



(12) 发明专利申请

(10) 申请公布号 CN 104656893 A

(43) 申请公布日 2015. 05. 27

(21) 申请号 201510061485. X

(22) 申请日 2015. 02. 06

(71) 申请人 西北工业大学

地址 710072 陕西省西安市友谊西路 127 号

(72) 发明人 高鹏 白晓亮 张树生 何卫平

杨森

(74) 专利代理机构 西北工业大学专利中心

61204

代理人 王鲜凯

(51) Int. Cl.

G06F 3/01(2006. 01)

G06K 9/00(2006. 01)

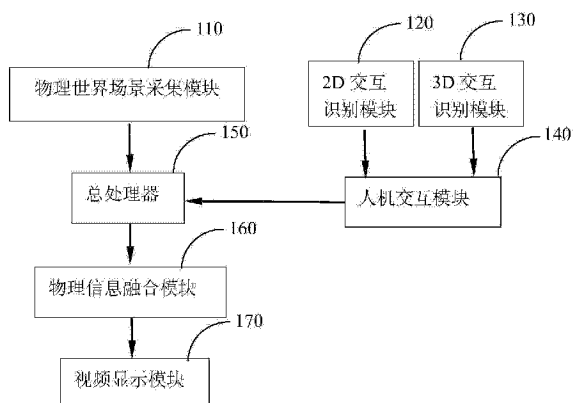
权利要求书2页 说明书6页 附图3页

(54) 发明名称

一种信息物理空间的远程交互式操控系统及方法

(57) 摘要

本发明涉及一种信息物理空间的远程交互式操控系统及方法,该系统包括 3D 交互识别模块、2D 交互识别模块、物理世界场景采集模块、总处理器、物理信息融合模块和视频显示模块;该方法可以远程的通过 2D 与 3D 交互相结合的方式,完成对信息物理融合场景下的物理信息对象进行实时的人机交互控制,与其他方法相比,该方法成本低廉,操作沉浸感强并且操作可以达到实时高效,对于未来的物理信息融合场景下人机交互有着非常广阔的应用前景。



1. 一种信息物理空间的远程交互式操控系统,其特征包括人机交互模块、物理世界场景采集模块、总处理器、物理信息融合模块和视频显示模块;所述的人机交互模块包括3D交互识别模块和2D交互识别模块;3D交互识别模块用于对用户手势动作进行识别,然后再对用户的手势动作数据进行处理,提取出用户的手势动作三维坐标值;2D交互识别模块用于进行3D交互识别模块操控的模式切换和对3D手势交互模块映射在虚拟场景范围进行操控范围的平移;物理世界场景采集模块用于远程拍摄物理世界工作场景,将相关物理对象世界坐标传输给处理器;物理信息融合模块用于将物理世界场景采集模块远程拍摄的相关物理对象世界范围与虚拟场景范围的虚拟场景坐标融合统一在一起;视频显示模块用于显示物理信息融合模块的物理信息场景和人机交互模块的人手参数化模型。

2. 根据权利要求1所述的信息物理空间的远程交互式操控系统,其特征包括所述的2D交互识别模块由一块2D的平面感应板所组成。

3. 根据权利要求1所述的信息物理空间的远程交互式操控系统,其特征包括所述的3D交互识别模块由至少一个微型RGB摄像头和至少一个深度摄像头组成。

4. 根据权利要求1所述的信息物理空间的远程交互式操控系统,其特征包括所述的物理世界场景采集模块由至少一个摄像头组成。

5. 根据权利要求1所述的信息物理空间的远程交互式操控系统,其特征包括所述的视频显示模块为手机显示屏、电脑显示屏或眼镜式微显示屏。

6. 一种信息物理空间的远程交互式操控方法,其特征包括下述步骤:

步骤1:物理世界场景采集模块获取世界物理场景;

步骤2:总处理器将获取的物理世界场景与信息对象融合在一起,转换成物理信息融合场景;

步骤3:人机交互模块捕获用户2D和3D的交互指令;

步骤4:总处理器接收用户2D和3D的交互指令,并将交互指令转化为控制指令;

步骤5:总处理器根据控制指令控制物理信息目标物体对象的变化;

步骤6:视频显示模块显示物理信息融合场景和人机交互模块控制对象变化的实时显示。

7. 根据权利要求6所述的信息物理空间的远程交互式操控方法,其特征包括所述的步骤3包括下述步骤:

步骤3a:人机交互模块中的2D交互识别模块识别用户2D交互指令;

步骤3b:人机交互模块中的3D交互识别模块识别用户3D交互指令;

步骤3c:2D交互识别模块将平面感应板坐标系下的一点坐标转换至物理信息融合场景下的图像平面坐标;

$$\begin{bmatrix} \mu'_a \\ \nu'_a \end{bmatrix} = \begin{bmatrix} s_x \times (n'_x x'_p + o'_x y'_p) + p'_x \\ s_y \times (n'_y x'_p + o'_y y'_p) + p'_y \end{bmatrix}$$

其中 (x'_p, y'_p) 为平面感应板坐标系下的一点坐标; (μ'_a, ν'_a) 为 (x'_p, y'_p) 映射到物理信息融合场景下的一点图像坐标;

s_x, s_y 分别为平面感应板上单位距离到物理信息融合场景中的显示图像在 X 轴和 Y 轴的放缩比例系数;

$\vec{n}' = [n'_x \ n'_y]^T$ 是平面感应板 X 轴在物理信息融合场景中的方向向量；

$\vec{o}' = [o'_x \ o'_y]^T$ 是平面感应板 Y 轴在物理信息融合场景中的方向向量；

$\vec{p}' = [\vec{p}'_x \ \vec{p}'_y]^T$ 是平面感应板坐标原点在物理信息融合场景中的位置；

步骤 4 :3D 交互识别模块将世界空间中任意一点三维坐标转换至物理信息融合场景中的图像坐标；

对相机进行内外参数的标定,计算出相机的内外参数相关矩阵参数;世界空间的任意一点,通过相机内外参数矩阵,映射到图像上一点对应的图像坐标,现实在虚拟场景之中;世界空间中任意一点三维坐标为 $A(x_{wa}, y_{wa}, z_{wa})$,映射到物理信息融合场景中的图像坐标为 $A'(\mu_a, v_a)$,则有以下关系:

$$\begin{bmatrix} u_a \\ v_a \\ 1 \end{bmatrix} = \begin{bmatrix} k_x \times \frac{(n_x x_{wa} + o_x y_{wa} + a_x z_{wa} + p_x)}{z_{ca}} + \mu_0 \\ k_y \times \frac{(n_y x_{wa} + o_y y_{wa} + a_y z_{wa} + p_y)}{z_{ca}} + \nu_0 \\ \frac{(n_y x_{wa} + o_y y_{wa} + a_y z_{wa} + p_y)}{z_{ca}} \end{bmatrix}$$

其中 $k_x = \sigma_x f$; $k_y = \sigma_y f$; σ_x, σ_y 分别为成像平面到图像平面上 x, y 两个方向上的缩放比例系数; f 为焦距;

(μ_0, ν_0) 为图像物理坐标系原点在图像坐标系中的坐标;

$\vec{n} = [n_x \ n_y \ n_z]^T$ 是 X 轴在相机坐标系中的方向向量;

$\vec{o} = [o_x \ o_y \ o_z]^T$ 是 Y 轴在相机坐标系中的方向向量;

$\vec{a} = [a_x \ a_y \ a_z]^T$ 是 Z 轴在相机坐标系中的方向向量;

$\vec{p} = [p_x \ p_y \ p_z]^T$ 是世界坐标系原点在相机坐标系中的位置;

d_x, d_y 为每一个像素点在 x, y 轴方向上的物理尺寸,则在虚拟场景显示图像与世界坐标的在 x, y 的比例系数分别为: $h_x = d_x \times \frac{u_a - u_b}{x_{wa} - x_{wb}}$; $h_y = d_y \times \frac{v_a - v_b}{y_{wa} - y_{wb}}$;

步骤 3d :总处理器将 2D 交互识别模块映射过来的图像平面坐标与物理信息融合模块内的物理信息融合场景的图像平面坐标进行匹配;

步骤 3e :总处理器将 3D 交互识别模块映射过来的图像平面坐标与物理信息融合模块内的物理信息融合场景的图像平面坐标进行匹配;

步骤 3f :总处理器将 2D 和 3D 交互识别模块进行融合匹配,完成物理信息融合场景下的人机交互。

一种信息物理空间的远程交互式操控系统及方法

技术领域

[0001] 本发明涉及物理信息融合环境下的人机交互技术,特别涉及一种人机交互系统与
人机交互方法。

背景技术

[0002] 物理信息融合技术,也就是虚拟现实是利用电脑模拟产生一个三维空间的虚拟世界,提供使用者关于视觉、听觉、触觉等感官的模拟,让使用者如同身历其境一般,可以及时、没有限制地观察三度空间内的事物。

[0003] 目前,在现有的虚拟现实系统中,只是将虚拟信息静态的叠加现实在虚拟现实场景中,而不能对物理信息融合场景下的虚拟现实模型进行实时的人机操控。

发明内容

[0004] 要解决的技术问题

[0005] 为了避免现有技术的不足之处,本发明提出一种在虚拟现实融合系统下的实时人机交互系统和方法。该系统可以远程的通过 2D 与 3D 交互相结合的方式,完成对信息物理融合场景下的物理信息对象进行实时的人机交互控制。

[0006] 技术方案

[0007] 一种信息物理空间的远程交互式操控系统,其特征在于包括人机交互模块、物理世界场景采集模块、总处理器、物理信息融合模块和视频显示模块;所述的人机交互模块包括 3D 交互识别模块和 2D 交互识别模块;3D 交互识别模块用于对用户手势动作进行识别,然后再对用户的手势动作数据进行处理,提取出用户的手势动作三维坐标值;2D 交互识别模块用于进行 3D 交互识别模块操控的模式切换和对 3D 手势交互模块映射在虚拟场景范围进行操控范围的平移;物理世界场景采集模块用于远程拍摄物理世界工作场景,将相关物理对象世界坐标传输给处理器;物理信息融合模块用于将物理世界场景采集模块远程拍摄的相关物理对象世界范围与虚拟场景范围的虚拟场景坐标融合统一在一起;视频显示模块用于显示物理信息场景融合模块的物理信息场景和人机交互模块的人手参数化模型。

[0008] 所述的 2D 交互识别模块由一块 2D 的平面感应板所组成。

[0009] 所述的 3D 交互识别模块由至少一个微型 RGB 摄像头和至少一个深度摄像头组成。

[0010] 所述的物理世界场景采集模块由至少一个摄像头组成。

[0011] 所述的视频显示模块为手机平板显示屏、电脑显示屏或眼镜式微显示屏。

[0012] 一种信息物理空间的远程交互式操控方法,其特征在于包括下述步骤:

[0013] 步骤 1:物理世界场景采集模块获取世界物理场景;

[0014] 步骤 2:总处理器将获取的物理世界场景与信息对象融合在一起,转换成物理信息融合场景;

[0015] 步骤 3:人机交互模块捕获用户 2D 和 3D 的交互指令;

[0016] 步骤 4:总处理器接收用户 2D 和 3D 的交互指令,并将交互指令转化为控制指令;

[0017] 步骤 5 :总处理器根据控制指令控制物理信息目标物体对象的变化 ;

[0018] 步骤 6 :视频显示模块显示物理信息融合场景和人机交互模块控制对象变化的实时显示。

[0019] 所述的步骤 3 包括下述步骤 :

[0020] 步骤 3a :人机交互模块中的 2D 交互识别模块识别用户 2D 交互指令 ;

[0021] 步骤 3b :人机交互模块中的 3D 交互识别模块识别用户 3D 交互指令 ;

[0022] 步骤 3c :2D 交互识别模块将平面感应板坐标系下的一点坐标转换至物理信息融合场景下的图像平面坐标 ;

$$[0023] \quad \begin{bmatrix} \mu'_a \\ \nu'_a \end{bmatrix} = \begin{bmatrix} s_x \times (n'_x x'_p + o'_x y'_p) + p'_x \\ s_y \times (n'_y x'_p + o'_y y'_p) + p'_y \end{bmatrix}$$

[0024] 其中 (x'_p, y'_p) 为平面感应板坐标系下的一点坐标 ; (μ'_a, ν'_a) 为 (x'_p, y'_p) 映射到物理信息融合场景下的一点图像坐标 ;

[0025] s_x, s_y 分别为平面感应板上单位距离到物理信息融合场景中的现实图像在 X 轴和 Y 轴的放缩比例系数 ;

[0026] $\vec{n}' = [n'_x \quad n'_y]^T$ 是平面感应板 X 轴在物理信息融合场景中的方向向量 ;

[0027] $\vec{o}' = [o'_x \quad o'_y]^T$ 是平面感应板 Y 轴在物理信息融合场景中的方向向量 ;

[0028] $\vec{p}' = [\vec{p}'_x \quad \vec{p}'_y]^T$ 是平面感应板坐标原点在物理信息融合场景中的位置 ;

[0029] 步骤 4 :3D 交互识别模块将世界空间中任意一点三维坐标转换至物理信息融合场景中的图像坐标 ;

[0030] 对相机进行内外参数的标定,计算出相机的内外参数相关矩阵参数 ;世界空间的任意一点,通过相机内外参数矩阵,映射到图像上一点对应的图像坐标,现实在虚拟场景之中 ;世界空间中任意一点三维坐标为 $A(x_{wa}, y_{wa}, z_{wa})$,映射到物理信息融合场景中的图像坐标为 $A'(\mu_a, \nu_a)$,则有以下关系 :

$$[0031] \quad \begin{bmatrix} u_a \\ v_a \\ 1 \end{bmatrix} = \begin{bmatrix} k_x \times \frac{(n_x x_{wa} + o_x y_{wa} + a_x z_{wa} + p_x)}{z_{ca}} + \mu_0 \\ k_y \times \frac{(n_y x_{wa} + o_y y_{wa} + a_y z_{wa} + p_y)}{z_{ca}} + \nu_0 \\ \frac{(n_y x_{wa} + o_y y_{wa} + a_y z_{wa} + p_y)}{z_{ca}} \end{bmatrix}$$

[0032] 其中 $k_x = \sigma_x f ; k_y = \sigma_y f ; \sigma_x, \sigma_y$ 分别为成像平面到显示图像上 x, y 两个方向上的放缩比例系数 ; f 为焦距 ;

[0033] (μ_0, ν_0) 为图像物理坐标系原点在图像坐标系中的坐标 ;

[0034] $\vec{n} = [n_x \quad n_y \quad n_z]^T$ 是 X 轴在相机坐标系中的方向向量 ;

[0035] $\vec{o} = [o_x \quad o_y \quad o_z]^T$ 是 Y 轴在相机坐标系中的方向向量；

[0036] $\vec{a} = [a_x \quad a_y \quad a_z]^T$ 是 Z 轴在相机坐标系中的方向向量；

[0037] $\vec{p} = [p_x \quad p_y \quad p_z]^T$ 是世界坐标系原点在相机坐标系中的位置；

[0038] d_x, d_y 为每一个像素点在 x, y 轴方向上的物理尺寸, 则在虚拟场景显示图像与世界坐标的在 x, y 的比例系数分别为: $h_x = d_x \times \frac{u_a - u_b}{x_{wa} - x_{wb}}$; $h_y = d_y \times \frac{v_a - v_b}{y_{wa} - y_{wb}}$;

[0039] 步骤 3d: 总处理器将 2D 交互识别模块映射过来的图像平面坐标与物理信息融合模块内的物理信息融合场景的图像平面坐标进行匹配；

[0040] 步骤 3e: 总处理器将 3D 交互识别模块映射过来的图像平面坐标与物理信息融合模块内的物理信息融合场景的图像平面坐标进行匹配；

[0041] 步骤 3f: 总处理器将 2D 和 3D 交互识别模块进行融合匹配, 完成物理信息融合场景下的人机交互。

[0042] 有益效果

[0043] 本发明提出的一种信息物理空间的远程交互式操控系统及方法, 该方法可以远程的通过 2D 与 3D 交互相结合的方式, 完成对信息物理融合场景下的物理信息对象进行实时的人机交互控制, 与其他方法相比, 该方法成本低廉, 操作沉浸感强并且操作可以达到实时高效, 对于未来的物理信息融合场景下人机交互有着非常广阔的应用前景。

附图说明

[0044] 图 1 为实施例提供的一种信息物理空间的远程交互式操控系统结构示意图

[0045] 图 2 为本发明实施例提供的一种信息物理空间的远程交互式操控方法步骤流程图

[0046] 图 3 为本发明实施例提供的人机交互模块的结构流程图

具体实施方式

[0047] 现结合实施例、附图对本发明作进一步描述：

[0048] 图 1 为本发明实施例提供的一种信息物理空间的远程交互式操控设备结构示意图。

[0049] 一种信息物理空间的远程交互式操控设备 100, 包括: 物理工作场景采集模块 110、物理信息融合模块 160、2D 交互识别模块 120、2D 交互识别模块 130、人机交互模块 140、总处理器 150、视频显示模块 170。

[0050] 物理工作场景采集模块 110: 用于拍摄物理工作场景。在本发明提供的实施例中, 物理工作场景采集模块 110 优选为有至少一个摄像头组成, 世界坐标内的物理对象通过摄像头捕获；

[0051] 2D 交互识别模块 120: 用于进行 3D 交互识别模块操控的模式切换和对 3D 手势交互模块映射在虚拟场景范围进行操控范围的平移。在本发明提供的实施例中, 2D 交互识别模块 120 优选为由平板电脑或手机组成, 对平板电脑或手机进行开发, 通过 socket 通信与

总处理器 150 进行数据传输；

[0052] 3D 交互识别模块 130 :用于对用户手势动作进行识别,然后再对用户的手势动作数据进行处理,提取出用户的手势动作三维坐标值。在本发明提供的实施例中,2D 交互识别模块 130 优选为由至少一个深度摄像头和至少一个 RGB 摄像头组成,用户的手势动作以及手的世界坐标被两摄像头捕获,经过 2D 交互识别模块 130 中的微型处理器处理后,将数据传输给总服务器,与物理信息融合模块 160 融合的物理信息融合场景进行坐标统一；

[0053] 人机交互模块 140 :用于将 2D 交互识别模块 120 和 2D 交互识别模块 130 的数据经过人机交互模块 140 内嵌的为处理器进行处理,然后传输给总处理器 150；

[0054] 总处理器 150 :用于计算和处理物理信息融合模块 160 融合的物理信息场景以及人机交互模块 140 的相关数据；

[0055] 物理信息融合模块 160 :用于将物理工作场景采集模块 110 远程拍摄的相关世界物理对象和世界物理场景范围的物理世界坐标和虚拟场景的虚拟信息对象和虚拟场景范围的虚拟场景坐标,经过总处理器 150 进行数据处理,然后将坐标融合统一在物理信息融合场景中；

[0056] 物理信息融合模块 160 :还用于将人机交互模块 140 的手势动作三维坐标值统一融合在物理信息融合场景中；

[0057] 视频显示模块 170 :用于显示物理信息场景融合模块 120 的物理信息场景和人机交互模块 140 的人手参数化模型。本发明提供的实施例中视频显示模块 170 优选为电脑显示屏或者眼镜式微显示屏,由物理工作场景采集模块 110 捕获的世界坐标内的物理对象和人机交互模块 140 捕获的人手参数化模型经视频显示模块 170 显示。可以理解,由上述 2D 交互识别模块 120 和 2D 交互识别模块 130 控制指令控制物理信息融合场景的物理信息对象的变化过程及结果也经视频显示模块 170 显示。

[0058] 请参阅图 2 为本发明实施例提供的一种信息物理空间的远程交互式操控方法步骤流程图,其包括下述步骤：

[0059] 步骤 S210 :物理工作场景采集模块 110 世界物理场景。在本发明提供的实施例中,物理工作场景采集模块 110 由至少一个摄像头组成,用于采集世界物理场景,将采集到的图像传输给总处理器 150 进行处理。

[0060] 步骤 S220 :总处理器 150 将获取的物理工作场景与信息对象融合在一起,转换成物理信息融合场景。在本发明实施例中,总处理器 150 将物理工作场景采集模块 110 捕获的世界物理场景和世界物理对象与虚拟信息场景和虚拟信息对象进行融合,通过目前常用的坐标转换方法,将世界物理场景和世界物理对象与虚拟信息场景和虚拟信息对象坐标范围统一,再经过物理信息融合模块 160,使得物理信息场景和对象融合在一起,通过视频显示模块 170 实时显示出来。

[0061] 步骤 S230 :人机交互模块 140 捕获用户 2D 和 3D 的交互指令。人机交互模块 140 是由 2D 交互识别模块 120 和 3D 交互识别模块 130 共同组成的。2D 交互识别模块 120 主要用于提供菜单进行交互模式切换、2D 的交互操作,如 2D 范围内的平移、旋转、缩放等；3D 交互识别模块 130 主要用于捕捉用户的手势动作以及用户人手的世界坐标位置信息,然后将捕捉到的数据传送给总处理器 150,与物理信息融合模块 160 中的物理信息对象进行人机交互；其中 3D 手势根据用户操作需要可以设定,如手掌张开与合拢手势完成对物理信息

融合模块 160 中的物理信息对象拾取与脱离,手掌上下左右平移完成对物理信息融合模块 160 中的物理信息对象的任意角度旋转等等;在本发明方法中,2D 交互识别模块 120 和 3D 交互识别模块 130 可以结合使用,共同组成人机交互模块 140;完成对物理信息融合模块 160 中的物理信息对象的任何设定的操作以及对操作范围的任意平移切换。

[0062] 步骤 S240:总处理器 150 接收用户的手势动作及其他交互信息,并将交互信息转化为控制指令。在本发明实施例中,总处理器 150 同时接收 2D 交互识别模块 120 和 3D 交互识别模块 130 的交互信息,并将其交互信息转化为控制指令。

[0063] 步骤 S250:总处理器 150 根据控制指令控制物理信息目标物体对象的变化。在本发明实施例中,控制指令由人机交互模块 140 实时的通过 socket 通讯与总处理器 150 进行传输。

[0064] 步骤 S260:视频显示模块 170 显示物理信息融合场景和人机交互模块 140 控制对象变化的实时显示。在发明实施例中,视频显示模块 170 可以实时的显示人机交互模块 140 控制物理信息对象接受控制指令,并且将物理信息对象执行人机交互模块 140 的控制指令,实时地显示出来。

[0065] 图 3 为本发明实施例提供的人机交互模块 140 的结构流程图。包括以下步骤:

[0066] 步骤 S310:识别用户 2D 交互指令。在本发明实施例中,2D 交互识别模块优选为平面感应板,用户通过平面感应板输入 2D 交互指令,控制物理信息融合模块 160 中的物理信息对象执行操作指令。例如,用户通过手指滑动平面感应板可以控制控制物理信息融合模块 160 中的物理信息操作范围进行平移,用户通过平面感应板的按钮可以控制 3D 交互识别模块 130 的操作模式的切换等;

[0067] 步骤 S320:3D 交互识别模块 130 识别用户 3D 交互指令。在本实施例中,3D 交互识别模块 130 优选为一个微型 RGB 摄像头和一个微型深度摄像头,用于捕捉人手的三维手势指令,然后通过 3D 交互识别模块 130 中的手势库完成 3D 交互指令的输入。

[0068] 步骤 S330:2D 操作的坐标映射关系转换至图像平面坐标。在本发明实施例中,通过权利要求书中的 2D 操作的坐标映射关系转换至物理信息融合场景下的图像平面坐标的方法,将 2D 交互识别模块 120 在平面感应板上的 2D 交互范围映射到物理信息融合模块 160 中 3D 空间的 2D 操作平面上。

[0069] 步骤 S340:3D 操作的坐标映射关系转换至图像平面坐标。在本发明实施例中,通过权利要求书中的 3D 操作的坐标映射关系转换至物理信息融合场景下的图像平面坐标的方法,将 3D 交互识别模块 130 在世界空间范围内的 3D 操作范围映射到物理信息融合模块 160 中的 3D 空间的 3D 操作空间。

[0070] 步骤 S350:通过总处理器 150,将 2D 交互识别模块 120 映射过来的图像平面坐标与物理信息融合模块 160 内的物理信息融合场景的图像平面坐标进行匹配,使 2D 交互识别模块 120 的 2D 交互范围与物理信息融合模块 160 内的物理信息融合场景中的 3D 中的 2D 操作范围进行统一。

[0071] 步骤 S360:通过总处理器 150,将 3D 交互识别模块 130 映射过来的图像平面坐标与物理信息融合模块 160 内的物理信息融合场景的图像平面坐标进行匹配,上述中的图像平面坐标与物理信息融合模块 160 内的物理信息融合场景的图像平面坐标进行匹配,使 3D 交互识别模块 120 的 3D 交互范围与物理信息融合模块 160 内的物理信息融合场景中的 3D

中的 3D 操作范围进行统一。

[0072] 步骤 S370 :总处理器 150 将 2D 和 3D 交互模块进行融合匹配,完成物理信息融合场景下的人机交互。在本发明实施例中,2D 和 3D 交互模块是相辅相成,相互协调统一的,因为在交互过程中,有些操作 2D 交互容易完成,而有些操作则用 3D 交互容易完成,将两者结合使用,可以非常高效的完成交互。

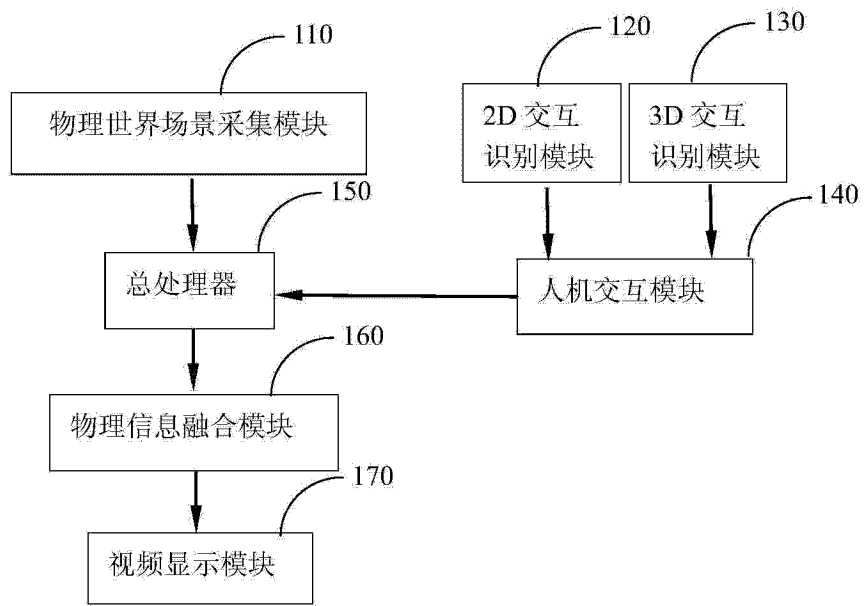


图 1

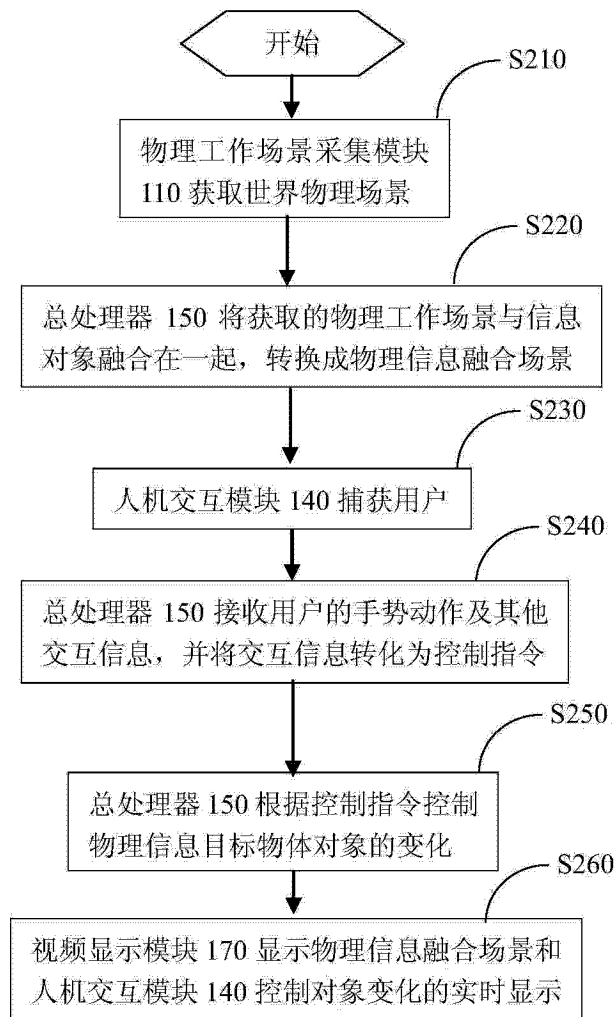


图 2

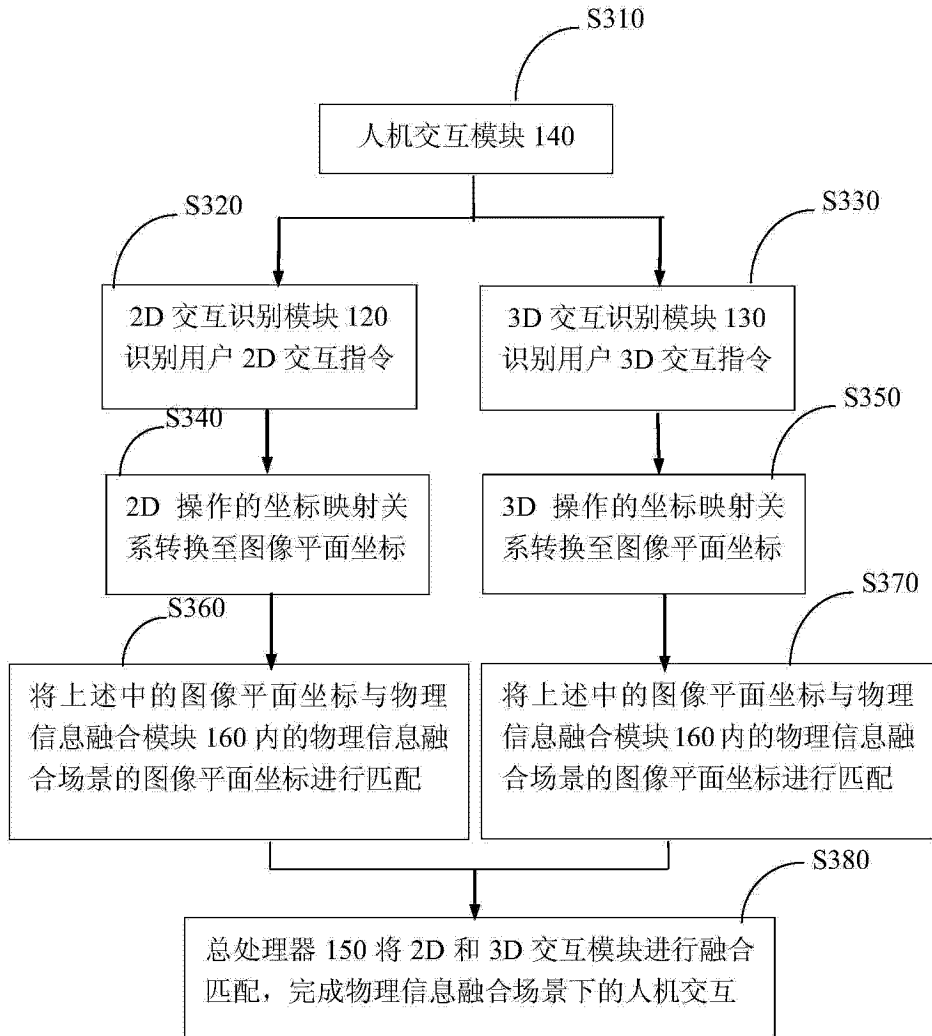


图 3