DIGITAL AUDIO WATERMARKING MENGGUNAKAN ANALISIS AUDIO CONTENT

Dean Fathony Alfatwa – NIM : 13503003

Program Studi Teknik Informatika, Institut Teknologi Bandung
Jl. Ganesha 10, Bandung
E-mail : if13003@students.if.itb.ac.id

Abstrak

Perkembangan teknologi jaringan saat ini membuat penyebaran data multimedia digital khususnya data audio menjadi semakin mudah. Banyak situs di internet yang menawarkan lagu-lagu dalam bentuk file digital seperti mp3, wav, dsb yang dapat di-*download* dengan mudah baik melalui pembayaran dengan kartu kredit maupun *download* secara gratis. Perlu adanya perlindungan terhadap hak cipta terhadap data audio digital agar tidak ada pihak yang dirugikan dalam penyebaran data semacam ini.

Perlindungan terhadap data audio dapat dilakukan dengan digital audio watermarking, dimana informasi yang tidak dapat didengar (bukan audio) dapat disimpan ke dalam data audio tersebut. Informasi yang dapat disimpan antara lain pencipta lagu tersebut dan pengguna yang telah membeli lagu tersebut secara sah.

Dalam makalah ini akan dijelaskan salah satu teknik dalam digital audio watermarking yang menekankan pada sinkronisasi terhadap penyerangan yang disebabkan oleh *audio editing* sederhana atau perusakan data audio, keduanya merupakan jenis serangan low-cost yang efektif terhadap teknik digital audio watermarking yang terdahulu. Skema digital audio watermarking dengan kompleksitas rendah digunakan di sini sebagai cara yang efektif untuk menghalangi pengguna dari penyebaran data audio secara ilegal. Skema ini didasarkan pada analisis *audio content* menggunakan filter gelombang pandu ketika watermark disisipkan pada domain perubahan Fourier.

Keywords: *Digital watermark, Audio watermark,* Gelombang pandu, *Audio Content Analysis, Fourier Transform Domain*

1. Pendahuluan

Digital audio watermarking, teknik penembunyian informasi pada jenis sinyal yang tidak dapat dilihat dalam data audio digital akhir-akhir ini semakin menjadi perhatian. Diantara berbagai macam penggunaan dalam *digital audio watermarking*, perlindungan terhadap hak cipta merupakan isu yang paling banyak diminati. Hal ini membuat penerapan watermark jenis ini pun semakin dibutuhkan.

Perkembangan teknologi jaringan yang semakin cepat dan semakin majunya teknik kompresi data audio membuka peluang untuk pendistribusian musik secara online menjadi industri yang menjanjikan. Tapi seperti yang diketahui, penduplikasian data dalam teknologi digital sangatlah mudah. Dengan adanya teknologi digital seperti saat ini, penyalinan dan penyebaran data secara ilegal menjadi jauh lebih mudah daripada sebelumnya. Hal ini menyebabkan pencipta lagu dan distributor ragu-ragu untuk melangkah ke dalam industri jenis ini. Sehingga teknologi perlindungan terhadap isi dari data audio secara tepat merupakan kunci untuk memunculkan industri jenis ini.

Enkripsi dan watermarking merupakan dua hal yang sangat penting dalam hal perlindungan terhadap isi dari data audio. Enkripsi melindungi isi data audio dari setiap orang yang tidak memliki kunci dekripsi yang tepat. Hal ini berguna untuk melindungi data audio ketika disadap saat pengiriman. Namun, setelah penerima yang dituju mendekripsi data audio tersebut dengan kunci yang benar, data audio tersebut masih dapat didistribusikan secara ilegal dan dapat pula disalahgunakan. Sementara itu, watermark tidak dapat dihilangkan dari data audio bahkan oleh penerima yang dituju. Sinyal watermark yang disisipkan akan tetap berada pada data audio secara permanen

meskipun data audio tersebut diproduksi atau didistribusi ulang. Sehingga sinyal ini dapat digunakan untuk melindungi hak cipta isi data audio dari penyalinan secara ilegal.

Penerapan lain dari *digital audio watermarking*, termasuk di dalamnya penyembunyian data untuk komunikasi yang tersembunyi, penyisipan data yang membantu pada proses pelabelan data audio, dan pendeteksian pengubahan data untuk membuktikan keaslian data. Penyembunyian data juga dapat digunakan sebgai pelengkap untuk enkripsi, yaitu untuk memeperketat keamanan kominikasi dengan menyembunyikan transmisi dari data yang sensistif (penting). Data bantu yang disembunyikan dapat berupa lirik lagu, deskripsi dari data audio yang membawanya, atau sebagai layanan penghubung ke basis data eksternal. Hilangnya watermark yang kurang kuat menandakan adanya pengubahan data secara tidak sah.

Penerapan watermarking yang berbeda, memiliki sekumpulan kebutuhan yang berbeda pula. Dalam makalah ini pembahasan akan difokuskan pada perlindungan terhadap hak cipta. Hal ini menjadi fokus karena isu ini yang paling membutuhkan kemampuan watermark agar data dapat bertahan dari serangan. Isu ini telah menjadi isu yang paling hangat saat ini. Pengguna diuntungkan dengan adanya data label yang disisipkan sementara *hacker* tidak mengetahui adanya data komunikasi yang tersembunyi. Jadi, watermark yang tersembunyi di dalam dua penerapan ini secara umum bukanlah sasaran dari serangan.

Makalah ini tersususn sebagai berikut. Apa saja yang dibutuhkan dalam sistem *audio watermarking* dipaparkan pada bab dua. Kegiatan sebelumnya dalam dunia *audio wtermarking* dijelaskan di bab tiga. Kegiaatan yang dilakukan saat ini dalam hal Fourier Domain Watermarking dijelaskan dalam bab empat. Hasil penellitian dan analisis dipaparkan dalam bab lima. Kesimpulan pada bab enam.

2. Kebutuhan dalam Sistem *Audio Watermarking*

Agar watermark yang disisipkan dapat secara efektif melindungi hak cipta terhadap data audio digital, telah disetujui secara umum bahwa skema wwtermarking yang bagus harus memenuhi ciri sebagai berikut:

1. Watermark yang disisipkan tidak boleh menghasilkan suara yang menyimpang dari kualitas suara yang dimiliki data audio aslinya.
2. Pendeteksian dan komputasi yang dibutuhkan dari wtermark yang disisipkan haruslah rendah. Kompleksitas dari pendeteksian wtermark haruslah rendah untuk memfasilitasi integrasinya ke dalam produk elektronik pengguna.
3. Pendeteksian waternark harus dilakukan tanpa mencari keterangan dari data aslinya. Hal ini disebut sebagai *blind detection*.
4. Watermark seharusnya tidak dapt terdeteki tanpa pengetahuan sebelumnya terhadap aliran dari watermark yang disisipkan. Hal ini mencegah penyerang dari cara pembalikan proses penyisipan untuk menghilangkan watermark.
5. Watermark yang disissipkan harus kuat melawan penyerangan yang umum melawan proses sinyal seperti *filtering*, *resampling*, dan kompresi.
6. Watermark harus bertahan dari serangan-serangan seperti *random cropping* dan penambahan bunyi-bunyian. Namun serangan kuat yang menghasilkan bunyi-bunyian yang sangat mengganggu dapat diabaikan dari tes ketahanan.

3. Metode Pendahulu dalam *Audio Watermarking*

Banyaknya variasi metode *audio watermarking* dengan karakteristik yang berbeda telah disusulkan. Metode-metode ini akan dijabarkan dalam bagian ini.

Kegiatan awal dalam penyisipan *audio watermark* menghasilkan data yang tidak dapat terdengar dengan menempatkan sinyal watermark pada bagian frekuensi yang tidak signifikan. Pilihan yang paling populer adalah pada bagian frekuensi yang lebih tinggi, dimana penurunan pendengaran manusia menyamai puncaknya sekitar 1 kHz. Pada beberapa sistem, sinyal watermark difilter secara ketat sebelum dimasukkan ke dalam suara yang asli. Pada sistem yang lain, besarnya koefisien perubahan Fourier (Fourier Transform) di atas rentang frekuensi 2.4 kHz sampai 6.4 kHz diganti dengan rangkaian watermark. Dalam sistem ini, data yang tidak dapat terdengar semakin dikembangkan dengan hanya menyisipkan wtermark pada segmen suara dimana komponen frekuensi rendahnya memiliki nilai energi yang lebih tinggi. Sinyal frekuensi rendah yang kuat di dalam suara asli dapat membantu untuk menutupi sinyal kuat dari wtermark yang disisipkan.

Daerah lain dari manusia yang tidak peka adalah koefisien perubahan fase Fourier (*Fourier Transform*). Telinga manusia secara umum tidak peka terhadap penyimpangan fase, dan lebih spesifik lagi kurang nya kemampuan merasakan nilai fase secara mutlak. Skema yang diajukan untuk mengganti fase dari sebuah segmen awal suara dengan fase yang ditunjuk yang merepresentasikan wtermark. Fase pada segment berikutnya diatur menjaga fase relatif diantara segmen. Di dalam sistem yang lain, koefisien fase perubahan Fourier (Fourier Transform) yang dipilih dalam frekuensi yang lebih tinggi diabaikan dan nilai baru diberikan berdasarkan koefisien ketetanggaan yang ditunjuk. Watermark direpresentasikan oleh fase relatif antara koefisien yang diplih dan tetangga mereka. Masalah dalam skema watermarking yang menyembunyikan sinyal watermark di dalam area yang tidak signifikan adalah kurangnya kekuatan untuk menahan pemrosesan sinyal dan beberapa serangan. Algoritma kompresi tidak melindungi area ini secara baik, sehingga hacker dapat melakukan serangan yang lebih kuat terhadap area ini. Tanpa menghasilkan bunyi yang mengganggu.

Kelas algoritma lain menyisipkan wtermark sebagai sinyal gema dari suara asli. Data yang tidak dapat terdengar dari penyembunyian gema didasarkan pada teori bahwa resonansi adalah hal yang biasa dalam lingkungan manusia dimana biasanya manusia tidak merasakannya sebagai bunyi yang gaduh. Di dalam algioritma ini sinyal wtermark sebenarnya merupakan versi sinyal asli yang ditunda dan dilemahkan. Rangkaian watermark direpresentasikan dengan jumlah penundaan yang diterima kembali dengan mengamati puncak otokorelasi dalam domain waktu (*time domain*) atau di dalam domain cepstrum (*cepstrum domain*).

Akhir-akhir ini beberapa penelitian menggunakan konsep yang dipinjam dari komunikasi spektrum tersebar dan menyisipkan watermark sebagai bunyi acak yang semu si dalam domain waktu (*time domain*). Hal ini dijamin oleh teori spektrum tersebar bahwa waermark yang dsisipkan secara statistik tidak dapat terdeteksi oleh hacker. Karena telinga manusia memliki perbedaan kepekaan untuk bunyi tambahan di dalam sekumpulan frekuensi yang berbeda, semua kegiatan yang diajukan menggunakan beberapa penyaring untuk membentuk bunyi semu acak seperti spektrum sehingga didapatkan data yang tidak dapat terdengar. Penyaring sederhana sudah digunakan dalam sebuah karya, dan sebuah penyaring *nonlinear* dipakai pada karya yang lain. Pada sistem yang lain, daripada menyaring bunyi semu acak, sebuah skema dikembangkan untuk membangkitkan sinyal *band-limited* watermak semu acak. Data yang tidak dapat terdengar dari watermark yang disisipkan lebuh jauh dapat dijamin dengan menggunakan efek penutupan dari sistem pendengaran manusia. Satu sistem mengunakan MPEG-I Audio Psychoacoustic Model 1 untuk membentuk sinyal wtermark seperti spektrum, sementara sistem yang lain menggunakan model penutupan dari MPEG-II AAC. Deteksi watermark dilakukan dengan menhitung korelasi antara sinyal suara yang diwatermark dan sinyal dari watermark. Dengan dilengkapi teori komunikasi spektrum tersebar, jenis watermarking ini biasanya bertahan cukup kuat melawan penyimpangan suara dan serangan-serangan. Namun sinkronisasinya sulit untuk diterapkan dan membutuhkan *harga* komputasi yang tinggi.

Tren lain dalam *digital audio wtermarking* adalah dengan mengkombinasikan penyisipan watermark dengan kompresi atau proses modulasi. Integrasi dari keduanya dapat meminimalkan hubungan saling mengganggu yang tidak menguntungkan antara watermarking dan kompresi, khususnya mencegah mencegah penhilangan watermark oleh kompresi. Di dalam sebuah skema, penyisipan watermark dilakukan pada saat penhitungan vektor. Watermark disisipkan dengan cara mengubah kode vektor yang dipilih atau mengubah penyimpangan faktor penahan di dalam proses pencarian. Kebutuhan suara asli untuk mengekstrak watermark secara garis besar membatasi penerapan dari skema ini. Algoritma yang lain menyisipkan watermark secara langsung di dalam modulasi sigma delta *bitstream* untuk menyisihkan kebutuhan taransformasinya ke dalam data PCM, untuk menjaga agar *harga* komputasi tetap rendah. Hal ini penting untuk sistem modulasi sigma delta dimana penghematan perangkat keras adalah tujuan utamanya. Di dalam skema yang lain, watermraking didintegrasikan dengan kompresi MPEG-II AAC. Watermark disisipkan dengan memodofikasi koefisien kompresi yang dipilih seperti faktor skala.

4. Algoritma yang Diusulkan

Meskipun metode-mrtode yang dideskripsikan pada bagian tiga mempinya ciri dan keunggulan masing-masing, ada beberapa masalah yang sama dari metode-metode tersebut. Masalah ini dapat dihasilkan dari peng-*editan* suara secara sederhana seperti

pemotongan segmen suara yang tidak diinginkan atau serangan yang disengaja seperti penghapusan secara acak atau penambahan sample ke dalam data audio yang telah diwatermark. Serangan *random sample cropping* ini sangat efektif untuk ikut campur dalam proses pendeteksian watermark terhadap algoritma yang telah disebutkan di atas. Serangan ini memiliki kompleksitas penghitungan yang sangat rendah. Disamping itu, jika dilakukan secara benar, serangan ini akan menghasilkan suara yang mengganggu terhadap sinyal suara yang asli. Ada argumen yang mungkin mengatakan bahwa sebuah serangan yang sangat kuat hanya dapat dilakukan oleh para profesional saja dan tidak oleh sebagian besar konsumen. Namun, sekali metode watermarking digunakan secara luas, hampir dapat dipastikan bahwa beberapa penyerang profesional akan membuat dan menyebarkan alat penyerang sehingga sebagian besar pengguna biasa akan mampu melakukan serangan yang kuat. Sebuah metode telah diusulkan untuk memecahkan masalah sinkronisasi, dimana exhautive search digunakan dan sinyal suara asli dibutuhkan. Sebagai hasilnya, kompleksitas perhitungannya menjadi terlalu tinggi, dan kebutuhan dari suara asli untuk pendeteksian watermark telah membatasi penerapan metode ini. Lebih jauh lagi, metode ini hanya mampu menangani serangan peng-*edita*-n yang biasa, bukan untuk serangan *random sample cropping*.

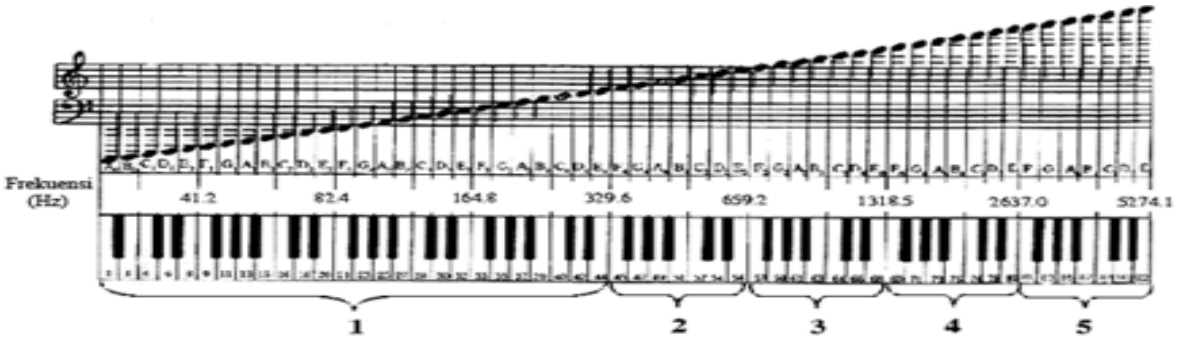
Dalam makalah ini diajukan solusi dengan kompleksitas rendah untuk mengatasi masalah sinkronisasi yang disebabkan oleh serangan yang biasa dan kuat. Solusi ini disusun oleh teknik ekstraksi yang menonjol dan prosedur penyisipan watermark *Fourier transform domain*. (Salient point) Ekstraksi melalui analisis *audio content* dilakukan pada saat penyisipan watermark dan proses deteksi sehingga sinkronisasi didapatkan pada masing-masing titik yang menonjol. Algoritma ekstraksi dirancang sehingga titik-titik penting tetap stabil setelah penyimpangan. Penyisipan watermark dan deteksi *Fourier Transform Domain* diambil karena informasi frekuensi domain sangat kecil terpengaruh oleh *sample cropping* dan *time domain*.

Salah satu karakteristik umum diantara algoritma *audio watermarking* yang ada adalah watermarknya disisipkan sepanjang keseluruhan sinyal suara. Namun, ini mungkin bukan cara yang paling efisien untuk menyisipkan dan mendeteksi watermark. Untuk penyerang yang mahir, perbedaan jumlah serangan dapat digunakan untuk segmen yang berbeda dari sinyal audio untuk mencegah timbulnya suara yang mengganggu. Sebagai contoh penghapusan secara acak sebuah *sample* dari seratus *sample* dalam segmen energi yang menggunakan daya suara yang tinggi dari sinyal suara akan menghasilkan suara yang dapat disadari, tetapi akibat dari melakukan ini di dalam segmen dengan energi yang rendah akan menghasilkan suara yang tidak dapat terdengar. Jadi, watermark yang disiipkan pada area yang dapat diserang secara kuat akan menghadapi serangan yang lebih kuat dan dapt juga dihancurkan oleh serangan tersebut. Kontribusi lain lain dari teknik ini adalah pengenalan dari daerah yang sensitif terhadap serangan melalui analisis *audio content.* Jika watermark hanya disisipkan dalam area yang sensitif terhadap serangan dimana serangan kecil yang akan dilakukan, kompleksitas penghitungan dari penyisipan dan deteksi watermark akan diuntungkan dari konsep yang sama.

4.1. Analisis *Audio Content* untuk Watermarking

Dalam sistem ini, analisis *audio content* dilakukan dengan tujuan ekstraksi poin menonjol dan identifikasi area yang sensitif terhadap serangan. Poin yang menonjol pada dalam sebuah sinyal suara mengizinkan deteksi watermark untuk sinkronisasi kembali pada lokasi-lokasi ini. Sinkronisasi oleh poin yang menonjol meiliki kompleksitas yang jauh lebih rendah dibandingkan *exhautive search* dan membuat deteksi blind watermark menjadi mungkin. Perlu dicatat bahwa di sini point yang menonjol tidak dimasukkan tetapi poin itu diekstrak dari suara mentah (*raw audio*) menggunakan analisis *audio content*. Pendekatan ini memiliki dua keuntungan disamping penysipan sinyal sinkronisasi. Keuntungan yang pertama yaitu pendekatan analisis *content* ini tidak menimbulkan penyimpangan dari sinyal suara asli karena di sini tidak ditambahkan sesuatu yang baru. Keuntungan yang lain adalah sinyal sinkronisasi yang ditambahkan secara jelas lebih mungkin untuk dipilih oleh penyerang.

Metode ekstraksi poin menonjol yang bagus seharusnya menghasilkan perkiraan aliran yang sama dari poin menonjol dalam sinyal suara sebelum dan setelah serangan seperti kompresi suara, penyaringan *low-pass*, dan penambahan suara. Untuk mendapatkan ini, poin menonjol diekstrak berdasarkan ciri suara yang peka terhadap telinga manusia. Dalam hal ini, jika penyerang ingin

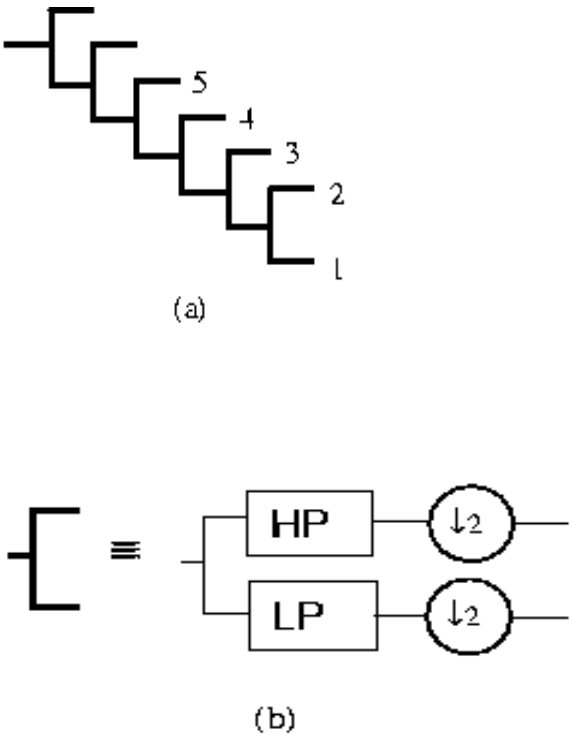


Gambar 1. Ilustrasi dari korespondensi antara not musik nilai frekuensi, dan 5 pembagian sub-*band*

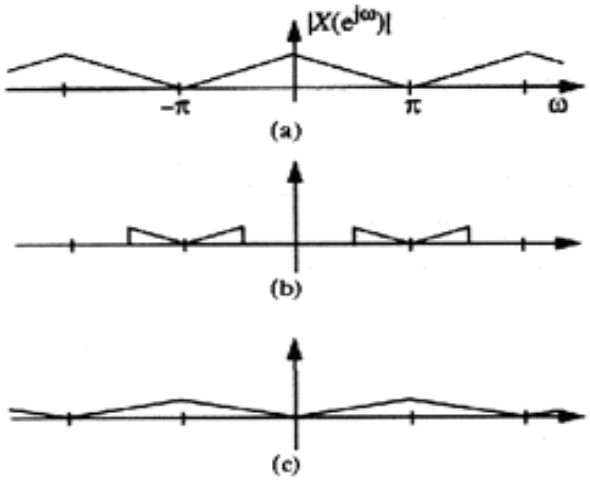
menghancurkan poin penting dia harus mengubah ciri ini dan menghasilkan penyimpangan yang disadari (noticable). Di sini dipilih variasi energi sebagai ciri utama untuk ekstraksi poin penting karena harga komputasinya yang rendah dan pengubahan dari ciri ini akan menghasilkan suara yang dapat didengar.

Skema dasarnya adalah untuk mengekstrak poin penting sebagai lokasi dimana energi sinyal suara secara cepat mencapai nilai puncaknya. Jika pendekatan ini bekerja dengan baik untuk musik sederhana denag sedikit instrumen, ini mempunyai dua masalah untuk musik yang lebih kompleks. Masalah yang pertama adalah variasi energi secara keseluruhan menjadi ambigu untuk musik yang rumit dimana banyak instrumen yang dimainkan bersama. Jadi, kestabilan dari point penting menurun. Masalah lain yaitu nilai ambang optimal yang berbeda untuk musik yang berbeda kompleksitasnya. Jika nilai ambang yang tinggi cocok untuk musik dengan variasi energi yang tajam, penerapan pada nilai yang sama untuk musik kompleks akan menghasilkan sangat sedikit poin penting.

Oleh karena itu, ada manfaatnya membagi musik yang kompleks menjadi beberapa musik yang sederhana sehingga kestabilan dari point penting dapat diperbaiki dan setiap ambang yang sama dapat diterapkan pada semua bagian musik. Musik yang komplek biasanya tersusun atas instrumen-instrumen yang frekuensi pokoknya menempati rentang frekuensi yang berbeda untuk membentuk sebuah harmoni. Gambar 1 memperlihatkan korespondensi antara not musik dengan nilai frekuensi. Pada gambar ini juga diperlihatkan pembagian untuk perancangan dalam makalah ini yang mengandung 5 rentang frekuensi. Perlu diketahui bahwa lebar frekuensi untuk masing-masing oktaf tidak sama. Interval frekuensi pada Gambar 1 berkorespondensi terhadap output dari 6-tingkat dekomposisis gelombang pandu *dyadic* di bawah dasar sample 44.1 kHz seperti yang diperlihatkan pada Gambar 2.



Gambar 2. Dekomposisi gelombang pandu *dyadic* 6-tingkat, dimana masing-masing cabang di (a) merepresentasikan struktur di (b) dan output dinomori berkorespondensi dengan *subband* pada gambar 1



Gambar 3. Efek dari pembalikan frekuensi akibat dari *downsampling* setelah penyaringan *high-pass* : (a) Spektrum sebelum penyaringan, (b) Spektrum setelah penyaringan *high-pass*, (c) Spektrum setelah *downsampling*, dimana frekuensi tertinggi pada (a) sekarang diperlihatkan dalam frekuensi terendah

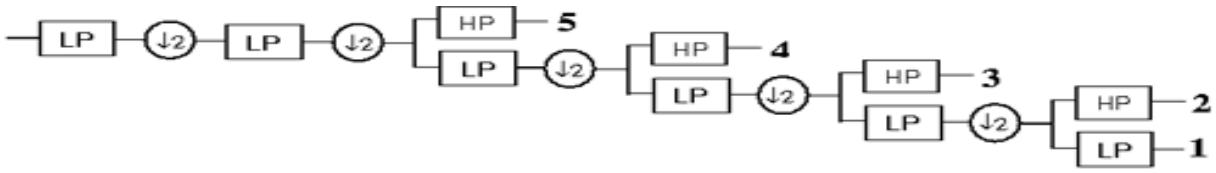
Untuk mencegah efek pembalikan frekuensi, yang disebabkan oleh penerapan dari *downsampling* untuk output dari penyaringan *high-pass* seperti yang diperlihatkan pada gambar 3, dekomposisi gelombang pandu *dyadic* pada Gambar 2 dimodifikasi menjadi seperti Gambar 4 dengan cara menyisihkan langkah *downsampling* setelah masing-masing penyaringan *high-pass*. Jadi, point penting diekstrak secara terpisah dari masing-masing lima output pada Gambar 4.

Prosedur dari identifikasi area yang peka terhadap serangan diarahkan pada penurunan kompleksitas penyisipan dan deteksi watermark. Jadi, hal ini penting bahwa proses identifikasi ini sendiri tidak membutuhkan banyak komputasi. Dalam makalah ini, proses identifikasi area yang peka terhadap serangan digabungkan ke dalam proses ekstraksi point penting sehingga hampir tidak adak komputasi tambahan yang dibutuhkan untuk identifikasi area yang peka terhadap serangan. Serangan yang dimaksud adalah serangan *random sample cropping*. Area peka serangan yang berhubungan adalah area nada energi tinggi. Karena poin penting yang dipilih oleh algoritma dalam makalah ini terletak pada posisi dimana sinyal energi suara cepat mencapai puncak, area yang mengikuti masing-masing poin penting akan mengandung energi yang tinggi. Secara mudah ini didefinisikan sebagai area peka serangan, sehingga dibutuhkan penghitungan tambahan.

4.2. Penyisipan dan Deteksi Watermark *Fourier Transform Domain*

Meskipun poin penting dipilih untuk memungkinkan kestabilan, sulit untuk mendapatkan poin penting yang sama secara tepat setelah beberapa pemrosesan suara seperti kompresi. Pemindahan sejumlah poin penting dalam lokasi merupakan hal yang biasa dan seharusnya dapat ditoleransi. Jika penyisipan dandeteksi watermark dilakukan dalam *time domain*, sangat jelas bahwa meskipun pemindahan poin penting hanya sejumlah kecil, hal ini akan menimbulkan masalah karena penyisipan dan deteksi tidak dapan disinkronisasikan. Namun masalah ini dapat dikurangi dengan mempertimbangkan besarnya koefisien dari DFT (*Discrete Fourier Transform*).

Properti ini dipaparkan pada Gambar 5 (pada halaman berikutnya), dimana $a(i)$, $i = 1, \ldots, 2^p$, adalah area yang diwatermark. Watermark disisipkan pada $|A(k)|$, $k = 1, \ldots, 2^p$, dimana $A(k)$ koefisien DFT (*DiscreteFourier Transform*) dari $a(i)$. Misalkan poin penting slah ditempatkan dalam proses deteksi, dan area yang diwatermark tertukar dengan area yang lain $b(i)$, $1 < i < 2^p$. Namun, hal ini merupakan properti yang diketahui bahwa $c(i)$ dibentuk dengan memindahkan bagian paling kanan dari $a(i)$ ke bagian paling kiri, lalu $c(i)$ dan $a(i)$ memilki besar koefisien DFT (*Discrete Fourier Transform*) yang sama, yaitu:



Gambar 4. Dekomposisi gelombang pandu *dyadic* yang telah dimodifikasi

$$|C(k)| = |A(k)|, k = 1, \ldots, 2^p \quad (1)$$

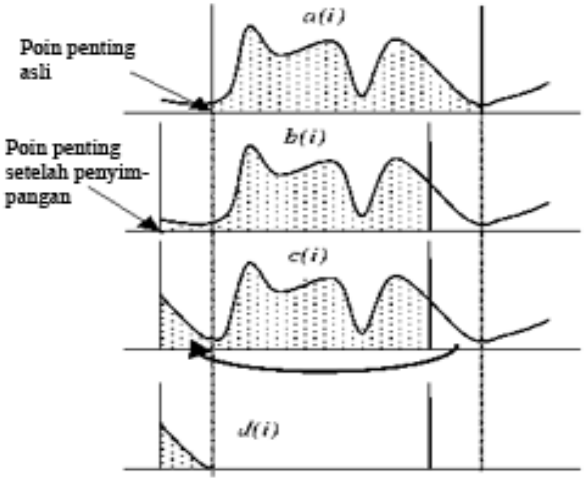
Ditunjukkan perbedaan antara $b(i)$ dan $c(i)$, dengan:

$$d(i) = c(i) - b(i), i = 1, \ldots, 2^p \quad (2)$$

Lalu, dimiliki:

$$|B(k)| \approx |C(k)| + |D(k)| = |A(k)| + |D(k)|, k = 1, \ldots, 2^p \quad (3)$$

Jadi, dari (3) dapat dilihat kesalahan yang disebabkan oleh poin penting adalah $|D(k)|$. Tidak ada efek salah sinkronisasi yang merugikan dalam *frequency domain*. Ketika jumlah pemindahan relatif kecil dibandingkan ukuran *window*, energi di dalam $|D(k)|$ juga kecil.



Gambar 5. Efek dari kesalahan penempatan poin penting pada watermarking DTF (*Discrete Fourier Transform*)

Agar watermark yang disisipkan menjadi tidak terdengar, hal ini biasa untuk memanfaatkan efek penyembunyian frekuensi sementara dari sitem pendengaran manusia. Penyembunyian sementara menunjuk pada efek abhwa sinyal yang lebih lemah segera sebelum dan sesudah sinyal yang lebih kuat mungkin tidak terdengar ketika penyembunyian frekuensi menunjuk pada efek dimana dua sinyal terjadi secara bersamaan dan memiliki frekuensi yang dekat, sinyal yang lebih kuat akan menyebabkan sinyal yang lebih lemah menjadi tidak terdengar.

Karena watermark dalam maklah ini hanya ditempatkan pada area yang peka terhadap serangan, yang mempunyai nilai energi yang tinggi, maka digunakan efek penyembunyian sementara. Sehingga watermark dengan energi yang lemah ditutupi oleh energi tinggi dari sample suara pada area ini. Untuk dapat mengambil keuntungan dari efek pnyembunyian frekuensi, skema yang diajukan hanya menyisipkan sinyal watermark di dalam rentang koefisien DFT (*Discrete Fourier Transform*) yang memiliki nilai yang luas.

Deteksi watermark dilakukan dengan cara mengkalkulasi koefisien korelasi rata-rata antara aliran watermark dan sinyal suara yang diwatermark dalam *Fourier Transform Domain* dam membandingkan ini dengan nilai ambang. Standar untuk memilih nilai ambang adalah untuk memperkecil harga yang diperkirakan untuk pendeteksian kesalahan. Perlu dicatat bahwa harga ketidaktepatan (yaitu kegagalan mendeteksi ketika da watermark) berbeda dengan tanda kesalahan (yaitu menyatakan ada deteksi padahal tidak ada watermark). Meskipun harga-harga ini bervariasi pada penerapan yang berbeda, secara umum benar bahwa harga tanda kesalahan lebih beasr daripada ketidaktepatan. Tanda kesalahan seharusnya sangat rendah karena ini mengurangi kredibilitas metode watermarking untuk membunktikan kepemilikan hak cipta. Sebaliknya batasan dari rating ketidaktepatan (atau gagal mendeteksi) tidak perlu diperketat, karena reting kegagalan mendeteksi dari 1% atau 10% mungkin mempunyai efek yang sama untuk menakuti orang agar tidak menggandakan data suara secara ilegal. Sebagai kesimpulannya ambang deteksi harus diatur relatif tinggi untuk meyakinkan tidak adanya kesalahan deteksi yang terjadi.

5. Hasil Penelitian

Properti seperti tidak dapat terdengar dan kuat dari skema watermarking yang diajukan dpertunjukkan dengan tiga bagian sinyal suara: komposisi Piano oleh Bach hanya dengan sebuah piano, simfoni “Bolero” oleh Ravel dengan terompet dan drum, dan lagu dengan vokal manusia dan musik latar belakang yang kompleks. Semua sinyal diambil contohnya pada frekuensi 44.1 kHz, dan masing-masing bagian berduarasi sekitar 30 detik.

5.1. Analisis *Audio Content*

Kefektifan dari analisis *audio content* yang diajukan diukur oleh kemampuannya untuk mengekstrak bagian sama dari poin penting pada sinyal audio sebelum dan sesudah penyerangan dan atau proses terhadap sinyal. Contoh dari perbandingan antara poin penting yang diekstrak dari file asli dan file yang telah diproses ditunjukkan pada Tabel 1 . seperti yang dapat dilihat dari contoh ini, hampir setiap poin penting lebih atau kurang bergeser beberapa poin. Namun, seperti yang dijelaskan ada bagian 4.2, hal ini tidak menyebabkan efek yang menimbulkan masalah pada deteksi watermark. Dalam kenyataanya penenmapatan yang tidak tepat pada poin kurang dari 100 menghasilkan penurunan yang sangat kecil untuk koefisien korelasi rata-rata pada deteksi watermark. Oleh karena itu, ini dapat dianggap sebagai ekstraksi poin penting yang berhasil. Beberapa poin penting mungkin hilang dan beberapa mungkin terciptakan setelah proses. Namun, lagi-lagi fenomena ini hanya menyebabkan kemunduran yang tipis dari hasil deteksi.

Rating kesuksesan dari ekstraksi secara benar terhadap poin penting dengan dan tanpa gelombang pandu dipaparkan pada Tabel 2. Serangan yang digunakan pada Tabel 2 adlah kompresi/dekompresi MP3. Penyimpangan seperti adanya suara gangguan tambahan, penyaringan *low-pass*, dan *downsampling* menyebabkan kesalahan penempatan poin penting yang lebih sedikit daripada kompresi/dekompresi MP3. Hal ini diamati bahwa semakin kompleks bagian musik, semakin rendah rating kesuksesannya. Namun, penggunaan dekomposisi gelombang pandu sangat efektif untuk menaikkan rating kesuksesan dari ekstraksi poin penting dari musik yang kompleks.

Tabel 1. Perbandingan antara poin penting yang diekstrak dari file asli dan file suara yang telah diproses dimana yang dicetak tebal dianggap sebagai kegagalan

Lokasi Poin Penting yang diekstrak dari file asli	Lokasi Poin Penting yang diekstrak dari file yang telah mengalami penyimpangan	Perubahan jumlah poin penting antara dua file
4401	**4557**	**-156**
6581	6581	0
14196	**Tidak ada**	
14463	**Tidak ada**	
19464	19471	-7
21092	21063	29
28657	28651	6
44152	44104	48
59635	59637	-2
91080	91126	-46
94883	94879	4
98548	98545	3
Tidak ada	**105946**	
112475	112471	4
127941	127958	-17
129319	129315	4
131028	131025	3
138454	138342	112
144478	**144489**	**-11**
145827	145823	4
153485	153484	1
185107	185056	49
192565	**192297**	**268**
216786	216784	2
224510	224555	-45
232790	232808	-18
242895	242878	17
264519	264518	1
271803	271803	0
273508	273507	1
297097	297097	0
304761	304760	1
320039	320013	26
335700	335700	0
343182	**343309**	**-127**
347048	**347233**	**-185**
351030	351003	27
359383	359351	32
382173	382186	-13
384255	384259	-4
389912	389914	-2
391882	391884	-2
397653	397654	-1
399407	**399526**	**-119**
406960	**Tidak ada**	
422233	422234	-1
426680	426682	-2
429936	**Tidak ada**	
437456	437373	-17
444820	444795	25
460640	460643	-3

Tabel 2. Rating kesuksesan dari ekstraksi poin penting dengan benar setelah melewati tiga operasi kompresi/dekompresi MPEG Layer III dengan *bit rate* 64 kbps

Tes Suara	Rating kesuksesan tanpa penyaringan gelombang pandu	Rating kesuksesan mengguna kan penyaringan gelombang pandu	Peningkatan rating kesuksesan
Piano tunggal	83.3 %	83.6 %	0.3 %
Musik dengan drum dan trompet	71.4 %	77.3 %	5.9 %
Vokal dengan musik latar belakang yang kompleks	63.0 %	73.1 %	10.1 %

5.2. Penyisipan Watermark

Kualitas dari metode watermark yang yang diajukan dievaluasi menggunakan tes *blind listening*. Dimana suara asli dan suara yang telah diwatermark diperdengarkan kepada pendengar tanpa diketahui suara mana yang telah diwatermark. Mereka ditanya untuk mengatakan suara mana yang memiliki kualitas yang lebih baik.

Persentasi suara mana yang lebih disukai diperlihatkan pada Tabel 3. Dari hasil yang didapatkan hampir setengah pendengar lebih menyukai suara yang telah diwatermark daripada suara yang asli. Sehingga dapat disimpulkan bahwa tidak ada penyimpangan suara yang terjadi dari penyisipan watermark.

Tabel 3. Tes blid listening terhadap bagian suara yang diwatermark

Tes Suara	Persentase yang lebih menyukai suara asli daripada yang diwatermark
Piano tunggal	45.5 %
Musik dengan drum dan trompet	54.5 %
Vokal dengan musik latar belakang yang kompleks	45.5 %

5.3. Deteksi *Blind* Watermark

Kekuatan dari algoritma *blind* watermark dites melawan beberapa jenis serangan termasuk penambahan suara, kompresi MPEG, *random cropping*, dan penyaringan *low-pass*. Kualitas dari deteksi watermark dievaluasi dengan perbandingan antara nilai kolerasi yang didapat dari ID pengguna yang benar kolerasi terbesar yang didaptakan dari 1000 ID pengguna lain secara acak. Perbandingan antara nilai korelasi dari ID pengguna yang benar dan korelasi terbesar yang didapatkan dari 1000 pengguna lain secara cak dirangkum dalam Tabel 4. Masing-masing jenis serangan menghasilkan perbedaan penurunan jumlah perbandingan puncak. Pengamatan yang dilakukan:

- Penambahan bunyi
  Bunyi dengan kekuatan 10 % dari sinyal suara ditambahkan. Bunyi pada tingkat ini dapat didengarkan dengan jelas, tapi hanya menyebabkan penurunan yang sedang pada perbandingan puncak.

- Kompresi MPEG
  Dalam aplikasi multimedia, kompresi meruapakan prosedur yang umum untuk efisiensi meningkatkan transmisi dan penyimpanan. Beberapa informasi dibuang pada saat proses kompresi, yang dapat menciptakan potensi bahya terhadap deteksi watermark. Untuk mengetes kekuatan dari pendekatan teknik watermark yang diajukan terhadap kompresi, sinyal suara dikompresi dan didekompresi menggunakan MPEG layer III denag *bit rate* 64 kbps. Seperti yang diperlihatkan pada Tabel 4, serangan ini lebih serius daripada yang lain. Namun,

waterark masih dapat terdeteksi secara benar.

- *Random Cropping*
  Secara acak memotong sebuah sampel dari 100 sampel yang mengakibatkan masalah sinkronisasi untuk metode watermarking *time domain*. Namun korelasi perbandingan puncak hanya sedikit menurun daripada metode yang diajukan.

- Penyaringan *low-pass*
  Dengan disisipkannya watermark dalam domain frekuensi, penyaringan *low-pass* dengan pemotongan frekuensi dapat menjadi efektif menghilangkan watermark yang disisipkan, namun karena watermark disisipkan pada pita frekuensi dengan energi yang paling tinggi, penyaringan terhadap watermark yang disisipkan juga sangat mempengaruhi kualitas suara.

Seperti yang diperlihatkan pada Tabel 4, korelasi perbandingan puncak setelah berbagai macam serangan diperkirakan antara 1.5-2.5. Nilai ini dapat ditingkatkan jika watermark disisipkan pada setiap bagian sinyal suara, atau jika suara asli digunakan dalam deteksi watermark. Namun krelasi perbandingan pada Tabel 4 sudah cukup tinggi untuk deteksi watermark yang tidak ambigu. Keefisienan yang didapatkan dengan deteksi *blind watermark* dan penyisipan pada area yang sensitif terhadap serangan sangat penting digunakan untuk watermark suara.

Tabel 4. Perbandingan antara korelasi puncak dengan ID pengguna yang benar dan korelasi pada 1000 percobaan acak yang terbesar

Jenis Serangan	**Piano Tunggal**	**Drum dan Trompet**	**Vokal dengan musk latar belakang yang kompleks**
Tidak ada serangan	2.63	2.56	2.17
Penambahan Suara	2.44	2.11	1.70
Kompresi MPEG	2.14	1.98	1.51
Random Cropping	2.25	2.08	1.76
Penyaringan *low-pass*	2.07	1.92	1.71

6. Kesimpulan

Perkembangan teknologi multimedia yang sangat cepat memfasilitasi produksi dan transmisi dari data media digital. Hal ini tidak hanya memberikan kesempatan tapi juga tantangan untuk perlindungan terhadap hak cipta. Skema watermarking suara yang memenuhi kekuatan dan kompleksitas penghitungan yang rendah mealuli analisis *audio content* telah dipaparkan dalam makalah ini. Analisis mengidentifikasikan area peka serangan yang cocok untuk menempatkan watermark, dan menyediakan hasil segmentasi suara yang konsisten sebelum dan sesudah serangan. Penyaringan gelombang pandu yang dimodofikasi digunakan untuk memperkuat hasil analisis untuk musik yang kompleks. Setelah analisi *audio content*, skema penyisipan watermark dikembankan dalam domain perubahan Fourier yang mempergunakan efek penyembunyian frekuensi sementara pada sistem pendengaran manusia. Watermark yang disisipkan tidak dapat didengar. Solusi watermark yang diajukan, yang mengkombinasikan sinkronisasi ciri dari analisis *audio content*, menyediakan kompleksitas yang rendah dalam mengahdapi serangan.

7. Daftar Pustaka

[1] Nedeljko Cvejic, Tapio Seppänen “Fusing Ditgital Audio
Watermarking and Authentication in Diverse Signal Domain”,
University Of Oulu, Finland.

[2] Chung-Ping Wu, Po-Chyi Su, C.-C. Jay Kuo, “Robust Audio
Watermarking for Copyright Protection”, University of Southern
California, Los Angeles, 1999.

[3] Chung-Ping Wu, Po-Chyi Su, C.-C. Jay Kuo, “Robust and Efficient
Digital Audio Watermarking”, University of Southern California,
Los Angeles, 2000.

[4] Audio Box, “Digital Audio Watermarking”,

http://www.ece.uvic.ca, 2006.

[5] Joong, Kim Hyoung. Audio Watermarking Techniques. Department of Control and Instrumentation Engineering Kangwon National University Chunchon 200-701, Korea.

[6] Munir, Rinaldi. Diktat Kuliah IF5054 Kriptografi. Departemen Teknik Informatika, Institut Teknologi Bandung. 2006