

ICS: 33.160.01

CCS: 16513

# 团 体 标 准

T/GDIOT 025—2024

## 基于神经网络的全景视频超分辨率重建质量 评价标准

Quality evaluation standard for super-resolution reconstruction of  
panoramic video based on neural network

2024-12-23发布

2024-12-23实施

广东省物联网协会 发布

## 目 录

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 客观评价 .....	2
5.1 一般要求 .....	2
5.2 评价指标 .....	2
6 主观评价 .....	3
6.1 一般要求 .....	3
6.2 评分方法和评分标度 .....	5
6.3 评价指标 .....	6

## 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。本文件的发布机构不承担识别专利的责任。

本文件由广东省物联网协会归口。

本文件起草单位：广东恒电信息科技股份有限公司、中山大学、南方医科大学、珠江水利委员会珠江水利科学研究所、北京师范大学、暨南大学、广东开放大学、广东工业大学、广东技术师范大学、广东海洋大学、岭南师范学院、广东第二师范学院、澳门科技大学、华南理工大学、广州卫生职业技术学院、广州番禺职业技术学院、广东理工职业学院、广州中医药大学、广东药科大学、广东省科技基础条件平台中心、佛山市中医院。

本文件主要起草人：高静、吴迪、戴少锋、杨晴宇、高钰芯、冯丙文、刘磊、陈炳丰、黄聃、雷方元、彭小红、余应淮、杨俊杰、邬依林、赵庆林、王和勇、吴力挽、汤非易、陈靖宇、汪静、陈浪城、张立泓、门韶洋、杨进、滕璐瑶、卢琰、龙超、关锦图、陈永恒、莫纯宇、尹钟萱。

# 基于神经网络的全景视频超分辨率重建质量评价标准

## 1 范围

本文件规定了基于神经网络的全景视频超分辨率质量的客观及主观评价标准。

本文件适用于对利用神经网络进行超分辨率重建的全景视频的质量评价。

## 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。

GY/T 134—1998 数字电视图像质量主观评价方法

DB11/T 384.5—2018 图像信息管理系统技术规范 第5部分：图像质量要求与评价方法

GY/T 340—2020 4K超高清视频图像质量主观评价用测试图像

## 3 术语和定义

下列术语和定义适用于本文件。

### 3.1

#### 客观评价 objective assessment

专业检测单位遵循规定的检测程序，按照规定的检测条件、项目和指标，使用专用测试设备，对图像质量进行物理的定量评价。

[来源： DB11/T 384.5—2018， 3.1.5]

### 3.2

#### 主观评价 subjective assessment

特定人员群组在设定的条件下，按判别项目和分级要求，以个体视觉感观对图像质量进行主观的定性评价。

[来源： DB11/T 384.5—2018， 3.1.4]

### 3.3

### 信噪比 Signal to Noise Ratio

表示信号功率与噪声功率之间的比值，通常以分贝（dB）为单位表示。

#### 4 缩略语

下列缩略语适用于本文件。

CPU：中央处理器（Central Processing Unit）

ARM：进阶精简指令集机器（Advanced RISC Machine）

SSIM：结构相似性指标（Structural similarity index measure）

JVET：国际标准组织联合视频探索小组（Joint Video Exploration Team,）

VRU：国内标准组织虚拟现实联盟（Virtual Reality Unity, VRU）

#### 5 客观评价

##### 5.1 一般要求

###### 5.1.1 数据集准备

应准备高质量的原始视频数据作为参考，以及经过超分辨率处理后的低分辨率视频数据。

###### 5.1.2 硬件平台

CPU主频大于2.5GHz、内存大于16GB的x86、x64、ARM或兼容架构计算机。

###### 5.1.3 操作系统

Windows、Linux、麒麟或其他兼容操作系统。

##### 5.2 评价指标

###### 5.2.1 峰值信噪比

应在球面上选取大量的均匀分布采样点，通常是655362个点。将采样点重新投影到参考图像和失真图像上，计算它们的PSNR值。再取平均求得S-PSNR，计算见式（1），（2）。

S-PSNR>40dB者，判为该项合格。

$$PSNR_i = 10 \times \log_{10} \left( \frac{MAX_i^2}{MSE_i} \right) \dots\dots\dots(1)$$

$$S - PSNR = \frac{1}{n} \sum_i^n PSNR_i \dots\dots\dots(2)$$

式子中：

$MAX_i$ 是测试序列中第*i*个采样点像素值的最大可能值；

$MSE_i$ 是原始序列第*i*帧和被测序列第*i*个采样点之间的均方误差；

$PSNR_i$ 是测试序列第*i*个采样点的峰值信噪比；

S-PSNR是测试序列的平均峰值信噪比。

### 5.2.2 结构相似性指数

应先计算序列的像素均值和标准差，再计算结构相似性指数SSIM，计算见式（3）。

SSIM>0.9者，判为该项合格。

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \dots\dots\dots(3)$$

式子中：

$\mu_x$  和  $\mu_y$  分别是原始序列x和被测序列y的像素均值；

$\sigma_x$  和  $\sigma_y$  分别是原始序列x和被测序列y的像素标准差；

$\sigma_{xy}$  是原始序列x和被测序列y的像素协方差；

$C_1$  和  $C_2$  是常数，用于稳定分母。

## 6 主观评价

### 6.1 一般要求

#### 6.1.1 观看条件

全景视频超分辨率重建质量的主观评价应该在以下所给出的观看条件下进行：

- a) 视场：设置小型绿色标志实时提醒用户视场中心，避免注意力偏移引发的噪声数据；
- b) 沉浸式环境：使用头戴式显示设备来提供全方位的视觉体验，确保用户在观看过程中不受到外界干扰。

#### 6.1.2 信号源

信号源直接提供基准视频序列信号，作为被测模型的输入，经过被测模型处理后，形成被测序列。

#### 6.1.3 测试素材

在进行序列制作测试时，应选用不少于20个全景视频。这些视频的来源可以是JVET、

VRU或其他知名视频平台。选定的视频应包含多种场景，以确保时空域动态范围的广泛性。

#### 6.1.4 观看员

观看员应是应邀参加主观评价的评分人员。观看员通常分为两种类型，即专业观看员和非专业观看员。一般由非专业观看员进行主观评价，当需要精确判断时，可由受过专业训练的专业观看员来进行评价和分析。

#### 6.1.5 评价测试阶段

在每个评价测试阶段开始时，应向观看员详细、正确地介绍评价方法和存在的质量损伤类型，并进行评价示范显示。每个观看员的观看时间应控制在30分钟以内，以避免疲劳和眩晕。

在正式测试开始前，应引入3至5个评价序列来稳定观看员的判别力，其结果数据不纳入测试结果的统计中。不同测试序列的显示顺序采用伪随机方式。为了检测相关性，有的测试可以重复进行，但要避免相同测试图像在相继的评价序列中出现。

评价周期的显示流程见图1。

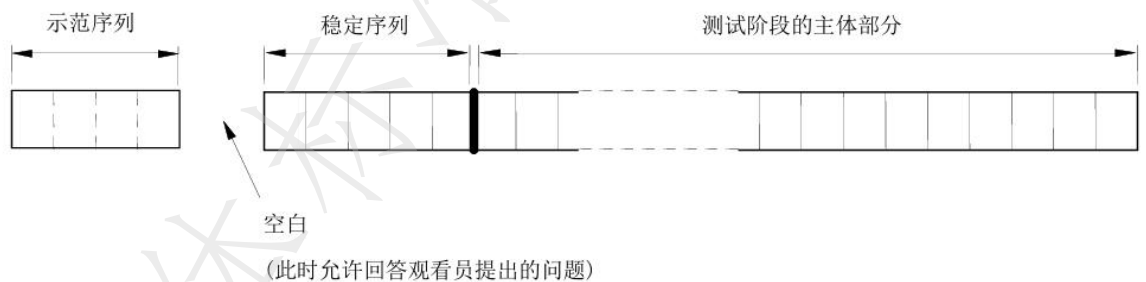


图1 评价周期的显示流程

#### 6.1.6 结果表达

对于每个测试图像，应给出评分值统计分布的平均值、标准偏差和95%置信度区间。

结果还必须包含以下的信息：

- a) 测试情况的说明；
- b) 测试素材的情况；
- c) 图像源的类型和显示监视器型号；
- d) 观看员的人数和类型；
- e) 使用的基准系统；
- f) 测试总的平均分；
- g) 如淘汰了某些观看员，要说明原始的和调整后的平均分和95%置信度区间。

## 6.2 评分方法和评分标度

### 6.2.1 评分方法

在双刺激连续质量标度法中，要求观看员观看每个测试序列的两种状态，即基准状态（直接来自信号源）和被测状态（经被测模型）的序列总体质量进行评价，并在评分表上作出标记。

### 6.2.2 评分标度

观看员应在垂直标尺上标出记号来确定每次演示视频的的总体质量。垂直标尺是成对的，对应每个测试视频的两次演示。为了防止量化误差，标尺提供了连续的评分机制，并分成了长度相等的5段，对应优、良、中、差、劣5个等级。图2给出了典型评分标度。

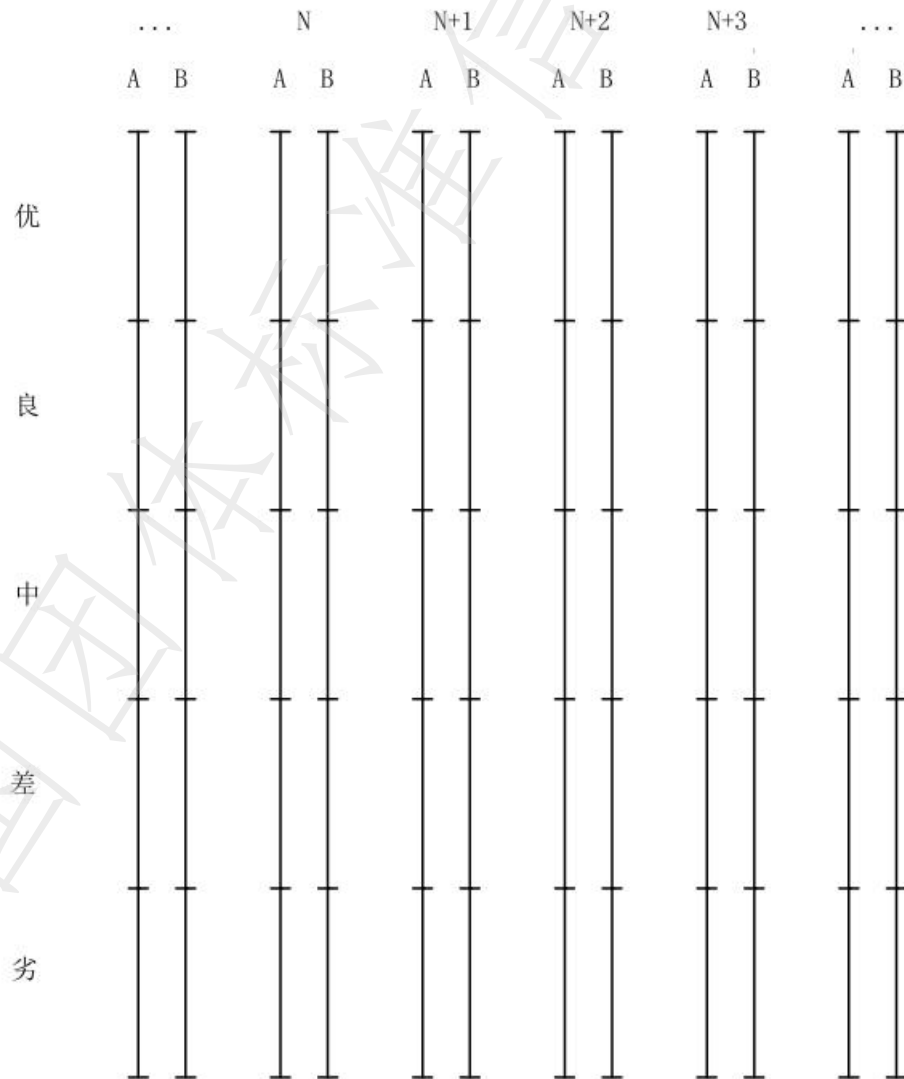


图2 典型评分标度

### 6.3 评价指标

#### 6.3.1 评分量化

将每个测试条件下的基准序列和被测序列从评分标度上的度量长度转换为0至100范围内的评分，然后计算基准序列与被测序列之间的差值。

#### 6.3.2 平均分计算

对评分结果进行分析的第一步是计算每一测试序列的平均评分，见式（4）：

$$\bar{u}_{jkr} = \frac{1}{N} \sum_i^N u_{ijk} \dots\dots\dots(4)$$

式中：

$u_{ijk}$ ——观看员*i*在测试条件*j*、测试序列*k*、重复*r*次情况下的评分；

$N$ ——观看员数量；

类似的，可计算出每一测试序列的总平均评分 $\bar{u}_k$ 。

#### 6.3.3 质量提升率

经质量重建后的全景视频质量提升率应高于20%，质量提升率计算见式（5）：

$$E = \left( \frac{b-a}{a} \right) \times 100\% \dots\dots\dots(5)$$

式中：

$a$ 表示基准序列主观质量评分；

$b$ 表示被测序列主观质量评分。