

视频帧率上转换检测技术综述

何天琦¹, 蒋兴浩^{1,2}, 孙钺锋^{1,2}

(1. 上海交通大学电子信息与电气工程学院, 上海 200240;
2. 信息内容分析技术国家工程实验室, 上海 200240)

摘要: 视频帧率上转换检测技术是视频取证技术的一种。为了系统阐述视频上转换检测领域现状, 合理导向后续研究, 对相关技术进行了综述。首先对相关研究历史和发展进程进行阐述, 总结上转换概念及技术框架。然后根据检测技术的目的, 对现有算法分类阐述。最后汇总介绍了视频帧率上转换检测领域的主要研究团队及其研究成果。从算法框架、检测结果等方面对比现有检测技术, 提出了两点展望。视频帧率上转换作为视频后处理技术的重要组成部分, 目前仍需进一步研究。

关键词: 视频篡改取证; 帧率上转换; 帧率上转换检测; 周期性分析; 篡改算法分类识别

中图分类号: TP391

文献标识码: A

doi: 10.11959/j.issn.2096-109x.2018085

Review of video frame rate up conversion detection

HE Tianqi¹, JIANG Xinghao^{1,2}, SUN Tanfeng^{1,2}

1. School of Electronic Information and Electrical Engineering, Shanghai Jiaotong University, Shanghai 200240, China
2. National Engineering Lab on Information Content Analysis Techniques, Shanghai 200240, China

Abstract: Video frame rate up-conversion detection technology is a kind of video forensics technology. In order to systematically elaborate the research progress in the field of video up-conversion detection and guide the follow-up research of video up-conversion detection algorithms, the main algorithms of up-conversion detection technology are summarized. First of all, elaborate the relevant research history and development process and summarize the up-conversion concept and technical framework. According to the purpose of the detection technology, classify and explain the existing algorithms. Finally, the main research teams in the field and their research results were summarized. From the aspects of algorithm framework and test results, compare the existing detection techniques and propose two prospects. Video frame rate up-conversion is an important part of video post-processing technology, and still need further research.

Key words: video forensics, frame rate up conversion, frame rate up conversion detection, periodic analysis, tamper algorithm identification and classification

1 引言

近年来, 随着互联网飞速发展, 人们获取信息的途径有了巨大的改变, 从文字、书本扩展到了图像、声音、图形、视频等感觉媒体, 而在网络时代,

它们通过网络传播、分享, 被统称为数字媒体信息。大部分数字媒体信息从产生到传播, 经历获取、处理、存储、管理、输出 5 个部分。随着数字媒体信息的广泛传播, 上述每个环节都存在安全风险。

在诸多数字媒体中, 图像以点阵的形式承载

收稿日期: 2018-08-07; 修回日期: 2018-09-17

通信作者: 何天琦, htq_2011@163.com

了二维空间上的所有信息，在数字媒体早期就成为主要传播媒介；而视频的传输需要更多的带宽资源^[1]，所以传播稍稍迟后，随着技术手段的更新、网络流量资费的下降，具备三维时空信息的视频格式得以广泛传播。正是由于视频领域的快速发展，数字媒体篡改的重心也随之倾斜，当前视频篡改的技术手段层出不穷，其真实性、可靠性时刻面临诸多风险，对社会造成了很多不利影响。视频帧率作为视频的基础组成，保证帧率的真实可靠是保证视频真实可靠的前提，因此，针对视频帧率的篡改检测研究^[2]尤为重要。

一方面，人们追求更高质量的观感体验，视频帧率越高越好，出现了多种上转换技术，篡改者利用该技术提升帧率，消除低帧率带来的停滞感^[3]；另一方面，随着监控视频成为重要证据，不法分子可以通过上转换技术消除截断视频中画面的跳跃感，让检查者忽视篡改痕迹。因此，针对视频帧率上转换技术的检测迫在眉睫。

研究证明，数字图像或者视频进行篡改时，一般作用于非编码域层面，因此对于压缩编码的视频，帧率上转换篡改至少会引起2次重新编码，即双压缩^[4]效应。为了实现应用场景更广、能够区分双压缩痕迹和上转换痕迹的检测方法，后来的研究人员致力于提取更加明显的痕迹来增加检测的顽健性，如帧间相似度的异常、边缘密度的异常等，根据这些人眼难以捕捉的痕迹，手工提取特征，不断完善上转换检测算法。这些方法的局限性在于仅判断视频是否经历过上转换，而并没有对篡改采用的上转换方法进行进一步划分。细致的分类可以针对性抵抗篡改攻击。直到2017年，Ding等^[5]提出检测上转换篡改的同时，对采

用的篡改算法进行分类，实现了进一步的检测。

针对不同的检测手段，国内外团队的研究成果如表1所示。

综合上述分析，国内外研究团队针对视频帧率上转换检测技术领域取得了不错的进展^[6]。近些年来，该领域主要研究以国内团队为主，包括中山大学^[7]、湖南大学^[5,8-10]、南京信息工程大学^[9]、中国矿业大学^[10]，国外研究团队包括韩国国防科学研究所^[11]它们通过改进算法准确性、丰富实验样本、拓展细粒度检测，为上转换检测领域贡献了丰富的研究成果和数据。但由于视频上转换技术不断更新，上述团队未对不同的上转换算法进行研究，所以还需要进一步完善；同时由于上转换领域缺乏系统的样本数据库，难以全面衡量算法的可靠性和有效性。

为了更好地完善上转换检测研究，本文涵盖了对上转换技术的介绍，汇总了典型的上转换技术。并且，通过查阅大量文献对近15年视频帧率上转换检测领域的研究成果进行了汇总总结。

2 上转换篡改技术简述

2.1 上转换概念

在理论研究中，假设未经过上转换的视频原始帧率为15帧/秒，当篡改2倍、4倍至帧率为30帧/秒、60帧/秒时，本质上是以2倍速、4倍速来播放，所以，为使新视频流畅播放，需要通过技术手段增加视频的帧数，即插帧技术，也称作帧率上转换技术。

2.2 上转换篡改算法框架

视频帧率上转换篡改技术难点和关键点在于生成插值帧。目前有2类主要的生成算法：基于

表1 国内外团队研究成果

年份	团队	成果	场景
2013年	米兰理工大学电子与信息学院 Paolo Bestagini ^[6] 团队	首次提出上转换痕迹检测 ^[6]	短片段视频、数帧数字图像集
2013年	中山大学黄继武 ^[7] 团队	首次明确对连续视频进行检测 ^[7]	特定软件获得的上转换视频
2015—2018年	湖南大学杨高波 ^[5,8-10] 团队、南京信息工程大学孙星明 ^[9] 团队、中国矿业大学李雷达 ^[10] 团队	采用更加广泛且典型的帧率篡改手段，获得更为置信的实验样本；提取各类痕迹特征	适用于推断原始帧率、定位篡改点的环境
2017年	韩国国防科学研究所 Dae-Jin Jung 团队联合 KAIST 的 Heung Kyu Lee ^[11] 团队	考虑到上转换操作对连续视频中目标和背景的运动，提出了新的痕迹特征	适用于包含目标运动场景的连续视频

非运动补偿的算法和基于运动补偿的算法, 它们区分明显。下文分别对 2 类算法进行阐述。

2.2.1 基于非运动补偿的算法

基于非运动补偿的算法^[1,12-14]是指采用帧复制或帧平均方式获得插值帧, 如图 1 所示, 插值帧获得公式如下。

$$f'(x, y, t) = \omega f(x, y, t-1) + \varphi f(x, y, t+1) \quad (1)$$

其中, $f'(x, y, t)$ 为插值帧, $f(x, y, t-1)$ 为前向参考帧, $f(x, y, t+1)$ 为后向参考帧, ω , φ 为前后参考帧的权重系数, 当其中某一项为零时, 退化成非运动补偿模式下帧率上转换最简单算法, 即帧复制; 此类方法优点是运算速度快, 可用在对高实时性要求的场景中, 缺点是它们没有考虑场景和目标的运动, 在运动区域多的视频中会产生运动模糊或虚影效应。

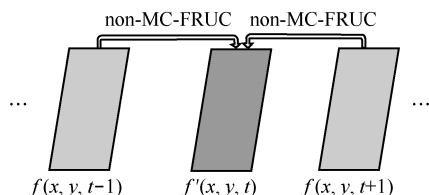


图 1 基于非运动补偿算法图解

2.2.2 基于运动补偿的算法

如图 2 所示, 基于运动补偿的算法考虑帧间运动信息^[15-23], 具体包括 2 个部分: 运动估计 (ME, motion estimation) 以及运动补偿插值 (MCI, motion compensation interpolation)。运动补偿指的是: 假定视频的特定宏块 (互不重叠) 内的每个像素有相同位移, 根据第 $n-1$ 帧的宏块信息在第 n 帧中找到与之最匹配的匹配块, 两者之间的位移即是 2 个宏块之间的运动向量。接下来进行运动补偿插值, 在已估计的运动向量基础上, 根据特定的插值方法生成插值帧。由于估计方法和插值方式的多样性, 形成了不同的帧率上采样算法。基于运动补偿的上转换算法可以实现更加自然的篡改视频; 该算法缺点在于计算复杂度高。上述 2 类篡改算法均改变了原始视频的信息, 在时空域上会产生篡改痕迹, 可作为检测是否经过上转换的依据。

2.3 上转换典型技术

目前, 较为典型的基于运动补偿方式的帧率上转换算法有以下几种, 这些算法由不同的运动估计

和补偿差值方式构成, 下面从两方面进行介绍。

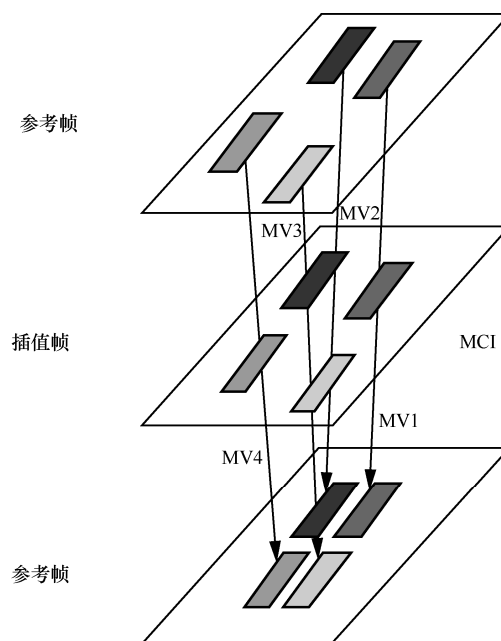


图 2 基于运动补偿算法图解

1) 扩展双边运动估计模式 (EBME, extended bilateral motion estimation)^[24]: 该方法采用扩展运动轨迹范围的双向估计以提高运动估计部分的准确度, 并利用运动向量修正模块对错误的运动向量加以修正、运动平滑模块对运动向量进行平滑处理, 在运动补偿部分, 采用基于加权的双向运动补偿差值模块来获取插值信息, 即重叠块运动补偿模式 (OBMC, overlapped block motion compensation)^[25], 最后获得插值帧。

2) 自适应重叠块运动补偿模式 (AOBMC, adaptive overlapped block motion compensation)^[26]: 该方法在重叠块运动补偿模式基础上, 根据邻近运动向量的相关性自适应调整重叠窗口的权重系数, 即采用自适应重叠块运动补偿技术获得效果更好的插值信息。相比传统的重叠块运动补偿技术, 此方法克服了过渡平滑和差分块的限制。

3) 双重运动估计模式 (DualME, dual motion estimation)^[27]: 在本方法中, 通过调整上一帧与当前帧中单双向块的比例, 获得更高准确度的运动估计信息, 从而得到高质量的插值信息, 在双重运动估计基础上, 结合动态纹理表征的视频显著性模式^[28]进一步提升上转换效果。

4) 多通道混合模式 (MCMP, multi-channel

mixed pattern)^[29]: 在运动估计模块中, 首先计算了图像块的亮度和色度多通道信息, 计算前后向运动向量场, 并利用时空域修正模块对获得的运动向量进一步修正。最后采用双权重重叠块运动补偿模块进行补偿, 获得插值帧信息。该方法的优点在于, 在保障插值帧质量的同时降低计算复杂度。

5) 选择方向运动估计模式 (DSME, direction select motion estimation)^[30]: 本文方法优化运动估计模块来获得更好的插值帧。通过评估前后向运动向量的可靠性, 选择其中最优的方向产生用于插值帧的运动信息。该方法的平均峰值信噪比提高了 5.31 dB, 且将插值帧的平均结构相似度降低至 0.053。

同时, 随着视频市场不断扩大, 视频帧率上转换领域有了很多效果不错的商业软件, 如 MSU、MVTtools2、YUVsoft、Respeedr Motion Perfect 等, 其中, MSU 与 YUVsoft 仅支持整数倍帧率上转换操作。

3 上转换篡改检测算法简述

3.1 篡改检测算法汇总

近年来, 国内外的上转换检测技术主要依据手工提取视频特征。到目前为止, 帧率上转换检测领域文献汇总情况如表 2 所示。

表 2 帧率上转换检测领域文献汇总

上转换篡改算法	篡改检测算法
基于非运动补偿	Bian ^[7] 等
基于运动补偿	Bestagini ^[6] 等、Yao ^[9] 等、Ming Xia ^[8,10] 等、Dae-Jin Jung ^[11] 等、Ding ^[5] 等

可以看出, 针对运动补偿的检测算法数量远超过针对非运动补偿的检测算法数量, 这与现有的上转换算法重点研究提升运动补偿效果相吻合。

3.2 篡改检测算法框架和分类

3.2.1 篡改检测算法框架

上转换检测算法中, 正确描述和提取篡改痕迹是检测算法的难点。算法框架分为 3 个部分。

1) 视频预处理。对获得的篡改视频的帧数、分辨率等进行预处理。

2) 时空域特征提取及预处理。根据上述对帧率上转换算法的阐述可知, 篡改过程主要是插值

帧的构造过程, 其信息来源于原始视频中前后帧信息, 该操作不可避免地会在时空域带来强的相关性, 这种相关性即成为篡改检测的痕迹特征。为去除特征的稀疏性、强化特征数值的差异性, 对特征进行预处理。

3) 基于特征的检测算法。检测算法分为 2 种。第一种是周期性检测^[31]: 插值帧的某些特征与相邻原始帧不同, 当连续对某段视频进行插值时, 该特征会出现周期性的极大值或者极小值, 可以通过判断特征是否具有周期性来检测篡改。同时, 特征出现极值表明当前帧是插值帧, 并通过统计极值周期推断原始帧率。第二种是基于特征相似性的匹配: 由于不同上转换算法生成的篡改视频的特征各自不同, 因此根据特征信息的匹配度进行检测。

3.2.2 篡改检测算法分类

总体来说, 视频帧率上转换检测算法根据实现的目标不同可以分成 2 类。

1) 检测目标不包含对篡改方式的分类识别。如图 3 所示, 典型算法主要包括: 基于残差序列分布的检测算法、基于帧间相似度分布的检测算法、基于边缘密度分布的检测算法、基于平均纹理变化分布的检测算法及基于运动效应分布的检测算法。

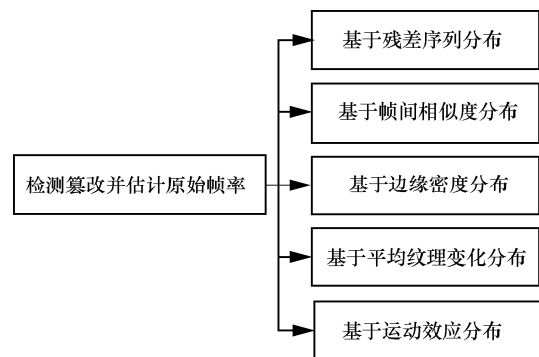


图 3 第一类检测算法

2) 检测目标包含对篡改方式进行分类识别。典型算法包括基于差值信息分布和集成分类器的检测算法。

下文首先阐述典型篡改检测算法, 然后从样本选择和检测准确率方面进行汇总总结。

3.3 检测目标不包含篡改方式识别

3.3.1 基于残差序列的算法

Bestagini 等^[6]首次对帧率上转换检测领域展

开研究, 该团队采用主动比对的方法, 提出了一个基于残差序列的检测模型。针对视频的每一帧, 根据其相邻帧计算该帧的估计帧, 并计算当前帧与估计帧的误差, 该方法发现如果当前帧是插值帧, 那么误差结果近乎为 0; 如果当前帧是原始帧, 那么误差结果较明显。通过对每一帧计算误差信息 $e_{ij}(wt)$, 构造一维均方误差特征 $e(wt)$ 。篡改视频的该特征在频域中呈现周期性, 根据归一化频域中的第一个尖峰信息 Δf 可以计算上转换倍率, 从而计算原始帧率。对于采用帧复制、连续两帧帧平均算法篡改的视频, 该方法检测准确性不高, 容易误判。

3.3.2 基于帧间相似度分布

为了更清晰地描述篡改痕迹, Bian^[7]等将描述相邻两帧的结构相似性衡量指数^[32]特征应用到上转换检测领域, 通过检测视频结构相似性特征对应的相似性序列的周期性, 实验发现在插值帧处出现相似度尖峰。由于运动缓慢的视频中可能存在临近帧相似度值与插值帧相似度相近, 造成对尖峰周期识别困难, 所以该方法分别对原始视频相似度和篡改视频相似度进行分析, 得出经验化参数, 作为区分原始帧与插值帧的阈值。同时研究发现, 与数字图像的插值特性^[33-38]类似, 插值参数 (即上转换倍率) 和经过离散傅里叶变换 (DFT, discrete Fourier transform) 后相似度在频域上的尖峰之间有如下关系。

$$\frac{rate_1}{rate_2} = 1 - f_p \text{ 或 } \frac{rate_1}{rate_2} = f_p \quad (2)$$

其中, $rate_1$ 为原始帧率, $rate_2$ 为篡改后帧率, f_p 为相似度信息在频域中峰值坐标; 由此可以计算插值参数并推断原始帧率。

Bian 团队在 Bestagini 团队的基础上全面对各类视频进行探究, 分辨率从 176×144 到 1 920×1 080 不等, 并且实验了 6 种帧率的上转换视频, 分别为 15 帧/秒、20 帧/秒、24 帧/秒、25 帧/秒、30 帧/秒、60 帧/秒, 共计 15 种上转换组合方式, 其中, 上转换视频作为正样本有 3 000 个, 原始视频作为负样本有 1 000 个。

该方法首次全面采集了各种内容的视频样本, 考虑了整数倍率和非整数倍率上转换, 以及含噪样本, 检测结果的置信度较高。而缺点为仅

采用 ImTOO^[10]单一转换软件产生正样本, 且仅采用帧复制方式, 算法的检测范围单一, 针对其他篡改手段的检测效果没有合理保障。

3.3.3 基于边缘密度的算法

2015 年, Yao^[9]团队研究发现, 帧率上转换篡改行为会破坏连续帧边缘密度在时域上的一致性, 使插值帧的边缘密度明显少于相邻原始帧, 当视频序列多倍率上转换时, 将引起边缘密度周期性的低谷, 根据上述研究, 该团队提出利用边缘密度特征来描述篡改痕迹。为获得每帧的边缘信息, 该方法首先将视频序列解成连续帧, 采用 Kirsch 边缘检测器^[39]提取帧的亮度边缘信息。

$$\rho_k = \frac{N_k}{S} \quad (3)$$

根据式(3)计算边缘密度 (其中, N_k 为边缘像素个数/帧, S 为像素总数/帧), 最终组成边缘密度特征序列。该方法的难点在于边缘算子的选择, 由于上转换视频来源包括帧复制或者帧平均算法, 因此不适合选择具备平滑去噪效应的算子。并且如图 4 所示, 实验证明 sobel 算子、Prewitt 算子等提取的边缘密度特征在整段视频序列中几乎一致, 而采用 Kirsch 算子获得的边缘密度特征具备明显的周期性波动。同时, 传统的 Kirsch 边

缘算子的计算结果为 $g = \begin{bmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}$ 矩阵按 45°

旋转后获得的 8 个方向的卷积核与帧亮度信息卷积的最大幅值, 而该团队创新性地将获得的 8 个方向的平方均值作为 Kirsch 边缘特征, 实验证明该特征具备更优的检测特性。

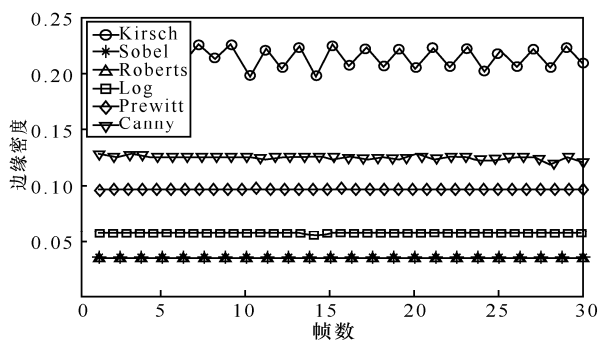


图 4 不同边缘检测算子效果比对

边缘密度特征在插值帧处形成低谷是相对于

邻近原始帧而言的,而不同插值帧的边缘密度特征并不相同,固定的阈值参数不能满足自动化检测的要求。因此,引入 Kaufman^[40]自适应移动均值方法综合考虑当前低谷和历史低谷的特征信息,形成自适应阈值曲线,区分插值帧和原始帧。该方法的优点在于不仅能够检测采用帧复制、帧平均方式的篡改行为,还能够有效检测出基于运动补偿方式的篡改行为,其平均检测准确率可以达到 94.5%。

3.3.4 基于平均纹理变化的算法

不同于上述方法,Xia^[10]团队从视频画面纹理的变化中发现了篡改痕迹。该团队分析发现,在基于运动补偿的帧率上转换篡改过程中,运动向量补偿模块中的权重选择机制本质上是选择相邻帧中最相似的图像块,该机制类似于图像去噪应用中的图像平均机制,均在提高视觉质量的同时消除了高频成分,因此会在插值帧纹理丰富的区域造成模糊,当采用帧平均纹理差值(ATV, average texture variation)^[10]来量化模糊程度时,发现插值帧的平均纹理差值普遍小于原始帧。

具体来说,首先对待检测的压缩编码视频解压,获得帧序列,然后针对帧序列提取亮度信息,并计算每帧的平均纹理差值,从而描绘全部视频的对应差值曲线。如图 5 所示,以 Foreman 视频为例,实验分析发现,原始视频的平均纹理差值曲线趋势平滑且不具备周期性,其变化区间仅为 $[0.022\ 8, 0.026\ 5]$,而上转换篡改后视频的曲线具有明显周期性波动,其中,采用基于运动补偿方式获得的上转换视频的差值变化区间最大,为 $[0.008\ 4, 0.026\ 5]$ 。

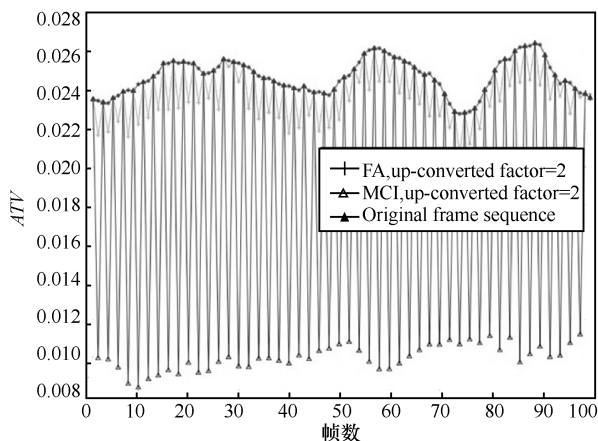


图 5 ATV 曲线效果 (以 Foreman 序列为例)

为了估计原始帧率,通过对帧平均纹理差值曲线进行离散傅里叶变换得到频谱曲线,再采用高通滤波器去除低频分量,获得的高频尖峰坐标即为原始帧率。同样,该算法通过设计合理的自适应阈值来区分原始帧段和插值帧,从而定位插值帧。

该方法重点考虑了基于运动补偿的帧率上转换篡改行为,当上转换倍率小于 4 时,检测准确率大于 93.5%,但当上转换倍率超过 4 时,检测准确率会产生不稳定波动。其优点在于可以从原始帧序列中自动区分上转换视频序列,并且有较低的时间复杂度。其缺点在于,当待检测视频序列包含快速场景切换时,平均纹理差值的异常会导致插值帧误判为原始帧,降低了检测准确率。

3.3.5 基于运动效应的算法

在国内对帧率上转换领域持续取得研究进展的同时,2017 年韩国国防科学研究所的 Jung^[11]团队提出了一种新的检测算法:基于运动轨迹周期性的检测,他们注意到上转换插值信息周期性体现在像素块位置的移动上,而基于运动补偿的上转换篡改视频的运动向量(如图 6 所示)满足特征要求。

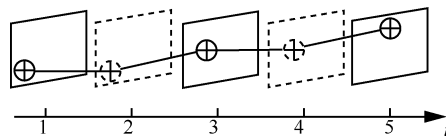


图 6 运动补偿中的运动向量

该方法分成 2 个部分,首先针对第 t 帧计算第 $t-1$ 帧和第 t 帧之间运动向量 $mv_1(i, j, t)$ 与第 t 帧和第 $t+1$ 帧之间的运动向量 $mv_2(i, j, t)$,然后利用运动修剪模块剔除异常运动向量对,最后利用运动向量的周期性实现上转换检测。实验发现,当帧率 24 帧/秒上转换至 25 帧/秒时,该方法检测准确率急剧降低。另外,该方法额外考虑了帧率下转换的情况,并对应进行了分析。由于在现实生活中,帧率下转换的应用场景很少,对其的研究并没有过多深入。

3.4 检测目标包含篡改方式识别

迄今为止,针对视频是否经过帧率上转换篡改问题,最高可实现接近 100% 的准确率。另外,

在帧率上转换篡改领域，针对基于运动补偿方式的研究越来越多，为了更细粒度的检测篡改，需要对篡改方式进行分类识别。

2017 年，Ding^[5]等提出了上转换视频盲检测系统，为了减少上转换视频在时域和空域上的冗余，算法首先计算当前帧与前后连续两帧像素域差值特征，进一步计算垂直、水平、主对角线、次对角线 8 个方向上的截断差值特征，从而统计 8 个方向上的一阶马尔可夫转移概率矩阵，对概率转移矩阵进行降维处理，得到 $(2T+1)^2$ 维特征向量，其中， T 为截断差值特征的截断阈值。对于经过压缩编码的视频，对亮度信息进行离散余弦变换获得离散余弦变换（DCT, discrete cosine transform）系数作为差值信息进行处理。差值信息来源于运动和纹理丰富的插值帧，而在原始帧或静止插值帧中，差值信息过于稀疏，零值占比接近 99.9%，在预处理模块中，去除差值信息零值占比大于 99.5% 的差值信息；同时由于快速场景变换会引起运动区域的显著变化，会对差值信息造成误筛选，因此预处理模块还包括场景变换检测部分。

基于运动补偿的上转换篡改算法由 2 部分组成：运动估计和运动补偿，不同的运动补偿算法主要差异在于上述 2 个部分的不同，其中运动估计包括 4 类：单向运动估计^[25,41]、双重运动估计^[24,26-27]、单双向运动估计^[29-30,42]、多假设运动估计^[43]；运动补偿主要包括以下 5 类：直接运动补偿^[44]、重叠块运动补偿^[25]、自适应重叠块运动补偿^[26]、多通道混合模式^[29]和选择方向运动估

计模式^[30]，其中，重叠块运动补偿能有效减轻块效应，应用广泛。

由于上述不同运动估计和运动补偿方式的不同组合，理论上可以对不同上转换篡改算法进行分类。该方法中，团队创新性地对 11 种不同的基于运动补偿的上转换篡改算法进行分类，采用基于 12 类集成分类器的误差纠错码^[32]将多维分类转化成 66 个二进制分类问题，一阶马尔可夫特征对应二进制向量，根据输出向量与码字之间的最小汉明距离实现分类。

4 检测算法对比与总结

4.1 样本汇总对比

算法样本汇总如表 3 所示。

综上所述，在视频帧率上转换检测领域，样本分辨率范围较一致，所选取的原始帧率和上转换帧率也较一致，但至今没有标准样本库，仍存在多种参数配置不一致的情形，不仅影响了各类算法的检测准确率，更阻碍了上转换检测领域的后续研究。因此，建立一个标准样本库显得尤为迫切。

4.2 算法准确率对比

由于缺乏统一标准的样本库，各类检测算法的检测准确率无法公正地对比衡量，本文分别对各类方法准确率进行汇总，具体如表 4 所示。

根据表 4，可得如下结论。

1) 针对未压缩视频：基于帧间相似度分布的检测准确率最高，为 99%。

2) 针对 H.264/AVC 压缩视频：基于帧间相似

表 3 算法样本汇总

算法	样本
基于残差序列分布	3 段通用影像传输格式（CIF, common intermediate format）视频：Foreman、Hall、Mobile
基于帧间相似度分布	100 段未压缩视频及 400 段 H.264/AVC 压缩编码格式视频；分辨率从 176×144 至 1920×1080
基于边缘密度分布	5 段未压缩视频，包含 4 类：单一目标强运动、单一目标弱运动、多目标强运动、多目标弱运动。分辨率从 176×144 至 1920×1080
基于纹理变化分布	30 段未压缩 CIF 视频，30 段 H.264/AVC 格式视频
基于运动效应分布	50 段 CIF 视频，分辨率从 176×144 至 1920×1080 不等，每段序列不超过 300 帧；包含 5 种压缩参数（100, 90, 80, 70, 60）的 H.264/AVC 格式视频
基于差值信息分布	15 段 CIF 视频和 10 段高清视频；对应采用 5 种不同图像组（GOP, group of picture）；量化参数配置获得 MPEG-2 和 H.264/AVC 格式视频

表 4 平均检测准确率列表

算法	平均准确率
基于残差序列分布	1) 上转换成 25 帧/秒, 45 帧/秒, 60 帧/秒, 90 帧/秒: 视频达到 50 帧, 准确率突变到 100%; 2) 上转换成 120 帧/秒: 视频达到 50 帧, 准确率激增至 70%, 随后在达到 600 帧时, 上升到 100%; 3) 上转换成 36 帧/秒: 视频帧数在 200 帧内, 准确率逐渐上升, 直至 100%
基于帧间相似度分布	1) 未压缩视频: 大于 99%; 2) H.264/AVC 压缩视频: 大于 97%
基于边缘密度分布	1) 未压缩视频: 94.05%; 2) H.264/AVC 压缩视频: 93.2%
基于平均纹理变化分布	1) OBMC 篡改方式: 最高为 95.2%; 2) 帧平均篡改方式: 最高为 93.5%; 3) 采用软件 MSU、MVTools: 准确率分别为 94.5%、94.4%
基于运动效应分布	1) 帧复制方式: 97%; 2) 帧平均方式: 78%; 3) 运动补偿方式: 90%
基于差值信息分布	1) 未压缩视频: 97%; 2) H.264/AVC 压缩视频: 95.16%; 3) MPEG-2 压缩视频: 83.24%; 4) 分类识别篡改方式: 81.27%

度分布的检测准确率最高, 为 97%。

3) 针对 MPEG-2 压缩视频: 基于差值信息分布的检测准确率最高, 为 83.24%。

4) 针对帧复制篡改: 基于运动效应分布的检测准确率最高, 为 97%。

5) 针对帧平均篡改: 基于运动效应分布的检测准确率最高, 为 93.2%。

6) 针对运动补偿篡改: 基于运动效应的检测准确率最高, 为 95.2%。

7) 针对分类识别篡改方式: 基于差值信息分布的分类准确率最高, 为 81.27%。

4.3 算法优缺点对比

检测目标不包含篡改方式识别和包含篡改方式识别的优缺点分析分别如表 5 和表 6 所示。

总体来说, 上述各类篡改检测算法有各自的适用场景和优缺点; 在上转换检测领域, 未来可以针对识别篡改方式进行深入研究, 扩大算法的适用场景、提升分类识别的准确率。

表 5 优缺点分析-检测目标不包含篡改方式识别

算法	优点	缺点
基于残差序列的算法	适用于小段视频检测	1) 对帧复制、帧平均的篡改视频检测准确率不高, 易误判; 2) 算法检测准确率与视频序列长度有较强相关性, 视频帧数增加时, 检测复杂度线性增加
基于帧间相似度的算法	样本全面, 置信度高	仅用 ImTOO ^[7] 单一转换软件, 且该软件仅采用帧复制方式
基于边缘密度的算法	可检测帧复制、帧平均、运动补偿方式的篡改视频	边缘检测算子选择困难
基于平均纹理变化的算法	1) 从原始帧序列中自动区分上转换视频序列; 2) 低时间复杂度	当待检测视频序列包含快速场景切换时, 帧平均纹理差值的异常会导致插值帧误判为原始帧, 降低了检测准确率
基于运动效应的算法	考虑了帧率下转换的情况, 并对应进行了分析	当帧率 24 帧/秒上转换至 25 帧/秒时, 该方法检测准确率急剧降低

表 6 优缺点分析-检测目标包含篡改方式识别

算法	优点	缺点
基于差值信息的算法	创新性地对 11 种不同的基于运动补偿的上转换篡改算法进行分类	检测算法复杂度较高

5 结束语

5.1 总结

随着视频成为主要的信息传播媒介, 帧率上转换领域逐渐兴起。在引言部分, 介绍了帧率上转换技术的研究发展过程、国内外主要研究团队及其研究情况。在视频帧率上转换概念与分类部分, 介绍了上转换算法的概念, 详细阐述了 3 类上转换算法框架与思路, 汇总了目前典型的上转换算法, 相较基于非运动补偿的上转换算法, 基于运动补偿的上转换算法研究体系更加完善, 所以针对后者的检测算法成为上转换检测领域的研究重点。在第 4 节, 根据是否包括对上转换算法的分类检测分别对近 15 年该领域的算法思路展开框架分析、优缺点分析、样本分析等。

本文创新性有两点。一方面, 本文聚焦视频帧率上转换检测领域, 系统综合全面地分析了该问题出现的背景及现状。从算法逻辑、检测范围、检测准确率等方面对现有检测算法进行了深度比较。另一方面, 本文作为帧率上转换检测领域的综述性文章, 总结概括了前人的研究成果, 为后来研究人员知悉该领域现状提供了巨大的便捷, 并且在参考文献部分, 囊括了上转换领域的绝大部分文章, 为后来者节省大量的资料搜集时间。

5.2 展望

虽然帧率上转换检测已经取得了相当高的准确率, 但由于应用场景的变化多样, 需要的检测指标也日趋复杂, 上转换检测领域仍有很多需要分析解决的问题。

1) 到目前为止, 全部的检测算法均依赖手工特征, 而手工特征更多地依赖专家经验, 这使上转换痕迹的提取过高地受专家知识限制, 忽略其他的潜在特征。

如今, 处理器等硬件设备不断发展、日趋革新, 这使智能的学习上转换痕迹成为可能。实验证明神经网络具备拟合任意复杂函数的特点, 可以构造相当复杂的非线性映射, 并且深度神经网络能产生足够高的假设空间维度, 具有强大的信息表征能力, 更重要的是, 其提取

的特征比人工提取的特征具备更优的泛化性能, 可以学习深层痕迹信息。将神经网络与上转换检测相结合, 不但能够学习到人类难以发现的信息, 还能够保证检测的顽健性, 因此, 如何将神经网络应用到上转换篡改痕迹的提取上具有重要意义。

2) 运动补偿方式各式各样, 使基于运动补偿方式的上转换算法也各具特点。对于需要获得原始视频的人而言, 知道篡改者采用的篡改方法会大幅度降低恢复原始视频的难度, 因此, 分类识别篡改方式显得尤为重要。实验证明, 深度学习在图像识别分类领域有着卓越的表现, 对于上转换篡改方式的分类问题, 可以构造篡改痕迹特征图作为图像样本, 针对不同上转换算法, 构造的特征图各不相同, 通过对特征信息的深度学习实现更加智能准确的分类结果。由于深度学习是通过对训练样本的学习得到分类模型, 这意味着需要足够大的样本量。对于上转换算法分类问题, 汇集各类篡改方式的视频, 建立专用视频样本库显得尤为重要, 另外, 通过视频解成帧等方法也是扩大样本量的有效途径。

参考文献:

- [1] GIRONI A, FONTANI M, BIANCHI T. A video forensic technique for detecting frame deletion and insertion[J]. IEEE International Conference on Acoustics, Speech and Signal Processing. 2014:6226-6230.
- [2] LUO W, QU Z, PAN F, et al. A survey of passive technology for digital image forensics[J]. Frontiers of Computer Science in China. 2008, 1(2):166-179.
- [3] 黄昆仑, 白蔚. 视频帧率上转换技术综述[J]. 数字通信世界, 2011(5):82-84.
HUANG K L, BAI W. Overview of video frame rate up-conversion technology[J]. Digital communication world, 2011(5):82-84.
- [4] LUO W, WU M, HUANG J. MPEG recompression detection based on block artifacts[J]. Proceedings of SPIE-The International Society for Optical. 2008, 6819: 68190X-68190X-12.
- [5] DING X, YANG G, Li R. Identification of motion-compensated frame rate up-conversion based on residual signal[J]. IEEE Transactions on Circuits & Systems for Video Technology, 2017, PP(99):1-1.
- [6] BESTAGINI P, BATTAGLIA S, MILANI S, et al. Detection of temporal interpolation in video sequences[C]// IEEE International Conference on Acoustics, Speech and Signal Processing.

- 2013:3033-3037.
- [7] BIAN S, LUO W, HUANG J. Detecting video frame-rate up-conversion based on periodic properties of inter-frame similarity[J]. *Multimedia Tools & Applications*, 2014, 72(1):437-451.
- [8] 夏明, 杨高波, 黄勤珍. 使用 PSNR 周期特性检测视频帧率上转换伪造[J]. *计算机应用研究*, 2016, 33(8): 2409-2412.
- XIA M, YANG G B, HUANG Q Z. Detecting frame-rate up-conversion forgery in videos using periodicity of PSNR change[J]. *Application Research of Computers*, 2016, 33(8): 2409-2412.
- [9] YAO Y, YANG G, SUN X. Detecting video frame-rate up-conversion based on periodic properties of edge-intensity[J]. *Journal of Information Security & Applications*, 2016, 26(C):39-50.
- [10] XIA M, YANG G, LI L. Detecting video frame rate up-conversion based on frame-level analysis of average texture variation[J]. *Multimedia Tools & Applications*, 2017, 76(6):8399-8421.
- [11] JUNG D J, LEE H K. Frame-rate conversion detection based on periodicity of motion artifact[J]. *Multimedia Tools & Applications*, 2017(1):1-22.
- [12] CHAO J, JIANG X, SUN T. A novel video inter-frame forgery model detection scheme based on optical flow consistency[M]. *Digital Forensics and Watermarking*. Springer Berlin Heidelberg, 2013:267-281.
- [13] WU Y, JIANG X, SUN T. Exposing video inter-frame forgery based on velocity field consistency[C]//*IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2014: 2674-2678.
- [14] LIU Y, HUANG T. Exposing video inter-frame forgery by Zernike opponent chromaticity moments and coarseness analysis[J]. *Multimedia Systems*, 2017, 23(2):223-238.
- [15] ORCHARD M T, SULLIVAN G J. Overlapped block motion compensation: an estimation-theoretic approach[J]. *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, 1994, 3(5):693-9.
- [16] KANG S J, CHO K R, KIM Y H. Motion compensated frame rate up-conversion using extended bilateral motion estimation[J]. *IEEE Transactions on Consumer Electronics*, 2008, 53(4):1759-1767.
- [17] CHOI B D, HAN J W, KIM C S. Motion-compensated frame interpolation using bilateral motion estimation and adaptive overlapped block motion compensation[J]. *IEEE Transactions on Circuits & Systems for Video Technology*, 2007, 17(4):407-416.
- [18] KANG S J, YOO S, KIM Y H. Dual motion estimation for frame rate up-conversion[J]. *IEEE Transactions on Circuits & Systems for Video Technology*, 2011, 20(12):1909-1914.
- [19] LI R, GAN Z, CUI Z. Multi-channel mixed-pattern based frame rate up-conversion using spatio-temporal motion vector refinement and dual-weighted overlapped block motion compensation[J]. *Journal of Display Technology*, 2014, 10(12):1010-1023.
- [20] YOO D G, KANG S J, KIM Y H. Direction-select motion estimation for motion-compensated frame rate up-conversion[J]. *IEEE/OSA Journal of Display Technology*, 2013, 9(10):840-850.
- [21] LIN H Y, TSAI T H. High visual quality particle based frame rate up conversion with acceleration assisted motion trajectory calibration[J]. *Journal of Display Technology*, 2012, 8(6):341-351.
- [22] LIU H, XIONG R, ZHAO D. Multiple hypotheses bayesian frame rate up-conversion by adaptive fusion of motion-compensated interpolations[J]. *IEEE Transactions on Circuits & Systems for Video Technology*, 2012, 22(8):1188-1198.
- [23] 陈伟. 基于运动估计和运动补偿的帧率上转换算法研究[D]. 哈尔滨工业大学, 2016.
- CHEN W. Research on frame rate up-conversion algorithm based on motion estimation and motion compensation[D]. Harbin Institute of Technology, 2016.
- [24] KANG S J, CHO K R, KIM Y H. Motion compensated frame rate up-conversion using extended bilateral motion estimation[J]. *IEEE Transactions on Consumer Electronics*, 2008, 53(4):1759-1767.
- [25] ORCHARD M T, SULLIVAN G J. Overlapped block motion compensation: an estimation-theoretic approach[J]. *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, 1994, 3(5):693-9.
- [26] CHOI B D, HAN J W, Kim C S, et al. Motion-compensated frame interpolation using bilateral motion estimation and adaptive overlapped block motion compensation[J]. *IEEE Transactions on Circuits & Systems for Video Technology*, 2007, 17(4):407-416.
- [27] KANG S J, YOO S, KIM Y H. Dual motion estimation for frame rate up-conversion[J]. *IEEE Transactions on Circuits & Systems for Video Technology*, 2011, 20(12):1909-1914.
- [28] 胡圆圆. 基于视觉显著性的视频关键帧提取与帧速率上转换[D]. 南京邮电大学, 2016.
- HU Y Y. Video key frame extraction and frame rate up conversion based on vision saliency[D]. Nanjing University of Posts and Telecommunications, 2016.
- [29] LI R, GAN Z, CUI Z. Multi-channel mixed-pattern based frame rate up-conversion using spatio-temporal motion vector refinement and dual-weighted overlapped block motion compensation[J]. *Journal of Display Technology*, 2014, 10(12):1010-1023.
- [30] YOO D G, KANG S J, KIM Y H. Direction-select motion estimation for motion-compensated frame rate up-conversion[J]. *IEEE/OSA Journal of Display Technology*, 2013, 9(10):840-850.
- [31] MAHDIAN B, SAIC S. Blind authentication using periodic properties of interpolation[J]. *IEEE Transactions on Information Forensics & Security*, 2008, 3(3):529-538.
- [32] DIETTERICH T G, BAKIRI G. Solving multiclass learning problems via error correcting output codes[J]. *Journal of Artificial Intelligence Research*, 1995, 2(1):263-286.
- [33] POPESCU A C, FARID H. Exposing digital forgeries by detecting traces of resampling[J]. *IEEE Transactions on Signal Processing*, 2005, 53(2):758-767.
- [34] VAZQUEZ-PADIN D, COMESANA P. ML estimation of the resampling factor[C]// *IEEE International Workshop on Information Forensics and Security*. IEEE, 2012:205-210.

- [35] KIRCHNER M. Fast and reliable resampling detection by spectral analysis of fixed linear predictor residue[C]// The Workshop on Multimedia & Security. ACM, 2008:11-20.
- [36] HUANG H C, FANG W C. Metadata-based image watermarking for copyright protection[J]. Simulation Modelling Practice & Theory, 2010, 18(4):436-445.
- [37] KOBAYASHI M, OKABE T, SATO Y. Detecting forgery from static-scene video based on inconsistency in noise level functions[J]. IEEE Transactions on Information Forensics & Security, 2010, 5(4):883-892.
- [38] LIAO D, YANG R, LIU H. Double H.264/AVC compression detection using quantized nonzero AC coefficients[J]. Proceedings of SPIE - The International Society for Optical Engineering, 2011, 7880(2):7880Q-7880Q-10.
- [39] TEKALP A M. Digital video processing[M]. Englewood Cliffs, NJ: Prentice Hall; 1995.
- [40] KAUFMAN P J. Smarter Trading: Improving Performance in Changing Markets[M]. McGraw-Hill; 1995: p. 138-51.
- [41] KIM U S, SUNWOO M H. New frame rate up-conversion algorithms with low computational complexity[J]. IEEE Transactions on Circuits & Systems for Video Technology, 2014, 24(3):384-393.
- [42] TSAI T H, Lin H Y. High visual quality particle based frame rate up conversion with acceleration assisted motion trajectory calibration[J]. Journal of Display Technology, 2012, 8(6):341-351.
- [43] LIU H, XIONG R, Zhao D. Multiple hypotheses Bayesian frame rate up-conversion by adaptive fusion of motion-compensated interpolations[J]. IEEE Transactions on Circuits & Systems for Video Technology, 2012, 22(8):1188-1198.
- [44] DAR Y, BRUCKSTEIN A M. Motion-compensated coding and frame rate up-conversion: models and analysis[J]. IEEE Transactions on Image Processing, 2015, 24(7): 2051-2066.

[作者简介]



何天琦（1995-），女，安徽池州人，上海交通大学硕士生，主要研究方向为视频帧率上转换篡改检测。



蒋兴浩（1976-），男，上海人，上海交通大学教授、博士生导师，主要研究方向为互联网内容安全、视频内容智能理解与网络传播控制、多媒体篡改检测分析取证技术、网络协议内容分析。

孙锁锋（1975-），男，上海人，上海交通大学副教授，主要研究方向为视频篡改分析视频水印技术。