

Description, Recognition and Analysis of Biological Images

Donggang Yu*, Jesse S. Jin*, Suhuai Luo*, Tuan D. Pham[†] and Wei Lai**

*University of Newcastle, NSW 2308, Australia

[†]The University of New South Wales, Canberra, ACT2600, Australia

**Swinburne University of Technology, Hawthorn, VIC3122, Australia

Abstract. Description, recognition and analysis biological images plays an important role for human to describe and understand the related biological information. The color images are separated by color reduction. A new and efficient linearization algorithm is introduced based on some criteria of difference chain code. A series of critical points is got based on the linearized lines. The series of curvature angle, linearity, maximum linearity, convexity, concavity and bend angle of linearized lines are calculated from the starting line to the end line along all smoothed contours. The useful method can be used for shape description and recognition. The analysis, decision, classification of the biological images are based on the description of morphological structures, color information and prior knowledge, which are associated each other. The efficiency of the algorithms is described based on two applications. One application is the description, recognition and analysis of color flower images. Another one is related to the dynamic description, recognition and analysis of cell-cycle images.

Keywords: High-content screening, morphological structures, image contours, shape recognition, image reconstruction, analysis of cell phases.

PACS: 43.35.Ei, 78.60.Mq

INTRODUCTION

Description, recognition and analysis biological images plays an important role for human to describe and understand the related biological information. In order to better understand the complex processes of cell nucleus division under new drug treatment [1]-[5], the cell images need to be described, recognized and analyzed. Another problem for identifying the size and shape of the cell nuclei is that they are touching each other. In this case, it is necessary that the images of touching cell nuclei are detected, separated and reconstructed automatically in order to determine the size and shape of the cell nuclei in the frame [6, 7, 8]. Also it is difficult for most recognition methods to deal with them, including both structural and statistical approaches [9], [10], [11], [12], [13], [14]. In structural recognition method [9], it is difficult to describe the structure patterns of touching cells. If optimized nearest neighbor classifier (ONNC) [11] is used, a lot of patterns of touching cell can influence training result, and it is not possible to find the shape of each cell which is touched with other cells.

The description of some objects can be described by one color in color biological images. In this case, color is one important description feature, and the algorithm, color reduction [15], is introduced to separate color biological images.

The efficiency of the algorithms is described based on two applications. One application is the description, recognition and analysis of color flower images. Another one is related

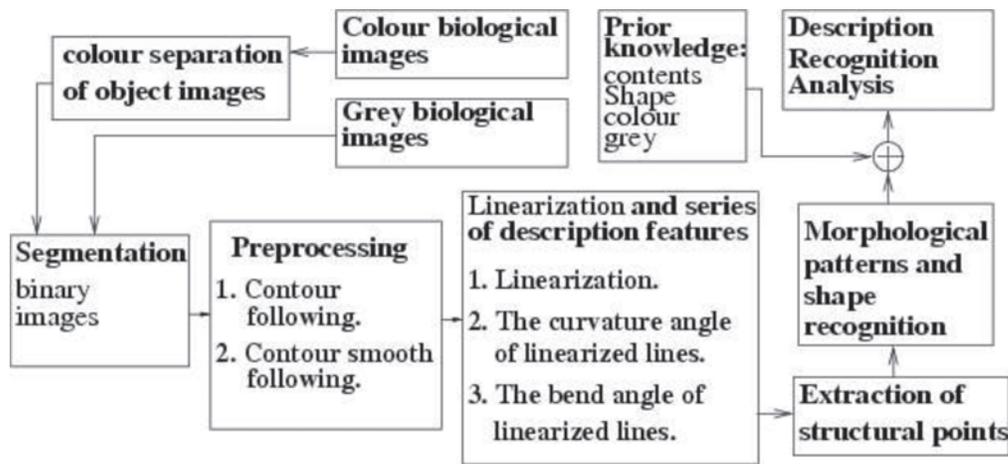


FIGURE 1. The processing diagram of description, recognition and analysis of biological images.

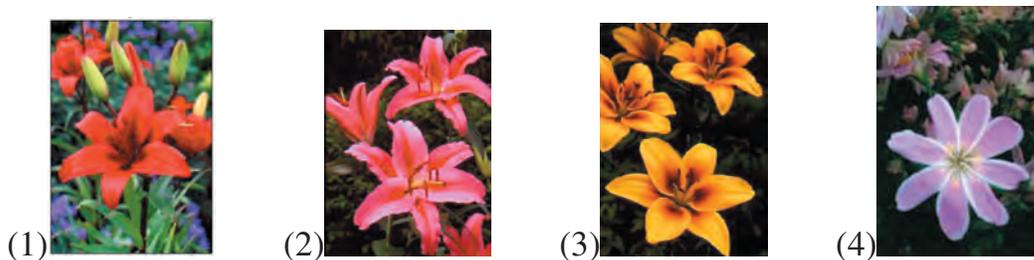


FIGURE 2. Four sample color flower images.

to the dynamic description, recognition and analysis of cell-cycle images. In Section 2, color separation of color images is introduced, smooth following and linearization are described based on difference chain codes. Secondly, morphological structural points are found in terms of smooth followed contours and linearized lines, and then the patterns morphological structural points and their properties are given. In Section 3, dynamic description, recognition and analysis of cell-cycle are described based on the description of shape structures, prior information, which are associated each other. Finally, a conclusion is given. The diagram of processing system is shown in Fig. 1.

COLOR IMAGE SEPARATION, EXTRACTION OF DESCRIPTION FEATURES AND SHAPE RECOGNITION

In order to describe, recognize and analyze biological images, we have to use all possible features. For example, if the set is “flower image”, the following procedures should be done: (a) all flower images are collected; (b) shape of flower is recognized; (c) determine which flower shape belongs based shape database. Also, it is possible that the same shape is different object in different background which is described by text contents and color. For example, the sample images in Fig. 2 could be recognized as Indo-Pacific star (starfish) in marine environment or a star in a painting etc.

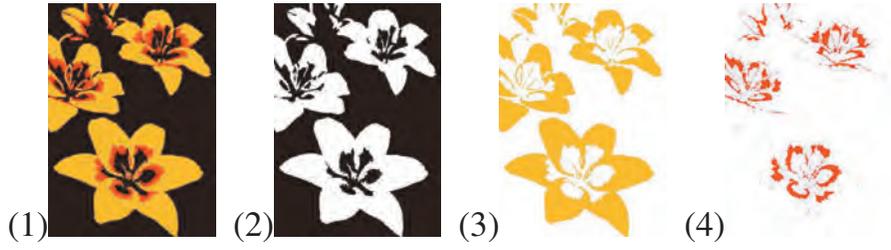


FIGURE 3. The color reduction and lay images of the image in Fig. 2(3)

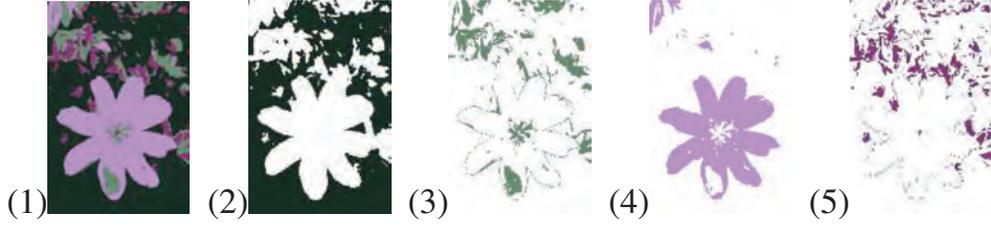


FIGURE 4. The color reduction and lay images of the image in Fig. 2(4).

Color image separation based on color reduction

In most of cases, it is possible that the object images can be represented and understood with a limited number of colors [15]. Especially, in many cases, the object images can be constructed by one color. For example, we can say the blue sky, white cloud, yellow lily flower, silver plan, red or gold sun and so on. However, the true type color images consist of more than 16 million different colors in a 24 bit full RGB color space. In order to separate the object image which consists of one color, it is necessary to reduce the number of colors. One method of color reduction [15] is used, which is multi-thresholding (adaptive color reduction) with SOFM neural network. Based on the above algorithm, the optimal number of thresholds of color reduction, c_n , is estimated. Let the color reduction set is C_R , then

$$C_R = \{c_r([r_0, g_0, b_0]), \dots, c_r([r_{c_n-1}, g_{c_n-1}, b_{c_n-1}])\} \quad (1)$$

where $c_r([r_i, g_i, b_i])$ is i -th color. For example, the number of colors of images in Figs. 2(3-4) and 2(1) is reduced, and the processed results are shown in (1), 4(1) and 5(1) respectively. Furthermore, each lay image can be got based on reduction colors. Let $I_{cr}[i]$ is the lay image which is found based on the $c_r[i]$ -th color, where the background color is supposed as white color in general cases. If the object image consists of white color

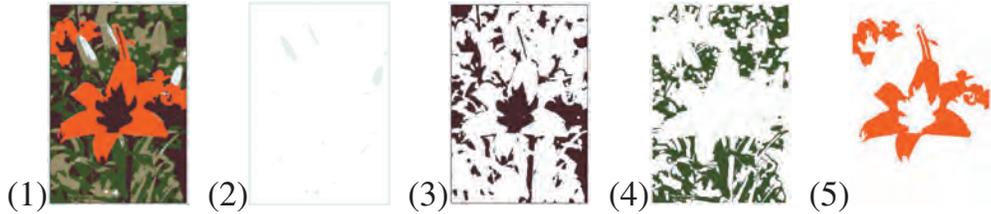


FIGURE 5. The color reduction and lay images of the image in Fig. 2(1).

based on prior information (text, color etc.), the background color is black color. Based on the above algorithm, the set of the lay images (I_{rc}) is

$$I_{rc} = \begin{cases} I_{rc}[0] & I_{xy}[r, g, b] = c_r[r0, g0, b0] \\ \cdot \\ \cdot \\ I_{rc}[i] & I_{xy}[r, g, b] = c_r[ri, gi, bi] \\ \cdot \\ \cdot \\ I_{rc}[c_n - 1] & I_{xy}[r[c_n - 1], g[c_n - 1], b[c_n - 1]] \\ & = c_r[r[c_n - 1], g[c_n - 1], b[c_n - 1]] \end{cases} \quad (2)$$

For the images in Figs. 3(1), 4(1) and 5(1), their lay images are shown in Figs. 3(2-4), 4(2-5) and 5(2-5) respectively based on the above algorithm.

The linearization and description features

All lay images can be transformed into binary images based on two colors of object and background. For example, if a yellow flower need to be retrieved or classified, the binary image is found from the yellow lay image in Fig. 3(3). Its binary image is shown in Fig. 6(1). line segment, curvature angle and bend angle between neighboring lines, and their convexity and concavity are useful to describe the shape of binary images [16]. Let the starting point of an binary image be the upper-left corner. The chain code set of contour k is represented as:

$$C_k = \{c_0, c_1 \dots c_i, \dots c_{n-1}, c_n\} \quad (3)$$

where i is the index of the contour pixels. The difference code is defined as:

$$d_i = c_{i+1} - c_i. \quad (4)$$

In the smoothed contour, $|d_i|$ equals 0 or 1 [17] [18].

The linearizing lines based on difference codes

The smoothed contour can be converted to a set of lines, which consist of ordered pixels. Suppose that chain code set of line ln of contour k is

$$c_k^{ln} = \{c_k^{ln}[0], c_k^{ln}[1] \dots c_k^{ln}[i], \dots c_k^{ln}[n_k^{ln} - 1]\}, \quad (5)$$

where n_k^{ln} is the total number of pixels in the line ln . A linearized line has following property:

if

$$d_{ij} = c_k^{ln}[i] - c_k^{ln}[j] \quad (i = 0, \dots [n_k^{ln} - 1]), (j = 0, \dots [n_k^{ln} - 1]), \quad (6)$$

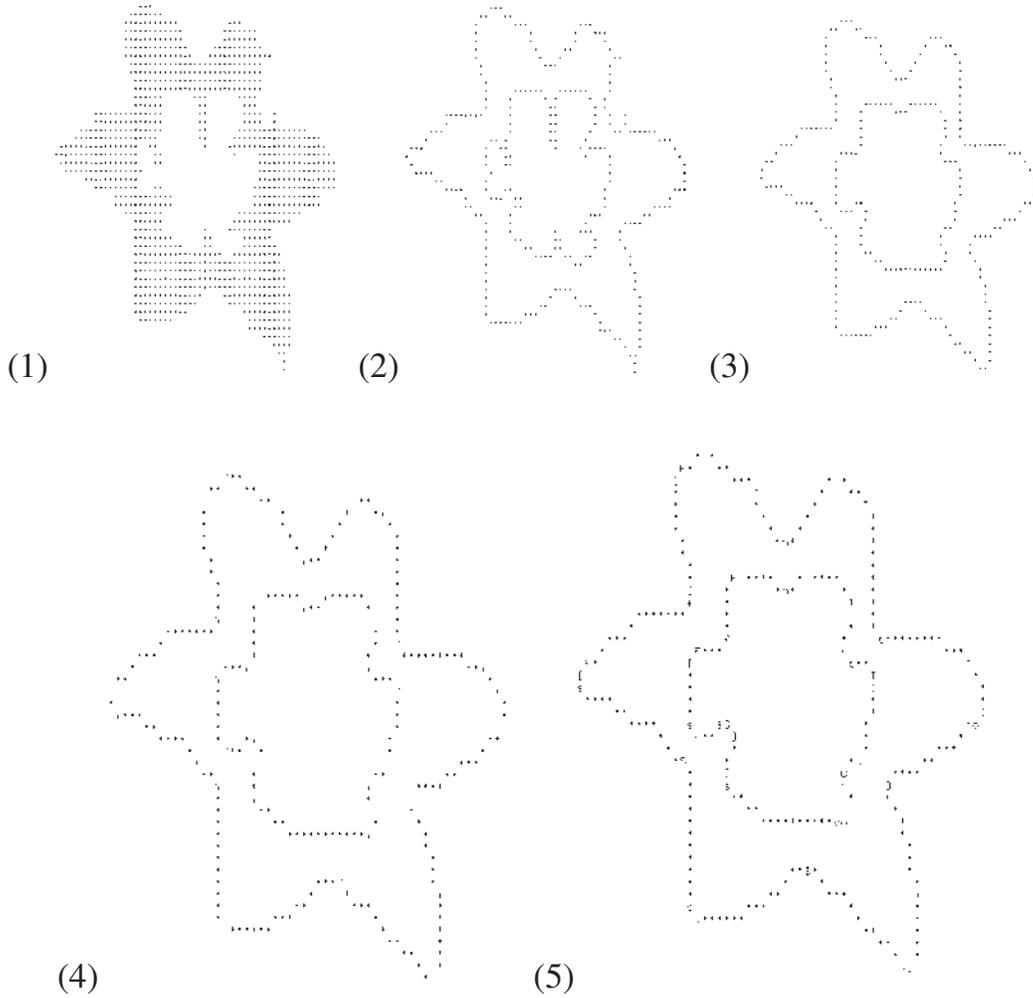


FIGURE 6. Binary image, contour and smooth following, linearization, and structural points of the image in Fig. 3(3).

then

$$|d_{ij}| \leq 1(\text{mod}8) \quad (i = 0, \dots, [n_k^{ln} - 1]), (j = 0, \dots, [n_k^{ln} - 1]). \quad (7)$$

Therefore, a linearized line contains only two element codes, representing with $cdir1$ and $cdir2$. Also, its sets of chain code (c_k^{ln}), x and y coordinates (x_k^{ln} and y_k^{ln}), can be found.

The following, smoothing and linearization results of two examples (Figs. 6(1) and 6(1)) are shown in Figs. 6(2-4) respectively, and the starting point of each line is represented by character “Y”.

The set of curvature and bend angles of linearized lines

1. The curvature angle of a linearized line

Let l_{se} is straight line between the starting and end points of each linearized line. The curvature angle is defined as the direction angle between the x coordinate axis and l_{se} , and the angle is formed with starting from the direction of the x coordinate

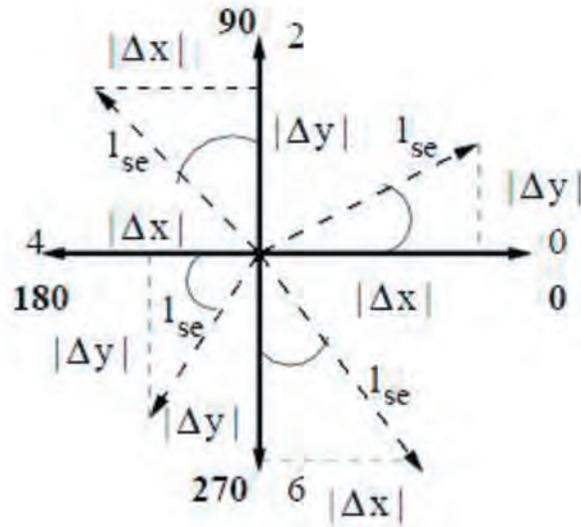


FIGURE 7. Finding the curvature angle of linearized lines.

axis to the direction of linearized line, which is determined by the line's element codes, in anti-clock. It can be found based on Fig. 7. Suppose \angle_{curve} is the curvature angle, then it is found as follows: (corresponding four quadrants):

(1) $cdir1$ and $cdir2$ being chain code 0, 1 or 2 (the first quadrant)

let \angle_{se} be the tangent angle of linearized line ln , then

$$\angle_{se} = \text{tag}^{-1}\left(\frac{\Delta y}{\Delta x}\right), \quad (8)$$

where Δx and Δy are the absolute values of differences of x and y coordinate between starting and end points of line, which are shown in Fig. 7.

Let \angle_{curve} be curvature angle of a line, then

$$\angle_{curve} = (180^\circ/\pi)\angle_{se} \quad (9)$$

based on Fig. 7 and Equation 8.

(2) $cdir1$ and $cdir2$ being chain code 4, 5 or 6 (the third quadrant)

In this case,

$$\angle_{curve} = 180^\circ + (180^\circ/\pi)\angle_{se} \quad (10)$$

based on Fig. 7 and Equation (8).

(3) $cdir1$ and $cdir2$ being chain code 2, 3 or 4 (the second quadrant)

In this case, \angle_{se} is

$$\angle_{se} = \text{tag}^{-1}\left(\frac{\Delta x}{\Delta y}\right), \quad (11)$$

and

$$\angle_{curve} = 90^\circ + (180^\circ/\pi)\angle_{se} \quad (12)$$

based on Fig. 7 and Equation 11

(4) $cdir1$ and $cdir2$ being chain code 6, 7 or 0 (the fourth quadrant)

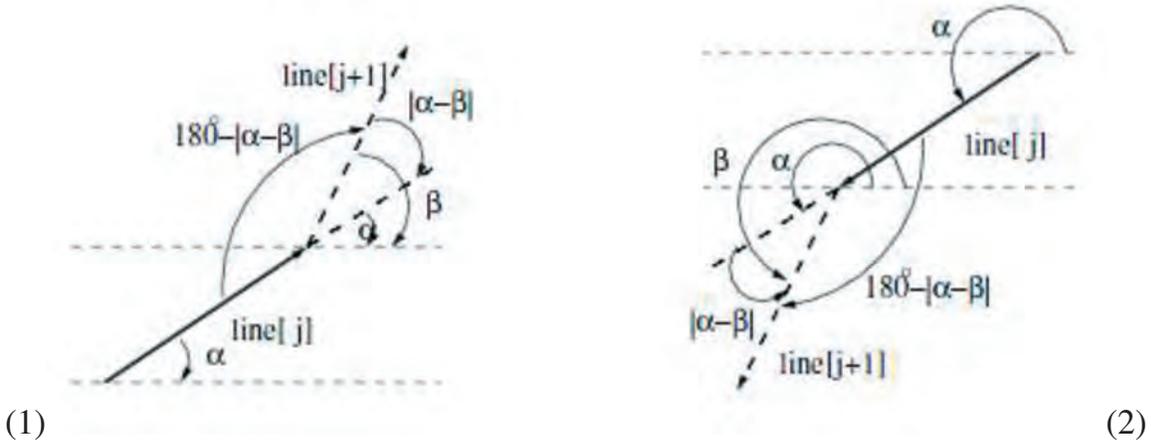


FIGURE 8. Finding the bend angle between two neighboring linearized lines.

ln	Curvature angle	Bend angle (convex or concave)
0	206°	104° convex
1	282°	97° concave
2	199°	109° convex
3	270°	111° convex
4	338°	111° concave
5	270°	75° convex
6	15°	131° concave
7	326°	65° convex
8	80°	141° convex
9	120°	83° concave

(1)

ln	Curvature angle	Bend angle (convex or concave)
10	23°	86° convex
11	116°	121° convex
12	174°	100° concave
13	96°	114° convex
14	161°	121° convex
15	219°	87° concave
16	127°	145° convex
17	161°	135° convex

(2)

Table 1: Series of morphological feature vectors of lily flower in Fig. 6.

In this case,

$$\angle curve = 270^\circ + (180^\circ/\pi)\angle se \quad (13)$$

based on Fig. 7 and Equation 11.

2. Bend angle and its property

The bend angle is defined as the angle between the lines j and $(j + 1)$.

(a) The bend angle of linearized lines

Let $\angle curve[j]$ and $\angle curve[j + 1]$ be the curvature angle of linearized lines j and $[j + 1]$ respectively, and $\Delta[j, j + 1]$ be their curvature angle difference, then

$$\Delta[j, j + 1] = \angle curve[j] - \angle curve[j + 1], \quad (14)$$

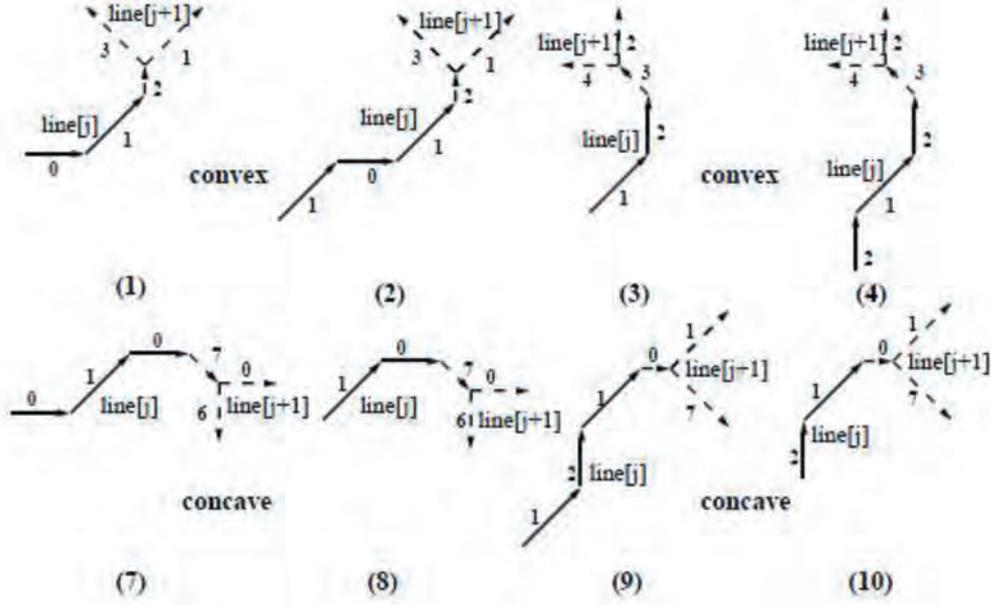


FIGURE 9. Detecting patterns of bend angle convexity or concavity.

where $\angle_{curve}[j]$ is α , $\angle_{curve}[j + 1]$ is β , and $\Delta[j, j + 1]$ is $|\alpha - \beta|$ in Fig. 9 respectively. Let $\angle_{angle}[j]$ be the bend angle, then:

$$\angle_{angle}[j] = 180^\circ - |\Delta[j, j + 1]| \quad (15)$$

based on Fig. 8.

(b) The property of bend angle

The bend angle property, convex or concave, can be determined based on the structure patterns of the element codes between two neighboring lines. Let $l_{cdir1}[j]$ and $l_{cdir2}[j]$ be the first and second element codes of the line j , and $l_{cdir1}[j + 1]$ and $l_{cdir2}[j + 1]$ be the first and second element codes of the line $(j + 1)$ respectively, then there are thirty two detection patterns of the bend angle property (convex or concave). The detection patterns (1-4) and (7-10) are shown in Fig. 8.

One detection rules can be described as following:

If $l_{cdir1}[j]$ is chain code 0 and $l_{cdir2}[j]$ is chain code 1 (see Figs. 9(1) and 9(7)), or $l_{cdir1}[j]$ is chain code 1 and $l_{cdir2}[j]$ is chain code 0 (see Figs. 9(2) and 9(8)), then

- If $l_{cdir1}[j + 1]$ is chain code 2 or $l_{cdir2}[j + 1]$ is chain code 2, then $\angle_{angle}[j]$ is convex (see Figs. 9(1-2)).
- If $l_{cdir1}[j + 1]$ is chain code 7 or $l_{cdir2}[j + 1]$ is chain code 7, then $\angle_{angle}[j]$ is concave (see Figs. 9(7-8)).

Similarly, other detection rules can be described based on other patterns.

Based on the above algorithms, the series set of the curvature angle, the series of bend angle (including their convexity or concavity) of two sample images, which are shown in Fig. 6(4), can be found from starting line to end line. For example, the vector of a series of curvature and bend angles of the outer contour in Fig. 6(4) is:

$206^\circ_{-104^\circ}(\text{convex}) \rightarrow 282^\circ_{-97^\circ}(\text{concave}) \rightarrow 199^\circ(\text{convex})_{-109^\circ} \rightarrow 270^\circ_{-111^\circ}(\text{convex}) \rightarrow$
 $338^\circ_{-111^\circ}(\text{concave}) \rightarrow 270^\circ_{-75^\circ}(\text{convex}) \rightarrow 15^\circ_{-131^\circ}(\text{concave}) \rightarrow 326^\circ_{-65^\circ}(\text{convex}) \rightarrow$
 $80^\circ_{-141^\circ}(\text{convex}) \rightarrow 120^\circ_{-83^\circ}(\text{concave}) \rightarrow 23^\circ_{-86^\circ}(\text{convex}) \rightarrow 116^\circ_{-121^\circ}(\text{convex}) \rightarrow$
 $174^\circ_{-100^\circ}(\text{concave}) \rightarrow 96^\circ_{-114^\circ}(\text{convex}) \rightarrow 161^\circ_{-121^\circ}(\text{convex}) \rightarrow 219^\circ_{-87^\circ}(\text{concave}) \rightarrow$
 $127^\circ_{-145^\circ}(\text{convex}) \rightarrow 161^\circ_{-135^\circ}(\text{convex})$ (see Table 1). These series sets of description features can be used to describe and recognize the shape of contours. Shape of the outer contour in Fig. 6(4) is hexangular because there are six groups of convex and concave changes. Therefore, if the color yellow is detected and it is a flower, the object image is yellow lily flower as there are six petals for most of lily flower. Also, it could possibly be recognized as Indo-Pacific star (starfish) in marine environment.

Structural points of smoothed contours

The structural points are some special points which can be used to represent convex or concave change in the direction of chain codes between two neighboring lines along the contour. Their definition and detection are based on the structure patterns of element codes of two lines. Assume that $line[ln]$ is the current line and that $line[ln - 1]$ is the previous line.

Definition 1: The convex point in the direction of code 4 (represented with the character “^”)

If the element codes 3, 4 and 5 occur successively as a group of neighborhood linearized lines, then one convex point can be found as follows:

if the first element code $cdir1$ of $line[ln]$ is code 4, the second element code $cdir2$ is code 5 and the direction chain code of the last pixel of $line[ln - 1]$ is code 3, then the first pixel of the current line $line[ln]$ is a convex point.

Definition 2: The concave point in the direction of code 4 (represented with the character “m”)

If the element codes 5, 4 and 3 occur successively as a group of neighborhood linearized lines, then one concave point can be found as follows:

if the first element code $cdir1$ of $line[ln]$ is code 4, the second element code $cdir2$ is code 3 and the direction chain code of the last pixel of $line[ln - 1]$ is code 5, then the first pixel of the current line, $line[ln]$, is a concave point.

Similar to Definitions 1-2, other structural feature points can be defined and found. These points are convex points “v”, “[”, “)”, “F”, “o”, “T”, “s”, and concave points “\$”, “[”, “(”, “f”, “O”, “t” and “S” which are shown in Figs. 10 respectively. These structural points describe the convex or concave change in different chain code directions along the contour, and they can therefore be used to represent the morphological structure of contour regions.

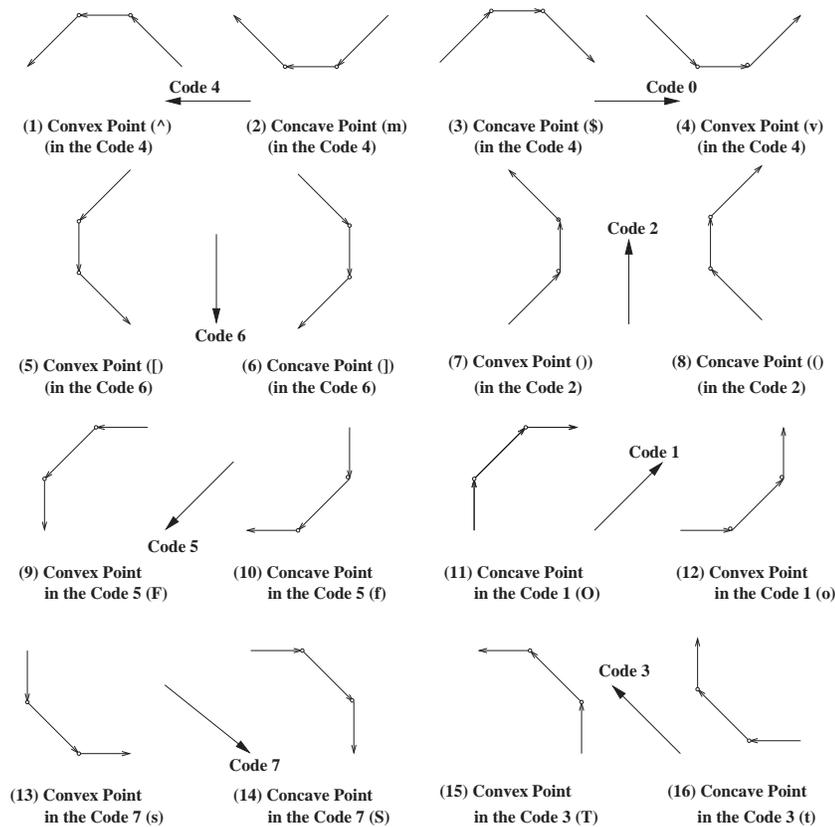


FIGURE 10. The pattern models of structural points.

Shape recognition based on structural points, curvature and bend angles

The sets of structural point, curvature and bend angles of contours can be found based on the above algorithms. They can be used to recognize the shape of contours. Based on the above method, all structural points of the contours in Fig. 6(4) can be found, and they are shown in Fig. 6(5).

For morphological structure of the outer contour of Fig. 6(5) there is a series of structural points:

“^” → “F” → “[” (convex change) → “]” → “f” (concave change) → “F” → “[” → “s” (convex change) → “S” (concave change) → “s” → “v” (convex change) → “\$” (concave change) → “v” → “o” → “(” (convex change) → “(” → “O” (concave change) → “o” → “)” → “T” (convex change) → “t” → “)” (concave change) → “)” → “T” → “^” (convex change) → “m” (concave change) → “^” (convex change).

It is clear, both outer contours are six angles because of six pairs of convex and concave

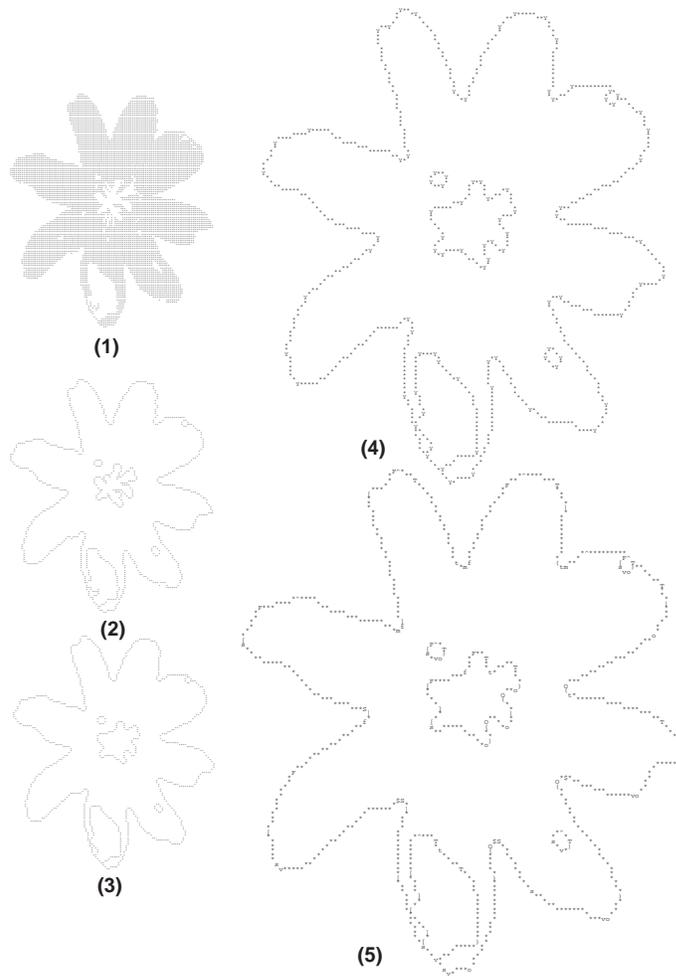


FIGURE 11. Binary image, contour and smooth following, linearization, and structural points of the image in Fig. 4(4).

change. For most sorts of lily flower there are six petals which are constructed by sixangles.

The binary image, contour and smooth following, linearization, and structural points of the image in Fig. 4(4) are shown in Fig. 11. For morphological structure of the outer contour of Fig. 11(5) there is a series of structural points:

“^” → “F” → “[” (convex change) → “]” → “f” → “m” (concave change) → “^” → “F” → “[” → “s” (convex change) → “S” → “]” → “f” (concave change) → “F” → “[” → “s” → “v” (conve change) → “\$” → “S” → “]” (concave change) → “[” → “s” → “v” → “o” (convex change) → “O” → “\$” → “S” (concave change) → “s” → “v” → “o” → “)” (convex change) → “(” → “O” → “\$” (concave change) → “v” → “o” → “)” → “T” (convex change) → “t” → “(” → “O” (concave change) → “o” → “)” → “T” → “^” (convex change) → “m” → “t” → “)” (concave change) → “)” → “T” → “^” → “F” (convex change) → “f” → “m” → “t” (concave change) → “T” → “^” (convex change).

For the image in Fig. 4(4), there are eight petals which are constructed by eight-

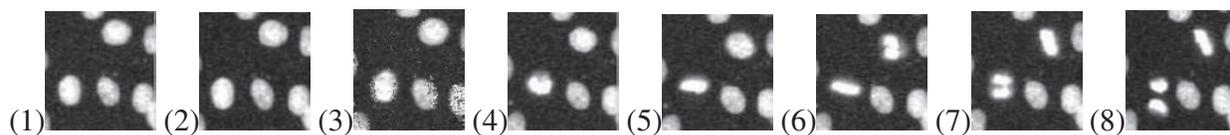


FIGURE 12. One example of cell-cycle.

square based on the above structural pattern. Each convex change consists of a group of structural points which can be used to describe convex change in the direction of a chain code. For example, the first group of structural points in Fig. 6(5) is structural points “^”, → “F” and → “[” which are all convex points to describe the shape of the convex contour. In some cases, the shape recognition of other lay images which are in the outer contour is needed. Their colors and shape can describe the property of object image. The series set of linearized lines, curvature and bend angles give the detail of contour description. Based on the above similar description, different object image structure patterns can be constructed. Based on these structural patterns, the shape of object images can be recognized. Therefore, if retrieved object image is lily flower, the images in Figs. 2(1-3) are selected. If red lily flower is searched, the images in Figs. 2(1-2) are got. Also, the image in Fig. 2(3) is yellow lily flower. The image in Figs. 2(4) is not lily flower based on the constructed structure patterns of the object images of flowers.

DYNAMIC DESCRIPTION, RECOGNITION AND ANALYSIS OF CELL-CYCLE

An essential task for cell-cycle screening is to measure cell cycle progression (inter phase, prophase, metaphase, and telophase) which can be identified by measuring nuclear changes. One example, the nuclear migration during cell division (eight series of times), is shown in Fig. 12, which are the section of big original microscope images.

The morphological structure of different cell phases

In order to find the morphological structure of different cell phases, we can use the methods which described in Sections 2, 3 and 4 to process and analyze these original microscope cell images. Many methods and algorithms are developed for the description of contours in the past [19] [20] [21] [22]. However, these descriptions cannot form series of sets, or the inner contour of a binary image can not be processed based on these algorithms [21], which make the analysis and understanding of contour shape difficult.

Binarization

The Otsu’s thresholding method [23] is used to separate binary cell images in cell-cycle screening and is shown in Fig. 13.

Smooth following and linearization

The contours in Fig. 12(1-8) are followed smoothly and linearized. They are shown in Fig. 14 where the spurious points in contours are removed and character "Y" is the first point of each linearized line.

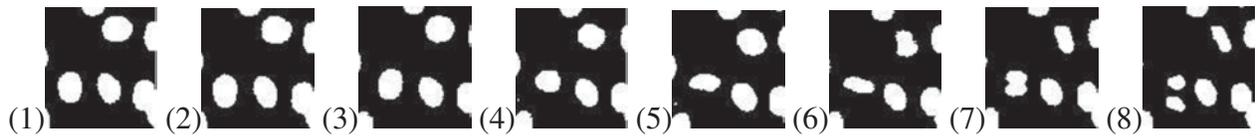


FIGURE 13. Binarization of cell images in one cell-cycle (Fig. 12).

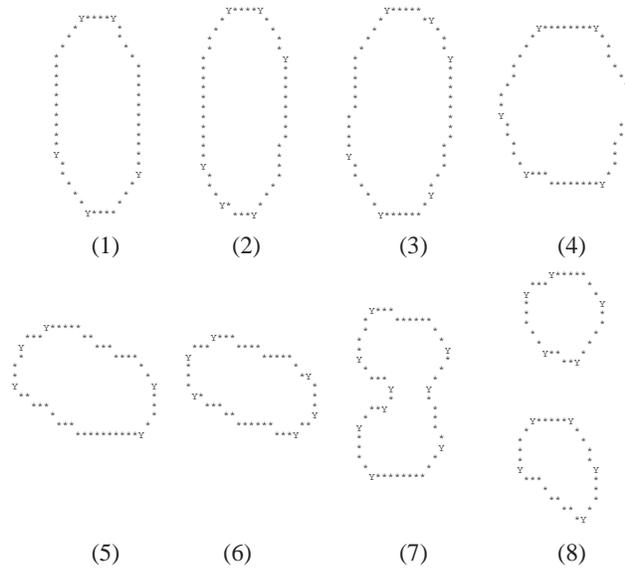


FIGURE 14. Smooth following and linearization of contours of the bottom-left cell in the cell-cycle (Fig. 12).

The series of structural points of cell images

The series of structural points of cell images (see Fig. 14) can be found and shown in Fig. 15 based on the above algorithm (see Section 4). For the outer contour in Fig. 15(1),

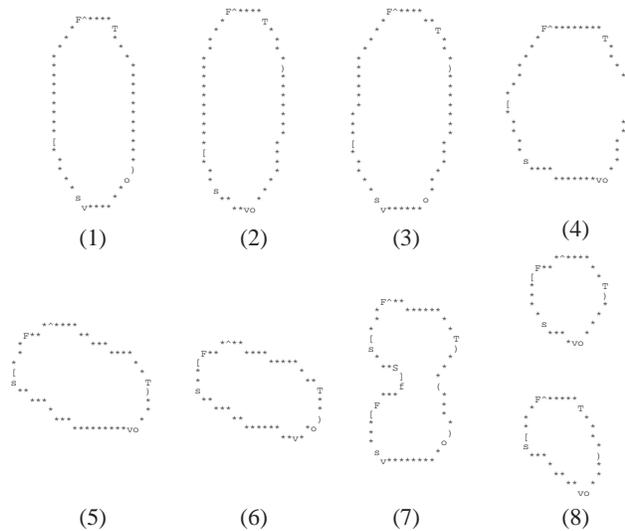


FIGURE 15. Cell morphological structures of different phases (the bottom-left cell in the cell-screen (Fig. 12).

the series of structural points is:



FIGURE 16. Binarization images of three touching cell images and one telophase cell image.

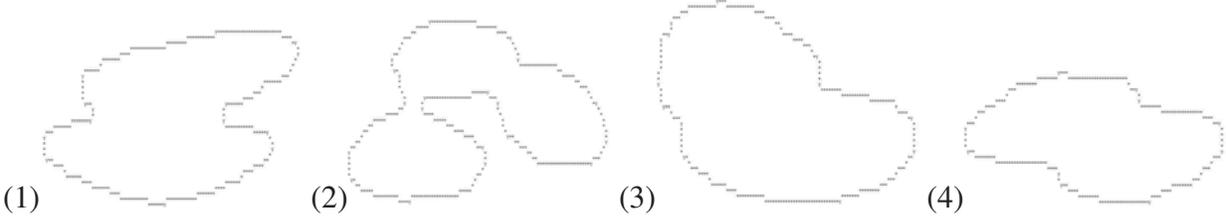


FIGURE 17. The results of smooth following and linearization for the images in Figs. 16(1-4).

“ \wedge ” \rightarrow “F” \rightarrow “[” \rightarrow “s” \rightarrow “v” \rightarrow “o” \rightarrow “)” \rightarrow “T”. It is clear that the contour shape is convex polygon, where it can be approximately defined as an ellipse.

For the outer contour in Fig. 15, the series of structural points is:

“ \wedge ” \rightarrow “F” \rightarrow “[” \rightarrow “s”(convex) \rightarrow “S” \rightarrow “]” \rightarrow “f” (concave) \rightarrow “F” \rightarrow “[” \rightarrow “s” \rightarrow “v” \rightarrow “o” \rightarrow “)” (convex) \rightarrow “(” (concave) \rightarrow “)” \rightarrow “T” (convex).

It is clear, that the above contour shape consist of of two pairs of convex and concave change.

Separation and reconstruction of touching cells

There are some images of touching cell in a frame of microscope image. We can see that there are some groups of touching cell images (see Fig. 12. There are nine groups of touching cell images in Fig. 12(1). Therefore, these groups of touching cell images make analyze and recognize the cell images of cell-cycle screening automatically very difficult. In order to result the problem, we need to determine which image is a group of touching cell images, and then find all separation points of the group of touching cell images. Finally, each touched cell image need to be reconstructed.

Morphology structures of touching cell images

Based on the prior knowledge, the cell shape of cell-cycle screening images can approximate as an ellipse before it is divided. Therefore, if two or more cells are touched, there are is one concave structural point at least on its outer contour. Also, its size is larger than that of one cell image as touching cell image consists of two or more cells. Some binarization images of touching cell images are shown in Figs. 16(1-3). Their smooth following and linearization results can be shown in Figs. 17(1-4) based on algorithms of smooth following, and linearization which are described in Sections 2 and 3. Furthermore, the structural points of the images in Figs. 17(1-4) can be extracted based on algorithm of extracting structural points (Section 4). We can see that there are some concave structural points on the contours of the images in Figs. 18(1-4). Based on the definition of structural points, one concave point means a concave change in the direc-

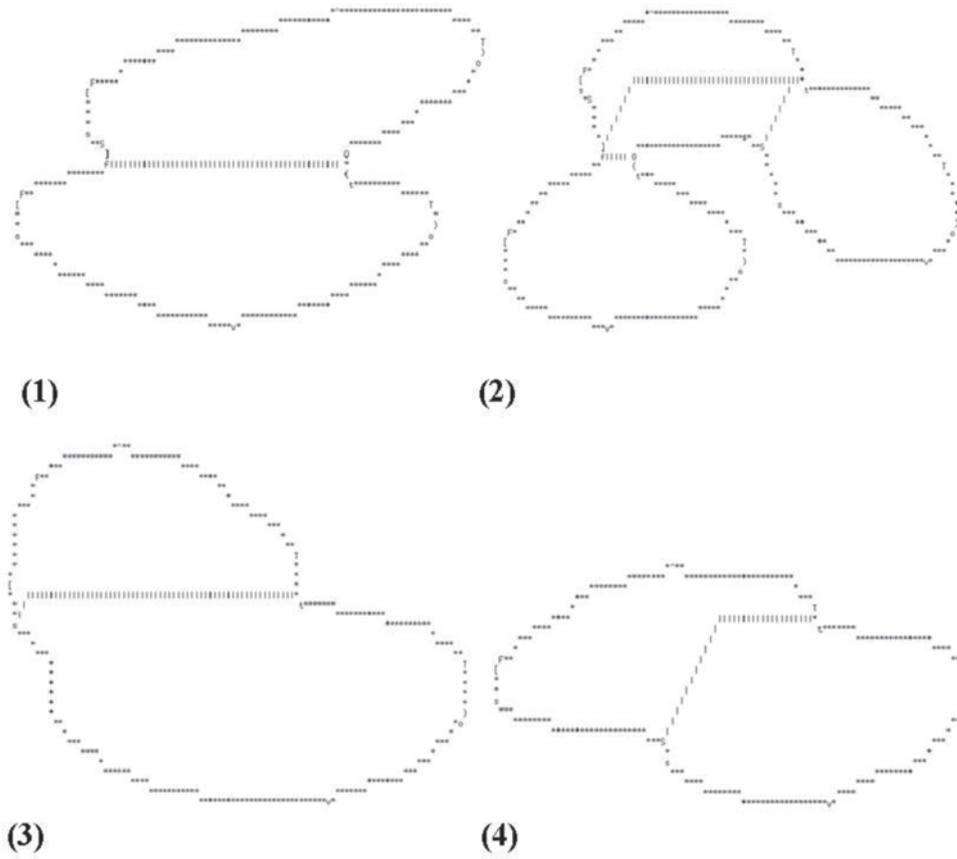


FIGURE 18. The extracting structural points of the images in Figs. 18(1-4).

tion of one chain code on the contour. Let a series of concave structural points on the outer contour of touching cell images is

$$S_{cc} = \{s_{cc}(0), s_{cc}(1) \dots s_{cc}(i), \dots s_{cc}(n-1), s_{cc}(n)\} \quad (16)$$

where $s_{cc}(i)$ is the structural point number of the i -th concave structural point on the contour, and there are n concave structural points on the contour. It is clear that $s_{cc}(i) < s_{cc}(i+1)$. In fact, one concave change on the contour may consists of several closest concave structural points. For example, if there exists $s_{cc}(i+1) - s_{cc}(i) = 1$ and $s_{cc}(i+2) - s_{cc}(i+1) = 1$, then that means one concave change consists of three concave structural points, $s_{cc}(i)$, $s_{cc}(i+1)$ and $s_{cc}(i+2)$. In this case, these three concave structural points should be merged into one group of concave structural points. After the above merging processing for S_{cc} , a series of groups of concave structural points (S_{cg}),

$$S_{cg} = \{s_{cg}(0), s_{cg}(1) \dots s_{cg}(i), \dots s_{cg}(k-1), s_{cg}(k)\} \quad (17)$$

can be found, where k is the number of groups and $k \leq n$. For example, nine concave structural points in Fig. 18(2) are merged into three groups of concave structural points. The morphological patterns of touching cell images can be determined based on the number of groups of concave structural points.

If $k = 1$ or $k = 2$, two cells are touched.

If $k = 3$, three cells are touched.

If $k = 4$, four cells are touched.

Separation points of touching cell images

The method of searching separation points can be described as follows.

Case 1 ($k = 1$):

If $k = 1$, there is one group of concave points, $s_{cg}(0)$. Suppose $s_{cg}(0)$ contains p concave points, $s_{cg0}(0), \dots, s_{cg0}(p-1)$ $p < 4$. For each concave point, find its match convex structural points which are defined as its corresponding convex structural points in the approximate reverse direction of chain code. For example, if $s_{cg0}(0)$ is concave structural point “^”, then its match convex structural points are “s”, “v” and “o”. Let the number of the corresponding match convex structural points for all $s_{cg0}(0), \dots, s_{cg0}(p)$ be q , and they are represented as $s_{cv}(0), \dots, s_{cv}(q-1)$. We can determine separation points which make minimum distance between one pair of one concave structural points in $\{s_{cg0}(0), \dots, s_{cg0}(p)\}$ and one convex structural point in $\{s_{cv}(0), \dots, s_{cv}(q-1)\}$. That is

$$\{s_{cg0}(m), s_{cc}(n)\} = \text{mini}\{|s_{cg0}(i), s_{cc}(j)|i < p, j < q\}, \quad (18)$$

where $s_{cg0}(m)$ and $s_{cc}(n)$ are selected separation points.

Case 2 ($k = 2$):

If $k = 2$, there are two groups of concave points, $s_{cg}(0)$ and $s_{cg}(1)$. Suppose the number of concave structural points in $s_{cg}(0)$ is p_0 , and in $s_{cg}(1)$ is p_1 respectively. In this case, we can determine separation points which make minimum distance between one pair of one concave structural point in $\{s_{cg0}(0), \dots, s_{cg0}(p_0)\}$ and one in $\{s_{cg1}(0), \dots, s_{cg1}(p_1)\}$. That is

$$\{s_{cg0}(m), s_{cg1}(n)\} = \text{mini}\{|s_{cg0}(i), s_{cg1}(j)|i < p_0, j < p_1\}, \quad (19)$$

where $s_{cg0}(m)$ and $s_{cg1}(n)$ are selected separation points.

Case 3 ($k > 2$):

If $k > 2$, there are more than two groups of concave points, $s_{cg}(0) \dots s_{cg}(l)$ $l > 2$. In this case, we can determine each pair of separation points which make minimum distance between each pair of one concave structural point in $\{s_{cgx}(0), \dots, s_{cgx}(p_x)\}$ and one in $\{s_{cgy}(0), \dots, s_{cgy}(p_y)\}$, where $\{s_{cgx}(0), \dots, s_{cgx}(p_x)\}$ and $\{s_{cgy}(0), \dots, s_{cgy}(p_y)\}$ are neighboring groups of concave structural points. That is

$$\{s_{cgx}(m), s_{cgy}(n)\} = \text{mini}\{|s_{cgx}(i), s_{cgy}(j)|i < p_x, j < p_y\}, \quad (20)$$

where $s_{cgx}(m)$ and $s_{cgy}(n)$ are selected separation points. For example, if $k = 3$, there are three pairs of groups of concave structural points, $s_{cg}(0)$ and $s_{cg}(1)$, $s_{cg}(1)$ and $s_{cg}(2)$, and $s_{cg}(2)$ and $s_{cg}(0)$ respectively.

Based on the above algorithm, we can find all separation points of images in Figs. 18(1-4). We can find related separation lines (see Figs. 18(1-4)) and the coordinate data of related arcs which are shown in Figs. 19(2,3), 20(2,3) and 21(2,3,4) based on these separation points and series of points of the contour. These contours of touching cell images are shown in Figs. 19(1), 20(1) and 21(1).

Reconstruction of touching cell images

We have found the coordinate data of all related arcs which are separated based on the above algorithm. As all cell shapes are approximately as an ellipse, touching cell images can be reconstructed based on these separated arcs. The reconstruction method is direct least square fitting of ellipses [24]. The the reconstructed cell images are shown



FIGURE 19. The contour, separated arcs and reconstructed ellipses of sample touching cell image 1 in Fig. 16(1).



FIGURE 20. The contour, separated arcs and reconstructed ellipses of sample touching cell image 2 in Fig. 16(2).

in Figs. 19(4,5), 20(4,5) and 21(5,6,7) based on the coordinate data of separated arcs respectively.

Dynamic analysis and recognition of different cell phases

In order to detect the morphological structures of cells, it is necessary to recognize the cell shape of different cell phases. We can see that there mainly two classes of cell shapes, ellipse and barbell, for our cell screening. The ellipse shapes can be three types, skew, horizontal and vertical based on relation between their major and minor axes. It is clear that here “skew”, “horizontal” and “vertical” are fuzzy. Some recognition models of cell shapes can be described as following based on their morphological structures:

If a cell shape is a ellipse, there are no concave structural points on the outer contour of the cell contour. Furthermore, four models of ellipse can be described as following. Calculate number of group of codes, (codes 0, 4, 5 and 1), (codes 5, 1, 2 and 6), (codes 6, 2, 7 and 3), (codes 7, 3, 0 and 4), on the outer contour of the cell image contour respectively.

Morphological model 1: Ellipse shapes $e_{(5,1,2,6)}$ and $e_{(6,2,1,5)}$.

For these shapes, the number of group of codes, (codes 5, 1, 2 and 6), is largest.

Let $c_{5,6,1,2}$ be the total number of codes 5, 6, 1 and 2, c_t be the total number of all codes, $c_{5,1}$ be the total number of codes 5 and 1, and $c_{6,2}$ be the total number of codes 6 and 2, on the outer contour of the cell image respectively. If (1) the above Condition is met;



FIGURE 21. The contour and separated arcs of sample touching cell image 3 in Fig. 16(3).



FIGURE 22. The reconstructed ellipses of sample touching cell image 3 in Fig. 16(3).

(2) its outer contour mainly consists of chain codes 5, 6, 1 and 2 ($c_{5,6,1,2} \geq \frac{1}{2}c_t$); (3) the number of chain codes 5 and 1 is more than that of chain codes 6 and 2 ($c_{5,1} \geq c_{6,2}$), then the cell image shape is recognized as the shape $e_{(5,1,2,6)}$, otherwise ($c_{5,1} < c_{6,2}$) the cell image is recognized as the shape $e_{(6,2,1,5)}$.

In this model, the cell image shape, $e_{(5,1,2,6)}$, is a skew ellipse in the direction of code 5 or 1, and the cell image shape, $e_{(6,2,1,5)}$, is a vertical ellipse mainly.

Morphological model 2: Ellipse shapes $e_{(7,3,0,4)}$ and $e_{(0,4,3,7)}$.

For these shapes, the number of group of codes, (codes 7, 3, 0 and 4), is largest.

Let $c_{7,0,3,4}$ be the total number of codes 7, 0, 3 and 4, c_t be the total number of all codes, $c_{0,4}$ be the total number of codes 0 and 4, and $c_{7,3}$ be the total number of codes 7 and 3, on the outer contour of the cell image respectively. If (1) the above Condition is met; (2) its outer contour mainly consists of chain codes 7, 0, 3 and 4 ($c_{7,0,3,4} \geq \frac{1}{2}c_t$); (3) the number of chain codes 7 and 3 is more than that of chain codes 0 and 4 ($c_{7,3} \geq c_{0,4}$), then the cell image is recognized as the ellipse shape $e_{(7,3,0,4)}$, otherwise ($c_{7,3} < c_{0,4}$) the cell image is recognized as the shape $e_{(0,4,3,7)}$.

In this model, the cell image shape, $e_{(7,3,0,4)}$, is considered as a skew ellipse in the direction of code 7 or 3, and the cell image shape, $e_{(0,4,7,3)}$, a horizontal ellipse.

Similarly, model 3, $e_{(6,2,7,3)}$ and $e_{(7,3,2,6)}$, and model 4, $e_{(5,1,0,4)}$ and $e_{(0,4,5,1)}$, can be defined respectively. **Morphological model 5:** Barbell shapes.

If (1) there are two concave structural changes; (2) there is one pair of corresponding concave structural points, " \wedge " and " $\$$ " (horizontal), "]" and "(" (vertical), "f" and "O" (skew), or "t" and "S" (skew), then cell image contour can be recognized as the barbell shape. There are four types of barbell shapes which can be defined based on which pair of concave structural points is found. The cell image contour in Fig. 15(7) can be recognized as a barbell shape, and its pair of corresponding concave structural points are structural points "]" and "(".

Also, other morphological structures of cell images can be described.

Similarly, other morphological structures of cell images in cell-screen can be constructed. Based on the above models, morphological structures of different cell phases can be analyzed as follows:

(1) if (a) the cell shapes are a ellipse shape; (b) the number of points in cell outer contour is large enough (threshold can be found from statistical analysis), then the cell is a normal changing one and its changing trend is that the rate between major and minor axes of the cell image is decreased with time (metaphase).

(2) if the rate between major and minor axes of the cell image is large enough (threshold can be found from statistical analysis), then the cell will be split soon (telophase).

(3) if the cell shapes are a barbell shape, then then the cell will be split (telophase).

(4) if (a) the cell shapes are a ellipse shape; (b) the rate between major and minor axes of the cell image is about 1:1; (c) the number of points in cell outer contour is little enough (threshold can be found from statistical analysis), then the cell is a newly changing one, and its changing trend is that the number of points in cell outer contour is increased with

time (prophase).

Based on the above analysis, the cells in Figs. 15(1-5) are at metaphase, the cells in Figs. 15(6-7) are at telophase soon, and the cells in Fig. 15(8) are at prophase.

In this way, cell-screening can be analyzed and recognized dynamically and automatically.

CONCLUSION

The color reduction algorithm is used to extract color feature of color biological images. Smooth following techniques are developed to smooth contours of images. The spurious pixels (points) of contours in images can be removed based on the structural patterns of difference codes and of spurious point groups. The algorithm can let the abstract value of difference codes between neighboring points be less than 2. This makes the linearization of contours possible.

A new and efficient algorithms about linearization smoothed contours is discussed. A series of linearized lines in a contour is formed based on element chain codes which are found in terms of difference chain codes. a series of description features of contours is formed, but not by other methods.

The morphological structural points are proposed. The morphological structural points are defined to describe morphological structures between neighboring linearized lines along contours. Extraction of structural points is based on structural patterns which are determined by element chain codes. Two applications are used to discuss how to use related features such as color, morphological structures and prior information for description, recognition and analysis of biological images.

These efficient and new algorithm have been developed for dynamic description, recognition and analysis of cell-cycle in detail. The algorithm of extracting structural features (structural points) is described based smooth followed contour, linearized line and difference chain codes. The morphological structures of cell contour shapes are constructed based on the structure points, difference chain code of linearized line and geometry features. The dynamic analysis and recognition of cell-cycle are approached based on morphological structure models and prior knowledge of cell images. The best useful contribution is that some series of structural features of linearized lines of biological image contours are found and morphological models of cell contour shapes are described based on our algorithm, but other methods not. Our method is efficient and new because morphological structure models of biological images are constructed, and these models simulate artificial intelligence.

Acknowledgement: The project is partially supported by ARC LP0347156 and DP0559647.

REFERENCES

1. S. Fox, "Accommodating cells in HTS", *Drug Discovery World*, **5**, pp. 21–30, 2003.
2. Y. Feng, "Practicing cell morphology based screen", *European Pharmaceutical Review*, **7**, pp. 7–11, 2002.

3. R. Dunkle, "Role of image informatics in accelerating drug discovery and development", *Drug Discovery World*, **5**, pp. 75–82, 2003.
4. J. c. Yarrow, et al., "Phenotypic screening of small molecule libraries by high throughput cell imaging", *Comb Chem High Throughput Screen*, **6**, pp. 279–286, 2003.
5. Y. Hiraoka, and T. Haraguchi, "Fluorescence imaging of mammalian living cells", *Chromosome Res.*, **4**, pp. 173–176, 1996.
6. X. Chen, X. Zhou, and S.T.C. Wong, "Automated segmentation, classification, and tracking cancer cell nuclei in time-lapse microscopy", *IEEE Trans. on Biomedical Engineering*, in press.
7. T.D. Pham, D. T. Tran, X. Zhou, S.T.C. Wong, "An automated procedure for cellphase imaging identification", *Proc. AI-2005 Workshop on Learning Algorithms for Pattern Recognition*, pp. 52–29, 2005.
8. T.D. Pham, D.T. Tran, X. Zhou, and S.T.C. Wong, "Classification of cell phases in time-lapse images by vector quantization and Markov models", *Neural Stem Cell Research*, ed. Erik V. Greer, Nova Science, New York, 2006.
9. J. Hu and H. Yan, "Structural primitive extraction and coding for handwritten numeral recognition," *Patt. Recog.*, **31**, (5), pp. 493-509, 1996.
10. M. Shi, Y. Fujisawa, T. Wakabayashi and F. Kimura, "Handwritten numeral recognition using gradient and curvature of gray scale image," *Patt. Recog.*, **35**, pp. 2051-2059, 1997.
11. H. Yan, "Prototype optimization of a nearest neighbor classifier using a multi-layer neural network," *Patt. Recog.*, **26**, pp. 317-324, 1993.
12. H. Yan, "Handwritten digit recognition using optimized prototypes," *Patt. Lett.*, **15**, pp. 207-211, 1994.
13. A. Malaviya and R. Klette, "A fuzzy Syntactic method for on-line handwriting recognition", *In Advance in Structural and Syntactical Pattern Recognition*, (P. Perner, P. Wang, and A. Rosenfeld, Eds.), Springer, 1996.
14. S. W. Lee, "Off-line recognition of totally unconstrained handwritten numerals using multilayer cluster neural networks", *IEEE Trans. Pattern Anal. Mach. Intell.*, **18**, (6), pp. 648-652, 1996.
15. N. Papamarkos, C. Strouthopoulos and I. Andreadis, "Multithresholding of Colour and Gray-level Image Through a Neural Network Technique," *Image and Vision Computing*, **18**, pp. 213-222, 2000.
16. N. Katzir, M. Lindenbaum and M. Porat, "Curvature Angle Segmentation under Partial occlusion," *IEEE Trans. Pattern Analysis Mach. Intell.*, **16**, (5), pp. 513-519, 1994.
17. D. Yu and H. Yan, "An efficient algorithm for smoothing binary image contours," *Pro. of ICPR'96.*, **2**, pp. 403-407, 1996.
18. D. Yu and H. Yan, "Separation of Touching Handwritten Multi-Numeral Strings Based on Morphological Structural Features," **34**, (3), pp.587 -599, 2001.
19. F. Moktarian, A. K. Mackworth, "A Theory of Multiscale Curvature-Based Shape Representation for Planer Curvature Angles. *IEEE Trans. Pattern Analysis Mach*", *Intell*, **14** (8), pp. 789–805, 1992.
20. A. M. N. Fu, H. Yan, "Huang K. A Curvature Angle Bend Function Based Method to Characterize Contour Shapes", *Patt. Recog*, **30** (10), pp. 1661–1671, 1997.
21. M. Sonka, V. Hlavac, and R. Boyle, "Image Processing, Analysis and Machine Vision", *Chapman & Hall Computing*, Cambridge, 1993.
22. Y. Yang, H. Yan, "An adaptive logical method for binarization of degraded document images", *Pattern Recognition*", **33**(5), pp. 787–807, 2000.
23. N. Ostu, "A thresholding selection method from gray level histogram", *IEEE Trans. Systems Man Cybernet*, **SMC 8**, pp. 62-66, 1978.
24. A. Fitzgibbon, M. Pilu, and R. B. Fisher, "Direct Least Square Fitting of Ellipse", *Term Analysis and Machine Intelligence*, **21**, pp. 476-480, 1999.

Copyright of AIP Conference Proceedings is the property of American Institute of Physics and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.