
퍼지모델을 이용한 유사성 기반의 동적 클러스터링

이한수* · 김수대* · 김용현* · 김성신*

*부산대학교

Similarity-based Dynamic Clustering Using Radar Reflectivity Data

Hansoo Lee* · Sudaek Kim* · Yonghyun Kim* · Sungshin Kim*

*Pusan National University

E-mail : {hansoo, kimsudaek, yhkim, sskim}@pusan.ac.kr

요 약

어떠한 객체의 움직임이나 상태변화를 추적하기 위해서 사용하는 방법으로는 칼만필터, 파티클 필터, 동적 클러스터링 등이 있다. 이 중 동적클러스터링 기법은 여러 프레임에 걸쳐 클러스터를 추적하고 변화 경향을 분석하는데 유용한 방법이다. 본 논문에서는 유사성 기반의 동적 클러스터링 방법을 제안하고 시뮬레이션 하여 검증하였다. 제안한 동적 클러스터링 방법은 연속된 각 프레임에 대해 유사한 특성을 가지는 클러스터를 인접한 프레임에 걸쳐 동일한 클러스터로 판단하는 방법이다. 각 정지 프레임에서의 클러스터의 특성을 이용하여 프레임의 변화를 분석하고 유사성이 높은 클러스터들을 동일 클러스터로 지정하였다. 유사성 판단 방법은 Mamdani 방식의 퍼지 모델을 제안하였다. 제안한 알고리즘은 시간에 대해 연속성을 가진 레이더 반사도 데이터에 적용하였고 시간의 흐름에 따른 클러스터의 변화를 관측할 수 있었다.

ABSTRACT

There are number of methods that track the movement of an object or the change of state, such as Kalman filter, particle filter, dynamic clustering, and so on. Amongst these method, dynamic clustering method is an useful way to track cluster across multiple data frames and analyze their trend. In this paper we suggest the similarity-based dynamic clustering method, and verifies it's performance by simulation. Proposed dynamic clustering method is how to determine the same clusters for each continuative frame. The same clusters have similar characteristics across adjacent frames. The change pattern of cluster's characteristics in each time frame is throughly studied. Clusters in each time frames are matched against each others to see their similarity. Mamdani fuzzy model is used to determine similarity based matching algorithm. The proposed algorithm is applied to radar reflectivity data over time domain. We were able to observe time dependent characteristic of the clusters.

키워드

Dynamic clustering, state estimation, similarity characteristic,
continuative frame, Fuzzy model

1. 서 론

동적 클러스터링 기법은 여러 프레임에 걸쳐 존재하는 데이터를 통해 클러스터링을 수행하고, 이를 통해 도출된 새로운 클러스터를 통해서 클러스터간의 유사성을 판단하고 변화경향을 추적하는 경우에 사용하는 기법이다. 이 기법을 사용

하여 연속적인 데이터에서 공통적인 요소를 추출하여 분석할 수 있기 때문에 네트워크나 데이터베이스 분석, 레이더 데이터 분석 등에 동적 클러스터링이 적용될 수 있다.

이 때, 클러스터 간의 유사성을 어떤 기준으로 판단할 것인가 하는 문제가 발생할 수 있다. 특히 클러스터의 형태가 시간이 지나더라도 급격하게

변하지 않는 경우에는 이 유사성의 기준이 클러스터링의 성능 척도가 된다. 이러한 경우에는 각 프레임에서 클러스터링 된 데이터의 중심 값의 시간에 따른 변화 혹은 시간에 따른 클러스터 크기의 변화 등과 같은 조건들을 사용할 수 있다. 본 논문에서는 이와 같은 유사도 판단을 퍼지 모델을 적용하여 수행하는 동적 클러스터링 기법을 제안한다.

II. 본 론

2.1 동적 클러스터링

본 논문에서 적용한 동적 클러스터링 기법은 입력으로 현 시간 T를 기점으로 하여 클러스터링을 시작하는 과거시간 T-k에 걸쳐서 존재하는 데이터를 이용하여 클러스터링을 하는 과정으로 이루어져 있다.

먼저 현 시간 T의 클러스터 C_i 가 iso-cluster 그룹에 속해 있는지 판단한다. 여기서 iso-cluster는 시간의 변화에 따라서 같은 범주로 묶을 수 있는 유사한 클러스터의 집합을 뜻한다. 만약 이 그룹에 속해 있지 않다면 클러스터 C_i 와 C_i 의 중심 값과 인접한 위치에 있는 T-1 시간대의 클러스터들 사이에서 유사도를 구한 후 이 값이 한계 값 이상이면 iso-cluster 그룹에 추가한다. 현 시간 T의 마지막 클러스터 C_n 까지 동일한 과정을 반복한다.

2.2 퍼지 모델

동적 클러스터링을 수행하면서 iso-cluster 그룹에 현 시간 T의 클러스터가 속하지 않을 때 퍼지 모델을 사용하게 된다.

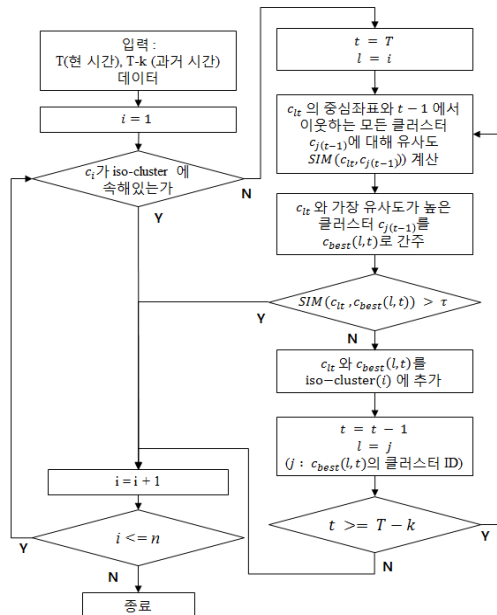


그림 1. 동적 클러스터링 알고리즘

표 1. 유사도 측정을 위한 퍼지 추론 규칙

$ \Delta x $	$ \Delta y $	$ \Delta z $	$ \Delta rsz $	Output
small	small	small	small	high
			medium	middle
			large	low
		large	small	middle
			medium	low
			large	low
	large	small	small	middle
			medium	low
			large	low
		large	small	low
			medium	low
			large	low
large	small	small	small	middle
			medium	low
			large	low
		large	small	low
			medium	low
			large	low
	large	small	small	low
			medium	low
			large	low
		large	small	low
			medium	low
			large	low

클러스터들 간의 유사도를 구하는 방식은 Mamdani 방식의 퍼지 모델을 설계하여 구현하였다. 구현한 퍼지 모델의 입력은 dx, dy, dz, rsz 네 가지 변수로 구성되어 있으며 출력이 하나인 모델이다. 추론 규칙의 구현 방식은 다음 수식 1에 나타나는 것과 같다.

$$\text{"If } |\Delta x| \text{ is small and } |\Delta y| \text{ is small and } |\Delta z| \text{ is small and } |\Delta rsz| \text{ is small then Output is high."} \quad (1)$$

이 퍼지 모델이 가지는 입력의 경우의 수를 고려하여 24가지의 경우에 대하여 퍼지 모델을 설정하였다. 각각의 경우의 수는 표 1에 정리하였으며, 이것을 통해서 도출되는 멤버쉽 함수와 출력 함수는 그림 1과 그림 2에 나타내었다.

dx, dy, dz는 각각 현 시각 T에서의 클러스터 C_i 의 중심값의 공간좌표(x_i,y_i,z_i)와 이전 시각 T-1의 인접한 클러스터 C_j 의 중심값의 공간좌표(x_j,y_j,z_j)의 유클리드거리 값이며, rsz는 C_i 와 T-1의 인접한 클러스터 C_j 의 사이즈 차이값이다. 출력 값은 무게중심법(COA : Center of Area) 비퍼지화기를 제안하였으며 그것에 해당하는 식은 아래와 같이 나타낼 수 있다.

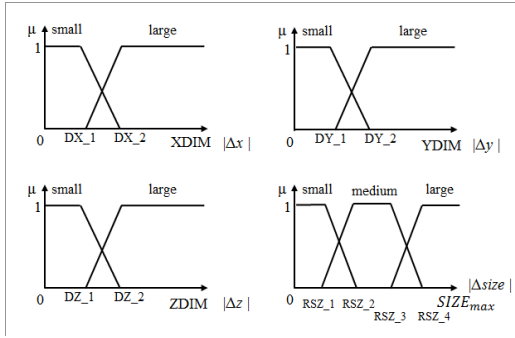


그림 2. 퍼지 모델의 입력부

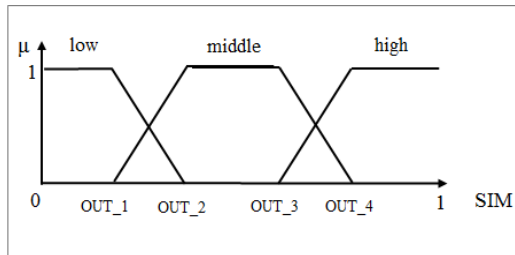


그림 3. 퍼지 모델의 출력부

$$z\ COA = \sum_z z \mu_A(z) / \sum_z \mu_A(z) \quad (2)$$

위의 수식을 이용하여 퍼지 모델에서 24개의 경우에 대하여 각각 결과값이 나오게 된다. x,y,z의 변화폭이 작고 크기의 변화도 작으면 유사성이 가장 큰 iso-cluster로 판단하고, x,y,z의 변화폭이 크고 크기의 변화도 크면 유사성이 낮은 iso-cluster로 판단한다.

III. 실험 및 결과

본 논문에서 제안한 유사성 기반의 동적 클러스터링 방법을 레이더 반사도 데이터에 적용해 보았다. 비교를 위해서 대조군으로 퍼지 모델을 적용하지 않고 유클리드 공간에서 거리 변화를 비교하는 방법을 채택하였다. 실험을 위해 사용한 레이더 반사도 데이터의 분포 형태는 그림 4와 같이 나타난다.

동적 클러스터링을 적용하기 위해서는 각각의 프레임에 대해서 정적 클러스터링이 적용되어 있어야 한다. 그리고 그 클러스터들을 이용하여 동적 클러스터링을 수행하게 되는데, 그림 5에 나와 있는 파란색으로 표현된 부분이 유사성 기반의 동적 클러스터링을 적용한 것이고, 빨간색으로 표현된 부분은 단순 거리 변화 클러스터링 방법을 적용한 것이다. 그리고 검게 표시된 부분은 두 가지 방식의 클러스터링이 적용될 때 공통적으로 시작하는 시작점이 된다.

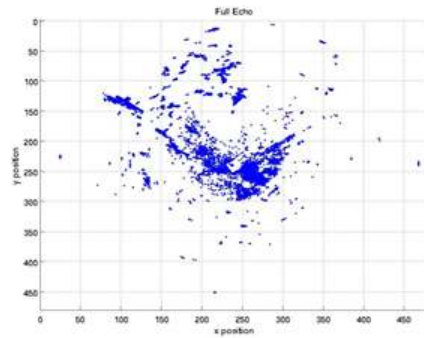


그림 4. 레이더 반사도 데이터

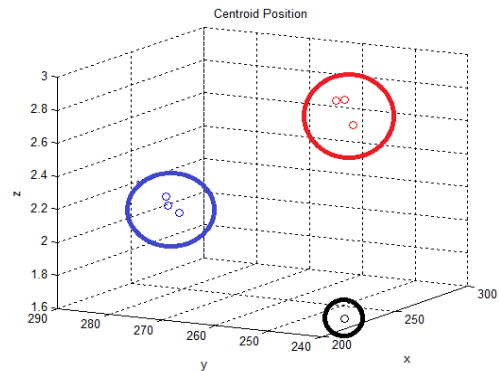


그림 5. iso-cluster의 centroid 분포

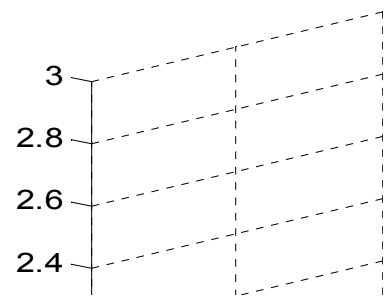


그림 6. 동적 클러스터링 적용 결과

표 2. 서로 다른 두 가지 클러스터링 결과

	$ \Delta x $	$ \Delta y $	$ \Delta z $	$ \Delta rsz $
유사성 기반	38.10	44.1	0.42	33
단순 거리 변화	46.17	14.34	1.13	1217

시작점을 기준으로 두 가지 서로 다른 동적 클러스터링 방식을 적용하면 그림 6과 같이 표현할 수 있다. 시작점에서 방법에 따라 크게 두 갈래로 나뉘어서 클러스터링이 진행되는 것을 볼 수 있다. 파란색은 유사도 기반 퍼지 모델 동적 클러스터링 방식이고, 빨간색은 단순 거리 변화 클러스터링 방식을 나타낸다. 이 두 가지 모델의 레이더 데이터 적용 결과는 그림 7에 나타내었다.

이에 대해 수행한 결과를 정리하면 표 2와 같이 나타낼 수 있다. 시작점에서의 시간을 T라고 할 때, 시작점의 중심값의 좌표 x,y,z 와 rsz 값과 T-1에서 적용한 방법에 따라 각각 iso-cluster라고 판단한 클러스터의 중심값의 좌표 x,y,z 와 rsz 값의 차이를 계산하여 정리한 것이다.

두 가지 방법에 대하여 유클리드 거리를 계산하면 3차원 거리 계산식에 의해서 유사성 기반 동적 클러스터링 방법은 58.28이, 단순 거리 비교 방법에서는 48.36이 나오게 된다. 하지만 크기 차이가 유사성 기반 방식의 경우에는 33으로 거의 비슷하게 표현되었지만 단순 거리 비교 방법에서는 1217로 굉장히 많은 차이가 나고 있다. 따라서 거리만 비교하였을 때와 유사성 기반 퍼지 모델을 사용하여 판단하였을 때의 결과가 다르게 나오게 된다.

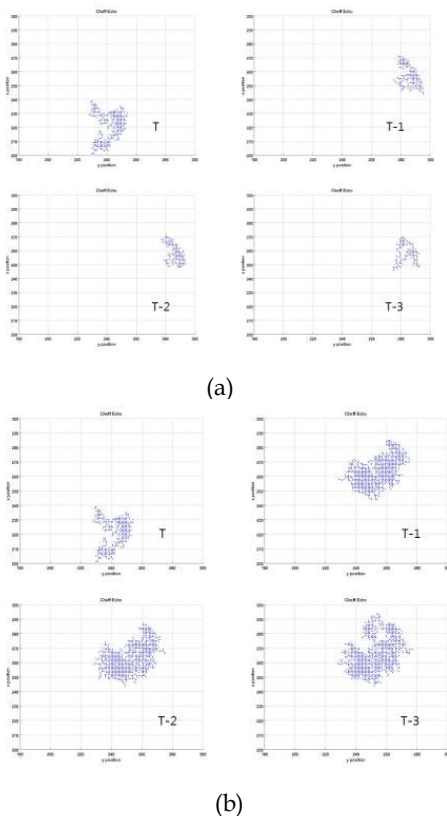


그림 7. 동적 클러스터링 적용 결과
 (a) 유사도 기반 퍼지 모델
 (b) 단순 거리 계산 모델

이 때, 거리는 조금 더 차이가 나지만 크기 변화가 작은 퍼지 모델을 이용한 동적 클러스터링이 올바른 결과가 된다.

IV. 결 론

본 논문에서는 퍼지 모델을 이용한 유사도 기반 동적 클러스터링 방법에 관하여 논의하였다. 대조 방식과 비교하였을 때 판단의 기준이 된 측면을 고려하여 결과가 효율적으로 도출되는 것을 확인할 수 있었다. 퍼지 규칙을 적용하게 되면 다양한 기준을 고려하여 동적 클러스터링을 수행할 수 있다. 데이터의 종류에 따라서 퍼지 규칙을 추가하면 해당 방식에 최적화된 동적 클러스터링을 수행할 수 있을 것이다.

서로 다른 위치에 클러스터가 존재하는데 시작점과의 거리 변화와 크기 변화의 폭이 같게 나올 때 어떤 것을 택하여 클러스터링을 계속해 나갈 것인가에 대한 부분은 추후 연구로 계속 진행되어야 할 부분이다.

감사의 글

본 연구는 지식경제부 및 정보통신산업진흥원의 특수환경 Navigation/Localization 기술연구센터 지원사업의 연구결과로 수행되었음 (NIPA-2011-C7000-1001-0004)

참고문헌

- [1] Yong Hyun Kim, Sudaek Kim, Hansoo Lee, Sungshin Kim, " Static and Dynamic Frame Clustering Algorithms for 3D Tracking of Deformable Structures Using Doppler Weather Radar," International Symposium on Advanced Intelligent Systems, 2011.
- [2] XIAOZHE WANG, ROB HYNDMAN, "Characteristic-Based Clustering for Time Series Data," DATA MINING AND KNOWLEDGE DISCOVERY, Vol.13, No.3, pp335-364, 2006.
- [3] A.M.Alonso, J.R.Berrendero, A.Hernandez, Ajustel, "Time Series clustering based on forecast densities," Computational Statistics & Data Analysis, Vol.51, No.2, pp762-776, 2006.
- [5] 전진호, 이계성, "시계열데이터의 모델기반 클러스터 결정," 한국콘텐츠학회논문지, Vol.7, No.6, pp22-30, 2007.
- [6] 김영인, 박영배, "시계열 패턴의 클러스터링을 위한 효율적인 다차원 동적 인덱스 구조," 한국정보과학회 학술발표논문집, Vol.21, No.1, pp85-88, 1994.