Bayesian reasoning for Laban Movement Analysis used in human-machine interaction

Jörg Rett* and Jorge Dias

Institute of Systems and Robotics,
University of Coimbra,
Polo II, 3030-290 Coimbra, Portugal
E-mail: jrett@isr.uc.pt
E-mail: jorge@isr.uc.pt
*Corresponding author

Juan-Manuel Ahuactzin

Probayes SAS 345, rue Lavoisier,
Inovalléé, 38330 Montbonnot, France
E-mail: Juan-Manuel.Ahuactzin@probayes.com

Abstract: We present the implementation of computational Laban Movement Analysis (LMA) for human-machine interaction using Bayesian reasoning. The research field of computational human movement analysis is lacking a general underlying modelling language, i.e., how to map the features into symbols. With such a semantic descriptor, the recognition problem can be posed as a problem to recognise a sequence of symbols taken from an alphabet consisting of motion-entities. LMA has been proven successful in areas where humans are observing other humans' movements. LMA provides a model for observation and description and a notational system (Labanotation). To implement LMA in a computer, we have chosen a Bayesian approach. The framework allows us to model the process, learn the dependencies between features and symbols and to perform online classification using LMA-labels. We have chosen the application 'social robots' to demonstrate the feasibility of our solution.

Keywords: Laban Movement Analysis; LMA; human movement analysis; computer vision; social robots; Bayesian approaches; Bayesian reasoning; human-machine interaction; human-machine interface.

Reference to this paper should be made as follows: Rett, J., Dias, J. and Ahuactzin, J-M. (2010) 'Bayesian reasoning for Laban Movement Analysis used in human-machine interaction', *Int. J. Reasoning-based Intelligent Systems*, Vol. 2, No. 1, pp.13–35.

Biographical notes: Joerg Rett is currently a PhD student at the University of Coimbra, Portugal. He received his MSc in Electrical Engineering/System Design and Technology from FH Darmstadt, Germany and University of Sunderland, UK. He also received his primary degree as Dipl.-Ing (FH) in Electrical Engineering/Automation at FH Darmstadt, Germany. He has worked mainly with human-robot interaction, social robots, Laban Movement Analysis, robot vision and Bayesian approaches.

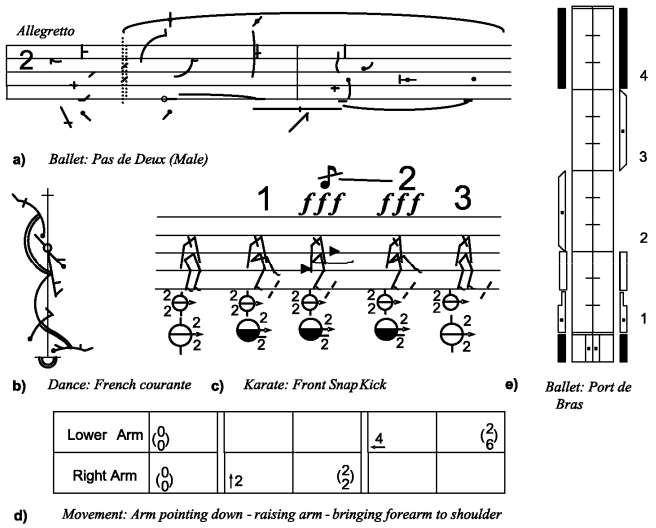
Jorge Dias received the Electrical Engineer Degree (specialisation in computers) from the Faculty of Sciences and Technology, University of Coimbra on 1984. He received his PhD in Electrical Engineering from the University of Coimbra, specialisation in control and instrumentation in 1994. His main research area is computer vision, with activities and contributions in the field since 1984.

Juan-Manuel Ahuactzin received his Bachelor's degree in Computer Systems Engineering from the Universidad de las Americas (UDLA), Mexico in 1989 and his MS and PhD in Computer Science from the Institut Polytechnique National de Grenoble (INPG), France in 1991 and 1994, respectively. In 2005, he joined the Probayes Company as Director of the Research. His interest includes robotics, motion planning, computational geometry, Bayesian programming and genetic algorithms.

1 Introduction

Human body movement is essentially the process of moving one or more body parts to a specific location along a certain trajectory. In some cases, a person observing the movement might be able to recognise it through the spatial pathway alone. Usually, the task of classification is supported by additional evidences which cannot be retrieved from the kinematic information alone. These evidences can be interpreted as the *expressiveness* of movements. Some aspects of expressive movements can be singled out and treated as 'gestures' (Kendon, 2004). Kendon (2004) holds the view that willingly or not, humans, when in co-presence, continuously inform one another about their intentions, interests, feelings and ideas by means of visible bodily action. Analysis of face-to-face interaction has shown that bodily action can play a crucial role in the process of interaction and communication. Kendon (2004) states that expressive actions like greeting, threat and submission often play a central role in social interaction.

Figure 1 Different notational systems for movements
(a) the Banesh Movement Notation
(b) the Beauchamp-Feuillet Notation
(c) the Sutton MovementWriting
(d) the Eshkol-Wachman Movement Notation and
(e) Labanotation



1.1 Notational systems

In order to approach the expressive content of movements scientifically, a notational system is needed. Early notational systems are known from the 17th century. Pierre Beauchamp and Raoul Auger Feuillet began in 1700 a program of publishing notated dances (Little and Marsh, 1992) (see Figure 1). The *Benesh Movement Notation* was invented in the late 1940s by Benesh and Benesh (1983) to document any form of dance or human movement. The *DanceWriting* was first developed in 1966 by Sutton (1982) and extended to a greater body of work called *MovementWriting*. The *Eshkol-Wachman Movement Notation* was developed by the choreographer Noa Eshkol and architect Abraham Wachman (Eshkol and Wachmann, 1958). It has been used to analyse animal behaviour (Golani, 1976) as well as dance. Rudolf Laban (1879–1958), was a notable central European dance artist and theorist, whose work laid the foundations for Laban Movement Analysis (LMA). Used as a tool by dancers, athletes, physical and occupational therapists, it is one of the most widely used systems of human movement analysis. A more detailed overview can be found in Rett (2008). The use of spatial descriptors is a common characteristic for all above mentioned notational systems. What makes the framework of LMA so special is its ability to describe an additional 'expression' that accompanies the spatial trajectory. The *Effort* component can be seen as the key descriptor to solve the task of analysing 'expressive movements'.

1.2 Movement analysis and robotics

Robotics has already acknowledged the evidence that human movements could be an important cue for human-robot interaction. Sato et al. (1996), while defining the requirements for 'human symbiosis robotics', state that those robots should be able to use non-verbal media to communicate with humans and exchange information. This skill could enable the robot to actively perceive human behaviour, whether conscious and unconscious. Fong et al. (2003) state in their survey on 'socially interactive robots' that the design of sociable robots needs input from research concerning social learning and imitation, gesture and natural language communication, emotion and recognition of interaction patterns. Otero et al. (2006) suggest that the interpretation of a person's motion within its environment can enhance human-robot interaction in several ways. They state that body motion and context provide in many situations enough information to derive the person's current activity.

1.3 LMA and believable social robots

In this section, we want to present arguments that lead to the conclusion that LMA will be a very useful 'skill' for social robots. For this, we will present a number of statements taken from a framework for 'socially intelligent agents' by Dautenhahn (1998) and relate them to LMA.

Dautenhahn (1998) argues that the interactions need to be 'acceptable' and 'comfortable' to humans, that humans are active agents who want to use their body and explore the environment and that the interfaces should serve the natural human needs of activity. We conclude that:

1 A social robot that is able to perceive human body movements provides a more comfortable interface to the person.

Dautenhahn (1998) states that 'cognitive technology' has to understand human perception in order to optimise cognitive fit of technologically constructed tools, that the study of biological life and living can further research on artificial agents and that research on social robots could learn from human factors (ergonomics) about the study of how humans

and machines interact in order to design technology that work well in 'human terms'. We conclude that:

2 LMA is natural because it is based on humans observing other humans' movements.

Dautenhahn (1998) suggests that a central set of mechanisms which constitutes 'social intelligence' in humans are 'stories', that if stories are fundamental to human (social) intelligence, then social robots have to be good at telling and listening to stories. In dance movements, the art of choreography can also be interpreted as an attempt to tell a story and whole choreographies are written using Labanotation. We conclude that:

3 LMA is natural because its symbols can be used to tell a story.

Dautenhahn (1998) argues that social robots should support individualised interactions, when personality, character, individual relationships are desirable (e.g., a personal assistant). A social robot which can adapt to our habits can be more robust through a smaller number of probable hypotheses. The style of movement is a personal attribute, thus, a descriptor of expressive movement is also a descriptor of the person itself. We conclude that:

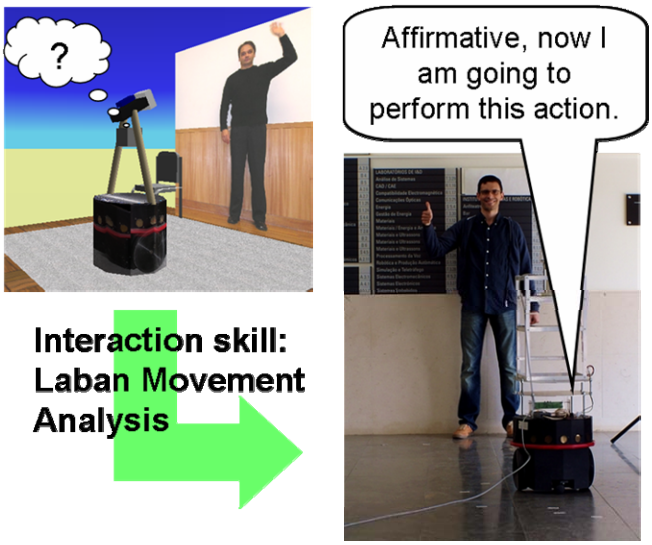
4 LMA is natural because it allows a personalisation of the system to the user.

Dautenhahn (1998) states that believable technology is 'familiar' to humans, it meets their cognitive and social typically human needs. We finally conclude that:

5 The skill of LMA appears to be natural, thus a social robot having this skill will also be believable.

Figure 2 summarises the goal of the work described in this article. LMA is implemented as a skill for human-robot interaction.

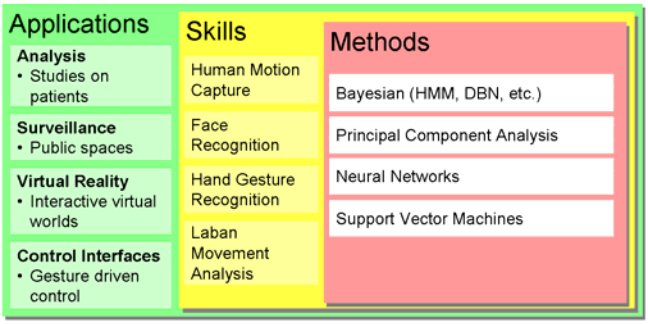
Figure 2 Creating a social robot through implementing an interaction skill, i.e., LMA (see online version for colours)



1.4 The problem – computational human movements analysis

This article is in the research area of 'computational human movement analysis' (C-HMA). While naming the important issues and giving references to the state of the art, we will show how computational LMA fits in. In literature, this area has been named 'visual analysis of human movements' (Gavrila, 1999), 'looking at people' (Pentland, 2000), 'human motion analysis' (Aggarwal and Cai, 1999) or 'vision-based human motion capture and analysis' (Moeslund et al., 2006). Gavrila (1999) points out that the ability to recognise humans and their activities by vision is a key for machine to interact intelligently and effortlessly with a human-inhabited environment.

Figure 3 Some frequently found application, 'skills' and methods in C-HMA (see online version for colours)



In order to organise and compare research in the area of C-HMA three issues are important:

1 the application that is targeted

2 the 'skill' that the system provides to support interaction with humans

3 the methods that are used to design and implement the system.

The area of C-HMA has a wide range of applications, thus the literature (Moeslund and Granum, 2001; Gavrila, 1999) clusters them in more or less the groups shown in Figure 3. The type of application also determines the 'skill' which the system needs to provide. *Human motion capture* is the task of tracking several body parts. *Face recognition* is a skill that is closely connected to human movements as it allows personalisation. *Gesture recognition* is the skill to classify a meaningful movement. *LMA* is the skill discussed in this article. There are certain methods that have been used frequently in C-HMA some of them are shown in Figure 3. Diard et al. (2003) predict that in the future, probabilistic reasoning will provide a new paradigm for understanding neural mechanisms and the strategies of animal behaviour. Further, that it will raise the performance of engineering artefacts to a point where they are no longer easily outperformed by the biological examples they are imitating. In the context of this work, Bayesian methods will be used to design the model for computational LMA, for learning and for classification of movements.

Recognition

We relate our work to a taxonomy suggested by Moeslund and Granum (2001) which divides the problem of C-HMA into the four processes:

1 The *initialisation process* establishes the model that is used for the observed object.

2 The *tracking process* holds the methods for segmentation of the subject in the image from the background, extracting 'good' features and finally correspondence of the segments of features between consecutive images.

3 The *pose estimation process* defines to which extend a model of the human body is used.

4 The *recognition process analyses* some features or variables to classify the movement.

Using this implementation oriented taxonomy the contribution of our work lies mainly in the 'recognition process'.

Sensor modalities

Two main streams of sensor modalities for gathering human movement data can be found. One relies on computer vision and sensor signals generated by some kind of camera, the other is based on devices that are placed on the subject which transmit or receive generated signals. In Moeslund and Granum (2001), the former is called 'passive', the latter 'active sensing'. Active sensing allows for simpler processing and is widely used when the applications are situated in well-controlled environments (Moeslund and Granum, 2001). Passive sensing is mainly used in situations where mounting devices on the subject is not an option. Given the benefits and drawbacks of the main streams, this work will follow a mixed approach. The system benefits from the simpler processing and higher precision of the active sensor during the recoding step, while using the attractive touch-free alternative of computer vision for the classification step.

Labelled datasets

Usually, movement data are labelled by low-level descriptors (e.g., *XYZ* position) or high-level concepts (e.g., 'writing on a black-board'). Mid-level labels are whether missing or 'invented' for a specific application. Moeslund and Granum (2001) conclude that the research field of C-HMA is lacking a general underlying modelling language. A semantic descriptor allows posing the classification task as a problem to recognise a sequence of symbols taken from an alphabet consisting of motion-entities. Systems which are based on such modelling language can use it as a ground truth for recoding and labelling training data. The inherent constraints of a modelling language can be used to make the task of movement recognition more tractable. This work poses the automatic movement classification task as a problem to recognise a sequence of symbols taken from an alphabet consisting of motion-entities. The alphabet and its underlying model is well defined though LMA. The LMA parameters serve as mid-level descriptors that can be produced and understood by the system.

1.5 Related works

Laban Movement Analysis

A long tradition in research on computational solutions for LMA has the group around Norman Badler, who already started in 1993 to re-formulate Labanotation in computational models (Badler et al., 1993). Zhao and Badler (2005) presented a computational model of gesture acquisition and synthesis which can be used to learn motion qualities from live performance. A more detailed version of their work can be found in the thesis of Zhao (2002), who based his work on the earlier implementation of the 3-D animation control module EMOTE (Chi et al., 2000). The work of Zhao (2002) is related with our work, particularly the relationship of LMA components with physically measurable entities. For classification of *Effort* qualities from the low-level features a three-layered feedforward neural network (NN) was used. Zhao's (2002) main contribution was the automatic classification of *Effort* qualities from movement data obtained from an active sensor as well as from a (stereoscopic) visual tracker. Our article goes one step further by learning the 2-D projections, which allows feeding in multiple single-camera data. In our case, Bayesian models and their representation as Bayesian nets are used, which offer the possibility to discuss the phenomenon in terms of dependencies, observations and probabilities. Zhao (2002) reported that, when feeding in data from the visual tracker the classification results decreased. This was partially due to the noise generated by occlusions and the tracker getting lost. Probabilistic approaches, like the one used in this work, usually perform better under such circumstances.

Nakata et al. (2002) reproduced expressive movements in a robot that could be interpreted as emotions by a human observer. The first part described how some parameters of LMA can be calculated from a set of low-level features. The critical point in Nakata et al. (2002) is the mapping of low-level features to LMA parameters. The computational model is closely tied to the embodiment of the robot which has only a low number of degrees of freedom. The major physical entities were chosen subjectively by the designer without experimental data for evaluation. This article investigates the framework of LMA as deeply as possible to choose 'good' candidates for low-level features.

Human movement analysis

There has been an interesting work which also used movement descriptors and a probabilistic framework.

Bregler (1997) introduced mid-level descriptors embedded in a thorough probabilistic framework that produced a robust classification for human movements. The concept of multiple hypotheses is kept from low-level motion clusters to high-level gait categories producing good classification results even for noisy and uncertain evidences in natural environments. Model parameters are learned from training data using the EM-algorithm. The work points towards the concept of atomic phonemes and words used in speech recognition. Bregler (1997) defines his 'movemes' as simple dynamical categories, i.e., a set of second order linear dynamical systems. Bregler (1997) used a probabilistic method for the same reasons it is used in our work:

1 noisy input images

2 spatial and temporal ambiguities

3 occlusion

4 cluttered environments

5 large variability.

The critical point in his approach were the 'movemes' themselves. The 'movemes' appear limited in their expressiveness. This might have been caused by their simplicity and that no relations are drawn to models and data of physiological studies of human movements. To overcome this weakness, the work presented in this article ties the descriptors to a well established notational framework: LMA.

In Rosales and Sclaroff (2000) 3-D data from an active sensor was used to obtain a set of movement sequences. Then 2-D projections from several orientations are generated. For the same orientations, projections of a 3-D model to images are created. Treating the two sets as input-output data a NN was trained. They achieved good results for training five sequences sampled at 32 orientations. As their system only provides the pose of a human body, the classification of movements still remained an open issue and consequently no descriptor was introduced. Also, online or real-time behaviour was not addressed in their work.

Very good classification results obtained from probabilistic methods was also demonstrated for the application of sign language recognition. Starner and Pentland (1995) based their system on real-time tracking of the hands using colour gloves and a single camera with five frames per second. Starner et al. (1998) later removed the constraint of using colour gloves and added relative displacement as a feature. The problem of personalisation was not addressed and it appears that the dataset may be recorded from a single person. Starner (1995) first used a five-state hidden Markov model (HMM) with three-skip transitions, later a four-state HMM with two-skip transitions (Starner and Pentland, 1995) and, finally, it reduced to one skip transition (Starner et al., 1998). This reflects a main problem when using HMM: the design of the 'best' topology. The topologies presented in this article are designed without explicit temporal transitions.

Social robots

The term *social robots* are strongly associated with anthropomorphic social behaviour (Breazeal, 2003). Scenarios in which social robots have been tested already are museums and exhibitions. The guide robots of Nourbakhsh et al. (2003), Burgard et al. (1998) and Siegwart et al. (2003) have already developed quite advanced methods for autonomous navigation and provide various output modalities for the interaction with the human. The fact that their input modalities mainly relies on pressing some buttons shows the need for a more intuitive way of interaction. Burgard et al. (1999) contribute this to a lack of a convincing methodology for 'intuitive' human-robot interaction where it is not desirable to expect the person to learn a large variety of control gestures. Our work suggests:

1 LMA as a convincing methodology

2 expressive movements as an intuitive modality for interaction.

1.6 The contribution of this work

The contribution of this work for the field of C-HMA encompasses the following points:

1 'This work provides semantic descriptors for the automatic analysis of human movements based on the framework of LMA'. The alphabet and its underlying model is well-defined though LMA. Systems which are based on this modelling language can use it as a ground truth for recoding and automatic labelling of training data.

2 'This work shows the design of probabilistic models which relate physically measurable entities obtained from movement tracking to the descriptors of LMA.' Both points 1 and 2 use a common language which allows discussing and incorporating knowledge gained in human science. This article provides a theoretical 'fit' in description and modelling.

3 'This work implements classifiers for human movements based on the descriptors of LMA'. By using the descriptors of LMA, especially *Effort*, the process of movement recognition can reach a higher level of sophistication (expressiveness of a movement).

4 'This work uses a Bayesian approach for the process of learning and classification'. The Bayesian approach provides the classifier with the ability to better deal with the always apparent incompleteness of the real-world data. Additionally, it creates a measure of certainty for human-robot interaction.

5 'This work follows a mixed approach for human movement tracking by recording the data with an active sensor and then mapping it to a camera plane'. Our system benefits from the simpler processing and higher precision of the active sensor during the recoding step, while using the attractive touch-free alternative of computer vision for the classification step.

1.7 The organisation of this article

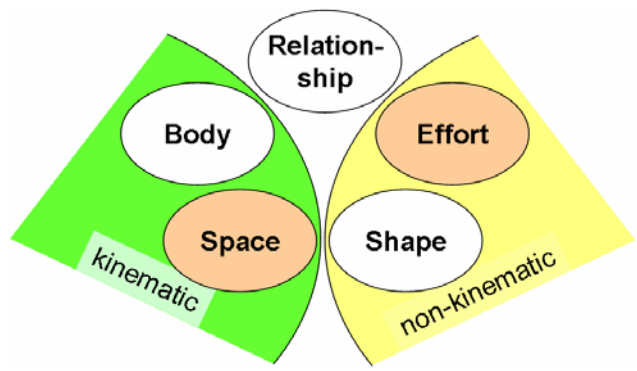
Section 2 will unfold the framework of LMA. The two components *Space* and *Effort* will be discussed in detail and prototypical movements will be presented. Section 3 presents the technique and models related with computational retrieval of human movement data. The correspondence between LMA and low-level features are shown and statistically evaluated. The concept of segmentation is presented. Section 4 introduces the Bayesian approach and presents the models for computational LMA. It shows the process of learning and online classification. In Section 5, the implementation of the system is presented. Results are compared for different types of classifiers. The application of the system for the social robot Nicole is presented. Section 6 ends this article with conclusions of the presented work and developments that are planned for the future.

2 Laban Movement Analysis

LMA is a method for observing, describing, notating and interpreting human movement. The general framework was described in 1980 by Irmgard Bartenieff a scholar of Rudolf Laban in Bartenieff and Lewis (1980). While being widely applied to studies of dance and application to physical and mental therapy (Bartenieff and Lewis, 1980), it has found little application in the engineering domain. Most notably, Badler et al. (1993) who already started in 1993 to reformulate Labanotation in computational models. Recently, researchers from neuroscience started to investigate the usefulness of LMA to describe certain effects on the movements of animals and humans. Foroud and Whishaw (2006) adapted LMA to capture the kinematic and non-kinematic aspects of movement in a reach-for-food task by human patients whose movements had been affected by stroke. It was stated that LMA places emphasis on underlying motor patterns by notating how the body segments are moving, how they are supported or affected by other body parts, as well as whole body movement.

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 Badler's group (Chi et al., 2000; Zhao, 2002) mention five major components as shown in Figure 4. *Relationship* describes modes of interaction with oneself, others and the environment (e.g., facings, contact and group forms). As *Relationship* appears to be one of the lesser explored components, some literature (Foroud and Whishaw, 2006) only considers the remaining four major components. *Body* specifies which body parts are moving, their relation to the body centre, the kinematics involved and the emerging locomotion. *Space* treats the spatial extent of the mover's *kinesphere* (often interpreted as reach-space) and what form is being revealed by the spatial pathways of the movement. *Effort* deals with the dynamic qualities of the movement and the inner attitude towards using energy. *Shape* is emerging from the *Body* and *Space* components and focused on the body itself or directed towards a goal in space. The interpretation of *Shape* as a property of *Body* and *Space* might have been the reason for Irmgard Bartenieff to mention only three major components of LMA. Like suggested in Foroud and Whishaw (2006) we have grouped *Body* and *Space* as kinematic features describing changes in the spatial-temporal body relations, while *Shape* and *Effort* are part of the non-kinematic features contributing to the qualitative aspects of the movement (see Figure 4). We expect that the strength of computational LMA can already be proven by regarding one component from each group. Thus, in the following, we will concentrate on the *Space* and *Effort* component and introduce the other components elsewhere (Rett, 2008).

Figure 4 The major components of LMA are *Body*, *Space*, *Effort*, *Shape* and *Relationship* (see online version for colours)



2.1 Space

The *Space* component presents the different concepts to describe the pathways of human movements inside a frame of reference, when 'carving shapes in space' (Bartenieff and Lewis, 1980). *Space* specifies different entities to express movements in a frame of reference determined by the body of the actor. The different entities which are specified by the *Space* component have been presented in Rett et al. (2008). Two entities are regarded in this article as shown in Figure 5(a):

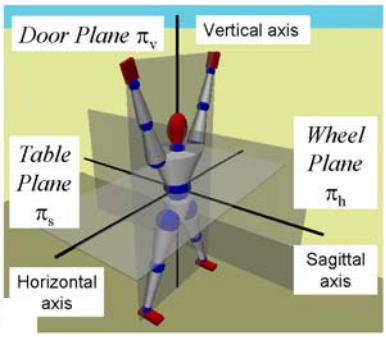
1 the *three axes* – vertical, horizontal and sagittal axis

2 the *three planes – door plane* (vertical) π_v, *table plane* (horizontal) π_h and the *wheel plane* π_s (sagittal) each one lying in two of the axes.

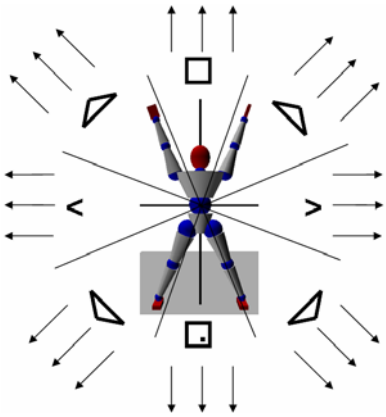
This work uses the concept of *vector symbols* (Longstaff, 2001) which is based on lines of motion as shown in Figure 5(b). In Rett et al. (2008), these *vector symbols* have

been projected to the *three planes* producing a two-dimensional representation. This work regards only those *vector symbols* related with the *door plane* π_v.

Figure 5 (a) Two concepts of the *Space* component: 1) *three axes* and *three planes* (b) *vector symbols* are based on lines of motion which are projected to, e.g., the *door plane* π_v (see online version for colours)



(a)



(b)

2.2 *Effort*

The *Effort* component describes the dynamic qualities of the movement and the inner attitude towards using energy. It consists of four *Effort* qualities: *Space*, *Weight*, *Time* and *Flow*. In Bartenieff and Lewis (1980) the underlying cognitive process was suggested and summarised in Rett (2008). Each quality is bipolar and can have values between two extremes. The values for the *Effort* qualities are shown in (1):

$$\begin{aligned} Space &\in \{direct, neutral, indirect\} \\ Time &\in \{sudden, neutral, sustained\} \\ Weight &\in \{strong, neutral, light\} \\ Flow &\in \{bound, neutral, free\} \end{aligned} \tag{1}$$

Movements are described and distinguished by those qualities that are close to an extreme, e.g., a *punch* has *strong weight*, *sudden time* and *direct space*. For this movement, the *flow* quality is considered to be neutral.

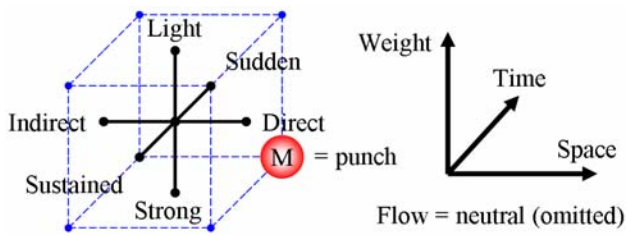
Combinations of all four qualities close to an extreme rarely occur as they produce extreme movements (e.g., tearing something apart) (Bartenieff and Lewis, 1980). Also single-quality movements are rare (Bartenieff and Lewis, 1980) and even for a trained Laban performer (i.e., Laban notator) difficult to perform (Zhao, 2002). Combinations of three qualities, with the fourth considered to be neutral, appear to be the most natural way to perform an action. With only *Flow*, *Space*, *Weight* or *Time* being neutral, the combinations are called *Action Drive*, *Spaceless*, *Timeless* or *Weightless*, respectively. Laban associated some *Basic Effort Actions* to the *Action Drive* which are shown in Table 1. The literature on LMA like Zhao (2002) often gives some exemplary movements like those shown in the second column. More examples are given in Rett (2008).

Table 1 Basic effort action drives

Action	*Example*	*Space*	*Weight*	*Time*
Punch	Punching	Direct	Strong	Sudden
Slash	Slashing	Indir.	Strong	Sudden
Dab	Lunging	Direct	Light	Sudden
Flick	Cleaning	Indir.	Light	Sudden
Press	Pushing	Direct	Strong	Sustained
Wring	Stretching	Indir.	Strong	Sustained
Glide	Erasing	Direct	Light	Sustained
Float	Spraying	Indir.	Light	Sustained

For the remaining three combinations (*Spaceless*, *Timeless* or *Weightless*), no explicit actions were defined but some examples were given in Bartenieff and Lewis (1980). Using the concept of *Effort* combinations, a movement can be defined by its position in one of the four 3-D spaces. The space of the *Action Drive* is often modelled as a cube where each vertex represents an action (see Figure 6). The edge length represents the distance between two extremes (e.g., sudden and sustained). Figure 6 shows the space of *Action Drive* with some movement M; in this case, a *Punch*.

Figure 6 The bipolar *Effort* qualities of the *action drive* represented as a cube (see online version for colours)



Notes: *Flow* = *neutral* (omitted). The position of the movement M (*punch*) indicates its qualities, i.e., *direct*, *sudden* and *strong*.

Movements with only two *Effort* qualities are called *Incomplete* or *Inner States* as they occur often during transitions between two three-quality combinations. They can also reflect a failure to produce a certain three-quality action (e.g., an attempt to perform a *punch* fails due to weakness).

2.3 Database of expressive movements

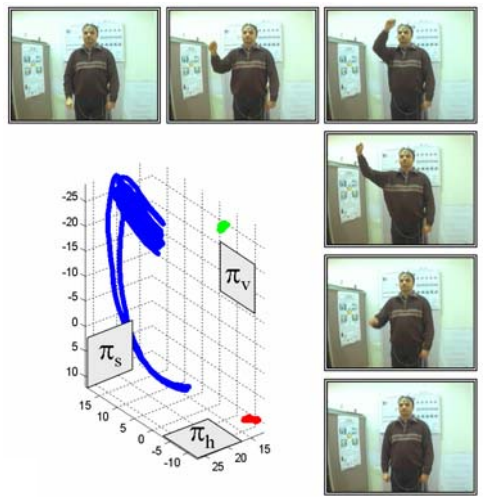
From the 32 possible movements with distinct *Effort* qualities that can be derived from the LMA definitions of *Action Drive*, *Spaceless*, *Weightless* and *Timeless* we have created a database of 'expressive movements'. The database extended the earlier 'gesture' set (Rett and Dias, 2007). Some of the movements are based on suggestions mentioned in Bartenieff and Lewis (1980) and Zhao (2002) others are commonly used gestures with anticipated *Effort* qualities.

From the database, sets of movements were put together. The first set is called 'expressive movements' and holds movements which show some interesting spatial patterns. This set is presented in greater detail in Rett et al. (2008). Though some of the 'expressive movements' already have some distinguishable *Effort* qualities a second set of movements ('bye-byes') was put together as shown in Table 2. The reason for this was to have spatially one single movement pattern (*byebye*) but four different performance 'flavours' (*Effort* combinations).

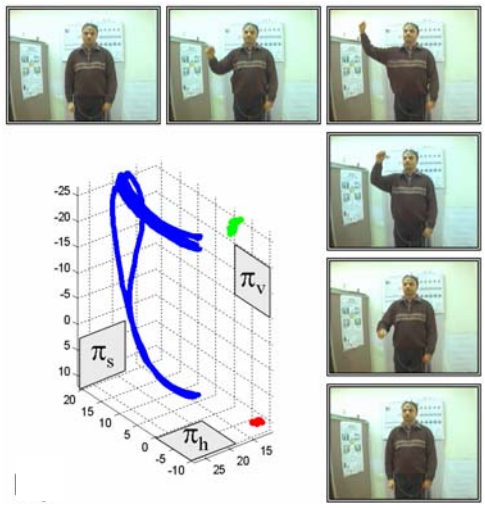
Table 2 Expressive movements from our human interaction database (HID) with *Effort* qualities, principal plane π_{prin} and principal phase Ph_{prin}

M	*Space*	*Weight*	*Time*	π_{prin}	Ph_{prin}
Dab	Direct	Light	Sudden	YZ	Stroke
Flick	Indir.	Light	Sudden	YZ	Stroke
Glide	Direct	Light	Sust.	YZ	Stroke
Float	Indir.	Light	Sust.	YZ	Stroke

Figure 7 Two movements from the set 'bye-byes' (a) bye-bye performed in a 'flick' way and (b) bye-bye performed in a 'glide way' (see online version for colours)

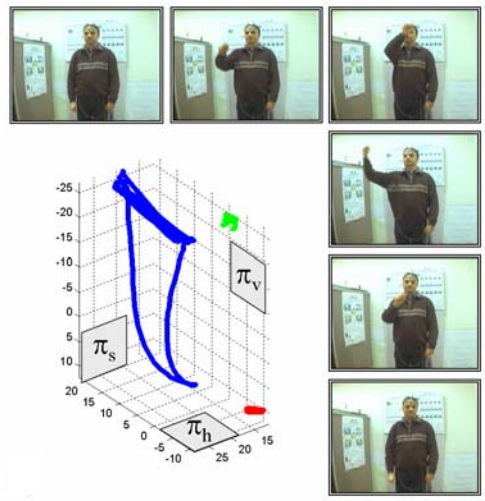


(a)

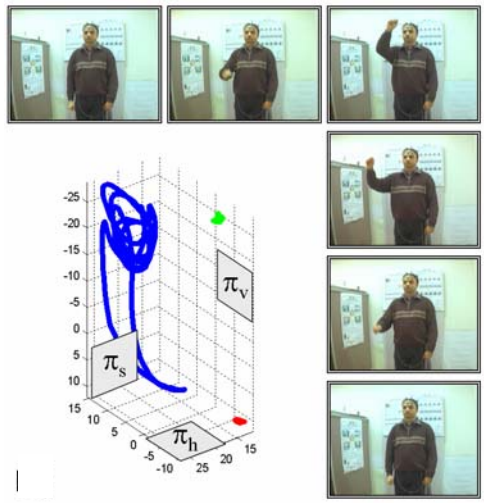


(b)

Figure 8 Two movements from the set 'bye-byes' (a) bye-bye performed in a 'dab' way and (b) bye-bye performed in a 'float way (see online version for colours)



(a)



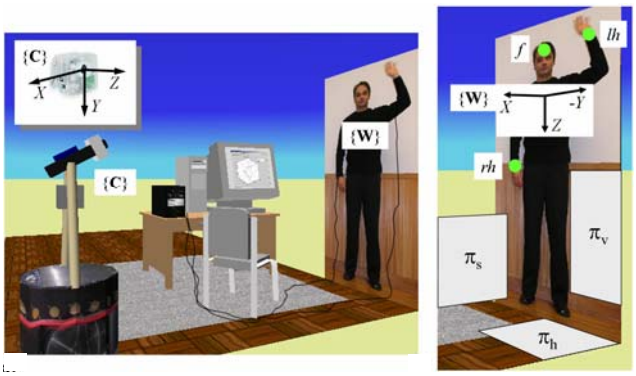
(b)

The *Space* quality is considered *neutral* and the *Weight* quality is *light* for all movements M. Table 2 reveals their distinct *Effort* qualities while the trajectories appear similar as shown in Figure 7. In this case, both movements have the spatial form of a *bye-bye* gesture and thus produce a similar sequence of *vector symbols*. The *Effort* qualities for the 'flick' way of performance will be *Space = indirect* and *Time = sudden* while the 'glide' way of performance will yield *Space = direct* and *Time = sustained*. Figure 8 show two more distinct cases in terms of *Effort*. The *Effort* qualities for the 'dab' way of performance will be *Space = direct* and *Time = sudden* while the 'float' way of performance will yield *Space = indirect* and *Time = sustained*.

3 Human movement tracking

The main idea transported in this section is a mixed approach concerning the sensor technology to obtain human movement data. By using an active sensor during the recoding step the system can benefit from simpler processing and higher precision. For the classification step the attractive touch-free alternative of computer vision can be used. Once the correspondence between the different sensor modalities is established the recorded and labelled, 3-D movement data obtained from the active sensor can be mapped to 2-D plane(s) aligned with the camera plane(s). After this, the same algorithm can be used to calculate low-level features and perform a temporal segmentation.

Figure 9 Example of a scene for human-robot interaction, (a) frames of reference can be defined for a camera $\{C\}$ and the world $\{W\}$ (b) world frame of reference $\{W\}$ with orientation of axes, principal planes and sensor positions (see online version for colours)



(a) (b)

The scenery for an interaction of a human with a robot or, more general, a machine can be described by frames of reference. Figure 9(a) shows an example with a mobile robot equipped with a camera to perform online classification of movements. Furthermore, the active sensor for the preceding step of 3-D movement recording can be seen. The world frame of reference $\{W\}$ might be placed at any position in the 3-D scenery. For the experimental setup used in this work, the world frame of reference $\{W\}$ coincides with the frame of reference of the active sensor. This is why $\{W\}$ is close to the centre of the *kinesphere* defined by the hands and the face. Figure 9(b) presents a closer look at $\{W\}$ by showing the orientation of the axes and the three principal planes (*door plane* π_v, *table plane* π_h and *wheel plane* π_s) as mentioned in LMA. The figure also indicates the position where the three active sensors are attached, i.e., right hand *rh*, left hand *lh* and face *f*.

The active sensor and the visual tracker are described in more detail in Rett (2008). The geometrical relationship between the two sensor modalities is shown in Rett et al. (2008) while the process of calibration is presented in Rett (2008). With this any 3-D position can be related to a point in the 2-D projection. The spatial concept of 2-D planes has also been shown earlier in Section 2 through the principal planes (*door plane* π_v, *table plane* π_h, *wheel plane* π_s).

3.1 Relating LMA and low-level features

Three paradigms guided the selection of features in this work. First, the features were chosen by interpreting the parameters of LMA through physical measurable entities that could describe them best. Second, the features and their cardinality were chosen by predicting an optimal performance when using a Bayesian method for learning and classification. Third, the features were chosen according to our interpretation of 'Ockham's Razor', that is simple features, low cardinality and a small number of them. The initial hypotheses of correspondences between LMA parameters and physical entities are expressed as shown in Table 3. The hypotheses were established with having our primary paradigm in mind.

Table 3 Initial hypotheses of correspondences between LMA parameters and physical entities

LMA parameter	*Physical entities*
Space	Displacement angle
Time.sudden	High acceleration, high velocity
Time.sustained	Low acceleration, low velocity
Space.direct	Small curvature, high angular vel.
Space.indirect	High curvature, high angular vel.
Weight.strong	Muscle tension, medium accel.
Weight.light	Muscle relaxed

For the *Space* component of LMA the sequence of displacement angles is used as a descriptor as shown in Rett et al. (2008). As the position data is projected to planes, each plane produces a sequence of displacement angles with a certain sampling rate and discretisation. For the *Effort* component of LMA, the assumption of a high acceleration when *Time.sudden* occurs seems to be a logical choice. The high velocity might follow as a consequence of the high acceleration. The inverse situation is assumed during *Time.sustained* when low acceleration and velocity is assumed. Interpreting *Space.direct* as reaching towards a

target we can assume a straight trajectory of the hand. This suggests taking the curvature into consideration as a measure of 'directness'. The mathematical definition of curvature though, requires a parameterised curve which is independent of time t. We decided to approximate the curvature by calculating the change of displacement angles (angular change; angular velocity). Zhao (2002) used in his computational LMA the displacement D, the estimated velocity $\hat{v}_i$, the estimated acceleration $\hat{a}_i$ at t_i and the average velocity and acceleration over a segment.

3.2 Computing the low-level features

All computations are based on the raw tracking data inside our HID. The tracking data consists of:

1 the 2-D or 3-D position $\mathbf{X}_{bp}$ of a point belonging to a body part bp

2 the timestamp t_i given by some timer function of the system.

The position is defined in a frame of reference ϕ indicated by ${}^{\phi}\mathbf{X}$. This usually indicates the sensor used for input like the camera $\{C\}$ or the active sensor $\{W\}$. With the sampling (frame) index i the sampling interval Δt_{i+1} can be calculated between two consecutive frames i and $i+1$. In order to treat 2-D and 3-D data equally, the first step is to project the 3-D data to some suitable planes. This work regards only the *door plane* (vertical) π_v and the right hand *rh*. To allow for a fast computation we are discretising the low-level features to a low cardinality as shown in (2).

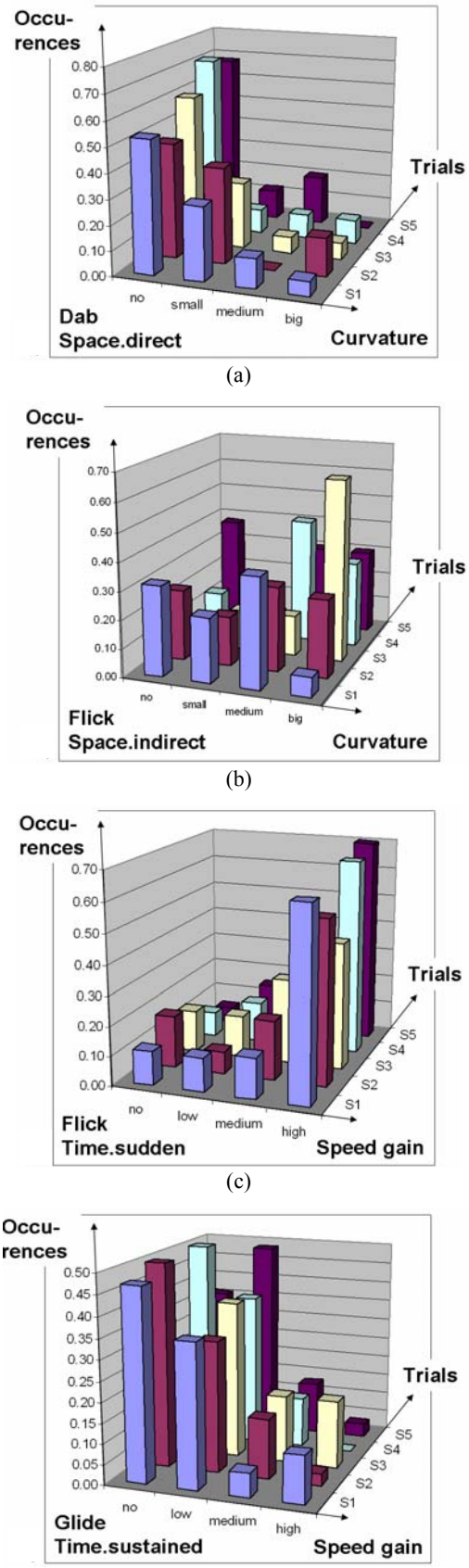
$$\begin{aligned} D &\in \{180, 135, 90, 45, 0, -45, -90, -135\}\ \langle 8 \rangle \\ Vel &\in \{slow, medium, fast\}\ \langle 3 \rangle \\ Acc &\in \{no, low, medium, high\}\ \langle 4 \rangle \\ K &\in \{zero, small, medium, big\}\ \langle 4 \rangle \end{aligned} \tag{2}$$

It can be seen that this approach used only four discrete variables per body part and plane.

3.3 Evaluating the variance between trials

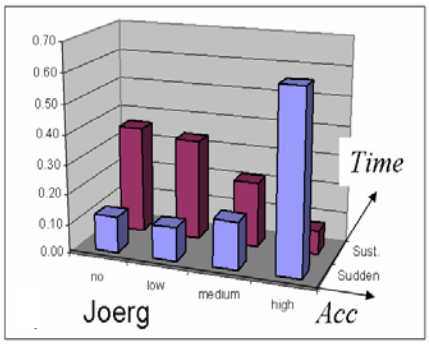
To test the variance of our low-level features over some trials for given *Effort* qualities we conducted the following experiment. By using the table of expressive movements (Table 2) we know the *Effort* qualities that are present for a certain movement ('ByeBye_dab' has *Space.direct*). The low-level feature values (*Vel*, *Acc* and K) are computed for each frame (i) and each trial S. A suitable representation is to create a histogram of the value space of each low-level variable. Figures 10(a) and 10(b) shows the evaluation of the low-level feature *curvature K* when used to describe the *Effort* quality *space* along five trials.

Figure 10 Evaluating the usefulness of the low-level features *curvature* K and *speed gain Acc* for the *Effort* qualities *Space* and *Time* on five trials (see online version for colours)

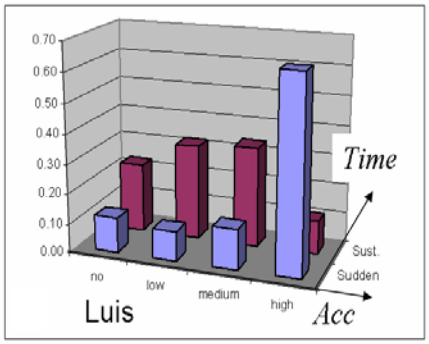


Through the tables of exemplary movements for the *Effort* qualities we can collect sets of movements with a certain *Effort* quality present. As an example all movements that are known to be *Time.sudden* were collected to one set. It can be seen that the majority of the *curvature* K has value *no* when performing the *Space.direct* 'byebye_dab'. In the case of a *Space.indirect* 'byebye_flick' performance, the majority of the *curvature* K has value *big* or *medium*. Another example is shown in Figures 10(c) and 10(d) for the evaluation of *speed gain Acc* for *Time*. It can be seen that nearly all *speed gain Acc* values are showing *high* when performing the *Time.sudden* 'byebye_flick'. In the case of *Time.sustained* represented by 'byebye_glide', the majority is with the *no* or *low* value. The diagrams show that the major tendencies and patterns across the trials allow a relation between the low-level features curvature K and *speed gain Acc* with the *Effort* qualities *Space* and *Time*.

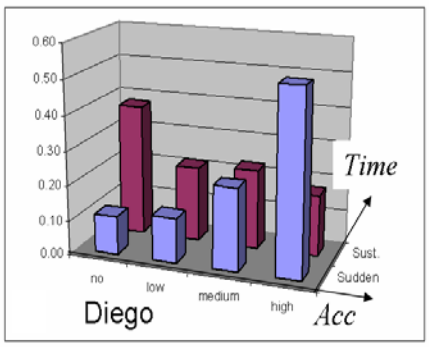
Figure 11 Feature histogram for *speed gain Acc* given by all trials with *Time.sudden* against all trials with *Time.sustained* (see online version for colours)



(a)



(b)



(c)

Note: Each diagram (a), (b) and (c) represents trials of a single person.

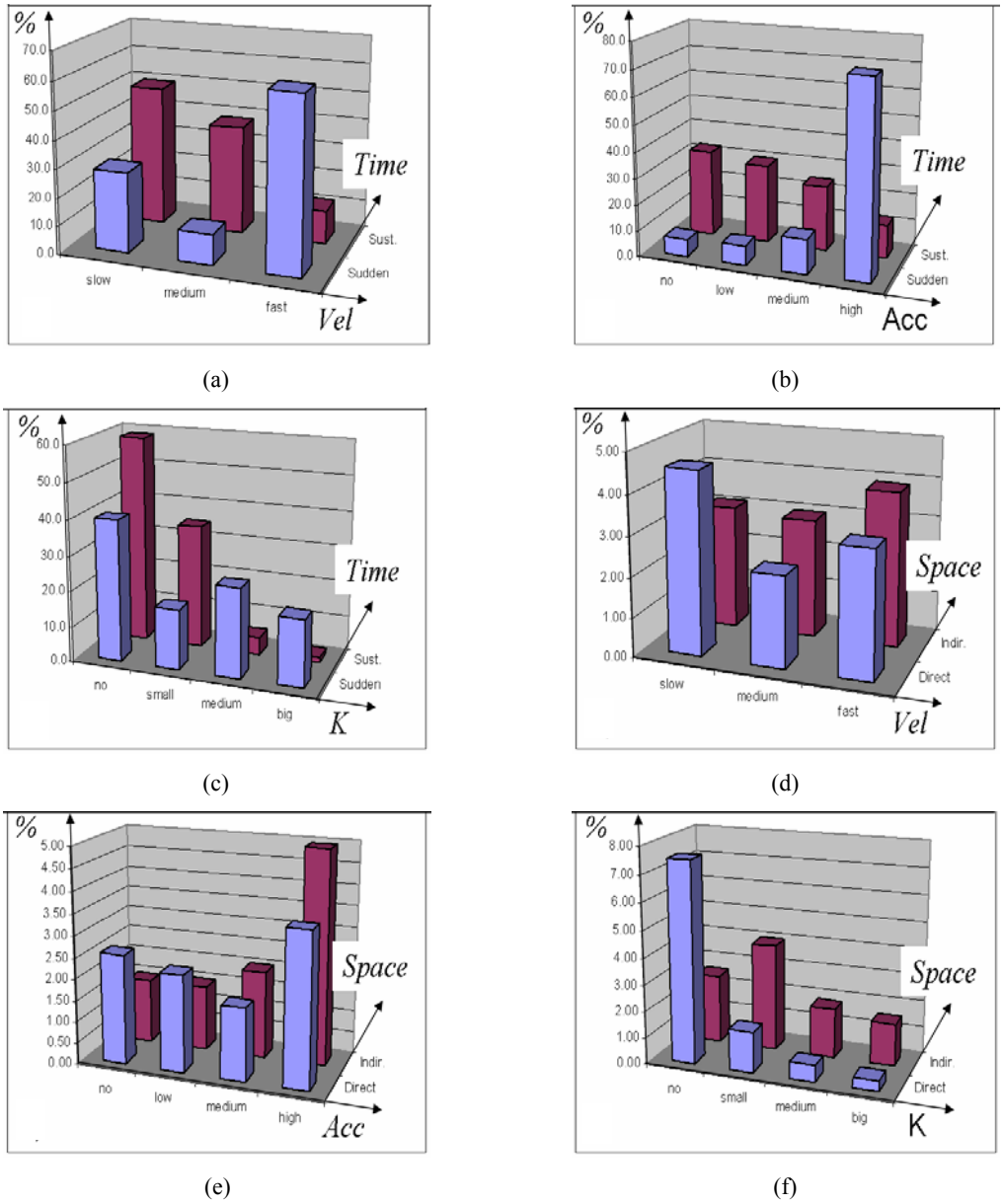
3.4 Evaluating the variance between persons

The diagrams of Figure 10 were performed by the same person ('Joerg'). The next experiment investigates the variance between persons. For this, the former trial histograms are summed to a single movement histogram and afterward movement histograms with the same *Effort* quality are created. Figure 11 shows the feature histogram for *speed gain Acc* as a fusion of all trials with *Time.sudden* (blue bars in the front) against all trials with *Time.sustained* (red bars in the back). Each Figures 11(a), 11(b) and 11(c) represents trials of a single person. It can be seen that the patterns for *Time.sudden* are similar for all three persons, while 'Diego' (Figure 11) shows a higher number of medium values. The patterns for *Time.sustained* appear different as 'Joerg' trials are monotonically decreasing from *no* values, 'Luis' trials are more 'Gaussian' with a mean between *low* and *medium* values and 'Diego' trials have a peak at *no* values. The patterns for *Space.direct* are similar for all three persons, without any significant distinction. Our conjecture is that some groups of patterns are more 'personal' than others. *Time.sudden* and *Space.direct* seem to have a pattern which is unaffected by the actor. For *Time.sustained* and *Space.indirect* qualitative changes can be observed depending on the performing person. It appears that histograms with 'Dirac'-like distributions are less affected by the variance of the performing person. Histograms with 'Gaussian' or monotonic distributions are in general affected by the actor. Though, we can see that an additional variance will be added when using a multiple person approach, the important distinction between opposite *Effort* qualities like *indirect* and *direct* and *sudden* and *sustained* still holds.

3.5 Evaluating the relation of low-level feature to Effort qualities

For the final evaluation, 'person' histograms are also fused. Figures 12(a)–12(c) shows the results for all trials, movements and persons for the *Effort* quality *Time*. The bars in the front (blue) show *Time.sudden* while the bars in the back (red) show *Time.sustained*. For *speed* (*Vel*) we can see a distinct pattern. *Time.sustained* is monotonically decreasing from the low values. For *Time.sudden* we can see that the majority of *speed* is at *fast* values. Our conjecture is that low-level feature *speed* can be used as an evidence for *Time*. For *speed gain* (*Acc*) shown in Figure 12(b), we can also see a monotonically decreasing pattern for *Time.sustained* and a monotonically increasing pattern for *Time.sudden*. These patterns are even more distinct for *speed gain Acc* and can also be used as an evidence for *Time*. The patters for *curvature* K are not distinct as both are monotonically decreasing as shown in Figure 12(c). This evidence can be assumed independent from *Time*.

Figure 12 Behaviour of the low-level features for the *Effort* quality (a)–(c) *Time* and (d)–(f) *Space* (see online version for colours)



Figures 12(d)–12(f) shows the results for the *Effort* quality *Space*. The bars in the front (blue) show *Space.direct* while the bars in the back (red) show *Space.indirect*. For *speed* (*Vel*), we can see no strongly distinct pattern where the values are more or less equally distributed as shown in Figure 12(d). This evidence can be assumed independent from *Space*. Equally distributed values can also be observed for *speed gain* (*Acc*) and *Space.direct* as shown in Figure 12(e). When having *Space.indirect* though, the distribution shows a monotonically increasing behaviour towards high *values*. As *indirect* movements often manifest as curved trajectories the linear acceleration changes constantly. Thus, *speed gain Acc* might be used as an evidence for *Space*. Finally, *curvature K* shows a strong pattern when in comes to *Space.indirect* as nearly all values are *no*. This distinguishes clearly from the 'Gaussian' shaped distribution around the *low* values when it comes to *Space.indirect*. Our conjecture is that the low-level feature *curvature K* can be used as an evidence for *Space*.

3.6 Movement segmentation

Effort qualities determine the most 'expressive' part of a movement. The beginning and the end might have completely different qualities than the main part, thus, it is useful to perform some kind of segmentation. In the domain of gesture analysis, an established model is that of dividing the movement into three phases (Rossini, 2004):

1 *pre-stroke* (preparation)

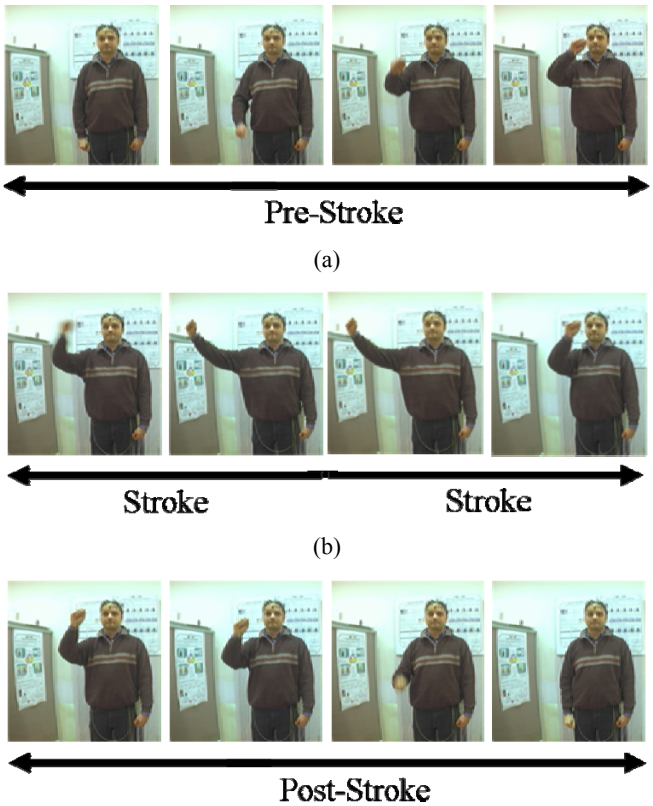
2 *stroke*

3 *post-stroke* (retraction).

We adopted this three-phase model for the segmentation of our expressive movements.

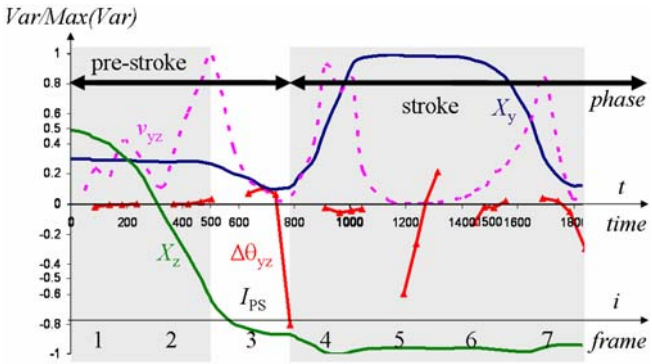
Figure 13 shows 'ByeBye_dab' segmented into three phases. The first four images belong to the *pre-stroke* phase where the hand moves from the rest position upward to position the hand and prepare for the waving. The following four images show the performance of the waving during the *stroke* phase. The final four images belong to the *post-stroke* phase where the hand moves back to the rest position. Gesture recognition systems have often adopted this temporal composition (Starner, 1995; Pavlovic, 1999). In Kettebekov et al. (2002) the phases are called 'phonemes' following the terms used in phonology to describe the principal sounds in human languages.

Figure 13 Gesture phases, (a) pre-stroke (b) stroke (c) post-stroke (see online version for colours)



From a computational viewpoint the problems is the detection of the frame i in which the movement passes from one phase to the next. In the case of the above mentioned three-phase model, we would need to detect two frames of inflection I_{ps} in which the movement passes from (p)re-stroke to (s)troke and I_{sp} where the movement passes from (s)troke to (p)ost-stroke. We constrain the problem in a way that each movement starts from an at ease position and that the movements conclude with the return of both hands to that position. Figure 14 shows the signals of the position X_y and X_z, the velocity v_{yz} and the angular change $\Delta\theta_{yz}$ which related also the first four images of Figure 13. All variables have continuous values and are taken from the *door plane* π_v, i.e., *YZ*. Through inspection of the angular change $\Delta\theta_{yz}$ represented by a red line in Figure 14 we can see that it produces high values during times of rest. Thus, $\Delta\theta_{yz}$ serves in our case as a trigger for I_{ps}. The detection of the second frame of inflection I_{sp} where the movement passes from the *stroke* to the *post-stroke* we whether decide after reaching the rest position what was the last 'peak' of $\Delta\theta_{yz}$.

Figure 14 Signals of the position X_y, the velocity v_{yz} and the angular change $\Delta\theta_{yz}$ (see online version for colours)



4 Bayesian models

The concepts of LMA and the characteristics of our system to track human movements can be mathematically and computationally modelled using a common framework. The Bayesian theory gives us the possibility to deal with incompleteness and uncertainty, make predictions on future events and, most important, provides an embedded scheme for learning. An over view on probabilistic reasoning and its two basic rules is presented in Bessière et al. (2008).

Included in the Bayesian framework are specialised models which have a long tradition in many areas. Some examples of these models are HMMs, Kalman filters and particle filters. Bayesian models have already been used in a broad range of technical applications (e.g., navigation, speech recognition, etc.). Especially in the closely related field of *gesture recognition*, these models have proven their usability (Starner, 1995; Pavlovic, 1999). Recent findings indicate that Bayesian models can also be useful in the modelling of cognitive processes. Researches on the human brain and its computation for perception and action, report that Bayesian methods have proven successful in building computational theories for perception and sensorimotor control (Knill and Pouget, 2004).

4.1 Entropy – a measure of uncertainty

Shannon (1949) extended the information theory by proposing a measure – *entropy* – wherein symbols have unequal probability of occurring. This measure associates information with uncertainty using the concept of

probability. Being X a discrete random variable over a sample space the entropy was defined as:

$$H(X) = -\sum_{X} P(X) \log P(X). \tag{3}$$

In equation (3), the logarithm's base determines in what unit entropy is measured. Hereafter, it will be also assumed the convention 0log0 = 0, since $x \log x \to 0$ as $x \to 0$, which means that adding terms with zero probability does not change the entropy. Entropy is a monotonic function and a formal measure of uncertainty. If all samples of a random experiment have the same probability, i.e., if $p_i = \frac{1}{n}$, wherein n is the number of possibilities (the cardinality of X), H is a monotonic increasing function of n. It can be proven that the entropy of a random variable with n possible outcomes verifies the condition:

$$0 \le H(X_u) \le \log n \tag{4}$$

4.2 Global model

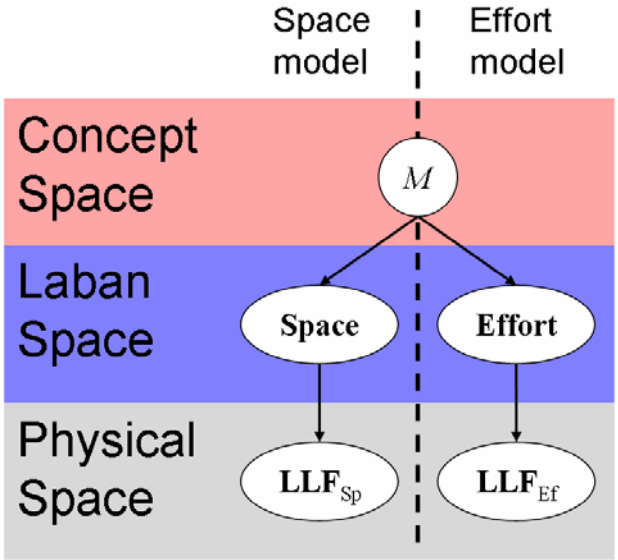
The *Global model* which describes the phenomenon of computational LMA is shown in Figure 15. Having the concept of a movement represented by the variable M certain characteristics will be exhibited through the sets of variables of LMA (**Space** and **Effort**). The sets of LMA can be observed through the set of low-level features **LLF**. This concept is accompanied by different levels of abstraction by introducing the *concept space*, the *Laban space* and the *physical space*. The nodes represent variables (e.g., movement M) and sets of variables (e.g., low-level features **LLF**). The arcs describe the dependencies between the nodes. The movement M represents the parent node which effects the child nodes in the *Laban space*. It appears that given the movement M the sets inside the *Laban space* are independent from each other.

Each of the nodes on the *Laban space* is a parent for the set of low-level features **LLF**. If we assume that for each LMA set n there exists an independent subset of the low-level feature set $\mathbf{LLF}_n$ we can decompose the *Global model* into a number of submodels as shown in Figure 15. The dependencies can also be expressed as a joint distribution and its decomposition as:

$$\begin{aligned} &P(M\ \mathbf{Space}\ \mathbf{Effort}\ \mathbf{LLF}) \\ &\quad = P(M)\, P(\mathbf{Space}|M)\, P(\mathbf{Effort}|M) \\ &\qquad P(\mathbf{LLF}_{Sp}|\mathbf{Space})\, P(\mathbf{LLF}_{Ef}|\mathbf{Effort}) \end{aligned} \tag{5}$$

In the following section the *Space* and *Effort* models will be discussed in detail.

Figure 15 Bayes-net of the *Global model* (see online version for colours)



Notes: Each component of LMA is represented as a node in the *Laban space* and contains a set of variables. The physical space holds the set of all low-level feature variables **LLF**. The variables of the **LLF** set depend on the variables of the *Laban space*. The high level concept of movement is represented by the variable M. By assuming n independent subsets for the low-level feature set $\mathbf{LLF}_n$ joint models can be composed from any combination of the n submodels.

4.3 Space model

The *Space* component of LMA is modelled using the concept of *vector symbols* and was already used in Rett et al. (2008). Referring to the *Global model* (see Section 4.2) the *Space model* is particular as the *vector symbols* represent both, the *Laban space* (*Space*) and the *Physical space* ($\mathbf{LLF}_{Sp}$). We can state:

$$\mathbf{Space} = \mathbf{LLF}_{Sp} \in \{B\} \tag{6}$$

It can be seen that in this work only the *vector symbols* B from the *door plane* π_v are used which has been referred to as a '2-D model' in Rett et al. (2008). In cases where the index bp is used, it corresponds to the *bodypart* like the right hand *rh*, the left hand *lh* and the face *f*. As we describe the spatial pathway of a movement by 'atomic' displacements, we refer to the *vectors symbols* sometimes as *atoms*. Movements are expressed as up, down, left and right resulting in the values *U*, *D*, *L* and *R* respectively and the indication of no movement 0 as shown in equation (7).

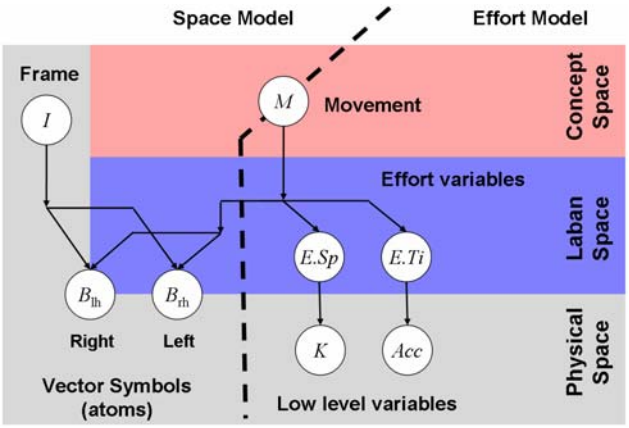
$$\begin{aligned} M &\in \{byebye\_dab, byebye\_glide, \\ &\quad byebye\_flick, byebye\_float\} \langle 4 \rangle \\ I &\in \{1, \ldots, I_{max}\} \langle I_{max} \rangle \\ B_{bp} &\in \{O, U, UR, R, DR, D, DL, L, UL\} \langle 9 \rangle \end{aligned} \tag{7}$$

The *Space model* assumes that each movement $M = m$ produces certain *atoms* $B_{bp} = b_{bp}$ at a certain point in time, i.e., frame $I = i$ and for a certain *Bodypart* bp. In this model, a certain movement m is 'causing' the *atoms* b at the frame i. The *evidences* that can be measured are the *atoms* b and the frame i. The model might be applied to any number of body parts bp which are treated as independent evidences as thus expressed through a product as shown in the joint distribution of equation (8).

$$P(M\ I\ B) = P(M)\ P(I) \prod_{bp} \{P(B_{bp} \mid M\ I)\} \tag{8}$$

The left side of Figure 16 shows the corresponding representation in a Bayes-net.

Figure 16 The movement *M* belongs to the *concept space* (see online version for colours)



Notes: *Space model* (left side): the *vector symbols* are part of both, the *Laban space* and the *Physical space*. Their instances are the left (*lh*) and right hand (*rh*). The frame *I* is associated with the physical space only. *Effort model* (right side): the variables *Effort Space E.Sp* and *Effort Time E.Ti* belong to the *Laban space*, while *curvature* K and *speed gain Acc* belong to the *Physical space*.

4.4 Effort model

The *Effort model* describes the dynamic aspects of the movement. It relates the low-level features *speed gain Acc* and *curvature* K to the *Effort* qualities *Time* (*E.Ti*) and *Space* (*E.Sp*). In order to not confuse the *Space* component from the previous section with the *Space* quality of the *Effort* component, all variable symbols of *Effort* have a leading *E.* before the quality. As defined in the *Global model* two sets of variables are used in the model:

$$\mathbf{LLF}_{Ef} \in \{Acc, K\} \quad \mathbf{Effort} \in \{E.Ti, E.Sp\} \tag{9}$$

The relation between the two sets has already been investigated, established and developed in Section 3. The *concept space* relates the *Effort* qualities to a specific movement M. This has been introduced in Section 2.2 and 2.3. The *Effort model* is related with a specific plane and body part where the *Effort* qualities can be detected best. The variables and their sample space are shown in (10):

$$\begin{aligned} Acc &\in \{no, low, medium, high\} \langle 4 \rangle \\ K &\in \{no, small, medium, big\} \langle 4 \rangle \\ E.Sp &\in \{direct, indirect\} \langle 2 \rangle \\ E.Ti &\in \{sudden, sustained\} \langle 2 \rangle \end{aligned} \tag{10}$$

Each movement M will produce a certain set of *Effort* qualities during a certain phase. Thus, we have a conditional dependency of *Effort Space E.Sp* and *Effort Time E.Ti* from the movement M as can be seen in the right side of Figure 16.

The *Effort* variables cannot be directly measured but observed through some low-level features (i.e., $\mathbf{LLF}_{Ef}$). Thus, there is a dependency of the non-observable variables from the *Effort* set and $\mathbf{LLF}_{Ef}$. The joint distribution can be expressed as:

$$\begin{aligned} &P(M\ E.Sp\ E.Ti\ Acc\ K) \\ &\quad = P(M)\ P(E.Sp \mid M)\ P(E.Ti \mid M) \\ &\qquad P(Acc \mid E.Ti)\ P(K \mid E.Sp). \end{aligned} \tag{11}$$

4.5 Joint model using Space and Effort

With the previously discussed models many combinations can be built, tested and compared with each other. As defined in the *Global model* four sets of variables are used in the model:

$$\begin{aligned} &\mathbf{Space} = \mathbf{LLF}_{Sp} \in \{B\} \\ &\mathbf{Effort} \in \{E.Ti, E.Sp\} \quad \mathbf{LLF}_{Ef} \in \{Acc, K\} \end{aligned} \tag{12}$$

Figure 16 shows the combination of the *Space* and *Effort model*. The *Space* and *Effort model* are coupled through the movement variable M. This means that a certain movement $M = m$ will exhibit certain *atoms* $B = b$ and certain *Effort* variables $E.Ti = e.Ti$, $E.Sp = e.Sp$ along the time. The temporal dependency is modelled through the variables frame I. The joint distribution can be expressed as:

$$\begin{aligned} &P(M\ I\ B\ E.Sp\ E.Ti\ Acc\ K) \\ &\quad = P(M)\ P(I) \prod_{bp} \{P(B_{bp} \mid M\ I)\} \\ &\qquad P(E.Sp \mid M) \quad P(E.Ti \mid M) \\ &\qquad P(Acc \mid E.Ti) \quad P(K \mid E.Sp). \end{aligned} \tag{13}$$

4.6 Learning of probability tables

The previous section concluded with a joint distribution made of several distributions, e.g., probability of *vector symbols* from *door plane* given *movement* and *frame* $P(B|M\,I)$. Some of those distributions need to be 'learned'. In our case, 'learning' means that trials with a known label are fed into the system which in return identifies the parameters of a chosen distribution.

One type of distribution that can be chosen is a histogram. The probability distribution can be learned by counting the observations a variable has a certain value. For a finite number of discrete values the process can be described as building a histogram. By dividing the counts for each value i of the variable $V(V=i)$ by the total number of samples n a probability distribution can be computed. The assumptions that apply are:

1 All samples n come from the same phenomenon.

2 All samples are from a single variable V.

3 The order of the samples is not important.

When learning a probability distribution through the histogram, some values of V might have zero probability simply because they have never been observed. Whenever these values occur in the later classification stage, the corresponding hypothesis(es) will receive also a zero probability. In continuous classifiers that are based on multiplicative update of beliefs, this leads to an immediate and definite out-rule of the hypothesis(es). One way of solving this is to use an equation which produces a minimum probability for non-observed evidences. Equation (14) is based on the *Laplace succession law*.

$$P(V=i)=\frac{n_i+1}{n+\lfloor V\rfloor} \tag{14}$$

The minimum probability which is produced when no observation has happened $(n_i=0)$ is $P_{min}=1/\left(n+\lfloor V\rfloor\right)$. Taking the *atom* variable B_{rh} which has nine values $\lfloor V\rfloor=9$ as an example, it can be seen that by learning from six samples $n=6$ each non-observed value will receive a probability of $P(V)=0.0667$ for all values i where $n_i=0$.

4.7 Questions for classification

Classification is the final step after the model has been established and the tables have been learned: *classification accesses the knowledge of model and learning through inference*. Given our joint distribution:

$$P(M\,I\,B\,E.Sp\,E.Ti\,Acc\,K) \tag{15}$$

we need to formulate a *question*, i.e., what we want to classify and what we can observe. The following set of equations shows particularly interesting *questions*.

Question (16) asks for the distribution of the variable movement M knowing the frame I and the *atom* B from the *door plane* π_v.

$$P(M|I\,B) \tag{16}$$

This represents the classification of a movement taking into account the *Space* component of LMA from a frontal view.

Question (17) asks for the distribution of the variable *Effort.Time E.Ti* knowing the *speed gain Acc*.

$$P(E.Ti\,|\,Acc) \tag{17}$$

This represents the classification of the *Effort* quality *Time* knowing the *speed gain*. With this *question* we are able to label a certain movement with the *Effort* quality *Time*. A similar question can be established for *Space* knowing the *curvature* K.

Question (18) asks for the distribution of the variable movement M knowing the two *Effort* variables *Effort.Space E.Sp* and *Effort.Time E.Ti*.

$$P(M\,|\,E.Sp\,E.Ti) \tag{18}$$

This represents the classification of a movement taking into account only the *Effort* component of LMA by using the two qualities *Space* and *Time*. With the latter *Question* (18) and the former *Question* (16) we are able to compare movement classifications based on components from the kinematic and the non-kinematic group (see Section 2).

Question (19) asks for the distribution of the variable movement M knowing the two *Effort* variables *Effort.Time E.Ti*, *Effort.Space E.Sp* and the frame I and the *atom* B from the *door plane* π_v.

$$P(M\,|\,I\,B\,E.Sp\,E.Ti) \tag{19}$$

This can be seen as the 'ultimate' *question* asked in this work: the classification of a movement taking into account components from the kinematic and the non-kinematic group of LMA represented by *Space* and *Effort*.

4.8 Continuous classification of movements

After the interesting *questions* have been defined the problem of continuous update will be tackled. Continuous update of believe is a desirable characteristic of human-machine interaction. With this the system can continuously refine its classification results through the newly incoming evidences.

The concept for *Question* (16) $P(M\,|\,I\,B)$ has already been developed in Rett et al. (2008) and yielded:

$$P(M_{n+1}\,|\,i_{1:n+1}b_{1:n+1})\propto P(M_n)P(b_{n+1}\,|\,M\,i_{n+1}) \tag{20}$$

We can see that the prior of step $n+1$ is the result of the classification of step n. Given a sufficient number of evidences (*atoms*) and assuming that the learned tables represent the phenomenon sufficiently good, the classification will converge to the correct hypothesis.

By adding the evidences from *Effort* we get the answer for *Question* (19)

$$P(M_{n+1} \mid i_{1:n+1} b_{1:n+1} E.Sp_{1:n+1} E.Ti_{1:n+1}) \propto \\ P(M_n) P(b_{n+1} \mid M\, i_{n+1}) \\ P(E.Sp_{n+1} \mid M) P(E.Ti_{n+1} \mid M) \tag{21}$$

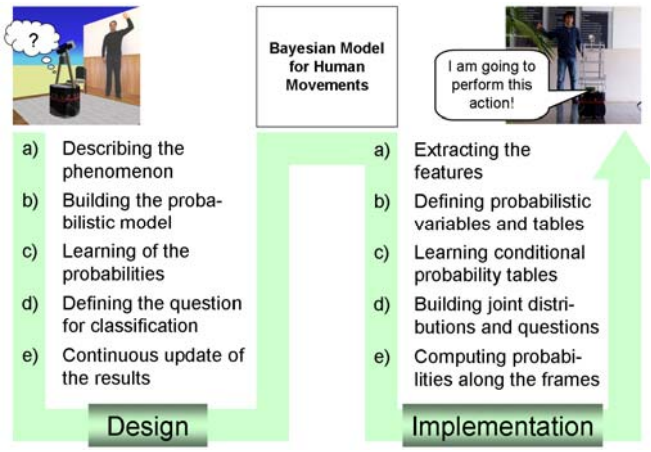
The final classification result is given by the *maximum a posteriori* (MAP) method. In some cases, it might be useful to produce a result before the final frame I_{max} is reached. In order to have a measure which determines the right moment, we have chosen the already presented *entropy* $H(M)$. For a given evolution of a probability distribution over time, the 'cut-off' frame is triggered where the entropy ratio $H(M) \mid H(M)_{max}$ drops under a threshold of 0.1. We define two levels of certainty:

- *Uncertain*: $1 \geq H(M) \mid H(M)_{max} < 0.1$.
- *Certain*: $0.1 \geq H(M) \mid H(M)_{max}$.

5 Movement recognition system

The previous sections of this article reflect the five steps of designing a probabilistic model as shown in Figure 17.

Figure 17 The five steps of designing a probabilistic model (shown in the previous sections) relate to the five steps of implementing the processes (presented in this section) (see online version for colours)



The implementation of the processes and its results can also be organised in five steps.

a The first step is the extraction and computation of the low-level features and was already addressed in Section 3.

b The second step is the definition of the probabilistic variables and conditional probability tables. Those definitions can be derived directly from the second step of the design process, i.e., building of the probabilistic model. Each step in implementation has, in fact, a corresponding step in design.

c Section 5.1 will present the third step, i.e., the process of learning and the resulting conditional probability tables. The discussion will show that the Bayesian approach allows a comparison of those learned tables, even before entering the classification stage.

d The fourth step, in which the joint distributions and questions are built can also be derived from the fourth step of the design process.

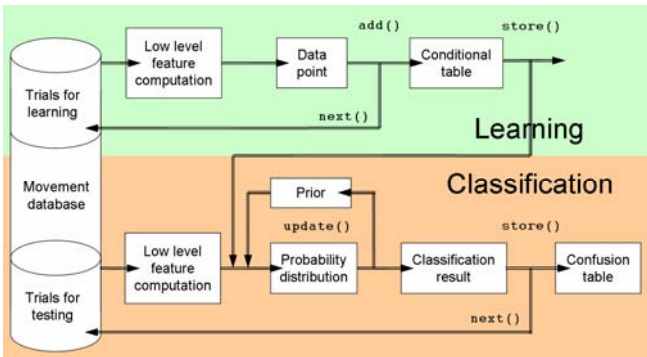
e Section 5.2 has its focus on the fifth step of implementation, i.e., the computation of probabilities along the frames (time). The discussion on the evolution of probabilities is presented and gives some insight on certainty for a specific trial.

With this the process of implementation is concluded. To emphasise the important characteristic of the system it is called online movement anticipation and recognition (OMAR) system.

5.1 Learning conditional probability tables

Figure 18 (top) shows the flow chart of the learning process. From the movement database (HID) a set of trials for learning is chosen and fed into the system for low-level feature extraction. The database consists of five trials per person and movement. Three trials are usually chosen for learning. Each trial produces one data point per feature and frame. Learning based on a histogram approach creates conditional probability tables simply by adding those points until all trials are processed.

Figure 18 (top) learning process*, (bottom) classification process** (see online version for colours)

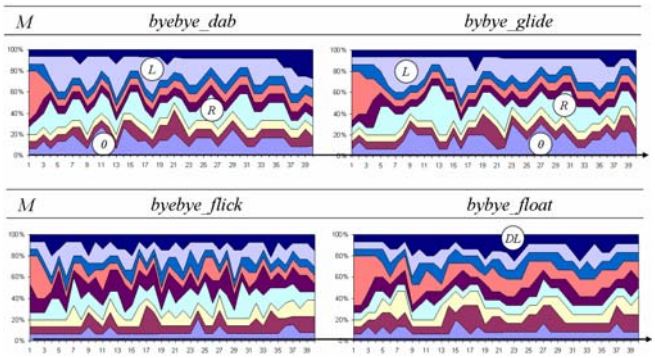


Notes: *From a trial for learning low-level features are extracted and 'added' to the histogram. The conditional table is 'stored' after all trials are processed.
**The low-level features are extracted from a trial for testing. Each frame the probability distribution is computed and the prior 'updated'. The final classification result is stored for each trial in a confusion table.

The set of movements under investigation is the 'byebye' set. When stacking the probabilities for each value one over each other, patterns can be observed along the time given by the frame I and between the movements M. Figure 19

shows the *Space* patterns for the movements' *byebye_dab*, *byebye_glide*, *byebye_flick* and *byebye_float*. The movement patterns look similar and especially *byebye_dab* and *byebye_glide* can be easily confused. Both have strong probabilities at left *L* and right *R* *atoms* together with always apparent zero *atoms* 0. This is different to the *byebye_flick* and *byebye_float* movements which have zero *atoms* only at minimum probability. It also appears that the other atoms are distributed more uniformly.

Figure 19 Learned conditional probability table $P(B \mid M\,I)$ for movement set 'byebye' (see online version for colours)



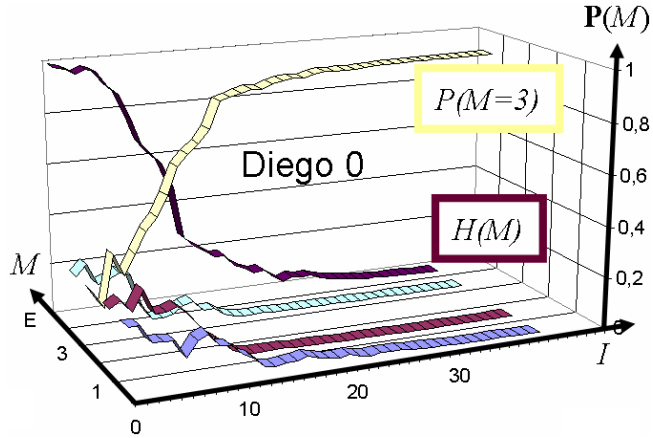
5.2 Continuous classification and certainty

The fifth and final step of implementation concerns the investigation of the evolution of probabilities and the confusion table that can be obtained for all trials of the test set. Figure 18 (bottom) shows the flow chart of the classification process. The inner loop of continuous update produces the evolution of probabilities. The outer loop of next trial produces the confusion table. Figure 18 (top) shows that classification uses the same process for the computation of low-level features as learning before. With the low-level features and the previously stored conditional probability table, it is possible to compute the desired probability distribution. This goes according to the defined joint distribution and the desired question. Through feeding in (replacing) the result of the probability distribution as the new prior, a continuous update of the classification results for all frames can be obtained. The result of the 'last' frame gives the final result and while looping through all trials for testing, confusion table can be built.

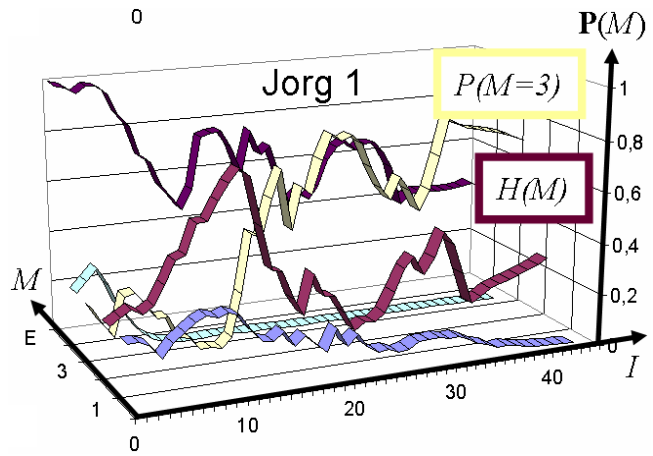
We can conclude that the two processes of learning and classification are based on the same type of observations as shown in Figure 18. The previously presented scheme starts by learning and, after the conditional probability table has been build, continues with classification. An important feature of Bayesian histogram learning is that it can be 'switched back' at any time to learn new and more data. When applied to, e.g., a social robot, the system creates a skill that can be expressed as 'lifelong learning'. When placed in a human-robot interaction scenario the robot can use this skill to continuously learn new data during his daily operation.

The evolution of movement probabilities and the certainty of the belief are shown in Figure 20. The two cases classify trials from the 'byebye' set. Figure 20(a) shows a case where the correct movement is classified fast and confident. The probability distribution for the four movements starts with a uniformly distributed prior (25%) and the maximum entropy ratio (1). In the 21st frame, the belief passes from *uncertain* to *certain*. The trial is finally correctly classified as *byebye_flick* with a probability of 100%.

Figure 20 Evolution of the movement probabilities $P(M)$ along the time (i) for two trials of the 'byebye' set (see online version for colours)



(a)



(b)

A different case is shown in Figure 20(b), where the belief never reaches the *certain* state though it concludes with the correct movement *byebye_flick*. It can be seen that the entropy was not monotonically decreasing and surpassed the minimum around the 13th frame. The behaviour of the system to actually change its belief while producing a peak of entropy is a natural characteristic of this type of Bayesian classifier.

5.3 Classification of movements using only Space

This section conducts the experiment of movement classification using the *Space model* from a single (*door plane* π_v) projection. The obtained results can then be

compared to models using *Effort*, to models using *Space* and *Effort*, and so on. The results for all trials are shown in Table 4. By using *Space* 15 of 60 trials are classified wrongly leaving a recognition rate of 75%. Though the number of hypotheses is low (only four movements) the recognition rate is also low. The confusion between *byebye_dab* and *byebye_glide* was already anticipated given the learned tables from the previous section. The 'byebye' set provides only few spatial distinctiveness.

Table 4 Confusion table for classification of movements using *Space*

Movement	*1*	*2*	*3*	*4*	Σ_e
1 dab	3	12			12
2 glide		15			0
3 flick	2		12	1	3
4 float				15	0
					15

5.4 Classification of Effort qualities for movements

This and the following sections will evaluate the effect on the performance of the OMAR system when the *Effort* model is introduced. Each of the four movements of the dataset 'bye-byes' has a special 'flavour' in the way it is performed. This flavour is defined by *basic effort actions* know from LMA (see Section 2.2). Table 5 shows a fragment of the *basic effort action drives* for the *Effort* qualities *Time* and *Space*. In order to get results on movement classification using *Effort*, first *Effort* itself needs to be classified. Thus, this section presents the performance of a classifier for *Effort* which will be used in the following section to classify movements.

Table 5 Subset of the *basic effort action drives*

Action	*Space*	*Time*
Dab	Direct	Sudden
Flick	Indirect	Sudden
Glide	Direct	Sustained
Float	Indirect	Sustained

For this the probabilities for the two *Effort* qualities will be calculated each frame a movement happens (non-zero velocity frame). Each *Effort* quality uses only one low-level descriptor: *Time E.Ti* is associated to *speed gain Acc* and *Space E.Ti* is associated to *curvature* K. The belief will be updated and converges to the hypothesis that explains the observations (*Acc* and K) best. Figure 21 shows the probabilistic evolution of the two *Effort* qualities for the trial *byebye_glide Jorg1*. In this case the qualities converge early to the correct hypotheses.

The results for all trials are shown in Table 6. For the *Effort* quality *Space* 9 results and for the *Effort* quality *Time* 3 results are classified wrongly. This yields a recognition rate of 90% for the former and 70% for the latter. This suggests that the low-level feature *speed gain Acc* represents a stronger descriptor for *Time* than *curvature* K for *Space*.

Figure 21 Evolution of the probabilities for *Effort*, *Time* and *Space* qualities by the submodel of the trial *bye-bye_glide Jorg1* (see online version for colours)

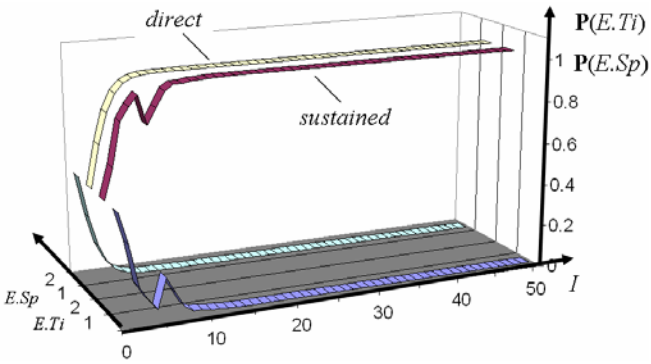


Table 6 Confusion table for classification of *Effort*

Movement	*1*	*2*	Σ_e
1 direct	25	5	5
2 indirect	4	26	4
1 sudden	29	1	1
2 sustained	2	28	2
			12

5.5 Classification of movements using only Effort

The next step is a model which classifies movements M given the results from the previous model as evidences. For this, the OMAR system uses the function best() provided by the probabilistic library. The function returns the variable value with the highest probability. With this the resulting *Effort* values will be used as certain (hard) evidences.

Figure 22 *Effort Time* and *Space* variables as a stream of hard evidences computed by the submodel of the trial *byebye_dab Diego2* (see online version for colours)

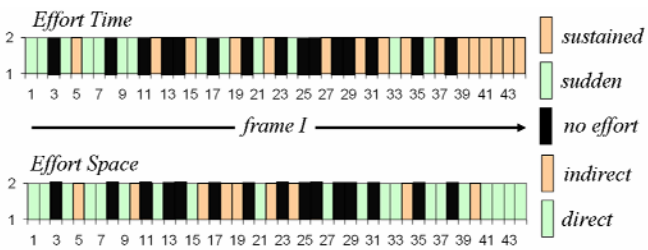
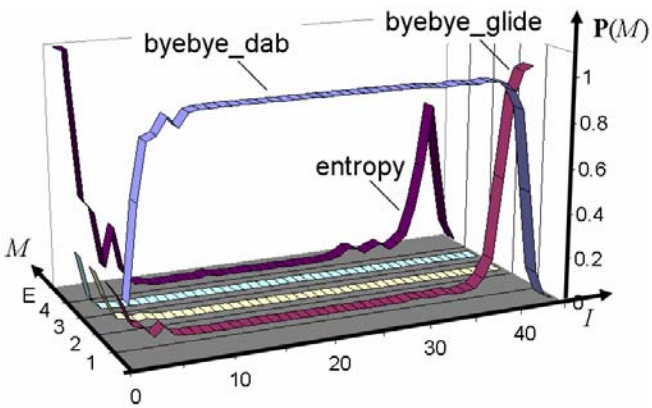


Figure 22 shows the *Effort Time* and *Space* variables as a stream of hard evidences computed by the submodel of the trial *byebye_dab Diego2*. It can be seen that for the trial *byebye_dab Diego2* not every frame the correct *Effort* qualities are calculated. The ideal stream should only consist of light green evidences, i.e., *Time.sudden* and *Space.direct*. The black rectangles represent frames with zero velocity where the probability remains constant. The computation of hard evidences using the function best()

obscures the uncertainty of the (soft) *Effort* qualities. Figure 22 also shows that *Effort Time* becomes more sustained (light red) close to the end of the trial. This will finally also result in a change of belief concerning the type of movement.

Figure 23 shows the probability evolution of the trial *byebye_dab Diego2*. The OMAR system classifies correctly $M = byebye\_dab$ most of the time and only changes its belief in favour of $M = byebye\_glide$ in the last eight frames. This behaviour can be observed for three more trials and seems to indicate the performers 'anticipation' of the end.

Figure 23 Evolution of the probabilities for movement M using *Effort* and *Space* of the trial *byebye_dab Diego2* (see online version for colours)



The results for all trials are shown in Table 7. By using *Effort* 14 of 60 trials are classified wrongly leaving a recognition rate of 77%. We can see that the recognition rate has improved slightly compared to using a *Space model*.

Table 7 Confusion table for classification of movements using *Effort*

Movement	*1*	*2*	*3*	*4*	Σ_e
1 dab	13	3	2	2	7
2 glide		15			0
3 flick			12	3	3
4 float	4			11	4
					14

5.6 Classification of movements using Space, Effort and entropy

The next step is to use the *Joint model* as presented in Section 4.5 to fuse the evidences from the *Space* and *Effort model*. The results shown in Table 8 where obtained by using not all *Effort* qualities and low-level features. By using both, *Space* and *Effort* 10 of 60 trials are classified wrongly leaving a recognition rate of 83%. It appears that with a *Joint model* of *Effort* and *Space* a better recognition rate can be obtained as with each of the single models.

Some of the trials that have been classified wrongly are of the type already shown in Figure 23. The system classifies correctly most of the time and only changes its belief during last frames. A reasonable solution is to let the system decide when it is 'certain' about its classification. The right measure for this certainty is the entropy. A value of 0.1 was chosen for the entropy to finish the classification. Table 9 shows the results for deciding based on a low entropy level (0.1). When deciding based on a low entropy level only five of 60 trials are classified wrongly leaving a recognition rate of 92%. The best recognition rate can be obtained by deciding at the first appearance of certainty.

Table 8 Confusion table for classification of movements using *Space* and *Effort*

Movement	*1*	*2*	*3*	*4*	Σ_e
1 dab	9	6			6
2 glide		15			0
3 flick			13	2	2
4 float		2		13	2
					10

Table 9 Confusion table for classification of movements using *Space* and *Effort* and entropy

Movement	*1*	*2*	*3*	*4*	Σ_e
1 dab	13	2			2
2 glide		15			0
3 flick			14	1	1
4 float		2		13	2
					5

5.7 Scenario Nicole@Play

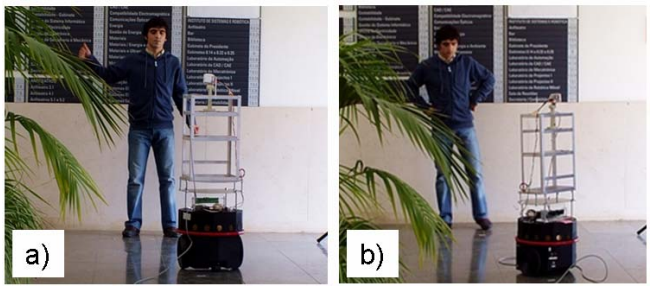
The current target application for the continuous classification of movements is the social robot 'Nicole'. The social robot 'Nicole' is designed as an autonomous platform with which human-robot interaction can be investigated. The vision system is using a static single camera. The system for continuous classification of movements runs on a notebook where the process of tracking can be observed. The navigation system is held by a PC board inside the Scout platform.

The complete scenario 'Nicole@Play' includes also stationary PC which runs the script interaction. It holds way-points for navigation, asks for the presence of a person and the result of the movement recognition. The different processes exchange information by using 'sockets'. Version 1.0 of the system architecture is a redefined version the gesture perception system (GP-System) presented in Rett and Dias (2005). Apart from the already discussed modules for perception the system architecture also includes the 'action planner' which controls the sequential execution of the tasks inside the interaction scenario. It holds the script that tells in which way the robot acts upon the perceptions.

In our first trials, Nicole was using audio outputs like asking for confirmation on the recognised commands and robot movements. Movies of the trials can be downloaded from the project's web page http://paloma.isr.uc.pt/nicole/.

The scenario was tested in a natural environment at the entrance hall of the Department of Electrical Engineering, University of Coimbra in June 2006. Figure 24 shows some of the states during the interaction. After Nicole has been called she navigates to the position where she expects the user (Phase 1: long distance approach). She will then, look around in search for a person (Phase 2: user search). The first person she detects will be approached (Phase 3: short distance approach). After taking the optimal interaction position she will greet the user and ask for a gesture (Phase 4: initiate interaction). In the next phase, Nicole will observe and anticipate the movement of the user's hand(s) (Phase 5: tracking and gesture recognition). After being certain about the perceived gesture Nicole will perform a related action (e.g., turning around) (Phase 6: action). After this, Nicole will end up in phase two or three start all over again.

Figure 24 Nicole V1.0 interacting at the entrance of our department (a) Phase 5: tracking and gesture recognition (b) Phase 6: action (see online version for colours)



6 Discussion and conclusions

6.1 Resume

The work presented in this thesis started with the premise that the field of C-HMA is in need of an annotated database fur human movements. Low-level features like 'acceleration' can easily be extracted by machines and action descriptors like 'dabbing paint on a canvas' can easily be understood by humans. A good descriptive language needed to be chosen that provides the labels on a medium level in between features and actions. Section 2 shows that LMA is a good choice for this descriptive language. The section presents a thorough overview over the properties and capabilities of each component and their relation to each other. This part concludes with examples from our HID to outline the applicability of LMA for an annotated database.

This work will contribute to applications like 'social robots', 'smart rooms' or 'rehabilitation'. The required technical solution brings together a single camera mounted on a mobile platform, multiple cameras mounted on the walls of a room and high precision data from an active sensor. Section 3 bases the computation of the low-level features on two very different sensor types, i.e., single camera and active sensor. The sensor data can be registered by calibrating the two devices which allows working with a database of rich 3-D position data and sensory input from 2-D projections. The low-level features are extracted from trials of our database and the evaluation shows that these features are useful as evidences for LMA descriptors.

To extract the LMA descriptors automatically the Bayesian framework is used as presented in Section 4. It:

1 presents the models as Bayesian nets which allows multidisciplinary evaluations

2 can be designed in a modular fashion so the influence of each component can be studied

3 takes into account that LMA is based on human observations where incompleteness and uncertainty are issues.

Section 5 presents the implementation which proves the feasibility of this approach. The probabilistic approach provides the learned data in a way that allows its visual inspection and evaluation. With this, expected results for classification can be anticipated. The chosen histogram-based approach for learning provides a simple way of adding data points. The characteristics to add data at any time opens the possibility for a continuously learning artificial agent. As a benefit of the modularity of the OMAR system results for movement classification can be presented and compared separately for *Space*, *Effort* and joining *Space* and *Effort*. Also a 'stand-alone' classification of the *Effort* qualities is possible. One observed characteristics of the implemented type of Bayesian classifier was that the system actually changed its belief while producing a peak of entropy. The evolution of the OMAR system for the 'bye-bye' set when using the different models is:

1 a recognition rate of 75% when using only *Space*

2 a recognition rate of 77% when using only *Effort*

3 a recognition rate of 83% when using both

4 a recognition rate of 92% when deciding based on a low entropy level, i.e., at the first appearance of certainty.

To prove that the system could also perform well in a natural indoor environment a series of demonstration of the social robot 'Nicole' was conducted since Summer 2006 at various expositions.

6.2 Future works

The main goals of the future research will be to establish LMA as a general tool for the evaluation of human movements and provide those communities that collect large amounts of experimental data with technical solutions for labelled datasets.

The research will be justified by showing that rehabilitation processes do benefit from evaluations based on LMA. That comparison of experimental data with very distinct experimental setups is possible by using the descriptors of LMA. Data from computational LMA opens the possibility to cluster motor deficits and neurological disorders that are similar with regards to LMA. A successful research towards these goals must be based on the creation of a large database, the implementation of systems to collect this data, an interface design that appears 'natural' to the patient and an intensive multidisciplinary discussion and cooperation.

Acknowledgements

The authors would like to thank Luis Santos from the Institute of Systems and Robotics, Coimbra for his work on the implementation. This work is partially supported by FCT-Fundação para a Ciência e a Tecnologia Grant #12956/2003 to J. Rett and by the BACS-project-6th Framework Programme of the European Commission contract number: FP6-IST-027140, Action line: Cognitive Systems to J. Rett, J.M. Ahuactzin and L. Santos.

References

Aggarwal, J.K. and Cai, Q. (1999) 'Human motion analysis: a review', *CVIU*, Vol. 73, No. 3, pp.428–440.

Badler, N.I., Phillips, C.B. and Webber, B.L. (1993) *Simulating Humans: Computer Graphics, Animation and Control*, Oxford Univ. Press.

Bartenieff, I. and Lewis, D. (1980) *Body Movement: Coping with the Environment*, Gordon and Breach Science, New York.

Benesh, R. and Benesh, J. (1983) *Reading Dance: The Birth of Choreology*, McGraw-Hill Book Company Ltd.

Bessière, P., Laugier, C. and Siegwart, R. (2008) 'Probabilistic reasoning and decision making in sensory-motor systems', *Springer Tracts in Advanced Robotics*, Vol. 46, Springer, Berlin.

Breazeal, C. (2003) 'Toward sociable robots', *Robotics and Autonomous Systems*, Vol. 42, pp.167–175.

Bregler, C. (1997) 'Learning and recognising human dynamics in video sequences', in *Conference on Computer Vision and Pattern Recognition*, San Juan, Puerto Rico.

Burgard, W., Cremers, A.B., Fox, D., Hahnel, D., Lakemeyer, G., Schulz, D., Steiner, W. and Thrun, S. (1998) 'The interactive museum tour-guide robot', in *AAAI/IAAI*, pp.11–18.

Burgard, W., Cremers, A.B., Fox, D., Hahnel, D., Lakemeyer, G., Schulz, D., Steiner, W. and Thrun, S. (1999) 'Experiences with an interactive museum tour-guide robot', *Artificial Intelligence*, Vol. 114, Nos. 1–2, pp.3–55.

Chi, D., Costa, M., Zhao, L. and Badler, N. (2000) 'The emote model for effort and shape', in *SIGGRAPH 00, Computer Graphics Proceedings, Annual Conference Series, ACM SIGGRAPH*, pp.173–182, ACM Press.

Dautenhahn, K. (1998) 'The art of designing socially intelligent agents: science, fiction and the human in the loop', *Applied Artificial Intelligence*, Vol. 12, pp.573–617.

Diard, J., Bessière, P. and Mazer, E. (2003) 'A survey of probabilistic models, using the Bayesian programming methodology as a unifying framework', in *Proc. of the Int. Conf. on Computational Intelligence, Robotics and Autonomous Systems*, Singapore (SG).

Eshkol, N. and Wachmann, A. (1958) *Movement Notation*, Weidenfield and Nicholson.

Fong, T., Nourbakhsh, I. and Dautenhahn, K. (2003) 'A survey of socially interactive robots', *Robotics and Autonomous Systems*, Vol. 42, pp.143–166.

Foroud, A. and Whishaw, I.Q. (2006) 'Changes in the kinematicstructure and non-kinematic features of movements during skilled reaching after stroke: a Laban Movement Analysis in two case studies', *Journal of Neuroscience Methods*, Vol. 158, pp.137–149.

Gavrila, D.M. (1999) 'The visual analysis of human movement: a survey', *CVIU*, Vol. 73, No. 1, pp.82–98.

Golani, I. (1976) 'Homeostatic motor processes in mammalian interactions: a choreography of display', *Perspectives in Ethology*, Vol. 2, pp.69–134, Plenum Press, New York.

Kendon, A. (2004) *Gesture: Visible Action as Utterance*, Cambridge University Press.

Kettebekov, S., Yeasin, M. and Sharma, R. (2002) 'Prosody based co-analysis for continuous recognition of coverbal gestures', in *International Conference on Multi-modal Interfaces (ICMI'02)*, Pittsburgh, USA, pp.161–166,.

Knill, D.C. and Pouget, A. (2004) 'The Bayesian brain: the role of uncertainty in neural coding and computation', *TRENDS in Neurosciences*, Vol. 27, pp.712–719.

Little, M. E. and Marsh, C. G. (1992) *La Danse Noble, An Inventory of Dances and Sources*, Broude Brothers Ltd.

Longstaff, J.S. (2001) 'Translating vector symbols from Laban's (1926) choreographie', in *26. Biennial Conference of the International Council of Kinetography Laban*, ICKL, Ohio, USA, pp.70–86.

Moeslund, T., Hilton, A. and Kruger, V. (2006) 'A survey of advances in vision-based human motion capture and analysis', *CVIU*, Vol. 103, Nos. 2–3, pp.90–126.

Moeslund, T.B. and Granum, E. (2001) 'A survey of computer vision-based human motion capture', *CVIU*, Vol. 81, No. 3, pp.231–268.

Nakata, T., Mori, T. and Sato, T. (2002) 'Analysis of impression of robot bodily expression', *Journal of Robotics and Mechatronics*, Vol. 14, pp.27–36.

Nourbakhsh, I., Kunz, C. and Willeke, T. (2003) 'The mobot museum robot installations: a five year experiment', in *IROS 2003*.

Otero, N., Knoop, S., Nehaniv, C., Syrda, D., Dautenhahn, K. and Dillmann, R. (2006) 'Distribution and recognition of gestures in human-robot interaction', in *The 15th IEEE International Symposium on Robot and Human Interactive Communication, 2006*, ROMAN 2006, pp.103–110,.

Pavlovic, V.I. (1999) 'Dynamic Bayesian networks for information fusion with applications to human-computer interfaces', PhD thesis, Graduate College of the University of Illinois.

Pentland, A. (2000) 'Looking at people: sensing for ubiquitous and wearable computing', *IEEE Transactions on PAMI*, Vol. 22, No. 1, pp.107–119.

Rett, J. (2008) 'Robot-human interface using Laban Movement Analysis inside a Bayesian framework', PhD thesis, University of Coimbra.

Rett, J. and Dias, J. (2005) 'Visual based human motion analysis: mapping gestures using a puppet model', in Bento, C., Cardoso, A. and Dias, G. (Eds.): *EPIA 05*, *LNCS (LNAI)*, Vol. 3,808, Springer, Heidelberg.

Rett, J. and Dias, J. (2007) 'Human-robot interface with anticipatory characteristics based on Laban Movement Analysis and Bayesian models', in *Proceedings of the 2007 IEEE 10th International Conference on Rehabilitation Robotics*, Noordwijk, The Netherlands.

Rett, J., Dias, J. and Ahuactzin, J.M. (2008) *Frontiers in Brain, Vision and AI, Chapter Laban Movement Analysis Using a Bayesian Model and Perspective Projections*, pp.108–210, I-Tech Education and Publishing, Vienna.

Rosales, R. and Sclaroff, S. (2000) 'Learning and synthesizing human body motion and posture', in *Fourth IEEE International Conference on Automatic Face and Gesture Recognition*, pp.506–511.

Rossini, N. (2004) 'The analysis of gesture: establishing a set of parameters', in Camurri, A. and Volpe, G. (Eds.): *GW 2003*, *LNCS (LNAI)*, Vol. 2915, pp.124–131, Springer, Heidelberg.

Sato, T., Nishida, Y. and Mizoguchi, H. (1996) 'Robotic room: symbiosis with human through behavior media', *Robotics and Autonomous Systems*, Vol. 18, pp.185–194.

Shannon, C. (1949) *The Mathematical Theory of Communication*, University of Illinois Press.

Siegwart, R. et al. (2003) 'Robox at Expo.02: a large-scale installation of personal robots', *Robotics and Autonomous Systems*, Vol. 42, Nos. 3–4, pp.203–222.

Starner, T. (1995) 'Visual recognition of American sign language using hidden Markov models', Master's thesis, MIT.

Starner, T. and Pentland, A. (1995) 'Visual recognition of American sign language using hidden Markov models', in *International Workshop on Automatic Face and Gesture Recognition*, Zurich, Switzerland, pp.189–194,.

Starner, T., Weaver, J. and Pentland, A. (1998) 'Real-time American sign language recognition using desk and wearable computer based video', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20, No. 12, pp.1371–1375.

Sutton, V. (1982) *Dance Writing Shorthand for Modern and Jazz Dance*, Center Sutton Movement Writing.

Zhao, L. (2002) 'Synthesis and acquisition of Laban Movement Analysis qualitative parameters for communicative gestures', *PhD thesis*, University of Pennsylvania.

Zhao, L. and Badler, N.I. (2005) 'Acquiring and validating motion qualities from live limb gestures', *Graphical Models*, Vol. 67, No. 1, pp.1–16.