**4视觉系统**

This section describes our vision system. In Section 4.1, we model an eye coordinate frame and eye behaviors such as saccades and pursuits. In Section 4.2, we describe how to determine the point of sight at which the character is looking. In Section 4.3, we discuss how to estimate the state of a moving object with our vision system.

本节介绍我们的视觉系统。在第4.1节中，我们建立了眼睛坐标系和眼睛行为的模型，如眼跳和追踪。在第4.2节中，我们描述了如何确定角色正在寻找的视点。在第4.3节中，我们讨论了如何用我们的视觉系统来估计运动物体的状态。

**4.1 眼部模型**

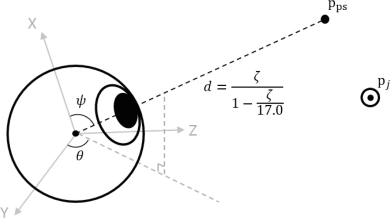
As shown in Figure [2](http://delivery.acm.org/10.1145/3370000/3360905/a3-eom.html?ip=122.205.5.129&id=3360905&acc=ACTIVE%20SERVICE&key=BF85BBA5741FDC6E%2ECC932049E1B2BA72%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35&__acm__=1574280850_ed68f5ad2990d4cd532b3ea91481679e#fig2), we adopt a spherical coordinate system to define the eye frame. In this frame, we use three parameters to specify an eye pose  ，Here θ and ψ are respectively the azimuthal and polar angles, which together represent a gaze direction, and ζ is the focal length. We bound each parameter with a normal range of eye movement that can be obtained from medical research [Serway et al. [2018](http://delivery.acm.org/10.1145/3370000/3360905/a3-eom.html?ip=122.205.5.129&id=3360905&acc=ACTIVE%20SERVICE&key=BF85BBA5741FDC6E%2ECC932049E1B2BA72%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35&__acm__=1574280850_ed68f5ad2990d4cd532b3ea91481679e#Bib0046); Shin et al. [2016](http://delivery.acm.org/10.1145/3370000/3360905/a3-eom.html?ip=122.205.5.129&id=3360905&acc=ACTIVE%20SERVICE&key=BF85BBA5741FDC6E%2ECC932049E1B2BA72%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35&__acm__=1574280850_ed68f5ad2990d4cd532b3ea91481679e#Bib0048)]: That is, −44.2∘≤θ≤44.2∘ for adduction and abduction, −47.1∘≤ψ≤27.7∘ for depression and elevation, and 0.0≤ζ≤17.0 (mm) for focal length. The focal length approaches 17.0mm when human looks at a point at infinity.

如图[2](http://delivery.acm.org/10.1145/3370000/3360905/a3-eom.html?ip=115.156.140.169&id=3360905&acc=ACTIVE%20SERVICE&key=BF85BBA5741FDC6E%2ECC932049E1B2BA72%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35&__acm__=1574148925_581c80d73e5154b164c4bbb251ae28a4" \l "fig2)所示，我们采用球坐标系来定义眼框。在此框架中，我们使用三个参数定义眼睛姿势，这里θ我和ψ分别是方位角和极角，它们共同代表凝视的方向，ζ 是焦距长度。我们用正常眼睛的移动范围（可以在医学研究上获取到）去界定每一个参数。其中：

内收和外展

凸显和隆起

其中17.0mm代表人看向无限远处的点时的焦距



**Fig. 2. Eye coordinate frame.**

We also bound eye movement to generate realistic gaze behaviors, saccades and pursuits, which are important characteristics of human eyes. Saccades and pursuits refer to rapid movements of eyeballs to find new objects and slow eyeball movements to track objects, respectively. As in Yeo et al. [[2012](http://delivery.acm.org/10.1145/3370000/3360905/a3-eom.html?ip=122.205.5.129&id=3360905&acc=ACTIVE%20SERVICE&key=BF85BBA5741FDC6E%2ECC932049E1B2BA72%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35&__acm__=1574280850_ed68f5ad2990d4cd532b3ea91481679e#Bib0063)], we adopt a simplified profile of saccades and pursuits and impose velocity constraints on these behaviors based on the results from Robinson et al. [[1965](http://delivery.acm.org/10.1145/3370000/3360905/a3-eom.html?ip=122.205.5.129&id=3360905&acc=ACTIVE%20SERVICE&key=BF85BBA5741FDC6E%2ECC932049E1B2BA72%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35&__acm__=1574280850_ed68f5ad2990d4cd532b3ea91481679e#Bib0042)], Meyer et al. [[1985](http://delivery.acm.org/10.1145/3370000/3360905/a3-eom.html?ip=122.205.5.129&id=3360905&acc=ACTIVE%20SERVICE&key=BF85BBA5741FDC6E%2ECC932049E1B2BA72%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35&__acm__=1574280850_ed68f5ad2990d4cd532b3ea91481679e#Bib0030)], Leigh and Zee [[2015](http://delivery.acm.org/10.1145/3370000/3360905/a3-eom.html?ip=122.205.5.129&id=3360905&acc=ACTIVE%20SERVICE&key=BF85BBA5741FDC6E%2ECC932049E1B2BA72%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35&__acm__=1574280850_ed68f5ad2990d4cd532b3ea91481679e#Bib0064)], and Itti et al. [[2006](http://delivery.acm.org/10.1145/3370000/3360905/a3-eom.html?ip=122.205.5.129&id=3360905&acc=ACTIVE%20SERVICE&key=BF85BBA5741FDC6E%2ECC932049E1B2BA72%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35&__acm__=1574280850_ed68f5ad2990d4cd532b3ea91481679e#Bib0018)]. Note that our system only uses velocity constraints for a natural gaze behavior and does not enforce explicit saccades or pursuits in relation to the object motion.

我们还通过眼睛的运动来产生真实的注视行为、扫视和追踪，这些都是人眼的重要特征。扫视和追踪分别是指眼球快速转动寻找新的物体以及眼球缓慢转动来追踪物体。借鉴Yeo等人所说的，[2012]，我们采用了一个对扫视和追逐的简化描述，并根据Robinson等人的研究结果对这些行为施加速度限制。注意，我们的系统只对自然凝视行为使用速度约束，不对于物体运动相关的明确的扫视或追踪行为进行（速度）限制。

In particular, we set the maximum saccade speed to 800∘/s and the maximum pursuit speed to 100∘/s considering only azimuthal and polar eyeball movements. Saccades repeat in every 200ms interval, each followed by a 200ms pause for recharge. Pursuits occur simultaneously even during the recharge times for saccades. We incorporate all of these in our vision system as follows:

where is the eyeball speed for azimuthal and polar movements and B is the upper bound of its magnitude.

特别地，只考虑方位角和极性眼球运动情况下，我们将最大扫视速度设定为800。/ s，最大跟踪速度为100。/s。每间隔200毫秒重复一次扫视，然后每间隔200毫秒暂停一次（进行休息）。即使在扫视的休息时间内，追逐也会同时发生。我们将所有这些都纳入我们的视觉系统，如下所示：

这里是眼球方位角和极角的移动速度，B是其大小的上限。

**4.3对象状态估计**

In this section, we discuss how to estimate the object state from a partial observation. To generate realistic gaze behaviors, we imitate the human vision system through estimating the object state under uncertainty instead of using the true full state of an object. First, we introduce an observation model that measures the object position from a true object state obtained through the physics-based simulation. Then, we discuss a process model that produces the predicted object state. Finally, we describe how to estimate the object state based on the belief update using a Kalman filter [Erez and Smart [**2012**](http://delivery.acm.org/10.1145/3370000/3360905/a3-eom.html?ip=122.205.5.129&id=3360905&acc=ACTIVE%20SERVICE&key=BF85BBA5741FDC6E%2ECC932049E1B2BA72%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35&__acm__=1574365151_2d97624c811c1d0a1b336e7c6e298cfa#Bib0008)].

在本节中，我们将讨论如何从部分观测值估计对象状态。为了产生真实的凝视行为，我们通过估计不确定条件下的对象状态来模拟人类视觉系统，而不是使用对象的真实完整状态。首先，我们介绍了一个观测模型，该模型通过基于物理的模拟获得真实的物体状态来测量物体的位置。然后，我们讨论了产生预测对象状态的过程模型。最后，我们描述了如何使用Kalman滤波器估计基于信念更新的目标状态[Erez和Smart 2012]。

Under the assumption that the character does not know the true state, our system makes an observation from the true state of an object. To construct an observation model that abstracts human visual sensors, we assume that humans perceive the positional information of a moving object to estimate its velocity instantaneously. Moreover, the positional information is not accurate in general, particularly when the object is distant from the point of sight. Under these assumptions, we formulate the observation model of our vision system as follows:

假设角色不知道真实的状态，我们的系统从对象的真实状态进行观察。为了抽象出人类视觉传感器的观测模型，我们假设人类（能）感知到运动物体的位置信息，从而能瞬时估计其速度。此外，位置信息通常不准确，特别是当物体远离视线时。在这些假设下，我们的视觉系统的观察模型如下：

Let  be the number of objects in the environment that are perceived by the vision system. Our observation model makes partial observation on the state of an object from its true state  that is composed of positon and velocity , considering observation noise ρρ, for  . H is the transformation matrix that maps the object state onto a noise-free partial observation, that is, the position of an object. Observation noise ρ has a multivariate Gaussian distribution with covariance Rj. η is the constant for the size of the fovea. Covariance matrix Rj is a diagonal matrix, where the value of diagonal elements depends on the Euclidean distance between the point of sight pps and the object position pj; specifically, the error of the observed object position increases as the character's visual attention becomes farther from the object [Erez and Smart [**2012**](http://delivery.acm.org/10.1145/3370000/3360905/a3-eom.html?ip=122.205.5.129&id=3360905&acc=ACTIVE%20SERVICE&key=BF85BBA5741FDC6E%2ECC932049E1B2BA72%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35&__acm__=1574365151_2d97624c811c1d0a1b336e7c6e298cfa#Bib0008)].

令为我们视觉系统中感知到的环境中对象的个数。考虑到观测噪声，对于j =1到，我们的观测模型，基于物体真实状态（含有噪声）中位置和速度组成的状态构成了观测（数据）。其中H是将对象状态映射到无噪声部分的观测值（即对象的位置）的变换矩阵。观测噪声ρ服从多元协方差高斯分布，η是中心凹大小的常数，是一个对角矩阵，其中对角元素的值取决于视点之间的欧几里得距离和物体位置; 具体地说，观察到的物体位置的误差随着人物的视觉注意力离物体越来越远而增加。

Under the assumption that the brain has prior knowledge on the dynamics of the object [McIntyre et al. [**2001**](http://delivery.acm.org/10.1145/3370000/3360905/a3-eom.html?ip=122.205.5.129&id=3360905&acc=ACTIVE%20SERVICE&key=BF85BBA5741FDC6E%2ECC932049E1B2BA72%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35&__acm__=1574365151_2d97624c811c1d0a1b336e7c6e298cfa#Bib0029)], we now formulate the process model that deals with the evolution of the object state as follows:

假设大脑对物体的动力学有先验知识[McIntyre等人。2001年），我们现在制定了处理对象状态演变的过程模型，如下所示：

