



(12) 发明专利

(10) 授权公告号 CN 109729263 B

(45) 授权公告日 2021.02.23

(21) 申请号 201811491384.6

H04N 13/275 (2018.01)

(22) 申请日 2018.12.07

(56) 对比文件

(65) 同一申请的已公布的文献号
申请公布号 CN 109729263 A

CN 102256061 A, 2011.11.23

CN 101521740 A, 2009.09.02

CN 105335718 A, 2016.02.17

(43) 申请公布日 2019.05.07

CN 103236082 A, 2013.08.07

(73) 专利权人 苏州中科广视文化科技有限公司
地址 212400 江苏省苏州市太仓市科教新城健雄路20号

CN 106331480 A, 2017.01.11

CN 104159120 A, 2014.11.19

CN 106027852 A, 2016.10.12

(72) 发明人 李兆歆 穆乐文 王兆其

CN 106101535 A, 2016.11.09

CN 107750451 A, 2018.03.02

(74) 专利代理机构 南京苏高专利商标事务所
(普通合伙) 32204

CN 108564554 A, 2018.09.21

US 2012120264 A1, 2012.05.17

代理人 柏尚春

US 2012307085 A1, 2012.12.06

(51) Int. Cl.

CN 106101535 A, 2016.11.09

US 2005093985 A1, 2005.05.05

H04N 5/232 (2006.01)

H04N 5/14 (2006.01)

H04N 13/106 (2018.01)

审查员 张述照

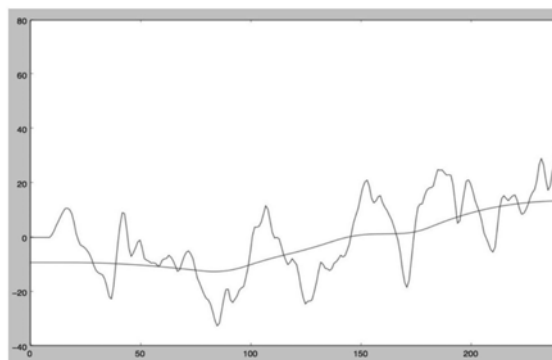
权利要求书1页 说明书4页 附图1页

(54) 发明名称

基于融合运动模型的视频除抖方法

(57) 摘要

本发明公开了一种基于融合运动模型的视频除抖方法,该方法包括以下步骤:(1)在视频的相邻帧之间计算光流,提取特征点并计算匹配结果;(2)根据特征点的匹配结果进行三维重建,恢复相机的三维相机位姿和场景的三维点云;(3)由光流建立基于网格的二维运动模型,描述两帧之间的变形关系;(4)平滑三维运动模型运动估计,求解网格模型的运动补偿;(5)平滑基于网格模型运动轨迹,求解运动补偿;(6)根据求得的运动补偿绘制稳定的图像帧。本方法融合了三维运动模型和二维运动模型,提升了视频除抖算法的视觉效果和鲁棒性,使用优化方法平滑运动轨迹,求解运动补偿,稳定视频内容,优化观看体验。



1. 一种基于融合运动模型的视频除抖方法,其特征在于:该方法包括以下步骤:

一、在视频的相邻帧之间计算光流,提取特征点并计算匹配结果;

二、根据特征点的匹配结果进行三维重建,恢复相机的三维相机位姿和场景的三维点云;

三、由光流建立基于网格的二维运动模型,描述两帧之间变形关系,得到视频内容运动轨迹;

四、平滑三维运动模型运动轨迹,求解基于网格模型的运动补偿,初步稳定步骤三中得到的运动轨迹,所述的平滑三维运动模型运动轨迹通过使用线性规划方法平滑相机位姿的各运动分量,求解相机位姿的运动补偿变换,根据变换与点云深度求解成像平面上的二维运动,根据此运动视为光流重复步骤三建立基于网格的二维运动模型,作为运动补偿稳定步骤三建立的二维运动模型运动轨迹;

五、平滑二维运动模型运动轨迹,求解运动补偿,对步骤四中所述的运动轨迹进行补充稳定;

六、根据求得的运动补偿绘制稳定的图像帧。

2. 根据权利要求1所述的基于融合运动模型的视频除抖方法,其特征在于步骤一中,所述的计算光流是通过对视频的相邻帧提取异构的匹配结果来估计场景运动。

3. 根据权利要求1所述的基于融合运动模型的视频除抖方法,其特征在于步骤二中,特征点匹配用于建立三维运动模型,估计相机在三维空间中的位姿变化。

4. 根据权利要求1所述的基于融合运动模型的视频除抖方法,其特征在于步骤三中,光流用于建立二维运动模型,估计视频相邻帧之间基于网格的变形。

5. 根据权利要求1所述的基于融合运动模型的视频除抖方法,其特征在于步骤五中,所述的平滑二维运动模型运动轨迹是在对基于网格的运动模型进行平滑时,对网格顶点的运动进行子空间分解,分解为基向量运动轨迹,使用线性规划方法平滑各基向量运动轨迹,求得对步骤四的补充运动补偿。

6. 根据权利要求1所述的基于融合运动模型的视频除抖方法,其特征在于步骤六中,最终用于绘制稳定视频帧的运动补偿,由步骤四平滑三维运动模型轨迹求解的运动补偿和由步骤五平滑二维运动模型轨迹求解的运动补偿相加得到。

基于融合运动模型的视频除抖方法

技术领域

[0001] 本发明涉及数字图像处理与计算机视觉领域,具体是一种通过平滑运动轨迹求解对图像帧的运动补偿,重绘制视频帧得到稳定平滑的视觉效果的一种基于融合运动模型的视频除抖方法。

背景技术

[0002] 随着互联网和消费电子产品的发展,视频已经成为记录和分享信息的主要载体,使用移动设备随时随地拍摄的视频内容数量激增,而手持设备拍摄的视频通常无法像专业设备拍摄效果一样有平滑的镜头轨迹,这导致观看体验下降。所以如何解决视频抖动问题成为了研究热点。

[0003] 数字防抖方案通过算法处理视频内容,由视频运动估计和运动平滑两部分组成。运动估计得到视频内容的运动轨迹,运动平滑将抖动的运动轨迹处理为理想的平滑运动轨迹。其中运动估计部分主要分为二维运动模型估计方法和三维运动估计方法。二维运动模型通过特征点匹配、光流法等二维运动估计方法建立运动模型,描述相邻帧之间的几何变形如仿射变换、单应性变换或基于网格的变形。估计过程速度快,效果鲁棒,但是存在场景景深描述能力不强的问题,所估计的运动误差较大。三维运动模型通过运动恢复结构等三维重建方法对视频内容进行三维结构重建,恢复相机位姿和场景深度,对场景和相机运动描述较为准确全面。但三维运动模型估计依赖的三维重建方法存在重建失败的情况,会导致视频除抖方法运行失败。

发明内容

[0004] 发明目的:为了解决现有技术的不足,本发明所述的一种基于融合运动模型的视频除抖方法,该方法结合二维运动模型方法和三维运动模型方法,解决三维重建可能失败时,基于三维运动模型方法的视频稳定算法运动失败的问题,同时利用了三维信息改善了只参考二维运动视频稳定方法的稳定效果。

[0005] 技术方案:为了实现上述目的,本发明所述的一种基于融合运动模型的视频除抖方法,该方法包括以下步骤:

[0006] 一、在视频的相邻帧之间计算光流,提取特征点并计算匹配结果;

[0007] 二、根据特征点的匹配结果进行三维重建,恢复相机的三维相机位姿和场景的三维点云;

[0008] 三、由光流建立基于网格的二维运动模型,描述两帧之间变形关系,得到视频内容运动轨迹;

[0009] 四、平滑三维运动模型运动轨迹,求解基于网格模型的运动补偿,初步稳定步骤三中得到的运动轨迹;

[0010] 五、平滑二维运动模型运动轨迹,求解运动补偿,对步骤四中所述的运动轨迹进行补充稳定;

[0011] 六、根据求得的运动补偿绘制稳定的图像帧。

[0012] 作为本发明的进一步优选,所述的步骤一中,所述的计算光流是通过对视频的相邻帧提取异构的匹配结果来估计场景运动。

[0013] 作为本发明的进一步优选,所述的步骤二中,特征点匹配用于建立三维运动模型,估计相机在三维空间中的位姿变化。

[0014] 作为本发明的进一步优选,所述的步骤三中,光流用于建立二维运动模型,估计视频相邻帧之间基于网格的变形。

[0015] 作为本发明的进一步优选,所述的步骤四中,所述的平滑三维运动模型运动轨迹通过使用线性规划方法平滑相机位姿的各运动分量,求解相机位姿的运动补偿变换,根据变换与点云深度可以求解成像平面上的二维运动,根据此运动视为光流重复步骤三建立基于网格的二维运动模型,作为运动补偿稳定步骤三建立的二维运动模型运动轨迹。

[0016] 作为本发明的进一步优选,所述的步骤五中,所述的平滑二维运动模型运动轨迹是在对基于网格的运动模型进行平滑时,对网格顶点的运动进行子空间分解,分解为基向量运动轨迹,使用线性规划方法平滑各基向量运动轨迹,求得对步骤四的补充运动补偿;

[0017] 作为本发明的进一步优选,所述的步骤六中,最终用于绘制稳定视频帧的运动补偿,由步骤四平滑三维运动模型轨迹求解的运动补偿和由步骤五平滑二维运动模型轨迹求解的运动补偿相加得到。

[0018] 作为本发明的进一步优选,该方法融合了三维运动模型和二维运动模型的运动补偿进行视频除抖。

[0019] 有益效果:本发明所述的一种基于融合运动模型的视频除抖方法,使用三维运动模型和二维运动模型融合的方法建立参数化模型稳定视频相邻帧之中场景的运动,其中使用线性规划优化方法平滑运动轨迹,求解稳定视频内容需要的运动补偿,使原本抖动的视频变为稳定的视频,优化观看体验。

附图说明

[0020] 图1为使用线性规划方法平滑运动分量轨迹示意图;

[0021] 图2为网格运动模型示意图;

[0022] 图3为网格顶点之间角度保护的几何关系示意图。

具体实施方式

[0023] 下面结合附图和具体实施例对本发明的技术方案作进一步的详细说明,但并不局限于此,凡是对本发明技术方案进行修改或者同等替换,而不脱离本发明技术方案的精神和范围,均应涵盖在本发明的保护范围中。

[0024] 实施例

[0025] 基于融合运动模型的视频除抖方法,按以下步骤实现:

[0026] 一、获取相邻视频帧之间的特征点匹配和光流:

[0027] 将图像转为灰度图像,提取SIFT特征点,使用k近邻方法为每个特征点在相邻帧中寻找在描述子欧式空间上最相近的两个特征点。计算这两个特征点与原图特征点描述子距离的差异,如果大于某阈值则认为正确的匹配,保留匹配结果。

[0028] 使用Lucas-Kanade光流算法估计每个像素在视频相邻帧之间的运动情况 $m = (u, v)$, u 为图像 x 轴坐标的移动, v 为图像 y 轴坐标的移动。

[0029] 二、根据特征点的匹配结果进行三维重建, 恢复相机的三维相机位姿和场景的三维点云;

[0030] 使用已有的由运动恢复结构的算法根据特征点的匹配结果求解相机位姿和场景

的稀疏点云。相机位姿为相机外参矩阵 $[R|t] = \begin{bmatrix} r_{1,1} & r_{1,2} & r_{1,3} & t_1 \\ r_{2,1} & r_{2,2} & r_{2,3} & t_2 \\ r_{3,1} & r_{3,2} & r_{3,3} & t_3 \end{bmatrix}$, 任一稀疏点云的三维空间齐

次坐标为 $P = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$, 每个稀疏点云在像平面上投影的坐标为 $C = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$, 相机内参矩阵为

$K = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$, 根据相机模型得到 $C = K[R|t]P$, 根据此关系构造最小二乘优化问题最小化

三维空间点云投影到像平面的投影误差, $\arg \min \sum_{i=1}^n \|C - K[R|t]P\|_2^2$, 其中 $K[R|t]P$ 取非齐次部分。最终求解稀疏点云和相机位姿。

[0031] 三、根据光流建立基于网格的二维运动模型, 描述两帧之间的变形关系:

[0032] 将每个图像帧用同样尺寸的网格等分, 每个网格都存在四个顶点, 对步骤一中估计的光流进行等间隔采样, 采样到的光流 m_k 所在的网格的四个顶点的位置为 V_k , 为了约束光流 m_k 的图像坐标 P_k 与网格顶点坐标 V_k 的关系最小化 $E_d = \sum_k \|w_k^T V_k - (P_k + m_k)\|^2$, w_k^T 根据 P_k 在网格中的位置对四个网格顶点进行坐标的加权, 此约束保证网格模型的运动拟合图像内容的运动, 除此之外, 网格模型还要保护本身的拓扑结构不变, 对任意三个相邻的网格顶点之间建立保角几何约束, 设 $V_1 V_2 V_3$ 为任意三个相邻的网格顶点, 其坐标符合 $V_1 = V_2 + u(V_3 - V_2) + vR_{90}(V_3 - V_2)$ 几何关系, 最小化 $E_s(V_1) = w_s \| |V_1 - (V_2 + u(V_3 - V_2) + vR_{90}(V_3 - V_2))| \|^2$ 保护任意三个相邻网格顶点的夹角关系, 网格的形状就得到了保护。使用最小二乘法求解 $\arg \min \sum_i E_d(V_i) + E_s(V_i)$, 求得网格顶点的运动。

[0033] 四、平滑三维运动模型运动轨迹, 求解基于网格模型的运动补偿:

[0034] 线性规划平滑相机位姿轨迹的各运动分量, x 、 y 、 z 轴的旋转、平移。设某运动分量的数据为 $T = [t_1, \dots, \dots]$, 设线性规划稳定的目标为最小化,

[0035] $e = (e^1, e^i, \dots, e^i = (e_1^i, \dots, e_n^i), c = (w_1, w_2, c)$ 为各阶残差的权重, e 为代表各阶残差的优化变量。设一阶残差 $r_n = t_n - R_n(p_{n-1}, p_n) = r_n + p_{n-1}$ 由此建立问题约束:

$$[0036] \quad \|R_n(p_{n-1}, p_n)\| \leq e_n^1$$

$$[0037] \quad \|R_{n+1}(p_n, p_{n+1}) - R_n(p_{n-1}, p_n)\| \leq e_n^2$$

$$[0038] \quad \|R_{n+2}(p_{n+1}, p_{n+2}) - 2R_{n+1}(p_n, p_{n+1}) + R_n(p_{n-1}, p_n)\| \leq e_n^3$$

$$[0039] \quad 0 \leq e_n^i$$

$$[0040] \quad |p_n| < h$$

[0041] 最终求解每个运动分量的运动补偿 $P = (p_1, \dots, p_n)$

[0042] 任意一帧的各运动分量的运动补偿可以计算相机位姿的变换矩阵 $[R|t]$ ，将在此帧可视范围内的点云 P 带入 $K[R|t]P$ 计算，取非齐次部分得到一组稀疏的二维运动。将其视作光流使用与步骤三相同的方法计算每一帧基于网格的运动补偿。与步骤三求得的基于网格的运动相加，得到由三维运动模型初步稳定的基于网格的运动模型轨迹。

[0043] 五、平滑基于网格模型运动轨迹，求解运动补偿：

[0044] 三维运动模型方法可能会出现失败，还有部分网格运动没有被稳定过。在二维的基于网格的模型上再做第二次运动补偿，提升方法的鲁棒性。基于网格的运动模型在时间维度上需要优化每个顶点的运动轨迹，在空间维度上需要保护顶点之间的结构不变不出现扭曲。这样导致随着视频帧数量的增加，优化项迅速增加，使优化变得困难。本发明使用矩阵分解，首先建立基于网格的运动模型的运动轨迹矩阵。

$$[0045] \quad M_{2n \times t} = \begin{bmatrix} x_1^1 & x_2^1 & \dots & x_t^1 \\ y_1^1 & y_2^1 & \dots & y_t^1 \\ \dots & \dots & \dots & \dots \\ x_1^n & x_2^n & \dots & x_t^n \\ y_1^n & y_2^n & \dots & y_t^n \end{bmatrix}$$

[0046] 矩阵描述了 n 个网格顶点的 x 轴 y 轴的运动分量，通过 SVD 分解。分解为 $M_{2n \times t} = C_{2n \times r} E_{r \times t}$ ，其中 $E_{r \times t}$ 为运动在子空间的分解轨迹，其中 r 为 9，将 $E_{r \times t}$ 中每行视作运动分量轨迹用步骤四中方法平滑。得到 $\hat{E}_{r \times t}$ 和平滑后的网格顶点运动轨迹 $\hat{M}_{2n \times t} = C_{2n \times r} \hat{E}_{r \times t}$ ，求得第二次运动补偿结果 $\hat{M}_{2n \times t} - M_{2n \times t}$ ，与三维运动模型方法求解的运动补偿相加，得到总的运动补偿。

[0047] 六、根据求得的运动补偿绘制稳定的图像帧；

[0048] 根据初始的网格顶点坐标和运动补偿可以求得新的网格顶点坐标，对网格顶点坐标加权可得到网格内每个像素的新坐标 $w_k^T V_k$ ，将像素值复制到新的坐标即完成稳定帧的绘制。

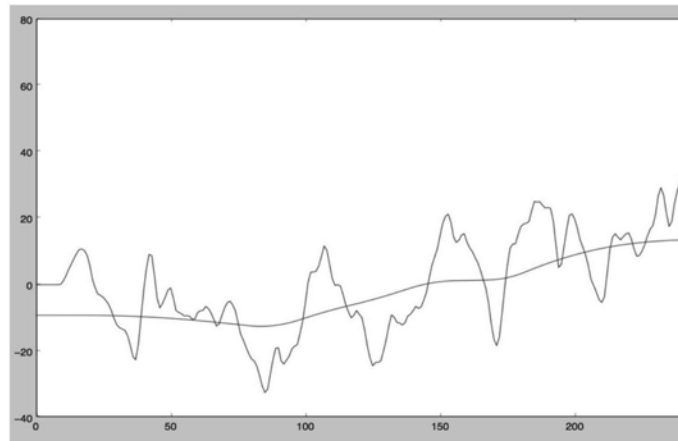


图1

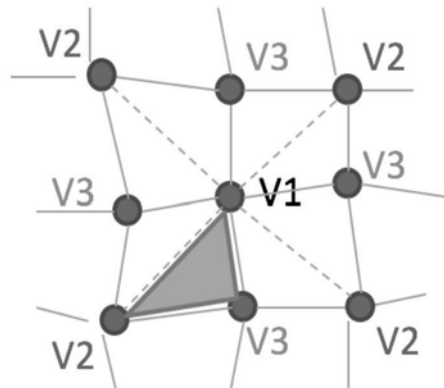


图2

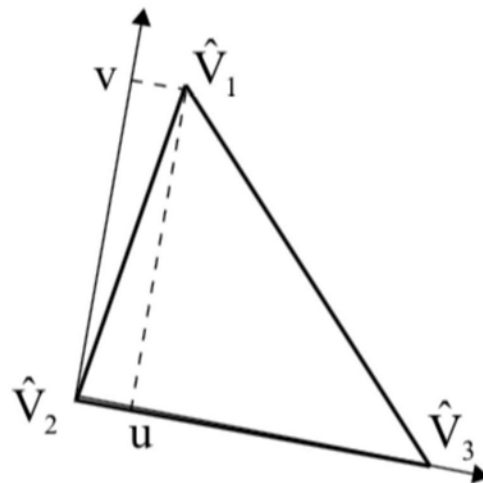


图3