

一种基于动态图片技术的高动态范围全景视频应用

马志友

深圳看到科技有限公司

DOI:10.12238/acair.v2i4.10367

[摘要] 在数字媒体和视觉艺术的领域,技术的每一次飞跃都极大地丰富了我们的视觉体验和社交互动方式,近年来全景视频和动态图正逐渐改变我们记录与分享世界的方式。随着技术的不断发展,这些先进的图像捕获方式为我们提供了全新的视角和体验。本文旨在探讨前沿的影像技术——动态图片与高动态范围(HDR)以及全景视频技术相结合的一种扩展应用,即全景HDR动态图。全景HDR动态图的应用带来了多重好处:首先,全景HDR动态图通过互动触摸方式增强了社交互动,因为朋友和家人可以通过动态效果更加真实地感受到照片中的场景。其次,全景HDR动态图通过高动态范围高清晰度且变化多端的全景特效,深度展示了内容的真实性,这对于新闻报道或证据记录尤为重要。最后,全景HDR动态图有助于保留那些难以用静态图片捕捉的进程瞬间,如孩子们的笑容变化或是宠物的活泼动作。通过人工智能技术,有助于在全景HDR动态图内容生成过程中更有力地进行关键内容呈现,从而改变我们记录和分享生活的方式与方法。

[关键词] 动态图片; 动态照片; 高动态范围; 全景视频; 全景HDR动态图; 全景特效; 人工智能
中图分类号: G255.71 **文献标识码:** A

A high dynamic range panoramic video application based on dynamic picture technology

Zhiyou Ma

Shenzhen Kandao Technology Co., Ltd.

[Abstract] In the fields of digital media and visual arts, every leap in technology has greatly enriched our visual experience and social interaction. In recent years, panoramic videos and Live Photo (or Motion Photo) images are gradually changing the way we record and share the world. With the continuous development of technology, these advanced image capture methods provide us with new perspectives and experiences. This article aims to explore an extended application of cutting-edge imaging technology – Live Photo combined with high dynamic range (HDR) and panoramic video technology, namely Panoramic HDR Live Photo (PHLP). The application of panoramic HDR Live Photo brings multiple benefits: firstly, Panoramic HDR Live Photo enhance social interaction through interactive touch, as friends and family can feel the scene in the photo more realistically through dynamic effects. Secondly, Panoramic HDR Live Photo with high dynamic range and high-definition panoramic effects deeply demonstrate the authenticity of the content, which is particularly important for news reporting or evidence recording. Finally, Panoramic HDR Live Photo help preserve process moments that are difficult to capture with static images, such as children's smiling changes or pets' lively movements. Through artificial intelligence technology, key content can be presented more effectively in the process of generating Panoramic HDR Live Photo images, thereby changing the way and methods we record and share our lives.

[Key words] Live Photo; Motion Photo; High Dynamic Range; Panoramic video; Panoramic HDR Live Photo; Panoramic Effects; Artificial Intelligence

引言

动态图集合了动画设计、平面设计和电影语言,它的表现形式各式各样^[1]。动态图片Live Photo是苹果公司2015年推出的一项创新功能,它巧妙地将静态图片与短视频结合在一起,创造出一种全新的视觉体验。当你按下快门的那一刻,iPhone不仅

捕捉了一张高分辨率的JPEG格式静态图片,还同时录制了一段简短的视频(通常为MOV格式)。2016年,谷歌专门给动态照片设计了一套MicroVideo的“单文件”标准,但一直没有正式公布,直至如今才正式发布改名为Motion Photo动态照片存储标准。区别于苹果系统的动态图片,Google旗下的Android系统提供的微

视频或者动态照片使用的是单个文件方式,即将图片和视频绑定在一起,其中图片在先视频在后。同年微软在Window10桌面版和移动版的照片应用新增了动态影像(动态图像)的功能。图片和视频的这种组合使得动态图片或者动态照片不再是一张单纯的静止的照片,而是一个能够“复活”的动态记忆。用户只需长按照片,就能触发这段动态视频的播放,仿佛时间倒流,重新体验那一刻的情感与氛围。这项技术极大地增强了照片的表现力,让回忆更加生动鲜活。此外,动态图片还促进了社交互动,因为它提供了更丰富的信息,使人们能够更准确地传达情感和故事背景。

动态图形运动往往与声音相配合,视听之间形成一种统合关联的“通感”审美^[4]。全景视频技术致力于提供360度无死角的视觉体验,让观众仿佛置身于场景之中。全景摄像机通过使用多个镜头或特殊的鱼眼透镜,能够捕捉到水平方向360度及垂直方向上的广阔视野,支持双目立体视觉,为用户带来3DOF和6DOF自由度的沉浸式观看体验^[5]。这种技术广泛应用于虚拟现实(VR)、旅游、房地产展示等领域,让用户能够远程探索世界的各个角落。全景视频的核心在于其高效的三维空间图像拼接算法^[6],通过光流拼接、加权融合、图像ISP校正算法等算法将多张不同角度拍摄的图片有机的无缝拼接成一个连续的画面,同时保持高质量的细节和色彩一致性。随着技术的进步,现在的全景视频不仅支持2K、4K、8K乃至12K以上的高清分辨率,还能实现实时拼接和播放,大大提升了用户体验。

1 全景HDR动态图

全景HDR动态图是在动态图片技术的基础上,进一步向全景领域拓展的产物。它包含一个主要的静态图片和一个全景视频文件,除此之外,它可以包含增益图(GainMap)作为次要图像,以增强图像的视觉效果。这种格式支持包括JPEG、HEIC在内的多种图片格式,而视频文件格式则需满足MP4或MOV规范,确保了高度的兼容性和灵活性。高动态范围(HDR)^{[2][3]}技术在超高清视频中的应用可实现更大曝光动态范围,HDR增益图有两种表现形式:一是Grayscale GainMap,即灰度增益图,它主要关注照片的亮度层次,以增强暗部和亮部的细节与纵深感。二是Color GainMap,即彩色增益图,它则更侧重于色彩的校正和平衡,以提升整体色彩的鲜艳度,接近自然。除此之外,全景动态图要求视频至少包含一个视频轨道,并且可以不包含或者包含最多一个音频轨道。这意味着制作者可以根据需要选择是否添加背景音乐或解说,从而创造出更加个性化和富有创意的作品。

全景HDR与动态图片的创新融合,标志着多媒体体验领域的一次重大飞跃。随着命名空间的增加,这一技术不仅扩展了其应用场景的边界,还极大地丰富了用户体验的多样性。全景动态图设置的引入,让用户能够在捕捉广阔景象的同时,记录下动态瞬间,从而打破了静态图片的限制,让回忆更加生动鲜活。而双目模式的加入,则进一步推动了沉浸式体验的发展,通过模拟人眼视觉原理,为用户带来前所未有的深度感知和立体视觉效果。

2 技术实现与挑战

全景HDR动态图的生成要求摄影设备在拍摄时就要同时处

理图像和视频数据,并确保两者同步。在平台上复现全景HDR动态照片则需要额外的开发工作来整合图片和视频显示。在技术实现上,全景HDR动态图在原有的动态照片基础上进行一些扩展。对全景HDR动态图的XMP描述子的扩展实现具体包括(但不限于)以下内容:(1)增加命名空间;(2)增加全景动态图设置;(3)增加双目模式;(4)增加全景展开模式;(5)增加全景特效模式。

全景HDR动态图技术通过将静态全景照片与动态视频元素相结合,创造出一种全新的高动态范围视觉呈现形式。用户在浏览全景图片时,能够感受到场景中动态元素的变化,如同置身于一个生动的三维空间之中。这种技术的应用不仅限于个人娱乐,也在旅游、教育、广告等多个领域展现出巨大潜力,为用户提供更加沉浸式的体验。

利用双目立体视觉原理,全景HDR动态图技术为用户带来更为真实的深度感和立体感。通过模拟人眼观察物体的方式,这种模式能够在平面图像上呈现出接近真实的三维效果,极大地提升了用户的视觉享受。在实际应用中,无论是观看自然风光、城市景观还是艺术作品,用户都能感受到前所未有的空间层次和细节丰富度,仿佛能够触摸到画面中的每一个元素。

在全景HDR动态图技术中,通过全景展开模式提供了多种选项以适应不同的视觉需求和使用场景。首先,默认的平面模式是最基本的展开方式,它保持图像的原始比例和结构,适用于需要保持画面真实性的场合。其次,全景常见的等距柱状投影模式则引入了等比例拉伸,使得图像在各个方向上都保持一定的平衡,适合于展示更加均衡的视角,适合整图整个画面查看。六方体贴图模式和等角度立方体贴图模式进一步优化了图像的展开,通过算法提高了数据保存效率;使得图像在保持整体协调的同时,细节部分也得到了增强,这对于提升观看体验尤为重要。

全景动态图技术的特效模式为内容创作带来了前所未有的视觉体验。动态小行星模式通过智能算法优化图像的色彩饱和度和对比度,突出画面关键要素且曲线优美;子弹时间模式则为用户带来一种复古而神秘的视觉效果,适合营造特定的氛围;移动延时模式创新性地模拟了时间逆流的效果,为静态图片增添了动态的时间维度;定点环绕模式让用户仿佛置身其中,提供了沉浸式的观看体验;大力超人模式允许用户从不同角度同时观察同一场景,极大地丰富了观者的感知层次;跨越地平线模式则以经典的黑白色调强化了图像的艺术感,适用于追求极简风格的创作者。

这些多样化的特效模式不仅增强了内容的视觉冲击力,也提升了观众的兴趣和参与度。例如,在旅游摄影领域,使用移动延时模式可以捕捉到游客在名胜古迹前的精彩瞬间,并赋予其流动的时间感;而在产品展示上,利用定点环绕模式可以让潜在客户全方位地了解商品特性,提高转化率。此外,对于艺术家而言,跨越地平线模式提供了一种全新的艺术表达方式,帮助他们探索更多可能性。总之,全景动态图技术的特效模式的创新应用正在不断拓展着多媒体体验的边界,让每个人都能成为创意无限的内容创造者。

3 人工智能的灵活应用

对于全景HDR动态图而言,高质量的图像拼接算法至关重要,

尤其是在处理大量数据时仍能保持高效率和低延迟。此外, 如何用AI算法实现切片等方式在有限的带宽下传输高分辨率的全景HDR动态图^[7], 以及各种手机等多种平台如何优化适配全景HDR动态图的播放器以支持各种交互操作, 乃至画面的Qos保障^[8]也是一个值得探讨的方向。本文对此暂不展开讨论, 主要聚焦AI在首图获取的应用。在全景HDR动态图具体应用中, 由于要实时控制时间和内容的冲突, 针对首图或者封面图的选取是拍摄者的技术挑战。一种简便粗暴的做法把首图获取点作为固定点, 无论是把拍摄视频中部作为首图获取点, 抑或是将拍摄结束时间作为首图获取点, 均属此类。

硬性规定首图的时间轴位置虽然带来了用户操作便利性, 缺点是所拍摄动态图可能不能获取到最佳首图画面效果。引入深度学习和人工智能技术在全景HDR动态图中的首图出图是一个不错的选择。在某种程度上, 首图的选择与视频编码中的关键帧选取也有类似之处。常见的关键帧用法在码率控制时也可以区分有固定GOP (Fixed GOP) 方式和动态GOP (Dynamic GOP) 方式。显然选择固定位置出首图的方式是倾向于用户来决定和判断最佳点, 而动态的首图方式则是用算法来辅助调整最佳首图位置^{[13][14]}。

$$D_n(x, y) = |f_{n+1}(x, y) - f_n(x, y)| \cap |f_n(x, y) - f_{n-1}(x, y)|$$

公式1 三帧差分法

人工智能在全景HDR动态图的首图选择上可以综合运用以下算法结果: (1) 计算视频帧间差分, 获取帧间差距的最大值(如公式1三帧差分法); (2) 计算视频光流量, 获取光流移动最小的帧(见公式2 Lucas-Kanade光流法); (3) 基于颜色直方图计算帧间颜色/亮度变化最大值; (4) 语音识别辅助决策^{[9][10]}; (5) 视频人脸识别手势识别辅助决策^{[11][12]}。

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n (I_{xi})^2 & \sum_{i=1}^n I_{xi}I_{yi} \\ \sum_{i=1}^n I_{yi}I_{xi} & \sum_{i=1}^n (I_{yi})^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^n I_{xi}I_{ti} \\ \sum_{i=1}^n I_{yi}I_{ti} \end{bmatrix}$$

公式2 Lucas-Kanade光流法

全景HDR动态图的多样性人工智能应用有助于保留那些难以用静态图片捕捉的进程瞬间, 如孩子们的笑容变化或是宠物的活泼动作。通过人工智能技术, 有助于在全景HDR动态图内容生成过程中更有力地进行关键内容呈现, 从而改变我们记录和分享生活的方式与方法。

4 结语

动态照片和全景视频技术的发展, 不仅为我们提供了记录生活的独特方式, 也极大地推动了多媒体内容的创作与分享。从简单的动态捕捉到复杂的全景体验, 这些技术正逐步成为我们日常生活的一部分, 让我们的世界变得更加丰富多彩。未来, 随着技术的不断成熟和创新, 我们可以期待更多令人兴奋的应用出现, 进一步拓宽我们的视野和想象空间。例如, 在虚拟现实 (VR) 内容的制作中, 开发者可以利用这些技术创建出令人身临其境的环境。通过专业的软件工具, 创作者可以将全景照片和全景视

频拼接成一幅活动的全景HDR动态图。这些全景HDR动态图不仅可以在微博、小红书等网站上分享, 还可以被iPhone分享到朋友圈供用户浏览。

随着技术的不断进步和市场需求的日益增长, 全景动态图技术正站在一个崭新的起点上, 展望未来, 其发展方向将更加多元且充满无限可能。一方面, 技术创新将持续推动全景动态图向更高质量的视觉体验迈进。例如, 通过更高分辨率的图像捕捉、更流畅的动态播放以及更细腻的色彩处理, 使得全景动态图能够呈现出更加逼真、生动的画面效果。同时, 结合人工智能与机器学习技术, 全景动态图有望实现更加智能化的内容生成与编辑, 为用户提供更加个性化、定制化的体验。

【参考文献】

- [1]刘妍彤,张阳.数字媒体界面中动态图形特征研究[J].数字通信世界,2022(5):121-123.
- [2]国家广播电视总局发布HDR技术标准[J].中国有线电视,2022(01):14.
- [3]顿胜堡,李新国,张志刚,等.HDR显示技术及标准化趋势[J].信息技术与标准化,2016(10):17-21.
- [4]朱文涛.动态图形视听设计的通感实验研究[J].创意与设计,2022(5):43-49.
- [5]曾焕强,孔庆玮,陈婧,等.沉浸式视频编码技术综述[J].电子与信息学报,2024,46(2):602-614.
- [6]宋聪聪,高策,张艳超,等.基于三维球面模型的全景视频实时拼接方法[J].光学学报,2023,43(10):65-78.
- [7]谢文京,王悦,张新峰,等.基于视点的全景视频编码与传输优化[J].扬州大学学报(自然科学版),2018,21(2):45-49.
- [8]杜丽娜,卓力,李嘉锋.全景视频QoE评价研究进展[J].信号处理,2022,38(9):1831-1842.
- [9]Speech emotion recognition:Features and classification models[J].Lijiang Chen;;Xia Mao;;YuLi Xue;;Lee Lung Cheng. Digital Signal Processing.2012.
- [10]马志友,杨莹春,吴朝晖.二次特征提取及其在说话人识别中的应用[J].电路与系统学报,2003,8(2):130-133.
- [11]陈钊,陈亮,夏天,等.基于计算机视觉与智能识别的电力施工安全监测技术[J].电子设计工程,2024,32(9):114-118.
- [12]杨纪争,冯筠.面向静态手势识别的边缘序列递归模型算法[J].计算机辅助设计与图形学学报,2017,29(4):599-606.
- [13]A spatial-temporal framework based on histogram of gradients and optical flow for facial expression recognition in video sequences[J].Xijian Fan;Tardi Tjahjadi.Pattern Recognition.2015.
- [14]Anadaptive classification system for video-based face recognition[J].Information Sciences,2010.

作者简介:

马志友(1976--),男,汉族,浙江省乐清人,硕士研究生,专业:计算机应用; 职称:高级工程师(副高),研究方向:音视频编解码与直播,全景相机。