

一种确定混沌伪随机序列复杂度的模糊关系熵测度*

陈小军^{1)†} 李赞¹⁾ 白宝明¹⁾ 蔡觉平²⁾

1)(西安电子科技大学综合业务网理论与关键技术国家重点实验室,西安 710071)

2)(西安电子科技大学微电子学院,西安 710071)

(2010年6月12日收到;2011年3月1日收到修改稿)

将模糊关系的概念引入混沌伪随机序列复杂度的测度方法之中,提出了一种新的混沌伪随机序列复杂度测度方法——模糊关系熵(fuzzy relationship entropy,简记为F-REn)测度方法,并推导了F-REn的两个基本性质.仿真结果表明,该测度方法能够有效测度混沌伪随机序列的复杂度,与近似熵(ApEn)测度方法和符号熵测度方法相比,F-REn测度具有更加好的对序列符号空间的适用性、更加小的对测量维度的敏感性和更加强的对分辨率参数的鲁棒性.

关键词: 混沌伪随机序列, 模糊理论, 复杂度

PACS: 42. 65. Sf

1. 引言

混沌伪随机序列是通过对混沌映射产生的序列进行量化而得到的伪随机序列. 由于其具有类噪声、对初始值敏感以及非周期等特性,因此,可以取代传统的伪随机序列应用于保密扩频通信之中^[1,2]. 由于混沌伪随机序列的复杂度是指它与随机序列的相似程度,是对利用序列的部分恢复出整体的难易程度的度量. 因此它是衡量保密通信系统中扩频序列抗干扰和抗截获能力的重要指标^[3-5].

复杂度问题的研究一直以来都被国内外学者广泛关注,并与非线性科学及其混沌动力学的复杂度研究交错在一起,形成了非线性科学和复杂度问题的研究热潮. 其研究大致经过了三个阶段:在第一阶段,在文献[6]中Kolmogorov提出了复杂度的原创思想,之后,Lempel和Ziv将Kolmogorov复杂度在计算机中实现^[7]. 在第二阶段,为了进一步提高复杂度测度方法的实用性,相关学者提出了采用线

性复杂度(LC)^[8-10]、线性复杂度曲线(LCP)^[11,12]以及 k 错线性复杂度(k -error LC)^[13,14]进行序列复杂度测度. 但是,由于这些方法都是针对常规伪随机序列提出的,而随着混沌伪随机序列的深入研究和应用推广,其难以胜任混沌伪随机序列复杂度的准确测度^[15]. 因此,在第三阶段,为了满足这种测度需求,文献[15]提出了基于近似熵(ApEn)的混沌伪随机序列复杂度测度方法. ApEn是一种基于边缘概率分布统计的量化时间序列复杂程度的方法^[16]. 已经被广泛应用于医学生理信号处理^[17,18]、气象预测^[19]和机电信号故障诊断^[20]等领域. ApEn测度方法虽然能够较为准确的测度不同伪随机序列的复杂度,但仍然存在测度结果受参数选取影响较大的问题. 在这种情况下,文献[21,22]提出了采用符号动力学方法测度混沌伪随机序列复杂度的方法(又称为符号熵方法). 该方法能够有效的测度不同复杂度的混沌伪随机序列,并且进一步降低了参数选取对测度结果的影响. 目前其已经被应用在事件相关电位分析^[23]等方面. 但是,由于符号动力

* 国家科技重大专项(批准号:2010ZX03006-002-04)、国家高技术研究发展计划(批准号:2009AA01Z237)、国家自然科学基金(批准号:61072070)、教育部新世纪优秀人才支持计划(批准号:NCET-07-0653)、高等学校学科创新引智计划(批准号:B08038)、教育部长江学者和创新团队发展计划(批准号:IRT0852)和陕西省自然科学基金基础研究计划项目(批准号:SJ08F09)资助的课题.

† E-mail: xjchen@mail.xidian.edu.cn

学方法中固有的粗粒化标示特点,其在序列符号空间元素较多或者窗口向量长度较大的情况下,计算复杂度很大,且在测度之前必须已知待测序列的符号空间大小,而在实际应用中要得到这个先验知识是很困难的.因此,目前需要研究一种新的能够在不同参数条件下准确完成混沌伪随机序列复杂度测度的理论.

1965年 Zadeh 提出了模糊集合的概念^[24,25],用数学的方法对现实世界的问题进行了模糊性和不确定性研究.自此,模糊理论作为一门重要的数学分支得到了深入的研究.在模糊理论的研究基础之上,Chen 等提出了使用模糊熵(fuzzy entropy,简记为 FuzzyEn)作为医学生理电信号的特征测度标准^[26,27],将模糊理论应用于生理电信号序列复杂度测度之中,得到了很好的测度效果.针对上述现有混沌伪随机序列复杂度测度的不足和问题,本文将模糊理论中关于模糊关系的概念^[24]引入到混沌伪随机序列复杂度的统计学测度方法之中.提出了模糊关系熵(fuzzy relation entropy,简记为 F-REn)概念,并将其作为混沌伪随机序列复杂度测度的新标准,有效克服了前述 ApEn 和符号动力学方法的局限,提高了混沌伪随机序列复杂度测度的准确性和实用性.

2. 基于 ApEn 和符号熵的混沌伪随机序列复杂度测度

为了便于对比和分析本文提出的 F-REn 测度方法的优越性,在本节首先简要介绍基于 ApEn 和符号熵的混沌伪随机序列复杂度测度定义.

2.1. 基于 ApEn 的混沌伪随机序列复杂度测度

定义 1^[16] 对于一个长度为 N 的序列样本空间 $[u(1), u(2), \dots, u(N)]$,首先定义 m 维向量组 $\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(i), \dots, \mathbf{x}(N - m + 1) \in R^m$, 其中

$$\mathbf{x}(i) = [u(i), u(i + 1), \dots, u(i + m - 1)]$$

$$(1 \leq i \leq N - m + 1). \quad (1)$$

m 维向量 $\mathbf{x}(i)$ 和 $\mathbf{x}(j)$ 的最大距离为

$$d[\mathbf{x}(i), \mathbf{x}(j)] = \max_{k=1,2,\dots,m} (|u(i + k - 1) - u(j + k - 1)|), \quad (2)$$

可得

$$C_i^{m,N}(r) = (\text{满足 } d[\mathbf{x}(i), \mathbf{x}(j)] \leq r \text{ 的 } j \text{ 的个数}) / (N - m + 1). \quad (3)$$

根据 $C_i^{m,N}(r)$, 定义

$$\Phi^{m,N}(r) = (N - m + 1)^{-1} \sum_{i=1}^{N-m+1} \ln C_i^{m,N}(r), \quad (4)$$

那么 ApEn 可以被定义为

$$S_{Ap}(m, r, N) = \Phi^{m,N}(r) - \Phi^{m+1,N}(r). \quad (5)$$

向量维数 m 的最大值由观察空间的长度 N 确定,当 m 越大时, ApEn 越接近测度熵^[16].文献[28]对二进制序列的 m 的选取进行了详细的讨论. Ornstein 和 Weiss 的研究表明:对于长度为 N 的 K 进制随机序列,为了测定长度为 m 的窗口向量出现的概率,所需观测序列长度的增长大于取值空间 K 的指数的增长,因此最小的观察区间长度满足 $K^{2^m} < N$, 即

$$m_{\max} = \max(m: K^{2^m} < N). \quad (6)$$

因此,可以看出, ApEn 方法的准确测度必须建立在准确的参数选取基础之上,在第 4 节将仿真验证参数选取对 ApEn 测度结果的影响.

2.2. 基于符号熵的混沌伪随机序列复杂度测度

为了识别一个混沌时间序列中的不同的动力学演化形式,一般地可以通过粗粒化的方法,将其转化为一个相应的符号序列,这样混沌时间序列其轨道演化的信息就通过符号序列达到了编码.混沌序列的粗粒化可以通过对相空间的分割而实现.

定义 2^[21,22] 给定一个由 t 个符号组成的符号集 $\{s_0, s_1, \dots, s_{t-1}\}$ 和一个 $t + 1$ 个临界点组成的集合 $\{C_0, C_1, \dots, C_t\}$, 可以用以下分割规则将时间序列 $\{u_1, u_2, u_3, \dots\}$ 转换为一个符号序列 $\{s(1), s(2), s(3), \dots\}$, 若 $C_k < u_j \leq C_{k+1}$, 则 $s(j) \equiv s_k$. 将所得符号序列 $\{s(1), s(2), s(3), \dots\}$ 分割成为长度为 m 的窗口向量,标记为

$$I_x(m, i) = \sum_{p=1}^m t^{m-p} s(p + i). \quad (7)$$

其中, t 是符号集 $\{s_0, s_1, \dots, s_{t-1}\}$ 中不同符号的个数, m 表示窗口向量的长度, i 表示窗口向量沿着符号序列 $\{s(1), s(2), s(3), \dots\}$ 从第 i 个符号开始.为了方便,通过将符号 s_k 用相应的整数 k 来代替,由(7)式可以看出每一个窗口向量都可以方便的用整数集

$$\{0, 1, 2, \dots, t^m - 1\} \quad (8)$$

的一个整数唯一的标记和辨别.用 P_{I_x} 表示特定窗口向量 I_x 出现的概率,它可以用窗口向量 I_x 在混沌符号序列 $\{s(1), s(2), s(3), \dots\}$ 中出现的次数除以所有窗口向量的总数计算.包含在符号序列中的信息量可以用下式引入的符号熵来定量表示:

$$S_{\text{ymb}}(m, t, N) = -\frac{1}{m} \sum_{l_x} P_{l_x} \ln P_{l_x}. \quad (9)$$

对于一个由混沌序列经过量化和判决得到的伪随机序列,伪随机序列的每一个值又可以看做一个符号,因此,混沌伪随机序列可以直接看成一个符号序列.将混沌伪随机序列分割成长度为 m 的窗口向量,通过用(7)和(8)式统计出不同窗口向量出现的次数,求出各窗口向量在混沌伪随机序列中出现的概率,就可以用(9)式直接计算出混沌伪随机序列的符号熵,根据符号熵的大小来判断序列的复杂度.

从(8)和(9)式可知,符号熵的计算中,窗口向量的状态集合由 t^m 确定,因此可知,当序列符号状态数较大的情况下,或者窗口向量较长的情况下,符号熵的计算复杂度会急剧增加,这就造成了符号熵在多进制混沌伪随机序列情况下的复杂度测度应用受到严格的限制.例如当 $t = 64, m = 4$ 时,窗口向量状态集合中元素数为 $t^m = 16,777,216$. 由此可知,对于如此大的状态空间进行符号熵值的计算实现非常困难,且时效性不好.

3. 基于模糊关系熵(F-REn)的混沌伪随机序列复杂度测度

针对上述测度方法的不足和问题,我们根据混沌伪随机序列的特性,构造出相适应的模糊关系,将模糊关系度量模块嵌入统计度量方法之中,有效提高了混沌伪随机序列复杂度测度的通用性和准确性.

3.1. 模糊关系熵的引入

通过定义1可以看出,(3)式为测度在 $1 \leq j \leq N - M + 1$ 情况下 m 维窗口向量 $x(i)$ 与 $x(j)$ 之间的最大距离在容限 r 之内的情况出现的频繁程度(当 $N \rightarrow +\infty$ 时,也就是此种情况出现概率值).显而易见,在 ApEn 序列复杂度测度过程中, $C_i^{m,N}(r)$ 对最终序列复杂度的确定起着决定性作用,从(3)式可知, $C_i^{m,N}(r)$ 又由分辨率参数 r 确定,当 $d[x(i), x(j)] \leq r$ 时 $C_i^{m,N}(r)$ 的分子部分计数加1,否则计数加0.如下式所示:

$$C_i^{m,N}(r) = \begin{cases} +\frac{1}{N-m+1}, & (d[x(i), x(j)] \leq r), \\ +0, & (d[x(i), x(j)] > r). \end{cases}$$

由此可知, $C_i^{m,N}(r)$ 作为单调递增函数,在度量分辨

率参数 r 确定的情况下, $C_i^{m,N}(r)$ 的分子部分每次的叠加取值“非0即1”.在这种“硬判决”机理下,容易出现不同窗口向量之间相似程度测度混淆的问题,从而丢失了一部分序列复杂度判决信息,导致序列复杂度测度受参数值影响较大,出现测度错误的情况.

为了解决这个问题,我们引入模糊理论中的模糊关系概念,并采用相应的模糊隶属函数作为不同向量维度之间的模糊关系度量,下面我们首先给出模糊关系的定义,并在此基础上,对模糊关系熵的进行定义.

3.2. 基于序列复杂度测度的模糊关系定义

为了将模糊关系的概念应用于不同维度向量距离的度量之中,首先给出基于混沌伪随机序列复杂度测度的模糊关系定义如下:

定义3 设长度为 m 的序列窗口向量为 $\{s^m(i) = [u(i), u(i+1), \dots, u(i+m-1)]\}$ 和 $\{s^m(j) = [u(j), u(j+1), \dots, u(j+m-1)]\}$, 论域 $X \in \{0, 1, \dots, 2^q - 1\}$ 为 $u(k), k \in [1, 2, \dots, N]$ 的取值空间, q 是混沌伪随机序列中每个序列值的二进制比特数,得到直积 $X \otimes X = \{(u(i), u(j)) \mid u \in X\}$, 那么 X 上的模糊关系 R 是这个直积的子集 $U = \{(u(i+k), u(j+k)) \mid u \in X, k \in [0, m-1]\}$ 上的模糊集,即 $R \in \mathcal{A}(U)$,

$$R: U \rightarrow [0, 1],$$

$R(u(i), u(j))$ 表示 $u(i)$ 与 $u(j)$ 具有 R 关系的程度,也称为 X 上的模糊关系(fuzzy relationship),其隶属函数为

$$R(u(i), u(j)) = 2^{-(1/|u(i)-u(j)|)^{1/r}} \quad (i, j = 1, 2, \dots, N - m + 1), \quad (10)$$

其中参数 r 为模糊关系的度量精度.模糊关系程度是一个相似性的“软度量”,其测度了两个集合之间具有 R 关系的程度,采用模糊隶属函数的形式对这种程度进行量化确定,因此能够较为准确的对其进行度量.

3.3. 基于模糊关系熵(F-REn)的混沌伪随机序列复杂度测度

根据(10)式的模糊关系概念,给出了基于 F-REn 的混沌伪随机序列复杂度测度方法.

定义4 令正整数 N 是混沌伪随机序列的观测长度,对于正实数 r 和正整数 $m, m \leq N - 2$. 给定非

负整数序列 $[u(1), u(2), \dots, u(N)]$, 其中, $u(i) \in \{0, 1, \dots, 2^r - 1\}$, $1 \leq i \leq N$, 形成 m 维非负整数序列空间 $s(1), s(2), \dots, s(N - m + 1) \in N^m$,

$$s^m(i) = [u(i), u(2), \dots, u(i + m - 1)] \quad (1 \leq i \leq N - m + 1). \quad (11)$$

根据(10)式给定模糊隶属函数, 可得到 m 维的第 i 个和第 j 个窗口向量之间的模糊关系度量

$$A_{ij}^m = \frac{1}{m} \sum_{k=0}^{m-1} R(u(i+k), u(j+k)) = \frac{1}{m} \sum_{k=0}^{m-1} 2^{-(|u(i+k) - u(j+k)|/r)^2}. \quad (12)$$

令不同窗口向量之间的距离为

$$d_{ij}^{m+1} = d[s^{m+1}(i), s^{m+1}(j)] = \max_{k=0,1,\dots,m} (|u(i+k) - u(j+k)|). \quad (13)$$

根据(10)式给定的模糊隶属函数, 可得到 $m+1$ 维的第 i 个和第 j 个窗口向量之间的模糊关系度量

$$B_{ij}^{m+1} = R(s^{m+1}(i), s^{m+1}(j)) = 2^{-(d_{ij}^{m+1}/r)^2}, \quad (14)$$

可得

$$C_i^m = (N - m)^{-1} \sum_{j=1, j \neq i}^{N-m+1} A_{ij}^m \quad (15)$$

和

$$C_i^{m+1} = (N - m - 1)^{-1} \sum_{j=1, j \neq i}^{N-m} B_{ij}^{m+1}. \quad (16)$$

可得

$$\Phi^m = (N - m + 1)^{-1} \sum_{i=1}^{N-m+1} C_i^m \quad (17)$$

和

$$\Phi^{m+1} = (N - m)^{-1} \sum_{i=1}^{N-m} C_i^{m+1}. \quad (18)$$

定义混沌伪随机序列的模糊关系熵(F-REn)为

$$S_{F-R}(m, r, N) = \ln \Phi^m(r) - \ln \Phi^{m+1}(r). \quad (19)$$

在(19)式中, m 为测量维度参数, 其确定了窗口向量的长度; N 为采样点参数, 其确定了待测序列的长度; r 为分辨率参数, 其确定了不同窗口向量之间距离的模糊关系度量精度.

根据(12)和(14)式可知, 模糊隶属函数 R 确定了不同窗口向量之间的模糊关系度量. 令 $d = |u(i+k) - u(j+k)|$, 则(10)式可写为

$$R(u(i), u(j)) = 2^{-(d/r)^2}. \quad (20)$$

根据(20)式可知, $R(u(i), u(j))$ 是 r 的增函数. 图 1 显示了(20)式构成的模糊隶属函数值(模糊关系程度)随着参数 r 变化的情况, 从图中可知,

随着参数 r 的增大, 相同窗口向量距离得到的模糊隶属函数值也增大, 因此, 可以进一步验证参数 r 确定了不同窗口向量之间距离被模糊化度量的精度. 例如, 当不同窗口向量之间距离 $d = 5$ 时, 在 $r = 2$ 的情况下, 该距离被模糊化度量为 0.01314, 而在 $r = 10$ 的情况下, 该距离被模糊化度量为 0.8409. 可以看出, 改变分辨率参数 r 能够使模糊化度量结果发生变化, 因此, 通过确定不同的分辨率参数能够灵活的调整模糊化度量的精度.

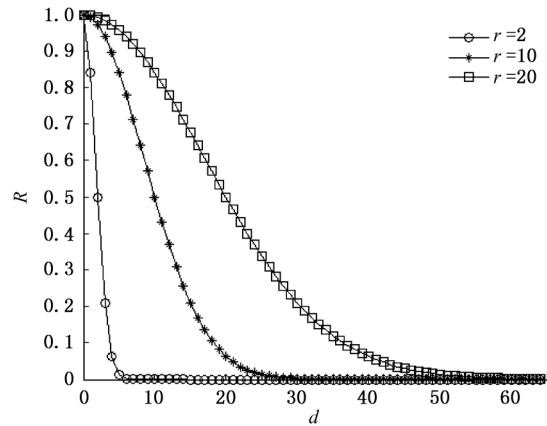


图 1 不同分辨率参数下模糊隶属函数值结果

因此, 可以看出, F-REn 是在 ApEn 测度方法中嵌入模糊度量模块, 使得统计测度结果更加的精确.

3.4. F-REn 测度的内涵分析

3.4.1. A_{ij}^m 的分析

根据(12)式可知, $u(i+k) - u(j+k)$ 是第 i 个窗口向量和第 j 个窗口向量相对应元素之间的差值, 所以 $2^{-(|u(i+k) - u(j+k)|/r)^2}$ 是定义 3 中确定的模糊隶属函数值, 其刻画了第 i 个和第 j 个窗口向量中相对应元素 $u(i)$ 与 $u(j)$ 所具有的模糊关系程度. 因此, 从(12)式可知, $A_{ij}^m = \frac{1}{m} \sum_{k=0}^{m-1} 2^{-(|u(i+k) - u(j+k)|/r)^2}$ 是第 i 个和第 j 个 m 维窗口向量对应元素模糊关系度量的平均值. 根据定义 4 可知, 在 F-REn 测度中将这个模糊关系度量的平均值作为 m 维窗口向量的第 i 个和第 j 个窗口向量之间的模糊关系度量值. 因此, A_{ij}^m 的内涵为第 i 个和第 j 个 m 维窗口向量之间的模糊关系度量值.

3.4.2. B_{ij}^{m+1} 的分析

从(13)式可知, d_{ij}^{m+1} 为第 i 个和第 j 个 $m+1$ 维窗口向量中对应元素差值的最大值. 因此, 从(14)式可知, B_{ij}^{m+1} 是该对应元素差值最大值的模糊隶属

函数值,也就是模糊关系度量值.根据定义 4 可知, B_{ij}^{m+1} 的内涵为第 i 个和第 j 个 $m+1$ 维窗口向量之间的模糊关系度量值.

3.4.3. Φ^m 和 Φ^{m+1} 的分析

根据(15)和(17)式可知, Φ^m 为序列中所有 m 维窗口向量之间模糊关系度量值的算术平均值.根据(16)和(18)式可知, Φ^{m+1} 为序列中所有 $m+1$ 维窗口向量之间模糊关系度量值的算术平均值.因此可知, Φ^m 内涵为序列中 m 维窗口向量之间相似程度的累计量,而 Φ^{m+1} 内涵为序列中 $m+1$ 维窗口向量之间相似程度的累计量.

因此,通过(19)式可知,F-REn 测度与 ApEn 测度相似,从统计学意义上描述了序列 m 维和 $m+1$ 维窗口向量之间相似的频繁程度的度量.但是不同之处在于,F-REn 测度采用了模糊关系的概念对窗口向量之间的相似程度进行了“软度量”,从而克服了 ApEn 对窗口向量之间相似程度“硬度量”带来的信息丢失问题,能够更加精确地完成混沌伪随机序列复杂度测度.

3.5. 基本性质证明

根据 FuzzyEn 的定义,分别得到 $r \rightarrow \infty$ 和待测序列为常数序列情况下的 2 个基本性质:

性质 1 对于任意混沌伪随机序列 $[u(1), \dots, u(N)]$, 其中 $u(i) \in \{0, 1, 2, \dots, 2^q - 1\}$, $1 \leq i \leq N$, r 为正实数, m 为正整数,且 $m \leq N - 2$, 有

$$\lim_{r \rightarrow \infty} S_{F-R}(m, r, N) = 0.$$

证明 根据 F-REn 定义可知

$$S_{F-R}(m, r, N) = \ln \Phi^m(r) - \ln \Phi^{m+1}(r). \quad (21)$$

将(17),(18)式代入(21)式中得到

$$S_{F-R}(m, r, N) = \ln \frac{(N - m + 1)^{-1} \sum_{i=1}^{N-m+1} C_i^m}{(N - m)^{-1} \sum_{i=1}^{N-m} C_i^{m+1}}. \quad (22)$$

我们能够进一步将(15),(16)式和(12),(14)式代入(22)式得到

$$S_{F-R}(m, r, N) = \ln \frac{[m(N - m + 1)]^{-1} \sum_{i=1}^{N-m+1} \sum_{j=1, j \neq i}^{N-m+1} \sum_{k=0}^{m-1} 2^{-(|u(i+k) - u(j+k)|/r)^2}}{(N - m - 1)^{-1} \sum_{i=1}^{N-m} \sum_{j=1, j \neq i}^{N-m} 2^{-(d_{ij}^{m+1}/r)^2}}. \quad (23)$$

对于任意确定的混沌伪随机序列, $|u(i+k) - u(j$

$+ k)|$ 均有界,因此有

$$\lim_{r \rightarrow \infty} 2^{-(|u(i+k) - u(j+k)|/r)^2} = 1 \quad (24)$$

和

$$\lim_{r \rightarrow \infty} 2^{-(d_{ij}^{m+1}/r)^2} = 1. \quad (25)$$

根据(23)式可得到

$$\lim_{r \rightarrow \infty} S_{F-R}(m, r, N) = \ln \frac{[m(N - m + 1)]^{-1} (N - m + 1) (N - m) m}{(N - m - 1)^{-1} (N - m) (N - m - 1)} = 0. \quad (26)$$

证毕.

性质 2 当被测序列为常数序列时, $S_{F-R}(m, r, N) = 0$.

证明 当被测序列为常数序列时,对于 $\forall i, j, k$, 都有 $|u(i+k) - u(j+k)| = 0$, 所以,

$$d_{ij}^{m+1} = \max_{k=0, 1, \dots, m} (|u(i+k) - u(j+k)|) = 0,$$

因此,根据(23)式可知

$$S_{F-R}(m, r, N) = \ln \{ [m(N - m + 1)]^{-1} \times (N - m + 1) (N - m) m \times \{ [(m - 1) (N - m - 1)]^{-1} \times (N - m) (N - m - 1) (m + 1) \}^{-1} \} = 0. \quad (27)$$

证毕.

4. 系统仿真与结果分析

为了对比分析方便,本节中将以文献[15,21]中采用的 logistic 映射和耦合映射格子(coupled map lattice)系统迭代产生的伪随机序列为例,验证说明 F-REn 作为混沌伪随机序列复杂度测度具有更加优越的对序列符号空间的适用性,更加小的对测量维度的敏感性和更加强的对分辨率的鲁棒性.

考虑 logistic 映射和耦合映射格子系统分别为

$$x_{n+1} = 4x_n(1 - x_n) \quad (0 \leq x_n \leq 1, n = 0, 1, 2, \dots) \quad (28)$$

和

$$y_{n+1}(1) = (1 - \varepsilon)f(y_n(1)) + \varepsilon g_n, \\ y_{n+1}(i) = (1 - \varepsilon)f(y_n(i)) + \varepsilon f(y_n(i + 1)) \quad (i = 2, \dots, 6; n = 0, 1, 2, \dots), \\ g_n = f(y_n(2)) \quad (0 \leq y_n \leq 1; n = 0, 1, 2, \dots), \quad (29)$$

其中边缘条件满足 $y_n(7) = y_n(1)$, 映射子系统

$f(x) = 4x(x - 1)$, 空间维数为 6, 耦合系数 $\varepsilon = 0.99$, 选择耦合信号 g_n 作为伪随机序列的生成函数.

在区间 $(0, 1)$ 上随机的选取迭代初始值进行

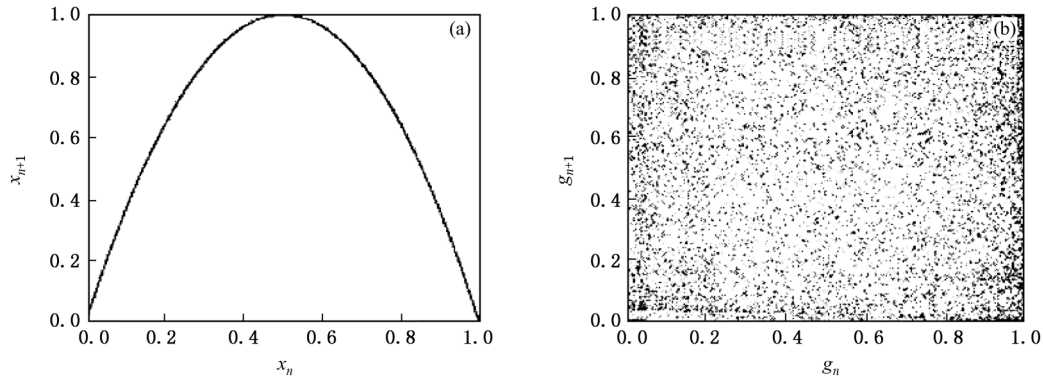


图2 混沌序列的相空间结构 (a) logistic 序列, (b) 耦合映射格子序列

采用如下判决公式进行量化:

$$\sigma_c(x) = j, \tag{30}$$

若

$$\sin^2 \left| \frac{j\pi}{2K} \right| < x < \sin^2 \left| \frac{(j+1)\pi}{2K} \right| \tag{31}$$

$(j = 0, 1, 2, \dots, K-1)$.

对 logistic 映射迭代和耦合映射格子系统产生的序列进行量化, 可以得到 $K = 2^q$ 进制的混沌伪随机序列 $[u_1, u_2, u_3, \dots]$, 其中 $u_i \in \{0, 1, 2, \dots, 2^q - 1\}$. q 为待测混沌伪随机序列值的 2 进制比特位数, 由于 $t = 2^q$, 故符号空间 t 越大, 所需要的 2 进制比特位数越多, 因此, 本文将 q 也作为符号空间参数.

通过上述说明可知, 符号空间参数 q 、测量维度参数 m 和分辨率参数 r 对序列复杂度测度的刻画具有重要意义. 对于符号空间参数, q 的增加将会增大待测序列的变化范围, 需要提高复杂度测度精度才能够得到准确的序列复杂程度刻画结果, 反之亦然. 对于测量维度, m 的增加将会增大测度的基本比较衡量单位(窗口向量)的长度, 随着窗口向量的变化, 该待测序列在不同测度方法中表现出来的复杂程度也会随之发生变化. 例如, 在 ApEn 测度中, (6) 式表示了完成准确测度条件下, 测量维度和采样点数的关系. 对于分辨率参数, r 的变化能够实现对测度精度的调节, 但是在某些测度方法中, 由于测度算法本身的缺陷, 会造成分辨率参数变化导致的测度错误, 例如, ApEn 测度就会出现这样的情况.

从上述阐述可知, 这些参数的变化都会影响到混沌伪随机序列复杂度测度的结果, 为了进一步验

证本文提出 F-REn 测度方法的性能, 针对上述两种已知复杂度的混沌伪随机序列, 分别采用 ApEn、符号熵和 F-REn 方法进行测度仿真, 分析和说明 F-REn 测度方法在对序列符号空间的适用性、对测量维度参数的敏感性和对分辨率参数的鲁棒性方面的特性.

4.1. 对序列符号空间的适用性

序列符号空间是指待测序列可选项元素集的大小, 例如, 对于二进制序列来说, 其符号空间为 2. 在实际序列复杂度测度过程中, 我们总是希望一种测度方法能够准确的测度任意符号空间大小混沌伪随机序列的复杂度. 因此, 本文将测度方法对序列符号空间的适用性定义为: 具有不同符号空间大小的混沌伪随机序列对序列复杂度测度结果的影响程度. 这种影响越大序列复杂度测度方法对序列符号空间大小的适用性越差. 反之, 测度方法对序列符号空间的适用性越好, 该测度方法能够完成复杂度准确测度的序列符号空间范围越大.

测度方法对序列符号空间的适用性存在两个问题: 1) 待测序列的符号空间是不是测度方法必须已知的信息? 2) 序列的符号空间大小变化对测度结果有何影响?

对于第一个问题, 从 (8) 和 (9) 式可知, 要确定符号熵值, 必须确定待测度序列的参数 l_x , 而由 (7) 式可知, 要确定参数 l_x 必须已知符号空间大小 t , 由于 $t = 2^q$, 即必须已知 q . 因此, 在符号熵方法中, 待测度序列的符号空间 q 作为先验知识, 必须是测度

之前的已知信息. 而从 ApEn 和 F-REn 测度方法的定义可知, 对于这两种测度方法来说, 并不需要已知待测序列的符号空间大小, 因此可以对任意未知符号空间大小的混沌伪随机序列进行复杂度测度.

对于第二个问题, 不同的测度方法对不同符号空间大小序列的测度适用性不同, 表 1 为序列符号

空间参数 q 变化情况下, ApEn, 符号熵和 F-REn 测度方法对 logistic 映射序列和耦合映射格子序列的测度结果, 测度中 ApEn, 符号熵和 F-REn 方法测度参数选取为: 序列长度 $N = 1000$, 测度分辨率参数 $r = 1$ (符号熵方法没有参数 r), 测量维度参数 $m = 3$. 本文对 3 种序列复杂度测度结果进行仿真.

表 1 3 种序列复杂度测度方法对不同符号空间大小(q)序列的测度结果

q	S_{Ap}		S_{symb}		S_{F-R}	
	logistic 映射	耦合映射格子	logistic 映射	耦合映射格子	logistic 映射	耦合映射格子
4	0.682	0.8313	1.3806	2.2328	1.8779	4.6015
5	0.6655	0.1636	1.5986	2.2954	2.0068	5.9713
6	0.6222	0.0198	1.8043	2.3005	2.0271	6.6682
7	0.5410	0.0003	1.9836	2.3019	2.0273	7.4025
8	0.3815	0.0004	2.1079	2.3019	1.9879	6.8923
9	0.2510	-0.001	溢出	溢出	2.09651	7.1681
10	0.1485	-0.001	溢出	溢出	2.104825	6.8162
11	0.0699	-0.001	溢出	溢出	2.269221	6.7940
12	0.0356	-0.001	溢出	溢出	2.1815	8.0007

参照图 2 中所示的序列复杂度情况可知, 表 1 中复杂度测度结果出现了错误或者数据溢出. 在 ApEn 测度方法中, 仅仅当 $q = 4$ 时, 两种序列的复杂度能够被正确测度, 当 $q > 4$ 时出现耦合映射格子序列测度值小于 logistic 映射序列测度值的情况, 这与实际序列复杂度相反, 因此无法正确进行混沌伪随机序列复杂度的正确测度. 在符号熵测度方法中, 当 $q = 4$ 时, 序列复杂度能够被正确测度, 但是随着 q 的不断增大, 由于序列状态空间较大, 造成符号熵值计算量急剧增大, 仿真结果出现数据溢出现象. 因此在该仿真参数条件下, 符号熵测度方法对 $q \geq 9$ 的混沌伪随机序列的复杂度太大, 实用性不强. 在 F-REn 测度方法中, 随着参数 q 的增大, 两种序列的测度值结果始终与其实际序列复杂度相匹配, 因此, 序列符号空间大小对 F-REn 测度方法影响很小, 该测度方法具有很好的序列适用性.

4.2. 对测量维度的敏感性

测量维度参数 m 是指待测序列的窗口向量长度. 由于在 ApEn、符号熵和 F-REn 测度方法中测量维度参数都是确定序列复杂度的一个重要参数, 其变化会对最终测度结果产生影响. 在实际测度中, 我们总是希望测量维度参数的变化对最终测度结果的影响越小越好, 所以测度方法对测量维度参数

的变化越敏感, 其出现测度误差和错误的可能性就越大. 也就是说, 我们总是希望测度方法对测量维度具有较低的敏感性.

为了进一步说明本文提出的 F-REn 测度方法的优越性, 图 3 表示了随着测量维度参数 m 的变化, 3 种不同的测度方法得到的测度结果. 仿真中参数选取为: $N = 2000, r = 1$ (符号熵没有参数 r). 其中, 图 3(a)–(c) 为 $q = 6$ 情况下, 3 种测度方法随着 m 变化所得到的测度结果, 图 3(d)–(f) 为 $q = 12$ 情况下, 3 种测度方法随着 m 变化所得到的测度结果.

从图 3(b) 中可以看出, 对 $q = 6$ 的序列进行复杂度测度时, 在 ApEn 测度中当 $m > 2$ 时就出现复杂度测度错误的现象. 从文献 [15] 的 (14) 式可知, 由于 q 增大, 能够准确测度的 m 取值进一步变小. 因此, 如图 3(e) 所示, 对于 $q = 12$ 的序列进行复杂度测度时, 只有 $m = 1$ 的情况下能够正确测度两种序列的复杂度.

如图 3(c) 和 (f) 所示, 当 $m = 1$ 时, 符号熵测度中的不同窗口向量只有 t 种, 即符号状态空间只有 t 个状态, 且窗口向量总数为序列长度 N , 因此各种窗口向量接近均匀分布, 符号熵方法不能够有效的测度混沌伪随机序列复杂度. 当 $m \neq 1$ 时, 在 $q = 6$ 情况下, 如图 3(c) 所示, 与 ApEn 测度相比, 符号熵测度能够很好的完成序列复杂度的测度, 但是随着

m 的逐渐增大,由于计算量增加,符号熵测度仿真出现数据溢出,当 $m > 4$ 时由于溢出无法得到测度值. 在 $q = 12$ 的情况下,如图 3(f)所示,由于序列状态空间进一步增大,导致无法实现序列复杂度的准确测度. $m = 2$ 时两种序列的符号熵复杂度测度值非常接近,出现混淆; $m > 2$ 时符号熵测度仿真出现数据溢出.

如图 3(a)和(d)所示,而对于 F-REn 测度来说,不仅对 $q = 6$ 的序列,在 $m \in [1, 10]$ 范围内均能够完成复杂度的准确测度,而且对于 $q = 12$ 的大符号空间序列,在 $m \in [1, 8]$ 范围内也能够完成复杂度的准确测度. 因此, F-REn 测度方法对测量维度参数的敏感性非常小,测量维度参数的变化对最终测度结果的影响很小.

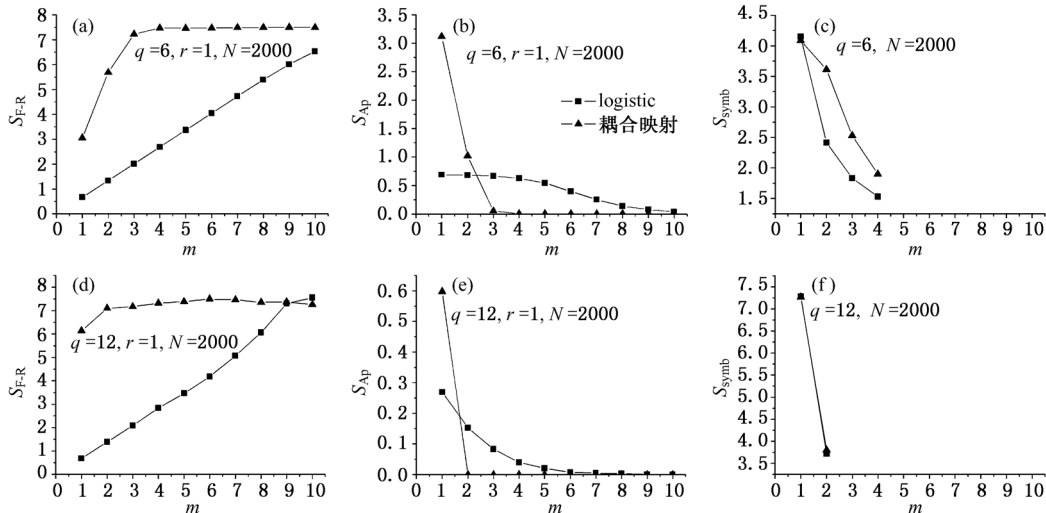


图 3 3 种不同测度方法的测度值随测量维度参数变化情况 (a) $q = 6$ 时 F-REn 测度, (b) $q = 6$ 时 ApEn 测度, (c) $q = 6$ 时符号熵测度, (d) $q = 12$ 时 F-REn 测度, (e) $q = 12$ 时 ApEn 测度, (f) $q = 12$ 时符号熵测度

4.3. 对分辨率的鲁棒性

由于符号熵测度方法中没有分辨率参数,所以在分辨率参数鲁棒性的仿真中仅对 ApEn 测度和 F-REn 测度进行对比分析.

在 ApEn 和 F-REn 测度方法中,分辨率性能是衡量待测序列不同截取窗口之间差异的门限水平,它确定了复杂度测度过程的精度.

从文献[15]可知,在基于 ApEn 的混沌伪随机序列复杂度测度中,分辨率参数通过将序列窗口向量之间的距离门限设定为 r ,从而实现了复杂度测度过程的精度调整.但是在实际测度过程中,总是希望通过改变分辨率参数灵活地调整测量结果的精度.但是在此基础上,更为重要的是,在不同测度精度条件下都能够正确地测度出混沌伪随机序列的复杂程度.也就是说,复杂度测度方法需要具有对分辨率参数的鲁棒性.由于本文提出的基于 F-REn 的混沌伪随机序列复杂度测度,采用了模糊

理论思想衡量不同窗口向量之间的距离.因此,能够进一步降低分辨率参数对测度结果的影响,提高复杂度测度中对分辨率的鲁棒性.

图 4 为 ApEn 和 F-REn 测度方法的测度值随着分辨率参数变化情况.其中从图 4(b)中可知, ApEn 测度方法在参数 r 取值过小会出现测度错误,而从图 4(a)可知, F-REn 测度结果受参数 r 的变化影响很小,没有出现测度错误的情况.随着待测度序列采样点数的降低,由于统计样本点减少,造成测度结果的出现不准确情况.从图 4(d)中可以看出,在待测序列采样点数 N 从 2000 降低到 500 的情况下, ApEn 测度方法能够准确测度的分辨率参数选取范围进一步缩小,在 $r < 2$ 的情况下不能够正确测度;而从图 4(c)可知,尽管采样点数降低了,但是在 F-REn 测度方法中分辨率参数的变化并没有引起测度结果的错误.因此,可以看出,与 ApEn 测度相比, F-REn 测度对分辨率参数变化的具有很好的鲁棒性.

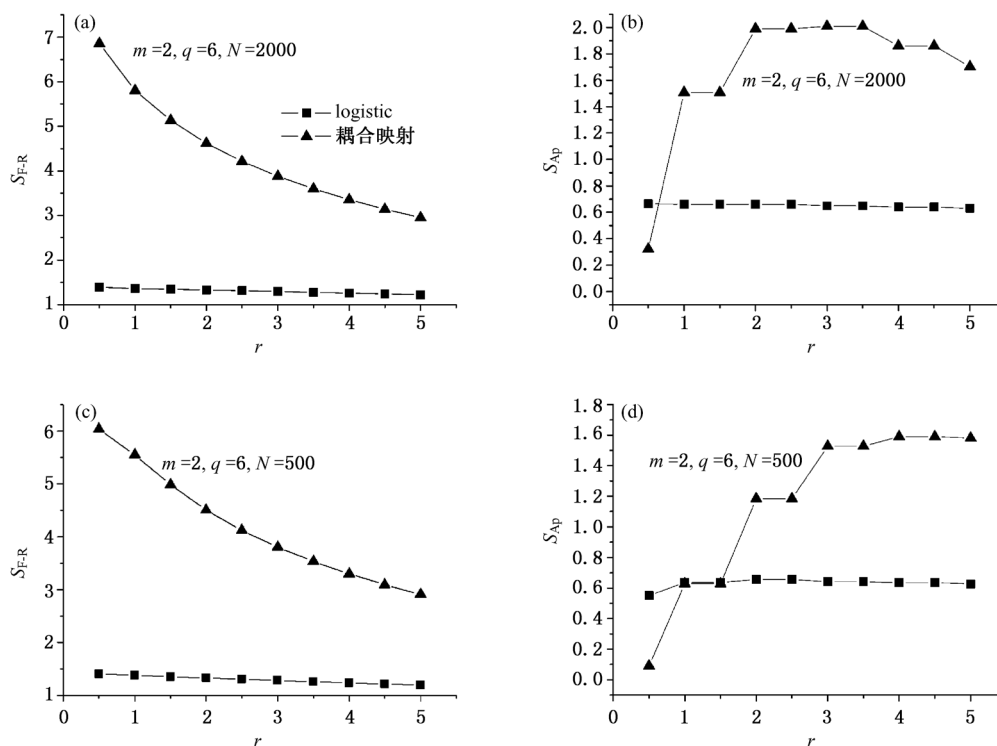


图4 ApEn 和 F-REn 测度值随着分辨率参数变化情况 (a) $N = 2000$ 时 F-REn 测度, (b) $N = 2000$ 时 ApEn 测度, (c) $N = 500$ 时 F-REn 测度, (d) $N = 500$ 时 ApEn 测度

5. 结 论

提出了一种新的基于模糊关系熵(F-REn)分析混沌伪随机序列复杂度的方法. 根据模糊理论思想构造了混沌伪随机序列复杂度测度的模糊隶属函数,并基于该隶属函数提出了用于混沌伪随机序列

复杂度测度的模糊关系熵的定义. 将 F-REn 测度、ApEn 测度和符号熵测度进行了性能比较和分析. 在相同数据要求条件下, 仿真结果显示, 与其他测度方法相比, F-REn 测度具有更加好的对序列符号空间的适用性、更加小的对测量维度的敏感性和更加强对分辨率参数的鲁棒性.

[1] Heidari-Bateni G, McGillem C D 1994 *IEEE Trans. Commun.* **42** 1524

[2] Heidari-Bateni G, McGillem C D 1992 *Wireless Communications 1992 Conference Proceedings* p437

[3] Zan Li, Jueping Cai, Xiaojun Chen, Xiaofeng Lu 2009 *Wireless Communications and Networking Conference*, 2009 IEEE, (4) 1—5

[4] Li Z, Cai J P, Lu X F, Si J B 2009 *Communications*, 2009. *ICC 09. IEEE International Conference on* (6) p1—5

[5] Li Z, Cai J P, Chang Y L 2009 *IEEE Trans. Commun.* **57** 812

[6] Kolmogorov A N 1965 *Problems of Information Transmission* **1** 4

[7] Lempel A, Ziv J 1976 *IEEE Trans. Inform. Theory* **22** 75

[8] Ruepple R A 1984 *Ph. D. Dissertation* (Zurich: Swiss Federal Institute of Technology)

[9] Xiao G Z, Wei S M, Lam K Y, Imamura K 2000 *IEEE Trans. Inform. Theory* **46** 2203

[10] Kolokotronis N, Kalouptsidis N 2003 *IEEE Trans. Inform. Theory* **49** 3047

[11] Griffin F, Shparlinski I E 2000 *IEEE Trans. Inform. Theory* **46** 2159

[12] Gutierrez J, Shparlinski I E, Winterhof A 2003 *IEEE Trans. Inform. Theory* **49** 60

[13] Kurosawa K, Sato F, Sakata T, Kishimoto W 2000 *IEEE Trans. Inform. Theory* **46** 694

[14] Lauder A G B, Paterson K G 2003 *IEEE Trans. Inform. Theory* **49** 273

[15] Cai J P, Li Z, Song W T 2003 *Acta Phys. Sin.* **52** 1871 (in Chinese) [蔡觉平、李 赞、宋文涛 2003 物理学报 **52** 1871]

[16] Pincus S M 1995 *Chaos* **5** 110

[17] Zhuang J J, Ning X B, Zou M, Sun B, Yang X 2008 *Acta Phys.*

- Sin.* **57** 2805 (in Chinese) [庄建军、宁新宝、邹 鸣、孙 飙、杨 希 2008 物理学报 **57** 2805]
- [18] Xie Y, Xu J X, Yang H J, Hu S J 2002 *Acta Phys. Sin.* **51** 205 (in Chinese) [谢 勇、徐健学、杨红军、胡三觉 2002 物理学报 **51** 205]
- [19] Wang G Q, Zhang Z P 2008 *Acta Phys. Sin.* **57** 1976 (in Chinese) [王启光、张增平 2008 物理学报 **57** 1976]
- [20] Cao B, Lü X Q, Zeng M, Wang Z M, Hang S S 2006 *Acta Phys. Sin.* **55** 1696 (in Chinese) [曹 彪、吕小青、曾 敏、王振民、黄石生 2006 物理学报 **55** 1696]
- [21] Xiao F H, Yan G R, Han Y H 2004 *Acta Phys. Sin.* **53** 1871 (in Chinese) [肖方红、阎桂荣、韩宇航 2004 物理学报 **53** 2877]
- [22] Azad R K, Rao J S, Ramaswamy R 2002 *Chaos, Soliton. Fract.* **14** 633
- [23] Liu X F, Yu W L 2008 *Acta Phys. Sin.* **57** 2587 (in Chinese) [刘小峰、俞文莉 2008 物理学报 **57** 2587]
- [24] Hu B Q 2004 *Fundamental of Fuzzy Theory* (Wuhan: Wuhan University Press) Chapt1—3 (in Chinese) [胡宝清 2004 模糊理论基础(武汉: 武汉大学出版社)第1—3章]
- [25] Zadeh L A 1965 *Information and Control* **8** 338
- [26] Chen W, Wang Z Z, Xie H B, Yu W X 2007 *IEEE Trans. Neural Syst. Rehabil. Eng.* **15** 266
- [27] Chen W T, Zhuang J, Yu W X, Wang Z Z 2009 *Med. Eng. Phys.* **31** 61
- [28] Ornstein D, Weiss B 1990 *Ann. Probab.* **18** 905

New complexity metric of chaotic pseudorandom sequences using fuzzy relationship entropy*

Chen Xiao-Jun^{1)†} Li Zan¹⁾ Bai Bao-Ming¹⁾ Cai Jue-Ping²⁾

1) (State Key Laboratory of Integrated Service Networks (ISN), Xidian University, Xi'an 710071, China)

2) (School of Microelectronics, Xidian University, Xi'an 710071, China)

(Received 12 June 2010; revised manuscript received 1 March 2011)

Abstract

A new complexity metric to evaluate the unpredictability of the chaotic pseudorandom sequences based on the fuzzy relationship entropy (F-REn) is proposed in the view of maximal randomness of the sequences with arbitrary length. On this condition, two basic properties of F-REn are proved. Simulations and analysis results show that, the proposed F-REn works effectively to discern the changing complexities of the chaotic pseudorandom sequences, and compared with complexity metric based on the approximate entropy (ApEn) and symbolic dynamics approach, F-REn works have obvious advantages in the applicability of symbolic space, the sensitivity of vector dimension and the robustness of resolution parameter.

Keywords: chaotic pseudorandom sequences, fuzzy theory, complexity.

PACS: 42. 65. Sf

* Project supported by the National Science and Technology Major Project of the Ministry of Science and Technology of China (Grant No. 2010ZX03006-002-04), the National High Technology Research and Development Program of China (Grant No. 2009AA01Z237), the National Natural Science Foundation of China (Grant No. 60706027), the Program for New Century Excellent Talents in University of Ministry of Education of China (Grant No. NCET-07-0653), the Programme of Introducing Talents of Discipline to Universities of China (Grant No. B08038), the Program for Changjiang Scholars and Innovative Research Team in University of Ministry of Education of China (Grant No. IRT0852), and the Natural Science Foundation of Shaanxi Province, China (Grant No. SJ08F09).

† E-mail: xjchen@mail.xidian.edu.cn