www.sensorsportal.com

Sensors & Transducers

http://www.sensorsportal.com

Research on the Covering of the Biomimetic Pattern Recognition of the High Dimensional Information Geometry

Youzheng ZHANG
Quzhou College of Technology, Quzhou, Zhejiang, 324000, China
Tel.: +8613957024768, fax: +865708068336
E-mail: youzhengzhang@yeah.net

Received: 15 April 2013 /Accepted: 20 July 2013 /Published: 30 July 2013

Abstract: Objective: To research on the application of high dimensional space theory to the biomimetic pattern recognition. Procedures: The sample was constructed into a k-dimensional simplex in a high dimensional space and 0.85 times of the average distance between the vertices was chosen as the threshold value thus a convex cell body covering the k-dimensional simplex was constructed. By determining whether the sample point could be covered by the convex cell body derived from a certain sample group, the sample was recognized and classified. Methods: The orthogonal complementary space of the subspace was constructed and the Euclidean distance between the point and the subspace was calculated and the distance from the point to the k-dimensional simplex was further calculated. Results: The study solved the problem of whether the sample point is covered by a certain convex cell body and its correctness was testified via human-face recognition. *Copyright © 2013 IFSA.*

Keywords: High dimension space, Face recognition, Biomimetic pattern recognition, k-dimensional simplex, Convex cell body cover.

1. Introduction

A gray image is mapped a point in $R^{m\times n}$ - Space. The traditional pattern recognition is to find an optimal partition in $R^{m\times n}$ - Space. Now commonly method used by Support Vector Machine (SVM) is through kernel function mapped $R^{m\times n}$ to R^k , and to find the best division in the mapping R^k - Space [1]. These methods were identified from non the same type things object "difference", the realization of algorithms, are focused on different things, "different", that is, distinguish a class samples from limited known samples of class. This form of human cognition of things is very different: people know things are a class of a class of understanding and attach importance to the links between similar things, or, one sample of unknown samples with the unlimited class distinction. Of "difference" as the starting point of traditional pattern recognition, inevitably result in the following two limitations: first, encountered new things for the first time that have not learned, easily mistaken for a certain type have been learned old things; Second, to learn a new things, they tend to upset the old knowledge, that undermine the already studied the identification of old things. This is the traditional pattern recognition theory in practice is difficult to make real the reason for desired effect.

Academician Wang Shoujue from the perspective of human cognition, re-examine the neural network pattern recognition, innovation proposed new pattern recognition – Biomimetic Pattern Recognition (BPR) [2], based on the multi-dimensional manifold topology theory. Emphasize "Cognition", use of "High dimensional space non hypersphere complex geometry coverage" for pattern recognition, shown in Fig. 1. High dimensional geometry space is a vehicle to achieve Biomimetic Pattern Recognition [3]. High dimensional geometry space theory and calculation process is from the concept of high-dimensional space geometry, using the set of points in the space to describe the subspace, Combination of simple operation by means of geometry, high dimensional space complex calculations [4].

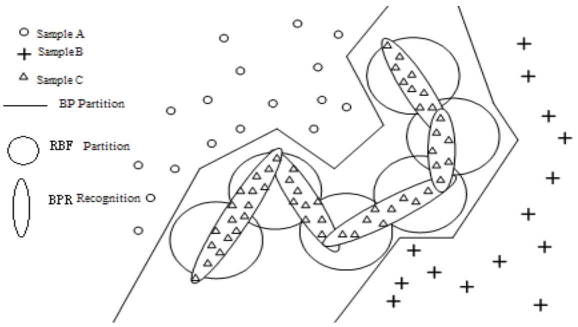


Fig. 1. BPR.

But research on the covering of the biomimetic pattern recognition of the high dimensional information geometry is still poor [5-8], In this paper through high-dimensional information geometry biomimetic pattern recognition, proposed using high-dimensional information geometry to construct a convex cell covering of biomimetic pattern recognition, judging if some high-dimensional space samples be covered by a convex cell body, the average distance of the k-dimensional simplex vertices, as the convex cell body threshold reference value, in order to build a more satisfactory coverage of the body, Finally, face recognition test the feasibility of the method.

2. High-dimensional Geometry Theory

Each element in R^n ($n \geq 2$) is called a point, and the establishment of the following concepts [6]:

Definition 1 If k vectors $\mathbf{a}_1 - \mathbf{a}_0$, $\mathbf{a}_2 - \mathbf{a}_0$, $\cdots$, $\mathbf{a}_k - \mathbf{a}_0$ in the vector space R^n are linearly independent, then the point set { $\mathbf{a}_0 , \mathbf{a}_1 , \cdots , \mathbf{a}_k$ } is called affine independent.

Definition 2 If k vectors $\mathbf{a}_1 - \mathbf{a}_0$, $\mathbf{a}_2 - \mathbf{a}_0$, $\cdots$, $\mathbf{a}_k - \mathbf{a}_0$ in the vector space R^n are linearly independent, that

$$S = \left\{ \mathbf{x} \mid \mathbf{x} = \sum_{i=1}^{k} \lambda_i (\mathbf{a}_i - \mathbf{a}_0), \lambda_i \in R \right\}$$

Is linear subspace in R^n , the base for the subspace is { $\mathbf{a}_1 - \mathbf{a}_0$, $\mathbf{a}_2 - \mathbf{a}_0$, $\cdots$, $\mathbf{a}_k - \mathbf{a}_0$ } $k = \dim S$, abbreviated

$$S = \langle\langle \mathbf{a}_1 - \mathbf{a}_0, \cdots, \mathbf{a}_k - \mathbf{a}_0 \rangle\rangle$$

Definition 3 Let $B = \{\mathbf{a}_0, \mathbf{a}_1, \cdots, \mathbf{a}_k\}$ is affine independent, record set

$$\Delta \mathbf{a}_0 \mathbf{a}_1 \cdots \mathbf{a}_k = \left\{ \sum_{i=0}^{k} \lambda_i a_i \mid \lambda_i \geq 0, 0 \leq i \leq k, \sum_{i=0}^{k} \lambda_i = 1 \right\} \tag{1}$$

to k-dimensional simplex, abbreviated as k- simplex.

Point to line segment distance: at the two cases, First: the pedal of point to the line segment AB in AB , then the distance shall be from point to pedal. Second: the pedal of point to the line segment AB out AB , the distance shall be from point to the endpoint that near the pedal [3] (Fig. 2).

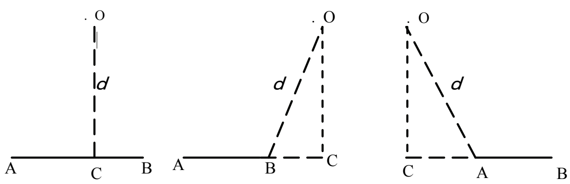


Fig. 2. Points to the distance line segment.

Point to the limited plane distance: the distance is similar from point to line segment, the pedal of point to the plane if in the limited plane; its distance is the distance from point to pedal; if outside of a limited plane, then the distance of point to plane distance of the minimum value of all points. If the limited plane is convex, then this distance must be a point to end point or a plane the distance between the boundary points [9].

The actual pattern recognition in order to determine whether a particular sample points are to be identified A class, A as a set, then the sample points in set A ($A \subset R^d$). The distribution of the concrete application of pattern recognition based object can be different dimensions of "manifolds"; it is necessary building a collection Multidimensional geometric shapes to cover A in feature space R^d.

Definition 4 Suppose V is a k-simplex in the feature space R^d, the shortest distance from point x to V be met:

$$d(x, V) = \min\left(d(x, y), \forall y \in V\right),$$

if U met:

$$U = \left\{ x \mid x \in d(x, V) < \theta, x \in R^d / V \right\}$$

constant $\theta > 0$, claimed U a k-simplex coverage is actually a cover of convex cell body.

BPR algorithm that: the sample points x to be recognized to a k- simplex is less than the distance threshold θ, that x belonging to one k-simplex coverage, then x belong to the k-simplex with the class.

A gray image, mapped a point of R^n Space, through feature extraction, will points in R^n Space projected into the feature R^d Space $(d << n)$, then $k+1$ images $a_0, a_1, \cdots, a_k$, projected on to R^d Space, as $k+1$ points, if it is still recorded as a_0, a_1 $\cdots$, a_k, then the vector $\mathbf{a}_1 - \mathbf{a}_0$, $\mathbf{a}_2 - \mathbf{a}_0$, $\cdots$, $\mathbf{a}_k - \mathbf{a}_0$ are linearly independent, generate k-dimensional linear subspace, denoted S ,on the other hand, the set

$$\mathbf{A} = \{\sum_{i=1}^{k} \lambda_{i-1}(\mathbf{a}_i - \mathbf{a}_0) \mid \lambda_i \geq 0, \sum_{i=1}^{k} \lambda_i = 1\} \tag{2}$$

constitute k-simplex, apparently $\mathbf{A} \in S$.

There were m classes, each class to take $k+1$ training samples, then $k+1$ images in the feature R^d space composition the k-simplex in its generated k-dimensional linear subspace, there were m linear subspaces and k-simplexes. Now, if there were l same type images to be recognized, these images in the feature R^d space generated *l-1* -dimensional linear subspace of the *l-1* -simplex. Identification problem is to judge the distance that *l-1* -simplex which in *l-1* -dimensional linear subspace and m patterns k-simplexes which in *k*-dimensional subspace, if the distance with a k-simplexs less than a given threshold θ , then the sample would fall into the class [10].

3. Distance Between Point and k-simplex

3.1. Distance Between Point and Line Segment (Fig. 2)

a) if $\langle \overrightarrow{AO}, \overrightarrow{AB} \rangle \cdot \langle \overrightarrow{BO}, \overrightarrow{BA} \rangle \geq 0$, then

$$d = \left\| \overrightarrow{AO} \right\| \sqrt{1 - \left(\frac{\langle \overrightarrow{AO}, \overrightarrow{AB} \rangle}{\left\| \overrightarrow{AB} \right\|} \right)^2} \tag{3}$$

b) if $\langle \overrightarrow{AO}, \overrightarrow{AB} \rangle < 0$, then

$$d = \left\| \overrightarrow{AO} \right\|, \text{ or } d = \left\| \overrightarrow{BO} \right\|$$

Here the operator $< \mathbf{a}, \mathbf{b} >$ was the inner product of the vectors.

3.2. Distance Between the Point x_0 and k- simplex $\Delta \mathbf{a}_0 \mathbf{a}_1 \cdots \mathbf{a}_k$

Firstly here should discuss the distance between the point x_0 and the k-dimension plane. The literature [6] had proposed the basic results of this topic; here the Theorem 1 and Theorem 2 had perfected it.

The k-dimension plane

$$S^k : \mathbf{a}_0 + \langle\langle \mathbf{a}_1 - \mathbf{a}_0, \cdots, \mathbf{a}_k - \mathbf{a}_0 \rangle\rangle$$

and the vector noted

$$\mathbf{v}_i = \mathbf{a}_i - \mathbf{a}_0, 1 \leq i \leq k \text{, let } \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$$

were linearly independent and $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ were the basis of the k-dimension sub-space $V = \langle\langle \mathbf{v}_1, \cdots, \mathbf{v}_k \rangle\rangle$, so that

$$S^k = \mathbf{a}_0 + V, \dim V = k \text{, where } 1 \leq k \leq n-1$$

Supposing $\{\mathbf{v}_{k+1}, \cdots, \mathbf{v}_n\}$ was the basis of $V^{\perp}$, then S^k could be expressed by $\langle \mathbf{x}, \mathbf{v}_j \rangle = b_j$, where $b_j = \langle \mathbf{a}_0, \mathbf{v}_j \rangle$ for $k+1 \leq j \leq n$, and point $\mathbf{x}_0 \in R^n \setminus S^k$, the k-dimension plane $S^k = \mathbf{a}_0 + V$, the distance between $\mathbf{x}_0$ and S^k was

$$d(x_0, S^k) = \min_{\mathbf{x} \in S^k} |\mathbf{x}_0 - \mathbf{x}| = \min_{\mathbf{v} \in V} |(\mathbf{x}_0 - \mathbf{a}_0) - \mathbf{v}|$$
$$= d(\mathbf{x}_0 - \mathbf{a}_0, V)$$

Geometrically, finding the distance between $\mathbf{x}_0$ and S^k was same as to find projection of the orthocomplement of the vector $\mathbf{x}_0 - \mathbf{a}_0$ to V, which denoted by $V^{\perp}$.

Theorem 1. Let $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ were n-dimension vectors, $k < n$ and linearly independent, let $A = [\mathbf{v}_1, \cdots, \mathbf{v}_k]^T_{k \times n}$, then $r(AA^T) = r(A) = k$, and $Q = A^T(AA^T)^{-1}$ was the generalized right inverse matrix of matrix A

Proof: Let $x \in Ker(AA^T)$, then $AA^T x^T = 0$ $xAA^T x^T = 0$, $(xA)(Ax)^T = 0$, $\|xA\| = 0$, i.e. $xA = 0$, $x \in Ker(A)$, so $Ker(AA^T) \in Ker(A)$, while

$$r(A) = n - \dim(Ker(A))$$
$$= n - \dim(Ker(AA^T)) = r(AA^T) = k,$$

this indicated that the matrix AA^T was reversible.

Moreover, $AQ = AA^T(AA^T)^{-1} = I_{k\times k}$.

Theorem 2. Let $P = A^T(AA^T)^{-1}A$, then the vector-group $P-I$ were orthogonal to the vector-group A.

Proof: for

$$AP = A\cdot[A^T(AA^T)^{-1}A] = A,$$
$$A(P-I) = 0,$$

Let $P-I = [\mathbf{u_1},\cdots,\mathbf{u_n}]^T$, for $(P-I)^T = P-I$

So $A(P-I) = [\mathbf{v}_1,\cdots,\mathbf{v}_k]^T\cdot[\mathbf{u}_1^T\cdots\mathbf{u}_n^T]_{n\times n} = \mathbf{O}$, so the vector-group $P-I$ were orthogonal to the vector-group A, and $(P-I)\subset Ker(A)$. In other words, that when $\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v}_k$ was the basis of $V = \langle\langle\mathbf{v}_1,\cdots,\mathbf{v}_k\rangle\rangle$.

The sub-space spanked by the vector-group $P-I$ was the orthocomplement of V noted by $V^\perp$, and one group basis of them was the basis of orthocomplement $V^\perp$.

If $\{\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v}_k\}$ was the orthogonal basis of V, than

$$P = \left(\frac{\mathbf{v}_1^T}{\|\mathbf{v}_1\|}\cdots\frac{\mathbf{v}_k^T}{\|\mathbf{v}_k\|}\right)\cdot\left[\frac{\mathbf{v}_1}{\|\mathbf{v}_1\|},\cdots,\frac{\mathbf{v}_k}{\|\mathbf{v}_k\|}\right]^T,$$

else Orthogonal normalized vector group $[\mathbf{u_1},\cdots,\mathbf{u_n}]$. Let vector

$$\boldsymbol{\beta} = \left(\frac{\langle\mathbf{x}_0-\mathbf{a}_0,\mathbf{u}_1^\mathbf{T}\rangle}{\|\mathbf{u}_1^\mathbf{T}\|},\frac{\langle\mathbf{x}_0-\mathbf{a}_0,\mathbf{u}_2^\mathbf{T}\rangle}{\|\mathbf{u}_2^\mathbf{T}\|},\cdots,\frac{\langle\mathbf{x}_0-\mathbf{a}_0,\mathbf{u}_n^\mathbf{T}\rangle}{\|\mathbf{u}_n^\mathbf{T}\|}\right),$$

where $if\ \ \mathbf{u}_j^\mathbf{T} = \mathbf{0}, then\ \frac{\langle\mathbf{x}_0-\mathbf{a}_0,\mathbf{u}_j^\mathbf{T}\rangle}{\|\mathbf{u}_j^\mathbf{T}\|} = 0$.

$\boldsymbol{\beta}$ is projection of the vector $\mathbf{x}_0-\mathbf{a}_0$ to the orthogonal complement space $V^\perp$, $\|\boldsymbol{\beta}\|$ is distance between $\mathbf{x}_0$ and S^k.

According to the theorems above, if

a) $\{\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v}_k\}$ were the basis of V, than

$$d(x_0,S^k) = \left\|\left(\frac{\langle\mathbf{x}_0-\mathbf{a}_0,\mathbf{u}_1^\mathbf{T}\rangle}{\|\mathbf{u}_1^\mathbf{T}\|},\cdots,\frac{\langle\mathbf{x}_0-\mathbf{a}_0,\mathbf{u}_n^\mathbf{T}\rangle}{\|\mathbf{u}_n^\mathbf{T}\|}\right)\right\| \quad (4)$$

b) $\{\mathbf{v}_{k+1},\cdots,\mathbf{v}_n\}$ were the basis of $V^\perp$, then

$$d(x_0,S^k) = \|(\mathbf{x}_0-\mathbf{a}_0)Q\|, \quad (5)$$

where $Q = B^T(BB^T)^{-1}B$, and

$$B = \left[\frac{\mathbf{v}_{k+1}}{\|\mathbf{v}_{k+1}\|},\cdots,\frac{\mathbf{v}_n}{\|\mathbf{v}_n\|}\right]^T.$$

If $\{\mathbf{v}_{k+1},\cdots,\mathbf{v}_n\}$ were the orthogonal basis of $V^\perp$, then

$$d(x_0,S^k) = \left\{\sum_{i=k+1}^{n}\left(\frac{\langle\mathbf{x}_0-\mathbf{a}_0,\mathbf{v}_i\rangle}{\|\mathbf{v}_i\|}\right)^2\right\}^{\frac{1}{2}} = \left\{\sum_{i=k+1}^{n}\left(\frac{(\langle\mathbf{x}_0,\mathbf{v}_i\rangle-b_i)^2}{\|\mathbf{v}_i\|^2}\right)^2\right\}^{\frac{1}{2}}, \quad (6)$$

where $b_i = \langle\mathbf{a}_0,\mathbf{v}_i\rangle$, and $k+1\le i\le n$.

The distance between the point x_0 and the k-dimension simplex $\Delta\mathbf{a}_0\mathbf{a}_1\cdots\mathbf{a}_k$ was the special case of the distance x_0 to the k-dimension plane, but there are two different:

a) To determine a pedal x_0 on the $\Delta\mathbf{a}_0\mathbf{a}_1\cdots\mathbf{a}_k$, for vectors $P-I$ and $I-P$ are opposite direction, so $\boldsymbol{\beta}$ by (1) as determined and $-\boldsymbol{\beta}$, one of them point to $\Delta\mathbf{a}_0\mathbf{a}_1\cdots\mathbf{a}_k$, and the other away.

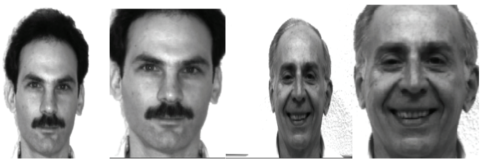
b) If the pedal is not in simplex, then its distance is point to the vertices of a simplex or one edge.

4. Realization of Biomimetic Pattern Recognition Based on High-dimensional Space and Face Recognition Test

There were m classes, each class to take $k+1$ training samples, then $k+1$ images in the feature R^d space composition the k-simplex in its generated k-dimensional linear subspace, there were m linear subspaces and k-simplexs. Now, if there were l same type images to be recognized, these images in the feature R^d space generated *l-1*-dimensional linear subspace of the *l-1*-simplex. Identification problem is to judge the distance that *l-1*-simplex which in *l-1*-dimensional linear subspace and m patterns k-simplexs which in k-dimensional subspace, if the distance with a k-simplexs less than a given threshold θ, then the sample would fall into the class [6].

Using the ORL [11] and YALE A [12] face database for the experimental subjects (Fig. 3).

(a) Part images of the ORL face database.



(b) Part images of the YALE a face database and the interception of a standardized image.

Fig. 3. Part images of the ORL and YALE a face database.

In the structure of k-simplex cover body – convex cell body, how choose the threshold θ is critical, too large may lead to increased error rate, is too small may result in rejection rate increased. As the choice of training samples with random, so to the average distance between the training sample and the training sample approximate the distance between to be recognized sample with the k-simplex. So you can take θ to the average distance α times between the training samples, α take 1 near the value.

In the ORL face database, select 30 persons in each of five of 150 face images as training samples, each of five face images to form a 4-simplex, the other a total of 150 images per 5 constitute the test sample A, the remaining 10 persons per 10, sum 100 images constitute the test samples B. PCA as feature extraction tool, PCA as feature extraction tool, principal component contribution ratio is 0.85, the dimension of feature space dimension is 47.

Experiment, takes

$$\theta = 0.85 \times \frac{1}{6} \sum_{1 \le j < i \le 4} d(a_i, a_j) \qquad (7)$$

that 0.85 times of the average distance between 4 samples .

We use above BPR were compared to the general pattern recognition method (SVM kernel nearest neighbor method) results in the following table:

Table 1. Experimental results.

	Test sample A			Test sample B	
	Correct rate	**Error rate**	**False rate**	**Corre ct rate**	**Error rate**
BPR	84.6 %	2.6 %	12.6 %	89.0 %	11.0 %
SVM	85.7 %	14.2 %	0	0	100 %

The YALE A face database we are also has been tested, first, face image using artificial interception of the face, then a unified size, and image histogram for the amendment, shown in Fig. 3 (B), taking each person four images as training samples, the contribution rate of principal component extraction is 0.85, the dimension of feature space is 20 dimensional, error recognition 13, rejection 1, correct recognition rate was 86.7 %.

5. Conclusions

Based on high-dimensional geometry, by the homology of the similar sample points, solve the threshold issue for k-simplex structure covered body, provides a way to structure the best coverage body, expand the Biomimetic Pattern Recognition method, and in face recognition experiments is valid.

Acknowledgements

Author gratefully acknowledges the Projects Supported by Qu Zhou City science and technology (20101059) for supporting this research.

References

[1]. Vapnik V., Statistical learning theory, *John Wiley &Sons, Inc.*, New York, 1998.

[2]. Wang Shoujue, Bionic (topologica1) pattern recognition - a new model of pattern recognition theory and its applications, *Aeta Electronica Sinica*, Vol. 30, Issue 10, 2002, pp. 1417 - 1420.

[3]. Wang Shoujue, Wang Bainan, Analysis and Theory of High-Dimension Space Geometry for Artificial Neural Networks, *Aeta Electronica Sinica*, Vol. 30, Issue 1, 2002, pp. 1-4.

[4]. Wang Shoujue, Lai Jiangliang, First step to Multi-dimensional space biomimetic informatics, *National Defese Industry Press*, Beijing, 2008, 1.

[5]. Wenming Cao, Clifford Manifold, Learning for Nonlinear Dimensionality Reduction, *Chinese of Journal Electronics*, Vol. 18, Issue 4, 2009, pp. 650-654.

[6]. Wenming Cao, Nengheng Zheng, Feng Hao, Hhigh-dimensional information geometry and speech analysis, *Science Press*, Beijing, 3, 2011.

[7]. Xiao Xiao, Xianbao Wang, Sunyuan Shen, Shoujue Wang, A Cognitive Model in Biomimetic Pattern Recognition and its Applications, *Cybernetics and Applications Lecture Notes in Electrical Engineering,* Vol. 107, Issue 1, 2012, pp. 953-960.

[8]. Youzheng Zhang, Wenhai Cheng, An optimized application based on double-weight neural network and genetic algorithm, *Communication in Computer and information Scienc,* Vol. 268, Issue 2, 2012, pp. 99-105.

[9]. Lai Jiangliang, Research on geometry analysis ofdots in high-dimensional space for pattern recognition, *Institute of Semiconductors Chinese Academy of Sciences,* Beijing, 2005.

[10]. Yang Guo-Wei, Wang Shoujue, Liu Yang-yang, Research on Two Key Technical Problems in Biomimetics Pattern Recognition, *Aeta Electronica Sinica*, Vol. 36, Issue 12, 2008, pp. 2490 - 2492.

[11]. http://www.cl.cam.ac.uk/research/dtg/attarchive/ /facedatabase.html

[12]. http://cvc.yale.edu/projects/yalefaces/yalefaces.htm
