분산 자료에 대한 초완비 표현 방법

*이상철, *박종우, **곽칠성
*재능대학 컴퓨터로봇과, **재능대학 디지털정보전자과
e-mail : sangch@mail.jnc.ac.kr

A method of overcomplete representation for distributed data

*Sang-Cheol Lee, *Jong-Woo Park, **Chil-Seong Kwak
*Dept. of Computer Robot, **Dept. of Digital Information & Electronics
JaiNeung College

Abstract

This paper propose a method for representing distributed data of sensor networks. The proposed method is based on a general distributed regression framework that models sensor data by fitting a global function to each of the local measurements and explores the possible extensions of distribution regression by using complex signal representations. In order to reduce the amount of processed data and the required communication, the signal model has to be as compact as possible, with the ability to restore the arbitrary measurements. To achieve this requirement, data compression step is included, where the basis function set is changed to an overcomplete set. This have better advantages in case of nonstationary signal modeling than complete base representation.

I. 서론

이 논문은 신호압축을 위해, 분산 회기 체계(distributed regression framework)에 기반을 둔 센서 네트워크 자료의 표현법을 제안한다. 제안된 방법은 기본적으로 2-단계의 자료 축소 처리방법으로서, 분산 회기에 의해 계산된 자료 모델이 초완비 기저(overcomplete basis)를 사용함으로써 다시 한번 압축되는 형태이다.

초완비 기저 표현법은 필요 이상의 신호성분(signal components) 을 사용하여 신호를 분해(decompose)하는 기법이다 [1],[2]. 이때 사용되는 기저는 일반적으로 1개 이상의 직교 기저(orthogonal bases)와 몇 몇 여분의 함수들로 이루어진다. 이러한 신호 분해 기법을 사용하는 이유는, 웨이브렛 변환에서의 직교 기저 확장이 임의의 신호를 모델링 하기 어렵다는 단점을 극복하기 위한 것이다. 하지만 초완비 기저 표현에 기반을 둔 알고리즘은 과도한 계산량 때문에 센서 네트워크에 바로 적용하기에는 적합하지 않다. 결국 초완비 기저 표현을 센서 네트워크에 적용하기 위한 다소 현실적인 대안으로 약한 초완비 기저(weakly overcomplete basis)를 사용하는 기법이 연구되었다 [3].

II. 센서 네트워크에서의 분산 회기 체계

분산 회기 체계는 센서 네트워크 자료를 모델링하기 위한 체계이다 [4]. 이 체계에서 기저 함수 hi의 집합이 주어지고, 그 매개변수 wi가 계측 자료에 연속으로 적용된다. 이러한 모델링 과정은 다음 식(1)과 같다.

$$\hat{f}(t) = \sum_i w_i h_i(t) \approx f(t) \qquad (1)$$

여기서, f(t)는 계측 신호이다.
선형 회기법에 의한 최소 제곱 최적화 (Least Squares Optimization)는 다음 식 (2)와 같다.

$$w = (H^T H)^{-1} H^T f = A^{-1} b \qquad (2)$$

여기서, H는 행(row)이 계측 값이고, 열(column)이 기저 함수인 기저행렬이다. 또한, f는 각각의 계측 값들이 하나의 원소를 갖는 계측 벡터이다. 계측 값은 시간과 공간의 함수f(x,y,t) 이므로, 기저 함수 또한 시간과 공간의 함수이다. 하지만, 단순화하기 위해서, 시간 종속 함수와 공간 종속함수를 분리하고 공간 종속함수

는 선형이라고 하면 기저함수는 식 (3)과 같이 표현할 수 있다.

$$\sum_i h_i(x,y,t) = x + y + \sum_{i=1} h_i(t) \tag{3}$$

선형 회기 계산은 계산하기 편리하며, 커널 회기(kernel regression)를 사용함으로써 분산 시킬 수 있다. 여기서, 회기 모델은 A행렬을 희박한 행렬(sparse matrix)로 만들면서, 국부 기저 함수(local basis function)의 하중 합(weighted sum)의 형태를 취한다. 알고리즘은 가우스 소거법(Gaussian elimination)의 분산 시스템 응용을 사용함으로써 전역 선형 시스템인 Aw=b의 해를 구한다. 몇몇의 알고리즘에 의해 해를 구할 수 있는 동적 계획법(dynamic programming)문제 이다 [5].

III. 분산 회기 체계에서의 초완비 기저

초완비 표현법의 경우에, 잘 알려진 초완비 기저들로부터 최적화 방법에 의해 찾을 수 있는 최적 기저 성분의 개수로 신호가 분해된다. 알고리즘의 적절한 선택은 기저 추적 방법(basis pursuit method) [2]의 개념에 기반을 둔 약한 초완비 기저 표현법[3]을 사용하는 것이다.

초완비 기저는 완비기저 [h1, h2, ... , hN]과 추가 함수 h0로 구성된다. 여기서, h0=[g1, g2, ... , gN]이다. 이 방법의 입력은 time-windowed 자료로서, 첫 단계에 기저 [h1, h2, ... , hN]로 변환된다. 만일 값을 모르는 하중인자 c에 의해 여분의 기저함수를 고려한다면, 입력 신호는 다음 식 (4)와 같이 표현될 수 있다.

$$[ch0,\ h1-cg1,\ h2-cg2,\ \ldots,\ hN-cgN] \tag{4}$$

L1 노름 최소화가 영이 아닌 계수의 최소 개수를 보장한다는 것이 알려져 있기 때문에, 식(4)의 계수 중 최소한 하나는 영이어야만 한다. 따라서 c는 {0, h1/g1, h2/g2, ... , hN/gN}중에 하나의 값을 갖게 된다. 이 값들을 c에 대입하여 최소 노름을 갖는 것이 가장 간결한 표현을 제공하게 된다.

이 방법을 분산 회기 체계에서 2단계로 구현시킨다. 첫 번째 단계는 가우스 소거법 기저함수 h1, h2, ... , hN+2의 계수 w1, w2, ... , wN+2를 계산하는 것이다. 두 번째 단계는 앞에서 설명된 방법에 의해 초완비 풀이를 계산하는 것이다. 이것은 직교기저 h1, h2, ... , hN에 여분의 기저함수 h0를 추가하고, 계수 w0, w1, ... , wN을 다시 계산함으로써 이루어진다.

상대적으로 작은 값을 갖는 계수를 영으로 고려하여 영이 아닌 기저함수들 만으로 신호를 표현함으로써 신호를 압축한다. 만일 계측된 신호가 기저함수와 동일하다면, 영이 아닌 계수는 단 하나만 존재하게 되며 이상적인 압축이 가능하게 된다.

IV. 결론 및 향후 연구 방향

이 논문에서 분산 회기 체계를 확장하는 방법을 제안하였다. 센서로부터 계측된 신호를 표현하기 위해서 약한 초완비 기저를 사용하는 방법이다. 이 신호 표현법은 그 특성상 계측값의 압축된 형태로 표현할 수 있기 때문에 통신 용량 및 전력을 고려해야 하는 센서 네트워크에 대해 적절한 해법을 제공해 준다.

요구되는 통신 용량은 이 해법에 의해 상당히 감소되기는 하지만, 개개의 노드들은 여전히 많은 계산량을 갖게 된다는 단점이 있다. 다만 대부분의 전력 소모가 통신에 의해 결정된다는 점을 감안하면 제안된 방법이 타당함을 알 수 있다. 앞으로의 연구방향은 주문형 반도체 등에 의해 개개의 노드에서의 계산비용마저 최소화 시키는 것이다.

참고문헌

[1] S. Mallat, Z. Zhang, "Matching pursuit in a time-frequency dictionary", IEEE Trans. on Signal Processing, Vol. 41, pp. 3397-3415, 1993.

[2] S. S. Chen, D. L. Donoho, M. A. Saunders, "Atomic Decomposition by Basis Pursuit", SIAM Journal on Sientific Computing, Vol. 20, pp. 33-61, 1998.

[3] A. R. Varkonyi-Koczy, M. Fek, "Recursive Overcomplete Signal Representations", IEEE Trans. on Instrumentation and Measurement, Vol. 50, pp. 1698-1703, 2001.

[4] C. Guestrin, et al., "Distributed Regression: an Efficient Framework for Modeling Sensor Network data", Proceedings of the Third International Symposium on Information Processing in Sensor Networks(IPSN 2004), Berkeley, 2004.

[5] M. A. Paskin, G. D. Lawrence, "Junction Tree Algorithms for Solving Sparse Linear Systems", Technical Report UCB/CSD-03-1271, University of California, Berkeley, 2003.