ABSTRACT
The blind image separation has been widely investigated nowadays. As a result, many algorithms of feature extraction have been developed for direct application of such image structures. One example of this, the separation of mixed fingerprints found in a crime scene, in which a mixture of two or more fingerprints may be gathered, for identification, they must be separated. In this paper, we propose a new technique for multiple mixed images separation based on modified Weibull distribution. We use an efficient method based on genetic algorithm and maximum likelihood for estimating the parameters of such score functions. Also the accuracy of this proposed distribution is measured, and we compare the algorithmic performance using the efficient approach with some other previous distributions. The numerical results show that the proposed distribution is flexible and has efficient results.

General Terms
Image processing, image denosing, parametric distribution

Keywords
Source separation, blind image separation, FastICA, Maximum likelihood, Genetic algorithm, Modified Weibull distribution.

1. INTRODUCTION
Recently the blind source separation (BSS) had more attention because it is considered as an advanced image/signal processing technique that has many applications such as: image, speech sound, biomedicine, and communication [1–4]. BSS aims to recover source images/signals from a mixture with little prior information. Many BSS algorithms have been employed from various viewpoints, including mutual information minimization [5], maximum likelihood [6], principle component analysis (PCA) [7], non-Gaussianity [8], tensors [9], and neural networks [10–12]. Regarding to BSS, the separation and optimization methods are the most important roles. Separation step is used as the separability measurement, and optimization step is used to get the optimum solution for the objective function which we get from separation mechanism. In the independent component analysis (ICA) framework, the challenging problem is to accurately estimate the statistical model of the sources. Practical BSS scenarios use difficult source distributions and even situations where many sources with variant probability density functions (pdf) mixed together. In recent literature, many parametric density models have been made available towards this direction. Examples of such models, the generalized Gaussian density (GGD) [13], the generalized gamma density (GGD) [15], the generalized alfa-beta distribution (AB-divergences) [16], the Pearson family of distributions [17], and even the so-called extended generalized lambda distribution (EGLD) [18] which is an extended parameterizations of the aforementioned generalized lambda distribution (GLD) and generalized beta distribution (GBD) models [19], and even combinations and generalizations such as super and generalized Gaussian mixture model (GMM) [14]. In this paper, we propose the Modified Weibull distribution (MWD) which is a modification -or we can say a generalization - of Weibull distribution. We have evaluated the accuracy of our proposed MWD and compare the algorithmic performance using many different previous distributions. The numerical results, shows that the MWD gives good results comparing with many different cases. The rest of this paper is organized as follows: In section 2, we present the BSS model and the independent component analysis (ICA) technique. In section 3, we will discuss the MWD (propose distribution). In section 4, we will use maximum likelihood to estimate the parameters of MWD based on genetic algorithm. In section 5, we will present the numerical results. Finally, we will present the experimental results and the computational efficient performance of our proposed technique.

2. BLIND SOURCE SEPARATION
Imagine a room where a musical band is playing, or some people are speaking with each other [20]. For simplifying suppose we have just two people are speaking with each other, and there are two microphones hold in different locations. We will get two recorded time signals, which we could denote by \( x_1(t) \) and \( x_2(t) \), with \( x_1 \) and \( x_2 \) the amplitudes, and \( t \) the time index. We find that each of the recorded signals is a weighted sum of the speech signals emitted by the two speakers, which we denote by \( s_1(t) \) and \( s_2(t) \). We could express this as a linear equation:

\[
\begin{align*}
    x_1(t) &= a_{11}s_1 + a_{12}s_2 \\
    x_2(t) &= a_{21}s_1 + a_{22}s_2
\end{align*}
\]

(1)

(2)

The \( a_{ij} \) are constant coefficients that give the mixing weights and they depend on the distances between the microphones and the speakers. They are assumed unknown, since we cannot know their values \( a_{ij} \), without knowing all the properties of the physical mixing system. The source signals \( s_i \) are unknown as well, since the major problem is that we cannot record them directly. What we want is to find the original signals from the mixtures, this is called the blind source separation problem. Blind means that we know very little information about the original sources. We can assume that the mixing coefficients \( a_{ij} \) are different enough to make the matrix which they form to be invertible. Thus, there exists a matrix \( W \) with coefficients \( w_{ij} \), such that:

\[
\begin{align*}
    s_1(t) &= w_{11}x_1 + w_{12}x_2 \\
    s_2(t) &= w_{21}x_1 + w_{22}x_2
\end{align*}
\]

(3)

(4)

Such a matrix \( W \) can be found as the inverse of the matrix that consists of the mixing coefficients \( a_{ij} \), one approach for solving this problem [20] would be to use some information on the statistical properties of the signals \( s_i \) to estimate the parameters \( a_{ij} \). Actually, it turns out that it will be enough to
assume that \( s_1(t) \) and \( s_2(t) \), at each time instant \( t \), are statistically independent.

3. INDEPENDENT COMPONENT ANALYSIS

Independent component analysis (ICA) is a method for finding underlying factors or components from multivariate (multi-dimensional) statistical data. What distinguishes ICA from other methods is that it looks for components that are both statistically independent, and non-Gaussian [25]. Now, assume that we observe \( n \) linear mixtures \( x_1 \ldots x_n \) of \( n \) independent components [20]

\[
x_j = a_{j1}s_1 + a_{j2}s_2 + \cdots + a_{jn}s_n \quad \text{for all } j
\]

The time index \( t \) has been dropped; in the ICA model [20], [25], it is assumed that each mixture \( x \) and each independent component is a random variable, instead of a proper time signal. The observed values \( x_j(t) \), e.g., the microphone signals, are then a sample of this random variable. As a preprocessing to simplify the calculation, we can assume that both the mixture variables and the independent components have zero mean. If not, then the observed variables \( x \) can always be centered by subtracting the sample mean, this makes the model zero-mean. It would be convenient to use a vector-matrix notation instead of the sums like in the previous equation. Let’s denote by \( x \) the random vector whose elements are the mixtures \( x_1 \ldots x_n \), and by \( s \) the random vector with element \( s_1 \ldots s_n \), and by \( A \) the matrix with elements \( a_{ij} \). The above mixing model can be written as

\[
x = As
\]

Also the model can be written as

\[
x = \sum_{i=1}^{n} a_is_i
\]

The statistical model in Eq. 6 is called independent component analysis model, or ICA model.

The ICA model is a generative model, this means, it describes how the observed data are generated by a process of mixing the components \( s_i \).

Starting point for ICA is very simple, assume that the components \( s_i \) are statistically independent. Also, they must have non-Gaussian distributions.

Why Gaussian variables are forbidden?

Assume that the mixing matrix is orthogonal and the \( s_i \) are Gaussian [25]. Then \( x_1 \) and \( x_2 \) are Gaussian, uncorrelated, and of unit variance, their joint density is given by

\[
p(x_1, x_2) = \frac{1}{2} \exp \left( -\frac{x_1^2 + x_2^2}{2} \right)
\]

The distribution of this function is completely symmetric. Therefore, it does not contain any information on the directions of the columns of the mixing matrix \( A \). This is why \( A \) cannot be estimated.

Measures of nongaussianity

To use non-gaussianity in ICA estimation [22], we must have a quantitative measure of non-gaussianity of a random variable, say \( y \). Simply, let us assume that \( y \) is centered (zero-mean) and has variance equal to one, in fact, one of the preprocessing in ICA algorithm is to make this simplification possible. There are many measures of non-gaussianity, the classical measure of non-gaussianity is kurtosis or the fourth-order cumulate, and another very important measure of non-gaussianity is Negentropy which based on the information theoretic quantity of (differential) entropy. Also, Minimization of Mutual Information, the most popular approach for estimating the ICA model is maximum likelihood estimation.

Preprocessing for ICA

It is usually very useful to do some preprocessing before applying an ICA algorithm on the data [25]:

Centering \( x \), that means making \( x \) a zero-mean variable by subtracting its mean, and this preprocessing is the most basic and necessary preprocessing, this is just for simplifying the ICA algorithm and does not mean that the mean could not be estimated. After estimating \( A \) -the mixing matrix- with centered data, the estimation can be completed by adding the mean vector of \( s \) back to the centered estimates of \( s \) [25].

Whitening, after centering and before the application of the ICA algorithm, we transform the observed vector \( x \) linearly to obtain a new vector \( \hat{x} \) which is white, i.e. its components are uncorrelated and their variances equal unity [20], [25].

The FastICA Algorithm

We introduced different measures of nongaussianity [21] i.e. objective functions for ICA estimation. In practice, also we need an algorithm for maximizing the contrast function, one of the most efficient algorithms of the ICA is the FastICA Algorithm, and this is what we will use in our new proposed method.

![FastICA flow chart](image)
W must be determined. Generally, the majority of BSS approaches perform ICA, by essentially optimizing the negative log-likelihood (objective) function with respect to the un-mixing matrix W such that
\[
L(u,W) = \sum_{i=1}^{N} E[\log p_{ai}(u_i)] - \log |\text{det}(W)|.
\] (11)

Where \(E[.\)] represents the expectation operator and \(p_{ai}(u_i)\) is the model for the marginal pdf of \(u_i\) for all \(l = 1,2,\ldots,N\). In effect, when correctly hypothesizing upon the distribution of the sources, the maximum likelihood (ML) principle leads to estimating functions, which in fact are the score functions of the sources [25]
\[
\varphi_l(u_i) = -\frac{d}{du_i} \log p_{ai}(u_i)
\] (12)

In principle, the separation criterion in (11) can be optimized by any suitable ICA algorithm where contrasts are utilized (see, e.g., [2]). The FastICA [3], based on
\[
W_{k+1} = W_k + D(E[\varphi(u)u^T]) - \text{diag}(E[\varphi(u)u_i]u_i)W_k,
\] (13)
where, as defined in [4],
\[
D = \text{diag}\left(\frac{1}{E[\varphi(u)u_i] - E[\varphi_i(u)]}\right),
\] (14)
where \(\varphi(t) = [\varphi_1(u_1), \varphi_2(u_2), \ldots, \varphi_n(u_n)]^T\), valid for all \(l = 1,2,\ldots,n\). In the following section, we propose MWD for image modeling.

5. MODIFIED WIEBULL DISTRIBUTION
Following [22] MWD is a new generalization of the two parameters Weibull distribution. The pdf of MWD is defined as:
\[
f(x) = (\alpha + \beta x^\gamma)^{\alpha -1} \times \exp\{-\alpha x - \beta x^\gamma\},
\] (15)
where \(x > 0\) cumulative distribution function of MWD is given by:
\[
F(x) = 1 - \exp\{-\alpha x - \beta x^\gamma\}, x \geq 0,
\] (16)
where \(\gamma > 0, \alpha, \beta \geq 0\) such that \(\alpha + \beta > 0\). It is clear that the MWD is very flexible. This is so since there are many other distributions that can be considered as special cases of MWD, by selecting the appropriate values of the parameters. These special cases include four distributions as shown in Table (I). In Figures (2-4) there are the distributions generated from MWD by changing the parameters.

6. PARAMETER ESTIMATION
To estimate the parameters of MWD, the maximum likelihood is used. Let \(X_1, X_2, \ldots, X_n\) be a sample of size \(N\) from an MWD. Then the log-likelihood function \(\mathcal{L}\) is given by:
\[
\ell = \prod_{i=1}^{n} f_i(x) = \prod_{i=1}^{n} [(\alpha + \beta x_i^\gamma)^{\alpha -1} \times \exp\{-\alpha x_i - \beta x_i^\gamma\}]
\] (17)

Hence, the log likelihood function \(\mathcal{L} = \ln \ell\) becomes
\[
\mathcal{L} = \log \ell = \log \left( \prod_{i=1}^{n} (\alpha + \beta x_i^\gamma)^{\alpha -1} \times \exp\{-x_i - \beta x_i^\gamma\} \right)
\] (18)
\[
\mathcal{L} = \sum_{i=1}^{n} \log ((\alpha + \beta x_i^\gamma)^{\alpha -1} \times \exp\{-x_i - \beta x_i^\gamma\})
\] (19)
\[ L = \log \ell = \sum_{i=1}^{n} \log(\alpha + \beta x_i r^{-1}) \] 

\[ + \sum_{i=1}^{n} \log(\exp(-\alpha x_i - \beta x_i r^{-1})) \] 

(20)

\[ L = \log \ell = \sum_{i=1}^{n} \log(\alpha + \beta x_i r^{-1}) + \sum_{i=1}^{n} (-\alpha x_i - \beta x_i r^{-1}) \] 

(21)

Therefore, maximum likelihood estimation of \( \alpha, \beta \) and \( \gamma \) are derived from the derivatives of \( L \). They should satisfy the following equations:

\[ \frac{\partial L}{\partial \alpha} = 0 \quad \frac{\partial L}{\partial \beta} = 0 \quad \frac{\partial L}{\partial \gamma} = 0 \] 

(22)

\[ \frac{\partial L}{\partial \beta} = \sum_{i=1}^{n} \frac{\alpha + \beta x_i r^{-1} - \alpha x_i}{} - \sum_{i=1}^{n} x_i r^{1} \] 

(23)

\[ \frac{\partial L}{\partial \gamma} = \sum_{i=1}^{n} \frac{\alpha + \beta x_i r^{-1}}{} - \sum_{i=1}^{n} \beta x_i r^{-1} \log x_i \] 

(24)

To estimate the value of parameters, the system of equations (22-24) must be solved. However, it is difficult to solve this system so, the genetic algorithm (GA) [23-24] will be used as an alternative numerical method to estimate the parameters. The GA optimization technique lies in the fact that it can minimize the negative of the log-likelihood objective function in (11), essentially without depending on any derivative information.

7. NUMERICAL RESULTS

Numerical experiments show that the GA method converges to an acceptably accurate solution with substantially fewer function evaluations. We have generated random number from MWD with parameters \( \alpha, \beta \) and \( \gamma \). By performing GA, we obtained the best estimation of parameters as in table (II).

### Table 1: Parameter estimation by using GA

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>( \beta )</th>
<th>( \gamma )</th>
<th>( \hat{\alpha} )</th>
<th>( \hat{\beta} )</th>
<th>( \hat{\gamma} )</th>
<th>Err</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>2.97</td>
<td>4.11</td>
<td>1.86</td>
</tr>
<tr>
<td>X2</td>
<td>5.2</td>
<td>6.8</td>
<td>2.5</td>
<td>5.27</td>
<td>6.80</td>
<td>2.42</td>
</tr>
<tr>
<td>X3</td>
<td>1.9</td>
<td>8.2</td>
<td>5.7</td>
<td>1.98</td>
<td>8.12</td>
<td>563</td>
</tr>
</tbody>
</table>

We resolve to FastICA algorithm for blind image separation. This algorithm depends on the estimated parameters and an unmixing matrix \( W \) which estimated by FastICA algorithm. By substituting (22) into (19) for the source estimates \( u_j \), it becomes clear that the proposed score function inherits a generalized parametric structure, which can be attributed to the highly flexible MWD parent model. So, a simple calculus yields the flexible BSS score function

\[ \phi_1(u) = -\frac{d}{dx} \log(\alpha + \beta x^\gamma(\gamma - 1)) \times \exp(\alpha x - \beta x^\gamma) \] 

(25)

In principle \( \phi_1(u) \) is capable of modeling a large number of signals as well as various other types of challenging heavy- and light-tailed distributions. Experiments were done to investigate the performance of our method through three applications (two in source separation and one in image denoising) when impulsive noise is presented. In all experiments, the performance of our method is compared with generalized gamma [26], tanh, skew, pow3 [25], and Gauss [15]. Our performance is measured by the peak-signal-to-noise ratio (PSNR), defined as:

\[ PSNR = 20\log_{10} \left( \frac{255}{MSE} \right) \] 

(26)

### Table 2 Image separation PSNR

<table>
<thead>
<tr>
<th>Distribution</th>
<th>First Image</th>
<th>Second Image</th>
<th>Third Image</th>
<th>Forth Image</th>
<th>Elapsed (in seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
<td>PSNR</td>
<td>MSE</td>
<td>PSNR</td>
<td>MSE</td>
<td>PSNR</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1208</td>
<td>57.0987</td>
<td>0.1113</td>
<td>57.6640</td>
<td>0.1907</td>
</tr>
<tr>
<td>Gauss</td>
<td>0.0016</td>
<td>76.0444</td>
<td>0.0085</td>
<td>68.8500</td>
<td>0.0454</td>
</tr>
<tr>
<td>Tanh</td>
<td>0.0016</td>
<td>76.0275</td>
<td>0.0086</td>
<td>68.8042</td>
<td>0.0551</td>
</tr>
<tr>
<td>Generalized</td>
<td>0.0053</td>
<td>70.9873</td>
<td>0.0055</td>
<td>68.7757</td>
<td>0.0909</td>
</tr>
<tr>
<td>Pow3</td>
<td>0.2731</td>
<td>53.7621</td>
<td>0.0099</td>
<td>69.7402</td>
<td>0.0949</td>
</tr>
<tr>
<td>Modified</td>
<td>0.0012</td>
<td>71.8994</td>
<td>0.0053</td>
<td>70.8845</td>
<td>0.0800</td>
</tr>
<tr>
<td>Weibull</td>
<td>0.5904</td>
<td>71.4094</td>
<td>0.0053</td>
<td>70.9781</td>
<td>0.0800</td>
</tr>
</tbody>
</table>

In this example, we illustrate the performance of our algorithm to denoise medical images taken from [24]. Where Figure (4-10) show the original images, noised images, and denoised images by different algorithms. After applying algorithms of skew, Gauss, tanh, generalized gamma, pow3 and, our algorithm MWD, the results are illustrated in Figure (8-13), also Table (IV) illustrates the performance of these algorithms. From table (IV) and Figure (4-10), the MWD is higher performance than other algorithms.
Fig. 5. Medical image denoising using Gauss filter: A, D are the source images, B, E are the noised images, C, F are the denoised images.

Fig. 6. Medical image denoising using Skew filter: G, J are the source images, H, K are the noised images, I, L are the denoised images.

Fig. 7. Medical image denoising using Pow3 filter: M, P are the source images, N, Q are the noised images, O, R are the denoised images.

Fig. 8. Medical image denoising using Tanh filter: A, D are the source images, B, E are the noised images, C, F are the denoised images.

Fig. 9. Medical image denoising using generalized gamma filter: G, J are the source images, H, K are the noised images, I, L are the denoised images.

Fig. 10. Medical image denoising using MWD filter: M, P are the source images, N, Q are the noised images, O, R are the denoised images.
8. CONCLUSION

In this paper, we introduced a new technique for image denoising and blind image separation based on exponentiated transmuted Weibull distribution. Our proposed technique outperforms existing solutions in terms of separation quality and computational cost. Using the GA to estimate the parameters of MWD gives small error. Also, the results of MWD are better than other algorithms.

9. REFERENCES


[26] Internet web: http://sipi.usc.edu/database/database.cgi

[27] Internet, web:http://www.lupusimages.com