

스타랩 세미나 발표

Camera based Motion Related Works

임양규 | lim0386@gmail.com





Single Camera Pointing Gesture Recognition Using Spatial Feature and Support Vector Machines

2007

Z. Černekova ´, N. Nikolaidis and I. Pitas

Department of Informatics, Aristotle University of Thessaloniki

특이점

연구의 특이점



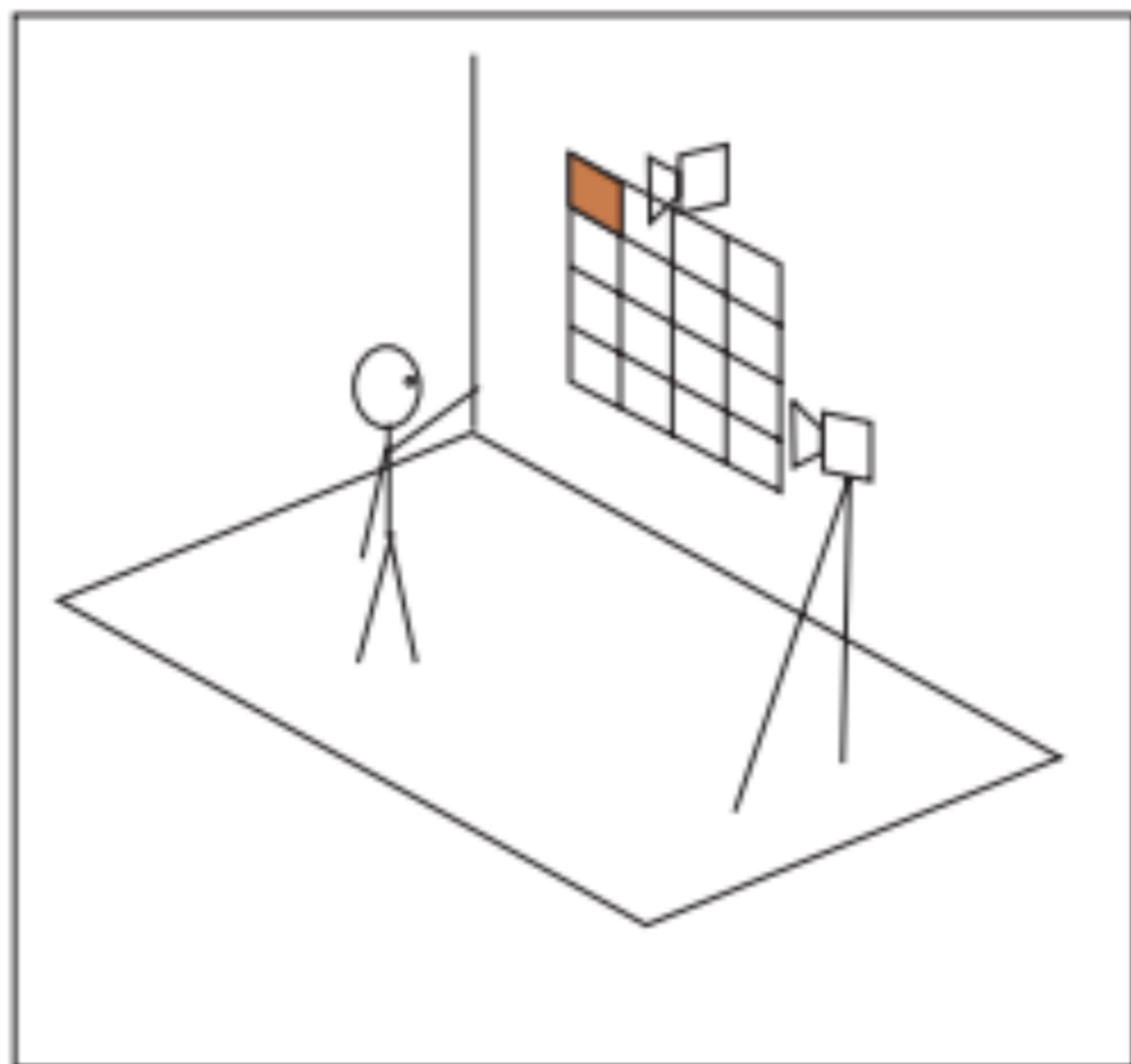
using Single Camera



SVM (Support Vector Machine)



Pointing Gesture



- 
- A man with a beard and a woman are seen from behind, looking at a wall covered in various papers, diagrams, and a large red letter 'Z'. The man is wearing a maroon shirt and a backpack, while the woman is wearing a green shirt. The wall has several orange labels and a small red typewriter illustration at the bottom left.
- 배경(노이즈) 필터링
 - GVF (Gradient Vector Flow) 윤곽선 따기



Camera 좌측방

1. 초기윤곽잡기 원, 네모 등등
2. 대략적인 윤곽선 결과
3. 윤곽선 도출

Comparison of the results between the distance potential snake and the GVF snake. Based on "Snakes, shapes, and gradient vector flow," by Xu et al., 1998.



Finger Detection for Sign Language Recognition

Ravikiran J, Kavi Mahesh, Suhas Mahishi, Dheeraj R, Sudheender S, Nitin V Pujari
Proceedings of the International MultiConference of Engineers and Computer Scientists 2009 Vol I




Camera 정면 상단

- 손가락의 모양
- 얼굴의 방향(시선?)

$$\overline{\mathbf{v}}_i^2 = [\mathbf{x}_h^i, \mathbf{x}_{ft}^i]$$

[머리상단, 손가락끝]

- 
- A man with a beard and a woman are seen from behind, looking at a wall covered in various papers, diagrams, and notes. The man is wearing a maroon shirt and a backpack, while the woman has long red hair and is wearing a green top. The wall features a large red letter 'Z', a small red typewriter illustration, and several orange labels with text like 'Team C' and 'Team D'.
- SVM (Support Vector Machine)
 - 패턴인식, 자료분석을 위한 지도학습모델
 - 비확률적 이진 선형 분류 모델 생성

Run	l-o-o	accuracy
1.	v_1	70.7 %
2.	v_2	79.1 %
3.	v_3	68.2 %
4.	m_1	72.6 %
5.	m_2	69.3 %

Run	l-o-o	accuracy
1.	z_1	98.2 %
2.	z_2	94.4 %
3.	z_3	95.6 %
4.	s_1	91.3 %
5.	s_2	92.6 %



Real-time American Sign Language Recognition with Convolutional Neural Networks

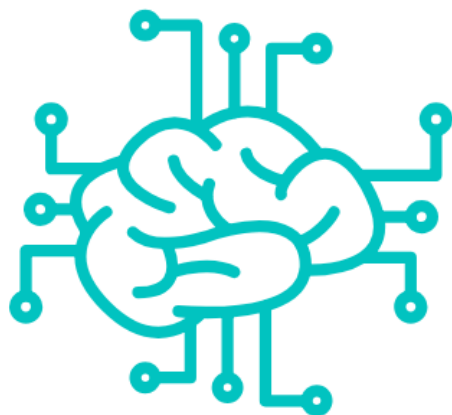
Brandon Garcia, Sigberto Alarcon Viesca
Stanford University

요약

연구의 특이점



ASL (American Sign Language)



CNN (Convolutional Neural Network)



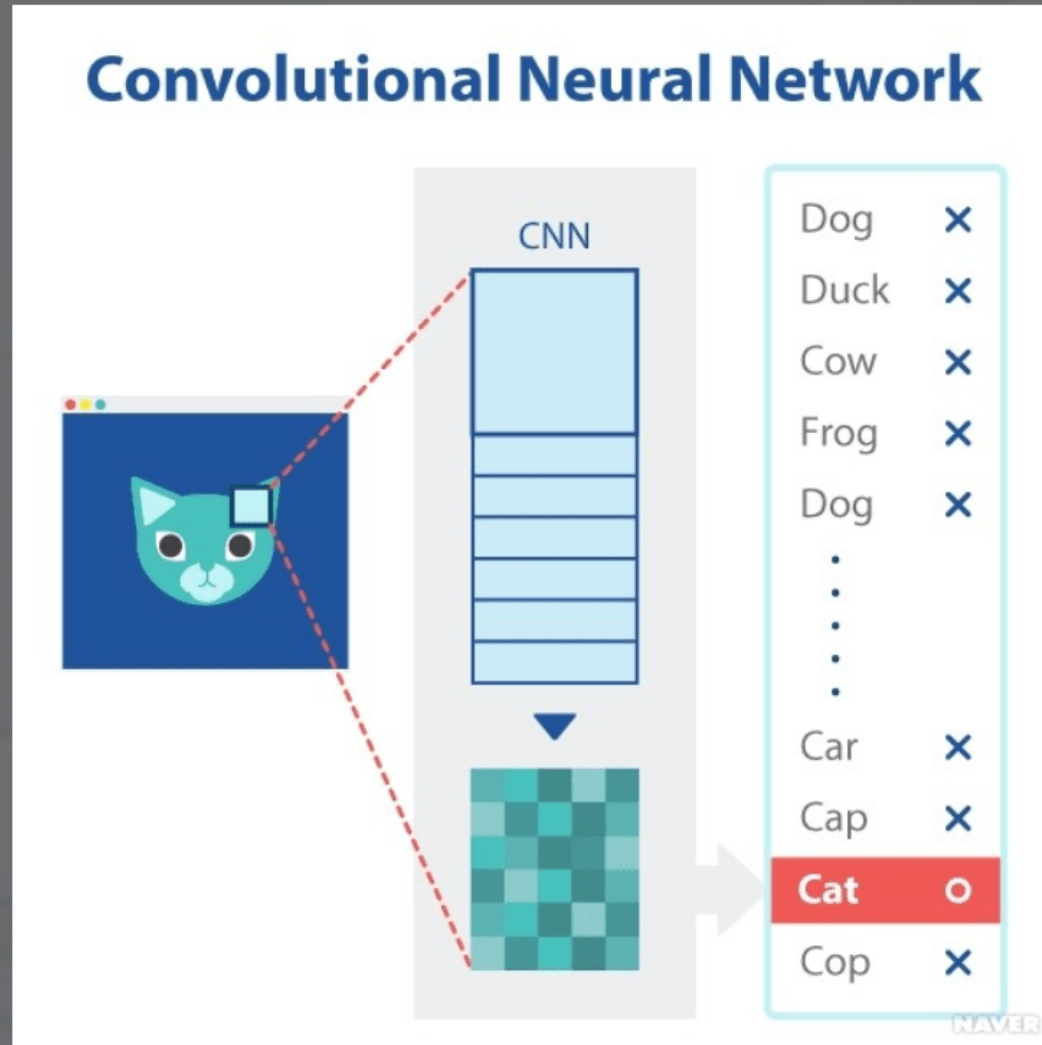
Text (Alphabet)

CNN: Convolutional Neural Network

- 심층 신경망 (Deep Neural Network)의 한 종류
- 하나 또는 여러 개의 컨볼루션 계층(convolutional layer)과 통합 계층(pooling layer), 완전하게 연결된 계층(fully connected layer)들로 구성된 신경망
- Backpropagation Algorithm을 통해 훈련
- 객체 분류, 탐지에 활용되는 DNN 모델

*컨볼루션 $y(t) = \int x(\tau)h(t - \tau) d\tau$ 신호 $h(t)$

CNN: Convolutional Neural Network



- 본인들의 데이터: 24개의 방향에서 촬영된 이미지 부족
- ILSVRC (large Scale Visual Recognition Challenge)2012 + Berkeley Vision + GoogLeNet
- 데이터 세트 별로 이미지 크기 다른 것을 맞추는 작업

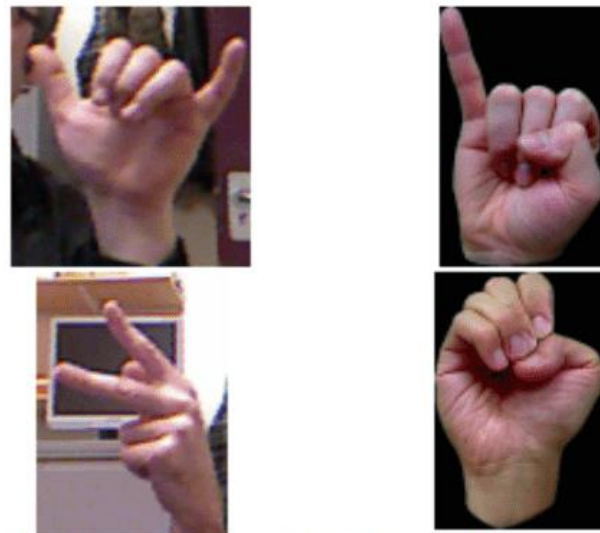

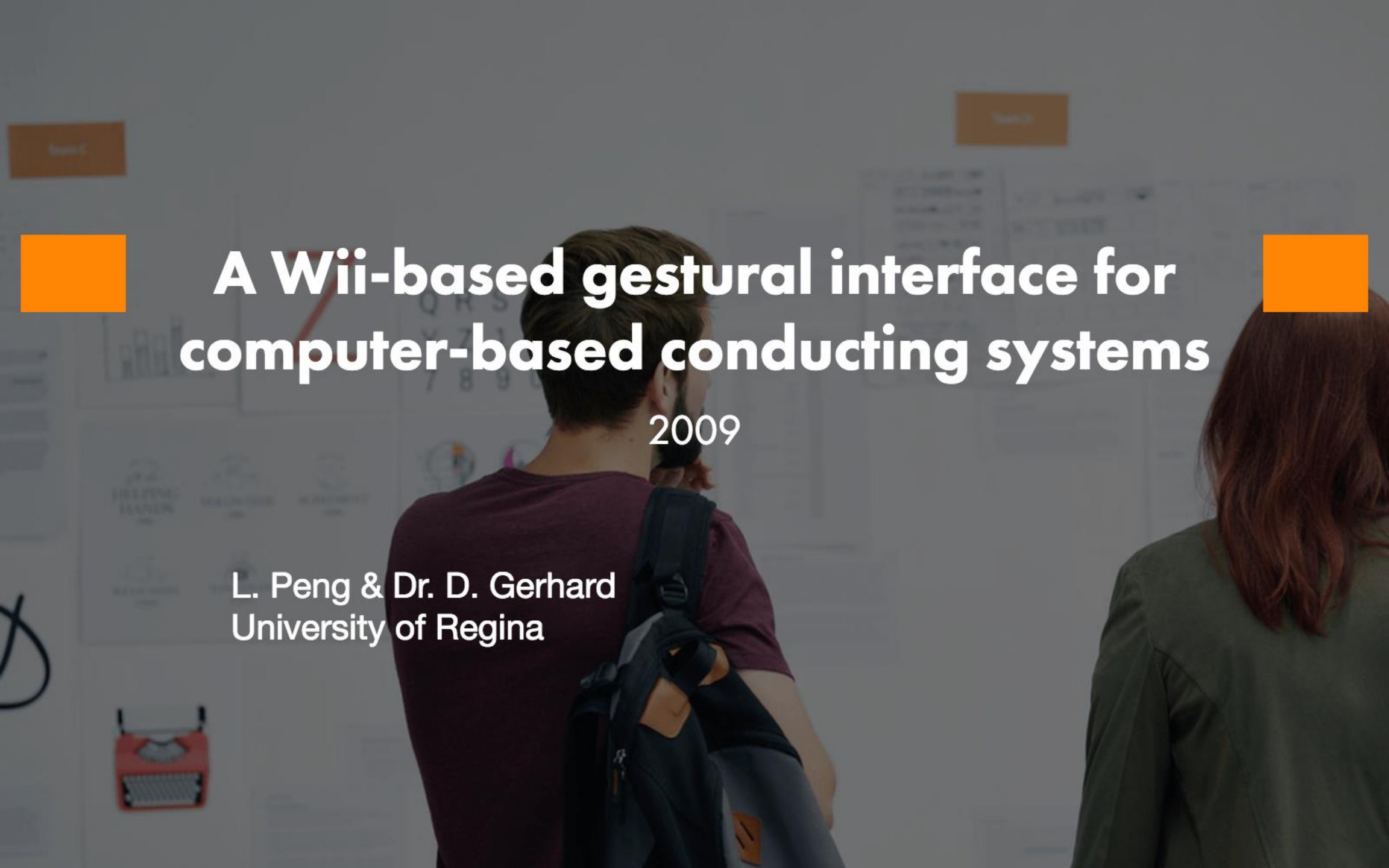


Fig. 1. Dataset examples. Left: Surrey University. Right: Massey University. Top left: *y*. Bottom left: *k*. Top right: *i*. Bottom right: *e*.

- 
- 본인들의 데이터: 24개의 방향에서 촬영된 이미지 부족
 - ILSVRC (large Scale Visual Recognition Challenge)2012 + Berkeley Vision + GoogLeNet
 - 데이터 세트 별로 이미지 크기 다른 것을 맞추는 작업

1. 레이어 1 재초기화 및 학습속도 증대
2. 레이어 2 재초기화 및 학습속도 증대
3. 레이어 1 재초기화 학습속도 증대 및 배치 크기 확대
4. 레이어1 재초기화, 균일 학습

Model	Top-1 Val Accuracy	Top-5 Val Accuracy
a - y [1_init]	0.6847	0.9163
a - y [2_init]	0.6585	0.9043
a - y [batch]	0.6965	0.9076
a - y [full_lr]	0.7200	0.9098
a - k [2_init]	0.7430	0.897
a - e [2_init]	0.9782	1.000



A Wii-based gestural interface for computer-based conducting systems

2009

L. Peng & Dr. D. Gerhard
University of Regina

요약

연구의 특이점



wii모트의 센서를 사용한 지휘 동작 검출



wii모트의 카메라를 사용함 (적외선)



2, 3, 4 박자의 패턴을 인식 시킴

문제점

연구로서의 문제점 (임양규 관점)



음악 지식의 부족


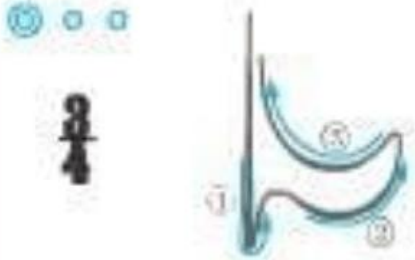
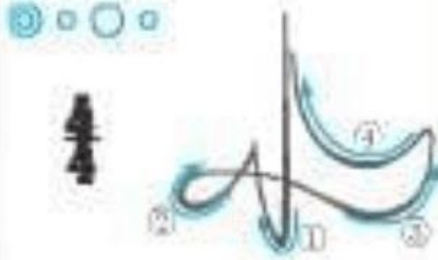
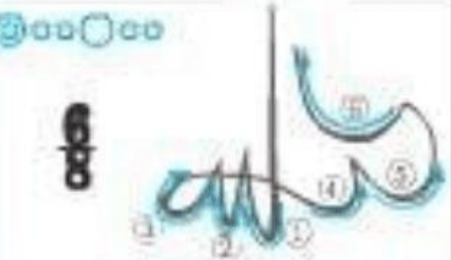

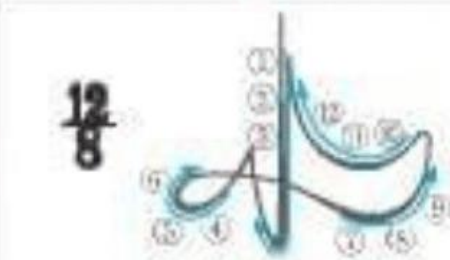


정형화 된 지휘 동작은 없음



결론: 음악과 관련된 사항이 아닐 수도 있음

박자와 지휘법

박자의 종류	지휘법		
홀박자			
짝박자			

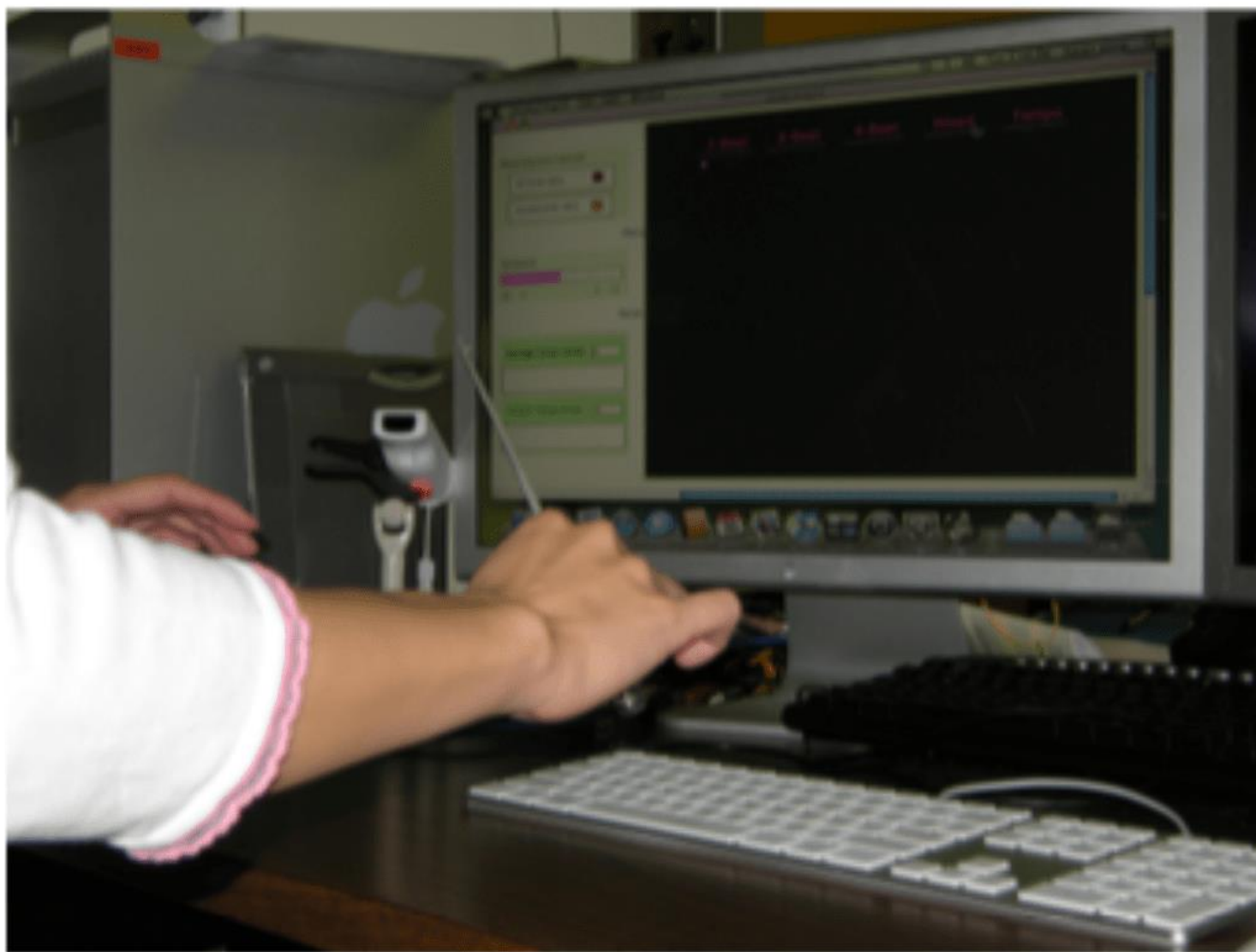
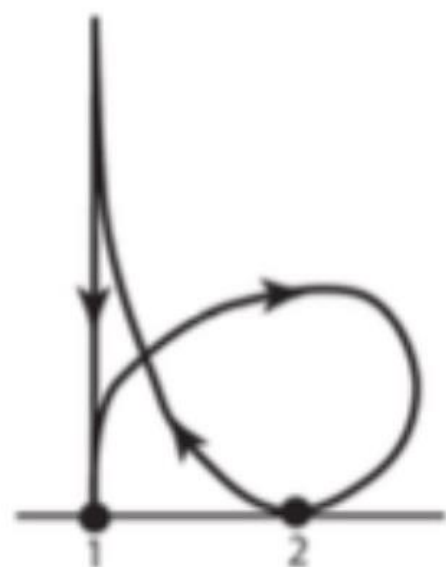
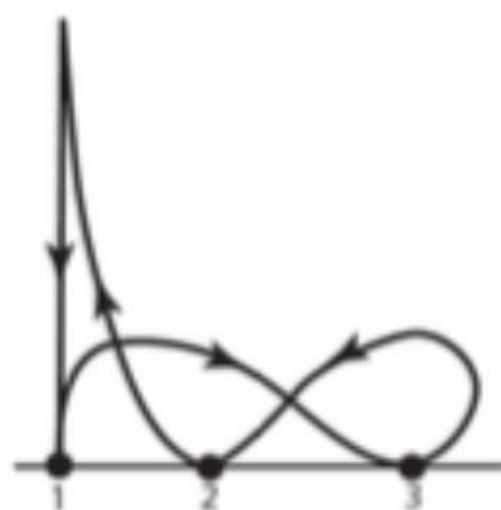


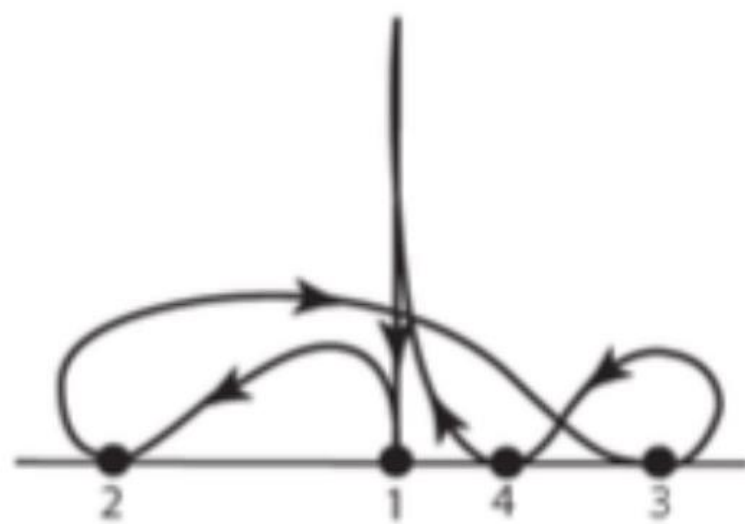
Figure 4. The setup



(a) 2-beat pattern



(b) 3-beat pattern



(c) 4-beat pattern

Figure 5. Three beat patterns

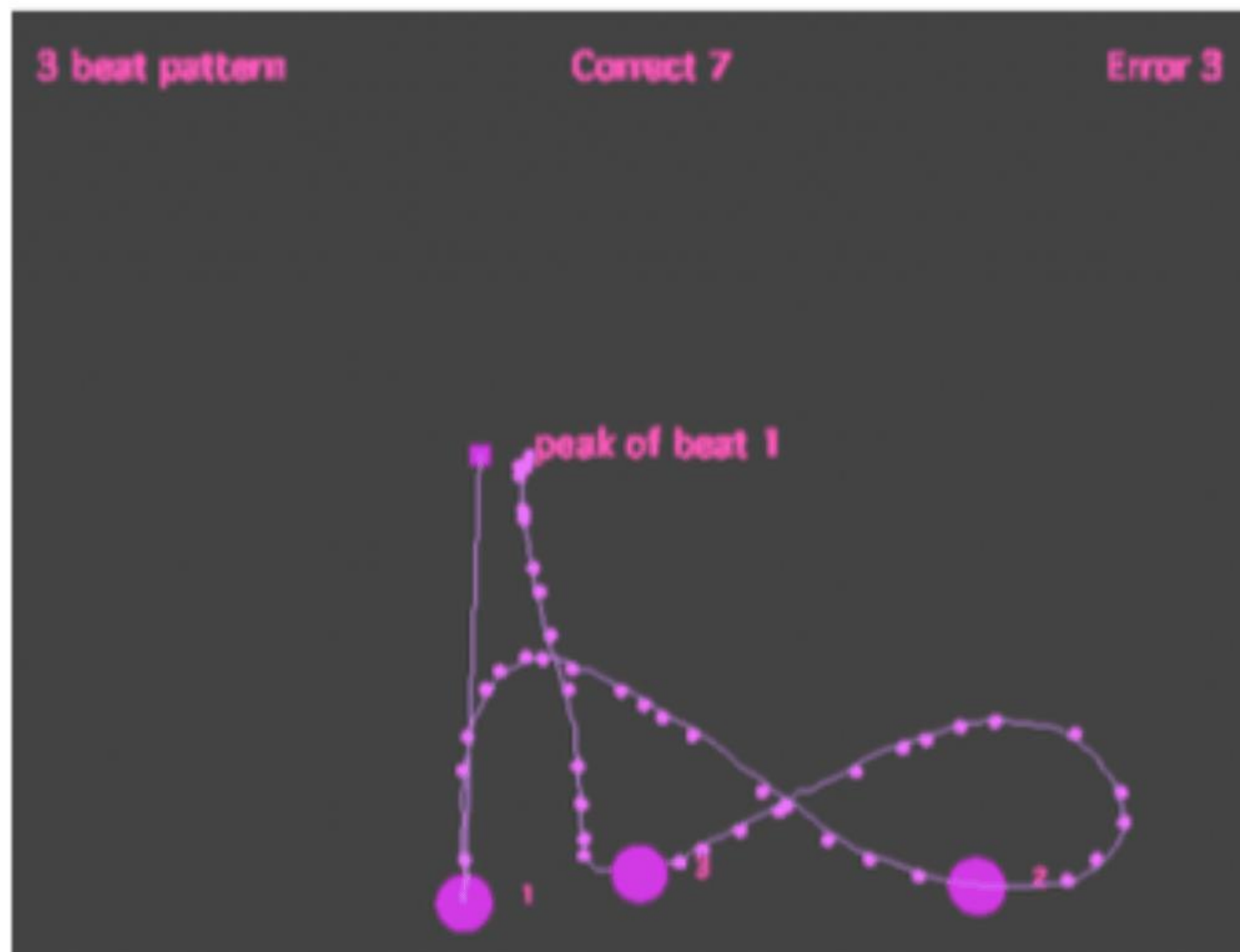


Figure 10. A beat pattern performed by the right hand

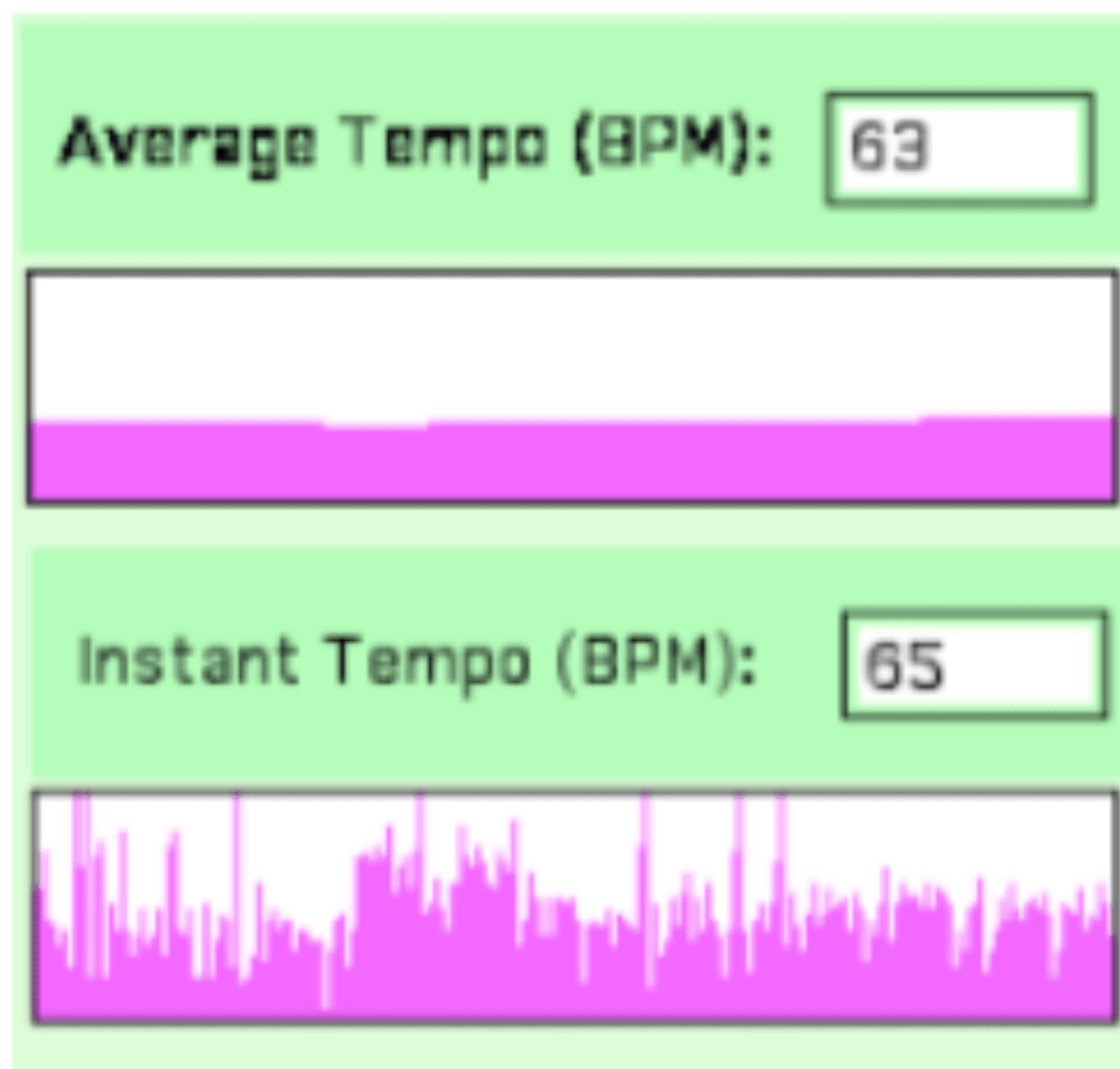


Figure 11. Visual representation of the value of tempo

요약

- 비록 지휘 동작은 정형화 되어 있지 않지만, 특정 동작을 똑같이 모사하는 것
- 동작의 모사 유사도 (동작의 움직임 위치, 시간 등등)에 의해 bpm속도를 측정
- tracking같은 실시간 지점 추적 같은 방식이긴 하나 특정 지점간의 연결점이 합쳐져 완성된 동작이 나올때 그것을 인식하고 음악의 속도에 반영하는 결과