

This is the published version of this work:

Pham, T., & Tran, D. (2006). Gaussian Mixture and Markov Models for Cell-Phase Classification in Microscopic Imaging. In M. Jamshidi (Ed.), Proceedings of the 2006 IEEE/SMC International Conference on System of Systems Engineering (pp. 328-333). USA: IEEE, Institute of Electrical and Electronics Engineers.
<https://doi.org/10.1109/SYSOSE.2006.1652316>

This file was downloaded from:

<https://researchprofiles.canberra.edu.au/en/publications/gaussian-mixture-and-markov-models-for-cell-phase-classification->

©2006 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works

Notice:

The published version is reproduced here in accordance with the publisher's archiving policy 2006.

Gaussian Mixture and Markov Models for Cell-Phase Classification in Microscopic Imaging

Tuan D. Pham

School of Information Technology
James Cook University
Townsville, QLD 4811
Australia
tuan.pham@jcu.edu.au

Dat T. Tran

School of Information Sciences and Engineering
University of Canberra
Canberra, ACT 2601
Australia
dat.tran@canberra.edu.au

Abstract – *Studies of drug effects on cancer cells are performed through measuring cell cycle progression such as inter phase, prophase, metaphase and anaphase in individual cells. Such studies require the processing and analysis of huge amounts of image data. Manual image analysis is very time consuming thus costly, potentially inaccurate and poorly reproducible. Stages of an automated cellular imaging analysis consist of segmentation, feature extraction, classification, and tracking of individual cells in a dynamic cellular population. Image classification of cell phases in a fully automatic manner presents the most difficult task of such analysis. We considered applying several versions of Gaussian mixture and Markov models for automating the classification of cell nuclei in different mitotic phases recorded over a period of twenty-four hours at every fifteen minutes with a time-lapse fluorescence microscopy. The experimental results have shown that the proposed methods are effective and have potential for higher performance.*

Keywords: Cellular imaging, identification, Gaussian mixture models, Markov models.

I. INTRODUCTION

High content screening by automated fluorescence microscopy is becoming one of the most widely used research tools to assist scientists in understanding the complex process of cell division or mitosis [1]-[3]. Its power comes from the sensitivity and resolution of automated light microscopy with multi-well plates, combined with the availability of fluorescent probes that are attached to specific subcellular components, such as chromosomes and microtubules, for visualization of cell division or mitosis using standard epifluorescence microscopy techniques [4]. By employing a carefully selected reporter probes and filters, fluorescence microscopy allows specific imaging of phenotypes of essentially any cell component [5]. With these probes we can determine both the amount of a cell component, and most critically, its distribution within the cell relative to other components. Typically, 3-4 different components are localized in the same cell using probes that excite at different wavelengths. Any change in cell physiology would cause a redistribution of one or more cellular components, and

this redistribution provides a certain cytological marker that allows for scoring of the physiological change.

Automated time-lapse fluorescence microscopy imaging provides an important method for the observation and study of cellular nuclei in a dynamic fashion [6], [7]. An essential task for high content screening is to measure cell cycle progression (inter phase, prophase, metaphase, and telophase) in individual cells as a function of time. Cell cycle progress can be identified by measuring nuclear changes. Stages of an automated cellular imaging analysis consist of segmentation, feature extraction, classification, and tracking of individual cells in a dynamic cellular population; and the classification of cell phases is considered the most difficult task of such analysis [8], [9], [10].

In time-lapse microscopy images are usually captured in a time interval of more than 10 minutes. During this period dividing nuclei may move far away from each other and daughter cell nuclei may not overlap with their parents. Given the advanced fluorescent imaging technology, there still remain technical challenges in processing and analyzing large volumes of images generated by time-lapse microscopy. The increasing quantity and complexity of image data from dynamic microscopy renders manual analysis unreasonably time-consuming. Therefore, automatic techniques for analyzing cell-cycle progress are of considerable interest in the drug discovery process.

To automate the process of identifying cellular phases using time-lapse fluorescence microscopic image sequences, we first apply a shape-and-size based method which merges the over-segmented nuclear fragments. Secondly we extract useful features to discriminate the shapes and intensities of different image cell phases. We then use these image features to build cell phase models using several computational algorithms. To identify an unknown cell, we extract its features then compare those with the phase models. The methods for image classification are Gaussian mixture and Markov models. We will present the proposed methodology and discuss the experimental results in the following sections.

II. GAUSSIAN MIXTURE MODELING

Gaussian mixture model (GMM) is an effective model capable of achieving high recognition accuracy for pat-

tern recognition. Gaussian mixture modeling is a statistical clustering method. A number of prototypes are generated from the training feature vectors by representing the feature space as a mixture of Gaussian distributions. Each prototype consists of a set of model parameters including mean vector, covariance matrix and mixture weight. Parameters are trained in an unsupervised classification method.

A. GMM procedure

Given a training set of T feature vectors $\mathcal{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T\}$, where each source vector $\mathbf{x}_t = (x_{t1}, x_{t2}, \dots, x_{tK})$ is of K dimensions. $\lambda = \{\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_N\}$ represents the set of N Gaussian model parameters $\mathbf{g}_n = \{w_n, \mu_n, \Sigma_n\}$, $n = 1, 2, \dots, N$, where w_n are mixture weights satisfying $\sum_{n=1}^N w_n = 1$, μ_n are mean vectors and Σ_n are covariance matrices.

In general, the GMM design can be stated as follows. Given a training set \mathcal{X} and the size N of the GMM, we seek to find the model λ such that an estimation function is optimised. We review in the subsequent section three estimation functions and the corresponding techniques which are maximum likelihood estimation, minimum fuzzy c -means squared-error estimation and minimum fuzzy entropy squared-error estimation for the GMM.

B. Maximum likelihood estimation

In order to maximise the likelihood function $P(X|\bar{\lambda})$, the following Q -function is maximised [11]

$$\begin{aligned} Q(\lambda, \bar{\lambda}; X) &= \sum_{n=1}^N \sum_{t=1}^T P(n|\mathbf{x}_t, \lambda) \log P(\mathbf{x}_t, n|\bar{\lambda}) \\ &= \sum_{n=1}^N \sum_{t=1}^T P(n|\mathbf{x}_t, \lambda) \log[\bar{w}_n G(\mathbf{x}_t, \bar{\mu}_n, \bar{\Sigma}_n)] \end{aligned} \quad (1)$$

where $P(n|\mathbf{x}_t, \lambda)$ is the *a posteriori* probability for the n th mixture, $n = 1, \dots, N$, $\bar{\lambda}$ is a reestimate of λ , and $G(\mathbf{x}_t, \mu_n, \Sigma_n)$ are the K -variate Gaussian component densities with mean vectors μ_n and covariance matrices Σ_n

$$G(\cdot) = \frac{1}{(2\pi)^{\frac{K}{2}} |\Sigma_n|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (\mathbf{x}_t - \mu_n)' \Sigma_n^{-1} (\mathbf{x}_t - \mu_n) \right\} \quad (2)$$

where $G(\cdot) = G(\mathbf{x}_t, \mu_n, \Sigma_n)$, $(\mathbf{x}_t - \mu_n)'$ is the transpose of $(\mathbf{x}_t - \mu_n)$, Σ_n^{-1} is the inverse of Σ_n , and $|\Sigma_n|$ is the determinant of Σ_n .

Maximizing the Q -function with respect to $\bar{\lambda}$ results in the following GMM algorithm

- 1) Given a training data set $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T\}$, where $\mathbf{x}_t = (x_{t1}, x_{t2}, \dots, x_{tK})$, $t = 1, 2, \dots, T$.
- 2) Initialize the *a posteriori* probability values $P(n|\mathbf{x}_t, \lambda)$, $1 \leq t \leq T$, $1 \leq n \leq N$, at random
- 3) Given $\epsilon > 0$ (small real number).
- 4) Set $i = 0$ and $D^{(i)} = 0$. Iteration:
 - a. Compute GMM parameters

$$\begin{aligned} \bar{w}_n &= \frac{1}{T} \sum_{t=1}^T P(n|\mathbf{x}_t, \lambda), & \bar{\mu}_n &= \frac{\sum_{t=1}^T P(n|\mathbf{x}_t, \lambda) \mathbf{x}_t}{\sum_{t=1}^T P(n|\mathbf{x}_t, \lambda)} \\ \bar{\Sigma}_n &= \frac{\sum_{t=1}^T P(n|\mathbf{x}_t, \lambda) (\mathbf{x}_t - \bar{\mu}_n) (\mathbf{x}_t - \bar{\mu}_n)'}{\sum_{t=1}^T P(n|\mathbf{x}_t, \lambda)} \end{aligned} \quad (3)$$

- b. Update the *a posteriori* probability values

$$P(n|\mathbf{x}_t, \bar{\lambda}) = \frac{P(\mathbf{x}_t, n|\bar{\lambda})}{\sum_{k=1}^N P(\mathbf{x}_t, k|\bar{\lambda})} = \frac{\bar{w}_n G(\mathbf{x}_t, \bar{\mu}_n, \bar{\Sigma}_n)}{\sum_{k=1}^N \bar{w}_k G(\mathbf{x}_t, \bar{\mu}_k, \bar{\Sigma}_k)} \quad (4)$$

- c. Compute $D^{(i+1)}$

$$\begin{aligned} D^{(i+1)} &= -\log P(X|\bar{\lambda}) = -\sum_{t=1}^T \log P(\mathbf{x}_t|\bar{\lambda}) \\ &= -\sum_{t=1}^T \log \sum_{n=1}^N \bar{w}_n G(\mathbf{x}_t, \bar{\mu}_n, \bar{\Sigma}_n) \end{aligned} \quad (5)$$

- d. If

$$\frac{|D^{(i+1)} - D^{(i)}|}{D^{(i+1)}} > \epsilon \quad (6)$$

set $D^{(i)} = D^{(i+1)}$, $i = i + 1$ and go to step (a).

C. Minimum fuzzy c -means squared-error estimation

The fuzzy c -means (FCM) squared-errors function [12] is used

$$J(U, \lambda, \bar{\lambda}; X) = \sum_{n=1}^N \sum_{t=1}^T u_{nt}^m d_{nt}^2 \quad (7)$$

where d_{nt} is redefined as follows

$$\begin{aligned} d_{nt}^2 &= -\log P(\mathbf{x}_t, n|\bar{\lambda}) \\ &= -\sum_{n=1}^N \sum_{t=1}^T P(n|\mathbf{x}_t, \lambda) \log[\bar{w}_n G(\mathbf{x}_t, \bar{\mu}_n, \bar{\Sigma}_n)] \end{aligned} \quad (8)$$

Minimizing the J -function with respect to $\bar{\lambda}$ results in the following FCM-GMM algorithm

- 1) Given a training data set $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T\}$, where $\mathbf{x}_t = (x_{t1}, x_{t2}, \dots, x_{tK})$, $t = 1, 2, \dots, T$.
- 2) Initialize the membership values u_{nt} , $1 \leq t \leq T$, $1 \leq n \leq N$, at random
- 3) Given $\epsilon > 0$ (small real number).
- 4) Set $i = 0$ and $D^{(i)} = 0$. Iteration:
 - a. Compute GMM parameters

$$\begin{aligned}\bar{w}_n &= \frac{\sum_{t=1}^T u_{nt}^m}{\sum_{t=1}^T \sum_{n=1}^N u_{nt}^m}, & \bar{\mu}_n &= \frac{\sum_{t=1}^T u_{nt}^m \mathbf{x}_t}{\sum_{t=1}^T u_{nt}^m}, \\ \bar{\Sigma}_n &= \frac{\sum_{t=1}^T u_{nt}^m (\mathbf{x}_t - \bar{\mu}_n)(\mathbf{x}_t - \bar{\mu}_n)'}{\sum_{t=1}^T u_{nt}^m}\end{aligned}\quad (9)$$

b. Update the fuzzy membership values

$$u_{nt} = \frac{1}{\sum_{k=1}^N (d_{nt}^2/d_{kt}^2)^{1/(m-1)}} \quad (10)$$

c. Compute $D^{(i+1)}$

$$D^{(i+1)} = J(U, \lambda; X) = \sum_{n=1}^N \sum_{t=1}^T u_{nt}^m d_{nt}^2 \quad (11)$$

d. If

$$\frac{|D^{(i+1)} - D^{(i)}|}{D^{(i+1)}} > \epsilon \quad (12)$$

set $D^{(i)} = D^{(i+1)}$, $i = i + 1$ and go to step (a).

D. Minimum fuzzy entropy squared-error estimation

The fuzzy entropy (FE) squared-errors function [12] is used

$$H(U, \lambda, \bar{\lambda}; X) = \sum_{n=1}^N \sum_{t=1}^T u_{nt} d_{nt}^2 + m_E \sum_{n=1}^N \sum_{t=1}^T u_{nt} \log u_{nt} \quad (13)$$

where d_{nt} is redefined as in (8), and $m_E > 0$ is a value which controls the degree of fuzzy entropy. Minimizing the H -function with respect to $\bar{\lambda}$ results in the following FE-GMM algorithm

- 1) Given a training data set $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T\}$, where $\mathbf{x}_t = (x_{t1}, x_{t2}, \dots, x_{tK})$, $t = 1, 2, \dots, T$.
- 2) Initialize the membership values u_{nt} , $1 \leq t \leq T$, $1 \leq n \leq N$, at random
- 3) Given $\epsilon > 0$ (small real number).
- 4) Set $i = 0$ and $D^{(i)} = 0$. Iteration:
 - a. Compute GMM parameters

$$\begin{aligned}\bar{w}_n &= \frac{1}{T} \sum_{t=1}^T u_{nt}, & \bar{\mu}_n &= \frac{\sum_{t=1}^T u_{nt} \mathbf{x}_t}{\sum_{t=1}^T u_{nt}}, \\ \bar{\Sigma}_n &= \frac{\sum_{t=1}^T u_{nt} (\mathbf{x}_t - \bar{\mu}_n)(\mathbf{x}_t - \bar{\mu}_n)'}{\sum_{t=1}^T u_{nt}}\end{aligned}\quad (14)$$

b. Update the fuzzy membership values

$$u_{nt} = \frac{e^{-d_{nt}^2/m_E}}{\sum_{k=1}^N e^{-d_{kt}^2/m_E}} \quad (15)$$

c. Compute $D^{(i+1)}$

$$\begin{aligned}D^{(i+1)} &= H(U, \lambda, \bar{\lambda}; X) \\ &= \sum_{n=1}^N \sum_{t=1}^T u_{nt} d_{nt}^2 + m_E \sum_{n=1}^N \sum_{t=1}^T u_{nt} \log u_{nt}\end{aligned}\quad (16)$$

d. If

$$\frac{|D^{(i+1)} - D^{(i)}|}{D^{(i+1)}} > \epsilon \quad (17)$$

set $D^{(i)} = D^{(i+1)}$, $i = i + 1$ and go to step (a).

III. MARKOV MODELING

We present herein a Markov-chain modeling method for cell-phase identification. The occurrences of phases in a sequence of cells can be regarded as a stochastic process and hence the cell sequence can be represented as a Markov chain where phases are states. The occurrence of the first phase in the sequence is characterized by the initial probability of the Markov chain and the occurrence of the other phase given the occurrence of its previous phase is characterized by the transition probability. Given a training set of cell sequences, the initial and transition probabilities for Markov chains representing cell sequences are calculated and the set of those probabilities is regarded as a Markov model for that sequence.

Given a training set of Q sequences $\mathcal{O} = \{\mathcal{O}_1, \mathcal{O}_2, \dots, \mathcal{O}_Q\}$, each of which is a sequence of T_q random variables $\mathcal{O}_q = \{O_{q1}, O_{q2}, \dots, O_{qT}\}$. Let $\mathcal{S} = \{s_1, s_2, \dots, s_M\}$ be the set of M states in a Markov chain.

We now develop a Markov chain-based modeling method to represent phase sequences found in a given set of cells. Define the following parameters

$$\pi = [\pi_i], \quad \pi_i = Pr(O_{q1} = s_i) \quad (18)$$

$$\mathbf{A} = [a_{ij}], \quad a_{ij} = Pr(O_{qt} = s_j | O_{q(t-1)} = s_i) \quad (19)$$

where $q = 1, 2, \dots, Q$, Q is the number of phase sequences, $t = 2, \dots, T_q$, T_q is the length of the sequence O_q , $i = 1, \dots, M$ and $j = 1, \dots, M$, M is the number of phases.

Phases are states in Markov chains and values in the vector π and matrix \mathbf{A} are state initial probabilities and state transition probabilities, respectively. The set $\lambda = (\pi, \mathbf{A})$ is called the Markov phase model that represents the phase sequences in the training cell set as the Markov chains.

We can obtain

$$\pi_i = \frac{n_i}{\sum_{k=1}^M n_k}, \quad a_{ij} = \frac{n_{ij}}{\sum_{k=1}^M n_{ik}} \quad (20)$$

IV. ALGORITHMS FOR CELL PHASE IDENTIFICATION

We present in this section the training and identification algorithms for the following proposed models: GMM, and GMM-Markov.

A. GMM-based algorithm

The training and classification procedures of this GMM-based algorithm are summarized as follows

Training:

- 1) Given \mathcal{X} as the universe of cell phases.
- 2) Train GMM-based phase models
 - a) Divide the set \mathcal{X} into P distinct subsets X^1, X^2, \dots, X^P , where each X^p contains cells of phase p .
 - b) For each subset X^p , train a GMM-based phase model λ_p of N mixtures using an estimation method presented in Section II

Identification:

- 1) Given a cell of an unknown phase \mathbf{x} .
- 2) Calculate the minimum distance between \mathbf{x} and $\lambda_p, p = 1, \dots, P$

$$d_p = -\log P(\mathbf{x}|\lambda_p) = -\log \sum_{n=1}^N w_{np} G(\mathbf{x}, \mu_{np}, \Sigma_{np}) \quad (21)$$

- 3) Assign \mathbf{x} to phase p that has the minimum distance:

$$p^* = \arg \min_p (d_p) \quad (22)$$

B. GMM-Markov-based algorithm

The training and classification procedures of this GMM-Markov algorithm are summarized as follows.

Training:

- 1) Given \mathcal{X} as the universe of cell phases.
- 2) Train GMM-based phase models
 - a) Divide the set \mathcal{X} into P distinct subsets X^1, X^2, \dots, X^P , where each X^p contains cells of phase p .
 - b) For each subset X^p , train a GMM-based phase model λ_p of N mixtures using an estimation method presented in Section II
- 3) Train Markov model for all phases

- a) Align cells in the set \mathcal{X} as sequences of cells
- b) Extract Q phase sequences $\mathcal{O}_1, \mathcal{O}_2, \dots, \mathcal{O}_Q$ from the set \mathcal{X}
- c) Using Q phase sequences, calculate π and \mathbf{A} according to (20)

Identification:

- 1) Given an unknown sequence of cells $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T\}$
- 2) Identify phase for the first cell \mathbf{x}_1 in the sequence as follows

- a) Calculate the minimum distance between \mathbf{x}_1 and $\lambda_p, p = 1, \dots, P$

$$d_p = -\log P(\mathbf{x}_1|\lambda_p) = -\log \sum_{n=1}^N w_{np} G(\mathbf{x}_1, \mu_{np}, \Sigma_{np}) \quad (23)$$

- b) Calculate the similarity score $S(\mathbf{x}_1, p)$

$$S(\mathbf{x}_1, p) = \frac{\pi_p}{\sum_{k=1}^P (d_p/d_k)^{1/(m-1)}} \quad (24)$$

where $m > 1$.

- c) Assign \mathbf{x}_1 to the phase p^* that has the maximum score:

$$p^* = \arg \max_p S(\mathbf{x}_1, p) \quad (25)$$

- 3) For each cell $\mathbf{x}_i, i = 2, \dots, T$, identify it as follows

- a) Calculate the minimum distance between \mathbf{x}_i and $\lambda_p, p = 1, \dots, P$

$$d_p = -\log P(\mathbf{x}_i|\lambda_p) = -\log \sum_{n=1}^N w_{np} G(\mathbf{x}_i, \mu_{np}, \Sigma_{np}) \quad (26)$$

- b) Calculate the similarity score $S(\mathbf{x}_i, p)$

$$S(\mathbf{x}_i, p) = \frac{a_{p^*p}}{\sum_{k=1}^P (d_p/d_k)^{1/(m-1)}} \quad (27)$$

where $m > 1$ and p^* is the identified phase of the previous cell.

- c) Assign \mathbf{x}_i to the phase p^* that has the maximum score:

$$p^* = \arg \max_p S(\mathbf{x}_i, p) \quad (28)$$

V. EVALUATION EXPERIMENTS AND RESULTS

A. Data Set

The data set contains 375841 cells in 892 nuclear sequences provided by the Department of Cell Biology at the Harvard Medical School to evaluate the presented modeling methods for the cell phase classification system. The average number of cells per sequence is 421. Imaging was performed by time-lapse fluorescence microscopy with a time interval of 15 minutes. Two types of sequences were used denoting drug treated and untreated. Cell cycle progress was affected by drug and some or all of the cells in the treated

sequences were arrested in metaphase. Cell cycle progress in the untreated sequences was not affected. Cells without drug treatment will usually undergo one division during this period of time.

B. Feature Extraction

After the nuclear segmentation has been performed, it is necessary to perform a morphological closing process on the resulting binary images in order to smooth the nuclear boundaries and fill holes inside the nuclei. These binary images are then used as a mask on applied the original image to arrive at the final segmentation. From this resulting image, features can be extracted. The ultimate goal for feature selection is to assign correct phase to cells via the training of some identification technique. In this work, a set of cell-nuclear features are extracted based on the experience of biologists. To identify the shape and intensity differences between different cell phases, a set of 7 features are extracted. These features include maximum intensity, mean, stand deviation, major axis, minor axis, perimeter, and compactness [8].

Because the feature values have different ranges, the scaling of features is therefore necessary by calculating the z -scores [8]:

$$z_{tk} = \frac{x_{tk} - \mu_k}{s_k} \quad (29)$$

where x_{tk} is the k -th feature of the t -th nucleus, μ_k the mean value of all T cells for feature k , and s_k the mean absolute deviation, that is

$$s_k = \frac{1}{T} \sum_{t=1}^T |x_{tk} - \mu_k| \quad (30)$$

C. Initialisation and constraints on parameters during training

It was shown in the literature that no significant difference in pattern recognition was found by using different initialisation methods [13], [14]. Therefore the following parameters of the GMM are initialised: Mixture weights, mean vectors, covariance matrices and fuzzy membership functions. The covariance matrices are diagonal, i.e. $[\Sigma_k]_{ii} = \sigma_k^2$ and $[\Sigma_k]_{ij} = 0$ if $i \neq j$, where σ_k^2 , $1 \leq k \leq K$ are variances. A variance limiting constraint was applied to all GMMs using diagonal covariance matrices. This constraint places a minimum variance value σ_{min}^2 on elements of all variance vectors in the GMM, that is, $\sigma_i^2 = \sigma_{min}^2$ if $\sigma_i^2 \leq \sigma_{min}^2$ [14]. In our experiments, $\sigma_{min}^2 = 10^{-2}$. The degree of fuzziness m and the degree of fuzzy entropy m_E were set to 1.1 and 0.1, respectively.

D. Experimental Results

There are 5 phases to be identified: interphase, prophase, metaphase, anaphase, and arrested metaphase. We divide the data set into 5 subsets for training 5 models and a subset for identification. Each of the 5 training sets for 5 phases contains 5000 cells, which are extracted from the

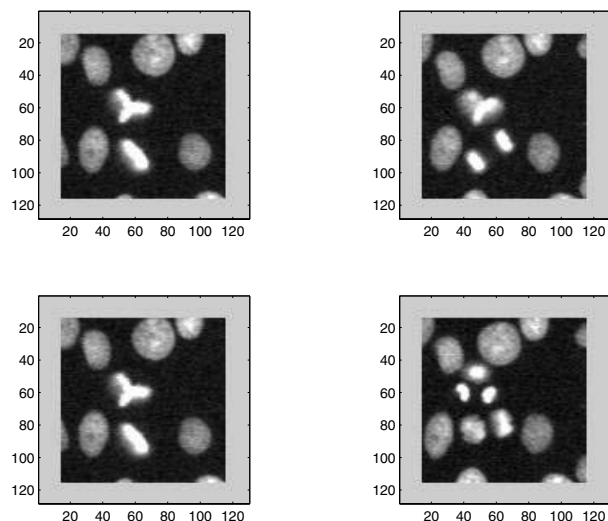


Fig. 1. Nuclear appearance changes during cell mitosis

cell sequences labeled from 590 to 892. These sequences are also used to calculate the Markov model parameters. The identification set contains sequences labeled from 1 to 589. There are 249,547 cells in this identification set. Figure 1 shows the nuclear appearance changes during cell mitosis (top row: split into two; bottom row: split into three). The consecutive image subframes from an image sequence show nuclear size and shape changes during cell mitosis.

Seven features extracted for each cell are considered as the feature vectors for training and evaluating VQ approaches based on the k -means [15], [16], LBG [17], fuzzy entropy (FE) [18], [19], [20] and fuzzy c -means (FCM) [12]; GMM approaches based on the maximum likelihood estimation, minimum fuzzy entropy squared-error estimation and the combinations of VQ and Markov models, and of GMM and Markov models.

Table 1 presents the classification results for Gaussian mixture models which include 8, 16, 32, 64 and 128 Gaussian mixtures, respectively. Table 2 shows the results obtained from the vector quantization (Q) based method [9], [10]. The combined minimum fuzzy c -means squared-error estimation and Markov modeling method achieved better classification rates than the Gaussian mixture modeling methods in all the model sizes and for all the estimation methods. In general, the GMM-based approach outperforms the VQ-based models. The experimental results can be generally noted that the combination of either the GMM or VQ with and Markov modeling method always increases the classification rates.

VI. CONCLUSIONS

We have applied several pattern recognition methods for the classification of cell phases using time-lapse fluorescence microscopic image sequences.

The incorporation of probabilistic analysis using Markov chains into the template matching using Gaussian mixture modeling approach helps improve the classification rates

TABLE I

CLASSIFICATION RESULTS (%) FOR DIFFERENT ESTIMATION METHODS
IN GAUSSIAN MIXTURE MODELING.

Estimation Method	GMM Size				
	8	16	32	64	128
Maximum likelihood	83.4	86.4	87.0	87.2	87.8
Maximum likelihood + Markov	85.7	87.3	88.0	88.2	88.6
Minimum fuzzy entropy	82.4	85.6	86.7	87.1	87.9
Minimum fuzzy entropy + Markov	84.7	86.5	87.7	88.2	88.6
Minimum fuzzy c-means	83.5	85.7	87.1	87.7	88.2
Minimum fuzzy c-means + Markov	85.6	86.9	88.2	88.6	88.9

TABLE II

CLASSIFICATION RESULTS (%) FOR DIFFERENT PARTITION METHODS IN
VQ MODELING.

Partition Method	VQ codebook size				
	8	16	32	64	128
<i>k</i> -means	82.7	84.7	85.5	86.2	85.3
<i>k</i> -means + Markov	84.2	85.5	86.2	86.8	86.7
LBG	83.3	85.8	85.6	86.1	85.5
LBG + Markov	84.4	86.3	86.3	86.8	86.6
Fuzzy entropy	85.0	85.5	85.6	85.9	86.1
Fuzzy entropy + Markov	85.5	86.1	86.6	86.9	87.3
Fuzzy c-means	84.0	85.3	86.5	87.0	88.2
Fuzzy c-means + Markov	85.6	86.2	86.6	87.0	88.4

with various clustering criteria and various estimation methods. From the experimental results, it can be seen that the fuzzy Gaussian mixture model (either fuzzy entropy or fuzzy *c*-means) is superior to the Gaussian mixture modeling method. The fuzzy *c*-means algorithm seeks to partition a data set into a specified number of fuzzy regions which are represented by the corresponding fuzzy prototypes. The degrees of each cellular-image feature vector that belong to different clusters are characterized by the corresponding fuzzy membership grades taking real values between 0 and 1. Thus, the use of the fuzzy *c*-means algorithm provides more effective analysis of the present problem where the image boundaries of different classes are vaguely defined.

The selection of useful features is a very important task for any classifier. In this work we have selected seven features of the nuclear images to train several classifiers. However, the issue of feature selection is worth investigating in our future research as new useful image features can certainly enhance the performance of the proposed classifiers, particularly the combined GMM and Markov modeling based classifier which has achieved the highest classification rate.

Molecular imaging is an exciting area of research in life sciences, which provides an outstanding tool for the study of diseases at the molecular or cellular levels. Some molecular imaging techniques have been implemented for clinical applications [21]. To contribute to this emerging imaging technology, we have presented and discussed several computational models for the identification of cellular phases based on fluorescent imaging data. This task is an important component for any computerized imaging system which

automates the screening of high-content, high-throughput fluorescent images of mitotic cells to aid biomedical or biological researchers to study the mitotic data at dynamic ranges for various applications including the study of the complexity of cell processes, and the screening of novel anti-mitotic drugs as potential cancer therapeutic agents.

Acknowledgments

The time-lapse fluorescence microscopic image data set was provided by our collaborator, the HCNR Center for Bioinformatics, Harvard Medical School. The first author wishes to acknowledge the financial support provided by the ARC Discovery Project Grant DP0665598.

REFERENCES

- [1] Fox, S., Accommodating cells in HTS, *Drug Discovery World*, **5** (2003) 21-30.
- [2] Feng, Y., Practicing cell morphology based screen, *European Pharmaceutical Review*, **7** (2002) 7-11.
- [3] Dunkle, R., Role of image informatics in accelerating drug discovery and development, *Drug Discovery World*, **5** (2003) 75-82.
- [4] Yarrow, J.C., et al., Phenotypic screening of small molecule libraries by high throughput cell imaging, *Comb Chem High Throughput Screen*, **6** (2003) 279-286.
- [5] Murphy, D.B., *Fundamentals of light Microscopy and Electronic Imaging*. Wiley-Liss, 2001.
- [6] Hiraoka, Y., and Haraguchi, T., Fluorescence imaging of mammalian living cells, *Chromosome Res*, **4** (1996) 173-176.
- [7] Kanda, T., Sullivan, K. F., and Wahl G. M., Histone-GFP fusion protein enables sensitive analysis of chromosome dynamics in living mammalian cells, *Current Biology*, **8** (1998) 377-85.
- [8] Chen, X., Zhou, X., and Wong, S.T.C.: Automated segmentation, classification, and tracking cancer cell nuclei in time-lapse microscopy. *IEEE Trans. on Biomedical Engineering*, in press.
- [9] Pham, T.D., Tran, D., Zhou, X., Wong, S.T.C.: An automated procedure for cell-phase imaging identification. *Proc. AI-2005 Workshop on Learning Algorithms for Pattern Recognition*, pp. 52-29, 2005.
- [10] Pham, T.D., Tran, D.T., Zhou, X., and Wong, S.T.C.: Classification of cell phases in time-lapse images by vector quantization and Markov models. *Neural Stem Cell Research*, ed. Erik V. Greer, Nova Science, New York, 2006.
- [11] Huang, X. D., Arik, Y. and Jack, M. A., *Hidden Markov Models For Speech Recognition*. Edinburgh University Press, 1990.
- [12] Bezdek, J.C., *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, New York and London, 1981.
- [13] Rabiner, L. R., Levinson, S. E. and Sondhi, M. M., On the application of vector quantisation and hidden Markov models to speaker-independent, isolated word recognition, *The Bell System Technical Journal*, **62** (1983) 1075-1105.
- [14] Reynolds, D. A. and Rose, R. C., Robust text-independent speaker identification using Gaussian mixture models, *IEEE Trans. Speech and Audio Processing*, **3** (1995) 72-83.
- [15] Cover, T.M. and Hart, P.E., Nearest neighbor pattern classification. *IEEE Trans. Information Theory*, **13** (1967) 21-27.
- [16] Duda, R.O., and Hart, P.E., *Pattern classification and scene analysis*, John Wiley & Sons, New York, 1973.
- [17] Linde, Y., Buzo, A., and Gray, R.M., An Algorithm for Vector Quantization", *IEEE Trans. Communications*, **28** (1980) 84-95.
- [18] Tran, D., and Wagner, M., Fuzzy entropy clustering, *Proc. FUZZ-IEEE 2000 Conf.*, vol. 1, pp. 152-157.
- [19] Tran, D. and Wagner, M., Fuzzy hidden Markov models for speech and speaker recognition, in *Proc. North American Fuzzy Information Society (NAFIPS'99)*, pp. 426-430, 1999.
- [20] Tran, D. and Wagner, M., Fuzzy entropy hidden Markov models for speech recognition, *Proc. Spoken Language Processing (ICSLP)*, vol. 1, pp. 421-424, 2000.
- [21] Special Issue on: Molecular Imaging: Emerging Technology & Biomedical Applications, *Proceedings of the IEEE*, **93**, number 4, 2005.