

具备上下文依赖及鲁棒性的视频水印算法^{*})

田捷¹ 张新访¹ 王同洋¹ 吴俊军¹ 程明²

(华中科技大学信息科学与技术研究所 武汉 430074)¹

(清华大学计算机软件与理论研究所 北京 100084)²

摘要 数字水印技术可用于内容的鉴定及防止伪造,被看作是对已发布内容避免其受到未经授权的复用的重要技术。所有权信息以水印的方式嵌入至数字内容中,从而能够提供版权保护、所有权声明、数码内容完整性核查及其在遭受攻击时提供版权侵害的证据,因而水印技术成为构造数字媒体版权备选的方案。本文介绍了水印方法与有信编码与嵌入的关联性,设计了一个有效的方法,使用码字子集来代表预插入的0或1数据位。此外,还提出了一个码字生成方案,能控制码间距,从而保证了水印的鲁棒性与信号的保真度。通过比较测试可以看出,我们提出的方法在视频压缩方面具备出色的稳健性。

关键词 数字水印,离散余弦变换,版权保护,有信编码

Robust Video Watermarking Algorithm with Correlative Context

TIAN Jie¹ ZHANG Xin-Fang¹ WANG Tong-Yang¹ WU Jun-Jun¹ CHENG Ming²

(Institute of Information Sciences and Technology, Huazhong University of Science and Technology, Wuhan 430074)¹

(Institute of Computer Software and Computing Theory, Tsinghua University, Beijing 100084)²

Abstract Digital watermarking, a popular technique for content authentication and forgery prevention, is viewed as an enabling technology to protect distributed content from unauthorized reuse. Digital watermarks of ownership embedded onto digital content offer copyright protection, ownership assertion, integrity checks for digital content, and can provide evidence of copyright infringement after an attack. Watermarking technologies have been envisioned as a potential means for establishing ownership on digital media objects. In this paper, a watermarking method that relies on informed coding and informed embedding is presented. We devise an efficient method, which uses a subset of various codewords to represent the 0 and 1 message bits to be embedded. We propose a codeword generation scheme that keeps control of the distance between codewords in order to secure fidelity and robustness of the watermark. When compared to existing video watermarking schemes, our method yields superior robustness to video compression.

Keywords Digital watermarking, DCT transform, Copyright protection, Informed coding

1 引言

数字水印已经成为对数字化内容实施版权保护的一项必备技术^[1]。这种技术通过将隐藏信息对原有多媒体内容做不可察觉的修改,用于对内容授权所有者的识别及对非法复制拷贝刻录设备的侦测。对于数字视频而言,水印技术不仅仅需要仔细的设计,使得数字视频在播放的时候不会察觉到水印的存在,同时应该对在 MPEG-II 数字视频内容压缩编码等应用背景下出现的经常性数据失真体现出鲁棒性。数字水印算法通常在频率域中嵌入二进制代码消息,其中一个最为成功的例子就是 Porter's^[2] 提出的扩频方法,它具有如下几个优点:把一系列伪随机特征码字嵌入到将要显示出的数据信息中;被插入到频率域部分的信号能量与原有的这些频率域部分信号能量不存在强关联性;当数字水印被散布于多个频率带的时候,该算法可以针对通常出现的信息失真提供一定的稳健性。针对 Porter's^[2] 方案的改进,算法 Mayer's^[3] 采用

有信编码与嵌入的方法。有信嵌入就是在对数字水印编码之前,首先对原有内容数据做分析检验,即在对于通常出现的信息失真而提出的鲁棒性能、数字水印对于视频图像逼真度与保真度造成的可感知的影响方面,这两者之间寻求一种恰当并且有效以期达到一种平衡与折衷状态的过程与方法。基于 Costa 的理论,提出有信编码并应用到水印系统中,宿主信号在编码器端或嵌入端是已知的,将宿主信号作为一种辅助信息(side-information),调整编码器的输出结果,使水印自适应于宿主信号,这种编码方法称为有信编码。亦即有信编码就是在算法执行处理阶段,通过对比原有内容数据信息,在多组备选的系统特征码字之中做出选择,使得这些特征码字被嵌入到视频数据信号中时,能使得原始图像显示信息达到最小的变形失真。Silva's^[4] 提出了基于 Mayer's^[3] 设计的有信编码与嵌入的修正方法,由于这种修正方案仅需要较少的计算量,因而它更适合于视频应用环境中。此外,在误码率和丢包率等性能的 QoS(网络服务质量)方面,方案 Silva's^[4] 相比较

^{*}湖北省自然科学基金(No. 2002ABA055);武汉市青年科技晨光计划(No. 20015005078)。田捷 博士,主要研究方向:网络信息安全、普适与网格计算、知识产权保护;张新访 教授,主要研究方向:工程数据管理、嵌入式操作系统智能卡;王同洋 副教授,主要研究方向:工程数据库管理系统三维 cad 系统、嵌入式操作系统;吴俊军 副教授,主要研究方向:嵌入式系统、信息安全;程明 硕士,主要研究方向:计算机图形压缩方法与优化。

于 Mayer's^[3] 而言更胜一筹, 因为其消息数据位通过格式编码时不会受到约束关联, 但是在面临相关类型的外部系统攻击情况下的报文差错率方面, 即当执行编码后的数据帧中存在至少一个报文数据位出现不正确数值的百分比, 方案 Silva's^[4] 却没有得到明显的改善。

在 Doerr's^[5] 方案中, 提出将稳健独立分量分析的方法 (Independent Components Analysis, 简称为 ICA) 引入到数字水印的领域里面来, 这种想法在方案 Hussein's^[6] 得到进一步的发展, 并且在 Hussein's^[6] 中提出了被称为 ICA-Trellis 视频水印且归属于扩展频谱范畴的专门解决方案, 它使用上述提到的 ICA 方法将各段的 1 帧 (Intra-frame) 降低关联性。这种解相关方法能够使得所有产生的数字信号都不相互关联, 从而大大降低由于引入数字水印而带来的不同程度的数据畸变, 同时 ICA 方法可以用来在数据报文插入一组符合统计规律的独立信号之前预测是否存在残差帧; 另外, 该方案中也提出了一种改良的格形编码方式, 能用来生成一系列特征码字阵列, 这些代码字能将最有效能的数字水印嵌入至视频帧中, 从而使整个数据帧产生最小的畸变。

我们在本文将介绍一种新的算法, 它能在视频数据压缩等方面相比方案 Hussein's^[6] 中提出的 ICA-Trellis 方法而体现出更好的效率与性能。我们的算法使用了一种代码字生成技术方案, 能较为有效地控制代码字间距, 从而在视频显示信号的保真度与被插入其间数字水印的鲁棒性之间达到一种动态平衡与折衷。

2 数字水印的设计算法

我们的方法联合了有信编码与嵌入式的技术。我们没有如同方案 Mayer's^[3] 那样把一幅图像看作是一个不可分的整体, 而是根据 8×8 像素块包含的图像独立特征分量来选取代表准备嵌入数据的特征码字, 这些码字体现为数据位 0 或 1, 被插入的数据信息是二进制的, 而选取出来的特征码字会代表各自对应隐藏在这 8×8 像素块中的数据信息位, 其过程如图 1 所示。

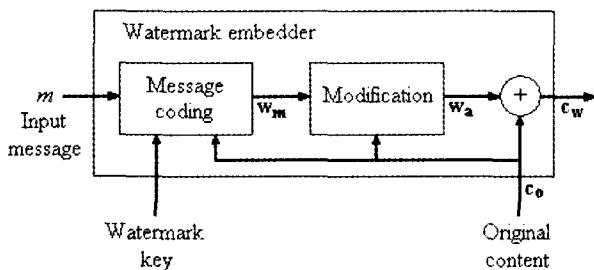


图 1 有信编码与嵌入式插入水印过程

因而, 数据信息的长度就不能超过一幅图像中所有数据块的数量。对每一个图像数据块实施离散余弦变换 (Discrete Cosine Transform, 简称为 DCT), 如图 2 所示, 修改第 L 个最小的交流系数, 使得它包含特征码字的成分。首先, 将每一个图像块的第 L 个系数复制到一个长度为 L 的一维矢量中, 我们称这个系数阵列为抽取矢量 v_0 , 然后去寻找一个与矢量 v_0 有最紧密关系的特征码字, 把这个代码字嵌入到对应的块中。如果没有找到非常满意的相关性, 那么矢量 v_0 就应做一定的修正, 直到这样的相关性满足为止。我们称这样被修改的

矢量为水印矢量 v_w 。在某些情况下, 这个矢量会与原矢量 v_0 相同。特征码字产生过程如下:

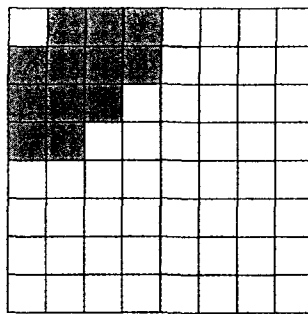


图 2 对 8×8 像素块实施 DCT 变换

以伪随机方式生成长度为 L 的 P 个特征码字组, 其中会有 $P/2$ 个代码字代表 0, 而其余则代表 1。我们可将其划分为两个子集, 用 A 与 B 表示: 代表值 0 的子集 $\textcircled{1}A = \{a_1, a_2, \dots, a_{P/2}\}$ 与代表值 1 的子集 $\textcircled{2}B = \{b_1, b_2, \dots, b_{P/2}\}$ 。我们知道, 构造码字子集的方式会对于数字水印对抵抗普通系统攻击的鲁棒性以及被嵌入水印的视频图像信息逼真度与保真度都有着极大的影响, 因此, 我们用如下方法来设计这些码字。首先, 我们使用只有用户才知道的水印密钥 K 作为生成码字 a_1 的依据并将其长度规范化, 将生成的连续两个码字间距作为参数 d_0 的值。如果把 d_0 与第一个码字 a_1 相加, 就会生成第二个码字 a_2 , 也将其字长规范化。这就是使用原始密钥 K 作为依据来随机生成矢量的方法, 其余的码字也以相同的方式创建出来, 并受到如下的约束: 在 i 从 2 至 $P/2$ 的闭区间上, 有 $\textcircled{3} |a_{i-1} - a_i| = d_0$, 其中的 a_{i-1} 与 a_i 字长也都规范化。另外, 子集 B 的构造方式如下: 在 i 从 1 至 $P/2$ 的闭区间上, 有 $\textcircled{4} b_i = -a_i$ 。从我们通过计算机模拟实验的分析数据来看, 相比原来的那种单一随机模式下产生的码字而言, 采用这种方法构造出来的特征码字能够同时在保真度与鲁棒性等方面体现出较强的性能特征。选择适当的 d_0 与 L, P 会导致两个码字子集在水印鲁棒性与图像保真度之间不断地出现如图 3~5 那样的变化而形成动态平衡。为了清楚说明, L 的值设定为 3 且不考虑 d_0 的值。这样一来, 如果 P 值越大, 则视频图像的畸变就越小。反之, 如果 P 值越小, 则视频图像的畸变就越大。这是因为码字分布的区间越大, 就会促使与矢量 v_0 逼近的概率越高。举个例子, 如果让 a_i 矢量代表准备插入的数据位, 那么就会有一个较大的数据源 a'_i 可供选择, 这就意味着整体图片有较好的逼真度。然而对于另外一个子集 B 而言, 集合越大却不是好事, 因为这会造成由于两个子集抽取出来的码字太过于接近而使得水印的鲁棒性能降低, 这样一来就使得 P 值不能太大。与此同时, 我们可以看到, 如果 d_0 的值设置得过小, 会使水印的鲁棒性能增强, 但同时会使 8×8 数据块的视觉效果质量被大大地破坏。这种情况在图 4 中体现出来, 图中 P 被设定为一个合理的值 71 且 d_0 被设定为一个较小的值 0.2 之后, 产生了相距甚大的两个子集。最后在图 5 中描述了当 P 被设定为一个合理的值 77 且 d_0 被设定为一个合理的值 1.3 之后的情况, 从中我们可以发现子集之间的码字间存在着适当的间距, 并且每个子集都有足够数量的码字。就矢量 v_0 的长度 L 而言, 每一个数据块中的水印都能修改系数的值, 同时由于这些水印都必须插值到较小的交流

频率系数中,因此我们推荐 L 的值低于 20 以下为好。

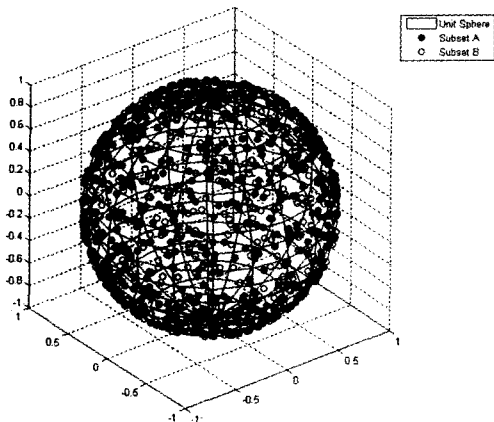


图3 $P=988, L=3$ 时高保真度、低鲁棒性

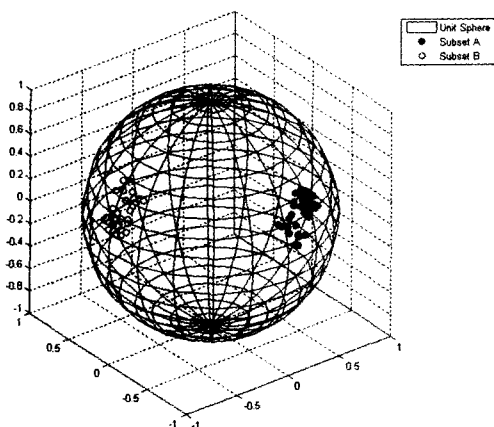


图4 $P=71, d=0.2$ 时低保真度、高鲁棒性

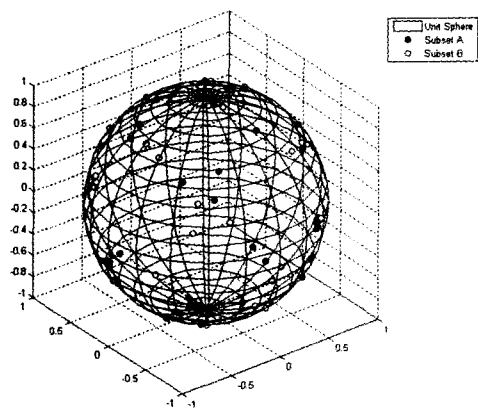


图5 $P=77, d=1.3$ 时保真度与鲁棒性平衡

一旦构建出了特征码字,随后就要把这些数字信息嵌入到图像中去。对于一个 8×8 的数据块而言,抽取矢量 v_0 与代表即将插入至某一特定块的 0 或 1 值设定位的 $P/2$ 个码字中的任一个之间的系数都会被计算出来。这里要注意的是:虽然所有的码字都被规范化,但是 v_0 是没有被规范化的,这是因为如果改变了 v_0 的大小,就会使我们提出的方法发挥不出它应有的性能。假设我们准备把被子集 A 所代表的位 0 值插入。在 i 从 1 至 $P/2$ 的闭区间上,设定码字与 v_0 间最相关系数为 a_{\max} ,有公式⑤ $a_{\max} = \max_i(a_i \cdot v_0)$ 。设定 a_{\max} 与 v_0 之间的相关系数减去 b_i 与 v_0 之间的相关系数为 R_{0i} ,有公式

⑥ $R_{0i} = (a_{\max} - b_i) \cdot v_0$ 。如果在 i 从 1 至 $P/2$ 的闭区间上,存在⑦ $R_{0i} > 0$ 并且如果没有图像的畸变,则条件⑦足以让解法器能够提取出正确的水印。然而,假设遇到系统攻击或者图像的畸变,条件⑦就无法保证一个具备鲁棒性数字水印仍然存在,因此我们需要确保 R_{0i} 是一个较大的正数。为了做到这一点,引入鲁棒性极限值 R_i ,使得在从 1 至 $P/2$ 的闭区间上所有的 i 而言,都有⑧ $R_{0i} > R_i$ 。这就是说:矢量 v_0 与子集 B 中有害的码字间存在足够大的间距。如果条件⑧无法实现,那么就需要重复修正 v_0 ,直到条件满足为止。设定被改进的矢量为 v_w ,根据方程式⑥,对于在数值区间中的所有 i 而言, R_{0i} 的值越大,码字 a_{\max} 被插入至块中的概率也就越大。同样地,对于每一个 b_i 而言,如果 R_{0i} 大于鲁棒性极限值 R_i ,那么在水印矢量 v_w 与子集 B 中有害的码字之间就存在足够大的间距。从另一个方面来说,如果 R_{0i} 的值小于 R_i ,那么 v_w 就必须作修正,直到每一个 R_{0i} 都大于极限值 R_i 。这样的修正过程是从第一次发现对于 R_{0i} 的最小值而言出现的所有有害的码字子集 B 开始的,亦即⑨ $R_{0\min} = \min_i R_{0i}$ 。为了对图像保真度的影响达到最小, $R_{0\min}$ 与相应的矢量 b_{\min} 被用来修正水印矢量 v_w ,从而使得在下次 R_{0i} 被计算特定的 b_{\min} 时, R_{0i} 能生成恰好相等于 R_i 的值且存在与先前 v_w 值之间的最小欧几里得距离。这样的修正可按照如下方式实现:

⑩ $v_w + (R_i - R_{0\min})(a_{\max} - b_{\min}) / |a_{\max} - b_{\min}| \rightarrow v_w$, 这个过程可反复执行,即重新计算 $R_{0\min}$ 与重复修正矢量 v_w ,直到 $R_{0\min}$ 比 R_i 大很多为止。其结果 v_w 将会成为水印矢量,亦即被修正的 8×8 数据块系数。最后的步骤就是将反离散余弦变换运用到这些数据块中,以获得空间数据信息。整个处理过程对于图像所有数据块都会反复执行。

当解码报文信息的时候,图像仍会被划分为 8×8 的数据块,且离散余弦变换也会被应用。用低频系数来获取水印矢量 v_{un} ,它是实施编码后 v_w 的带有噪音的译文。接着, v_{un} 与所有的 P 个码字之间的系数都会计算出来。带有最大系数的码字就会被解法器选作为特征码字 a_{\max} 并且从 a_{\max} 所属的子集中选取出的数据位也会被选定为原始数据位。

3 算法性能评价

我们已成功地将设计出的水印算法应用于多个视频序列中,并通过对 MPEG-II 的压缩攻击来测试设计方案的鲁棒性。我们将与 Hussein's^[6] 提出的 ICA-Trellis 水印方案在性能方面做一个较为详细的比较。一共有五个视频序列(container, hall monitor, mother and daughter, news, suzie),每一个序列包含有 150 帧供我们作测试。我们使用规格为 15 的图像群且只将每帧的亮度 Y 帧实施水印插入。按照 QCIF 图像格式(176×144)中规定的次序,使用我们提出的方法能将最大为 396 个信息数据位嵌入到每帧当中去。然而,为了与 ICA-Trellis 方案能承受的有效载荷保持大致相同,我们决定只进行每帧当中的 44 个信息数据位的嵌入,这 44 个信息数据位共分 9 次插入到同一帧当中。由于存在冗余信息,我们能使用一个较小的鲁棒性极值 R_i ,因此也会实现较小帧信息畸变的目的。给予 ICA-Trellis 方案的有效载荷仅有每帧 31 个信息数据位。对于两种方法而言,图像质量都能保持大致相同(峰值信噪比 PSNR=42dB)。通过图 6 我们可以看到视频序列中的名为 Suzie 一帧在实施水印插入前后的视觉效果图。

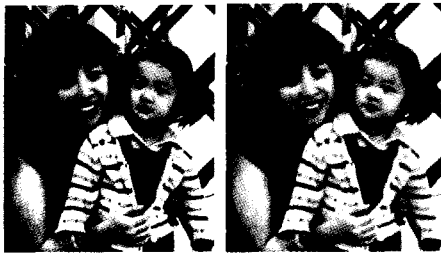


图6 视频帧 suzie 加入水印前后对比图

通过设定 P 的值为 128, 创建一个代表位 0 且含有 64 个码字的子集, 及另外一个代表位 1 且含有 64 个码字的子集。这样, 足够多的码字可以确保嵌入了水印的视频有较高的鲁棒性能, 同时也使得在编解码阶段的系统承担合理的计算量。其中的码字被规范化后的长度 L 为 16, 且设定 0.4 为间距 d_0 的值(图 7)。

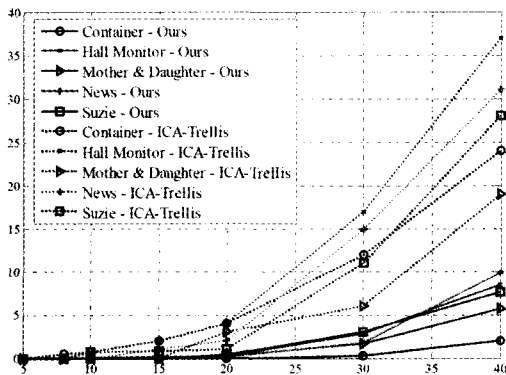


图7 实施 MPEG-II 压缩后鲁棒性比较

与 ICA-Trellis 方案比较所有五个视频数据帧实施不同的压缩率后体现出来的性能, 发现在压缩比为 20:1, 亦即相对应的码率达到 758kbps 的时候, 我们的方法数据位误码率仅有 0.2%, 而 ICA-Trellis 方案的数据位误码率却达到了 2%~4% 的程度, 这意味着它的平均数据位误码率为 3%, 相当于平均下来就有 15 倍之多的数据位错误。而当压缩比率达到 30:1, 亦即相对应的码率达到 505kbps 的时候, 采用我们的方法平均仅会有少于 2% 的嵌入数据信息被丢失, 而相比之下的 ICA-Trellis 方案的数据位误码率却达到了 6%~

(上接第 246 页)

结合, 成功地实现了目标模型根据当前跟踪环境的自动组合, 克服了 CamShift 跟踪算法不能适应目标外观变化的固有缺陷, 证明了用多模型思想解决变外观目标跟踪问题是可行的。然而, 本文仅考虑了目标做平移运动时的情形, 对更复杂运动形式还暂未考虑; 另外, 多模型的引入等于放宽了模型的约束条件, 这在一定程度上减小了模型对目标的识别能力, 如何通过引入目标边缘、形状等信息来提高算法的抗干扰能力? 这些都是今后需进一步研究的内容。

参考文献

- 1 Bradski G, Clara S. Computer vision face tracking for use in a perceptual user interface[J]. International Technology Journal, 1998, 2:1~15
- 2 Comaniciu D, Ramesh V, Meer P. Kernel-based object tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(5): 564~577

17% 的程度, 这意味着它的平均数据位误码率为 12.2%, 相当于平均就有 6 倍之多的数据位错误。最后, 当压缩比率达到 40:1, 亦即相对应的码率达到 379kbps 的时候, 采用我们方法而产生的数据位误码率范围在 2%~9.8%, 而相比之下的 ICA-Trellis 方案的数据位误码率却达到了 19%~37% 的程度, 这意味着它的平均数据位误码率为 27.8%, 平均就有 4 倍之多的数据位错误。

结论 本文提出了一个与内容关联的视频水印方案。一个每帧 44 位的二进制数据信息被嵌入视频图像中。经过水印嵌入之后的视频序列仍具备值为 42dB PSNR 的保真度。通过对 MPEG-II 压缩比率改变的测试, 不难发现我们的方法所具备的鲁棒性能优于近期被认为是最优视频水印插入算法的 ICA-Trellis 方案。分别在压缩比率为 20:1, 30:1 及 40:1 的时候, 我们提出的方法与 ICA-Trellis 相比, 体现出对应的 15 倍、6 倍及 4 倍的性能优化, 这对于数字水印技术的发展与完善有着极其重要的应用价值。

参考文献

- 1 Toyota Y, Mizuki T, Sone H. A study on security assessment of digital watermarking[J]. In: Record of Electrical and Communication Engineering, 2006. 547~548
- 2 Porter J, Rajan P. Image adaptive watermarking techniques using models of the human visual system[J]. In: IEEE Proceedings of the 38th Southeastern Symposium on System Theory, 2005. 354~357
- 3 Mayer J, Silva R. Efficient informed embedding of multi-bit watermark[J]. In: IEEE International Conference on Acoustics, Speech, and Signal Processing, 2004. 389~392
- 4 Silva R, Mayer J. Informed embedding for multibit watermarking [A]. In: Proceedings 16th Brazilian Symposium on Computer Graphics and Image Processing[C]. Sao Carlos, Brazil: IEEE Comput. Soc, 2003; 214~21
- 5 Gwenaél D, Ingemar J, Miller L. An efficient algorithm for informed embedding of dirty-paper trellis codes for watermarking [J]. In: IEEE International Conference on Image Processing, 2005. 697~700
- 6 Hussein J, Franck D. An ICA Based Algorithm for Video Watermarking[J]. In: IEEE International Conference on Acoustics, Speech and Signal Processing, 2005. 805~808
- 7 Mayer J, Bermudez J. Multi-bit informed embedding watermarking with constant robustness[A]. In: International Conference on Image Processing[C]. Piscataway, NJ, USA, IEEE, 2005. 669~72
- 8 Yamashita T, Sato A, Yamashita K. Digital watermarking using DCT fractal coding without original image information[J]. Electrical Engineering, 2006. 48~55
- 9 Stern H, Efron B. Adaptive color space switching for tracking under varying illumination[J]. Image and Vision Computing, 2005, 23(3): 353~364
- 10 Peng N S, Yang J, Liu Z. Mean shift blob tracking with kernel histogram filtering and hypothesis testing[J]. Pattern Recognition Letters, 2005, 26(5): 605~614
- 11 Nguyen H T, Ji Q, Arnold W M. Spatio-temporal context for robust multitarget tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(1): 52~64
- 12 Black M J, Jepson A D. Eigen Tracking: Robust matching and tracking of articulated objects using a view-based representation [J]. International Journal of Computer vision, 1998, 26(1): 63~84
- 13 Ho J, Lee K C, Yang M H, et al. Visual tracking using learned linear subspace[C]. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2004
- 14 Collins R, Liu Y X. Online Selection of Discriminative Tracking Features[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(10): 1631~1643