

Image Based Monitoring for Combustion Systems

J. Chen, *Member, IAENG* and T.-Y. Hsu

Abstract— A novel method of on-line flame detection in video is proposed. It aims to early detect the current state of the combustion system and prevent the system from further degradation and occurrence of failure. The proposed method consists of hidden Markov model (HMM) and multiway principal component analysis (MPCA). MPCA is used to extract the cross-correlation among spatial relationships in the low dimensional space while HMM constructs the temporal behavior of the sequence of the spatial features. The probability distribution of the normal status can be trained by the images collected from the normal operation processes. The proposed method can generate simple probability monitoring charts to track the progress of the current transition state sequence and monitor the occurrence of the observable upsets. To demonstrate the performance of the proposed method, data from the monitoring practice in the real combustion systems are conducted.

Index Terms— Combustion Systems; Hidden Markov Model; Multiway Principal Component Analysis; Statistical Process Control

I. INTRODUCTION

The industrial furnaces, such as fossil fuel-fired furnaces, are widely used to produce electrical power by combusting different fuels [1]. Many combustion problems, including poor flame stability, low combustion efficiency and high pollutant emissions, should be considered. Therefore, the monitoring of combustion processes have become highly desirable to monitor, estimate and even reduce the levels of pollutant emissions [2,3].

One traditional way to monitor the combustion process is flame watch by experienced workers. Nevertheless, the drawbacks are that it is hard to quantify the combustion performance only by human experience and the working environment is usually too poor [4]. Temperature is a general measurement index to monitoring the combustion process. The conventional temperature monitoring devices, such as thermocouples and pyrometers, are single-point measurement and cannot provide the temperature distribution of the flame. With the current development of optical sensing and digital image processing techniques, on-line continuous monitoring of the flame is becoming possible by a spatial and

cost-effective way [3].

Some researchers developed statistical sensitivity methods to monitor the combustion process. Yan et al. (2002) extracted the geometrical and luminous features of the pulverized coal flames to quantify the combustion process and did sensitivity analysis [1]. They demonstrated that some image features could help detect the change of combustion processes. However, most research focused only on the point estimation without considering the region of operating distribution. Whereas an incipient fault occurs in the operation combustor, fault information is often weak; it is difficult to extract weak fault information using direct analysis of the image data because in practice, the flame flicker frequency is not constant and it varies with time. Therefore, it is necessary to create a new process control method to monitor the combustion process.

Hidden Markov models (HMM) are recognized as being appropriate for time sequence data. In a HMM model, each observation is the data sequence depending on the previous elements in the sequence. An HMM is capable of characterizing a doubly embedded stochastic process with an underlying stochastic process that can be observed through another set of stochastic processes because the state of the system is not observable directly [5]. With the inspiration of the HMM framework, a statistical model of spatiotemporal dynamic process variables for the sequence images is developed using HMM and MPCA in this paper. The dynamic data array is constructed by incorporating both static and dynamic image characteristics using the prior and the current images. The strongest relations of the scores are extracted by MPCA. With the extracted scores, the HMM model can be integrated into the existing MPCA method to enhance the capability of statistical process monitoring. Thus, the control chart for a two-dimensional image is developed in this research. The chart is not only an analysis of the two-dimensional spatiotemporal measurements but also capture the statistical characteristics of the practical data.

II. MPCA-HMM BASED MODELING FOR DYNAMIC IMAGES

Flame images exhibit spatially homogeneous repetitions and temporally dynamic patterns. The repetition is the pattern of the burning flame that changes regularly its amplitude, form and color. The sequences of images show a regular pattern repetition and dynamic behavior because this repetition can be found in time as well as in space. This means that each flame image is seen as a point in a given subspace, following a trajectory as time evolves. The modeling dynamic images can locate an appropriate subspace to represent the trajectory and identify the stochastic

This work was sponsored in part by National Science Council, R.O.C. and in part by the project Toward Sustainable Green Technology in the Chung Yuan Christian University, Taiwan, under grant CYCU-98-CR-CE

J. Chen is with Department of Chemical Engineering, Chung-Yuan Christian University, Chung-Li, Taiwan 320, Republic of China (corresponding author to provide fax: 886-3-265-4199; e-mail: jason@wavenet.cycu.edu.tw).

T.-Y. Hsu is with Department of Chemical Engineering, Chung-Yuan Christian University, Chung-Li, Taiwan 320, Republic of China.

trajectory with a method of dynamic stochastic theory. In this research, an integrated image-based model that makes use of two statistic techniques, MPCA and HMM, to describe the spatial and temporal behavior is proposed for monitoring the operation combustors. MPCA is used to extract the spatial features of the trajectory in a smaller space while HMM is used to train temporal behavior using a dynamic probability model.

A. Spatial Feature Extraction: MPCA

The data set of a sequence of τ images can be arranged as a two-dimensional array as shown in Fig. 1. Assume each color image of the size is $N_1 \times N_2$ and the original color encoding is *RGB*. It can be reshaped into a sequence of column vectors of the length $M = N_1 \times N_2 \times 3$. With a sequence of τ color images, the column vectors ($\mathbf{x}_k \in R^M$, $k=1, L, \tau$) are ordered from top to the bottom in a two-dimensional array $\mathbf{X} = [\mathbf{x}_1 \ L \ \mathbf{x}_\tau]^T \in R^{M \times \tau}$, where lower indices indicate that the matrix collects frames from 1 to τ . The temporal mean of each pixel is computed and subtracted from \mathbf{X} , and a new matrix \mathbf{Y} is obtained. To get the information on the variability among the images, MPCA is performed,

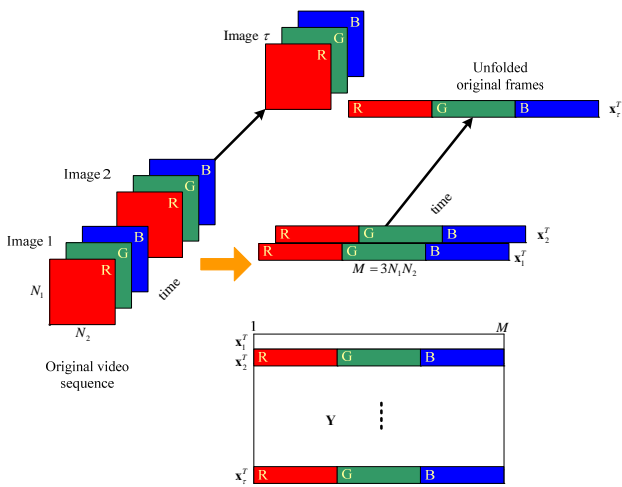


Fig. 1. Unfolded structures of the dynamic textures

$$\mathbf{Y} = \sum_{r=1}^R \mathbf{t}_r \mathbf{p}_r + \mathbf{E} = \mathbf{TP}^T + \mathbf{E} \tag{1}$$

where R is the number of principal components, \mathbf{E} is a residual matrix, and \mathbf{T} represents the score matrix

$$\mathbf{T} = \begin{bmatrix} \mathbf{t}_1^T \\ M \\ \mathbf{t}_k^T \\ M \\ \mathbf{t}_\tau^T \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & L & t_{1R} \\ & & & M \\ t_{k1} & t_{k2} & & t_{kR} \\ & & & M \\ t_{\tau 1} & t_{\tau 2} & L & t_{\tau R} \end{bmatrix} \tag{2}$$

and $\mathbf{P} = [\mathbf{p}_1 \ \mathbf{p}_2 \ L \ \mathbf{p}_R]$ is the loading matrix. This means that dimension reduction is performed by retaining the first R single value of the decomposition. Thus each frame is written as

$$\mathbf{y}_k^T = \hat{\mathbf{y}}_k^T + \mathbf{e}_k = t_{k1}\mathbf{p}_1^T + t_{k2}\mathbf{p}_2^T + L + t_{kR}\mathbf{p}_R^T + \mathbf{e}_k \tag{3}$$

This means that $\hat{\mathbf{y}}_k^T$ is a linear combination of the columns of matrix \mathbf{P} , which contains the basis of the subspace onto which the image vector \mathbf{y}_k^T is projected. The columns of matrix \mathbf{P} contain the eigenvectors whose linear combination generates one flame image. The score vector \mathbf{t}_k^T fixes the weights of the linear combination, so the features of image frames changed in time are extracted. \mathbf{e}_k represents the residual error.

B. Temporal Feature Extraction: HMM

HMM illustrated in Fig. 2 is a double stochastic model, where the non-boxed circles represent a sequence of observations. When HMM is applied to the combustor system, the method for modeling sequential image data is the chain structure that captures interaction between a sequence of inputs,

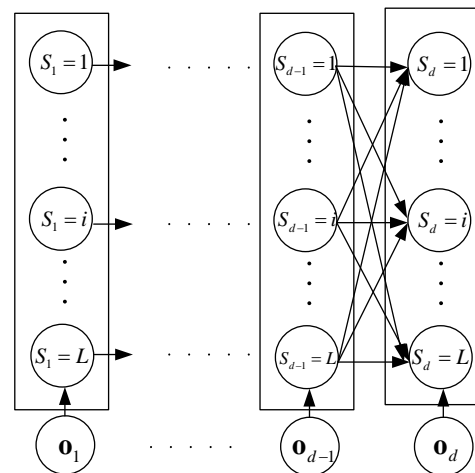


Fig. 2. Graphical model for HMM with the sequential variables

$$\mathbf{o}_k = [t_{1,k} \ t_{2,k} \ L \ t_{R,k} \ v_k]^T \tag{4}$$

where \mathbf{o}_k consists of extracted $t_{r,k}$, $r=1,2,L,R$ over the first R components of \mathbf{T} at time point k and the

corresponding residual variation at time point k , which is defined as

$$v_k = \mathbf{e}_k^T \mathbf{e}_k \quad (5)$$

To capture the interaction in the sequence data, the number of the past τ observation is used to train HMM. First the feature data from the MPCA decomposition, $\{\mathbf{o}_1 \ \mathbf{o}_2 \ \dots \ \mathbf{o}_\tau\}$ is collected. In Fig. 2, there is a set of hidden variables, $\mathbf{s} = \{S_1, \dots, S_d\}$, and a fixed number of states (L). The probability $P(\mathbf{o})$ can be written as a sum of term $P(\mathbf{o}|\mathbf{s})$ over all the possible state sequences,

$$\begin{aligned} P(\mathbf{o}) &= \sum_{\text{all } \mathbf{s}} P(\mathbf{o}|\mathbf{s}) \\ &= \sum_{\mathbf{s}} P(\mathbf{o}|\mathbf{s}) \cdot P(\mathbf{s}) \\ &= \sum_{\mathbf{s}} \prod_{k=1}^K P(\mathbf{o}_k | S_k) P(S_k | S_{k-1}) \end{aligned} \quad (6)$$

The transition distribution ($P(S_k | S_{k-1})$) is defined as

$$P(S_k = j_1 | S_{k-1} = j_2) = a_{j_1 j_2}, \quad 1 \leq j_1, j_2 \leq L \quad (7)$$

Thus the probability of the transition, $\mathbf{A} = \{a_{j_1 j_2}\}$, is an $L \times L$ state transition probability matrix. Note that at the starting point, the initial state distribution, $\Pi = \{\pi_j\}$ and $\pi_j = P(S_1 = j)$. The observation distribution ($P(\mathbf{o}_k | S_k)$) is characterized by a mixture of Gaussian density

$$\begin{aligned} P(\mathbf{o}_k | S_k = j) &= \sum_{h=1}^H c_{jh} N(\boldsymbol{\mu}_{jh}, \boldsymbol{\Sigma}_{jh}) \\ \text{and } \sum_{h=1}^H c_{jh} &= 1, \quad c_{jh} \geq 0, \quad j = 1, 2, \dots, L \end{aligned} \quad (8)$$

where H is the number of mixtures. c is the weight for each mixture component but it should sum up to 1 for each state. Each Gaussian probability density function is an elliptically symmetric density with the mean vector $\boldsymbol{\mu}_{jh}$ and the covariance matrix $\boldsymbol{\Sigma}_{jh}$. The observation probability parameters are $\mathbf{B} = \{c_{jh}, \boldsymbol{\mu}_{jh}, \boldsymbol{\Sigma}_{jh}\}$. For detail in mixture of Gaussian distribution, please refer to [6]. Thus, the HMM model parameters are $\lambda = [\mathbf{A} \ \mathbf{B} \ \Pi]$. The detail of the training algorithm can be found in [7].

Given the above optimization for determining the HMM parameters, $\lambda = [\mathbf{A} \ \mathbf{B} \ \Pi]$, it is necessary to return to the assumption that the window size (d) is known. To determine the window size, the HMM problem should be solved in Eq. (6) for several values of d , with the value of d giving the lowest dispersion of a stochastic process,

$$\min_d H(d) = -\min_d \int p(\mathbf{o}_k) \log(\mathbf{o}_k) d\mathbf{o}_k \quad (9)$$

being the best estimate of the window size. The above equation, so called entropy, is a unified probability measure of uncertainty quantification. The main advantage of entropy is that it provides a general description of stochastic system without constraints of certain distribution [8]. This approach, which is essentially a search in one direction, is required because the variable (d) is discrete, so it is not possible to determine the analytical derivation of $H(d)$ with respect to the window size. The technique is iterative in nature and the search procedure is terminated when minimal relative improvement ($(H(d) - H(d-1))/H(d-1)$) is less than a preset value.

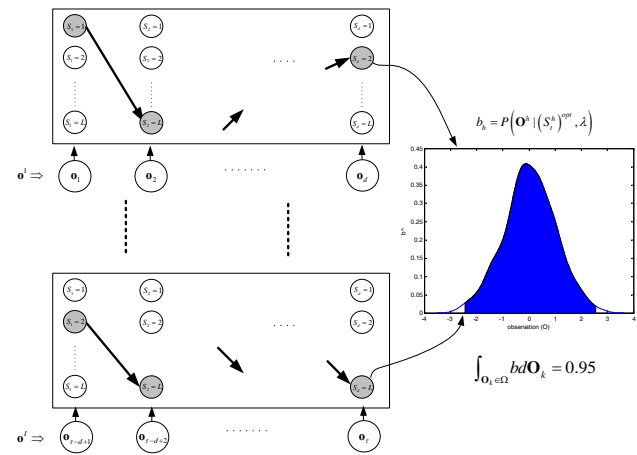


Fig. 3. The optimal state sequence in bold line in the trained model for each data set and the probability of the selected optimal state at the end time point of the moving window forms a distribution.

III. MONITORING CONTROL CHART BASED ON MPCA-HMM

Once the HMM model is trained, the most likely state sequence for an observed sequence could be found. For a given model and an observation sequence $\mathbf{O}^k = \{\mathbf{o}_{k-d} \ \dots \ \mathbf{o}_{k-1} \ \mathbf{o}_k\}$, the most likely underlying state sequence ($(s_1^{d+1})^* = \{s_1^* \ s_2^* \ \dots \ s_{d+1}^*\}$) can be found

$$(s_1^{d+1})^* = \{s_1^* \ s_2^* \ \dots \ s_{d+1}^*\} = \arg \max_{s_1^{d+1}} P(\mathbf{O}^k, s_1^{d+1} | \lambda) \quad (10)$$

A recursive Viterbi algorithm with dynamic programming can be used to find the most likely state sequence [9]. The same procedure is repeated for all the collected sets to compute the most likely state sequences for each set. Fig. 3 shows the single best state path for each set. Not all the sets follow the same best path. Different selected states for different sets account for the fact that the observed values of

the process variables do not perfectly conform to the deterministic model. This variability is resulted from the uncertain variations and disturbances among the lurking variables that affect the system.

For on-line monitoring of the observation images, the control limits should be built for normal operation. As the new image is taken, the image data from the beginning till the current time point (k) are available ($\mathbf{Y}(k)$)

$$\mathbf{Y}(k) = [\mathbf{y}_{k-d+1}^T \quad \mathbf{L} \quad \mathbf{y}_{k-1}^T \quad \mathbf{y}_k^T] \quad (11)$$

The score matrix by projecting the data ($\mathbf{Y}(k)$) onto the loading matrix (\mathbf{P}) can be computed:

$$\mathbf{T}(k) = \mathbf{Y}(k)\mathbf{P} \quad (12)$$

$$\mathbf{T}(k) = \begin{bmatrix} \mathbf{t}_{k-d}^T \\ \mathbf{M} \\ \mathbf{t}_{k-1}^T \\ \mathbf{t}_k^T \end{bmatrix} = \begin{bmatrix} t_{k-d,1} & t_{k-d,2} & \mathbf{L} & t_{k-d,R} \\ & & & \mathbf{M} \\ t_{k-1,1} & t_{k-1,2} & \mathbf{L} & t_{k-1,R} \\ t_{k,1} & t_{k,2} & \mathbf{L} & t_{k,R} \end{bmatrix} \quad (13)$$

The portion of the data ($\mathbf{Y}(k)$) corresponding to the $M - R$ smallest singular values forms the residual variation

$$\mathbf{v}(k) = \begin{bmatrix} v_{k-d} \\ \mathbf{M} \\ v_{k-1} \\ v_k \end{bmatrix} = \begin{bmatrix} \mathbf{e}_{k-d}^T \mathbf{e}_{k-d} \\ \mathbf{M} \\ \mathbf{e}_{k-1}^T \mathbf{e}_{k-1} \\ \mathbf{e}_k^T \mathbf{e}_k \end{bmatrix} \quad (14)$$

$$\begin{bmatrix} \mathbf{e}_{k-d}^T \\ \mathbf{M} \\ \mathbf{e}_{k-1}^T \\ \mathbf{e}_k^T \end{bmatrix} = (\mathbf{I} - \mathbf{P}\mathbf{P}^T)\mathbf{Y}(k) \quad (15)$$

Then concatenate the score vector \mathbf{t}_n^T and the variation v_n , $n = k-d, L, k-1, k$ sequentially to make a sequence pool $\mathbf{O}^k = \{\mathbf{o}_{k-d} \quad \mathbf{L} \quad \mathbf{o}_{k-1} \quad \mathbf{o}_k\}$. With the developed probability distribution that reflects the normal operation, the control limit for the selected state at each time point is required to detect any departure of the process from its standard behavior

$$\int_{P(\mathbf{o}_k | (s_k)^*, \lambda) > P^{th}} P(\mathbf{o}_k | (s_k)^*, \lambda) d\mathbf{o}_k \leq 0.95 \quad (16)$$

This integral is approximated by bootstrapping the original sequential data set to generate enough data and obtain truly quality-related variables from the sequential process data. For more information on the bootstrapping, please refer to work by [10]. The data corresponding to 95% of the confidence limit is taken to get a likelihood threshold (P^{th}).

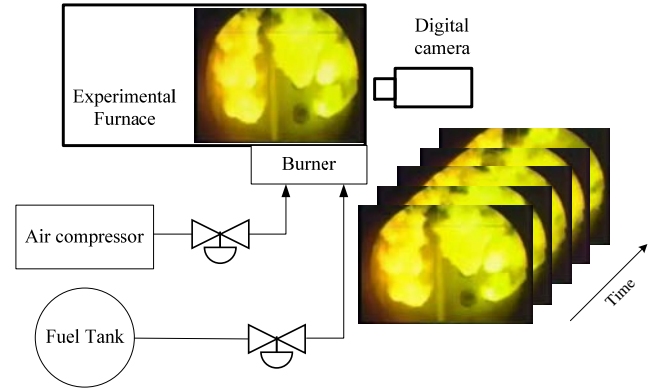


Fig. 4. The experimental combustion system.

IV. EXAMPLES

A schematic diagram of the experimental furnace with a flame-monitoring system is shown in Fig. 4. The burner of the experimental system uses industrial heavy oil as fuels. The air source of combustion comes directly from a direct-driven air compressor. A digital color camera is installed to capture the images of flames in the furnace. The camera with the specifications of 658×492 pixels and a resolution of 24 bits per pixel is capable of capturing flame images at 73 frames per second. The output signal of the camera is sent to a computer in IEEE-1394a interface and the imaging sample rate is set as one frame per four seconds considering the computer loading.

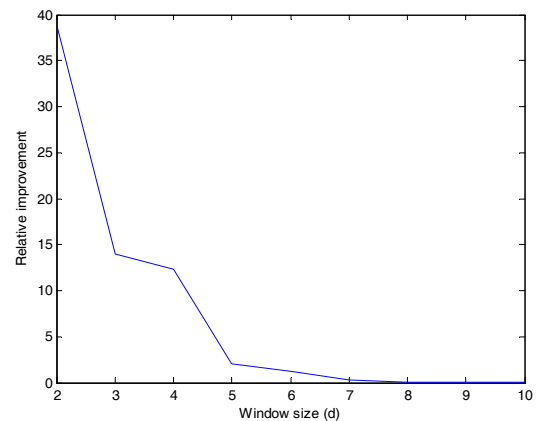


Fig. 5. Entropy of the observation probability

The proposed method is implemented real-time on a lap-top and tested for a large variety of conditions in comparison with the method utilizing only the color variation information. A total of 2000 image data which is not used as a testing set in this case study is collected to build the model when no deterministic disturbances are identified in the operating log. In MPCA-HMM, the time-sequential image data set is unfolded in a two-dimensional array (shown in Fig. 1). Then this matrix is decomposed by MPCA, and via cross-validation, it is found that 7 principal components are needed to describe the data set. This model captures 98.64%

of the variation in the process data set. Then, temporal sequential data are constructed using the extracted features from PCA model. During training HMM, the different window sizes of HMM are applied. Fig. 5 shows the entropy of the observation probability decreases with the increasing number of window sizes. Thus, the selected window size of HMM is 7. The control charts of the models for the normal operation are shown in Fig. 6, where the control limit of $\log P(\mathbf{o}_k | \lambda)$ is obtained from the proposed MPCA-HMM. The control limits representing 95% and 99% confidence limits are also shown in Fig. 6.

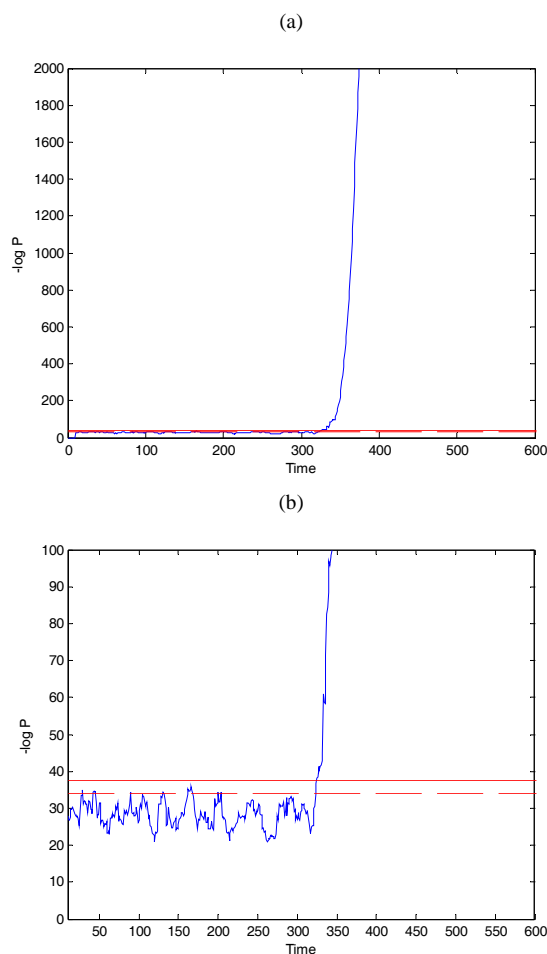


Fig. 6. (a) The testing data monitored by PCA-HMM control charts of the normal operating condition with 95% (dash line) and 99% (solid line) confidence limits. The first 300 points are normal and the rest (from 301 to 600), abnormal; (b) zooming in on the data around the control limits.

Two abnormal conditions, which are two different levels of feed air rates linearly decreasing during two different periods of time, are used for testing. For easy visualization and comparisons, time-sequential images in one normal condition and two abnormal ones are shown in Fig. 7. However, differences between the normal (Fig. 7(a)) and the minor abnormal (Fig. 7(b)) operating behaviors can not be distinguished from the image data directly. The spatiotemporal correlations in the image are not apparent because flame flicker frequency which is not constant incurs the variations in flame pixels. To save space limitation, only the minor fault condition is used for testing. The set of 600

observed images, including the first 300 images in the normal condition and the rest 300 in the abnormal one, is used for detection. For the abnormal condition, the monitoring outcomes of the proposed model are shown in Fig. 6. The initial time of the fault is induced at the 301th point. As shown in Fig. 6, even though the abnormal process with small disturbance exists, the fault can still be detected after MPCA-HMM is applied to characterizing the temporal dynamics between images. $P(\mathbf{o}_k | S_k^{opt})$ has been increased remarkably and it has fallen outside of the 95% confidence limit after the 320th time point. The small detection lag is due to the transport delay of the air flow.

If the complete training data structure is directly applied to HMM without variable reduction, the number of observed variables are more than 971208 at each time point. The HMM model with 3 hidden states would result in 5827257 parameters, which are a prohibitively large number. It is impossible to train the model. The problem is usually encountered when we have a moderately large number of observed variables. It can be alleviated by the proposed method. In this study, the number of parameters in MPCA-HMM is 51, which is significantly less than that in HMM only. The smaller the number of parameters is fitted, the faster the model training would be converged and the less risk of overfitting the training data would have. Thus, it explains why HMM is used to construct the temporal behavior of the variables after MPCA in the proposed method.

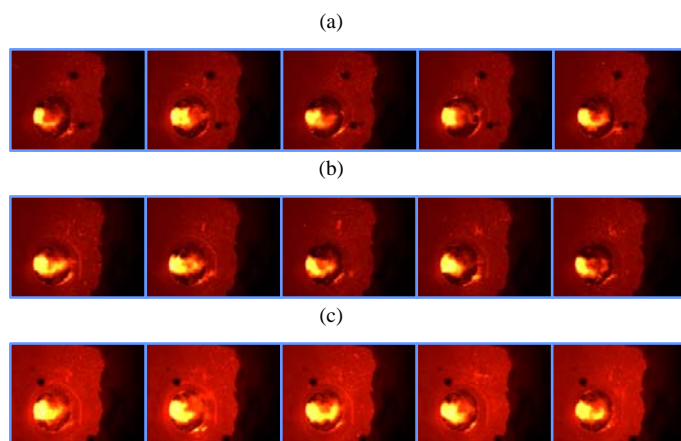


Fig. 7. Time-sequential images where the interval between the images is three minutes: (a) normal; (b) abnormal with the minor fault condition; (c) abnormal with the major fault condition

V. CONCLUSION

Abundant real-time digital image data can be gathered in the automatic system. Particularly, in this work, the digital RGB flame images collected from the combustor contain the complexity and uncertainty of the process behavior. The volume of image data being generated and processed keeps growing and there appears no end in sight to this trend, so mining and fusing a series of universal applicable quantities from the historical data to monitor the operation status is essential. This will help combustion control as well as reduce

fuel consumption. In this paper, HMM based monitoring is proposed. However, the use of HMM in on-line image analysis has up to now been limited to the area of the text and speech recognition. Since the sequences of images from combustors can be regarded as simply two-dimensional signals with the independent variables repeated in the time frame. In this research, the conventional multivariate MPCA are extended by incorporating the HMM structure to solving the detection problem for the two-dimensional repetition observation. In order to quantitatively discriminate the operation distribution, MPCA-HMM is applied at the modeling probability distribution stage. The spatiotemporal patterns can be trained. The corresponding control limits that evaluate the current process can be estimated. Even though the sequence image has a two-dimensional spatial relationship and a temporal behavior, the comparisons of quantitative and qualitative results of the proposed method are made through a real experimental problem for investigating the feasibility of the proposed MPCA-HMM method. The proposed method may be incorporated with different surveillance systems for image monitoring and early detection.

REFERENCES

- [1] Y. Yan, G. Lu, and M. Colechin, "Monitoring and characterisation of pulverized coal flames using digital imaging techniques," *Fuel*, 81, pp.647-656. 2002.
- [2] J.M. Beer. "Combustion technology developments in power generation in response to environmental challenges," *Prog. Energy Combust Sci.*, 26, pp. 301-327, 2000.
- [3] L.Gang, Y. Yong and C. Mike, "A digital imaging based multifunctional flame monitoring system," *IEEE Trans Instrum Meas.*, 53, pp.1152-1158. 2004.
- [4] H. Zhang, Z. Zou, J. Li and X. J. Chen, "Flame image recognition of alumina rotary kiln by artificial neural network and support vector machine methods," *Cent South Univ Technol.*, 15, pp.39-43, 2008.
- [5] L.R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, 76, pp.257-286, 1989
- [6] T. Mitchell, *Machine Learning*, McGraw Hill, New York, 1997.
- [7] J.M. Lee, S.J. Kim, Y. Hwang and C.S. Song, "Diagnosis of mechanical fault signal using continuous hidden Markov model," *J Sound Vibrat.*, 276, pp.1065-1080, 2004.
- [8] A. Papoulis, *Probability, random variables and stochastic processes*, 3rd ed., New York: McGraw-Hill, 1991.
- [9] M.D. Moore, "Most Likely state sequence speech reconstruction using a generalized Hidden semi-Markov model with two distinct regeneration times applied to English," Ph.D dissertation, Rensselaer Polytechnic Institute, 2004.
- [10] B. Efron and R. J. Tibshirani, *An introduction to the bootstrap*, Chapman & Hall, London, 1993.