

研究简报

MPEG-2视频编码的运动估计与实现¹⁾

陈维强 高文

(哈尔滨工业大学计算机科学与工程系 哈尔滨 150001)

关键词 运动图象压缩, MPEG-2, 运动估计.

1 引言

MPEG (Moving Picture Expert Group) 专家组于1994年11月新加坡会议上提出了 MPEG-2标准的主体部分(系统、视频^[1]和音频). 可以说 MPEG-2标准将成为(或正在成为)包括从宽带通信、广播电视到计算机多媒体与娱乐等应用领域的一项共性技术, 其应用范围是非常广泛. 然而, 与 MPEG-1相比, MPEG-2的编码器实时实现更为复杂、难度更大, 且其中表现最为突出的部分之一就是 MPEG-2视频编码的运动估计.

只就 MPEG-2 MP@ML 广播级而言, 即支持 CCIR-601格式(720pixels × 480lines × 30frames/sec 或相当), 其采样密度相当于是 MPEG-1所支持的 SIF 格式的四倍, 并且 MPEG-2还同时允许支持多种运动估计类型. 例如在编码时针对帧结构图象序列: 允许以整个宏块(16 × 16像素)进行运动估计(与 MPEG-1相同); 同时允许将宏块按顶、底场分成两个 16 × 8块进行运动估计; 而且当编码序列中没有定义 B 类型图象时, 还允许做双基(dual prime)运动估计.

由此可见, MPEG-2视频编码运动估计的实时实现是具有挑战性. 目前这方面的报道还很少^[2]. 鉴于 MPEG-2视频编码标准是一个新标准, 其运动估计与 MPEG-1相比有较大改进, 并且在 MPEG-2视频编码标准中只是定义了运动估计的类型, 而对具体如何实现 MPEG-2视频编码运动估计并没有作定义. 针对这些问题, 本文阐述了 MPEG-2视频编码标准所支持的运动估计, 并给出了其软件实现, 分析了基于全搜索算法的 MP@ML 广播级 MPEG-2视频编码运动估计的计算量, 最后通过实验得出了一些简化实现 MPEG-2视频编码运动估计的对比实验结果. 本文对 MPEG-2视频编码运动估计的 VLSI 实现研究具有参考价值, 而且还可作为 MPEG-2视频编码标准关于运动估计的一个补充.

1) 国家863计划智能计算机系统主题和国家“九五”攻关计划的支持项目.

2 MPEG-2视频编码运动估计与软件实现

2.1 相关概念

MPEG-2编码图象通常被分成三种类型:即 I 图象、P 图象、B 图象,其中 I 图象一般只做帧内编码,P 图象做前向预测编码,B 图象做双向预测编码. 从一个 I 图象起始到下一个 I 图象之前称为一个图象组(GOP),GOP 包含的图象帧数记为 N,相邻参考帧之间的距离记为 M.

MPEG-2视频编码标准支持三种输入图象序列编码格式:①输入图象按场格式编码,记为场编码方式;②输入图象按帧/场格式编码,即在编码时可进行自适应帧、场方式运动补偿与 DCT 变换,记为帧/场编码方式;③输入图象按帧格式编码,记为帧编码方式. 在场编码方式中,允许进行 16×16 (场)运动估计、 16×8 上、下运动估计和双基运动估计;在帧/场编码方式中,允许进行 16×16 (帧)运动估计、 16×8 顶、底运动估计和双基运动估计;在帧编码方式中,只允许进行 16×16 (帧)运动估计. 其中双基运动估计只在编码图象序列中没有定义 B 类型图象时(即 $M=1$)采用.

2.2 运动补偿类型判决方法

针对多种运动估计结果,本文给出一种运动补偿类型两步判决方法. 设待估计图像的数据块(Y 分量)用数组 $data[k][l]$ 表示,参考块用 $data'[k][l]$ 表示,其中 $k \in [0, K-1]$, $l \in [0, L-1]$,在 MPEG-2 视频编码标准中 $K=16$, $L=16$ 或 8 . 设数据块间的绝对值误差和记为 TAD(Total Absolute Difference),平方误差和记为 TSD(Total Squared Difference).

定义宏块的像素方差(Var)

$$\text{Var} = \sum_{k=0}^{15} \sum_{l=0}^{15} data^2[k][l] - \frac{\left(\sum_{k=0}^{15} \sum_{l=0}^{15} data[k][l] \right)^2}{256}. \quad (1)$$

显然,Var 表征了宏块数据 $data[.][.]$ 的自相关程度,TAD 与 TSD 表征了宏块数据与参考数据之间的互相关程度. 如果 Var、TAD 与 TSD 的值越小,则各自所表征的相关性就越强.

考虑到类似性,这里仅以场编码方式 B 图象宏块的运动补偿类型判决为例. 设宏块的像素方差为 Var, 16×16 (场)双向预测的 TAD 值为 d_i , 16×16 (场)前向预测的 TAD 值为 d_f , 16×16 (场)后向预测的 TAD 值为 d_b , 16×8 上、下块双向预测的 TAD 值为 d_{8i} , 16×8 上、下块前向预测的 TAD 值为 d_{8f} , 16×8 上、下块前向预测的 TAD 值为 d_{8b} . 判决方法如下:

1) 一次判决:根据 TAD 值选择最佳的运动补偿类型. 即选择 $d_i, d_f, d_b, d_{8i}, d_{8f}$ 与 d_{8b} 中的最小值者所对应的运动补偿类型为最佳运动补偿类型.

2) 二次判决:根据 TSD 与 Var 值判决是否采用运动补偿. 根据最佳运动补偿类型所对应的运动向量计算宏块的 TSD 值. 如果 $TSD > \text{Var}$ 且 $TSD \geq 9 \times 256$ 时,宏块类型应为 I 块,不做运动补偿;否则采用运动补偿.

2.3 运动估计的软件实现

MPEG-2视频编码标准支持三种编码方式,即本文所定义的场编码方式、帧/场编码方式和帧编码方式.图1是这三种编码方式所支持的P、B图象宏块运动估计的软件实现方块图,当为B图象宏块时,需进行双向运动估计.实现时所有运动估计都是采用FSA算法,并包括半像素级(half pixel)搜索,且匹配准则为绝对值误差和TAD(注:双基运动估计进行在异侧参考场中的搜索时匹配准则采用平方误差和TSD).限于篇幅,这里忽略了包括半像素搜索、双基运动估计等在内的MPEG-2视频编码运动估计的软件实现具体步骤.

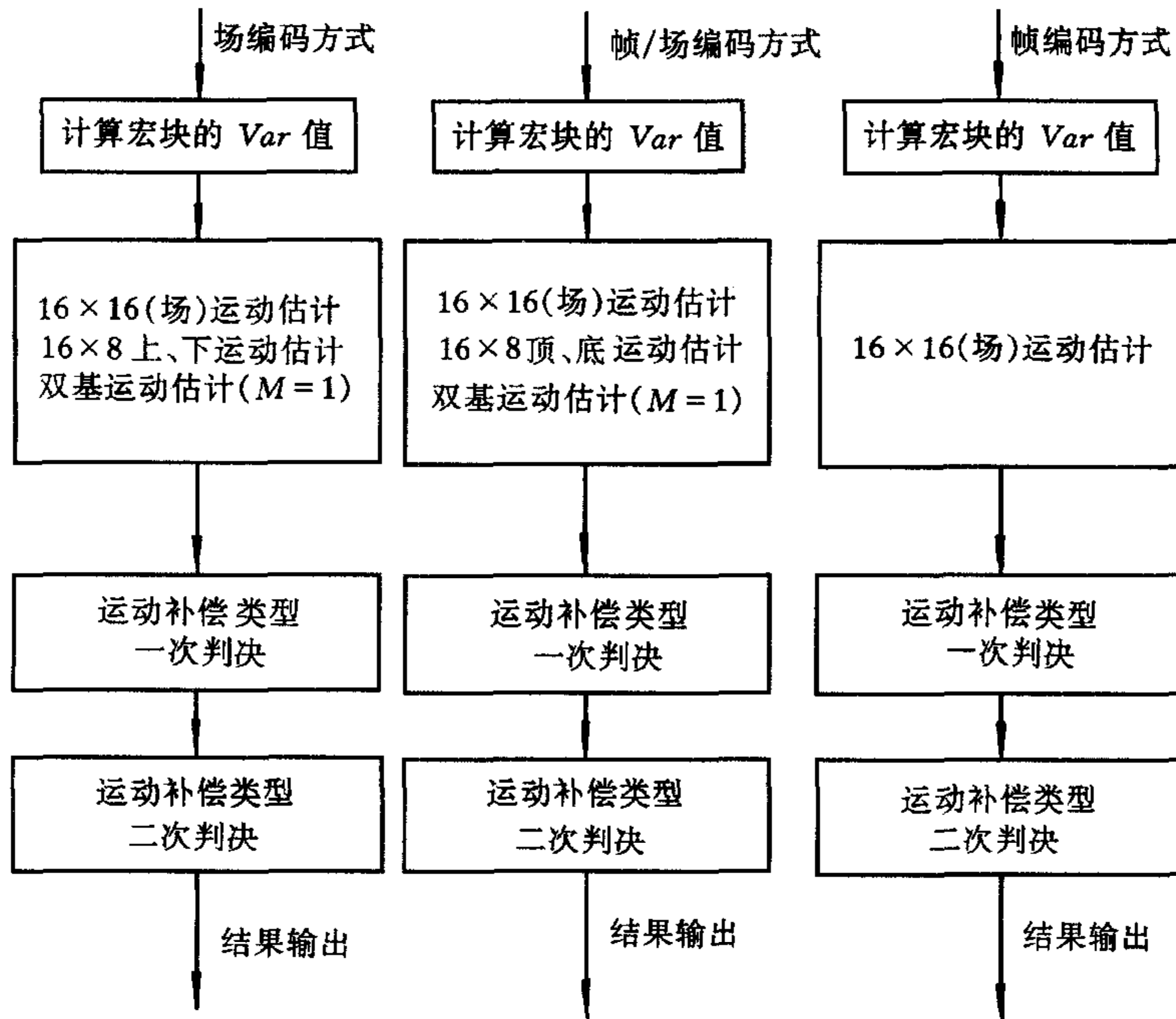


图1 P,B 图象宏块的运动估计

3 MPEG-2视频编码运动估计复杂度分析

MP@ML 广播级是支持 CCIR-601格式,即720pixels × 480lines × 30frames/sec 或相当.这里考查两种 GOP 结构下 MPEG-2编码器的每秒运动估计计算复杂度,一种是 N = 15、M = 3,另一种是 N = 15、M = 1(即不包含 B 类型图象).显然每帧图象的宏块数为 1350,每场图象的宏块数为 675,且每秒有两个整 GOP.根据不同的编码方式,可以得出各自的每秒运动估计计算量.设 16 × 16 块运动估计的运动向量在水平与垂直方向的最大位移为 ±V_{max},设 V_{max} = 16,则根据从上各式可具体求出各编码方式的运动估计复杂度(见表1).进一步在表2中给出了帧/场编码方式的整像素搜索、半像素搜索、类型判决以及双基搜索(M = 1)在总计算量中所占的百分比,由表2可知除整像素搜索外,半像素与双基运动估计(M = 1)的运算量显然也占有一定的份量.

4 模拟实验结果

本节主要考查一些简化实现 MPEG-2运动估计方案的性能,所得结果可以为整个 MPEG-2运动估计单元乃至编码器的实时实现提供参考.模拟实验是在标准测试序列 football、flower garden 与 mobile&calender 的各自前31帧上进行, $V_{\max}=16$,且图象组 GOP 的结构为 $N=15, M=3$ 或1.进行比较时采用的指标是三组测试序列 P、B 图象运动估计的平均 $\overline{\text{MSE}}$ 和解码图象总平均峰值信噪比 $\overline{\text{PSNR}}$.编码时设输出的比特速率为 8 Mbvits/s.

表1 $V_{\max}=16$ 时运动估计复杂度($\times 10^6/\text{sec}$)

| | 场编码方式 | | 帧/场编码方式 | | 帧编码方式 | |
|--------|-------------|-------------|-------------|-------------|-------------|-------------|
| | $N=15, M=3$ | $N=15, M=1$ | $N=15, M=3$ | $N=15, M=1$ | $N=15, M=3$ | $N=15, M=1$ |
| 加/减法数量 | 73 963.8 | 44 814.7 | 73 923.8 | 43 727.1 | 36 731.6 | 21 422.7 |
| 乘/除法数量 | 543.4 | 596.9 | 434.1 | 512.9 | 168.7 | 102.5 |
| 绝对值运算量 | 37 057.0 | 21 777.3 | 36 542.4 | 21 318.0 | 18 211.7 | 10 615.4 |
| 比较运算量 | 34.9 | 20.8 | 34.4 | 20.4 | 17.2 | 9.7 |

表2 百分比(帧/场编码方式)

| | $N=15, M=3$ | | | $N=15, M=1$ | | | |
|--------|-------------|--------|-------|-------------|--------|------------|-------|
| | 整像素 | 半像素 | 类型判决 | 整像素 | 半像素 | dual prime | 类型判决 |
| 加/减法数量 | 97.56% | 2.15% | 0.29% | 96.21% | 2.21% | 1.39% | 0.28% |
| 乘/除法数量 | 0.00% | 91.72% | 8.28% | 0.00% | 45.28% | 49.05% | 6.67% |
| 绝对值运算量 | 98.87% | 1.09% | 0.04% | 98.87% | 1.09% | 0.00% | 0.04% |
| 比较运算量 | 95.97% | 4.03% | 0.00% | 94.44% | 4.44% | 1.22% | 0.00% |

实验分四种情况进行,结果为:1)不同编码方式下的 $\overline{\text{MSE}}$ 比较($M=3$),帧编码方式的运动估计计算量是最小的,但实验结果表明帧编码方式的 $\overline{\text{MSE}}$ 普遍高于场编码方式与帧/场编码方式,即其运动估计效果最差, $\overline{\text{MSE}}$ 值约高出45左右;而场编码方式与帧/场编码方式的总平均 $\overline{\text{MSE}}$ 比较接近.2)帧/场编码方式与帧编码方式下的 $\overline{\text{PSNR}}$ 与 $\overline{\text{MSE}}$ 比较($M=1$), $M=1$ 时,帧/场编码方式带双基、帧/场编码方式不带双基与帧编码方式的 $\overline{\text{PSNR}}$ 与 $\overline{\text{MSE}}$ 比较结果:从 $\overline{\text{PSNR}}$ 上看,不带双基的帧/场编码方式比帧编码方式高0.31dB,而带双基的帧/场编码方式比不带双基的帧/场编码方式又高出0.27dB;从 $\overline{\text{MSE}}$ 上看,不带双基的帧/场编码方式比帧编码方式低39.17,而带双基的帧/场编码方式比不带双基的又低7.52.3)带半像素与不带半像素的 $\overline{\text{PSNR}}$ 与 $\overline{\text{MSE}}$ 比较($M=3$),当 $M=3$ 时带半像素与不带半像素的 $\overline{\text{PSNR}}$ 与 $\overline{\text{MSE}}$ 实验结果是:从 $\overline{\text{PSNR}}$ 上看,帧/场编码方式下不带半像素搜索时降低了3.74dB,帧编码方式下不带半像素搜索时降低了2.17dB;从 $\overline{\text{MSE}}$ 上看,帧/场编码方式下不带半像素搜索时提高了49.13,帧编码方式下不带半像素搜索时提高了50.06.4)以解码重建图象做参考与以原始图象做参考的 $\overline{\text{PSNR}}$ 比较($M=3$),在帧/场编码方式下,以

原始图象做参考时信噪比PSNR降低了1.59dB,帧编码方式下也降低了1.25dB.

参 考 文 献

- 1 ISO/IEC 13818-2:Information Technology-Generic Coding of Moving Pictures and Associated Audio: Video. ISO/IEC JTC1/SC29/WG11. 1994
- 2 Akiyama T, Aono H *et al.* MPEG-2 video codec using image compression DSP. *IEEE Transactions on Consumer Electronics*. 1994, **40**(3):466-472

MOTION ESTIMATION FOR MPEG-2 VIDEO CODING AND ITS IMPLEMENTATION

CHEN WEIQIANG GAO WEN

(Department of Computer Science and Engineering, Harbin Institute of Technology, Harbin 150001)

Key words Moving picture compress, MPEG-2, motion estimation.

学术活动简讯

首届全国“技术过程的故障诊断与安全性”学术会议 (中国 SAFEPROCESS' 99)征文通知

“技术过程的故障诊断与安全性”近20多年来已发展成为热点研究方向之一,理论与应用成果大量涌现.为推动我国这一领域研究工作的开展,并与国际自控联的相应机构接轨,中国自动化学会已于1997年10月批准成立“技术过程的故障诊断与安全性”专业委员会.首届学术会议拟定于1999年5月在北京清华大学举行.具体事宜如下:

一 征文范围

石油化工、电力、冶金、航空航天、电子、机械等行业中,涉及故障检测、故障诊断、可靠性、安全控制、监测控制、容错控制等方面的新理论、新方法、新技术和新的应用成果.

二 征文要求

1. 全文不超过7000字;2. A4单页小4号字激光打印(四周各留2.5厘米);3. 按《自动化学报》的论文格式排版;4. 一式两份(不管录用与否一律不退稿).

三 征文截稿日期:1998年12月31日.

四 录用通知发出日期:1999年1月31日.

五 来稿请寄:北京 100084 清华大学自动化系 周东华教授收

主办单位:中国自动化学会技术过程的故障诊断与安全
性专业委员会

清华大学自动化系(代章)