PROBABILISTIC LMA-BASED HUMAN MOTION ANALYSIS BY CONJUGATING FREQUENCY AND SPATIAL BASED FEATURES

Kamrad Khoshhal [1] *, Hadi Aliakbarpour* [1] *,Joao Quintas* [1] *,Martin Hofmann* [2] *,and Jorge Dias* [1]

[1] University of Coimbra, Institute of Systems and Robotics, Coimbra, Portugal

[2] Technische Universität München, Institute for Human-Machine Communication, Munich, Germany

ABSTRACT

This paper presents an approach to analyse human motions using Laban Movement Analysis (LMA) system. LMA is a known descriptor for analysing human body motion by using different components that it has. We attempt to present how the frequency and spatial-based features can be applied in LMA concepts, and then we present that how the conjugating of these two kinds of different domains features can be fused by a Bayesian Network (BN). The point is that, the human motions based on the type of body parts motion, can be recognized by some aspects or features better than the others. By combining those different aspects (LMA components) or features, more precise estimation for more various human motions can be obtained. The results approve that our idea to recognize human motion is very successful.

1. INTRODUCTION

This paper presents an approach to analyse human motion based on the LMA notation concepts in different domains. LMA notation is based on different components. This work aims to study about human body dynamics (LMAs Effort) in frequency domain and changing the human body blob (Shape component) in spatial domain.

Human actions emerge from a wide range of signals, most significantly facial expressions, voice and body motion [11]. In a daily basis persons judge each other’s behaviour through the way the others move. This work tries to use the richness of human motion to create a sustainable basis for developing human motion model.

In this approach, human body part positions are the selected data to analyse human motion. We obtain them using a motion tracker suit which gives us several body part poses related to a world reference at a constant frame rate. Then we use linear transformation to project each body parts position related to the body centre reference. Based on the egocentric position data, the acceleration signal of each body part can be obtained.

To analyse Effort component, which is related to the dynamic properties or quantity of energy of human body parts motion, this work proposes to use power spectrum technique. By studying motion in the frequency domain, it is possible to find interesting signal (acceleration) properties (Section 2.1).

To analyse Shape component which is related to the changing of the human body blob, a simple feature to analyse one of the important parameters of Shape component in our case application (Section 2.2) was used.

A three abstraction level Bayesian Net (BN) was used to analyse the data, justified by previous related work and also for its ability to fuse different features from different domains (such as frequency and spatial features) into one global model, which is an essential characteristic in our approach using LMA concepts.

The results of current approach are compared and analysed with the previous work results, given the expectation that using combination of frequency and spatial features are better at the task of estimating different human motions.

1.1. PREVIOUS RELATED WORKS

In our life, we receive a lot of information about people by the messages emerging from their body. Usually, people are able to perceive a lot of information about others without the existence of verbal communication. For example, when a person who is sitting on a chair and shaking his/her foot, we understand that it might be under stress, or in another case, when someone is doing something very fast, we realize it is in a hurry. There are several related interesting works, like the COBOL European project which investigates about human communication understanding through emotional body language by neuro and computer scientists. For instance, the groups of Gelder, Flash and Giese attempted to estimate human behaviours like joy, sadness, anger and fear [5],[6] and [10] by the analysis of their dynamic body expressions.

Analysing human motion is a prerequisite for understanding any human activities, such as human-robot or human-human interaction, etc. Analysis of human activities can be investigated in different levels. Bobick in [3] presented a survey about the different levels of human motion definitions, such as human movement, activity and action. In Bobick's terminology, the movements are the lowest level of

human motions which do not need any contextual or previous knowledge to be identified, but to understand human action or behaviour, a sequence of human movements or states, related to the environment or scenario is necessary.
Several surveys were published about human motion analysis such as Aggarwal et al. in [1] which covered various methods used in articulated and elastic non-rigid motion. Pentland in [12] touched on several interesting topics in human motion analysis and its applications. Moeslund and Granum [8] presented a survey of computer-vision-based human motion capture problems (initialization, tracking, pose estimation and recognition). Poppe in [13] described the characteristics of human motion analysis. The study divided the analysis into modeling and estimation phases.
There are many different kinds of human movements that researchers try to recognize. Each of them can be identified with position variations of one or more involved human body parts and changing human body shape as a blob, without the need to know the underlying movement concept.
We intend to use the advantages of human dynamic characteristics and human body shape characteristics, which can improve the estimation results.
There are some works to analysis different human motions based on their body motion dynamics [4], and human body shape deformations [15], and also [14] that combined both the characteristics to analyse human gestures, but they used just spatial features to obtain all the parameters.
In our previous work [7] the sufficient result for analysing dynamics characteristics of human motion based on the frequency features was presented. In this paper we present that by combining frequency-based and spatial-based features by a BN, more reliable results can be achieved.
Section 2 explains the feature extraction of human movement based on the LMA components. Section 3 describes how the idea was modeled by Bayesian approach. Experimental results are shown in section 4, and Section 5 closes with a conclusion and an outlook for future works.

2. LABAN MOVEMENT ANALYSIS

Laban Movement Analysis (LMA) is a well-known method for observing, describing, notating, and interpreting human movement [2], [4]. The theory of LMA consists of several major components, though the available literature is not in unison about their total number. The works of Norman Badlers group [4] mention five major components; *Body*, *Effort*, *Space*, *Shape* and *Relationship*.

2.1. EFFORT COMPONENT CHARACTERIZATION

One of the most important components of LMA is *Effort* that this work characterizes in the frequency domain. *Effort* is a component dealing with the dynamics of motion for understanding the more subtle characteristics about the way a movement is performed. The difference between punching someone in anger and reaching for a glass is slight in terms of the physical body organization as they both rely on extension of the arm. However the control and timing of the motion in each of these cases are very different. *Effort* has four quantities, and each of them has two bi-polar states. The relation of the Low Level Features (LLFs) and *Effort* properties is discussed in [7].
In [7] *Effort.time*, one of the Effort qualities, was characterized by frequency-based features, which resulted from 3D acceleration signals of several body parts related to a world reference. The presented results showed some false LMA parameters estimations which affected the final classification results. Thus, obtaining LMA-based human motion parameters related to a local reference (human center as an origin of the coordinate system) will provide more accurate data based on LMA definitions, and less data redundancy, resulting in an improvement in human motion description. For clarification purposes lets analyze the walking action: by analyzing the data relative to body center, the computed energy of body parts signals, related to the body-center is reduced.
The Power Spectrum (PS) (obtained by Fast Fourier Transform (FFT)) of the acceleration sequence data can be calculated by equation (1) (as can be seen in Figure 1):

$$\Phi_{bp}(\omega) = \left| \frac{1}{\sqrt{2\pi}} \sum_{-\infty}^{\infty} {}^{c}f_{bp} e^{-i\omega n} \right|^2 = \frac{F_{bp}(\omega) F_{bp}^{*}(\omega)}{2\pi} \qquad (1)$$

Where ω denotes the angular frequency and $F_{bp}(\omega)$ is the continuous Fourier transform of ${}^{c}f_{bp}$, and $F_{bp}^{*}(\omega)$ denotes its complex conjugate.

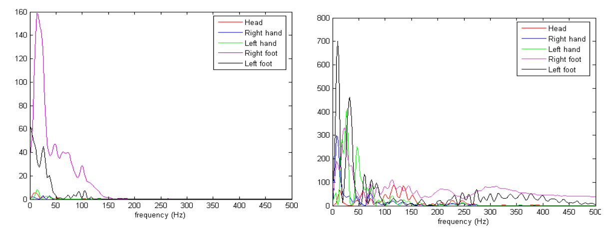

Figure 1: Some samples of PS result of acceleration signals of the some parts of body for (left) walking and (right) running movements.

Having all these PS signals for each selected body part acceleration in different movements, and by collecting first four coefficients (peak) of the extracted signals [7], we will have sufficient features in our application. Thus, four features for each of the five body parts define the feature space to be used for classifying various movements in Effort aspect. We call the coefficients as $Max\{{}^{c}f_{bp}^{i}\}$ which is maximum content of each *i* subdomain frequency of the PS signal for each *bp*. The set of *bp* and subdomain frequency are defined as follows:

$bp \in$ {Head, Right hand, Left hand, Right foot, Left foot}, Frequency $\in$ {(0 - 10), (11 - 20), (21 - 30), (31 - 40)}.

2.2. SHAPE COMPONENT CHARACTERIZATION

The Shape model describes the changes in the shape of the human body through the movements of the limbs in three planes; vertical, horizontal and sagittal. There are some quantities for discretizing of changing shape on each of the three planes:

$$Vertical\ Plane \in \{Sinking, Still, Rising\}$$
$$Horizontal\ Plane \in \{Enclosing, Still, Spreading\}$$
$$Sagital\ Plane \in \{Retreating, Still, advancing\}$$

In this paper, the changing of human shape in vertical plane was modeled, which can be extended to the other planes easily. Difference height of human head and feet related to the body center can be used as a feature to analyse the changing of shape in the vertical plane:

$$\Delta H = \sum_{t=2}^{n} \left(\left({}^{z}X^{t}_{head} + {}^{z}X^{t}_{Feet} \right) - \left({}^{z}X^{t-1}_{head} + {}^{z}X^{t-1}_{Feet} \right) \right), \quad (2)$$

where ${}^{z}X^{t}_{head}$ and ${}^{z}X^{t}_{Feet}$ denote position of head and feet related to the body center in z axis at time t, respectively. *n* denotes the number of frame inside of the window signal, and ΔH denotes the difference height of human body between the first and end of the window signal in centimeter unit. By having this equation result, $Shape_{ver}$ qualities can be estimated by discretizing of the ΔH :

$$Shape_{ver} = \begin{cases} Rising & \Delta H > 0.2 \\ Still & 0.2 \geq \Delta H \geq -0.2 \\ Sinking & \Delta H < -0.2 \end{cases} \quad (3)$$

3. BAYESIAN-BASED MODELING

Bayesian Networks (BN) presents many advantages on using prior knowledge and modeling the dynamic dependencies between parameters of object states. One of the key characteristic of BN is its ability to fuse different types of data in one model. This work developed a classification model based on BN, justified by this flexibility characteristic. Thus, applying Bayesian theory will allow the preparation of a general LMA-based human motion analysis framework to fuse multimodal data, e.g. data emerging from frequency and spatial domains.
Figure 2 presents the BN with two levels in three abstraction layers. In the first level of the BN set, we provided six independent parallel BNs for the five body parts in Effort aspect and one quantity (vertical plane) in Shape aspect.
The set of LLFs emerge from $\{Max\{{}^{c}f^{i}_{bp}\}\}$ which has four states {No, Low, Medium and High}, are obtained using the same threshold definition as in [7]. *Effort.time* component { $Ef.T^{bp}$ } for a body part has two states: Sustained and Sudden. And ΔH which was discretized in 3 states {Up, Still, Down} by Eq.3, to respect to the Shape states which were explained before.
A one second window-frame which moves on the signals at a half of second step was defined. That is a common rule to not lose information through the long signal data. To get a more reliable result, the learning dataset is provided by using the most representative one second of each action sample.

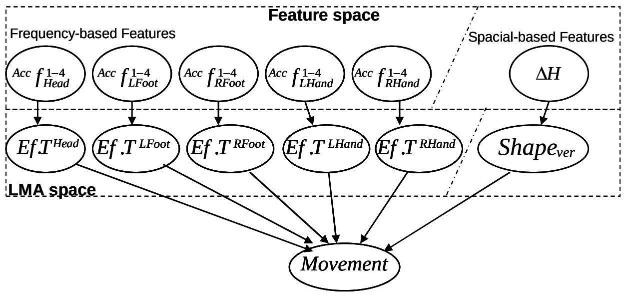

Figure 2: Global Bayesian network

After learning the model, we will be able to answer the Bayesian questions of the probability of LMA parameters given the observed features { $P(LMA \mid LLFs)$ }, and the probability of a movement given the LMA parameters { $P(M \mid LMA)$ }. In [7] can be seen more detail about the BN.

4. EXPREMENTAL RESULTS

The experiments are done for different kinds of human Movements (at least 10 times for each type of movement in different durations). A motion tracker suit was used to obtain human body parts position and acceleration signals. One second window-frame which can shift an half second, was defined on the signals, to collect the features from the sequence of data [7].
A free Bayesian toolbox which was provided by Kervin Murphy and Berkeley [9] is used. The structure which is presented on Figure 2 was implemented using the toolbox.
The proposed method is applied on several collected samples.

Table 1: Classification result using frequency-based feature (Effort)

	Walk	Run	Sit &Bend	Rise	Fall	Stand	%
Walk	102	4				3	93.57
Run	1	67					98.53
Sit & Bend			6			20	23.07
Rise						10	0
Fall	17	4	8		5	4	13.16
Stand						105	100

Table 1 presents the results for walking, running and standing activities, with considerably high positive matches based only on the *Effort.Time* component, but in others activities which are more understandable in the Shape aspect, it was completely failed [7]. Table 2 presents the results of our idea by conjugating the Effort and Shape components.

Table 2: Classification result using frequency and spatial based features (Effort and Shape)

	Walk	Run	Sit &Bend	Rise	Fall	Stand	%
Walk	56						100
Run		60			2		96.77
Sit & Bend			41			1	97.6
Rise				24			100
Fall				1	22		95.7
Stand						109	100

The results show that these two components are sufficient to distinguish the selected human movements. For instance; Effort can distinguish between walking and running, and sitting and falling movements. And Shape is the right component to distinguish between sitting and rising and all activities which occur in standing state (walking and running and standing).

5. CONCLUSION AND FUTURE WORK

In this paper, we presented a novel approach to understand human motions based on the LMA descriptor by conjugating two different types of feature domains, using a 2-step probabilistic approach. Frequency-based features which are extracted by the PS of acceleration signals, are used to analyse the Effort component of LMA which relies on the quantification of energy during human motion [7]. And for the Shape component which relies on change of human shape on the vertical plane, the variation of human height is applied.

A BN has been used in a couple of steps and in different domains (frequency and spatial) to obtain the probabilities of the LMA parameters and human motions. It was presented that to analyse human body motions, not only the dynamics of human body motions are necessary but also the change of human body in space are important. Thus, having a combination of frequency-based features to analyse dynamic of human body motion [7], and special-based feature to analyse deformation of human body in the space, will be sufficient to estimate the performed human movements as can be seen in the results.

We intent to extend the method to analyse human activities and behaviours by estimating the relationship between individual egocentrical body motion features and parameters of environment, which can include other performers (human interaction).

6. ACKNOWLEDGMENT

This work has been supported by the European Union within the FP7 Project PROMETHEUS, www.prometheus-FP7.eu. Kamrad Khoshhal Roudposhti and Hadi Ali Akbarpour are supported by the FCT (Portuguese Foundation for Science and Technology).

7. REFERENCES

[1] J.K. Aggarwal, Q. Cai, W. Liao, and B. Sabata. Articulated and elastic non-rigid motion: a review. pages 2_14, 1994. Proceedings of the IEEE Workshop on Motion of Non-Rigid and Articulated Objects.

[2] N.I. Badler, C.B. Phillips, and B.L. Webber. Simulating Humans: Computer Graphics,Animation, and Control. Oxford Univ. Press, 1993.

[3] A.F. Bobick. Movement, activity and action: the role of knowledge in the perception of motion. Philosophical Transactions of the Royal Society B:Biological Sciences, 352(1358):1257-1265, 1997.

[4] Diane Chi, Monica Costa, Liwei Zhao, and Norman Badler. The emote model for Effort and Shape. In Proceedings of the 27th SIGGRAPH '00, pages 173-182, NY, USA, 2000.

[5] Marina Fridin, Avi Barliya, Edna Schechtman, Beatrice de Gelder, and Tamar Flash. Computational model and the human perception of emotional body language (EBL). In Proccedings of the Symposium on Mental States, Emotions and their Embodiment, pages 16-19. Edinburgh, 2009.

[6] J. Grezes, S. Pichon, and B. de Gelder. Perceiving fear in dynamic body expressions. Neuroimage, 35(2):959-967, 2007.

[7] Kamrad Khoshhal, Hadi Aliakbarpour, Joao Quintas, Paulo Drews, and Jorge Dias. Probabilistic LMA-based classification of human behaviour understanding using power spectrum technique. In 13th International Conference on Information Fusion2010, EICC Edinburgh, UK, July 2010.

[8] Thomas B. Moeslund, Adrian Hilton b, and Volker Kruager. A survey of advances in vision-based human motion capture and analysis. Computer Vision and Image Understanding, pages 90_126, 2006.

[9] Kervin Murphy and Berkeley. Bayesnet toolbox. In http://bnt.insarouen.fr/, 2005.

[10] Lars Omlor and Martin A. Giese. Extraction of spatio-temporal primitives of emotional body expressions. Neurocomput., 70(10-12):1938_1942, 2007.

[11] Maja Pantic, Alex Pentland, Anton Nijholt, and Thomas Huang. Human computing and machine understanding of human behavior: a survey. In proceedings of the ICMI '06, pages 239_248, NY, USA, 2006. ACM.

[12] Alex Pentland. Looking at people: Sensing for ubiquitous and wearable computing. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22:107_119, 2000.

[13] R. Poppe. Vision-based human motion analysis: An overview. Computer Vision and Image Understanding, 108:4-18, 2007.

[14] Joerg Rett. ROBOT-HUMAN Interface using LABAN Movement Analysis Inside a Bayesian framework. PhD thesis, University of Coimbra, 2008.

[15] Dilip Swaminathan, Harvey Thornburg, Jessica Mumford, and et al. A dynamic bayesian approach to computational laban shape quality analysis. Advances in HCI, 2009:17, 2009.