COMPUTER ANIMATION AND VIRTUAL WORLDS
*Comp. Anim. Virtual Worlds* 2016; **27**:311–320
Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/cav.1714

SPECIAL ISSUE PAPER

# Continuous body emotion recognition system during theater performances

Simon Senecal[1]*, Louis Cuel[1], Andreas Aristidou[2] and Nadia Magnenat-Thalmann[1]

[1] MIRALab, University of Geneva, Geneva, Switzerland
[2] University of Cyprus, Nicosia, Cyprus

## ABSTRACT

Understanding emotional human behavior in its multimodal and continuous aspect is necessary for studying human machine interaction and creating constituent social agents. As a first step, we propose a system for continuous emotional behavior recognition expressed by people during communication based on their gesture and their whole body dynamical motion. The features used to classify the motion are inspired by the Laban Movement Analysis entities [11] and are mapped onto the well-known Russell Circumplex Model [4]. We choose a specific case study that corresponds to an ideal case of multimodal behavior that emphasizes the body motion expression: theater performance. Using a trained neural network and annotated data, our system is able to describe the motion behavior as trajectories on the Russell Circumplex Model diagram during theater performances over time. This work contributes to the understanding of human behavior and expression and is a first step through a complete continuous emotion recognition system whose next step will be adding facial expressions. Copyright © 2016 John Wiley & Sons, Ltd.

**KEYWORDS**
emotion recognition; Laban Movement Analysis; behavior; motion capture; animation; interaction; nonverbal communication

***Correspondence**
Simon Senecal, MIRALab, University of Geneva, Geneva, Switzerland.
E-mail: senecal@miralab.ch

## 1. INTRODUCTION

The study of human multimodal emotional behavior can lead to the creation of virtual characters with different moods, emotional states, or personalities. In order to create social emotional agents and allow them to interact, it is necessary to understand the emotion expressed through the various communication channels, such as body motion, facial expression, and voice. A number of studies on emotion recognition, across different research domains, have been published. The majority of the first studies were focused on facial expression. Inspired by the work of psychologists and behaviorists such as Ekman [1] during the 1970s, systems that allow discrete recognition of emotion were developed and even integrated in some application like with interactive robots. More recently, emotion recognition from body motion was possible owing to the recent advances in motion capture systems. Using visual pattern recognition at the origin, new techniques using statistical analysis and machine learning give the possibility to consider the dynamics of the body and not only static postures or gestures. Their main difficulty is to overcome the fact that a large number of parameters may influence the recognition as the classification method and the quality and quantity of the data. However, very high recognition rate can be obtained from specific sets of data.

In order to achieve a good simulation for the body language, a simple but efficient body movement description system is needed: the Laban Movement Analysis (LMA) system fulfills these requirements. The LMA system draws on Rudolph Laban's theories and allows to describe, interpret, and document human movements. It is a multidisciplinary system, incorporating contributions from psychology, kinesiology, and anatomy. It is one of the most used languages to analyze human movement and has been regularly used to document and describe choreographies and dance over the last century.

The relation between movement and emotion has been studied extensively in psychology. Several emotion models were developed over the last 30 years, but one of the most used in computer applications is the continuous Russell Circumplex Model (RCM), which maps the different emotions following two dimensions: the *pleasure/displeasure* value and the *activation/deactivation* value. Coming from a long comparative work and experimentation, its classification allows the emotion to be placed on a 2D diagram for visualization.

Taking into consideration the broad range of classification criteria for emotion recognition, it is difficult to imagine a system that can adapt to any possible situation, and/or for different performers. Indeed, looking at the body movements, the meaning may vary significantly regarding the characteristic of the person itself: internal moods, social context, stress, culture, energy, and so on. Among other possibilities of body expression situation, theater performance is a good compromise between natural and clear expression, which is closely related to emotion. In this work, we aim at recognizing the emotional state of theater actors by mapping their motion, with regard to the expressed emotion, onto the Russell bi-dimensional space. This can be considered as a first step over a complete emotion recognition system. We integrate neural networks, because of their ability to approximate multivariate functions, to form the mapping. The neural networks are trained using the attribution of an emotion to a performance and the corresponding measured LMA features. It is then possible to have a correspondence with LMA features and coordinates of the emotion that vary in a continuous manner, allowing us to study and visualize emotion's dynamic.

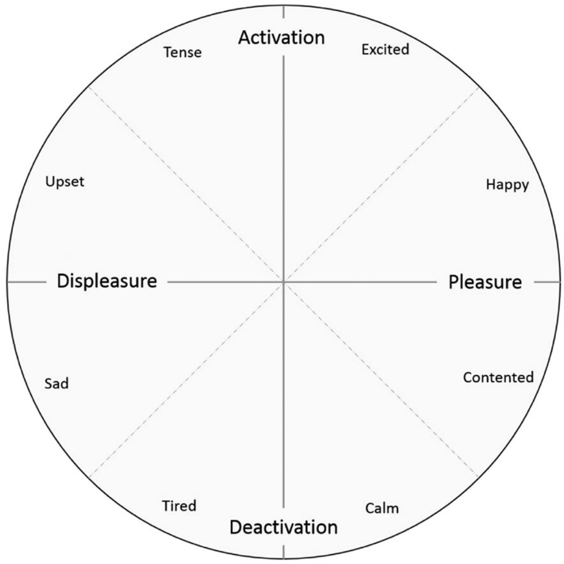


**Figure 1.** The Russell bi-dimensional model of emotion from 2003. The two axes correspond to activation or *intensity* and *pleasure* or valence. The sub-division of this space shows main sub-categories, but it remains bi-dimensional.

## 2. PREVIOUS WORK

Emotion recognition and automatic emotion recognition is a domain started mainly by a psychologist, which begins with facial analysis based on the Facial Action Coding System, which is as an emotion description language, from the studies of Ekman [1]. Much software were released, such as [2,3], that used various technologies (image processing, depth sensor) to identify the emotion of a character in real time. Although the quality of facial emotion recognition is quite high and allows to release commercial application, we do not see on the market emotion analysis based on body movement.

From the 1950s, several discrete models of emotion have been developed. Then, the work of James A. Russell and other psychologists during the 1970s started a new family of continuous model that classify the emotion onto few dimensions. A bi-dimensional proposed by Russell [4] or tri-dimensionals made by Mehrabian [5] have been used more recently by the artificial intelligence and cognitive science community. In addition, the Ortony *et al.* [6] model is an attempt from the cogniticians' angle to have a representation of emotion following fewer dimension. Using experimental results and dimension reduction, Russell proposed to classify emotions following two dimensions: (i) the *activation* dimension corresponds to the energy embedded into the expression. It varies from high intensity to low intensity. (ii) The *pleasure* dimension is related to the pleasure or nonpleasure you feel during the expression (obviously *sad* has a negative pleasure). Then he placed the obtain diagram points representing emotions (for example, *happy* has a high *activation* and is positive *pleasure*). More recently, a new paper by Russell [7] shows an updated version of his previous model (Figure 1) by separating the emotion space into sub-categories. These models are well suited for computation analysis as it is the result of comparative works and experimentation and reduce the possible variable in two dimensions that allow better visualization.

Body expression is also important within the expression of emotion as explained by the early work of Stanisvlaski in his books [8,9], showing the importance of emotion expression during theater performances. A more recent study by Kipp *et al.* [10] of emotion in old movies shows that universal association of gesture handedness with association of *pleasure* and *activation* as well as theater data is well suited for analysis because of the wide range of emotions displayed and emotion inherent of the actor's message.

To understand the body ways of expression, the work of Laban [11] proposed a body motion description system that is originally focused on dance. This LMA system was continued and developed further by Bartenieff [12] and other scientists. LMA has been studied by the computer graphics community for over a decade, and a collection of different features and measurements has been proposed to extract its components. It has been integrated to several body emotion recognition systems. The EMOTE system, introduced by Chi *et al.* [13], synthesizes gesture, for motion parameterization and expression, based on the LMA effort quality. Later, Zhao and Badler [14] used the EMOTE system and a neural network for gesture animation. LMA has been successfully used in many applications of computer graphics, such as motion analysis [15,16] or motion retrieval [17], all with a common denominator, the attempt to extract the style of motion. Santos and Dias [18] presented a tool to describe basic human behavior patterns using LMA, while

 *Comp. Anim. Virtual Worlds* 2016; **27**:311–320 © 2016 John Wiley & Sons, Ltd.
DOI: 10.1002/cav

others employed a subset of LMA components to explain bodily expressions of a human-form robot [19–21]. The LMA concept has been also utilized to quantify the expressive content of gestures with regard to emotion [22,23]. Recently, Truong *et al.* [24] introduced a set of 3D gesture descriptors based on an LMA model, to recognize the gesture and emotional content of orchestra conductors. In addition, Zacharatos *et al.* [25] used a set of body motion features, based on the LMA effort component, to provide sets of classifiers for emotion recognition in a game scenario. Morita *et al.* [26] conducted an experiment with the relation between body motion and emotion; they found a good correlation between LMA features and mood rating. A straight-forward application of Laban theories is dance; Shiratori *et al.* [27] used the Laban theory for synthesizing dance motion matched to music. Fourati *et al.* [28] shows relevant body cues for the analysis of emotional behavior during daily actions. Kim *et al.* [29] uses Kinect and LMA to propose an emotional representation for robotics. In [30], Aristidou and Chrysanthou used a variety of LMA features to classify acted dance performances with different emotions, and they also provided a brief analysis of how these features change on movements with different emotions, finding movement similarities between the different emotional states. Some experiments like those of Valstar *et al.* [31] try to map automatic facial emotion recognition to emotion space. In this paper, we will use LMA as a body motion description as it proved its efficiency on a variety of case studies.

As the complex aspect of body expression is compared with that of the facial one, there is for the moment multiple approaches for describing, classifying, and recognizing body emotion, which can show good result but in a limited range of application. In this work, we are using the Aristidou method [32] of LMA features extraction to map emotion as it is expressed onto the RCM diagram. We will use neural network as the mapping function. Research on neural networks skyrocketed recently with respect to their efficiency and ability to solve complicated problems that are difficult to model or consisting of a huge amount of variable. A very interesting tool to approximate function and space clustering, it has been used in the domain of speech synthesis and recognition, and image processing. The ability of unsupervised learning made it appropriate for experimentation. Many scientists used neural networks in emotion recognition [33–35], and the work of Bouchard *et al.* [36] on body motion segmentation was the first attempt to justify their classification criteria. As the neural network shows their efficiency and adaptability for facial and body expression recognition, we believe it is the appropriate tool for body emotion recognition. Our contribution is to analyze expressed emotion during theater performances using extracted LMA entities mapped by a neural network to the bi-dimensional Russell emotion model. After a training phase, we can obtain a function that can make the correspondence motion/emotion over time on a 2D diagram that allows visualization. Beyond the possibility to study emotion dynamic, we demonstrate that LMA features correlate with the Russell Circumplex Model and help achieve automatic emotion mapping.

## 3. LMA FEATURES DESCRIPTION

Human motion analysis is particularly challenging, especially when stylistic characteristics are of high importance. The difficulty is even more pronounced when motion is used to describe and classify human emotions or behaviors. In this work, we utilize a human full-body motion analysis that is based on the LMA principles, aiming to identify those factors that describe the movement signature of the performer.

LMA is a language for interpreting, describing, visualizing, and notating human movement; it offers a clear documentation of the human motion, and it is divided into four main components:

- BODY, which describes the structural and physical characteristics of the human body;
- EFFORT, which describes the intention and the dynamic quality of the movement, the texture, the feeling tone, and how the energy is being used on each motion;
- SHAPE, which analyzes the way the body changes shape during movement; and
- SPACE, which describes the movement in relation with the environment.

The EFFORT component, which is generally related to the changes of mood or emotion, is further divided into four sub-categories, each having two polarities named EFFORT factors:

- *Space* addresses the quality of active attention to the surroundings. It has two polarities, direct (focused and specific) and indirect (multi-focused and flexible attention).
- *Weight* is a sensing factor, sensing the physical mass and its relationship with the gravity. It is related to the movement impact and has two dimensions: strong (bold, forceful) and light (delicate, sensitive).
- *Time* is the inner attitude of the body toward the time, not the duration of the movement. Time polarities are sudden (has a sense of urgency, staccato, unexpected, and isolated) and sustained (has a quality of stretching the time, legato, and leisurely).
- *Flow* is the continuity of the movement; it is related to the feelings and progression. The flow dimensions are bound (controlled, careful, and restrained movement) and free (released, outpouring, and fluid movement).

In order to extract movement characteristics that discriminate human behaviors with regard to emotion, we utilize the LMA framework described by Aristidou *et al.* [32] for full-body motion analysis. We used a 35-frame sliding window with a 1-frame step (our motion data are sampled at 30 fps) to extract the LMA features. The features and their respective measurements are summarized in Table I.

**Table I.** The features and measurements used to extract the movement characteristics based on the LMA components.

| | Features | | Measurements | | | |
|---|---|---|---|---|---|---|
| | $f^i$ | Description | $f^i_{max}$ | $f^i_{min}$ | $f^i_{\sigma}$ | $f^i_{\mu}$ |
| BODY | $f^1$ | Feet-hip distance | $\phi_1$ | $\phi_2$ | $\phi_3$ | $\phi_4$ |
| | $f^2$ | Hands-shoulder distance | $\phi_5$ | $\phi_6$ | $\phi_7$ | $\phi_8$ |
| | $f^3$ | Hands distance | $\phi_9$ | $\phi_{10}$ | $\phi_{11}$ | $\phi_{12}$ |
| | $f^4$ | Hands-head distance | $\phi_{13}$ | $\phi_{14}$ | $\phi_{15}$ | $\phi_{16}$ |
| | $f^5$ | Hands-hip distance | $\phi_{17}$ | $\phi_{18}$ | $\phi_{19}$ | $\phi_{20}$ |
| | $f^6$ | Hip-ground distance | $\phi_{21}$ | $\phi_{22}$ | $\phi_{23}$ | $\phi_{24}$ |
| | $f^7$ | Hip-ground minus feet-hip | $\phi_{25}$ | $\phi_{26}$ | $\phi_{27}$ | $\phi_{28}$ |
| | $f^8$ | Centroid-ground distance | $\phi_{29}$ | $\phi_{30}$ | $\phi_{31}$ | $\phi_{32}$ |
| | $f^9$ | Gait size | $\phi_{33}$ | $\phi_{34}$ | $\phi_{35}$ | $\phi_{36}$ |
| EFFORT | $f^{10}$ | Head orientation | $\phi_{37}$ | | $\phi_{38}$ | $\phi_{39}$ |
| | $f^{11}$ | Deceleration peaks | | | | $\phi_{40}$ |
| | $f^{12}$ | Pelvis velocity | $\phi_{41}$ | | $\phi_{42}$ | $\phi_{43}$ |
| | $f^{13}$ | Hands velocity | $\phi_{44}$ | | $\phi_{45}$ | $\phi_{46}$ |
| | $f^{14}$ | Feet velocity | $\phi_{47}$ | | $\phi_{48}$ | $\phi_{49}$ |
| | $f^{15}$ | Pelvis acceleration | $\phi_{50}$ | | $\phi_{51}$ | |
| | $f^{16}$ | Hands acceleration | $\phi_{52}$ | | $\phi_{53}$ | |
| | $f^{17}$ | Feet acceleration | $\phi_{54}$ | | $\phi_{55}$ | |
| | $f^{18}$ | Jerk | $\phi_{56}$ | | $\phi_{57}$ | |
| SHAPE | $f^{19}$ | Volume (5 joints) | $\phi_{58}$ | $\phi_{59}$ | $\phi_{60}$ | $\phi_{61}$ |
| | $f^{20}$ | Volume (upper body) | $\phi_{62}$ | $\phi_{63}$ | $\phi_{64}$ | $\phi_{65}$ |
| | $f^{21}$ | Volume (lower body) | $\phi_{66}$ | $\phi_{67}$ | $\phi_{68}$ | $\phi_{69}$ |
| | $f^{22}$ | Volume (left side) | $\phi_{70}$ | $\phi_{71}$ | $\phi_{72}$ | $\phi_{73}$ |
| | $f^{23}$ | Volume (right side) | $\phi_{74}$ | $\phi_{75}$ | $\phi_{76}$ | $\phi_{77}$ |
| | $f^{24}$ | Torso height | $\phi_{78}$ | $\phi_{79}$ | $\phi_{80}$ | $\phi_{81}$ |
| | $f^{25}$ | Hands level | | | | $\phi_{82}$-$\phi_{84}$ |
| SPACE | $f^{26}$ | Total distance | | | | $\phi_{85}$ |
| | $f^{27}$ | Total area | | | | $\phi_{86}$ |
| | $f^{28}$ | Total volume | | | | $\phi_{87}$ |

Each feature is decomposed to four measurements: maximum ($f_{max}$), minimum ($f_{min}$), standard deviation ($f_{\sigma}$), and average ($f_{\mu}$), computed on a 35-frame time window.

## 4. PROTOCOL

Theater performances are an interesting case of interaction wherein one protagonist is rather passive (the public) and the other one very active in its proficiency on the communication. During theater performances, actors use a variety of ways to communicate emotion and story with the public. Body expression is therefore of high importance and plays a key role. The interaction aspect though can be placed under consideration as the actor does not have real feedback from the public, but the action of performing will make the actor emphasize on its most expressive capabilities. Notice that to have a powerful expressivity does not mean to express something of high intensity (as "very happy") but rather to be easily understood. It is important to understand here the subtle difference between the internal emotional state of a person (what he or she is thinking) and the expressed emotion that is seen externally. The difference can be misleading, and our study is focused only on the second case.

We firstly record motion capture data of people expressing emotions corresponding to the Russell classification. We chose eight emotions that are distributed along the four quadrants of the model: *happy, excited, afraid, annoyed, sad, bored, tired*, and *relaxed*. As we want a context highlighting a behavior that is interactive, we chose theater

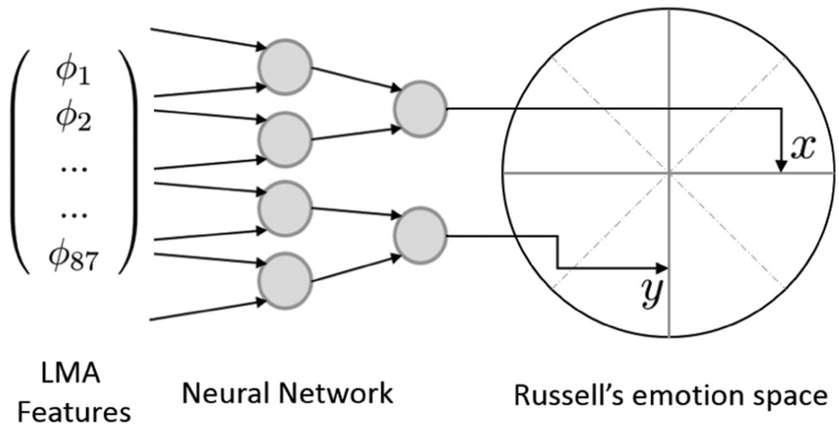


**Figure 2.** Pipeline of mapping motion to the Russell space. This system allows continuous visualization of motion in terms of *activation* and *pleasure*. From 87 LMA features extracted from motion, we have 2D coordinates of emotion.

expression as it is a good case of emotional communication. We invited a number of actors to perform motion sequences, while they express only one emotion at a time. Using the Aristidou method, we extracted the LMA features by averaging measures on a sliding window of 35 frames (on data captured at 30 fps). Then, in order to map the input motion data to an emotional space, we need a multivariable function that outputs a dual of values (Russell coordinates) from a high number of entries (86 LMA values). To do this, we introduced a neural network to approximate the coordinates of the input motion onto the RCM, managing at the same time with a large number of the LMA variables. The proposed pipeline is illustrated in Figure 2.

## 5. EXPERIMENTAL RESULTS

### 5.1. Data Acquisition

In our experiments, we used a Kinect for Windows v2 as a motion capture device. It is a mobile, fast, and efficient device that tracks the skeleton data directly with an internal processing. Although it has low resolution in space and time compare with that of professional motion capture systems, the discreet characteristic presented some advantages concerning the naturalness aspect of recorded motion. Data acquisition problems came mostly from the fact that the Kinect is a depth camera and not a 3D sensing system. As a result, Kinect faces occlusion problems because the recorded person rotates itself, sometimes facing the camera. A person presenting his or her side is more difficult to detect. It leads to some requirements on the actor's performance. Operating at 30 fps with a resolution of $512 \times 424$ pixels for the depth camera, although noncommunicated, the accuracy for skeleton body tracking is sufficient to recognize basic gestures and small spatial displacement. We used a smoothing function during the motion capture for the joint tracking to have better measurements and reduce the subsequent problems. It is important though to note the variability of the data owing to the several factors that influence the performance, such as emotional state and current moods, actor's skills and experience, actor's personality and idiosyncrasies, cultural references, current environment, and reaction of public. The most difficult aspect of our study is that expression of an emotion remains anyway subjective. Thus, we have tried to reduce the potential influence of external factors that affect the quality of the motion during the capturing procedure. As the Kinect device is lightweight and noninvasive, compared to the traditional motion capture system, the performers were more confident about themselves and could behave more naturally. The only technical requirement is to face in a relatively short angle the device. To avoid face-to-camera situations, we asked the remaining actors to face the performer as if there was a real public. Ten different actors from a professional theater school were playing eight different emotional behaviors (Figure 3). The performers were professional actors, one man and nine women, and their age ranges between 18 and 27 years. The actors were asked to perform an emotional state for 40–50 seconds, improvised from their range of acting repertoire. Most of the time, they ask other actors to launch their behavior with a dialogue that triggered an emotional response. In total, we end up with 80 performances, with approximately 53 minutes (96 000 frames) of motion; the data set is sufficiently large, having in mind that in the experiments we do not use individual theater motion but the entire sequence, where the performer acts during 40 seconds without repeating his or her movements. To avoid corrupting the data, the information communicated to the performers was very limited and did not contain the analysis criteria. We stored and used the data as a BioVision Hierarchy (BVH) format, mapped to the same 3D character for uniform processing and analysis.

### 5.2. Results

#### 5.2.1. Training.

We use a neural network that has 86 inputs and two outputs with a total of 10 hidden layers. Of the data, 70% were

**Figure 3.** Professional actor performing theater's expression of emotion and Kinect for Windows v2.

used for training using the Levenberg–Marquardt method and 15% for validation. The 86 measurements proposed by Aristidou *et al.* [32] are used as input values, whereas the output values are the $x, y$ coordinates of the RCM diagram, indicating the pleasure and activation correspondence of the emotion. The remaining 15% data served as test data on which we could use the trained network. All inputs are not normalized, but they have the same factor. We chose 10 layers as they show the best results. It took 20 iterations to have a sufficiently trained network.

After data processing and training of the neural network, we were able to extract a mathematical function that takes 86 values and gives two values. Note that this function is similar to the dimensional reduction method, such as principal component analysis. We try this function first with the 15% remaining test data and second with another motion sequence, where the actor has been asked to change the expressed feeling between a number of different emotions within the same performance.

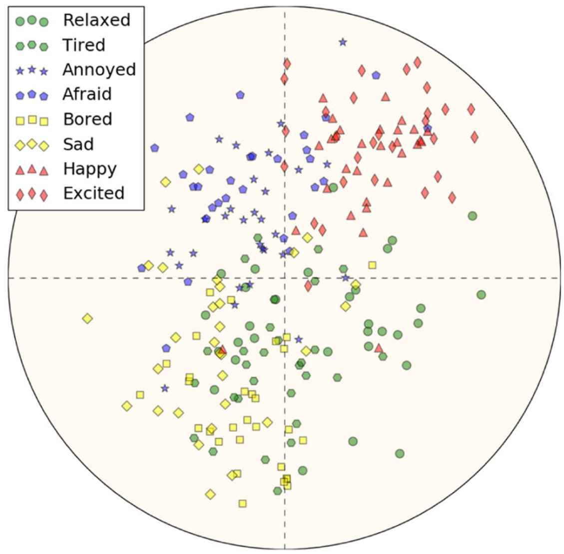


**Figure 4.** Results of the neural network function on the test data after training phase. Each color corresponds to one emotion from all performers. Each point corresponds to LMA measurements performed within a 35-frame window.

#### 5.2.2. Recognition Results.

We can see in Figure 4 the result of the trained neural network. Each color corresponds to one emotion coming from all actors. The dispersion of the clouds highlights the diversity of movement produced by different people. However, most of the points for each cloud are contained in a defined area located in the right Russell quadrant. Please note that we averaged the data for better visualization.

Looking at the clouds individually and separating the people, as for the emotion *happy* in Figure 5, it is obvious that the distribution is almost the same and shows a good acting homogeneity among the actors. Additionally, the covered area is still placed on the correct quadrant of the Russell model.

To improve the visualization of the results, we draw the mean values and the standard deviation of the clouds and overlay it to the RCM, as shown in Figure 6. One can observe the tendency to the center. This is due to the averaging effect of the neural network. It can be reduced by increasing the quality of the data. Another point is that the clouds seem to be aligned along a diagonal. In fact, the difference between *activation* and *deactivation* is easier to produce as an actor and to detect via motion capture than *pleasant* and *unpleasant* aspects. The experiment showed us that this dimension may be more carried by the facial expression. However, it can be observed that the values are quite close to their reference area. They are indeed in their respective part of the angular space, except for *tired* that shows a tendency to be seen as more *unpleasant* than it should be. This can be explained by the difficulty to express emotion that have more *deactivation*: as the actor does not move much, it has to be a very subtle motion.


*Comp. Anim. Virtual Worlds* 2016; **27**:311–320 © 2016 John Wiley & Sons, Ltd.
DOI: 10.1002/cav

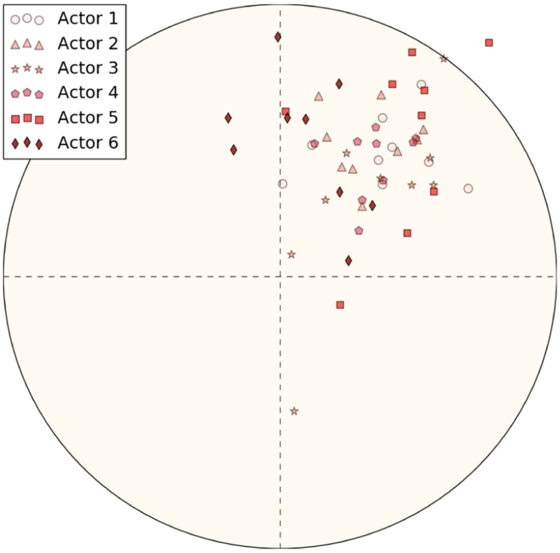


**Figure 5.** Output of the neural network for only the happy emotion. The distribution of each actor is shown. Each point corresponds to LMA measurements performed within a 35-frame window.

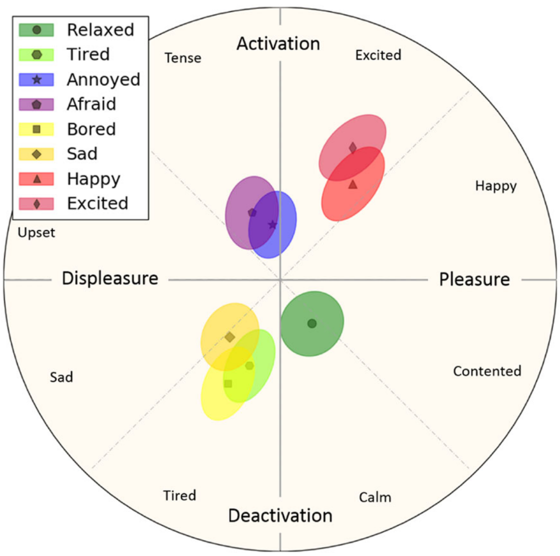


**Figure 6.** Mean value and standard deviation for each emotion, from all performers. For such recognition system, it is difficult to present a percentage of success because of its continuity aspect: we can however see a qualitative feedback as the order along the two axis with respect to the Russell area references.

### 5.2.3. Continuous Recognition.

To test the performance and relevance of the trained neural network for continuous emotion analysis, we registered a sequence of movements that embed four emotions coming from four different quadrant of the Russell model (*sad, afraid, happy*, and *calm*). The actor was asked to perform 10 seconds of each emotion in the previously mentioned order and switch from one to the other in a continuous

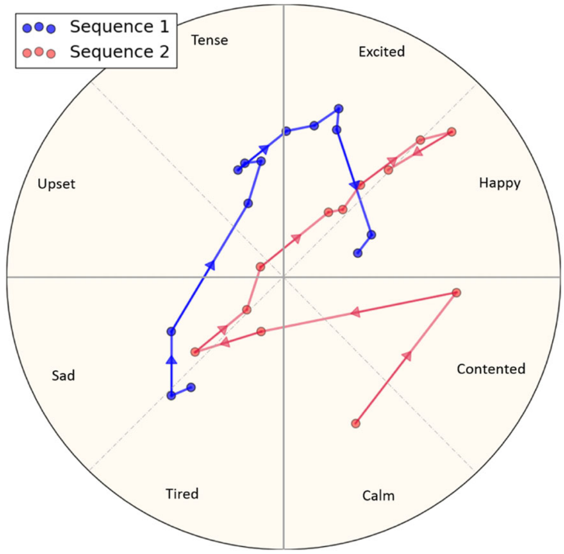


**Figure 7.** Output of the neural network for two sequences (blue and red) of successive emotions. The actors express a series of specific emotions (calm, sad, afraid, and happy, in this order) during a 40-second recording (10 seconds each). We can see the trajectory of the expressed emotion over time. The flow of time is indicated by arrows.

way. We recorded his or her motion and extracted the LMA components for the whole sequence. It represents around 60 LMA vector values. Then we used the trained neural network to visualize the corresponding output. The results are presented in Figure 7, where the trajectories of the detected emotion are shown over time (the arrows mean the direction of time). We can visualize the emotion transition as a trajectory on the diagram. Although some part of the trajectory can vary much, by averaging, there is a clear continuous variation along the quadrants. This continuous detection of the emotion's variation allows us therefore to know that the performers changed his or her emotional state. The presence of noise obliged us to average still in a strong manner, but we can imagine having more fine transition detection with better processing. Our method validate the continuous body emotion recognition by a neural network for theater performances.

## 6. CONCLUSION AND FUTURE WORK

In this work, we have proposed a framework that uses LMA components and machine learning to map an input motion to the Russell Circumplex Model. Results show that we can recognize multiple emotions and place them into the Russell 2D diagram with sufficient accuracy that respects the initial classification. Thus, we show that LMA and neural networks are suitable for continuous analysis of emotion and estimation in terms of *intensity* and *valence*. Analysis of the sequence of emotion shows a progressive change on the diagram, which is an illustration of the continuous shift in emotion over time. We show that we can perform a continuous analysis of emotion in terms of trajectory of the Russell diagram. This can be seen as the study of emotion's dynamic. Beyond the ability to see the emotional trajectory of an agent, it is possible to use this framework in the context of animation on a motion synthesizing perspective to achieve more natural emotion behavior.

To improve the reliability of the recognition, a first idea is to improve the quality of the input data by reducing the inherent problem of depth cameras. A future work would be to improve the capabilities of our capturing system by using a multi-device architecture to deal with larger angles and occlusions, such as the work Kitsikidis *et al.* [37] work. Making a statistic analysis of the impact of the noise on the quality of the data would also help. In addition, as previously mentioned, the *pleasure* dimension is difficult to recognize owing to the multimodality of the context of theater using other channels such as the facial expression made by the performers. Additionally, as our system seems better at picking up the activation dimension, facial expressions will add much information toward discriminating of the pleasure dimension. Therefore, we will integrate facial expressions and combine it with the existing body analysis to obtain a better emotion recognition that leads to a more complete multimodal emotional behavior analysis.

## REFERENCES


1. Ekman P, Friesen WV. *Facial Action Coding System*. Consulting Psychologists Press, Stanford University: Palo Alto, 1977.
2. EMOVU. Available from: www.emovu.com [Accessed on February 2016].
3. NVISO. Available from: www.nviso.ch [Accessed on February 2016].
4. Russell JA. A circumplex model of affect. *Journal of Personality and Social Psychology* 1980; **39**: 1161–1178.
5. Mehrabian A. Pleasure–arousal. Dominance: a general framework for describing and measuring individual differences in temperament. *Current Psychology* 1996; **14**(4): 261–292.
6. Ortony A, Clore GL, Collins A. *The Cognitive Structure of Emotions*. Cambridge University Press: Cambridge, New York, 1994. Autre tirage: 1999.
7. Russell JA. Core affect and the psychological construction of emotion. *Psychological Review* 2003; **110**(1): 145–72.
8. Stanislavsky K. *Building a Character*, Theatre Arts Books. Theatre Arts Books, 1948.
9. Stanislavski K. *An Actor's Work: A Student's Diary*. Taylor & Francis, 1938.
10. Kipp M, Martin J-c. Gesture and emotion: can basic gestural form features discriminate emotions? In *Proceedings of the 3rd International Conference on Affective Computing and Intelligent Interactions*, De Rode Hoed Amsterdam, Netherlands, 2009.
11. von Laban R, Ullmann L. *The Mastery of Movement*. Northcote House, 1988.
12. Bartenieff I, Lewis D. *Body Movement: Coping with the Environment*. Gordon and Breach Science Publishers, 1980.
13. Chi D, Costa M, Zhao L, Badler N. The EMOTE model for effort and shape. In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques—SIGGRAPH '00*, Ernest N. Morial Convention Center New Orleans. Louisiana USA, 2000; 173–182.
14. Zhao L, Badler NI. Acquiring and validating motion qualities from live limb gestures. *Graphical Models* 2005; **67**(1): 1–16.
15. Torresani L, Hackney P, Bregler C. Learning motion style synthesis from perceptual observations. *Neural Information Processing Systems* 2006; **19**: 1393–1400.
16. Luo P, Neff M. In *Proceedings of Motion in Games (MIG'12)*, Rennes, France, 2012; 254–265.
17. Wakayama Y, Okajima S, Takano S, Okada Y. In *Knowledge-Based and Intelligent Information and Engineering Systems: 14th International Conference, KES 2010, Cardiff, UK, 8–10 September 2010, Proceedings, Part IV*, Setchi R, Jordanov I, Howlett RJ, Jain LC (eds). Springer Berlin Heidelberg: Berlin, Heidelberg, 2010; 251–260.
18. Santos L, Dias J. In *Emerging Trends in Technological Innovation: First IFIP WG 5.5/SOCOLNET Doctoral Conference on Computing, Electrical and Industrial Systems, doCEIS 2010, Costa de Caparica, Portugal, 22–24 February 2010. Proceedings*, Camarinha-Matos LM, Pereira P, Ribeiro L (eds). Springer Berlin Heidelberg: Berlin, Heidelberg, 2010; 187–194.
19. Nakata T, Mori T, Sato T. Quantitative analysis of impression of robot bodily expression based on Laban Movement Theory. *Journal of the Robotics Society of Japan* 2001; **19**(2): 252–259.
20. Masuda M, Kato S, Itoh H. Emotion detection from body motion of human. In *The 12th International Conference Principles of Practice in Multi-Agent Systems*, Nagoya, Japan, 2009; 322–334.
21. Masuda M, Kato S. Motion rendering system for emotion expression of human form robots based on Laban Movement Analysis. In *Proceedings—IEEE International Workshop on Robot and Human Interactive Communication*, Congress Center Il Principino Viareggio, Italy, 2010; 324–329.

22. Fagerberg P, Stå hl A, Höök K. Designing gestures for affective input: an analysis of shape, effort and valence. In *MUM 2003: Proceedings of the 2nd International Conference on Mobile and Ubiquitous Multimedia, 10-12 December, 2003, Norrköping, Sweden*. ACM, Norrköping, Sweden, 2003; 57–65.
23. Hartmann B, Mancini M, Pelachaud C. Implementing expressive gesture synthesis for embodied conversational agents. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Vol. 3881 LNAI. Springer, 2006; 188–199.
24. Truong A, Boujut H, Zaharia T. Laban descriptors for gesture recognition and emotional analysis. *The Visual Computer* 2016; **32**(1): 83–98.
25. Zacharatos H, Gatzoulis C, Chrysanthou Y, Aristidou A. Emotion recognition for exergames using Laban Movement Analysis. In *Proceedings of the Motion on Games*, New York, NY, USA, 2013; 39–44. ACM.
26. Morita J, Nagai Y, Moritsu T. Relations between body motion and emotion: analysis based on Laban Movement Analysis. In *The Annual Meeting of the Cognitive Science Society*, Berlin, Germany, 2013; 1026–1031.
27. Shiratori T, Nakazawa A, Ikeuchi K. Dancing-to-music character animation. *Computer Graphics Forum* 2006; **25**(3): 449–458.
28. Fourati N, Pelachaud C. Relevant body cues for the classification of emotional body expression in daily actions. In *Proceedings of the Sixth International Conference on Affective Computing and Intelligent Interaction (ACII2015)*, Xi'an, China, 2015; 267–273.
29. Kim WH, Park JW, Lee WH, Chung MJ, Lee HS. LMA based emotional motion representation using RGB-D camera. In *HRI '13: Proceedings of the 8th ACM/IEEE International Conference on Human–Robot Interaction*, Tokyo, Japan, 2013; 163–164.
30. Aristidou A, Chrysanthou Y. Motion indexing of different emotional states using LMA components. In *SIGGRAPH Asia 2013 Technical Briefs on—SA '13*, Hong Kong, 2013; 1–4.
31. Valstar M, Schuller B, Smith K, *et al.* AVEC 2013: the continuous audio/visual emotion and depression recognition challenge. In *Proceedings of the 3rd ACM International Workshop on Audio/Visual Emotion Challenge (AVEC '13)*, Barcelona, Catalunya, Spain, 2013; 3–10.
32. Aristidou A, Charalambous P, Chrysanthou Y. Emotion analysis and classification: understanding the performers' emotions using the LMA entities. *Computer Graphics Forum* 2015; **34**(6): 262–276.
33. Levine DS. Neural network modeling of emotion. *Physics of Life Reviews* 2007; **4**(1): 37–63.
34. Lotfi E, Akbarzadeh-T MR. Practical emotional neural networks. *Neural Networks* 2014; **59**: 61–72.
35. Thenius R, Zahadat P, Schmickl T. EMANN—a model of emotions in an artificial neural network. *Advances in Artificial Life, ECAL 2013* 2013; **12**(SEPTEMBER): 830–837.
36. Bouchard D. Segmenting motion capture data using a qualitative analysis. In *MIG '15: Proceedings of the 8th ACM SIGGRAPH Conference on Motion in Games*, Paris, France, 2015; 23–30.
37. Kitsikidis A, Dimitropoulos K, Douka S, Grammalidis N. Dance analysis using multiple Kinect sensors. In *VISAPP2014*, Lisbon, Portugal, 2014; 789–795.

## AUTHORS' BIOGRAPHIES

**Simon Senecal** is currently a PhD student and Marie-Curie fellow in the MIRALab laboratory, within the University of Geneva. His research interests include emotion recognition, human behavior, human machine interaction, and motion analysis.

**Louis Cuel** received his PhD in Applied Mathematics from the University of Grenoble, France, in 2016. His research interests include application in 3D mathematical processing, surface reconstruction, 3D visualization, mapping 3D objects and clothes positioning.

**Andreas Aristidou** is a post-doc researcher at the University of Cyprus; his research interests lie on motion analysis and classification, motion synthesis, and involve motion capture and Inverse Kinematics. He had been a Cambridge European Trust fellow, at the Department of Engineering, University of Cambridge, where he obtained his PhD. Andreas has a BSc in Informatics and Telecommunications from the National and Kapodistrian University of Athens, and he is an honor MSc graduate of Kings College London.

**Nadia Magnenat-Thalmann** has pioneered research into virtual humans over the last 30 years. She obtained several Bachelor's and Master's degrees in various disciplines (Psychology, Biology, and Biochemistry) and a PhD in Quantum Physics from the University of Geneva in 1977. From 1977 to 1989, she was a professor at the University of Montreal in Canada. From 1989 to 2010, she was a professor at the University of Geneva where she founded the interdisciplinary research group MIRALab. Together with her PhD students, she has published more than 550 papers on virtual humans and virtual worlds with applications in 3D clothes, hair, body modeling, emotional virtual humans and social robots, and medical simulation of articulations. In 2009, she received a Dr Honoris Causa from the Leibniz University of Hanover in Germany and, in May 2010, the Distinguished Career Award from the Eurographics. In June 2010, she received an Honorary Doctorate of the University in Ottawa in Canada. She is presently a professor and director of the Institute for Media Innovation (IMI) at Nanyang Technological University, Singapore, and the director of MIRALab at the University of Geneva, Switzerland. She is a life member of the Swiss Academy of Engineering Sciences? (SATW).


DOI: 10.1002/cav