Chapter 2

FRAR Based CBIR
CHAPTER 2

FRAR MODEL FOR CBIR

2.1. Introduction

Generally, image processing and image analyses are performed under one of the following techniques: (i) Transform based and (ii) Model based. The transform based techniques are widely used in image processing such as image coding and restoration. Various transform based techniques and their efficiencies are reported in various literature. Among the existing transform based techniques, the most widely used methods are Discrete Fourier Transform (DFT), Discrete Cosine Transform (DCT), Discrete Sine Transform (DST), Wavelets Transform, Karhunen-Loeve Transform (KLT), Walsh-Hadamard Transform, Harr Transform and Orthogonal Polynomials Transform. Each transform based technique has its own specific features. Recently, the Wavelets transform based image coding technique has become popular due to its efficiency. Most of the transform based techniques demand more computations and some of them require large amount of memory. Compared to model based techniques, the applications of the transform based techniques are limited to low-level image processing and analysis.

Model based techniques are appropriate to handle the problems effectively which involves large amount of data such as image filtering,
object recognition, etc. To handle this huge volume of data, it would be preferable to have an underlying model that explains the dominant statistical characteristics of the given data. The different classes of models that are suggested in literature exploit the statistical properties among the neighbouring pixels for low-level image processing and analysis. The statistical models attracted many researchers, because of its wide range of applications at low-level image processing that include texture analysis, image smoothing, enhancement, restoration, segmentation, edge detection and image data compression.

In General, the model based techniques are not usually made explicit, but are made implicit by the adoption of assumptions that incorporate certain model assumptions within them. Most of the algorithms use the assumption that an image can be processed randomly with wide sense of stationary properties, linear dependency, uncorrelated white noise, etc. In that sense, Markov Random Field (MRF) and Auto Regressive (AR) models are most appropriate to almost all the low-level image processing. Many researchers have explored the efficiency of the MRF and AR models for low-level image processing. The pixels in a two-dimensional image are spatially equal interval of distance in row wise and column wise as in time series (equal interval of time) and the pixels in the images satisfy the sampling properties and satisfy stationarity, linear dependency and white noise uncorrelation. As discussed in Chapter 1, different types of stochastic models that are used for image processing and analysis include, Auto Regressive (AR), Moving Average (MA), Auto Regressive Moving Average (ARMA) and Auto Regressive Integrated
Moving Average (ARIMA) with various assumptions about images. The assumptions, Random Field, Markov Random Field, Gibbs Field and σ-Field etc, are made on the basis of the nature of images. A brief review of the related literature, under the different approaches, is given below.

The main advantage of the AR model over the other models is that it is regenerative. Also it represents all the information in an image of size \((N \times N)\) with two sets of parameters, one set containing a minimum number of parameters having most of the information while another set containing \(N^2\) parameters, the so-called residuals, having the remaining information. The residuals can be stored with minimum number of bits than the original image pixels without sacrificing any accuracy [Kash80]. With the use of stored parameters and the residual values, the original image can be reconstructed with good quality. The textured images can be generated only with the use of stored parameters of the model having no compromise with the quality of the image.

In order to analyse the content of an image, many researchers have investigated various statistical models such as co-occurrence matrix method, MRF, AR models, Principal Component Analysis (PCA), etc. It is observed that the MRF and AR models play a significant role in all the low-level image processing. Moreover, in these models the degree of accuracy of parameter estimation plays a significant role in obtaining satisfactory results in image processing. The parameter estimation techniques are generally grouped into three classes, viz., sampling theory (or classical approach), Bayesian approach and iterative procedure. The classical methods depend on the degree of accuracy starting with the
most precise method and continuing in the order of precision are Maximum Likelihood method, Least Square method, Appropriate Least Square method and Yule-Walker method.

With the consideration of all the points aforesaid, a statistical model called a new Family of Full Range Auto Regressive model is developed and proposed in this thesis. The main advantage of the proposed model is that it works on both textured and untextured images, unlike the existing models. The common framework is based on a new family of Full Range Auto Regressive (FRAR) model that captures the low-level features viz., texture, edge, shape and color and used to retrieve the target image and its relevant images from the image database.

2.2. Regression model

In statistical techniques, the Regression model is used to predict one variable from one or more other variables. Regression models provide the scientist with a powerful tool, allowing predictions about past, present, or future events to be made with information about past or present events. In order to construct a regression model, both the information which is used to make the prediction and the information which is to be predicted must be obtained from sample objects or individuals. The relationship between the two pieces of information is then modeled with a linear transformation. In future, the first information is necessary and the regression model is used to transform this information into the predicted. In other words, it is necessary to have information on both variables before the model can be constructed. The
goal of regression procedure is to create a model where the predicted and observed values of the variable are similar.

2.3. Auto Regressive model

The auto regressive model is one of a group of linear prediction formulas that attempt to predict an output $y[n]$ of a system based on the previous outputs $(y[n-1], y[n-2], ...)$ and inputs $(x[n], x[n-1], x[n-2], ...)$. Deriving the linear prediction model involves determining the coefficients $a_1, a_2, ...$ and $b_0, b_1, b_2, ...$ in the equation:

$$y[n](estimated) = a_1 \cdot y[n-1] + a_2 \cdot y[n-2] + ... + b_0 \cdot x[n] + b_1 \cdot x[n-1] + ...$$

... (2.1)

A model which depends only on the previous outputs of the system is called an Auto Regressive model (AR), while a model which depends only on the inputs to the system is called a Moving Average model (MA) and of course a model based on both inputs and outputs is an Auto Regressive Moving Average Model (ARMA). Note that by definition, the AR model has only poles while the MA model has only zeros. Several methods and algorithms exist for calculating the coefficients of the AR model. An auto regressive model is defined as below:

$$x_t = \sum_{i=1}^{N} a_i x_{t-i} + \varepsilon_t$$

... (2.2)

where $a_i$ are the auto-regressive coefficients, $x_t$ is the series under investigation, and $N$ is the order (length) of the filter which is generally very much less than the length of the series. The noise term or residue, epsilon in the above equation is assumed to be Gaussian noise.
Verbally, the current term of the series can be estimated by a linear weighted sum of previous terms in the series and the weights are referred as auto regressive coefficients. The problem in AR analysis is to derive the "best" values for $a_i$ given a series $x_t$. The majority of methods assume that the series $x_t$ is linear and stationary. By convention the series $x_t$ is assumed to be zero mean, if not this is simply another term $a_0$ in front of the summation in the equation (2.2).

A number of techniques exist for computing AR coefficients. The main two categories are Least Squares method and Burg method. The most common least squares method is based upon the Yule-Walker equations. In this method the coefficients are derived by multiplying the above definition by $x_{t-d}$, by considering the expectation values and normalizing gives a set of linear equations called the Yule-Walker equations that can be written in matrix form as

$$
\begin{pmatrix}
1 & r_1 & r_2 & r_3 & r_4 & \ldots & r_{N-1}
\end{pmatrix}
\begin{pmatrix}
a_1 \\
a_2 \\
a_3 \\
\vdots \\
a_N
\end{pmatrix}
= 
\begin{pmatrix}
r_1 \\
r_2 \\
r_3 \\
\vdots \\
r_N
\end{pmatrix}
$$

where $r_d$ is the auto correlation coefficient at delay $d$.

Ordinary Regressive analysis has strict assumptions including normality, independence, homogeneity, etc and these assumptions have to be met. However, when using time series data in regressive, one must always check to make sure that all the assumptions of the classical linear regressive model are met. Auto correlation is a serial correlation, which is the correlation of a series of data with its own lagged values. It is
a violation of independent assumptions that commonly occurs when data are taken over time. This tells that observations are not independent. Accordingly, the residuals from Ordinary Least Square regressive model won’t be independent either, as shown above. This dependency is a violation of the NII (Normal, Independent, Identically distributed) assumptions about residuals required by the Ordinary Least Square regressive model. It is important to identify the presence of auto correlation in the data and to appropriately account for it in the modeling. Different definitions of auto correlation depend on the field of study which is being considered. In some fields, the term is used interchangeably with auto-covariance.

2.3.1. Auto correlation function

Self-similarity describes the phenomenon where certain properties are preserved irrespective of scaling in space or time. Network Traffic process is a second-order self-similarity and it describes the property that the correlation structure of a time-series is preserved irrespective of time aggregation. This correlation is captured by the auto correlation function (ACF), which measures the similarity between a series $X_t$, and a shifted version of itself, $X_{t+k}$. Simply applying the auto correlation function of a second-order self-similar time-series is the same across multiple aggregation levels.

In statistics, the auto correlation function (ACF) of a random process describes the correlation between the processes at different points in time. Let $X_t$ be the value of the process at time $t$ (where $t$ may be
an integer for a discrete-time process or a real number for a continuous-time process). If $X_t$ has mean $\mu$ and variance $\sigma^2$ then the definition of the ACF is

$$R(t, s) = \frac{\mathbb{E}[(X_t - \mu)(X_s - \mu)]}{\sigma^2} \quad \ldots (2.4)$$

The expression is not well-defined for all time series or processes, since the variance $\sigma^2$ may be zero (for a constant process) or infinite. If the function $R$ is well-defined its value must lie in the range $[-1, 1]$, with 1 indicating perfect correlation and -1 indicating perfect anti-correlation.

If $X_t$ is second-order stationary then the ACF depends only on the difference between $t$ and $s$ and can be expressed as a function of a single variable. This gives

$$R(k) = \frac{\mathbb{E}[(X_t - \mu)(X_{t+k} - \mu)]}{\sigma^2} \quad \ldots (2.5)$$

where $k$ is the lag, $|t - s|$. It is common practice in many disciplines to drop the normalization by $\sigma^2$ and use the term auto correlation interchangeably with auto-covariance.

For a discrete time series of length $n \{X_1, X_2, \ldots, X_n\}$ with known mean and variance, an estimate of the auto correlation may be obtained as

$$\hat{R}(k) = \frac{1}{(n-k)\sigma^2} \sum_{t=1}^{n-k} [X_t - \mu][X_{t+k} - \mu] \quad \ldots (2.6)$$

for any positive integer $k < n$. When the true mean $\mu$ is known, this estimate is unbiased. However, if the true mean and variance of the process are not known, and $\mu$ and $\sigma^2$ are replaced by the standard
formulae for sample mean and sample variance, then this estimate is biased. An alternative way, a periodogram based estimate replaces $n-k$ in the above formula with $n$. This estimate is always biased; however, it usually has a smaller mean square error. In image processing, the above definition is often used without the normalization, that is, without subtracting the mean and dividing by the variance. When the auto correlation function is normalized by mean and variance, it is sometimes referred to as the auto correlation coefficient.

Multi dimensional auto correlation is defined similarly. For example, in three dimensions the auto correlation of a square-summable discrete signal would be

$$R(j,k,l) = \sum_{n,q,r}(x_{n,q,r})(x_{n-j,q-k,r-l})$$  \hspace{1cm} \text{(2.7)}

When mean values are subtracted from signals before computing an autocorrelation function, the resulting function is usually called an auto-covariance function.

### 2.4. Full Range Auto Regressive (FRAR) Model

Let $X$ be a random variable that represents the intensity value of a pixel at location $(k, l)$ in an image of size $(L \times L)$. It is also assumed that $X$ may have noise and is considered as independently and identically distributed Gaussian random variable with discrete time space and continuous state space with mean zero and variance $\sigma^2$ and is denoted as $\varepsilon(k, l)$. That is., $\varepsilon(k, l) \sim N(0, \sigma^2)$. 

30
Since \( \{X_t; t \in S\} \) is a stochastic process, where 
\( S = \{t(k - l); 1 \leq k, l \leq N\} \), \( \{X(t)\} \) can be considered as a Markov process if it has the conditional probability

\[
P\left\{ X(t_n) = i_n \mid X(t_k) = i_k; k = 0, 1, 2, \ldots, n - 1 \right\} \quad \text{... (2.8)}
\]

\[
P\left\{ X(t_n) = i_k \mid X(t_{n-1}) = i_{n-1} \right\} \quad \text{... (2.9)}
\]

\( \forall i_k, k = 0, 1, 2, \ldots, n-1 \) and \( t_k \) belonging to the state space \( S \) and \( t_0 < t_1 < \ldots < t_n \).

The two dimensional monochrome images are modeled \( \{X(k,l), 1 \leq k, l \leq N\} \) by means of discrete spatial interval of equal distance with Gaussian Markov Random Field (MRF). Thus, the proposed equation (2.10), a family of models by a discrete-time stochastic process \( \{X(t)\}, t = 0, \pm 1, \pm 2, \pm 3, \ldots \), is called the Full Range Auto Regressive (FRAR) model, by the difference equation

\[
X(k,l) = \sum_{p=-l}^{l} \sum_{q=-l}^{l} \frac{K \sin(\theta) \cos(\rho)}{a'} X(k + p, 1 + q) + \epsilon(k,l)
\]

\[
= \sum_{p=-l}^{l} \sum_{q=-l}^{l} \Gamma_r X(k + p, 1 + q) + \epsilon(k,l) \quad \text{... (2.10)}
\]

where \( \Gamma_r = \frac{K \sin(\theta) \cos(\rho)}{a'} \), and \( r = |p + q| \).

In the above equation (2.10), \( X(k + p, l + q) \) accounts for the spatial variation owing to image properties and \( \epsilon(k,l) \) is the spatial variation.
owing to additive noise and the model coefficient \( \Gamma_r = \frac{K \sin(r\theta) \cos(r\phi)}{\alpha^r} \) is the \( r^{th} \) coefficient of variation among the low-level primitives in the small image region. The coefficients are interrelated. The interrelationship is established through the model parameters \( K, \alpha, \theta, \) and \( \phi \) which are real. The model parameters are estimated on the intensity values of sub image with the concept of Bayesian approach. The model coefficients \( \{r_s\} \) in the equation (2.10) are functions of \( K, \alpha, \theta, \phi \) as well as \( n \). i.e.,

\[
\Gamma_r = \Gamma_r(K, \alpha, \theta, \phi) = \frac{K}{\alpha^r} \sin(r\theta) \cos(r\phi) \quad \text{where } K \in \mathbb{R}: \text{set of real numbers;}
\]

\( \alpha > 1; \theta, \phi \in [0, 2\pi] \) and \( n \in \{1, 2, \ldots\} \).

The proposed model is employed to analyse a two-dimensional discrete monochrome image of size \((L \times L)\). The given input image is assumed to be a Gaussian Markov Random Field. The image is partitioned into various sub images of size \((M \times M)\), \((M < L)\), to locally characterise the nature of the image. With the Markovian assumption, the conditional probability of \( X(t) \) given all other values only depends upon the nearest neighbourhood values, i.e., \( P(X(t)/X(t-j)); j = 1, 2, \ldots, M. \)

The initial assumptions about the parameters are \( K \in \mathbb{R}, \alpha > 1, \) and \( \theta, \phi \in [0, 2\pi] \). Further restriction on the range of parameters is placed by examining the identifiability of the model. It is interesting to note that some of the models used in the previous works, such as, white noise, auto regressive finite order and infinite order auto regressive models can be regarded as special cases of the proposed model. Thus
(i) if we set $\theta = 0$, then the FRAR model reduces to the white noise process.

(ii) when $\alpha$ is large, the coefficients $r_s$ become negligible as $r$ increases. So the FRAR model is reduced to AR ($p$) model approximately, for a suitable value of $p$.

(iii) when $\alpha$ is chosen less than one, then the FRAR model becomes an explosive infinite order AR model.

The fact that $X(t)$ has regressive on its neighbourhood pixels gives rise to the terminology of auto regressive process. However, in this case, the dependence of $X(t)$ on neighbourhood values may be true to some extent. In fact, the process is Gaussian under the assumption that the $\epsilon(k, l)$ is Gaussian and in this case its probabilistic structure is completely determined by its second order properties. Second order properties meant for the proposed FRAR model is asymptotically stationary up to order two, provided $1 - \alpha < K < \alpha - 1$. Finally the range of the parameters of the model are set with the constraints $K \in \mathbb{R}, \alpha > 1, 0 < \theta < \pi, 0 < \phi < \pi/2$ and $\epsilon(k, l)$ are independent and identically distributed Gaussian random variables with mean zero and variance $\sigma^2$.

### 2.4.1. The proposed FRAR model for CBIR

Existing computer vision technologies can readily extract features at low-level and intermediate levels (color, texture, depth, edge, region, simple motion and shape, etc). Recent works are also attempted to incorporate user feedbacks to improve the relevance of the image retrieval
results. However, the field is still young and the state-of-the-art has been quite primitive. In an effort to enable efficient retrieval in a multimedia database, we have been developing a multi-level data modeling and retrieval system to facilitate CBIR. The system integrates the low-level descriptions of the image and their associated confidence factors.

The characteristics of the proposed FRAR model are: (a) the exploration of CBIR from largely reduced image data, (b) the exploitation of intrinsic image features that are most effective for CBIR, and (c) the CBIR applicable to all possible low-level features. New and better methods of Modeling and Matching are essential for effective and efficient CBIR in image databases. Hence, a new family of Full Range Auto Regressive model has been proposed to design a CBIR system for the efficient retrieval of images. The usage of the proposed FRAR model based framework that facilitates texture analysis and edge detection are presented in Chapter 3 and Chapter 4 of the thesis. The auto correlation function is derived from the auto correlation coefficients of the proposed model and used for texture analysis and detecting edges for the untextured images.

2.4.2. Feature extraction of images with auto correlation function

One of the problems in designing an efficient CBIR technique is extracting a set of appropriate features from the image. Hence, there is no effective theory, the task of finding effective features for retrieval of an image is still an open issue, but there is a possibility to state some desirable properties of features of an image:
• The features should be informative.
• The dimensionality of features should be as low as possible.
• The features should have high discriminating power.
• The features should convey some meaning to human perception.

Since the textural properties of image appear to carry useful information for discrimination purpose, it is important to develop features for texture based on suitable quantitative representation of some of these properties. Different models have been developed to solve the low-level texture analysis problems. The auto regressive model for texture analysis exploits this linear dependence and the auto correlation function is a feature that tells about the size of the tonal primitives. Hence the auto correlation coefficients are much useful to extract the features of texture images.

In order to retrieve images, we must be able to efficiently compare two images to determine the similar contents. An efficient matching scheme further depends upon the discriminatory information contained in the feature vectors of the images. Let \( I = \{I_1, I_2, \ldots, I_N \} \) be the sequence of all gray level images in an image database (IDB). Thus \( I_i \) \((M, N)\) denotes a gray-scale image of size \((M \times N)\) with pixel values in the range from 0 to 255 for the image \( I_i \).

Let \( z \) represents a mapping from the image set \( I \) onto a \( n \)-dimensional feature space \( F \) given by

\[ z:I \rightarrow F \]
where

\[
F = \begin{bmatrix}
\alpha_i \\
\alpha_j \\
\vdots \\
\alpha_n
\end{bmatrix}, \quad \alpha_j \in \mathbb{R}; \quad i = 1, 2, \ldots, N; \quad j = 1, 2, \ldots, n
\]  \quad \ldots (2.11)

and \( \mathbb{R} \) is the set of real numbers.

Under this mapping, each image \( I_i \) is represented by a unique feature vector

\[
F_i = \begin{bmatrix}
\alpha_i \\
\alpha_j \\
\vdots \\
\alpha_n
\end{bmatrix}, \quad \alpha_j \in \mathbb{R}; \quad j = 1, 2, \ldots, n
\]  \quad \ldots (2.12)

The following section contains a description of the proposed texture feature vector extraction for effective retrieval with auto correlation coefficients.

Let \( \{ F(x, y); x = 1, 2, \ldots, X, \; y = 1, 2, \ldots, Y \} \) be a two-dimensional image pixel array. For color images \( F(x, y) \) denotes the color value at pixel \((x, y)\) i.e., \( F(x, y) = \{ F_R(x, y), F_B(x, y), F_B(x, y) \} \). For black and white images \( F(x, y) \) denotes the grayscale intensity value of pixel \((x, y)\).

The problem of retrieval can be stated as: For a query image \( Q \), we find image \( T \) from the image database, that the distance between corresponding feature vectors is less than specified threshold, i.e.,

\[
D(\text{Feature}(Q), \text{Feature}(T)) \leq t
\]
2.4.3. Auto correlation function based on the proposed model

The Auto correlation function evaluates the linear spatial relationships between primitives based on distance measures. In order to extract the features of the textures, colors, edges and shapes, the undermentioned set of autocorrelation coefficients are used:

\[
C_{jj}(p, q) = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N-q} f(i, j) f(i+p, j+q)}{\left( \sum_{i=1}^{M-p} \sum_{j=1}^{N-q} f^2(i, j) \right)^{\frac{1}{2}}}
\]

... (2.13)

where \(p, q\) are the positional difference in the \(i, j\) direction and \(M, N\) are image dimensions.

The FRAR Model has been used to extract the feature vectors with auto correlation coefficients. The auto correlation coefficients are calculated as mentioned in equation (2.13) with the positional differences \((p, q)\). The calculated auto correlation coefficients are converted into autonums with a simple transformation. In the proposed FRAR model, the extracted autonums are considered as the feature vectors of the input image to generate the feature set and is represented below:

\[
F = \{F_1, F_2, \ldots, F_r\}
\]

... (2.14)

where \(F_r\) denotes the feature vectors of the input image.

The feature sets are generated with the low level features such as color, texture, edges and shapes using the auto correlation coefficients as mentioned in the equation (2.10) and the feature database of the image.
database (IDB) is established. Then the target image has been retrieved from the image database with the feature sets of the feature database.

The detailed procedures employed to extract textual features, color features, edge and shape features with the proposed FRAR model and their usage in retrieval of images are discussed in the subsequent chapters.