

面向机器视觉安防摄像机图像 质量评测方法研究报告



中国安全防范产品行业协会

通讯地址：北京市海淀区西三环北路87号国际财经中心C座1401

联系电话：010-68730250

网址：<http://www.21csp.com.cn>

公安部安全与警用电子产品质量检测中心

通讯地址：北京市海淀区首都体育馆南路一号（北京市2808信箱47分箱）

联系电话：010-68773375、68773780、68773591

网址：www.tcsbj.com；www.ga.net.cn/jczx

中国安全防范产品行业协会
公安部安全与警用电子产品质量检测中心

目录

1 安防摄像机成像技术趋势及质量评价体系的演进	02
1.1 安防摄像机成像技术趋势	02
1.2 图像质量评价体系的演进	05
2 面向机器视觉的图像质量评价及分析方法	08
2.1 整体质量评价	08
2.2 图像质量分析维度	09
2.3 基于机器视觉任务的图像评价与图像质量维度关系	17
3 评测场景构建	20
3.1 安防摄像机关键应用场景	20
3.2 室外场景室内化设计	21
3.3 测试设备	22
3.4 室内模拟场景	24
4 测试流程及评价方法	29
4.1 测试流程	29
4.2 评分方法	31

1 安防摄像机成像技术趋势及质量评价体系的演进

1.1 安防摄像机成像技术趋势

1.1.1 摄像机从给人看到给机器看为主

随着 AI 技术、算力和数据量的快速发展，使得大量视频数据的“受众”已不再是人，而是机器和算法。通过人脸识别、车辆识别、行人再识别 ReID、行为分析等算法，使得视频中嵌入的信息发挥了越来越大的价值。安防行业正在从“给人看”转变成“给机器看”为主。

“给机器看”可以实现海量视频的快速分析，大幅提升效率及发现更多目标细节。对于摄像机拍摄的海量数据，机器处理将是后续发展的唯一选择，因为单纯依靠人力识别需要海量的人力投入且效率低下，而机器算法可以 7×24 小时全天不间断分析，大幅减少操作人员浏览视频数据的时间，快速自动定位敏感信息并进行结构化的存储，减少疲劳等原因造成的审查疏忽，从而极大的提升效率；同时，依靠算法模型的优化和海量数据训练的优势，智能算法已经表现出超越人眼的性能，在目标检测、目标识别等多项指标上都达到新的高度，因此利用机器可以发现更多人眼忽略的细节，找出更多隐藏的目标。

在这种背景下，摄像机系统的整体设计也在从面向人眼视觉到面向机器视觉任务的系统进行重构，覆盖从采集成像，图像 ISP，水印增强及编码增强关键端到端环节，包括针对机器视觉特性的光学系统和传感器特殊设计、针对机器视觉关注目标区域的图像增强、编码增强等。



E2E 环节	给人看到给机器看的架构变化
采集环节	镜头及 sensor 从可见光到多光谱支持 大幅提升红外 QE 性能 宽动态技术
图像 ISP	从简单面向人眼视觉到面向机器视觉任务增强图像质量 现有 ISP 大部分环节都是面向人眼视觉的，需要针对机器视觉识别率重构或增强
编码	从传统面向人眼感知模型的编码到面向机器感知模型的编码。传统面向人眼感知模型的编码降低识别率，需要面向机器感知模型的编码保证识别率不下降。
质量评价	从面向人眼到面向机器视觉任务，从主观到客观

1.1.2 AI 算力换图像，提升目标图像质量

自从 Intel 发表了“learning to see in the dark”的论文后，学术界及业界关于 AI 算力应用与提升图像质量的研究如火如荼。这篇论文中提出了一个大的方向，就是 AI ISP，论文中提出一个模型，一个以 RAW bayer 为输入，RGB 输出为输出，论文主要关注图像噪声，示例中噪声被很好的去除，尤其是彩色噪声部分让大家有深刻印象，在消除噪声方面取得了很好的效果。2020 年，苏黎世联邦理工学院（ETHZ）发表了一篇论文“Replacing Mobile Camera ISP with a Single Deep Learning Model”，其中提出一个新的算法 PyNet，只需单个端到端深度学习模型，就能替代手机的 ISP。它调教出的相机算法可以从一个手机移植到另一个手机上，而不必使用手机的 ISP 更是将 AI ISP 推向了一个新的高度。如下图是一个采用 AI ISP 一个采用传统 ISP 的差异。



大家可以看出 AI ISP 图像效果在噪声、颜色、清晰度等方面都有了大幅提升。但该方案的一个关键问题是需要大量的 AI 算力及内存带宽。无疑，这方面芯片公司会走在前面，因为他们可以结合芯片算力及硬件设计实现 ISP 的部分或全部 AI 化，消除这些关键瓶颈。

当然，在芯片实现之前，结合现有芯片的有限算力及内存带宽，很多厂商采用了降低算力及带宽的方式比如针对部分帧或针对目标 ROI 区域进行 AI 图像增强，提升图像效果从而提升目标识别率也有了很好的实践。

1.1.3 高质量成像与环保无光污染兼顾

最近几年，随着人们对于环保要求越来越高，越来越多的人注意到身边的监控摄像机，特别是因为夜间的光污染问题会导致人们的反感。为了获得高质量的图像质量，摄像机需要配备足够强的补光灯，但在夜间的黑暗背景下人眼对强光非常敏感，运动进入监控拍摄范围时因为适应时间不够，会引起明显不适，甚至会产生眩晕和短时致盲，威胁人身健康和道路安全。



夜间光污染严重

750nm 补光也存在一定光污染

针对强烈的用户需求，摄像机行业也在快速推出相应的技术。总结起来有这么四种：一是部分摄像机改进补光方式，通过频闪补光、脉冲补光、补光透镜优化等方式减少同量级补光下对人眼的刺激程度；二是摄像机通过大光圈加高感度 sensor 来实现；三是部分摄像机通过 AI 来提升图像质量；四是采用多

光谱的方式，同时获取彩色图像及高信噪比的红外图像，通过图像融合的方式得到高质量的彩色图像。红外补光不会被人感知，从而可以做到“零光污染”。

综上，在新的技术趋势下，安防摄像机成像从只给人看到给机器看为主，同时环保无光污染也是关键技术趋势。图像质量评价也迫切需要对现有以给人主观评价为主的进行变革，结合摄像机成像技术针对给机器看制定新的评价标准体系。

1.2 图像质量评价体系的演进

在监控摄像机行业的图像评测领域，近年主要使用的行业标准为《GA/T 1127-2013 安全防范视频监控摄像机通用技术要求》和《GA/T 1128-2013 安全防范视频监控高清晰度摄像机测量方法》，里面图像测试主要包括分辨力、图像尺寸、最大亮度鉴别等级、最低可用照度、照度适应范围、亮度信号信噪比、色彩还原误差等。这些指标主要是直接计算一些特定的物理参数。通常是在较好的光照下，对二维图卡进行测试分析。这种方式与实际场景的光照、拍摄对象都不相符，不能很好地反映真实使用场景下的图像质量。同时也不能反映机器视觉角度的图像质量。特别地，很多安防摄像机针对目标如人脸、车、车牌等重点进行图像质量增强，按照传统图卡等方式无法真实评价。

很多读者会问，安防摄像机能否采用手机领域的 DXO 标准来评价。我们知道 Dxomark 按照视觉感知模型对成像评价指标权重进行设定，评价体系中还加入部分主观评测分数，给人看为主而不是机器看。同时其评价对象手机拍摄的主要目标和安防摄像机的目标要求也存在很大差异。

其实，结合机器视觉任务进行图像质量评价最近几年也在快速成熟。大家知道在自动驾驶领域，业界开始制定图像质量的主观和客观测试方法，即给人看和给机器看的标准体系。国际电气电子工程师学会在 2018 年发布白皮书《IEEE P2020 Automotive Imaging White Paper》，分为给人看的标准 IQ for viewing standard 和给机器看的标准 IQ for Computer vision standards。在给机器看的标准里，建议将图像采集系统作为一个整体来评价。当前针对机器视觉考虑的 KPI 评价指标包括基于检出目标物的评价 Contrast detection probability (CDP) 和基于颜色区分度的评价 Color separation probability (CSP)。

针对安防关键应用场景，结合机器视觉任务进行图像质量评价是当前的一个关键方向。

1.3 面向机器视觉任务的图像质量评价概述

研究面向机器视觉任务的图像质量评价方法，对改进面向特定机器视觉任务的成像能力，有着非常重要的作用和指导价值。不同的机器视觉任务对图像质量的要求各有侧重，如人脸识别任务，对图像的清晰度、对比度等有较高的要求，对颜色偏差等要求相对较低；而行人再识别任务，则对图像的颜色表达的准确性具有较高的需求。

本文重点研究面向人脸识别任务的图像质量评价方法。

我们知道，人脸识别性能容易受到光照、姿态、表情、分辨率、遮挡，年龄等因素的影响 [1]。这些因素可以分为内部因素和外部因素。内部因素是指人脸生理变化产生的图像差异，如年龄变化等；外部因素与环境交互有关，如光照、姿态、表情等；根据与摄像机成像能力的关系，又可以进一步将外部因素分为与摄像机成像能力相关和不相关的因素，如光照、分辨率、清晰度等与摄像机成像能力直接相关，而姿态、遮挡等因素，更多的是与人脸与摄像机的相对位置有关，与摄像机成像能力关联并不紧密。本文旨在研究面向机器视觉任务的摄像机成像能力，因此重点考虑与摄像机成像能力直接相关的质量因素（如清晰度、噪声、亮度、色彩饱和度等），而忽略其他外部因素的影响。

人脸图像相似度是进行人脸识别任务的重要依据。粗略来说，人脸图像相似度可认为由两部分组成，一是人脸身份的相似度，二是图像质量的相似度（见公式 1-1）。

$$S(I_1, I_2) = S_{id}(I_1, I_2) \cdot S_{quality}(I_1, I_2) \quad (1-1)$$

其中 $S(I_1, I_2)$ 是图像 I_1 和 I_2 的整体相似度（即人脸识别任务的相似度）， $S_{id}(I_1, I_2)$ 是图像 I_1 和 I_2 的身份信息相似度， $S_{quality}(I_1, I_2)$ 是图像 I_1 和 I_2 的质量相似度。由于我们评价的是图像质量，那么采用人脸相似度评价与图像质量评价是否等价呢？当两张人脸图像身份越相似，图像质量越相似，则整体相似度越高。当两张图像来自于同一个人的时候， $S_{quality}(I_1, I_2)$ 为 1，此时 $S(I_1, I_2) = S_{quality}(I_1, I_2)$ ，即图像识别任务相似度与图像质量相似度一致。因此，若采用图像整体相似度对类内样本（即图像来自于同一个人）的质量相似度进行描述，即可保证图像质量评价与图像识别任务的一致性，实现面向机器视觉任务的图像质量评价方法。

Bowyer 等人 [4] 对不同质量图像对的相似度分数进行了统计分析。如图 1.1 所示，图中横坐标是图像对的相似度，纵坐标表示的样本出现概率。从右图中可以看出，类间相似度都比较低，且不同质量的图像对类间相似度分布并无明显差别。而不同质量图像的类内相似度分数则有明显的差异（如左图所示）。质量高的图像类内相似度相对较高（绿色曲线），质量差的图像类内相似度相对偏低（蓝色曲线），质

量最差的图像类内相似度最低（红色曲线）。因此，采用类内图像对的相似度能够对图像样本质量做较好的区分。

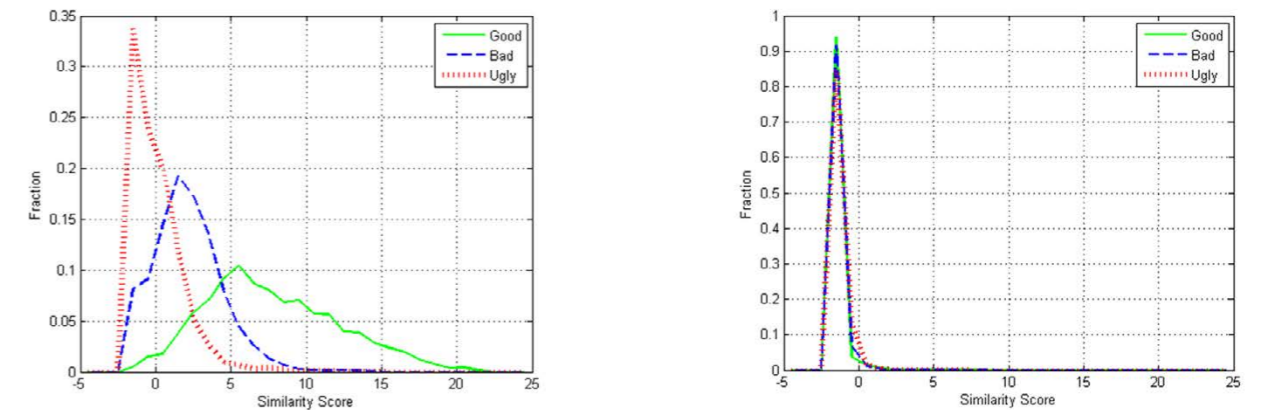


图 1.1 不同质量的类内样本相似度分布（左）和类间样本相似度分布（右），摘自文献 [4]

2 面向机器视觉的图像质量评价及分析方法

常见的图像质量评价方法可以分为全参考方法、半参考方法和无参考方法 [6]。文献 [4] 中指出，人脸图像质量差异来自于图像对，而非单张图像本身，而本方法需要对单张人脸图像质量进行打分。借鉴全参考方法，提出一种类全参考的图像质量评价方法，引入标准参考图像，将待评价图像质量与标准参考图像作比较，得到该图像的质量评价分数，将图像对的差异转换成单张图像的质量评价。因标准参考图像与待评价图像不是严格逐像素对齐，为与全参考方法做区别，故称之为“类全参考方法”。

本方法拍摄画面包含多张人脸图像，因此需要从拍摄图像中抠取所需人脸图像作为评价时的目标图像。图 2.1 为人脸目标图像示例。



图 2.1 人脸目标图像示意图

2.1 整体质量评价

如第一章所述，本方法采用人脸识别任务的类内人脸特征相似度分数作为图像整体质量评价基础。计算待评价图像与参考图像的人脸特征相似度，即得到待评价图像的整体质量评价。

各种人脸识别算法由于训练数据分布、权重选择等各方面原因，其对于同一副图像评价也不同，和主观评价类似也有算法自己的评价喜好。实操上，建议结合几家的算法一起对图像进行评价，相对客观。ArcFace 开源人脸识别算法 [7] 是最近业界比较好的开源算法，本文采用它来计算人脸特征相似度进行演示说明，评测步骤如下：

- 1) 使用 ArcFace 开源人脸识别模型，计算目标图像的人脸特征向量 $\bar{\epsilon}$ 。
- 2) 按公式 (2-1) 计算待评价目标图像与对应参考目标图像的人脸特征相似度 S 。对应参考目标图像是指与待评价目标图像拍摄时位置最接近的参考目标图像。

$$S = \frac{\bar{\epsilon}_1 \cdot \bar{\epsilon}_2}{\|\bar{\epsilon}_1\| \|\bar{\epsilon}_2\|} \quad (2-1)$$

式中：

ϵ_1 ——表示对应参考目标图像的人脸特征

ϵ_2 ——表示待评价目标图像的人脸特征

需要说明的是，人脸识别算法存在距离不对齐问题。也就是说，系统对有些人的人脸特征学习得很好，所以类间距很小，而对于那些学得不太好的人，类间距就会很大。为了减小波动，提升用户体验，在实践上，需要结合误报率指标进行距离修正，降低方差，从而显著的提升识别算法的准确度。

2.2 图像质量分析维度

基于类内人脸特征相似度的评价方法由于神经网络的黑盒特点，其评价结果很难对图像质量调优起到指导意义，这个也是 AI 领域的关键挑战之一。在本文中我们尝试结合常用图像质量评价因子对图像质量进行更细粒度的分析，并与基于类内人脸特征相似度的评价方式关联，从而对图像质量调测提供一定意义上的指导。

在图像质量指标的选择上，公安行业标准《安全防范 人脸识别应用 静态人脸图像采集规范》一文中指出“要求图像清晰，五官及面部轮廓无明显畸变，无明显阴影，肤色呈自然真实色调”，并对色彩、亮度和锐度处理提出了要求。公安行业标准《安防人脸识别应用系统 第 2 部分：人脸图像数据》中对人脸图像脸部区域作出了如下要求：“人脸完整、轮廓清晰、人脸长宽比例不失真，光照均匀且无阴影”。Karahan 等人 [2] 在《How Image Degradations Affect Deep CNN-based Face Recognition?》一文中从图像清晰度、噪声、遮挡、颜色失真等方面研究图像退化对人脸识别性能的影响。Abhishek Dutta 等人 [3] 从人脸姿态、高斯噪声和运动模糊角度研究了图像质量对识别分数的影响。

从摄像机成像能力角度来看，清晰度，噪声和颜色是三个最受关注的质量维度。基于这三个维度进行细化，提出 6 个质量指标进行分析，包括清晰度、噪声、对比度、亮度、色彩还原、色彩饱和度，给出各单项的质量评价及与整体质量的相关度，用于指导调优。

质量指标计算时对图像局部信息比较敏感，而不同次拍摄时人脸在图像中的相对位置有差异，取人脸图像时容易出现偏差。因此，通过人脸识别模型找到人脸五点位置，利用五点位置抠取出人脸五点区域，从而保证待评价目标图像与对应参考目标图像的评测范围对齐。图 2.2 为目标图像五点区域示意图。

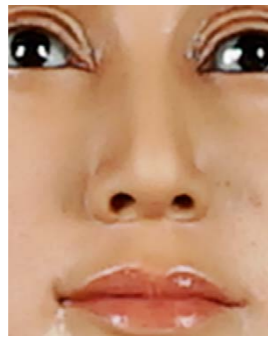


图 2.2 目标图像五点区域示意图

下文详细列出 6 种细分图像指标的计算方法。

2.2.1 清晰度

按如下步骤计算：

a) 将目标图像五点区域转换为灰度图，在灰度空间上计算清晰度：

$$F = \frac{1}{W \times H} \sqrt{\sum_{x=1}^W \sum_{y=1}^H (R(x,y) - R'(x,y))^2} \quad (2-2)$$

式中：

$R(x, y)$ ——表示目标图像五点区域第 x 行第 y 列的像素值

$R'(x, y)$ ——表示目标图像五点区域高斯低通滤波结果中第 x 行第 y 列的像素值

W, H ——表示目标图像五点区域的宽、高

F ——表示目标图像的清晰度

b) 用公式 (2-3) 计算待评价目标图像相对于参考目标图像的清晰度分数：

$$Sc^1 = 1 - \max((F^s - F^t), 0) / F^s \quad (2-3)$$

式中：

F^s ——表示参考目标图像的清晰度

F^t ——表示待评价目标图像的清晰度

Sc^1 ——表示待评价目标图像的清晰度分数

清晰度由高到低的图像示例如下：



图 3.3 不同清晰度水平图像示例

2.2.2 噪声

按如下步骤计算：

a) 根据人脸五点位置找到人眼上方的额头平滑区域，将该区域转换为灰度图，在灰度空间上计算噪声水平：

$$Av = \frac{1}{W \times H} \sqrt{\sum_{x=1}^W \sum_{y=1}^H R(x,y)} \quad (2-4)$$

$$N = \text{VAR}(R(x,y) \times \frac{100}{Av}) \quad (2-5)$$

式中：

$R(x, y)$ ——表示目标图像五点区域第 x 行第 y 列的像素值

Av ——表示目标图像五点区域的均值

W, H ——表示目标图像五点区域的宽、高

N ——表示目标图像噪声水平

b) 用公式 (2-6) 计算待评价目标图像相对于参考目标图像的噪声分数：

$$Sc^2 = 1 - \max((N^t - N^s), 0) / N_{th} \quad (2-6)$$

式中:

N^s ——表示参考目标图像的噪声水平

N^t ——表示待评价目标图像的噪声水平

N_{th} ——表示预设的噪声阈值

Sc^2 ——表示待评价目标图像的噪声分数

噪声由低到高的图像示例如下:

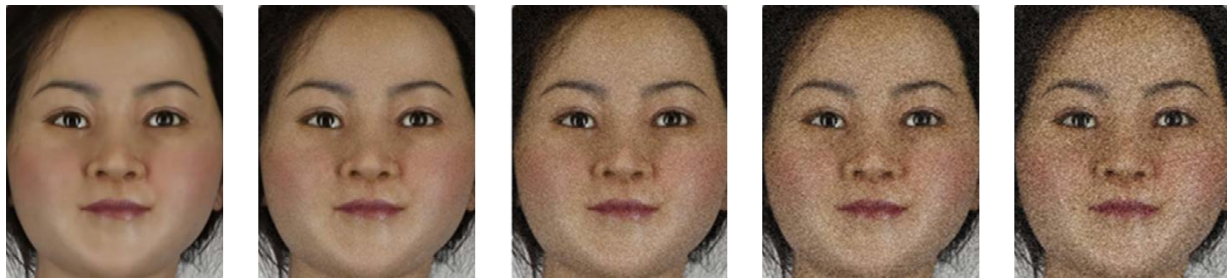


图 3.4 不同噪声水平图像示例

2.2.3 对比度

按如下步骤计算:

a) 将目标图像五点区域转换为灰度图, 在灰度空间上计算对比度:

$$D1_{con} = R(x,y) - \frac{1}{W \times H} \sqrt{\sum_{x=1}^W \sum_{y=1}^H R(x,y)} \quad (2-7)$$

$$D2_{con} = \frac{1}{W \times H} \sqrt{\sum_{x=1}^W \sum_{y=1}^H R(x,y) - R(x,y)} \quad (2-8)$$

$$D1_{con} = \text{where}(D1_{con} > 0, D1_{con}, 0) \quad (2-9)$$

$$D2_{con} = \text{where}(D2_{con} > 0, D2_{con}, 0) \quad (2-10)$$

$$Con = \max(Av_{D1}, Av_{D1}) \quad (2-11)$$

式中:

$R(x, y)$ ——表示目标图像五点区域第 x 行 y 列的灰度值

W, H ——表示目标图像五点区域的宽、高

$D1_{con}$ ——表示目标图像五点区域与其灰度均值求差后, 保留差值图中的正数值, 将负数值取零的结果

$D2_{con}$ ——表示目标图像五点区域与其灰度均值求差后, 将差值图中的正数值取零, 将负数值取绝对值的结果

Av_{D1}, Av_{D1} ——表示 $D1_{con}, D2_{con}$ 的图像均值

b) 用公式 (2-12) 计算待评价目标图像相对于参考目标图像的对比度分数:

$$Sc^3 = 1 - |Con^s - Con^t| / \max(Con^s, Con^t) \quad (2-12)$$

式中:

Con^s ——表示参考图像的对比度

Con^t ——表示待评价图像的对比度

Sc^3 ——表示待评价图像的对比度分数

对比度由高到底的图像示例如下:



图 2.5 不同对比度水平图像示例

2.2.4 亮度

按如下步骤计算：

a) 将目标图像五点区域转换为灰度图，并将该图均分为 8×8 个区块，在灰度空间上计算亮度：

$$Lt_i = \frac{1}{W \times H} \sum_{x=1}^{x=W} \sum_{y=1}^{y=H} R_i(x,y) \quad (2-13)$$

式中：

W, H ——表示目标图像五点区域的宽、高

$R_i(x, y)$ ——表示目标图像五点区域第 i 区块第 x 行第 y 列的像素值

Lt_i ——表示目标图像五点区域第 i 区块的亮度均值

b) 用公式 (2-14) 计算待评价目标图像相对于参考目标图像的亮度分数：

$$Sc^4 = 1 - \frac{1}{64} \sum_{i=1}^{i=64} |Lt_i^s - Lt_i^t| / Lt_i^s \quad (2-14)$$

式中：

Lt_i^s ——表示参考目标图像五点区域第 i 区块的亮度

Lt_i^t ——表示待评价目标图像五点区域第 i 区块的亮度

Sc^4 ——表示待评价目标图像的亮度分数

亮度由高到低的图像示例如下：



图 2.6 不同亮度水平图像示例

2.2.5 色彩还原

将目标图像五点区域转到 CIELAB 色彩空间，并将该图均分为 8×8 个区块，计算色彩还原分数：

$$\Delta E_i = 1 - \sqrt{(L_i^s - L_i^t)^2 + (A_i^s - A_i^t)^2 + (B_i^s - B_i^t)^2} \quad (2-15)$$

$$Sc^5 = \frac{1}{64} \sum_{i=1}^{i=64} \Delta E_i \quad (2-16)$$

式中：

L_i^s ——表示参考目标图像五点区域第 i 区块的 L 分量均值

L_i^t ——表示待评价目标图像五点区域第 i 区块的 L 分量均值

A_i^s ——表示参考目标图像五点区域第 i 区块的 A 分量均值

A_i^t ——表示待评价目标图像五点区域第 i 区块的 A 分量均值

B_i^s ——表示参考目标图像五点区域第 i 区块的 B 分量均值

B_i^t ——表示待评价目标图像五点区域第 i 区块的 B 分量均值

ΔE_i ——表示待评价目标图像五点区域第 i 区块的色彩还原分数

Sc^5 ——表示待评价目标图像的色彩还原分数

色彩还原由高到低的图像示例如下：

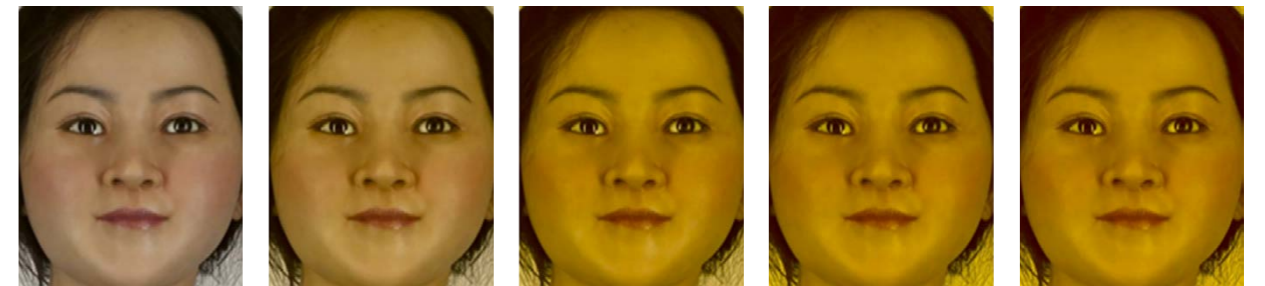


图 2.7 不同色彩还原水平图像示例

2.2.6 色彩饱和度

按如下步骤计算：

a) 将目标图像五点区域均分为 8×8 个区块，按公式 (2-17) 计算目标图像各区块色彩饱和度。

$$Sat_i = \begin{cases} 0 & \text{if } \max_i(R,G,B)=0 \\ 1 - \frac{\min_i(R,G,B)}{\max_i(R,G,B)} & \text{otherwise} \end{cases} \quad (2-17)$$

式中：

Sat_i ——表示目标图像五点区域第 i 区块的色彩饱和度

$\min_i(R,G,B)$ ——表示目标图像五点区域第 i 区块 R,G,B 三通道最小值

$\max_i(R,G,B)$ ——表示目标图像五点区域第 i 区块 R,G,B 三通道最大值

b) 用公式 (2-18) 计算待评价目标图像相对于参考目标图像的色彩饱和分数：

$$Sc^6 = 1 - \frac{1}{64} \sum_{i=1}^{64} |Sat_i^s - Sat_i^t| / Sat_i^s \quad (2-18)$$

式中：

Sat_i^s ——表示参考目标图像五点区域第 i 区块的色彩饱和度

Sat_i^t ——表示待评价目标图像五点区域第 i 区块的色彩饱和度

Sc^6 ——表示待评价目标图像的色彩饱和度分数

色彩饱和度由高到低的图像示例如下：



图 2.8 不同色彩饱和度水平图像示例

2.3 基于机器视觉任务的图像评价与图像质量维度关系

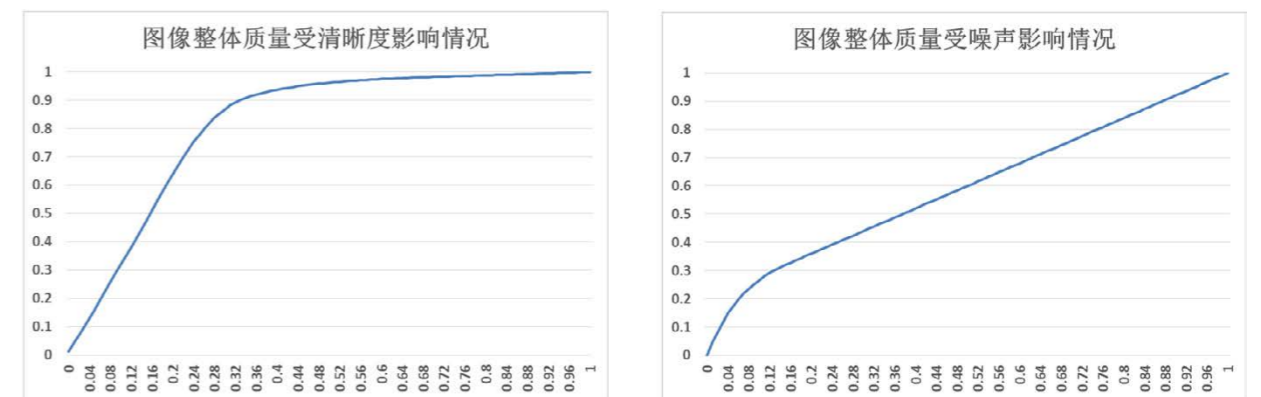
本部分重点研究细分图像指标与图像整体质量的关联关系，试图探寻其内在规律，从而达到指导摄像机成像的优化的目的。

实验方法如下：

根据真实图像退化原理进行仿真，得到模拟真实退化的图像。利用仿真图像，可以计算图像整体质量评价和细分指标分数，从而探索两者之间的关系。实验数据来自 2800 个不同人的证件照，分别对这些照片进行 6 种图像退化处理，得到退化图片集合。部分退化处理后的图像例子如图 2.3—图 2.8 所示。具体退化方法如下：

- 清晰度退化：设置不同程度的高斯模糊核，通过卷积方式对图像添加模糊。
- 噪声退化：设置不同噪声方差，对图像添加服从高斯分布的噪声。
- 对比度退化：转换图像到 CIELAB 颜色空间，调节 L 通道各像素相对于其全图均值的偏差大小，调节后将图像转换回 RGB 颜色空间。
- 亮度退化：R、G、B 三通道同步缩小或放大。
- 色彩还原退化：转换图像到 CIELAB 颜色空间，调节色彩通道数值大小，调节后将图像转换回 RGB 颜色空间。
- 色彩饱和度退化：逐像素调节 R、G、B 数值与 RGB 均值的偏差大小。

计算每种退化图像与原始图像之间的特征相似度，整理数据获得相似度与单项细分指标分数的变化关系。



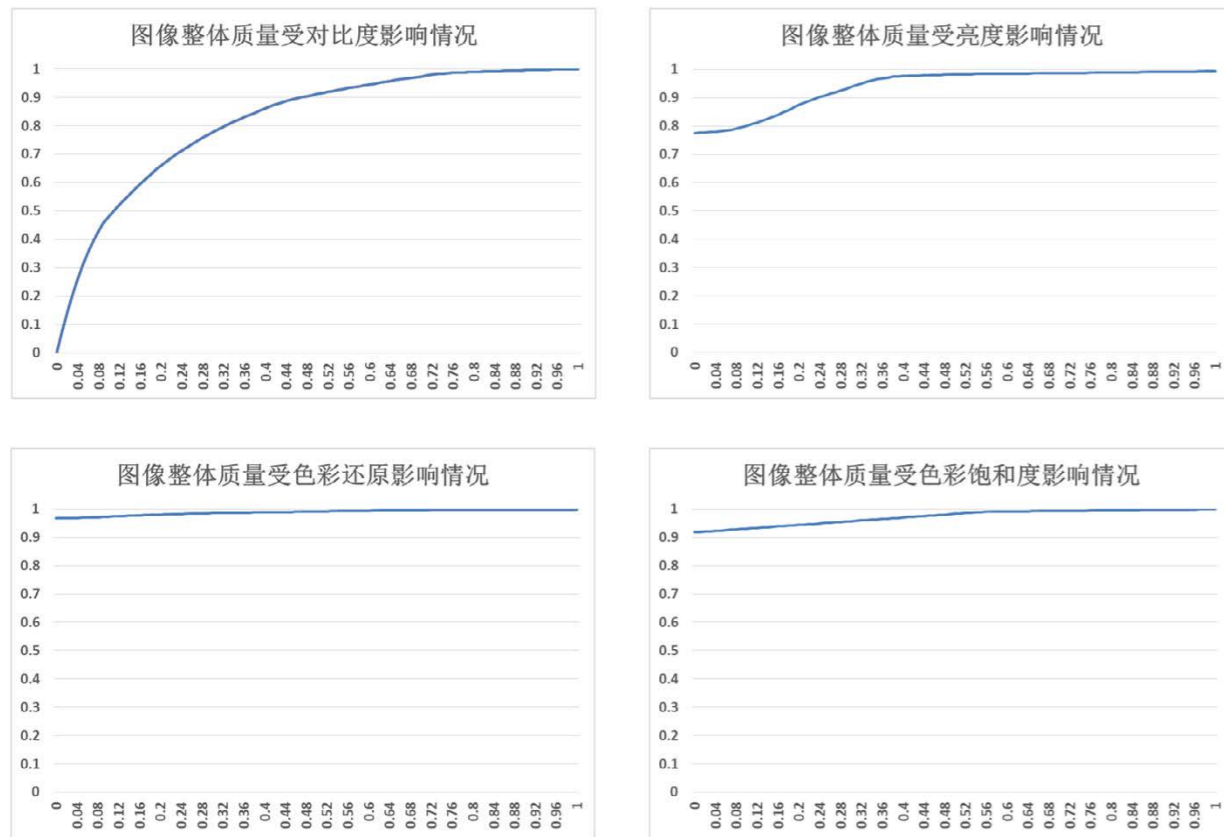


图 2.9 相似度受细分指标分数影响情况

上图横轴代表目标图像的细分指标分数，纵轴代表整体质量分数。从图中可以看出：

- 清晰度、噪声、对比度对图像整体质量有较明显的影响，用这三个质量维度来描述人脸识别任务的图像质量比较合理；色彩还原、色彩饱和度、亮度对人脸识别性能影响较弱。
- 对于部分细分图像指标，当分数低于一定阈值时图像整体质量会发生陡降情况，实际选择摄像机时要尽量避免此类情况的应用。
- 不同细分图像指标对图像整体质量的重要程度不同，每个维度不同区间上对图像整体质量的影响也不同。需要进一步研究针对不同退化图像，各细分图像指标对识别性能的影响程度。

对于摄像机细分图像指标，本方法从 1) 各细分图像指标本身分数和 2) 细分图像指标对整体图像质量的影响两方面进行描述。首先，各细分图像指标是对图像质量的一种客观评价。一般而言，图像质量等级越高，各细分指标分数也相对较高。然而这也不是绝对的，图像整体质量是由多维度纠缠复合的结果，细分指标相对于图像整体质量分数不存在严格的线性关系。尤其是针对特定机器视觉任务，各指标的影响程度差别较大。

针对特定的细分图像指标，使用图 2.9 中的各维度分布曲线，可以对应得到该质量分数对整体图像质量的影响程度，从而可以得到图像质量的改进方向优先级。质量指标对相似度影响越大，则表明该指标调优的优先级越高，调优后对整体图像质量的改善最明显。

下面采用雷达图的形式来展现细分图像指标和各维度指标对图像整体质量的影响。以图 2.10 为例，该图像在清晰度上分数最低，色彩还原和噪声分数分列倒数二三位。而从图像指标机器视觉影响分布图可以看到，噪声对图像整体质量的影响最大，而色彩还原基本没有影响，体现了面向机器视觉任务时细分图像指标对机器识别相似度影响的差异性。

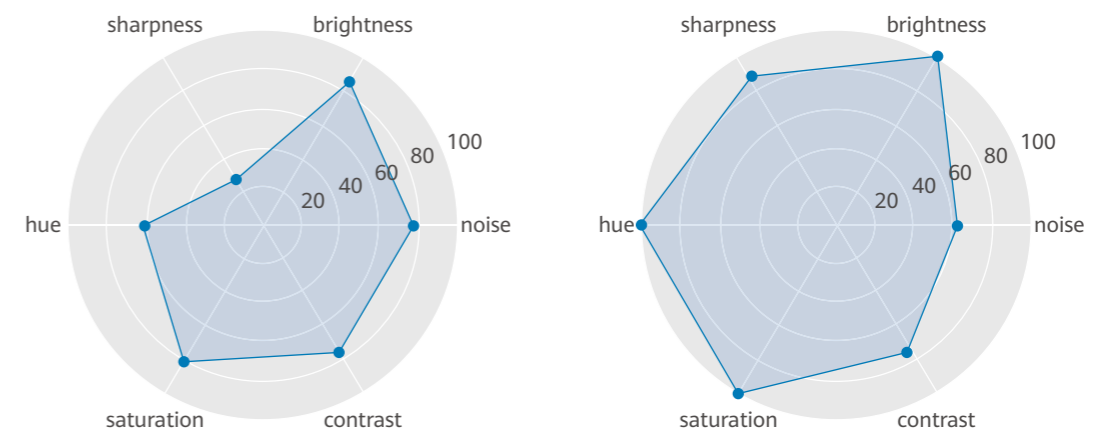


图 2.10 细分图像指标分布（左）和图像指标机器视觉影响分布（右）

3 评测场景构建

3.1 安防摄像机关键应用场景

随着安防摄像机技术的发展，用户对于摄像机在全天候的清晰成像提出了越来越高的要求，特别是低照、强光、逆光等宽动态场景。下面给出些示例

1) 白天均匀光场景；

这个现在各家厂商基本差距很小。



2) 白天背光场景；

安防中逆光背光环境是难以避免的。当摄像机处于逆光环境中拍摄时，画面会出现黑色的图像，目标清晰成像也是摄像机关键技术挑战之一。



3) 白天侧光场景；

侧光会造成阴阳脸等问题。



4) 夜晚低照场景；

黑夜中光线不好、照度低，摄像机看不清人脸，满足不了清晰成像的要求是关键痛点。



5) 人车混行场景。

人车混行场景是治安防控重点建设场景之一。该场景的挑战是画面中远近目标无法清晰呈现，车牌这类高反光的物体也很容易过曝，人、车等目标无法兼顾，无法真正满足视频结构化的应用需求。



3.2 室外场景室内化设计

室外场景室内化是实现图像质量评价标准化可重复的关键之一。

采用室外场景室内模拟化的方式搭建上述场景，模拟构建人脸识别任务实际应用：（1）使用多类别、多角度布光模拟真实室外场景的光照条件；（2）使用高仿真模具模拟待测对象，使用旋转支架装置模拟监控画面中的对象运动；

整体布局安排如图 3.1 所示。以测试支架为中心，在支架前方设置补光光源，在支架后方设置背光源。人头模具安装在测试支架上。摄像机（包括待测摄像机和参考相机）正对测试支架拍摄，拍摄视频通过数据线传输到图形工作站用于分析。

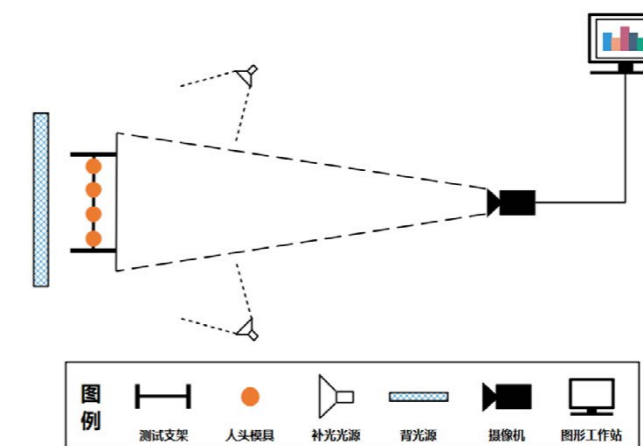


图 3.1 整体布局图

3.3 测试设备

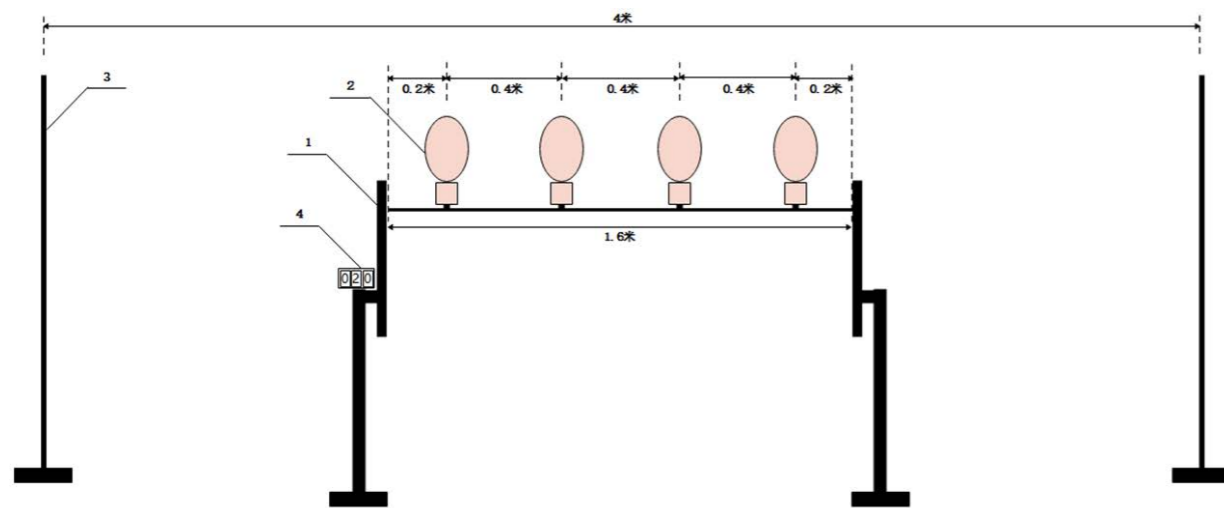
3.3.1 测试支架

测试支架由旋转装置和标定装置构成，外观如图 3.2、图 3.3 所示。

旋转装置用于控制人头模具进行匀速圆周运动。该装置以两个对称旋转臂提供旋转动力，带动人头模具运动。旋转臂之间用支撑架连接，支撑架上固定人头模具。通过机械设计，控制旋转过程中模具保持直立。

标定装置包括位置标定杆和角度显示装置。位置标定杆用于对齐待测摄像机的拍摄画面，放置于旋转装置两侧，间隔为 4 米。

角度显示装置对旋转角度进行实时测量和显示，用于在不同视频中选出相同旋转位置的视频片段。角度显示位数为 4 位，角度测量精度为 $\pm 0.1^\circ$ 。人头模具旋转到正上方时角度为 0° ，旋转到正前方时角度为 90° 。



说明：
1——旋转装置 2——人头模具 3——位置标定杆 4——角度显示装置

图 3.2 测试支架正视示意图

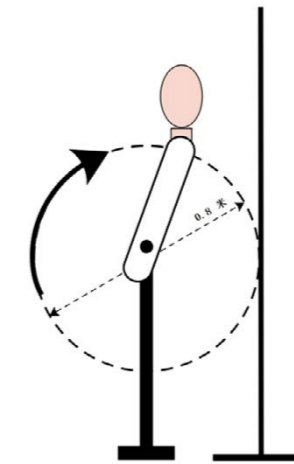


图 3.3 测试支架侧视示意图

3.3.2 人头模具

人头模具使用真人 1:1 比例的仿真头模，从左到右分别为老年男性，青年男性，青年女性，老年女性四种类别。模型皮肤使用硅胶材料，头发使用真人头发。如图 3.4 所示，4 个人头模具并排放置在支撑架上。

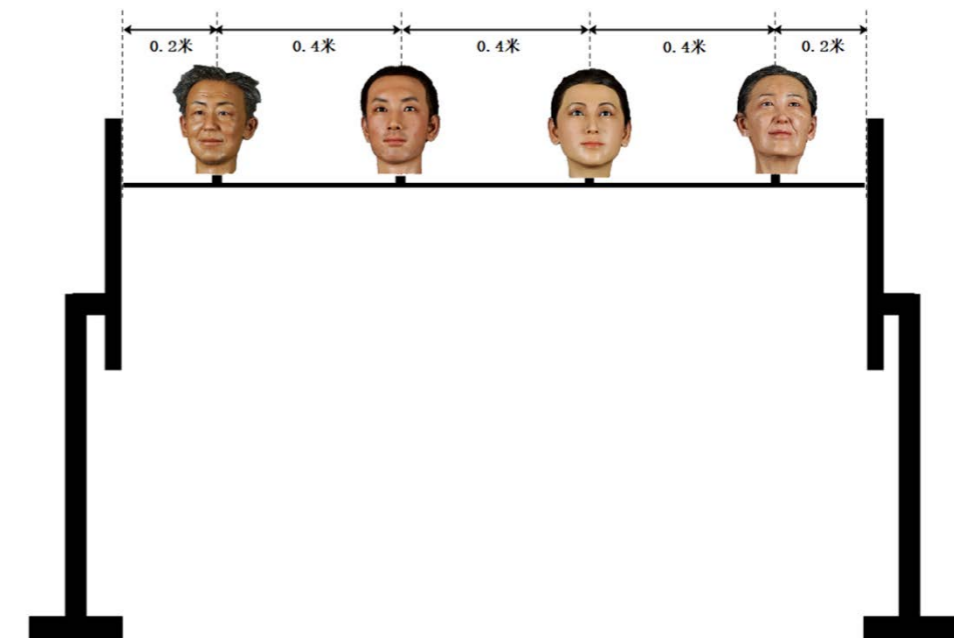


图 3.4 人头模具布置图

拍摄时的参考相机分辨率符合 D 类要求，色彩还原误差应符合 3 级要求，几何失真应不大于 3%，宽动态能力的综合评价评价应不低于 120 分。

3.4 室内模拟场景

本测试内容模拟上文中提到的摄像机的 5 种关键应用场景，即（1）白天均匀光场景；（2）白天逆光场景；（3）白天侧光场景；（4）夜晚低照场景；（5）夜晚低照混行场景。

3.4.1 白天均匀光场景模拟

白天均匀光模拟场景，测试内容为不同色温对图像质量的影响。如图 3.5 所示，在测试支架的 $\pm 45^\circ$ 方向布置补光光源，头模旋转到正前方时，各头模位置色温和照度如下表所示，共测试 3 个光照情况。

表 4.1 白天均匀光模拟场景待测光照表

照度 /lux	色温 /K
1000 ± 100	6500 ± 200
1000 ± 100	3000 ± 200
1000 ± 100	10000 ± 200

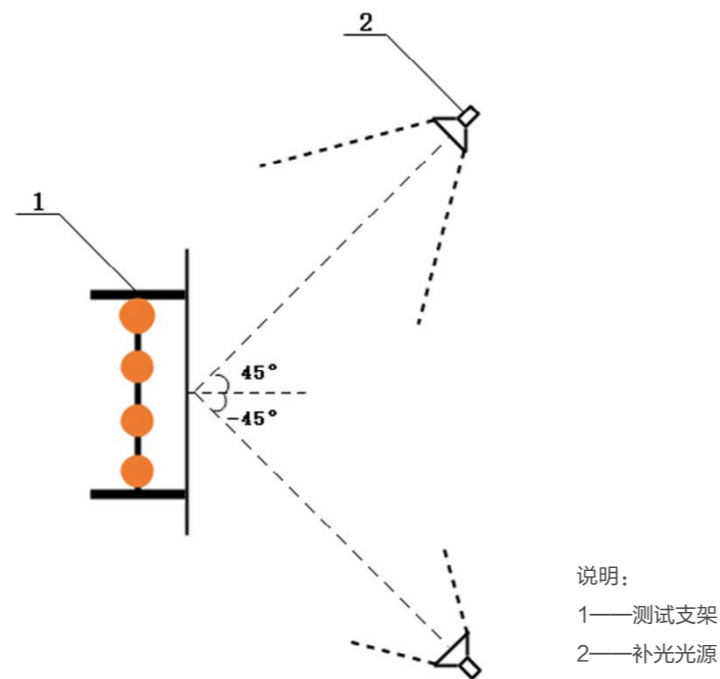


图 3.5 白天均匀光模拟场景布光示意图

3.4.2 白天逆光场景模拟

白天逆光模拟场景，测试内容为不同逆光比例对图像质量的影响。光照布置上，分为均匀光和逆光两部分，如图 3.6 所示。均匀光方面，在测试支架的 $\pm 45^\circ$ 方向布置补光光源；逆光方面，在测试支架后方布置背光源，背光源表面照度均匀度为 90%。待测试色温和照度如表 3.2 所示，共测试 5 个光照情况。

表 4.2 白天逆光模拟场景待测光照表

背光源表面照度 /lux	均匀光照度 /lux	色温 /K
400 ± 40	200 ± 20	5000 ± 200
800 ± 80	200 ± 20	5000 ± 200
1600 ± 160	200 ± 20	5000 ± 200
3200 ± 320	200 ± 20	5000 ± 200
6400 ± 640	200 ± 20	5000 ± 200

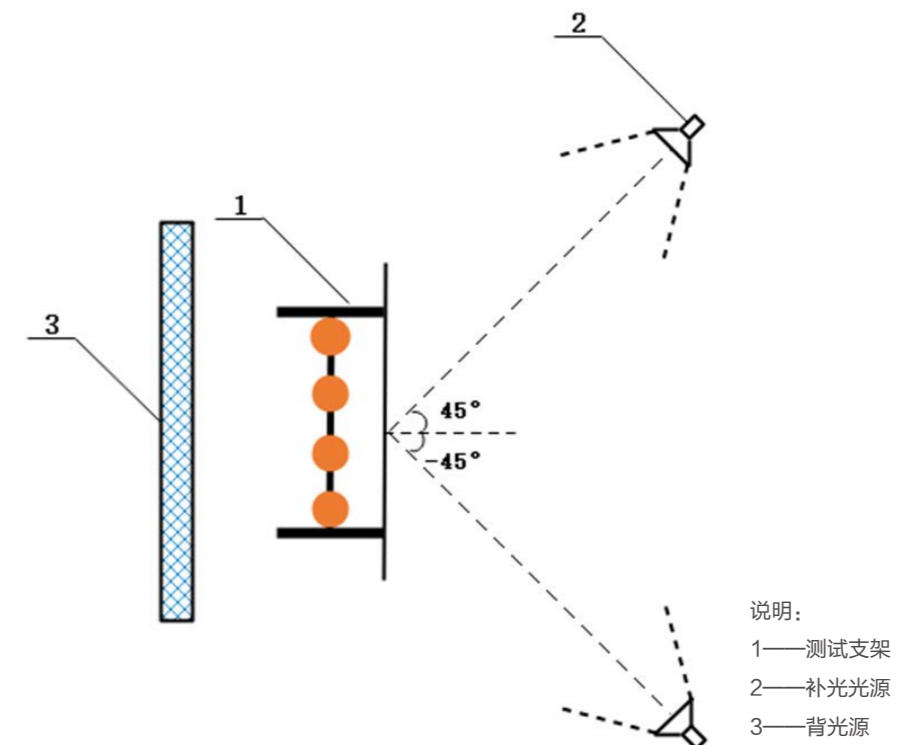


图 4.6 背光源示意图

3.4.3 白天侧光场景模拟

白天侧光模拟场景，测试内容为非均匀侧面光照对图像质量的影响。如图 3.7 所示，在测试支架的 $\pm 75^\circ$ 方向布置补光光源，控制两侧光源照度实现不同侧光情况。头模旋转至正前方时，以四个头模的中心位置为基准点，面向 $\pm 90^\circ$ 方向的色温和照度如表 3.3 所示，共测试 4 个光照情况。

表 3.3 白天侧光模拟场景待测光照表

90° 方向		-90° 方向	
照度 /lux	色温 /K	照度 /lux	色温 /K
1000 ± 100	6500 ± 200	250 ± 25	6500 ± 200
1000 ± 100	6500 ± 200	120 ± 12	6500 ± 200
1000 ± 100	6500 ± 200	60 ± 6	6500 ± 200
1000 ± 100	6500 ± 200	30 ± 3	6500 ± 200

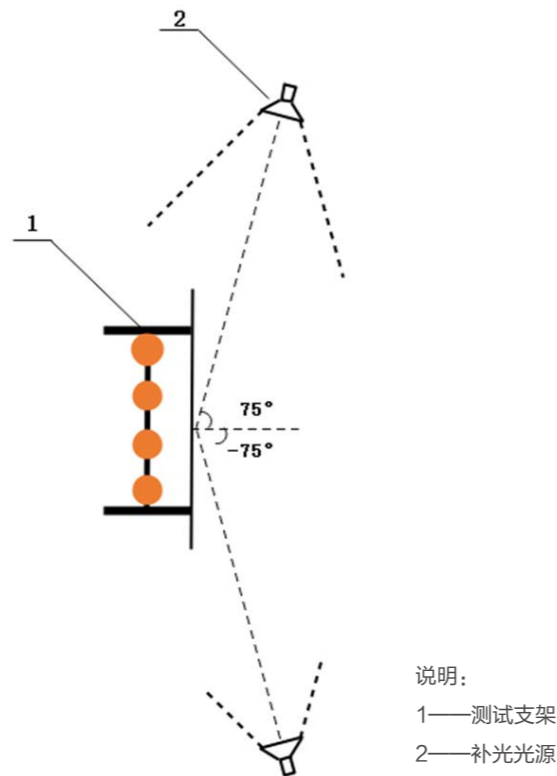


图 3.7 白天侧光模拟场景布光示意图

3.4.4 夜晚低照场景模拟

夜晚低照模拟场景，测试内容为不同低照度情况对图像质量的影响。如图 3.8 所示，在测试支架的 $\pm 45^\circ$ 方向布置补光光源。头模旋转至正前方时，各头模位置面向正前方的色温和照度如表 3.4 所示，计划共测试 7 个光照情况。

表 3.4 夜晚低照模拟场景待测光照表

照度 /lux	色温 /K
0.5 ± 0.05	4000 ± 200
1 ± 0.1	4000 ± 200
2 ± 0.2	4000 ± 200
5 ± 0.5	4000 ± 200
10 ± 1	4000 ± 200
20 ± 2	4000 ± 200
40 ± 4	4000 ± 200

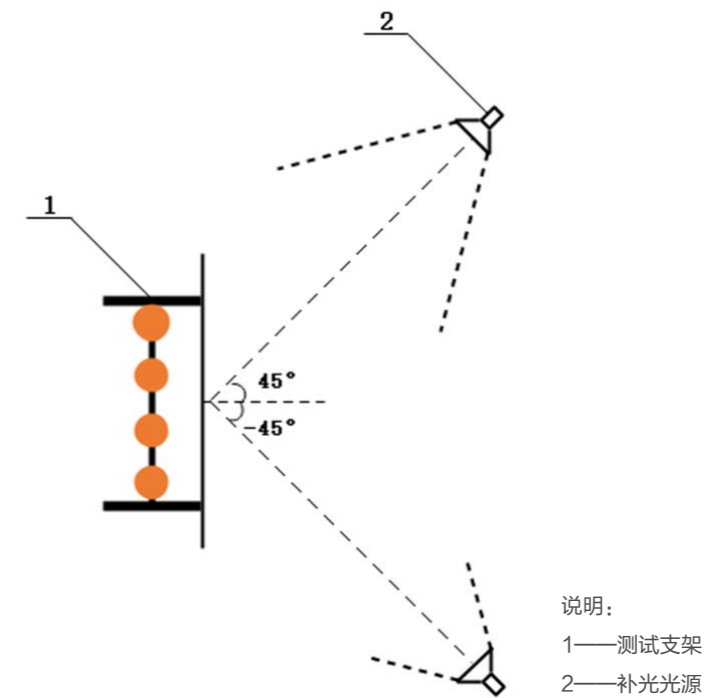


图 3.8 夜晚低照模拟场景布光示意图

3.4.5 夜晚人车混行场景模拟

夜晚人车低照混行模拟场景，测试内容为含有汽车大灯的低照度情况对图像质量的影响。如图 3.9 所示，在测试支架侧面放置一组汽车大灯，表面照度 >1 万 lux。在测试支架的 $\pm 45^\circ$ 方向布置补光光源。头模旋转至正前方时，各头模位置面向正前方的色温和照度如下表所示，共测试 3 个光照情况。

表 3.4 夜晚低照模拟场景待测光照表

照度 /lux	色温 /K
5 ± 0.5	4000 ± 200
10 ± 1	4000 ± 200
20 ± 2	4000 ± 200

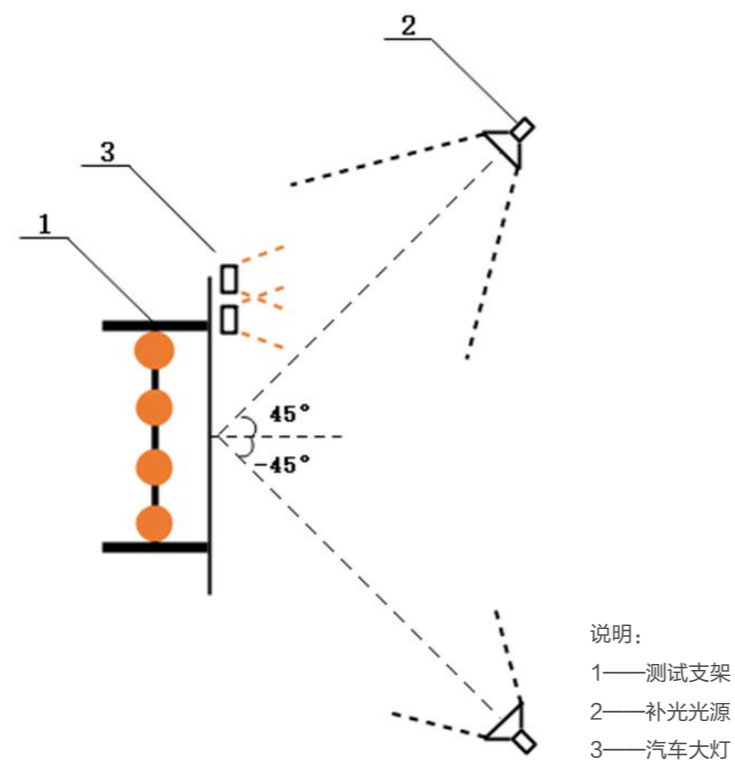


图 3.9 夜晚低照模拟场景布光示意图

4 测试流程及评价方法

4.1 测试流程

测试过程主要包括两个部分：（1）参考目标图像的获取和（2）待评价目标图像的获取。

4.1.1 参考目标图像获取步骤

参考目标图像获取步骤如下：

- 使用参考相机对测试支架进行拍摄。调节相机姿态位置，使得拍摄画面边缘与测试支架的位置标定杆对齐。
- 参考正文的设定，环境光照调节为指定色温及光强。
- 调节测试支架旋转位置，控制支架静止在特定角度后，使用参考相机拍摄参考照。参考照拍摄的支架旋转角度包括 0° 、 10° 、 20° 、 30° 、 40° 、 50° 、 60° 、 70° 、 80° 、 90° 。
- 使用人脸检测算法在各角度参考照中检出参考目标图像。



图 4.1 参考目标图像获取流程

4.1.2 视频评测目标图像获取步骤

视频评测时，待评价目标图像获取步骤如下：

- 使用待测摄像机对测试支架进行拍摄。调节摄像机姿态位置，使得拍摄画面边缘与测试支架的位置标定杆对齐。

- b) 启动测试支架，等待 2 秒以保证进入匀速转动状态。
- c) 环境光照调节为指定色温及光强，使用待测摄像机进行视频拍摄。拍摄视频时长应大于一个旋转周期。
- d) 针对上述视频，截选出支架旋转角度 0-90° 范围的视频图像。使用人脸检测算法对上述视频图像进行逐帧检测，记录不同旋转角度下的待测目标图像坐标信息，同时输出该场景的待测目标图像序列。
- e) 依次开启定义的各场景光照，使用待测摄像机进行拍摄。拍摄视频时长应大于一个旋转周期。
- f) 针对上述各场景视频，截选出支架旋转角度 0-90° 范围的视频图像。根据第 4 步得到的不同旋转角度下的待测目标图像坐标信息，从各场景视频中选出相应场景的待测目标图像序列。

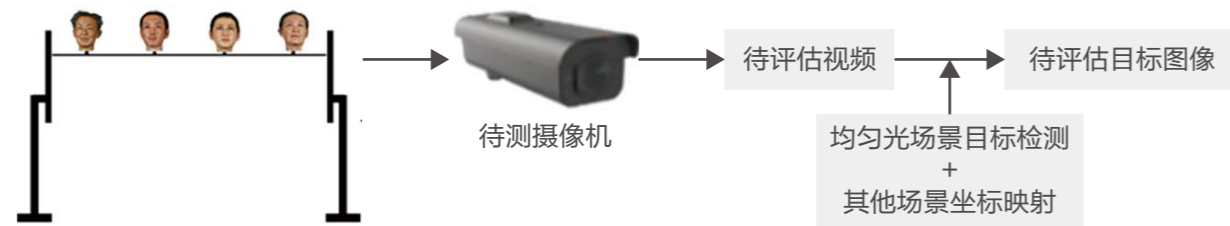


图 4.2 视频评测目标图像获取流程

4.2 评分方法

图像质量评分计算方法如下：

- a) 对单个目标图像，计算整体质量评价，计算各维度细分图像指标评价。
- b) 对于单场景中单个头模的所有目标图像，按公式 (4-1) 和 (4-2) 计算单场景单头模质量评价：

$$VFS = \frac{1}{t} \sum_{f=1}^{f=t} S_f \tag{4-1}$$

$$VF^j = \frac{1}{t} \sum_{f=1}^{f=t} Sc_f^j \tag{4-2}$$

式中：

VFS —— 表示单场景单头模整体质量评价

VF^j —— 表示单场景单头模第 j 个维度的细分图像指标评价

S_f —— 表示单场景中单个头模的第 f 个待评价目标图像的整体质量评价，上限 t 表示该场景中该头模一共有 t 个目标图像

Sc_f^j —— 表示单场景中单个头模的第 f 个待评价目标图像的第 j 个维度的细分图像指标评价，上限 t 表示该场景中该头模一共有 t 个目标图像

注：针对开源及第三方算法，需要基于误报率对于模型输出的目标相似度进行距离修正从而降低波动。此时上述公式 4-1,4-2 中需要增加一个 $f(x)$ 用于距离修正

- c) 按公式 (4-3) 和公式 (4-4) 计算单个场景质量评价。

$$VS = \frac{1}{4} \sum_{i=1}^{i=4} VFS_i \tag{4-3}$$

$$V^j = \frac{1}{4} \sum_{i=1}^{i=4} VF_i^j \tag{4-4}$$

式中：

VS —— 表示单个场景整体质量评价

V^j —— 表示单个场景第 j 个维度的细分图像指标评价

VFS_i —— 表示单场景中第 i 个头模的整体质量评价

VF_i^j —— 表示单场景中第 i 个头模的第 j 个维度的细分图像指标评价

4.1.3 抓拍评测目标图像获取步骤

抓拍评测时，待评价目标图像获取步骤如下：

- a) 使用待测摄像机对测试支架进行拍摄。调节摄像机姿态位置，使得拍摄画面边缘与测试支架的位置标定杆对齐。
- b) 启动测试支架，等待 2 秒以保证进入匀速转动状态。
- c) 依次开启定义的各场景光照，使用待测摄像机进行抓拍拍摄。若相机直接输出目标图像，每个场景针对每个头模抓取 10 张目标图像；若相机抓拍输出为全画面图像，每个场景抓取 10 张全画面图像，使用人脸检测算法对抓取图像进行检测，获得待测目标图像。

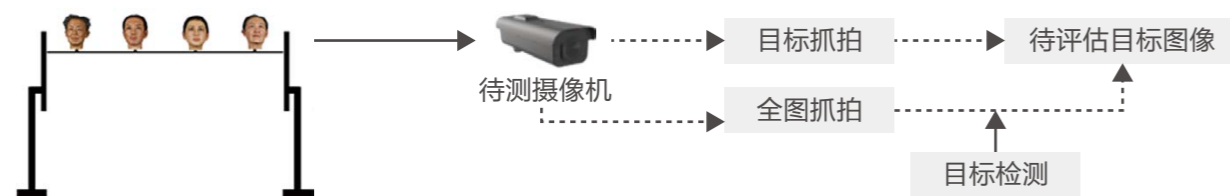


图 4.3 抓拍评测图像获取流程

d) 对各个场景的质量评价分数进行加权评价, 得出整体质量评价分数 MVB。

e) 计算单个场景细分图像指标对其对图像整体质量的影响

$$\text{IMP}^i = f_x(V^i) \quad (4-5)$$

式中:

IMP_j —— 表示单个场景第 j 个维度对图像整体质量的影响分数

V_j —— 表示单个场景第 j 个维度的细分图像指标评价

$f_x()$ —— 表示查表函数

这样, 我们利用单个场景各维度细分图像指标评价和细分指标对整体质量的影响分数, 绘制细分图像指标分布图和图像指标机器视觉影响分布图。

参考文献

- [1] Jain, Anil K., and Stan Z. Li. Handbook of face recognition. Vol. 1. New York: Springer, 2011.
- [2] Karahan, Samil, et al. "How image degradations affect deep cnn-based face recognition?." 2016 International Conference of the Biometrics Special Interest Group (BIOSIG). IEEE, 2016.
- [3] Dutta, Abhishek, Raymond Veldhuis, and Luuk Spreeuwens. "Can facial uniqueness be inferred from impostor scores?." arXiv preprint arXiv:1310.6376 (2013).
- [4] Aggarwal, Gaurav, et al. "Predicting good, bad and ugly match pairs." 2012 IEEE Workshop on the Applications of Computer Vision (WACV). IEEE, 2012.
- [5] Phillips, P. Jonathon, et al. An introduction to the good, the bad, & the ugly face recognition challenge problem. IEEE, 2011.
- [6] Bosse S, Maniry D, Müller KR, Wiegand T, Samek W. Deep neural networks for no-reference and full-reference image quality assessment. IEEE Transactions on Image Processing. 2017 Oct 10;27(1):206-19.
- [7] Deng J, Guo J, Xue N, et al. Arcface: Additive angular margin loss for deep face recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2019: 4690-4699.