

# **Removable Watermarking Sebagai Pengendalian Terhadap Cyber Crime Pada Audio Digital**

## **Removable Watermarking As Cyber Crime Control In Digital Audio**

**Reyhani Lian Putri \*, Mona Renasari, Gelar Budiman**

*Prodi S1 Teknik Telekomunikasi, Fakultas Teknik Elektro, Universitas Telkom  
Jalan Telekomunikasi No. 1, Sukapura, Bojongsoang, Bandung, Jawa Barat 40257*

---

### **Abstrak**

Perkembangan teknologi informasi yang pesat menuntut penggunaannya untuk lebih berhati-hati seiring semakin meningkatnya *cyber crime*. Banyak pihak telah mengembangkan berbagai teknik perlindungan data digital, salah satunya adalah *watermarking*. Teknologi *watermarking* berfungsi untuk memberikan identitas, melindungi, atau menandai data digital, baik audio, citra, ataupun video, yang mereka miliki. Akan tetapi, teknik tersebut masih dapat diretas oleh oknum-oknum yang tidak bertanggung jawab. Pada penelitian ini, proses *watermarking* diterapkan pada audio digital dengan menyisipkan *watermark* yang terdengar jelas oleh indera pendengaran manusia (*perceptible*) pada audio *host*. Hal ini bertujuan agar data audio dapat terlindungi dan apabila ada pihak lain yang ingin mendapatkan data audio tersebut harus memiliki "kunci" untuk menghilangkan *watermark*. Proses *removable watermarking* ini dilakukan pada data *watermark* yang sudah diketahui metode penyisipannya, agar *watermark* dapat dihilangkan sehingga kualitas audio menjadi lebih baik. Dengan menggunakan metode ini diperoleh kinerja *audio watermarking* pada nilai distorsi tertinggi dengan rata-rata nilai SNR sebesar 7,834 dB dan rata-rata nilai ODG sebesar -3,77. Kualitas audio meningkat setelah *watermark* dihilangkan, di mana rata-rata SNR menjadi sebesar 24,986 dB dan rata-rata ODG menjadi sebesar -1,064 serta nilai MOS sebesar 4,40.

**Kata kunci:** *removable watermarking*, audio, *audio watermarking*, perlindungan hak cipta, *perceptible*

---

### **Abstract**

*The rapid growth of information technology requires its users to be more careful due to the increase of cyber crime. Various techniques for digital data protection have been developed, one of which is watermarking. Watermarking technology serves to provide an identity, protect by marking of digital data including image, audio, and videos. However, this technique can still be hacked by irresponsible users. In this study, watermarking process is applied to digital audio by inserting watermark audio that can be sensed by human (perceptible) to the audio host. This process is aimed to protect the audio data, so that if there are other parties who wish to get the original audio, it required a "key" to remove the watermark. Removable watermarking process which performed on the data is already known so that the watermark can be removed and the resulting audio quality will be better. The obtained audio watermarking performance at the highest distortion value with an average value of 7.834 dB SNR and the average value of -3.77 ODG. The audio quality is increased after the watermark is removed, which on average amounted to 24.986 dB SNR and an average ODG amounted to -1.064, so that MOS value is 4.40.*

**Keywords:** *removable watermarking*, audio, *audio watermarking*, copyright protection, *perceptible*

---

## **I. PENDAHULUAN**

Pesatnya perkembangan teknologi dewasa ini memudahkan penggunaannya untuk mengakses dan menyebarkan berbagai informasi yang tersedia di dunia maya secara bebas. Kebebasan untuk mengakses informasi tersebut juga menuntut pengguna agar berhati-hati terhadap penyalahgunaan produk digital seperti pencaplokan hak cipta. Untuk itu, diperlukan suatu metode untuk melindungi data digital dari oknum tertentu. Salah satu metode yang banyak digunakan adalah *watermarking*. *Watermarking* adalah suatu metode penyisipan informasi tertentu, baik dalam

bentuk umum atau rahasia, ke dalam data digital (audio, gambar, atau video), namun keberadaannya tidak dapat dideteksi oleh indera manusia, baik indera pendengaran maupun indera penglihatan, dan mampu bertahan dari berbagai pengujian ketahanan seperti *cropping*, kompresi, *resampling*, dll, sampai tahap tertentu. Bagi para responden, perbedaan antara data digital yang telah dimasukkan *watermark* ataupun yang belum dimasukkan *watermark* tidak akan terlalu mengetahui dengan jelas [1]. *Host* yang dapat digunakan dalam *watermarking* ada tiga, yaitu audio, citra, dan video.

Dalam *watermarking*, ada beberapa syarat yang harus dipenuhi, yaitu *payload*, *robustness*, dan *imperceptibility*. *Payload* adalah jumlah bit *watermark* yang disisipkan. Semakin banyak bit yang disisipkan, maka audio ter*watermark* akan semakin mengganggu

---

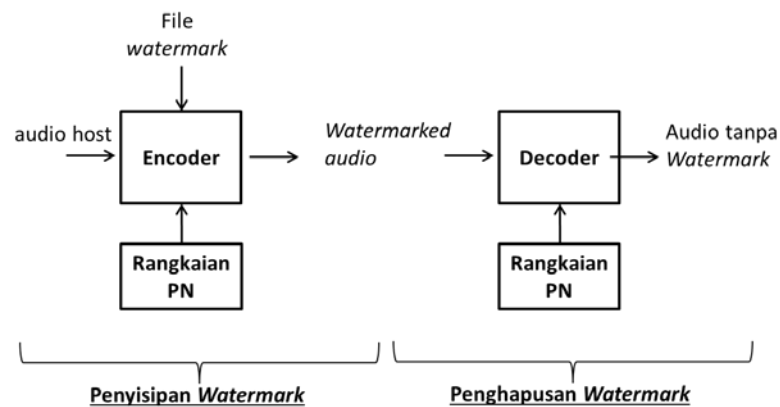
\* Corresponding Author.

Email: reyhani.lian@gmail.com

Received: December 17, 2016 ; Revised: May 05, 2017

Accepted: June 19, 2017 ; Published: August 31, 2017

© 2017 PPET - LIPI

Gambar 1. Skema Proteksi Audio Digital dengan *Removable Watermarking*

indera pendengaran manusia. *Robustness* adalah ketahanan yang dimiliki oleh *watermark*, dimana *watermark* harus tahan terhadap serangan yang disengaja dan manipulasi data. *Imperceptibility* adalah syarat *watermark* agar tidak terasa oleh indera manusia, baik indera pendengaran maupun indera penglihatan [1].

Penelitian terkait *removable watermarking* telah dilakukan pada [2]-[9]. Pada [2]-[3] dibahas proses *removable watermarking* dengan merepresentasikan skema untuk proses penyisipan *watermark* terlebih dahulu, dilanjutkan proses ekstraksi, dan selanjutnya adalah proses *removal*. *Watermark* pada *audio host* diekstraksi dan *watermark* yang masih tersisa pada *watermarked audio* dihilangkan untuk mendapatkan data audio dengan kualitas yang lebih baik. Proses *removable watermarking* pada *host* citra yang terlihat (*visible*) telah diterapkan pada [4]. Pada [5], audio yang telah diberi *watermark* dikembalikan ke bentuk semula dengan menghilangkan *inaudible watermark* yang disisipkan. Implementasi audio *watermark* juga telah dilakukan pada [6]-[9] dengan memanfaatkan transformasi DCT [6] dan FFT [7]-[9].

Pada penelitian ini, akan diimplementasikan teknik digital *watermarking* untuk perlindungan hak cipta pada data digital. Data digital yang digunakan berupa audio sebagai media *host* yang akan disisipi dengan *audio watermark* yang terdengar jelas oleh indera pendengaran manusia (*perceptible*). Akan tetapi proses penyisipan akan berpengaruh pada menurunnya kualitas audio *host* [6]. Walaupun demikian, hal ini dilakukan untuk perlindungan hak cipta media *host*. Kualitas audio *host* akan baik seperti semula jika *watermark* dihilangkan dan untuk menghilangkannya dibutuhkan “kunci”. Ini akan menjamin hak cipta dari audio *host* sehingga hanya pihak-pihak yang memiliki “kunci” saja yang dapat menikmati audio tanpa *watermark*.

## II. PERANCANGAN REMOVABLE WATERMARKING PADA FILE AUDIO

*Audio watermarking* adalah metode penyisipan informasi (umum atau rahasia) ke dalam data audio, yang keberadaannya tidak dapat dideteksi oleh indera pendengaran manusia (*inaudible*) dan mampu bertahan dari berbagai serangan pengolahan sinyal sampai tahap tertentu. *Watermark* pada audio dapat dihilangkan

dengan menggunakan suatu “kunci” yang ditentukan berdasarkan metode penyisipan informasi yang telah diketahui sebelumnya. Umumnya informasi yang disisipkan menggunakan *imperceptible watermarking*, yaitu menggunakan informasi yang tidak dapat dideteksi oleh indera manusia, baik indera pendengaran (*inaudible*) maupun indera penglihatan (*invisible*). Namun, *perceptible watermarking* dapat digunakan untuk melindungi data digital dengan cara yang lebih terbuka sehingga keberadaan *watermark* dapat langsung diketahui.

Skema *audio watermarking* yang dilakukan pada penelitian ini dapat dilihat pada Gambar 1. Di sisi *encoder*, *audio host* yang ingin diproteksi disisipi *watermark* dengan memanfaatkan rangkaian PN untuk memilih frekuensi *carrier* secara acak. Pada sisi *decoder*, *watermark* dihilangkan dari *audio watermarked* untuk memperoleh sinyal audio dengan kualitas yang lebih baik.

### A. Penyisipan Watermark

Salah satu metode penyisipan *watermark* yang sering digunakan adalah metode *spread spectrum* karena memiliki ketahanan (*robust*) yang tinggi. Pada penelitian ini digunakan metode *Frequency Hopping Spread Spectrum* (FHSS) yang merupakan salah satu pendekatan metode *spread spectrum*. FHSS memodulasi audio *host* dengan *Pseudorandom-sequence* atau *pseudonoise* atau rangkaian PN untuk memilih secara acak frekuensi yang telah ditentukan dan frekuensi ini digunakan sebagai *carrier frequency* dari sinyal data. *Carrier frequency* melompat (*hop*) di spektrum pada pola ditentukan oleh urutan PN.

Pada sisi *encoder*, dilakukan proses penyisipan *watermark* ke dalam *audio host* dengan menggunakan metode FHSS. *Audio host* dalam format \*.wav diubah ke domain frekuensi dengan FFT dan data *watermark* dalam format \*.txt (teks) diubah ke dalam bentuk biner. Rangkaian PN selanjutnya dibangkitkan. Dua koefisien FFT dipilih secara acak berdasarkan rangkaian *pseudonoise* (PN) yang telah dibangkitkan. Nilai rata-rata dari semua koefisien *magnitude* FFT dihitung. Data *watermark* dalam bentuk biner disisipkan ke dalam koefisien *magnitude* frekuensi yang telah dipilih secara acak. Apabila yang disisipkan adalah bit 1, maka nilai koefisien FFT pada frekuensi yang lebih rendah

dijadikan  $K$  dB di atas nilai rata-rata, dan nilai koefisien FFT yang kedua dijadikan  $K$  dB di bawah nilai rata-rata. Sebaliknya, jika yang disisipkan adalah bit 0, maka nilai koefisien FFT pada frekuensi yang lebih rendah dijadikan  $K$  dB di bawah nilai rata-rata, dan nilai koefisien FFT yang kedua dijadikan  $K$  dB di atas nilai rata-rata. Lakukan *inverse* FFT untuk mendapatkan kembali audio dalam domain waktu. Keluaran yang dihasilkan adalah audio dengan *watermark* (*watermarked audio*).

## B. Penghapusan Watermark

Untuk menghilangkan *watermark* dari *watermarked audio*, dibutuhkan kunci yang sama dengan yang digunakan pada proses penyisipan. Proses ekstraksi dan penghapusan *watermark* dapat dilihat pada Gambar 2. Proses penghapusan watermark diawali dengan FFT yang diterapkan pada *watermarked audio* untuk mengubah sinyal ke dalam domain frekuensi. Dua koefisien FFT dipilih secara acak berdasarkan rangkaian PN yang telah dibangkitkan, dimisalkan *magnitude* koefisien frekuensi rendah adalah  $F_{k1}$  dan *magnitude* koefisien frekuensi tinggi adalah  $F_{k2}$ . Nilai deteksi ( $D$ ) merupakan selisih kedua *magnitude* tersebut sebagaimana dirumuskan pada persamaan (1).

$$D = F_{k1} - F_{k2} \quad (1)$$

Tanda dari nilai  $D$  di atas menunjukkan nilai bit yang diekstraksi. Jika nilai  $D$  adalah positif, maka yang diekstrak adalah bit 1. Namun jika nilai  $D$  adalah negatif, maka yang diekstrak adalah bit 0. Lakukan *inverse* FFT untuk mengembalikan bit biner ke dalam domain waktu. Proses diteruskan hingga tidak ada lagi bit *watermark* yang akan diekstraksi.

Setelah *watermark* diekstraksi, selanjutnya *watermark* dihilangkan. Langkah awal adalah menghitung nilai rata-rata *magnitude* koefisien FFT pada subband. Nilai rata-rata ini selanjutnya dipetakan ke masing-masing nilai  $F_{k1}$  dan  $F_{k2}$ . Lakukan *inverse* FFT untuk mendapatkan kembali audio dalam domain waktu. Jika masih ada bit *watermark* yang belum dihilangkan maka proses *removal* akan terus berlanjut, didapatkan audio tanpa *watermark*.

## C. Parameter Pengujian

Parameter pengujian yang digunakan untuk melihat performansi hasil perancangan adalah *Mean Opinion*

*Score* (MOS), *Signal to Noise Ratio* (SNR) dan *Perceptual Evaluation of Audio Quality* (PEAQ).

### 1) Mean Opinion Score (MOS) [10]

MOS adalah salah satu parameter pengujian secara subjektif berdasarkan indera pendengaran manusia untuk mengukur tingkat *inaudibility* terhadap perbandingan audio host dan audio yang sudah diberi *watermark*. Namun dalam penelitian ini, MOS digunakan untuk menguji kualitas audio terhadap audio host dan audio yang sudah tidak memiliki *watermark*. Parameter MOS ditentukan dengan melakukan survei terhadap tiga puluh responden penelitian secara acak. Hasil penelitian terhadap responden diukur dengan kriteria sebagaimana ditunjukkan pada Tabel 1.

TABEL 1  
KRITERIA PENILAIAN MOS

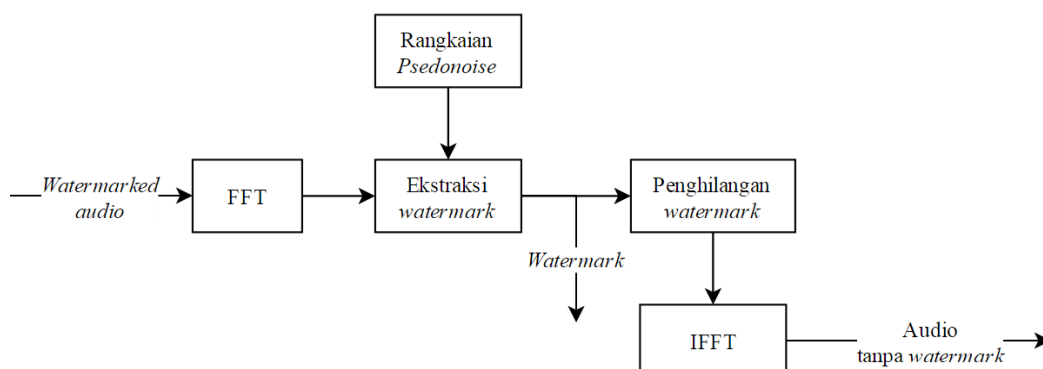
Skala MOS	Kualitas Audio	Level Distorsi
5	Amat Baik	Tidak ada watermark dan audio terdengar jelas
4	Baik	Watermark masih terasa sedikit namun tidak mengganggu audio
3	Cukup	Watermark masih terasa namun sedikit mengganggu audio
2	Kurang	Watermark mengganggu namun audio masih dapat didengar
1	Buruk	Watermark mengganggu dan audio tidak terdengar

### 2) Signal to Noise Ratio

*Signal to Noise Ratio* (SNR) adalah nilai yang menyatakan tingkat *noise* atas audio yang telah disisipi pesan *watermark*. Semakin rendah SNR, maka semakin jelas *watermark* yang disisipkan. SNR digunakan untuk mengukur kualitas audio secara objektif. Menurut *International Federation of the Phonographic Industry* (IFPI), SNR sinyal audiowatermarking yang baik adalah lebih besar dari 20 dB. Nilai SNR dapat dihitung dengan Persamaan (2) [1].

$$SNR = 10 \log_{10} \frac{\sum_{t=0}^{N-1} f^2(n)}{\sum_{t=0}^{N-1} (g^2(n) - f^2(n))} \quad (2)$$

dimana  $N$  merupakan panjang audio,  $f(n)$  adalah sampel sinyal audio asli dan  $g(n)$  adalah sampel sinyal audio *watermark*.



Gambar 2. Skema Ekstraksi dan Watermark Removal

### 3) *Perceptual Evaluation of Audio Quality (PEAQ)* [11]

*Perceptual Evaluation of Audio Quality (PEAQ)* adalah salah satu parameter dari *Objective Difference Grade (ODG)*. Berdasarkan ITU-R BS.1387, PEAQ merupakan salah satu standar yang menjelaskan pengujian kualitas audio secara objektif. Parameter ini membandingkan perbedaan bunyi audio dengan melakukan pengukuran terhadap data audio asli dan data audio yang diujikan, dalam hal ini adalah audio yang *watermark*-nya telah dihilangkan. Parameter PEAQ dapat dilihat pada skala pengukuran seperti Tabel 2.

TABEL 2  
SKALA PENGUKURAN PEAQ

Skala PEAQ	Deskripsi
0	Tidak terdeteksi
-1	Terdeteksi namun tidak mengganggu
-2	Cukup mengganggu
-3	Mengganggu
-4	Sangat mengganggu

### III. ANALISIS DAN PENGUJIAN SISTEM

Pengujian pada penelitian ini menggunakan lima sampel audio berbeda dengan durasi masing-masing adalah 10 detik dan frekuensi *sampling* sebesar 44100 Hz. Setiap audio diuji dengan nilai distorsi  $K = 0$  dB dan  $K = 25$  dB. Nilai SNR dan ODG hasil penyisipan dan ekstraksi dihitung sebagai parameter pengukuran kualitas audio. Tabel 3 dan Tabel 4 di bawah ini menunjukkan pengaruh dari nilai distorsi  $K$  terhadap kualitas audio berdasarkan parameter SNR dan ODG/PEAQ.

Secara keseluruhan, pada perbandingan nilai SNR, semua audio host memiliki nilai SNR diatas 20 dB saat diberikan nilai  $K = 0$  dB. Namun pada saat diberikan nilai  $K = 25$  dB, nilai SNR menurun menjadi dibawah 20 dB. SNR paling rendah ditunjukkan oleh *host* audio “host – perang.wav” yang berarti audio tersebut kualitasnya paling buruk dibandingkan *host* audio lainnya. Setelah audio yang telah diberikan *watermark* diekstraksi dan *watermark* dihilangkan, nilai SNR meningkat, baik dengan nilai  $K = 0$  dB atau  $K = 25$  dB.

TABEL 3  
PERBANDINGAN NILAI SNR

Audio Host	SNR Penyisipan (dB)		SNR Ekstraksi (dB)	
	K = 0 dB	K=25 dB	K=0 dB	K=25 dB
host – bola.wav	24,12	12,45	24,70	25,12
host – metal.wav	24,76	12,73	25,14	22,96
host – perang.wav	23,42	3,93	23,90	23,25
host – percakapan.wav	26,32	5,46	26,45	25,31
host – pop.wav	22,94	4,76	26,74	26,35

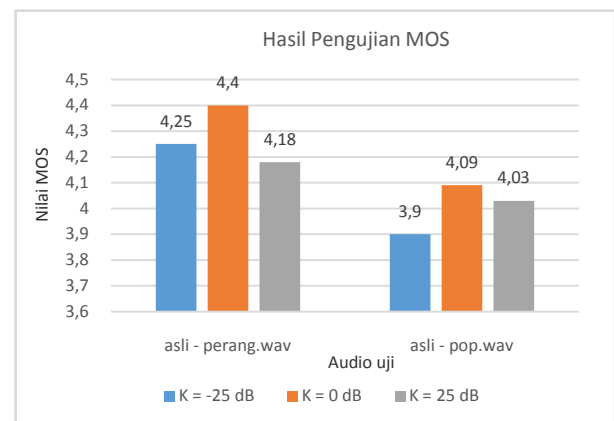
Selain menghitung nilai SNR, parameter lainnya yang bersifat objektif adalah ODG/PEAQ. Dari pengujian yang dilakukan, maka didapatkan hasil seperti pada Tabel 4. Nilai ODG sebanding dengan nilai SNR. Saat proses penyisipan, nilai ODG saat  $K = 0$  dB

lebih besar daripada  $K = 25$  dB. Audio *host* “host – pop.wav” menunjukkan nilai paling rendah setelah *watermark* disisipkan. Setelah proses ekstraksi dan penghilangan *watermark*, nilai ODG pada saat nilai  $K = 25$  dB meningkat.

TABEL 4  
PERBANDINGAN NILAI ODG/PEAQ

Audio Host	ODG Penyisipan		ODG Ekstraksi	
	K = 0 dB	K=25 dB	K = 0 dB	K = 25 dB
host – bola.wav	-0,64	-3,82	-0,64	-0,72
host – metal.wav	-1,17	-3,45	-1,17	-1,25
host – perang.wav	-0,27	-3,78	-0,11	-0,20
host – percakapan.wav	-1,04	-3,89	-1,04	-1,58
host – pop.wav	-1,29	-3,91	-1,29	-1,57

Pengujian sistem secara subyektif dilakukan dengan melakukan tes audio kepada minimal tiga puluh orang responden secara. Tes audio ini dinamakan dengan *Mean Opinion Score (MOS)* di mana responden diminta untuk membandingkan perbedaan kualitas dari *watermarked audio* dengan audio yang *watermark* telah dihilangkan.



Gambar 3. Hasil Pengujian MOS

Berdasarkan Gambar 3, audio hasil perbaikan dari *watermarked audio* baik pada audio “asli – perang.wav” maupun “asli – pop.wav” dengan nilai distorsi  $K = 0$  dB memiliki nilai MOS yang lebih tinggi dibandingkan  $K = -25$  dB dan  $K = 25$  dB. Jika dibandingkan, “asli – perang.wav” memiliki nilai MOS yang lebih tinggi dari “asli – pop.wav” dari semua nilai distorsi  $K$ . Hal ini menunjukkan audio perbaikan dari suara dengan latar suara tembakan dan ledakan terdengar lebih mirip dengan audio aslinya.

### KESIMPULAN

Pada penelitian ini telah dilakukan perancangan proteksi audio digital dengan removable watermarking. Proses penyisipan menggunakan metode FHSS karena memiliki tingkat ketahanan tinggi. Nilai distorsi  $K$  pada proses penyisipan *watermark* dapat menurunkan kualitas *watermarked audio*. Hasil ODG juga sebanding dengan nilai SNR. Nilai ini menjelaskan bahwa kenaikan nilai distorsi  $K$  sebanding dengan kualitas audio yang disisipi *watermark*. Semakin tinggi nilai distorsi  $K$ , maka kualitas audio yang disisipi *watermark* juga akan semakin menurun.

### UCAPAN TERIMA KASIH

Penelitian ini difasilitasi oleh Jurusan Teknik Telekomunikasi, Fakultas Teknik Elektro, Telkom University. Untuk itu, penulis sangat berterima kasih kepada dosen untuk bimbingan dan arahan serta fasilitas yang diberikan.

### DAFTAR PUSTAKA

- [1] A. D. Setiarsih, T. A. Bw, and A. T. Wibowo, "Watermarking audio dengan menggunakan metode Frequency Hopping Spread Spectrum (FHSS)," S.T. final project, Institut Teknologi Telkom, Bandung, Indonesia, 2010.
- [2] M. Löytynoja, N. Cvejic, and T. Seppänen, "Audio protection with removable watermarking," in *Proc. of the 6th International Conference on Information, Communications and Signal Processing (ICICS)*, pp. 1-4, 2007.
- [3] A. Anastasijević and D. Čoja, "Frequency hopping method for audio watermarking," *Telfor Journal*, Vol. 4, no. 2, pp. 155–160, 2012.
- [4] Y. Hu, S. Kwong, and J. Huang, "An algorithm for removable visible watermarking," *IEEE Transactions on Circuits And Systems for Video Technology*, vol. 16, no. 1, pp. 129–133, 2006.
- [5] M. Unoki and R. Miyauchi, "Reversible watermarking for digital audio based on cochlear delay characteristics," in *Proc. of 7th International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP)*, 2011, vol. 99, no. Cd, pp. 314–317.
- [6] M. K. Dutta, P. Gupta, and V. K. Pathak, "Perceptible audio watermarking for digital right management control," in *Proc. of 7th International Conference on Information, Communications and Signal Processing (ICICS)*, 2009.
- [7] N. Yusuf, "Analisis ketahanan audio watermarking di domain frekuensi pada ambient mode dengan menggunakan frequency masking method," in *e-Proceeding of Engineering*, 2016, Vol. 3, no.1, pp. 291.
- [8] N. Fajria, Suwandi, and H. B. D. Kusumaningrum, "Aplikasi Fast Fourier Transform (FFT) untuk deteksi kelainan pita suara berbasis android application Fast Fourier Transform (FFT)," S.T. final project, Universitas Telkom, Bandung, Indonesia, 2016.
- [9] G. Budiman, A. B. Suksmono, and D. Danudirdjo, "Fibonacci sequence based FFT and DCT performance comparison in audio watermarking," *Pertanika J. Sci. Technol.*, vol. 24, no. 1, pp. 1–10, 2016.
- [10] *Series P:Telephone Transmission Quality, Telephone Installations, Local Line Networks, Methods For Objective And Subjective Assessment Of Quality, Mean Opinion Score (MOS) Terminology*, ITU-T Recommendation P.800.1, 2003.
- [11] P. Kabal, "An Examination and Interpretation of ITU-R BS. 1387: Perceptual Evaluation of Audio Quality," McGill Univ., 2003, pp. 1–96.