

团 体 标 准

T/INFOCA 12—2024

全景视频拼接技术指南

Technical guidelines for panoramic video stitching

2024-××-××发布

2024-××-××实施

中关村现代信息消费应用产业技术联盟 发布

目 次

前 言	II
引 言	III
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 缩略语	2
4 全景视频拼接技术建议	2
4.1 输入建议	2
4.2 输出建议	3
4.3 速度和质量建议	3
参考文献	5

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

本文件的某些内容可能涉及专利。本文件的发布机构不承担识别这些专利的责任。

本文件由中关村现代信息消费应用产业技术联盟提出并归口。

本文件起草单位：北京大学、大连理工大学、哈尔滨工业大学

本文件主要起草人：张启、刘宇擎、熊瑞勤、王苜社、王兴涛、范晓鹏、贾棋、冯晓美

引 言

全景视频是未来视频产业发展的重要趋势之一。全景视频拼接是全景视频制作过程中的必要环节，直接影响全景视频内容的制作效率、质量。因此，对全景视频拼接技术的设计与实现进行指导具有重要意义。

本文件的目的是对全景视频拼接过程的输入输出提出技术指导，并考虑实际应用需求，给出对拼接速度和拼接质量之间对应关系的建议。

本文件的制定对全景视频拼接技术满足功能性、效率和质量方面的要求具有指导作用，有助于促进全景视频产业发展。

全景视频拼接技术指南

1 范围

本文件提出了对全景视频拼接技术的指导。

本文件适用于全景视频拼接技术设计与实现。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 9813.1 计算机通用规范第1部分：台式微型计算机

GB/T 9813.2 计算机通用规范第2部分：便携式微型计算机

GB/T 9813.3 计算机通用规范第3部分：服务器

GB/T 38259 信息技术虚拟现实头戴式显示设备通用规范

GY/T 340-2020 超高清清晰度电视图像质量主观评价方法双刺激连续质量标度法

GY/T 307 超高清清晰度电视系统节目制作和交换参数值

GY/T 315 高动态范围电视制作和交换图像参数值

GY/T 356 VR视频系统节目制作和交换用视频参数值

3 术语和定义

下列术语和定义适用于本文件。

3.1

全景视频 panoramic video

指能够覆盖水平和垂直方向上的大部分甚至全部视角的视频，用户可以随意调节观看角度进行观看。

3.2

180度全景视频 180-degree panoramic video

能够覆盖到某个场景在水平方向上 180 度景象的全景视频。

3.3

360度全景视频 360-degree panoramic video

能够覆盖到某个场景在水平方向上 360 度景象的全景视频。

3.4

立体全景视频 stereo panoramic video

使观看者能够感受到立体景象的全景视频。

3.5

全景视频拼接 panoramic video stitching

将多个内容部分重叠或相互衔接的视频合并成全景视频的技术过程。

3.6

位深 bit depth

图像或视频每个像素的数据所占的二进制位数。

3.7

全景映射 panorama projection

将球面或柱面上的全景视频图像投影至二维平面的技术过程。

3.8

等距柱状映射 equirectangular projection (ERP)

一种全景映射方式，将全景视频图像的球面等距地投影到平面上，经线投影为等距的垂直线，纬线投影为等距的水平线。

3.9

平均主观得分 mean opinion score (MOS)

一种评价图像或视频质量的主观指标，让人类受试者在一个标准化的量表上对他们感知到的质量进行评分，然后将这些评分求平均值。

4 缩略语

下列缩略语适用于本文件。

ERP：等距柱状映射（equirectangular projection）

MOS：平均主观得分（mean opinion score）

4 全景视频拼接技术建议

4.1 输入建议

4.1.1 相机数目

4.1.1.1 180度全景视频相机数目

采集用相机数目宜不低于2台。

4.1.1.2 360度全景视频相机数目

采集用相机数目宜不低于3台。

4.1.1.3 360度立体全景视频相机数目

采集用相机数目宜不低于4台。

4.1.2 相机分辨率

采集用相机分辨率宜支持如下规格：1920*1280、3840*2160、6048*4032。

4.1.3 相机帧率

采集用相机帧率宜支持如下规格：60、90、120。

4.1.4 相机色彩位深

采集用相机色彩位深宜达到或高于10位。

4.2 输出建议

4.2.1 全景视频分辨率

4.2.1.1 180度全景视频分辨率

拼接后的180度全景视频分辨率宜达到或高于3840*3840。

4.2.1.2 360度全景视频分辨率

在全景映射格式为ERP的情况下，拼接后的360度全景视频分辨率宜达到或高于7680*3840。

4.2.1.3 360度3D全景视频分辨率

在全景映射格式为ERP的情况下，拼接后的360度立体全景视频分辨率宜达到或高于7680*3840。

4.2.2 全景视频帧率

拼接后的全景视频帧率宜达到或高于50。

4.2.3 全景视频色彩位深

拼接后的全景视频色彩位深宜达到或高于10位。

5.3 速度和质量建议

5.3.1 概述

不同的应用场景对全景视频拼接速度要求不同，拼接速度和拼接质量之间存在权衡关系。在直播场景下，全景视频拼接为了满足实时性要求，通常采用较简单的算法，降低了拼接质量。在离线内容制作场景下，通常不对拼接速度进行要求，可以采用较复杂的拼接算法，达到最佳的拼接质量。因此，对全景视频拼接速度和质量统一进行技术参数上的建议，适用于直播、离线内容制作等多种应用场景。

5.3.2 全景视频拼接速度和质量测试配置建议

对全景视频拼接速度和质量进行技术参数建议前，本文件给出用于测试拼接速度和质量配置建议，包含测试硬件配置建议、主观评价测试方法建议和测试源配置建议。

5.3.2.1 测试硬件配置建议

全景视频拼接速度和质量测试基于头戴式设备、计算机和服务器等硬件。头戴式设备宜遵循GB/T 38259-2019的规定。计算机和服务器宜遵守GB/T 9813.1、GB/T 9813.2或GB/T 9813.3的规定。

测试时头戴式设备的显示分辨率宜达到或高于1440*1440，显示频率宜达到或高于72赫兹。

5.3.2.2 主观评价测试方法建议

全景视频拼接质量受到采集相机素质、采集环境、拼接算法性能等多种因素影响，本文件建议使用主观评价方法测试全景视频拼接质量。

全景视频拼接质量主观评价的测试阶段、测试图像的显示和评分标度宜遵循GY/T 340-2020中5.5、5.6、5.7的规定。主观评价人员配置情况宜满足：人员不低于15人，分多组进行评价。

5.3.2.3 测试源配置建议

全景视频拼接速度和质量测试需要源采集视频和源全景视频，其中源全景视频由源采集视频经人工拼接后生成，不存在明显的噪声、亮度差异、拼缝等影响源全景视频质量的缺陷。

测试用的源采集视频内容格式宜符合GY/T 307或GY/T 315中的要求，源全景视频内容格式宜符合GY/T 356的要求，视频内容、时长、数目等其它配置宜符合GY/T 340-2020中5.3的要求。

5.3.3 速度和质量指标建议

当应用场景对拼接速度没有要求时，在5.3.2所描述的测试配置下，全景视频拼接质量的MOS评分宜达到95分。

当应用场景对拼接速度有要求时，全景视频拼接速度和质量宜达到如下规格：

- 1) 当拼接后的全景视频分辨率为3840*1920时，拼接速度宜达到每秒处理120帧，在4.3.2.1-4.3.2.3所描述的测试配置下，全景视频拼接质量的MOS评分宜达到85分。
- 2) 当拼接后的全景视频分辨率为5760*2880时，拼接速度宜达到每秒处理50帧，在4.3.2.1-4.3.2.3所描述的测试配置下，全景视频拼接质量的MOS评分宜达到85分。
- 3) 当拼接后的全景视频分辨率为7680*3840时，拼接速度宜达到每秒处理30帧，在4.3.2.1-4.3.2.3所描述的测试配置下，全景视频拼接质量的MOS评分宜达到85分。

参考文献

- [1] ITU-R BT.500-11, Methodology for the subjective assessment of the quality of television pictures.
- [2] ITU-T P.800.1 Mean Opinion Score (MOS) terminology.
- [3] ITU-T P.919 Subjective test methodologies for 360° video on head-mounted displays.