[UCI]I804:24011-200000242123

2011 년 8 월

박사학위 논문

# Automatic Recognition of Apoptotic cells in Fluorescence Microscopic Image

조선대학교 대학원

정보통신공학과

유 해 릉

# Automatic Recognition of Apoptotic cells in Fluorescence Microscopic Image

# 형광현미경 영상에서 아폽토시스 세포들의 자동화된 인식

2011 년 8 월 25 일

조선대학교 대학원

정보통신공학과

유 해 릉

# Automatic Recognition of Apoptotic cells in Fluorescence Microscopic Image

지도교수 신영숙

이 논문을 공학 박사학위신청 논문으로 제출함

2011 년 4 월


조선대학교 대학원

정보통신공학과

유 해 릉

# 유 해 룡의 박사학위논문을 인준함

위원장 전북대학교 교 수 오일석 (인)

위 원 조선대학교 조교수 권구락 (인)

위 원 조선대학교 조교수 한송이 (인)

위 원 조선대학교 부교수 김태형 (인)

위 원 조선대학교 부교수 신영숙 (인)

2011 년 6 월

조선대학교 대학원

# Table of Contents

# List of figures

# List of tables

# 요약

## 형광현미경 영상에서 아폽토시스 세포들의 자동화된 인식

유해릉
지도교수: 신영숙
정보통신공학과
조선대학교 대학원

컴퓨터 기술의 진보로 디지털 의료 영상들은 임상 진단 및 치료에 중요한 역할을 담당한다. 본 연구는 아폽토시스 이미지 전처리 기법, 이미지 세분화 기법, 특징 추출 기법 및 자동 분류 및 인식 기법을 포함하는 형광 현미경 영상을 사용하여 자동 인식에 사용되는 주요 기술들에 중점을 둔다. 각 기술 설명에서 사용된 현존하는 방법들을 검토하였으며, 아폽토시스자동 인식에서 좋은 성능을 이루는 몇가지 새롭고 강인한 처리 알고리즘들을 제안하였다.

영상 전처리 단계에서 현미경 영상안에서 영상의 질 저하 요소들과 잡음의 일반적인 유형들을 분석하였다. 영상의 특성에 따라 그라디언트 기반 이방성 필터링 알고리즘을 제안하였으며, 이것은 배경 잡음 및 펄스 잡음을  필터링 할

수 있으며 더불어 윤곽선 정보를 유지할 수 있다.

영상안에서 작은 조각들에 의한 조악한 성능을 해결하기 위해 잡음 검출 방법에 기반된 정보 엔트로피와 적응형 중간 필터링 알고리즘이 제안되었다. 마지막으로, 윤곽선 개선 알고리즘에 기반된 이웃 화소 대비가 영상에서 흑백 대비를 개선하기 위해 제안되었다. 영상 분할 단계에서 현존하는 일반적으로 사용된 기술들이 검토되었으며, 스네이크 모델의 개념과 응용, 개선된 방법들이 상세하게 기술되었다. GVF 모델은 초기 윤곽선 위치가 좀 더 자유로우며 오목한 경계선을 검출할 수 있는 성능으로 다른 스네이크 모델보다 큰 장점을 가진다. 그러나 광원주위 고리의 검출 및 낮은 대비 경계선과 잡음에 노출된 경계선을 검출하는 데는 약점이 있다. 이 문제를 해결하기 위하여, 스네이크의 움직임을 제한하는 크기-기반 GVF 스네이크 모델을  제안하였다. 스네이크 모델은 분리된 세포들의 경계선과 중첩된 세포들의 경계선을 검출하기 위해 사용되었으며, 연속적인 분리 과정이 상세하게 기술되었다.

특징 추출 단계에서 기존 주요 기법을 검토 하고 본 실험에서 사용된 다양한 특징들을 기술하였다.

인식 단계에서는 현미경 영상 인식에 대한 기본적인 구조들을 소개하였다. 배깅 의사 결정 트리 알고리즘이 10 특징들과 분류의 정확성을  측정하기 위해 사용되었다. 마지막으로, 실험 결과들이 비교 되었으며 아폽토시스 인식에 적용된 제안된 알고리즘들에 대한 우수성을 기술하였다.

# I. Introduction

## A. Research Bachground

Clinical examination is an important basic disease diagnosing technique in medicine. It analyzes blood, body fluid, secretion through organoleptic inspection, physiological inspection, biochemical inspection or microscopic inspection, observes the change of state and diagnoses the disease. Microscopic image is a key method for clinical examination, which observes and analyzes the change of numbers, morphologies and states of objects in image to justify the physiological diagnosis. Microscopic image based diagnosis went through three main stages of development: (1) manual observation. In this stage, from the manipulation of microscope, the acquisition of experimental data, the analysis of data to the acquisition of final diagnosis are completely finished manually; (2) Semiautomatic observation. In this stage microscope manipulation, data acquisition can be processed automatically but the data analysis and final diagnosis still have to be implemented manually (3) automatic observation. In this stage, the whole process metioned before can be implemented automatically. It is evident that fully automatic diagnosing technique is quite laborsaving and objective compared to previous ones. Hence, this is an active research topic nowadays and many scientists have made some initiatory achievements, however, no mature system appears. In this thesis, I will propose a series of improvements of key techniques for the automatic recognition and classification

of objects in microscopic images. Although I use apoptosis as my target object in this thesis, these algorithms are also applicable to other smilar experiments.

Apoptosis, or programmed cell death, is a highly regulated process that allows a cell to self-destruction in order to eliminate unwanted or dysfunctional cells in the body. Light and electron microscopes have identified various morphological changes that occur during apoptosis [1]. In the early process of apoptosis, cell shrinkage and pyknosis are visible [2]. During cell shrinkage, the cells are smaller, the cytoplasm is dense, and the organelles are more tightly packed. Pyknosis is the result of chromatin condensation, and this is the most characteristic feature of apoptosis. Extensive plasma membrane blebbing occurs followed by karyorrhexis and the separation of cell fragments into apoptotic bodies during a process called “budding”. Apoptotic bodies consist of cytoplasm with tightly packed organelles with or without a nuclear fragment.

In clinical applications and research, two major problems were addressed in manually detecting and counting apoptosis: (1) The discrimination between live cancer cell and apoptosis; (2) The estimation of the number of dead cells. The first problem is due to the complexity of microscopic images and the limited experience of biomedical professionals. Inherent variability of many factors such as environment illumination conditions, dye duration, film thickness and film inhomogeneities will result in different image luminance and color distribution, which will greatly increase the complexity of images [3]. That is why some new imaging techniques are used nowadays such as holography images that give up using staining [4]. Considering the difference of experience and subjectivity of

specific biomedical professionals, accurate discrimination between a live cancer cell and apoptosis is a very challenging task. The second problem furtherly increases the difficulty in that it is very difficult to estimate the number of dead cells even each single apoptotic fragment has been detected precisely. Usually, this procedure is implemented by esitimation by naked eyes. The two problems mentioned above greatly reduce the accuracy of manual counting of apoptosis. In our experiments, apoptotic fragments appear as small, condensed, and bright particles. I first detect the apoptotic fragments according to the techniques of image processing and pattern recognition, and then try to group the apoptotic fragments into one intact cell using the average size of live cancer cell, which stands for one dead cell induced by the anticancer drug, Oxaliplatin, in our experiments. The key procedure is the accurate recognition of live cancer cell and apoptotic fragments.

## B. Key Techniques Research Status of Microscope Imaging

In order to achieve automatic recognition and analysis of apoptosis, several key techniques should be emphasized. The common framework of an image based recognition system is described by Fig.1

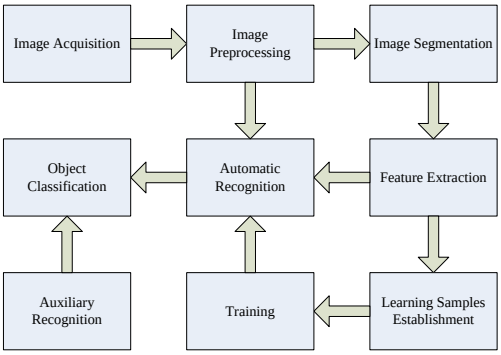


Fig 1 Common framework of image-based recognition system

This paper will focus on the techniques of microscopic image preprocessing, single cell segmentation, microscopic image feature extraction and automatic classification.

## 1. Microscopic image preprocessing research status

The visual inspection of specimens is one of the most common techniques used for learning and diagnosis in medicine. Due to the influence of many external factors in image acquisition process, images are always contaminated by noises. There are also several know factors that cause spatial intensity heterogeneity in microscopic images, such as (a) photo bleaching. (b) fluorescent attenuation. In order to improve the quality of images and to remove the influence of noises, many image filtering techniques and contrast enhancement algorithm are researched.

### a. Microscopic image filtering research status

Given the complexity of factors introducing noise and intensity heterogeneity, there are two approaches for improving the quality of microscopic images. First, one could focus on improving specimen preparation and imaging conditions, for example improved fluorescent dyes. Second, one could attempt to correct pixel intensities after image acquisition, as it is the case for image restoration. Apart from some traditional filtering algorithm such as self-adaptive median filtering algorithm [5,6], hybrid filtering algorithm[7], anisotropic filtering algorithm [8], feature fusion filtering algorithm [9,10,11],wavelet based filtering algorithm

[12,13,14,15] etc, some specific filtering and restoration algorithms were also proposed such as the empirical correction methods for intensity loss [27], constant thresholding[28], iterative correction methods [29], 2D histogram or estimations of intensity decay function [30]. Although these methods work well for certain category of microscopic image, most of them require the image meet some basic requirements such as the photo bleaching is a spatially homogeneous in a lateral plane, which greatly limit the universality and availability.

### b. Microscopic image enhancement research status

In the past, histogram equalization (HE) [24-26] was most commonly used in microscopic images. HE leads to a uniform global intensity distribution in output image. However, it cannot effectively enhance local intensity variation due to its global property. To address this problem, adaptive histogram equalization (AHE) has been used to adjust intensity variation locally by computing local histograms with spatially different windows [31]. A major problem with AHE is high sensitivity to noise, which results in amplification of undesired noise values. An improved approach to adjust local intensity variation is contrast limiting adaptive histogram equalization (CLAHE) [32]. It reduces noise amplification due to AHE by setting clipping limits and so removes boundary artefacts by background. Other image enhancement algorithm used for microscopic image includes neighborhood-based contrast enhancement algorithm [16-18], Wiener filtering algorithm [19, 20], blind deconvolution algorithm [21-23] and histogram equalization algorithm [24-26]. However neighborhood-based contrast algorithm is widely used in

application because it is simple to implement and can restrain noise during enhancement.

## 2. Microscopic image segmentation research status

In microscopic images, the most demanding procedure is segmentation, which separates our target objects in image, such as nuclei, cytoplasm, certain interior structure, fragments of apoptosis etc. Properly identifying objects (nuclei) that are well dispersed, non-confluent, and bright relative to the background is straightforward by applying a simple threshold to the image such as Otsu [43] and Mixture of Gaussians [44], traditional edge detector [45-49]. This can obtain satisfactory results but usually fails when nuclei are touching. However, for most microscopic images, at least some nuclei are touching. The traditional algorithms used for touching cell sepration include watershed algorithm [32], morphological segmentation [33], segmentation combining intensity, edge and shape information [34] etc. In recent years, some other new algorithms such as gradient flow tracking [35], graph-cut method [36], concave points and ellipse fitting [37] etc were proposed. No matter what method you select, it always involves procedures like parameter selection and rules setting. Every segmentation method is always applicable to a specific kind of image and there has been no all-purpose segmentation algorithm. In this thesis, I focused on the introduction of a GVF snake model with area constraint and its application in detecting the boundary of a single cell, which should be robust to complex noise, uneven distribution etc.

## 3. Microscopic feature research status

The effective extraction of robust features is the basis of automatic recognition of targets in microscopic images. Features can be devided into three categories: shape feature, texture feature and color feature. Color feature can also be regarded as a kind of special texture feature.

Commonly used shape descriptors include moment description method [50-52], polygonal fitting method [53-55], edge singular point descriptior [56, 57], and Fourier coefficient descriptor [58, 59]. Hu first proposed the concept of moment, applied it to shape recognition and inferred a series of basic properties. Moment has some properties like translation invariance, scale invariance and rotational invariance. Teague came up with the concept of orthogonal moment, which can construct arbitrarily high order moment and is an integral computation, robust to noise [51]. Shen proposed the concept of wavelet moment method [60]. Polygonal fitting computes similar polygon on object boundary through which to describe the object shape. Edge singular descriptor designs an algorithm to look for angular points on boundary and describe the object shape by computing the distribution of angular points. Fourier coefficient method is a classical shape descriptor, which computes the Fourier transform of points on boundary and describes the shape.

Commonly used texture descriptor includes moment invariants descriptor [61-63], graph theory based topology analysis [64, 65], texture spectrum descriptors [66-68], transform domain besed descriptor [69, 70] and neural networks based method [71, 72]. Moment invariants are widely used that include

cooccurrence matrix, Zernike moment etc. Topology analysis computes the internal granulometric distribution to describe object texture. Texture spectrum is a newly developed texture description method that enhances the object textures through a series of transform, obtain the texture spectrum and compute the texture features using the spectrum. Transform domain used spatial transform, frequency transform and wavelet transform to describe texture in a transform domain. Neural networks method use training to learn the characteristic of texture feature and then ouput description methods of texture. In this thesis, I analyze and choose a series of robust features after obtaining the properties and characteristics of cells.

## 4. Pattern recognition and classification research status

Pattern recognition and classification is an assignment of input samples to be recognized into a given set of class according to its properties and characteristics through a set of rules. There are many recognition methods at present, which includes statistical pattern recognition, syntactic pattern recognition, fuzzy theory based pattern recognition, and neural networks based pattern recognition.

Statistical pattern recognition: this kind of techniques has very mature theory and many approaches are usually effective. Most commonly used techniques include clustering analysis, statistical decision, nearest neighbor method. Commonly used algorithms include C-Mean clustering [73-76], K-Nearest Neighbor [77-80] and decision tree [81, 82].

Syntatic pattern recognition [83]: also called structural pattern recognition, decomposes the object into several basic units. We can describe the objects

using these units and their interrelationship by character string and graph. Then we analyze them using formal language theory and classify them into a specific category if they correspond to the grammer of that category.

Fuzzy theory based pattern recognition [84, 85]: this technique applies fuzzy mathematics theory to solve the problem of pattern recognition, which is especially applicable to the situation of classification fuzziness. The effectiveness depends on the goodness of membership function.

Neural networks [86-88]: Neural networks utilize the internal connections of a large number of single basic units-neurons to compose a complicated nonlinear dynamic system, which possess some certain characteristics of biological neural networks and have strong self-studying ability, self-organizing ability and fault-tolerant ability. It can be used for association, recognition and decision. A dramatically difference between neural networks and the above mentioned ways is that neural networks has automatic feature extracting ability during studying process.

## C. Main Difficulties in Research

### 1. Preprocessing step

The main problem is how to establish a quality degradation model of noise pollution, how to solve the contradiction between image filtering and image enhancement and how to improve processing speed. In my research, the accurate measurement of the size distribution of live cell and apoptosis is vital to the success of the accurate recognition. How to measure this size precisely is

another difficulty.

## 2. Segmentation step

Every segmentation algorithm has its own speciality and is only applicable to specific situation. The characteristics of microscopic images, complex background, big individual differences and uneven distribution etc, all increases the difficulties greatly. It is impossible to obtain a good result if a global segmentation algorithm is used. How to select and design an appropriate segmentation method and how to set parameter automatically is still a great difficult in research.

## 3. Feature extraction step

Feature extraction comes after image segmentation. According to the above analysis, it is known that 100% percent segmentation is impossible, so the extractiong of robust features against noise and deformation is dramatically important.

Finding the best and most robust features that have the maximum discrimination power between classes is vitally important to the results of the following classification step.

## 4. Pattern Recognition step

What recognition technique to choose, how to establish a training library and how to improve training convergence rate and recognition accuracy are key problems in this step.

## D. Materials Used In the Experiments

### 1. Cell culture and drug treatment

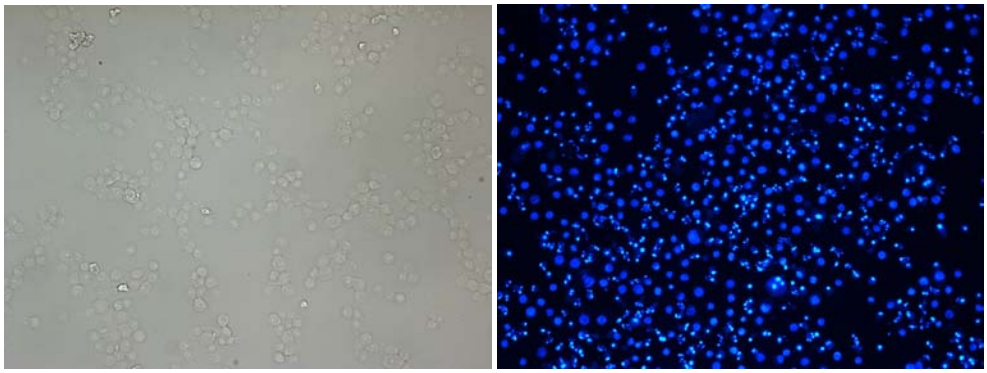
HCT116 human colon adenocarcinoma cell line obtained from American Type Culture Collection (Manassas, USA) was grown in RPMI 1640 medium (Invitrogen, USA) supplemented with 10% (v/v) fetal bovine serum (FBS; Invitrogen) and 1% penicillin-streptomycin (Welgene, Korea) in a 37°C humidified incubator in an atmosphere of 5% CO2. Drug treatment of cells was performed by adding 50 μM oxaliplatin (L-OHP; Boryung Pharmaceutical, Korea) or vehicle (for control sample) to the culture medium and incubating for 48 h.

### 2. Hoechst 33342 (HO) staining

As nuclear condensation and fragmentation is best well-known features of apoptosis and this is also very simple and easy way to detect apoptosis. Therefore, it is a well-established method widely used in the detection of apoptosis that the nuclei are stained to observe their alteration. In our experiments, the cells were incubated with 1 μg/ml HO for the final 10 min of drug treatment in the 37°C incubator and then, both floating and attached cells were collected by centrifugation. The pooled cell pellets were washed with ice-cold phophate-buffered saline (PBS), fixed in 3.7% formaldehyde on ice, washed again with PBS, and a fraction of the suspension was centrifuged in a cytospinner (Thermo Shandon, Pittsburgh, PA).

## 3. Microscope and camera

The slides were air-dried, mounted in an anti-fade solution, and images were analyzed using a DM5000 fluorescence microscope (Leica, Germany) at excitation/emission wavelengths of 340/425 nm. For images capture, microscope connected Leica DFC480 camera and Application Suite software were used. The original image size captured is 2560 × 1920. Fig.2 shows the light image and the Hoechst image in our experiment. In Fig.2 (b), it is easy to see that: (1) live cancer cells are of low brightness, big size and circular shape and apoptotic fragments are of high brightness, tiny size and near-circular shape.



(a) The reference light image (b) The Hoechst image

Fig 2 The reference light image and the Hoechst image

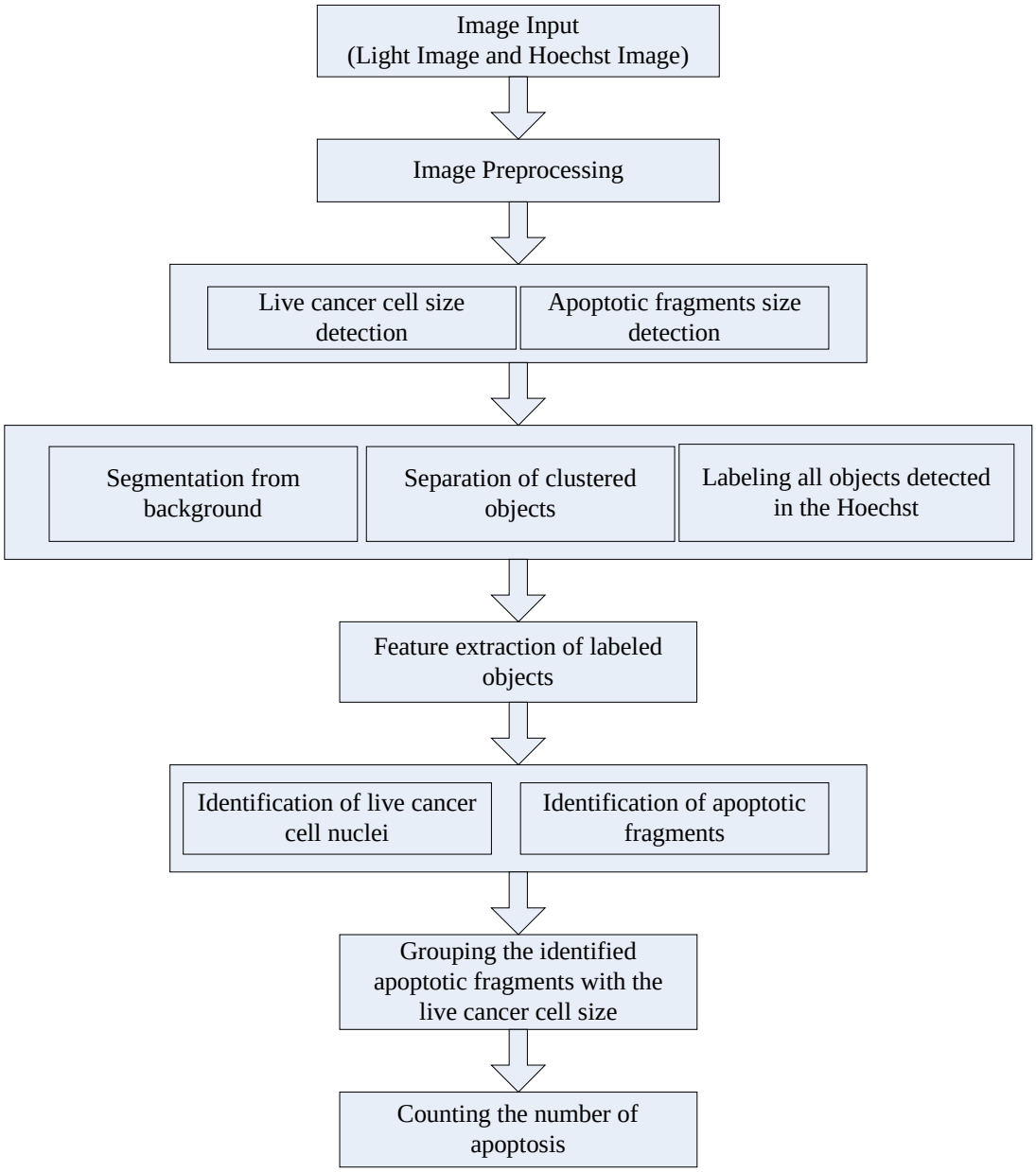



Fig 3 System Framework

# II. Image Preprocessing

In image acquisition process, the quality of microscopic images will be degraded by electrical noise, quantizing noise, light illumination etc. Hence, image preprocessing is necessary and important to improve the quality. In this thesis, I first analyzed the quality degradation factors of microscopic images and types of noise, and then I proposed a basic procedure of preprocessing of microscopic images. In order to overcome the influence of background noise and pulse noise, a gradient-based anisotropic filtering algorithm was proposed, which can filter out the background noise while preserve object boundary effectively. Due to the bade performance of gradient-based method to tiny particles in image, an information engropy based noise detection and adaptive median filtering algorithm was addressed to solve this problem. Lastly, a local contrast imformation based enhancement method was proprosed to improve boundary contrast.

## A. Noise in Medical Images

All imaging modalities, but especially those that are relevant for medical imaging, generate image noise, whether due to stability of a low-flip angle MRI acquisition [38],ultrasound speckle [39], quantum noise in an X-ray [40] or out of field counts in a PET scan [41].Virtually all imaging system also perform filtering on the image acquisition data both at an electronic level prior to reconstruction as well as during the image reconstruction phase. Indeed, much

recent advancement in reconstruction techniques for 3D imaging focus on including noise removal as part of the reconstruction optimization process[42].

## B. Quality of Microscopic Image

The optical microscopy imaging system uses light to convert a molecular distribution (the specimen) into a representative distribution of intensities (the image). In that sense, neglecting optical artifacts, any measurable difference in image intensity is related to information about specimen. When a specimen is probed with light, the physics of light gathered by lenses and the damaging effects of irradiation fundamentally limit the quality of light available for image formation and the spatial dimensions over which intensity differences can be discriminated. The limitation of the quality of the microscopeic images can be attributed to image intensity levels, image resolution, image pixel size and dynamic range. The former two factors are related to optical system and the latter two factors are related to detector system. Fig.4 illustrates the influences of image quality by these factors.

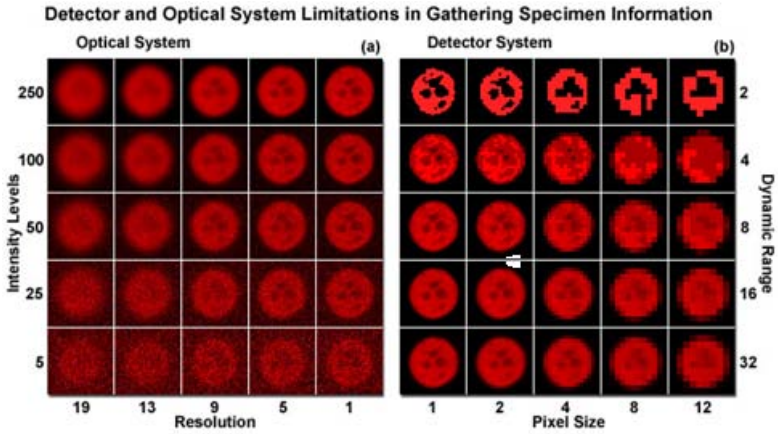


Fig 4 Quality limitation of microscopic image

The specimen is a nucleus of a human cervical carcinoma cell. We can see that as light intensity increases (Left: bottom to top), the photon counting noise becomes less apparent. Similarly, as resolution increases (Left: left to right), the general features of the nucleus appear more distinct. Note how intensity and resolution work together. In the right figure, the detector system (camera) takes the information captured in the microscope and samples the data into discrete light levels (bottom to top) and spatial elements (left to right). As pixel size increases (Right: left to right), feature details are lost until the identity of the image is lost. Similarly, as the number of intensity levels (dynamic range) decreases (Right: bottom to top), shading is lost and only highly contrasted features remain. Note that the two factors work together in providing information.

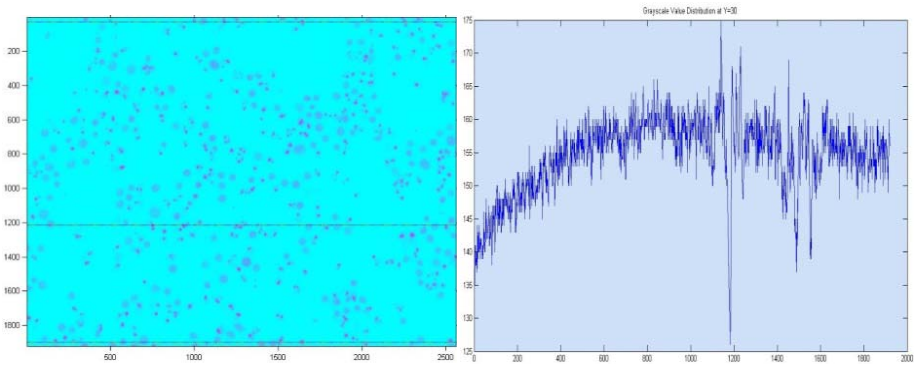
## C. Noise in Microscopic Images

### 1. Quality degradation analysis

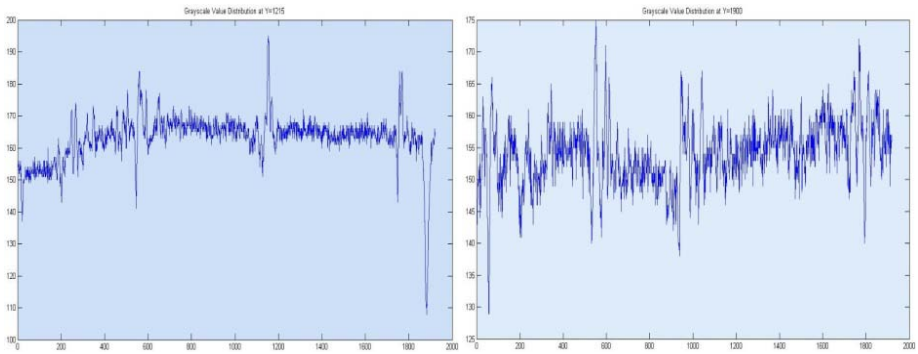
Due to the influence of image quantization, adding quantization noises exist in microscopic images and the whole image background shows nonuniform distribution. In Fig.5, I computed the gray level distribution of three places indicated by three horizontal scanning lines in positions Y=30, 1215 and 1900. It is evident that the gray level distribution in image is not a constant and some vibrating variations exist, which is smaller in background regions and larger in boundary regions.

In microscopic images, objects are always in motion. Due to the amplification

effect of microscope, the image is sensitive to focal length and usually the images appear blurry near the boundaries and halo-type boundaries appear. In this case, the precise detection of boundary becomes very difficult. Figure.6 demonstrates three typical morphologies of cells in the image, which all have blurry and halo-type boundaries.



(a) Scanning lines (b) Y=30



(c) Y=1215 (d) Y=1900

Fig 5 Noise Distribution Graph

▪

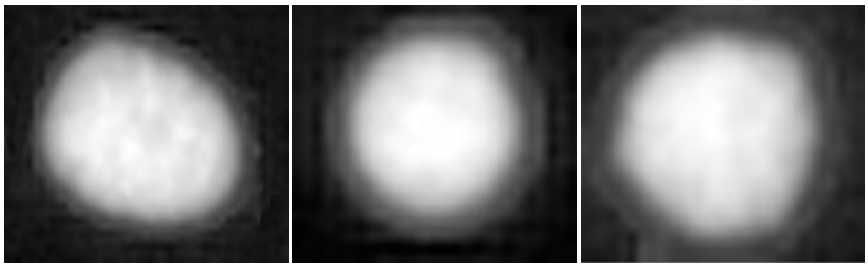


Fig 6 Cells with Blurry Boundaries

Because microscopic images are fluid images, apart from the objects, there exists some other objects like bubbles, sediments etc.

## 2. Quality degradation model

Normally in image denoising techniques, noise can be ideally divided into adding noise and multiplicative noise. Suppose the degraded image as $g$, the ideal image as $f$, and the noise as $n$. The adding model and the multiplicative model can be described in formula (2-1) and (2-2). Background noise in microscopic images can normally be regarded as adding Gaussian noise and low magnification lens tiny particles as adding pulse noise.

$$g = f + n \tag{2-1}$$

$$g = f + \mathrm{f} \times n \tag{2-2}$$

Besides noises, there are some other quality degradation factors like inappropriate focusing, uneven illumination etc. For better cell segmentation and recognition, the system should design good preprocessing algorithm to overcome noise as well as other factors.

The classical filtering algorithm includes linear spatial filtering and nonlinear

spatial filtering. Linear spatial filtering consists of multiplying each pixel in the neighborhood by a corresponding coefficient and then summing up the results to obtain the response at each point $(x, y)$. Nonlinear spatial filtering is also based on neighborhood operations, however, different from linear spatial filtering operation, which is based on linear operation, computing the sum of products, nonlinear spatial filtering is based on nonlinear operations involving the pixels of a neighborhood. Common filtering algorithm includes mean filter and median filter. In this paper, considering the characteristic of microscopic image, I researched on the adaptive gradient-based and anisotropic diffusion equation based adaptive noise-filtering algorithm. Considering the influence of tiny particles to gradient-based filtering, an information entropy based noise detection and adaptive median filtering algorithm was discussed as well.

## D. Noise Equalization Method in Medical Images

Most initial attempts at removing image noise focus on “smoothing” the pixel or voxel data by performing some sort of local averaging function. For example, Gaussian smoothing is an easily implemented smoothing algorithm; however is is clearly not desirable to locally smooth a data set in all cases (effectively removing high frequency and highly spatially localized image components). Therefore increasingly “smart filters” based on techniques such as anisotropic diffusion, which smoothes the image to different extents in the direction of the intensity gradient (across a boundary).and along the boundary, or wavelets, which are very useful because they can remove noise from an image wihle recognizing that certain noise-like components need to be preserved.

## E. Proposed Algorithm for Image Proprecessing

### 1. Adaptive gradient-based and anisotropic diffusion equation filtering algorithm

In order to filter out the noise while preserve the boundary, also to accelerate filtering speed, the filter should first make judgement to each pixel in image and implement filtering manipulation on pixels that are judged as noises, otherwise, no filtering manipulation. Then we can achieve noise filtering, boundary preservation and fast processing speed.

#### a. Gradient based background noise detection

From previous images, it is known that in background area, the grayscale changes less but in boundary area, it changes more. Therefore, we can compute the variation in image $I(x,y)$ and window $w(x,y)$ to judge the properties of pixel background, boundary or noise. To accelerate processing rate, in this thesis I only implement filter on possible noise candidate pixel. Suppose the gradient of each direction as $\nabla I_i(x,y)$, $i$ denotes the all eight directions. The judgement rule is described as follows

$$\nabla I(x,y) = \frac{\sum_{i=0}^{7} \nabla I_i(x,y)}{8} \tag{2-3}$$

$$\begin{cases} Boundary\ area\ \nabla I(x,y) \geq T \\ Background\ area\ \nabla I(x,y) < T \end{cases} \tag{2-4}$$

After the computation of all pixels, I implemented the denoising manipulation on the pixels that are judged as boundary and noise.

## b. Anisotropic filtering algorithm

According to the concept of divergence field, anisotropic diffusion equation can be described as follows:

$$\mathrm{I_t} = div(c(t,x,y)\nabla I) \tag{2-5}$$

In equation (2-5), $div$ is divergence operator, $\nabla$ is gradient operator, $I$ is the function of $(x,y)$, c $(t,x,y)$ is spatial scale function, namely the diffusion coefficient, which is the nonnegative monotonic decreasing function, $\mathrm{I_t}$ is the derivative of $I$ to $t$, $t$ is the time of thermal diffusion.

The selection of diffusion coefficient c $(t,x,y)$ will directly influence the filtering effect and in general, diffusion coefficient c $(t,x,y)$ is given a value of the norm of a vector $E(t,x,y)$, which has the following attributes:

(1) In interior area $E(t,x,y)$=0

(2) In the boundary area, $E(t,x,y)=\varepsilon \cdot e(t,x,y)$,$e(t,x,y)$ is the unit vector of the gradient of point (x, y) in the boundary area. $\varepsilon$ is the strength difference at each side of the boundary.

We can determine $c$ $(\mathrm{t},x,y)$ as

$$c(t,x,y) = g(\|E\|) \tag{2-6}$$

The classical $g(x)$ can be described as

$$g(x) = \frac{1}{1+(x/k)^2} \tag{2-7}$$

or

$$g(x) = \exp(-(x/k)^2) \tag{2-8}$$

Here, $k$ is the controlment coefficient of diffusion strength, $x$ is the gradient of the diffusion point. According to the definition of diffusion coefficient, in different directions different coefficient is adopted. In addition, I take the monotonic decreasing function in different directions as the diffusion coefficient, in background area or interior area of image, the gray level values are similar and gradient is very small, so the diffusion coefficient is large to implement smoothing. On the contrary, in the area of boundary and noise, gray level value changes much and according the gradient increases, diffusion coefficient is very small to preserve the boundary information. Here we consider the processing in eight directions in Fig.7.

The algorithm is achieved using iterations that are described in Table.1 as the discretization of the thermal diffusion process. The detailed algorithms are as follows:

Table 1 Anisotropic filtering algorithm description

1: Set the number of iteration N,

2: Compute the gradients in eight different directions

$\nabla_i \mathrm{I}(\mathrm{x},\mathrm{y}), i = 0,1,2,3,4,5,6,7$

3: Compute the diffusion coefficient in eight different directions

$\mathrm{c}_i(\mathrm{x},\mathrm{y}), i = 0,1,2,3,4,5,6,7$

4: Compute the grayscale value at each point after filtering

$$I'(x,y) = I(x,y) + \Delta_i \left\{ \sum_{i=0}^{7} \left( \frac{\mathrm{c}(\mathrm{x},\mathrm{y}) + \mathrm{c}_i(\mathrm{x},\mathrm{y})}{2} I_i(\mathrm{x},\mathrm{y}) \right) - \frac{8\mathrm{c}(\mathrm{x},\mathrm{y}) + \sum_{i=0}^{7} \mathrm{c}_i(\mathrm{x},\mathrm{y})}{2} I(x,y) \right\} \quad (2\text{–}9)$$

$\Delta_i$ is the iteration step size, $I_i$ is the gray level value

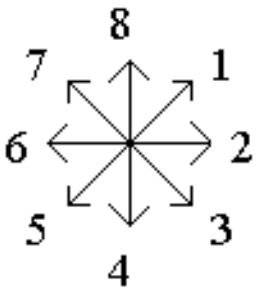


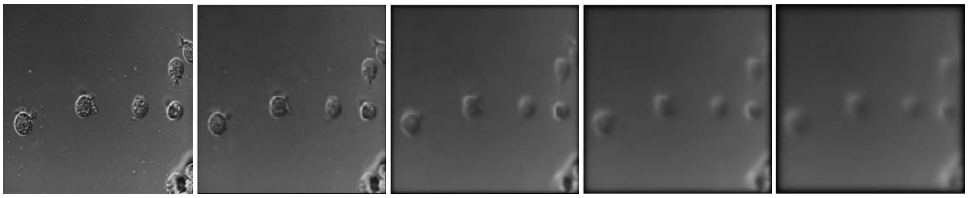

Fig 7 Diffusion direction diagram

## c. Experimental Results

I experiment the above algorithm of gradient-based background noise detection and anisotropic filtering. The diffusion coefficient is set as 10, iteration time as 5,10,15,20 and thresholding value as 30. A sample microscopic image included in Matlab was used, which demonstrates a good example of microscopic images with adding Gaussian background noise, pulse noise and tiny particle noise. I also compare the results of my algorithm with that of pure mdian filtering. It is very clear that, the filtering algorithm proposed above can obtain satisfactory results in removing background noise as well as preserving object boundaries. In addition, by using noise detection algorithm, the computational speed improved by 3 times compared with using anisotropic filtering algorithm alone.

Fig.8 demonstrates the experimental results. The top row shows the filtering results of anisotropic filtering algorithm using different iteration times. The bottom row shows the filtering results of traditional median filtering algorithm using different padding size. It is evident that our algorithm obtained better performance

in removing noises and preserving boundary details.



(a) Original (b) T=5 (c) T=10 (d) T=15 (e) T=20

(f) Original (g) S=4 (h) S=8 (i) S=12 (j) S=16

Fig 8 Experimental results using gradient-based noise detection and anisotropic filtering algorithm

A quantitative measure of noise reduction can be obtained by computing the signal-to-noise ratio (often abbreviated SNR or S/N), which is a measure used in science and engineering to quantify how much a signal has been corrupted by noise. There are many definition of SNR, however in image processing, the SNR is usually calculated as the ratio of the mean pixel value to the standard deviation of the pixel values.

$$\mathrm{SNR} = \frac{\mu}{\sigma} \qquad (2\text{-}10)$$

Sometimes SNR is defined as the square of the definition above.

Fig.9 demonstrates the variation of SNR of original microscopic image filtered by anisotropic filtering algorithm mentioned above and median filtering algorithm

respectively. It is evident that the anisotropic filtering algorithm proposed outperforms the median filtering algorithm in improving the SNR.

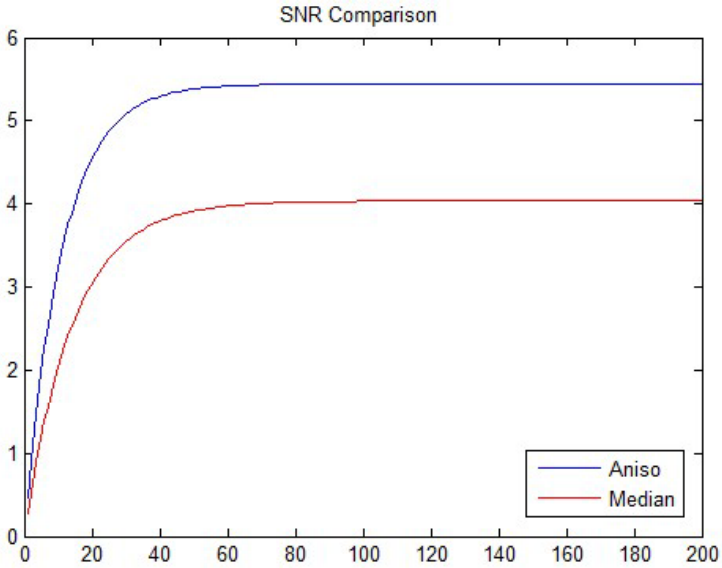



Fig 9 SNR comparison of Aniso and Median filtering

To demonstrate the advantages of our algorithm to other filtering algorithm, I compared the experimental results of other four widely used filtering algorithms in medical image and our proposed one. They are homomorphic filtering algorithm (Homo) [109], DCT-based filtering algorithm [110], wavelet based filtering algorithm (Wavelet) [111], isotropic diffusion based filtering algorithm (Iso) [112] and our anisotropic diffusion based filtering algorithm (Aniso). It is very easy to see that our anisotropic filtering algorithm outperforms others.

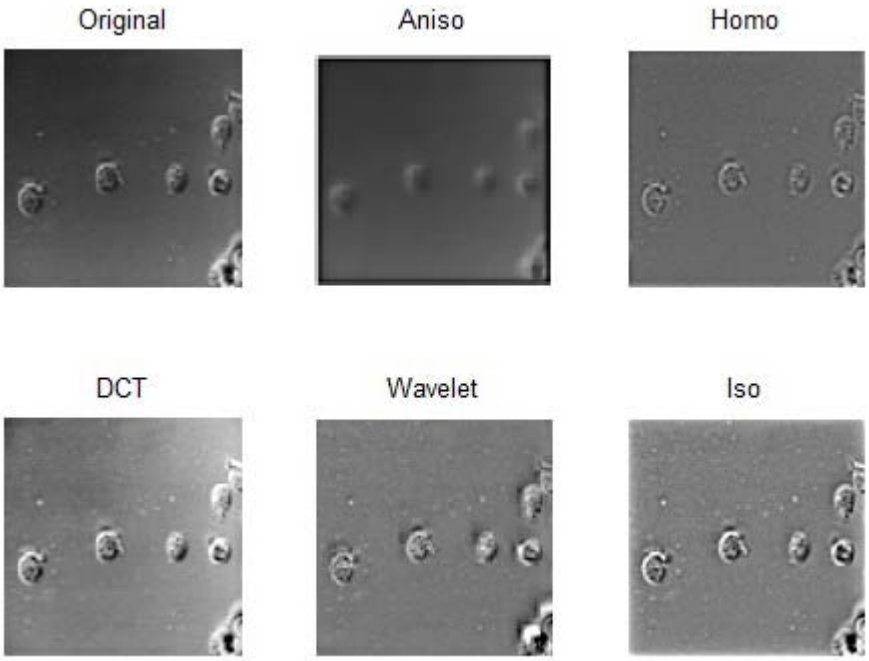
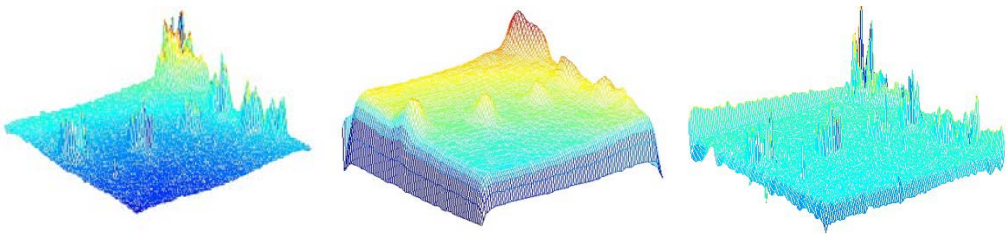



Fig 10 Performance of different filtering algorithm

Fig.11 demonstrates the surface map of the original image and the filtered image using anisotropic diffusion filtering algorithm and other four algorithms. Surface map draws a wireframe mesh with color determined by the intensity value in the image and the color is proportional to the surface height. It is evident that our proposed anisotropic diffusion filtering algorithm can obtain better filtering performance in removing noise and preserving boundary information while the performance of other four algorithms are not very satisfactory in our experiment.



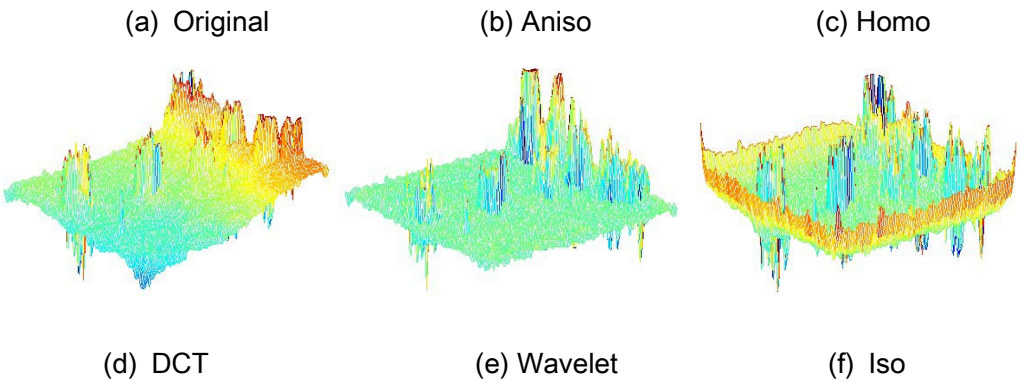



Fig 11 Comparison of surface map of different fitering algorithm

## 2. Information Entropy based noise detection and adaptive median filting algorithm

In information theory, entropy is a measure of the uncertainty associated with a random variable, which quantifies the expected value of the information contained in a message, usually in units such as bits. In image processing, information entropy is normally used as a metrics to measure the effect of a certain segmentation algorithm. Here I introduce this conception to the detection of noise in that the appearance of noise in image displays some extent of uncertainty, which can be described by the information entropy. Median filtering is widely used in filtering out pulse noise and particle noise. Considering the complexity of microscopic images, to achieve a satisfactory result of filtering, an adaptive algorithm should be used. In this thesis, an information entropy based noise detection and adaptive median filtering algorithm was proposed. The experimental results showed that the algorithm could filter out pulse noise and particle noise while preserve the detail information in image.

## a. Noise detection

To a specific pixel in image, whether or not it is polluted by noises should be determined by other pixels, especially the statistical information of its neighborhood pixels. Therefore, here information entropy of neighborhood contrast is used as the standard to determine whether the pixel is polluted by noise.

According to the theory of information, entropy is a probility function, which is defined as:

$$H = -\sum_{i=1}^{k} p_i \log p_i \tag{2-11}$$

In terms of a image, if $H$ denotes information entropy, then $p_i$ denotes the probility of appearance of different pixel grayscale values. Now I will generalize the definition of $p_i$ to probility of pixel local contrast.

As to an image of size $M \times N$, each pixel $I_{x,y}$, the subscript of $x$ and $y$ denotes the spatial position of this pixel, $I_{x,y}$ is the grayscale value. Here I define the neighborhood of the pixel, centers at the point of $(x, y)$, which includes an area of pixel of $(2\mathrm{n}+1) \times (2\mathrm{n}+1) - 1$. $i$, $\mathrm{j} = \pm 1, \pm 2, \ldots \pm n$ and can not be zero at the same time. Then the local contrast of pixel $I_{x,y}$ is defined as:

$$C_{x,y} = \frac{\left| I_{x,y} - \bar{I}_{x,y} \right|}{\bar{I}_{x,y}} \quad x = n, n+1, \ldots, M-n-1, y = n, n+1, \ldots, N-n-1 \tag{2-12}$$

Here

$$\bar{I}_{x,y}, = \frac{1}{(2n+1)(2n+1)-1} \left( \sum_{i=-n}^{n} \sum_{j=-n}^{n} I_{x+i,y+i} - I_{x,y} \right) \tag{2-13}$$

$I_{x,y}$ is the mean value of the pixel of neighborhood.

If we denote

$$\Delta I_{x,y} = \left| I_{x,y} - \bar{I}_{x,y} \right| \tag{2-14}$$

Then the local contrast can be denoted as

$$, \quad C_{x,y} = \frac{\Delta I_{x,y}}{\bar{I}_{x,y}} \quad x = n, n+1, \ldots, M-n-1, y = n, n+1, \ldots, N-n-1 \tag{2-15}$$

Then we can get the local contrast of each pixel:

$$P_{x,y} = \frac{C_{x,y}}{\sum_{i=n}^{M-n-1} \sum_{j=n}^{N-n-1} C_{i,j}} \quad x = n, n+1, \ldots, M-n-1, y = n, n+1, \ldots, N-n-1 \tag{2-16}$$

Finally, entropy of local contrast information is

$$H_c = \sum_{x=n}^{M-n-1} \sum_{y=n}^{N-n-1} P_{x,y} \log P_{x,y} \tag{2-17}$$

According to the theory of information entropy, if $H_c$ is small, the uncertainty of existence of noise is small. We can decide whether it is a noise according to its probility $P_{x,y}$. If $H_c$ is large, the uncertainty of existence of noise is large.

Based on the above analysis, we set a criterion of noise point judgement:

If $$P_{x,y} > P_{threshold} \tag{2-18}$$

The pixel is polluted by noise and it requires filtering, otherwise, it does not require filtering.

Here $P_{threshold}$ is the thresholding value of noise, which is defined as the probility of maximum entropy uncertainty.

$$P_{threshold} = \frac{1}{(M-2n) \times (N-2n)} \tag{2-19}$$

## b. Adaptive median filtering

After determining the noise points, to accelerate the computational speed and preserve boundaries, an adaptive weighting median filtering algorithm will be used to do filtering.

The weighting value has relationship to the value of local contrast. If local

contrast is small, it is less likely for this point to be noise, so the weighting value should take a large value. If the grayscale values of neighborhood are more close to median value, the weighting value should take larger value. Accordingly, if the grayscale values of neighborhood are less close to median value, the weighting value should take smaller value.

Based on the above criterion, we define the weighting coefficient of neighboring pixel as:

$$w_{x,y} = \frac{1/(1+(C_{x,y}I_{x,y} - C_M I_{M_{x,y}})^2)}{\sum_{i=-n}^{n}\sum_{j=-n}^{n}(1/(1+(C_{x+i,y+i}I_{x+i,y+i} - C_M I_{M_{x,y}})^2))} \quad (2\text{–}20)$$

$$x = n, n+1, \ldots, M-n-1, y = n, n+1, \ldots, N-n-1$$

Here $c_{x,y}$ denotes the grayscale value contrast of point $(x, y)$ and its neighborhood points.

$$C_{x,y} = \frac{\left| I_{x,y} - \frac{1}{(2n+1)(2n+1)-1}(\sum_{i=-n}^{n}\sum_{j=-n}^{n} I_{x+i,y+i} - I_{x,y}) \right|}{\frac{1}{(2n+1)(2n+1)-1}(\sum_{i=-n}^{n}\sum_{j=-n}^{n} I_{x+i,y+i} - I_{x,y})} \quad (2\text{–}21)$$

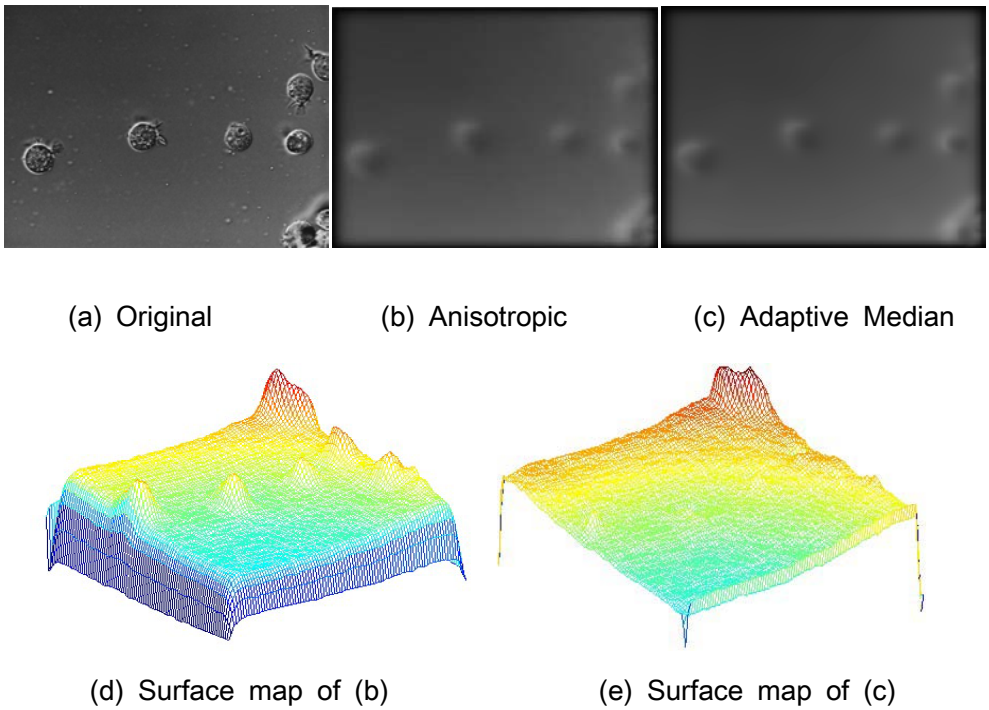
$c_M$ denotes the contrast of median value of current neighborhood to the average value of pixels in neighborhood area. $I_{M\,x,y}$ denotes the median value of current neighborhood area.

After obtaining the weighting coefficient of each pixel in a neighborhood area, multiply the coefficient with the according pixel grayscale value and then accumulate the value in the neighborhood as the output value after filtering.

$$\hat{f}(I_{x,y}) = \sum_{i=-n}^{n}\sum_{j=-n}^{n} w_{x+i,y+i} I_{x+i,y+i} \quad (2\text{–}22)$$

### c. Experimental results

The proposed method, gradient base noise detection and anisotropic filtering algorithm, works not satisfactorily in case of tiny particles and image boundaries in that they have large gradient and may be judged as object boundaries. Fig.9 shows the experimental results of filtering method above compared to gradient based and anisotropic filtering algorithm.



(a) Original (b) Anisotropic (c) Adaptive Median

(d) Surface map of (b) (e) Surface map of (c)

Fig 12 Experimental results of information entropy based noise detection method and adaptive median filtering algorithm

In Fig.9, the experimental results of the two filters proposed above were compared. (a) is the original image. (b) is the result image by anisotropic filter and (c) is the result image by adaptive median filter. In order to demonstrate the results clearly, gray level distributions of two results are showed in (d) and (e). It is evident that adaptive median filter is more capable to remove tiny particles than

anisotropic filter.

## 3. Neighborhood based contrast enhancement algorithm

Due to the factor of focusing distance, illumination environment and object motion, some boundaries of objects appear semitransparent and have less contrast between the backgrounds. If common contrast enhancement methods like histogram equalization and histogram stretching are used, many false objects and object adhesions will definitely appear. In this paper, I designed a nonlinear enhancement function based on the neighborhood contrast information to achieve the contrast enhancement of microscopic images.

### a. Algorithm description

## The definition of neighborhood contrast

Taking the pixels to process $I_{x,y}$ as the central points, we define two regions $\mathrm{R}_1$ and $\mathrm{R}_2$, and $\mathrm{R}_1$ is included inside $\mathrm{R}_2$, that is$\mathrm{R1} \subset \mathrm{R2}$. See Fig.13

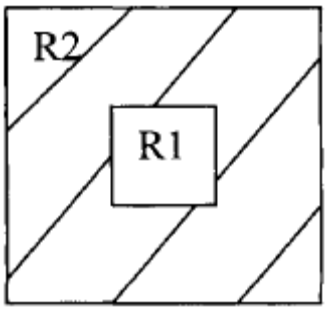



Fig 13 Neighborhood relationship diagram

Then we compute the average grayscale value of two regions respectively and denote them as $\mathrm{A}_1$ and $\mathrm{A}_2$ and the pixels in $\mathrm{A}_1$ are not included when we

compute $A_2$. The neighborhood contrast of the current pixel is defined as

$$C(x,y)=\frac{|A_1-A_2|}{|A_1+A_2|} \tag{2-23}$$

## Enhancement function

According to the definition of neighborhood contrast, the range should be between 0 and 1. In order to achieve neighborhood enhancement, a nonlinear monotonic function is adopted, that is

$$C^{'}(I_{x,y})=F(C(I_{x,y})) \tag{2-24}$$

$F(\cdot)$ is the enhancement function and what value it takes will directly determine the enhancement effects. Considering the characteristics of cell images, several important prerequisites should be met:

(1) Nonliner, which ensures the function to makes different extents of enhacement to different regions.

(2) The relationship between the two constrasts before and after enhancement

$$C^{'}(I_{x,y})\geq C(I_{x,y}) \tag{2-25}$$

According to the characteristics of microscopic images I used in the experiments, I choosed the size of $R_1$ $3\times 3$ and the size of $R_2$ $9\times 9$. We choose the enhancement function as square root function, that is

$$C^{'}(I_{x,y})=\sqrt{C(I_{x,y})} \tag{2-26}$$

### b. Boundary enhancement

After obtaining the enhanced constrast, according to the relationship between

$R_1$ and $R_2$ , the following formula will be used to achieve the boundary enhancement:

$$f^{'}(I_{x,y}) = \begin{cases} A_2 \times \dfrac{1+C^{'}(I_{x,y})}{1-C^{'}(I_{x,y})} & A_1 \geq A_2 \\ A_2 \times \dfrac{1-C^{'}(I_{x,y})}{1+C^{'}(I_{x,y})} & A_1 < A_2 \end{cases} \tag{2-27}$$

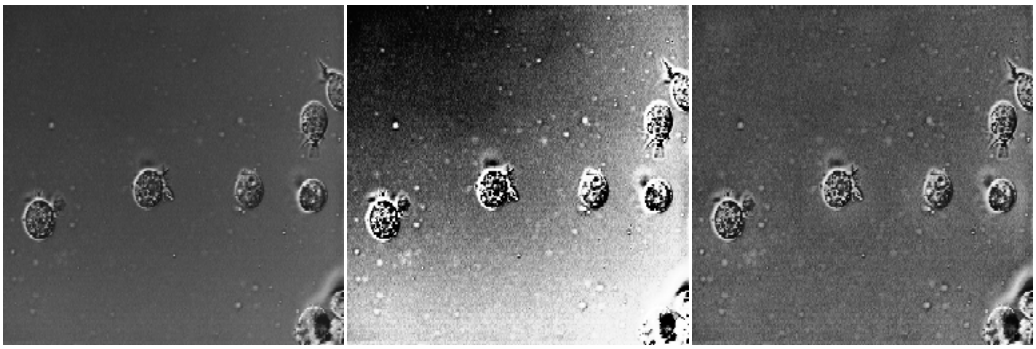
When $A_1 \geq A_2$, $R_1$ is regarded as including no boundary because

$$\frac{1+C'(I_{x,y})}{1-C'(I_{x,y})} > \frac{1+C(I_{x,y})}{1-C(I_{x,y})} \tag{2-28}$$

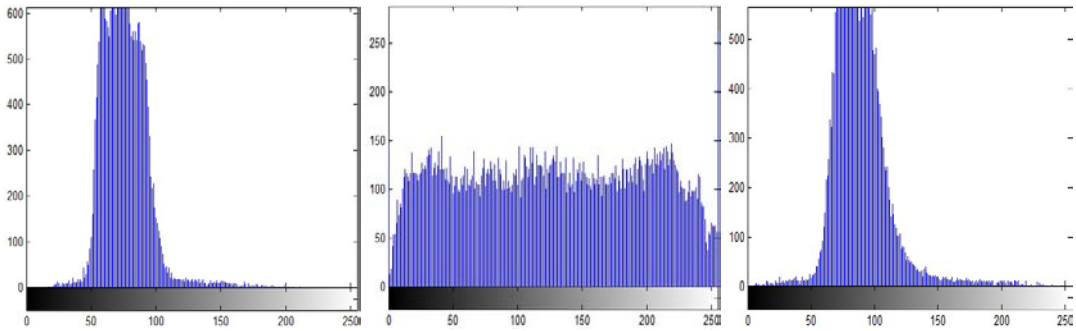
Therefore, the grayscale value increases. Likewise, when $A_1 < A_2$ , $R_1$ is regarded as including boundaries, so the grayscale value decreases. Thereby we achieve the classification of grayscale and increase the contrast of the whole image.

### c. Experimental Results

The experimental results of neighborhood based contrast enhancement algorithm are showed in Fig.11. (a) is the original image. (b) is the result image enhanced by histogram equalization (HE) method. (c) shows the result image using neighborhood contrast enhancement algorithm.



(a) Original (b) Results by HE (c) Results by our method



(d) Original Histogram (e) Histogram by HE (f) Histogram by our method

Fig 14 Experimental results using neighborhood based

contrast enhancement algorithm

## d. Algorithm discussion

It has been mentioned above that neighborhood based contrast enhancement algorithm can obtain good performance in improve the local contrast of an image. However, this algorithm requires great computational amounts.

In order to increase the processing speed, we can fit the original data with linear polynomial or simple quadratic polynomial in the allowable error scope. For instance, for the original image in Fig.8, instead of using the formula (2-26), a

linear polynomial and a quadratic polynomial were used to compute $C^{'}(I_{x,y})$ . Formula (2-29) fits the original data using a linear polynomial. The root mean square error (RMSE) using linear polynomial is 0.1018. Fig.14 demonstrates the performance of fitting.

$$C^{'}(I_{x,y}) = 0.05417 * C(I_{x,y}) + 4.548 \qquad (2\text{-}29)$$

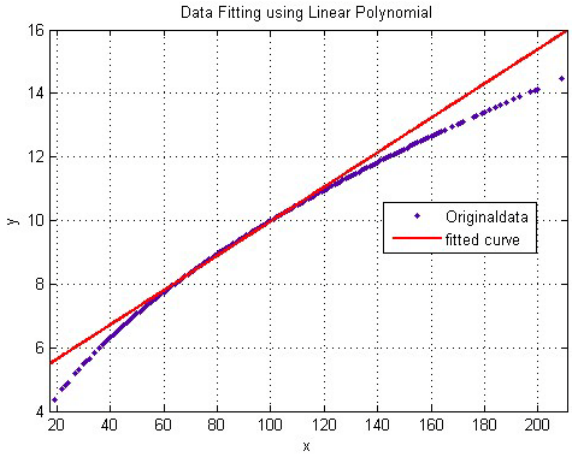



Fig 15 Data Fitting using Linear Polynomial

Formula (2-30) fits the original data using a quadratic polynomial, which obtain a good fitting performance of RMSE 0.02226. Fig.15 demonstrates the performance of this fitting.

$$C^{'}(I_{x,y}) = \text{-}0.0001337 * C^{2}(I_{x,y}) + 0.07838 * C(I_{x,y}) + 3.521 \qquad (2\text{-}30)$$

The advantages of data fitting include two aspects:

(1) It use the fitted data instead of full data to carry out algorithm.

(2) The fitted data make use of polynomial, which makes it possible to process the computation parallelly on different compuatational units.

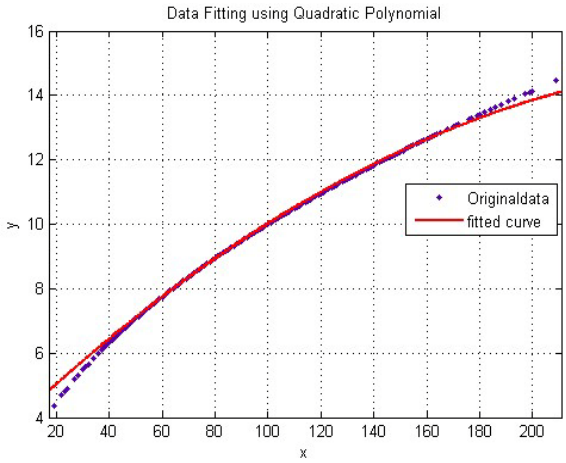



Fig 16 Data Fitting using Quadratic Polynomial

# F. Pixel Classification

Pixel classification consists of determining each pixel of the image, a class among background or target objects that we have interests in.

## 1. Pixel Classification Project Steps

1. Collect images, each containing pixels from only one class of interests
2. Extract samples (small windows surrounding pixels of interest) from images
3. Calculate derived features
4. Train classifier to distinguish between "background" and "not background" classes
5. Apply learned classifier to test images containing pixels from both classes

## 2. Pixel Classification Methods

To realize this classification, several classification methods are used in this work including Bayes, K-means and SVM.

## Bayes Method:

The Bayes classifier is based on the Bayesian decision theory. It is a supervised statistical approach to pattern classification which assumes that the decision problem is expressed in probabilistic terms. In our experiments, since the Bayes is dealing with color images, a mixture of Gaussian (MoG) distribution model is used. The mixture of Gaussians model uses a number of Gaussians to create a more robust description of the target class. For each element $x$, the class that maximizes the probability to contain this element is searched.

$$f(x,i) = -\frac{1}{2}(x-\mu_i)^T \sum_i^{-1}(x-\mu_i) - \frac{1}{2}\log\left|\sum_i\right| - \log p_i + \frac{n}{2}\log 2\pi \tag{2-30}$$

where $n$ is the number of classes, $u_i$ the mean attribute vector, $\sum_i$ is the conditional covariance matrix and $p_i$ the prior probability of class $i$

## K-means Method:

K-means is one of the simplest unsupervised learning algorithms that solve the well-known clustering problem. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) fixed a priori. This algorithm aims at minimizing an objective function, in this case a squared error function. The objective function

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n} \left\| x_i^{(j)} - c_j \right\|^2 \tag{2-31}$$

where $\left\| x_i^{(j)} - c_j \right\|^2$ is a chosen distance measure between a data point $x_i^{(j)}$ and the cluster center $c_j$, is an indicator of the distance of the $n$ data points from their respective cluster centers.

## SVM Method:

The Support Vector Machine (SVM) performs classification by constructing an N-dimensional hyperplane that optimally separates the data into two categories. SVM are learning systems that use hypothesis space of linear functions by projecting the data into a higher dimensional feature space. The use of kernel function $k(.,.)$ implicitly performs a non-linear mapping to a high dimension feature space reducing the training of a SVM to maximizing a convex quadratic from subject to linear constraints.

$$K(x_i, x_j) = \exp\left( \frac{-\left\| x_i - x_j \right\|^2}{2\sigma^2} \right) \tag{2-32}$$

$$\varphi(x) = \sum_{i \in SV} \alpha_i y_i k(x_i, x) + b \tag{2-33}$$

Our final classification results find that K-means algorithm obtained the best results. K-means algorithm was used to cluster our image elements into two parts marked by two different colors: background by red and target objects by green. To demonstrate the result image clearly, we just show the magnified

image of the sub-image in Fig.17

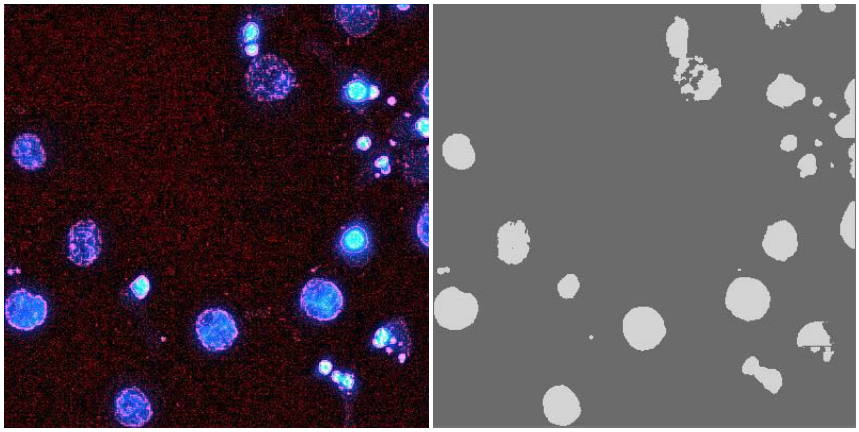


Fig 17 Results of pixel classification

(a) Magnified parts of original Hoechst image (b) Magnified part of Hoechst image after image classification stage

## G. Size Distribution Detection

Granulometries are well-known morphological tools that were first introduced by Matheron [89] that consist of an iterative sequence of morphological operations. The simplest granulometry consists of a series of structural openings which we use in our experiments. A morphological opening, Sop, of an image S by a structuring element (*SE*) B, is the image consisting of the union of every translation x of B that is totally contained within the image set in S, giving a coarse representation of the original image.

$$S_{op} = S_{OB} = \bigcup \{ (B + x) \mid B + x \subseteq S \} \tag{2-34}$$

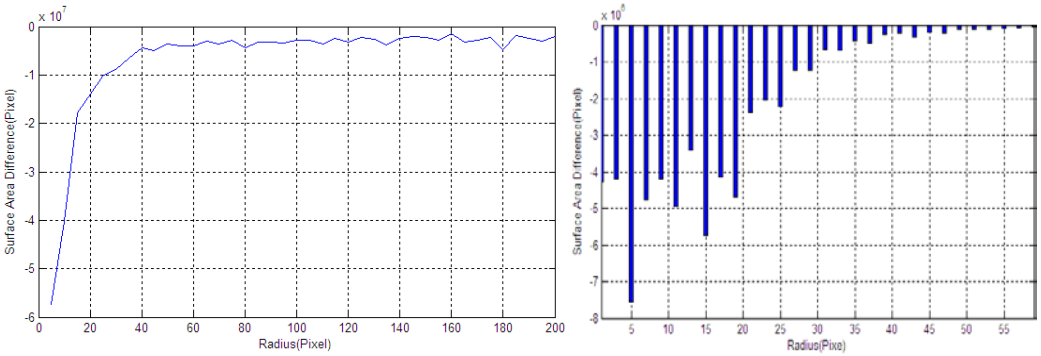
An opening-granulometry uses a series of SEs of increasing size $r$, that remove all image particles within which the SE cannot fit grains of smaller size are

removed from the image by small-scale SEs and larger grains are removed sequentially as the SE is increased in size until no image content remains.

The size distribution detection procedure in our experiment includes two main steps: (1) Size distribution detection of live cancer cell in reference light image; (2) Size distribution detection of apoptotic fragments in Hoechst image. In our experiment, we use surface area to describe the granulometric function as Eq. (2-21)

$$A(k) = \mathrm{U}\left(\gamma_{s_k}(f)\right) \tag{2-35}$$

where $f$ represents the original image and $S_k$, k=1,2,⋯ is a sequence of opening structuring elements of increasing size. $\gamma_{s_k}(f)$ is the surface area. Fig. 8 is the size distribution we obtained in the experiment. We use difference of surface area to measure the distribution of our target objects, live cancer cell in continuous curve and apoptotic fragments in histogram. We first get the size distribution of live cancer cell, which includes nuclei, cytoplasm, chromosome, membrane etc. From the Fig.12 (a), we can see that live cancer cell sizes are almost less than 60 pixels in radius. From the Fig.13 (b), we can see that objects number gets a maximum of radius around 5 pixels and we think they are apoptotic fragments. Because Hoechst image only shows the nuclei and apoptotic fragments so we use the 60 pixels as the coordinate limits in Fig.13 (b).




(a) Live cancer cell size (b) Apoptotic fragments size

Fig 18 Size distribution

# III. Image segmentation algorithm

Image Segmentation is the process of partitioning a digital image into multiple regions or sets of pixels [90, 91]. The result of image segmentation is a set of regions that collectively cover the entire image, or a set of contours extracted from the image. All of the pixels in a region are similar with respect to some characteristics or computed properties, such as color, intensity, or texture. Adjacent regions are significantly different with respect to the same characteristics. Edge detection is one of the most frequently used techniques in digital image processing. The boundaries of object surfaces in a scene often lead to oriented localized changes in intensity of an image, called edges. This observation combined with a commonly held belief that edge detection is the first step in image segmentation, has fueled a long search for a good edge detection algorithm to use in image processing [92]. This search has constituted a principal area of research in low-level vision and has led to a steady stream of edge detection algorithms published in the image processing journals over the last two decades. Even recently, new edge detection algorithms are published each year.

## A. Commonly Used Segmentation Methods

For intensity images, three popular approaches are commonly used: thresholding methods, edge-based methods, region-based techniques.

## 1. Thresholding method

Thresholding methods, which make decisions based on local pixel information, are effective when the intensity levels of the objects fall squarely outside the range of levels in the background. Because spatial information is ignored, however, blurred region boundaries can create havoc. Commonly used thresholding techniques include the following:

(1) Basic Global Thresholding

(2) Basic Adaptive Thresholding

(3) Double Thresholding

## 2. Edge-based methods

Edge detection techniques transform images to edge images benefiting from the changes of grey tones in the images. Edges are the sign of lack of continuity, and ending. Because of this transformation, edge image is obtained without encountering any changes in physical qualities of the main image. Objects consist of numerous parts of different color levels. In an image with different grey levels, despite an obvious change in the grey levels of the object, the shape of the image can be distinguished.

Three most frequently used edge detection methods are compared. They are (1) Roberts Edge Detection; (2) Sobel Edge Detection and (3) Prewitt edge detection. The details of methods as follows:

## The Roberts Detection:

The Roberts Cross operator performs a simple, quick to compute, 2-D spatial gradient measurement on an image. It thus highlights regions of high spatial frequency that often correspond to edges. In its most common usage, the input to the operator is a grayscale image, as is the output. Pixel values at each point in the output represent the estimated absolute magnitude of the spatial gradient of the input image at that point. See Fig.14

| | |
|---|---|
| +1 | 0 |
| 0 | -1 |

Gx

| | |
|---|---|
| 0 | +1 |
| -1 | 0 |

Gy

Fig 19 Roberts Mask

## The Prewitt Detection:

The prewitt edge detector is an appropriate way to estimate the magnitude and orientation of an edge. Although differential gradient edge detection needs a rather time consuming calculation to estimate the orientation from the magnitudes in the x and y-directions, the compass edge detection obtains the orientation directly from the kernel with the maximum response. The prewitt operator is limited to 8 possible orientations, however experience shows that most direct orientation estimates are not much more accurate. This gradient-based edge detector is estimated in the 3x3 neighbourhood for eight directions. All the

eight convolution masks are calculated. One convolution mask is then selected, namely that with the largest module. See Fig.15

| | | |
|---|---|---|
| -1 | +1 | +1 |
| -1 | -2 | +1 |
| -1 | +1 | +1 |

0°

| | | |
|---|---|---|
| +1 | +1 | +1 |
| -1 | -2 | +1 |
| -1 | -1 | +1 |

45°

Fig 20 Prewitt Mask

## The Sobel Detection:

The Sobel operator performs a 2-D spatial gradient measurement on an image and so emphasizes regions of high spatial frequency that correspond to edges. Typically, it is used to find the approximate absolute gradient magnitude at each point in an input grayscale image. In theory at least, the operator consists of a pair of 3x3 convolution kernels as shown in Figure 4. One kernel is simply the other rotated by $90^0$.This is very similar to the Roberts Cross operator. The convolution masks of the Sobel detector are given below. See Figure.16

| | | |
|---|---|---|
| -1 | 0 | +1 |
| -2 | 0 | +2 |
| -1 | 0 | +1 |

Gx

| | | |
|---|---|---|
| +1 | +2 | +1 |
| 0 | 0 | 0 |
| -1 | -2 | -1 |

Gy

Fig 21 Sobel Mask

The weakness of the above-mentioned boundary based methods in connecting together broken contour lines make them, too, prone to failure in the

presence of blurring. Active contour can also be included in this category and I will describe it in detail in the part.

## 3. Region-based method

Region-based is a segmentation method of segmentation by finding coherent regions (pixel similarities) according to common image properties rather than finding boundaries. These image properties consist of

(1) Intensity values from original images, or computed values based on an image operator

(2) Textures or patterns that is unique to each type of region

(3) Spectral profiles that provide multidimensional image data

This approach has specific advantages over boundary-based methods:

(1) It is guaranteed (by definition) to produce coherent regions. Linking edges, gaps produced by missing edge pixels, etc. are not an issue.

(2) It works from the inside out, instead of the outside in. The question, which object a pixel belongs to, is immediate, not the result of point-in-contour tests.

However, it also has drawbacks:

(1) Decisions about region membership are often more difficult than applying edge detectors.

(2) It cannot find objects that span multiple disconnected regions. (Whereas edge-based method can be designed to handle "gaps" produced by occlusion—the Hough transform is one example.)

## B. Medical Image Segmentation

Medical image segmentation refers to the segmentation of known anatomic structures from medical images.

Structures of interest include organs or parts thereof, such as cardiac ventricles or kidneys, abnormalities such as tumors and cysts, as well as other structures such as bones, vessels, brain structures etc. The overall objective of such methods is referred to as computer-aided diagnosis; in other words, they are used for assisting doctors in evaluating medical imagery or in recognizing abnormal findings in a medical image.

In contrast to generic segmentation methods, methods used for medical image segmentation are often application-specific; as such, they can make use of prior knowledge for the particular objects of interest and other expected or possible structures in the image. This has led to the development of a wide range of segmentation methods addressing specific problems in medical applications.

Some methods proposed in the literature are extensions of methods originally proposed for generic image segmentation. In a modification of the watershed transform is proposed for knee cartilage and gray matter/white matter segmentation in magnetic resonance images (MRI). This introduces prior information in the watershed method via the use of a previous probability calculation for the classes present in the image and via the combination of the watershed transform with atlas registration for the automatic generation of markers.

Other methods are more application specific; for example in, segmentation

tools are developed for use in the study of the function of the brain, i.e. for the classification of brain areas as activating, deactivating, or not activating, using functional magnetic resonance imaging (FMRI) data. The method of performs segmentation based on intensity histogram information, augmented with adaptive spatial regularization using Markov random fields. The latter contributes to improved segmentation as compared to non-spatial mixture models, while not requiring the heuristic fine-tuning that is necessary for non-adaptive spatial regularization previously proposed.

## C. General Introduction of Traditional GVF Model and Its Application to Cell Images

### 1. Traditional snake model and GVF model

The snake model represents a kind of energy minimization curve that can move under the influence of different force terms. The internal force term restrains its shape and the external force term restrains its movements while the image force term guides it to prominent features of the image. The traditional snake model can be defined as a parametric curve

$$X(s) = [x(s), y(s)] \qquad s \in [0 \;\; 1] \tag{3-1}$$

that moves through the spatial domain of an image to minimize the energy functional defined as

$$E_{snake} = E_{\text{int}} + E_{image} + E_{con} = \int_0^1 (E_{\text{int}}(v(s)) + E_{image}(v(s)) + E_{con}(v(s)))\, ds \tag{3-2}$$

Where $E_{int}$ represent the internal energy of the curve due to bending, $E_{image}$

gives rise to the image forces, and $E_{con}$ gives rise to the external constraint forces.

### a. Internal Energy

The internal curve energy can be written

$$E_{int} = (\alpha(s)|v_s(s)|^2 + \beta(s)|v_{ss}(s)|^2)/2 \quad (3\text{-}3)$$

The curve energy is composed of a first-order term controlled by $\alpha(s)$ and a second-order term controlled by $\beta(s)$. The first-order term makes the snake act like a membrane and the second-order term makes it act like a thin plate. Adjusting the weights $\alpha(s)$ and $\beta(s)$ controls the relative importance of the membrane and thin-plate terms.

### b. Image Force

In order to make snakes useful for application we need energy functional that attract them to salient features in images. Normally, three different energy functionals attract a snake to lines, edges, and terminations. The total image energy can be expressed as a weighted combination of the three energy functionals

$$E_{image} = w_{line}E_{line} + w_{edge}E_{edge} + w_{term}E_{term} \quad (3\text{-}4)$$

### c. External Force

External force is given to snake model externally and empirically, which leads snake model to an estimated direction. Different external forces lead snake model to different salient features in images.

The traditional snake model has several limitations: firstly, the model fails if the object has a concave or convex contour with a big curvature; secondly, it requires the initial contours closely approximate the real boundaries of objects, namely, the model does not have a wide detection range.

Classical improvements to traditional snake model can be roughly divided into three classes:

(1) Improvements in external force, such as Balloon force, GVF Snake

(2) Improvements in outline, such as B-spline Snake

(3) Improvements in energy optimization method, such as neural networks method, dynamic prograttuning etc.

## d. GVF snake Model

To solve the problem of limited capture range and poor convergence, Xu and Prince [93] proposed GVF as a new external force for snake. The GVF field $\boldsymbol{v}(x,y) = [u(x,y), v(x,y)]$ defined as the equilibrium solution of the following system of partial differential equations:

$$\boldsymbol{v}_t = \mu\nabla^2\boldsymbol{v} - (\boldsymbol{v} - \nabla f)|\nabla f|^2, \qquad \boldsymbol{v}_0 = \nabla f \tag{3-5}$$

Where $\boldsymbol{v}_t$ denotes the partial derivative of $\boldsymbol{v}$ with respect to $t$, and $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$ is the Laplacian operator. $f$ is an edge map derived from the image and defined to have larger values at the features of interest. A typical choice is given as follows:

$$f(x,y) = -E_{ext}(x,y) = |\nabla(G_\sigma(x,y) * I(x,y))^2| \tag{3-6}$$

For step edge

$$f(x,y) = -E_{ext}(x,y) = -G_\sigma(x,y) * I(x,y) \quad (3\text{-}7)$$

When $|\nabla f|$ is small, the solution of (3) is dominated by $\boldsymbol{v}_t = \mu\nabla^2\boldsymbol{v}$ , which implies homogeneous linear diffusion yielding a slowly varying field of $\boldsymbol{v}$. Whereaswhen $|\nabla f|$ is large, the second term in (3–5) is dominant and produces the effect of keeping $\boldsymbol{v}$ nearly equal to$|\nabla f|$. The parameter $\mu$ regulates the tradeoff between the first term and the second term in the equation and should be set according to the amount of noise present in the image (larger $\mu$ for higher noise).

As the GVF field is calculated as a diffusion of the gradient vectors of a gray–level or binary edge map derived from the image, it greatly increases the capture range of the snake and its capability to move into boundary concavities.

### e. GVF snake applied to microscopic images

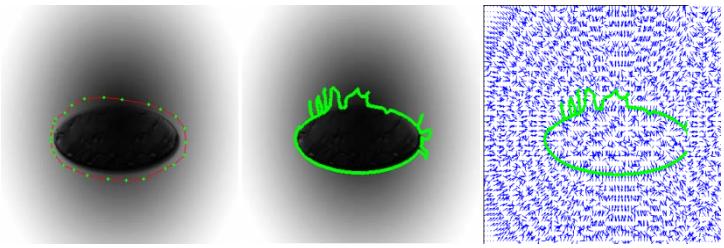
Here I will not go into details of the already known advantages of GVF snake model such as larger capture area, capability in detecting concave boundaries of cells etc. I will focus my experiment on the boundary detection ability of GVF snake on two types of images considering the characteristics of microscopic images, which include image with halo–type boundaries and noise pollution, which are two common image states for microscopic images. Here I assume the shape of objects in microscopic images circular or elliptical shape. The experimental results are shown in Fig.17

According to the experimental results above, it is easy to see that GVF snake model still have some drawbacks applied to boundary detection application for microscopic images:

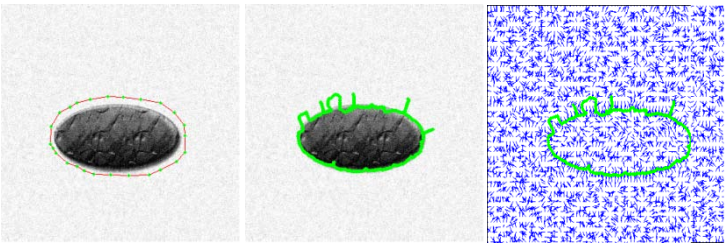
Firstly, GVF snake model is very sensitive to the characteristics of background pixel distribution, such as shadowed background, haloed background and unevenly distributed background distribution, which are very common case in microscopic images.

Secondly, GVF snake model is very sensitive to background noises no matter they are Gaussian distribiton adding noises or speckle noises.

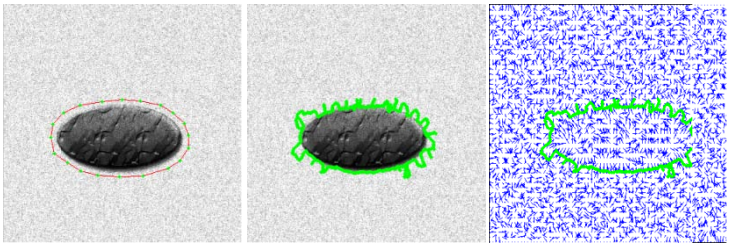
(a) Original elliptical shape cell



(b) Cell image with halo-type shadows



(c) Cell image added with Gaussian distributed noise



(d) Cell image added with speckle noise

Fig 22 GVF snake applied to microscopic images

## D. Proposed GVF Snake Model with Size Constraint

Considering the drawbacks of GVF snake model, I give some improvements proposals:

(1) Maintain the basic model of traditional or GVF snake model and try to provide a high quality initial contour for snake model.

(2) Make some modifications to snake model to demonstate its features of flexibility and deformability.

### 1. Provide a high quality initial contour for snake model

There are two methods to make an initial contour: manual method and automatic method.

Manual method requires the manipulator to produce the initial contour according to his judgements of the characteristic of the object size, shape and snake model capture area.

Automatic method is based on the segmentation results of traditional boundary detection operators like Roberts operator, Sobel operator, Canny operator etc, which can obtain satisfactory performance in simple images. For images with complex background, some posterprocessing like cavity filling and boundary thinning are needed.

## 2. GVF snake model with size constraint

As described above, haloed boundary and noise can greatly influence the boundary detection performance of GVF snake. As to noise, we can use the anisotropic filter mentioned above to remove the noise. For haloed boundary, it is not easy to deal with.

I propose a GVF snake model with area constraint to overcome the influence of haloed boundary. The proposed external energy functional is described as follows:

$$E_{size} = w\big(S(x,y)\big)E_{ext}(x,y) \tag{3-8}$$

Where $E_{ext}(x,y)$, the standard external force of GVF and $w\big(S(x,y)\big)$ is the constraint coefficient of the area $S(x,y)$ at $(x,y)$. The constraint coefficient is employed to require the movement of snake to meet a certain requirement. The constraint is designed to be close to one if the movement of snake meets the requirement and zero if not. The constrait coefficient is given as follows:

$$w\big(S(x,y)\big) = 1 - e^{\gamma} \tag{3-9}$$

$$\gamma = -(\overline{R(x,y)} - \alpha_R)^2/a^2 \tag{3-10}$$

$$(\overline{R(x,y)} = \int_0^1 \sqrt{R_x(s,x(s))^2 + R_y(s,y(s))^2} \tag{3-11}$$

$$R_x\big(s,x(s)\big) = x(s) - \int_0^1 x(r)dr \tag{3-12}$$

$$R_y\big(s,y(s)\big) = y(s) - \int_0^1 y(r)dr \tag{3-13}$$

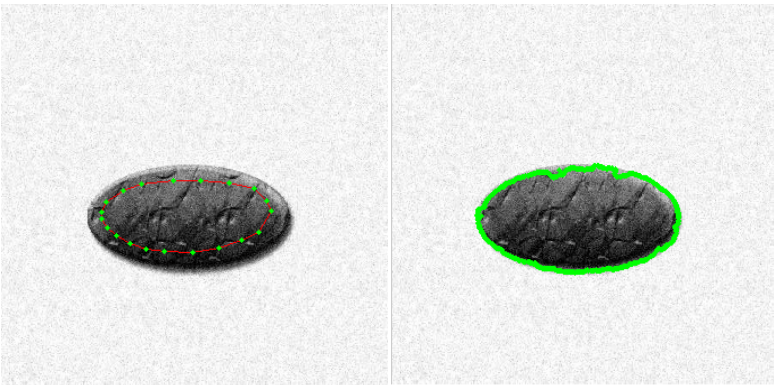
Where $a$ is a adjusting parameter, which take different values according to the extent of haloed boundary. $\overline{R(x,y)}$ is the average radius of snake contour. $\alpha_R$ is the mean radius of cell, which can be computed using the size distribution obtained previously. When $\overline{R(x,y)} \to \alpha_R, \gamma \to zero$, $E_{size}$ will go to zero too, the speed can be adjusted by $a$.

The Euler equations that characterize the solution $(x,y)$ to minimize $w\big(S(x,y)\big)$ are given as follows:

$$\frac{\big(\overline{R(x,y)} - \alpha_R\big)\big(x(s) - \int_0^1 x(r)dr\big)}{\sqrt{R_x(s,x(s))^2 + R_y(s,y(s))^2}} = 0 \tag{3-14}$$

$$\frac{\big(\overline{R(x,y)} - \alpha_R\big)\big(y(s) - \int_0^1 y(r)dr\big)}{\sqrt{R_x(s,x(s))^2 + R_y(s,y(s))^2}} = 0 \tag{3-15}$$

The experimental results are shown in Fig.18



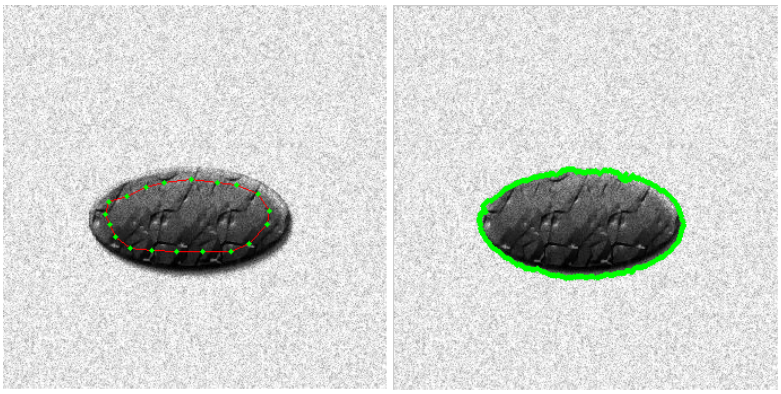


Fig 23 Boundary detection performance of

GVF snake model with size constraint

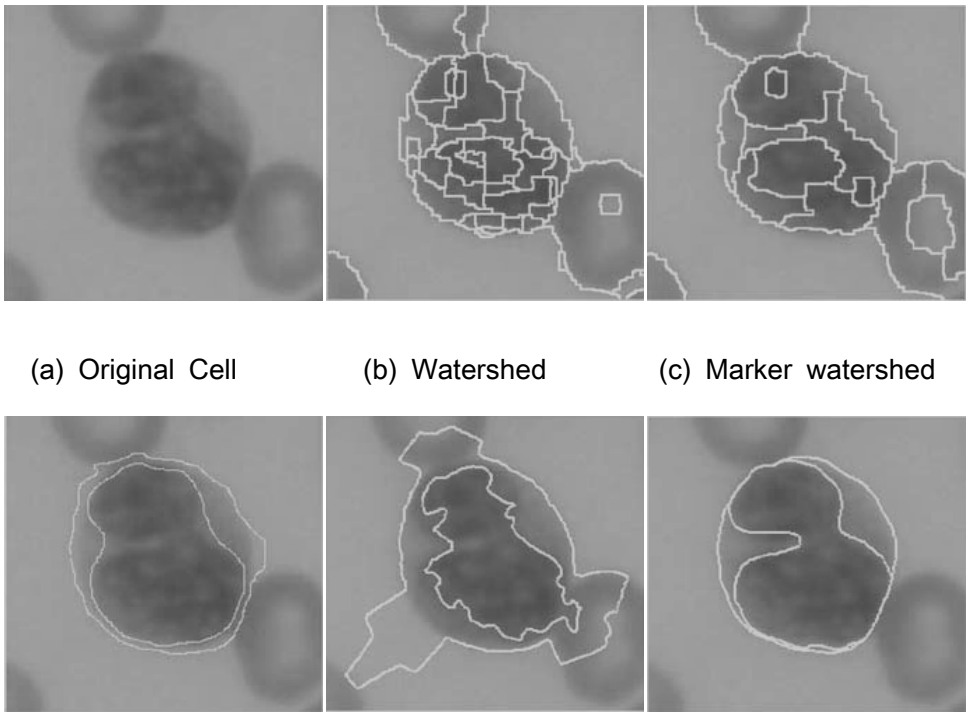
## a. Comparison of segmentation performance

To demonstrate the advantages of the GVF snake model with size constraint in segmentation application, several widely used segmentation algorithm, such as traditional watershed, marker-controled watershed, traditional region merging, FLD (Fisher Linear Discriminant), original GVF snake are described and compared with our method.

Watershed transform use the idea of ridge and catchment basin to solve the image segmentation problems. Marker-controled watershed can overcome the problem of oversegmentation of traditional watershed.Fisher linear discriminate is a common algorithm in pattern recognition, which can also be used in image-segmentation problems by treating segmentation as a classification problem. Region growing has been introduced above.

The segmentation comparison is carried out using cell of lymphoma with clear boundary and concave texture structure, which is a very satisfactory sample

to test our algorithm. The results demonstrated that our algorithm not only can detect the concave boundaries but also can overcome the influence of uneven distribution and noise.



(a) Original Cell (b) Watershed (c) Marker watershed

(d) FLD (e) GVF (f) Our method

Fig 24 Performance comparison of different segmentation algorithm

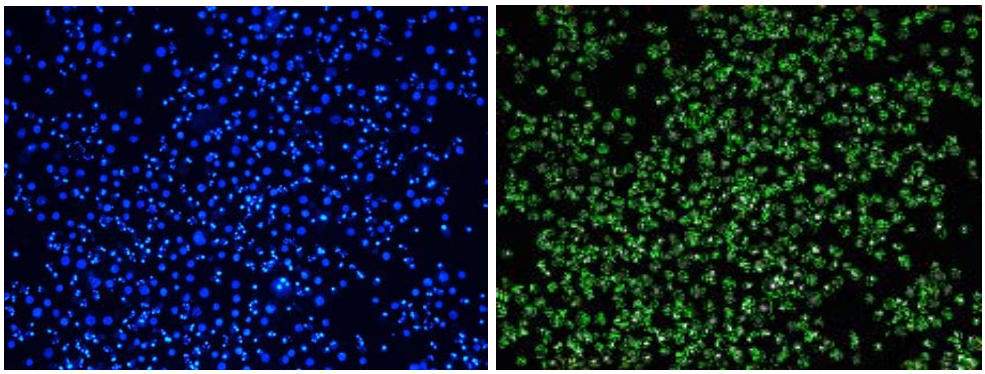
## b. Algorithm discussion

It has been discussed above that our GVF snake model with size constraint can obtain good performance compared with traditional GVF model in terms of images of noise and uneven distribution. However, the performance of this algorithm is not very satisfactory in case of the irregular object shape because the algorithm makes use of the average radius of the objects. For objects of irregular

shape, the use of average radius will bring much error. So, before we use this algorithm, we should first make sure that the objects in the image have the regular shape, such as round shape and elliptical shape.

## E. Separation of Overlappig Cells

I achieve this procedure in four steps: First, I transform the preprocessed Hoechst image into binary image via the thresholding method. After the binarization process, many cell centers were labeled as background, and I use the classical morphological algorithms to fill in these center holes. Second, I determine the detected objects as individual or overlapping objects according to their sizes compared with the size of live cancer cell nuclei and apoptotic fragments. Third, the dividing lines were drawn between overlapping objects. In this work, the traditional watershed algorithm using distance transform was used to achieve this goal. Lastly, the target objects are marked in our result image.Figure.19 demonstrate the results of segmentation step



(a) The original Hoechst image (b) live cell and apoptotic fragments

Fig 25 Performance of cell segmentation

# Ⅳ. Feature extraction and analysis

Feature extraction in pattern recognition is based on finding mathematical methods for reducing dimensionality of pattern representation. A lower-dimensional representation based on patter descriptors is a so-called feature. It plays a crucial role in determining the separating properties of pattern classes. The choice of features, attributes, or measurements has an important influence on: (1) accuracy of classification, (2) time needed for classification, (3) number of examples needed for learning and (4) cost of performing classification.

A good feature should remain unchanged if variations take place within a class, and it should reveal important differences when discriminating between patterns of different classes. In other words, patterns are described with as little loss as possible of pertinent information.

There are four known categories in the literature for extracting features [94]:

(1) Nontransformed structural characteristics: moments, power, amplitude information, energy, etc

(2) Transformed structural characteristics: frequency and amplitude spetra, subspace transformation methods, etc

(3) Structural descriptions: formal languages and their grammers, parsing techniques, and string matching techniques

(4) Graph descriptors: attributed graphs, relational graphs, and semantic networks

In our experiments, for each case, 10 features were estimated automatically

from illumination intensity, morphological and textural nuclear features [95]. Information about nuclear size and shape was captured by morphological features, which constituted measurements of nuclear area, roundness and concavity [95]. The feature of concavity, attempts to measure the severity of concavities, or the indentations of a nucleus [95].

The remaining two features were textural features that encoded chromatin distribution of the cell nucleus. These features were estimated by means of nuclear histograms and the co-occurrence matrix [96]. Nuclear chromatin-texture quantification has been examined in several studies, and has proved to carry significant diagnostic information in the analysis of pathologic material [97-99]. To quantify texture properties of nuclei, textural features were formed from first order statistics and from spatial gray tone co-occurrence probability matrices [96, 100]. The gray level co occurrence matrix was used for second order texture information extraction from cell nuclei. A co-occurrence matrix P is an estimate of the second order joint conditional probability density function (PDF) $P(i,j|d,\phi)$ , $\phi^0 = \{0^0, 45^0, 90^0, 135^0\}$. Each $P(i,j|d,\phi)$ is the probability of transition from gray level i to gray level j, given an inter-sample spacing of d, and the direction is given by the angle $\phi$. Numerous features can be extracted from co-occurrence matrices and a large number of such features has been proposed [96, 100]. Most authors, however, agree that in practice only a few of these are independent. In this work, we adopted the following features, which appear to be the most effective in texture discrimination, as it is also reported by other researchers [98, 101, 102].

Table 2 Typical feature list

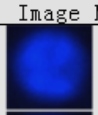
| Num | Image | MeanInt | MedInt | StdInt | MajAxL | MinAxL | Peri | Area | Eccen | Contr | Entr |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 |  | 0.298 | 0.315 | 0.086 | 309.54 | 296.14 | 1028.12 | 71906 | 0.29 | 0.17 | 2.39 |
| 2 |  | 0.319 | 0.333 | 0.07 | 254.59 | 242.17 | 866.33 | 48240 | 0.31 | 0.24 | 2.42 |
| 3 |  | 0.376 | 0.392 | 0.044 | 295.55 | 208.87 | 895.51 | 47877 | 0.71 | 0.23 | 1.75 |
| 4 |  | 0.318 | 0.329 | 0.027 | 268.75 | 251.15 | 891.15 | 53179 | 0.36 | 0.24 | 2.47 |
| 5 |  | 0.285 | 0.293 | 0.052 | 294.53 | 286.89 | 994.41 | 66592 | 0.23 | 0.21 | 2.29 |
| 6 |  | 0.342 | 0.371 | 0.066 | 268.23 | 253.94 | 897.08 | 53418 | 0.32 | 0.26 | 2.23 |
| 7 |  | 0.468 | 0.447 | 0.097 | 146.19 | 133.73 | 496.21 | 15153 | 0.41 | 0.28 | 2.54 |
| 8 |  | 0.466 | 0.454 | 0.121 | 183.65 | 159.58 | 584.65 | 23069 | 0.49 | 0.33 | 2.53 |
| 9 |  | 0.489 | 0.481 | 0.135 | 145.96 | 133.81 | 476.75 | 15331 | 0.41 | 0.34 | 2.65 |
| 10 |  | 0.472 | 0.452 | 0.103 | 129.31 | 111.13 | 410.02 | 11308 | 0.51 | 0.39 | 2.73 |
| 11 |  | 0.523 | 0.511 | 0.129 | 146.52 | 140.36 | 487.16 | 16172 | 0.29 | 0.32 | 2.51 |
| 12 |  | 0.503 | 0.497 | 0.112 | 132.01 | 106.45 | 405.78 | 11057 | 0.59 | 0.41 | 2.63 |

Table.2 lists some typical images of live normal cells, apoptotic fragments and their respective feature values. I used ten features in our experiments and the descripition for each feature is listed in Table.3.

Table 3 Feature description

| No | Name | Description |
|---|---|---|
| 1 | MeanInt | Mean of pixel intensity values |
| 2 | MedInt | Median of pixel intensity values |
| 3 | StdInten | Standard deviation of pixel intensity values |
| 4 | MajAxL | The length (in pixel) of the major axis of the ellipse shape cell |
| 5 | MinAxL | The length (in pixel) of the major axis of the ellipse shape cell |
| 6 | Peri | The total number of pixels around the boundary of each cell |
| 7 | Area | The actual number of pixels inside the cell region |
| 8 | Eccen | The eccentricity of the ellipse shape cell (Value between 0 and 1. 0 is actually a circle and 1 is line segment) |
| 9 | Contr | Co-occurrence texture feature contrast |
| 10 | Entr | Co-occurrence texture feature entropy |

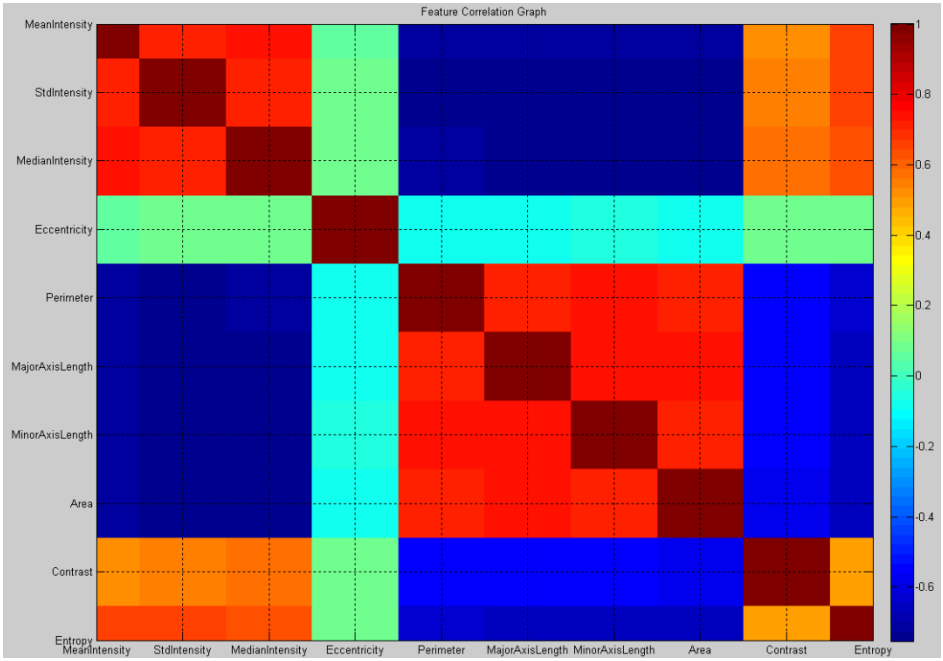



Fig 26 Features Correlation Graph

# V. Pattern Recognition Algorithm Research

The key point of pattern recognition is the design of a classification rule or function according to the training samples, which will cause the minimum loss or cost of error when we classify our objects.

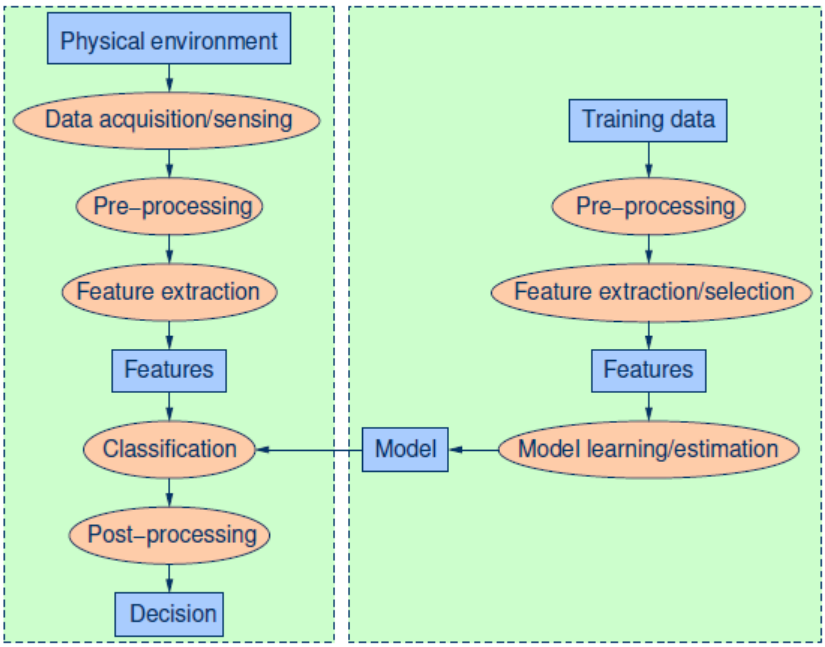



Fig 27 Pattern recognition and classification flow chart

The most important steps of the pattern recognition procedure are feature extraction/selection and model learning/estimation. Feature extraction/selection, or sometimes is called dimensionality reduction, plays an important role in classification performance. A recognition system is designed using a finite set of inputs. While the performance of this system increases if we add additional features, at some point a further inclusion leads to performance degradation.

Dimensionality reduction is accomplished based on either feature selection or feature extraction. Feature selection is based on omitting those features from the available measurements wich do not contribute to class separability.Feature extraction, on the other hand, considers the whole information content and maps the useful information content into a lower dimensional feature space.

In this chapter, I will check the importance of all features I selected for the classification of apoptosis and measure the classification accuracy of them. The classification algorithm used here is decision tree algorithm.

## A. Decision Tree Algorithm

Decision tree algorithm has been used broadly for several years. It is an approximate discrete function method and can yield many useful expressions. It is one of the most important methods for classification. This algorithm's terms followe the "tree" metaphor.It has a root, which is the first split point of the data attribute for building a decision tree. It also has leaves, so that every path from root to leaf will form a rule that is easily understood.

Since the dicision tree is built by given data, the data value and character will be more important. For example, the amount of data will affect the result of the tree building procedure. The type of attribute value will also affect the tree model. Decision trees need two kinds of data: training and testing. Training data, which are usually the bigger part of data, are used for constructing trees. The more training data collected, the higher the accuracy of the results. The other group of data, testing, is used to get the accuracy rate and misclassification rate of the decision tree. Many decision tree algorithms have been developed. One of the

most famous is ID3 [104,105], whose choice of split attribute is basd on information entropy. C4.5 is an extension of ID3 [106],It improves computing efficiency, deals with continuous values, handles attributes with missing values, avoids overfitting, and performs other functions.

## B. Bagging Algorithm

In data mining, an approach to make decisions more reliable is to combine the outputs of different models. Several machine learning techniques do this by learning an ensemble of models and using them in combination: prominent among these is a scheme called “bagging” [107].

Bagging predictor is a method for generating multiple versions of a predictor and using these to get an aggregated predictor. The aggregation averages over the versions when predicting a numerical outcome and performs a plurality vote when predicting a class. The multiple versions are created by making bootstrap replicate of the learning set and using these as new learning sets. Tests on real and simulated data sets using classification and regression trees and subset selection in linear regression have shown that bagging can provide substantial gains in accuracy [108]

Bagging attempts to neutralize the instability of learning methods by simulating the process using a given training set. Instead of sampling a fresh, independent training dataset each time, the original training data is altered by deleting some instances and replicating others. For the replacement, instances are randomly sampled from the original dataset to create a new one of the same size. This sampling procedure inevitably replicates some of the instances and

deletes the others. Here is the algorithm for bagging

Table 3 Bagging decision tree algorithm

```
Model generation
training data
the number of models
For i=1 to k do // create k models
    Create bootstrap sample, D_i, by sampling with
    Replacement from training data
    Apply learning algorithm to the sample.
    Store the result of model, M_i
Endfor
Classification
For i=1 to k do
    Predict/Classify a testing data X using model M_i
Endfor
Return class that is predicted most often
```

## C. Experiment and Discussion

### 1. Dataset description

The dataset includes 10 categories of features extracted from 25000 cells in 25 microscopic images. The total number of data is 10× 25000.

### 2. Experiments

In order to compare the experimental results of Bagging decision tree algorithm, another two classification methods were also applied here, decision

tree algorithm and Bagging Naïve Bayes, All experiments were implemented under Matlab 2010B platform and using Matlab statistical toolbox. The first step to construct the classification ensemble will be to find a good leaf size for the individual trees. Here I tried sizes of 4, 6 and 10. I used the number of trees, 30. For reproducibility and fair comparisons, I reinitialize the random number generator, which is used to sample with replacement from the data, each time we build a classifier. The error are comparable for the three leaf-size options. I will therefore work with a leaf size of 10 because it results in leaner trees and more efficient computations. Here I did not split the data into training and test subsets. This is done internally, it is implicit in the sampling procedure that underlines the method. At each bootstrap iteration, the bootstrap replica is the training set, and any customers left out ("out-of-bag") are used as test points to estimate the out-of-bag classification error reported above.

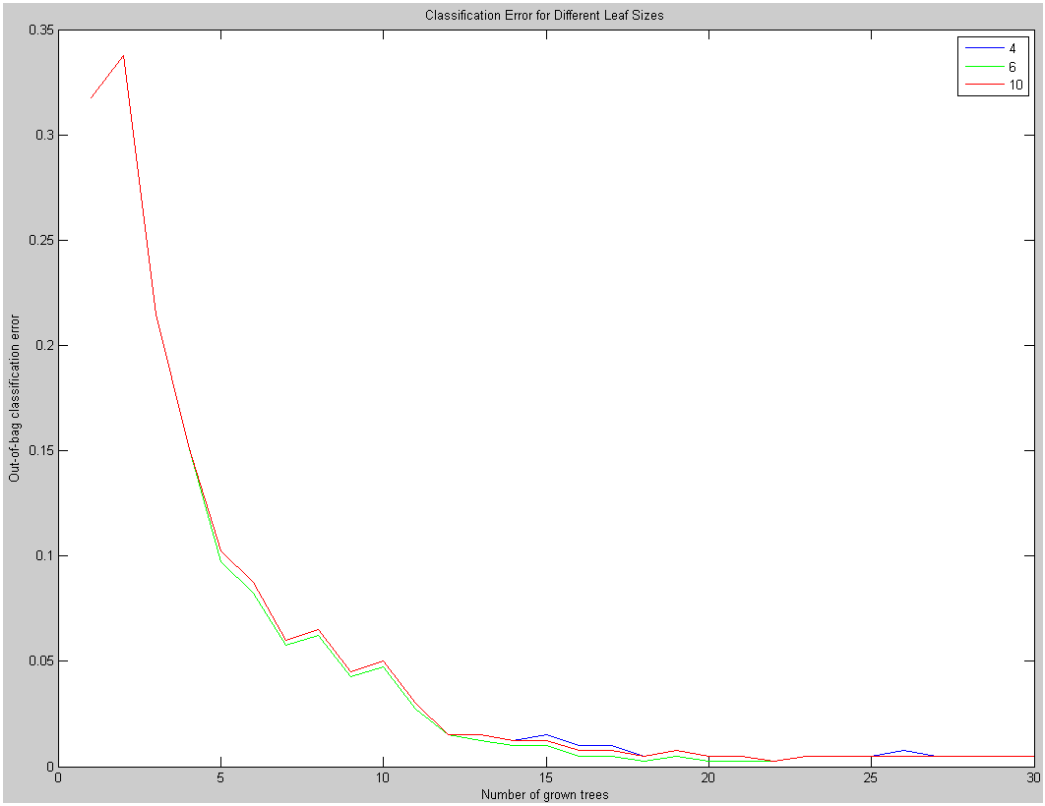



Fig 28 Classification error for different leaf size

Next, I want to find out whether all the features are important for the accuracy of our classifier. I do this by turning on the feature importance measure, and plot the results to visually find the most important features. It is easy to see that feature 3 (MedianIntensity), 7 (MinorAxisLength) and 8 (Area) stand out of the rest while feature 4 (Eccentricity), 9 (Contrast) and 10 (Entropy) has the least ability to seperate apoptosis from normal live cell. Generally speaking, illumination features and size features has most separation ability while texture features has less separation ability, which exactly demonstrate the intuitive feeling that apoptosis

and live cell give to us.

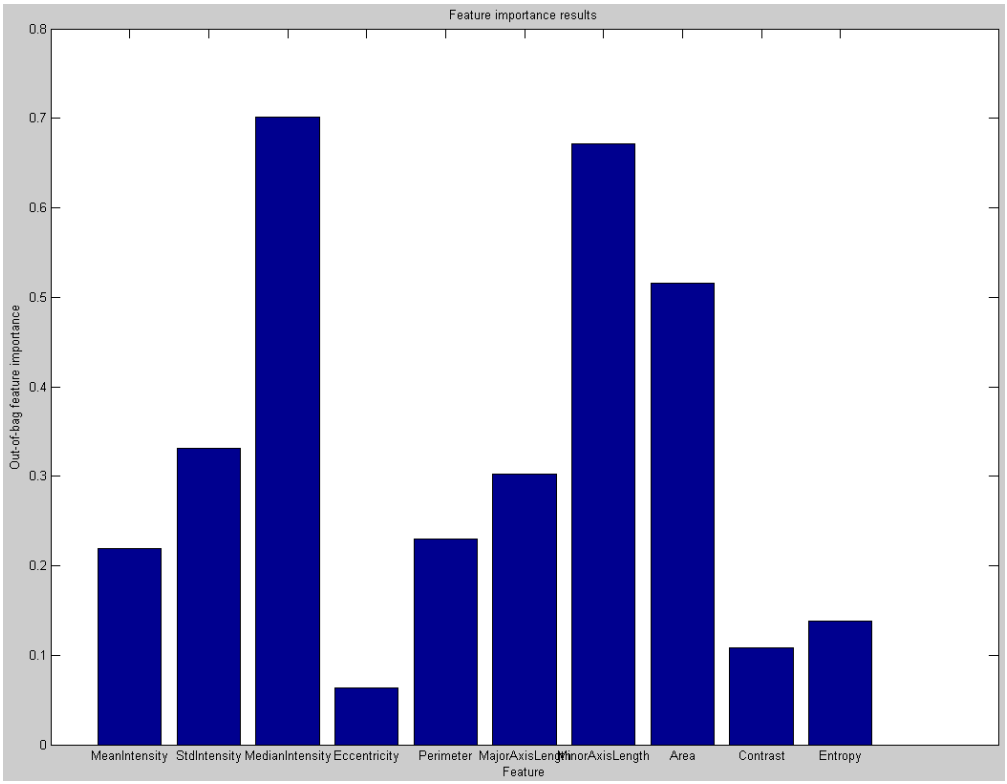



Fig 29 Feature importance graph

The following figure shows comparison of the classification of using all features and only 3 most contributing features.It is evident that when the number of trees is not very big, the classification results using 3 features are comparable to thre results using all ten features.

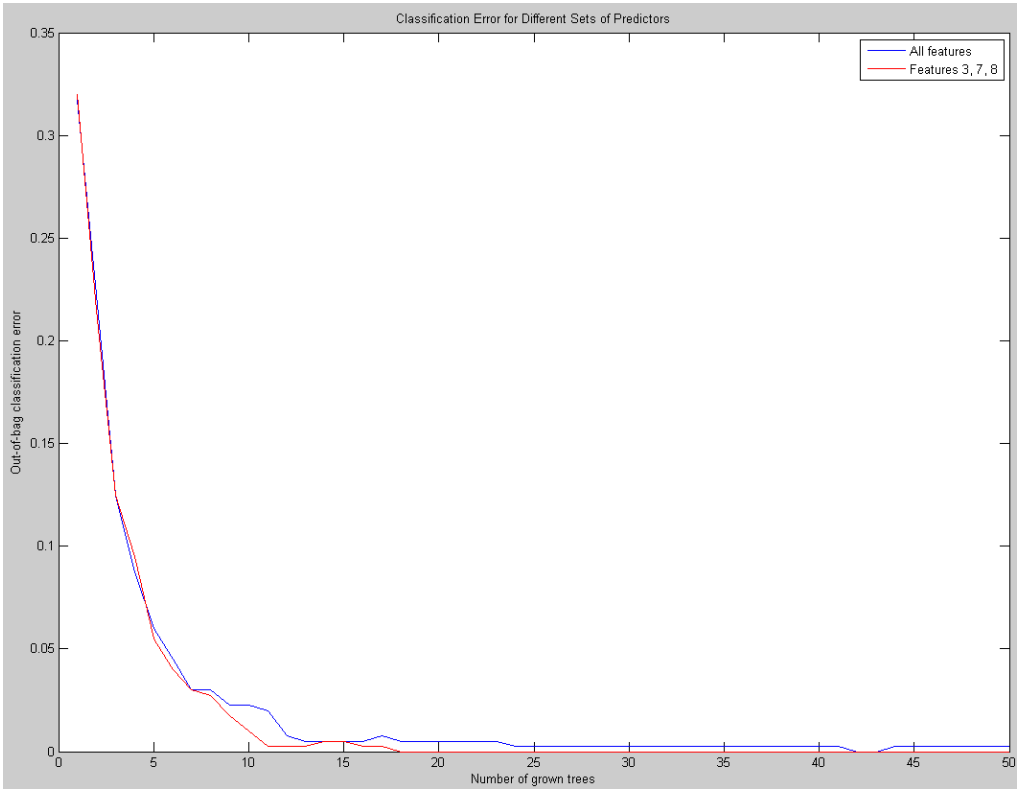



Fig 30 Classification error using most important features

## 3. Performance measurement

Many different metrics are used in machine learning and data mining to build and evaluate models. I employed four performance measures: precision, recall, F-measure and ROC curve.

### a. Confusion Matrix

A distinguished confusion matrix is obtained to calculate the four measures. Confusion matrix is a matrix representation of the classification results. The upper left cell denotes the number of samples classifies as true while there are true,

called true positive (TP). The lower right cell denotes the number of smaples classified as false while they are actually false, called true negative (TN). The other two cells, lower left cell and upper right cell, denote the number of samples misclassified. Specifically, the lower left cell denoting the number of smaples misclassified as false while they actually are true,called false negative (FN), and the upper right cell denoting the number of samples misclassified as true while they actually are false, called false positive (FP). In this thesis, the confusion matrix can be described in following table

Table 4 A two-by-two confustion matrix

| | Classified as live cell | Classified as apoptosis |
|---|---|---|
| Actual live cell | TP | FN |
| Actual apoptosis | FP | TN |

Once the confusion matrixes were constructed, the precision, recall, F-measure are easily calculated as:

$$Recall = \frac{TP}{TP+FN} \tag{5-1}$$

$$Precision = \frac{TP}{TP+FP} \tag{5-2}$$

$$F\_measure = \frac{2\times TP}{2\times TP+FP+FN} \tag{5-3}$$

Less formally, *precision* measures the percentage of the actual live cell among the cells that were classified as live cell. *Recall* measures the percentage of the actual live cells that were discovered. *F_measure balances* between precision and recall. According to the feature importance results, it is known that

MedianIntensity, MinorAxisLength and Area are three features with best classification ability. I experimented on classification results using these three features and the remaining features repectively. The training data is 320 and test data is 80. Table 5 shows the performance results of three methods.

Table 5 The classification performance results using 7 secondary features

| Seven features | Decision tree | Bagging decision tree | Bagging Naïve Bayes |
|---|---|---|---|
| Precision | 89.60% | 90.89% | 91.51% |
| Recall | 86.59% | 87.24% | 88.34% |
| F_measure | 84.96 | 85.12% | 86.27% |

Table 6 The classification performance results using 3 primary features

| Three features | Decision tree | Bagging decision tree | Bagging Naïve Bayes |
|---|---|---|---|
| Precision | 97.34% | 97.25% | 98.31% |
| Recall | 95.08% | 96.11% | 97.21% |
| F_measure | 96.43 | 97.18% | 98.35% |

## b. ROC curve

A ROC graph is another way besides confusion matrix to examine the performance of classifiers. A ROC graph is a plot with the false positive rate on the $X$ axis and true positive rate on the $Y$ axis.The point $(0,1)$ is the perfect classifier and it classifies all positive cases and negative cases correctly. It is point $(0,1)$ because the false positive rate is $0$ (none), and the true positive rate

is 1 (all).

$$TPR = \frac{TP}{TP+FN} \tag{5-4}$$

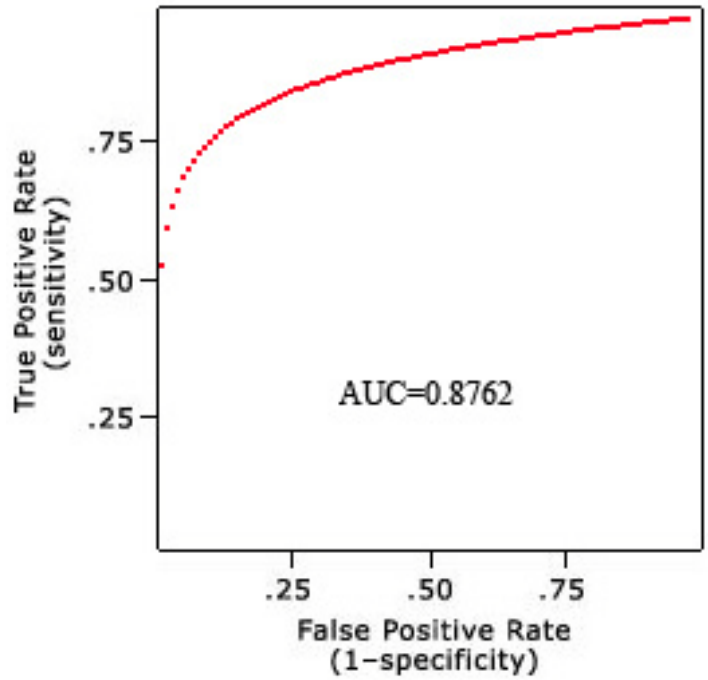
$$FPR = \frac{FP}{FP+TN} \tag{5-5}$$

## Features of ROC curve

- An ROC curve or point is independent of class distribution or error costs
- An ROC graph encapsulates all information contained in the confusion matrix, since *FN* is the complement of *TP*, *TN* is the complement of *FP*
- ROC curve provide a visual tool for examining the tradeoff between the ability of a classifier to correctly identify positive cases and the number of negative cases that are incorrectly classified.

## Area-based accuracy measure

The area beneath an ROC curve (AUC) can be used as a measure of accuracy in many applications. The area measures discrimination ability. The AUC is useful in that it aggregates performance across the entire range of trade-offs.The higher the AUC, the better. Fig.24 shows the ROC curve of bagging decision tree algorithm.The upper figure is the ROC curve of bagging decision tree algorithm using 7 secondary features and the beneath one is the ROC curve using 3 primary features. It is evident that the latter classifier obtained the better classification performance.



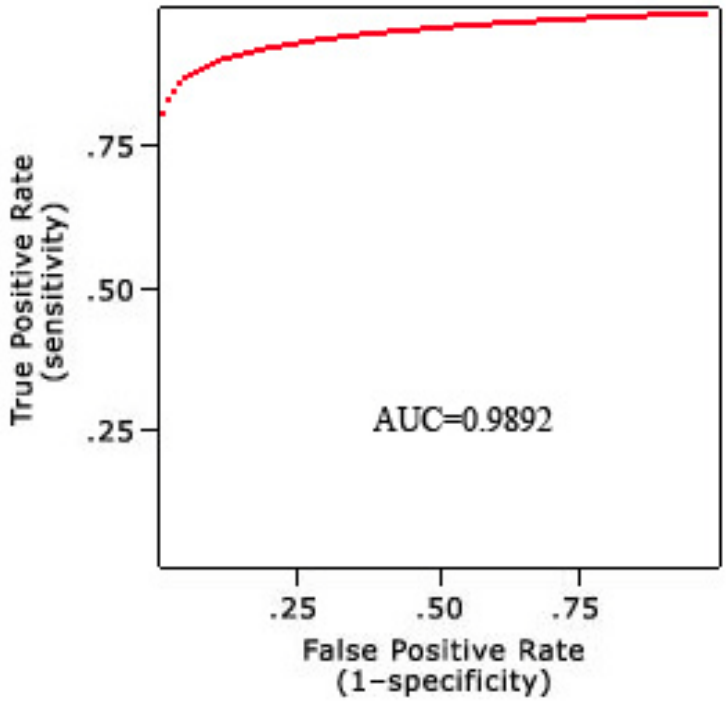




Fig 31 ROC curve of bagging decision tree algorithm

# VI. Experimental Results

## A. Detection Accuracy

In order to evaluate quantitatively the results of experimental results of segmentation and edge detection for live cells and apoptotic fragments, I compared the results obtained by using the algorithm proposed in chapter 3 with visual inspection by two experts. Five different images with different numbers of apoptotic fragments are used, on average, around 1500 interested objects are included in each image.

Here, area overlap measurement [103] is used to measure the results of segmentation and edge detection results by comparing the results I obtained using the algorithm proposed in Chapter 3 with the results obtained by manual segmentation. The area overlap of two different results are defined as

$$O(R_m, R_s) = \frac{S(R_m \cap R_s)}{(S(R_m) + S(R_s))/2} \tag{6-1}$$

Where Rm is the manually labeled region and Rs is the region extracted by the snake model. The ∩ operator takes the intersection of two regions, where $S(\cdot)$ is the area of the region. In our experiment, 10 randomly selected apoptotic bodies were selected and the edges were detected by two experts and compared with our results using the area overlap measurement. The result is shown in Fig.26. The terms, Expert1 and Expert2, measure the degree of accordance of our results against the results obtained manully. “Between Experts” represents the

area overlap measurement between the two experts. It is clear that our proposed algorithm can obtain good performance.

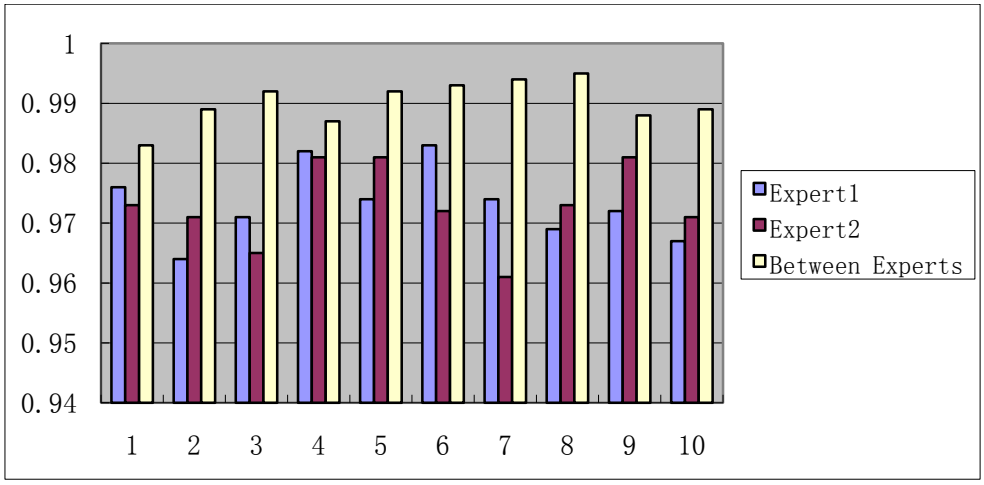



Fig 32 Area overlap measurement.

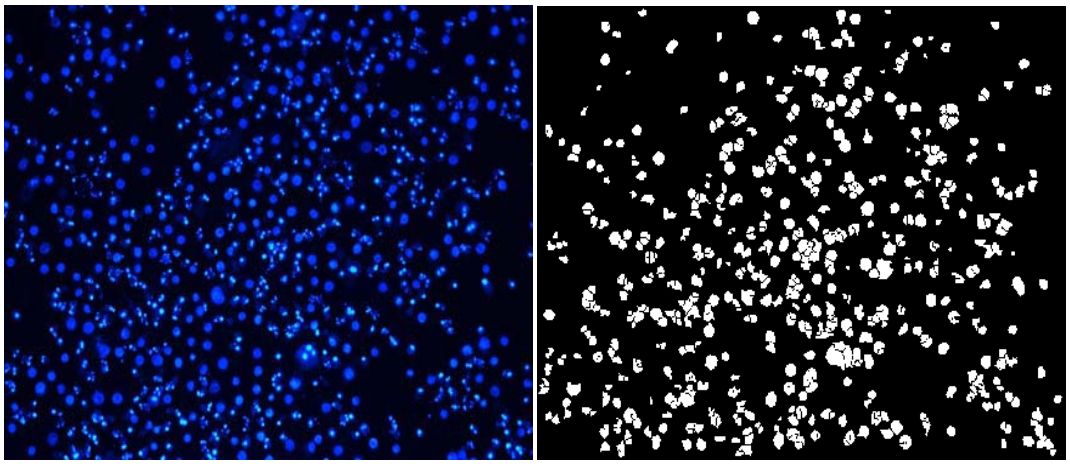
## B. Classification Accuracy

In Chapter 5, I have gone into details of the classification method I used, bagging dicision tree and the performance evaluation of this classifier applied to our experiments. In order to compare the experimental results, I used five different images with different numbers of object. The comparison of classification results are shown in Table.7. In this table, I listed the total number of cells in each image, the number of apoptotic recognized automatically by classifier used in our experiments, the number of apoptotic fragments recognized by four experts manually, the average number of apoptotic fragments by experts, the standard deviation between experts, and the standard deviation calculated between our results and experts' results. The results show that our method can

not only obtain good recognition performance, but can help to find some apoptotic fragments that are not easy to identify merely by naked eyes.

Table 7 Comparison between automatic classification and manual classification

| Image Index | Total cell number | Results by our methods | Results by Expert1 | Results by Expert2 | Results by Expert3 | Results by Expert4 | Average of Experts | S.D. between experts | S.D. between our results and experts |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1468 | 341 | 328 | 330 | 321 | 327 | 326 | 3.39 | 7.52 |
| 2 | 1521 | 352 | 344 | 348 | 346 | 349 | 346 | 2.06 | 4.24 |
| 3 | 1598 | 389 | 377 | 379 | 379 | 381 | 379 | 1.41 | 7.07 |
| 4 | 1524 | 362 | 366 | 368 | 364 | 366 | 366 | 1.41 | 2.83 |
| 5 | 1577 | 394 | 388 | 390 | 389 | 386 | 388 | 1.50 | 4.24 |



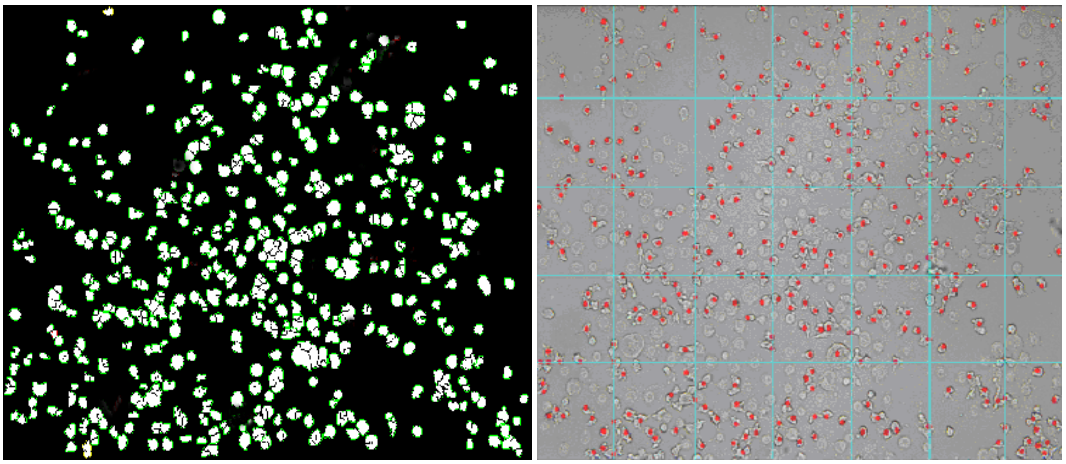
(a) Orginal Hoechst Image (b) The apoptotic fragments

Fig 33 The performce of apoptotic fragments detection

## C. Apoptosis Counting

Here, I use the size distribution of live cancer cell obtained in Chapter 2 to group apoptotic fragments into intact cells that represent dead cells induced by anti-cancer drug. Fig.34 demonstrates the performance comparison of automatic

apoptosis counting and manual counting. Fig.35 demonstrates more examples of automatic counting in our experiments.



(a) Apoptosis counted automatically (b) Apoptosis counted manually

Fig 34 The performance comparsion of apoptotic counting and manual counting

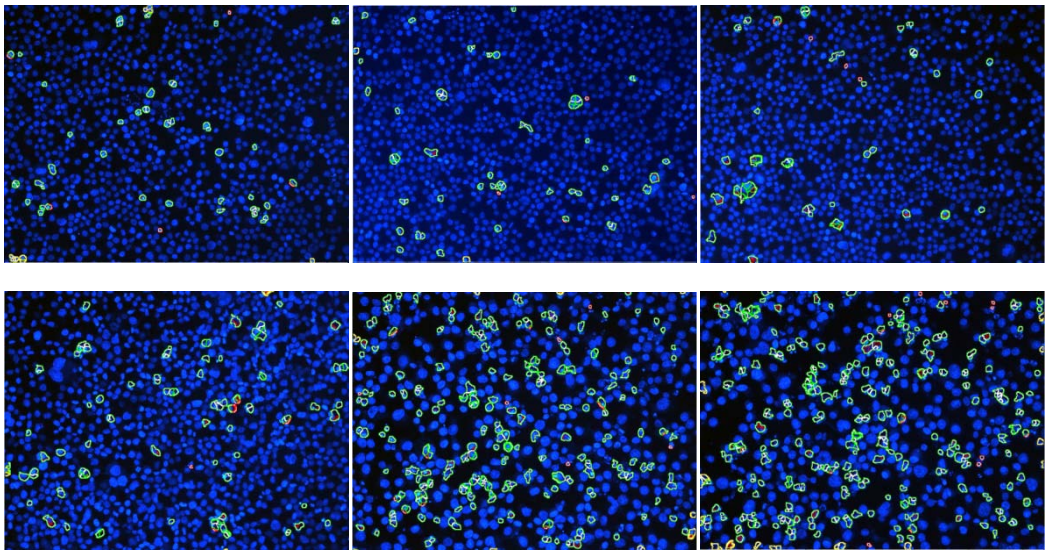


Fig 35 More examples of automatic apoptosis counting

# VII. Conclusions and Contributions

In this thesis, to implement automatic recognition of apoptosis in microscopic images, a series of image processing algorithms were proposed and analyzed, some important features for apoptosis were extracted, different classification methods were compared, and the results showed that our algorithm could obtain very satisfactory performance for automatic recogntion of apoptosis.

In image preprocessing stage, I analyzed the noise distribution in microscopic images and summarized the common filtering methods used in medical images. In order to remove background noise and preserve boundaries, an adaptive gradient-based and anisotropic diffusion equation-filtering algorithm was proposed. Several fitering algorithm was compared and it turned out that our algorithm outperforms others.To solve the limitatation of gradient-based filter, an information entropy based noise detection and adaptive median filtering algorithm was proposed and the results demonstrated that it can furtherly remove particle noise which can be left out by gradient-based filter. In image enhancement, a neighborhood based contrast enhancement algorithm was proposed to solve the uneven distribution of microscopic images.

In image segmentation stage, some widely used segmentation algorithm was summarized and compared. The shortcomings of GVF model for microscopic images were analyzed. In order to detect the boundary of cell in microscopic images, a GVF snake model with size constraint was proposed. The segmentation performance of several algorithm including ours were compared

and it turned out that our model works very well in detecting cell boundaries.

In feature extraction stage, ten features including illumination, size, texture were extracted analyzed.

In cell recognition and classification stage, decision tree algorithm and bagging algorithm were used to analyze classification accuracy and feature importance. It turned out that our classification algorithm can obtain very good performance in recognizing live cell and apoptotic cell.

Although our algorithm was initially designed for the automatic recognition of apoptosis, they are still applicable to other microscopic image related applications so long as some specific parameters are adjusted. In the future, my research work will concentrate on the design of new feature descriptors that is robust to complex conditions of microscopic images.

# Bibliography

[1] Hacker G. “The morphology of apoptosis”, Cell Tissue Res 301:5-17, 2000

[2]Kerr JF, Wyllie AH, Currie AR, “Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics”, Br J Cancer 26: 239-57,1972

[3]Brinkmann M, New technologies for automated cell counting based on optical image analysis, Cytotechnology 38, 119-127, 2002

[4]A. MOLDER, M. SEBESTA, Non-invasive, label-free cell counting and quantitative analysis of adherent cells using digital holography, Journal of Microscopy, 232, 240-247, 2008

[5]Khryashchev V, Priorov A L, Image denoising using adaptive switching median filter. IEEE International Conference on Image Processing, 2005:117-120

[6]NallaperUmal K, Varghese J, Saudia S. Adaptive rank-ordered switching median filter for salt & pepper impulse noise reduction. IEEE Annual India Conference, 2006: 1-6

[7]Wang Y C, Liang D Q, Ma H and Wang Y. An algorithm for image denoising based on mixed filter. Proceedings of the 6th World Congress on Intelligent Control and Automation, IEEE 2006: 9690-9693

[8]Yu Yongjian, Acton Scott T. Speckle reducing anisotropic diffusion. IEEE Transaction on Image Processing, 2002, 11(11): 1260-1269

[9]Izquierdo M E, Ghanbari M. Texture smoothing and object segmentation using feature-adaptive weighted Gaussian filtering. IEEE 1998: 650-655

[10]Basu M. Gaussian-based edge-detection methods- A survey. IEEE Transactions on System, Man and Cybernetics, 2002, 3(32): 252-260

[11]Fu S J, Ruan Q Q, Geng Y L and Wang W Q. Feature-oriented coupled bidirection flow for image denoising and edge sharping. Tencon 2005 IEEE Region 10: 1-5

[12]Ye Zhen. Wavelet domain multiresolution markov models for image segmentation and denoising applications. Proquest information and learning company, 2005: 78-90

[13]Portilla J, Strela V. Image denoising using scaled mixtures of Gaussians in the wavelet domain. IEEE Transactions on image processing, 2004, 11(12): 1338-1351

[14]Liu J, Moulin P. Image denoising based on scale-space mixture modeling of wavelet coefficient. IEEE International Conference on Image Processing, Kobe, Japan, 1999, 1:386-390

[15]Dongwook C, Bui T D. Multivariate statistical approach for image denoising. IEEE International Conference on Acoustics, Speech and Signal Processing, 2005: 589-592

[16]Lamberti F, Montrucchio B, Sanna A. CMBFHE: A novel contrast enhancement technique based on cascaded multistep binomial filtering histogram equalization. IEEE Transaction on Consumer Electronics, 2006, 52 (3): 949-952

[17]Matz S C , de Figueiredo R J P. Nonlinear image contrast enhancement based on munsell's scale. IEEE ICASSP 2006: 949-952

[18]Chen Z Y, Abidi B R. A generalized and automatic image contrast enhancement using gray level grouping. IEEE ICASSP 2006 965-968

[19]William K, Farmarz D. Fast computation techniques for pseudoinverse and Weiner image restoration. IEEE Transactions on Computer, 1997,26 (6):571-580

[20]Kazubek M. Wavelet domain image denoising by thresholding and Weiner Filtering. IEEE Signal Processing, 2003,10(11): 324-326

[21]Ayers G R, Dainty J C. Iterative blind deconvolution method and its application. Optics Letter, 1998, 13(7):547-549

[22]Chan T F, Wong C K. Total variation blind deconvolution. IEEE Transactions on Image Processing, 1998, 17(3):370-375

[23]Kunder D, Hatzinakos D. A novel blind deconvolution scheme for image restoration using recursive filtering. IEEE Transactions on Signal Processing, 1998,46(2)

[24]Zimmerman J B, Pizer S M. An evaluation of the effectiveness of adaptive histogram equalization for contrast enhancement. IEEE Transactions on Medical Imaging, 1998, 7(4): 304-312

[25]Zhang Y J. Improving the accuracy of direct histogram specification. Electronics Letters, 1992, 28 (3): 213-214

[26]Yang X, Xiao Q, Raafat H. Direct mapping between histogram: An improved interactive image enhancement method, Proc. Of IEEE Int. Conf. on Systems, Man and Cybernetics, Charlottesville, Virginia, 1991: 243-247

[27]Rodenacker. K.. Aubele. P. Groping for quantitative digital 3-D image analysis: an approach to quantitative fluorescence in situ hybridization in thick tissue sections of prostate carcinoma. Anal.Cel. Pathol.15,19-29l

[28]Irinopoulo, T., Vassy, J., Beil, M., Nicolopoulo, P., Encaoua, D. & Rigaut, J. (1997) 3-D DNA image cytometry by confocal scanning lasermicroscopy in thick tissue blocks of prostatic lesions. Cytometry, 27, 99-105.

[29]Visser, T., Groen, F. & Brakenhoff, G. (1991) Absorption and scattering correction in fluorescence confocal microscopy. J. Microsc. 163, 189-200.

[30]Liljeborg, A., Czader, M. & Porwit, A. (1995) A method to compensate for light attenuation with depth in 3D DNA image cytometry using a confocal scanning laser microscope. J. Microsc. 177, 108-114.

[31]Pisano, E., Zong, S., Hemminger, M. et al. (1998) Contrast limited adaptive histogram equalization image processing to improve the detection of simulated spiculations in dense mammograms. J. Digital Imaging, 11, 193-200.

[32]Malpica, N., de Solorzano, C. O., Vaquero, J. J., Santos, A., Vallcorba,I., Garcia-Sagredo, J. M., and del Pozo, F. (1997). Applying watershed algorithms to the segmentation of clustered nuclei. Cytometry 28, 289-297.

[33]Meyer, F., and Beucher, S. (1990). Morphological segmentation. J Visual Communication and Image Representation 1, 21-46.

[34]Wahlby, C., Sintorn, I. M., Erlandsson, F., Borgefors, G., and Bengtsson, E. (2004). Combining intensity, edge and shape information for 2D and 3D segmentation of cell nuclei in tissue sections. J Microsc 215, 67-76.

[35]Li G, Liu T, 3D cell nuclei sgementation based on gradient flow tracking. BMC Cell Biol. 2007 Sep 4; 8:40

[36]Carlos O, Arrate M, Segmentation of touching cell nuclei using a two-stage graph cut model. 16th Scandinavian Conference on Image Analysis, Berlin, Heidelberg: Springer-Verlag, 2009,pp. 410-419

[37]Xiangzhi Bai, Changming Sun, Touching cells splitting by using concave points and ellipse fitting, 2008 Digital Image Computing: Techniques and Applications, 2008,pp.271-278

[38]Hashemi RH, Nrandley WG. MRI the basic. Lippincott Williams and Wilkins, 2002

[39]Wagner RF. Statistics of speckle in ultrasound B-scans. IEEE Trans Sonics Ultrosonics 1983; 30:156-63

[40]Guido V. Artificial neural networks for the segmentation of medical images. Technical Report, 1994, (2)

[41]Cherry SR, Sorenson JA, Phelps ME. Physics in nuclear medicine. 2003

[42]Mallat S G. A theory for multiresolution signal decomposition: The wavelet representation. IEEE Transaction on Pattern Analysis and Machine Intelligence, 1989, 11(7): 927-938

[43]Nobuyuki Otsu (1979). "A threshold selection method from gray-level histograms". IEEE Trans. Sys., Man., Cyber. 9 (1): 62-66.

[44]J. Bilmes: A Gentle Tutorial on the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models. Technical Report, University of Berkeley, ICSI-TR-97-021, 1997.

[45]K. Engel (2006). Real-time volume graphics,. pp. 112-114

[46]J. P. Costella, 2011, A superior edge detection gradient operator

[47]R. Boyle and R. Thomas *Computer Vision: A First Course*, Blackwell Scientific Publications, 1988, pp 50 - 51

[48]LS. Davis, "A survey of edge detection techniques", Computer Graphics and Image Processing, vol 4, no. 3, pp 248-260, 1975

[49]Canny, J., A Computational Approach To Edge Detection, IEEE Trans. Pattern Analysis and Machine Intelligence, 8(6):679-698, 1986

[50]Hu M K. Visual pattern recognition by moment invariants. IRE Transactions on Information Theory, 1962, vol: IT-8:179-187

[51]Teague M. Image analysis via the general theory of moments. J. Opt. Soc. Amer. 1980, 70 (8):920-930

[52]Leu J G. Computing a shape moments from its boundary. Pattern Recognition. 1991, 10: 949-957

[53]Sklanskv J, Gonzalez V. Fast polygonal approximation of digitized curves. Pattern Recognition, 1980, 12: 327-331

[54]Pertz J C, Vidal E. Optimum polygonal approximation of digitized curves. Pattern Recognition Letters, 1994, 15: 743-750

[55]Sarkar D. A simple algorithm for detection of significant vertices for polygonal approximation of chain-coded curves. Pattern Recognition Letters, 1993, 14(12): 959-964

[56]Rajan P K, Davidson J M. Evaluation of corner detection algorithms. IEEE 1989: 29-33

[57]Reche P, Urdiales C. Corner detection by means of contour local vectors. Electronics Letters, 2002, 38(14): 699-701

[58]Persoon E, Fu K S. Shape discrimination using Fourier descriptors, IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986, 8: 388-397

[59]Kauppien H, Sepanen T. An experiment comparison of autoregressive and Fourier-based descriptors in 2D shape classification. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1995,2: 201-207

[60]Shen D, Horace H S. Discriminative wavelet shape descriptors for recognition of 2-D patterns. Pattern Recognition, 1999,32(2):151-165

[61]The C H, Chin R T. Image analysis by the method of moments. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1980, 10 (4): 496-513

[62]Belkasim S O, Shridhar M and Ahmadi M. Pattern Recognition with moment invariant: A comparative study and new results. Pattern Recognition, 1991, 24:1117-1138

[63]Zahn C T. Graph-theoretic methods for detecting and describing gestalt clusters. IEEE Transaction on Computing, 1971, 20: 68-86

[64]Chang C C, Buehrer D J. A shape recognition scheme based on relative distance of feature points from the centroid. Pattern Recognition,1991, 24(11): 1053-1063

[65]Wu Z, Leahy R. An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993,15 (11): 1101-1113.

[66]Ojata T, Pietikainen M and Maenpaa T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 1(42): 215-228

[67]He D C, Wang L. Texture Unit, Texture spectrum, and texture analysis. IEEE Transaction on Geosciences and Remote Sensing, 1990, 28(4): 509-512

[68]Khouzani J K, Zadeh S H. Rotation-invariant multiresolution texture analysis using Radon and Wavelet transforms. IEEE Transactions on Image Processing, 2005,14(6): 783-795

[69]Kokare M, Biswas P K, Chatterji B N. Texture image retrieval using new rotated complex wavelet filters. IEEE Transactions on System, Man and Cybernetics-Part B, 2005, 35(6):1168-1178

[70]Selesnick I W, Baraniuk R G and Kingsbury N G. The dual-tree complex wavelet transforms. IEEE Signal Processing Magazine, 2005, 22(6): 123-151

[71]Rowley H A, Baluja S and Kanade T. Neural network-based face detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(1): 23-38

[72]Li W D, Ong S K and Nee A Y. A hybrid method for recognizing interacting machining features. International Journal of Production Research, 2003,41 (9): 1887-1980

[73]MacQueen J .Some Methods for classification and analysis of multivariate observations. Proc. 5th Berkley Symp. Math. Statist, Prob,1967,1:281-297

[74]Huang Z. Extensions to the K-means algorithm for clustering large data sets with categorical values. Data Mining and Knowledge Discovery, 1998, 2:283-304

[75]Khan S S, Ahmad A. Cluster center initialization algorithm for K-Means clustering. Pattern Recognition Letters, 2004(25): 1293-1302

[76]Xu R, Wursch D. Survey of clustering algorithm. IEEE Transactions on Neural Networks, 2005, 16(3): 655-662

[77]Hart P E. The condensed nearest neighbor rule, IEEE Transactions on Information Theory, 1968, 14(3): 515-516

[78]Pan J S, Qiao Y L and Sun S H. A fast k nearest neighbors classification algorithm. IEEE Transactions on Fundamentals, 2004, E87-A(4): 961-963

[79]Zhang B, Srihari S N. Fast K-Nearest Neighbor using clustering-based tree. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(4): 525-528

[80]Chen L, Yap K H. A fuzzy K-Nearest-Neighbor algorithm to blind image deconvolution. IEEE International Conference on Systems, Man and Cybernetics, 2003-03:2049-2054

[81]Quinlan J R. Induction of decision tree. Machine Learning, 1986 (2): 81-106

[82]Quinlan J R. Programming for machine learning. CA: SanMateo, 1993

[83]Aldefeld B. An automatic recognition of 3D structure from 2D representations. Computer-Aided Design, 1983, 15(2): 59-64

[84]Shokri Z S, Kamel M S. The mathematical and numerical properties of the fuzzy C-Means algorithm. Fuzzy Sets and Systems, 1992, (49): 181-191

[85]Happner F, Klawonn F, Kruse R. Fuzzy cluster analysis: Methods for classification. Data Analysis, and Image Recognition, New York, Wiley, 1999

[86]Shin C, Yun U, Kim H and Park S. A hybrid approach of neural networks and memory-based learning to data mining. IEEE Transactions on Neural Networks, 2000, 11(3): 637-646

[87]Hopfield J J. Neurons with graded response have collective computational properties like those of two state neurons. Proc. Natl. Acad. Sci. USA, 1984, 81 (8): 3088-3092

[88]Chua L O, Yang L. Cellular neural networks: Theory. IEEE Transactions on Circuits and Systems, 1988, 35 (10): 1257-1272

[89]G. Matheron, Randoms Sets and Integral Equation, Wiley, New York, 1978

[90]Hacker G, The morphology of apoptosis, Cell Tissue Res 301:5-17, 2000

[91]Hailing Liu, “Automatic recognition of apoptosis in cancer cell images”, presented at the Symposium onf Emerging Topics in Contral and Modeling Biomedical Systems 201, Urbana_Champaign, Urbana, IL USA

[92]Yang L, Meer P, “Unsupervised segmentation based on robust estimation and color active contour models” IEEE Trans. Inf. Technol. Biomed. 9, 475-486

[93]C. Xu and J.L.Prince, “Snakes, shapes, and gradient vector flow”, IEEE Trans. Image Processing, vol. 7,pp. 359-369, Mar.1998

[94]E. Ciaccio, S. Dunn, and M. Akay. Biosignal pattern recognition and interpretation systems: Part i. IEEE Eng. In Med. Biol. 13:89-97, 1993

[95]Spyridonos P, Ravazoula P, Cavouras D, Berberidis K, Nikiforidis G. Computer based grading of heamatoxylin-eosin stained tissue sections of urinary bladder carcinomas. Med Inform Internet Med 2001;26(3):179—90.

[96]Harralick R, Shanmugam K. Textural features for image classification. IEEE Trans Syst Man Cybernet 1973;3(6): 610—21.

[97]Der Poel HV, Schaafsma HE, Vooijs PG, Debruyne FM, Schalken JA. Review article. Quantitative light microscopy in urological oncology. J Urol 1992; 148:1-13.

[98]Van Velthoven R, Petein M, Zlotta A, Oosterlinck W, Meijden A, Zandona C, et al. Computer-assisted chromatin texture characterization of feulgen-stained nuclei in a series of 331 transitional bladder cell carcinomas. J Pathol 1994; 173: 235—42.

[99]Walker RF, Jackway PT, Lovell B. Cervical cell classification via co-occurrence and Markov random field features. In: Proceedings of digital image computing: techniques and applications; 1995.

[100]Ohanian P, Dubes R. Performance evaluation for four classes of textural features. Pattern Recogn 1992; 25(8):819—33.

[101]Choi H-K, Vasko J, Bengtsson E, Jarkrans T, Malmstrom U, Wester K, et al. Grading of transitional cell bladder carcinoma by texture analysis of histological sections. Anal Cell Pathol 1994;6:327—43.

[102]Walker R, Jackway P, Longstaff I. Improving co-occurrence matrix feature discrimination. In: Maeder A, Lovell B, editors. Digital image computing: techniques and applications. Brisbane, Australia; 1995, p. 643—8.

[103]Liu. T, Young. G, “Space analysis of grey matter diffusivity: methods and applications.” NeuroImage 31, 51-65, 2006

[104]Quinlan, J.R., “Learning effective classification procedures and their application to chess ending games.” Machine Learning: An Artificial Intelligence Approach. Ed. M. Kaufmann. Vol. III.1983.463-482

[105]Quinlan, J.R., “Induction of Decision Trees,” Machine Learing. Vol.1. 1986.81-106

[106]Prather, J.C., Lobach.D.F., “Medical data ming knowledge discovery in a clinical data warehouse.” Proc AMIA Annual Fall Symposium. 1997:p.101-105

[107]H.W. Ian, E.F., “Data mining: Practical machine learning tools and techniques.” 2005: Morgan Kaufmann

[108]B. Leo, “Bagging Predictors, in Machine Learning.” 1996.p.123-14

[109]S.E. Umbaugh, “Computer Imaging: Digital Image Analysis and Processing”. CRC Press. Florida, 2005

[110]L. Yaroslavsky, M. Eden, Fundamentals of Digital Optics, Birkhauser, Boston, 1996

[111] T. Zhang, B. Fang, Y. Yuan, Y.Y. Tang, Z. Shang, D. Li, and F. Lang, “Multiscale Facial Structure Representation for Face Under Varying Illumination, Pattern Recognition, vol. 42, no. 2, pp. 252-258, Recognition February 2009.

[112]P. Perona, “Steerable-scalable kernels for edge detection and junction analysis”, Image Vis Comput., vol. 10, pp.663-672,1992

[113]Kervrann, C., Legland, D. & Pardini, L. (2004) Robust incremental compensation of the light attenuation with depth in 3D fluorescence microscopy.

J. Microsc. 214, 297-314.

# ABSTRACT

## Automatic Recognition of Apoptotic cells in Fluorescence Microscopic Image

By Liu Hailing

Advisor: Prof. Shin YoungSuk, Ph.D.

Department of Information and Communication Engineering

Graduate School of Chosun University

With the advance of computer techniques, digital medical images play an important role in clinical diagnosis and treatment. In this paper, I focused on the description of the key techniques used in the automatic recognition of apoptosis using fluorescence microscopic images, which include image preprocessing technique, image segmentation technique, feature extraction technique and automatic classification and recognition technique. In the description of each technique, I reviewed the existing methods used and proposed some new and robust processing algorithms, which achieve good performance in automatic recognition of apoptosis.

In image preprocessing step, I analyzed the main quality degradation factors and common types of noise in microscopic images. According to characteristics of image, I proposed a gradient-based anisotropic filtering algorithm, which is able to filter out background noise and pulse noise and preserve the information of boundaries as well. Considering its bad performance for tiny particles in image, an information entropy based noise detection method and adaptive median filtering algorithm were proposed to address this problem. Lastly, a neighborhood contrast based boundary enhancement algorithm was proposed to improve the contrast of images.

In image segmentation step, the existing commonly used techniques for image segmentation were reviewed and the conception and application of snake model and its improved edition, GVF snake model were described in detail. GVF model has great advantage to other snake models in more freedom of initial contour position and capability to detect concaved boundary, however it is weak in detecting haloed or low contrast boundaries and noise-polluted boundaries. To address this problem, I proposed a size based GVF snake model to limit the movement of snake. Snake model was used to obtain the boundary of separated cells and for overlapping cells, a seriers of separation manipulation was also stated in detail.

In feature extraction step, I reviewed the existing main techniques and described various features used in our experiment.

In recognition step, I introduced the basic framework of microscopic image recognition. Bagging decision tree algorithm was used to measure the

classification ability of 10 featues and classification accuracy.

Lastly, the experimental results were compared and analyzed, which demonstrated the success of our algorithms applied to recognition of apoptosis.