

面向残障人士的基于眼动交互的绘画系统

李 珊, 潘 纲, 李石坚

(浙江大学计算机学院, 浙江杭州 310027)

摘 要: 本文建立了一个面向严重肢体残障人士的眼动交互绘画系统. 为此, 首先给出了眼动数据的预处理与分析方法; 为减弱米达斯(Midas)接触问题对眼动行为判别的影响, 提出了基于时域约束的注视点聚类算法, 可提取出注视点判断用户当前的视点中心, 并据此进一步判断用户眼睛行为; 针对眼动原始数据中的噪声影响, 设计了窗口长度自适应的视点信号平滑处理方法, 可减弱视点抖动产生的噪声, 实现自由绘画较为平稳的绘画轨迹. 在此基础上, 设计并实现了基于 ASL 眼动仪的眼动交互式绘画系统.

关键词: 眼动交互; 驻留时间; 眼睛行为

中图分类号: TP37 **文献标识码:** A **文章编号:** 0372-2112 (2011) 3A-163-05

Eye-Controlled Painting System for Disabled

LI Shan, PAN Gang, LI Shi-jian

(Department of Computer Science, Zhejiang University, Hangzhou, Zhejiang 310027, China)

Abstract: The paper describes an eye-controlled painting system which enables people with severe mobility impairments to draw just by moving their eyes. We design and implement the system, and also propose some algorithms to improve the performance of eye-tracking interaction. We present a time-domain constrained eye-tracking clustering algorithm to detect fixations, and then identify eye movement behaviors, which can reduce the “Midas touch” problem. We also present a window-adaptive discrete Gaussian smoothing algorithm, which can filter noise and smooth the path of eye-movement in a reasonable way.

Key words: eye-tracking interaction; dwell time; eye movement behavior

1 引言

由于严重肢体残障人士的肢体自由移动受限, 与环境的交互与沟通就极其不便, 例如无法像正常人一样通过鼠标键盘使用电脑. 根据 2006 年我国第二次全国残疾人抽样调查^[2], 全国各类残障人士 8296 万人, 其中肢体残疾 1233 万人, 多重残疾 1352 万人. 其中严重肢体残障者虽然无法正常移动肢体, 但大部分可以正常控制眼睛运动. 通过眼睛运动(眼动)与外界进行交互与沟通正成为严重肢体残障者的主要交互方式之一. 当前的眼动交互主要存在两个问题: (1) 眼动噪声问题, 由于眼动获取设备存在噪声, 眼睛本身也存在抖动、眨眼等干扰行为, 使得眼动数据存在干扰, 需要去噪^[11]; (2) 米达斯(Midas)接触问题^[7,8,10], 由于用户视线运动的随意性而造成计算机对用户意图识别的困难, 即用户的每次眼睛定位都可能引发一条并非想要的计算机命令. 这两个问题的存在阻碍了眼动交互应用的实用化.

本文面向严重肢体残障者, 提出并实现了一个基于

眼动交互的绘画系统, 是作者 PCC09 会议论文^[13]的杂志扩展版. 主要工作有: (1) 提出了基于时域约束的视点聚类算法, 从原始视点数据流中提取出注视点, 判断出用户当前的视点中心, 并进一步判断用户眼睛行为; (2) 利用眼动窗口长度自适应的高斯滤波算法, 实现了有效的自由绘画, 可得出较为平稳的绘画轨迹; (3) 在眼动仪设备的基础上实现了一个可同时支持自由绘画与任务选择的眼动交互绘画软件系统.

2 相关工作

在眼动交互方面, 主要有完全眼动交互, 眼动辅助交互, 眼动打字和眼动绘画等方面的研究. 其中眼动绘画主要有两类: 自由眼动绘画与任务选择的眼动绘画. (1) 自由眼动绘画: 无论用户看哪里, 视点轨迹都会被画下来. 相关研究有 Tchalenko^[3]和 EagleEyes^[4]. 自由绘画可以激发创造力和想象力, 但受米达斯接触问题所限, 无法区分“看”和“绘画”两种行为, 最后的作品只能是连续的线条; (2) 任务选择的眼动绘画: Anthony 等人研究

的 EyeDraw^[5]中,用户可以选择当前任务,光标可以在“看”和“操作”两种状态下切换,切换的依据是视觉驻留时间^[9],当眼睛凝视某处超过一定时间阈值时,光标状态发生切换.该系统只有确定几何图形的绘画,无法根据眼动轨迹画图. Alvin 等人提出了注视分析模型 GEM(Gaze Estimation Model)^[6],其利用目光移动的方式判断是“寻找”还是“绘画”,这一模型通过判断前一点与当前点的距离和停留时间计算得出用户的兴趣中心.当眼动注视点之间距离很远且在屏幕上很分散时,用户行为基本判断为“寻找”;相反,当眼动相对集中在某一兴趣中心周围时,判断用户行为为“绘画”.但该判断规则易混淆两种行为,难以实际应用.

在多通道绘画方面,国内学者也做了若干相关研究,如中科院软件所设计实现了一个基于笔式交互和语音交互的面向儿童的多通道交互系统^[8],系统中定义了一些基于笔和语音的交互,可支持孩子们以自然的方式,通过笔和语音同系统进行交互.

3 眼动绘画的若干关键算法

我们首先介绍两个重要眼动概念:眼动状态和眼动行为:(1)眼动状态:在生理学上指眼睛当前运动的状态,不涉及用户意识,包括注视(fixations)、跳动(Saccades)和平滑尾随跟踪(Smooth pursuit)三种^[12].本文中主要用到了前两种.从生理学角度看,眼动是由一系列对观察目标的注视和注视之间的跳动组成的,其中注视表现为视点在被观察目标上的停留,这些停留一般至少持续 100ms~200ms 以上;跳动则指视点在注视点间的飞速跳跃,视角为 1°~40°,持续时间为 30ms~120ms,最高速度为 400°/s~600°/s,跳动一般发生在对目标的搜寻过程中;(2)眼动行为:在本文中指与眼动交互命令发布相关、用户有意识的眼动交互操作,包括视觉驻留行为(Dwell)和移动行为(Moving).视觉驻留指眼睛有意识地凝视待操作对象的行为,而移动表示视点运动是无规则的.

本文通过视点信号的处理从原始的眼动数据中提取出注视和跳动两大类眼动状态,在识别出注视点的基础上进行眼动行为的识别判断,当眼睛注视某个兴趣区域超过驻留时间阈

值时,表示一个有意义的眼动行为,从而引发绘画操作.其涉及视点处理的三个关键算法:注视点提取算法、眼动行为识别算法、视点序列平滑算法.图 1 为眼动交互绘画的基本流程,首先采用基于时域约束的注视点聚类算法对原始眼动视点序列进行聚类,将同时满

足一定时间和空间约束的眼动视点序列聚类为一个注视聚类;然后通过眼动行为识别算法对注视聚类中心点序列进行处理,判断结果是否满足特定的驻留空间阈值,从而判断当前眼动行为.当眼动行为为视觉驻留时,往往代表当前绘画任务的改变(如绘画开始或结束等);当眼动行为为移动时,通过视点平滑算法进行视点的平滑滤波处理,执行当前绘画任务(如曲线绘画等).通过这三种算法,可有效从原始眼动数据中去除噪声,提取出眼动状态,进一步得到特定的眼动行为,从而达到用眼睛控制绘画操作的目的.

3.1 基于时域约束的注视点聚类与眼动状态提取

视点采集设备获取的原始视点信号可表示为一个视点序列 G :

$$G = \{g_i | i = 1, \dots, n\}, g_i = (x_{g_i}, y_{g_i}, t_{g_i})$$

每个视点 g_i 可用平面上的二维坐标 (x_{g_i}, y_{g_i}) 外加时间维 t_{g_i} 表示.原始视点信号 G 无法直接表示用户的操作意图,若要进行眼动交互绘画,需对这些原始视点信号进行分析.

虽然已经存在一些在线或数据流的聚类算法,但对眼动绘画的案例,它们并不十分合适^[1].其一,由于聚类的个数是未知的,连续的 K-means 算法并不合适;其二,由于眼动聚类的实时性要求,而在线层次聚类算法^[10]需要得到一段数据后再更新,也不适用于眼动聚类;其三,对于一般聚类个数未知的聚类算法^[10],一般先设定一个 K 值,再使用二次计算,合并或分裂距离过小或过大的类.但对于眼动绘画,有实时性要求,不能撤销之前的眼动操作.

为此,本文设计了基于时域约束的注视点聚类算法,即在连续时间内同时满足一定时间和空间约束的眼动视点序列聚类为一个注视聚类.该算法同时考虑了视点序列时间上和空间上的相似性,只有在时间和空间综合计算后才能得出视点序列的内部距离.当视点之间的时间间隔很短并且空间距离较近时,视点之间的距离必定落在聚类阈值内部.当视线由于眼睛自身的抖动短暂的跳出目标范围之外时,仅仅参考空间位置的聚类方法将出现较大的空间距离,很可能暂时大于聚类阈值范围,而本算法通过加入时间权重,总的视点距离依然可以落在阈值内部.只有当视线较长时间的落在聚类判断阈值范围外时,视点聚类距离才会很大,表明当前眼睛并不是盯着目标,而是随意寻找,眼睛状态为跳动,此时开启新的聚类.

定义一种新的融合时间与空间的视点距离度量 ξ :

$$\xi(g_i, g_j) = \frac{k_1 * \sqrt{(x_{g_i} - x_{g_j})^2 + (y_{g_i} - y_{g_j})^2} + k_2 * |t_{g_i} - t_{g_j}|}{k_1 + k_2} \quad (1)$$

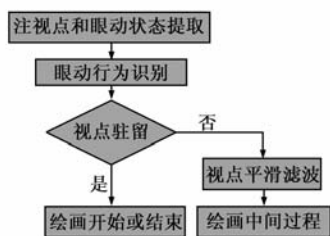


图1 眼动交互绘画基本流程

其中 k_1 与 k_2 为权重系数,用于调整时间域与空间域在视点距离中的重要程度.基于该视点距离度量,我们建立了一种新的基于时域约束的注视点聚类算法,从而可提取注视点并计算眼动状态.其基本思想如下:输入当前视点,判断同时满足一定时间和空间阈值的眼动视点序列作为一个注视点聚类 G_1 ,每读入一个点判断是否在当前注视聚类的范围内,是则更新当前聚类中心,否则更新潜在聚类中心.判断时间,若连续 T 时间都落在当前注视聚类中,则当前状态为注视.具体算法如图 2 所示. c_1 为当前聚类中心, s_0 为当前视点下标, $\text{State}(c_1)$ 为 c_1 的状态, T_{fix} 和 D_{fix} 是聚类的时间和空间阈值.

算法 1 基于时域约束的注视点聚类算法

```

0: 初始化:  $c_1 = \text{NULL}$ 
1: 输入:当前视点  $g_k$ 
2: if  $c_1 = \text{NULL}$  then
3:    $c_1 = g_k; s_0 = k$ 
4:    $\text{State}(c_1) = \text{Saccades}$ 
5: else
6:   if  $\xi(g_k, c_1) < D_{\text{fixation}}$  then
7:     更新聚类中心
8:     
$$c_1 = \left( \frac{\sum_{i=s_0}^k x_{g_i}}{k - s_0}, \frac{\sum_{i=s_0}^k y_{g_i}}{k - s_0}, \frac{\sum_{i=s_0}^k t_{g_i}}{k - s_0} \right)$$

9:     if 持续时间  $t_{g_k} - t_{g_{s_0}} > T_{\text{fix}}$  then
10:       $\text{State}(c_1) = \text{Fixation}$ 
11:     else  $\text{State}(c_1) = \text{Saccades}$ 
12:   endif
13: else
14:   保存聚类中心  $c_1$  及其对应的状态  $\text{State}(c_1)$ //新产生一个注
15:   视点
16:    $c_1 = g_k; s_0 = k$ 
17:    $\text{State}(c_1) = \text{Saccades}$ 
18:   endif
19:   输出:聚类中心  $c_1$  和  $\text{State}(c_1)$ ;
20:   goto 1

```

图 2 基于时域约束的注视点聚类算法

3.2 眼动行为识别

得到注视点及眼动状态(State)后,还需要根据状态识别出眼动行为,从而触发相应操作.本文中利用视觉驻留时间阈值判断眼动行为,若持续的注视点落在同一个目标范围内,且满足驻留时间阈值,则为视觉驻留行为,否则为移动行为.具体算法如图 3 所示.这里 c_0 为上一个注视聚类中心, c_1 为当前注视聚类中心, D_{dwell} 为在同一个驻留兴趣范围的时空阈值, T_{dwell} 为驻留时间.若当前眼动状态为跳动,则眼动行为是移动行为.否则,若当前眼动状态为注视,且持续的注视点落在同一个区间内,时间大于驻留时间,则眼动行为为视觉驻留.

算法 2 眼动行为识别算法

```

1: 输入:当前注视聚类中心  $c_1$  及上一个注视聚类中心  $c_0$ 
2: if 当前  $\text{State}(c_1) = \text{Saccades}$  then
3:   眼动行为 = Moving
4: else
5:   if  $\xi(c_0, c_1) < D_{\text{dwell}}$  and  $t_{c_1} - t_{c_0} > T_{\text{dwell}}$  then
6:     眼动行为 = Dwell
7:   else
8:     眼动行为 = Moving
9: 输出:眼动行为

```

图 3 眼动行为识别算法

3.3 视点平滑滤波

在眼动自由绘画时,为防止过多的抖动,需要对视点进行平滑.具体方法如下:在视点序列集 $G = \{g_i | i = 1, \dots, n\}$ 上取平滑时间窗口 N ,平滑算法从当前点回溯前 N 时间内的视点,求加权均值作为平滑视点 δ ,窗口随时间向前推进.

窗口大小的选取要综合考虑两个问题:(1)去除抖动和尖锐点;(2)保留原始信息,及时响应.选取较小的窗口能让光标及时跟随眼动运动,但是抖动和尖锐点较明显;相反的,较大的窗口能使光标平稳,但是对于光标位置改变的响应较慢,得到的平滑结果与原始信息区别较大.因此,本文选择了基于速度和眼动状态的窗口大小选取方法,在眼动状态为注视时,较大的时间窗口,减少注视过程中可能的尖锐点产生的抖动;在眼动状态为跳动时,窗口随速度增大而减小,在平滑的同时也保证光标及时的响应.视点平滑滤波具体见图 4.

算法 3 窗口长度自适应的视点平滑滤波算法

```

1: 输入:当前视点  $g_n$ 
2: if ( $g_n \in \text{Fixation}(c_1)$ )
3:    $N = \min(\text{smoothMax}, \text{fixSample})$ ;
4: else if ( $g_n \notin \text{Fixation}(c_1)$ )
5:    $N = \max(\text{smoothMax} * \exp(-V_{g_n}^2 / 2\epsilon^2), \text{smoothMin})$ 
6: endif
7: 
$$\delta_n = \frac{\sum_{i=0}^{N-1} e^{-(g_{n+1-N+i} - g_n)^2 / 2\sigma^2} g_{n+1-N+i}}{\sum_{i=0}^{N-1} e^{-(g_{n+1-N+i} - g_n)^2 / 2\sigma^2}}$$

8: 输出:平滑视点  $\delta_n$ 
9: goto 1

```

图 4 窗口长度自适应的视点平滑滤波算法

我们将三角滤波^[14], Kalman 滤波^[15]与本文提出的窗口长度自适应的视点平滑滤波算法进行了比较.在一般情况下,本文滤波算法的结果与三角滤波、Kalman 滤波的结果都较相近;当眼动速度较快的时候,本文算法要优于其他两个算法,如图 5 所示.

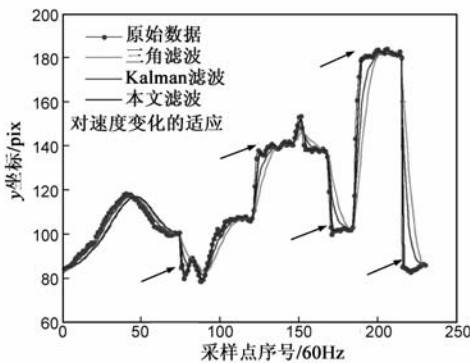


图5 速度变化较剧烈时视点的平滑处理结果比较

4 眼动绘画系统的实现及结果

4.1 眼动仪设备

本文的眼动绘画系统采用由美国 ASL(Applied Science Laboratories)实验室研制的 H6 眼动仪^[8]作为眼动数据采集设备.该设备包括 ASL 头戴式光学系统,ASL 控制单元两部分.

4.2 绘画任务选择状态机

本文中绘画系统主要有两种眼动任务:观察和绘画,两种任务以视觉驻留作为依据互相转换^[5,6,9].其中绘画又分为图章绘图、几何绘图和自由曲线三种形式.根据不同的绘画任务,用户有不同的操作流程.具体绘画任务选择状态机如图 6 所示.

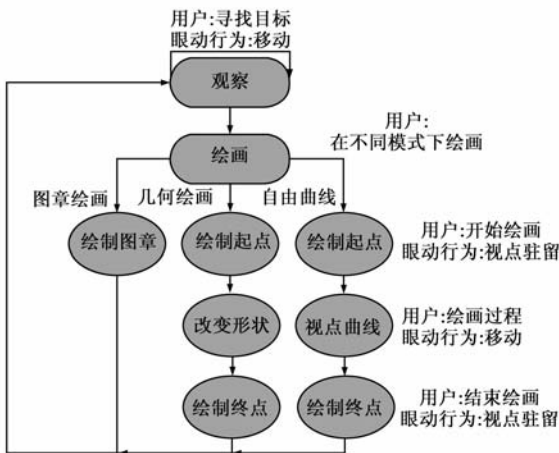


图6 绘画任务选择状态机

准备自由曲线的眼动绘图时,用户可随意的看,此时眼动行为是移动,当选定绘画目标点以后凝视目标,当达到视觉驻留时间阈值后系统检测到视觉驻留行为,触发绘画任务.例如图章绘图状态下,用户选定图章,凝视目标点到达驻留时间.在几何绘图状态下,用户凝视起点到达驻留时间,拖动光标改变形状的大小,再凝视终点到达驻留时间.在自由曲线绘图状态下,用户凝视某个起点,开启自由绘图,中间的眼动轨迹都在平滑滤波后被画下来,最后凝视终点结束自由绘图.在

绘画任务完成后用户移动视点到的位置,切换为观察任务.

4.3 眼动绘画系统运行结果

本系统功能主要有 4 部分:绘画模式选择、绘画颜色选择、系统命令选择和操作说明选择.绘画模式能绘制自由曲线,图案图章,几何图形等不同方式;绘画颜色能选择不同颜色的画笔和填充.系统界面如图 7 所示.用户操作场景及部分绘画结果如图 8 图 9.

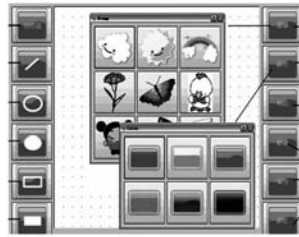


图7 系统功能及界面



图8 用户操作场景



图9 部分绘画结果展示

4.4 系统可用性评估

为了评估系统可用性,我们设计了如下调查问卷.问卷中涉及用户的舒适性,使用过程中疲劳程度,绘画过程中的响应程度等方面的评价.我们选取了 3 名普通用户模拟肢障者,仅用眼睛对该系统使用一段时间后填写问卷,评价该系统.

表 1 用户体验问卷调查打分(1分:最差,5分:最好)

题目/用户	A	B	C	平均
1.舒适性	4	4	4	4.0
2.使用没有疲劳	2	4	3	3.0
3.使用后没有眼干涩	5	4	5	4.7
4.及时响应程度	4	5	5	4.7
5.眼睛定位容易程度	2	4	4	3.3
6.画线容易程度	4	4	4	4.0
7.选颜色容易程度	3	4	3	3.3
8.趣味性	5	4	5	4.7
总体评价	4	5	4	4.3

总体上看,用户在该系统的使用过程中及之后对系统的评价良好,用户感觉使用过程中舒适性好,但使用时间过长后有眼干涩的情况.在绘画过程中,用户认为响应及时,画线和选取颜色都较容易,但有用户认为眼睛定位较为吃力,这与眼睛本身的抖动特性有关.最后,用户认为该眼动绘画系统趣味性较强,使用起来比较有意思.

5 结语

本文基于眼动交互的理念,面向严重肢体残障人士,设计并实现了一个基于眼动交互的绘画系统.为了从原始视点信号中得到当前的眼动行为,本文提出了基于时域约束的视点聚类算法,能从原始数据流中提取出注视点,判断出用户当前的注视点,以此可进一步判断用户绘画的行为.为了减弱米达斯接触问题影响,本文利用驻留时间的判断区分“观察”和“绘画”两种用户行为.最后为了实现自由绘画,本文设计了基于眼动状态的平滑滤波方法,可针对不同的绘图状态,消除原始视点数据的抖动问题.

参考文献

- [1] Mahamed G H Omran, Andries P Engelbrecht, Ayed Salman. An overview of clustering methods[J]. Intelligent Data Analysis, 2007, 11(6): 538 – 605.
- [2] 第二次全国残疾人抽样调查主要数据公报[R]. 国家统计局, 2006.
- [3] Tchalenko J. Drawing and cognition free-eye drawing point [J]. Art & Design Research Journal, 2001, (11): 36 – 41.
- [4] James Gips, Peter Olivier. EagleEyes: an eye control system for persons with disabilities[A]. The 11th International Conference on Technology & Persons with Disabilities[C]. Los Angeles, 1996. 3.
- [5] Anthony J Hornof, Anna Cavender. EyeDraw: enabling children with severe motor impairments to draw with their eyes[A]. Proceedings of the SIGCHI conference on Human factors in computing systems[C]. USA: ACM Press, 2005. 161 – 170.
- [6] Chiu Po Chan, Alvin W Yeo. Gaze estimation model for eye drawing[A]. Proceedings of ACM CHI'06 extended abstracts on Human factors in computing systems [C]. USA: ACM Press, 2006. 1559 – 1564.
- [7] Jacob R J K. What you look at is what you get: Eyemovement based interaction techniques[A]. Proceedings of the SIGCHI conference on Human factors in computing systems: Empowering people[C]. USA: ACM Press, 1990. 11 – 18.
- [8] 李杰, 田丰, 王维信, 戴国忠. 面向儿童的多通道交互系统[J]. 软件学报, 2002, 13(9): 1846 – 1851.
Li Jie, Tian Feng, Wang Weixin, Dai Guozhong. A multimodal interaction system for children[J]. Journal of Software, 2002, 13(9): 1846 – 1851. (in Chinese)
- [9] Chris Lankford. Effective eye-gaze input into Windows[A]. Proceedings of the 2000 symposium on Eye tracking research & applications[C]. USA: ACM Press, 2000. 23 – 27.
- [10] Guha S, Meyerson A., Mishra N, Motwani R, O'Callaghan L. Clustering data streams: Theory and practice [J]. IEEE

Transactions on Knowledge and Data Engineering, 2003, 15 (3): 515 – 528.

- [11] 冯成志, 沈模卫. 视线跟踪技术及其在人机交互中的应用[J]. 浙大学报, 2002, 29(2): 225 – 232.
Feng Chengzhi, Shen Mowei. Application of gaze tracking in human computer interaction[J]. Journal of Zhejiang University (Science Edition), 2002, 29(2): 225 – 232. (in Chinese)
- [12] Carpenter, Roger H S. Movements of the Eyes[M]. Pion Ltd, 1988.
- [13] 李珊, 潘纲, 李石坚. 面向残障人士的基于眼动交互的绘画系统[A]. 第 5 届全国普适计算学术会议论文集[C]. 西安, 2009.
Li Shan, Pan Gang, Li Shi-jian. Eye-controlled painting system for disabled[A]. PCC[C]. Xi'an, 2009. (in Chinese)
- [14] M S Arulampalam, S Maskell, N Gordon, T Clapp. Tutorial on particle filters for online nonlinear/non-gaussian Bayesian tracking[J]. IEEE Transactions on Signal Processing, 2002, 50 (2): 174 – 188.
- [15] M Kumar, J Klingner, R Puranik, T Winograd, A Paepcke. Improving the accuracy of gaze input for interaction[A]. Proceedings of Eye Tracking Research & Applications[C]. Savannah: ACM, 2008. 65 – 68.

作者简介



李 珊 女, 1987 年 10 月出生于浙江金华. 现为浙江大学计算机学院研究生, 研究方向为普适计算.

E-mail: lishan33@zju.edu.cn



潘 纲(通讯作者) 男, 1976 年 10 月出生于浙江永嘉. 现为浙江大学计算机学院副教授, 研究方向为普适计算、计算机视觉与模式识别.

E-mail: gpan@zju.edu.cn



李石坚 男, 1979 年 12 月出生于湖南益阳. 现为浙江大学计算机学院讲师, 研究方向为普适计算、传感器网络.

E-mail: shijianli@zju.edu.cn