

Representation and Symbolization of Motion Captured Human Action by Locality Preserving Projections

Sang Ryong Lee, Geun Sub Heo and Choon-Young Lee*

School of Mechanical Engineering, Kyungpook National University, 1370 Sangyeok-dong, Buk-gu, Daegu, 702-701, Republic of Korea

Received: 15 Jul. 2013, Revised: 28 Nov. 2013, Accepted: 29 Nov. 2013

Published online: 1 Jan. 2014

Abstract: Human motion analysis and assessment are important in determining Parkinsons disease and stroke, or in measuring skill quality in basic motions. Reduced space is useful in representing motion segments and finding basic behavioral patterns for humanoid robot control using the modularized approach. In the current paper, we represent motion-captured data of human action in a reduced space of nonlinear degrees of freedom in which the original motion is characterized. First, we represent high-dimensional data, such as motion sequence of the position of joints in Cartesian space, in a reduced space using the locality preserving projection (LPP) method. Second, we find a similarity measure between the actions. Finally, we assess human motions using a similarity measure to find the most similar one. The LPP is a linear dimensionality reduction algorithm that builds a graph for neighborhood information and maps data points to a reduced space. The reason for using LPP in our study is that it is defined globally, and any new data element can be mapped in the reduced space. Our method includes the generation of symbolic code sequence corresponding to complex, high-dimensional motion. Interdisciplinary synergy combined with information technology and wearable sensor systems can broaden the possible future applications in rehabilitation engineering.

Keywords: Motion Analysis, Locality Preserving Projection (LPP), Motion Capture Data, Dimension Reduction

1 Introduction

Dimension reduction creates a meaningful representation within low-dimensional space for high-dimensional data, such as large pixel data of an image scene or the human bodys joint data for a specific task [1][2]. Local metric information and optimization technique can determine the underlying manifold of a data set. The intrinsic dimension of a data set is obtained by empirical deduction or by maximum likelihood estimation. Sanger studied human arm movements and found that low dimensionality might be an efficient way to describe a large class of arm movements [3]. Allen defined a low-dimensional subspace called eigengrasps to approximate the hand motion required for a given grasping task and designed control algorithms to operate in the reduced dimension [4]. Explicit representations for dynamic shape manifold were studied on moving humans to recognize actions using low-dimensional embedding by locality preserving projections (LPP) [5]. Mao proposed the concept of kinematic synergies to address the dimensionality reduction problem for hand movement control and coordination [6]. Optimization technique was applied to

find the weights and firing timings of extracted synergies, decomposing basic motions. Low-dimensional embeddings were used to represent muscle activations and motion primitives from electromyographic signals, which were recorded by upper limb motion [7]. He and Niyogi introduced a new linear dimensionality reduction algorithm called LPP, which has a transformation matrix that maps data points into a subspace [8]. The LPP has several interesting features, such as locality preservation with objective functions to find linear approximation rather than nonlinear manifold similar to Locally Linear Embedding. Moreover, input space is globally defined, and LPP can locate new data in the reduced dimension. Statistically, LPP is linear mapping that preserves the structure of the underlying distribution best in the L^2 sense. Human motions are sets of time-trajectories for body joints. The complexity of spatiotemporal motion trajectories makes using dimensional reduction methods necessary for further analysis and assessment of the underlying information. Human motion assessment in gait rehabilitation can be found in literature [9]. The representation of movement primitives for the joint

* Corresponding author e-mail: cylee@knu.ac.kr

trajectories was conducted via a principal component analysis of motion captured data. The method is computationally efficient by representing an arbitrary motion as a linearly combined basis functions [18]. To derive sets of human motion primitive as atomic-level and meta-level behavior primitives, a spatio-temporal intrinsic dimension reduction technique was applied using a manually segmented human motion capture data [19]. The extraction of trajectories through sensor systems has been well studied by a number of researchers. However, investigations and further analyses on the comparison of given trajectories are rare [10]. In the field of surveillance systems, research has focused mainly on the classification of motion trajectories generally represented in two-dimensional (2D) [11]. Using the obtained trajectories, path modeling can recognize activity or detect anomaly. Numerous methods can measure similarity between trajectories in low-dimensional space. The simplest method is to use the Euclidean distance between two trajectories having the same length. Other advanced measures for trajectory distance include spatial similarity by Hausdorff distance [12], fitting by Hidden Markov Model (HMM), and the degree of alignment between two trajectory sequences by Longest Common Subsequence distance or Dynamic Time Warping. Many algorithms on analyzing trajectories on the 2D or three-dimensional plane have been reported. The current paper uses low-dimensional embeddings of high-dimensional human motions to obtain meaningful comparisons between two time trajectories of the motion to be assessed. The approach focuses on the representation and abstraction of highly complex joint motions in the reduced space using motion capture data by CMU Graphics Lab Motion Capture Database [13]. This paper is organized as follows. One of useful nonlinear dimension reduction algorithm, LPP, will be described in the next section. Using the low-dimensional mapping, we describe the symbolization process to represent the original high-dimensional data as a sequence of symbols. The experimental results on the comparison of human actions in daily life will be given in the following section with discussions on the popular trajectory analysis methods. We finally conclude our works with further research directions.

2 Low Dimensional Representation and Abstraction

2.1 Dimension Reduction by LPP

The generic problem of low-dimensional embedding is as follows. Assume a data set of m -elements exists in n -dimensional space, which can be written as

$$\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m \in \mathbb{R}^n \quad (1)$$

and we want to find a transformation matrix $\mathbf{W} \in \mathbb{R}^{n \times r}$ connecting the data points to a new set of points in the reduced dimension $r \leq n$, where

$$\mathbf{y}_i = \mathbf{W}^T \mathbf{x}_i \quad (2)$$

The LPP is a linear approximation of the Laplacian Eigenmap [8]. First, using ε -neighborhood in the Euclidean distance sense or k -nearest neighborhood, we construct the adjacency graph. In this graph, nodes i and j are connected by an edge if \mathbf{x}_i and \mathbf{x}_j are close neighbors. Second, we determine sparse symmetric $m \times m$ weight matrix \mathbf{S} , with S_{ij} representing the weight connecting nodes i and j . The kernel of radial basis function can be used in \mathbf{S} as follows:

$$S_{ij} = e^{-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\sigma}} \quad (3)$$

where σ is a constant. We then compute the generalized eigenvector problem to determine the eigenvectors and eigenvalues using

$$\mathbf{X} \mathbf{L} \mathbf{X}^T \mathbf{w} = \lambda \mathbf{X} \mathbf{D} \mathbf{X}^T \mathbf{w} \quad (4)$$

where \mathbf{D} is a diagonal matrix with elements in column-wise (or row-wise) sums of \mathbf{S} , i.e., $D_{ii} = \sum_j S_{ji}$. The Laplacian matrix is defined as $\mathbf{L} = \mathbf{D} - \mathbf{S}$. The symbol \mathbf{X} is a matrix the j -th column of which is \mathbf{x}_j . The column vectors $\mathbf{w}_0, \mathbf{w}_1, \dots, \mathbf{w}_{r-1}$ are the solutions of Eq. (4) ordered as corresponding eigenvalues in a non-decreasing manner, which can be denoted as $0 \leq \lambda_0 \leq \lambda_1 \leq \dots \leq \lambda_{r-1}$. The cost function to be minimized for the solution is the same as the following form with similarity matrix:

$$\sum_{i,j} (\mathbf{y}_i - \mathbf{y}_j)^2 S_{ij} \quad (5)$$

The S_{ij} is used to preserve locality, which attempts to ensure that if two points are in one neighborhood, then the resulting points in the reduced dimension are also close. The optimization of the Eq. (5) can be rewritten as

$$\frac{1}{2} \sum_{i,j} (\mathbf{y}_i - \mathbf{y}_j)^2 S_{ij} = \frac{1}{2} \sum_{i,j} (\mathbf{w}^T \mathbf{x}_i - \mathbf{w}^T \mathbf{x}_j)^2 S_{ij},$$

which is solved by standard graph theory to find the eigenvectors by imposing a constraint, like $\mathbf{y}^T \mathbf{D} \mathbf{y} = 1$.

The LPP is derived by preserving local information and is thus less sensitive to noises like outliers. LPP also has more discriminating ability than the conventional Principal Component Analysis since it maintains the embedded structure of training samples unlike the PCA.

2.2 Abstraction by Symbolization

Let the high-dimensional trajectory data be represented on the r -dimensional space. We rewrite the vector-valued

sequence at instant k as the combined sequence by a weighted Euclidean norm as

$$\mathbf{y}_k = [y_1(k) \ y_2(k) \ \dots \ y_r(k)]^T \quad (6)$$

$$y_{res}(k) = \sqrt{\mathbf{y}_k^T \mathbf{A} \mathbf{y}_k} \quad (7)$$

where $\mathbf{A} \in \mathbb{R}^{r \times r}$ is a diagonal matrix whose elements are the weights for each axis of representation. The weights are determined by the user. The magnitude of the combined signal $y_{res}(k)$ at a specific instant k is partitioned into $p_0(k)$ -regions. Each region is assigned a value between zero and $p_0(k) - 1$. If we use the q -step template, we have partitions in a time order as $p_0(k), p_1(k), \dots, p_{q-1}(k)$. The total number of symbols N is

$$N = p_0 \times p_1 \times \dots \times p_{q-1} = \prod_{i=0}^{q-1} p_i \quad (8)$$

We can make the sequence of symbols on the combined signal, i.e., Eq. (7), using the partition template, i.e., Eq. (8). Fig. 1 shows a brief illustration of the symbolization procedure.

2.3 The Size of Reduced Dimension

The question on the size of dimension is a fundamental one and it plays an important role on the accuracy and robustness of the result. The dimension of the reduced space is a key parameter determined by the user or by an estimator. If the dimension is smaller than the intrinsic one, some important information or features can be missed. Otherwise, a too large dimension produces a noise-laden output. No explicit agreement exists to determine the intrinsic dimension. Bickel proposed a new method to estimate the intrinsic dimension of a data set using the principle of maximum likelihood estimation (MLE) on the distances between close neighbors [14]. In the current paper, MLE-based determination of intrinsic dimension was used before the dimension reduction was conducted by LPP. The basic idea to find the size of low-dimensional space is to treat the observations of a fixed point as a homogeneous Poisson process in a small sphere. The MLE-based algorithm requires also the choice of the number of neighbors in the calculation. The negative bias problem also can be considered as a minor if the intrinsic dimension of the real data is small.

3 Experiments

3.1 Representation of Activity in Daily Life

Motion capture data has played a significant role in character animation in the multimedia industry. Kaneko proposed an image-based user interface for retrieving

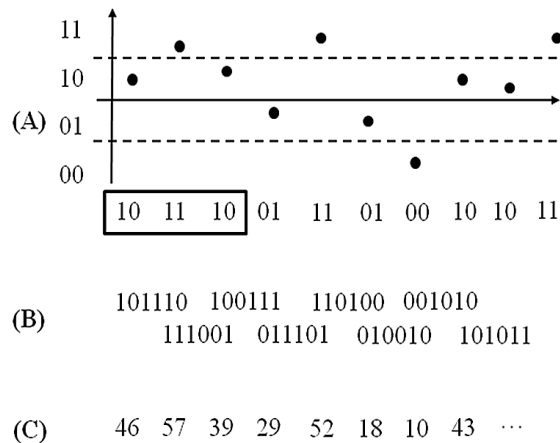


Fig. 1: Process of symbolization: the horizontal axis is related to time; the vertical axis is the value of data points marked as black dots. (A) Partition of data points into four possible codes for each time instant. (B) Using k -step template (e.g., 3 step, in this case), words are formed by moving the template along the time axis. (C) The decimal codes are generated as the final sequence of symbols for the given series.

motion data using a self-organizing map (SOM) called a Motion Map [15]. Li presented an efficient motion data indexing and retrieval method based on SOM and the Smith-Waterman string similarity metric [16]. Motion data were obtained from the mocap database by CMU. The categories are human interaction, interaction with environment, locomotion, physical activity and sports, situation and scenarios, and test motions. The physical activities and sports category includes motions from basketball, dance, gymnastics, acrobatics, martial arts, racquet sports, soccer, boxing, golf, Frisbee, and general exercise and stretching.

The upper body arm motions, including those of the shoulder, elbow, and wrist joints, are characterized by seven degrees of freedom (DOF). Using the kinematic relation of the body structure, the arm posture can also be represented using positions for each joint. MLE estimation of the intrinsic dimension for the upper body arm produces two or three dimensions of the arm postures from the CMU database. As shown in Fig. 2, the physical intuition on the intrinsic dimensions is that arm posture can be characterized by folding/stretching and lowering/holding up. To obtain the linear transformation matrix of Eq. (2), the motion capture data by subject no. 2 in the CMU database were used. The data consist of 10 motion clips for various expressions and human behaviors like walking, running, jogging, punching, bending, lifting, washing, and the combination of similar activities.

Fig. 3 shows the result of mapping eight sample upper arm motions projected on the reduced dimension. The motion sequence of drinking water is mapped as a trajectory consisting of elements coinciding with the

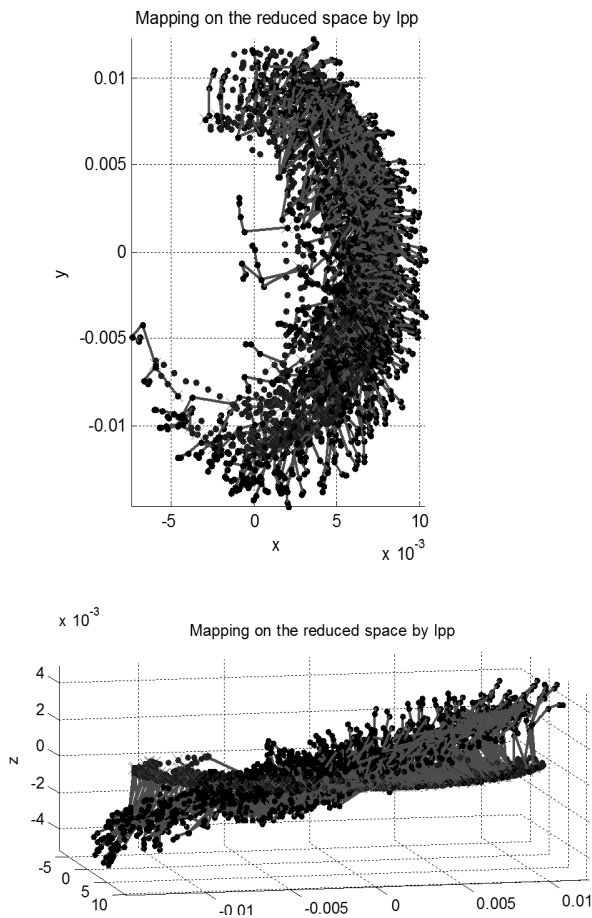


Fig. 2: Upper arm motion is represented on the reduced dimension of two. The intrinsic dimension was determined by MLE, and the dimension reduction technique used was LPP. The dots represent the corresponding arm posture drawn in lines.

training data. However, the sequence of a swimming stroke is represented as a trajectory with a few new elements not used in the training process of obtaining the transformation matrix. That is, new posture can be mapped onto a point, which is an interpolated one between the postures close to it.

These trajectories on the reduced dimension characterize specific activities. Further analysis on the trajectory will provide useful information on the activity, and the characteristic properties, such as the similarity between two limbs, can be assessed. Trajectories in high-dimensional space mapped on the reduced dimension of the intrinsic DOF can be classified and recognized efficiently through abstraction and representation. A similarity check was conducted for each pair of activity motion by applying histogram analysis on the abstracted trajectory, which is a coded sequence. The

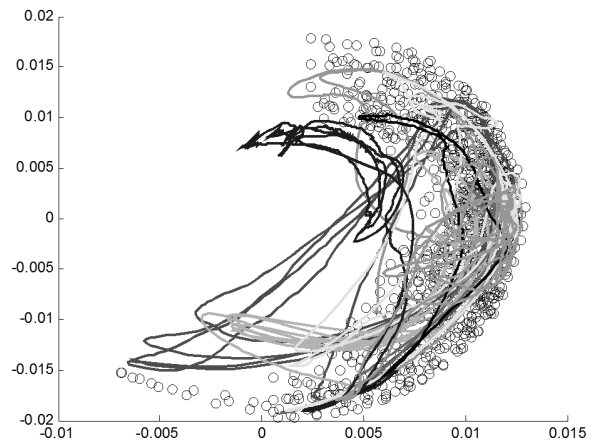


Fig. 3: Typical results of the mapping sample motion sequences on the reduced dimension. The eight motions plotted are chopping wood, swimming, digging, hand shaking, playing drums, drinking water, weight lifting, and rowing.

Table 1: Similarity analysis between the sample motion sequences on the reduced dimension. Eight motions are denoted by alphabet characters as A (chopping wood), B (swimming), C (digging), D (hand shaking), E (playing drums), F (drinking water), G (weight lifting), and H (rowing) [13]. Values are in percent (%).

	A	B	C	D	E	F	G	H
A	-	63.7	66.5	31.6	79.4	64.4	67.6	76.2
B	63.7	-	86.2	58.1	60.0	74.4	66.4	73.2
C	66.5	86.2	-	45.5	61.8	71.7	66.4	75.4
D	31.6	58.1	45.5	-	33.1	63.7	59.1	50.7
E	79.4	60.0	61.8	33.1	-	59.0	81.1	83.6
F	64.4	74.4	71.7	63.7	59.0	-	64.4	72.8
G	67.6	66.4	66.4	59.1	81.1	64.4	-	88.2
H	76.2	73.2	75.4	50.7	83.6	72.8	88.2	-

similarity index is as follows:

$$SI = 1 - \frac{\sum_i |h_A(i) - h_B(i)|}{\sum_i |h_A(i) + h_B(i)|} \quad (9)$$

where h_A , h_B are histograms of the abstracted sequence A and B, respectively. Table 1 shows the result of the similarity analysis between two of the eight motions.

3.2 Discussion

To analyze a time series trajectory represented on the reduced dimension, we assume the following notation:

$$T_a : \{(\mathbf{y}_{a1}, t_1), (\mathbf{y}_{a2}, t_2), \dots, (\mathbf{y}_{aN}, t_N)\} \quad (10)$$

where N is the duration of trajectory. A number of measures can compare two trajectories. The simplest

L -norm distance between two trajectories, T_a and T_b , is

$$D_1(T_a, T_b) = \frac{1}{N} \sum_{j=1}^N d_j^L \quad (11)$$

$$d_j^L = \left(\sum_{i=1}^r \{y_{aj}(i) - y_{bj}(i)\}^L \right)^{1/L}$$

where $y_{aj}(i)$ is the i -th component of the vector \mathbf{y}_{aj} . Eq. (11) is the average distance between the two corresponding points on the trajectories. Distance measure D_1 is limited to the equal duration trajectories. The spatial similarity between two trajectories can be obtained by the Hausdorff distance.

$$D_2(T_a, T_b) = \max\{D_3(T_a, T_b), D_3(T_b, T_a)\} \quad (12)$$

$$D_3(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\|$$

Hausdorff distance D_2 cannot distinguish the direction of the trajectories and suffers from peak noise data. Using a probabilistic model, a trajectory expressed as Eq. (10) can be projected into parametric space, which is a set of HMM parameters. Thus, the shortcomings of metrics such as Eq. (11) or (12) can be overcome [17].

$$D_4(T_a, T_b) = |L(T_a, \lambda_a) + L(T_b, \lambda_b) - L(T_a, \lambda_b) - L(T_b, \lambda_a)| \quad (13)$$

where $L(T_i, \lambda_j)$ is the likelihood of the trajectory T_i to the HMM model λ_j . The model λ_j is fitted to the given trajectory T_j by defining the structure of topology, the number of states, and the probability model using a priori knowledge. These numerous free parameters make applying the HMM difficult in the analysis.

The objectives of comparing the trajectories are mainly within the following categories: (1) classification, (2) recognition, and (3) assessment. We focused on the assessment of trajectory by comparing similarity with others. In such process, the similarity of trajectories with different durations and the variable sampling time are challenging, and they should be pursued. Table 1 shows the similarity analysis, in which we can assess the motion characteristics. Digging and swimming are similar motions. Weight lifting and rowing motions are also similar. Hand shaking has the least similarity with most of the other test motions. The relative similarity between motions can be found without complex calculations using the histogram, regardless of the difference in motion duration. This method can be applied to the automatic investigation of motion improvement in rehabilitation engineering. The integrated control of robotic manipulators can be combined with the adaptation of trajectories in low-dimensional space for the tool motion [20] [21].

4 Conclusions

In this paper, we proposed a new approach to assess human motions using similarity analysis of motion

trajectories. The motion capture database of human action was used to illustrate the technique. High-dimensional data are not appropriate for trajectory analysis; thus, the dimension reduction approach was adopted to obtain tractable simplified data while preserving important features of the original data. Among the dimension reduction algorithms, the locality preserving projection method was used because of its linearity and interpolation property on the reduced space. The motions of human activities are represented in the intrinsic dimension space and are characterized by the sequence of symbols. The comparison of trajectories requires a technique imitating humans intuitive reasoning. Previous methods suffer from constraints or limitations on the duration, ambiguity in direction, or outliers, whereas our method is based on the spatiotemporal features of the trajectory; therefore it is easily applied to compare trajectories with different lengths. The proposed method can also be used for motion retrieval system in computer graphics.

The control of robot system with redundant degree of freedom can utilize the method of low-dimensional embedding to find feasible solutions in complex environment. Human arm has an extra degree of freedom to follow the given trajectory in three-dimensional space without colliding with obstacles. An investigation on the property and the generalization of the method to various human motions will be pursued. The application to the intelligent control algorithm in the humanoid robot system is another area of interest. We focus on the assessment and evaluation of control performance for the movement evolution during the rehabilitation training using the proposed method of representing complex and high dimensional data.

Acknowledgement

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIP) (2010-0020089, 2011-0023437, 2011-0030075) and Kyungpook National University Research Fund, 2012.

The authors are grateful to the anonymous referee for a careful checking of the details and for helpful comments that improved this paper.

References

- [1] J. B. Tenenbaum, V. D. Silva, and J. C. Langford, *Nature*, **290**, 2319 (2000).
- [2] S. T. Roweis, and L. K. Saul, *Nature*, **290**, 2323 (2000).
- [3] T. D. Sanger, *The J. of Neuroscience*, **20**, 1066 (2000)
- [4] M. Ciocarlie, C. Goldfedr, P. Allen, *Proceedings of the IEEE/RSJ Intelligent Robots and Systems*, San Diego, USA, (2007).
- [5] L. Wang, and D. Suter, *IEEE Tr. Image Proc.*, **16**, 1646 (2007).

- [6] R. Vinjamuri, M. Sun, C. Chang, H. Lee, R. J. Scabassi, and Z. Mao, *IEEE Tr. Biomed. Engin.*, **57**, 284 (2010).
- [7] P. K. Artemiadis, and K. J. Kyriakopoulos, *IEEE Tr. Robotics*, **26**, 393 (2010).
- [8] X. He and P. Niyogi, *Proceedings of the International Conference on Advances in Neural Information Processing Systems*, Whistler, Canada, December, 8-13 (2003).
- [9] H. Sadeghi, P. Allard, F. Prince, and H. Labelle, *Gait and Posture*, **12**, 34 (2000).
- [10] F. Porikli, *Proceedings of IEEE European Conference on Computer Vision (ECCV)*, Workshop on PETS, Prague, Czech Republic, (2004).
- [11] Z. Zhang, K. Huang, T. Tan, *Proceedings of the 18th International Conference on Pattern Recognition (ICPR06)*, Hong Kong, China, (2006).
- [12] I. Junejo, O. Javed, M. Shah, *Proceedings of the International Conference on Pattern Recognition (ICPR04)*, Cambridge, UK, **2**, 716 (2004).
- [13] CMU Graphics Lab Motion Capture Database, <http://mocap.cs.cmu.edu/>, (2003).
- [14] E. Levina, A. Mi, P. Bickel, *Ann Arbor MI*, **48109**, 2005-2008 (2004).
- [15] Y. Sakamoto, S. Kuriyama and T. Kaneko, *Proceedings of Eurographics/ACM SIGGRAPH Symposium on Computer Animation*, Grenoble, France, (2004).
- [16] S. Wu, S. Xia, Z. Wang, C. Li, *Proceedings of Computer Graphics International*, Victoria, Canada, (2009).
- [17] Z. Fu, W. Hu, T. Tan, *Proceedings of the International Conference on Image Processing (ICIP05)*, Genova, Italy, **2**, 602 (2005).
- [18] B. Lim, S. Ra, F. C. Park, *Proceedings of the International Conference on Robotics and Automation (ICRA05)*, Barcelona, Spain, 4630 (2005).
- [19] O. C. Jenkins, M. J. Matari, *Proceedings of the IEEE/RSJ Intelligent Robots and Systems (IROS02)*, Lausanne, Switzerland, 2551 (2002).
- [20] H. T. Le, S. R. Lee, C. Y. Lee, *International Journal of Innovative Computing, Information and Control*, **7**, 3255 (2011).
- [21] C. Y. Lee, J. J. Lee, *IEEE Tr. Ind. Electron.*, **52**, 320 (2005).



Sang Ryong Lee received his BS degree in the department of mechanical engineering from Seoul National University, Korea in 1980. He received his MS degree in the department of mechanical engineering from Korea Advanced Institute of Science and Technology, in 1982. He received his Ph. D degree in the department of mechanical engineering from Georgia Institute of Technology, U.S.A. in 1989. He joined the school of mechanical engineering at Kyungpook National University, in 1982, where he is currently a Professor. His research interests are design and control of mobile robots, intelligent systems and automation.



Geun Sub Heo received his BS and MS degrees, both in the school of mechanical engineering from Kyungpook National University, Korea, in 2009 and 2011, respectively. He is currently pursuing Ph. D. degree in the school of mechanical engineering, Kyungpook National University from 2011. His research interests are intelligent control, robotic systems, and embedded systems.



Choon-Young Lee received his B.S. degree in electronic engineering from Hanyang University, Seoul, Korea, in 1996 and the M.S. and Ph. D degrees in electrical engineering from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 1998 and 2003, respectively. He joined the school of mechanical engineering, Kyungpook National University, in 2005, where he is currently a Associate Professor. From 1998 to 2000, he had developed a gait training robot system as a researcher on the project sponsored by the Ministry of Health and Welfare, Korea. He also participated in the development of an emotional robot from 2001 to 2002. His research interests are in the field of neural network application on the design of control system, ITS, sensor fusion, and robotics. Prof. Lee is a member of ICROS and IEEE.