AN AUTOMATIC VISION-BASED MONITORING SYSTEM FOR ACCURATE VOJTA-THERAPY

Muhammad Hassan Khan, Julien Helsper, Cong Yang and Marcin Grzegorzek

Research Group for Pattern Recognition, University of Siegen, Germany.

ABSTRACT

Vojta-therapy is a useful technique to treat the disorders in the central nervous and musculoskeletal system. During the therapy, a specific stimulation is given to the patients in order to cause the patient's body to perform certain reflexive pattern movements. The repetition of this stimulation ultimately brings forth the previously blocked connections between the spinal cord and brain, and after a few sessions, patients can perform these movements without any external stimulation. In this paper we proposed an automatic infant's (i.e., patient) detection and recognition of specific movements in various body parts during the therapy process, using RGB-D data. First, A robust template matching based algorithm is exploited for infant's detection using his/her head location. Second, various features are computed to capture the movements of different body parts during the therapy. In the classification stage, a multi-class support vector machine (mSVM) is used to classify the accurate movements of infant during the therapy process, which ultimately reveals the correctness of the given treatment. The proposed algorithm is evaluated on our challenging dataset, which was collected in a children hospital. The detection and classification results show that the proposed method is highly useful to recognize the correct movement pattern either in hospital or in-home therapy systems.

Index Terms— Template matching, Region growing, Vojta-therapy, Movement pattern, Microsoft Kinect

1. INTRODUCTION

Vojta-therapy (VT) is based on the principle of reflex locomotion, i.e., the patient's central nervous system can be activated by giving the correct simulation, assuming that it is still partially intact. Reflex locomotion is a combination of Reflex Creeping in prone lying position and Reflex Rolling from a supine and side lying positions, which enable the elementary patterns of movement in patients. According to vojta, one can observe motor reactions occurring throughout the patient's body when a specific stimulation is given to him/her, while lying in one of the above mentioned positions. Therefore, the therapists exploit a combination of 10 different zones on a human's body by putting light pressure on this area and resistance to the current movement (e.g., the tendency to rotate the head during reflex creeping) to cause the patient's body to perform certain reflexive movement patterns. Repeating this activity many times, previously blocked connections between the patient's spinal cord and brain become available, and the patient is able to perform similar movements without any external stimulation. VT has been effectively used to treat various diseases like cerebral palsy, peripheral paralysis of the arms/legs, hip joint dysplasia and others. More detail about the VT is available in [1].

VT can be applied to patients of any age group but it is extremely effective for young babies of less than 6 months because most of the developmental changes take place in the early stage of a child's life. For a treatment to be successful, the therapy session of 5-20 minutes should be performed several times in a day or week over a time period of several weeks or months. Therefore, the therapists may recommend an in-home continuation of the therapy. An automatic vision based system is required to analyze the accurate movements of a patient (i.e., correct treatment), during the therapy session. The aim of home-based therapy evaluation is to provide an accurate in-home therapy alternative to in-hospital therapy. The therapy process at home is not only helpful for the quick recovery of a patient but also quite useful for those who do not have access to a local hospital offering said treatment. Moreover, the frequent visits to the therapist's clinic add an economic burden as well.

In this paper, we proposed an automatic method for infant's detection and accurate recognition of the movement patterns in his/her body parts during VT, using RGB-D data. The proposed method operates in two steps. First, it searches a pre-defined infant's head template in depth image using the sum of square differences (SSD) and cross correlation (CC) based matching algorithms in an efficient way. The detected head locations are fitted into a 3D head model for verification. We apply a region growing algorithm on verified locations to extract the entire infant's body region from the scene. Second, the recorded movements are analyzed and classified using a multi-class Support Vector Machine (mSVM). During therapy, a specific movement can be observed in the upper and lower limbs of an infant while lying in one of the aforementioned positions. We computed various features to capture the movements of upper and lower limbs and an mSVM is ex-

ploited to classify the accurate movements, which ultimately leads us to the correctness of the given treatment. The proposed method is evaluated on our challenging dataset, which was collected in the children hospital [2]. The dataset contains more than 8,000 RGB and depth frames from Kinect. Experimental results show that the proposed method performs rather well and is highly useful for at home-based therapy systems.

2. RELATED WORK

2.1. Human Detection

A large number of algorithms have been proposed for human detection in RGB and depth images. Numerous RGB based techniques generally used either object features such as histogram of oriented gradient (HoG) descriptors [3] or motion features like spatio-temporal interest points (STIPs) [4]. Some other techniques exploited shape-based features [5] and the interest points extraction from the scene such as scale-invariant feature transform (SIFT) [6]. Although most of these approaches claim high accuracy in human-detection, visual data still remains a challenging task due to difficulties caused by illumination changes, occlusions, as well as complex and cluttered background. Alternatively, the depth data has several advantages over the visual data as it provides 3D structural information of the scene. Moreover, the layered structure of the depth data offers more details and important cues for human detection and their action recognition.

Due to these advantages, many algorithms for human detection have also been proposed in depth images. Zhu et al. [7] proposed a human detection by tracking his/her head and torso in depth images. They proposed the circle and box fitting to detect the head and torso respectively. In [8] a support vector machine has been trained to discriminate the human head and shoulders for detection purposes. The human upper body template is proposed in [9] to slide over the depth image and detect the human region using the Euclidean distance. Xia et al. [10] proposed a human-head template for the detection of a human in depth images using the 2D chamfer matching method. Stahlschmidt et al. [11] proposed a human detection technique using a matched filter known as Mexican hat wavelet to the segmented foreground information of depth image. This technique is only applicable when the camera is mounted in top-view position. Compared to these methods, the proposed algorithm does not require any statistical learning, motion information, and is computationally inexpensive. The proposed method exploits a fast template matching algorithm to detect the infant's body region using his/her head location and segments it with more accuracy.

2.2. Rehabilitation

Industrial motion sensors have been utilized as a physical rehabilitation tool [12, 13]. However, these methods require wearing a number of sensors on the human body, causing discomfort. Virtual reality and motion-based games have also been exploited for rehabilitation [14, 12]. Admittedly, the evidence indicates that these methods provide an interactive, engaging, and effective environment for physical therapy, but they require expensive hardware and software. They are generally designed to suit a very specific group of patients and cannot be useful in cases of newly born babies because they cannot interact with such systems. With the invention of a very low cost device, i.e., Microsoft Kinect, a new tool is now to be considered in similar rehabilitation, as well as in assessment and monitoring systems. A Kinect camera is quite cheap and can be extended far beyond gaming [15].

Several researchers have utilized the Kinect as assistive technology at ambient assisted living and rehabilitation places. In [16], authors proposed a system for in-home rehabilitation using Dynamic Time Warping (DTW) algorithm and fuzzy logic. The evaluation is performed on the trajectory of joints and the time duration in the completion of designated exercise. Exell et al. [17] utilized the skeletal tracking information to analyze the rehabilitation in the upper limbs. Chang et al. [18] proposed a system that used the motion tracking of 6 upper limbs for rehabilitation. For validation they used the outputs of OptiTrack as ground truth and compared them to the outputs of Kinect. Yao-Jen et al. [19] developed a Kinect-based rehabilitation system to assist the therapists in their work to treat the children suffering from motor disabilities. They used the motion tracking data to analyze the rehabilitation standards and to allow the therapist to view the rehabilitation progress. Gama et al. [20], proposed an interface for adults to monitor the the correct description of therapeutic movements. The recent surveys on various therapy techniques for rehabilitation using Kinect are presented in [21, 22]. All of the above mentioned techniques used skeleton information from Kinect, which either could not be made available in the case of infants or which was problematic if some body parts were occluded. Moreover, due to the nature of VT, some of the patient's body parts are always occluded, as shown in Fig. 4-5.

Compared to the aforementioned methods, the proposed algorithm exploits only depth information for infant's detection and used the segmented body region in combination with visual information to classify the accurate movements in their various body parts.

3. OVERVIEW OF THE PROPOSED METHOD

This section provides an overview of the major steps in the proposed method. The detailed implementation of each step is presented in Section 4.

First, the infants are detected in depth data using the location of their head position. In a given 2D array of depth data from Microsoft Kinect, we reduced the noise and smoothed the array for later process. The edges information in the depth

Fig. 1: Overview of the proposed method

data is utilized to locate the candidate positions of the infant's region. The proposed method scans across the whole image and exploits the relationship between the sum of squared differences (SSD) and cross correlation (CC) based matching algorithms, to locate the possible regions of the infant's head. Each of the detected head region is verified through a 3D head-model, which utilized both the edge and relational depth change information from original depth array. After verification, a Region Growing algorithm is exploited on the detected location to segment the infant's body region from the captured scene. We performed the calibration of depth and RGB images using the camera's extrinsic and intrinsic parameters to segment the infant's body region from RGB image as well.

Second, We extract the various features in the segmented infant's body region to capture the movements in upper and lower limbs. The visual segmented information is utilized to identify the infant's lying position during the treatment process. In the classification stage, a muti-class support vector machine is used to classify the accurate movements of the infant during the treatment.

4. PROPOSED METHOD

4.1. Infant's Detection

The 3D depth sensor of the Kinect camera provides the depth information of the captured scene as a two-dimensional (2D) array of pixel values, known as depth map. The sensor can capture the distance information in the range of 0.8-3.5 meters. In the scene, all the locations where the sensor is unable to measure the depth, are filled with the offset value zero in the array. These pixels can be viewed as random black spots in the depth image. We considered these pixels as noise and had to recover their true depth values in order to avoid their interference. Keeping in mind that an image is a continuous space, the depth value of these missing pixels are recovered from the neighboring pixels using the nearest neighbor interpolation algorithm. Then, a median filter of size 4×4 is applied on the depth array to make the data smoother.

The next task after the aforementioned pre-processing (i.e., recovery of missing pixels) is to locate the position of the infant's head in a depth map. The edges information of the depth map is utilized for this purpose. A Canny edge detector is used to find all the edges in the depth map and the edges smaller than a pre-defined threshold are dropped to reduce the calculation and the disturbance from the irregular objects in the surrounding. The binary template used for the infant's head detection, is shown in Fig. 2(c). For infant's head detection, the template is traversed over the entire edge image from top-left to bottom-right. The matching measure is determined on each pixel of the edge image which lies under the template. For the matching of the infant's head, we utilized the SSD and CC based matching algorithms. Let us define the edge image as E and template image T with size $h \times w$. The equation of SSD can be written as:

$$S(x,y) = \sum_{i=0}^{h}\sum_{j=0}^{w}\Big(E(x+i,y+j) - T(i,j)\Big)^2, \quad (1)$$

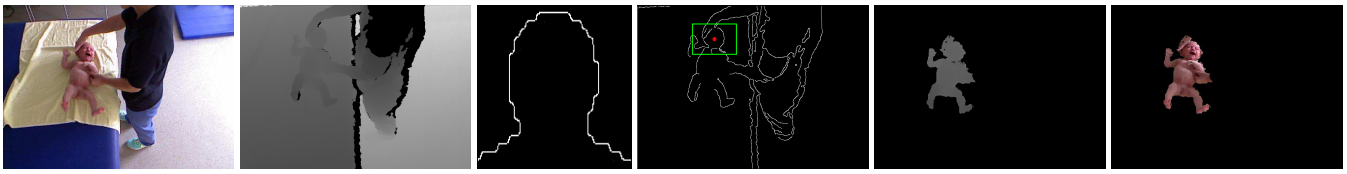
Where S represents the sum of squared difference, computed at (x,y) position of E and i,j representing the x,y locations of pixels in T. SSD can also be viewed as the Euclidean distance between image patch E and template T. Expanding Eq. (1) yields,

$$\begin{aligned} S(x,y) = &\sum_{i=0}^{h}\sum_{j=0}^{w}E^2(x+i,y+j) + \sum_{i=0}^{h}\sum_{j=0}^{w}T^2(i,j) \\ &- 2\sum_{i=0}^{h}\sum_{j=0}^{w}E(x+i,y+j)T(i,j) \end{aligned} \quad (2)$$

In Eq. (2), the first term is the sum of squared values in the edge image, the second term belongs to the template image and the third term is twice the correlation between the image patch and template. Note that the term $\sum_{i=0}^{h}\sum_{j=0}^{w}T^2(i,j)$ is constant. Let us assume that the term $\sum_{i=0}^{h}\sum_{j=0}^{w}E^2(x+i,y+j)$ is approximately constant, so the remaining term (i.e., cross correlation) is,

$$C(x,y) = \sum_{i=0}^{h}\sum_{j=0}^{w}E(x+i,y+j)T(i,j), \quad (3)$$

where C represents the cross correlation. It can be easily observed from Eq. (2) that the Euclidean distance between the patch of edge image and template decreases as the correlation (i.e., similarity) between the image patch and template increases. This also provides an intuition to use correlation to measure the similarity. The regions of image where the correlation is high are the places where template and image match well. The cross correlation can be computed by taking the inner product of image patch and template. This calculation is computationally quite expensive [23]. For a search area of size $M \times M$ and a template of size $N \times N$ require $N^2(M-N+1)^2$ additions and $N^2(M-N+1)^2$ multiplications [24]. In order to reduce the computational overhead, we compute the cross correlation in frequency domain



(a) (b) (c) (d) (e) (f)

Fig. 2: Example of infant detection (a) color image (b) depth image (c) template used for head detection, (d) detected head's location after verification (e) and (f) segmented infant's body region in depth and RGB images respectively.

using the Fourier frequency transformation (FFT) and correlation theorem. The theorem states that multiplying the Fourier transform of one function (i.e., our edge image) by the complex conjugate of the Fourier transform of the other (i.e., the head template) gives the Fourier transform of their correlation. That is,

$$C(u,v) = \mathcal{F}^{-1}\Big(\mathcal{F}^{*}(E(x,y))\mathcal{F}(T(i,j))\Big), \qquad (4)$$

where $\mathcal{F}$ represents the Fourier transform, $\mathcal{F}^{*}$ represents the complex conjugate and $\mathcal{F}^{-1}$ is the inverse Fourier transform. The computational cost in frequency domain with the same aforementioned search area and template size is, $12M^2log_2M$ real multiplications and $18M^2log_2M$ real additions [24]. Although there is no difference in the detection results of the proposed method either in computing correlation in spatial domain or frequency domain, it is shown in [23] that correlation computation in frequency domain is 2.5 to 12 times faster than in spatial domain.

The size of the head varies from child to child and also depends on the distance of the child from the camera during the therapy. Although the distance and angle orientation of the camera from the table surface (where the patient is lying for treatment) is defined to capture the best quality data with minimum occlusion, we apply the proposed matching algorithm in a multi-resolution fashion which makes the algorithm robust to scale in change. Three different sizes of templates are exploited with a sampling rate of $\frac{1}{4}$ for pyramids construction, to detect head location in depth image. Using the above mentioned method, few locations with high correlations are marked as a possible infant's head. However, all the detected locations do not necessarily contain the infant's head. Therefore, the detected locations are verified through a 3D head model fitting technique. A 3D hemisphere model is used for verification of the detected head regions. We exploited the 3D head model fitting technique proposed in [10] for this purpose. Once a detected infant's head is confirmed, we applied Region Growing algorithm on the verified head location to extract the infant's body region from the depth image. The Region Growing algorithm compares the similarity of a pixel to its neighboring pixels and grows itself by including them in the region if the similarity difference is lower than the predefined threshold. As a result, we segmented a child from the captured scene, as shown in Fig. 2(e).

Next, we performed the calibration of depth image with RGB image using Kienct's extrinsic and intrinsic parameters and segmented the infant's body region in RGB image as well, which is required in the next stage of our proposed method.

4.2. Pose Representation

The therapy is given to infants in prone, supine, and side lying positions. In our work we primarily focus on the treatment in prone and supine lying positions. We carefully captured the RGB and depth frames containing the movements of upper and lower limbs, while lying in one of the aforementioned positions. It is very difficult to detect the lying position of an infant during the therapy process using only depth data. Therefore, we utilized the calibrated RGB segmented region for this purpose. To classify the movements of upper and lower limbs, we divided the dataset into four classes. Upper limbs movement in supine lying position, lower limbs movement in supine lying position, upper limbs movement in prone lying position and lower limbs movement in prone lying position. Let us define these classes as N_1, N_2, N_3 and N_4 respectively. The movement patterns are described using the 9-dimensional features vector f. The first attribute f_1 is the lying position of an infant, which can be detected using a face detection algorithm on segmented RGB image. For the other attributes, we used the bounding box of the whole segmented region from depth image in two equally high sub-boxes to define the movements in upper and lower regions separately in a more discriminative way (see Fig. 3). Let us define the subscript u is for the boundary box in the upper region and l is for the boundary box in the lower region. The remaining attributes are computed as follows:

$$f_2 = \frac{h_u}{w_u} \qquad f_3 = \frac{h_l}{w_l} \qquad f_4 = \frac{h}{w}$$
$$f_5 = A_u \qquad f_6 = A_l \qquad f_7 = A_u + A_l$$
$$f_8 = L_u \qquad f_9 = L_l$$

where h_u, h_l are the height, w_u, w_l are the width, L_u, L_l are the length of the contours and A_u, A_l represent the area of the extracted silhouette in upper and lower bounding boxes respectively, as shown in Fig. 3. The purpose for the selection of these features is to capture the movements of upper and lower limbs during the treatment process. Moreover, we have scaled each feature attribute to the range [-1, +1]. The main advantage of scaling is to avoid elements in greater numeric

ranges dominating those in smaller numeric ranges. Another advantage is to avoid the numerical difficulties during the calculation, for example in SVM, kernel values usually depend on the inner products of feature vectors (e.g., the linear kernel and the polynomial kernel), large attribute values might cause numerical problems [25].

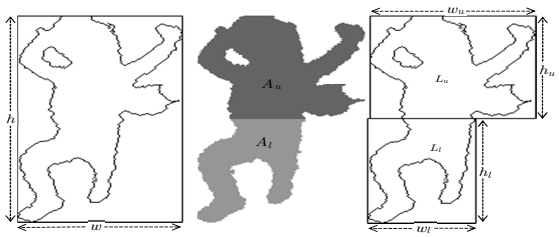


Fig. 3: Example of feature computation. Shape bounding box and equally high sub-boxes (i.e., $h_u = h_l$) used for feature extraction.

5. EXPERIMENTS AND RESULTS

We captured more than 8,000 frames containing the movements of 10 children of various ages (2 weeks to 6 months) and both genders (8 males and 2 females). The camera was put on a tripod, at the height of 2 meters and with an angle of 45° from the table surface (where the infant is lying for therapy). The setting was chosen to be in accordance with the recommendation for capturing the best data quality (i.e., movements patterns in upper and lower limbs) with minimum occlusion. For each patient the therapy session usually lasted between 15-20 minutes, both the RGB and depth frames are recorded at a 640×480 resolution.

5.1. Detection Performance

The proposed detection algorithm is applied on our collected challenging dataset for infant's detection using only depth information. We achieved the detection rate of 83.29 % with a small false alarm of 1.27%. The reason for the drop in detection rate is due to the occlusion at the location of the infant's head. Although it was pre-decided with the therapists to use their hands with the minimum occlusion of the infant's head structure, the situation was particularly often faced when the treatment was given in supine lying position, as shown in Fig. 4(b).

5.2. Classification

In this paper, we used the implementation of SVM in the $LIBSVM$ library [26]. We used the C-Support Vector Classification formulation of the library with the soft margin parameter C and, γ when RBF kernels are used as meta-parameter. A 5-fold cross-validation is performed to validate

(a) (b)

Fig. 4: Example of head poses during the therapy. (a) an infant with clear head structure in prone lying position, (b) occlusion in the infant's head structure can be observed in supine lying position due to the therapist's hand.

the model with the selection of the above mentioned meta-parameters prior to it training the actual model on the full training dataset. Tab. 1 summarized the classification results of SVM for accurate movements in upper and lower limbs during the treatment. One can see that the movement of lower limbs in prone position is the set of activities (i.e., N_4) where the most confusion occurs because the required portion of the lower limbs is occluded by the therapist's arm. The lying position of an infant, however, is detected but the estimation of lower limbs movement might become unstable due to the therapy technique requirement, as shown in Fig. 5(d). Obtaining a classification rate of more than 75% lets us believe in this solution and encourages us to plan future work in this research area by considering some other features to identify the correct patterns of movement.

(a) (b) (c) (d)

Fig. 5: Example of an infant's limbs movement during therapy

Table 1: Classification rate of the proposed algorithm

	N_1	N_2	N_3	N_4	Accuracy(%)
N_1	354	46	0	0	88.50
N_2	70	329	1	0	82.25
N_3	1	3	301	95	75.25
N_4	1	0	176	223	55.75
μ					75.44

6. CONCLUSION AND FUTURE RESEARCH

In this paper, we proposed an automatic method for infant's detection and accurate recognition of movement patterns during VT. The proposed method works in two steps. First, we detect and segment an infant's body region from

depth image using his/her head location. Second, the recorded movements in the infant's body are analyzed and classified using SVM. The experimental results show that the proposed method performs well and can be used for in-home based vojta-therapy systems to recognize the accuracy of the treatment. In future, we plan to implement the complete algorithm in depth data by including a few other body parts of an infant for detection, utilize the layered structure of depth date to identify the infant's lying position and to include a few other motion features of an infant's limbs like trajectories for accurate classification of movement patterns.

7. REFERENCES

[1] "Vojta-therapy," `http://www.vojta.com`, Accessed: 2015-12-30.

[2] "Red cross children's hospital, siegen, germany," `http://www.drk-kinderklinik.de`, Accessed: 2015-12-30.

[3] C. Thurau, "Behavior histograms for action recognition and human detection," in *Human Motion–Understanding, Modeling, Capture and Animation*, 2007.

[4] L. Gorelick, M. Blank, E. Shechtman, M. Irani, and R. Basri, "Actions as space-time shapes," in *IEEE Tran. on Pattern Analysis and Machine Intelligence*, 2007, pp. 2247–2253.

[5] C.R. Wren, A. Azarbayejani, T. Darrell, and A.P. Pentland, "Pfinder: real-time tracking of the human body," in *IEEE Tran. on Pattern Analysis and Machine Intelligence*, 1997, pp. 780–785.

[6] DG. Lowe, "Object recognition from local scale-invariant features," in *the Proc. of the 7th IEEE Int. Conf. on Computer vision*, 1999, pp. 1150–1157.

[7] Y. Zhu and K. Fujimura, "Constrained optimization for human pose estimation from depth sequences," in *IEEE Asian Conf. on Computer Vision*, 2007, pp. 408–418.

[8] M. Rauter, "Reliable human detection and tracking in top-view depth images," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition Workshops*, 2013, pp. 529–534.

[9] O. Hosseini Jafari, D. Mitzel, and B. Leibe, "Real-time rgb-d based people detection and tracking for mobile robots and head-worn cameras," in *IEEE Int. Conf. on Robotics and Automation*, 2014.

[10] L. Xia, CC. Chen, and JK Aggarwal, "Human detection using depth information by kinect," in *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition Workshops*, 2011, pp. 15–22.

[11] C. Stahlschmidt, A. Gavriilidis, J. Velten, and A. Kummert, "Applications for a people detection and tracking algorithm using a time-of-flight camera," in *Multimedia Tools and Applications*, 2014, pp. 1–18.

[12] S. Arteaga et al., "Low-cost accelerometry-based posture monitoring system for stroke survivors," in *Proc. of Int. ACM SIGACCESS Conf. on Computers and accessibility*, 2008, pp. 243–244.

[13] C.H. Shih, M.L. Chang, and C.T. Shih, "A limb action detector enabling people with multiple disabilities to control environmental stimulation through limb action with a nintendo wii remote controller," in *Research in Developmental Disabilities*, 2010, pp. 1047–1053.

[14] D. Jack et al., "A virtual reality-based exercise program for stroke rehabilitation," in *Proc. of Int. ACM conference on Assistive technologies*, 2000, pp. 56–63.

[15] J. Han, L. Shao, D. Xu, and J. Shotton, "Enhanced computer vision with microsoft kinect sensor: A review," in *IEEE Tran. on Cybernetics*, 2013.

[16] C. Bryanton et al., "Feasibility, motivation, and selective motor control: virtual reality compared to conventional home exercise in children with cerebral palsy," *Cyberpsychology & behavior*, vol. 9, pp. 123–128, 2006.

[17] T. Exell et al., "Goal orientated stroke rehabilitation utilising electrical stimulation, iterative learning and microsoft kinect," in *IEEE Int. Conf. on Rehabilitation Robotics*, 2013, pp. 1–6.

[18] C. Chang et al., "Towards pervasive physical rehabilitation using microsoft kinect," in *IEEE 6th Int. Conf. on Pervasive Computing Technologies for Healthcare*, 2012, pp. 159–162.

[19] Y. Chang, S. Chen, and J. Huang, "A kinect-based system for physical rehabilitation: A pilot study for young adults with motor disabilities," *Research in developmental disabilities*, vol. 32, no. 6, pp. 2566–2570, 2011.

[20] A. D. Gama et al., "Guidance and movement correction based on therapeutics movements for motor rehabilitation support systems," in *IEEE 14th Symposium on Virtual and Augmented Reality*, 2012, pp. 191–200.

[21] A. D. Gama et al., "Motor rehabilitation using kinect: A systematic review," *Games for Health Journal*, vol. 4, no. 2, pp. 123–135, 2015.

[22] H. Mousavi Hondori and M. Khademi, "A review on technical and clinical impact of microsoft kinect on physical therapy and rehabilitation," *Journal of Medical Engineering*, vol. 2014.

[23] D. Lyon, "The discrete fourier transform, part 6: Cross-correlation," in *Journal of Object Technology*, 2010, vol. 9, pp. 18–22.

[24] J. P. Lewis, "Fast normalized cross-correlation," in *Vision interface*, 1995, vol. 10, pp. 120–123.

[25] Hsu, Chih-Wei, C. Chang, and Chih-Jen Lin, "A practical guide to support vector classification," 2003.

[26] Chang, Chih-Chung, and C. Lin, "Libsvm: A library for support vector machines," in *ACM Tran. on Intelligent Systems and Technology*, 2011.