

PERANCANGAN DAN ANALISIS SINKRONISASI PADA WATERMARKING AUDIO STEREO BERBASIS QIM DENGAN TEKNIK GABUNGAN SWT-DST-QR-CPT

DESIGN AND SYNCRONIZATION ANALYSIS ON AUDIO STEREO WATERMARKING BASED ON QIM WITH COMBINED TECHNIQUES OF SWT-DST-QR-CPT

Syarahbil Pawellang¹, Gelar Budiman, S.T., M.T.², Azizah, S.T.,M.T.³

^{1,2,3}Electrical Engineering Faculty of Telkom University, Bandung, Jawa Barat 40257

¹abil.pawellang@gmail.com, ²gelarbudiman@telkomuniversity.ac.id, ³azizah@telkomuniversity.ac.id

ABSTRAK

Audio watermarking adalah suatu teknik menyisipkan suatu informasi kedalam suatu *file audio* sehingga orang lain tidak menyadari adanya informasi tambahan tersebut. Pada *paper* ini penulis akan merancang *audio watermarking* berbasis *Quantization Index Modulation (QIM)* dengan teknik gabungan *Stationary Wavelet Transform (SWT)*, *Discrete Sine Transform (DST)*, *QR Decomposition* dan *Cartesian Polar Transform (CPT)* menggunakan sinkronisasi berbasis *Spread Spectrum (SS)*. Pada metode *SWT*, *host audio* akan didekomposisi untuk dipilih *subband* yang akan digunakan dan memilih koefisien *SWT*. Hasil domain frekuensi *SWT* akan diubah kedalam domain waktu menggunakan metode *DST*. Selanjutnya matriks yang akan dipilih akan diubah kedalam matriks ortogonal dan matriks segitiga menggunakan metode *QR*. Koefisien akan ditransformasi menggunakan *CPT* agar dapat tahan terhadap berbagai serangan. Untuk mendeteksi posisi awal watermark akibat serangan sinkronisasi, penulis memasukan bit sinkronisasi. Setelah itu, proses *embedding* dilakukan dengan menggunakan metode *QIM*. Penelitian ini menghasilkan watermarking yang memiliki nilai rata-rata *SNR* sebesar 32.718 dB dan *BER* sebesar 0.2058.

Kata Kunci : *Audio Watermarking*, Sinkronisasi, *Quantization Index Modulation (QIM)*, *Stationary Wavelet Transform (SWT)*, *Discrete Sine Transform (DST)*, *Quick Response (QR)*, *Cartesian Polar Transform (CPT)*, *Spread Spectrum (SS)*.

ABSTRACT

Audio watermarking is a technique of inserting information into an audio file so humans are not aware of the existence of such additional information. In this paper that we propose audio watermarking based on *Quantization Index Modulation (QIM)* with a combined techniques of *Stationary Wavelet Transform (SWT)*, *Discrete Sine Transform (DST)*, *QR Decomposition*, *Cartesian and Polar Transform (CPT)* using synchronization based on *Spread Spectrum (SS)*. On the *SWT* decomposition, *host audio* will be decomposed for selected *subbands* and *coefficients* which is used to next method. The results of the time domain will be changed into frequency domain using *DST* method. Furthermore the matrix which is selected will be transformed into a orthogonal matrix and triangular matrix using *QR* method. Coefficient will be transformed using the *CPT* in order to against various attacks. To detect the beginning location of watermark due to sychronization attack, we insert synchronization bit. After that, the process of embedding using *QIM*. This research produces watermarking which have average value of *SNR* 32.718 dB and *BER* 0.2058.

Keywords : *Audio Watermarking*, Synchronization, *Quantization Index Modulation (QIM)*, *Stationary Wavelet Transform (SWT)*, *Discrete Sine Transform (DST)*, *Quick Response (QR)*, *Cartesian Polar Transform (CPT)*, *Spread Spectrum (SS)*.

1. PENDAHULUAN

Pada era globalisasi ini, kemudahan memperoleh informasi sangatlah mudah terutama melalui *internet*. Siapapun dapat saling bertukar informasi melalui *internet*. Kemudahan itu sendiri menjadi tantangan dalam melakukan tindakan kejahatan. Salah satu contoh kejahatan tersebut adalah kepemilikan hak cipta. Untuk mecegah kepemilikan hak cipta tersebut ada suatu teknik yang disebut *watermarking*. *Digital watermarking* adalah salah satu metode efektif yang digunakan untuk mencegah hak cipta dengan cara menyembunyikan data digital dalam bentuk data multimedia seperti gambar, *audio*, *video* tanpa merusak kualitas dari data yang akan disisipkan [1]. *Audio watermarking* adalah suatu teknik menyisipkan suatu informasi kedalam suatu *file audio* sehingga orang lain tidak menyadari adanya informasi tambahan tersebut. Banyak penelitian sebelumnya membahas tentang *audio watermarking*.

Pada jurnal ini akan mengimplementasikan sebuah aplikasi *watermarking* pada *host audio* dengan menggunakan metode yang telah digabungkan dengan tujuan mengambil kelebihan dan mengurangi kelemahan dari masing-masing metode tersebut. Metode *audio watermarking* yang akan dibuat berbasis QIM dengan teknik gabungan SWT, DST, QR, CPT menggunakan sinkronisasi berbasis SS.

2. DASAR TEORI

2.1. Audio Watermarking

Audio watermarking adalah suatu teknik menyisipkan suatu informasi kedalam suatu *file audio* sehingga orang lain tidak menyadari adanya informasi tambahan tersebut. *Watermark* merupakan solusi yang menjanjikan untuk melindungi dari hak kepemilikan suatu media dengan cara menyisipkan suatu data yang hanya pemilik mengetahuinya [3]. Dalam *audio watermarking* terdapat proses yaitu proses *embedding* dan ekstraksi. Proses *embedding* yaitu menyisipkan suatu data watermark kedalam audio yang di-*watermark* (*host audio*). Setelah data ter-*watermark*, data akan diambil kembali yang disebut proses ekstraksi. Namun, dalam pengambilan data kembali, data tersebut kemungkinan akan mendapatkan serangan.

2.2. Quantization Index Modulation

Quantization Index Modulation (QIM) merupakan metode yang pertama kali diusulkan oleh Chen dan Wornell dimana metode ini menerapkan kuantisasi ke dalam *sample host audio* dalam proses penyisipan *watermarking* [4]. QIM merupakan salah satu metode yang memiliki tingkat *robustness* yang baik. Berikut adalah persamaan untuk penyisipan dengan metode QIM :

$$F'(0) = \begin{cases} A_k, & \text{jika } w = 0 \text{ dan } \arg \min |F(0) - A_k| \\ B_k, & \text{jika } w = 1 \text{ dan } \arg \min |F(0) - B_k| \end{cases} \quad (2.1)$$

Dimana :

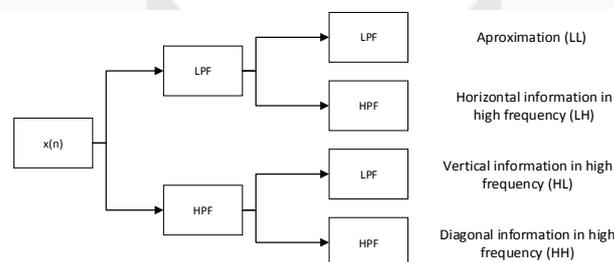
$$A_k = \left(2k + \frac{1}{2}\right)\Delta ; B_k = \left(2k - \frac{1}{2}\right)\Delta ; \text{ dengan } k = 0, \pm 1, \pm 2, \dots$$

Untuk proses ekstraksi, persamaan QIM sebagai berikut :

$$\tilde{V}(k) = \text{mod} \left(\text{ceil} \left(\frac{F'(0)}{\Delta} \right), 2 \right) \quad (2.2)$$

2.3. Stationary Wavelet Transform

Stationary Wavelet Transform (SWT) adalah algoritma transformasi wavelet yang dirancang untuk mengatasi kekurangan dari *Discrete Wavelet Transform* (DWT) yaitu kurangnya *translation invariance*. *Translation invariance* dapat dicapai dengan menghapus *downsampler* dan *upsampler* di DWT dan meningkatkan koefisien filter dalam algoritma dengan faktor 2^{j-1} pada tingkat j_{th} . SWT merupakan skema yang berhubungan dengan redundansi karena output dari setiap level SWT mengandung jumlah sampel yang sama dengan inputnya. Jadi untuk dekomposisi N level terdapat redundansi N pada koefisien *wavelet* [5].



Gambar 2.1 Diagram SWT *filter bank* 2 level [6]

2.4. Discrete Sine Transform

Discrete Sine Transform (DST) sangat identik dengan *Shifted Discrete Fourier Transform* yang merupakan modifikasi permutasi tertentu dari sinyal asli. Berikut merupakan persamaan dari DST [7]:

$$y(k) = \sum_{n=0}^{N-1} x(n) \sin \left(\frac{\pi kn}{N+1} \right), \quad k = 0, \dots, N-1 \quad (2.3)$$

Persamaan ini juga memiliki sifat pemadatan energi yang bisa digunakan untuk audio watermarking. Dibawah ini merupakan persamaan dari *Inverse Discrete Sine Transform* (IDST) [7]:

$$x(n) = \frac{2}{N+1} \sum_{k=0}^{N-1} y(k) \cdot \sin\left(\frac{\pi kn}{N+1}\right), \quad n = 0, \dots, N-1 \quad (2.4)$$

2.5. QR Decomposition

QR *Decomposition* atau juga disebut QR *factorization* merupakan suatu operasi yang digunakan untuk membagi sebuah matriks menjadi matriks ortogonal dan matriks segitiga. Sebuah matriks A, difaktorisasi menggunakan QR *decomposition* dengan persamaan sebagai berikut [8]:

$$A = Q * R \quad (2.5)$$

2.6. Cartesian-Polar Transform

Cartesian-Polar Transform (CPT) adalah suatu transformasi konformasi dari sistem koordinat *Cartesian* ke sistem koordinat *Polar*. Perhatikan sistem koordinat *polar* (r, θ), di mana r menunjukkan jarak antara titik asal dan titik yang ditentukan, dan θ menunjukkan sudut antara garis referensi. Transformasi dari koordinat *Cartesian* ke koordinat *Polar* diberikan oleh persamaan berikut [9]:

$$r = \sqrt{x^2 + y^2}, \quad \theta = \tan^{-1}\left(\frac{y}{x}\right) \quad (2.6)$$

Dimana (x, y) adalah titik dalam sistem koordinat *Cartesian*. Transformasi dari koordinat *Polar* ke koordinat *Cartesian* dapat dituliskan sebagai berikut [9]:

$$x = r \cos \theta, \quad y = r \sin \theta \quad (2.7)$$

2.7. Synchronization Code

Salah satu masalah dalam sinkronisasi yaitu mengikat *watermark* dengan audio invarian. Maka digunakan kode sinkronisasi dengan menyinkronkan posisi dari penyisipan dan deteksi *watermark* sehingga *watermark* dapat terhindar dari *synchronization error* [10]. *Synchronization* berfungsi untuk mendeteksi posisi awal *watermark* ketika diserang *synchronization attack*.

2.8. Spread Spectrum

Spread Spectrum merupakan suatu teknik *digital watermarking* dimana *watermark* disebar ke seluruh spektrum sinyal host sehingga energi sinyal frekuensi yang digunakan sangat kecil. Berikut merupakan persamaan dari *spread spectrum*:

$$s_w(n) = s_0(n) + \alpha \cdot w_n \quad (2.8)$$

$s_w(n)$ merupakan *audio watermarked*, $s_0(n)$ merupakan sinyal *host audio*, α merupakan penguat agar mendapatkan *robustness* dan *imperceptibility* yang lebih baik, w_n dan merupakan sinyal *watermark*.

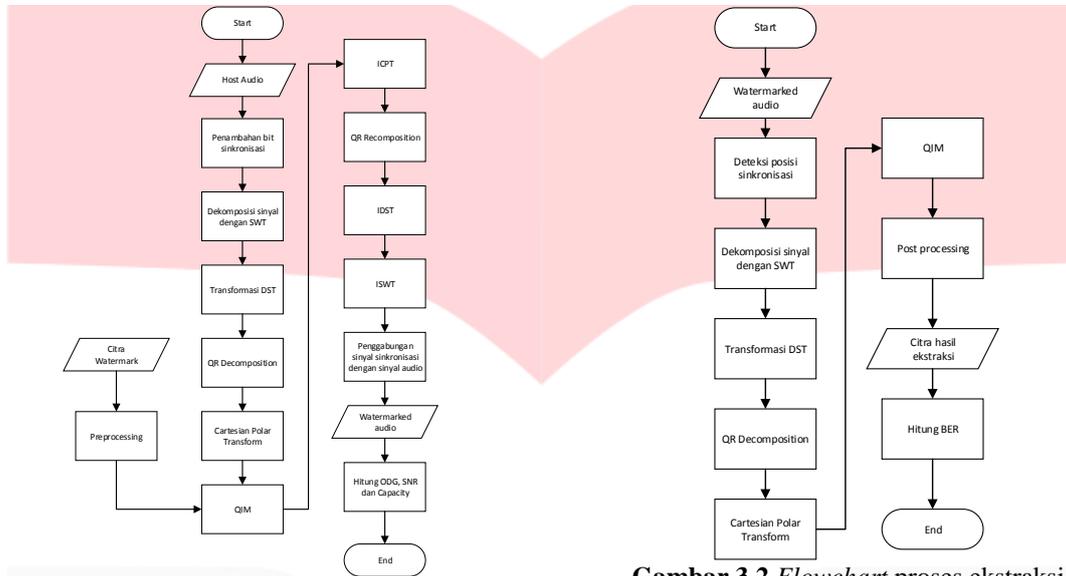
3. PERANCANGAN SISTEM

Perancangan sistem pada penelitian ini dibagi menjadi dua bagian, yaitu proses *embedding* dan proses ekstraksi. Pertama dalam proses *embedding host audio* akan disisipkan sebuah data (*watermark*). Dalam proses tersebut menggunakan beberapa metode yang diharapkan dapat meningkatkan *robustness* dari *audio watermarking*. Setelah *audio* telah di-*watermark*, *audio* akan diserang dengan serangan. Lalu dilakukan proses ekstraksi yaitu dengan mengambil *audio watermark* yang telah diserang. Kemudian *audio* akan diukur apakah tahan terhadap serangan atau tidak dan dihitung nilai BER dan SNR.

3.1 Proses Embedding

Proses *embedding* merupakan suatu tahapan menyisipkan suatu data *watermark* kedalam *file host audio*. Proses ekstraksi merupakan suatu tahapan pengambilan data kembali untuk data yang sudah disisipkan (*watermarked*). Penulis menggunakan dekomposisi SWT untuk memilih *subband*. Proses dekomposisi akan dibagi menjadi beberapa tingkat SWT kemudian dipilih *subband* yang memiliki daya terbesar. Metode DST digunakan untuk mengubah sinyal domain waktu menjadi sinyal domain frekuensi. *Output* dari proses DST akan menerapkan QR *Decomposition* bahwa *frame* akan didekomposisi menjadi matriks ortogonal dan matriks segitiga. Selanjutnya, penulis mengubah komponen dengan menggunakan transformasi CPT. *Audio* akan disisipkan bit sinkronisasi untuk mendeteksi lokasi awal *watermark*.

Penulis menggunakan metode QIM untuk proses *embedding* dan ekstraksi. Parameter akan diubah menjadi dalam metode QIM. Akhirnya, penulis merekonstruksi sinyal dari satu matriks dimensional menjadi dua matriks dimensional dan menghitung SNR, ODG dan C. Untuk lebih jelasnya lihat gambar 3.1 untuk proses *embedding*. Proses selanjutnya untuk proses ekstraksi hampir sama seperti *embedding* hanya dalam proses ini menggunakan audio ter-*watermark*. *Watermark* berupa citra akan diekstraksi untuk dicari nilai BER.



Gambar 3.1 Flowchart proses *embedding*

Gambar 3.2 Flowchart proses ekstraksi

4. ANALISIS

A. Tingkat Robustness

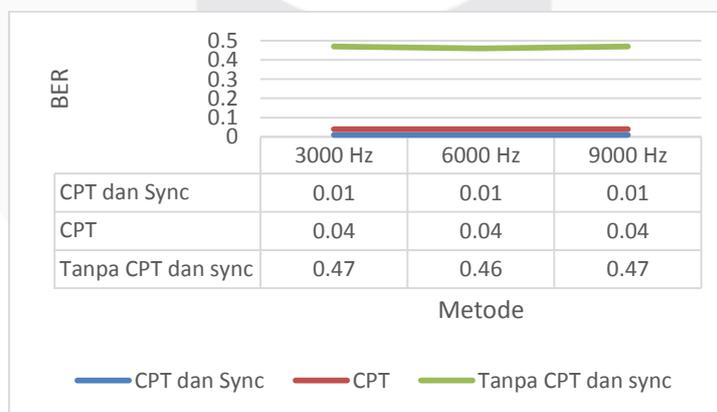
Setelah mencari pengaruh parameter kinerja untuk mendapatkan parameter optimal / terbaik, audio akan diuji tingkat ketahanan dengan adanya CPT dan sync. Penelitian ini akan diuji menggunakan CPT dan *synchronization bit*, CPT saja, dan tanpa CPT dan tanpa *synchronization bit*. Dalam analisis bagian ini, penulis mengambil Host.wav untuk dijadikan sampel grafik pengaruh adanya CPT dan *sync* karena Host.wav memiliki perubahan yang signifikan.

Tabel 4.1 Parameter optimal

N	N Frame	nbit	thr	posisiw	typew	alpha	nblock	Nbsi
4	2048	5	0,016	0	6	0,006	6	8

▪ Serangan LPF

Serangan LPF (*Low Pass Filter*) merupakan salah satu serangan yang berubah secara signifikan terhadap dengan ada CPT dan *sync*. Berikut merupakan grafik pengaruh CPT dan *sync* terhadap serangan LPF :

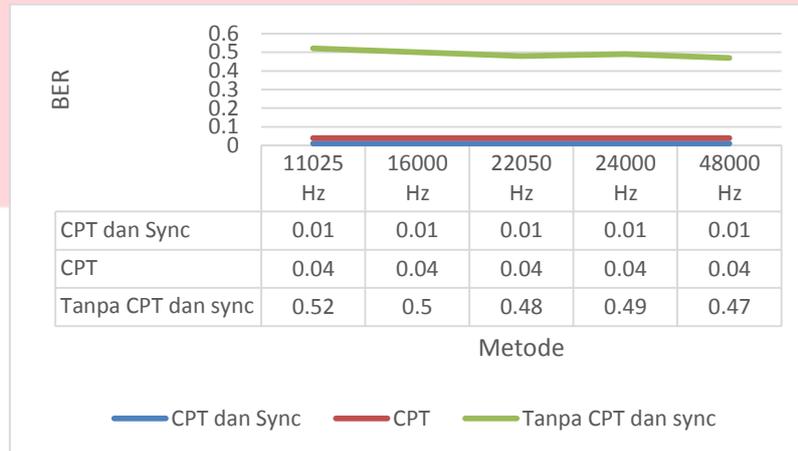


Gambar 4.1 Pengaruh CPT dan *sync* terhadap serangan LPF

Berdasarkan grafik diatas, dengan adanya CPT dan *sync audio watermarked* memiliki tingkat *robustness* yang baik terhadap serangan LPF karena memiliki nilai BER yang rendah.

- Serangan Resampling

Serangan *Resampling* merupakan salah satu serangan yang berubah secara signifikan terhadap dengan ada CPT dan *sync*. Berikut merupakan grafik pengaruh CPT dan *sync* terhadap serangan *Resampling* :

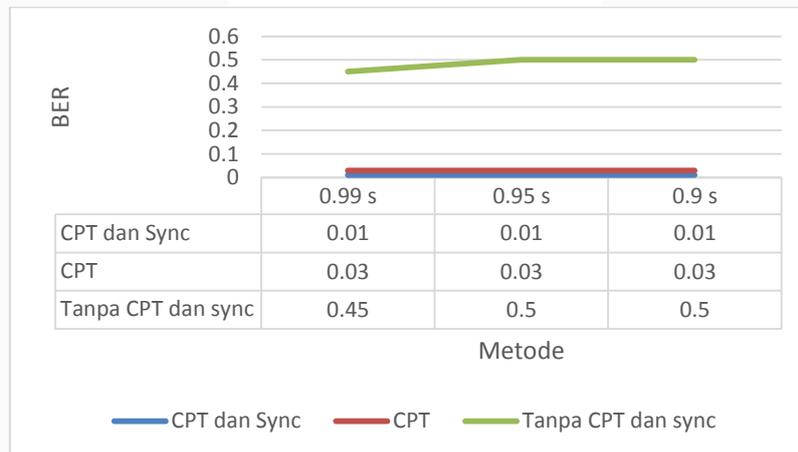


Gambar 4.2 Pengaruh CPT dan *sync* terhadap serangan *Resampling*

Berdasarkan grafik diatas, dengan adanya CPT dan *sync audio watermarked* memiliki tingkat *robustness* yang baik terhadap serangan *Resampling*.

- Serangan Speed Change

Serangan *Speed Change* merupakan salah satu serangan yang berubah secara signifikan terhadap dengan ada CPT dan *sync*. Berikut merupakan grafik pengaruh CPT dan *sync* terhadap serangan *Speed Change* :

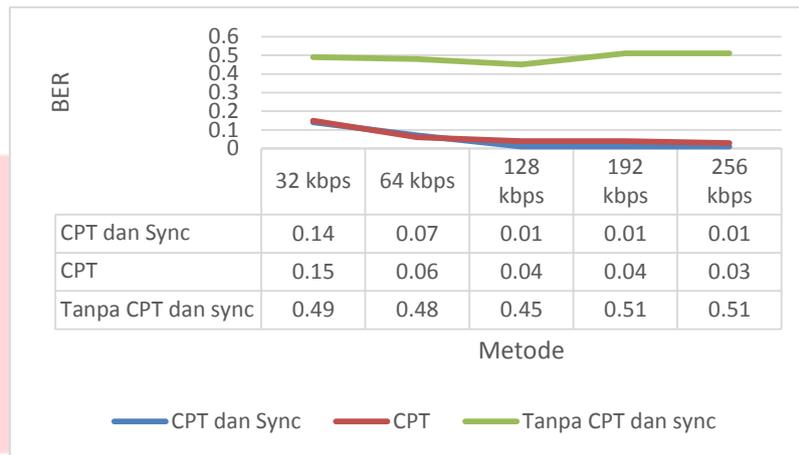


Gambar 4.3 Pengaruh CPT dan *sync* terhadap serangan *Speed Change*

Berdasarkan grafik diatas, dengan adanya CPT dan *sync audio watermarked* memiliki tingkat *robustness* yang baik terhadap serangan *Speed Change*.

- Serangan Mp3 Compression

Serangan *Mp3 Compression* merupakan salah satu serangan yang berubah secara signifikan terhadap dengan ada CPT dan *sync*. Berikut merupakan grafik pengaruh CPT dan *sync* terhadap serangan *Mp3 Compression* :

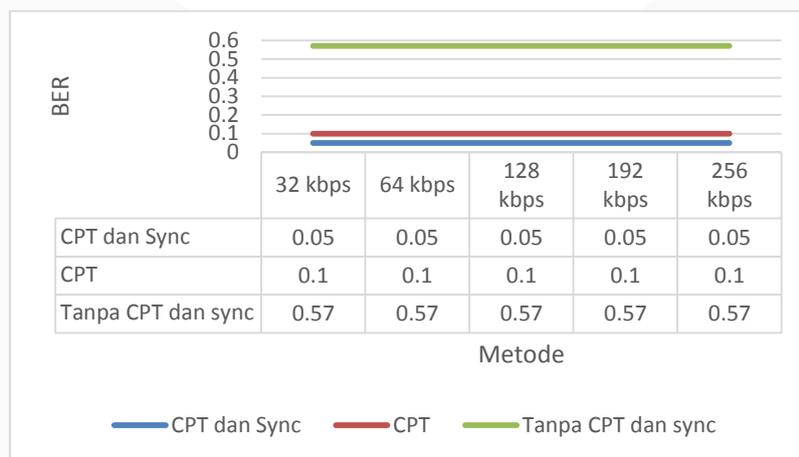


Gambar 4.4 Pengaruh CPT dan sync terhadap serangan Mp3 Compression

Berdasarkan grafik diatas, dengan adanya CPT dan sync, *audio watermarked* memiliki tingkat *robustness* yang baik terhadap serangan Mp3 Compression.

- Serangan Mp4 Compression

Serangan Mp4 Compression merupakan salah satu serangan yang berubah secara signifikan terhadap dengan ada CPT dan sync. Berikut merupakan grafik pengaruh CPT dan sync terhadap serangan Mp4 Compression :

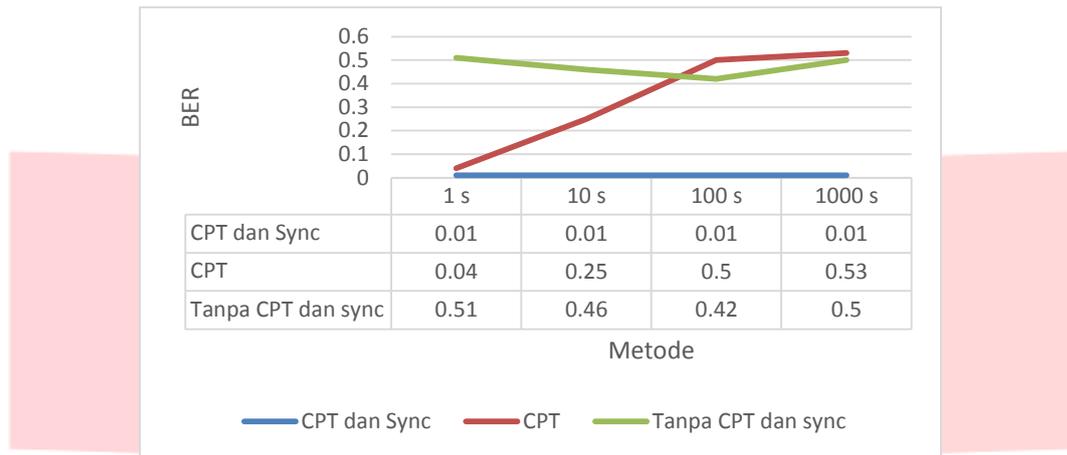


Gambar 4.7 Pengaruh CPT dan sync terhadap serangan Mp4 Compression

Berdasarkan grafik diatas, dengan adanya CPT dan sync, *audio watermarked* memiliki tingkat *robustness* yang baik terhadap serangan Mp4 Compression.

- Serangan Delay

Serangan Delay merupakan salah satu serangan yang berubah secara signifikan terhadap dengan ada CPT dan sync. Berikut merupakan grafik pengaruh CPT dan sync terhadap serangan Delay :



Gambar 4.8 Pengaruh CPT dan *sync* terhadap serangan *Delay*

Berdasarkan grafik diatas, dengan adanya *sync*, *audio watermarked* memiliki tingkat *robustness* yang baik terhadap serangan *Delay*.

B. Kualitas Audio

Kualitas audio merupakan hal terpenting dari *audio watermarking*. Kerusakan pada audio merupakan hal yang tidak bagus untuk *audio watermarking*. Data watermark jangan sampai terasa oleh pendengar. Parameter tersebut berdasarkan pada nilai ODG. Kejernihan suara ketika audio telah diberi *watermark* merupakan salah satu parameter dari SNR. Berikut merupakan hasil uji dari kualitas *audio watermarked* :

Tabel 4.1 Kualitas audio

Host	SNR			ODG		
	Dengan <i>sync</i> , dengan <i>cpt</i> dan parameter terbaik	Tanpa <i>sync</i> , dengan <i>cpt</i> , dengan parameter terbaik	Tanpa <i>sync</i> , tanpa <i>cpt</i> , dengan parameter terbaik	Dengan <i>sync</i> , dengan <i>cpt</i> dan parameter terbaik	Tanpa <i>sync</i> , dengan <i>cpt</i> , dengan parameter terbaik	Tanpa <i>sync</i> , tanpa <i>cpt</i> , dengan parameter terbaik
Pop.wav	29,8816	30,8404	91,7397	-0,4652	-0,42802	0,20512
Rock.wav	31,5276	31,426	95,2508	0,18518	0,1872	0,20578
EDM.wav	30,2703	30,0992	94,4653	0,099812	0,12458	0,20318
Jazz.wav	34,7799	33,9614	80,6672	-1,0395	-1,4912	-0,74641
Host.wav	37,1495	36,593	81,7727	0,17054	0,17045	0,18621

Berdasarkan dari tabel tersebut, semua audio memiliki kualitas audio yang sangat baik. Jika menggunakan *sync*, CPT dan parameter optimal, Host.wav memiliki nilai SNR yang paling tinggi dibanding *host audio* yang lainnya. EDM.wav memiliki nilai ODG yang paling bagus dibanding *host audio* yang lainnya karena mendekati nilai 0. Pada *host audio* Jazz.wav memiliki nilai ODG -1.0395 diimana *watermark* terasa namun tidak mengganggu kualitas dari audio. Dengan tanpa *sync*, tidak ada perubahan yang signifikan terhadap nilai SNR dan ODG dari semua *host audio*.

5. KESIMPULAN

1. Terdapat pengaruh dari setiap parameter terhadap hasil kinerja dari *watermarking audio stereo* berbasis QIM dengan teknik gabungan SWT, DST, QR dan CPT menggunakan sinkronisasi berbasis SS. Parameter tersebut adalah N (level dekomposisi), N Frame (jumlah sampel per *frame* untuk diproses), nbit (jumlah bit kuantisasi QIM), thr (*threshold* untuk *wavelet*), posisiw (posisi penyisipan

- watermark*), type *w* (tipe *wavelet*), *alpha* (penguat *header sync*), *nblock* (panjang kode *header 1 PNcode*), dan *Nbsi* (panjang bit *header*).
2. Dengan adanya CPT dan bit sinkronisasi sangat berpengaruh terhadap ketahanan (*robustness*) *audio watermarked* jika diserang dengan beberapa serangan.
 3. Dalam penelitian ini, sistem dapat menahan beberapa serangan, seperti LPF (*Low Pass Filter*), *Resampling*, *Speed Change*, *Mp3 Compression*, *Mp4 Compression* dan *Delay*. Namun, tidak kuat menahan serangan BPF (*Band Pass Filter*), *Noise*, *Time Scale Modification* dan *Pitch Shifting*.
 4. Sistem menggunakan CPT dan sinkronisasi pada saat diserang LPF (*Low Pass Filter*), *Resampling*, *Speed Change*, dan *Delay*, *audio watermarked* memiliki nilai BER rata-rata 0.01. Ketika tanpa menggunakan sinkronisasi, *audio watermarked* tidak tahan terhadap serangan *Delay*.
 5. *Audio watermarked* dengan menggunakan CPT dan *synchronization* memiliki nilai SNR rata-rata 32.718 dB. *Audio watermarked* menggunakan CPT dan tanpa *synchronization* memiliki nilai SNR rata-rata 32.584 dB. *Audio watermarked* tanpa menggunakan CPT dan *synchronization* memiliki nilai rata-rata 88.779 dB.
 6. *Audio watermarked* memiliki nilai index MOS rata-rata Host.wav = 4.3, Rock.wav = 4.4, EDM.wav = 4.4, Jazz.wav = 4.3, Host.wav = 4.3.

6. REFERENSI

- [1] Aparna J R and S. Ayyappan, "Comparison of digital watermarking techniques," *2014 Int. Conf. Comput. Power, Energy, Inf. Commun.*, pp. 87–92, 2014.
- [2] A. El Bahi and A. Adib, "A High Capacity Quantization-Based Audio Watermarking Technique Using the DWPT," *Int. Conf. Multimed. Comput. Syst.*, pp. 1–5, 2014.
- [3] J. Bajpai and A. Kaur, "A literature survey - Various audio watermarking techniques and their challenges," *Proc. 2016 6th Int. Conf. - Cloud Syst. Big Data Eng. Conflu. 2016*, pp. 451–457, 2016.
- [4] N. Khademi, M. A. Akhaee, S. M. Ahadi, M. Moradi, and A. Kashi, "Audio watermarking based on Quantization Index Modulation in the frequency domain," *ICSPC 2007 Proc. - 2007 IEEE Int. Conf. Signal Process. Commun.*, no. November, pp. 1127–1130, 2007.
- [5] C. Delong, L. Qirui, Y. Guilan, and X. Jianbin, "Content-Based Audio Watermarking Method To Resist De-Synchronization Attacks," *Int. Conf. Inf. Netw. Secur.*, 2014.
- [6] S. D. R. Mirajkar Pradnya P., "Image Fusion Based On Stationary Wavelet Transform," *Int. J. Adv. Eng. Res. Stud. E-ISSN2249–8974 Res.*, vol. 2013, no. L, pp. 99–101, 2013.
- [7] H. Yassine, B. Bachir, and K. Aziz, "A Secure and High Robust Audio Watermarking System for Copyright Protection," *Int. J. Comput. Appl.*, vol. 53, no. 17, pp. 33–39, 2012.
- [8] M. Hemis, B. Boudraa, and T. Merazi-meksen, "Intelligent Audio Watermarking Algorithm using Multi-objective Particle Swarm Optimization," *Int. Conf. Electr. Eng.*, no. 1, pp. 0–4, 2015.
- [9] P. K. Dhar and T. Shimamura, "Audio Watermarking in Transform Domain Based on Singular Value Decomposition and Quantization," *Asia-Pacific Conf. on Communications*, pp. 516–521, 2012.
- [10] B. Lei, I. Y. Soon, and E. L. Tan, "Robust SVD-based audio watermarking scheme with differential evolution optimization," *IEEE Trans. Audio, Speech Lang. Process.*, vol. 21, no. 11, pp. 2368–2378, 2013.
- [11] J. Li and T. Wu, "Robust audio watermarking scheme via QIM of correlation coefficients using LWT and QR decomposition," *ICCCS 2015 - Proc. 2015 Int. Conf. Inf. Cybern. Comput. Soc. Syst.*, no. 4, pp. 1–6, 2015.