

电子稳像中稳像质量评价方法研究

黄 晨¹, 王建军¹, 高 昕¹, 丁 晟²

(1. 北京跟踪与通信技术研究所, 北京 100094; 2. 装甲兵工程学院控制工程系, 北京 100072)

摘 要: 在使用电子稳像算法对抖动视频序列进行稳像后, 如何对稳定后的图像序列进行评估, 目前尚未有一个统一的评价标准。对目前常用的主、客观稳像图像评价方法进行介绍分析, 考虑不同的场景使用不同的评价方法。分析人眼视觉特性, 认为结合人眼视觉特性的评价方法能较为全面客观的反映稳像图像质量是未来图像质量评价的发展方向。为进一步研究稳像图像质量评价提供理论基础。

关键词: 电子稳像; 稳像评价; 保真度; 峰值信噪比; 人眼视觉特性

中图分类号: TP391.41 **文献标识码:** A **DOI:** 10.3969/j.issn.1001-5078.2013.05.001

Study on video stabilization performance assessment in electronic image stabilization

HUANG Chen¹, WANG Jian-jun¹, GAO Xin¹, DING Sheng²

(1. Beijing Institute of Tracking and Telecommunications Technology, Beijing 100094, China;

2. Department of Control Engineering, Academy of Armored Force Engineering, Beijing 100072, China)

Abstract: The main subjective and objective image stabilization evaluation methods are described and analyzed. Different evaluation methods should be used for different scenarios. By analyzing the human visual characteristics, the evaluation method combined with the human visual characteristics could objectively reflect the image quality of the stabilized image. It is the development trend of the image quality assessment. The study provides a theoretical basis for further research of stabilized image quality evaluation.

Key words: electronic image stabilization; stabilizing algorithm evaluation; fidelity; PSNR; human visual characteristics

1 引 言

稳像技术指隔离外界对成像系统的扰动, 保证图像信息采集、输出的完整性、流畅性, 其主要分为机械稳像、光学稳像、电子稳像。电子稳像^[1] (electronic image stabilization, EIS) 技术利用数字图像处理方法确定图像序列的偏移并进行补偿。与传统的稳像方法相比, 具有高精度、小体积、易操作等优点, 目前广泛应用于机器人遥控系统、无人车辆导航系统、视频侦察系统等。随着电子信息技术的发展, 利用电子稳像技术实现视频图像序列的稳定成为现代稳像技术的主要发展方向。

电子稳像系统一般包括三个主要过程, 即运动估计、运动滤波和运动补偿, 对应的关键技术为全局

运动估计、运动矢量滤波、运动补偿校正, 系统框图如图 1 所示。

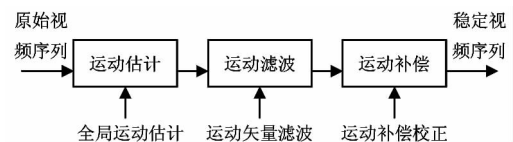


图 1 电子稳像系统框图

电子稳像算法的目的是消除视频中图像全局运动的随即抖动, 并同时保留摄像机的有意运动, 如水平、垂直方向的扫动。对于经过电子稳像处理后输

作者简介: 黄 晨(1987 -), 男, 硕士, 主要从事图像处理, 光学设备总体, 辐射特性测量, 电子稳像等研究。E-mail: raulhuang2010@gmail.com

收稿日期: 2012-10-16; **修订日期:** 2012-10-26

出的稳定视频序列的评价,即对稳像算法效果的评价是一个重要的环节,客观统一标准的稳像质量评价将有利于稳像算法的改进,促进该项技术的发展。但是由于视频不同的运动场景、运动方式及视频图像特征和其他客观环境因素的影响,使得目前稳像图像序列的评价无法有一个统一的标准。对现有的稳像图像质量评价方法进行介绍与分析,对静态背景和具有扫描运动或含有运动物体的图像序列的场景使用不同的评价方法。电子稳像是一个视觉寻优的过程,对抖动图像进行稳像处理,目的是为了保证视频序列的稳定性和人眼观察的流畅性,稳像质量的评价应考虑客观对抖动的滤除和人眼主观观测的特点。所列举的方法亦对数字图像质量评价有所参考。

2 主观评价方法

主观评价是最为直接的稳像评价方法。以个人作为观察者对某一抖动的视频序列,对其稳像前和稳像后的效果及图像序列品质进行评定。选取的观察者人数越多,统计样本越多,其统计结果就越能反映稳像效果。但此方法存在的不足在于评价过程繁琐,耗时长,人力成本高。个体观察者在评价时,由于受到周围环境、文化背景、个体差异等影响,人眼对同一图像序列中的不同区域有不同的兴趣程度,即人眼视觉特性中的视觉注意特性,使得个体的评价往往带有较强的主观性。

3 保真度(Fidelity)评价方法

3.1 均方差(MSE)方法^[2]

均方差(MSE)是两帧图像间相对应像素灰度的偏差值。它反映了图像序列变化的快慢和变化量的大小,其定义如下:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \| I_1(i, j) - I_0(i, j) \|^2 \quad (1)$$

式中, n, m 分别表示参考图像帧和当前帧的像素; $I_1(i, j)$ 和 $I_0(i, j)$ 分别表示图像序列补偿后的当前帧和参考帧在 (i, j) 点处像素的灰度值。对于进行稳像后的相邻两幅图像, $MSE(I_1, I_0)$ 值越小,说明两幅图像重合度越高。 $MSE(I_1, I_0) = 0$ 时,则表明两幅图像完全重合,但在实际应用中由于噪声、运动估计误差、光照变化影响等原因, MSE 值为0的情况不会出现。

3.2 差分图方法

基于均方差的原理,提出差分图方法,指利用相邻两帧的灰度图像进行差分相减,得到差分图。对于背景静止的抖动视频,若相邻帧的抖动被完全补

偿时,差分图的灰度值为0。同峰值信噪比一样,在稳像过程总存在其他影响因素,使得差分图的灰度值不为0。稳像后的相邻帧进行差分处理,差分图上残留的灰度值越小则表明稳像补偿越完全,稳像效果越好。基于差分图的原理,文献[3]提出一种方法:对待处理的视频进行稳定处理之后得到的视频称为回答视频,观察回答视频每一帧和正解视频中相应帧的差别,用正解帧以帧单位对回答帧进行合格与否的判定,求出矫正率,如果这一视频的矫正率达到90%,则这一待处理视频被此稳定算法成功的稳定。矫正率的计算方法为:

矫正率 = 评价的视频序列的合格帧数/评价的视频序列的全帧数。

合格帧(OK):在正解帧和回答帧的差分图像中没有出现边缘宽度超过3像素的对象。

不合格帧(NG):有上述边缘对象的帧。

OK数 + NG数 = 评价的动画的全帧数。

其中,以正解帧和回答帧的同一像素位置对等的亮度和两个色度值Y, Cb, Cr欧几里得距离差为32以上的像素为边缘。边缘宽度在3像素以下为合格。该评价方法只适合于静态背景下稳像效果评价。

3.3 峰值信噪比(PSNR)方法

峰值信噪比^[4](peak signal to noise ratio, PSNR),反映的是参考图像和当前图像之间的峰值信噪比,目前广泛应用在衡量图像质量。此方法本质上与均方差(MSE)方法相同,其定义为:

$$PSNR(I_1, I_0) = 10 \log \frac{255^2}{MSE(I_1, I_0)} \quad (2)$$

根据公式,对于灰度图像其灰度最大值为255,帧间的PSNR值越大代表图像的灰度差别越小,稳像的效果也越好。PSNR对于静态背景有很好的效果评价,但对于具有扫描运动或含有运动物体的图像序列,即使完全补偿图像的晃动量,稳像后图像仍不能重合,因此无法采用PSNR方法进行评价。此外PSNR方法未考虑到人眼的视觉特性,因为人眼的视觉对于误差的敏感度并不是绝对的,其感知结果会受到许多因素的影响而产生变化,如人眼对空间频率较低的对比差异敏感度比较高,对亮度对比差异的敏感度较高等。这些人眼视觉特性使得采用PSNR的评价无法和人眼看到、感知的视觉品质完全一致,可能出现在评价时PSNR较高的稳像图像较PSNR较低的视觉感受差。

3.4 帧间变换保真度(ITF)

帧间变换保真度(inter-frame transformation fi-

delity, ITF) 是基于峰值信噪比 PSNR 计算获得的, 其定义如下式:

$$\text{ITF} = \frac{1}{N-1} \sum_{i=1}^{N-1} \text{PSNR}(i) \quad (3)$$

文献[5]根据 ITF 对稳像结构进行客观评估, 实验结果表明适合于静态背景下的稳像评价。

3.5 帧间变换保真度差异 (DITF)

针对背景运动的情况下, 考虑连续帧间变化的差异来衡量稳像的结果, 基于帧间变换保真度 (ITF) 原理, 文献[5]提出使用帧间变换保真度差异 (difference of inter-frame transformation fidelity, DITF) 来对背景运动下稳像结果进行评估, 其定义如下:

$$\text{DITF}(I_t) = |\text{PSNR}(I_{t-1}, I_t) - \text{PSNR}(I_t, I_{t+1})| \quad (4)$$

其中, I_t 为参考帧; I_{t-1}, I_{t+1} 分别为 I_t 前一帧和后一帧。根据式(4), 若视频中摄像机运动越平滑, 连续的帧间变化越小, DITF 值应该越小, 由此考察稳像算法对摄像机运动中所含抖动量的滤除和平滑效果, 但对于含有变焦、目标尺度变化的场景并不适用。

4 结构相似 (SSIM) 法

Wang^[6]等人认为人眼视觉系统的主要功能是提取视场中的结构信息, 而基于结构失真的视频质量评价法—结构相似 (SSIM) 法, 该方法对两幅图像的亮度、对比度和相似度进行比较, 其公式定义为:

$$l(i, j) = \frac{2\mu_i\mu_j + C_1}{\mu_i^2 + \mu_j^2 + C_1} \quad (5)$$

$$c(i, j) = \frac{2\sigma_i\sigma_j + C_2}{\sigma_i^2 + \sigma_j^2 + C_2} \quad (6)$$

$$s(i, j) = \frac{\sigma_{ij} + C_3}{\sigma_i\sigma_j + C_3} \quad (7)$$

$$\text{SSIM}(i, j) = [l(i, j)]^\alpha [c(i, j)]^\beta [s(i, j)]^\gamma \quad (8)$$

式中, i, j 分别代表原始图像和失真图像; μ 是图像的均值; σ 是标准差; C_1, C_2, C_3 是常数, SSIM 使得测量结构信息的改变与感知图像质量的变化非常接近, 与主观评价有较好的相关性。此后, 很多学者对其进行不同的改进, 以适应不同的领域。文献[7]提出了基于灰度投影结构相似性 (GPASSIM) 的电子稳像质量评价方法。GPASSIM 评价的具体计算步骤如下:

(1) 将参考帧与当前帧的图像分成不重叠的若干子块, 计算每个子块的行、列灰度投影值, a 和 b

分别表示参考帧和当前帧中对应的子块。计算得到参考帧和当前帧子块的行、列灰度投影的均值 $\mu_{a1}, \mu_{b1}, \mu_{a2}, \mu_{b2}$ 和标准方差 $\sigma_{a1}, \sigma_{b1}, \sigma_{a2}, \sigma_{b2}$; 同时计算出参考帧和当前帧子块的灰度投影协方差 $\sigma_{ab1}, \sigma_{ab2}$ 。

(2) 根据公式(9)~式(11), 分别计算子块内行、列两个方向的亮度、对比度和结构的比较函数, 其比较函数可表示为:

亮度比较函数:

$$l(a, b) = \frac{2\mu_a\mu_b + D_1}{\mu_a^2 + \mu_b^2 + D_1}, D_1 = (K_1 Z)^2 \quad (9)$$

对比度比较函数:

$$c(a, b) = \frac{2\sigma_a\sigma_b + D_2}{\sigma_a^2 + \sigma_b^2 + D_2}, D_2 = (K_2 Z)^2 \quad (10)$$

结构比较函数:

$$s(a, b) = \frac{2\sigma_{ab} + D_2/2}{\sigma_a\sigma_b + D_2/2} \quad (11)$$

(3) 计算参考帧与当前帧的每个对应子块的行、列结构相似度, 同理可得算的到参考帧和当前帧两帧图像之间的行、列结构相似度。最后, 参考帧与当前帧图像的结构相似度可定义为:

$$\text{GPASSIM}(A, B) =$$

$$\sqrt{\text{GPASSIM}_1(A, B) \times \text{GPASSIM}_2(A, B)} \quad (12)$$

实验表明, GPASSIM 评价方法较 PSNR 具有更好的灵敏度, 有效提高了稳像视频质量评价的主客观性。

5 标准差方法

文献[8]提出了利用标准差的稳像算法效果评价的方法, 其原理是通过计算稳像补偿后图像场景运动的标准差, 以此作为稳像精度标准, 反映电子稳像算法的精确程度。

标准差在概率统计中作为统计分布程度 (Statistical Dispersion) 上的测量, 定义为方差的算术平方根, 表示反映组内个体间的离散程度。标准差越高, 表示实验数据越离散, 稳像效果越不理想。反之, 标准差越低, 代表稳像补偿越精确, 稳像效果越好, 其定义如下:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (13)$$

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}, n \in N, n > T \quad (14)$$

式(13)表示抖动差值数据的均值 $\bar{x}$, 式(14)表示抖动差值数据的标准差 σ 。实验结果表明该方法能够解决了电子稳像中在某些特殊情况下 (如相机扫描

运动或目标含运动小物体),利用均方差方法、峰值信噪比及差分图法等客观方法无法准确对稳像效果进行评价的问题。

6 基于随机性检验的稳像评价方法

文献[9]认为摄像机随机抖动所引起的图像全局运动属于一种无规律的、随机性的变化,这和连续、平滑的摄像机有意扫描运动所引起的视频全局运动有着明显的区别。因此,所拍摄视频的全局运动形式体现了摄像机运动的形式和程度,即视频全局运动的随机程度越强,则摄像机抖动越强,视频的稳定性越差,反之则摄像机抖动越小,视频的稳定性越好。基于此,提出根据稳定后视频全局运动的随机程度(随机性)来评价视频的稳定程度。由于现有的随机性检测对象是二进制序列,对于稳像图像评价需要将视频的全局运动转化为二进制序列,并且根据评价目标选择合适的检测指标来衡量视频中随机抖动的程度。为此需进行全局运动编码,其关键在于将编码所得到的二进制序列依然能反映出全局运动变化的随机程度。视频相邻图像的全局运动向量包含三个分量:旋转分量、水平平移分量、垂直平移分量,对全局运动向量的三个分量分别编码。具体的编码规则如下列所示:

(1)视频第一帧图像全局运动矢量的分量编码为0;

(2)如果全局运动矢量的分量为正,则相应分量的编码为1;

(3)如果全局运动矢量的分量为负,则相应分量的编码为0;

(4)如果全局运动矢量的分量为0,则相应分量的编码和前一帧图像同一分量的编码相同。

根据全局运动向量任一分量的符号对全局运动进行编码,反映了视频全局运动的随机变化。稳像后的视频越趋于稳定,则图像全局运动矢量的变化应该呈现一致、连续的状态,即全局运动中各分量的符号应该趋近一致,反之则呈现随机变化的状态。

7 基于人眼视觉特性的评价方法

文献[10]认为对于不同场景不同形式的抖动视频,其稳像后图像的散度(divergence)和抖动(jitter)之间存在平衡。结合人眼视觉特性,根据符合人眼视觉特性的抖动截止频率确定期望误差值,将1 Hz作为适合于人眼的截止频率。通过高通/低通滤波和截止频率 c ,将不需要的抖动运动分解为散度和抖动。低频部分是图像序列 i 的期望误差值,其平方代表了散度:

$$D_c = \frac{1}{\text{frames}} \sum_{i=1}^{\text{frames}} \{e_f \leq c(i)\}^2 \quad (14)$$

散度(divergence)是低频部分的平方,其平方根表示期望的偏移量:

$$E\{|e|\} \approx \sqrt{D} \quad (15)$$

同样的,高频部分的平方表示抖动量:

$$J_c = \frac{1}{\text{frames}} \sum_{i=1}^{\text{frames}} \{e_f \leq c(i)\}^2 \quad (16)$$

偏移量 e 表示了 x, y 轴、旋转角、缩放上理想位置参数与实际位置参数的差别。如,式(14)和式(16)可同时用光谱能量强度(PSD)函数计算。

$$\text{Jitter_attenuation} = J_{c_stabilized} / J_{c_original} \quad (17)$$

衰减系数 J 表示残留的抖动与原始抖动的关系,提供一个更不受原有运动约束的参考值来表述稳像后视频图像的的稳定程度。

8 分析和总结

分析了现有的稳像图像评价方法,考虑不同场景各种方法的使用范围:对于静态背景的稳像图像评价,均方差(MSE)、峰值信噪比(PSNR)、帧间变换保真度(ITF)、差分图、灰度投影结构相似性(GPASSIM)有较好的客观评价效果;对于具有扫描运动或含有运动物体的图像序列,帧间变换保真度差异(DITF)、标准差方法、随机性检验方法可以反映出动态场景稳像图像的质量。进一步的提出结合人眼视觉特性,根据符合人眼视觉特性的抖动截止频率确定期望误差值,定义稳像前后的抖动量之比定义为抖动衰减系数来衡量稳像图像质量。综上所述,考虑到稳像图像序列输出最终经过接收者和观察者的人眼系统,未来研究方向应进一步分析考虑人眼视觉特性,包括前庭系统机理、视觉暂留机制、人眼视觉注意机制(人眼视觉兴趣)、人眼视觉采样机制等对稳像图像序列的影响,改善现有的评价机制,提高稳像图像质量评价的准确性及适用性。

参考文献:

- [1] Du Dengchong, Jiang Xiaoyu, Yao Jun. Electronic image stabilization algorithm based on phase correlation[J]. Laser & Infrared, 2009, 39(1): 64-66. (in Chinese)
杜登崇, 蒋晓瑜, 姚军. 基于相位相关的电子稳像算法研究[J]. 激光与红外, 2009, 39(1): 64-66.
- [2] Zhang Yongxiang. Study on electronic image stabilization technology for the image sequences of the ship-borne camera system based on the Gyro[D]. Beijing: Chinese Academy of Science, 2005. (in Chinese)
张永祥. 基于陀螺仪的舰载电子稳像技术研究[D]. 北

- 京:中国科学院研究生院,2005.
- [3] Meng Long. Video stabilizing for digital camera[D]. Beijing:Tsinghua University,2004. (in Chinese)
孟龙. 数码相机拍摄视频的稳像处理[D]. 北京:清华大学,2004.
- [4] Morimoto C,Chellappa R. Evaluation of image stabilization algorithms[J]. IEEE,1998,0-7803-4428-6/98.
- [5] Wu Si. Research on techniques for video motion information analysis[D]. Beijing:Chinese Academy of Science, 2005. (in Chinese)
吴思. 视频运动信息分析技术研究[D]. 北京:中国科学院研究生院,2005.
- [6] Wang Zhou,Lu Ligang,Bovik A C. Video quality assessment based on structural distortion measurement[J]. Signal Processing: Image Communication, 2004, 19(2): 121-132.
- [7] Yang Guang,Wang Rui,Cui Yuzhu. Electronic image stabilization and video quality assessment in 2-dimension-array[J]. Infrared and Laser Engineering, 2011, 40(7): 1365-1369. (in Chinese)
杨光,王睿,崔玉柱. 二维面阵的电子稳像与评价方法[J]. 红外与激光工程,2011,40(7):1365-1369.
- [8] Lü Gaojie,Zhang Guohua,Che Hong. A method for evaluating the accuracy of electronic image stabilization system and algorithm[J]. Electronics Optics & Control,2011,18(3):77-79. (in Chinese)
吕高杰,张国华,车宏. 电子稳像算法精度评价方法研究[J]. 电光与控制,2011,18(3):77-79.
- [9] Zhang Yuefei. Research on digital image stabilization for in-car video camera[D]. Chengdu:University of Electronic Science and Technology of China,2011. (in Chinese)
张跃飞. 车载摄像机数字稳像技术研究[D]. 成都:电子科技大学,2011.
- [10] Matti Niskanen,Olli Silven,Marius Tico. Video stabilization performance assessment[J]. IEEE,2006,1-4244-0367-7/06.