Technological module for unsupervised, personalized cardiac rehabilitation exercising

Anargyros Chatzitofis*, Dimitrios Zarpalas*, Dimitrios Filos†, Andreas Triantafyllidis†, Ioanna Chouvarda†, Nicos Maglaveras†, Petros Daras*

*Information Technologies Institute / Centre for Research and Technology Hellas - 6th km Charilaou-Thermi, Thessaloniki, 57001
†Institute of Applied Biology / Centre for Research and Technology Hellas - 6th km Charilaou-Thermi, Thessaloniki, 57001

Abstract—Cardiac Rehabilitation (CR) can significantly improve mortality and morbidity rates from Cardiovascular Diseases (CVD). Nevertheless, traditional CR is diminished by low subsequent adherence rates. Thus, in this paper, an e-Health technological module for human motion analysis and user modelling is proposed, in order to address the requirements of unsupervised, tele-rehabilitation systems for CVD, by evaluating and personalizing prescribed physical CR programs. The proposed module consists of a) an exercise capturing and evaluation component, and b) a user modelling and decision support system for personalization of cardiac rehabilitation programs. In particular, the module monitors and analyses the body movements of the patient when exercising in real-time, while based on this analysis and the heart-rate measurements, it is capable of short-term and long-term CR session adaptation. The proposed module constitutes a significant tool for internet-enabled sensor-based home exercise platforms.

Index Terms—e-health, CVD module, motion capture, exercise evaluation, user modelling, decision support system

I. INTRODUCTION

Patient adherence to prescribed exercises is crucial to disease prevention and rehabilitation, since physical inactivity is the fourth leading risk factor for global mortality [1]. More than one billion adults world-wide suffer from Cardiovascular Diseases (CVD), Coronary Heart Disease and other preventable non-communicable cardiac diseases. Effective Cardiac Rehabilitation (CR) can significantly improve mortality and morbidity rates, leading to longer independent living and a reduced use of health care resources. However, uptake of traditional community-based long-term (phase III) cardiac rehabilitation is low (approx. 11%) and is further diminished by low subsequent adherence rates due to severe lack of programs, travel time, scheduling issues, lack of peer mentoring, and low self-efficacy associated with poor exercise technique and perceived poor body image.

Taking advantage of the technological and e-Health achievements, patients can be motivated to be more active and improve the quality of their lives through user-friendly, home-based rehabilitation systems, modelling and capturing their needs and preferences related to their condition and physical activity status.

In this paper, a human motion analysis (HMA) and decision support system (DSS) module is proposed, addressing the needs of CVD patient-oriented, home-based rehabilitation systems by assessing and personalizing prescribed physical exercises. In particular, HMA serves motion analysis, capturing and evaluation, while DSS achieves user modelling and cardiac rehabilitation session adaptation according to personalized characteristics.

Human motion capture (MoCap) is performed, providing real-time motion capture data, allowing for exercise evaluation and, therefore, patient exercising monitoring. In particular, MoCap data, providing instantaneous local and global indications concerning the performed exercises, are analyzed, resulting in patient feedback, i.e., exercise repetition detection and evaluation.

On the other hand, DSS is constructed aiming to a) model and capture users needs and preferences related to their condition and physical activity status and b) adapt the exercise sessions according to the personalised characteristics of the individuals in the axes of types of physical activities, exercise intensity and duration, and personal schedules. The decision horizon of the DSS is both short-term, i.e., pertaining to corrective modifications within the exercise session, and long-term, i.e., suggesting adaptations in exercises on a day by day or week by week basis.

II. RELATED WORK

Motion capturing and evaluation of the performed actions, in combination with decision support systems, prevail in the vast majority of the recently devised rehabilitation frameworks, both for home-based and for hospital-based systems. Kinect-RehabPlay project, proposed by Vieira et al. [2], is a home-based cardiac rehabilitation system, consisting of a virtual reality environment, and a monitoring software package, while motion capturing is performed with the aid of a Kinect sensor. The patients are shown the prescribed exercises by a virtual coach, being also able to see their own avatar on the screen. Joint angles are used for the evaluation of the performed actions and real-time feedback is provided, both concerning user placement in space, and the exercises themselves. The system presents two progressive levels of intensity, with the second being established 3 months after the beginning of the program, and, contrary to the proposed module, the only real-time adaptations that can be made are related to the trainer speed and the number of repetitions to be performed.
Lu et al. propose CAROLS [3], a system for home-based cardiac rehabilitation. The system aims to assist health personnel in providing patients with personalized exercise programs and motivating them to adhere to these programs, facilitating their health recovery. Alike to the proposed module, the system combines data acquired with the aid of a motion sensing and a wearable sensor device so that vital signs (e.g., ECG, breathing rate), perceived exercise intensity and compliance, can be monitored and transmitted to a decision-support system in real-time. Thus, based on the analysis of these data, personalized instructions are returned to patients as feedback, and to clinicians for applying personalized changes to future programs based on patient rehabilitation progress. However, the exercising evaluation is posture-based and not exercise-based, giving less evidence of the way a patient performs the exercises, while the decision support system adaptation is exclusively focused on patient safety, not on further CR program personalization.

III. MOTION ANALYSIS AND USER MODELLING MODULE

For a successful design of an unsupervised and personalized cardiac rehabilitation program, through the use of a technological module, two main requirements must be considered. The first one must be based on the personalization of the exercises during the execution of the exercise session, while the second one is more centralized on the modification of the long term parameters of the CR program. The reason that this approach is defined is that the patient must be on one hand well-guided to perform the exercise session correctly, exercising within the beneficial zones [4]. In other words, during an exercise session, the exercises that the patient performs, must focus on the activation of several muscle groups while the cardiorespiratory output must be the maximum possible, but emphasizing on and prioritising the safety of the patient. It is worth mentioning that to do so, in an unsupervised environment, motion capture and exercise evaluation is required to feed the system with information concerning the patient physical activity. On the other hand, the patient must be compliant with the exercise program and, thus, the DSS should also focus on the personalization of the program taking into account the exercise type, as well as the patients needs and preferences.

There are general principles to achieve the best possible effects in exercise-based rehabilitation. According to clinical guidelines [5], the structure of physical rehabilitation programs needs to be parameterised based on the Frequency, Intensity, Type, and Time duration of exercise (FITT). Therefore, best effects can be achieved by the regulation of the above-mentioned parameters and through detection of patient performance based on his/her physiological response, e.g., heart rate (HR) values along with other important subjective parameters such as the rate of perceived exertion (RPE) [6].

A. Module Requirements

In order to achieve the development of such a module, based on the aforementioned principles, motion data and vital signs are required for HMA and DSS components. More specifically:

1) Motion capture data: Information regarding the human body joints positions and orientations in time is required for HMA. MoCap can be achieved with the aid of the Microsoft Kinect device that constitutes a marker-less, thus user-friendly, motion capture solution [7]. Furthermore, fusion of inertial data from Wireless Inertial Measurement Units [8] (WIMUs) with the Kinect produced skeleton, has been developed, offering a more accurate motion capture solution as an extra option for the module, allowing for appropriate selection based on the required intricacy.

2) Heart rate data: In order to realise the real-time DSS functionality of the DSS components, sensing devices for HR monitoring in real-time such as a wrist-worn band [9], e.g. Microsoft Band 2, must be utilised as well. Real-time HR values must be collected and analyzed after each set of exercise repetitions.

3) Blood pressure data: A blood pressure monitor device must be used to measure blood pressure for the DSS pre-screening component. Ideally, a blood pressure device that can return measurements, when required, can be used with the module.

B. Module Architecture - Data Flow

The proposed module has been built based on these principles, consisting of four main components:

- **Motion Capturing**, for capturing the performed exercises,
- **Exercise Evaluation**, for exercise repetition detection and evaluation,
- **Online DSS**, for real-time adaptation of the exercising sessions,
- **Offline DSS**, for long-term adaptation of the exercises.

![Fig. 1. High-level diagram of the module (cyan components) integrated in a rehabilitation platform.](image)

The proposed technology has been developed to be a significant e-Health module, utilized and integrated in CVD rehabilitation systems. For instance, although there are several
extra modules and features that a CVD rehabilitation platform can include, the integration of the proposed module with two other components, namely a rehabilitation application and a data storage module, constitute the core of a CVD rehabilitation platform, including the basic functionality to run exercise class (Exercell) or exergame [10] sessions. In Figure 1, a simplified diagram of the module and the aforementioned components is shown, where there exist 3 different types of interaction, direct, internal (direct-hidden) and indirect (through the core rehabilitation application or the data storage component). In particular, the interactions are enumerated and described below:

1) (direct) Motion data are provided to the core rehabilitation application.
2) (internal) Motion data are transferred to the exercise evaluation. Event-based architecture has been implemented in order to avoid latency due to the core application frame-rate.
3) (direct) The evaluation data — (repetitions, cumulative accuracy) — are given to the core application.
4) (direct) Data extracted using the module as well as HR measurements are stored in a data storage module, allowing for further usage from the Offline DSS.
5) (direct) Online DSS gets the HR values for real-time adaptation.
6) (indirect) Online DSS retrieves the exercise evaluation results through the core application. Thus, taking into account the exercise accuracy and the HR results (real-time adaptation), DSS selects the next exercise.
7) (indirect) Offline DSS gets and analyzes the aforementioned stored data of the performed sessions per patient, updating then the patient profiles for the next sessions (long-term adaptation).

C. Human motion analysis

Focusing on HMA (motion capturing and exercise evaluation components), the diagram in Figure 2 illustrates the internal interaction flow. The motion capturing component acquires and processes / filters the raw motion capture data, in order to feed the exercise evaluation component. Pre-recorded training motion data have been considered as reference and used to train the latter, so that it can detect the time interval of the performed repetitions. Thus, feeding the exercise evaluation component with captured motion data in real-time results in exercise repetition detection and, comparing the cropped motion data of the detected repetitions to reference data, in exercise evaluation. Motion data collection during exercising, allows for cumulative accuracy extraction at the end of the every specific exercise in order to be used for real-time adaptation from Online DSS (Section III-D).

1) Motion Capturing: The Motion Capturing component uses the skeleton tracking algorithm of the official MS Kinect SDK 2.0, thus the users 3D body information, i.e., the global position and orientation of 25 human body joints (as they are shown in Figure 3a), is given per frame. In particular, considering the Kinect sensor position and rotation as origin, the global 3D position and orientation (in quaternion) for each j-joint are given for each frame. However, given that the Kinect skeleton could be problematic due to self-occlusions or fast movements, further processing is required to improve the motion capture data quality. 3D Kalman filtering is applied to achieve smooth frame-to-frame transition and prevent erroneous estimates of Kinect skeleton tracking [11]. Erroneous estimates of joint positions (especially under circumstances like self-occlusion) can be corrected by imposing inter-frame correlation of joint positions via filtering. The state transition matrix is set based on the velocity definition, while the measurement vector corresponds to the estimated 3D joint position, and is modelled as the actual position plus zero-mean Gaussian white noise. Applying Kalman on joint 3D positions and spherical linear interpolation on the joint quaternions, the extracted body information per frame is given in order to be used for exercise evaluation and virtual coach avatar animation. Virtual coach is required in order to instruct a patient how to correctly perform the prescribed exercises of the CR program.

Further motion capture solutions have been investigated in order to address some of the Kinect tracking limitations. Fusing Kinect skeleton with inertial data from WIMUs (acceleration, angular velocity, magnetic field data), allows for reliable motion tracking. In particular, the WIMUs are utilized in order to enhance the accuracy of the skeleton tracking, mainly when severe self-occlusions are present, while they increase the tracking stability by removing possible jumps of the Kinect estimated joint positions during fast movements or temporary occlusions.

Thus, a new real-time approach for skeleton tracking is offered by fusing the Kinect skeleton tracking with the WIMU
orientations, extracted by applying Madgwick [12] orientation filter on inertial data. More specifically, the orientations provided by the WIMUs are spatially aligned / calibrated to the Kinect body structure and are used to animate the bones, on which the sensors are placed (Figures 3b, 3c). The novelty of this approach is that it prevents the WIMU orientation from drifting by applying an on-the-fly update/calibration of the WIMU orientations based on the Kinect joint orientations, once the Kinect skeleton is considered stable (i.e., the tracking confidence of the limb end-joints, hands, thumbs, hand tips, feet - is high).

Taking advantage of the WIMUs placed on the human body, an inertial sensor \( s \) measures the orientation of its local coordinate system \( F_s^T \) with respect to a fixed global frame of reference \( F^T \). Consider the inertial sensor \( s \) associated with a rigid-body part \( p \), rotated by \( B_p^T \), equal to the Kinect joint orientations, when assuming that the tracking is stable. The aim is to rotate \( p \) according to the subsequent rotational information provided by the sensor \( s \). Let the quaternion \( q_{B_s}^T \) be the rotational offset occurring from \( F_s^T \) to \( B_p^T \):

\[
q_{B_s}^T = F_s^T * \otimes B_p^T
\]  

(1)

\( \otimes \) denotes the quaternion product and * denotes the quaternion conjugate. The rotational offset \( q_{B_s}^T \) between \( s \) and \( p \) is given. Then, by multiplying the quaternion \( q_{B_s}^T \) with the quaternion \( F_s^T \), given from sensor \( s \) at a time instance \( t \), the orientation of the rigid body part \( p \), \( B_p^T \), is calculated. Finally, it is worth noting that skeleton structure (bone lengths and joint positioning) as well as the root joint position (SpineBase) are estimated directly from the Kinect tracking implementation since the WIMUs are not directly capable of translation estimation. The method has been designed to function with an arbitrary number of WIMUs to minimize the associated equipment cost and complexity, depending on the motion intricacy to be captured. For instance, to increase the accuracy using only 3 WIMUs, the sensors can be placed on the thorax and the forearms, as depicted in Figure 3b. In Figure 3c, the Kinect tracking is confused due to the body part stacking (hands to head), thus the fused skeleton gives better results.

2) Exercise Evaluation: The exercise evaluation component is based on the detection of a priori known action / exercise instances within a signal of motion data, so that they can subsequently be analyzed and evaluated, providing evidence of whether the right exercise has been performed and how good the performance was.

Initially, efficient algorithms - based on Adaptive Boosting (AdaBoost) and Random Forest Regression (RFR) for real-time reliable discrete and continuous gesture recognition, respectively - using motion data were studied and used. In particular, boosting algorithms in general combine multiple weak classifiers (like base learners), each one performing classification based on a simple rule, in order to form a strong classifier producing results taking all the above rules into account. They are usually the best solution when a wide variety of data need to be predicted with high accuracy. More specifically, Adaptive Boosting initially fits a weak learner, namely Decision stump, on the entire dataset, equally weighting all samples. The misclassified samples are identified, higher weights are assigned to them and a new learner is added aiming for the correct classification of these specific samples. In this way, new learners are iteratively added, until the desired accuracy or number of learners is reached. Similarly, random forests also combine multiple weak models (i.e., trees) to form a powerful one, the output result of which, in case of regression, is the average of the outputs of the different trees. Random forest algorithm can be used both for classification and regression, and it can handle large, high dimensional datasets effectively. Training set samples and features are randomly sampled with replacement, building different decision trees. Multiple such samplings are performed and finally the decisions of all these trees are combined in one.

However, due to the complexity of CVD rehabilitation exercises [13], gesture recognition could not meet the requirements with respect to physical exercising evaluation, required for exercise repetition detection. The term exercise repetition detection is used for a) recognizing that a repetition of a specific exercise was performed and b) detecting the time interval during which it was actually performed (the time interval is necessary in order to crop and evaluate only this specific part of the motion data sequence). For this purpose, the aforementioned algorithms have been extended to serve repetition detection in complex CVD rehabilitation program exercises, e.g. trunk rotation with knee lift every other turn. In particular, weighting the frames of the motion data, key-frames (special poses) and their order are manually defined, expecting from the patients to perform these poses with same order during exercising. To do so, the component algorithm consists of three main steps:

- Initially, a reference weighting signal \( W(f) \) is given for each exercise based on the manual weighting, where \( W \) is the weight and \( f \) the frame of the reference sequence.
- Based on \( W(f) \) and the aforementioned algorithms, the motion data frames are evaluated (confidence estimation per frame) in order to recognize whether the special poses have been performed or not, resulting also in a confidence signal \( C(f) \), where \( C \) is the confidence and \( f \) the frame of the sequence.
- Finally, if the confidence signal satisfies the pre-defined conditions (signals correlation, similarity and duration) in comparison with \( W(f) \), the repetition is detected and the evaluation of the motion data can follow.

Furthermore, for some continuous exercises (e.g. running on the spot), two repetitions, each performed by different limbs, are usually considered as one. For instance, a walking on the spot repetition is considered completed when the subject performs two steps. To this end, slight modifications were made to the detection algorithms, so that a full repetition can be triggered after the detection of two partial repetitions performed by opposite limbs.

As already mentioned, repetition detection is followed by
exercise evaluation, which depends on the accurate temporal signal synchronization, performed with the aid of multidimensional Dynamic Time Warping (MD-DTW) a widely used technique for non-linear signal alignment and comparison under certain constraints [14]. The temporal alignment of the capture with the corresponding reference signal is followed by their spatial alignment, separately performed for the upper and lower body through Singular Value Decomposition (SVD). Sequentially, the exercise evaluation algorithm relies on the comparison of the 3D-skeleton motion data of the patient with the reference pre-captured exercises using human body motion features as the joints 3D position \( p_j \), orientation \( \mathbf{q}_j \), normalized linear velocity \( \mathbf{v}_j \), angular velocity \( \mathbf{\dot{q}}_j \) and anthropometric angle \( \theta_j \).

3) Technical details: Regarding the technical implementation, the HMA components have been developed as .NET 3.5 C# Dynamic Link Libraries (Unity3D 5 x64 compatible), providing programmatic access to:

- Select and configure the motion capturing method (method configurations, sensor parameters, etc.).
- Initialization of the motion capturing component.
- Event-driven architecture for body frame acquisition (plus the Kinect RGB-D frames).
- Initialization of the exercise evaluation component.
- Provide the instantaneous feedback per repetition by feeding the exercise ID (in order to initially load the exercise reference motion data), the user’s body frame (from motion capturing) and the reference motion frame index.
- Provide the cumulative accuracy per exercise (after finishing a set of repetitions).

D. Decision Support System

The DSS component must be designed systematically following clinical guidelines and recommendations of experts in exercise-based cardiac rehabilitation and health psychology [15]. Extensibility and scalability of the DSS system are ensured through the deployment of a service-oriented architecture.

As mentioned, the DSS employs diversified operations which can be grouped according to their time horizon (either short-term or long-term). These groups are reflected in the two main components of the DSS, the Online DSS (short-term horizon) and the Offline DSS (long-term horizon).

1) Online DSS: The ultimate goal of the Online Decision Support is to evaluate the patients clinical status before the start of an exercise session and enable the personalization of the session according to patients performance and needs. In other words the aim of the component is two-fold:

1) To evaluate the patients clinical status prior to the execution of the current session of the CR program, in order to allow the user to proceed to start exercising with safety.
2) To evaluate the performed exercises in terms of execution accuracy and the actual cardiac response to the exercise. In this respect, a personalized modification on

the exercise session is achieved, enabling the patient to take full advantage of the beneficial effects of the current exercise session of the CR program.

This component is rule based. The structure is based on the deterministic rules in the format of conditionaction (IF-ELSE). In this component, the data sources include:
- the HR as captured by the wearable sensors;
- the cumulative accuracy, as computed by the Exercise Evaluation component. This parameter is based on the evaluation of the detected repetitions for each exercise;
- answers to questionnaires (before and after the execution of the exercise session);
- information related to the previous exercise sessions, such as the perceived exertion, and pleasure, the actual performance of the previous sessions, the patients vital signs, as well as the evaluation results during the previous pre-screening phases.

The Online DSS component is a sub-system enhanced with suitable communication interfaces to communicate with a core rehabilitation application (client). It consists of well-defined web service-based APIs to achieve efficient data exchange with respect to the required DSS input information and the generated outputs. The APIs implemented were designed in such a way so that both performance, in terms of response time, and the quality of the DSS outcome are both guaranteed.

2) Offline DSS: The purpose of this component is to evaluate patient specific information in order to modify the prescribed CR program as well as to address specific goals of the patient. The Offline DSS component requires the combination of data from different sources which include diverse type of information. This information ranges from the performance of specific exercises, vital signs, answers to questionnaires as provided by the user and system use. This information is then used to assess patients needs and preferences in order to reach the main goal of the rehabilitation system, which is the beneficial execution of the CR program. In particular, the Offline DSS has two main functionalities:

- Evaluation of the already performed exercise sessions of the CR program in association with the patients clinical status in order to modify the next sessions of the CR program.
- Evaluation of the user input in order to personalize the functionality of the system.

By nature, the Offline DSS is a complex component able to be involved in numerous UCs, including UCs related to a) the activity and the personalization of the CR program and b) the motivation of the patient for a more beneficial use of the rehabilitation system. As in the Online DSS case, the Offline DSS interfaces are implemented as a RESTful web service with specific endpoints. This component, as in the Online DSS case, is based on deterministic rules of condition-action. Taking into account the complexity of the Offline DSS functionality, this component is further divided into three separate components.

1) Performance capturing component.
2) Program customization component.
3) Behavioural change and notification component.

IV. DISCUSSION AND NEXT STEPS

In this paper the design of a modular approach for unsupervised and personalized cardiac rehabilitation program was described. The design and development of a system targeted at leveraging modern information technology in the construction of the next-generation of cardiac rehabilitation programs is presented. These programs provide personalised and reliable information to the CVD patients, facilitating their health management, while reducing the cost for healthcare resources. In this context, the proposed module will be able to provide personalized exercise sessions, tailored to patient needs and performance characteristics, helping patients to be more active and motivated to improve the quality of their lives.

The modular approach makes it appropriate and eligible for use in any application that deal with the personalization of the exercise session. In this respect, this architecture has been adopted in the context of the EU project, PATHway [16], main goal of which is to provide individualized rehabilitation programs to CVD patients, following Phase III of such programs. The initial tests seem to be promising, making the proposed module as appropriate for such systems. The PATHway system has been already implemented and integrated, while a Randomized Control Trial (RCT) [17] has been already designed. The primary aim of the RCT is to check whether tele-rehabilitation services could lead to an increase of the physical activity uptake.

ACKNOWLEDGEMENTS

This work was supported by the EU funded project PATHway under contract 643491.

REFERENCES