rate of 10 FPS. With the Python-based prototype, update rates between ≈4 FPS and ≈7 FPS can be realistically achieved.

The main preview area (panel 1 in Figure 7) has the same shape as the depth image. For the parts on the left and right sides that are not captured by the RGB image, the edges identified on the depth image are displayed instead. The outline of the patient's torso does not extend beyond the edges of the RGB image. In panel 2 of the preview screen (Figure 7), several recording and camera parameters, such as the frame rate in FPS, exposure time τ in ms, etc., are shown, along with the intermediate parameters computed for automatic-exposure control and color correction. In panel 3, the full set of edges identified on the current depth image is displayed. The two vertical lines delineate the area of the depth image that is covered by the color image.

The prototype for real-time recording of the 3D torso surface patches, as well as for postprocessing and calibration, was implemented in Python version 3 using a recent version of NumPy and SciPy [39]. The librealsens version 2 library [40] was used to control the acquisition, convert the depth values into a point cloud, and compute the corresponding texture uvs map for the RGB image. The OpenCV library [41] was used to generate the preview display, and the generation and cleaning of the 3D meshes were accomplished using the open3D library [27]. The most computationally demanding components, the depth-edge detection (Section 2.2.3), automatic white balancing (Section 2.2.1), and patient-locked auto-exposure control (Section 2.2.2), were converted into Python-C modules using Cython [42].

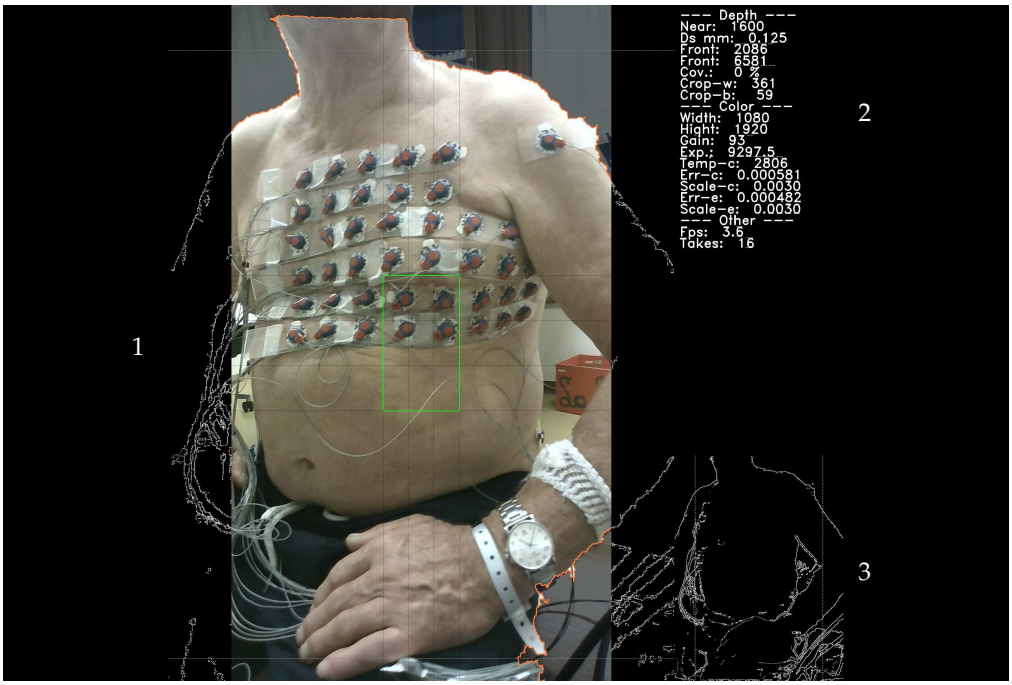

Figure 7. The preview screen is divided into three panels. The main panel (1) displays the image recorded by the color sensor of the 3D DS camera. The parts of the 3D image for which no color information could be captured are replaced with the edges extracted from the depth image shown in panel 3. The current values of the color temperature, exposure time, frames per second, and other process parameters are displayed in panel 2. The two vertical lines indicate the area where the views of both cameras overlap.

In total, five male subjects between 38 and 70 years of age participated in the present study. Each subject was seated on a chair or examination bed, depending on the available space. After applying the ECG electrodes to the chest and back, the subjects were instructed to maintain the posture described in Section 2.4. The measurement of the torso surface and the recording of a 30 -min long ECG with 67 channels took about 30 min to 45 min. After each measurement, the data were analyzed and the prototype improved accordingly.

The data set recorded from the first subject turned out to be quite limited and, therefore, is not included in the presented results, as it was affected by the automatic white balancing and exposure control of the color sensor, which could not cope well with the diverse and complex lighting conditions. Further, the 3D points recorded by the depth sensor were directly transformed to match the color image captured by the color sensor. This posed several challenges related to occluded surface parts causing undesirable distortions and the introduction of noncausal surfaces. Starting with the data for the second subject, the direct mapping was replaced with the texture mapping approach, which yielded better results and allowed for the implementation of the algorithms for occlusion management and the removal of noncausal triangles, as described in Section 2.2.4.

For each patient, 12 to 15 views were recorded. Each of the views contained a 3D surface described by ≈170,000 vertices and ≈300,000 triangles. As shown in Table 1, between 7 and 21 iterations of the symmetric ICP algorithm were necessary to align the surfaces. The maximum correspondence distance between the points of the surface pairs was reduced in every iteration step, starting from 7 cm–12 cm and reaching 0.7 mm–1.2 mm. More iterations were necessary to align the surfaces joining the frontal and dorsal views on the left side of the torso. In cases where the available space around the subject was insufficient, the number of iterations required to align the surfaces was increased. In the most challenging scenario, the proper alignment of the surfaces was not possible at all. This situation was encountered in the data set recorded from subject 5, where part of the torso surface on the

left side was obscured by the backrest of the chair. Among other challenges, this required an increased number of 21 iterations to align the leftmost frontal and dorsal views.

Table 1. The symmetric ICP alignment metrics obtained from the data sets of four out of five subjects participating in the study. For an average of 14 angular views, the computation of the pairwise transformation matrix was repeated between 7 and 21 times, with an average repetition rate of 9.7 repeats per surface pair. The average initial distance between corresponding points was 7 cm, the average root mean square error was 0.7 mm, and the average final correspondence distance was 1 mm.

Subj.	# Views	# $\frac{repeats}{pair}$			Correspondence Distance (mm)						Final Rmse (mm)		
					Initial			Final					
		Min.	Mean	Max.	Min.	Mean	Max.	Min.	Mean	Max.	Min.	Mean	Max.
2	12	9	10.1	12	37.4	71.6	124.9	0.7	0.9	1.2	0.5	0.6	0.8
3	14	7	9.4	17	51.4	76.1	105.0	0.9	1.0	1.1	0.6	0.7	0.7
4	14	8	9.4	17	42.9	61.2	105.2	0.9	1.0	1.1	0.6	0.7	0.8
5	15	8	9.9	21	50.0	71.3	117.2	0.9	1.0	1.1	0.6	0.7	0.7
Mean	14	7	9.7	21	37.4	70.0	124.9	0.7	1.0	1.2	0.5	0.7	0.8

Across all subjects, a final root mean square error between consecutive surfaces of 0.7 mm was achieved. Using the proposed approach, 12 to 15 surfaces per patient were registered within 13 min. As shown in Table 2, the extraction of the electrode marker points and the computation and labeling of the electrode positions were completed after another ≈8 min.

Table 2. Performance measures obtained from four out of five subjects participating in the study. The positions mounted on a patient's torso were extracted from 14 recorded angular views, on average, within ≈21.8 min. The symmetric ICP-based pairwise alignment of the corresponding surfaces of about 139,527 vertices and 254,351 triangles took approximately ≈2/3 of this time.

Subj.	# Views	# Vertices			# Triangles			Time (min.)		
		Min.	Mean	Max.	Min.	Mean	Max.	ICP	Pos.	Total
2	12	63,204	101,446	132,813	112,547	178,654	235,112	6.58	5.75	12.51
3	14	131,090	198,732	232,417	252,514	380,923	448,766	17.75	9.94	28.28
4	14	155,025	189,075	223,063	292,232	357,555	426,261	14.62	7.78	22.94
5	15	172,548	203,327	241,228	322,431	384,722	458,542	15.24	7.79	23.48
Mean	14	63,204	176,301	241,228	112,547	331,879	458,542	13.55	7.82	21.80

The recording sessions were part of a larger clinical pilot study investigating the prognostic value of index arrhythmias with respect to the outcome of pulmonary vein ablation, for which the participants provided informed consent. Apart from the 3D camera and ECG recordings, this study was based on clinical data recorded during the patient's clinical treatment. Therefore, CT recordings and other independent means of recording the electrode positions relative to the torso were included. To assess the accuracy of electrode localization, the electrode positions were backprojected onto the individual views of the torso and marked on the corresponding color images. Examples are shown in Figures 3, 4b,c and 8.

The annotated RGB images were presented to an expert who used the cross-hair tool shown in Figure 8b to manually adjust the position of each marker. In order to facilitate this task, two markers were used: the green marker indicates the backprojected position of the marker and the red marker corresponds to the manually adjusted position. All positions were checked during this process and if necessary, they were moved to better reflect the perceived center positions on each view. When finished, all positions were reprojected onto a 3D space.

For the set of corrected positions of each electrode, the mean point, as well as the mean distance and standard deviation to this mean point, were computed. The resulting values are shown in Table 3, along with the mean and standard deviations of the computed electrode positions with respect to the manually determined mean. Both sets of results were influenced by the accuracy of the registration process and the fact that no unique solution exists for the backprojection of the electrode positions onto the individual views. In addition, the mean and standard deviation of the registration errors and the error between the mean and standard deviation of the distances between the individual projections and their mean point are listed.

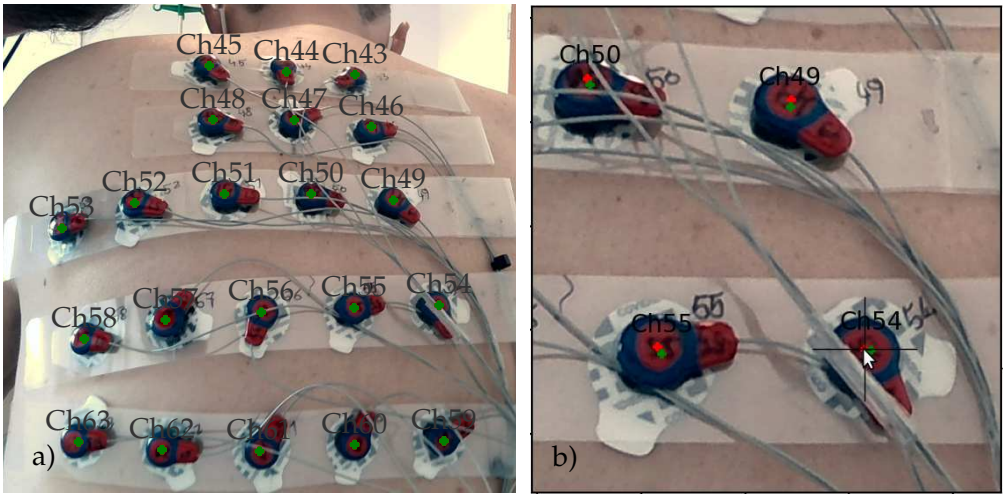

Figure 8. Manual evaluation of the proposed approach for locating the electrode positions on a patient's torso. The electrode positions are backprojected onto each recorded torso surface segment (**a**). An electrode position can be moved by clicking on the corresponding green cross-shaped graphical marker displayed on the texture image. Its new position is selected by pointing and clicking on it (**b**). In case the position pointed to is not backed by a valid surface triangle, the new point (red cross) is moved to the closest possible position. By right-clicking on an electrode marker, it can be disabled and/or enabled on the presented view. Any disabled markers are not considered suitable for further evaluation.

Table 3. The electrode positions computed using the proposed approach and defined by manually marking the clip center on each view deviated from each other, on average, by ≈1.9 mm ± 1.5 mm. In addition, the distance between the computed electrode position and the mean point obtained by the projection of each electrode onto each individual view, as well as the average position found by manually marking the center of the clip on each view, are provided. Along with the variation resulting from the ICP-based surface alignment, both values allow for the assessment of how well the proposed approach can approximate the true positions of the electrodes.

Subj.	Manual (mm)		$\overline{Manual} - marker$ (mm)		Mapping (mm)		Registration (mm)	
	Mean	std.	Mean	std.	Mean	std.	Mean	std.
2	1.4	0.9	1.7	1.4	0.6	0.9	0.6	0.2
3	2.1	1.4	1.6	1.5	0.9	1.4	0.6	0.2
4	2.9	1.8	2.2	1.5	1.1	1.8	0.7	0.2
5	2.7	1.6	2.4	1.8	0.9	1.6	0.6	0.2
Mean	2.3	1.4	2.0	1.5	0.9	1.4	0.6	0.2

The corrected electrode positions deviated, on average, by 2.3 mm ± 1.4 mm from the mean point, and the computed electrode positions deviated from the mean point by [2.0 mm ± 1.5 mm]. This is in accordance with the limitations posed by the backprojection,

where the reprojected points deviated from the computed position by 0.9 mm $\pm$ 1.4 mm, and the ICP registration resulted in an average deviation between corresponding points of 0.6 mm $\pm$ 0.2 mm. Given the amount of data to be processed per subject, the overall time of 22 min. required to extract and align the electrode positions is quite impressive, considering that only the computations of the asymmetric ICP and the HDBSCAN algorithms are implemented as part of the native open3D library and as Cython scripts, respectively. The rest of the implementation was carried out in Python using NumPy arrays only. In contrast, the expert required between 30 min. and 45 min. to point and place the electrode markers on the 14 views of a single data set.

4. Discussion

The results are promising given the fact that the torso is a far less rigid structure compared to the skull. Further, the limited space conditions and adverse environmental conditions typically found in clinical settings, e.g., outpatient and local practitioner clinics, are quite challenging. This is evident in the results shown in Table 3 for subjects 4 and 5. In both cases, nearby obstacles such as backrests or furniture limited access to the patient's left side, resulting in increased positional variations of 2.2 mm $\pm$ 1.5 mm and 2.4 mm $\pm$ 1.8 mm in relation to the mean of the manually defined electrode positions. This is compared to 1.7 mm $\pm$ 1.4 mm and 1.6 mm $\pm$ 1.5 mm for subjects 2 and 3, respectively.

These values are still in the range reported for recently proposed approaches for localizing electrodes mounted on the human body. As shown in Table 4, few studies exist that evaluate the use of 3D DS cameras [19,20] and photogrammetry methods [18] for localizing ECG electrodes on the torso. The achieved results varied between 1.16 mm and 11.8 mm, depending on the metrics and positional references used. The authors of [20] used the Hausdorff metric to compare the positions obtained from a Microsoft Kinect 3D DS camera to positions found on MRI or CT scans. On average, they achieved a positional error of 11.8 mm, which is an order of magnitude larger than the error between 1.16 mm and 2.5 mm achieved by Schulze et al. [18], Alioui et al. [19] and the present study, all of which used the Euclidean metric instead.

The majority of studies proposed methods for the localization of EEG sensors mounted on the scalp. Apart from Homölle and Oostenveld [8], the achieved average positional errors ranged from 1.5 mm [12] to 3.26 mm [14] using various reference measurements, including the mean of manually placed marks [12,14] and positional references generated using a magnetic digitizer [8,13,16] such as the Polhemus Fastrak. Comparing the positional error of 9.4 mm achieved by Homölle and Oostenveld [8] to all other results, it can be assumed that this was mainly caused by unavoidable inaccuracies when taking the magnetic digitizer measurements.

Considering that the positions of ECG electrodes mounted on the torso are directly affected by any movements, the positional error of 2.0 mm achieved in the present study is a clear indication that the active engagement and participation of the patient in the measurement is essential. The instructions on how the patient can easily maintain a posture that facilitates the recording of the electrode positions have a huge impact on the outcome of the measurements. If the instructions are not clearly defined by the measurement protocol, or not properly understood or followed by the patient, the positional error will increase. For example, subject 4 (see Table 3) changed the position of his arms during the measurement twice. This immediately resulted in an increased positional error of 2.2 mm $\pm$ 1.5 mm.

In addition to the limited space, the lighting conditions encountered in the clinical environment, as well as tight schedules, have a direct impact on the average positional error. Varying lighting conditions, including multiple light sources with differing light temperatures, on the other hand, can have a negative impact on photogrammetric approaches and 3D DS camera-based measurements of the torso surface and the electrode positions thereon. Algorithms for automatic white balancing and exposure control have been adopted to improve color constancy across multiple 3D views of the torso and maintain a constant exposure of the torso independent of the viewing direction and angle. In combination with

the developed calibration method, this resulted in increased accuracy in identifying those pixels representing the color markers.

Table 4. Comparison of the proposed approach for localizing ECG electrodes on the torso using a 3D DS camera with recent developments. In contrast to the large number of publications addressing the localization of EEG electrodes, only a few could be found using 3D DS cameras. The results obtained in the present study are within the ranges found by other studies.

Source	Sensor/Method	Multi-View Registration	Ref.	# El.	μ (mm)	σ (mm)	max. (mm)
			ECG				
Presented	3D DS	Symmetric ICP	manual mean	67	2.0	1.5	–
			reprojection		0.9	1.4	–
	manual marking		manual mean		2.3	1.4	–
Perez E. (2018) [20]	3D DS	Kinect Software	MRI/CT	128	11.8	–	64.8
Alioui S. (2017) [19]	3D DS	only one view	N/A	3	2.5	–	–
Schulze W. (2014) [18]	Photogrammetry	Least Squares	marker plate	80	1.16	0.97	–
			EEG				
Chen S. (2019) [14]	3D DS	Least Squares	multiple repeats	30	3.26	1.05	–
Homölle S. (2019) [8]	3D DS	Scanner Software	Magnetic digitizer	61	9.4	–	10.9
Cline C. C. (2018) [13]	IR-Scanner	Scanner Software	Magnetic digitizer/	128	1.73	0.37	–
	Magnetic digitizer		IR-Scanner		2.98	0.89	–
	VR-Digitizer				3.74	0.71	–
Clausner T. (2017) [43]	Photogrammetry	Photogrammetry	Face Scan	68	1.30	0.6	–
	Magnetic digitizer	software			7.80	2.1	–
Butler R. 2017 [9]	MRI	–	manual mean	63	0.5	–	–
Dalal S. S. (2014) [12]	Magnetic digitizer	–	FaceScan +	68	6.8	–	13.3
	Flying Triangulation		manual mean		1.5	–	2.9
Kössler L. (2010) [16]	Laser scanner	Scanner software	Magnetic digitizer	68	1.83	1.16	–

Time, in particular, is a very limited resource, which largely limits the routine use of magnetic digitizers within clinical environments. For precise positional measurement, the exact placement of the magnetic probe on each electrode and manual triggering of the measurement are required. An experienced user requires about 15 min. to accomplish this task. Any attempt to reduce this time can only be achieved by the less accurate placement of the probe on each electrode, which can result in increased positional errors of 7.8 mm and higher, as encountered by Clausner et al. [43].

In general, keeping the required human interactions and number of related errors as low as possible is one key goal for establishing NICE-based tools and procedures in clinical environments. The time required to localize the electrode positions on the human torso, as well as the amount of ionizing radiation the patient is exposed to, are key factors that can either prevent or facilitate a successful uptake. Alternative approaches currently used to obtain the electrode positions include manually placing markers on CT and MRI scans [9,12,19]. and automatically segmenting and pointing a magnetic digitizer probe to each individual electrode [8,13,16]. These approaches require a significant amount of time (about 45 min.) to point to each electrode, which is more than the 15 min. required for magnetic probe-based measurements. The mentioned approaches suffer from an additional bias related to the individual human perception of the electrode and marker shapes, as well as inaccuracies in the way the pointing probes are placed onto the electrode.

In contrast, the proposed 3D DS camera-based approach is not affected by these kinds of errors. When implemented on a tablet computer, the presented approach will enable clinicians to acquire the electrode positions and torso surfaces within 10 min. Therefore,

average positional errors of less than 2.5 mm will be feasible even under limited spatial conditions and tight schedules.

Some aspects essential for the successful clinical uptake of the presented approach still have to be addressed. On all color sensors, the raw signals recorded for red, green, and blue channels have to be converted into the RGB color space before they can be used. If the required parameters are not properly calibrated, the resulting images may show a bluish hue that can not be corrected by any white-balancing algorithm. This was the case for subject 5 shown in Figure 3, and caused the additional peak (4) in the calibration heat map shown in Figure 6. During the preparation of future studies, it is necessary to establish an appropriate procedure for verifying and optimizing the settings for these parameters before the first measurement and at regular intervals.

Each 3D DS camera data set also provides a point cloud representation of the torso surface. This is used in current studies to build electroanatomical models for electrocardiographic noninvasive imaging methods from clinical cardiac CT slices only. Further applications of the proposed approach are currently being investigated for enhanced electrical impedance tomography.

5. Conclusions

In the presented work, a complete 3D DS camera-based system was developed for localizing 67 ECG electrodes identified by color markers. Issues such as varying lighting conditions, including multiple light sources with different light temperatures, and the alignment of individual 3D views were addressed. The implemented recording protocol provides precise rules on how to seat the patient and includes well-defined instructions for the patient to easily maintain a specific isometric posture while all views are recorded. The resulting active engagement and participation of the patient in the measurement helped to minimize positional errors caused by the patient moving during the measurement. In combination with the symmetric ICP algorithm implemented, average positional errors of 2.3 mm or less could be achieved for each measurement.

The implemented prototype system localizes the electrodes on the torso with minimal human interaction. It can handle diverse lighting conditions and operate in narrow spaces, as encountered in clinical settings such as outpatients of local practitioner clinics.

Author Contributions: Conceptualization, J.B. and C.H.; methodology, J.B.; software, J.B. and C.H.; validation, J.B., C.H. and H.B.; formal analysis, H.B. and C.S.; data curation, C.H. and H.B.; writing—original draft preparation, C.H.; writing—review and editing, J.B., C.H., H.B. and C.S.; visualization, C.H.; supervision, C.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Linz Center of Mechatronics.

Institutional Review Board Statement: This study was conducted in accordance with the Declaration of Helsinki, and approved by the Ethics Committee of the Medical Faculty of Johannes Kepler University Linz (protocol code 1190/2020 and date of 16 September 2020).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data and software are made available in the data and software repository of Johannes Kepler university. Until public access to these reporsitories is available in its full extent data developed data and software are provided on request by the institute for Biomedical Mechatronics at Johannes Kepler University, mmt@jku.at.

Acknowledgments: The authors would like to thank the patients who participated in the PAIP study, as well as the team in the Department of Cardiology at the Kepler University Hospital for providing the required resources and for their support. Further, the authors would like to thank Werner Baumgartner and colleagues from the Biomedical Mechatronics Institute at Johannes Kepler University for their support during the development of the prototype and the preparation of this manuscript.

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

Abbreviations

The following abbreviations are used in this manuscript:

EEG	Electroencephalogram
ECG	Electrocardiogram
EIT	Electrical Impedance Tomography
ECT	Electrical Capacitance Tomography
MRI	Magnetic Resonance Tomography
CT	Computed Tomography
3D DS camera	3D Depth-Sensing camera
RGB	Red, Green, Blue
AWB	Automatic White Balancing
FOV	Field of View
AE	Auto-Exposure
ICP	Iterative Closest Point

References

1. Ogawa, R.; Baidillah, M.R.; Darma, P.N.; Kawashima, D.; Akita, S.; Takei, M. Multifrequency Electrical Impedance Tomography With Ratiometric Preprocessing for Imaging Human Body Compartments. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 1–14. [CrossRef]
2. Lymperopoulos, G.; Lymperopoulos, P.; Alikari, V.; Dafogianni, C.; Zyga, S.; Margari, N. Applications for Electrical Impedance Tomography (EIT) and Electrical Properties of the Human Body. *Adv. Exp. Med. Biol.* **2017**, *989*, 109–117. [PubMed]
3. Gupta, S.; Lee, H.J.; Loh, K.J.; Todd, M.D.; Reed, J.; Barnett, A.D. Noncontact Strain Monitoring of Osseointegrated Prostheses. *Sensors* **2018**, *18*, 3015. [CrossRef] [PubMed]
4. Rudy, Y.; Lindsay, B.D. Electrocardiographic Imaging of Heart Rhythm Disorders: From Bench to Bedside. *Card. Electrophysiol. Clin.* **2015**, *7*, 17–35. [CrossRef] [PubMed]
5. Zhang, J.; Sacher, F.; Hoffmayer, K.; OH́ara, T.; Strom, M.; Cuculich, P.; Silva, J.; Cooper, D.; Faddis, M.; Hocini, M.; et al. Cardiac Electrophysiological Substrate Underlying the ECG Phenotype and Electrogram Abnormalities in Brugada Syndrome Patients. *Circulation* **2015**, *131*, 1950–1959. [CrossRef] [PubMed]
6. Seger, M.; Hanser, F.; Dichtl, W.; Stuehlinger, M.; Hintringer, F.; Trieb, T.; Pfeifer, B.; Berger, T. Non-invasive imaging of cardiac electrophysiology in a cardiac resynchronization therapy defibrillator patient with a quadripolar left ventricular lead. *Europace* **2014**, *16*, 743–749. [CrossRef]
7. Ettl, S.; Rampp, S.; Fouladi-Movahed, S.; Dalal, S.S.; Willomitzer, F.; Arold, O.; Stefan, H.; Häusler, G. Improved EEG source localization employing 3D sensing by "Flying Triangulation". In Proceedings of the Optical Metrology, Munich, Germany, 13–15 May 2013.
8. Homölle, S.; Oostenveld, R. Using a structured-light 3D scanner to improve EEG source modeling with more accurate electrode positions. *J. Neurosci. Methods* **2019**, *326*, 108378. [CrossRef]
9. Butler, R.; Gilbert, G.; Descoteaux, M.; Bernier, P.M.; Whittingstall, K. Application of polymer sensitive MRI sequence to localization of EEG electrodes. *J. Neurosci. Methods* **2017**, *278*, 36–45. [CrossRef]
10. Marino, M.; Liu, Q.; Brem, S.; Wenderoth, N.; Mantini, D. Automated detection and labeling of high-density EEG electrodes from structural MR images. *J. Neural Eng.* **2016**, *13*, 056003. [CrossRef]
11. Ma, Y.; Mistry, U.; Thorpe, A.; Housden, R.J.; Chen, Z.; Schulze, W.H.W.; Rinaldi, C.A.; Razavi, R.; Rhode, K.S. Automatic Electrode and CT/MR Image Co-localisation for Electrocardiographic Imaging. In *Proceedings of the Functional Imaging and Modeling of the Heart*; Ourselin, S., Rueckert, D., Smith, N., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 268–275.
12. Dalal, S.S.; Rampp, S.; Willomitzer, F.; Ettl, S. Consequences of EEG electrode position error on ultimate beamformer source reconstruction performance. *Front. Neurosci.* **2014**, *8*, 42. [CrossRef]
13. Cline, C.C.; Coogan, C.; He, B. EEG electrode digitization with commercial virtual reality hardware. *PLoS ONE* **2018**, *13*, e0207516. [CrossRef]
14. Chen, S.; He, Y.; Qiu, H.; Yan, X.; Zhao, M. Spatial Localization of EEG Electrodes in a TOF+CCD Camera System. *Front. Neuroinform.* **2019**, *13*, 21. [CrossRef]
15. Ghanem, R.N.; Ramanathan, C.; Jia, P.; Rudy, Y. Heart-surface reconstruction and ECG electrodes localization using fluoroscopy, epipolar geometry and stereovision: Application to noninvasive imaging of cardiac electrical activity. *IEEE Trans. Med. Imaging* **2003**, *22*, 1307–1318. [CrossRef]
16. Kössler, L.; Maillard, L.; Benhadid, A.; Vignal, J.P.; Braun, M.; Vespignani, H. Spatial localization of EEG electrodes. *Neurophysiol. Clin. = Clin. Neurophysiol.* **2007**, *37*, 97–102. [CrossRef]
17. He, Y.; Qiu, H.; Gu, Y.; Chen, S. EEG Electrode Localization based on Joint ToF and CCD Camera Group. In Proceedings of the 2019 IEEE 5th International Conference on Computer and Communications (ICCC), Chengdu, China, 6–9 December 2019; pp. 1771–1776. [CrossRef]

18. Schulze, W.H.W.; Mackens, P.; Potyagaylo, D.; Rhode, K.; Tülümen, E.; Schimpf, R.; Papavassiliu, T.; Borggrefe, M.; Dössel, O. Automatic camera-based identification and 3-D reconstruction of electrode positions in electrocardiographic imaging. *Biomed. Tech. Biomed. Eng.* **2014**, *59*, 515–528. [CrossRef]
19. Alioui, S.; Kastelein, M.; van Dam, E.M.; van Dam, P.M. Automatic registration of 3D camera recording to model for leads localization. In Proceedings of the 2017 Computing in Cardiology (CinC), Rennes, France, 24–27 September 2017; pp. 1–4. [CrossRef]
20. Perez-Alday, E.A.; Thomas, J.A.; Kabir, M.; Sedaghat, G.; Rogovoy, N.; van Dam, E.; van Dam, P.; Woodward, W.; Fuss, C.; Ferencik, M.; et al. Torso geometry reconstruction and body surface electrode localization using three-dimensional photography. *J. Electrocardiol.* **2018**, *51*, 60–67. [CrossRef]
21. Intel RealSense^{TM} Depth Camera SR300 Series Product Family. 2021. Available online: https://www.intelrealsense.com/wp-content/uploads/2019/07/RealSense_SR30x_Product_Datasheet_Rev_002.pdf (accessed on 24 April 2023).
22. Zabatani, A.; Surazhsky, V.; Sperling, E.; Moshe, S.B.; Menashe, O.; Silver, D.H.; Karni, Z.; Bronstein, A.M.; Bronstein, M.M.; Kimmel, R. Intel RealSense^{TM} SR300 Coded Light Depth Camera. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *42*, 2333–2345. [CrossRef]
23. Cohen, N. A Color Balancing Algorithm for Cameras. EE368 Digital Image Processing. 2011. Available online: https://stacks.stanford.edu/file/druid:my512gb2187/Cohen_A_New_Color_Balancing_Method.pdf (accessed on 24 April 2023).
24. Chen, W.; Li, X. Exposure Evaluation Method Based on Histogram Statistics. In Proceedings of the 2017 2nd International Conference on Electrical, Automation and Mechanical Engineering (EAME 2017), Shanghai, China, 23–24 April 2017; Atlantis Press: Amsterdam, The Netherlands, 2017; pp. 290–293. [CrossRef]
25. Wang, B.; Fan, S. An Improved CANNY Edge Detection Algorithm. In Proceedings of the 2009 Second International Workshop on Computer Science and Engineering, Tianjin, China, 3–6 August 2014; Volume 1, pp. 497–500. [CrossRef]
26. Rusinkiewicz, S. A Symmetric Objective Function for ICP. *ACM Trans. Graph.* **2019**, *38*. [CrossRef]
27. Zhou, Q.Y.; Park, J.; Koltun, V. Open3D: A Modern Library for 3D Data Processing. *arXiv* **2018**, arXiv:1801.09847.
28. Blanco, J.L.; Rai, P.K. Nanoflann: A C++ Header-Only Fork of FLANN, a Library for Nearest Neighbor (NN) with KD-Trees. 2014. Available online: https://github.com/jlblancoc/nanoflann (accessed on 24 April 2023).
29. Muja, M.; Lowe, D.G. Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration. In Proceedings of the International Conference on Computer Vision Theory and Applications, Lisboa, Portugal, 5–8 February 2009.
30. McInnes, L.; Healy, J.; Astels, S. hdbscan: Hierarchical density based clustering. *J. Open Source Softw.* **2017**, *2*, 205. [CrossRef]
31. Ester, M.; Kriegel, H.P.; Sander, J.; Xu, X. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD'96), Portland, OR, USA, 2–4 August 1996.
32. Byrd, R.H.; Lu, P.; Nocedal, J.; Zhu, C. A Limited Memory Algorithm for Bound Constrained Optimization. *SIAM J. Sci. Comput.* **1995**, *16*, 1190–1208. [CrossRef]
33. Zhu, C.; Byrd, R.H.; Lu, P.; Nocedal, J. Algorithm 778: L-BFGS-B: Fortran Subroutines for Large-Scale Bound-Constrained Optimization. *ACM Trans. Math. Softw.* **1997**, *23*, 550–560. [CrossRef]
34. Bernardini, F.; Mittleman, J.; Rushmeier, H.; Silva, C.; Taubin, G. The ball-pivoting algorithm for surface reconstruction. *IEEE Trans. Vis. Comput. Graph.* **1999**, *5*, 349–359. [CrossRef]
35. Digne, J. An Analysis and Implementation of a Parallel Ball Pivoting Algorithm. *Image Process. Line* **2014**, *4*, 149–168. [CrossRef]
36. Hintermüller, C.; Seger, M.; Pfeifer, B.; Fischer, G.; Modre, R.; Tilg, B. Sensitivity and Effort-Gain Analysis: Multi-Lead ECG Electrode Array Selection for Activation Time Imaging. *IEEE Trans. Biomed. Engeneering* **2006**, *53*, 2055–2066. [CrossRef]
37. Kambhatla, N.; Leen, T. Classifying with Gaussian Mixtures and Clusters. In *Proceedings of the Advances in Neural Information Processing Systems*; Tesauro, G., Touretzky, D., Leen, T., Eds.; MIT Press: Cambridge, MA, USA, 1994; Volume 7.
38. Reynolds, D. Gaussian Mixture Models. In *Encyclopedia of Biometrics*; Li, S.Z., Jain, A.K., Eds.; Springer: Boston, MA, USA, 2015; pp. 827–832. [CrossRef]
39. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17*, 261–272. [CrossRef]
40. Intel RealSense^{TM} Cross Platform API. 2022. Available online: https://intelrealsense.github.io/librealsense/doxygen/ (accessed on 1 April 2022).
41. Itseez. Open Source Computer Vision Library. 2015. Available online: https://github.com/itseez/opencv (accessed on 24 April 2023).
42. Behnel, S.; Bradshaw, R.; Citro, C.; Dalcin, L.; Seljebotn, D.S.; Smith, K. Cython: The Best of Both Worlds. *Comput. Sci. Eng.* **2011**, *13*, 31–39. [CrossRef]
43. Clausner, T.; Dalal, S.S.; Crespo-García, M. Photogrammetry-Based Head Digitization for Rapid and Accurate Localization of EEG Electrodes and MEG Fiducial Markers Using a Single Digital SLR Camera. *Front. Neurosci.* **2017**, *11*, 264. [CrossRef]