

# ADVANCED AUGMENTED REALITY FEEDBACK FOR TEACHING 3D TOOL MANIPULATION

**Tobias Blum, Tobias Sielhorst and Nassir Navab**

Chair for Computer Aided Medical Procedure and Augmented Reality, TU Munich

Blum/sielhors/navab@in.tum.de

**Abstract:** In this chapter we describe a novel technique for automatically providing elaborate feedback to AR/VR trainees. Our focus is on teaching 3D manipulation of tools in an AR enhanced simulator.

Feedback is a critical part of any learning procedure. This is also true for VR and AR teaching systems. We propose a method for providing feedback without time consuming authoring of possible mistakes and advice on improvements. We have implemented the proposed method and show our first results.

Since 3D movements cannot be trivially compared to each other the dynamic synchronization of expert and trainee performances is a crucial part of our system. Showing a synchronized replay of the expert and the trainee's movement, our system reveals differences without requiring the presence of the expert. Furthermore, we use synchronization for statistical analysis of several expert performances for averaging out unimportant variations in the manipulation path. We visualize the variance of expert movements and provide the trainee with appropriate feedback depending on such variance.

## 1. Introduction

Virtual and augmented reality techniques in education have shown to be promising for elevating the quality of training and for coping with increasing financial pressure. In fact, the main cost of education, in particular medical training, is due to the expensive time spent by many experts to provide continuous feedback to trainees. When looking at existing systems, two main advantages over conventional methods can be identified. On one hand, AR/VR systems visualize instructions and learning content without the attendance of a teacher, which for complex tasks is superior to other presentation methods, and also reduces workload from human teachers resulting in cost savings. On the other hand, AR and VR enable us to simulate situations and environments for training, which would be very expensive to realize in reality or in some cases even impossible to realize. However we believe that in addition to giving instructions for performing the task, providing trainees with feedback on their execution of the task could also be realized in a much more vivid, effective and even economic way than it is done today.

### 1.1 Feedback

In her examination of feedback research Mory [1] defines feedback as presented information that allows comparison between an actual outcome and a desired outcome. We will use this definition throughout our work. The importance of feedback for learning is well known and has for instance been expressed by Annett [2]: "Informing the learner of the outcome of each response or trial (knowledge of response, or KR) typically gives the most rapid learning, whereas no-KR, or practice-only conditions generally show poor learning or none at all. The rate and extent of learning is sensitive to the amount of information given - the more detailed the KR the better the learning." Besides giving feedback in written form, which is usual for written tests, feedback normally is given verbal, which is certainly in most cases the best way a learner can be informed about his performance, mistakes and how to correct them. However an automated approach has also important advantages. As the attendance of an expert is not necessary, which can be problematic especially for doctors in medical training, the costs could be lowered and the trainees could use the self dependent system whenever they want. Another benefit comes up in particular for learning tasks which involve motor skills. One problem here is that motor skill is conventionally distinguished from verbal knowledge and the former is often inaccessible to the latter [2]. One example Annett gives on this, is that skilled swimmers cannot answer factual questions about the breast-stroke any faster or more accurately than novices. He also mentions that the use of kinematic data may be helpful to identify critical features of the performance, whereas when using e.g. video recordings the need for an expert interpretation could arise again.

Our approach is to some extent related to giving feedback by analyzing kinematic data using video. However we use Augmented Reality for visualization not only of kinematic data, but to give feedback in a more sophisticated way. We track tools used to accomplish a task both for the expert's and the trainee's

performance. To give feedback, we visualize these simultaneously and synchronously, which is described in detail later. Using statistical analysis we also derive additional information from the expert's performances, which is utilized for giving more intelligent feedback. Doing this in a reasonable way is certainly an objective that can not be achieved for all different educational exercises with one general approach. So in this article, we want to show the importance, potentials and one possible solution for a series of educational challenges, which include the teaching of subtle movements in a tool manipulation task.

After looking at related work in the next subsection we will describe the technical setup of our current system in section 2. In section 3 we explain the importance and the mathematical fundamentals of synchronizing multidimensional data such as 3D movements. Section 4 shows different examples of providing feedback with an Augmented Reality training system that makes use of recorded and intelligently processed expert data. Section 5 discusses the current results and future directions.

### *1.2 Related Work*

Liu [3] uses a haptic interface and on-screen visualization for medical education. While simulating diagnostic peritoneal lavage, feedback is given in different ways. During the simulation, the trainee has to select the correct next step of the procedure from a randomized list. At the end incorrect choices are reported and so an evidence of the procedural knowledge is given. Feedback is also given when the trainee selects the site to make the incision, by providing information why an incorrect chosen site is not suitable. During the insertion of the cannula, entry angle as well as location and depth of insertion are compared with instructor-defined optimal values. Excessive deviation (also instructor defined) is reported as an error, and the trainee is asked to try again. This feedback certainly accomplishes the requirements of this simulator, but in contrast to our approach it had to be authored for this specific task and would not be applicable for teaching complex 3D manipulation, which is our objective.

The assessment and evaluation of skills is a field of active research and does also incorporate knowledge about the performance of the trainee. Biggs [4] defines assessment as summative evaluation of student performance after a teaching episode, usually quantitatively conceived. As automated assessment mostly only needs one quantitative value as result, which reflects the skill of the learner, normally not enough information is available to provide the learner with elaborated feedback and it is only of smaller use for the assessed one. So Eyal [5] and Nizard [6] used the duration of performance as indicator for learning progress in virtual reality based and computer aided learning. The ICSAD [7,8] uses the time taken, the number of movements made, the speed of travel, and the distance traveled by each instrument to assess the performance. Also statistical methods have been used [9,10] whereas the methods also aim at segmenting the task in order to achieve a better assessment of the performance. But all these methods only provide a quantitative assessment and do not provide any data which helps the trainee to identify and learn from his mistakes.

## **2. Description of our AR-based teaching system**

For our setup, we use a combination of AR visualization and a physical simulator. With this combination we make use of the advanced visualization possibilities of an AR/VR system and the superior haptic properties of a physical simulator as opposed to simulated haptics. The decision in favor of an AR system as opposed to pure Virtual Reality is based on the fact that we can use the realistic look and feel of physical simulators without modeling and rendering them. Therefore we have to only render the additional elements, e.g. in case of a human delivery simulator the head of the baby.

### *2.1 AR system*

The augmented reality system we use is the research system RAMP. It has been developed by Siemens Corporate Research (SCR) for real time augmentation in medical procedures [11]. It is optimized for accurate and real time alignment of the real and the virtual scene. The system features a high resolution video see through HMD and infrared inside-out tracking. Its accuracy, high resolution, high update rate and minimal lag is currently state of the art.

In order to visualize the performance of an expert, which has been recorded previously, the 6DOF of the instruments must be tracked. There are different tracking solutions. For two reasons we decided not to take the same tracking system as the AR system does. First, there is the line of sight problem. Some of the targets for the single camera tracking of the AR system are likely to be occluded [12]. Note, that the system has to see at least eight markers for accurate, reliable and robust tracking. Second, the error function of

targets of the single camera tracking is unequally distributed in space [12]. The least accuracy is in the viewing direction. For real-time augmentation, this is fine, because less accuracy is needed in viewing direction for a satisfactory overlay that results in a minimal alignment error in the 2D image [13]. We intend to record the movements of the instrument and visualize it from potentially any direction. Therefore we have chosen an external tracking system that tracks with a more equally distributed error function and uses a minimal set of markers per target in order to estimate position and orientation. Practically we use a multiple-camera infrared tracking system by A.R.T., which uses retro-reflective markers like the inside out tracking of the AR system. This makes it simple to register one tracking system to the other. Both tracking systems can therefore be registered using a common reference target. Thus, we can provide the instrument's position in a common coordinate system. In order to transfer the coordinates of the instruments, which are provided by the outside-in tracking system, into the coordinate system of the marker frame, a simple equation can be applied. This is possible because both tracking systems use the same model data for the marker frame, and both systems generate the coordinate system in the same way from the model data. With this combination, the system is highly dynamic and robust. With this setup we can move the cameras of the external tracking system while the system is running, without any need for re-calibration.

## *2.2 Targeted simulator*

The AR/VR learning system we describe in this chapter is not a stand-alone application. It is meant as a valuable addition to existing simulators. Our studies are directed to the human delivery simulator of Burgkart, Obst et al. [14]. The delivery simulator is a full featured simulator that consists of a hardware and a software model. The hardware model is a female body phantom with the baby's head on a robot arm inside. The baby is reduced to its head, which is acceptable for delivery simulation. The use cases include uncomplicated as well as many pathological deliveries, because the head is usually the most important birth obstacle. The head of the baby phantom has a force/torque sensor for haptic interaction with the user. The simulator moves the head of the baby accordingly to the birth process.

The software model comprises of two parts. The physiological part provides real time values of blood pressure, heart rates, pain and oxygen supply. The values are generated using the contraction on oxytocine, fatigue of mother and fetus, oxygen supply to the child, and individual boundary conditions. These boundary conditions can be e.g. mother's oxytocine production, stress and pain sensitivity, heart volume, as well as many fetal parameters. The biomechanical values are calculated into the physiological model, too, where the forces may cause e.g. pain or blood loss.

The biomechanical model provides values in real time for the position of the baby's head based on contraction forces, friction in the birth channel, tissue forces, and the user applied forces. The contraction forces are updated by the physiological model.

In former studies [15] we have proven the concept of a collaboration between both systems, the delivery simulator on one hand and the AR system on the other hand. We aim at giving automated feedback on the performance of a forceps delivery that can be taught with the delivery simulator. The forceps is a medical instrument that is used by gynecologists for extraction of the baby in the critical case of birth stop. The setup can be seen in figure 1. In order to focus on the concepts, in this chapter we present a more general approach to the problem of giving automated feedback to the trainee/student.

## *2.3 Omnidirectional visualization of the experts performance*

For optimal learning of complex spatial movements and tasks, the desirable procedure is an expert demonstrating actions and giving feedback to the practicing trainees immediately. In general, the schedule of experienced experts is tight. This makes it difficult to demonstrate the task to each individual trainee and provide him/her feedback. In addition, it would be desirable to allow the majority to learn from the best international experts in their field. So we are tracking the movement of a tool used by an expert while he/she uses a simulator or performs a real (often complicated) task. This information is reproduced for demonstration to trainees using Augmented Reality. Using a computer model of the tracked instrument and the movement which has been recorded, we are able to show the experts performance to the trainees. A major advantage of doing so is the omnidirectional visualization. Trainees can study the expert's movements from any direction and are also able to stop, rewind or slow down the replay. This is already an important step towards a training system, which does not rely on the permanent attendance of an experienced teacher. However only studying the expert's movement is not enough. Without the possibility to get feedback on their own training runs, students still need the presence of an expert teacher.



*Figure 1. The AR delivery simulator setup; user working with human delivery system[14] wearing the RAMP HMD system.*

### **3. Synchronization**

#### *3.1 Motivation for synchronization*

As already mentioned, Mory [1] defined feedback as presented information that allows comparison between actual and desired outcome. When the feedback is given by authoring typical mistakes and comments for improvement, it will be very difficult to give feedback according to this definition. All possible relevant mistakes had to be authored in order to give appropriate feedback. We want to avoid authoring in our system and give automated feedback in accordance to Mory's definition. So we need to compare the desired outcome, which is the expert's performance, to the trainee's performance. Furthermore, we do not want only an assessment by one value but rather a visualization of this comparison to the trainee in a meaningful way. Comparing two movements in space is a non-trivial task. First of all we have to cope with the problem that both movements most likely have not been carried out with the same speed, and so a comparison is very difficult. Note, that teachers often stop in the middle of manipulation and explain some issues and then continue the manipulation. To allow a reasonable comparison, we have to synchronize both performances.

In figure 2 the trajectories of a tool which has been tracked during two attempts to perform the same movement is shown. It is obvious that both movements are similar, but when looking at the movement in x-direction over time, it can be seen that there is a big temporal difference. A simple approach to deal with this problem would be to scale them to the same length. Unfortunately this does not lead to satisfying result, because the speed at which both tasks are performed changes during execution. An example of this can be seen in figure 3. We need a synchronization of both movements so that variations in time are removed and similar steps are shown synchronously.

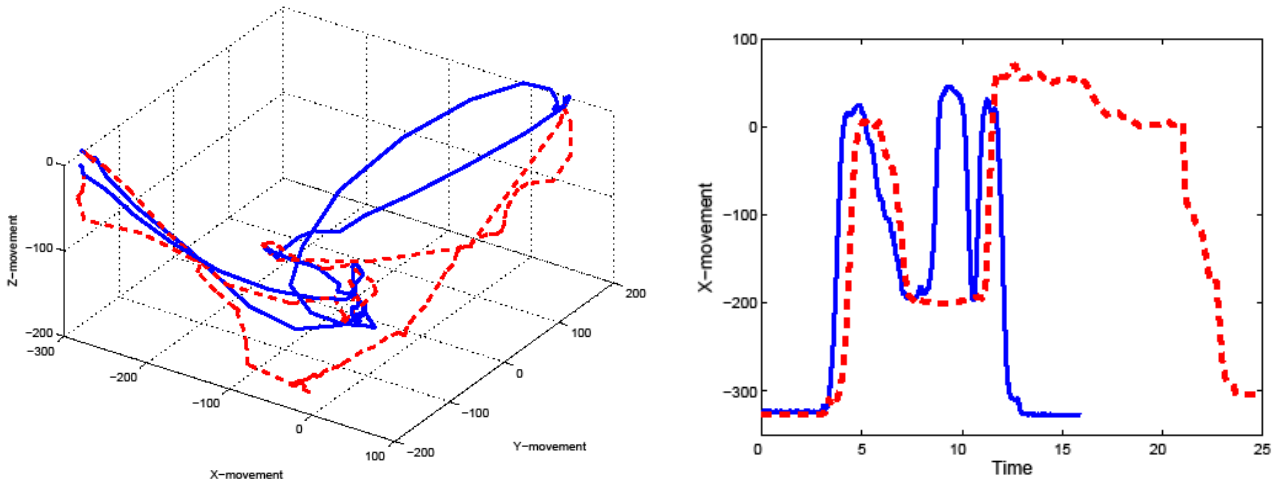


Figure 2. shows two performances of the same movements in space, the right figure show the x-movement of the same performances over time

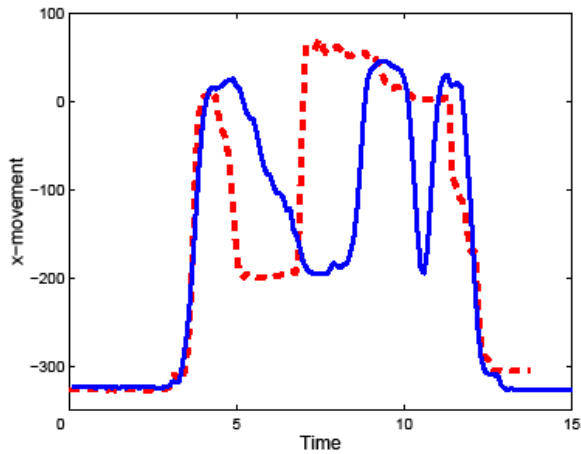


Figure 3. Both movements have been scaled to the same length.

This can be posed as a problem where we have two trajectories  $A = (a_1, K, a_m)$  and  $B = (b_1, K, b_n)$ ,  $a_1$  being the data we get at a certain time, in our case containing a time stamped location and rotation of the tracked object. We need a monotonous mapping between the points of both,  $W = ((i(1), j(1)), K, (i(K), j(K)))$  such that one trajectory is synchronized to the other one. Functions  $i$  and  $j$  define the mapping between the elements of the two series. This mapping  $w$  can also be seen as a warping function or warping-path that is applied to the time-axis of one trajectory and synchronizes this to the other one. There have been different attempts to cope with similar problems. For example, by searching for distinct characteristics in both trajectories a landmark-based synchronization can be achieved. This has been done e.g. by Wang [16] in order to analyze growth curves of different children. The approach suffered from the problem of determination of landmarks, which is time consuming and error-prone and has been done manually in [16]. Another method is the Longest Common Subsequence (LCSS) that has been used for defining similarity measures between mobile trajectories [17] by counting the number of points, which are similar for both trajectories. In doing so LCSS also provides matching points. We have tested the LCSS for our purposes and discovered that this method by definition has some problems, when both trajectories are too distant or the update rate of the tracking system changes. Instead we propose the use of the Dynamic Time Warping (DTW), which has originally been introduced for speech recognition [18], nonparametric sample curve synchronization [19] and signature verification [20].

### 3.2 Dynamic Time Warping

Let  $A = (a_1, K, a_i)$  and  $B = (b_1, K, b_j)$  be the two sequences of data, which need to be synchronized. The DTW is defined recursively on A and B as follows.

**DEFINITION:** Given a distance-function  $d(x,y)$ , the  $DTW(A,B)$  is recursively defined as:

$$DTW(A, B) = d(a_m, b_n) + \min(DTW(A_{m-1}, B_{n-1}), DTW(A_{m-1}, B_n), DTW(A_m, B_{n-1}))$$

The result is a similarity measure between A and B in regard to the distance function  $d(x,y)$ . While computing this similarity measure the DTW has to match every point from each trajectory with at least one point from the other trajectory. This can be seen in figure 4. This matching is done in a way so that the summed up distance between both trajectories,

$$\sum_{k=1}^K d(i(k), j(k))$$

, which also defines the similarity measure between both, is minimized. In each recursion step the last points of both trajectories have to be matched to each other. Either both points, or only one of them is left out in the next step, depending on what produces the lowest result. DTW can be computed using dynamic programming and takes time of  $O(\delta(n+m))$  when using a matching window of  $\delta$ , and another step with complexity  $O(n)$  afterwards to obtain the warping path  $w$  between A and B. The recursive definition is top-down and starts the computation on both complete trajectories lessening them in each step. The dynamic programming approach works bottom-up filling up a  $n$  by  $m$  matrix row or column wise, starting at the upper left, where in each step  $(i,j)$  the  $DTW = (A_i, B_j)$  is computed based on the minimum of  $DTW = (A_{i-1}, B_{j-1})$ ,  $DTW = (A_{i-1}, B_j)$  and  $DTW = (A_i, B_{j-1})$  by adding  $d(a_i, b_j)$ . The final result is the summed up distance between all matched points. At the end every cell  $(i,j)$  of the matrix contains the result of  $DTW(A_i, B_j)$ . The matchings can be obtained by backtracking, starting at the lower-right of the matrix. In each of the backtracking steps the points  $a_i$  and  $b_j$  are stored as correspondent and one proceeds either with field  $(a_{i-1}, b_{j-1})$ , field  $(a_{i-1}, b_j)$  or field  $(a_i, b_{j-1})$  whichever is smaller. These matchings can also be seen as warping-path or warping-function synchronizing both trajectories with each other. In figure 5(a) two trajectories are shown in x and y dimension over time, where the dashed lines show the mapping. In figure 5(b) trajectory B is synchronized to trajectory A and figure 5(c) show the warping-path derived from the mappings.

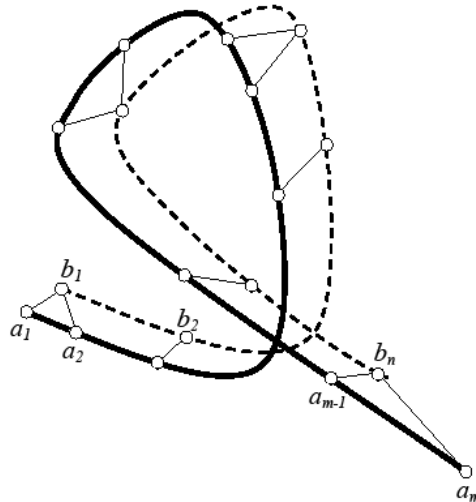


Figure 4. DTW matches all point while minimizing the summed up distances

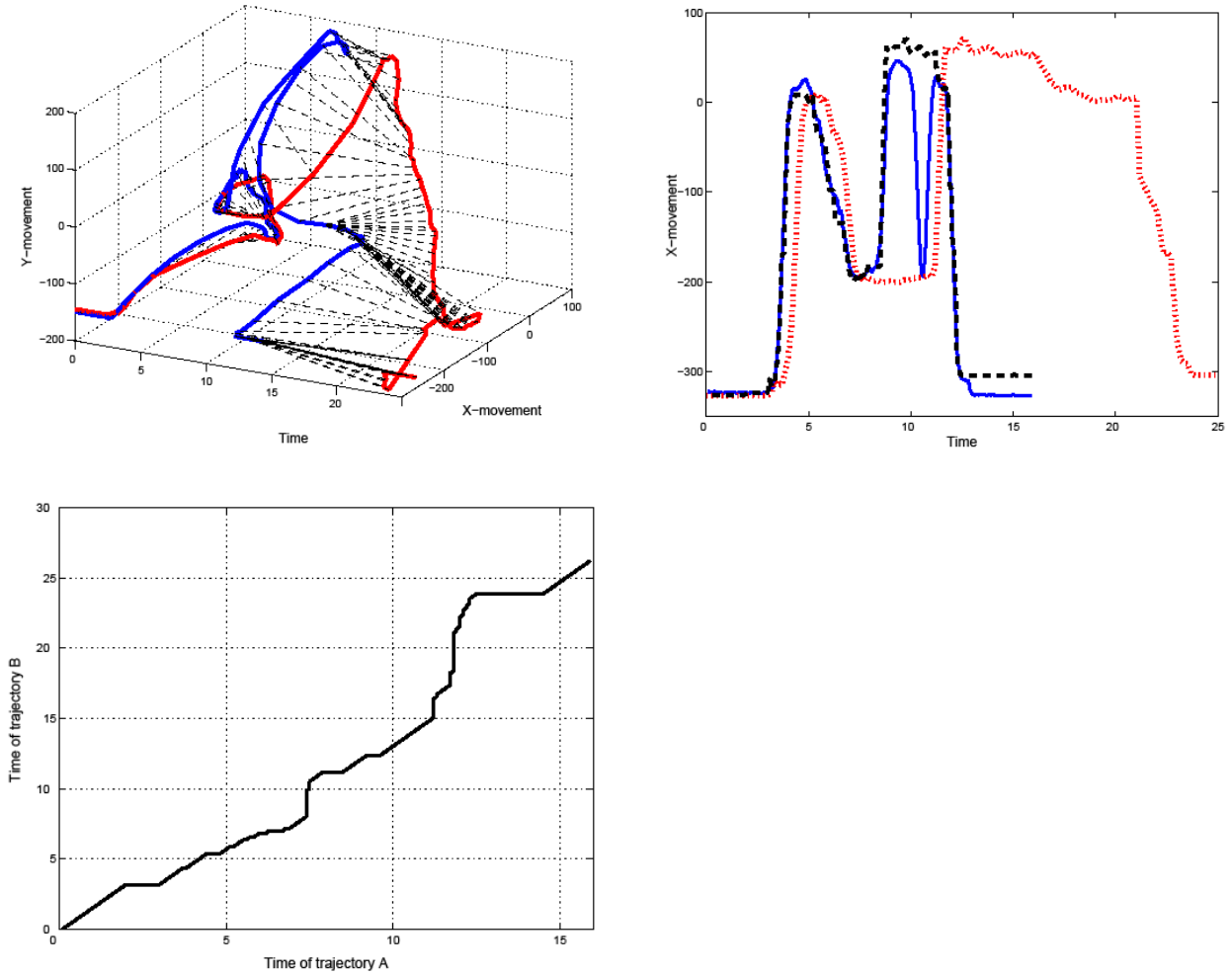


Figure 5. (a) shows the mapping computed with DTW, (b) shows both trajectories and one synchronized to the other, (c) shows a warping path that synchronizes B to A

To understand the power of the DTW appropriately, one has to take a closer look at the distance function. For the algorithm it does not matter on which data it is working, as long as a distance function is defined. So one could use the DTW also on multidimensional data including not only position and orientation of one object but also multiple objects, translational and angular velocity, force and other data, as long as a meaningful distance measure is defined on this multidimensional data. In our case, we only use the Euclidean distance. First attempts to also use orientation and translational velocity have not shown significant differences but should be examined more exhaustively in the future. As tracking errors resulting in a huge distance measure for some points could strongly influence the synchronization, we used a robust similarity measure limiting the maximum distance value.

#### 4. Feedback based on comparison to the expert performance

Being able to synchronize different performances to each other we introduce different methods for giving feedback. We aim mainly at providing visual and qualitative, dynamic feedback, which can be interpreted by a trainee without requiring presence of an expert. And as already mentioned we want to avoid authoring but rather use techniques which account for all types of possible mistakes. We implemented and investigated several ways to use the synchronized trajectories in order to give elaborate feedback, which are presented in detail in the following subsections. These include the simultaneous replay of both movements as well as a presentation of further statistical information seamlessly integrated into the AR visualization interface. Note, that since in all cases we visualize the result using AR, all advantages of the omnidirectional viewing are preserved. In addition the proposed strategy is applicable not only to 3D motion but also to force or any combination of position, orientation and force feedback.

In order to prove the concept we have built a demonstrator. It consists of our AR system and a tracked medical tool that has to be led through a metal wire. We simulate the behavior of experts and novices by the following: For experts' actions we keep the metal wire visible to the user (see figure 6). For trainees' actions we cover or remove the wire. Therefore we have a defined expert action that can vary a little, but that is not obvious to the trainee.

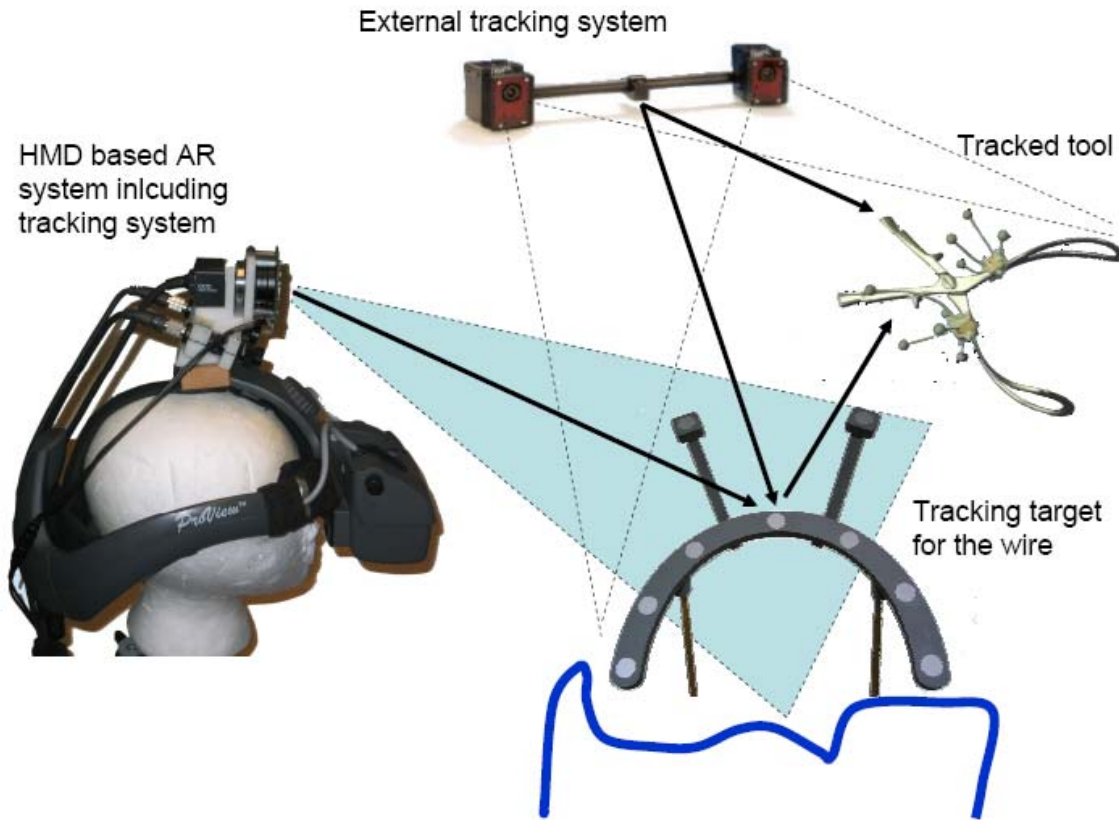


Figure 6. Setup of our system. The wire is in place of the physical simulator here only. We want to keep the setup for explanation as simple as possible

#### 4.1 Synchronized replay of both performances

In order to give the trainee the chance of comparing his performance with the expert's one we can show the recorded actions in a synchronized way. As already shown synchronizing is important for a direct visual comparison. Both movements can be replayed at the speed of the expert or the trainee or a slowed down version of each. The replay can also be stopped or rewinded in order to study the situation in detail. The problem is not as simple as one person being always faster or slower than the other one. For each person the speed of manipulation varies dynamically during the performance. One performance can be faster than the expert's in the beginning but slower in the end. Showing both movements synchronized to one speed of performance reveals differences in space, but note that the temporal difference is not lost. The warping function which we derive from the DTW contains information about the temporal difference. We present this temporal differences using progress bars (see figure 7). The left image shows just a simple replay of two movements and so the two forceps are not visualized at the same place. The right images shows the synchronized replay and therefore also the bars show a different progress.

By this means we visualize differences for cases where time is critical. Using the speedometer, which can already be seen on the pictures, we plan to extend this so that not only the overall progress but also the difference of the momentary speed is shown.

At first view one could think that this way to provide feedback would correspond with the definition of feedback we use. Both the actual and the expert's performances are shown to the trainee is now able to visually discover spatial and differences. Using the speedometer temporal differences can also be presented. But this is not accurate because the expert's performance we are showing is only one recording of his/her



performance. It might be afflicted by small, unwanted deviations. The next section introduces a solution to this problem.

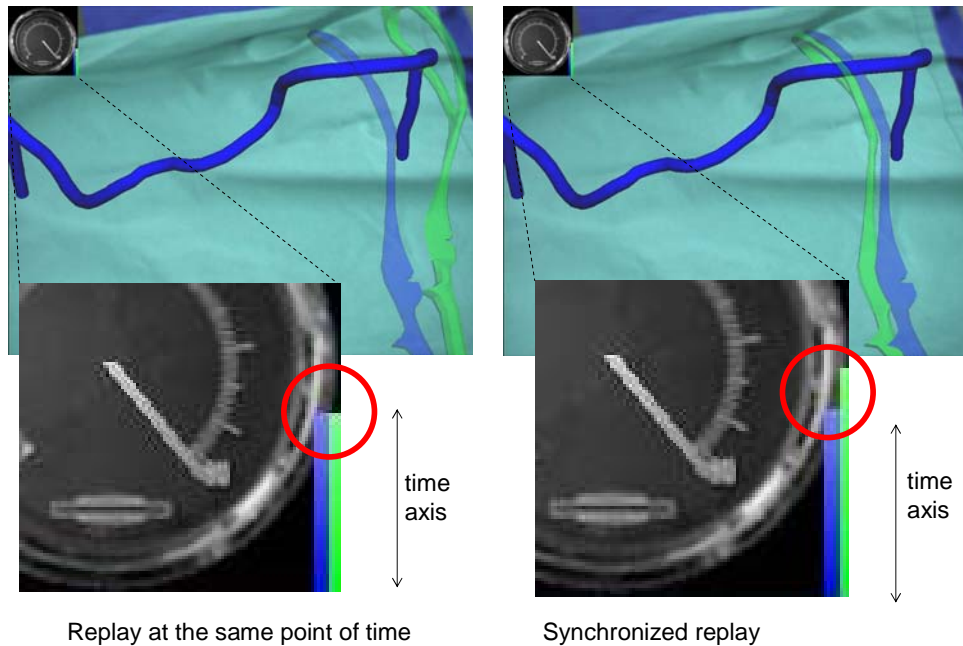


Figure 7. Comparing two actions: Synchronization interface for different performances

#### 4.2 Using many expert's actions

Theoretically, we would need an objective perfect or at least representative performance of an expert in order to compare it to the trainee's action. In practice one would subjectively rate several performances of one or even many experts in order to find the best one. But a subjectively method is not satisfying, especially as even the expert may not be able to assess his own performance, which is true for motor skills [2].

Our approach is to average several expert's actions or even experts' actions. Using a large number of performances we would get us closer to a objective comparison, whereas the problem arises how to present such a comparison of one to multiple movements to trainee. Again, the synchronized data proves to be useful since a simple average without synchronization does not provide the desired result (see figure 8).

We use the method Gasser and Wang [19] proposed for warping a larger number of different trajectories to a common time base in order to do further statistical analysis. After recording the movement of an expert doing the same motions several times, we apply this method to our data. By computing the mean movement, we can average out noise and little differences. Although for synchronization only Euclidean distance has been considered, the structural average does also include the orientation so that we get a real average of the whole task, which can now be used as a reference.

An example of this can be seen in figure 9. We recorded the forceps ten times while following the wire. The average trajectory we computed is shown in the upper right.

So instead of showing a synchronized replay of the trainee's and one expert's performance we can now use the average of performance of multiple experts. We are therefore able to give feedback by visualizing the differences between desired and actual outcome. We can also get further benefits from averaging. This is described in the next section.

#### 4.3 Using variance for feedback

The methods described above allow us to visualize the difference between desired and actual outcome. The trainees now can identify where they have deviations from the desired outcome, but they do not know if a deviation is critical or not. We use statistical analysis to provide such feedback.

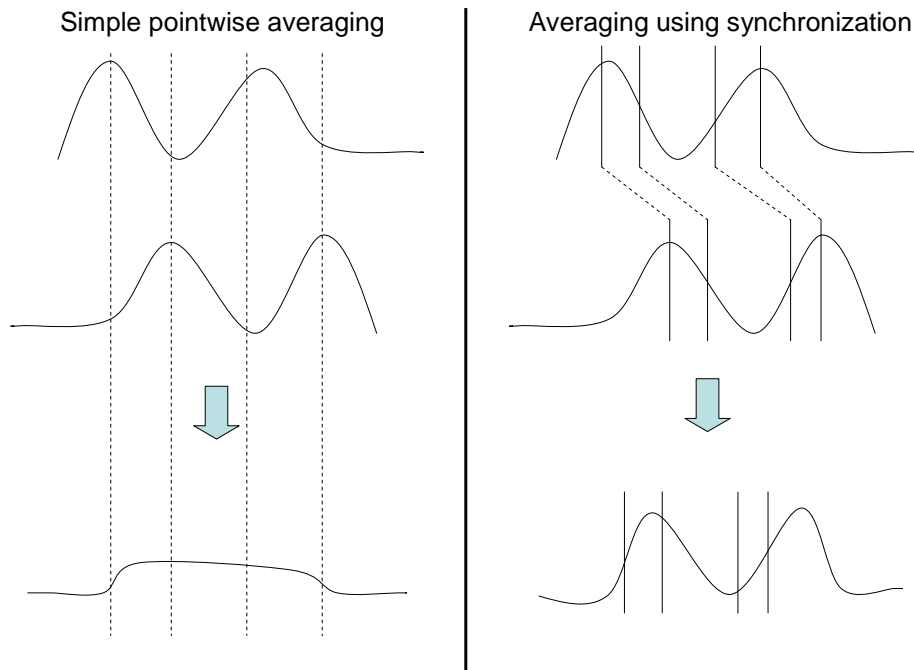


Figure 8. Averaging of curves can be also done after synchronization

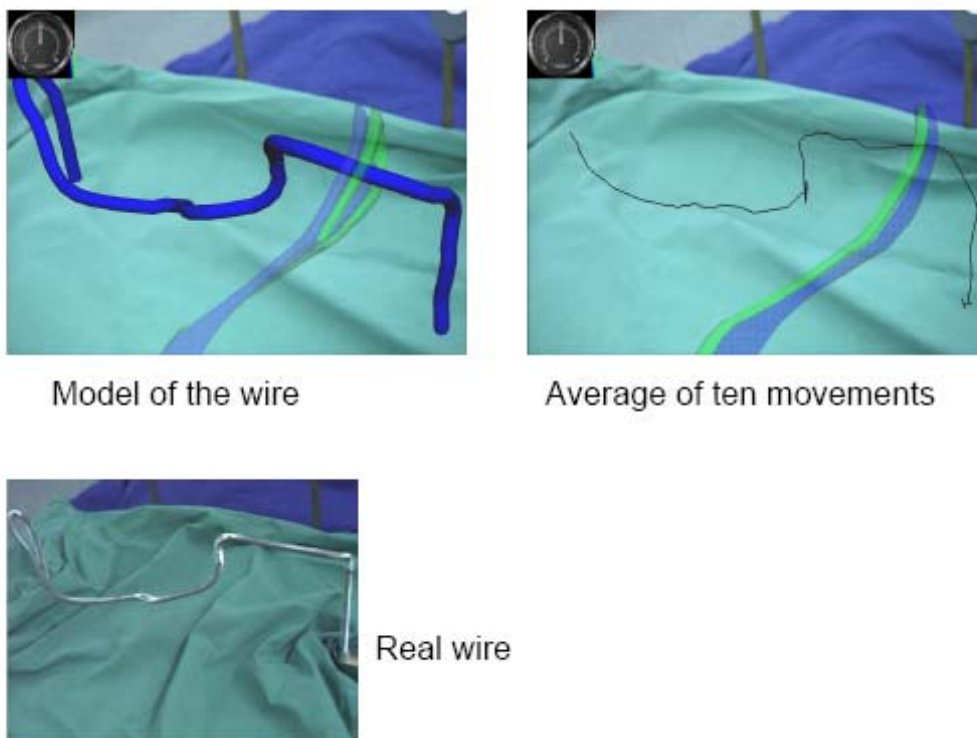


Figure 9. Visualization of an average movement computed by averaging ten exemplary movements

We derive from this synchronized set of performances the variance. In figure 10 we visualize it with ellipsoid. We assume that experts performing the same action several times using the same method will likely have a low variance at the critical parts while less important parts of the action are performed with a higher variance. This assumption is based on the idea that critical parts get more attention by the experts than non critical ones. For statistical comparison of different experts this assumption makes even more sense since critical parts must be the same by definition and non critical parts are likely to change because of personal preferences. This is certainly not true when there are different ways of accomplishing the same task.

But as we intend to teach only one method at a time, we can restrict the system to learn and teach one method.

The covariance of movements in each point is valuable information. It can be used to return a meaningful value of difference of two points via the Mahalanobis distance, which takes the multidimensional covariance into account. It means that if the trainee is off the track but the variation of the experts is high it does not matter.

As this direct visualization of the covariance is not very intuitive we are using the Mahalanobis distance to offer a color coding of the tool indicating whether the difference is severe or not (figure 11). Note that this enhancement enables the system to automatically indicate the quality of performance at each point of the trajectory and visualize it. This feature might also be used for assessment, but this not the scope of this chapter. The emphasis here is on real-time visualization of such feedback from any viewpoint in space. The trainee's attention can be drawn by this means of color coding of Mahalanobis distance to the average trajectory to improve his/her performance where crucial deviations occur. This can be indicated e.g. using colors by tinting the tools in red for critical or in green for non-critical deviations.

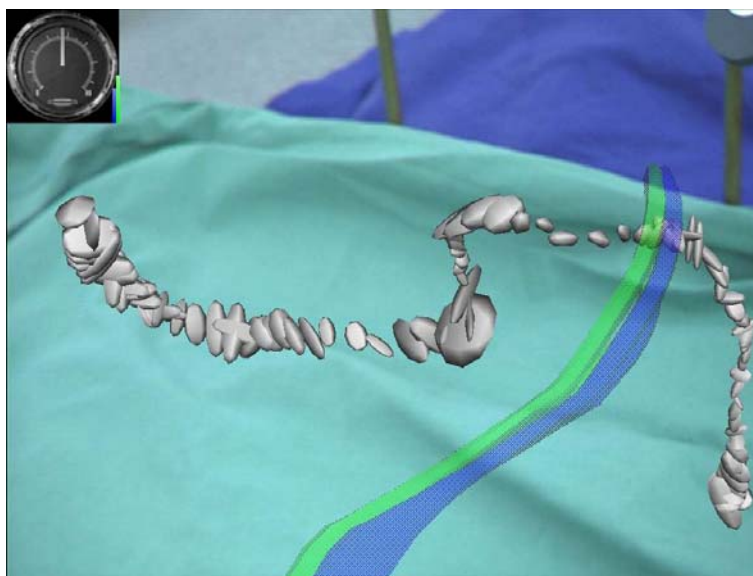


Figure 10. Ellipse visualization of covariances at each point

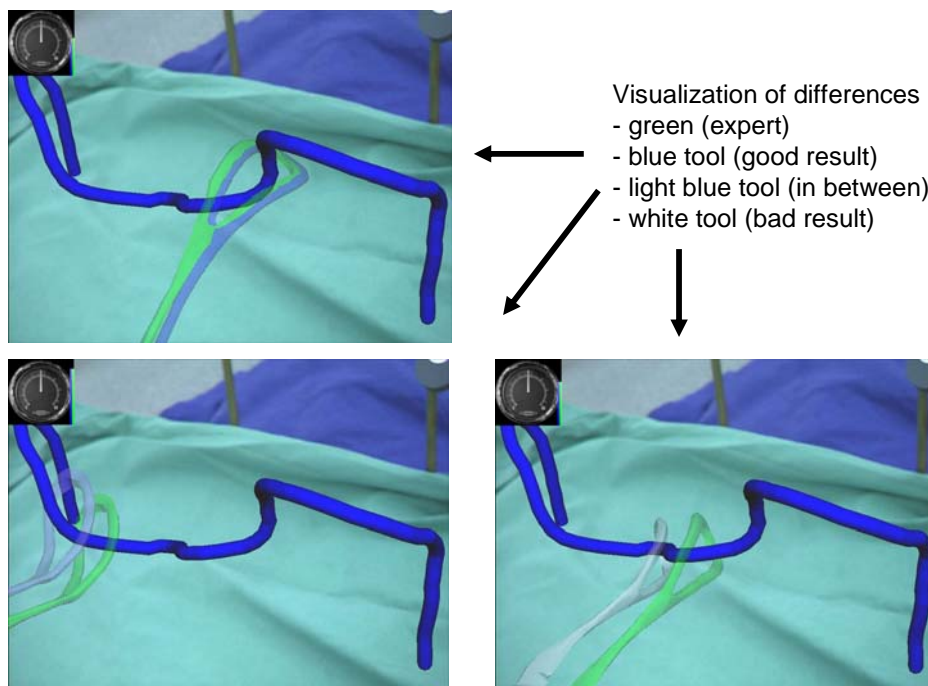


Figure 11. Giving direct feedback via color coding

#### 4.4 Online synchronization

Another interesting issue is online synchronization. The proposed synchronization method was considered up to now as an offline process. This means that the trainee performs a full action and afterwards the synchronization is done. Therefore a trainee will get only feedback after the performance. The reason is that the DTW does not allow for mapping between the two trajectories until the whole matrix is filled up during the computation, and the backtracking is done. One possible method for online synchronization is to simply suppose that the trainee will follow the exact same trajectory as the expert with the same speed. While the real data of the trainee's manipulation is acquired, the system modifies the predicted path. This allows gradual synchronization, while filling in the missing data through simple prediction. In first trials we managed to get such an online synchronization. As expected this is not as exact as the offline synchronization using DTW.

Using such an online synchronization, one can give feedback during the performance. This can also be used to trigger certain actions during the performance. We associated an action with one point in the expert's trajectory. When the trainee reaches this point, which is determined using the online synchronization, we triggered this action. So instructions, additional information or relevant data can be automatically shown at the right time or other scripted effects can be triggered. This could help the student in associating events that he/she needs to pay attention to during the 3D manipulation. Furthermore, in this way the student could also be reminded of actions to perform at different moments.

### 5. Discussion and future work

This chapter focuses on the presentation of new ideas for providing feedback to students and trainees using an AR system. The system relies on tracking experts and students, while manipulating tools, e.g. surgical instruments or forceps for delivery assistance. As soon as the expert's and student's actions are recorded, the system automatically synchronizes them and provides different feedback to the student. The system has been demonstrated to the scientific community at IEEE International Symposium on Mixed and Augmented Reality (ISMAR 05) and more than fifty persons have tested the DTW synchronization between their action and the reference one. The introduction of the use of DTW for synchronization of expert's and trainee's action has proved efficient and successful.

In this work we only made use of spatial data, i.e. tracked movements, for the synchronization. A tool has usually six degrees of freedom. The DTW can handle any number of dimensions since it uses a distance function that melts down differences of states to one distance value. The choice of such distance function is crucial for correct synchronization of different actions. In a sophisticated simulator, we usually have access to more data than the position and orientation of tools. For example the delivery simulator incorporates haptic sensors that provide key information on success of expert's and student's performance. We plan to integrate such information and define new similarity measures for advanced synchronization and providing more feedback to students. The system is currently being integrated into the human delivery simulator [14] at hospital rechts der Isar in Munich, Germany.

### References

- [1]. E.H. Mory. *Feedback Research Revisited*. Handbook of Research on Educational Communications and Technology. 2004; 745-783.
- [2]. J. Annett. *Motor Skills*. Learning and Skills. 1995; 56-75.
- [3]. A. Liu, C. Kaufmann, T. Ritchie. *A computer-based simulator for diagnostic peritoneal lavage*. Medicine meets virtual reality. 2001; 279-285.
- [4]. J.B. Biggs, P.J. Moore. *The process of learning*. 1993.
- [5]. R. Eyal, F. Tendick. *Spatial Ability and Learning the Use of an Angled Laparoscope in a Virtual Environment*. Medicine meets virtual reality. 2001; 146-152.
- [6]. J. Nizard, M. Duyme, Y. Ville. *Teaching ultrasound-guided invasive procedures in fetal medicine: learning curves with and without an electronic guidance system*. Ultrasound in Obstetrics and Gynecology. 2002; 19(3): 274-277.
- [7]. A. Dosis, R. Aggarwal, F. Belloa, K. Moorthy, Y. Munz, D. Gillies, A. Darzi. *Synchronized video and motion analysis for the assessment of procedures in the operating theater*. Archives of surgery. 2005; 140(3): 293-299.
- [8]. J. Torkington, S.G.T. Smith, I. Rees, A. Darzi. *Skill transfer from virtual reality to a real laparoscopic task*. Surgical Endoscopy. 2001; 15(10): 1076-1079.

- [9]. H.C. Lin, I. Shafran, T.E. Murphy, A.M. Okamura, D.D. Yuh, G.D. Hager. *Automatic Detection and Segmentation of Robot-Assisted Surgical Motions*. Medical Image Computing and Computer-Assisted Intervention. 2005; 802-810.
- [10]. J. Rosen, L. Chang, J.D. Brown, B. Hannaford, M. Sinanan, R. Satava. *Minimally Invasive Surgery Task Decomposition*. Medicine meets virtual reality. 2003; 295-301.
- [11]. F. Sauer, A. Khamene, S. Vogt. *An Augmented Reality Navigation System with a Single-Camera Tracker: System Design and Needle Biopsy Phantom Trial*. Medical Image Computing and Computer-Assisted Intervention. 2002; 116-124.
- [12]. S.Vogt, A. Khamene, F. Sauer, H. Niemann. *Single Camera Tracking of Marker Clusters: Multiparameter Cluster Optimization and Experimental Verification*. IEEE and ACM International Symposium on Mixed and Augmented Reality. 2002; 127-136.
- [13]. W.A. Hoff, T.L. Vincent. *Analysis of Head Pose Accuracy in Augmented Reality*. IEEE Trans. Visualization and Computer Graphics. 2000.
- [14]. T. Obst, R. Burgkart, E. Ruckhäberle, R. Riener. *Geburtensimulator mit multimodaler Interaktion - Delivery Simulator with Multimodal Interaction*. Automatisierungstechnik. 2000; 54: 280-287.
- [15]. T. Sielhorst, T. Obst, R. Burgkart, R. Riener, N. Navab. *An Augmented Reality Delivery Simulator for Medical Training*. International Workshop on Augmented Environments for Medical Imaging - MICCAI Satellite Workshop. 2004.
- [16]. T. Gasser, A. Kneip, A. Binding, A. Prader, L. Molinari. *The dynamics of linear growth in distance, velocity and acceleration*. Annals of Human Biology. 1991; 18(3): 187-205.
- [17]. D. Gunopulos. *Discovering Similar Multidimensional Trajectories*. Proceedings of the 18th International Conference on Data Engineering. 2002.
- [18]. H. Sakoe, S. Chiba. *Dynamic programming algorithm optimization for spoken word recognition*. IEEE Trans. Acoust. Speech Signal Process. 1978; 43-49.
- [19]. K. Wang, T. Gasser. *Synchronizing sample curves nonparametrically*. Annals of Statistics. 1999; 27(2): 439-460.
- [20]. M.E. Munich, P. Perona. *Camera-Based ID Verification by Signature Tracking*. Proceedings of the 5th European Conference on Computer Vision. 1998; 782-796.