

Penggunaan *Independent Component Analysis (ICA)* untuk Pembuangan Noise dan Artefak pada Sinyal Campuran

Riwinoto¹, Benyamin Kusumoputro²

¹Program Studi Teknik Informatika, Politeknik Batam
Park way st, Batam centre, Batam
Telp. (0778) 469856 Fax. (0778) 563620
E-mail: riwi@polibatam.ac.id

²Departemen Teknik Elektro, Fakultas Teknik, Universitas Indonesia
Jakarta
Email:kusumo@eng.ui.ac.id

ABSTRAK

Pengambilan sinyal campuran (dari beberapa sinyal asli) yang diamati sering memberikan hasil yang tidak akurat karena adanya distorsi terhadap sinyal campuran. Distorsi tersebut bisa berupa noise atau artefak. Jika seluruh distorsi pada sinyal campuran dapat dibuang maka sinyal campuran tersebut dapat dikatakan sebagai sinyal campuran yang murni bebas dari distorsi.

Noise adalah bagian dari sinyal asli yang mempunyai karakteristik distribusi gaussian dan memiliki periode waktu hidup yang sama dengan waktu hidup sinyal asli. Kurtosis adalah cumulant orde keempat yang merupakan perhitungan statistik orde tinggi. Nilai kurtosis pada sinyal berdistribusi gaussian adalah 0. Jadi untuk sinyal noise, terlihat nilai kurtosis akan mendekati 0.

Artefak adalah sinyal transient pendek yang mengganggu sinyal asli. Artefak biasanya terjadi karena ketidaksempurnaan proses perekaman atau akuisisi sinyal pada alat atau sensor perekam. Untuk mendeteksi sebuah sinyal termasuk artefak, sinyal dibagi menjadi beberapa blok. Setiap blok kemudian dihitung nilai variansinya. Kemudian dihitung kembali nilai variansi (variance of variance) dari himpunan variansi tersebut. Sinyal artefak ditunjukkan dengan nilai variance of variance yang lebih besar dari variance of variance sinyal biasa.

Independent Component Analysis (ICA) dapat digunakan untuk memperbaiki sinyal campuran yang mengandung noise dan artefak dengan menggunakan nilai kurtosis dan variance of variance komponen pembentuk sinyal campuran. Komponen sinyal yang bukan noise dan artefak kemudian direkonstruksi kembali menjadi sinyal campuran bebas noise dan artefak. Dari hasil percobaan, sinyal campuran yang mengandung noise dan artefak berhasil rekonstruksi kembali menjadi sinyal campuran yang bebas noise dan artefak.

Percobaan berhasil menemukan threshold kurtosis sebesar 1.7 dan threshold variance of variance sebesar 9. Sebuah sinyal diidentifikasi sebagai noise jika kurtosis sinyal lebih kecil dari threshold kurtosis sebesar 1.7 sedangkan sinyal diidentifikasi sebagai artefak jika nilai variansi sinyal lebih besar dari threshold variansi 9.

Kata Kunci: ICA, noise, artefak, kurtosis, variance of variance

1. PENDAHULUAN

Pengolahan sinyal suara telah banyak diteliti dan ditulis pada banyak paper khususnya pemisahan sinyal satu dengan sinyal yang lain. Pemisahan tersebut banyak diaplikasikan untuk kasus biomedik seperti memisahkan detak jantung bayi[3,4] dalam kandungan dengan detak jantung ibu bayi, dalam dunia telekomunikasi misalnya pemisahan antara sinyal audio dengan sinyal video dan dalam pengolahan citra untuk memisahkan noise yang terdapat pada image dengan image asli.

Metoda yang diusulkan untuk pemisahan tersebut yang paling populer adalah Independent Component analysis (ICA) yang diperkenalkan oleh hyvarinen [2]. ICA merupakan metode untuk memisahkan komponen sinyal sumber (asli) dari penggabungan sinyal-sinyal sumber tersebut melalui mekanisme perekaman dari sejumlah sensor (alat) perekaman. Input dari ICA adalah sinyal mixture sinyal sumber.

Pada sinyal-sinyal suara yang ditangkap oleh beberapa alat perekam seringkali terjadi adanya gangguan pada sinyal tersebut. Gangguan tersebut dapat berupa *noise* maupun *artefak*. Menurut Taigang [1], noise didefinisikan sebagai bagian dari sinyal yang membuat distorsi sinyal. Akibatnya terjadi kesalahan analisis dari sinyal tersebut. Contoh dalam perekaman detak jantung bayi dalam kandungan, gerakan teratur dari otot perut ibu bayi membuat distorsi sinyal detak jantung bayi dan detak jantung ibu yang sedang diobservasi oleh sensor.

Menurut taigang [1], Artefak adalah distorsi yang terjadi pada sinyal yang disebabkan oleh proses perekaman suara seperti gerakan elektroda (sensor) yang dipasang di badan ibu bayi. Adanya artefak menyebabkan sinyal tercampur hasil rekaman menjadi cukup terganggu apalagi jika gangguan artefak tersebut terjadi pada waktu (time) yang menentukan misalnya pada kasus deteksi jantung terjadi serangan jantung. Gangguan tersebut menyebabkan adanya kesulitan dalam analisa detak jantung pasien.

Oleh karena itu, paper ini memaparkan percobaan untuk menemukan sinyal-sinyal sumber asli yang terbebas dari noise dan artefak dari sinyal tercampur yang diobservasi menggunakan beberapa perekaman.

2. METODE PENELITIAN

2.1 Independent Component Analysis (ICA)

ICA adalah teknik perhitungan statistik untuk menemukan faktor-faktor tersembunyi yang ada pada sekumpulan variabel acak atau sinyal-sinyal. Salah satu penggunaan metode ICA adalah untuk memisahkan sinyal-sinyal tercampur yang berasal dari beberapa sumber yang saling bebas statistik satu sama lain dan distribusi sumber tersebut bersifat non gaussian

2.1.1 Model dasar

ICA menggunakan persamaan sederhana yaitu sebagai berikut:

$$X=AS \quad (1)$$

Dimana S merupakan kumpulan m Sinyal source, matrik A merepresentasikan mekanisme penggabungan sinyal-sinyal source dan X merepresentasikan sinyal mixture hasil penggabungan tersebut. Contoh sederhana dari rumus diatas terdapat dalam *cocktail problem* dimana terdapat 2 orang yang berhitung dengan bahasa masing-masing. Dua buah microphone merekam suara kedua orang tersebut dan menghasilkan suara campuran. Jadi terdapat dua buah suara campuran dari kedua orang tersebut.

Ide utama dari ICA adalah menemukan source asli S dengan asumsi bahwa source asli tersebut saling independent satu sama lain secara statistik. Artinya *joint probabilistic density function* komponen-komponennya (pdf) adalah product dari densitas seluruh sinyal asli.

$$P(s)=\prod p(s_i) \quad (2)$$

Di mana $p(s_i)$ adalah pdf dari sinyal asli dan $P(s)$ adalah *joint density function*.

Dengan sebuah vektor V, tujuan dari ICA adalah menemukan matrik U sedemikian sehingga $V=UX$

Dimana V adalah estimasi dari source sinyal S. Dari persamaan diatas diketahui bahwa matrik U adalah estimasi dari invers matrik A.

2.1.2 Algoritma

Estimasi model data ICA biasanya menggunakan fungsi objective misal *mutual information* atau *negentropy* kemudian dilakukan optimalisasi. Pada paper ini, estimasi model data ICA menggunakan algoritma JADE yang dikembangkan oleh Cardoso [3]. Algoritma JADE mendasarkan atas diagonalisasi matrik cummulant bersama. Algoritma JADE adalah sebagai berikut:

1. Inisialisasi

Dilakukan proses whitening W dan temukan $Z=WX$

Definisikan matrik covariance $R_x=E(X X^T)$ dimana E adalah Fungsi Ekspektasi.

D adalah matrik diagonal dari eigenvalue, H adalah matrik eigen vector.

$$W=HD^{-1/2}H^T$$

2. Estimasi himpunan {Qz} dari matrik cummulant

Jika terdapat vector z berukuran n X 1 dan matrik M berukuran n X n, matrik cummulant adalah

$$Q_z(M) = E\{(z^T M z) z z^T\} - R_z \text{tr}(M R_z) - R_z M_z - R_z M^T R_z \quad (3)$$

Dimana tr menyatakan trace dari matrik.

3.Optimisasi orthogonal contrast

Menemukan matrik rotasi U dimana matrik cummulant seditagonal mungkin.

4. Pemisahan

Estimasi A sebagai $V=UW^{-1}$ dan sinyal asli sebagai $V=U^{-1}X$.

Jika estimasi sinyal asli menggunakan ICA telah menghasilkan estimasi sinyal maka dilakukan identifikasi sinyal asli apakah merupakan sinyal noise, artefak atau memang sinyal sesungguhnya yang diinginkan. Rekonstruksi sinyal campuran dilakukan dengan mengeset seluruh row yang merepresentasikan sinyal gangguan noise atau artefak dengan 0.

$$X'=UV' \quad (4)$$

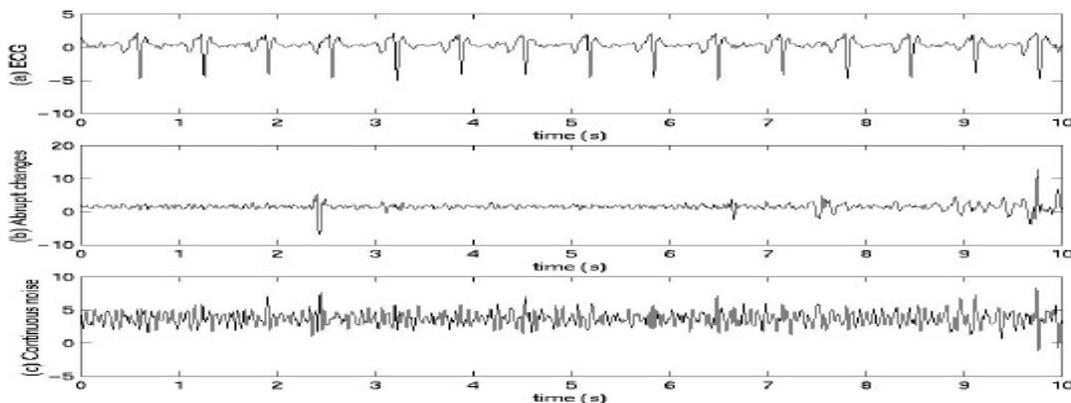
Dengan V'

$$\begin{bmatrix} V_{11} & V_{12} & V_{1N} \\ 0 & 0 & 0 \\ V_{31} & V_{32} & V_{3N} \end{bmatrix}$$

Misalkan row kedua adalah noise atau artefak dan matrik $V_{ij}(i,j=1,..N)$ adalah elemen dari matrik V dan N merepresentasikan jumlah sinyal.

2.2 Identifikasi komponen noise dan artefak

Berdasarkan morfologi bentuk sinyal dapat diketahui bahwa terdapat 3 jenis sinyal yaitu normal, noise kontinyu dan *abrupt change*. Seperti terlihat pada gambar sinyal ECG (*Electro Cardiogram*). Untuk mengidentifikasi noise kontinyu digunakan kurtosis sedangkan untuk mengidentifikasi artefak (*abrupt change*) digunakan variansi,



Gambar 1. Kategori Sinyal

2.2.1 Kurtosis

Kurtosis adalah cumulant orde keempat yang merupakan perhitungan statistik orde tinggi. Untuk sinyal X, rumus kurtosis adalah sebagai berikut:

$$Kur(X_{z^T} - 3[E(x^2)]^2) \quad (5)$$

Nilai kurtosis pada sinyal gaussian adalah 0. Untuk sinyal noise nilai kurtosis akan mendekati 0. Oleh karena itu, untuk mengidentifikasi apakah sinyal adalah noise atau tidak, nilai treshold ditentukan. Taigang[2]

menyatakan bahwa diperlukan beberapa kali percobaan untuk mencari treshhold kurtosis noise. Jika nilai kurtosis sinyal di bawah treshhold maka sinyal tersebut dapat dikategorikan sebagai noise.

2.2.2 Variansi

Untuk sinyal X dengan N sample, variansi dihitung sebagai:

$$\text{Var}_x = \sum_{n=1}^N [x(n) - \overline{x(n)}]^2 \quad (6)$$

Dimana $\overline{x(n)}$ adalah nilai mean dari x(n).

Abrupt change seringkali merupakan sinyal yang terjadi dalam waktu yang pendek (*short transient*). Untuk mendeteksi adanya sinyal termasuk *abrupt changes*, sinyal dibagi menjadi beberapa blok. Setiap blok kemudian dihitung nilai variansi. Kemudian dihitung kembali Nilai variansi dari himpunan variansi tersebut (var_{var}). Sinyal abrupt change ditunjukkan dengan nilai var_{var} yang lebih besar dari var_{var} sinyal biasa. Oleh karena itu perlu ditentukan nilai treshhold untuk var_{var} tersebut. Taigang [2] menyatakan nilai var_{var} dari abrupt changes lebih besar dari nilai treshhold yang ditentukan dari percobaan.

3. PERCOBAAN

3.1 Eksprimental Setup

3.1.1 Tujuan percobaan

1. Membuang komponen sinyal tercampur yang diidentifikasi sebagai noise dan melakukan rekonstruksi sinyal tercampur dari sinyal-sinyal yang bukan noise
2. Membuang komponen sinyal tercampur yang diidentifikasi sebagai artefak dan melakukan rekonstruksi sinyal tercampur dari sinyal-sinyal yang bukan artefak

3.1.2 Data

Penyiapan data dalam penelitian meliputi penyiapan sinyal asli yang diketahui serta proses pencampuran sinyal-sinyal asli ditambah dengan sinyal *gaussian noise* atau sinyal artefak. Data awal adalah suara musik (file source 5.wav), sirine ambulans (file source4.wav), noise (file noise.wav) yang dibangkitkan oleh fungsi matlab dengan power 0 db, serta sinyal pengganggu singkat sebagai artefak (file artefak.wav). Sinyal-sinyal tersebut mempunyai periode waktu 5 detik dengan frekuensi sampling 8 KHz.

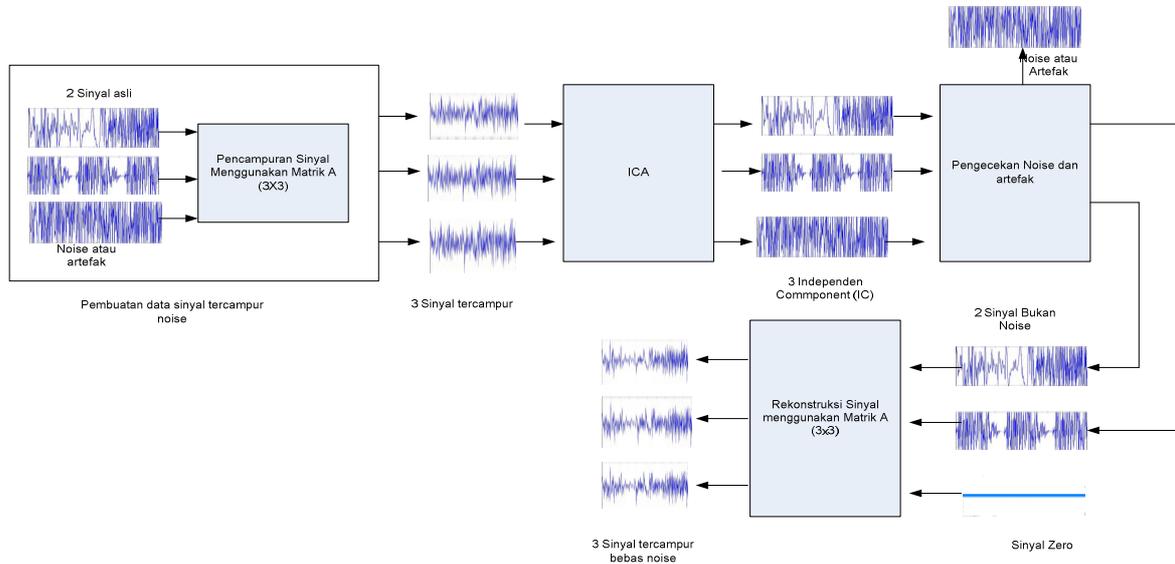
Pembuatan sinyal tercampur

Proses pencampuran 3 sinyal tersebut dilakukan dengan menggunakan matrik pencampuran berukuran 3X 3 yang dibangkitkan secara acak. Berdasarkan tujuan percobaan, ditentukan dua jenis pencampuran sinyal yaitu:

1. Pencampuran sinyal musik dan sirine ambulan ditambah sinyal noise
Pencampuran tersebut menghasilkan tiga buah sinyal tercampur noise. Nilai treshhold kurtosis dapat ditentukan dengan beberapa percobaan pengamatan hasil ICA pada sinyal tercampur dengan matrik pencampur yang berbeda-beda.
2. Pencampuran sinyal musik dan sirine ambulan ditambah sinyal artefak
Pencampuran tersebut menghasilkan tiga buah sinyal tercampur artefak. Nilai treshhold artefak dapat ditentukan dengan beberapa percobaan pengamatan hasil ICA pada sinyal tercampur dengan matrik pencampur yang berbeda-beda.

3.1.3 Blok Diagram Percobaan

Berikut adalah blok diagram untuk eksperimental setup pembuangan noise dan artefak.

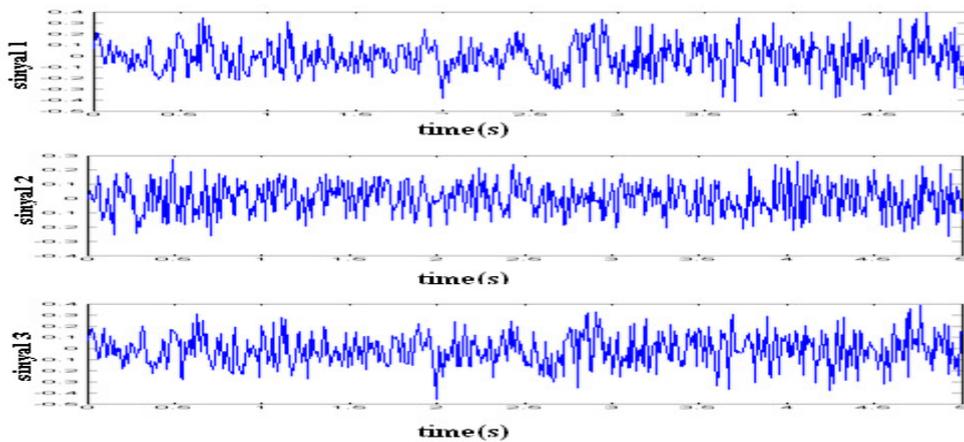


Gambar 2. Blok Diagram Eksperimental Setup Pembuangan Noise dan Artefak

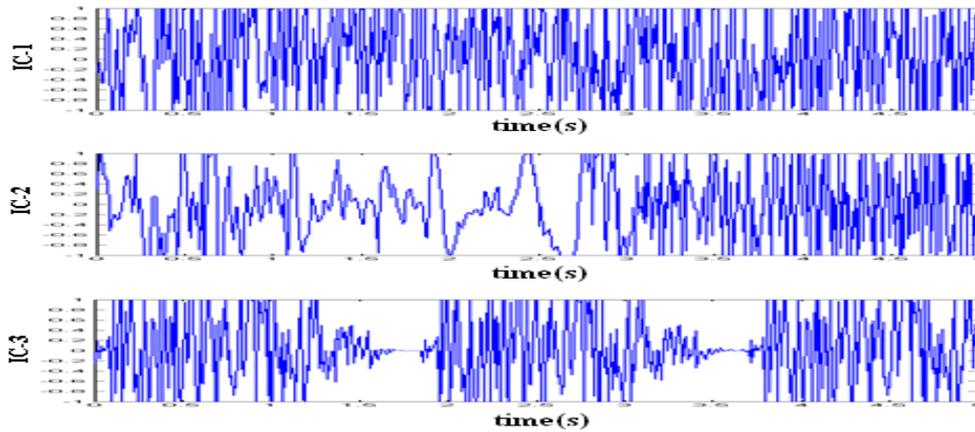
3.2 Hasil Eksperimen

3.2.1 Contoh rekonstruksi sinyal tercampur yang mengandung noise gaussian

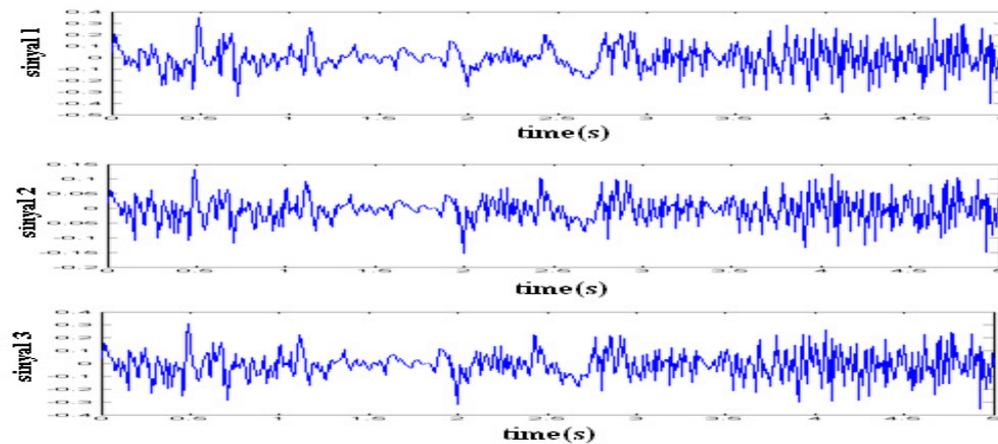
Nilai treshold kurtosis 1.7. Nilai tersebut didapatkan dari beberapa percobaan.



Gambar 3. Sinyal mixture awal mengandung noise



Gambar 4. ICA Component hasil estimasi percobaan 1



Gambar 5. Sinyal mixture bebas noise hasil rekonstruksi

Tabel 1. Kurtosis dan Var of var estimasi ICA percobaan 1

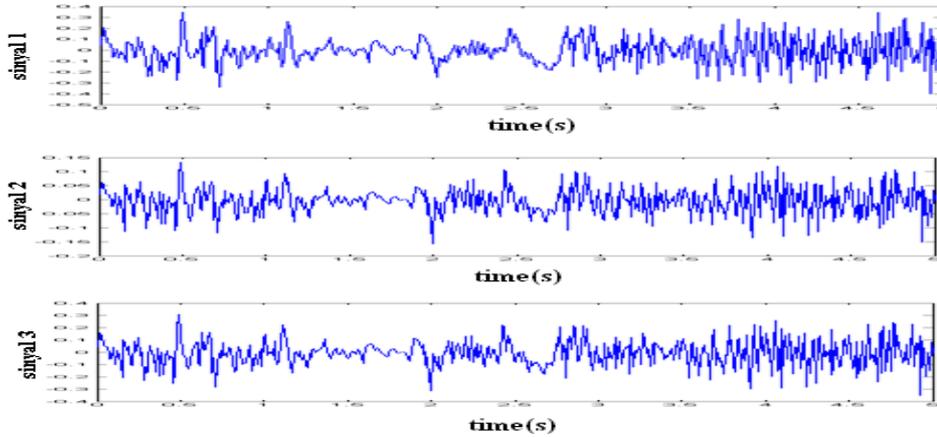
Estimasi hasil ICA	Kurtosis estimasi	Var of var estimasi
IC -1	1,61	1,52
IC-2	3,945	0.34
IC-2	4,27	0.28

Berdasarkan Tabel 1, sinyal *independent component* (IC) ke 1 pada Gambar 4 merupakan sinyal noise karena nilai kurtosisnya 1.6. Hasil rekonstruksi seperti Gambar 5 memperlihatkan bahwa sinyal tercampur yang bebas noise lebih jelas dibandingkan sinyal tercampur yang mengandung noise seperti Gambar 3. Berdasarkan persepsi suara manusia sinyal tercampur hasil rekonstruksi juga lebih jernih dibandingkan dengan sinyal tercampur awal yang masih mengandung noise.

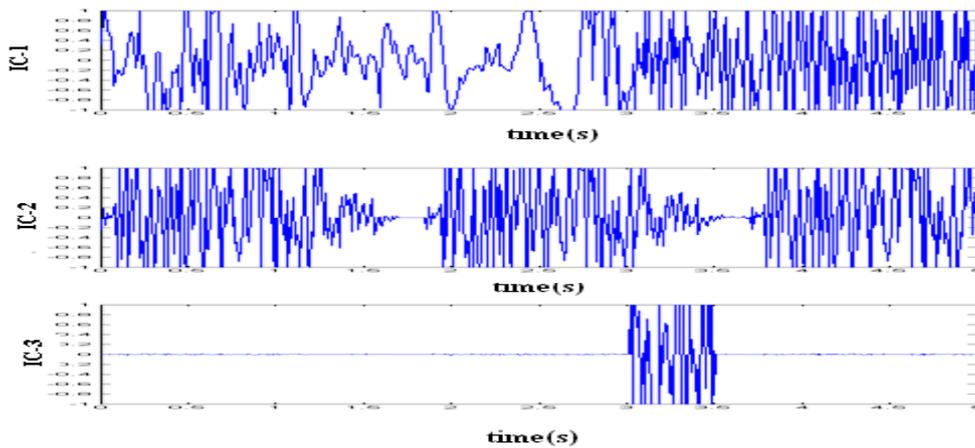
Threshold kurtosis secara umum dapat ditentukan 1.7 karena power noise yang ditambahkan pada percobaan adalah 0 db. Semakin besar power maka nilai kurtosis semakin kecil. Tabel 1 memperlihatkan bahwa estimasi noise dari ICA mempunyai power yang lebih besar dibandingkan sinyal noise asli.

3.2.2 Contoh rekonstruksi Sinyal tercampur yang mengandung Artefak

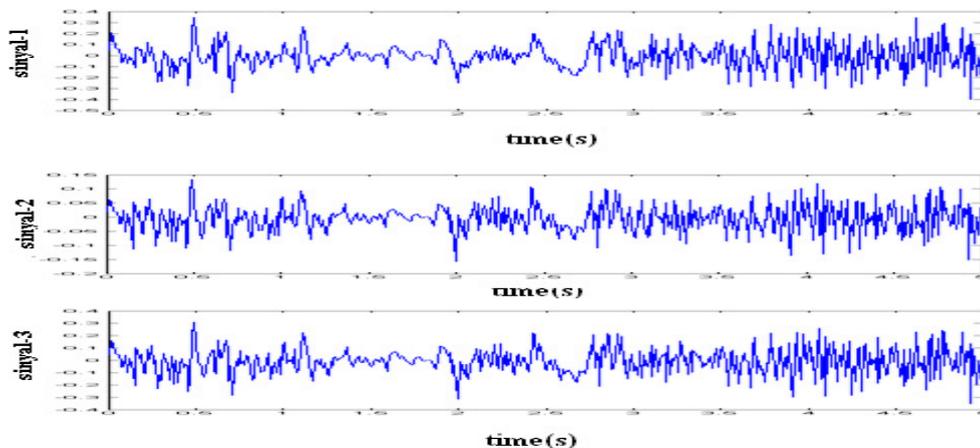
Nilai treshold var of var= 9



Gambar 6. Sinyal Mixture Awal mengandung artefak



Gambar 7. . ICA Component hasil estimasi percobaan 2



Gambar 8. Sinyal mixture bebas artefak hasil rekonstruksi

Tabel 2. Kurtosis dan Var of var estimasi ICA percobaan 2

Estimasi hasil ICA	Kurtosis estimasi	Var of var estimasi
IC -1	4.09	0.25
IC-2	6.14	0.64
IC-2	93.64	9.14

Berdasarkan Tabel 2, sinyal *independent component* (IC) ke 3 pada Gambar 7 merupakan sinyal artefak karena nilai var of var 9.14. Hasil rekonstruksi seperti Gambar 8 memperlihatkan bahwa sinyal tercampur yang bebas artefak mempunyai amplitudo yang lebih kecil daripada sinyal tercampur yang mengandung artefak seperti Gambar 6. Sinyal artefak menyebabkan amplitudo sinyal tercampur menjadi lebih besar dari sinyal tercampur tanpa artefak karena pengaruh power sinyal artefak pada saat pencampuran sinyal.

Tabel 3. Variansi sinyal pada periode 3-3.5 detik

	variansi periode (3-3.5 s)	
	Sinyal tercampur artefak	Sinyal rekonstruksi bebas artefak
Sinyal 1	0.0276	0.0114
Sinyal 2	0.0158	0.0014
Sinyal 3	0.0242	0.0079

Berdasarkan Tabel 3, pada pengamatan di periode 3-3.5 detik (periode dimana sinyal artefak mulai mengganggu) juga terlihat bahwa variansi sinyal tercampur artefak lebih besar dari pada sinyal tercampur bebas artefak.

Threshold artefak yang ditemukan 9 tergantung juga dari perbandingan variansi artefak dengan sinyal yang lain. Jika power artefak lebih besar maka sinyal tersebut dapat diidentifikasi sebagai artefak, namun sebaliknya jika lebih kecil dibandingkan dengan sinyal yang lain maka sinyal tersebut akan gagal ditentukan sebagai artefak. Jadi untuk threshold artefak pada kasus satu bisa berbeda dengan threshold pada kasus yang lain.

4. KESIMPULAN

1. Penghilangan noise pada sinyal tercampur yang mengandung noise gaussian berhasil dilakukan. Nilai threshold kurtosis yang didapatkan dari hasil percobaan adalah 1.7
2. Penghilangan artefak pada sinyal tercampur yang mengandung artefak berhasil dilakukan. Nilai threshold artefak untuk data yang tersedia pada percobaan adalah 9.

5. REFERENSI

- [1]. Taigang He, Gari Clifford and Lionel Tarassenko, Application of ICA in Removing Artefacts from the ECG, University of Oxford, Oxford, UK
- [2]. Aapo Hyvärinen and Erkki Oja, Independent Component Analysis: Algorithms and Applications, Helsinki University of Technology, Neural Networks, 13(4-5):411-430, 2000.
- [3]. Klaus Nordhausen, Jean-Francois Cardoso, Hannu Oja, Esa Ollila, JADE and ICA performance criteria.
- [4]. M. Sheikh M. Algunaidi¹, M. A. Mohd Ali¹, K. B. Gan¹, E. Zahedi², Fetal Heart Rate Monitoring Based on Adaptive Noise Cancellation and Maternal QRS Removal Window, ISSN 1450-216X Vol.27 No.4 (2009), pp.565-575, 1. Faculty of Engineering and Built Environment, University Kebangsaan Malaysia (UKM), Bangi, Malaysia, 2. SHARIF University of Technology Tehran, Iran.