

# 时域和酉空间中基于最大相关熵准则的非线性噪声处理

姜 骁<sup>2\*</sup>, 马文涛<sup>1</sup>, 曲 桦<sup>2</sup>

(1. 西安交通大学 电子与信息工程学院, 西安 710049; 2. 西安交通大学 软件学院, 西安 710049)

(\* 通信作者电子邮箱 kiimm\_pri@163.com)

**摘 要:** 针对非线性噪声处理的问题, 考虑到信号的高阶统计量以及在酉空间可以很好地处理非高斯噪声, 提出了在时域和酉空间中基于最大相关熵准则(MCC)的噪声处理算法。结合 MCC 和梯度下降算法, 设计出了时域中非线性噪声的滤波算法。同时将该算法推广到酉空间中噪声处理, 给出了酉空间中基于 MCC 的滤波算法。通过仿真研究发现, 在时域和酉空间中, 基于 MCC 的滤波算法相对于传统的基于最小均方差(LMS)的滤波算法在处理非高斯噪声的问题时有着显著优势, 以更快的收敛速度达到能够较完整地保留信号特征的效果。

**关键词:** 自适应滤波器; 酉空间; 最大相关熵准则; 最小均方差算法; 非线性噪声

**中图分类号:** TN911.7 **文献标志码:** A

## Max correntropy criteria-based nonlinear noise processing in time domain and unitary space

JIANG Xiao<sup>2\*</sup>, MA Wen-tao<sup>1</sup>, QU Hua<sup>2</sup>

(1. School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an Shaanxi 710049, China;

2. School of Software Engineering, Xi'an Jiaotong University, Xi'an Shaanxi 710049, China)

**Abstract:** Considering the problems for nonlinear-noise processing and taking account of that higher-order statistics of the signal and unitary space can be a good deal with non-Gaussian noise, the noise processing algorithm based on Max Correntropy Criteria (MCC) in the time domain and the unitary space was proposed. Combining the MCC and gradient descent algorithm, a nonlinear-noise filtering algorithm in the time domain was designed. At the same time, extending the algorithm to the noise processing in the unitary space, the unitary space filtering algorithm based on the MCC was put forward. The simulation study shows that the algorithm based on the MCC algorithm has significant advantages compared with the traditional Least Mean Square (LMS) based filtering algorithm, which means it can achieve more complete signal characteristics by faster convergence.

**Key words:** adaptive filter; unitary space; Max Correntropy Criteria (MCC); Least Mean Square (LMS) algorithm; nonlinear noise

## 0 引言

在现实的应用系统中,要处理的信号噪声往往是具有非线性非高斯特性的。目前对于噪声的滤波问题,已经有了很多的研究<sup>[1-3]</sup>,然而这些算法大部都是基于传统的基于最小均方差算法(Least Mean Square, LMS)算法的,而此算法都是在对于噪声是线性和高斯假设下进行设计的。因此,在实际环境中出现的非高斯非线性的噪声处理往往具有一定的局限性。因为它通常是以均方误差(Mean Squared Error, MSE)准则为代价函数的,其只考虑了信号的二阶统计量,在信号处理中不能准确地体现出信号的统计特征。近年来文献[4]中提出了一种新的相似度方法——广义相关熵(Correntropy),该方法考虑到了信号的高阶统计量特征,可以更加准确地评估输入输出信号的误差,基于该相似度函数制定出了一种新的评价准则——最大相关熵准则(Max Correntropy Criteria, MCC)。由于该准则可以很好地处理非线性非高斯环境中的问题,已经得到了广泛的应用<sup>[5-7]</sup>。考虑到 MCC 的特性,将其应用于非线性非高斯噪声的语音信号自适应滤波中,提出了基于 MCC 的非线性噪声滤波算法。同时若将信号通过酉

变换映射到酉空间中进行处理,可以很好地解决具有非线性非高斯特性的噪声滤波问题<sup>[8]</sup>。本文将 MCC 应用于酉空间中的信号处理,给出了一种酉空间中基于 MCC 的非线性噪声滤波算法。相比传统的 LMS 算法,基于 MCC 的算法在非线性的信号处理中具有快速收敛和鲁棒性的特点。

## 1 基本理论

### 1.1 LMS 算法

LMS 算法在 1959 年被引进自适应系统来训练一个未知系统的参数。算法的核心是要使得滤波器的输出信号与期望信号之间误差的均误差达到最小。其算法进行参数调整的迭代公式如式(1)所示:

$$\mathbf{W}(n+1) = \mathbf{W}(n) + \mu * e(n) * \mathbf{X}(n) \quad (1)$$

其中:  $\mathbf{W}(n)$  为自适应系统参权重数,  $\mathbf{X}(n)$  为系统的输入。由于 LMS 算法具有简单的计算复杂度,因而其在实际中得到了广泛的应用<sup>[9-10]</sup>。

### 1.2 MCC

熵是一种利用数学的方式来实现信息量化的方式,可以较完整地表征整个信息的结构和信息,可作为自适应系统中

收稿日期: 2012-06-08; 修回日期: 2012-07-11。

基金项目: 国家无线重大专项(2010ZX03004-001, 2010ZX03004-002, 2011ZX03002-001); 国家自然科学基金资助项目(61071126)。

作者简介: 姜骁(1990-), 男, 新疆哈密人, 硕士研究生, 主要研究方向: 复杂网络; 马文涛(1979-), 男, 陕西宝鸡人, 博士研究生, 主要研究方向: 通信网、机器学习; 曲桦(1961-), 男, 陕西杨凌人, 教授, 博士生导师, 博士, 主要研究方向: 现代通信网、计算机网络体系结构。

的代价函数,用于训练系统权值。Correntropy 是利用高阶统计量来定义两个随机变量之间相似度,在自适应滤波器中即理想输出信号与滤波器实际输出信号之间的信号相似度的计算衡量方法。

Correntropy 作为衡量任意两个随机变量  $X$  和  $Y$  之间局部相似度的方法,其定义如式(2)所示:

$$V(X, Y) = E[K(X - Y)] \quad (2)$$

其中:  $K(\cdot)$  是定义的核函数,通常选择常用的高斯核函数,其中参数核大小  $\sigma$  对是影响其性能的主要因素<sup>[11]</sup>,同泰勒级数展开可知包含两个变量误差  $(X - Y)$  的高阶统计量。在实际应用中只有有限数量的一些变量  $X$  和  $Y$  样本  $\{(X_i, Y_i)\}_{i=1}^N$  是可用的。同时,使用样本估计是对期望操作更常见的一种方式。使用 Parzen 窗估计可以得到 Correntropy 的一种近似估计,取其最大值就导出 MCC,如式(3)所示:

$$\begin{cases} \bar{V}(X, Y) = \frac{1}{N} \sum_{i=1}^N K(X_i - Y_i) \\ \text{MCC} \Leftrightarrow \max \bar{V}(X, Y) \end{cases} \quad (3)$$

MCC 作为代价函数有着如同 LMS 算法的简单计算复杂度<sup>[12-13]</sup>。选择常规的高斯核函数来计算 Correntropy,可得式(4):

$$J_n = \frac{1}{\sqrt{2\pi}\sigma} * \frac{1}{N} \sum_{i=n-N+1}^n \exp\left(-\frac{(d_i - y_i)^2}{2\sigma^2}\right) \quad (4)$$

如果滤波器在时刻  $n$ ,当时的权值为  $W_n$ ,那么代价函数可以如下表示:

$$J_n = \frac{1}{\sqrt{2\pi}\sigma} * \frac{1}{N} \sum_{i=n-N+1}^n \exp\left(-\frac{(d_i - W_n^T X_i)^2}{2\sigma^2}\right) \quad (5)$$

分析可得,找到一个权值  $W_n$  的集合去取代代价函数  $J$  的最大值是很困难的,在此也采用如 LMS 算法的梯度下降的方法来进行迭代运算,从而得到系统的最优权值。其中梯度如式(6)所示:

$$\begin{aligned} J_n &= -\frac{1}{\sqrt{2\pi}\sigma^3} * \frac{1}{N} \sum_{i=n-N+1}^n \exp\left(-\frac{(d_i - W_n^T X_i)^2}{2\sigma^2}\right) (d_i - \\ &W_n^T X_i) = -\frac{1}{\sqrt{2\pi}\sigma^3} * \frac{1}{N} \sum_{i=n-N+1}^n \exp\left(-\frac{e_i^2}{2\sigma^2}\right) e_i \end{aligned} \quad (6)$$

其中  $e_i = d_i - W_n^T X_i$ 。

从而可以得到具体的迭代公式如式(7)所示:

$$W_{n+1} = W_n - \mu \nabla J_n \quad (7)$$

其中  $\nabla J_n$  为  $J_n$  的梯度。

将式(5)代入计算,可以得到:

$$W_{n+1} = W_n + \frac{\mu}{\sqrt{2\pi}\sigma^3} * \frac{1}{N} \sum_{i=n-N+1}^n \left[ \exp\left(-\frac{e_i^2}{2\sigma^2}\right) e_i X_i \right] \quad (8)$$

若取一个点的样本均值( $n=1$ ),权值迭代式可简化为:

$$W_{n+1} = W_n + \frac{\mu}{\sqrt{2\pi}\sigma^3} * \exp\left(-\frac{e_n^2}{2\sigma^2}\right) e_n X_n \quad (9)$$

### 1.3 酉变换

在酉空间中,与 Euclid 空间的正交变换相平行的概念就是酉变换<sup>[14]</sup>。酉变换即是在酉空间  $V$  的一个线性变换  $\sigma$ ,若对于任意的  $\alpha, \beta \in V$ ,都有  $[\sigma(\alpha), \sigma(\beta)] = [\alpha, \beta]$ 。任意的 Hermite 矩阵都“酉相似”于一个实对角矩阵,即若  $A$  是一个  $n$  阶的 Hermite 矩阵,则存在一个  $n$  阶酉矩阵  $U$ ,使得  $U^H A U = U^{-1} A U$  是一个实对角矩阵。因此可以通过 Hermite 矩阵获得一个酉矩阵。其中酉变换可以简化为式(10):

$$T(x) = Ax; x \in C^n(R^n) \quad (10)$$

其中  $A$  为酉矩阵将非线性的矩阵向量在酉空间中作线性的

变换。对矩阵进行计算和特征分解<sup>[15]</sup>,利用信号与噪声子空间的正交特性来处理非高斯的信号噪声。

若  $A$  为 Hermite 矩阵,可以对其进行 Hessenberg 分解,即  $A$  可写为:

$$A = P * H * P'$$

其中:  $P$  为酉矩阵,  $H$  为 Hessenberg 矩阵。酉矩阵  $P$  可作为我们将数据映射到酉空间中的变化。

## 2 在时域和酉空间中对应的算法

### 2.1 在时域中基于 MCC 的滤波算法

在时域中,要处理的信号为有效信号和噪声的混合信号,对给定步长大小的 MCC,选择合适大小的核函数,进行权值的更新。时域中基于 MCC 的算法如下:

步骤 1 设置参数及初始化,迭代次数  $e_{max}$ ,滤波器抽头数  $M$ ,收敛因子  $mu$ ,核大小  $\sigma$ ,权值向量  $W(0) = [0, 0, \dots, 0]$ 。

步骤 2 迭代训练。迭代次数  $k$  小于  $e_{max}$  时,进行如下迭代:

$$Y_k = X_k^T W_k$$

$$e_k = S_k - Y_k$$

$$W_{k+1} = W_k + \frac{mu}{\sqrt{2\pi}\sigma^3} * \exp\left(-\frac{e_k^2}{2\sigma^2}\right) e_k X_k$$

其中:  $Y_k, X_k$  为第  $k$  次迭代时的输出和输入;  $S_k$  为期望的输出信号;  $W_k$  为第  $k$  次迭代得到的权值向量;  $e_k$  为第  $k$  次的误差。

步骤 3 根据训练好的权值向量进行含噪信号的滤波处理:

$$Y = W * X$$

其中:  $W$  为训练好的权值系数,  $X$  为滤波器输入,  $Y$  为滤波器输出。

### 2.2 酉空间基于 MCC 的滤波算法

在酉空间中,首先要对输入信号进行酉变换,将信号投影到酉空间中,同时要系统对权值也进行对应的酉变换,对给定步长大小的 MCC,选择合适大小的核函数,在酉空间中进行权值的更新。基于酉变换 MCC 的算法如下:

步骤 1 设置参数及初始化,迭代次数  $e_{max}$ ,滤波器抽头数  $M$ ,收敛因子  $\mu^U$ ,核大小  $\sigma$ , Hermite 矩阵  $H$ ,权值向量  $W(0) = [0, 0, \dots, 0]$ 。

步骤 2 对 Hermite 矩阵进行 Hessenberg 分解得到酉变换矩阵  $P$ 。

步骤 3 迭代训练。迭代次数  $k$  小于  $e_{max}$  时,进行如下迭代。

对输入向量和权值向量进行酉变换得到酉变换后的结果:

$$X^U(k) = X(k) * P$$

$$W^U(k) = W(k) * P$$

$$Y_k^U = X_k^{UT} W_k^U$$

$$e_k = S_k - Y_k^U$$

$$W_{k+1}^U = W_k^U + \frac{\mu^U}{\sqrt{2\pi}\sigma^3} * \exp\left(-\frac{e_k^2}{2\sigma^2}\right) e_k X_k^U$$

其中:  $Y_k^U, X_k^U$  为第  $k$  次迭代时的输出和输入;  $S_k$  为期望的输出信号;  $W_k^U$  为第  $k$  次迭代得到的权值向量;  $e_k$  为第  $k$  次的误差。

步骤 4 根据训练好的权值向量进行含噪信号的滤波处理:

$$Y = W * X$$

### 3 仿真结果分析

#### 3.1 基于 MCC 自适应滤波器设计架构

基于 MCC 自适应滤波器是能够根据输入信号自动调整性能进行数字信号处理的数字滤波器。如图 1 所示为基于 MCC 的滤波器框架。以输入和输出信号的相关熵统计特征为依据,采用特定的 MCC 算法自动地调整滤波器的系数,使其达到滤波特性最优。自适应滤波器对输入信号序列的  $U(k)$  的每一个样值,按照给定的算法,更新、调整系数权值  $W(n)$ ,使其输出序列  $Y(k)$  和期望输出序列  $S(k)$  之间的误差最小,即输出信号序列  $Y(k)$  尽可能地逼近期望信号序列  $S(k)$ 。

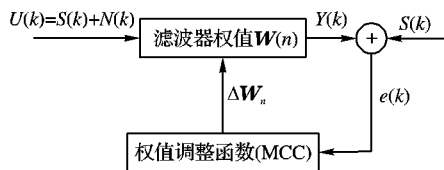


图1 基于 MCC 的滤波器框架

#### 3.2 实验环境及原始数据

本文在 CPU 为 2 GHz, 内存为 2 GB 的计算机中, 以 Matlab2011b 为仿真平台, 进行噪声干扰抵消滤波仿真实验。在仿真中以作者所录的人声朗读音频文件 (read.wav) 为基本数据, 以采样率为 8 kHz 进行采样得到源信号  $S$ 。同时我们主要针对非高斯有色噪声的干扰进行研究, 而有色噪声可通过将白噪声进行非线性动态处理来产生, 所以采用式 (11) 作为有色噪声产生模型来得到仿真中所需的噪声  $colN$ 。

$$colN(k) = \frac{4 \sin(n(k-1)) * n(k-1)}{1 + n^2(k-1)} \quad (11)$$

其中  $n$  表示产生的高斯白噪声。

根据采样得到的源信号  $S$  和有色噪声信号  $colN$  进行叠加得到含噪信号  $S'$ , 如式 (12) 所示:

$$S' = S + colN \quad (12)$$

#### 3.3 仿真结果分析

为了验证算法的有效性, 每 50 次滤波实验为一组进行了 8 组实验, 根据 3.2 节中的描述得到待处理的信号, 同时对时域和酉空间下的 LMS 算法与 MCC 算法进行了统计结果对比, 包括结果输出、误差输出和均方差的统计。

##### 3.3.1 基于 MCC 的自适应滤波器收敛性能分析

算法的收敛性是保证算法性能的主要特性, 为了研究所提算法的收敛性, 本文以正弦信号作为原始信号, 噪声模型采用式 (11) 来得到含噪的信号, 分别采用 LMS 算法和基于 MCC 的自适应滤波算法进行噪声抵消实验, 通过计算它们的均方误差得到了其收敛结果, 如图 2 所示。

从图中可以看到两种算法的性能差别。LMS 算法在收敛过程中, 当迭代到 300 次时逐渐收敛到最优值, 其值约在 0.22。同时基于 MCC 的算法明显收敛速度快于 LMS 算法, 仅仅不到 100 次的迭代过程就将算法的精度收敛到了 0.2 以下。基于 MCC 的滤波算法不但收敛速度快而且有更更高的精度。

##### 3.3.2 时域中基于 MCC 的滤波算法结果分析

根据 2.1 节中算法的描述, 首先对算法中的参数进行初始化, 其中滤波抽头数  $M = 10$ , 收敛因子  $\mu = 0.05$ , 核大小  $\sigma = 0.5$ , 权值向量  $W$  初始化为 0。图 2 给出了时域中两种算法滤波结果, 其中图 3(a)、(b) 分别为混合噪声之后的信号和期望信号, 而图 3(c)、(d) 分别为经过 LMS 算法和基于 MCC

的滤波算法后得到的结果, 从图 3(d) 中可以看出经过 MCC 滤波处理后的结果, 相对来说可以很好将原始信号的特征保留得更明显一些, 尤其在  $0 \sim 2 \times 10^4$  个采样点的区间。

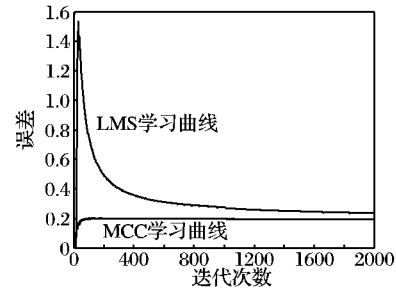


图2 LMS 算法和基于 MCC 算法的学习曲线

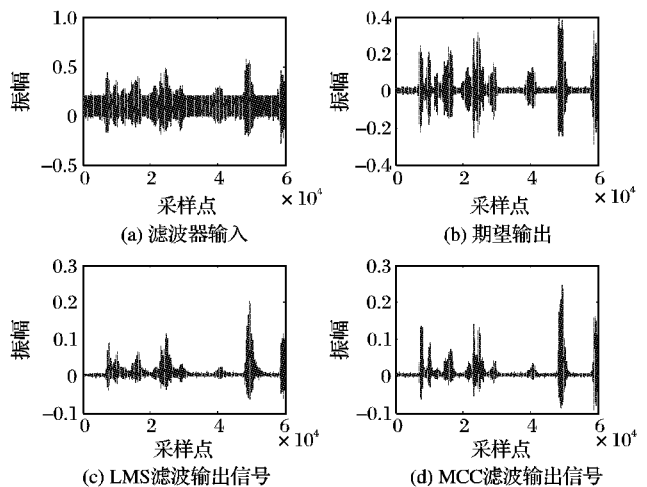


图3 在时域中两种算法滤波结果

从两种算法的滤波输入输出误差角度对两种算法的性能进行对比, 如图 4 所示。可以看出, 在时域中两者的误差信号大致相同。但是通过计算 LMS 算法与 MCC 对应的均方差数值分别为 0.0011 与  $9.6782 \times 10^{-4}$ , 可知 MCC 相对较好。

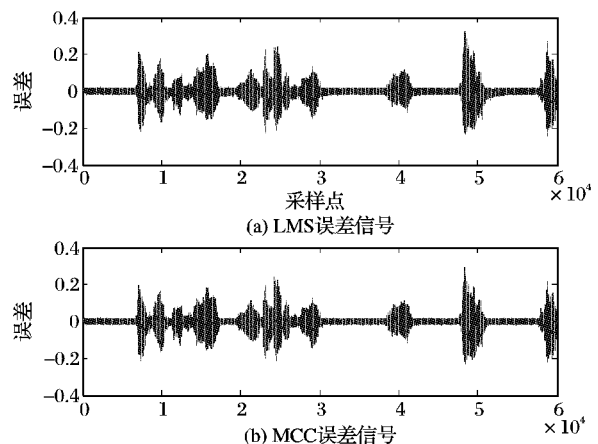


图4 在时域中两种算法的误差对比

##### 3.3.3 酉空间中基于 MCC 的滤波算法结果分析

在酉空间中对信号进行处理已经成为一种研究的主要方向, 本文将 MCC 应用于酉空间中的噪声滤波问题, 使用 3.2 节中所提的算法进行了在酉空间中的噪声滤波仿真, 其中参数设置为滤波抽头数  $M = 10$ , 收敛因子  $\mu = 0.05$ , 核大小  $\sigma = 0.5$ , 权值向量  $W$  初始化为 0, 使用 Matlab 中的函数 hilb 产生一个 10 阶的 Hilbert 矩阵, 并对其进行 Hessenberg 分解可得到酉变换矩阵  $P$ 。如图 5 所示为酉空间中两种算法滤波结果, 从图 5(c)、(d) 所示, 经过 LMS 算法滤波后结果与 MCC

滤波后的信号对比, MCC 滤波后的信号在信号的振幅和波峰的特征都远好于 LMS 算法。

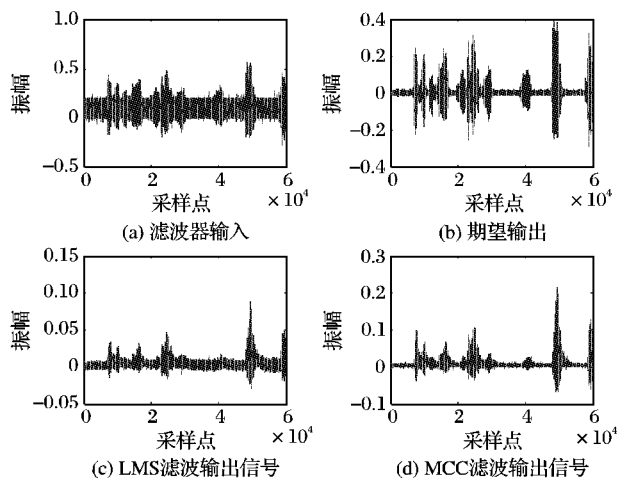


图 5 在酉空间中两种算法滤波结果

对两者的误差信号进行统计并进行对比,如图 6 所示。可以明显看到 MCC 的误差要小于 LMS 算法的滤波器,也意味着基于 MCC 的滤波器有着更高的精度。

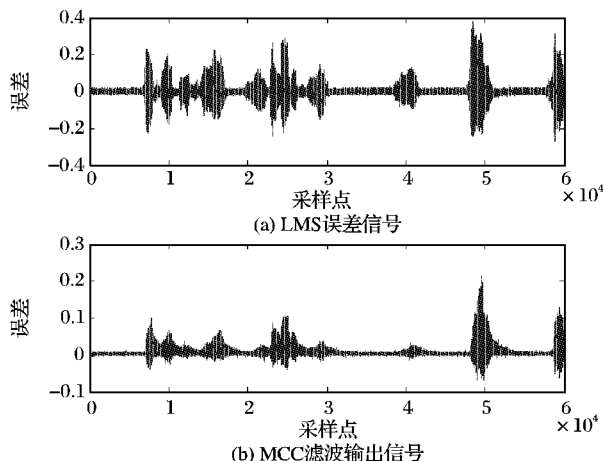


图 6 在酉空间中两种算法的误差对比

### 3.3.4 核大小 $\sigma$ 对算法性能的影响分析

由于 MCC 的函数性能直接受步长和核函数大小的影响<sup>[16]</sup>,在步长一定的情况下,需要研究不同的核大小对基于 MCC 的算法的影响。实验分别在时域和酉空间中进行了 8 组实验,核大小分别选取为 0.1~0.8 共 8 个数据进行实验,得到不同  $\sigma$  下的滤波器输出与期望信号的均方误差结果,如图 7 所示。均方误差是在每个相应大小的核函数下取得的统计量,在图中可以明显看到相同输入在不同的核大小对算法的影响。同时也直观地看到了相同信号在不同处理域中的差异,在时域中核大小取 0.4 时,可以达到误差最小,而在酉空间中核大小则应取 0.2。

## 4 结语

本文对输入信号在时域和酉空间中分别进行了有色噪声的滤波实验,同时进行了 MCC 和 LMS 算法的性能比较。通过仿真分析可知 MCC 作为代价函数,在时域和酉空间中对非线性有色信号进行滤波,无论在时域还是酉空间中,算法的性能都是优于 LMS 的。尤其在酉空间中,优势更加明显,同时其收敛速度和收敛精度比 LMS 有着显著优势。由此可知基于 MCC 的滤波算法对于有色噪声干扰滤波不失为一种有效

的方法。

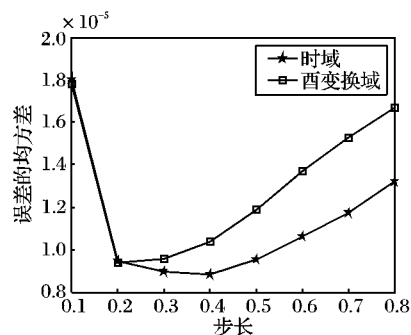


图 7 在时域和酉空间中步长对 MCC 性能的影响

### 参考文献:

- [1] 刘明远,徐松涛,刘凯,等.改进的变步长 LMS 自适应滤波算法分析[J].火力与指挥控制,2010,35(12):110-118.
- [2] SHI KUN, MA XIAOLI. Transform domain LMS algorithms for sparse system identification[C]// 2010 IEEE International Conference on Acoustics Speech and Signal Processing. New York: IEEE, 2010:3714-3717.
- [3] GÓRRIZ J M, RAMÍREZ J, ALVAREZ S C, *et al.* Novel LMS algorithm applied to adaptive noise cancellation[J]. IEEE Signal Processing Letters, 2009, 16(1): 34-37.
- [4] 杨南海,黄明明,赫然,等.基于最大相关熵准则的鲁棒半监督学习算法[J].软件学报,2012,23(2):279-288.
- [5] SINGH A, PRINCIPE J C. Using correntropy as a cost function in adaptive filters[J]. International Joint Conference on Neural Networks. Piscataway: IEEE, 2009:2950-2955.
- [6] ZHAO SONGLIN, PRINCIPE J C. A nonparametric information theoretic approach for change detection in time series[C]// The 2011 International Joint Conference on Neural Networks. Piscataway: IEEE, 2011: 1281-1284.
- [7] CHOUZENOUX E, MOUSSAOUI S, IDIER J, *et al.* Optimization of a maximum entropy criterion for 2D nuclear magnetic resonance reconstruction[C]// 2010 IEEE International Conference on Acoustics Speech and Signal Processing. Piscataway: IEEE, 2010: 4154-4157.
- [8] 张杰,廖桂生,王珏.空间相关色噪声下基于酉变换的信号源数目估计[J].电子学报,2005,9(9):1581-1584.
- [9] CHEN YILUN, GU YUANTAO, HERO A O. Sparse LMS for system identification[C]// Proceedings of the 2009 IEEE International Conference on Acoustics, Speech and Signal Processing. Washington, DC: IEEE Computer Society, 2009:3125-3128.
- [10] TING L K, COWAN C F N, WOODS R F. LMS coefficient filtering for time-varying chirped signals[J]. IEEE Transactions on Signal Processing, 2004, 52(11): 3160-3169.
- [11] LIU WEIFENG, PRINCIPE J C. Correntropy: Properties and applications in non-Gaussian signal processing[J]. IEEE Transactions on Signal Processing, 2007, 55(11): 5286-5298.
- [12] 李明,杨成梧.基于模糊推理的变步长 LMS 自适应滤波算法[J].控制工程,2006,16(3):237-240.
- [13] 林川,冯全源.模糊步长 LMS 算法及其性能分析[J].系统工程与电子技术,2007,29(6):967-971.
- [14] 牟建超,高梅国,白锦良.基于酉变换的改进 GSO 波束形成器[J].北京理工大学学报,2010,30(7):826-830.
- [15] 詹耀华.关于在欧氏空间和酉空间里的正定、半正定矩阵的对比研究[J].沈阳大学学报,2010,22(3):24-26.
- [16] 李爱国,汪保男.一种非线性新相关信息熵定义及其性质、应用[J].信息与控制,2011,40(3):401-407.