Emotion Expression in Human Body Posture and Movement: A Survey on Intelligible Motion Factors, Quantification and Validation

Mehdi-Antoine Mahfoudi, Alexandre Meyer, Thibaut Gaudin, Axel Buendia, Saida Bouakaz

▶ **To cite this version:**

Mehdi-Antoine Mahfoudi, Alexandre Meyer, Thibaut Gaudin, Axel Buendia, Saida Bouakaz. Emotion Expression in Human Body Posture and Movement: A Survey on Intelligible Motion Factors, Quantification and Validation. IEEE Transactions on Affective Computing, 2023, 14 (4), 10.1109/TAFFC.2022.3226252 . hal-03899236

HAL Id: hal-03899236
https://hal.science/hal-03899236v1

Submitted on 5 May 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

© 2022 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Digital Object Identifier (DOI): 10.1109/TAFFC.2022.3226252

Emotion expression in human body posture and movement: a survey on intelligible motion factors, quantification and validation

Mehdi-Antoine Mahfoudi, Alexandre Meyer, Thibaut Gaudin, Axel Buendia and Saida Bouakaz

Abstract—Many areas in computer science are facing the need to analyze, quantify and reproduce movements expressing emotions. This paper presents a systematic review of the intelligible factors involved in the expression of emotions in human movement and posture.
We have gathered the works that have studied and tried to identify these factors by sweeping many disciplinary fields such as psychology, biomechanics, choreography, robotics and computer vision. These researches have each used their own definitions, units and emotions, which prevents a global and coherent vision. We propose a meta-analysis approach that cross-references and aggregates these researches in order to have a unified list of expressive factors quantified for each emotion.
A calculation method is then proposed for each of the expressive factors and we extract them from an emotionally annotated animation dataset: Emilya. The comparison between the results of the meta-analysis and the Emilya analysis reveals high correlation rates, which validates the relevance of the quantified values obtained by both methodologies. The analysis of the results raises interesting perspectives for future research in affective computing.

Index Terms—Emotion, Body Expression, Survey, Meta-analysis, Human motion, Posture, Human Movement, Motion analysis, Emilya

1 Introduction

Emotion is a complex phenomenon that is difficult to formalize. According to the American Psychological Association, emotional experiences involve three components: a subjective experience, a physiological response and a behavioral or expressive response. The interpretation of an emotion is subjective, as two different people can perceive and interpret the same emotion differently [1]. Likewise, perception of emotion creates expressions which change from one culture to another [2], [3]. Furthermore, the complexity of an expression increases even more as humans express it through different channels such as facial expressions, speech, postures and movement. Thus, extensive research in multiple research areas has been conducted to understand how humans express emotions motivated by many applications such as disease detection and therapy in psychology or medicine [4], [5]; expression recognition in computer vision [6]; or synthesis of emotional gestures in computer graphics [7]. Abundant literature exists working on different input modalities such as speech, text, physiological signals, images, video, skeleton, etc. using techniques from different areas such as signal processing, statistics, machine learning, etc.

Verbal expression provides an important medium for the expression of emotion [8], nevertheless it is now widely accepted that nonverbal behavior constitutes another important way of communication in addition to speech. Even if facial expression has been widely studied in the area of nonverbal expression [9], several studies from various domains have shown that body expressions are as powerful as facial expressions [10]. In computer science, with the growth and easy access of devices that track 3-dimensional body, like the Kinect [11] or accelerometer based motion capture system [12], body movements become common data. Our assumption is that many applications would benefit from any advances in body expression understanding in order to provide more natural interactions in video games, robotics, human-computer interaction, artistic creation, etc. Some of these applications may favor fully automated approaches, such as when only an input/output emotion label is needed [13], [14]. Other applications might need a set of intuitive parameters associated with values for each expression, in order to edit or analyze them [15], [16], [17], [18], [19].

A movement is commonly represented in computer science by a succession of poses. Each pose is represented by a three dimensional position of each joint or an angle between each successive body parts. Understanding where an emotion is hiding in such data is a non trivial task. In this study we focus on finding parameters that remain intelligible to a human. The key challenge is to find a list of parameters able to encode the emotions for any action such as running, jumping, kicking, etc. A computer animator might want to change one of these parameters to produce a more pronounced expression of a cartoon character [20]. A virtual reality application could analyze the emotional contents of user gestures [21] in order to produce a symmetric expression to a non human virtual character. The movements of a robot might suggests an emotion using such parameters, even if its morphology is limited [22]. We expect to identify understandable parameters such as the tilt of the

- *M. Mahfoudi, A. Meyer and S. Bouakaz are with Univ Lyon, UCBL, CNRS, INSA Lyon, LIRIS, UMR5205, F-69622 Villeurbanne, France.*
- *M. Mahfoudi, T. Gaudin and A. Buendia are with SpirOps, 75011 Paris, France.*
- *A. Buendia is with CÉDRIC, Conservatoire National des Arts et Métiers, 75003 Paris, France.*

spine or a kind of energy measure of each joint movement. Intuitively, a person leaning forward or with low energy motion will reflect a more negative feeling than a person with a bulging torso or with a dynamic motion.

In order to understand how the emotion is expressed in human posture and motion, many research fields such as psychology, arts, computer vision and computer graphics have proposed many comprehensible factors based on intuitive concepts, sometimes calling them parameters or descriptors. But in all publications, these descriptors are completely heterogeneous, either in their definition or units. In this paper, we propose to make a systematic review of all the comprehensive descriptors described in those fields. We have developed a method that unifies these heterogeneous results into a quantified list of expressive descriptors. The same descriptors have been measured in a motion capture database [23], [24]. We finally compared the results of the two analyses, and quantified the correlation between the two approaches.

2 Expressing emotion through body motion

Several studies from various domains have shown that body expressions are as powerful as facial expressions [10], [25] and this have been verified in many situations such as education [26], [27], sport [28] or communication [29]. According to [25], [30], body language includes different types of non-verbal indicators embodied in the postures and in the gestures: hands, head positions or movement; opening or closure of the shoulders and the torso; tempo, repetitions, energy of a gesture; etc. All these small imperceptible indicators change the way a person realizes an action such as walking, sitting, standing, catching objects or lying down, etc. It is commonly accepted [31] that it provides some cues about the inner emotional state of a person.

More formally, the psychology domain defines what are called general movement protocols, which consist of a set of cues for each emotion. These cues include textual description, often accompanied by values presented in a table. The computer vision domain talks about descriptors used by algorithms of classification for automatic emotion recognition. The field of computer graphics refers to the parameters, often provided to an animator in order to edit an animation. These three concepts called cues, descriptors and parameters all refer to the factors responsible for the bodily expression of emotion. The field of psychology seeks to understand the mechanism of the expression and the factors influencing the perception. The field of art has also contributed to formalize the relationship between movement and expressiveness [32], [33], [34]. Historically, the field of affective computing has used this knowledge to provide automatic approaches to recognize emotion expressions [35]. There is now a growing interest in approaches based on machine learning and more specifically on deep learning that relies on data to automatize this process. These approaches are powerful and strongly improve the recognition rates [14], [36], [37], [38]. In computer graphics, machine learning approaches are able to produce realistic animations [39], [40], and to automatically transfer an expression from one gesture to another [41], [42], [43]. All these approaches tackle an emotion as a high-level information, mostly a label. In emotion recognition, the approaches find the label in real time with very little error. In computer animation, a user choose the label of an emotion and get a realistic result. Nevertheless, many applications would benefit from having more granularity than a simple label. The use of intermediate parameters allows a user to refine the results proposed by an algorithm to handle more subtle emotions. Finding such intermediate parameters remains an open question in the field of computer animation. Few studies in this field have sought to exploit expressive models based on broadly applicable and intelligible parameters [7], [20], [44].

In this study our goal is to propose a list of comprehensible and intelligible features responsible for the expression of emotion in human motion. Our contribution focusing on comprehensible and broadly applicable features could be exploited in multiple research areas: as input parameters to enhance existing user-controllable animation systems [39], [40], in machine-learning-based approaches to build better latent spaces from motion capture data [45], or in psychology to unify different parameters.

2.1 Related Surveys

During the last years, several surveys have listed works related to expression, emotions, styles in body posture, gestures, animations, often focusing on the recognition aspect. The first surveys dealing with emotion recognition do not address a specific modality [46], [47], [48]. They cover in the same article a wide spectrum of cues: voice, face, physiological signals, etc. Emotion recognition based on posture and body movements is only a small part of these surveys.

Kleinsmith et al. in [10] were the first to propose a review entirely dedicated to the perception and recognition of affective body expressions. They study the question of how human factors such as culture affect expression, perception and recognition. They also review psychological studies that have examined body configurations to assess whether specific features contribute to the recognition of specific affective states. Karg et al. in [49] extend the review by looking at the generation or synthesis of such motions, including expressions. Zacharatos et al. in [50] focus on the importance of movement segmentation, since several gestures and emotion may appear in a long movement. Fripp in [51] complete the review by integrating methods for recognizing expression in static posture. Larboulette et al. in [52] provide a review of computable expressive descriptors in human motion. They provide formulas that can be applied to 3D motion capture data.

Recently, Noroozi et al. have updated previous surveys [53]. They review the complete framework for automatic recognition of emotional body gestures, from the capture of static and dynamic pose to the learning process dedicated to the recognition task. They also discuss multi-modal approaches that combine speech or face with body gestures to improve emotion recognition. Ribet et al. extend the review further by studying the notion of style in the context of data representation, motion capture and recognition [38]. As define in [54], they focus on style by defining them with the notion of verbs and adverbs. They study methods working on parameterized motions: the style in motion is defined as

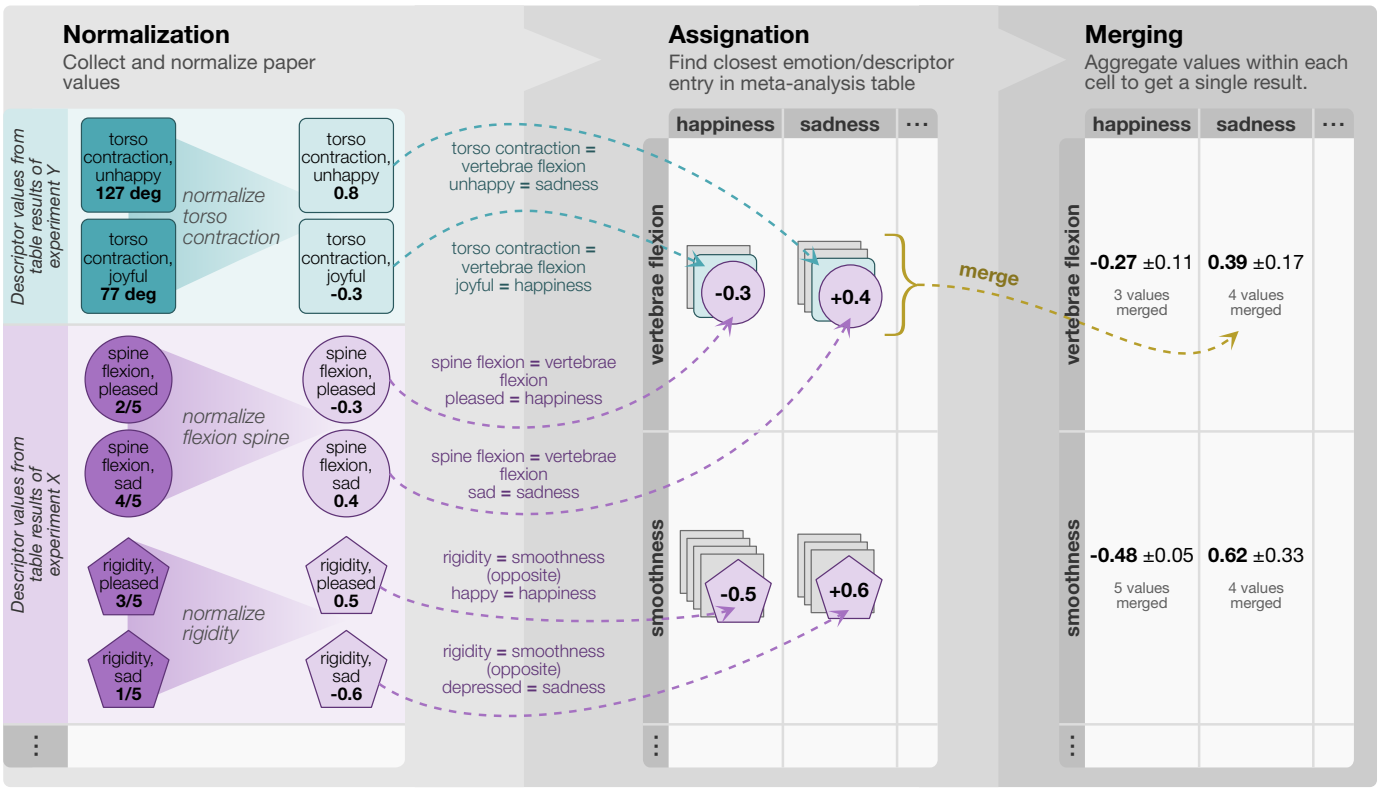


Fig. 1. Overview of the aggregation strategy employed for the meta-analysis and decomposed in three steps: normalization, assignation and merging. Numerical results above are only given as examples.

verbs being parameterized by adverbs. For example, walk was a verb and happy and angry were some of its adverbs. Lastly Deligianni et al. in [55] focus on the body expression recognition in the healthcare context. They review methodologies for gait analysis to detect mood disorders.

By matching results from two different methodologies, our work complements recent surveys on several aspects. First of all, we identify the comprehensible descriptors used in the research to differentiate the expression of emotions. To our knowledge, there is no prior work which attempted to build such a list in a systematic review. We also propose a quantification of these descriptors for a number of emotions. The originality of our approach lies in the fusion of different research papers into unified values through a meta-analysis. In addition to this quantitative approach we provide a way to extract from motion capture data a large part of the descriptors. Larboulette et al. proposed a similar review of descriptor computation. We add some necessary details to ensure perfect reproducibility of the computation and we use anthropomorphic tables to ensure that motion capture data is standardized for any skeleton topology. We finally present a methodology to validate the reliability of each of the emotionally unified values. This validation is obtained by comparing the value of the meta-analysis with those extracted from a motion capture database. All these steps taken together help to determine which comprehensible descriptors are most relevant, and for which emotions, with the aim of improving scientific understanding of bodily expression of emotion.

3 Parameters and Quantification Resulting from the Meta-Analysis

This section describes the systematic review done on previous works studying the expression of emotion through body movement in order to extract and unify all descriptors. The meta-analysis provides a list of expressive descriptors associated with a range of values for most common emotions. This section explains the methodology, the difficulties, and the resulting list of parameters and values.

An exploration of the state of the art in different research areas lead to identify studies providing comprehensible descriptor measures for each emotion. (Section 3.1). Despite the diversity of the collected studies, we can observe their results are often represented in tables. The most common representation uses rows for expressive factors and columns for emotions. Each line represents how a given expressive factor is modulated between emotions. We identified recurring expressive factors and emotions by comparing tables from different papers. Even if labels do not match, it is often possible to find correspondences using the textual description. This allows to obtain unified results from heterogeneous data in most cases. The *reference* lists provided in Sections 3.3 and 3.2 is the result of this process. Using such reference lists makes possible the aggregation of heterogeneous studies. The result takes the form of a big table, using rows as our reference expressive factors and columns as our reference emotions. It is subsequently named *meta-analysis table*.

An overview of our aggregation strategy is represented in Fig. 1. It can be decomposed into three main steps:

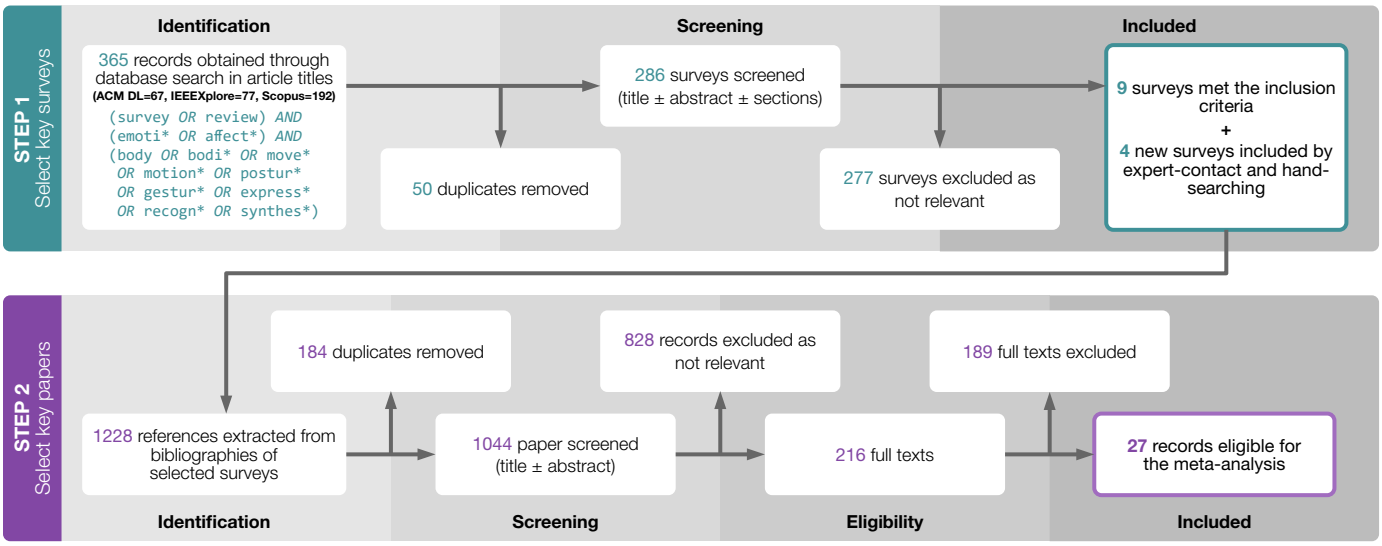


Fig. 2. Methodology used to identify relevant studies.

1) the *normalization* step ensures that all extracted values coming from different sources are expressed in the same range and units (Section 3.4);
2) the *assignation* step fills meta-analysis table with the normalized values (Section 3.5);
3) the *merging* step merges values within each meta-analysis cell (Section 3.6).

The last part of this section presents the quantified results of the computed meta-analysis (Section 3.7).

3.1 Identification of relevant studies

This meta-analysis focuses on the quantification of comprehensible descriptors responsible for the expression of emotions in the human body and movement. The methodology for identifying relevant studies is illustrated in Fig. 2. It is based on the PRISMA reporting guidelines [56].

For the first step, three databases were queried leading to 365 records with publications dating up to 2022. A survey met the inclusion criteria if it contained at least one section dealing with the expression of emotions in the human movement and/or posture, regardless of the field of application (recognition, synthesis, medical applications, etc.). The final process in step 1 led to identify 9 surveys covering our research field. 4 additional surveys have been included as a result of expert-contact and hand-searching. The publication dates of the included surveys ranges up to 2019.

For the second step, a number of 1228 references have been extracted from the bibliography sections of the relevant surveys. The screening process led to find 216 articles dealing with emotions in the movement and posture of a human body. The inclusion was based on the following criteria:

1) The article provided at least one descriptor extractable from human body movement and/or posture. This excluded descriptors based on physiological, neurological, face, voice or music signals, widely covered in the literature [57], [58], [59], [60].
2) At least one of the descriptors meeting criterion 1 could be easily understood by a non-expert and represented as a real number. This excluded state of the art methods based on complex mathematical transformations, or deep neural networks [61], [62], [63].
3) The article provided measures for at least one descriptor meeting criterion 2 and for at least two different emotions (not including neutral). In the context of this meta-analysis, an emotion is defined as a phenomena of short duration, intense and which can change quickly [49], [64], [65]. We therefore excluded results focusing on *moods* (phenomena of longer duration), *pain* (which can be considered as a separate emotional phenomenon [66]), and more generally affective phenomena related to the healthcare domain [4].

TABLE 1
Listing of surveys and studies included with the PRISMA procedure.

Step 1 : 13 surveys included

[67]* [48]* [46]* [49]* [10]* [50]* [68]* [53]* [69]* [47]† [15]† [51]† [55]†

Origin	No. of surveys
* Database search	9
† Expert contact and hand-searching	4

Step 2 : 27 studies included in meta analysis

[24][1,a] [70][1,a] [71][1,a] [72][1,b] [73][1,b] [74][1,b] [75][1,b] [76][1,b] [77][1,b] [78][1,b] [79][2,a] [80][2,a] [81][2,a] [82][2,a] [83][2,a] [84][2,a] [5][2,b] [85][3,b] [86][3,b] [87][3,b] [88][3,c] [89][3,c] [90][4,a] [91][4,a] [92][4,b] [93][5,a] [94][5,a]

Research fields	No. of studies
[1] Affective computing	10
[2] Psychology	7
[3] Kinesiology	5
[4] Arts	3
[5] Robotics	2
Measurement method	**No. of studies**
[a] Perceptual assessments	13
[b] Feature extraction from motion capture or video	12
[c] Both (perceptual + extraction)	2

The TABLE 1 lists the 27 articles included in the meta-analysis, with publications dating up to 2018. These articles differ both in their field of research and in the way the values are calculated. There is an almost equal proportion of studies based on the extraction of motion capture data and those based on perceptual assessments. The dominant research area is related to affective computing (emotion recognition, computer vision), followed by psychology, kinesiology (study of human movements), arts (dance and music performances), and robotics.

To conclude this section, it should be noted that we have chosen to extract articles from bibliographies of relevant surveys instead of making a direct systematic search of relevant papers. We acknowledge the limitation of the process as this may have left out some papers. A direct systematic search (i.e. the same query as in the identification process of step 1 in Fig. 2, without the boolean expression "survey OR review") leads to a very large amount of records (> 5000). Even if not all such studies may have been relevant to this paper, a more inclusive analysis may provide further insights in the subject.

3.2 Reference list of expressive descriptors

The reference list of expressive descriptors is one of the main contribution of this paper. Each of expressive descriptor in the list is made of a label, a description, a list of studies referencing it, and two opposite poles. Such a bipolar representation allows determining the level of influence of an expressive descriptor for a given emotion, relative to a neutral state, which is determined statistically (see Section 3.4). The expressive descriptors families are described in the following subsections.

3.2.1 Biomechanical features

Biomechanical features designate the lowest level primitives in human body that hold expressivity: these are the joints movements like neck flexion, shoulder abduction, etc. Among all the expressive descriptors, these are the most difficult to label. In our study set (see the "References" column in table of Fig. 3), many naming conventions can be observed such as Euler Axes, body parts orientation, etc. This leads to a lot of confusion, and sometimes hardly reproducible experiments.

In order to prevent such kind of reproductibility issue, we used knowledge from the field of functional anatomy [95], [96], [97] to design Fig. 3. It provides an illustration of all the joint rotations referenced in the meta-analysis, along with the studies that measure them.

In the rest of this study, the biomechanical features listed in Fig. 3 will sometimes be postfixed by the *amplitude* keyword in order to make the distinction between instantaneous measurements (e.g. elbow flexion) and delta measurements over a time window (e.g. elbow flexion *amplitude*).

3.2.2 High level features

In contrast to biomechanical features that are intended to describe body movements in a very factual way, higher level features allow to describe movements and postures in a more conceptual way. They may seem subjective, but they can actually be defined as a combination of lower level factors (see Section 4.1.3). Moreover, these features are scalable, in the sense that they can be used to describe the movements of a limb as well as a whole body.

In the first part of TABLE 2 are listed all gait features (walk stride length, walk speed, etc.), descriptive adjectives (fluidity, regularity, etc.) and movement properties (movement activity, movement amplitude, etc.) that we encountered in the meta-analysis study set.

The second part of TABLE 2 provides a listing of the main components of the Laban's effort and shape factors along with the studies that measure them. Rudolf Laban is a choreographer from the 20th century who developed notation systems to describe human movements. His work is used in many research fields such as psychology, robotics, computer animation or motion analysis. An important part of his work is dedicated to the research of factors allowing to communicate and to transcribe the expressive qualities of a movement. Laban groups these factors into two distinct categories:

- *effort* for dynamic factors (how movement progresses in a time sequence);
- *shape* for postural factors (how body form changes).

Effort is subdivided into space, time, weight and flow components. Shape can be related to effort by expressing it with the same components [98].

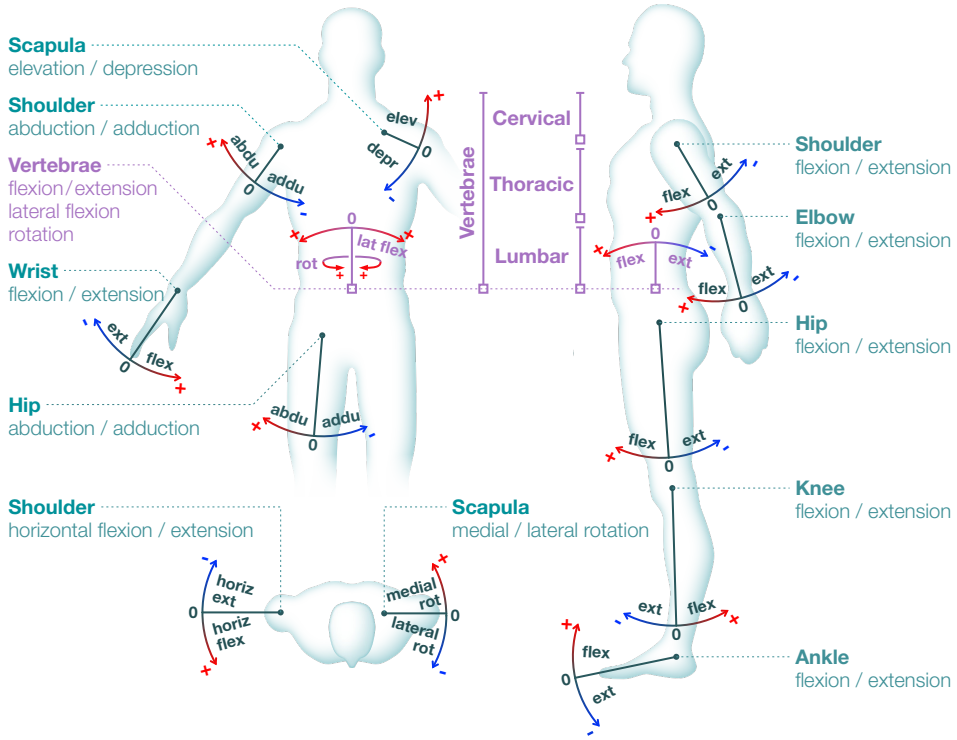
The definitions provided in TABLE 2 are based on Laban's books [32], [33] and writings from certified Laban Movement Analysts [98], [99], [100]. Although a substantial part of the studies listed in TABLE 2 do not refer to the Laban factors explicitly, their descriptions closely match. The Laban time effort descriptor refers to the speed/duration of a movement, which can be described in many ways, such as 'slow/fast' [84], 'velocity' [77], [92], etc. The Laban weight effort relates to the amount of energy injected into the movement and match many descriptions such as 'energy' [82], [89], 'powerful' [71], etc. The Laban space effort refers to the spatial expansiveness of a movement and is really close to the notions of 'directness' [81], 'spatial extension' [80], etc.

3.3 Reference list of emotions

Designing a reference set of emotions is not a trivial task. A small set would aggregate too many different labels into the same bin, resulting in a lack of affect diversity. On the other hand, using a larger set decreases aggregates, leading to fuzzier and less reliable values.

Many researchers have attempted to define an universal set of perceived emotions [101] and there is now scientific consensus on at least six of them: *happiness*, *sadness*, *disgusts*, *fear*, *anger*, or *surprise*. A lot of considered studies are referencing these "basic" emotions. Though, there are also references to more subtle affective states such as *shame* [24], [80], [81], *boredom* [71], [80] or *anxiety* [24], [71], [89], [102].

In order to classify a wide range of emotion types whilst keeping a compact representation, we used a lookup table of 38 affect categories proposed by Scherer [64], and presented in TABLE 3. A category is made of a central label and stems referring to synonyms of that label. A stem can occur in multiple categories, as it may have different meaning depending on the context.



Name*	References
ankle flexion	[88]
cervical vert. flex.	[88][80][79][76][5][74]
cervical vert. lat. flex.	[80][74]
cervical vert. rot.	[74]
elbow flexion	[88][87][76]
hip flexion	[88][87][76]
knee flexion	[88][87][76]
scapula elevation	[88][80]
scapula medial rot.	[87][80][75]
shoulder abduction	[75]
shoulder flex.	[88][87][76]
shoulder horiz. flex.	[75]
thoracic vert. flex.	[88]
vert. flex.	[88][75][79][76][83]
vert. lat. flex.	[88]
vert. rot.	[88]
wrist flexion	[88]

*Notes:
vert. stands for *vertebrae*,
flex. stands for *flexion*,
lat. stands for *lateral* ,
rot. stands for *rotation*,
horiz. stands for *horizontal*

Fig. 3. Biomechanical features extracted from meta analysis: schematic and table representations are presented side by side.

Our modifications and additions to Scherer's original work are motivated by the following points:

- Some studies include *neutral* performances. The new neutral category catches these kinds of specific entries.
- Usage of *joy* is often ambiguous, as it can refer to happiness as well as elation/euphoria. We renamed the Scherer's *joy* category to the less ambiguous word *elation* to prevent any misinterpretation, and we placed the *joy** stem in the *elation* and *happiness* categories.
- All other bolded stems are not referenced in Scherer's original work, but do appear in the meta-analysis study set. For each of these stems, we chose the closer affect categories simply using the Merriam-Webster English dictionary.

3.4 Normalizing results from heterogeneous studies

Distinct studies may have made reference to the same expressive factor using completely different units: studies based on perceptual assessments mostly use n-item Likert scales, whereas computational studies provide measurable data, such as angles or speed.

As mentioned above, every study result from our meta-analysis set (listed in TABLE 1) can be represented in the following table form:

$$\begin{array}{c|ccc} & e_1 & \dots & e_j \\ \hline f_1 & x_{1,1} & \dots & x_{1,j} \\ \vdots & \vdots & \ddots & \vdots \\ f_i & x_{i,1} & \dots & x_{i,j} \end{array} \quad (1)$$

where each cell $x_{i,j} \in \mathbb{R}$ contains the value of an expressive feature labelled f_i when performing emotion labelled e_j. These labels are the ones used in the original paper.

We define as $\mathbf{\Omega}$ the set of non-significant values. A *non-significant* $x_{i,j}$ expresses that feature f_i is unnoticeable regarding emotion e_j. These values are either those labelled as *non-significant* ('ns') in the source paper such as some values in [81], either those which are blank cells such as some values in [76].

Any other computation employed to determine the significance of specific cells, such as ANOVA statistical method, have been ignored. We believe that evicting some values is only relevant on their own datasets, whereas this study aims to aggregate as much value as possible without any assumption. Therefore, we take all raw results of all cells, provided that a value is explicitly specified.

Some studies are using *unsigned* features that are part of a more general *signed* feature. For example, [80] is using *Shoulders up* which is the positive part of the *signed* feature *Shoulder Elevation*. We denote $\mathbf{U}$ this set of *unsigned* features.

For any input set of values $\mathbf{V}$, its cardinality is denoted $\mathrm{card}(\mathbf{V})$. The arithmetic mean, standard deviation, root-mean-square and relative standard deviation (also known as coefficient of variation) of $\mathbf{V}$ are respectively defined as:

TABLE 2
High level features extracted from meta analysis

Features based on the effort and shape factors of Laban				
Name	Description	Negative pole	Positive pole	References
laban flow effort	Degree of control, resistance or introversion during movement progression	Bound, tense movement progression, feeling of control.	Free, loose movement progression, feeling of relaxation.	[88][84][89][90]
laban flow shape	Amount of self-centeredness in body posture configuration	Closing, shrinking, contracting, orienting inward	Opening, growing, expanding, orienting outward	[88][89][79][72]
laban space effort	Level of expansion or contraction during movement performance	Direct, straight line direction, tapered movement progression, feeling of narrowness	Indirect, undulating line direction, flexible movement progression, feeling of large space	[88][80][84][92][81][71][82][89][90][72][93][94]
laban space shape	Occupation of limb on transversal plane	Enclosing, narrowing, decreasing transversal occupation	Spreading, widening, increasing transversal occupation	[88][24][89][93][94][81]
laban time effort	Time impression during movement performance	Sudden, high speed, short time period, ephemeral feeling	Sustained, low speed, long time period, sensation of eternity	[88][84][92][91][81][71][82][77][24][89][93][85][94][74][73]
laban time shape	Orientation of body bosture on lateral plane	Retreating, hallowing, orienting backward	Advancing, bulging, orienting forward	[81][24][72][93][94]
laban weight effort	Quantity of force, energy and gravity perceived during movement performance	Strong, sensation of weight, high gravity resistance, energy liberation	Light, feeling of weightlessness, low gravity resitance	[88][80][84][92][81][71][82][24][89][90][72][93][85][94][83]
laban weight shape	Configuration of body posture on vertical axis	Sinking, shortening, orienting downard	Rising, lengthening, orienting upward	[80][81][24][93][94][79]
Features based on gait properties, descriptive adjectives and general movement properties				
Name	Description	Negative pole	Positive pole	References
approach	Overwhole body orientation toward an object or situation	Avoiding, turning away	Approching, turning toward	[79][70]
exaggeration	Amplification of timing/speed/amplitude of movement	Unexagerrated (tends to neutral movement)	Highly exagerrated and stereotyped	[90]
movement activity	Overall impression of movement quantity, considering both amplitude and frequency	Tends to immobility	A lot of movements	[80][84][82][72][78][24]
movement amp.	Spatial and/or angular range of movement primitive	Small amplitude	Large amplitude	[91][89]
movement anticipation	Preparation time before action execution. Anticipation usually consists in movement in the opposite direction of the main action	Lower anticipation time	Higher anticipation time	[89]
movement exertion	Needed time for main action progression. Exertion is easily recognizable in a jump movement as the airborne phase	Lower exertion time	Higher exertion time	[89]
movement settle	Stabilization phase right after exertion. Preceeds rest pose or new movement execution	Lower settle time	Higher settle time.	[89]
regularity	Amount of changes in phrasing, tempo and speed during overall performance	Irregular	Regular	[91][24][72]
smoothness	General impression of fluency during movement performance	Jerky and with a lot of stop and go	Smooth/fluent movements	[84][91][24][90][72][78][70]
walk arm swing	Amplitude of front/back arms sway during walk	Higher swing amplitude	Lower swing	[5][83]
walk lat. body sway	Lateral sway of upper body during walk, left and right movements are undifferentiated	Reduced sway, impression of rigid walk	Exagerated sway, a feeling of "drunk" walk	[5]
walk speed	Translationnal velocity	Lower velocity	Higher velocity	[88][86][5]
walk step frequency	Number of footstep during a specific time period	Lower frequency	Higher frequency	[88][86]
walk stride duration	Duration between two consecutive footsteps	Lower duration	Higher duration	[86]
walk stride length	Distance between two consecutive footsteps	Lower stride length	Higher stride length	[88][86][83]
walk vertical head sway	Up-Down head sway during walk	Lower vertical sway	Higher vertical sway	[5]

$$\bar{\mathbf{V}} = \frac{1}{\mathrm{card}(\mathbf{V})} \sum_{v \in \mathbf{V}} v \tag{2}$$

$$\sigma(\mathbf{V}) = \sqrt{\frac{1}{\mathrm{card}(\mathbf{V})} \sum_{v \in \mathbf{V}} (v - \bar{\mathbf{V}})^2} \tag{3}$$

$$\mathrm{rms}(\mathbf{V}) = \sqrt{\frac{1}{\mathrm{card}(\mathbf{V})} \sum_{v \in \mathbf{V}} v^2} \tag{4}$$

$$\mathrm{rsd}(\mathbf{V}) = \frac{\sigma(\mathbf{V})}{|\bar{\mathbf{V}}|} \tag{5}$$

Let denote $\mathbf{X}_i = \{x_{i,1}, \ldots, x_{i,j}\}$ the value set of the i^{th} expressive feature for each emotion. We can define the set of significant values as follows:

$$\mathbf{S}_i = \{x \in \mathbf{X}_i \mid x \notin \mathbf{\Omega}\} \tag{6}$$

We define a *normalization* function which expresses each value relative to the mean $\bar{\mathbf{S}}_i$ and then scale it using $\sigma(\mathbf{S}_i)$. it is defined as follows:

$$\eta_i(x) = \begin{cases} \mathrm{ignored}, & \text{if } x \in \mathbf{\Omega} \text{ or } \sigma(\mathbf{S}_i) = 0 \text{ or } \mathrm{rms}(\mathbf{S}_i) = 0 \\ \dfrac{x}{\mathrm{rms}(\mathbf{S}_i)}, & \text{if } x \in \mathbf{U} \\ \dfrac{x - \bar{\mathbf{S}}_i}{\sigma(\mathbf{S}_i)}, & \text{otherwise} \end{cases} \tag{7}$$

Each raw table (1) is normalized by applying $\eta_i(x)$ on each cell. If $\eta_i(x) = \mathrm{ignored}$, the cell value is ignored for

TABLE 3
Affect categories and associated stems suggested by Scherer [64]. A stem may be followed by a *, suggesting all possible variations (e.g. *joy** gives *joyfulness*, *joyous*, etc.). Our changes are framed

Category Name	Matching Patterns
admiration/awe	admir*, ador*, awe*, dazed, dazzl*, enrapt*, enthrall*, fascina*, marveli*, rapt*, reveren*, spellbound, wonder*, worship*
amusement	amus*, fun*, humor*, laugh*, play*, rollick*, smil*
anger	anger, angr*, cross*, enrag*, furious, fury, incens*, infuriat*, irate, ire*, mad*, rag*, resent*, temper, wrath*, wrought*, hot anger, cold anger, aggressiv*, attack*
anxiety	anguish*, anxi*, apprehens*, diffiden*, jitter*, nervous*, trepida*, wari*, wary, worried*, worry*
being touched	affect*, mov*, touch*
boredom	bor*, ennui, indifferen*, languor*, tedi*, wear*, tired*, exhaust*
compassion	commiser*, compass*, empath*, pit*, sympath*
contempt	contempt*, denigr*, deprec*, deris*, despi*, disdain*, scorn*
contentment	comfortabl*, content*, satisf*
desperation	deject*, desolat*, despair*, desperat*, despond*, disconsolat*, hopeless*, inconsol*, helpless*
disappointment	comedown, disappoint*, discontent*, disenchant*, disgruntl*, disillusion*, frustrat*, jilt*, letdown, resign*, sour*, thwart*, defeat*
disgust	abhor*, avers*, detest*, disgust*, dislik*, disrelish, distast*, loath*, nause*, queas*, repugn*, repuls*, revolt*, sicken*, antipath*
dissatisfaction	dissatisf*, unhapp*
envy	envious*, envy*
fear	afraid*, aghast*, alarm*, dread*, fear*, fright*, horr*, panic*, scare*, terror*
feeling	love, affection*, fond*, love*, friend*, tender*
gratitude	grat*, thank*
guilt	blame*, contriti*, guilt*, remorse*, repent*
happiness	cheer*, bliss*, delect*, delight*, enchant*, enjoy*, felicit*, happ*, merr*, joy*
hatred	acrimon*, hat*, rancor*
hope	buoyan*, confident*, faith*, hop*, optim*
humility	devout*, humility, inferior*
interest/enthusiasm	absor*, alert, animat*, ardor*, attenti*, curi*, eager*, enrapt*, engross*, enthusias*, ferv*, interes*, zeal*, alert*, concentrat*, excit*
irritation	annoy*, exasperat*, grump*, indign*, irrita*, sullen*, vex*, antipath*, defiance
jealousy	covetous*, jealous*
elation (joy)	ecstat*, elat*, euphor*, exalt*, exhilar*, exult*, flush*, glee*, joy*, jubil*, overjoyed, ravish*, rejoic*, triumph*
longing	crav*, daydream*, desir*, fanta*, hanker*, hark*, homesick*, long*, nostalg*, pin*, regret*, wish*, wistf*, yearn*
lust	carnal, lust*, climax, ecsta*, orgas*, sensu*, sexual*
pleasure/enjoyment	enjoy*, delight*, glow*, pleas*, thrill*, zest*
pride	pride*, proud*, conceit*
relaxation	ease*, calm*, carefree, casual, detach*, dispassion*, equanim*, eventemper*, laid-back, peace*, placid*, poise*, relax*, seren*, tranquil*, unruffl*
relief	relie*
sadness	chagrin*, deject*, dole*, gloom*, glum*, grie*, hopeles*, melancho*, mourn*, sad*, sorrow*, tear*, weep*, depress*
shame	abash*, asham*, crush*, disgrace*, embarras*, humilia*, shame*, inferior*
surprise	amaze*, astonish*, dumbfound*, startl*, stunn*, surpris*, aback, thunderstruck, wonder*
tension/stress	activ*, agit*, discomfort*, distress*, strain*, stress*, tense*
positive	agree*, excellent, fair, fine, good, nice, positiv*
negative	bad, disagree*, lousy, negativ*, unpleas*, negati*, reject*, denial*
neutral	neutral*

all subsequent steps of the meta-analysis. The set of N normalized tables obtained from the meta-analysis study set is denoted $\mathbf{T}_{norm} = \{\boldsymbol{t}_1, \boldsymbol{t}_2, \ldots, \boldsymbol{t}_N\}$

It is worth noticing that $\eta_i(x)$ uses $\bar{\mathbf{S}}_i$ as a reference point. Carefully choosing this point is crucial for the merging step. Most of the studies in the meta-analysis set introduce a neutral emotion for comparison purpose, and some authors use it as a reference point [76], [87], [92]. We argue that values coming from neutral labels are often biased: this state has no clear definition, and can be influenced by many factors such as the performer mood, personality or morphology.

On the other hand, arithmetic mean is highly dependent on the chosen emotion set. For instance, computing mean on the expressive feature set $\{x_{happy}, x_{sad}, x_{angry}\}$ probably won't give the same result as $\{x_{sad}, x_{shameful}, x_{bored}\}$. We observed that such extreme cases never occur in practice because studies use sparse and distinct emotions such as $\{x_{happy}, x_{sad}\}$ [5] or $\{x_{angry}, x_{fear}, x_{happy}, x_{sad}\}$ [76], leading to a stable mean.

Though we seem to find a satisfying reference point for our meta-analysis, there is no doubt that research on affective computing would benefit from a deeper, dedicated study on this topic.

3.5 Filling the meta-analysis table

Once our studies are converted into normalized tables, we have to put their values into their corresponding row/column entries in the meta-analysis table $\tilde{\mathbf{M}}$, using our reference lists of emotions and features. Let denote $\tilde{\mathbf{E}} = \{\tilde{e}_1, \ldots, \tilde{e}_{39}\}$ the reference emotion list (first column in TABLE 3) and $\tilde{\mathbf{F}} = \{\tilde{f}_1, \ldots, \tilde{f}_F\}$ the reference feature list (in Fig. 3 and TABLE 2).

For each value $x_{i,j}$ in a given normalized table $\boldsymbol{t}_k$, we identify a subset of reference emotions in $\tilde{\mathbf{E}}$ and a reference expressive feature in $\tilde{\mathbf{F}}$ that respectively match the labels e_j and f_i of the original paper. This gives the corresponding cells to update in the meta-analysis table. The whole procedure is expressed by the following algorithm:

Algorithm 1 Assignation procedure

Input set of normalized tables $\mathbf{T}_{norm}$, set of reference emotions $\tilde{\mathbf{E}}$ and set reference expressive features $\tilde{\mathbf{F}}$
Output meta-analysis table $\tilde{\mathbf{M}}$

$\tilde{\mathbf{M}} = \text{createEmptyMetaAnalysisTable}(\tilde{\mathbf{F}},\ \tilde{\mathbf{E}})$
for each $\boldsymbol{t}_k$ in $\mathbf{T}_{norm}$ **do**
  for each $feature$ in $\boldsymbol{t}_k$.getFeatures() **do**
    for each $emotion$ in $\boldsymbol{t}_k$.getEmotions() **do**
      $\tilde{\mathbf{E}}_{match} = \text{matchEmotionSet}(emotion, \tilde{\mathbf{E}})$
      $\{\tilde{f},\ isOpposite\} = \text{matchFeature}(feature, \tilde{\mathbf{F}})$
      $x_{norm} = \boldsymbol{t}_k.\text{getCell}(feature,\ emotion)$
      if $x_{norm} \neq$ ignored **then**
        if $isOpposite$ **then**
          $x_{norm} = -x_{norm}$
        for each $\tilde{e}_{match}$ in $\tilde{\mathbf{E}}_{match}$ **do**
          $\tilde{\mathbf{M}}.\text{appendValueToCell}(x_{norm},\ \tilde{f},\ \tilde{e}_{match})$
return $\tilde{\mathbf{M}}$

3.5.1 Emotion label matching

The function matchEmotionSet used in Algorithm 1 return a subset $\tilde{E}_{match}$ from any input original label e with the following algorithm:

Algorithm 2 *matchEmotionSet*

Input emotion label $emotion$, set of reference emotions $\tilde{\mathbf{E}}$
Output set of matching reference emotions $\tilde{\mathbf{E}}_{match}$

$\tilde{\mathbf{E}}_{match} = \{\emptyset\}$
for each $\tilde{e}$ in $\tilde{\mathbf{E}}$ **do**
  for each $stem$ in $\tilde{e}$.getStems() **do**
    if $\text{matchStem}(stem, emotion)$ **then**
      $\tilde{\mathbf{E}}_{match}.\text{appendValue}(\tilde{e})$
      break
return $\tilde{\mathbf{E}}_{match}$

Where $\text{matchStem}(stem,\ emotion)$ is a wildcard string matching function (e.g. *joyfulness* matches *joy**, *fearful* matches *fear**, etc.).

The Algorithm 2 only returns the empty set if the input label does not match any of our affect categories. In these special cases, values are simply ignored and discarded from the meta-analysis table. This occurs for some labels in [79] and [71] for which we did not find any satisfying affective category in Scherer table (such as scrutiny, puzzlement, etc.).

3.5.2 Expressive descriptor matching

Unlike emotions, the expressive descriptor matching called matchFeature in Algorithm 1 is not an automatically performed step. It has been done during the result transcription of our study set. Each of the expressive parameters labels (i.e. each f_i described in Section 3.4) have been manually associated with one reference parameter from $\tilde{\mathbf{F}}$, by carfully reading the definition provided in the original paper. This association may be subject to a sign inversion ($isOpposite$ property in Algorithm 1). For example, some authors may use *extension* angles as positive pole, while our reference list favor *flexion* angles.

In cases where the definition of the parameter provided in the paper seemed unintuitive or difficult to associate with one of our reference parameters in $\tilde{\mathbf{F}}$, we have chosen to ignore the values. This is for example the case for the notions of *periodicity* [72], *motion of the barycentre* [72], *excursion* [78], *frequency* [89], *pleasant* [82], *illustrator* [80] etc.

3.6 Merging values inside each meta-analysis cell

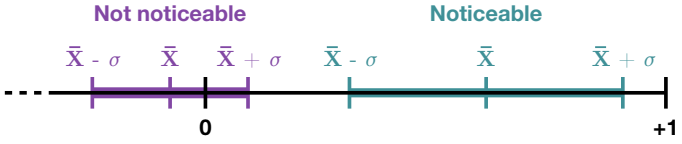
After the assignation step described in the previous section, we obtain a meta-analysis table $\tilde{\mathbf{M}}$. Each cell of $\tilde{\mathbf{M}}$ contains a set of values $\mathbf{X}$ coming from various studies but referring to the same feature and emotion. That value set need to be merged to obtain an exploitable scalar representation. For each cell $\mathbf{X}$, we compute its arithmetic mean $\bar{\mathbf{X}}$ and its standard deviation $\sigma(\mathbf{X})$.

3.7 Meta-analysis results

For all cells in the meta-analysis table $\tilde{\mathbf{M}}$, we propose a method to identify the most noticeable ones. The proposed method tries to encode both the level of agreement between papers and the ability to represent the emotion. The noticeability of a given cell $\mathbf{X}$, is determined using the opposite notion:

$$\text{unnoticeable}(\mathbf{X}) = \begin{cases} \infty, & \text{if } \bar{\mathbf{X}} = 0 \\ 1, & \text{if } \bar{\mathbf{X}} \neq 0 \text{ and } \text{card}(\mathbf{X}) = 1 \\ \text{rsd}(\mathbf{X}), & \text{otherwise} \end{cases} \quad (8)$$

A value is not relevant to represent the emotion if $\bar{X}$ is close to zero. On the other hand, a cell is considered noticeable if $\text{unnoticeable}(\mathbf{X})$ is strictly less than one. This occurs when at least two measures coming from independent experiments have been merged, and the majority of measures are sharing the same sign. This does not necessarily mean that cells with only one measure ($\text{unnoticeable}(\mathbf{X}) = 1$) should not be trusted, since the original experiment may have found a significant difference between emotions for this single value.



In order to have a more direct view of relevant results, we also propose a method to sort emotions and descriptors by level of noticeability. The noticeability level of a given descriptor (row) or emotion (column) in the meta-analysis table is computed using only the noticeable cells:

$$\text{noticeableLevel} = \sum 1 - \text{unnoticeable}(\mathbf{X}_{noticeable}) \quad (9)$$

For a given descriptor, a high level of noticability can be interpreted as a high level of agreement on its ability to characterize different emotions. For a given emotion, a high level of noticability indicates a high level of agreement on its ability to be expressed through different descriptors. The meta-analysis results are presented in TABLE 4. Emotions without noticeable cells have been discarded from the table for layout issues (irritation, humility, amusement, etc.). In the following subsections, $\oplus$ will refer to positivity and $\ominus$ to negativity of the associated value.

TABLE 4

Meta-analysis results: within each cell, the first value stands for the mean and the value inside parenthesis stands for standard deviation. None parenthesis indicates that the cell is the result of a single value. Noticeable cells are underlined ($\mathrm{unnoticeable}(\mathbf{X})$ strictly less than one). A cell is empty if none study have provided a numerical value. Each descriptor and emotion label is followed by a number inside square brackets which stands for the noticeability level.

	Sadness [14.3]	Anger [12.3]	Happiness [10.0]	Elation [7.3]	Neutral [5.2]	Shame [3.5]	Pride [2.2]	Contentment [2.1]	Disgust [2.1]	Fear [1.9]	Desper. [1.7]	Contempt [1.5]	Surprise [1.5]	Boredom [1.2]	Relax. [0.9]	Anx. [0.7]
walk speed [4.7]	-1.3 (0.20)	+1.1 (0.09)	+1.0 (0.01)	+1.0 (0.02)	-0.7 (0.02)			-0.0								
laban time effort [4.6]	+1.3 (0.29)	-1.2 (0.43)	-0.5 (0.59)	-0.5 (0.61)	+0.3 (0.42)	+1.2 (0.10)	-0.5 (0.54)	+0.7 (0.41)	-0.5 (0.46)	-0.3 (0.71)		-1.1	-0.1 (0.28)	+0.8 (0.14)	+0.6 (0.06)	+0.4 (0.38)
walk step frequency [4.5]	-1.4 (0.02)	+1.1 (0.04)	+0.9 (0.05)	+0.9 (0.05)	-0.8 (0.23)			+0.2								
cervical vert. flex. [3.6]	+1.1 (0.48)	+0.3 (0.37)	-0.9 (0.11)	-1.0 (0.22)	-0.6	+0.9 (0.46)	-0.9 (0.91)	-0.8	+0.8 (1.12)	-0.2 (0.49)	+0.7 (0.10)	-1.6		-0.7 (1.08)		
vertebrae flexion [3.5]	+0.9 (0.31)	+0.6 (0.35)	-1.0 (0.66)	-0.9 (0.31)	-0.6	+0.7 (0.00)	-1.5 (0.00)	-0.6		+0.8	+0.7			+0.7		
shoulder flex. amp. [3.4]	-1.0 (0.24)	+1.3 (0.15)	+0.9 (0.00)	+0.9	-0.8			+0.1		-0.7 (0.16)						
laban weight shape torso [3.4]	-1.3 (0.31)	-0.5	+1.0	+0.4 (0.61)		-1.5 (0.36)			-0.9 (0.13)	-0.4 (0.25)	-0.2	-0.5 (0.16)	+1.2	-1.1		
laban flow shape torso [3.3]	-1.5 (0.39)	+0.5 (0.09)	+1.1 (0.28)	+1.1 (0.28)	-0.1		+0.5	+0.5 (0.33)								-1.5
walk stride length [3.3]	-1.5 (0.14)	+1.0 (0.09)	+0.6 (0.65)	+1.0 (0.04)	-0.4 (0.21)		+0.8	-0.2								
laban weight effort [3.3]	+1.1 (0.38)	-1.1 (0.77)	-0.4 (0.81)	-0.5 (0.63)	+0.6 (0.17)	+0.9 (0.40)	-0.4 (0.62)	+0.5 (0.69)	+0.2 (0.72)	-0.6 (0.50)	-0.6	+0.0 (0.70)	-0.2 (0.42)	+0.6 (0.36)	+0.5	+0.6 (0.42)
movement activity [3.1]	-1.2 (0.32)	+0.5 (0.68)	+0.0 (0.54)	+1.1 (0.87)	-1.3	+0.3	+0.6 (0.89)		-0.7 (0.42)	+0.5 (0.39)	+0.5 (0.08)	-2.1	+1.2 (0.31)	-0.7		-1.0
laban space shape [2.9]	-0.9 (1.11)	+0.7 (0.07)	+0.7 (0.86)	+0.7 (0.86)	-0.3 (0.44)	-1.5	+0.9 (0.24)	+0.4 (0.01)		-0.2			-1.1		-0.1	-0.8 (0.56)
laban weight shape [2.8]	-1.4 (0.20)	-0.2 (0.32)	+1.0 (0.37)	+1.0 (0.37)	+1.1	-0.9 (0.25)	+1.5		-1.0	-0.5		-1.1	+1.1		+0.9	-0.2
elbow flexion amplitude [2.6]	-1.3 (0.12)	+1.0 (0.19)	+0.7 (0.09)	+0.8	-1.0			+0.4		-0.8						
laban time shape [2.5]	+0.4 (0.97)	+0.6 (0.76)	-0.4 (0.15)	-0.4 (0.19)	-1.3	+0.2 (0.39)	-1.7		-0.4 (0.13)	+0.5 (1.38)			-0.8 (0.18)		-1.0	+0.8
hip flexion amplitude [2.3]	-1.1 (0.06)	+1.1 (0.25)	+0.9 (0.35)	+1.2	-1.2			+0.2		-0.9						
knee flexion amplitude [2.2]	-0.9 (0.15)	+1.4 (0.20)	+0.5 (0.24)	+0.3	+0.3			-1.1		-1.2						
laban time effort head [2.1]	+1.1 (0.57)	-1.2 (0.15)	-0.4 (0.12)		+1.3					-0.4						
laban flow effort [2.0]	-0.0 (0.56)	-1.5 (0.14)	+0.5 (0.63)	+0.1 (0.69)	+1.1 (0.28)		-0.7	+1.0 (0.61)								-0.7
laban space effort [2.0]	-0.1 (1.05)	+0.1 (1.20)	+0.1 (1.10)	+0.5 (1.07)	+0.9 (0.27)	-0.7	-0.3 (0.15)	+0.1 (0.47)	-0.2 (0.49)	+0.4 (0.68)	+0.6	-0.9 (0.16)	+0.3 (0.70)	-0.4 (0.58)	-1.3	-0.9 (1.33)
approach [1.6]	-1.2 (0.28)	+0.4 (0.40)	+1.2 (0.69)	+1.2 (0.69)		-1.5				-0.2 (0.54)		-1.5	+0.7		+0.7	
movement amplitude [1.5]	-0.9 (0.62)	+1.2 (0.45)	+0.5 (0.46)	-0.2	-0.4		+0.7	-0.0		-1.0 (0.55)						-1.0
walk arm swing [1.5]	-1.4 (0.36)	+0.8	+0.8 (0.22)				+0.4									
smoothness [0.8]	+0.9 (1.02)	-0.8 (0.92)	-0.2 (0.36)	-0.1 (0.22)	+1.1 (0.23)	+0.1	+0.6 (0.71)		+0.8	-0.6 (0.64)	-1.5		+1.0			+0.5 (0.81)
scapula elevation [0.7]	+0.3	+1.2 (0.35)	-0.2 (0.73)	+0.4 (1.36)	+1.0	+0.7	+0.9	-1.4	+0.4	+0.9 (0.19)	+0.5	+0.7				
scapula medial rotation [0.6]		-0.3		-0.7 (0.29)		-0.2	-0.5		-1.7	-1.1 (0.36)	-1.5					
laban flow shape [0.1]	-0.6 (0.57)	+0.3 (0.45)	+1.1 (0.63)			-0.8 (0.00)	+1.3 (0.00)		-0.4 (0.71)	-1.2 (0.79)	-0.8	+1.3	+0.7 (0.12)	-0.8		
regularity [0.0]	+0.6 (0.87)	-0.7 (0.86)	-0.1 (0.46)	+0.1	+1.7	-0.4	+1.3		-1.0	-0.3 (0.86)			+1.7			-0.4
cervical vert. lat. flex. [0.0]	+0.4	+1.1 (0.61)				+1.3	+0.9		+0.4	+0.4				+1.7		
laban space shape arm [0.0]	-0.3		+1.4	+1.4		-0.4			-0.5 (0.06)	-1.5		-1.2				
movement activity arm [0.0]	-1.6	+1.2	+1.3	+1.3	-0.8	-1.2	+0.3			+0.6						+0.1
shoulder horiz. flex. [0.0]				-1.2 (0.15)												
movement anticipation [0.0]	+2.2	+0.7	-0.9	-0.9	-0.2		-0.4	-0.7								-0.7
movement exertion [0.0]	-2.0	+0.8	+0.4	+0.4	+0.7		+0.4	-1.0								+0.7
movement settle [0.0]	+1.0	-1.5	+0.2	+0.2	-0.8		-0.2	+1.7								-0.5
ankle flexion amplitude [0.0]	-1.6	-0.6	+0.4	+0.4	+1.4			+0.4								
thoracic vert. flex. [0.0]	+1.0	+0.0	-1.1	-1.1	+1.2			-1.1								
vert. lat. flex. amp. [0.0]	-1.6	+1.1	+1.1	+1.1	-0.3			-0.3								
vert. rot. amp. [0.0]	-1.4	+1.0	+1.0	+1.0	-1.0			+0.5								
wrist flexion amplitude [0.0]	-0.7	+0.2	+1.9	+1.9	-0.7			-0.7								
cervical vert. flex. amp. [0.0]	+0.7	+1.0	-0.1		-1.6											
cervical vert. lat. flex. amp. [0.0]	+0.3	+1.0	+0.3		-1.7											
cervical vert. rot. amp. [0.0]	-0.2	+0.6	+1.1		-1.5											
elbow flexion [0.0]	-1.4	+0.7	-0.5							+1.2						
hip flexion [0.0]	-0.6	+0.4	-1.2							+1.4						
knee flexion [0.0]	-1.0	-0.1	-0.6							+1.7						
laban time effort torso [0.0]	+1.5	-1.2	-0.5							+0.2						
laban weight effort leg [0.0]	+0.6	-1.7	+0.8				+0.3									
laban weight shape leg [0.0]		-1.0								-1.0	+1.0			+1.0		
movement amplitude head [0.0]	-1.2	-0.2	+1.6							-0.2						
movement amplitude torso [0.0]	-1.0	+0.9	+1.1							-1.0						
regularity head [0.0]	+1.5	-0.9	+0.4							-1.0						
regularity torso [0.0]	+1.6	-1.1	-0.1							-0.5						
shoulder abduction [0.0]				+1.3												
shoulder flexion [0.0]	+1.1	+0.4	+0.1							-1.6						
smoothness head [0.0]	+1.5	-1.1	+0.1							-0.5						
smoothness torso [0.0]	+1.6	-1.2	-0.3							-0.1						
walk stride duration [0.0]	+1.5	-0.9	-0.9	-0.9	+0.3											
exaggeration [0.0]	-1.0		+1.0													
walk lateral body sway [0.0]	+1.0		-1.0													
walk vertical head sway [0.0]	-1.0		+1.0													

3.7.1 Noticeable emotions

The underlined cells of a given emotion (column) highlight its noticeable descriptors. For example, according to the meta-analysis results, the expression of pride is characterized by three noticeable factors: extension of the vertebral column (*vertebrae flexion* $\ominus$), limbs taking up a lot of space on the transversal plane (*Laban space shape* $\oplus$) and large and flexible movements (*Laban space effort* $\oplus$). The three emotions with the highest noticeability level are sadness, anger and happiness. The meta-analysis results tends to prove that the bodily expression of these 'basic' emotions is well covered and understood by the literature.

Sadness is characterized by decreased movement speed (*walk speed* $\ominus$, *Laban time effort* $\oplus$), decreased movement energy (*Laban weight effort* $\oplus$, *movement activity* $\ominus$), decreased movement amplitude (*shoulder flexion amplitude* $\ominus$, *movement amplitude* $\ominus$) and contracted posture (*vertebrae flexion* $\oplus$, *Laban weight shape* $\ominus$).

Anger is characterized by increased movement speed (*walk speed* $\oplus$, *Laban time effort* $\ominus$), increased movement energy (*Laban weight effort* $\oplus$), increased movement amplitude (*shoulder flexion amplitude* $\ominus$, *movement amplitude* $\ominus$), tense movements (*Laban effort flow* $\ominus$), contracted spine (*vertebrae flexion* $\oplus$) and limbs taking up a lot of space on the transversal plane (*Laban space shape* $\oplus$).

Happiness is characterized by increased movement speed (*walk speed* $\oplus$, *Laban time effort* $\ominus$), increased movement amplitude (*shoulder flexion amplitude* $\ominus$, *movement amplitude* $\ominus$) and expanded spine (*vertebrae flexion* $\ominus$, *Laban weight shape* $\oplus$).

The 'neutral' column has seven noticeable cells with a mean far from 0, though this central value should represent a neutral state. This reinforces the idea suggested in Section 3.4 that the notion of *neutral* is biased.

3.7.2 Noticeable descriptors

The underlined cells of a given descriptor (row) highlight its noticeable emotions. For example, the results of *vertebrae flexion* suggest that this descriptor has a noticeable contribution to the expression of five emotions: sadness, anger, happiness, elation and pride.

The five descriptors with the highest level of noticeability are related to movement speed (*walk speed*, *walk step frequency* and *Laban time effort*), and spinal flexion (*vertebrae flexion* and *cervical vertebrae flexion*).

The importance of movement speed to the bodily expression of emotion is highlighted by the three first descriptors *walk speed*, *walk step frequency* and *Laban time effort*. The high level of noticeability of the two gait descriptors *walk speed* and *walk step frequency* is explained by the very low standard deviations. The *Laban time effort* is extensively referenced in the literature (see TABLE 2), leading to eight noticeable emotions.

The next two descriptors are related to spinal flexion: *vertebrae flexion* and *cervical vertebrae flexion*. They could seem intuitively correlated, however it can be observed that these two descriptors do not share the same noticeable emotions. For example, anger and pride have low standard deviations for *vertebrae flexion* and high standard deviations for *cervical vertebrae flexion*. This could suggest that the cervical vertebrae have a low contribution to the expression of these emotions, but the lower vertebrae are important. Our hypothesis is that the different parts of the vertebral column do not characterize the same emotions.

According to the noticeable cells of TABLE 4, the three descriptors related to movement speed and the two ones related to spinal flexion seem important in expressing sadness, anger and happiness. According to [103], these three emotions are scattered in the arousal-valence space. Using the results from TABLE 4, these three emotions can be defined as:

	valence	**arousal**	**spinal flex.**	**mvt. speed**
sadness	$\ominus$	$\ominus$	$\oplus$	$\ominus$
anger	$\ominus$	$\oplus$	$\oplus$	$\oplus$
happinness	$\oplus$	$\oplus$	$\ominus$	$\oplus$

We can observe that the arousal dimension is associated with movement speed and the valence dimension with spine flexion. Further studies similar to [104] could be carried out to confirm the hypothesis that some descriptors can be mapped to some dimensions of the arousal-valence space.

3.7.3 Explanation of low and zero noticeability levels

Zero and low noticeability levels are observed for the parameters at the bottom of the table and the emotions on the right side of the table. Low levels are observed because these descriptors and emotions have only cells containing a single value ($\mathrm{unnoticeable}(\mathbf{X}) = 1$). These are therefore not noticeable from a meta-analysis perspective because they have not been sufficiently studied and quantified. The meta-analysis show that the understanding and the quantification of the expression of the emotions after the *neutral* column is still an open issue. In the same way, every descriptors after *Laban flow shape* have not been explored enough.

4 Analyzing Expressive Motions in the Emilya Database

The meta-analysis is the aggregation of highly heterogeneous data, which may seem risky. For this reason we computed similar descriptors directly on an emotionally annotated motion capture database, with the aim of validating the results of the meta-analysis.

Many databases are available in the field of motion analysis [38]. We chose the Emilya database [23], [24] as it provides a large number of animation files with a wide variety of performed actions and emotions, which can lead to results that are free of a specific motion. The Emilya database contains more than 7000 motions of 7 actions: walking, sitting down, knocking at a door, lifting and throwing an object with one hand and moving objects with two hands. It includes 8 emotional expressions (sadness, anger, happiness, fear, pride, anxiety, shame and neutral). The expressive motions are performed by 11 actors (6 females and 5 males) with an average age of 26 years.

Kleinsmith et al. caution about the influence that spontaneous or acted actions can have on the relevance of analyses [10]. Regarding this issue, we argue that most of the papers used in the meta-analysis are using recorded motions of acted data. Moreover, we think that the range of values we

obtained will be particularly useful for animation synthesis applications: in this case acted emotions which are more exaggerated and stereotyped does not seem to be a problem and may even be a required quality.

4.1 Feature extraction methodology

This section describes all the steps required to extract expressive descriptors from an animation. A similar work has been conducted by Larboulette et al. [52] but our proposed framework is a complement to their work for several reasons. It introduces the use of anthropomorphic tables to ensure that motion capture data is standardized for any human. Anthropomorphic tables are also used to adequately weight the contributions of the joints and body segments. This section also addresses the question of which coordinate system should be used to represent positions and rotations (either local or absolute, depending on the use case). It also proposes a simple method to segment motion data into movement primitives (instead of using arbitrary fixed time windows). This section finally provides a dedicated part on the measurement of biomechanical joint angles. All these additions and clarifications aim at proposing a framework allowing a total reproducibility of the experiments. It also allows to compare the results of different motion capture databases or different actors.

4.1.1 File preprocessing

Each of the BVH files in the Emilya database is composed of the following items:

- A *reference skeleton*, which is a hierarchy of named joints. It is represented in Fig. 4: short labels in square brackets are used in equations to keep a compact representation. For this skeleton a *reference pose* is also defined: it corresponds to the position of each joint relative to its parent considering zero rotation.
- *Successive poses* representing the animation. Each pose is defined as a time series of joint transforms (translations and rotations) which must be applied to the *reference pose*.

Describing how to obtain such data from the BVH files is beyond the scope of this topic, and documentation can be found in [105].

Fig. 4 represents the Emilya *reference skeleton* and how it is associated with a segmented body representation. Each of the body segments are referenced in De Leva's anthropomorphic table [106]. We use the average of the female and male values as reference.

From this Section to Section 4.1.4, M will refer to the number of frames in the animation and N to the number of joints in the hierarchy. The tilde notation $\tilde{s}$ above any symbol s will refer to relative data (in opposition to absolute data).

Each animation coming from the Emilya database can be represented as a succession of M poses over a period of time $t_M - t_1$, where $t_{i+1} - t_i = \Delta t$ is a fixed constant. We represent a pose at time t_i as a hierarchy of N joints, defined by a set $\tilde{\mathbf{X}}$ of N relative positions constant at every time t_i (except for the Pelvis joint) and a set $\tilde{\mathbf{E}}(t_i)$ of N varying relative ordered angles:

$$\tilde{\mathbf{X}}(t_i) = \{\tilde{\mathbf{x}}_1, \ \dots \ , \tilde{\mathbf{x}}_N\}(t_i) \tag{10}$$

$$\tilde{\mathbf{E}}(t_i) = \{\tilde{\mathbf{e}}_1, \ \dots \ , \tilde{\mathbf{e}}_N\}(t_i) \tag{11}$$

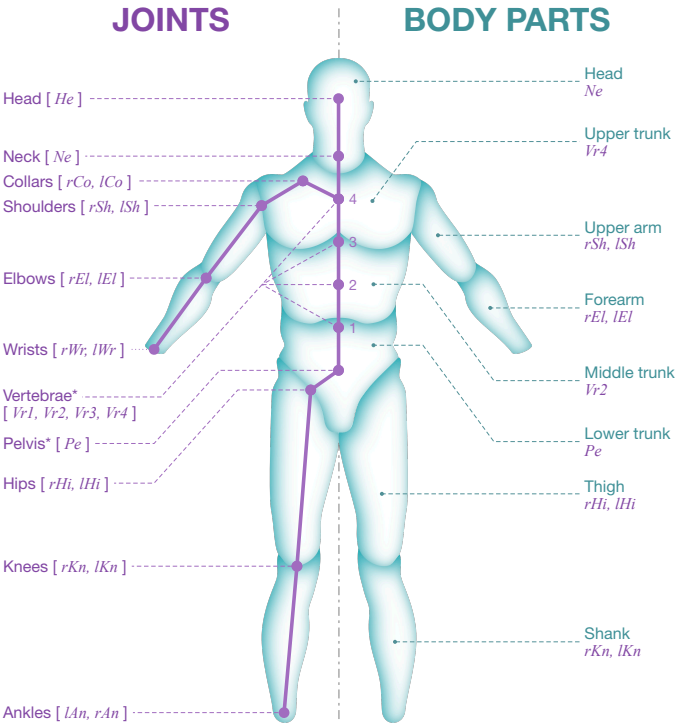


Fig. 4. Our reference skeleton: On the left, violet dots and labels refer to joints from the Emilya BVH files. The asterisk symbol * indicates a label that has been changed for readability issues. Short labels in square brackets are used in equations to keep a compact representation. On the right, blue labels refer to anatomical segments from De Leva's anthropomorphic table. Below each segment, its parent joint is indicated in violet.

where $\tilde{\mathbf{x}}_{Pe}(t_i)$ and $\tilde{\mathbf{e}}_{Pe}(t_i)$ are expressed relative to the world.

The raw input data is transformed as follows:

$$\tilde{\mathbf{X}}'(t_i) = \{scale \, . \, \tilde{\mathbf{x}}_1, \ \dots \ , \ scale \, . \, \tilde{\mathbf{x}}_N\}(t_i) \tag{12}$$

$$\tilde{\mathbf{E}}'(t_i) = \{\mathcal{F}_{BW3}(\tilde{\mathbf{e}}_1), \ \dots \ , \ \mathcal{F}_{BW3}(\tilde{\mathbf{e}}_N)\}(t_i) \tag{13}$$

where $\mathcal{F}_{BW3}$ represents a 3rd order Butterworth filter of a cut frequency of one tenth of the sampling frequency applied on each Euler angle. The choice of this specific filtering procedure have been motivated by [107]. Therefore, this produces smooth signal even at third order derivative whilst keeping great motion fidelity.

The value $scale$ is applied to positions in order to get standardized measurements expressed relative to a human skeleton of a reference height:

$$scale = \frac{AnthropomorphicHeight}{EmilyaHeight} \tag{14}$$

where $AnthropomorphicHeight$ refers to the sum of *trunk* (suprasternale to middle hip joint centers), *thigh* (hip joint center to knee joint center) and *shank* (knee joint center to lateral malleolus) heights from De Leva's table [106] and where $EmilyaHeight$ is computed as the height difference between the middle of the shoulders and the middle of the ankles.

4.1.2 Motion measurements and analysis toolbox

In this section are presented measurements and analysis techniques used for the extraction of high level expressive features of motion.

• **Opposite notion:** it should be noted that many descriptors are computed using the opposite notion of the original formulation to stick to the reference descriptors provided in Section 3.2. For example, a positive value of smoothness corresponds to the additive inverse of a jerk value. The opposite notion of a feature is denoted as:

$$Feature_{\ominus} = -Feature \tag{15}$$

• **Parent joint:** the direct parent of a joint k is denoted $\text{parent}(k)$.

• **Segment length:** for each joint k, the associated segment length is defined as:

$$l_k = \begin{cases} 0, & \text{if } k = Pe \\ scale \,.\, \|\tilde{\mathbf{x}}_k\|, & \text{otherwise} \end{cases} \tag{16}$$

Bone lengths are defined using $scale$ and t_0 in order to ensure constant and standardized values (i.e. animation without scaling factor).

• **Relative joint quaternions:** the set of relative joint quaternions at time t_i is defined as:

$$\tilde{\mathbf{Q}}(t_i) = \{\tilde{\mathbf{q}}_1,\ \tilde{\mathbf{q}}_2,\ \dots\ ,\ \tilde{\mathbf{q}}_N\}(t_i) \tag{17}$$

These are obtained by converting the set of filtered ordered Euler angles $\tilde{\mathbf{E}}'(t_i)$.

• **Absolute joint quaternions:** the set of absolute joint quaternions at time t_i is defined as:

$$\mathbf{Q}(t_i) = \{\mathbf{q}_1,\ \mathbf{q}_2,\ \dots\ ,\ \mathbf{q}_N\}(t_i) \tag{18}$$

These are obtained using a forward kinematic algorithm on the skeleton hierarchy with the set of relative quaternions $\tilde{\mathbf{Q}}(t_i)$.

• **Quaternion operations:** given two rotation quaternions $\mathbf{q}(t_a)$ and $\mathbf{q}(t_b)$, the delta rotation quaternion from time t_b to t_a is defined as:

$$\Delta\mathbf{q}(t_a, t_b) = \mathbf{q}(t_a) * \overline{\mathbf{q}}(t_b) \tag{19}$$

where the symbol $*$ denotes the quaternion product and $\overline{\mathbf{q}(t_b)}$ is the conjugate of $\mathbf{q}(\mathbf{t_b})$.

The angle of a quaternion $\mathbf{q}$ is obtained from its w component:

$$\text{angle}(\mathbf{q}) = 2 \,.\, \text{acos}(\mathbf{q}_w) \tag{20}$$

• **Absolute joint positions:** the set of absolute joint positions at time t_i is defined as:

$$\mathbf{X}(t_i) = \{\mathbf{x}_1,\ \mathbf{x}_2,\ \dots\ ,\ \mathbf{x}_N\}(t_i) \tag{21}$$

These are obtained using a forward kinematic algorithm on the skeleton hierarchy using the set of relative positions $\tilde{\mathbf{X}}'(t_i)$ and the set of relative joint quaternions $\tilde{\mathbf{Q}}(t_i)$.

• **Local end effector positions:** whenever it is impossible to analyze movements in world space coordinates (e.g analysing arm swing during a walk animation), end effectors positions are expressed relative to a specific position and orientation in order to obtain exploitable data.

Such spaces are defined using a three-dimensional origin vector $\mathbf{o}$ and a triplet of vectors $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ representing a left/up/front orthonormal basis.

The normalization function applied to any vector $\mathbf{x}$ is defined as:

$$\mathbf{unit}(\mathbf{x}) = \frac{\mathbf{x}}{\|\mathbf{x}\|} \tag{22}$$

Feet and torso are expressed relative to a *Legs* orthonormal basis:

$$\begin{cases} \mathbf{o}_{Legs}(t_i) = \mathbf{x}_{Pe}(t_i) \\ \mathbf{v}_{Legs}(t_i) = \mathbf{unit}(\ \mathbf{x}_{Vr1}(t_i) - \mathbf{x}_{Pe}(t_i)\ ) \\ \mathbf{w}_{Legs}(t_i) = \mathbf{unit}(\ (\mathbf{x}_{lHi}(t_i) - \mathbf{x}_{rHi}(t_i)) \times \mathbf{v}_{Legs}(t_i)\ ) \\ \mathbf{u}_{Legs}(t_i) = \mathbf{v}_{Legs}(t_i) \times \mathbf{w}_{Legs}(t_i) \end{cases} \tag{23}$$

Hands and head are expressed relative to the *Arms* orthonormal basis:

$$\begin{cases} \mathbf{o}_{Arms}(t_i) = \mathbf{x}_{Vr4}(t_i) \\ \mathbf{v}_{Arms}(t_i) = \mathbf{unit}(\ \mathbf{x}_{Ne}(t_i) - \mathbf{x}_{Vr4}(t_i)\ ) \\ \mathbf{w}_{Arms}(t_i) = \mathbf{unit}(\ (\mathbf{x}_{lCo}(t_i) - \mathbf{x}_{rCo}(t_i)) \times \mathbf{v}_{Arms}(t_i)\ ) \\ \mathbf{u}_{Arms}(t_i) = \mathbf{v}_{Arms}(t_i) \times \mathbf{w}_{Arms}(t_i) \end{cases} \tag{24}$$

The set of end effectors identifiers is defined as:

$$\mathbf{EndEff} = \{lWr,\ rWr,\ lAn,\ rAn,\ Vr4,\ He\} \tag{25}$$

with $\{lWr,\ rWr\}$ representing hands, $\{lAn,\ rAn\}$ representing feet, $\{He\}$ representing Head and $\{Vr4\}$ representing Torso.

The set of end effector positions expressed in their corresponding space (either *Arms* or *Legs*) is defined as:

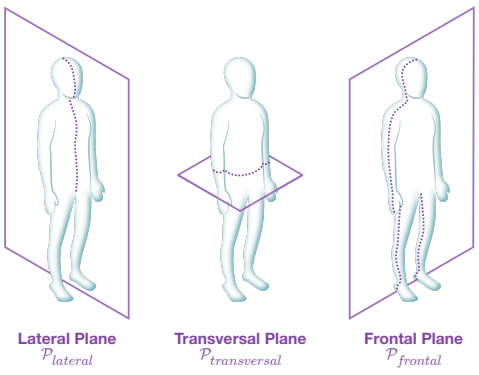
$$\tilde{\mathbf{Eff}}(t_i) = \{\tilde{\mathbf{eff}}_k(t_i) \mid k \in \mathbf{EndEff}\} \tag{26}$$

• **Anatomical planes:** any plane can defined with a coordinate $\mathbf{o}$ and a direction $\mathbf{n}$ representing the normal vector to the plane with $\|\mathbf{n}\| = 1$:

$$\mathcal{P} = \{\mathbf{o}, \mathbf{n}\}$$

For a given time t_i the three orthogonal anatomical planes are defined as follows:

$$\begin{aligned} \mathcal{P}_{lateral}(t_i) &= \{\mathbf{o}_{Legs}(t_i),\ \mathbf{u}_{Legs}(t_i)\} \\ \mathcal{P}_{transversal}(t_i) &= \{\mathbf{o}_{Legs}(t_i),\ \mathbf{v}_{Legs}(t_i)\} \\ \mathcal{P}_{frontal}(t_i) &= \{\mathbf{o}_{Legs}(t_i),\ \mathbf{w}_{Legs}(t_i)\} \end{aligned} \tag{27}$$



• **Distance to plane:** The signed and unsigned distance of a point $\mathbf{p}$ to a plane $\mathcal{P} = \{\mathbf{o}, \mathbf{n}\}$ are respectively defined as:

$$\text{sdist}(\mathbf{p}, \mathcal{P}) = (\mathbf{p} - \mathbf{o}) \cdot \mathbf{n} \tag{28}$$

$$\text{dist}(\mathbf{p}, \mathcal{P}) = |\text{sdist}(\mathbf{p}, \mathcal{P})| \tag{29}$$

• **Travelled distance:** for any varying position $\mathbf{p}(t)$ at time t, its instantaneous travelled distance is defined as:

$$\text{travel}(\mathbf{p}, t) = \|\mathbf{p}(t) - \mathbf{p}(t - \Delta t)\| \tag{30}$$

• **Translational joint velocity, acceleration and jerk:** for any varying position $\mathbf{p}(t_i)$ at time t_i, its velocity, acceleration and jerk are defined using finite central differences as in [52]:

$$\mathbf{vel}(\mathbf{p}, t_i) = \frac{\mathbf{p}(t_{i+1}) - \mathbf{p}(t_{i-1})}{2 \,.\, \Delta t} \tag{31}$$

$$\mathbf{acc}(\mathbf{p}, t_i) = \frac{\mathbf{p}(t_{i+1}) - 2 \,.\, \mathbf{p}(t_i) + \mathbf{p}(t_{i-1})}{\Delta t^2} \tag{32}$$

$$\mathbf{jerk}(\mathbf{p}, t_i) = \frac{\mathbf{p}(t_{i+2}) - 2 \,.\, \mathbf{p}(t_{i+1}) + 2 \,.\, \mathbf{p}(t_{i-1}) - \mathbf{p}(t_{i-2})}{2 \,.\, \Delta t^3} \tag{33}$$

These formulas can be applied to any scalar expression $s(t_i)$ and any varying vector of arbitrary dimension.

• **Curvature:** for any varying position $\mathbf{p}(t_i)$ at time t_i, its curvature and radius of curvature are respectively defined as:

$$\text{curv}(\mathbf{p}, t_i) = \frac{\|\mathbf{vel}(\mathbf{p}, t_i) \times \mathbf{acc}(\mathbf{p}, t_i)\|}{\|\mathbf{vel}(\mathbf{p}, t_i)\|^3} \tag{34}$$

$$\text{rcurv}(\mathbf{p}, t_i) = \frac{1}{\text{curv}(\mathbf{p}, t_i)} \tag{35}$$

and for any varying scalar $s(t_i)$ at time t_i, the curvature and radius of curvature are simplified to:

$$\text{curv}(s, t_i) = \frac{|\text{acc}(s, t_i)|}{(1 + \text{vel}(s, t_i)^2)^{\frac{3}{2}}} \tag{36}$$

$$\text{rcurv}(s, t_i) = \frac{1}{\text{curv}(s, t_i)} \tag{37}$$

• **Rotational joint velocity, acceleration and jerk:** for any varying rotation quaternion $\mathbf{q}(t_i)$ at time t_i, its rotational velocity, rotational acceleration and rotational jerk are respectively defined as:

$$\text{rotvel}(\mathbf{q}, t_i) = \frac{\text{angle}(\Delta\mathbf{q}(t_{i+1}, t_{i-1}))}{2 \,.\, \Delta t} \tag{38}$$

$$\text{rotacc}(\mathbf{q}, t_i) = \frac{\text{angle}(\Delta\dot{\mathbf{q}}(t_{i+1}, t_{i-1}))}{\Delta t^2} = \frac{\text{angle}(\Delta\mathbf{q}(t_{i+1}, t_i) * \overline{\Delta\mathbf{q}(t_i, t_{i-1})})}{\Delta t^2} \tag{39}$$

$$\text{rotjerk}(\mathbf{q}, t_i) = \frac{\text{angle}(\Delta\dot{\mathbf{q}}(t_{i+2}, t_i) * \overline{\Delta\dot{\mathbf{q}}(t_i, t_{i-2})})}{\Delta t^2} \tag{40}$$

These equations are obtained by adapting finite central differences to quaternions, assuming that a delta quaternion $\Delta\mathbf{q}(t_a, t_b)$ defined in (19) is equivalent to the following relation in cartesian space:

$$\Delta\mathbf{p}(t_a, t_b) = \mathbf{p}(t_a) - \mathbf{p}(t_b)$$

• **Kinetic energy of human movement:** if the human body is represented as a poly-articulated system of S rigid segments with a mass (represented in Fig. 4), then its total kinetic energy can be defined as the sum of the kinetic energies of its segments [108]:

$$E_{body}(t_i) = \sum_{k=1}^{S} E_k(t_i) \tag{41}$$

The kinetic energy of a single rigid segment k at time t_i is equal to the sum of its translational and rotational kinetic energies:

$$E_k(t_i) = E_{trans}(k,\, t_i) + E_{rot}(k,\, t_i) \tag{42}$$

The position of the body center of mass (CoM) at time t_i is computed from s segments as follows:

$$\mathbf{CoM}_{Body}(t_i) = \frac{\sum_{k=1}^{S} m_k \,.\, \mathbf{CoM}_k(t_i)}{\sum_{k=1}^{S} m_k} \tag{43}$$

where m_k is the k^{th} segment mass. and $\mathbf{CoM}_k(t_i)$ denotes the CoM absolute position of the k^{th} segment. It is calculated using the segment's CoM longitudinal position.

Segment masses and CoM longitudinal positions are extracted from De Leva's table [106], taking the mean of female and male values.

The position of the k^{th} segment center of mass relative to the body center of mass is defined as:

$$\mathbf{CoM}_{k/Body}(t_i) = \mathbf{CoM}_k(t_i) - \mathbf{CoM}_{Body}(t_i) \tag{44}$$

Translational energy of segment k at time t_i is defined as:

$$E_{trans}(k,\, t_i) = \frac{1}{2} \,.\, m_k \,.\, \|\mathbf{vel}(\mathbf{CoM}_{k/Body}, t_i)\|^2 \tag{45}$$

Rotational energy of segment k at time t_i is defined as the sum of the rotational energies on the $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ axes:

$$E_{rot}(k,\, t_i) = \sum_{\mathbf{axis} \in \{\mathbf{x},\mathbf{y},\mathbf{z}\}} \frac{1}{2} \,.\, I_{k,\mathbf{axis}} \,.\, \omega_{k,\mathbf{axis}}(t_i)^2 \tag{46}$$

where $I_{k,\mathbf{axis}}$ is the moment of inertia of segment k around $\mathbf{axis}$ and $\omega_{k,\mathbf{axis}}(t_i)$ is the absolute angular velocity of segment k around $\mathbf{axis}$ at time t_i.

Each moment of inertia is calculated as follows:

$$I_{k,\mathbf{axis}} = m_k \,.\, r_{k,\mathbf{axis}}^2 \tag{47}$$

where $r_{k,\mathbf{axis}}$ is the radius of gyration around $\mathbf{axis}$. It is extracted from De Leva's table [106], taking the mean of female and male values.

Each angular velocity is defined as:

$$\omega_{k,\mathbf{axis}}(t_i) = \frac{\Delta\theta_{k,\mathbf{axis}}(t_i)}{2 \,.\, \Delta t} \tag{48}$$

where $\Delta\theta_{k,\mathbf{axis}}(t_i)$ is the delta euler angle around $\mathbf{axis}$.

Delta euler angles are obtained from a delta quaternion by using an Euler to quaternion conversion in XYZ order. This delta quaternion is denoted $\Delta\mathbf{q}(t_{i+1}, t_{i-1})$ and corresponds to the absolute quaternion of the parent joint associated to segment k (see Fig. 4). The delta is defined between t_{i+1} and t_{i-1}, which explains the division by $2 \,.\, \Delta t$.

• **End effectors motion segmentation:** Most expressive descriptors are computed frame by frame and do not need motion segmentation. On the other hand, some movement

analysis techniques need to split the end effector animation into meaningful primitives. There are a lot of advanced segmentation techniques [109] [110]. We developed an intuitive and computationally efficient approach well suited for our dataset analysis. Our segmentation approach is based on the observation that meaningful end effector movement phases are often characterized by a significantly higher radius of curvature and at least one noticeable velocity peak, as illustrated in Fig. 5.

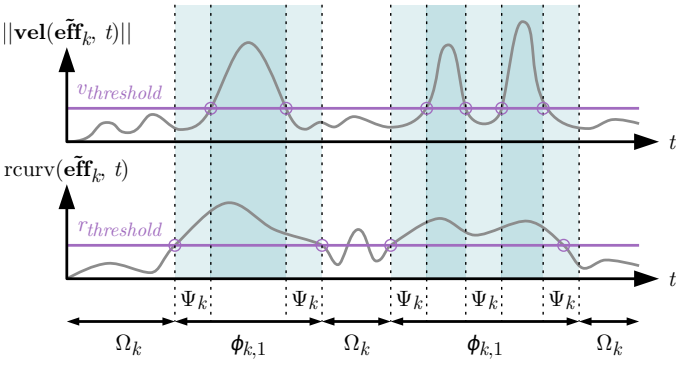


Fig. 5. Signal motion segmentation using thresholds on velocity and radius of curvature.

Segmentation allows to identify the intervals of *movement* and *pause*. This is done by looking at the variations of the end effectors positions.

The set of movement intervals of the k^{th} end effector is defined as:

$$\mathbf{\Phi}_k = \{\phi_{k,1},\ \phi_{k,2},\ \ldots,\ \phi_{k,i}\} \tag{49}$$

Each $\phi_{k,i}$ is defined as a set of successive time values starting at t_{start}, ending at t_{end} and with a time step of Δt:

$$\phi_{k,i} = \{t_{start} + \Delta t,\ t_{start} + 2\,.\,\Delta t,\ \ldots,\ t_{end}\} \tag{50}$$

Any movement interval $\phi_{k,i}$ has the following properties:

- The radius of curvature $\mathrm{rcurv}(\tilde{\mathbf{eff}}_k, t)$ is always above the threshold $r_{threshold} = 0.05\ m$ for any time in the interval $\phi_{k,i}$
- The velocity magnitude $\|\mathbf{vel}(\tilde{\mathbf{eff}}_k, t)\|$ is above the threshold $v_{threshold} = 0.3\ m/s$ for at least one time in the interval $\phi_{k,i}$

The threshold values $v_{threshold}$ and $v_{threshold}$ have been determined empirically for the Emilya database.

Let denote $\mathbf{A} = \{t_1, t_2, \ldots, t_M\}$ the set of each successive time values in the whole animation. The set of *idle* intervals for joint k is then defined as:

$$\mathbf{\Omega}_k = \mathbf{A} - \mathbf{\Phi}_k \tag{51}$$

Finally, we denote $\mathbf{\Psi}_k$ the set of time values which belong to a $\mathbf{\Phi}_k$, but where the speed does not reach the velocity threshold $v_{threshold}$. These correspond to the light blue portions in Fig. 5.

• **Result based on the aggregation of multiple joint or segment values:** many high level descriptors in the following subsections are expressed as a single scalar for the whole body. This is based on the assumption that a body aggregation value is equal to the weighted sum of its subpart contributions:

$$value_{body} = \sum_{subpart=1}^{subpartCount} weight_{subpart}\ .\ value_{subpart} \tag{52}$$

This formulation is regularly employed in the field of motion analysis [52], [85], [102], and we will make the same assumption for this study. Since the meta-analysis study set contains many studies based on perceptual assessments, it is expected that any feature $value_{body}$ computed from the dataset should match the way a human would rate it.

In order to match $value_{body}$ with an experiment based on perceptual assessment, two questions have to be addressed:

- Choose the appropriate coordinate system to calculate the subpart contribution $value_{subpart}$.
- Choose the appropriate weight $weight_{subpart}$ according to the measured descriptor.

Regarding the first point, we have noticed that a lot of papers in motion analysis do not explicitly specify a coordinate system (such as in [52], [85], [102], etc.). However, this choice greatly influences the result of $value_{body}$. Let consider the kinematic chain $A \rightarrow B \rightarrow C$. It is possible to express C either (1) relative to a reference point such as world origin or center of mass, (2) relative to root A or (3) relative to its direct parent B. This choice depends on the computed feature. We argue that option (3) is preferable as soon as we consider the joints of kinematic chains: it allows measuring the movements of C independently of the influence of its parents. In this case, it should be noted that the position of C relative to B is equivalent to the relative rotation applied to C and weighted by the length between B and C.

The question of which weight to apply to $value_{subpart}$ is also worth discussing. In the specific case of energy, the weight $weight_{subpart}$ is equal to the segment mass, but it is not necessarily relevant for all measurements. For example, we can wonder if the computation of the fluidity of joints should not be weighted by the segment length rather than its mass. Only a few work which attempted to find extraction formulas validated by perceptual assessments [85], [111] . Making such a validation is out of the scope of this study, as it would require a whole study dedicated to this topic. Instead, for each of the expressive features involving a weighted sum we have empirically chosen the most appropriate weight:

- mass m_k for all the features representing movement dynamics (energy or force).
- segment length l_k for all the features representing kinematics motion qualities.
- all features representing aggregations of specific positions or trajectories (such as the end effectors) are not weighted ($weight_{subpart} = 1$).

4.1.3 Extracting high level features

• **Smoothness:** using the definition provided in [52] the smoothness is linked to the opposite notion of jerk. We

calculate it as the opposite notion of the average of the angular jerk over the whole animation and for all the joints:

$$Smoothness_{\ominus} = \frac{1}{M \,.\, N} \sum_{i=1}^{M} \sum_{k=1}^{N} l_k \,.\, |\mathrm{rotjerk}(\tilde{\mathbf{q}}_{\mathrm{parent}(k)}, t_i)| \tag{53}$$

The measure of $TorsoSmoothness$ is the $Smoothness$ calculated on a reduced set of joints $\{Vr1,\ Vr2,\ Vr3,\ Vr4\}$. $HeadSmoothness$ is the measure of $Smoothness$ on the single element $\{He\}$.

• **Regularity:** it is a complex notion involving both movement continuity (trajectory, speed, etc.) or the repetition of movement patterns [24]. The pattern repetition tracking is based on signal autocorrelation [112] which is not particularly suitable for a large database analysis such as Emilya, where repetitions are mostly due to the performed gestures (e.g. knocking, walking, etc.). As a consequence, we define the regularity solely on its continuity component. The regularity of the k^{th} end effector is calculated as the opposite notion of its velocity curvature (high velocity curvature is synonym of low regularity):

$$reg_k \ominus (t_i) = \mathrm{curv}(\|\mathbf{vel}(\tilde{\mathbf{eff}}_k, t_i)\|, t_i) \tag{54}$$

The global regularity is defined as the mean of the end effector regularities over the whole animation:

$$Regularity_{\ominus} = \frac{1}{M \,.\, \mathrm{card}(\mathbf{EndEff})} \sum_{i=1}^{M} \sum_{k \in \mathbf{EndEff}} reg_k \ominus (t_i) \tag{55}$$

The measure of $TorsoRegularity$ and $HeadRegularity$ are defined as $Regularity$ but calculated respectively on $\{Vr4\}$ and $\{He\}$.

• **Movement activity:** it is calculated using $\mathbf{\Phi}_k$, $\mathbf{\Psi}_k$, $\mathbf{\Omega}_k$ and $v_{threshold}$ described in Fig. 5.

The level of activity of the k^{th} end effector is defined as:

$$\begin{aligned} Active_k \ = & \sum_{\{t \in \mathbf{\Phi}_k \mid t \notin \mathbf{\Psi}_k\}} \mathrm{travel}(\tilde{\mathbf{eff}}_k, t) \\ & + \sum_{\{t \in \mathbf{\Phi}_k \mid t \in \mathbf{\Psi}_k\}} \frac{\|\mathbf{vel}(\tilde{\mathbf{eff}}_k, t)\|}{v_{threshold}} \,.\, \mathrm{travel}(\tilde{\mathbf{eff}}, t) \end{aligned} \tag{56}$$

The idle level of the k^{th} end effector is defined as:

$$\begin{aligned} Idle_k \ = & \sum_{t \in \mathbf{\Omega}_k} v_{threshold} \,.\, \Delta t \\ & + \sum_{\{t \in \mathbf{\Phi}_k \mid t \in \mathbf{\Psi}_k\}} \left(1 - \frac{\|\mathbf{vel}(\tilde{\mathbf{eff}}_k, t)\|}{v_{threshold}}\right) \,.\, \mathrm{travel}(\tilde{\mathbf{eff}}_k, t) \end{aligned} \tag{57}$$

The measure of movement activity of the k^{th} end effector is then defined as:

$$Activity_k = \frac{Active_k}{Idle_k + Active_k} \tag{58}$$

And the measure of body movement activity is defined as:

$$Activity = \frac{1}{\mathrm{card}(\mathbf{EndEff})} \sum_{k \in \mathbf{EndEff}} Activity_k \tag{59}$$

The measure of $ArmActivity$ is defined as $Activity$ but calculated on the end effectors subset $\{lWr,\ rWr\}$.

• **Laban factors:** The work of Laban is extensively used in the field of motion analysis. The major drawback comes from the fact that these factors are very subjective: they are based on personal interpretation, and require a certification in Laban Movement Studies. TABLE 5 gives a review of existing formulations for each of the Laban effort/shape factors. Due to high subjectivity of the Laban parameters, this leads to various computation methods in scientific literature. Bolded entries are the definitions we have selected (most dominant). Selected methods are detailed in the rest of this section.

• **Laban space effort:** it is defined as the level of directness of the end effectors. First, we need to segment the end effectors motions into meaningful movement primitives using the method described in Section 4.1.2.

Then from any movement interval $\phi_{k,i}$ of the k^{th} end effectors, the trajectory length is computed as:

$$TrajectoryLength_{k,i} = \sum_{t \in \phi_{k,i}} \mathrm{travel}(\tilde{\mathbf{eff}}, t) \tag{60}$$

and the straight line length is defined as:

$$StraightLength_{k,i} = \|\tilde{\mathbf{eff}}(t_{end}) - \tilde{\mathbf{eff}}(t_{start})\| \tag{61}$$

where t_{start} and t_{end} are the boundaries of the given interval $\phi_{k,i}$.

The measure of trajectory directness for the k^{th} end effectors is defined as:

$$Directness_k = \frac{\sum_{i=1}^{\mathrm{card}(\mathbf{\Phi}_k)} TrajectoryLength_{k,i}}{\sum_{i=1}^{\mathrm{card}(\mathbf{\Phi}_k)} StraightLength_{k,i}} \tag{62}$$

The space effort is calculated as the mean of end effectors directness:

$$SpaceEffort = \frac{1}{\mathrm{card}(\mathbf{EndEff})} \sum_{k \in \mathbf{EndEff}} Directness_k \tag{63}$$

• **Laban time effort:** it is defined as the opposite notion of the average of the angular acceleration over the whole animation and for all the joints:

$$TimeEffort_{\ominus} = \frac{1}{M \,.\, N} \sum_{i=1}^{M} \sum_{k=1}^{N} l_k \,.\, |\mathrm{rotacc}(\tilde{\mathbf{q}}_{\mathrm{parent}(k)}, t_i)| \tag{64}$$

where l_k is the segment length as in (16) and rotacc is the angular acceleration as in (39).

• **Laban weight effort:** it is calculated as the opposite notion of the mean of the body kinetic energy over the whole animation:

$$WeightEffort_{\ominus} = \frac{1}{M} \sum_{i=1}^{M} E_{body}(t_i) \tag{65}$$

where $E_{body}(t_i)$ is the body kinetic energy as in (41).

• **Laban flow effort:** there is no clear consensus on the formulation of Laban flow effort (TABLE 5), but following the most dominant definition, it is related to jerk. As a result, it is calculated in the same way as the smoothness in (53):

$$FlowEffort_{\ominus} = \frac{1}{M \,.\, N} \sum_{i=1}^{M} \sum_{k=1}^{N} l_k \,.\, |\mathrm{rotjerk}(\tilde{\mathbf{q}}_{\mathrm{parent}(k)}, t_i)| \tag{66}$$

TABLE 5
List of existing computation methods for the extraction Laban effort/shape factors: most popular methods are represented as bolded entries.

Laban factor	Computation methods	References
flow effort	**Jerk**	[113] [85] [52] [114]
	Curvature	[115] [116]
	Velocity	[116] [117]
	Acceleration	[116] [117]
	Distance between minimum jerk trajectory and measured trajectory	[118]
	Kinetic Energy	[117]
	Frequency and rhythm of motion/pause phases	[117]
	Distance from wrists to pelvis	[119]
flow shape	**Bounding volume (Bounding box or Convex Hull)**	[52] [120] [114]
space effort	**Ratio between measured trajectory length and straight line segment length**	[117] [119] [85] [118] [52] [114]
	Level of similarity between end effector directions	[115] [85] [120] [94]
space shape	**Measure of body occupation using projection on *transversal* plane**	[52] [120]
time effort	**Acceleration**	[119] [116] [115] [85] [118] [52] [120]
	Movement duration	[117] [113]
	Movement Frequency (Movement number per time unit)	[94]
	Velocity	[116]
	Movement / Pause ratio	[113]
	Duration of pause phases	[113]
	Jerk	[111]
time shape	**Measure of body occupation using projection on *lateral* plane**	[52] [120]
weight effort	**Kinetic energy**	[102] [94] [115] [85] [52] [120]
	Acceleration	[113] [114] [117]
	Velocity	[113]
	Vertical component of wrist's position (relative to pelvis)	[119]
	Electromyogram	[111]
weight shape	**Measure of body occupation using projection on *frontal* plane**	[52] [120]

This does not mean that smoothness and Laban flow effort are two equivalent concepts, since the formula for the Laban flow effort have been chosen despite a lack of scientific consensus.

• **Laban space shape:** is the mean of joint distances to the lateral plane over the whole animation:

$$SpaceShape = \frac{1}{M \,.\, N}\sum_{i=1}^{M}\sum_{k=1}^{N} \mathrm{dist}(\mathbf{x}_k(t_i),\ \mathcal{P}_{lateral}(t_i)) \tag{67}$$

where $\mathcal{P}_{lateral}(t_i)$ is the lateral plane as in (27) and dist is the distance point-plane as in (29).

The measure of the space shape component on the arms is denoted $ArmSpaceShape$. The formula is the same as $SpaceShape$, except that it is calculated on a reduced set of joints $\{lSh,\ rSh,\ lEl,\ rEl,\ lWr,\ rWr\}$.

• **Laban time shape:** it is calculated as the mean of signed joint distances to the frontal plane over the whole animation:

$$TimeShape = \frac{1}{M \,.\, N}\sum_{i=1}^{M}\sum_{k=1}^{N} \mathrm{sdist}(\mathbf{x}_k(t_i),\ \mathcal{P}_{frontal}(t_i)) \tag{68}$$

• **Laban weight shape:** is the mean of joint distances to the transversal plane over the whole animation:

$$WeightShape = \frac{1}{M \,.\, N}\sum_{i=1}^{M}\sum_{k=1}^{N} \mathrm{dist}(\mathbf{x}_k(t_i),\ \mathcal{P}_{transversal}(t_i)) \tag{69}$$

The measure of the weight shape component on the torso is denoted $TorsoWeightShape$. The formula is the same as $WeightShape$ except that it is calculated on a reduced set of joints $\{Vr1,\ Vr2,\ Vr3,\ Vr4,\ lCo,\ rCo,\ lSh,\ rSh\}$.

• **Laban flow shape:** it is calculated as the mean of the skeleton bounding volume over the whole animation:

$$FlowShape = \frac{1}{M}\sum_{i=1}^{M} \mathrm{boundingvolume}(t_i) \tag{70}$$

where $\mathrm{boundingvolume}(t_i)$ is the volume of the bounding box calculated using the set of joint positions at time t_i expressed in the *Legs* orthonormal basis (23).

• **Laban torso flow shape:** using the set of absolute joint positions $\mathbf{X}(t_i)$, let denote $\mathrm{trianglearea}(lSh, rSh, Pe, t_i)$ the triangle area formed by the shoulders and the first vertebra at time t_i. The torso flow shape is defined as the mean of the triangle area over the whole animation:

$$TorsoFlowShape = \frac{1}{M}\sum_{i=1}^{M} \mathrm{trianglearea}(lSh, rSh, Pe, t_i) \tag{71}$$

4.1.4 Extracting Biomechanical features

Biomechanical features (prefixed by *angle* or *amplitude angle*) are straightforward to define and calculate when based on formal definitions of functional anatomy.

• **Biomechanical joint angle:** According to the degrees of freedom of a joint, its biomechanical angles are measured on one, two or three planes.

To obtain such a decomposition, a segment (starting from the joint position to its unique child) is projected onto three orthogonal planes. These planes are defined using the joint position and an orthonormal basis $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ which differs depending on the considered joint to obtain a zero angle in rest position. The correct orthonormal bases can easily

be deduced from Kapandji's illustrations in [95], [96], [97]. Such a decomposition ensures that the angles will remain in the range $[-\pi, \pi]$.

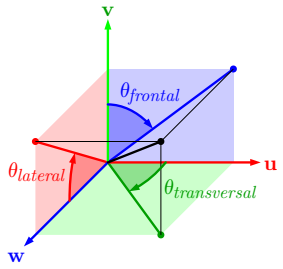


Fig. 6. Decomposition of a joint rotation using plane projection

For example, as shown in Fig. 6, in the shoulder joint breaks down into 3 biomechanical angles: the flexion/extension angle $\theta_{lateral}$ on the lateral plane, the abduction/adduction $\theta_{frontal}$ angle on the frontal plane and the horizontal flexion/extension angle $\theta_{transversal}$ on the transversal plane.

Let denote $\theta(t_i)$ any biomechanical angle (such as shoulder flexion angle, scapula medial rotation angle, etc.) measured at time t_i. The mean biomechanical angle over the whole animation is defined as follows:

$$\bar{\theta} = \frac{1}{M}\sum_{i=1}^{M}\theta(t_i) \tag{72}$$

where $\bar{\theta}$ gives information about the postural tendency of a given joint (e.g. is the shoulder rather oriented forward or backward).

• **Biomechanical joint angle amplitude:** A number of $ExCount$ local extrema are extracted from the discrete time series $\theta(t_i)$ of M frames. As represented in Fig. 7, each extremum e_i is either a local minimum or maximum of $\theta(t_i)$. The first extremum e_1 is positioned at t_1.

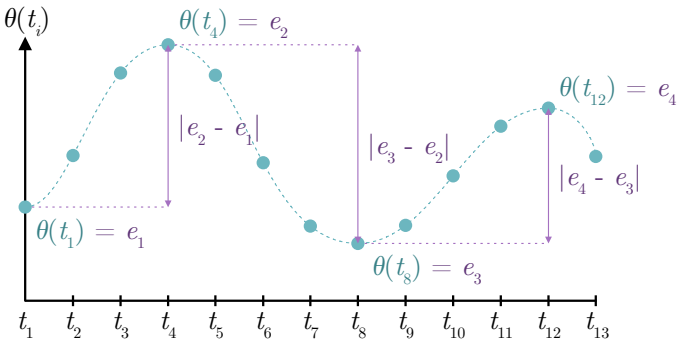


Fig. 7. Extrema sampling on a discrete set of measures

The biomechanical angle amplitude for a given animation is defined as:

$$\theta_{amp} = \frac{1}{ExCount - 1}\sum_{i=2}^{ExCount}|e_i - e_{i-1}| \tag{73}$$

4.2 Results of the descriptors extraction in the Emilya database

All the equations provided from Section 4.1.1 to Section 4.1.4 are given for a single animation file. The extraction results of all these animations need to be aggregated in order to obtain a table comparable to the one calculated in the meta-analysis.

The Emilya dataset is divided into 8 emotions: neutral, anger, anxiety, joy, panic fear, pride, sadness and shame. Let denote $\mathbf{Files}_{emotion}$ the set of $N_{emotion}$ animation files for a given emotion in the Emilya dataset.

For a given expressive feature and a given emotion of the Emilya dataset, let denote $\mathbf{M}_{feature,emotion}$ the set of $N_{emotion}$ measures extracted from $\mathbf{Files}_{emotion}$. The aggregation of a given $feature$ for a given $emotion$ is equal to the mean and the standard deviation of the measurement set $\mathbf{M}_{feature,emotion}$.

The extraction results in TABLE 6 are obtained by computing such aggregation for every features and every emotions. It should be noticed that the measurements are expressed in their original units.

Some features of the meta-analysis have not been computed in the Emilya analysis because they were too specific to a given animation (such as walk speed, walk vertical head sway, etc.) or they needed advanced gesture knowledge (such as approach, exaggeration, movement anticipation, etc.).

5 Comparison between meta-analysis and Emilya dataset extraction

The results of the meta-analysis and the Emilya analysis share the same features, but the values have been calculated using completely different methodologies. Evaluating the similarity rate between the two tables allows validating certain results of the meta-analysis but also to determine the reliability of the expressive and emotional parameters.

5.1 Comparison method

For any feature in the Emilya analysis table (TABLE 6) a comparable feature can be found in the meta-analysis (TABLE 4), and for any emotion in the Emilya analysis table, one or more corresponding emotions can be found in the meta-analysis table.

We obtain the following correspondences using the emotion label matching method described in Section 3.5.1:

Emilya analysis	Meta-analysis
Neutral	→ Neutral
Anger	→ Anger
Anxiety	→ Anxiety
Joy	→ Happiness, Elation
Panic fear	→ Fear
Pride	→ Pride
Sadness	→ Sadness
Shame	→ Shame

For a given comparable feature f, an emotion of the meta-analysis table is said to be comparable if its corresponding cell is not empty and a correspondence exists with one of the emotions of the Emilya analysis table.

Let denote $\mathbf{X}_{f,meta}$ and $\mathbf{X}_{f,emilya}$ the value sets containing all comparable emotion values for a given feature f.

TABLE 6
Results of the Emilya dataset extraction: each cell contains an arithmetic mean and a standard deviation in parentheses. Some expressive features have been scaled for reading purposes, as specified in parentheses in the first column. All angles (such as shoulder flexion, scapula elevation or knee flexion amplitude, etc.) are expressed in degrees.

Expressive feature	Anger	Anxiety	Joy	Neutral	Panic fear	Pride	Sadness	Shame
cervical vertebrae flexion	7.884 (2.298)	8.084 (2.789)	7.325 (2.582)	7.404 (2.043)	7.015 (2.305)	7.211 (2.029)	10.440 (2.785)	9.644 (2.808)
cervical vertebrae lateral flexion	1.476 (0.893)	1.409 (1.031)	1.403 (0.942)	1.039 (0.806)	1.684 (1.146)	1.391 (0.877)	1.371 (1.063)	1.467 (1.280)
elbow flexion	46.514 (19.436)	52.600 (24.048)	49.967 (25.738)	38.060 (22.280)	54.831 (26.361)	41.419 (22.465)	48.354 (28.555)	47.493 (26.590)
elbow flexion amplitude	0.548 (0.377)	0.380 (0.365)	0.439 (0.323)	0.324 (0.309)	0.359 (0.324)	0.336 (0.268)	0.243 (0.254)	0.245 (0.229)
hip flexion	14.709 (28.000)	11.463 (27.418)	10.267 (24.217)	6.668 (25.238)	12.069 (24.958)	9.038 (25.088)	8.693 (27.878)	10.959 (27.376)
hip flexion amplitude	0.346 (0.541)	0.302 (0.543)	0.311 (0.515)	0.290 (0.516)	0.283 (0.481)	0.242 (0.443)	0.249 (0.519)	0.239 (0.442)
knee flexion	34.158 (29.289)	36.391 (30.963)	34.810 (27.010)	30.829 (31.344)	39.738 (28.174)	31.616 (29.749)	30.827 (31.434)	34.713 (32.250)
knee flexion amplitude	0.439 (0.461)	0.424 (0.446)	0.501 (0.457)	0.363 (0.425)	0.424 (0.439)	0.402 (0.460)	0.301 (0.408)	0.323 (0.383)
laban flow effort ($\times 0.01$)	-6.728 (3.918)	-4.474 (3.316)	-5.398 (3.401)	-2.372 (1.866)	-6.043 (4.606)	-2.994 (2.147)	-1.676 (1.408)	-1.990 (1.636)
laban flow shape	0.352 (0.106)	0.328 (0.106)	0.383 (0.141)	0.314 (0.093)	0.360 (0.116)	0.333 (0.106)	0.297 (0.092)	0.290 (0.098)
laban flow shape torso	0.083 (0.003)	0.083 (0.003)	0.083 (0.003)	0.082 (0.003)	0.083 (0.003)	0.083 (0.003)	0.082 (0.003)	0.082 (0.003)
laban space effort	1.098 (0.217)	1.071 (0.118)	1.087 (0.160)	1.118 (0.594)	1.075 (0.184)	1.097 (0.243)	1.075 (0.158)	1.064 (0.101)
laban space shape	0.167 (0.030)	0.148 (0.026)	0.158 (0.032)	0.151 (0.026)	0.155 (0.027)	0.151 (0.029)	0.142 (0.022)	0.136 (0.024)
laban space shape arm	0.216 (0.031)	0.196 (0.029)	0.211 (0.039)	0.202 (0.028)	0.206 (0.033)	0.204 (0.033)	0.185 (0.029)	0.183 (0.031)
laban time effort	-19.018 (10.518)	-13.025 (9.347)	-15.903 (9.452)	-7.416 (5.703)	-16.934 (12.368)	-9.377 (6.549)	-5.260 (4.433)	-6.124 (4.985)
laban time shape	0.061 (0.054)	0.058 (0.049)	0.052 (0.049)	0.042 (0.048)	0.062 (0.046)	0.038 (0.048)	0.062 (0.063)	0.064 (0.052)
laban weight effort ($\times 0.1$)	-0.005 (0.005)	-0.003 (0.003)	-0.005 (0.004)	-0.002 (0.002)	-0.005 (0.005)	-0.003 (0.002)	-0.001 (0.001)	-0.001 (0.002)
laban weight shape	0.325 (0.031)	0.322 (0.031)	0.329 (0.029)	0.325 (0.026)	0.322 (0.030)	0.330 (0.027)	0.320 (0.035)	0.320 (0.032)
laban weight shape torso	0.363 (0.017)	0.360 (0.019)	0.363 (0.018)	0.363 (0.017)	0.357 (0.019)	0.367 (0.016)	0.356 (0.022)	0.358 (0.020)
movement activity	0.467 (0.193)	0.344 (0.192)	0.457 (0.190)	0.300 (0.189)	0.390 (0.199)	0.357 (0.197)	0.236 (0.175)	0.240 (0.163)
movement activity arm	0.666 (0.213)	0.437 (0.254)	0.632 (0.225)	0.392 (0.213)	0.487 (0.274)	0.502 (0.227)	0.279 (0.205)	0.266 (0.211)
regularity	-23.457 (6.105)	-20.744 (7.120)	-22.216 (6.222)	-14.622 (5.560)	-23.577 (7.720)	-17.071 (5.261)	-12.864 (5.403)	-14.744 (5.937)
regularity head	-27.549 (9.591)	-23.828 (9.693)	-26.344 (10.799)	-16.820 (8.352)	-27.501 (10.685)	-19.625 (7.787)	-14.193 (7.540)	-16.306 (8.048)
regularity torso	-25.023 (9.966)	-20.749 (10.236)	-23.660 (10.025)	-13.589 (8.439)	-24.803 (11.935)	-16.474 (8.199)	-11.178 (7.387)	-12.984 (8.038)
scapula elevation	-7.259 (2.377)	-7.626 (2.163)	-7.508 (2.643)	-9.099 (2.039)	-6.500 (2.839)	-8.020 (1.778)	-8.544 (2.971)	-7.892 (2.735)
scapula medial rotation	1.745 (2.392)	1.444 (2.578)	2.126 (2.969)	2.577 (2.501)	1.373 (2.775)	2.621 (2.601)	1.590 (2.762)	1.238 (3.120)
scapula medial rotation amplitude	1.924 (1.370)	1.308 (1.071)	1.854 (1.524)	1.203 (1.585)	1.664 (1.441)	1.287 (1.249)	0.830 (1.022)	0.839 (0.813)
shoulder abduction	17.822 (10.306)	12.466 (10.222)	16.685 (13.526)	11.311 (7.615)	13.929 (10.108)	13.578 (11.831)	9.549 (10.007)	8.307 (9.261)
shoulder flexion	16.806 (14.886)	14.700 (13.665)	16.889 (15.545)	12.330 (10.403)	12.596 (14.367)	13.954 (13.549)	20.485 (15.830)	18.170 (12.636)
shoulder flexion amplitude	0.343 (0.326)	0.220 (0.205)	0.361 (0.379)	0.197 (0.220)	0.243 (0.241)	0.244 (0.217)	0.149 (0.163)	0.151 (0.240)
shoulder horizontal flexion	32.564 (27.385)	37.970 (30.287)	35.346 (29.061)	36.826 (31.965)	28.877 (36.216)	34.896 (32.565)	56.242 (24.306)	56.721 (29.756)
smoothness ($\times 0.01$)	-6.728 (3.918)	-4.474 (3.316)	-5.398 (3.401)	-2.372 (1.866)	-6.043 (4.606)	-2.994 (2.147)	-1.676 (1.408)	-1.990 (1.636)
smoothness head ($\times 0.1$)	-18.857 (11.508)	-11.115 (7.876)	-13.241 (7.919)	-5.322 (3.178)	-16.102 (11.505)	-7.280 (4.080)	-4.181 (2.907)	-5.241 (3.633)
smoothness torso ($\times 0.1$)	-22.370 (14.547)	-13.857 (10.242)	-18.688 (12.910)	-7.258 (5.690)	-21.378 (18.206)	-9.814 (7.461)	-5.553 (4.542)	-6.400 (4.952)
vertebrae flexion	12.853 (7.659)	14.621 (8.523)	11.905 (8.640)	12.304 (7.568)	16.736 (8.553)	8.390 (8.091)	15.369 (9.829)	15.744 (8.268)
vertebrae lateral flexion amplitude	0.029 (0.022)	0.020 (0.016)	0.028 (0.024)	0.016 (0.014)	0.024 (0.018)	0.024 (0.021)	0.015 (0.013)	0.016 (0.017)

The correlation coefficient between those two value sets is defined as follows:

$$\mathrm{corr}(\mathbf{X}_{f,meta}, \mathbf{X}_{f,emilya}) = \frac{\mathrm{cov}(\mathbf{X}_{f,meta}, \mathbf{X}_{f,emilya})}{\sigma(\mathbf{X}_{f,meta}) \,.\, \sigma(\mathbf{X}_{f,emilya})} \tag{74}$$

where cov is the covariance between two value sets $\mathbf{A}$ and $\mathbf{B}$ of a number of C correspondences and is defined as:

$$\mathrm{cov}(\mathbf{A}, \mathbf{B}) = \frac{1}{C} \sum_{i=1}^{C} (a_i - \bar{\mathbf{A}})(b_i - \bar{\mathbf{B}}) \tag{75}$$

5.2 Comparison results

The comparison between the results of the meta-analysis and the Emilya analysis are presented in Fig. 8. Each value in the spider graph is standardized as:

$$x'_i = \frac{x_i - \bar{\mathbf{X}}}{\sigma(\mathbf{X})}$$

where $\mathbf{X}$ is either $\mathbf{X}_{f,meta}$ or $\mathbf{X}_{f,emilya}$.

The extraction methods in the Emilya analysis have been carefully adjusted to match the sign of the descriptor definitions provided in the meta-analysis (see opposite notion in (15)). Therefore, only high positive correlations can be interpreted as high matching values. Looking at the resulting graphs, we determined empirically that a correlation coefficient higher than 0.8 can be considered as a "good" match between the analyses.

5.2.1 Notable results

First of all, none of the spider graphs has a circular shape, which confirms the hypothesis that all the presented descriptors are able to encode the diversity of the addressed emotions. This is even more obvious when considering a set of several descriptors.

19 out of 33 descriptors have a spider graph labelled as a "good" match ($\mathrm{corr} > 0.8$). 14 of them even have "excellent" correlation levels above 0.9. This means that for the majority of the expressive descriptors, the values from both meta-analysis and Emilya analysis have good reliability. We hope that the primary purpose of this study is therefore fulfilled.

Among the 14 "excellent" descriptors, 6 of them have a low level of noticeability (i.e not studied enough) in the meta-analysis: *shoulder flexion*, *knee flexion*, *movement activity arm*, *regularity torso*, *smoothness head* and *smoothness torso*. Our results suggest that these descriptors should be considered for future research.

By tracking value origins, we know that some of the meta-analysis spider graphs are only the result of the aggregation of studies based on perceptual assessments. Although further studies are needed to confirm this hypothesis, the high correlation levels for *smoothness head/torso*,

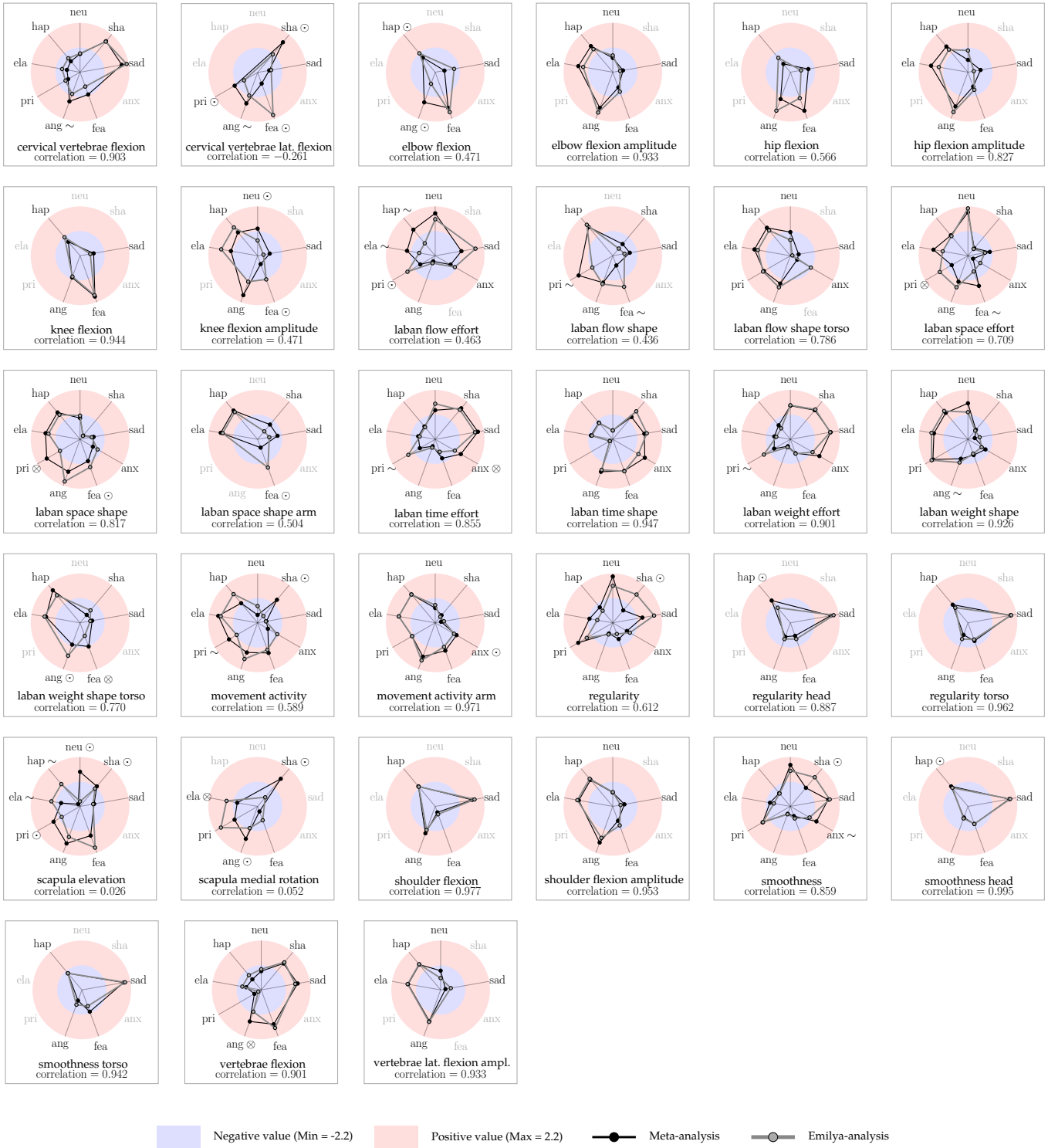


Fig. 8. Comparison between the values of the meta-analysis and the Emilya analysis. Each emotion whose values have opposite signs is followed by one of the three following symbols: $\sim$ indicates a meta-analysis cell which is not noticeable (unnoticeability level above one), $\odot$ indicates a meta-analysis cell containing only the value of a single experiment and $\otimes$ indicates all other cases (noticeable cells that mismatches with the Emilya analysis).

regularity head/torso and *movement activity arm* seem to indicate that we have found extraction methods in the Emilya analysis that fit the way humans perceive these notions. The case of *movement activity arm* is particularly noteworthy because the meta-analysis values come exclusively from a an experiment based on perceptual assessment conducted by Fourati et al. on the Emilya database [24]. Despite the use of two different methodologies on the same database, our results and those of Fourati et al. are more than 90% correlated.

The hypothesis formulated in Sections 3.4 and 3.7 that the *neutral* is rarely placed at zero is confirmed by the comparisons. We therefore recommend considering the neutral point as a regular emotion rather than a reference point in future experiments.

A last notable observation is that the spider graphs of the weight and time efforts are very similar, while they should illustrate two independent notions according to Laban's theory. Our comparisons suggest that *light* movements are always *sustained*, and *strong* movements are always *sudden*. However, Laban provides examples in [33] that contradict our results: one example among others is the action of *pressing* which correspond to a *strong* and *sustained* movement. The hypothesis which seems to us the most probable is that choreographed movements have much more diversity than movements of everyday life. This diversity is not represented in the meta-analysis and the Emilya dataset as it almost exclusively refers to everyday actions such as walking, knocking on a door, sitting down, etc. Our study reveals a high similarity between weight and time efforts for this kind of movements, but this does not mean that the two factors are equivalent in every situation.

5.2.2 Hypotheses for mismatching values

In order to explain the "weak" matching graphs (i.e. $\mathrm{corr} \leq 0.8$), it is necessary to look at the emotions whose values have opposite signs. There are several factors and hypotheses to explain these mismatches (42 mismatches out of 216 comparisons).

22 mismatches result from an Emilya analysis value which is compared with a meta-analysis cell containing only the value of a single experiment (represented with $\odot$ in Fig. 8). This can be problematic, especially if the original study uses the performances of only few actors [91] [84] [81], or if the author claimed they did not found significant differences between emotions for this value. Investigating the less studied emotions and descriptors would certainly be beneficial in future works.

14 mismatches occur when the meta-analysis cell has its unnoticeability strictly above one (see Section 3.7), which means that this cell is the mean of multiple experiment but is not noticeable (represented with $\sim$ in Fig. 8). As previous studies do not reach a consensus for this value, it is not surprising that the value extracted from the Emilya analysis could not be comparable.

The last 6 value mismatches do not have an obvious explanation (represented with $\otimes$ in Fig. 8). This is for example the case of pride which have three noticeable values in the meta-analysis (Laban space effort, Laban space shape and vertebrae flexion in TABLE 4), but two of them have does not match with the Emilya analysis (Laban space shape et Laban space effort). Our hypothesis is that an emotion can be performed in different manners, even with contradictory values for a specific descriptor. Tracy et al. [121] show that two expressions of pride can be identified either with arms raised or hands on hips, both corresponding to contradictory values for the Laban space shape.

Having one or two mismatches on a spider graph lowers significantly the correlation level. Such mismatches do not imply that the given descriptor is not relevant for all the emotions. A particularly noteworthy case concerns the expression of fear, which lowers the correlation levels of certain graphs (such as cervical vertebrae lateral flexion, Laban space shape arm, etc.). In many cases, the comparison for the expression of fear gives diametrically opposed results. Fear manifests itself either through an expanded posture (similar to surprise), or through a contracted body posture (protection against an external threat). These two archetypal body configurations are valid, but we cannot differentiate between these two cases using our methodology unless we use two distinct labels. We know that the Emilya dataset mainly contains contracted body postures for the performance of fear, but many studies in the meta-analysis do not provide this information, making impossible the usage of distinct labels.

The relevance of a descriptor or its calculation method can be questioned when the correlation rate is very low (scapula elevation and scapula medial rotation) or even negative (cervical vertebrae lateral flexion). The calculation method of these descriptors is probably not to blame since it is limited to simple angle measurements. For the scapulation elevation, one possible interpretation could be that this descriptor is relevant only for a subset of emotions: sadness, anger and fear. The spider graph shapes of the two other descriptors (cervical vertebrae lateral flexion and scapula medial rotation) could even lead to question their relevance, at least in the context of this study.

Our last observation concerns the weak matching graph obtained for the Laban flow effort. This result is disappointing, considering that this factor is highly referenced in the literature (Tables 2 and 5). The Laban flow effort is defined in the meta-analysis as a tense/bound movement progression on the negative pole and a loose/free movement progression on the positive pole (TABLE 2). As suggested by previous studies (see TABLE 5), the Laban flow effort is computed in the Emilya analysis using the opposite notion of jerk (66): discontinuous progression on the negative pole (high jerk, bound) and continuous progression on the positive pole (low jerk, free). The spider graph for the Laban flow effort tends to indicate that the calculation approach based only on the jerk does not reflect correctly this notion. All the other approaches based on "elementary" notions (curvature, velocity, acceleration, etc.) have also been computed without significantly improving the results. Some authors [116], [117] suggests that the Laban flow effort is in fact a combination of several of these elementary notions, and even that positive and negative poles are not described with the same notions. Laban effort components are complex notions anyhow, and further studies would be needed to determine extraction methods whose results better coincide with experiments based on perceptual assessments. Such investigation is out of the scope of this study, but the works

of Fdili Alaoui et al. [111] or Samadani et al. [85] could be good tracks to follow.

6 Conclusions & Future Work

The main purpose of this study was: (1) to establish a broad list of expressive descriptors responsible for the perception of emotions in human movements and (2) to quantify the contribution of each of these descriptors in a variety of emotions. To this end, we conducted a meta-analysis on other studies addressing this question. The results of this meta-analysis were compared with a computational approach on the Emilya dataset. The comparison reveals good matching levels between the two employed methodologies, which validates the majority of our quantified values. It is important to note that the Emilya dataset and the vast majority of studies in the meta-analysis are based on acted situations. Therefore, our results are necessarily related to this specific context. Further experimentation would be needed to determine whether the use of non-acted recordings produces less clear correlations.

The wide diversity of expressive descriptors we have extracted from the literature reveals how complex the question of bodily expression of emotion is. However, this study only focused on the question of movement and posture features. Additional work could be carried out to obtain quantified values for gestural descriptors, such as scratching the head, crossing the arms, etc. We would thus have the full range of criteria to analyze and synthesize emotions in human movements.

Our meta-analysis shows that basic emotions are relatively well covered in the literature. We do hope, however, that the missing entries revealed by our meta-analysis will encourage future research to explore more subtle and less documented emotions.

We also observed that the extraction of certain expressive descriptors produced results very similar to those obtained during experiments based on perceptual assessment. It would be interesting to deepen this approach in order to better model the perceptual schemes responsible for the identification of an emotion. Such perceptual schemes could be the building blocks of an artificial intelligence able to mimic the way a human perceives and expresses emotion in body motion. Our future research will focus on the exploitation of these results in the fields of automatic emotion recognition and animation synthesis.

Acknowledgments

This work was partly funded by an ANRT (http://www.anrt.asso.fr/en) CIFRE scholarship between the LIRIS research lab (CNRS) and the SpirOps company. The authors would like to thank Quentin Prevost and William Ledoux for their valuable comments and suggestions, Catherine Pelachaud who allowed us to use the Emilya database in the scope of this publication, and the anonymous reviewers for their helpful comments.

References

[1] R. A. Khan, A. Crenn, A. Meyer, and S. Bouakaz, "A novel database of children's spontaneous facial expressions (LIRIS-CSE)," *Image and Vision Computing*, vol. 83–84, pp. 61–69, Mar. 2019.

[2] A. Kleinsmith, P. R. De Silva, and N. Bianchi-Berthouze, "Cross-cultural differences in recognizing affect from body posture," *Interacting with Computers*, vol. 18, no. 6, pp. 1371–1389, Dec. 2006.

[3] G. Bijlstra, R. W. Holland, R. Dotsch, and D. H. J. Wigboldus, "Stereotypes and prejudice affect the recognition of emotional body postures," *Emotion*, vol. 19, no. 2, pp. 189–199, 2019.

[4] T. A. Olugbade, A. Singh, N. Bianchi-Berthouze, N. Marquardt, M. S. H. Aung, and W. Amanda C. De C, "How Can Affect Be Detected and Represented in Technological Support for Physical Rehabilitation?" *ACM Transactions on Computer-Human Interaction (TOCHI)*, Jan. 2019.

[5] J. Michalak, N. F. Troje, J. Fischer, P. Vollmar, T. Heidenreich, and D. Schulte, "Embodiment of Sadness and Depression—Gait Patterns Associated With Dysphoric Mood," *Psychosomatic Medicine*, vol. 71, no. 5, pp. 580–587, Jun. 2009.

[6] T. Randhavane, U. Bhattacharya, K. Kapsaskis, K. Gray, A. Bera, and D. Manocha, "Identifying Emotions from Walking using Affective and Deep Features," *arXiv:1906.11884 [cs]*, Jan. 2020.

[7] A. Aristidou, Q. Zeng, E. Stavrakis, K. Yin, D. Cohen-Or, Y. Chrysanthou, and B. Chen, "Emotion control of unstructured dance movements," in *Proceedings of the ACM SIGGRAPH / Eurographics Symposium on Computer Animation*, ser. SCA '17. New York, NY, USA: Association for Computing Machinery, Jul. 2017, pp. 1–10.

[8] K. Schuller, *The Biopolitics of Feeling: Race, Sex, and Science in the Nineteenth Century*. Duke University Press, Jan. 2018.

[9] S. Li and W. Deng, "Deep Facial Expression Recognition: A Survey," *IEEE Transactions on Affective Computing*, pp. 1–1, 2020.

[10] A. Kleinsmith and N. Bianchi-Berthouze, "Affective Body Expression Perception and Recognition: A Survey," *IEEE Transactions on Affective Computing*, vol. 4, no. 1, pp. 15–33, Jan. 2013.

[11] R. Lun and W. Zhao, "A Survey of Applications and Human Motion Recognition with Microsoft Kinect," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 29, no. 05, p. 1555008, Aug. 2015.

[12] I. Arun Faisal, T. Waluyo Purboyo, and A. Siswo Raharjo Ansori, "A Review of Accelerometer Sensor and Gyroscope Sensor in IMU Sensors on Motion Capture," *Journal of Engineering and Applied Sciences*, vol. 15, no. 3, pp. 826–829, Nov. 2019.

[13] U. Bhattacharya, T. Mittal, R. Chandra, T. Randhavane, A. Bera, and D. Manocha, "STEP: Spatial Temporal Graph Convolutional Networks for Emotion Perception from Gaits," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 02, pp. 1342–1350, Apr. 2020.

[14] A. Crenn, A. Meyer, H. Konik, R. A. Khan, and S. Bouakaz, "Generic Body Expression Recognition Based on Synthesis of Realistic Neutral Motion," *IEEE Access*, vol. 8, pp. 207 758–207 767, 2020.

[15] J. Huang and C. Pelachaud, "Expressive body animation pipeline for virtual agent," in *International Conference on Intelligent Virtual Agents*. Springer, 2012, pp. 355–362.

[16] Y. Ding, K. Prepin, J. Huang, C. Pelachaud, and T. Artières, "Laughter animation synthesis," in *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-Agent Systems*, 2014, pp. 773–780.

[17] M. Neff and C. Pelachaud, "Animation of Natural Virtual Characters," *IEEE Computer Graphics and Applications*, vol. 37, no. 4, pp. 14–16, 2017.

[18] T. Randhavane, A. Bera, K. Kapsaskis, R. Sheth, K. Gray, and D. Manocha, "EVA: Generating Emotional Behavior of Virtual Agents using Expressive Features of Gait and Gaze," in *ACM Symposium on Applied Perception 2019*, ser. SAP '19. New York, NY, USA: Association for Computing Machinery, Sep. 2019, pp. 1–10.

[19] T. Beinema, D. Davison, D. Reidsma, O. Banos, M. Bruijnes, B. Donval, Á. F. Valero, D. Heylen, D. Hofs, G. Huizing, R. B. Kantharaju, R. Klaassen, J. Kolkmeier, K. Konsolakis, A. Pease, C. Pelachaud, D. Simonetti, M. Snaith, V. Traver, J. van Loon, J. Visser, M. Weusthof, F. Yunus, H. Hermens, and H. op den Akker, "Agents United: An Open Platform for Multi-Agent Conversational Systems," in *Proceedings of the 21st ACM International Conference on Intelligent Virtual Agents*. New York, NY, USA: Association for Computing Machinery, Sep. 2021, pp. 17–24.

[20] D. Chi, M. Costa, L. Zhao, and N. Badler, "The EMOTE model for effort and shape," in *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, ser. SIGGRAPH '00.

USA: ACM Press/Addison-Wesley Publishing Co., Jul. 2000, pp. 173–182.

[21] R. S. Ziegelmaier, W. Correia, J. M. Teixeira, and F. P. M. Simões, "Laban Movement Analysis applied to Human-Computer Interaction," in *2020 22nd Symposium on Virtual and Augmented Reality (SVR)*, Nov. 2020, pp. 30–34.

[22] K. Takahashi, M. Hosokawa, and M. Hashimoto, "Remarks on designing of emotional movement for simple communication robot," in *2010 IEEE International Conference on Industrial Technology*, Mar. 2010, pp. 585–590.

[23] N. Fourati and C. Pelachaud, "Emilya: Emotional body expression in daily actions database," in *9th International Conference on Language Resources and Evaluation (LREC 2014)*, Reykjavik, Iceland, May 2014.

[24] N. Fourati and C. Pelachaud, "Perception of Emotions and Body Movement in the Emilya Database," *IEEE Transactions on Affective Computing*, vol. 9, no. 1, pp. 90–101, Jan. 2018.

[25] J. K. Burgoon, V. Manusov, and L. K. Guerrero, *Nonverbal Communication*, 2nd ed. New York: Routledge, Sep. 2021.

[26] F. BAMBAEEROO and N. SHOKRPOUR, "The impact of the teachers' non-verbal communication on success in teaching," *Journal of Advances in Medical Education & Professionalism*, vol. 5, no. 2, pp. 51–59, Apr. 2017.

[27] A. J. Hale, J. Freed, D. Ricotta, G. Farris, and C. C. Smith, "Twelve tips for effective body language for medical educators," *Medical Teacher*, vol. 39, no. 9, pp. 914–919, Sep. 2017.

[28] P. Furley and G. Schweizer, "Body Language in Sport," in *Handbook of Sport Psychology*. John Wiley & Sons, Ltd, 2020, ch. 59, pp. 1201–1219.

[29] M. Schlögl and C. A. Jones, "Maintaining Our Humanity Through the Mask: Mindful Communication During COVID-19," *Journal of the American Geriatrics Society*, vol. 68, no. 5, pp. E12–E13, May 2020.

[30] A. W. Siegman and S. Feldstein, Eds., *Nonverbal Behavior and Communication*, 2nd ed. New York: Psychology Press, Dec. 2013.

[31] E. De Stefani and D. De Marco, "Language, Gesture, and Emotional Communication: An Embodied View of Social Interaction," *Frontiers in Psychology*, vol. 10, p. 2063, 2019.

[32] R. von Laban and F. C. Lawrence, *Effort*. Macdonald & Evans, 1947.

[33] R. von Laban, *The Mastery of Movement*. Macdonald and Evans, 1980.

[34] F. Thomas and O. Johnston, *Disney Animation: The Illusion of Life*. Abbeville Press, 1981.

[35] A. Crenn, R. A. Khan, A. Meyer, and S. Bouakaz, "Body expression recognition from animated 3D skeleton," in *2016 International Conference on 3D Imaging (IC3D)*. Liège, Belgium: IEEE, Dec. 2016, pp. 1–7.

[36] T. Tao, X. Zhan, Z. Chen, and M. van de Panne, "Style-ERD: Responsive and Coherent Online Motion Style Transfer," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 6593–6603.

[37] D. Holden, I. Habibie, I. Kusajima, and T. Komura, "Fast Neural Style Transfer for Motion Data," *IEEE Computer Graphics and Applications*, vol. 37, no. 4, pp. 42–49, 2017.

[38] S. Ribet, H. Wannous, and J.-P. Vandeborre, "Survey on Style in 3D Human Body Motion: Taxonomy, Data, Recognition and its Applications," *IEEE Transactions on Affective Computing*, pp. 1–1, 2019.

[39] S. Starke, H. Zhang, T. Komura, and J. Saito, "Neural state machine for character-scene interactions," *ACM Transactions on Graphics*, vol. 38, no. 6, pp. 209:1–209:14, Nov. 2019.

[40] S. Starke, Y. Zhao, T. Komura, and K. Zaman, "Local motion phases for learning multi-contact character movements," *ACM Transactions on Graphics*, vol. 39, no. 4, pp. 54:54:1–54:54:13, Jul. 2020.

[41] S. Xia, C. Wang, J. Chai, and J. Hodgins, "Realtime style transfer for unlabeled heterogeneous human motion," *ACM Transactions on Graphics (TOG)*, vol. 34, no. 4, p. 119, 2015.

[42] M. E. Yumer and N. J. Mitra, "Spectral style transfer for human motion between independent actions," Jul. 2016.

[43] K. Aberman, Y. Weng, D. Lischinski, D. Cohen-Or, and B. Chen, "Unpaired motion style transfer from video to animation," *ACM Transactions on Graphics*, vol. 39, no. 4, pp. 64:64:1–64:64:12, Jul. 2020.

[44] F. Durupinar, M. Kapadia, S. Deutsch, M. Neff, and N. I. Badler, "Perform: Perceptual approach for adding ocean personality to human motion using laban movement analysis," *ACM Transactions on Graphics (TOG)*, vol. 36, no. 1, p. 6, 2017.

[45] C. Beyan, S. Karumuri, G. Volpe, A. Camurri, and R. Niewiadomski, "Modeling Multiple Temporal Scales of Full-body Movements for Emotion Classification," *IEEE Transactions on Affective Computing*, pp. 1–1, 2021.

[46] H. Gunes, B. Schuller, M. Pantic, and R. Cowie, "Emotion representation, analysis and synthesis in continuous space: A survey," in *Face and Gesture 2011*. Santa Barbara, CA, USA: IEEE, Mar. 2011, pp. 827–834.

[47] R. A. Calvo and S. D'Mello, "Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications," *IEEE Transactions on Affective Computing*, vol. 1, no. 1, pp. 18–37, Jan. 2010.

[48] Z. Zeng, M. Pantic, G. I. Roisman, and T. S. Huang, "A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 1, pp. 39–58, Jan. 2009.

[49] M. Karg, A.-A. Samadani, R. Gorbet, K. Kuhnlenz, J. Hoey, and D. Kulic, "Body Movements for Affective Expression: A Survey of Automatic Recognition and Generation," *IEEE Transactions on Affective Computing*, vol. 4, no. 4, pp. 341–359, Oct. 2013.

[50] H. Zacharatos, C. Gatzoulis, and Y. L. Chrysanthou, "Automatic Emotion Recognition Based on Body Movement Analysis: A Survey," *IEEE Computer Graphics and Applications*, vol. 34, no. 6, pp. 35–45, Nov. 2014.

[51] B. Stephens-Fripp, F. Naghdy, D. Stirling, and G. Naghdy, "Automatic Affect Perception Based on Body Gait and Posture: A Survey," *International Journal of Social Robotics*, vol. 9, no. 5, pp. 617–641, Nov. 2017.

[52] C. Larboulette and S. Gibet, "A Review of Computable Expressive Descriptors of Human Motion," *Proceedings of the 2nd International Workshop on Movement and Computing (MOCO'15)*, Aug. 2015.

[53] F. Noroozi, D. Kaminska, C. Corneanu, T. Sapinski, S. Escalera, and G. Anbarjafari, "Survey on emotional body gesture recognition," *IEEE Transactions on Affective Computing*, 2018.

[54] C. Rose, M. Cohen, and B. Bodenheimer, "Verbs and adverbs: Multidimensional motion interpolation," *IEEE Computer Graphics and Applications*, vol. 18, no. 5, pp. 32–40, Sep. 1998.

[55] F. Deligianni, Y. Guo, and G.-Z. Yang, "From Emotions to Mood Disorders: A Survey on Gait Analysis Methodology," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 6, pp. 2302–2316, Nov. 2019.

[56] D. Moher, A. Liberati, J. Tetzlaff, and D. G. Altman, "Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement," *Annals of Internal Medicine*, vol. 151, no. 4, pp. 264–269, Aug. 2009.

[57] A. Dessai and H. Virani, "Emotion Classification using Physiological Signals: A Recent Survey," in *2022 IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems (SPICES)*, vol. 1, Mar. 2022, pp. 333–338.

[58] P. Santhiya and S. Chitrakala, "A Survey on Emotion Recognition from EEG Signals: Approaches, Techniques & Challenges," in *2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN)*, Mar. 2019, pp. 1–6.

[59] M. Moolchandani, S. Dwivedi, S. Nigam, and K. Gupta, "A survey on: Facial Emotion Recognition and Classification," in *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*, Apr. 2021, pp. 1677–1686.

[60] S. Latif, R. Rana, S. Khalifa, R. Jurdak, J. Qadir, and B. W. Schuller, "Survey of Deep Representation Learning for Speech Emotion Recognition," *IEEE Transactions on Affective Computing*, pp. 1–1, 2021.

[61] N. M. Khair, H. Muthusamy, S. Yaacob, and S. N. Basah, "Recognition of Emotions in Gait Patterns Using Discrete Wavelet Transform," *International Journal of Biomedical and Clinical Engineering (IJBCE)*, vol. 1, no. 1, pp. 86–93, Jan. 2012.

[62] M. Daoudi, S. Berretti, P. Pala, Y. Delevoye, and A. Del Bimbo, "Emotion Recognition by Body Movement Representation on the Manifold of Symmetric Positive Definite Matrices," in *Image Analysis and Processing - ICIAP 2017*, ser. Lecture Notes in Computer Science, S. Battiato, G. Gallo, R. Schettini, and F. Stanco, Eds. Cham: Springer International Publishing, 2017, pp. 550–560.

[63] P. Barros, D. Jirak, C. Weber, and S. Wermter, "Multimodal emotional state recognition using sequence-dependent deep hi-

erarchical features," *Neural Networks*, vol. 72, pp. 140–151, Dec. 2015.

[64] K. R. Scherer, "What are emotions? And how can they be measured?" *Social Science Information*, vol. 44, no. 4, pp. 695–729, Dec. 2005.

[65] R. Cowie, N. Sussman, and A. Ben-Ze'ev, "Emotion: Concepts and Definitions," in *Emotion-Oriented Systems: The Humaine Handbook*, ser. Cognitive Technologies, R. Cowie, C. Pelachaud, and P. Petta, Eds. Berlin, Heidelberg: Springer, 2011, pp. 9–30.

[66] G. Gilam, J. J. Gross, T. D. Wager, F. J. Keefe, and S. C. Mackey, "What Is the Relationship between Pain and Emotion? Bridging Constructs and Communities," *Neuron*, vol. 107, no. 1, pp. 17–21, Jul. 2020.

[67] C. L. Bethel and R. R. Murphy, "Survey of Non-facial/Non-verbal Affective Expressions for Appearance-Constrained Robots," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 38, no. 1, pp. 83–92, Jan. 2008.

[68] C. Vinola and K. Vimaladevi, "A Survey on Human Emotion Recognition Approaches, Databases and Applications," *ELCVIA : Electronic Letters on Computer Vision and Image Analysis*, pp. 00 024–44, 2015.

[69] M. Imani and G. A. Montazer, "A survey of emotion recognition methods with emphasis on E-Learning environments," *Journal of Network and Computer Applications*, vol. 147, p. 102423, Dec. 2019.

[70] P. A. Wilson and B. Lewandowska-Tomaszczyk, "Affective Robotics: Modelling and Testing Cultural Prototypes," *Cognitive Computation*, vol. 6, no. 4, pp. 814–840, Dec. 2014.

[71] H. Gunes and M. Piccardi, "Observer annotation of affective display and evaluation of expressivity: Face vs. face-and-body," in *Proceedings of the HCSNet Workshop on Use of Vision in Human-Computer Interaction - Volume 56*, ser. VisHCI '06. Canberra, Australia: Australian Computer Society, Inc., Nov. 2006, pp. 35–42.

[72] S. Piana, A. Stagliano, A. Camurri, and F. Odone, "A set of full-body movement features for emotion recognition to help children affected by autism spectrum condition," in *IDGEI International Workshop*, 2013.

[73] H. Zacharatos, C. Gatzoulis, Y. Chrysanthou, and A. Aristidou, "Emotion Recognition for Exergames using Laban Movement Analysis," in *Proceedings of Motion on Games*, ser. MIG '13. Dublin 2, Ireland: Association for Computing Machinery, Nov. 2013, pp. 61–66.

[74] C. Busso, Z. Deng, M. Grimm, U. Neumann, and S. Narayanan, "Rigid Head Motion in Expressive Speech Animation: Analysis and Synthesis," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15, no. 3, pp. 1075–1086, Mar. 2007.

[75] A. Kleinsmith, N. Bianchi-Berthouze, and A. Steed, "Automatic Recognition of Non-Acted Affective Postures," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 41, no. 4, pp. 1027–1038, Aug. 2011.

[76] C. L. Roether, L. Omlor, A. Christensen, and M. A. Giese, "Critical features for the perception of emotion from gait," *Journal of Vision*, vol. 9, no. 6, pp. 15–15, Jun. 2009.

[77] F. E. Pollick, H. M. Paterson, A. Bruderlin, and A. J. Sanford, "Perceiving affect from arm movement," *Cognition*, vol. 82, no. 2, pp. B51–B61, Dec. 2001.

[78] D. Glowinski, N. Dael, A. Camurri, G. Volpe, M. Mortillaro, and K. Scherer, "Toward a Minimal Representation of Affective Gestures," *IEEE Transactions on Affective Computing*, vol. 2, no. 2, pp. 106–118, Apr. 2011.

[79] W. T. James, "A Study of the Expression of Bodily Posture," *The Journal of General Psychology*, vol. 7, no. 2, pp. 405–437, Oct. 1932.

[80] H. G. Wallbott, "Bodily expression of emotion," *European Journal of Social Psychology*, vol. 28, no. 6, pp. 879–896, 1998.

[81] M. de Meijer, "The contribution of general features of body movement to the attribution of emotions," *Journal of Nonverbal Behavior*, vol. 13, no. 4, pp. 247–268, Dec. 1989.

[82] H. G. Wallbott and K. R. Scherer, "Cues and channels in emotion recognition," *Journal of Personality and Social Psychology*, vol. 51, no. 4, pp. 690–699, 1986.

[83] J. M. Montepare, S. B. Goldstein, and A. Clausen, "The identification of emotions from gait information," *Journal of Nonverbal Behavior*, vol. 11, no. 1, pp. 33–42, Mar. 1987.

[84] J. Montepare, E. Koff, D. Zaitchik, and M. Albert, "The Use of Body Movements and Gestures as Cues to Emotions in Younger and Older Adults," *Journal of Nonverbal Behavior*, vol. 23, no. 2, pp. 133–152, Jun. 1999.

[85] A.-A. Samadani, S. Burton, R. Gorbet, and D. Kulic, "Laban Effort and Shape Analysis of Affective Hand and Arm Movements," in *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, Sep. 2013, pp. 343–348.

[86] G. E. Kang and M. M. Gross, "The effect of emotion on movement smoothness during gait in healthy young adults," *Journal of Biomechanics*, vol. 49, no. 16, pp. 4022–4027, Dec. 2016.

[87] L. Omlor and M. A. Giese, "Extraction of spatio-temporal primitives of emotional body expressions," *Neurocomputing*, vol. 70, no. 10, pp. 1938–1942, Jun. 2007.

[88] M. M. Gross, E. A. Crane, and B. L. Fredrickson, "Effort-Shape and kinematic assessment of bodily expression of emotion during gait," *Human Movement Science*, vol. 31, no. 1, pp. 202–221, Feb. 2012.

[89] M. M. Gross, E. A. Crane, and B. L. Fredrickson, "Methodology for Assessing Bodily Expression of Emotion," *Journal of Nonverbal Behavior*, vol. 34, no. 4, pp. 223–248, Dec. 2010.

[90] S. Brownlow, A. R. Dixon, C. A. Egbert, and R. D. Radcliffe, "Perception of movement and dancer characteristics from point-light displays of dance," *The Psychological Record*, vol. 47, no. 3, pp. 411–422, Jul. 1997.

[91] S. Dahl and A. Friberg, "Visual Perception of Expressiveness in Musicians' Body Movements," *Music Perception: An Interdisciplinary Journal*, vol. 24, no. 5, pp. 433–454, Jun. 2007.

[92] M. Sawada, K. Suda, and M. Ishii, "Expression of Emotions in Dance: Relation between Arm Movement Characteristics and Emotion," *Perceptual and Motor Skills*, vol. 97, no. 3, pp. 697–708, Dec. 2003.

[93] M. Masuda, S. Kato, and H. Itoh, "Emotion Detection from Body Motion of Human Form Robot Based on Laban Movement Analysis," in *Principles of Practice in Multi-Agent Systems*, ser. Lecture Notes in Computer Science, J.-J. Yang, M. Yokoo, T. Ito, Z. Jin, and P. Scerri, Eds. Berlin, Heidelberg: Springer, 2009, pp. 322–334.

[94] T. NAKATA, "Analysis of Impression of Robot Bodily Expression," *Journal of Robotics and Mechatronics*, vol. 14, no. 1, pp. 27–36, 2002.

[95] A. Kapandji, *Anatomie Fonctionnelle. Volume 1. Membre Supérieur.*, 7th ed. Maloine, 2018, vol. 1.

[96] A. Kapandji, *Anatomie Fonctionnelle. Volume 2. Membre Inférieur.*, 7th ed. Maloine, 2018, vol. 2.

[97] A. Kapandji, *Anatomie Fonctionnelle. Volume 3. Tête et Rachis.*, 7th ed. Maloine, 2019, vol. 3.

[98] L. Bishko, "Animation principles and Laban movement analysis: Movement frameworks for creating empathic character performances," in *Nonverbal Communication in Virtual Worlds: Understanding and Designing Expressive Characters*. Pittsburgh, PA, USA: ETC Press, Jan. 2014, pp. 177–203.

[99] L. Bishko, "Empathy in virtual worlds: Making characters believable with Laban movement analysis," in *Nonverbal Communication in Virtual Worlds: Understanding and Designing Expressive Characters*. Pittsburgh, PA, USA: ETC Press, Jan. 2014, pp. 223–234.

[100] L. Bishko, "Our empathic experience of believable characters," in *Nonverbal Communication in Virtual Worlds: Understanding and Designing Expressive Characters*. Pittsburgh, PA, USA: ETC Press, Jan. 2014, pp. 47–59.

[101] P. Ekman and D. Cordaro, "What is Meant by Calling Emotions Basic," *Emotion Review*, vol. 3, no. 4, pp. 364–370, Oct. 2011.

[102] D. Glowinski, M. Mortillaro, K. Scherer, N. Dael, G. Volpe, and A. Camurri, "Towards a minimal representation of affective gestures," in *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, 2015, pp. 498–504.

[103] J. Russell, "A circumplex model of affect," *Journal of Personality and Social Psychology*, vol. 39, no. 6, pp. 1161–1178, Dec. 1980.

[104] A. Kleinsmith and N. Bianchi-Berthouze, "Recognizing Affective Dimensions from Body Posture," in *Affective Computing and Intelligent Interaction*, ser. Lecture Notes in Computer Science, A. C. R. Paiva, R. Prada, and R. W. Picard, Eds. Berlin, Heidelberg: Springer, 2007, pp. 48–58.

[105] M. Meredith and S. Maddock, "Motion capture file formats explained," *Department of Computer Science, University of Sheffield*, vol. 211, pp. 241–244, 2001.

[106] P. de Leva, "Adjustments to Zatsiorsky-Seluyanov's segment inertia parameters," *Journal of Biomechanics*, vol. 29, no. 9, pp. 1223–1230, Sep. 1996.

[107] D. A. Winter, *Biomechanics and Motor Control of Human Movement*. John Wiley & Sons, Oct. 2009.

[108] M. L. Latash and V. M. Zatsiorsky, *Biomechanics and Motor Control: Defining Central Concepts*. Elsevier Inc., Oct. 2015.

[109] D. Weinland, R. Ronfard, and E. Boyer, "A survey of vision-based methods for action representation, segmentation and recognition," *Computer Vision and Image Understanding*, vol. 115, no. 2, pp. 224–241, Feb. 2011.

[110] R. T. Arn, P. Narayana, T. Emerson, B. A. Draper, M. Kirby, and C. Peterson, "Motion Segmentation via Generalized Curvatures," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 12, pp. 2919–2932, Dec. 2019.

[111] S. Fdili Alaoui, J. Françoise, T. Schiphorst, K. Studd, and F. Bevilacqua, "Seeing, Sensing and Recognizing Laban Movement Qualities," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, ser. CHI '17. New York, NY, USA: Association for Computing Machinery, May 2017, pp. 4009–4020.

[112] S. Fdili Alaoui, F. Bevilacqua, B. Bermudez Pascual, and C. Jacquemin, "Dance interaction with physical model visuals based on movement qualities," *International Journal of Arts and Technology*, vol. 6, no. 4, pp. 357–387, Jan. 2013.

[113] A. Truong, "Analyse du contenu expressif des gestes corporels," Theses, Institut National des Télécommunications, Sep. 2016.

[114] M. Kapadia, I.-k. Chiang, T. Thomas, N. I. Badler, and J. T. Kider, "Efficient motion retrieval in large motion databases," in *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*, ser. I3D '13. Orlando, Florida: Association for Computing Machinery, Mar. 2013, pp. 19–28.

[115] T. Giraud, F. Focone, V. Demulier, J. C. Martin, and B. Isableu, "Perception of Emotion and Personality through Full-Body Movement Qualities: A Sport Coach Case Study," *ACM Transactions on Applied Perception*, vol. 13, no. 1, pp. 2:1–2:27, Oct. 2015.

[116] L. Santos, J. A. Prado, and J. Dias, "Human Robot interaction studies on laban human movement analysis and dynamic background segmentation," in *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct. 2009, pp. 4984–4989.

[117] G. Volpe, "Computational models of expressive gesture in multimedia systems," Ph.D. dissertation, University of Genova, 2003.

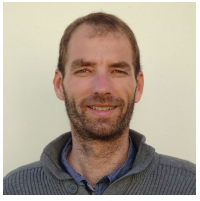
[118] E. Van Dyck, P.-J. Maes, J. Hargreaves, M. Lesaffre, and M. Leman, "Expressing Induced Emotions Through Free Dance Movement," *Journal of Nonverbal Behavior*, vol. 37, no. 3, pp. 175–190, Sep. 2013.

[119] H. M. Mentis and C. Johansson, "Seeing movement qualities," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '13. New York, NY, USA: Association for Computing Machinery, Apr. 2013, pp. 3375–3384.

[120] K. Hachimura, K. Takashina, and M. Yoshimura, "Analysis and evaluation of dancing movement based on LMA," in *ROMAN 2005. IEEE International Workshop on Robot and Human Interactive Communication, 2005.*, Aug. 2005, pp. 294–299.

[121] J. L. Tracy and R. W. Robins, "Show Your Pride: Evidence for a Discrete Emotion Expression," *Psychological Science*, vol. 15, no. 3, pp. 194–197, Mar. 2004.

Mehdi-Antoine Mahfoudi received the engineering degree in computer sciences from École d'ingénieur IMAC and the master's degree in computer imaging from Université Gustave Eiffel in 2016. He is a PhD candidate at the University Claude Bernard Lyon 1 with the LIRIS research lab and the SpirOps company. His research interests are focused on the analysis and synthesis of expressivity in human movements.



Alexandre Meyer received the Ph.D. degree in computer science from Université Grenoble 1, France, in 2001. From 2002 to 2003, he was a Postdoctoral Fellow with the University College London. Since 2004, he has been working as an Associate Professor with Université Claude Bernard Lyon 1, France, and a member of the LIRIS research lab. His current research concerns computer animation and computer vision of characters.

Thibaut Gaudin received his Master Degree of Engineering from the French Graduate Engineering School ESIEA in 2000. He has a French Master degree in Simulation and Virtual Reality from the French Engineer School Arts et Métiers ParisTech in 2001. Specialized in procedural character animation, he worked as a R&D Engineer and as a freelance for different research lab and companies. Since 2009, he is a R&D Engineer for the French company SpirOps.

Axel Buendia received a PhD in artificial intelligence in 2003. He is one of the founders of SpirOps, a company which provides since 2004 advanced AI services and tools. He has worked on several important collaborative research projects in different areas such as video games, robotics, urban simulation. Associate Professor at Cnam since 2008, he is teaching Artificial Intelligence at Cnam-Enjmin. In 2020, he is nominated full Professor at Cnam of the Interactive Digital Media chair. His researches focus on decisional systems.

Saida Bouakaz received the Ph.D. degree from Joseph Fourier University, Grenoble, France. She is currently a Full Professor with the Department of Computer Science, Université Claude Bernard Lyon 1, France. Her research interests include computer vision and graphics including motion capture and analysis, gesture recognition, and facial animation.