In Proceedings of Intelligent Virtual Agents (IVA), Boston, 2014
(This version is the authors' manuscript, the final publication is available at link.springer.com)

Planning Motions for Virtual Demonstrators

Yazhou Huang and Marcelo Kallmann

University of California, Merced

Abstract. In order to deliver information effectively, virtual human demonstrators must be able to address complex spatial constraints and at the same time replicate motion coordination patterns observed in human-human interactions. We introduce in this paper a whole-body motion planning and synthesis framework that coordinates locomotion, body positioning, action execution and gaze behavior for generic demonstration tasks among obstacles.

Keywords: virtual trainers, motion planning, intelligent virtual humans.

1 Introduction

Virtual humans and embodied conversational agents are promising in the realm of human-computer interaction applications. One central goal in the area is to achieve virtual assistants that can effectively interact, train, and assist people in a wide variety of tasks. The need to demonstrate objects and procedures appears in many situations; however, the underlying motion synthesis problem is complex and has not been specifically addressed before. Simple everyday demonstrations involve a series of coordinated steps that a virtual agent needs to replicate. The agent needs to walk while avoiding obstacles along the way, stop at an appropriate demonstration location with clear view to the target and observer, interact with the object (e.g. point to it and deliver information), and also maintain visual engagement with the observer. This work addresses such harmonious multi-level orchestration of actions and behaviors (see Figure 1).

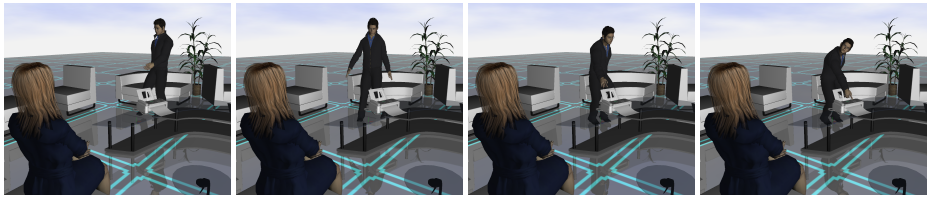

Fig. 1. Our PLACE planner synthesizes whole-body demonstrations for arbitrary targets, observers, obstacles, and visual occluders.

The proposed model was built from experiments with human subjects where participants were asked to freely approach target objects at different positions and to deliver object information to observers at various locations. These experiments provided

ground truth data for defining a coordination model that is able to orchestrate the involved pieces of a demonstration task. The result is a whole-body motion planning framework, called PLACE, that addresses the five main pieces of the problem in an unified way:
• Placement: optimal character placement is essential for addressing target and observer visibility, locomotion accessibility, and action execution constraints;
• Locomotion: locomotion synthesis among obstacles and towards precise placements allows the character to position itself in order to perform a demonstration;
• Action: realistic action execution needs to address arbitrary object locations and to avoid nearby obstacles when needed;
• Coordination: coordination is important for well transitioning from locomotion to the upper-body demonstrative action; and
• Engagement: observer engagement is obtained with a gaze behavior that interleaves attention to the observer and the target in order to achieve effective demonstrations.

The realism of the solutions is addressed at two levels. At the behavioral level, placement, coordination and engagement are solved following models extracted from experiments with human subjects. At the motion synthesis level, locomotion and actions are synthesized from collections of motion capture clips organized for efficient synthesis and coordination. The techniques were developed such that solutions can be computed at interactive rates in realistic environments. See Figure 2 for examples.

Contributions: The main contribution of this work is the overall definition, modeling and effective solution of whole-body demonstrative tasks. The proposed techniques are the first to address the overall problem in an integrated fashion.

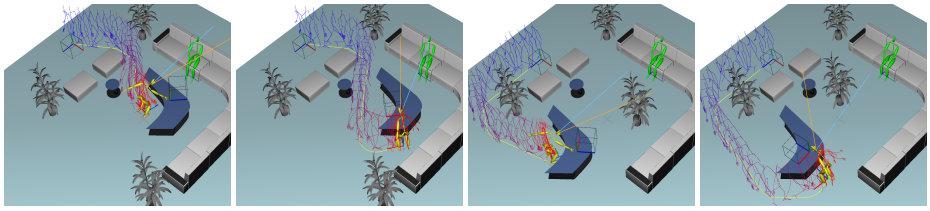

Fig. 2. From left to right: in the first two scenarios the computed demonstrations reasonably face the observer, while in the last two cases a visual occluder (the house plant) leads to solutions with non-trivial placements. The orange and blue lines respectively represent the head and the eye gaze orientations, at the demonstration action stroke point. The resulting gaze always reaches eye contact with the observer.

2 Related Work

Many works have focused on upper-body gesture and action modeling, including stroke-based blending [25], action synthesis with varied spatial constraints [17, 24], motion style control [6, 19], and search in interpolated motion graphs [20]. Our approach for action synthesis relies on available motion interpolation techniques [9,19] but providing

a new collision avoidance mechanism in blending space in order to successfully address realistic scenarios with obstacles.

With respect to data-based locomotion methods, several techniques have been proposed for achieving realistic and controllable locomotion synthesis [7,14–16,26]. While several of these methods can probably be extended to address departures and arrivals with position and orientation constraints, such an extension is not trivial. Our solution is based on a specific organization of locomotion clips that allows for fast locomotion synthesis ensuring such constraints. The locomotion planning problem becomes even more challenging when it has to be coordinated with an upper-body action. Previous work [5, 22] has addressed the combination of arm planning (reaching or grasping) on top of locomotion, however without a coordination model.

When it comes to whole-body motion synthesis that involves the scheduling and synchronization of upper- and lower-body, methods have been developed for splicing upper-body actions from one motion to another [2, 8], and more recently object manipulations have been coordinated with locomotion [1]. However these methods have not addressed a coordination model for transitioning from locomotion into a demonstration action, a specific situation that involves different types of constraints.

The fact that demonstrations have to be performed with respect to an observer also distinguishes our overall motion synthesis problem from previous work. Addressing an observer is important for achieving realistic solutions and as well effective interactions with virtual humans. For instance, it has been shown that visual engagement improves the amount of information memorized by an audience observing robotic storytellers [18] and narrative virtual agents [3]. Although previous work has focused on modeling gaze behavior in great detail [4, 27], little attention has been given to integration with full-body motion synthesis. In computer animation simple solutions have been employed [25,28] based on pre-defined points of interest, however not associating with a complete set of events observed from human subjects during action execution and locomotion. With respect to modeling placement for action execution, Scheflen and Ashcraft [21] present a pioneering work introducing the concept of *territoriality* in human-human interactions, but unfortunately without computational models.

In conclusion, the proposed PLACE planner addresses the problem of whole-body demonstration at multiple levels and uniquely integrates behavioral models from human subjects with realistic data-based motion synthesis.

3 Modeling Demonstrative Tasks

We have modeled the overall problem of synthesizing humanlike demonstrations with the help of experiments with human subjects. Our setup follows the approach in [10], but extending it for extracting complete motion models for demonstration tasks. Four human participants were recruited to perform a variety of pointing tasks with full-body motion capture. Six small target objects T_i, $i \in \{1, \ldots, 6\}$, were placed on a horizontal coarse mesh grid and participants were asked to perform demonstration actions towards each T_i for a human observer O_j standing at five different positions around the targets ($j \in \{1, \ldots, 5\}$). Each action consisted of pointing and delivering a short information about an object. Each configuration $\{T_i, O_j\}$ represented one trial per participant and

generated one motion. A total of 30 distinct motions were generated per participant, each motion consisting of a complete pointing action with the associated locomotion and gaze behavior. The gaze typically moved several times between T_i and O_j. Each participant started from about 4 feet away from the mesh grid before walking towards the grid to point and describe T_i (see Figure 3). The sequence of target selection was random and the targets were of similar size in order to reduce possible side effects related to their size [23].

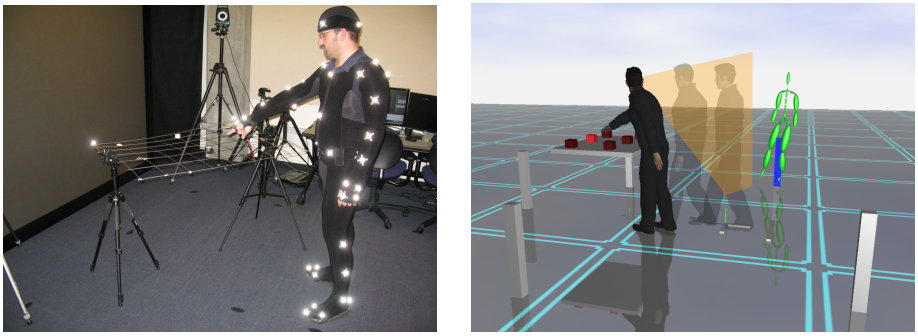

Fig. 3. Left: experiment setup. Right: illustration of one reconstructed motion. The observer location is represented with the green character and the maximum head orientation performed in the direction of the observer is shown with the orange plane.

The full-body capture data was annotated manually with an annotation tool specifically developed to mark and extract the parameters and timings of all relevant events in each trial. One of the most important behaviors observed was the chosen positioning that each participant used to perform the pointing action. The chosen position ensured that the target and the observer were visible, and that the head rotation needed for eye contact with the observer was feasible. The position also ensured a successful execution of the action and with a fluid transition from locomotion. We now derive a generic placement model based on the observed data.

For each trial in the collected motion data we extracted the corresponding target position $\mathbf{p_t}$, the position of the observer $\mathbf{p_o}$, and the demonstrator position $\mathbf{p_d}$. Position $\mathbf{p_d}$ is the position used to perform the demonstration action, and is defined as the position when the locomotion is detected to completely stop, since there is a period when the action execution overlaps with the locomotion. Figure 4 plots locomotion trajectories and their corresponding final demonstration positions. The 5 distinct colors represent the 5 different observer locations. Each color appears 6 times, one for each target T_i. It is possible to observe that the demonstration positions do not show an obvious structure in global coordinates.

A local 2D coordinate system with origin at $\mathbf{p_t}$ is then used to derive our model. The coordinate system is illustrated with the XZ frame in Figure 5. The XZ frame can have arbitrary orientation, however it is more intuitive when the Z axis is orthogonal to the table border closest to the target. We can now use angles to locally encode all

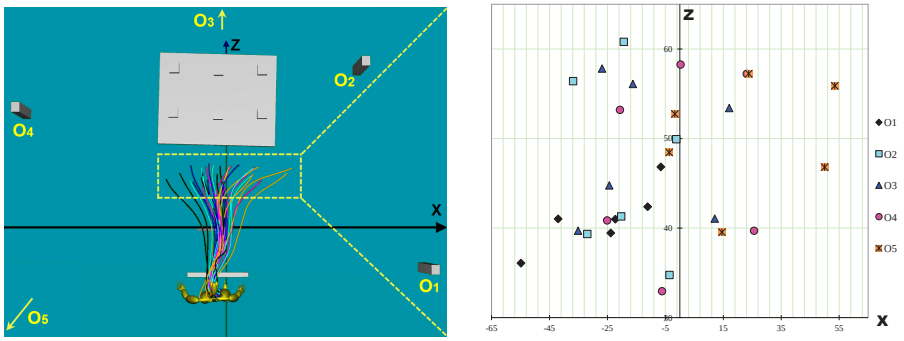

Fig. 4. Locomotion trajectories for each participant (left) and their ending positions in a closer look (right). Motions with the same color are with respect to a same observer.

relevant placement parameters with respect to the target. The used local angles will not model the proximity of the demonstrator to the target, since this is a parameter that is action-dependent and we leave it as a free parameter in our model. For example a pointing motion can be executed with arbitrary distance to the target while this is not the case in a manipulation task. The angles considered by our model are the following (see Figure 5): the observer position α with respect to Z, the demonstrator position β with respect to -Z, the demonstrator's body orientation θ with respect to Z, and the maximum head rotation ϕ (in respect to Z) performed towards the observer.

The approach of expressing placement locally with respect to the action target correlates with the *axis concept* used for describing interaction connections [21]. By expressing the collected parameters with respect to our local coordinate system, the plots of their values nicely fit into clusters with good structure. This is shown in Figure 6. Since the proposed placement model shows good structure, we then performed non-linear regressions in order to be able to estimate β, θ and ϕ as a function of an arbitrary input value for α. After smoothing the raw measurements with a least-squares Savitzky-Golay filter, quadratic and cubic polynomial functions were fitted for β, θ and ϕ (see Appendix for details).

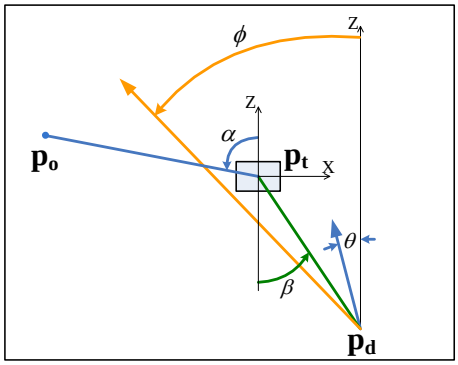

Fig. 5. Local coordinate system of the placement model. Angle α encodes the observer location, β the demonstrator location, θ the body orientation at action start, and ϕ encodes the maximum head rotation towards the observer.

The overall demonstration problem is then modeled as follows: given an upper-body demonstrative action A to be performed, the corresponding target object position $\mathbf{p_t}$, and the position of the observer $\mathbf{p_o}$, the goal of the PLACE planner is to synthe-

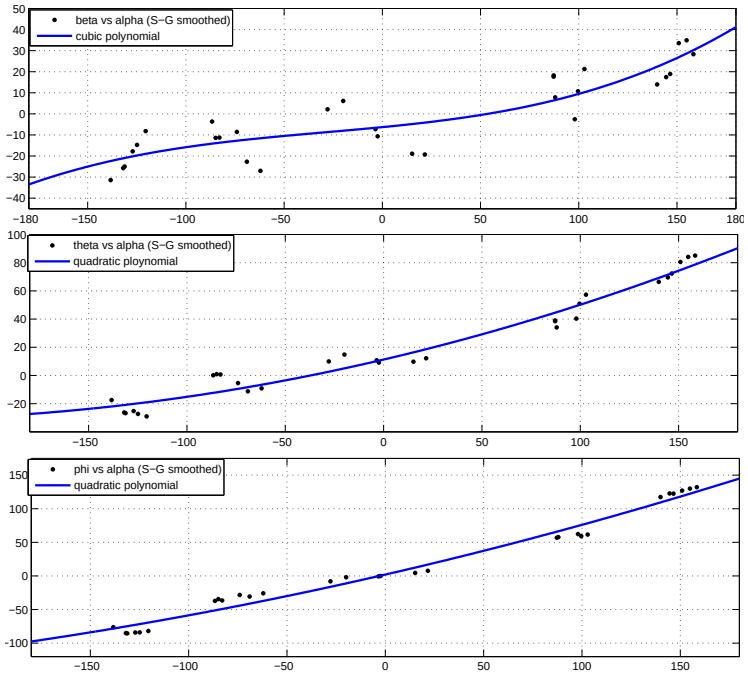

Fig. 6. Parameters of the placement model fitted with non-linear regression on filtered data points. The horizontal axis is the α value and the vertical axis, from top to bottom, represents β, θ, and ϕ respectively. Values are in degrees.

size a full-body motion for approaching the target and performing A with respect to $\mathbf{p_t}$ and for the observer located at $\mathbf{p_o}$. The planner solves the problem with the following steps. First, a suitable demonstration position $\mathbf{p_d}$ and body orientation $\mathbf{q_d}$ are determined by using the placement model and taking into account visual occluders, action feasibility and locomotion accessibility. Then, a locomotion sequence $L(\mathbf{p_d}, \mathbf{q_d})$ is synthesized for the character to walk from its current position to the demonstration placement $(\mathbf{p_d}, \mathbf{q_d})$. Action $A(\mathbf{p_t})$ is then synthesized and coordinated with the locomotion L. Finally, the head and eyes are animated to replicate the same gaze behavior patterns observed in the collected motions. These steps represent the five main components of PLACE and they are explained in the following sections.

4 Placement

Given the demonstrative action A, the target object position $\mathbf{p_t}$, and the observer position $\mathbf{p_o}$, the placement module will determine the optimal body position and orientation $(\mathbf{p_d}, \mathbf{q_d})$ for performing A. First, the action synthesis module (described in Section 6) is queried for its preferred distance d_{pref} to execute $A(\mathbf{p_t})$. This distance denotes the preferred Euclidean distance between $\mathbf{p_t}$ and $\mathbf{p_d}$ so that the character will be more likely to succeed in performing the action. The computation of d_{pref} is action-dependent, it may

be automatically selected according to reachability constraints if the demonstration has to achieve object contact, or in other cases (like in pointing), it can be a user-specified parameter fixed or dependent on features of the environment (like the size of the target).

The local reference frame of our placement model is then placed with origin at $\mathbf{p_t}$ and with the Z axis set to be orthogonal to the closest edge of the supporting table. At this point our local placement model can be applied in order to estimate a first candidate placement $(\mathbf{p_d^0}, \mathbf{q_d^0})$, where $\mathbf{p_d^0}$ is obtained by combining the estimated angle β with the preferred distance d_{pref}, and $\mathbf{q_d^0}$ represents the orientation estimated by θ in global coordinates. If $\mathbf{p_t}$ lies between $\mathbf{p_o}$ and $\mathbf{p_d^0}$, the Z axis of the local placement coordinate frame is re-oriented towards $\mathbf{p_o}$ and the initial placement $(\mathbf{p_d^0}, \mathbf{q_d^0})$ is re-computed. This will make the initial placement to directly face the observer, a desired property in a placement.

Given a candidate placement, the placement is considered valid if: it does not lead to collisions with the environment, if $A(\mathbf{p_o})$ can be successfully executed from it, and if there is a collision-free path with enough clearance for the character to reach it. If a candidate placement is not valid due a collision, it is tested again a few times with slightly perturbed values and d_{pref} distances, thus increasing the chances of finding valid placements by local adjustment of the generated positions. Several placements may be valid and therefore we will search for the optimal one with respect to visibility, head rotation comfort, and distance to the target.

Starting from $(\mathbf{p_d^0}, \mathbf{q_d^0})$ we determine several valid placements $(\mathbf{p_d^k}, \mathbf{q_d^k})$ by adjusting the Z axis of the local model to new orientations around $\mathbf{p_t}$, for example by rotation increments of five degrees in both directions (see Figure 7). For each new orientation, the respective estimated placement is computed and tested for validity (adjusted if needed) and stored if valid. The search for valid placements may be set to be exhaustive with respect to the used rotation increment or to stop after a certain number of valid samples is found. The result of this phase is a set of K valid placements $(\mathbf{p_d^k}, \mathbf{q_d^k})$, $k \in \{1, \ldots, K\}$. We then sort the placements in this set with respect to the following ranking cost function f_c:

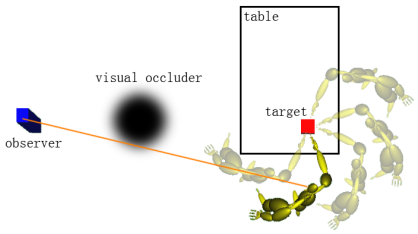

Fig. 7. Valid placements around the target are identified and ranked for selection.

$$f_c = e_{vis} * w_v + e_{neck} * w_n + e_{action} * w_a,$$

where e_{vis} is a measure of how occluded the observer is from the placement, e_{neck} is the amount of neck rotation required for reaching eye contact with the observer, e_{action} is the absolute difference between d_{pref} and the actual distance from $\mathbf{p_d^k}$ to $\mathbf{p_t}$, and the scalar weights are constants used to adjust the relative contributions of each term.

The weights are adjusted such that the contribution of e_{vis} is significant, since uncomfortable (but feasible) placements are preferable to placements with bad visibility. The house plant in the scenarios of Figure 2 is an example of an object modeled with

partial occlusion set to $e_{vis} = 50\%$. Candidate placements with zero visibility are discarded, as shown in Figure 7. The result is a sorted list of valid placements that can be used for executing A. The placement with minimum f_c is selected as the target demonstration location $(\mathbf{p_d}, \mathbf{q_d})$ to be used. Since this placement has been already checked for validity, it can be safely passed to the motion synthesis modules described in the next sections. In case simplified validity tests are employed, the motion synthesis may happen to not be successful at some later stage, in which case the next placement in the list can be used as the next alternative.

5 Locomotion Synthesis

The locomotion synthesis module has to be able to address three main requirements:
• to be able to quickly check for locomotion accessibility to candidate placements when queried by the body placement module (Section 4), in order to allow for quick rejection of placements that offer no accessibility;
• to be able to synthesize motions that can navigate through narrow passages and with precise departure and arrival positions and orientations; and
• to produce purposeful motions resembling the ones observed in our experiments with human subjects, which consistently had sharp turnings (with small turning radius) at the departure and arrival locomotion phases.

With respect to the first requirement, accessibility queries are computed with an efficient algorithm for computing paths [12, 13], which is used to determine path feasibility with clearance under a few milliseconds of computation in our scenarios. With respect to the second and third requirements, we adopt a path following approach in order to be able to safely navigate through narrow passages of the environment.

In order to address precise departure and arrival placements, and to achieve the observed behavior of sharp turns during path following, we propose an optimized locomotion synthesis method based on a specific organization of locomotion clips from motion capture. Three specific types of locomotion sequences were collected with a full-body motion capture system: departure motions, arrival motions, and walk cycles. Each motion is then parameterized to cover a small area around its original trajectory, and to depart or arrive with a parameterized orientation at the first or final frame. As a result we obtain parameterized motions that cover all possible required initial and final orientations, and that concatenate into walk cycles during the path following phase. Additional details on the locomotion module are omitted due lack of space, but will be made available from the website of the authors. Figure 8 illustrates one example.

6 Action Synthesis

The synthesis of the demonstrative action is performed with blending operations in a cluster of example motions. Given the target position $\mathbf{p_t}$ to be addressed by the end-effector at the stroke point of the action, blending weights are determined by inverse blending optimization [9] in order to address the target precisely. See Figure 9.

Collision avoidance has shown to be important. It not only increases the ability to find solutions in cluttered environments but it also improves the number of successful

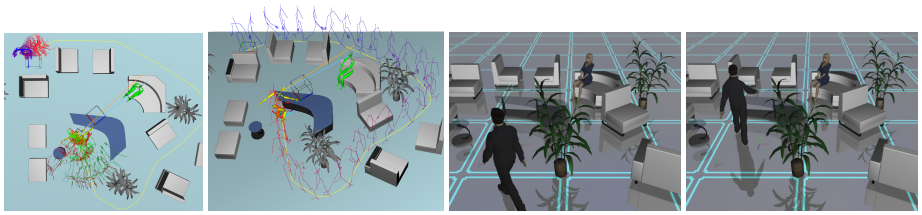

Fig. 8. Illustration of one particular locomotion sequence planned. From left to right: the departure and arrival clips nearby the path to be followed, skeleton trails illustrating the whole motion obtained, and two snapshots of the final result.

placements to be considered for action execution. We have developed a collision avoidance method that operates on the blending space of the example motions defining an action. Blending space operations have been employed before in motion planning [11], but here we develop a faster collision avoidance procedure that does not require expensive planning around the obstacles. We create repulsive force fields in 3D and compute a scalar potential of collision P_c that encodes the potential of collision between the agent's end-effector E and the obstacles. Instead of computing the force field in discretized 2D cells [29], we approximate the bounding volume of the nearby obstacles with small spheres S_i (see Figure 9-center), so that $P_c = exp(-\sum distance(E, S_i))$.

Let $\mathbf{p_t}$ be the target object to be addressed by action A. First, blending weights $\mathbf{w_t}$ that generate action $A(\mathbf{p_t})$ are computed by inverse blending. The produced motion can be re-written as a sequence of frames $F_i(\mathbf{w}_i)$, $i \in \{1, \ldots, n\}$, and initialized with $\mathbf{w}_i = \mathbf{w_t}$, $\forall i \in \{1, \ldots, n\}$. Next, we make a single pass from F_1 to F_n and adjust intermediate frames F_j at a given discretization resolution. The resolution is relative to the distance covered by the end-effector. Given a frame $F_j(\mathbf{w}_j)$ being visited, if its corresponding posture collides or is detected to be too close to an object, $\mathbf{w}_j$ is adjusted by inverse blending in order to minimize P_c at F_j, essentially shifting F_j away from the obstacles. Collisions are checked intermittently at mesh level and the process moves on to F_{j+1} when F_j becomes collision-free. Each time a frame is adjusted, the weights of the nearby frames (according to a smoothness window) are updated so that the overall sequence of weights is smooth, producing a new final motion $A(\mathbf{p_t})$ that smoothly avoids the nearby obstacles.

The method typically solves action synthesis under 300 milliseconds, with most computation time spent on mesh collision checking. The approach is able to control how much deformation is allowed, thus controlling the balance between action motion quality, action adaptation to obstacles, and body placement search time.

7 Locomotion-Action Coordination

The locomotion transition into the action requires special attention in order to generate realistic results. We start by using a transition window of 0.58 seconds that is the average window observed from our studies with human subjects. The window tells how early, before finishing the locomotion, the action should start to be executed. The ac-

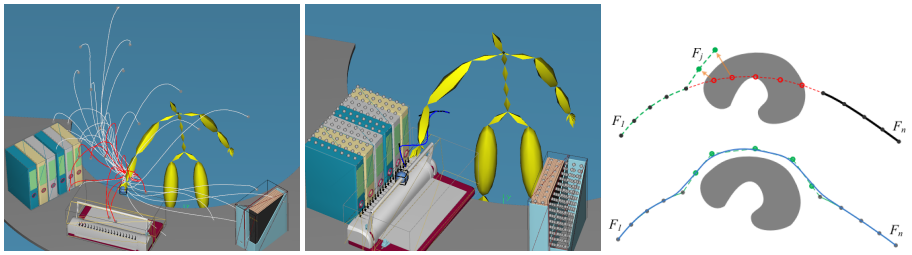

Fig. 9. Left: trajectories of one pointing database with 10 blended motions for one solution marked in red. Center: spheres are used to approximate the nearby objects and to compute P_c. Right: If intermediate frames collide their blending weights are adjusted to remove the collision.

tion will start gradually taking control over the upper-body and will achieve full control when the locomotion stops. An important coordination problem that we address here is to ensure that the resulting arm swing pattern during the transition remains realistic.

Let S_l^{end} be the arm swing at the end of the locomotion sequence, and S_a be the arm swing direction of the action. In our examples the right arm is used by the action but the presented analysis can be equally employed to both arms. Two main cases may happen: 1) the arm swings can be codirectional, in which case a natural transition is automatically achieved, or 2) the arm swings can be contradirectional, what would cause a sudden change in the arm swing direction during the transition window. Sudden changes in the arm swing were never observed in our experiments with human subjects, who were very good at achieving coherent final steps with clear final arm swings. We therefore fix contradirectional cases in two possible ways. If the final locomotion swing S_l^{end} slightly overlaps into the action arm swing S_a, it is then shortened to match S_a and without having it to return to its target rest position. This is accomplished by repeating the final frames of S_l^{end}, skipping the same amount of initial frames of S_a, then smoothly blending into the latter. If however the final locomotion swing shows a significant overlap, S_l^{end} is then dropped and the previous swing cycle S_l^{prev} is extended to override S_l^{end}, before blending into S_a. We examine the swing velocities generated from both and the one showing better consistency is applied. Figure 10 illustrates the process.

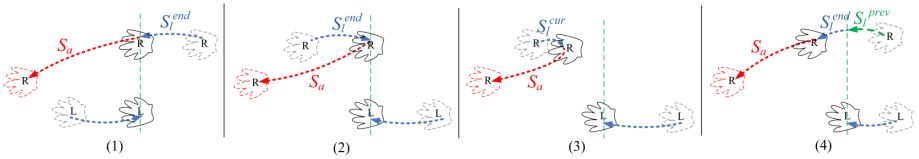

Fig. 10. Overlapping transition of the final arm swing of the locomotion S_l^{end} towards the arm swing direction generated by the action S_a. Codirectional cases can be directly blended (1), however contradirectional cases (2) have to be adjusted either by shortening the final locomotion swing (3) or by overriding it with the previous swing (4).

8 Engagement Module

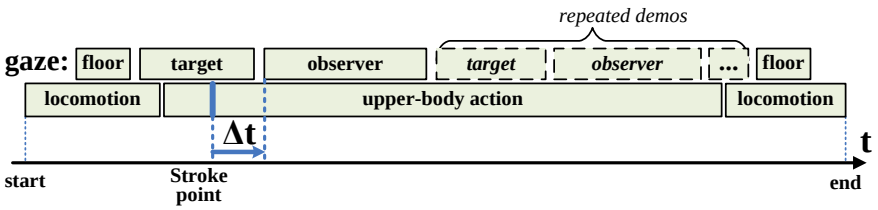

Fig. 11. Synchronization between gaze and motion events.

The final step of PLACE includes a gaze model that follows the observed behavior in our experiments with human subjects. We observed that each demonstration trial consisted of a series of largely consistent gaze and gaze-related events where participants first gazed at the floor during the locomotion, and then gazed at the target and the observer during the upper-body action. Our gaze model generates gaze events that follow these observed patterns. We use a temporal delay Δt between the action stroke point and the start of the object gaze event that is correlated with the agent's head rotation angle ϕ. When the observer is not in the center of the agent's field of view, the gaze towards the observer starts before the action reaches the stroke point, resulting in a negative Δt. The gaze behavior also incorporates gaze durations that decline over time across repeated demonstrations, an observed behavior in our experiments with human subjects. Figure 11 illustrates the main events of the gaze model.

9 Results and Discussion

Results are presented in Figures 1, 2, 12, 13 and 14. Additional results are presented in the accompanying video to this paper, which is available at `http://graphics.ucmerced.edu`. The planner is capable of synthesizing entire sequences in a range from 100 to 400 milliseconds, depending on the complexity of the environment and the collision avoidance settings.

Our results demonstrate that body placements are always well chosen and lead to positions clearly well addressing all involved constraints. The coordination of the swing arm trajectories has also shown to always produce good results. In terms of limitations, our planner leaves out facial expressions and other behaviors that are specific to the context of the scenario being simulated. Our model was also only designed to handle demonstrations for a single observer, although we believe that multiple observers can be easily incorporated if behavioral data exploring relevant possible configurations is obtained. The collected motion capture actions used in our blending procedures are available from the following project website: `http://graphics.ucmerced.edu/software/invbld/`.

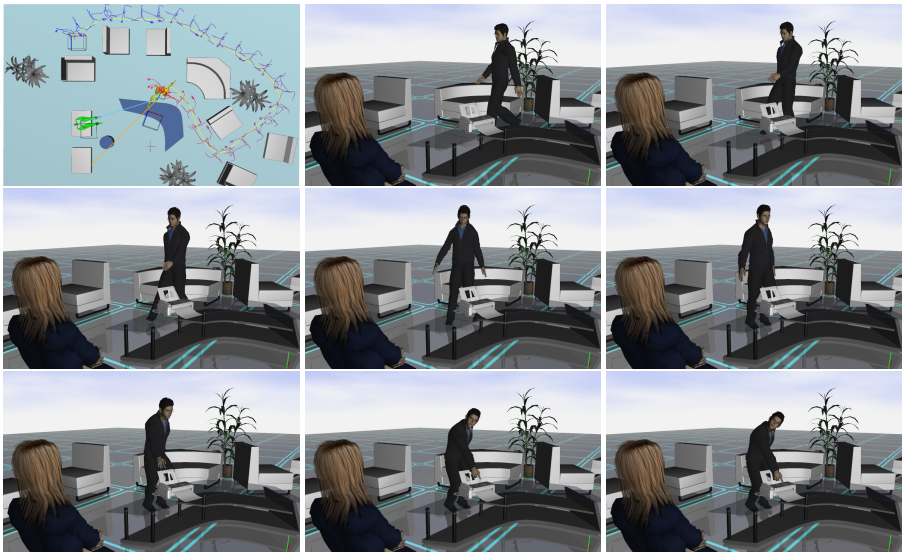

Fig. 12. Example of a solution produced by PLACE. The top-left image shows the planning scenario and the solution placement for execution of the demonstration. The following sequence of snapshots shows the arrival locomotion seamlessly transitioning into the demonstration action pointing at the fax machine with coordinated gaze towards the observer.

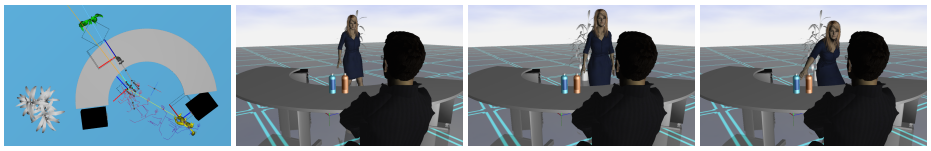

Fig. 13. Short-range solution suitable for pointing and describing the blue bottle.

10 Conclusion

We have introduced a new behavioral and motor planning model for solving demonstrative tasks. Our proposed PLACE planner uniquely explores body placement trade-offs involving visibility constraints, action feasibility, and locomotion accessibility. The proposed techniques can be computed at interactive rates and are suitable to several applications relying on interactive virtual humans as virtual trainers.

Acknowledgments This work was partially supported by NSF Award IIS-0915665.

Appendix

Polynomial (cubic and quadratic) functions were chosen over other types of fitting functions such as Gaussian and Fourier for better extrapolations when $\alpha > 150$ and

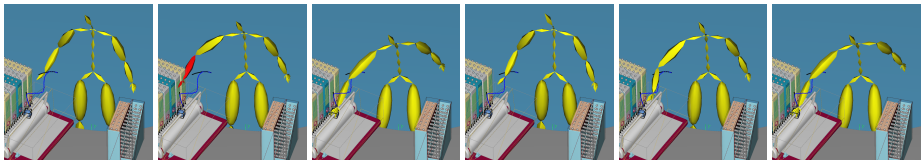

Fig. 14. Action synthesis corrected by force fields in blending space. The left three images show an action that originally produced collisions with obstacles; and the following three images show a collision avoidance solution (along the blue trajectory) for removing the collisions. Without the collision avoidance the action would not be feasible and a new placement would be necessary.

$\alpha <$ -150. Details are given below.

• $\beta = f(\alpha) = p_1\alpha^3 + p_2\alpha^2 + p_3\alpha + p_4$
Coefficients (with 95% confidence bounds):
$p_1 = 2.392\text{E}{-6}(-6.74\text{E}{-6}, 1.152\text{E}{-5})$,
$p_2 = 0.0003056(-0.0004444, 0.001056)$,
$p_3 = 0.1145(-0.04067, 0.2697)$,
$p_4 = -6.062(-15.42, 3.294)$.
Goodness of fit: $SSE = 5386$, $R^2 = 0.6156$,
$AdjustedR^2 = 0.5713$, $RMSE = 14.39$.

• $\phi = f(\alpha) = p_1\alpha^2 + p_2\alpha + p_3$
Coefficients (with 95% confidence bounds):
$p1 = 0.0006673(0.0001145, 0.00122)$,
$p2 = 0.6736(0.6315, 0.7158)$,
$p3 = 2.073(-5.167, 9.312)$.
Goodness of fit: SSE : 3381, R^2 : 0.9785,
$AdjustedR^2$: 0.9769, $RMSE$: 11.19.

• $\theta = f(\alpha) = p_1\alpha^2 + p_2\alpha + p_3$
Coefficients (with 95% confidence bounds):
$p_1 = 0.0006228(0.000262, 0.0009837)$,
$p_2 = 0.3267(0.2991, 0.3542)$,
$p_3 = 11.29(6.564, 16.02)$.
Goodness of fit: $SSE = 1441$, $R^2 = 0.9635$.
$AdjustedR^2 = 0.9608$, $RMSE = 7.304$.

References

1. Bai, Y., Siu, K., Liu, C.K.: Synthesis of concurrent object manipulation tasks. ACM Trans. Graph. 31(6), 156:1–156:9 (Nov 2012)
2. van Basten, B.J.H., Egges, A.: Flexible splicing of upper-body motion spaces on locomotion. Computer Graphics Forum 30(7), 1963–1971 (2011)
3. Bee, N., Wagner, J., André, E., Vogt, T., Charles, F., Pizzi, D., Cavazza, M.: Discovering eye gaze behavior during human-agent conversation in an interactive storytelling application. In: Int'l Conference on Multimodal Interfaces and Workshop on Machine Learning for Multimodal Interaction. pp. 9:1–9:8. ICMI-MLMI '10, ACM, New York, NY, USA (2010)
4. Cullen, K.E., Huterer, M., Braidwood, D.A., Sylvestre, P.A.: Time course of vestibuloocular reflex suppression during gaze shifts. Journal of Neurophysiology 92(6), 3408–3422 (2004)
5. Esteves, C., Arechavaleta, G., Pettré, J., Laumond, J.P.: Animation planning for virtual characters cooperation. ACM Trans. Graph. 25(2), 319–339 (Apr 2006)
6. Grochow, K., Martin, S., Hertzmann, A., Popović, Z.: Style-based inverse kinematics. ACM Transactions on Graphics (Proceedings of SIGGRAPH) 23(3), 522–531 (2004)
7. Heck, R., Gleicher, M.: Parametric motion graphs. In: Proceedings of the symposium on Interactive 3D Graphics and Games (I3D). pp. 129–136. ACM Press, New York, NY, USA (2007)

8. Heck, R., Kovar, L., Gleicher, M.: Splicing upper-body actions with locomotion. In: Proceedings of Eurographics 2006 (september 2006)
9. Huang, Y., Kallmann, M.: Motion parameterization with inverse blending. In: Proceedings of the Third International Conference on Motion In Games. Springer, Berlin (2010)
10. Huang, Y., Kallmann, M., Matthews, J.L., Matlock, T.: Modeling gaze behavior for virtual demonstrators. In: Proceedings of the 11th International Conference on Intelligent Virtual Agents (IVA) (2011)
11. Huang, Y., Mahmudi, M., Kallmann, M.: Planning humanlike actions in blending spaces. In: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (2011)
12. Kallmann, M.: Shortest paths with arbitrary clearance from navigation meshes. In: Proceedings of the Eurographics / SIGGRAPH Symposium on Computer Animation (SCA) (2010)
13. Kallmann, M.: Dynamic and robust local clearance triangulations. ACM Transactions on Graphics 33(5) (2014)
14. de Lasa, M., Mordatch, I., Hertzmann, A.: Feature-Based Locomotion Controllers. ACM Transactions on Graphics 29(3) (2010)
15. Lee, Y., Wampler, K., Bernstein, G., Popović, J., Popović, Z.: Motion fields for interactive character locomotion. ACM Trans. Graph. 29(6), 138:1–138:8 (Dec 2010)
16. Levine, S., Wang, J.M., Haraux, A., Popović, Z., Koltun, V.: Continuous character control with low-dimensional embeddings. ACM Trans. Graph. 31(4), 28:1–28:10 (Jul 2012)
17. Mukai, T., Kuriyama, S.: Geostatistical motion interpolation. In: ACM SIGGRAPH. pp. 1062–1070. ACM, New York, NY, USA (2005)
18. Mutlu, B., Hodgins, J.K., Forlizzi, J.: A storytelling robot: Modeling and evaluation of human-like gaze behavior. In: Proceedings of HUMANOIDS'06, 2006 IEEE-RAS International Conference on Humanoid Robots. IEEE (December 2006)
19. Rose, C., Bodenheimer, B., Cohen, M.F.: Verbs and adverbs: Multidimensional motion interpolation. IEEE Computer Graphics and Applications 18, 32–40 (1998)
20. Safonova, A., Hodgins, J.K.: Construction and optimal search of interpolated motion graphs. ACM Transactions on Graphics (SIGGRAPH 2007) 26(3) (Aug 2007)
21. Scheflen, A.E., Ashcraft, N.: Human Territories: How We Behave in Space-Time. Prentice-Hall, Englewood Cliffs, NJ, USA (1976)
22. Shapiro, A., Kallmann, M., Faloutsos, P.: Interactive motion correction and object manipulation. In: ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games (I3D). Seattle (2007)
23. Soukoreff, R.W., MacKenzie, I.S.: Towards a standard for pointing device evaluation: Perspectives on 27 years of fitts' law research in hci. International Journal of Human-Computer Studies 61, 751–789 (2004)
24. Sumner, R.W., Zwicker, M., Gotsman, C., Popović, J.: Mesh-based inverse kinematics. ACM Trans. Graph. 24(3), 488–495 (2005)
25. Thiebaux, M., Marshall, A., Marsella, S., Kallmann, M.: Smartbody: Behavior realization for embodied conversational agents. In: Seventh International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS) (2008)
26. Treuille, A., Lee, Y., Popović, Z.: Near-optimal character animation with continuous control. In: Proceedings of ACM SIGGRAPH. ACM Press (2007)
27. Van Horn, M.R., Sylvestre, P.A., Cullen, K.E.: The brain stem saccadic burst generator encodes gaze in three-dimensional space. J. of Neurophysiology 99(5), 2602–2616 (2008)
28. Yamane, K., Kuffner, J.J., Hodgins, J.K.: Synthesizing animations of human manipulation tasks. ACM Transactions on Graphics (Proceedings of SIGGRAPH) 23(3), 532–539 (2004)
29. Zhang, L., LaValle, S.M., Manocha, D.: Global vector field computation for feedback motion planning. In: Proceedings of the 2009 IEEE International Conference on Robotics and Automation. pp. 3065–3070. IEEE Press, Piscataway, NJ, USA (2009)