

빅데이터 기반의 실시간 생체 신호 모니터링을 이용한 분석시스템 : 야구 수비능력 측정을 중심으로

오영환*

An Analysis System Using Big Data based Real Time Monitoring of Vital Sign: Focused
on Measuring Baseball Defense Ability

Young-Hwan Oh*

요 약

빅데이터(Big data)는 제4차 산업혁명 시대를 맞이하여 과학, 기술, 산업, 사회분야에서 사물인터넷(IoT), 인공지능(AI), 클라우드(Cloud)와 더불어 공공분야와 민간분야를 아우르는 곳에서 중요한 키워드가 되고 있다. 빅데이터 기반의 서비스는 교통, 기상, 의료, 마케팅 등의 다양한 분야에서 제공되고 있다. 특히 스포츠 분야에서는 병원이나 재활센터가 아닌 훈련이나 일상 생활에서 생체 신호(Vital sign)를 측정할 수 있는 웨어러블 장치(Wearable device)의 등장으로 여러 형태의 생체 신호를 수집, 관리할 수 있게 되었다. 하지만 아직까지 스포츠분야, 즉 야구선수의 훈련(training)과 재활(rehabilitation)을 위한 웨어러블 장치에서 추출된 생체 신호를 가지는 빅데이터에 대한 연구가 활성화되지 못하고 있다. 따라서 본 논문에서는 야구선수에 대한 훈련, 특히 내야와 외야 수비수수에 대한 운동량 측정 생체신호를 빅데이터 기반으로 저장하고 분석할 수 있는 시스템에 대한 연구를 제안한다.

ABSTRACT

Big data is an important keyword in World's Fourth Industrial Revolution in public and private division including IoT(Internet of Things), AI(Artificial Intelligence) and Cloud system in the fields of science, technology, industry and society. Big data based on services are available in various fields such as transportation, weather, medical care, and marketing. In particular, in the field of sports, various types of bio-signals can be collected and managed by the appearance of a wearable device that can measure vital signs in training or rehabilitation for daily life rather than a hospital or a rehabilitation center. However, research on big data with vital signs from wearable devices for training and rehabilitation for baseball players have not yet been stimulated. Therefore, in this paper, we propose a system for baseball infield and outfield players, especially which can store and analyze the momentum measurement vital signals based on big data.

키워드

Big Data, Baseball Defense Ability, Wearable Device, Vital Sign, Accelerometer Sensor
빅데이터, 야구 수비 능력, 웨어러블 기기, 생체 신호, 가속도 센서

* 나사렛대학교 정보통신학과 교수(yhoh@kornu.ac.kr) • Received : Nov. 27, 2017, Revised : Jan. 06, 2018, Accepted : Feb. 15, 2018
• 접수일 : 2017. 11. 27
• 수정완료일 : 2018. 01. 06
• 게재확정일 : 2018. 02. 15
Dept. of Information and Communication, Korea Nazarene University,
Email : yhoh@kornu.ac.kr

1. 서 론

빅 데이터란 종래의 데이터베이스 환경에서 처리하던 데이터에 비해 그 크기와 양이 거대하고 생명주기도 짧고 종류와 형태에서도 문자, 숫자 데이터뿐 아니라 영상, 음성, 신호 데이터를 포함하는 방대한 양의 데이터를 의미한다. 단지 데이터의 크기가 거대하다고 빅 데이터라고 하지는 않는다. 데이터 양(Data volume), 데이터 속도(Data velocity) 그리고 데이터 다양성(Data variety)의 3차원적 특징이 모두 나타날 때를 의미한다[1-2]. 스포츠 경기에서 빅데이터 활용은 필수이다. 빅 데이터는 최근 전 세계 스포츠업계의 주요 관심사로서 축구, 야구, 농구 등의 이미 다양한 분야에서 많이 활용되고 있다[3]. 머니볼 이론(Moneyball theory)이란 영화 ‘머니볼(Money ball)’을 통하여 널리 알려지게 된 미국 프로야구단 오클랜드 애슬레틱스(Oakland Athletics)에서 구단주 빌리 빈(William Lamar Beane)이 경기 데이터를 세밀하게 분석하여 오직 데이터만으로 선수들을 배치하여 승률을 높였다는 게임 이론이다[4]. 리그에서도 최하위 팀을 4년 연속 포스트시즌(Post season)에 진출시키고, 메이저 리그(Major league) 최초로 20연승이라는 신기록을 세우기도 하였다. 최근 들어서 정보통신과 영상 기술의 발달로 인하여 세밀하고 방대한 야구 경기 데이터의 수집이 가능하여졌으며, 트랙맨(Track man)과 같은 군사기술과 스탯캐스트(Statcast)와 같은 수리통계학을 이용하여 투수의 투구 궤적, 타구 방향과 속도, 수비수의 움직임 그리고 주루플레이 등 야구의 모든 경기를 데이터화할 수 있게 되었다. 기존의 문자, 숫자 형태의 정형 데이터뿐만 아니라 센서 신호와 같은 아날로그 형태의 비정형 데이터를 획득, 저장, 분석 그리고 활용을 하여, 최근 야구경기의 흥행을 가져다 주는 빅 데이터의 중요성이 더욱 커지고 있다[5].

2017년 현재 대한야구협회(Korea Baseball Association)에 등록된 고교야구는 75개팀이 있으며 고교야구 주말리그, 평가전, 지역대회 등 많은 경기수와 과도한 훈련으로 부상과 밀접한 관계를 가지고 있다. 고교야구 주말리그제는 문화체육관광부, 대한야

구협회가 2011년부터 시행한 제도로서 ‘공부하는 학원 스포츠’라는 취지를 가지고, 전국을 8개 권역으로 나눠 7-8개 팀을 배정하여 주말과 공휴일에 열리고 있다¹⁾. 주중에는 일반 학생들과 함께 정규 수업을 듣고, 방과 후에 훈련과 시합을 진행하는 것이 원칙이다. 그러나 현재 고교야구 선수들의 진로에 있어서 제한된 프로야구 입단과 대학 진학을 위한 훈련시간이 현저히 줄어들어 따라 고강도 훈련과 경기로 인하여 항상 부상에 노출되어 있는 실정이다[6]. 고교야구선수들은 대학선수, 프로선수들과의 격차를 줄이기 위해 과거에 비해 체격, 체력 및 기술적인 모든 측면에서 비약적인 발전을 이루어 왔다. 그러나 야구선수의 기초자료가 되는 순발력, 지구력, 운동신경능력 등의 경기력 평가와 평가기준에 대한 연구들이 프로야구 선수를 대상으로 국한되어 왔으며, 고교야구나 대학야구 선수들에 대한 기본적인 연구들은 미비한 실정이다. 야구 운동 평가에 있어 근력, 지구력, 순발력 평가는 선수와 지도자 및 트레이너에게 훈련 계획을 수립하고 훈련결과를 평가하는데 있어 매우 중요한 정보로서 활용될 수 있다.

생체 신호란 인간의 신체에서 측정되는 신호의 순간 값을 의미하며, 각 개인의 질병과 건강에 대해서 의료기관이나 국가 지정 병원에서 치료와 재활 목적을 가지고 정보로 활용되는 경우가 많다[7]. 그 중에서 주로 관리되는 생체 신호는 혈압(Blood pressure), 맥박(Pulse), 심전도(Electrocardiogram) 등이 있고, 병원이나 재활원, 요양원에서 이를 모니터링하는 연구와 더불어 스포츠와 레저 분야에서 간편하게 모니터링하는 기기에 대한 관심도 점점 증가하고 있다. 생체 신호 데이터는 크게 2가지의 데이터 종류로 나눌 수 있다. 첫 번째, 체온(Body temperature), 혈압, 맥박 등과 같이 단순한 값을 가지는 비연속, 비정형적인 데이터 형태이다. 두 번째, 심전도, 가속도(Accelerometer), 호흡(Respiration) 등과 같이 연속적으로 유입되는 데이터이다. 연속적인 데이터의 경우 그 데이터 양이 방대하기 때문에 종래의 데이터베이스에 저장하기에 적합하지 않다. 불연속적인 자료인 경우에도 단순 값을 가지지만 주기가 매우 짧아 질수록, 이를 측정하는 경우 데이터 양이 많아지게 되어 기존의 DBMS(Database

1) 대한야구소프트볼협회, <https://www.korea-baseball.com/>

Management System)에 저장하기에 적합하지 않다. 종래의 데이터베이스에 저장하기에 적합하지 않은 방대한 양의 데이터를 빅 데이터로 분류하여 관리하여야 한다. 생체신호는 원 데이터(Raw-Data)를 가지고 다양한 알고리즘을 이용하여 특정 값을 추출하고, 분석이 가능한 정보로 저장/관리해야 한다.

이러한 패러다임의 변화에 따라 야구선수의 훈련과 재활에서 만들어지는 생체신호 빅데이터를 이용하여 데이터를 저장하고 분석 및 처리하는 기술이 있어야 한다. 야구선수의 빅데이터는 선수들이 경기 훈련 전에 몸에 센서(Sensor)를 부착하는 방식으로 수집된다. 훈련 도중 발생하는 선수들의 심전도와 맥박, 혈압 등의 데이터를 수집한다. 또한 가속도 센서(Acceleration sensor)와 자이로 센서(gyro sensor)로 만들어지는 움직임과 거리를 환산하여 유도되는 순발력 지수를 이용하여 각 선수를 위한 수비 분석 플랫폼을 통해 선수들의 기량을 분석하고 전략을 짜는데 활용한다.



그림 1. 야구 수비 동작
Fig. 1 Baseball defensive motion

다음 그림 1.은 야구 수비동작에 대한 예를 보여준다. 야구는 투수놀음이라는 말이 있듯이 좋은 투수와 훌륭한 타격 선수가 있는 팀이 강한 팀이고, 플레이오프나 한국시리즈에서 높은 성적을 낼 수 있는 중요한 요소이다. 그러나 작은 실수는 대량 실점으로 이어져 패배로 나타나는 경우가 적지 않다. 야구 선수의 수비 범위와 능력은 승패의 결정적 요인이다. 아직까지 수비에 대한 야구 선수에 대한 운동량 분석이나 평가를 위한 측정 기준이 거의 없는 편이다. 야구선수 포지션 중에 투수의 투구 동작(직구, 커브)이나 프로야구 선수의 타격 동작에 대한 운동학적 분석에 대한 연구는 있다[8]. 그러나 야구선수의 수비 동작에 대한 운동학

적, 생체학적 분석을 다룬 연구는 거의 없다. 선수들의 움직임을 분석할 수 있는 가속도 센서와 생체의 변화를 모니터링 할 수 있는 심박수 센서 착용을 허용하려는 움직임도 있다. 그 이유는 운동량 측정과 경기력 향상을 위한 방안이 되기도 하지만, 선수들을 부상이나 사고로 인한 경기력 부재를 탈피하려는 시도로 해석된다.

본 논문에서는 야구선수의 훈련과 재활을 위하여 웨어러블 장치에서 추출된 생체 신호를 가지고 이를 빅데이터로 관리하며 야구 수비선수에 대한 운동량 측정 생체신호를 저장하고 분석할 수 있는 시스템에 대한 연구를 진행하였다. 아래 그림 2.는 야구 수비선수 포지션 이동에 관한 예를 보여 준다.

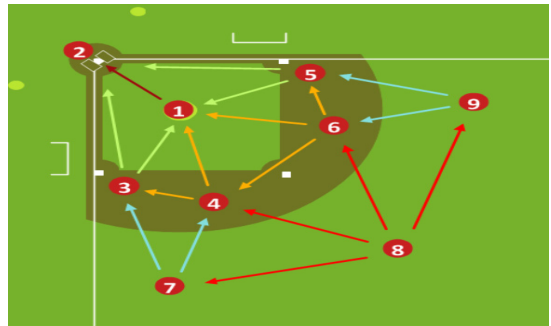


그림 2. 야구 수비선수 포지션 이동
Fig. 2 Position shift of baseball defensive player

II. 웨어러블 기기를 이용한 생체신호 획득

모바일 기기(Mobile device)와 스마트 기기(Smart device)는 인터넷과 전자통신기술의 유래없는 발전으로 말미암아 새로운 패러다임의 전환을 이끌고 있으며, 신체에 부착 가능한 형태의 웨어러블 기기(Wearable device)는 기존의 디바이스로 불가능하였던 새로운 부가가치 창출을 가속시키고 있다. MIT 미디어랩(Media Lab)에서는 웨어러블 디바이스는 신체에 부착하여 컴퓨팅 행위를 할 수 있는 모든 것을 지칭하며 일부 컴퓨팅 기능을 수행할 수 있는 애플리케이션을 포함하고 있다고 정의한다²⁾. 웨어러블 기기는 인간 신체의 생리학적 신호를 감지하기 위한 다수의 센서들이 사용자에게 불편함없이 부착되고, 연속적

으로 실시간 신호 측정이 되도록 한다. 이렇게 측정된 신호는 블루투스(Bluetooth), 지그비(Zigbee)와 같은 무선 통신기술을 통하여 서버로 전송되며, 획득된 대부분 신호는 심박수, 체온, 혈압, 맥박 등의 생체 데이터이다.

현재 웨어러블 기기를 각종 스포츠 선수 인체에 부착하여 웨어러블 기기의 다양한 센서를 통하여 신체 안의 생체 데이터를 획득, 저장, 분석하여 새로운 서비스와 가치를 창출하고, 사용자의 운동량과 신체 데이터를 수집할 수 있도록 확장한 후 유/무선 네트워크를 통해 클라우드에 저장하고 이를 방대한 양의 빅데이터 분석을 하도록 지원한다. 이후 분석된 빅데이터를 기반으로 하여 사용자 위치, 심전도, 근전도, 운동량, 스트레스 정도와 선수의 체온, 체중 등의 변화 등에 맞추어 사용자가 원하는 기능까지 예측해서 서비스를 제공할 수 있다[9].



그림 3. 웨어러블 기기의 종류
Fig. 3 Types of wearable devices

III. 야구선수의 생체신호 분석

야구선수의 개인적인 수비 능력은 주로 운동 신경(Nerve ability), 체격(Physical status), 체력(Physical strength), 집중력(Concentration)이 이를 결정한다. 특히 수비 움직임(Defensive movements)의 보폭은 걸을 때나 달릴 때의 크기와 신체의 길이에 따라 다르다. 본 연구에서는 각 야구선수의 수비 능력을 가속도 센서를 이용하여 움직임 속도 및 운동량을 측정한다.

3.1 실험 방법

본 실험 참여자는 K고등학교 야구부에서 운동하고 있는 내야와 외야선수 8명을 임의로 선택하여 운동량 측정을 위한 실험을 하였다. 이들의 평균 나이는 17.9세이고, 평균 키는 178.3cm 그리고 평균 몸무게는 73.4kg이었다. 측정된 운동량분석은 감독으로 하여금 학생들의 수비 위치와 운동능력을 위한 포지션 결정을 하는 데 도움을 주는 목적이 있음을 알려주었다. 아래 그림 4는 야구 선수의 수비능력 측정에 대한 분포 예를 보여준다.

3.2 실험 장비

가속도 센서를 이용하여 일상생활 및 운동, 게임시장에서 인간의 행동을 인식하고 보행수와 주행 형태를 측정하여 여러 분야에 활용하려는 시도가 다양하게 진행되고 있다[10-11]. 가속도 센서와 자이로센서를 동시에 사용하면 사람이 이동하거나 달리는 경로를 추적할 수 있다. 또한 이동거리도 함께 계산이 가능하다.

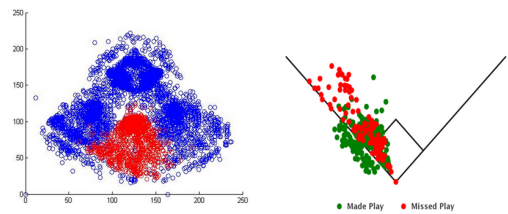


그림 4. 수비능력 측정
Fig. 4 Defensive ability measurement

본 연구에서 사용되는 웨어러블 기기는 16bit MCU(Micro Controller Unit)와 가속도 센서(자이로 센서 포함) 그리고 블루투스 통신 모듈로 구성되어 있다. 이 모듈을 야구 선수의 신체에 부착하고 어떠한 방향으로든지 움직이게 되면, 움직인 경로와 거리를 환산하여 데이터로 변환할 수 있다. 즉, 가속도 센서는 각 방향별 이동거리를 계산하고 자이로 센서는 이동방향을 보다 정확하게 계산할 수 있다.

2) The Wearable Computing Market:A Global Analysis, GigaOm Pro



Part #	Gyro Full Scale Range	Gyro Sensitivity	Gyro Rate Noise	Accel Full Scale Range	Accel Sensitivity	Digital Output	Logic Supply Voltage	Operating Voltage Supply
UNITS:	(°/sec)	(LSB/°/sec)	(dps/√Hz)	(g)	(LSB/g)		(V)	(V +/-5%)
 MPU-6000	±250 ±500 ±1000 ±2000	131 65.5 32.8 16.4	±2 ±4 ±8 ±16	16384 8192 4096 2048		PC or SPI VDD	2.375V~ 3.46V	4x4x0.9
 MPU-6050	±250 ±500 ±1000 ±2000	131 65.5 32.8 16.4	±2 ±4 ±8 ±16	16384 8192 4096 2048		PC 1.8V±5% or VDD	2.375V~ 3.46V	4x4x0.9

그림 5. MCU(Micro Controller Unit)
Fig. 5 MCU(Micro Controller Unit)

3.3 측정 방법

선수들은 심장근처의 몸 표면과 오른쪽 팔 상박에 가속도 센서 모듈을 부착하고, 중력 가속도의 신호를 X, Y, Z축 방향으로 각각의 아날로그 신호를 획득하여, 이를 디지털 신호로 변환한다. 이 가속도 센서를 이용하면 획득된 센서 데이터를 이용하여 운동에 따른 근육의 사용도와 힘에 미치는 영향을 운동별로 데이터 측정을 할 수 있다. 본 논문에서는 가속도 센서를 이용하여 운동별 센서 데이터의 차이점과 활동량에 대해서 측정을 하여 분석을 한다. 이를 이용하여 신체 움직임의 강도와 칼로리 소모량을 측정할 수 있다. 운동부하 검사(Exercise stress test)는 야구 선수에게 수비 연습을 시켜 운동량을 늘린 상태에서 심전도와 혈압, 맥박을 측정하여 생체 신호 데이터를 얻는다. 먼저 운동 전 휴식 상태에서 심전도, 맥박, 혈압을 측정한다. 운동을 시작하여 운동 강도 단계를 조절하면서 검사 동안 지속적으로 심박동수와 심전도가 모니터에 기록되고, 약 10분 정도의 간격으로 혈압을 측정한다.

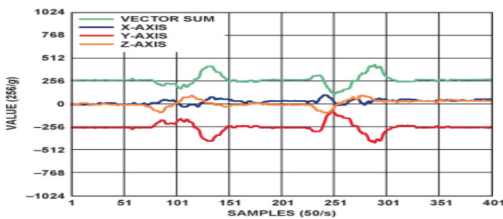


그림 6. 가속도 센서의 측정 값
Fig. 6 Measured value of acceleration sensor

3.4 데이터 분석

가속도 센서는 중력 가속도를 포함하고 타임스탬프와 x, y, z 축에 작용하는 모든 가속도를 출력하기 때

문에 이 정보를 그대로 적용하면 큰 오차를 가진 값을 갖게 된다. 그래서 자이로 센서를 추가시켜 출력 데이터를 보정하고 이를 이용하여 보행 수와 거리를 산출한다. 야구 선수는 생체 신호 획득을 위한 웨어러블 디바이스를 착용한 후 각각의 내야와 외야 수비 동작에 대하여 50분의 1초 단위로 250개씩 데이터를 샘플링한다. 가속도 센서를 이용한 보폭측정에는 중력가속도 성분 출력으로 인한 적분 오차 때문에 한계가 있기 때문에 이를 보정하기 위하여 가속도 센서의 타임스탬프와 x, y, z축 출력을 하나의 대표 값으로 처리하는 벡터합산(Vector Sum)를 사용한다.

IV. 빅데이터 기반의 생체신호 모니터링 시스템

본 장에서는 웨어러블 기기에서 측정되는 신호 중심전도와 가속도, 맥박, 혈압 신호에 대하여 빅데이터 전송과 저장을 설명한다. 생체신호는 HDFS(Hadoop Distributed File System)에 저장되며 비정형의 생체신호를 분석 가능한 데이터로 변환하는 맵 리듀스(Map/Reduce)를 이용한다.

4.1 생체신호 특정값 추출

웨어러블 기기로 측정이 가능한 생체신호의 종류는 심전도, 가속도, 혈압, 맥박 등의 신호가 있다. 심전도(ECG)는 심장근처의 몸의 표면에 부착하여 기록한 것으로 심장의 상태를 진단하는 중요한 수단으로 활용된다. 심장 주위에서 측정된 ~mV 범위의 심전도는 LPF(Low pass filter)를 통하여 필터링되고, IA(Instrumentation Amplifier)의 입력된 후 증폭되어 심전도 파형이 1Hz~35Hz 범위의 심전도 신호로 출력된다. 아날로그 형태의 심전도 파형은 MCU의 ADC(Analog_to_Digital Conveter)에서 디지털 신호로 변환된다. 디지털 신호로 변환된 심전도 신호는 블루투스 통신 모듈을 거쳐서 데이터로 저장된다. 본 연구에서 1초에 보통 200개 정도의 데이터를 생산하는 심전도신호는 1시간 측정했을 때 720,000개의 데이터가 생성된다. 따라서 측정되는 심전도 신호의 데이터는 소수점 이하의 큰 길이를 가지므로 빅데이터 성격을 갖는다.

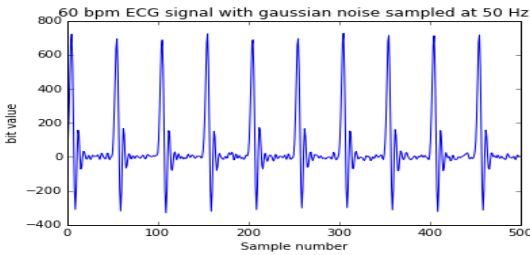


그림 7. 심전도(ECG) 신호
Fig. 7 Electrocardiogram(ECG) signal

가속도(ACC)는 야구 선수의 수비 움직임을 측정할 수 있는 신호로 X축, Y축, Z축으로 출력된다. 웨어러블 기기로부터 추출되는 가속도 신호는 1초당 60개의 신호를 저장하는데 1시간 측정 시에 216,000개의 데이터가 만들어 진다.

$$\begin{aligned}
 x &= A \sin(\omega t) \\
 \dot{x} &= \omega A \cos(\omega t) \\
 \ddot{x} &= -\omega^2 A \sin(\omega t)
 \end{aligned}$$

Displacement (unit: m) ↓

Velocity (unit: m/s) ↓

Acceleration (unit: m/s²) ↓

Differential ↑

Integral ↑

그림 8. 가속도, 속도, 변위의 미적분
Fig. 8 Calculus of acceleration, velocity and displacement

가속도센서로 측정된 가속도 시간신호(time signal)는 수학적으로는 한번 적분하면 속도신호로, 두 번 적분하면 변위신호로 변환될 수 있다. 그러나 실제로는 가속도 시간신호를 적분해서 속도 시간신호, 변위 시간신호를 만드는 것은 어려우며, 그 이유는 가속도신호 측정과정에서 측정 노이즈(noise)가 발생하여 시간 적분과정에서 변환된 신호를 왜곡시키기 때문이다. 가속도 신호를 이용하여 획득할 수 있는 특징 값으로는 보행수, 칼로리, 이동거리 등이 있으며 걷기, 달리기를 분석할 수 있다.

혈압과 맥박은 운동선수의 중요한 생체신호로서 운동을 할시, 심장의 활동으로 혈압도 오르고 맥박도 빨라진다. 그러나 혈관에 문제가 있을 시 맥박과 혈압에 문제가 발생한다. 성인의 맥박수는 분당 60~80회이다. 나이가 적을수록 많아져서 고등학교 학생 나이에 해당하는 맥박수는 평균 80~100회이다. 본 실험에서 측정된 신호는 1초에 보통 300개 정도의 생체 데이터를

생성하므로 방대한 데이터와 짧은 주기를 갖는 선형 구조의 형태이므로 빅 데이터의 속성을 가지고 있다.

심전도, 가속도, 혈압, 맥박 등의 신호로부터 추출된 특정 값은 정형화되어 있으나 종래의 데이터베이스로 관리하기에는 부적합하다. 따라서 추출된 신호는 양과 가치에 의해 빅데이터로서의 저장과 분석이 필요하다.

4.2 생체신호의 빅데이터 저장

아파치 하이브(Apache Hive)는 하둡(Hadoop)에서 실행되는 데이터 웨어하우스(Data Warehouse) 하부 구조로서 데이터의 생성, 질의 그리고 분석 기능을 제공한다[12]. 이는 Apache HDFS이나 Apache HBase와 같은 저장 시스템을 이용하여 대용량 데이터 집합들을 분석한다. HiveQL은 DBMS의 SQL(Structure Query Language)과 같이 데이터 생성 및 질의, 제어 언어를 제공하며 Map/Reduce의 기능도 지원한다. 그림 9.에서는 학생 야구선수 데이터를 분석하기 위한 HiveQL 쿼리문을 작성하였다. 참고로 HiveQL은 대소문자를 구분하지 않는다.

```

hive> CREATE TABLE students(ID INT, name STRING, Height FLOAT,
Weight FLOAT)
> PARTITIONED BY (ver timestamp)
> ROW FORMAT DELIMITED FIELDS TERMINATED BY ';'
> LINES TERMINATED BY '\n'
OK
Time taken: 0.461 seconds
    
```

그림 9. HiveQL 데이터 정의어(DDL)
Fig. 9 HiveQL Data Definition Language(DDL)

다음 그림 10.은 Hive는 local file system에 있는 데이터와 HDFS에 저장된 데이터를 모두 업로드한다. Hive CLI에서 다음과 같이 LOAD DATA를 입력한다.

OVERWRITE INTO 절은 중복된 데이터가 있어도 무시하고 입력한다. PARTITION 절은 파티션 키인 ver 값을 '2017-10-12'로 설정해 데이터를 입력하는 설정이다. 생체신호 Raw-data로부터 여과와 정제 과정을 통하여 추출된 파일을 Table 형식에 맞게 저장한 후, HiveQL을 이용하여 Hive에 업로드한다.

```
hive> LOAD DATA LOCAL INPATH
'/home/user/hadoop/study/baseball/K20171012.csv'
> OVERWRITE INTO TABLE students
> PARTITION (ver='2017-10-12');
#결과#
Loading data to table default.students partition (ver='20171012)
OK
Time taken: 1.345 seconds
```

그림 10. HiveQL 데이터 저장
Fig. 10 HiveQL data store

4.3 생체신호의 빅데이터 추출

하둡의 HDFS 저장소에는 학생 야구선수의 가속도 센서 값이 timestamp, x축, y축, z축을 가지고 웨이블 형태로 저장된다. 이는 의미가 없는 비정형의 수치값으로 원데이터 형태이며 분석이 가능한 정보로 변환을 할 필요가 있는 경우 Map/Reduce 처리를 통하여 분석가능한 특정 값 데이터로 추출한다. 다음 그림 11.은 Map/Reduce로 변환된 가속도 신호의 샘플을 보여준다.

```
1602184302300, 6.106235, 1.23234566, 7.3443424
1602184302310, 6.345345, 0.93785444, 7.2034576
1602184302320, 5.834034, 0.74345655, 8.7425045
1602184302330, 5.967543, 0.82395656, 8.6023458
1602184302340, 6.158355, 1.18507834, 8.1340789
1602184302350, 6.530476, 1.05468325, 7.9673589
1602184302360, 5.997458, 0.9325674, 7.8904568
```

그림 11. Map/Reduce로 변환된 가속도 신호
Fig. 11 Converted acceleration signal to map/reduce

V. 결론 및 향후개선 방향

본 논문에서는 고교야구 선수에 대한 훈련과 평가, 특히 내야와 외야 수비수수에 대한 운동량 측정 생체 신호를 빅데이터 기반으로 저장하고 분석할 수 있는 시스템을 제안하였다. 야구 선수들은 공격과 수비에서 타격, 타율, OPS, 어시스트 등의 정량적 데이터 측정에 의존하여 평가되어 왔다. 향후 빅 데이터 스포츠 분석과 전략 시스템을 이용하여 야구 선수의 운동 신경 능력과 순발력, 근력을 포함하는 인지적 생체 데이터까지 변환하여 수집, 저장, 관리 그리고 분석하는 기술로 발전할 여지가 많다. 이는 야구감독과 트레이너, 코치진이 생체신호 모니터링 분석 시스템을 이용

하여 야구 선수들의 체계적인 훈련량을 조절할 수 있고, 재활에 대해서도 도움을 줄 수 있다.

감사의 글

이 논문은 2017년도 나사렛대학교 학술연구비 지원에 의해 연구되었음

References

- [1] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, and A. H. Byers, *Big data: The next frontier for innovation, competition, and productivity*, Chicago: McKinsey Global Institute, 2011.
- [2] S. Yoon, H. Namgung, S. Yang, and H. Kim, "Big Data Driven Semantic Web Technology Trends large," *J. of the Korean Institute of Communications and Information Sciences*, vol. 29, no. 11, 2013, pp. 24-29.
- [3] J. Cho, "Utilization and Prospect of Sport Big Data," *The Korean Journal of Measurement and Evaluation in Physical Education and Sports Science*, vol. 14, no. 3, 2012, pp. 01-11.
- [4] E. Wassermann, D. R. Czech, M. J. Wilson and A. B. Joyner, "An Examination of the Moneyball Theory: A Baseball Statistical Analysis," *The Sport Journal*, vol. 19, no. 1, 2005, pp. unpaginated.
- [5] W. Gin, "Big data and labor: What baseball can tell us about information and inequality," *J. of Information Technology & Politics*, vol. 15, no. 1, 2017, pp. 66-79.
- [6] S. Kim, K. Han, and S. Zhang, "Norm-Referenced Criteria for Isokinetic Strength of the Lower Limbs for the Korean High School Baseball Players," *The Korean Journal of Sports Medicine*, vol. 34, no. 1, 2016, pp. 48-56.
- [7] M. Joo, D. Ko, and H. Kim, "Development of Smart Healthcare Wear System for Acquiring Vital Signs and Monitoring Personal Health," *J.*

- of Korea Multimedia Society, vol. 19, no. 5, 2016, pp. 808-817.
- [8] Y. Lee, "The Kinematic Analysis of the Pitching motion for the Straight and Curve ball," Master's thesis, Chang-won National University, 2001.
- [9] W. Raghupathi and V. Raghupathi, "Big data analytics in healthcare: promise and potential," *Health Information Science and Systems*, vol. 2, no. 3, 2014, pp. unpaginated.
- [10] S. Kim, "A Way of Unusual Gait Cognition for Life Safety," *J. of the Korea Institute of Electronic Communication Sciences*, vol. 11, no. 2, Feb. 2016, pp. 215-222.
- [11] D. Suh and Y. Oh, "A Novel Way of Safety Awareness on the Walking with Single Sensor," *J. of the Korea Institute of Electronic Communication Sciences*, vol. 11, no. 2, Feb. 2016, pp. 197-202.
- [12] P. Zikopoulos, C. Eaton, D. DeRoos, T. Deutsch, and G. Lapis, *Understanding Big Data - Analytics for Enterprise Class Hadoop and Streaming Data*, New Delhi: McGraw-Hill Osborne Media, 2012.

저자 소개



오영환(Young-Hwan Oh)

1991년 인하대학교 전자계산공학과 졸업(공학사)

1997년 인하대학교 대학원 전자계산공학과 졸업(공학석사)

2001년 인하대학교 대학원 전자계산공학과 졸업(공학박사)

2002년 나사렛대학교 정보통신학과 교수

※ 관심분야 : 빅데이터, 상황인지, 사물인터넷