STUDI DAN IMPLEMENTASI *WATERMARKING* CITRA DIGITAL DENGAN MENGGUNAKAN FUNGSI *HASH*

Fahmi

Program Studi Teknik Informatika, Sekolah Teknik Elektro dan Informatika, Institut Teknologi Bandung
Jl. Ganesha No. 10 Bandung
e-mail: huicassava@yahoo.com

ABSTRAK

Penyisipan *watermark* digital dilakukan sebagai salah satu cara untuk menandai kepemilikan hak cipta atas sebuah citra digital. Salah satu cara penyisipan *watermark* yang dapat dilakukan adalah penyisipan pada ranah frekuensi. Citra digital ditransformasikan dengan DCT (*Discrete Cosine Transform*), diubah koefisien-koefisiennya dan kemudian ditransformasikan kembali dengan IDCT (*Inverse Discrete Cosine Transform*). Pengubahan dilakukan terhadap beberapa koefisien frekuensi menengah yang terpilih untuk mendapatkan keseimbangan antara *robustness* dan *fidelity*. Untuk menghasilkan sinyal *watermark* yang aman, diperlukan ketergantungan pola bit *watermark* terhadap citra asli. Ketergantungan ini dapat diperoleh dengan menggunakan fungsi *hash*. Untuk menghasilkan ketergantungan terhadap citra asli, digunakan fungsi *hash* MD5 yang digabungkan dengan algoritma enkripsi DES (*Data Encryption Standard*) untuk menghasilkan masukan untuk pembangkit bilangan acak (PRNG, *Pseudo Random Number Generator*). Bilangan acak yang dihasilkan kemudian digunakan dalam pengolahan data *watermark*. Hasil pengujian menunjukkan bahwa teknik penyisipan *watermark* yang diterapkan dapat memenuhi kriteria *security* dan *fidelity*. Selain itu, *watermark* yang disisipkan memiliki *robustness* terhadap operasi-operasi manipulasi citra sampai batas tertentu.

Kata kunci : *watermark*, DCT, ranah frekuensi, fungsi *hash*, MD5, DES, PRNG.

Makalah diterima 20 Juli 2007. Revisi akhir 20 Juli 2007.

1. PENDAHULUAN

Perkembangan komputer memberikan banyak kemudahan di bidang multimedia digital yang memberikan keungulan dibandingkan dengan multimedia konvensional. Di antaranya adalah kemudahan dalam penyalinan dan penyebaran sebuah arsip multimedia digital. Selain memberikan nilai positif, hal ini juga menimbulkan dampak negatif, yaitu jika penyalinan dan penyebaran arsip multimedia dilakukan secara ilegal. Ini merupakan ancaman terhadap hak kekayaan intelektual dari pemilik arsip multimedia yang bersangkutan. Salah satu cara untuk melindungi hak kekayaan intelektual tersebut adalah dengan melakukan *watermarking*.

Untuk meningkatkan keamanan, data *watermark* terlebih dahulu diacak dan disebarkan pada data arsip penampung. Hal ini dilakukan dengan menggunakan pola bilangan acak. Pola bilangan acak dibuat bergantung kepada citra asli dengan menggunakan fungsi *hash* dalam proses menghasilkan *seed*, yaitu bilangan yang digunakan sebagai nilai awal dalam pembangkitan bilangan acak. Fungsi *hash* digunakan karena setiap bit yang dihasilkan fungsi *hash* bergantung kepada setiap bit pada pesan asli.

2. *DIGITAL WATERMARKING*

Digital watermarking adalah penambahan data rahasia (*watermark*) ke dalam sebuah arsip digital. *Watermark* berisi informasi yang berkaitan dengan arsip penampungnya. *Watermark* dapat berupa data teks, citra, maupun suara. Penyisipan *watermark* pada arsip citra dapat dilakukan pada ranah spasial maupun pada ranah frekuensi. Penyisipan pada ranah frekuensi memiliki keunggulan dibandingkan dengan penyisipan pada ranah spasial. Pada ranah frekuensi, sebuah modifikasi akan mempengaruhi keseluruhan *pixel* dalam blok. Dengan begitu, kemungkinan rusaknya *watermark* oleh manipulasi citra akan menjadi lebih kecil.

Ada beberapa kriteria yang perlu dipenuhi dalam *digital watermarking*, yaitu:

- *Robustness*, yaitu ketahanan *watermark* terhadap manipulasi yang dilakukan pada arsip penampungnya.
- *Fidelity*, yaitu perbandingan antara kualitas arsip penampung setelah penyisipan *watermark* dengan kualitas arsip semula. Pada penyisipan yang baik, perubahannya tidak dapat dikenali oleh manusia.

- *Recovery*, yaitu pengungkapan terhadap data yang disembunyikan. *Watermark* yang disisipkan harus dapat diambil kembali.
- Security, yaitu keamanan *watermark*. *Watermark* tidak boleh terdeteksi oleh pihak lain, sekalipun algoritma penyisipannya bersifat publik.

Pengukuran *fidelity* dapat dihitung dengan menghitung nilai MSE (*Mean Squared Error*) dan PSNR (*Peak Signal to Noise Ratio*). MSE dan PSNR dapat dihitung dengan persamaan (1) dan (2). Pada persamaan (1), $I(x,y)$ adalah nilai *grey-level* citra asli di posisi (*x,y*), *I'* adalah nilai derajat keabuan citra yang telah diberi *watermark* di posisi (*x,y*), *X* dan *Y* adalah ukuran panjang dan lebar. Pada persamaan (2), m adalah nilai maksimum yang mungkin dimiliki oleh sebuah *pixel*. Sebagai contoh, untuk data citra 8 bit, nilai maksimumnya adalah 255.

$$MSE = \frac{1}{XY}\sum_x\sum_y\left[I(x,y) - I'(x,y)\right]^2 \qquad (1)$$

$$PSNR = 10\log\frac{m^2}{MSE} \qquad (2)$$

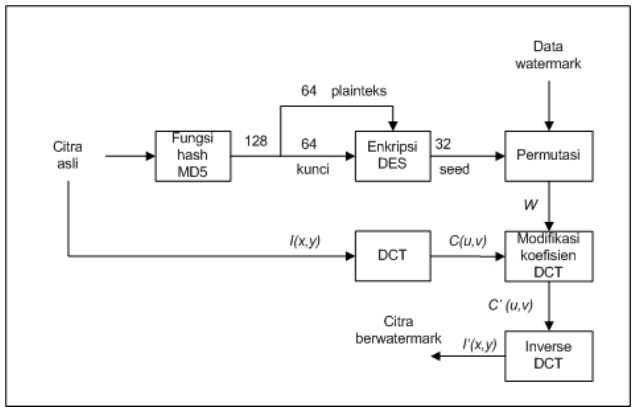
Penghitungan nilai PSNR hanya terdefinisi dengan baik pada pengukuran *luminance* (intensitas cahaya, misalnya pada citra *grayscale*). Untuk citra berwarna, pendekatan yang dapat dilakukan adalah menghitung nilai PSNR masing-masing kanal dan kemudian menghitung nilai rata-ratanya.

3. METODE

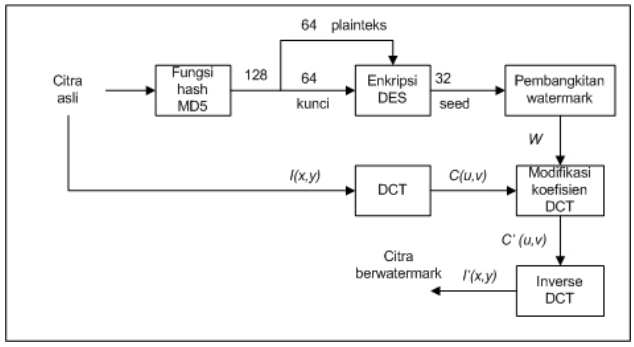
Pada makalah ini, penyisipan *watermark* dilakukan pada ranah frekuensi. Transformasi citra dilakukan dengan menggunakan DCT (*Discrete Cosine Transform*), sehingga dapat dikatakan bahwa penyisipan dilakukan pada ranah DCT. Penyisipan dilakukan terhadap citra bitmap dengan kedalaman warna 24 bit.

Pada makalah ini, ada dua metode yang digunakan untuk menyisipkan *watermark*. Perbedaan kedua metode ini adalah pada data *watermark* yang disisipkan. Metode pertama menggunakan citra biner sebagai *watermark*, sedangkan metode kedua menggunakan *watermark* yang dibangkitkan secara acak. Diagram kedua metode ini disajikan pada gambar 1 dan gambar 2. *I* melambangkan matriks nilai piksel citra asal, *C* melambangkan matriks koefisien DCT citra asal, *W* melambangkan data *watermark* yang disisipkan, *C'* melambangkan matriks koefisien DCT yang sudah dimodifikasi, dan *I'* melambangkan matriks nilai piksel sesudah penyisipan *watermark*.

Khusus untuk *watermark* yang dibangkitkan secara acak, ukuran *watermark* yang disisipkan kira-kira setara dengan ukuran citra biner yang memiliki panjang dan lebar masing-masing setengah dari panjang dan lebar citra asli.



Gambar 1. Penyisipan *watermark* dari citra biner



Gambar 2. Penyisipan *watermark* yang dibangkitan secara acak

Seperti yang terlihat pada gambar, citra asli ditransformasikan ke ranah DCT, dan kemudian dilakukan modifikasi koefisien DCT. Setelah itu citra ditransformasikan kembali ke ranah spasial dengan menggunakan IDCT.

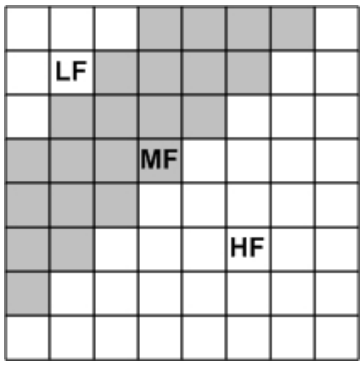
Modifikasi koefisien DCT dilakukan berdasarkan persamaan di bawah ini:

$$c_i' = c_i\,(1 + \alpha w_i) \qquad (3)$$

c_i' adalah koefisien DCT setelah modifikasi, c_i koefisien DCT sebelum modifikasi, α adalah kekuatan penyisipan *watermark*, dan w_i adalah nilai bit *watermark* yang disisipkan. Nilai w_i adalah -1 jika bit yang akan disisipkan adalah 0, dan 1 jika bit yang akan disisipkan adalah 1. Nilai kekuatan penyisipan $\alpha = 0{,}1$ jika nilai absolut dari c_i lebih besar atau sama dengan 15, dan $\alpha = 0{,}5$ untuk nilai c_i yang lain.

Penyisipan bit *watermark* dilakukan pada koefisien frekuensi menengah. Hal ini dilakukan sebagai *trade-off* antara *fidelity* dan *robustness*. Penyisipan pada frekuensi

rendah dapat menyebabkan perubahan yang terlalu besar, sedangkan penyisipan pada frekuensi tinggi dapat menyebabkan *watermark* mudah rusak. Koefisien yang dimodifikasi adalah koefisien-koefisien bernilai tertinggi sebanyak bit yang akan disisipkan pada sebuah blok koefisien DCT. Pengelompokan frekuensi pada sebuah blok DCT dapat dilihat pada gambar 3. LF menyatakan koefisien frekuensi rendah, MF menyatakan koefisien frekuensi menengah, dan HF menyatakan koefisien frekuensi tinggi.



Gambar 3. Pengelompokan koefisien DCT

Bit-bit *watermark* yang disisipkan adalah hasil dari proses permutasi (untuk *watermark* dari citra biner) maupun proses pembangkitan *watermark*. Kedua proses ini menggunakan pola bilangan acak. Pembangkitan pola bilangan acak dilakukan dengan menggunakan metode *linear congruential generator* (LCG) seperti yang terdapat pada [2]. Persamaan yang digunakan adalah sebagai berikut:

$$I_{j+1} = (aI_j + c) \bmod m \qquad (4)$$

dengan a=16807, c=0, dan m=2147483647.

Untuk menghasilkan *watermark* yang aman, maka pola bilangan acak harus bergantung kepada citra asli. Untuk menghasilkan ketergantungan ini digunakan fungsi *hash*. Seperti yang dapat dilihat pada gambar 1 dan gambar 2, citra asli di-*hash* dengan algoritma MD5. Nilai *hash* yang dihasilkan dibagi dua dan kemudian di lakukan enkripsi DES dengan menggunakan bagian pertama sebagai plainteks dan bagian kedua sebagai kunci. 32 bit pertama dari hasil enkripsi ini dijadikan sebagai *seed* untuk pembangkit bilangan acak. *Seed* yang dihasilkan ini bergantung kepada citra asli, sehingga pola bilangan acak yang dihasilkan juga bergantung kepada citra asli.

Ekstraksi dilakukan dengan melakukan transformasi DCT terhadap citra ber-*watermark* dan membandingkannya dengan koefisien citra asli untuk mendapatkan pola bit. Pola bit yang terdeteksi dibandingkan dengan pola bit yang disisipkan untuk mengetahui keberadaan *watermark*.

Perbandingan antara *watermark* yang terdeteksi dengan *watermark* asli dilakukan dengan menghitung koefisien korelasi. Perhitungan dilakukan dengan persamaan berikut ini:

$$C(W,W') = \frac{\sum w_i w_i'}{\sqrt{\sum w_i^2}\sqrt{\sum w_i'^2}} \qquad (5)$$

C(W,W') adalah koefisien korelasi, *W* adalah *watermark* asli, *W'* adalah *watermark* terdeteksi, w_i adalah bit *watermark* asli ke-*i*, w_i' adalah bit *watermark* terdeteksi ke-*i*.

4. PENGUJIAN

Watermarking dilakukan terhadap beberapa arsip citra dengan berbagai ukuran dengan ukuran *watermark* yang berbeda-beda. Beberapa hasil penyisipan *watermark* dapat dilihat pada gambar 4 dan gambar 5.

Beberapa skenario pengujian dilakukan untuk menguji metode *watermarking* yang digunakan. Skenario-skenario tersebut adalah:

- Penggunaan normal.
- Ekstraksi dari citra yang tidak diberi *watermark*.
- Ekstraksi dengan citra asli yang dimodifikasi.
- Dua kali penyisipan *watermark*.
- Ketahanan terhadap *cropping*.
- Ketahanan terhadap *scaling*.
- Ketahanan terhadap rotasi.
- Ketahanan terhadap kompresi JPEG.
- Ketahanan terhadap *screen capture*.

5. KESIMPULAN

Hasil pengujian menunjukkan bahwa untuk ukuran citra yang sama, penyisipan *watermark* yang lebih besar menghasilkan perubahan yang lebih besar. Hal ini terlihat dari penurunan nilai PSNR seiring dengan peningkatan ukuran *watermark*.

Watermark yang disisipkan sudah memenuhi kriteria keamanan. *Watermark* tidak dapat dideteksi ketika citra asli diubah meskipun hanya 1bit. Artinya fungsi *hash* dapat digunakan untuk menghasilkan *watermark* yang aman.

Penyisipan w*atermark* juga sudah memenuhi kriteria *fidelity*, karena perubahan yang diakibatkan tidak dapat dikenali oleh mata manusia.

Watermark memiliki *robustness* yang terbatas terhadap masing-masing modifikasi citra. Secara umum, *robustness* ditentukan oleh jenis dan besarnya manipulasi yang dilakukan terhadap citra.

Pada *watermark* yang berasal dari citra biner, hasil deteksi masih dapat dikenali ketika nilai koefisien

(a) citra asli

(b) citra *watermark*

(c) citra ter-*watermark*

Gambar 4. *Watermarking* dengan *watermark* dari citra biner

(a) citra asli

(b) citra ter-*watermark*

Gambar 5. *Watermarking* dengan *watermark* yang dibangkitkan secara acak

korelasinya lebih besar atau sama dengan 0,40. Nilai ini dapat dijadikan nilai batas untuk menentukan keberadaan *watermark*, terutama untuk *watermark* yang dibangkitkan secara acak.

REFERENSI

[1] Reka Major, Valentin Deac, Monica Borda, "An Application of the Hash Functions in Digital Watermarking", 2002.

[2] W. H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, "Numerical Recipes in C. The Art of Scientific Computing", Cambridge University Press, 1997.

[3] Jiri Fridrich, "Application of Data Hiding in Digital Images", 1998.

[4] Juan R. Hernandez, "DCT-Domain Watermarking Techniques for Still Images: Detector Performance Analysis and a New Structure", 2000.