SEMANTIC ANALYSIS AND DATA STORAGE OF SKILLS.

Carlo Alberto Avizzano, Emanuele Ruffaldi, Massimo Bergamasco.

Scuola Superiore S.Anna
Piazza Martiri Libertà, 56100 Pisa, Italy
+39 050 883974
{carlo,e.ruffaldi,bergamasco}@sssup.it

Abstract: The analysis of complex motor skills for the transfer of expertise to trainees requires not only advancements in capturing and rendering technologies, but also research in the area of modeling, representation and storage. This talk presents the aspects of skills processing related to the encoding of the digital skill expressed in terms of decomposition of fundamental elements of the skill. The Multimodal Data System (MMDS) is introduced as a framework for the skill processing, with details related to storage and distributed processing of data.

Key words: skills analysis, digital representation, machine learning, multimodal, training.

1- Introduction.

The improvements in multimodal feedback technologies, capturing systems and machine learning give the opportunity for exploiting areas of research that has been typically limited to expensive virtual environment setups. The analysis of human motor skills, and the techniques for improving them, are aspects discussed in several virtual reality simulators although with the focus on specific domains. The analysis and the capacity of transfer a skills in humans is an important element of the future evolution of the Internet, that has already a strong role in the process of learning and exchanging information.

In the area of digital representation of physical entities, a lot of effort have been devoted to representing the basic senses, like vision and audio, supported by advancements in technologies for visualization and sound generation. Other senses have received less attention, or the research is more recent, like the case of haptics. There is anyway a forgotten aspect of human sensing that is receiving a lot of attention because of improvements in sensing technologies and analysis methods, it is the sense of movement **[B1]**. Research on gesture recognition and in general on human movements has improved along time, and it has reached now commercial level in various areas like entertainment and medicine.

The higher level understanding of human actions, aimed at the representation of human skills, has received less attention in terms of encoding and common mathematical representation. Research has been devoted mostly to the creation of simulators for the training of specific skills with less attention in finding a common model.

The paper is structured as follows. First we present the research context in which this work is collocated, then we discuss the problem of skill representation in multiple layers. The fourth section presents an architecture for the storage and processing of skills.

2- Research context.

The concept of digitally assisted system for training is quite long. At the end of 60's flight simulators recreated simulation environments to allow training pilots. At this stage, simulators were not addressing the full training problem, but remained at a level of pure simulation to serve as a test platform to improve performances independently. Only recently some simulators integrated training components within the simulation interaction **[MS1]** with applications in a variety of applications, from space training **[LK1]** to more recent trainers for robotic surgery **[NO1]**.

In initial solutions for training, the focus have been on the realism of simulation and interaction, providing an assessment of the overall task performance. This approach has been superseded by an analysis of user movements and actions, allowing objective evaluations of quality of the task performance **[NO1] [DS1]**. Such analysis can be easily performed in tool mediated contexts like robotic surgery **[RB1]**, or can be supported by high quality motion sensing technologies in free body contexts **[CC1]**. Because the objective of these systems is the transfer of the human skill from virtual to real environment, it is necessary to evaluate this transfer and understand how the multimodal technology is able to perform such transfer **[LR1]**.

The improvements of multimodal technologies, in particular haptic feedback, have given the opportunity of not only creating more realistic environments, but also for changing the stimuli that drive the learning process **[SF1]**. The potentialities of exploiting multimodal feedback as an active part of the learning process were pioneered by Yokokohji in 1996 **[YH1]** who proposed an approach, the "What you see is What you Feel"' (WYSIWYF) paradigm, that exploited haptic interfaces to record and playback to users proper motion-force cues of a given task.

With time haptics proved to be a dominant modality when complex sensori-motor skills are considered. However their effectiveness is diminished when contact is only partially present (in time or task) and the skills requires a fully multimodal interaction with the environments. An augmented reality environment was proposed by **[YO1]** to assist assembling procedures by exploiting iconic messaging. Visual information are particularly relevant since they can be used for a wide set of interaction information: semiotic, ergodic and epistemic. An example has been offered by Weidenhausen who, using a mixed reality environment, set up a full system that drives operators in learning assembling/disassembling of automobile motors **[WK2]**.

Learning from audio stimuli is very typical in music performancing and, of course, during verbal teaching. Less usual is the exploitation of the capabilities of the auricular subsystem to provide training information. Gutierrez **[AG1]** combined audio-haptic for the vocational training of visual impaired in different fields, economic, logistic and mathematics. Given the reduced weight and the easiness of integration **[TS1]** proposed a motion training system for karate skills while **[BM1]** proposed a full choreography assistant in theatrical performances.

The research on these training systems is mostly domain specific and it is not following a common approach in the representation of the skills. A skills is a human ability that can be learn and it is expressed by specific characteristics **[P1]**: reduced mental load, repetitiveness of performance, reduced attention requirements, energy and errors. The interest of this work is toward motor skills, that are the expression of a specific type of human memory, the enactive memory, that is the memory of doing **[FB1]**. Like any type of memory it should be possible to encode and represent it for retrieval and transfer.

In terms of encoding most of the research focuses on the representation of low level movements **[JC1]**, with specific focus on the area of music performance. At the same time there are semantic representations of human tasks, but these two approaches have not yet found an integrated approach **[WK1]** **[L1]**. This paper presents the current research toward an approach to digital representation of skills that can be applied in the creation of multimodal skills trainers, and in particular some examples will be related to current research on a rowing demonstrator.

3- Encoding digital skills.

As many kinds of data, skills can be modeled accordingly to several principles, and already some experiments to achieve training representations have been tried in past. Notwithstanding this, none of the existing representation is valid enough to design methodologies for encoding. The main issue for this lies in the fact that skills represents a complex set of aspects that difficultly could be caught out with a simple representation. Encoding a skill requires therefore a multi-structured approach that distinguishes among these aspects.

3.1- Data capture.

At a first level we distinguished among skill and task. A task is usually more richer of details with respect to a skill, it is bounded to a context and may depend from one or several objectives (e.g. driving a bicycle), while a skill usually express a much more focused set of abilities (moving linearly) that makes easier the analysis of underlying processes. In terms of mathematical representation, a skill can be analyzed by taking into account a reduced set of variables with respect to those required for interpreting the task. For instance gesture recognition of user hands usually takes into account complete description of all limbs movements since gestures may differ for details expressed at any level of the motion. Conversely similarities among skills are based on the effect produced onto the task objective and not on the whole set of elements which are included in the task.

As a matter of fact, the distinction among skills and task can be performed by the choice of a subset of data (the variables) which are necessary for the identification of the skills features. This choice in existing recognition/training systems is usually made implicitly, or driven by empirical principles: similarities, existence of mathematical tools for simplification **[JM1]** **[HE1]**, statistical properties **[BC1]**, black box approaches **[SM1]**, invariance and stability **[PC1]**, etc.

The correct choice of variables will influence the following capabilities of encoding. It is important that the chose variables will preserve relevant features of the skills under examination, for instance even the selection of the coordinate system may affect the quality of the results. the brain makes different internal use of reference coordinates **[B1]** to map different part of the skills: external reference coordinate systems (allocentric) and relative (egocentric) coordinate systems maps to actions. Any choice at this stage would encode the information into a different skills map on which the effects of disturbances, variability, styles, accuracy are retrievable at different level of complexity. Euclidean coordinates can be used for analysis techniques but they depend on the anthropometric measures of the users, and in addition algorithms based on linear interpolation as PCA or Independent Component Analysis **[HK1]** tend to lose the constraints given by the body segments lengths. Angular coordinates are user independent but are more difficult in the case of visualization. Among these the Euler angles have the additional problem of the gimbal lock during interpolation and for this reason it is preferred to use the quaternion for this type of operations.

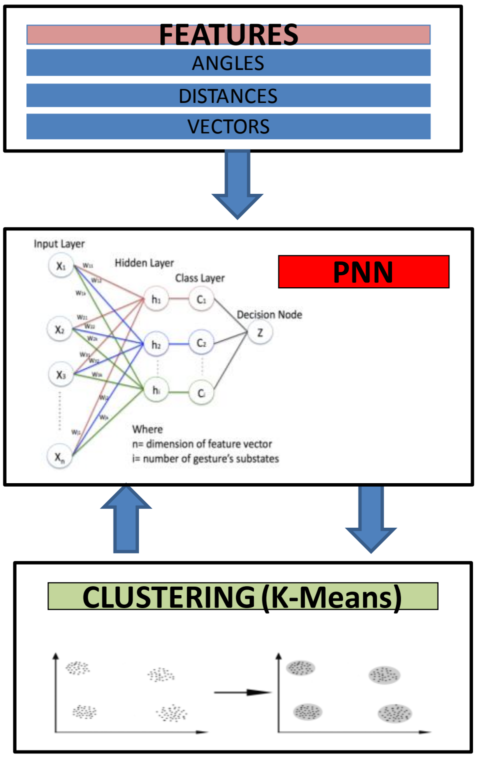

Figure 1: Raw data, angles, distances and differential vectors are grouped into sets with the help of clustering algorithms. Then programmed into neural network who can selectively discriminate relevant information and provide synthetic measurements to the recognition system

In Portillo **[PS1]** we showed the relationships among the properties of an encoding set of variables, which was mixed out of egocentric and allocentric variables, and the properties of skills encoding. In particular experiments evidenced that the role of a particular set of variables may be more or less significant with respect the specific phase of the skill. A 'strictly-redundant' variables set may therefore be useful when properties of the skill are not stationary with time.

In particular we showed how using a combination of Classification algorithms and probabilistic neural networks **[S1]** it is possible to setup an instrument that selects and optimize the relevant information from data without losing the ability to work in real-time as required for interactive environments.

3.2 – Interpretation.

The concept classifying the information existent into raw data into small (usually meaningless) groups, lead us toward the concept of features **[L1]** and the interpretation steps which aim to extract a semantic from the existing data. It is useful to think to skills in the same manner in which we think about music, speech or other form of human expressions that requires dynamics and interaction. Differently from pictures, the digital representation of these form of knowledge introduce an intermediate level between the raw data and the content analysis. As for letters in handwriting, notes in music or phonemes in speech, the skills variables could be aggregated and classified into richer and structured sets that codify the complexity of expression.

Several approaches to skills recognition make use of machine learning tools (Self organizing maps, Principal Component Analysis, Learning Vector quantization) or simpler quantizer vectors to cluster raw data streams in sequence of organized features. Clusterization into features introduce a precision loss in the data stream named average distortion. Distortion is sometimes useful since it allows to neglect, by approximation, the errors related to noise in execution. However, such a distortion should carefully handled in the feature design in order not to essential information of the task.

The creation of a set of features is the first step toward interpretation. Features represent abstract data that summarize properties of the action performed. Several elements are required to such representation for being able to properly maintain information for skill representation:

- Tolerance to noise and small errors in performance execution, it is likely that all the skill acquisition will be affected by instrumental errors, and any repetition of a similar skill will be differently acquired in data both due to human and environmental factors. However, when repeating a skill, we expect that it will be decomposed always in analogous sequences of features. The presence of noise may alter the value of some of them, but not impact on the overall structure;
- Filtering main invariances of the skills, it is useful if typical deformations in representation, such as time factors, scale factors, translation, rotation or projection onto different representation spaces, could be already eliminated by feature representation. This will allow the modeling and classification algorithms only to focus on the essence of skills without being affected from environmental or acquisition factors.
- Segmentation, features should not overlap in time different phases of a given action which may have different motion properties and objectives. Short time analysis, wavelet analysis, acceleration decomposition are all valid methods to preserve this property during decomposition.

The quality of features and the validity of the "vocabulary" can be validated still at this stage by using classification to features and numerical computation of the classification error. In Figure 2 we show the results of an experiment we carried out during tennis-table recordings of a player. Input data of human arm (angles for shoulder, elbow and wrist) were collected together the relative ball-racket position. Then data were processed into a computing pipeline (segmentation using PCA and mexican hat convolution, dynamic time warping to uniform length, and K-means quantizer) to reduce the trajectories into a sequence of feature symbols.

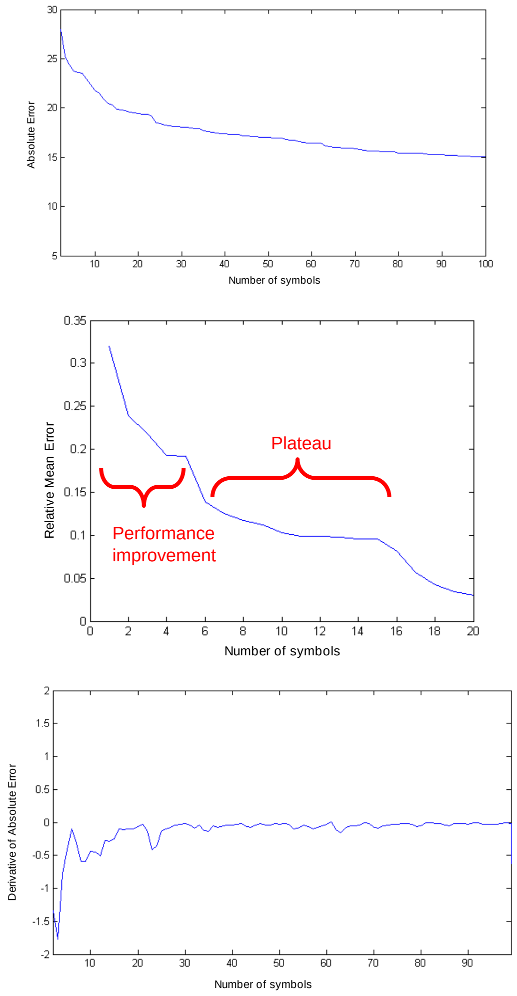

Figure 2: Reconstruction error computed comparing original data with those reconstructed from symbols identified by feature decomposition sequences, the X-axis always represent the number of clusters used, while the Y-axis represents: a) absolute error; b) relative error; c) differential error with respect the number of clusters.

The Integral of Time and Absolute Error (ITAE), the normalized error, and its derivative among the original trajectories and those represented by symbols centers (the centers of each K-means group) were then used to estimate the relationships among the number of clusters and the accuracy of the simplified representation. For this case we concluded that a relevant threshold of 20 is a good representation to distinguish among the basic element of motion and the errors introduced by repetition of actions. The relative distortion achieved with a reconstruction provided by 20 samples was less than 3% of the trajectory.

Interpretation tools will typically involve four basic types of actions: Semantic interpretation, Complexity reduction, Data classification and Data quantification. At the end of interpretation process a classification set and a quantification set of information will be available.

3.3 – Sub-skills.

Once a semantic organization of the feature cluster is achieved, data classes and the quantification information associated to them should serve as guidance in creating an association from content to skills. Again, if we think to speech, being able to distinguish the sequence of words is not enough to reconstruct the discourse and the dialogue abilities that could be in the behind. A wider analysis on the structured use of terms, on the verbal inflexions, of rhythmic and other is sometimes required to extract precise information. Similarly, in order to model similar aspects of sensori-motor and cognitive properties of the skill, two complementary methodology of representation are required:

- A sensori-motor model of the interaction; and
- A flexible task description

Ruffaldi in **[R1]** proposed a model to encapsulated both the sensori-motor model and the task description into a unique structure that may represent and classify the user interaction. The scope of this representation is not, as typically found in literature, to achieve an interpreter of human skill, but to focus interpretation within an interactive loop that conveys training information to the user.

Based on the Warren **[W1]** representation of human environment interaction, the enactive trainer shown in Figure 3 collects both data from sensori-motor observation (which flows to the trainer from the digital observation function λ, as well as the knowledge on the task history that is stored into the digital trainer memory (second loop).

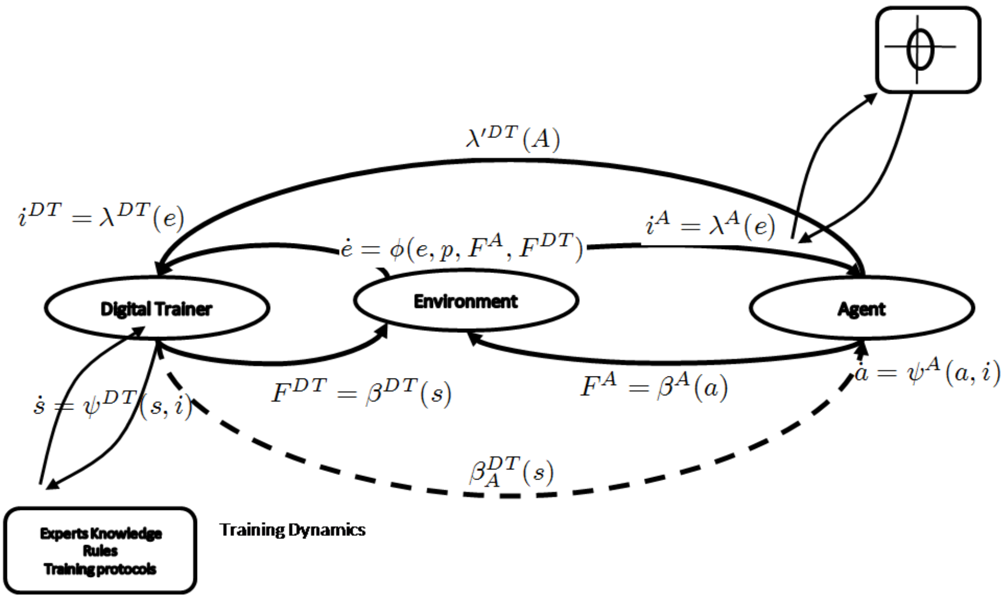

Figure 3: Information transfer model in the Enactive Trainer.

3.4 – Skills.

Skills are retrieved from top-down analysis of the task components, by identifying the quality of abilities that are required to accomplish task action and correlating them to a given set of subskill (both cognitive and sensori-motor) who have been classified **[BA1]** accordingly to a systematic analysis of relevant classes of interoperation.

Subskills decomposition is not orthogonal in the sense that each skill is performed in parallel with other ones and its execution affects also the performances of other ones, however for each subskill the following features will be observed: the will reflect the relevant variables that will be captured in the feature sets; they behave a complementarily of interaction modalities required; subskills do not break the fundamental relationships of the task; it exist a scientific background to assist the analysis of the subskill and to provide evidence for training.

The organization of a skill into sub-skills allows to understand the commonalities between tasks, to allow training on relevant components, and moreover to exploit the similarities in the learning process. Each sub-skills can be described in terms of involved sensor modalities, information bandwidth, learning methods and relationships with other sub-skills. This information is useful later for characterizing the representation of the sub-skill itself.

Each the decomposition of skills in subskills should provide at least two types of descriptive variables: observational and control variables. Observational variables are those elements such as quality of motion, time history, time descriptors, who are almost stable across exercises and which can be used to establish a metrics in the skill; for example when considering the bimanual coordination sub-skill, a possible descriptor is based on the periodic description of movement and the use of the Hilbert transform for representing the phase shifts **[RT1]**; Control variables will describe those attentional/perceptional stimuli that affects the voluntary execution of the skills without influencing the overall dynamics and perception of the task. Control variables are enhanced with the adoption of multimodal feedback and can also be shown with cross modality information (e.g. audio feedback for positional errors). It is important that a determined set of control variables will be generated accordingly to each category of subskill.

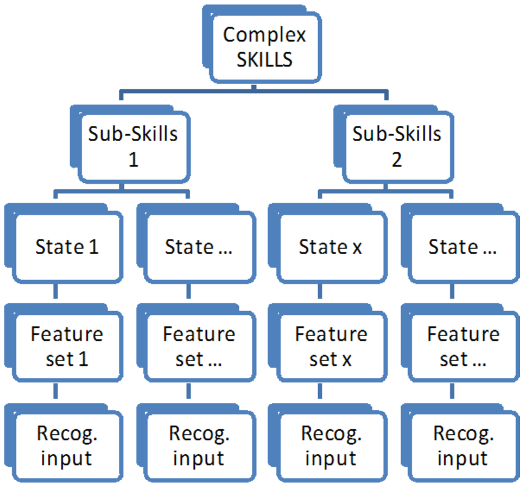

Figure 4: Logic decomposition of a skill into structured classes that can be handled by subskills architecture and analysis techniques.

4- Multimodal data system.

The multimodal nature of skill data requires the possibility of storing different types of information, and moreover to take into account the multiple levels of representation discussed above. For this reason the storage of the skills and in particular the recordings of capturing sessions of experts' data have been organized around a new architecture for data processing, the MultiModal Data System (MMDS).

The main problem that MMDS wants to address is the management of the variety of data sources that are involved in the multimodal capture for an experiment. Multiple video streams, motion capture data, biometrical information, tactile and force sensors are example of such information. Each of these modalities can be recorded with a data format that in most of the cases is standardized, like for audio and video, but these formats are not able to capture the spatial and temporal relationship between each modality. Moreover it is important to record not only the data itself, but also the characteristics of the data source, because they are important for later analysis of the experiment. For a video stream the types of cameras used, or for a haptic interface its characteristics and internal parameters.

The MMDS is organized into two fundamental blocks, the storage component and the processing component as shown in Figure 5.

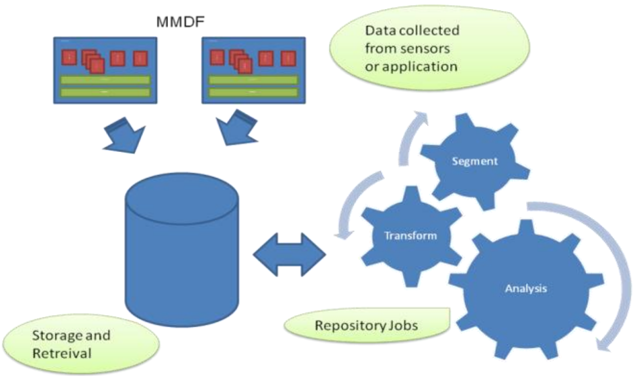

Figure 5: Organization of the MultiModal Data System for skill storage and processing.

4.1 – Storage.

The storage level manages the representation of the skills in a database that keeps both the low level information and the metadata necessary to describe the information. This level stores information in the database following the concepts of the MPEG-7 **[CS1]** standard, with specific extensions dealing with features and sub-skills. The choice of the MPEG-7 allows exchanging data in the database with the external world using XML and in addition to support variety of data types, as binary data and ontology-based information.

The MPEG-7 schemas are flexible and can be extended for new applications. In particular the current focus of MPEG-7 is in the direction of audio and video encoding. For the storage of skills we are extending MPEG-7 adding new modules in several areas, both at low level and at high level. At the low level it is necessary to integrate concepts from haptics, motion capture, physics and annotation of real scenes. The skill is encoded by specific descriptors (in the sense of MPEG-7 schema) that describe the decomposition in tasks and sub-skills. Finally specific descriptors of sub-skills are necessary for their encoding, and this requires a reference to specific algorithms adopted for analyzing and representing the specific sub-skills. These specialized schemas are collected under the name of Multimodal Data Format (MMDF).

The structured information from the enhanced MPEG-7 document, and the binary data, is stored in a relational database management system (RDMS) **[WK3]** in a hierarchical form, with binary data stored in compact form. In this way it is possible to take advantage of the fundamental features of relational databases as persistence, reliability and distributed features.

The storage level is directly connected to data acquisition module, allowing storing data coming from sensors directly into the database. This operation can be performed offline using XML as interchange format or in online mode. In the online mode an experiment template is created with empty slots in which information will be completed coming from the different sensors. In this way the channels of information coming from network connections (like as VRPN **[HS1]** or raw TCP) are dispatched to the correct locations in the structured experiment data.

4.2 –Computation.

The processing phase is based on a data-flow based approach in which multiple nodes are connected for creating a computing network. This network of computation complements the features of existing solutions, like Simulink, with specific modules dedicated to data segmentation, clustering and skill analysis, and it produces information integrated and annotated in the storage component.

The major feature of this component is the retrieval of the recorded skill, at low level as happens in motion capture databases [MR1] and at higher level adopting the specific representation of the skills.

5– Conclusions

This work presented the problem of skills analysis and encoding in terms of multiple layer decomposition and processing model. There are several open issues in the area of encoding, and more components that need to be investigated for a general approach in the representation of skills.

6– Acknowledgements.

This work has been carried out with the financial assistance of the European Commission which co-funded the integrated project "SKILLS" (IST-2006-035005). The authors are grateful to the EU for the support given in carrying out these activities.

7- References

[AG1] Avizzano C.A., Gutierrez T., Casado S. and others, Grab: Computer graphics access for blind people through a haptic and audio virtual environment. In Proceedings of Eurohaptics, 2003

[B1] Berthoz A. The Brain's Sense of Movement. Harvard University Press, 2000.

[BA1] Bardy B., Avizzano C.A., Gopher D. and Bergamasco M., Unified Framework for the Acquisition and the Transfer of Skill. Technical Report of EU Skill Integrated Project FP6-*IST*-2005-035005-IP, 2008

[BC1] Billard, A.G. and Calinon, S. and Guenter, F., Discriminative and adaptive imitation in uni-manual and bi-manual tasks. In Robotics and Autonomous Systems, 54(5): 370-384, 2006

[BM1] Baalman M. A.J., Moody-Grigsby D. and C. Salter, Schwelle: sensor augmented, adaptive sound design for live theatrical performance. In Proceedings of the 7th international conference on New interfaces for musical expression, pp 178-184,2007.

[CC1] Chua P.T., Crivella R., Daly B., Hu N., Schaaf R., Ventura D., Camill T., Hodgins J. and Pausch R. , Training for physical tasks in virtual environments: Tai Chi. In Proceedings of IEEE Virtual Reality, pp 87-94, 2003.

[CS1] Chang S.F., Sikora T. and Purl A., Overview of the MPEG-7 standard. In IEEE Transactions on Circuits and Systems for Video Technology, 11(6): 688-695, 2001

[DS1] Darzi A., Smith S. and Taffinder N. , Assessing operative skill: Needs to become more objective. In BMJ: British Medical Journal, 318(7188): 887, 1999

[FB1] Foley M.A., Bouffard V., Raag T. and DiSanto-Rose M., The effects of enactive encoding, type of movement, and imagined perspective on memory of dance. In Psychological Research, 53(3): 251-259, 1991

[HE1] Harding P. R. G. and Ellis T. J. Recognizing hand gesture using Fourier descriptors. In International Conference on Pattern Recognition (ICPR), 3: 286-289, 2004.

[HK1] Hyvarinen A., Karhunen J. and Oja E., Independent Component Analysis. In IEEE Transactions on Biomedical Engineering, 14: 21-30, 2000

[HS1] Hudson T.C., Seeger A., Weber H., Juliano J. and Helser A.T., VRPN: a device-independent, network-transparent VR peripheral system. In Proceedings of the ACM symposium on Virtual reality software and technology, pp 55-61, 2001

[JC1] Jensenius A.R., Camurri A., Castagne N., Maestre E., Malloch J., McGilvray D., Schwarz D. and Wright M., Panel: the need of formats for streaming and storing music-related movement and gesture data. In Proceedings of the International Computer Music Conference, 2007

[JM1] Jiang X. and Motai Y., Learning by observation of robotic tasks using on-line PCA-based Eigen behavior. In IEEE International Symposium on Computational Intelligence in Robotics and Automation, pp. 391-396, 2005

[L1] Lee C., Learning Reduced-Dimension Models of Human Actions, PhD Thesis, CMU, 2000

[LK1] Loftin R.B. and Kenney P. Training the Hubble space telescope flight team. In IEEE Computer Graphics and Applications, 15(5): 31-37, 1995

[LR1] Lehmann K.S., Ritz J.P., Maass H., Cakmak H.K., Kuehnapfel U.G., Germer C.T., Bretthauer G. and Buhr H.J., A Prospective Randomized Study to Test the Transfer of Basic Psychomotor Skills From Virtual Reality to Physical Reality in a Comparable Training Setting. In Annals of Surgery, 241(3): 442, 2005

[MR1] Muller M., Roder T., Clausen M., Efficient Content-Based Retrieval of Motion Capture Data. In SIGGRAPH, 2005

[MS1] de Moraes R.M. and dos Santos Machado L., Hidden Markov Models for Learning Evaluation in Virtual Reality Simulators. In International journal of computers & applications, 25(3): 212-215, 2003

[NO1] Narazaki K., Oleynikov D. and Stergiou N. Robotic surgery training and performance. In Surgical Endoscopy, 20(1): 96-103, 2006

[P1] Posner M.I., Components of Skilled Performance. In Science, 152(3730): 1712-1718, 1966.

[PC1] Parameswaran V. and R. Chellappa , View invariance for human action recognition. In International Journal of Computer Vision, 66(1): 83-101,2006

[PS1] Portillo-Rodriguez O., Sandoval-Gonzalez O., Avizzano C.A., Ruffaldi E., Vercelli D. and Bergamasco M., Development of a 3D Real Time Gesture Recognition Methodology for Virtual Environment Control. In 17th International Symposium on Robot and Human Interactive Communication, 2008

[R1] Ruffaldi E., Mathematical Modeling of Skills. In workshop on Skills Capture and Transfer in RO-MAN08, Munich, Germany, 2008

[RB1] Rosen J., Brown J.D., Barreca M., Chang L., Hannaford B. and Sinanan M., The Blue DRAGON-A System for Monitoring the Kinematics and the Dynamics of Endoscopic Tools in Minimally Invasive Surgery for Objective Laparoscopic Skill Assessment. In Studies in Health Technology and Informatics-Medicine Meets Virtual Reality, 85: 412-418, 2002

[RT1] Rosenblum L.D. and Turvey M. T., Maintenance tendency in co-ordinated rhythmic movements: relative fluctuations and phase. In Neuroscience, 27(1): 289-300, 1988.

[S1] Specht D. F., Probabilistic neural networks. In Neural Networks, 3(1): 109-118, 1990

[SF1] Sano A., Fujimoto H. and Matsushita K. Machine mediated training based on coaching. In Proceedings of IEEE Systems, Man, and Cybernetics SMC'99, 2, 1992

[SM1] Solis J., Marcheschi S., Frisoli A., Avizzano C.A. and Bergamasco M., Reactive robot system using a haptic interface: an active interaction to transfer skills from the robot to unskilled persons. In Advanced Robotics, 21(3): 267-291, 2007

[TS1] Takahata M., Shiraki K. Sakane Y. and Takebayashi Y., Sound feedback for powerful karate training. In NIME '04: Proceedings of the 2004 conference on New interfaces for musical expression, pp 13-18, 2004

[Y1] Yuan M.L., Augmented reality for assembly guidance using a virtual interactive tool. In International Journal of Production Research, 1: 1-23, 2007

[YH1] Yokokohji Y., Hollis R.L., Kanade T., Henmi K. and Yoshikawa, T. Toward machine mediated training of motor skills. Skill transfer from human to human via virtual environment. In 5th IEEE International Workshop on Robot and Human Communication, pp. 32-37, 1996

[W1] Warren W. H., The dynamics of perception and action. In Psychological Review, 113(2): 358-389, 2006

[WK1] Wong S.F., Kim T.K. and Cipolla R., Learning Motion Categories using both Semantic and Structural Information. In Proceedings of Computer Vision and Pattern Recognition, pp. 1-6, 2007

[WK2] Weidenhausen J., Knoepfle C. and Stricker D., Lessons learned on the way to industrial augmented reality applications, a retrospective on ARVIKA. In Computer & Graphics, 27(6): 887-891, 2003

[WK3] Westermann U. and Klas W., An analysis of XML database solutions for the management of MPEG-7 media descriptions. In ACM Computing Surveys, 35(4): 331-373, 2003