



## (12) 发明专利申请

(10) 申请公布号 CN 104484034 A

(43) 申请公布日 2015. 04. 01

(21) 申请号 201410711947. 3

G06T 7/20(2006. 01)

(22) 申请日 2014. 11. 27

(71) 申请人 韩慧健

地址 250014 山东省济南市二环东路 7366  
号

(72) 发明人 韩慧健 张锐 贾可亮 刘峥  
梁秀霞

其他发明人请求不公开姓名

(74) 专利代理机构 北京元本知识产权代理事务  
所 11308

代理人 秦力军

(51) Int. Cl.

G06F 3/01(2006. 01)

G06K 9/62(2006. 01)

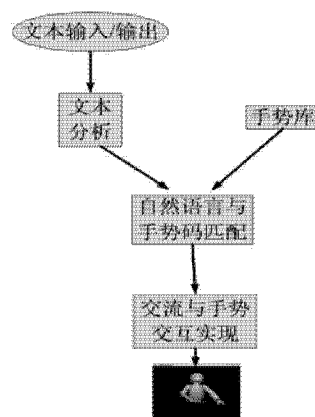
权利要求书2页 说明书8页 附图1页

### (54) 发明名称

一种基于手势识别的手势运动基元过渡帧定位方法

### (57) 摘要

一种基于手势识别的手势运动基元过渡帧定位方法,包括利用视频或图像采集装置获取手势图像,进行预处理,将人物图像按照轮廓、颜色、纹理分割为轮廓图像、颜色图像和纹理图像;将经过处理的轮廓图像、颜色图像和纹理图像进行再合成,建立手势图像数据库;基于手势图像数据库中的手势图像,定位到过渡帧,通过关键帧插值进行平滑连接合成连贯的手势动作,建立与手势图像匹配的手势运动基元模型。基于手势识别的手势运动基元过渡帧定位方法基于真人图像建立,画面生动真实,并且虚拟人模型灵活多变,能够直观、生动的表现,消除了运动跨度过大导致运动过渡的难题并有效地实现了运动细节的增强与保持。



1. 一种基于手势识别的手势运动基元过渡帧定位方法,其特征在于,依次包括以下步骤:

(1) 利用视频或图像采集装置获取一个或多个分辨率为  $N \times M$  的手势图像,对一个或多个手势图像进行预处理,将一个或多个手势图像中的人物图像提取出来,删除除了人物图像部分之外的背景图像;

(2) 对提取后的人物图像进行去噪处理,消除传输和采集过程中的噪声干扰,之后将去噪后的人物图像按照轮廓、颜色、纹理分割为轮廓图像、颜色图像和纹理图像;

(3) 对轮廓图像进行修正,进行平滑处理;

(4) 将颜色图像按照 RGB 图像的 R、G、B 三个分量通道进行分解,分别将三个分量通道中的每一个像素与其邻近的 8 个像素构成  $3 \times 3$  的像素块,按照如下公式分别得到处理后的每一个像素分别对应的 R、G、B 三个分量的值:

$$P_{ij}(R) = \left( \frac{P_{(i-1)(j-1)} + P_{(i-1)j} + P_{(i-1)(j+1)} + P_{(i+1)(j-1)} + P_{(i+1)j} + P_{(i+1)(j+1)} + P_{(i-1)(j-1)} + P_{(i-1)j} + P_{(i-1)(j+1)} + P_{(i+1)(j-1)} + P_{(i+1)j} + P_{(i+1)(j+1)} + P_{(i-1)(j-1)} + P_{(i-1)j} + P_{(i-1)(j+1)} + P_{(i+1)(j-1)} + P_{(i+1)j} + P_{(i+1)(j+1)}}{4} + P_{ij} \right) / 2$$

$$P_{ij}(G) = \left( \frac{P_{(i-1)(j-1)} + P_{(i-1)j} + P_{(i-1)(j+1)} + P_{(i+1)(j-1)} + P_{(i+1)j} + P_{(i+1)(j+1)} + P_{(i-1)(j-1)} + P_{(i-1)j} + P_{(i-1)(j+1)} + P_{(i+1)(j-1)} + P_{(i+1)j} + P_{(i+1)(j+1)} + P_{(i-1)(j-1)} + P_{(i-1)j} + P_{(i-1)(j+1)} + P_{(i+1)(j-1)} + P_{(i+1)j} + P_{(i+1)(j+1)}}{4} + P_{ij} \right) / 2$$

$$P_{ij}(B) = \left( \frac{P_{(i-1)(j-1)} + P_{(i-1)j} + P_{(i-1)(j+1)} + P_{(i+1)(j-1)} + P_{(i+1)j} + P_{(i+1)(j+1)} + P_{(i-1)(j-1)} + P_{(i-1)j} + P_{(i-1)(j+1)} + P_{(i+1)(j-1)} + P_{(i+1)j} + P_{(i+1)(j+1)} + P_{(i-1)(j-1)} + P_{(i-1)j} + P_{(i-1)(j+1)} + P_{(i+1)(j-1)} + P_{(i+1)j} + P_{(i+1)(j+1)}}{4} + P_{ij} \right) / 2$$

$P_{ij}(R)$ 、 $P_{ij}(G)$ 、 $P_{ij}(B)$  分别为经过处理后的 R、G、B 三个分量通道中的像素值,  $P_{ij}$  为 R、G、B 三个分量通道中对应第  $i$  行、第  $j$  列的像素值,其中  $i = 1, 2, \dots, N$ ,  $j = 1, 2, \dots, M$ ;

(5) 将经过处理后的 R、G、B 三个分量通道中的像素值进行分配,合成经过处理后的颜色图像;

(6) 对纹理图像进行修正,进行平滑处理;

(7) 将经过处理的轮廓图像、颜色图像和纹理图像进行再合成,建立手势图像数据库;

(8) 基于手势图像数据库中的手势图像,按照身体的生理结构以及各个关节运动时的约束关系,以关键帧的方式建立手势动作的基本单位,并且描述所述基本单位的时序变化信息,具体包括如下步骤:

对两个待合成手势运动基元  $\Phi_i$  和  $\Phi_j$ , 假设长度分别为  $m$  和  $n$ ;

a. 提取手势运动基元  $\Phi_i$  运动恢复时段的关键帧序列 F1, 假设长度为  $len1$ ;

b. 提取手势运动基元  $\Phi_j$  运动准备时段关键帧序列 F2, 假设长度为  $len2$ ;

c. 对运动关键帧序列 F1, 从起始关键帧位置 1, 开始移动到末尾关键帧位置  $len1$ , 对每一个位置  $k$ , 依次循环计算关键帧序列 F2 中第  $len2-i$  个关键帧与运动关键帧序列 F1 中第  $k$  关键帧的相似度  $Sim$ , 如果存在相似度小于阈值  $\varepsilon$ , 则终止循环, 否则直至循环结束, 定位到过渡帧;

(9) 通过关键帧插值进行平滑连接合成连贯的手势动作, 建立与手势图像匹配的手势运动基元模型;

(10) 将手势运动基元模型和图像数据库中的手势图像对应合成, 建立虚拟人数据库;

(11) 将虚拟人数据库中的数据上传至云数据库单元, 所述云数据库单元将接收到的数据存储后作为备份数据, 供远程客户端下载; 远程客户端下载云数据库单元中存储的数据, 进行自定义编辑, 将自定义编辑后的数据回传至虚拟人数据库, 对虚拟人数据库中的数据

进行更新。

2. 如权利要求 1 所述的基于手势识别的手势运动基元过渡帧定位方法,其特征在于:所述步骤(8)中以关键帧的方式建立手势动作的基本单位,并且描述所述基本单位的时序变化信息步骤中,还包括为手势运动基元模型中的关键帧添加 flag 标记位,将手势运动基元按具有时间次序关键帧标注“start”,“prepare”,“achieve\_goal”,“release\_goal”,“end”,并且将每个手势运动基元分解为四个运动时段:准备阶段、伸出阶段、后伸出阶段、恢复阶段。

3. 如权利要求 1 所述的基于手势识别的手势运动基元过渡帧定位方法,其特征在于:所述约束关系采用四元数表示关节旋转。

4. 如权利要求 1 所述的基于手势识别的手势运动基元过渡帧定位方法,其特征在于:对手势关键帧  $G_i$  与  $G_j$  相似度  $\text{Sim}(G_i, G_j)$  计算如下: $\text{Sim}(G_i, G_j) = \text{Dist}(p_i, p_j) + v\text{Dist}(v_i, v_j)$ , 其中  $\text{Dist}(p_i, p_j)$  表示各个关节角度加权距离,  $\text{Dist}(v_i, v_j)$  表示各关节旋转加速度的加权距离,参数  $v$  表示不同关节位置对速度差异的影响权重。

5. 如权利要求 1 所述的基于手势识别的手势运动基元过渡帧定位方法,其特征在于:还包括步骤(12):在手势与语言协同表达的手势动画合成过程中,根据转写规则在手势运动基元全集上存在一个运动基元覆盖路径,在以上运动基元覆盖路径下实现基于运动基元连接的手势动画合成。

6. 如权利要求 5 所述的基于手势识别的手势运动基元过渡帧定位方法,其特征在于:通过 MacSHAPA 获取手势与自然语言的映射匹配。

## 一种基于手势识别的手势运动基元过渡帧定位方法

### 技术领域

[0001] 本发明涉及手势动画研究领域,更具体地说涉及到手势动画合成中基于手势识别的手势运动基元过渡帧定位方法。

### 背景技术

[0002] 语言伴随性手势合成是以自然语言理解为基础,使用手势作为虚拟人语言信息表达的辅助输出方式。已有手势合成研究主要集中在以自然文本汉语词与手语词映射为基础的手语动画合成,并未考虑语言伴随手势,使得虚拟人类似于机器人一样机械运动,无法满足人机交互的智能性和交互性要求,在交互性和真实感方面还有待提升。在人机交互过程中,通过语言伴随性手势合成研究能够为人机交互中手势和语音协同表达提供解决方案,提高手势动画的真实感,实现虚拟人生动、形象、直观的视觉运动效果。

[0003] 手势识别主要是通过分析手势具有的各种特征实现对手势的判断,这些特征主要包括颜色特征、文理特征、形状特征、轮廓特征及动态手势具有的运动特征等。目前,手势识别系统主要分为基于数据手套的识别系统和基于视觉的识别系统。基于数据手套的识别系统主要通过传感器获取人手运动的信号,这种方式更有利于获取手势运动信息;然而,这些设备价格非常昂贵而且在使用这些设备的同时往往给用户带来诸多不便。但是,基于视觉的识别系统由于不需要价格昂贵的硬件设备,因此越来越适合人机交互技术。早期的基于视觉的手势识别系统往往需要对手势进行标记然后再进行图像处理;然而,目前基于视觉的手势识别系统主要是利用手势特征如手势颜色、手势形状、手势轮廓或其他深度信息进行识别。基于视觉的手势识别系统可以识别输入视频中的不同手势,并且用这些手势作为计算机的人工输入命令,并且可以建立包括有完整人物及其手势的图库、并且这种手势识别系统由于设备简单,因而使用方便有效。基于视觉的手势识别系统为人机交互提供了一种直观的交流方式。

[0004] 语言伴随性手势作为一种动作语言,是手和手臂运动的统一,需要计算机动画来实现具体动作。人体的外形主要有皮肤和所附着骨骼肌肉运动决定的,因此,在描述手势动作之前必须先建立一个静态几何模型,通过激活具有大量关节的复杂模型,根据决定运动的几个主关节的最终角度可以确定手臂运动,从而创建精确灵活的基本手势动作单元——手势运动基元。根据伴随语言学 and 统计学分析,基于语言连贯表达的手势运动基元连接能够实现手势和有声语言的协同表达,提高运动合成效果和真实感。然而,目前虚拟人并非基于真人图像建立,不够生动真实,并且虚拟人模型固定,无法根据环境和需要进行变换。

[0005] 此外,手势和语言是时序并行输出的,手势运动要素会随着语言的连贯表达而时刻变化。当人们进行口语交流时形成连续不断的语流,同样手势姿态连续表达会生成相应的手势动作,语言和手势之间存在相互的层级关系。作为生成手势动画的基础,如何实现基本的手势动作在语言中的定位成为语言伴随学和虚拟现实技术的研究难点。随着计算机技术的进一步发展,数字媒体技术的出现为手势与语言映射提供了解决方案——音视频数据分析。通过对音视频数据进行分析,将其中的语言和手势进行转写,并标记对应位置,将能

够有效地实现手势与语言的协同表达。

[0006] 然而,现有的方法存在诸多不足源于所处理的主要是非结构化运动模型之间的过渡,运动过渡关键帧只能实现时序性,单纯靠关键帧数值相似度计算进行运动平滑容易造成运动细节丢失。本发明利用具有结构化和实例化的手势运动模型不但消除了运动跨度过大导致运动过渡难以实现问题,还有效地实现了运动细节的增强与保持。

## 发明内容

[0007] 本发明的目的在于克服现有技术的不足,提供一种基于真人图像建立,画面生动真实,并且虚拟人模型灵活多变,能够直观、生动的方式表现手势动作,而且消除了运动跨度过大导致运动过渡的难题并有效地实现了运动细节的增强与保持的基于手势识别的手势运动基元过渡帧定位方法,并且云数据库单元的使用,使得数据备份份数更多,更有保障,利用了云计算端大的存储容量,优越的运算能力和更广的数据网络,功能更加强大,结合远程客户端,可以对数据进行自定义编辑后更新,操作和查询更加方便,可控性更强,更加灵活多变,适应性 更强。

[0008] 本发明提供了一种基于手势识别的手势运动基元过渡帧定位方法,依次包括以下步骤:

[0009] (1) 利用视频或图像采集装置获取一个或多个分辨率为  $N \times M$  的手势图像,对一个或多个手势图像进行预处理,将一个或多个手势图像中的人物图像提取出来,删除除了人物图像部分之外的背景图像;

[0010] (2) 对提取后的人物图像进行去噪处理,消除传输和采集过程中的噪声干扰,之后将去噪后的人物图像按照轮廓、颜色、纹理分割为轮廓图像、颜色图像和纹理图像;

[0011] (3) 对轮廓图像进行修正,进行平滑处理;

[0012] (4) 将颜色图像按照 RGB 图像的 R、G、B 三个分量通道进行分解,分别将三个分量通道中的每一个像素与其邻近的 8 个像素构成  $3 \times 3$  的像素块,按照如下公式分别得到处理后的每一个像素分别对应的 R、G、B 三个分量的值:

$$[0013] \quad P_{ij}(R) = \left( \frac{P_{(i-1)(j-1)} + P_{(i-1)j} + P_{(i-1)(j+1)} + P_{(i+1)(j-1)} + P_{(i+1)j} + P_{(i+1)(j+1)} + P_{(i-1)(j-1)} + P_{(i-1)j} + P_{(i-1)(j+1)} + P_{(i+1)(j-1)} + P_{(i+1)j} + P_{(i+1)(j+1)} + P_{(i-1)(j-1)} + P_{(i-1)j} + P_{(i-1)(j+1)} + P_{(i+1)(j-1)} + P_{(i+1)j} + P_{(i+1)(j+1)}}{3} + P_{ij} \right) / 2$$

$$[0014] \quad P_{ij}(G) = \left( \frac{P_{(i-1)(j-1)} + P_{(i-1)j} + P_{(i-1)(j+1)} + P_{(i+1)(j-1)} + P_{(i+1)j} + P_{(i+1)(j+1)} + P_{(i-1)(j-1)} + P_{(i-1)j} + P_{(i-1)(j+1)} + P_{(i+1)(j-1)} + P_{(i+1)j} + P_{(i+1)(j+1)} + P_{(i-1)(j-1)} + P_{(i-1)j} + P_{(i-1)(j+1)} + P_{(i+1)(j-1)} + P_{(i+1)j} + P_{(i+1)(j+1)}}{3} + P_{ij} \right) / 2$$

$$[0015] \quad P_{ij}(B) = \left( \frac{P_{(i-1)(j-1)} + P_{(i-1)j} + P_{(i-1)(j+1)} + P_{(i+1)(j-1)} + P_{(i+1)j} + P_{(i+1)(j+1)} + P_{(i-1)(j-1)} + P_{(i-1)j} + P_{(i-1)(j+1)} + P_{(i+1)(j-1)} + P_{(i+1)j} + P_{(i+1)(j+1)} + P_{(i-1)(j-1)} + P_{(i-1)j} + P_{(i-1)(j+1)} + P_{(i+1)(j-1)} + P_{(i+1)j} + P_{(i+1)(j+1)}}{3} + P_{ij} \right) / 2$$

[0016]  $P_{ij}(R)$ 、 $P_{ij}(G)$ 、 $P_{ij}(B)$  分别为经过处理后的 R、G、B 三个分量通道中的像素值,  $P_{ij}$  为 R、G、B 三个分量通道中对应第 i 行、第 j 列的像素值,其中  $i = 1, 2, \dots, N$ ,  $j = 1, 2, \dots, M$ ;

[0017] (5) 将经过处理后的 R、G、B 三个分量通道中的像素值进行分配,合成经过处理后的颜色图像;

[0018] (6) 对纹理图像进行修正,进行平滑处理;

[0019] (7) 将经过处理的轮廓图像、颜色图像和纹理图像进行再合成,建立手势图像数据库;

[0020] (8) 基于手势图像数据库中的手势图像,按照身体的生理结构以及各个关节运动时的约束关系,以关键帧的方式建立手势动作的基本单位,并且描述所述基本单位的时序变化信息,具体包括如下步骤:

[0021] 对两个待合成手势运动基元  $\Phi_i$  和  $\Phi_j$ ,假设长度分别为  $m$  和  $n$ ;

[0022] a. 提取手势运动基元  $\Phi_i$  运动恢复时段的关键帧序列  $F1$ ,假设长度为  $len1$ ;

[0023] b. 提取手势运动基元  $\Phi_j$  运动准备时段关键帧序列  $F2$ ,假设长度为  $len2$ ;

[0024] c. 对运动关键帧序列  $F1$ ,从起始关键帧位置 1,开始移动到末尾关键帧位置  $len1$ ,对每一个位置  $k$ ,依次循环计算关键帧序列  $F2$  中第  $len2-i$  个关键帧与运动关键帧序列  $F1$  中第  $k$  关键帧的相似度  $Sim$ ,如果存在相似度小于阈值  $\varepsilon$ ,则终止循环,否则直至循环结束,定位到过渡帧;

[0025] (9) 通过关键帧插值进行平滑连接合成连贯的手势动作,建立与手势图像匹配的手势运动基元模型;

[0026] (10) 将手势运动基元模型和图像数据库中的手势图像对应合成,建立虚拟人数据库;

[0027] (11) 将虚拟人数据库中的数据上传至云数据库单元,所述云数据库单元将接收到的数据存储后作为备份数据,供远程客户端下载;远程客户端下载云数据库单元中存储的数据,进行自定义编辑,将自定义编辑后的数据回传至虚拟人数据库,对虚拟人数据库中的数据进行更新。

## 附图说明

[0028] 图 1 语义手势匹配流程示意图

[0029] 图 2 手势运动基元中各运动阶段中各关节运动特征图

## 具体实施方式

[0030] 下面详细说明本发明的具体实施,有必要在此指出的是,以下实施只是用于本发明的进一步说明,不能理解为对本发明保护范围的限制,该领域技术熟练人员根据上述本发明内容对本发明做出的一些非本质的改进和调整,仍然属于本发明的保护范围。

[0031] 图 1 中自然语言与手势码匹配需要借助于自然语言与手势协同映射关系,通过对音视频数据进行分析,将其中的语言和手势进行转写,并标记对应位置,能够有效地实现手势与语言的协同表达,通过 MacSHAPA 对视频转录获取手势与 自然语言的映射匹配,并通过具有语言连贯表达特性的运动基元进行拼接合成连续手势动画。

[0032] 在交流中伴随语言产生的手势有些是相对固定的,而大多数的手势都具有较大的随机性特点,但是相对而言,由于人的生理结构等约束和限制,人的手势又是相对固定的,因此,研究中将实现语言与手势之间的非线性匹配,即相同的交流话题中可能有不同的手势出现,或者是同一手势运动的幅度不同。

[0033] 由于所关注的焦点是人与虚拟人“语言”交流过程中语言伴随的手势动画合成,因此在转写话语过程中仅限于交流过程中虚拟人一方所说的“言语”,只包括虚拟人在人机对话过程中进行“语言”交流时有意义且含有主谓的言语单位或言语片段,不包括输入的交互语言、其他情感和停顿等信息,针对智能咨询员(虚拟人)在对一款软件进行介绍时的言语

信息和手势转写如表 1。

[0034] 表 1 自然语言与手势运动基元映射

[0035]

号	带转写标注的自然语言	手 势	
		运动基元 码	手势运动基元的语义描述
	[在这将能看到系统的操作界面], /	(G11)	((右手臂平伸指向系统屏幕))
	在界面的[这个位置]/是	(G11)	((右手臂平伸指向系统屏幕))过渡到((伸出手臂双手食指
	[盒子状]的图标, /	, G12)	和拇指在身体前方相对呈现正方形))
	您觉得这款软件界面[设计的怎么样呢? ]	(G13)	((双手臂向前上方伸展并作耸肩状))

[0036] 采用转写规则,其中“/”表示一个言语节,“[]”表示手势协同表达的话语,“G11, G12, G13”表示不同的手势运动基元编码,“(())”表示所做手势的语义描述。其中第二个语句中涉及到两个手势运动基元的连接,如果是基于同一话题的整个交流中,这样的手势连续表达现象比较明显。

[0037] 本发明根据手势识别和手势运动特点,以及运动任务目标原理,提出了一种基于手势识别的手势运动基元过渡帧定位方法,依次包括以下步骤:

[0038] (1) 利用视频或图像采集装置获取一个或多个分辨率为  $N \times M$  的手势图像,对一个或多个手势图像进行预处理,将一个或多个手势图像中的人物图像提取出来,删除除了人物图像部分之外的背景图像;

[0039] (2) 对提取后的人物图像进行去噪处理,消除传输和采集过程中的噪声干扰,之后将去噪后的人物图像按照轮廓、颜色、纹理分割为轮廓图像、颜色图像和纹理图像;

[0040] (3) 对轮廓图像进行修正,进行平滑处理;

[0041] (4) 将颜色图像按照 RGB 图像的 R、G、B 三个分量通道进行分解,分别将三个分量通道中的每一个像素与其邻近的 8 个像素构成  $3 \times 3$  的像素块,按照如下公式分别得到处理后的每一个像素分别对应的 R、G、B 三个分量的值:

$$[0042] \quad P_g(R) = \left( \frac{P_{(i-1)(j-1)} + P_{(i-1)j} + P_{(i-1)(j+1)} + P_{(i)(j-1)} + P_{(i)j} + P_{(i)(j+1)} + P_{(i+1)(j-1)} + P_{(i+1)j} + P_{(i+1)(j+1)}}{3} + P_g \right) / 2$$

$$[0043] \quad P_g(G) = \left( \frac{P_{(i-1)(j-1)} + P_{(i-1)j} + P_{(i-1)(j+1)} + P_{(i)(j-1)} + P_{(i)j} + P_{(i)(j+1)} + P_{(i+1)(j-1)} + P_{(i+1)j} + P_{(i+1)(j+1)}}{3} + P_g \right) / 2$$

$$[0044] \quad P_g(B) = \left( \frac{P_{(i-1)(j-1)} + P_{(i-1)j} + P_{(i-1)(j+1)} + P_{(i)(j-1)} + P_{(i)j} + P_{(i)(j+1)} + P_{(i+1)(j-1)} + P_{(i+1)j} + P_{(i+1)(j+1)}}{3} + P_g \right) / 2$$

[0045]  $P_{ij}(R)$ 、 $P_{ij}(G)$ 、 $P_{ij}(B)$  分别为经过处理后的 R、G、B 三个分量通道中的像素值,  $P_{ij}$  为 R、G、B 三个分量通道中对应第  $i$  行、第  $j$  列的像素值, 其中  $i = 1, 2, \dots, N$ ,  $j = 1, 2, \dots, M$ ;

[0046] (5) 将经过处理后的 R、G、B 三个分量通道中的像素值进行分配, 合成经过处理后的颜色图像;

[0047] (6) 对纹理图像进行修正, 进行平滑处理;

[0048] (7) 将经过处理的轮廓图像、颜色图像和纹理图像进行再合成, 建立手势图像数据库;

[0049] (8) 基于手势图像数据库中的手势图像, 按照身体的生理结构以及各个关节运动时的约束关系, 以关键帧的方式建立手势动作的基本单位, 并且描述所述基本单位的时序变化信息, 具体包括如下步骤:

[0050] 对两个待合成手势运动基元  $\Phi_i$  和  $\Phi_j$ , 假设长度分别为  $m$  和  $n$ ;

[0051] a. 提取手势运动基元  $\Phi_i$  运动恢复时段的关键帧序列 F1, 假设长度为  $len1$ ;

[0052] b. 提取手势运动基元  $\Phi_j$  运动准备时段关键帧序列 F2, 假设长度为  $len2$ ;

[0053] c. 对运动关键帧序列 F1, 从起始关键帧位置 1, 开始移动到末尾关键帧位置  $len1$ , 对每一个位置  $k$ , 依次循环计算关键帧序列 F2 中第  $len2-i$  个关键帧与运动关键帧序列 F1 中第  $k$  关键帧的相似度  $Sim$ , 如果存在相似度小于阈值  $\varepsilon$ , 则终止循环, 否则直至循环结束, 定位到过渡帧;

[0054] (9) 通过关键帧插值进行平滑连接合成连贯的手势动作, 建立与手势图像匹配的手势运动基元模型;

[0055] (10) 将手势运动基元模型和图像数据库中的手势图像对应合成, 建立虚拟人数据库;

[0056] (11) 将虚拟人数据库中的数据上传至云数据库单元, 所述云数据库单元将接收到的数据存储后作为备份数据, 供远程客户端下载; 远程客户端下载云数据库单元中存储的数据, 进行自定义编辑, 将自定义编辑后的数据回传至虚拟人数据库, 对虚拟人数据库中的数据进行更新。

[0057] 所述步骤 (8) 中以关键帧的方式建立手势动作的基本单位, 并且描述所述基本单位的时序变化信息步骤中, 还包括为手势运动基元模型中的关键帧添加 flag 标记位, 将手势运动基元按具有时间次序关键帧标注 “start”, “prepare”, “achieve\_goal”, “release\_goal”, “end”, 并且将每个手势运动基元分解为四个运动时段: 准备阶段、伸出阶段、后伸出阶段、恢复阶段。

[0058] 所述约束关系采用四元数表示关节旋转。

[0059] 对手势关键帧  $G_i$  与  $G_j$  相似度  $Sim(G_i, G_j)$  计算如下:  $Sim(G_i, G_j) = Dist(p_i, p_j) + vDist(v_i, v_j)$ , 其中  $Dist(p_i, p_j)$  表示各个关节角度加权距离,  $Dist(v_i, v_j)$  表示各关节旋转加速度的加权距离, 参数  $v$  表示不同关节位置对速度差异的影响权重。

[0060] 此外, 还包括步骤 (12): 在手势与语言协同表达的手势动画合成过程中, 根据转写规则在手势运动基元全集上存在一个运动基元覆盖路径, 在以上运动基元覆盖路径下实现基于运动基元连接的手势动画合成。

[0061] 优选地, 通过 MacSHAPA 获取手势与自然语言的映射匹配。

[0062] 本发明所述基于手势识别的手势运动基元过渡帧定位方法的插值手势动画合成



可形式表示为：

[0063] 令  $I$  为手势运动库中运动基元全集,假设在手势与语言协同表达的手势动画合成过程中,根据转写规则在  $I$  子集  $U$  ( $U \subset I$ ) 上存在一个运动基元覆盖路径：

$$[0064] \quad S = \{\Phi_1, \Phi_2, \dots, \Phi_M\} \quad (1)$$

[0065] 其中  $\bigcup_{j=1}^M \Phi_j \subseteq U$   $M \geq 1$ ,  $S$  中的运动基元  $\Phi_j$  的次序是固定的。

[0066] 要想在以上运动基元覆盖路径下实现基于运动基元连接的手势动画合成,任意相邻运动基元的  $\Phi_m$  和  $\Phi_{m+1}$  需要且必须满足以下条件：

$$[0067] \quad \text{Sim}(\Phi_m(G_i), \Phi_{m+1}(G_j)) < \varepsilon \quad (2)$$

[0068] 其中,  $\varepsilon$  表示关键帧相似度阈值。即运动基元  $\Phi_m$  和  $\Phi_{m+1}$  在满足该阈值约束的情况下可以通过关键帧插值进行平滑连接合成连贯的手势动作。

[0069] 关键帧相似度计算,由于虚拟人手臂运动位置和运动速度在手势合成前后是紧密相关的,在计算过渡帧相似时充分考虑手臂关节位置与运动速度的影响,对手势关键帧  $G_i$  与  $G_j$  相似度 (距离差异)  $\text{Sim}(G_i, G_j)$  计算如下：

$$[0070] \quad \text{Sim}(G_i, G_j) = \text{Dist}(p_i, p_j) + v \text{Dist}(v_i, v_j) \quad (3)$$

[0071] 其中  $\text{Dist}(p_i, p_j)$  表示各个关节角度加权距离,  $\text{Dist}(v_i, v_j)$  表示各关节旋转加速度的加权距离,参数  $v$  表示不同关节位置对速度差异的影响权重。

[0072] 通过对手势运动基元运动时段划分进行统计分析发现:在手势运动基元的整个运动过程中都包含有手臂 (前臂和小臂) 运动,并且手掌关节 (主要指手指关节) 运动仅在伸出阶段和后伸出阶段出现,而在运动准备阶段和运动恢复阶段不包含手指运动,附图 2 描述了手势运动基元中各运动阶段中各关节运动特征。

[0073] 在进行基于运动基元插值的手势动画合成过程中主要考虑上肢的肩关节、肘关节和腕关节等的旋转角变化幅度对手势动作平滑度影响。由于肩关节和肘关节的旋转幅度大小直接反映在肘关节和腕关节的空间坐标位置的变化幅度。故可利用前后两个关键帧中肘关节和腕关节的欧氏距离的加权和衡量两帧之间关节角度距离：

$$[0074] \quad \text{Dist}(p_{ik}, p_{jk}) = \max(D_{ijk}^{\text{left}}, D_{ijk}^{\text{right}})$$

$$[0075] \quad D_{ijk}^{\text{left}} = \alpha \sqrt{(x_{i0}^{\text{left}} - x_{j0}^{\text{left}})^2 + (y_{i0}^{\text{left}} - y_{j0}^{\text{left}})^2 + (z_{i0}^{\text{left}} - z_{j0}^{\text{left}})^2} +$$

$$\beta \sqrt{(x_{i1}^{\text{left}} - x_{j1}^{\text{left}})^2 + (y_{i1}^{\text{left}} - y_{j1}^{\text{left}})^2 + (z_{i1}^{\text{left}} - z_{j1}^{\text{left}})^2}$$

$$[0076] \quad D_{ijk}^{\text{right}} = \alpha \sqrt{(x_{i0}^{\text{right}} - x_{j0}^{\text{right}})^2 + (y_{i0}^{\text{right}} - y_{j0}^{\text{right}})^2 + (z_{i0}^{\text{right}} - z_{j0}^{\text{right}})^2} +$$

$$\beta \sqrt{(x_{i1}^{\text{right}} - x_{j1}^{\text{right}})^2 + (y_{i1}^{\text{right}} - y_{j1}^{\text{right}})^2 + (z_{i1}^{\text{right}} - z_{j1}^{\text{right}})^2} \quad (4)$$

[0077] 其中,  $D_{ijk}^{\text{left}}, D_{ijk}^{\text{right}}$  分别表示左右手臂关节距离差异,  $(x_{1k}, y_{1k}, z_{1k})$ ,  $l \in \{i, j\}$ ,  $k \in \{0, 1\}$  分别表示肘关节和腕关节的三个空间坐标。 $\alpha, \beta$  表示两关节的权重系数满足以下条件 (根据实验验证  $0.6 < \alpha < 0.8$  时运动效果较好)：

$$[0078] \quad \alpha + \beta = 1$$

$$[0079] \quad \alpha > \beta \quad (5)$$

[0080] 手势图像采集是也本发明方法的关键,一般通过视频或图像采集设备对手势图

形进行采集,从而提取手势视频流或手势图像。手势图像预处理是手势分割与手势分析的前提与基础。一般手势图像预处理主要是为了消除图像在传输和采集过程中所产生的干扰,从而为后续工作打好基础。

[0081] 计算机的色彩显示基本原理与彩色的电视机显示色彩基本原理一样,都是通过 R、G、B 相加混色原理,通过发射不同强度的三种电子束,使屏幕内覆盖的蓝、红、绿的荧光材料通过发光而产生色彩。该色彩表示法称为 RGB 颜色空间表示。在多媒体计算机技术里,RGB 色彩空间表示是最常用的颜色空间。RGB 颜色空间是基于自然界中三种基色光的加法混合原理,将 Red、Green、Blue 三种基色按照从黑色(灰度值为 0)到白色(灰度值为 255)的亮度值在三个颜色通道中分配,从而确定色彩,由于 RGB 颜色空间中的三个分量 R、G、B 有 255 个等级,通过 RGB 三个分量的组合可以产生  $256 \times 256 \times 256$  种颜色,因此,RGB 颜色空间是广泛应用于数字图像处理与存储的颜色空间之一。在 RGB 色彩空间中,任意色光 F 都可以用 R、G、B 三种分量相加混合而成,其中,当 R、G、B 三基色分量最弱(为 0)时,混合为黑色光,当 R、G、B 三基色分量最强(为 255)时,混合为白色光。RGB 颜色空间中三个轴分别对应 R、G、B 三分量,原点对应 R、G、B 值为 0 的点即黑色,与原点对应的顶点为纯白色。过原点的对角线对应从黑色到白色的灰度值变化。而立方体内(包含立方体表面)的点都是  $256 \times 256 \times 256$  种颜色中的一种,用从原点到该点的矢量表示。本发明通过特殊的图像处理方式,使得经过处理后的手势图像,颜色失真更小,颜色更加真实,具体通过将颜色图像按照 RGB 图像的 R、G、B 三个分量通道进行分解,分别将三个分量通道中的每一个像素与其邻近的 8 个像素构成  $3 \times 3$  的像素块,按照如下公式分别得到处理后的每一个像素分别对应的 R、G、B 三个分量的值:

$$[0082] \quad P_{ij}(R) = \left( \frac{P_{(i-1)(j-1)} + P_{(i-1)j} + P_{(i-1)(j+1)} + P_{(i-1)(j+1)} + P_{(i+1)(j-1)} + P_{(i+1)j} + P_{(i+1)(j+1)} + P_{(i+1)(j+1)}}{3} + \frac{P_{(i-1)(j-1)} + P_{(i-1)j} + P_{(i-1)(j+1)} + P_{(i+1)(j-1)} + P_{(i+1)j} + P_{(i+1)(j+1)}}{3} + \frac{P_{(i-1)(j-1)} + P_{(i-1)j} + P_{(i-1)(j+1)} + P_{(i+1)(j-1)} + P_{(i+1)j} + P_{(i+1)(j+1)}}{3} + P_{ij} \right) / 2$$

$$[0083] \quad P_{ij}(G) = \left( \frac{P_{(i-1)(j-1)} + P_{(i-1)j} + P_{(i-1)(j+1)} + P_{(i-1)(j+1)} + P_{(i+1)(j-1)} + P_{(i+1)j} + P_{(i+1)(j+1)} + P_{(i+1)(j+1)}}{3} + \frac{P_{(i-1)(j-1)} + P_{(i-1)j} + P_{(i-1)(j+1)} + P_{(i+1)(j-1)} + P_{(i+1)j} + P_{(i+1)(j+1)}}{3} + \frac{P_{(i-1)(j-1)} + P_{(i-1)j} + P_{(i-1)(j+1)} + P_{(i+1)(j-1)} + P_{(i+1)j} + P_{(i+1)(j+1)}}{3} + P_{ij} \right) / 2$$

$$[0084] \quad P_{ij}(B) = \left( \frac{P_{(i-1)(j-1)} + P_{(i-1)j} + P_{(i-1)(j+1)} + P_{(i-1)(j+1)} + P_{(i+1)(j-1)} + P_{(i+1)j} + P_{(i+1)(j+1)} + P_{(i+1)(j+1)}}{3} + \frac{P_{(i-1)(j-1)} + P_{(i-1)j} + P_{(i-1)(j+1)} + P_{(i+1)(j-1)} + P_{(i+1)j} + P_{(i+1)(j+1)}}{3} + \frac{P_{(i-1)(j-1)} + P_{(i-1)j} + P_{(i-1)(j+1)} + P_{(i+1)(j-1)} + P_{(i+1)j} + P_{(i+1)(j+1)}}{3} + P_{ij} \right) / 2$$

[0085]  $P_{ij}(R)$ 、 $P_{ij}(G)$ 、 $P_{ij}(B)$  分别为经过处理后的 R、G、B 三个分量通道中的像素值,  $P_{ij}$  为 R、G、B 三个分量通道中对应第 i 行、第 j 列的像素值,其中  $i = 1, 2, \dots, N$ ,  $j = 1, 2, \dots, M$ ;

[0086] 过渡关键帧快速定位算法,对运动基元进行标注处理并且手势动作基元都满足起始运动位置约束,能够减少大量的运动基元间比较操作,但是运动基元间过渡关键帧的定位仍然需要计算关键帧之间相似度,选择满足阈值约束的关键帧作为过渡帧。为了在保证运动合成质量的同时,提高计算效率,本发明提出了基于逆向关键帧序列的过渡关键帧快速定位算法。对两个待合成手势运动基元  $\Phi_i$  和  $\Phi_j$ ,假设长度分别为 m 和 n,过渡关键帧快速定位算法如下(多个手势运动基元合成可按照同样的方法依次进行处理):

[0087] (1) 提取手势运动基元  $\Phi_i$  运动恢复时段的关键帧序列 F1,假设长度为 len1;

[0088] (2) 提取手势运动基元  $\Phi_j$  运动准备时段关键帧序列 F2,假设长度为 len2;

[0089] (3) 对运动关键帧序列 F1,从起始关键帧位置 1(关键帧标号)开始移动到末尾关键帧位置 len1。对每一个位置 k,依次循环计算关键帧序列 F2 中第 len2-i 个关键帧与运动关键帧序列 F1 中第 k 关键帧的相似度 Sim,如果存在相似度满足阈值  $\varepsilon$ (公式 2) 则终止

循环,否则直至循环结束,定位到过渡帧。

[0090] 尽管为了说明的目的,已描述了本发明的示例性实施方式,但是本领域的技术人员将理解,不脱离所附权利要求中公开的发明的范围和精神的情况下,可以在形式和细节上进行各种修改、添加和替换等的改变,而所有这些改变都应属于本发明所附权利要求的保护范围,并且本发明要求保护的产品各个部门和方法中的各个步骤,可以以任意组合的形式组合在一起。因此,对本发明中所公开的实施方式的描述并非为了限制本发明的范围,而是用于描述本发明。相应地,本发明的范围不受以上实施方式的限制,而是由权利要求或其等同物进行限定。

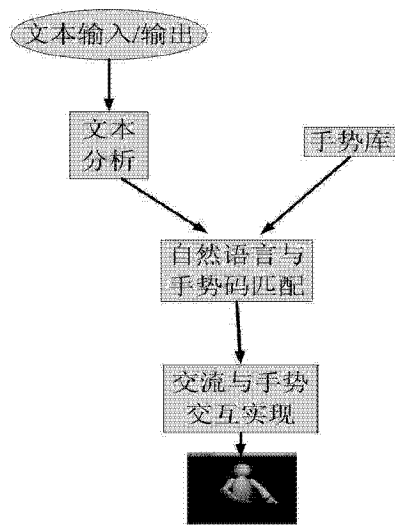


图 1

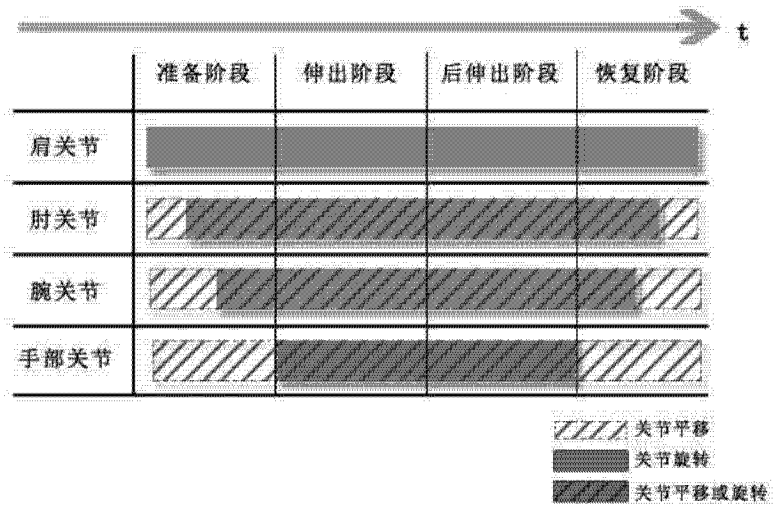


图 2