

Data Compression Technology of Distributed Cooperative Passive Location

Zhiwei Zhuang¹, Yi Zhan², Jian Qian¹

¹School of Communication Engineering, Hangzhou Dianzi University, Hangzhou

²Science and Technology on Communication Information Security Control Laboratory, No. 36 Research Institute of CETC, Jiaxing

Email: zhuang_zhiwei@126.com

Received: Oct. 22nd, 2014; revised: Nov. 16th, 2014; accepted: Nov. 20th, 2014

Copyright © 2014 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

In data compression technology of distributed cooperative passive location, TDOA has the advantages of high positioning accuracy, positioning speed and strong anti-interference ability. But, TDOA location needs different sensors to transmit the received signal to the same sensor to make TDOA estimation, which needs massive data transmission. In order to solve the massive data transmission, data compression is used in location. For getting higher compression ratio, vector quantization is used to deal with the signal. Through simulating the different vector dimension and different codebook size to find what affect the TDOA estimation accuracy. Experiments prove that vector quantization can obtain high compression ratio without loss of the accuracy of TDOA estimation, and solve the problem of massive data transmission in TDOA effectively.

Keywords

Vector Quantization, TDOA, Data Compression

分布式协同无源定位中的数据压缩技术

庄智威¹, 詹毅², 钱建¹

¹杭州电子科技大学通信工程学院, 杭州

²中国电子科技集团36所, 通信信息控制和安全技术重点实验室, 嘉兴

Email: zhuang_zhiwei@126.com

收稿日期：2014年10月22日；修回日期：2014年11月16日；录用日期：2014年11月20日

摘要

分布式协同无源定位技术中，到达时差定位技术(Time Differences of Arrive, TDOA)具有定位精度高、定位速度快、抗干扰能力强的优点。但是，TDOA定位需要不同位置的传感器节点将采样值传送到同一个节点进行时差估计，需要大量的数据传输。为了避免大规模的数据传输，数据压缩技术被应用到定位之中。为了达到较高的压缩比，对信号波形进行矢量量化(Vector Quantization, VQ)。仿真不同维数、不同码书大小对TDOA估计精度的影响。实验证明矢量量化能够取得巨大的压缩比，且不失TDOA估计精度，有效的解决TDOA定位技术的大规模数据传输问题。

关键词

矢量量化，到达时差定位，数据压缩

1. 引言

分布式协同无源定位中，需要不同位置的传感器节点把接收到的信号样值传送到同一节点，利用相关法估计达到时间差来确定目标源的位置。由于无线传感网络的节点通信能力较弱、节点的能量有限，这种大规模的数据传输使分布式协同定位技术面临着巨大的挑战，而数据压缩是解决此类问题最常用的方法之一。目前，针对无源定位的数据压缩技术领域较新的研究有：Mark L. Fowler 和 Chen Mo 提出的基于费舍尔信息的数据压缩方法[1] [2]。该方法折衷信息传输率、能耗、估计精度，在保证较高的TDOA估计精度的条件下，达到高效的压缩。李协在无线传感网中面向到达时差估计的数据压缩方法研究中[3]，对费舍尔信息的数据压缩进一步研究，提出新的比特分配方法代替运算效率低的拉格朗日乘子法，同时提出基于费舍尔信息的数据压缩方法中幅度谱不携带时延信息，只传送相位谱得到进一步的压缩。以上无源定位的数据压缩技术研究均是基于标量量化，而矢量量化技术比标量量化能获得更大的压缩比，且矢量量化的解码算法简单。该文对信号波形进行矢量量化，传送矢量索引，在接收端根据索引利用码书还原信号波形，并与接收到的信号作相关运算，得到TDOA估计值。

2. 分布式协同无源定位技术

TDOA 定位技术具有精度高、速度快等优点，是无源定位技术最常用的方法之一。假设目标发射的信号为随机信号 $s(t)$ ，两个分散在不同位置的传感器节点收到的信号 $s_i(t)$ 和 $s_k(t)$ 分别为

$$s_i(t) = s(t) + n_i(t) \quad (1)$$

$$s_k(t) = As(t - \tau_{ik}) + n_k(t) \quad (2)$$

τ_{ik} 为两个传感节点接收到信号的时差。假设接收到的信号中噪声统计独立。则两个信号的相关函数[4]为：

$$\begin{aligned} R_{ik}(\tau) &= E\{s_i(t)s_k(t+\tau)\} \\ &= AE\{s(t)s(t+\tau-\tau_{ik})\} + AE\{n_i(t)s(t+\tau-\tau_{ik})\} + E\{s(t)n_k(t)\} + E\{n_i(t)n_k(t)\} \\ &= AE\{s(t)s(t+\tau-\tau_{ik})\} \\ &= AR_{ss}(\tau-\tau_{ik}). \end{aligned} \quad (3)$$

式(3)中, $R_{ss}(\tau)$ 是 $s(t)$ 的自相关函数。所以 $R_{ik}(\tau)$ 最大值出现在 $\tau = \tau_{ik}$ 处。即 $\hat{\tau}_{ik} = \arg \max_{\tau} \{ |R_{sisk}(\tau)| \}$ 。假设电波在空气中传播速率为光速 c , 时差转化为距离差, 通过目标源到两个传感器的距离差可以获得一组双曲线。三个不同位置的节点可以获得三条双曲线, 任取其中的两条双曲线, 求其交点就是目标源的位置。

3. 矢量量化

3.1. 码书设计

码书设计、码字搜索、码字的索引分配是矢量量化的三大关键技术。码书的性能直接影响矢量量化的性能。文中选取欧几里得测度作为失真测度:

$$d(\mathbf{X}, \mathbf{Y}) = \sqrt{\sum_{i=0}^{k-1} |x_i - y_i|^2} \quad (4)$$

在量化复数序列时, 式(4)中的绝对值为模。码书设计遵循两条准则[5]: (1) 最近邻条件, 给定码书后, 训练矢量的最佳划分通过每个训练矢量到离它最近的码字得到。假设码书 $\mathbf{C} = \{\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_{N-1}\}$, 训练矢量 $\mathbf{X} = \{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{M-1}\}$, 则最佳划分 $\mathbf{S} = \{s_0, s_1, \dots, s_{N-1}\}$ 满足: $S_i = \left\{ \mathbf{v} \mid d(\mathbf{v}, \mathbf{y}_i) = \min_{0 \leq j \leq N-1} d(\mathbf{v}, \mathbf{y}_j), \mathbf{v} \in \mathbf{X} \right\}$ 。(2) 质心条件, 给定训练矢量划分, 最优码书中的码字应为各聚类的矢量质心。给定划分为 $\mathbf{S} = \{s_0, s_1, \dots, s_{N-1}\}$, 使平均失真达到最小, 码字 \mathbf{y}_i 须为相应 s_i 的质心, 即质心为: $\mathbf{y}_i = \frac{1}{|s_i|} \sum_{\mathbf{v} \in s_i} \mathbf{v}$ 。 $|s_i|$ 为集合 s_i 中矢量的个数。

由 Linde, Buzo 和 Gray 三人提出一种有效的矢量量化码书设计算法——LBG 算法, 就是上述两个准则反复迭代的过程, 从初始码书获得最优码书。LBG 算法的流程如下:

(1): 给定初始码书 $\mathbf{C}^{(0)} = \{\mathbf{y}_0^{(0)}, \mathbf{y}_1^{(0)}, \dots, \mathbf{y}_{N-1}^{(0)}\}$, 迭代次数 n 设为 0, 平均失真 $D^{(-1)} \rightarrow \infty$, 相对误差门限 $\varepsilon (0 < \varepsilon < 1)$ 。

(2): 码书中的码字作为质心, 根据最近邻条件把训练矢量 \mathbf{X} 划分为 $\mathbf{S}^{(n)} = \{s_0^{(n)}, s_1^{(n)}, \dots, s_{N-1}^{(n)}\}$ 。

(3): 计算平均失真: $D^{(n)} = \frac{1}{M} \sum_{i=0}^{M-1} \min_{0 \leq j \leq N-1} d(\mathbf{x}_i, \mathbf{y}_j^{(n)})$, 是否满足 $\left| \frac{D^{(n-1)} - D^{(n)}}{D^{(n)}} \right| \leq \varepsilon$ 。若满足, 停止迭代, $\mathbf{C}^{(n)}$ 是最优码书。否则, 进行步骤(4)。

(4): 根据质心条件, 求质心 $\mathbf{y}_i^{(n+1)} = \frac{1}{|s_i^{(n)}|} \sum_{\mathbf{v} \in s_i^{(n)}} \mathbf{v}$, N 个质心构成新码书, $n = n + 1$, 转步骤 2。

3.2. 码字搜索

码书建立后, 对随机输入矢量 \mathbf{x} , 按照最近邻条件寻找最小失真码字, 即 $d(\mathbf{x}, \mathbf{y}_i) = \min_{0 \leq j \leq N-1} d(\mathbf{x}, \mathbf{y}_j)$ 。

穷尽搜索法是最直观的最近邻码字搜索算法, 需要计算输入矢量与所有码字的失真, 找出最小失真码字, 并传输该码字的索引 i , 在接收端根据码书寻找 i 所代表的码字 \mathbf{y}_i , 用 \mathbf{y}_i 代替输入矢量。

4. 仿真分析

信源信号为随机序列的 FM 调制的窄带信号, 载波为 10 MHz, 采样频率为 100 MHz, 带宽为 0.1 MHz。传输信号的复包络在传感器节点进行 $K = 16, 24, 32, 40$ 维数的矢量量化, 码书大小为 128、256、512。

图 1 中码书大小为 128, 各维数在每个信噪比下做 1000 次蒙特卡洛仿真求得的 TDOA 均方误差。图 2 中矢量维数 $K = 32$, 不同码书大小在每个信噪比下做 1000 次蒙特卡洛仿真求得的 TDOA 均方误差。

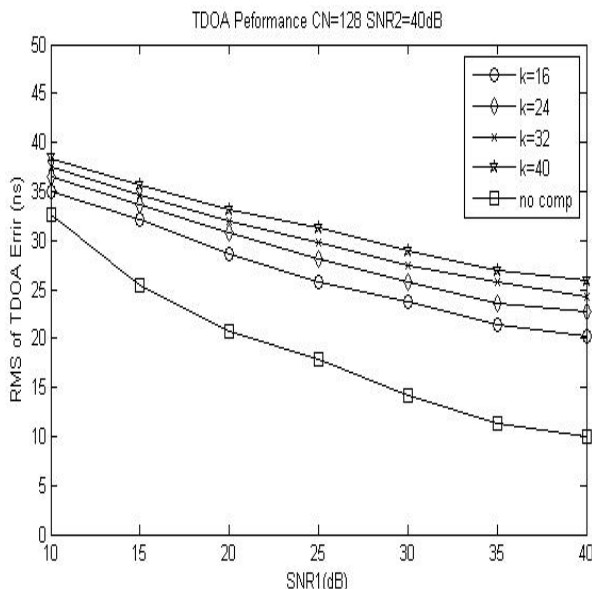


Figure1. The mean square error of TDOA in each dimension with CN = 128

图 1. CN = 128 各维数下的 TDOA 均方误差

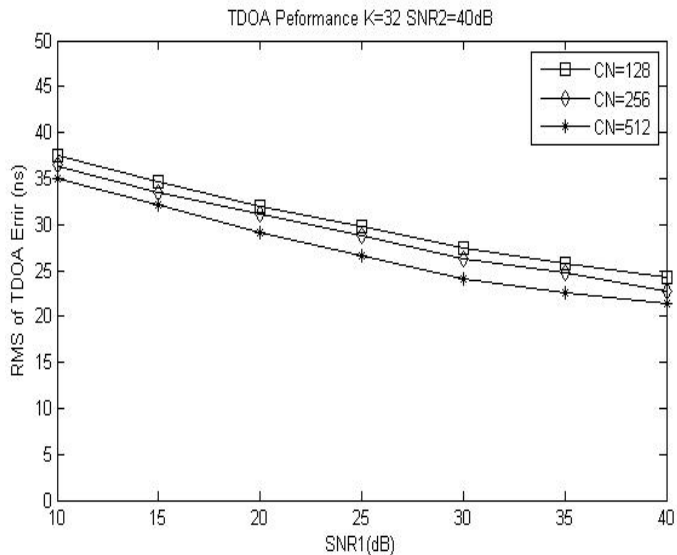


Figure2. The mean square error of TDOA in different codebook sizes with K = 32

图 2. K = 32 不同码书大小下的 TDOA 均方误差

假设每个采样点为一个浮点数，传送 $K \times M$ 个采样点，需要 $32 \times K \times M$ 个比特。对信号经过矢量量化后，需要传 M 个索引到参考节点，若码书大小为 CN 则每个索引需要 $\log_2 CN$ 个比特编码，总共传输 $M \times \log_2 CN$ 个比特，假设信道无失真传输。如图 1 中所示， $CN = 128$ ，各维数矢量量化后所能达到的压缩比在 70 倍以上，并且 TDOA 值的均方误差值与没有压缩的相差并不大。基于费舍尔信息的数据压缩方法中，压缩比达到 4 倍时，TDOA 估计值与没有压缩条件下相差甚微。虽然矢量量化在估计精度上没有标量化来的高，但在估计精度需求并不是很高的条件下，矢量量化所能获得的压缩比是标量化所达不到的。从图 1 可以看出，码书大小一定时，随着维数的增加，压缩比增大，但是编码速率下降，信

号失真变大, 所得到的 TODA 估计性能也变差。在图 2 中, 矢量量化维数 $K = 32$, 码书大小分别为 128、256、512, 能达到的压缩比分别为 146 倍, 128 倍, 114 倍, 比较看出, 码书大小越大, TDOA 的估计精度越好。

5. 结束语

矢量量化技术应用于分布式无源定位的数据压缩技术中, 不但能够获得很高的压缩比, 定位精度也不会损失太多。其中需要选择适当的维数、码书大小折衷压缩比和计算复杂度。进一步的工作可以对信号进行变换域矢量量化, 对信号进行正交变换后得到变换域系数, 舍去其中的高频分量, 不但缩减了量化维数, 还能取得更好的压缩效果。

参考文献 (References)

- [1] Chen, M. and Fowler, M.L. (2010) Data compression for multi-parameter estimation for emitter location. *IEEE Transactions on Aerospace and Electronic Systems*, **46**, 308-322.
- [2] Fowler, M.L. and Chen, Mo. (2005) Fisher-information-based data compression for estimation using two sensors. *IEEE Transactions on Aerospace and Electronic Systems*, **41**, 1131-1137.
- [3] 李协, 张效义, 汪子嘉 (2012) 无线传感网中面向到达时差估计的数据压缩方法研究. *信号处理*, **9**.
- [4] 刘聪锋 (2011) 无源定位与跟踪. 西安电子科技大学出版社, 西安, 207-233.
- [5] 孙圣和, 陆哲明 (2002) 矢量量化技术及应用. 科学出版社, 北京, 53-61.