A Statical Color Image Segmentation Using a Diagonal Of The Modified Riesz Mixture Model

Taher Ben Arab¹,², Mourad Zribi¹, and Afif Masmoudi²

¹Laboratoire d’Informatique Signal et Image de la Cte d’Opale (LISIC-EA 4491), ULCO, 50 Rue Ferdinand Buisson BP 719, 62228 Calais Cedex, France.
²Laboratoire de Probabilités et Statistique, Faculté des Sciences de Sfax, Université de Sfax. B. P. 1171 Sfax, Tunisie.

This paper describes a new approach to adapted estimation of parametric mixture model based on diagonal of the modified Riesz distribution (DMRD) defined in $\mathbb{R}^r$, $r \geq 2$. The DMRD can model accurately a withe variety of color image. This parameters index a family of distribution which include the bivariate Gamma and the convolution product between bivariate Gamma and univariate Gamma. In our work, we have applied the bivariate Gamma distribution to mixture model. We use the Expectation Maximization (EM) algorithm to estimate the model parameters of the color image data and the number of mixture components is estimated by using $K$-means clustering algorithm. The $K$-means clustering algorithm is also utilized for developing the initial estimates of the EM-algorithm. This study investigates the DMRD in unsupervised color image segmentation. Experiments with real color images are described which verify the validity of method and its implementation.

Index Terms—Diagonal of the modified Riesz distribution, Expectation Maximization algorithm, $K$-means clustering algorithm, Unsupervised color image segmentation.

I. INTRODUCTION

IMAGE segmentation is the important problem in image analysis. It consists of partitioning an image into segment (class) having a homogenous description, generally in terms of a parametric model of the image. The image segmentation methods can be divided into two categories depending upon the type of image. The images can be broadly categorized into two types namely, gray level images and color images. A gray level image is usually characterized by pixel intensity. But in color images the color is a perceptual phenomenon related to human response to different wavelengths in the visible electro-magnetic spectrum. In color images the features that represent the image pixel are highly influenced by three feature descriptions namely, intensity, color and texture. Among these features color is the most important one in segmenting the color images since intensity and texture features also be embedded in color features. A better color space than the Read Green and Blue (RGB) space in representing the colors of human is perception the Hue Saturation and Intensity (HSI) space, in which the color information is represented by Hue and Saturation (HS) values. Thus the human perception of image can be characterized through a bivariate random variable consisting of HS which can be measured using generic structure of a color appearance model.

In model based image segmentation the whole image is divided into different image regions and each image region is characterized by a suitable probability distribution. For ascribing a probability model to the feature vector of the pixels in the image region, it is needed to study the statistical characteristics of the feature vector.

In image segmentation it is customary to consider that the whole image is characterized by a finite DMRD mixture model. In particular case, the diagonal of the Riesz distribution is represent the bivariate Gamma distribution (BGD) definite in $\mathbb{R}^2$. That is, the feature vector of each image region follows a Gamma distribution. But in many color images the feature vector represented by HS are having finite values. Hence, to have an accurate image segmentation of these sorts of color images it is needed to develop and analyze image segmentation methods based on bivariate mixture distributions.

Here, it is assumed that the feature vector in different image regions follows a BGD distribution and the feature vector of the whole image is characterized by a finite bivariate Gamma mixture model.

In our work, the number of image regions is obtained by $K$-means algorithm for which the initial value of the number of components is identified from the number of peaks in the image histogram. The model parameters are estimated by using Expectation Maximization (EM) algorithm. The EM-algorithm is one of the most preferred method of estimating the model parameters in mixture distributions. The EM-algorithm requires the updated equations of the model parameters which are derived for the bivariate Gamma mixture model. The initialization of the model parameters for carrying the EM-algorithm is done through feature vector of the pixel intensities of the image regions obtained through $K$-means clustering and moment method of estimation. An image segmentation algorithm with component likelihood maximization under Bayesian frame work is also developed and analyzed. The rest of this paper is organized as follows. Section II presents some definition of the DRMD and its properties. Section III describes the DMRD segmentation algorithm. Section IV is devoted to experimental results and performance.
Section VI contains a conclusion.

II. MODEL OF THE DIAGONAL OF THE MODIFIED RIESZ DISTRIBUTION

Under the approach based on the theory of Jordan algebras, a family of distributions on symmetric cones, termed the Riesz distributions, was first introduced by Hassairi and Lajmi [12] under the name of Riesz natural exponential family (Riesz NEF); it was based on a special case of the so-called Riesz measure from Faraut and Koranyi [9], going back to Riesz [14]. This Riesz distribution generalizes the matrix multivariate Gamma and Wishart distributions, containing them as particular cases [1].

A. Riesz distribution

The study of Riesz distributions originates from the paper [14] of Marcel Riesz, where he introduces a family of convolution operators as a generalization of the classical Riemann- Liouville operators on the positive reals to the Lorentz cone in $\mathbb{R}^r$.

Let $E$ be the Euclidean space of $(r, r)$ real symmetric matrices equipped with the scalar product $x, y > = tr(x y)$ and $\Omega$ denotes the cone of $(r, r)$-symmetric positive definite matrices. For $x = (x_{ij})_{1 \leq i, j \leq r}$ in $E$ and $1 \leq k \leq r$, we define the sub-matrices $P_k(x) = (x_{ij})_{1 \leq i, j \leq k}$ and $\Delta_k(x)$ denotes the determinant of the $(k, k)$ matrix $P_k(x)$. Then the generalized power of $x$ in the cone $\Omega$ of positive definite elements of $E$ is defined, for $s_1 = (s_1, s_2, \ldots, s_r) \in \mathbb{R}^r$, by

$$\Delta_{s_k}(x) = \Delta_1(x)^{s_1 - s_k} \Delta_2(x)^{s_2 - s_3} \ldots \ldots \Delta_r(x)^{s_r}. \quad (1)$$

For $s_1 = (s_1, s_2, \ldots, s_r)$ satisfying the conditions $s_i > \frac{i - 1}{2}$, the absolutely continuous Riesz measure is defined by

$$R_{s_k}(dx) = \frac{\Delta_{s_k}-\frac{p}{x}}{\Gamma\left(\Omega(s_k)\right)} 1_{\Omega}(x) dx, \quad (2)$$

where $1_{\Omega}(x)$ is the indicator function defined on $\Omega$, $n = \frac{r(r+1)}{2}$ is the dimension of $E$, $\Gamma(\Omega(p)) = (2\pi)^{r(r-1)/2} \prod_{i=1}^{r} \Gamma(s_i - 1/2) (i - 1))$ and $\Gamma(a) = \int_{0}^{+\infty} e^{-x}x^{-a}dx, a > 0$.

Then, a result due to Gindikin [10] says that for all $\theta \in -\Omega$, the Laplace transform of measure $R_{s_k}$ is defined by

$$L_{R_{s_k}}(\theta) = \int_{E} e^{\theta(x)} R_{s_k}(dx) = \Delta_{s_k}(\theta^{-1}). \quad (3)$$

and for $s_k$ satisfying the conditions $s_i > \frac{i - 1}{2}$, the Riesz distribution with shape parameter $s_k$ and scale parameter $\sigma$ in $\Omega$ is given by

$$R(s_k, \sigma)(dx) = \frac{e^{-\sigma x} \Delta_{s_k}(x)}{\Gamma(s_k) \Delta_{s_k}((\sigma)^{-1})} 1_{\Omega}(x) dx. \quad (4)$$

When $s_1 = s_2 = \ldots = s_r = p > \frac{r - 1}{2}$, $R(s_k, \sigma)$ is reduced to the Wishart distribution

$$W(p, \sigma) = \frac{1}{\Gamma(p)} \det(\sigma^{-p}) e^{-\sigma x} \det(x)^{p - \frac{r}{2}} 1_{\Omega}(x). \quad (5)$$

B. The proposed DMRD

In this section, we introduce a model of the DMRD. The model is based on Laplace transform of Riesz distribution. In our work, we study the case of $r = 2$.

Let $X$ be a random Riesz $(2, 2)$-matrix with parameters $s_2 = (s_1, s_2) \in [0, +\infty[ \times \frac{1}{2}, +\infty[$ and $\theta$ in $I_2 - \Omega$, where $I_2$ is the identity matrix of order 2. Then, according to equation (3), the Laplace transform of $X$ is defined by

$$L_X(\theta) = \int_{E} e^{\theta(x)} R_{s_k}(dx) = \Delta_{s_k}((I_2 - \theta)^{-1}). \quad (6)$$

To obtain our model, we use an affine transformation to the random $X$ given by $\bar{X} = \Sigma X$ where $\Sigma = \left(\begin{array}{cc} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{array}\right)$, $\Sigma_{11} > 0$, $\Sigma_{22} > 0$ and $\Sigma_{11} \Sigma_{22} - \Sigma_{12}^2 > 0$.

The proposed model of the diagonal of the modified Riesz is given by $Y = (Y_1, Y_2) = diag(\bar{X})$ (the diagonal elements of the matrix $\bar{X}$).

To define the Laplace transform $L_Y(\theta)$ of $Y$, we introduce the following proposition.

**Proposition II.1.** Let $s_2 = (s_1, s_2) \in [0, +\infty[ \times \frac{1}{2}, +\infty[$ and $\theta \Sigma \in I_2 - \Omega$, then

$$L(\theta) = [(1 - \theta_1 \Sigma_{11})(1 - \theta_2 \Sigma_{22}) - \theta_1 \theta_2 \Sigma_{12}^2]^{-s_1} \times \left[1 - \theta_2 \Sigma_{22}\right]^{-s_2}. \quad (7)$$

**Proof**

Firstly, let $Y = diag(\bar{X})$ be a random vector of $\mathbb{R}^2$ whose elements are the diagonal elements of the matrix $\bar{X}$. Then, for all $\theta \in I_2 - \Omega$, we have

$$L_X(\theta) = \Delta_{s_k}((I_2 - \theta \Sigma)^{-1}) = [1 - \theta_2 \Sigma_{22}]^{-s_2} \left[(1 - \theta_1 \Sigma_{11})(1 - \theta_2 \Sigma_{22}) - \theta_1 \theta_2 \Sigma_{12}^2\right]^{-s_1}. \quad \square$$

In the remainder of this section, we will define the distribution of $Y$ in both case $(s_1 = s_2)$ and $(s_2 > s_1)$.

In this two cases, we define in propositions (II.2) and (II.5) the probability density function (pdf) $f_{x,y}$ of $Y$.

**B.1. Case I: $s_1 = s_2$.**

**Proposition II.2.** Let $s_1 = s_2, \theta \in -\Omega$ and $\Sigma$ is a positive definite matrix. The distribution of the random variable $Y$ is a BGD with pdf is defined by

$$f_{1,y}(y_1, y_2) = \exp(-\frac{\Sigma_{22}y_1 + \Sigma_{11}y_2}{\Sigma_{11} \Sigma_{22} - \Sigma_{12}^2}) \times \frac{(y_1 y_2)^{y_1-1}}{\Sigma_{11} \Sigma_{22} - \Sigma_{12}^2} \Gamma(s_1), \quad (y_1 y_2)^{y_1-1} \times f_{s_1}(\theta_1 y_1, \theta_2 y_2) I_{[0, +\infty[}(y_1, y_2), \quad (8)$$

where $I_{[0, +\infty[}(y_1, y_2)$ is the indicator function defined on $[0, +\infty[\times [0, +\infty[\times [0, +\infty[\times [0, +\infty[\times [0, +\infty[\times [0, +\infty[$ and $f_{s_1}(z)$ is given by

$$f_{s_1}(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma((s_1 + k)k)!}, \quad \forall z \in \mathbb{R}. \quad (9)$$


We present in Fig.1 a BGD for different values of parameters.

\[ L(\theta) = [(1 - \theta_{i1}\Sigma_{i1})(1 - \theta_{i2}\Sigma_{i2}) - \theta_{i1}\theta_{i2}\Sigma_{i1}^2 - \Sigma_{i2}^2]^{-s_i}. \]

According to [3], if \( \theta \in -\Omega \), then this Laplace transform is related to a BGD with pdf defined in Eq.4.

\[ \text{Proposition II.4.} \]

We can express the bivariate random variable \( Y \) defined by

\[ f_{Y_i}(y_i) = \left( \frac{y_i}{\Sigma_{ii}} \right)^{s_i-1} \frac{\exp\left( - \frac{y_i}{\Sigma_{ii}} \right)}{\Sigma_{ii} \Gamma(s_i)} I_{[0, +\infty)}(y_i). \]  

where \( s_i \) > 0 represent the shape parameter and \( \Sigma_{ii} > 0 \) called the scale parameter.

To define the parameters of the random variable \( Y \), we introduce the following proposition.

\[ \text{Proposition II.3.} \]

The marginal distribution \( Y_i, \ i = 1, 2 \), is distributed according to a univariate gamma distribution with pdf defined by

\[ f_{Y_i}(y_i) = \frac{y_i^{s_i-1} \exp\left(-\frac{y_i}{\Sigma_{ii}}\right)}{\Sigma_{ii} \Gamma(s_i)} I_{[0, +\infty)}(y_i). \]  

This different moments cited in Proposition II.4 can be obtained by differentiating this expression of the Laplace transform defined in Eq.3, where \( s_1 = s_2 \), with respect to \( \theta_{i,j}, \ i, j = 1, 2 \).

\[ \text{B.2. Case 2:} \ s_2 > s_1. \]

The Laplace transform of \( Y \) according to the DMRD, given in Eq.3 is the Laplace transform of a convolution product of two independent random variables \( Z \) and \( T \), where:

- \( T \) is a univariate gamma distribution with shape parameter \( s_2 - s_1 \) and scale parameter \( \Sigma_{i2} \) with pdf \( f_T(t) \) defined in Eq.5.
- \( Z = (Z_1, Z_2) \) is a BGD with parameters \( (\Sigma_{i1}, \Sigma_{i2}, \Sigma_{j1}, \Sigma_{j2}) \), where \( \Sigma_{i1} = \Sigma_{i2} - \Sigma_{j2} \) with pdf \( f_{Z_i}(z_i) \) defined in Eq.4.

We can express the bivariate random \( Y \) by \( Y = (Z_1, Z_2 + T) \).

\[ \text{Proposition II.5.} \]

Let \( 0 < s_1 < s_2 \) and \( \Sigma \) is a positive definite matrix. In this case, the distribution of the random variable \( Y \) is defined by

\[ f_{2,Y}(y_1, y_2) = \frac{1}{\Sigma_{i1} \Sigma_{i2}} \left( \text{var} Y_1 \right)^{-s_1} e^{-\frac{y_1^2}{\Sigma_{i1}}} \frac{\exp\left(-\frac{y_2}{\Sigma_{i2}}\right)}{\Gamma(s_2)} \Phi_3 \left( \begin{array}{c} y_1 - s_1 \Sigma_{i1} \Sigma_{i2}^{-1} \Sigma_{i2} \\ s_2 - s_1 \end{array} ; s_2, y_2, y_1 \right), \]

where \( h = \delta \Sigma_{i1} \Sigma_{i2}^{-1} \Sigma_{i2} \) and \( \Phi_3 \) is the so-called Horn function.

\[ \text{Proof} \]

Firstly, by using the independence assumption between \( Z \) and \( T \), the density of \( Y \) can be expressed as

\[ f_{2,Y}(y_1, y_2) = \int_0^{\infty} f_{Z_1, T}(y_1, y_2, u) f_T(u) du. \]

\[ \text{Secondly, by making the following change of variable} \ u = \frac{y_2}{\Sigma_{i2}}, \ \text{in Eq.7} \] and by using the following series expansion \( e^t = \sum_{j \geq 0} \frac{t^j}{j!} \), we have the desire result.

To define the parameters of the random variable \( Y \), we introduce the following proposition.

\[ \text{Proposition II.6.} \]

The moments and the covariance of the random variable \( Y \), with pdf defined in Eq.4, are given by

\[ E[Y_1] = s_1 \Sigma_{i1}, \ E[Y_2] = s_2 \Sigma_{i2} \]

\[ \text{var}(Y_1) = s_1 \Sigma_{i1}^2, \ \text{var}(Y_2) = s_2 \Sigma_{i2}^2 \]

\[ \text{cov}(Y_2, Y_2) = s_1 \Sigma_{i1}^2. \]

This different moments cited in Proposition II.6 can be obtained by differentiating this expression of the Laplace transform defined in Eq.3, where \( s_2 > s_1 > 0 \), with respect to \( \theta_{i,j}, \ i, j = 1, 2 \).
C. Parameters estimation of the DMRD

In this section, we estimate the parameters of the DMRD in both cases \((s_1 = s_2\) and \(s_2 > s_1\)). We assume that \((s_1, s_2)\) is a known parameters. Therefore, the DMRD is characterized by unknown parameters \(\Pi = (\Sigma_{11}, \Sigma_{22}, \Sigma_{12}, \delta)\), where \(\Sigma_{12} = \Sigma_{11} \Sigma_{22} - \Sigma_{12}^2\).

Let \(Y = (Y_1, \ldots, Y_n)\), where \(Y_i = (Y_{i1}, Y_{i2})\), \(n\) vectors distributed according to a DMRD with the unknown parameter vector \(\Pi\).

C.1. Case 1: \(s_1 = s_2 = s\)

C.1.1. Method of Maximum Likelihood (MML): The likelihood function of a sample bivariate observations \(Y_1, Y_2, \ldots, Y_n\) of density defined in Eq.4 is given by

\[
l(Y; \Pi) = -ns \ln(\Sigma_{12}) - n\Sigma_{22} \bar{Y}_1 - \frac{n}{s} \Sigma_{12} \bar{Y}_2 + (s_1 - 1) \sum_{i=1}^{n} \ln(y_{i1}y_{i2}) - \log \Gamma(s_1) - n\Sigma_{11} \bar{Y}_2 + \sum_{i=1}^{n} \ln(f_s(\delta y_{i1}y_{i2})),
\]

(12)

where \(\bar{Y}_j = \frac{1}{n} \sum_{i=1}^{n} Y_{ij}, j = 1, 2\), \(\Sigma_{12} = \Sigma_{11} \Sigma_{22} - \Sigma_{12}^2\) and \(\delta = \frac{\Sigma_{12}}{\Sigma_{12}}\).

By taking the differential, with respect to \(\Sigma_{11}\), \(\Sigma_{22}\) and \(\Sigma_{12}\), one has

\[
-ns + n\Sigma_{22} \bar{Y}_1 - \frac{n}{s} \Sigma_{22} \bar{Y}_2 + n\Sigma_{11} \bar{Y}_2 - \frac{2 \Sigma_{12}^2}{\Sigma_{12}} H = 0,
\]

(13)

\[
-ns + n\Sigma_{11} \bar{Y}_2 - \frac{n}{s} \Sigma_{11} \bar{Y}_1 + n\Sigma_{22} \bar{Y}_1 - \frac{2 \Sigma_{12}^2}{\Sigma_{12}} H = 0,
\]

(14)

\[
ns - n\Sigma_{22} \bar{Y}_1 - n\Sigma_{11} \bar{Y}_2 + \frac{\Sigma_{12}^2}{\Sigma_{12}} + \Sigma_{11} \Sigma_{22} - \Sigma_{12}^2 H = 0,
\]

(15)

with \(H = \sum_{i=1}^{n} y_{i1}y_{i2} f_{s+1}(\delta y_{i1}y_{i2})/f_s(\delta y_{i1}y_{i2})\) and \(f_{s+1} = f_s\).

From Eq.9, Eq.10 and Eq.11, the ML estimators of \(\hat{\Sigma}_{11}\) and \(\hat{\Sigma}_{22}\) are defined by

\[
\hat{\Sigma}_{11} = \frac{\bar{Y}_1}{s}, \quad \hat{\Sigma}_{22} = \frac{\bar{Y}_2}{s}
\]

(16)

By replacing \(\Sigma_{11}\) and \(\Sigma_{22}\) by their estimators in Eq.8, then the estimator of \(\hat{\Sigma}_{12}\) is the root of the following function

\[
\phi(\hat{\Sigma}_{12}) = n - \frac{s}{\bar{Y}_1 \bar{Y}_2 - s^2 \Sigma_{12}} \sum_{i=1}^{n} y_{i1}y_{i2} f_{s+1}(\delta y_{i1}y_{i2}) = 0,
\]

(17)

where \(\hat{\delta} = \frac{s \Sigma_{12}^2}{(\bar{Y}_1 \bar{Y}_2 - s^2 \Sigma_{12})^2}\).

A closed-form solution of Eq.13 does not exist the parameter \(\hat{\Sigma}_{12}\). We can obtain a solution by using a Newton-Raphson procedure. The convergence of the Newton-Raphson procedure is generally obtained after few iterations.

C.1.2. Method of Moments (MM): The estimator of \(\Pi = (\Sigma_{11}, \Sigma_{22}, \Sigma_{12})\) by the MM is the solution of the following system

\[
\begin{cases}
\bar{Y}_1 = E[Y_1], & \bar{Y}_2 = E[Y_2] \\
\frac{1}{n} \sum_{i=1}^{n} (Y_{1i} - \bar{Y}_1)^2 (Y_{2i} - \bar{Y}_2) = cov(Y_2, Y_2)
\end{cases}
\]

Consequently,

\[
\hat{\Sigma}_{11} = \frac{\bar{Y}_1}{s}, \quad \hat{\Sigma}_{22} = \frac{\bar{Y}_2}{s}
\]

(18)

\[
\hat{\Sigma}_{12} = \frac{1}{ns} \sum_{i=1}^{n} (Y_{1i} - \bar{Y}_1)^2 (Y_{2i} - \bar{Y}_2).
\]

(19)

In this case, we constate that the parameters \(\Sigma_{11}\) and \(\Sigma_{22}\) have the same estimators obtained by the MML and the MM.

C.2. Case 2: \(s_2 > s_1\)

C.2.1. Method of Maximum Likelihood: The log-likelihood function of \(Y\) is given by

\[
\begin{align*}
l(Y; \Pi) &= -ns_1 \ln \Sigma_{12} + n(s_1 - s_2) \ln(\Sigma_{22}) - n\Sigma_{22} \bar{Y}_1 + \frac{n}{s_1} \Sigma_{22} \bar{Y}_2 + n\Sigma_{11} \bar{Y}_2 - \frac{2 \Sigma_{12}^2}{\Sigma_{12}} H = 0, \\
-ns_1 + n\Sigma_{22} \bar{Y}_1 - \frac{n}{s_2} \Sigma_{22} \bar{Y}_2 + n\Sigma_{11} \bar{Y}_2 - \frac{2 \Sigma_{12}^2}{\Sigma_{12}} H = 0,
\end{align*}
\]

(20)

with \(H = \sum_{i=1}^{n} y_{i1}y_{i2} f_{s+1}(\delta y_{i1}y_{i2})/f_s(\delta y_{i1}y_{i2})\) and \(f_{s+1} = f_s\).

By taking the differential, with respect to \(\Sigma_{11}\), \(\Sigma_{22}\) and \(\Sigma_{12}\), one has

\[

\begin{align*}
-ns_1 + n\Sigma_{22} \bar{Y}_1 - \frac{n}{s_2} \Sigma_{22} \bar{Y}_2 + n\Sigma_{11} \bar{Y}_2 - \frac{2 \Sigma_{12}^2}{\Sigma_{12}} H &= 0, \\
-ns_1 + n\Sigma_{11} \bar{Y}_2 - \frac{n}{s_2} \Sigma_{11} \bar{Y}_1 + n\Sigma_{22} \bar{Y}_1 - \frac{2 \Sigma_{12}^2}{\Sigma_{12}} H &= 0,
\end{align*}
\]

\[
ns_1 - n\Sigma_{22} \bar{Y}_1 - n\Sigma_{11} \bar{Y}_2 + \frac{\Sigma_{12}^2}{\Sigma_{12}} + \Sigma_{11} \Sigma_{22} - \Sigma_{12}^2 H = 0,
\]

(21)

with \(H = \sum_{i=1}^{n} y_{i1}y_{i2} f_{s+1}(\delta y_{i1}y_{i2})/f_s(\delta y_{i1}y_{i2})\) and \(f_{s+1} = f_s\).

A closed-form solution of Eq.13 does not exist the parameter \(\hat{\Sigma}_{12}\). We can obtain a solution by using a Newton-Raphson procedure. The convergence of the Newton-Raphson procedure is generally obtained after few iterations.

C.2.2. Method of Moments (MM): The estimator of \(\Pi = (\Sigma_{11}, \Sigma_{22}, \Sigma_{12})\) by the MM is the solution of the following system

\[
\begin{cases}
\bar{Y}_1 = E[Y_1], & \bar{Y}_2 = E[Y_2] \\
\frac{1}{n} \sum_{i=1}^{n} (Y_{1i} - \bar{Y}_1)^2 (Y_{2i} - \bar{Y}_2) = cov(Y_2, Y_2)
\end{cases}
\]

Consequently,

\[
\hat{\Sigma}_{11} = \frac{\bar{Y}_1}{s}, \quad \hat{\Sigma}_{22} = \frac{\bar{Y}_2}{s}
\]

(18)

\[
\hat{\Sigma}_{12} = \frac{1}{ns} \sum_{i=1}^{n} (Y_{1i} - \bar{Y}_1)^2 (Y_{2i} - \bar{Y}_2).
\]

(19)

In this case, we constate that the parameters \(\Sigma_{11}\) and \(\Sigma_{22}\) have the same estimators obtained by the MML and the MM.

A closed-form solution in Eq.17, Eq.18 and Eq.19 does not exist for the \(\hat{\Sigma}_{11}, \hat{\Sigma}_{22}\) and \(\hat{\Sigma}_{12}\) parameters. We can obtain
a solution by using a Newton-Raphson procedure.

C.2.2. Method of Moments: The estimators of \( \Pi = (\Sigma_{11}, \Sigma_{22}, \Sigma_{12}^2) \) by MM is the solution of the following system
\[
\begin{align*}
\bar{Y}_1 &= E[Y_1], \quad \bar{Y}_2 = E[Y_2] \\
\frac{1}{n} \sum_{i=1}^{n} (Y_1^{(i)} - \bar{Y}_1)^t (Y_2^{(i)} - \bar{Y}_2) &= \text{cov}(Y_2, Y_2)
\end{align*}
\]

Consequently,
\[
\hat{\Sigma}_{11} = \frac{\bar{Y}_1}{s_1}, \quad \hat{\Sigma}_{22} = \frac{\bar{Y}_2}{s_2}
\]
\[
\hat{\Sigma}_{12}^2 = \frac{1}{ns_1} \sum_{i=1}^{n} (Y_1^{(i)} - \bar{Y}_1)^t (Y_2^{(i)} - \bar{Y}_2)
\]

D. Some Simulations

In order to compare the performance of the MM estimator and the ML estimator, we propose any simulations.

D.1. Case 1: \( s_1 = s_2 = s \).
In this case, we generate a random vector \( Y \) according to a BGD with different parameters

i) \( s = 3, \Sigma_{11} = 0.3, \Sigma_{22} = 0.5, \)

ii) \( s = 3, \Sigma_{11} = 10, \Sigma_{22} = 5. \)

The comparative study of the MM and ML is characterized by the MSE as a function of \( n \), where \( n \) is the size of the sample. The number of resampling is \( N = 1000 \). In Fig.3 and Fig.4 we present respectively in Fig.3 and Fig.4 the MSE of the estimator of the parameter \( \Sigma_{12}^2 \) in two case (\( \Sigma_{12}^2 = 0.105 \) and \( \Sigma_{12}^2 = 45 \)). The circle curves correspond to the estimator of MM whereas the triangle curves correspond to the estimator of ML. We observe from the two figures that the ML method is more efficient than the MM method.

\[
\Sigma_{12}^2 = 0.105
\]

Fig. 3. MSE versus \( n \) for parameters \( \Sigma_{12}^2 (s = 3, \Sigma_{11} = 5, \Sigma_{22} = 10) \).

\[
\Sigma_{12}^2 = 45
\]

Fig. 4. MSE versus \( n \) for parameters \( \Sigma_{12}^2 (s = 3, \Sigma_{11} = 0.3, \Sigma_{22} = 0.5) \).

D.2. Case 2: \( s_2 > s_1 \).

In this case, we generate two random vectors \( Z \) and \( T \). The random vector \( Z = (Z_1, Z_2) \) follows the BGD with parameters \( s_1, \Sigma_{11}, \Sigma_{22}, \Sigma_{12}^2 \). The random vector \( T \) follows the univariate gamma distribution with parameters \( s_2 - s_1, \Sigma_{22} \). This two variables \( Z \) and \( T \) are independent. The random variable \( Y \) according to the DMRD is obtained from the two variables \( Z \) and \( T \).

The comparative study of the MM and ML is characterized by the MSE as a function of \( n \), where \( n \) is the size of the sample. The number of resampling is \( N = 1000 \). In Fig.5, Fig.6 and Fig.7 we present the variation of the MSE depending of \( n \), where \( n \) is the size of the sample, for different case of the parameters \( (s_1, s_2, \Sigma_{11}, \Sigma_{22}, \Sigma_{12}^2) \). Fig.5 presents the variation of the MSE for different case of the parameter \( \Sigma_{12}^2 \) where the parameters \( s_1, s_2, \Sigma_{11}, \Sigma_{22} \) are fixed. Fig.6 presents the variation of the MSE for different case of the parameter \( \Sigma_{11} \) where the parameters \( s_1, s_2, \Sigma_{22}, \Sigma_{12}^2 \) are fixed. Fig.7 shows the variation of the MSE for different case of the parameter \( \Sigma_{22} \) where the parameters \( s_1, s_2, \Sigma_{11}, \Sigma_{12}^2 \) are fixed. The circle curves correspond to the estimator of MM whereas the triangle curves correspond to the estimator of ML. We observe from the three figures that the ML method is more efficient than the MM method.

\[
\Sigma_{12}^2 = 14.14
\]

\[
\Sigma_{12}^2 = 63.63
\]

\[
\Sigma_{12}^2 = 35.35
\]

Fig. 5. MSE versus \( n \) for parameters \( \Sigma_{12}^2 (s_1 = 1, s_2 = 2, \Sigma_{11} = 10, \Sigma_{22} = 5) \).
In continuation in our work we will generalize the DMRD to the mixture model.

III. DMRD SEGMENTATION ALGORITHM

A. DMRD mixture model

A crucial problem here is the choice of the mixture pdf. Generally the Gaussian distribution is considered \([17]\), yet recent works have show that other distributions may provide better modelling capabilities. Among these distributions that we shall consider in this work, we have the DMRD for a known value of \(s_1 = s_2\) which represent the BGD defined in Proposition II.2. It is important to note that for a known value of \((s_1, s_2)\), a BGD is fully characterized by \(\Pi = (\Sigma_1, \Sigma_2, \Sigma_{12})\), where \(\Sigma_{12} = \Sigma_1 \Sigma_2 - \Sigma_{11}^2\).

Since the entire image is a collection of regions, which are characterized by BGD defined in Proposition II.2, it can be characterized through a K-component finite BGD and its pdf is of the form

\[
h(y_1, y_2) = \sum_{k=1}^{K} \alpha_k f_k(y_1, y_2; \Pi),
\]

where \(0 < \alpha_k < 1\), \(\sum_{k=1}^{K} \alpha_k = 1\) and \(f_k(y_1, y_2)\) is the pdf of the BGD given by Eq.4.

B. Estimation of the parameters mixture by EM-Algorithm

The problem of estimating the parameters which determine a mixture has been the subject of diverse studies. During the last two decades, the MML has become the most common approach to this problem. Of the variety of iterative methods which have been suggested as alternatives to optimize the parameters of a mixture \([3]\), the one most widely used is expectation maximization (EM). EM was originally proposed by \([6]\) for estimating the ML of stochastic models. The algorithm employs an iterative procedure and the practical form is usually simple.

To obtain the estimation of the model parameters, we utilized the EM-algorithm by maximizing the expected likelihood function for carrying out the EM-algorithm. The log-likelihood function of bivariate observations \((y_{11}, y_{21}), \ldots, (y_{1n}, y_{2n})\) drawn from an image with probability density function

\[
h(y_{11}, \ldots, y_{2n}) = \sum_{k=1}^{K} \alpha_k f_k(y_{11}, \ldots, y_{2n}; \Pi)
\]

is

\[
\log L(\Pi) = \frac{1}{n} \sum_{j=1}^{n} \log \left( \sum_{k=1}^{K} \alpha_k \exp \left( -\frac{\sum_{1}^{2} \sum_{1}^{2} y_{1j} + \sum_{1}^{2} y_{2j} }{\Sigma_{11} \Sigma_{22} - \Sigma_{12}^2} \right) \right)
\]

where \(\delta_k = \frac{\sum_{1}^{2} \sum_{1}^{2} - \Sigma_{12}^2}{\Sigma_{11} \Sigma_{22} - \Sigma_{12}^2}\).

The model parameters

\(\Pi = (\alpha_1, \ldots, \alpha_K, \Sigma_{11}, \Sigma_{22}, \Sigma_{12}, \ldots, \Sigma_{11}, \Sigma_{22}, \Sigma_{12}, \ldots, \Sigma_{11}, \Sigma_{22}, \Sigma_{12})\)

is estimated by using the EM algorithm.

The EM algorithm is decomposed in the two following steps:

**Step E:** The updated equation of the parameter \(\alpha_k\) is

\[
\alpha_k^{(l+1)} = \frac{1}{n} \sum_{j=1}^{n} \frac{\tau_{k,j} y_{1j} y_{2j} ; \Pi^{(l)}}{\sum_{k=1}^{K} \alpha_k f_k(y_{1j}, y_{2j}; \Pi^{(l)})}
\]

**Step M:** The updated equations of \(\Sigma_{11}, \Sigma_{22}, \Sigma_{12}\) at \((l + 1)th\) iteration is

\[
\sum_{j=1}^{n} \tau_{k,j} y_{1j} y_{2j} ; \Pi^{(l)} \left[ -s + \frac{\sum_{1}^{2} \sum_{1}^{2} y_{1j} y_{2j} ; \Pi^{(l)}}{\Sigma_{12}^2} \right] f_{x+1}(\delta_k y_{1j} y_{2j}) = 0,
\]

where \(f_{x}(y_{1j}, y_{2j}; \Pi^{(l)})\) is given by Eq.4.

**Step M:** The updated equations of \(\Sigma_{11}, \Sigma_{22}, \Sigma_{12}\) at \((l + 1)th\) iteration is

\[
\sum_{j=1}^{n} \tau_{k,j} y_{1j} y_{2j} ; \Pi^{(l)} \left[ -s + \frac{\sum_{1}^{2} \sum_{1}^{2} y_{1j} y_{2j} ; \Pi^{(l)}}{\Sigma_{12}^2} \right] f_{x+1}(\delta_k y_{1j} y_{2j}) = 0,
\]

where \(f_{x+1}(y_{1j}, y_{2j}; \Pi^{(l)})\) is given by Eq.4.

A closed-form solution does not exist for the \(\Sigma_{11}, \Sigma_{22}, \Sigma_{12}\), \(k = 1, 2, \ldots, K\), parameters. This is achieved by using a Newton-Raphson procedure initialized by estimator \((\Sigma_{11}, \Sigma_{22}, \Sigma_{12})\) defined in equations Eq.25, Eq.26 and Eq.27. The convergence of the Newton-Raphson procedure
is generally obtained after few iterations. The efficiency of the EM-algorithm in estimating the parameters is heavily dependent on the number of regions in the image. The number of mixture components initially taken for K-means algorithm is by plotting the histogram of the pixel intensities of the whole image. The number of peaks in the histogram can be taken as the initial value of the number of regions K. Usually the mixing parameter $\alpha_k$ and the region parameters $(\Sigma_{11,k}, \Sigma_{22,k}, \Sigma_{12,k}^2)$ are unknown. A commonly used method in initialization is by drawing a random sample in the entire image data. This method perform well only when the sample size is large, and the computation time is heavily increased. When the sample size is small it is likely that some small regions may not be sampled. To overcome this problem, we use K-means algorithm [16] to divide the whole image into various homogeneous regions. We obtain the initial estimates of the parameters $(\Sigma_{11,k}, \Sigma_{22,k}, \Sigma_{12,k}^2)$ for each image region using the method of moment estimators for BGD and for the parameters $\alpha_k$ as $\alpha_k = \frac{1}{K}$, for $k = 1, 2, \ldots, K$. Therefore the initial estimates of $(\Sigma_{11,k}, \Sigma_{22,k}, \Sigma_{12,k}^2)$ can be obtained by the MM presented in section II.C.

C. Application of the DMRD mixture in segmentation

After estimating the parameters of the DMRD mixture the prime step in image segmentation is allocating the pixels to the segments of the image. The image segmentation steps are the following:

Step 1) Plot the histogram of the whole image.

Step 2) Obtain the initial estimates of the model parameters using K-means algorithm and moment estimators as discussed in section IV.

Step 3) Obtain the refined estimates of the model parameters by using the EM-algorithm with the updated equations given in section II.A.

Step 4) Assign each pixel into the corresponding $k^{th}$ region (segment). That is,

$$j(y) = \arg\max_{1 \leq k \leq K} \left( \alpha_k f_k(y) \right).$$

where $j(y)$ represents the label of the class of the pixel $y = (y_1, y_2, ..)$.

IV. EXPERIMENTATION

To demonstrate the utility of the image segmentation algorithm developed in this paper, an experiment is conducted with four colors satellites images. A random sample of this images is taken the feature vector consisting of hue and saturation for each pixel of the each image is computed utilizing HSI color space. In HSI color space the hue and saturation values are computed from the values of RGB for each pixel in the image using the formula

$$\text{Hue} = H = \cos^{-1}\left( \frac{(R - G) + (R - B)}{2\sqrt{(R - G)^2 + (R - B)(G - B)}} \right), \quad B < G$$

$$\text{Saturation} = S = 1 - \frac{\min(R, G, B)}{I}, \quad \text{where} \quad I = \frac{R + G + B}{3}.$$

With the feature vector $(H, S)$ each image is modelled by using the two component mixture BGD. The number of segments in each of the four colors images considered for experimentation is determined by the histogram of pixel intensities. The histograms of the four images are shown in Fig.8.

The initial estimates of the number of regions $K$ in each image are obtained and given in Table 1.

### Table 1

<table>
<thead>
<tr>
<th>Image</th>
<th>sat 3</th>
<th>sat 4</th>
<th>sat 7</th>
<th>sat 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimate of $K$</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

After assigning these initial values of $K$ to each image data set, the K-means algorithm is performed. Using these initial parameters estimates $(\alpha_k, \mu_{1,k}, \mu_{2,k}, \sigma_{1,k}, \sigma_{2,k}, \rho_k)$ for a bivariate Gaussian mixture (BGAUM) and $(\alpha_k, \Sigma_{11,k}, \Sigma_{22,k}, \Sigma_{12,k}^2)$ for the bivariate Gamma mixture (BGM), where $k = 1, \ldots, K$, by the K-means algorithm to the EM algorithm. The computed values of the initial estimates and the final estimates of the two models parameters for each image are shown in tables II, III, IV and V.

Substituting the final estimates of the model parameters, the pdf of the feature vector of each image are estimated. Using the estimated pdfs and the image segmentation algorithm given in section III.B, the image segmentation is done for each of the four colors images under consideration. After developing the image segmentation method it is needed to verify the utility of segmentation in model building of the image for image retrieval. Using the estimated probability density function of the images under consideration the retrieved images are obtained.

The original, segmented and retrieved images are shown in Fig.9, Fig.10, Fig.11 and Fig.12 presented in Appendix A.

The performance evaluation of the retrieved image is done by subjective image quality (SIQ) testing or by objective image quality (OIQ) testing. The OIQ testing methods are often used since the numerical results of an objective measure are readily computed and allow a consistence comparison of
From the Table VI, it is observed that all the image quality metrics for the six images are meeting the standard criteria. This implies that using the proposed algorithm the images are retrieved accurately. A comparative study of proposed algorithm with that of algorithm based on BGM model and BGAUM model with K-means reveals that the MSE of the proposed model is less than the BGAUM. Based on all other quality metrics also it is observed that the performance of the proposed model in retrieving the images is better than the BGAUM.

V. CONCLUSION

In this paper, we proposed a segmentation algorithm adapted to color image by the use of the distribution of the diagonal of the modified Riesz distribution. Here, it is assumed that the color image characterized by HSI color space, in which HS values are non negative. They are characterized by bivariate Gamma mixture model which is a special case of the DMRD. The model parameter are estimated using the EM-algorithm. The initialization and the number of image segments are determined through $K$-means algorithm and moment method and presented in Table VI.

Different algorithms. There are SIQ measures available for performance evaluation of the Image Segmentation method. An extensive survey of quality measures is given by [8]. For the performance evaluation of the developed segmentation algorithm, we consider the IQM namely Mean Square Error (MSE), Signal to Noise Ratio (SNR) and Maximum Distance (MD), are computed for all the four images with respect to the developed method and earlier methods and presented in Table VI.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Initial by K-means</th>
<th>Final by EM Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>k1</td>
<td>k2</td>
<td>k3</td>
</tr>
<tr>
<td>BGM</td>
<td>s</td>
<td>9</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>1/3</td>
<td>1/3</td>
</tr>
<tr>
<td>$\Sigma_{11}$</td>
<td>0.17</td>
<td>0.035</td>
</tr>
<tr>
<td>$\Sigma_{12}$</td>
<td>0.19</td>
<td>0.017</td>
</tr>
<tr>
<td>$\Sigma_{22}$</td>
<td>0.003</td>
<td>0.0003</td>
</tr>
<tr>
<td>BGAUM</td>
<td>s</td>
<td>1</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>1/3</td>
<td>1/3</td>
</tr>
<tr>
<td>$\Sigma_{11}$</td>
<td>0.015</td>
<td>0.045</td>
</tr>
<tr>
<td>$\Sigma_{12}$</td>
<td>0.29</td>
<td>0.0442</td>
</tr>
<tr>
<td>$\Sigma_{22}$</td>
<td>0.0005</td>
<td>0.0018</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Images</th>
<th>Quality Metrics</th>
<th>BGM</th>
<th>MBGAUM</th>
</tr>
</thead>
<tbody>
<tr>
<td>sat 3</td>
<td>Mean Square Error</td>
<td>0.1101</td>
<td>0.5402</td>
</tr>
<tr>
<td></td>
<td>Signal to Noise Ratio</td>
<td>8.9124</td>
<td>7.1417</td>
</tr>
<tr>
<td></td>
<td>Maximum Distance</td>
<td>0.5499</td>
<td>0.8915</td>
</tr>
<tr>
<td>sat 4</td>
<td>Mean Square Error</td>
<td>0.5928</td>
<td>0.9516</td>
</tr>
<tr>
<td></td>
<td>Signal to Noise Ratio</td>
<td>5.2988</td>
<td>0.4964</td>
</tr>
<tr>
<td></td>
<td>Maximum Distance</td>
<td>0.6194</td>
<td>0.9313</td>
</tr>
<tr>
<td>sat 7</td>
<td>Mean Square Error</td>
<td>0.578</td>
<td>0.8238</td>
</tr>
<tr>
<td></td>
<td>Signal to Noise Ratio</td>
<td>5.4817</td>
<td>2.1839</td>
</tr>
<tr>
<td></td>
<td>Maximum Distance</td>
<td>0.7313</td>
<td>0.8779</td>
</tr>
<tr>
<td>sat 9</td>
<td>Mean Square Error</td>
<td>0.5339</td>
<td>0.6642</td>
</tr>
<tr>
<td></td>
<td>Signal to Noise Ratio</td>
<td>6.2761</td>
<td>4.0913</td>
</tr>
<tr>
<td></td>
<td>Maximum Distance</td>
<td>0.5837</td>
<td>0.7520</td>
</tr>
</tbody>
</table>
APPENDIX A

Fig. 9. Sat 3 segmented with MBGM and MBGAUM

Fig. 10. Sat 4 segmented with MBGM and MBGAUM

Fig. 11. Sat 7 segmented with MBGM and MBGAUM

Fig. 12. Sat 9 segmented with MBGM and MBGAUM
REFERENCES