PROFESSIONAL ENRICHMENTS ON INDEPENDENT COMPONENT ANALYSIS FOR BLIND SOURCE SEPARATION

A PARIMALAGANDHI
KIT-Kalaignar Karunanidhi Institute of Technology, Coimbatore, Tamil Nadu, India
parimalagandhi@yahoo.com

S VIJAYAN
Surya Engineering College, Erode, Tamil Nadu, India.
vijayansurya@gmail.com

Abstract: *Independent Component Analysis (ICA) and its variants were proposed for Blind Source Separation (BSS) wherein most of the algorithms assumed that the sources are independent, non-negative and well-grounded. Some variants of ICA made the independence assumption unnecessary by utilizing information from theoretically based metrics; however, these methods are very slow in the process. In this paper an efficient Clustering of Mutual Information based least dependent Component Analysis (CMILCA) is proposed to cluster based least dependent components in a computationally efficient way by utilizing a squared-loss variant of mutual information. However, CMILCA provides better results just for a less number of sources and observations. In order to overcome this, Clustering in Conjugate Gradients of Mutual Information based least dependent Component Analysis in Riemannian manifold (CCGMILCA) is proposed to achieve convergence faster. The Riemannian directional derivative is used for local minimum efficiency, which results in the estimation of a weight matrix by giving independent components corresponding to various sources of mixture signals. The results of our experiments show that proposed algorithms take less time and high signal to noise ratio than the existing Blind Source Separation techniques.*

Key words: *Blind Source Separation (BSS), Independent Component Analysis (ICA), K-Nearest Neighbor (KNN), Mutual Information based Least-dependent Component Analysis (MILCA), Convergence gradient method*

1. Introduction

Blind Source Separation (BSS) [1] is the method for separating a set of source signals from combined signals while not knowing any information concerning source signals or the mixing methodology. In the past 20 years, BSS has developed speedily and has found wide application in wireless communication, microwave radar signal process, image signal process, speech signal process, medicine signal process, unstable wave detection, etc. BSS is a much preferred analytical topic in the signal processing field. Earlier, several techniques were provided for separating time based signals like audio. Nowadays, BSS is developed for multi-dimensional information like pictures and tensors which may not have time as one of the dimensions. In BSS, the mixtures are taken as linear mixtures of the sources with tiny mathematical noises, and also the task in BSS is reduced to solely approximating the source signals. Nonnegative matrix factorization (NMF) and independent component analysis (ICA) are two of the acquainting approaches for BSS. From the literature outlined, there is a tendency that once the quantity of sources increases, then the performances of ICA and its variants decline, and also the reconstructed images made by NMF algorithms are similar to the corresponding average images. The analyses of issues in ICA variants and proposing an approach to resolve these issues are the most significant objectives of this paper.

ICA locates the components which are statistically independent or as independent as possible which is an appropriate assumption for BSS [2][3]. Commonly, ICA algorithm is performed based on two principles such as measures of non-Gaussianity of estimated components by utilizing some objective function and ICA estimation to find uncorrelated components [4]. FastICA and InfoMax [5] are well-liked approaches for minimizing mutual information and maximizing non-Gaussianity for generating independent components. In Non-negative Independent Component Analysis (NICA)[6], sources are assumed as independent, nonnegative and well-grounded. These assumptions make NICA possible to have a global convergence guarantee. Hence InfoMax, FastICA and NICA are not fortifying to decompose least independent mixed components. MILCA [7] gets least independent components from mixed components by measuring the independency of components using Mutual information rather than assuming the components is independent. However, the convergence of MILCA algorithm is very slow. The BSS algorithm has to be competent to decompose a large number of least independent mixed components with very less time.

Information-minimization clustering is acquired in an unsupervised manner so that the mutual information between independent components in different cluster assignments is minimized. However, Mutual information based clustering still involves non-convex optimization problems, and therefore finding a good local optimal solution is not

a simple practice. The square loss variant of Mutual Information SMI [8] is convex under moderate conditions and as a result, it improves the non-convexity of mutual knowledge regularization. The improvement of efficiency in the Clustering method of Mutual Information based least dependent Component Analysis (CMILCA) using SMI establishes an overall component dependent generalization error bound. The clustering effectively upgrades in MILCA lowering the processing time and develops the quality of BSS. However, the efficiency of CMILCA is affected when the number of sources and observations are increased.

Furthermore, the convergence issue due to a number of sources and observations in CMILCA is solved by utilizing a Riemannian conjugate gradient algorithm. The independent components of tangent space vectors are considered as Riemannian manifold, that is, to a smooth space. The minimization of mutual information between components is improved by the derivative of Riemannian gradient manifolds. A non-linear Conjugate Gradients (CG) method in a Riemannian manifold provides an alternative search direction to solve minimization of objective efficiently.

The remaining part of this paper is organized as follows: Section 2 is concerning literature review, Section 3 explains the concepts about the existing technique for blind source separation. Section 4 presents completely on the proposed methods and newly framed algorithms and Section 5 is regarding the performance evaluation outcomes of the proposed method in comparison with other existing methods. Finally, Section 6 concludes with the future research work.

2. Literature Review

Zibulevsky, M., et al [4] presented multi-node sparse representation for blind source separation technique. In their paper the main issue of blind source separation is regarded from a set of instantaneous linear combinations, with the unknown mixing matrix. The proposed system in their paper improves the separation through utilization of the structure of signals where a number of subsets of the wavelet packet coefficients have better sparsity and separability than others. The best subset consistent with global distortion is chosen to restore the mixing matrix with the aid of the Infomax algorithm or clustering.

Plumbley, M. D., [24] proposed algorithms for non-negative independent component analysis. The nonnegative ICA approach illustrated in their paper depends vigorously on the protection of the zero estimation of the information and source signals. If a bias has been additional throughout the information generation method, or the mean value has been subtracted throughout the observation method, this might cause failure of the algorithmic program, with either “slack” within the resolution (a set of potential rotations all with nonnegative reconstruction error), or no rotation abilities to produce zero reconstruction error. One potential approach to beat this drawback may well be incorporated with an adaptable offset within the nonnegative ICA analysis, modifying the algorithmic program to optimize a weighted sum of the initial error. The mean square error of ICA is minimized by using the axis pair rotation and geodesic algorithm. Additionally, the various geometry concerns are derived via geodesic search technique.

Vigliano, D., & Uncini, A., [31] presented a versatile ICA resolution for nonlinear blind source separation. In their paper, a novel learning method and structure is projected for decisive non-linear problem in blind source separation. For this reason, Combination of knowledge theory and adaptive neural network are taken into account. This mixing model is assumed as post nonlinear block followed by convolutive mixing channel. The flexibleness of this methodology is generated by means of spline-SG neurons achieving an online computation for score functions.

Szu, H., & Kopriva, I., [30] considered a deterministic method to blind source separation of the space variant imaging system. In their paper, the positivity constraints in blind source separation are recovered through the probabilistic method by considering independence among sources with utilization of the entire pixel information. In addition, this method is accomplished for resolving the space variant issues by means of reducing the second law of thermodynamics based contrast function. However, the estimated error is high. Bronstein, A. M., et al [4] provided the sparse ICA method for achieving the blind separation of broadcast and reflected images. In their paper, the sparse ICA method is extended for concerning the issue of separating the image not including any prior information regarding its structure. The modern approaches in sparse ICA are investigated. This technique mostly utilized different combinations of source images which is complicated.

Zhu, X. L., et al [35] proposed adaptive nonlinear PCA procedures for blind source separation without including pre-whitening. In their paper, modified nonlinear principal component analysis is described. In addition, least mean square algorithm and recursive least squares algorithm are presented. This method can achieve online blind source separation by means of directing the un-whitened annotations. However, this technique is not useful for whitened signals. Saruwatari, H., et al [27] proposed fast convergence process based on the combination of ICA and beam forming for blind source separation.

In their paper, frequency domain ICA and null beam forming including direction-of-arrival (DOA) evaluation are proposed. Also, the combination of ICA and null beam forming with DOA is explained based on the diversity algorithm to improve the convergence. This method is used for rapid and huge in conventional K-means clustering method. This weighted K-means clustering method provides rapid and precise convergence optimization which is offline. Qingming, Y., [25] proposed weighted K-means clustering for blind source separation. In their paper, the weighted K-means clustering is developed to avoid the accuracy problem estimations which are not suitable for noisy images.

Chan, T. H., et al [5] described convex analysis structure for blind source separation of non-negative source signals. The convex analysis structure is deterministic which requires no independent source postulation. This convex framework is developed based on the local dominance and various standard statements. Extreme-point finding method is explained for satisfying convex analysis of mixtures of non-negative sources derived by using either linear programming or simplex geometry. This method has high effectiveness to solve the optimization problem, but the degradation in accuracy is high. Ozerov, A., & Fevotte, C., [22] described multi-channel nonnegative matrix factorization in convolutive combinations for audio source separation. In their paper, the estimation of the mixing and source factors is addressed based on two methods. The first method called expectation maximization method is proposed for increasing the precise mutual likelihood of the multi-channel information. The second method, named multiplicative method from NMF technique is proposed to increase the amount of individual likelihoods of entire channels. However, signal-to-distortion ratio for this method is high.

Hsieh, H. L., & Chien, J. T., [9] investigated the non-negative matrix factorization for independent component analysis, which is used for blind source separation. In this method, the sources are converted through their cumulative distribution functions and nonparametric quantization is achieved for generating the nonnegative matrix. This method is used to separate music and speech signals effectively. Mustafi, A., & Ghorai, S. K., [20] proposed blind source separation method by means of fractional Fourier transform for de-noising medical images. Their method gave improved visual transparency to experts for utilizing the images during diagnosis. This method is also used for eliminating the most important component of additive or multiplicative noise from obtaining the image. De-noising the images without the knowledge about noise PDF is performed by using genetic algorithm and Gaussian filtering method. This method has better performance but the observation in time domain is very complex. Chien, J. T., & Hsieh, H. L., [6] described convex divergence independent component analysis for blind source separation. In their paper, convex divergence is achieved through applying the convex functions to Jensen's inequality. Here, the convex divergence ICA is generated and nonparametric C-ICA procedure is derived, including different convexity factors wherein the non-Gaussian source signals are categorized by means of the Parzen window based allocation. The speed of convergence is increased through the scaled natural gradient process. However, the degradation accuracy was more.

Bobin, J., et al [3] presented about sparsity and Adaptivity for blind source separation of partially correlated sources. In their paper, sparsity enforcing blind source separation technique coined adaptive morphological component analysis was proposed for retrieving sparse and partially correlated sources. This method provided an adaptive re-weighting technique for favouring samples according to their rank of correlation. Their method was highly robust to the inequitable correlation of sources, but the signal-to-distortion ratio for this method was high. Benahmed, A., et al [2] presented the estimation of direction of arrival (DOA) based blind source separation for smart antenna. This method is used to direct the antenna radiation pattern towards the required radio frequency emitter. The different antenna patterns are observed and separated by using discrete Fourier transform model. ICA based separation is provided to extract the independent components and mixing matrix. This method will not be effective in the presence of co-channel interference. Meganem, I., et al [19] proposed a linear quadratic method for blind source separation based on the nonnegative matrix factorization. In their paper, the spectral unmixing for urban hyper-spectral images is specifically concerned. Different linear quadratic models based on NMF are developed and the performances are evaluated. However, convergence problems were occurring. Hattay, J., et al [11] described non-negative matrix factorization based on wavelet transform for blind source separation. This method utilized an adaptive quincunx lifting scheme, according to the wavelet decomposition for pre-processing the input information followed by NMF. The unmixed images are restructured by means of inverse adaptive quincunx lifting scheme. However, peak signal-to-noise ratio of this method was high.

Wang, X., et al [33] provided fast nonlinear principal component analysis technique which is utilized for blind source separation. This new method is explored by means of mixing the most favorable step size including the most favorable energy factors which are obtained through the decrement of the cost function of nonlinear principal

component analysis method. The convergence speed of this method is high, but the distance index is low. Guo, X., et al [8] proposed linear predictive coding error clustering criterion for nonnegative matrix factorization based blind source separation. In their paper, the clustering algorithm is developed by mixing the frequency and time activation of NMF to improve the accuracy of clustering. This improved clustering algorithm was based on the linear predictive coding error and the factors of NMF. However, this method may fail for percussive sources.

3. Existing Methods.

An Efficient Clustering Of Mutual Information Based Least Dependent Component Analysis (MILCA)

In MILCA method, K-nearest neighbor technique is used to estimate dependencies among randomly generated independent components from mixed signals. A hierarchical clustering algorithm is used to cluster the output based on component interdependencies.

Let X and Y are continuous random variables with joint density f(x, y).

Marginal densities fx (x) = $\int dy\, f(x,y)$ and fy (y) = $\int dx\, f(x,y)$

MI is I(X, Y) = $\iint dx\, dy f(x,y) \log \frac{f(x,y)}{f_x(x) f_y(y)}$ (1)

Mutual information is measured in terms of entropies. The entropies are represented as follows:

$H(X) = -\int dx\, f_x(x) \log f_x(x)$ (2)

$H(Y) = -\int dy\, f_y(y) \log f_y(y)$ (3)

$H(X,Y) = \iint dx\, dy f(x,y) \log f(x,y)$ (4)

It can be written as

$H(X,Y) = H(X)+H(Y)- H(X,Y)$ (5)

A pre-whitening using principle component transformation and rescaling procedures is applied on mixture signals to make the covariance matrix as isotropic. The minimization of mutual information contrast function in MILCA is done by a pure rotation function. For any number M of random variables, the MI is defined as

$I(X1,X2,X3,\ldots XM) = \sum_{m=1}^{M} H(X_M) - H(X_1, X_2, X_3, \ldots X_M)$ (6)

The grouping property of M dimensional MI is defined as

$I(X,Y,Z) = I((X,Y),Z)+I(X,Y)$ (7)

Here, I ((X, Y),Z) is the MI between the two variables Z and (X, Y).

For any set of random variables and any hierarchical clustering of this set into disjoint groups, the complete MI may also be hierarchically decomposed into MIs between groups and MIs inside each group.

The Kozachenko-Leonenko estimate for Shannon entropy,

$\hat{H}(X) = -\psi(k) + \psi(N) + \log c_d + \frac{d}{N}\sum_{i=1}^{N} \log \varepsilon(i)$ (8)

Where $\psi(x)$ is the digamma function, $\varepsilon(i)$ is twice the distance from xi to its k-th neighbor, d is the dimension of x and cd is the volume of the d-dimensional unit ball.

The estimate for MI is,

$I(X,Y) = -\psi(k) - \frac{1}{k} - \langle \psi(n_x) + \psi(n_y) \rangle + \psi(N)$ (9)

The MI estimation for m random variables is,

$I(X1,X2,\ldots Xm) = -\psi(k) - (m-1)\frac{1}{k} + (m-1)\psi(N) - \langle \psi(n_x) + \psi(n_y) + \cdots \psi(n_m) \rangle$ (10)

In a simple representation of ICA, one observes m random variables x1(t), x2(t), . . . ,xm(t) which are assumed to be linear mixtures of n unknown independent components $s_1(t), s_2(t), \ldots, s_n(t)$.

$s(t) = (s_1(t), s_2(t), \ldots, s_n(t))$ (11)

$x(t) = [x_1(t), x_2(t), \ldots, x_m(t)]^T$ (12)

$x(t) = As(t), t = 1, \ldots., T$ (13)

In Equation (13), A denotes the $n \times n$ non singular mixing matrix in which the quantity of sources is identical with the quantity of measured components. The decomposition into independent components by means of inverse transformation is defined as follows:

$\hat{s}(t) = W\, x(t)\ \ where\ W = A^{-1}$ (14)

The matrix W is decomposed into two elements as follows:

$W = RV$ (15)

In the above Equation (15), pre-whitening matrix V converts the covariance matrix into $C' = VCV^T = 1$ and R refers the pure rotation, which means orthogonal transformation for achieving non-Gaussianity. The aim of ICA is now to minimize I(X1 . . .Xm) under a pure rotation R. Any rotation R is denoted as follows:

$R = \prod_{i,j} R_{i,j}(\emptyset)$ (16)

$$R_{i,j}(\emptyset)(x_1 \dots x_i \dots x_j \dots x_n) = (x_1 \dots x_i' \dots x_j' \dots x_n) \quad (17)$$

$$x_i' = \cos\emptyset\, x_i + \sin\emptyset\, x_j \;\&\; x_j' = \sin\emptyset\, x_i + \cos\emptyset\, x_j \quad (18)$$

For this orthogonal rotation or orthogonal transformation Mutual Information (MI) is expressed as follows:

$$I(R_{i,j}(\emptyset)X) - I(X) = I(X_i', X_j') - I(X_i, X_j) \quad (19)$$

The optimal angle $\emptyset$ for equation (19) is computed using Equation

$$\hat{I}_{ij}(\emptyset) = \hat{I}(X_i', X_j') \quad (20)$$

The optimal angle $\emptyset$ which minimizes Equation (19) is found and estimates Si from X′.

However, MILCA approaches nonetheless contain non-convex optimization problems, and for this reason discovering a just right local optimum resolution isn't easy to observe. So alternative squared-loss variant mutual information [30] is utilized in CMILCA. Squared-loss mutual information based on marginal probability between two continuous random variables a and b for BSS is described as:

$$SMI(X,Y) = \frac{1}{2}\iint f(x)f(y)\left(\frac{f(x,y)}{f_x(x)f_y(y)} - 1\right)^2 dx\, dy \quad (21)$$

Equation (21) is the marginal density based square loss variant of mutual information.

The equation can be written as

$$SMI(X,Y) = \frac{1}{2}\iint f(x)\, f(y)\left(\frac{f(x,y)}{f_x(x)f_y(y)}\right)^2 dx\, dy - \frac{1}{2}\iint f(x)\, f(y)\frac{f(x,y)}{f_x(x)f_{by}(y)}\, dx\, dy + \frac{1}{2} \quad (22)$$

Equation (22) is again rewritten as,

$$SMI(X,Y) = \frac{1}{2}\iint f\left(\frac{y}{x}\right)\, f(x)\frac{f\left(\frac{y}{x}\right)}{f_x(x)f_y(y)}\, dx\, dy - \frac{1}{2} \quad (23)$$

Let us approximate the posterior probability $f\left(\frac{x}{y}\right)$ by the following kernel model.

$$f(y|x;\alpha) := \sum_{i=1}^{n} \alpha_{yi} K(x, x') \quad (24)$$

where $\alpha = (\alpha 1,1, \dots, \alpha c,ni)T$ s the parameter vector, $(\,)T$ denotes the transpose, and $K(x, x')$ denotes a kernel function [31] with a kernel parameter

$$K(x_i, x_j) = \begin{cases} \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma_i\sigma_j}\right) & \text{if } x_i \in N_t(x_j) \text{or } x_j \in N_t(x_i), \\ 0 & \text{Otherwise,} \end{cases} \quad (25)$$

where Nt(x) is the set of t nearest neighbors for xi which t is the kernel parameter, σi is the local scaling factor defined as,

σi= ‖ xi−x(t)i ‖ , and x(t)i is the tth nearest neighbor of xi.

The SMI approximation is as follows:

$$\widehat{SMI} := \frac{1}{2n}\sum_{y=1}^{c}\frac{1}{\pi_y}\alpha_y^T K^2 \alpha_y - \frac{1}{2}, \quad (26)$$

$$SMI(R_{i,j}(\emptyset)X) - SMI(X) = SMI(X_i', X_j') - SMI(X_i, X_j) \quad (27)$$

$$\widehat{SMI}_{ij}(\emptyset) = \widehat{SMI}\quad (X_i', X_j') \quad (28)$$

The optimal angle $\emptyset$ which minimizes Equation (27) is found and estimate si from X′.

4. Proposed Methods

Proposed Algorithm 1: Clustering of Mutual Information based least dependent Component Analysis (CMILCA)

Input: Mixture images or signals of n independent sources

Output: n independent sources

1. Denoising and whitening of input images or signals
2. For each pair of randomly generated variables find the angle $\emptyset$which minimizes
3. $\widehat{SMI}_{ij}(\emptyset) = \widehat{SMI}(X_i', X_j')$
4. Until $\widehat{SMI}(X_i', X_j')$ is not converged repeat Step 2.
5. $\hat{s}_i = X_i'$are the estimates for the sources

CMILCA algorithm gives better results than MILCA in terms of convergence rate. However the algorithm's performance is to be increased when the number of sources and observations are more. So the conjugate gradient on Riemannian manifold is further included in CMILCA which in the name of conjugate gradient based MILCA technique is discussed as follows:

Proposed Algorithm 2: Clustering in Conjugate Gradients of Mutual Information based least dependent Component Analysis (CCGMILCA)

The convergence of CMILCA is further improved for a larger number of sources and observations by using a conjugate gradient algorithm. Consider random vector z which refers to the whitened linear mixture independent component of sources

$(s_1(t), s_2(t), \ldots, s_n(t))$ and the permutation of sources is defined as $y = Wx$, where W is constrained to the square orthogonal matrix. The gradient algorithm with constrained convergence is obtained as follows:

Since W constructed from Equation (15) is orthogonal, the cost function is denoted as,

$$J(w_1, \ldots, w_n) = \sum_{i=1}^{n} E\{\min(0, y_i)^2\}, y_i w_i^T x \quad (29)$$

In the above Equation (29), w_i^T denotes the rows of matrix W.

Unlike classical ICA algorithms employing optimization methods on the Riemannian manifold which implies the orthogonality constraint on the independent components (ICs), the present approach relies on the unit-norm constraint on the ICs, i.e., the constraint surface is the Riemannian manifold.

In this paper, W in Equation (15) is obtained by solving Equation (27) with searching on minimal $\emptyset$ value. However, to increase convergence for obtaining optimal W, explicit derivative information, the discriminancy property of the contrast function is to be solved on manifold fields. Absil et al. [32] defined Riemannian manifold based derivatives to solve optimization problems.

Riemannian Manifolds are sets which can be manipulated with patches of $\mathbb{R}n$. These manipulations are called as charts. A set of compatible charts which covers the complete set is called as atlas for that particular set. The set and the atlas jointly constitute a manifold. Let $\mathcal{M}$ be a set. A chart of $\mathcal{M}$ is a pair (U,φ) where $U \in \mathcal{M}$ and φ is a bisection between U and an open set of Rn. Uis the chart's domain and n is the chart's dimension. Given $p \in U$, the elements of φ(p) = (x1,……,xn) are called the coordinates of p in the chart(U, φ).

A set $A = \{(U_i, \varphi_i), i \in I\}$ of pairwise smoothly compatible charts such that $\cup_{i\in I} U_i = \mathcal{M}$ is a smooth atlas of $\mathcal{M}$. A smooth manifold is a pair $\mathcal{M}$ = ($\mathcal{M}$,A+), where $\mathcal{M}$ refers a set and A+ is a maximal atlas of $\mathcal{M}$.

Tangent spaces for manifolds embedded in $\mathbb{R}n$. Let $\mathcal{M} \in \mathbb{R}n$ be a smooth manifold. The tangent space at $x \in \mathcal{M}$, noted Tx $\mathcal{M}$, is the linear subspace of $\mathbb{R}n$ defined by:

$$T_x \mathcal{M} = \left\{ \begin{matrix} v \in \mathbb{R}^n,\ v = c'(0) \text{for a smooth c:} \\ \mathbb{R} \to \mathcal{M} \ \text{such that } c(0) = x \end{matrix} \right\}$$

The dimension of Tx $\mathcal{M}$ is the dimension of a chart of $\mathcal{M}$ containing x.

The tangent space to $\mathcal{M}$at p, noted Tp $\mathcal{M}$, is the quotient space

$$Tp\ \mathcal{M} = Cp/\sim f[c] = \{[c]: c \in C_p\} \quad (30)$$

Given c ∈Cp, the equivalence class [c] is an element of Tp $\mathcal{M}$ called a tangent vector $\mathcal{M}$ at p.

The derivative of a scalar field f on $\mathcal{M}$ at $p \in \mathcal{M}$ in the direction ξ= [c] ∈ Tp $\mathcal{M}$ is the scalar

$$D\ f[p][\ \xi] := \frac{d}{dt} f(c(t))|_{t=0} = (f \circ c')\ (0) \quad (31)$$

The equivalence relation over Cp is specifically derived such that this definition does not rely on the selection of c, the representative of the equivalence class ξ. In the above notation, the brackets around ξ are a convenient way of denoting that ξ is the direction. They do not mean that some sort of equivalence class of ξ is being considered.

Riemannian manifold is a pair ($\mathcal{M}$, g), in which$\mathcal{M}$ refers to the smooth manifold and g refers to the Riemannian metric. A Riemannian metric is a smoothly varying inner product defined on the tangent spaces of $\mathcal{M}$, that is, for each $p \epsilon \mathcal{M}$, gp(.,.) = $\langle .,. \rangle_p$ is an inner product on Tp $\mathcal{M}$.

Let f be a scalar field on a Riemannian manifold $\mathcal{M}$. The gradient of f at p, denoted by grad f(p), is defined as theunique element of Tp $\mathcal{M}$ satisfying:

$$D\ f\ [p][\ \xi] = \langle \text{grad } f(p), \xi \rangle_p\ , \forall\ \xi \in T_p\ \mathcal{M} \quad (32)$$

Thus, grad f : $\mathcal{M} \to T\ \mathcal{M}$ is a vector field on M.

The gradient depends on the Riemannian metric but directional derivatives fail. For a scalar field f in a Euclidean space, grad f is the usual gradient, which is noted ∇f. Remarkably Similar to the Euclidean case, the gradient narrated above is the steepest-ascent vector field and the norm ‖ grad f(p) ‖p is the steepest slope of f at p. More precisely,

$$\| \text{grad } f(p) \|p\ = \max_{\xi \in Tp\ \mathcal{M}} D\ f\ [p][\ \xi] \quad (33)$$

A retraction in a manifold M is a smooth mapping R from the tangent bundle T$\mathcal{M}$ onto $\mathcal{M}$ with the following properties.

For all x in $\mathcal{M}$, let Rx denote the restriction of R to Tx $\mathcal{M}$. Then, Rx(0) = x, where 0 is the zero element of Tx $\mathcal{M}$, and The differential (DRx)0 : T0(Tx $\mathcal{M}$) ≡Tx$\mathcal{M}$ →Tx$\mathcal{M}$is the identity map on Tx $\mathcal{M}$, that is, (DRx)0 = Id (local rigidity). Vector transport in a manifold $\mathcal{M}$ is a smooth mapping

$$\text{Transp: } T\mathcal{M} \oplus T\mathcal{M} \longrightarrow T\mathcal{M}\text{: } (\eta,\xi\ )\ \text{Transp}(\eta \longmapsto \xi) \quad (34)$$

A candidate demixing matrix, Xk $\epsilon \mathcal{M}$, of order

D×D is constructed by normalizing the columns of a constructed matrix W, which can be written as

Xk =W× ddiag(WTW) -1/2 at iteration k = 0,1,..n. the contrast (objective) function value f and the conjugate gradient ∇f(Xk) are evaluated at Xk using (27). Since this algorithm evolves on the Riemannian feasible set $\mathcal{M}$, ∇f(X0) ought to be projected onto the tangent space,$T_{X_k}\mathcal{M}$ to obtain the respective Riemannian gradient

$$\text{gradf(Xk)} := \nabla f(Xk) - Xk\ \text{ddiag}\ (X_k^T \nabla f(Xk)) \quad (35)$$

$$\min_{x \in \mathbb{R}^n} f(x) \quad (36)$$

Such that f is continuously differentiable and Steepest descent (SD) or gradient descent method is arguably one of the unique and most well-known algorithms available. Given an initial guess or initial iterate x0 ∈ ℝn, it attempts to develop its predicament by following the most promising direction iteratively. Now elegantly, it generates a sequence of iterates:

x0, x1,……..∈ ℝn according to the update equation

$$x_{k+1} = x_k + \alpha_k d_k \quad (37)$$

A sophistication layer is added to this simple algorithm by constructing an alternative search direction dk which is a carefully derived linear combination of both ∇f(Xk) and the previous search direction d_{k-1} based on the non-linear conjugate gradients (CG) method. Thus incorporating a form of inertia in the search procedure is:

$$d_k = -\nabla f(X_k) + d_{k-1} \quad (38)$$

From the update Equation (37) and the search direction Equation (38), it is obvious that the CG method relies on the vector space structure of ℝn, by composing points and vectors using linear combinations. Though this dependency is not fundamental, where both equations can be altered so that they will make the optimization problems which are of the form (36) where the search space $\mathcal{M}$ is a Riemannian manifold. The Riemannian structure is not needed to have a notion of gradient.

The update Equation (37) produces xk+1, a new point on the search space, by moving away from xk along the direction $\alpha_k d_k$. The notion of retraction embodies this very same idea and suggests the more general update formula:

$$xk+1 = Rx_k\ (\alpha_k d_k) \quad (39)$$

where $d_k \in T_{x_k}\mathcal{M}$ is a tangent vector at xk+1. Similarly, the search direction Equation (38) produces the tangent vector d_k by combining two vectors:-grad f(xk+1) and d_{k-1}, where the former is the Riemannian gradient of f at x_k. Those are respectively, tangent vectors at x_k and x_{k-1}. As a result, they cannot be combined directly and do not belong to the same subspace. This issue can be solved by transportingd_{k-1} to $x_k$using a vector transport:

$$d_{k-1}^{+} = \text{Trans}\ p_{x_k \leftarrow x_{k-1}}\ (d_{k-1}) \quad (40)$$

The search direction equation then becomes:

$$d_k = -\ \text{grad}\ (f(x_k)) + \beta_{k-1} d_{k-1}^{+} \quad (41)$$

A standard trick to accelerate the CG algorithm is to precondition the iterations by operating a change of variables on the tangent spaces $x_k\ \mathcal{M}$ as mentioned by Hager.,et al [33]. This change of variable should be chosen so such that to reduce the condition number of the cost function.

Normally, this is achieved by change of variables closely related to the inverse of the Hessian nearby or at a critical point. Certainly a change of variables on $Tx_k\ \mathcal{M}$ amounts to a change of Riemannian metric g_{x_k}, so that it is theoretically adequate to describe a CG method on Riemannian manifolds without explicitly allowing for preconditioning. In practice though, it is convenient to separate the work of describing manifolds (giving them a Riemannian structure, defining retractions, geodesics, projectors, etc.) and that of describing a cost function. Since the pre-conditioner depends on the cost function, explicit preconditioning of the Riemannian CG method is allowed for, with the following pre-conditioner:

$$\text{Precon}\ f(x) : T_x\mathcal{M} \longrightarrow T_x\mathcal{M} \quad (42)$$

The linear operator Precon f(x) must be symmetric w.r.t. the Riemannianmetric, positive definite and, ideally be some kind of cheap approximation of (Hess f(x))-1. The search direction equation now reads:

$$d_k = -\text{Preconf}(x_k)\ [\text{grad}\ f(x_k)] + \beta_{k-1} d_{k-1}^{+} \quad (43)$$

Notice that if Precon f(x) = (Hess f(x))-1 and β_{k-1} = 0, this is a Newtonstep. When no pre-condition is available or necessary, it is replaced by the identity operator. The step size α_k is chosen by a line search algorithm which approximately solves the one-dimensional optimization problem:

$$\min_{\alpha>0} \phi(\alpha) := f(\ Rx_k(\alpha d_k\ ) \quad (44)$$

If d_k is a descent direction for f (which is typically enforced), then $\phi'(0)$<0and it is necessarily possible to decrease ϕ (and hence f) with a positive stepsize. It does not matter whether (44) is solved exactly or not. Typically, it is sufficient to compute a large enough step size such that a sufficient decrement is obtained according to the Armijo criterion:

$$f(x_{k+1}) = \phi(\alpha_k) \leq \phi(0) + c_{\text{decrease}}\ \alpha_k\ \phi'(0) = f(x_k)$$

$+ c_{decrease}$ .D $f(x_k)$ $[\alpha_{kd_k}]$ (45)

The constant 0 < cdecrease < 1 is the sufficient decrease parameter. The simple backtracking line search in the above methods guarantees that this condition is satisfied.

Proposed Algorithm 2:

//Conjugate gradient based MILCA method

1. Perform steps 1 to 3 in algorithm 1
2. If W is orthogonal matrix
3. Compute the minimum cost function by using equation (35 and 36)
4. Compute the gradient with respect to $f(x_k)$ and update W for eqn(29)
5. Until $\widehat{SMI}(X_i', X_j')$not converged step 2.
6. Estimate $\hat{s}_i = X_i'$ for the sources

5. Results and Performance Comparison

The performances of Fast ICA, NICA, MILCA and CCGMILCA are evaluated and compared by selecting audio files and images from McGill Calibrated Color Image Database [34] and Centre for speech technology research (CSTR) database [35]. Sources are selected from databases and mixtures (observations) are generated by mixing sources with the linear mixing model. To generate mixture matrix A, vectorized sources are appended into the column of matrix W. Nonnegative mixing matrix H is generated randomly with normalized columns. The result of multiplying W with H is mixture matrix. Hence H is a normalized column matrix, so that each column of A contains a linear mixing of all the sources.

Signal-to-Noise Ratio (SNR), Structural Similarity Index (SSIM), Universal Image Quality Index (UIQI), and Computational time are used to evaluate to prove the better performance of the proposed blind source separation techniques on image dataset. Signal-to-Noise Ratio (SNR), Signal-to-distortion ratio (SDR), signal-to-interference ratio (SIR), SAR (signal to artifacts ratio) and magnitude ignal-to-Error Ratio (mSER) are used to evaluate the strength of the proposed blind source separation on Audio dataset. The experimental works were performed using Matlab ® R2015b on an Intel Pentium i5 3.4 GHz processor with 8 GB memory.

Three source images are taken from McGill Calibrated Color Image Database as shown in the first row of Figure 1 (b). Figure 1 (a) shows four observed images generated by mixing three sources. The second, third ,fourth, fifth and sixth rows of Figure 1(b) illustrate that the recovered images of BSS techniques FastICA, NICA, MILCA, CMILCA and CCGMILCA respectively. The MSE (Mean Square Error) values between recovered images by BSS approaches and source images are compared; the MSE value for proposed two techniques CMILCA and CGMILCA are reduced considerably than the other classical BSS approaches.

$$MSE = 1/MN\left\|X - \hat{X}\right\|_F^2 \quad (46)$$

where X and $\hat{X}$ are recovered and original image respectively. M, N is the width and height of image.

Three audio files are selected from Centre for Speech Technology Research (CSTR) database. Five mixture signals are generated by the following equation:

$$x_{ij}(t) = a_{ij}s_j(t) \quad (47)$$

where sj(t) are source signals, aij positive mixing matrix which is generated as like image mixture generation which is described in the earlier part of the Result and Discussion chapter. Figure 1 (a) shows the generated six observed signals. The selected audio files from database are in .wav format.

Figure 1 (b) illustrates the recovered signals by using all BSS techniques FastICA, NICA, MILCA, CMILCA and CGMILCA. Spatial distortion and interference components are computed by least-squares projection of the separated source signals onto the corresponding source signal subspaces [36].

$$e_{ij}^{spat}(t) = P_j^L \hat{s}ij(t) - sij(t) \quad (48)$$

$$e_{ij}^{interf}(t) = P_{all}^L \hat{s}ij(t) - P_j^L sij(t) \quad (49)$$

$$e_{ij}^{artif}(t) = \hat{s}ij(t) - P_{all}^L sij(t) \quad (50)$$

where P_j^L is the least-squares projector onto the subspace spanned by skj (t-τ) ,1≤ k ≤ I, 0≤ τ ≤ L-1, and P_{all}^L is the least-squares projector onto the subspace spanned by skl (t-τ), 1≤ k ≤ I, 1≤ l ≤ j, 0≤ τ ≤ L-1. The filter length L was set to 512 (32 ms), which was the maximal tractable length. The performance metrics are derived by the Equations (48) to (50).

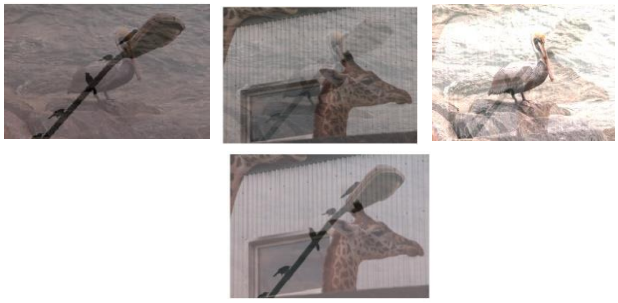
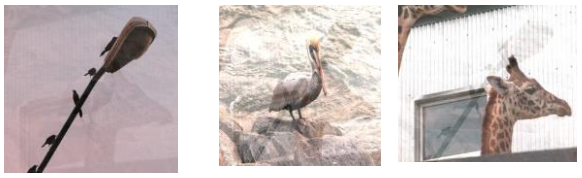


Fig. 1(a). Observed images

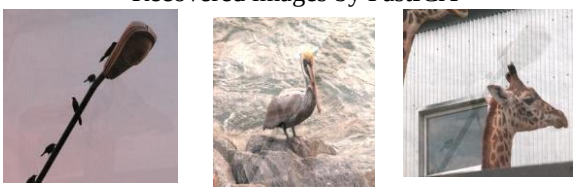
Source 1 Source 2 Source 3

Original Images



MSE=0.215 MSE=0.121 MSE=0.106

Recovered images by FastICA



MSE=0.104 MSE=0.112 MSE=0.102

Recovered images by NICA

MSE=0.085 MSE=0.099 MSE=0.0740

Recovered images by MILCA

MSE=0.083 MSE=0.071 MSE=0.091

Recovered images by CMILCA

MSE=0.071 MSE=0.051 MSE=0.088

Recovered images by CCGMILCA

Fig. 1(b). Original sources and separated sources using Fast ICA, NICA, MILCA, CMILCA and CCGMILCA

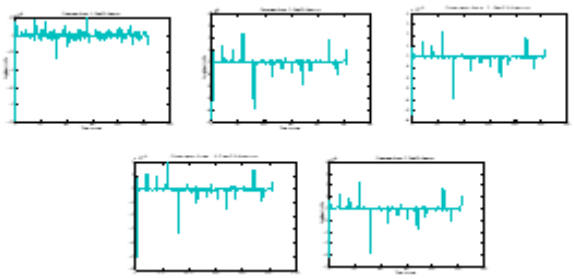


Fig. 2(a). Observed Audio Signals

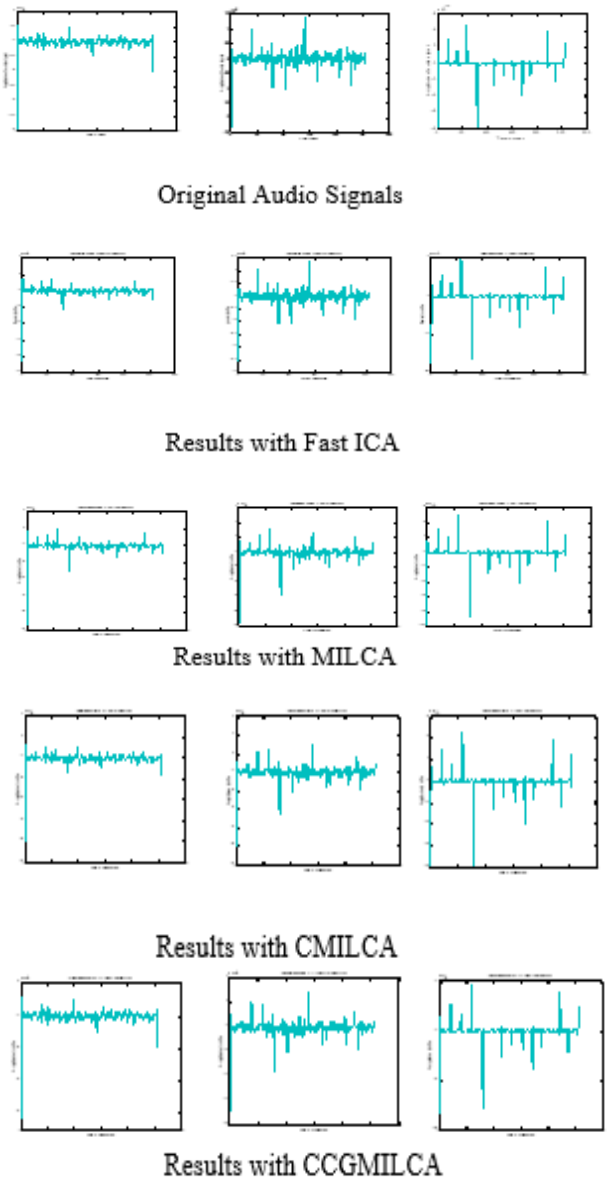


Fig. 2(b). Original sources and separated sources using Fast ICA, MILCA, CMILCA and CCGMILCA

5.1 Signal-to-Noise Ratio (SNR)

Signal-to-Noise Ratio (SNR) is given as,

$$\mathrm{SNR} = -10\log_{10}\frac{\|\mathrm{W}-\mathrm{B}\|_2^2}{\|\mathrm{W}\|_2^2} \quad (51)$$

In the above equation, every column of W includes the original image or audio signal and every column of B includes the related reconstructed image or audio signal. The SNR value for all methods for images is shown in Fig 3. SNR value of CCGMILCA method is higher than all other methods regardless of the number of sources. For a higher number of sources i.e. for 10, the SNR of FastICA is 30.20dB, NICA is 32.40 dB, MILCA 34dB and CMILCA and CCGMILCA are providing higher value of SNR as 36.2dB and 40dB respectively.

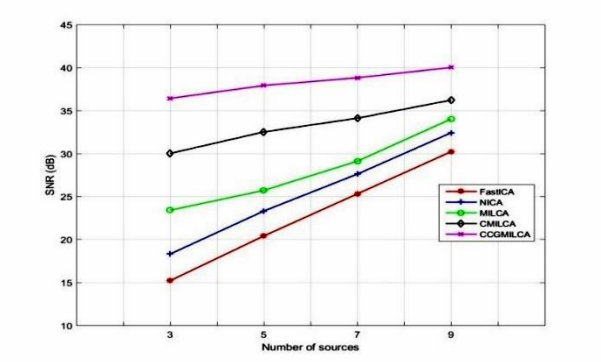


Fig. 3. Comparison of SNR value for image dataset (in dB)

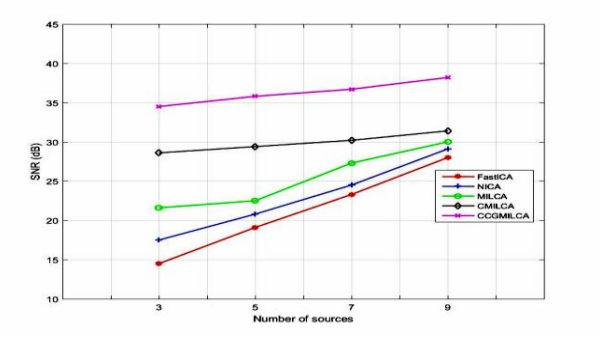


Fig. 4. Comparison of SNR value for Audio dataset (in dB)

Fig 4 shows the comparison of SNR for audio signals. The result shows that the SNR value of CCGMILCA method is higher than the other methods. While considering the number of source signal as 9, the SNR of FastICA is 27.8dB, NICA is 29.4dB, MILCA is 30dB and when CMILCA and CCGMILCA are applied the value of SNR is improved to 33dB and 38dB respectively.

5.2 Structural Similarity Index (SSIM)

It is defined as the similarity value between the original and recovered images s(x, y). It is given as follows:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y+c_1)(2\sigma_{xy}+c_2)}{(\mu_x^2+\mu_y^2+c_1)(\sigma_x^2+\sigma_y^2+c_2)} \quad (52)$$

In the above equation, μ_x, μ_y are averages of x, y respectively. σ_x^2, σ_y^2 are variances of x, y respectively, and c_1, c_2 are constants. σ_{xy} is the covariance of x and y.

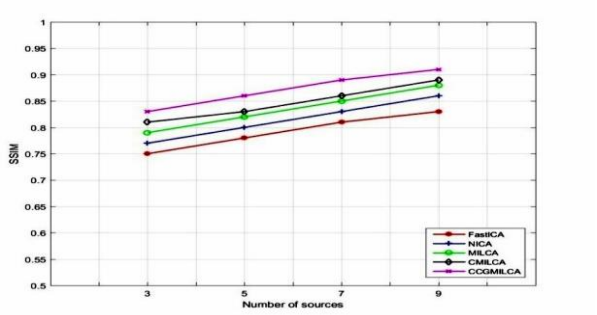


Fig. 5. Comparison of SSIM value for Image dataset

Fig 5 illustrates the comparison of SSIM for colour image database. The result shows that the SSIM value of CCGMILCA method is higher than the other methods. While considering the number of observations as 10, the SSIM of FastICA was 0.83, NICA was 0.86, MILCA was 0.88 and when CMILCA and CCGMILCA are applied, the value of SSIM is improved to 0.89 and 0.91 respectively.

5.3 Universal Image Quality Index (UIQI)

It is referred as the overall normalized quality value of the

$$\mathrm{UIQI} = \frac{1}{\mathrm{N}\times\mathrm{M}}\sum_{i=1}^{\mathrm{N}}\sum_{j=1}^{\mathrm{M}}\mathrm{Q}_{ij} \quad (53)$$

In the above equation, N and M are the row and column numbers of the image or audio signal respectively. Q_{ij} refers the local quality index at (i, j).

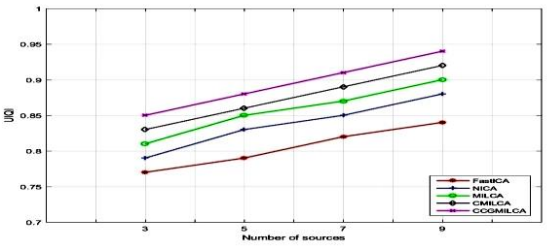


Fig. 6: Comparison of UIQI value for Image dataset

Fig 6 illustrates the comparison of UIQI for colour image database. The result shows that the UIQI value of CCGMILCA method is higher than the other methods. While considering the number of observations as 10, the UIQI of FastICA was 0.84, NICA was 0.88, MILCA was 0.90 and when CMILCA and CCGMILCA are applied the value of UIQI was improved to 0.92 and 0.94 respectively.

5.4 Signal-to-distortion ratio (SDR)

It is the ratio between Signal power and the Distortion power. Signal to Distortion ratio is obtained by using Equation (54).

$$SDR = 10 \log_{10} \frac{\sum_{i=1}^{I} \sum_{t} s_{ij}(t)^2}{\sum_{i=1}^{I} \sum_{t} (e_{ij}^{spat}(t) + e_{ij}^{interf}(t) + e_{ij}^{artif}(t))^2} \quad (54)$$

Fig .7 shows the comparison of SDR for audio signal. The result shows that the SDR value of CCGMILCA method is higher than all the other methods. For any number of observations, CMILCA and CCGMILCA based Blind source separation results from 18 dB to 20 dB. The results show that the proposed approaches are separating signals from the mixture signals with lesser distortion than classical Blind source separation. CMILCA method shows almost equal performance when compared to CCGMILCA method performance.

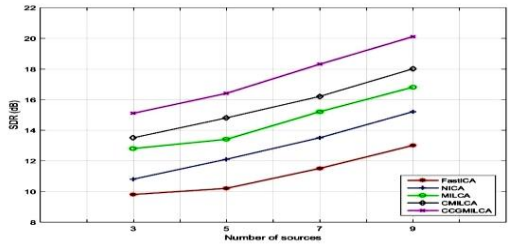


Fig. 7. Comparison of SDR for Audio Dataset

5.5 signal-to-interference ratio (SIR)

$$SIR = 10 \log_{10} \frac{\sum_{i=1}^{I} \sum_{t} (s_{ij}(t) + e_{ij}^{spat}(t))^2}{\sum_{i=1}^{I} \sum_{t} e_{ij}^{interf}(t)^2} \quad (55)$$

A noise due to mis-separation is called interference (For instance, while extracting the lyrics from a song, this might be a residual of the background tune).

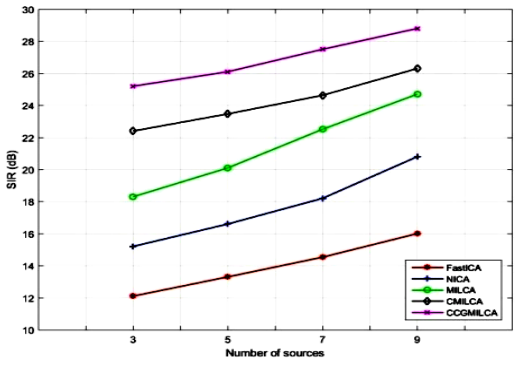


Fig. 8. Comparison of SIR for Audio Dataset

Fig 8 shows the comparison of SIR for audio signals. The result shows that the SIR value of CCGMILCA method is higher than the other methods. SIR value obtained for CMILCA is improved 9% than MILCA. CCGMILCA results around 15% improvement than MILCA method.

5.6 SAR (signal to artifacts ratio)

$$SAR = 10 \log_{10} \frac{\sum_{i=1}^{I} \sum_{t} (s_{ij}(t) + e_{ij}^{spat}(t) + e_{ij}^{interf}(t))^2}{\sum_{i=1}^{I} \sum_{t} e_{ij}^{artif}(t)^2} \quad (56)$$

A noise due to the reconstruction algorithm itself is known as artifacts. Fig 9 shows the comparison of SAR for audio signals. The result shows that the SIR value of CCGMILCA method is higher than the other methods. SIR value obtained for CMILCA is 6% improved than MILCA. CCGMILCA is obtained with more than 18% improvement than MILCA.

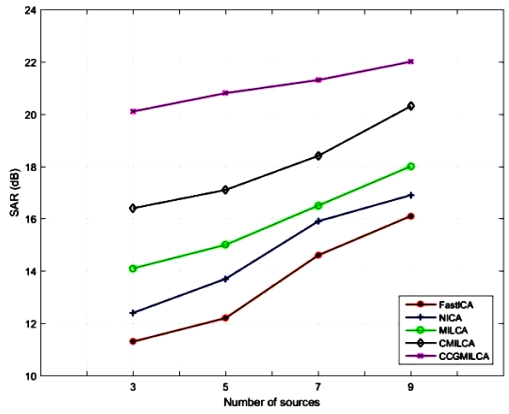


Fig. 9. Comparison of SAR for Audio Dataset

5.7 Computational Time

It refers to the time needed for finishing the computational process. Figure.10 illustrates the comparison of computational time for colour image database. The results show that the computational time value of CCGMUILCA method is lesser than the other methods. While considering the number of observations as 10,the computational time of FastICA was with 1 hr.

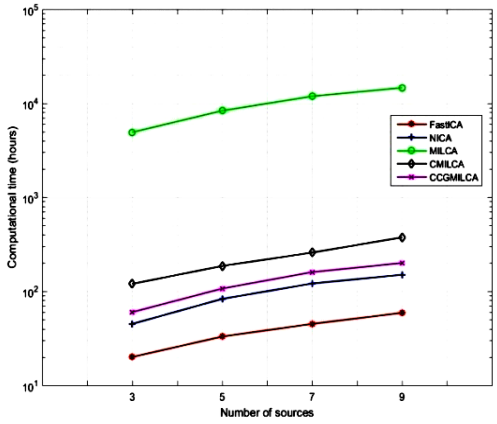


Fig. 10. Comparison of computation time for Image Dataset (hrs)

NICA was 2 hrs, MILCA was 29hrs and when CMILCA and CCGMILCA are applied, the value of computational time was reduced to 4hrs and 2.5hrs respectively. Comparatively MILCA method takes many hours compared to all other methods.

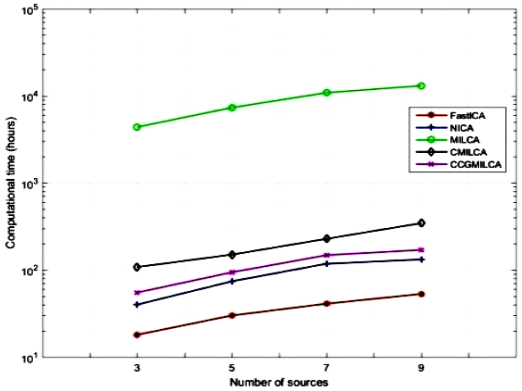


Fig. 11. Comparison of computational time Audio Dataset (hrs)

Fig 11 shows the comparison of computational time for audio signals. The result shows that the computational time value of CCGMILCA method is lesser than the other methods. While considering the number of observations as 10, the computational time of FastICA was 50 mins, NICA was 2.2hrs, MILCA was 25hrs and when CMILCA and CCGMILCA are applied the value of computational time is reduced to 5 hrs and 2.8 hrs respectively.

6. Conclusion and Future work

In this paper, the problem of Independent Component Analysis (ICA) in Blind Source Separation (BSS) is analyzed. From the analysis, it is found that MILCAmethod performs better blind source separation, but it takes much time to converge. Hence, Clustering of Mutual Information based Least-dependent Component Analysis (CMILCA) technique is developed to overcome the problem of MILCA.To supportthe number of observations and sources, Clustering in Conjugate Gradients of Mutual Information based Least-dependent Component Analysis (CCGMILCA) method is used. Thus the proposed BSS methods are used to reduce the considerable computation time of MILCA.The experiment results proved that the proposed techniques perform better in terms of SNR,SSIM for image and SDR,SIR,SAR for Audio dataset. In future the problem with NMF based Blind source separation techniques will be analyzed to give better results and compared with the proposed techniques.

References

1. Absil, P.-A., Mahony, R., & Sepulchre, R.: Optimization Algorithms on Matrix Manifolds. Princeton University Press.(2008).

2. Benahmed, A., Zenkouar, L., & Hamzaoui, E.: DOA estimation for automatic steering of a smart antenna radiation pattern using the conventional ICA-based blind source separation method. In Proceedings of 2014 Mediterranean Microwave Symposium (MMS2014), 1-5 (2014). doi: 10.1109/MMS.2014.7089001.

3. Bobin, J., Rapin, J., Larue, A., & Starck, J. L.: Sparsity and adaptivity for the blind separation of partially correlated sources. IEEE Transactions on Signal Processing, 63(5), 1199-1213. doi: 10.1109/TSP.2015.2391071.

4. Bronstein, A. M., Bronstein, M. M., Zibulevsky, M., & Zeevi, Y. Y.: Sparse ICA for blind separation of transmitted and reflected images. International Journal of Imaging Systems and Technology, 15(1), 84-91 (2005). doi: 10.1002/ima.20042.

5. Chan, T. H., Ma, W. K., Chi, C. Y., & Wang, Y.: A convex analysis framework for blind separation of non-negative sources. IEEE Transactions on Signal Processing, 56(10), 5120-5134 (2008). doi: 0.1109/TSP.2008.928937.

6. Chien, J. T., & Hsieh, H. L.: Convex divergence ICA for blind source separation. IEEE Transactions on Audio, Speech, and Language Processing, 20(1), 302-313 (2012). doi: 10.1109/TASL.2011.2161080.

7. Gang Niu,Wittawat Jitkrittum ,Bo Dai,Hirotaka Hachiya& Masashi Sugiyama.: Squared-loss Mutual Information Regularization:Information-theoretic Approach to Semi-supervised Learning. Proceedings of the 30 th International Conference on Machine Learning, Atlanta, Georgia, USA, 2013.

8. Guo, X., Uhlich, S., & Mitsufuji, Y. (2015, April). NMF-based blind source separation using a linear predictive coding error clustering criterion. In 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 261-265 (2015). doi: 10.1109/ICASSP.2015.7177972.

9. H. L. Hsieh and J. T. Chien: "A new nonnegative matrix factorization for independent component analysis," IEEE International Conference on Acoustics, Speech and Signal Processing, Dallas, TX, 2026-2029 (2010). doi: 10.1109/ICASSP.2010.5494945.

10. Hager, W.W., & Zhang, H.: A survey of nonlinear conjugate gradient methods. Paci_c journal of Optimization, 2(1), 35-58 (2006).

11. Hattay, J., Belaid, S., & Naanaa, W.: Non-negative matrix factorisation for blind source separation in wavelet transform domain. IET Signal Processing, 9(2), 111-119 (2014). doi:10.1049/iet-spr.2013.0409.

12. http://tabby.vision.mcgill.ca/html/Animals1.html

13. http://www.cstr.ed.ac.uk/projects/eustace/download.html

14. https://en.wikipedia.org/wiki/Blind_signal_ separation

15. Hyvärinen, A.: Independent component analysis: recent advances. Phil. Trans. R. Soc. A, 371(1984), 20110534(2013). doi: 10.1098/rsta.2011.0534.

16. Hyvrinen A, Oja, E.: Independent component analysis: algorithms and applications.Elsevier Ltd. Neural Networks 13(4), 411–430 (2000)

17. J.D. Victor, Phys. Rev. E 66, 051903-1 (2002)

18. Langlois, D., Chartier, S., & Gosselin, D. (2010). An introduction to independent component analysis: InfoMax and FastICA algorithms. Tutorials in Quantitative Methods for Psychology, 6(1), 31-38, (2010). doi: 10.20982/tqmp.06.1.p031.

19. Meganem, I., Deville, Y., Hosseini, S., Deliot, P., & Briottet, X: Linear-quadratic blind source separation using NMF to unmix urban hyperspectral images. IEEE Transactions on Signal Processing, 62(7), 1822-1833 . (2014). doi: 10.1109/MMS.2014.7089001.

20. Mustafi, A., & Ghorai, S. K.: A novel blind source separation technique using fractional Fourier transform for denoising medical images. Optik-International Journal for Light and Electron Optics, 124(3), 265-271 (2013). doi: 10.1016/j.ijleo.2011.11.052.

21. Oja, E., & Plumbley, M.: Blind separation of positive sources by globally convergent gradient search. Neural Computation, 16(9), 1811-1825(2004).doi: 10.1162/0899766041336413.

22. Ozerov, A., & Févotte, C.: Multichannel nonnegative matrix factorization in convolutive mixtures for audio source separation. IEEE Transactions on Audio, Speech, and Language Processing, 18(3), 550-563 (2010). doi: 10.1109/TASL.2009.2031510.

23. Parra, L., & Sajda, P.: Blind source separation via generalized eigenvalue decomposition. Journal of Machine Learning Research, 1261-1269 (2003). doi: 10.1162/jmlr.2003.4.7-8.1261.

24. Plumbley, M. D.: Algorithms for nonnegative independent component analysis. IEEE Transactions on Neural Networks, 14(3), 534-543(2003). doi: 0.1109/TNN.2003.810616.

25. Qingming, Y.: Blind source separation by weighted K-means clustering. Journal of Systems Engineering and Electronics, 19(5), 882-887(2008). doi: 10.1016/S1004-4132(08)60168-1.

26. R.L. Somorjai, “Methods for Estimating the Intrinsic Dimensionality of High-Dimensional Point Sets”, in Dimensions and Entropies in Chaotic Systems, G. Mayer-Kress,Ed. (Springer, Berlin 1986).

27. Saruwatari, H., Kawamura, T., Nishikawa, T., Lee, A., & Shikano, K.: Blind source separation based on a fast-convergence algorithm combining ICA and beamforming. IEEE Transactions on Audio, speech, and language processing, 14(2), 666-678 (2006). doi: 10.1109/TSA.2005.855832.

28. Stögbauer, H., Kraskov, A., Astakhov, S. A., & Grassberger, P. : Least-dependent-component analysis based on mutual information. Physical Review E, 70(6), 066123(2004). doi: 10.1103/PhysRevE.70.066123.

29. Suzuki, T., Sugiyama, M., Kanamori, T., and Sese, J. Mutual information estimation reveals global associations between stimuli and biological processes. BMC Bioin-formatics, 10(1):S52, 2009. doi: 10.1186/1471-2105-10-S1-S52.

30. Szu, H., & Kopriva, I.: Deterministic blind source separation for space variant imaging. In Fourth International Symposium on Independent Component

Analysis and Blind Signal Separation, (2003). doi: 10.1.1.640.4763.

31. Vigliano, D., & Uncini, A.: Flexible ICA solution for nonlinear blind source separation problem. Electronics letters, 39(22), 1616-1617(2003). doi: 10.1049/el:20031033.

32. Vincent, E., Gribonval, R., F´evotte, C.: Performance measurement in blind audio source separation. IEEE Trans. on Audio, Speech and Language Processing 1462–1469 (2006). doi: 10.1109/TSA.2005.858005.

33. Wang, X., Ou, S., Gao, Y., & Guo, X. : A new fast nonlinear principal component analysis algorithm for blind source separation. In Fuzzy Systems and Knowledge Discovery (FSKD), 2015 12th International Conference, 1626-1630 (2015). Doi: 10.1109/FSKD.2015.7382188.

34. Zelnik-Manor, L., & Perona, P.: Self-tuning spectral clustering. Advances in Neural Information Processing Systems 17,1601–1608 (2005). Cambridge, MA, USA: MIT Press. Doi: 10.1.1.84.7940.

35. Zhu, X. L., Zhang, X. D., Ding, Z. Z., & Jia, Y.: Adaptive nonlinear PCA algorithms for blind source separation without prewhitening. IEEE Transactions on Circuits and Systems I: Regular Papers, 53(3), 745-753 (2006). doi: 10.1109/TCSI.2005.858489.

36. Zibulevsky, M., Kisilev, P., Zeevi, Y. Y., & Pearlmutter, B. A.: Blind source separation via multinode sparse representation. Advances in Neural Information Processing Systems 14, 1049-1056(2002). doi: 10.1109/ICIP.2001.958086.

37. Parimala Gandhi A, Vijayan S. Various feasible maneuvers for the analysis of an image for retrieval of information and optimization. Int J Innovat Eng Techol. 2016;7(3):537-42.