

병렬구조 퍼지시스템을 이용한 카오스 시계열 데이터 예측 Chaotic Time Series Prediction using Parallel-Structure Fuzzy Systems

김민수 · 공성곤

Min-Soo Kim and Seong-Gon Kong

승실대학교 전기공학과

요 약

이 논문에서는 병렬구조 퍼지시스템(PSFS)에 기초한 카오스 시계열 데이터의 예측 알고리즘에 대해 연구하였다. 병렬구조 퍼지시스템은 병렬로 연결된 여러개의 퍼지시스템에 의하여 구성되어 있다. 병렬구조 퍼지시스템을 구성하고 있는 각 퍼지시스템은 다른 임베딩 차원과 시간지연을 가지고 과거의 데이터를 이용하여 동일한 데이터를 독립적으로 예측한다. 퍼지시스템은 입력력 데이터를 클러스터링하여 모델링되는 MISO Sugeno 퍼지규칙에 의하여 특징지어 진다. 각 퍼지시스템에 대한 최적 임베딩차원은 주어진 시간지연값에 대해서 최적의 성능을 갖도록 선정된다. 병렬구조 퍼지시스템은 각 구성요소 퍼지시스템들의 예측값 중에서 최대값과 최소값을 가지는 예측결과를 제외하고 나머지 값들을 평균하여 최종 예측결과를 얻는다.

ABSTRACT

This paper presents a parallel-structure fuzzy system(PSFS) for prediction of time series data. The PSFS consists of a multiple number of fuzzy systems connected in parallel. Each component fuzzy system in the PSFS predicts the same future data independently based on its past time series data with different embedding dimension and time delay. The component fuzzy systems are characterized by multiple-input single-output(MISO) Sugeno-type fuzzy rules modeled by clustering input-output product space data. The optimal embedding dimension for each component fuzzy system is chosen to have superior prediction performance for a given value of time delay. The PSFS determines the final prediction result by averaging the outputs of all the component fuzzy systems excluding the predicted data with the minimum and the maximum values in order to reduce error accumulation effect.

1. 서 론

어떤 시스템의 과거 행동에 근거하여, 그 시스템의 미래의 행동에 대하여 예측하는 것은 과학의 본질이다[1]. 예측방법에는 크게 두가지 접근방법이 있는데, 하나는 모델에 기초한 접근방법(model-based approach)이고 다른 하나는 통계적인 방법이다. 모델에 기초한 예측방법은 충분한 이전의 정보를 구할 수 있다면, 그것에 기초하여 예측을 위한 정확한 수학적 모델을 구성할 수 있다고 가정한다. 일단 의미있는 모델이 구성되면, 그 모델에 기초하여 시스템의 미래행동을 예측할 수 있다. 반면에 통계적인 예측방법은 어떤 시스템에 의하여 생성된 일련의 관측값들을 직접적으로 분석하여 미래의 행동을 예측하고자 한다. 모델에 기초한 접근방법으로는 사용되는 모델이 충분히 정확하지 않다면 의미있는 예측결과를 기대할 수 없다. 복잡한 시스템의 정확한 모델링은 실제적으로 많은 제약이 존재한다. 모델에 기초한 접근방법은 정확한 모델을 구성하는 것이 반드시 가능하지 않다는 어려움을 가지

고 있다. 비선형성이 강한 시스템은 많은 양의 자유도(degree of freedom)를 가진 아주 복잡하고 불규칙적인 신호를 발생시키는 경향이 있다. 통계적인 방법은 관측된 데이터의 통계적인 분석방법에 기초하고 있다. 통계적인 방법으로 실제 시스템들의 충분한 복잡도를 표현하는 것은 불가능하지만, 정확한 수학적인 모델을 구성하는 어려움 때문에 많은 현대 예측이론들이 통계적인 접근방법에 기초하여 개발되어 왔다.

많은 실제적인 시스템들로부터 발생하는 신호는 물리적 시스템의 고유의 비선형적인 특성 때문에 카오스적인 행동을 나타낸다[2]. 카오스 신호는 자연현상과 비슷한 특성으로 초기 조건 및 작은 오차에 민감하여 장기예측이 어려운 신호이다. 카오스 시계열 데이터가 초기값의 변동에 따라서 민감하게 영향을 받으므로 카오스 시계열 예측은 매우 어려운 문제이다. 카오스 신호가 작은 변동에 대해서 매우 큰 차이를 보이므로 카오스 시계열 데이터의 장기예측은 거의 불가능하다는 것이 잘 알려져 있다. 모델에 근거한 접근방법이나 통계적 특성에 기초한 접근 방법 모두 카오

스 신호를 정확히 예측한다는 것은 불가능하나 통계적 특성에 기초해 선형예측 방법이나 신경망을 이용한 학습 방법[3], 적응 알고리즘을 이용한 방법[4] 등을 이용하여 장기 예측 보다는 단기 예측이나 경향 분석에 중점을 두고 예측하고 있다.

이 논문에서는 병렬구조 퍼지시스템(Parallel-Structure Fuzzy System)을 사용하여 카오스 시계열 데이터의 예측에 적용하였다. 이 접근방법은 시스템 출력의 과거 관측 데이터에 기초하여 예측결과를 발생하므로 통계적인 접근방법에 해당된다. 병렬구조 퍼지시스템은 병렬로 연결된 여러 개의 구성요소 퍼지시스템들에 의하여 구성되어 있다. 병렬구조 퍼지시스템에서 각 구성요소 퍼지시스템은 다른 임베딩차원과 시간지연값을 가지고 과거 시계열 데이터에 기초하여 독립적으로 동일한 미래 데이터를 예측한다. 임베딩차원은 각 구성요소 퍼지시스템의 입력데이터의 개수를 결정한다. 시간지연에 따라서 각 퍼지시스템은 다른 시간 간격을 가지는 입력값들을 받아들인다. 병렬구조 퍼지시스템은 각 구성요소 퍼지시스템들이 예측한 결과들 중에서 2가지 극한값, 즉 최대값과 최소값을 제외시키고 나머지 예측결과들을 평균하여 최종 예측결과를 생성함으로써 오차 누적효과를 감소시킨다.

각 구성요소 퍼지시스템은입출력 학습데이터를 클러스터링하여 얻어지는 적은 수의 다입력단출력(MISO) Sugeno 퍼지규칙[5]을 포함한다. 퍼지시스템은 복잡한 물리적 시스템의 행동에 포함되어 있는 불확실성을 기존의 예측 알고리즘[6,7]에 비해 더 수월하게 표현할 수 있다. 많은 경우에 있어서, 퍼지규칙들은 인간 전문가의 경험 또는 공학적인 상식에 의하여 결정된다. 그러나 체계적인 지식이 존재하지 않고, 많은 실제 경우에서처럼 수치적인 입출력 데이터만이 존재하는 경우에는 적응 클러스터링 알고리즘[8]들이 자동적으로 퍼지규칙을 발생시키는 데 사용될 수 있다.

Subtractive 클러스터링 알고리즘[9,10]은 카오스 시계열 데이터의 예측을 위한 퍼지규칙을 발생시키는 데 이용된다. 병렬구조 퍼지시스템의 입력들은 시간지연과 임베딩차원에 의해서 결정된다. 각 주어진 시간지연에 대해서 최적 임베딩차원은 1에서부터 10까지 순차적으로 1씩 증가시키면서 한단계이후 예측결과와 실제 데이터와의 평균제곱오차(MSE)와 최대절대오차(MAE)에 의한 오차성능척도를 최소화시키는 값으로 구해진다. 이 과정은 시간지연을 1부터 N, 즉 구성요소 퍼지시스템의 개수까지 증가시키면서 반복한다. 병렬구조 퍼지시스템의 최종 예측결과는 모든 구성요소 퍼지시스템에 의하여 예측된 예측결과들을 평균하여 발생시킨다. 이때 각 구성요소 퍼지시스템의 예측결과

들 중에서 최대값과 최소값을 제외함으로써 오차누적효과를 감소시킨다. 시뮬레이션을 통하여 5개의 구성요소 퍼지시스템을 가진 병렬구조 퍼지시스템이 하나의 퍼지시스템에 비해 더 정확하게 카오스 시계열 데이터를 예측할 수 있었음을 확인하였다.

2. 카오스 시계열 데이터의 예측

통계적 접근방법에 기초한 카오스 시계열 데이터의 예측방법은 시계열 데이터의 미래 행동이 과거 데이터의 함수적 표현에 의해서 나타낼 수 있다는 가정에 기초하고 있다. 어떤 시계열 데이터 $x(k)$ 의 k 번째 시점에서 m 개의 과거 데이터가 주어졌을 때, τ 이후 시계열 데이터 $x(k + \tau)$ 는 다음 식과 같이 표현된다.

$$x(k + \tau) = P[x(k), x(k - \tau), x(k - 2\tau), \dots, x(k - (m - 1)\tau)] \quad (1)$$

여기에서 $P[\cdot]$ 는 통계적인 시계열 데이터의 예측과정의 입출력 관계를 나타내는 함수를 의미한다. 양의 정수 m 은 임베딩 차원이라고 하며, τ 는 시간지연을 나타낸다. 시계열 데이터의 미래값은 $m\tau$ 개의 과거 데이터의 선형 또는 비선형 함수 $P[\cdot]$ 의 출력에 의해서 예측될 수 있다.

시계열 데이터의 예측방법은 이전 단계에서 예측된 값이 이후의 시점에서 다시 입력으로 이용되는가에 따라 한 단계이후 예측(one-step-ahead prediction)과 단기예측(short-term prediction) 또는 장기예측(long-term prediction)으로 구분된다. 한 단계이후 예측에서는 식 (2)와 같이 예측된 값 $\hat{x}(k + \tau)$ 이 시간지연이 τ 인 m 개의 입력 데이터에 의해서 표현된다.

$$\hat{x}(k + \tau) = P[x(k), x(k - \tau), x(k - 2\tau), \dots, x(k - (m - 1)\tau)] \quad (2)$$

단기예측이나 장기예측에서는 이전에 예측된 값이 이후의 단계에서 다시 입력으로 사용된다. 식 (3)은 시계열 데이터의 단기예측 또는 장기예측 과정을 나타낸다. 미래 데이터 $\hat{x}(k + \tau)$ 는 과거에 예측된 데이터 $\hat{x}(k), \hat{x}(k - \tau), \hat{x}(k - 2\tau), \dots, \hat{x}(k - (m - 1)\tau)$ 에 의해서 표현된다. 작은 초기의 오차가 미래값에서 커다란 오차 누적효과를 나타내므로 과거에 예측된 데이터에 기초하여 카오스 시계열 데이터를 장기예측하는 것은 매우 어려운 과제이다.

$$\hat{x}(k + \tau) = P[\hat{x}(k), \hat{x}(k - \tau), \hat{x}(k - 2\tau), \dots, \hat{x}(k - (m - 1)\tau)] \quad (3)$$

최적 임베딩 차원 m 은 특정한 시간지연 τ 에 대해서

결정된다. 시간지연 τ 값이 주어졌을때 학습 데이터와 검증(validation) 데이터에 대한 오차 성능척도를 계산한다. 검증 데이터는 학습에 사용되지는 않지만 학습 결과가 제대로 얻어졌는지 나타내는 데 사용된다. 오차 성능척도는 한 단계이후 예측값과 실제 결과와의 차이에 의해서 계산되는 평균제곱오차(MSE)와 최대절대오차(MAE)로 정하였다. 주어진 시계열 데이터에 대해서 MSE와 MAE는 고정된 시간지연에 대해서 가능한 임베딩 차원값을 증가시키면서 계산된다. 이 과정은 학습데이터와 검증데이터에 대해서 반복된다. 최적 임베딩 차원은 학습데이터 및 검증데이터 모두에 대해서 오차 척도가 최소가 되는 값에 해당된다.

3. 병렬구조 퍼지시스템

3.1 병렬구조 퍼지시스템의 구성

병렬구조 퍼지시스템은 다른 수의 그리고 다른 샘플들에 기초한 여러 가지 예측 메카니즘을 이용하여 미래 데이터를 예측한다. 병렬구조 퍼지시스템은 병렬로 연결된 여러개의 구성요소 퍼지시스템으로 구성되어 있다. 그림 1은 병렬구조 퍼지시스템의 구조를 나타낸다. 병렬구조 퍼지시스템은 N 개의 구성요소 퍼지시스템 FS_1, FS_2, \dots, FS_N 에 의해서 구성되어 있다. 각 구성요소 퍼지시스템은 과거의 데이터에 기초하여 독립적으로 시점 $k+r$ 에서의 데이터 $\hat{x}(k+r)$ 을 예측한다. 병렬구조 퍼지시스템은 N 개의 구성요소 퍼지시스템의 예측결과 $\hat{x}_1(k+r), \hat{x}_2(k+r), \dots, \hat{x}_N(k+r)$ 에 의

하여 최종 예측값을 생성한다.

병렬구조 퍼지시스템에 의한 시계열 데이터 예측은 2개의 파라미터 τ 와 m 에 의해 특징 지어진다. 임베딩 차원 m 은 각 구성요소 퍼지시스템에 들어가는 입력의 수를 정의하고, 시간지연 τ 는 각 구성요소 퍼지시스템에 들어가는 입력 데이터의 간격을 정의한다. 시간지연 τ 는 1부터 N , 즉 구성요소 퍼지시스템의 개수까지의 값을 가질 수 있다. 각 구성요소 퍼지시스템의 최적 임베딩 차원은 주어진 시간지연에 대해서 계산된다. 예를들어 N 개의 구성요소 퍼지시스템을 가지는 병렬구조 퍼지시스템을 고려해 보자. i 번째 구성요소 퍼지시스템 FS_i 에 대한 임베딩차원을 m_i , 시간지연을 τ_i 라고 하면, 퍼지시스템 FS_i 는 (τ_i, m_i) 에 의해서 특징 지어진다. 이 논문에서는 r 을 1로 하였고, 퍼지시스템 FS_i 는 시간지연을 $\tau_i = i$ 로 정하였다. 일반적으로 각 구성요소 퍼지시스템은 임의로 다른 시간지연을 가질 수 있다. 따라서 퍼지시스템 FS_i 는 다음 식 (4)와 같이 m_i 개의 과거 데이터에 의해서 $\hat{x}_i(k+1)$ 을 예측한다.

$$\hat{x}_i(k+1) = P[\hat{x}(k+1-i), \hat{x}(k+1-2i), \dots, \hat{x}(k+1-m_i)] \quad (4)$$

그림 2는 병렬구조 퍼지시스템이 데이터 $\hat{x}(k+1)$ 을 예측하는데 사용된 입력 데이터의 예를 나타낸다. 이 병렬구조 퍼지시스템은 3개의 퍼지시스템 FS_1, FS_2, FS_3 으로 구성되어 있다. 각 퍼지시스템의 임베딩 차원 m 은 시간지연 1, 2, 및 3에 대해 각각 3, 4, 및 3으로 계산되었다. 따라서 이 병렬구조 퍼지시스템은 3개의 파라미터 쌍 (τ_i, m_i) 은 (1,3), (2,4), 그리고 (3,3)가 된다. 이것은 구성요소 퍼지시스템 FS_1 이 데이터 $\hat{x}(k), \hat{x}(k-1), \hat{x}(k-2)$ 에 의해서 예측결과 $\hat{x}_1(k+1)$ 을 만들어 내고, FS_2 는 $\hat{x}(k-1), \hat{x}(k-3), \hat{x}(k-5)$, 그리고 $\hat{x}(k-7)$ 에 의해 예측결과 $\hat{x}_2(k+1)$ 을 생성한다. 그리고 FS_3 는 $\hat{x}(k-2), \hat{x}(k-5)$, 그리고 $\hat{x}(k-8)$ 에 의해 예측결과 $\hat{x}_3(k+1)$ 을 결정한다. 병렬구조 퍼지시스템에서는 각 퍼지시스템의 예측결과 $\hat{x}_1(k+1), \hat{x}_2(k+1), \hat{x}_3(k+1)$ 의 평균값을 계산하여 최종 예측결과 $\hat{x}(k+1)$ 로 정한다. 일반적으로 N 개의 구성요소 퍼

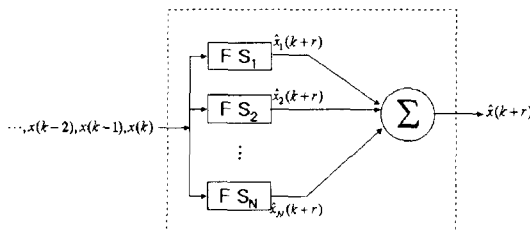


그림 1. 병렬구조 퍼지시스템의 구성

Fig. 1. Structure of the parallel-structure fuzzy system(PSFS)

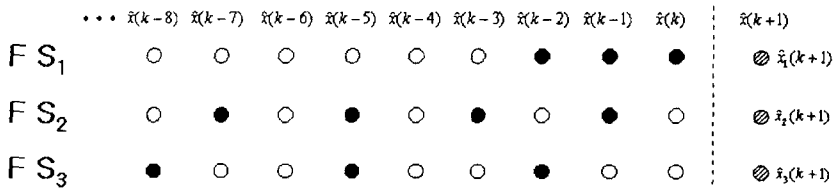


그림 2. $\hat{x}(k+1)$ 의 예측을 위한 병렬구조 퍼지시스템의 입출력 관계

Fig. 2. Inputs of each component fuzzy system in the PSFS for prediction of $\hat{x}(k+1)$

지시스템이 있을 때, 각 퍼지시스템의 예측결과 $\hat{x}_1(k+1), \hat{x}_2(k+1), \dots, \hat{x}_N(k+1)$ 을 평균하여 최종 예측결과 $\hat{x}(k+1)$ 은 각 구성요소 퍼지시스템에 의한 예측결과들의 평균에 해당한다.

$$\hat{x}(k+1) = \frac{1}{N-2} \left[\sum_{i=1}^N \hat{x}_i(k+1) + \max_i \hat{x}_i(k+1) - \min_i \hat{x}_i(k+1) \right] \quad (5)$$

일반적으로 $(k+r)$ 번째 시점에서의 r -단계 이후 예측값 $\hat{x}(k+r)$ 을 예측하기 위한 병렬구조 퍼지시스템을 고려한다. r 번째 구성요소 퍼지시스템은 과거 데이터 $\hat{x}(k+1-i), \hat{x}(k+1-2i), \dots, \hat{x}(k+1-m_i)$ 에 기초하여 예측결과 $\hat{x}_i(k+r)$ 을 생성한다. 병렬구조 퍼지시스템은 모든 구성요소 퍼지시스템의 출력들 $\hat{x}_1(k+r), \hat{x}_2(k+r), \dots, \hat{x}_N(k+r)$ 중에서 최대값 및 최소값을 가지는 예측결과를 제외한 나머지 결과들을 평균하여 최종 출력을 결정한다. 단기예측 또는 장기예측에서는 작은 양의 예측오차가 누적되어 여러 반복횟수이후에는 매우 큰 오차값이 된다. 병렬구조 퍼지시스템은 양 극단값을 제외한 모든 구성요소 퍼지시스템의 출력들을 평균함으로써 오차 누적효과를 감소시킨다.

$$\hat{x}(k+r) = \frac{1}{N-2} \left[\sum_{i=1}^N \hat{x}_i(k+r) + \max_i \hat{x}_i(k+r) - \min_i \hat{x}_i(k+r) \right] \quad (6)$$

여기서

$$\hat{x}_i(k+r) = P[\hat{x}(k+r-i), \hat{x}(k+r-2i), \dots, \hat{x}(k+r-m_i)] \quad (7)$$

이고, $i=1, \dots, N$ 이다. 임베딩 차원 m 은 구성요소 퍼지시스템의 입력의 수를 결정하고, 각 퍼지시스템에 대한 입력은 시간지연에 의해서 특징지어진다. 시계열 예측을 위한 병렬구조 퍼지시스템을 설계하기 위해서는 특정한 시간지연에 대해서 임베딩차원을 결정하여야 한다. 주어진 시간지연에 대해서 오차 성능척도는 한 단계 이후 예측결과와 실제 데이터와의 차이로부터 계산된다. 각각 선택된 시간지연에 대해서 최적 임베딩차원은 1부터 10사이에서 예측오차의 MSE와 MAE가 일정한 값에 해당한다. 가장 작은 MSE와 MAE를 가지는 상수 m 이 주어진 시간지연에 대한 최적 임베딩차원으로 결정된다. 학습데이터는 퍼지시스

템을 모델링하는데 사용되었고, 검증데이터는 고정된 시간지연에 대해서 최적 임베딩차원을 결정하기 위해서 사용된다. 이 과정은 τ 가 1부터 N 까지 변화할 때까지 반복된다. 시뮬레이션에서 구성요소 퍼지시스템은 5로 선택하였다. 구성요소 퍼지시스템의 수가 증가하면 보다 정확한 예측결과를 얻을 수 있으나 계산량이 증가하게 된다. 병렬구조 퍼지시스템은 최대 및 최소 예측결과를 제외하고 평균함으로써 최종 예측결과를 생성한다.

3.2 구성요소 퍼지시스템

퍼지 시스템은 전문가의 경험이나 지식을 “IF <전건부> THEN <후건부>” 형태의 언어 규칙을이용하여 불확실한 조건하에서 추론결과를 생성한다. 퍼지시스템 모델링은 주어진 학습데이터에 대해서 구조와 내부 파라미터를 식별하는 과정을 포함한다. Sugeno 퍼지모델[5]에서는 퍼지규칙의 전건부 변수는 퍼지집합으로, 후건부 변수는 입력변수들의 함수형태로 표현시킨다. 이 방법은 Mamdanin 퍼지모델[6]과는 달리 후건부를 입출력 변수들의 선형 또는 비선형 함수로 표현함으로써 기존의 근사화 방식에 비해 적은 수의 규칙으로도 비선형 특성을 갖는 입출력 관계를 정확하게 표현할 수 있다는 장점을 가진다. n 개의 입력변수 x_1, \dots, x_n 과 1개의 출력변수 y 를 갖는 Sugeno 퍼지 모델에서 각 규칙은 퍼지분할의 입출력 관계에 해당하며 다음 식과 같이 표현된다.

$$\text{IF } x_1 \text{ is } A_{i1} \text{ and } \dots \text{ and } x_n \text{ is } A_{in} \text{ THEN } y_i = f_i(x_1, \dots, x_n) \quad (8)$$

여기서 $i=1, \dots, M$ 이고 A_{ij} 는 멤버쉽함수가 $A_{ij}(x_j)$ 인 언어라벨이고, $f_i(\cdot)$ 는 입력과 출력을 대응시키는 함수를 나타낸다. M 은 퍼지시스템의 규칙의 수를 나타낸다. Sugeno 퍼지모델은 실제 과정에서 언어지는 수치적인 입출력 데이터로부터 퍼지규칙을 쉽게 발생시킬 수 있다. $A_{ij}(x_j)$ 는 i 번째 퍼지규칙에서 입력변수 x_j 에 할당된 퍼지 멤버쉽 함수이며, Gaussian 함수의 형태일 경우에는 식 (9)와 같이 표현된다. y_i 는 i 번째 퍼지규칙의 출력을 나타내며, 식 (10)처럼 입력변수들의 선형결합으로 표현될 수 있다.

$$A_{ij}(x_j) = \exp\left(-\frac{1}{2} \left(\frac{x_j - c_{ij}}{w_{ij}}\right)^2\right) \quad (9)$$

$$f_i(x_1, \dots, x_n) = a_{0i} + a_{1i}x_1 + \dots + a_{ni}x_n \quad (10)$$

이때 c_{ij} 는 Gaussian 멤버쉽함수 $A_{ij}(x_j)$ 의 중심의 위치, 그리고 w_{ij} 는 $A_{ij}(x_j)$ 의 폭을 나타내는 파라미터이

며, 계수 $a_{0i}, a_{1i}, \dots, a_{mi}$ 는 입력력 학습데이터로부터 최적화 과정을 통하여 결정된다. 단순화된 추론방법 [7]에서는 M 개의 규칙을 가지는 퍼지시스템의 출력 y 는 다음과 같이 표현된다.

$$y = \frac{\sum_{i=1}^M \mu_i f_i(x_1, \dots, x_n)}{\sum_{i=1}^M \mu_i} \quad (11)$$

여기서 μ_i 는 i 번째 규칙의 전건부 적합도를 나타낸다.

$$\mu_i = \prod_{j=1}^n A_{ij}(x_j) \quad (12)$$

$$= \exp\left(-\frac{1}{2} \sum_{j=1}^n \left(\frac{x_j - c_{ij}}{w_{ij}}\right)^2\right) \quad (13)$$

Sugeno 퍼지모델에서는 후건부 변수가 입력변수들의 함수형태로 표현되므로 다음과 같은 이점을 가지고 있다. 첫째, 전문가의 지식이나 공학적인 상식으로 부터 규칙을 추출하는 복잡한 과정은 주어진 입력력 데이터로부터 계수 $a_{0i}, a_{1i}, \dots, a_{mi}$ 를 최적화하는 작업으로 귀결되어 쉽게 퍼지규칙을 생성시킬 수 있게 된다. 둘째, 후건부에 선형관계식을 사용함으로써 적은 제어규칙으로도 실제 시스템의 복잡한 비선형 특성을 나타낼 수 있다.

3.3 클러스터링에 기초한 구성요소 퍼지시스템의 모델링

구성요소 퍼지시스템의 파라미터는 클러스터링 알고리즘을 사용하여 구할 수 있다. Subtractive 클러스터링 알고리즘[13]은 각 데이터에서 다른 데이터들과의 거리에 반비례하는 함수로 정의되는 포텐셜값을 계산하고, 가장 큰 포텐셜 값을 가지는 데이터를 클러스터 중심 $x_i^* = (x_{1i}^*, \dots, x_{mi}^*)$ 으로 결정한다. 포텐셜 값은 데이터 사이의 거리에 반비례하므로, 데이터가 밀집되어 있으면 포텐셜 값이 증가하고, 따라서 클러스터 중심이 많이 분포한다는 것을 의미한다. 포텐셜 값이 가장 큰 데이터를 첫 번째 클러스터의 중심으로 결정하고, 이후의 클러스터 중심은 방금 구한 클러스터 중심의 영향을 제거한 후에 결정한다. 이러한 과정은 포텐셜 값이 미리 지정한 값 이하가 될 때까지 반복한다.

입력이 n 차원 벡터 x_1, x_2, \dots, x_m 이고 출력이 1차원 y_1, y_2, \dots, y_m 라고 하면, 입력력 데이터는 $(n+1)$ 차원 입력력 공간 상에서 m 개의 입력력 벡터 $X_1, X_2,$

\dots, X_m 으로 표시할 수 있다. 이러한 데이터에 대해서 Subtractive 클러스터링 알고리즘에 의한 데이터 클러스터링 과정, 즉 클러스터 중심의 결정방법은 다음과 같다.

Step 1: 모든 데이터를 $[0,1]$ 로 정규화 한다.

Step 2: 각 데이터마다 포텐셜 P_i 를 계산한다.

$$P_i = \sum_{j=1}^m \exp(-\alpha \|X_i - X_j\|^2), \quad i = 1, 2, \dots, m \quad (14)$$

여기서 $\alpha = 4/r_a^2$ 이며, 포텐셜 값에 영향을 주는 데이터의 간격을 결정한다. 즉 반경 $r_a < 1$ 밖에 존재하는 데이터는 포텐셜 값에 실질적으로 영향을 주지 못한다.

Step 3: P_i 에서 가장 큰 포텐셜 값을 P_1^* 이라 하고 이것에 해당하는 데이터를 첫번째 클러스터 중심 X_1^* 으로 정한다.

Step 4: 첫 번째 클러스터 중심의 영향을 제거한 포텐셜 값 P_i' 을 계산한다.

$$P_i' = P_i - P_1^* \exp(-\beta \|X_i - X_1^*\|^2) \quad (15)$$

여기서 $\beta = 4/r_b^2$ 이며, $r_b > r_a$ 로 하여 두 번째 클러스터 중심이 첫 번째 클러스터 중심의 근처에 생기지 않도록 함으로써 클러스터 중심들이 밀집되는 것을 방지한다. 첫 번째 클러스터 중심 근처에는 많은 데이터들이 몰려있게 되므로 이와같이 방금 구한 클러스터 중심의 영향을 제거하는 과정이 필요하다.

Step 5: P_i' 에서 가장 큰 포텐셜 값을 P_2^* 이라 하고 이것에 해당하는 데이터를 두 번째 클러스터 중심 X_2^* 으로 정한다. 일반적으로 k 번째 클러스터 중심 X_k^* 의 영향을 제거한 포텐셜 값 P_i' 을 계산하고 가장 큰 포텐셜 값에 해당하는 데이터를 클러스터 중심 X_{k+1}^* 으로 정한다.

$$P_i' = P_i - P_k^* \exp(-\beta \|X_i - X_k^*\|^2) \quad (16)$$

Step 6: 계산된 클러스터 중심을 받아들일 수 있는지 결정한다. 만일 $P_k^*/P_1^* \geq \bar{\epsilon}$ 이거나 $P_k^*/P_1^* > \underline{\epsilon}$ 이고 $d_{\min}/r_a + P_k^*/P_1^* \geq 1$ 이면 클러스터 중심으로 결정하고 Step 5를 반복한다. 여기서 d_{\min} 은 이제까지 결정된 클러스터 중심들 $X_1^*, X_2^*, \dots, X_k^*$ 중에서 가장 짧은 거리를 나타낸다. 만일 $P_k^*/P_1^* > \underline{\epsilon}$ 이고 $d_{\min}/r_a + P_k^*/P_1^* < 1$ 이면 그 때의 X_k^* 를 0으로 하고 다음으로 높은 포텐셜 값을 선정한 후, $d_{\min}/r_a + P_k^*/P_1^* \geq 1$ 이면 그 데이터를 새로운 클러스터 중심으로 결정하고 Step 5를 반복한다. 만일 $P_k^*/P_1^* \leq \underline{\epsilon}$ 이면 중단한다. 클러스터 중심을 결정하는 과정에서 상하 기준 $\bar{\epsilon}$ 과 $\underline{\epsilon}$ 을 뚫으로써 포텐셜 값이 낮더라도 d_{\min} 이 크면 클러스터 중심으로 선정되도록 한다.

클러스터링 과정을 통하여 입력력 공간에서 M 개의 클러스터 중심 $X_1^*, X_2^*, \dots, X_M^*$ 을 결정하고 이것으로부터 M 개의 규칙을 갖는 퍼지시스템을 설계하는 과정은 다음과 같다. 클러스터 중심을 입력 및 출력 공간으로 분할하여 분할된 입력 공간은 전건부 퍼지집합으로 매핑한다. $n + 1$ 차원 클러스터 중심 X_i^* 에서 음 n 개의 값은 n 차원 입력공간의 $x_i^* = (x_{i1}^*, \dots, x_{in}^*)$ 로 분할되고 각 원소값은 전건부 퍼지규칙들이 가지는 멤버십함수의 중심을 결정한다. 이 논문에서는 Gaussian 함수 형태의 멤버십 함수를 사용하였으므로 $c_{ij} = x_{ij}^*$ 가 된다. 멤버십 함수의 폭 w_{ij} 는 다음과 같이 일정한 값으로 고정한다.

$$w_{ij} = \frac{r_a \|\max_i(x_i^*) - \min_i(x_i^*)\|}{\sqrt{M}} \quad (17)$$

여기에서 M 은 클러스터 중심의 수, 그리고 $\|\max_i(x_i^*) - \min_i(x_i^*)\|$ 는 클러스터 중심들의 거리중에서 최대값을 나타낸다. 클러스터 중심의 수는 퍼지규칙의 수에 해당한다. 다음은 Sugeno 퍼지모델의 i 번째 규칙의 후건부 출력 y_i 를 계산하기 위한 퍼지규칙의 후건부 파라미터 $a_{0i}, a_{1i}, \dots, a_{ni}$ 를 최적화하는 과정에 대해 살펴본다. 분할된 출력 공간은 다음과 같이 입력변수들의 선형 함수로 정의된다.

$$y_i = a_{0i} + a_{1i}x_1 + a_{2i}x_2 + \dots + a_{ni}x_n \quad (18)$$

$$= [a_{1i}a_{2i} \dots a_{ni}] \mathbf{x} + a_{0i} \quad (19)$$

$$= \mathbf{a}_i^T \mathbf{x} + a_{0i} \quad (20)$$

파라미터 \mathbf{a}_i 와 a_{0i} 를 선형 최소자승법(linear least-squares estimation)에 의해 계산하면, Sugeno 퍼지모델의 최종출력 y 는 다음과 같이 구할 수 있다.

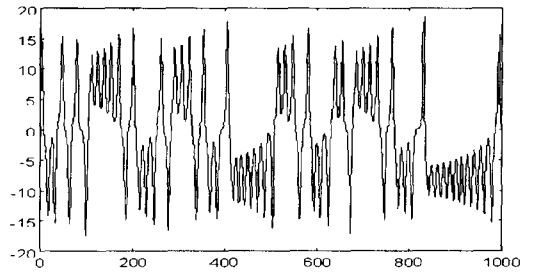
$$y = \frac{\sum_{i=1}^M \mu_i (\mathbf{a}_i^T \mathbf{x} + a_{0i})}{\sum_{i=1}^M \mu_i} \quad (21)$$

이 값은 각 구성요소 퍼지시스템의 최종 출력이 된다.

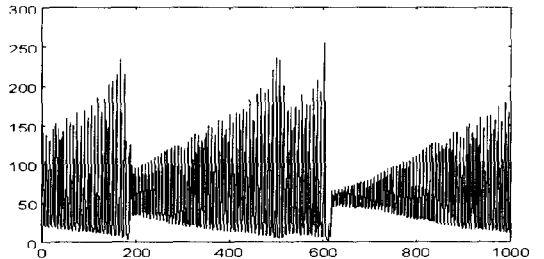
4. 시뮬레이션

4.1 카오스 시계열 데이터

병렬구조 퍼지시스템을 이용한 시계열 예측은 두가지 시계열 데이터, Lorenz 시계열 데이터[1]와 laser에서 발생된 신호[12]를 이용하여 시뮬레이션하였다. Lorenz 시계열 데이터는 다음 식으로 주어지는



(a) Lorenz 시계열 데이터(x-성분)



(b) Laser 시계열 데이터

그림 3. Lorenz와 laser 시계열 데이터의 모양
Fig. 3. The Lorenz and the Laser time series data

Lorenz 방정식의 x -성분 해를 사용하였다.

$$\begin{aligned} \dot{x} &= a(y - x) \\ \dot{y} &= bx - y - xz \\ \dot{z} &= -cz + xy \end{aligned} \quad (22)$$

여기서 각 파라미터는 $a = 10, b = 20, c = 20$ 으로 하였고, 샘플링 간격은 0.05이다. 그림 3은 병렬구조 퍼지시스템을 이용한 시계열 예측에서 사용된 x -성분 Lorenz 시계열 데이터와 laser 시계열 데이터 1000샘플을 나타낸다.

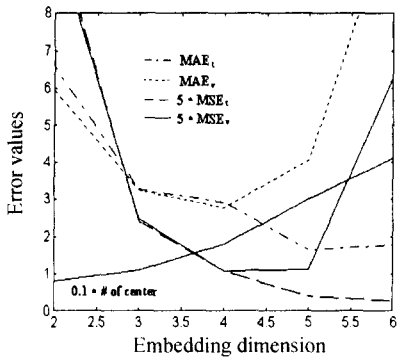
퍼지시스템 모델링을 위한 클러스터링에서 사용된 파라미터들은 $r_a = 0.3, r_b = 0.75, \bar{\epsilon} = 0.3$, 그리고 $\epsilon = 0.1$ 이다. 5개의 구성요소 퍼지시스템을 가진 병렬구조 퍼지시스템($N = 5$)을 데이터가 1,000 및 10,000개인 경우에 대해 적용하였다. 병렬구조 퍼지시스템에 의한 예측결과는 최적의 임베딩차원과 시간지연을 가지는 하나의 퍼지시스템에 의한 예측결과와 비교하였다. 데이터의 수가 1,000인 경우 500개의 데이터는 병렬구조 퍼지시스템을 학습시키는데, 300개 데이터는 검증 을 하는데, 그리고 나머지 200개의 데이터는 결과를 확인하는데 사용하였다. 10,500개의 데이터를 사용하는 경우 학습에 10,000개를 사용하였고, 300개는 검증에, 그리고 200개의 데이터를 예측결과 확인에 사용하였다.

4.2 단일 퍼지시스템에 의한 예측

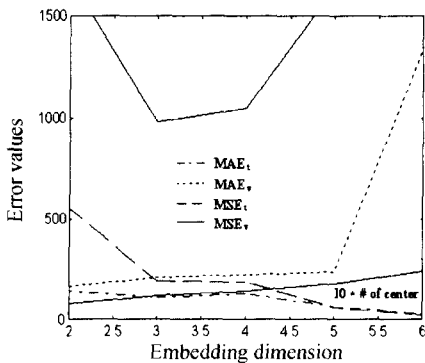
시계열 데이터 예측을 위한 병렬구조 퍼지시스템을 구성하기 위해서는 우선 특정한 시간지연 τ 에 따른 임베딩차원 m 이 결정되어야 한다. 주어진 τ 에 대해서 학습데이터와 검증데이터로 학습된 한 단계이후 예측값과 실제값의 차이에 의하여 오차 성능 척도를 계산한다. 임베딩 차원값을 1부터 시작하여 1씩 증가시키면서 오차 평가척도 값이 최소가 되는 점을 최적 임베딩차원으로 결정한다. 오차 평가지표는 학습데이터와 검증데이터의 MSE와 MAE를 사용한다.

그림 4는 주어진 시간지연 τ 에 대해서 최적 임베딩차원을 결정하는 방법을 나타낸 것이다. m 값을 차례로 변화시키면서 학습데이터 및 검증데이터에 적용하여 구한 MAE_L , MSE_L , MAE_V , MSE_V 를 각각 나타내었다. Lorenz 시계열 데이터인 경우 τ 가 4일때 m 가 4에서 평가지표가 최소값을 가지며 Laser 시계열 데이터인 경우 τ 가 3일 때 m 가 3에서 최소임을 알 수 있다.

이와 같은 방법으로 구한 파라미터 쌍을 (τ, m) 으



(a) Lorenz



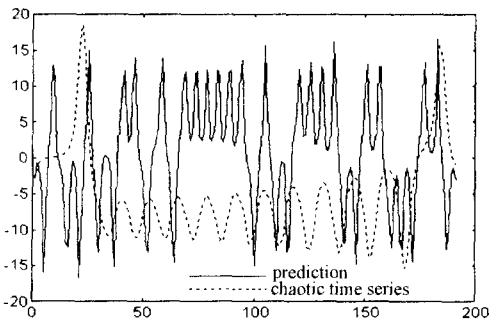
(b) Laser

그림 4. 최적 임베딩차원의 결정

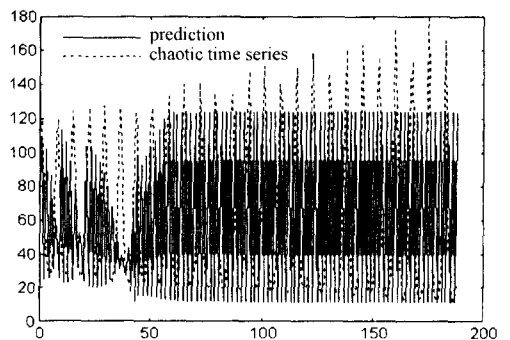
Fig. 4. Determination of the optimal embedding dimension

로 표현하였을 때, Lorenz 시계열 데이터는 (1,3), (2,3), (3,3), (4,4) 그리고 (5,5)로 구해졌으며, Laser 시계열 데이터에서는 (1,2), (2,4), (3,3), (4,3), 그리고 (5,3) 이었다. 이렇게 구한 시간지연과 임베딩차원의 쌍 중에서 Lorenz 시계열 데이터에서의 (4,4)와 Laser 시계열 데이터에서 (3,3)은 Fraser[13]가 제안한 상호 정보량(mutual information)을 근거로 하여 설정한 시간지연과 일치하며 False nearest neighbor (FNN) 알고리즘을 적용하여 구한 최소 임베딩 차원과 일치됨을 알 수 있었다.

그림 5는 최적의 임베딩 차원과 시간지연을 사용한 단일 퍼지시스템의 예측결과를 나타낸다. 고정된 τ 값에 대한 최적 임베딩차원 m 은 학습 및 검증데이터에 대한 한 단계이후 예측값과 실제값과의 오차 척도 MSE와 MAE가 가장 작은 값에 해당한다. 검증데이터는 퍼지시스템을 구성하는데 사용하지 않은 데이터로서 한 단계이후 예측에 적용하여 최적 임베딩차원을 결정하는데 사용하였다. 그림 5(a)는 $m=4$ 와 $\tau=4$ 일 때 Lorenz 시계열 데이터에 대해 단일 퍼지시스템을 이용하여 예측한 결과이다. 그림 5(b)는 $m=3$ 과 $\tau=3$ 일 때 Laser 시계열 데이터에 대해서 구한 단일



(a) Lorenz ($\tau=4, m=4$)



(b) Laser ($\tau=3, m=3$)

그림 5. 단일 퍼지시스템에 의한 예측결과

Fig. 5. Prediction result of a single fuzzy system

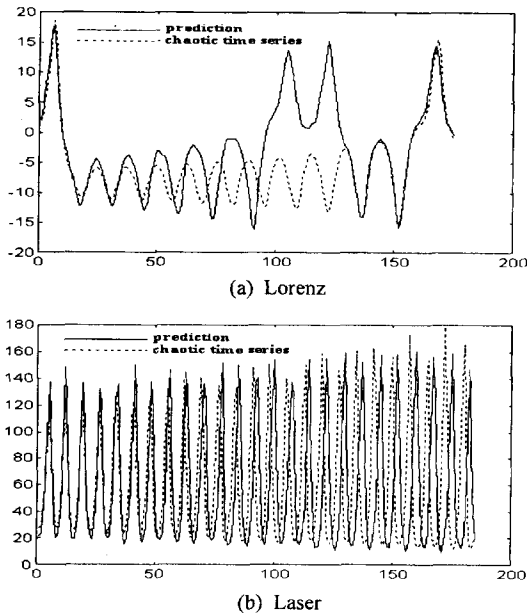


그림 6. 500개의 학습데이터를 이용하여 학습한 병렬구조 퍼지시스템의 예측결과
 Fig. 6 Prediction result of the PSFS with 500 training samples

퍼지시스템의 예측결과이다. 단일 퍼지시스템은 만족스러운 예측결과를 생성하지 못했으며, 카오스 시계열 데이터 예측에 적당하지 않았다. 시뮬레이션 결과를 통하여 단일 퍼지시스템은 최적의 임베딩차원과 시간 지연 파라미터를 가지고도 충분한 정확도를 가지는 예측결과를 얻을 수 없었다.

4.3 병렬구조 퍼지시스템에 의한 예측

다음은 이 논문에서 제안한 병렬구조 퍼지시스템을 카오스 시계열 데이터의 예측에 적용하였다. 여기서 사용한 병렬구조 퍼지시스템은 병렬로 연결된 5개의 구성요소 퍼지시스템을 포함한다($N=5$). 각 구성요소 퍼지시스템은 각 시간지연에 대해서 최적의 임베딩차원이 할당된다. 모든 구성요소 퍼지시스템에 의해서 생성된 5개의 예측결과들은 각 단계마다 최대값과 최소값을 가지는 예측결과를 제외하고 평균되어 최종 예측값을 얻는다. 최대값과 최소값을 가지는 예측결과들은 예측오차를 증가시키는 경향이 있으므로 최종 예측결과를 계산하는데 있어서 고려하지 않았다.

그림 6은 500개의 학습 데이터를 이용하여 구성된 병렬구조 퍼지시스템의 예측결과이다. 500개의 학습 데이터로는 단기 및 장기예측을 위한 병렬구조 퍼지시스템을 구성하는데 적절하지 못하였다. 그림 7은 10,000개의 학습데이터를 이용하여 학습된 병렬구조

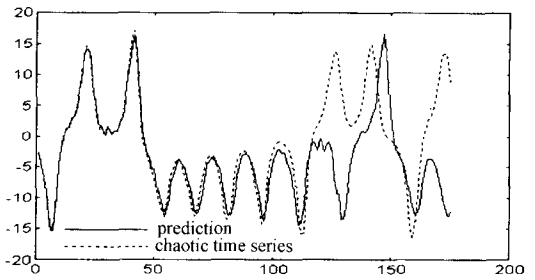


그림 7. 10,000개의 학습데이터를 이용하여 학습한 경우 병렬구조 퍼지시스템의 예측결과
 Fig. 7. Prediction result of the PSFS with 10,000 training samples

퍼지시스템의 시계열 데이터 예측결과를 나타낸다. 학습에 사용하지 않은 200개의 데이터를 예측한 결과 약 120샘플 정도까지 비교적 정확한 예측결과를 얻을 수 있었다. 이것은 충분한 수의 데이터를 이용하여 학습된 병렬구조 퍼지시스템은 카오스 시계열 데이터를 성공적으로 예측할 수 있다는 사실을 의미한다.

5. 결 론

이 논문에서는 카오스 시계열 데이터의 예측을 위하여 병렬구조 퍼지시스템을 제안하였다. 병렬구조 퍼지시스템은 병렬로 연결된 복수개의 구성요소 퍼지시스템들로 구성된다. 각 구성요소 퍼지시스템은 수치 입력력 데이터로부터 정보를 추출해내기 용이한 다입력 단출력 Sugeno 퍼지규칙들을 포함하고 있다. 시계열 데이터 예측을 위한 구성요소 퍼지시스템의 퍼지규칙을 생성하기 위하여 입력력 데이터를 클러스터링하여 모델링하는 방법을 사용하였다.

각 구성요소 퍼지시스템은 각각 다른 값의 임베딩차원과 시간지연들을 가지며 동일한 시점의 미래 데이터를 예측한다. 임베딩차원은 각 구성요소 퍼지시스템의 입력의 수를 결정하고, 시간지연은 각 구성요소 퍼지시스템의 입력 샘플들을 결정한다. 각 시간지연에 대해서 최적 임베딩차원은 한 단계 이후 예측결과와 실제값과의 평균제곱오차와 최대절대오차를 최소화하는 값으로 구한다. 이 논문에서 구성요소 퍼지시스템의 수는 5로 선택하였다. 구성요소 퍼지시스템의 수가 증가함에 따라 보다 정확한 예측결과들을 얻을 수 있으나, 이에 따라 계산량이 증가하므로 가능한 한 구성요소 퍼지시스템의 수를 적게 유지하여야 한다. 병렬구조 퍼지시스템은 각 퍼지시스템들이 출력한 예측결과들 중에서 최대값과 최소값을 제외하고 평균을 취해서 최종 예측결과를 생성한다. 컴퓨터 시뮬레이션을

통하여 학습데이터와 검증데이터를 이용하여 모델링된 병렬구조 퍼지시스템은 x -성분 Lorenz 시계열 데이터와 laser에서 발생한 카오스 시계열 데이터를 성공적으로 예측하였다. 또한 충분한 수의 데이터로 학습된 병렬구조 퍼지시스템은 Lorenz 시계열 데이터에 대해서 만족할 만한 정확도를 가진 예측결과를 생성하였다.

참고문헌

[1] A. S. Weigend and N. A. Gershenfeld, eds. *Time Series Prediction: Forecasting the Future and Understanding the Past*, Addison-Wesley Pub., pp. 175-193, 1994.

[2] M. Casdagal, "Nonlinear Prediction of Chaotic Time Series," *Physica D*, pp. 335-356, 1989.

[3] D. Lowe and A. R. Webb, "Time Series Prediction by Adaptive Networks: A Dynamical Systems Perspective," *IEE Proceedings F*, Vol. 1, pp. 17-24, 1991.

[4] D. S. Broomhead and D. Lowe, "Multi-variable Functional Interpolation and Adaptive Networks," *Complex Systems*, pp. 262-303, 1988.

[5] M. Sugeno, *Industrial Applications of Fuzzy Control*, Elsevier Science Pub., 1985.

[6] E. H. Mamdani, and S. Assilian, "An Experiment in Linguistic Synthesis with a Fuzzy Logic Controller," *Int. J. of Man Machine Studies*, Vol. 7, No. 1, pp. 1-13, 1975.

[7] R. R. Yager and D. P. Filev, *Essentials of Fuzzy Modeling and Control*, John Wiley & Sons, pp. 246-264, 1994.

[8] J. Bezdek, R. Hathaway, M. Sabin, and W. Tucker, "Convergence Theory for Fuzzy c-means: Counter Examples and Repairs," *The Analysis of Fuzzy Information*, J. Bezdek (ed.), CRC Press, Vol. 3, Ch. 8, 1987.

[9] J.-S. R. Jang, and C.-T. Sun, "Neuro-Fuzzy Modeling

and Control," *Proceedings of the IEEE*, March 1995.

[10] S. Chiu, "Fuzzy Model Identification Based on Cluster Estimation," *Journal of Intelligent & Fuzzy Systems*, Vol. 2, No. 3, Sept. 1994.

[11] S. Shao, "Fuzzy Self-organizing Control and its Applications for Dynamical Systems," *Fuzzy Sets and Systems*, Vol. 26, pp. 151-164, 1988.

[12] U. Huebner, N. B. Abraham, and C. O. Weiss, "Dimensions and entropies of chaotic intensity pulsations in a single-mode far-infrared NH₃ laser," *Physical Review A*, Vol. 40, pp. 6354, 1989.

[13] A. M. Fraser and H. L. Swinney, "Independent Coordinates for Strange Attractors from Mutual Information," *Physical Review A*, Vol. 33, pp. 1134-1140, 1986.

[14] M. B. Kennel et al., "Determining Embedding Dimension for Phase Space Reconstruction using Geometrical Construction," *Physical Review A*, Vol. 45, pp. 3403-3411, 1992.

김민수 (Min-Su Kim)

1995년 : 송실대학교 전기공학과(공학사)
 1997년 : 송실대학교 대학원 전기공학과(공학석사)
 1997년~현재 : 송실대학교 대학원 전기공학과 박사과정



공성곤 (Seong-Gon Kong) (정회원)

1982년 : 서울대학교 전기공학과(공학사)
 1987년 : 서울대학교 대학원 전기공학(공학석사)
 1991년 : 미국 University of Southern California (공학박사)
 1992년 3월~현재 : 송실대학교 전기공학과 부교수

주관심분야 : 지능시스템이론, 영상 및 신호처리, 로봇틱스, 진화연산, 인공지능명