数据的多尺度表达及一类正交变换的构造*

丁 玮[1] 齐东旭[1,2]

[1](中国科学院计算技术研究所CAD开放研究实验室 北京 100080)

[2](北方工业大学CAD研究中心 北京 100041)

E-mail: J. Ding@writeme.com QiDongxu@unet.net.cn

摘要 着重研究基于{0,1}码变量表示的函数描述问题.首先,举例说明在多尺度分析观点下,信号处理中某些函数定义的新方法;进而提出"坐标分量分离方法",并由此发现了平面区域上点集坐标数值的自相似结构;在揭示了如此自相似结构的基础上,构造了高维单纯形上的正交完备的二值函数系统;丰富与推广了Rademacher函数及Walsh函数的理论及应用.

关键词 多尺度分析,正交变换,坐标分量分离.

中图法分类号 TP391

信息以数据形式表达,从而便于记录、存储、处理与传输,数据又以数的表达方式为根本.关于数的本质的认识、数的表达方式等成果展现了人类认识自然的漫长而艰难的历程.

在古典分析中讨论的函数,并不强调其自变量取值以何种记数形式.在不特别申明时都约定俗成地认为十进制是自然而方便的记数方式.然而,当我们在数字计算机上作信号处理时,几乎任何数据皆转化成为{0,1}码.将信息与{0,1}码联系起来,信息处理转化为{0,1}码的操作,这就是数字化技术的核心内容.[1]

对于任意的非负整数 p,总可以找到 $e_0\in\{0,1,\dots,a-1\}$,使 $p=p_1a+e_0$,有如此性质的最小的数字符号集合 $S=\{0,1,\dots,a-1\}$ 称为模 a 的完全剩余类.据此,经有限步得到

$$p=e_ka^k+e_{k-1}a^{k-1}+\dots+e_1a+e_0,$$

其中 a 称为基,p 亦可以简记为

$$p=(e_ke_{k-1}\dots e_1e_0).$$

通常,当取 $a=2$ 时,$e_j\in S=\{0,1\}$;$a=3$ 时,$e_j\in S=\{0,1,2\}$ 或 $\{0,1,-1\}$;$a=10$ 时,$e_j\in S=\{0,1,\dots,9\}$.实际上,取 a 为负整数也是可以的,当取负整数 a 作为基时,用 $|a|$ 个数字符号 $\{0,1,\dots,|a|-1\}$ 就可以把所有整数表示出来.值得注意的是,基的选取与相应的数字符号集合 S 中元素的个数之间并不是必然相关的.打破这个限制,我们可以首先认为数字符号集合 S,因数字化技术的方便,仅有两个元素,即 $S=\{0,1\}$.至于基,它可以相当宽松,甚至可以取其为虚数.[2]这样一来,本文讨论的{0,1}码技术,并非仅指二进制技术,虽然二进制技术仍然是它的主要内容.

我们讨论的变量 t,常常写成如下形式:如果 $0\leqslant t<1$,则

$$t=0.e_0e_1e_2\dots e_k\dots,\quad e_j\in\{0,1\},\ j=0,1,2,\dots \tag{1}$$

本文着重研究式(1)基于{0,1}码变量表示的函数描述问题,首先举例说明在多尺度分析观点之下,信号处理中某些函数定义的新方法;进而提出"坐标分量分离方法",并由此发现了平面区域上点集坐标数值的自相似结构;进而,在揭示了自相似结构的基础上,构造了高维单纯形上的正交完备的二值函数系统;丰富与推广了Rademacher函数及Walsh函数的理论及应用.

1 函数按自变量{0,1}码表示的构造性定义

用{0,1}码的变量表达形式(1),可以定义各种各样的复杂函数,这样给出的许多种函数在数字信号处理中有特殊的用途.

---

* 本文研究得到国家自然科学基金资助.作者丁玮,1971年生,博士生,主要研究领域为计算机辅助几何设计,计算机图形学,图象处理.齐东旭,1940年生,教授,博士导师,主要研究领域为计算机辅助几何设计,计算机图形学,数值分析,图象处理.

本文通讯联系人:丁玮,北京100080,中国科学院计算技术研究所CAD开放研究实验室

本文1998-02-28收到原稿,1998-04-20收到修改稿

(1) $f_0(t)=1, f_k(t)=e_k,\ k=1,2,\ldots$

实际上，$f_k(t)=(R_k(t)+1)/2, k=1,2,\ldots$，其中 $R_k(t)$ 为 Rademacher 函数. 有时，$\{f_k(t)\}$ 在应用中更直接方便.

(2) $g_0(t)=e_0, g_k(t)=e_{k-1}\oplus e_k,\ k=1,2,\ldots$

其中 $\oplus$ 表示模 2 加法.

(3) $h_0(t)=e_0, h_k(t)=e_0\oplus e_1\oplus\ldots\oplus e_k, k=1,2,\ldots$

$h_k(t)$ 可以表示成 Walsh 函数的线性组合，例如，$h_k(t)=(Wal_j(t)+1)/2$，k 与 j 的对应关系与 Walsh 函数的次序排列规定有关.

(4) $l_0(t)=e_0, l_k(t)=\frac{1}{k+1}(e_0\oplus e_1\oplus\ldots\oplus e_k), k=1,2,\ldots$

图 1 为前 4 个函数的示意图.

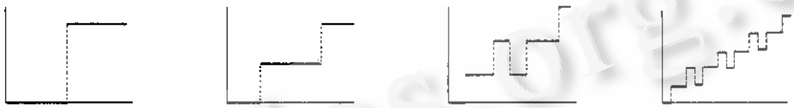


图1　一个多尺度函数的自相似结构示意图

(5) 著名的 Rademacher 函数可以表示为 $R_n(t)=\mathrm{sgn}(\sin 2^n\pi t)$，其中 sgn 表示符号函数. 它亦可表示为

$$R_0(t)=1, 0\leqslant t<1$$

$$R_1(t)=\begin{cases}1, & 0\leqslant t<\frac{1}{2}\\ -1, & \frac{1}{2}\leqslant t<1\end{cases}$$

$$R_n(t)=R_n(t+1), R_n(t)=R_1(2^{n-1}t).$$

如果注意变量 t 的 $\{0,1\}$ 码表示方式(1)，则 Rademacher 函数可以表述成[3,4]

$$R_0(t)=1, R_k(t)=(-1)\uparrow e_{k-1},\quad t\in[0,1), k=1,2,\ldots$$

这里 $A\uparrow\alpha$ 表示 A^α(下同).

在上面所举的典型例子中，我们强调的共同之处在于自变量的表达形式(1). 由于 $\{e_k\}$ 反映了二值结构和尺度特性，因而它在数字信号处理中有特别的优越性.

2　重心坐标系之下多尺度细化及坐标分量分离方法

[0,1]区间上的二值正交函数在信号处理中有重要的应用. 对高维情形，张量积型区域上的正交函数族的定义可以从一维情况类比推广过去.[5]但是，对单纯形区域上正交函数族的定义，却不是简单类比能够完成的. 以平面三角域上的 Walsh 函数族的定义来说，首见于文献[6]，它给出了 Hadamard 次序的定义. 之后，关于 Paley 次序的定义在文献[7,8]中给出. 本文继续研究这一问题，强调{0,1}码表示函数及式(1)的作用.

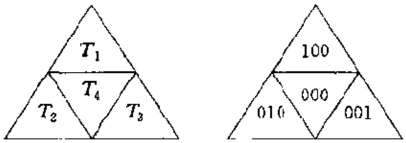
我们采用重心坐标. 为此，令三角形区域 T 为坐标三角形，对任意 $P\in T$，记 P 的重心(面积)坐标为 (u,v,w)，$u\geqslant 0, v\geqslant 0, w\geqslant 0, u+v+w=1$. 将 3 个坐标的数值写成{0,1}码，这里是二进制形式

$$u=0.u_0u_1u_2\ldots,\quad v=0.v_0v_1v_2\ldots,\quad w=0.w_0w_1w_2 \tag{2}$$

这里，$u_j, v_j, w_j\in\{0,1\}$.

下面我们考虑不同尺度之下平面上点的所谓坐标分离问题. 这里所说的坐标分离是指，我们在表达形式(2)中首先取 u_0, v_0, w_0，由它们作成的三元组 (u_0, v_0, w_0) 可以认为是 (u,v,w) 的第 1 次近似，简写成紧凑形式 $u_0v_0w_0$；接着取 u_1, v_1, w_1，将三元组 (u_1, v_1, w_1) 写成紧凑形式 $u_1v_1w_1$；依次类推.

对于坐标三角形上任意一个点 $P=(u,v,w)$ 来说，一旦取定之后，其由形式(2)给出的 $u_jv_jw_j$ 都是确定的 0 与 1 的排列. 且根据 P 在坐标三角形 T 上的位置不同，$u_jv_jw_j$ 的排列也有所不同. 但若将 T 做相似的子分割，则可以发现在子区域上的 $u_jv_jw_j$ 分布的规律. 为说明这一规律，首先连结 T 的各边中点，得 4 个小三角形，记为 T_1, T_2, T_3, T_4(如图 2(a)所示). 在这个分割之下，当 $P\in T_1$ 时，$u_0v_0w_0=100$；当 $P\in T_2$ 时，$u_0v_0w_0=010$；当 $P\in T_3$ 时，$u_0v_0w_0=001$；当 $P\in T_4$ 时，$u_0v_0w_0=000$(如图 2(b)所示).

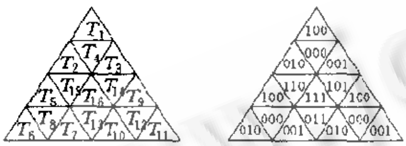


(a)　　(b)

图2　三角形区域第1次子分割的 $u_0v_0w_0$ 分布

当我们考察 $u_1v_1w_1$ 的分布时，将在第 1 次子分割的基础上作第 2 次分割. 容易验证，在第 2 次子分割下，相应于图 3(a)所示分割次序的 $u_1v_1w_1$ 分布为如图 3(b)所示.

进而观察 $u_2v_2w_2$ 分布，这时 T 被分割为 64 个子区域. 将图 3(b)的分布填在图 2(a)中 T_1,T_2,T_3 所示区域上，而在 T_4 上呈现图 3(b)中数值的反码. 一般说来，在第 j 次分割之下，有 $u_{j-1}v_{j-1}w_{j-1}$ 的分布. 为了得到在第 j 次分割之基础上的 $u_jv_jw_j$ 分布，要对第 j 次分割再作细化，再把 $u_{j-1}v_{j-1}w_{j-1}$ 的分布数据填到如图 2(a)所示的 T_1,T_2,T_3 上，而 T_4 则是 $u_{j-1}v_{j-1}w_{j-1}$ 分布数值的反码(当 $j=2$ 时，如图 4 所示).



(a)　　(b)

图3　三角形区域第2次子分割的 $u_1v_1w_1$ 分布

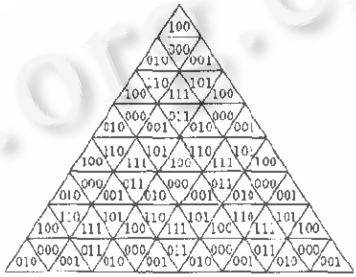


图4　三角形区域第3次子分割的 $u_2v_2w_2$ 分布

3　图象处理中一类正交二值函数新的构造性定义

由前述的坐标分离方式，利用表达式(2)，我们容易定义三角域上的 Rademacher 函数. 为此，引入记号

$$\xi_j=(-1)\uparrow u_j,\eta_j=(-1)\uparrow u_j,\zeta_j=(-1)\uparrow w_j,$$

这样，Rademacher 函数可以表示为

$$R_0(u,v,w)=1\ ,$$

$$R_1(u,v,w)=\eta_0\zeta_0,R_2(u,v,w)=\xi_0\zeta_0,R_3(u,v,w)=\xi_0\eta_0.$$

如果 $j(m)$ 表示 j 之后连续写 m 个零的数码序列，那么，对于 $m=0,1,2,\ldots$，有

$$R_{1(m)}(u,v,m)=\eta_m\zeta_m,R_{2(m)}(u,v,w)=\xi_m\zeta_m,R_{3(m)}(u,v,w)=\xi_m\eta_m.$$

对 $m=0,1$ 的情形如图 5 所示. 这里，黑色表示该区域上函数值定义取 1，白色表示函数值取 0.

R_0　R_1　R_2　R_3　R_{10}　R_{20}　R_{30}

图 5　三角域上的 Rademacher 函数

下面定义三角域上的 Walsh 函数. 所谓 Walsh 函数，是由可数无穷多个函数组成的函数族. 但如何对这个函数族中的函数进行排序，是理论研究与实际应用都十分关心的问题. 最早在文献[6]中通过 Hadamard 矩阵给出了三角域上的 Walsh 函数定义，那便是所谓 Hadamard 次序的 Walsh 函数. 它的生成过程如下.

令
$$H_2=\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix},H_{2^{k+1}}=H_{2^k}\otimes H_2,k=1,2,\ldots$$

其中⊗表示矩阵的 Kronecker 乘积. 按取定的 2^k 阶 Hadamard 矩阵的任意行(列)，将 1 或 −1 向三角域子分割的每个子域，以图 2(a)($k=2$)或图 3(a)($k=4$)所示的次序做分配. 从而完成三角域上 Hadamard 次序的 Walsh 函数的定义. 问题是，相应于单变量情形，Paley 次序如何向三角域推广. 我们根据前面 Rademacher 函数的表达结果，可以利用 Rademacher 函数写出 Paley 次序的 Walsh 函数. 为此，对任意非负整数 n，把它写成四进制形式

$$n=(\ldots n_pn_{p-1}\ldots n_2n_1),n_j\in\{0,1,2,3\}$$

注意十进制与四进制的对应关系

1,	2,	3,	4,	5,	6,	7,	8,	9,	…,	20,	21,	22,	23,	24,	25,	…
1,	2,	3,	10,	11,	12,	13,	20,	21,	…,	110,	111,	112,	113,	120,	121,	…

得出三角域上 Paley 次序的 Walsh 函数 $\{W_k=W_k(u,v,w),k=0,1,2,\ldots\}$，列表如下.

表 1　三角域上用 Rademacher 函数表示的 Walsh 函数

W_0	W_1	W_2	W_3	W_4	W_5	W_6	W_7	W_8	W_9	W_{10}	W_{11}	W_{12}	W_{13}
R_0	R_1	R_2	R_3	R_{10}	$R_{10}R_1$	$R_{10}R_2$	$R_{10}R_3$	R_{20}	$R_{20}R_1$	$R_{20}R_2$	$R_{20}R_3$	R_{30}	$R_{30}R_1$

W_{14}	W_{15}	W_{16}	W_{17}	W_{18}	W_{19}	W_{20}	W_{21}	W_{22}	W_{11}
$R_{30}R_2$	$R_{30}R_3$	R_{100}	$R_{100}R_1$	$R_{100}R_2$	$R_{100}R_3$	$R_{100}R_{10}$	$R_{100}R_{10}R_1$	$R_{100}R_{10}R_2$	$R_{100}R_{10}R_3$

……

它的一般表达式为

$$W_n(u,v,w)=\prod_{i=j}^{1}R_{n_i(i-1)}(u,v,w),\quad n=0,1,2,\ldots$$

其中 $n=(n_jn_{j-1}\ldots n_2n_1)_4$，$n_i=\{0,1,2,3\}$，$n_j\neq 0$ 且当 $i>j$ 时，$n_i=0$. 它的前 16 个函数示于图 6 中.

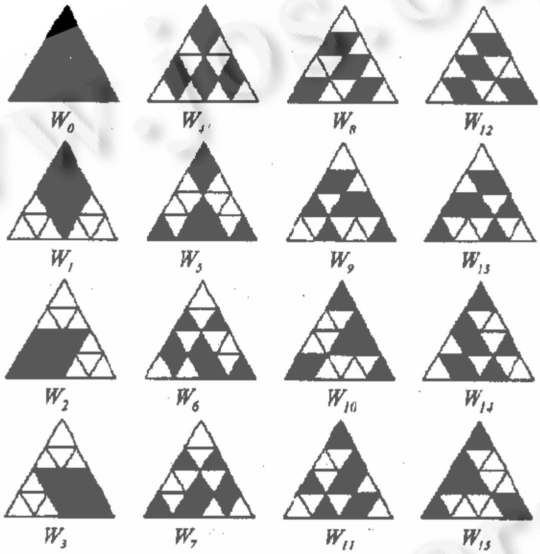


图 6　三角域上的 Paley 次序的前 16 个 Walsh 函数

参考文献

1　尼葛洛庞帝. 数字化生存. 海口：海南出版社，1997
(Negroponte N. Being Digital. Haikou: Hainan Publishing House, 1997)

2　Katai I, Szabo J. Canonical number system for complex integers. Acta Scientiarum Mathematicarum, 1975,37:255～260

3　柳重堪. 正交函数及其应用. 北京：国防工业出版社，1982
(Liu Chong-kan. Orthogonal Function and Its Application. Beijing: National Defense Industry Publishing House, 1982)

4　郑维行，苏维宜，任福贤. 沃尔什函数理论与应用. 上海：上海科技出版社，1983
(Zheng Wei-xing, Su Wei-yi, Ren Fu-xian. Walsh Function Theory and Its Application. Shanghai: Shanghai Science and Technology Publishing House, 1983)

5　哈尔姆斯. 序率理论基础与应用. 北京：人民邮电出版社，1980
(Harmuth H F. Sequency Theory Foundations and Applications. Beijing: People's Post and Telecommunications Publishing House, 1980)

6　Feng Yu-yu, Qi Dong-xu. On the Haar and Walsh system on a triangle. Journal of Computational Mathematics, 1983,1(3): 223～232

7　齐东旭. 三角域上的 Walsh 函数. 科学通报，1988,33(9):715～716
(Qi Dong-xu. Walsh function on the triangular area. Chinese Science Bulletin, 1988,33(9):715～716)

8 齐东旭. 一类自相似结构及 Walsh 函数新的定义. 数学年刊, 1991, 12(增刊A): 103~105
(Qi Dong-xu. New definition for a kind of self-similar structure and Walsh function. Chinese Annals of Mathematics, 1991, 12 (Supplement A): 103~105)

Multi-scale Expression of Data and Construction of a Kind of Orthogonal Transformation

DING Wei[1] QI Dong-xu[1,2]

[1](*CAD Laboratory Insitute of Computing Technology The Chinese Academy of Sciences Beijing* 100080)
[2](*CAD Research Center North China University Beijing* 100041)

Abstract In this paper, the authors mainly discuss about how to express functions based on {0,1} code. At first, the examples are given to explain a new method in signal processing, in which functions are defined from the viewpoint of multi-scale analysis. Then the authors present "coordinate heft separate" method, discover self-similar structure of points set on plain area; based on this self-similar structure, construct orthogonal complete 2-value function system of high-dimension simplicity; enrich and extend the theory and application of Rademacher function and Walsh function.

Key words Multi-scale analysis, orthogonal transformation, coordinate heft separate.