IJCTA, 10(07), 2017, pp. 79-85
© International Science Press

Human Action Recognition Using Dynamic Time Warping Algorithm and Reproducing Kernel Hilbert Space for Matrix Manifold

Wanhyun Cho* Sangkyoon Kim and Soonyoung Park****

Abstract : In this paper, we propose new human action recognition rule using the Dynamic Time Warping (DTW) algorithm and the Reproducing Kernel Hilbert Space (RKHS) for matrix manifold. First, we have reviewed the Dynamic Time Warping algorithm and the RKHS with positive definite kernel function. Second, we define RKHS for a manifold of covariance feature matrix extracted from each frame image of human action video, and then we calculate the distance between two sequences of feature metrics given from two human action videos using a positive definite kernel function. Third, we recognize human action by applying a classifier which is the combination of DTW and k-nearest neighbours (*k*-NN) method to the distance matrix given by two sequence of feature matrices. Finally, we assess a performance of the proposed method using KTH human action data set. Experimental results reveal that the proposed method performs very well on this data set.

Keywords : Human action classification, Dynamic time warping, Reproducing kernel Hilbert space, K-nearest neighbours method, KTH dataset.

1. INTRODUCTION

The automatic recognition and localization of human actions in video sequence has become an important research topic in computer vision, which are very useful for various applications. Among those, several key applications are including advanced human-computer interaction (HCI), assisted living, physical therapy, movies, 3DTV & animations, gesture-based interactive game, sport motion analysis, smart environments, surveillance, and video annotation [1-3]. Here, the human action recognition task are largely consists of three areas which are image representation, human action recognition, and software developments.

First, we have reviewed several research papers for human behaviour recognition using the DTW method. Blackburn and Riberio [4] demonstrated the applicability of Isomap for dimensionality reduction in human motion recognition. And they show how an adapted dynamic time warping algorithm can be successfully used for matching motion patterns of embedded manifolds. Sempena et al [5] built feature vector from joint orientation along time series that invariant to human body size. They applied the dynamic time warping to the resulted feature vector to recognize various human actions. Huu et al [6] presented a human action recognition method using dynamic time warping and voting algorithms on 3D human skeletal models. Gong and Medioni [7] proposed a Spatial-Temporal Manifold model to analyze non-linear multivariate time series with latent spatial structure and apply it to recognize actions in the joint-trajectories space.

Second, several research papers have been published for using the reproducing kernel Hilbert space (RKHS) theory on the human action recognition. Danafar et al [8] introduced two novel approaches

* Dept. of Statistics, Chonnam National University, South Korea. *E-mail : whcho@chonnam.ac.kr*

** Dept. of Electronics Engineering, Mokpo National University, South Korea. *E-mail : narciss76@mokpo.ac.kr, sypark@mokpo.ac.kr*

outperforming state of the art algorithms. They are an unsupervised nonparametric kernel based method exploiting the Maximum Mean Discrepancy test statistic, and a supervised method based on Support vector Machine with a characteristic kernel specifically tailored to histogram-based information. Gaidon et al [9] modeled actions as time series of per-frame representations and proposed a kernel specifically tailored for the purpose of action recognition. Harandi et al [10] proposed to embed the Grassmann manifolds into reproducing kernel Hilbert spaces and then tackle the problem of discriminant analysis on such manifolds.

In this paper, we propose new human action recognition rule, which is a combination of the Dynamic Time Warping (DTW) algorithm and the Reproducing Kernel Hilbert Space (RKHS) for matrix manifold. First, we extract the covariance feature matrix from each image of video sequence representing human action. Next, we embed these feature matrices into RKHS by using a positive definite kernel function, and then we compute the distance on RKHS between two feature matrices. Finally, we recognize the human actions by using DTW and K-nearest neighbours based on derived RKHS distances for two sequence of images.

2. HUMAN ACTION RECOGNITION

Here, we have reviewed the Dynamic Time Warping technique and RKHS theory. Next, we suggest new human action recognition method using theoretical backgrounds of these methods.

A. Dynamic Time Warping

First, DTW algorithm has earned its popularity by being extremely efficient as the time-series similarity measure which minimizes the effects of shifting and distortion in time by allowing elastic transformation of time series in order to detect similar shapes with different shapes [11]. Given two time series

$X = (x_1, ..., x_N)$, and $Y = (y_1, ... , y_M)$ represented by the sequences of values, DTW yields optimal solution in the O(NM) time which could be improved further through different techniques such as multi-scaling.

Algorithm starts by building the distance matrix

$C \in \Re^{N \times M}$ representing all pairwise distances between X and Y. This distance matrix called be local cost matrix for the alignment of two sequences X and Y:

$$C = \{c_{ij} : c_{ij} = \|, (i, j) = (1, 1), ... , (N, M)\} \in \Re^{N \times M}$$

Once the local cost matrix built, the algorithm finds the alignment path which runs through the low-cost areas-“valleys” on the cost matrix. This alignment path (or warping path) defines the correspondence of an element $x_i \in X$ to $y_j \in Y$ following the boundary condition which assigns first and last elements of X and Y to each other.

Formally speaking, the alignment path built by DTW is a sequence of points $p = (p_1, ... , p_K)$ with $p_l = (p_i, p_j)$, $(i, j) \in [1 : N] \times [1 : M]$ for $l \in [1 : K]$ which must satisfy to the following criteria: Boundary condition, Monotonicity condition, Step size condition. The cost function associated with a warping path computed with respect to the local cost matrix will be:

$$c_p(X, Y) = \sum_{l=1}^{K} c(x_{nl}, y_{ml})$$

The warping path which has a minimal cost associated with alignment called the optimal warping path. We will denote this path P^*. To find this path P^*, DTW employs the Dynamic Programming – based algorithm with complexity only O(NM).

B. Reproducing Kernel Hilbert Space

Second, we will also briefly review how to construct a RKHS form any positive definite function and its properties [12-13]. We first define inner product.

Definition 1. Let H be a vector space over $\Re$. A function $<. , .>_{\mathrm{H}} : \mathrm{H} \times \mathrm{H} \to \Re$ is said to be an inner product on H if it must satisfy the following conditions:

1. **Symmetry :** $\langle \varphi, \psi \rangle_{\mathrm{H}} = <\varphi, \psi\rangle_{\mathrm{H}}, \forall\ \varphi, \psi \in \mathrm{H}$.
2. **Bi-linearity :** $\langle \alpha\varphi + \beta\psi, \phi\rangle_{\mathrm{H}} = \alpha<\varphi, \phi\rangle_{\mathrm{H}}, + \beta(\varphi, \phi)_{\mathrm{H}}, \forall\ \phi, \varphi, \psi \in \mathrm{H}, \forall\ \mathrm{H}, \alpha, \beta \in \Re$.
3. **Positive definite:** $\langle \varphi, \varphi\rangle_{\mathrm{H}} \geq 0$, and $\langle \varphi, \varphi\rangle_{\mathrm{H}} = 0$ if and only if $\varphi = 0$.

Now, we define kernel function, and we mention several properties of positive definite kernel related with Hilbert space.

Definition 2. Let X be a non-empty set. A function $k : \mathrm{X} \times \mathrm{X} \to \Re$ is a positive definite kernel if k is symmetric, and k satisfy the following condition:

$\sum_i \sum_j c_i c_j k(x_i, x_j) \geq 0$, for all $(c_1, ..., c_m) \in \Re^m$ and $(x_1, ..., x_m) \in \mathrm{X}^m$.

First, the Gaussian kernel function is positive definite is equivalent with a kernel function is negative definite.

Theorem 3. Let X be a nonempty set and $k : \mathrm{X} \times \mathrm{X} \to \Re$ be a kernel. The kernel $\exp(-\gamma k(x, y))$ is positive definite for all $\gamma > 0$ if and only if the kernel k is negative definite.

Second, a negative definite kernel presents the squared norm of a Hilbert space under some conditions stated by the following theorem.

Theorem 4. Let X be a nonempty set and $k : \mathrm{X} \times \mathrm{X} \to \Re$ be a negative definite kernel. Then, there exists a Hilbert space H and a mapping $\varphi : \mathrm{X} \to \mathrm{H}$ such that

$$k(x, y) \ = \ \| \varphi(x) - \varphi(y) \| + h(x) + h(y)$$

where $h : \mathrm{X} \to \Re$ is a function which is nonnegative whenever k is. Furthermore, if $k(x, x = 0)$ for all $\mathrm{x} \in \mathrm{X}$, then $h = 0$.

Third, the following lemma says the relationship between a feature map and kernel function.

Lemma 5. Let X be a nonempty set, H be an inner product space, and $\varphi : \mathrm{X} \to \mathrm{H}$ be a function. Then, a function $k : \mathrm{X} \times \mathrm{X} \to \Re$ defined by $k(x, y) \equiv \| \varphi(x) - \varphi(y)\|_{\mathrm{H}}$ is negative definite.

Next, we note that every inner product in Hilbert space is a positive definite function, and more generally.

Lemma 6. Let H be any Hilbert space, X be a non-empty set and a map $\varphi : \mathrm{X} \to \mathrm{H}$. Then, $k(x, y) \equiv \langle \varphi(y)\rangle_{\mathrm{H}}$ is a positive definite function.

Now, we define a reproducing kernel function and a reproducing kernel Hilbert space.

Definition 7. Let H be a Hilbert space of $\Re$-valued functions defined on a non-empty set X. A function $k : \mathrm{X} \times \mathrm{X} \to \Re$ is called a reproducing kernel of H, and is a reproducing kernel Hilbert space (RKHS), if a function k satisfies

1. **Reproducing Property :** $\varphi(x) = \langle \varphi(\cdot), k(\cdot , x)\rangle_{\mathrm{H}}, \forall x \in \mathrm{X}, \forall \varphi \in \mathrm{H}$.
2. ***k* span H :** $\mathrm{H} = \{\varphi(\cdot) : \varphi(\cdot) = \sum_i \alpha_i k(\cdot, x), x \in \mathrm{X}\}$.

Finally, we can have a very fundamental results.

Theorem 8 (Moore-Aronszajn). Every positive definite kernel k is associated with a unique RKHS H.

From the results of consideration until now, we can see the following results. In general, the degree of similarity between two patterns in the pattern recognition problem is represented by the distance between the two patterns. But, the feature vectors that extracted from each of the patterns belong to the non-linear manifold. Therefore, it is not appropriate to apply the computer vision algorithms which are primarily developed for data points lying in Euclidean space directly to points on the non-linear manifold.

To overcome this problem, one could think of embedding a point in the manifold into a point in the high dimensional reproducing kernel Hilbert space (RKHS) using kernel methods. Here, according to Mercer’s theorem, the positive definite kernel can transform a point in manifold into a point in RKHS.

Formally, the pattern X in manifold X is first mapped into a function space H using feature map $\varphi : X \to H$, and then positive definite kernel $k : X \times X \to \Re$ is define on space X × X. Hence, to compare two patterns X and X in a manifold, we have to using an inner product $\langle \varphi(x), \varphi(x') \rangle$ in a function space H. But, to avoid working in the potentially high-dimensional function space H, we can replace an inner product $\langle \varphi(x), \varphi(x') \rangle$ in a function space as a positive definite kernel function $k(x, x')$ defined on a pattern space X by mean of the "kernel trick":

$$\langle \varphi(x), \varphi(x') \rangle = k(x, x')$$

In conclusion, in order to compare the degree of similarity between two patterns in manifold X, we have to calculate the distance $\|x - x'\|^2$ of the two patterns. By the way, this can be approximated by the distance $\|\varphi(x) - \varphi(x')\|^2$ in the feature space. This can also be computed using the kernel trick as the following equation:

$$\begin{aligned} \|\varphi(x) - \varphi(x')\|^2 &= \langle \varphi(x) - \varphi(x'), \varphi(x) - \varphi(x') \rangle \\ &= \langle \varphi(x) - \varphi(x) + \varphi(x') - \varphi(x') \rangle - 2 \langle \varphi(x), \varphi(x') \rangle \\ &= k(x, x) + k(x', x') - 2k(x, x') \end{aligned}$$

Finally, to compare two patterns, we only compute the positive definite kernel function.

C. Human Action Recognition

We will suggest new algorithm that can recognize the human action using a combination of DTW and RKHS.

In our action recognition system, we first express two sequences of feature matrices given from two videos of human actions as two time series of feature matrices X and Y:

$$X = [X_1, \dots, X_i, \dots X_{T1}],$$

and

$$Y = [Y_1, \dots, X_j, \dots X_{T2}],$$

where each feature matrices X_i and Y_j are given from $(i, j)^{th}$ frame images of two video data Ω_X and Ω_Y. Here, we will use a feature matrix X_i or Y_j given from each frame image as the covariance matrix descriptor C_X or C_Y defined as following manner.

Definition 9. Given a region of interest R_X or R_Y of each frame image I_X or I_Y, let $X_i \in \Re^d$ or $Y_i \in \Re^d$, for $i = 1, \dots N$ be feature vectors form R_X or R_Y. Then, the covariance matrix descriptor C_X or C_Y is respectively defined as

$$C_X = \frac{1}{N-1} \sum (x_i - \mu_x)(x_i - \mu_x),$$

$$\rightarrow \qquad C_X = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \mu_X)(x_i - \mu_X)^T$$

where μ_x or μ_y is the mean vector, and N is the number of pixels in region R_X or R_Y. Here, we note that the covariance matrix descriptor becomes to be a set of symmetric positive definite (SPD) matrices Sym_d^+, and also they lies on a connected Riemannian manifold.

The DTW algorithm is applied to find the warping path satisfying the conditions minimizing the warping cost. Here, the warping path reveals the similarity matching between two time series input data. Moreover, the similarity between two input action matrices is computed by using the Euclidean distance between the covariance matrix descriptor C_X or C_Y.

But, since two matrixes belong to the Riemannian manifold as mentioned on above, we have to embed a point in the manifold into a point in the high dimensional reproducing kernel Hilbert space (RKHS) using kernel methods. In this case, we will consider the metric on the SPD matrices Sym_d^+ that can be used to define positive definite Gaussian kernel. In particular, we focus on the log-Euclidean distance which is a true geodesic distance on the Riemannian manifold Sym_d^+. Under the log-Euclidean framework, the geodesic distance between $X_i = C_{x_i}$ and $Y_j = C_{y_j}$ for X_i, $Yj \in \mathrm{Sym}_d^+$, can be expressed as,

d_g $(\mathrm{X}_i, \mathrm{Y}_j) = \|\log(\mathrm{X}_i) - \log(\mathrm{Y}_j)\|_{\mathrm{F}}$, where log (·) denotes the usual matrix logarithm operators, and $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius matrix norm induced by the Frobenius matrix inner product $\langle\cdot, \cdot\rangle_{\mathrm{F}}$. Here, the log-Euclidean distance defines a true geodesic distance that has proven an effective distance measure on Sym_d^+. Moreover, it yields a positive definite Gaussian kernel as mentioning in the following Theorem.

Theorem 10. Let $(\mathrm{Sym}_d^+, d_{\mathrm{F}})$ be a metric space and define $k : \mathrm{Sym}_d^+ \times \mathrm{Sym}_d^+ \rightarrow \Re$ be define by

$$k(\mathrm{X}_i, \mathrm{Y}_j) \equiv \exp(-d_g^{\,2}(\mathrm{X}_i, \mathrm{Y}_j)/2\sigma^2).$$

Then, k is a positive definite kernel for all $\sigma > 0$ if and only if there exists an inner product space Z and a function log : $\mathrm{Sym}_d^+ \rightarrow \mathrm{Z}$ such that

$$d_g(\mathrm{X}_i, \mathrm{Y}_j) = \|\log(\mathrm{X}_i) - \log(\mathrm{Y}_j)\|_{\mathrm{F}}.$$

Therefore, in order to calculate the distance matrix for two time series data required by the DTW algorithm, we calculate the distances between every corresponding two component matrices of two time series data by using a symmetric positive definite kernel function as the following form.

$$\begin{aligned} d_g(\mathrm{X}_i, \mathrm{Y}_j) &= \|\log(\mathrm{X}_i) - \log(\mathrm{Y}_j)\|_{\mathrm{F}} \\ &= k(\mathrm{X}_i, \mathrm{X}_i) + k(\mathrm{Y}_j, \mathrm{Y}_j) - 2k(\mathrm{X}_i, \mathrm{Y}_j) \\ &= 2(1 - \exp(-\mathrm{Tr}(\log(\mathrm{X}_i)^{\mathrm{T}} \log(\mathrm{Y}_j))/2\sigma^2)). \end{aligned}$$

And then, by applying DTW algorithm with an obtained distance matrix, we find an alignment warping path with minimum cost, and computed the minimum distance for its warping path.

Finally, in order to recognize or classify a testing human action, we use k-NN method. First, we compute the minimum warping distances between a testing action sample and all training action samples by using DTW. Second, we select k nearest neighbourhood samples that have the smaller distances with testing sample. Third, we count the number k_i, $i = 1, ..., \mathrm{C}$ of training sample that belong to i-th class C_i among a selected k nearest neighbourhood tanning samples, and then we calculate the posterior classification rates,

$$p(\mathrm{C}_i \mid \text{testing sample}) = k_i / k, \; i = 1, \ldots \mathrm{C}.$$

Therefore, if the class C_k has the maxmum of posterior classification rates, we classify a testing sample into the class C_k.

3. EXPERIMENTAL RESULTS

A. KTH Dataset

In order to evaluate the performance of the proposed method, we used the KTH human action dataset. This dataset contains 25 people performing six action classes, namely: walking, running, jogging, hand waving, boxing, and hand clapping. Each video sequence contains one actor performing an action. In order to train the proposed model, we used the KTH human action dataset. This dataset consisted of 384 video sets in which 16 people performed 6 action classes a total of 4 times each. Each frame image of a given video consisted of a black and white image with (160 × 120) resolution. To test the performance of our model, we have also used another dataset of 216 video sequences consisting of 9 people performing 6 different human behavioral tasks a total of 4 times each.

B. Procedure of Human Action Recognition

Figure 1 shows the overall processing of the proposed human action recognition system. This system goes through three steps as follows. In the first step, we computed the covariance descritors using the following feature vectors x given from each image of video representing some human action:

$$x = \left(x, y \,|\, \mathrm{I}_x \,|, |\, \mathrm{I}_y \,|, \sqrt{\mathrm{I}_x^2 + \mathrm{I}_y^2}, |\, \mathrm{I}_{xx} \,|, |\, \mathrm{I}_{yy} \,|, \arctan\left(\frac{|\, \mathrm{I}_x \,|}{|\, \mathrm{I}_y \,|} \right) \right)$$

where x, y are pixel positions, and I_x, I_y are intensity derivatives. The covariance matrix for an image path of arbitrary size is a (8 × 8) symmetric positive definite matrix. And we transform a video sequence for given human action into a sequence of covariance matrices. In the second step, we apply DTW algorithm to matching two time sequences of covariance matrices representing two human actions. In this case, we use a positive definite kernel function to compute the distance between two covariance matrices. In the third step, to recognize a given human action, we use k-NN algorithm for the minimum warping distances obtained by DTW algorithm. That is, we select k nearest neighborhood samples that have the smaller distances with testing sample. We count the number k_i, i = 1 ... , C of training sample that belong to the class C_i among selected k nearest neighborhood tanning samples. And then we classify a testing sample into the class C_k that has the largest number.

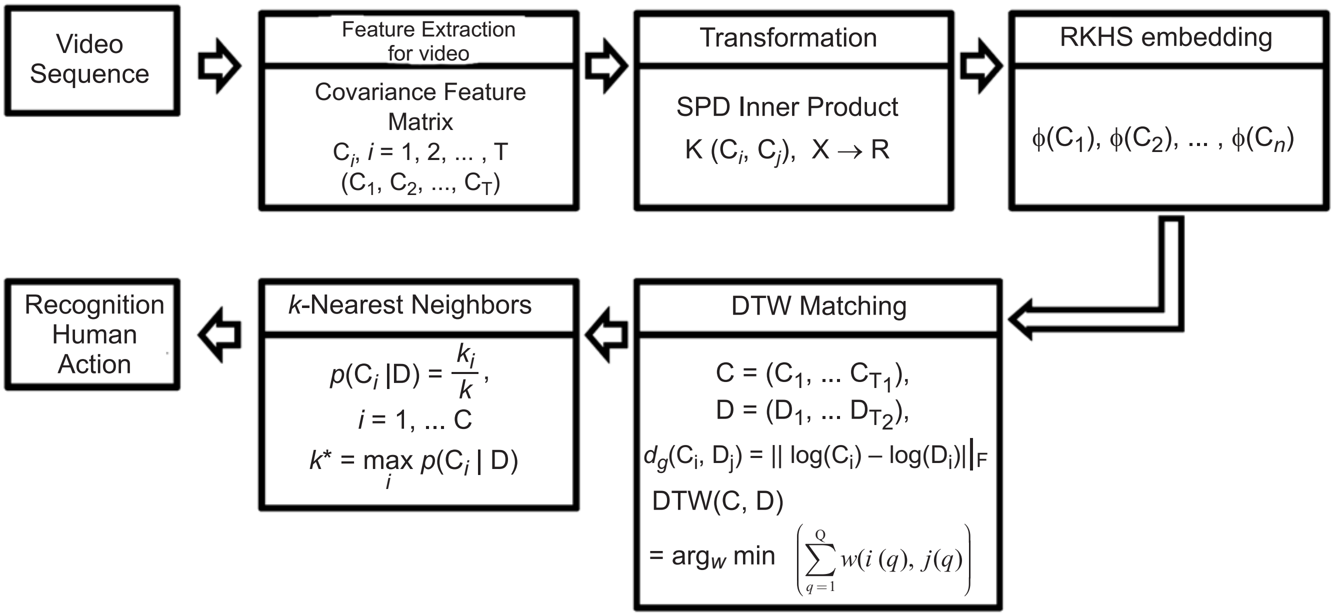
Video Sequence
Feature Extraction for video
Covariance Feature Matrix
C_i, i = 1, 2, ... , T
(C_1, C_2, ..., C_T)
Transformation
SPD Inner Product
K (C_i, C_j), X → R
RKHS embedding
$\phi(C_1), \phi(C_2), \ldots, \phi(C_n)$
DTW Matching
C = (C_1, ... C_{T1}),
D = (D_1, ... D_{T2}),
$d_g(C_i, D_j) = \| \log(C_i) - \log(D_i) \|_F$
DTW(C, D)
$= \arg_w \min \left(\sum_{q=1}^{Q} w(i(q), j(q)) \right)$
k-Nearest Neighbors
$p(C_i | D) = \frac{k_i}{k}$,
i = 1, ... C
$k^* = \max_i p(C_i | D)$
Recognition Human Action

Figure 1: Procedure of Human Action Recognition

C. Recognition Result

Table 1 shows the classification rates for the six human behaviors obtained by the proposed method represented in matrix form. From the results in Table 1, we noted that six human behaviors can be mainly divided into two categories with similar behavior. The first category of similar actions includes boxing, hand-clapping, and hand-waving. The second category of similar behavior includes jogging, running, and walking. Here, we can see, a little misclassification happening a little between the actions belonging to the same category. Consequently, we note that the correct classification rate of the proposed method totally appears to be 84.65% on average.

Table 1
Classification rate

Classification rate	*Boxing*	*Hand clapping*	*Hand waving*	*Jogging*	*Running*	*Walking*
Boxing	1.00	0	0	0	0	0
Hand-clapping	0.14	0.86	0	0	0	0
Hand-waving	0.08	0	0.92	0	0	0
Jogging	0	0	0	0.75	0.25	0
Running	0	0	0	0.17	0.83	0
Walking	0	0	0	0.28	0	0.72

4. CONCLUSION

In this paper, we propose new human action recognition method using the Dynamic Time Warping (DTW) algorithm and a symmetry positive definite kernel function defined on the Reproducing Kernel Hilbert Space (RKHS). First, we computed the covariance feature matrices from human action video. Second, we calculated the distance matrix between two sequences of covariance feature matrices using a symmetry positive definite kernel function. Third, we recognized the testing action video by applying the DTW and *k*-NN method for both training samples and testing sample.

To evaluate a performance of the proposed method, we conducted the experiment using KTH data. Experimental results show that our method performs very well on public video datasets more than others. Our future work will extend the proposed method to other video recognition problems such as 3D human action recognition, gesture recognition, and surveillances system.

5. ACKNOWLEDGMENT

This work was supported by Korea Research Foundation (2016905794).

6. REFERENCES

1. M. B. Holte, C. Tran, M. M. Trievedi, and T. B. Moeslund, “Human Pose Estimation and Activity Recognition from Multi-View Videos: Comparative Explorations of Recent Developments,” IEEE Journal of selected topics in signal processing, vol. 6(5), pp. 538-552, 2012.
2. X. Xu, J. Tang, X. Zhang, X. Liu, H. Zhang, and Y. Qiu, “Exploring Techniques for Vision Based Human Activity Recognition: Methods, Systems and Evaluation,” Sensors, vol. 13, pp. 1635-1650, 2013.
3. G. Cheng, Y. Wan, A. N. Saudagar, K. Namuduri, and B. P. Buckles, “Advances in Human Action Recognition: A Survey,” Proceeding of Computer Vision and Pattern Recognition, Jan. 2015, pp. 1-30.
4. J. Blackburn and E. Ribeiro, “Human Action recognition Using Isomap and Dynamic Time Warping,” LNCS 4814, pp. 285-298, 2007.
5. S. Sempena, N. U. Maulidevi, and P. R. Aryan, “Human Action Recognition Using Dynamic Time Warping,” 2011 International Conference on Electrical Engineering and Informatics, 17-19 July, 2011, Bandung, Indonesia.
6. P. C. Huu, L. Q. Khanh, and L. Thanh Ha, “Human Action Recognition Using Dynamic Time Warping and Voting Algorithm,” Journal of Science: Comp. Science & Com. Eng., vol. 30(3), pp. 22-30, 2014.
7. D. Gong and G. Medioni, “Dynamic Manifold Warping for View Invariant Action Recognition,” Proc. Of the IEEE 13th International Conference on Computer Vision (ICCV 2011), Barcelona, Spain, Nov. 2011.
8. S. Danafar, A. Giusti, J. Shmidhuber, “Novel Kernel Based Recognizers of Human Actions,” EURASIP Journal on Advances in Signal Processing – Special issue on video analysis for human behavior understanding, vol. 2010, Feb. 2010.
9. Gaidon, Z. Harchaoui, and C. Schmid, “A time series kernel for action recognition,” Proc. of BMVC 2011 – British Machine Vision Conference, Aug. 2011, Dundee, United Kingdom.
10. M. T. Harandi, C. Sanderson, S. Shirazi, and B. C. Lovell, “Kernel Analysis on Grasmann Manifolds for Action recognition,” Preprint submitted to Pattern Recognition Letters.
11. P. Senin, “Dynamic Time Warping Algorithm Review,” Information and Computer Science Department, University of Hawaii at Manoa, Honolulu, USA, December, 2008.
12. S. Jayasumana, R. Hartley, M. Salzmann, and M. Harandi, “Kernel Methods on Riemannian Manifolds with Gaussian RBF Kernels,” arXiv:1412.0265v2 [cs.CV] 17 Mar. 2015.
13. Y. Wu, Y. Jia, P. Li, J. Zhang, and J. Yuan, “Manifold Kernel Sparse Representation of Symmetric Positive Definite Matrices and Its Application,” IEEE Transactions on Image Processing, vol. 24(11), Nov. 2011.