

Independent Component Analysis (ICA) dan Sparse Component Analysis (SCA) dalam Pemisahan Vokal dan Instrumen pada Seni Geguntangan

Angga Pramana Putra^{a1}, Ni Wayan Wiantari^{a2}, Ni Putu Mira Novita Dewi^{b3},
I Dewa Made Bayu Atmaja Darmawan^{b4}

^aProgram Studi Teknik Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana

Jalan Raya Kampus Unud, Jimbaran, Bali, 80361, Indonesia

¹sunhangga@gmail.com

²wayan_wiantari@yahoo.co.id

³miranovitad@gmail.com

⁴dewabayu@cs.unud.ac.id

Abstrak

Geguntangan adalah pesantian dalam upacara keagamaan yang diiringi dengan gamelan. Indra pendengaran manusia cenderung memiliki keterbatasan, yang menyebabkan tidak semua vokal yang tercampur dengan gamelan bisa didengar jelas. Oleh karena itu diperlukan suatu sistem yang dapat digunakan untuk memisahkan vokal dengan gamelan pada geguntangan. Pemisahan sumber suara ini dikategorikan sebagai Blind Source Separation (BSS) atau disebut juga Blind Signal Separation yang artinya sumber tidak dikenal. Algoritma yang digunakan untuk menangani BSS adalah algoritma Independent Component Analysis (ICA) dan Sparse Component Analysis (SCA) dengan berfokus pada pemisahan sinyal suara pada file suara berformat wav. Algoritma SCA dan ICA digunakan untuk proses pemisahan suara dengan parameter nilai yang digunakan adalah Mean Square Error (MSE) dan Signal to Interference Ratio (SIR). Dari hasil simulasi menunjukkan Hasil perhitungan MSE dan SIR dengan menggunakan mixing matriks [0.3816, 0.8678], [0.8534, -0.5853] didapatkan hasil untuk metode ICA, nilai MSE adalah 3.60×10^{-5} untuk vokalnya dan 1.71×10^{-6} untuk instrumennya, memiliki SIR 44.84784936 untuk vokalnya dan 59.89866528 untuk instrumennya. Untuk metode SCA nilai MSE adalah 1.25×10^{-4} untuk vokalnya dan 4.21×10^{-6} untuk instrumennya, memiliki SIR 40.00942668 untuk vokalnya dan 54.70771013 untuk instrumennya.

Keywords: *Blind Sources Separation, Independent Component Analysis, Sparse Component Analysis, Audio Signal Processing, Mean Square Error, Signal to Interference Ratio*

1. Pendahuluan

Geguntangan adalah pesantian dalam upacara keagamaan yang diiringi dengan gamelan. Geguntangan sering dipakai dalam upacara keagamaan untuk mengiringi jalannya upacara dan juga untuk hiburan masyarakat. Namun, suara yang tercampur antara vokal dengan gamelan menyebabkan sulitnya dalam mempelajari vokal pada geguntangan.

Indra pendengaran manusia cenderung memiliki keterbatasan, yang menyebabkan tidak semua vokal yang tercampur dengan gamelan bisa didengar jelas. Oleh karena itu diperlukan suatu sistem yang dapat digunakan untuk memisahkan vokal dengan gamelan pada geguntangan. Pemisahan sumber suara ini dikategorikan sebagai *Blind Source Separation* (BSS) atau disebut juga *Blind Signal Separation* yang artinya sumber tidak dikenal [10].

BSS merupakan suatu cara untuk memisahkan sinyal tercampur menjadi sejumlah sinyal pembentuknya, tanpa informasi mengenai jumlah sumber sinyal, atau proses tercampurnya sinyal-sinyal tersebut. BSS memanfaatkan perbedaan sifat sinyal sebelum terdeteksi sensor dan informasi yang diperoleh akibat perbedaan sudut datang dan jarak tempuh pada sensor. Sensor yang digunakan adalah *microphone*. Terdapat dua teknik dalam merekam suara yaitu, *single channel* dan

multi channel. *Single channel* adalah teknik perekaman yang menggunakan sensor tunggal, dan *multi channel* adalah teknik perekaman yang menggunakan sensor lebih dari satu.

Algoritma yang digunakan untuk menangani BSS adalah algoritma *Independent Component Analysis (ICA)* dan *Sparse Component Analysis (SCA)* dengan berfokus pada pemisahan sinyal suara pada file suara berformat wav yang memiliki dua atau lebih suara yang tercampur sehingga didapatkan hasil suara terekam terpisah. Alasan digunakannya format wav adalah file berformat wav berisi suara yang tidak mengalami kompresi [9].

Algoritma ICA digunakan pada penelitian ini, karena algoritma ICA dapat membedakan unsur atau komponen penyusun campuran sinyal secara independen [2]. Sedangkan, algoritma SCA digunakan dengan alasan algoritma SCA mampu memanfaatkan sedikit sinyal untuk mengekstraksi sumber dan memiliki presisi yang lebih tinggi dalam pemisahan sinyal suara [10], yang mana pada penelitian ini hanya menggunakan 2 sumber suara, yaitu sinyal suara vokal dan instrumen. Sebelum melakukan proses ICA atau SCA, pertama kali yang harus dilakukan adalah beberapa proses pendahuluan (*pre-processing*). *Pre-processing* yang dimaksud adalah *centering* dan *whitening*.

Penerapan algoritma ICA dan SCA selain digunakan dalam *Blind Source Separation*, bisa juga digunakan untuk memperbaiki kualitas suara pada data audio yang tercampur dengan *noise* atau suara yang mengganggu atau yang tidak diinginkan seperti penelitian yang dilakukan oleh Indra, dkk dimana penelitian yang dilakukannya mengenai pemisahan sinyal dengan *noisanya* untuk memperbaiki kualitas suatu audio atau *noise reduction*.

Dalam pemisahan sinyal suara, selain menggunakan algoritma ICA dan SCA, ada juga algoritma lain yang biasa digunakan dalam *Blind Source Separation* yaitu *Non-negative Matrix Factorization (NMF)*. NMF dapat memproses data dalam skala besar menggunakan model faktorisasi matriks dibandingkan dengan algoritma klasik. Beberapa kelebihan dari NMF adalah implementasi yang mudah, interpretabilitas hasil dekomposisi yang baik, dan ruang penyimpanan yang kecil [3].

Algoritma SCA dan ICA digunakan untuk proses pemisahan suara dengan parameter nilai yang digunakan adalah *Mean Square Error (MSE)* atau melihat kesamaan antara hasil output dengan input dipilih untuk menguji hasil output suara. Selain itu parameter *Signal to Interference Ratio (SIR)* digunakan untuk mengetahui rasio interferensi output suara dibandingkan dengan input. Pada pengenalan sinyal suara, proses mendengarkan secara lisan hasil output dibandingkan dengan input suara digunakan untuk mengukur parameter keberhasilan. Selain itu, dengan membandingkan sinyal input dan output, dapat diketahui hasil proses pemisahan suara.

2. Metode Penelitian

2.1. Rancangan Penelitian

Penelitian yang dilakukan merupakan penelitian eksperimental dengan membandingkan MSE dan SIR pada algoritma ICA dan SCA dalam pemisahan vokal dengan gamelan pada geguntangan. File geguntangan yang diperoleh dalam bentuk wav akan dilakukan proses *centering* dan *whitening* untuk mendapatkan sebuah vektor baru yang variansnya sama dengan satu, terakhir melakukan proses ICA dan SCA terhadap vektor hasil *centering* dan *whitening*.

2.2. Lokasi dan Waktu Penelitian

Penelitian ini dilakukan di Laboratorium Jaringan Program Studi Teknik Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana selama 3 bulan.

2.3. Data dan Metode Pengumpulan Data

Alat yang digunakan pada penelitian ini meliputi alat perekam menggunakan recorder Tascam DR-40. perekaman pertama dilakukan pada suara vokal yang dilanjutkan dengan merekam suara instrumennya, *python* sebagai bahasa pemrograman untuk mengolah sinyal suara geguntangan, dan laptop sebagai alat untuk menggunakan perangkat lunak yang akan digunakan. Data yang digunakan pada penelitian ini meliputi rekaman vokal serta instrumen geguntangannya dalam format file wav. Jumlah data yang digunakan sebanyak 80 data. Data vokal sebanyak 40 dengan narasumber yang berbeda, yaitu 2 laki-laki dan 2 perempuan. Data instrumen sebanyak 40 yang merupakan iringan dari vokalnya. Pengambilan data dilakukan dari pengambilan vokal yang dilanjutkan dengan

pengambilan data instrumen yang telah disesuaikan tempo dengan vokalnya. Untuk menyesuaikan vokal dengan instrumennya dengan cara menggunakan headset pada penabuhnya, sehingga iringan dapat selaras dengan vokalnya.

2.4. Tahapan Penelitian

Tabel 1. Tahapan Penelitian

Deskripsi	Luaran	Indikator Capaian
Pengumpulan dan determinasi data rekaman.	Informasi dari rekaman geguntangan yang memiliki sedikit gangguan suara lain.	Tersedianya data rekaman geguntangan yang memiliki sedikit gangguan suara lain.
<i>Pre-processing centering</i> pada data rekaman geguntangan.	Data rekaman dengan <i>noise</i> yang minimal.	Didapat matriks dari data rekaman dengan <i>noise</i> yang minimal.
<i>Pre-processing whitening</i> pada matriks hasil <i>centering</i> .	Matriks identitas hasil transformasi sinyal rekaman.	Didapatnya matriks identitas dari hasil transformasi sinyal rekaman.
Proses ICA dan SCA pada matriks hasil <i>pre-processing</i> .	Sinyal rekaman terpisah antara vokal dan gamelan.	Didapatnya sinyal rekaman yang terpisah antara vokal dan gamelan.
Proses pengujian data hasil proses ICA dan SCA dengan MSE dan SIR.	Nilai MSE dan SIR sebagai parameter keberhasilan pemisahan sinyal rekaman.	Didapatnya nilai MSE dan SIR sebagai parameter keberhasilan dalam pemisahan sinyal rekaman.
Analisis data.	Informasi nilai MSE dan SIR.	Didapatnya informasi algoritma terbaik dalam pemisahan sinyal.

2.5. *Pre-processing Data*

2.5.1. *Pre-processing Centering* Rekaman Geguntangan

Sinyal rekaman geguntangan dengan rata-rata sinyal itu sendiri untuk membuat rata-rata suatu data menjadi nol. Perhitungan menggunakan fungsi dari *python*. Hasil dari proses *centering* adalah berupa matriks.

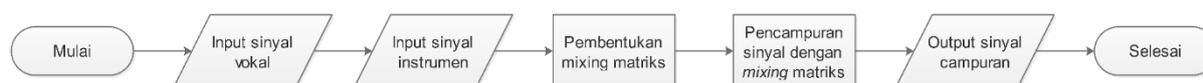
2.5.2. *Pre-processing Whitening* Rekaman Geguntangan

Matriks dari hasil *centering* akan dilakukan proses transformasi data ke bentuk data baru agar didapat nilai matriks kovariansnya berupa matriks identitas menggunakan fungsi yang terdapat pada *python*.

2.6. Proses ICA dan SCA

a. Proses *mixing* matriks

Sebelum sinyal diproses menggunakan algoritma ICA dan SCA, sebelumnya sinyal dicampurkan menggunakan *mixing* matriks. *Mixing* matriks dipulihkan dengan cara acak yang nantinya digunakan untuk mencampurkan sinyal vokal dan sinyal instrumennya. Output yang dihasilkan dari proses ini berupa sinyal vokal dengan instrumen yang tercampur.



Bagan 1. Diagram alir *mixing* matriks

b. Proses ICA

Proses ICA dapat digambarkan dengan diagram alir seperti yang terlihat pada bagan 1 dengan input sinyal campuran yang telah dibuat dengan *mixing* matriks.



Bagan 2. Diagram alir algoritma ICA

Pada proses ini didapatkan nilai Y sebagai sinyal estimasi dan W sebagai *unmixing* matriks. Proses pemisahan suara dilakukan dengan persamaan sebagai berikut:

$$Y(t) = W Xn(t) \quad (1)$$

Untuk menentukan komponen dari algoritma ICA diperlukan metode *symmetrical*, berikut langkah-langkahnya:

1. Menentukan jumlah komponen yang independen atau variabel m .
2. Memilih nilai awal vektor kompleks w , dengan y adalah perkalian w dengan xn .
3. Mencari nilai g (ketidaklinieritas) dan g' menggunakan persamaan berikut:

$$\begin{aligned} g(y) &= y \exp(-y^2/2) \\ g'(y) &= (1-y^2)\exp(-y^2/2) \end{aligned} \quad (2)$$

4. Menghitung nilai w baru dengan persamaan:

$$w \leftarrow E\{zg(w^T z)\} - E\{g'(w^T z)\}w \quad (3)$$

5. Melakukan proses iterasi seperti pada persamaan berikut:

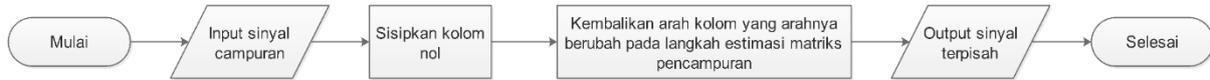
$$\begin{aligned} W &\leftarrow \frac{W}{\|W\|} \\ W &\leftarrow \frac{3}{2}W - \frac{1}{2}WW^T W \end{aligned} \quad (4)$$

6. Didapatkan nilai W atau *unmixing* matriks.

Berkutnya dilakukan persamaan $Y(t) = W Xn(t)$ dengan memasukkan nilai W untuk mendapatkan sinyal estimasi atau Y .

c. Proses SCA

Proses SCA dapat digambarkan dengan diagram alir seperti yang terlihat pada bagan 2.



Bagan 3. Diagram alir algoritma SCA

Pada proses ini didapat sinyal sumber dengan mencari estimasi matriks lalu dilakukan pemulihan sinyal sumber. Untuk mencari estimasi matriks campuran digunakan persamaan:

$$X_i = k_1 \bar{X}_j + k_2 \bar{X}_k \quad (5)$$

Setelah didapatkan estimasi matriks campuran, selanjutnya dilakukan pemulihan dengan menggunakan persamaan $Q_j = L^t \times (L \times L^T)^{-1} \times X_j$ dengan kondisi:

$$Q_j = (Q_{1j}, Q_{2j}, \dots, Q_{n-1,j})^T, \text{ then } \begin{cases} S_{kj} = Q_{kj} & (k < j_1) \\ S_{kj} = 0 & (k = j_1) \\ S_{(k+1)j} = Q_{kj} & (j_1 < k < j_2) \\ S_{(k+2)j} = 0 & (k = j_2) \\ S_{(k+2)j} = Q_{kj} & (k > j_2) \end{cases} \quad (6)$$

2.7. Pengujian Metode

2.7.1. Pengujian Kualitas Data Hasil Proses ICA dan SCA dengan MSE dan SIR

a. *Mean Square Error (MSE)*

MSE merupakan suatu metode untuk mengukur perbedaan antara estimator (sinyal rekonstruksi) dan nilai sebenarnya (sinyal *baseline*) dari kualitas yang diperkirakan. Secara garis besar dengan menghitung nilai MSE, maka akan diperoleh selisih pergeseran yang diperoleh antara sinyal asli dan sinyal rekonstruksi, yang dapat ditunjukkan pada persamaan di bawah ini [1]. Variabel yang dibutuhkan untuk pengujian MSE adalah jumlah data, sinyal asli, dan sinyal estimasi.

$$MSE = \frac{1}{n} \int_i^t (S - S_n)^2 dt \quad (7)$$

b. *Signal to Inference Ratio (SIR)*

SIR adalah perbandingan daya sinyal terhadap daya interferensi dan digunakan untuk menilai kualitas sinyal terhadap gangguan interferensi akibat penggabungan sinyal antara vokal dengan instrumennya. Interferensi adalah gangguan selain *noise* yang dapat menyebabkan kualitas suatu sinyal menurun. Semakin tinggi nilai SIR, maka kualitas sinyal semakin baik, begitu juga sebaliknya. Variabel yang dibutuhkan untuk pengujian SIR adalah nilai dari MSE. Rumus untuk mencari nilai SIR adalah sebagai berikut:

$$SIR = -10 \log_{10}(MSE) \quad (8)$$

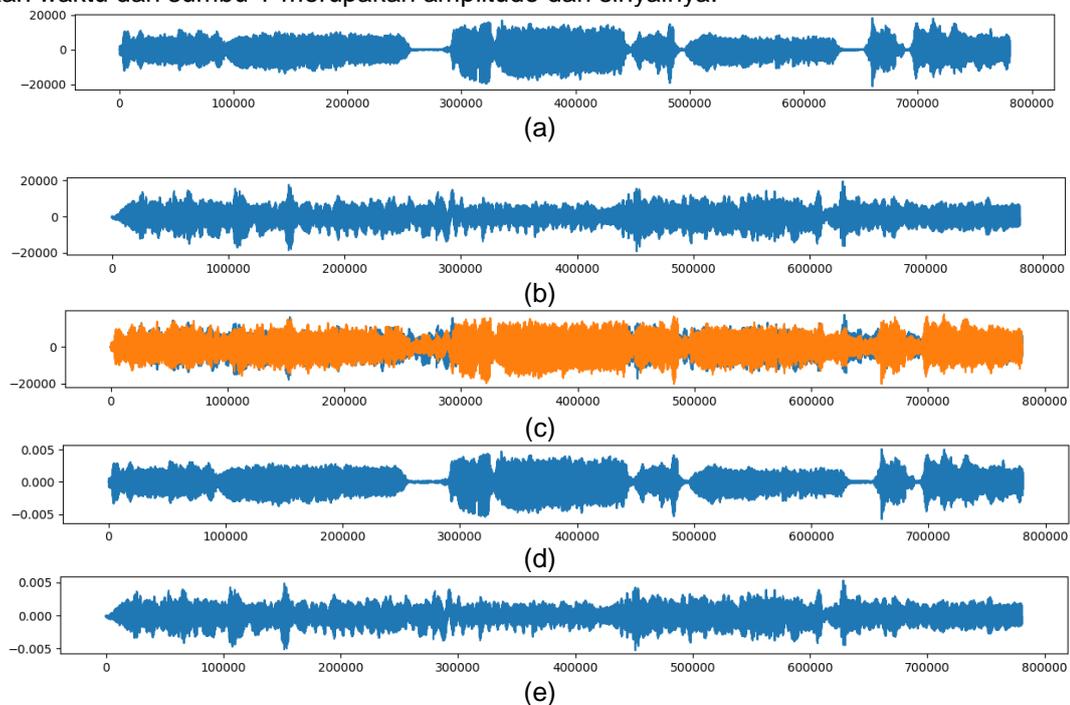
2.7.2. Analisis Data

Data yang diperoleh dari hasil perhitungan MSE dan SIR akan menentukan kualitas hasil pemisahan. Bila nilai MSE semakin mendekati nol menunjukkan sinyal hasil pemisahan semakin mirip dengan sinyal aslinya, dan bila nilai SIR semakin besar maka kualitas sinyal hasil pemisahan akan semakin bagus.

3. Hasil dan Pembahasan

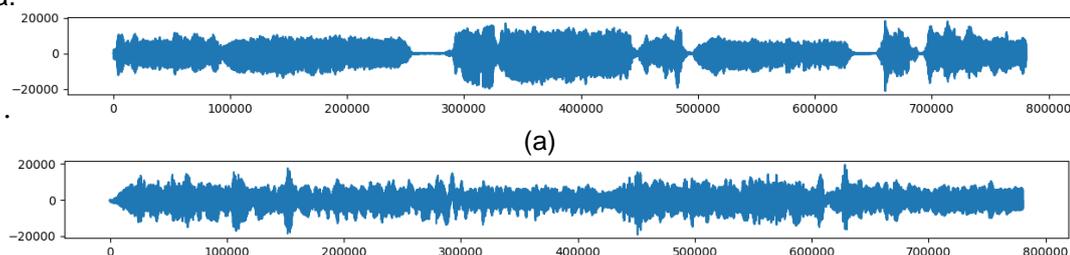
3.1. Analisa Hasil Simulasi *Blind Source Separation*

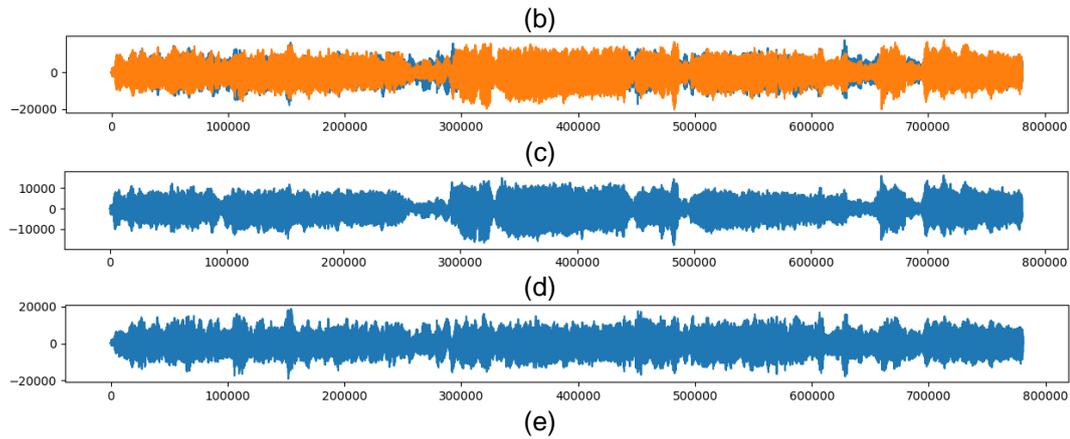
Untuk mendapatkan hasil berupa sinyal pemisahan dari penerapan *Blind Source Separation* dilakukan dengan simulasi *python*. Gambar 1 menunjukkan sinyal sumber, campuran dan hasil pemisahan dari suara wanita pertama dengan instrumennya menggunakan metode ICA. Sumbu X merupakan waktu dan sumbu Y merupakan amplitudo dari sinyalnya.



Gambar 1. Hasil plotting sinyal sumber vokal (a), sinyal sumber instrumen (b), sinyal campuran (c), sinyal vokal hasil pemisahan (d) dan sinyal instrumen hasil pemisahan (e)

Selanjutnya adalah implementasi dari metode SCA, pada gambar 2 ditunjukkan sinyal sumber, campuran dan hasil pemisahan dari suara wanita pertama dengan instrumennya menggunakan metode SCA. Dengan sumbu X melambangkan waktu dan sumbu Y merupakan amplitudo dari sinyalnya.





Gambar 2. Hasil plotting sinyal sumber vokal (a), sinyal sumber instrumen (b), sinyal campuran (c), sinyal vokal hasil pemisahan (d) dan sinyal instrumen hasil pemisahan (e)

3.2. Analisa MSE dan SIR pada penerapan Algoritma BSS

Pengujian dari kedua metode menggunakan parameter MSE untuk menguji *error* antar sinyal dengan cara membandingkan sinyal sumber dengan sinyal hasil pemisahan dan SIR untuk menguji kekuatan sinyal terhadap sinyal interferensi akibat penggabungan sinyal. Semakin kecil nilai MSE dan semakin besar nilai SIR maka metode yang digunakan dalam pemisahan sinyal semakin baik, begitu juga sebaliknya. Hasil perhitungan MSE dan SIR dengan menggunakan *mixing* matriks [0.3816, 0.8678], [0.8534, -0.5853]. Selanjutnya diimplementasikan terhadap 80 data dan didapatkan kesimpulan dengan mencari rata-rata dari tiap hasil pengujian yang didapat. Berikut disajikan rata-rata dari hasil pengujian 80 data:

Tabel 2. Rata-rata hasil perhitungan MSE dan SIR metode ICA dan SCA

	Rata-rata							
	ICA				SCA			
	MSE		SIR		MSE		SIR	
	Vokal	Instrumen	Vokal	Instrumen	Vokal	Instrumen	Vokal	Instrumen
Wanita 1	2.73E-05	4.57E-06	45.8225729	55.1981151	6.48E-05	7.53E-06	43.545176	52.7577786
Wanita 2	3.86E-05	8.36E-07	44.4312811	61.1268276	1.43E-04	3.41E-06	38.73923	55.0189437
Laki-laki 1	2.44E-05	6.49E-07	46.2381967	61.9981791	9.04E-05	2.65E-06	40.545783	55.8913833
Laki-laki 2	5.38E-05	7.97E-07	42.8993468	61.2715393	2.00E-04	3.26E-06	37.207518	55.1627349

Dari tabel 1, dapat disimpulkan untuk metode ICA, nilai MSE adalah 3.60×10^{-5} untuk vokalnya dan 1.71×10^{-6} untuk instrumennya, memiliki SIR 44.84784936 untuk vokalnya dan 59.89866528 untuk instrumennya. Untuk metode SCA nilai MSE adalah 1.25×10^{-4} untuk vokalnya dan 4.21×10^{-6} untuk instrumennya, memiliki SIR 40.00942668 untuk vokalnya dan 54.70771013 untuk instrumennya. Apabila nilai MSE semakin mendekati nol menunjukkan sinyal hasil pemisahan semakin mirip dengan sinyal asli sedangkan semakin besar SIR menunjukkan semakin bagus kualitas sinyal hasil pemisahan.

4. Kesimpulan

Kesimpulan dari tulisan ini adalah:

- a. Hasil perhitungan MSE dan SIR dengan menggunakan *mixing* matriks [0.3816, 0.8678], [0.8534, -0.5853] didapatkan hasil untuk metode ICA, nilai MSE adalah 3.60×10^{-5} untuk vokalnya dan 1.71×10^{-6} untuk instrumennya, memiliki SIR 44.84784936 untuk vokalnya dan 59.89866528 untuk instrumennya. Untuk metode SCA nilai MSE adalah 1.25×10^{-4} untuk vokalnya dan 4.21×10^{-6} untuk instrumennya, memiliki SIR 40.00942668 untuk vokalnya dan 54.70771013 untuk instrumennya.
- b. Dilihat dari hasil nilai estimasi *error* dari MSE dan SIR, maka metode ICA lebih baik dari SCA dalam pemisahan antara vokal dengan iringan musik gegungtangannya.

Referensi

- [1] Anda A R, "Penggunaan Frekuensi Sesaat Untuk Deteksi Pola Suara Kerusakan Motor Listrik", Institut Teknologi Sepuluh November Surabaya, 2006.
- [2] Chan T F, S. J. 2005. *Image Processing and Analysis: Variational PDE, Wavelet and Scholastic Methods*. Philadelphia.
- [3] Ganesh, Wenwu Wang. 2014. "Blind Source Separation Advance in Theory, Algorithms and Applications". Springer Heidelberg New York Dordrecht London.
- [4] Hayati, Dian Nur. 2010. "Penerapan Independent Component Analysis (ICA) Untuk Pemisahan Sinyal Suara Mesin Berputar di PT. Gresik Power Indonesia "THE LINDLE GROUP". ITS.
- [5] J. Benesty, J. Chen, Y. Huang, "Microphone Array Signal Processing" *Conventional Beamforming Techniques* pp. 39-46, 2008.
- [6] Nurlaili, I. 2011. "Pendeteksi Suara Jantung S1 dan S2 Menggunakan High Frequency Signatures". Bandung: IT Telkom.
- [7] Sari, "Pemisahan Sinyal Akustik Bawah Air Menggunakan Blind Separation of Source (BSS)". Tugas Akhir, Jurusan Teknik Elektro, Institut Teknologi Sepuluh Nopember Surabaya. 2011.
- [8] V. Capdevielle, C. Serviere, and J. Lacoume, "Blind separation of wideband sources in the frequency domain," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, pp. 2080–2083, Detroit, Mich, USA, May 1995.
- [9] Westner, Alex, V. Michael Bove Jr. 1999. "Blind Separation of Real World Audio Signals Using Overdetermined Mixtures". Cambridge : MIT Media Laboratory.
- [10] Xianchuan, Dan Hu, Jindong Xu. 2014. "Blind Source Separation Theory and Applications". China: Beijing Normal University.