

团 体 标 准

T/INFOCA 1—2024

全景视频增强处理技术规范

Technical specifications for panoramic video enhancement processing

（征求意见稿）

（本稿完成日期：2024 年 3 月 25 日）

××××-××-××发布

××××-××-××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	III
引 言	IV
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
5 全景视频增强处理技术基本架构	3
5.1 概述	3
5.2 基本架构	3
6 技术要求	4
6.1 全景视频画质提升	4
6.2 全景视频投影变换及视点选择	6
6.3 全景视频编辑	7
7 测试方法	7
7.1 全景视频画质提升	8
7.2 全景视频投影变换及视点选择	10
7.3 全景视频编辑	11
附录 A （规范性） 测试配置	12
A.1 测试系统配置	12
A.1 测试源配置	12
附录 B （规范性） 技术指标	13
B.1 主观评价	13
B.2 客观测试	13
参考文献	18

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。

请注意本文件的某些内容可能涉及专利，具体名称及专利号在本文件的注释中给出。

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：中国科学院大学，中兴通讯股份有限公司，河南广播电视台，哈尔滨工业大学，中国传媒大学

本文件主要起草人：张新峰，安泓宇，张远建，贾中原，范晓鹏，赵德斌，万照麟，张远，叶龙，冯晨曦

引 言

随着5G、超高清、虚拟现实等新兴技术的发展，为满足用户对美好视听的新需求，业界提出了“5G高新视频”概念，其中，“高”是指视频融合4K/8K、3D、虚拟现实/增强现实/混合现实、高帧率、高动态范围、广色域等技术格式；“新”是指具有新奇视觉体验的创新应用场景，能够吸引观众兴趣。全景视频是超高清视频、互动视频、沉浸式视频、云游戏、虚拟现实、增强现实和混合现实等技术应用的重要载体，通过对全景视频进行增强处理，可以提供高质量沉浸式视觉体验。

针对“更高技术格式、更新应用场景、更美视听体验”的高新视频新业态发展需求，全景视频逐渐步入到人们的日常生活当中，广泛应用于数字创意、互动娱乐、智慧文旅、智能驾驶、安防监控和视频会议等场景。本文件规定的全景视频增强处理技术规范能够应用于评估新的全景视频增强处理技术，对相关技术在行业应用前景进行评估，形成一套适用性、可操作性强，同时具备全面性、规范性、系统性、科学性的标准化指导意见，企业可以根据本文件对全景视频增强处理技术的功能和性能进行规范化评估和测试，也为设计全景视频增强处理技术时，系统可实现用户高质量沉浸式视觉体验提供有效依据和指导。

为推动全景视频增强处理技术面向用户体验更新优化，完善相应评价体系，满足实际应用需求，促进全景视频增强处理技术标准化建设和规范化发展，特制定此标准。

全景视频增强处理技术规范

1 范围

本文件规定了全景视频增强处理技术的功能要求、性能要求以及测试方法。

本文件适用于全景视频增强处理技术的设计、开发、测试和应用。

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 38259-2019 信息技术虚拟现实头戴式显示设备通用规范

GY/T 155 高清晰度电视节目制作及交换用视频参数值

GY/T 307 超高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 340—2020 超高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

T/UWA 010-2022 智能视频处理系统通用技术规范

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

由全方位摄像机拍摄得到，覆盖整个 360° 空间的视频类型。

3.2

全景视频增强处理技术 panoramic video enhancement processing

对全景视频进行视频优化的技术。

注：视频优化包括但不限于视频画质提升、视频投影变换失真降低、视频编辑效率提升等，该技术可以基于云平台

或者终端进行实现。

3.3

全景视频投影变换 panoramic video projection

全景视频从三维球形格式到二维平面格式或从二维平面格式到三维球形格式的转换映射方法。

3.4

全景视频视点图像 panoramic video viewpoint image

全景视频中人眼实际观看到的部分区域（视角 90° 左右）。

3.5

噪声 noise

视频图像上的颗粒状图像损伤。

注：噪点包括雪花噪声、椒盐噪声、高斯噪声、伪影、块效应、模糊和斑点等。

3.6

特定子图像 specific subimage

视频图像中具有特定含义的图像区域。

3.7

辨识度 degree of identification

人眼主观判断图像中标识物（数字、字母、人、物体等）的容易程度。

4 缩略语

下列缩略语适用于本文件：

AR：增强现实（Augmented Reality）

CMP：立方体投影（Cubemap Projection）

EAP：等面积矩形投影（Equi-Area Projection）

ERP：等距柱状投影（Equi-Rectangular Projection）

HDR：高动态范围（High Dynamic Range）

ISP：正二十面体投影（Icosahedron Projection）

OHP：正八面体投影（Octahedron Projection）

PSNR：峰值信噪比（Peak Signal to Noise Ratio）

SSIM：结构相似性（Structural Similarity）

SSP：球面条带投影（Segmented Sphere Projection）

TSP: 金字塔棱台投影 (Truncated Square Pyramid Projection)

VMAF: 视频质量多方法融合评价 (Video Multimethod Assessment Fusion)

VR: 虚拟现实 (Virtual Reality)

5 全景视频增强处理技术基本架构

5.1 概述

全景视频增强处理技术,在传统的视频增强处理技术基础上针对全景视频成像过程和用户观看习惯等特点做了相应的优化改进,同时集成了基于人工智能的视频处理算法,主要应用于低质全景视频增强、用户端画质提升和内容再生产等场景。

5.2 基本架构

全景视频增强处理技术基本架构如图 1 所示,主要包括全景视频画质提升、全景视频投影变换及视点选择和全景视频编辑三个功能模块。全景视频画质提升主要完成超分辨率、插帧、噪声去除、偏色矫正、去模糊、去抖动和 HDR 转换,实现综合画质提升。全景视频投影变换及视点选择主要包括全景视频投影、全景视频反投影和全景视频视点选择。全景视频编辑主要包括特定子图像去除、画面主体选择和字幕生成。

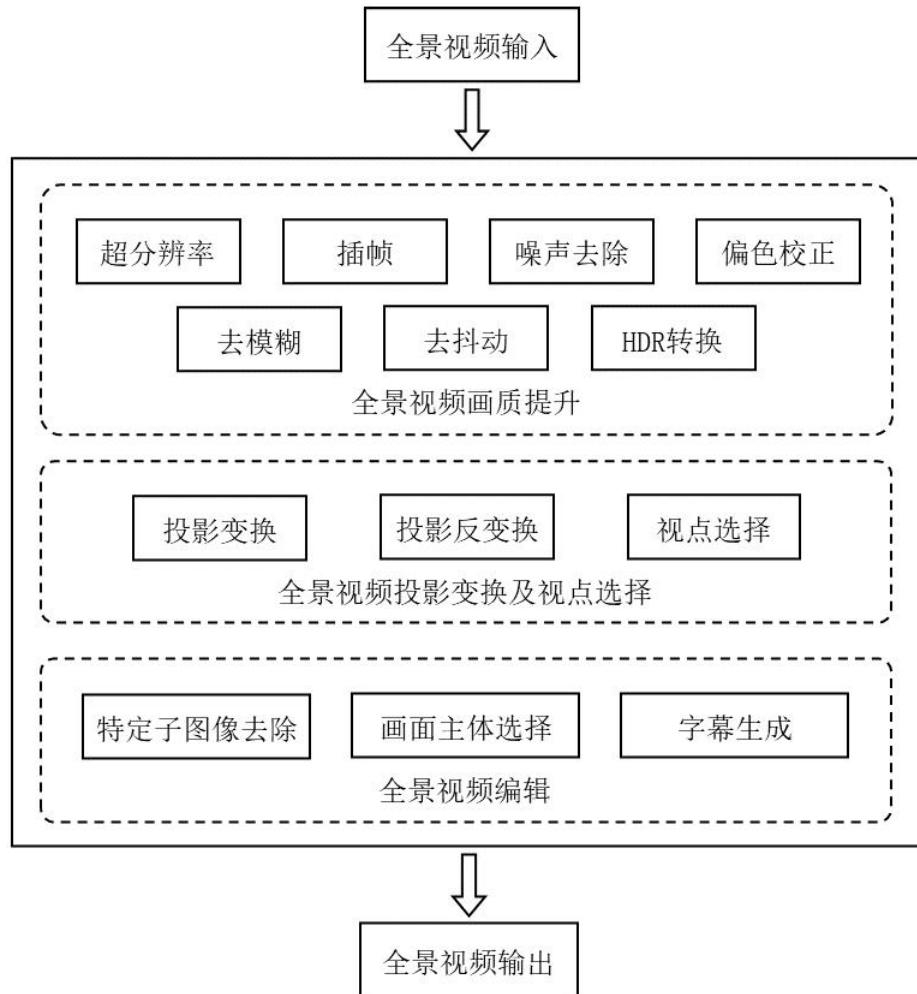


图1 全景视频增强处理技术基本架构

6 技术要求

6.1 全景视频画质提升

6.1.1 超分辨率

全景视频超分辨率技术支持对单帧或多帧低分辨率全景视频进行高分辨率重建处理，并满足以下要求：

- 应至少支持 2 倍、4 倍的分辨率提升；
- 应支持高清 HD（1920×1080）全景视频转 UHD 超高清 4K（3840×2160）全景视频；
- 应支持高清 HD（1920×1080）全景视频转 UHD 超高清 8K（7680×4320）全景视频；
- 应支持超分辨率重建倍率可调节。

6.1.2 插帧

全景视频插帧技术支持帧率提升，并满足以下要求：

- a) 应至少支持 2 倍、4 倍插帧；
- b) 应支持 UHD 超高清 8K（7680×4320）全景视频适配插帧至 120FPS。

6.1.3 噪声去除

全景视频噪声去除技术支持对噪声的修复，噪声类型至少包括以下类型：随机噪声、周期性噪声（网格、条纹噪声等）、压缩噪声（块效应噪声）和边缘锐化，并满足以下要求：

- a) 应支持去噪后全景视频主观清晰度提升，细节保留，未出现过锐现象，辨识度未明显下降；
- b) 应支持输出的全景视频比源视频的图像质量提升率应不低于 5%，图像质量提升率应按附录 B.2.5 中公式计算；
- c) 应支持去噪强度可调节；
- d) 应支持噪声强度智能检测，自适应进行去噪。

6.1.4 偏色校正

全景视频偏色校正技术支持调整整体色调偏差的全景视频，并满足以下要求：

- a) 应支持校正后全景视频的色彩更接近于真实；
- b) 应支持输出的全景视频比源视频的图像质量提升率应不低于 5%，图像质量提升率应按附录 B.2.5 中公式计算；
- c) 应支持偏色校正强度可调节；
- d) 应提供可视化环境，支持人工辅助调整偏色校正结果。

6.1.5 去模糊

全景视频去模糊技术支持处理镜头聚焦不准、目标高速运动等产生模糊的全景视频，并满足以下要求：

- a) 应支持去模糊后全景视频的辨识度优于输入视频；
- b) 应支持输出的全景视频比源视频的图像质量提升率应不低于 5%，图像质量提升率应按附录 B.2.5 中公式计算；
- c) 应支持对大于 16×16 像素的模糊目标进行去模糊处理，模糊目标可识别率应不低于 90%，可识别率应按附录 B.2.6 中公式计算。

6.1.6 去抖动

全景视频去抖动技术支持处理图像画面整体抖动的全景视频，并满足以下要求：

- a) 应支持去抖动后全景视频抖动程度减弱或消失；

- b) 应支持输出的全景视频比源视频的图像质量提升率应不低于 5%，图像质量提升率应按附录 B.2.5 中公式计算；
- c) 应支持输出全景视频视场角损失不会过大，全景视频视场角损失应不高于 10° 。

6.1.7 HDR 转换

全景视频 HDR 转换技术应满足 T/UWA 010-2022 中 6.1.4 的要求。

6.1.8 综合画质提升

经过超分辨率、插帧、噪声去除、偏色校正、去模糊、去抖动和 HDR 转换中一种或多种人工或智能增强处理技术处理全景视频后，满足以下四项要求中的任意三项：

- a) 输出的全景视频比源视频的图像质量提升率应不低于 10%，图像质量提升率应按附录 B.2.5 中公式计算；
- b) 输出的全景视频比源视频的图像客观评价指标 VMAF 平均分至少提高 3 分；
- c) 输出的全景视频比源视频的图像客观评价指标 PSNR 提升应不低于 1dB，SSIM 提升应不低于 0.1，PSNR 应按附录 B.2.1 中公式计算，SSIM 应按附录 B.2.3 中公式计算；
- d) 输出的全景视频 ERP 投影格式比源视频 ERP 投影格式的全景图像客观评价指标 WS-PSNR 提升应不低于 1dB，WS-SSIM 提升应不低于 0.1，WS-PSNR 应按附录 B.2.2 中公式计算，WS-SSIM 应按附录 B.2.4 中公式计算；

6.2 全景视频投影变换及视点选择

6.2.1 投影变换

全景视频投影变换技术满足以下要求：

- a) 应支持三维球形全景视频转换到二维平面 ERP、EAP、CMP、ISP、OHP、TSP 或 SSP 格式全景视频中的任意一种；
- b) 应支持投影变换后的全景视频可以直接对其进行增强处理；
- c) 应支持投影变换后全景视频无明显模糊、扭曲、卡顿、抖动和重影现象。

6.2.2 投影逆变换

全景视频投影逆变换技术满足以下要求：

- a) 应支持二维平面 ERP、EAP、CMP、ISP、OHP、TSP 或 SSP 格式全景视频中的任意一种转换到三维球形全景视频；
- b) 应支持投影变换后全景视频用户可以直接观看；
- c) 应支持投影变换后全景视频无明显模糊、扭曲、卡顿、抖动和重影现象；

- d) 应支持全景视频反变换拼接后接缝处对齐，无明显块效应。

6.2.3 视点选择

全景视频视点选择技术满足以下要求：

- a) 应支持对全景视频高频视点区域进行检测选取；

注：全景视频高频视点区域为用户实际观感兴趣的全景视频局部区域。

- b) 应支持对全景视频高频视点区域进行增强处理；
- c) 应支持对全景视频高频视点区域进行评价测试。

6.3 全景视频编辑

6.3.1 特定子图像去除

全景视频特定子图像去除技术满足以下要求：

- a) 应支持特定子图像检测，特定子图像智能检测召回率不低于 95%，召回率应按附录 B.2.7 中公式计算；
- b) 应支持指定区域的特定子图像去除。

6.3.2 画面主体选择

全景视频画面主体选择技术满足以下要求：

- a) 应支持全景视频画面主体随用户头部和眼部运动跟随运动；
- b) 应支持画面主体移动或场景切换时，输出全景视频平稳不抖动。

6.3.3 字幕生成

全景视频字幕生成技术满足以下要求：

- a) 应支持视频中的语音自动生成字幕，语音自动生成字幕识别字错误率不高于 10%，字错误率应按附录 B.2.8 中公式计算；
- b) 应支持 ASS 或 SRT 字幕格式文件。

7 测试方法

全景视频增强处理技术测试基于测试源全景视频进行增强处理，通过测试系统验证相应技术功能和技术指标，测试设备及测试源配置见附录 A，测试环境及测试条件见附录 B.1。

7.1 全景视频画质提升

7.1.1 超分辨率

全景视频超分辨率技术测试步骤如下：

- a) 将附录 A.2 的 4K 转 25fps 高清源视频进行超分辨率，输出 UHD 超高清 4K (3840×2160) 目标视频，目标视频能保持图像细节清晰；
- b) 将附录 A.2 的 4K 转 25fps 高清源视频进行超分辨率，输出 UHD 超高清 8K (7680×4320) 目标视频，目标视频能保持图像细节清晰；
- c) 将附录 A.2 的 4K 转 25fps 高清源视频进行随机倍率点（2 倍、4 倍、8 倍、16 倍）超分辨率，输出对应倍率点分辨率目标视频，目标视频能保持图像细节清晰；
- d) 步骤 a) 输出的目标视频符合预期，则判定满足 2 倍超分辨率要求；
- e) 步骤 b) 输出的目标视频符合预期，则判定满足 4 倍超分辨率要求；
- f) 步骤 c) 输出的目标视频符合预期，则判定满足重建倍率可调节超分辨率要求。

7.1.2 插帧

全景视频插帧技术测试步骤如下：

- a) 将附录 A.2 的 4K 转 25fps 高清源视频进行插帧，输出 50fps 目标视频，目标视频能保持画面流畅、无明显闪烁抖动现象；
- b) 将附录 A.2 的 4K 转 15fps 高清源视频进行插帧，输出 60fps 目标视频，目标视频能保持画面流畅、无明显闪烁抖动现象；
- c) 将附录 A.2 的 4K 转 30fps 高清源视频进行超分辨率，输出 UHD 超高清 8K (7680×4320) 30fps 目标视频，进行插帧，输出 120fps 目标视频，目标视频能保持画面流畅、无明显闪烁抖动现象；
- d) 使用附录 A.1.2 的媒体分析工具检测步骤 a) 输出的目标视频，目标视频帧率为 50fps，则判定满足 2 倍插帧要求；
- e) 使用附录 A.1.2 的媒体分析工具检测步骤 b) 输出的目标视频，目标视频帧率为 60fps，则判定满足 4 倍插帧要求；
- f) 使用附录 A.1.2 的媒体分析工具检测步骤 c) 输出的目标视频，目标视频帧率为 120fps，则判定满足 UHD 超高清 8K (7680×4320) 全景视频适配 120FPS 要求。

7.1.3 噪声去除

全景视频噪声去除技术测试步骤如下：

- a) 将附录 A.2 的含噪声源视频进行噪去除, 输出目标视频, 使用附录 A.1.2 的媒体分析工具检测目标视频, 目标视频能保留画面细节, 提升主观清晰度, 无过锐现象;
- b) 比对目标视频和源视频, 图像质量画质提升主观评价的测试阶段、测试图像的显示和评分标度遵守 GY/T 340-2020 中 5.5、5.6、5.7 的规定。结果分析、结果说明遵守 GY/T 340-2020 中 5.8、5.9 的规定。用双刺激连续质量标度法得到源视频图像评分与目标视频图像评分, 计算画质提升率, 图像质量提升率应按附录 B.2.5 中公式计算;
- c) 针对同一含噪声源视频, 输出不同处理强度目标视频, 可进行人工选择, 则判定满足去噪强度可调节;
- d) 选取不同噪声强度源视频, 可进行自动智能去噪, 则判定满足自适应去噪。

7.1.4 偏色校正

全景视频偏色校正技术测试步骤如下:

- a) 将附录 A.2 的含偏色源视频进行偏色校正, 输出目标视频, 使用附录 A.1.2 的媒体分析工具检测目标视频, 目标视频更加真实清晰;
- b) 依照 7.1.3 b) 中的测试步骤比对目标视频和源视频, 计算画质提升率;
- c) 针对同一偏色源视频, 输出不同处理强度目标视频, 可进行人工选择, 则判定满足偏色校正强度可调节;
- d) 在可视化环境下进行偏色校正结果调整。

7.1.5 去模糊

全景视频去模糊技术测试步骤如下:

- a) 将附录 A.2 的 4K 转标清源视频进行去模糊, 输出目标视频, 使用附录 A.1.2 的媒体分析工具检测目标视频, 目标视频能保留画面细节, 提升辨识度;
- b) 依照 7.1.3 b) 中的测试步骤比对目标视频和源视频, 计算画质提升率;
- c) 计算目标视频大于 16×16 像素的模糊目标可识别率, 符合 6.1.5 b) 的要求, 可识别率应按附录 B.2.6 中公式计算。

7.1.6 去抖动

全景视频去抖动技术测试步骤如下:

- a) 将附录 A.2 的 4K 转标清源视频进行去抖动, 输出目标视频, 使用附录 A.1.2 的媒体分析工具检测目标视频, 目标视频能保留画面细节, 减少抖动程度;
- b) 依照 7.1.3 b) 中的测试步骤比对目标视频和源视频, 计算画质提升率;
- c) 使用附录 A.1.2 的媒体分析工具检测目标视频的视场角, 符合 6.1.5 b) 的要求。

7.1.7 HDR 转换

全景视频 HDR 转换技术测试步骤遵守 T/UWA 010-2022 中 7.1.4 的规定。

7.1.8 综合画质提升

对源视频经过超分辨率、插帧、噪声去除、偏色校正、去模糊、去抖动和 HDR 转换中一种或多种人工或智能增强处理技术处理后，输出目标视频。

- a) 依照 7.1.3 b) 中的测试步骤比对目标视频和源视频，计算画质提升率；
- b) 使用附录 A.1.2 的计算工具计算输出的目标视频和源视频的 VMAF 平均分；
- c) 使用附录 A.1.2 的计算工具计算输出的目标视频和源视频的平均 PSNR 和平均 SSIM，目 PSNR 应按附录 B.2.1 中公式计算，SSIM 应按附录 B.2.3 中公式计算；
- d) 使用附录 A.1.2 的计算工具计算输出的目标视频 ERP 投影格式和源视频 ERP 投影格式的平均 WS-PSNR 和平均 WS-SSIM，WS-PSNR 应按附录 B.2.2 中公式计算，WS-SSIM 应按附录 B.2.4 中公式计算；
- e) 经过增强处理技术处理后的目标全景视频，满足 6.1.8 a)、b)、c)、d) 四项要求中的任意三项，则判定满足全景视频综合画质提升要求。

7.2 全景视频投影变换及视点选择

7.2.1 投影变换

全景视频投影变换技术测试步骤如下：

- a) 将三维球形源视频进行二维 ERP、EAP、CMP、ISP、OHP、TSP 或 SSP 格式投影变换，输出目标视频；
- b) 对目标视频进行 7.1 中各项增强处理，符合 7.1.8 中的综合画质提升要求；
- c) 使用附录 A.1.2 的媒体分析工具检测目标视频，目标视频未出现明显模糊、扭曲、卡顿、抖动和重影现象，则判定满足投影变换失真较小。

7.2.2 投影逆变换

全景视频投影逆变换技术测试步骤如下：

- a) 将二维 ERP、EAP、CMP、ISP、OHP、TSP 或 SSP 格式源视频进行投影变换，输出三维球形目标视频；
- b) 使用附录 A.1.2 的媒体分析工具检测目标视频，目标视频未出现明显模糊、扭曲、卡顿、抖动和重影现象，则判定满足投影变换失真较小。
- c) 使用附录 A.1.2 的媒体分析工具检测目标视频，全景视频投影逆变换后边界处未出现明显接

缝，未出现明显块效应，则判定满足拼接对齐要求。

7.2.3 视点选择

全景视频视点选择技术测试步骤如下：

- a) 对附录 A.2 的 4K 转 25fps 高清源视频进行高频观看视点选择划分，输出目标区域视频；
- b) 对目标区域视频进行 7.1 中各项增强处理，符合 7.1.8 中的综合画质提升要求。
- c) 使用附录 A.1.2 的媒体分析工具检测目标视频，目标视频细节清晰。

7.3 全景视频编辑

7.3.1 特定子图像去除

全景视频特定子图像去除技术测试步骤如下：

- a) 将源视频进行特定子图像去除，输出目标视频；
- b) 对比源视频和目标视频，记录水印检测和去除数据；
- c) 计算特定子图像检测交并比，大于 0.5 为检测正确，交并比应按附录 B.2.9 中公式计算；
- d) 结合步骤 c) 的检测结果，计算特定子图像检测召回率，符合 6.3.1 a) 的要求，召回率应按附录 B.2.7 中公式计算。

7.3.2 画面主体选择

全景视频画面主体选择技术测试步骤如下：

- a) 结合主观头部和眼部运动数据，输出目标区域视频，目标区域视频画面主体应完整，无闪烁抖动；
- b) 画面主体移动或场景切换时，目标区域视频应自动切换主体，平稳无抖动。

7.3.3 字幕生成

全景视频字幕生成技术测试步骤如下：

- a) 将源视频进行语音自动生成字幕，并生成 ASS 或 SRT 格式的字幕；
- b) 对比语音自动生成字幕信息与源视频字幕信息，计算语音自动生成字幕字错误率，符合 6.3.4 a) 的要求，字错误率应按附录 B.2.8 中公式计算。

附录 A

(规范性)

测试配置

A.1 测试系统配置

A.1.1 硬件配置

全景视频增强处理技术测试基于显示器、头戴式设备、计算机和服务器等硬件。显示器遵守 GY/T 340-2020 表 2 的规定。头戴式设备遵守 GB/T 38259-2019 的规定。计算机遵守 GB/T 9813.1、GB/T 9813.2 或 GB/T 9813.3 的规定。

A.1.2 软件配置

全景视频增强处理技术测试基于媒体分析工具、客观指标计算工具等软件。媒体分析工具示例：如 MediaInfo 等媒体分析软件。客观指标计算工具示例：如 FFmpeg 等客观指标计算软件。

A.1 测试源配置

全景视频增强处理技术测试需要源全景视频，由 4K 源全景视频制作 4K 转 25fps 高清、4K 转 15fps 高清、4K 转 30fps 高清、含噪声、含偏色、含模糊、含抖动、特定子图像和字幕等子测试数据集。4K 源全景视频遵守 T/UWA 010-2022 附录 C 的规定。

特别地，对于全景视频，人眼实际观看时聚焦于空间中的部分感兴趣视点区域，需对其进行选取划分，制作全景视频高频视点序列。高频视点序列配置示例：如基于 SURF 局部特征检测高频视点区域生成全景视频高频视点序列。

附录 B

(规范性)

技术指标

全景视频增强处理技术的测试包括主观评价和客观测试，主观评价和客观测试关联互补。主观评价基于人眼主观对全景视频的感受给出全景视频质量评价，评价经过增强处理前后的全景视频主观质量提升。客观测试验证增强处理功能实现和全景视频各项客观评价指标。

B.1 主观评价

全景视频经过增强处理前后主观评价的测试阶段、测试图像的显示和评分标度遵守 GY/T 340-2020 中 5.5、5.6、5.7 的规定。全景视频主观评价配置示例：参考主观评价人员共 10 人，分 10 组进行评价，每组 1 人，评价观看条件如表 B.1 所示。

表 B.1 主观评价观看条件

序号	项目	4K UHD TV (SDR) 参数值
1	观看距离	1.6 倍图像高度
2	监视器屏幕的峰值亮度	102.4 cd/m ²
3	监视器屏幕黑场亮度与峰值亮度之比	<0.0001
4	背景亮度与峰值亮度之比	<0.005
5	室内环境光照明	低
6	背景光光源色温	D ₆₅
7	观看员的座位布局	水平方向在中垂线±30° 内

B.2 客观测试

全景视频经过增强处理前后客观测试按本节技术指标公式进行计算比较。

B.2.1 PSNR

PSNR 按公式 (B.1) (B.2) 计算，数值越大图像失真越小。

$$MSE = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W (X(i,j) - Y(i,j))^2 \quad (B.1)$$

$$PSNR = 10 \log_{10} \left(\frac{MAX^2}{MSE} \right) \quad (B.2)$$

式中：

MSE ——测试图像 X 和参考图像 Y 的均方误差；

H, W ——图像的高度和宽度；

MAX ——图像数据类型动态范围最大值，对于 uint8 类型图像数据，取值范围是[0, 255]，即 $MAX=255$ 。

B. 2. 2 WS-PSNR

WS-PSNR 按公式 (B. 3) (B. 4) (B. 5) 计算，数值越大图像失真越小，在 PSNR 的基础上基于公式 (B. 3) 进行纬度相关的加权。

$$L_{i,j} = \cos((i - (H/2) + 0.5)\pi/H) \quad (B. 3)$$

$$WS_{MSE} = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W L_{i,j} (X(i,j) - Y(i,j))^2 \quad (B. 4)$$

$$WS_{PSNR} = 10 \log_{10} \left(\frac{MAX^2}{WS_{MSE}} \right) \quad (B. 5)$$

式中：

$L_{i,j}$ ——图像 (i,j) 坐标位置的权重值；

WS_{MSE} ——测试图像 X 和参考图像 Y 的均方误差；

H, W ——图像的高度和宽度；

MAX ——图像数据类型动态范围最大值，对于 uint8 类型图像数据，取值范围是[0, 255]，即 $MAX=255$ 。

B. 2. 3 SSIM

SSIM 计算时首先将整张图像划分为像素块，像素块大小一般为 11×11 ，每个像素块按公式 (B. 6) (B. 7) (B. 8) (B. 9) (B. 10) 计算对应 SSIM，最后所有像素块 SSIM 值取平均，取值范围是[0, 1]，数值越大图像失真越小。

$$\mu_X = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W X(i,j) \quad (B. 6)$$

$$\sigma_X^2 = \frac{1}{H \times W - 1} \sum_{i=1}^H \sum_{j=1}^W (X(i,j) - \mu_X)^2 \quad (B. 7)$$

$$\sigma_{XY} = \frac{1}{H \times W - 1} \sum_{i=1}^H \sum_{j=1}^W (X(i,j) - \mu_X)(Y(i,j) - \mu_Y) \quad (B. 8)$$

$$l(X,Y) = \frac{2\mu_X\mu_Y + C_1}{\mu_X^2 + \mu_Y^2 + C_1}, c(X,Y) = \frac{2\sigma_X\sigma_Y + C_2}{\sigma_X^2 + \sigma_Y^2 + C_2}, c(X,Y) = \frac{\sigma_{XY} + C_3}{\sigma_X\sigma_Y + C_3} \quad (B. 9)$$

$$SSIM(X, Y) = l(X, Y) \cdot c(X, Y) \cdot s(X, Y) \quad (B.10)$$

式中：

H, W ——像素块的高度和宽度；

μ_X, μ_Y ——测试像素块 X 和参考像素块 Y 的均值；

σ_X, σ_Y ——测试像素块 X 和参考像素块 Y 的方差；

σ_{XY} ——测试像素块 X 和参考像素块 Y 的协方差；

C_1, C_2, C_3 ——常数，为了避免分母为 0，通常取 $C_1 = (K_1 \cdot L)^2$ ， $C_2 = (K_2 \cdot L)^2$ ， $C_3 = 0.5C_2$ ，一般地， $K_1=0.01$ ， $K_2=0.03$ 。

B.2.4 WS-SSIM

WS-SSIM 计算时首先将整张图像划分为像素块，像素块大小一般为 11×11 ，每个像素块按公式 (B.11) (B.12) (B.13) (B.14) (B.15) (B.16) 计算对应 SSIM，最后所有像素块 SSIM 值取平均，取值范围是 $[0, 1]$ ，数值越大图像失真越小，在 SSIM 的基础上基于公式 (B.11) 进行纬度相关的加权。

$$L_{i,j} = \cos((i - (H/2) + 0.5)\pi/H) \quad (B.11)$$

$$\mu_X = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W L_{i,j} X(i, j) \quad (B.12)$$

$$\sigma_X^2 = \frac{1}{H \times W - 1} \sum_{i=1}^H \sum_{j=1}^W L_{i,j} (X(i, j) - \mu_X)^2 \quad (B.13)$$

$$\sigma_{XY} = \frac{1}{H \times W - 1} \sum_{i=1}^H \sum_{j=1}^W L_{i,j} (X(i, j) - \mu_X)(Y(i, j) - \mu_Y) \quad (B.14)$$

$$l(X, Y) = \frac{2\mu_X\mu_Y + C_1}{\mu_X^2 + \mu_Y^2 + C_1}, c(X, Y) = \frac{2\sigma_X\sigma_Y + C_2}{\sigma_X^2 + \sigma_Y^2 + C_2}, s(X, Y) = \frac{\sigma_{XY} + C_3}{\sigma_X\sigma_Y + C_3} \quad (B.15)$$

$$WS-SSIM(X, Y) = l(X, Y) \cdot c(X, Y) \cdot s(X, Y) \quad (B.16)$$

式中：

$L_{i,j}$ ——像素块 (i, j) 坐标位置的权重值；

H, W ——像素块的高度和宽度；

μ_X, μ_Y ——测试像素块 X 和参考像素块 Y 的均值；

σ_X, σ_Y ——测试像素块 X 和参考像素块 Y 的方差；

σ_{XY} ——测试像素块 X 和参考像素块 Y 的协方差；

C_1, C_2, C_3 ——常数，为了避免分母为 0，通常取 $C_1 = (K_1 \cdot L)^2$ ， $C_2 = (K_2 \cdot L)^2$ ， $C_3 = 0.5C_2$ ，一般地， $K_1=0.01$ ， $K_2=0.03$ 。

B.2.5 图像质量提升率

图像质量提升率 P 按公式(B. 17)计算。

$$P = \frac{b - a}{b} \quad (\text{B. 17})$$

式中：

a, b ——源全景视频和增强处理后全景视频图像主观评价得分。

B. 2.6 可识别率

可识别率 R 按公式(B. 18)计算。

$$R = \frac{T}{T + F} \quad (\text{B. 18})$$

式中：

T ——正确识别的目标对象数目；

F ——错误识别和未识别到的目标对象数目。

B. 2.7 召回率

召回率 $Recall$ 按公式(B. 19)计算。

$$Recall = \frac{TP}{TP + FN} \quad (\text{B. 19})$$

式中：

TP ——检测正确的目标对象数目；

FN ——未检测正确的目标对象数目。

B. 2.8 字错误率

字错误率 W 按公式(20)计算。

$$W = \frac{S + D + I}{S + D + I + C} \quad (\text{B. 20})$$

式中：

S ——被错误替换的字数；

D ——未识别到的字数；

I ——识别出的验证数据中不存在的字数；

C ——识别正确的字数。

B. 2.9 交并比

交并比 IoU 按公式(21)计算。

$$IoU = \frac{A \cap B}{A \cup B} \quad (\text{B. 21})$$

式中：

A, B ——目标对象和非目标对象面积;

参考文献

- [1] 5G高新视频-沉浸式视频技术白皮书（2020）[J]. 国家广播电视总局科技司，2020.
-