

Article

A Kinect-Based System for Upper-Body Function Assessment in Breast Cancer Patients

Rita Moreira ^{1,2}, André Magalhães ^{2,3} and Hélder P. Oliveira ^{1,2,*}

¹ Faculdade de Engenharia da Universidade do Porto, R. Dr. Roberto Frias, Porto 4200-465, Portugal; E-Mail: aritacmoreira@gmail.com

² Instituto de Engenharia de Sistemas e Computadores, Tecnologia e Ciência - INESC TEC, R. Campus da FEUP, Dr. Roberto Frias, Porto 4200-465, Portugal; E-Mail: amag1976@gmail.com

³ Faculdade de Medicina da Universidade do Porto, Alameda Prof. Hernâni Monteiro, Porto 4200-319, Portugal

* Author to whom correspondence should be addressed; E-Mail: helder.f.oliveira@inesctec.pt; Tel.: +351-22-209-4306.

Academic Editors: Philip Morrow, Kenji Suzuki and Gonzalo Pajares Martinsanz

Received: 16 September 2015 / Accepted: 30 October 2015 / Published: 5 November 2015

Abstract: Common breast cancer treatment techniques, such as radiation therapy or the surgical removal of the axillary lymphatic nodes, result in several impairments in women's upper-body function. These impairments include restricted shoulder mobility and arm swelling. As a consequence, several daily life activities are affected, which contribute to a decreased quality of life (QOL). Therefore, it is of extreme importance to assess the functional restrictions caused by cancer treatment, in order to evaluate the quality of procedures and to avoid further complications. Although the research in this field is still very limited and the methods currently available suffer from a lack of objectivity, this highlights the relevance of the pioneer work presented in this paper, which aims to develop an effective method for the evaluation of the upper-body function, suitable for breast cancer patients. For this purpose, the use of both depth and skeleton data, provided by the Microsoft Kinect, is investigated to extract features of the upper-limbs motion. Supervised classification algorithms are used to construct a predictive model of classification, and very promising results are obtained, with high classification accuracy.

Keywords: quality of life; breast neoplasms/surgery; range of motion; articular; upper extremity/physiopathology; edema/epidemiology; postoperative complications/epidemiology; lymph node excision/adverse effects

1. Introduction

Breast cancer is the most common cancer both in developed and developing regions, representing a total of 23% [1]. Nevertheless, due to effective earlier diagnosis methods and effective adjuvant therapies, the mortality rate is about 27% in the more developed regions [1]. These high rates of long-term disease survival indicate that a large proportion of patients needs to live daily with the consequences of the treatment. Consequently, the interest in the psychological adaptation and quality of life (QOL) of women after treatment has been growing.

About 25% of women diagnosed with breast cancer present cancer cells in the axillary lymph node system [2]. In these cases, treatments can include the clearance of axillary lymph nodes or axillary radiotherapy. Both of these procedures may cause serious damage of shoulder function and interruption of lymph drainage of the upper limb, which results in accumulation of lymph in the arm tissue (lymphedema) [3]. Therefore, several upper-limb problems are normally verified in these patients, including restricted shoulder mobility, lymphedema and arm/shoulder pain. According to the study of Nesvold *et al.* [4], at a five-year follow-up, the prevalence of arm/shoulder pain is 30% to 40%, of lymphedema 10% to 15% and of restricted arm/shoulder mobility 15% to 30%. These upper-body morbidities are highly correlated with decreased QOL, since they disrupt valuable daily life activities [5]. Furthermore, they can be considered as an aesthetic problem [6,7]. However, there is no clinical standard for the diagnosis of these problems, and the methods currently used are non-practical for clinical settings and have problems regarding the lack of objectivity or inaccuracy.

An objective evaluation of the upper-body motion restrictions can be a helpful tool in the determination and quantification of these sequelae related to breast cancer treatment. This functional evaluation is essential to identify which procedures have the least morbidity associated. Moreover, with an early identification and diagnosis of impairments, therapeutic interventions can lead to greater success in managing upper-limb morbidity, with improved outcomes for breast cancer survivors [8]. Thus, the lack of a practical and standard clinical tool in this area leads to the need for a simple, accurate, low-cost and reproducible method for the assessment of the upper-body function (UBF).

An early diagnosis of upper-limb impairments is also important to have an identification of adequate therapies that can lead to greater success in managing functional morbidity, as well as the prevention of progression, with improved outcomes and QOL for breast cancer survivors. The benefits of physiotherapy and physical activity on the recovering of upper-extremity range-of-motion (ROM), strength and function for women's QOL have been well documented [9–13]. Therefore, a comprehensive model of care to identify exercise prescription and to guide the rehabilitation of breast cancer-related physical impairments would be also useful.

The present work proposes an effective method for the objective evaluation of the upper-body functional status, suitable for breast cancer patients. For this purpose, the use of low-cost equipment,

the Microsoft Kinect (Microsoft Corp., Redmond, WA, USA), which provides RGB and depth data and allows a simplified skeleton tracking, is investigated. With this system, features of the patients' upper-limb motion were extracted, and a predictive classification model was built using supervised classification algorithms. The results obtained were very promising, which indicate that the methodology proposed is efficient in the evaluation of the UBF status in breast cancer patients. The work presented in this paper is an extension of research available in the article [14].

The paper is organized as follows: Section 2 presents a review of all of the topics related to upper-body functional assessment. Section 3 provides some insights and describes the developed methodology regarding the proposed system. The main results obtained can be consulted in Section 4. Finally, Section 5 serves as a conclusion to the present research.

2. Related Work

Decreased UBF caused by breast cancer treatment is correlated with a higher arm volume (edema) and restricted ROM. Despite its influence on women's QOL, there is no clinical standard for the diagnosis of these impairments. Moreover, current technologies are not suitable for clinical settings and have problems of accuracy and reproducibility.

2.1. Upper-Limb Volume Measurements

Lymphedema is one of several morbidities that adversely affect the upper-body function status of breast cancer patients. It is regarded as an incurable, progressive, disfiguring and disabling disorder [15]. Only an early detection of edema indicators allows an early intervention, reducing the risk of chronic lymphedema [16].

Ideally, edema evaluation should be performed by the comparison between measurements of the limb before and after the treatment. However, there is no habit for performing these examinations before the surgery; thereby, bilateral limb comparisons are usually made [17].

2.1.1. Traditional Methods

Common methodologies currently used for lymphedema detection include water displacement, circumference measurements, bioelectrical impedance spectroscopy (BIS) or imaging techniques (see Table 1 [18–22]).

When performed properly, water displacement is accurate [20]; still, it is time consuming, non-portable and non-hygienic [19]. Regarding the multiple circumferential measurements, sources of error arise from the assumption of the circular cross-section of the limb [20]. Additionally, arm circumferences can be obtained with the use of a Perometer[®] [21], with reliable and highly-reproducible measurements [19], but the size and the cost of the equipment limit its usability. Imaging techniques, such as computed tomography (CT), are also used to evaluate limb swelling [22], but they require equipment that is not commonly available in clinics. On the other hand, BIS has been previously well described as a lymphedema measure [20]. Nevertheless, this method involves the use of several electrodes placed along the arm, which leads to a high lifetime operational cost.

Table 1. Comparison of the most significant methods for limb edema assessment. Adapted from [18]. BIS, bioelectrical impedance spectroscopy.

System	Time to Operate	Home/Travel	Accuracy	Cost	Complex
Water Displacement: The limb is immersed into a container, and the amount of the displaced water represents the volume [19].	High	No	High	Low	Medium
Circumferential Measurements: The volume can be estimated assuming cylindrical/conic volumes between several measures taken along the limb [20].	High	Yes	Low	Low	Low
Perometer ®: The device scans the limb with IR light and assesses the volume at small intervals [21].	Medium	No	High	Medium	High
CT: Determines the overall cross-section area and quantifies the density of the tissues [22].	Low	No	High	High	High
BIS: A small current passes through the body; it measures volumes by comparing the impedance values of both arms [20].	Medium	No	Medium	High	Medium

2.1.2. Three-Dimensional Scanners

Some of the methods described above can provide an objective and accurate measure of the limb volume; however, they are time consuming, complicated or expensive. In this way, research is being done to test the application of the arising three-dimensional (3D) technologies on body volume estimation. Low-cost, non-invasive and easy-to-use 3D body-surface scanners are transforming the ability to accurately measure the body size, shape and skin surface area. These features make them appealing for widespread clinical applications. In recent years, several authors [23,24] validated the used of these systems for limb volume assessment. Examples, as the case of Polhemus FastSCAN™ and the Insignia™ laser scanner, are reported. These methods combine precision and reproducibility; therefore, they may be useful to measure lymphedema. Nonetheless, they are expensive devices and are quite difficult to handle by non-professionals. On the other hand, Lu *et al.* [18] explored the use of a low-cost system, the Microsoft Kinect, for arm modeling and volume measurement. However, the proposed method is quite complex, since it depends on gyroscopes and accelerometers attached to the sensor.

2.2. Upper-Limb Motion Evaluation

The assessment of functional status after breast cancer treatment should also give attention to UBF impairments, activity limitations and participation restriction, deduced from limited shoulder ROM. In

that way, as for edema detection, there is a need for an objective, reproducible and low-cost method for the evaluation of upper-body motion.

2.2.1. Subjective Methods

Normally, UBF evaluation relies on subjective measurements of patients’ experiences and function limitation [25]. Thus, several generic self-report questionnaires have been developed to capture the effects of injury on the upper-body function (see Table 2 [25–29]).

Table 2. Self-report scales used for upper-body function assessment of breast cancer patients. Adapted from [25]. PSFS, Patient-Specific Functional Scale; DASH, Disabilities of the Arm, Shoulder and Hand; UEFI, Upper Extremity Functional Index; KAPS, Kwan’s Arm Problem Scale.

Scale	Paper	Type of Measure	Description	Clinical interpretation	Comments
PSFS	Stratford <i>et al.</i> [26]	Clinical measure of function	3 items; 11-point scale	Higher score, better function	For use in the clinical setting; measures the change in function specific to the individual survivor.
DASH	Hudak <i>et al.</i> [27]	Pain-related upper extremity disability	30 items; 5-point scale	Higher score, poorer function	Has not been validated in breast cancer patients.
UEFI	Stratford <i>et al.</i> [28]	Upper-body function	20 items; 5-point scale	Higher score, better function	Valid and sensitive to changes in the breast cancer population.
KAPS	Kwan <i>et al.</i> [29]	Upper-body symptoms and function	13 items; 5-point scale	Higher score, poorer function	Developed to identify shoulder and arm problems during breast cancer treatment.

Specific scales for UBF assessment include the Patient-Specific Functional Scale (PSFS) [26], developed as a clinical measure of functional status limitations related to the effect of a treatment/intervention. The Disabilities of the Arm, Shoulder and Hand (DASH) scale [27], in turn, was designed to measure physical function and symptoms in patients with any disorder affecting the upper extremities. This self-report covers symptoms, such as pain, weakness and numbness, and the degree of difficulty performing several tasks [25]. The Upper Extremity Functional Index (UEFI) [28] aims to evaluate patients’ upper-extremity functional status in a variety of activities. On the other hand, Kwan’s Arm Problem Scale (KAPS) [29] was developed to identify shoulder and arm problems in breast cancer patients, including arm/shoulder motion, pain, stiffness, swelling and impairments with respect to performing daily activities. Although easy-to-use and useful to provide reference data, subjective methods are generally not accurate and have problems related to impartiality and poor reproducibility.

2.2.2. Objective Methods

For the objective assessment of upper-extremity functional limitations, there are several methods that include tests of flexibility, strength and endurance. From these tests, the most common for the evaluation of UBF in breast cancer patients is goniometry [30], used to assess active and passive shoulder ROM in all planes. Comparing measurements of the affected and unaffected limb, it is possible to detect restricted mobility or impaired shoulder function. Other approaches can include the use of isometric and isokinetic dynamometry and/or maximal performance of a set of tasks/exercises using the repetition maximum method [31]. Therefore, strength and endurance can be measured by means of an incremental exercise protocol where each stage lasts one minute in duration and increments are made by increasing speed of movement and weight held [31]. On the other hand, the assessment of the grip strength with a standard hand dynamometer can also be relevant, since this is an important requisite for good arm function [31]. However, the use of these methods is not a clinical standard in breast cancer, and so, its sensitivity to functional limitations related to the treatment is not well known.

2.3. Summary

As can be verified in this section, methods of the assessment of upper-body functional status are non-practical for clinical use, time consuming, not reproducible and/or not accurate. Therefore, the lack of a practical and standard clinical tool in this area leads to the need for a simple, portable and objective system for the assessment of UBF in breast cancer patients. Recently, advances in 3D image acquisition technologies have made them an attractive tool with potential applications in this field.

3. An Upper-Body Function Evaluation System

This paper addresses the development of a Kinect-based system for the evaluation of the upper-body function. For this purpose, Microsoft Kinect capabilities are used, to extract features that characterize the upper-limb motion. In this manner, this work included the development of a database of breast cancer patients acquired using the Kinect device. This database comprises color, depth and skeleton data of the patients performing adduction and abduction movements of the upper-limbs. The complete medical information is available for each patient, including the ones diagnosed with lymphedema and data regarding the type of treatments used.

The proposed system can be implemented as a sequence of a few high-level operations (see Figure 1). Firstly, an effective segmentation of the patient was accomplished, combining the skeleton information and the depth data. Furthermore, using the patient silhouette and the position of the shoulder and elbow joints, it was possible to perform the detection of the arm contour, important for the extraction of a movement's features. Thereafter, it is possible to obtain features that relate the ROM of both shoulders, the volume ratio, the hand height and width, the elbow flexion and the movement acceleration. Finally, a predictive model is built, using a machine learning methodology, to evaluate the upper-body functional status.

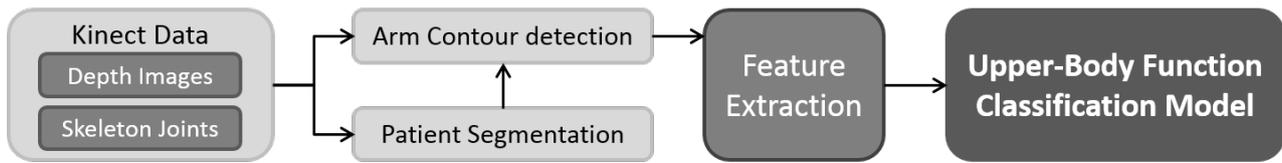


Figure 1. System flowchart for the upper-body function evaluation of the proposed method.

3.1. Depth-Map Noise Reduction

Kinect is able to capture real-time depth maps using an infra-red projector and camera. However, these data present a significant amount of noise-related issues, due to the limits of working distance, occlusions, multiple reflections or scattering. This leads to missing regions and unstable boundaries in depth maps. Thus, this work included the implementation of a method to address the issue of noise around object boundaries. For this purpose, a bilateral filter is used. This type of filter combines a domain kernel (σ_d), which gives priority to pixels that are close to the target pixel, with a range kernel (σ_r), which gives priority to pixels similar to the target pixel [32]. With this approach, it is possible to obtain image smoothing and to preserve edges at the same time.

3.2. Patient Segmentation

The conditions of data acquisition are not constant from patient to patient. Ideally, this would be carried out in a uniform background, but occasionally, there is also the presence of outlier objects. Thus, captured images are normally composed of the background (wall) and the patient, but also with non-desirable objects (see Figure 2a,b). Therefore, global traditional threshold methods, such as Otsu’s, are not the most appropriate in this case, as previously verified [33]. In this manner, the available information of the skeleton joints can be very useful, since each joint indicates a position in the depth frame where the body of the patient will be positioned. Thereby, the pixel values in these positions are assessed, and a double threshold can be defined using the maximum and minimum values of intensity.

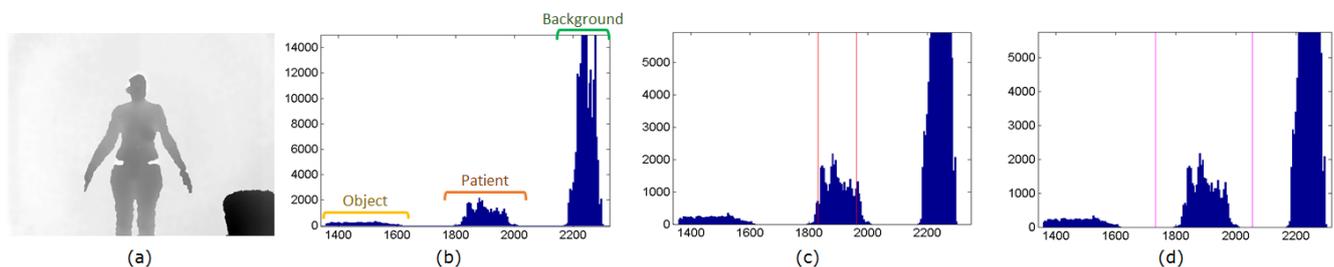


Figure 2. (a) Normalized depth image, where it is possible to visualize the presence of a non-desirable object; (b) corresponding histogram, where it is possible to identify the background (green), the patient (orange) and an outlier object (yellow); (c) the maximum and minimum values of intensity at the joints’ positions represented by the red lines; (d) 5% of the maximum and minimum values of intensity, represented by the pink lines, used as a double threshold for segmentation

However, using only these thresholds, several points that probably belong to the body will be considered as background (see Figure 2c). Therefore, a margin of 5% above and below the maximum and minimum, respectively, is used in order to consider all of the objects in the segmentation (see Figure 2d). Using this double threshold, the patient’s mask is successfully obtained. To remove small holes that may be present, a fill holes filter is applied. Further, to get a smoothed body shape, a median filter (11×11) is used.

The parameters of the segmentation algorithm were experimentally tuned and based on relative values, which minimizes the influence in the subsequent tasks.

3.3. Arm Point Detection

This work also included the detection of the points that delimit the upper arm. This region of the upper limb is bounded superiorly by the shoulder, inferiorly by the elbow and communicates medially with the axilla [34]. In this fashion, there are a set of points that should be found (see Figure 3): the armpit (A_1), the medial (E_1) and the lateral (E_2) elbow point, the medial shoulder point (S_2) and the lateral shoulder point (S_1). With this in mind, we start with the detection of the patient contour, accomplished by a morphological dilation of the mask, using a 3×3 structuring element, and by an exclusive disjunction operation:

$$I_{XOR} = I_M \oplus I_D \tag{1}$$

where I_M represents the binary body mask, I_D the outcome of the dilation and I_{XOR} the resultant silhouette. After this process, we have a set of pixels (the contour), which can be iteratively assessed, in order to find the points of interest. For this purpose, we took advantage of the shoulder and elbow joints given by the Kinect skeleton, represented by a S and E in Figure 3. To detect the point A_1 , we will search the pixels that are below the point S , since the armpit will always be beneath the Kinect shoulder joint, and select the closer one. Then, the medial elbow point (E_1) will correspond to the pixel that is closer to the elbow joint E . To guarantee that the point selected is indeed E_1 , it is only accepted as valid if the line segment $\overline{A_1E_1}$ does not intersect \overline{SE} . On the other hand, to detect E_2 , a line perpendicular to \overline{SE} , that contains E_1 is first determined. Then, we search the silhouette pixels that are members of this line. To choose a valid point E_2 , it is guaranteed that \overline{SE} intersects $\overline{A_1E_2}$. To find S_2 , we search for the pixel that forms with the shoulder S a line parallel to $\overline{E_1E_2}$. Besides, the intersection of $\overline{S_2E_1}$ and \overline{SE} is also a condition. Finally, S_1 will be the point of intersection between $\overline{E_1A_1}$ and $\overline{S_2S}$. This process is done for the right and left arm and repeated at each frame available for each patient.



Figure 3. Kinect shoulder and elbow joints (red dots) and upper-arm detected points (green dots).

3.4. Feature Extraction

With the Kinect skeleton data and the upper-arm segmentation, it was possible to perform a selection of features to assess, in order to evaluate the upper-body mobility. Therefore, it is possible to compare between both limbs: the upper-arm volume, the shoulder ROM, the maximum hand height and width, the acceleration of the movement and the elbow flexion. In order to minimize the influence of errors in the previous tasks (e.g., skeleton detection, segmentation and arm detection), with an impact on feature values, median values are used in order to avoid outliers, reducing the impact of the prediction of upper-body mobility issues.

3.4.1. Volume

To assess the volume of both limbs, the depth information is used, since it allows the quantification of the volume ratio between the two upper arms. In this case, a plan of reference of the upper-limb is first determined, and then, the volume is defined as the sum of unit values between the estimated plan and the arm surface (see Figure 4a). As a reference, the maximum depth value on the arm (which corresponds to the most distant point) is considered, represented by a red block in Figure 4a. Since it corresponds to the most trustworthy arm segmentation, the volume values are only considered for the frames where shoulder ROM is between 45 and 135 degrees. Moreover, only values between a range of 40% of the median are considered, in order to discard possible outliers. An average value of each arm is obtained, and their ratio is computed (μVOL).

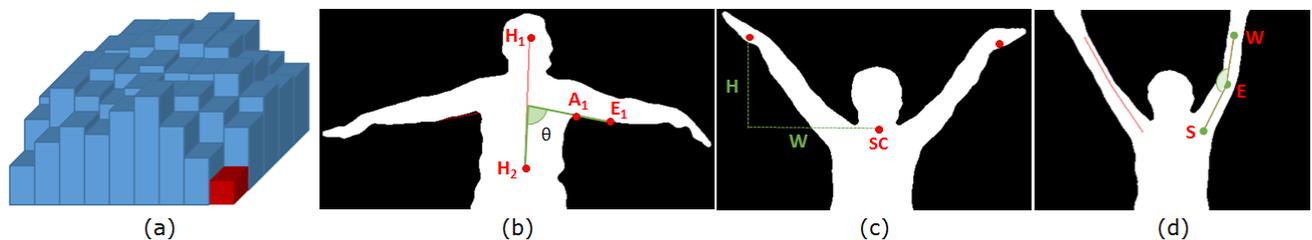


Figure 4. (a) Illustration of the volume measurement based on voxels; (b) angle θ measured to evaluate the shoulder range-of-motion (ROM); (c) hand height (H) and width (W) measurements; and (d) the angle computed to detect the elbow flexion.

3.4.2. Range of Motion

The decreased ROM of the shoulder is a common consequence observed in patients with limited UBF. This feature can be obtained by measuring the angle between the upper limb and the patient body. In other words, the angle θ between the line segment $\overline{A_1E_1}$ is computed, and the line $\overline{H_1H_2}$ is defined by the head and hip points of the Kinect skeleton (see Figure 4b). Therefore, the ROM of the right and left shoulder is obtained, for each frame of each patient, and their ratio is computed. From these data, two features are selected:

- the ratio of the maximum ROM, obtained at the maximum height achieved by the hand ($maxROM$).

- the average of the ROM ratio along all movement (μROM).

3.4.3. Hand Height and Hand Width

The reduced motion of the limb can also be detected comparing the hand height and, consequently, the width, that the patient reaches during the movement. Therefore, the x and y coordinates of the hands joints are analyzed, and the shoulder center (SC) joint is used as a reference point. This means that, for all of the frames, the difference between the hand coordinates (x and y) and the SC point is computed, represented, respectively, by the H and W in the Figure 4c. Thereafter, for each frame, the ratio of the y values of both hands is computed, as well as the the ratio of x values. As with the ROM, two features are obtained for the height and the width, for a total of four:

- the ratio of the maximum height achieved by the hand ($maxH$).
- the width at the maximum height point ($maxW$).
- the average of the height ratio along all movement (μH).
- the average of the width ratio along all movement (μW).

3.4.4. Elbow Flexion

The behavior of the elbow during the abduction/adduction movement was also evaluated. In this way, with the knowledge of the position of the shoulder, elbow and wrist joints, it is possible to calculate the angle between the elbow-wrist (\overline{EW}) and elbow-shoulder (\overline{ES}) lines (see Figure 4d) and, therefore, to assess the elbow flexion. This angle is obtained for the right and left elbow for each frame, and their ratio is computed. Once again, the average of the ratio values is used as a feature ($\mu ElbF$).

3.4.5. Hand Acceleration

The reduced acceleration of the movement can also be indicative of a motion problems, so the dynamics of each limb was also computed (see Figure 5). This is accomplished measuring the variation of the hand position over time. The joint chosen was the hand, since it has the biggest variation of position during the movement.

Hence, the instantaneous velocity of motion can be calculated by Equation (2), where T is the sampling interval, n is the total number of frames used to evaluate the movement and x, y, z represent the hand coordinates, in meters.

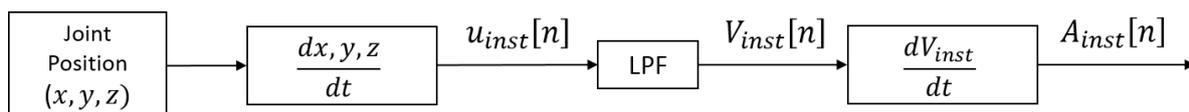


Figure 5. Flowchart of the process to obtain the hand’s instantaneous acceleration. (x, y, z) represent the joint coordinates of the hand in meters.

To reduce the generated noise from the environment and small body movements, a low-pass filter with a cutoff frequency of 2 Hz is applied to $u_{inst}[n]$, since the low frequencies are the main interest. In this way, an instantaneous velocity $V_{inst}[n]$ is obtained, with no phase distortion. The instantaneous

acceleration of the movement is then defined by Equation (3). Thus, it is possible to compute the average of the movement acceleration for each hand and to extract their ratio as a feature (μAcc).

$$u_{inst}[n] = \frac{dx, y, z}{dt} \Big|_{t=nT} \quad (m/s) \tag{2}$$

$$A_{inst}[n] = \frac{V_{inst}}{dt} \Big|_{t=nT} \quad (m/s^2) \tag{3}$$

3.5. Classification Models

A total of nine features were extracted (see Table 3), which allowed the use of machine learning techniques to build a predictive classification model.

Table 3. Features used in the classification models.

-	Acronym	Description
1	μVOL	Average of the volume ratio.
2	$maxROM$	ROM at the maximum height achieved by the hand.
3	$maxHeight$	Maximum height achieved by the hand.
4	$maxWidth$	Width at the maximum height point.
5	$\mu ElbF$	Average of the right and left ratio of the elbow angle.
6	μAcc	Average of the hand instantaneous acceleration.
7	μH	Average of the height ratio along the movement.
8	μW	Average of the width ratio along the movement.
9	μROM	Average of the ROM ratio along the movement.

The development of the models calls for a gold standard for result comparison (a ground truth, GT). The patients can be divided into two different classes, the ones that present a reduction of upper-limb mobility and the ones that do not notice any significant difference. In this manner, a binary GT is used based on the diagnosis of patients with and without lymphedema, performed by a medical expert. Different supervised learning algorithms were tested: the Fisher linear discriminant analysis (LDA), the naive Bayes classifier and support vector machines (SVMs). In order to explore all of the possibilities and to compare performances, the models were trained and tested for all of the possible combinations of features, which leads to 2^9 different subsets. In this manner, the model was designed using the selected features, considering all of the possible subsets, by means of a leave-one-out (LOO) scheme [35].

4. Results

4.1. Database

4.1.1. Database Acquisition

The present research depends on the availability of training and testing examples used in the development of the models. Therefore, this research included the collection of training data with the Microsoft Kinect. For this purpose, an acquisition application was developed using the Microsoft SDK. With this application, it is possible to record RGB and depth frames (resolution of 640×480 pixels) each 15 frames per second (fps), as well as information of the skeleton joints. In Figure 6, it is possible to visualize the graphical user interface (GUI) of the developed application, where a color and a depth stream are present. Also, the user is allowed to introduce the patient’s name, pertinent comments and to control the Kinect tilt. Moreover, the complete Kinect skeleton can also be visualized, along with a graphical representation of the X and Y positions of a joint chosen by the user.

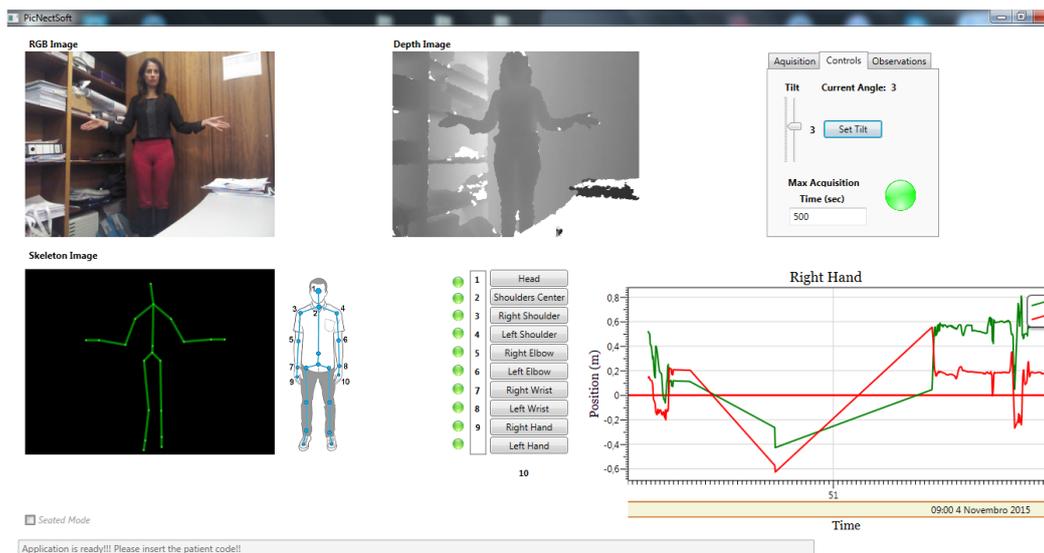


Figure 6. Graphical user interface of the application developed for data acquisition.

During the protocol of acquisition, the patients are placed at two meters from the camera and asked to perform simple movements of abduction and adduction with the upper limbs (see Figure 7). In this manner, the acquired data include several color and depth images with a resolution of 640×480 pixels, as well as the position of the skeleton joints, while performing those exercises. Currently, the database includes 48 patients, from which 24 present lymphedema.

4.1.2. Database Analysis

A ground truth is needed to build and compare the models. In this case, we use the medical diagnosis of lymphedema for this purpose, since we consider that most patients with UBF restrictions are diagnosed with some degree of lymphedema. To validate this assumption, a subjective evaluation of the patients’ upper-body functional status was also performed using the UEFI self-report questionnaire.

This questionnaire consists of 20 items that are rated on a five-point Likert scale and inquires about the patient’s current upper functional status in a variety of activities. UEFI was chosen, since it is relatively simple and fast to fill out and has been demonstrated to be valid and sensitive to changes in the breast cancer population [25]. However, this report has weaknesses when the location of the affected limb is on the opposite side of the dominant hand, since the person never used that arm before to perform some of the tasks evaluated. Moreover, there is minimal information available about the scoring, interpretation and functional implications of the results of this self-report. Nevertheless, since the diagnosis of patients with lymphedema is used as the GT for this purpose, the UEFI score can be useful in this study to validate the reference data used to build the classification models.

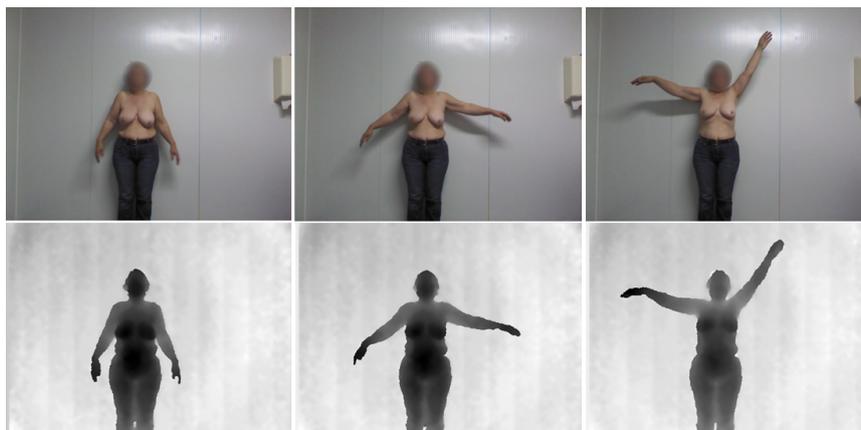


Figure 7. Example of the data acquired using the Kinect device: color and corresponding depth map frame.

Thus, the patients were divided into two classes according to the threshold 0.6 of the UEFI score ratio, defined taking into account the patients’ diagnosis and treatments and resorting to the opinion of a medical expert in breast cancer. In this manner, to better understand the agreement between the evaluation of lymphedema and the UEFI questionnaire, the point-biserial correlation coefficient was assessed (see Equation (4)). This coefficient is a special case of Pearson’s correlation in which one variable is quantitative (ratio) and the other variable is dichotomous, as in the present study.

$$r_{pb} = \frac{M_1 - M_0}{s_n} \sqrt{\frac{n_1 n_2}{n^2}} \tag{4}$$

where s_n is the standard deviation:

$$s_n = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2} \tag{5}$$

M_1 being the mean value on the continuous variable X for all data points in Group 1 and M_0 the mean value on the continuous variable X for all data points in Group 2. Further, n_1 is the number of data points in Group 1; n_0 is the number of data points in Group 2; and n is the total sample size. Point-biserial values range from -1 to 1 , where positive scores closer to one are desirable, since this implies higher reliability. In this fashion, a correlation score of $r_{pb} = 0.538$ was obtained, which means that the disagreement between the self-report and the medical diagnosis is relatively low. This result

confirms the validity of using the diagnosis of lymphedema as reference data, since these patients are the ones that present higher impairments in the upper-body functional status, as implied by the UEFI score.

4.2. Depth Map Noise Reduction

To remove the noise of the unstable boundaries of the depth maps, a bilateral filter was implemented. With this type of filter, a Gaussian kernel and the intensity difference on the pixel are combined to obtain image smoothing, thanks to the domain component of the filter, preserving edges at the same time, thanks to the range component. The filter was tested with different parameters, varying the domain kernel (σ_d), the range kernel (σ_r) and, also, the window size. It was only possible to perform the evaluation of these methods by visual observation of the patients' point cloud (see Figure 8). Nevertheless, very interesting results were obtained with $\sigma_r = 5$, $\sigma_d = 15$ and with a 15×15 window: the filter was capable of smoothing the raw depth data and removed the staircase effect presented in the original depth map. In this way, it is possible to smooth the image and also to preserve the important details. With a small window, the staircase effect is also removed, but the appearance of several artifacts on the depth map is verified. On the other hand, with a σ_r excessively high, important features are over-smoothed.

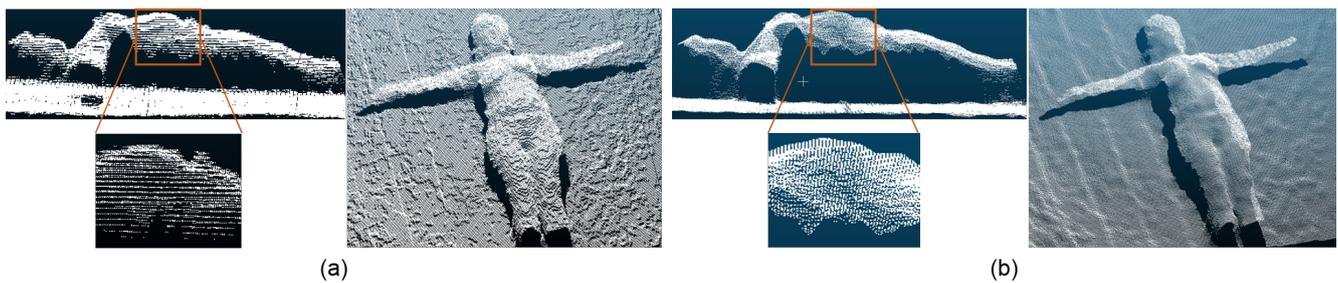


Figure 8. (a) Original raw depth data and (b) the result of the bilateral filter with $\sigma_r = 5$, $\sigma_d = 15$ and a 15×15 window.

4.3. Patient Segmentation

In order to validate the automatic method used for patient segmentation, a manual segmentation was performed in 48 frames that were randomly chosen, using image editor software. Furthermore, a median filter (11×11) was applied for smoothing. Therefore, it was possible to evaluate the automatic method calculating two different similarity indexes: the Dice coefficient and the Jaccard index (see Table 4, Equations (6) and (7)).

$$D(A, B) = \frac{2(A \cap B)}{(A \cap B + A \cup B)} \tag{6}$$

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \tag{7}$$

Moreover, the similarity between the two contours of the masks is evaluated by the Hausdorff and average distance. The Hausdorff distance is defined as the maximum distance of a set to the nearest point in the other set (see Equation (8)). Roughly speaking, it captures the maximum separation between the manual and the automatic contours. Consider the sets of points A and B . First, for each point in A ,

the minimum distance to all points in B is obtained. Then, the directed Hausdorff distance $h(A, B)$ will be the maximum of this set of minimum distances (see Equation (8)).

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\| \tag{8}$$

Table 4. Similarity indexes used to evaluate the body and arm segmentation.

Name	Equation	-
Dice coefficient (D)	Equation (6)	Measures the extent of spatial overlap between two binary images. It gives more weighting to instances where the two images agree. Its values range between 0 (no overlap) and 1 (perfect agreement).
Jaccard index (J)	Equation (7)	Measures similarity between finite sample sets and is defined as the number of attributes shared divided by the total number of attributes present in either of them. Its values range between 0 (no similarity) and 1 (equal).

The results obtained are presented on Tables 5 and 6. Regarding the Dice and Jaccard coefficients, the indexes are almost one, which means that the proposed method presents a high similarity and a high overlap with the ground truth. The average distance is relatively low, and the Hausdorff distance, which represents the worst case scenario, has an mean around four pixels. In Figure 9, it is possible to consult the automatic segmentation and the ground truth for different patients. With the analysis of these results, it is possible to say that the method used appears to be valid for the patient segmentation in Kinect depth images.



Figure 9. Ground truth (GT) and automatic body segmentation examples.

Table 5. Dice coefficient and Jaccard index results of body segmentation and arm segmentation.

-	Body Segmentation		Arm Segmentation	
	Mean	SD	Mean	SD
Dice coefficient (D)	0.999	0.003	0.783	0.048
Jaccard index (J)	0.997	0.005	0.646	0.066

Table 6. Body and arm contour detection error (pixels) evaluated by the Hausdorff (*h*) and average distance (*Avg*).

-		Detected ⇒ GT				GT ⇒ Detected			
		Mean	SD	Max	Min	Mean	SD	Max	Min
Body	<i>h</i>	3.99	7.31	42.76	0	3.60	7.42	38.60	0
	<i>Avg</i>	0.09	0.21	1.28	0	0.11	0.33	2.16	0
Arm	<i>h</i>	12.90	2.80	18.60	6.00	12.70	4.25	22.00	5.83
	<i>Avg</i>	4.82	1.03	7.28	2.67	4.91	1.18	7.20	2.83

4.4. Arm Segmentation

For the validation of the method used for upper-arm detection, the upper arms were also manually segmented to have a GT. As with the patient segmentation, the Dice coefficient and the Jaccard index, as well as the average and Hausdorff distances were computed (see Tables 5 and 6). The similarity indexes present reasonably good results. On average, the Dice coefficient is around 78%, which represents a relatively high overlap between the ground truth and the detected arm, and a similarity of 65%, represented by the Jaccard index. The average distance presents a mean value of almost five pixels, while the Hausdorff distance has a higher value of around 13 pixels.

With these results, it is possible to conclude that, when compared to the body segmentation, this method is more error prone, but still quite satisfactory (see Figure 10). It is important to take into account that the manual segmentation carried out in this case is harder to control and operator dependent, therefore a big source of errors. Furthermore, the automatic detection of the arm points is done based on the Kinect skeleton, which can also influence the final result.

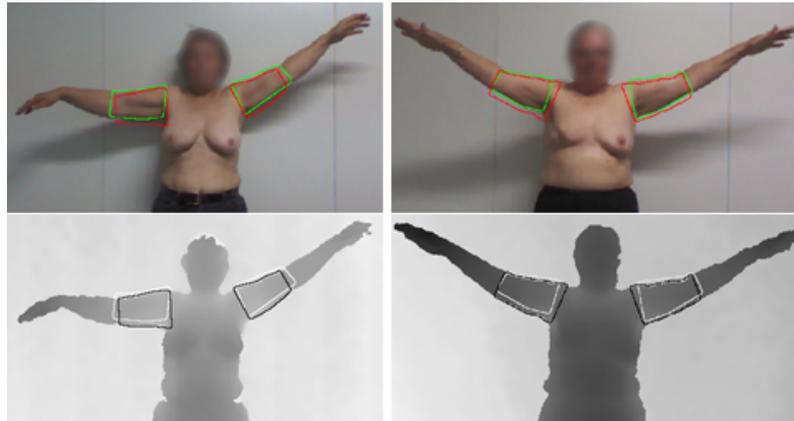


Figure 10. Arm contour detection examples. Detected contour (red/black) and ground truth (green/gray).

4.5. Upper-Body Functional Evaluation

For the creation of a predictive classification model, different supervised learning classifiers were tested: LDA, naive Bayes and SVMs. The models were designed using the extracted features (see Table 3), considering all of the possible subsets, by a LOO scheme [35]. Models considering binary classes were trained, classifying the patients with and without UBF impairments. SVM training was performed with linear, polynomial and radial basis function (RBF) kernels. For all of these cases, exponentially-growing sequences of C were tested, from $C = 2^{-2}$ until $C = 2^6$. For the polynomial kernel, the order varied from two to six, while γ , for the RBF kernel, was tested with: $\gamma = 0.25, 0.5, 0.75, 1$.

The classification results can be consulted in Table 7. For each classifier, the parameters and the set of features that lead to lower misclassification errors are presented. In some of the tests, more than one subset of features obtained the same misclassification error. For those cases, the ones with less complexity are chosen. Furthermore, it is possible to consult the confusion matrix for the models with the best performance (see Table 8), as well as the precision, recall and specificity results for all of the classifiers (see Table 9). The precision, recall and specificity are defined by Equations (9–11), where TP is the true positive values, FP the false positive values, FN the false negative values and TN the true negative values.

$$Precision(P) = \frac{TP}{TP + FP} \tag{9}$$

$$Recall(R) = \frac{TP}{TP + FN} \tag{10}$$

$$Specificity(S) = \frac{TN}{TN + FP} \tag{11}$$

Table 7. Classification results for the different classifiers tested, using the lymphedema diagnosis as the ground truth (GT).

Classifier	Kernel	C	Order	γ	MER	Feat. Set
LDA	-	-	-	-	0.29	[maxROM, maxWidth]
Naive Bayes	-	-	-	-	0.27	[$\mu ElbF$, μAcc , μROM]
SVM	Linear	2^{-1}	-	-	0.25	[maxROM, $\mu ElbF$, μAcc]
SVM	Polynomial	2^3	4	-	0.19	[maxWidth, μROM]
SVM	RBF	2^4	-	0.75	0.19	[μVOL , maxWidth, μROM]

The best performance was verified for the SVM classifier, either with the polynomial or the RBF kernel, with a misclassification error of 0.19. On the other hand, the maximum error was 0.29 for the LDA classifier. As expected, the SVMs classifiers presented better results, since they proved to be a good option when there is a limited amount of data available for training.

Table 8. Confusion matrix for the SVM model, either with a polynomial of an RBF kernel.

True Prediction	Reduced UBF	Normal UBF
Reduced UBF	18	6
Normal UBF	3	21

Table 9. Precision, recall and specificity results for each classifier.

Classifier	Precision	Recall	Specificity
LDA	0.92	0.46	0.96
Naive Bayes	0.92	0.50	0.96
Linear SVM	1.00	0.50	1.00
Polynomial SVM	0.86	0.75	0.88
RBF SVM	0.86	0.75	0.88

Regarding now the precision and recall, it is possible to compare, for all of the models, how many of the positive predictions are actually correct (precision) and how many of the positive-labeled instances were actually matched (recall). High precision values were obtained, which is deduced from a big proportion of positive cases that are correctly assigned. However, it can also be noted that the most common misclassification cases are lymphedema patients that are not considered, implied by a lower recall value. Either way, considering the models with the best performance (polynomial and RBF SVMs), the results are still very promising, since the models have been proven to be capable of performing the

correct classification of patients with reduced UBF, overlooking only six cases. As a result, the method developed appears to be a suitable solution for the functional evaluation of the upper body of breast cancer patients. The most common features used in almost all of the selected models are the *maxW* and the μROM , as well as the *maxROM*, the $\mu ElbF$ and the μAcc . This means that these features are probably the ones that have a greater influence on the functional assessment.

Overall, with this analysis, it was possible to train and test different supervised classification algorithms (LDA, naive Bayes and SVMs) for all of the possible subsets of features, in order to explore all of the possibilities and to compare the performances. Unfortunately, there is no similar work in the literature to compare the approaches, but the results are still very promising. This notwithstanding, in future work, it would be important to increase the precision of this prediction model to reduce FN classifications. Therefore, the methodology proposed appears to be suitable for the evaluation of the upper-body functional status in breast cancer patients, but improvements are still needed.

5. Conclusions

Despite its influence on patients' QOL, there is no clinical standard for the diagnosis of the decreased motion of UBF caused by breast cancer treatment. The most reported methods are operator dependent and non-practical for clinical settings, while incorrect use limits their accuracy, and they are non-reproducible.

The main goal of this research was the study of a Kinect-based method for the diagnosis of the UBF impairments, suitable for breast cancer patients. The work included the collection of a dataset, using the Kinect device, composed of patients with normal and decreased UBF. Using the depth and skeleton data available, it was possible to obtain features of the movement performed by each patient and, therefore, to test the behavior of several supervised learning algorithms, in order to build a predictive classification model. The best classification performance was obtained using the SVM classifier, with a misclassification error of 0.19. Therefore, these findings suggest that the methodology proposed is efficient in the evaluation of the upper-body function. It is important to take into account that there is no similar work in the literature, so it is not possible to compare the results with other methods. Henceforth, more research is needed in order to progress in this direction and to validate the results.

After breast cancer treatment, it is essential for women to maintain continuous physical activity in order to recover the upper-limb mobility. In that way, a home-based exercise program is normally recommended, but the patients do not always perform the exercises as they should. This highlights the importance of a surveillance rehabilitation model for breast cancer patients, to promote and support physical activity and exercise behaviors. In the future, we intend to investigate a rehabilitation model using the previously-mentioned method. Taking advantage of the Kinect device, an application could be developed in order to instruct the patient on how to execute the exercises and to perform an evaluation of their performance.

Acknowledgments

The work leading to these results has received funding from the European Community's Seventh Framework Programme [Grant No. FP7-600948].

Author Contributions

Rita Moreira, André Magalhães and Hélder P. Oliveira conceived of and designed the experiments. Rita Moreira and Hélder P. Oliveira designed the application for data acquisition. Rita Moreira and André Magalhães performed the experiments. Rita Moreira and Hélder P. Oliveira analyzed and processed the data. Rita Moreira wrote the paper under the guidance of Hélder P. Oliveira.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Ferlay, J.; Shin, H.; Bray, F.; Forman, D.; Mathers, C.; Parkin, D. GLOBOCAN 2008 v2.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 10. Available online: <http://globocan.iarc.fr> (accessed on 27 October 2010).
2. Goyal, A.; Newcombe, R.; Chhabra, A.; Mansel, R. Morbidity in Breast Cancer Patients with Sentinel Node Metastases Undergoing Delayed Axillary Lymph Node Dissection (ALND) Compared with Immediate ALND. *Ann. Surg. Oncol.* **2008**, *15*, 262–267.
3. Donker, M.; van Tienhoven, G.; Straver, M.E.; Meijnen, P.; van de Velde, C.J.H.; Mansel, R.E.; Cataliotti, L.; Westenberg, A.H.; Klinkenbijn, J.H.G.; Orzalesi, L.; *et al.* Radiotherapy or surgery of the axilla after a positive sentinel node in breast cancer (EORTC 10981-22023 AMAROS): A randomised, multicentre, open-label, phase 3 non-inferiority trial. *Lancet Oncol.* **2014**, *15*, 1303–1310.
4. Nesvold, I.; Reinertsen, K.; Fossa, S.; Dahl, A. The relation between arm/shoulder problems and quality of life in breast cancer survivors: A cross-sectional and longitudinal study. *J. Cancer Surviv.* **2011**, *5*, 62–72.
5. Stanton, A.; Krishnan, L.; Collins, C. Form or function? Part 1. Subjective cosmetic and functional correlates of quality of life in women treated with breast-conserving surgical procedures and radiotherapy. *Cancer* **2001**, *91*, 2273–2281.
6. Oliveira, H.P.; Cardoso, J.S.; Magalhaes, A.; Cardoso, M.J. Methods for the aesthetic evaluation of Breast Cancer Conservation Treatment: A Technological Review. *Curr. Med. Imaging Rev.* **2013**, *9*, 32–46.
7. Cardoso, M.J.; Oliveira, H.P.; Cardoso, J.S. Assessing Cosmetic Results after Breast Conserving Surgery. *J. Surg. Oncol.* **2014**, *110*, 37–44.
8. Bulley, C.; Coutts, F.; Blyth, C.; Jack, W.; Chetty, M.; Barber, M.; Tan, C. Prevalence and Impacts of Upper Limb Morbidity after Treatment for Breast Cancer: A Cross-Sectional Study of Lymphedema and Function. *Cancer Oncol. Res.* **2013**, *1*, 30–39.
9. McNeely, M.; Campbell, K.; Ospina, M.; Rowe, B.; Dabbs, K.; Klassen, T.; Mackey, J.; Courneya, K. Exercise interventions for upper-limb dysfunction due to breast cancer treatment. *Cochrane Database Syst. Rev.* **2010**, *6*, doi:10.1002/14651858.CD005211.pub2.
10. De Backer, I.; Schep, G.; Backx, F.; Vreugdenhil, G.; Kuipers, H. Resistance Training in Cancer Survivors: A Systematic review. *Int. J. Sports Med.* **2009**, *30*, 703–712.

11. Cheema, B.; Gaul, C.; Lane, K.; Singh, M. Progressive resistance training in breast cancer: A systematic review of clinical trials. *Breast Cancer Res. Treat.* **2008**, *109*, 9–26.
12. McNeely, M.; Campbell, K.; Rowe, B.; Klassen, T.; Mackey, J.; Courneya, K. Effects of exercise on breast cancer patients and survivors: A systematic review and meta-analysis. *Can. Med. Assoc. J.* **2006**, *175*, 34–41.
13. Speck, R.; Courneya, K.; Mâsse, L.; Duval, S.; Schmitz, K. An update of controlled physical activity trials in cancer survivors: A systematic review and meta-analysis. *J. Cancer Surviv.* **2010**, *4*, 87–100.
14. Moreira, R.; Magalhães, A.; Oliveira, H.P. A Kinect-Based System to Assess Lymphedema Impairments in Breast Cancer Patients. In Proceedings of the 7th Iberian Conference on Pattern Recognition and Image Analysis, Santiago, Spain, 17–19 June 2015; pp. 228–236.
15. Hayes, S.; Janda, M.; Cornish, B.; Battistutta, D.; Newman, B. Lymphedema After Breast Cancer: Incidence, Risk Factors, and Effect on Upper Body Function. *J. Clin. Oncol.* **2008**, *26*, 3536–3542.
16. Armer, J.; Radina, M.; Porock, D.; Culbertson, S. Predicting Breast Cancer-Related Lymphedema Using Self-Reported Symptoms. *Nurs. Res.* **2003**, *52*, 370–379.
17. Bulley, C.; Gaal, S.; Coutts, F.; Blyth, C.; Jack, W.; Chetty, M.; Barber, M.; Tan, C. Comparison of Breast Cancer-Related Lymphedema (Upper Limb Swelling) Prevalence Estimated Using Objective and Subjective Criteria and Relationship with Quality of Life. *BioMed Res. Int.* **2013**, *2013*, doi:10.1155/2013/807569.
18. Lu, G.; DeSouza, G.; Armer, J.; Anderson, B.; Shyu, C. A system for limb-volume measurement using 3D models from an infrared depth sensor. In Proceedings of IEEE Symposium on Computational Intelligence in Healthcare and e-Health, Singapore, Singapore, 16–19 April 2013; pp. 64–69.
19. Taylor, R.; Jayasinghe, U.; Koelmeyer, L.; Ung, O.; Boyages, J. Reliability and validity of arm volume measurements for assessment of lymphedema. *Phys. Ther.* **2006**, *86*, 205–214.
20. Stanton, A.; Badger, C.; Sitzia, J. Non-invasive assessment of the lymphedematous limb. *Lymphology* **2000**, *33*, 122–135.
21. Pero-System Messgerate GmbH. Available online: <http://www.pero-system.de/> (accessed on 3 November 2015).
22. Dixon, A.; Wheeler, T.; Lomas, D.; Mackenzie, R. Computed tomography or magnetic resonance imaging for axillary symptoms following treatment of breast carcinoma? A randomized trial. *Clin. Radiol.* **1993**, *48*, 371–376.
23. Vukotich, C.; Geyer, M.; Erdeljac, F. Use of a Laser Scanning System to Measure Limb Volume in Chronic Edema. Available online: http://www.resna.org/sites/default/files/legacy/conference/proceedings/2011/RESNA_ICTA/vukotich-69743.pdf (accessed on 3 November 2015).
24. Trombetta, C.; Abundo, P.; Felici, A.; Ljoka, C.; di Cori, S.; Rosato, N.; Foti, C. Computer Aided Measurement Laser (CAML): Technique to quantify post-mastectomy lymphoedema. *J. Phys. Conf. Ser.* **2012**, *383*, doi:<http://dx.doi.org/10.1088/1742-6596/383/1/012018>.
25. Campbell, K.; Pusic, A.; Zucker, D.; McNeely, M.; Binkley, J.; Cheville, A.; Harwood, K. A prospective model of care for breast cancer rehabilitation: Function. *Cancer* **2012**, *118*, 2300–2311.

26. Stratford, P. Assessing Disability and Change on Individual Patients: A Report of a Patient Specific Measure. *Physiother. Can.* **1995**, *47*, 258–263.
27. Hudak, P.; Amadio, P.; Bombardier, C.; Beaton, D.; Cole, D.; Davis, A.; Hawker, G.; Katz, J.; Makela, M.; Marx, R.G.; *et al.* Development of an upper extremity outcome measure: The DASH (disabilities of the arm, shoulder, and head). *Am. J. Ind. Med.* **1996**, *29*, 602–608.
28. Stratford, P.; Binkley, J.; Stratford, D. Development and initial validation of the upper extremity functional index. *Physiother. Can.* **2001**, *53*, 259–267.
29. Kwan, W.; Jackson, J.; Weir, L.; Dingee, C.; McGregor, G.; Olivotto, I. Chronic Arm Morbidity After Curative Breast Cancer Treatment: Prevalence and Impact on Quality of Life. *J. Clin. Oncol.* **2002**, *20*, 4242–4248.
30. Freitas-Silva, R.; Conde, D.; Freitas-Júnior, R.; Martinez, E. Comparison of quality of life, satisfaction with surgery and shoulder-arm morbidity in breast cancer survivors submitted to breast-conserving therapy or mastectomy followed by immediate breast reconstruction. *Clinics* **2010**, *65*, 781–787.
31. Hayes, S.; Rye, S.; Battistutta, D.; DiSipio, T.; Newman, B. Upper-body morbidity following breast cancer treatment is common, may persist longer-term and adversely influences quality of life. *Health Qual. Life Outcomes* **2010**, *8*, 92–98.
32. Tomasi, C.; Manduchi, R. Bilateral filtering for gray and color images. In Proceedings of the Sixth International Conference on Computer Vision, Bombay, India, 4–7 January 1998.
33. Oliveira, H.; Cardoso, J.; Magalhaes, A.; Cardoso, M. A 3D low-cost solution for the aesthetic evaluation of breast cancer conservative treatment. *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.* **2014**, *2*, 90–106.
34. Drake, R.; Vogl, A.; Mitchell, A. *Gray's Anatomy for Students*; Churchill Livingstone: London, UK, 2009.
35. Duda, R.O.; Hart, P.E.; Stork, D.G. *Pattern Classification*, 2nd ed.; Wiley Interscience Publication: Hoboken, NJ, USA, 2000.