HAL
open science

Towards the design of augmented feedforward and feedback for sensorimotor learning of motor skills

Paraskevi Kritopoulou, Sotiris Manitsaris, Fabien Moutarde

To cite this version:

Paraskevi Kritopoulou, Sotiris Manitsaris, Fabien Moutarde. Towards the design of augmented feedforward and feedback for sensorimotor learning of motor skills. 3rd International Symposium on Movement and Computing (MOCO'2016), Jul 2016, Thessalonique, Greece. pp.38, 10.1145/2948910.2948959 . hal-01437136

HAL Id: hal-01437136

https://hal.science/hal-01437136

Submitted on 19 Jan 2017

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

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

Towards the design of augmented feedforward and feedback for sensorimotor learning of motor skills

Paraskevi Kritopoulou
Centre for Robotics - MINES ParisTech, PSL Research University
60, Bd Saint-Michel
75272 Paris, France
paraskevi.kritopoulou@mines-paristech.fr

Sotiris Manitsaris
Centre for Robotics - MINES ParisTech, PSL Research University
60, Bd Saint-Michel
75272 Paris, France
sotiris.manitsaris@mines-paristech.fr

Fabien Moutarde
Centre for Robotics - MINES ParisTech, PSL Research University
60, Bd Saint-Michel
75272 Paris, France
fabien.moutarde@mines-paristech.fr

ABSTRACT
Creating a digital metaphor of the "in person transmission" of manual-crafting motor skills, is an extremely complicated and challenging task. We are aiming to achieve the above by creating a mixed reality environment, supported by an interactive system for sensorimotor learning that relies on pathing techniques. The gestural instruction of a person, the Learner, arises from the reference gesture of an Expert. The concept of the system is based on the simple idea of guiding with the projection of a gesture depicting path in 2D space and in real time. The path is projected either as a feedforward that describes the gesture that has to be executed next, either as a feedback that amends the gesture while taking into account the time needed to correct the mistake. This projection takes place in the exact area where the object lies and the Learner is being trained, to avoid any distraction from the crafting task.

Author Keywords
Pathfinding; gesture recognition; sensorimotor learning; feedback; feedforward; scene analysis.

ACM Classification Keywords
Algorithms.

INTRODUCTION
Traditionally, the transmission of practical knowledge motor skills can be passed on only by demonstration from master to apprentice. The apprentice unconsciously improves existing skills and acquires knowledge through repetitive crafting [10]. Yet, the technological advancement offers new interactive possibilities of learning, not necessarily based on imitation or verbal communication, and with means yet unexplored. This research results to realising the educational potential of pathfinding algorithms and their possible contribution on self-training, on motor skills.

STATE OF THE ART
Most developed systems that aim on motion training, are situation based or knowledge focused and not gestural. Yet gestural correction with educational or training purposes, has been implemented in certain fields. These fields have as a common characteristic the fact that, their nature renders the guidance by an Expert of the field a necessity, so that a learner can improve his/her dexterity on this very field. These systems have been implemented either on Virtual Reality Environment (VE), either on Augmented Reality Environment (AE).

The existing approaches in real time, focus mainly on the superimposition of the Learners avatar on the ideal avatar of an Expert, a technique initially proposed by the "Just Follow Me" project. The superimposition has been fabricated, by wired models either on screen, either on HMDs [13]. Variations on the concept of superimposition are trying to improve the effectiveness of the training system through feedbacks, such as using 3D avatars of the Expert and the Learner, projected on screen by various angles in real time [9]. Supplementary, it has been attempted in real time, the coloring of the body members of the avatar corresponding to the learner that decline from the gesture. Simultaneously, after the completion of the gesture, this system provided instructions that emphasize on the declination of the already performed gesture from the ideal one in slow motion, along with a score report [2]. There is also an approach that targets novice users, and focuses on mapping an instruction video of self-images where the gesture of the expert is combined with the gesture of the learner. The result is the image of the learner performing the gestures with the accuracy of an expert [3]. In an attempt for feedforward instruction the movement of the Expert was demonstrated in advance at each frame so that the user can copy the movement [7]. Although superimposition seems to be a good solution for training on fields requiring physical tasks such as dancing and martial arts, in crafting fields that are related to objects, prove to be distracting.

Other implementations are trying to take advantage of the human senses of sight, hearing and touch. By using color coded arrows that indicate the quality of the movement and

are projected on the limbs of the user, a technique proposed by the LightGuide system [11], with acoustic tones and vibrations, the strive to achieve fruitful instruction, proceeds. Multimodal research has resulted in a system which uses an optical stimulus representing distance to instruct the user as both feedforward and feedback, providing in addition, a special acoustic feedback that keeps the user alert in the case of the declination of the gesture [4]. A similar implementation to the LightGuide, projects an optical signal lengthwise the users limb, and in the position where the limb should be placed, using supplementary vibration sensors to notify the user about any declination. In parallel, it uses an acoustic signal to achieve the timing of the exercise [12]. With a simplified projection of the stance of the body of the user, accompanied with ribbons at the end of the limbs that guide the movement of the trainee, and the concurrent resonance of acoustic stimulus whose purpose is to preserve the rhythm of the movements, YouMove functions as a corrective mirror. The avatar of the trainee, is projected on a screen in real-size. Furthermore, the system has the ability to replay the parts of the performed gesture where there was declination from the optimal [1].

The above techniques concluded that virtual or augmented reality environments are as good as and not significantly better than the traditional training methods, and yet much better than the training provided by instruction videos. Feedforward attempts are scarce and provide insufficient information about the quality and the progression of the gesture to the learner. Yet, for learning to take place, internal and external sources of information about motor performance is essential [5]. To achieve this, we propose an alternative way of instruction based on pathfinding algorithms.

METHODOLOGY OVERVIEW

The implementation of the presented system will be aiming to improve pottery skills. Figure 1 illustrates the process.

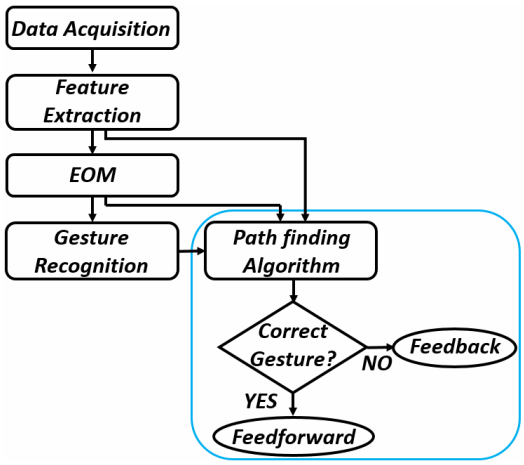

Figure 1. Pipeline

Initially, the data concerning the coordinates of the Learner are acquired through motion capture. Ensuing, through feature extraction we define the coordinates of the gesture. These features are used at first to detect the position of the Learner, and at second, to recognize his/her gesture, in order to select subsequently and depending on the gesture that is currently performed, the appropriate Expert Operational Model (EOM) [6]. The EOM is a model that encapsulates all the information of the said gesture, as performed by an Expert. Therefore, the model includes the subtle details of the quality of the gesture, which can distinguish an Expert from a Learner. Next, and as soon as the EOM is selected, the system verifies if the gesture of the Learner that is recognized is similar to the respective gesture of the Expert. Finally, the instruction of the Learner through displayed paths takes place.

The concept of the path guidance is based on the need to guide the user at every stage of the tutoring process. Thus, the system functions in two modes that use the “driving car” metaphor, which is inspired by automotive navigation systems. The first mode is a feedforward that instructs the Learner about the way the gestures should be performed towards completion, while the second one is generated as a reaction to a detected gestural declination of the learner. In both modes, as starting point of the path is considered the position in which the Learner is currently placed, and the terminal point is the position the Expert was placed in a more advanced point of the gesture.

Regarding the definition of gestures in the crafting fields, the complexity of the skill to be learned determines the number of component parts that the skill is consisted of [5]. In the pottery use case we segmented the series of movements that are used to create a simple clay object, a porringer, into 4 gestures. Each gesture is related to a basic clay formation and to the correlation of the clay with the wheel. The first gesture is about centering the clay on the wheel and opening its bottom, the second is raising walls, the third is a first configuration of the object and then comes the final configuration along with the object removal from the wheel.

Feedforward

The purpose of this system is motor learning, and to achieve this, it is essential for the learner to practice correctly [5]. In the augmented environment this principle is implemented by projecting the next gesture that has to be performed by the learner, by using the EOM. The model is given in the form of a set of mathematical equations, each one of which contains the exact description of the gesture in time [6]. So, by extracting the coordinates, the movement that stepwise shapes the gesture is extracted. The EOM has been designed to calculate the position on the 3D space, yet the information is displayed to the Learner as an augmented path in 2D space; a 2D line that describes the gesture which has to be performed next, as illustrated in Figure 2 for the pottery-learning environment. This way the learner knows which is his following action and consciously follows the instructions.

The EOM model produces each coordinate, of a temporary position of the gesture. Due to the nature of pottery crafting, the EOM used, is concerning the coordinates of the palms since the position of these parts of the hand controls the

shape of the clay. For example the X coordinate of the palm that describes a certain gesture, is calculated from the equation:

$$X_t = \alpha_{10} - \alpha_{11} \cdot Z_{t-1} - \alpha_{12} \cdot Y_{t-1} + sv_X \cdot \left(X_{t-1} - X_{t-2}\right) + \alpha_{13} \cdot X_{t-1}$$

where t is the time indicator, α are the coefficients developed by state-space estimation methodology, and sv is the state variable for the X equation – one of the calculated coefficients that varies through time, and this way achieves the dynamic description of the gesture. The model is acquired by using the coordinates of three continuous timestamps from the feature extraction of the Learner. The gesture is extracted from its beginning to its end, regardless of the moment in which it is constructed. Furthermore, each limb that participates in the gesture has its own EOM model and its position variables X, Y and Z, are mutually depended.

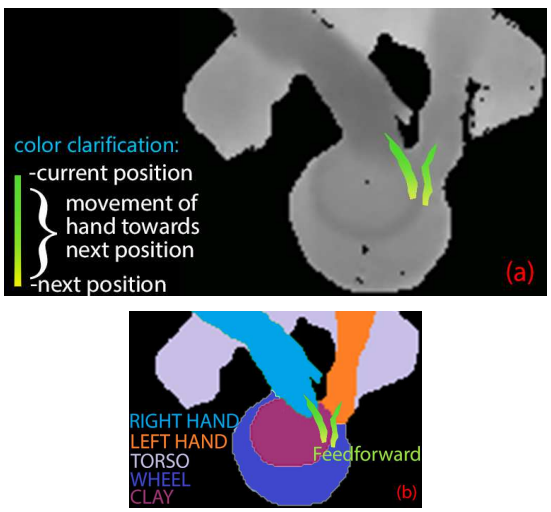

Figure 2. Graphical representation of the (a) feedforward mode, and (b) the explanation of the grayscale image

During the craftsmanship, the user is immersed on the object, therefore, a mixed-reality approach that will rectify without obstructing the learning process is desirable. Thus, color code is used, to denote the feedforward (green) and the feedback (red) modes, together with how big the declination of the gesture is. Figures 2 and 3 depicts the color code and the way the information is displayed to the learner.

Feedback

In our pursuit, real time feedback holds the importance of immediate testimony in the progress of training. Therefore, as soon as an unacceptable declination in the optimal position of the hand occurs, a path that aims at the correction of the gesture is calculated, by taking into consideration that the obstruction of the object is dependent both on time and the progression of the gesture itself. This is achieved again with the help of the EOM. The terminal point of this path is the ideal position in which the Expert would be placed after a certain time period which is reliant on the interfering nature of the object, and is extracted by the EOM.

Figure 3 illustrates the paths that are projected when declination occurs in a pottery-learning environment, and in the event of both hands failing. In No 1 path, an interposing obstruction is displayed through color code (blue), thus, the learner should make adjustments on the Z axis (high of the hand).

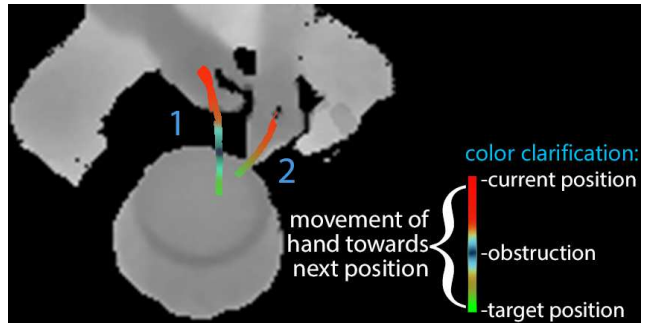

Figure 3. Graphical representation of the feedback mode (1) with obstruction, (2) without obstruction

The system should be able to calculate the best path the deteriorated gesture of the Learner must follow, in order to be corrected, in regard to the stance of the Expert. This is achieved by using the A* pathfinding algorithm designed to return certain waypoints that highlight the optimal path, and are used to reconstruct a curve by Polynomial Interpolating. This curve is the feedback, which is a path projected simply to the immediate environment of the Learner, to avoid any distraction from the learning process.

A*(star) algorithm

A* is a graph search algorithm, designed to discover the optimal path that connects two specific points in space [7]. To achieve that, the space is depicted as a grid that allows 8 possible directions of movement on the x-y plane, while the vertical adjustment of the z axis is done separately since all objects on the crafting field are bypassed vertically. Yet, the continuous environment is not limited by certain angles or directions, as occurs in the case of the discrete environment that a squared network like the grid creates. Therefore, polynomial processing is applied on the defined path, so that the Learner may be led towards any possible direction.

Polynomial Implementation

We propose an ameliorative process which uses alternatively the A*, and in which the A* algorithm returns some nodal points which define the beginning, the terminal, and the path around the object. These points are used to construct a polynomial curve in which the information is enclosed compressed, in the form of an equation. As a result, the storage space is minimised and naturalness in movement is achieved, without receding from the optimal path.

IMPLEMENTATION

Our instruction system towards implementation, is consisted of three main components. The first one is a short-range Time-of-Flight sensor, a camera that captures depth-maps. The second component is a personal computer in which the gestural recognition part, processes the input data from the

gesture control and then selects the mode and calculates the path. The path is displayed to the user with the third component, an image projector. The projector displays the feedforward or feedback path into the same coordinate space where the learner acts. A description of the environment in which the training takes place is displayed in figure 4.

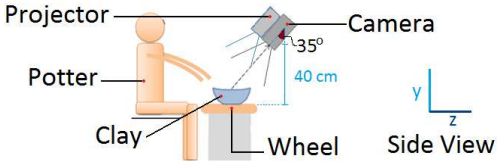

Figure 4. Augmented Environment set-up

CONCLUSION

We propose the implementation of pathfinding techniques as a sensorimotor feedforward and feedback to a person who aims to self-improve his/her motor skills in object-crafting fields. In order to achieve that we use the EOM as the core of the training method, a model that describes the gesture, as performed by an Expert craftsman. Due to the nature of the craftsmanship, the formulation of a mixed-reality environment, where gesture recognition and instruction will take place, is essential to achieve our purpose. The selected case study where we intend to apply and test our methodology, is the crafting art of pottery. We expect that during testing, the interactive system will produce high results due to the fact that the instruction is clearly and directly projected, and not revealed after a series of predefined actions given at each step at a time in the form of an achievement. This way, the system does not compel the optimal solution to the user, further, allowing enough autonomy and the space to improvise.

ACKNOWLEDGMENTS

Paraskevi Kritopoulou has received scholarship by the Bodossaki Foundation.

REFERENCES

1. Fraser Anderson, Tovi Grossman, Justin Matejka, George Fitzmaurice. 2013. YouMove : Enhancing Movement Training with an Augmented Reality Mirror. *In Proceedings of the 26th annual ACM symposium on User interface software and technology* (UIST 2013), http://doi.acm.org/10.1145/2501988.2502045
2. Jacky C.P. Chan, Howard Leung, Jeff K.T. Tang and Taku Komura. 2011. A Virtual Reality Dance Training System Using Motion Capture Technology. IEEE Transactions on Learning Technologies, v.4 n.2, p.187-195, April 2011, http://doi.acm.org/10.1109/TLT.2010.27
3. Minoru Fujimoto, Masahiko Tsukamoto, Tsutomu Terada. 2012. A Dance Training System that Maps Self-Images onto an Instruction Video. *In Proceedings of the 5th International Conference on Advances in Computer-Human Interactions* (Achi 2012), pp.309-314. http://www.thinkmind.org/download.php?articleid=achi_2012_12_30_20280 PDF
4. Alina Glushkova, Sotiris Manitsaris. 2015. Gesture Recognition Technologies for Gestural Know-how Management – Preservation and Transmission of Expert Gestures in Wheel Throwing Pottery. In *Proceedings of the 7th International Conference on Computer Supported Education* (CSEDU 2015), 405-410, http://doi.acm.org/10.5220/0005475904050410
5. Tim Hopper. 2014. PRINCIPLES OF LEARNING MOTOR SKILLS. Retrieved May 4, 2016 from http://web.uvic.ca/~thopper/WEB/archive247/term2/week3/principleslearning.htm
6. Sotiris Manitsaris, Alina Glushkova, Eleni Katsouli, Athanasios Manitsaris and Christina Volioti. 2015. Modeling gestural know-how in pottery based on state-space estimation and system dynamic simulation. In *Proceedings of the 6th International Conference on Applied Human Factors and Ergonomics* (AHFE 2015),http://doi.acm.org/10.1016/j.promfg.2015.07.833
7. Otniel Portillo-Rodriguez, Oscar O. Sandoval-Gonzalez, Emanuele Ruffaldi, Rosario Leonardi, Carlo Alberto Avizzano and Massimo Bergamasco. 2008. Real-Time Gesture Recognition, Evaluation and Feed-Forward Correction of a Multimodal Tai Chi Platform. *In Proceedings of the 3rd international workshop on Haptic and Audio Interaction Design* (HAID 2008), http://doi.acm.org/10.1007/978-3-540-87883-4 4
8. Amit Patel. 2009. Amit's A* Pages. Retrieved November 22, 2016 from http://theory.stanford.edu/~amitp/GameProgramming/
9. Philo Tan Chua, R. Crivella, B. Daly, Ning Hu, R. Scaaf, D. Ventura, T. Camill, J.Hodgins, R. Pausch. 2003. Training for physical tasks in virtual environments: Tai Chi. *In Proceedings of the IEEE Virtual Reality Conference* (VR 2003), http://doi.acm.org/10.1109/VR.2003.1191125
10. Michael Polanyi. 1973. *Personal Knowledge-Towards a Post-critical Philosophy*. London, Routledge & Kegan Paul.
11. Rajinder Sodhi, Hrvoje Benko and Andrew D. Wilson. 2012. LightGuide: Projected Visualizations for Hand Movement Guidance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, http://doi.acm.org/10.1145/2207676.2207702
12. Joao Vieira, Mauricio Sousa, Artur Arsenio and Joaquim Jorge. 2015. Augmented Reality for Rehabilitation Using Multimodal Feedback. *In Proceedings of the 3rd 2015 Workshop on ICTs for improving Patients Rehabilitation Research Techniques* (REHAB 2015), http://doi.acm.org/10.1145/28389544.2838954
13. Ungyeon Yang. 2002. Implementation and Evaluation of "Just Follow Me": An Immersive, VR-Based, Motion-Training System. In *Proceedings of the International Conference on Virtual Systems and Multimedia* (VSMM 2002), http://doi.acm.org/10.1162/105474602317473240