NHẬN DẠNG CHUYỂN ĐỘNG QUAY DỰA TRÊN MÔ HÌNH MARKOV ẨN VÀ ĐẠI SỐ HÌNH HỌC BẢO GIÁC

ROTATION RECOGNITION BASED ON HIDDEN MARKOV MODEL AND CONFORMAL GEOMETRIC ALGEBRA

Nguyễn Năng Hùng Vân[1], **Phạm Minh Tuấn**[1], **Tachibana Kanta**[2]

[1]*Trường Đại học Bách khoa, Đại học Đà Nẵng; Email: nguyenvan@dut.udn.vn, pmtuan@dut.udn.vn*
[2]*Trường Đại học Kogakuin; Email: kanta@cc.kogakuin.ac.jp*

Tóm tắt – Ngày nay, các nhà nghiên cứu đã phát triển đại số hình học (Geometric Algebra - GA) để biểu diễn các đối tượng trong không gian 3 chiều (3D) một cách chính xác và hiệu quả. Vì vậy, GA có thể ứng dụng vào các lĩnh vực nhận dạng vật thể hay nhận dạng các hành vi của đối tượng 3 chiều. Trong bài báo này, tác giả đã đề xuất mô hình Markov ẩn kết hợp với mật độ xác suất trên không gian đại số hình học bảo giác (Conformal Geometric Algebra - CGA) nhằm tính toán xác suất của trạng thái ẩn đối với quá trình chuyển đổi trạng thái của cổ tay. Từ kết quả thực nghiệm, phương pháp đề xuất sử dụng CGA Gauss trong việc tính toán mật độ xác suất của trạng thái ẩn sẽ cho kết quả tốt hơn nhiều so với sử dụng hàm mật độ Gauss thông thường.

Từ khóa – đại số hình học; học máy; mô hình xác suất; mô hình Markov ẩn; mật độ Gauss.

Abstract – Nowadays, many researchers have developed mathematical tools of Geometric Algebrato representing objects in the 3D space accurately and effectively. So GA can be applied to the field of object recognition or identification of the behavior of 3D objects. In this paper, the authors propose a Hidden Markov Model combined with the probability density on the Conformal Geometric Algebra space to calculate the probability of hidden states for the transition state of the wrist. From the experimental results, the proposed method using the CGA Gaussian distribution to calculate the probability density of the hidden state is better than using the conventional Gaussian distribution.

Key words – geometric algebra, machine learning, probabilistic model, hidden Markov model, Gaussian distribution.

1. Giới thiệu

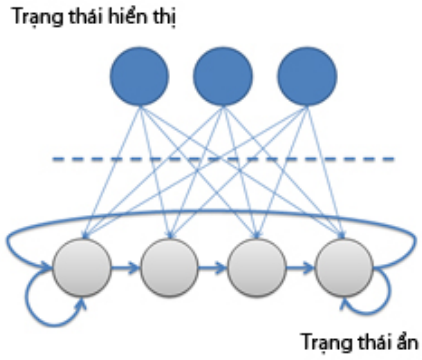
Ngày nay, có rất nhiều nghiên cứu liên quan đến nhận dạng vật thể hay nhận dạng hành vi trong không gian 3D [1]. Những nghiên cứu này, thường sử dụng mô hình vectơ trong không gian 3D [2], các vectơ thường độc lập với nhau nên không thể hiện được các mối liên kết hình học trong không gian. Một số nghiên cứu khác xây dựng mô hình liên kết các vectơ bằng các lý thuyết xác suất. Thông thường hàm Gauss được sử dụng rộng rãi trong tất cả các loại mô hình này. Tuy nhiên, hàm Gauss cũng không thể hiện hết tất cả các liên kết hình học trong không gian 3D như một phân bố cong nào đó.

Hiện nay, các nhà nghiên cứu đã phát triển một công cụ đại số hình học có khả năng biểu diễn các đối tượng trong không gian một cách chính xác và dễ dàng. GA hay còn gọi là đại số Cliford, là một mô hình toán học phát triển từ sự kết hợp giữa đại số và hình học [3][4][5]. GA có thể biểu diễn các vectơ hay các mối liên kết của chúng trong 3D một cách đơn giản và chính xác. Vì vậy, GA bắt đầu được các nhà nghiên cứu trong lĩnh vực công nghệ thông tin quan tâm tới. Có rất nhiều ứng dụng của GA như mô hình xử lý tính hiệu, xử lý ảnh sử dụng không gian GA số phức [6][7][8] hay quaternions [9][10][11]. Hơn nữa, mô hình xác suất sử dụng GA trong lĩnh vực học máy (Machine Learning - ML) là một lĩnh vực hoàn toàn mới và hiệu quả mang lại rất cao để ứng dụng vào các lĩnh vực nhận dạng 3D và nhận dạng các hành vi của đối tượng [12][13][14][15].

2. Phương pháp đề xuất

2.1. Mô hình Markov ẩn

Mô hình Markov [16] ẩn là một phương pháp sử dụng xác suất để mô hình hóa dữ liệu theo thời gian một cách có trình tự. Do đạt được độ chính xác cao và có khả năng thay đổi cấu trúc dễ dàng nên mô hình này ngày càng được sử dụng rộng rãi trong nhiều lĩnh vực, đặc biệt là trong lĩnh vực nhận dạng mẫu tiếng nói, hình ảnh và các đối tượng vật thể. Hình 1 là ví dụ cho mô hình HMM.



Hình 1: *Mô hình HMM*

HMM được xác định bởi mô hình xác suất:

$$\lambda = (\mathrm{N}, \mathrm{M}, \mathrm{A}, \mathrm{B}, \Pi)$$

Trong đó:

- N là số lượng trạng thái ẩn của mô hình, ký hiệu các trạng thái $\mathrm{S} = \{\mathrm{S}_1, \mathrm{S}_2 \ldots, \mathrm{S}_\mathrm{N}\}$ và trạng thái ở thời điểm t là q_t.
- M là số lượng tín hiệu có thể quan sát được, ký hiệu các tín hiệu quan sát $\mathrm{V} = \{\mathrm{v}_1, \mathrm{v}_2 \ldots, \mathrm{v}_\mathrm{M}\}$ là tín hiệu quan sát được ở thời điểm t là O_t.
- A là ma trận xác suất chuyển đổi trạng thái ẩn trong mô hình, $\mathrm{A} = \{\mathrm{a}_{\mathrm{i,j}}\}$ với: $\mathrm{a}_{\mathrm{i,j}} = \mathrm{p}(\mathrm{q}_{\mathrm{t}+1} = \mathrm{S}_\mathrm{j} | \mathrm{q}_\mathrm{t} = \mathrm{S}_\mathrm{i})$, $1 \leq \mathrm{i,j} \leq \mathrm{N}$ thỏa mãn điều kiện $\sum_{\mathrm{j}=1}^{\mathrm{N}} \mathrm{a}_{\mathrm{i,j}} = 1$.
- B là ma trận xác suất đầu ra của các trạng thái ẩn đối với các tín hiệu quan sát $\mathrm{B} = \{\mathrm{b}_\mathrm{j}(\mathrm{k})\}$ với $\mathrm{b}_\mathrm{j}(\mathrm{k}) = \mathrm{p}(\mathrm{v}_\mathrm{t} \text{ at } \mathrm{t} | \mathrm{q}_\mathrm{t} = \mathrm{S}_\mathrm{j})$, $1 \leq \mathrm{j} \leq \mathrm{N}, 1 \leq \mathrm{k} \leq \mathrm{M}$ thỏa mãn

điều kiện $\sum_{k=1}^{M} b_j(k) = 1$. Với $b_j(k)$ là xác suất đầu ra của trạng thái ẩn j đối với tín hiệu quan sát k. Trong mô hình xác suất liên tục thì B chính là hàm mật độ xác suất của các trạng thái.

- π là xác suất khởi đầu của mỗi trạng thái $\pi = \{\pi_i\}$ với $\pi_i = p(q_1, = S_i), 1 \leq i \leq N$ thỏa mãn điều kiện $\sum_{i=1}^{M} \pi_j = 1$.

Có 3 bài toán kinh điển áp dụng HMM vào các ứng dụng phức tạp trong thực tế sau:

Bài toán 1. Cho trước chuỗi tín hiệu quan sát được ở thời điểm t là $O = O_1, O_2, \ldots, O_t$ và mô hình HMM đại diện bởi bộ tham số $\lambda = (A, B, \pi)$. Làm sao để tính toán một cách hiệu quả $p(O|\lambda)$ và xác suất phát sinh O từ mô hình λ.

Bài toán 2. Cho trước chuỗi tín hiệu quan sát được ở thời điểm t là $O = O_1, O_2, \ldots, O_t$ và mô hình HMM đại diện bởi bộ tham số $\lambda = (A, B, \pi)$. Cần tìm ra chuỗi trạng thái tối ưu nhất $Q = q_1, q_2, \ldots, q_t$ đã phát sinh ra O? Trong đó q_t là số thứ tự của các trạng thái ẩn.

Bài toán 3. Cho trước chuỗi tín hiệu quan sát được ở thời điểm t là $O = O_1, O_2, \ldots, O_t$. Làm thế nào để xác định các tham số mô hình $\lambda = (A, B, \pi)$ sao cho cực đại hóa xác suất $p(O|\lambda)$. Đây chính là bài toán huấn luyện mô hình.

Bài báo này, tập trung nghiên cứu bài toán 3 trong mô hình HMM với xác suất đầu ra của trạng thái ẩn đối với tín hiệu quan sát là liên tục [17]. Nội dung chính của bài báo là đề xuất mật độ xác suất trên không gian CGA thay cho hàm Gauss trên không gian thực nhằm tăng độ chính xác của xác suất đầu ra trong các trạng thái ẩn đối với các tín hiệu quan sát.

2.2. Đại số hình học bảo giác

Đại số hình học bảo giác (CGA) [3] là một phần của GA. GA định nghĩa không gian bằng cách định nghĩa p+q vectơ cơ sở vuông góc nhau $O = \{e_1, \ldots, e_p, e_{p+1}, \ldots, e_{p+q}\}$, trong đó ${e_i}^2 = +1, \forall i \in \{1, \ldots, p\}$ và ${e_i}^2 = -1, \forall i \in \{p+1, \ldots, q\}$. Ở đây, ta ký hiệu không gian được định nghĩa bởi O là $G_{p,q}$. Như vậy không gian thực m chiều $\Re^m$ có thể được biểu diễn bởi $G_{m,0}$ trong GA.

Không gian CGA mở rộng từ không gian thực $\Re^m$ với việc thêm 2 vectơ cơ sở. Do đó, một không gian CGA được xác định bởi m + 2 vectơ cơ sở $\{e_1, \ldots, e_m, e_+, e_-\}$, trong đó e_+ và e_- được định nghĩa như sau:

$$\begin{aligned} {e_+}^2 &= e_+.e_+ = 1, \\ {e_-}^2 &= e_-.e_- = -1, \\ e_+.e_- &= e_+.e_i = e_-.e_i = 0, \forall i\{1, \ldots, m\}. \end{aligned} \tag{1}$$

Do đó, một không gian CGA có thể được ký hiệu bởi $G_{m+1,1}$. Ta định nghĩa thêm 2 vectơ cơ sở sau:

$$e_0 = 1/2(e_- - e_+), e_\infty = (e_- + e_+). \tag{2}$$

Từ (1) và (2) chúng ta có

$$\begin{aligned} e_0.e_0 &= e_\infty.e_\infty = 0, \\ e_0.e_\infty &= e_\infty.e_0 = 0, \\ e_0.e_i &= e_\infty.e_i = 0, \forall i \in \{1, \ldots, m\}. \end{aligned} \tag{3}$$

Theo đề xuất của Hestenes [4], vectơ thực $x = \sum_i^m x_i e_i \in \Re^m$ có thể biểu diễn bởi một điểm $P \in G_{m+1,1}$ trên không gian CGA như sau:

$$P = x + \frac{1}{2}||x||^2 e_\infty + e_0 \tag{4}$$

Một hình cầu được biểu diễn như một vectơ bảo giác (conformal vector) trong không gian CGA:

$$S = P - 1/2r^2 e_\infty = x + 1/2\{||x||^2 - r^2\}e_\infty + e_0, \tag{5}$$

Ở đây S chính là biểu diễn của một mặt cầu có tâm là x, bán kính r trong không gian thực $\Re^m$. Khi nội tích (inner product) $S.Q = 0$, Q sẽ là điểm nằm trên mặt cầu S. Từ (4) và (5), chúng ta nhận thấy rằng một điểm bất kỳ là một hình cầu với bán kính $r = 0$ trong không gian CGA $G_{m+1,1}$.

Không gian CGA cũng biểu diễn mặt phẳng như một vectơ bảo giác như sau:

$$L = n + de_\infty \tag{6}$$

Trong đó n là một vectơ, d là một hệ số vô hướng của mặt phẳng $n.x + d = 0$ trong không gian thực $\Re^m$. Nếu nội tích $L.Q = 0$ thì khi đó Q là một điểm nằm trên mặt phẳng L.

Một vectơ bảo giác S trong $G_{n+1,1}$ được viết dưới dạng tổng quát:

$$S = s + s_\infty e_\infty + s_0 e_0. \tag{7}$$

$s = \sum_i^m s_i e_i$ là một vectơ trong không gian thực $\Re^m$. s_∞ và s_0 là tham số của các vectơ cơ sở e_∞ và e_0.

Trong bài báo này, chúng tôi sử dụng nội tích giữa điểm P và một vectơ bảo giác S như là khoảng cách giữa chúng trong không gian CGA. Từ (3), (4) và (7), nội tích giữa P và Q là:

$$\begin{aligned} d(P, S) \propto P.Q &= (x + 1/2||x||^2 e_\infty + e_0) \cdot (s + s_\infty e_\infty + s_0 e_0) \\ &= x.s - s_\infty - 1/2||x||^2 s_0 \end{aligned} \tag{8}$$

2.3. Hàm mật độ xác suất của trạng thái trong mô hình HMM

2.3.1. Mật độ Gauss

Mật độ Gauss là một phân bố xác suất cực kỳ quan trọng trong nhiều lĩnh vực. Đối với mô hình HMM thì mật độ xác suất của trạng thái i được xác định bởi:

$$\begin{aligned} b_i(x) &= N\left(x; \mu; \sum\right) \\ &= \frac{1}{(2\pi)^{\frac{d}{2}} ||\sum_i||^{\frac{1}{2}}} \exp\left(-\frac{(x - \mu_i)^T (x - \mu_i)}{2\sum_i}\right) \end{aligned} \tag{9}$$

Trong đó:

- x là một vector quan sát.
- μ_i là kỳ vọng toán học của x.
- $\sum_i$ là ma trận hiệp phương sai.
- d là số chiều trong không gian.

2.3.2. Mật độ CGA Gauss:

Mật độ CGA Gauss được xây dựng trên cự ly giữa một điểm P và một vectơ bảo giác S_l trên không gian CGA kết hợp với mật độ Gauss. Hàm mật độ CGA Gauss [13] sẽ là:

$$b_i(x) = \prod_{l}^{m} b_i(x|\lambda_l) \quad (10)$$

$$b_i(x|\lambda_l) = \frac{1}{\sqrt{2\pi\lambda_l}} \exp\left(-\frac{d^2(P, S_l)}{2\lambda_l}\right) \quad (11)$$

- x là là một vectơ quan sát.
- S_l là vectơ bảo giác riêng (eigen conformal vector) thứ l trong không gian CGA.
- P là 1 điểm được xác định bởi:

$$P = x + 1/2||x||^2 e_\infty + e_0.$$

- Λ_l là phương sai hay giá trị riêng (eigen value) thứ l của quan sát x trong không gian CGA.

3. Kết quả nghiên cứu

3.1. So sánh hàm mật độ Gauss và CGA Gauss với dữ liệu phân bố trên đường cong

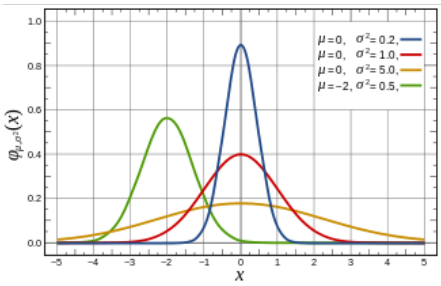
- Hàm mật độ Gauss trong không gian 1 chiều:

$$f(x; \mu; \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (12)$$

Nếu biến ngẫy nhiên x có phân phối này, $\mu = 0$ và $\sigma = 1$, phân phối được gọi là chuẩn và hàm mật độ rút gọn thành:

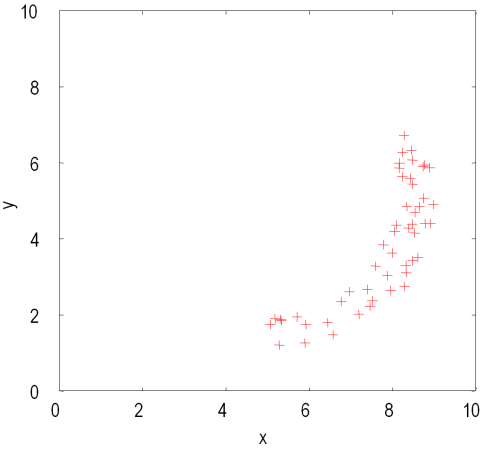
$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \quad (13)$$

Hình 2 biểu diễn hàm mật độ xác suất Gauss với các tham số khác nhau. Dễ dàng nhận thấy rằng mật độ Gauss có hình dạng "núi" nên chỉ có thể xấp xỉ những dữ liệu gom cụm với nhau. Tức là ở gần tâm của dữ liệu thì có phân bố dày và ở xa tâm thì sẽ có phân bố thưa hơn. Đối với những phân bố phức tạp như hình cung Hình 3 thì sẽ khó sử dụng mật độ Gauss để xấp xỉ.

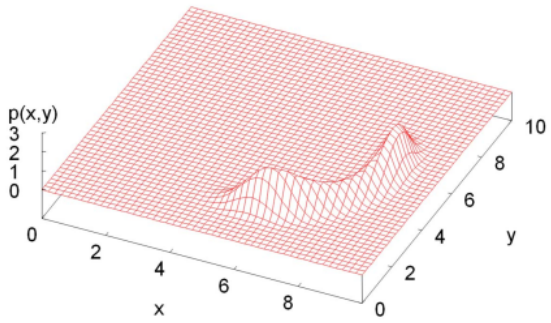


Hình 2: *Hàm mật độ Gauss*

Sử dụng công thức (10), ta có thể tính mật độ xác suất của dữ liệu trong không gian CGA. Hình 4 là kết quả của việc tìm mật độ xác suất CGA Gauss cho dữ liệu được phân bố ở Hình 3.



Hình 3: *Phân bố dữ liệu trên hình cung*

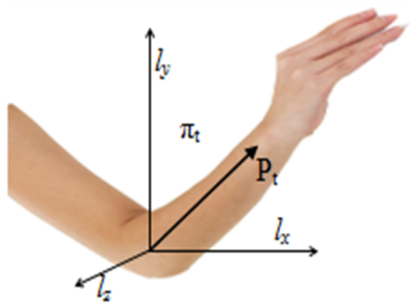


Hình 4: *Mật độ xác suất hình cung*

Từ kết quả trên, ta có thể nhận thấy được rằng việc sử dụng hàm mật độ CGA Gauss sẽ tốt hơn với mật độ Gauss trong các trường hợp dữ liệu có phân bố trên các mặt cong. Ví dụ như các bộ phận của cơ thể người khi quay quanh một khớp nào đó sẽ phân bố trên một mặt cầu trong không gian 3D nên sử dụng hàm mật độ CGA Gauss là thích hợp nhất.

3.2. So sánh hàm mật độ Gauss và CGA Gauss trên mô hình Markov ẩn

Giả sử có một cánh tay được quay trong không gian 3 chiều với một góc quay tự do. Thời gian quay của mỗi chiều quay là $T = 50$ đơn vị thời gian. Số lần thay đổi chiều quay là $N = 2$ và 3 loại góc quay ứng với từng đơn vị thời gian là $\alpha = \{0.05, 0.1, 0.2\}$ (rad). Báo cáo tiến trình thực nghiệm trên mô hình HMM với xác suất quan sát được dựa theo hàm Gauss và hàm CGA Gauss bằng cách huấn luyện cho mô hình HMM dựa trên các quan sát của cổ tay. Ở đây cho trạng thái ẩn bằng số lần thay đổi chiều quay.



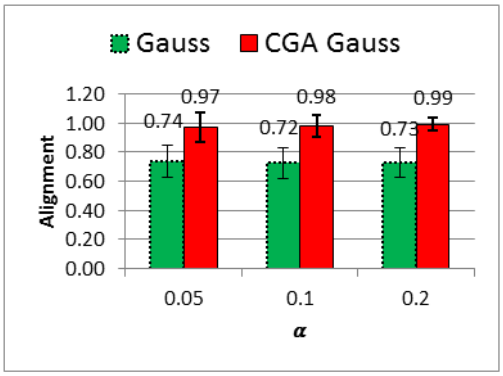
Hình 5: *Mô phỏng các chuyển động cánh tay*

Tiếp theo, báo cáo quan sát quá trình chuyển đổi trạng thái ẩn của 2 mô hình HMM đã huấn luyện khi thực hiện việc quay cánh tay. Quá trình này được thực hiện 100 lần và mỗi lần các chiều quay được thay đổi một cách ngẫu nhiên. Hàm liên kết (alignment fuction) được sử dụng để đánh giá mức độ giống nhau giữa trạng thái quay trên thực tế của cổ tay và trạng thái ẩn của mô hình HMM:

$$A(Y_t, Y_e) = \frac{\sum_{kl} y_{kl;t} y_{kl;e}}{\sqrt{\sum_{kl} y^2_{kl;t}} \sqrt{\sum_{kl} y^2_{kl;e}}} \in [0, 1]$$

Trong đó, $Y_t = [y_{kl;t}]$ và $Y_e = [y_{kl;e}]$ là ma trận trạng thái thực và trạng thái ẩn của mô hình HMM. Với y_{kl} được định nghĩa như sau:

$$y_{kl} = \begin{cases} 1 & (y_k = y_l) \\ 0 & (y_k \neq y_l) \end{cases}$$



Hình 6: *Trung bình và độ lệch chuẩn của liên kết trong 2 mô hình HMM*

Hình 6 là kết quả thực nghiệm quan sát quá trình chuyển đổi trạng thái của cổ tay. Từ kết quả thực nghiệm, dễ dàng nhận thấy sử dụng CGA Gauss để tính toán mật độ xác suất của trạng thái ẩn sẽ cho kết quả tốt hơn nhiều so với sử dụng hàm mật độ Gauss thông dụng. Từ đó có thể kết luận rằng phương pháp đề xuất là một mô hình hữu hiệu trong việc nhận dạng hành vi đối tượng trong không gian 3D.

4. Kết luận

Bài báo này đề xuất một phương pháp nghiên cứu mới nhận dạng các hành vi và các đối tượng 3D sử dụng mô hình xác suất GA và mô hình Markov ẩn. Ưu điểm chính của phương pháp này là biểu diễn các đối tượng trong không gian rất chính xác, đặc biệt là trong việc nhận biết các hành vi. Kết quả nghiên cứu có ý nghĩa khoa học và xã hội cao, góp phần mở ra hướng nghiên cứu mới, nhận dạng các hành vi và các đối tượng 3D, hướng đến xây dựng một hệ thống nhận dạng các hành vi con người.

Tài liệu tham khảo

[1] James D. Foley, Andries van Dam, Steven K. Feiner, John F. Hughes, *Computer graphics: Principles and practice (2nd ed.)*, Addision – Wesley Longman Plublishing Co., Inc., Boston, MA, 1990.

[2] Eberlt, D.H. *3D Game Engine Design: A Practical Approach to Real-Time Computer Graphics*, Morgan Kaufmann Publishers. 2001.

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

[4] D. Hestenes, *New foundations for classical mechanics*, Dordrecht, 1996.

[5] 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.

[6] 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.

[7] A. Hirose, Complex – Valued Neural Network: Theories and Applications, *Series on Innovative Intelligence*, Vol. 5, 2006.

[8] T. Nitta, An Extension of the Back – Propagation Algorithm to Complex Numbers, *Neural Network* , Volume 10, Number 8, pp. 1391 – 1415 (25), November 1997.

[9] 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.

[10] S. Buchholz and N. LeBihan, Optimal separation of polarized signals by quaternionic Neural Network, *14th European Signal Processing Conference*, OUSIPCO 2006, Stember 4-8, Florence Italya, 2006.

[11] M.Jordan, J.Kleinberg, B.Scholkopf, *InformationScienceandStatistics*, 2006.

[12] D. Hestenesand G. Sobczyk. *Clifford Algebra to Geometric Calculus: Aunified language formathematics and physics*, Reidel, 1984.

[13] P. M. Tuấn, A clustering method for geometric data based on approximation using conformal geometric algebra, *2011 IEEE International Conference on Fuzzy Systems*, pp. 2540 – 2545, 2011.

[14] Hildenbrand D., Fontijne D., Perwass Ch., Dorst L., Geometric Algebra an its Application to Computer Graphics, T*otorial notes of the EUROGRAPHICS conference*, 2004, Grenoble.

[15] E.M.S. Hitzer Ecilidean Geometric Objects in the Cliford Geometric Algebra of Origin, 3-space, *Infinity Bulletin of the Belgian Mathematical Society – Simon Stevin*, 2004.

[16] Kristie Seymore, Andrew McCallum, and Roni Rosenfeld.Learning Hidden Markov Model Structure for Information Extraction, *AAAI 99 Workshop on Machine Learning for Information Extraction*, 1999.

[17] Lawrence R. Rabiner, Fellow, IEEE, *A Tutorial on Hidden Markov Models and Selected Applicatios a Speech Recognition*, 1989

(BBT nhận bài: 22/12/2013, phản biện xong: 29/12/2013)