

Tensor Data Feature Extraction: The Result of Wayang Golek Menak Dance Capture Motion by Vector Quantization Method

^{1,2}Joko Sutopo, ¹Adhi Susanto, ¹P. Insap Santosa and ¹Teguh Barata Adji

¹Department of Electrical Engineering and Information Technology,
Universitas Gajah Mada, Yogyakarta, Indonesia

²Departement of Electrical Engineering, Universitas Teknologi Yogyakarta, Maltai, Indonesia

Abstract: Capturing the motion of Wayang Golek Menak dance from Yogyakarta using Kinect sensor that produce motion data with format *BVH (Biovision Hierarchy). The motion data have coordinate x, y, z then, the reduction is conducted by Principal Component Analysis (PCA) methods the result of data reduction by PCA is conducted using fitting process. Fitting is extraction feature that looking for the characteristic of feature so it can be classified. Fitting methods which is applied in this research is Vector Quantization (VQ). This method is for compressing feature data by making code book then, it's changed into code vector that can use to classified on the next process. The result from Vector Quantization (VQ) method with code vector which the value is 00, 11, 11, 00 with the code book 0.25; 0.5; 0.75 and 1.

Key words: Extraction feature, dance motion, Vector Quantization (VQ), code book, code vector, PCA

INTRODUCTION

The art of puppetry, puppet Menak (Setyowibowo, 1991) or Beksa Golek Menak (Smith, 2002; Holland, 2008; Sukistono, 2010) consists of a series of dance movements that have coordinated and integrated with regard to flexibility, vibration (Giovanni *et al.*, 2012), translational and geometric transformations (Hassani, 2008; Supandi, 1979; Supriyanto, 1992).

Kinect technology has the ability to interpret gestures or specific gestures by utilizing RGB camera and projector infrared microchip to record 3D video data in certain lighting (Han *et al.*, 2013; Held *et al.*, 2012; Pham *et al.*, 2013) by synchronizing the Kinect sensor (Warade *et al.*, 2012) in a pose (Aggarwal and Cai, 1997) motion on the body skeleton (Han *et al.*, 2013; Liu *et al.*, 2009; Xu and Lee, 2014) with different variations (Chen *et al.*, 2013).

A dataset motion capture (mocap) is a digital data representation of the results of body motion capture techniques at certain times (Guerra, 2005) into a data format Biovision Hierarchy (BVH) (Chen *et al.*, 2013; Liu *et al.*, 2009, 2010) in the development of data (Han *et al.*, 2013; Liu *et al.*, 2009; Xu and Lee, 2014). The tensor is essentially the general form of the scalar and vector which has a geometric relationship between the

scalar and vector operations in a process of multi-dimensional (Liu *et al.*, 2009) to help model the human body in general (Chen *et al.*, 2013).

This research using tensor data feature extraction of wayang golek dance motion Menak, Yogyakarta from the capturing result by Kinect sensor. Feature extraction using Vector Quantization (VQ) method or known as fitting is purpose to find feature characteristic which will be used for feature classification. Position of tensor data is x, y, x from dancer which is captured by Kinect sensor. Motion position data will be in reduction processed and feature extraction that purpose to accelerate the computation process. The result of VQ method is code vector, it will be used in the next level for classifying.

Literature review

Motion capture technique: Motion capture data is representation of digital data from the result of capturing technique to the actor or figure motion on real time mode (Yoon *et al.*, 2010) through the Kinect sensor (Chen *et al.*, 2013). Digital data which is in format motion data or position or coordinate orientation (points) from the position of body motion on certain time (Guerra, 2005). The display of mocap data format is shown on 3D format (X, Y, Z) related with figure character. Mocap data format

Table 1: Matrix to high order tensor

Matrik	High order tensor
(Columns/Rows) A_i, A_j	(Fibers) X_{ij}, X_{ik}, X_{jk} (slices) $X_{i\cdot}, X_{\cdot j}, X_{\cdot k}$ Mode-3,2,1; horizontal, lateral, frontal slices
$A:B = \sum_{i=1}^n \sum_{j=1}^m A_{ij} B_{ij}$	$X:Y = \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^p X_{ijk} Y_{ijk}$
$\ A\ = \sqrt{\sum_{i=1}^n \sum_{j=1}^m A_{ij}^2}$	$\ X\ = \sqrt{\sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^p X_{ijk}^2}$
(Rank one matrix) $A = ab^T; A_{ij}=a_i b_j$	(Rank one tensor) $X = a \circ b \circ c; X_{ijk} = a_i b_j c_k$
(symmetric) $A = A^T$	(Supersymmetric: cubical+symmetry) $X \in R^{1 \times 1 \times 1}$ X_{ijk} is constant when permuting i, j, k
$A = I, A_{ij} = \delta_{ij}$	$X = T, X_{ijk} = \delta_{ijk}$

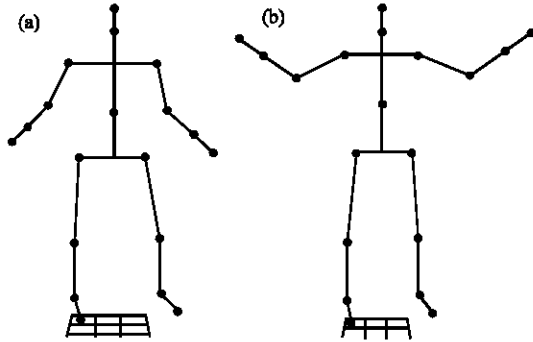


Fig. 1: Skeleton structure BVH: a) the initial position of the frame skeleton and b) position the frame skeleton-1

consist of skeleton part (Obdrzalek *et al.*, 2012) which is total representation of character motion, bone as the basic entity of skeleton which is the funtion as transformation subject, channel or Degree of Freedom (DOF) as the parameter for transformation from bone (translation, rotation (Chen *et al.*, 2013; Liu *et al.*, 2009; Xu and Lee, 2014) move orientation) also the frame as collection channel/DOF information for each bone in a pose animation consist of some frame where each animation consist of 260 frame per second (fps) maximum, evenly the frame taken consist of 30 untill 60 fps.

Biovision Hierarchy (BVH) used as mocap data format which is produced by Kinect, BVH have supporting data which is compatible for import and export at some 3D software (Chen *et al.*, 2013). BVH format data consist of 2 part, the first part is information about hierarchy structure from the bone and the second is parameter data information in each channel. BVH is used for data standard of animation movement based on human body structure also as technique supporter of animation mocap and other supporting software inside the development data (Xu and Lee, 2014).

BVH skeleton structure in Fig. 1 and 2 in the form of dots and dashes that represented a process of continued movement of the bones and joints of the body. With the data format of BVH mocap results can be used to import and export animation data in a variety of application programs include: 3ds Max, Motion Builder, Blender, Daz Studio, Maya, Poser and Lightwave 3D (Izani *et al.*, 2004).

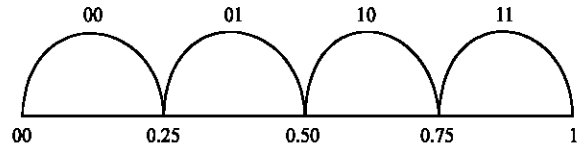


Fig. 2: Code book and code vector

Tensor data: Basically, tensor is a general form from scalar vector and matrix (Barmpoutis *et al.*, 2007; Kolda and Bader, 2009; Yoon *et al.*, 2010). In tensor, scalar is order-0, the vector is an order-1 and the matrix is order-2. Tensor with orders of >3 is a high order tensor (Ueltschi, 2014) (Table 1). Scalar, vector, matrix and rank tensor (3.29) from the table above can define by:

$$\begin{aligned}
 a &= \partial_i e_i b = b_j \tilde{e}_j c = c_k \hat{e}_k \text{ (Vector spaces)} \\
 A &= a \circ b = \underbrace{a_i b_j}_{=A_{ij}} e_i \circ \tilde{e}_j \circ \text{ (Matrix, second order tensor)} \\
 A &= a \circ b \circ c = \underbrace{a_i b_j c_k}_{=X_{ijk}} e_i \circ \tilde{e}_j \circ \hat{e}_k \text{ (Third, order tensor)}
 \end{aligned}$$

Tensor data is data that has three or more point coordinates (i.e., coordinates x, y, z). Data dance Wayang Golek Menak the catch (Motion Capture) with Kinect sensor already has a data tensor (x, y, z). The tensor data that will be used in the next process (Laghdir *et al.*, 2008).

Feature extraction: Knowing or classifying object at picture should extration picture features first then those features are used inside of patter to get classifier final class. Extraction feature (Kapsouras *et al.*, 2013; Liu *et al.*, 2009) purpose to looking for significant feature area at the picture. Those area can be defined in global environment or local and can be differented by form, texture, size, statistic trait and others (Gen and Runwei, 2000). Local features extraction method can be divided into being suitable intensity and based structure. Method intensity based on pattern analyze of local intensity to find the area which have uniqueness desired or stability criteria. Extraction feature consider to identify the characteristics which can be made good representation from object so it can differentiate object categoried with the variant tolerance (Goldberg, 1989).

Vector Quantization (VQ): Vector Quantization (VQ) is lossy data compression method based on coding block principle. This transform using code book. In some case, code book can be used to code entropy of discrete value at the same way, it will produce value code that encoded length-variable as the output first (Lu and Chang, 2010). Code book which is made is synchronized by motion data which produce code vector. Code vector is the result of feature reduction then, the result will be used in the next step introducing wayang golek motion dance Menak, Yogyakarta (Fig. 2).

Principal Component Analysis (PCA): Data tensor dance Wayang Golek Menak have dimensions large matrix so that required data reduction tensor dance. The purpose of this tensor data reduction process to become lighter computing. The method used to reduce the data tensor is Principal Component Analysis (PCA) (Smith, 2002; Holland, 2008). If be defined a matrix A with x an eigenvector and λ is the eigenvalue then to get the eigenvector and eigenvalue can use any general equations of PCA:

$$Ax = \lambda x$$

$$(A - \lambda I)x = 0$$

The results of the reduction in the dance motion sensor data is used to classify the dance movement into two classes, namely Jogetan and Sabetan.

MATERIALS AND METHODS

The methods of this research is observation and direct interview, the informant is expert dancer of wayang golek Menak, Yogyakarta to knowing basic motions of wayang golek dance Menak. Below this is the process of data feature extraction wayang golek dance Menak, Yogyakarta (Fig. 3).

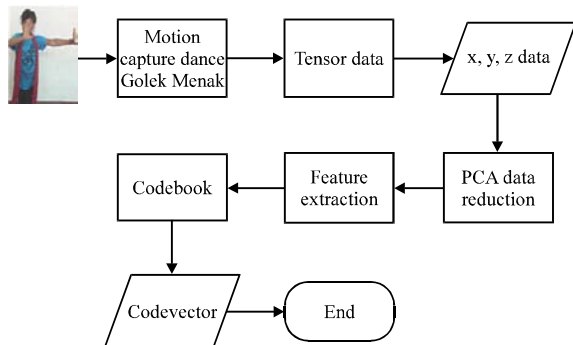


Fig. 3: Flowchart feature data extraction of Wayang Golek Menak dance motion, Yogyakarta

RESULTS AND DISCUSSION

Capturing motion of wayang golek dance Menak, Yogyakarta by kinect sensor will be produce motion data with *BVH (Biovision Hierarchy) format. That motion data is tensor data such as point x, y, z so the position happen is hr, hφ, hθ. The next step is doing data reduction by Principal Component Analysis (PCA) method then, the future extraction is held by Vector Quantization (VQ). Below this is the process of future extraction (Table 2 and 3). Matrix of head motion data from wayang golek motion dance Menak, Yogyakarta with 50×183 dimension.

Table 2: BVH dance motion data

Features	1	2	...	182	183
1	49,1575	-459,115	...	-460,3857	112,4401
2	55,8285	50,4966	...	52,4788	-463,826
3	-458,651	54,602	...	113,0160	47,37443
4	49,1111	-459,7830	...	-461,5900	112,4401
5	55,9189	50,4966	...	52,4788	-463,8257
6	-458,6820	54,6020	...	113,0160	45,6038
7	49,1111	-459,7830	...	-461,5900	111,9682
8	55,9189	50,4966	...	52,4788	-464,6176
9	-458,6820	54,6020	...	113,0160	44,7971
10	49,2492	-459,7830	...	-461,5900	111,8297
11	55,901	50,6734	...	52,4788	-465,200
12	-458,812	55,0423	...	113,016	44,6894
13	49,2492	-460,004	...	-461,590	111,826
14	55,901	50,7066	...	52,4788	-465,253
15	-458,812	54,8777	...	113,016	44,6894
16	49,2492	-460,027	...	-461,590	111,826
17	55,901	50,7865	...	52,9786	-465,253
18	-458,812	54,6818	...	113,299	44,6894
...
49	49,6509	459,328	...	-463,537	112,7761
50	55,5504	47,3258	...	47,37443	-462,896

Table 3: The result of data reduction with PCA

Features	1	2
1	0,987	0,002
2	1,000	0,991
3	0,001	0,999
4	0,987	0,000
5	1,000	0,991
6	0,001	0,999
7	0,987	0,000
8	1,000	0,991
9	0,001	0,999
10	0,987	0,000
11	1,000	0,991
12	0,000	0,999
13	0,987	0,000
14	1,000	0,991
15	0,000	0,999
16	0,987	0,000
17	1,000	0,992
18	0,000	1,000
19	0,987	0,000
20	1,000	0,992
...
49	0,988	0,001
50	0,999	0,985

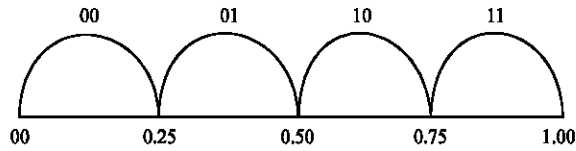


Fig. 4: Code book

Table 4: The result of extration feature on code vector

Features	1	2	-----Code vector -----		
1	0,987	0,002	11	00	0011
2	1,000	0,991	11	11	1100
3	0,001	0,999	00	11	1111
4	0,987	0,000	11	00	1100
5	1,000	0,991	11	11	1111
6	0,001	0,999	00	11	0011
7	0,987	0,000	11	00	1100
8	1,000	0,991	11	11	1111
9	0,001	0,999	00	11	0011
10	0,987	0,000	11	00	1100
11	1,000	0,991	11	11	1111
12	0,000	0,999	00	11	0011
13	0,987	0,000	11	00	1100
14	1,000	0,991	11	11	1111
15	0,000	0,999	00	11	0011
16	0,987	0,000	11	00	1100
17	1,000	0,992	11	11	1111
18	0,000	1,000	00	11	0011
19	0,987	0,000	11	00	1100
20	1,000	0,992	11	11	1111
...
49	0,988	0,001	11	00	1100
50	0,999	0,985	11	11	1111

Data reduction by Principal Component Analysis (PCA) method. The function of this data reduction is accelerating computation process. The result of this data reduction by PCA method (Table 3). Then, make code book based on value from data reduction result. code book of this reserach (Fig. 4). Using the code book, the reduction result is changed into code vector. The result in Table 4.

CONCLUSION

The conclusion of this research is extracting character by Vector Quantization method will be produse code vector fiture character so it can be used for fiture classification process.

REFERENCES

Aggarwal, J.K. and Q. Cai, 1997. Human motion analysis: A review. Proceedings of the IEEE Workshop on Nonrigid and Articulated Motion, June 16, 1997, IEEE, San Juan, Puerto Rico, ISBN:0-8186-8040-7, pp: 90-102.

Barmpoutis, A., B.C. Vemuri, T.M. Shepherd and J.R. Forder, 2007. Tensor spline for interpolation and approximation of DT-MRI with applications to segmentation of isolated rat Hippocampi. *IEEE. Trans. Med. Imaging*, 26: 1537-1546.

Chen, Y., Z. Liu and Z. Zhang, 2013. Tensor-based human body modeling. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 23-28, 2013, IEEE, Portland, Oregon, pp: 105-112.

Gen, M. and C. Runwei, 2000. Genetic Algorithm and Engineering Optimization. John Wiley & Son Inc, New York, USA., Pages: 450.

Giovanni, S., Y. Choi, J. Huang, E. Khoo and K. Yin, 2012. Virtual Try-On using Kinect and HD Camera. In: Motion in Games, Kallmann, M. and K. Bekris, (Eds.). Springer, Berlin, Germany, pp: 55-65.

Goldberg, D.E., 1989. Genetic Algorithm in Search, Optimization and Machine Learning. Addison Wesley Publication Co., USA., ISBN-13: 9780201157673, Pages: 412.

Guerra, F.G., 2005. Optical motion capture: Theory and implementation. *RITA.*, 12: 61-90.

Han, J., L. Shao, D. Xu and J. Shotton, 2013. Enhanced computer vision with microsoft kinect sensor: A review. *IEEE. Trans. Cybern.*, 43: 1318-1334.

Hassani, S., 2008. Mathematical Methods for Students of Physics and Related Fields. Springer, Berlin, Germany, Pages: 831.

Held, R., A. Gupta, B. Curless and M. Agrawala, 2012. 3D puppetry: A kinect-based interface for 3D animation. Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology, October 7-10, 2012, ACM, New York, USA., pp: 423-434.

Holland, S.M., 2008. Principal Components Analysis (PCA). Department of Geology, University of Georgia, Athens, GA., USA., May 2008, pp: 1-10.

Izani, M., A.R. Eshaq, N. Zainuddin and A. Razak, 2004. A study on practical approach of using motion capture and keyframe animation techniques. Proceedings of the 8th International Conference on Information Visualisation, July 16, 2004, IEEE, London, England, ISBN:0-7695-2177-0, pp: 849-852.

Kapsouras, I., S. Karanikolos, N. Nikolaidis and A. Tefas, 2013. Feature Comparison and Feature Fusion for Traditional Dances Recognition. In: Engineering Applications of Neural Networks, Iliadis, L., H. Papadopoulos and C. Jayne (Eds.). Springer, Berlin, Germany, pp: 172-181.

Kolda, T.G. and B.W. Bader, 2009. Tensor decompositions and applications. *J. SIAM. Rev.*, 51: 455-500.

Laghdar, A., Y. Fortin, D.L. Cruz, C. Mariella and R.E. Hernandez, 2008. Development of a technique to determine the 3D elasticity tensor of wood as applied to drying stress modeling. *Wood Sci. Technol.*, 10: 35-44.

- Liu, J., B. Zuo, X. Zeng, P. Vroman and B. Rabenasolo, 2010. Nonwoven uniformity identification using wavelet texture analysis and LVQ neural network. *Expert Syst. Appl.*, 37: 2241-2246.
- Liu, M., Z. Miao, J. Li and Z. Xu, 2009. Design and implementation of a vision-based motion capture system. *Proceedings of the 5th International Conference on Image and Graphics*, September 20-23, 2009, IEEE, Beijing, China, ISBN:978-1-4244-5237-8, pp: 716-721.
- Lu, T.C. and C.Y. Chang, 2010. A survey of VQ code book generation. *J. Inform. Hiding Multimedia Signal Process.*, 1: 190-203.
- Obdrzalek, S., G. Kurillo, F. Ofli, R. Bajcsy and E. Seto *et al.*, 2012. Accuracy and robustness of Kinect pose estimation in the context of coaching of elderly population. *Proceedings of the 2012 IEEE Annual International Conference on Engineering in Medicine and Biology Society (EMBC)*, August 28-September 1, 2012, IEEE, Berkeley, California, ISBN: 978-1-45 77-1787-1, pp: 1188-1193.
- Pham, H.T., J.J. Kim and G.W. Young, 2013. A low cost system for 3D motion analysis using microsoft kinect. *Appl. Mech. Mater.*, 284: 1996-2000.
- Setyowibowo, W., 1991. [The story of Wayang Menak in Javanese song tembang]. Master Thesis, University of Michigan, Ann Arbor, Michigan. (In Indonesian).
- Smith, L.I., 2002. A Tutorial on Principal Component Analysis. Cornell University, New York, USA.
- Supandi, A., 1979. [Gending accompaniment puppet, dance accompaniment, kilingan: Presentation of Sundanese gamelan]. Gedung A Kementerian Pendidikan dan Kebudayaan, Central Jakarta, Indonesia. (In Sundanese)
- Supriyanto, 1992. [Form of presentation of golek dance classification menak Gaya Yogyakarta: Research report]. Sekolah Tinggi Seni Indonesia, Bandung, Indonesia.
- Ueltschi, T., 2014. Third-order tensor decompositions and their application in quantum Chemistry. Master Thesis, University of Puget Sound Tacoma, Tacoma, Washington.
- Warade, S., J. Aghav, P. Claude and S. Udayagiri, 2012. Real-time detection and tracking with Kinect. *Proceedings of the International Conference on Computer and Information Technology (ICCIT'2012)*, June 16-17, 2012, Bangkok University, Bangkok, Thailand, pp: 86-89.
- Xu, W. and E.J. Lee, 2014. Face recognition using wavelets transform and 2D PCA by SVM classifier. *Intl. J. Multimedia Ubiquitous Eng.*, 9: 281-290.
- Yoon, S.M., M. Scherer, T. Schreck and A. Kuijper, 2010. Sketch-based 3D model retrieval using diffusion tensor fields of suggestive contours. *Proceedings of the 18th ACM International Conference on Multimedia*, October 25-29, 2010, ACM, New York, USA., ISBN:978-1-60558-933-6, pp: 193-200.