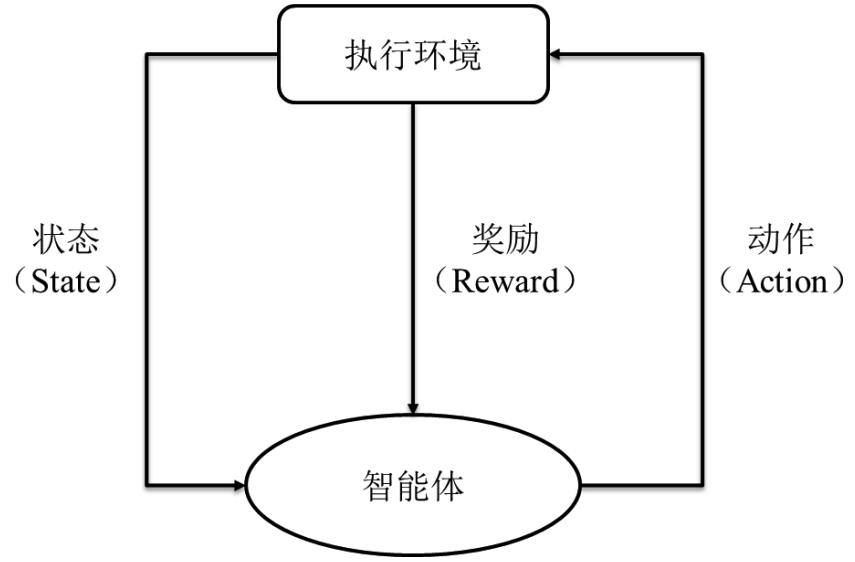

执行环境
状态
（State）
奖励
（Reward）
动作
（Action）
智能体

图 1

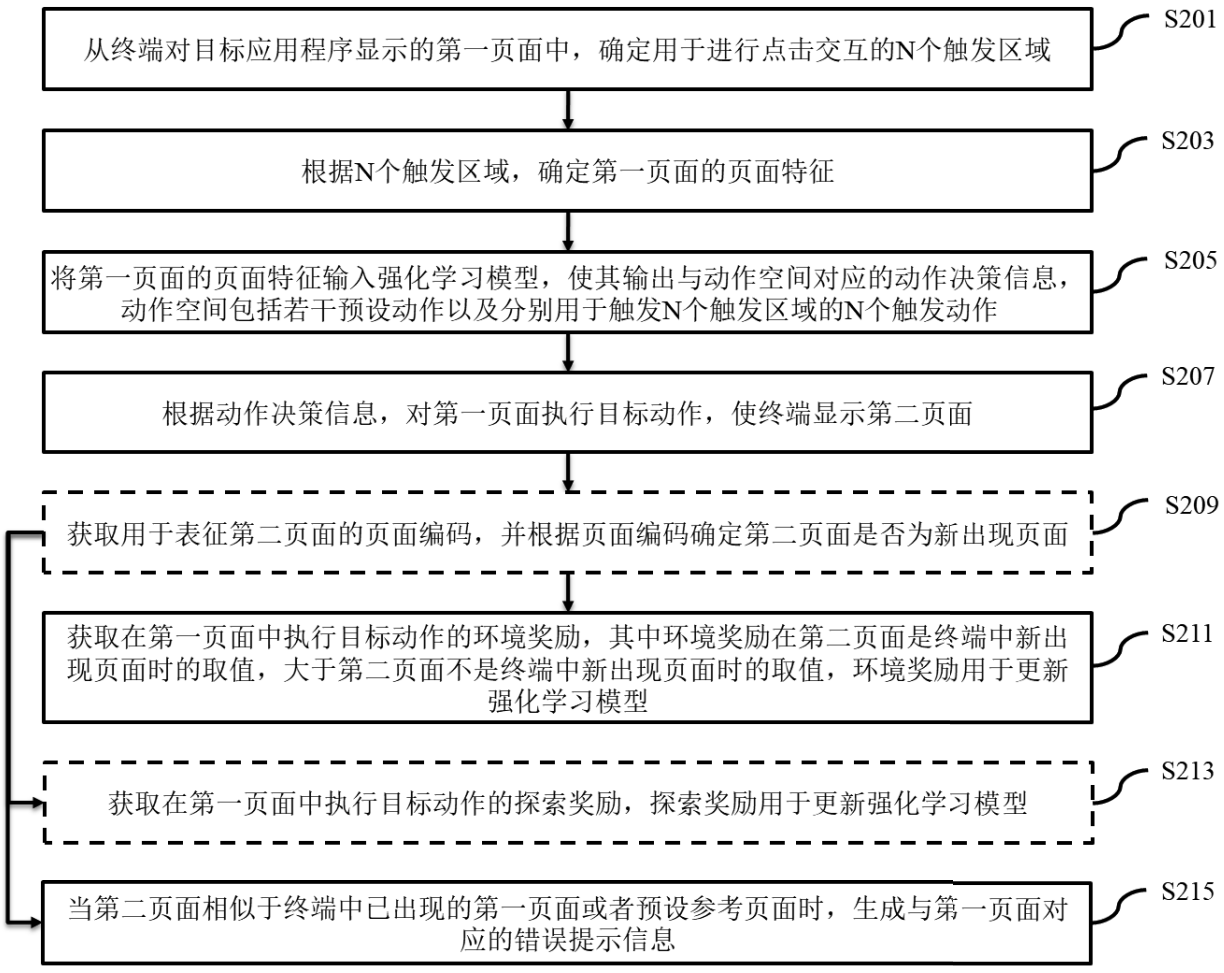

从终端对目标应用程序显示的第一页面中，确定用于进行点击交互的N个触发区域
S201
根据N个触发区域，确定第一页面的页面特征
S203
将第一页面的页面特征输入强化学习模型，使其输出与动作空间对应的动作决策信息，动作空间包括若干预设动作以及分别用于触发N个触发区域的N个触发动作
S205
根据动作决策信息，对第一页面执行目标动作，使终端显示第二页面
S207
获取用于表征第二页面的页面编码，并根据页面编码确定第二页面是否为新出现页面
S209
获取在第一页面中执行目标动作的环境奖励，其中环境奖励在第二页面是终端中新出现页面时的取值，大于第二页面不是终端中新出现页面时的取值，环境奖励用于更新强化学习模型
S211
获取在第一页面中执行目标动作的探索奖励，探索奖励用于更新强化学习模型
S213
当第二页面相似于终端中已出现的第一页面或者预设参考页面时，生成与第一页面对应的错误提示信息
S215

图 2

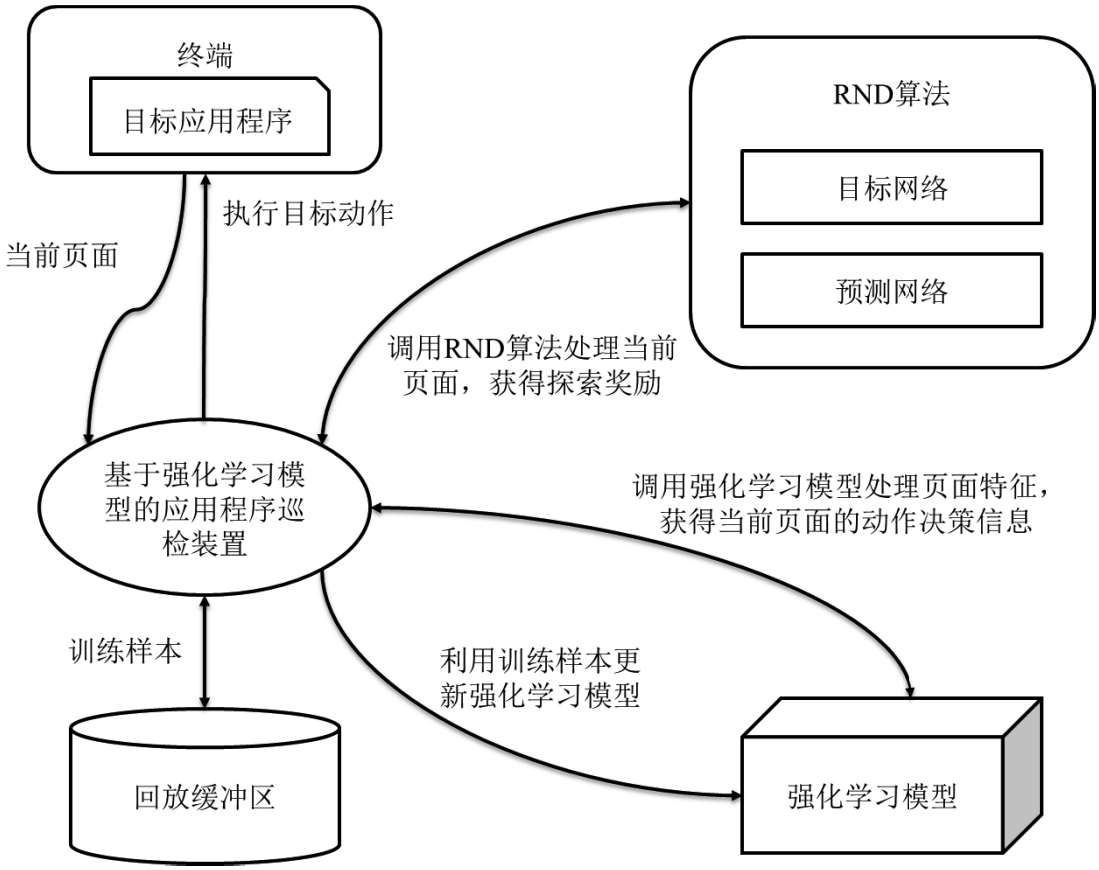

终端
目标应用程序
RND算法
目标网络
预测网络
执行目标动作
当前页面
调用RND算法处理当前
页面，获得探索奖励
基于强化学习模
型的应用程序巡
检装置
调用强化学习模型处理页面特征，
获得当前页面的动作决策信息
训练样本
利用训练样本更
新强化学习模型
回放缓冲区
强化学习模型

图 3

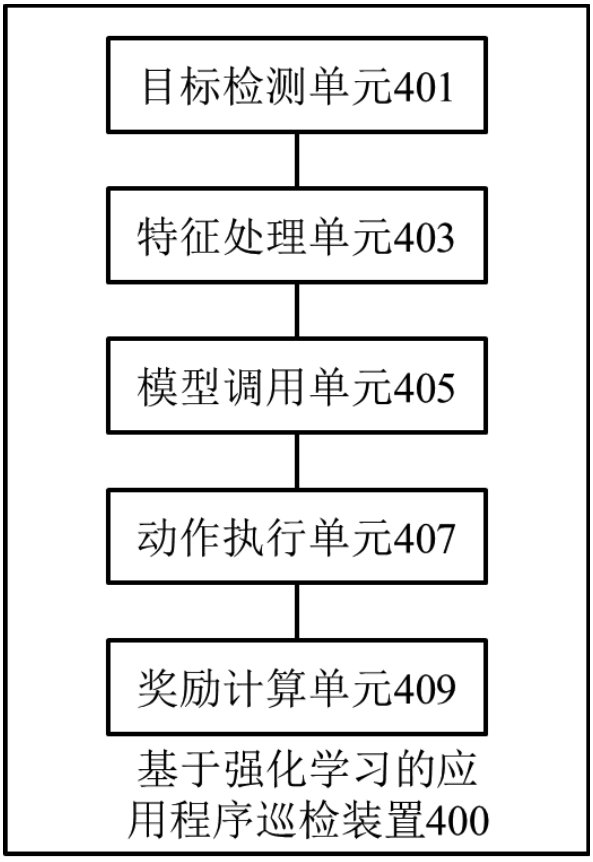

目标检测单元401
特征处理单元403
模型调用单元405
动作执行单元407
奖励计算单元409
基于强化学习的应
用程序巡检装置400

图 4

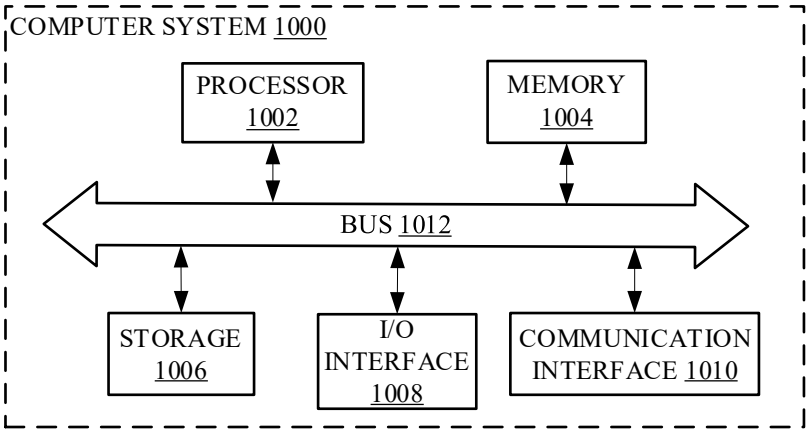

COMPUTER SYSTEM 1000
PROCESSOR 1002
MEMORY 1004
BUS 1012
STORAGE 1006
I/O INTERFACE 1008
COMMUNICATION INTERFACE 1010

图 5