To appear in ACM TOG 31(4).

Eyecatch: Simulating Visuomotor Coordination for Object Interception

Sang Hoon Yeo Martin Lesmana Debanga R. Neog Dinesh K. Pai

Sensorimotor Systems Laboratory, University of British Columbia*

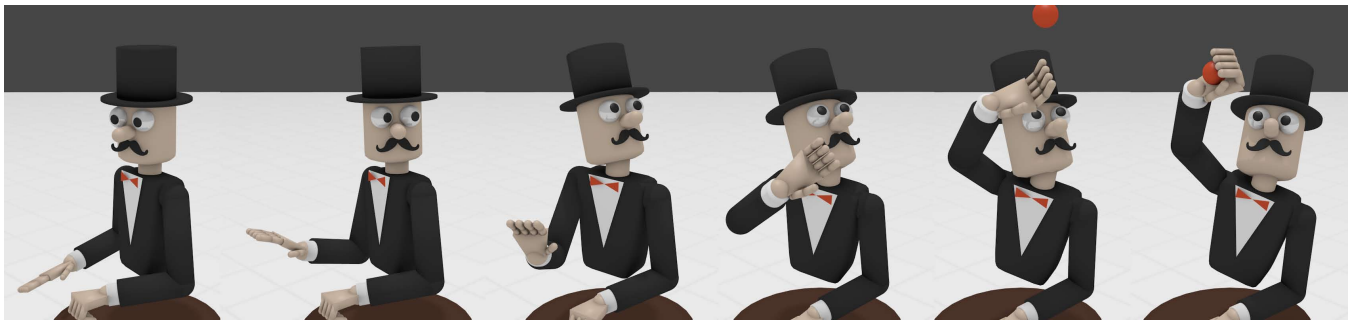


Figure 1: *Catching a thrown ball. The movement depends on visual estimates of the ball's motion, which trigger shared motor programs for eye, head, arm, and torso movement. The gaze sets the goal for the hand. Initially the movements are reactive, but as visual estimates improve predictive movements are generated to the final catching position.*

Abstract

We present a novel framework for animating human characters performing fast visually guided tasks, such as catching a ball. The main idea is to consider the coordinated dynamics of sensing and movement. Based on experimental evidence about such behaviors, we propose a generative model that constructs interception behavior online, using discrete submovements directed by uncertain visual estimates of target movement. An important aspect of this framework is that eye movements are included as well, and play a central role in coordinating movements of the head, hand, and body. We show that this framework efficiently generates plausible movements and generalizes well to novel scenarios.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Animation

Keywords: human animation, sensorimotor computation, eye movements, catching, interception, movement coordination

1 Introduction

Fast interception movements, such as catching thrown balls, are important in computer animation, especially of sports. At first glance these appear to be straightforward problems: the trajectory of the ball is a simple parabola, and once a point of interception is picked, the character's hand could be moved to the point of interception using a variety of animation techniques. Such an animation, however, will appear very unrealistic and robotic because it ignores several crucial features of real human movement. In this paper, we propose a generative model of such movements that takes into account these features and measurements of human catching behavior. By emphasizing the role of gaze as a prime coordinator of the motion, the model can efficiently generate realistic interceptive movements without requiring too many ad hoc assumptions.

How humans catch. To understand how normal, untrained people actually perform interception tasks, it is helpful to look at some data. We measured catching behavior of an untrained subject, simultaneously measuring the motion of the ball, hand, head, and body using a motion capture system, and eye movements using a head mounted eye tracker. See Fig. 2; details of the measurement system are given in Sec. 5.

First, the ball trajectory may be simple but is not known to the human catcher who has to quickly estimate it, primarily using vision. This immediately suggests that eye and head movements that enable clear vision of the moving ball are an important part of real catching behavior. This can be seen even by informal observation and in Fig. 2(a). This is why we pay special attention to eye movements, and not just to body movements as is the usual practice in full body animation (i.e., other than in face animation). Two types of eye movements are observed (see Fig. 3). (1) *smooth pursuit* eye movements which attempt to continuously track the ball, especially prominent after the ball has passed the apex of its trajectory. (2) *catch-up saccades*, which are very fast movements, especially at the beginning when the quality of pursuit is not good.

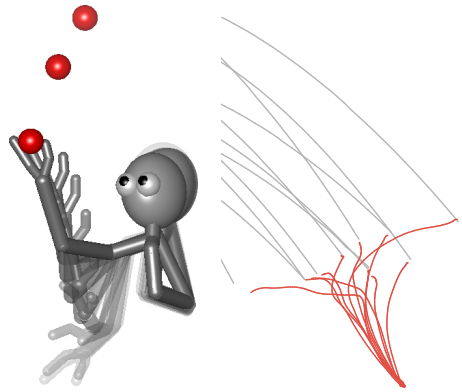
Second, it is not uncommon for the flight duration to be very short, about a second, and human vision is relatively slow, so there is very little time to obtain an accurate estimate and plan the catch before moving the hand. Instead the hand begins to move very early, when only a crude estimate is known, and appears to have an initial "open loop" phase, followed by a "closed loop" or "homing" phase. This was observed a century ago by Woodworth [1899], who proposed this "two component" model. Thus hand movements are not straight, preplanned trajectories, but are curved, with a stereotypical initial phase, followed by individual corrections to the target. See Fig. 2(b).

Third, much experimental evidence suggests that human movements are not planned in their entirety but generated by blending together discrete, short-duration submovements. Some of the evidence for and against this discrete submovement hypothesis is reviewed in Sec. 2. The example data shown in Figs. 2(b) and 6 show that real hand trajectories appear to have discontinuous higher derivatives, consistent with the hypothesis. This hypothesis is particularly appealing for computer animation, as we can leverage standard tools for constructing smooth curves from basis functions.

Finally, somewhat surprisingly, we observed close synchronization

*{shyeo,martinle,debanga,pai}@cs.ubc.ca

between the submovements detected in the hand movements and catch-up saccades of the eyes. This phenomenon can be seen in Fig. 4 and consistently in other trials. Saccades and the peak hand velocity appear to be very well synchronized. Furthermore, if we decompose the hand velocity profile into submovements, we find that the start times of the submovements are also well synchronized to saccades. We can also qualitatively observe from decomposition results in Fig. 6 that the three dimensional hand trajectory is well represented by piecewise linear segments that correspond to the submovement directions.



(a) (b)

Figure 2: *Captured catching behavior:* ***(a)*** *Temporal change of the upper body configuration.* ***(b)*** *Captured hand trajectories for eleven trials (red) and corresponding ball trajectories (gray).*

Figure 3: *Eye movements for three different ball trajectories, as seen in a reference frame fixed at the initial position of the eye. The ball (gray) goes upward initially and then downward after passing the apex (green circle); note that due to perspective projection, the true apex differs from that of the projected trajectory. Gaze position (red) are captured every 10 ms (dots).*

This analysis suggests an appealing hypothesis about movement generation, at least for visually guided interceptive movements: the eye and hand share the same *motor program* that is triggered by sensory events. Analogous hypotheses have been proposed by others, in slightly different contexts, e.g., [Starkes et al. 2002; Johansson et al. 2001; Hayhoe and Ballard 2005]. This primacy of gaze in locomotion is neatly summarized by Berthoz [2000] in the phrase: " 'Go where I'm looking', not 'Look where I'm going.' " It also underlies the coach's admonition to "keep your eyes on the ball" to improve performance. This hypothesis deserves more detailed study in the movement sciences, presumably using a minimalistic highly controlled experiment design, as in the work cited above, but that is not our goal here. Rather, our goal is to show that we can exploit it to efficiently generate realistic animations of goal directed behavior of a complex 3D character for computer graphics.

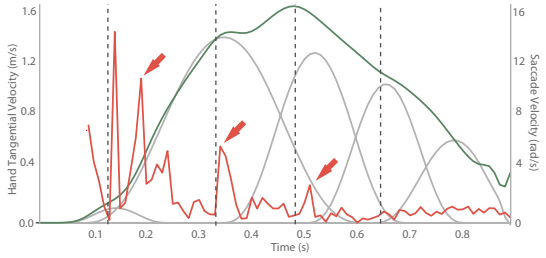


Figure 4: *Tangential velocity of captured hand (green) and eye movements (red). Red arrows are observed saccades of which the estimated trigger times are drawn as dotted lines. Gray curves shows the decomposed submovements.*

Contributions. We propose a new generative model for simulating a large class of fast interceptive movements. The model is grounded in experimental observations of human behavior, both in our lab and reported in the literature. A key feature of the model is that it is based on simultaneous measurements of eye movements along with traditional motion capture of hand and body movements; this has not been previously seen in computer animation, to the best of our knowledge. We combine a model of active vision with a model of movement generation using short duration discrete submovements, based on experimental observations. More importantly, we propose that submovements of the eyes, head, hand, and body are tightly synchronized, which provides an exceptionally convenient way to produce natural looking coordination. To our knowledge this is also new in computer animation. Several other known features of human catching behavior are included in the model as well. Even though the model captures significant features of real human behavior, several tradeoffs were made to make the simulation very efficient and practical, taking only a few seconds to generate realistic catching behavior. Since it is not based on raw motion capture data but on the underlying principles, the model generalizes very well to novel scenarios, such as catching with poor visibility and sudden changes in trajectory.

The paper is organized as follows. We first briefly review related work in Sec. 2. Sec. 3 introduces a practical model for visual estimation of target movement based on plausible statistics of human vision. Sec. 4 first shows how movements are composed of submovements. It then describes how gaze movements are first generated (Sec. 4.1), followed by synchronized movements of the head (Sec. 4.2), the hand (Sec. 4.3), and the body (Sec. 4.4). Sec. 5 and the video show the results. Sec. 6 concludes with a discussion of the limitations and benefits of our approach.

2 Related Work

Due to its importance, object interception and its related topics have been investigated not only in computer graphics, but also in neuroscience and robotics. We briefly review the most relevant work here.

Computer Graphics.
There have been very few papers written specifically on the topic of interception movements. Gillies and Dodgson [1999] proposed a method to solve the classical "baseball outfielder problem" of how to run to catch a ball, but no implementation results were reported. More recently, several character animation papers included object interception as an example [Francik and Szarowicz 2005; Abe and Popović 2006; Cooper et al. 2007], but did not consider gaze. Papers on object manipulation have either not included gaze (e.g., [Pollard and Zordan 2005; Liu 2009]) or included gaze as a

post-process for realism (e.g, [Tsang et al. 2005; Yamane et al. 2004]).

Gaze has been an important part of virtual characters [Lee et al. 2002; Itti 2003; Pelachaud and Bilvi 2003; Garau et al. 2003; Yamane et al. 2004; Gu and Badler 2006; Peters and Qureshi 2010], and crowd simulation [Shao and Terzopoulos 2005; Grillon and Thalmann 2009]. Perhaps the closest to our view of coupling perception and movement are the seminal papers by Terzopoulos and co-workers (e.g., [Tu and Terzopoulos 1994; Terzopoulos and Rabie 1997; Lee and Terzopoulos 2006]).

Movement Neuroscience.
Visuomotor coordination and object interception have been extensively studied in the movement sciences. A recent review by Zago et al. [2009] provides an excellent overview of this area. Other relevant works are cited in context throughout the rest of the paper. Even though some one dimensional models have been proposed (e.g., [Dessing et al. 2005]), much of this work is descriptive and can not be directly used for 3D character animation.

Submovements. A long standing hypothesis is that human (and animal) movement is not generated continuously, but produced by combining discrete building blocks, called submovements. There has been considerable experimental support for this hypothesis, starting from the seminal work of Soechting and Terzuolo [1987] and Vallbo and Wessberg [1993]. Although there are criticisms [Sternad and Schaal 1999] that this may not apply to rhythmic movements, the concept of submovements continues to have wide support and is used in recent analyses on hand movement [Novak et al. 2002] and clinical treatment after stroke [Dipietro et al. 2009].

Robotics.
Several impressive visuomotor control models for catching have been proposed in robotics [Hove and Slotine 1991; Riley and Atkeson 2002; Bauml et al. 2011; Birbach et al. 2011]. Since their main goal is to generate robust high performance catching, the proposed models are optimized for the robot system itself, rather than for producing human-like movement.

3 Visual Estimation of Target Movement

Human vision is remarkably complex and the subject of intense ongoing study. For computer animation we need a relatively simple model that captures the relevant features of human vision, but by no means all the complexities. It is commonly hypothesized that the brain is able to use the internal models of object dynamics [Wolpert et al. 1995] to estimate and predict the object's state over time. We use a Kalman filter for predicting the state of ball, using plausible models of noise introduced by visual sensing.

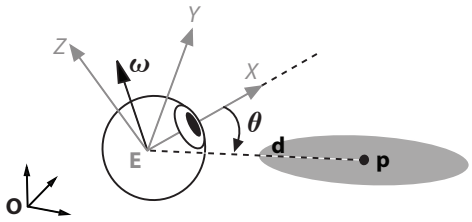


Figure 5: *Eye coordinate frame and the corresponding uncertainty of an object. Uncertainty is represented as a multivariate normal distribution. The error covariance ellipsoid around the true object location p is calculated by its deviation θ from the gaze direction (X-axis) and the axis of rotation ω.*

Fig. 5 defines a spatial coordinate frame attached to the eye, with its origin at the center of the globe, X-axis aligned with the visual axis, and Z-axis vertical in a reference position looking straight ahead. For ball catching, we define the perceived position and velocity of the ball as a six dimensional state vector in the eye coordinate frame, $\mathbf{x} = \begin{pmatrix} \mathbf{p} \\ \dot{\mathbf{p}} \end{pmatrix}$.

Table 1: *Parameters for error standard deviations of foveal vision: units for last three rows are identical to their original unit*

	Parameter	Value
Retinal Position	$\sigma_{q_z},\ \sigma_{q_y}$	$2.9 \times 10^{-4}\ rad$
Retinal Velocity	$\sigma_{\dot{q}_z}, \sigma_{\dot{q}_y}$	$0.05\lvert\dot{q}\rvert$
Depth	σ_x	$0.03d$
Depth Velocity	$\sigma_{\dot{x}}$	$0.5\lvert\dot{d}\rvert d$

3.1 Vision

We approximate the probability of the state of a moving point seen by the eye as a multivariate normal distribution in the eye coordinate frame. As shown in Fig. 5, the probability of the perceived position $\mathbf{p}$ and velocity $\dot{\mathbf{p}}$ of the target in the eye-fixed frame is:

$$\begin{bmatrix} \mathbf{p} \\ \dot{\mathbf{p}} \end{bmatrix} = N\left(\begin{bmatrix} \bar{\mathbf{p}} \\ \dot{\bar{\mathbf{p}}} \end{bmatrix},\ \mathbf{A}\boldsymbol{\Sigma}^2\mathbf{A}^T \right) \tag{1}$$

$$\boldsymbol{\Sigma} = \frac{1}{\alpha(\theta)}\ diag([\sigma_x, \sigma_y, \sigma_z, \sigma_{\dot{x}}, \sigma_{\dot{y}}, \sigma_{\dot{z}}]), \tag{2}$$

$$\mathbf{A} = \begin{bmatrix} e^{[\omega]\theta} & \mathbf{O}_{3\times3} \\ \mathbf{O}_{3\times3} & e^{[\omega]\theta} \end{bmatrix}, \tag{3}$$

where $\bar{\mathbf{p}}$ and $\dot{\bar{\mathbf{p}}}$ are the true position and the velocity of the ball in 3 dimensional space, $\boldsymbol{\Sigma}$ is the error covariance matrix in eye frame and $\mathbf{A}$ the corresponding transformation matrix from the target to the eye center. The block diagonal elements of $\mathbf{A}$ represent the rotation needed for the eye to foveate the target and $[\omega]$ is a skew symmetric matrix of the axis of this rotation.

The noise parameters in Eq. 2 are chosen as follows and are summarized in Table 1. The location of a point on the retina is represented using spherical coordinates, sequential rotation around Z and Y axes of the eye frame, called retinal coordinates: $q = (q_z, q_y)^T$. We set the standard deviations σ_{q_z} and σ_{q_y} to correspond to the spatial resolution of standard $20/20$ vision. Using the small angle approximation, we can convert the error from retinal to eye coordinates: $(\sigma_z, \sigma_y)^T = d(\sigma_{q_z}, \sigma_{q_y})^T$, where d is the distance from eye to the object.

The ability to discriminate the difference in retinal velocity is roughly proportional to the actual retinal velocity. The Weber fraction of velocity detection is known to be around 5% [McKee 1981]. Therefore, if the probability of perceived velocity is also assumed to be normally distributed over its true velocity, the standard deviation of the error, $(\sigma_{\dot{q}_z}, \sigma_{\dot{q}_y})^T$, can be chosen to be proportional to the actual velocity. Similar to the position error, the velocity error represented in eye coordinate will be $(\sigma_{\dot{z}}, \sigma_{\dot{y}})^T = d(\sigma_{\dot{q}_z}, \sigma_{\dot{q}_y})^T$.

Our ability to measure depth or change in depth is known to be much worse than detecting the retinal location and is perceived using binocular vision and other cues. Simulating stereo vision is beyond the scope of this paper; we set the depth resolution, σ_x and $\sigma_{\dot{x}}$, to be much higher and also proportional to the depth to approximate the vergence angle. Finally, the factor α accounts for the drop in acuity with distance to fovea, the central region of highest acuity.

3.2 Internal Model and Bayesian State Estimation

Using the vision model described above, a noisy observation of the target states is obtained. For simulation, we discretize the state update with a time step $\Delta t = 20\ ms$.

The internal model of the ball's dynamics is represented as

$$\mathbf{x}_{k+1} = \mathbf{F}\mathbf{x}_k + \mathbf{b}, \tag{4}$$

$$\mathbf{F} = \begin{bmatrix} \mathbf{I}_{3\times3} & \mathbf{I}_{3\times3}\Delta t \\ \mathbf{O}_{3\times3} & \mathbf{I}_{3\times3} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} \mathbf{O}_{3\times1} \\ \mathbf{g}\Delta t \end{bmatrix}, \tag{5}$$

where $\mathbf{g}$ is gravity vector, which we assume to be constant. In other words, we assume that the brain has prior knowledge about gravity (see, e.g., [McIntyre et al. 2001]).

Using the vision model described above, we can then model the observer. If $\mathbf{R}_k$ is the orientation and $\mathbf{r}_k$ the position of the eye frame at time k, the observed position and velocity of the ball $\mathbf{z}_k$ is:

$$\mathbf{z}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{h}_k + \mathbf{v}_k, \tag{6}$$

where $\mathbf{v}_k$ is the observation noise defined in Eq. 3 while $\mathbf{H}_k$ and $\mathbf{h}_k$ are the transformation from the spatial to the eye frame:

$$\mathbf{H}_k = \begin{bmatrix} {\mathbf{R}_k}^T & \mathbf{O}_{3\times3} \\ \mathbf{O}_{3\times3} & {\mathbf{R}_k}^T \end{bmatrix}, \quad \mathbf{h}_k = \begin{bmatrix} -{\mathbf{R}_k}^T\mathbf{r}_k \\ \mathbf{O}_{3\times1} \end{bmatrix}. \tag{7}$$

We use a standard Kalman filter to implement the Bayesian inference with internalized dynamics in the brain. For completeness, the predict-update filtering algorithm is given below.

- Predict:

$$\mathbf{x}_{k|k-1} = \mathbf{F}\mathbf{x}_{k-1|k-1} + \mathbf{b} \tag{8}$$

$$\mathbf{P}_{k|k-1} = \mathbf{F}\mathbf{P}_{k-1|k-1}\mathbf{F}^T \tag{9}$$

- Update:

$$\mathbf{y}_k = \mathbf{z}_k - \mathbf{H}_k\mathbf{x}_{k|k-1} - \mathbf{h}_k \tag{10}$$

$$\mathbf{S}_k = \mathbf{H}_k\mathbf{P}_{k|k-1}{\mathbf{H}_k}^T + \mathbf{A}_k{\mathbf{\Sigma}_k}^2{\mathbf{A}_k}^T \tag{11}$$

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}{\mathbf{H}_k}^T{\mathbf{S}_k}^{-1} \tag{12}$$

$$\mathbf{x}_{k|k} = \mathbf{x}_{k|k-1} + \mathbf{K}_k\mathbf{y}_k \tag{13}$$

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{k|k-1}, \tag{14}$$

where $\mathbf{A}_k$ and $\mathbf{\Sigma}_k$ are previously defined in Eq. 3. Since we assume that the subject initially has no prior knowledge about the states, the initial value of the prior estimate covariance, $P_{0|0}$, is made sufficiently large and $\mathbf{x}_{0|0}$ is chosen to be the zero vector.

The performance of the Kalman filter in object tracking is known to deteriorate seriously if the tracked object changes its trajectory suddenly, especially when the error covariance is small, such as the case when the ball we are tracking bounces unexpectedly. Therefore, in order to implement human vision's ability to track an object with a sudden trajectory change, we should endow an adaptive behavior to the filter. This is typically achieved by resetting the prior error covariance matrix $\mathbf{P}_{k|k-1}$ to a sufficiently high value when an abnormal tracking error is recognized. We choose the reset decision variable as a Mahalanobis measure of the innovation: ${\mathbf{y}_k}^T{\mathbf{S}_k}^{-1}\mathbf{y}_k$, which indicates the abnormality of the error between prediction and observation.

4 Movement Generation

Based on the visually estimated target movement, we can then generate the gaze movement to the target (Sec. 4.1), followed by synchronized movements of the head (Sec. 4.2), the hand (Sec. 4.3), and the body (Sec. 4.4).

To make this complex whole body movement tractable, the movements are generated kinematically, rather than using a dynamic model of human biomechanics. This is an approximation we make primarily for efficiency, but one that is well supported by human movement research. Even though dynamic properties of the body are important for the ultimate control of the movement, the brain appears to represent external movement goals kinematically at a high level. This can be seen by the fact that one's handwriting looks the same regardless of posture or writing tool used; this phenomenon was termed "motor equivalence" by Donald Hebb [1949]. Since then, several kinematic invariants and laws have been proposed for movement planning, though the topic remains controversial, as are most topics in neuroscience. The dynamics must influence the movement at some level, at least in the lower level controllers. We take this evidence pragmatically to mean that a kinematic approximation is plausible, and helps to improve the efficiency of the animation system.

As discussed in Sec. 1 we assume that movements are produced by blending short duration submovements together. Each submovement is smooth, with a bell-shaped velocity[1] profile. Different smooth shapes have been proposed, including those that minimize jerk (time derivative of acceleration) [Flash and Hogan 1985] and the more flexible delta lognormal shape [Plamondon 1995]. We chose the minimum jerk profile for simplicity, as it has fewer parameters. The tangential velocity $v(t^0, t^f, t)$ with unit displacement is defined in a time interval $t^0 < t < t^f$ as:

$$v(t^0, t^f, t) = \frac{30}{(t^f - t^0)^5}(t - t^0)^2(t - t^f)^2. \tag{15}$$

The velocity of a movement $\dot{\mathbf{u}}(t)$ is represented as a superposition of submovement velocity profiles $\dot{\mathbf{u}}_i$:

$$\dot{\mathbf{u}}(t) = \sum_{i=1}^{n} \dot{\mathbf{u}}_i(t) = \sum_{i=1}^{n} \mathbf{b}_i\ v(t_i^0, t_i^f, t), \tag{16}$$

where $\mathbf{b}_i$ is the basis vector representing the direction and magnitude of the submovement. Very fast movements, such as saccades, can consist of single submovement, and slower movements of head and hand can consist of multiple submovements. Care should be taken in case of superimposed submovements: when the destination of the new submovement is given, $\mathbf{b}_i$ is not the difference between the destination and current position, but rather, it should be the difference between the new destination and the previous submovement destination.

Fig. 6 shows the velocity profile and decomposed submovement of some measured hand trajectories. Decompositions are done by nonlinear least squares optimization with pre-chosen number of submovements: finding a set of submovement parameters, $\mathbf{b}_i, t^0, t^f$, that minimizes the error between captured and composed trajectories. Note that this submovement decomposition may not be unique and several different combination can exist depending on the initial condition and selection of the basis [Rohrer and Hogan 2003].

Using this framework, animation of body movement, including head movement and hand movement becomes very clear: for each update of visual information, corresponding submovements of each body part is determined and superimposed on the current motion.

[1] More accurately, *speed*, but we follow the usage in the human movement sciences.

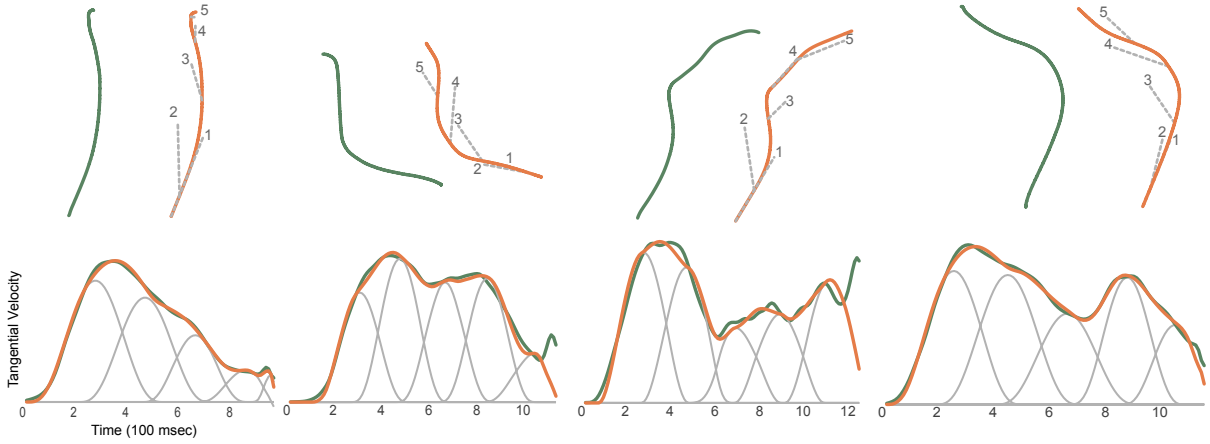


Figure 6: *Decomposed submovement of hand trajectory: [Top row] Real (green) and synthesized (orange) submovements shown in three dimensional space. Dotted lines show the direction of each submovement. [Bottom row] Decomposed submovements shown in tangential velocity*

4.1 Gaze Movements

Given the estimated ball state, we first determine gaze, that is, where to look. We define (binocular) gaze as a point in space where the eyes are looking. This determines the orientations of the eyes in space. As described earlier, a distinguishing feature of our approach to animation is that the gaze is determined prior to determining body movement. In humans and other animals, gaze is not significantly affected by body movement unless the eyes reach a biomechanical limit. Indeed, several fast low-level reflexes, including the vestibulo-ocular and vestibulo-collic reflexes, stabilize gaze in space and cancel disturbances due to body movements.

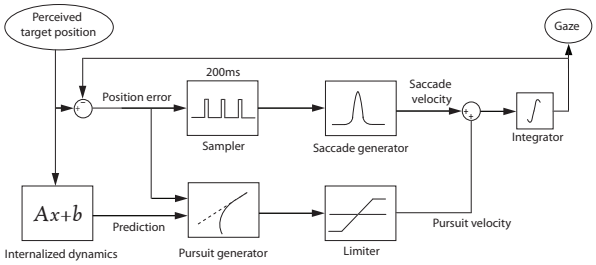


Figure 7: *Overall control structure of gaze*

To foveate an object of interest, two types of gaze movements are used, called saccade and pursuit. The saccade is a very fast, impulsive eye movement (up to 800°/s in humans) [Leigh and Zee 1999], but the high speed comes at a cost: vision is very poor during a saccade. Humans and other primates can use smooth pursuit eye movements to continuously track moving objects, but only at relatively low speeds (an order of magnitude slower than saccades).

The properties of saccades and pursuit are well studied [Young and Stark 1963; Robinson et al. 1986], in particular [Orban de Xivry and Lefèvre 2007] provide a recent perspective. Our goal here is not to precisely represent the decision to saccade or pursue, but rather to have a simple model that captures the statistical regularities observed in the data. The typical time required for two consecutive saccades is known to be around 200 ms [Young and Stark 1963; Robinson 1965], corresponding to the 200 ms recharge duration of the saliency map in the model proposed by Itti et al. [2006]. Tracking fast moving objects is challenging, and is likely to require a rapid sequence of catch-up saccades. As shown in Fig. 4, we also observe this in our experimental data, with catch-up saccades triggered at approximately 200 ms interval. Therefore, we assume 200 ms as the visuomotor update time that regulates the general movement including head and hand movement. The overall block diagram of the eye movement velocity generator, including saccade and pursuit, is described in Fig. 7.

The saccade amplitude is determined by current error between gaze and target angles. Given its amplitude, the likely duration of a saccade can be well estimated from the "main sequence" relationship [Leigh and Zee 1999]. The velocity of saccade is chosen as in Eq. 15; real saccades have a more asymmetric velocity profile, but the difference is imperceptible for computer animation.

Since pursuit is a smooth movement, we set its velocity to be updated at every time step Δt with $\mathbf{v} = (\bar{\mathbf{p}} - \mathbf{p})/\Delta t$. Here, $\mathbf{p}$ is the current gaze and $\bar{\mathbf{p}}$ is the position of the target at the next step as predicted by the internal model. The maximum pursuit velocity is limited to 100°/s as per human oculomotor limits [Meyer et al. 1985], restricting it to follow only low speed targets.

When an object is detected, the estimated positions in the early stages will be very noisy. Therefore, if the saccade is simply triggered to the estimated position, the initial eye movement will be very inaccurate and unrealistic. The brain avoids this problem by suppressing the initial saccade until the likelihood of position of the object exceeds some significance criterion [Carpenter and Williams 1995]. We relate this decision criterion to the innovation observed by the perceptual system, using the error covariance of innovation that is defined in Eq. 11: the initial saccade is triggered when the estimated error of the object position in the retinal coordinate, which is the Y-Z coordinate of the eye frame, is reduced below a certain threshold.

4.2 Head Movements

We assume that the main goal of head movement in target tracking is to assist the eye in tracking the target more easily. Therefore, the movement for the head is defined similarly to the gaze: we define head movement by setting the "head gaze", a spatial point that the head is facing. Since the rotation of the head follows Donders' law [Crawford et al. 2003], which states that three dimensional head rotation can sufficiently be described by two parameters, the longitudinal rotation followed by latitudinal rotation, the rotation matrix

of the head, $\mathbf{R}_h$, is determined by the head gaze. If the head gaze is represented with spherical coordinates, $[q_z, q_y]$, with respect to the head frame in the home configuration, $\mathbf{R}_{h0}$, the corresponding head orientation is $\mathbf{R} = \mathbf{R}_{h0}\mathbf{R}_z(q_z)\mathbf{R}_y(q_y)$.

It is known that the same brain areas are involved in eye and head movement in both saccades [Freedman et al. 1996] and pursuit [Lanman et al. 1978]. Therefore, we assume that the head also generates saccades and pursuit in the same way as the eye, whereas the head saccade is a slower, longer lasting bell shaped velocity profile compared to the eye. Given a gaze shift, the corresponding saccade amplitude of the head is linearly related with a 20° dead zone [Leigh and Zee 1999]. There is a more sophisticated model proposed by Freedman [2001] for the kinematic relationship between the head and eye movement, but this model is not directly applicable to 3-dimensional, large angle movements since it is focused on horizontal movements with a limited range.

Once the change of head gaze is determined, the peak velocity is also determined by its well known linear relationship to the amplitude [Leigh and Zee 1999], with slope varying between $4s^{-1}$ to $8s^{-1}$. If we apply this relationship to the submovement shape function in Eq. 15, we get a constant submovment duration around $400\ ms$. Taken together, we can define the submovement parameters in Eq. 16 for saccade velocity of the head gaze, $\dot{\mathbf{g}}$:

$$\xi = \max(0, \theta - 0.349) \quad (17)$$

$$\mathbf{b}_i = (e^{[\omega_i]\xi} - I)(\mathbf{g}_{i-1} - \mathbf{p}) \quad (18)$$

where $[\omega_i,\ \theta_i]$ is the axis-angle representation of the displacement from the last head gaze destination $\mathbf{g}_{i-1}$ to the current eye gaze destination with respect to the current head position $\mathbf{p}$, and ξ is the required angular displacement of the head saccade. Here θ is the non-negative magnitude of the rotation, with the sign of the rotation absorbed in ω. Note that this definition will not generate a minimum jerk angular velocity profile, but it is sufficiently close when the gaze is not very near. The pursuit velocity of the head, which is identical to that of the eye, is then added.

4.3 Hand Movements for Interception

With predetermined gaze behavior, the strategy of manual interception can be modeled as a simple algorithm. When the gaze is successfully tracking the ball, we can simply move the hand towards the gaze. In other words, if the hand is "latched" to the eye frame while the gaze is latched to the target, the hand will be driven sufficiently close to the ball when it reaches the body. This strategy can be observed when we try to catch a fly. Before we trigger the final snatch, we move our hand in the same way that we move our eyes and head.

When the hand is synchronized to the gaze, the remaining decision will be to determine the distance of the catch. The data shown in Fig. 8 strongly suggests that each subject has a preferred interception distance from the head frame. This distance may be affected by considerations such as the manipulability of the arm or the effective compliance produced by muscle properties, but for simplicity we choose a fixed distance as a decision variable for catching.

If we draw a sphere whose radius is the preferred distance around head, the catch position is determined as an early intersection point between the sphere and the predicted ball trajectory. Note that this interception planning is only valid when we have a sufficiently accurate prediction of the object trajectory. For this decision, we use the determinant of the current posterior error covariance of the state, defined in Eq. 14, as a decision variable.

Note that a decision made too early can cause a serious mistake,

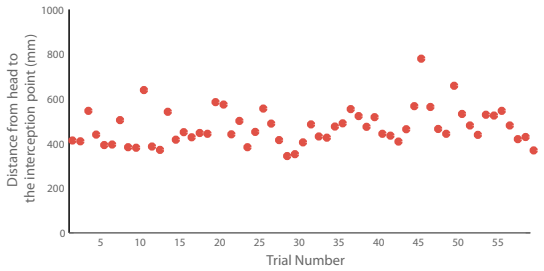


Figure 8: *Distance from head to the interception point summarized from 59 ball catching trials for one subject*

whereas one that is too late gives too little time to move. Therefore, we empirically searched for the best decision timing and corresponding value of the decision variable. Interestingly, from the simulation result, we see that in most cases a good decision is made around the apex of the ball trajectory as illustrated in Fig. 9. This agrees with the transition from catch-up saccades to pursuit that is previously shown in Fig. 3, and also corresponds to the large change of the hand trajectory observed in the data. Fig. 10 shows the captured hand trajectory for different catching trials and we can see that a clear branch of the trajectory toward the target occurs around the apex of the ball.

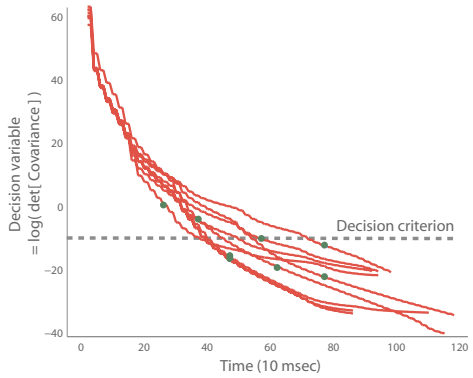


Figure 9: *Decision variables and ball apex: Lines show the decision variable, the determinant of the posterior error covariance matrix, for various catching simulations. For each trial, the time when the ball reaches its apex is marked as a green dot. Gray dotted line shows the criterion used for the prediction.*

The reason why the decision is made around the apex can be intuitively understood: since the velocity of the ball reaches its minimum at the apex and the gaze usually catches up with the ball by saccades and start to pursue, the quality of the position and velocity estimation is considerably improved around the apex. This corresponds to the fact skilled jugglers only look at the juggled balls at their apex [Beek and Lewbel 1995].

Based on this hypothesis, we divide the interception strategy into two parts: a reactive phase and a proactive phase, similar to Woodworth's two component model. The strategies for composing submovements for each phase are as follows:

Reactive Phase. A submovement to drive the hand to the gaze direction with preferred catching distance is generated. The destination of the submovement is not exactly set to the gaze center to avoid the occlusion of the target due to the hand. For a right-handed subject, the destination is set to be the 45 degrees lower right point with respect to the gaze that corresponds to the captured configura-

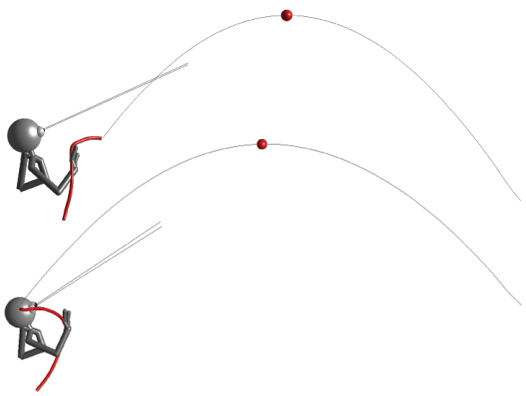


Figure 10: *Transition of the hand trajectory around the apex (from real data). The subject's pose corresponds to the ball apex. Red curves show the entire hand trajectories. It can clearly be seen that significant modification of the trajectory occurs around the apex.*

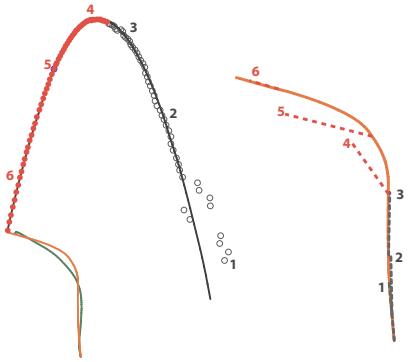


Figure 11: *Submovement composition of the hand trajectory for a given ball trajectory (black curve). Left: Circles show the perceived position of the ball over time. The simulated hand trajectory (orange curve) is compared to the captured data (green curve). As the catch phase changes from reactive (gray circles) to proactive (red circles), corresponding submovements are generated. Right: dotted lines (gray for reactive and red for proactive phase) are directions and amplitudes of submovements triggered at each visuomotor update (numbers match to the estimated ball position on the trajectory).*

tion in Fig. 10. We name this position as the *ready pose*, $\mathbf{p}_{ready}$, that is fixed with respect to the eye frame.

If $\mathbf{R}_e$ and $\mathbf{p}_e$ are the orientation and position of the eye frame, and $\tilde{\mathbf{p}}_m$ is the destination of the previous submovement, direction vector $\mathbf{b}$ in Eq. 16 can be determined as follows:

$$\mathbf{b} = \mathbf{R}_e \mathbf{p}_{ready} + \mathbf{p}_e - \tilde{\mathbf{p}}_m \tag{19}$$

We limit the maximum velocity to $2m/s$ as observed in the data. We also assume from the data that the duration of the submovement is approximately $400\ ms$ so that it reaches the maximum velocity when the next submovement is triggered. This is consistent with the observation of Novak et al. [2002].

Proactive Phase. When the subject has a sufficiently accurate prediction of the ball trajectory, the observed ball position and velocity can be extrapolated into the future to determine the interception point. Here we limit the prediction to a specific time window, $400\ ms$, that reflects the prediction limit of the internal model. When the ball penetrates the sphere of the preferred distance, the interception point and remaining time is determined and a corresponding submovement is generated. Otherwise, the hand goes in the direction of the closest point on the predicted trajectory with maximum $2m/s$ velocity. The magnitude of the submovement is also adjusted so that the final hand position does not exceed the biomechanical limit of reach. Using this strategy, we can generate the distinctive curved path of the hand trajectory observed from the data. Fig. 11 shows the resultant hand trajectory and submovements.

The orientation and grasping angle of the hand can also be determined in this phase by applying a relationship observed in the data. From the data, we observe that the angle of attack between the ball and the palm remains at 90 degrees and grasping starts with the last submovement, when the decision to catch is made.

4.4 Body Movements

As previously described, we focus on generating coupled motion between gaze, head, and hand. Body movements are not the focus of this work as they have been extensively studied in computer animation. For a given trajectory of the hand and head, we solve for the configuration of the rest of the body using inverse kinematics (IK). The character model used consists of a torso that is connected to the ground by a 3 DOF rotational joint and a 7 DOF (right) arm model. The position of the head is fixed on the torso. The input provided to the IK algorithm consists of the position and the normal vector of the palm that is adjusted to satisfy the preferred angle of attack between the ball and the hand.

Although general treatment for IK is not our main focus, its solution is not trivial due to kinematic redundancy, with biomechanical constraints and learned styles contributing to realistic solutions. We first implemented the IK solution proposed by [Grochow et al. 2004], but it did not generalize well when the movement approaches or crosses the boundary of the training set where the learned movements need to be extrapolated. More recent latent variable models [Urtasun et al. 2008; Wang et al. 2007] may fare better, but we instead used a simpler approach of combining example configurations with weighted pseudo-inverse of the Jacobian to obtain our IK solution (for all 3+7 DOF).

Our approach first solves for the torso and arm joint angles according to the specified hand position, we then solve for the relative hand orientation according to the ball velocity. To solve for the torso and arm configuration, we first search in our training set (derived from motion capture data) for configurations with hand position most similar to the target position. The average of these configurations is then blended with the previous timestep configuration and a weighted pseudo-inverse of the Jacobian is used to correct the resulting configuration so that its hand position matches the target. During the reactive phase, the rotation of the hand is interpolated from its home orientation to a predetermined *ready pose* orientation with a set velocity. Subsequently, in the proactive phase, we find the rotation that gives the nearest interpolation from the previous palm normal vector to the estimated ball velocity vector and apply it to the hand. Finally, we use a canned animation to generate the finger grasping motion during the final moments of interception.

5 Results

Using the proposed methods, we developed a fully generative 3D character animation model of the ball catching, shown in Fig. 1 and the accompanying video.

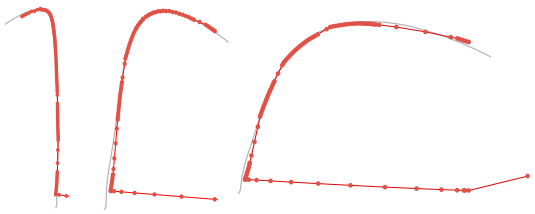


Figure 12: *Simulated gaze behavior for a given ball trajectory. Colors and data rates are the same as Fig. 3.*

Figure 13: *Comparison between real (green) and simulated (red) hand trajectories for the same ball catching tasks. Submovements vectors are plotted as a gray dotted lines. Note that in the bottom right trials, the simulated character failed to catch the ball since it is subjected to a very extreme condition.*

Validation. We compared the results of our simulation to measurements of human catching, to see if there is a qualitative agreement. Note that the measurements were *not* used for estimating the parameters of the model, so qualitative agreement is the best one can expect. We recorded the movement of upper body using an 8-camera Vicon MX motion capture system (Vicon, Los Angeles), and eye movements using a head mounted C-ETD eyetracker (Chronos Vision, Berlin) to capture the head-unrestrained eye movements. Both systems recorded at 100 Hz, which was sufficient for our needs. Subjects were seated on a comfortable stool and instructed to catch a $5cm$ diameter ball thrown in the subject's direction. The gaze behavior for different ball trajectories was previously shown in Fig. 3. The simulated data generated using our system is shown in Fig. 12 and looks qualitatively similar. Comparison of measured and simulated hand trajectories in Fig. 13 shows that the proposed algorithm is sufficient to generate qualitative details of real hand movement as well.

Generalization. We simulated ball catching for many different trajectories. The character is able to catch the ball successfully in 98 percent of the trials. To test the model capability to generalize beyond simple trials, we introduce several variations in the catching scenario. In one scenario, we assume that the ball unexpectedly bounces off a transparent wall. We found that the Kalman filter with adaptive resetting successfully tracks the ball after the bounce. The character often fails to catch if the time to recover the prediction is too short, just as humans would. In another scenario, we simulated poor vision by increasing the standard deviation of the error in Table 1 by a factor of 10. As expected, the resultant motion of character showed a slow start-up followed by an abrupt catching with an 81.2 percent success ratio.

Performance. By design, the model is very efficient. We implemented the entire algorithm in MATLAB (version 7.9.0, R2009b, MathWorks Inc.) without using externally compiled subroutines. The software ran on a 3.40 GHz Intel Core i7 computer. The computation for generating the entire catching movement, including IK, for 100 frames, with a 1 second ball trajectory, takes 5.93 seconds (16.9 fps).

6 Conclusions

We have proposed a framework of animating visually guided interception based on the neurophysiology of the visuomotor system. The framework includes novel features such as gaze based motor coordination and submovement composition, that have not been previously introduced in computer animation. Since the model is grounded in human behavior, the simulated catching movements look very human-like. Interestingly, we found that the use of submovements produces very subtle behaviors such as discontinuities and hesitation, which are clearly different from continuous, robotic motions.

Our framework currently has several limitations. These are mostly due to this paper's focus on constructing a complete and practical generative model for computer animation, which required simplifying many of the known complexities of the human sensorimotor system. First, our vision model is a highly simplified approximation of the human visual system. Our linear Kalman filter could be improved using more sophisticated (and complex) Bayesian filters. Although a visuomotor update interval of about 200 ms is observed in both our data and in previous studies, it is likely that there is some variability as a result of a probabilistic decision process. A more accurate model of this decision process may be able to account for the variability. In general, it may also be desirable to model context-dependent and individual variability in the coordination pattern. Finally, we currently rely on kinematic control, but it is clear that a dynamic model, based on realistic biomechanics and neural control (e.g., [Sueda et al. 2008; Lee et al. 2009]), may be able to better account for movements at the limits of human performance.

Notwithstanding the limitations listed above, we believe our framework provides a new approach to computer animation, based on principles that may be used by the human brain to control movement, such as visual motion estimation, use of gaze to control body movements, and generating movements using overlapping submovements. Since we base our animations on these principles, rather than on a corpus of raw data, the method is very efficient and generalizes in plausible ways to novel scenarios.

Acknowledgements This work was supported in part by the Canada Research Chairs Program, the Peter Wall Institute for Advanced Studies, NSERC, ICICS, and the Canada Foundation for Innovation, with additional support from the Human Frontier Science Program.

References

ABE, Y., AND POPOVIĆ, J. 2006. Interactive animation of dynamic manipulation. In *Proceedings of the 2006 ACM SIGGRAPH/Eurographics symposium on Computer animation*, Eurographics Association, 195–204.

BAUML, B., SCHMIDT, F., WIMBOCK, T., BIRBACH, O., DIETRICH, A., FUCHS, M., FRIEDL, W., FRESE, U., BORST, C., GREBENSTEIN, M., ET AL. 2011. Catching flying balls and preparing coffee: Humanoid rollin'justin performs dynamic and sensitive tasks. In *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, IEEE, 3443–3444.

BEEK, P., AND LEWBEL, A. 1995. The science of juggling. *Scientific American 273*, 5, 92–97.

BERTHOZ, A. 2000. *The brain's sense of movement*. Harvard Univ Pr.

BIRBACH, O., FRESE, U., AND BAUML, B. 2011. Realtime perception for catching a flying ball with a mobile humanoid. In *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, IEEE, 5955–5962.

CARPENTER, R., AND WILLIAMS, M. 1995. Neural computation of log likelihood in control of saccadic eye movements. *Nature 377*, 6544, 59–62.

COOPER, S., HERTZMANN, A., AND POPOVIĆ, Z. 2007. Active learning for real-time motion controllers. In *ACM Transactions on Graphics (TOG)*, vol. 26, ACM, 5.

CRAWFORD, J., MARTINEZ-TRUJILLO, J., AND KLIER, E. 2003. Neural control of three-dimensional eye and head movements. *Current opinion in neurobiology 13*, 6, 655–662.

DESSING, J., PEPER, C., BULLOCK, D., AND BEEK, P. 2005. How position, velocity, and temporal information combine in the prospective control of catching: Data and model. *Journal of cognitive neuroscience 17*, 4, 668–686.

DIPIETRO, L., KREBS, H., FASOLI, S., VOLPE, B., AND HOGAN, N. 2009. Submovement changes characterize generalization of motor recovery after stroke. *Cortex 45*, 3, 318–324.

FLASH, T., AND HOGAN, N. 1985. The coordination of arm movements: an experimentally confirmed mathematical model. *The journal of Neuroscience 5*, 7, 1688–1703.

FRANCIK, J., AND SZAROWICZ, A. 2005. Character animation with decoupled behaviour and smart objects. In *6th International Conference on Computer Games CGAIMS, Louisville, Kentucky, USA*.

FREEDMAN, E., STANFORD, T., AND SPARKS, D. 1996. Combined eye-head gaze shifts produced by electrical stimulation of the superior colliculus in rhesus monkeys. *Journal of neurophysiology 76*, 2, 927–952.

FREEDMAN, E. 2001. Interactions between eye and head control signals can account for movement kinematics. *Biological cybernetics 84*, 6, 453–462.

GARAU, M., SLATER, M., VINAYAGAMOORTHY, V., BROGNI, A., STEED, A., AND SASSE, M. 2003. The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, 529–536.

GILLIES, M., AND DODGSON, N. 1999. Ball catching: An example of psychologically-based behavioural animation. Eurographics UK.

GRILLON, H., AND THALMANN, D. 2009. Simulating gaze attention behaviors for crowds. *Computer Animation and Virtual Worlds 20*, 2-3, 111–119.

GROCHOW, K., MARTIN, S. L., HERTZMANN, A., AND POPOVIĆ, Z. 2004. Style-based inverse kinematics. In *ACM SIGGRAPH 2004 Papers*, ACM, New York, NY, USA, SIGGRAPH '04, 522–531.

GU, E., AND BADLER, N. 2006. Visual attention and eye gaze during multiparty conversations with distractions. In *Intelligent Virtual Agents*, Springer, 193–204.

HAYHOE, M., AND BALLARD, D. 2005. Eye movements in natural behavior. *Trends in cognitive sciences 9*, 4, 188–194.

HEBB, D. 1949. *The organization of behavior: A neuropsychological theory*. Lawrence Erlbaum.

HOVE, B., AND SLOTINE, J. 1991. Experiments in robotic catching. In *American Control Conference, 1991*, IEEE, 380–386.

ITTI, L. 2003. Realistic avatar eye and head animation using a neurobiological model of visual attention. Tech. rep., DTIC Document.

ITTI, L. 2006. Quantitative modelling of perceptual salience at human eye position. *Visual cognition 14*, 4-8, 959–984.

JOHANSSON, R., WESTLING, G., BÄCKSTRÖM, A., AND FLANAGAN, J. 2001. Eye–hand coordination in object manipulation. *the Journal of Neuroscience 21*, 17, 6917–6932.

LANMAN, J., BIZZI, E., AND ALLUM, J. 1978. The coordination of eye and head movement during smooth pursuit. *Brain Research 153*, 1, 39–53.

LEE, S., AND TERZOPOULOS, D. 2006. Heads up!: biomechanical modeling and neuromuscular control of the neck. In *ACM Transactions on Graphics (TOG)*, vol. 25, ACM, 1188–1198.

LEE, S., BADLER, J., AND BADLER, N. 2002. Eyes alive. In *ACM Transactions on Graphics (TOG)*, vol. 21, ACM, 637–644.

LEE, S., SIFAKIS, E., AND TERZOPOULOS, D. 2009. Comprehensive biomechanical modeling and simulation of the upper body. *ACM Transactions on Graphics (TOG) 28*, 4, 99.

LEIGH, R., AND ZEE, D. 1999. *The neurology of eye movements*. No. 55. Oxford Univ Pr.

LIU, C. 2009. Dextrous manipulation from a grasping pose. In *ACM Transactions on Graphics (TOG)*, vol. 28, ACM, 59.

MCINTYRE, J., ZAGO, M., BERTHOZ, A., LACQUANITI, F., ET AL. 2001. Does the brain model newton's laws? *Nature Neuroscience 4*, 7, 693–694.

MCKEE, S. 1981. A local mechanism for differential velocity detection. *Vision Research 21*, 4, 491–500.

MEYER, C., LASKER, A., AND ROBINSON, D. 1985. The upper limit of human smooth pursuit velocity. *Vision Research 25*, 4, 561–563.

NOVAK, K., MILLER, L., AND HOUK, J. 2002. The use of overlapping submovements in the control of rapid hand movements. *Experimental Brain Research 144*, 3, 351–364.

ORBAN DE XIVRY, J., AND LEFÈVRE, P. 2007. Saccades and pursuit: two outcomes of a single sensorimotor process. *The Journal of Physiology 584*, 1, 11–23.

PELACHAUD, C., AND BILVI, M. 2003. Modelling gaze behavior for conversational agents. In *Intelligent Virtual Agents*, Springer, 93–100.

PETERS, C., AND QURESHI, A. 2010. A head movement propensity model for animating gaze shifts and blinks of virtual characters. *Computers & Graphics 34*, 6, 677–687.

PLAMONDON, R. 1995. A kinematic theory of rapid human movements. *Biological Cybernetics 72*, 4, 295–307.

POLLARD, N., AND ZORDAN, V. 2005. Physically based grasping control from example. In *Proceedings of the 2005 ACM SIGGRAPH/Eurographics symposium on Computer animation*, ACM, 311–318.

RILEY, M., AND ATKESON, C. 2002. Robot catching: Towards engaging human-humanoid interaction. *Autonomous Robots 12*, 1, 119–128.

ROBINSON, D., GORDON, J., AND GORDON, S. 1986. A model of the smooth pursuit eye movement system. *Biological Cybernetics 55*, 1, 43–57.

ROBINSON, D. 1965. The mechanics of human smooth pursuit eye movement. *The Journal of Physiology 180*, 3, 569.

ROHRER, B., AND HOGAN, N. 2003. Avoiding spurious submovement decompositions: a globally optimal algorithm. *Biological cybernetics 89*, 3, 190–199.

SHAO, W., AND TERZOPOULOS, D. 2005. Autonomous pedestrians. In *Proceedings of the 2005 ACM SIGGRAPH/Eurographics symposium on Computer animation*, ACM, 19–28.

SOECHTING, J., AND TERZUOLO, C. 1987. Organization of arm movements. motion is segmented. *Neuroscience 23*, 1, 39–51.

STARKES, J., HELSEN, W., AND ELLIOTT, D. 2002. A menage a trois: the eye, the hand and on-line processing. *Journal of sports sciences 20*, 3, 217–224.

STERNAD, D., AND SCHAAL, S. 1999. Segmentation of endpoint trajectories does not imply segmented control. *Experimental Brain Research 124*, 1, 118–136.

SUEDA, S., KAUFMAN, A., AND PAI, D. K. 2008. Musculotendon simulation for hand animation. *ACM Trans. Graph. (Proc. SIGGRAPH) 27*, 3, 83:1–83:8.

TERZOPOULOS, D., AND RABIE, T. 1997. Animat vision: Active vision in artificial animals. *Videre. Journal of Computer Vision Research 1*, 1, 2–19.

TSANG, W., SINGH, K., AND FIUME, E. 2005. Helping hand: an anatomically accurate inverse dynamics solution for unconstrained hand motion. In *Proceedings of the 2005 ACM SIGGRAPH/Eurographics symposium on Computer animation*, ACM, 319–328.

TU, X., AND TERZOPOULOS, D. 1994. Artificial fishes: Physics, locomotion, perception, behavior. In *Proceedings of the 21st annual conference on Computer graphics and interactive techniques*, ACM, 43–50.

URTASUN, R., FLEET, D. J., GEIGER, A., POPOVIĆ, J., DARRELL, T. J., AND LAWRENCE, N. D. 2008. Topologically-constrained latent variable models. In *Proceedings of the 25th international conference on Machine learning*, ACM, New York, NY, USA, ICML '08, 1080–1087.

VALLBO, A., AND WESSBERG, J. 1993. Organization of motor output in slow finger movements in man. *The Journal of physiology 469*, 1, 673.

WANG, J. M., FLEET, D. J., AND HERTZMANN, A. 2007. Multifactor gaussian process models for style-content separation. In *Proceedings of the 24th international conference on Machine learning*, ACM, New York, NY, USA, ICML '07, 975–982.

WOLPERT, D., GHAHRAMANI, Z., AND JORDAN, M. 1995. An internal model for sensorimotor integration. *Science 269*, 5232, 1880.

WOODWORTH, R. 1899. Accuracy of voluntary movement. *The Psychological Review: Monograph Supplements 3*, 3, i.

YAMANE, K., KUFFNER, J., AND HODGINS, J. 2004. Synthesizing animations of human manipulation tasks. In *ACM Transactions on Graphics (TOG)*, vol. 23, ACM, 532–539.

YOUNG, L., AND STARK, L. 1963. Variable feedback experiments testing a sampled data model for eye tracking movements. *Human Factors in Electronics, IEEE Transactions on*, 1, 38–51.

ZAGO, M., MCINTYRE, J., SENOT, P., AND LACQUANITI, F. 2009. Visuo-motor coordination and internal models for object interception. *Experimental Brain Research 192*, 4, 571–604.