

사용자 행동인식을 위한 적응적 경계 보정기반 Particle Swarm Optimization 알고리즘

허성욱^{***}, 권용진^{**}, 강규창^{**}, 배창석^{**}

^{*}과학기술연합대학원 컴퓨터소프트웨어및공학전공

^{**}한국전자통신연구원 BigData소프트웨어연구소
{h75304, scoco, k2kang, csbae}@etri.re.kr

Adaptive Boundary Correction based Particle Swarm Optimization for Activity Recognition

Seonguk Heo^{***}, Yongjin Kwon^{**}, Kyuchang Kang^{**}, Changseok Bae^{**}

^{*}Dept. of Computer Software & Engineering, University of Science & Technology

^{**}BigData Software Research Laboratory, ETRI

요 약

본 논문은 사용자 행동인식을 위해 기존 PSO (Particle Swarm Optimization) 알고리즘의 경계선을 통한 데이터 분류에서 데이터의 수집환경에 의해 발생하는 문제를 벡터의 길이비교를 이용한 보정을 통해 보완한 알고리즘을 제안한다. 기존의 PSO 알고리즘은 데이터 분류를 위해서 데이터의 최소, 최대값을 이용하여 경계를 생성하고, 이를 이용하여 데이터를 분류하였다. 그러나 PSO를 이용하여 행동인식을 할 때 행동이 수집되는 환경에 따라서 경계에 포함되지 못해 행동이 분류되지 못하는 문제가 있다. 이러한 분류의 문제를 보완하기 위해 경계를 벗어난 데이터와 각 행동을 대표하는 데이터의 벡터 길이를 계산하고 최소길이를 비교하여 분류한다. 실험결과, 기존 PSO 방법에 비해 개선된 방법이 평균적으로 약 1%, 걸기 7%, 서기 7%의 개선된 결과를 얻었다.1)

1. 서론

건강상태를 검진받기 위해 병원에 가서 진찰을 받으며, 이때 병원에서는 우리의 생체정보 뿐만 아니라 일상생활의 여러 정보들을 요구한다. 의사는 생체정보와 일상생활 정보를 함께 고려하여 환자의 상태를 검진을 한다. 이러한 병원에서 요구하는 생체정보와 일상에서의 데이터를 이용하여 환자의 건강상태를 파악하는데 도움을 주는 연구가 진행되고 있다 [1].

또한, 일상생활에서의 정보를 디지털화하여 객관적으로 표현하기 위한 디지털문진표에 대한 연구가 진행되고 있다 [2]. 디지털문진표란 기존의 종이형태의 문진표를 디지털형태로 만든 것으로, 이를 이용하면 의사가 환자의 상태를 좀 더 쉽게 진단할 수 있을 것이다. 디지털문진표에는 다양한 항목이 존재하며, 사용자 행동정보는 환자의 활동과 연관된 자료로 사용된다. 문진표에서 활동과 연관된 정보들로는 “하루에 얼마나 걷고 있습니까?”, “하루에 얼마나 자리에 앉아계십니까?”, “규칙적으로 운동을 하시고 계십니까?” 와 같이 활동과 관련된 여러 정보들을 수집된 행동정보를 이용하여 객관적으로 표현이 가능하다. 의사는 수치화된 행동정보를 이용하여 기존의 문진표보다 더 쉽고 정확하게 환자의 상태를 파악할 수 있을 것이다.

본 논문에서는 디지털 문진표에서 사용가능한 사용자

의 행동정보를 장시간동안 지속적으로 수집하는 알고리즘을 제안한다. 사용자 행동을 장시간 수집하고 수집된 데이터를 분류하기 위해서는 수집환경에 민감하지 않는 분류 방법을 사용하여야 한다. PSO (Particle Swarm Optimization) 알고리즘은 데이터 분류하는 문제를 해결하는데 최적화된 알고리즘이다 [3-4]. 또한, PSO 알고리즘은 구현이 간단하고 최적화된 문제에 도입하기 편리하다. 이러한 이유로 우리는 사용자의 행동인식을 위해서 이 알고리즘을 이용한다. PSO를 통해서 데이터를 분류할 때 데이터의 특징과 수집환경에 따라서 데이터 분류의 문제가 발생하였는데, 본 논문에서는 이러한 문제를 해결하기 위해 벡터길이를 이용하여서 문제를 개선하였다.

본 논문은 서론을 시작으로 하여 2장에서는 PSO 알고리즘에 대한 설명과 사용자 행동정보를 분류하는데 발생하는 문제를 해결하기 위한 개선된 방법을 제시한다. 3장에서는 실험을 위한 데이터 분석을, 4장에서는 제안한 알고리즘에 대한 실험을 진행한다. 결론에서는 실험결과를 바탕으로 제안된 방법을 증명한다.

2. 적응적 경계 보정기반 PSO 알고리즘

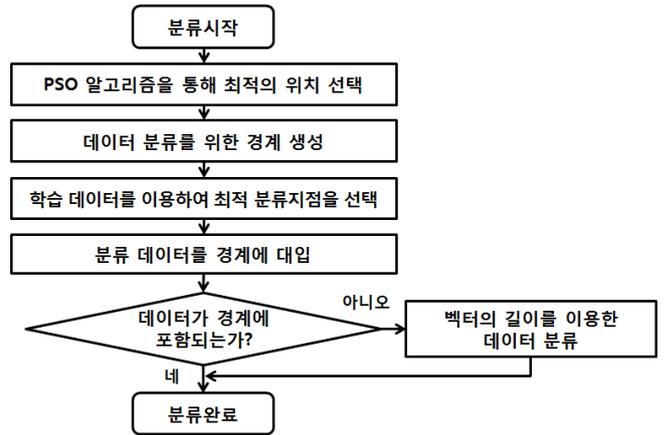
적응적 경계 보정기반 PSO 알고리즘은 기존의 PSO 알고리즘을 이용하여 행동인식을 할 때 발생한 경계선을 벗어나는 데이터의 분류오류를 개선하기 위해 제안된 알고리즘이다. 제안된 PSO 알고리즘은 기존의 PSO 알고리즘 방식에서 경계를 벗어나는 경우 벡터의 길이를 이용하

1) 이 논문은 2012년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. 2012-0001002)

여 보정하는 방법이다.

2.1 Particle Swarm Optimization

PSO (Particle Swarm Optimization)는 각 개체들의 이동경험을 집단 전체에 공유하여 그 정보를 바탕으로 최적의 위치를 찾아가는 알고리즘이다 [3]. 알고리즘에서 파티클 (Particle)은 데이터 객체를 의미하며, 이 파티클들이 모여서 군집화 (Swarm)를 이룬다. PSO에서 각각의 파티클은 공유된 정보를 이용하여 속도를 결정하고 이를 이용하여 다음 위치를 결정한 후, 자신의 정보를 다시 공유한다. 이러한 과정은 최적의 위치를 찾을 때까지 반복된다. PSO는 반복적인 과정을 통해 얻어진 최적의 위치를 이용하여 데이터를 분류한다.



(그림 1) 보완된 PSO 알고리즘 순서

2.2 적응적 경계 보정기반 Particle Swarm Optimization

PSO 알고리즘을 통해 얻어진 최적의 위치는 입력되는 다른 정보들을 분류하는데 사용된다. 본 논문에서는 데이터를 분류하기 위해 각 행동의 학습 데이터에서의 값의 최소, 최대값과 최적화 위치를 이용하여 분류를 위한 경계를 만들어 행동을 인식한다. 분류를 위해서 생성된 여러 경계 중 학습 데이터를 이용하여 분류에 가장 큰 영향을 주는 경계를 찾아 이를 이용하여 데이터를 분류한다. 전체 특징 경계를 이용하는 것이 아닌 가장 큰 영향을 주는 부분을 이용하여 데이터를 분류하는 이유는 데이터를 분류하는데 있어 빠르고 정확한 분류를 위해서 이러한 방법을 사용하고 있다. 이 방법 [5]은 데이터의 분류 정확성을 높이고, 분류의 필요한 시간을 줄여준다.

하지만 사용자 행동인식을 위해서 이 방법을 적용하였을 때 문제가 발생한다. 일반적으로 분류 데이터들은 각 분류별 데이터의 특징이 유사하므로 경계선을 이용하여 쉽게 분류할 수 있다. 반면, 행동인식의 경우 데이터를 수집하는 환경에 따라서 같은 행동이더라도 데이터가 갖는 크기와 특징이 차이가 날 수 있다. 이러한 데이터를 PSO를 통해 분류할 때, 분류경계에 포함되지 못하여 행동인식에 오류가 발생한다.

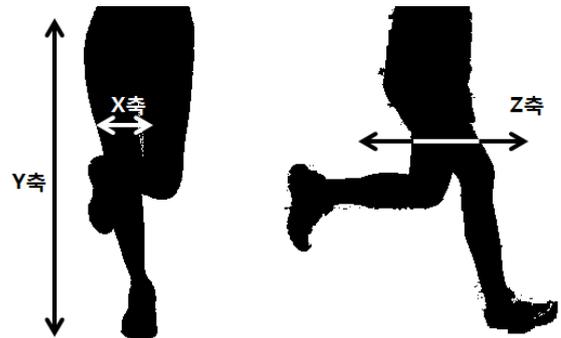
우리는 수집 환경에 의해 경계를 벗어나는 문제를 해결하기 위해서 벡터길이를 이용한 개선방안을 제안한다. 수집환경에 따라서 데이터의 크기가 달라지기는 하지만 데이터의 특징이 남아 있기 때문에 이 점을 이용하여 분류를 보완한다. 경계를 벗어나는 데이터는 분류의 기준이 되는 위치의 벡터 길이를 계산하여 최소화 길이를 이용하여 데이터를 분류한다. (그림 1)은 제안된 적응적 경계 보정기반 PSO 알고리즘의 순서도이다.

3. 사용자 행동 데이터 분석

3.1 실험을 위한 데이터 수집

제안된 적응적 경계 보정 기반 PSO 알고리즘을 위해서 실제 사용자의 행동정보를 이용하여 실험을 진행한다.

실험을 위해서 스마트폰의 가속도 센서를 이용하여 사용자의 행동정보를 수집하였으며, 다양한 환경을 만들기 위해서 일상생활에서의 데이터를 수집하였다. 또한, 수집주기를 다양화하여서 환경에 따른 데이터의 변화를 주었다. 사용자는 (그림 2)처럼 스마트폰을 주머니에 넣고, 이때 스마트폰의 가속도 센서가 지속적으로 사용자의 움직임 정보를 수집, 기록하는 것으로 가정한다. 수집된 데이터의 크기는 가속도 센서의 값이기 때문에 m/s^2 의 단위를 가진다.



(그림 2) 스마트폰의 가속도 센서 방향

3.2 행동정보의 특징

수집된 데이터를 이용하여 분류실험을 하려면, 분류의 대상이 있어야 한다. 일반적으로 가속도 센서를 이용한 실험에서는 다양한 이유로 사용자의 많은 행동을 수집하고 있다 [6]. 일상에서 가장 큰 영향을 주는 행동과 그리고 기존의 연구 [6]를 참고하여 우리는 사용자의 행동 중 앉기, 걷기, 서기를 분류한다. 다음은 수집된 데이터를 이용한 통계적 데이터이다. 통계적 데이터를 이용하여 행동을 정의한다. 표준편차를 기준으로 앉기와 서기의 경우에 걷기에 비해 상대적으로 낮은 표준편차를 가지고 있는데 이는 앉기와 서기가 정적인 행동임을 나타낸다. 이와 다르게 걷기는 표준편차가 다른 행동보다 크기 때문에 동적인 행동으로 정의한다.

<표 1> 각 행동별 통계적 데이터

	앉기	걷기	서기
X축평균(μ_x)	-1.04	-1.45	-1.79
Y축평균(μ_y)	1.89	9.52	0.55
Z축평균(μ_z)	9.5	0.82	-0.21
X축편차(σ_x)	0.28	0.82	0.14
Y축편차(σ_y)	0.21	1.35	0.04
Z축편차(σ_z)	0.08	1.18	0.24

3.3 알고리즘을 위한 특징추출

행동분류를 위해서 본 논문에서는 3 개의 특징을 사용하며, 가속도 센서가 3 개의 축을 가지기 때문에 총 9개의 특징점을 이용하여 데이터를 분류한다. 실험에 사용되는 특징은 시간축면에서 평균, 표준편차를 이용하며 주파수 축면에서 Energy 특징을 이용하였다. Energy 특징은 다음 식을 이용하여 계산된다.

$$Energy = \frac{\sum_{i=1}^{|w|} |x_i|^2}{|w|} \quad (1)$$

여기서 ω 는 윈도우 크기를, x_i 는 i 번째 데이터의 크기를 의미한다. 사용되는 특징들은 가속도 센서를 이용한 연구들 [7]에서 일반적으로 이용되는 방법이다. 우리는 특징추출을 통해서 얻어진 9 개의 특징점을 적응적 경계 보정 기반 PSO 알고리즘을 증명하는데 사용한다.

4. 실험결과

기존의 PSO 알고리즘과 개선된 알고리즘을 비교하기 위해서 일상생활에서의 사용자 행동정보를 2시간동안 기록하여 이용한다. 3장에서 설명한 환경에서 데이터를 수집하였으며, 특히 수집 주기를 다양화하여 개선된 방법이 기존 알고리즘에서 발생한 문제를 해결하는지를 확인한다. 비교에 사용된 수집 주기는 기준이 되는 16Hz, 4Hz, 1초, 4초, 8초, 16초, 32초 이며, 16Hz가 기준 데이터가 된다. 또한 PSO 알고리즘만 아니라, 다른 분류알고리즘의 경우 제시된 문제에서 얼마나 정확한 행동인식을 보이는지 같이 비교한다. 비교에 사용되는 알고리즘은 MLP (MultiLayer Perceptron), C4.5 (의사결정트리), SVM (Support 벡터 Machine)을 이용한다.

실험결과는 일차적으로 각 행동별로 구별하여 수집 주기에 따라 얼마나 행동인식에 차이가 나는지를 확인하고, 이차적으로 결과들의 평균을 이용하여 알고리즘을 비교한다.

4.1 앉기 결과에 대한 비교

각 알고리즘을 이용하여 앉기를 인식한 결과이다. 먼저 PSO 알고리즘과 개선된 PSO 알고리즘 결과를 비교하면 수집주기가 1초일 때 까지 분류차이가 0.1% 미만으로

거의 차이가 나지 않는다. 그러나 4초 이후부터 차이가 나기 시작하여 32초일 때 개선된 PSO 알고리즘이 약 5% 정도 높은 인식률을 보이고 있다. 개선된 PSO 알고리즘은 다른 타 알고리즘과 비교하여도 가장 높은 인식률을 보인다.

<표 2> 앉기 행동에 대한 알고리즘별 인식 결과

	PSO	개선된 PSO	MLP	C4.5	SVM
16Hz	99.99%	99.99%	99.98%	99.88%	99.99%
4Hz	99.99%	99.98%	99.93%	99.85%	99.97%
1초	99.99%	99.90%	99.70%	99.65%	99.82%
4초	99.55%	99.84%	98.73%	98.52%	99.29%
8초	98.67%	99.46%	97.44%	97.01%	98.44%
16초	97.66%	98.74%	94.83%	93.97%	96.84%
32초	92.57%	97.89%	89.41%	87.65%	92.94%

4.2 걷기 결과에 대한 비교

<표 3> 걷기 행동에 대한 알고리즘별 인식 결과

	PSO	개선된 PSO	MLP	C4.5	SVM
16Hz	87.91%	88.30%	89.26%	90.37%	97.95%
4Hz	93.25%	93.55%	94.75%	94.55%	99.19%
1초	89.20%	91.96%	98.00%	97.20%	96.40%
4초	76.07%	84.11%	96.88%	90.63%	85.94%
8초	69.77%	79.53%	90.91%	87.88%	72.73%
16초	59.60%	70.00%	81.25%	81.25%	50.00%
32초	50.00%	67.50%	66.67%	77.78%	44.44%

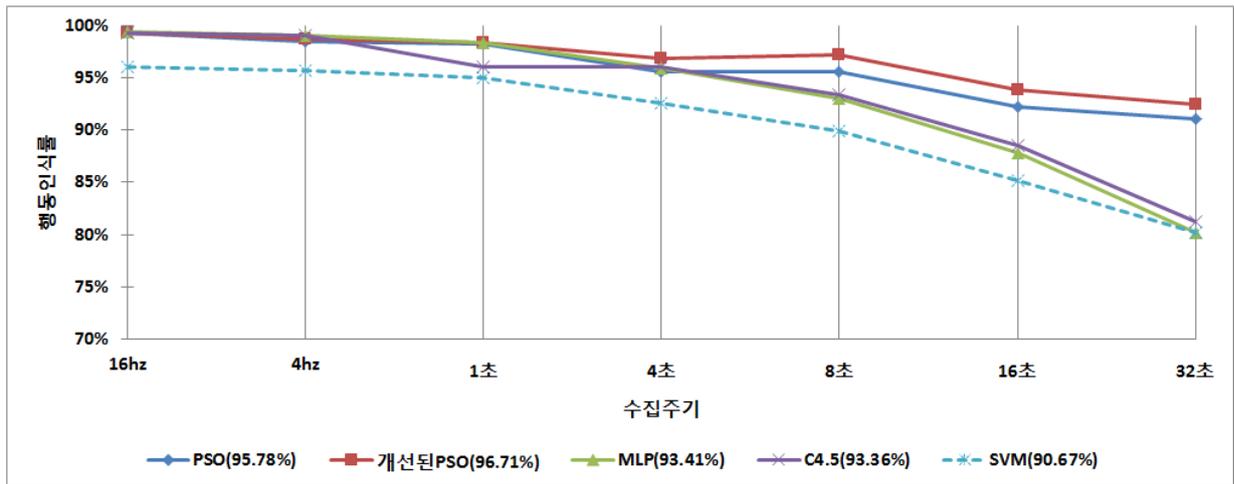
걷기의 경우 개선된 PSO 알고리즘이 모든 수집구간에서 PSO 알고리즘보다 더 나은 결과를 보이고 있다. 특히 32초일 때 최대 17%의 인식 차이를 보이고 있다. 하지만, 앉기의 경우 개선된 PSO 알고리즘보다 C4.5나 MLP가 더 나은 결과를 보이고 있다.

4.3 서기 결과에 대한 비교

<표 4> 서기 행동에 대한 알고리즘별 인식 결과

	PSO	개선된 PSO	MLP	C4.5	SVM
16Hz	97.69%	98.20%	98.25%	97.59%	65.83%
4Hz	94.49%	94.89%	94.31%	94.61%	62.68%
1초	90.12%	90.78%	88.52%	89.72%	58.21%
4초	79.19%	82.49%	74.73%	79.03%	44.09%
8초	74.10%	81.67%	60.22%	67.74%	31.18%
16초	66.57%	74.57%	38.30%	51.06%	10.64%
32초	41.88%	69.38%	17.39%	34.78%	0.00%

서기의 경우 가장 큰 비교결과를 보이고 있다. 1초까지는 1% 미만의 차이를 보이고 있으나 그 이후 인식의 차이가 커지며, 특히 32초일 때 개선된 PSO 알고리즘 결과가 28% 더 나은 결과를 보이고 있다. 특히, 다른 알고리즘과 비교 하더라도 아주 높은 분류 결과를 보이고 있다.



(그림 3) 각 알고리즘별 결과의 평균 그래프

4.4 전체평균 결과

(그림 3)은 각 알고리즘 별 결과의 평균을 그래프로 보여주고 있다. 결과를 통해 개선된 PSO 알고리즘은 기존의 PSO 알고리즘에 비해서 앞기 1%, 걷기 7%, 서기 7% 정도의 개선된 결과를 얻었다. 또한, 다른 알고리즘과 비교를 하더라도 수집환경에 따른 인식의 결과가 높다. 이 점은 스마트폰을 이용하여 사용자의 행동을 인식할 때 고려되는 수집시간, 데이터 길이와 같은 수집환경에서 발생하는 문제점들을 해결할 수 있을 것으로 고려된다.

5. 결론 및 향후 연구

우리는 사용자 행동인식을 하는데 있어 수집 환경에 변화에 따라 PSO 알고리즘이 경계를 벗어나는 데이터를 분류하지 못하는 문제점을 발견하여, 이를 백터길이를 통해 보정하는 알고리즘을 제안하였으며, 실험을 통해 이를 증명하였다. 수집주기를 다양화 하여 사용자 행동을 기록하는 환경을 변화 시켰으며, 이때 수집된 데이터를 이용하여 PSO 알고리즘과 일반적으로 사용되는 다른 알고리즘 과도 인식률을 비교하였다. 결과를 통해서 개선된 PSO 알고리즘이 기존의 PSO 알고리즘에 비해서 모든 행동에서 더 나은 결과를 보였다. 이를 통해 백터 길이를 이용하여 경계를 벗어나는 데이터를 분류하는 방법이 행동인식의 결과를 개선하는데 도움이 된다는 것을 증명하였다. 또한, 다른 알고리즘과 비교하더라도 수집 환경이 변화하더라도 인식의 정확도가 많이 차이 나지 않음을 보인다. 이는 일 상과 같이 수집환경이 다양하게 변하는 환경에서 다른 알고리즘보다 더 나은 결과를 얻을 수 있음을 의미한다.

향후 연구로는 걷기의 분류 결과를 더 개선시킬 방법을 찾는 것이다. 현재 개선된 PSO 알고리즘의 결과는 81.99%로 C4.5와 비교하여 약 7% 낮은 결과를 보이고 있다. 이는 걷기의 데이터와 서기의 데이터가 유사하여 걷기의 데이터를 서기로 잘못 비교하거나, 앞기에서 걷기로 변화되는 지점에서 데이터를 명확하게 인식하지 못하여 생기는 오차이다. 이러한 부분에서 오차를 줄일 수 있는 방

법을 다음 연구에서 진행할 것이다. 또한, 현재는 앞기, 걷기, 서기의 3가지 행동을 인식하고 있으나, 추후 연구에서는 뛰기, 계단 오르고 내리기와 같은 추가적인 행동과 예외적인 행동(전화받기, 어플리케이션 동작시키기 등)을 인식하여 처리하는 방법을 고려하고자 한다.

참고문헌

- [1] T. Ito, T. Ishihara, Y. Nakamura, S. Muto, M. Abe, Y. Takagi, "Prospects for Using Lifelogs in the Medical Field", NTT Technical Review, Vol.9, no.1, 2011. 1.
- [2] Kyuchang Kang, Seonguk Heo, Changseok Bae, Dongwon Han, "Mobile Health Screening Form Based on Personal Lifelogs and Health Records", in Proc. Information Technology Convergence and Services, pp. 557-565, 2011.
- [3] Kennedy, J., Eberhart, R., "Particle Swarm Optimization", in Proc. International Conference on Neural Network, pp.1942-1948, 1995.
- [4] Shi, Y., Eberhart, R., "A Modified Particle Swarm Optimizer", in Proc. International Conference on Evaluation Computing, pp.69-73, 1998.
- [5] Changseok Bae, Wei-Chang Yeh, Noorhaniza Wahid, Yuk Ying Chung, Yao Liu, "A new simplified Swarm Optimization (SSO) Using Exchange Local Search Scheme", International Journal of Innovative Computing, Information and Control, vol. 8, no. 7, pp.1-11, 2012.
- [6] Khan, A.M, Young-Koo Lee, Lee, S.Y, Tae-Seong Kim, "A Triaxial Accelerometer-Based Physical-Activity Recognition via Augmented-Signal Features and a Hierarchical Recognizer", IEEE Trans. Information Technology in Biomedicine, vol. 14, no. 5, 2010.
- [7] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, "Activity Recognition from Accelerometer Data", in Pro. American Association for Artificial Intelligence, pp. 1541-1546, 2005.