

文章编号: 1001-0920(2006)04-0371-05

## 基于内容分类的乘性水印最优检测

王金伟, 刘光杰, 戴跃伟, 王执钊  
(南京理工大学 自动化系, 南京 210094)

**摘 要:** 提出了一种 DCT 域内基于内容分类的最优检测方法. 该方法利用 DCT 变换中图像块低频能量的差异将 DCT 系数划分成平滑和纹理两类, 并分别对这两类数据进行参数估计以实现基于最大似然比的最优检测. 实验表明该方法较原有的整体参数估计和基于频段分类的参数估计在性能上有明显的提高.

**关键词:** 水印; 最优检测; 分类; 纹理; 平滑  
**中图分类号:** TP391 **文献标识码:** A

## Multiplicative Watermark Optimum Detection Based on Content Classification

WANG Jin-wei, LIU Guang-jie, DAI Yue-wei, WANG Zhi-quan

(Department of Automation, Nanjing University of Science and Technology, Nanjing 210094, China. Correspondent: WANG Jin-wei, E-mail: wjwei\_2004@163.com)

**Abstract:** A novel optimum detection method based on content classification is proposed, which uses the difference of energy sum of DCT coefficients in low and middle frequency to classify the image block into two classes. The optimum detection based on the maximum likelihood is realized through the parameter estimation of two kinds of image block DCT coefficients. The experiments show the performance of the new method is remarkably improved comparing with global and local parameter estimation based methods.

**Key words:** Watermarking; Classification; Optimum detection; Texture; Plain

### 1 引 言

近年来随着网络及多媒体的高速发展, 商业多媒体产品的版权保护等问题日益受到重视. 数字水印技术的发展为解决版权问题提供了有利的手段, 同时提出了大量用于版权保护的数字鲁棒数字水印算法<sup>[1-3]</sup>, 但是这些算法主要通过设计嵌入端的嵌入算法来提高水印的鲁棒性, 以达到抵抗各种可能针对水印的攻击. 本文讨论的提高鲁棒性的办法是在提取端设计基于最大似然比的最优检测方法实现针对乘性扩频水印的鲁棒性.

水印检测器的鲁棒性可以归结为在攻击下水印检测值的大小, 具体而言, 对固定嵌入算法和嵌入强

度的嵌入端, 更为鲁棒的检测器应该具有更大的检测值. 目前的研究主要是基于相关公式<sup>[11]</sup>及其变形公式<sup>[2,3]</sup>的检测. 基于检测理论的最优检测思想, 将水印的提取问题转化为信号的检测问题, 这样可以方便地设计盲检测的水印系统, 更为关键的是在这种模式下水印系统的检测性能得到更多的提高.

最优检测依赖系数分布模型的建立, 因此模型越精确检测的可信度越高. 文献[4~7]中指出, DCT, DWT 和金字塔变换的系数有着接近广义高斯分布的分布模型, DFT 的变换系数的幅值则接近于韦伯分布<sup>[8,9]</sup>. 因此在检测时, 关键是对这些拟定好的分布进行参数估计, 模型越接近真实分布可信

收稿日期: 2005-03-07; 修回日期: 2005-06-15

基金项目: 国家自然科学基金项目(60374066); 江苏省自然科学基金项目(BK2002101); 博士点基金项目(20020288052).

作者简介: 王金伟(1978—), 男, 内蒙古海拉尔人, 博士生, 从事多媒体数字水印及其模型的研究; 王执钊(1939—), 男, 武汉人, 教授, 博士生导师, 从事大系统的容错控制理论与应用、混沌控制理论等研究.

性就越高

## 2 基于图像块内容的分类

提高模型精确度的最直接办法是将数据集划分为几个子类,用几个不同的子模型分别描述各个子类的分布情况,例如根据频段将DFT系数划分为几个子类<sup>[9]</sup>来提高模型的精确度.将这种方法引入DCT域,即按照频段对数据集进行划分证明可以在一定程度上提高检测的性能,然而这种方法建立在不同频段系数集合的分布应该具有不同形态的假设之上,并没有确定的理论依据

本文基于DCT变换中不同纹理程度的图像块具有不同的分布参数的思想,将图像块按照平滑程度的不同进行分类.结合DCT变换的原理,从DCT变换的AC系数可以对图像块进行分类,因为具有较强纹理的图像具有较大的能量值,而较平滑的图像块其能量值较小

设图像块大小为  $m \times n$ , 定义该图像块的能量为

$$E = \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} f^2(i, j), \quad (1)$$

$i = 0, j = 0$

其中  $f(i, j)$  为图像  $(i, j)$  位置的DCT变换系数

图像的高频部分往往包含可能存在的高频噪声干扰而不能准确反映纹理的状况,因此在计算  $E$  时,选择如图 1 所示的中频AC系数进行计算

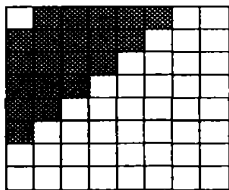


图 1 能量系数位置的选择

基于能量  $E$  的分类涉及阈值的选择,使用固定的阈值进行分类,可能会导致区分出现很大的偏差,而自适应的选择阈值会取得更好的结果,即根据不同的图像选择不同的能量值

设图像块根据DCT变换所计算的能量上限为  $E_{max}$ , 下限为 0; 方差的上限为  $\sigma_{max}^2$ , 下限为 0. 对于整幅图像而言,需要计算能量的均值  $\mu$ . 从平均值  $\mu$  的大小与能量上限  $E_{max}$  的比较初步判断这幅图像是平滑还是纹理,如果均值  $\mu$  大则说明这副图像是纹理丰富的,否则就是平滑的.然后,对这幅图像中所有图像块的能量值求方差  $\sigma^2$ , 并与方差的上限  $\sigma_{max}^2$  比较,这样做的目的是判断能量值的变化范围有多大.方差  $\sigma^2$  越大说明图像的纹理区域和平滑区域的差别越明显,而当图像的能量方差  $\sigma^2$  太小时,考虑到参数估计样本的个数,无论均值  $\mu$  的大小为多少,不

再考虑分类

阈值  $T_r$  根据每一幅图像均值  $\mu$  和方差  $\sigma^2$  的大小进行调节,使其在对图像分类时变换系数的概率分布能够达到更好效果.为此,阈值  $T_r$  分为以下 5 种情况:

$$T_r = \begin{cases} 0, \sigma^2 < \phi_{max}^2; \\ \mu, \mu < \gamma E_{max}, \phi_{max}^2 & \sigma^2 < \varrho_{max}^2; \\ \mu + \theta\sigma, \mu < \gamma E_{max}, \varrho_{max}^2 & \sigma^2 < \alpha_{max}^2; \\ \mu - \theta\sigma, \mu > \gamma E_{max}, \phi_{max}^2 & \sigma^2 < \varrho_{max}^2; \\ \mu, \mu > \gamma E_{max}, \varrho_{max}^2 & \sigma^2 < \alpha_{max}^2. \end{cases} \quad (2)$$

其中:  $\theta$  是调节因子,  $\phi, \varrho$  和  $\gamma$  是根据实验得到的经验值

## 3 模型拟合度的分析

为了表述简单,令不做任何划分的方法,即整体方法为方法 1; 令划分频段的方法为方法 2; 令本文的方法为方法 3

关系熵定义<sup>[10]</sup>如下:

$$R_E = \sum_{i=0}^{l-1} p_i(x) \log \frac{p_i(x)}{q_i(x)}. \quad (3)$$

其中  $p(x)$  和  $q(x)$  为两种概率分布.两种分布之间的关系熵通常情况下是非负的,而当两种概率分布完全相同,即  $p(x) = q(x)$  时,关系熵  $R_E$  为 0; 当两种概率分布差异大时,关系熵  $R_E$  较大; 当用  $p(x)$  表示原始的概率分布,  $q(x)$  表示拟合后的概率分布时,关系熵越小,说明  $q(x)$  拟合  $p(x)$  的程度越高,反之,则拟合程度越低

图 2 中给出了 3 种方法的原始概率分布和拟合后的曲线图

表 1 中列出了使用 3 种方法后所得关系熵.在实验中令阈值  $T_r$  简单地等于所有图像块的能量均值  $\mu$ . 由表 1 可以看出,方法 3 的参数估计最好,方法 2 次之,方法 1 较差.这说明划分能够有利于模型参数的精确估计,而基于内容的划分更有利于系数概率分布的拟合

## 4 乘性水印的最优检测

在乘性水印算法中,最优检测主要分为两类:一类是以固定强度嵌入的数字水印检测,且检测时需要知道嵌入强度的大小;另一类是考虑了人类视觉模型和攻击的变强度的数字水印检测,此方法不需要知道嵌入强度.这些结论都是在DCT域中的AC变换系数符合广义高斯分布的假设前提之下得到的.定义广义高斯函数为

$$f(x) = A \exp(-|\beta(x - \mu)|^p), \quad (4)$$

$$\beta = \frac{1}{\sigma} \left( \frac{\Gamma(3/p)}{\Gamma(1/p)} \right)^{1/2}, \quad (5)$$

表 1 3 种方法的关系熵

	方 法 1	方 法 2	方 法 3	方 法 3	方 法 3
Lena	0.117 61	0.087 169	0.049 104	0.030 412	0.069 65
Tiffany	0.162 72	0.147 07	0.127 25	0.054 201	0.000 696 33
Splash	0.116 85	0.119 68	0.020 798	0.075 198	0.003 668 8
Plane	0.087 239	0.071 278	0.027 348	0.025 16	0.044 654
Peppers	0.140 78	0.116 6	0.121 18	0.0425 08	0.044 523

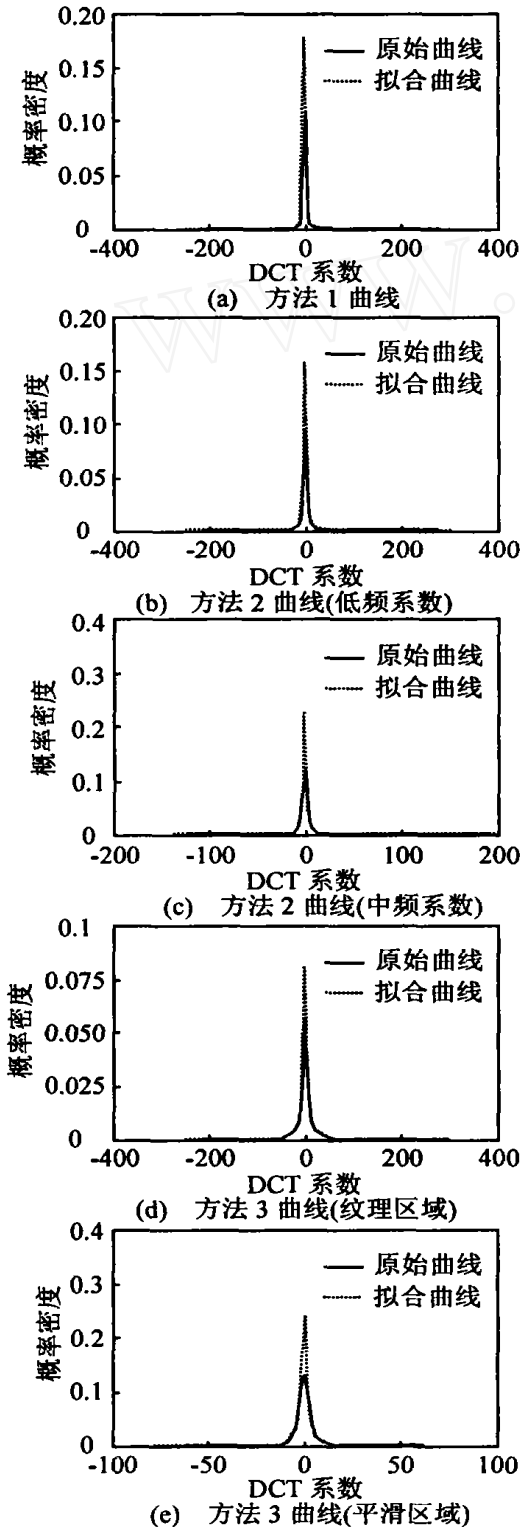


图 2 3 种方法的原始概率分布和拟合后曲线

$$A = \beta\rho/2\Gamma(1/\rho). \tag{6}$$

其中

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt,$$

$\mu$  和  $\sigma^2$  分别是广义高斯分布的期望和方差,  $\rho$  是函数的形状参数. 拉普拉斯分布和高斯分布分别是  $\rho=1$  和  $\rho=2$  的情况. 由于 DCT 变化的基函数正交, 因此其 AC 系数的均值近似为 0, 即  $\mu=0$ .

本文均假设水印序列  $W = \{w_i, i=1, 2, \dots, N\}^T$  的概率密度函数服从  $[-1, 1]$  的均值分布. 令  $X = \{x_i, i=1, 2, \dots, N\}$  表示未嵌入数字水印的 AC 系数,  $Y = \{y_i, i=1, 2, \dots, N\}$  表示嵌入数字水印的 AC 系数,  $\alpha = \{\alpha_i, i=1, 2, \dots, N\}$  表示水印的嵌入强度. 系数  $X$  和系数  $Y$  的关系由下式确定:

$$y_i = x_i(1 + \alpha w_i). \tag{7}$$

#### 4.1 固定强度的水印最优检测

在最大似然估计检测中, 根据 Bayes 假设检验理论, 给出以下两个假设条件: 1)  $y$  中未嵌入数字水印  $w^*$ ; 2)  $y$  中嵌入数字水印  $w^*$ .

如果接受假设 2) 成立, 那么要求对数极大似然比  $z(y)$  大于对数决策阈值  $\eta$ , 否则拒绝假设 2, 接受假设 1. 对数极大似然比  $z(y)$  和对数决策阈值  $\eta$  的表达式如下:

$$z(y) = \prod_{i=1}^N |\beta_i y_i|^{\rho_i} (1 - (1 + \alpha w_i^*)^{-\rho_i}), \tag{8}$$

$$\eta = \sqrt{2} \sigma \text{erfc}^{-1}(2\bar{P}_{fp}) + \mu_z, \tag{9}$$

$$\mu_z = \prod_{i=1}^N \frac{1}{\rho_i} (1 - (1 + \alpha w_i^*)^{-\rho_i}), \tag{10}$$

$$\sigma_z^2 = \prod_{i=1}^N \frac{1}{\rho_i} (1 - (1 + \alpha w_i^*)^{-\rho_i})^2. \tag{11}$$

其中

$$\text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} \exp(-t^2) dt,$$

$$\bar{P}_{fp} = \frac{1}{2} \text{erfc}\left(\frac{\eta - \mu_z}{\sqrt{2} \sigma_z}\right).$$

最优检测原理及公式推导的详细介绍参见文献[11, 12].

#### 4.2 可变强度的水印最优检测

在考虑攻击和人类视觉模型(HVS)后, 假设检

验给出如下的假设条件: 1)  $\alpha = 0$ ; 2)  $\alpha > 0$

如果能够检测出水印的存在, 即意味着水印嵌入强度不等于 0, 接受假设 2), 那么要求最优检测统计  $z(y)$  大于等于决策阈值  $\eta$  否则, 拒绝假设 2), 接受假设 1).

最优检测统计  $z(y)$  和决策阈值  $\eta$  分别为

$$z(y) = \frac{1}{N} \sum_{i=1}^N \rho_i |\beta_i y_i|^\rho w_i, \quad (12)$$

$$\eta = \sqrt{2} \sigma_e \text{erfc}^{-1}(2P_{fp}) + \mu_z, \quad (13)$$

$$\mu_z = \frac{1}{N} \sum_{i=1}^N w_i, \quad (14)$$

$$\sigma_e^2 = \frac{1}{N^2} \sum_{i=1}^N \rho w_i^2. \quad (15)$$

在文献[4~6]中给出了可变强度的水印最优检测的详细推导过程

### 5 实验分析

实验中载体对象是尺寸为 512 × 512 的图像, 如Lena, Plane, Tiffy, Peppers, Splash 等 同时对本文提出的方法, 即方法 3 与方法 1 和方法 2 进行比较, 并对图像进行 8 × 8 的分块DCT 变换

为便于说明 3 种方法的优劣性, 对检测值和阈值进行标准化处理 标准化后的示意图如图 3 所示, 图中  $z_{n1}$  和  $z_{n2}$  分别是标准化后的两种方法的检测值 ( $V_D$ ),  $\eta_t$  是标准化后的阈值 ( $T_D$ ). 标准化后两种方法的性能比较显而易见 如图中所标识的  $z_{n1}$  和  $z_{n2}$ , 由于  $z_{n1} > z_{n2}$ , 可以判断  $z_{n1}$  的方法优于  $z_{n2}$  的方法 为了评价两种检测方法的性能, 给出平均距离为

$$\bar{D} = V_D - T_D, \quad (16)$$

当  $\bar{D} > 0$  时接受假设 2), 且其值越大检测成功率越高; 当  $\bar{D} < 0$  时接受假设 1).

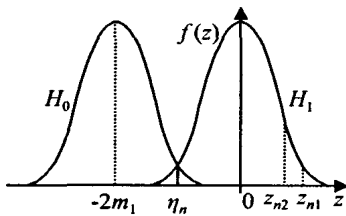


图 3 标准化示意图

#### 5.1 固定强度的水印检测

若以固定强度嵌入水印, 则  $\alpha = \alpha_i, i = 1, 2, \dots, N$ . 在实验中选取  $\alpha = 0.15$ , 3 种方法的实验数据如表 2 所示

计算上述 3 种方法的检测均值的平均距离: 方法 1 的平均距离  $\bar{D}_1 = 8.15 - 4.75 = 3.40$ , 方法 2 的平均距离  $\bar{D}_2 = 9.09 - 4.75 = 4.34$ , 方法 3 的平均距离  $\bar{D}_3 = 11.59 - 4.75 = 6.84$  可以判断出方法

3 明显好于方法 1 和方法 2, 而方法 2 又好于方法 1, 这与第 3 节中通过关系熵得到的结论一致

表 2 3 种方法固定强度的水印检测实验数据

	方法 1	方法 2	方法 3
Lena	8.87	10.38	11.16
Peppers	9.16	10.27	12.97
Tiffy	8.40	8.77	11.90
Plane	9.82	10.86	11.83
Splash	4.49	5.15	10.08
均值	8.15	9.09	11.59

#### 5.2 可变强度的水印检测

对于可变强度的水印检测, 由于需要考虑 HVS, 这时水印的嵌入强度依赖于图像载体 在这里 HVS 选取文献[7]中所提出的模型, 实验数据见表 3.

表 3 3 种方法标准化后的实验数据

	方法 1	方法 2	方法 3
Lena	20.94	22.87	24.36
Peppers	32.63	37.58	38.03
Tiffy	21.76	22.96	26.57
Plane	18.23	20.38	20.78
Splash	19.30	20.67	25.89
均值	22.57	24.89	27.13

上述 3 种方法检测均值的平均距离计算结果如下: 方法 1 的平均距离  $\bar{D}_1 = 22.57 - 4.75 = 17.82$ , 方法 2 的平均距离  $\bar{D}_2 = 24.89 - 4.75 = 20.14$ , 方法 3 的平均距离  $\bar{D}_3 = 27.13 - 4.75 = 22.38$

可以判断出方法 3 明显好于方法 1 和方法 2, 而方法 2 好于方法 1. 平均距离显示的结果同第 3 节通过关系熵得到的结论一致

### 6 结 论

本文给出了一种 DCT 域内基于内容分类的最优检测算法 本算法主要针对提高变换系数分布模型的参数估计而提出, 因为参数估计的越准确, 分布模型越接近真实情况, 相应的检测结果也会越精确 通过本文提出的算法同以前的算法在实验中的比较表明, 本算法确实提高了最优检测的结果

本文在分类上仅采用了最简单的方法, 即能量, 从分类的精度而言 DCT 域能量分类是比较粗糙的, 但是计算复杂度较低思想简单容易理解, 因此, 若要继续提高最优检测的结果, 可以考虑使用更好的分类方法划分平滑区域和纹理区域, 或者把图像划分为几个区域而不仅仅是两个区域

#### 参考文献 (References)

[1] Langelaam G, Setyawan I, Legendijk R. Water-marking Digital Image and Video Data [J]. IEEE Signal Processing Magazine, 2000, 17(5): 20-46  
[2] Cox IJ, Kilian J, Keighton T, et al. Secure Spread

- Spectrum Watermarking for Multimedia [J]. *IEEE Trans on Image Processing*, 1997, 6(12): 1673-1687.
- [3] Kunder D, Hatzinakos D. A Robust Digital Image Watermarking Method Using Wavelet-based Fusion [A]. *Proc IEEE Int Conf on Image Processing* [C]. California, 1997: 544-547.
- [4] Cheng Q, Huang T S. Robust Optimum Detection of Transform Domain Multiplicative Watermarks [J]. *IEEE Trans on Signal Processing*, 2003, 51 (4): 906-924
- [5] Cheng Q, Huang T S. Optimum Detection of Multiplicative Watermarks Using Locally Optimum Decision Rule [A]. *IEEE Int Conf on Multimedia and Expo* [C]. Tokyo, 2001: 309-312
- [6] Hernández J R, Amado M, Pérez-González F. DCT-domain Watermarking Techniques for Still Images: Detector Performance Analysis and a New Structure [J]. *IEEE Trans on Image Processing*, 2000, 9(1): 55-68
- [7] Buccigrossi R W, Simoncelli E P. Image Compression Via Joint Statistical Characterization in the Wavelet Domain [J]. *IEEE Trans on Image Proceeding*, 1999, 8(12): 1688-1701.
- [8] Cheng Q, Huang T S. Optimum Detection and Decoding of Multiplicative Watermarks in DFT Domain [A]. *IEEE Int Conf on Acoustics, Speech, and Signal Processing* [C]. Orlando, 2002: 3477-3480
- [9] Barni M, Bartolini F, De Rosa A, et al. A New Decoder for the Optimum Recovery of Nonadditive Watermarks [J]. *IEEE Trans on Image Processing*, 2001, 10 (5): 755-766
- [10] Cachin C. An Information-theoretic Model for Steganography [A]. *Proc of 2nd Workshop on Information Hiding* [C]. Oregon: Springer-Verlag, 1998: 306-318
- [11] Ming N T, Garg H K. Wavelet Domain Watermarking Using Maximum-likelihood Detection [A]. *Proc of SPIE-IS and T Electronic Imaging on Security, Steganography, and Watermarking of Multimedia Contents VI* [C]. California: San-Jose, 2004: 816-826
- [12] Kwon S G, Lee S H, Kwon K K, et al. Watermark Detection Algorithm Using Statistical Decision Theory [A]. *IEEE Int Conf on Multimedia and Expo* [C]. Lausanne, 2002: 561-564

(上接第 370 页)

## 参考文献(References)

- [1] Tao G, Kokotovic P V. Adaptive Sliding Control of Plants with Unknown Dead-zone [J]. *IEEE Trans on Automatic Control*, 1994, 39(1): 59-68
- [2] Tao G, Kokotovic P V. Discrete-time Adaptive Control of Systems with Unknown Dead-zone [J]. *Int J of Control*, 1995, 61(1): 1-17.
- [3] Cho H Y, Bai E W. Convergence Results for an Adaptive Dead Zone Inverse [J]. *Int J of Adaptive Control and Signal Processing*, 1998, 12(3): 451-466
- [4] Wang X S, Hong H, Su C Y. Model Reference Adaptive Control of Continuous Time Systems with an Unknown Dead-zone [J]. *IEE Proc - Control Theory Application*, 2003, 150(3): 261-266
- [5] Wang X S, Hong H, Su C Y. Robust Adaptive Control a Class of Nonlinear Systems with an Unknown Dead-zone [J]. *Automatica*, 2004, 40(3): 407-413
- [6] Zhang T, Ge S S, Hang C C. Stable Adaptive Control for a Class of Nonlinear Systems Using a Modified Lyapunov Function [J]. *IEEE Trans on Automatic Control*, 2000, 45(1): 129-132
- [7] 张天平. 基于一种修改的李亚普诺夫函数的自适应模糊滑模控制 [J]. *自动化学报*, 2002, 28(1): 137-142  
(Zhang T P. Adaptive Sliding Mode Control Based on a Modified Lyapunov Function [J]. *Acta Automatica Sinica*, 2002, 28(1): 137-142 )
- [8] Zhang T P. Stable Direct Adaptive Fuzzy Control for a Class of MIMO Non-linear Systems [J]. *Int J of Systems Science*, 2003, 34(6): 375-388
- [9] Nussbaum R D. Some Remarks on the Conjecture in Parameter Adaptive Control [J]. *System Control Letters*, 1983, 1(3): 243-246
- [10] Ye X D, Jiang J P. Adaptive Nonlinear Design without a Priori Knowledge of Control Directions [J]. *IEEE Trans on Automatic Control*, 1998, 43 (11): 1617-1621.
- [11] Ge S S, Hong F, Lee T H. Adaptive Neural Control of Nonlinear Time-delay Systems with Unknown Virtual Control Coefficients [J]. *IEEE Trans on Systems, Man and Cybernetics - Part B*, 2004, 34(1): 499-516