

面向立体视频的自适应数字水印算法

朱仲杰¹⁾ 王玉儿¹⁾ 蒋刚毅²⁾ 梁丰^{b)}

¹⁾(浙江万里学院宁波市 DSP 重点实验室, 宁波 315100) ²⁾(宁波大学电路与系统研究所, 宁波 315211)

摘要 鉴于立体视频是未来视频技术的一个重要发展方向,对其产品的知识产权进行保护需要提前解决,以避免盗版和侵权现象。为此,在前期数字水印和立体视频研究工作的基础上,通过详细分析立体视频的特点后,给出了一种面向立体视频的自适应数字水印新算法。新算法是将水印信息添加到增强层码流的视差矢量数据上,这种水印信息的嵌入不仅对基本码流没有影响,而且不会影响到基本码流重建后的视频图像质量。该水印算法采用自适应策略,使得嵌入水印后的图像质量损失达到最小。其水印的嵌入和提取操作均基于压缩域进行,无需视频码流的完全解码,而且水印的提取使用盲检法,无需视频原图像。整个算法简单、快速、高效。通过不同测试序列的仿真实验结果显示,水印的嵌入对原图像影响很小,使用自适应算法进行水印嵌入后图像的峰值信噪比下降值均小于非自适应算法。

关键词 立体视频水印 视差矢量 自适应算法 压缩域 盲检

中图分类号: TP309 **文献标识码:** A **文章编号:** 1006-8961(2007)01-0068-05

Adaptive Digital Watermarking Algorithm for Stereo Video

ZHU Zhong-jie¹⁾, WANG Yu-er¹⁾, JIANG Gang-yi²⁾, LIANG Feng¹⁾

¹⁾(Ningbo Key Laboratory of DSP, Zhejiang Wanli University, Ningbo 315100)

²⁾(Institute of Circuits and Systems, Ningbo University, Ningbo 315211)

Abstract Stereo video is an important future trend in video technologies and intellectual property rights protection for stereo video should be resolved in advance to avoid piracy. In this paper, based on our previous research work on both digital watermarking and stereo video, a novel adaptive digital watermarking algorithm for stereo video is proposed after analyzing the characteristics of stereo video. The watermark is embedded into disparity vectors of enhancement-layer stream, without affecting the structure of base-layer stream. The proposed algorithm is an adaptive one, which keep the quality degradation of the overall reconstructed video images being the smallest. The whole watermarking process is mainly performed in compressed domain without the need of the fully decoding of compressed stream. And watermark detection can be blindly performed without the need of the original video sequence. As a whole, the proposed algorithm is fast, simple. Experimental results with several stereo sequences show that the embedding of watermark has little affection on image quality and the degradation of PSNR with the adaptive algorithm is smaller than that with non-adaptive algorithm.

Keywords stereo video watermarking, disparity vector, adaptive algorithm, compressed domain, blind detection

1 引言

众所周知,立体视频是未来视频技术的一个重要发展方向^[1-3]。它是基于人眼的双目视差原理,

即双目各自独立地接收来自同一场景但有微小角度差异的两幅不同图像来形成双目视差以获得立体视觉,由于大脑能感受到图像的深度信息,使欣赏到的图像有强烈的逼真感,从而可体验到立体视觉效果。虽然到目前为止,立体视频技术还没有能真正的实

基金项目:浙江省自然科学基金项目(Y105592)

收稿日期:2005-10-19; 改回日期:2006-02-08

第一作者简介:朱仲杰(1976~),男,2004年3月获浙江大学信电系博士学位。现任浙江万里学院内聘教授,学科带头人。主要研究兴趣为2D与3D视频编码与传输、信息隐藏与新型数字水印技术、视频语义分析与检索等。近年来发表学术论文20余篇,其中SCI/EI检索论文10余篇。E-mail:zhongjiezh@hotmail.com

用化,离大规模的应用还有较长的路要走,但立体视频产品的知识产权保护是一个需要提前解决的问题,因为这个问题的尽快解决可以避免未来立体视频产品的盗版和侵权现象。

数字水印技术是近年发展起来的一种有效的用于数字产品版权保护和真伪性认证的新技术^[4,5],但目前面向立体视频的数字水印算法还很少。与图像水印相比,视频水印有更高的要求^[6],如:水印的嵌入和提取应直接基于压缩域;水印的嵌入不应影响视频码流的完整性;水印算法应简单、快速,应尽可能满足实时性的要求等。目前基于压缩域的视频水印技术研究非常活跃。在压缩域中嵌入水印,由于无需对视频码流进行重新解码,因此具有复杂度低、易于实现等优点。本文在笔者前期数字水印和立体视频研究工作的基础上^[7,8],通过详细分析立体视频的特点后,给出了一种面向立体视频的压缩域数字水印新算法。该新算法具有以下特点:(1)由于水印信息被直接加到视差矢量上,因此水印信息的嵌入对基本码流没有影响,也不会影响到基本码流重建后的视频图像质量;(2)水印的嵌入是采用自适应策略,即算法能根据水印信号和立体视频数据的特点,通过自适应调整水印的嵌入方案,以尽可能少地改变原视频数据,使嵌入水印后的图像质量损失达到最小,由于人眼的视觉特性,使增强层的视频图像质量微小下降不会影响到整个立体视频图像的视觉效果;(3)水印的嵌入和提取是直接基于压缩域进行,无需对整个视频流进行重新解码。另外,水印的提取使用盲检法,无需视频原图像。最后,水印的提取和检测只需要对视差矢量进行简单的计算和判断即可完成,因此算法简单、快速,可满足实时性的要求,且易于软硬件的实现。

2 立体视频水印的嵌入方案

在立体视频的编码中,为了达到高效编码的目的,一般采用 MCP(motion compensated prediction)与 DCP(disparity compensated prediction)相结合的编码方案(如图 1 所示)。这样的编码方案不但可以去除通道内部前后帧图像之间的时域冗余和通道间图像对之间的空域冗余,而且 MCP 与 DCP 相结合还可以对由于运动而引起的“遮挡”和“暴露”区域进行编码^[9-11]。

采用上述方案编码压缩后的视频流由基本层和增

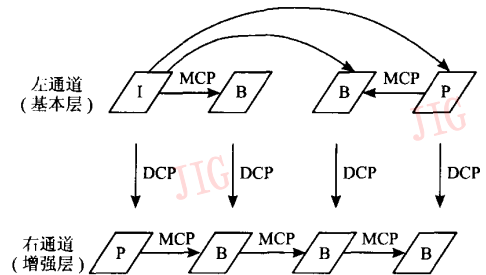


图 1 MCP 与 DCP 相结合的立体视频图像编码方案
Fig. 1 MCP and DCP combined scheme of stereo video coding

强层码流构成。其中,增强层码流可以看作由视频流头信息、DCT(discrete cosine transform)数据、运动矢量数据和视差矢量数据 4 个部分组成(如图 2 所示)。

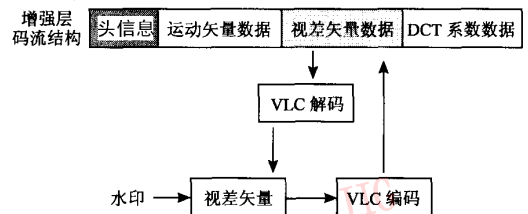


图 2 水印嵌入方案
Fig. 2 Watermark embedding scheme

在本文的水印算法中,水印信息被加到增强层码流的视差矢量数据上。该方法的优点是:①水印信息的嵌入不仅对基本码流没有影响,而且不会影响到基本码流重建后的视频图像质量;②由于对视差矢量的操作比较容易,且对码流的改动较小,因此不但不会影响码流的完整性,而且算法简单、快速。

3 立体视频水印新算法

设 (D_x^B, D_y^B) 表示某一编码宏块 B 的视差矢量,该视差矢量特征值 ρ_B 定义为

$$\rho_B = \text{mod}(2D_x^B, 2) + 2\text{mod}(2D_y^B, 2) \quad (1)$$

式中, $\text{mod}(\cdot, 2)$ 表示对 \cdot 取模 2 运算, ρ_B 的取值为 $\{0, 1, 2, 3\}$ 。

对每一个要嵌入水印的视差矢量 (D_x^B, D_y^B) , 设待嵌入的水印比特为 w_B , 根据视差矢量特征值 ρ_B , 总共有以下 6 种基本嵌入方案:

(1) 当 $w_B = 0$ 时, 如果 $\rho_B \neq 0$, 则改变 D_x^B, D_y^B , 使得 $\rho_B = 0$; 当 $w_B = 1$ 时, 如果 $\rho_B \neq 1$, 则改变 D_x^B, D_y^B ,

使得 $\rho_B = 1$ 。

(2) 当 $w_B = 0$ 时, 如果 $\rho_B \neq 0$, 则改变 D_x^B, D_y^B , 使得 $\rho_B = 0$; 当 $w_B = 1$ 时, 如果 $\rho_B \neq 2$, 则改变 D_x^B, D_y^B , 使得 $\rho_B = 2$ 。

(3) 当 $w_B = 0$ 时, 如果 $\rho_B \neq 0$, 则改变 D_x^B, D_y^B , 使得 $\rho_B = 0$; 当 $w_B = 1$ 时, 如果 $\rho_B \neq 3$, 则改变 D_x^B, D_y^B , 使得 $\rho_B = 3$ 。

(4) 当 $w_B = 0$ 时, 如果 $\rho_B \neq 1$, 则改变 D_x^B, D_y^B , 使得 $\rho_B = 1$; 当 $w_B = 1$ 时, 如果 $\rho_B \neq 2$, 则改变 D_x^B, D_y^B , 使得 $\rho_B = 2$ 。

(5) 当 $w_B = 0$ 时, 如果 $\rho_B \neq 1$, 则改变 D_x^B, D_y^B , 使得 $\rho_B = 1$; 当 $w_B = 1$ 时, 如果 $\rho_B \neq 3$, 则改变 D_x^B, D_y^B , 使得 $\rho_B = 3$ 。

(6) 当 $w_B = 0$ 时, 如果 $\rho_B \neq 2$, 则改变 D_x^B, D_y^B , 使得 $\rho_B = 2$; 当 $w_B = 1$ 时, 如果 $\rho_B \neq 3$, 则改变 D_x^B, D_y^B , 使得 $\rho_B = 3$ 。

改变 D_x^B, D_y^B 是指进行 $D_x^B = D_x^B + 0.5$ 或 $D_y^B = D_y^B + 0.5$ 的运算, 或者同时改变。自适应算法的思想是, 根据待视差矢量的特点, 通过选择最佳的嵌入方案来使得嵌入后的图像质量损失最小。定义 $S = \{s_0, s_1, s_2\} = \{000, 001, 010, 100, 101, 110\}$ 为 6 种嵌入方案的二进制代码, 其分别表示 6 种嵌入方案。

设 $N_i (i = 0 \sim 10)$ 表示图像组中嵌入到第 i 帧图像中的水印比特总数, $W_i = \{w_{i,l}\} (l = 0 \sim N_i - 1)$ 表示要嵌入到第 i 帧图像的水印比特流, $(D_{i,x}^B, D_{i,y}^B) (l = 0 \sim N_i - 1)$ 表示第 i 帧图像中第 l 个待嵌入水印的宏块 B_l 的视差矢量, $\rho_{i,B_l} (l = 0 \sim N_i - 1)$ 表示相应的视差矢量的特征值。定义 $c_1, c_2, c_3, c_4, c_5, c_6$ 分别表示采用 6 个方案嵌入水印时需要改变的视差矢量的个数, c_{\min} 定义如下:

$$c_{\min} = \min \{c_1, c_2, c_3, c_4, c_5, c_6\} \quad (2)$$

c_{\min} 表示分别采用上述 6 种嵌入方案嵌入水印时, 最优方案所需改动的视差矢量的个数。本文的自适应算法就是根据 c_{\min} 值来选择每一帧图像的水印嵌入方案。嵌入方案代码作为水印的头信息一起嵌入到图像中, 即实际嵌入的水印序列 \hat{W}_i 为

$$\hat{W}_i = \{s_{i,0}, s_{i,1}, s_{i,2}\} \cup \{w_{i,l}\} (l = 0 \sim N_i - 1) \quad (3)$$

水印信息并不是加到所有的视差矢量上, 而是选择其中绝对值较大的视差矢量来嵌入水印信息。给定阈值 T , 对于某个视差矢量 (D_x^B, D_y^B) , 如果满足以下条件:

$$|D_x^B| + |D_y^B| \geq T \quad (4)$$

则在该视差矢量上嵌入水印, 否则不嵌入水印。视频水印的嵌入新算法简要流程如图 3 所示。

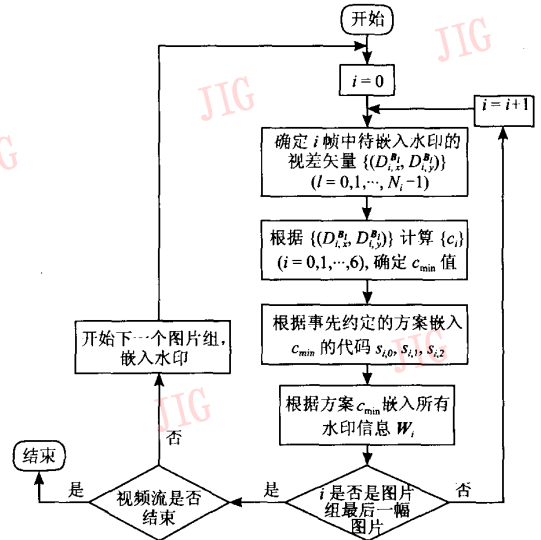


图 3 水印嵌入流程图

Fig. 3 Flow chart of watermark embedding

上述的水印嵌入算法只需要对视差矢量进行简单的判断和加减运算即可嵌入水印, 其不仅对数据流的改动很小, 且码流比特率基本保持不变。由于其计算复杂度很小, 因此能够进行水印的实时嵌入。

水印提取是水印嵌入的逆过程, 其算法不仅非常简单, 而且具备盲检功能, 也无需视频原图像。

4 实验仿真与结果

为了检验算法的有效性, 还对水印算法进行了计算机仿真测试, 仿真测试使用了以下标准立体视频测试序列:

(1) Train_and_Tunnel 图像序列, 共 192 帧, 图像尺寸为 720×576 。

(2) Herve 图像序列, 共 22 帧, 图像尺寸为 512×512 。

(3) Pascal 图像序列, 共 22 帧, 图像尺寸为 512×512 。

(4) lml 图像序列, 共 22 帧, 图像尺寸为 512×512 。

设 δ 和 WDR 分别表示加入水印后的视频图像的平均信噪比损失和水印正确检出率 (watermark detect ratio, WDR)^[8]。图 4 ~ 图 6 和表 1 ~ 表 3 给

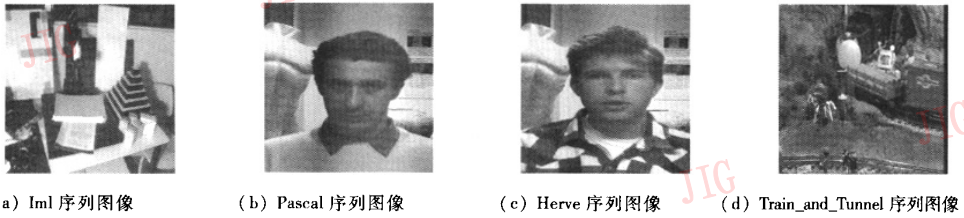


图 4 未加水印的增强层图像

Fig. 4 Images of enhancement layer before watermarking

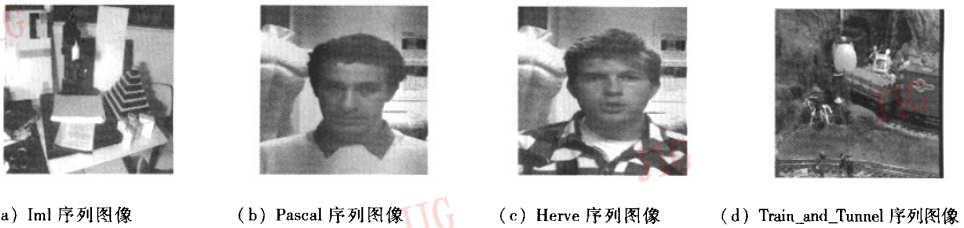


图 5 嵌入水印后的图像

Fig. 5 Watermarked images

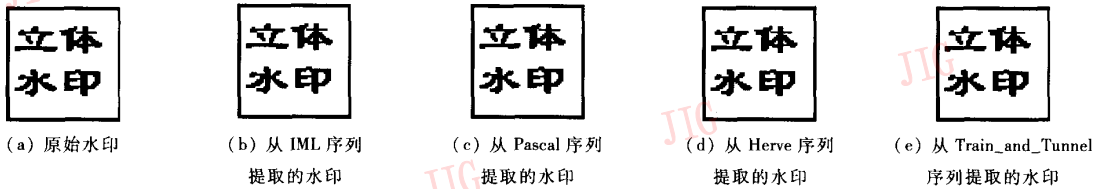


图 6 水印检测结果

Fig. 6 Watermark detection results

表 1 嵌入水印后的图像质量(平均信噪比)下降值和水印检出率

Tab.1 Degradations of watermarked images and WDRs

指标	序列			
	Iml	Herve	Pascal	Train_and_Tunnel
δ (dB)	0.023 5	0.107 8	0.041 3	0.148 9
水印检出率	1.000 0	1.000 0	1.000 0	1.000 0

表 2 未加水印的视频图像水印检测结果

Tab.2 Detection results of non-watermarked video sequences

指标	序列			
	Iml	Herve	Pascal	Train_and_Tunnel
水印检出率	0.002 3	0.042 3	0.014 2	0.032 5

出了 4 组图像序列的实验结果,图 4 是未嵌入水印之前的增强层图像;图 5 是嵌入水印后的对应图像;

表 3 分别采用自适应方案和非自适应方案嵌入水印后的图像质量下降值(δ)

Tab.3 Comparing of degradations(δ) of watermarked images using adaptive and non-adaptive schemes respectively

水印嵌入方案	不同序列嵌入水印后的信噪比下降值 (dB)			
	Iml	Herve	Pascal	Train_and_Tunnel
自适应方案	0.023 5	0.107 8	0.041 3	0.148 9
非自适应方案	0.077 2	0.109 1	0.093 6	0.216 3

图 6 是从嵌入水印后的增强层图像中提取出的水印图像。表 1 给出了水印嵌入前后的视频图像的平均信噪比下降值 δ 和水印正确检出率。

从图 5、图 6 以及表 1 可以看出,嵌入水印后图像的信噪比下降很小,即水印的嵌入具有不可见性。4 幅图像的水印检出率都为 1,这说明水印算法是有

效的。表2给出了算法对未嵌入水印信息的视频图像进行检测的结果。由于视频图像未加水印,表中的水印正确检出率很小,这表明在没有加入水印的视频图像中提取的不是水印信息,而可以看作噪声信息。表3给出了自适应算法和非自适应算法的水印嵌入结果,从中可以看出,采用自适应算法嵌入水印后的视频图像质量下降值要小于非自适应算法的质量下降。

5 结论

对立体视频产品的知识产权保护是一个需要提前解决的问题,因为这个问题的提早解决可以避免未来立体视频产品的盗版和侵权现象。本文提出的面向立体视频的数字水印新算法是将水印信息嵌入到增强层的视频码流中,这种水印嵌入对基本层的视频码流没有影响。该算法采用自适应策略,并根据水印信号和立体视频数据的特点,自适应地调整嵌入方案,通过尽可能少的改变原视频数据,以使嵌入水印后的图像质量损失达到最小。其水印的嵌入和提取操作基于压缩域进行,且水印检测具有盲检功能。整个算法简单、快速。实验结果显示本文提出的新算法是有效的。

参考文献 (References)

- 1 Rhys Hawkins. Digital Stereo Video: Display, Compression and Transmission[D]. Ph. D. thesis, the Australian National University, Canberra, Australian, 2002.
- 2 Park J, Yang K H, Iwadata Y. Efficient representation and compression of Multi-view images [J]. IEICE Transactions on Information and Systems, 2000, **E83-D(3)**: 2186 ~ 2188.
- 3 Bas P, Chassery J M, Macq B. Geometrically invariant watermarking using feature points [J]. IEEE Transactions on Image Processing, 2002, **11(9)**: 1014 ~ 1028.
- 4 Voloshynovskiy S, Deguillaume F, Pun T. Multi-bit digital watermarking robust against local nonlinear geometrical distortion [A]. In: Proceedings of the 2001 International Conference on Image Processing [C], Thessaloniki, Greece, 2001: 999 ~ 1002.
- 5 Dong P, Brankov J G, Galatsanos N. Geometric robust watermarking based on a new mesh model correction approach [A]. In: Proceedings of the 2002 International Conference on Image Processing [C], Rochester, New York, USA, 2002: 493 ~ 496.
- 6 Langelaar G C. Real-time Watermarking Techniques for Compressed Video Data [D]. Ph. D. thesis, Delft University of Technology, Delft, Holland, 2000.
- 7 Zhu Z J, Jiang G Y, Yu M, et al. New adaptive watermarking algorithm [A]. In: Proceedings of the 2002 IEEE Region 10 Conference on Computers, Communications, Control and Power Engineering [C], Beijing, China, 2002: 152 ~ 155.
- 8 Zhu Zhong-jie, Jiang Gang-yi, Yu Mei, et al. New algorithm for video watermarking in MPEG-2 compressed domain [J]. Acta Electronica Sinica, 2003, **32(1)**: 21 ~ 24. [朱仲杰, 蒋刚毅, 郁梅等. MPEG-2 压缩域的视频数字水印新算法 [J]. 电子学报, 2003, **32(1)**: 21 ~ 24.]
- 9 Zhu Z J, Jiang G Y, Yu M, et al. Fast disparity estimation algorithm for stereo video coding [A]. In: Proceedings of the 2002 IEEE Region 10 Conference on Computers, Communications, Control and Power Engineering [C], Beijing, China, 2002: 285 ~ 288.
- 10 Sun J, Zheng N N, Shum H Y. Stereo matching using belief propagation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, **25(7)**: 787 ~ 800.
- 11 Davis J, Ramamoorthi R, Rusinkiewicz S. Space-time stereo: A unifying framework for depth from triangulation [A]. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition [C], Nice, France, 2003: 359 ~ 366.