Multimodal Motion Guidance: Techniques for Adaptive and Dynamic Feedback

Christian Schönauer[1,2]
schoenauer@ims.tuwien.ac.at

Kenichiro Fukushi[1,3]
fukushi@isl.titech.ac.jp

Alex Olwal[1]
olwal@media.mit.edu

Hannes Kaufmann[2]
kaufmann@ims.tuwien.ac.at

Ramesh Raskar[1]
raskar@media.mit.edu

MIT Media Lab[1]
Vienna University of Technology[2]
Tokyo Institute of Technology[3]

Figure 1: a) Multimodal feedback guides a user to follow directed motion from a remote teacher. b) Motion capture allows full-body feedback. c) Vibrotactor hardware d) for continuous feedback on speed and direction.

ABSTRACT

The ability to guide human motion through automatically generated feedback has significant potential for applications in areas, such as motor learning, human-computer interaction, telepresence, and augmented reality.

This paper focuses on the design and development of such systems from a human cognition and perception perspective. We analyze the dimensions of the design space for motion guidance systems, spanned by technologies and human information processing, and identify opportunities for new feedback techniques.

We present a novel motion guidance system, that was implemented based on these insights to enable feedback for position, direction and continuous velocities. It uses motion capture to track a user in space and guides using visual, vibrotactile and pneumatic actuation. Our system also introduces motion retargeting through time warping, motion dynamics and prediction, to allow more flexibility and adaptability to user performance.

Categories and Subject Descriptors

H.5.2 [**Information interfaces and presentation**]: User Interfaces - Input devices and strategies

ICMI'12, October 22–26, 2012, Santa Monica, California, USA.

Keywords

Multimodal feedback, motion guidance, tactile feedback

1. INTRODUCTION

The development and training of motor skills can greatly benefit from efficient interaction techniques and multimodal feedback. While a teacher manually guides a student's movement in a classical learning scenario, technical systems have the potential to substitute and enhance this process, e.g., in Tai Chi [16]. One of the most important aspects in these systems, as well as for motor learning in general, is real-time motion feedback with respect to a target motion [12]. This naturally extends to other human-computer interaction (HCI) scenarios where feedback is critical for interacting with virtual content. An augmented movie set, for example, could have actors and director interact with computer-generated graphics using different modalities, as illustrated in Figure 1a and 2.

In this work, we distinguish between motor learning and motion guidance. While one of our overall goals is to teach new skills to a user, a technical system specifically designed for motor learning has to take many aspects into account, such as the different stages in the learning process [6] and attention in various phases [22]. We, instead, focus on multimodal feedback for motion guidance, which is an important aspect of acquiring new skills.

We propose a framework for manifold motion guidance and feedback through a motion guidance system (MGS), which includes interfaces for mixed reality scenarios and telepresence applications. Our modular framework, allows reuse and combination of feedback modalities and concepts for different applications. We analyse the design space for motion guidance and discuss the implementation of our framework, which uses full body tracking data and generates real-

Figure 2: An augmented movie set. The director (right) gives tactile feedback through a GUI to an actor (left) that interacts with a virtual object.

time visual, vibrotactile and pneumatic feedback to guide an individual's movement towards a desired motion.

Our **contributions** include:

- Analysis of the **design space** for motion guidance.
- Implementation of **multiple modules** for full-body 3D feedback that combines motion capture with tactile and visual stimuli.
- Dynamically adapting feedback based on a user's motion through **motion retargeting**, employing *time warping, motion dynamics, speed indication* and *movement prediction* for precise timing of feedback.

2. DESIGN SPACE

This section discusses the design space of multimodal feedback for motion guidance and how it relates to the human perceptual system, psychophysiological factors and information processing in the brain.

2.1 Human Information Processing

Human information processing is executed in three phases:

Stimulus identification describes the process of perceiving information. In the context of multimodal interaction, it is important to consider that perception varies under physical or cognitive workload [2] and that the perception of a feedback modality can be influenced by a parallel stimuli.

Response selection is the cognitive mechanism that converts the stimulus (perception) to a response (action). The human body provides a vast amount of possible responses to a stimulus (body movements) with its many degrees-of-freedom (DOFs). A large number of possible responses increases the response time, however, due to the associated cognitive process. Therefore, when considering stimulus-response (S-R) translations, it is often necessary to limit the set of responses and take into account S-R compatibility (e.g., spatial proximity between feedback and action).

Response programming, or "motor execution", executes the action and is affected by factors such as S-R frequency, movement direction, or sequence length. Furthermore, an action can alter the user's focus of attention and perception process.

Attention ([10] p33) and motor centers are tightly linked in the human brain and attention is crucial for motion guidance feedback, as it can lead to more effective motor performance and learning [22]. Selective attention preserves processing capabilites and processes only a subset of information.

2.2 Feedback Modalities

We perceive our environment with multiple senses simultaneously, as real-world feedback is multimodal. To simulate this in a virtual system, different feedback modalities and actuators have to be employed. We provide a brief overview of modalities and attributes relevant for motion guidance. Emphasis is placed on haptics, as an in-depth discussion of all modalities is beyond the scope of this paper.

Visual feedback is often considered the dominant form of feedback [10] (p112) and can provide users with objective observation of their motion (errors), augmented awareness for correction (e.g., emphasized body parts indication), superimposed target trajectories, or scores [3, 15].

Usability and user experience depends on the display hardware and properties, including latency, update rate, size, resolution, mobility and stereoscopy [18, 15].

Spatial Audio can be used to communicate 3D positions or direction. Dozza et al., for example, encode information using sinusoids with variable frequency and loudness [4].

Haptic feedback is composed of skin sensation and proprioception [10] (p194). **Skin sensation** is perceived by sensory receptors that enable us to localize and recognize individual tactile cues. Tactile displays can often indicate body part positions more intuitively than visual or audio feedback [20, 12]. Robustness and low weight has made vibration motors popular for tactile actuation.

Baek et al. [1] present a training system for fencing and dancing, and Rosenthal et al. explore foot step navigation in dance [17]. Spelmezan et al. [20] study spontaneous reactions to tactile patterns, while TIKL [12] gives a user visual information and vibrotactile feedback on the deviation of arm posture. McDaniel et al. [14] propose a framework for mapping vibrotactile patterns to motion.

Pneumatic pressure could potentially provide well localizable high-intensity stimuli, but is less popular due to size and tubing complexity of the required compressor.

Skin stretching with customized actuators may work well for indicating directions as pin-based actuators enable fine-grained feedback to the human skin [10] (p198). Constructions are, however, often bulky.

Feedback affecting **proprioception** (i.e., the sense of body parts' positions) has been shown, for example, using mechanical configurations with exoskeletons and electrogalvanic stimulation [10] (p194). In this work, we focus on modalities that do not directly alter the user's posture as active user engagement and minimal interference with the user's motion is preferred in typical motor learning scenarios [19].

2.3 Feedback vs. Support

Stimuli that are provided for motion guidance can be classified into feedback and support [14]. Feedback observes and takes into account user performance through real-time measurements in a feedback loop. Support is provided based on a predefined sequence regardless of user actions. Continuity of the feedback coupling to the user's performance varies with systems and modalities. Visual feedback, for example, can be used for continuous quantitative feedback on the deviation from desired pose using arrows drawn for each limb [18] or a target pose and qualitative verbal feedback [3].

2.4 Type of Feedback

Various options exist to encode information within the feedback modalities. The **user state** can, e.g., be visually represented by a mirrored virtual avatar in the user's pose [16]. The presentation of **intended pose** through, e.g., a teacher or ghost avatar [1], can be considered either support or feedback, e.g., if the teacher adapts to user performance.

Intended movement encodes how a user is supposed to move, while **motion dynamics** can communicate motion speed, for example. **Timing** and **performance** (e.g., the deviation from an intended pose) can be quantified and encoded in more abstract notions using audio [3]. Finally, the **state of virtual objects or other users** can be incorporated, such as collision with virtual objects in augmented reality or telepresence applications.

2.5 Level of Abstraction

The level of abstraction with which the stimuli are presented to the user has a large impact on the response selection phase in human information processing. Preattentive processing handles stimuli in a highly parallel manner without focused attention, for example, when extracting simple visual features and primitives. Therefore, a lower level of abstraction might trigger a quicker response even without training [22], while abstract feedback requires more cognitive processing to select the appropriate response.

Portillo-Rodriguez et al. [16] provide direct feedback using a push/pull metaphor where two vibration motors on the left and right hand indicate hand distance by varying vibration intensity. McDaniel et al. [14] explore temporal and spatial factors in vibrotactile patterns to encode direction, rotation or timing. They exploit the Sensory Saltation tactile illusion [8], where a series of sequentially activated vibrotactile motors can provide the perception of an interpolated, continuously moving linear stimulus, rather than one that appears and disappears in discrete locations. Israr et al. [9] equip a chair with twelve vibration motors to let users experience the illusion of moving touches on their back. This spatio-temporal pattern display technique enables more informative indication. In contrast, symbolic feedback with a learned meaning, such as Braille or pin-based displays [10] are highly abstract and require significant cognitive processing and training. Many examples of different levels of abstraction for visual and audio feedback can be found throughout HCI. For motion guidance, we can consider, e.g., a virtual hand to follow vs. a symbolic arrow, or a sinusoid played back in stereo headphones vs. a spoken command, to indicate direction.

Level of abstraction can also be viewed from the perspective of shift in attention [10] (p34) [22]. The sudden appearance or disappearence of a stimulus causes an exogeneous, automatic shift in attention. Symbolic cues (e.g., words or signs), on the other hand, cause an endogeneous (performer-driven) shift of attention and will in general require more processing capabilities, reaction time and training. A shift in attention also depends on the perceptual expectation of the user [10]. Therefore, training affects the S-R phase in information processing as well as the stimulus identification. However, generalizing attributes for low/high levels of abstraction is difficult, due to limitations of human cognition and potential interference between modalities and actuators.

2.6 Complexity

Complexity, or the amount of cognitive effort on the user, is not only affected by the level of abstraction, but also by a number of other factors. **The size of the response set**, or the number of movements conveyed through the feedback, should be kept to a minimum or be reduced by combining triggered movements into sequences. **The length of a movement sequence** does, on the other hand, also increase response selection time. An increased **S-R frequency** can, however, be exploited to decrease response selection time, for example, in training. Each application thus needs to find an optimal trade-off between these factors, based on its specific requirements.

2.7 Spatial Locality

Spatial compatibility [6] between a stimulus and response is crucial for an effective feedback system. There are usually advantages in applying a stimulus in spatial proximity to the limbs participating in the intended movement. More recently, this has been explained with a shift in attention [10] (p36) towards the designated region. If a (trained) response set exists, it usually causes an action, an effect also called precuing [10] (p15), which works better with decreasing complexity. Spatial locality is crucial for haptic feedback, but applies also to other modalities, e.g., for visual feedback on the side of the screen that affects that side of the user's body.

Spatial locality for positioning a display for visual feedback is also an important factor. If the user is expected to move around in the physical space, then the display can, for example, be worn (e.g., HMD [18]) or follow the user (e.g., robotic display [15]).

Psychophysichal interference between actuators, such as spatial masking effects for vibrotactors [5], has also to be taken into account. The spatial location of potential distractors thus has to be considered in a trade-off between spatial compatibility and interference.

2.8 Synchronicity & Timeliness

Synchronicity is relevant between user performance and intended pose/movement, as well as, between different modalities. While timing and speed is important for many motion guidance applications, systems often accept the delay between intended and user movement. Feedback could, however, be given ahead of time using **movement prediction** in certain situations, to compensate for system delay and user reaction. Furthermore, the feedback loop usually only compares the state of a user's movement with that of a teacher, which may result in unintended fast or intense corrective motions by the user. To our knowledge, no feedback system takes into account body dynamics. Techniques for prediction and interpolation exist in motion animation [21], but these have not been applied to motion guidance.

Furthermore, while a system should provide stimuli from different modalities synchronously, it needs to take into account that close temporal proximity may cause stimuli to interfere [5].

2.9 Quantification of Performance

The difference between intended and actual movement has to be evaluated to calculate deviation or error for the feedback in a motion guidance system. How this is done, depends on application and the feedback's abstraction level. It can be based on the state or the dynamics of a motion's basic parameters, considering spatial and temporal attributes.

Dynamic real-time feedback is constantly adapted in a closed loop [16], as opposed to offline feedback or pure sup-

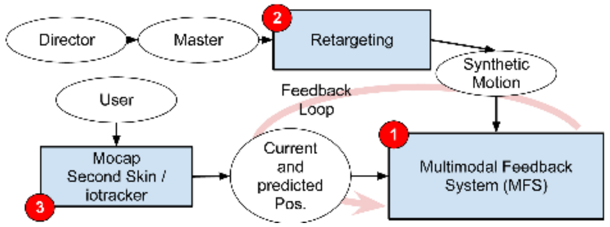


Figure 3: Our multimodal guidance system is based on (1) multimodal feedback, (2) dynamic motion retargeting, and (3) motion capture.

port systems. Motion capture is often used to acquire and compare a user's posture to a prerecorded representation. Visual, haptic or audio feedback can be given for the current state (e.g., virtual avatar [1]) or angular state-differences to a teacher position [12]. However, current systems don't consider dynamics and timing in their feedback loop.

A user will be able to follow feedback better, if it is not opposed to his current body dynamics. **Motion retargeting** has been used for animations to generate physically plausible motions e.g., [21, 13]. However, limitations on processing time are inherently different in real-time MGSs.

Quantitative performance can be measured based on angular joint errors [12], target positions [17] or paths, intra-body constraints, body volume [3] or derived parameters like speed, smoothness and energy efficiency.

Dynamic time warping through scaling in the time domain can provide an error measurement for an offline feedback process after the trial [1], where extreme positions in joint angles are used for alignment. While motions can be represented as a sequence of states [16], we are interested in using the timing information based on motion trajectories to improve student/teacher synchronizity, just as a human coach would adapt to the skill and speed of a student.

3. MOTION GUIDANCE SYSTEM

We have designed and implemented a Motion Guidance System (MGS) based on visual and tactile feedback, with a number of modules that form a subset of the described design space. The components in our system can be flexibly combined into new applications thanks to their modular nature.

Our implemented system consists of three classes of components, as illustrated in Figure 3:

- Motion tracking. We use motion capture systems to track the user's pose, such that appropriate feedback can be generated based on limb or body position and orientation.

- Dynamic motion retargeting. We provide automatically generated feedback based on current and predicted user motion, which can be merged with interactive manual control from an experimenter or director.

  Motion retargeting techniques (detailed in 3.3 and 3.4) enable our system to generate target motion adapting to user motion and state. In order to generate dynamic feedback we don't solely rely on the user's pose at a given time, but are using methods usually applied in motion retargeting for animation to predict future postures based on body dynamics. The user's required effort is minimized by finding the optimal point in time, where user and target trajectories can be matched and aligned.

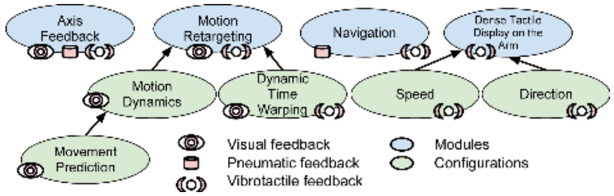


Figure 4: Overview of the implemented modules, their configurations and available modalities

- Multimodal feedback. Our modules generate visual, vibrotactile and pneumatic feedback for generic use with arbitrary limbs, or for specific tasks like arm movement or navigation (See Figure 4 for an overview).

3.1 Position and Direction

As previously discussed, it is often important to provide continuous feedback with a low abstraction level, for quick response selection by the user, as well as having an emphasis on high spatial locality. The rotation of a joint in the human skeleton has a maximum of three DOFs and in a low-level, bottom-up approach, we could consider giving feedback on these three DOFs to each limb in a similar and flexible manner.

The basic concept is to use two tactors to indicate rotation around one axis by activating a tactor on one side to push or pull the limb in a direction [12, 11].

Our implemented module is used to guide a user's limb into a predefined, or interactively specified, posture using two actuators per DOF, to indicate directional movement, as illustrated in Figure 5.

An error function calculates the feedback strength and direction based on current and target rotation, and a director can adjust the feedback, or add/remove DOFs for different limbs through our GUI (Figure 6).

Besides the primary vibrotactile feedback for this module, we also support pneumatic actuation, which provides a stronger tactile sensation. More complex mechanical configuration, however, limits its scalability to a larger number of DOFs.

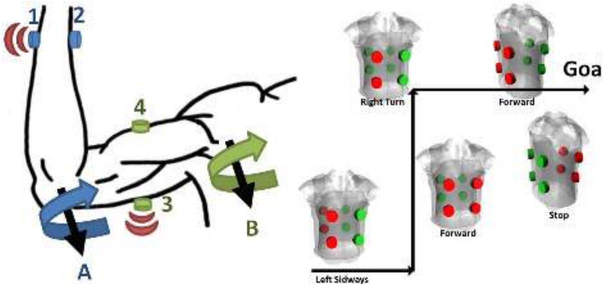
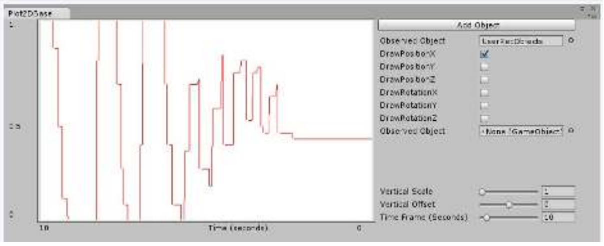


Figure 5: (Left) Vibrotactors 1-2 and 3-4, indicate rotations around axes A and B, respectively. (Right) Activation patterns for navigation with pneumatic feedback vest (active tactors = red).



Observed Object [illegible]
DrawPositionX ☑
DrawPositionY ☐
DrawPositionZ ☐
DrawRotationX ☐
DrawRotationY ☐
DrawRotationZ ☐
Observed Object <None (GameObject)>
Vertical Scale 1
Vertical Offset 0
Time Frame (Seconds) 10

Figure 6: Our GUI provides straightforward access for controlling and adjusting tactile feedback.

These approaches provide high spatial locality and a low level of abstraction. However, they also generate a vast amount of information for multiple limbs, dramatically increasing complexity and cognitive processing time, which might cause psychophysiolocial interference.

Visually, axis feedback can be implemented as arrows drawn over the virtual representation of a user to indicate movement directions for certain limbs.

Similarly to [18], our implemented visual arrows show deviation from intended pose on the user's avatar compared to the teacher avatar (See Figure 7, left). The added arrows increase the level of abstraction, since they need to be interpreted by the user. Nevertheless, they should improve performance when it is difficult to identify small pose differences.

We also support prerecorded target movements, where the feedback dynamically adapts to the user's motion and the current state of the played back sequence.

The position of the end effector (e.g., hand [11]) can be sufficient for many applications, which allows a reduction of the response set, as suggested in Section 2.6. If a larger number of DOFs need to be taken into account (e.g., [12]), then spatial locality and possible resulting interferences become an important factor.

3.2 Continuous Feedback and Velocity

Interaction in many applications is focused on either hands or arms, and it is therefore useful to encode more information in these areas. This is challenging with the low abstraction level from the previously described two-tactor approach for position and direction, due to the limited area and resulting spatial and temporal interference.

Vibrotactors are, however, robust, relatively small and lightweight, consume little power, and can be used wirelessly and in spatial locality. Controllability of frequency and amplitude with quick actuation allows for implementations of different levels of abstraction and information encoding. This has made them popular for many applications, including our implementation of a dense tactile display for an arm.

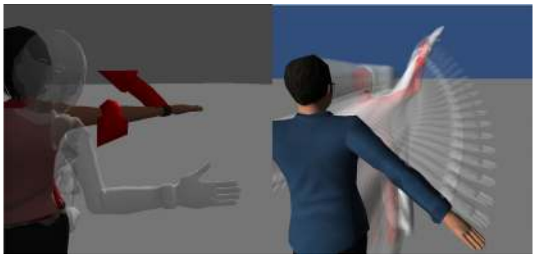


Figure 7: (Left) Motion guidance using ghost avatar and directional arrows. (Right) Motion sequence with transparent avatars, where the red avatar indicate a compulsory pose.

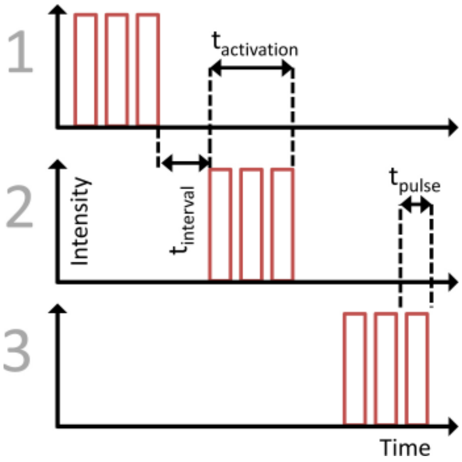


Figure 8: Sequential pulsing of three vibrotactors to indicate directional speed.

The sensory saltation effect can also be employed to add information like vectors or speed of intended movement. While the saltation effect has already been used to indicate rotations of the arm [12] or direction and rotation in a planar setting [9], indication of a vector in three dimensions on the arm, and especially speed, have been less explored.

Our implemented module is designed to provide movement speed sensation in three directions by employing a dense tactile display, where speed is indicated by triggering the vibrotactors in sequence, as shown in Figure 8.

By controlling burst durations and onset times, perceived stroking movements can be generated at a desired target speed. The actuators are turned on for a pulse duration (t_{pulse}) of 20-200 ms, where 20 ms was chosen as the minimum speed which subjects could perceive as a moving tactile stroke in a pilot study. With a t_{pulse} of 200 ms, a single loop of indication would take approximately two seconds, which was chosen as a practical maximum for our applications. $t_{activation}$ is the sum of pulses of a single tactor. The actuation intervals are calculated from user anatomy (i.e., arm length) and target velocity (v_{target}) using the equations:

$$t_{interval} = \frac{\frac{armlength}{v_{target}} - (3 * t_{activation})}{2} \tag{1}$$

$$t_{interval,calibrated} = t_{interval} * factor_{calib} \tag{2}$$

Preliminary experiments with five study participants on perceived absolute speed, indicate individual differences that can be corrected with a calibration factor ($factor_{calib}$), which we plan to explore further in future work.

Directional speed can only be triggered serially as shown in Figure 10, through sequential tactor activation, and depends on a presented prerecorded teacher movement. In the current implementation the sequences are independent of the user's performance and can be considered support rather than feedback.

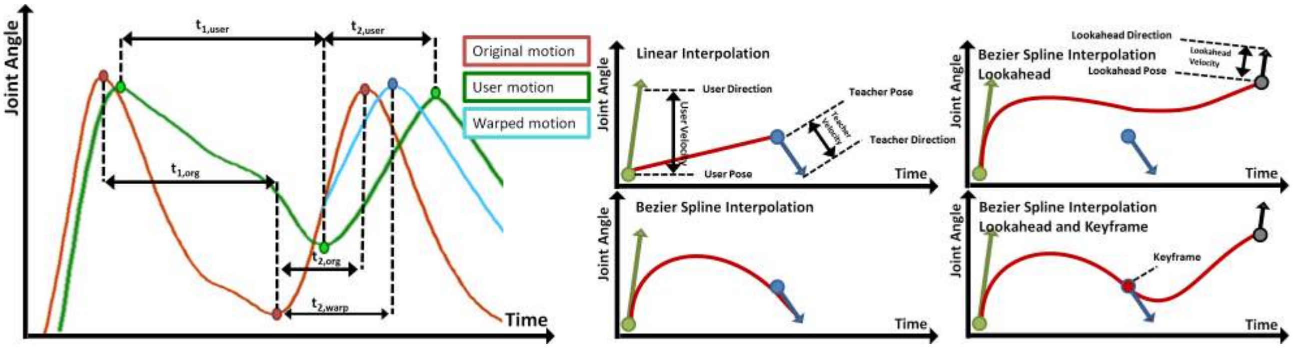


Figure 9: Dynamic Motion Retargeting. (Left) Target motion is time warped to user performance. (Middle) Path interpolation: Linear vs Bezier spline. (Right) Lookahead and keyframe indication.

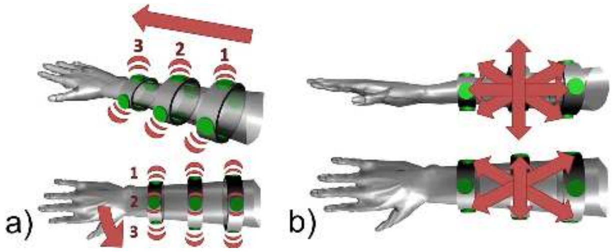


Figure 10: Sequential triggering of vibrotactors can be used to indicate a) speed and b) direction vectors.

The module implementing the rich vibrotactile display can also be configured to present translational forearm-directions through sequentially triggered tactors in seven different directions (See Figure 10). It allows the guidance towards target poses, where target speed can be chosen according to factors such as desired loop time or optimal user perception.

3.3 Dynamic Time Warping

In some situations, where complex feedback prevents a user from reacting in a timely manner, it might be a viable alternative to adapt the timing. When the quantification of performance is based purely on the user's state (e.g., joint angles) and the error needs to be minimized, the intended movement's speed can be matched with the user's movements.

We implemented dynamic time warping for non time-critical sequences, to adjust the teacher's speed and target pose to user performance, using two different approaches.

First, we extend Lieberman and Breazeal's work [12], where an error function is used for evaluation, and apply it to real-time feedback. The error function is calculated for each frame based on joints' angular (and velocity) differences and increases with deviating user motion, such as when user lags behind or performs an incorrect movement. When the error reaches a configurable threshold, the teacher is continuously slowed down towards a static posture or a defined minimal velocity, to allow the user to catch up or correct the motion. Once the error function estimates that the user is acceptably close to the target posture, speed is gradually increased until original velocity is reached.

Our second approach analyzes and searches for minima and maxima for target and user's joint angles' curves, as shown in Figure 9. The curves are matched for each joint axis and the time duration between a number of recent extreme points is measured. Figure 9 shows a warping example where the original motion is slowed down by the factor $t_{1,user}/t_{1,org}$.

An average speed-adjustment factor is calculated based on matches and factors of multiple joints and DOFs to match the teacher and student motion in time. The procedure is repeated as soon as a new local extreme point is discovered in the user's motion for any DOF. The algorithm requires careful parameterization and works only with a limited number of DOFs. The main challenges are correct detection and matching of extreme points over multiple DOFs and a combination of the resulting factors to single warp factor.

In future work, we plan to extend our approach by using subgestures to determine the timing differences between states [16].

3.4 Motion Dynamics and Prediction

Prediction or extrapolation of the user's movement, and consideration of future states of the intended movement, can be used to guide the user towards smoother and more natural movements. The prediction of a user's motion can be based on the movement speed and direction.

Motion dynamics are used to interpolate between user motion and intended movement for feedback on the required corrective path to produce smooth and natural user movement. Bezier spline interpolation is applied between user and teacher joint angles, taking into account rotational speed for each limb, as illustrated in Figure 9. The user's current movement dynamics are also compared to the teacher's future motions. If a teacher pose and velocity is found in the (parameterizable) near future that better matches the user's state than the current target state, then the user may be redirected towards that motion instead. The future state and its velocity is then added as an additional control point and parameterization in the Bezier interpolation.

It might sometimes be necessary to direct the user's selective attention to specific poses or movements. The intended pose might even be more important than the intended movement in certain choreographies, requiring a mechanism to deal with these situations. We have therefore implemented the option for a posture to be interactively marked as a keyframe by an experimenter/director.

Keyframes inform users of important poses that need to be executed and cannot be skipped due to deviating user motion (Figure 9). These keyframes are also emphasized when calculating trajectories by changing the weight of the control points in the Bezier spline. This enables a trade-off between accurate and smooth motion. If less keyframes are annotated beforehand, the system suggests overall movement. More keyframes, on the other hand, require accurate and detailed motion.

Figure 9 shows the possible configurations for the motion dynamics module in a one-dimensional example. Linear and Bezier spline interpolations are used to calculate movement paths with and without taking velocities into account. It also illustrates how a future (lookahead) pose and movement velocity can affect the movement path shown to the user.

When feedback has to be given on more complex movements, visual feedback works well [1, 16]. Presenting a ghost or teacher avatar to show the intended pose and another avatar animated with the current state of the user, is a well-established method. While the representation has a higher level of abstraction, the complexity can be handled because following and imitating the movements of others is a natural learning concept.

Currently, we use visual feedback through ghost target (teacher) and user avatars for motion dynamics and prediction. The avatars consist of a flexible number of animated body parts, which can be aligned and presented for optimal spatial locality based on the scenario (e.g., mirrored or watched from behind).

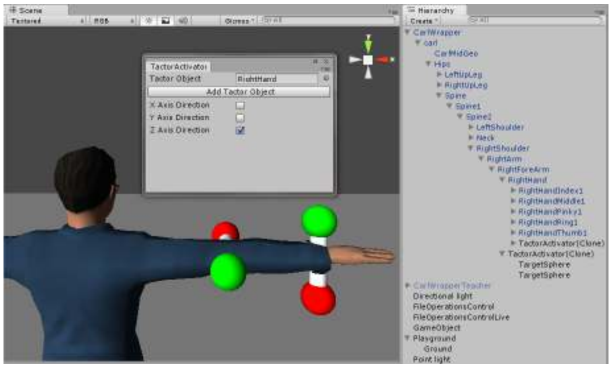
Furthermore, we have implemented motion guidance paths for feedback on movement dynamics. Instead of a single teacher state (e.g., [16, 18]), we incorporate a movement sequence (e.g., up to a few seconds) into our feedback loop. The visualization consists of multiple transparent avatars in different postures, rendered behind each other to visualize not only a single state of the teacher, but a whole motion sequence (See Figure 7, right). Keyframes are visualized as avatars colored in red and move towards the user during the interpolation.

3.5 Haptic Feedback

Multimodal feedback can also be given to communicate the state of objects in a virtual scene. From visual feedback alone, interaction with virtual objects is often difficult, since depth is hard to judge even with stereoscopic displays, when the sense of touch is missing.

In a movie set, for example, where virtual content is to be merged with video material in a green room environment, actors are usually not aware of their exact position relative to the virtual objects. Current systems deliver visual feedback on the mixed reality configuration, but require the actors to look at screens mounted around the set. Our system could provide tactile feedback as the primary modality, which could be embedded in the actor's clothing. Multiple tasks are useful in this scenario, including feedback on pose, path, reaching [11] or multiple DOFs tasks for one or multiple actors.

Our implemented module augments interactions with virtual objects by using physics simulations to detect collisions with the user's body and generates 3 DOF directional tactile feedback. Our GUI (See Figure 11) makes it straightforward to visually map the feedback to different limbs of a virtual avatar and to arbitrary feedback axes on the body. This module calculates direction and intensity of the feedback and activates the corresponding tactor, which can be used flexibly for different body parts and objects due to its low level of abstraction and complexity.



Tactor Object: RightHand
X Axis Direction ☐
Y Axis Direction ☐
Z Axis Direction ☑

Figure 11: Our graphical interface allows direct editing of haptic feedback on an avatar in a virtual 3D environment.

3.6 Guided Navigation

Feedback on intended movement (path) or pose (target position) of the user in space is important for spatial navigation. Most scenarios primarily deal with a user moving in a plane and we therefore consider only the 2D position and orientation of the user in this plane. For tactile feedback, the torso provides a relatively large input area and interference between tactors is limited. Although providing feedback for navigation to the torso isn't spatially local to the legs, multiple studies show promising results (e.g., [14]).

We steer tracked users towards a target position through generated tactile feedback, as shown in Figure 5, right . Our feedback patterns control 3 DOFs; forward/backward translation, sideways translation, and rotation around the up-axis. Our implementation uses pneumatic feedback since vibrotactile and rotational patterns are less clearly perceived during physically and cognitively demanding tasks [20]. We have, however, also implemented vibrotactile feedback to allow comparisons in a future evaluation.

Directions of translation and rotation are presented sequentially to avoid temporal interference, and preliminary experiments with multiple activation patterns indicate that pulsing at approximately 3 Hz gives clear feedback for our pneumatic actuators. We plan additional experiments to investigate this further.

3.7 Implementation

Our current applications employ the iotracker [19] or Second Skin [7] motion capture systems to track the user's 6 DOFs pose.

Vibrotactile feedback is provided using tactors from Audioacoustics Engineering. Our dense tactile display uses twelve vibrotactors on the arm, as shown in Figure 1 c). Our custom control board drives the tactors at 250 Hz, the recommended frequency for human skin perception [5], and communicates wirelessly with the host PC using bluetooth. We use pulsed vibrations, instead of continuous, as it is better for perception, as shown in related work [2] and confirmed in our early experiments. We have experimented with different activation patterns and tactor configurations. Figure 10 shows the positions where vibrotactile stimuli are applied for our dense tactile display.

Pneumatic feedback is provided using a 3RD Space Vest from TN Games, which applies pressure using four actuators on the chest and four on the back, as shown in Figure 5.

The modules and feedback loop are implemented in C# with the Unity 3.5 game engine. The visual feedback and virtual world is generated with Unity's renderer, while our C/C++ plugins control the tactile feedback modules, and interface with the motion capture systems using the Opentracker middleware. The system and tracking server run on a single PC (Intel i7-2600K, 16 GB RAM).

4. CONCLUSION & FUTURE WORK

In this paper, we discussed the design space, challenges and solutions, potential scenarios, and our implementation of a multimodal motion guidance system. Our system integrates multiple modules that generate feedback based on a user's performance and intended movement. We have integrated a number of different feedback modalities (vibrotactile, pneumatic, and visual) and shown their application for different scenarios and use cases.

Due to the vast nature of feedback options and complexity of human cognition and motor system, we restricted this work to 3D motor movements that are tracked by a motion capture system, and implemented a set of feedback modules, which appeared most promising and generic, after an analysis of the design space.

We are currently extending our real-time authoring framework to better support prototyping, as well as scientific evaluation of motion guidance applications. We also plan to present results from our ongoing evaluation of individual feedback modalities and their combinations.

It would be interesting to explore adding modules for finer feedback granularity (e.g., feedback on finger movement) and integrate other technologies in the system (e.g., HMD for augmented reality, or heat feedback). Finally, we will extend our work from motion guidance to motor learning, focusing on cognitive aspects and processes that are relevant in motion feedback.

5. ACKNOWLEDGEMENTS

This work was funded in part by the European Union within the PROFITEX project (FP7-NMP2-SE-2009-228855). Alex Olwal was supported by a Swedish Research Council Fellowship and a Blanceflor Foundation Scholarship.

6. REFERENCES

[1] S. Baek, S. Lee, and G. J. Kim. Motion retargeting and evaluation for VR-based training of free motions. *Science And Technology*, 2003.

[2] W. Borchers, J. and Prinz. *Design and Recognition of Tactile Feedback Patterns for Snowboarding.* PhD thesis, RWTH Aachen University, 2008.

[3] E. Charbonneau, A. Miller, and J. J. LaViola. Teach me to dance: exploring player experience and performance in full body dance games. In *Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology - ACE '11*, pages 43:1–43:8, Nov. 2011.

[4] M. Dozza, F. B. Horak, and L. Chiari. Auditory biofeedback substitutes for loss of sensory information in maintaining stance. *Experimental brain research*, 178(1):37–48, Mar. 2007.

[5] J. B. F. V. Erp, H. a. H. C. V. Veen, C. Jansen, and T. Dobbins. Waypoint navigation with a vibrotactile waist belt. In *ACM Transactions on Applied Perception*, volume 2, pages 106–117, 2005.

[6] P. Fitts and M. Posner. *Human performance.* Brooks/Cole., Oxford, England, 1967.

[7] K. Fukushi, J. Zizka, and R. Raskar. Second Skin: Motion Capture with Actuated Feedback for Motor Learning. In *ACM SIGGRAPH 2011*, page 4503, 2011.

[8] F. A. Geldard. *Sensory Saltation: Metastability in the Perceptual World.* John Wiley & Sons Inc, 1975.

[9] A. Israr and I. Poupyrev. Tactile brush: Drawing on Skin with a Tactile Grid Display. In *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11*, page 2019. ACM Press, May 2011.

[10] J. A. Jacko. *The Human Computer Interaction Handbook.* CRC Press, third edition, 2012.

[11] M. Klapdohr, B. Wöldecke, D. Marinos, J. Herder, C. Geiger, and W. Vonolfen. Vibrotactile Pitfalls: Arm Guidance for Moderators in Virtual TV Studios. *Information Systems*, pages 72–80, Dec. 2010.

[12] J. Lieberman and C. Breazeal. TIKL: Development of a Wearable Vibrotactile Feedback Suit for Improved Human Motor Learning. *IEEE Transactions on Robotics*, 23(5):919–926, Oct. 2007.

[13] C. K. Liu and Z. Popovic. Synthesis of Complex Dynamic Character Motion from Simple Animations. *ACM Trans. Graph.*, 21(3):408–416, 2002.

[14] T. McDaniel, D. Villanueva, S. Krishna, and S. Panchanathan. MOVeMENT: A framework for systematically mapping vibrotactile stimulations to fundamental body movements. In *Haptic Audio-Visual Environments and Games (HAVE), 2010 IEEE International Symposium on*, pages 1–6, 2010.

[15] A. Nakamura, S. Tabata, T. Ueda, S. Kiyofuji, and Y. Kuno. Dance training system with active vibro-devices and a mobile image display. In *Intelligent Robots and Systems, 2005.(IROS 2005). 2005 IEEE/RSJ International Conference on*, pages 3075–3080. IEEE, 2005.

[16] O. Portillo-Rodriguez, O. O. Sandoval-Gonzalez, E. Ruffaldi, R. Leonardi, C. A. Avizzano, and M. Bergamasco. Real-Time Gesture Recognition, Evaluation and Feed-Forward Correction of a Multimodal Tai-Chi Platform. *Haptic and Audio Interaction Design*, 5270:30–39, 2008.

[17] J. Rosenthal, N. Edwards, D. Villanueva, S. Krishna, T. McDaniel, and S. Panchanathan. Design, Implementation, and Case Study of a Pragmatic Vibrotactile Belt. *IEEE Trans. on Instrumentation and Measurement*, 60(1):114–125, 2011.

[18] Y. Ryota Sakamoto, Yuke Yoshimura, Tokuhiro Sugiura and Nomura. A motion instruction system using head tracking back perspective. *World Automation Congress*, 2010.

[19] C. Schönauer, T. Pintaric, and H. Kaufmann. Full Body Interaction for Serious Games in Motor Rehabilitation. In *ACM Augmented Human International Conference*, 2011.

[20] D. Spelmezan, M. Jacobs, A. Hilgers, and J. Borchers. Tactile motion instructions for physical activities. In *Proceedings of the 27th international conference on Human factors in computing systems - CHI '09*, page 2243, New York, USA, Apr. 2009.

[21] S. Tak and K. Hyeong-Seok. A Physically-Based Motion Retargeting Filter. *ACM Transactions on Graphics (TOG)*, 24(1):98–117, 2005.

[22] G. Wulf. *Attention and motor skill learning.* Human Kinetics, 2007.