



(12)发明专利申请

(10)申请公布号 CN 111600676 A

(43)申请公布日 2020.08.28

(21)申请号 202010506845.3

(22)申请日 2020.06.05

(71)申请人 上海特金无线技术有限公司

地址 201114 上海市闵行区新骏环路245号
第6层E612室

(72)发明人 李瀚 姜化京 姜维

(74)专利代理机构 上海慧晗知识产权代理事务
所(普通合伙) 31343

代理人 李茂林

(51)Int.Cl.

H04K 3/00(2006.01)

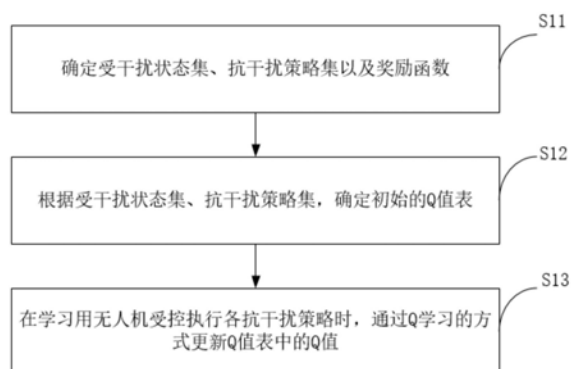
权利要求书3页 说明书15页 附图10页

(54)发明名称

Q值表确定方法、抗干扰方法、装置及设备

(57)摘要

本发明提供了一种无人机通信Q值表确定方法、无人机通信抗干扰方法、装置及设备。其中的Q值表确定方法,包括:确定受干扰状态集、抗干扰策略集以及奖励函数;根据受干扰状态集、抗干扰策略集,确定初始的Q值表;在学习用无人机受控执行各抗干扰策略时,通过Q学习的方式更新Q值表中的Q值;所更新的Q值与学习用无人机所采用的抗干扰策略、所接收到的干扰信号的功率信息与信道信息,以及奖励函数相关联。其中,不同的受干扰状态信息表征了学习用无人机接收到的干扰信号的功率信息所处的不同功率区间和/或不同的信道,可以更加全面反映外部环境对无人机通信的影响,进而所确定的抗干扰策略更加符合实际飞行情况。



1. 一种无人机通信Q值表确定方法,应用于控制端,其特征在于,包括:

确定受干扰状态集、抗干扰策略集以及奖励函数;

根据所述受干扰状态集、所述抗干扰策略集,确定初始的Q值表;

在学习用无人机受控执行各抗干扰策略时,通过Q学习的方式更新所述Q值表中的Q值;所更新的Q值与所述学习用无人机所采用的抗干扰策略、所接收到的干扰信号的功率信息与信道信息,以及所述奖励函数相关联;

其中,所述Q值表记载了每一种受干扰状态信息下执行不同抗干扰策略对应的Q值;所述受干扰状态集表征了预先定义的受干扰状态信息的集合,所述抗干扰策略集为学习用无人机可执行的抗干扰策略的集合,所述奖励函数用于表征所述学习用无人机所执行的不同抗干扰策略、执行对应抗干扰策略之前的受干扰状态信息,以及奖励值之间的函数关系;不同的所述受干扰状态信息表征了所述学习用无人机接收到的干扰信号的功率所处的不同功率区间和/或不同的信道;所述抗干扰策略被执行时,以调整所述学习用无人机的通信信道和/或通信带宽。

2. 根据权利要求1所述的方法,其特征在于,通过Q学习的方式更新所述Q值表,包括:

获取所述学习用无人机在任意第K个时隙感知的第一受干扰状态信息;

确定所述学习用无人机在第K个时隙的第一抗干扰策略;

根据所述奖励函数,确定所述学习用无人机在所述第一受干扰状态信息下执行所述第一抗干扰策略获得的第一奖励值;

获取所述学习用无人机在第K+1个时隙感知的第二受干扰状态信息;

根据第K个时隙对应的Q值表、所述第一受干扰状态信息、所述第一抗干扰策略、所述第二受干扰状态信息以及所述第一奖励值,更新Q值表,得到第K+1个时隙对应的Q值表。

3. 根据权利要求2所述的方法,其特征在于,确定所述学习用无人机在第K个时隙的第一抗干扰策略,包括:

根据第K个时隙对应的Q值表,确定与所述第一受干扰状态信息对应的抗干扰策略子集;

确定所述抗干扰策略子集中Q值最大的抗干扰策略作为所述第一抗干扰策略。

4. 根据权利要求2所述的方法,其特征在于,确定所述学习用无人机在第K个时隙的第一抗干扰策略,包括:

获取所述学习用无人机在第K-1个时隙感知的第三受干扰状态信息;

根据第K-1个时隙对应的Q值表,确定与所述第三受干扰状态信息对应的抗干扰策略子集;

根据所述抗干扰策略子集中的每个抗干扰策略对应的Q值,通过玻尔兹曼随机策略确定所述第一抗干扰策略。

5. 根据权利要求2-4任意一项所述的方法,其特征在于,根据所述奖励函数,确定所述学习用无人机在所述第一受干扰状态信息下执行所述第一抗干扰策略获得的第一奖励值之前,还包括:

确定所述学习用无人机在所述第一受干扰状态信息下执行所述第一抗干扰策略之后,所述学习用无人机接收通信信号的信噪比;

若所述信噪比小于或等于预设的所述学习用无人机的信号解调门限,则:确定所述奖

励函数的函数值为一预设奖励值；

若所述信噪比大于预设的所述信号解调门限，则：根据信号吞吐量与信道切换开销确定所述奖励函数；所述信号吞吐量表征所述学习用无人机接收所述通信信号的能力，所述信道切换开销表征所述学习用无人机切换通信信道产生的损失。

6. 根据权利要求5所述的方法，其特征在于，所述信噪比根据以下公式确定：

$$\eta(a_t, c_t, B_t^1, B_t^2) = \frac{P_1(t)}{N_0 B_t^1 + P_2(t)};$$

其中：

t表示时隙；

$\eta(a_t, c_t, B_t^1, B_t^2)$ 表示在第t个时隙的信噪比；

a_t 表示在第t个时隙学习用无人机的通信信道；

c_t 表示在第t个时隙干扰信号信道；

B_t^1 表示在第t个时隙学习用无人机的通信信道的带宽；

B_t^2 表示在第t个时隙干扰信号带宽；

$P_1(t)$ 表示在第t个时隙学习用无人机接收到的通信信号功率；

$P_2(t)$ 表示在第t个时隙学习用无人机接收到的干扰信号的功率。

7. 一种无人机通信抗干扰方法，其特征在于，包括：

确定所述无人机在当前时刻的受干扰状态信息，不同的受干扰状态信息表征了对应无人机接收到的干扰信号的功率所处的不同功率区间和/或不同的信道；

根据所述当前时刻的受干扰状态信息以及根据权利要求1-6任意一项所述的方法确定的Q值表，确定所述无人机在当前时刻所需执行的抗干扰策略；

执行所述当前时刻所需执行的抗干扰策略，以调整所述无人机的通信信道和/或通信带宽。

8. 一种无人机通信Q值表确定装置，其特征在于，包括：

第一模块，用于确定受干扰状态集、抗干扰策略集以及奖励函数；

第二模块，用于根据所述受干扰状态集、所述抗干扰策略集，确定初始的Q值表；

第三模块，用于在学习用无人机受控执行各抗干扰策略时，通过Q学习的方式更新所述Q值表中的Q值；所更新的Q值与所述学习用无人机所采用的抗干扰策略、所接收到的干扰信号的功率信息与信道信息，以及所述奖励函数相关联；

其中，所述Q值表记载了每一种受干扰状态信息下执行不同抗干扰策略对应的Q值；所述受干扰状态集表征了预先定义的受干扰状态信息的集合，所述抗干扰策略集为学习用无人机可执行的抗干扰策略的集合，所述奖励函数用于表征所述学习用无人机所执行的不同抗干扰策略、执行对应抗干扰策略之前的受干扰状态信息，以及奖励值之间的函数关系；不同的所述受干扰状态信息表征了所述学习用无人机接收到的干扰信号的功率所处的不同功率区间和/或不同的信道；所述抗干扰策略被执行时，以调整所述学习用无人机的通信信道和/或通信带宽。

9. 根据权利要求8所述的装置，其特征在于，所述第三模块，包括：

第一信息获取单元,用于获取所述学习用无人机在任意第K个时隙感知的第一受干扰状态信息;

第一策略确定单元,用于确定所述学习用无人机在第K个时隙的第一抗干扰策略;

第一计算单元,用于根据所述奖励函数,确定所述学习用无人机在所述第一受干扰状态信息下执行所述第一抗干扰策略获得的第一奖励值;

第二信息获取单元,用于获取所述学习用无人机在第K+1个时隙感知的第二受干扰状态信息;

第二计算单元,用于根据第K个时隙对应的Q值表、所述第一受干扰状态信息、所述第一抗干扰策略、所述第二受干扰状态信息以及所述第一奖励值,更新Q值表,得到第K+1个时隙对应的Q值表。

10. 根据权利要求9所述的装置,其特征在于,所述第一策略确定单元,包括:

第一策略子集确定子单元,用于根据第K个时隙对应的Q值表,确定与所述第一受干扰状态信息对应的抗干扰策略子集;

第一策略选择子单元,用于确定所述抗干扰策略子集中Q值最大的抗干扰策略作为所述第一抗干扰策略。

11. 根据权利要求9所述的装置,其特征在于,所述第一策略确定单元,包括:

信息获取子单元,用于获取所述学习用无人机在第K-1个时隙感知的第三受干扰状态信息;

第二策略子集确定子单元,用于根据第K-1个时隙对应的Q值表,确定与所述第三受干扰状态信息对应的抗干扰策略子集;

第二策略选择子单元,用于根据所述抗干扰策略子集中的每个抗干扰策略对应的Q值,通过玻尔兹曼随机策略确定所述第一抗干扰策略。

12. 根据权利要求9-11任意一项所述的装置,其特征在于,还包括:

第三计算单元,用于确定所述学习用无人机在所述第一受干扰状态信息下执行所述第一抗干扰策略之后,所述学习用无人机接收通信信号的信噪比;

若所述信噪比小于或等于预设的所述学习用无人机的信号解调门限,则:确定所述奖励函数的函数值为一预设奖励值;

若所述信噪比大于预设的所述信号解调门限,则:根据信号吞吐量与信道切换开销确定所述奖励函数;所述信号吞吐量表征所述学习用无人机接收所述通信信号的能力,所述信道切换开销表征所述学习用无人机切换通信信道产生的损失。

13. 一种电子设备,其特征在于,包括处理器与存储器,

所述存储器,用于存储代码和相关数据;

所述处理器,用于执行所述存储器中的代码用以实现权利要求1-6任一项所述的方法。

14. 一种存储介质,其上存储有计算机程序,该程序被处理器执行时实现权利要求1-6任一项所述的方法。

Q值表确定方法、抗干扰方法、装置及设备

技术领域

[0001] 本发明涉及无人机领域,并且更具体地,涉及一种无人机通信Q值表确定方法、无人机抗干扰方法、装置及设备。

背景技术

[0002] 随着无人机技术的不断发展,越来越多的消费级无人机被应用在普通人的日常生活。

[0003] 无人机在飞行过程中,恶意干扰对无人机通信安全威胁极大,为了应对恶意干扰威胁,现有相关技术中,当无人机通信受到干扰时,通过感知干扰信号的信道,之后将无人机的通信信道切换至与干扰信号的信道不同的信道,这样可以避免干扰信号干扰无人机通信。

[0004] 可见,现有技术中,仅通过获取无人机外部环境中的干扰信号的信道,不能全面反映出外部环境对无人机通信的影响,并且,现有相关技术仅通过切换无人机通信信道的方式抗干扰,抗干扰策略比较单一,灵活性比较差。

发明内容

[0005] 本方面提供了一种无人机通信Q值表确定方法、抗干扰方法、装置及设备,以解决现有的无人机抗干扰技术不能全面反映外部环境对无人机通信的影响以及抗干扰策略比较单一、灵活性差的问题。

[0006] 根据本发明的第一方面,提供了一种无人机通信Q值表确定方法,应用于控制端,包括:

[0007] 确定受干扰状态集、抗干扰策略集以及奖励函数;

[0008] 根据所述受干扰状态集、所述抗干扰策略集,确定初始的Q值表;

[0009] 在学习用无人机受控执行各抗干扰策略时,通过Q学习的方式更新所述Q值表中的Q值;所更新的Q值与所述学习用无人机所采用的抗干扰策略、所接收到的干扰信号的功率信息与信道信息,以及所述奖励函数相关联;

[0010] 其中,所述Q值表记载了每一种受干扰状态信息下执行不同抗干扰策略对应的Q值;所述受干扰状态集表征了预先定义的受干扰状态信息的集合,所述抗干扰策略集为学习用无人机可执行的抗干扰策略的集合,所述奖励函数用于表征所述学习用无人机所执行的不同抗干扰策略、执行对应抗干扰策略之前的受干扰状态信息,以及奖励值之间的函数关系;不同的所述受干扰状态信息表征了所述学习用无人机接收到的干扰信号的功率所处的不同功率区间和/或不同的信道;所述抗干扰策略被执行时,以调整所述学习用无人机的通信信道和/或通信带宽。

[0011] 可选的,通过Q学习的方式更新所述Q值表,包括:

[0012] 获取所述学习用无人机在任意第K个时隙感知的第一受干扰状态信息;

[0013] 确定所述学习用无人机在第K个时隙的第一抗干扰策略;

[0014] 根据所述奖励函数,确定所述学习用无人机在所述第一受干扰状态信息下执行所述第一抗干扰策略获得的第一奖励值;

[0015] 获取所述学习用无人机在第K+1个时隙感知的第二受干扰状态信息;

[0016] 根据第K个时隙对应的Q值表、所述第一受干扰状态信息、所述第一抗干扰策略、所述第二受干扰状态信息以及所述第一奖励值,更新Q值表,得到第K+1个时隙对应的Q值表。

[0017] 可选的,确定所述学习用无人机在第K个时隙的第一抗干扰策略,包括:

[0018] 根据第K个时隙对应的Q值表,确定与所述第一受干扰状态信息对应的抗干扰策略子集;

[0019] 确定所述抗干扰策略子集中Q值最大的抗干扰策略作为所述第一抗干扰策略。

[0020] 可选的,确定所述学习用无人机在第K个时隙的第一抗干扰策略,包括:

[0021] 获取所述学习用无人机在第K-1个时隙感知的第三受干扰状态信息;

[0022] 根据第K-1个时隙对应的Q值表,确定与所述第三受干扰状态信息对应的抗干扰策略子集;

[0023] 根据所述抗干扰策略子集中的每个抗干扰策略对应的Q值,通过玻尔兹曼随机策略确定所述第一抗干扰策略。

[0024] 可选的,根据所述奖励函数,确定所述学习用无人机在所述第一受干扰状态信息下执行所述第一抗干扰策略获得的第一奖励值之前,还包括:

[0025] 确定所述学习用无人机在所述第一受干扰状态信息下执行所述第一抗干扰策略之后,所述学习用无人机接收通信信号的信噪比;

[0026] 若所述信噪比小于或等于预设的所述学习用无人机的信号解调门限,则:确定所述奖励函数的函数值为一预设奖励值;

[0027] 若所述信噪比大于预设的所述信号解调门限,则:根据信号吞吐量与信道切换开销确定所述奖励函数;所述信号吞吐量表征所述学习用无人机接收所述通信信号的能力,所述信道切换开销表征所述学习用无人机切换通信信道产生的损失。

[0028] 可选的,所述信噪比根据以下公式确定:

$$[0029] \quad \eta(a_t, c_t, B_t^1, B_t^2) = \frac{P_1(t)}{N_0 B_t^1 + P_2(t)};$$

[0030] 其中:

[0031] t表示时隙;

[0032] $\eta(a_t, c_t, B_t^1, B_t^2)$ 表示在第t个时隙的信噪比;

[0033] a_t 表示在第t个时隙学习用无人机的通信信道;

[0034] c_t 表示在第t个时隙干扰信号信道;

[0035] B_t^1 表示在第t个时隙学习用无人机的通信信道的带宽;

[0036] B_t^2 表示在第t个时隙干扰信号带宽;

[0037] $P_1(t)$ 表示在第t个时隙学习用无人机接收到的通信信号功率;

[0038] $P_2(t)$ 表示在第t个时隙学习用无人机接收到的干扰信号的功率。

[0039] 根据本发明的第二方面,提供了一种无人机通信抗干扰方法,包括:

[0040] 确定所述无人机在当前时刻的受干扰状态信息,不同的受干扰状态信息表征了对应无人机接收到的干扰信号的功率所处的不同功率区间和/或不同的信道;

[0041] 根据所述当前时刻的受干扰状态信息以及根据本发明第一方面及其可选方案涉及的方法确定的Q值表,确定所述无人机在当前时刻所需执行的抗干扰策略;

[0042] 执行所述当前时刻所需执行的抗干扰策略,以调整所述无人机的通信信道和/或通信带宽。

[0043] 根据本发明的第三方面,提供了一种无人机通信Q值表确定装置,包括:

[0044] 第一模块,用于确定受干扰状态集、抗干扰策略集以及奖励函数;

[0045] 第二模块,用于根据所述受干扰状态集、所述抗干扰策略集,确定初始的Q值表;

[0046] 第三模块,用于在学习用无人机受控执行各抗干扰策略时,通过Q学习的方式更新所述Q值表中的Q值;所更新的Q值与所述学习用无人机所采用的抗干扰策略、所接收到的干扰信号的功率信息与信道信息,以及所述奖励函数相关联;

[0047] 其中,所述Q值表记载了每一种受干扰状态信息下执行不同抗干扰策略对应的Q值;所述受干扰状态集表征了预先定义的受干扰状态信息的集合,所述抗干扰策略集为学习用无人机可执行的抗干扰策略的集合,所述奖励函数用于表征所述学习用无人机所执行的不同抗干扰策略、执行对应抗干扰策略之前的受干扰状态信息,以及奖励值之间的函数关系;不同的所述受干扰状态信息表征了所述学习用无人机接收到的干扰信号的功率所处的不同功率区间和/或不同的信道;所述抗干扰策略被执行时,以调整所述学习用无人机的通信信道和/或通信带宽。

[0048] 可选的,所述第三模块,包括:

[0049] 第一信息获取单元,用于获取所述学习用无人机在任意第K个时隙感知的第一受干扰状态信息;

[0050] 第一策略确定单元,用于确定所述学习用无人机在第K个时隙的第一抗干扰策略;

[0051] 第一计算单元,用于根据所述奖励函数,确定所述学习用无人机在所述第一受干扰状态信息下执行所述第一抗干扰策略获得的第一奖励值;

[0052] 第二信息获取单元,用于获取所述学习用无人机在第K+1个时隙感知的第二受干扰状态信息;

[0053] 第二计算单元,用于根据第K个时隙对应的Q值表、所述第一受干扰状态信息、所述第一抗干扰策略、所述第二受干扰状态信息以及所述第一奖励值,更新Q值表,得到第K+1个时隙对应的Q值表。

[0054] 可选的,所述第一策略确定单元,包括:

[0055] 第一策略子集确定子单元,用于根据第K个时隙对应的Q值表,确定与所述第一受干扰状态信息对应的抗干扰策略子集;

[0056] 第一策略选择子单元,用于确定所述抗干扰策略子集中Q值最大的抗干扰策略作为所述第一抗干扰策略。

[0057] 可选的,所述第一策略确定单元,包括:

[0058] 信息获取子单元,用于获取所述学习用无人机在第K-1个时隙感知的第三受干扰状态信息;

[0059] 第二策略子集确定子单元,用于根据第K-1个时隙对应的Q值表,确定与所述第三

受干扰状态信息对应的抗干扰策略子集；

[0060] 第二策略选择子单元,用于根据所述抗干扰策略子集中的每个抗干扰策略对应的Q值,通过玻尔兹曼随机策略确定所述第一抗干扰策略。

[0061] 可选的,该装置还包括:

[0062] 第三计算单元,用于确定所述学习用无人机在所述第一受干扰状态信息下执行所述第一抗干扰策略之后,所述学习用无人机接收通信信号的信噪比;

[0063] 若所述信噪比小于或等于预设的所述学习用无人机的信号解调门限,则:确定所述奖励函数的函数值为一预设奖励值;

[0064] 若所述信噪比大于预设的所述信号解调门限,则:根据信号吞吐量与信道切换开销确定所述奖励函数;所述信号吞吐量表征所述学习用无人机接收所述通信信号的能力,所述信道切换开销表征所述学习用无人机切换通信信道产生的损失。

[0065] 根据本发明的第四方面,提供了一种电子设备,包括处理器与存储器,所述存储器,用于存储代码和相关数据;

[0066] 所述处理器,用于执行所述存储器中的代码用以实现本发明第一方面及其可选方案涉及的方法。

[0067] 根据本发明的第五方面,提供了一种存储介质,其上存储有计算机程序,该程序被处理器执行时实现本发明第一方面及其可选方案涉及的方法。

[0068] 本发明提供的无人机通信Q值表确定方法、无人机通信抗干扰方法、装置及设备,其中确定的Q值表包括每一种受干扰状态信息下执行不同抗干扰策略对应的Q值,由于不同的受干扰状态信息表征了学习用无人机接收到的干扰信号的功率信息所处的不同功率区间和/或不同的信道,相对现有技术,能够更加全面反映外部环境对无人机通信的影响,进而所确定的抗干扰策略更加符合实际飞行情况;由于抗干扰策略被执行时选择调整无人机的通信信道和/或通信带宽,使得抗干扰策略的选择更加灵活。

附图说明

[0069] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动性的前提下,还可以根据这些附图获得其他的附图。

[0070] 图1是本发明一实施例中确定Q值表的学习场景图;

[0071] 图2是本发明一实施例中无人机通信Q值表确定方法的流程图一;

[0072] 图3是本发明一实施例中无人机通信Q值表确定方法的流程图二;

[0073] 图4是本发明一实施例中无人机通信Q值表确定方法的流程图三;

[0074] 图5是本发明一实施例中无人机通信Q值表确定方法的流程图四;

[0075] 图6是本发明一实施例中无人机通信Q值表确定方法的流程图五;

[0076] 图7a是干扰的增益系数为0时的示意图。

[0077] 图7b是干扰的增益系数为1时的示意图。

[0078] 图7c是干扰的增益系数为 $\frac{B_l^1}{B_l^2}$ 时的示意图。

- [0079] 图8是本发明一实施例中无人机飞行的场景示意图；
- [0080] 图9是本发明一实施例中无人机通信抗干扰方法的流程图；
- [0081] 图10是本发明一实施例中无人机通信抗干扰装置的模块示意图；
- [0082] 图11是本发明一实施例中无人机通信Q值表确定装置的模块示意图一；
- [0083] 图12是本发明一实施例中无人机通信Q值表确定装置的模块示意图二；
- [0084] 图13是本发明一实施例中无人机通信Q值表确定装置的模块示意图三；
- [0085] 图14是本发明一实施例中无人机通信Q值表确定装置的模块示意图四；
- [0086] 图15是本发明一实施例中无人机通信Q值表确定装置的模块示意图五；
- [0087] 图16是本发明一实施例中电子设备的构造示意图。
- [0088] 附图标记说明：
- [0089] 11-学习用无人机；
- [0090] 12-通信用无人机；
- [0091] 13-干扰无人机；
- [0092] 14-控制端；
- [0093] 21-无人机；
- [0094] 22-干扰源；
- [0095] 23-通信站；
- [0096] 31-状态获取模块；
- [0097] 32-策略确定模块；
- [0098] 33-策略执行模块；
- [0099] 41-第一模块；
- [0100] 42-第二模块；
- [0101] 43-第三模块；
- [0102] 431-第一信息获取单元；
- [0103] 432-第一策略确定单元；
- [0104] 4321-第一策略子集确定子单元；
- [0105] 4322-第一策略选择子单元；
- [0106] 4323-信息获取子单元；
- [0107] 4324-第二策略子集确定子单元；
- [0108] 4325-第二策略选择子单元；
- [0109] 433-第一计算单元；
- [0110] 434-第二信息获取单元；
- [0111] 435-第二计算单元；
- [0112] 436-第三计算单元；
- [0113] 51-处理器；
- [0114] 52-总线；
- [0115] 53-存储器。

具体实施方式

[0116] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0117] 本发明的说明书和权利要求书及上述附图中的术语“第一”、“第二”、“第三”“第四”等(如果存在)是用于区别类似的对象,而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换,以便这里描述的本发明的实施例能够以除了在这里图示或描述的那些以外的顺序实施。此外,术语“包括”和“具有”以及他们的任何变形,意图在于覆盖不排他的包含,例如,包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元,而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

[0118] 下面以具体地实施例对本发明的技术方案进行详细说明。下面这几个具体的实施例可以相互结合,对于相同或相似的概念或过程可能在某些实施例不再赘述。

[0119] 为了便于描述,下述先对本申请实施例涉及的部分名词进行简介:

[0120] 奖励函数:在本申请中奖励函数用于表征无人机执行抗干扰策略后获得的直接反馈(也即基于奖励函数确定的奖励值)。进一步来说,奖励函数用于表征无人机所执行的不同抗干扰策略、执行对应抗干扰策略之前的受干扰状态信息,以及奖励值之间的函数关系。

[0121] Q值:Q值根据Q(Quality)函数确定,本申请中Q值表征无人机在一种受干扰状态(信息)下,执行某一抗干扰策略,根据Q函数确定的反馈,其中,每一种受干扰状态对应的一系列抗干扰策略中,对应Q值最大的抗干扰策略为该种受干扰状态下最优的抗干扰策略。

[0122] Q值表:Q值表可以理解为每一种受干扰状态信息下执行每一抗干扰策略对应的Q值的集合,即上述Q值的集合。

[0123] 图1是本发明一实施例中确定Q值表的学习场景图。

[0124] 请参考图1,该场景所关联的设备可包括:学习用无人机11、通信用无人机12、干扰无人机13以及控制端14。

[0125] 需要说明的是,本发明实施例中通过强化学习的方式确定Q值表,本发明实施例中的学习用无人机11处于动态时变的受干扰环境中,针对动态时变环境中的抗干扰问题,本发明实施例通过马尔科夫决策过程建模,为设计抗干扰算法提供理论支撑。

[0126] 请参考图1,学习用无人机11与通信用无人机12可进行实时通信,干扰无人机13用于发射干扰信号,干扰信号能够影响学习用无人机11与通信用无人机12之间的通信。学习用无人机11可接收控制端14的控制指令,该控制指令进一步可包括通信抗干扰指令。控制端14可以是学习用无人机11外部的控制端,控制端14也可以集成在学习用无人机11内部。

[0127] 图2是本发明一实施例中无人机通信Q值表确定方法的流程图一。

[0128] 请参考图2,一种实施方式中,无人机通信Q值表确定方法,应用于控制端14,包括:

[0129] S11:确定受干扰状态集、抗干扰策略集以及奖励函数;

[0130] S12:根据受干扰状态集、抗干扰策略集,确定初始的Q值表;

[0131] S13:在学习用无人机11受控执行各抗干扰策略时,通过Q学习的方式更新Q值表中的Q值;所更新的Q值与学习用无人机11所采用的抗干扰策略、所接收到的干扰信号的功率

信息与信道信息,以及奖励函数相关联;

[0132] 其中,Q值表记载了每一种受干扰状态信息下执行不同抗干扰策略对应的Q值;受干扰状态集表征了预先定义的受干扰状态信息的集合,抗干扰策略集为学习用无人机11可执行的抗干扰策略的集合,奖励函数用于表征学习用无人机11所执行的不同抗干扰策略、执行对应抗干扰策略之前的受干扰状态信息,以及奖励值之间的函数关系;不同的受干扰状态信息表征了学习用无人机11接收到的干扰信号的功率信息所处的不同功率区间和/或不同的信道;抗干扰策略被执行时,以调整学习用无人机11的通信信道和/或通信带宽。

[0133] 本发明实施例中的无人机通信Q值表确定方法,其中确定的Q值表包括每一种受干扰状态信息下执行不同抗干扰策略对应的Q值,由于不同的受干扰状态信息表征了学习用无人机11接收到的干扰信号的功率信息所处的不同功率区间和/或不同的信道,相对现有技术,能够更加全面反映外部环境对无人机通信的影响,进而所确定的抗干扰策略更加符合实际飞行情况;由于抗干扰策略被执行时选择调整无人机的通信信道和/或通信带宽,使得抗干扰策略的选择更加灵活。

[0134] 本发明实施例中,功率区间可以理解为依据一设定的功率间隔划分的区间,例如:将学习用无人机11接收到的干扰信号的功率划分为L个连续不重叠的范围,则对应L种接收功率区间(L种接收功率区间也可以理解为L种功率等级),定义L个连续不重叠的范围依次为:

[0135] $h_1 < h_2 < \dots < h_l < h_{l+1} < \dots < h_{L+1}$;

[0136] 当 $P_l(t) \in [h_l, h_{l+1})$ 时,对应的接收功率区间为l, h_l 为对应的l功率区间对应的下边界, h_{l+1} 对应的l功率区间对应的上边界,其中, $1 \leq l \leq L$, $P_l(t)$ 为学习用无人机11在t时隙接收的干扰信号的功率。

[0137] 本发明实施例中,预定义的受干扰状态集可表征为:

[0138] $S = \{s_1(l_1, c_1), s_2(l_2, c_2) \dots s_X(l_X, c_X)\}$;

[0139] 其中,

[0140] S的大小(即S包括的受干扰状态信息数)为 $X = L \times N$;N为干扰信号的信道数;

[0141] s_X 表示第X个受干扰状态信息;

[0142] l_X 表示第X个受干扰状态信息对应的功率区间;

[0143] c_X 表示第X个受干扰状态信息对应的信道。

[0144] 本发明实施例中,抗干扰策略可以是:学习用无人机11通信信号的信道与带宽的组合。抗干扰策略集为所有可能的抗干扰策略的集合。

[0145] 抗干扰策略集可表征为:

[0146] $G = \{g_1(B_1, a_1), g_2(B_2, a_2) \dots g_Y(B_Y, a_Y)\}$;

[0147] 其中,

[0148] G的大小(即G包括的抗干扰策略数)为 $Y = B \times M$,B为学习用无人机11可调的带宽数,M为学习用无人机11的通信信道数;

[0149] g_Y 表示第Y个抗干扰策略;

[0150] B_Y 表示第Y个抗干扰策略对应学习用无人机11通信信号的带宽;

[0151] a_Y 表示第Y个抗干扰策略对应学习用无人机11通信信号的信道。

[0152] 其他实施方式中,抗干扰策略还可以是:表征学习用无人机11调整当前通信信号

的信道和带宽的动作,例如,抗干扰策略 g_1 可表征为:

[0153] $g_1 = g((B_1 \rightarrow B_2), (a_1 \rightarrow a_2))$, 抗干扰策略 g_1 表示学习用无人机11需要将当前信道 B_1 调整至信道 B_2 ,将当前带宽 a_1 调整至带宽 a_2 ;此种情况下, G 的大小为 $Y = B^2 \times M^2$ 。

[0154] 初始的 Q 值表可以是全0矩阵,大小为 $X \times Y$,具体可表征为:

$$[0155] \quad Q_0 = \begin{bmatrix} 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix}.$$

[0156] 图3是本发明一实施例中无人机通信 Q 值表确定方法的流程图二。

[0157] 请参考图3,一种实施方式中,通过 Q 学习的方式更新 Q 值表,即步骤S13,包括:

[0158] S131:获取学习用无人机11在任意第 K 个时隙感知的第一受干扰状态信息;

[0159] S132:确定学习用无人机11在第 K 个时隙的第一抗干扰策略;

[0160] S133:根据奖励函数,确定学习用无人机11在第一受干扰状态信息下执行第一抗干扰策略获得的第一奖励值;

[0161] S134:获取学习用无人机11在第 $K+1$ 个时隙感知的第二受干扰状态信息;

[0162] S135:根据第 K 个时隙对应的 Q 值表、第一受干扰状态信息、第一抗干扰策略、第二受干扰状态信息以及第一奖励值,更新 Q 值表,得到第 $K+1$ 个时隙对应的 Q 值表。

[0163] Q 值的更新公式可表征为:

$$[0164] \quad Q_{k+1}(s_k, g_k) = Q_k(s_k, g_k) + \tau (R_k + \gamma Q_{k+1} - Q_k(s_k, g_k));$$

$$[0165] \quad \chi_{k+1} = \max Q_k(s_{k+1}, \bar{g});$$

[0166] 其中,

[0167] Q_k 表示第 K 个时隙对应的 Q 值表;

[0168] Q_{k+1} 表示第 $K+1$ 个时隙对应的 Q 值表;

[0169] s_k 表示第一受干扰状态信息;

[0170] g_k 表示第一抗干扰策略;

[0171] R_k 表示第一奖励值;

[0172] s_{k+1} 表示第二受干扰状态信息;

[0173] $\tau \in (0, 1]$,表示学习步长,用于控制 Q 学习的收敛速度;

[0174] $\gamma \in (0, 1]$,表示折扣因子,用于评价未来回报值对当前决策的影响权重;

[0175] \bar{g} 表示状态 s_{k+1} 下所有可能的抗干扰策略。

[0176] 图4是本发明一实施例中无人机通信 Q 值表确定方法的流程图三。

[0177] 请参考图4,一种实施方式中,确定学习用无人机11在第 K 个时隙的第一抗干扰策略,即步骤S132,包括:

[0178] S1321:根据第 K 个时隙对应的 Q 值表,确定与第一受干扰状态信息对应的抗干扰策略子集;

[0179] S1322:确定抗干扰策略子集中 Q 值最大的抗干扰策略作为第一抗干扰策略。

[0180] 确定与第一受干扰状态信息对应的抗干扰策略子集,可以理解为根据 Q_k 确定第一受干扰状态信息下所有可能的抗干扰策略的集合。

[0181] 图5是本发明一实施例中无人机通信Q值表确定方法的流程图四。

[0182] 请参考图5,一种实施方式中,确定学习用无人机11在第K个时隙的第一抗干扰策略,即步骤S132,包括:

[0183] S1323:获取学习用无人机11在第K-1个时隙感知的第三受干扰状态信息;

[0184] S1324:根据第K-1个时隙对应的Q值表,确定与第三受干扰状态信息对应的抗干扰策略子集;

[0185] S1325:根据抗干扰策略子集中的每个抗干扰策略对应的Q值,通过玻尔兹曼随机策略确定第一抗干扰策略。

[0186] 通过玻尔兹曼随机策略确定第一抗干扰策略可表征为:

[0187] $W(k) = [w_1(k), w_2(k), \dots, w_y(k), \dots, w_Y(k)]$;

$$w_y(k) = \frac{\exp(\zeta Q(s_k, g_y))}{\sum_{y \in B \times M} \exp(\zeta Q(s_k, g_y))};$$

[0189] 其中,

[0190] ζ 为玻尔兹曼更新系数, $w_y(k)$ 为第K时隙无人机选择抗干扰策略 g_y 的概率。

[0191] 其他实施方式中,确定学习用无人机11在第K个时隙的第一抗干扰策略还可以通过例如 ϵ -greedy策略或者高斯策略确定。

[0192] 图6是本发明一实施例中无人机通信Q值表确定方法的流程图五。

[0193] 请参考图6,一种实施方式中,根据奖励函数,确定学习用无人机11在第一受干扰状态信息下执行第一抗干扰策略获得的第一奖励值之前,即步骤S133之前,还包括:

[0194] S136:确定学习用无人机11在第一受干扰状态信息下执行第一抗干扰策略之后,学习用无人机11接收通信信号的信噪比;

[0195] 若信噪比小于或等于预设的学习用无人机11的信号解调门限,则:确定奖励函数的函数值为一预设奖励值;

[0196] 若信噪比大于预设的信号解调门限,则:根据信号吞吐量与信道切换开销确定奖励函数;信号吞吐量表征学习用无人机11接收和/或传输通信信号的能力,信道切换开销表征学习用无人机11切换通信信道产生的损失。

[0197] 以上方案中,信号解调门限可以理解为,学习用无人机11能够解调收到的通信信号的的门限值,具体可以是表征学习用无人机11接收的信号的信噪比的数值。当学习用无人机11接收信号的信噪比过低,即信噪比小于或等于设定的信号解调门限,则表明此时学习用无人机11无法使用接收到的通信信号,进而可以认为此时无人机无法工作,奖励函数的函数值设定为一预设奖励值,该预设奖励值表示对学习用无人机11在执行第一抗干扰策略之后的反馈,由于执行第一抗干扰策略之后学习用无人机11并不能解调通信信号,预设奖励值可以是表征负反馈(消极反馈)的数值,例如可以是负数值,也可以是零值。

[0198] 以上方案中,信道切换开销表征学习用无人机11切换通信信道产生的损失(也可以理解为奖励函数的修正项),切换信道产生的损失例如可以包括时间开销、执行切换动作所述消耗的能量,这些统一用切换开销表征,实际应用中,具体的切换开销可根据经验确

定,也可根据Q学习的结果对切换开销进行多次调整。

[0199] 在一种实施方式中,奖励函数可以表征为:

$$[0200] \quad \mu_s(a_t, a_{t-1}, c_t, B_t^1, B_t^2) = \begin{cases} B_t^1 \log(1 + \eta(a_t, c_t, B_t^1, B_t^2)) - \kappa_c(1 - \delta(a_t, a_{t-1})), & \eta(a_t, c_t, B_t^1, B_t^2) > \lambda \\ 0, & \eta(a_t, c_t, B_t^1, B_t^2) \leq \lambda \end{cases};$$

[0201] 其中,

[0202] $\mu_s(a_t, a_{t-1}, c_t, B_t^1, B_t^2)$ 为 t 时隙的奖励函数(也可称为效用函数);

$B_t^1 \log(1 + \eta(a_t, c_t, B_t^1, B_t^2))$ 表示信号吞吐量;

[0203] $\kappa_c(1 - \delta(a_t, a_{t-1}))$ 表示信道切换开销;

[0204] λ 为信号解调门限;

[0205] a_{t-1} 表示在第 t-1 个时隙学习用无人机11的通信信道;

[0206] a_t 表示在第 t 个时隙学习用无人机11的通信信道;

[0207] c_t 表示在第 t 个时隙干扰信号信道;

[0208] B_t^1 表示在第 t 个时隙学习用无人机11的通信信道的带宽;

[0209] B_t^2 表示在第 t 个时隙干扰信号带宽;

[0210] $\eta(a_t, c_t, B_t^1, B_t^2)$ 表示在第 t 个时隙的信噪比;

[0211] κ_c 表示信道切换开销系数,具体为一常数,该常数可使用经验值,也可根据多次Q学习的结果确定;

$$[0212] \quad \delta(a_t, a_{t-1}) \text{ 为指示函数, } \delta(a_t, a_{t-1}) = \begin{cases} 1, & a_t = a_{t-1} \\ 0, & a_t \neq a_{t-1} \end{cases};$$

[0213] 信噪比可以表征为:

$$[0214] \quad \eta(a_t, c_t, B_t^1, B_t^2) = \frac{P_1(t)}{N_0 B_t^1 + P_2(t)};$$

[0215] 其中:

[0216] $P_1(t)$ 表示在第 t 个时隙学习用无人机11接收到的通信信号功率;

[0217] $P_2(t)$ 表示在第 t 个时隙学习用无人机11接收到的干扰信号的功率;

[0218] N_0 表示单位带宽的噪声功率。

[0219] 学习用无人机11在实际飞行中, $P_1(t)$ 、 $P_2(t)$ 均可通过其自身的传感器或其他功率检测设备直接获取。

[0220] $P_1(t)$ 具体可以表征为:

[0221] $P_1(t) = p \omega_1(t)$;

[0222] $\omega_1(t) = d_1^\alpha \varepsilon_1$;

[0223] 其中:

[0224] p 表示通信用无人机12的信号发射功率;

[0225] $\omega_1(t)$ 表示在第t个时隙学习用无人机11接收通信信号的信道增益；

[0226] d_1 表示在第t个时隙学习用无人机11与通信用无人机12之间的距离；

[0227] α 表示路径衰落因子；

[0228] ε_1 表示瞬时衰落系数,服从单位均值的指数分布；

[0229] $P_2(t)$ 具体可以表征为：

[0230] $P_2(t) = J \omega_2(t)$ ；

[0231] $\omega_2(t) = d_2^\beta \varepsilon_2 f(a_t, c_t, B_t^1, B_t^2)$ ；

[0232]
$$f(a_t, c_t, B_t^1, B_t^2) = \begin{cases} 1, & B_t^1 > B_t^2, \quad a_t = c_t \\ \frac{B_t^1}{B_t^2}, & B_t^1 < B_t^2, \quad a_t = c_t ; \\ 0, & a_t \neq c_t \end{cases}$$

[0233] 其中：

[0234] J表示干扰源发射的干扰信号的功率；

[0235] $\omega_2(t)$ 表示在第t个时隙学习用无人机11接收干扰信号的信道增益；

[0236] d_2 表示在第t个时隙学习用无人机11与干扰源(干扰无人机13)之间的距离；

[0237] β 表示路径衰落因子；

[0238] ε_2 表示瞬时衰落系数,服从单位均值的指数分布；

[0239] $f(a_t, c_t, B_t^1, B_t^2)$ 在第t个时隙干扰的增益系数。

[0240] 图7a是干扰的增益系数为0时的示意图。

[0241] 图7b是干扰的增益系数为1时的示意图。

[0242] 图7c是干扰的增益系数为 $\frac{B_t^1}{B_t^2}$ 时的示意图。

[0243] 图7a、图7b、图7c的纵坐标表示信号功率谱,横坐标表示频率。

[0244] 请参考图7a,当 $a_t \neq c_t$ 时,干扰信号与通信信号信道不重合,干扰信号不会干扰通信信号,此时干扰信号的增益系数为0。

[0245] 请参考图7b,当 $a_t = c_t$,并且 $B_t^1 > B_t^2$ 时,学习用无人机11可以接收全部的干扰信号,此时干扰信号的增益系数为1。

[0246] 请参考图7c,当 $a_t = c_t$,并且 $B_t^1 < B_t^2$ 时,学习用无人机11可以接收全部的干扰信

号,接收的干扰信号的比例(可以理解为此时的干扰信号的增益系数)为 $\frac{B_t^1}{B_t^2}$ 。

[0247] 图8是本发明一实施例中无人机飞行的场景示意图。

[0248] 请参考图8,该场景所关联的设备可包括:无人机21、干扰源22、通信站23。

[0249] 其中,无人机21与通信站23进行通信,通信站23可以例如是地面固定的通信站(例如远程控制中心),也可以例如是地面移动的通信站(例如车载通信站),也可以例如是另一

无人机。

[0250] 干扰源22连续在飞行区域发射干扰信号,干扰信号能够干扰无人机21与通信站23的通信。干扰源22可以例如是地面固定的干扰源,也可以例如是地面移动的干扰源,也可以例如是另一发射干扰信号的无人机。干扰源22可以进一步理解为正常工作的可发射无线信号的电子设备,该电子设备发射的无线信号能够对无人机21的通信造成干扰。

[0251] 图9是本发明一实施例中无人机通信抗干扰方法的流程图。

[0252] 请参考图9,一种无人机通信抗干扰方法,包括:

[0253] S21:确定无人机21在当前时刻的受干扰状态信息,不同的受干扰状态信息表征了对应无人机21接收到的干扰信号的功率信息所处的不同功率区间和/或不同的信道;

[0254] S22:根据当前时刻的受干扰状态信息以及上述方案涉及的方法确定的Q值表,确定无人机在当前时刻所需执行的抗干扰策略;

[0255] S23:执行当前时刻所需执行的抗干扰策略,以调整无人机21的通信信道和/或通信带宽。

[0256] 本发明实施例中,确定无人机21在当前时刻的受干扰状态信息可以进一步理解为确定当前时刻的无人机21接收或感知到的干扰信号的功率以及干扰信号的信道,根据感知到的干扰信号的功率确定该功率对应的功率区间。一种实施方式中,干扰信号的功率及信道可通过频谱检测技术或宽带频谱感知技术确定。

[0257] 以上方案中,通过结合无人机21当前时刻飞行环境的干扰信号的功率以及干扰信号的信道以及已确定的Q值表确定在当前时刻所需执行的抗干扰策略,由于无人机21的所需执行的抗干扰策略同时考虑了当前时刻的干扰信号的功率以及信道,相对现有技术,能够更加全面反映外部环境对无人机21通信的影响,所确定的抗干扰策略更加符合实际飞行情况;由于考虑了当前时刻的干扰信号的功率以及信道,相应的抗干扰策略被执行时选择调整无人机21的通信信道和/或通信带宽,使得抗干扰策略的选择更加灵活。

[0258] 图10是本发明一实施例中无人机通信抗干扰装置的模块示意图。

[0259] 请参考图10,一种无人机通信抗干扰装置,包括:

[0260] 状态获取模块31,用于确定无人机在当前时刻的受干扰状态信息,不同的受干扰状态信息表征了对应无人机接收到的干扰信号的功率信息所处的不同功率区间和/或不同的信道;

[0261] 策略确定模块32,用于根据当前时刻的受干扰状态信息以及已确定的Q值表,确定无人机在当前时刻所需执行的抗干扰策略;Q值表记载了每一种受干扰状态信息下执行不同抗干扰策略对应的Q值;

[0262] 策略执行模块33,用于执行当前时刻所需执行的抗干扰策略,以调整所述无人机的通信信道和/或通信带宽。

[0263] 图11是本发明一实施例中无人机通信Q值表确定装置的模块示意图一。

[0264] 请参考图11,一种无人机通信Q值表确定装置,包括:

[0265] 第一模块41,用于确定受干扰状态集、抗干扰策略集以及奖励函数;

[0266] 第二模块42,用于根据受干扰状态集、抗干扰策略集,确定初始的Q值表;

[0267] 第三模块43,用于在学习用无人机11受控执行各抗干扰策略时,通过Q学习的方式更新Q值表中的Q值;所更新的Q值与学习用无人机11所采用的抗干扰策略、所接收到的干扰

信号的功率信息与信道信息,以及奖励函数相关联;

[0268] 其中,Q值表记载了每一种受干扰状态信息下执行不同抗干扰策略对应的Q值;受干扰状态集表征了预先定义的受干扰状态信息的集合,抗干扰策略集为学习用无人机11可执行的抗干扰策略的集合,奖励函数用于表征学习用无人机11所执行的不同抗干扰策略、执行对应抗干扰策略之前的受干扰状态信息,以及奖励值之间的函数关系;不同的受干扰状态信息表征了学习用无人机11接收到的干扰信号的功率所处的不同功率区间和/或不同的信道;抗干扰策略被执行时,以调整学习用无人机11的通信信道和/或通信带宽。

[0269] 本发明一实施例中的无人机通信Q值表确定装置,其中确定的Q值表包括每一种受干扰状态信息下执行不同抗干扰策略对应的Q值,由于不同的受干扰状态信息表征了学习用无人机11接收到的干扰信号的功率信息所处的不同功率区间和/或不同的信道,相对现有技术,能够更加全面反映外部环境对无人机通信的影响,进而所确定的抗干扰策略更加符合实际飞行情况;由于抗干扰策略被执行时选择调整无人机的通信信道和/或通信带宽,使得抗干扰策略的选择更加灵活。

[0270] 图12是本发明一实施例中无人机通信Q值表确定装置的模块示意图二。

[0271] 请参考图12,一种实施例中,第三模块43,包括:

[0272] 第一信息获取单元431,用于获取学习用无人机11在任意第K个时隙感知的第一受干扰状态信息;

[0273] 第一策略确定单元432,用于确定学习用无人机11在第K个时隙的第一抗干扰策略;

[0274] 第一计算单元433,用于根据奖励函数,确定学习用无人机11在第一受干扰状态信息下执行第一抗干扰策略获得的第一奖励值;

[0275] 第二信息获取单元434,用于获取学习用无人机11在第K+1个时隙感知的第二受干扰状态信息;

[0276] 第二计算单元435,用于根据第K个时隙对应的Q值表、第一受干扰状态信息、第一抗干扰策略、第二受干扰状态信息以及第一奖励值,更新Q值表,得到第K+1个时隙对应的Q值表。

[0277] 图13是本发明一实施例中无人机通信Q值表确定装置的模块示意图三。

[0278] 请参考图13,一种实施例中,第一策略确定单元432,包括:

[0279] 第一策略子集确定子单元4321,用于根据第K个时隙对应的Q值表,确定与第一受干扰状态信息对应的抗干扰策略子集;

[0280] 第一策略选择子单元4322,用于确定抗干扰策略子集中Q值最大的抗干扰策略作为第一抗干扰策略。

[0281] 图14是本发明一实施例中无人机通信Q值表确定装置的模块示意图四。

[0282] 请参考图14,一种实施例中,第一策略确定单元432,包括:

[0283] 信息获取子单元4323,用于获取学习用无人机11在第K-1个时隙感知的第三受干扰状态信息;

[0284] 第二策略子集确定子单元4324,用于根据第K-1个时隙对应的Q值表,确定与第三受干扰状态信息对应的抗干扰策略子集;

[0285] 第二策略选择子单元4325,用于根据抗干扰策略子集中的每个抗干扰策略对应的

Q值,通过玻尔兹曼随机策略确定第一抗干扰策略。

[0286] 图15是本发明一实施例中无人机通信Q值表确定装置的模块示意图五。

[0287] 请参考图15,一种实施例中,该装置还包括:

[0288] 第三计算单元436,用于确定学习用无人机11在第一受干扰状态信息下执行第一抗干扰策略之后,学习用无人机11接收通信信号的信噪比;

[0289] 若信噪比小于或等于预设的学习用无人机11的信号解调门限,则:确定奖励函数的函数值为一预设奖励值;

[0290] 若信噪比大于预设的信号解调门限,则:根据信号吞吐量与信道切换开销确定奖励函数;信号吞吐量表征学习用无人机11接收通信信号的能力,信道切换开销表征学习用无人机11切换通信信道产生的损失。

[0291] 可选的,信噪比根据以下公式确定:

$$[0292] \quad \eta(a_t, c_t, B_t^1, B_t^2) = \frac{P_1(t)}{N_0 B_t^1 + P_2(t)};$$

[0293] 其中:

[0294] t表示时隙;

[0295] $\eta(a_t, c_t, B_t^1, B_t^2)$ 表示在第t个时隙的信噪比;

[0296] a_t 表示在第t个时隙学习用无人机11的通信信道;

[0297] c_t 表示在第t个时隙干扰信号信道;

[0298] B_t^1 表示在第t个时隙学习用无人机11的通信信道的带宽;

[0299] B_t^2 表示在第t个时隙干扰信号带宽;

[0300] $P_1(t)$ 表示在第t个时隙学习用无人机11接收到的通信信号功率;

[0301] $P_2(t)$ 表示在第t个时隙学习用无人机11接收到的干扰信号的功率。

[0302] 综上,本发明一实施例提供的无人机通信Q值表确定装置,策略确定模块32通过结合无人机当前时刻飞行环境的干扰信号的功率以及干扰信号的信道以及已确定的Q值表确定在当前时刻所需执行的抗干扰策略,由于无人机的所需执行的抗干扰策略同时考虑了当前时刻的干扰信号的功率以及信道,相对现有技术,能够更加全面反映外部环境对无人机通信的影响,所确定的抗干扰策略更加符合实际飞行情况;由于考虑了当前时刻的干扰信号的功率以及信道,相应的抗干扰策略被执行时选择调整无人机的通信信道和/或通信带宽,使得抗干扰策略的选择更加灵活。

[0303] 图16是本发明一实施例中电子设备的构造示意图。

[0304] 请参考图16,一种电子设备,包括处理器51与存储器53,

[0305] 存储器53,用于存储代码和相关数据;

[0306] 处理器51,用于执行存储器53中的代码用以实现本发明第一方面及其可选方案涉及的方法。

[0307] 处理器51能够通过总线52与存储器53通讯。

[0308] 本发明一实施例还提供了一种存储介质,其上存储有计算机程序,该程序被处理器执行时实现本发明第一方面及其可选方案涉及的方法。

[0309] 最后应说明的是:以上各实施例仅用以说明本发明的技术方案,而非对其限制;尽管参照前述各实施例对本发明进行了详细的说明,本领域的普通技术人员应当理解:其依然可以对前述各实施例所记载的技术方案进行修改,或者对其中部分或者全部技术特征进行等同替换;而这些修改或者替换,并不使相应技术方案的本质脱离本发明各实施例技术方案的范围。

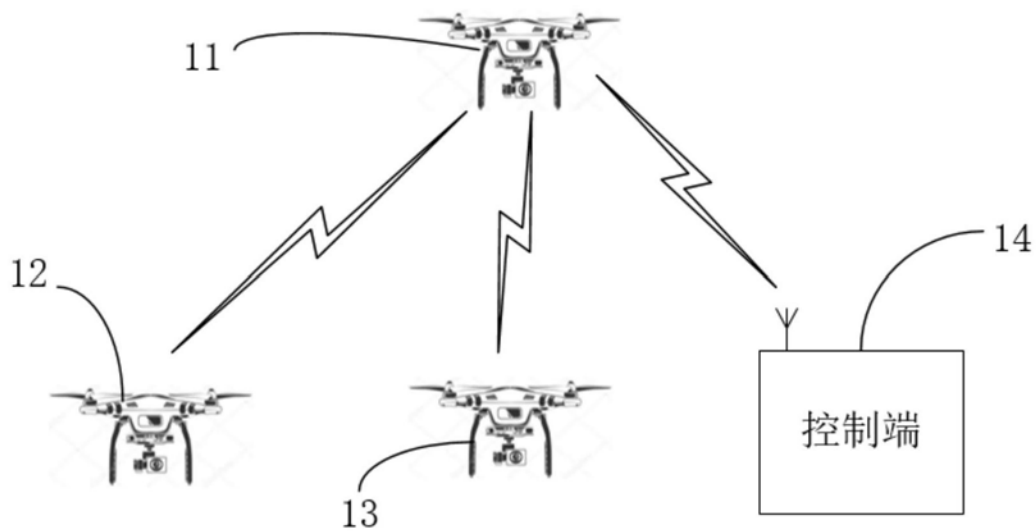


图1

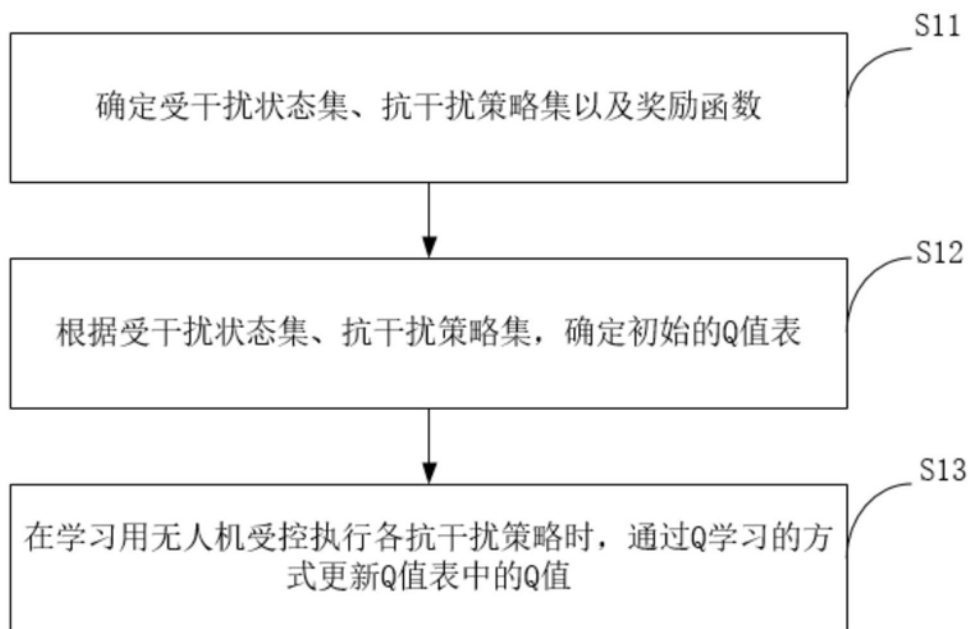


图2

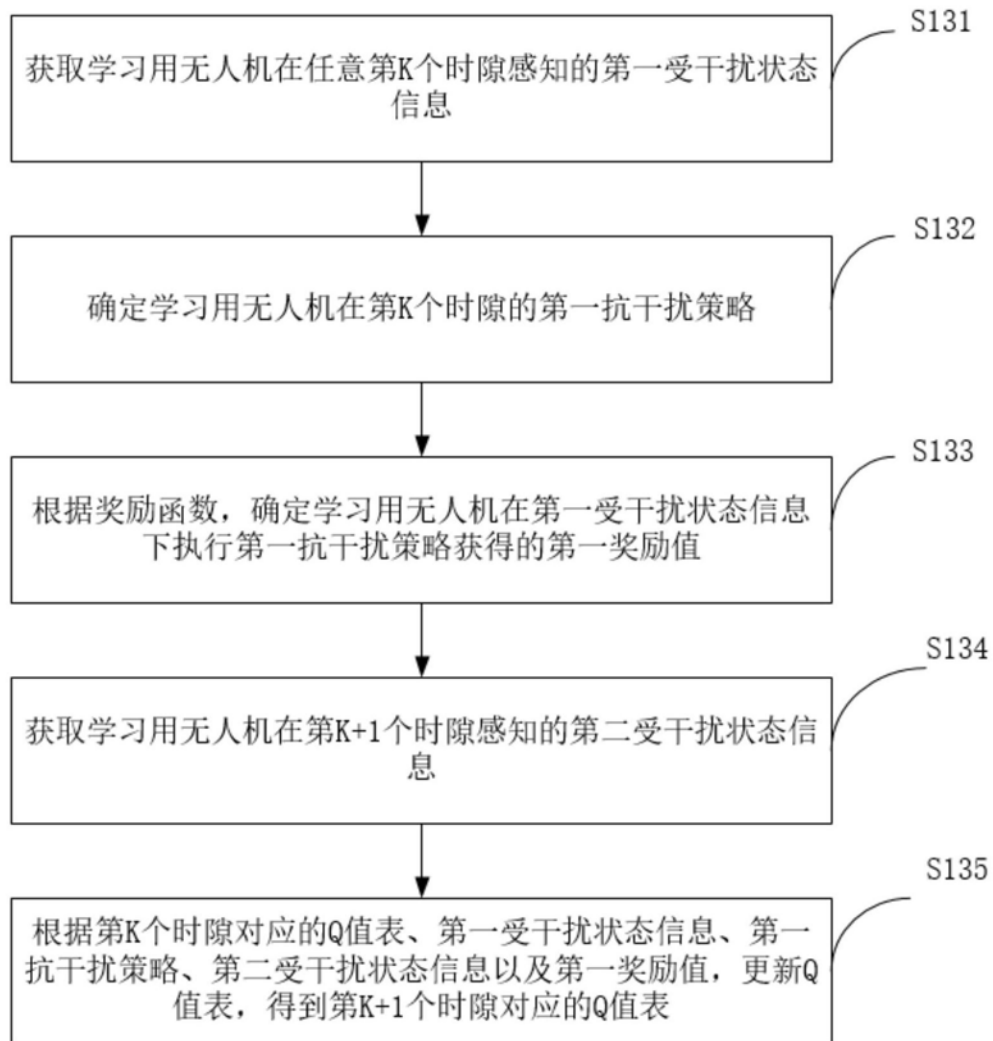


图3

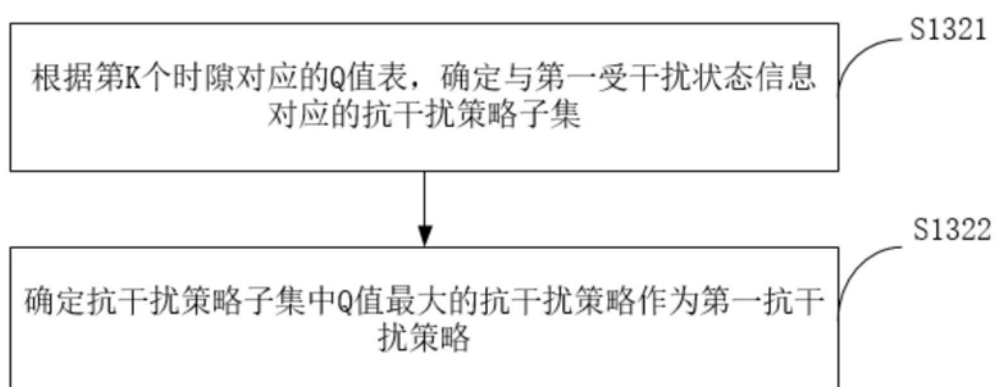


图4

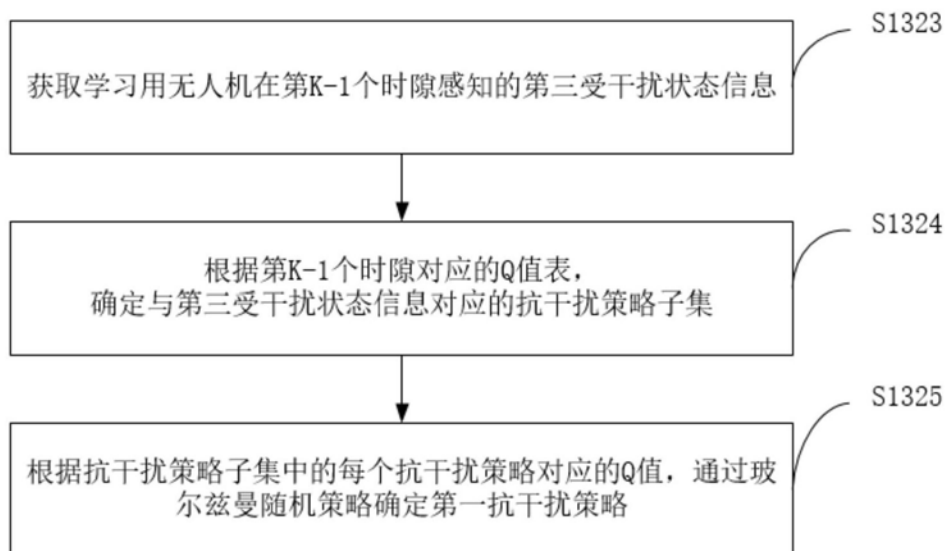


图5

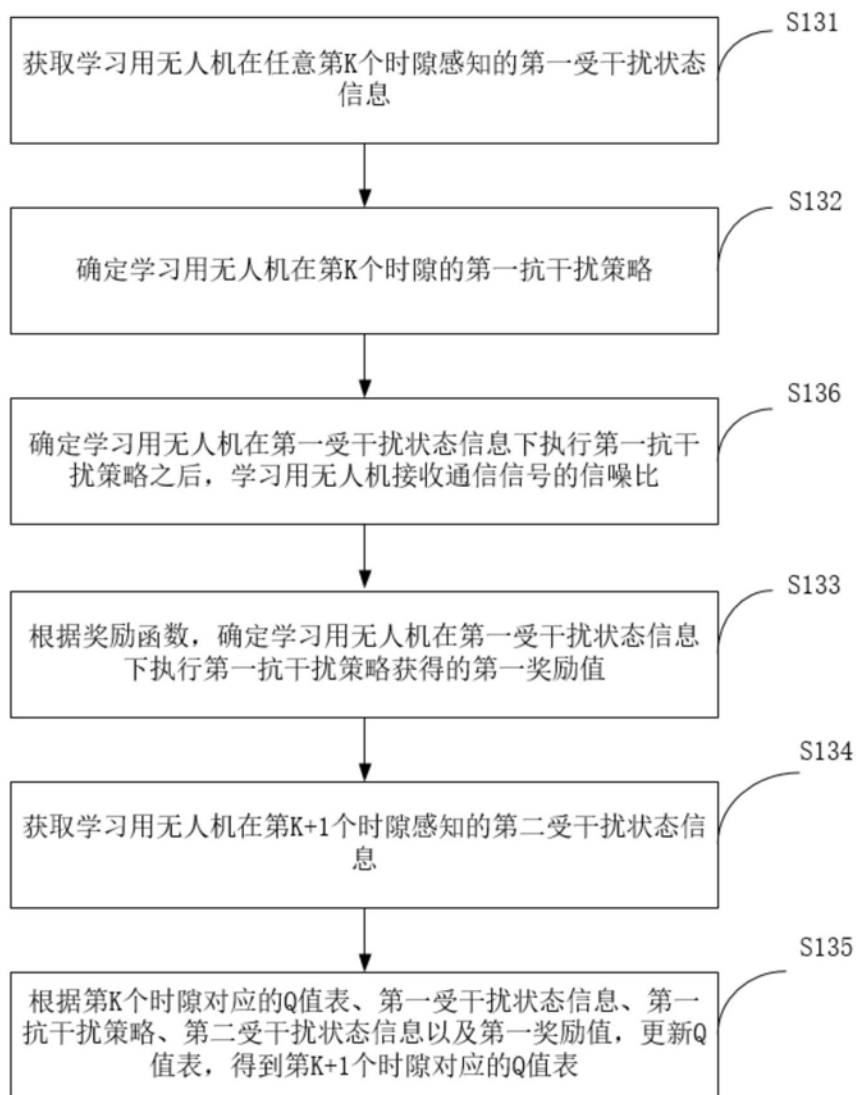


图6

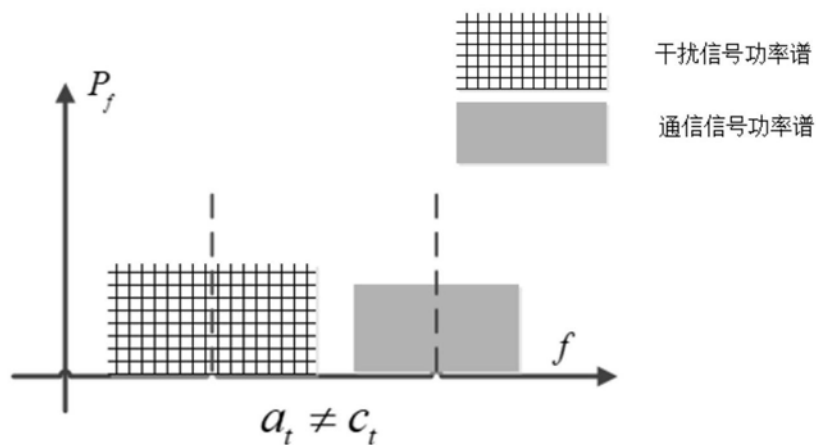


图7a

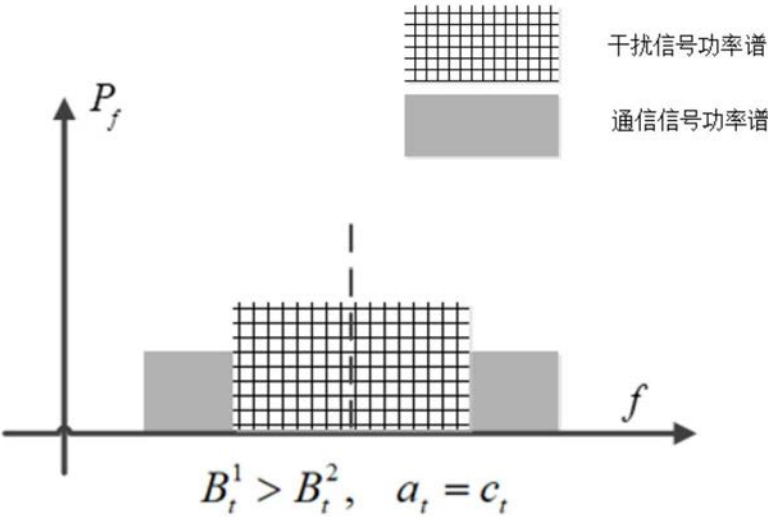


图7b

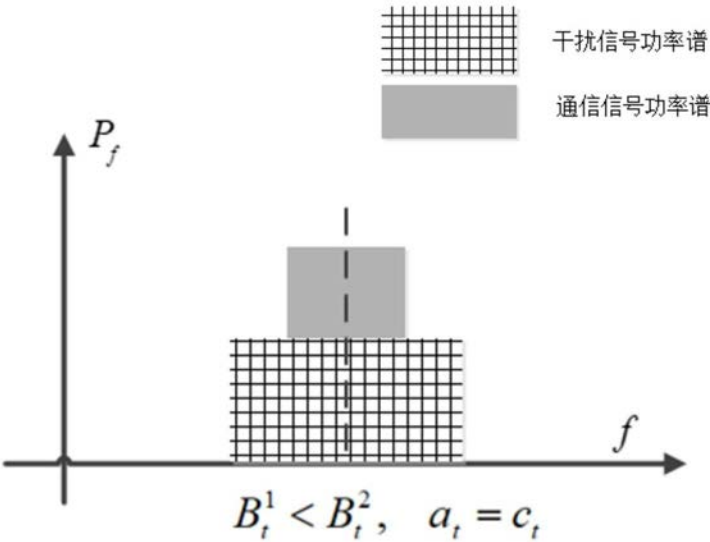


图7c

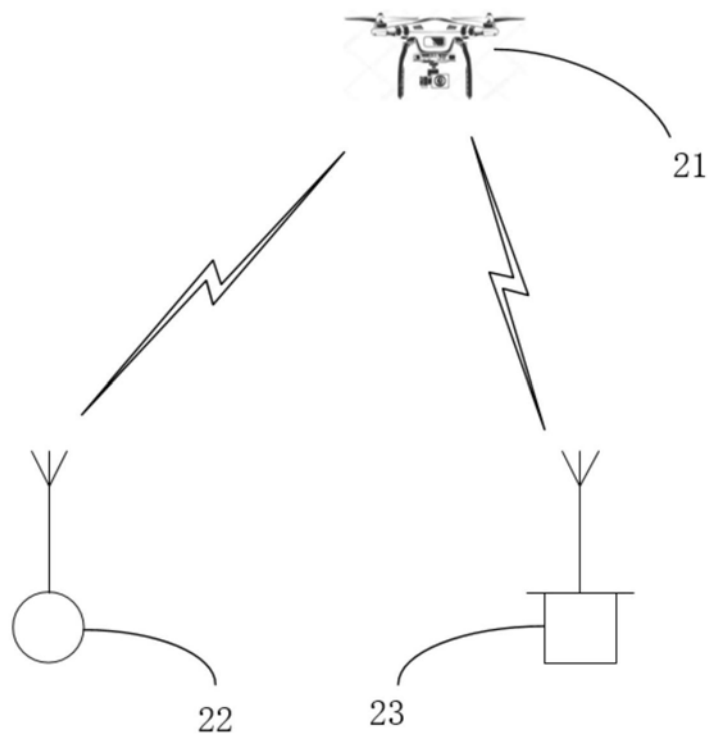


图8

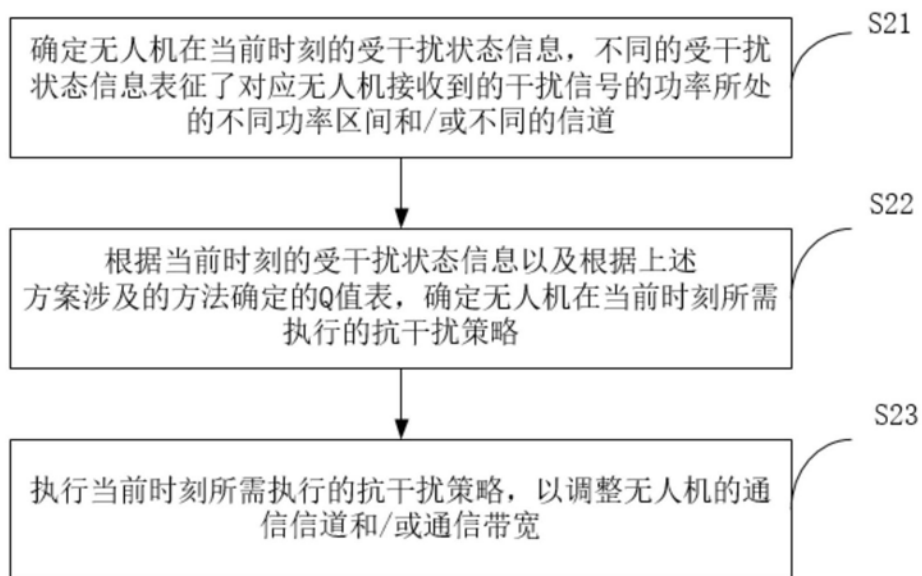


图9

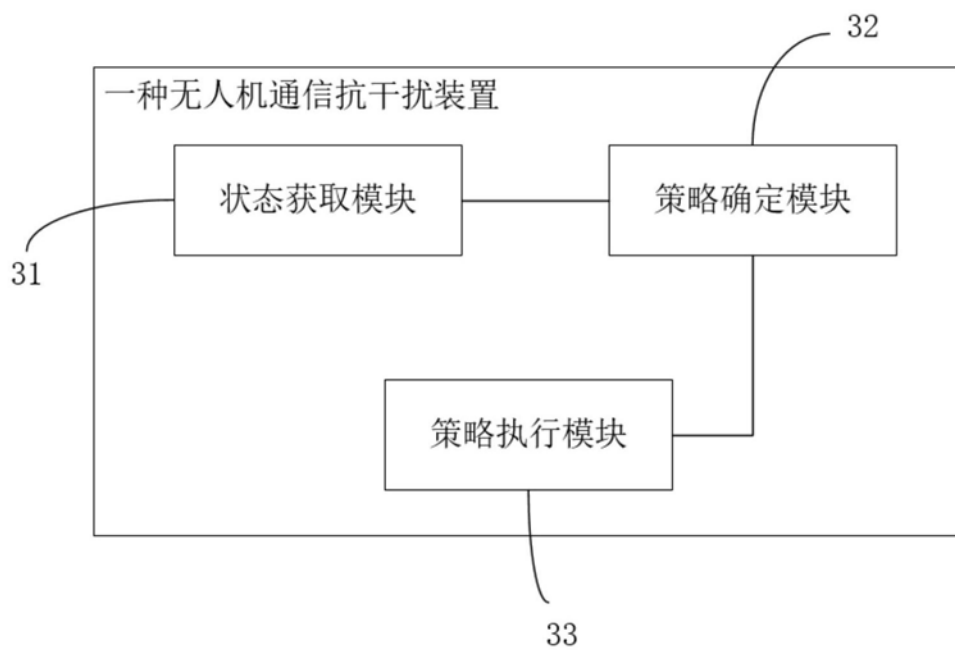


图10

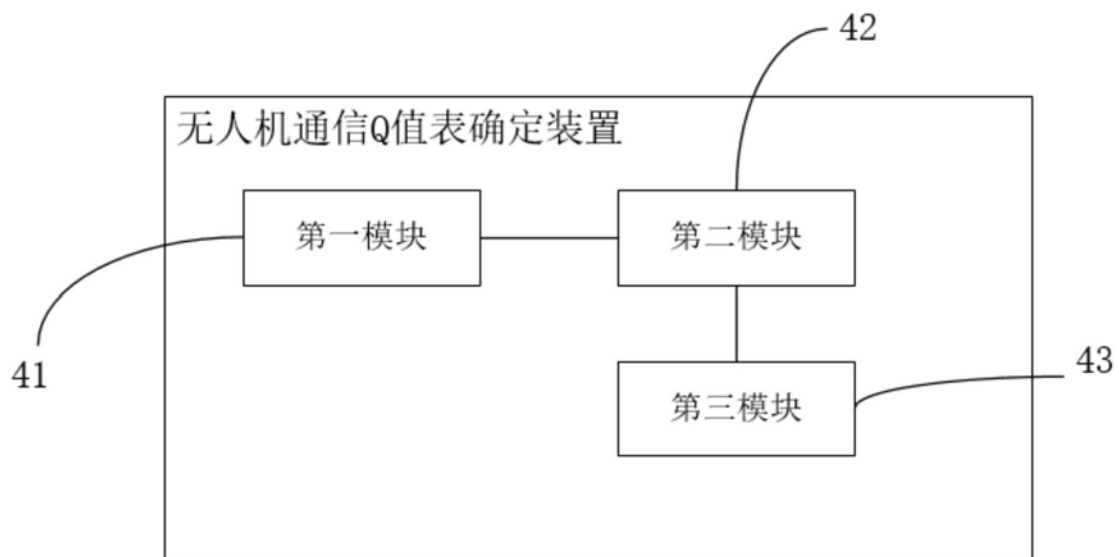


图11

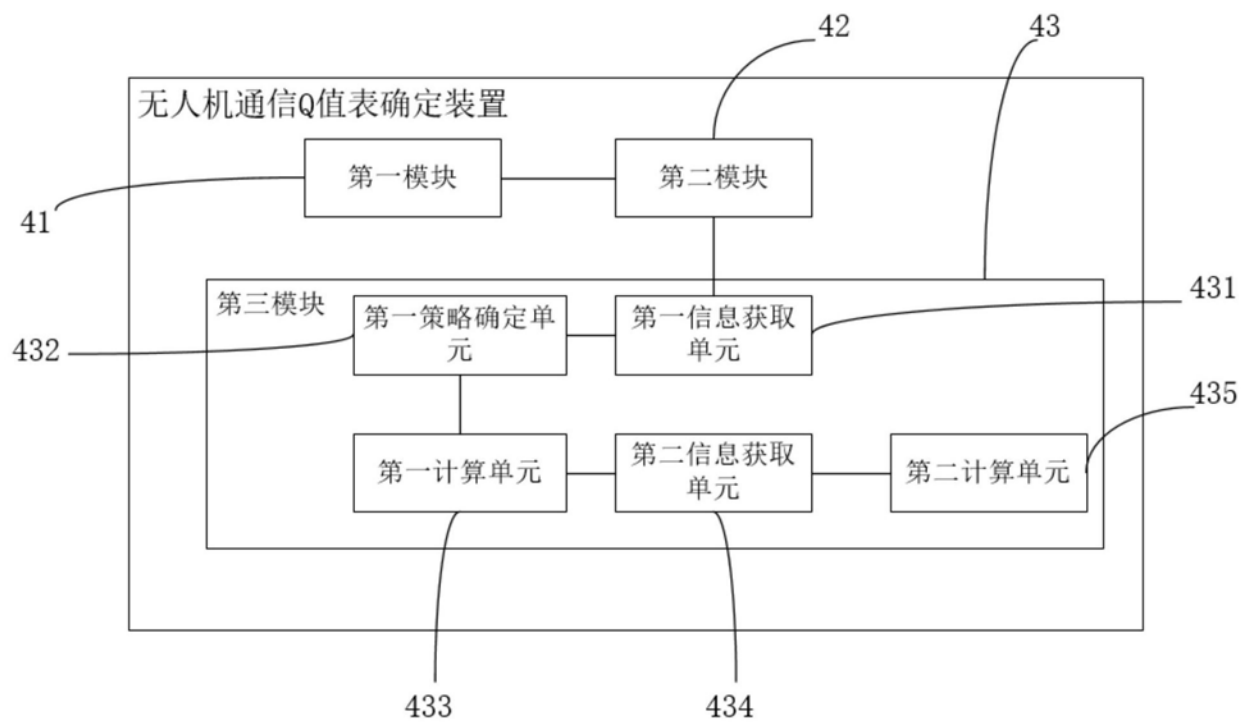


图12

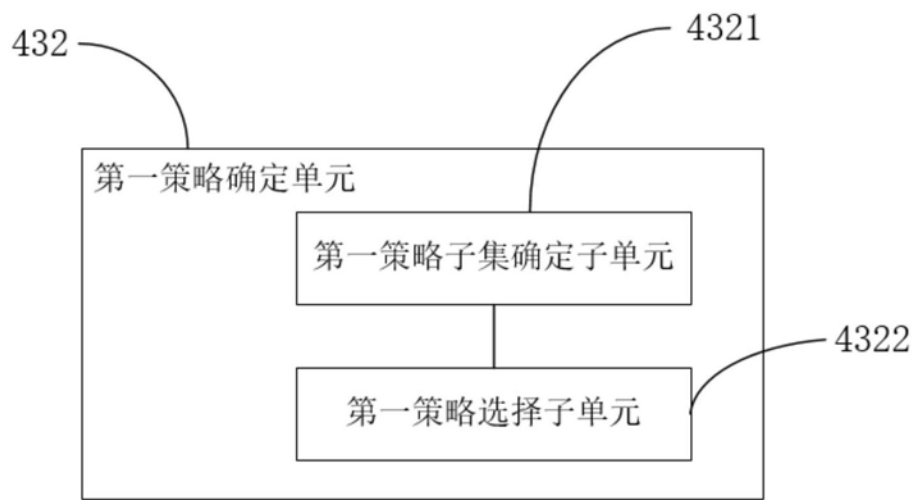


图13

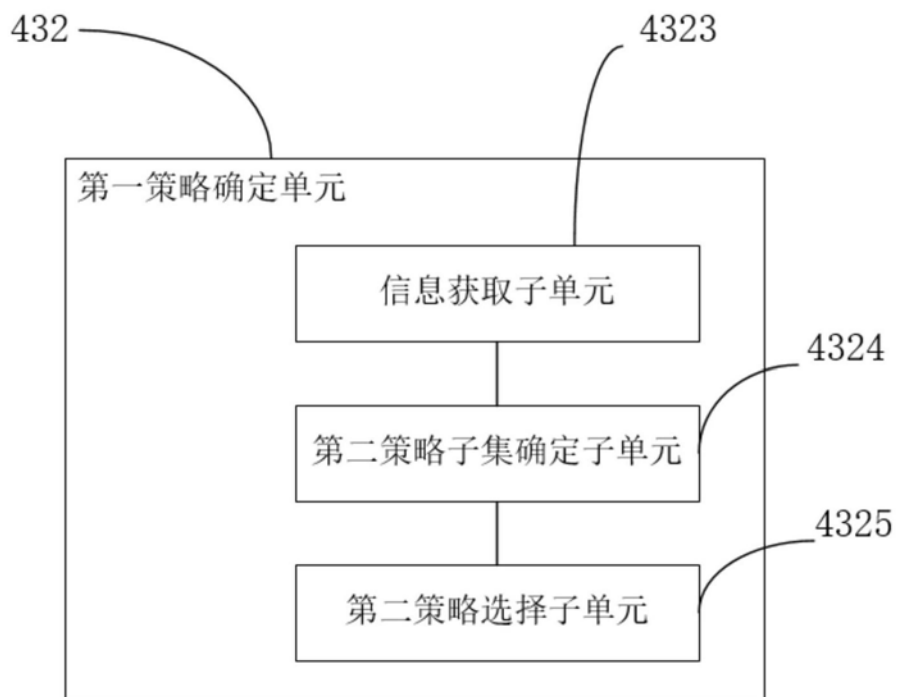


图14

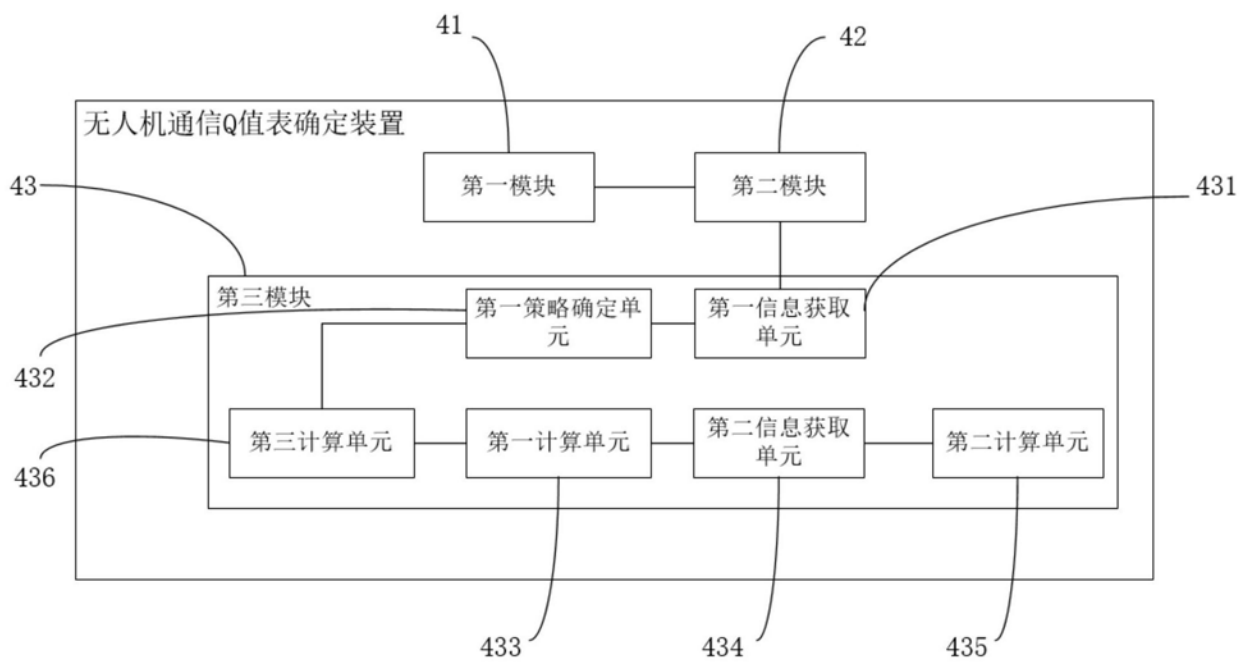


图15

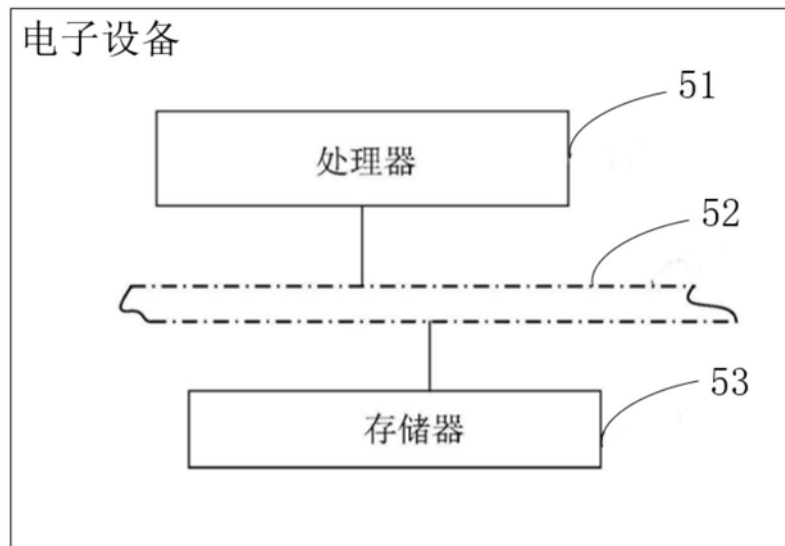


图16