



一种鱼眼视频全景拼接中的亮度补偿算法

黎吉国^{1,4}, 王悦², 张新峰³, 马思伟^{2*}

1. 中国科学院计算技术研究所, 北京 100190, 中国

2. 北京大学信息科学技术学院, 北京 100871, 中国

3. School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798, Singapore

4. 中国科学院大学计算机与控制学院, 北京 100049, 中国

* 通信作者. E-mail: swma@pku.edu.cn

收稿日期: 2018-01-07; 接受日期: 2018-01-31; 网络出版日期: 2018-03-16

国家高技术研究发展计划 (批准号: 2015AA015903) 资助项目

摘要 随着鱼眼相机逐渐成为全景视频的主要采集设备, 面向鱼眼图像的高效全景拼接技术也得到越来越多的关注. 为了解决镜头间亮度差异造成的拼接视频质量下降的问题, 本文提出了一种基于非均匀采样直方图的亮度补偿算法. 该算法根据全景图像在不同纬度空间采样率变化的特点对相邻图像的重叠区域进行非均匀采样, 构成统计直方图, 通过直方图匹配实现图像间的亮度补偿. 此外, 考虑到图像亮度分布对拼接图像质量的影响, 本文根据重叠区域图像亮度直方图提出了一种新的自适应基准图像选择方法, 基于该方法得到的基准图像可以有效提高全景视频的质量, 该方法充分考虑了鱼眼镜头的畸变特性, 不要求拼接的图像之间精确配准, 能够适应不同的光照条件. 实验表明, 本文提出的算法在保证全景视频质量的同时, 有效地降低了亮度补偿的复杂度, 从而可以应用于基于鱼眼镜头的全景实时视频拼接系统.

关键词 图像处理, 虚拟现实, 直方图匹配, 鱼眼镜头, 视频拼接

1 引言

全景视频直播因其 360° 的视角和沉浸式的体验, 而在近几年快速发展. 国内外的各大视频网站都已经逐步支持全景视频的观看, 越来越多的体育赛事和娱乐节目也开始使用全景直播. 由于鱼眼镜头的宽视角可以大大降低全景视频拍摄所需的相机个数, 因此, 基于鱼眼镜头的全景相机逐渐成为了全景视频的主要获取设备, 如三星 Gear 360, 诺基亚 OZO 等. 同时, 基于鱼眼镜头的实时全景视频拼接技术也成为了提高全景视频服务质量的关键技术. 鱼眼镜头靠人为引入桶形畸变获得广角成像的大视野, 图像除了画面中心处物体形状不变外, 其余部分均有一定的扭曲, 如图 1(b) 所示. 为了得到高质量的全景视频, 一般需要拼接四路以上的鱼眼视频, 常见的基于鱼眼 GoPro 的全景相机架如图 1(a)

引用格式: 黎吉国, 王悦, 张新峰, 等. 一种鱼眼视频全景拼接中的亮度补偿算法. 中国科学: 信息科学, 2018, 48: 261–273, doi: 10.1360/N112017-00243

Li J G, Wang Y, Zhang X F, et al. A luminance compensation method for fisheye video panorama stitching (in Chinese). Sci Sin Inform, 2018, 48: 261–273, doi: 10.1360/N112017-00243

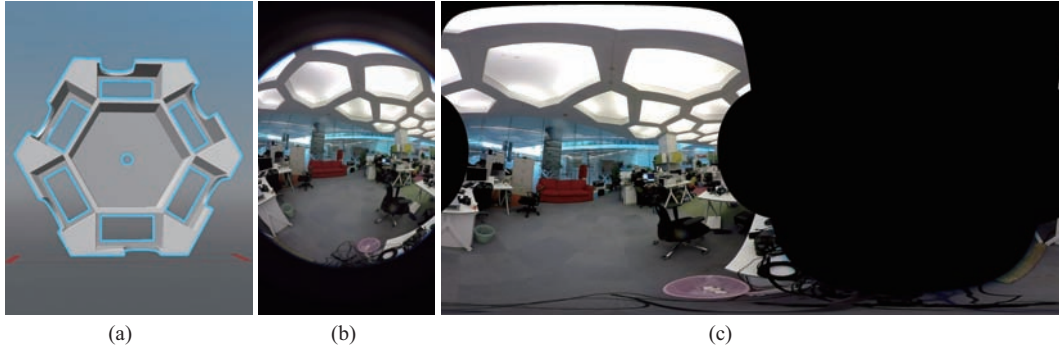


图 1 (网络版彩图) 典型的基于鱼眼相机的视频采集与拼接

Figure 1 (Color online) Fisheye-based video caption and stitching system. (a) Frame for 6 GoPro cameras; (b) fisheye image; (c) equirectangular projection

所示. 同时, 由于鱼眼图像的视角 (field of view, FOV) 往往达到甚至超过 180° , 使得它不能展开在平面上, 通常借助等距圆柱投影 (equirectangular projection, ERP) 将展开的鱼眼图片投影在平面上 (如图 1(c) 所示). 鱼眼图像全景视频拼接的目标是将多个鱼眼图像, 根据其重叠区域内容将其拼接在一起构成全景视频图像. 近期, 针对鱼眼全景相机的高效视频拼接研究得到越来越多的关注^[1,2], 为了满足实时性的要求, 研究人员尝试了使用 GPU^[3] 或者 FPGA^[4] 并行加速的视频拼接方案. 这些工作针对鱼眼相机的标定, 鱼眼图像的对齐, 甚至渐晕的处理, 都提出了相应的解决方案. 但是由于拍摄角度问题, 不同相机图像间不可避免的会存在亮度差异, 如果不进行有效处理, 这种差异会显著降低合成视频的主观质量. 然而, 目前的研究对于不同鱼眼相机图像之间的亮度补偿却鲜有涉及.

因此, 本文提出了一种新的面向鱼眼视频拼接的高效亮度补偿方法. 首先, 利用鱼眼镜头变化的空间采样率, 提出了利用非均匀采样构建相邻图像重叠区域直方图, 通过直方图调整实现拼接图像的亮度补偿的方法. 该方法可以在保持拼接质量的条件下有效地减少均匀采样直方图方法的冗余计算. 此外, 为了提高全景图像质量, 在亮度补偿中, 提出了基于重叠区域归一化直方图的自适应基准图像决策方法. 上述两个步骤都是通过非均匀采样的直方图实现, 从而有效地降低了算法复杂度. 实验表明, 本文提出的亮度补偿算法在复杂度和拼接质量中取得了很好的折中, 可以实现高质量的全景视频实时拼接.

本文组织结构如下: 第 1 部分介绍了亮度补偿的相关工作和鱼眼镜头的特点, 第 2 部分介绍亮度补偿和鱼眼图像的相关工作, 第 3 部分依次介绍了提出的基于非均匀采样直方图的亮度补偿和自适应基准图像选择算法, 第 4 部分通过实验分析论证了算法的有效性, 第 5 部分对本文进行了总结.

2 相关工作

2.1 亮度补偿

在视频/图像拼接中, 亮度补偿是利用待拼接的两幅图像的重叠部分, 得到两幅图像之间的亮度关系 $I_i = f(I_{i+1})$, 然后对其中一幅做亮度变换 $I'_i = f(I_i)$, 使得最终拼接的全景图像具有近似接近的曝光亮度, 在重叠区域平滑自然过度. 其中 f 是亮度变换函数, I_i 和 I_{i+1} 表示两幅相邻的图像. 根据亮度变换函数的形式, 可以将亮度补偿分为线性补偿^[5~8] 和非线性补偿^[9,10]. 线性补偿中亮度变换函

数是一个线性变换或者仿射变换, 可表示为

$$\begin{aligned} y &= Mx, \\ y &= (r_2, g_2, b_2)^T, \\ x &= (r_1, g_1, b_1, 1)^T, \end{aligned} \quad (1)$$

其中 r, g, b 分别表示图像的 3 个通道的像素值. 此时 M 是一个 3×4 的矩阵, 由两幅图像重叠部分对应的点或者块来估计. 线性变换具有明确的数学形式, 但是在实际的图像中, 亮度差异不一定满足线性关系, 使得线性变换的效果有限. 而且线性变换模型参数的求解需要使用对齐之后的像素对, 使得其对噪声比较敏感. 非线性变换则没有确定的数学形式, 往往使用一个查找表来表示亮度变换函数. 其中比较典型的的就是基于直方图匹配的亮度补偿. 它在图像拼接中的应用可简单描述如下.

图像 I_i 与 I_{i+1} 的重叠部分为 $R_{i,i+1}$, 图像 I_{i+1} 与 I_i 的重叠部分为 $R_{i+1,i}$, 则直方图匹配的过程如下:

- (1) 由式 (2) 得到 $R_{i,i+1}$ 和 $R_{i+1,i}$ 的统计直方图 $h_{R_{i,i+1}}$ 和 $h_{R_{i+1,i}}$,

$$h_r[v] = \sum_{m,n \in R} \delta(v, y(m, n)), \quad (2)$$

其中 R 表示一个区域, δ 是 Kronecker 函数, y 表示亮度分量, (m, n) 表示图像上点的二维坐标, v 表示离散的亮度值, 当亮度位深是 8 bit 时, v 的取值是 0 到 255.

- (2) 通过累计直方图, 计算从 $h_{R_{i,i+1}}$ 到 $h_{R_{i+1,i}}$ 的直方图匹配函数 T :

$$h_{R_{i,i+1}} = T(h_{R_{i+1,i}}). \quad (3)$$

- (3) 对 I_{i+1} 中每个像素的亮度通道应用变换 T :

$$I'_{i+1} = T(I_{i+1}). \quad (4)$$

该方法实现简单, 鲁棒性好, 但是由于鱼眼镜头的非线性变换特性, 使得即使相同曝光条件下, 相邻图像的重叠区域也会呈现不同的直方图分布, 所以该方法不能直接应用在鱼眼图像的拼接中.

根据估计亮度变换函数的点的采样密度, 亮度补偿又可以分为密集采样^[5,9]和稀疏采样^[8,10]. 其中密集采样直接使用有效图像区域 (整幅图像, 重叠区域或者重叠区域中选择一个区域) 的全部像素来进行亮度变换函数的求解. 它利用统计来增强算法的鲁棒性, 往往不要求拼接的图像精确对齐, 但是会带来计算的冗余. 稀疏采样对有效区域进行降采样或者提取特征点^[11], 然后利用较少的点来估计亮度变换函数, 在一定程度上降低了计算量. 但是降采样的方式也损失了一定的鲁棒性; 特征点方式在特征点寻找和匹配上也会引入大量的计算, 使它并不适合对速度有较高要求的场合.

2.2 鱼眼图像

鱼镜头是一种近似鱼眼变换 (fish-eye transform, FET)^[11] 的超广角镜头. 为了达到甚至超过 180° 的视角, 鱼镜头拍摄的图像带有强烈的畸变 (如图 1(b) 所示), 使得同一区域在两幅图像中呈现不同的形状, 即使他们处于相同的曝光条件, 也会呈现不同的直方图分布. 如图 2(c) 的灰色重叠区域具有不同统计直方图分布, 如图 2(d) 所示.

由于鱼镜头的超广视角, 矫正后的图像需要借助全景图像投影变换^[12,13]才能呈现在一个平面上, 最常用的是等距圆柱投影 (如图 2(b) 所示). 等距圆柱投影将一个球面投影在一个长宽比为 2 : 1

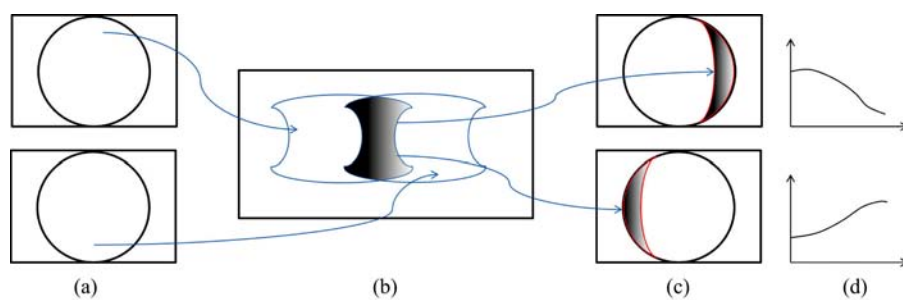


图 2 (网络版彩图) 等距圆柱投影前后的重叠部分

Figure 2 (Color online) Overlap region before and after equirectangular projection. (a) Origin image; (b) equirectangular projection result and the overlap region; (c) overlap region in the origin image; (d) histogram of the two regions

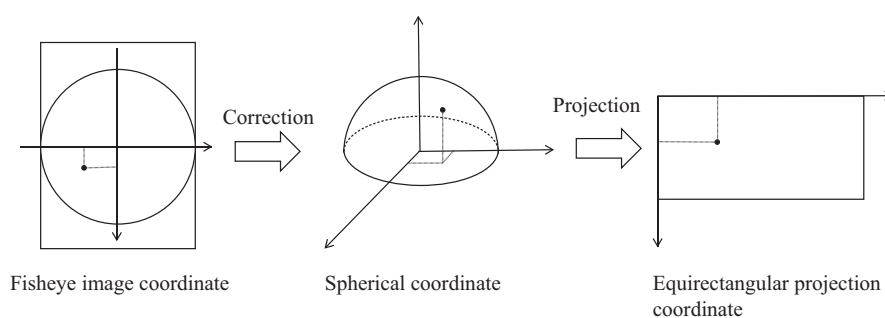


图 3 鱼眼图像预处理

Figure 3 Preprocessing of the fisheye image

的矩形平面上, 球面上的经线和纬线对应于平面上的竖直线和水平线. 等距圆柱投影在全景视频中应用广泛, 但是它在顶部和底部过采样十分严重, 使得针对正常图像的图像处理算法在这里会有大量的计算冗余.

3 算法介绍

针对鱼眼图像的特性, 结合传统亮度补偿的思想, 设计出一种适合于鱼镜头全景视频拼接的亮度补偿算法. 首先对鱼眼图像进行畸变矫正, 并展开投影到平面上, 然后对重叠部分进行非均匀采样, 最后自适应地选择一幅图像作为基准图像, 使用非均匀采样得到的点依次进行直方图匹配, 达到亮度补偿的目的. 算法假设已经得到了拼接的各项参数, 包括相机的内参、外参和鱼镜头的畸变参数. 这样便可以得到图像的重叠部分. 算法的输入是原始图像和拼接参数, 输出是经过亮度补偿的图像, 用于全景图的生成.

3.1 预处理

鱼眼图像预处理主要包括鱼眼图像展开 (warp)^[14]、畸变矫正、等距圆柱投影和计算重叠部分. 如图 3 所示, 先根据鱼眼变换将鱼眼图像展开到一个球面上, 同时使用畸变参数对鱼眼图像进行畸变矫正, 然后使用等距圆柱投影将图像从球面投影到平面上, 最后根据相机之间的相对位置关系便可以得到投影图像之间的重叠区域. 这些步骤都是视频拼接的必要步骤. 经过这些步骤后, 可以得到每一幅

输入图在平面上的展开图和每幅图像与相邻图像的重叠区域 (如图 2(b) 所示), 以便于后续基于重叠区域进行的计算.

3.2 基于鱼眼展开图的直方图统计

传统的基于直方图的亮度补偿都是对原始图像进行直方图统计, 但是对于鱼眼图像, 该方案是不适用的. 这是因为理想情况下, 如果不考虑成像畸变以及不同视角物体形状的变化, 在相同的曝光条件下, 有

$$\text{hist}(P_1(R)) = \text{hist}(P_2(R)), \quad (5)$$

其中 R 是空间中一特定区域, P_1 和 P_2 是小孔成像过程, hist 是直方图函数. 所以曝光条件一致和直方图一致在理论上是等效的.

然而鱼眼图像经过了一个鱼眼变换^[11], 鱼眼变换是一个非线性变换, 变换的结果与 R 在鱼眼图像中的位置有关, 所以会造成即使在相同的曝光条件下, 也会有

$$\text{hist}(\text{FET}(P_1(R), p_1)) \neq \text{hist}(\text{FET}(P_2(R), p_2)), \quad (6)$$

其中 FET 表示鱼眼变换, p_1 和 p_2 表示区域在两幅图像中的位置, 通常情况下 $p_1 \neq p_2$. 即在鱼眼图像中, 曝光条件一致与直方图一致在理论上并不等效. 所以不能直接使用原始图像的重叠部分计算直方图, 而应该使用展开后的图像. 这里本文使用等距圆柱投影得到的展开图像如图 2 所示, 对其重叠部分进行直方图统计.

3.3 非均匀采样

直方图统计过程中, 将面临如何对重叠区域采样的问题. 密集采样虽然可以提高算法的鲁棒性, 但是引入了大量的冗余计算. 降采样则会损失一定的精度和鲁棒性. 为了有效降低算法计算量的同时保证拼接图像质量, 本文提出了非均匀采样的直方图统计方法.

针对等距圆柱投影在顶部和底部过采样的问题, Sun 等^[15] 提出一种带权重的峰值信噪比 (weighted to spherically uniform peak signal-to-noise ratio, WS-PSNR), 对不同位置的均方误差加一个权重因子, 来更好的估计 360° 视频的质量, 受此启发, 本文对鱼眼全景展开图的不同位置使用不同的采样密度, 在降低计算量的同时最大限度地保证拼接图像质量.

如图 4(a) 所示, 参考 WS-PSNR 中对不同位置的像素使用不同的权重和采样密度 (如图 4(b) 所示), 以此来降低计算量. 这里本文使用横向采样间隔来控制不同位置的采样密度. 根据图像上某一点的 WS-PSNR 尺度因子, 得到横向采样间隔

$$\text{stripe} = \text{MIN} + (1 - \text{scale})(\text{MAX} - \text{MIN}), \quad (7)$$

$$\text{scale} = \cos \phi, \quad (8)$$

其中 stripe 代表一行中相邻两个采样点的采样间隔, MAX 和 MIN 表示采样间隔的最大最小值, 缺省分别为 256 和 1, scale 表示 WS-PSNR 中的尺度因子, 变化范围为 $[0, 1]$, ϕ 表示球面维度.

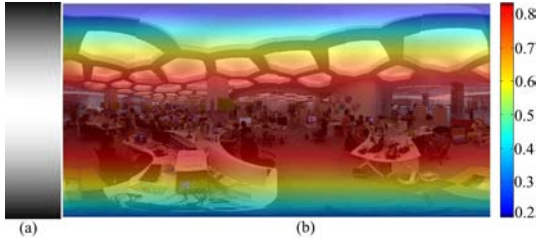


图 4 (网络版彩图) (a) WS-PSNR 的权重因子分布图; (b) 一幅等距圆柱投影全景图上的采样密度分布

Figure 4 (Color online) (a) Weight distribution map of WS-PSNR; (b) density distribution of an equirectangular projection panorama image

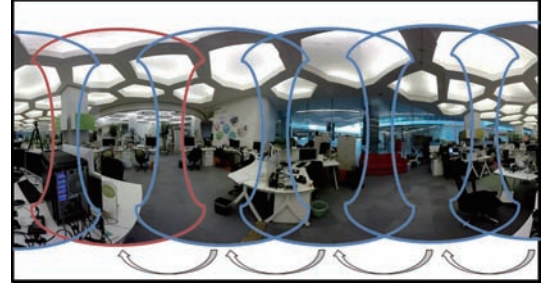


图 5 (网络版彩图) 利用相邻图像的重叠部分的直方图匹配依次进行亮度补偿, 左侧第一个箭头指向的红色边框图像是基准图像

Figure 5 (Color online) Luminance compensation by histogram matching with overlapping parts, where the image with red border is the reference image

3.4 自适应基准图像决策

直方图匹配是一个简单而有效的亮度补偿方法^[16]. 从一幅图像开始, 依次利用重叠部分进行直方图匹配即可实现多幅图像的亮度补偿. 我们把第一次匹配的目标图像称为整个匹配过程的基准图像(如图 5 所示). 由于计算复杂度的限制, 匹配只进行一遍, 所以基准图像就决定了所有图像的整体亮度, 而如何选择基准图像就成了提高全景视频图像质量的关键因素.

根据最大熵原理, 在没有任何先验的情况下, 一个分布的期望应该是均匀分布. 直观上, 图像的亮度集中分布于某一区域会导致图像的对比度变差, 影响图像的主观质量. 所以, 使用图像重叠部分直方图与均匀分布的距离来度量一个图像适合作为基准图像的程度. 这里用得分来表示

$$s[i] = \frac{1}{\frac{N(R_{i,j})}{N(R_{i,j})+N(R_{i,k})}D(H(R_{i,j})) + \frac{N(R_{j,k})}{N(R_{i,j})+N(R_{i,k})}D(H(R_{i,k})))}, \quad (9)$$

$$D(H) = \sum_{i=0}^{255} |H(i) - U(i)|, \quad (10)$$

其中 $s[i]$ 表示第 i 幅图像的得分, $R_{i,j}$ 表示第 i 幅图像和第 j 幅图像的重叠区域. 一般有 $j = i - 1$ 以及 $k = i + 1$, 分别用于表示与第 i 幅图像相邻的两幅图像. H 表示归一化直方图函数, N 表示一个区域采样点的个数, D 表示一个分布与均匀分布之间的距离, U 表示均匀分布. 式 (9) 中分母项表示 $D(H(R_{i,j}))$ 和 $D(H(R_{i,k}))$ 按照重叠区域点的个数进行加权, 取倒数是为了使得 $s[i]$ 的大小与改图作为基准图像的合适程度一致. 选择得分最高的图像作为基准图像, 可以避免出现过暗或者过亮的输出图像, 从而提高算法的鲁棒性, 保证拼接输出视频的质量.

4 实验结果

本文使用由 6 个 GoPro 搭建的实时全景视频拼接系统来测试本文提出的亮度补偿算法的性能. 测试系统的主要模块如图 6 所示. 其中, 用于视频采集的 GoPro 输出分辨率为 1920×1080 , 帧率为 30 fps, 每个镜头的视场角 (FOV) 大约为 195° . 六路来自 GoPro 的视频流通过两个采集卡 (每个采集卡最多支持四路高清视频流的采集) 输入主机, 用于视频拼接. 亮度补偿部分使用 CPU 实现, 实时拼

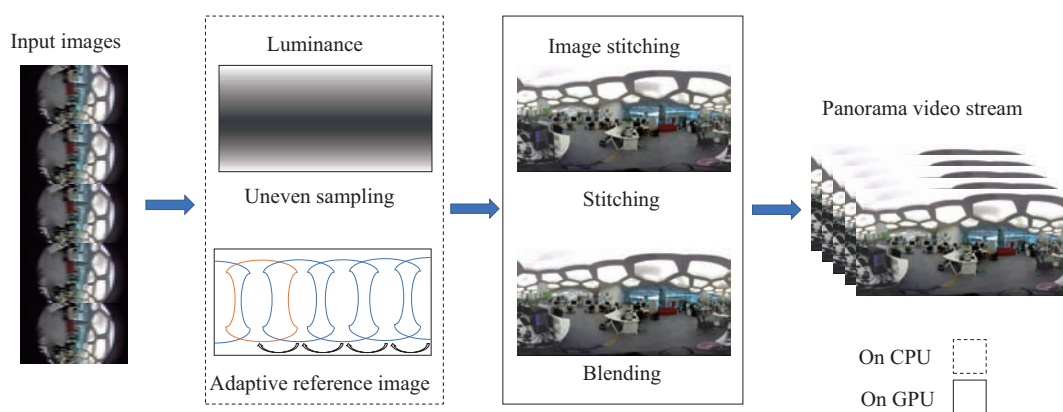


图 6 (网络版彩图) 实时全景视频拼接系统

Figure 6 (Color online) Real-time panorama video stitching system

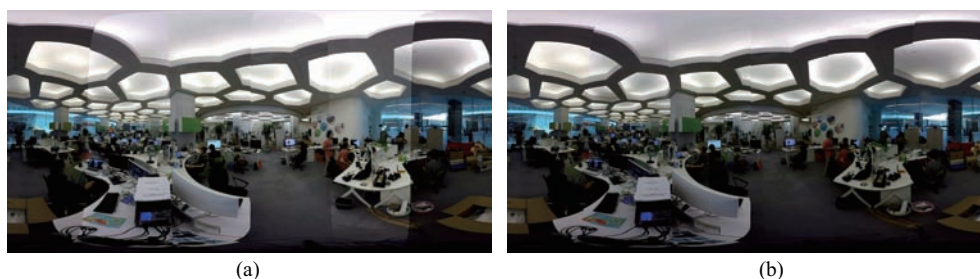


图 7 (网络版彩图) 两种直方图统计方法的亮度补偿结果

Figure 7 (Color online) The luminance compensation results based on histogram statistics with (a) original images, (b) corrected images

接部分使用 OpenCL 进行 GPU 并行加速, 输出的全景视频分辨率为 3840×1920 . 主机的配置为 Intel I7 6700K, 8 G 内存.

亮度补偿的目的是得到更加自然的拼接输出, 所以本文使用拼接模块输出图像的质量来评价亮度补偿的性能. 同时单独测量亮度补偿模块所消耗的时间来评测其效率.

4.1 基于鱼眼展开图的直方图统计

由 2.1 小节分析可知, 直接基于原始图像的直方图统计进行亮度补偿^[16]并不能真实反映场景中的亮度分布, 使用基于鱼眼展开图的直方图统计才能得到与场景亮度分布一致的统计结果, 从而可以通过直方图匹配达到亮度补偿的效果. 这里我们对比两者的补偿效果, 来验证前面的分析结果. 为避免基准图像对实验结果的影响, 实验中使用相同的基准图像.

由图 7 可以看出, 基于原始图像直方图统计的亮度补偿并不能很好地校正不同图像之间的亮度差异 (如图 7(a) 所示), 相邻图像依然有明显的亮度差异, 而基于鱼眼展开图的直方图统计的亮度补偿效果则非常明显, 不同图像的亮度已经基本一致 (如图 7(b) 所示). 实验结果很好地验证了 2.1 小节的理论分析.

表 1 不同的最小最大采样间隔对输出全景图像质量和亮度补偿效率的影响

Table 1 Effectiveness and efficiency of luminance compensation with different min/max sampling intervals

Index	(min, max)	Time (ms)	PSNR	WS-PSNR	Index	(min, max)	Time (ms)	PSNR	WS-PSNR
1	(1, 1024)	9.08	31.4385	31.5229	9	(2, 256)	10.05	34.2231	34.1798
2	(1, 512)	9.93	31.9630	32.0129	10	(4, 256)	9.11	34.4783	34.4546
3	(1, 256)	11.57	32.8647	32.8514	11	(8, 256)	8.24	35.7685	35.7989
4	(1, 128)	13.30	33.9658	33.8869	12	(16, 256)	7.51	38.0210	37.9278
5	(1, 64)	14.99	34.5430	34.4906	13	(32, 256)	6.97	37.1566	37.1456
6	(1, 32)	17.54	35.7429	35.6996	14	(64, 256)	6.58	34.8895	35.1714
7	(1, 16)	20.52	37.0278	36.8767	15	(128, 256)	6.38	34.4281	34.0717
8	(1, 8)	24.31	39.6580	39.5388					

4.2 非均匀采样

可以确定的是, 降采样之后, 一定会带来拼接图像质量的下降, 亮度补偿效率的提升. 这里主要通过实验来研究两者之间的关系. 本文使用密集采样得到的全景图作为源图像, 使用非均匀采样得到的全景图作为受损图像, 使用 PSNR 和 WS-PSNR 来度量拼接输出图像的质量, 统计多次拼接过程中亮度补偿的时间, 来度量其效率, 并考察不同的最大最小采样间隔对拼接图像质量和亮度补偿效率的影响. 为避免基准图像不同对实验结果造成影响, 实验中都选择相同的图像作为基准图像. 实验结果如表 1 所示, 其中 (min, max) 表示最小和最大的采样间隔, 密集采样的时间消耗为 29.63 ms, 所有时间测量都为多次测量后的平均值. 通过表 1 可得到如下结论:

(1) 非均匀采样可以大大减少亮度补偿的计算量, 同时只有少量的质量损失, 即使最大采样间隔设置为 1024, PSNR 依然高达 31.4385 dB.

(2) 由表 1 的 1~8 可以看出, 在固定最小采样间隔的情况下, 最大采样间隔越大, 模块所消耗的时间越少, 同时输出图像的质量越差.

(3) 由表 1 的 9~15 可以看出, 在固定最大采样间隔的情况下, 最小采样间隔与效率正相关, 但是图像质量 (PSNR) 的变化并不是单调的, 随着最小采样间隔的增大, PSNR 先增加后减少. 采样点减少了, 而图像质量却变好了, 说明本文使用的非均匀采样方案并不是最优的, 还有改进的空间.

4.3 自适应基准图像决策

自适应基准图像决策可以根据环境光照自动选择合适的基准图像, 以得到最自然的拼接输出, 能够大大提高亮度补偿算法的鲁棒性. 由于拼接输出图像并没有可以参考计算 PSNR 的源图像, 所以这里主要测试 (1) 选择不同得分的图像作为基准图像对输出图像质量的影响; (2) 使用自适应基准图像的计算消耗. 为避免采样间隔对实验结果造成影响, 实验中采用相同的采样间隔, 最大值为 256, 最小值为 1. 实验结果如图 8, 9 和表 2 所示. 其中在图 8 中 img0, img1, img5 的得分分别为 1.5063, 1.2881, 0.8527. 由图 8, 9 和表 2 可以看出:

(1) 由图 8 可以看出, 直方图分布较“均匀”的图像, 如 img0, 有较高的得分, 而直方图分布集中于一个区间的图像, 如 img5, 有较低的得分, 所以式 (9) 可以比较好地衡量图像直方图分布的“均匀”程度, 从而衡量这幅图像是否适合作为基准图像.

(2) 由图 9 可以看出, 基准图像的选择对拼接输出图像的质量有很大的影响. 选择一个得分较低的基准图像会造成一些区域出现不自然的现象 (如图 9(a) 所示), 而选择一个得分较高的基准图像则

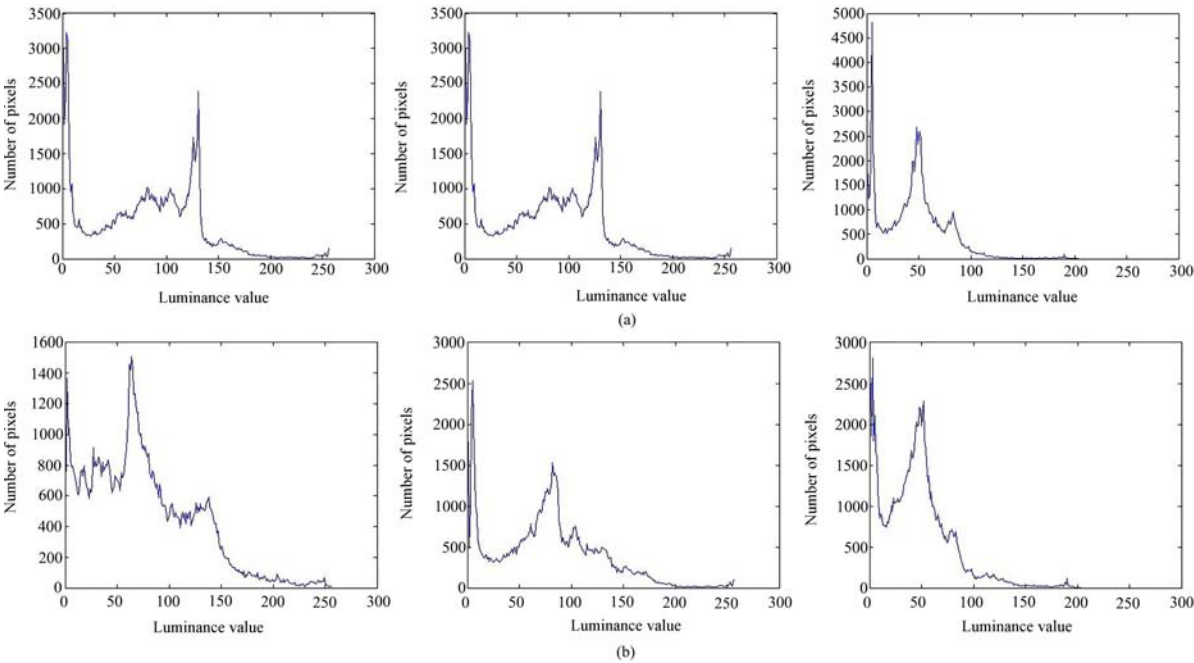


图 8 三幅输入图像的直方图统计结果, 从左到右分别为 `img0`, `img1`, `img5`

Figure 8 Histograms of three input images. From left to right are `img0`, `img1`, `img5`. (a) Histogram of overlap region with previous image; (b) histogram of overlap region with next image

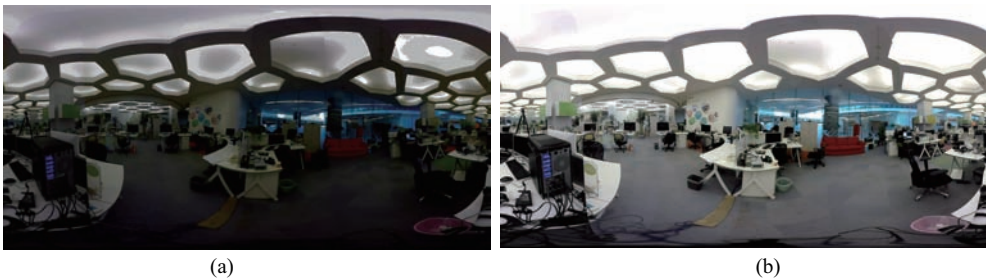


图 9 (网络版彩图) 不同基准图像对拼接输出结果的影响

Figure 9 (Color online) Different results of different reference images. (a) Selecting `img5` (score=0.8527) as reference image; (b) Selecting `img0` (score=1.5063) as reference image

表 2 时间复杂度对比实验 (100 帧)

Table 2 Experiment results for efficiency (100 frames)

Method	Time (ms)
[5], YUV	2173
Proposal method (without adaptive reference image)	1115
Proposal method (with adaptive reference image)	1143

能够得到更加自然的拼接输出 (如图 9(b) 所示).

(3) 由表 2 可以看出, 自适应基准图像方法只引入了很小的计算量, 非常适合使用在对速度有较高要求的实时视频拼接应用中.

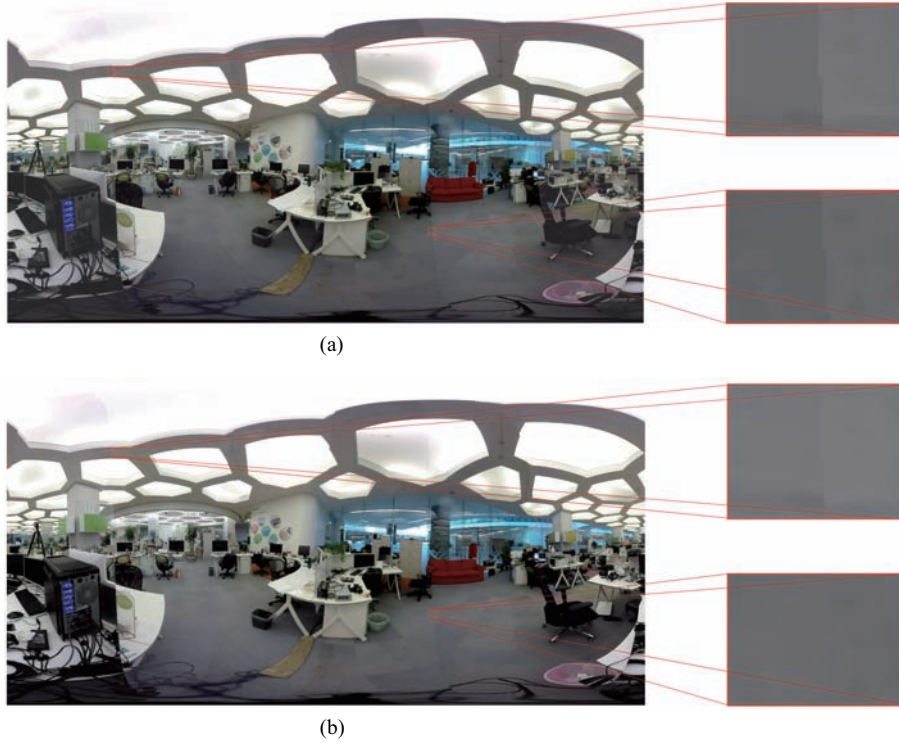


图 10 (网络版彩图) 亮度补偿效果对比

Figure 10 (Color online) Experiment results of referenced method and proposal method. (a) Ref. [5], YUV color space; (b) proposal method

4.4 对比实验

当前应用于全景图像/视频亮度补偿的算法主要有文献 [5, 8~10] 等. 其中, 文献 [5] 介绍一种在 $l\alpha\beta$ 空间进行图像风格化的方法, 可将其应用在亮度补偿中. 文献 [8] 介绍了一种暗角和亮度补偿的方法, 其中亮度补偿的方法与文献 [5] 相同. 文献 [9] 使用的是最基础的直方图亮度补偿的方法, 是本文算法改进前的版本. 文献 [10] 需要对每一帧进行 SIFT 特征提取, 并优化能量函数, 复杂度太高, 不适合应用在实时场景中. 这里本文选择文献 [5] 中的方法进行对比实验.

在对比实验的配置上, 文献 [5] 中原始算法在 $l\alpha\beta$ 空间^[17] 进行, 需要将输入图像 (YUV 空间) 变换到 $l\alpha\beta$ 空间, 复杂度较高, 并且在实验中发现, 直接在 YUV 空间中进行变换并未降低效果甚至效果更好, 所以在对比实验中, 本文直接使用文献 [5] 中算法在 YUV 空间的 Y 分量进行变换, 在实时处理的情况下考察其亮度补偿效果. 文献 [5] 的实现如下:

$$Y'_{\text{cur}} = \frac{\text{var}_{\text{ref}}}{\text{var}_{\text{cur}}} (Y_{\text{cur}} - \text{mean}_{\text{cur}}) + \text{mean}_{\text{ref}}, \quad (11)$$

其中, mean_{ref} , var_{ref} 表示前一帧图像与当前图像重叠部分的均值和方差, mean_{cur} , var_{cur} 表示当前图像与前一帧图像重叠部分的均值和方差. Y_{cur} , Y'_{cur} 表示当前图像校正前和校正后的亮度分量值. 本文算法与文献 [5] 中算法的复杂度对比如表 2 所示, 亮度补偿效果对比如图 10 所示. 其中图 10 选择放大的区域, 是图像对齐较好, 并且灰度比较均匀的拼缝区域, 放大时只展示了亮度分量. 从图 10 和表 2 可以看出:

(1) 文献 [5] 中算法已经得到了较好的亮度补偿效果, 但是本文算法依然在细节上有所提升, 在拼接处有更小的亮度跳变.

(2) 相比于文献 [5], 本文算法更高效, 只需要大约一半的时间. 主要原因在于, 本文算法可以很容易完全使用整数运算实现, 同时不损失精度, 而文献 [5] 中算法对每个像素进行浮点运算, 否则将引入整数截断的精度损失.

5 结论

针对基于鱼眼相机的全景视频拼接, 本文提出一种新的基于直方图匹配的亮度补偿方法. 首先将鱼眼图像展开, 并进行畸变矫正, 使用等距圆柱投影将其投影在平面上, 然后使用非均匀采样对相邻图像的重叠部分进行采样, 最后基于非均匀采样得到的采样点, 自适应选择基准图像, 从基准开始, 依次进行直方图匹配, 达到对所有图像进行亮度补偿的效果. 实验表明, 非均匀采样大大降低了计算复杂度, 而只有少量的图像质量损失, 自适应基准图像在引入少量计算的情况下, 可以大大提高亮度补偿算法的鲁棒性, 避免出现不自然的拼接输出.

虽然在理论上, 依次进行直方图补偿保证了所有输入图像的直方图都向基准图像对齐, 但是实践中依然发现, 依次补偿的方法会在一定程度上出现累积误差, 从而造成首尾相连和多处重叠部分有不自然的过度痕迹, 这也是该算法有待改进的地方. 同时, 全景视频拼接的实时性随着输入视频流的个数的增加和输出视频的分辨率的提升而面临更大的挑战, 亮度补偿的 CPU 实现将成为实时性的瓶颈, 所以如何将该算法在 GPU 上并行实现, 将是下一步工作的重点.

参考文献

- 1 Deng X M, Wu F C, Wu Y H, et al. Automatic spherical panorama generation with two fisheye images. In: Proceedings of the 7th World Congress on Intelligent Control and Automation, Chongqing, 2008. 5955–5959
- 2 Ho T, Budagavi M. Dual-fisheye lens stitching for 360-degree imaging. In: Proceedings of International Conference on Acoustics, Speech and Signal Processing, New Orleans, 2017
- 3 Liao W S, Hsieh T J, Liang W Y, et al. Real-time spherical panorama image stitching using OpenCL. In: Proceedings of International Conference on Computer Graphics and Virtual Reality, Las Vegas, 2011. 113–119
- 4 Xu Y, Zhou Q H, Gong L W, et al. High-speed simultaneous image distortion correction transformations for a multicamera cylindrical panorama real-time video system using FPGA. IEEE Trans Circ Syst Video Technol, 2014, 24: 1061–1069
- 5 Reinhard E, Adhikhmin M, Gooch B, et al. Color transfer between images. IEEE Comput Graph Appl, 2001, 21: 34–41
- 6 Tian G Y, Gledhill D, Taylor D, et al. Colour correction for panoramic imaging. In: Proceedings of the 6th International Conference on Information Visualisation, London, 2002. 483–488
- 7 Brown M, Lowe D G. Automatic panoramic image stitching using invariant features. Int J Comput Vision, 2007, 74: 59–73
- 8 Kim Y N, Sim D G. Vignetting and illumination compensation for omni-directional image generation on spherical coordinate. In: Proceedings of the 16th International Conference on Artificial Reality and Telexistence-Workshops, Hangzhou, 2006. 413–418
- 9 Fecker U, Barkowsky M, Kaup A. Histogram-based prefiltering for luminance and chrominance compensation of multi-view video. IEEE Trans Circ Syst Video Technol, 2008, 18: 1258–1267
- 10 Yamamoto K, Oi R. Color correction for multi-view video using energy minimization of view networks. Int J Autom Comput, 2008, 5: 234–245
- 11 Basu A, Licardie S. Alternative models for fish-eye lenses. Pattern Recogn Lett, 1995, 16: 433–441

- 12 Salomon D. Transformations and Projections in Computer Graphics. Berlin: Springer, 2007
- 13 Ye Y, Alshina E, Boyce J. Algorithm descriptions of projection format conversion and video quality metrics in 360Lib version 54. In: Joint Video Exploration Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Document: JVET-H1004, Macau, 2017
- 14 Szeliski R, Shum H Y. Creating full view panoramic image mosaics and environment maps. In: Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques. New York: ACM, 1997. 251–258
- 15 Sun Y, Lu A, Yu L. Ahg8: WS-PSNR for 360 video objective quality evaluation. In: Joint Video Exploration Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Document: JVET-D0040, Chengdu, 2016
- 16 Chen J L, Liu J L, Ye J H, et al. Luminance and chrominance correction for multi-view video using overlapped local histogram matching. Chinese J Image Graph, 2007, 12: 1992–1999 [陈建乐, 刘济林, 叶建洪, 等. 多视点视频中基于局部直方图匹配的亮度和色差校正. 中国图象图形学报, 2007, 12: 1992–1999]
- 17 Ruderman D L, Cronin T W, Chiao C C. Statistics of cone responses to natural images: implications for visual coding. J Opt Soc Am, 1998, 15: 2036–2045

A luminance compensation method for fisheye video panorama stitching

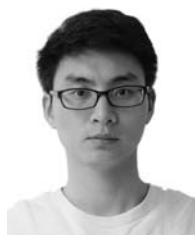
Jiguo LI^{1,4}, Yue WANG², Xinfeng ZHANG³ & Siwei MA^{2*}

1. *Institute of Computing Technology, Chinese Academy of Sciences, 100190, China;*
2. *School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China;*
3. *School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798, Singapore;*
4. *School of Computer and Control Engineering, University of Chinese Academy of Sciences, Beijing 100049, China*

* Corresponding author. E-mail: swma@pku.edu.cn

Abstract The fisheye camera has becoming the most popular panoramic video capturing device. In fact, the stitching for fisheye images has drawn more and more attention. To solve the artifacts resulting from the exposure difference of several cameras, this paper proposes a luminance compensation method based on an uneven sampling histogram. According to the transition of the spatial sampling rate with the latitude's increasing in panorama, the proposed method samples the overlapped region of two adjacent images to calculate the histograms, and matches the histograms one by one to balance the luminance. Moreover, in consideration of the influence of luminance distribution to panorama video quality, we present a method for adaptive reference image selection. The selected reference image based on the proposed method can improve the overall quality of the panorama effectively. Due to the consideration of the distortion characteristics of fisheye images, the proposed method does not require an accurate alignment, and it is also well-adapted to luminance variations. Experimental results show that the proposed method can reduce the time complexity for panorama video stitching significantly without obvious quality degeneration, and it is suitable for fisheye lens-based real-time panoramic video stitching.

Keywords image processing, virtual reality, histogram matching, fisheye lens, video stitching



Jiguo LI was born in 1995. He received his bachelor degree from Huazhong University, Wuhan, China, in 2016. Currently he is a Ph.D. candidate at the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China. His research interests include video stitching and cross-media intelligence.



Yue WANG was born in 1984. He received his B.E. degree from Tsinghua University in 2006, and Ph.D. degree from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2012. He recently founded VsceneVideo Technology Co., Ltd. His research interests include video encoding, processing and transmission.



Xinfeng ZHANG was born in 1983. He received a B.S. degree in computer science from the Hebei University of Technology, Tianjin, China, in 2007, and a Ph.D. degree in computer science from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2014. From July 2014 to October 2017, he was a research fellow in Nanyang Technological University, Singapore. He is currently a post-doctoral fellow with the Ming Hsieh Department of Electrical Engineering, University of Southern California, Los Angeles, USA. His research interests include image and video processing and image and video compression.



Siwei MA was born in 1979. He received his Ph.D. degree from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2005. Currently, he is a professor at the School of Electrical Engineering and Computer Science, Peking University. His research interests include video encoding, processing and transmission.