Assessment of rehabilitative exercises by humanoid robot

Mikhail Simonov¹ and Gianpiero Delconte²

Abstract. This article describes an approach in which the rehabilitative exercise prepared by healthcare professional (human) is encoded as formal knowledge and used by humanoid robot to assist patients in residential settings without involving other care actors. The authors are researching on the new cognitive capability enabling robots to judge about the correctness of the rehabilitative exercise performed by patients following the robot’s indications. The proposed method uses the Dynamic Time Warping functionality comparing the correct sequence (encoded in the Knowledge Base) with the human actions being observed by the robot’s eyes. The proposed approach is an enabler of better sustainable rehabilitative care services in remote residential settings because of lowering the need of human care.

1 INTRODUCTION

The demand of remote care services for elderly population is steadily growing because of the demographic developments reducing the availability of caregivers. For this reason the care stakeholders are actively looking at the new enablers of remote Assistive Living services (Fig. 1).

To offer valuable and sustainable care services using the telemedicine approach in general, the key aspects are (a) the monitoring of Activities of Daily Living (ADL), (b) the activity recognition and classification, and (c) rule-based Decision Support System capable to trigger the relevant situations/events.

In the known art, the rehabilitative exercises are prepared by physiotherapists and offered to patients (individually or in small groups) in specialized care centers. Whenever it is impossible to the patient to reach any care center because of the physical impairments, the healthcare professionals offer the necessary care in their residential settings. Because of the steadily reducing support ratio due to demographic developments, the sustainability of rehabilitative care at home could be achieved by using non human healthcare actors, for example by anthropomorphic robots.

The use of humanoid robot as tutor to offer the rehabilitation at home relies on a number of cognitive capabilities, such as the capability (a) to reach the position nearby patient at right moment/time, (b) to persuade the patient to start doing exercises, (c) to show the correct motions to follow, (d) to observe how the patient is doing the exercise, and (e) to make a judgment about the compliance of the exercise being made by patient to the one prepared by physiotherapist(s). Upon them, the robot can decide the most appropriate automated reaction to stimulate patient(s) continuing the exercises or to report to human healthcare actors the non-compliant or abnormal situations.

The advent of robotics and the recent developments of advanced cognitive capabilities make possible using anthropomorphic robots in remote residential settings for care purposes. In the current state-of-the-art, the robot can navigate till the position in front of the patient, can gain the patient’s attention, can speak inviting to follow the physical exercise, and can show the correct sequence of movements [1]. The mobile robots can observe the activities performed by humans (Fig.2) using the eyes [2] and produce the sequences of numeric data describing the skeleton/node’s movements. For example, the sequence $N(t) = \{x(t_k), y(t_k)\}$ describing the motion of the leftmost vector (node) could be used to assess the motion of the right hand while doing the rehabilitative exercise.

Unfortunately the capability to judge about the quality of the exercises performed by patients - following the robot’s indications - is not available yet. The authors investigates on the pattern matching techniques aiming to implement said capability to observe the happenings (actions performed by human), to compare them with the intended and formalized behavior, and to take a decision about the match and/or mismatch. The ultimate goal is to set up the pervasive remote care rehabilitative service in which the robot will filter automatically the situations non-needling specific human interventions in order to concentrate the available

¹ M. Simonov is with the Istituto Superiore Mario Boella (ISMB), Via P.C. Boggio 61, Torino, 10138, Italy, (corresponding author, e-mail: simonov@ismb.it; phone: +39-011-22761; fax: +39-011-2276299).
² Gianpiero del Conte is with the Politecnico di Torino, email: gianpiero.delconte@ymail.com
limited workforce of healthcare professionals to the only planned interventions and the situations classified by the robot as “abnormal”.

Many research studies in care domain speak about the use of more or less intrusive wearable and environmental sensors for automated monitoring of people living independently. A review of the literature available in [3] highlights the focus on monitoring techniques, the use and placement of wearable sensors, the methods of data collection and elaboration techniques. An example of the selective activity monitoring is reported in [4].

2 PROJECT BASELINE

One of the possible approaches to the use of robots in the rehabilitative care is shown in [5]. However, their use at home seems not yet widely discussed. KSERA project uses the humanoid robot to assist sufferers of Chronic Obstructive Pulmonary Disease (COPD) living independently at their homes. COPD is the occurrence of commonly co-existing diseases of the lungs in which the airways narrow over time causing shortness of breath (dyspnea). This limitation is poorly reversible and usually gets progressively worse over time. For these reasons the COPD care does include a pulmonary rehabilitation program which can be done nearly anywhere. The physical and breathing exercises build the muscle strength and endurance to reduce shortness of breath.

Accordingly the care guidelines, the difficulty levels of pulmonary rehabilitation programs can vary depending on the SpO2 levels. In KSERA, each pulmonary rehabilitation exercise is developed by healthcare professionals and encoded by engineers in software program executable by the robot. The correct sequences are stored in the Knowledge Base. The system keeps awareness about the duration of each element included in the sequence.

The Ubiquitous Monitoring sub-system (UMS) of KSERA detects the conditions triggering the rehabilitative exercises. The measurement of the blood’s SpO2 level using the wearable oxi-pulsiometer governs the choice of an appropriate exercise. The humanoid robot starts the rehabilitation program. It detects where the patient is, it moves nearby the patient, it stops in front of him/her. The robot invites the patient to follow the exercise and starts showing the pre-programmed sequence of actions. During the exercise, the robot remains in the position permitting to meter the human activities by direct observation.

3 PROPOSITION

In order to express a judgment about the quality of pulmonary rehabilitation exercise, the authors use the robot eyes to acquire the sequence of video frames at a standard rate. The video sequence is submitted to the video-analysis software in order to transform it in the motion vectors. This step is made using the commercially available libraries containing the Skeleton class of Kinect software. A similar approach is documented in [6]. The motion vector is processed in order to extract the time constrains corresponding to each activity element. To simplify the task, in this study we measure the moving arm/leg duration and the time intervals between the mobility periods (Fig.3).

Let us trace the positions of the same skeleton’s node Ni in time dimension. It gives the sequence of N(t). Let us denote the spatial distances between two time points as \( d(Ni, tk, ti) = |x(tk) - x(ti)| + |y(tk) - y(ti)|^2 \). for each node Ni, having the coordinates x(t), y(t), and z(t). Let us denote the time interval [tk, ti] as \( \Delta ti \). Until the spatial distance between the positions of the same node varying/neighborhood in time dimension remains small, the whole time segment is classified as “no action” or Pause. The time intervals between two adjacent Pauses are classified as DoingActivity.

At this point the annotated patient’s mobility becomes a sequence \( S1 = (\text{Action}(t1, t1t1), \text{Pause}(t1t1, t1t2), \text{Action}(t1t2, t1t3), ... \}

Another similar sequence corresponding to the storyboard of the exercise encoded for the humanoid robot is available and it could be retrieved from the Knowledge Base. Let us call it as \( S2 \). It could be used as the known pattern for the comparisons between \( S1 \) and \( S2 \). The original problem about judging about the quality of exercises is simplified and replaced by the question AreSimilar(S1, S2)? Two sequences have the same structure, so the comparison in terms of similarity and timing is possible. Based on the results of comparison, one can trigger the cases requiring human care interventions. The authors implement the filtering and the corresponding control action as IF AreSimilar(S1, S2) = TRUE THEN Tell(“Ok”) ELSE AdvertDoctor().

As pointed by [6], some delay between all the \( t \) terms in \( S2 \) and the corresponding – if any – \( t \) terms in \( S1 \) exist because the robot starts first. The normal delay between two actions is \( d1 = |tk - ti| \). The patient can decide to follow the robot or to do something else. By following the pulmonary rehabilitative program shown by robot, each human action could have different duration compared with the robot’s one. For each term in \( S1 \), being compared with the corresponding term in \( S2 \), we calculate the time warping coefficients \( k_i = (t_{i+1} - t_i)/((t_{i+1} - t_i)) \). These coefficients compare the speed of doing the same movement by the therapist (represented by the humanoid robot) and by the patient. Our ambition is the new cognitive capability to let robots judging the compliance between the pairs \( S1, S2 \). To operate with variable durations we have encoded the concept of the correct timing (Fig. 4). Until the duration of the segments stays in the allowed range, e.g. 0.8<k_i<1.2, the timing is judged as correct.
Although the DTW is designed for full sequence matching (the exercise is a sequence composed by several segments/actions), it can be also used for sub-sequence matching (for example moving the left arm up). Our algorithm checks the full match first (Fig. 9). It stops when no matches were found. Elsewhere, it looks for more sub-sequence occurrences by using a sliding time window. The starting timestamp of the first cycle is always used to characterize the exercise being executed. To collate adjacent windows as a single process, the timestamps are compared: If the comparison \((t_c - t_l)\) gives more than 10 seconds, the authors decide to abort the sequence.

Fig. 4. Fuzzy definition of the correct length of the segments.

Fig. 5. The rehabilitation program being encoded in software.

Fig. 6. Artificial physiotherapist shows the exercise.

Fig. 7. Patient follows the activity shown by robot. The robot observes human activities.

Fig. 8. The durations of rehabilitation exercise made by robot/human.

Fig. 9. Using DTW to compare human-made and robot-made sequences.

Once the enumeration of classifications \(\text{Event}_1 = (\text{moving}, t_4); \text{Event}_2 = (\text{moving}, t_7); \text{Event}_3 = (\text{moving}, t_9)\), … becomes available, the initial \(E(t)\) sequence could be unloaded from the memory. The main classification loop produces the sequence of \([\{xxx, t_4\}, \{xxx, t_7\}, \{xxx, t_9\}, \ldots\]\) going to be stored in the knowledge base. The role of DTW in this process is fundamental because the duration of patient’s mobility segments is never known a-priori.

Now let us forget any numeric data: the patterns observed by the robot are the patterns of human behavior. It is important to note that robot is not expected to comment on the correctness of each segment to avoid the noisy dialogs. Despite the precision of this technology is lower than 100%, the above functionality could operate with real time data in a satisfactory manner because it is sufficient to react few times during/after the exercise. This example illustrates the implementation of DTW when two sequences are strings in discrete time-space. The distance \(d(x, y) = |x - y|\) was used to determine the timestamp and the duration of the doing exercise processes.

The experiment was performed in batch mode with completed sequences only. Once the full \(S(t)\) is acquired by humanoid, it is supplied as input to the new algorithm.

The authors have collected some sequences from few patients following the robot in doing exercises. To verify the applicability of the approach to the wider targets, the simulated datasets were produced using these real data as the correct (Identity) and incorrect (Null) situations. The real time-varying sequences (doing correctly/ incorrectly/nothing) were elaborated applying stretching of amplitude/time, and linear translations to obtain additional
simulated datasets. The larger data collection (Table 1) is used to compare two sequences using the DTW algorithm.

### Table 1. Rehabilitation exercise using a robot.

<table>
<thead>
<tr>
<th>Set</th>
<th>Patient activity</th>
<th>Correct duration</th>
<th>Duration by DTW</th>
<th>Similarity coeff.</th>
<th>Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.0*tAmplitude</td>
<td>10</td>
<td>10</td>
<td>1.00</td>
<td>Y</td>
</tr>
<tr>
<td>2</td>
<td>3.0*tAmplitude</td>
<td>10</td>
<td>10</td>
<td>1.00</td>
<td>Y</td>
</tr>
<tr>
<td>3</td>
<td>10.0*tAmplitude</td>
<td>10</td>
<td>10</td>
<td>1.00</td>
<td>Y</td>
</tr>
<tr>
<td>4</td>
<td>0.5*tAmplitude</td>
<td>10</td>
<td>10</td>
<td>1.00</td>
<td>Y</td>
</tr>
<tr>
<td>5</td>
<td>0.3*tAmplitude</td>
<td>10</td>
<td>10</td>
<td>1.00</td>
<td>Y</td>
</tr>
<tr>
<td>6</td>
<td>0.1*tAmplitude</td>
<td>10</td>
<td>10</td>
<td>1.00</td>
<td>Y</td>
</tr>
<tr>
<td>7</td>
<td>2*time</td>
<td>10</td>
<td>20</td>
<td>0.50</td>
<td>N</td>
</tr>
<tr>
<td>8</td>
<td>3*time</td>
<td>10</td>
<td>30</td>
<td>0.33</td>
<td>N</td>
</tr>
<tr>
<td>9</td>
<td>1.2*time</td>
<td>10</td>
<td>12</td>
<td><strong>0.83</strong></td>
<td>Y</td>
</tr>
<tr>
<td>10</td>
<td>1.1*time</td>
<td>10</td>
<td>11</td>
<td>0.91</td>
<td>Y</td>
</tr>
<tr>
<td>11</td>
<td>0.9*time</td>
<td>10</td>
<td>9</td>
<td>1.11</td>
<td>Y</td>
</tr>
<tr>
<td>12</td>
<td>0.8*time</td>
<td>10</td>
<td>8</td>
<td><strong>1.25</strong></td>
<td>Y</td>
</tr>
<tr>
<td>13</td>
<td>0.7*time</td>
<td>10</td>
<td>7</td>
<td>1.43</td>
<td>N</td>
</tr>
<tr>
<td>14</td>
<td>0.6*time</td>
<td>10</td>
<td>6</td>
<td>1.67</td>
<td>N</td>
</tr>
<tr>
<td>15</td>
<td>0.5*time</td>
<td>10</td>
<td>5</td>
<td>2.00</td>
<td>N</td>
</tr>
<tr>
<td>16</td>
<td>0.4*time+6</td>
<td>10</td>
<td>4</td>
<td>2.50</td>
<td>N</td>
</tr>
<tr>
<td>17</td>
<td>0.2*time+1</td>
<td>10</td>
<td>2</td>
<td>5.00</td>
<td>N</td>
</tr>
<tr>
<td>18</td>
<td>0.6t+5 and 6A</td>
<td>10</td>
<td>10</td>
<td>1.00</td>
<td>N</td>
</tr>
<tr>
<td>19</td>
<td>Null</td>
<td>10</td>
<td>20</td>
<td>0.50</td>
<td>N</td>
</tr>
<tr>
<td>20</td>
<td>Identity</td>
<td>10</td>
<td>10</td>
<td>1.00</td>
<td>Y</td>
</tr>
<tr>
<td>21</td>
<td>0.3+7 and 0.5A</td>
<td>10</td>
<td>5</td>
<td>2.00</td>
<td>N</td>
</tr>
<tr>
<td>22</td>
<td>1.7+3 and 0.1A</td>
<td>10</td>
<td>17</td>
<td>0.59</td>
<td>N</td>
</tr>
<tr>
<td>23</td>
<td>0.7+3 and 0.1A</td>
<td>10</td>
<td>7</td>
<td>1.43</td>
<td>N</td>
</tr>
<tr>
<td>24</td>
<td>1.5+6 and 0.1A</td>
<td>10</td>
<td>15</td>
<td>0.67</td>
<td>N</td>
</tr>
<tr>
<td>25</td>
<td>0.9+7 and 0.1A</td>
<td>10</td>
<td>9</td>
<td>1.11</td>
<td>Y</td>
</tr>
<tr>
<td>26</td>
<td>0.7+3 and 0.5A</td>
<td>10</td>
<td>7</td>
<td>1.43</td>
<td>N</td>
</tr>
<tr>
<td>27</td>
<td>0.5+4 and 0.1A</td>
<td>10</td>
<td>5</td>
<td>2.00</td>
<td>N</td>
</tr>
<tr>
<td>28</td>
<td>0.8+8 and 0.1A</td>
<td>10</td>
<td>8</td>
<td>1.25</td>
<td>Y</td>
</tr>
</tbody>
</table>

Based on the collected and simulated data, one can observe that the DTW detects correctly both positive and negative deviations from the intended timing of the exercise. Assuming +/-20% as a limit of the allowed imprecision while repeating the activities, the DTW triggers said condition using a similarity coefficient 1.04 ca. Therefore the condition ($S_{Robot} = S_{Patient}$) could be evaluated through $\text{DTW}(S_{Robot}, S_{Patient}) \leq \{0.83, 1.25\}$.

### 5 CONCLUSIONS AND FUTURE WORK

This paper has described an engineering proof of concept to set up remote rehabilitative care of patients living independently at their homes. In the current demographic scenario, it exemplifies a possible approach to reduce the percentage of the human rehabilitative care.

The experimental results support the feasibility of the approach/method, but the numbers have limited validity because the sample is not statistically representative. Instead, the simulation covered almost all possible situations in which patients repeat the rehabilitation program fully correctly, partially correctly, not correctly, or not at all.

The monitoring of the activities performed by patients following the robot could be set up without additional and/or specialized sensors [9, 10] because the activities are not critical. Therefore, the robot eyes could be sufficient.

Compared with other signal processing techniques (the correlation is an example), the use of Dynamic Time Warping is beneficial because let to detect the instances of doing exercises having variable durations.

We observed that in typical conditions in which the patient followed the robot in doing the exercise, it keeps the same or similar behavior during more days. Repeating the same exercise during different days, taking annotated datasets stored in the data warehouse with the coefficients characterizing the normality or deviations, one can analyze the time series and judge about deviations, anomaly, or disease progression.

It appears possible to conclude that the abnormal durations detected by the robot could be useful as an indicator of the normal/abnormal ADLs.

The proposed approach is cheap in terms of computational power. By contrary, the hardware installation is very expensive because of the humanoid robot.

The future work is to refine the method and to experiment it at wider scale.

### 6 Acknowledgment

The research leading to these results has received funding from the European Community’s 7th Framework Program FP7/2007-2013 under grant agreement 248085- KSERA.

### REFERENCES