# MULTIVARIATE IMAGE ANALYSIS FOR INFERENTIAL SENSING: A FRAMEWORK

**Honglu Yu , John F. MacGregor**

*McMaster Advanced Control Consortium*
*Department of Chemical Engineering*
*McMaster University, Hamilton, On. Canada*

Abstract: This paper presents a framework for developing vision-based inferential sensors. This framework not only gives a summary of existing methodologies, but also combines the methods used in other areas, such as traditional machine vision, multivariate image analysis and multivariate data analysis, and gives a broad vision for future developments of the area. 

Keywords: Inferential Sensor, Multivariate Image Analysis, Principal Component Analysis, Partial Least Squares, Machine Vision

## 1. INTRODUCTION

One of the most important elements in the successful monitoring and control of systems is fast, reliable and inexpensive on-line sensors. However, not all the important variables can be measured on-line. In this situation, a strategy often used is to estimate the unknown variable from other easily measured variables through a model, known as inferential control.

In this paper, attention is focused on a new generation of inferential sensing techniques, where process or product property variables difficult to measure are estimated using information extracted from online acquired images. This type of sensor can greatly improve the quality control performance in at least two areas: i) industries that produce solid products such as food, polymers, pulp and paper, where product properties can only be measured periodically in a laboratory; and ii) processes whose performance is related to their visual appearance, such as flame in a combustion process.

The petrochemical industry made rapid advances in multivariable model predictive control largely because they had the availability and abundance of inexpensive and informative sensors such as thermocouples, pressure transducers, flow meters, pH and ion-specific meters and gas chromatographs. This is a direct result of the fact that the major streams in petrochemical processes consist of well mixed gases and liquids which made the use of such sensors very easy. On the other hand, the solids processing industry has had much less success at implementing advanced control precisely because of the lack of such sensors. In industries that produce solid products, the product properties are generally measured periodically by manually collecting samples and then analyzing them in the laboratory. The analysis procedure may require that the samples be destroyed and the procedures are time consuming and manpower intensive. Therefore, it is usually impossible to obtain on-line quality measurements and even the simplest automatic control algorithm can hardly be implemented. In this situation, an inferential sensor based on an imaging system is very attractive. By taking images of the products, the samples remain untouched and the sampling rate can be very high. The product properties can then be predicted in real time from image data.

In fact, using imaging techniques for industrial production is not a new concept. Machine vision has been developed for more than 30 years. However, there are three major differences between conventional machine vision and vision-based inferential sensor. First, the objectives and the types of the problems dealing with are different. A large amount of the problems in machine vision area involve locating and identifying objects, measuring the size of products, sorting different types of products and detecting defects. These imaging systems are designed as a replacement of human vision and decision system. On the other hand, the objective of an inferential sensor is to quantitatively predict uneasily measured property variables for online monitoring and feedback control, which is actually beyond the capability of human vision system. Second, most images used in machine vision are greyscale or binary images while most images used for vision-based inferential sensors are multivariate images, such as multispectral images and color images. This is because multivariate images, compared with greyscale or binary images contain much more information. Third, since the problem type and image type are different, so do the methodologies developed. In machine vision area, most methods are used for greyscale images and to compute geometric characteristic of the objects. In the case of building inferential sensors, one is interested in extracting the subtle information (may or may not be seen) from multivariate images and relating the extracted information with some chemical or physical properties.

Research on multivariate image analysis started in 1970s mainly in remote sensing area. Afterwards, it has been developed in many other scientific areas, such as medical imaging in medical science and microscope multispectral imaging in chemistry. Most of these studies focus on how to efficiently display a multivariate image (e.g. visualisation of a hyperspectral image with 250 bands), on how to segment an image into different areas (e.g. segmentation of a satellite map with different features: road, river, land and grass) and on how to detect certain features (e.g. detection of cancer cells in an image obtained from Magnetic Resonance Imaging (MRI)). In other words, the purpose of these studies is to develop a human assistance system rather than an automatic system because the output is generally an image with enhanced appearance.

Recently, more and more attention has been paid to the development of vision-based inferential sensor. A series of papers have been published to solve several different problems, including seasoning level prediction on snack food (Yu and MacGregor, 2003a), pulp properties estimation using NIR imaging (Bharati et al., 2002) and environmental monitoring system for combustion process using flame color images (Yu and MacGregor, 2003b).

In this paper, a general framework for developing vision-based inferential sensor is presented. This framework not only gives a summary of the approaches published in literature, but also gives other possible options by combining the methods developed in traditional machine vision and multivariate image analysis areas. This framework can be used as a guideline for future applications.

## 2. MULTIVARIATE IMAGES

A multivariate image is a set of congruent images. The definition for congruence in imaging is given by Geladi and Grahn (1996): two or more images are congruent if they can be stacked so that for each pixel in one image there is a corresponding pixel in the other image(s) that can be referred to the same position in the object or scene depicted.

From the data storage point of view, a multivariate image is a three-way matrix. Two of the ways in this three-way matrix are the geometrical image coordinates which describe the image scene plane and usually treated as a pair. The third way is the 'variable' way. Hence one could also view a multivariate image as a two-way array of pixel intensity vectors. Multivariate image data often contain highly correlated information among the variables because the individual images within one multivariate image are congruent.

A multivariate image can be obtained from varied sources. A common way is to stack the images of an object at different radiation energies or wavelengths. This is a multispectral image. Magnetic Resonance Imaging (MRI) is another imaging technique to obtain multivariate images. By varying the parameters (most common parameters are the spin-echo time and the relaxation delay time), a set of images can be obtained. A multivariate image can also be constructed by combining images obtained from different instruments. A multi-temporal image is a type of multivariate image as well, which consists of the images of the same scene but taken at different times. A color image can be also considered as a 3-variable multivariate image because a color image consists of 3 color channels: red, green and blue. Another technique of creating multivariate images is to combine a grayscale image with the copies of it that are derived from different spatial filtering operations.

## 3. VISION-BASED INFERENTIAL SENSOR

A vision-based inferential sensor consists of three basic elements: an imaging system (e.g. camera), a computer and a model which predicts the property variables using information from the images. In this paper, we mainly consider the model building problem.

Though technically it is possible to build a vision-based inferential sensor based on fundamental model, most sensors are data-based. Like building any data-based model, the first thing is to collect a good training dataset, since the quality of the model is largely dependent on the quality of the training dataset. This training dataset should cover a large range of the operation range and the imaging condition should be kept as constant as possible. Both the images and the corresponding product properties need to be collected. If the dataset is collected in an online production mode, it is very crucial to ensure they have consistent time stamps in both images and property measurements.

Obviously, no regression methods can directly relate a set of three-way image matrices to a set of property variables. In most applications, a two-step procedure is adopted. In this two-step procedure, first a feature vector is extracted from each image, and then the extracted feature vectors are used as predictor to be regressed against properties variables. The feature extraction step is the key to build a successful model, not only because it largely reduces the problem dimension, but also, more importantly, it concentrates the related information. There are many different ways to extract features. In the framework proposed in the next section, we summarized the feature extraction methods into four categories. However, the choice of features is very case dependent and in many situations an iteration procedure is needed.

To predict the property variable for a new image, one needs to first compute its feature vectors and then calculate the estimation through the regressed model. Figure 1 shows a basic scheme for model building and for new image prediction.

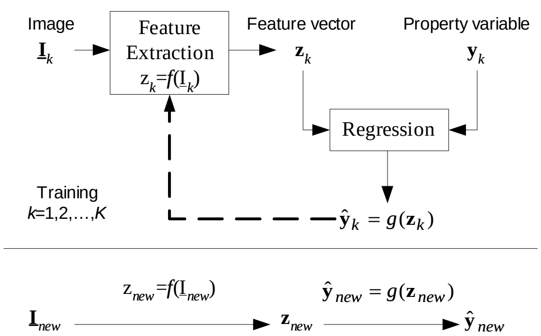


Figure 1: A scheme for model building and prediction

## 4. A FRAMEWORK FOR DEVELOPING VISION-BASED INFERENTIAL SENSOR

In this section, a framework is proposed to give a guideline for building an inferential sensor using multivariate images. This framework is originally a summary of several applications. However, it has been extended to give more options by combining many other approaches used in the traditional machine vision, multivariate image analysis and multivariate data analysis.

Figure 2 shows the proposed framework. In this framework, there are three steps: preprocessing, feature extraction and regression.

### *4.1 Preprocessing*

Though other preprocessing techniques, such as removing noise and improving contrast may be used, here the preprocessing step mainly refers to variable dimension reduction.

Variable dimension reduction could be a significant preprocessing procedure in dealing with multivariate images, especially for images with high variable dimension. Because the individual bands of a multivariate image are often highly correlated, the proper variable dimension reduction can not only remove or reduce this variable redundancy and increase the computation efficiency, but can also often result in a better signal to noise ratio and concentrate on the useful information.

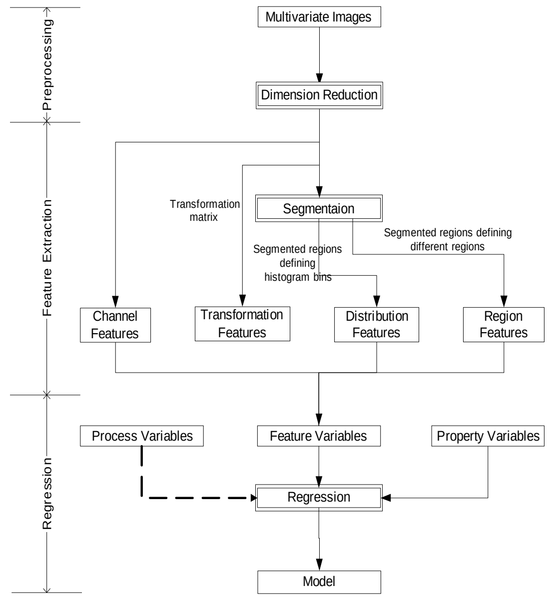


Figure 2 Proposed framework for predicting process/product properties using multivariate images

In both the snack food application (Yu and MacGregor , 2003a) and the flame application (Yu and MacGregor, 2003b), Principal Component Analysis (PCA) is used to reduce variable dimensions.

To perform PCA on an *M*-band multivariate image (often referred to as multi-way PCA), one first unfolds the image matrix $\underline{\mathbf{I}}$, where each pixel is considered as an observation. Normal PCA can be then performed on the unfolded matrix **I**.

$$\mathbf{I} = \sum_{a=1}^{A} \mathbf{t}_a \mathbf{p}_a{}^T + \mathbf{E}$$

where $A$ is the number of principal components, the $\mathbf{t}_a$'s are score vectors and the corresponding $\mathbf{p}_a$'s are loading vectors and **E** is the residual.

Score $\mathbf{t}_a$'s can be refolded back into the original image size and form a new multivariate image with dimension $A$ (the number of principal components), which could be much smaller than $M$ (the original dimension). This new lower-dimensional multivariate image is sent to the feature extraction steps.

There are many other options for variable dimension reduction. These optional methods have been used in the traditional multivariate image analysis and data analysis.

PCA belongs to a family of linear projection approaches. In these linear projection approaches, a low dimensional set of projection directions are defined by using certain criteria and the multivariate image data is then projected to this low dimensional subspace. The differences among the methods are different criteria. PCA, finds the subspace that explains the largest percentage of the variation, Discriminant Analysis (DA, including Fisher's Discriminant Analysis, Canonical Component Transform and PLS Discriminant Analysis), finds the subspace that has the largest discriminating power (Maxwell, 1976; Lied and Esbensen, 2001), Independent Component Analysis (ICA) finds the subspace from which the projections of the data onto each of the basis vectors are independent (Kaarna et al., 2000), Projection Pursuit (PP) finds the subspace that captures the non-Gaussian distributed information (Ifarraguerri and Chang, 2000) and Decision Boundary Feature Extraction (DBFE) finds the subspace defined by the decision boundary (Lee and Landgrebe, 1993). Among these projection methods, PCA is the simplest one in computation and the one that has been most widely used.

Other than linear projection methods, several non-linear transformation methods, such as Self-Organizing Map (Kohonen, 1984) and Sammon's mapping (Sammon, 1969) can also be considered. However, the main drawback of these nonlinear mapping methods is that they often require long computation time (Bonnet et al., 1997).

### *4.2 Feature extraction*

Feature extraction methods are summarized into four categories: channel features, transformation features, distribution features and region features. These features can be used independently or used in a combination manner.

<u>Channel features</u>

Channel features can be statistical measurements for each individual channel image, such as average values, standard deviations, moments and histogram. One example is found in Yu and MacGregor (2003a) where average colors of the snack food images were used as feature variables to predict seasoning level on the snack food product. Another example is given in Hätönen et al. (1999) where mean, standard deviation, skewness and kurtosis values of different RGB channels were used as feature variables to predict mineral concentrations in the flotation froth.

When computing channel features, each channel image is processed one at a time as a grayscale image. Therefore, many image processing methods for grayscale images can be used to extract certain information, such as the use of wavelet transforms on each channel image to extract texture information.

<u>Transformation features</u>

Transformation features refer to the transformation matrix obtained in variable dimension reduction step, such as the loading vectors $\mathbf{p}_a$'s when use PCA. In Bharati et al. (2002), where several properties of pulp were predicted using NIR multispectral images, the first principal component loading vectors were used as feature vectors. It has been pointed out in Yu and MacGregor (2003a) that if one performs PCA on an image data without mean-centering the first PCA loading vector is similar to the normalized average spectral (variable) response.

<u>Distribution features</u>

Pixel intensity histogram is an important tool and has been used extensively in grayscale image analysis. If in the dimension reduction step the variable dimension has been reduced to one, a histogram can be easily calculated and used as features.

However, if reduced variable dimension still larger than one, distribution of the spectral (variable) response needs to be described using a multi-dimensional histogram. Evidently, to obtain such a multi-dimensional histogram could be a very computation intensive work since the amount of calculation increases in an exponential manner as the number of dimensions increases. Furthermore, even if one does obtain such a high dimensional histogram as feature variables, he could encounter problems to obtain a good model in the regression step. In one of the feature extraction methods presented in Yu and MacGregor (2003a), all the snack food color images are first projected to a two-dimensional $\mathbf{t}_1$-$\mathbf{t}_2$ subspace and then an unfolded 32×32 histogram was used as feature variables to represent each image. The results showed that the model obtained using these feature variables had good prediction ability for the images having the same size with the training images but could have large errors when predicting from images with smaller size..

Yu and MacGregor (2003a) proposed another idea to further obtain a one-dimensional histogram. In their approach, $\mathbf{t}_1$-$\mathbf{t}_2$ score space was divided into several bins. Each bin was defined in a way that pixels falling into each bin were assumed having similar seasoning level. The counts of the pixels falling into each bin then formed a one-dimensional histogram.

Follow up this idea; one can obtain a one-dimensional histogram for a multivariate image by first segmenting the image into different parts and then counting the number of pixels falling into each part. Pixels falling into each part should have somehow common properties. In this way, the computation difficulties caused by high dimension are bypassed. However, how to segment the images becomes a new problem. More multivariate image segmentation approaches will be discussed in the 'region features' section.

The histogram is not the only distribution descriptor. It has been shown in Yu and MacGregor (2003a) that using cumulative histogram instead of histogram often can have better signal-to-noise ratio and result in more reasonable parameter estimates.

Region features

Sometimes, the imaging scene consists of several distinguished regions with different characteristics. In this situation, a natural thinking is to segment these regions first and then compute the features for each region, which can be channel features, transformation features and/or distribution features.

One example is the feature extraction method developed in Yu and MacGregor (2003b) for the flame application. In this application the objective is to predict several process properties, including heat of combustion of the waste feed and concentration of $NO_x$ and $SO_2$ in the off-gas, using RGB flame images. To extract feature information from rapidly changing flame images, a flame luminous region is first separated from non-luminous region and nine feature variables are then computed for these two regions respectively. The separation of the flame luminous region and the non-luminous region was done by first projecting pixels onto the $\mathbf{t}_1$-$\mathbf{t}_2$ PCA score space and then defining a polygon mask in score space that indicates the flame luminous region. The boundary of this mask was obtained by a trial and error process, whereby one selects a mask area in the score plot, selects the pixels lying under it and highlights them in the image space, and iterates until one obtains a mask that segments the feature of interest. This segmentation approach was first introduced in 1989 by Esbensen and Geladi, and has become a very effective method. In Yu and MacGregor (2003a), a product mask was defined to separate snack food product pixels and conveyor belt pixels. However, in their work, the mask is not defined by a full manual trial-and-error procedure, but based on the computation of some covariance property of the score combinations.

As mentioned in the introduction section, segmentation is one of the major topics in traditional multivariate image analysis studies. Therefore, many other multivariate image segmentation techniques have been presented.

Basically, the segmentation approaches for multivariate images can be divided into two main categories. In the first category, only spectral (variable) information of each pixel in the image plane is considered. In another words, pixels are treated independently of one another. Therefore, a pixel can be treated as an observation in multivariate data analysis. The classical classification methods, such as K-means (Duda and Hart, 1973), fuzzy C-means (Boudraa et al., 2000), neural networks (Reddick et al. 1997), maximum likelihood (Liang et al. 1994) and discriminant analysis, can be directly used. Under the same category, segmentation can also be obtained by thresholding the pixel intensity values based on histogram information. This has been extensively used in the segmentation of grayscale images. For multivariate images, one option is to find a proper thresholding value for each individual band (variable) image and combine the results (Raya, 1990). However, in this way, the nature of the multivariate image has been ignored. Another option is to find a thresholding range in the $M$-dimensional space for an $M$-band image. For $M$=2, the thresholding range can be defined by the methods mentioned above (choosing a polygon mask). However, the computation of the $M$-dimensional histogram and choosing a proper range of the thresholding area will be very difficult for the case where there are more than two dimensions ($M$>2). A compromise solution is to choose polygon masks for each pair of the variable combinations.

In the second category, spatial information as well as spectral (variable) information is used for classification. This is also known as multivariate image texture analysis. There are three general approaches to segment an image based on textural information. In the first approach, spatial and spectral information are considered simultaneously. The approaches presented are often extended from grayscale image texture analysis. Several segmentation methods based on Markov Random Fields (MRF) models and maximizing a posteriori distribution probability (MAP), which utilize both spectral and spatial information to model the local correlation structure of an image, have been presented (Kartikeyan et al., 2002). Kovalev et al. (2001) proposed a texture analysis method for multivariate image that is based on extended co-occurrence matrices. Extraction and classification of homogeneous objects (ECHO) (Kettig and Landgrebe, 1976), is another texture segmentation method which is based on the spatial and spectral homogeneity of the regions. The listed methods above are only a few examples. However, methods considering both spatial and spectral information at the same time generally involve iterative computation and/or optimization that require intensive computation time. In the second approach, one or several representative images are generated by variable reduction or some other operation (e.g. using ratios). Then the texture methods for grayscale images can be applied on them individually. In the third approach, the images of one or several channels in a multivariate image are filtered or transformed. The newly filtered or transformed images are expected to contain the spatial information. A new multivariate image is then constructed by stacking the new images (and the original image) together. The segmentation methods in the first catalogue can then be used for classification. Liu and MacGregor (2002) studied both of the latter two approaches by using wavelets to extract the spatial information.

### *4.3 Regression*

After extracting feature variables, the regression step is performed as in normal data analysis to build a model. Technically, any data based regression methods can be used in this step. The methods can be linear regression approaches, such as Multivariate Linear Regression (MLR), Partial Least Squares (PLS), Ridge Regression (RR). PLS is a method which can well handle highly correlated feature variables and has shown to be a simple and effective approach. Although PLS is a linear regression methods, transformation or non-linear terms can be used to handle the possible non-linearity. For example, in Bharati et al. (2002), one of the pulp properties, DCM Resin, was transformed into ln (DCM Resin). The regression methods can be also nonlinear regression approaches, such as neural networks. In Wang et al. (2002), a neural network was used to predict the $NO_x$ emissive concentration of a boiler using features extracted from color flame images.

In some situation, besides the feature variables extracted from images, other process measurements can also be used to build the model. In Yu and MacGregor (2003b), a PLS model was built to predict the product of heat of combustion of waste stream and the fuel flow rate. The heat of combustion was then obtained by dividing the PLS model prediction by the liquid fuel flow rate. Other than the feature variables, the flow rates of liquid fuel and natural gas were also used as the predictors. This is shown in equation:

$$\hat{Q} = \left[F_{lf} \;\; F_{ng} \;\; \mathbf{v}\right] \cdot \hat{\boldsymbol{\gamma}}, \text{ where } Q = H_{Lf} \cdot F_{Lf}$$

$$\hat{H}_{lf} = \hat{Q} / F_{lf}$$

where $\mathbf{v}$ is the feature vector extracted from the image data, $F_{ng}$ is the natural gas flow rate, $F_{lf}$ is the flow rate of liquid fuel, $\hat{\boldsymbol{\gamma}}$ is the model regression coefficient vector and $H_{Lf}$ is the heat of combustion.

## 5. SUMMARY AND CONCLUSIONS

A framework is proposed for building inferential sensors using multivariate image analysis for process monitoring and control. This framework not only gives a summary of existing methodologies, but also combines the methods used in other areas, such as traditional machine vision, multivariate image analysis and multivariate data analysis, and therefore gives a broad vision for the future development.

## REFERENCES


Bharati, M. H., J. F. MacGregor, Marc Champagne and Marc Barrette (2002), Using NIR Multivariate Image Regression Techniques to Predict Pulp Properties, Control Systems 2002, Stockholm, Sweden, June 3-5

Bonnet, N., M. Herbin, and P. Vautrot (1997), “Multivariate Image Analysis and Segmentation in Microanalysis”, *Scanning Microscopy*, **11**, 1

Boudraa, A., S. M. R. Dehak, Y. Zhu, C. Pachai, Y. Bao, and J. Grimaud (2000), “Automated Segmentation of Multiple Sclerosis Lesions in Multispectral MR Imaging Using Fuzzy Clustering,” *Computers in Biology and Medicine*, **30**, 23

Duda, R. D., and P. E. Hart (1973), “Pattern Classification and Scene Analysis,” New York, John Wiley & Sons

Esbensen, K., and P. Geladi (1989), “Strategy of Multivariate Image Analysis (MIA),” *Chemometrics and Intelligent Laboratory Systems*, **7**, 67

Geladi, P., and H. Grahn (1996), Multivariate Image Analysis, John Wiley & Sons, Chichester, UK

Hätönen, J., H. Hyötyniemi, J. Miettunen, and L. E. Carlsson (1999), “Using Image Information and PLS for Predicting Mineral Concentrations in the Flotation Froth,” Proceedings of he Second International Conference on Intelligent Processing and Manufacturing of Materials (IPMM'99), July 10–15, Hawaii, USA

Ifarraguerri, A. and C. Chang (2000), “Unsupervised Hyperspectral Image Analysis with Projection Pursuit”, *IEEE Trans. on Geoscience and Remote Sensing*, **38**, 2529

Kaarna, A., P. Zemcik, H. Kälviäinen, and J. Parkkinen (2000), “Compression of Multispectral Remote Sensing Images Using Clustering and Spectral Reduction,” *IEEE Transactions on Geoscience and Remote Sensing*, **38**, 1073

Kartikeyan, B., K. L. Majumder, D. K. Pal, A. Sarkar, M. K. Biswas, and V. Kumar (2002), “A MRF Model-based Segmentation Approach to Classification for Multispectral Imagery,” *IEEE Transactions on Geoscience and Remote Sensing*, **40**, 1102

Kettig, R. L. and D. A. Landgrebe (1976), “Classification of Multispectral Image Data by Extraction and Classification of Homogeneous Objects,” *IEEE Transactions on GeoScience Electronics*, **GE4**, 19

Kohonen, T. (1984), Self-organization and Associative Memory, Springer, Berlin

Kovalev, V. A., F. Kruggel, H. Gertz, and D. Y. Von Gramon (2001), “Three-Dimensional Texture Analysis of MRI Brain Datasets,” *IEEE Transactions on Medical Imaging*, **20**, 424

Lee, C. and D. A. Landgrebe (1993), “Decision Boundary Feature Extraction for Nonparametric Classification,” *IEEE Transactions on Systems, Man, and Cybernetics*, **23**, 433

Liang, Z., J. R. MacFall, and D. P. Harrington (1994), “Parameter Estimation and Tissue Segmentation from Multispectral MR Images,” *IEEE Transactions on Medical Imaging*, **13**, 441

Lied, T. T. and K. H. Esbensen (2001), “Principles of MIR, Multivariate Image Regression I: Regression Typology and Representative Application Studies,” *Chemom. Intell. Lab. Syst.*, **58**, 213

Liu, J. and J. F. MacGregor (2002), “Multiresolutional Multivariate Image Analysis and its Application to Color Texture Analysis,” *submitted to* Pattern Recognition

Maxwell, E. L. (1976), “Multivariate Systems Analysis of Multispectral Imagery”, *Photogrammetric Engineering and Remote Sensing*, **42**, 189

Raya, S. P. (1990),“Low-Level Segmentation of 3-D magnetic Resonance Brain Images—A Rule-Based System,” IEEE Trans. on Medical Imaging, **9**, 327

Reddick, W. E., J. O. Glass, E. N. Cook, T. D. Elkin and R. J. Deaton (1997), “Automated Segmentation and Classification of Multispectral Magnetic Resonance Images of Brain Using Artificial Neural Networks,” *IEEE Transactions on Medical Imaging*, **16**, 911

Sammon, J. W. (1969), “A Nonlinear Mapping for Data Structure Analysis”, *IEEE Trans. Comput*. **C18**, 401

Wang, F., X. J. Wang, Z. Y. Ma, J. H. Yan, Y. Chi, C. Y. Wei, M. J. Ni, and K. F. Cen (2002), “The Research on the Estimation for the NOx Emissive Concentration of the Pulverized Coal Boiler by the Flame Image Processing Technique,” *Fuel*, **81**, 2113

Yu, H. and J. F. MacGregor (2003a). Multivariate Image Analysis and Regression for Prediction of Coating Content and Distribution in the Production of Snack Foods. *Chemom. Intel. Lab. Syst*. **67**, 125

Yu, H. and J. F. MacGregor (2003b). Monitoring Turbulent Nonpremixed Flames in an Industrial Boiler Using Multivariate Image Analysis (MIA). to appear in *AIChE J.*