

# 沉浸式虚拟现实系统中头部位姿预测方法

魏青<sup>1,2</sup>, 刘钊铭<sup>1,2</sup>, 刘乃龙<sup>1,2</sup>, 崔龙<sup>1</sup>, 张峰<sup>1</sup>

1. 中国科学院沈阳自动化研究所机器人学国家重点实验室, 辽宁 沈阳 110016; 2. 中国科学院大学, 北京 100049

基金项目: 科技部磁约束核聚变专项(2012GB102005)

通信作者: 魏青, weiqing@sia.cn 收稿/录用/修回: 2017-12-26/2018-04-02/2018-07-17

## 摘要

为了最小化运动画面延时对沉浸式虚拟现实系统的影响, 提出了一种具有运动画面延时抖动补偿功能的卡尔曼滤波算法, 用于对用户头部位姿进行预测. 首先对运动画面延时进行分析和预测, 然后根据头部运动相关性分析结果选择位姿预测方式, 并在头部运动建模和位姿预测时分别利用延时预测结果对延时抖动进行补偿. 实验结果表明, 相比设备原有预测方法, 其在一般情况下的姿态预测绝对误差的均值分别减小 45.74%、47.25% 和 40.96%, 最大值分别减小 11.49%、26.34% 和 36.79%; 位置预测绝对误差的均值分别减小 35.94%、45.90% 和 55.81%, 最大值分别减小 1.05%、25.60% 和 44.74%.

## 关键词

虚拟现实  
运动画面延时预测  
头部位姿预测  
卡尔曼滤波  
中图法分类号: TP37  
文献标识码: A

## Head Pose Prediction Method for Immersive Virtual Reality Systems

WEI Qing<sup>1,2</sup>, LIU Zhaoming<sup>1,2</sup>, LIU Nailong<sup>1,2</sup>, CUI Long<sup>1</sup>, ZHANG Feng<sup>1</sup>

1. State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China;

2. University of Chinese Academy of Sciences, Beijing 100049, China

## Abstract

To minimize the impact of motion-to-photo latency, we propose a Kalman filter algorithm with motion-to-photon latency jitter compensation to predict the user's head pose. First, we analyze and predict the motion-to-photon latency. Second, we select a head pose prediction method according to the correlation analysis results of the head motion. Further, we use the latency prediction result to compensate for the delay jitter in the head motion modeling and pose prediction. Compared to the original prediction method for the equipment, the average of the absolute error of attitude prediction decreases by 45.74%, 47.25%, and 40.96%. The maximum decreases by 11.49%, 26.34%, and 36.79. Moreover, the average of the absolute error of the position prediction decreases by 35.94%, 45.90%, and 55.81% and the maximum decreases by 1.05%, 25.60% and 44.74%.

## Keywords

virtual reality;  
motion-to-photon latency  
prediction;  
head pose prediction;  
Kalman filter

## 0 引言

作为解决遥操作时延问题最具潜力的方法, 虚拟现实技术目前已经在遥操作机器人预仿真和辅助操作等领域发挥了重要应用<sup>[1-3]</sup>. 沉浸式虚拟现实系统利用具有大视场角的头盔显示器(helmet-mounted display, HMD)来欺骗大脑, 从而让用户具有较好的沉浸感<sup>[4]</sup>. 然而, 为了更好的用户体验, 虚拟现实系统还需要能够实时响应用户的头部位姿<sup>[5]</sup>. 虽然用户的头部运动可以通过多种传感器进行测量, 但是受传输和计算等时间的限制, 头盔显示器并不能立即把相应画面显示给用户, 从而会产生运动画面延时(motion-to-photo latency, 即从开始采集用户头部位姿到显示出相应画面的延时), 如图1所示. 该延时会造成视

觉运动感知与前庭运动感知的差异, 不仅会降低虚拟现实系统的临场感, 还会加重大脑负担并引起眩晕与恶心<sup>[6-9]</sup>.

到目前为止, 运动画面延时是造成虚拟现实系统眩晕的最主要原因<sup>[10]</sup>. 运动画面延时虽然可以通过对各环节进行优化而减小, 但受处理顺序的限制不可能完全消除. 因此, 对用户头部运动进行预测并根据预测的画面显示时的头部位姿进行渲染是最小化延时影响的有效途径. 扩展卡尔曼滤波(EKF)、无迹卡尔曼滤波(UKF)和粒子滤波(PFs)是常用的时延补偿预测滤波算法, 其中EKF在头部位姿预测时具有一定的优势<sup>[11-12]</sup>. Himberg等提出了一种基于旋转四元数增量的头部姿态预测方法<sup>[13]</sup>, 其利用EKF和四元数增量对旋转速度进行估计, 进而预测用户头部姿态, 该方法的预测精度与基于四元数的预测方法相差

不大,但通过对旋转运动的解耦减小了计算量。Himberg 还提出了一种基于测量协方差自适应辨识的 R-Adaptive EKF<sup>[14]</sup>,该方法可以降低预测结果中的噪声,但在剧烈运动时会产生较大的预测误差。Drouard<sup>[15]</sup>和 Kang<sup>[16]</sup>都提出了基于图像处理的头部姿态估计方法,这些方法需要的设备简单、鲁棒性高,不过当应用到 HMD 上时,其精度和实时性还需要提高。Choi 等采用线性回归方法对头部姿态的欧拉角进行多步预测<sup>[17]</sup>,该方法的算法简单且可以补偿较大的时延,但是该方法进行长程预测时会有较大误差且易引起死锁现象。Horn 等提出了一种基于动力学模型的头部姿态预测方法<sup>[18]</sup>,并将其应用到了人形机器人控制上,该方法把人体躯干和头部的连接等效为弹性连接,并利用最小二乘来辨识模型参数,不过其只考虑了对平面旋转运动的预测。除有效性和精度不足外,当应用到实际 HMD 上时,现有预测算法还缺乏对两个问题的分析和考虑:

1) 受处理器负载等多种因素的影响,运动画面延时总是变化的;

2) 人体骨骼肌肉模型比较复杂,头部运动会受到一定的运动学约束<sup>[10]</sup>,即头部在各个轴向上的运动间可能会存在耦合,而仅当头部在不同轴向上的运动相互独立时,才适合对运动进行解耦预测以获得较低的算法复杂性。

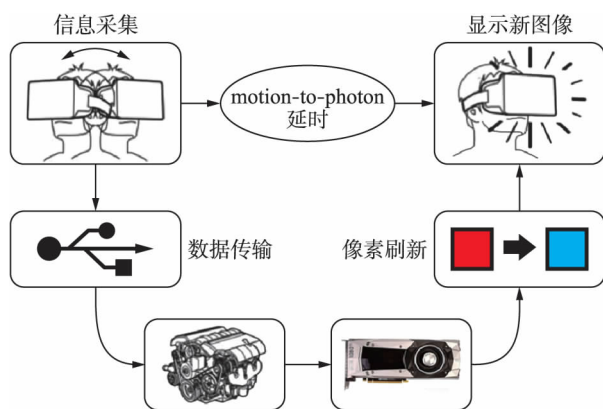


图1 运动画面延时产生过程

Fig.1 Motion-to-photon latency generation process

基于此,本文提出了一种具有延时抖动补偿功能的卡尔曼滤波算法对用户头部位姿进行预测。相比现有预测方法,所提算法增加了对运动画面延时和头部运动相关性的分析与运用,并在头部运动建模和位姿预测时分别对延时抖动进行了补偿。实验结果表明,该算法在多种运动模式下均可显著提高头部位姿的预测精度。

## 1 算法

### 1.1 运动画面延时的分析与预测

对于 HMD,每一帧的运动画面延时都由多个延时环节构成,而每个延时环节又受处理器负载、总线负载以及环境等多种因素的影响,因此运动画面延时的构成比较复杂。图2是从 Oculus DK2 上采集到的运动画面延时数据及其概率分布(Oculus DK2 的屏幕刷新频率为 75 Hz,一个垂

直同步周期大约 13.3 ms,当渲染管线经过优化后,从开始位姿采集到显示出相应画面需要两个垂直同步周期)。

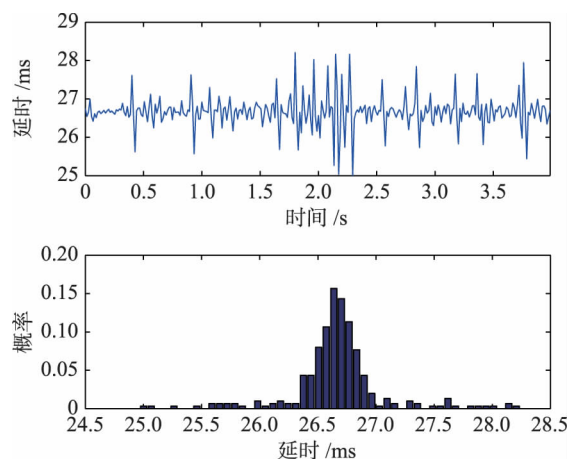


图2 运动画面延时及其概率分布

Fig.2 Motion-to-photon latency and its probability distribution

选择长度为 10 的滑动窗口分析当前延时与历史延时值的相关性,结果如图3所示。

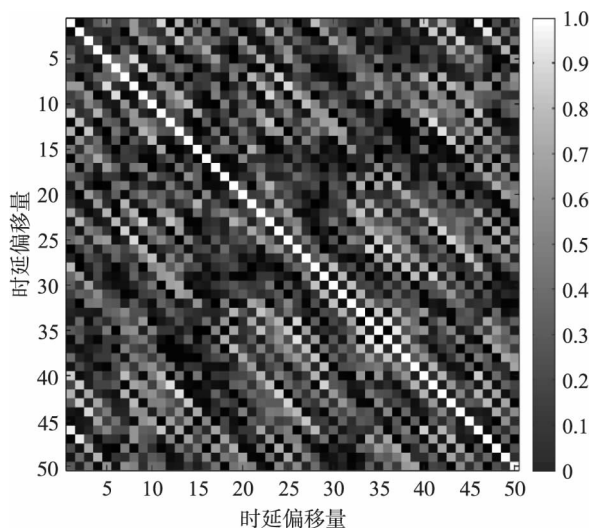


图3 运动画面延时的相关性图

Fig.3 Correlation diagram of motion-to-photon latency

从相关性分析结果可以看出,当前的运动画面延时不仅与距离其较近延时相关,还可能与距离其较远的延时值有较大的相关性且规律性较弱。基于此,本文采用基于稀疏矩阵的多元线性回归算法<sup>[19-20]</sup>对其进行预测,该算法充分利用了数据间的相关性,只选择和当前数据相关性足够大的历史数据来预测当前数据,不仅降低了回归算法的时间复杂性,还避免了数据的过拟合。

假设已测得的运动画面延时序列为  $x_1, x_2, \dots, x_n$ , 首先选定滑动窗口的长度  $m$  ( $m < n$  且延时抖动越剧烈,窗口的长度越短),并构造回归向量  $X_i = [x_{i-m+1}, x_{i-m+2}, \dots, x_i]^T$  和完全矩阵  $X = [X_m, X_{m+1}, \dots, X_{n-1}]$ . 然后计算矩阵  $X$  中的所有列与  $X_n$  的相关系数,并选择与其相关系数

最大的  $k$  列构成回归矩阵  $R = [\mathbf{1}_{m \times 1}, X_p, \dots, X_q, \dots, X_r]_{m \times (k+1)}$ ,  $k$  的选择方法与  $m$  类似.

定义回归方程为  $X_n = b_0 R_1 + b_1 R_2 + \dots + b_k R_{k+1}$ , 其中  $b_i$  为回归系数, 并根据回归矩阵和回归向量计算回归系数的最小二乘估计值  $\hat{B}$ :

$$\hat{B} = \begin{bmatrix} \hat{b}_0 \\ \hat{b}_1 \\ \vdots \\ \hat{b}_k \end{bmatrix} = (R^T R)^{-1} R^T X_n \quad (1)$$

最后, 利用前面构建的回归向量构造新的回归矩阵  $R' = [\mathbf{1}_{m \times 1}, X_{p+1}, \dots, X_{q+1}, \dots, X_{r+1}]_{m \times (k+1)}$ . 延时的预测过程为  $\hat{X}_{n+1} = R' \hat{B}$ ,  $\hat{X}_{n+1} = [\hat{x}_{n-m+2}, \hat{x}_{n-m+3}, \dots, \hat{x}_{n+1}]_{m \times 1}^T$ . 在实际计算中, 考虑到矩阵求逆的时间复杂度较高, 本文通过求解方程组  $(R^T R) \hat{B} = R^T X_n$  来计算回归系数的估计值.

## 1.2 头部运动相关性分析

一般情况下, 当人期望向某一个方向运动时, 经常是同时执行几个方向的运动进行运动合成 (比如当你想看低一点的东西时经常是探一下身体, 而不是直接下蹲), 因此头部在部分轴向上的运动间会存在耦合. 首先对采集到的用户头部运动数据进行相关性分析, 结果如图 5 所示 (坐标系如图 4 所示<sup>[10]</sup>,  $p_x, p_y, p_z$  分别代表  $x, y, z$  方向的平动,  $R_x, R_y, R_z$  分别代表  $x, y, z$  方向的转动. 为了便于直观显示, 以下的图形和数字结果中都采用固定角表示旋转).

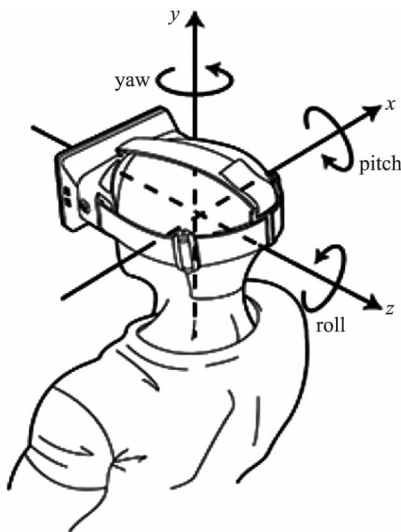


图4 Oculus DK2 坐标系

Fig.4 Coordinate system of Oculus DK2

从图 5 可以看出, 头部在部分轴向上的运动间确实存在一定的耦合 (比如当想向前移动时, 经常是向前探一下身体, 同时头部上转, 眼睛高度也会下降, 因此  $p_z$  与  $R_x$  和  $p_y$  具有较强的相关性; 当低头看时, 头部下转的同时眼睛位置也在下降, 因而  $R_x$  与  $p_y$  也具有明显的相关性; 当想观察左边或右边物体时, 左右平动时躯干和颈部也常会跟着

转动, 但竖直方向的移动一般不大, 因此  $p_x$  与  $R_y$  具有明显的相关性, 而与  $p_y$  的相关性较弱), 因此更适合对头部运动进行整体预测, 而不是对各个轴向的分量单独考量.

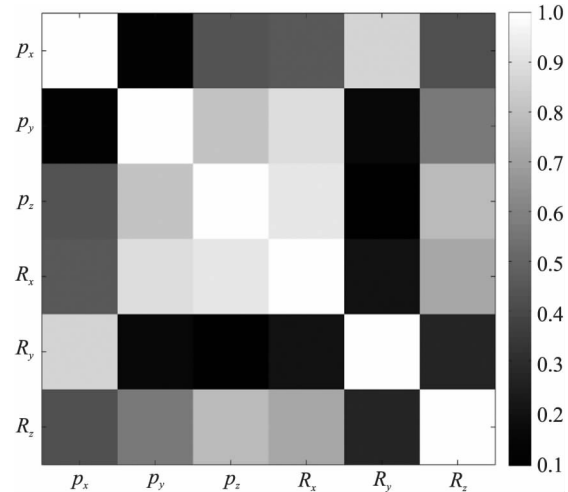


图5 头部各方向运动的相关性图

Fig.5 Correlation diagram of head motion in different directions

## 1.3 头部运动建模与预测

由于头部在部分方向上的运动间存在耦合, 因此更适合对其位姿进行整体预测, 定义系统状态  $x = [x_1 \ x_2]^T$ , 其中, 姿态分量  $x_1 = [q \ \omega]$ ,  $q$  为头部姿态的四元数,  $\omega$  为旋转角速度; 位置分量  $x_2 = [p \ v]$ ,  $p$  为头部位置,  $v$  为头部线速度, 则用户头部位姿演变过程可表示为

$$\begin{aligned} x(k) &= Ax(k-1) + W(k-1) \\ z(k) &= Hx(k) + V(k) \end{aligned} \quad (2)$$

其中,  $W(k)$  为角加速度和线加速度引起的过程噪声,

$V(k)$  为测量噪声,  $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ ,  $H = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}$ . 根据旋转四元数的定义, 头部姿态的演变过程可表示为

$$q(k) = \Delta q(\omega(k), \tau(k)) \cdot q(k-1) \quad (3)$$

而头部位置的变化过程可表示为

$$p(k) = \Delta p(v(k), \tau(k)) + p(k-1) \quad (4)$$

其中  $\tau(k)$  为第  $k$  帧运动画面延时的预测值, 用于对延时抖动进行补偿. 在固定速度模型下, 有:

$$\Delta q(\omega(k), \tau(k)) = \begin{bmatrix} \cos \frac{\theta}{2} \\ \frac{\omega_0(k)}{\|\omega(k)\|} \cdot \sin \frac{\theta}{2} \\ \frac{\omega_1(k)}{\|\omega(k)\|} \cdot \sin \frac{\theta}{2} \\ \frac{\omega_2(k)}{\|\omega(k)\|} \cdot \sin \frac{\theta}{2} \end{bmatrix} \quad (5)$$

其中,  $\|\omega(k)\| = \sqrt{\omega^T(k) \cdot \omega(k)}$ ,  $\theta = \tau(k) \cdot \sqrt{\omega^T(k) \cdot \omega(k)}$ , 分别为速度大小和角度增量. 而头部位置增量可表示为

$$\Delta p(v(k), \tau(k)) = v(k) \cdot \tau(k) \quad (6)$$

下面计算模型参数:

$$A_{12} = \frac{\partial x_1(k)}{\partial x_2(k-1)} = \frac{\partial (q(k) \omega(k))}{\partial (p(k-1) v(k-1))} = 0 \quad (7)$$

$$A_{11} = \frac{\partial \mathbf{x}_1(k)}{\partial \mathbf{x}_1(k-1)} = \begin{bmatrix} \frac{\partial \mathbf{q}(k)}{\partial \mathbf{q}(k-1)} & \frac{\partial \mathbf{q}(k)}{\partial \boldsymbol{\omega}(k-1)} \\ \frac{\partial \boldsymbol{\omega}(k)}{\partial \mathbf{q}(k-1)} & \frac{\partial \boldsymbol{\omega}(k)}{\partial \boldsymbol{\omega}(k-1)} \end{bmatrix} \quad (8)$$

结合式(3)可得:

$$A_{11} = \begin{bmatrix} \Delta \mathbf{q}(\boldsymbol{\omega}(k), \tau(k)) & \frac{\partial \mathbf{q}(k)}{\partial \boldsymbol{\omega}(k-1)} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \quad (9)$$

其中,

$$\frac{\partial \Delta \mathbf{q}(\boldsymbol{\omega}(k), \tau(k))}{\partial \boldsymbol{\omega}_i(k-1)} = \begin{bmatrix} -\frac{\tau}{2} \cdot \Delta \mathbf{q}_{i+1} \\ \frac{\boldsymbol{\omega}_0}{2 \|\boldsymbol{\omega}\|^2} (\tau \cdot \boldsymbol{\omega}_i \cdot \Delta \mathbf{q}_0) - \frac{\boldsymbol{\omega}_0}{\|\boldsymbol{\omega}\|^2} \cdot \Delta \mathbf{q}_{i+1} + \delta_{i,0} \cdot \frac{\Delta \mathbf{q}_{i+1}}{\boldsymbol{\omega}_i} \\ \frac{\boldsymbol{\omega}_1}{2 \|\boldsymbol{\omega}\|^2} (\tau \cdot \boldsymbol{\omega}_i \cdot \Delta \mathbf{q}_0) - \frac{\boldsymbol{\omega}_1}{\|\boldsymbol{\omega}\|^2} \cdot \Delta \mathbf{q}_{i+1} + \delta_{i,1} \cdot \frac{\Delta \mathbf{q}_{i+1}}{\boldsymbol{\omega}_i} \\ \frac{\boldsymbol{\omega}_2}{2 \|\boldsymbol{\omega}\|^2} (\tau \cdot \boldsymbol{\omega}_i \cdot \Delta \mathbf{q}_0) - \frac{\boldsymbol{\omega}_2}{\|\boldsymbol{\omega}\|^2} \cdot \Delta \mathbf{q}_{i+1} + \delta_{i,2} \cdot \frac{\Delta \mathbf{q}_{i+1}}{\boldsymbol{\omega}_i} \end{bmatrix} \quad (11)$$

其中, 角速度  $\boldsymbol{\omega} = \boldsymbol{\omega}(k)$ ,  $\Delta \mathbf{q} = \Delta \mathbf{q}(\boldsymbol{\omega}(k), \tau(k))$ , 步长  $\tau = \tau(k)$ ,  $\delta = \mathbf{I}_{3 \times 3}$ .

$$A_{21} = \frac{\partial \mathbf{x}_2(k)}{\partial \mathbf{x}_1(k-1)} = \frac{\partial [\mathbf{p}(k) \mathbf{v}(k)]}{\partial [\mathbf{q}(k-1) \boldsymbol{\omega}(k-1)]} = \mathbf{0} \quad (12)$$

$$A_{22} = \frac{\partial \mathbf{x}_2(k)}{\partial \mathbf{x}_2(k-1)} = \begin{bmatrix} \frac{\partial \mathbf{p}(k)}{\partial \mathbf{p}(k-1)} & \frac{\partial \mathbf{p}(k)}{\partial \mathbf{v}(k-1)} \\ \frac{\partial \mathbf{v}(k)}{\partial \mathbf{p}(k-1)} & \frac{\partial \mathbf{v}(k)}{\partial \mathbf{v}(k-1)} \end{bmatrix} \quad (13)$$

设  $\mathbf{a}(k)$  为过程噪声中的线加速度部分, 则线速度预测结果为  $\mathbf{v}(k) = \mathbf{v}(k-1) + \mathbf{a}(k-1) \cdot \tau(k-1)$ , 然后结合式(4)和式(6)可得:

$$\mathbf{p}(k) = (\mathbf{v}(k-1) + \mathbf{a}(k-1) \cdot \tau(k-1)) \cdot \tau(k) + \mathbf{p}(k-1) \quad (14)$$

根据式(13)和式(14)可得:

$$A_{22} = \begin{bmatrix} \mathbf{I} & \tau(k) \cdot \mathbf{I} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \quad (15)$$

由于 HMD 直接测量到的是头部的旋转四元数和位置  $\mathbf{z}(k) = \begin{bmatrix} \mathbf{q}(k) \\ \mathbf{p}(k) \end{bmatrix} + \mathbf{V}(k)$ , 因此有  $H_{12} = \mathbf{0}$ ,  $H_{21} = \mathbf{0}$ . 进一步可以得到:

$$H_{11} = \frac{\partial \mathbf{z}(k)}{\partial \mathbf{x}(k)} = \begin{bmatrix} \frac{\partial \mathbf{q}(k)}{\partial \mathbf{q}(k)} & \frac{\partial \mathbf{q}(k)}{\partial \boldsymbol{\omega}(k)} \\ \frac{\partial \mathbf{p}(k)}{\partial \mathbf{q}(k)} & \frac{\partial \mathbf{p}(k)}{\partial \boldsymbol{\omega}(k)} \end{bmatrix} = [\mathbf{I} \quad \mathbf{0}] \quad (16)$$

$$H_{22} = \frac{\partial \mathbf{z}(k)}{\partial \mathbf{x}(k)} = \begin{bmatrix} \frac{\partial \mathbf{p}(k)}{\partial \mathbf{p}(k)} & \frac{\partial \mathbf{p}(k)}{\partial \mathbf{v}(k)} \\ \frac{\partial \mathbf{v}(k)}{\partial \mathbf{p}(k)} & \frac{\partial \mathbf{v}(k)}{\partial \mathbf{v}(k)} \end{bmatrix} = [\mathbf{I} \quad \mathbf{0}] \quad (17)$$

当利用卡尔曼滤波器对用户头部位姿进行预测时, 设  $\mathbf{W}(k)$  和  $\mathbf{V}(k)$  的协方差 (covariance) 分别为  $\mathbf{Q}$  和  $\mathbf{R}$ , 由于  $\mathbf{Q}$  是由加速度引起的协方差, 因此其对运动画面延时变化比较敏感, 利用延时预测结果对其进行自适应调整:

$$\mathbf{Q}_{\text{adj}} = \begin{bmatrix} \mathbf{Q}_{\text{adj}1} \\ \mathbf{Q}_{\text{adj}2} \end{bmatrix} \quad (18)$$

其中,  $\mathbf{Q}_{\text{adj}i} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \frac{f_v \cdot \tau(k)}{2} \mathbf{I} \end{bmatrix} \mathbf{Q}_i$ ,  $i = 1, 2$ ;  $f_v$  为虚拟现实设备的垂直同步频率. 当对用户头部运动进行建模时, 利用

$$\frac{\partial \mathbf{q}(k)}{\partial \boldsymbol{\omega}(k-1)} = \begin{bmatrix} \frac{\partial \mathbf{q}(k)}{\partial \boldsymbol{\omega}_0(k-1)} & \frac{\partial \mathbf{q}(k)}{\partial \boldsymbol{\omega}_1(k-1)} & \frac{\partial \mathbf{q}(k)}{\partial \boldsymbol{\omega}_2(k-1)} \end{bmatrix}$$

根据式(3)有:

$$\frac{\partial \mathbf{q}(k)}{\partial \boldsymbol{\omega}(k-1)}_i = \frac{\partial \Delta \mathbf{q}(\boldsymbol{\omega}(k), \tau(k))}{\partial \boldsymbol{\omega}(k-1)}_i \cdot \mathbf{q}(k-1) \quad (10)$$

设  $\mathbf{w}(k)$  是旋转加速度引起的过程噪声, 则速度预测结果为  $\boldsymbol{\omega}(k) = \boldsymbol{\omega}(k-1) + \mathbf{w}(k-1) \cdot \tau(k-1)$ , 对式(10)的系数进行求解可得:

运动画面延时预测结果来直接修正头部位姿预测结果不同, 该部分主要用来修正头部角速度和线速度的预测结果, 从而修正下一步的头部位姿预测结果.

当利用文[10]中给出的人体简化模型进行头部位置预测时, 需要先根据实测的位姿数据在线辨识出人体简化模型的参数, 然后利用辨识出的模型和姿态预测结果来预测用户头部位置 (该法在文[10]中已有介绍, 且其 C++ 源码可通过 <https://developer.oculusvr.com/> 下载, 在此不再累述).

## 2 实验与分析

本文构建的实验平台包括一台式计算机 (CPU: Intel Core i5-4590, GPU: Nvidia GeForce GTX 970) 和一台 HMD (Oculus DK2, 如图 6 所示). 实验场景为遥操作机器人更换聚变堆真空室内隔热瓦块的仿真场景, 如图 7 所示 (开发环境为 VC++ 2010, 程序中利用着色程序对显示画面进行了预处理, 用以补偿镜片畸变和颜色散射).



图6 Oculus DK2

Fig.6 Oculus DK2

用户佩戴 HMD, 并在虚拟场景中执行隔热瓦片检查及更换任务. 根据前期调试经验, 选择延时预测参数为  $n = 50$ ,  $m = 10$ ,  $k = 5$ , 同时记录运动画面延时的实际值和预测值, 以及用户头部位姿的实际值和预测值.

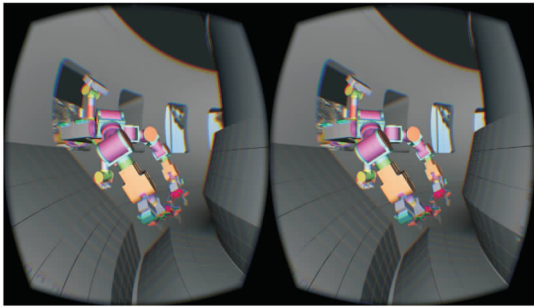


图7 虚拟现实场景  
Fig.7 The virtual reality scene

2.1 运动画面延时的预测结果

直接采集的运动画面延时的实际值如图 2 所示,运动画面延时的预测结果如图 8 所示(真实的运动画面延时误差为相邻两帧运动画面延时的差,预测误差为运动画面延时的实测值与预测值的差)。

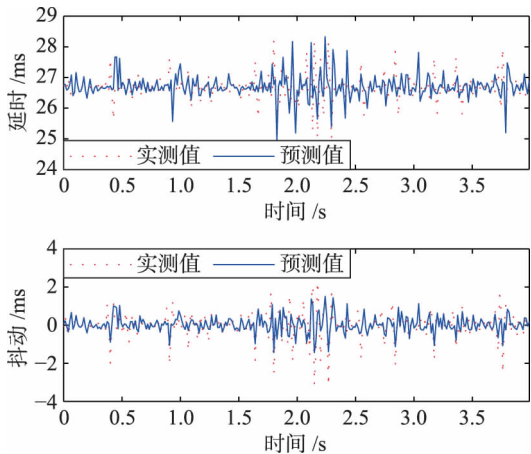


图8 运动画面延时的预测结果  
Fig.8 Prediction results of motion to photon delay

当不对运动画面延时进行预测时,运动画面延时绝对误差的均值为 0.328 ms,最大值为 3.164 ms. 利用本文所提方法对运动画面延时进行预测和补偿后,其绝对误差的最大值下降到 1.531 7 ms,减小了 51.62%;绝对误差的均值下降到 0.244 1 ms,减小了 25.65%。

从图 8 中还可以看出,虽然运动画面延时的预测结果在大部分情况下都可对运动画面延时的抖动进行补偿,但在个别时刻仍会出现较大的预测误差. 这是因为运动画面延时的构成比较复杂,其本身具有一定的随机性,一般对随机量的预测只能实现期望上的最优,而不能保证每个值都准确,因此预测结果中会出现个别预测误差较大的情况.

2.2 头部姿态的预测结果

从 HMD 上采集到的用户头部姿态数据和姿态的预测误差分别如图 9 和图 10 所示. 图 10 中 np 为不加预测时运动画面延时造成的姿态误差,ca 为文[10]给出的设备自带预测算法的预测误差,kf 为本文所提算法的预测误差,其对应的数值统计结果如表 1 所示.

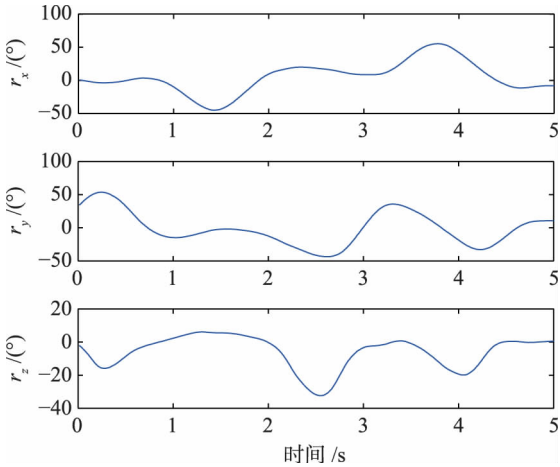


图9 从 Oculus DK2 上采集的头部姿态数据  
Fig.9 Head posture data collected from Oculus DK2

从图 9 可以看出,用户头部在 3 个轴向上的最大转动速度分别达到了 130°/s、199°/s 和 107°/s,比一般情况下人在观察物体时的头部转动速度还要大.

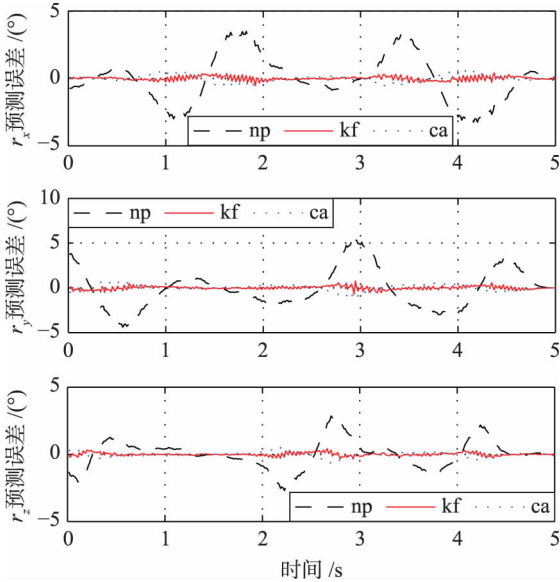


图10 头部姿态的预测误差  
Fig.10 Prediction errors of head posture

表 1 头部姿态预测误差统计表			
Tab.1 Statistical table of the head posture prediction errors			
方法	X 向误差 /(°) 均值/最大值	Y 向误差 /(°) 均值/最大值	Z 向误差 /(°) 均值/最大值
np	1.350/3.488	1.778/5.314	0.796/2.874
ca	0.208/0.617	0.274/1.044	0.133/0.669
kf	0.113/0.546	0.144/0.769	0.079/0.423

从表 1 中可以看出,相比文[10]给出的预测方法,本文所提算法可显著提高用户头部姿态的预测精度,预测误差的绝对平均值分别减小 45.74%、47.25% 和 40.96%,最大值分别减小 11.49%、26.34% 和 36.79,对误差均值的优



化程度要好于最大值。

### 2.3 头部位姿的预测结果

从 HMD 上采集到的用户头部位姿数据和位置的预测误差分别如图 11 和图 12 所示。图 12 中 np 为不加预测时运动画面延时造成的位置误差, ca 为文 [10] 给出的设备自带算法的位置预测误差, kf 为本文所提算法的预测误差, 其对应的数值统计结果如表 2 所示。

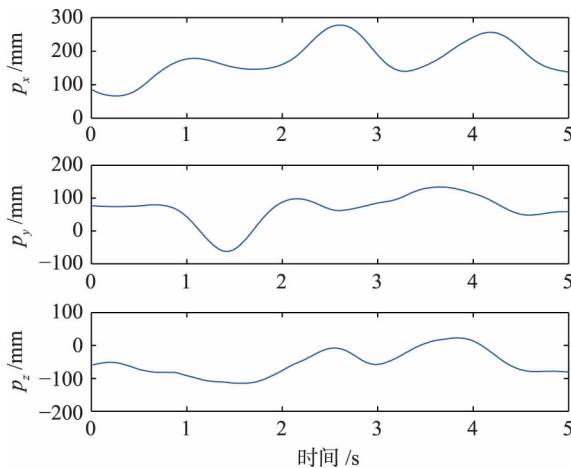


图 11 从 Oculus DK2 上采集的头部位姿数据

Fig.11 Head position data collected from Oculus DK2

从图 11 可以看出, 用户头部在 3 个方向上的最大移动速度分别达到了 363.6 mm/s、387.5 mm/s 和 233.8 mm/s, 要大于一般情况下人在观察物体时的头部移动速度。

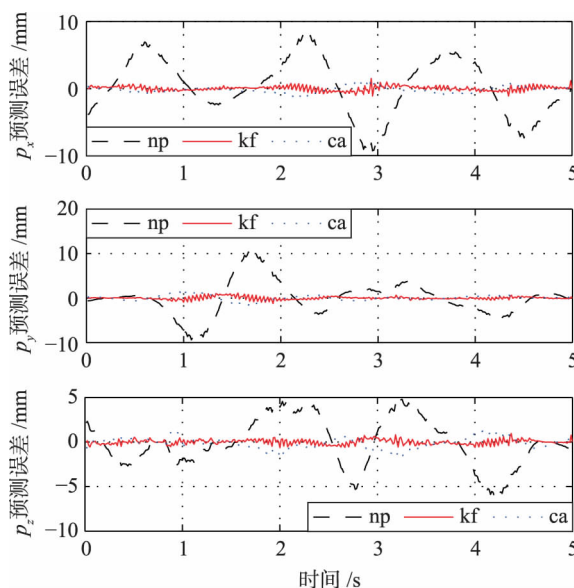


图 12 头部位姿的预测误差

Fig.12 Prediction errors of head position

从表 2 中可以看出, 相比本文提出的预测算法, 文 [10] 中给出的基于人体运动学约束的预测方法并没有提高位置预测的精度, 笔者认为这与人体模型比较复杂有关, 人的颈部和背部均为复杂的柔性多关节机构, 用刚性

连杆简化本来就会存在误差, 因此基于该模型进行位置预测时也会存在一定的预测误差。相比文 [10] 给出的设备自带预测方法, 本文所提算法位置预测误差的绝对平均值分别减小 35.94%、45.90% 和 55.81%, 最大值分别减小 1.05%、25.60% 和 44.74%, 与姿态预测结果类似, 本文所提算法对误差平均值的优化程度也要好于最大值。

表 2 头部位姿预测误差统计表

Tab.2 Statistical table of the head position prediction errors

方法	X 向误差 /mm 均值/最大值	Y 向误差 /mm 均值/最大值	Z 向误差 /mm 均值/最大值
np	3.557/9.708	2.741/10.345	2.225/6.244
ca	0.442/1.516	0.442/1.643	0.493/1.714
kf	0.283/1.500	0.239/1.223	0.218/0.947

### 2.4 剧烈运动下的姿态预测结果

在前述实验场景下, 不执行任何操作任务, 仅是以最快的速度沿 Y 轴转动头部, 此时头部运动状态和对头部姿态的预测结果如图 13 所示。

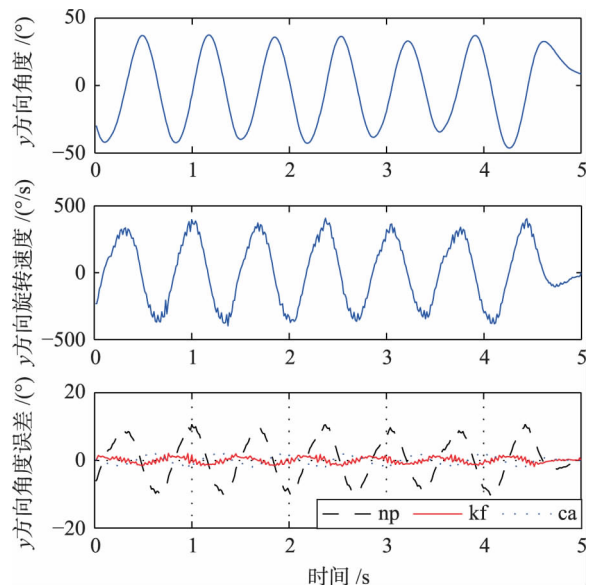


图 13 剧烈运动下的姿态预测结果

Fig.13 Posture prediction results under violent motion

从图 13 中可以看出, 此时头部沿 Y 轴的最大转动速度达到了  $400^\circ/\text{s}$ , 且最大加速度超过了  $2\,000^\circ/\text{s}^2$ , 此时对姿态预测误差的统计如表 3 所示。

表 3 剧烈运动下的姿态预测误差

Tab.3 Posture prediction errors under violent motion

误差 / (°)	np	ca	kf
均值	5.497	1.024	0.767
最大值	10.626	2.614	2.028

从表 3 中可以看出, 当头部做剧烈运动时, 所提算法仍可以显著提高位姿预测的精度, 相比文 [10] 给出的设备自带预测方法, 本文所提算法姿态预测误差的绝对平均值

减小了 25.14%，最大值减小了 22.42%；相比无预测补偿的情况，误差更是减小了 80% 以上。

## 2.5 实验结果分析

现有头部位姿预测算法基本都是认为运动画面延时是固定的，但是受处理器负载和系统实时性等多种因素的影响，这种理想状态在实际 HMD 上往往难以实现，因此本文在对头部运动进行建模和预测时都利用延时预测结果对延时抖动进行了补偿。在建模阶段，通过对延时抖动进行补偿可降低非等间隔采样的影响，从而直接提高头部位姿的预测精度。而在预测阶段，通过对延时抖动进行补偿可提高对头部运动速度的预测精度，从而提高算法鲁棒性并间接提高头部位姿的预测精度。

虽然人类的头部运动具有很大的随意性，但是受反应时间和骨骼肌肉强度限制，人体无法在极短的时间内连续改变运动方向，本文所提算法在对头部位姿进行预测时不仅考虑了头部的移动和旋转速度，还利用了头部运动加速度（有时会很大）和运动画面延时预测结果对位姿预测结果进行修正，因此当头部进行连续运动时会有较好的预测效果。另一方面，当在一个延时周期内头部运动方向刚好发生改变时，此时虽然可能有较大的加速度，但由于采样周期很短，因此头部运动速度和位姿在一个周期内并不会发生很大的改变，再加上有加速度的超前修正，故位姿预测误差也不会太大。而且在体验沉浸式虚拟现实系统时还发现，人类对持续的抖动比较敏感，偶尔一次的抖动并不会引起用户的眩晕与恶心。

当利用本文算法对头部位姿进行预测时，对误差均值的优化程度要好于最大值，笔者认为这与运动画面延时的预测结果有关，对运动画面延时的预测只能实现期望上的最优且个别时刻仍会出现预测误差较大的情况，而当这些时刻头部运动速度也刚好较大时便会体现到位姿预测误差上，因此对位姿预测误差平均值的优化要好于最大值。

在实时性方面，本文算法进行一次完整预测大约需要 1 ms，其中延时预测过程大约 0.2 ms，位姿预测过程大约 0.8 ms，而当对各方向上的运动进行解耦后，位姿预测过程耗时在 0.1 ms 以下。当对渲染管线进行后，头部位姿预测和图像渲染需要在同一个垂直同步周期（约 13.3 ms）内完成，而一个比较复杂场景（约 90 万个三角面）的渲染时间约 4.5 ms。因此，虽然本文算法比已有方法的时间复杂度高，但实时性仍然完全满足虚拟现实系统的要求（本文实验平台仅满足虚拟现实要求的最低硬件配置，当硬件性能提升后，算法实时性还会进一步提升）。

## 3 结论

为了最小化运动画面延时对沉浸式虚拟现实系统的影响，本文提出了一种具有延时抖动补偿功能的卡尔曼滤波算法对用户头部位姿进行预测，并在 Oculus DK2 上对其有效性进行了验证。相比现有的头部位姿预测算法，所提算法增加了对运动画面延时和头部运动相关性的分析与运用，文中先根据头部运动相关性的分析结果选择了合适的预测方式，然后在头部运动建模和位姿预测时分别利用延时预测结果对延时抖动进行了补偿。实验结果表明，本文所提算法可显著提高头部位姿的预测精度（相比文[10]给出的设备自带预测算法，在一般情况下，其姿态预测绝对误差的均值分别减小 45.74%、47.25% 和 40.96%，最大值分别减小 11.49%、26.34% 和 36.79%；位置预测绝对误差的均值分别减小 35.94%、45.90% 和 55.81%，最大值分别减小 1.05%、25.60% 和 44.74%），但是受运动画面延时随机性的影响，其对误差最大值的优化程度要低于均值。受实验条件限制，本文没能够根据骨骼肌肉的生理学特性建立出精确的头部位姿约束模型，如果能够利用该模型对头部位姿预测过程进行约束和优化，理论上可以进一步提高头部位姿的预测精度。

## 参考文献

- [1] Bo L, Fuyi C, Jinfeng L. The teleoperation simulation system based on VR[C]//2014 International Conference on Virtual Reality and Visualization. Piscataway, NJ, USA: IEEE, 2014: 7–11.
- [2] Cichon T, Roßmann J. Robotic teleoperation: Mediated and supported by virtual testbeds[C]//2017 IEEE International Symposium on Safety, Security and Rescue Robotics (SSRR). Piscataway, NJ, USA: IEEE, 2017: 55–60.
- [3] Rakshit S M, Banerjee S, Hempel M, et al. Fusion of VR and teleoperation for innovative near-presence laboratory experience in engineering education[C]//2017 IEEE International Conference on Electro Information Technology (EIT). Piscataway, NJ, USA: IEEE, 2017: 376–381.
- [4] White P J, Byagowi A, Moussavi Z. Effect of viewing mode on pathfinding in immersive virtual reality[C]//37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Piscataway, NJ, USA: IEEE, 2015: 4619–4622.
- [5] Boustila S, Bechmann D, Capobianco A. Effects of stereo and head tracking on distance estimation, presence, and simulator sickness using wall screen in architectural project review[C]//2017 IEEE Symposium on 3D User Interfaces (3DUI). Piscataway, NJ, USA: IEEE, 2017: 231–232.
- [6] So R H, Lo W T, Ho A T. Effects of navigation speed on motion sickness caused by an immersive virtual environment[J]. Human Factors, 2001, 43(3): 452–461.
- [7] Aldaba C N, White P J, Byagowi A, et al. Virtual reality body motion induced navigational controllers and their effects on simulator sickness and pathfinding[C]//39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Piscataway, NJ, USA: IEEE, 2017: 4175–4178.
- [8] Rebenitsch L, Owen C. Review on cybersickness in applications and visual displays[J]. Virtual Reality, 2016, 20(2): 101–125.

(下转第 64 页)

- [13] Cai T T, Silverman B W. Incorporating information on neighboring coefficients into wavelet estimation[J]. Sankhya: The Indian Journal of Statistics, Series B, 2001, 63(2): 127–148.
- [14] Chen G, Zhu W P. Signal denoising using neighboring dualtree complex wavelet coefficients[J]. Signal Processing et, 2012, 6(2): 143–147.
- [15] Qin Y, Tian L, Wang C. Power quality disturbances detection and location based on multi-wavelet and neighboring coefficient denoising[J]. International Conference on Electric Utility Deregulation and Restructuring and Power Technologies, 2011: 424–427.
- [16] Dai J. Image denoising based on combining neighboring wavelet coefficients[J]. International Congress on Image & Signal Processing, 2009: 1–3.
- [17] 薛坚, 于盛林. 基于邻域相关性新阈值函数的提升小波域信号降噪法[J]. 信息与控制, 2008, 37(6): 665–669.  
Xue J, Yu S L. Lifting wavelet-domain signal denoising based on a new threshold function with neighbor dependency[J]. Information and Control, 2008, 37(6): 665–669.
- [18] Barri A, Dooms A, Schelkens P. The near shift-invariance of the dual-tree complex wavelet transform re-visited[J]. Journal of Mathematical Analysis and Applications, 2013, 389(2): 1303–1314.
- [19] Kennedy J, Eberhart R C. Particle swarm optimization[C]//Proceedings of the IEEE International Conference on Neural Networks. Piscataway, NJ, USA; IEEE, 1995: 1942–1948.
- [20] 吴一全, 纪守新. 基于混沌粒子群优化的图像 Contourlet 阈值去噪[J]. 光子学报, 2010, 39(9): 1645–1651.  
Wu Y Q, Ji S X. Image contourlet threshold denoising based on chaotic particle swarm optimization[J]. Acta Photonica Sinica, 2010, 39(9): 1645–1651.
- [21] 谭跃. 具有混沌局部搜索策略的粒子群优化算法研究[D]. 长沙: 中南大学, 2013: 45–49.  
Tan Y. Research on particle swarm optimization with chaotic local search[D]. Changsha: Central South University, 2013: 45–49.
- [22] 周燕, 刘培玉, 赵静, 等. 基于自适应惯性权重的混沌粒子群算法[J]. 山东大学学报(理学版), 2012, 47(3): 27–32.  
Zhou Y, Liu P Y, Zhao J, et al. Chaos particle swarm optimization based on the adaptive inertia weight[J]. Journal of Shandong University: Natural Science, 2012, 47(3): 27–32.

## 作者简介

宋立业(1972–), 男, 硕士, 副教授. 研究领域为电力系统自动化装置.

周 乐(1989–), 男, 硕士. 研究领域为小波去噪, 基于数据驱动的故障诊断.

刘昕明(1984–), 男, 博士, 讲师. 研究领域为复杂工业建模、优化和诊断.

(上接第 56 页)

- [9] Bos J E, Bles W, Groen E L. A theory on visually induced motion sickness[J]. Displays, 2008, 29(2): 47–57.
- [10] LaValle S M, Yershova A, Katsev M, et al. Head tracking for the oculus rift[C]//2014 IEEE International Conference on Robotics and Automation. Piscataway, NJ, USA: IEEE, 2014: 187–194.
- [11] LaViola J J. A comparison of unscented and extended Kalman filtering for estimating quaternion motion[C]//American Control Conference. Piscataway, NJ, USA: IEEE, 2003: 2435–2440.
- [12] Rhijn A V, Liere R V, Mulder J D. An analysis of orientation prediction and filtering methods for VR/AR[C]//IEEE Proceedings VR 2005 Virtual Reality. Piscataway, NJ, USA: IEEE, 2005: 67–74.
- [13] Himberg H, Motai Y. Head orientation prediction: Delta quaternions versus quaternions[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 2009, 39(6): 1382–1392.
- [14] Himberg H, Motai Y, Barrios C. R-adaptive kalman filtering approach to estimate head orientation for driving simulator[C]//2006 IEEE Intelligent Transportation Systems Conference. Piscataway, NJ, USA: IEEE, 2006: 851–857.
- [15] Drouard V, Horaud R, Deleforge A, et al. Robust head-pose estimation based on partially-latent mixture of linear regressions[J]. IEEE Transactions on Image Processing, 2017, 26(3): 1428–1440.
- [16] Kang M J, Lee H Y, Kang J W. Head pose estimation using random forest and texture analysis[C]//2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference. Piscataway, NJ, USA: IEEE, 2016: 1–4.
- [17] Choi S W, Seo M W, Kang S J. Prediction-based latency compensation technique for head mounted display[C]//2016 International SoC Design Conference. Piscataway, NJ, USA: IEEE, 2016: 9–10.
- [18] Horn M, Sreenivasa M, Mombaur K. Optimization model of the predictive head orientation for humanoid robots[C]//2014 IEEE-RAS International Conference on Humanoid Robots. Piscataway, NJ, USA: IEEE, 2014: 767–772.
- [19] Dan C, Xiuhui F, Wei D, et al. Shifted gamma distribution and long-range prediction of round trip timedelay for internet-based teleoperation[C]//2008 ROBIO 2008 IEEE International Conference on Robotics and Biomimetics. Piscataway, NJ, USA: IEEE, 2009: 1261–1266.
- [20] 魏青, 崔龙. 基于时延预测的遥操作机器人预测显示方法[J]. 机器人, 2017, 39(3): 298–306.  
Wei Q, Cui L. Predictive display for telerobot based on time-delay prediction[J]. Robot, 2017, 39(3): 298–306.

## 作者简介

魏 青(1988–), 男, 博士. 研究领域为遥操作机器人, 人机交互, 虚拟现实.

刘钊铭(1989–), 男, 博士生. 研究领域为遥操作机器人, 人机交互, 机器学习.

刘乃龙(1989–), 男, 博士生. 研究领域为遥操作机器人, 人机交互, 机器人控制.