



(12) 发明专利

(10) 授权公告号 CN 116543838 B

(45) 授权公告日 2023.09.05

(21) 申请号 202310814363.8

G16B 50/00 (2019.01)

(22) 申请日 2023.07.05

G06N 3/0464 (2023.01)

G06N 3/08 (2023.01)

(65) 同一申请的已公布的文献号

申请公布号 CN 116543838 A

(56) 对比文件

(43) 申请公布日 2023.08.04

CN 110246544 A, 2019.09.17

CN 115620854 A, 2023.01.17

(73) 专利权人 苏州凌点生物技术有限公司

CN 112419295 A, 2021.02.26

地址 215000 江苏省苏州市自由贸易试验

CN 1052142 A, 1991.06.12

区苏州片区苏州工业园区星湖街328

CN 111931754 A, 2020.11.13

号创意产业园10-401-001单元

CN 108596227 A, 2018.09.28

CN 113874494 A, 2021.12.31

(72) 发明人 张玲芳

审查员 武茹茹

(74) 专利代理机构 北京艾格律诗专利代理有限

公司 11924

专利代理师 谢毅

(51) Int. Cl.

G16B 30/10 (2019.01)

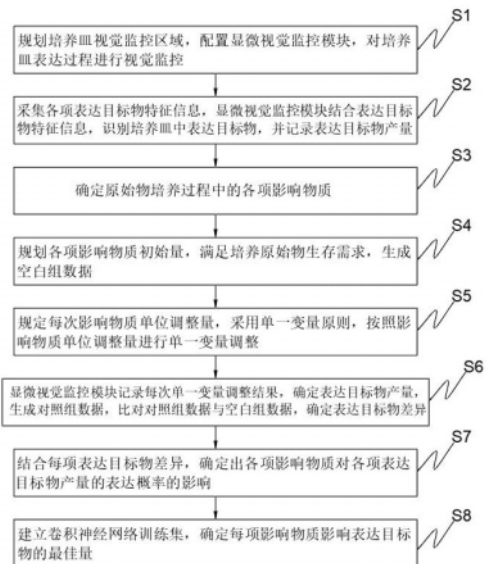
权利要求书2页 说明书6页 附图4页

(54) 发明名称

一种生物基因选择表达概率的数据分析方法

(57) 摘要

本发明涉及生物基因检测技术领域,具体地说,涉及一种生物基因选择表达概率的数据分析方法。其包括结合每项表达目标物差异,确定出各项影响物质对各项表达目标物产量的表达概率的影响;建立卷积神经网络训练集,确定每项影响物质影响表达目标物的最佳量。本发明确定表达目标物差异,建立卷积神经网络训练集,确定每项影响物质影响表达目标物的最佳量,在验证过程中通过各项影响物质浓度变化确定对表达目标物表达概率影响程度,为后期进行原始物培养提供参考依据,根据表达目标物产量峰值点以及对应的当前影响物质浓度,确定出各项影响物质浓度对表达目标物产量的影响区间,比对各组对照组数据,确定表达目标物产量最佳时对应的各项影响物质浓度。



1. 一种生物基因选择表达概率的数据分析方法,其特征在于,包括如下步骤:

S1、规划培养皿视觉监控区域,配置显微视觉监控模块,对培养皿表达过程进行视觉监控;

S2、采集各项表达目标物特征信息,显微视觉监控模块结合表达目标物特征信息,识别培养皿中表达目标物,并记录表达目标物产量;

S3、确定原始物培养过程中的各项影响物质;

S4、规划各项影响物质初始量,满足培养原始物生存需求,生成空白组数据;

S5、规定每次影响物质单位调整量,采用单一变量原则,按照影响物质单位调整量进行单一变量调整;

S6、显微视觉监控模块记录每次单一变量调整结果,确定表达目标物产量,生成对照组数据,比对对照组数据与空白组数据,确定表达目标物差异;

S7、结合每项表达目标物差异,确定出各项影响物质对各项表达目标物产量的表达概率的影响;

S8、建立卷积神经网络训练集,确定每项影响物质影响表达目标物的最佳量。

2. 根据权利要求1所述的生物基因选择表达概率的数据分析方法,其特征在于:所述S1中规划培养皿视觉监控区域的方法包括如下步骤:

S1.1、确定表达目标物集中位置,标记为集中监控区域;

S1.2、根据显微视觉监控模块反馈画面清晰度,调整监控点高度。

3. 根据权利要求1所述的生物基因选择表达概率的数据分析方法,其特征在于:所述S2中采集各项表达目标物特征信息的方法包括如下步骤:

S2.1、建立各项表达产物特征信息库;

S2.2、根据表达产物特征信息库确定每项表达产物对应特征,通过特征进行表达产物区分。

4. 根据权利要求1所述的生物基因选择表达概率的数据分析方法,其特征在于:所述S3中影响物质包括温度因子、PH值因子、存在性因子、营养液浓度、水溶液浓度以及参考溶液浓度。

5. 根据权利要求1所述的生物基因选择表达概率的数据分析方法,其特征在于:所述S4中规划各项影响物质初始量的方法包括如下步骤:

S4.1、确定原始物初始培养量;

S4.2、采集原始物存在本体内所必须的各项影响物质含量,生成原始物影响物质含量数据库;

S4.3、结合原始物初始培养量,规划出培养皿中各项影响物质初始量。

6. 根据权利要求1所述的生物基因选择表达概率的数据分析方法,其特征在于:所述S5中单一变量原则包括如下步骤:

S5.1、确定需要验证的影响物质,标记为变量影响物质;

S5.2、确定变量影响物质调整前其余各项影响物质量,标记为恒量影响物质;

S5.3、规定表达目标物变化量正常范围,按照表达目标物变化量正常范围规划变量影响物质单位调整量。

7. 根据权利要求6所述的生物基因选择表达概率的数据分析方法,其特征在于:所述

S5.3中规划变量影响物质单位调整量的方法采用平均分算法,其算法公式包括如下步骤:

步骤一、确定表达目标物变化量正常范围,标记为 $[0, N]$;

步骤二、确定变量影响物质初始量状态下初始表达目标物量 N_{IQ} ,并确定叠加单位调整量 M 后形成的叠加变量影响物质质量状态下的叠加表达目标物量 N_{SQ} ,计算叠加表达目标物量 N_{SQ} 与初始表达目标物量 N_{IQ} 的差值,标记为 N_1 ;

步骤三、当 $N_1 \in [0, N]$,输出单位调整量为 M ;

当 $N_1 \notin [0, N]$,输出单位调整量为 $M/2$,确定变量影响物质初始量叠加单位调整量 $M/2$ 产生的中值表达目标物量 N_{MD} ;

步骤四、重复步骤三,直至中值表达目标物量 $N_{MD} \in [0, N]$,输出对应的叠加单位调整量。

8. 根据权利要求7所述的生物基因选择表达概率的数据分析方法,其特征在于:所述S7中确定出各项影响物质对各项表达目标物产量的表达概率的影响的方法包括如下步骤:

第一步、通过所述步骤二确定各项影响物质初始量 M_0 与其对应单位调整量 M 的比值率 $M_{ratio} = \frac{M}{M_0}$;

第二步、确定各项影响物质叠加位调整量 M 后产生的叠加表达目标物量 N_{SQ} ,计算出各项影响物质影响值 $N_{effect} = N_{SQ} \times M_{ratio} = N_{SQ} \times \frac{M}{M_0}$;

第三步、比对各项影响物质影响值大小,对各项影响物质影响表达目标物产生进行排序。

9. 根据权利要求8所述的生物基因选择表达概率的数据分析方法,其特征在于:所述S8中卷积神经网络训练集包括输入层、卷积层以及输出层。

一种生物基因选择表达概率的数据分析方法

技术领域

[0001] 本发明涉及生物基因检测技术领域,具体地说,涉及一种生物基因选择表达概率的数据分析方法。

背景技术

[0002] 生物基因选择表达是生物体在个体发育的不同时期、不同部位,通过基因水平、转录水平等的调控,表达基因组中不同的部分,其结果是完成细胞分化和个体发育,在细胞分化中,基因在特定的时间和空间条件下有选择表达的现象,其结果是形成了形态结构和生理功能不同的细胞。

[0003] 细胞培养皿是一种支持细胞生长及保持细胞稳定性的特殊溶液,是化学、生物和药理科学研究中的重要体系,也是细胞活性和新药开发的重要基础,细胞培养皿的主要成分有水、营养成分和存在性因子。

[0004] 在进行生物基因选择表达概率的数据预测过程中,现有的验证方式大多数通过人工进行实验操作,在细胞培养皿加入维持细胞存活的影响物质,通过改变影响物质的浓度确定不同影响物质对生物基因选择表达概率的影响,而细胞培养皿所添加的影响物质种类较多,且不同细胞所需的影响物质不同,通过人工进行逐步分析实验,任务量过于繁重,效率较低。

[0005] 为了应对上述问题,现亟需一种生物基因选择表达概率的数据分析方法。

发明内容

[0006] 本发明的目的在于提供一种生物基因选择表达概率的数据分析方法,以解决上述背景技术中提出的问题。

[0007] 为实现上述目的,提供了一种生物基因选择表达概率的数据分析方法,包括如下步骤:

[0008] S1、规划培养皿视觉监控区域,配置显微视觉监控模块,对培养皿表达过程进行视觉监控;

[0009] S2、采集各项表达目标物特征信息,显微视觉监控模块结合表达目标物特征信息,识别培养皿中表达目标物,并记录表达目标物产量;

[0010] S3、确定原始物培养过程中的各项影响物质;

[0011] S4、规划各项影响物质初始量,满足培养原始物生存需求,生成空白组数据;

[0012] S5、规定每次影响物质单位调整量,采用单一变量原则,按照影响物质单位调整量进行单一变量调整;

[0013] S6、显微视觉监控模块记录每次单一变量调整结果,确定表达目标物产量,生成对照组数据,比对对照组数据与空白组数据,确定表达目标物差异;

[0014] S7、结合每项表达目标物差异,确定出各项影响物质对各项表达目标物产量的表达概率的影响;

- [0015] S8、建立卷积神经网络训练集,确定每项影响物质影响表达目标物的最佳量。
- [0016] 作为本技术方案的进一步改进,所述S1中规划培养皿视觉监控区域的方法包括如下步骤:
- [0017] S1.1、确定表达目标物集中位置,标记为集中监控区域;
- [0018] S1.2、根据显微视觉监控模块反馈画面清晰度,调整监控点高度。
- [0019] 作为本技术方案的进一步改进,所述S2中采集各项表达目标物特征信息的方法包括如下步骤:
- [0020] S2.1、建立各项表达产物特征信息库;
- [0021] S2.2、根据表达产物特征信息库确定每项表达产物对应特征,通过特征进行表达产物区分。
- [0022] 作为本技术方案的进一步改进,所述S3中影响物质包括温度因子、PH值因子、存在性因子、营养液浓度、水溶液浓度以及参考溶液浓度。
- [0023] 作为本技术方案的进一步改进,所述S4中规划各项影响物质初始量的方法包括如下步骤:
- [0024] S4.1、确定原始物初始培养量;
- [0025] S4.2、采集原始物存在本体所必须的各项影响物质含量,生成原始物影响物质含量数据库;
- [0026] S4.3、结合原始物初始培养量,规划出培养皿中各项影响物质初始量。
- [0027] 作为本技术方案的进一步改进,所述S5中单一变量原则包括如下步骤:
- [0028] S5.1、确定需要验证的影响物质,标记为变量影响物质;
- [0029] S5.2、确定变量影响物质调整前其余各项影响物质量,标记为恒量影响物质;
- [0030] S5.3、规定表达目标物变化量正常范围,按照表达目标物变化量正常范围规划变量影响物质单位调整量。
- [0031] 作为本技术方案的进一步改进,所述S5.3中规划变量影响物质单位调整量的方法采用平均分算法,其算法公式包括如下步骤:
- [0032] 步骤一、确定表达目标物变化量正常范围,标记为 $[0, N]$;
- [0033] 步骤二、确定变量影响物质初始量状态下初始表达目标物量 N_{IQ} ,并确定叠加单位调整量 M 后形成的叠加变量影响物质量状态下的叠加表达目标物量 N_{SQ} ,计算叠加表达目标物量 N_{SQ} 与初始表达目标物量 N_{IQ} 的差值,标记为 N_1 ;
- [0034] 步骤三、当 $N_1 \in [0, N]$,输出单位调整量为 M ;
- [0035] 当 $N_1 \notin [0, N]$,输出单位调整量为 $M/2$,确定变量影响物质初始量叠加单位调整量 $M/2$ 产生的中值表达目标物量 N_{MD} ;
- [0036] 步骤四、重复步骤三,直至中值表达目标物量 $N_{MD} \in [0, N]$,输出对应的叠加单位调整量。
- [0037] 作为本技术方案的进一步改进,所述S7中确定出各项影响物质对各项表达目标物产量的表达概率的影响的方法包括如下步骤:

[0038] 第一步、通过所述步骤二确定各项影响物质初始量 M_0 与其对应单位调整量 M 的比值率 $M_{ratio} = \frac{M}{M_0}$;

[0039] 第二步、确定各项影响物质叠加位调整量 M 后产生的叠加表达目标物量 N_{SQ} ,计算出各项影响物质影响值 $N_{effect} = N_{SQ} \times M_{ratio} = N_{SQ} \times \frac{M}{M_0}$;

[0040] 第三步、比对各项影响物质影响值大小,对各项影响物质影响表达目标物产生进行排序。

[0041] 作为本技术方案的进一步改进,所述S8中卷积神经网络训练集包括输入层、卷积层以及输出层。

[0042] 与现有技术相比,本发明的有益效果:

[0043] 该生物基因选择表达概率的数据分析方法中,确定表达目标物差异,建立卷积神经网络训练集,确定每项影响物质影响表达目标物的最佳量,在验证过程中通过各项影响物质浓度变化确定对表达目标物表达概率影响程度,为后期进行原始物培养提供参考依据,根据表达目标物产量峰值点以及对应的当前影响物质浓度,确定出各项影响物质浓度对表达目标物产量的影响区间,比对各项对照组数据,确定表达目标物产量最佳时对应的各项影响物质浓度。

附图说明

[0044] 图1为本发明的整体流程步骤图;

[0045] 图2为本发明的整体运行示意图;

[0046] 图3为本发明的规划培养皿视觉监控区域的方法步骤图;

[0047] 图4为本发明的采集各项表达目标物特征信息的方法步骤图;

[0048] 图5为本发明的规划各项影响物质初始量的方法步骤图;

[0049] 图6为本发明的单一变量原则步骤图。

具体实施方式

[0050] 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。

[0051] 请参阅图1-图6所示,提供了一种生物基因选择表达概率的数据分析方法,包括如下步骤:

[0052] S1、规划培养皿视觉监控区域,配置显微视觉监控模块,对培养皿表达过程进行视觉监控;

[0053] S2、采集各项表达目标物特征信息,显微视觉监控模块结合表达目标物特征信息,识别培养皿中表达目标物,并记录表达目标物产量;

[0054] S3、确定原始物培养过程中的各项影响物质;

[0055] S4、规划各项影响物质初始量,满足培养原始物生存需求,生成空白组数据;

[0056] S5、规定每次影响物质单位调整量,采用单一变量原则,按照影响物质单位调整量

进行单一变量调整；

[0057] S6、显微视觉监控模块记录每次单一变量调整结果，确定表达目标物产量，生成对照组数据，比对对照组数据与空白组数据，确定表达目标物差异；

[0058] S7、结合每项表达目标物差异，确定出各项影响物质对各项表达目标物产量的表达概率的影响；

[0059] S8、建立卷积神经网络训练集，确定每项影响物质影响表达目标物的最佳量。

[0060] 具体使用时，在进行生物基因选择表达概率验证过程中，首先规划培养皿视觉监控区域，配置显微视觉监控模块，对培养皿表达过程进行视觉监控，采集各项表达目标物特征信息，显微视觉监控模块结合表达目标物特征信息，识别培养皿中表达目标物，并记录表达目标物产量，这些均为表达目标物特征信息，通过各项特征信息定位表达目标物，确定原始物培养过程中的各项影响物质，规划各项影响物质初始量，满足培养原始物生存需求，生成空白组数据，规定每次影响物质单位调整量，采用单一变量原则，按照影响物质单位调整量进行单一变量调整，即每次调整一项影响物质，其余影响物质浓度不变，且每次调整单位调整量浓度的影响物质，显微视觉监控模块记录每次单一变量调整结果，确定表达目标物产量，生成对照组数据，比对对照组数据与空白组数据，确定表达目标物差异，随后结合每项表达目标物差异，确定出各项影响物质对各项表达目标物产量的表达概率的影响，建立影响物质递增减区间与对应的表达目标物产量比对库，确定表达目标物产量变化点（例如从递增区间变化为递减区间），重新规划单位调整量，确定出表达目标物产量峰值点以及对应的当前影响物质浓度，从而建立卷积神经网络训练集，确定每项影响物质影响表达目标物的最佳量。

[0061] 此外，S1中规划培养皿视觉监控区域的方法包括如下步骤：

[0062] S1.1、确定表达目标物集中位置，标记为集中监控区域；

[0063] S1.2、根据显微视觉监控模块反馈画面清晰度，调整监控点高度。

[0064] 在进行视觉监控区域规划过程中，由于不同的表达目标物在表达过程中，其表达规则与表达数量不同，其集中位置也会随之改变，此时通过确定各项表达目标物表达规则与表达数量，确定表达目标物集中位置，标记为集中监控区域，并将显微视觉监控模块配置在集中监控区域，同时根据显微视觉监控模块反馈画面清晰度，调整监控点高度，使得反馈的监控画面能够清晰的观察出各项表达目标物。

[0065] 进一步的，S2中采集各项表达目标物特征信息的方法包括如下步骤：

[0066] S2.1、建立各项表达产物特征信息库；

[0067] S2.2、根据表达产物特征信息库确定每项表达产物对应特征，通过特征进行表达产物区分。

[0068] 在进行表达目标物识别过程中，首先建立各项表达产物特征信息库，例如表达目标物在显微状态下显现的状态、颜色以及形状等，随后根据表达产物特征信息库确定每项表达产物对应特征，通过特征进行表达产物区分，用于后期进行表达产物量统计。

[0069] 再进一步的，S3中影响物质包括温度因子、PH值因子、存在性因子、营养液浓度、水溶液浓度以及参考溶液浓度，其中温度因子为培养皿培养原始物的温度；

[0070] PH值因子为培养皿培养原始物的PH值；

[0071] 存在性因子：用于维持原始物的健康，调节原始物细胞膜的通透性；

- [0072] 营养液浓度:包括氨基酸以及碳水化合物等营养物质,用于给予原始物生存所需能量;
- [0073] 水溶液浓度:为培养皿中添加的水溶液浓度,用于维持原始物活力;
- [0074] 参考溶液浓度:为不同原始物所需的特殊物质,例如抗生素。
- [0075] 具体的,S4中规划各项影响物质初始量的方法包括如下步骤:
- [0076] S4.1、确定原始物初始培养量;
- [0077] S4.2、采集原始物存在本体时内所必须的各项影响物质含量,生成原始物影响物质含量数据库;
- [0078] S4.3、结合原始物初始培养量,规划出培养皿中各项影响物质初始量。
- [0079] 在进行各项影响物质初始量规划过程中,由于需要验证的原始物脱离了其本体,需要在培养皿中添加原始物存在本体时所必须的各项影响物质,首先确定原始物初始培养量,随后采集原始物存在本体时内所必须的各项影响物质含量,生成原始物影响物质含量数据库,结合原始物初始培养量,规划出培养皿中各项影响物质初始量,作为维持该原始物生存的各项影响物质质量,形成空白组数据,以供后期进行对照实验,单一分析出各项影响物质质量对该原始物的表达目标物的影响。
- [0080] 此外,S5中单一变量原则包括如下步骤:
- [0081] S5.1、确定需要验证的影响物质,标记为变量影响物质;
- [0082] S5.2、确定变量影响物质调整前其余各项影响物质质量,标记为恒量影响物质;
- [0083] S5.3、规定表达目标物变化量正常范围,按照表达目标物变化量正常范围规划变量影响物质单位调整量。
- [0084] 在进行变量调整过程中,为了保证单一变量性,首先确定需要验证的影响物质,标记为变量影响物质,随后确定变量影响物质调整前其余各项影响物质质量,标记为恒量影响物质,恒量影响物质将作为无关变量,其各项恒量影响物质质量始终保持不变,规定表达目标物变化量正常范围,按照表达目标物变化量正常范围规划变量影响物质单位调整量,例如当变量影响物质单位调整量过大时,其表达目标物变化量急剧增加或者导致原始物停止表达,表明此时规划变量影响物质单位调整量的方法失效,无法达到验证效果,只有当表达目标物处于正常变化状态,才能确定出对原始物表达过程中产生的影响。
- [0085] 进一步的,S5.3中规划变量影响物质单位调整量的方法采用平均分算法,其算法公式包括如下步骤:
- [0086] 步骤一、确定表达目标物变化量正常范围,标记为 $[0, N]$;
- [0087] 步骤二、确定变量影响物质初始量状态下初始表达目标物量 N_{IQ} ,并确定叠加单位调整量 M 后形成的叠加变量影响物质质量状态下的叠加表达目标物量 N_{SQ} ,计算叠加表达目标物量 N_{SQ} 与初始表达目标物量 N_{IQ} 的差值,标记为 N_1 ;
- [0088] 步骤三、当 $N_1 \in [0, N]$,输出单位调整量为 M ;
- [0089] 当 $N_1 \notin [0, N]$,输出单位调整量为 $M/2$,确定变量影响物质初始量叠加单位调整量 $M/2$ 产生的中值表达目标物量 N_{MD} ;

[0090] 步骤四、重复步骤三,直至中值表达目标物量 $N_{MD} \in [0, N]$,输出对应的叠加单位调整量。

[0091] 再进一步的,S7中确定出各项影响物质对各项表达目标物产量的表达概率的影响的方法包括如下步骤:

[0092] 第一步、通过步骤二确定各项影响物质初始量 M_0 与其对应单位调整量M的比值率 $M_{ratio} = \frac{M}{M_0}$;

[0093] 第二步、确定各项影响物质叠加位调整量M后产生的叠加表达目标物量 N_{SQ} ,计算出各项影响物质影响值 $N_{effect} = N_{SQ} \times M_{ratio} = N_{SQ} \times \frac{M}{M_0}$;

[0094] 第三步、比对各项影响物质影响值大小,对各项影响物质影响表达目标物产生进行排序。

[0095] 在进行各项影响物质对各项表达目标物产量的表达概率的影响确定过程中,首先确定各项影响物质初始量 M_0 与其对应单位调整量M的比值率 $M_{ratio} = \frac{M}{M_0}$,随后确定各项影响物质叠加位调整量M后产生的叠加表达目标物量 N_{SQ} ,计算出各项影响物质影响值 $N_{effect} = N_{SQ} \times M_{ratio} = N_{SQ} \times \frac{M}{M_0}$,最后比对各项影响物质影响值大小,对各项影响物质影响表达目标物产生进行排序,依次确定各项影响物质的影响程度。

[0096] 此外,S8中卷积神经网络训练集包括输入层、卷积层以及输出层,其中:

[0097] 输入层:用于输入原始物初始培养量、空白组数据、恒量影响物质、变量影响物质、初始表达目标物量 N_{IQ} 、单位调整量M、影响物质初始量 M_0 以及表达目标物变化量正常范围 $[0, N]$;

[0098] 卷积层:规划叠加表达目标物量 N_{SQ} 计算步骤,确定规划叠加表达目标物量 N_{SQ} 是否属于表达目标物变化量正常范围 $[0, N]$,规划比值率计算算法 $M_{ratio} = \frac{M}{M_0}$,生成影响物质影响值计算算法 $N_{effect} = N_{SQ} \times M_{ratio} = N_{SQ} \times \frac{M}{M_0}$;

[0099] 输出层:根据各项计算结果,输出对应数值,并绑定各项变量数据。

[0100] 以上显示和描述了本发明的基本原理、主要特征和本发明的优点。本行业的技术人员应该了解,本发明不受上述实施例的限制,上述实施例和说明书中描述的仅为本发明的优选例,并不用来限制本发明,在不脱离本发明精神和范围的前提下,本发明还会有各种变化和改进,这些变化和改进都落入要求保护的本发明范围内。本发明要求保护范围由所附的权利要求书及其等效物界定。

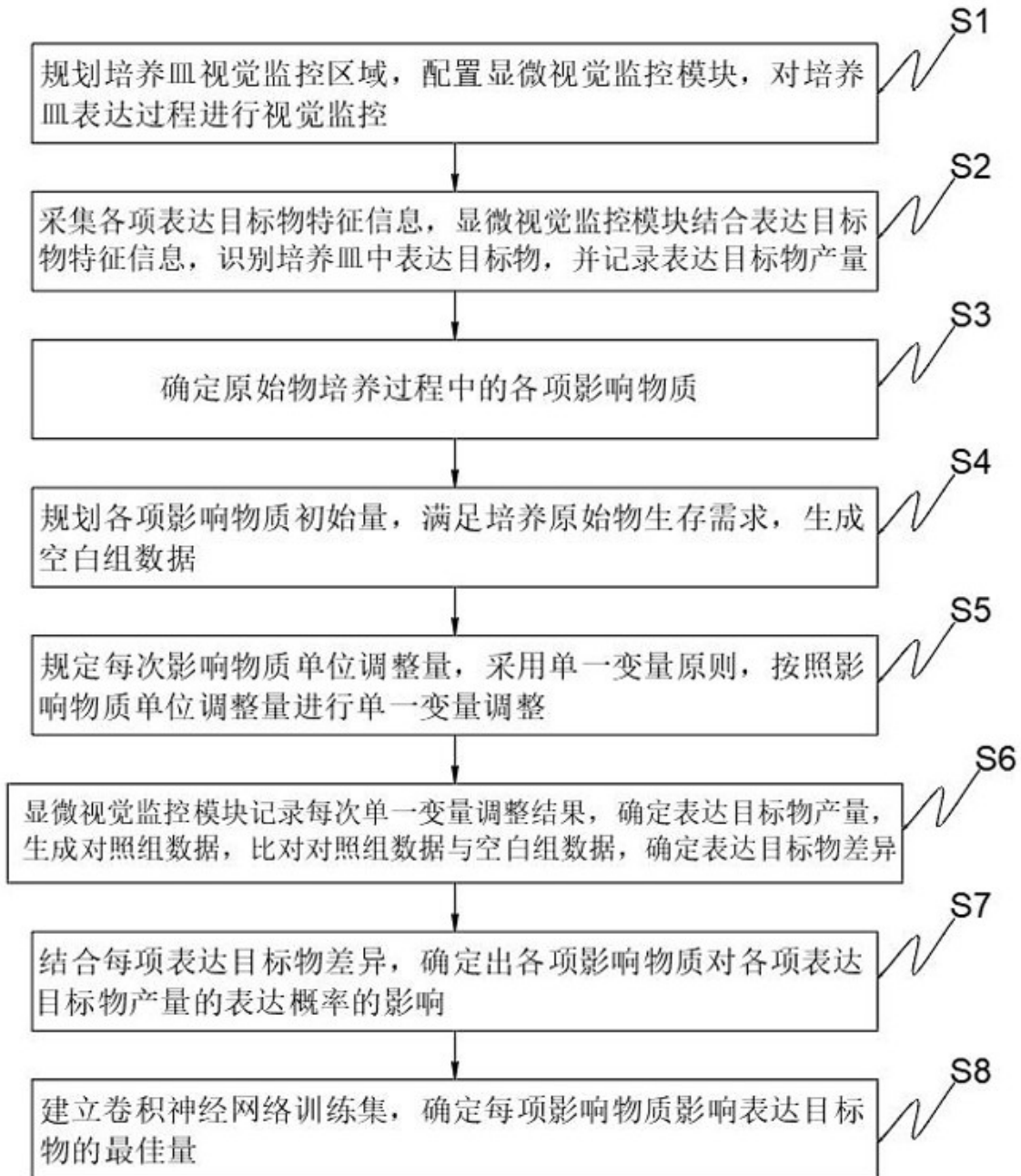


图 1

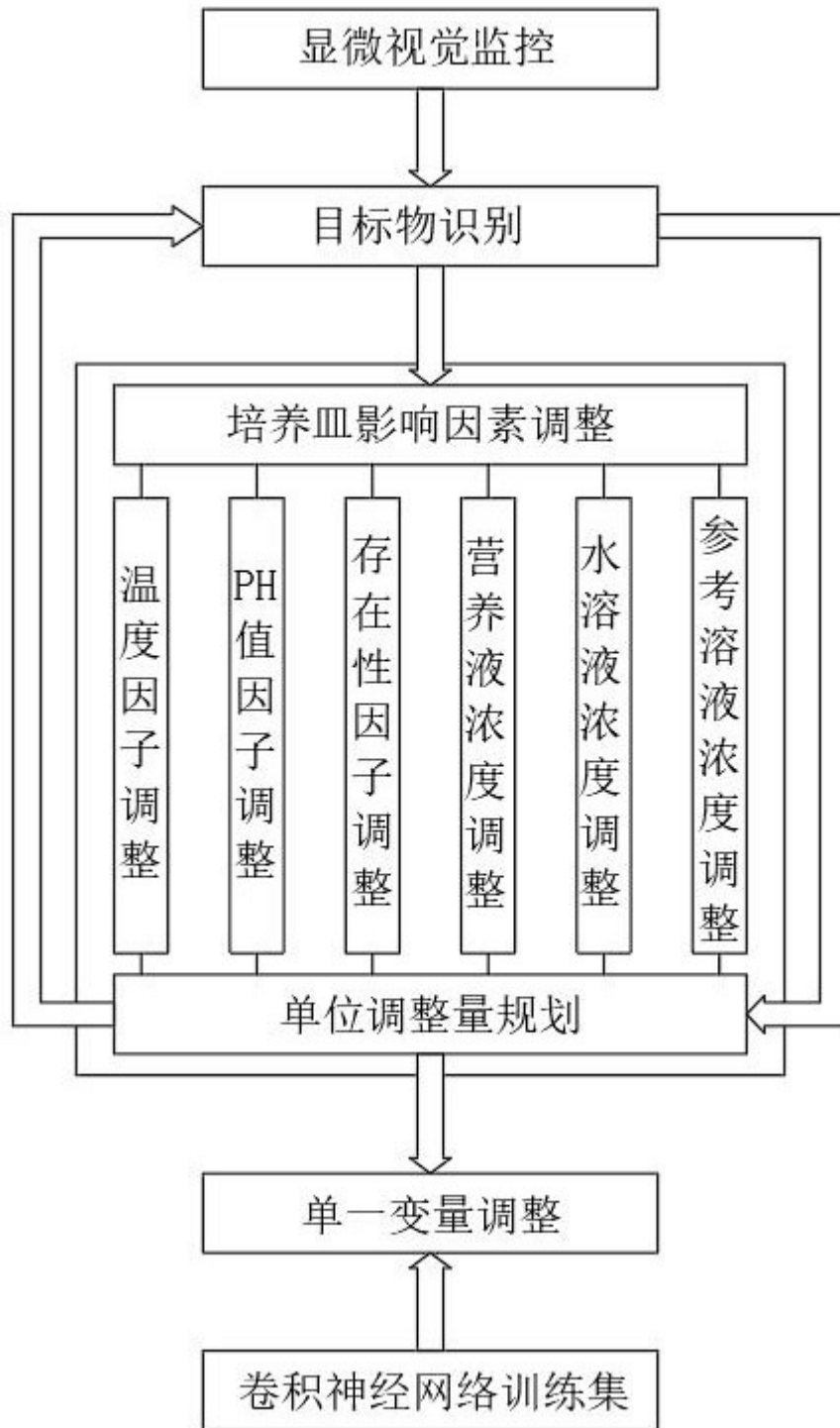


图 2

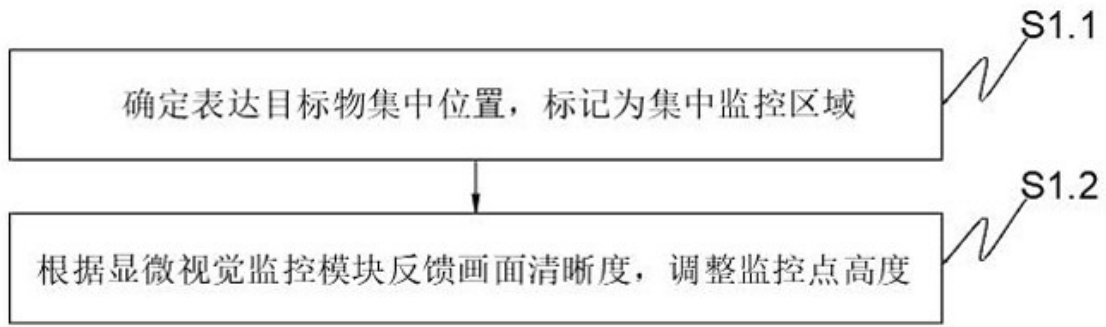


图 3

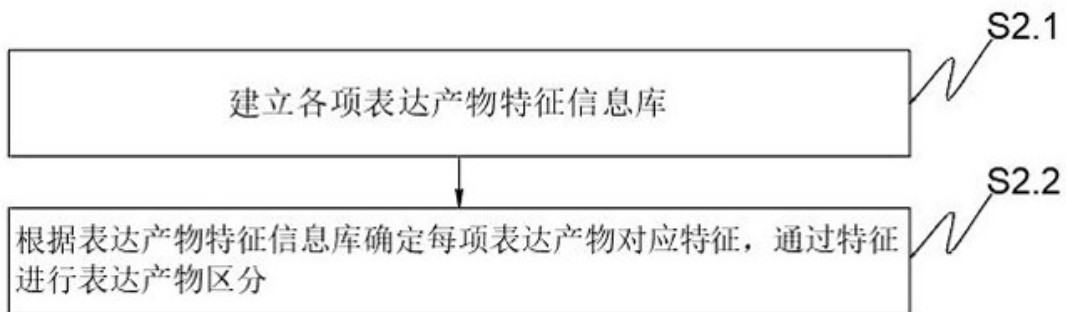


图 4

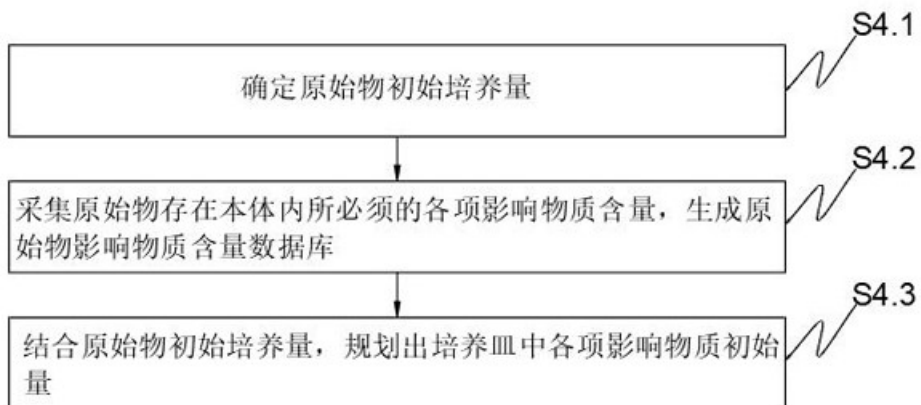


图 5

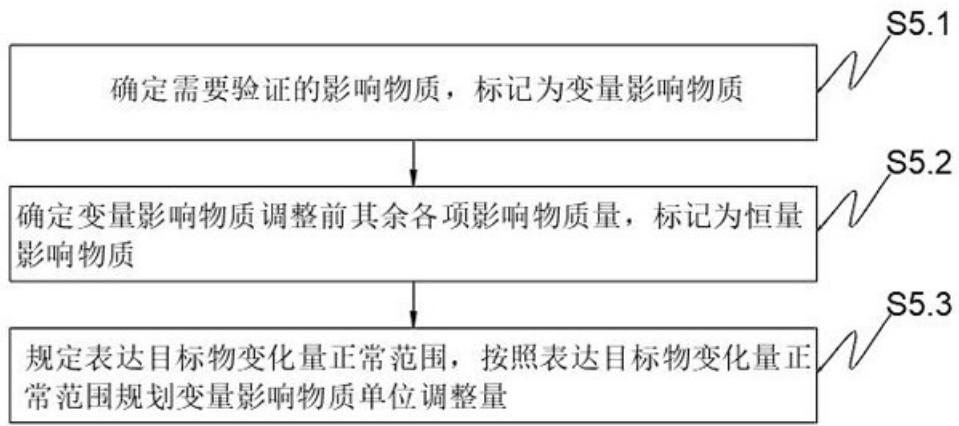


图 6