

南京邮电大学
Nanjing University of Posts and Telecommunications

数字图像与视频处理

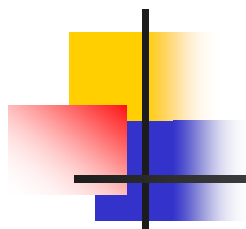
第9章 图像与视频的质量评价





本章学习目标

- 掌握人眼视觉特性的知识，包括对比敏感度和掩盖效应的概念。
- 掌握图像与视频质量的主观评价方法。
- 掌握全参考图像质量的客观评价方法，包括基于信号保真度的均方误差和峰值信噪比，基于结构相似性的质量评价，以及基于信息保真度准则的评价方法。
- 了解半参考、无参考图像质量评价方法，包括自然场景统计特性的原理和建模方法，以及常见的空域与频域特征提取方法。
- 了解全参考、半参考和无参考视频质量评价方法。



第9章 图像与视频的质量评价

- 9.1 常见的图像与视频失真类型
- 9.2 图像与视频质量的主观评价
- 9.3 图像质量的客观评价
- 9.4 视频质量的客观评价
- 9.5 MATLAB编程实例
- 9.6 小结



9.1常见的图像与视频失真类型

基本概念:

- **参考图像:** 也称为标准图像或无失真图像, 即原始没有受到任何失真的图像。一般指通过图像采集设备获取还未经过压缩等处理的原始图像。
- **失真图像:** 指参考图像在经过压缩等处理过程中受到不同类型或不同程度失真后产生的待评价图像。



9.1 常见的图像与视频失真类型

基本概念:

- **图像质量**: 有两方面的含义
 - **图像保真度**: 即失真图像与参考图像之间的相似程度或信息保持程度, 反映人眼观察图像时视觉感知的舒适性, 人眼视觉感知越舒适, 则认为图像质量越好。
 - **图像可懂度**: 即图像向人或机器提供有效信息的程度。
- 本书中的图像质量是指**图像保真度**, 主要讲述针对图像保真度的测量方法。
- **图像质量评价 (IQA)** 与**视频质量评价 (VQA)** 就是通过主观或客观的方式对失真图像或视频进行评分, 以准确反映失真图像或视频的视觉质量。

图像失真类型

■ (1) 图像编码产生的压缩失真

- 图像的数据量巨大，在绝大多数应用环境中都需要先经过编码（压缩）再进行存储或传输，在编码过程中，常使用量化来减少数据量，从而产生压缩失真。
- 压缩技术不同，产生的失真视觉效果也不相同。



a) 参考图像



b) JPEG 压缩失真



c) JPEG2000 压缩失真

图 9-1 压缩失真示意图

图像失真类型

- (2) 图像采集时由于镜头器件缺陷产生的噪声失真
 - 在采集图像时，由于镜头内部物理器件的缺陷有时会产生噪声失真，根据噪声特性是否与信号相关，常分为加性噪声与乘性噪声。
 - 加性噪声的特性与信号无关，而乘性噪声的特性与信号相关，其幅度会受到信号的调制。一般将噪声视为加性高斯白噪声（AWGN）进行处理和分析。



a) 参考图像

b) 加入高斯白噪声后的图像

图 9-2 噪声失真示意图

图像失真类型

- (3) 图像采集时由于镜头抖动或散焦产生的模糊失真
 - 模糊失真也是很常见的一种失真类型，根据成因一般可分为散焦模糊或由于镜头与景物之间的相对运动产生的运动模糊。
 - 对模糊失真特性建模时，为简单起见常用圆对称的高斯模糊来模拟。



a) 参考图像

b) 模糊失真图像

图 9-3 模糊失真示意图

图像失真类型

- (4) 码流在易错信道中传输由于比特误码产生的传输失真
 - 编码后的图像码流在易错信道中传输，有可能由于网络环境（如快速衰落的瑞利信道）的影响而造成**比特误码**，虽然先进的编码器中会有**错误弹性机制**，但比特误码产生的失真仍然无法完全消除。



a) 参考图像

b) 传输失真图像

图 9-4 传输失真示意图



视频失真类型

- (1) 视频编码产生的压缩失真
 - 根据目前常用的视频编码标准，压缩失真主要包括MPEG-2压缩失真、MPEG-4/H. 264AVC压缩失真、HEVC压缩失真与AVS压缩失真等。
 - 压缩失真在视觉上表现为空域上的方块效应、假轮廓、细节模糊等。
- (2) 视频压缩后的码流在易错信道中传输时由于数据丢失产生的传输失真
 - 传输失真主要与信道条件相关，比如在无线信道中传输产生的误码失真、在IP网络中传输产生的数据丢包失真等。
 - 传输失真在视觉上既表现为空域上的信息错乱，又表现为时域上的运动补偿不匹配、拖拽效应、抖动效应等。
 - 一帧图像的部分信息丢失会影响当前帧周围信息甚至相邻多帧信息的正确解码，大大影响视频质量，这就是视频中的误码传播现象。

视频失真类型



a) 参考视频



b) MPEG-2 压缩失真

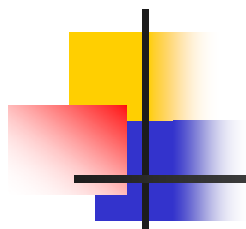


c) H.264 压缩失真



d) 传输失真

图 9-5 视频失真示意图



第9章 图像与视频的质量评价

- 9.1 常见的图像与视频失真类型
- **9.2 图像与视频质量的主观评价**
- 9.3 图像质量的客观评价
- 9.4 视频质量的客观评价
- 9.5 MATLAB编程实例
- 9.6 小结



9.2.1 对比敏感度与视觉掩盖效应

- 人类视觉系统 (Human Visual System, HVS)
 - 是大多数图像处理应用的最终接收者。
 - 压缩失真在视觉上表现为空域上的方块效应、假轮廓、细节模糊等。
- 人眼视觉特性 (Human Visual Characteristics)
 - HVS对图像与视频质量的感知有其内在的规律和特点。
 - HVS对视觉信息的处理机制是一个极其复杂的过程，涉及到视觉神经学、生理学、心理学、认知学等多学科知识，目前还没有被完全的认识和理解。
 - 目前更多地是使用工程化的方法对输入输出关系进行建模。



9.2.1 对比敏感度与视觉掩盖效应

- 1. 对比敏感度 (Contrast Sensitivity)
 - 对比度决定了人眼对亮度变化的感受程度，**对比度阈值** (Contrast Threshold) 是指人眼能觉察到的亮度变化的临界值；
 - 对比敏感度定义为对比度阈值的倒数。对比度阈值越低，则对比敏感度越高，即人眼能分辨亮度变化的能力越强。
 - 对于视觉信号中不同的空间频率内容，HVS具有不同的对比敏感度。
 - **对比敏感度函数** (*Contrast Sensitivity Function, CSF*) 就描述了对比敏感度随空间频率的变化特性，其数学表达式为：

$$CSF(f) = 2.6 \times (0.0192 + 0.114f) \times \exp[-(0.114f)^{1.1}]$$

- 其中 f 是图像的空间频率，单位是周期/度 (cycles/degree)。

9.2.1 对比敏感度与视觉掩盖效应

随着空间频率由小到大，人眼对比敏感度先由小变大，再由大变小，大约在空间频率为**8**周期/度的时候取得最大值，当空间频率增大到**60**周期/度后对比敏感度趋近于零，即对亮度变化不再敏感。

CSF具有带通特性，对中间频率敏感度大，对高频与低频敏感度相对较小。

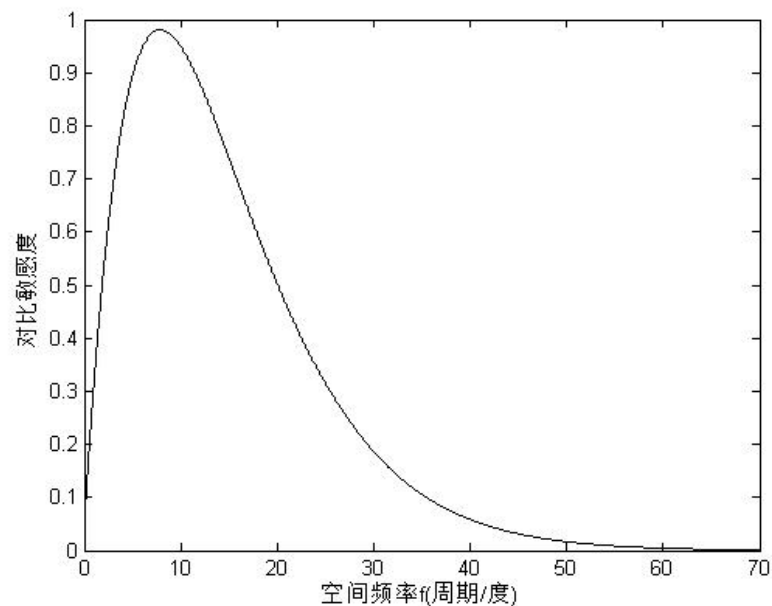
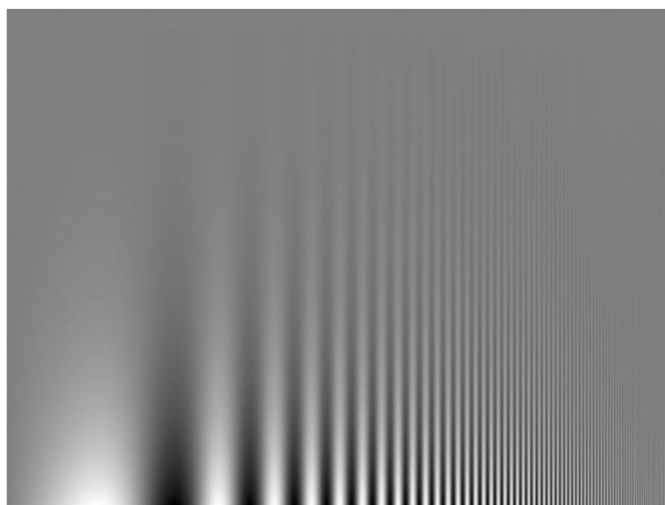


图9-6 对比敏感度函数示意图



9.2.1 对比敏感度与视觉掩盖效应

- 2. 视觉掩盖效应 (Visual Masking Effect)
 - 视觉掩盖效应 (也称视觉掩模)：指人眼对视觉信号的感知能力（或视觉信号的可见性）会因为周围视觉信号的存在而减弱的现象，其强弱可根据掩盖信号出现与否所导致的视觉信号可见性的变化程度来衡量。
 - 一般而言，当视觉信号与周围掩盖信号具有相近的空间位置和运动情况、相似的频率变化情况与方向模式时，掩盖效应更强。
 - 常见的掩盖效应包括对比度掩盖效应、纹理掩盖效应、运动掩盖效应等。
 - 掩盖效应会使视觉信号中失真的可见性发生变化，很多最近的IQA方法也都利用了视觉掩盖效应，以定量地识别不同类型的失真被掩盖的程度。

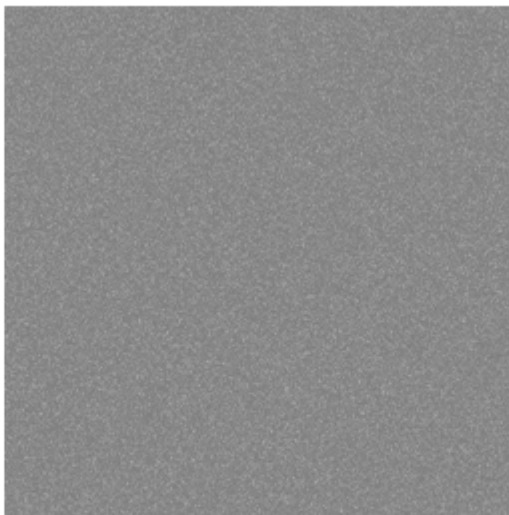
9.2.1 对比敏感度与视觉掩盖效应

从**IQA**的角度来分析，图**(b)**是参考图像与失真图像之间的误差信号，误差强度在空间上分布均匀，但是从图**(c)**可以看出，随着局部空间位置上纹理模式和复杂程度的不同，误差的可见性也明显不同。

Lena图中的帽穗处，空间频率很高，纹理模式复杂，高斯噪声失真被掩盖。



a) 参考图像



b) 高斯白噪声图像



c) 失真图像

图 9-7 视觉掩盖效应示意图



9.2.2 电视图像质量的主观评价方法

- 主观意见评分（ Mean Opinion Score, MOS）
 - 图像与视频质量的主观评价方法类似，都是通过人工来观察图像，并对感知到的图像或视频质量进行打分，最后对多人打分的结果进行统计平均，得到图像或视频的平均评分。
 - 为了得到统计上有意义的主观质量评价结果，对观察者的特性、观察时的实验环境（比如显示器的大小、亮度和对比度、人眼观察距离以及环境亮度等）、打分标准、具体评价方法与流程等都有较严格的规定。
 - 在2012年1月公布的建议书BT. 500-13 “电视图像质量的主观评价方法”和BT. 710-2 “高清晰度电视图像质量的主观评价方法”中对评价方法和实验环境等因素给出了指导性建议。



9.2.2 电视图像质量的主观评价方法

- 1. 对观察者的要求
 - 人数、观察者的图像处理经验、视力校正、打分规则说明等。
- 2. 对实验环境的要求
 - 显示器的亮度、对比度、分辨率、观察距离、环境亮度与照度等。
- 3. 打分标准（5级打分法）
 - 损伤尺度：有图像处理经验的专业人员更容易发现图像与视频中的失真，宜采用损伤尺度。
 - 质量尺度：没有图像处理经验的一般人员宜采用质量尺度。

质量尺度	得分	损伤尺度	得分
非常好	5	无察觉	5
好	4	刚察觉	4
一般	3	轻微讨厌	3
差	2	讨厌	2
非常差	1	难以观看	1

9.2.2 电视图像质量的主观评价方法

■ 4. 具体评价方法

- 一个评价阶段应在半小时以内完成，以防观察者出现视觉疲劳。
- 对于静态图像，展示3~4秒并重复5次（最后两次用于评价）比较合适。
- 对具有时变特性的视频序列，每个序列展示10秒并重复2次（第二次用于评价）比较合适。

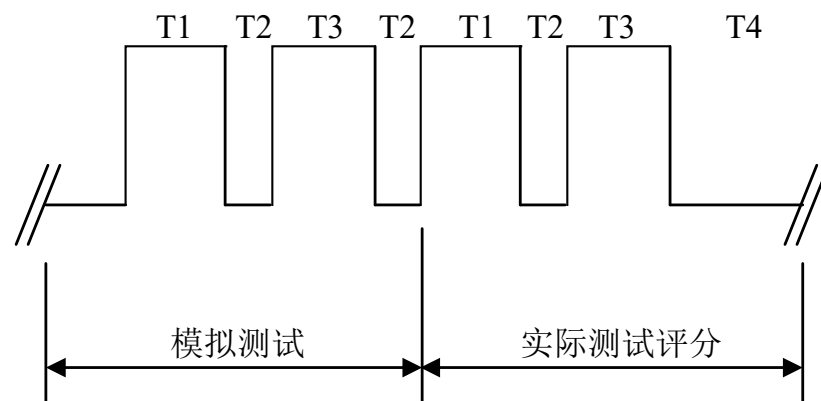


图9-8 评价流程示意图



9.2.2 电视图像质量的主观评价方法

■ 4. 具体评价方法

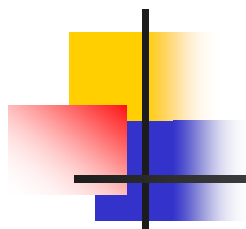
- 评价方法常采用双刺激连续质量尺度（Double Stimulus Continuous Quality Scale, DSCQS）法。
- 所谓“双刺激”是指由参考图像与相应的失真图像组成图像对，观察者观看以随机顺序出现的一系列图像对（图像对中参考图像与失真图像出现的顺序是随机的，且观察者不被告知哪个是参考图像，哪个是失真图像，以避免观察者打分时带有偏见），并对两者的质量都按照连续尺度的5级打分法给出评价。
- 最后根据不同观察者对同一测试图像或视频的主观评分，进行归一化和统计分析，得到主观评价结果。



9.2.2 电视图像质量的主观评价方法

■ 主观质量评价的优缺点

- **优点：**准确可靠，因为大多数图像处理系统的最终接收者都是人眼，因此其评价结果常作为基准来判断客观质量评价方法的预测性能。
- **缺点：**
 - 1) 主观评价需要的观察者数目较多，**会耗费大量人力和时间**，实际操作起来很不方便且综合成本较大。
 - 2) 主观评价有时会受到观察者个人偏好、观察经验、当时情绪等不确定心理因素的影响，从而对打分结果产生一定影响。
 - 3) 主观评价只能做事后评价，无法应用于需要对图像与视频质量**实时监控**的场合。
 - 4) 最重要的是，主观评价无法进行“自动的”质量评价，更无法作为质量目标来指导图像与视频系统的优化设计。



第9章 图像与视频的质量评价

- 9.1 常见的图像与视频失真类型
- 9.2 图像与视频质量的主观评价
- **9.3 图像质量的客观评价**
- 9.4 视频质量的客观评价
- 9.5 MATLAB编程实例
- 9.6 小结



9.3.1 图像质量客观评价方法的分类

- 1. 基于参考图像的可用性进行分类
 - 1) 全参考 (Full Reference, FR) IQA: 对失真图像进行评价时根据需要可以利用参考图像的所有信息。
 - 2) 半参考 (Reduced Reference, RR) IQA: 也称为减少参考或部分参考质量评价, 对失真图像评价时利用了参考图像的部分信息, 这部分信息通常是从参考图像中提取的具有代表性的特征。
 - 3) 无参考 (No Reference, NR) IQA: 也称为盲 (Blind) 图像质量评价, 即对失真图像进行评价时不利用对应的参考图像做对比, 仅仅根据失真图像自身的信息进行质量评价。。



9.3.1 图像质量客观评价方法的分类

- 2. 基于失真图像是否包含彩色信息进行分类
 - 1) 灰度图像 (Gray Image) 质量评价: 失真图像仅包含亮度通道, 不包含彩色信息。
 - 2) 彩色图像 (Color Image) 质量评价: 失真图像包含彩色信息, 根据不同格式具有不同的彩色通道, 如RGB、YUV、HSV等。
- 3. 基于应用范围进行分类
 - 1) 通用的质量评价: 是指设计的质量评价方法可应用于不同的领域和失真类型。
 - 2) 专用的质量评价: 所设计的质量评价方法专门针对某些应用领域或特定的失真类型, 如针对JPEG压缩失真的质量评价、针对模糊程度 (Blur) 或锐度 (Sharpness) 的质量评价、监控网络视频流质量的评价方法等。



9.3.2 全参考图像质量评价

- 本章主要针对单通道的灰度图像（暂不涉及彩色图像质量评价）；
- 假定待评价的失真图像与参考图像尺寸一致，且在像素空间坐标位置上是严格对齐的，即两者之间仅存在灰度值的差异，而不存在位置上的偏移（如平移、转换、缩放等）。
 - 这符合大部分图像处理系统的应用场景，比如参考图像经过压缩编码（JPEG、JPEG2000）、噪声污染或信道误码后引入的失真都表现为像素值的变化而图像像素位置不变。



9.3.2 全参考图像质量评价

- 1. 均方误差 (Mean Squared Error, MSE)
 - 基于信号保真度（或误差信号敏感性）的IQA方法，认为失真图像是由参考图像加上误差信号得到，通过测量误差信号的视觉感知强度来评价图像失真的程度。
 - 对于数字化的尺寸为 $M \times N$ 个像素的二维灰度图像，设 $f(i, j)$ 与 $g(i, j)$ 分别表示参考图像与失真图像在 (i, j) 位置的像素值，则二者之间的均方误差（MSE）定义为对应位置像素灰度值误差的平方的平均值。

$$MSE = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N [f(i, j) - g(i, j)]^2$$



9.3.2 全参考图像质量评价

- 1. 均方误差 (Mean Squared Error, MSE)
 - MSE计算每一像素位置上灰度值的误差，之后做平方与平均化处理，因此可以写成如下等价的一维信号误差统计的形式：

$$MSE = \frac{1}{K} \sum_{k=1}^K [f(k) - g(k)]^2$$

- 其中 $K=M \times N$ ， $k=\{1, 2, \dots, K\}$ 表示像素索引， $f(k)$ 与 $g(k)$ 分别表示参考图像与失真图像的第 k 个像素值。
- 这相当于将二维图像信号按行或按列拉成一维向量的形式进行误差统计。



9.3.2 全参考图像质量评价

- 1. 均方误差 (Mean Squared Error, MSE)
 - MSE根据信号误差的统计特性来表达图像的失真程度，
 - MSE值越大，表示两个图像在所有像素位置上平均的误差平方值越大，即失真图像偏离参考图像的程度越大，其失真就越大，质量越低；
 - 反之，MSE值越小，失真越小，质量越高。
 - 极端情况下，当失真图像与参考图像在每一像素位置上的灰度值都相同时，MSE获得最小值0。
 - 至今，MSE仍然广泛应用于大量的图像处理与质量评价系统中。



9.3.2 全参考图像质量评价

■ 1. 均方误差的优点:

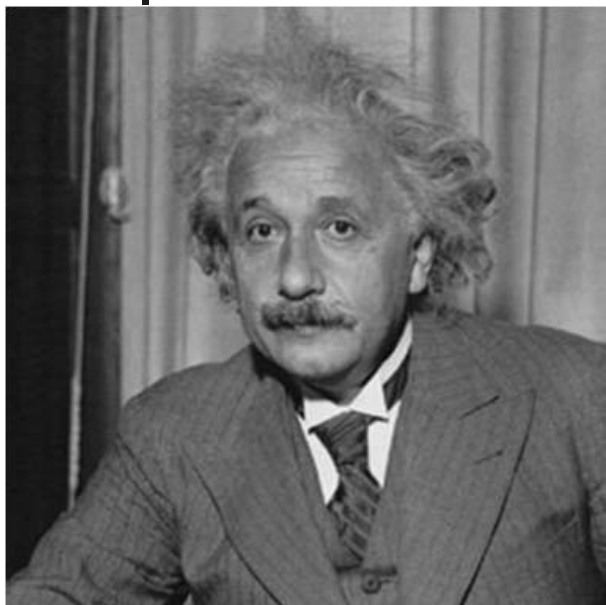
- 1) 具有明确的物理意义, 易于理解, 就是表示所有像素平均意义上的误差, 即误差信号的能量。
- 2) 计算简单, 只需要极小的计算量, 在所有的质量评价算法中是最高效的。
- 3) 是可微分的, 便于数学运算和分析, 因此适合作为质量优化的指标嵌入到图像处理系统中来指导优化算法的设计。
 - 比如在图像编码与视频编码系统 (H. 264、HEVC) 中, 常使用MSE作为失真测量并结合码率模型进行编码参数的优化设计, 以达到最优的率失真优化 (Rate Distortion Optimization, RDO) 性能, 即用最小的码率获得最小的失真。



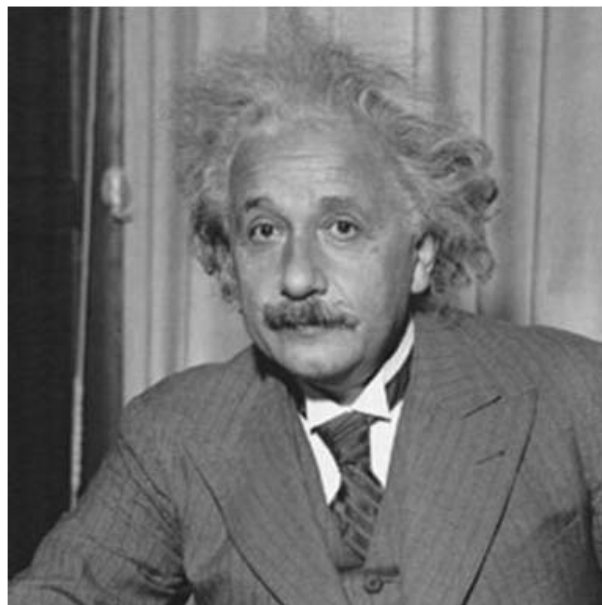
9.3.2 全参考图像质量评价

■ 1. 均方误差的缺点:

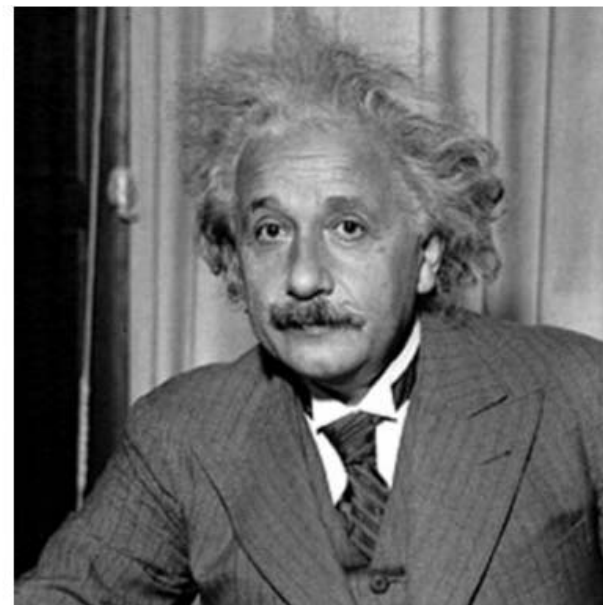
- 1) MSE完全忽视了二维图像信号内部像素之间的空间相关性即结构特性，将二维图像信号当作一维信号来处理。
- 2) MSE将误差信号与图像信号完全割裂开来，忽略了图像的局部特征对误差信号可见性的影响，即认为误差信号在不同的图像区域具有相同的视觉重要性。
- 3) MSE没有反映人眼观察图像的过程，与人眼评价结果的一致性较低。MSE相同的失真图像，其主观质量可能差别很大，反之，主观质量相似的失真图像，MSE可能差别很大。



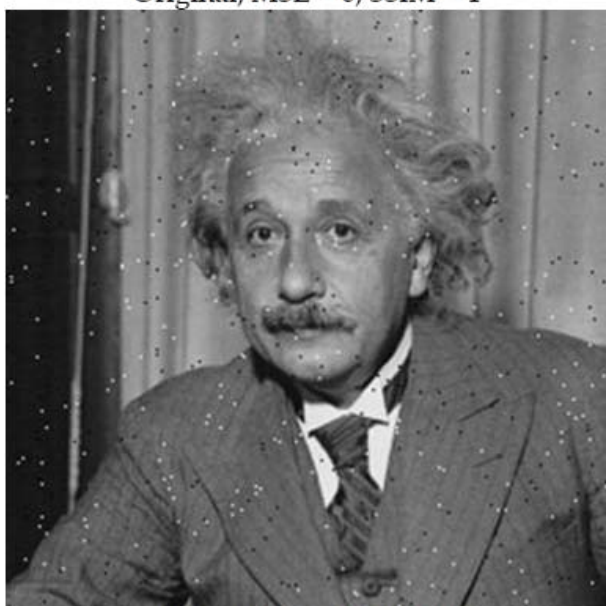
Original, MSE = 0; SSIM = 1



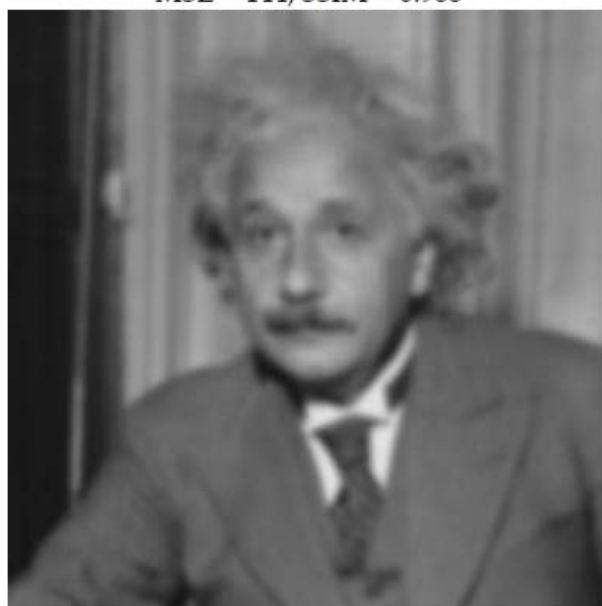
MSE = 144, SSIM = 0.988



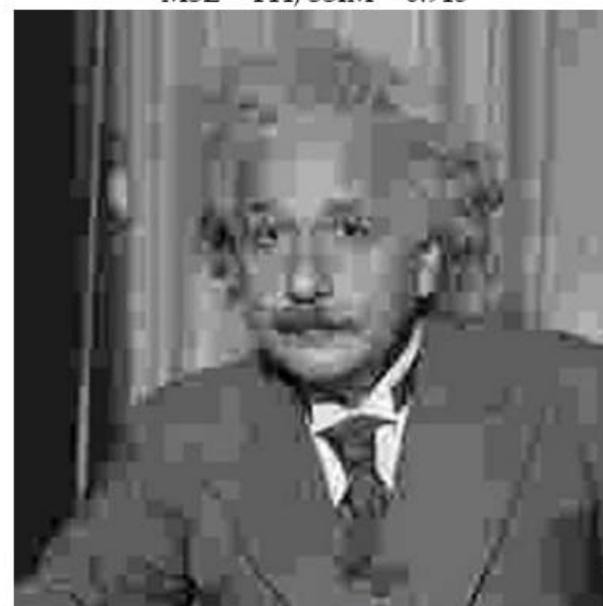
MSE = 144, SSIM = 0.913



MSE = 144, SSIM = 0.840



MSE = 144, SSIM = 0.694



MSE = 142, SSIM = 0.662



9.3.2 全参考图像质量评价

- 2. 峰值信噪比 (Peak Signal-to-Noise Ratio, PSNR)
 - 定义为信号最大可能的峰值功率与噪声信号的功率之比，因此可以看作是对MSE从量纲上的一种转换。
 - 设L表示灰度图像最大的像素值，对于常见的8比特 (bit) 量化的灰度图像， $L=2^8-1=255$ ，则PSNR的计算公式为：

$$PSNR = 10 \cdot \lg \left(\frac{L^2}{MSE} \right) = 10 \cdot \lg \left(\frac{L^2}{\frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N [f(i, j) - g(i, j)]^2} \right)$$

- 其中 $\lg()$ 表示以10为底的对数，PSNR的单位是分贝 (dB) 。

9.3.2 全参考图像质量评价

■ 2. 峰值信噪比

- PSNR与MSE成反比，与图像质量成正比。失真图像的MSE越低，PSNR越高，图像质量越好。

- 一般而言，当失真图像的PSNR在35dB以上时，人眼几乎觉察不到失真，图像质量较高；
- 当PSNR在28dB到35dB之间时，失真图像会呈现出一定程度的差异，图像质量一般；
- 当PSNR在28dB以下时，图像质量的降质较为明显，人眼观察会感觉到不舒适感。

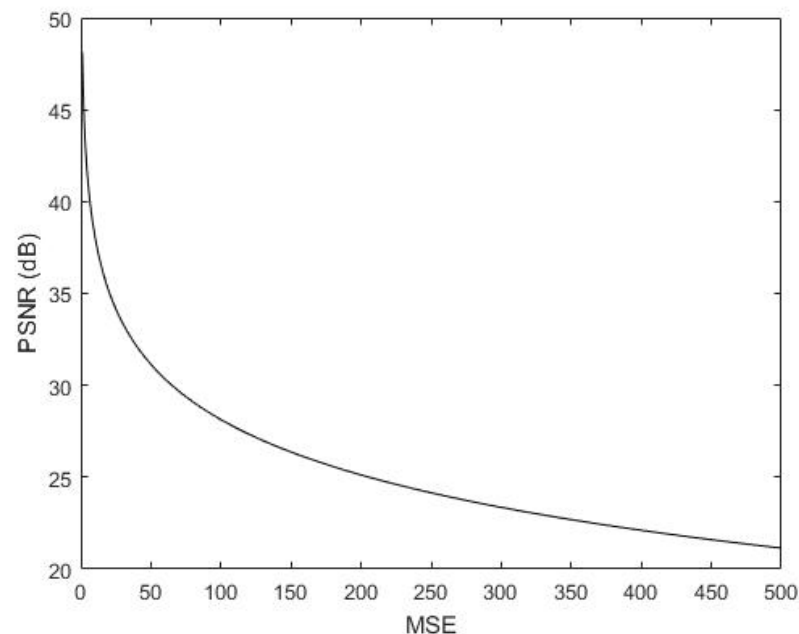


图9-9 PSNR与MSE的关系曲线



9.3.2 全参考图像质量评价

- 3. 结构相似性测量 (Structure Similarity, SSIM)
 - SSIM于2004年由王舟等人提出，**主要基于人眼对图像局部结构信息变化敏感的特性**，同时考虑了图像的亮度与对比度变化，结合三方面的相似性测量进行质量预测。
 - 根据图像结构的局部化特性，一般是先对图像分块（如 8×8 块）测量其结构相似性，再对所有块加权平均得到整个图像的SSIM值。

9.3.2 全参考图像质量评价

■ 3. 结构相似性测量 (Structure Similarity, SSIM)

- 设x和y分别为参考图像X与失真图像Y对应位置的划分图像块，设图像块大小相同且其中像素的个数为N，则定义：

亮度相似性：

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}$$

$$\mu_x = \frac{1}{N} \sum_{i=1}^N x_i$$

对比度相似性：

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}$$

$$\sigma_x = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)^2}$$

结构相似性：

$$s(x, y) = \frac{2\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}$$

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y)$$

- 结合亮度、对比度与结构相似性得到SSIM：

$$SSIM(x, y) = [l(x, y)]^\alpha \cdot [c(x, y)]^\beta \cdot [s(x, y)]^\gamma$$



9.3.2 全参考图像质量评价

■ 3. 结构相似性测量 (Structure Similarity, SSIM)

$$SSIM(x, y) = [l(x, y)]^\alpha \cdot [c(x, y)]^\beta \cdot [s(x, y)]^\gamma$$

- 其中 α 、 β 、 γ 是大于零的常数用于调整亮度、对比度与结构信息对质量评价的相对重要程度。
- 可认为三者的重要性相同，设置 $\alpha=\beta=\gamma=1$ 并且 $C_3 = C_2/2$ ，则上式为：

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

- 常数 C_1 与 C_2 根据经验设置为 $C_1 = (K_1L)^2$ 及 $C_2 = (K_2L)^2$ ，其中 L 为图像最大可取的像素值，对于8比特量化的灰度图像，其值为255，而 K_1 与 K_2 的取值分别为0.01和0.03。



9.3.2 全参考图像质量评价

■ SSIM的性质:

- 1) 对称性: $SSIM(x, y) = SSIM(y, x)$ 。
- 2) 有界性: $SSIM(x, y) \leq 1$ 。
- 3) 具有唯一的最大值: 当且仅当 $x=y$ 时, 才有 $SSIM(x, y)=1$ 。

■ 整个图像的SSIM计算:

- 当对应位置图像块的SSIM计算之后, 整个图像X与Y之间的SSIM可根据各图像块的重要程度加权平均来计算.

$$SSIM(X, Y) = \frac{\sum_{j=1}^M \omega_j(x_j, y_j) \cdot SSIM(x_j, y_j)}{\sum_{j=1}^M \omega_j(x_j, y_j)}$$

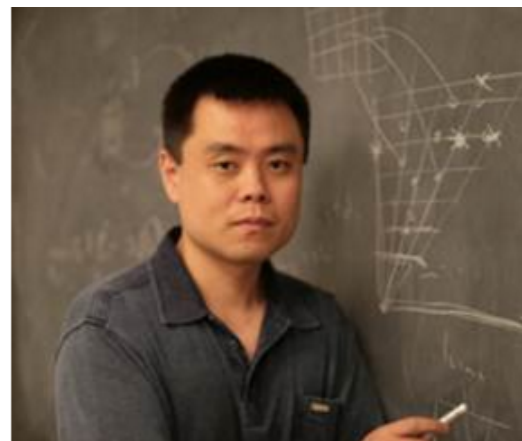
$$SSIM(X, Y) = \frac{1}{M} \sum_{j=1}^M SSIM(x_j, y_j)$$

- M为图像块数, $\omega_j(x_j, y_j)$ 为第j个图像块的重要程度。

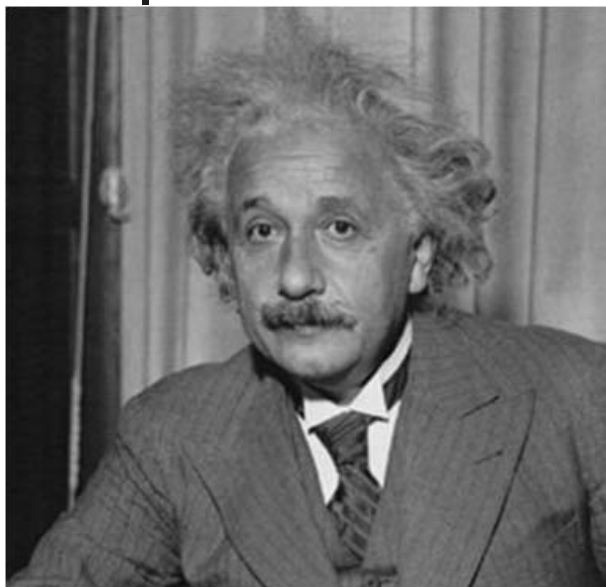
9.3.2 全参考图像质量评价

- **SSIM的优点：**在质量预测的准确性和计算高效性方面取得了较好的平衡，自提出以来获得了广泛认可，被应用到图像处理的多个领域。
 - 1) 基于自顶向下的方式利用了HVS对图像质量的感知规律。
 - 2) 相比MSE和PSNR，其评价结果具有更高的准确性。
 - 3) 具有较低的复杂度，便于嵌入到图像处理系统中来评价质量或优化算法。

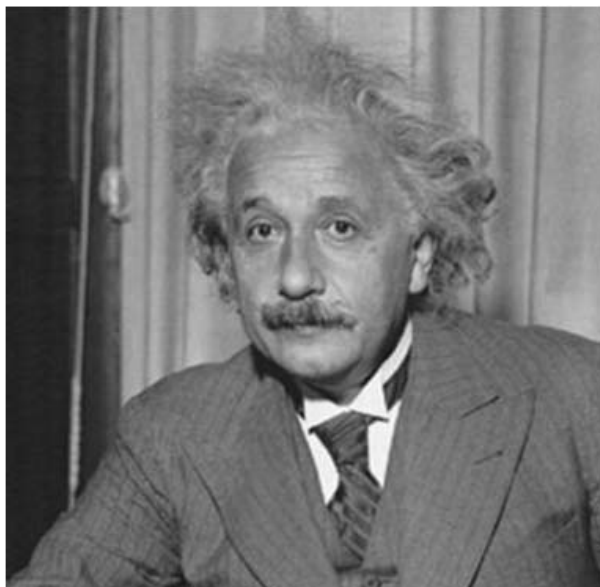
Zhou Wang, PhD, PEng, FIEEE, FCAE
University Research Chair, Professor
[Dept. of Electrical & Computer Engineering](#)
[University of Waterloo](#)
Office: E5-5113
Phone: (519)888-4567 ext. 35301
Fax: (519)746-3077
Email: zhou.wang@uwaterloo.ca
Mailing Address:
Dept. of ECE, Univ. of Waterloo
200 University Ave. W.
Waterloo, ON N2L 3G1, Canada



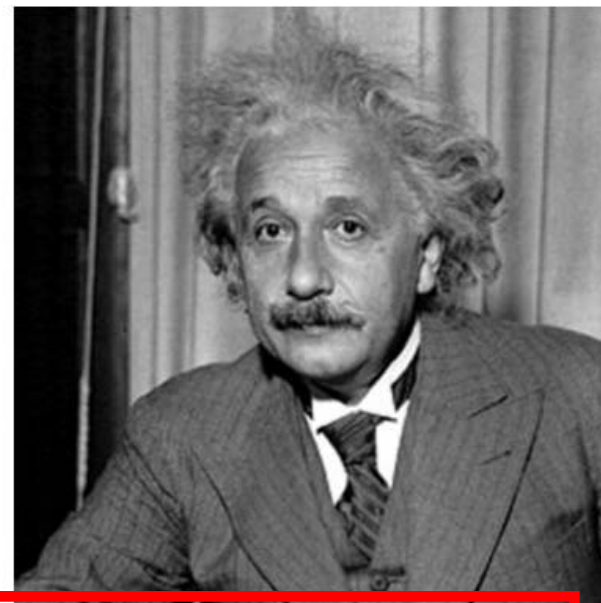
2014 - **Fellow of IEEE** (IEEE, for contributions to perceptual image processing and quality assessment)



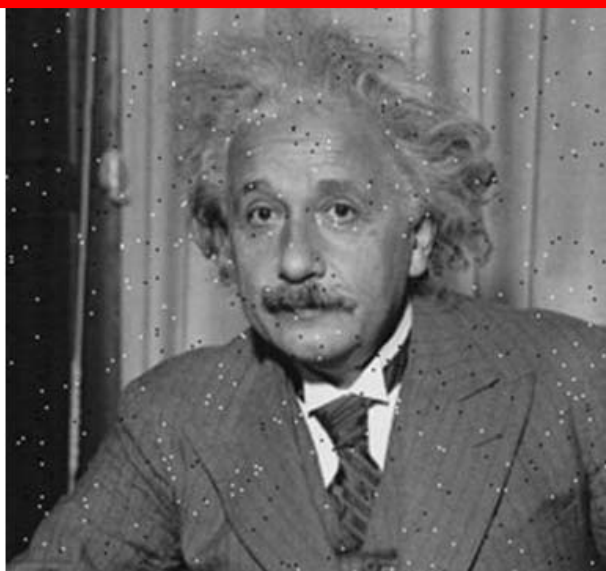
Original, MSE = 0; SSIM = 1



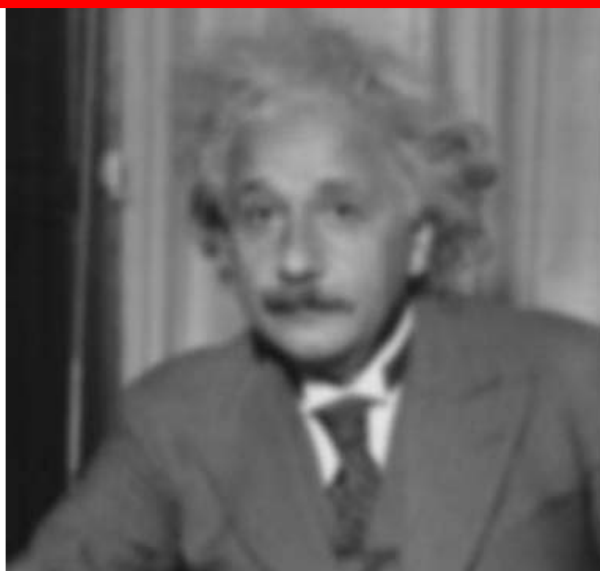
MSE = 144, SSIM = 0.988



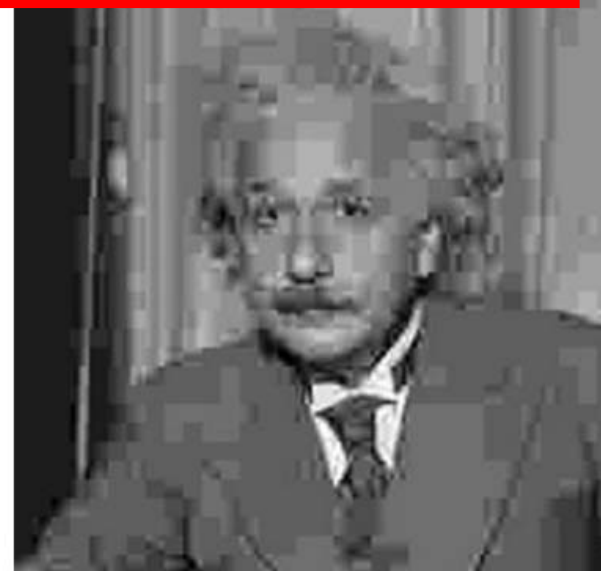
MSE = 144, SSIM = 0.913



MSE = 144, SSIM = 0.840



MSE = 144, SSIM = 0.694



MSE = 142, SSIM = 0.662



9.3.2 全参考图像质量评价

■ 4. SSIM算法的扩展

- (1) 多尺度结构相似性 (Multi-Scale SSIM)
- 考虑观察距离对图像细节或质量感知的影响，在相似性测量之前，先将图像变换到不同尺度，以模拟不同的观察距离。
- 生成多种尺度图像的方法，一般是通过低通滤波再下采样的方式得到类似于“金字塔”的多尺度图像描述。
- 将原始图像大小作为尺度1，先进行高斯低通滤波或平均滤波，再对滤波后的图像进行因子为2的下采样，即宽度和高度都降为原来的一半，得到尺度2图像，以此进行K-1次，可得到尺度K图像。

9.3.2 全参考图像质量评价

■ 4. SSIM算法的扩展

- (1) 多尺度结构相似性 (MS-SSIM)
- 在参考图像与失真图像的第j个尺度上，计算对比度与结构信息的相似性测量，而只在尺度K上计算亮度相似性。
- 最后联合不同尺度的相似性测量得到失真图像的MS-SSIM测量值。

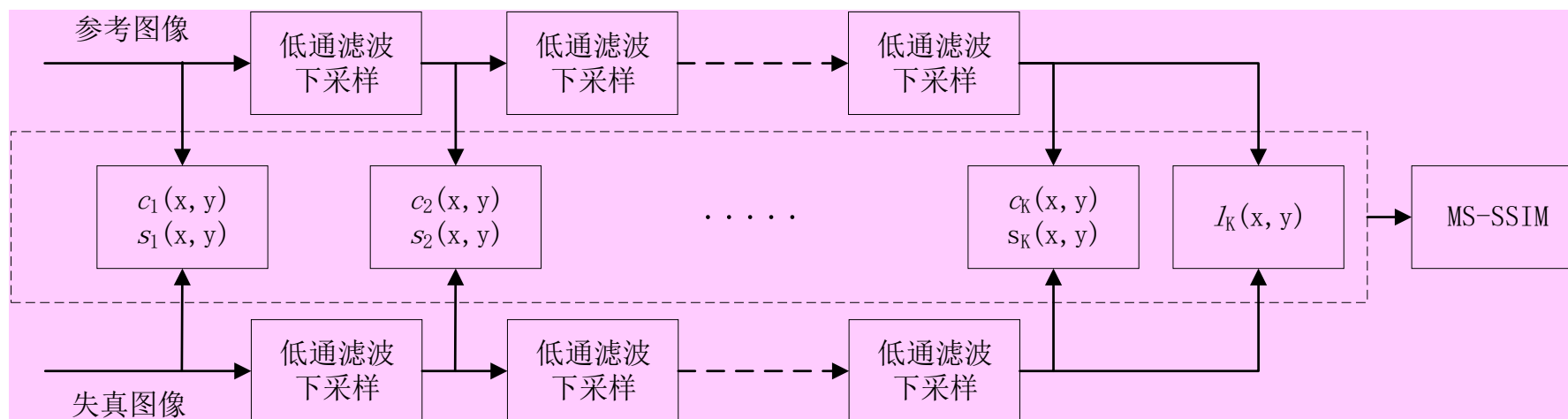


图9-10 MS-SSIM算法框架



9.3.2 全参考图像质量评价

- 4. SSIM算法的扩展
 - (1) 多尺度结构相似性 (MS-SSIM)
 - 计算公式:

$$MS-SSIM(X, Y) = [l_K(X, Y)]^{\alpha_K} \cdot \prod_{j=1}^K [c_j(X, Y)]^{\beta_j} [s_j(X, Y)]^{\gamma_j}$$

- 以K=5为例，将不同尺度的结构信息权重之和规范化为1，根据实验确定的不同尺度的各分量权重经验值为 $\beta_1=\gamma_1=0.0448$ ， $\beta_2=\gamma_2=0.2856$ ， $\beta_3=\gamma_3=0.3001$ ， $\beta_4=\gamma_4=0.2363$ ， $\alpha_5=\beta_5=\gamma_5=0.1333$ 。
- 可见，中间尺度的权重较大，其余尺度的权重较小，这与HVS特性是相符的，即人眼对中间尺度的图像信息最为敏感，随着尺度变大或变小，人眼对图像信息的敏感性逐渐变小。



9.3.2 全参考图像质量评价

■ 4. SSIM算法的扩展

- (2) 复小波结构相似性 (Complex Wavelet SSIM)
- 为了克服失真图像与参考图像的像素位置空间不对齐时（比如存在平移、旋转、缩放等几何失真时）评价结果不准确的问题；
- CW-SSIM首先对参考图像与失真图像进行复数小波的可操作金字塔变换 (Steerable Pyramid Transform)，将图像分解为多个子带，对各个子带的复数小波系数分块（如 7×7 块）测量相似性，最后对所有子带的分块相似性进行平均得到总体的相似性值。

$$CW-SSIM(c_x, c_y) = \frac{2 \left| \sum_{i=1}^N c_{x,i} c_{y,i}^* \right| + K}{\sum_{i=1}^N |c_{x,i}|^2 + \sum_{i=1}^N |c_{y,i}|^2 + K}$$



9.3.2 全参考图像质量评价

■ 4. SSIM算法的扩展

- (2) 复小波结构相似性 (Complex Wavelet SSIM)
- 其中, $\mathbf{c}_x = \{c_{x,i} | i=1,2,\dots,N\}$ 与 $\mathbf{c}_y = \{c_{y,i} | i=1,2,\dots,N\}$ 分别是参考图像与失真图像经复小波变换后从相同子带的相同位置分块提取的小波系数, c^* 表示系数 c 的复数共轭, K 是为防止分母接近于零测量值不稳定而设置的小的正数。
- CW-SSIM的计算形式与SSIM类似, 是从空域到复小波域的扩展。
- 相比MSE与SSIM, CW-SSIM在失真图像与参考图像空间位置不对齐时仍能得到稳定准确的质量评价结果, 在进行质量评价之前不需要采用复杂的算法完成图像配准, 有效降低了质量评价的复杂度和难度。



9.3.2 全参考图像质量评价

■ 4. SSIM算法的扩展

- (2) 复小波结构相似性 (Complex Wavelet SSIM)
- CW-SSIM对几何失真鲁棒的依据在于：
 - 1) 空域中轻微的平移、旋转和缩放在复小波域中表现为全部小波系数一致的相位变化 (Phase Changes) ；
 - 2) 相比频域系数的幅度，频域系数的相位模式携带了更多关于图像局部结构的信息。



9.3.2 全参考图像质量评价

- 5. 信息保真度准则 (Information Fidelity Criterion, IFC)
 - 从中高层语义级进行质量判定，与人眼感知机理较为一致，是另一类有代表性的质量评价方法。
 - 这类方法从信息论的观点出发，将图像失真的过程看作是参考图像的信息经过易错信道传输后信息丢失的过程。
 - 在信道传输过程中，引入的失真越大，则信息丢失的越多，接收到的图像中保留的信息越少，图像质量越低；反之亦然。

9.3.2 全参考图像质量评价

■ 5. 信息保真度准则（IFC）

- 设信道的输入（发送端）为无失真的参考图像X，信道的输出（接收端）为失真图像Y；
- 基于IFC的方法首先对图像信源特性与信道失真特性合理建模，之后通过测量失真图像Y与参考图像X之间的互信息（Mutual Information）的大小来定量地确定失真图像的质量。

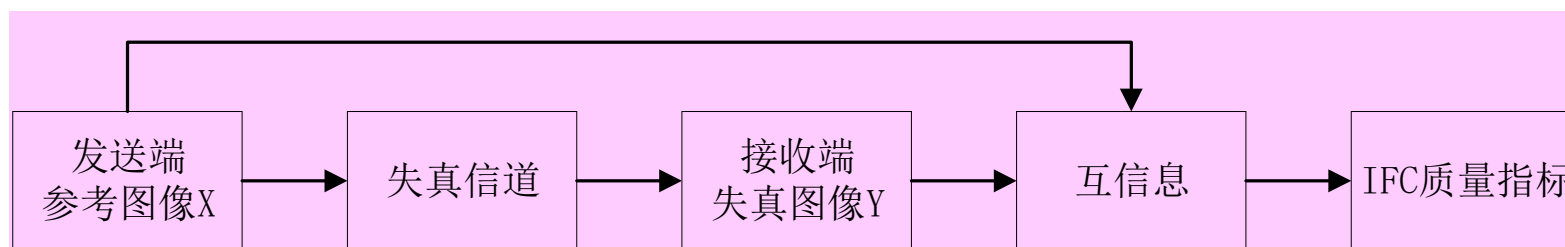


图9-11 基于信息保真度质量评价方法框架



9.3.2 全参考图像质量评价

■ 5. 信息保真度准则（IFC）

- 1) 图像信源模型：
- 对图像的建模是基于自然场景统计特性（Natural Scene Statistics, NSS），即自然场景中的图像与视频信号在所有可能的信号空间中仅仅占据很小的一个子空间，这个子空间具有一定的结构，可以通过构建适当的模型及其统计特征进行描述。
- 自然无失真的图像在某些统计属性或特征参数上具有规律性，这种规律性不依赖于具体的图像内容；
- 而失真图像在此模型上将使统计规律发生变化，从而偏离自然图像所构成的子空间。



9.3.2 全参考图像质量评价

■ 5. 信息保真度准则（IFC）

- 1) 图像信源模型：
- 先对图像进行多尺度多方向的小波变换以模拟HVS对图像信号的分解，之后使用高斯尺度混合（GSM）模型在小波变换域对具有自然图像统计特性的子带系数进行建模。
- 记对参考图像X做金字塔小波分解后子带k的小波系数向量表示为 $C_k = \{C_{ik} | i=1, 2, \dots, N_k\}$ ，其中 N_k 表示子带k的小波系数的个数， C_k 服从 GSM 分布；

$$C_k = S_k U_k = \{S_{ik} U_{ik}, i = 1, 2, \dots, N_k\}$$

- 其中 S_k 是正数组成的随机向量， U_k 是服从均值为零方差为 $\sigma_{U_k}^2$ 的高斯分布的随机向量， S_k 与 U_k 相互独立。



9.3.2 全参考图像质量评价

■ 5. 信息保真度准则（IFC）

- 2) 信道失真模型:
- 在小波变换域对子带系数用简单的信号衰减与加性高斯噪声来模拟;

$$D_k = G_k C_k + V_k = \{g_{ik} C_{ik} + V_{ik} \mid i = 1, 2, \dots, N_k\}$$

- C_k 表示参考图像X小波分解后子带k的系数随机向量,
- D_k 表示失真图像Y小波分解后相应子带k的系数随机向量,
- G_k 是一个确定性的随机向量来模拟信号衰减, 对应于子带系数能量衰减导致的细节模糊失真,
- V_k 是均值为零方差为 $\sigma_{V_k}^2$ 的随机向量来模拟加性高斯噪声, 对应于图像的噪声失真。



9.3.2 全参考图像质量评价

■ 5. 信息保真度准则（IFC）

- 根据条件互信息和高斯分布信息熵的定义，计算一个子带内条件互信息的公式为：

$$\begin{aligned} I(C_k; D_k | s_k) &= \sum_{i=1}^{N_k} I(C_{ik}; D_{ik} | s_{ik}) \\ &= \sum_{i=1}^{N_k} (h(D_{ik} | s_{ik}) - h(D_{ik} | C_{ik}, s_{ik})) \\ &= \frac{1}{2} \sum_{i=1}^{N_k} \log_2 \left(1 + \frac{g_{ik}^2 s_{ik}^2 \sigma_{U_k}^2}{\sigma_{V_k}^2} \right) \end{aligned}$$

- 失真图像的总体质量指标（IQI）由所有子带条件互信息之和得到；

$$IQI(X, Y) = I(C; D | s) = \sum_{k=1}^K I(C_k; D_k | s_k)$$



9.3.2 全参考图像质量评价

- 5. 信息保真度准则（IFC）的优缺点
 - IQI是质量属性，而不是失真属性。
 - 其最小值为0，表示失真图像丢失了参考图像的所有信息；最大值可以为无穷大，表示失真图像没有丢失任何信息。因此IQI的值越大，则失真图像的质量越好。
 - IQI的质量评价性能相比MSE与PSNR要好很多，比单尺度结构相似性（SSIM）稍好。
 - 但IQI的计算复杂度稍大，主要计算花费在图像的多尺度多方向的小波分解上，总体计算时间大概是SSIM的5~10倍。

9.3.3 半参考图像质量评价

- 1. 半参考IQA仅利用参考图像的部分特征
 - 半参考IQA利用参考图像的部分特征或参数作为评价依据，其准确性一般低于全参考IQA方法，且依赖于所参考的特征数据量的大小。
 - 将图像失真看作参考图像通过易错信道信息丢失的过程，从参考图像提取的特征通过辅助信道传输到接收端，接收端通过提取失真图像的特征并与接收到的特征作比较，即可评价失真图像的质量。

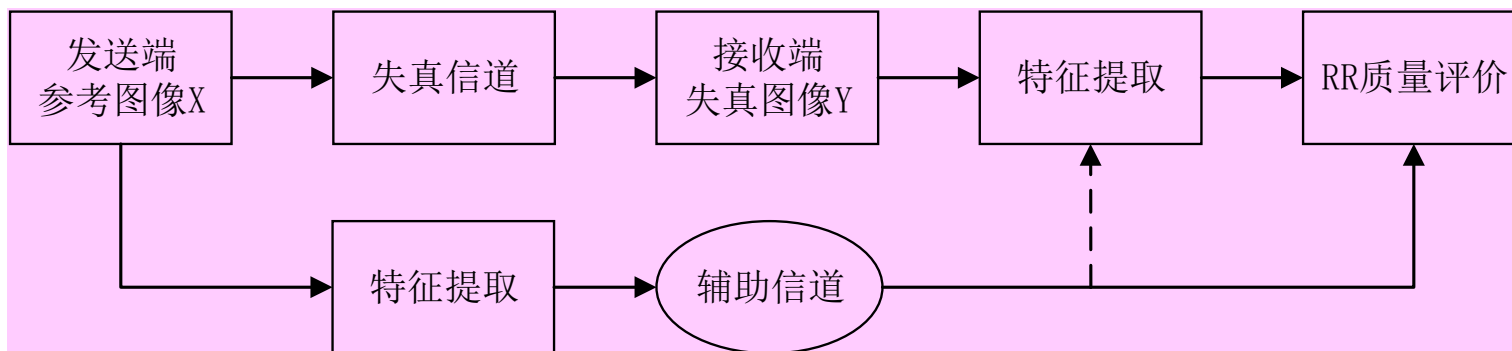


图9-12 半参考IQA方法框架

9.3.3 半参考图像质量评价

■ 1. 半参考IQA



- 在评价半参考IQA方法的性能时，常把质量预测的准确性和所参考的特征数据量的大小综合考虑，即希望利用较少的特征数据量来得到较好的质量预测准确性。
- 半参考IQA中所提取的特征要满足下列特性：
 - 1) 与人眼感知质量紧密相关；
 - 2) 能准确高效地表达参考图像与失真图像的信息；
 - 3) 对不同类型不同程度的失真具有不同的敏感性。



9.3.3 半参考图像质量评价

- 1. 基于熵差分（ED）的半参考IQA
 - 使用高斯尺度混合（GSM）模型对图像小波分解后的子带系数进行建模，通过测量失真图像与参考图像相应子带之间的小波系数熵的差分来评价失真图像的质量。
 - 主要步骤：
 - 1) 将参考图像X与失真图像Y做多尺度多方向的小波分解；
 - 2) 对小波系数加上高斯噪声来模拟人眼观察图像时由于视觉神经的处理而引入的不确定性；
 - 3) 计算参考图像与失真图像子带k块m的条件熵；
 - 4) 计算相应小波子带之间的熵差分作为半参考IQA的指标。

$$RRED_k^{M_k} = \frac{1}{L_k} \sum_{m=1}^{M_k} \left| \gamma_{mk}^r h(C'_{mk} | s_{mk}) - \gamma_{mk}^d h(D'_{mk} | t_{mk}) \right|$$



9.3.3 半参考图像质量评价

- 1. 基于熵差分（ED）的半参考IQA
 - 使用高斯尺度混合（GSM）模型对图像小波分解后的子带系数进行建模，通过测量失真图像与参考图像相应子带之间的小波系数熵的差分来评价失真图像的质量。
 - 主要步骤：
 - 1) 将参考图像X与失真图像Y做多尺度多方向的小波分解；
 - 2) 对小波系数加上高斯噪声来模拟人眼观察图像时由于视觉神经的处理而引入的不确定性；
 - 3) 计算参考图像与失真图像子带k块m的条件熵；
 - 4) 计算相应小波子带之间的熵差分作为半参考IQA的指标。



9.3.4 无参考图像质量评价

- 1. 无参考IQA比全参考、半参考IQA更具实用价值；
 - 对计算机来讲，无参考IQA是最难的；但对人来讲，很容易。
 - 无参考IQA的基本思路：
 - 认为无失真的自然图像在空域或变换域（如小波域、DCT域等）具有某些统计特征，而不同类型不同程度的失真将使这些统计特征发生相应的变化，因此通过对失真图像提取有效的特征，并与人工标定好的主观质量分（MOS）一起组成训练集，使用机器学习的方法得到训练集中特征与质量分的映射关系；
 - 之后对要评价质量的测试图像，通过提取相应的特征，用训练好的映射关系即可得到其主观评价分。

9.3.4 无参考图像质量评价

- 1. 基于DCT系数特征的无参考IQA方法主要步骤：
 - 1) 在训练阶段，由失真图像及其主观评价分组成训练集。对失真图像，先进行高斯低通滤波和下采样，生成多尺度的图像表达，之后对图像分块（如 5×5 块）做2维的DCT变换；
 - 2) 对DCT系数块分别按照方向特性与频率特性划分为不同的子带，按方向划分可捕获方向特性，按频率划分可捕获低频、中频与高频信息的特性；对子带系数用广义高斯分布（GGD）拟合；

DC	C ₁₂	C ₁₃	C ₁₄	C ₁₅
C ₂₁	C ₂₂	C ₂₃	C ₂₄	C ₂₅
C ₃₁	C ₃₂	C ₃₃	C ₃₄	C ₃₅
C ₄₁	C ₄₂	C ₄₃	C ₄₄	C ₄₅
C ₅₁	C ₅₂	C ₅₃	C ₅₄	C ₅₅

(a) 按方向划分

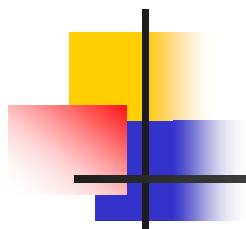
DC	C ₁₂	C ₁₃	C ₁₄	C ₁₅
C ₂₁	C ₂₂	C ₂₃	C ₂₄	C ₂₅
C ₃₁	C ₃₂	C ₃₃	C ₃₄	C ₃₅
C ₄₁	C ₄₂	C ₄₃	C ₄₄	C ₄₅
C ₅₁	C ₅₂	C ₅₃	C ₅₄	C ₅₅

(b) 按频率划分



9.3.4 无参考图像质量评价

- 1. 基于DCT系数特征的无参考IQA方法主要步骤：
 - 3) 对每一个DCT块提取DCT频域系数幅度的变化特征；
 - 4) 对每一个DCT块提取高频子带与中频及低频子带能量比率特征，来描述局部块内子带间的能量变化；
 - 5) 特征提取后，与主观质量分组成训练特征集，用机器学习的方法学习特征与质量分的函数关系；
 - 6) 在测试阶段，对要评价的失真图像，使用与训练阶段相同的方法提取特征，根据训练好的特征与质量分的函数关系，得到失真图像的质量评分。
- 目前有些无参考IQA的准确性已经超过简单的MSE与PSNR，甚至能达到或超过SSIM的预测性能。



第9章 图像与视频的质量评价

- 9.1 常见的图像与视频失真类型
- 9.2 图像与视频质量的主观评价
- 9.3 图像质量的客观评价
- **9.4 视频质量的客观评价**
- 9.5 MATLAB编程实例
- 9.6 小结



视频质量评价的难点

- 1. 视频质量评价（VQA）比IQA更加困难，体现在以下三点：
 - 1) 处理的信号更加复杂，视频比图像多了一维时域信息（即运动信息），视频信号本身特性的建模更为困难；
 - 2) 视频中既存在空域失真（如压缩方块效应、振铃效应、模糊等）又存在时域失真（如抖动效应、鬼影效应等），这些失真之间还存在相互影响；
 - 3) 人眼对视频信号的感知机理相比对静态图像信号的感知机理复杂的多，人眼如何理解视频尤其是其中的运动信息目前还没有足够精确的结论。
 - 因此，目前VQA方法的研究还于初级阶段，其质量评价的准确性远不及IQA的准确性高。



9.4.1 全参考视频质量评价

- VQA关键问题：是对运动信息进行有效的提取和利用。
 - 1) 显式运动信息提取：
 - 光流法：提取的运动信息精度高，但计算复杂；
 - 运动估计：提取的运动信息精度稍低，但可接受，且计算简单。
 - 2) 隐式运动信息提取：
 - 将视频看作2维图像加上1维时间组成的3维信号，对视频数据进行空时3维滤波来提取运动信息，并分析其统计特性。
 - 在VQA中，显式与隐式利用运动信息的应用都很多，也有各自优缺点，在VQA时可根据需要进行选择。

9.4.1 全参考视频质量评价

■ 1. 基于运动信息加权的VQA

- 视频可看作由一帧帧的图像及其之间的运动信息所构成，VQA可利用IQA的方法先对每一帧局部空域的质量进行测量，之后使用时域运动信息作为权重因子来调节块层与帧层的空域质量分，最后进行融合得到序列的评价质量。

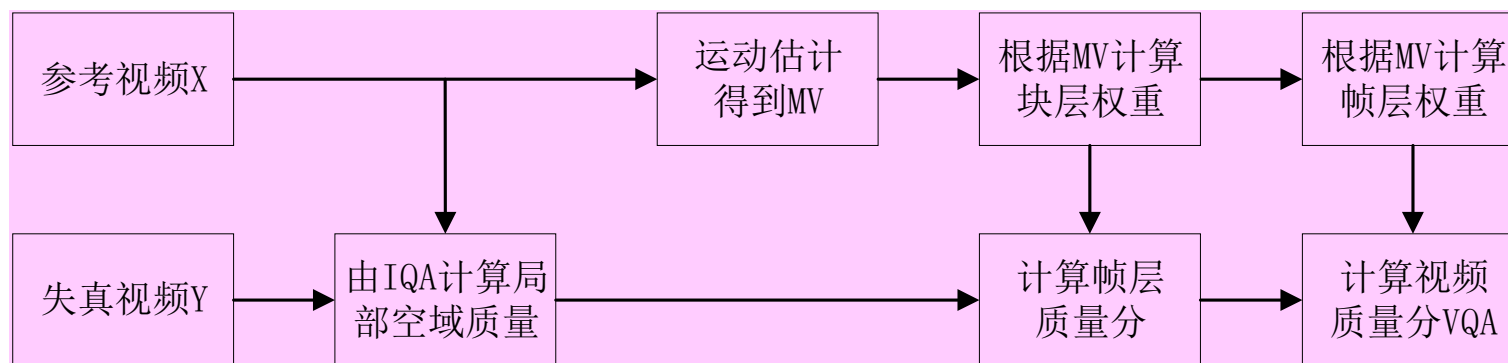


图9-15 基于运动信息加权的VQA框架

9.4.1 全参考视频质量评价

■ 1. 基于运动信息加权的VQA主要步骤:

- ① 对参考视频X与失真视频Y，用针对图像的IQA计算失真视频各帧的空域质量，常见的IQA方法都可使用；
- ② 基于参考视频X，用运动估计或光流法来计算相邻帧之间的运动信息；
- ③ 根据HVS对运动感知的特性，由块层的运动信息来计算块层视觉权重；如 $W_{i,j} = \sqrt{(MV_{i,j,x})^2 + (MV_{i,j,y})^2}$
- ④ 根据块层质量的加权和来计算帧层的质量评分； $VF_i = \frac{\sum_{j=1}^M W_{i,j} MSE_{i,j}}{\sum_{j=1}^M W_{i,j}}$
- ⑤ 用类似的方法根据运动信息计算帧层的权重 W_i ，失真视频的质量评分对所有帧的质量分加权后得到。

$$VQ = \frac{\sum_{i=1}^N W_i \cdot VF_i}{\sum_{i=1}^N W_i}$$

9.4.1 全参考视频质量评价

■ 2. 基于运动的视频空时质量评价

- 综合利用空域和时域的信息，并根据从参考视频计算出的运动轨迹来评估运动造成的视觉失真及其对空域与时域失真的影响。

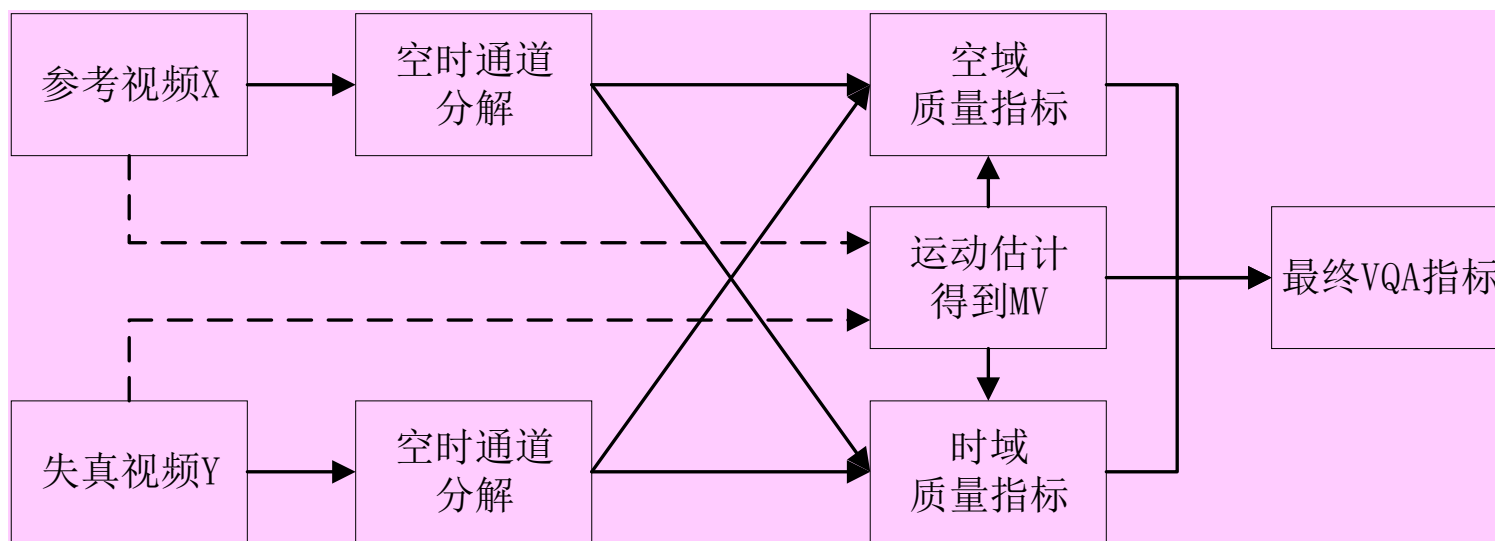


图9-16 基于运动的视频空时VQA框架



9.4.1 全参考视频质量评价

■ 2. 基于运动的VQA主要步骤:

- 1) 对参考视频X与失真视频Y, 用空时滤波器分解为不同的空时通道, 比如小波分解、Gabor分解等, 以模拟HVS对视频的分析过程;
- 2) 用光流法或运动估计得到参考视频X与失真视频Y各自的运动信息, 计算运动失真;
- 3) 根据参考视频X与失真视频Y经空时分解后的系数, 考虑运动对空域失真的调节作用, 计算空域质量指标;
- 4) 联合参考视频X与失真视频Y经空时分解后的系数以及运动信息, 计算时域失真;
- 5) 最后, 结合视频空域失真、时域失真以及运动失真得到最后的视频质量评价指标。



9.4.2 半参考视频质量评价

■ 半参考VQA的难点

- 半参考VQA：利用参考视频的部分信息或特征来辅助完成失真视频的质量评价。
- 半参考VQA设计的方法学类似于半参考IQA，但除了提取空域特征，还必须提取时域特征以反映帧间失真的情况。
- 可以选择参考视频中的某些帧（或者帧中的某些子带以及子带块）作为辅助信息，实现半参考的VQA。
- 由此可见，半参考VQA的准确性依赖于所提取的空域与时域特征的质量和数量。



9.4.2 半参考视频质量评价

- 基于熵差分的半参考VQA方法主要步骤：
 - 1) 计算空域熵差分质量指标 (SRRED)：对参考视频X和失真视频Y的各帧分别做多尺度多方向的小波分解，并在各子带内分块计算空域的基于熵差分的质量指标；
 - 2) 计算时域熵差分质量指标 (TRRED)：对帧差分做多尺度多方向的小波分解，在各子带内分块计算时域的基于熵差分的质量指标；
 - 3) 联合空域与时域的基于熵差分的质量指标，得到最终的质量评价指标 (STRRED)。

$$STRRED_k^{M_k} = SRRED_k^{M_k} \cdot TRRED_k^{M_k}$$



9.4.3 无参考视频质量评价

■ 无参考VQA的难点

- 无参考VQA：是指仅依据失真视频本身进行质量的预测，没有参考视频的任何信息可供利用。
- 无参考VQA方法的设计原理与无参考的IQA方法类似，重点在于从失真视频中提取哪些特征来有效地描述空域与时域信息，以及如何将这些空时域特征与主观质量分联系起来。

9.4.3 无参考视频质量评价

■ 基于DCT特征的无参考VQA

- 同时利用显式和隐式的方式提取运动信息，通过对相邻帧差分做DCT变换来隐式分析运动特性，用运动估计的方法得到MV来显式描述运动的连贯性。

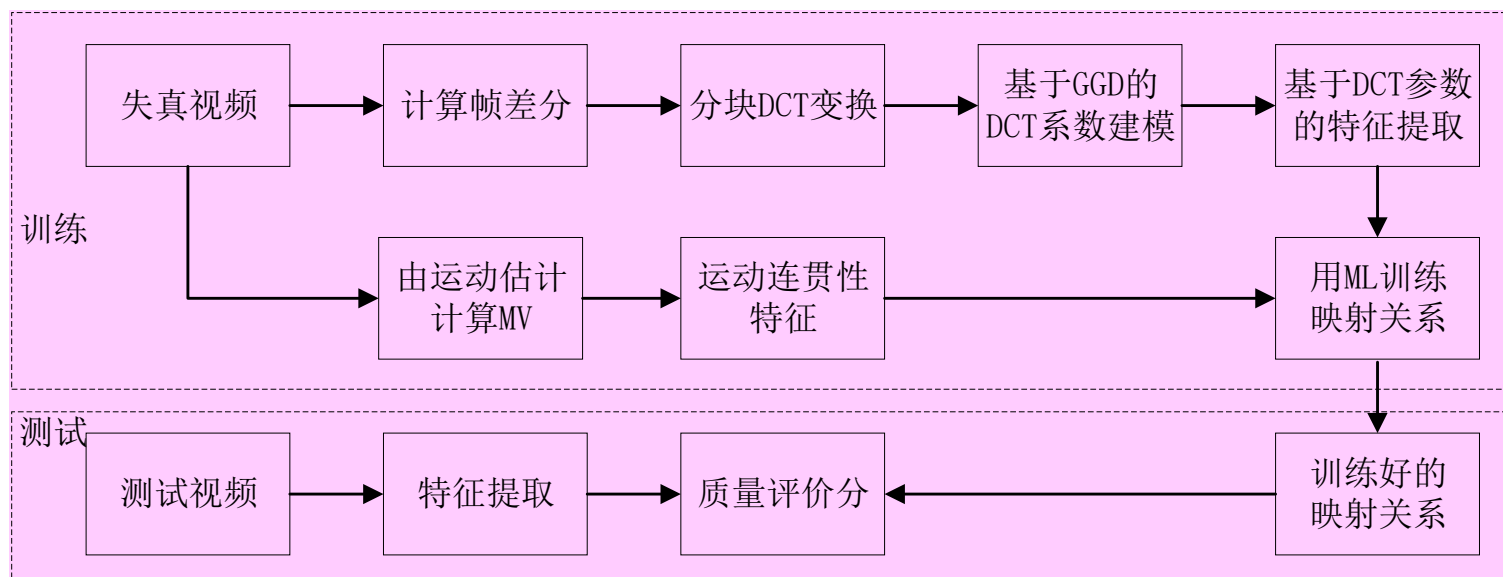
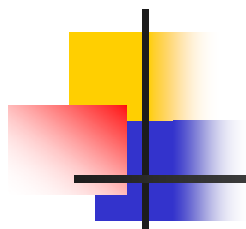


图9-17 无参考视频质量评价方法框架



9.4.3 无参考视频质量评价

- 基于DCT特征的无参考VQA主要步骤：
 - 1) 在训练阶段，由一些失真视频及其主观评价分组成训练集，计算帧差分的DCT变换系数，用广义高斯分布（GGD）进行建模；
 - 2) 用运动估计的方法计算块层的MV，提取运动连贯性特征；
 - 3) 将空时域特征与相应的主观质量分形成训练特征集，用机器学习（ML）的方法训练空时域特征与主观质量分的函数关系；
 - 4) 在测试阶段，对要测试的失真视频先提取相应的特征，之后使用训练好的函数关系来得到失真视频的质量评价分。



第9章 图像与视频的质量评价

- 9.1 常见的图像与视频失真类型
- 9.2 图像与视频质量的主观评价
- 9.3 图像质量的客观评价
- 9.4 视频质量的客观评价
- **9.5 MATLAB编程实例**
- 9.6 小结



9.5 MATLAB编程实例

- 在Matlab编程环境中调试执行SSIM的代码，并测试质量预测的准确性和计算复杂度；
- 1. SSIM的函数实现代码（`ssim.m`）
- 2. SSIM函数测试与结果可视化代码（`ssim_test.m`）



9.6 小结

- 本章首先介绍了人眼视觉特性中的对比敏感度和视觉掩盖效应的概念，以及常见的图像与视频失真类型。
- 然后，重点介绍了全参考IQA，包括基于信号保真度的均方误差和峰值信噪比，基于结构相似性的质量评价，以及基于信息保真度准则的评价方法。同时也介绍了半参考、无参考IQA，包括自然场景统计特性的原理和建模方法，以及常见的空域与频域特征提取方法。
- 最后，介绍了全参考、半参考以及无参考VQA，包括常见的时域特征提取方法。