

一种高效的数字笔迹多维数据编码算法*

李俊峰⁺, 戴国忠

(中国科学院 软件研究所 人机交互技术与智能信息处理实验室,北京 100080)

An Efficient Digital Ink Multi-Dimension Data Coding Algorithm

LI Jun-Feng⁺, DAI Guo-Zhong

(Human Computer Interaction and Intelligent Information Processing Laboratory, Institute of Software, The Chinese Academy of Sciences, Beijing 100080, China)

+ Corresponding author: Phn: +86-10-62540434, E-mail: solemnizeljf@hotmail.com, <http://www.iscas.ac.cn>

Li JF, Dai GZ. An efficient digital ink multi-dimension data coding algorithm. *Journal of Software*, 2006, 17(9):1860-1866. <http://www.jos.org.cn/1000-9825/17/1860.htm>

Abstract: An efficient digital ink data coding algorithm IWPSP (integer wavelet packet based hierarchical set partitioned) is proposed in this paper. The algorithm compresses digital ink multi-dimension data losslessly using three approaches: integer wavelet packet transform, hierarchical set partitioned, significant bits combination code and fast adaptive arithmetic code. The experiments show that the IWPSP algorithm is efficient.

Key words: digital ink multi-dimension data coding; integer wavelet packet transform; hierarchical set partitioned; fast adaptive arithmetic code

摘要: 提出了一种高效的数字笔迹数据编码算法 IWPSP(integer wavelet packet based hierarchical set partitioned).该算法通过引入整数小波包变换、层次性集合分裂、重要位组合编码和快速自适应算术编码等方法,无损地压缩了数字笔迹多维数据.实验证明,提出的 IWPSP 算法是高效的.

关键词: 数字笔迹多维数据编码;整数小波包变换;层次性集合分裂;快速自适应算术编码

中图法分类号: TP391 文献标识码: A

笔式交互由于其基于纸笔隐喻、接近人们日常生活、自然高效的交互方式,已成为人机交互领域的一个热点.在笔式交互过程中,用户通过数字笔输入数字笔迹,完成交互任务.从数据角度看,数字笔迹由若干笔划组成,每个笔划是数字笔在一次相继地按下和抬起之间所经过的笔迹;笔划由若干采样点组成,采样点一般有 X 坐标值、Y 坐标值、压力值、时间值等属性.如果用整数表示各个属性,则需要 16 个字节描述一个采样点.由于数字笔的采样频率达到 100Hz~133Hz,用户如果连续书写 1 分钟,数字笔迹的存储空间就需要 1.6KB.庞大的数据量对计算机容量和网络传输技术都提出了很高的要求,这成为笔式交互深入发展的瓶颈.因此,数字笔迹数据压缩成为笔式交互技术中的重要课题.

* Supported by the National Natural Science Foundation of China under Grant Nos.60273024, 60033020 (国家自然科学基金); the National Grand Fundamental Research 973 Program of China under Grant No.2002CB312103 (国家重点基础研究发展规划(973)); the Key Innovation Project from Institute of Software, the Chinese Academy of Sciences (中国科学院软件研究所创新基金重大项目)

Received 2005-07-06; Accepted 2005-10-11

为了方便数字笔迹的传输、处理、表示,W3C(world wide web consortium)制定了 InkML^[1]标准,使数字笔迹可以在不同的应用之间实现互换.InkML 规范了数字笔迹存储格式,可用作手写识别、签名验证等功能.InkML 支持完整的手写笔迹表示方式,除了数字笔的位置,也可以记录笔的倾角、压力等.但在 InkML 中,数字笔迹数据压缩并没有被深入研究.

胡云飞^[2]设计了一种基于曲线拟合进行数字笔迹存储的方法.该方法采用三阶 B 样条曲线作为拟合使用的曲线,综合了等距加权采样算法以及折点识别算法等算法的优点.对数字笔迹进行了压缩,平均压缩比为 40%.微软公司采用两种方式对数字笔迹进行压缩^[3]:一种是直接压缩;另一种是转换为图像格式(GIF)存储.但是,这些压缩方法都属于有损压缩,而且仅仅适用于采样点属性只有 X 坐标值和 Y 坐标值的情况.

为了无损地压缩数字笔迹的多维数据,本文提出一种编码算法 IWPSP(integer wavelet packet based hierarchical set partitioned).该算法利用可逆整数小波包变换去除数字笔迹数据在时间维度上的强相关性,然后从变换系数自身特点出发,通过层次性集合分裂的编码方法,对变换系数进行嵌入式编码.同时,部分编码结果用快速自适应算术编码再次压缩.实验结果证明,算法能够无损地、高效率地压缩数字笔迹的多维数据.

1 数字笔迹多维数据编码算法

数字笔迹多维数据 I 在数据角度上看由笔划组成, $I=\{S_i|i \in [1,n],n$ 为笔划数}.笔划 S_i 由采样点 L_j^i 组成, $S_i = \{L_j^i(x_j^i, y_j^i, p_j^i, t_j^i) | j \in [1, k_i], k_i$ 为采样点数}.由于数字笔的采样频率一般达到 100Hz~133Hz,所以数字笔迹多维数据在时间维度上具有强相关性.图 1 给出了数字笔迹截图及相应的多维数据.

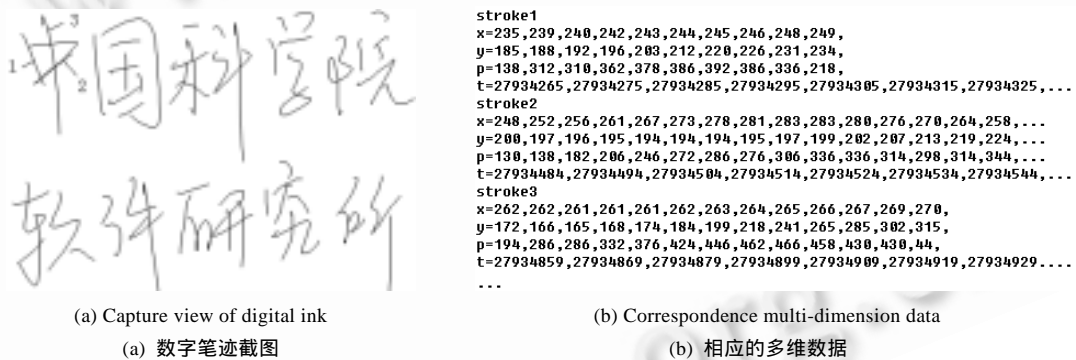


Fig.1 The capture view and multi-dimension data of the digital ink

图 1 数字笔迹截图及相应的多维数据

从图 1 可以看到:同一笔划内的属性具有相关性,而且不同笔划间的相同属性值也具有相关性,特别是时间属性.为了最大限度地去除数字笔迹属性值在时间维度上的相关性,编码算法将多个笔划 S_i 合为一个笔划 S , $S = \{L_j(x_j, y_j, p_j, t_j) | j \in [1, k], k = \sum_{i=1}^n k_i\}$.为了解码需要,将原笔划包含的点数进行记录, $K=\{k_i|i \in [1,n]\}$.由于采样点不同属性之间不存在相关性,所以编码算法会对 K, X, Y, P, T 分别进行压缩,其中 X, Y, P, T 分别是数字笔迹 x 坐标、 y 坐标、压力 p 、时间 t 属性值的集合.为了方便起见,仅对编码算法压缩 X 集合进行叙述, $X=\{x_j | j \in [1, k]\}$,其余集合的压缩类似于 X 集合的压缩.

IWPSP 数字笔迹多维数据编码算法过程如图 2 所示.

由于整数小波包变换具有优异的时频局部能力以及良好的去相关能力,编码算法首先利用整数小波包变换得到变换系数.然后从变换系数自身特点出发,通过层次性集合分裂的编码方法对变换系数进行嵌入式编码.同时,部分编码结果用快速自适应算术编码再次压缩.

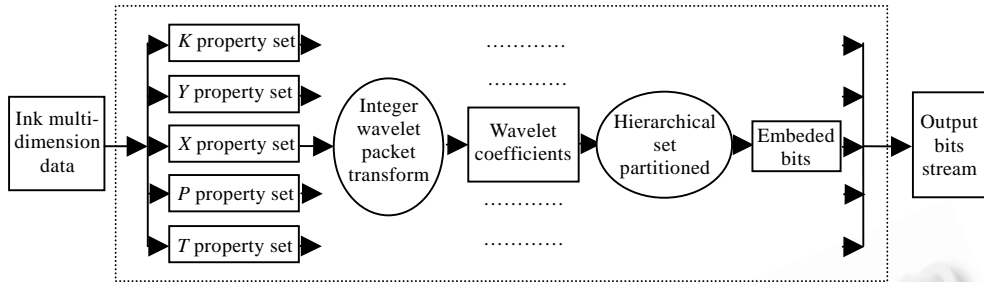


Fig.2 IWPSP digital ink multi-dimension data code diagram

图2 IWPSP 数字笔迹多维数据编码算法流程图

1.1 整数小波包变换(integer wavelet packet transform)

自 20 世纪 80 年代末期 Mallat^[4]首次将快速小波变换引入数据处理以来,小波变换得到了广泛的应用,并取得了良好的效果.但在实际应用中,快速小波变换有其自身的弱点:(1) 需要作卷积运算,计算复杂;(2) 对数据的尺寸有要求.为此,Swelden^[5]提出了基于提升模式的小波变换,它使人们用一种简单的方法来解释小波的基本理论,并且在提升模式下建立的整数小波变换具有变换可逆、运算简单、整数存储等优点,在图像压缩领域^[6]得以深入的研究.

基于提升模式的整数小波变换分为分裂(split)、预测(predict)和更新(update)3 个步骤.

分裂:将数据 X 分裂成两个互不相交的子集,通常是奇偶分裂,即 $E^0(n)=X(2n); O^0(n)=X(2n+1)$.

预测:采用预测算子 P 作用于偶信号,得到奇信号的预测值 E_p^0 ,再将 O^0 与预测值相减得到预测误差 O^1 ,即

$$O^1(n) = O^0(n) - \text{int}(E_p^0(n)).$$

更新:采用更新算子 U 作用于预测误差 O^1 ,得到偶信号的更新值 O_U^1 ,再将 E^0 与更新值相减得到 E^1 ,即

$$E^1(n) = E^0(n) - \text{int}(O_U^1(n)).$$

其中, $\text{int}(\cdot)$ 为取整函数.

经过上述 3 个步骤后,数据实际上被分解为两个频率部分:偶信号对应低频;奇信号对应高频.然而,在小波变换中,分解层数是一个模糊的数量,由编码算法任意规定.更重要的是,进一步的分解只作用于低频部分,高频部分保持不动,这使得小波变换在处理高频部分复杂的数据时出现效果不理想等情况.

小波包变换^[7]是小波变换的推广.与小波变换不同的是,数据的高频部分也和低频部分一样被递归分解,所以它能够提供更灵活的、更适合于数据时频域特性的分解方式.为了选择合适的小波包基,人们开展了大量的研究工作,大体上可分为全局搜索和局部搜索两种.其中,全局搜索算法^[8]必须等小波包完全分解后才能进行搜索,这无疑会大大增加计算量,影响数据的编码速度;而在局部搜索算法^[9]中,通过定义费用指标,一边分解,一边判别父节点费用与子节点费用和的关系,对不满足条件的节点将停止分解,这使得编码速度有了大幅度的提高.

本文给出一种新的整数小波包变换算法, LSS 为待分解集合列表, LDS 为已分解集合列表,定义费用指标为

$$F(Z) = \sum_{x_j \in Z} |x_j|, Z \text{ 为 } X \text{ 的子集.}$$

(1) 初始化 LSS 为 X 集合, LDS 为空集

(2) 处理 LSS 中的每一个集合 Z ,直到 LSS 为空集

- 计算集合 Z 的费用 $C = F(Z) = \sum_{x_j \in Z} |x_j|$
- 利用整数小波变换,将 Z 分解为两个子集合 Z_e, Z_o
- 计算子集合费用和 $C' = F(Z_e) + F(Z_o)$
- 如果 $C' \leq C$,将 Z_e, Z_o 加入到 LSS 中,从 LSS 中去除 Z

- 否则,将 Z 从 LSS 移至 LDS

显然,整数小波包变换在具备整数小波变换快捷、能够在当前位置上完成计算、能够对任意尺寸数据进行变换等优点的同时,由于采用了变换系数幅值和的费用指标,使得变换结果的幅值和大大幅度降低,为随后的嵌入式编码打下了良好的基础。

1.2 层次性集合分裂编码(hierarchical set partitioned coding)

为了更好地对整数小波包变换后的变换系数进行编码,IWPHSP 借鉴了图像编码中嵌入式编码的思想,尤其是 Asad 和 Pearlman 提出的 SPECK^[10]算法,该算法具有完全嵌入性、渐进传输性、低计算复杂性、较高的信噪比和较好的图像复原质量等优点。然而,理论分析和实验结果表明,该算法也存在一些不足,具体表现为:当某一集合 Z 在当前量化步长下是非重要的,它会被留在 LIS 中,随着量化步长的更新, Z 会不断地被测试,直到在某一量化步长下成为重要集合而被进一步分解为止。在每一次测试中,由于 Z 是非重要集合,SPECK 编码器会输出“0”作为解码标志。由于 IWPHSP 将会对数字笔迹多维数据进行无损压缩,如果直接应用 SPECK 算法的基本步骤,那么上述情况会大量出现,必然导致编码器的低效。而且在整数小波包变换后, LDS 中的集合分布于各级分解层,已经不适合采用八带分裂的方式,因此,本文给出层次性集合分裂编码方法 HSP,通过改变 SPECK 编码算法的基本步骤,使得编码方法适合于整数小波包变换的结果,同时改进了整体编码算法的性能。

HSP 编码算法的核心思想是:用“0,1,2,4,...,2ⁿ”取代“1,2,4,...,2ⁿ”作为编码过程中的量化阈值, n 是初始化过程中的最大位面。对应于每个量化阈值,编码算法建立待分解集合列表 LSS ,集合的重要性测试、分裂等操作只当前 LSS 进行,最大限度地适应于小波系数的幅值特点,提高了编码性能。

集合 X 进行整数小波包变换后,已分解集合存在于 LDS 中, $LDS=\{Z_i|i=1,2,\dots,l\}$ 。设 LBP 为待定位面集合列表,定义集合 Z 的重要性判定函数 $S_n(Z)$ 、位面计算函数 $P(Z)$ 、尺寸函数 $C(Z)$ 分别为

$$S_n(Z) = \begin{cases} 1, & 2^n \leq \max_{x_i \in Z} \{|x_i|\} < 2^{n+1} \\ 0, & \text{else} \end{cases},$$

$$P(Z) = \begin{cases} \left\lceil \log_2 \left(\max_{x_i \in Z} |x_i| \right) \right\rceil + 1, & 0 < \max_{x_i \in Z} |x_i| \\ 0, & 0 = \max_{x_i \in Z} |x_i| \end{cases},$$

$$C(Z) = \text{size}(Z) = \text{集合 } Z \text{ 包含的变换系数的数目。}$$

具体算法步骤如下:

(1) 初始化

- 计算 LDS 中每一集合的位面 $n_i=P(Z_i)$,将 Z_i 加入到 LBP 中
- 输出 $n = \max_{1 \leq i \leq l} (n_i)$
- 对应于每个量化阈值,建立待分解集合列表 LSS_j ,其中 $0 \leq j \leq n$
- 如果 $0 < n$
 - 执行 $ProcessLBP()$
 - 清空 LBP

(2) 排序与细化

- 以集合尺寸递增的顺序处理当前 LSS 中的每一个集合 Z
 - 如果 $C(Z)=1, ProcessOne(Z)$
 - 否则, $ProcessS(Z)$
- 如果 $1 < n$
 - 执行 $ProcessLBP()$
 - 清空 LBP

(3) 量化步长更新

- 如果 $n=0$,退出
- 否则,将 $n-1$,转移至(2)

HSP 算法中用到的方法有 $ProcessS(),ProcessOne(),ProcessLBP()$:

$ProcessS(Z)$

{

- 利用二叉树分裂方法,将集合 Z 分裂为两个子集 Z_1, Z_2
- 组合输出 $S_n(Z_1), S_n(Z_2)$
- 处理子集 $Z_m(m=1,2)$
 - 如果 $S_n(Z_m)=1$
 - 如果 $C(Z_m)=1, ProcessOne(Z_m)$
 - 否则, $ProcessS(Z_m)$
 - 否则
 - 如果 $1 < n$, 将 Z_m 加入到 LBP 中

}

$ProcessOne(Z)$

{

- 如果 Z 被首次处理,输出 Z 包含的变换系数 x_i 的符号
- 输出 $|x_i|$ 的第 n 个最重要位,更新 x_i
- 如果 $1 < n$, 将 Z 加入到 LBP 中

}

$ProcessLBP()$

{

- 处理 LBP 中的每一个集合 Z , 计算 Z 的位面 $P(Z)$
- 利用二进制自适应算术编码输出 LBP 中集合的位面
- 处理 LBP 中的每一个集合 Z , 将 Z 加入到 $LSS_{P(Z)}$ 中

}

在 HSP 编码算法中:

(1) 重要位组合编码

将集合 Z 分裂为两个子集 Z_m 后,可以利用上下文组合输出 $S_n(Z_m)$. 因为如果 Z 集合分裂,那么该 Z 集合必然是重要的. 所以,在新生成的两个子集 Z_m 中,必然有一个是重要的. 换句话说,组合编码中不会出现“00”的情况. 因此,如果第 1 个子集已经输出“0”,则最后一个“1”无须输出.

(2) 二进制自适应算术编码

处理 LBP 中的集合 Z , 计算 Z 的位面 $P(Z)$, 然后利用二进制自适应算术编码输出 LBP 中集合的位面. 相对于直接输出 $P(Z)$, 二进制自适应算术编码无疑是较慢的. 但通过分析可知 $0 \leq P(Z) \leq n$, 而数据经过整数小波包变换后,一般满足 $n \leq 20$. 在这种情况下,利用快速算术编码^[11]可以得到比较满意的编码效率. 基于算法复杂度、编码效率等因素的考虑,算法没有对符号位、重要位等类型的输出位进行算术编码.

2 实验结果与分析

实验在整数小波变换中采用 $5-3^{[12]}$ 提升模式,边缘采用镜面对称处理. 表 1 给出 IWPSHP 算法对不同数字笔迹压缩的字节率以及编解码时间,表 2 给出不同属性值在压缩流中所占的比例.

实验结果表明: IWPSHP 能够无损地、高效地压缩数字笔迹多维数据,其中对时间属性的压缩最为有效,对压力属性的压缩效果比较差. 这是由于时间属性相比压力属性而言更为连续,这使得整数小波包变换对时间属

性的相关性去除得更加有效,在相同的采样点数条件下,字节率与数据的相关性成反比,即数据的相关性越大,字节率越低.数据的相关性对应于数字笔迹书写的密集程度,如实验中的第2组和第3组(如图3所示),由于第3组的密集性较大,则字节率较低.

而且,IWPSHP的编解码所需的时间很少,这使得IWPSHP适合在实时性系统中应用.如果数字笔迹含有更多的属性,如数字笔与书写平面所成的角度,IWPSHP同样可以进行高效的压缩.

Table 1 The compression byte rate and code/decode time on different digital ink data

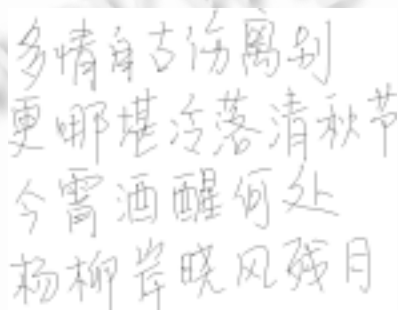
表 1 不同数字笔迹压缩的字节率以及编解码时间

Number of sample points	770	2 603	2 739	3 839	7 076	9 801
Coding data (Byte)	2 107	6 739	6 296	6 595	16 672	23 525
Byte per sample point	2.736	2.589	2.30	1.72	2.35	2.40
Time of coding (s)	0.047	0.219	0.219	0.250	0.578	0.843
Time of decoding (s)	0.047	0.204	0.188	0.219	0.484	0.750

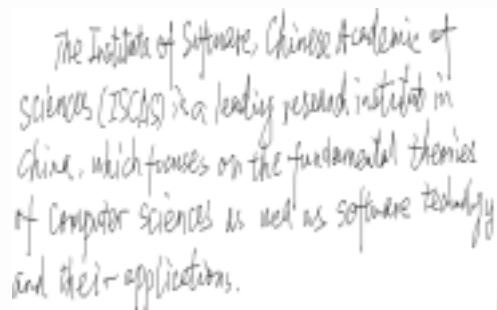
Table 2 The compress proportion of sample points attributes

表 2 对采样点属性的压缩比例

Number of sample points	0.022 801	0.014 493	0.022 809	0.023 745	0.022 92	0.019 329
Coordinate property	0.421 612	0.406 602	0.414 646	0.403 216	0.423 905	0.409 741
Pressure property	0.378 842	0.406 968	0.350 06	0.372 858	0.373 353	0.383 43
Time property	0.176 746	0.171 937	0.212 485	0.200 18	0.179 822	0.187 5



(a) The second digital ink data
(a) 第2组数字笔迹



(b) The third digital ink data
(b) 第3组数字笔迹

Fig.3 The digital ink data with different density

图 3 不同密集程度的数字笔迹

3 结 论

本文提出一种数字笔迹多维数据编码算法 IWPHSP(integer wavelet packet based hierarchical set partitioned),其优越性主要体现在以下几个方面:

- (1) 全面引入整数小波包变换,通过定义新的费用指标局部搜索最佳基.编码算法在具备整数小波变换优点的同时,使得变换结果的幅值和大幅度降低,为随后的嵌入式编码打下了良好的基础;
- (2) 充分分析小波系数幅值特点,采用层次性集合分裂方法,使得编码性能有了较大的提高;
- (3) IWPHSP 利用组合编码和快速自适应算术编码,改进了比特分配策略;
- (4) IWPHSP 在提高编码性能的同时,改进集合分裂、测试重要性程序,降低了算法复杂度.

IWPHSP 算法在整数小波变换中单独采用 5-3 提升模式.但是,通过实验可知:在同一提升模式下,不同的属性值具有不同的压缩效果.所以,我们希望通过实验选用更加适合数字笔迹多维数据的提升模式,甚至不同属性值选用不同的提升模式,使压缩效果更佳.另一方面,IWPHSP 算法属于无损压缩算法,这使得 IWPHSP 算法不能应用于有损压缩的情况.我们希望能对 IWPHSP 算法进行改进,使之能够满足有损、无损压缩等多种需求.

本文的算法充分结合了整数小波包变换与嵌入式集合分裂编码的优点,对未来数字笔迹多维数据编码算

法有积极的意义.

References:

- [1] Chee YM, Magana JA, Franke K, Froumentin M, Russell G, Madhvanath S, Seni G, Tremblay C, Yaeger L. Ink markup language W3C working draft. 2004. <http://www.w3.org/TR/2004/WD-InkML-20040928/>
- [2] Hu YF, Qin YY, Dai GZ. Method of stroke storing and rendering based on curve-fitting. Journal of Computer Engineering and Science, 2003, 25(5):42-45 (in Chinese with English abstract).
- [3] Microsoft Corporation. Save (PresistenceFormat, CompressionMode). 2004. <http://msdn.microsoft.com/library/default.asp?url=/library/en-us/tpcsdk10/lonestar/microsoft.ink/Classes/ink/Methods/save0.asp>
- [4] Mallat S. A theory for multiresolution signal decomposition: The wavelet representation. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1989,11(7):674-693.
- [5] Sweldens W. The lifting scheme: A construction of second generation wavelets. SIAM Journal of Mathematical Analysis, 1998, 29(2):511-546.
- [6] Zhang LB, Wang K. An image compression algorithm based on integer wavelet transform. Journal of Software, 2003,14(8): 1433-1438 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/14/1433.htm>
- [7] Colifman R, Wickauser M. Entropy-Based algorithm for best basis selection. IEEE Trans. on Information Theory, 1992,38(2): 909-996.
- [8] Xiong Z, Ramchandran K, Orchard MT. Wavelet packet image coding using space-frequency quantization. IEEE Trans. on Image Processing, 1998,7(6):160-174.
- [9] Chen YY, Zhang B. Composite wavelet packet and best basis. Journal of Software, 1998,9(3):161-168 (in Chinese with English abstract).
- [10] Islam A, Pearlman WA. An embedded and efficient low-complexity hierarchical image coder. In: Proc. of the Int'l Society fore Optical Engineering Conf. on. Visual Communications and Image Processing. 1999. 294-305.
- [11] Bodden E, Clasen M, Kneis J. Arithmetic coding revealed. In: Proc. of the Proseminar Datenkompression 2001, Chair for Computer Science VI. University of Technology Aachen/RWTH. 2003. 100-157.
- [12] Adams MD, Kossentini F. Reversible integer-to-integer wavelet transforms for image compression: Performance evaluation and analysis. IEEE Trans. on Image Processing, 2000,9(6):1010-1024.

附中文参考文献:

- [2] 胡云飞,秦严严,戴国忠.基于曲线拟合的笔迹存储和绘制方法的研究.计算机工程与科学,2003,25(5):42-45.
- [6] 张立保,王珂.一种基于整数小波变换的图像编码算法.软件学报,2003,14(8):1433-1438. <http://www.jos.org.cn/1000-9825/14/1433.htm>
- [9] 陈玉宇,张钊.混合小波包与最佳基.软件学报,1998,9(3):161-168.



李俊峰(1976 -),男,辽宁盖州人,博士,主要研究领域为人机交互技术,模式识别.



戴国忠(1944 -),男,研究员,博士生导师,CCF高级会员,主要研究领域为人机交互技术,实时智能,软件工程.