

# PENGINDERAAN KOMPRESIF BERBASIS SPARSITY AVERAGING REWEIGHTED ANALYSIS UNTUK KOMPRESI SINYAL AUDIO DAN SPEECH

Rahmawan Ilham Al Fatha<sup>1</sup>, Jangkung Raharjo<sup>2</sup>, Jaspas Hasudungan<sup>3</sup>

<sup>1,2,3</sup> Universitas Telkom, Bandung

rahmawaniham@student.telkomuniversity.ac.id<sup>1</sup>,  
jangkungraharjo@telkomuniversity.ac.id<sup>2</sup>, jhmanurun@telkomuniversity.ac.id<sup>3</sup>

---

## Abstrak

Data audio di era digital ini sangat umum digunakan baik untuk kebutuhan personal maupun industri. Namun, ukuran data dalam proses transmisi dan penyimpanan adalah bagian yang penting karena masalah ukuran datanya yang sering kali memiliki ukuran yang besar. Kompresi data audio yang efisien menyelesaikan masalah ukuran data untuk kebutuhan penyimpanan dan transmisi ini. *Compressed sensing* (CS) memperkenalkan teknik akuisisi sinyal yang melampaui efisiensi teori *Nyquist* pada umumnya dalam pengambilan sampel. Salah satu metode CS adalah *sparsity averaging reweighted analysis* (SARA) yang diusulkan untuk meningkatkan kinerja dari metode *basis pursuit denoise* (BPDN). Kemudian, terdapat metode *source separation via reweighted analysis* (SSRA) diusulkan untuk data sinyal satu dimensi. Metode SSRA tidak melakukan rata-rata *basis sparsity* seperti yang dilakukan pada metode SARA, sehingga jurnal ini melakukan analisis untuk kinerja SARA yang dimodifikasi mengikuti prinsip SSRA pada sinyal satu dimensi.

Jurnal ini menganalisis kinerja SARA pada data sinyal senandung suara dan sinyal musik hasil segmentasi yang terdiri dari *reff/chorus* dan *verse* dari musik bergenre *jazz* dan *reggae* dengan jenis file wav. Dimana nantinya akan dilakukan pengujian terhadap 4 file audi yaitu *jazz.wav* dan *reggae.wav* yang mewakili data music, kemudian *ep2.wav* dan *ep3.wav* untuk jenis file senandung suara. SARA terdiri dari 2 tahap, yaitu tahap awal untuk penentuan *basis sparsity* dengan mengusulkan basis baru dari hasil rata-rata dari beberapa basis dan proses *reweighted* dari metode rekonstruksi BPDN. Hasil data audio yang paling bagus diperoleh file audio *Jazz* dengan nilai *Signal to Noise Ratio* (SNR) sebesar 56.75 dB, nilai *Object Different Grade ODG* yang dihasilkan sebesar -1.643 dan nilai *Structural Similary Index Measure* (SSIM) sebesar 0.9997. Sedangkan, performansi data *speech* terbaik diperoleh file *Ep2* dengan nilai SNR sebesar 45.10 dB, nilai ODG sebesar -3.948 dan SSIM sebesar 0.9981.

**Keywords:** *Compressed sensing*, *sparsity averaging*, *reweighted analysis*, *basis pursuit denoise*, *spread spectrum*, *wavelet*

---

## 1. Pendahuluan

Pengolahan data audio memainkan peran penting dalam kehidupan modern saat ini baik itu untuk penggunaan pribadi maupun dalam bidang industri. Namun, ukuran data pada proses transmisi maupun penyimpanannya merupakan bagian yang kritis karena masalah pada besarnya ukuran penyimpanan data

---

*Email addresses:* rahmawaniham.student@telkomuniversity.ac.id (Rahmawan Ilham Al Fath),  
jangkungraharjo@telkomuniversity.ac.id (Jangkung Raharjo), jhmanurun@telkomuniversity.ac.id (Jaspas Hasudungan)

May 30, 2021

audi dengan mempertahankan kualitas yang baik. Karena banyaknya file yang dibutuhkan membuat pengelolaan data audio menjadi tantangan [1]. Kompresi data audio yang efisien dapat menyelesaikan masalah ukuran data audio dan transmisi ini.

Berdasarkan pada penelitian sebelumnya, ada banyak transformasi *wavelet* yang dapat di gunakan untuk melakukan dekomposisi pada sinyal. Beberapa contoh *mother wavelet* adalah *Haar, Daubechies, Symlet, Biorthogonal, Meyer, Coiflets, Morlet* dll dimana di temukan bahwa jenis *wavelet Daubechies* memiliki performa terbaik. Sehingga TA ini melakukan analisis kinerja SARA pada data sinyal audio dan *speech*. SARA terdiri dari 2 tahap, yaitu tahap awal untuk penentuan *basis sparsity* dengan mengusulkan proses penghitungan rata-rata dari beberapa *frame basis wavelet Daubechies* dan proses *reweighted* dari metode rekonstruksi BPDN. Pengujian parameter SARA pada Tugas Akhir ini dilakukan dengan membandingkan BPDN versi *unreweighted* atau BPSA dengan algoritma BPDN versi *Reweighted* yaitu SARA seperti yang dilakukan pada penelitian terkait [2] untuk melihat performa penggunaan BPSA terhadap metode SARA.

## 2. Dasar Teori

### 2.1. Compressive Sensing

*Compressive Sampling* (CS), yang dikenal juga sebagai *Compressive Sensing, Compressed Sensing* atau *Sparse Sampling*, merupakan metode baru dalam proses pencuplikan sinyal. CS merupakan metode pencuplikan sinyal dimana sinyal dapat direkonstruksi dari nilai frekuensi sampling yang jauh lebih kecil dari yang disyaratkan pada teorema Nyquist-Shannon. Sesuai dengan namanya, CS melakukan proses sampling terhadap sinyal yang telah dikompresi. Tujuan dari proses kompresi adalah memperkecil ukuran sinyal sehingga lebih efisien [1].

### 2.2. Basis Pursuit Denoise (BPDN)

Basis Pursuit Denoising (BPDN) berfungsi untuk menerima *noise* dalam pengukuran. BPDN sama dengan BP hanya saja fungsi constraint nya yang berbeda. Berikut persamaan  $l_2$ -norm yang dibatasi dengan  $\epsilon$

$$\hat{x} = \arg \min_x \|x\|_1; \text{ subject to } \frac{1}{2} \|y - A \cdot x\|_2^2 \leq \epsilon \quad (1)$$

Dimana,  $l_2$  juga dikenal sebagai norm euclidean, mewakili panjang atau ukuran vektor.  $A$  dapat berupa *sparse rectangular matrix* untuk menghitung vektor.

### 2.3. Sparsity Averaging Reweighted Analysis (SARA)

SARA terdiri dari 2 proses yaitu proses rata-rata basis  $\Psi$  dari beberapa frame basis *wavelet* dan proses *reweighted* BPDN.

#### 2.3.1. Rata-rata basis Sparsity

Rata-rata *basis sparsity*  $\Psi_i$  tersusun dari rangkaian  $q$  frame ( $1 \leq i \leq q$ ) menghasilkan *Parseval frames*  $\Psi \in \mathbb{C}^{N \times D}$ , dengan dimensi  $M < D$  sesuai persamaan [3]:

$$\Psi = \frac{1}{\sqrt{q}} [\Psi_1, \Psi_2, \dots, \Psi_q]. \quad (2)$$

Dari persamaan (2), selanjutnya dengan analisis norm  $l_0$  menghasilkan rata-rata *sparsity* sesuai persamaan:

$$\|\Psi^\dagger \bar{x}\|_0 = \sum_{i=1}^q \|\Psi^\dagger \bar{x}_i\|_0 \quad (3)$$

2.3.2. Analisis Reweighted BPDN

Sinyal dikatakan jarang (*sparse*) jika hanya mengandung  $K$  koefisien bukan nol dalam penguraiannya, dengan  $K \ll N$ , atau dapat dimampatkan jika rangkaian koefisien yang dipesannya meluruh dengan cepat dan sinyal dapat didekati dengan baik hanya dengan koefisien  $K$  pertama. Persamaan *inverse*:

$$y \equiv \Theta \alpha + n, \text{ dengan } \Theta \equiv \Phi \Psi \in \mathbb{C}^{M \times N}, \tag{4}$$

dengan matriks  $\Theta$  mengidentifikasi basis penginderaan seperti yang terlihat dari ketersebarannya. Biasanya  $M < N$  sehingga proses *inverse* adalah problem yang sulit dipecahkan solusinya.

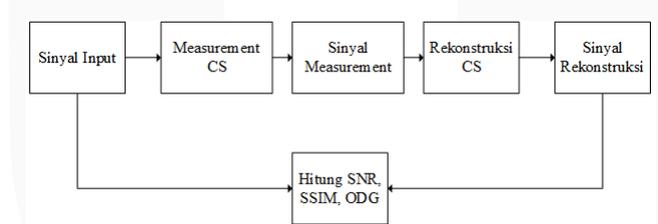
Teori ini juga menawarkan berbagai cara untuk merancang matriks penginderaan yang sesuai, yang secara khusus menunjukkan bahwa sejumlah kecil pengukuran diperlukan relatif terhadap pengambilan sampel Nyquist-Shannon yang naif:  $M \ll N$ . Perhatikan bahwa, dalam teori, minimalisasi  $l_0$  yang eksplisit membutuhkan lebih sedikit pengukuran,  $M \approx 2K$  [4]. Selanjutnya, mari kita tentukan persamaan *weighted l1*:

$$\begin{aligned} \min_{\tilde{x} \in \mathbb{R}^N} & \quad \|\mathbf{W}\Psi^T \tilde{\mathbf{x}}\|_1 \\ \text{subject to } & \quad \|\mathbf{y} - \Phi \tilde{\mathbf{x}}\|_2 \leq \epsilon \\ & \quad \tilde{\mathbf{x}} \geq 0 \end{aligned} \tag{5}$$

3. Sistem Model

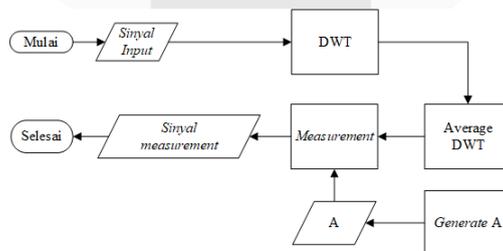
3.1. Deskripsi Sistem

Penelitian ini melakukan kompresi sinyal audio dan speech menggunakan CS. Proses dimulai dengan pemilihan sinyal audio atau speech untuk menjadi sinyal input. Proses CS terdiri dari dua proses utama, yaitu *measurement* CS dan rekonstruksi CS. *Measurement* CS merupakan proses pengukuran yang menghasilkan sinyal *measurement* dimana dimensi dari sinyal *measurement* akan lebih kecil dari segi sampel dibandingkan sinyal input. Rekonstruksi CS merupakan proses untuk mengembalikan sinyal *measurement* ke ukuran semula sesuai sinyal input, proses ini menghasilkan sinyal rekonstruksi. Ilustrasi dapat dilihat pada Gambar 1 dimana hasil parameter SNR, SSIM, dan ODG dihitung dari sinyal input dan sinyal rekonstruksi.

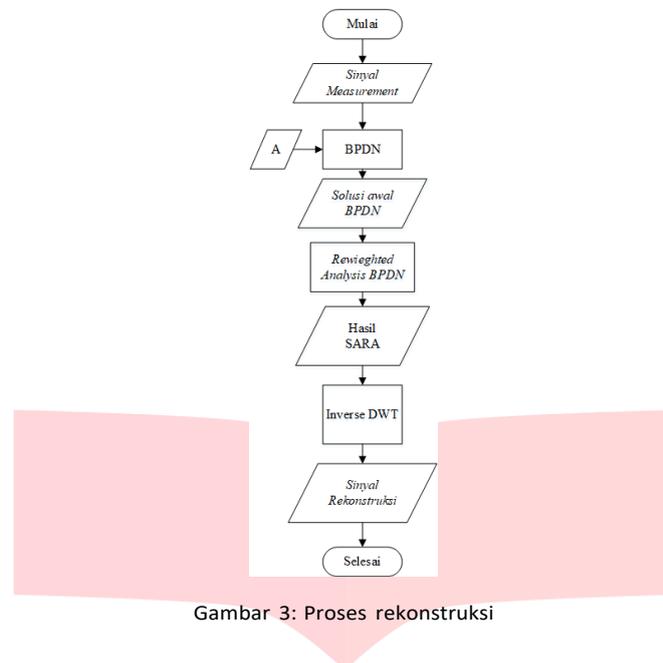


Gambar 1: Blok diagram sistem

3.2. Kompresi Citra dengan CS



Gambar 2: Proses kompresi audio dengan CS



Gambar 3: Proses rekonstruksi

Proses kompresi audio pada penelitian ini menggunakan CS berbasis transformasi DWT. Sinyal audio atau speech sebagai input awal proses transformasi DWT digunakan untuk mendapatkan nilai koefisien frekuensi. DWT digunakan untuk memisahkan frekuensi menjadi frekuensi rendah dan tinggi. Hasil transformasi ini adalah sinyal sparse yang akan menjadi sinyal yang dikompresi. Proses kompresi terjadi pada proses measurement dimana nilai sampling dari sinyal sparse direduksi menjadi ukuran matrik yang lebih kecil sesuai dengan ukuran matriks A. Matriks A adalah matriks pengukuran atau measurement matrix. Matriks A dibangkitkan secara acak. Ilustrasi dapat dilihat pada Gambar 2.

### 3.3. Proses rekonstruksi

Proses untuk mendapatkan Sinyal audio kembali adalah proses *reconstruction*. Proses ini bertujuan untuk melakukan rekonstruksi sinyal yang didapatkan dari sinyal hasil pengukuran CS atau *measurement sampling*. Pada penelitian ini sinyal yang akan direkonstruksi adalah sinyal sparse. Sinyal sparse adalah hasil transformasi DWT dari sinyal audio dan *measurement sampling* menjadi *compressed audio*. Penelitian ini menggunakan metode BPDN dan reweighted BPDN untuk metode rekonstruksi sinyal. Ilustrasi dapat dilihat pada Gambar 3.

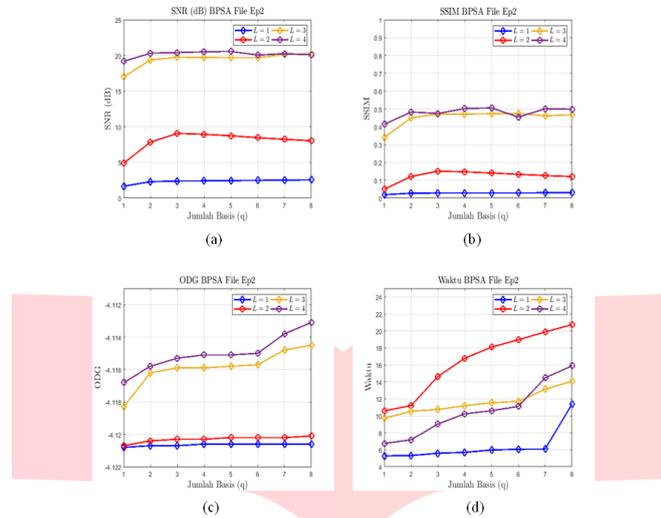
## 4. Simulasi Dan Analisis

### 4.1. Pengaruh Parameter Sparsity Averaging (SA)

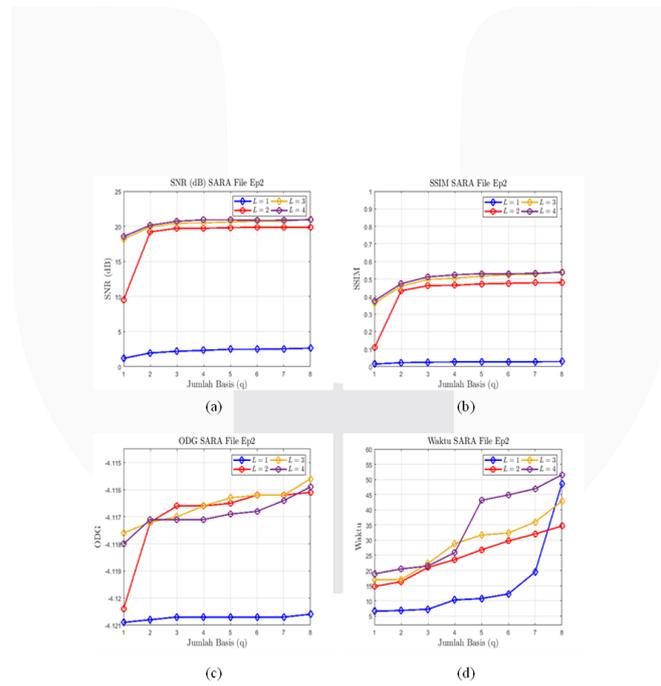
Proses SA pada jurnal ini terdiri dari proses pemilihan basis rata-rata yang terdiri dari jumlah basis dan level dekomposisi. Untuk membatasi jumlah kombinasi parameter, parameter yang digunakan tetap yaitu  $MR = 0,1$  atau 10%, parameter BPDN ( $\gamma = 0,1$ ,  $i_{MAX}^A = 10$ ,  $i_{MAX}^R = 50$ ), dan parameter SARA ( $i_{MAX}^{RW} = 5$ ,  $\epsilon = 0,001$ ). Hasil dari pengujian terdiri dari hasil rekonstruksi Basis Pursuit Sparsity Averaging (BPSA) dan SARA.

#### 4.1.1. Pengaruh Jumlah Basis ( $q$ ) dan Level Dekomposisi ( $L$ ) pada file Ep2

Pengujian ini melakukan analisis pengaruh jumlah basis ( $q$ ) dan level dekomposisi ( $L$ ) terhadap SNR, SSIM, ODG, dan waktu komputasi. Jumlah basis yang diujikan yaitu ( $q = 1, 2, \dots, 8$ ) dan level dekomposisi ( $L = 1, 2, 3, 4$ ). Dari variasi tersebut, hasil performansi pengujian dilaporkan pada Gambar 4 untuk hasil BPSA dan Gambar 5 untuk hasil SARA.



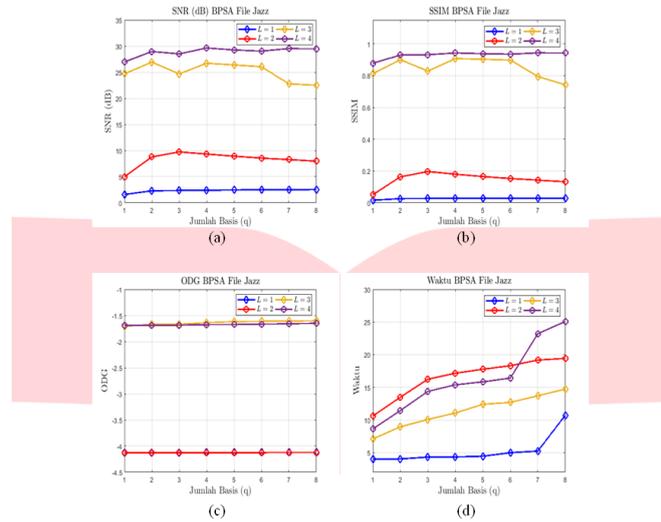
Gambar 4: Hasil BPSA Pengujian Pengaruh Jumlah Basis ( $q$ ) dan Level Dekomposisi ( $L$ ).  
 (a) Hasil SNR. (b) Hasil SSIM. (c) Hasil ODG. (d) Hasil waktu komputasi.



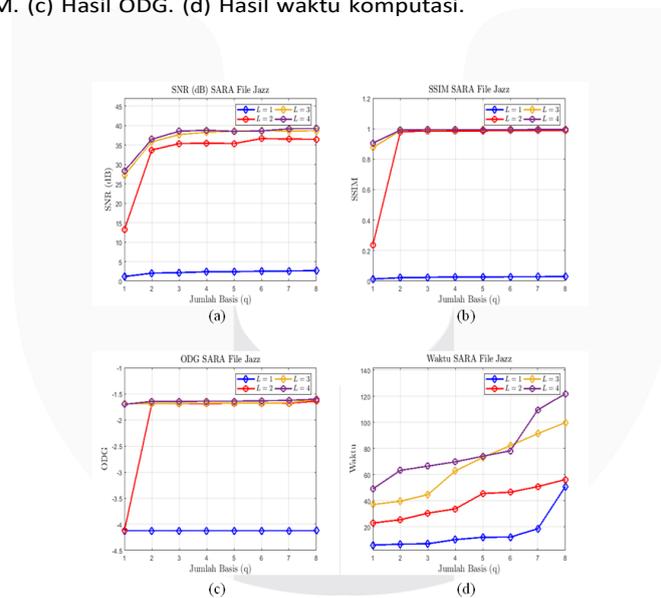
Gambar 5: Hasil SARA Pengujian Pengaruh Jumlah Basis ( $q$ ) dan Level Dekomposisi ( $L$ ).  
 (a) Hasil SNR. (b) Hasil SSIM. (c) Hasil ODG. (d) Hasil waktu komputasi.

4.1.2. Pengaruh Jumlah Basis ( $q$ ) dan Level Dekomposisi ( $L$ ) pada file Jazz

Pengujian ini melakukan analisis pengaruh jumlah basis ( $q$ ) dan level dekomposisi ( $L$ ) terhadap SNR, SSIM, ODG, dan waktu komputasi. Jumlah basis yang diujikan yaitu ( $q = 1, 2, \dots, 8$ ) dan level dekomposisi ( $L = 1, 2, 3, 4$ ). Dari variasi tersebut, hasil performansi pengujian dilaporkan pada Gambar 6 untuk hasil BPSA dan Gambar 7 untuk hasil SARA.



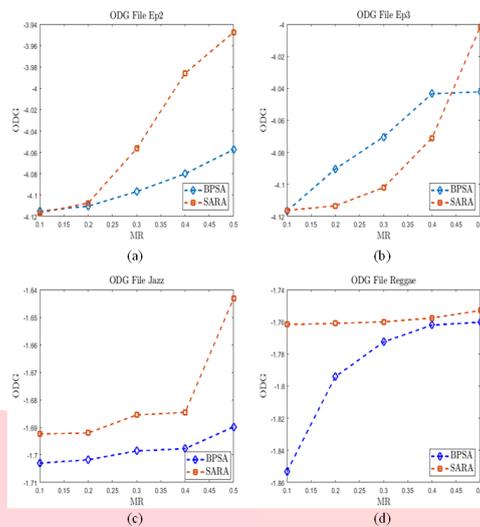
Gambar 6: Hasil BPSA Pengujian Pengaruh Jumlah Basis ( $q$ ) dan Level Dekomposisi ( $L$ ). (a) Hasil SNR. (b) Hasil SSIM. (c) Hasil ODG. (d) Hasil waktu komputasi.



Gambar 7: Hasil SARA Pengujian Pengaruh Jumlah Basis ( $q$ ) dan Level Dekomposisi ( $L$ ). (a) Hasil SNR. (b) Hasil SSIM. (c) Hasil ODG. (d) Hasil waktu komputasi.

4.2. Pengaruh Measurement Rate

Pada pengujian pengaruh ini, jumlah basis basis yang digunakan adalah  $q = 8$ , level dekomposisi  $L = 4$  dan jumlah data audio 4 (2 data *speech* dan 2 data audio). Pengujian tetap dilakukan dengan membandingkan SARA dengan BPSA. Hasil pengujian ini dilaporkan dalam SNR, SSIM, ODG dan waktu komputasi.



Gambar 8: Hasil ODG Pengujian MR.

Tabel 1: Hasil ODG Pengujian MR.

MR	Ep2		Ep3		Jazz		Reggae	
	BPSA	SARA	BPSA	SARA	BPSA	SARA	BPSA	SARA
0.1	-4.115	-4.117	-4.117	-4.116	-1.703	-1.692	-1.853	-1.762
0.2	-4.110	-4.107	-4.090	-4.113	-1.702	-1.692	-1.794	-1.761
0.3	-4.097	-4.056	-4.070	-4.102	-1.699	-1.685	-1.772	-1.760
0.4	-4.080	-3.986	-4.043	-4.071	-1.698	-1.685	-1.762	-1.758
0.5	-4.057	-3.948	-4.042	-4.002	-1.690	-1.643	-1.760	-1.753

4.2.1. ODG

Pengujian ODG dilakukan dengan mengubah nilai MR sebesar 0.1 sampai 0.5. Hasil pengujian dilaporkan pada Gambar 8 dan lebih detail pada Tabel 1. Hasil menunjukkan algoritma SARA mengungguli algoritma lain yang menjadi dasar dari CS pada sinyal audio yaitu algoritma BPSA.

4.2.2. SNR

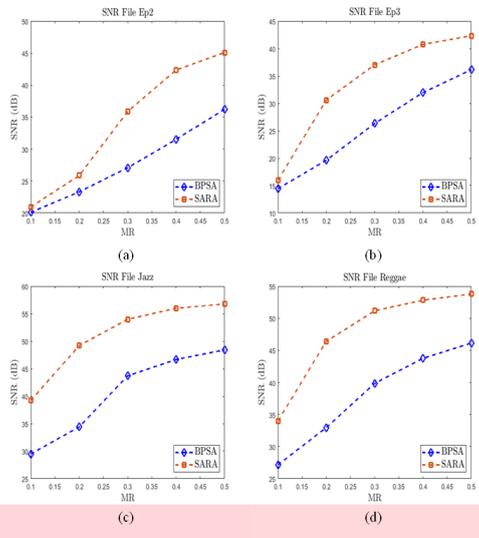
Pengujian SNR dilakukan dengan mengubah nilai MR sebesar 0.1 sampai 0.5. Hasil pengujian dilaporkan pada Gambar 9 dan lebih detail pada Tabel 2. Hasil menunjukkan algoritma SARA mengungguli algoritma lain yang menjadi dasar dari CS pada sinyal audio yaitu algoritma BPSA.

4.2.3. SSIM

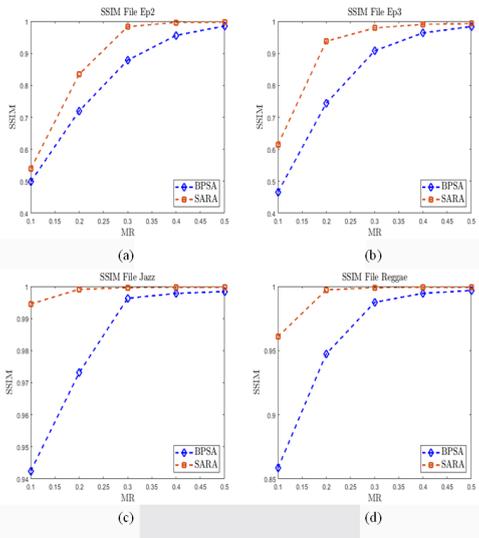
Pengujian SSIM dilakukan dengan mengubah nilai MR sebesar 0.1 sampai 0.5. Hasil pengujian dilaporkan pada Gambar 10 dan lebih detail pada Tabel 3. Hasil menunjukkan algoritma SARA memiliki performa yang lebih baik pada setiap nilai MR yang diuji jika dibandingkan dengan algoritma BPSA.

Tabel 2: Hasil SNR Pengujian MR.

MR	Ep2		Ep3		Jazz		Reggae	
	BPSA	SARA	BPSA	SARA	BPSA	SARA	BPSA	SARA
0.1	20.11	20.95	14.45	15.97	29.48	39.20	27.15	33.98
0.2	23.31	25.87	19.63	30.61	34.40	49.19	32.95	46.41
0.3	27.08	35.87	26.33	37.03	43.70	53.97	39.85	51.16
0.4	31.51	42.36	31.97	40.76	46.67	56.00	43.75	52.82
0.5	36.23	45.10	36.14	42.34	48.42	56.75	46.13	53.78



Gambar 9: Hasil SNR Pengujian MR.

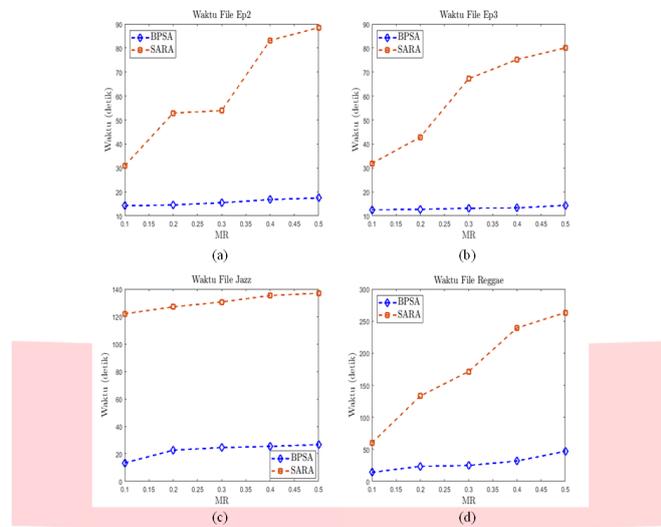


Gambar 10: Hasil SSIM Pengujian MR.

Tabel 3: Hasil SSIM Pengujian MR.

MR	Ep2		Ep3		Jazz		Reggae	
	BPSA	SARA	BPSA	SARA	BPSA	SARA	BPSA	SARA
0.1	0.4985	0.5390	0.4659	0.6143	0.9424	0.9945	0.8584	0.9608
0.2	0.7198	0.8348	0.7439	0.9379	0.9731	0.9991	0.9474	0.9973
0.3	0.8784	0.9836	0.9077	0.9794	0.9963	0.9996	0.9876	0.9990
0.4	0.9561	0.9965	0.9639	0.9907	0.9978	0.9997	0.9946	0.9993
0.5	0.9854	0.9981	0.9838	0.9932	0.9984	0.9997	0.9968	0.9994

4.2.4. Waktu Komputasi



Gambar 11: Hasil Waktu Komputasi Pengujian MR.

Pengujian waktu komputasi dari semua data saat menggunakan SARA pada  $MR = 0.1$  sampai dengan  $MR = 0.5$  didapat nilai yang lebih tinggi jika dibandingkan dengan BPSA. Hal ini ditunjukkan secara detail oleh Tabel 4 dan Gambar 11. Hasil ini menunjukkan bahwa SARA mempunyai performa waktu komputasi yang lebih lama dibandingkan BPSA karena adanya proses *reweighted* saat sinyal direkonstruksi.

Tabel 4: Hasil Waktu Komputasi (Detik) Pengujian MR

MR	Ep2		Ep3		Jazz		Reggae	
	BPSA	SARA	BPSA	SARA	BPSA	SARA	BPSA	SARA
0.1	14.22	30.89	12.53	31.89	13.43	121.96	14.07	60.11
0.2	14.53	52.89	12.72	42.84	22.66	127.02	23.43	133.12
0.3	15.48	53.91	13.20	67.20	24.60	130.55	24.72	171.24
0.4	16.76	83.17	13.32	75.19	25.46	135.27	31.76	239.22
0.5	17.48	88.49	14.37	80.07	26.77	136.94	46.84	263.24

4.3. Pengujian Mean Opinion Score (MOS)

Mean Opinion Score (MOS) merupakan parameter subjektif untuk menilai kualitas pemrosesan dengan berdasarkan panca indra manusia, sehingga subjektifitas sangat tergantung pada persepsi responden dalam memberikan penilaian. Penilaian dilakukan dengan mendengarkan audio sebelum dan sesudah pemrosesan dengan parameter terbaik kepada 25 orang responden untuk masing masing metode yaitu BPSA dan SARA. Dengan skala nilai 1 untuk sangat buruk dan 5 untuk sangat baik, berikut hasil rata rata nilai MOS yang di dapat dari ke 25 responden :

Tabel 5: Hasil Rata-Rata Pengujian MOS.

Metode	File Audio			
	Ep2	Ep3	Jazz	Reggae
BPSA	2,52	2,12	3,68	3,24
SARA	3,52	3,08	4,52	4,08

## 5. Kesimpulan

Pada penelitian ini, telah dirancang suatu sistem untuk kompresi data audio menggunakan metode CS berbasis algoritma SARA. Adapun kesimpulan yang didapat dari hasil pengujian adalah

1. Hasil ODG, SNR dan SSIM menunjukkan parameter terbaik adalah menggunakan jumlah basis  $q = 8$ , level dekomposisi  $L = 4$ . Sedangkan, waktu komputasi yang didapat linier terhadap  $q$  dan  $L$ , dimana waktu komputasi semakin lama untuk nilai  $q$  dan  $L$  yang semakin besar.
2. Metode *reweighted* BPDN menggunakan SARA mengungguli metode *unreweighted* -nya yaitu BPSA pada nilai  $MR = 0.1$  sampai  $MR = 0.5$  jika dilihat pada hasil ODG, SNR dan SSIM.
3. Penggunaan metode SARA pada data Jazz dan Reggae yang berjenis audio memiliki performa ODG, SNR dan SSIM yang lebih baik jika dilakukan kompresi menggunakan CS dibandingkan data Ep2 dan Ep3 yang berjenis *speech*. Hasil data audio yang paling bagus diperoleh audio Jazz dengan nilai SNR sebesar 56.75 dB, nilai ODG yang dihasilkan sebesar -1.643 dan nilai SSIM sebesar 0.9997. Sedangkan, performansi data *speech* terbaik diperoleh file Ep2 dengan nilai SNR sebesar 45.10 dB, nilai ODG sebesar -3.948 dan SSIM sebesar 0.9981.
4. Pada pengujian MOS metode SARA memiliki kualitas yang relatif lebih baik dari pada metode BPSA, dimana setiap file audio yang diproses menggunakan metode SARA memiliki peningkatan nilai parameter MOS yang lebih baik sehingga noise yang terdengar semakin berkurang dan audio yang terdengar lebih jelas.

## REFERENSI

- [1] D. L. Donoho, Compressed sensing, IEEE Transactions on Information Theory 52 (4) (2006) 1289–1306.
- [2] R. E. Carrillo, J. D. McEwen, D. Van De Ville, J.-P. Thiran, Y. Wiaux, Sparsity averaging for compressive imaging, IEEE Signal Processing Letters 20 (6) (2013) 591–594.
- [3] R. E. Carrillo, J. D. McEwen, Y. Wiaux, Sparsity Averaging Reweighted Analysis (SARA): a novel algorithm for radio-interferometric imaging, Monthly Notices of the Royal Astronomical Society 426 (2) (2012) 1223–1234.
- [4] E. J. Candes, T. Tao, Near-optimal signal recovery from random projections: Universal encoding strategies?, IEEE Transactions on Information Theory 52 (12) (2006) 5406–5425.