

논문 2021-1-11 <http://dx.doi.org/10.29056/jsav.2021.06.11>

# 행동 시계열 데이터와 $k$ -평균 군집화를 통한 젖소의 일일 행동패턴 검출

이성훈\*, 박기철\*, 박재화\*\*†

## Daily Behavior Pattern Extraction using Time-Series Behavioral Data of Dairy Cows and $k$ -Means Clustering

Seonghun Lee\*, Gicheol Park\*, Jaehwa Park\*\*†

### 요 약

지난 동안 낙농업계에서는 다양한 센서 기술과 ICT 응용이 도입되어왔으며 축적된 낙농 데이터를 토대로 과학적인 낙농생산관리가 가능해졌다. 그러나 이러한 시도들은 젖소의 출산이나 우유 생산량과 같은 낙농 생산성에 직접적으로 관여하는 요인들에 대해서만 집중적으로 이루어졌으며 이러한 결과에 근본적으로 관여하는 생리학적 혹은 동물심리학적 요인에 대해서는 연구가 더딘 실정이다. 이 논문에서는 이러한 연구의 일환으로서 젖소의 시간별 행동 데이터로부터 일일 행동패턴을 검출하는 기초적인 방안을 제시하였다.  $k$ -평균 군집화를 통해 한 젖소의 1594일간 행동을 네 개의 군집으로 구분하였으며 각 군집에 속한 데이터와 군집의 대표값을 시각화하여 군집 형성의 합리성을 확인하였다. 또한 개체의 일별 군집 변화를 토대로 군집 개수의 적정성을 판단하였다. 이 연구 결과가 향후 젖소의 이상상태나 질병징후의 포착 연구에 기여하기를 기대한다.

### Abstract

There are continuous and tremendous attempts to apply various sensor systems and ICTs into the dairy science for data accumulation and improvement of dairy productivity. However, these only concerns the fields which directly affect to the dairy productivity such as the number of individuals and the milk production amount, while researches on the physiology aspects of dairy cows are not enough which are fundamentally involved in the dairy productivity. This paper proposes the basic approach for extraction of daily behavior pattern from hourly behavioral data of dairy cows to identify the health status and stress. Total four clusters were grouped by  $k$ -means clustering and the reasonability was proved by visualization of the data in each groups and the representatives of each groups. We hope that provided results should lead to the further researches on catching abnormalities and disease signs of dairy cows.

**한글키워드** : 다변량 시계열, 행동 분석, 군집화, 시각화, 스마트팜

**keywords** : multivariate time-series, behavior analysis, clustering, visualization, smartfarm

\* 중앙대학교 컴퓨터공학과

\*\* 중앙대학교 소프트웨어학과

† 교신저자: 박재화(email: jaehwa@cau.ac.kr)

접수일자: 2021.06.04. 심사완료: 2021.06.19.

게재확정: 2021.06.20.

## 1. 서론

낙농업계에서의 가장 큰 목표는 낙농 생산성

을 증가시키는 것이며 특히 낙농 생산성 중 대부분을 차지하는 젓소의 우유 생산량을 증대시키기 위해서는 젓소의 건강 상태나 젓소가 먹는 음식물에 대한 영양학적이고 질병학적인 분석이 필요하다. 이러한 부분에 대해 과거에는 목장주나 수의사의 노동집약적인 분석 방법을 통해 낙농 생산성을 증대시켜왔으나 모든 젓소를 상시 모니터링하고 관리할 수 없다는 점에서 지속적인 한계를 보여왔다.

한편 지난 10년간 다양한 센서 기술과 ICT 응용이 낙농업계에 폭발적으로 도입되었고 그 결과 센서 데이터에 기반한 과학적인 사양관리가 보편화되었다[1-5]. 이로 인해 과거에는 시도되지 못했던 젓소의 일별 섭취량이나 행동량, 위치정보 등을 결합할 수 있게 되었고, 이에 다변량 분석이나 멀티모달 분석을 수행하는 것이 현재 낙농 분야에서의 주요 연구 현황이다[6-7]. 하지만 이러한 연구들은 우유 생산량이나 되새김질 시간, 운동 시간과 같이 낙농 생산량에 직접적으로 영향을 미치는 요소들에 대해서만 목표로 한다는 한계를 가지며 젓소 자체의 쾌적도나 행동 양태를 분석하려는 연구는 충분히 이루어지고 있지 않다.

예를 들어 Nils Zehner의 연구[8]에서는 가속도계를 이용하여 되새김질 시간을 추정하는 모델을 제안하였으며 Rombach의 연구[9]에서는 턱의 움직임을 통해 되새김질 시간을 추정한다. 이러한 연구들은 다양한 센서 데이터로부터 젓소의 특정 행동을 포착하려는 제한적인 분석이지 젓소의 전체적인 상태나 행동패턴을 목표로 하는 것이 아니라는 한계가 있다. 몇몇 연구[10-11]는 이에 더 나아가 젓소의 전반적인 행동을 추정하고자 하였으나 이들 연구는 예측 시간이 불연속적이며 자료 또한 제한적이기 때문에 젓소의 하루 동안의 행동 양식에 대해 파악하기에는 충분하지 않다.

본 논문에서는 기존에 다루지지 않았던 젓소의 일일 행동 패턴을 추출하고 분석하는 방법을 제시하며 이를 통해 젓소의 이상 행동이나 질병 징후 분석에 대한 기초적인 분석 방법을 제시한다. 이를 위해 젓소의 시간별 행동 데이터를 바탕으로 젓소의 일별 행태를 군집화하고 이로서 일반적인 생활양태를 파악한다. 데이터 수집은 충청남도의 한 농가에 설치된 로봇착유기로부터 약 1504일간 이루어졌으며 이를  $k$ -평균 군집화 알고리즘으로 군집화하여 일별 행동패턴을 찾아내었다. 주성분분석을 통해 고차원의 다변량 시계열 데이터를 차원축소시켰으며 그 결과를 다양한 시각화 알고리즘을 통해 검증하였다. 그 결과 데이터로부터 총 네 개의 행동패턴을 파악할 수 있었으며 각 행동패턴 간 분명한 차이를 파악하여 군집이 합리적인 수준에서 나뉘었음을 확인할 수 있었다.

이 논문에서 제시한 방법을 통해 특정 젓소의 일반적인 행동양태를 파악할 수 있으며 이는 곧 젓소의 이상행동이나 질병징후를 포착할 수 있는 중요한 단서가 된다. 또한 제시한 방법을 다수의 젓소에 적용하면 우군에 대한 일반적인 행동양태를 추출할 수 있으며 이를 토대로 행동 데이터를 기반으로 한 우군 분석이 가능하다. 이 연구 결과가 보다 많은 영역에서 다양한 후속 실험으로 이어지기를 기대한다.

## 2. 데이터 및 실험 방법

### 2.1 데이터

젓소의 행동 시계열 데이터는 충청남도에 위치한 농가의 젓소들로부터 수집되었으며 귀에 부착된 가속도계 센서(CowManager SensOor, Agis Automatisering BV, Harmelen, the Netherlands)를 통해 수집되었다. 이 센서는 내장

된 배터리를 통해 5년 동안 동작하며 수집된 데이터는 농가에 설치된 라우터를 거쳐 주기적으로 CowManager 시스템의 데이터베이스로 저장된다. 이 데이터는 CowManager 시스템에서 제공하는 웹 인터페이스를 통해 확인할 수 있다.

이 연구에서는 실험을 위해 사육중인 젓소들 중 세 마리만을 선택하였고 2017년 1월 2일부터 2021년 5월 14일까지 총 1594일 동안 데이터를 수집하였다. 선택된 세 젓소 중 한 개체는 2011년에, 다른 두 개체는 2012년에 출생하였다.

데이터 수집 시스템은 센서의 가속도 정보를 가공하여 추정된 5개의 시간별 행동 비율을 제공한다. 행동 비율은 해당 시간대에 얼마나 해당 행동을 했는지를 나타내며 각각의 행동은 행동 없음(No-Active), 약한 활동성(Active), 강한 활동성(High-Active), 섭식(Eating), 되새김질(Luminating)로 정의된다. 이러한 행동의 근거가 되는 센서값이나 알고리즘에 대한 것은 시스템에서 공개되어있지 않으며 이 연구에서는 이 시스템에서 제공하는 가공된 데이터만을 사용하여 젓소의 일일양태를 분석한다.

수집된 데이터는 24시간동안 5가지 행동에 대한 시계열 데이터로서 사용하였으며 전처리 또한 하루를 기준으로 수행되었다. 이 연구에서 수집된 데이터는 시스템에서 가공된 데이터이기 때문에 높은 수준의 정확성이 보장되어있다. 그러나 센서 교체 등으로 인한 시스템 외부적인 요인에 의해 약 1% 가량의 데이터가 불량으로 선별되어 폐기되었다.

## 2.2 방법

시계열 데이터로부터 일일 행동패턴을 검출하기 위해 시계열 데이터에 대한 군집화를 수행하였다. 이 때 일단위의 데이터는 한 시간 동안 5개의 행동 간 비율이 24시간동안 연속된 것이므로 120개의 특징을 가지는 벡터로 나타내어진다.

그러나 이같이 높은 차원의 데이터는 데이터의 특성을 이해하고 그에 맞는 데이터 간 유사도(Similarity)를 사용하지 않는 한 의미 있는 군집화가 어렵다. 이 논문에서는 이에 대한 예비적 실험으로 주성분분석기법을 이용한 차원축소기법을 사용하였다.

주성분분석은 데이터를 데이터의 공분산행렬로부터 얻어낸 고유벡터의 선형결합으로 나타내는 방법이다. 이 방법을 이용하면 데이터로부터 분산이 가장 크게 나타나는 축들을 찾아낼 수 있으며 이를 응용해 어느 정도의 정보 손실을 감안하며 현재 데이터가 가지는 차원보다 더 적은 차원으로 데이터를 표현할 수 있다. 그림 1은 한 개체의 데이터에 대해 주성분 개수에 따른 설명 가능한 분산량을 나타낸 것으로 39개의 주성분을 사용했을 때 90% 정도의 설명 가능한 분산량을 가지는 것을 확인할 수 있다. 이 연구에서는 설명 가능한 분산량이 전체 데이터 대비 90% 정도인 수준으로 데이터의 차원을 축소시켰으며 이는 세 젓소 개체에 대해 각각 39, 33, 42 차원에 대응한다.

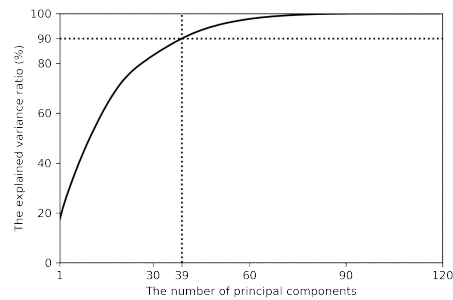


그림 1. 행동 시계열데이터의 주성분분석 시 정보 보존량 그래프

Fig. 1. Conservation of information graph from the principle component analysis of behavioral time-series data

데이터에 대한 탐색적 분석 결과 데이터의 군집화 알고리즘으로는 중심기반(center-based) 군

집화 기법을 사용하는 것이 타당하다고 판단하였다. 비록 데이터의 형태와 분포를 분석하여 젓소의 행동 데이터를 보다 잘 군집화할 수 있는 알고리즘을 찾아낼 수 있으나, 이 논문에서는 일일 행동패턴을 검출할 수 있는 방법론을 제시하는 것을 목표로 하였다.

$K$ -평균( $k$ -means) 알고리즘과  $k$ -중심( $k$ -medoids) 알고리즘은 중심기반 군집화 기법 중 널리 사용되는 방법으로 데이터를 군집 중심들로부터 가장 가까운 군집에 할당하는 방법이다.  $k$ -평균 알고리즘이 군집의 중심으로 평균(Mean)을 사용하는 것과 다르게  $k$ -중심 알고리즘은 군집의 대푯값으로 중앙값(Median)을 사용한다. 이러한 까닭에  $k$ -중심 알고리즘은  $k$ -평균 알고리즘보다 이상값에 둔감하다는 장점이 있다. 하지만 이 연구에서 사용하는 데이터는 데이터 수집 시스템으로부터 가공되어 나오며 이상 데이터가 관측되더라도 여기엔 생리적 또는 동물심리학적 의미가 있다. 그러므로 이 연구에서는  $k$ -평균 군집화 알고리즘을 사용하여 젓소의 행동 데이터를 군집화하였다.

한편 중심기반 군집화 기법의 낮은 초기 수렴 속도를 개선하기 위해  $k$ -평균 알고리즘을 개량한  $k$ -평균++ 알고리즘을 사용하였다[12].  $k$ -평균++ 알고리즘은 샘플과 군집 대푯값 사이의 거리를 이용하여 대푯값으로 이미 선정된 샘플로부터 가까운 샘플들을 초기 군집의 대푯값으로 선정되기 어렵게 하는 방법을 사용한다.

$k$ -평균 알고리즘에서의 데이터 간 유사도로서 유클리드 거리(Euclidean Distance)를 사용한다. 유클리드 거리는 두 벡터의 차이에 대한 L2 노름(norm)으로 계산되며 이는 두 벡터  $x, y$ 에 대해

$$ED(x, y) = \sqrt{\sum_{i=1}^N (x_i - y_i)^2} \quad \dots\dots (1)$$

과 같이 정의된다. 이때,  $N$ 은 벡터의 차원수이다.

한편  $k$ -평균 알고리즘을 위한 적절한 군집의

개수를 설정하여야 하는데 이를 위해 일반적으로는 군집의 에너지(Energy)에 대한 엘보우 분석(Elbow Method)이나 군집의 내부 평가 기법인 실루엣 점수(Silhouette Score)를 사용한다. 하지만 엘보우 분석은 분석가의 자의적인 판단이 개입하기 때문에 신뢰도가 낮으며 실루엣 점수 또한 고차원의 데이터에 대해서는 효과성이 떨어진다. 이에 일간 행동의 군집 변화 정도를 나타내는 클러스터 변동률이라는 새로운 지표를 고안하였다.

클러스터 변동률은 군집이 변화한 일수를 전체 일수로 나눈 것으로 정의된다. 좋은 클러스터 변동률이란 무엇인지 일반화하여 정의하기는 어려우나 클러스터 변동률이 클수록 군집화 결과의 낮은 안정성을 나타내기 때문에 어느 정도 낮은 수준의 클러스터 변동률을 가지는 것은 군집 개수가 적절히 설정되었음을 의미한다. 한편 젓소의 일시적인 특이 행동은 클러스터 변동률의 해석을 교란한다. 이에 클러스터 변동률을 계산함에 있어 특정 기간 동안 클러스터의 변동이 지속되어야 클러스터가 변동한 것으로 취급하였다. 이 연구에서는 실험을 통해 3일 정도의 윈도우를 사용하는 것이 적절하다고 판단하였다.

클러스터링 결과를 2차원에서 다양한 방법으로 시각화하기 위해 주성분분석과  $t$ -분포 확률적 임베딩 기법( $t$ -Stochastic Nearest Embedding,  $t$ -SNE)[13], 균등 다면체 근사 및 사영 기법(Uniform Manifold Approximation and Projection, UMAP)[14]을 적용하였다. 주성분분석이 데이터 공분산행렬의 고유벡터를 이용한 선형 차원축소 방법인 반면, 다른 두 알고리즘은 기계학습을 이용하여 비선형으로 차원을 축소한다.  $t$ -SNE는 2차원에서의 시각화를 위해 데이터 간 유사도를 확률분포로 모델링하며 확률분포 간 KL 발산(Kullback-Leibler Divergence)을 최적화한다. 하지만  $t$ -SNE는 무작위 시작점(Random

Initializer)과 경사 하강법(Gradient Descent)을 기반으로 최적화하기 때문에 항상 동일한 시각화 결과를 얻지는 못한다는 단점이 있다. 반면 UMAP은 리만 기하학(Riemannian Geometry)과 위상수학(Topology)을 차원축소에 접목시킨 방법으로  $t$ -SNE 대비 연산량이 적다는 장점이 있다. 또한  $t$ -SNE가 2차원으로만 차원을 축소할 수 있는 반면 UMAP은 임의의 저차원으로 차원을 축소할 수 있다. 하지만 UMAP 알고리즘은 이해가 무척 어려우며 신뢰성있는 결과를 확보하기 위해서는 비교적 복잡한 전제들을 고려해야 한다는 문제가 있다. 이 연구에서는 각 차원축소기법의 효과성이나 효율성을 평가하기보다는 여러 방법을 통한 데이터 시각화 결과를 비교하는 차원에서 시각화를 계획하였다.

### 3. 실험 결과

그림 2는 이 연구에서 수집한 세 젓소 데이터 중 1번 개체에 대한 군집 개수별 에너지 그래프이며 다른 개체들 또한 그림 2와 비슷한 형태를 나타내었다. 그림 2의 두 그래프는 각각 주성분 분석을 통해 차원축소된 데이터와 원본 데이터에 대한 군집 에너지를 나타낸다. 두 그래프를 동일한 척도에서 비교하기 위해 각 방법을 통해 얻은 군집 할당정보를 원본 데이터에 적용하여 군집 에너지를 계산하였다. 이처럼 동일한 환경에서는 일반적으로 군집화가 더 잘 이루어졌을 때 더 낮은 군집 에너지를 가지는 것으로 판단한다. 한편, 군집 에너지는 각 점과 그에 대응하는 군집 중심 사이의 거리에 대한 제곱합의 누계로 계산된다.

그림 2에서 보여지듯 주성분분석을 통한 차원 축소는 군집 에너지를 의미있는 수준으로 낮추지 못했음을 확인하였다. 보다 더 저차원으로 차원 축소가 일어난 경우에는 오히려 그림 2에서와 달

리 원본 데이터로 군집화한 경우가 군집 에너지가 더 낮았다. 그러나 데이터의 특징 개수가 1/3 수준으로 감소했기 때문에 군집화 속도 개선에 영향을 미치는 것으로 확인되었다.

분석에 사용할 적절한 군집 개수를 선정하기 위해 그림 2와 3에서 나타난 군집 에너지 그래프와 클러스터 변동률 그래프를 비교하였다. 실험 결과 세 개체 모두 군집이 네 개일 때 클러스터 변동률이 10%에 가장 근접하며 군집 에너지 또한 그래프의 엘보우 구간에 해당함을 확인하였다. 이에 따라 이후의 분석은 군집을 네 개로 설정하여 진행하였다. 한편 클러스터 변동률이 10%라 함은 변동이 주기적인 경우 10일에 한 번 주기로 군집 변화가 나타나는 것을 의미한다.

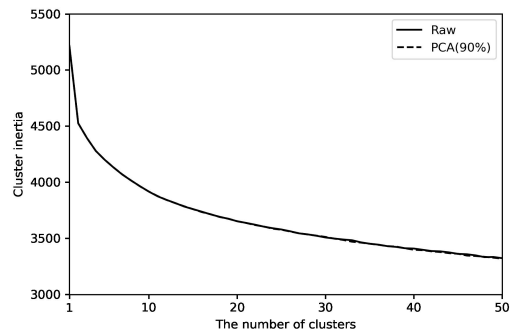


그림 2. 군집의 개수에 따른 군집 에너지 그래프  
Fig. 2. Cluster inertia graph according to the number of clusters

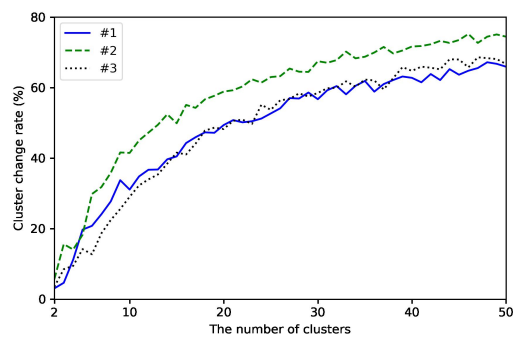


그림 3. 군집 개수에 따른 클러스터 변동률  
Fig. 3. Cluster change ratio according to the number of clusters

1번 개체를  $k$ -평균 군집화를 이용해 네 개의 군집으로 군집화했을 때 각 군집의 중심은 그림 4와 같이 나타내어진다. 군집은 (a)부터 (d)까지 크기순으로 정렬되었고 군집의 크기는 각각 559(35.3%), 528(33.3%), 308(19.4%), 189(11.9%)이다.

그림 4에 따르면 각 군집은 행동에 따른 차이가 나타난다. 군집 (a)와 (b)의 경우 전체 행동의 약 68.6%를 차지하는 만큼 일반적인 행동패턴으로 생각할 수 있으며 그래프에서도 유사한 패턴을 확인할 수 있다. 다만 군집 (a)가 (b)에 비해 되새김질(Luminating)을 많이 하고 밥을 조금 덜 먹으며 활동량도 적은 것을 알 수 있다. 군집 (c)의 경우 매우 활동적이며 되새김질의 비율이 모든 군집 중 가장 낮다. 반면 군집 (d)는 모든 군집 중 가장 정적인 패턴을 나타내며 밥 또한 가장 적게 먹는 것으로 보여진다.

그림 5는 1번 개체에 대한 세 가지의 시각화 결과이다. 그림의 각 점들은 시각화 알고리즘에 의해 2차원으로 차원축소된 데이터를 의미하며 각 점이 가지는 모양과 색은 군집화 알고리즘에 의해 데이터에 할당된 군집 정보를 나타낸다. 비록 각 그림에서 군집 간 경계가 명확하게 식별되는 것은 아니나 군집 자체는 합리적으로 구분되었음을 확인할 수 있다. 특히 그림 6의 3차원 주성분분석 결과에 의하면 그림 5에서의 주성분분석 결과에 비교했을 때 군집 간 경계가 더 명확하게 설정된 것을 확인할 수 있는데 이를 통해 현재 수집한 데이터를 2차원 공간에 나타내기 위해서는 상당한 정도의 정보손실을 감안해야 함을 알 수 있다.

한편 1번 개체를 제외한 나머지 두 개체에서도 1번 개체와 비슷한 수준의 분석 결과가 나오음을 확인하였다. 비록 군집의 크기나 중심의 모습이 다른 개체들과 어느 정도 차이를 보이거나 1번 개체에서 보여지는 것처럼 합리적인 수준에서 일일

행동양태가 추출되는 것을 확인하였다.

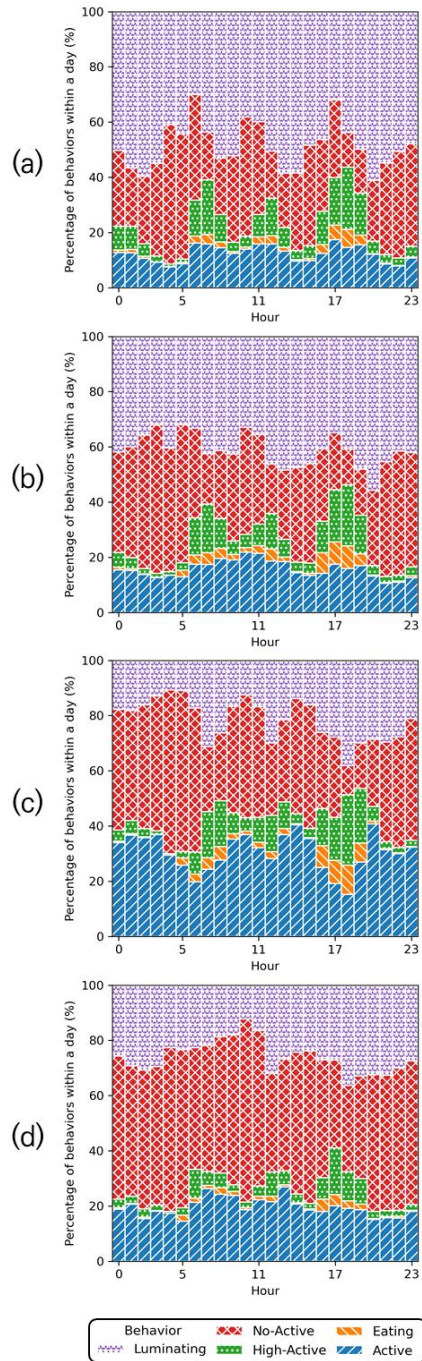


그림 4. 군집별 대표 생활 패턴  
Fig. 4. Representative daily behavior patterns according to each clusters

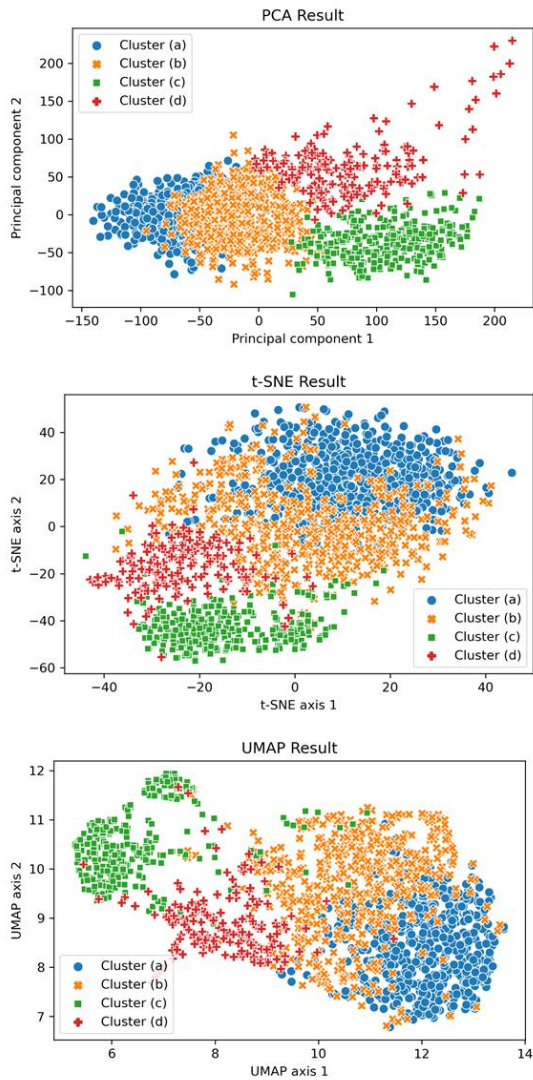


그림 5. 데이터 시각화 도식  
Fig. 5. Data visualization diagram

#### 4. 고찰

이 연구에서는 데이터 수집의 현실적인 어려움을 감안하여 기성 데이터 수집 시스템에서 제공하는 행동 추정 데이터를 기반으로 젓소의 일

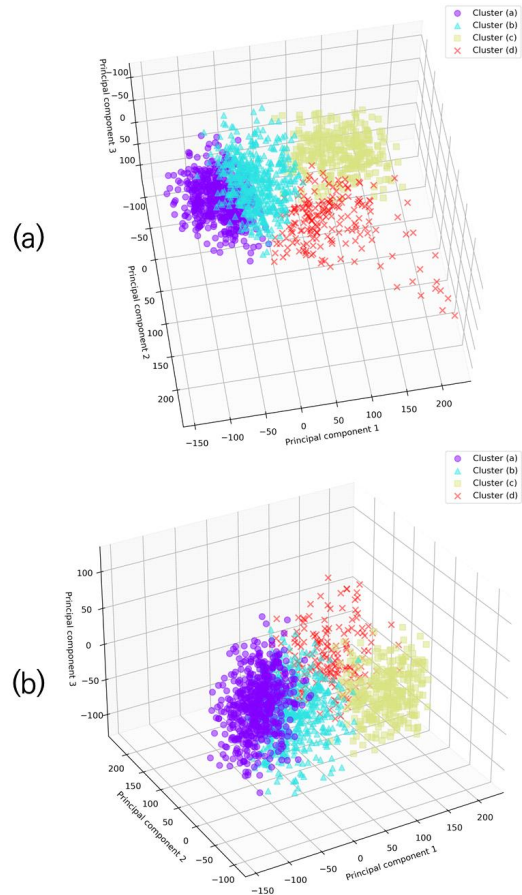


그림 6. 주성분분석을 이용한 삼차원 시각화 도식  
Fig. 6. 3D data visualization diagram based on PCA

일 행동패턴을 분석하였다. 그 결과 일일 행동 패턴이 유의미한 수준으로 추출됨을 확인할 수 있었으며 향후 데이터 특성에 맞는 특징 추출 기법이나 유사도, 군집화 기법을 통해 연구를 고도화할 수 있을 것으로 기대한다.

한편 데이터가 120개의 고차원 특징을 가지는 만큼 효과적인 특징 추출 방안에 대한 연구가 필수적이다. 이에 특징 추출을 위한 휴리스틱 모델을 설계하는 방법을 생각할 수 있으며 반대로 오토인코더(Autoencoder)와 같이 비선형 차원축소와 군집화를 함께 수행하는 방법도 존재한다[15].

또한 이 연구에서는 착유기 시스템이 센서 데이터로부터 추정하여 제공하는 행동 데이터가 진리값(Ground Truth)이라고 전제하였다. 더 정확한 연구를 위해서는 교육받은 감독자나 영상장치를 통해 젓소의 매시간 행동에 대한 진리값을 수집하여야 하는데 여기에는 많은 현실적인 어려움이 따른다. 다행히도 많은 문헌들에서 이 연구에서 사용한 데이터 수집 시스템을 다방면으로 분석하고 검증하였으며 이를 이 논문에서 제시한 방법과 결합하여 후속 연구를 설계할 수 있을 것이다.

한편 이 연구는 예비적 실험의 일환으로서 하나의 데이터 수집 시스템으로부터 얻을 수 있는 데이터를 다각적으로 분석하였지만, 농가에서는 한 젓소에 대해 동시에 여러 종류의 센서를 사용하는 일이 흔하다. 이러한 데이터들을 데이터 통합을 통해 패키징하고 이를 이 연구에 사용된 행동 정보와 결합한다면 멀티모달 방법을 통해 기존에 발견되지 않은 군집별 주요 특징들을 정의할 수 있을 것으로 기대한다.

#### 4. 결론

세 마리 젓소 각각에 대한 1594일간 행동 데이터를 사용하여 각 젓소의 일일 행동양태를 추출하였다. 이를 위해 주성분분석을 이용하여 고차원의 데이터를 차원축소하였으며 이를  $k$ -평균 군집화 알고리즘을 통해 군집화하였다. 군집의 개수는 군집 에너지와 클러스터 변동률을 고려하여 결정되었으며 다양한 시각화 기법을 사용하여 추출된 일일 행동양태들이 합리적인 수준에서 형성됨을 확인하였다.

이 연구에서 제시하는 방법은 젓소의 행동 데이터로 특정 행동만을 추정하는 기존의 방식을 넘어 젓소가 가지는 일반적인 행동양태를 추출한

다. 이 방법은 많은 후속 연구를 통해 보다 다양한 데이터와 결합되어 사용될 것이며 결과적으로 더 구체적이고 정확한 행동양태를 제공할 것이다. 이 연구 결과가 젓소의 생리적 또는 심리학적 분석과 결합하여 보다 고도화된 젓소의 과학적 사양관리에 기여할 수 있기를 기대한다.

본 결과물은 농림축산식품부의 재원으로  
첨단생산기술개발사업의 지원을 받아  
연구되었음(318005-4).

#### 참고 문헌

- [1] Andreas Kamilaris, Andreas Kartakoullis, Francesc X. Prenafeta-Boldu, "A review on the practice of big data analysis in agriculture", in *Computers and Electronics in Agriculture*, vol. 143, 23-37, 2017. DOI: <https://doi.org/10.1016/j.compag.2017.09.037>
- [2] Manlio Bacco, Paolo Barsocchi, Erina Ferro, Alberto Gotta, Massimiliano Ruggeri, "The Digitisation of Agriculture: a Survey of Research Activities on Smart Farming", in *Array*, vol. 3-4, 100009, 2019. DOI: <https://doi.org/10.1016/j.array.2019.100009>
- [3] Sjaak Wolfert, Lan Ge, Cor Verdouw, Marc-Jeroen Bogaardt, "Big Data in Smart Farming - A review", in *Agricultural Systems*, vol. 153, 69-80, 2017. DOI: <https://doi.org/10.1016/j.agry.2017.01.023>
- [4] C. J. Rutten, A. G. J. Velthuis, W. Steeneveld, H. Hogeveen, "Invited review: Sensors to support health management on dairy farms", in *Journal of Dairy Science*, vol. 96, 4, 1928-1952, 2013. DOI: <https://doi.org/10.3168/jds.2012-6107>
- [5] K. N. Dominiak, A. R. Kristense,

- “Prioritizing alarms from sensor-based detection models in livestock production – A review on model performance and alarm reducing methods”, in *Computers and Electronics in Agriculture*, vol. 133, 46–67, 2017. DOI: <https://doi.org/10.1016/j.compag.2016.12.008>
- [6] Kaitlynn M. Abell, Miles E. Theurer, Robert L. Larson, Brad J. White., David K. Hardin, Richard F. Randle, “Predicting bull behavior events in a multiple-sire pasture with video analysis, accelerometers, and classification algorithms”, in *Computers and Electronics in Agriculture*, vol. 136, 221–227, 2017. DOI: <https://doi.org/10.1016/j.compag.2017.01.030>
- [7] Yingqi Peng, Naoshi Kondo, Tateshi Fujiura, Tetsuhito Suzuki, Samuel Ouma, Wulandari, Hidetsugu Yoshioka, Erina Itoyama, “Dam behavior patterns in Japanese black beef cattle prior to calving: Automated detection using LSTM-RNN”, in *Computers and Electronics in Agriculture*, vol. 169, 105178, 2017. DOI: <https://doi.org/10.1016/j.compag.2019.105178>
- [8] Nils Zehner, Joël J. Niederhauser, Matthias Schick, Christina Umstatter, “Development and validation of a predictive model for calving time based on sensor measurements of ingestive behavior in dairy cows”, in *Computers and Electronics in Agriculture*, vol. 161, 62–71, 2019. DOI: <https://doi.org/10.1016/j.compag.2018.08.037>
- [9] M. Rombach, A. Münger, J. Niederhauser, K.-H. Südekum, F. Schori, “Evaluation and validation of an automatic jaw movement recorder (RumiWatch) for ingestive and rumination behaviors of dairy cows during grazing and supplementation.”, in *Journal of Dairy Science*, vol. 101, 3, 2463–2475, 2018. DOI: <https://doi.org/10.3168/jds.2016-12305>
- [10] Jamie Barwick, David W. Lamb, Robin Dobos, Mitchell Welch, Mark Trotter, “Categorising sheep activity using a tri-axial accelerometer”, in *Computers and Electronics in Agriculture*, vol. 145, 289–297, 2018. DOI: <https://doi.org/10.1016/j.compag.2018.01.007>
- [11] A. Rahman, D. V. Smith, B. Little, A. B. Ingham, P. L. Greenwood, G. J. Bishop-Hurley, “Cattle behaviour classification from collar, halter, and ear tag sensors”, in *Information Processing in Agriculture*, vol. 5, 1, 124–133, 2018. DOI: <https://doi.org/10.1016/j.inpa.2017.10.001>
- [12] David Arthur, Sergei Vassilvitskii, “ $k$ -means++: the advantages of careful seeding”, in *Proceedings of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms*, 1027–1035, 2007. URL: <https://dl.acm.org/doi/10.5555/1283383.1283494>
- [13] Laurens van der Maaten, Geoffrey Hinton. “Visualizing Data using  $t$ -SNE”, in *Journal of Machine Learning Research*, vol. 9, 2579–2605, 2008. URL: <http://jmlr.org/papers/v9/vandermaaten08a.html>
- [14] Leland McInnes, John Healy, James Melville, “UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction”, arXiv, 2020. URL: <https://arxiv.org/abs/1802.03426>
- [15] Hinton, G. E. and Salakhutdinov, R. R., “Reducing the Dimensionality of Data with Neural Networks”, in *Science*, vol. 313, 504–507, 2006. DOI: <https://doi.org/10.1126/science.1127647>

— 저 자 소 개 —



이성훈(Seonghun Lee)

2017.2 중앙대학교 컴퓨터공학부 졸업  
2019.2 중앙대학교 컴퓨터공학과 석사  
2019.3-현재 : 중앙대학교 컴퓨터공학과  
박사과정  
<주관심분야> 시계열분석, 데이터분석



박기철(Gicheol Park)

2020.2 중앙대학교 컴퓨터공학부 졸업  
2020.3-현재 : 중앙대학교 컴퓨터공학과  
석사과정  
<주관심분야> 패턴인식, 인공지능



박재화(Jaehwa Park)

1989 한양대학교 전자공학과 졸업  
1991 한양대학교 전자공학과 석사  
2000 Univ. at Buffalo 박사  
2000-2003 Motorola, Inc. 연구원  
2003-현재 : 중앙대학교 교수  
<주관심분야> 패턴인식, 인공지능