

볼츠만 머신을 이용한 인간 모션 학습 및 생성

Learning and Generation of Human Motion using Boltzmann Machines

이충연¹, 김지섭¹, 김은솔¹, Karinne Ramirez Amaro², Michael Beetz², 장병탁¹

한국정보과학회 제38회 추계학술발표회
2011. 11. 26

¹ 서울대학교 컴퓨터공학부 바이오지능연구실

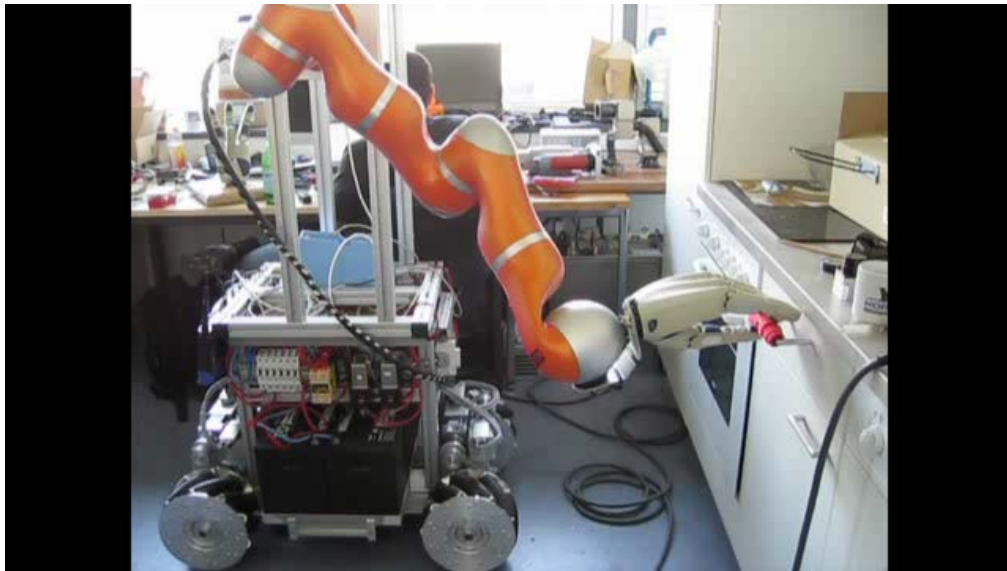
² Technische Universität München, Intelligent Autonomous System Group

목차

- 연구동기
- 연구목표
- 모션 데이터
- 모션 데이터 학습 및 생성
- 실험결과
- 결론 및 향후연구

연구동기

- 기계학습 기반의 로봇 모션 생성
 - 상황에 따른 유연한 로봇 제어가 필요
 - 기계학습 알고리즘을 통해 인간 모션을 학습 및 생성
 - 인간의 행동 학습 메커니즘을 로봇에게 적용



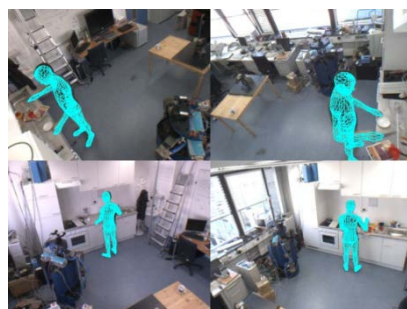
TUM Rosie



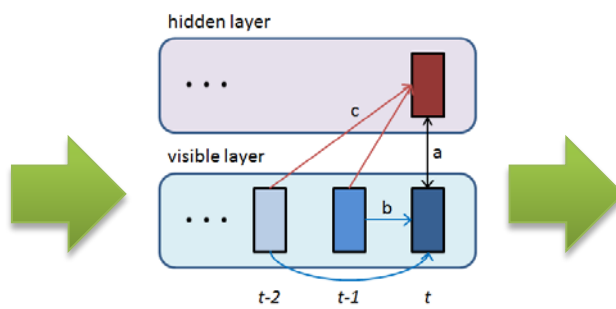
Fujitsu HOAP-3

연구목표

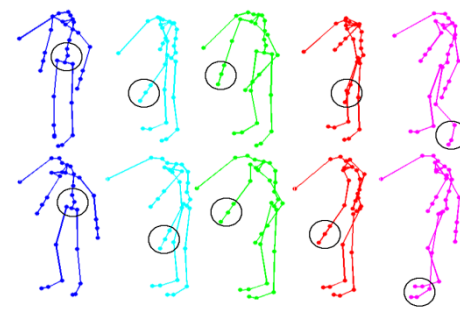
- 인간 모션 데이터 획득
 - 주방 환경과 같은 실제 일상 생활에서의 인간 모션 데이터 획득
 - 로봇이 카메라를 통해 학습할 수 있도록, **markerless 방식의 모션캡처 시스템을 활용**
- 기계학습 알고리즘을 이용하여 인간 모션 학습 및 생성
 - **Conditional Restricted Boltzmann Machine (cRBM)**을 사용하여 획득한 모션을 학습
 - 동일한 스타일의 모션을 생성해낼 수 있는 nonlinear generative model
 - 실수값의 이동 좌표 및 관절의 회전각도로 구성된 모션 데이터를 학습 가능



인간 모션 데이터 획득



모션 데이터 학습

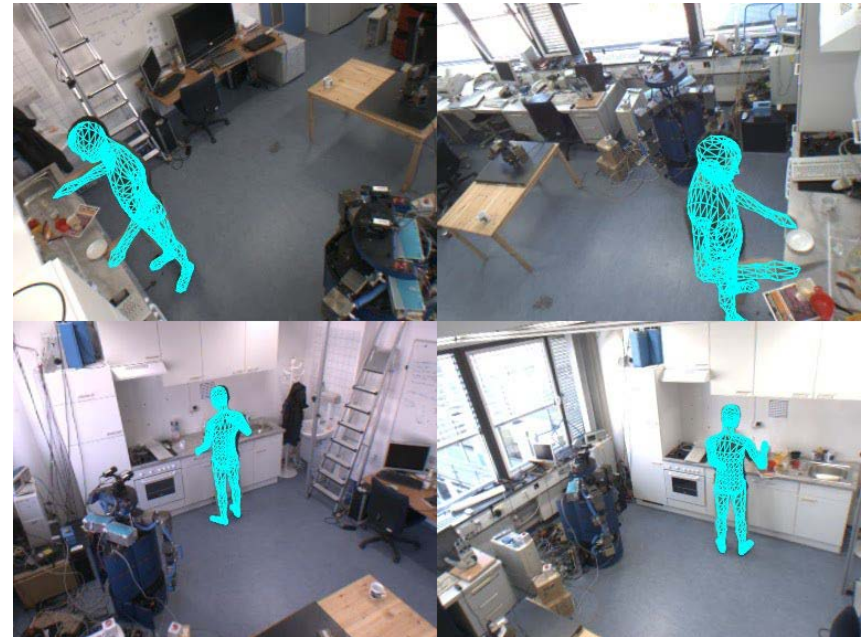
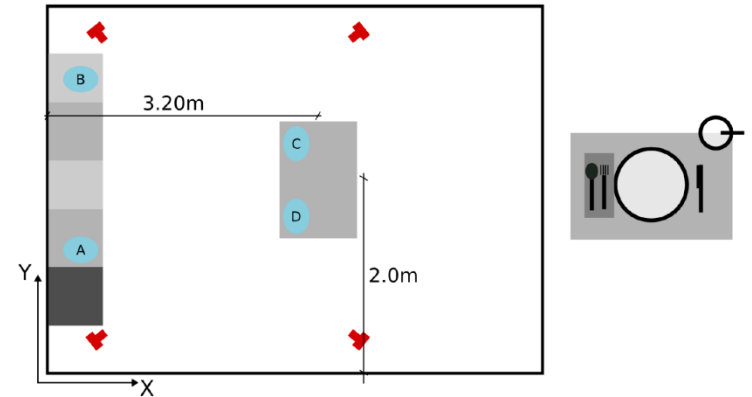


모션 데이터 생성

모션 데이터(1/2)

■ TUM Kitchen Data Set

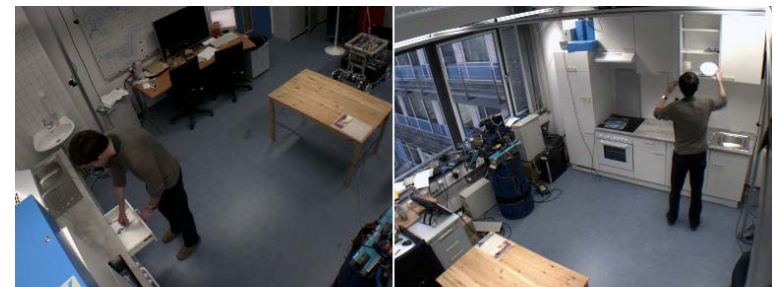
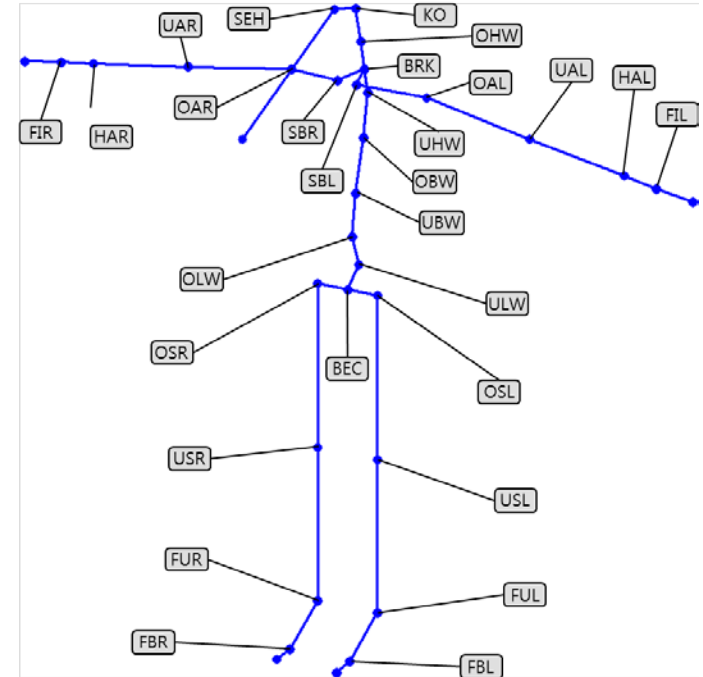
- 주방 환경에서 테이블을 세팅하는 동작
- 마커를 부착하지 않은 피험자를 촬영
- 촬영한 영상으로부터 모션 데이터 획득
- <http://ias.in.tum.de/software/kitchen-activity-data>



모션 데이터(2/2)

■ 데이터 명세

- BioVision사의 BVH 포맷
- 총 33개 관절에 대한 모델 구조 (skeleton hierarchy) 및 프레임별 관절들의 회전값을 포함
- 프레임 구성 정보
 - 관절 회전각 (Euler angles, 1-3 DOF)
 - 중심관절(BEC, 미저골)의 회전각 및 위치값
 - 나머지 관절들의 위치는 BEC와의 상대적인 오프셋으로 표현됨
- 단순 모션 학습을 위해 총 3명의 피험자가 선반과 서랍을 여는 동작을 각각 추출



(a) 서랍을 여는 동작

(b) 선반을 여는 동작

모션 데이터 학습 및 생성(1/2)

■ cRBM에서의 실수값 처리

- 기존의 RBM은 Stochastic binary units만을 처리
- 실수값의 Gaussian visible units를 사용할 수 있도록 함으로써,
각 관절의 회전각이 실수값으로 이루어진 모션캡처 데이터를 학습시키는 것이 가능

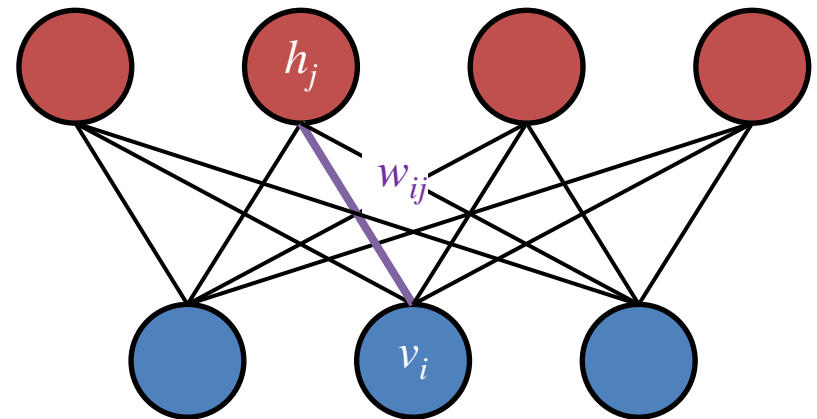
$$p(v, h) = \frac{\exp(-E(v, h))}{Z}$$

$$E(v, h) = \sum_i \frac{(v_i - a_i)^2}{2\sigma_i^2} - \sum_{ij} W_{ij} \frac{v_i}{\sigma_i} h_j - \sum_j b_j h_j$$

$$\Delta W_{ij} \propto \langle v_i h_j \rangle_{data} - \langle v_i h_j \rangle_{model}$$

$$p(v_i | h) = N\left(a_i + \sigma_i \sum_j W_{ij} h_j, \sigma_i^2\right)$$

$$p(h_j = 1 | v) = \frac{1}{1 + \exp\left(-b_j - \sum_i W_{ij} \frac{v_i}{\sigma_i}\right)}$$



모션 데이터 학습 및 생성(2/2)

■ cRBM에서의 모션 학습 및 생성

- 이전 visible variable들과의 추가적인 directed connections로부터 가중치를 갱신함으로써
모션 데이터의 temporal dependency를 학습

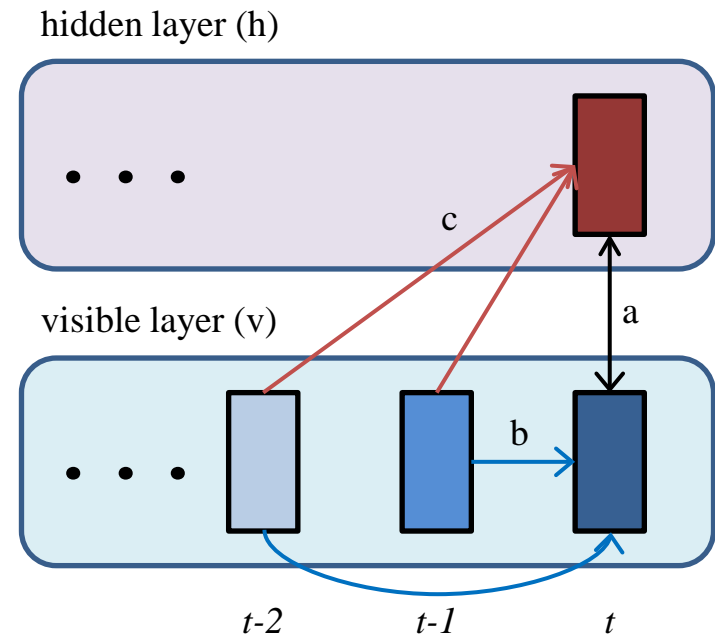
$$p(v_{i,t} | h_t, v_{<t}) = N\left(\hat{a}_{i,t} + \sum_j W_{ij} h_{j,t}, 1\right)$$

$$p(h_{j,t} = 1 | v_t, v_{<t}) = \frac{1}{1 + \exp(-\hat{b}_{j,t} - \sum_i W_{ij} v_{i,t})}$$

$$\begin{aligned} \hat{a}_{i,t} &= a_i + \sum_k A_{ki} v_{k,<t} & \hat{b}_{j,t} &= b_j + \sum_k B_{kj} v_{k,<t} \end{aligned}$$

■ Energy function:

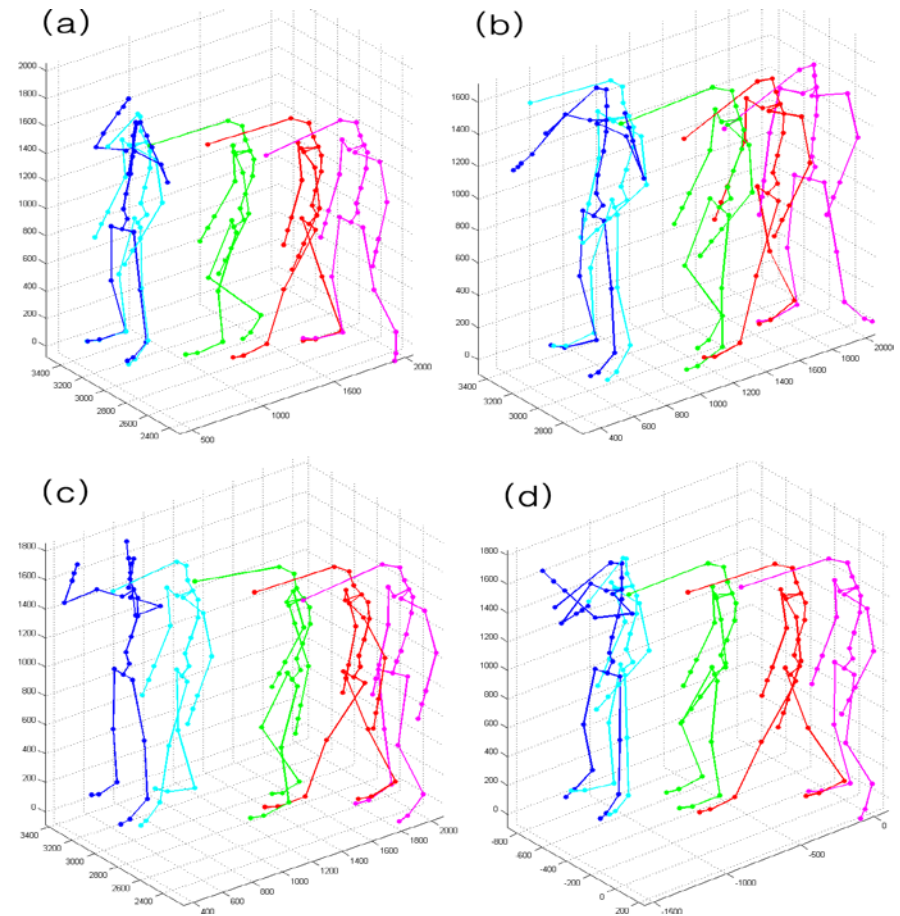
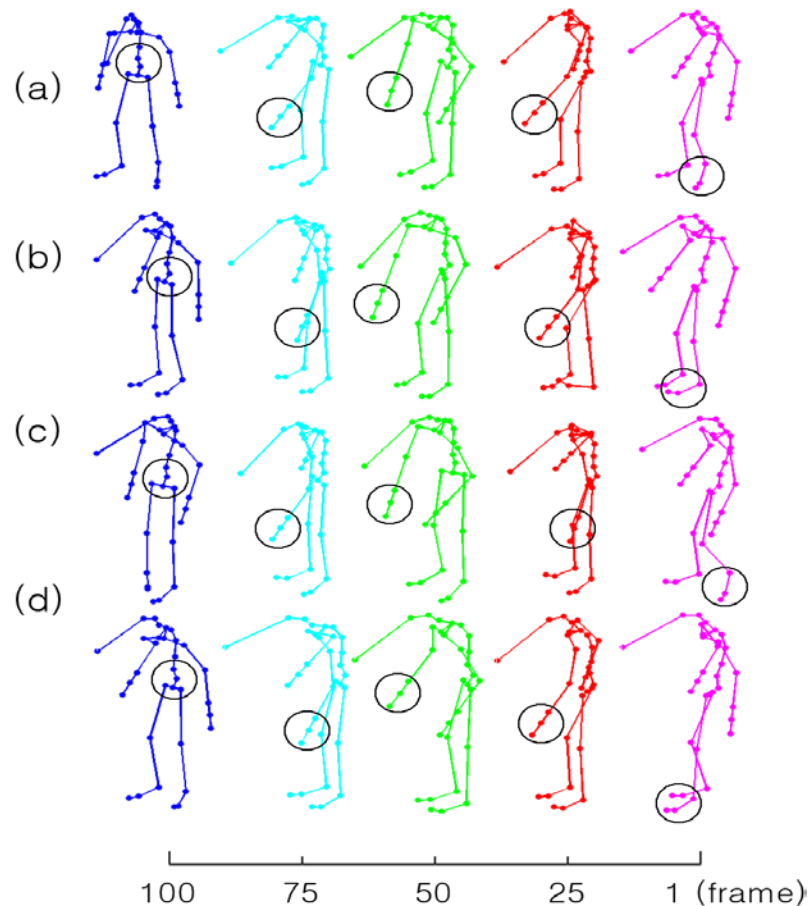
$$E(v_t, h_t | v_{<t}) = \frac{1}{2} \sum_i (v_{i,t} - \hat{a}_{i,t})^2 - \sum_{ij} W_{ij} v_{i,t} h_{j,t} - \sum_i \hat{b}_{i,t} h_{i,t}$$



* G. W. Taylor, G. E. Hinton and S. Roweis, "Modeling human motion using binary latent variables," NIPS, Vol. 19, 2006.

실험 결과(1/4)

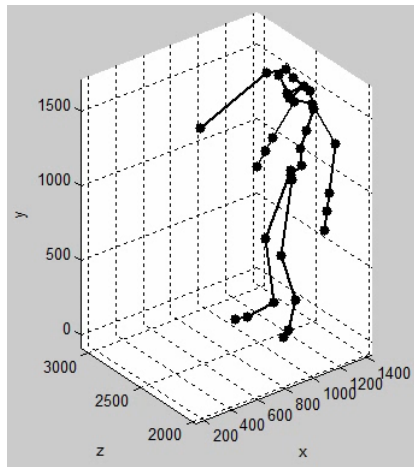
■ 모션 데이터 생성 결과



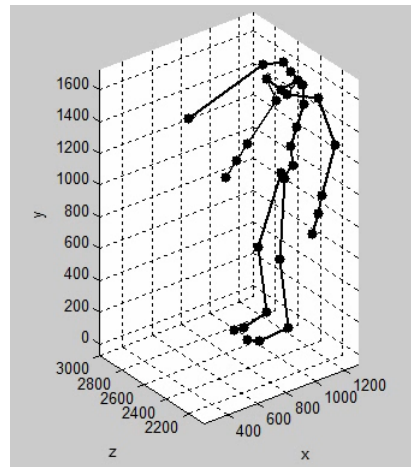
실험 결과(2/4)

■ 모션 데이터 생성 결과 (모션A: 서랍을 여는 동작)

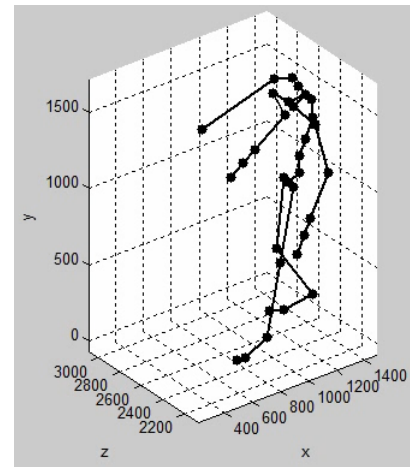
Train 1



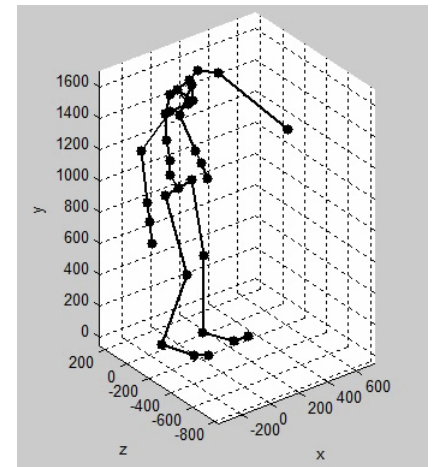
Train 2



Train 3



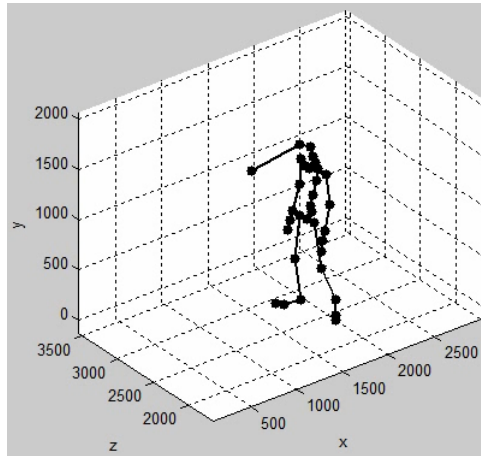
Generated



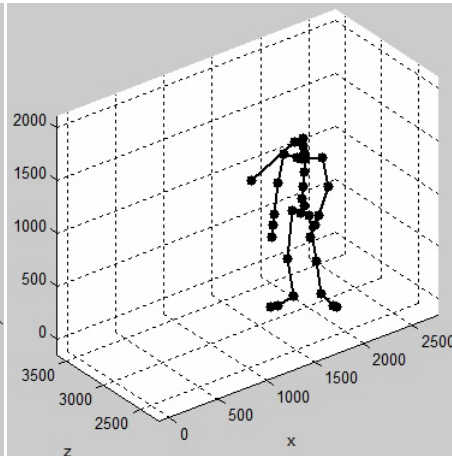
실험 결과(3/4)

- 모션 데이터 생성 결과 (모션B: 찬장에 있는 식기를 꺼내 테이블에 올려두는 동작)

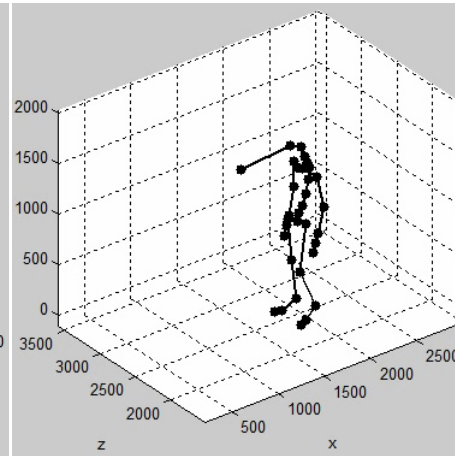
Train 1



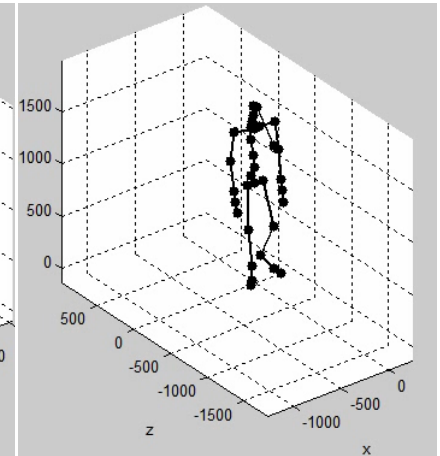
Train 2



Train 3

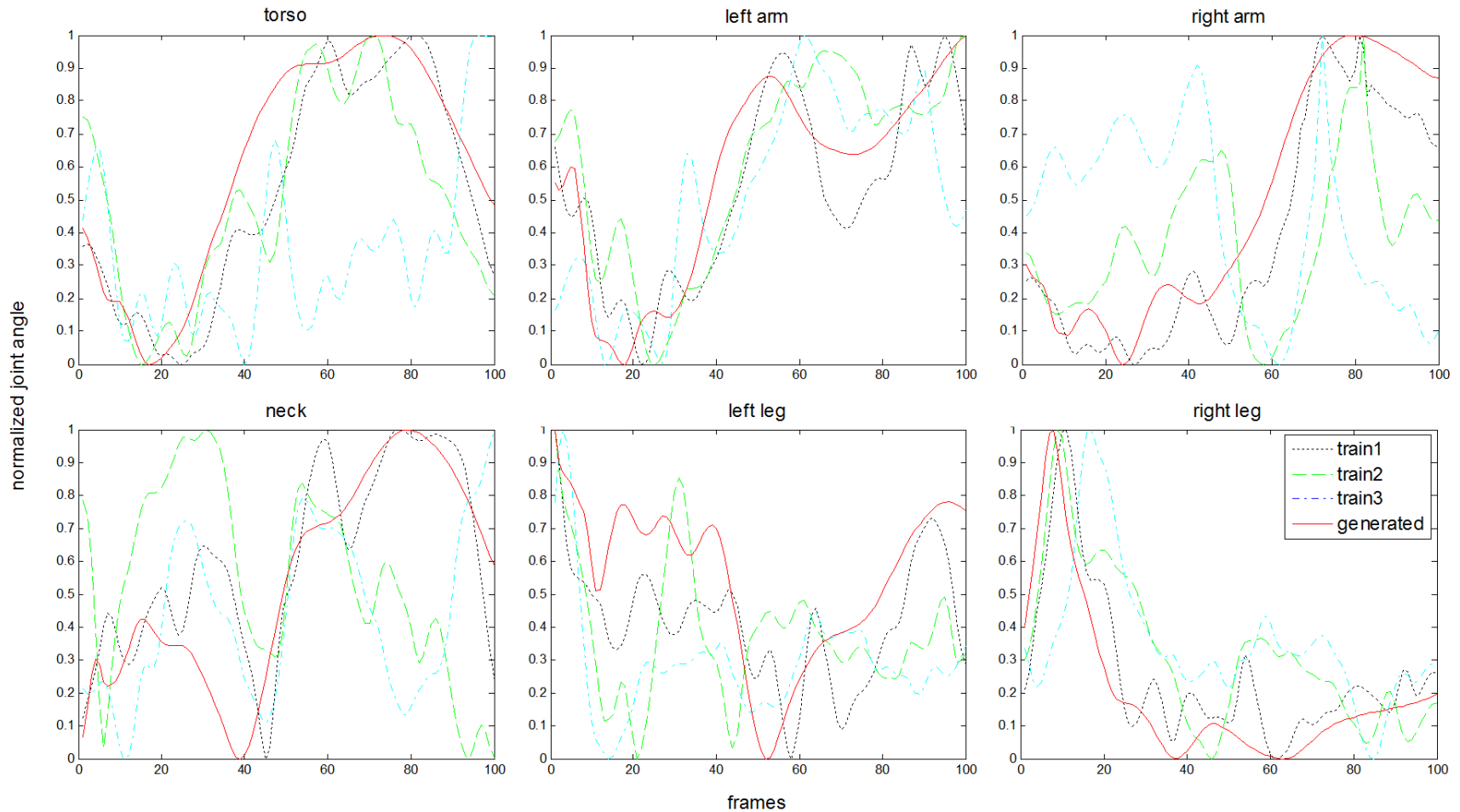


Generated



실험 결과(4/4)

■ 신체부위별 회전각 변화



결론 및 향후연구

■ 결론

- Markerless 모션 캡처 시스템을 이용하여 주방 공간에서의 인간 행동 데이터를 수집
- 모션 학습 및 생성에 적합한 nonlinear generative model인 cRBM을 이용하여 모션 학습
 - 원본 모션들과 동일한 작업을 수행하는 새로운 모션을 자동으로 생성
 - 생성된 모션 데이터를 신체 부위별로 분석하여 학습 데이터와의 유사성을 파악
- 실험 결과는 로봇이나 아바타가 **인간의 모션 데이터를 활용하여 점차적으로 인간 수준의 유연한 움직임을 학습**해나갈 수 있음을 시사함

■ 향후 연구

- 비디오 데이터를 이용한 기계학습 기반의 모션 학습 및 생성 기법을 통해 **상황에 따른 정확하고 자연스러운 로봇 모션 제어**
- 생성한 모션 정보를 prior knowledge로 사용하여 **모션 추적 및 인식 성능을 고도화**

감사합니다!

* 본 연구는 교육과학기술부의 재원으로 한국연구재단의 지원을 받아 수행된 연구(0421-20110032, 지능형 추천 서비스를 위한 인지기반 기계학습 및 추론기술, Videome), (2010-0018950, 뇌정보처리 기반 사용자 의도 변화 모델링 및 예측 기술 개발), (2010-0018950, 로봇팔의 유연한 모션 생성을 위한 기계학습 연구)이며, 교육과학기술부의 BK21-IT사업에 의해 일부 지원되었습니다.