CGA CLUSTERING BASED VECTOR QUANTIZATION APPROACH FOR HUMAN ACTIVITY RECOGNITION USING DISCRETE HIDDEN MARKOV MODEL

Nguyen Nang Hung Van[1], Pham Minh Tuan[1], Tachibana Kanta[2]

[1]*Danang University of Science and Technology; nguyenvan@dut.udn.vn, pmtuan@dut.udn.vn*

[2]*Kogakuin University; kanta@cc.kogakuin.ac.jp*

Abstract - Activity recognition has been taken great consideration by many scientists all over the world. However, the conventional research results need to be improved because of the complexity and unstability of object recognition. Especially with human activity recognition (HAR) in 3-dimensional space, the vector quantization based on k-means was not able to cluster two objects rotating around a common point but on a different plane because they have the same cluster centroid. In this paper, we propose a new method of vector quantization (VQ) performance optimally distribute VQ codebook components on Hidden Markov Model (HMM) state. This proposed method is carried out through two steps. First, the proposed method use Conformal Geometric Algebra (CGA) clustering algorithms to optimize VQ. Then, the proposed method uses discrete HMM to recognize the human activity. The experimental result with the CMU graphics lab motion capture database shows that the proposed method is more effective than conventional method.

Key words - Hidden Markov Model; vector quantization; clustering; k-mean; conformal geometric algebra.

1. Introduction

Human activity recognition is one of the important areas of computer vision research. Its applications include intelligent security monitoring system, health care systems, intelligent transportation systems, and a variety of systems that involve interactions between people and electronic devices such as human computer interfaces. Today there are many researches on human activity recognition area. For example, the discrete HMM (DHMM) is one of the most common recognition models and it is applied in many human activity recognitions such as human activity recognition using monocular camera [1] or speech recognition system [2].

In this paper, we focus on clustering algorithm based VQ for DHMM [3]. In conventional methods, the *k*-means is usually used to quantize a vector before applying to DHMM.

The advantage of *k*-means algorithm is simple and easy to understand and install. It is able to apply to assign the data to groups using Euclidean distance. However, using Euclidean distance is also the disadvantage of *k*-means algorithm in the case of 3-dimensional data. For example, when we have two objects rotating around a common point but is not same a plane, we can not cluster the coordinates of two objects correctly because two cluster centers of k-means will be same. Therefore, the result of *k*-means based vector quantization for the 3-dimensional rotation data such as human activity is not good. So, this paper proposed to use CGA clustering to quantize a vector for DHMM. CGA is a part of (Geometric Algebra) GA and is also called Clifford Algebra. CGA is the GA constructed over the resultant space of a projective map from an *m*-dimensional Euclidean or pseudo-Euclidean base space $\mathcal{R}^m$ into $\mathcal{G}_{m+1,1}$. This allows operations on the *m*-dimensional space, including rotations, translations and reflections to be represented using versors of the GA [4]; and it is found that points, lines, planes, circles and spheres gain particularly natural and computationally amenable representations [5, 6]. And, there are many applications of GA as signal processing model, using image processing of complex spatial GA [7] or quaternions [8].

In this paper, we present a new CGA clustering approach to improve the accuracy of the DHMM on HAR system based on VQ by implementing the optimal distribution of the codebook of HMM states. This technique, which has been named the distributed VQ of HMM is done through two steps. The first is to use CGA clustering algorithm to optimize the VQ, the next step will be to conduct HMM parameter estimation and classification of action [9].

The paper is structured as follows. The first is the introduction of this paper. The second presents the related research. Section 3 reports conformal geometric algebra and describes CGA clustering approach for DHMMs. Section 4 reports the comparative results of the proposed method using CGA clustering and conventional methods using *k*-means. Finally, the 5th section summarizes this paper.

2. Related research

This section presents the basic of a VQ for the discrete hidden Markov models (DHMMs). This section summarises a k-means based VQ and review DHMMs.

2.1. K-means based vector quantization

Vector quantization is a process of the mapping of a sequence of m-dimensional continuous vectors [10, 11] $\boldsymbol{O} = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_T\}, \boldsymbol{v}_t \in \mathbf{R}^m$ to a discrete, one dimensional sequence of codebook indices $\widehat{\boldsymbol{O}} = \{\widehat{\boldsymbol{v}}_1, \cdots, \widehat{\boldsymbol{v}}_T\}, \widehat{\boldsymbol{v}}_t \in \mathbf{N}$ where a codebook $\boldsymbol{C} = \{\boldsymbol{c}_1, \cdots, \boldsymbol{c}_K\}, \boldsymbol{c}_k \in \mathbf{R}^m$ and K is the number of centroids $\boldsymbol{c}_i$. The assignment of the continuous sequence to the codebook indices is a minimum distance search if the codebook $\boldsymbol{C}$ is generated,

$$\widehat{\boldsymbol{v}}_i = \underset{k}{\operatorname{argmin}}\, d(\boldsymbol{v}_i, \boldsymbol{c}_k), \forall i \in [1, \cdots, T]$$

where $d(\boldsymbol{v}_i, \boldsymbol{c}_k) = \|\boldsymbol{v}_i - \mathbf{c}_k\|^2$ is the squared Euclidean distance. There are many ways to generate the codebook. This section describes a basic method to generate $\boldsymbol{C}$ using *k*-means clustering.

Given a training set $\boldsymbol{S}_{\boldsymbol{train}} = \{\boldsymbol{O}_1, \cdots, \boldsymbol{O}_N\}$, where $\boldsymbol{N} = |\boldsymbol{S}_{\boldsymbol{train}}|$ is the number of samples $\boldsymbol{O}_i = \{\boldsymbol{v}_{i,1}, \cdots, \boldsymbol{v}_{i,T}\}, \boldsymbol{v}_{i,t} \in \mathbf{R}^m$. A codebook $\boldsymbol{C}$ is calculated by minimizing the following problem,

$$\min_{u,\mathbf{c}} \sum_{k=1}^{K} \sum_{i=1}^{N} \sum_{t=1}^{T} u_{k,i,t}\, d(\boldsymbol{v}_{i,t}, \boldsymbol{c}_k)$$

$$\text{s.t.} \quad \sum_{k=1}^{K} u_{k,i,t} = 1\ , \qquad u_{\,k,i,t} \in \{0,1\},$$

where $d(\boldsymbol{v}_{i,t}, \boldsymbol{c}_k) = \|\boldsymbol{v}_{i,t} - \mathbf{c}_k\|^2$ is the squared Euclidean distance between the vector $\boldsymbol{v}_{i,t}$ and the k^{th} codebook centroids $\boldsymbol{c}_k$. The *k*-means clustering algorithm to calculate the codebook $\boldsymbol{C}$ is described as follows.

```
algorithm kmeans_codebook()
input:
  v[N][T_N]: training set
  K: number of centroids
output:
  u[N][T_N]: memberships
  C[K]: array of codebook centroids
begin
  δ ← 1
  while (δ>0)
    δ ← 0
    for k from 0 to K-1 do
      C_new[k] ← 0 //Zero vector
      C_size[k] ← 0
    endfor

    for i from 0 to N-1 do
      for t from 0 to T(i)-1 do
        dmin ← ∞
        n ← 0
        for k from 0 to K-1 do
          d ← |V[i][t] - C[k]|
          if d<dmin then
             dmin ← d
             n ← k
          endif
        endfor
        if u[i][t] ≠ n then
           δ ← δ + 1
           u[i][t] ← n
        endif
        C_new[n] ← C_new[n] + V[i][t]
        C_size[n] ← C_size[n] + 1
      endfor
    endfor

    for k from 0 to K-1 do
      C[k] ← C_new[k] / C_size[k]
    endfor
  endwhile
end
```

2.2. Discrete HMMs

HMMs are the important methods to model temporal and sequence data. They are especially known for their application in real time pattern recognition such as handwriting digits recognition, speech recognition [12] and human ativity recognition. HMMs attempt to model such systems and allow:

(1) to infer the most likely sequence of states that produced a given output sequence,

(2) to infer which will be the most likely next state,

(3) to calculate the probability that a given sequence of outputs originated from the system.

This paper focuses on ability (3) of discrete HMMs. It means that this paper uses DHMMs for sequence classification. The DHMMs have been defined by the following set of parameters,

$$\lambda = \{A, B, \pi\}$$

where A is the state transition probability distribution given in the form of a matrix $A = \{a_{ij}\}$. B is the observation symbol (codebook index) probability distribution given in the form of a matrix $B = \{b_j(k)\}$ and π is the initial state distribution. Figure. 1 shows an example of Hidden Markov model.

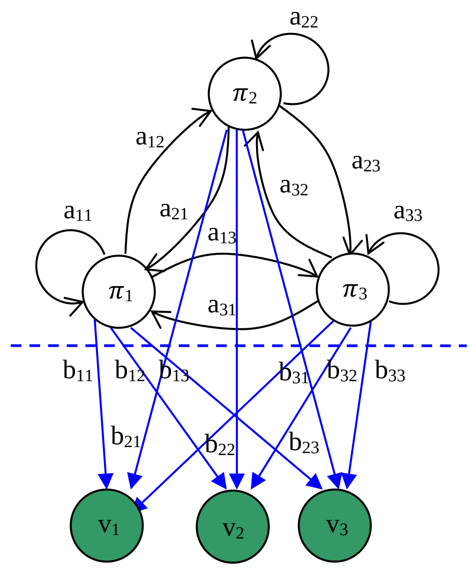

***Figure 1.** An example of Hidden Markov model*

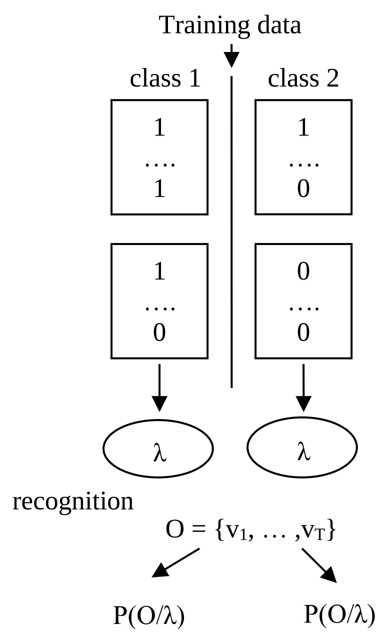

***Figure 2.** Recognition via HMM*

In order to use DHMMs, the continuous observations $\boldsymbol{O} = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_T\}, \boldsymbol{v}_t \in \mathbf{R}^m$ are vector quantized yielding discrete observation sequences $\widehat{\boldsymbol{O}} = \{\widehat{\boldsymbol{v}}_1, \cdots, \widehat{\boldsymbol{v}}_T\}, \widehat{\boldsymbol{v}}_t \in \mathbf{N}$

using the codebook. In the case of classification problem, we need create a set of models and specialise each model to recognize each of the separated classes. The parameters λ_i of the i[th] model can be trained with the well known EM algorithm [13]. After all models have been trained, the probability of the unknown-class sequence can be computed for each model. As each model specialised in a given class, the one which outputs the highest probability can be used to determine the most likely class for the new sequence, as shown in Figure 2.

3. Proposed method

This paper proposed a new vector quantization approach for DHMMs based on conformal geometric algebra clustering. This section reviews conformal geometric algebra and describes conformal geometric algebra clustering approach for DHMMs.

3.1. Conformal Geometric Algebra

Conformal Geometric Algebra is a part of Geometric Algebra [14] and is also called Clifford Algebra. GA defines the signature $p+q$ orthonormal basis vector $\mathcal{O}=\{e_1,\dots,e_p,e_{p+1},\dots,e_{p+q}\}$, such as $e_i^2=1,\forall i\in\{1,\dots,p\}$ and $e_i^2=-1,\forall i\in\{p+1,\dots,q\}$. GA denotes $\mathcal{O}$ by $\mathcal{G}_{p,q}$. For example, m-dimensional Euclidean vector space $\mathcal{R}^m$ is denoted by $\mathcal{G}_{m,0}$.

A CGA space is extended from the real Euclidean vector space $\mathcal{R}^m$ by adding 2 orthonormal basis vector. Thus, a CGA space is defined by $m+2$ basis vectors $\mathcal{O}=\{e_1,\dots,e_m,e_+,e_-\}$, where e_+ and e_- are defined as following:

$$e_+^2 = e_+\cdot e_+ = 1,$$
$$e_-^2 = e_-\cdot e_- = -1,$$
$$e_+\cdot e_- = e_+\cdot e_i = e_-\cdot e_i = 0, \forall i\in\{1,\dots,m\}.$$

Thus, a CGA can be expressed by $\mathcal{G}_{m+1,1}$. In addition, CGA defined:

$$e_0=\frac{1}{2}(e_- - e_+)$$
$$e_\infty=(e_- + e_+).$$

It is easy to see that:

$$e_0.e_0=e_\infty.e_\infty=0,$$
$$e_0.e_\infty=e_\infty.e_0=0,$$
$$e_0.e_i=e_\infty.e_i=0,\forall i\in\{1,\dots,m\}.$$

A conformal vector S is generally written in the following:

$$S=\boldsymbol{s}+s_\infty e_\infty+s_0e_0$$

Where $\boldsymbol{s}=\sum_i^m s_ie_i$ is a real vector in the Euclidean space $\mathcal{R}^m$. And s_∞, s_0 are the scalar coefficients of the basic vector e_∞ and e_0.The CGA can express a point, a sphere or a plane based on S. For example, sphere is represented as a following conformal vector:

$$S=x+\frac{1}{2}\{\|\boldsymbol{x}\|^2-r^2\}e_\infty+e_0,$$

where the sphere has center $\boldsymbol{x}$ and radius r in real Euclidean space $\mathcal{R}^m$. Note that the inner product $S\cdot Q$ is 0 for any point Q on the surface of the sphere S.

3.2. CGA clustering based vector quantization

This paper proposes a new vector quantization approach for DHMMs based on conformal geometric algebra clustering.

Given a training set $\boldsymbol{S_{train}}=\{\boldsymbol{O}_1,\cdots,\boldsymbol{O}_N\}$, where $\boldsymbol{N}=|\boldsymbol{S_{train}}|$ is the number of samples $\boldsymbol{O}_i=\{\boldsymbol{v}_{i,1},\cdots,\boldsymbol{v}_{i,T}\},\boldsymbol{v}_{i,t}\in\mathbf{R}^m$. This paper converts all samples $\boldsymbol{O}_i$ to a set of points $\boldsymbol{P}_i=\{\boldsymbol{p}_{i,1},\cdots,\boldsymbol{p}_{i,T}\}$, $\boldsymbol{p}_{i,t}=\boldsymbol{v}_{i,1}+\frac{1}{2}\|\boldsymbol{v}_{i,1}\|^2e_\infty+e_0\in\mathcal{G}_{m+1,1}$ in CGA space. The codebook is defined by a set of vector $\boldsymbol{C}=\{\boldsymbol{c}_1,\cdots,\boldsymbol{c}_K\}$, $\boldsymbol{c}_k=\boldsymbol{s}_k+s_{k,\infty}e_\infty+s_{k,0}e_0\in\mathcal{G}_{m+1,1}$ and is calculated by minimizing the following problem,

$$\min_{u,\mathbf{c}}\sum_{k=1}^{K}\sum_{i=1}^{N}\sum_{t=1}^{T}u_{k,i,t}d(\boldsymbol{p}_{i,t},\boldsymbol{c}_k)$$
$$\text{s.t.}\quad\sum_{k=1}^{K}u_{k,i,t}=1\ ,\qquad u_{\ k,i,t}\in\{0,1\},$$

where $d(\boldsymbol{p}_{i,t},\boldsymbol{c}_k)=\left(\boldsymbol{v}_{i,t}\cdot\mathbf{s}_k-s_{k,\infty}-\frac{1}{2}\|\boldsymbol{v}_{i,1}\|^2s_{k,0}\right)^2$ is the squared distance between the point $\boldsymbol{p}_{i,t}$ and the k[th] codebook centroids $\boldsymbol{c}_k$ in CGA space.

CGA based clustering proceeds by alternating between two steps:

- Assignment step: Assign each observation to the cluster whose mean yields the least within-cluster sum of squared distance in CGA space;
- Update step: Calculate the new means to be the centroids of the observations in the new clusters.

The centroid $\boldsymbol{c}_k$ can be calculated by minimization the following L fuction:

$$L=\sum_{i=1}^{N}\sum_{t=1}^{T}u_{k,i,t}\left(\left(\boldsymbol{v}_{i,t}\cdot\mathbf{s}_k-s_{k,\infty}-\frac{1}{2}\|\boldsymbol{v}_{i,1}\|^2s_{k,0}\right)^2-\lambda(\|\mathbf{s}_k\|^2-1)\right)$$

The CGA clustering algorithm to calculate the codebook $\boldsymbol{C}$ is described as following.

```
algorithm cgaclustering_codebook()
input:
  v[N][T_N]: training set
  K: number of centroids
output:
  u[N][T_N]: memberships
  C[K]: array of codebook centroids in CGA
means
Begin
  δ ← 1
  while (δ>0)
    δ ← 0
    for i from 0 to N-1 do
      for t from 0 to T(i)-1 do
        dmin ← ∞
        n ← 0
        for k from 0 to K-1 do
```

```
          d ← |V[i][t] - C[k]|
          if d<dmin then
             dmin ← d
             n ← k
          endif
        endfor
        if u[i][t] ≠ n then
           δ ← δ + 1
           u[i][t] ← n
        endif
      endfor
    endfor

    for k from 0 to K-1 do
      C[k] ← argminL(k)
    endfor
  endwhile
end
```

4. Experiment Result

The results are achieved by conducting feature selection, discretization data and experiments of HAR using the CMU graphics lab motion capture database

4.1. CMU graphics lab motion capture database

CMU graphics lab motion capture database [15] includes the data set made by a Vicon motion capture system consisting of 12 infrared MX-40 cameras, each of which is capable of recording at 120 Hz with images of 4 megapixel resolution. Motions are captured in a working volume of approximately 3m x 8m. The capture subject wears 41 markers and a stylish black garment. Vicon software will create two data files: ASF file and AMC file.

- In the ASF file (Acclaim Skeleton File), a base pose is defined for the skeleton that is the starting point for the motion data. ASF has information: length, direction, local coordinate frame, number of Dofs, joint limits and hierarchy, connections of the bone.
- The AMC file (Acclaim Motion Capture) contains the motion data for a skeleton defined by an ASF file. The motion data is given a sample at a time. Each sample consists of a number of lines, a segment per line, containing the data.

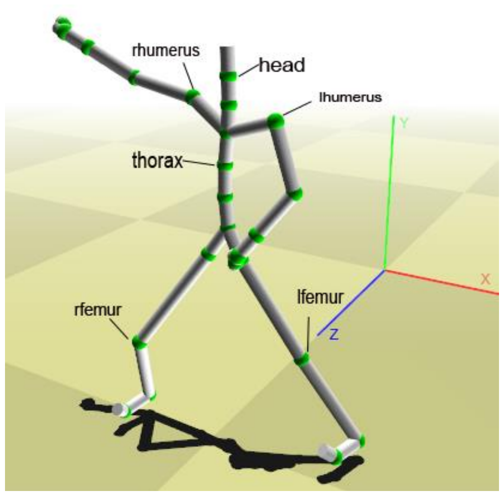

Figure 3. *Name list of 6 segments.*

In this experiment, we defined a sequence of m-dimensional continuous vectors $\boldsymbol{O}$ by using the direction of 6 segments. The name list of 6 segments shows as Figure. 3.

4.2. Experiment result

In this section, we demonstrate the performance of our proposed method by using running data and walking data downloaded from the CMU graphics lab motion capture database website. We compare our proposed method with k-means based VQ. Figure. 3 shows the training model and recognition using CGA Clustering based VQ approach for DHMM human activity recognition.

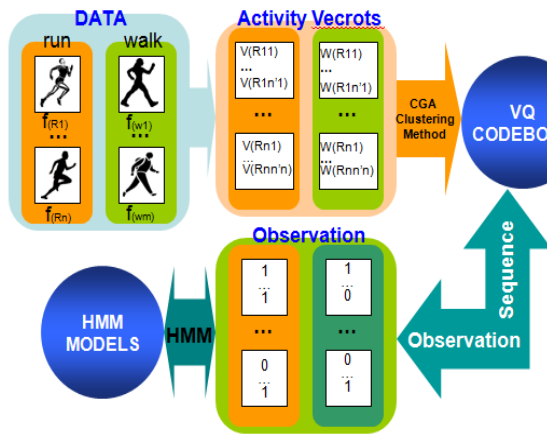

Figure 4. *Discrete data model*

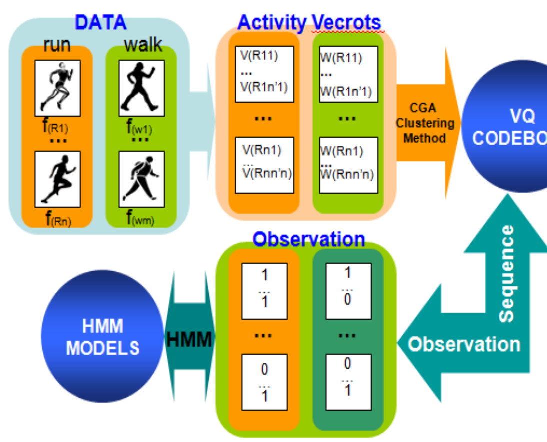

Figure 4. *Discrete data model*

For training and testing the model, we have a data set with 46 trials of the human running action and 131 trials of the human walking action. We separated the data in two equal parts by randomly. The first part is used for training and the second is used for testing.

Table 1. *Comparison results between K-means based VQ and CGA clustering based VQ.*

Class num	Frames per second	K-means (%)	CGA clustering (%)
2	24	55 (%)	**86 (%)**
3	24	71 (%)	**91 (%)**
4	24	53 (%)	**94 (%)**
5	24	49 (%)	**94 (%)**
2	12	47 (%)	26 (%)

3	12	60 (%)	**82 (%)**
4	12	54 (%)	**77 (%)**
5	12	68 (%)	63 (%)

The Table. 1 shows that the accuracy of human activity recognition. The result shows that the proposed method using CGA clustering based VQ was better than *k*-means based VQ. The accuracy of recognition using proposed method was the highest (94%) in the case of class number is 4 and using the data with 24 frames per second.

5. Conclusions

This paper presented the basic of a VQ for DHMMs. This paper also summarised a k-means based VQ and reviewed DHMMs. Then, this paper proposed a new approach of CGA Clustering based VQ for HMM. The experiment result of human activity recognition using CMU graphics lab motion capture database showed that the proposed method is better than conventional *k*-means based VQ method.

From this result, we can use CGA clustering algorithm to instead of k-means algorithm in the field of speech recognition, action recognition of objects. At the same time, the results of research opens up a new research direction in recognition theory, automatic control systems [16] and a variety of systems that involve interactions between people and electronic devices such as human computer interfaces.

REFERENCES

[1] Hoang Le Uyen Thuc, Pham Van Tuan, and Shian-Ru Ke, "Comparison of human action recognitions in monocular videos using DTW and HMM",journals, *UD*, 2014.

[2] Lawrence R.Rabiner, "A Tatorial on Hidden Markov Models and Selected Applications in Speech Recognition," *Fellow*, IEEE.

[3] M. Debyeche, J.P Haton, and A. Houacine "A New Vector Quantization front-end Process for Discrete HMM Speech Recognition System", *International Science Index, Vol: 1*, No: 6, 2007.

[4] Chris Doran anhd Anthony Lasenby *"Physical Applications of Geometric Algebra"*.

[5] C. Doran and A. Lasenby, *Geometric Algebra for Physicists*, Camgridge University Press, 2003.

[6] L. Dorst, D. Fontijne, and S. Mann, *Geometric Algebra for Computer Science: An Object – oriented Approacg to Geometric (Morgan Kaufmann Series in Computer Graphics)*, 2007.

[7] I. Sekita, T. Kurita, and N. Otsu, Complex Autoregressive Model for Shape Recognition, *IEEE Trans. On Pattern Analysis and Machine Intelligence*, Vol. 14, No. 4, 1992.

[8] N. Matsui, T. Isokawa, H. Kusamichi, F. Peper, and H. Nishimura, Quaternion Neural Network wwith geometric operators, *Journal on Intelligent and Fuzzy Systems*, Volume 15, Numbers 3-4, pp 149-164, 2004.

[9] F. Lefevre, "Non parametric probability estimation for HMM-based automatic speech recognition", *Computer Speech and Language*, vol. 17, pp. 113–136, 2003.

[10] J. Makhoul, S. Roucos, and H. Gish, "Vector Quantization in Speech Coding", *Proceedings of IEEE*, vol. 73. No. 11. 1551-1558, 1985.

[11] JeffA.Bilmes *"A Gentle Tutorial of the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixtureand Hidden Markov Models"*, U.C.Berkeley TR-97-021 April1998.

[12] Q. Huo and C. Chan, "Contextual vector quantization for speech recognition with discrete hidden Markov model", *Pattern recognition*, vol. 28 no. 4, pp. 513–517, 1995.

[13] R.M. Gray, *" Vector Quantization*", IEEE ASSP Magazine, 1984

[14] Minh Tuan Pham and Kanta Tachibana,*"An Algorithm for Fuzzy Clustering based on Conformal Geometric Algebra"*.

[15] The Carnegie Mellon University Motion Capture Database.

[16] Eduardo Bayro-Corrochano, Luis Eduardo Falcon-Morales and Julio Zamora-Esquivel *"Visually Guided Robotics Using Conformal Geometric Computing"*, 2007.

(The Board of Editors received the paper on 26/10/2014, its review was completed on 30/10/2014)