Pottery gestures style comparison by exploiting Myo sensor and forearm anatomy

Dimitrios Ververidis, Sotirios Karavarsamis, Spiros Nikolopoulos and Ioannis Kompatsiaris
Centre for Research and Technology Hellas
6km Charilaou-Thermi, 57001 Thermi, Greece
ververid@iti.gr

ABSTRACT
In this paper we propose a set of Electromyogram (EMG) based features such as muscles total pressure, flexors pressure, tensors pressure, and gesture stiffness, for the purpose of identifying differences in performing the same gesture across three pottery constructions namely bowl, cylindrical vase, and spherical vase. In identifying these EMG-based features we have developed a tool for visualizing in real-time the signals generated from a Myo sensor along with the muscle activation level in 3D space. In order to do this, we have introduced an algorithm for estimating the activation level of each muscle based on the weighted sum of the 8 EMG signals captured by Myo. In particular, the weights are calculated as the distance of the muscle cross-sectional volumes at Myo plane level from each of the 8 Myo pods, multiplied by the muscle cross-section volume. Statistics estimated on an experimental dataset for the proposed features such as mean, variance, and percentiles, indicate that gestures such as "Raise clay" and "Form down cyclic clay" exhibit differences across the three vase types (i.e. bowl, cylinder, and sphere), although perceived as identical.

Author Keywords
Pottery; EMG; Myo; WebGL; Forearm; Anatomy; stylistic variations.

ACM Classification Keywords
H.5.2 User interfaces: Input devices

INTRODUCTION
The digitization of cultural heritage is an interesting field which, most of the times, involves the 3D scanning of museum artifacts. However, the Intangible Cultural Heritage (ICH), a term introduced by UNESCO, includes traditions or living expressions such as oral traditions, performing arts, social practices, rituals and festive events that is also important to preserve [12]. Our work targets at capturing the intangible characteristics of pottery art performances, but it can be also used for other kinds of arts with hands. In particular, our goal is to identify differences between the pottery gestures that may be perceived as apparently identical and are difficult to isolate and transmit when safeguarding ICH. Towards this direction, for the purpose of capturing arts involving hand movements, a gesture recognition technique is needed that will be able to cope with the problem of fingers occlusion and non physical availability of sensors (i.e. it is impossible to place accelerometers at finger tips).

Today, the hardware for gesture recognition is based on infrared (IR) sensors, Electromyograms (EMGs), and Inertial Motion Sensors (IMS) or combinations of them. Some IR sensors are based on projecting an infrared pattern in the room and capturing the pattern with an IR camera. A common pattern is a grid of dots. If the dots captured on a surface are dense, the surface is near the camera, and vice versa. Such devices are LEAP motion sensor [7] and MS Kinect (Version 1) [5]. The methods based on EMGs classify movement according to the electricity on the surface of the skin, which reflects the activity of the muscles. Myo sensor is such a device [13]. Methods based on IMS(es) are exploiting several wearable sensors worn by the user in several places in his body. The device we have employed is Myo that combines EMG and IMS sensors in the forearm, and therefore, it can be used non-intrusively for pottery sessions.

Most of the bibliography in EMG gesture recognition techniques involves a sensor made by a certain lab that it is not available for the community and therefore experiments cannot be reproduced accurately. Myo sensor is a relatively new device as it was released to the public at 2013 and offers a benchmark device at low cost ($200). It is a wireless device connected to desktops or mobiles offering a good user experience. Myo is ideal for art performances such as pottery, since the other options of IR and IMS based sensors can not be used due to physical limitations. In this work, we focus on gesture recognition at the level of fingers from EMG(s). In particular, we wish to extract information about the finger-gestures, which is a particularly valuable piece of information for the learner in order to improve his performance in pottery arts. In the literature, the gesture recognition methods that rely on EMG typically refer to hand-gestures. Although we do not treat the same task, we have reviewed some of these methods so as to highlight the most interesting components.
The process of gesture recognition is split into two parts, namely feature extraction and classification. A feature proposed in [6] is the entropy of splines fitted on the EMG signal and the entropy of splines fit on the residuals of every fit. This process leads to a dimensionality reduction of the original signal where entropy can be estimated to characterize the low frequencies of the signal. The classifier employed was based on linear discriminant analysis (LDA) using the Mahalanobis distance in order to classify signs for the deaf people with an industrial research EMG sensor. In [19], the Hidden Markov Models (HMMs) and decision trees were used to separate signs language from EMG and acceleration values by using as features the duration of the sign, the autoregressive coefficients of fourth order applied to short-term frames of 250ms, and the mean values of the signals in the short-term frames. For this purpose a custom EMG sensor was used. In [17], the mean value, the variance, the fourth order coefficient, and the sample entropy on short-term frames of 250ms were used to classify gestures to seven primitive classes, e.g. grasp, point, tip, rest and so on by using two industrial EMG sensors. LDA was used as a classifier to achieve more than 97% correct classification rate. The resulting decision was graphically depicted by using a VRML model of the palm, where each finger was represented by the concatenation of cylinders. Motivated from this work, we have foreseen the necessity of visualizing the arm in 3D space, so that experiments are performed in emulated conditions. However compared to [17], instead of using VRML and Matlab we have preferred WebGL and javascript which are more accessible to the community and they can lead to a ready to use solution.

In [20], a custom armband sensor is presented that comprises of 8 transmitters and 8 receivers of electrical signals at 40Hz, which is able to measure the electrical impedance of a forearm cross-section area. A tomography of the forearm is achieved that is able to predict the inner muscles activation level. The combination of the 8 sensors to 2 pairs achieves a feature vector of 28 impedance coefficients that are updated every 100ms. These 28 coefficients are the electrical impedances from each of the 8 emitters to each of the 8 receivers around the forearm. Electrical impedance changes with respect to the tension of the inner muscles because tensed muscles have different impedance than the relaxed muscles. Thus, by the cartography of impedance coefficients in a 2D map, the tension of the inner and outer muscles can be found. The Support Vector Machine (SVM) based classifier achieved about 95% correct classification rate for 11 gestures. Following a similar approach, our work was also based on the notion of cross-sections of the forearm, and was further augmented by projecting the virtual models of the muscles. Other recent achievements in myoelectric interfaces can be found in the review paper by Hakonen [4].

Finally, it is important to mention that most of the aforementioned research is performed with a toolkit for experiments, i.e. Matlab, SPSS, WEKA, LabView, etc. In this work, we have relied on HTML-5, Javascript and WebGL so as to develop a cross-platform and ready to use eco-system for myoelectric analysis, that can be easily used for every application. An additional advantage of the employed architecture is that the computational burden of handling 3D graphics lie on the client side, where WebGL is taking advantage of the available GPU. Similarly, most of the Myo signal recording processes are performed by webworkers that run on a parallel thread to avoid overloading the GUI thread [15].

The outline of this paper is as follows. The Database and Annotation tool section presents the data recordings and the procedure that has been employed to annotate them. In the Methodology section, the proposed 3D visualization tool and the type of features that have been introduced for our analysis are described. Details of the implementation are provided in Implementation Context for Myoelectric analysis section. The experimental results are shown next, where the proposed features indicate differences across the three types of vases for the same gesture. Finally, conclusions are drawn in the last section.

DATABASE AND ANNOTATION TOOL

For the purposes of our experimental study, we have assembled a dataset of Myo recordings during 9 pottery art performances from a master performer, namely 3 bowls, 3 cylindrical vases, and 3 spherical vases where each one lasts about 100 secs. Each vase type is shown in Figure 1.

![Figure 1. Data consists of Myo recordings of the right hand during three vase classes, namely (a) bowl, (b) cylindrical, and (c) spherical vases.](image)

The video of each performance was captured to help annotating the signals. The video frames were recorded with a custom C++ program synchronized to start with the Myo recorder. Each video frame is named according to the current unix timestamp, and each Myo sample is also associated with its unix timestamp. The sampling rate for EMG is 200Hz and for IMU is 50Hz.

Subsequently, the Myo recordings were annotated into basic pottery gestures using a custom made annotation tool with a graphic user interface. The tool was written in Python 3 using tkinter as a GUI constructor. The tool as shown in Figure 2, visualizes, in synchronization with the video frames, the 8 EMG signals from each Myo pod and the 9 IMU signals, namely acceleration, gyroscope, orientation for each of the x, y, z axis. The playline in the timeline can be dragged with the mouse and the video frame is updated on-the-fly. Each of the three types of
vases requires about 15 gestures and the total number of gesture classes is 33. The gestures for making a bowl and a cylindrical vase as annotated with our tool, are shown in Table 1.

<table>
<thead>
<tr>
<th>Bowl</th>
<th>Cylindrical</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2, Take clay</td>
<td>0.9, Take clay</td>
</tr>
<tr>
<td>4.4, Throw clay</td>
<td>2.7, Throw clay</td>
</tr>
<tr>
<td>5.2, Wash hands</td>
<td>3.5, Take water</td>
</tr>
<tr>
<td>8.1, Centering clay</td>
<td>4.8, Centering clay</td>
</tr>
<tr>
<td>10.5, Regulate speed</td>
<td>6.5, Raise clay</td>
</tr>
<tr>
<td>12.3, Raise clay</td>
<td>9.8, Form down cyclic</td>
</tr>
<tr>
<td>15.6, Form down cyclic</td>
<td>20.5, Take water</td>
</tr>
<tr>
<td>26.3, Take water</td>
<td>22.0, Form palms facing</td>
</tr>
<tr>
<td>27.8, Form palms facing</td>
<td>25.7, Form palms open</td>
</tr>
<tr>
<td>34.8, Form palms open</td>
<td>27.5, Take water</td>
</tr>
<tr>
<td>36.7, Take water</td>
<td>29.1, Sidewalls cyl1</td>
</tr>
<tr>
<td>39.3, Sidewalls 1</td>
<td>32.2, Take water</td>
</tr>
<tr>
<td>50.7, Sidewalls 2</td>
<td>34.1, Sidewalls cyl2</td>
</tr>
<tr>
<td>54.2, Pointer fix</td>
<td>51.3, Pointer finger</td>
</tr>
<tr>
<td>56.3, Take instrument</td>
<td>55.2, Sidewalls cyl2</td>
</tr>
<tr>
<td>59.3, Form with instrument</td>
<td>57.3, Pointer finger</td>
</tr>
<tr>
<td>74.4, Tiny finger</td>
<td>58.0, Take instrument</td>
</tr>
<tr>
<td>82.5, Leave instrument</td>
<td>60.4, Form with instrument</td>
</tr>
<tr>
<td>84.5, Take rope</td>
<td>77.0, Tiny finger</td>
</tr>
<tr>
<td>86.2, Pass rope under</td>
<td>78.5, Pointer finger</td>
</tr>
<tr>
<td>89.7, Remove vase</td>
<td>81.4, Form with instrument</td>
</tr>
<tr>
<td>96.7, End</td>
<td>89.0, Leave instrument</td>
</tr>
<tr>
<td></td>
<td>90.9, Take rope</td>
</tr>
<tr>
<td></td>
<td>92.7, Pass rope under</td>
</tr>
<tr>
<td></td>
<td>96.2, Remove vase</td>
</tr>
<tr>
<td></td>
<td>103.1, End</td>
</tr>
</tbody>
</table>

Table 1. Gestures required to make a bowl and a cylindrical vase with timestamps in seconds for two annotated sessions

METHODOLOGY

In order to analyze the correspondence between the EMG-based muscle activation and the pottery gestures, we need to visualize the forearm in 3D space and interactively observe if the proposed theory is reflecting the actual muscle-driven process.

Virtual forearm model resources

The muscles that are located at the level of Myo pods are investigated by the use of an anatomy atlas that also provides for free the 3D models of the bones and the muscles in Wavefront OBJ format [8]. The involved anatomical objects are 2 bones (radius and ulna) and 15 muscles that are enlisted in Table 2. Some muscles have two heads that are connected to different bones, e.g. humerus and ulnar heads. The functionality of each muscle is found in [11]. The cross-sectional volume is the volume of the mesh formed by intersecting a 0.5 cm thick plane with each muscle at Myo cross-sectional level. The result is the forearm scene that is depicted in Figure 3.

More specifically, the muscles at Myo level as indicated by the gray plane intersecting the forearm are the electrical sources contributing to the EMG signals. The intersection of the plane with each muscle produces a muscle tomography which we will call slice. The slice is found with a known algorithm in 3D graphics called binary space partitioning [14]. The volume and the center of mass of the slice are also calculated. Volume is calculated with summing the signed volume of tetrahedrons formed by the triangles (faces) topped at the origin [18]. Let us name \( m_i \) the slice of the \( i \)th muscle. The center of mass of the slice, \( \zeta_{m_i} = [x_{m_i}, y_{m_i}, z_{m_i}] \), is calculated by dividing the centers of all of its faces, \( \zeta_f \), multiplied by their face area, \( a_f \), and afterwards divide each vector with the total area of its surfaces, e.g.

\[
\zeta_{m_i} = \frac{\sum_{f \in m_i} a_f \zeta_f}{\sum_{f \in m_i} a_f}.
\]

The center of mass is a weighted sum of the faces’ origin depending on their area and it is not affected if the mesh has a high number of vertices in certain spots. For example, in Figure 4, a comparison of the weighted average of faces’ vertices (center of mass) versus the plain average of faces’ vertices (center of geometry) is shown. It is seen that the center of geometry is biased to the left because vertices are densely populated, whereas the center of mass is an unbiased estimator of the center of the mesh because it takes into consideration the area of the faces formed by the vertices.

Muscle activation level

The negative values of the EMG signals were clipped as muscle activation is denoted by the positive spike values, i.e. the discharges of the muscles. Negatives are the charges of the muscles that occur as a physical re-action to the discharge. Next, an average mean filter of the past 25 values was used to remove high frequency artifacts. Each signal of the 8 Myo sensors is a weighted sum of the electrical signals of the 15 muscles, where the weights depend on the distance of the muscle from the sensor and the volume of the muscle. It is assumed that the weight is proportional to the volume of the muscle, i.e. the bigger the muscle is, the more electricity it produces. However, the weight is inversely proportional to the distance, that is the further the muscle is from the sensor, the lowest is its contribution to the sensor’s signal. Let \( e_{m_i}(t) \) be the electrical activity signal of the muscle \( m_i \), where \( i = 1, 2, \ldots, 15 \), then

\[
e_{m_i}(t) = V_{m_i} \sum_{j=1}^{8} \frac{s_j(t)}{d_{ij}},
\]

where \( V_{m_i} \) is the volume of the muscle \( m_i \), \( s_j(t) \) is the signal of the \( j \)th Myo pod and \( d_{ij} \) is the distance of each pod to each muscle, i.e.

\[
d_{ij} = \sqrt{(x_{s_j} - x_{m_i})^2 + (y_{s_j} - y_{m_i})^2 + (z_{s_j} - z_{m_i})^2}.
\]

The numbering Myo Pods starts from -45 degrees as shown in Figure 5 and continues counter-clockwise until 270 degrees.

Features extraction
The features proposed in this work are based on the muscle activation in certain muscle groups, which in the general case are found by:

\[
f = \frac{1}{N} \sum_{t=1}^{N} \sum_{m \in M} e_m(t),
\]

where \(M\) is the muscle group. Depending on the muscle group \(M\) we define the following features.

The \(f_{\text{total}}\) found when \(M\) is the set of all muscles of Table 2 denotes the total muscle pressure of the gesture.

The \(f_{\text{flex}}\) is the flexion tension of the gesture which is found by the flex muscles \(M=\{\text{brachioradialis}, \text{palmaris longus}, \text{ext. car. rad.}, \text{Uln. hea. lef. fle. car. uln.}, \text{Hum. hea. lef. fle. car. uln.}, \text{ext. dig. sup.}, \text{ext. dig. pro.}\}\) denoting the force put towards the anterior of the forearm. In the same manner extension tension \(f_{\text{ext}}\) of the gesture is found by the 5 extensors, i.e. \(M=\{\text{ext. dig.}, \text{ext. car. uln.}, \text{ext. car. rad. lon.}, \text{ext. dig. min.}, \text{ext. car. rad. bre.}\}\) denoting the force put towards the posterior of the forearm. The supination feature \(f_{\text{susp}}\) is defined by using only the supinator muscle, and the pronation feature \(f_{\text{pron}}\)
Figure 3. Intersecting the muscles with a 0.5cm thick plane in WebGL environment.

Figure 4. Comparison between (a) the center of geometry, and (b) center of mass.

is defined by using the \{“Uln. hea. lef. pro. ter.”, “Hum. hea. lef. pro. ter.”\} set of pronation muscles.

Finally, we denote as stiffness $f_{stif}$ the steadiness of the palm towards forearm that is caused by balanced force of the extensors/flexors and suspinators/pronators, i.e. the opposite muscle groups that put opposed forces to make the hand stiff to be defined as

$$f_{stif} = f_{flex}f_{ext} + f_{susp}f_{pron}. \quad (5)$$

IMPLEMENTATION CONTEXT FOR MYOELECTRIC ANALYSIS

A 3D engine was necessary to render the muscles for visualizing the whole forearm. The reason for favoring a solution based on web-browsers was due to the following: a) they are accessible from everywhere without any posing any software requirements; b) they offer a rich development environment as they include a console that can be used to interactively experiment with the loaded web page; and c) they support the WebGL of HTML5 that renders 3D graphics with the client’s Graphic Processing Unit (GPU). However, in order to cope with the problem that WebGL has low level commands that are not easy to use for experiments, we have employed the Three.js framework [2] which is a javascript framework for WebGL 3D graphics that has high level commands for rendering scenes.

On top of the aforementioned infrastructure, several javascript libraries were used in order to transform the web page into a full Myo observation console. In order to port Myo signals to a web page the Myo.js library was used [15]. The library ports values from the Myo desktop drivers to the web page through websocket technology. The Myo signals were plotted into the web page through a jquery library named as flot [10]. The signal values were also shown in an interactive user interface with the help of dat.gui library [3].

In this way, using the web browser the muscles electrical activity can be observed in real-time as shown in Figure 5. A video demonstrating the functionality of the tool can be found in [16]. In this video, the muscles electrical activity and the Myo pods signal values are represented with a yellow overlay on the respective objects.

EXPERIMENTAL RESULTS

Pottery gestures such as “Raise clay”, “Form down cyclic”, “Form with little finger”, “Form with palms open”, “Form with palms closed” are in common among the 3 types of vases, and therefore, constitute an interesting case to be analyzed using the proposed features. Other gestures in common such as “Take water”, “Take instrument”, “Leave instrument”, “Take rope”, “Pass rope under” do not exhibit style differences as they are very simple to perform. The features used in our experiments are the pressure, stiffness, flexors pressure (flexion), and extensors pressure (extension). Statistics such as the mean, median, the variance, and the 25/75 percentiles were estimated on these features across all instances per class (whereas by class we refer to the type vase being constructed, i.e. bowl, cylinder, sphere). The results are plotted in box plots [1].

In Figure 6, “Raise clay”, i.e. the gesture to pull clay upwards while it is rotating on the table, is compared across the 3 vase types with respect to the total muscle pressure feature. In the bowl vases the user seems to put high pressure across all of his attempts. In cylinder vases the pressure is much less, whereas in spherical vases the pressure varies across the attempts with a tendency to the high values. Based on these results, we can infer that cylindrical vases do not need much pressure when raising up the clay.

In Figure 7, the stiffness of the forearm when performing the “Tiny finger” gesture, i.e. the formulation of the upper rim of the pot with the little finger, is depicted. It is seen that when making a spherical vase the stiffness of the little finger should be high, which might be explained by the sensitivity of the narrow rim of the vase which requires more precise formulation with a stiff tool.

In Figure 8, the flexion of the forearm when performing the “Centering clay” gesture, i.e. the transformation of the raw clay into a rounded mass on the spinning table,
Figure 5. Live view of Myo signals and muscle electrical activity in the cross-section model of the muscles.

is depicted. The flexion of the forearm is an indication of press down and centering force put on the clay which is higher in the cylindrical and the spherical vases. This might be due to the higher height of cylinder and sphere that require a more condensed clay that can stay still when it is pulled up.

In Table 3, we present the mean value of 4 features for 4 gestures across all attempts per each of the 3 vase types. It is seen that significant differences exist across vase types within each gesture. For example, for “Raise clay” the stiffness and the total muscle pressure are higher in bowl than in the other vases. Over-viewing the values across gestures, it is validated that “Tiny finger” gesture has smaller muscle activation level than the “Raise

Figure 6. Total muscle pressure during “Raise clay” gestures across three types of vases.

Figure 7. Stiffness during “Tiny finger” gestures across three types of vases.

Figure 8. Flexion feature during “Centering clay” gesture across three types of vases.
The developed framework, is not just limited to forearm but it can be extended to any part of the body and it could be potentially connected to any EMG device that has the proper middle-ware (websocket interface). In the future, the accelerometer and the gyroscope of Myo will be connected to our framework, so that the 3D forearm can move in the virtual environment and offer a more insightful visualization. In addition, the muscles will be animated so that the electrical activity will drive the visualization of the finger movements and the forearm.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Union 7th Framework Programme (FP7/2011-9) under grant agreement 600676 corresponding to project i-Treasures, “Intangible Treasures - Capturing the Intangible Cultural Heritage and Learning the Rare Know-How of Living Human Treasures”, http://i-treasures.eu.

REFERENCES