摘　要

视频去隔行技术致力于改善人类视觉系统或机器视觉系统的主观体验，其主要处理对象是电视系统中广泛存在的隔行扫描格式。视频去隔行研究涉及到对运动图像的分析与理解，需要综合运用数字视频信号处理、模式识别、计算机视觉等领域的知识；同时，视频去隔行技术的不断成熟也能对相应的学科建设起到很好的推动作用。此外，以视频去隔行技术为代表的视频处理技术的发展对视频压缩、视频监控等相关领域的研究也有着很好的支撑作用。

在过去的数十年中，国内外对视频去隔行技术的研究取得了一系列重要的进展，研究成果遍及运动估计、运动分析、时空域滤波（插值）、质量评价等几乎所有的重要方向。遗憾的是，现有的研究仍存在三方面的问题：一是偏重于对显示效果的改善，而几乎没有考虑去隔行技术对视频压缩等重要领域的意义；二是专注于通用情况下的研究，对具体应用场景的先验知识利用得不够；三是高性能算法多走计算密集型路线，难以得到产业界的认同。

本文针对上述问题，从改善显示效果、提高压缩效率、利用先验知识、优化算法实现等方面开展了一系列的深入研究，主要贡献如下：

1、视频去隔行技术综述

最近的一篇视频去隔行综述文章发表于 1998 年，而且仅从时空域滤波算法的角度进行了归纳总结。此后的文献中虽然或多或少有一些综述内容，但也都没有能够摆脱这一局限性。本文首次分技术模块对视频去隔行的国内外研究现状进行了综述：在运动估计部分，综述了块匹配法、相位相关法、梯度光流法、贝叶斯法等方法；在运动分析部分，对运动检测和运动类型分析方法进行了讨论；在时空域滤波部分，分线性算法、非线性无运动补偿算法和基于运动补偿的算法三大部分进行了介绍。

2、多运动向量自适应选择算法

针对隔行视频序列的运动估计算法一般只在具有相同奇偶性的场之间进行运动估计。这样可以有效地预防误差传递，但在以垂直运动为主的区域会造成明显的性能下降。本文既考虑了相同奇偶性的场图像之间的运动向量，也考虑了相反奇偶性的场图像之间的运动向量，并基于高斯噪声模型提出了一种在多个运动向量之间进行自适应选择的可靠性判别准则，然后从数学上严格证明了该准则的最大概率意义。此外，本文还进一步提出了一种运动向量映射方法，在充分利用已有结果和有效预防误差传递之间取得了较好的折中。为了验证算法的性能，本文还提出了一种基于边缘的亚像素时空中值滤波算法。实验结果表明，本文提出的场运动估计算法相比于传统算法有将近 1 dB 的平均 PSNR 增益。

3、面向实时应用的自适应运动分析算法

大量的实验表明，没有哪种时空域滤波算法是万能的。一般说来，基于运动补偿的算法的性能好于无运动补偿的算法，但该类算法也存在着误差传递风险和计算量大的问题。为了与具体插值算法的特性相适应，结合实时系统要求快速判定运动类型的需求，本文提出了一种面向实时应用的自适应运动分析算法——DAMA 算法。为了实现解码端的实时去隔行，本文进一步提出用 MBP 类型信息和帧序重排特性来区别处理 I 图像、B 图像和 P 图像。客观实验表明，DAMA 算法的引入使得时空域滤波算法的平均 PSNR 性能提升了 1 dB 以上；主观实验表明，DAMA 算法能够在多种时空域滤波算法之间“取长补短”，显著提升去隔行的主观视觉效果；实时解码实验表明，DAMA 算法有助于增强时空域滤波算法对不准确的运动向量的鲁棒性。

4、细粒度混合运动补偿插值算法

运动补偿插值是视频去隔行技术的最后一个步骤，也是最关键的一个步骤。最负盛名的运动补偿插值算法之一是递归算法，这类算法的总体性能在所有算法中是最好的，但容易出现误差传递。本文首先提出了一种块尺寸自适应可变的运动分析框架——BAMA 框架，然后基于该框架提出了一种以递归的时空中值滤波算法为基础的细粒度混合运动补偿插值算法，从而既充分参考了前向的插值结果，又很好地抑制了误差传递。实验结果表明，细粒度插值可以给性能一般的中值滤波算法带来高达 1.83 dB 的平均 PSNR 增益，与此同时计算复杂度仅增加了约 30%。

5、公共研发平台上的性能验证

在编码端，去隔行技术可以大大增加图像之间的时空相关性，从而提高预测的精度。本文首先提出了一种编码性能验证方案，然后通过 AVS 和 H.264/AVC 等公共研发平台上的编码实验指出，去隔行所增加的数据量可以在很大程度上被新增的时空相关性抵消，因而在低码率下借助去隔行技术往往可以获得更好的压缩性能，甚至在高码率下有时亦如此。进一步的实验和分析表明，本文提出的视频去隔行关键技术比常见的算法更适用于视频压缩、视频转码等实时环境。去隔行技术的发展实际上也为视频压缩和视频转码提供了一种新的研究思路，具有重要的实际意义。

关键词：视频去隔行；运动向量可靠性；DAMA 运动分析；BAMA 运动分析框架；细粒度插值；视频压缩

Study on Key Techniques for Video Deinterlacing

Qian Huang (Computer Application Technology)
Supervised by Professor Wen Gao

Video deinterlacing, which converts the widely used interlaced scanning format in television systems to the progressive scanning format, devotes to improving the subjective experience of the human visual system or machine visual system. Research on video deinterlacing involves analyzing and understanding moving pictures, as well as synthesizing methods in digital video signal processing, pattern recognition, and computer vision. At the same time, the development of video deinterlacing techniques has also an impulse on corresponding scientific domains. In addition, the deinterlacing techniques also support the research on video compression and video surveillance.

In the past decades, great progresses have been made in the research areas of video deinterlacing, which include the techniques of motion estimation, motion analysis, spatio-temporal filtering (interpolation), and video quality assessment. However, some issues of deinterlacing techniques need to be further studied. Firstly, most researchers pay more attentions on the quality improvement for display applications, while paying less effort to make deinterlacing techniques for benefiting other applications such as video compression and transcoding. Secondly, most studies on deinterlacing research focus on general cases, without making full use of pre-knowledge existing in the video signals. Thirdly, high performance algorithms are typically computation-intensive, therefore, these algorithms are difficult to be implemented by the industry for practical applications.

In this dissertation, we address the above problems by focusing on enhancing the display quality, improving the compression performance, utilizing pre-knowledge, and optimizing algorithm implementations. The main contributions of this thesis are composed of:

1、Techical Survey on Video Deinterlacing

The latest survey on video deinterlacing was published in 1998, which in fact surveyed only the spatio-temporal filtering algorithms. After that, many publications also reviewed techniques of video deinterlacing. However, almost all of them failed to get rid of the same restriction. In this dissertation, we give a technical survey of video deinterlacing for the first time: in the motion estimation part, block-matching methods, phase-correlation methods, differential optical flow methods, and Bayesian methods are reviewed; in the motion analysis part, motion detection methods and motion type analysis methods are discussed; in the spatio-temporal filtering part, current algorithms are classified into linear algorithms, non-linear non-motion-compensated algorithms, and motion-compensated algorithms.

2. Adaptive Selection among Multiple Motion Vectors

Motion estimation algorithms for interlaced videos are usually conducted among same-parity fields only. In this way, the error propagation can be effectively prevented. However, the performance degrades in regions where vertical motions are dominant. In this dissertation, motion vectors among both same-parity fields and opposite-parity fields have been considered; and a new method has been proposed for determing the reliability of motion vectors. A theoretical proof is provided to show that the selected motion vector has the maximum probability. Moreover, we further introduce a motion vector mapping method that can achieve a good compromise between fully utilizing previous results and effectively preventing error propagation. To evaluate the proposed adaptive selection scheme, an edge-based sub-pixel spatio-temporal median filtering algorithm has also been proposed. Experimental results have shown that an average PSNR gain of about 1 dB can be achieved over a typical algorithm.

3. Adaptive Motion Analysis for Real-time Applications

Extensive experiments show that there is not any omnipotent deinterlacing algorithm. In general, motion-compensated (MC) algorithms outperform non-MC ones. However, the MC algorithms suffer from latent error propagations and heavy computation load. To adapt to the characteristic of specific interpolation algorithms, this dissertation proposes a directional adaptive motion analysis (DAMA) algorithm for real-time applications. To realize real-time deinterlacing at the decoder side, we further propose to separately process I pictures, B pictures, and P pictures utilizing the MBP information and the characteristic of frame reordering. Objective experiments show that, DAMA enhances the performance of spatio-temporal filtering for more than 1 dB. Subjective experiments show that, DAMA can make the best out of multiple spatio-temporal filtering algorithms so as to greatly improve the subjective quality. The real-time deinterlacing at decoder side shows that, DAMA helps to improve the robustness when motion vectors are unreliable.

4. Fine-Grained Hybrid Motion-Compensated Interpolation

Motion-compensated interpolation (MCI) is the last and the most crucial step of video deinterlacing. One of the most famous kinds of MCI methods is the recursive one, which generally has the best performance but suffers from latent error propagations. In this dissertation, we first present a block-size adaptable motion analysis (BAMA) framework, and then propose a fine-grained hybrid motion-compensated interpolation method based on several motion-compensated recursive median filters. In this way, we can make full use of previously deinterlaced results while restraining error propagation. Experimental results show that we can gain as much as 1.83 dB in terms of PSNR for ordinary median filtering methods, with only

marginal increase in computational complexity.

5. Performance Validation on Public Research & Development Platforms

Video deinterlacing techniques can greatly increase the spatio-temporal correlations so that the prediction precision in the encoder can be improved. In this dissertation, we propose a scheme for video coding performance validation. It should be noted that although deinterlacing doubles the data to be encoded, at the same time it also greatly increases the spatio-temporal correlations, which can compensate the encoding performance to a great extent according to extensive experiments on AVS and H.264/AVC, especially at low bitrates. Sometimes this is also true for high bitrate cases. Further experimental results indicate that the proposed key techniques for video deinterlacing are friendlier to video compression and video transcoding than conventional algorithms. In fact, the development of deinterlacing techniques also provides a new angle of view for research on video compression and video transcoding.

Keywords: Video Deinterlacing, Motion Vector Reliability, Directional Adaptive Motion Analysis, Block-size adaptive motion analysis, Fine-grained interpolation, Video Compression