

单体型组装问题计算模型的分析与比较^{*}

周伟 王建新 谢民主 陈建二

(中南大学信息科学与工程学院 长沙 410083)

摘要 单体型检测在遗传病基因的定位、药理反应的研究、个体识别等方面有极其广阔的应用前景。单体型组装问题指如何利用个体的基因测序片断数据,根据不同的优化准则确定该个体单体型的计算问题。对 MSR, MFR, MEC, WMLF, MEC/GI 等单体型组装模型做了详细的分析比较,得出了如下结论:在没有引入测序误差情况下,上述模型的重构精度基本一致。随着测序误差的增加,MEC/GI 模型的容错性最好,重构精度最高;MSR 模型受测序误差的影响最大,只适用于测序误差极小的情形。

关键词 单核苷酸多态性,基因型,单体型,NP 难问题

Survey on Computing Models for Haplotype Assembly Problem

ZHOU Wei WANG Jian-xin XIE Min-zhu CHEN Jian-er

(School of Information Science and Engineering, Central South University, Changsha 410083, China)

Abstract Haplotype detection has expansive application in inherited gene's orientation, medicine reaction's research and individual identification. The haplotype assembly problem is the computing problem of inducing a individual's haplotypes based on several optimal criteria from one's DNA fragments sequencing data. This paper made detailed research on such models as MSR, MFR, MEC, WMLF, MEC/GI and got conclusion as follows: in the case of no reading errors, the reconstruction rate of these models is basically identical. As the reading errors increasing, the MEC/GI model suffers reading errors' influence least and has the highest reconstruction rate; MSR model suffers reading errors' influence most and is only suitable for the case of low reading errors.

Keywords Single-nucleotide polymorphisms, Genotype, Haplotype, NP-hard problem

1 引言

不同的人具有不同的外貌、体格,对疾病有不同抵抗能力,对药物有不同的敏感性,从遗传上说,这是因为不同个体(除了同卵双胞胎外)的基因组不完全相同。两个人之间的 DNA 差异约占基因组的 0.1%,单核苷酸多态性 SNPs(single-nucleotide polymorphisms)为人类染色体某个位点上的碱基变化。SNPs 广泛分布在人类基因组中,在整个人类基因组中大约有 340 万个 SNPs^[1]。

单核苷酸多态性是一个物种中不同个体表型的主要遗传来源。识别 SNPs,对基因的精确定位、了解基因功能很有帮助,对遗传病等疾病的诊断和药物研究有重要作用。SNPs 可用于个体识别、亲子鉴定,亦可用于人类各群体的遗传关系分析^[2,3]。Stephens 等采用个体单体型问题变异的方法研究人类 313 个基因中的 3899 个 SNPs 后进行连锁不平衡分析,其结果支持了人类群体在近代扩张的说法^[2]。Horikawa 等根据 SNPs 进行关联分析,在墨西哥裔美国人中把 2 型糖尿病基因定位在 2 号染色体长臂,并发现 CAPN10 基因的 3 个 SNPs 和 2 型糖尿病相关^[3]。

一个 SNP 位点指的是在一个物种的基因组 DNA 序列中不同个体可能出现不同碱基的位置。对于人类等二倍体生物,染色体是成对存在,每一对染色体的 DNA 序列除 SNP 位

点外都是一样的。在一条染色体 SNP 位点上的碱基序列叫做单体型(Haplotype)。而一对染色体上的某一区域的 SNP 位点上的混合的碱基序列叫做基因型(genotype)。对于任何一个二倍体生物,都有一对单体型。单体型在 SNPs 的上述应用中扮演着重要的角色。不幸的是,在当前的实验技术下,直接测定个体的单体型既费钱又费时间,因此利用计算机技术来确定个体的单体型有极其重要的现实意义。确定单体型的计算问题可以分为两大类:单体型组装问题和单体型推断问题^[4]。本文主要对单体型组装问题各计算模型的性能进行比较研究。

2 单体型组装模型

单体型组装问题就是给定一组来自某对同源染色体的由 DNA 测序方法得到的 DNA 片段数据,根据片段上的 SNP 值组装出两条单体型^[5]。一对同源染色体在对应的 SNP 位点上的值可以相同,也可以不同,因此 SNP 值可以用两个字符 0 和 1 来表示,而不必用真正的碱基 A, C, G 和 T,以减少计算复杂度。这样 DNA 片断的数据集可以表示为在 $\{0, 1, -\}$ 上的一个 $m \times n$ 的矩阵,叫做 SNP 矩阵 M ^[5],其中矩阵 M 的 m 行表示 m 个片断, n 列表示按在染色体上的次序从左到右的 n 个 SNP 位点, $M_{i,j}$ 的值表示第 i 个片断在第 j 个 SNP 位点上的取值,其值可以是 0, 1 或 -, 其中 '-' 表示片断在该位点

^{*}国家自然科学基金重点项目:生物信息学中的相关组合理论和算法研究(60433020)。周伟 硕士研究生,主要研究领域为参数计算、计算机理理论;王建新 博士,教授,博士生导师,主要研究领域为计算机算法、网络优化理论、生物信息学;谢民主 硕士研究生,主要研究领域为参数计算、计算机理理论;陈建二 博士,长江学者特聘教授,博士生导师,主要研究领域为生物信息学、计算机理论、计算复杂性及优化。

的取值为空(其值未能确定),或者说该片断在这个 SNP 位点上有一个洞(hole)。

对于 SNP 矩阵 M 的两行,如果它们在列 j 上的值不相等,且都不为‘-’,则这两行在列 j 上冲突;如果这两行在所有的列上均不冲突,则这两行互相兼容。

如果 M 的所有行可以分成 2 个不相交的子集,且每个子集中的所有行都相互兼容,则称 M 可行。

理想的无任何测序误差过程得出的片段数据所对应的 SNP 矩阵显然是可行的,因为来自同一染色体的片段显然是相互兼容的。可是,由于当前技术的局限性,DNA 测序存在错误和直接测序片断长度较短,从而很难确定哪些 DNA 片断来自哪条染色体。Lancia 等对这个问题进行了研究,首先提出了下面 2 个计算模型^[5]:

最少片段删除(Minimum Fragment Removal, MFR):去掉最少的片段(即 SNP 矩阵的行),使得由该 SNP 矩阵可以组装出两个单体型,即使得该 SNP 矩阵可行;

最少 SNP 删除(Minimum SNP Removal, MSR):去掉最少的 SNP 位点(即 SNP 矩阵的列),使得 SNP 矩阵可行。

为了避免丢失片断或者位点信息,Lippert 等进一步提出了第 3 个计算模型^[6]。

最少错误修正(Minimum Error Correction, MEC)或叫做 MLF(Minimum Letter Flips):该模型要求修改最少的 SNP 矩阵元素,使得 SNP 矩阵可行。

在 MEC 模型的基础上,结合基因型的信息,文献[7]进一步提出以下模型。

MEC/GI(Minimum Error Correction/ Genotype Information):该模型要求修改最少的 SNP 矩阵元素,使得 SNP 矩阵可行且与基因型兼容。

基因型用字符集 $\{0, 1, 2\}$ 上的字符序列来表示,对基因型 $g = (g_1, g_2, \dots, g_n)$,以及两单体型 $h_1 = (h_{11}, h_{12}, \dots, h_{1n})$, $h_2 = (h_{21}, h_{22}, \dots, h_{2n})$,如果 $g_j \neq 2$,则令 $h_{1j} = h_{2j} = g_j$;如果 $g_j = 2$,求解 h_{1j}, h_{2j} 则与 MEC 模型相同。

类似地,为了利用位点的可信度权值信息,Greenberg 等提出了带权的 MLF 模型^[4]。

最少权值和翻转(weighted Minimum Letter Flips, WM-LF):给定片断数据以及位点的可信度权值,翻转某些位点的取值(0 翻转到 1 或从 1 翻转到 0),使得 SNP 矩阵可行,同时使得翻转的位点所关联的权值总和最小。

本文对单体型组装现有的 5 种主要计算模型 MFR, MSR, MEC, MEC/GI, WMLF 进行大量的对比实验,对这些模型的单体型重构的精确度进行比较,分析影响模型精确度的原因,从而揭示出各模型的内在特性,为进一步改进单体型组装模型提供依据。

3 单体型组装模型评价指标

单体型组装过程中,片断的处理分为以下 3 个步骤:最初为片断生成器根据初始单体型所生成的初始片断,然后引入了测序误差,最后是经过算法处理后所保留的片断信息,记为 f_c 。为了方便后面的描述,我们把由初始片断的数据集所对应的 SNP 矩阵记为 M ,由 f_c 的数据集所对应的 SNP 矩阵记为 P ,根据 P 的片断信息组装得到的单体型称为重构单体型。令 $x, y \in \{0, 1, -\}$,记

$$d(x, y) = \begin{cases} 1, & \text{if } x \neq -, y \neq - \text{ and } x \neq y \\ 0, & \text{otherwise} \end{cases}, g(x) = \begin{cases} 1, & \text{if } x \neq - \\ 0, & \text{otherwise} \end{cases}$$

$$E(x, y) = \begin{cases} 1, & \text{if } x \neq y \\ 0, & \text{otherwise} \end{cases}。以上定义将出现在后面对实验指标的$$

定义中。

本文通过以下指标对上述模型进行实验比较和分析:

①不同错误率下 MSR 模型的位点删除率、MFR 模型的片断删除率及 MEC, MEC/GI, WMLF 模型的位点修正率。各模型在进行单体型重构时,对片断或位点的处理具有其自身的特点。MSR 模型在处理过程中删除了片断所包含的某些位点信息,因此针对 MSR 模型提出了位点删除率的概念,即删除的单体型位点数与单体型长度的比值。MFR 模型则删除了整个片断信息,因此针对 MFR 模型提出了片断删除率的概念,即删除的片断数占有片断总数的比值。比较而言,MEC, MEC/GI, WMLF 模型不删除任何信息,而通过修改位点信息对片断进行处理,所以针对这 3 个模型提出了位点修正率的概念,即翻转的 SNP 位点数目(矩阵 P 与矩阵 M 中对应位点处取值不同的位点数目)占矩阵 M 中 0, 1 位点总数的百分比。

②片断错误率^[8]。片断错误率是 SNP 矩阵 P 中的错误信息占 P 中所有 0, 1 信息的比率,可表示为 $\sum_{ij} d(P[i, j], M[i, j]) / \sum_{ij} g(P[i, j])$ 。

③信息丢失率,信息丢失率表明了矩阵 M 与矩阵 P 之间的差异率,可表示为 $\sum_{ij} E(P[i, j], M[i, j]) / \sum_{ij} g(M[i, j])$ 。

根据片断错误率和信息丢失率,我们可以从分析片断数据的角度来分析各模型的特点及其对单体型重构率的影响。

④单体型重构率。文献[7]给出了它的具体定义:对任意两条片断 $m_i = (m_{i1}, m_{i2}, \dots, m_{in})$ 以及 $m_k = (m_{k1}, m_{k2}, \dots, m_{kn})$,定义 $HD(m_i, m_k)$ 为片断 m_i 和片断 m_k 相冲突的 SNP 位点数目,即 $HD(m_i, m_k) = \sum_{j=1}^n d(m_{ij}, m_{kj})$ 。令初始单体型 $h = (h_1, h_2)$, 重构单体型 $\bar{h} = (\bar{h}_1, \bar{h}_2)$, 令 $r_{i,j} = HD(h_i, \bar{h}_j)$, $i=1, 2, j=1, 2$, 则单体型重构率可表示为 $RR(h, \bar{h}) = 1 - (\min\{r_{11} + r_{22}, r_{12} + r_{21}\} / 2n)$ 。

由于没有考虑到重构单体型中存在的空隙,文献[7]提出的单体型重构率仅局限于重构单体型中不存在空隙的情况。为了更精确地揭示模型特性,必须考虑重构单体型中存在的空隙。这是因为,首先在某些情况下片断中存在空隙;其次,在某些模型进行单体型重构时,会丢失大量的信息,比如 MFR 模型会删除片断信息,MSR 模型会删除位点信息。这样,导致重构出来的单体型中某些 SNP 位点的信息为‘-’,即丢失了位点信息。为此,我们提出单体型空隙率来衡量重构单体型中空隙所占的比例,其定义为重构单体型中空隙的个数与单体型长度的比值。显然,文献[7]定义的单体型重构率与单体型空隙率的差值才是各模型真实的单体型重构率。因此在文献[7]的基础上,本文扩展了对单体型重构率的定义:单体型重构率 = 文献[7]定义的单体型重构率 - 单体型空隙率。

单体型重构率表明了重构单体型与初始单体型的近似度,是衡量各模型优劣的标准。

4 实验分析

4.1 实验环境

由于原始的 DNA 测序片段数据很难得到,大量文献利用计算机模拟真实生物数据的特征生成测试数据集,进行个体单体型各种算法的实验测试^[8, 11]。本文采用与上述文献相同的方法和参数来生成测试数据。模拟数据生成的方法如

下:首先随机生成指定长度的单体型,根据指定的两个单体型的差异率来随机生成另一个单体型。然后根据指定的测序误差、片段的覆盖率、片段的最小长度和最大长度来随机生成片段数据集。实验室的 DNA 测序误差为 3%~5%^[8],片段的覆盖率为 10 左右^[12]。为了使模拟生成的片段数据能很好地反映真实情况,根据文献^[13]的测试方法,首先采用著名的 shotgun 测序模拟片断生成器 Celsim 生成一系列的片断数据,生成参数设置为两个单体型的差异率为 20%,测序误差为 5%~20%,生成的普通片段的最小长度为 3,最大长度为 7,生成的 mate-pair 是由相隔指定 SNP 位点的来自于同一个单体型长度为 6 的一对普通片段构成,单体型长度设为 200,片段覆盖率等于 10,生成片断数的设置则按单体型长度 * 片断覆盖度/片断平均长度设置。模拟片断生成器的详细情况请参照文献^[8,13]。在一台有 4 个 Intel Xeon 3.6G CPU,内存为 4G 的安装了 Linux 操作系统的服务器上,通过一系列的模拟数据对前面介绍的各个模型进行了测试。以下各图中每一个点均为 100 次重复测试的平均值。

通过实验分析,我们发现计算模型与片断是否存在空隙之间没有关联,因此本文在片断不存在空隙的情况下,采用精确算法对各模型的单体型重构精度进行评估^[9,10]。

4.2 实验结果

图 1 说明了不同错误率下 MSR 模型的位点删除率, MFR 模型的片断删除率及 MEC, WMLF, MEC/GI 模型的位点修正率。由图 1 可知,随测序误差的增大,MSR 模型的位点删除率和 MFR 模型的片断删除率急剧增加。其中,MSR 模型在错误率为 5% 时,其位点删除率已达到 28%,这说明其单体型重构率最多也不会超过 72%,因此 MSR 与 MFR 模型的容错性能较差。MEC 模型的位点修正率最低,这说明其翻转了最少数目的位点,实验结果与其翻转最小位点使得 SNP 矩阵可行的模型特性相吻合。WMLF 与 MEC/GI 模型的位点修正率和测序误差基本保持一致,且高于 MEC 模型,说明这 2 个模型比 MEC 能更好地识别出错误的位点信息,具有良好的容错性。

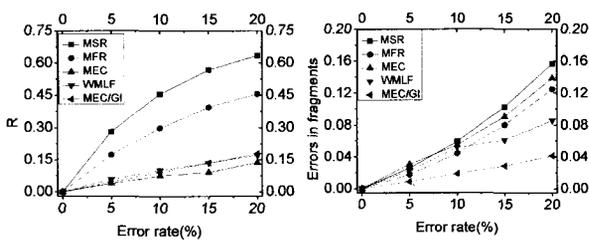


图 1 不同错误率下 MSR 模型的位点删除率, MFR 模型的片断删除率及 MEC, WMLF, MEC/GI 模型的位点修正率

图 2 揭示了不同错误率下各模型的片断错误率。由图 2 可知,MEC/GI 模型最低,在错误率较大时, WMLF 模型的片断错误率也比较低,其它模型的片断错误率则比较接近。MSR 模型在删除位点信息时,其宗旨为删除最少的位点,而不能保证保留下来的位点都是正确的,所以会产生以下情况:正确的位点被删除了,而错误的位点被保留下来,导致其片断错误率的产生。MFR 模型在删除片断时会发生如下情况:正确的片断被删除了,而存在测序误差的片断被保留在最后的输出矩阵中,导致其片断错误率的产生。由于 MSR 与 MFR 模型删除了大量信息,这些被删除的信息包含了大量错误信

息,因此其片断错误率远低于位点删除率和片断删除率。MEC 模型在修正位点值时缺乏附加信息的指导,会发生把正确位点翻转成错误位点的情况,在错误率较大时这种情况尤为明显,所以其片断错误率较高。WMLF 模型利用位点的可信度权值指导位点修正,在错误率较大时,其片断错误率低于 MEC。MEC/GI 模型的位点修正率与测序误差基本相同,且利用基因型的信息来指导位点修正,使得引入了测序误差的片断经过位点修正之后重新成为正确的片断,故其片断错误率低。

图 3 说明了不同错误率下各模型的信息丢失率。由图 3 可知,MSR 与 MFR 模型的信息丢失率远大于引入的测序误差,而 MEC, WMLF 与 MEC/GI 模型的信息丢失率较低。这是因为 MSR 模型的位点删除率很高,且在删除某个位点时,会影响覆盖此 SNP 位点的所有片断,从而丢失了大量信息。MFR 模型的片断删除率稍低于 MSR 模型的位点删除率,且在删除某条片断时,删除了片断所包含的所有位点信息,故其信息丢失率高。MEC, WMLF 与 MEC/GI 模型的信息丢失率则远小于测序误差,这是因为这几个模型不会删除位点或片断的信息,从而保持了初始 SNP 矩阵信息的完整性,且模型本身具有根据正确的位点信息以及附加信息自适应修正错误位点的行为,因此大大减少了测序误差所带来的负面影响,降低了信息丢失率。

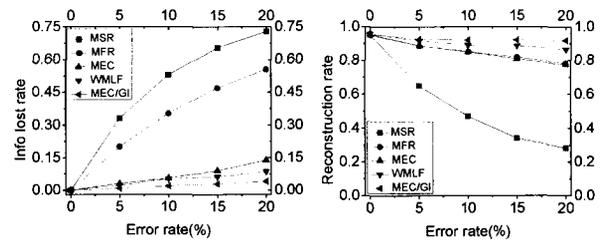


图 3 不同错误率下各模型的信息丢失率

图 4 揭示了不同错误率下各模型的单体型重构率。当测序误差为 0 时,上述 4 个模型的重构精度基本一致,达到 95% 以上。这也证明使用上述计算模型来计算重构单体型的合理性。但在测序误差为 0 的情况下,单体型重构率并没有达到 100%。这是因为 SNP 位点矩阵可组装成多对不同的单体型,不一定恰好组装成原始单体型。

由图 4 可知,随测序误差的增大,MSR 模型的单体型重构率急剧下降。这是因为 MSR 模型在进行单体型重构时直接删除位点信息,从而导致空隙的产生,其位点删除率决定了单体型空隙率的大小。在删除位点信息时,其宗旨为删除最少的位点,而不能保证保留下来的都是正确的位点,所以会产生以下情况:正确的位点被删除了,而发生了错误的位点被保留下来,从而导致了片断错误率的产生,这进一步降低了 MSR 模型的单体型重构率。随测序误差的增大, MFR 模型的单体型重构率呈线性下降的趋势。这是因为,为使 SNP 矩阵可行, MFR 模型在单体型重构时需要删除片断信息,且随着测序误差的增大,冲突的片断数目越来越多,删除的片断数也随之增加。随着片断删除率的增加,其信息丢失率也急剧增大,导致单体型重构率下降。随测序误差的增大, MEC 模型的单体型重构率与 MFR 基本一致。尽管 MEC 模型本身具有根据正确的位点信息来判断和修正错误信息的能力,但当错误率较大时,覆盖某个位点的错误片断数目可能大于正确片断数目,此时正确位点反而会翻转成错误位点,导致重构

单体型的对应位点发生错误。WMLF 与 MEC/GI 模型受测序误差的影响最小,具有最高的重构精度。这是因为 WMLF 在 MEC 模型的基础上利用了位点可信度权值的信息来修正 SNP 位点的取值,MEC/GI 则在 MEC 模型的基础上利用了基因型的信息来修正 SNP 位点的取值,与 MEC 模型启发式的自适应修正位点取值比较,降低了测序误差的影响,具有更高的准确性。

结束语 单体型的组装问题是生物信息学中一类重要的计算问题。本文讨论了组装问题主要计算模型的特点,其中 MFR 是基于错误片断删除;MSR,MEC,WMLF 是基于冲突片断修复;MEC/GI 则是一种“混合”的个体单体型模型,即同时利用了 SNP 片断及基因型数据的信息。在此基础上着重分析了各模型特点和单体型重构精度的关系,并给出了详细的实验数据。

图理论在单体型组装模型的研究中起到了关键的作用。给定 SNP 矩阵 M 以及冲突图 $G(V, E)$, M 的行集 $\{m_1, \dots, m_n\}$ 对应图 G 的顶点集 $V\{v_1, \dots, v_n\}$, 行 m_i 与 m_n 冲突意味着顶点 v_i 与 v_n 之间存在边相连。使 SNP 矩阵可行的过程可看成冲突图向二分图转化的过程,区别仅在于各模型对片断中错误位点处理的方式不同,从而导致了图操作的不同。除 MEC/GI 模型,本文介绍的其它模型都是直接从冲突图推导出的模型。在实际应用中,使用哪个模型取决于对给定数据预先的了解。

为了模型之间比较的可行性和正确性,本文采用的都是对应于各模型的精确算法。注意到绝大多数单体型组装模型的 APX-hard 特性,精确算法尽管能够保证解的全局最优性,但是当问题规模较大时,其指数级时间复杂度使得算法在实际应用中不可行。因此即使对同一模型,针对不同的情况需要不同的算法设计。由于单体型组装模型的组合特性,因此很难设计出好的具有常数因子的近似算法。为此许多学者进行了启发式算法的研究,代表性的有遗传算法^[7,15]和人工神经网络^[16]方面的研究。尽管这些方法不能提供所得解的近似度或求得精确解,但理论上利用启发式算法可收敛到局部最优解。

对各计算模型而言,最重要的是如何提高其重构精度。现有研究表明,通过利用基因型或位点的可信度权值信息可以有效地提高单体型重构精度,减少测序误差的影响。通过分析现有的单体型组装计算模型,我们可得出设计单体型组装计算模型时应尽量利用附加信息,如基因型信息和位点的可信度权值信息。同时,应尽量避免丢失位点或片断信息,而且模型应具有根据正确的位点信息来纠正错误信息的能力,以提高容错性能。

参 考 文 献

[1] Miller T P, Gu Z, Li Q, et al. Overlapping genomic sequences: a treasure trove of single-nucleotide polymorphisms. *Genome Res.*, 1998, 8: 748-754

[2] Stephens J C, Schneider J A, Tanguay D A, et al. Haplotype variation and linkage disequilibrium in 313 human genes. *Science*, 2001, 293: 489-493

[3] Horikawa Y, Oda N, Cox N J, et al. Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus. *Nature Genetics*, 2000, 26:163-175

[4] Clark. Inference of haplotypes from PCR-amplified samples of diploid populations. *Mol. Biol. Evol.*,1990,7(2): 111-122

[5] Lancia G, Bafna V, Istrail S, et al. SNPs problems, complexity and algorithms//Proceedings of the 9th Annual European Symposium on Algorithms. 2001:182-193

[6] Lippert R, Schwartz R, Lancia G, et al. Algorithmic strategies for the SNPs haplotype assembly problem. *Briefings in Bioinformatics*, 2002, 3(1): 23-31

[7] Wang R S, Wu L Y, Li Z P, et al. Haplotype reconstruction from SNP fragments by minimum error correction. *Bioinformatics*, 2005, 21(10): 2456-2462

[8] Panconesi A, Sozio M. Fast Hare: a Fast Heuristic for Single Individual SNP Haplotype Reconstruction//Proceedings of WABI, LNCS. Bergen, Norway: Springer-Verlag, 2004

[9] Xie M Z, Chen J E, Wang J X. Parameterized algorithm of the individual haplotyping MSR problem with mate-pairs. *Journal of Software*, 2007, 18(9): 2070-2082

[10] Research on Parameterized Algorithms of The Individual Haplotyping Problem. *Journal of Bioinformatics and Computational Biology (JBCB)*, 2007, 5(3): 795 - 816

[11] Li L, Khuri S. A Comparison of DNA Fragment Assembly Algorithms//Valafar F, Valafar H, eds. Proceedings of the International Conference on Mathematics and Engineering Techniques in Medicine and Biological Sciences. Nevada: CSREA Press, 2004: 329-335

[12] International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature*, 2001, 409: 860-921

[13] Myers G. A dataset generator for whole genome shotgun sequencing//Lengauer T, Schneider R, Bork P, et al., eds. Proceedings of 7th International Conference Intelligent Systems for Molecular Biology. California: AAAI Press, 1999: 202-210

[14] Greenberg H J, Hart W E, Lancia G. Opportunities for combinatorial optimization in computational biology. *INFORMS J. Comput.*, 2004, 14: 211-213

[15] Goldberg D E. Genetic Algorithms in Search, Optimization and Machine Learning. Reading: Addison-Wesley, 1989

[16] Zhang X S. Neural Networks in Optimization. Kluwer Academic Publishers, 2000

(上接第 127 页)

文采用小波分析和自适应算法对其进行了研究,深入研究了小波自适应算法的理论,并且进行了计算机仿真,结果表明,小波自适应算法可以有效地提高传感器信号的信噪比,从而可以提高信息融合的效果。

参 考 文 献

[1] 徐凯,纪红,乐光新. 一种改进的变步长自适应滤波器 LMS 算

法[J]. 电路与系统学报, 2004(4)

[2] 张子瑜,吴镇扬,任永川,等. 离散小波变换域 LMS 自适应消噪[J]. 电路与系统学报, 2000(1)

[3] 高鹰,谢胜利. 一种基于相关函数的 LMS 算法[J]. 数据采集与处理, 2004(1)

[4] 国良,强文义. 移动机器人信息融合技术研究[J]. 哈尔滨工业大学学报, 2003(7)

[5] 华志,谢存禧,曾德怀. 多传感器信息融合在移动机器人导航中的应用[J]. 组合机床与自动化加工技术, 2003(9)