形状分析ツールの開発とその応用
ー 浮世絵に描かれた役者の同定と分類 ー

A Shape Analyzing Tool and Its Applications
ー Identification and Classification of Actors Drawn in Ukiyoe Pictures ー

モハメド・アラミン・ブイヤン†、阿古　弥寿章†、濱　裕光†、松平　進††
MD. Al-amin BHUIYAN†, Yasuaki AKO†, Hiromitsu HAMA†, Susumu MATSUDAIRA††

† 大阪市立大学工学部　〒558 大阪市住吉区杉本 3-3-138
†† 甲南女子大学文学部　〒658 神戸市東灘区北町 6-2-23
† Fac. of Eng., Osaka City Univ., Sugimoto 3-3-138, Sumiyoshiku, Osaka, JAPAN 558
†† Fac. of Lit., Kounan Jyoshi Univ., Kitamachi 6-2-23, Higashinadaku, Kobe City, JAPAN 658

あらまし　多くの絵画の分析では色彩が中心に扱われてきたが、ここでは形状を扱うための手法を開発し、浮世絵（役者絵）へ応用した結果を示す。本手法は、最初は、イメージインデッキシングの例として、浮世絵における人物の同定に使うことを念頭に置いて開発されてきた。しかし、絵画データベース（ＤＢ）に限らず、より広い意味で線画イメージによるＤＢ検索、さらには３次元物体（遺物・遺跡など）の解析にも利用できるものである。ここでは、フリーハンド曲線（３次元の場合は平面）をベゼ曲線を用いて近似し、制御点のマッチングにより浮世絵に描かれた役者の同定と分類を行う。顔の各部分（左右の目、鼻、横顔の輪郭など）および全体から人物の類似度を計算し、その値を用いて同じ役者かどうかの判定を行った。その結果、適当なしきい値を設定することにより、うまく役者を識別することができた。また、本稿では紙面の都合上、報告できないが、同じ手法が正面写真に限れば、人物同定にも利用できるという結果も得られている。ここで、類似度の計算は、制御点間の距離を基本にしており、輪郭線全体にわたる必要性がないので、高速かつ柔軟なマッチングを行える利点を持つ。用いたデータ数は少ないが、本手法の可能性と有用性は計算機実験を通して確認できた。

Abstract　This paper presents a person identification and classification system of the actors drawn in Ukiyoe pictures. The system is based on visual features of the face from the image database files and is organized as a set of classifiers whose outputs are integrated after a normalization step. Line profile from the picture has been extracted in this investigation and has been approximated by Bèzier curves. The effectiveness of our method has been confirmed through computer simulation. The method developed here can be expanded to one of three dimensional shape analyzing tools. Furthermore, we hope that it has many applications in the field of not only natural science but also humanities.

キーワード： 浮世絵、人物同定、ベゼ曲線、曲線マッチング、アフィン変換
Keywords: Ukiyoe pictures, Person identification, Bèzier curve, Curve fitting, Affine transformation

1. Introduction

The identification of a person interacting with computers represents an important task for automatic systems in the area of information retrieval, automatic banking, control of access to security areas and so on. The need for a reliable identification of interacting with users is obvious. The possibility of integrating multiple identification cues, such as password, identification card, voice, face, fingerprints and so on, will in principle enhance the security of a system to be used by a selected set of people. Person identification through face recognition is the most familiar among the possible identification strategies. Several automatic or semi-automatic systems were realized in the early seventies. Different techniques were proposed, ranging from the geometric description of salient facial features to the expansion of a digitized image of the face on appropriate basis of images[1].

The objective of this research work is to develop an identification and classification system of Ukiyoe actors[2] based on visual face features. The strategy for the identification system approaches the comparison of the selected regions of the face. A set of regions, respectively encompassing the eyes and nose of the actors to be identified are being approximated by Bézier curves and the locations of the control points of these Bézier curves are computed[3]. The locations of the control points are compared with the corresponding locations of the control points stored in the database file for each reference actor.

Ukiyoe actors are being classified as fat, moderate or slim depending on the operator sigma as the ratio of the area of the face outline to the length of the outline, i.e., $\sigma = S/l$, where S is the area of the face outline and l is the length of the outline.

2. System Description

In order to be able to recognize faces by computers, face images are analyzed and then normalized, that is, are transformed to some invariant representation to size and direction. The fundamental steps employed for the face recognition process are shown in Fig. 1.

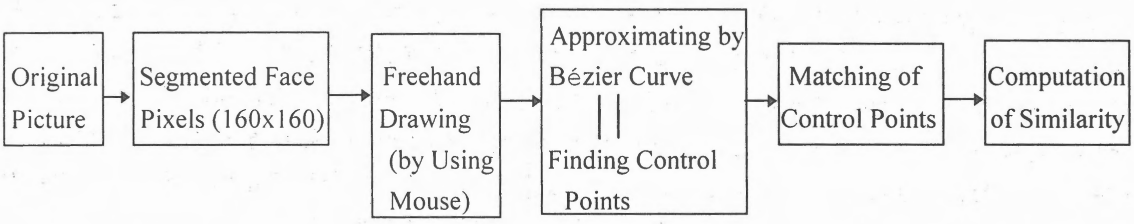


Fig. 1 Schematic diagram of the face identification system.

(1) Acquisition of Data

Acquisition of a number of Ukiyoe actors faces are done by using a scanner (Canon canoScan 600), as shown in Fig. 2. In this investigation, Ukiyoe portrait has been employed for image database in TIFF file mode. The TIFF file mode has been chosen because of its suitability for a wide range of applications and its independence of computer's architecture, operating system, graphics hardware and so on. It is reasonably compact and handles black-and-white, grayscale, and color images well, allowing the user to adjust for the unique characteristics of a scanner, monitor, or printer.

The whole images are being scanned and only the face parts are being taken out. During scanning operation, standardization and geometrical normalization (size and direction) of the images have been performed. Faces are stored in 10x7.5 cm size and during analysis these are being made as 160x160 pixel resolution, as shown in Fig. 3.

(2) Line Profiles

Line profiles of the nose and eyes are extracted from the pictures by using free hand drawing (by tracing the face outlines using mouse), as shown in Fig. 4, and approximated with Bézier curve, as shown in Fig. 5.

(3) Control Points

The control points of the Bézier curves are being computed and are stored in the respective data files.

(4) Preprocessing for Matching

This is for comparison of the location of the control points with the images stored in the database. In order to compare the image of an Ukiyoe actor with those stored in the database, different parts of the face, such as nose and eyes are drawn by freehand (traced by mouse) as binarized image. These images are then translated, scaled, and rotated so that the co-ordinates of a set of reference points take corresponding standard values.

No. 1　No. 2　No. 3　No. 4　No. 5　No. 6　No. 7

Fig. 2　Ukiyoe actors.

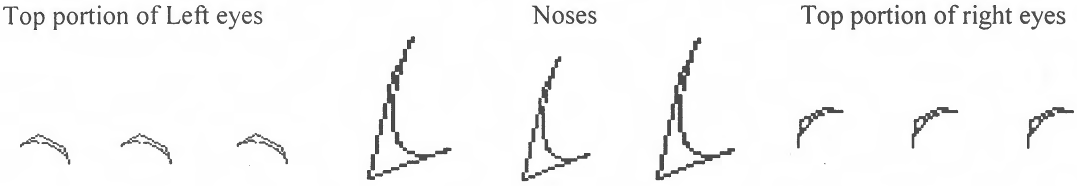
No. 1　No. 2　No. 3　No. 4　No. 5　No. 6　No. 7

Fig. 3　Faces of Ukiyoe actors as 160x160 pixel size.

No. 1　No. 2　No. 3　No. 4　No. 5　No. 6　No. 7

Fig. 4　Face outlines as traced by using mouse.

Top portion of Left eyes　　Noses　　Top portion of right eyes



No. 1　No. 2　No. 3　No. 1　No. 2　No. 3　No. 1　No. 2　No. 3

Fig. 5　Face outlines of different actors are compared by control points of approximating Bézier curves

3. Approximation Algorithm by Bézier Curve and Curve Fitting

A Bézier curve $Q(t)$ of degree n can be defined in terms of a set of control points P_i ($i=0,1,2,..,n$) and is given by[4]:

$$Q(t) = \sum_{i=0}^{n} P_i B_{i,n}(t) \tag{1}$$

where each term in the sum is the product of a blending function $B_{i,n}(t)$ and a control point P_i. The $B_{i,n}(t)$ are called Bernstein polynomials and are defined by:

$$B(t) = C_i^n t^i (1-t)^{n-i} \tag{2}$$

We can evaluate a Bézier curve in terms of its important properties: movement of the control points around influences the shape of the curve. The curve does pass through the two endpoints (P_0 and P_n) and is tangent at the endpoints to the corresponding edge of the polygon of the control points[5]. In order to find the control points of the third order Bézier curve, we applied a two stage (slope following and learning) algorithm[6].

1st Stage:

This stage is dedicated for finding the approximate location of the control points. So at this stage, the 2nd and 3rd control points of the Bézier curve are assumed to be located on the tangent lines made at the end points of the curve. Therefore, the 2nd and 3rd control points are moved along this line and the location of the $\frac{n}{5}$th and (n-$\frac{n}{5}$)th points obatined from the curve are compared with their corresponding points generated from the newly assumed control points. The reason behind this choice is two folds: (i) Since this is an approximation stage, so comparison of the $\frac{n}{5}$th and (n-$\frac{n}{5}$)th points with their corresponding points, instead of all the points with their corresponding points, obviously reduces a large amount of computational cost, and (ii) It has been found from observation that the more complex the shape of the curve or the less the number of points obtained from the given curve the better is the choice of the comparison points near the ending points. The choice of the comparison points just near the end points, however, boils down to another problem of finding the exact location of the control points. The algorithm for this approximation stage is given below.

[Step 1] Compute average tangent for both end points. Starting from the end points, move the 2nd and 3rd control points along the tangent line.

[Step 2] Compute error for the $\frac{n}{5}$th and (n-$\frac{n}{5}$)th points. If error is large enough, then goto Step 1, else goto the end of the 1st stage.

2nd Stage:

This stage is dedicated for the exact location of the control points. In this stage, first the control points P_1 and P_2 are changed alternatively by a large step and the error, that is, difference between the points obtained from the given curve and their corresponding points generated from the newly found control points are computed. As the contribution of the assumed control points in the error calculation reduces, the step size is made smaller. This process is continued until the error becomes less than 0.00001 (minimum permissible error).

[Step 1] Initialization: Count=0 (Learning time),

$\delta_x=\delta_y=1$ (Variable displacement for the movement of the 2nd and 3rd control points),

MAX=50 (Maximum Learning time),

ε=0.000001 (Minimum permissible error or difference between the corresponding points)

[Step 2] Count++;

if $E(x+\delta_x, y)$ is minimum, then goto Step 3

if $E(x-\delta_x, y)$ is minimum, then goto Step 3

if $E(x, y+\delta_y)$ is minimum, then goto Step 4

if $E(x, y-\delta_y)$ is minimum, then goto Step 4

if $E(x, y)$ is minimum, then $\delta_x=½\delta_x$, $\delta_y=½\delta_y$, goto Step 2

[Step 3] Searching minimum error for x-direction

While $(E(x+\delta_x, y)<E(x,y))$ {Count ++, $x=x+\delta_x$}

While $(E(x-\delta_x, y)<E(x,y))$ {Count ++, $x=x-\delta_x$}

if (count>MAX || $E(x,y)<\varepsilon$), then goto the end of the 2nd stage

else goto Step 2

[Step 4] Searching minimum error for y-direction

While $(E(x, y+\delta_y)<E(x,y))$ {Count++, $y=y+\delta_y$}

While $(E(x, y-\delta_y)<E(x,y))$ {Count++, $y=y-\delta_y$}

if(count>MAX || $E(x,y)<\varepsilon$), then goto the end of the 2nd stage

else goto Step 2

The data points obtained from the different curves of the face outline (nose and eyes) of the Ukiyoe actors are being fitted with piecewise geometrically continuous cubic Bézier curves. For curve fitting our particular interest is in the problem of converting bit-mapped images into its outlined representations[7]. In this case one piece of cubic Bézier curve is approximated to fit the complete outline.

4. Affine Transformation

The Affine transformation is the core of all geometric manipulations. It allows concise expression of the graphics transformations, such as, scaling, translation, and rotation needed to be performed on objects. The Affine transformation compensates for various viewing conditions.

Scaling is the process of expanding or compressing the dimensions of an object[8]. Points can be scaled by s_x along the x axis and by s_y along the y axis into new points by the multiplications:

$$x' = s_x x, \qquad y' = s_y y \tag{3}$$

In 3x3 matrix form, this can be written as[9]:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{4}$$

In translation, an object is displaced a given distance and direction from its original position. Points in the (x,y) plane can be translated to new positions by adding translation amounts to the coordinated of the points. For each point $P(x,y)$ to be moved by t_x units parallel to the x axis and t_y units parallel to the y axis to the new point $P(x,y)$, we can write:

$$x' = x + t_x, \qquad y' = y + t_y \tag{5}$$

In the 3x3 matrix form, this can be written as:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{6}$$

In rotation, the object is rotated θ about the origin. Points can be rotated through an angle θ about the origin given by:

$$x' = x\cos\theta - y\sin\theta, \qquad y' = x\sin\theta + y\cos\theta \tag{7}$$

In the 3x3 matrix form, this can be written as:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{8}$$

5. Development of the Database Management System (DBM)

A computerized database management system has been developed for the expected Ukiyoe actors. This menu driven software has been developed with control over the storage, retrieval and statistical analysis of data and information. The database has been furnished with 42 fields such as, name of the actor, photograph, control points of the left top eye, left bottom eye, right top eye, right bottom eye and nose. Once the actor information has been stored using the DBM , all or any part of it can be retrieved and manipulated by the system with the given options: add new records, edit a record, delete a record, search a record[10].

6. Experimental Results for Similarity

For the identification of different Ukiyoe actors, their face outline (nose and eyes) are approximated by Bézier curves and the control points are matched part by part basis and as a whole. The result of computation of similarity between different actors as a whole and part by part basis are shown in Table 1 and Table 2, respectively.

If the threshold value is chosen as 180 for Table 1, then we can declare that persons 1, 2 and 3 are the same and persons 6 and 7 are the same and all other persons are different. Again, by choosing the threshold value of 24 for Table 2(a), we can declare that the nose of persons 1, 2 and 3 are similar and that of persons 6 and 7 are similar, whereas the nose of all other persons are different. For the eye information, as shown in Table 2(b), 2(c), 2(d) and 2(e), although person 2 is similar to that of persons 1 and 3, butnevertheless yield some dissimilar results. Since the photograph of person 2 is right sided (all other persons are being left sided), so similar number of data points are being compared first and then his left eye is being compared with the right eye and right eye is being compared with the left eye of different

Table 1　Similarity of different actors as a whole.

Person	1	2	3	4	5	6	7
1	0	143	108	636	558	307	326
2	143	0	156	657	523	326	331
3	108	156	0	693	577	338	348
4	636	657	693	0	210	287	291
5	558	523	577	210	0	236	227
6	307	326	338	287	236	0	126
7	326	331	348	291	227	126	0

Table 2(a)　Similarity of different actors as their nose information.

Person	1	2	3	4	5	6	7
1	0	21.3	23.7	33.6	46.7	34.8	37.5
2	21.3	0	23.5	30.5	46.8	27.5	31.3
3	23.7	23.5	0	30.5	46.3	30.4	32.8
4	33.6	30.5	30.5	0	72.0	111.4	122.3
5	46.7	46.8	46.3	72.0	0	192.6	186.7
6	34.8	27.5	30.4	111.4	192.6	0	22.7
7	37.5	31.3	32.8	122.3	186.7	22.7	0

Table 2(b)　Similarity of different actors as top portion of the left eye.

<table>
<tr><td></td><td colspan="2">1</td><td colspan="2">2</td><td colspan="2">3</td><td colspan="2">4</td><td colspan="2">5</td><td colspan="2">6</td><td colspan="2">7</td></tr>
<tr><td>1</td><td colspan="2">0</td><td>8.0</td><td>19.8</td><td colspan="2">15.1</td><td colspan="2">34.3</td><td colspan="2">66.4</td><td colspan="2">48.6</td><td colspan="2">49.3</td></tr>
<tr><td>2</td><td>8.0</td><td>19.8</td><td colspan="2">0</td><td>9.3</td><td>22.7</td><td>30.1</td><td>36.2</td><td>58.3</td><td>73.8</td><td>40.2</td><td>56.4</td><td>41.7</td><td>53.2</td></tr>
<tr><td>3</td><td colspan="2">15.1</td><td>9.3</td><td>22.7</td><td colspan="2">0</td><td colspan="2">35.2</td><td colspan="2">61.4</td><td colspan="2">50.3</td><td colspan="2">52.8</td></tr>
<tr><td>4</td><td colspan="2">34.3</td><td>30.1</td><td>36.2</td><td colspan="2">35.2</td><td colspan="2">0</td><td colspan="2">50.6</td><td colspan="2">40.7</td><td colspan="2">43.7</td></tr>
<tr><td>5</td><td colspan="2">66.4</td><td>58.3</td><td>73.8</td><td colspan="2">61.4</td><td colspan="2">50.6</td><td colspan="2">0</td><td colspan="2">26.1</td><td colspan="2">27.9</td></tr>
<tr><td>6</td><td colspan="2">48.6</td><td>40.2</td><td>56.4</td><td colspan="2">50.3</td><td colspan="2">40.7</td><td colspan="2">26.1</td><td colspan="2">0</td><td colspan="2">7.8</td></tr>
<tr><td>7</td><td colspan="2">49.3</td><td>41.7</td><td>53.2</td><td colspan="2">52.8</td><td colspan="2">43.7</td><td colspan="2">27.9</td><td colspan="2">7.8</td><td colspan="2">0</td></tr>
</table>

Table 2(c) Similarity of different actors as top portion of the right eye.

	1	2	3	4	5	6	7
1	0	4.5 9.6	5.2	14.8	8.7	12.6	12.9
2	4.5 9.6	0	4.4 8.3	12.5 16.2	7.6 11.3	10.5 16.7	11.3 17.8
3	5.2	4.4 8.3	0	14.7	9.1	12.3	13.1
4	14.8	12.5 16.2	14.7	0	14.0	20.7	22.1
5	8.7	7.6 11.3	9.1	14.0	0	17.8	18.3
6	12.6	10.5 16.7	12.3	20.7	17.8	0	5.3
7	12.9	11.3 17.8	13.1	22.1	18.3	5.3	0

Table 2(d) Similarity of different actors as bottom portion of the left eye.

	1	2	3	4	5	6	7
1	0	5.3 8.5	7.3	12.6	22.7	24.3	26.1
2	5.3 8.5	0	6.1 10.5	9.8 14.8	15.3 26.5	20.7 36.5	22.8 35.3
3	7.3	6.1 10.5	0	12.3	20.8	24.7	25.8
4	12.6	9.8 14.8	12.3	0	13.6	19.8	20.7
5	22.7	15.3 26.5	20.8	13.6	0	27.5	25.3
6	14.3	20.7 36.5	24.7	19.8	27.5	0	6.7
7	26.1	22.8 35.3	25.8	20.7	25.3	6.7	0

Table 2(e) Similarity of different actors as bottom portion of the right eye.

	1	2	3	4	5	6	7
1	0	2.3 6.3	3.9	13.2	8.9	12.3	13.1
2	2.3 6.3	0	2.1 5.3	8.7 15.4	7.5 10.3	10.7 14.3	11.4 13.9
3	3.9	2.1 5.3	0	12.9	8.6	12.1	12.9
4	13.2	8.7 15.4	12.9	0	8.3	14.8	15.3
5	8.9	7.5 10.3	8.6	8.3	0	8.9	9.3
6	12.3	10.7 14.3	12.1	14.8	8.9	0	3.4
7	13.1	11.4 13.9	12.9	15.3	9.3	3.4	0

Table 3 Computation of similarity between different actors.

Person	S	l	$\sigma=S/l$	$d_1:d_2$	Comments	
1	14534	28498	0.51	$d_1<d_2$	Moderate	Left-sided
2	11089	22630	0.49	$d_1>d_2$	Moderate	Right-sided
3	12973	25973	0.50	$d_1<d_2$	Moderate	Left-sided
4	9736	36059	0.27	$d_1<d_2$	Slim	Left-sided
5	13065	23807	0.57	$d_1<d_2$	Fat	Left-sided
6	13570	25603	0.53	$d_1<d_2$	Moderate	Left-sided
7	13631	26213	0.52	$d_1<d_2$	Moderate	Left-sided

persons respectively. These results demonstrate that eyes of persons 1, 2 and 3 are similar and those of persons 6 and 7 are similar.

7. Experimental Results for Classification of Different Actors

Actors are being classified as fat, moderate and slim, according to the ratio of their face areas to the length of the face outlines. Their facing is being obtained by computing the difference between the location of the 2nd control point of the nose to the 1st control point of the left eye (d_1) and the 1st control point of the right eye (d_2). The result of computation of similarity, which is defined as the summation of the distance of the corresponding control points measured after Affine transformation between different actors, as their face outline is shown in Table 3. Persons are being classified depending on the threshold value of 0.53 for fat and 0.47 for slim.

8. Conclusion

The human face is an elastic object. So changing in facial expressions may change the position of the control points, reflecting the elasticity of the face and relative independence of its parts. So recognition of human faces is a problem that appears time.

In this investigation, Ukiyoe pictures has been analyzed, searched and classified. Detection of individual features of the actors such as eyes, nose and head outline have also been performed and the face models are being defined by the relative position of the features and their size. We have justified the system on a face image database of seven Ukiyoe actors with a success rate of 100%. Experimental results demonstrate that our system is reliable and no special tuning is required if the face occupies the same type of impression of the image.

In this investigation, the third order Bézier curve has been used to identify the face outlines. The adoption of Bézier curves has reduced the computational cost, as only four control points are sufficient to express one curve. But this type of image recognition may also be developed by using B-spline curves, Fourier descriptor or contour lines. For the classification of the Ukiyoe actors, we used $\sigma=S/l$ as a characteristic operator and chosen some threshold value of 0.53 or above to be the fat and 0.47 or less to be the slim person. But development of region features may also be employed for this purpose. So our next approach is the automatic identification of human faces using these above mentioned features.

References:

[1] R. Brunelli and D. Falavigna, "Person Identification Using Multiple Cues", IEEE transaction on Pattern Analysis and Machine Intelligence, Vol. 17, No. 10, 1995.
[2] Jun Takabuchi, "Image Retrieval for Ukiyoe DB", Graduation Thesis, Osaka City University, 1997.
[3] Md. Al-Amin Bhuiyan and Hiromitsu Hama, "An Accurate Method for Finding the Control Points of Bézier Curves", Memoirs of the Faculty of Engineering, Osaka City University, 1997 (submitted).
[4] Alan Watt and Mark Watt, "Advanced Animation and Rendering Techniques", ACM press, New York, USA, 1995.
[5] M. Newman and F. Sproull, "Principles of Interactive Computer Graphics", McGraw-Hill book company, 1979.
[6] Md. Al-Amin Bhuiyan and Hiromitsu Hama, "An Algorithm for Finding the Control Points of Bézier curves", Kansai Shibu Rengo Taikai `97, G16-6, 1997.

[7] Lejun Shao and Hao Zhou, "Curve fitting with Bézier cubics", Graphical Models and Image Processing, Vol. 58, No. 3, pp. 223-132, 1996.

[8] Roy A. Plastock and Gordon Kalley, "Computer Graphics", Mc-Graw-Hill, 1986.

[9] James D. Foley, Andries Van Dan, Steven K. Feiner, John F. Hughes, "Computer Graphics", Addison-Wesley, 1990.

[10] Md. Sakil Ahmed, Shahida Rafiq and Md. Al-Amin Bhuiyan, "Development of a Database Management System for the Expected Manpower Export in Bangladesh", The Dhaka University Journal of Science", Vol. 44, No. 1, 1996.