ABSTRACT
The present work focuses on perceptual control of haptic manipulation during high frequency interaction with mobile objects.

In particular in this work we focused on the analysis of the control and perceptual issues in the throwing and catching in juggling. A training multimodal system that exploits the concepts of co-located visuo-haptic feedback and encountered interfaces was implemented. Using such a system the user juggles with a number of virtual balls that are in contact with him from catch to throw. The control design of the system has been supported by a psychometric validation of the catch contact.

Index Terms: D.2.6 [Interactive environments]; F.1.2 [Interactive and reactive computation]; H.5.2 [User Interfaces]: Haptic I/O;

1 INTRODUCTION
Haptic interfaces have the great capability of enhancing the interaction in virtual environments introducing the sense of touch, providing feedback during the interaction with virtual objects. While most applications separate the real world interaction with the virtual one, it is possible to effectively put together the visual channel and the haptic one. The problem of co-location gives indeed the possibility of introducing augmented haptic reality [2]. At the same time it is interesting to highlight the role of the haptic interface during the interaction. In standard haptic interfaces the contact with the interface is constant, and it is suitable when the interaction metaphor is based on a tool. A full range of interfaces, called encounter interfaces [12, 17], display contact with the user only when there is a virtual contact to be simulated. In both cases the interaction has a slow dynamic, mostly related to applications in the area of manipulation or exploration.

Juggling is a scientific challenging example of task that needs haptic interfaces with high dynamics and mixed reality technologies. New high performance interaction paradigms must then be studied. This work is placed indeed in the general context of a juggling training system based on haptics. In particular there are two important element of this task that are addressed here, catching and throwing. The goal of this work is to present control and perceptual aspects in catching and throwing of virtual balls using a co-located encountered interface.

These aspects are being discussed through the introduction of a system in which the user interacts with virtual balls that are haptically simulated by real balls attached to an haptic interface. The result is an augmented reality setup in which the user touches real balls and sees their virtual equivalent by means of co-located projection.

2 RELATED WORK
The generation of the contact feedback is an open issue in haptics, in particular when the aim is the perception of a stable contact. In the case of interfaces in which the user holds a tool a possible approach is the event-based haptics [5] that generates an open loop feedback. Another approach is the one pursued by encountered type haptic interfaces, separately discovered by [17] and [12], and later improved in several direction, with video feedback [20], multiple finger feedback [21], grasping [4] or surface properties feedback [15].

When a skill like juggling is being analyzed it is discovered how throwing and catching are fundamental tasks [13], and at the same time it is important for the person to perceive the contact with a real ball. In other haptic applications for gaming like table tennis the interaction is tool mediated [9] and the haptic interface can be tool based. The first claim of this work is indeed the adoption of a dynamic haptic encountered interface for simulating the catch and throw of the balls. For understanding the effectiveness of the catch phase we have conducted a psychometric measurement of catch’s quality both for design and evaluation purposes.

Juggling training is one of the application in which the separation of the haptic space and the visual space is not good, because in that case the user would learn the task with an offset, as it is typical in mixed reality trainers based on video only [11]. For this reason
it is important to provide a co-located display of the two channels [16, 18] that is typical of haptic augmented reality setups [2]. This paper follows in this way, a novel approach of co-located haptics for juggling training, grounded by the multiple object strategy and dynamic encountered interface.

3 ARCHITECTURE

In this section the architecture of the proposed juggling training system is presented. In particular the overall objective of the juggling is introduced and all the parts composing the Mixed Reality system presented. A brief introduction to the proposed interaction paradigm is also discussed.

3.1 Haptic Interface

The robotic device used for the interaction is the GRAB Haptic interface [7], a large workspace 3-DOF haptic interface that has been previously used for single or multiple fingers interaction. In particular the device has a box workspace of 400 mm depth, 400 mm high and 600 mm wide. This haptic device is able to generate continuous forces at the end-effector of 4 N in the worst condition while the peak forces are up to 20 N, the device has a position accuracy at zero load less than 1 percent, i.e. 1mm over 100 mm. In its standard configuration the GRAB device has a thimble as end-effector allowing the user to directly interact with virtual objects. In this system the thimble has been replaced by a standard tennis ball (3.5 cm of radius). The size and the shape of the contact element have been selected for providing a good interaction with the whole hand, being satisfactory for the training application. The system was composed by two GRAB robotic arms (reaching a workspace 1200 mm wide) that where placed inside an L-shaped projection environment composed of two large projection screens: one frontal, the other as a walkable floor. The environments was also equipped with a VICON Mx (OMG,UK) infrared camera system, working at 300hz, used as head tracking system to grant correct stereo projection to the user immersed in the virtual reality simulation. In particular such system was composed of 7 cameras that track position of retro-reflective markers. The markers were attached to a pair of stereo glasses so that was possible to track the user head’s movement. Such glasses had INFITEC (Ulm, DE) filters to decouple stereo images that are sent to the frontal and bottom screen of L-system by four projector. The projectors themselves have special INFITEC filters mounted on their lens. This projection system was necessary to grant the co-location between the virtual environment and the user/haptic interface, resulting in a complete mixed reality application.

3.2 Graphics

All the graphics and rendering where generated with the eXtreme Virtual Reality(XVR) framework [3] that provides basic facilities for graphics and spatialized audio. In order to simplify the management of the L-shaped environment the virtual juggling application uses a module of XVR, called XVR Network Renderer [10]. Exploiting the cluster rendering technique, the Network Renderer synchronizes the visualization of the whole environment with the developed XVR application, that can run on a separated machine, linked to the system via Ethernet.

The L-shaped environment is an immersive scenario composed of two big screens (2x2.70 meters wide each), one frontal and the other, walkable, placed under the user. The resolution of each screen is 1440x1050 with a refresh rate of nearly 60Hz.

The overall system (figure 2) is composed by five computer, one dedicated to haptic control, one for the tracking system and three for the stereo projection and graphical raterization being one machine used as master graphical unit and the other two as slaves that take from the master the generated graphical packets and display the correct images trough four projectors.

Figure 2: Architecture of the juggling trainer

3.3 Interaction Paradigm

Interaction logic is quite simple. Two robot controlled tennis balls are grasped and launched by human hands. Such balls, during launch, directly control the motion in an immersive and co-located virtual environment of two analogous digital objects.

There are three main phases during the juggling of one ball: grasp, “manipulation-launch”, and fly. The cinematic constraints introduced in this scenario by the physical presence of the robots (workspace and interference) have been solved with three changes in the interaction paradigm:

1. The robot collocation has been organized such that each robot covers a proper amount of each hand workspace.

2. During flight, position of robot attached balls and virtual balls may differ.

3. At grasping the robot which is closer to the grasping hand serves the perceptual representation.

3.3.1 Juggling

In juggling the user’s hands move along more or less elliptical trajectories. The balls are released at the inside of the ellipses and caught at the outside. Timing and sequencing of balls trajectories impose a consistent rhythm, so that frequency and phase locking between the hands must be stably sustained to comply with the task. Location, direction and velocity of toses serve as an anchor for the global coordination, and the balls trajectories get smoother with expertise.

A juggler has to accommodate rather severe task constraints in order to juggle successfully. Time constraints and sequencing are formulated by the Shannon theorem of juggling [14]:

\[(F + D)H = (V + D)N\]

where \(F\) is the time that a ball spends in air, \(D\) the time that it spends in a hand, \(V\) the time a hand is vacant, \(N\) represents the number of balls and \(H\) the number of hands.

3.3.2 Dynamic encountering

A given amount of complexities arise when passing from static interaction to dynamic one. In detail the interaction with mobile objects requires additional issues to the haptic devices adopted.

The correct representation of Mass/Inertia properties is fundamental to improve transparency condition of the environment during manipulation phases. The control of reflected inertia is possible
by means of a closed loop feedback on the motor torque which can partially compensate for or enhance the real inertia of the haptic device.

On the other side, as we will see, tracking of velocities of both haptic device and objects are of high importance for a clean rendering in the instant of contact of the user hand with the virtual objects. At such instant the controller should dissipate onto the human hand the same kinetic energy which exists in the virtual moving object. A particular attention should be given to energy transfer during impact when the high frequency of the velocity change may affect the real perception of masses. The “Mechanical bandwidth” of these systems is well below the typical frequencies expressed during impact. It is then required that compensation of the impact energy should be performed by an explicit pre-warping of the haptic velocity to maintain the consistence of the perception. Section 4 will describe in more detail how this pre-warping should be estimated.

Workspace itself it is an added constraint. Both virtual object, haptic devices and the user should share the same workspace and during contact/impact the location of them should be coherent. The analysis of the phase plan during the juggling phases allow us to formulate an alternate policy of control for the haptic interface that better matches the application requirements while keeping into account the workspace constraints.

Dynamic collision prediction is also crucial for the correct control in an encountered haptics based Virtual Environment. It is necessary not only to represent in real time physical phenomena that can be observed from the simulation but also to predict the physical condition that will be shown to the user.

4 Perceptual Evaluation

To study the perception of users during dynamic contact with a ball simulated trough an encountered haptic interface, we carried out several psychometric tests. To describe the relation between the physical intensity of a stimulus and the correspondent intensity perceived by an observer a psychometric function is usually generated fitting a sigmoidal function to the experimental data. In a typical psychometric experiment with forced-choice design stimuli are repeated in a number of discrete trials, in random order, being one of them the target stimulus. To analyze different stimulus conditions, thresholds (that specify the psychometric function’s location along the stimulus axis) are usually compared [6].

The first test performed was intended to investigate perception threshold of users with real tennis balls. In particular we adopted a 2 alternative forced choice(2-AFC) design choosing the reference stimulus to be a ball falling down from an height of 15 cm to the palm of observer’s hand. We repeated the experiment to evaluate repetability of the results, and performed another session setting the new reference height to 20 cm. The algorithm used to perform the psychometric analysis is the QUEST proposed by Watson and Pelli[19]. The QUEST is a Bayesian adaptive method to evaluate thresholds of psychometric functions. In particular it uses a probability density function representing the initial guess about the location of the threshold. Bayes theorem is used to update the prior after each response and to choose an optimal next stimulus to be presented to the observer. Result of our first test session showed that observer could recognize the presented reference stimulus with an error of nearly the 15%-25% (this error is about the energies not the velocities). Once known the observers perceptions of real balls weights we could perform the intended perception experiment. Our scope was to establish the difference of perception between real weight of objects and simulated weights of objects happening during dynamic contacts in the specific case of braking a falling ball. So we decided to use a 2-AFC design where one stimulus was real and one given by an haptic device with a real ball mounted as end-effector. In particular we proposed a two interleaved QUEST method during a single trial session. In one QUEST the reference

![Figure 3: Stimulus presented in two sessions of the interleaved test](image3)

stimulus was given by the ball falling from 15 cm of height with respect to the observer hand and the comparison stimulus was given by the haptic device seeking the suggestion of the QUEST procedure. In the other QUEST instead we used the haptic stimulus as the reference one and the real ball stimulus as the comparison one. When the stimulus was performed by the haptic interface we changed the impact velocity in consequence of the relative height suggested by the bayesian method. When the comparison stimulus was given by the real ball we simply changed the height from which the balls had to fall into the observer's hand. From this test we could retrieve important results. As previously suggested by Avizzano at al.[1] the impact perception at regrasp is much stronger than in real cases. The reason why is that robot control do not compensate for the mechanical inertia of linkages at impact. In particular the corrective factor on velocity is a function of mass ratio and velocity change.

\[ k = \sqrt{\frac{M_{obj}}{M_{hi}} \left(1 - \frac{M_{obj}}{M_{hi}} \right)\left(\frac{v_{post}}{v_{obj}}\right)^2} \]

where \(M_{hi}\) is the haptic device equivalent mass expressed in the Cartesian space accordingly to [8]; the modulus takes into account impacts between objects moving in the same direction; and \(v_{post}\) can be determined by analysing the impact properties (momentum conservation). In the particular case of ball juggling \(v_{post} \approx 0\) so that \(k = \sqrt{\frac{M_{obj}}{M_{hi}}}\). From the calculus the ratio between the ball mass and the haptic mass reported at the end-effector was about 4. The experiment showed exactly that the perception of impact force of the haptic device was 4 time bigger than the equivalent real impact force (figure 3). After this result we adopted the \(k\) factor correction in the control algorithm and repeated the tests. As shown in figure 4 the new percentage error is again in the range of the one obtained by the first test session about the observer perception error with real balls. The test were repeated to validate the results as shown in the figure.

5 Interacting with Multiple Objects

This section describes the control strategy necessary for simulating the interaction with multiple virtual balls and how this approach can be extended to more balls and more complex configurations.

5.1 Control structure

In the trainer specific setup we have two virtual objects (2 balls) and two haptic devices that should perform the dynamic encounter
with the user caused by the interaction with the flying balls. To proper feedback the correct sensations to the user we needed to find a solution to the spatial and time scheduling of the devices. In the case of a “two versus two” setup we resolved this issue as it will be now introduced. In general, during the simulation we may have two spatial trajectories obtained by applying simple physical motion laws to the virtual balls and two devices that can keep track of such trajectories if they are free of movement. The device are free if they are not in contact with the user or in other words if the user has already launched the real ball attached to the device’s end-effector.

In this preliminary setup we decided not to track the hands of the users to concentrate on dynamical properties of the simulator. The result is to introduce a reference threshold to distinguish between a launched ball or a ball still in the hand of the user. In the future this constraint will be removed and replaced by the distance with the user hand computed by finer calibration between user’s hand and device’s end-effector.

For what concerns the spatial scheduling, the selection of the encounter device for a ball is the one with the minimal distance to the final position of each ball. In the virtual space, X axis of the simulation is partitioned along the robot absolute reference system x axis (“Grab Abs” in figure 6). An hemispace is served by the device on the right of the user and the other one served by the left device. To choose instead which ball to serve if two balls are coming on the same X partition of the space they are scheduled based on the time of flight or exactly by the time it takes to reach the hand position that in this case corresponds to the threshold one. The resulting schedule is a shortest time, first served. In this manner it is possible to grant all the falling balls to be served if the balls are launched with the typical temporal schedule of juggling. In particular two state-flow charts are responsible for the trajectory planning of the balls. In these charts final positions as well as flight time and velocities and positions profiles are calculated based on the initial launch velocity and position read by the encoders and reported at the end effector. All this information is sent to the spatial scheduling part, and from there is forwarded to the blocks that perform the time analysis choosing the trajectory to be served at every time instants.

The presented schema is then responsible to generate the correct reference trajectories for the graphics and for the control referred to an absolute reference system with the origin in the center of the workspace. These references are forwarded to the control loop of each GRAB arm. Being the local reference system different from the absolute one, each control properly manage the coordinates.

In particular for each arm both a force control algorithm as well as a position and velocity control loops have been developed. If we look at the schematic chart of the control (Figure 5) we can see that each device can be in one of five different states.

The initial state is the Manipulate one. In this state the user is free to move the corresponding interface with his hands. For perceptual reason in this state is applied only a force control to feedback to the user the correct weight of a tennis ball as if the ball was not attached to an haptic interface at all.

The condition to pass to the next logical state is the detaching of the ball from the user hand that means the throw of the ball. For our particular setup this is translated in the end-effector exceeding the vertical threshold as we supposed that every launch phase will finish exactly or at least under the vertical threshold. In this way the control moves to the FlyUp state. In this state the device moves with constant velocity along the y-axis to reach a starting vertical position chosen not to interfere with user hands and arms movements. This state last only 0.05 seconds. The third state is a Wait state where the interface stops its movements waiting a reference trajectory to seek. As soon as a reference exists the device goes in Rendezvous state. In this state a reference for the position control on X and Z axis is given so that the device could re-attach to the virtual ball at the impact with the user hand. The control remain in this state until a proper amount of time is passed. This check on the time is needed to have a timing consistence between the virtual ball and the device. When it is time to fall to reach the user hand at the final time, calculated by the trajectory control discussed above, the interface pass in the FlyDown state. In this state is applied to the control a reference constant velocity calculated using the $K$ factor introduced in the perceptive test. In this manner the effective impact is perceived by the user as the one of a real ball. This state last with the grasping by the user of the ball and the interface goes again in the Manipulate state. Again in the presented simplified simulation this condition is not given by the effective grasp but by the interface reaching the vertical threshold.

For co-location purpose the head of the user has been tracked and the stereo projection adapted to the correct perspective of the user. The virtual ball lives in the 3d space and its movements are regulated by the trajectory produced by the control algorithm. In particular during the launch and grasp phase of the juggling it is necessary to keep the exact co-location of the virtual ball, the device’s end-effector and the user’s hand.

To perform this exact co-location we need to apply transformations between several coordinate reference systems and to calibrate them. The coordinate systems involved are shown in Figure 6.

The coordinates of each device’s end-effector are reported to the GRAB absolute reference systems. The head tracking is trans-
formed from the VICON system to the XVR (graphics) one. Finally the GRAB absolute system and the graphical one are made coincident with a coordinate transformation.

5.2 Multiple Objects and Multiple devices

When the number of virtual objects to be simulated or the number of devices present in the workspace increases the control schema needs to be adjusted taking into account more complex issues. The growth of the number of virtual objects sets a timing constraint for the proper functioning of the system requiring fast response devices. It also introduces some conceptual difficulties. For example, allowing in the trainer 3 balls to be juggled and having only two device to represent them, introduces an incoherence during some time intervals. In particular this occurs at the very beginning of the simulation when the user should have two balls on one hand but the interface is only capable of making him feel one single ball. Apart from this starting situation, if the juggle is performed correctly, the user could still train himself and learn a correct juggling movement because only two real balls are in contact with user’s hands at the same time. In this case we must provide a memory for taking track, for each ball, of the correspondent hand where it lays.

Adding instead other devices on the same workspace we could fix the coherence problem at the cost of additional complexity of the spatial partition, trajectory profiles and obstacle avoidance strategy. However mathematics comes in help in solving this multi constraints problem. From the previously stated Shannon theorem is possible to calculate exactly the timing and the needs of a multiple objects, multiple devices setup.

Indeed the introduction of a multi ball dynamic system comprised of several devices will be an argument that will be discussed as a future evolution of this juggling trainer.

6 Evaluation of the system

The key factor in juggling is the tendency of two limbs to move at the same frequency in sync. The particular type of coordination displayed by juggling hands depends on the juggling pattern. Figure 7 show the most common ones. In the cascade pattern the effectively crossing of the balls between the hands demands that one hand catches at the same rate that the other hand throws.

The fountain pattern, in contrast, can be stably performed in two ways: by throwing and catching the balls simultaneously with both hands or by throwing a ball with one hand and catching one with the other at the same time so out of sync.

What is important is that the developed virtual haptic juggling trainer allows users to perform all this kind of patterns. Even if the presented patterns are specifically thought for at least three balls the capability of the system to perform these patterns is essential for a proper training and the proposed system appear to be easily extensible.

In figure 8 is possible to see the fountain pattern performed by a non expert juggler (only few minutes of juggling). The plot itself shows how each device serves only the virtual trajectories belonging to its workspace spatial partition. As it is visible, the haptic interfaces end-effector vertical positions remain always under 0.1 meters while the corresponding virtual ball trajectory can reach really higher heights without the restriction of the devices workspace.

Figure 9 instead, reports the positions of the virtual balls and devices’ end-effectors during a juggling trial, performed by a non expert user, seeking the cascade pattern. From the plot is visible how a ball thrown with an haptic interface is instead caught with the other.

The system is quite versatile and can be used with ease to learn the juggling skill in the subset of two balls juggling. The system use, as learning facility, a different gravitational acceleration that the Earth’s one so that is possible to slow down the entire juggling process and concentrate in timing and trajectory planning, increasing again the acceleration when the user becomes able enough.

Having at disposal the head tracking system is also possible to use such informations in the assessment phase. In particular the system can log the head orientation and check if the user is looking...
his hands, the devices end-effectors or the virtual balls in the air. This is important to understand what’s best to look at during a juggling exercise.

7 Conclusions and Future Work
This paper has approached the problem of juggling training from the point of view of two fundamental issues, throwing and catching. These two aspects have been addressed using an encountered haptic interface that allows to keep the basic interaction with the balls. In this work we have covered both the control aspect and the perceptual aspect both for design and validation of the feedback scheme.

There are two relevant aspects to be addressed in future work. The first is the evaluation of the effectiveness of this system for training purposes, eventually comparing it with another solution without haptic feedback. The second aspect is related to the evaluation and involves the management of more than two virtual balls as discussed above.

Acknowledgements
This work has been carried out within the context of the EU Integrated Project SKILLS, that deals with the use of MultiModal system for the capture model and rendering of data signals for the transfer of skills in different application fields.

The authors would like to acknowledge the European Commission for sponsoring this research inside the project SKILLS-IP. More information about the IP-SKILLS project may be found at http://www.skills-ip.eu or by contacting the authors.

References