A Computational Method to Automatically Detect the Perceived Origin of Full-Body Human Movement and its Propagation

Olga Matthiopoulou
University of Genoa
4689588@studenti.unige.it

Benoit Bardy
University of Montpellier
benoit.bardy@umontpellier.fr

Giorgio Gnecco
IMT Lucca
giorgio.gnecco@imtlucca.it

Denis Mottet
University of Montpellier
denis.mottet@umontpellier.fr

Marcello Sanguineti
University of Genoa
marcello.sanguineti@unige.it

Antonio Camurri
University of Genoa
antonio.camurri@unige.it

ABSTRACT

The work reports ongoing research about a computational method, based on cooperative games on graphs, aimed at detecting the perceived origin of full-body human movement and its propagation. Compared with previous works, a larger set of movement features is considered, and a ground truth is produced, able to assess and compare the effectiveness of each such feature. This is done through the use of the Shapley Value as a centrality index. An *Origin of Movement Continuum* is also defined, as the basis for creating a repository of movement qualities.

CCS CONCEPTS

• **Human-centered computing**; • **Computing methodologies**;

KEYWORDS

Full-body movement analysis, perceived origin of movement, automatic detection, game theory, graph theory

ACM Reference Format:
Olga Matthiopoulou, Benoit Bardy, Giorgio Gnecco, Denis Mottet, Marcello Sanguineti, and Antonio Camurri. 2020. A Computational Method to Automatically Detect the Perceived Origin of Full-Body Human Movement and its Propagation. In *Companion Publication of the 2020 International Conference on Multimodal Interaction (ICMI '20 Companion), October 25–29, 2020, Virtual event, Netherlands.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3395035.3425971

1 INTRODUCTION

The perceived origin of movement is the point at which a movement appears to originate from the point of view of an observer. A first proposal of a computational method to detect the perceived origin of movement was made in [1] and further investigated in [2]. The method is based on a mathematical game built over a suitably-defined graph structure representing the human body. Our objective here is to further develop this computational method for the analysis of expressive full-body movement qualities, by extending the findings of [1, 2] and providing a refined implementation. Differently from such works, we consider several movement features in the implementation of the method, namely speed, tangential acceleration, and angular momentum (see Section 3), in order to see which feature is best at predicting the origin of movement when this is computed using Shapley values (a well-known solution concept in cooperative game theory). So, we evaluate how the perceived origin of movement is affected by different movement qualities.

The remainder of this paper is structured as follows. Section 2 details the state of the art and the motivations for this research. Next, Section 3 covers the details of our approach, focusing on a cooperative game index used to estimate the importance of joints, on the description of our dataset, and also outlining the new movement features we introduce in this work. Then, in Section 4, we discuss the ongoing annotation process and an online tool we developed to collect ground truth. Section 5 presents our first results. In Section 6, we indicate our ongoing work about creating an innovative taxonomy of movement. Lastly, we conclude and discuss possible future extensions of this research.

ICMI '20 Companion, October 25–29, 2020, Virtual event, Netherlands

ACM ISBN 978-1-4503-8002-7/20/10…$15.00
https://doi.org/10.1145/3395035.3425971

2 STATE OF THE ART

As detailed in the next paragraphs, the motivation for this research is threefold: 1) the role of body expressions in movement analysis; 2) the connections between neuroscience and movement analysis; 3) the leading joint hypothesis [3].

First of all, several studies (see for example the survey paper [4]) have investigated the role of full-body expressions as a tool for communication. Results showed that full-body expressions convey a large amount of information, much greater than previously thought for non-verbal communication. This suggests that both form and motion information are highly important in the perception of affect from full-body expressions.

Secondly, cognitive neuroscience and full-body movement analyses are closely related. The brain organises and categorises human movement in order to extract meaning from it [5]. Different areas of the brain are associated to specific roles when it comes to the perception of movement features. It has been shown that the brain organises the perception of body movements through form and features of the movement rather than semantically.

Lastly, the most important motivation for this research comes from the literature on the leading joint hypothesis on limb motion, which states that "there is one leading joint that creates a dynamic foundation for motion of the entire limb" [3]. In other words, joints of a limb all have different roles when it comes to the production of a movement, according to how they are organised in the joint linkage structure. For example, when considering a punching motion, seeing arm and hand in motion does not necessarily mean that these joints are the origin of movement. In fact, when throwing a punch, the energy and power may originate from the shoulder joint, or from one foot. Hence, in this instance, the shoulder joint (or a foot joint) would be considered as the leading joint.

Based on this hypothesis, following [1, 2], we can define what we mean by the Origin of Movement. Gestures have different meanings based on their origin of movement, i.e., the exact position (joint) of the body which initiates this movement . We split the definition of the Origin of Movement into two scales. The *acted origin of movement* refers to how the muscle works at the leading joint, i.e., focus is given to the biomechanical level. The *perceived origin of movement* is considered from the viewpoint of the observer and is defined as the point on the body from which a movement appears to originate. Movement perception from external observation is an important topic that has been investigated, e.g., also in [6, 7].

This research starts from the results obtained in [1, 2] about the automatic detection of the perceived origin of movement. Their approach is based on a cooperative game-theoretical model, specifically a transferable-utility game, built over a suitably-defined graph structure representing the human body. The players of this game are graph vertices, forming a subset of body joints. Each vertex contributes to a shared goal, that goal being the way a specific movement feature is transferred among the joints. This cooperative game-theoretical model is constructed in such a way to allow one to measure the relevance of various segments in human movement when performing full-body movement analysis. Each group of players, called a coalition, has an associated utility, which represents their joint contribution to the common task. The adopted approach looks at movement features associated with each joint, namely speed in the case of [1, 2]. The results from [1] and the additional validation from [2] show that this approach allows one to automatically detect the origin of movements performed by dancers.

As stated previously, our main aim is to extend this approach in order to encompass more movement features. We also establish a ground truth for each movement feature based on annotations made on our dataset by experts and non experts.

3 DETAILS OF THE METHOD

3.1 Shapley Value

The Shapley value is a popular solution concept in game theory. It is a measure of the value of each player in a transferable utility game [8]. The general idea behind the Shapley value is that the importance of each player equals its average marginal contribution to the payoff of a coalition when the player joins that coalition (the average being computed with respect to a suitable probability measure on the set of such possible coalitions). The Shapley value represents a fair way to allocate to the players the utility of the coalition made of all of them (this is called the grand coalition).

In our application, we use the Shapley value to measure the *relevance* or *importance* of a vertex (joint) in the graph-based transferable utility game proposed in [1]. In simple terms, for such a game, the joint with the maximum Shapley value at a given moment is an estimate of the leading joint. In this way, a prediction of the origin of movement is provided. Moreover, for a specific movement, the Shapley values of all the joints provide a way to rank those joints according to their different degrees of leadership.

3.2 Dataset and Marker Set

Our dataset consists of 36 movement fragments from multiple subjects, recorded with the Qualisys motion capture system and synchronized via SMPTE with two videocameras (front and side views available). The videos consist of the subjects, some of which are expert dancers, performing simple sequences of movement with a clear perceived origin. Accompanying these videos and their MoCap position data are annotations for each recording as to which parts of the body are moving, the sequence they move in, as well as the intended origin of movement. These expressive movements represent normal full-body movements characterised, in any case, by a clear perceived origin. For example, a video with the annotation "Left shoulder pulls body making it turn" would be accompanying a video where the subject performs several trials of this movement, the shoulder being the origin and the rest of the body following. As such, when using these videos to predict the origin of movement using a specific movement feature, we can have a specific baseline against which we can compare the effectiveness of these features.

Using the Motion Capture technology, the subjects were fitted with 62 reflective markers, from which specific position data *(x, y, z)* were extracted per video and per frame. This constitutes the *full marker set*. This full marker set was then transformed into a *reduced marker set* considering each cluster obtained by combining several markers from the full marker set as one joint in the reduced marker set. This was done by creating a mapping from the full to the reduced marker set, and by calculating the barycentre of the joints in each cluster. For example, in the reduced marker set, the joint named "left hand" corresponds to the barycentre of the markers associated with the left palm and the left fingers.

As a result of this reduction of the maker set, there could be some information lost in this mapping. Thus we defined a *Mass Distribution feature*. The aim of this feature is to determine how much information is lost when moving from the full to the reduced marker set. Our assumption is that each cluster in the full marker set (which becomes a node in the reduced marker set) behaves approximately like a rigid body. We defined this Mass Distribution feature as the root mean square distance of the markers in each cluster from its barycenter, i.e., as $\sqrt{\frac{\sum_{i=1}^{n}\|\underline{x}_i-\bar{\underline{x}}\|^2}{n}}$, where n is the number of joints in a cluster, $\underline{x}_i$ the position vector of each joint in the cluster (this information comes from the motion capture data), and $\bar{\underline{x}}$ is the barycentre position vector of that cluster. The same mass is assumed for each joint of the same cluster.

The output is a scalar for each cluster (at each frame), expressed in millimetres. The purpose of this feature is to observe if the mass distribution in each given cluster remains constant over time. If indeed we see no significant change over time, then we can assume that the joints defining each cluster in the full model are almost rigid with respect to one another. Hence, the reduced model can be thought of as a good approximation of the full model in that part of the body. However, if we see that there is a lot of change over time, this would mean that some information is lost when moving from the full to the reduced marker set. We then computed for each cluster the *Coefficient of Variation* of this feature - the ratio of its empirical standard deviation with respect to its empirical mean - to evaluate how much dispersion exists in the data over time. The results are shown below in Table 1.

Table 1: Coefficient of Variation for each cluster of joints

head	0.08	**shoulder centre**	0.09	**hip centre**	0.06	**spine**	0.09
left elbow	0.02	**right elbow**	0.35	**left foot**	0.55	**right foot**	0.65
left hand	0.12	**right hand**	0.13	**left hip**	0.27	**right hip**	0.35
left knee	0.03	**right knee**	0.03	**left wrist**	0.06	**right wrist**	0.24

As we can see, for some clusters the variation is minimal, thus we can assume these to act as rigid bodies, while others, such as the feet or hand markers, have a high empirical standard deviation relative to the respective empirical mean. It is important to remark that this outcome could have been influenced by possible measurement errors while the data were being recorded. The markers on the extremities of the body could have been displaced from their original positions, since these moved a lot as the participants performed the movements. As such, we ignore this variation for the extremities of the body, and assume these to also act as rigid bodies.

3.3 Movement Features

With respect to [1, 2], here we consider a larger number of movement features (only speed was considered therein). In this way, it is possible to compare the results obtained using different features.

Speed. We are currently using at each joint several movement features and estimating their effectiveness in determining the origin of movement in a given frame. The first of these features considered is *speed* (i.e., the magnitude of the tangential velocity), which was also used in [1, 2], allowing a direct comparison with existing research. We calculate this speed as the norm of the velocity vector (exploiting the *np.linalg.norm* method in Python).

Tangential Acceleration. The second movement feature we consider is *tangential acceleration* (i.e., the derivative of speed).

Angular Momentum. Another movement feature we consider is *angular momentum*. Let us consider a possible scenario in which we have a movement where different parts of the body are rotating with respect to different axes passing through the centre of mass of the body. In order for these segments to have the same (or similar) direction of the angular momentum, they would have to be rotating together. The aim here is to find clusters of joints with similar direction of angular momentum. In this case, as a measure of similarity of this feature, we are currently considering Cosine Similarity. We calculate the Cosine Similarity between any two adjacent nodes, e.g., between hand and wrist. By doing this, it is possible to compare the directions of the angular momenta associated with the two nodes, and thus evaluate if two nodes are associated with (nearly) the same direction.

However, Cosine Similarity is just a measure of similarity, and it can be also negative, having the range $[-1, 1]$. In order to apply the game-theoretical framework proposed in [1] to a generic movement feature, we need both a non-negative dissimilarity measure and a non-negative similarity measure, since both are used to define the utility of each coalition (details are in [1]). We are currently dealing with this issue by just adding 1 to the Cosine Similarity. Hence, pairs of nodes that have nearly the same direction of the angular momentum will have a "corrected" Cosine Similarity close to 2.

4 ONLINE ANNOTATION TOOL

At this stage, we have developed an online tool that will be used to acquire annotations on our dataset from both experts and non-experts. The participants view each video from our dataset and select what they believe to be the leading joint. We have also extended the original online tool presented in [2], so that the participants can select also a second leading joint. In order to compare the results of the analysis with the ground truth provided by the participants, we consider the two joints with the largest Shapley value(s), not only a single joint with the largest Shapley value. This is because, as shown in [2], it sometimes occurs that there are two nodes with the largest Shapley value (or one with the largest Shapley value, and the second one with a nearly identical Shapley value), and these nodes are typically connected by an edge in the graph.

As a result, by using this approach of having the participants choose what they believe to be the origin of movement, we reach a sort of ground truth that possibly changes from participant to participant. Finding several frames in a recording for which the inter-participant agreement is strong suggests that in these frames there is a clear perception of the origin of movement.

5 RESULTS

Thus far, results of our analysis are in a preliminary phase, as we have not yet concluded their validation via the online annotation tool. The main analysis will come from examining the results of the survey, arriving at a ground truth and extracting relevant frames from this.

However, there are some initial results we can describe at this point. As stated before, the dataset consists of videos accompanied by annotations on the movements performed. These primary expert annotations can be used, on a given fragment of a video, to measure the accuracy (in percentage) of a movement feature in predicting the perceived origin of movement. These initial expert annotations comprise the perceived origin of movement for each fragment, which can be compared, e.g., to the one predicted by considering the two (adjacent) joints with the two largest Shapley values.

For each fragment, manual annotations of the origin of movement were recorded. Using the MATLAB code we have developed, for each fragment and for each frame, we obtained the joints with the 10 largest (possibly repeated) Shapley values. In other words, we used a specific movement feature in combination with the index (1^{st}, 2^{nd}, and so on) of the Shapley value in order to determine the accuracy of that feature in predicting the origin of movement. However, as we were not interested in all 10 joints, we focused our attention on the top two joints.

Hence, we can compare the perceived origin of movement from the manual annotations to the origin of movement predicted by a certain movement feature. This comparison is performed in three ways. First of all, we compare the ground truth against a random choice. This is to create a base case, as we expect that our results yield better accuracy than the random choice. We also assume that in the case of the random choice, the accuracy will be approximately 5% since 20 joints are considered in total.

The second comparison is between the ground truth coming from the manual annotations against the joint with the largest Shapley value. Thirdly, we compare the joints with the first and second Shapley values to the perceived origins of movement and we obtain an accuracy score, conditioned on the difference between the first and second Shapley values being less than 0.05.

As an example, Table 2 shows the results of this comparison for a specific fragment in our dataset, based on all movement features.

Table 2: Summary of results for a single fragment

Fragment t_028.3	Speed	Acceleration	Ang. Momentum
Random Choice	5.81%	5.44%	5.43%
First Largest Shapley Value	19.42%	13.07%	19.61%
First & Second Largest Shapley Values	80.36%	51.18%	50.89%

Table 2 details, for a specific fragment (t_028.3), and for each of the three movement features considered in this work, the results of the comparison between the labeled perceived origin of movement and its prediction obtained using each of the three methods described above (random choice, first largest Shapley value, first and second largest Shapley values). We can see, as expected, that the application of the random choice method yields a low accuracy score, close to 5% for all the movement features. From the table, we can also see that speed produces a rather good accuracy score for the third comparison. Indeed, overall the third comparison yields better results as it uses a more relaxed matching between Shapley values and ground truth. Thus, we can deduce that speed seems to be a somewhat good determinant for the perceived origin of movement, while there is still room for improvement for the other two features. However, one has to take into account that these conclusions hold only for one specific fragment, hence it is not currently possible to extend them to other fragments.

These results, while they can give us an initial idea of the performance of each movement feature, do not tell us much about the relevance of certain frames. Also, the annotations which comprise the perceived origins of movement were made by only one human, so are bound to produce incomplete and preliminary results.

6 ONGOING AND FUTURE WORK

A main focus is on the refinement of the *origin of movement* concept, as a foundation for the design of a rich multi-modal repository of movement qualities.

Further discussions with choreographers and other movement experts are leading to an improved model of this quality.

In particular, a refinement of the concept of origin of movement is the following *Origin of Movement Continuum* model. We define two different types of origin of movement (external and internal) as boundaries of this continuum, as follows:

(i) External: a joint that continuously leads the movement for a time interval, and maintains the energy in this joint. All other joints follow, which causes the movement of the surrounding joints. An example is a full-body movement where the hand is drawing a shape in the surrounding personal space. As a result, the energy at this joint is maintained with the other adjacent joints following, and the hand is the evident origin of movement. An origin of movement of this type may be clear to perceive, and changes across the body from one joint to another is most often clearly visible to an observer.

(ii) Internal: on the other boundary of the continuum, the origin of movement can be thought of a sudden “sparkle of energy” originating at a single joint. An example is a sudden right hip tangential acceleration that causes a propagation of a spiral movement in the body. This sort of “gush“ emerging from the body then propagates to the surrounding joints. This second type of origin of movement is therefore hardly predictable. It may have no obvious preparation movement, it may not be explicitly communicating, and may emerge from an intimate, inner behavior. An observer perceives this internal origin of movement as a sudden source of energy that starts from a joint and propagates in the body.

We are currently working at the creation of a rich repository of different cases of origin of movement. This implies the creation of a repository including a taxonomy of movements, to help in solidifying the proposed continuum model, but also to support further research activities as it can be used as a core element to construct a protocol for multimodal recordings. In essence, we aim to design a taxonomy capturing the quality of movement that will be useful for the validation and evaluation of the computational model investigated.

There are several ways to structure this taxonomy of movements: in terms of the body parts involved; in terms of proximal to distal movements; in terms of specific actions that are asked to the performer. These research questions will be faced in future activities of the EnTimeMent project.

Possible future extensions of this research include how multiple temporal scales can be incorporated. For example, we could look at a fast temporal scale at the very first moment of the origin of movement, and also at a slower temporal scale where we could analyse the origin of movement at a higher level. Thus, multiple temporal scales would be useful in prediction and analysis, resulting in higher level analyses.

Finally, another extension considers small groups of people and examines the emergence of the origin of movement in groups. In this case, a group would be considered as a single organism. Thus, in the context of graph and game theory, instead of considering joints as players, we would use similar techniques at a higher level, that of the individuals (and not the joints), and analyse the concept of origin of movement in terms of leadership in a team.

7 ACKNOWLEDGEMENTS

The work received support by the 4-year European Project n. 824160 FET PROACTIVE EnTimeMent (http://entimement.dibris.unige.it). G. Gnecco and M. Sanguineti are members of GNAMPA-INdAM (Gruppo Nazionale per l’Analisi Matematica, la Probabilità e le loro Applicazioni - Istituto Nazionale di Alta Matematica). The work of G. Gnecco was supported in part by the Italian Project ARTES 4.0 – Advanced Robotics and enabling digital TEchnology & Systems 4.0, funded by the Italian Ministry of Economic Development (MISE). The work of M. Sanguineti was supported in part by the National Research Council of Italy (CNR), where is Research Associate under Project PDGP 2018/20 DIT.AD016.001 (Technologies for Smart Communities).

REFERENCES

[1] K. Kolykhalova, G. Gnecco, M. Sanguineti, A. Camurri, and G. Volpe. Graph-restricted game approach for investigating human movement qualities. In *Proceedings of the 4th International Conference on Movement Computing*, MOCO '17, 2017.

[2] K. Kolykhalova, G. Gnecco, M. Sanguineti, A. Camurri, and G. Volpe. Automated analysis of the origin of movement: An approach based on cooperative games on graphs. *IEEE Transactions on Human Machine Systems*, to appear, 2020.

[3] N. Dounskaia. Control of human limb movements: The leading joint hypothesis and its practical applications. *Exercise and sport sciences reviews*, 38:201–208, 2010.

[4] A. Kleinsmith and N. Bianchi-Berthouze. Affective body expression perception and recognition: A survey. *IEEE Transactions on Affective Computing*, 4(1):15–33, 2013.

[5] M. J. Vaessen, E. Abassi, M. Mancini, A. Camurri, and B. de Gelder. Computational feature analysis of body movements reveals hierarchical brain organization. *Cerebral Cortex*, 29(8):3551–3560, 2019.

[6] A. Koul, M. Soriano, B. Tversky, C. Becchio, and A. Cavallo. The kinematics that you do not expect: Integrating prior information and kinematics to understand intentions. *Cognition*, 182:213–219, 2019.

[7] S. Sepp, S. J. Howard, S. Tindall-Ford, S. Agostinho, and F. Paas. Cognitive load theory and human movement: Towards an integrated model of working memory. *Educational Psychology Review*, 31:293–317, 2019.

[8] S. Tijs. *Introduction to game theory*. Hindustan Book Agency, 2011.