Team name	YNL
Team leader name	Yusuke GOUTSU
Team leader address, phone number and email	7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan goutsu@ynl.t.u-tokyo.ac.jp
Rest of team members	Wataru TAKANO Yoshihiko NAKAMURA
Team website URL (if any)	http://www.ynl.t.u-tokyo.ac.jp/index.html

Title of the contribution	Gesture Recognition using Hidden Markov Models with Fisher Vector (FV-HMM)
General method description	We proposed the hybrid generative-discriminative approach. The idea is to merge the description abilities of Hidden Markov Model (HMM), which is a generative models in general, with the discriminative skills of Support Vector Machine (SVM) with Fisher Vector (FV) that directly model the posterior probability and focus on the class separability. Fisher Kernel (FK) was proposed as a general way of mixing generative and discriminative models for classification. In this process, a generative model is used to define feature vectors and gestures are projected to the resulting feature space. The FK approach measures the relation between the gestures by comparing them in the feature space induced by HMM, which is considered as a point in the Riemannian manifold. This space have a benefit to measure geodesic distances between points along the manifold. In practice, each gesture is represented by a FV, whose components are called Fisher Scores (FSs), defined by derivatives of the log-likelihood of the generative model with respect to all parameters. A kernel can be defined in various ways in the resulting space and the inner product is used in FK. Fisher information matrix derived from the FK was used to normalize the FV.
References	M. Bicego, M. Cristani, V. Murino, E. Pekalska, and R.P.W. Duin: "Clustering-based construction of Hidden Markov Models for Generative Kernels", Energy Minimization Methods in Computer Vision and Pattern Recognition, pp.466-479, 2009.

Describe data preprocessing techniques applied (if any)	• **Normalization of skeleton information** • **Inverse kinematics using marker position data captured by Kinect**
Describe features used or data representation model (if any)	HMM with FV
Data modalities used, i.e. depth, rgb, skeleton... (if any)	Skeleton
Fusion strategy applied (if any)	
Dimensionality reduction technique applied (if any)	

Temporal clustering approach (if any)	Ward's method (hierarchical clustering) for HMMs
Temporal segmentation approach (if any)	The moment when hand marker velocity exceeds a threshold is a segmentation point.
Gesture representation approach (if any)	HMM
Classifier used (if any)	SVM
Large scale strategy (if any)	

Transfer learning strategy (if any)	
Temporal coherence and/or tracking approach considered (if any)	
Other technique/strategy used not included in previous items (if any)	
Method complexity analysis	

Qualitative advantages of the proposed solution	**It is already a proven point that fisher vector is useful method for object recognition in computer vision. We applyed this method to gesture recognition.**
Results of the comparison to other approaches (if any)	
Novelty degree of the solution and if is has been previously published	・There is almost no research of combination of HMM and FV. We also use this method in application strategy such as gesture recognition. ・Not yet published

Language and implementation details (including platform, memory, parallelization requirements)	**· Microsoft Visual C++ 2010 (x64) for Windows** **· Dell Precision T7500 (CPU:2.40GHz, Memory:48GB, OS:64bit)** **· More than 30GB of memory is required.**
Human effort required for implementation, training and validation?	Not at all.
Training/testing expended time?	· 6 and 12 hours for training HMM and SVM respectively over about 7000 gestures · 15 minutes for testing over about 3500 gestures
General comments and impressions of the challenge	· Thanks for organizing this interesting multi-modal gesture competition. I hope that this competition will be held next year and the research of this field becomes more and more popular!