A Virtual Glove System for the Hand Rehabilitation based on Two Orthogonal LEAP Motion Controllers

Giuseppe Placidi¹, Luigi Cinque², Andrea Petracca¹, Matteo Polsinelli¹ and Matteo Spezialetti¹

¹A²VILab, Department of Life, Health & Environmental Sciences, University of L’Aquila, Via Vetoio, L’Aquila, Italy
²Department of Computer Science, Sapienza University, Via Salaria, Rome, Italy
giuseppe.placidi@univaq.it, cinque@di.uniroma1.it, andrea.petracca@graduate.univaq.it,
matteo.polsinelli@student.univaq.it, matteo.spezialetti@graduate.univaq.it

Keywords: Hand Rehabilitation, Virtual Glove, LEAP Motion Controller.

Abstract: Hand rehabilitation therapy is fundamental in the recovery process for patients suffering from post-stroke or post-surgery impairments. Traditional approaches require the presence of therapist during the sessions, involving high costs and subjective measurements of the patients’ abilities and progresses. Recently, several alternative approaches have been proposed. Mechanical devices are often expensive, cumbersome and patient specific, while virtual devices are not subject to these limitations, but, especially if based on a single sensor, could suffer from occlusions. In this paper a novel multi-sensor approach, based on the simultaneous use of two LEAP motion controllers, is proposed. The hardware and software design is illustrated and the measurements error induced by the mutual infrared interference is discussed. Finally, a calibration procedure, a tracking model prototype based on the sensors turnover and preliminary experimental results are presented.

1 INTRODUCTION

The hand is a fundamental organ of the human body and fulfill several complex tasks corresponding to a deep brain involvement. For patients suffering from post-stroke or post-surgery residual impairments the recovery of the hand functions is extremely important for accelerating the rehabilitation process. The possibility of the recovery is strictly related to the frequency, the duration and the quality of the rehabilitation sessions (Kopp et al., 1999; Liepert et al., 2000; Hallett, 2001; Arya et al., 2011). In the traditional rehabilitation approach, a therapist is required to follow the patient during one-to-one expensive sessions and subjective evaluations (the therapist, basing on his experience, evaluates the results). Over the last years, several automated (tele)rehabilitation tools, based on mechanic or virtual devices have been presented, in order to allow patients to execute the therapy in a domestic environment being followed by therapists through Internet (Burgar et al., 2000; Kahn et al., 2006; Placidi, 2007; Franchi et al., 2009; Franchi et al., 2010; Zimmerli et al., 2013; Placidi et al., 2013; Llorëns et al., 2015).

Mechanical approaches usually involve the employment of gloves equipped with pressure sensors and pneumatic actuators for assisting and monitoring the hand movements and to apply forces to which the fingers have to oppose (Lum et al., 2012; Maciejasz et al., 2014). Despite the measurement accuracy they are capable to achieve, these systems are expensive, cumbersome (they limit the patients spontaneity and freedom of movements) and patient specific (limiting the possibility of reusing). For replacing mechanical components, existing virtual approaches are based on...
evaluation of videos from RGB or depth sensing cameras to calculate hand kinematic information in real time, (Rusak et al., 2011; Avola et al., 2013; Chaudhary et al., 2013; Placidi et al., 2014; Charles et al., 2014; Placidi et al., 2015).

The LEAP motion controller (http://www.leapmotion.com, 2016) is a recently presented low-cost and non-bulky hand tracking device, characterized by high-resolution and high-reactivity (Bachmann et al., 2015) and represents a good system to be used for virtual reality applications and rehabilitation (Petracca et al., 2015; Polsinelli, 2015). It is a stereo vision device composed by 3 LED sources and 2 infrared cameras, that suffers the limitations of a monocular system being unable to recognize every position of the hand (occlusions can frequently occur). In order to address the occlusions issue, in this work a multi-sensor approach based on the use of two orthogonal LEAPs is proposed. The rest of the paper is organized as follows: Section 2 describes the proposed strategy, by showing the hardware structure (2.1), the software architecture (2.2) and the evaluation of the infrared interferences when using two sensors (2.3); Section 3 shows the tracking model prototype and some hand tracking preliminary results (3.3).

2 MATERIALS AND METHODS

The system has been designed to obtain simultaneous information from two LEAPs, orthogonally placed each other. The reason is that a single LEAP sensor is not able to compute with accuracy the hand position if the palm (and the fingers) is not visible, that is if the vector orthogonal to the palm is approximately parallel to the sensor plane (see Figure 1 third row). Using two orthogonal sensors should ensure that at least one of them is able to get the correct position. One of the major issues to be addressed was the devices connection: the driver and the API are not able to manage multiple instances of the LEAP on the same machine. For this reason an architecture including a virtual machine has been designed. Moreover, in order to build a consistent hand model using information from both sensors, their reciprocal position had to be computed with high accuracy through a calibration procedure.

2.1 Hardware Setup

A LEAP Motion Controller is a small device, designed to be placed above the surface of a desk. It can be connected to the PC through a USB port. The hardware includes two monochromatic infrared cameras and three LEDs emitting static IR beams (wavelength 850 nm). Considered the 3-Dimensional reference system used by the LEAP to represent the objects, shown in Figure 2.a, the fields of view of the sensor is 150° along the X axis and 120° along the Z axis. The intensity of the emitted light, limited by the power supplied by the USB port, allows to observe an object within a distance of about 0.6 m from the sensor. Objects in the field of view of the LEAP are enlightened by the LEDs and the reflected light is captured by the cameras, producing a couple of grey-scale images that are used, at software level, to identify the objects positions, thanks to a stereo vision algorithm (Weichert et al., 2013). The sensor is able
to track up to two hands simultaneously, by using a 3D model of the hand that includes the positions of the fingers joints, but can also recognize objects like sticks and pencils. In this work, two LEAPs have been fixed to a support, in such a way as to be orthonormal and distant 0.25 m from the center of the scene, in order to create a wide area in which the hand can freely move and can be tracked by both sensors (Figure 2.b and Figure 3).

2.2 Software Architecture

LEAP is equipped with an SDK that allows to the communication with the sensor both through a native and a web-socket based interface. By using the latter it is possible to obtain a JSON data structure containing tracking data (for more details, please refer to (http://developer.leapmotion.com, 2016)). Unfortunately, the SDK does not allow the creation of more than one instance of the device on a single machine. In order to address this issue, a virtual machine had to be included in the software setup (Figure 4 shows the diagram of the architecture). The virtual machine (Slave) has been installed on the physical machine (Master), for assigning each of the connected sensors to a different SDK; a javascript router was used to carry data from both LEAPs to a single server.

![Diagram of software architecture](image)

**Figure 4:** The software architecture: a virtual machine (Slave) has been installed on the physical machine (Master), for assigning each of the connected sensors to a different SDK; a javascript router was used to carry data from both LEAPs to a single server.

In order to address this issue, a virtual machine had to be included in the software setup (Figure 4 shows the diagram of the architecture). The virtual machine (Slave) has been installed on the physical machine (Master): in this way, plugging both sensors, one of them has been assigned to Master and the other to Slave allowing to the machines to instantiate their own driver. On each machine, data provided by the SDK through the websocket (the port number was fixed) were captured by a javascript router (executed on an instance of Node.js (http://nodejs.org, 2016)) and rerouted towards a Node.js server (hosted on the Master machine). In this way, using again websockets, the server was able to send data of both devices to one or more clients running on Master.

2.3 Infrared Interferences Evaluation

The described setup has been used to evaluate the infrared interferences that the sensors could cause each other. The main idea was to record the position of the same set of points (by tracking the tip of a stick, see Figure 3), using both sensors, in different conditions: just the LEAP used for the measure switched on; both LEAPs switched on. The first case represented what we defined the “ideal condition” used for estimating the “positioning error” (it contained just the error intrinsic to the sensor). However, being the considered ideal condition also a measure, what we calculated was a “distance” between measurements and not properly an “error”. In order to place the stick in different positions, a circular wood disk containing a series of holes on concentric circumferences was used (Figure 3). The tool was designed in order to be kept parallel to the X-Y plane of the LEAPs and to place a wood stick at different positions along the Z axis by making it move through the holes. The stick (0.36 m long) was marked along its length in order to know the distance between the tip and the sur-
face of the disk. A set of 84 different positions of the
stick was measured from both sensors and repeated
for both conditions. A total of 3 and 7 points were
not considered during the test of the horizontal and
the vertical LEAP respectively, due to the impossibil-
ity of getting the positions (probably due to the closer
position of the points with respect to the sensor). It
has to be noted that the issue did not depend on the
infrared interferences, because the phenomenon hap-
pened in both conditions. Figure 5 shows the spatial
distribution of the points, measured with the LEAPs
without reciprocal interferences. The absolute values
of the variations in the measurements are represented
in Figure 6, for both LEAPs and separately for each
axes. Figure 7 also shows the distance between the
same points measured in different conditions.

Table 1 reports average values, standard deviations
and maximum values of the distance along all
the axes and in the space, between points measured in different conditions.

<table>
<thead>
<tr>
<th></th>
<th>HOR. LEAP</th>
<th>VER. LEAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>X distance</td>
<td>AVG 0.11</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>STD 0.17</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>MAX 0.80</td>
<td>0.90</td>
</tr>
<tr>
<td>Y distance</td>
<td>AVG 0.28</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>STD 0.36</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>MAX 2.30</td>
<td>1.50</td>
</tr>
<tr>
<td>Z distance</td>
<td>AVG 0.79</td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td>STD 1.75</td>
<td>1.48</td>
</tr>
<tr>
<td></td>
<td>MAX 14.20</td>
<td>10.70</td>
</tr>
<tr>
<td>3D distance</td>
<td>AVG 0.96</td>
<td>1.03</td>
</tr>
<tr>
<td></td>
<td>STD 1.73</td>
<td>1.44</td>
</tr>
<tr>
<td></td>
<td>MAX 14.20</td>
<td>10.70</td>
</tr>
</tbody>
</table>

3 TRACKING MODEL PROTOTYPE

3.1 Calibration Procedure

The proposed strategy was based on the proper rigid
transformation (roto-translation) of the vertical LEAP
information. In order to minimize the errors in the
tracking procedure, the position and the orientation
of one LEAP with respect to the other had to be eval-
The calibration procedure aimed to compute the roto-translation transformation that links the vertical reference system to the horizontal, starting from the coordinates of the same set of points in the space observed from the viewpoints (only one sensor was turned on during each measure). The problem of estimating 3D body transformations that aligns two sets of points with known correspondences is well known in the computer vision field (Sabata and Aggarwal, 1991; Egger et al., 1997). The corresponding set registration strategy used to compute the transformation (Besl and McKay, 1992) started from two sets of points $A = \{\vec{a}_i: 1 \leq i \leq m\}$ and $B = \{\vec{b}_i: 1 \leq i \leq m\}$ and found a rotation matrix $R$ and a translation vector $\vec{t}$ such that it was possible to move from a reference system to the other:

$$\vec{b} = R\vec{a} + \vec{t},$$

minimizing the error:

$$err = \frac{1}{m} \sum_{i=1}^{m} ||R\vec{a}_i + \vec{t} - \vec{b}_i||^2$$

For each set, the Center Of Mass (COM) was computed:

$$C_A = \frac{1}{m} \sum_{i=1}^{m} \vec{a}_i \quad \text{and} \quad C_B = \frac{1}{m} \sum_{i=1}^{m} \vec{b}_i$$

and used to center the clouds on the origin:

$$A^+ = \{\vec{a}_i^+: \vec{a}_i^+ = \vec{a}_i - C_A\}$$

$$B^+ = \{\vec{b}_i^+: \vec{b}_i^+ = \vec{b}_i - C_B\}.$$ 

Then, the $3 \times 3$ cross-covariance matrix between the sets $A^+$ and $B^+$ was computed:

$$H = \frac{1}{m} \sum_{i=1}^{m} \vec{a}_i^+ \vec{b}_i^+ \vec{b}_i^+ \vec{a}_i^+ = T$$

and the Singular Value Decomposition (SVD) was applied, in order to decompose $H$ in the vector of transformations $[U, S, V] = \text{SV D}(H)$, such that

$$H = USV^T.$$ 

where $U$ and $V$ were orthogonal matrices and $S$ was a non-negative diagonal matrix. Since the desired transformation $R$ was a rigid rotation, it could be computed as:

$$R = UV^T.$$ 

The sign of the determinant $\Delta$ of $V$ had not to be negative: in this case the third column of $V$ was multiplied for $-1$ and $R$ was computed again. The translation was derived as follows:

$$\vec{t} = -R \vec{C}_A + \vec{C}_B.$$ 

Lastly, the roto-translation was summarized with a unique matrix (represented in homogeneous coordinates):

$$W = \begin{pmatrix} R_{1,1} & R_{1,2} & R_{1,3} & t_1 \\ R_{2,1} & R_{2,2} & R_{2,3} & t_2 \\ R_{3,1} & R_{3,2} & R_{3,3} & t_3 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$ 

The calibration phase was implemented in MATLAB to run off-line. Given the set of points $P$, and following a cross validation strategy, the described point

![Figure 7: The distances (expressed in mm) between points measured in different conditions for the horizontal (up panel) and the vertical (down panel) sensors.](image)
registration process could be iterated \( k \) time, random selecting, at each step, the training subset \( T_k \subset P \) (with \( |T_k| = s \)) used for the registration and computing the average distance (between a point seen from the horizontal sensor and the roto-translation of the same point seen from the vertical sensor), over the remaining set \( P \setminus T_k \). At the end, the training set characterized by the lowest distance was selected and the corresponding roto-translation matrix stored in order to be used by the server in run-time. This choice was made to avoid that points affected by not negligible measurement errors could affect negatively the calibration process. \( P, k \) and \( s \) were the input of the calibration function.

### 3.2 Tracking Module

The proposed approach could be easily employed in a hand tracking system. The process to obtain the positions, illustrated in Figure 8, was based on a control switching approach: both LEAPs acquired their frames stream, but only one, depending on the rotation of the hand with respect to the horizontal LEAP references system, is used to represent the hand (Figure 9). Both LEAPs were simultaneously turned on and remained in this state for the whole session because the time necessary to switch on and off a sensor would be too high for real-time purposes. At each time, only one LEAP was selected as “active” and the corresponding frame was acquired and used to track the hand. The vector \( v \), orthogonal to the palm and usually used by the sensor to estimate the hand orientation, was used to compute the roll \( r \) of the hand, that is the angle between the X axis and the projection of the vector on the X-Y plane, with respect to the horizontal LEAP reference system. As shown in Figure 9, if \( r \) was in the range from \( 225^\circ \) to \( 315^\circ \) (the palm was facing downwards) or in the range from \( 45^\circ \) to \( 135^\circ \) (the palm was directed upwards) the horizontal LEAP was selected as active, while the vertical was “paused” (it was still turned on, but its frame rate was reduced to the minimum in order to save computational resources). Out of these ranges, the vertical and the horizontal LEAPs had to act in the opposite way: the vertical was active and the horizontal was paused. The prototype has been developed following the architecture scheme shown in Figure 4: the server component received data from routers and manipulated the information from the vertical LEAP (that is the data...
from the Slave router) performing a roto-translation to obtain the coordinates with respect to the horizontal LEAP reference system. The server was responsible to check which LEAP was active, to send only the information received from it to the client and, if needed, to change the active status of the sensors. The client was obtained by using a modified version of an example application included in the LeapJS framework (http://github.com/leapmotion/leapjs, 2016), "Threejs-bones.html".

3.3 Preliminary Results

In order to observe the online behavior of the system, the calibration procedure was performed by using a cloud of 68 points \( P = 68 \). The size \( s \) of \( T_k \) was set to 5. The registration was iterated \( k = 3 \times 10^4 \) times. The points was acquired by means of the same tool used for the infrared interference evaluation (Figure 3). The average distance of the selected training set was \( err = 3.62 \text{mm} \), with standard deviation \( \sigma = 2.89 \text{mm} \) and a maximum value \( M = 17.36 \text{mm} \). As it is possible to observe from Figure 10, like in the infrared interference analysis, there was a small group of points with large distance values and the “error” was concentrated on the Z axis. This suggested that distance was mostly due to the intrinsic imprecision of the sensors along the Z axis than to a calibration lack in accuracy. Anyway, also in this case, the average distance was not too large with respect to the size of the fingertips and joints and the accuracy was acceptable. Figure 11 shows a set of hand positions and the corresponding model reconstructions, obtained with the proposed approach. The accuracy and the fluidity of the tracking process were adequate (about 25 frames per second) and the change of perspective did not produce any jumps or other visually observable effects. Moreover the hand was correctly tracked also in those positions that would be critical in the single sensor scenario.

4 CONCLUSIONS

A multi-sensor approach for hand tracking, based on the use of two orthogonal LEAP sensors, was presented. The hardware and software design choices and the resulting development was illustrated and discussed. A calibration procedure, aiming to align the reference system of the LEAP has been proposed and evaluated in term of accuracy. The detected discrepancy was not enough to compromise the operation of the system. Experimental data, to observe the behavior of the system in presence of infrared interferences, showed an adequate tolerance to the disturb. An interesting result regarded the displacement distribution along the axes: the Z direction seemed to be more affected by disturb with respect to the others (this occurred also during calibration). This suggested that, both in the infrared interferences and in the reference system roto-translation experiments, most of the discrepancy was not introduced by the proposed approaches, but was due to an intrinsic lack of accuracy of the LEAP along the Z direction. In order to explore the phenomenon, future developments will con-

![Figure 10: The distance computed along the axes during the calibration procedure.](image-url)
cern the use of a refined and graduated setup tool or an external position sensor ((Placidi et al., 2007)) to be used for the definition of a “real” reference system, in order to measure the accuracy of the LEAP and take it into consideration during the assessment of the proposed system effectiveness. Regarding the proposed hand tracking prototype, it could be improved by substituting the turnover strategy with the fusion of the models from the two LEAPs, instead of using just one of them at once. In this case also residual occlusions could be better overcome. Moreover, synchronization issues should be studied and managed. Finally, the proposed system will be tested for rehabilitation purposes.

REFERENCES


