**文献总结**

该论文介绍了一种集成到Mini社交机器人中的响应式的联合注意系统（RJAS），实时计算用户的头部的方向、面部位置和手臂指向，机器人根据用户的注视和指向方向进行运动。当检测到干扰时，机器人会使用言语表达来提醒用户，从而增强对机器人响应能力的感知。设计了一个互动游戏场景“奇偶游戏”的用户实验，研究评估了联合注意系统激活与否对用户对机器人的社交存在感（Social Presence）的影响。

**本文主要贡献：**

1.联合注意系统设计：创建“兴趣区域”（Regions of Interest），让机器人根据信息的上下文调整其注意焦点，从而构建更丰富的交互框架。

2.人性化响应机制：模仿人类生物学中的行为（前庭系统在联合注意事件中的动作机制），使机器人在响应用户注意时表现出更细节的动作。

3.用户视角的评估：通过标准化问卷调查，从用户的角度评估机器人在社交存在感各维度上的表现，证明系统激活时用户能明显感知到机器人的响应行为。

**机器人联合注意机制研究现状：**

非语言线索：如眼球运动，头部方向和身体方向在互动中至关重要，有助于建立和维持个体之间的联系。

模仿与轮流：模仿指机器人模仿他人的动作或表情，从而增强彼此间的理解和社会凝聚力；轮流机制确保互动双方能交替承担信号发出者和响应者的角色，从而实现平衡互动和相互参与。

联合注意的不同阶段：建立相互凝视吸引注意力，然后通过凝视将焦点指向感兴趣的对象，重新建立眼神接触确保联合注意。

**响应式联合注意系统的架构：**

1.探测器模块（Detectors）： 收集用户信息（用户头部方向、面部位置和身体姿势等）。

2.视觉注意模块（VAM）： 基于用户的运动和正在进行的活动（与机器人互动的游戏）动态分配每个刺激的强度值，从而确定当前的FoA。

3.仿生行为模块（BBM）： 负责计算机器人的逆向运动学以调整其朝向FoA的姿态，并产生符合自然行为的动作（如眼睛、头部、颈部和躯干的运动）。生成口头回应通知用户机器人已感知到其提示。

探测器模块：

1.面部探测器： 基于SqueezeNet架构，使用边界框坐标计算图像中每个用户的面部质心。通过采用深度图像中质心的X和Y坐标，检测器获得从机器人到每个用户的面部的距离

2.头部方向探测器： 输入面部探测器检测到的用户面部的边界框，输出用户头部的旋转角度，包括三个欧拉角：偏航（yaw）、俯仰（pitch）和滚转（roll）。基于这些角度信息，头部探测器可以计算出头部姿态的（X, Y）。

公式1计算头部姿态。

图2b头部方向向量的角度被离散化为八个方向加一个正面方向，使机器人可以通过判断用户注视的方向，模拟用户的视线方向。

公式2～5 进行额外的修正来确定用户是否在看向机器人

3.手臂方向探测器： 使用PoseNet模型输入检测输入图像的用户身体上的17个关键点（头部、肩膀、肘部、手腕等部位的坐标）计算手臂的方向，特别是计算肩膀和手腕之间的角度

公式6计算手腕与肘部之间的相对角度来确定手臂的方向

4.手部探测器： 使用MediaPipe模型检测检测到用户手部的21个参考点（手掌的中心点、每个手指的关节点）捕捉手部的形状和姿势

视觉注意模块：

1.通过对来自不同探测器的信息进行处理，确定各个刺激的优先级。优先级定义为在某一时刻和空间内最重要的刺激。每个刺激都被分配一个强度值，范围为 [0, 1]。

优先级强度更新规则：用户的视线方向和手臂信号：当用户转动头部或用手臂指向某个方向时，这些刺激的强度值为1；否则，强度值为0。

用户的面部位置：其强度值固定为0.5，以确保机器人始终跟踪用户的面部位置。

2.将所有刺激按照强度值排序，并将强度值最高的刺激确定为共同关注点（FoA），将实际位置发送给控制器计算机器人的关节位置，并调整机器人以对准FoA。

在确定FoA后，系统会减少该刺激的强度值（将其设为0），以便机器人可以重新聚焦在用户的面部位置，从而模拟视线的转移机制。

用户面部的强度值始终为0.5，确保机器人始终跟踪用户的面部，并只在检测到用户提示或活动信号时才转移注意力。

仿生行为模块：

负责机器人的运动控制和语言互动，以实现机器人对共同关注事件（FoA）的对准，并向用户传达机器人感知到用户分心并主动参与联合注意过程的信号。

1.运动控制器：计算机器人自由度的逆运动学并产生自然运动来将机器人朝向FoA定向。

图5 控制器通过逆向运动学计算每个关节（如眼睛、颈部、头部和躯干）的目标角度，以对准FoA。公式7 计算转动角度

如果FoA在图像的可视范围内，机器人会首先通过调整眼睛的位置对准FoA。

如果眼睛达到极限，机器人会转动头部和颈部，同时应用反向运动来修正眼睛的位置

当头部对准FoA后，机器人将移动躯干，以确保整个身体都对准目标。

2.反应性人机交互：当机器人对用户的行为做出反应时，它通过口头交互做出评论，以便用户感知机器人正在对他们的行为做出反应。

如果机器人察觉到用户因转动头部而使得FoA超出了机器人的视野范围，机器人会转动头部，跟随用户的分心方向。

如果用户持续分心超过2秒，机器人会通过语言提示用户。

一旦用户重新关注到机器人或活动，机器人会继续进行当前的互动，保持互动的自然连续性。

**RJAS在Mini机器人中的集成：**

Mini是一款具有五个自由度的机器人，包括基础、手臂、颈部和头部。配备了集成在面部、心脏和嘴巴的LED灯，能够展现丰富的情感表达。内置扬声器和麦克风支持语音交互，电容传感器则用于检测用户触摸。配有一个平板显示器，用于展示多媒体内容和菜单，增加与用户互动的乐趣。使用RGB-D深度摄像头捕捉用户信息。

图7展示了Mini的整体软件架构，a为探测器模块。感知管理器PM处理来自检测器的信息，使用相机的局部位置作为坐标原点，计算图像内的像素坐标和真实世界坐标中的每个检测的质心位置，然后将此信息发送到b中的响应式的联合注意系统RJAS。HRIM为人机交互管理器，DMS为决策系统。视觉注意模块VAM会自动计算每个探测器的强度值，并根据c中技能模块Skills的用户动作和活动（如“奇偶游戏”）调整优先级。仿生行为模块BBM接收到VAM计算出的FoA位置后，命令d中执行器模块调整机器人的运动，使其向FoA方向移动。

**用户实验方法：**

假设参与者在RJAS活动与关闭条件下互动时，是否会感知到机器人具有更高的社交存在感（SP）的情况。机器人在实验中指导用户进行游戏根据RJAS是否活动，对用户的分心做出反应。

奇偶游戏规则：“奇偶游戏”是一款基于运气的游戏，其中两个玩家在每轮开始前选择“偶数”或“奇数”。倒计时后，玩家同时展示伸出的手指，手指的总和决定胜负。如果总和为偶数，选择“偶数”的玩家获胜；否则，“奇数”玩家获胜。

Table1是在整个交互过程中出现的刺激列表，包括时间、类型和描述。

Table2是社交存在感量表（SPI）。共同存在感（Co-Presence）指的是与另一个实体共享空间时的在场感，这是SP体验的核心。注意力分配（Attentional Allocation）反映了个体在互动中对他人关注的分布程度，表示个体在与他人互动时的关注程度。还评估了信息理解、行为相互依赖、情感理解和情感相互依赖等维度。

H1～H4是响应性联合注意在机器人交互中社交存在感的作用假设

**用户实验结果：**

图9比较了在参与者和活跃/非活跃RJAS的机器人进行交互，测量自我感知以及机器人感知。

图10比较了RJAS活跃比非活跃条件下，a共同存在感和b行为相互依赖具有更高的社交存在感，c信息理解方面RJAS的激活未能带来显著的改善，d注意力分配差异较小。

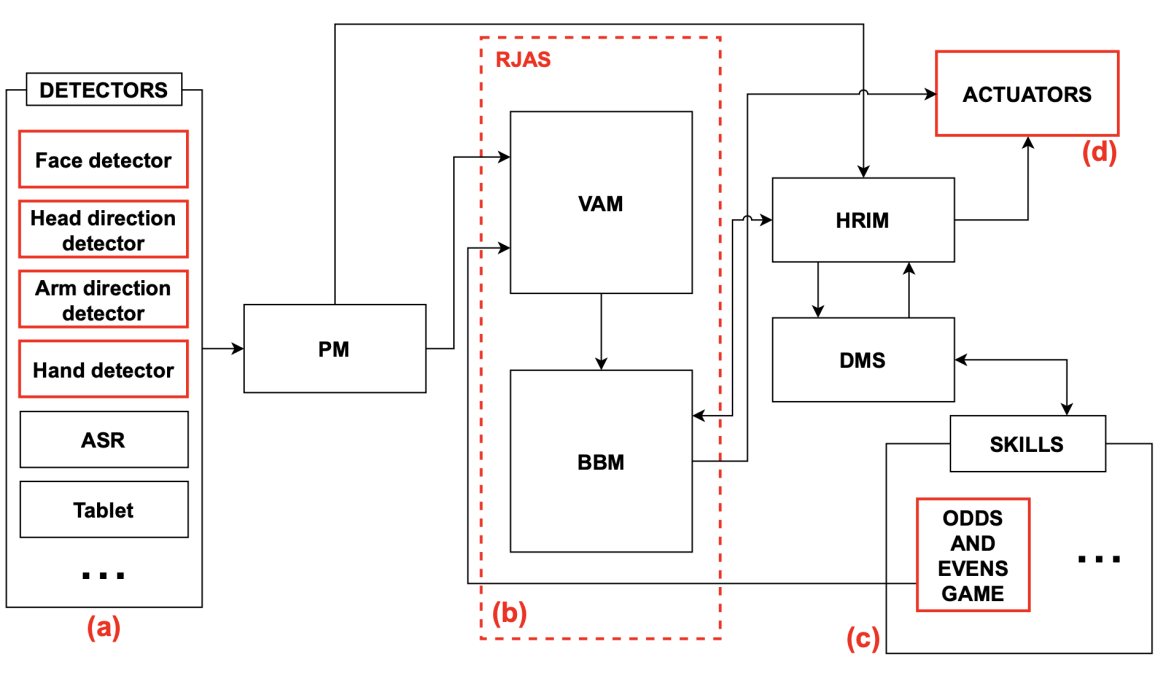
**未来工作：**

实验中去除了情感理解和情感相互依赖维度、手臂方向检测没有设计实验、本研究仅在一个机器人平台上进行测试、没有聚焦于特定人群

**Review**

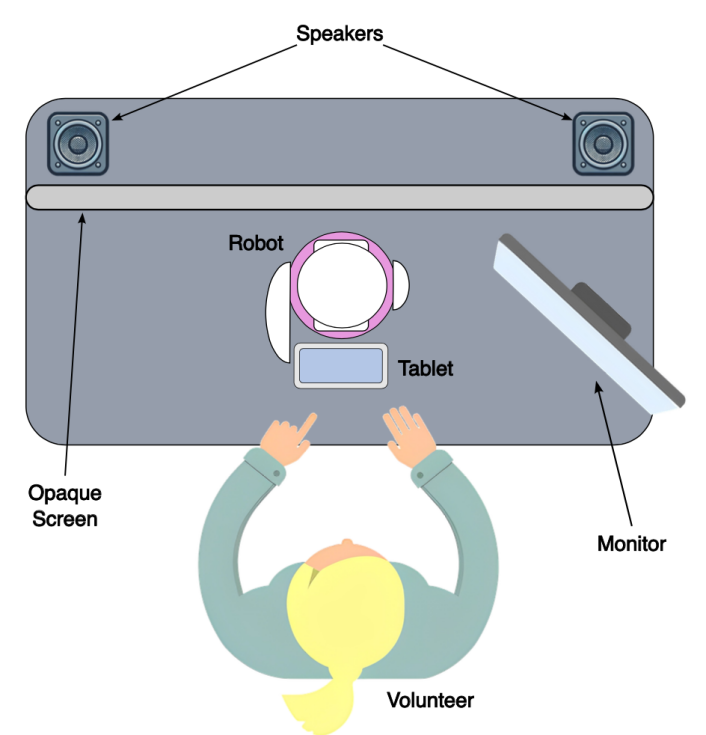
**问题1:**

图7展示了Mini软件架构中响应式联合注意系统的集成，可以标注并详细说明检测器检测到图像输入到RJAS的VAM和BBM模块的具体参数与计算公式。

****

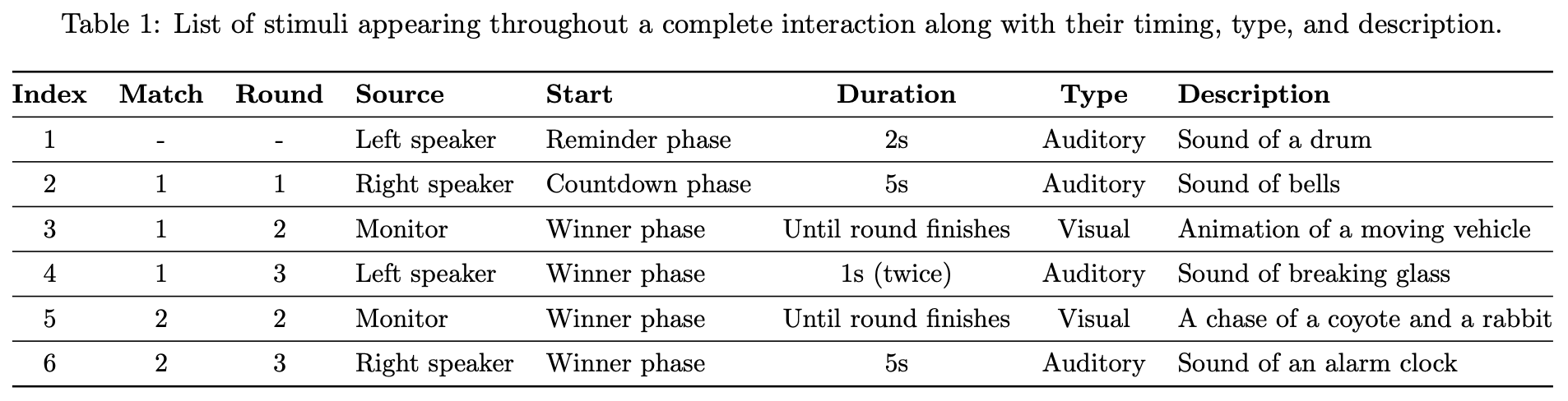
**问题2:**

图8是否可以采用参与者第一视角进行绘制，以更好地显示实验环境，包括实际场景中Mini机器人在桌子上的位置，机器人前方平板，不透明屏幕后面的两个扬声器，右侧桌面上的显示器。

****

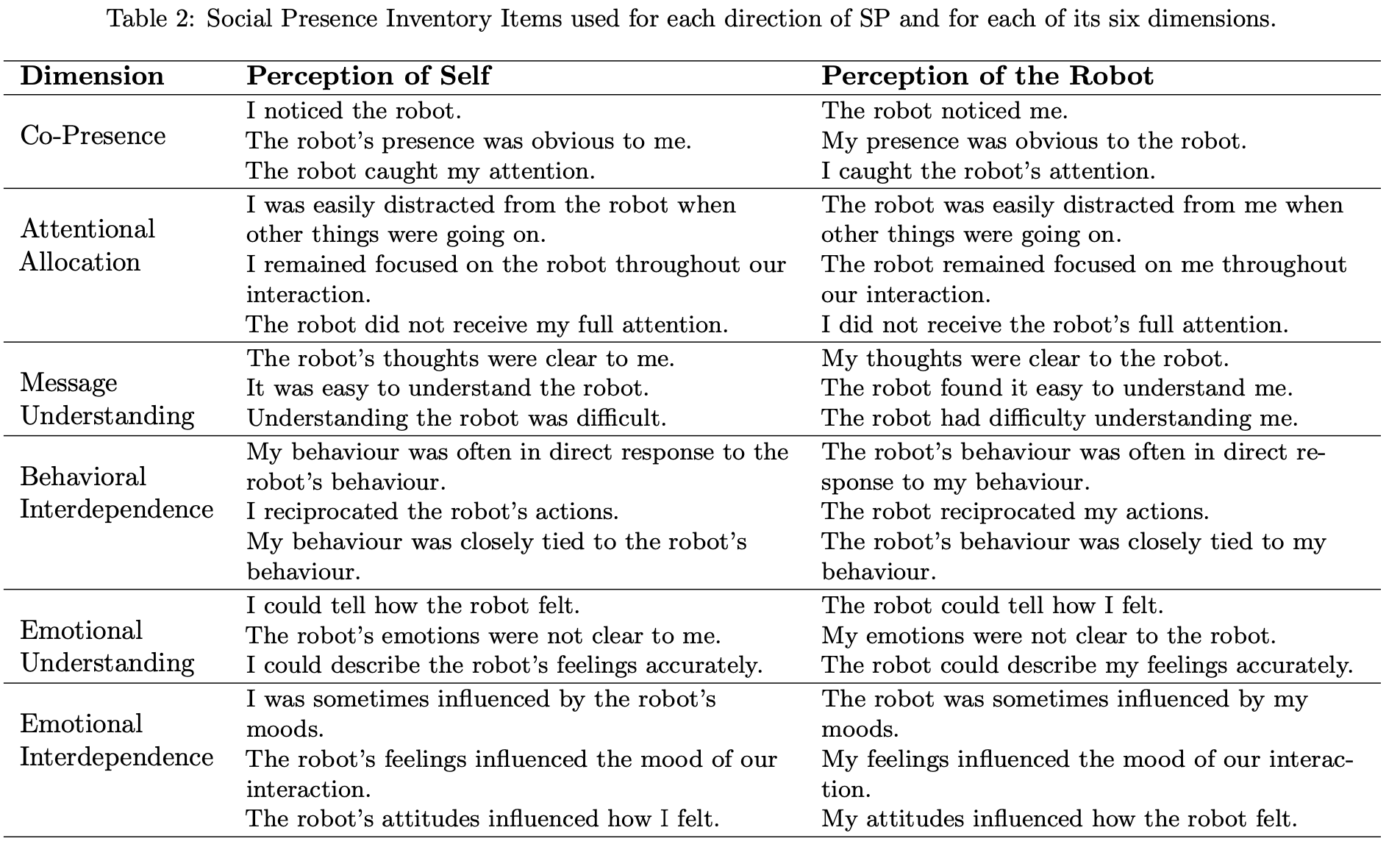
**问题3:**

实验Table1中显示交互过程中出现的是声音和视觉刺激干扰，可以考虑其他更复杂的干扰如声音与视觉同时干扰，并提高刺激持续时间。



**问题4:**

Table2讨论了多个社交存在感维度的影响，可以设计实验分析共同存在感与行为相互依赖等社交存在感维度关联性，比较在活跃/非活跃RJAS的机器人进行交互的实验结果。



**问题5:**

实验采用奇偶游戏作为互动场景，可以采用更多互动任务或者设计更复杂的交互场景，考察机器人在不同场景下的泛化能力，同时使机器人手臂方向检测功能、手部手势检测功能得到使用。