

DENGAN WATERMARKING PADA DOMAIN WAVELET

Ady Priyoyudo, Aris Sugiharto, Indriyati
Jurusan Matematika FMIPA UNDIP
Jl. Prof. Soedarto, Kampus UNDIP Tembalang, Semarang

Abstrak: *Watermarking* merupakan salah satu solusi dalam memecahkan permasalahan pembuktian kepemilikan data digital. Pada penelitian ini teknik *watermarking* pada citra digital menggunakan *wavelet* sebagai media transformasinya (*DWT*). Citra original ditransformasi menggunakan *wavelet* menjadi empat area frekuensi LL, LH, HL, dan HH. Bit-bit *watermark* ditanam pada area LH dan HL. Kualitas citra terwatermark diamati berdasarkan nilai *Peak Signal of Noise Ratio (PSNR)*.

Kata Kunci : *watermarking, watermark, wavelet, DWT, PSNR*

PENDAHULUAN

Perkembangan teknologi komputer saat ini telah membawa perubahan pada zaman ini. Sehingga zaman ini sering disebut sebagai Zaman Digital, atau Dunia Digital atau Dunia Bit (disebut demikian karena istilah bit sangat mendasar dalam masalah digital). Dengan perkembangan komputer digital dan perangkat-perangkat lainnya yang serba digital, telah membuat data digital semakin banyak digunakan dan mudah diduplikasi. Sehingga seringkali menimbulkan konflik. Konflik yang sering timbul adalah adanya sengketa antara beberapa pihak yang mengklaim bahwa pihaknya adalah pemilik sah dari sebuah citra digital.

Melihat hal ini, telah banyak praktisi Pengolahan Citra yang berusaha menyelesaikan permasalahan ini. Salah satu tekniknyanya adalah dengan menggunakan teknik *Watermarking* pada citra digital.

Watermark adalah suatu sinyal atau *pattern* yang disisipkan ke dalam citra asli (*original image*). *watermark* ini mempunyai tujuan yang sama yaitu memproteksi terhadap pelanggaran *watermarked image (original image yang telah disisipi watermark)*, misalnya penggandaan dan pembajakan *copyright*.

Penelitian ini ditujukan untuk mengetahui seberapa tingkat kualitas citra digital setelah disisipi *watermark* dengan teknik *watermarking* pada domain *wavelet*.

TINJAUAN PUSTAKA

Watermarking

Watermarking merupakan suatu bentuk dari *Steganography* (Ilmu yang mempelajari bagaimana menyembunyikan suatu data pada data yang lain) [4], dalam mempelajari teknik-teknik bagaimana penyimpanan suatu data (digital) ke dalam data *host* digital yang lain (Istilah *host* digunakan untuk data/sinyal digital yang ditumpangi).

Watermarking ini memanfaatkan kekurangan-kekurangan sistem indera manusia seperti mata dan telinga. Dengan adanya kekurangan inilah, metoda *watermarking* ini dapat diterapkan pada berbagai media digital.

Jadi *watermarking* merupakan suatu cara untuk penyembunyian atau penanaman data/informasi tertentu (baik hanya berupa catatan umum maupun rahasia) ke dalam suatu data digital lainnya, tetapi tidak diketahui kehadirannya oleh indera manusia (indera penglihatan atau indera pendengaran), dan mampu menghadapi proses-proses pengolahan sinyal digital sampai pada tahap tertentu [4].

Semua aplikasi dari *watermarking* tersebut, menuntut parameter yang berbeda dari penerapan metoda *watermarking*. Parameter-parameter yang perlu diperhatikan dalam penerapan metoda *watermarking*:

1. Jumlah data (*bitrate*) yang akan disembunyikan.
2. Ketahanan (*robustness*) terhadap proses pengolahan sinyal.



Gambar 1. *Trade-Off* dalam *Watermarking*

Terdapat dua proses dalam *watermarking*, yaitu proses penyisipan dan proses pengekstrakan. Proses penyisipan adalah proses menyisipkan *watermark* ke dalam media digital yang akan disisipi (dalam penelitian ini digunakan citra digital). Untuk menyisipkan suatu *watermark* ke dalam citra digital, diperlukan adanya bilangan kunci (*key*). Bilangan kunci ini terdiri dari beberapa digit bilangan yang akan membangkitkan bilangan random tertentu sebagai kunci dalam proses *watermarking*. Bilangan kunci ini berfungsi sebagai kunci dalam proses pengekstrakan kembali *watermark* yang telah tertanam.

Gambar 2. Proses Penyisipan *Watermark*

Sedangkan proses pengekstrakan adalah proses mengekstrak kembali *watermark* yang telah tertanam pada citra terwatermark.

Untuk proses pengekstrakan ini diperlukan bilangan kunci yang sama dengan bilangan kunci yang dipakai dalam proses penyisipan. Bilangan kunci ini akan membangkitkan bilangan random yang sama dengan bilangan random pada saat proses penyisipan.

Selain membutuhkan bilangan kunci dalam proses pengekstrakan, juga membutuhkan *watermark* asli dari citra terwatermark sebagai pembanding ukuran dalam membentuk kembali piksel-piksel *watermark* yang telah tersembunyi dalam citra terwatermark tersebut.

Gambar 3. Proses Pengekstrakan *Watermark*

Sifat dan Manfaat Watermarking

Untuk mendapatkan suatu teknik *digital watermarking* yang baik, maka teknik tersebut harus dapat memenuhi sifat-sifat di bawah ini [4] :

1. Elemen dari suatu data digital dapat secara langsung dimanipulasi dan informasi dapat ditumpangkan ke dalam data digital.
2. Penurunan kualitas dari data digital setelah dibubuhkan *watermark*, dapat seminimal mungkin.
3. *Watermark* dapat dideteksi dan diperoleh kembali meskipun setelah data digital diubah sebagian, dikompresi, ataupun di-*filter*.
4. Struktur dari *watermark* membuat penyerang sulit untuk mengubah informasi yang terkandung di dalamnya.
5. Proses untuk membubuhkan *watermark* dan mendeteksinya cukup sederhana.
6. Jika *watermark* dihapus, maka kualitas dari data digital yang ditumpanginya akan berkurang jauh atau bahkan rusak sama sekali.
7. Informasi *watermark* yang diselipkan dalam isi data digital dapat dideteksi ketika dibutuhkan.
8. Label hak cipta yang unik mengandung informasi pembuatan, seperti nama, tanggal, dan sebagainya, atau sebuah kode hak cipta seperti halnya ISBN (*International Standard for Book Notation*) pada buku-buku.
9. *Watermark* tidak dapat diubah atau dihapus (*robustness*) secara langsung oleh orang lain atau dengan menggunakan *software* pengolahan sinyal sampai tingkatan tertentu.
10. *Watermarking* yang diberikan lebih dari satu kali dapat merusak data digital aslinya. Cara ini dilakukan supaya orang lain tidak dapat melakukan pelabelan berulang terhadap data yang telah dilabel.

Ada beberapa manfaat yang ingin dicapai dari penggunaan *watermarking*, sebagai suatu teknik penyembunyian data pada data digital lain [4] , yaitu :

1. *Tamper-proofing*

Watermarking digunakan sebagai alat indikator yang menunjukkan apakah data digital yang asli telah mengalami perubahan dari aslinya (mengecek integritas data).

2. *Feature location*

Watermarking sebagai alat identifikasi isi dari data digital pada lokasi-lokasi tertentu, misalnya penamaan suatu objek tertentu dari beberapa objek yang ada pada suatu citra digital.

3. *Annotation/caption*

Watermark berisi keterangan tentang data digital itu sendiri, misalnya pada *broadcast monitoring* pada penayangan iklan di stasiun TV (*Cox, 2000*). Selain itu, *watermark* juga dapat digunakan untuk mengirimkan pesan rahasia.

4. *Copyright-Labeling*

Watermarking digunakan sebagai metoda untuk menyembunyikan label hak cipta pada data digital atau sebagai bukti autentik kepemilikan atas dokumen digital tersebut.

Transformasi Wavelet Diskrit

Transformasi pada *watermarking* digunakan sebagai penyederhanaan proses. *Wavelet* yang digunakan berupa basis dari ruang vektor dua dimensi. Dalam hal ini citra digital direpresentasikan dalam bentuk matrik $M \times N$ [3].



dengan $f(x,y)$ merupakan fungsi :



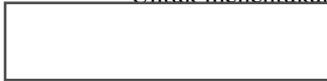
Transformasi *Wavelet* Diskrit / *Discret Wavelet Transform (DWT)* merekonstruksi matrik citra ke dalam empat daerah koefisien wavelet LL, LH, HL, dan HH.

Gambar 4. Transformasi Wavelet Diskrit

Peak Signal to Noise Ratio (PSNR)

Peak Signal to Noise Ratio (PSNR) merupakan nilai (rasio) yang menunjukkan tingkat toleransi *noise* tertentu terhadap banyaknya *noise* pada suatu sinyal video/citra. *Noise* adalah kerusakan sinyal pada bagian tertentu dalam sebuah video/citra sehingga mengurangi kualitas sinyal tersebut. Dengan kata lain *PSNR* merupakan suatu nilai yang menunjukkan kualitas suatu sinyal video/citra [5].

Untuk menentukan nilai *PSNR* digunakan rumus :



Sedangkan *MSE (Mean Square Error)* adalah tingkat kesalahan sinyal-sinyal video atau piksel-piksel citra hasil pemrosesan sinyal terhadap sinyal/citra original. Rumus untuk menghitung *MSE* pada citra digital adalah [4]:



(untuk Citra Grayscale)



(untuk Citra RGB)

Dimana :



: Piksel citra hasil pemrosesan.



: Piksel citra original.

i : index matriks (*Red = 1, Green = 2, dan Blue = 2*)

PEMBAHASAN

Metode

Pada penelitian ini semua bahan yang digunakan adalah citra digital yang mudah diperoleh di berbagai media. Metode yang digunakan dapat digambarkan sebagai berikut (gambar 5) :

Pertama, citra original disisipi citra *watermark* menghasilkan citra terwatermark. Kualitas citra terwatermark ini kemudian diuji dengan parameter yang digunakan adalah *Peak Signal to Noise Ratio (PSNR)*.

Selanjutnya citra terwatermark yang telah diketahui nilai *PSNR*, diekstrak menghasilkan citra *watermark* terekstrak. *Watermark* terekstrak ini hanya digunakan sebagai pembandingan kemiripan secara visual dengan citra *watermark* asli.

Metode ini dilakukan berulang-ulang dengan konstanta transformasi yang berbeda-beda. Citra digital dengan nilai *PSNR* tertentu dapat dikategorikan ke dalam 5 kategori sebagaimana tabel 1 [3]:

Tabel 1 Kategori *PSNR*

Gambar 5. Metode Uji Kualitas Citra Terwatermark

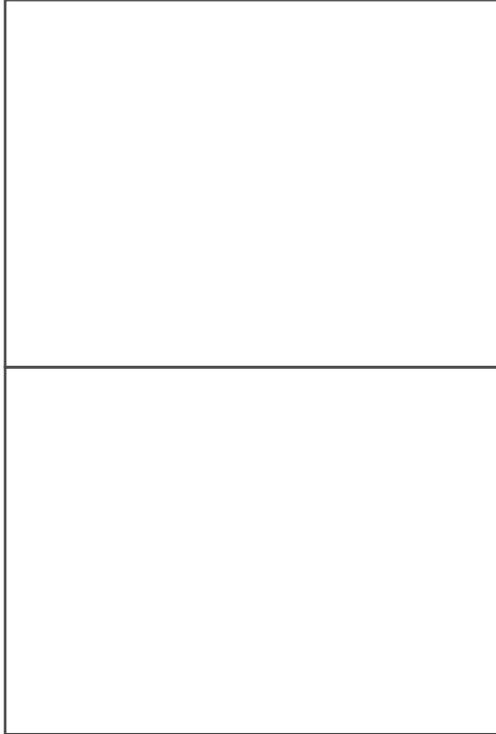
Hasil

Citra yang disimulasikan adalah citra *Mandril256RGB.bmp*, berukuran 256 x 256 piksel, dengan jenis citra *RGB* 24 bit. Sedangkan citra *watermark*-nya adalah citra *watermark1.bmp* berjenis *grayscale* 8 bit, berukuran 20 x 50 piksel.

(a) (b)

Gambar 6. (a) Mandril256RGB.bmp 256x256 piksel, (b) watermark1.bmp 20x50 piksel

Kedua citra di atas kemudian disimulasikan dalam sistim watermarking dengan menggunakan *Graphical User Interface (GUI) MATLAB 7.1* [2] yang diperlihatkan pada Gambar 7.



(a)

(b)

Gambar 7. (a)Penyisipan/Penanaman *Watermark*, (b)Pengekstrakan *Watermark*

Gambar 7(a) adalah simulasi proses penyisipan citra *watermark* ke dalam citra original. *PSNR* citra terwatermark (citra original yang telah tersisipi *watermark*) dapat terlihat pada panel *Catatan Proses* pada tampilan simulasi. Sedangkan gambar 7(b) adalah simulasi proses pengekstrakan kembali *watermark* yang telah tersisipi pada citra terwatermark.

Pengujian dengan menggunakan *wavelet Biortogonal 1.3, Daubechies 2, dan Dmeyer*, serta konstanta transformasi (*alpha*) 0.1, 0.5, 1.0, 1.5, dan 2.0 diperoleh hasil *PSNR* dan *watermark* terekstrak sebagaimana pada Tabel 2.

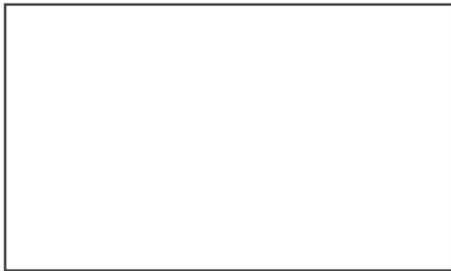
PSNR yang diperoleh dapat digambarkan dalam bentuk grafik sebagai berikut :





(a)

(b)



(c)

Gambar 8. (a) Grafik PSNR Wavelet Biortogonal 1.3, (b) Grafik PSNR Wavelet Daubechies 2, (c) Grafik PSNR Wavelet Dmeyer

Tabel 2. Nilai PSNR dan *Watermark* Terekstrak

ALPHA	WAVELET Biortogonal 1.3		Daubechies 2		DMeyer	
	PSNR*	Watermark Terekstrak	PSNR*	Watermark Terekstrak	PSNR*	Watermark Terekstrak
0.5	41.2053		41.32		41.2922	
1.0	35.2215		35.3441		35.3072	
1.5	31.7239		31.8478		31.8142	
2.0	29.2507		29.3747		29.3391	

PSNR* : PSNR untuk citra terwatermark

Grafik di atas menyatakan hubungan antara konstanta transformasi (*alpha*) dengan *Peak Signal to Noise*

Ratio (PSNR). Angka-angka pada absis X menunjukkan skala alpha, sedangkan pada ordinat Y menunjukkan skala *PSNR*.

Dari grafik terlihat garis bergerak menurun dari kiri ke kanan. Hal ini berarti bahwa semakin besar nilai *alpha* yang digunakan dalam proses penyisipan *watermark*, maka akan berpengaruh pada penurunan nilai *PSNR*. Dengan kata lain semakin besar nilai *alpha* yang digunakan, maka semakin menurun kualitas citra terwatermark.

Sedangkan pada proses ekstraksi, secara *visual* dapat terlihat bahwa nilai *alpha* berpengaruh pada *watermark* terekstrak. Semakin besar nilai *alpha* yang digunakan, maka akan semakin mirip *watermark* yang terekstrak dengan *watermark* asli.

PENUTUP

Dari hasil yang diperoleh pada penelitian ini maka dapat disimpulkan bahwa semakin besar konstanta transformasi (*alpha*) yang digunakan dalam proses transformasinya, berakibat pada semakin menurunnya kualitas citra terwatermark, tetapi secara visual citra *watermark* terekstrak semakin mirip dengan citra *watermark* asli.

DAFTAR PUSTAKA

- 1] Li Tan, Choo., *Still Image Compression using Wavelet Transform*, School of Information Technology and Electrical Engineering, The University of Queensland. Queensland, 2001.
- 2] Littlefield, Bruce, and Duane Hanselman, *MATLAB Bahasa Komputasi Teknis*, Andi and Pearson Education Asia Pte, Ltd. Yogyakarta, 2000.
- 3] Munir, Rinaldi, *Pengolahan Citra Digital dengan Pendekatan Algoritmik*, Informatika Bandung, 2004.
- 4] Supangkat, H., dkk, *Paper : Watermarking sebagai Teknik Penyembunyian Label Hak Cipta pada Data Digital*, Institut Teknologi Bandung, 2000.
- 5] www.cctv-information.co.uk/constant2/sn_ratio.html
- 6] www.mathworks.com

KOMPRESI CITRA DIGITAL DENGAN FRAKTAL SEBAGAI TEKNIK KOMPRESI ALTERNATIF

Aris Sugiharto ¹ dan Agus Harjoko ²
1. Jurusan Matematika FMIPA UNDIP
2. Jurusan Fisika FMIPA UGM

Abstrak: Kompresi citra digital merupakan sebuah metode yang bertujuan untuk meminimalisir adanya redundansi. Beberapa metode kompresi telah dikembangkan antara lain adalah kompresi JPEG, wavelet, serta neural network. Sedangkan pada penelitian ini dikembangkan metode kompresi dengan menggunakan fraktal. Pada metode ini citra digital dibagi menjadi dua blok yaitu blok range yang berukuran 4x4, 8x8 atau 16x16 dan blok domain yang berukuran dua kali dari blok range. Dasar dari kompresi ini adalah mencari kemiripan bagian citra yang ada diantara blok range dengan blok domain dalam bentuk transformasi. Koefisien-koefisien transformasi inilah yang selanjutnya disimpan sebagai ukuran file hasil kompresi. Ukuran file kompresi fraktal pada umumnya sangat kecil akan tetapi kualitas citra yang dihasilkan memiliki nilai PSNR yang rendah dan waktu untuk kompresi yang relatif lama.

Kata Kunci: kompresi, fraktal, blok range, blok domain, PSNR

PENDAHULUAN

Citra digital saat ini banyak digunakan dalam berbagai bidang. Mulai dari keperluan sehari-hari seperti cetak foto, pemetaan hutan, identifikasi forensik, rekam medis dengan menggunakan citra kedokteran (*medical images*) sampai pada citra satelit. Hampir semua citra digital memerlukan media penyimpanan (*storage*) yang cukup besar. Sehingga hal ini menimbulkan masalah jika citra disimpan dalam database yang memiliki keterbatasan media penyimpanan. Masalah lain adalah jika diinginkan untuk mengirimkan citra digital dengan menggunakan jalur komunikasi atau internet. Dengan ukuran yang besar maka citra digital juga memerlukan waktu pengiriman yang lama. Sehingga diupayakan suatu teknik yang dapat mereduksi besarnya ukuran file citra digital dengan kompresi. Hampir semua teknik kompresi memiliki keuntungan seperti di atas yaitu dapat mereduksi ukuran file citra digital, sehingga dapat meminimalisir semua kendala di atas. Akan tetapi dibalik keuntungan ini ada sisi lain yang merugikan yaitu turunnya kualitas citra. Banyak metode kompresi yang sudah dikembangkan antara lain adalah JPEG dan JPEG 2000. Pada penelitian ini akan dikembangkan metode kompresi dengan fraktal sebagai teknik kompresi alternatif.

TINJAUAN PUSTAKA

Kompresi

Saat ini kebanyakan aplikasi menginginkan representasi citra digital dengan menggunakan kebutuhan memori yang seminimal mungkin. Kompresi citra (*images compression*) mempunyai tujuan meminimalkan kebutuhan memori untuk merepresentasikan sebuah citra digital. Prinsip umum yang digunakan pada proses kompresi citra digital adalah mengurangi duplikasi data di dalam citra sehingga memori yang dibutuhkan untuk merepresentasikan citra menjadi lebih sedikit dari pada citra digital aslinya.

Terdapat dua proses utama dalam permasalahan kompresi citra digital.

a. Kompresi citra (*images compression*)

Pada proses ini citra digital dalam representasinya yang asli (belum dikompres) dikodekan dengan representasi yang meminimumkan kebutuhan memori. Citra dengan format bitmap pada umumnya tidak dalam bentuk kompresan. Citra yang sudah dikompres disimpan ke dalam arsip dengan menggunakan format tertentu.

b. Dekompresi citra (*images decompression*)

Pada proses dekompresi, citra yang sudah dikompresi harus dapat dikembalikan lagi menjadi representasi citra seperti citra aslinya. Proses ini diperlukan jika citra ingin ditampilkan ke layar atau disimpan ke dalam arsip dengan format yang tidak terkompres.

Beberapa Metode Kompresi

Kompresi JPEG

Teknik kompresi JPEG (Joint Photographic Expert Group) menggunakan transformasi DCT untuk merubah nilai – nilai pixel dalam domain spasial menjadi koefisien – koefisien dalam domain frekuensi. Dalam kompresi JPEG, mula – mula citra yang akan dikompresi dibagi menjadi blok-blok tertentu berukuran 8x8 yang selanjutnya ditransformasi dengan DCT untuk mendapatkan koefisien – koefisien dalam domain frekuensi [11]. Langkah selanjutnya adalah melakukan proses kuantisasi dan pengkodean dengan metode zig-zag menggunakan RLE, Huffman. Setelah itu dilakukan konversi vector kuantisasi menjadi bit-bit untuk mengkontruksi file JPEG.

Kompresi Wavelet

Selain JPEG teknik lain yang berkembang adalah kompresi wavelet dengan menggunakan transformasi wavelet DWT [1]. Transformasi DWT digunakan untuk mentransformasi citra menjadi beberapa koefisien wavelet. Selanjutnya koefisien wavelet ini ditanam kembali dengan menggunakan metode EZW [9].

Kompresi dengan Neural Network

Teknik kompresi lain yang sedang diteliti adalah teknik kompresi menggunakan neural network dimana pada teknik ini mula-mula citra ditransformasi dengan DWT atau WP (*Wavelet Packet*) sehingga menghasilkan koefisien-koefisien wavelet, selanjutnya koefisien ini direpresentasikan melalui neural network sehingga menghasilkan koefisien baru dengan space minimal melalui pelatihan – pelatihan neuron – neuron [5].

PEMBAHASAN

Menyimpan citra digital dengan menggunakan kumpulan pixel akan membutuhkan media penyimpanan (storage) yang sangat besar, namun apabila citra tersebut yang disimpan adalah transformasi affine-nya, maka kebutuhan tersebut akan dapat dikurangi secara signifikan. Dari latar belakang inilah muncul gagasan untuk mengkodekan citra dengan nisbah kompresi yang tinggi.

Kesulitan yang paling besar dalam upaya melakukan kompresi citra dengan menggunakan fraktal adalah menemukan bagian citra yang mirip dengan keseluruhan citra. Untuk itu diperlukan sebuah teknik khusus yang dapat digunakan untuk menemukan bagian-bagian yang mirip antara bagian citra dengan bagian yang lainnya.

PIFS (Partitioned Iterated Function System)

Citra secara alamiah umumnya hampir tidak pernah memiliki sifat self-similarity secara keseluruhan. Karena itu, citra alamiah secara umum juga tidak memiliki transformasi affine terhadap dirinya sendiri. Tetapi citra

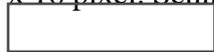
pada umumnya memiliki self-similarity lokal yaitu memiliki bagian – bagian citra yang mirip dengan bagian citra lainnya (Gambar 1).

Gambar 1. Self-similarity lokal
(Sumber : [3])

Secara umum kompresi menggunakan fraktal dilakukan dengan membagi citra asli menjadi beberapa blok yang tidak saling beririsan (*non overlapping*) yang dinamakan dengan blok range yang berukuran 4x4 atau 8x8 dan biasanya untuk mempermudah diambil blok yang berbentuk persegi. Selanjutnya dibuat juga beberapa blok domain yang ukurannya diambil 2 kali blok range [7]. Untuk blok domain yang diambil dapat beririsan ataupun tidak beririsan. Keuntungannya adalah jika digunakan yang tidak beririsan maka jumlah blok domainnya menjadi lebih sedikit dan waktu pencocokannya menjadi lebih cepat. Namun tentu saja hasilnya tidak sebaik jika digunakan blok domain yang beririsan. Karena dengan menggunakan blok domain yang beririsan jumlah blok domainnya semakin banyak sehingga kemungkinan self-similarity lokalnya juga tinggi namun waktu pencocokannya menjadi lebih lama. Selanjutnya untuk setiap blok range dicari bagian citra yang memiliki kemiripan (self-similarity) tinggi dengan blok domain. Tingkat kemiripan ini diukur dengan menggunakan RMS [4].



Untuk blok range dengan ukuran 8 x 8, maka akan diambil blok domain dengan ukuran 16 x 16 pixel. Sehingga jika terdapat sebuah citra dengan ukuran 256 x 256 pixel, maka dapat dibagi



menjadi = 1024 buah blok yang tidak saling beririsan. Sedangkan untuk blok domain yang dapat dibentuk adalah $(256 - 16 + 1)^2 = 58.081$ blok yang saling beririsan.

Sebelum setiap blok yang ada dalam blok range dicocokkan, maka setiap blok dalam blok domain harus terlebih dahulu diskalakan menjadi $\frac{1}{2}$ bagian. Penskalaan ini dimaksudkan untuk menjaga agar jarak antara blok domain dan blok range menjadi lebih mudah dihitung. Penskalaan dapat dilakukan dengan cara menjadikan 2x2 buah pixel menjadi satu buah pixel. Nilai satu buah pixel tersebut adalah rata-rata dari nilai keempat pixel. Jika ditemukan blok range dan blok domain yang memiliki kemiripan tinggi maka dilakukan transformasi affine w_i untuk memetakan blok domain ke blok range. Transformasi affine yang digunakan adalah [4] :



Dengan pemetaan w_i di atas, maka intensitas tiap pixel juga diskalakan dan digeser yaitu :

$$z' = s_i z + o_i$$

Parameter s_i menyatakan faktor kontras pixel, jika s_i bernilai 0 maka pixel menjadi gelap dan jika s_i sama dengan 1 maka kontrasnya tidak mengalami perubahan, sedangkan jika bernilai antara 0 sampai 1 maka kontrasnya menjadi berkurang dan jika lebih besar dari 1 maka kontrasnya akan bertambah. Parameter o_i menyatakan offset kecerahan (*brightness*) pixel. Nilai o_i positif akan mencerahkan gambar dan jika nilai o_i negatif maka akan menjadi gelap. Kedua parameter s_i dan o_i dapat memetakan secara akurat blok domain yang berskala abu-abu ke blok range yang berskala

abu-abu juga.

Untuk menjamin efek kontraktif dalam arah spasial, maka blok domain harus berukuran lebih besar dari pada blok range. Untuk alasan sederhana maka biasanya blok domain diambil 2 x blok range. Jadi jika terdapat sebuah citra dengan blok range yang berukuran $n \times n$ maka dapat diambil ukuran blok domainnya $2n \times 2n$ pixel. Perbandingan inil membuat persamaan transformasi affine menjadi lebih sederhana yaitu :



Parameter e_i dan f_i menyatakan pergeseran sudut kiri blok domain ke sudut kiri blok range yang bersesuaian. Sedangkan s_i dan o_i dihitung dengan menggunakan rumus regresi berikut :



DRMS =

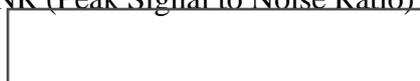
Selanjutnya transformasi affine w_i diuji terhadap blok domain D_i untuk menghasilkan blok uji $T_i = w_i(D_i)$. Jarak antara T dan R_i dihitung dengan DRMS. Transformasi affine yang terbaik adalah transformasi w yang meminimumkan jarak antara R_i dan T .

Runtutan pencarian dilanjutkan untuk blok-blok range berikutnya sampai semua blok range sudah dipasangkan dengan blok domain dengan transformasi affine-nya. Hasil dari proses kompresi adalah sejumlah IFS lokal yang dinamakan dengan PIFS. Seluruh parameter PIFS dipak dan disimpan ke dalam berkas eksternal. Parameter PIFS yang perlu disimpan hanyalah parameter e_i , f_i , s_i , o_i dan jenis operasi simetri terhadap setiap blok range. Dalam kenyataannya, parameter e_i dan f_i diganti dengan posisi blok domain yang dipetakan ke blok range. Sedangkan parameter a_i , b_i , c_i dan d_i tidak perlu disimpan karena nilainya sudah tetap yaitu $\frac{1}{2}$ untuk a_i dan d_i serta 0 untuk b_i dan c_i .

Untuk merekonstruksi citra setelah dikompres, maka dilakukan proses iterasi PIFS dari citra awal sebarang. Karena IFS adalah lokal kontraktif baik kontraktif dalam domain intensitas maupun domain spasial maka iterasinya akan konvergen ke citra tetap PIFS. Kontraktif intensitas penting untuk menjamin konvergensi ke citra semula (original). Sedangkan kontraktif spasial penting dalam menjamin rincian pada citra untuk setiap skala. Jika PIFS yang ditemukan selama proses kompresi baik, yaitu gabungan dari transformasi seluruh blok domain yang dekat dengan citra semula, maka titik tetap PIFS juga dekat dengan citra semula.

Selama proses pemulihan setiap IFS lokal mentransformasikan sekumpulan blok domain menjadi sekumpulan blok range. Karena blok range tidak saling beririsan dan mencakup seluruh pixel dalam citra, maka gabungan seluruh blok range akan menghasilkan citra titik tetap yang menyerupai citra semula.

Untuk mengetahui adanya perubahan kualitas citra sebagai akibat proses kompresi fraktal, digunakan ukuran PSNR (Peak Signal to Noise Ratio) [2] :





Metode Penelitian

Metode yang digunakan dalam kompresi fraktal adalah sebagai berikut :

a. Tahapan proses kompresi

Gambar 3. Tahapan Proses Kompresi

b. Tahapan proses De-kompresi

Gambar 4. Tahapan Proses De-Kompresi

Adapun algoritma untuk kompresi citra digital dengan fraktal adalah :

- a. Baca citra asli
- b. Menentukan ukuran matriks citra asli
- c. Menentukan ukuran blok range

- d. Menentukan ukuran blok domain
- e. Blok domain diskalakan ukurannya menjadi $\frac{1}{2}$ kali ukuran semula.
- f. Untuk setiap blok range
 - f.1 Dicari kemiripan antara blok range ke i dengan semua blok domain dengan menggunakan RMS.
 - f.2 Hitung transformasi affine untuk antara blok range ke i dengan blok domain yang terpilih.
 - f.3 Simpan koefisien transformasi affine ke i.
- g. Simpan semua parameter dalam PIFS

Sedangkan untuk algoritma de kompresi adalah :

- a. Baca citra asli (pembanding)
- b. Menentukan jumlah iterasi
- c. Mengambil semua parameter PIFS
- d. Untuk setiap blok range ke i
 - Transformasi affine dengan menggunakan parameter dari blok domain ke blok range ke i.
- e. Blok range hasil transformasi adalah citra hasil kompresi.

Simulasi dan Hasil

Simulasi kompresi citra digital dengan menggunakan fraktal dilakukan dengan menggunakan perangkat lunak Matlab 6.5. Adapun citra yang digunakan adalah mandril128.bmp yang berukuran 128 x 128.

(a)

(b)

Gambar 5. Simulasi kompresi dengan menggunakan citra Mandril (a). blok range 8x8 dan jumlah iterasi 4, (b) blok range 4x4 dan jumlah iterasi 8.

Setelah dilakukan beberapa simulasi diperoleh hasil sebagai berikut :

Tabel 1. Kompresi Fraktal Citra Mandril 128x128 dengan blok range 4x4

NO	Iterasi	Time (sec)	PSNR(dB)	Size (byte)
1.	2	515.922	25.072	5120
2.	4	522.594	25.759	5120
3.	8	513.172	25.803	5120
4.	16	509.688	25.800	5120
5.	32	524.281	25.789	5120

Tabel 2. Kompresi Fraktal Citra Mandril 128x128 dengan blok range 8x8

NO	Iterasi	Time (sec)	PSNR(dB)	Size (byte)
1.	2	37.813	22.962	1280
2.	4	35.297	23.742	1280
3.	8	36.625	23.701	1280
4.	16	37.468	23.701	1280
5.	32	38.047	23.701	1280

Tabel 3. Kompresi Fraktal Citra Mandril 128x128 dengan blok range 16x16

NO	Iterasi	Time (sec)	PSNR(dB)	Size (byte)
1.	2	2.672	20.912	320
2.	4	2.843	21.115	320
3.	8	2.859	21.11	320
4.	16	3.047	21.110	320
5.	32	3.407	21.110	320

KESIMPULAN

Dari uraian sebelumnya dapat disimpulkan bahwa teknik kompresi dengan fraktal menghasilkan ukuran file yang relatif kecil akan tetapi terjadi penurunan kualitas citra yang signifikan. Hal ini dapat dilihat dari nilai PSNR yang berkisar 20.912 hingga 25.803. Selain itu dengan kompresi fraktal waktu yang dibutuhkan untuk proses kompresi relatif lama berkisar 2.672 sampai 524.281 detik.

UCAPAN TERIMA KASIH

Terima kasih kepada Program Hibah Kompetisi A2 jurusan matematika FMIPA UNDIP tahun 2006 yang telah mendanai program magang penelitian ini.

DAFTAR PUSTAKA

- [1] Antonini, M., et. al., Images Coding Using Wavelet Transform, *IEEE Trans. On Image Processing*, Vol. 1, pp. 205 – 220, 1992.
- [2] Chen C., *On The Selection of Image Compression Algorithms*, NSC Grant Departemen of Computer Science National Tsing Hua University, Taiwan, 1998.
- [3] Fisher, Y., *Fractal Image Compression : Theory and Application* , Spinger – Verlag, 1994.
- [4] Galabov, M., Fractal Image Compression, *International Confrence On Computer Science System and Technologies – CompSysTech*, University Of Veliko Turnovo, 2003.
- [5] Gillespie, W., *Still Image Compression Using Neural Networks*, Utah State University, Logan Utah, 2003.
- [6] Gonzalez , R.C. and Woods, R.E., *Digital Image Processing*, Addison Wesley Publishing, 1993.
- [7] Khalid Kamali, *Fractal Video Compression*, Bachelor of Engineering (Computer Science) Faculty of Engineering and Surveying University of Southern Queensland, 2005.
- [8] Munir R, *Pengolahan Citra Digital dengan Pendekatan Algoritmik*, Penerbit Informatika, Bandung, 2004.
- [9] Shapiro, J.M., Embedded Image Coding Using Zerotree of Wavelet Coefficient, *IEEE Trans. On Image Processing*, Vol. 41, pp. 3445 – 3462, 1993.
- [10] Sugiharto Aris, *Pemrograman GUI dengan MATLAB*, Penerbit Andi Yogyakarta, 2006.
- [11] Wallace C.J., The JPEG Still Picture Compression Standart, *Communication ACM*, Vol. 34, pp. 31 – 44, 2005.

E-LEARNING SEBAGAI MODEL PROSES PEMBELAJARAN BERBASIS TEKNOLOGI INFORMASI

Djalal Er Riyanto, Eko Adi Sarwoko, Kushartantya

Program Studi Ilmu Komputer Jurusan Matematika FMIPA UNDIP
Jl. Prof. Soedarto, Kampus UNDIP Tembalang, Semarang

Abstrak: Dalam perkembangan teknologi global, e-learning merupakan bentuk teknologi informasi yang dapat diterapkan dalam bidang pendidikan. e-Learning merupakan suatu transformasi proses belajar mengajar yang ada di sekolah ke dalam bentuk digital. Dengan menggunakan teknologi e-Learning proses belajar mengajar dapat berlangsung secara live ataupun virtual. Diharapkan e-Learning ini dapat lebih merangsang siswa untuk mengeksplorasi pengetahuan, dibandingkan dengan hanya dibombardir dengan ilmu pengetahuan. Tulisan ini membahas pentingnya proses belajar mengajar dengan memanfaatkan teknologi e-Learning.

Kata Kunci : e-Learning, teknologi informasi

PENDAHULUAN

Istilah e-Learning dapat didefinisikan sebagai sebuah bentuk teknologi informasi yang diterapkan di bidang pendidikan dalam bentuk sekolah maya. Definisi e-Learning sendiri sebenarnya sangatlah luas, bahkan sebuah portal yang menyediakan informasi tentang suatu topik dapat tercakup dalam lingkup e-Learning ini. Namun, istilah e-Learning lebih tepat ditujukan sebagai usaha untuk membuat sebuah transformasi proses belajar mengajar yang ada di sekolah ke dalam bentuk digital yang dapat dijumpai oleh teknologi Internet.

Dalam teknologi e-Learning, semua proses belajar mengajar yang biasa didapatkan di dalam sebuah kelas dilakukan secara live namun virtual, artinya pada saat yang sama seorang guru mengajar di depan sebuah computer yang ada di suatu tempat, sedangkan siswa mengikuti pelajaran tersebut dari computer lain di tempat yang berbeda. Dalam hal ini, secara langsung guru dan siswa tidak saling berkomunikasi namun secara tidak langsung mereka saling berinteraksi pada waktu yang sama.

Harus diakui bahwa fokus e-Learning lebih pada efisiensi proses belajar mengajar, cara pengajaran maupun materi ajar masih dapat mengacu pada kurikulum nasional. Siswa lebih pasif dan berposisi sebagai konsumen pengetahuan. Guru sebagai otoritas yang pengetahuannya didukung oleh sistem perpustakaan dan metoda penyampaian. Konsep Knowledge Management, belajar mandiri yang berbasis pada kreativitas siswa dan mendorong siswa melakukan analisa hingga sintesa pengetahuan menghasilkan tulisan, informasi, dan pengetahuan sendiri menjadi focus yang lebih mengarah ke masa depan. Siswa tidak lagi dibombardir dengan doktrin ilmu pengetahuan tetapi lebih dirangsang untuk mengeksplorasi pengetahuan dan menjadi bagian integral proses pemurnian pengetahuan itu sendiri. Pada tahap ini, metoda belajar mengajar dan kurikulum nasional perlu mengalami perombakan lumayan banyak. Terlepas dari adanya konsep yang sangat revolusioner ini dan perlu diujicobakan piranti yang nantinya digunakan di e-Learning maupun Knowledge Management tidak berbeda jauh.

e-Learning merupakan sebuah bentuk teknologi informasi yang diterapkan di bidang pendidikan dalam bentuk sekolah maya, yang memuat sebuah transformasi proses belajar mengajar yang ada di sekolah ke dalam bentuk digital dengan memanfaatkan teknologi Internet.

Dalam teknologi e-Learning ini, diharapkan terjadi efisiensi proses belajar mengajar, mendorong siswa belajar mandiri yang berbasis pada kreativitas siswa dan mendorong siswa melakukan analisa dan sintesa pengetahuan. Siswa tidak lagi dibombardir dengan doktrin ilmu pengetahuan tetapi lebih dirangsang untuk mengeksplorasi pengetahuan dan menjadi bagian integral proses pemurnian pengetahuan itu sendiri. Penelitian ini bertujuan untuk mengembangkan sistem e-Learning yang dapat meningkatkan efisiensi dan produktivitas proses belajar mengajar.

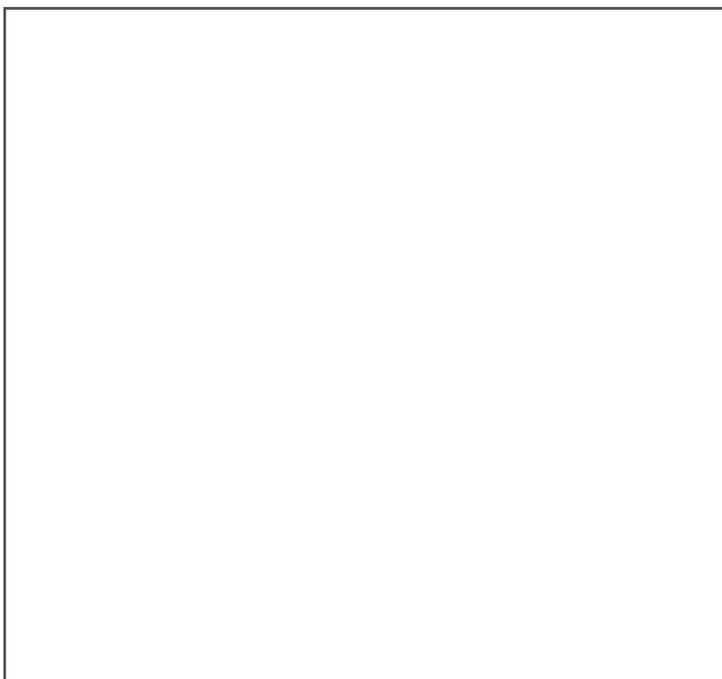
E-LEARNING DAN LINGKUNGANNYA

Sebagai suatu sistem real, sistem e-Learning tidak dapat lepas dari interaksi dengan lingkungan di sekelilingnya. Lingkungan e-Learning perlu diidentifikasi, dan dianalisis untuk mengetahui seberapa besar tingkat kontribusinya terhadap e-Learning.

Secara garis besar lingkungan e-Learning tersebut ialah terdiri atas: pertama **Sumber belajar**, yang memberikan isi dari e-Learning, menciptakan suasana yang kondusif dalam pemanfaatan e-Learning. Kedua **Fasilitative Tools**, yang terkait dengan teknologi yang dimanfaatkan untuk meningkatkan kemudahan pemanfaatan, dan kualitas layanan e-Learning dan terakhir **Isu dan Kebijakan**, baik yang berasal dari pemerintah, tuntutan publik, maupun tantangan yang sifatnya global.

Pemakai Sistem

Pemakai sistem (*user*) ialah mereka yang berinteraksi langsung, dan mendapatkan layanan dari sistem e-Learning. Layanan bervariasi tergantung pada jenis pemakai. e-Learning di dalam proyek ini terkait dengan empat pemakai, yaitu: siswa (individu), kelompok siswa (kelas), guru atau instruktur, dan manajer sekolah (lihat Gambar 1).



Gambar 1. Pemakai e-Learning

Layanan yang diberikan atau Interaksi yang dilakukan masing-masing pemakai diberikan pada Tabel 1. Penggunaan teknologi, khususnya teknologi jaringan komputer memberikan kualitas interaksi yang diselenggarakan

diantara pemakai menjadi lebih baik dibandingkan dengan sistem tradisional (menggunakan mekanisme *face-to-face*).

Sumber Belajar

Sumber belajar dapat dirumuskan sebagai segala sesuatu yang dapat memberikan kemudahan kepada peserta didik dalam memperoleh sejumlah informasi, pengetahuan, pengalaman dan ketrampilan, dalam proses belajar mengajar [3]. Atau sumber belajar dapat juga dikatakan sebagai daya yang bisa dimanfaatkan guna kepentingan proses belajar mengajar, baik secara langsung maupun tidak langsung, sebagian atau secara keseluruhan (Sujana, 2001).

Tabel 1. Layanan e-Learning

No.	Pemakai	Layanan atau Interaksi
1.	Siswa	Akses sumber belajar. Komunikasi (siswa-siswa, siswa-guru). Diskusi elektronik. Meminta layanan konsultasi.
2.	Kelompok siswa	Akses sumber belajar. Diskusi elektronik. Meminta layanan konsultasi.
3.	Guru atau Instruktur	Mensupply bahan ajar. Menerima layanan konsultasi. Merespon permintaan konsultasi.
4.	Manajer Sekolah	Menerima umpan balik.

Dalam pengertian yang sempit sumber belajar berupa bahan ajar (buku atau bahan cetak yang lain), sedangkan pengertian yang lebih luas sumber belajar berupa pengalaman. Di dalam e-Learning sumber belajar tersebut disimpan dalam media elektronik (CD-ROM, disket, atau hardisk).

Menurut Mulyasa [3], sumber belajar dibedakan menjadi lima kelompok, yaitu : **Manusia**, yaitu orang yang menyampaikan pesan secara langsung, seperti: guru, konselor, administrator, yang diniati secara khusus dan disengaja untuk kepentingan belajar. Kedua, **Bahan**, yaitu sesuatu yang mengandung pesan pembelajaran. Ketiga, **Lingkungan**, yaitu ruang dan tempat dimana sumber-sumber dapat berinteraksi dengan para peserta didik. Keempat, **Alat dan peralatan**, yaitu sumber belajar untuk produksi dan/atau memainkan sumber-sumber lain, dan terakhir **Aktivitas**, yaitu sumber belajar yang biasanya merupakan kombinasi antara suatu teknik dengan sumber lain untuk memudahkan belajar.

Sedangkan kegunaan sumber belajar secara umum ialah: merupakan pembuka jalan dan pengembangan wawasan terhadap proses belajar mengajar yang akan ditempuh, kedua : merupakan pemandu secara teknis dan operasional untuk penelusuran menuju penguasaan keilmuan secara tuntas, ketiga memberikan berbagai macam ilustrasi dan contoh-contoh yang berkaitan dengan aspek-aspek bidang keilmuan yang dipelajari, keempat : memberikan petunjuk dan gambaran kaitan bidang keilmuan yang sedang dipelajari dengan berbagai bidang keilmuan lainnya, dan menginformasikan sejumlah penemuan baru yang pernah diperoleh orang lain yang berhubungan dengan bidang keilmuan tertentu, serta menunjukkan berbagai permasalahan dalam suatu bidang keilmuan, yang menuntut adanya kemampuan pemecahan dari orang-orang yang mengabdikan diri dalam bidang tersebut.

Isu dan Kebijakan

Pada awal abad ini, dunia pendidikan di Indonesia menghadapi tiga tantangan besar, yaitu: (a) sebagai akibat dari krisis ekonomi, dunia pendidikan dituntut untuk mempertahankan hasil-hasil pembangunan pendidikan yang telah dicapai, (b) untuk mengantisipasi era globalisasi, dunia pendidikan dituntut untuk mempersiapkan sumber daya manusia yang kompeten agar mampu bersaing dalam pasar kerja global, dan (c) sejalan dengan diberlakukannya otonomi daerah, sistem pendidikan nasional dituntut untuk melakukan perubahan dan penyesuaian sehingga dapat mewujudkan proses pendidikan yang lebih demokratis, memperhatikan keberagaman kebutuhan / keadaan daerah dan peserta didik, serta mendorong peningkatan partisipasi masyarakat.

Terdapat dua hal penting yang sangat mempengaruhi e-Learning, yaitu: kebijakan pemerintah dibidang penyelenggaraan pendidikan, dan kurikulum. Pertama, Kebijakan penyelenggaraan pendidikan nasional. Salah satu dampak dari kebijakan-kebijakan tersebut ialah pemberian kepercayaan yang lebih luas kepada sekolah untuk mengoptimalkan sumber daya yang

tersedia bagi tujuan pendidikan yang diharapkan. Kedua, Kurikulum Berbasis Kompetensi, dalam rangka mempersiapkan lulusan pendidikan memasuki era globalisasi yang penuh tantangan dan ketidakpastian, diperlukan pendidikan yang dirancang berdasarkan kebutuhan nyata di lapangan. Untuk itu pemerintah memprogramkan Kurikulum Berbasis Kompetensi (KBK) sebagai acuan dan pedoman bagi pelaksanaan pendidikan untuk mengembangkan berbagai ranah pendidikan (pengetahuan, ketrampilan, dan sikap) dalam seluruh jenjang dan jalur pendidikan. Kompetensi diartikan sebagai pengetahuan, ketrampilan dan kemampuan yang dikuasai oleh seseorang yang telah menjadi bagian dari dirinya, sehingga ia dapat melakukan perilaku-perilaku kognitif, afektif, dan psikomotorik dengan baik. Kompetensi mencakup tugas, ketrampilan, sikap, dan apresiasi yang harus dimiliki oleh peserta didik untuk dapat melaksanakan tugas-tugas pembelajaran sesuai dengan jenis pekerjaan tertentu. Kompetensi merupakan perpaduan dari pengetahuan, ketrampilan, nilai dan sikap yang direfleksikan dalam kebiasaan berpikir dan bertindak.

Berdasarkan pengertian kompetensi, KBK dapat diartikan sebagai suatu konsep kurikulum yang menekankan pada pengembangan kemampuan melakukan (kompetensi) tugas-tugas dengan standar performansi tertentu, sehingga hasilnya dapat dirasakan oleh peserta didik, berupa penguasaan terhadap seperangkat kompetensi tertentu. Enam karakteristik kurikulum berbasis kompetensi, yaitu: (a) sistem belajar dengan modul, (b) menggunakan keseluruhan sumber belajar, (c) pengalaman lapangan, (d) strategi individual personal, (e) kemudahan belajar, dan (f) belajar tuntas. KBK diharapkan mampu memecahkan berbagai persoalan bangsa, khususnya dalam bidang pendidikan, dengan mempersiapkan peserta didik, melalui perencanaan, pelaksanaan, dan evaluasi terhadap sistem pendidikan secara efektif, efisien, dan berhasil guna. Tujuan utama KBK adalah memandirikan atau memberdayakan sekolah dalam mengembangkan kompetensi yang akan disampaikan kepada peserta didik, sesuai dengan kondisi lingkungan. Implementasi KBK menuntut kerjasama yang optimal diantara para pengajar. KBK adalah suatu konsep kurikulum yang menekankan pada pengembangan kemampuan melakukan (kompetensi) tugas-tugas dengan standar performansi tertentu.

Dengan demikian, implementasi kurikulum dapat menumbuhkan tanggung jawab, dan partisipasi peserta didik untuk belajar menilai dan mempengaruhi kebijakan umum, serta memberanikan diri berperan serta dalam berbagai kegiatan, baik di sekolah maupun di masyarakat. KBK memberikan keleluasaan kepada sekolah untuk menyusun dan mengembangkan silabus mata pelajaran sesuai dengan potensi sekolah, kebutuhan dan kemampuan peserta didik serta kebutuhan masyarakat disekitar sekolah. Silabus KBK dikembangkan oleh tiap sekolah, sehingga dimungkinkan beragamnya kurikulum antarsekolah atau antar wilayah tanpa mengurangi kompetensi yang telah ditetapkan dan berlaku secara nasional (standar akademis).

Manajemen Pengetahuan (Knowledge Management)

Manajemen Pengetahuan (MP) ialah serangkaian proses yang mengatur kreasi, penyebaran, dan pemanfaatan dari pengetahuan. MP menaruh perhatian dengan seluruh proses dari penemuan dan kreasi pengetahuan, dan memfasilitasi transformasi pengetahuan implisit ke dalam pengetahuan eksplisit yang dapat diakses yang dapat membawa ke lahirnya situasi pemecahan masalah yang relevan. Skenario dari MP dimulai dari kreasi pengetahuan untuk ditrapkan, digunakan, dan pengaruhnya terhadap kehidupan. Kreasi tersebut kemudian ditangkap untuk dapat disimpan, disusun dan ditransformasikan ke dalam bentuk yang dapat disebar, dan dipakai bersama. Pengetahuan yang sudah disusun dengan baik dan dapat diakses dengan mudah akan dapat digunakan untuk meningkatkan kualitas layanan, peningkatan daya saing produk, melahirkan teknologi, dan pengetahuan baru. Hubungan antar komponen di dalam MP diberikan pada Gambar 2.

Tujuan akhir dari MP ialah dalam menyusun, berbagi, dan menempatkan pengetahuan untuk menciptakan suatu nilai pokok didalam pengetahuan; menyesuaikan dengan cepat pada perubahan lingkungan untuk mendorong efisiensi dan meningkatkan kemampuan dalam memberikan layanan. Untuk melayani kebutuhan yang terus berkembang, perlu peningkatan keluwesan dan penyesuaian MP, kemampuannya dalam menangkap informasi, menciptakan

pengetahuan, meningkatkan nilai, dan kemampuannya dalam melindungi aset intelektual. Perlu dilakukan kreasi, peremajaan, ketersediaan, kualitas dan penggunaan pengetahuan oleh semua komponen sistem.

SISTEM DAN METODE PENGEMBANGAN E-LEARNING

Sistem e-Learning terdiri atas sekumpulan komponen yang kompleks dan dinamis, mencakup manusia (manajer sekolah, guru, siswa, masyarakat), informasi, proses administrasi, dan lain sebagainya. Salah satu aspek dari manajemen ialah untuk memfasilitasi efektivitas dan efisiensi berfungsinya komponen-komponen sistem.

Sistem e-Learning merupakan salah satu bentuk MP, dan dikembangkan dengan memanfaatkan teknologi, khususnya teknologi komputer dan komunikasi. Teknologi akan mendukung tersedianya layanan pencarian informasi dan pengetahuan secara cepat, akan tetapi teknologi juga dapat digunakan untuk mengirimkan pengetahuan ke sistem untuk dapat diakses bersama. Teknologi memungkinkan diwujudkannya berbagai pengetahuan yang ada agar dapat dimanfaatkan bersama, dan pada gilirannya akan dapat menciptakan pengetahuan yang baru.



Gambar 2 Skenario Manajemen Pengetahuan

Metode Pengembangan e-Learning

Dengan fungsi-fungsi dan harapan yang diberikan kepada e-Learning, maka diperlukan suatu tahapan pengembangan sistem e-Learning, yaitu : pertama : pengetahuan, Pengetahuan menyediakan kemampuan untuk merespon situasi yang baru. Pengetahuan merupakan hasil proses pembelajaran yang harus dikelola dengan efektif, untuk memastikan tujuan dasar untuk keberadaannya tercapai sebesar mungkin. Meskipun mempunyai peran yang penting, pengetahuan masih merupakan aset yang banyak diabaikan. Pengetahuan dapat disajikan secara eksplisit (dalam bentuk basis data informasi, kebijakan-kebijakan, dan prosedur-prosedur), atau dapat pula disajikan secara implisit (pengetahuan yang tidak diucapkan, budaya organisasi, kebiasaan, dan lain-lain). Untuk masing-masing individu pengetahuan dapat melahirkan suatu ide, pendapat, bakat, *relationship*, perspektif, dan konsep. Sedangkan secara bersama-sama pengetahuan dapat ditransformasikan kedalam bentuk teknologi, produk, layanan, kebiasaan, fasilitas, dan sistem. e-Learning memerlukan pengetahuan yang diperoleh oleh seseorang maupun sekelompok orang atau komunitas untuk dikelola, sehingga dapat dimanfaatkan dalam jangkauan yang lebih luas. Kedua, Basis data :

Salah satu bentuk penyajian pengetahuan ialah dengan menggunakan basis data. Basis data pengetahuan merupakan bentuk dasar dari suatu e-Learning. Dengan menggunakan perangkat lunak, pemakai akan dapat melakukan pencarian dengan menggunakan index dan panduan dengan pertanyaan, melalui perintah selangkah-demi-selangkah untuk mengerjakan tugas spesifik. Pencarian biasanya disajikan secara cukup interaktif, yang berarti pemakai dapat suatu kata kunci atau ungkapan untuk pencarian basis data, atau membuat suatu pilihan dari suatu daftar alphabet. Ketiga, Perangkat lunak : Analisis perangkat lunak mencakup tahapan untuk menentukan perangkat lunak mana yang sudah tersedia di pasar, dan perangkat lunak mana yang harus dikembangkan sendiri. Untuk perangkat lunak yang akan dikembangkan sendiri kemudian dianalisis untuk menentukan bahasa pemrograman, maupun tool-tool yang diperlukan. Berdasarkan analisis yang dilakukan dibutuhkan perangkat-perangkat lunak adalah pembuatan halaman web : Dreamweaver MX atau FrontPage XP dari Microsoft, gambar/Animasi : Flash, editor : NotePad dan komunikasi : NetMeeting. Khusus untuk NetMeeting merupakan perangkat lunak yang dapat diperoleh secara gratis. Keempat, Perangkat keras untuk server maupun untuk pemakai : Tujuannya ialah untuk mendapatkan spesifikasi minimum perangkat keras yang paling tepat untuk mendukung pemanfaatan e-Learning yang dihasilkan. Dari sisi server, akan memberikan kemampuan layanan yang paling optimal (kecepatan akses, jaminan keamanan). Dari sisi pemakai akan diperoleh informasi yang akan membantu pemakai dalam merencanakan penyediaan perangkat keras untuk memanfaatkan sistem yang dihasilkan. Secara teoritis e-Learning dapat berbasis CD-ROM, jaringan komputer, Intranet, maupun Internet. e-Learning untuk proyek ini akan dikembangkan dengan berbasis Intranet, sehingga dapat diperoleh model sistem yang dapat diakses oleh banyak pengguna di dalam suatu lingkungan pembelajaran yang secara geografis berada pada suatu lingkup tertentu (gedung, kompleks sekolah). Dengan model yang dihasilkan akan dapat didemonstrasikan kelebihan e-Learning dari sistem pembelajaran konvensional dalam memberikan layanan untuk akses basis data pengetahuan, tutorial, dan umpan balik berupa respon dan evaluasi pembelajaran, serta sebagai media untuk komunikasi antar siswa, ataupun siswa dengan guru. Kelima, Prosedur atau aturan-aturan yang diperlukan : Dengan prosedur dan aturan yang tersedia dan dikomunikasikan secara luas, akan mempermudah pengoperasian dan pemeliharaan sistem. Selain itu juga untuk memperlancar komunikasi atau hubungan antar pengelola dengan pemakai, atau antar pemakai dengan pemakai. Keenam, Sumber daya manusia (kualifikasi dan kompetensi) : Dasar analisis ialah untuk mendapatkan persyaratan-persyaratan pemakai yang paling minimum, sehingga sistem yang dihasilkan akan dapat dimanfaatkan dalam lingkup yang lebih luas sesuai dengan kelompok pemakai. Secara fungsional, sumber daya manusia yang berhubungan dengan e-Learning dikelompokkan berdasarkan kepentingan penggunaan e-Learning dan kewenangan. Pengelompokan tersebut ialah: *Pemakai Biasa*, yaitu pemakai yang memanfaatkan e-Learning untuk keperluan pembelajaran atau layanan pembelajaran. Kelompok pemakai ini memerlukan antar-muka layanan sistem yang paling efisien, paling menarik, dan paling mudah digunakan, dan *Pengelola Sistem*, yaitu orang-orang yang ditunjuk untuk mengelola e-Learning agar selalu *up-to-date* dan sesuai dengan kebutuhan. Pengelola sistem perlu memiliki kewenangan dan kemampuan teknis yang memadai agar dapat melaksanakan tugasnya. Dan terakhir, Sumber daya beaya untuk operasional dan pemeliharaan sistem : Untuk menjamin keberlanjutan sistem perlu dilakukan analisis beaya operasional dan pemeliharaan, untuk kemudian dirumuskan pola-pola untuk mendapatkan sumber beaya. Dengan pemakai dan berbagai kebutuhan yang bersifat personal maupun institusional, e-Learning diharapkan memfasilitasi berbagai kebutuhan pemakai untuk peningkatan kualitas proses belajar mengajar. Untuk keperluan tersebut perlu diciptakan suatu lingkungan pembelajaran yang kondusif, yaitu: tersedianya komunikasi diantara berbagai pemakai dengan mudah; koordinasi kegiatan pemakai; dan kolaborasi diantara grup pemakai pada kreasi, modifikasi dan penyebaran pengetahuan, hasil kecerdasan manusia dan produk.

Untuk mewujudkan lingkungan tersebut e-Learning memerlukan fungsi spesifik yang berhubungan ke komunikasi (e-mail dan forum diskusi); koordinasi; kolaborasi (dapat berbagi-pakai pengetahuan atau hasil kecerdasan manusia, dan ruang kerja); dan kontrol. Sistem e-Learning memerlukan dukungan untuk banyak fungsi informasi, mencakup mendapatkan dan peng-indeks-an; menangkap dan mengarsipkan; mencari dan mengakses; menciptakan dan pencatatan; kombinasi; menyusun dan memodifikasi; dan penelusuran.

KESIMPULAN

e-Learning merupakan sebuah bentuk teknologi informasi yang diterapkan di bidang pendidikan dalam bentuk sekolah maya. e-Learning ini dapat lebih merangsang siswa untuk mengeksplorasi pengetahuan, dibandingkan dengan hanya dibombardir doktrin ilmu pengetahuan.

Di dalam e-Learning sumber belajar tersebut disimpan dalam media elektronik (CD-ROM, disket, atau hardisk), dengan *Fasilitative Tools* yang merupakan aplikasi elektronik yang digunakan di dalam e-Learning sebagai bagian dari transformasi proses belajar mengajar seperti: program chat, e-mail, streaming audio, streaming video,

halaman web, dan sebagainya. e-Learning diharapkan dapat dikembangkan dengan menyediakan bahan ajar berbasis web dalam bentuk CD yang nantinya dapat dipasarkan kepada masyarakat luas.

DAFTAR PUSTAKA

- [1]. Alan Neibauer, *Small Business Solutions for Networking*, Terjemahan oleh BM Adam, PT Elex Media Komputindo, Jakarta, 2001.
- [2]. Arikunto, Suharsimi, *Dasar-dasar Evaluasi Pendidikan*, PT Bumi Aksara, Jakarta, 2002.
- [3]. Mulyasa, *Kurikulum Berbasis Kompetensi, Konsep, Karakteristik, dan Implementasi*, PT Remaja Rodakarya, Bandung., 2003.
- [4]. ---, *Propenas 2000-2004, UU Nomor 25 Tahun 2000 tentang Program Pembangunan Nasional Tahun 2000-2004*, Sinar Grafika, Jakarta, 2003.
- [5]. Spector, J. Michael, and Gerald S. Edmonds, *Knowledge Management in Instructional Design*, Publications of Eric Digest, Syracuse University, September 2002.

PENGGUNAAN MODEL MAXIMAL FREQUENT SEQUENCES dan Model Ruang Vektor DALAM SISTEM TEMU-KEMBALI INFORMASI TEKS BERBAHASA INDONESIA

Dwi Astuti Aprijani
Jurusan Matematika FMIPA Universitas Terbuka
E-mail: dwi_aprijani@yahoo.com

Abstrak: Pertumbuhan jumlah informasi tekstual yang tersedia secara elektronik meningkatkan kebutuhan kinerja temu-kembali (*retrieval*) yang lebih baik. Penggunaan frase telah terbukti dapat meningkatkan kinerja temu-kembali dibandingkan dengan model lain yang mengabaikan aspek sekuensial dari kemunculan kata dalam dokumen. Penelitian ini bertujuan untuk melihat penggunaan frase sebagai indeks dalam dua jenis koleksi dokumen berbahasa Indonesia. Hal lain yang diteliti adalah apakah penggunaan frase --- yang memperhatikan aspek sekuensial dari kemunculan kata dalam dokumen --- mempunyai pengaruh dalam kinerja temu-kembali dibandingkan pengindeksan berdasarkan kata yang diberi bobot. Frase yang digunakan dalam penelitian ini adalah *Maximal Frequent Sequences* (MFS). Kekuatan MFS ini adalah pertama, dapat membentuk indeks yang sangat solid dari kemunculan bersama sepasang kata dengan menoleransi adanya kata pemisah di antara pasangan kata tersebut. Kedua, MFS diekstrak jika dan hanya jika MFS tersebut muncul lebih sering dari *frequency threshold* yang diberikan, hal ini untuk menghindari penyimpanan frase yang tidak signifikan. Ketiga, jumlah istilah yang digunakan sebagai indeks sedikit. Hasil uji coba yang dilakukan terhadap 1162 dokumen ilmiah dan 3000 dokumen berita menunjukkan bahwa pada koleksi dokumen ilmiah, MFS dengan nilai *threshold* yang lebih kecil relatif mampu memperbaiki nilai presisi, baik presisi rata-rata sesudah memproses 20 dokumen maupun presisi rata-rata sesudah memproses dokumen dengan jumlah yang sesuai dengan jumlah dokumen relevan. Sedangkan pada koleksi dokumen berita, MFS dengan nilai *threshold* yang kecil relatif memperbaiki nilai presisi rata-rata sesudah memproses dokumen dengan jumlah yang sesuai dengan jumlah dokumen relevan, namun *threshold* yang lebih besar justru memperbaiki presisi rata-rata sesudah memproses 10 dokumen.

Kata Kunci: frase, maximal frequent sequences, pengindeksan, sekuen kata non-contiguous, temu-kembali informasi

PENDAHULUAN

Sistem temu-kembali berkinerja tinggi menjadi kebutuhan yang tak terelakkan, sejalan dengan ledakan

jumlah dokumen elektronik yang terjadi pada dasawarsa terakhir. Diperkirakan fenomena ini masih akan terus berlanjut di masa-masa mendatang. Banyak usaha yang telah dilakukan untuk menciptakan sistem temu kembali berkinerja tinggi tersebut, salah satunya adalah memperbaiki sistem pengindeksan dokumen [6].

Penelitian pada penggunaan frase untuk pengindeksan dokumen dalam temu-kembali informasi telah dimulai sejak 25 tahun yang lalu. Hasil awal amat menjanjikan, namun ternyata penggunaan frase menunjukkan hasil yang mengecewakan [8,11]. Menurut Mitra, terdapat dua alasan mengapa penggunaan frase kurang maksimal. Pertama, frase mewakili hanya satu aspek dari kemungkinan kueri yang multi aspek karena pada umumnya kata-kata di dalam frase tidak diperhatikan urutan kemunculannya dalam suatu dokumen. Kedua, frase tidak memungkinkan adanya kata-kata lain yang muncul diantaranya. Misalkan, frase “XML document retrieval” dan “XML retrieval”, dua frase ini akan dianggap berbeda, padahal dalam *natural language* mempunyai kandungan informasi yang sama [8].

Penelitian ini bertujuan untuk mengetahui sejauh mana peningkatan yang dihasilkan oleh sistem temu-kembali apabila pengindeksannya menggunakan frase yang memperhatikan urutan kemunculan kata dan menoleransi adanya kata pemisah diantaranya. Frase yang dapat mengakomodasi masalah ini adalah *Maximal Frequent Sequences* (MFS). Menurut penelitian yang dilakukan Doucet terhadap koleksi dokumen INEX (*Initiative for the Evaluation of XML Retrieval*), MFS menawarkan model *natural language* yang lebih realistis. Keuntungan lain adalah jumlah frase yang menjadi indeks lebih sedikit [7]. Kesimpulan ini perlu diuji untuk koleksi dokumen yang lain. Di sini digunakan koleksi dokumen berbahasa Indonesia.

MODEL RUANG VEKTOR

Prinsip dasar model ruang vektor adalah dokumen direpresentasikan dengan menggunakan vektor kata. Oleh sebab itu, setiap dokumen dinyatakan oleh , yaitu vektor berdimensi yang diisi dengan bobot, yang mengatakan seberapa penting kata tersebut dalam dokumen, dimana W adalah himpunan kata unik yang diperoleh dalam pengindeksan. Bobot kata unik ini dihitung menggunakan rumus *tf* *term-weight component* [9], yaitu:

$$w_{ij} = \frac{f_{ij}}{\sum_{j \in W} f_{ij}} \cdot \frac{1}{\sqrt{\sum_{i \in D} f_{ij}^2}}$$

dimana

$$w_{ij}$$

adalah frekuensi kata w

$$N$$

adalah jumlah total dokumen dalam koleksi

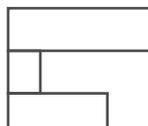
$$n_i$$

adalah jumlah dokumen yang mengandung w

W adalah himpunan akar kata (*stem*) berlainan yang diperoleh

MAXIMAL FREQUENT SEQUENCES

MFS adalah sekuen kata yang frekuen dalam koleksi dokumen dan tidak merupakan bagian dari sekuen lain yang lebih panjang yang juga frekuen. Suatu sekuen dikatakan frekuen apabila dia muncul minimal dalam θ (dokumen, dimana θ adalah *frequency threshold* yang diberikan. Misalkan S adalah himpunan dokumen, dan setiap dokumen mengandung sekuen-sekuen kata.



Definisi 1. Suatu sekuen adalah subsekuen dari sekuen q bila semua item , muncul dalam q dan item-item tersebut muncul dalam urutan yang sama seperti dalam p . Jika sekuen p adalah subsekuen dari sekuen q , dapat juga dikatakan

bahwa p muncul dalam q .

Definisi 2. Sekuen p disebut frekuensi dalam S jika p adalah subsekuen dari paling tidak τ dokumen dalam S , dimana τ adalah *frequency threshold* yang diberikan.



Definisi 3. Suatu sekuen p adalah *maximal frequent subsequence* dalam S jika tidak ada sekuen lain dalam



S sedemikian sehingga p adalah subsekuen dari q dan frekuensi dalam S .

Algoritma untuk menemukan MFS menggunakan metode *bottom-up* dan *greedy* [1,2], namun untuk penerapannya dalam dokumen teks berbahasa Indonesia, algoritma ini dimodifikasi sehingga pendekatannya *non-greedy* [3].

Kekuatan utama MFS dapat mengekstrak semua sekuen kata maksimal yang sering muncul dalam suatu koleksi dokumen secara efisien. Selain itu, kata-kata pemisah di antara sepasang kata diperbolehkan, sehingga kata-kata tidak harus muncul berurutan di dalam kalimat. Hal ini meningkatkan kualitas frase karena *processing* memperhatikan variasi *natural language*. Dengan adanya kata-kata pemisah, memungkinkan mencakup semua variasi penulisan yang ada karena untuk menulis suatu kalimat dengan maksud yang sama, setiap orang mempunyai gaya penulisan sendiri. Kelebihan lain dari teknik ini adalah kemampuan mengekstrak MFS dengan panjang berapapun. Sehingga dapat diperoleh deskripsi dokumen yang sangat solid.

EVALUASI DOKUMAN

Untuk mencari relevansi dokumen dengan kueri digunakan dua nilai *Retrieval Status Value* (RSV), yaitu: RSV fitur kata dan RSV model MFS, kemudian dua nilai RSV ini diintegrasikan menjadi satu skor tunggal untuk tiap dokumen yang sesuai dengan kueri yang dimaksud [7].

RSV Fitur Kata

Model ruang vektor memberikan kemudahan dalam menghitung similaritas antara dokumen dan kueri. Ada beberapa teknik untuk membandingkan dua vektor dalam temu-kembali informasi, diantaranya adalah *Euclidean distance*, *similaritas Jaccard* dan *similaritas cosine*, namun di sini digunakan *similaritas cosine* karena efisien dalam komputasi.



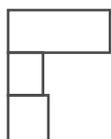
Dengan normalisasi vektor, yang dilakukan dalam fase pengindeksan, *cosine* disederhanakan menjadi perkalian



vektor .

RSV MFS

Tahap pertama adalah membuat indeks MFS untuk koleksi dokumen [3]. Kemudian melakukan prosedur untuk menyesuaikan (*matching*) MFS (frase yang merepresentasikan dokumen) dengan *keyphrase* (frase yang ditemukan dalam kueri). Metodenya adalah penguraian (dekomposisi) *keyphrase* menjadi pasangan. Jadi, suatu kueri yang terdiri dari beberapa kata, akan dipecah-pecah menjadi kombinasi dua kata. Tiap pasangan diberi skor yang menyatakan kuantitas relevansinya. Kuantitas relevansi dari suatu pasangan kata mengatakan seberapa besar dia membuat suatu dokumen relevan dengan memasukkan pasangan kata ini. Nilai ini bergantung pada kekhususan pasangan (*inverted document frequency*) dan koefisien *adjacency*. Kuantitas relevansi suatu pasangan yang dibentuk oleh dua kata yang tidak berdampingan nilainya lebih kecil.



Definisi 4. Misalkan D adalah koleksi dari N dokumen dan k adalah *keyphrase* berukuran m . Misalkan d_i dan k_j adalah dua

[Redacted]

kata berurutan dari sekuen s , dan n adalah jumlah dokumen dimana ditemukan. Kuantitas relevansi dari pasangan

[Redacted]

didefinisikan sebagai berikut

[Redacted]

[Redacted]

dimana menyatakan kekhususan dari s dalam koleksi D :

[Redacted]

[Redacted]

[Redacted]

adalah skor modifier untuk pasangan kata yang dibentuk dari kata-kata yang tidak berdampingan (*non adjacent*), dan

[Redacted]

[Redacted]

[Redacted]

menunjukkan jumlah kata yang muncul di antara s dan t , mengatakan bahwa s dan t berdampingan (*adjacent*).

[Redacted]

Menghitung RSV

Dalam prakteknya, beberapa kueri tidak mengandung suatu frase, dan beberapa dokumen tidak mengandung suatu MFS. Namun, pasti ada jawaban yang benar untuk kueri-kueri seperti ini, dan dokumen-dokumen tersebut tentunya relevan dengan kueri-kueri itu. Semua dokumen yang mengandung frase *matching* yang sama memperoleh RSV MFS yang sama. Oleh karena itu, sangat penting mendapatkan suatu cara untuk membedakannya. Ukuran similaritas cosine berdasarkan kata sangat cocok untuk masalah tersebut.

Pendekatan lain adalah menguraikan kembali pasangan-pasangan menjadi kata-kata tunggal dan membentuk vektor-vektor dokumen yang sesuai. Namun, hal ini tidak memuaskan karena kata-kata yang munculnya paling jarang hilang akibat algoritma ekstraksi MFS. Oleh sebab itu harus dihitung RSV lain menggunakan model ruang vektor fitur kata.

Untuk menghitung dua RSV menjadi satu skor tunggal, pertama-tama kedua RSV harus

dapat dibandingkan dengan memetakannya ke *common interval* [7].

EVALUASI KINERJA SISTEM TEMU-KEMBALI INFORMASI

Kriteria yang digunakan untuk menilai kualitas STI dalam penelitian ini adalah terpenuhinya kebutuhan pengguna. Hal ini dapat dilihat dari *recall* dan *precision*. *Recall* merupakan ukuran banyaknya dokumen relevan yang terambil dari kumpulan dokumen pada saat kueri diterapkan. Sedangkan *precision* merupakan ukuran ketepatan atau korelevansian hasil kueri [10].

Namun rumusan Salton yang digunakan untuk penilaian relevansi dalam penelitian ini sedikit dimodifikasi karena jumlah dokumen yang relevan untuk setiap kueri dalam koleksi dokumen yang digunakan tidak sama. Misalkan ada satu kueri yang memiliki dua dokumen relevan, dan dokumen yang menjadi keluaran sistem dibatasi hanya sampai peringkat 20, maka *precision*-nya adalah 0,1 padahal dokumen di peringkat 1 dan 2 adalah dokumen yang relevan, dengan demikian seharusnya *precision*-nya adalah 1. Untuk itu digunakan



R-Precision dengan rumus sebagai berikut: $R\text{-Precision} = \frac{a}{b}$; dimana a = jumlah dokumen yang relevan dan b = jumlah dokumen terproses yang sesuai dengan jumlah dokumen relevan dalam koleksi dokumen.

METODOLOGI DAN IMPLEMENTASI

Koleksi Data

Koleksi data yang dipergunakan dalam penelitian ini terdiri dari dua set corpus, yaitu corpus ilmiah dan corpus berita. Corpus ilmiah adalah koleksi dokumen hasil penelitian yang dilakukan dalam lingkungan institusi Badan Tenaga Atom Nasional (BATAN), terdiri dari 1162 buah dokumen, yang merupakan hasil penelitian dalam rentang waktu antara tahun 1985 sampai dengan tahun 1994 [4,6]. Sedangkan corpus berita merupakan kumpulan artikel yang dimuat antara Januari dan Juni 2002 dalam surat kabar harian Indonesia, Kompas on line, terdiri dari 3000 buah dokumen [5].

Implementasi Sistem

Seluruh aktivitas yang dilakukan dalam penelitian ini dilaksanakan pada komputer PC yang menjalankan sistem operasi Linux (distribusi Fedora Core 4) dengan prosesor Pentium IV 2.4 GHz dan memori sebesar 512 Mbytes.

Bahasa pemrograman yang dipergunakan secara ekstensif untuk seluruh implementasi dalam penelitian ini adalah Python. Python adalah bahasa berorientasi obyek (*Object Oriented Programming Language*) yang modular dan merupakan salah satu bahasa pemrograman tingkat tinggi. Dipilihnya bahasa pemrograman ini karena Python memiliki sintaks yang sederhana dan mudah dibaca, serta dapat berjalan di beberapa sistem yang berlainan, misalnya Windows maupun UNIX/Linux. Versi Python yang dipergunakan dalam penelitian ini adalah versi 2.4.1, yang dikeluarkan pada bulan September 2005. Bahasa ini dapat diambil dari situs utama <http://www.python.org>.

Prapengolahan Dokumen

Tujuan dari kegiatan ini adalah untuk menyiapkan dan merapikan data koleksi dokumen sehingga koleksi tersebut dapat dipergunakan secara mudah untuk proses-proses selanjutnya dalam penelitian ini. Aktivitas dalam kegiatan ini secara garis besar dapat dibagi menjadi 3 bagian, yaitu pengindeksan kembali, pemfilteran kata-kata tak bermakna (*stopword*) dan pengkodean dokumen (*encoding*).

Setiap dokumen dari koleksi data diindeks kembali agar setiap dokumen memiliki identitas unik berupa suatu bilangan integer. Untuk setiap dokumen, proses parsing dilakukan untuk mengambil judul dokumen, nama pengarang beserta isi dokumen. Proses filterisasi dilakukan untuk menghilangkan punctuasi dan kata-kata yang hanya terdiri dari bilangan saja atau yang hanya memiliki satu huruf saja, dan menghilangkan kata-kata tak bermakna.

Dokumen yang telah terindeks dan terfilter kemudian dikode sehingga tiap kata dalam dokumen diwakili oleh bilangan integer. Dari sini diterapkan algoritma pencarian MFS dari Ahonen-Myka [1] yang telah dimodifikasi kembali [3].

UJICOBA DAN ANALISA

Uji coba untuk pengindeksan berdasarkan MFS ini dilakukan dengan mengaplikasikan kueri S1-S10 pada koleksi dokumen ilmiah dan kueri number 1-number 20 pada koleksi dokumen berita.

Pengindeksan Berdasarkan MFS

MFS yang digunakan sebagai indeks dalam sistem temu-kembali di sini adalah MFS yang diperoleh dari nilai *threshold* 4 dan 7 untuk dokumen ilmiah dan nilai *threshold* 7 dan 10 untuk dokumen berita [3]. Jumlah indeks berdasarkan MFS ini jauh lebih sedikit, hanya 31% bila dibandingkan dengan jumlah indeks berdasarkan frekuensi kata untuk koleksi dokumen ilmiah. Perbandingan selengkapnya jumlah indeks berdasarkan frekuensi kata dan MFS dapat dilihat pada Tabel 1. Jumlah indeks yang sedikit memberikan keuntungan karena mempercepat proses pencarian dokumen yang sesuai dengan kueri.

Tabel 1. Perbandingan jumlah indeks pada koleksi dokumen

Jenis koleksi dokumen	Jumlah istilah indeks berdasarkan frekuensi kata	Jumlah istilah indeks berdasarkan MFS
Ilmiah	9636	3022
Berita	34159	10328

Hasil uji coba terhadap koleksi dokumen ilmiah menunjukkan bahwa presisi rata-rata sesudah memproses 20 dokumen (*Precision@20*) paling baik adalah 0,0850. Sedangkan presisi rata-rata sesudah memproses dokumen dengan jumlah yang sesuai dengan dokumen yang relevan (*R-Precision*) paling baik adalah 0,2135.

Tabel 2. Perbandingan kinerja temu-kembali koleksi dokumen ilmiah berdasarkan MFS dengan dua nilai threshold

Metode pengindeksan (jarak kata, nilai threshold)	<i>Precision@20</i>	<i>R-Precision</i>
MFS (g = 2, (= 4)	0,0800	0,2135
MFS (g = 2, (= 7)	0,0500	0,0587
MFS (g = 3, (= 4)	0,0850	0,2135
MFS (g = 3, (= 7)	0,0500	0,0630

Hasil uji coba terhadap koleksi dokumen berita menunjukkan bahwa presisi rata-rata sesudah memproses 10 dokumen (*Precision@10*) paling baik adalah 0,3150. Sedangkan presisi rata-rata sesudah memproses dokumen dengan jumlah yang sesuai dengan dokumen yang relevan (*R-Precision*) paling baik adalah 0,2448.

Tabel 3. Perbandingan kinerja temu-kembali koleksi dokumen berita berdasarkan MFS dengan dua nilai threshold

Metode pengindeksan (jarak kata, nilai threshold)	Precision@10	R-Precision
MFS (g = 2, (= 7)	0.3150	0.2448
MFS (g = 2, (= 10)	0.2950	0.2101
MFS (g = 3, (= 7)	0.2800	0.2302
MFS (g = 3, (= 10)	0.3000	0.2410

Pengindeksan Berdasarkan Term Frequency Cosine

Pengindeksan berdasarkan *term frequency cosine* (tfc) ini adalah pengindeksan berdasarkan kata-kata yang diberi bobot menggunakan rumus Salton [9] dan sistem temu-kembali menggunakan model ruang vektor dihitung berdasarkan similaritas cosine.

Hasil uji coba terhadap koleksi dokumen ilmiah memberikan *Precision@20* sebesar 0,1350 dan *R-Precision* sebesar 0.2417. Sedangkan untuk koleksi dokumen berita, *Precision@10* sebesar 0,3600 dan *R-Precision* sebesar 0.4315.

Pengindeksan Gabungan

Untuk lebih meningkatkan kinerja sistem temu-kembali, maka dicoba melakukan pengindeksan gabungan, antara tfc dan MFS. Hasilnya dapat dilihat pada Tabel 4.

Tabel 4. Perbandingan kinerja temu-kembali koleksi dokumen ilmiah dengan metode pengindeksan gabungan

Metode pengindeksan (jarak kata, nilai threshold)	Precision@20	R-Precision
tfc+MFS (g = 2, (= 4)	0,1400	0.2417
tfc+MFS (g = 2, (= 7)	0,1300	0.2417
tfc+MFS (g = 3, (= 4)	0,1450	0.2417
tfc+MFS (g = 3, (= 7)	0,1550	0,3073

Khusus untuk koleksi dokumen ilmiah, karena tidak didapatnya data yang meyakinkan mengenai relevansi dokumen terhadap kueri (ada 2 versi relevansi, dan keduanya banyak perbedaan), maka nilai-nilai presisi yang ditunjukkan pada tabel di atas tidak dapat dibandingkan secara langsung dengan presisi milik Indra Budi [6]. Tidak adanya data yang pasti mengenai data relevansi yang dipergunakan oleh Indra Budi membuat adanya kemungkinan perbedaan relevansi dokumen dengan kueri antara penelitian Indra Budi dan penelitian ini, sehingga nilai presisinya pun akan berbeda pula. Nilai *R-Precision* tertinggi dihasilkan oleh metode tfc+MFS dengan g = 3 dan (= 7). Pada koleksi dokumen ilmiah ini, nilai *threshold* yang lebih kecil relatif mampu memperbaiki nilai presisi, baik *Precision@20* maupun *R-Precision*. Koleksi dokumen ilmiah adalah koleksi dokumen yang relatif lebih homogen daripada koleksi dokumen berita. Selain bentuknya yang mengikuti gaya bahasa yang lebih baku, dokumen-dokumen dalam koleksi dokumen ilmiah ini juga ditulis oleh pengarang-pengarang tertentu saja, sehingga gaya penulisan pada koleksi ini mirip antara satu dokumen dengan dokumen lainnya. Oleh karena itu, penggunaan nilai *threshold* yang lebih kecil dapat menangkap esensi (atau MFS) dari tiap-tiap dokumen tersebut dengan lebih baik.

Pada koleksi dokumen ilmiah, MFS tidak dapat meningkatkan *Precision@20* maupun *R-Precision* secara signifikan. MFS dapat meningkatkan presisi apabila kueri berupa *natural language*, sedangkan kueri yang dipergunakan pada koleksi ilmiah ini sudah berupa kata-kata atau istilah khusus. Tidak adanya efek yang signifikan menunjukkan bahwa banyak dari kueri-kueri yang berupa kata-kata khusus tersebut yang tidak terwakili pada MFS dari koleksi dokumen ilmiah.

Hasil uji coba terhadap koleksi dokumen berita dengan menggunakan pengindeksan gabungan tfc+MFS dapat dilihat pada Tabel 5. Kolom *Precision@10* menunjukkan presisi rata-rata sesudah memproses 10 dokumen, sedangkan kolom *R-Precision* menunjukkan presisi rata-rata sesudah memproses dokumen dengan jumlah yang sesuai dengan dokumen yang relevan.

Tabel 5. Perbandingan kinerja temu-kembali koleksi dokumen berita dengan metode pengindeksan gabungan

Metode pengindeksan (jarak kata, nilai threshold)	Precision@10	R-Precision
tfc+MFS (g = 2, (= 7)	0.3700	0.4507

tfc+MFS (g = 2, (= 10)	0.3750	0.4510	
tfc+MFS (g = 3, (= 7)	0.3550	0.4428	
tfc+MFS (g = 3, (= 10)	0.3550	0.4399	

Untuk metode MFS, nilai *Precision@10* dan *R-Precision* tidak terlalu bisa dipergunakan untuk menunjukkan keakurasian sistem. Hal ini dikarenakan oleh banyaknya dokumen keluaran untuk suatu kueri dari kedua sistem tersebut yang memiliki nilai skor yang sama (baik yang relevan maupun yang tidak), sehingga dokumen-dokumen dengan nilai skor yang sama hanya diperingkat berdasarkan urutan kemunculannya dalam koleksi dokumen.

Untuk metode tfc dan gabungan (tfc+MFS), *Precision@10* dan *R-Precision* merupakan nilai keakurasian sistem, karena tidak ada satu dokumen pun yang memiliki nilai skor yang sama untuk suatu kueri.

Nilai *Precision@10* dan *R-Precision* tertinggi didapat dari metode tfc+MFS dengan $g = 2$, $(= 10$. Nilai *threshold* yang kecil membuat lebih banyak variasi dan jumlah MFS yang dihasilkan. Nilai *threshold* yang kecil memperbesar sebaran dari MFS, sehingga jumlah MFS yang mewakili tiap dokumen akan lebih banyak. Dari situ dapat ditarik kesimpulan bahwa penggunaan *threshold* yang kecil dapat menambah akurasi dan presisi dari suatu sistem temu-kembali informasi. Namun apabila dilihat dari nilai-nilai pada tabel di atas, nilai *threshold* yang kecil tidaklah selalu memperbaiki performa dari sistem temu-kembali informasi. Penggunaan *threshold* yang lebih kecil memang relatif memperbaiki nilai *R-Precision*, namun *threshold* yang lebih besar justru memperbaiki *Precision@10*.

KESIMPULAN DAN SARAN

Berdasarkan hasil uji coba penerapan MFS untuk pengindeksan dalam STI teks berbahasa Indonesia terhadap koleksi dokumen ilmiah dan koleksi dokumen berita, diperoleh kesimpulan sebagai berikut:

- **Jumlah istilah pada pengindeksan berdasarkan MFS jauh lebih sedikit dibandingkan jumlah istilah pada pengindeksan berdasarkan kata yang diberi bobot tfc, yaitu 31% untuk koleksi dokumen ilmiah, 30% untuk koleksi dokumen berita.**
- **Pada sistem temu-kembali untuk koleksi dokumen ilmiah, MFS dengan nilai *threshold* yang lebih kecil relatif mampu memperbaiki nilai presisi, baik *Precision@20* maupun *R-Precision*.** Namun MFS tidak dapat meningkatkan *Precision@20* maupun *R-Precision* secara signifikan. Hal ini menunjukkan bahwa banyak dari kueri-kueri yang berupa kata-kata khusus yang tidak terwakili pada MFS.
- **Pada koleksi dokumen berita, nilai *threshold* yang kecil dapat menambah akurasi dan presisi dari suatu STI.** Namun, nilai *threshold* yang kecil tidak selalu memperbaiki kinerja dari STI. Penggunaan *threshold* yang lebih kecil memang relatif memperbaiki nilai *R-Precision*, namun *threshold* yang lebih besar justru memperbaiki *Precision@10*.
- **Semakin homogen suatu koleksi dokumen, efek dari *threshold* yang kecil akan semakin baik.**
- Metode pengindeksan berdasarkan kata yang diberi bobot tfc (*term frequency cosine*) yang dipergunakan sebagai komplemen pada sistem MFS dapat meningkatkan nilai *precision* pada temu-kembali dokumen-dokumen yang relevan.
- Efektivitas dari penggunaan MFS pada sistem temu-kembali sangat ditentukan oleh bentuk dari kueri yang dipergunakan, apakah berbentuk istilah saja atau *natural language*.

Berdasarkan hasil penelitian, ada beberapa saran untuk pengembangan dan penyempurnaan sistem lebih lanjut, yaitu:

- Untuk mendapatkan nilai presisi yang optimum, penggunaan MFS untuk STI tidak berdiri sendiri karena istilah-istilah dan pasangan-pasangan kata yang berada di bawah *threshold* secara otomatis tidak akan terwakili dalam indeks MFS. Oleh karena itu, STI dengan MFS ini harus dikomplemen dengan suatu sistem yang dapat mengindeks seluruh istilah. Sistem tfc yang dipergunakan sebagai komplemen pada sistem MFS dalam penelitian ini terbukti dapat meningkatkan nilai presisi pada temu-kembali dokumen-dokumen yang relevan.
- **Mencari metode yang lebih baik lagi dalam memanfaatkan**

MFS untuk STI. Metode tersebut harus lebih *sophisticated* daripada sistem kuantitas relevansi yang dipakai dalam sistem ini.

- **Memperbaiki data dan nilai relevansi dari dokumen ilmiah BATAN sehingga dapat dipergunakan secara meyakinkan untuk keperluan penelitian STI.**

ACKNOWLEDGEMENT

The author would like to thank Jelita Asian for providing her Indonesian corpus used in this paper. The author also would like to thank Indra Budi for the BATAN corpus. Lastly, the author also thanks Hidayat Trimarsanto for his help in understanding Python.

DAFTAR PUSTAKA

- [1] Ahonen-Mika, Helena, Finding All Maximal Frequent Sequences in Text. In *Proceedings of the 16th International Conference on Machine Learning ICML-99 Workshop on Machine Learning in Text Data Analysis*, Ljubljana, Slovenia, pp. 11-17, J. Stefan Institute, eds. D. Mladenic and M. Grobelnik, 1999.
- [2] Ahonen, Helena, *Knowledge Discovery in Documents by Extracting Frequent Word Sequences*. Department of Computer Science at the University of Helsinki, Finland, 2000.
- [3] Aprijani, D. A., *Pengindeksan Berdasarkan Maximal Frequent Sequences dalam Sistem Temu Kembali Teks Berbahasa Indonesia*, Tesis S2, Depok: Fasilkom UI, 2006.
- [4] Aribawono, A.B., *Pendekatan Multi-dimensi Dokumen dalam Sistem Temu-kembali Informasi Menggunakan Model Spreading Activation*, Tesis S2, Depok: Fasilkom UI, 2001.
- [5] Asian, et. al., A Testbed for Indonesian Text Retrieval, In *Proceedings of the 9th Australasian Document Computing Symposium*, Melbourne, Australia, 2004.
- [6] Budi, I., *Pengindeksan dan Kemiripan Dokumen dalam Sistem temu-kembali Informasi*, Tesis, Depok: Fakultas Pasca Sarjana Universitas Indonesia, 2003.
- [7] Doucet, Antoine and Helena Ahonen-Myka, *Non-Contiguous Word Sequences for Information Retrieval*. Departement Of Computer Science, University of Helsinki, Finland, 2004.
- [8] Mitra, M., et. Al., An Analysis of Statistical and Syntactic Phrases, In *Proceedings of RIAO97, Computer-Assisted Information Searching on Internet*, pp. 200-214, 1997.
- [9] Salton, Gerard and Christopher Buckley, Term-weighting Approaches in Automatic Text Retrieval, *Information Processing & Management*, Vol. 24, No. 5, pp. 513-523, 1988.
- [10] Salton, Gerard, *Automatic Text Processing: The Transformation, Analysis, and Retrieval of Information by Computer*. New York: Addison-Wesley Publishing Company, 1989.
- [11] Turpin, A. and A. Moffat, Statistical Phrase for Vector-Space Information Retrieval. In *Proceedings of the 22nd ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 309-310, 1999.

MEKANISME SISTEM IDENTIFIKASI BIOMETRIK

Eko Adi Sarwoko

Program Studi Ilmu Komputer Jurusan Matematika FMIPA UNDIP
Jl. Prof. Soedarto, Kampus UNDIP Tembalang, Semarang

Abstrak: Dalam perkembangan teknologi global, identifikasi merupakan bagian penting dalam terjaminnya kerahasiaan personal/data. Tahap kerahasiaan ini akan terjamin dengan memanfaatkan identifikasi personal dalam pengaksesan dan pelayanannya. Tulisan ini membahas proses identifikasi personal dan perkembangan teknologi identifikasi personal, semula hanya sidik jari sebagai alat identifikasinya. Sekarang telah berkembang beberapa teknik untuk biometrik, seperti mata, tangan, tanda-tangan, ucapan dan mimik wajah manusia serta gabungan diantara aspek biometrik tersebut.

Kata Kunci: identifikasi personal, biometrik

PENDAHULUAN

Dalam teknologi terkini ditawarkan adanya beberapa kemudahan, seperti akses, pelayanan, dan sistem informasi. Kemudahan tersebut dapat dirasakan seperti pada mekanisme pengambilan uang melalui ATM (Anjungan Tunai Mandiri), mekanisme memperoleh sistem informasi (internet). Mekanisme tersebut diperlukan adanya jaminan kerahasiaan sehingga tidak dapat ditiru oleh user yang bukan berhak. Salah satu alat untuk menjamin bahwa yang berhak mendapatkan layanan itu harus memberikan data identifikasi. Sistem identifikasi tersebut bersifat otomatis dengan memberikan inputan identifikasi personal.

Beberapa teknik identifikasi personal yang telah berkembang adalah identifikasi sidik jari. Identifikasi sidik jari telah diakui sebagai alat identifikasi personal yang canggih. Hal ini telah diteliti bahwa sidik jari masing-masing manusia sangat unik, berbeda satu dengan yang lainnya. Mekanisme identifikasi ini telah digunakan dalam bidang forensik, untuk keperluan investigasi kriminal, identifikasi anggota badan, dsb.

Selain itu, pemanfaatan identifikasi personal ini dapat dirasakan pada bidang sipil, seperti ID kewarganegaraan, SIM (Surat Ijin Mengemudi). Ataupun pada bidang komersial seperti ATM, kartu kredit, telepon selular, kontrol akses, dsb. Pendek kata pemanfaatan identifikasi personal telah merambah pada berbagai layanan kebutuhan publik terkini.

Tulisan ini membahas arah perkembangan identifikasi personal terkini, dengan tidak hanya memanfaatkan sidik jari seperti sebelumnya, tetapi telah berkembang kepada identifikasi biometrik, seperti mata, tangan, tanda-tangan, ucapan dan mimik wajah manusia. Selain itu bagaimana tahapan mekanisme sistem identifikasi personal tersebut dilakukan?.

TAHAPAN PROSES IDENTIFIKASI PERSONAL

Mekanisme identifikasi personal pada dasarnya dapat dibedakan atas dua tahapan, yaitu tahapan proses verifikasi dan tahapan proses pengenalan. Pada tahapan verifikasi merupakan cara mengetahui autentikasi (keaslian) personal. Pada tahap proses verifikasi ini dilakukan dengan memasukkan token, seperti kartu paspor, SIM, kartu kredit, kunci pintu, dsb. Sedang pada tahapan proses pengenalan merupakan cara untuk mengidentifikasi. Pada tahap ini memberikan pengenalan knowledge/pemahaman identifikasi "siapa saya", dengan menggunakan password

(kata kunci), PIN, dsb. [1]

Sebagai contoh dalam mekanisme pengambilan uang ATM, akan melalui dua tahapan dasar tersebut. Pertama tahapan proses verifikasi identifikasi personal dapat dilakukan dengan adanya keberadaan kartu ATM yang harus dimasukkan, kemudian dilanjutkan dengan tahapan proses pengenalan identifikasi personal yaitu dengan nomor PIN (Personal Identification Number) yang harus dimasukkan.

Tahapan proses ini mempunyai beberapa kelemahan mendasar, seperti pada tahap pemasukan verifikasi identifikasi personal dapat dimungkinkan hilang, lupa, dan salah menempatkannya. Sedangkan pada tahap pengenalan identifikasi personal dapat terjadi lupa terhadap yang valid (pernah diubah), dapat ditebak oleh impostor (orang yang tidak berhak). Bahkan terdapat data bahwa 25% terjadi pencurian uang di ATM [1].

Seperti telah dijelaskan di depan bahwa identifikasi personal telah berkembang dengan pesat. Identifikasi personal yang semula hanya mengandalkan kemampuan sidik jari manusia, telah bergeser dengan memanfaatkan identifikasi biometrik seperti mata, tangan, tanda-tangan, ucapan dan mimik wajah manusia[1]. Bahkan telah diteliti bahwa gabungan identifikasi wajah dikombinasikan dengan ucapan manusia [3], serta penggabungan identifikasi wajah dengan tanda-tangan sebagai alat identifikasi personal [2].

Pemanfaatan identifikasi biometrik ini diharapkan dapat merupakan solusi atas kelemahan proses identifikasi personal tersebut, sehingga dapat memberikan pelayanan dan kemudahan seperti yang diharapkan oleh para pengguna teknologi.

MEKANISME SISTEM IDENTIFIKASI BIOMETRIK

Dalam tahap identifikasi biometrik dapat mengidentifikasi individu-individu berdasarkan perbedaan lingkup karakteristik behaviour/psikologi (biometric identifier). Hal ini dimungkinkan bahwa karakteristik psikologi/behaviour setiap manusia berbeda-beda. Selain itu identifier biometrik dianggap lebih reliable dibandingkan berdasarkan pemasukan token dan pengenalan knowledge.

Mekanisme sistem biometrik dapat digambarkan dengan beberapa fase, pertama fase penggolongan (enrollment). Pada fase ini masukan akan di pindai (scan) oleh sensor biometrik, yang merupakan representasi karakteristik digital. Selanjutnya fase pencocokan, dalam fase ini inputan database akan dicocokkan dengan identifikasi data. Dapat dimungkinkan adanya reduksi, sehingga dihasilkan representasi digital. Hasil ini akan diproses dengan ekstraktor ciri untuk menghasilkan suatu representasi yang ekspresif dalam bentuk template. Bergantung aplikasinya template dapat disimpan dalam database di sistem biometrik atau dapat direkam pada kartu magnetik (atau smartcard). Sedang pada fase pengenalan, karakteristik individu dibaca oleh pembaca biometrik (reader). Selanjutnya dikonversi dengan format digital, untuk diproses sebagai ekstraktor ciri (template). Hasil template ini selanjutnya dicocokkan dengan

identifikasi individu [1]. Lihat gambar 1.

Sistem biometrik belumlah sempurna, karena suatu saat masih dapat melakukan kesalahan dengan menerima impostor sebagai individu yang juga valid (terjadi kesalahan pencocokan), sebaliknya terjadi penolakan terhadap individu yang valid (terjadi kesalahan ketidakcocokan). Untuk menjamin terhindarnya kesalahan seperti itu, sesuai referensi [3] memadukan ciri biometrik wajah dengan ucapan, serta dari referensi [2] memadukan biometrik wajah dengan ciri tanda-tangan. Selain itu dalam penerapannya ukuran database template sangatlah besar, bahkan dalam database perbankan pusat pernah terjadi bottleneck saat proses identifikasi [1].

Sistem biometrik yang ideal, diharapkan mempunyai karakteristik sebagai berikut : pertama aspek universal, artinya ciri ini dapat berlaku secara umum (bahwa setiap manusia mempunyai karakteristik), kedua aspek unik (tidak ada dua manusia yang mempunyai karakteristik yang sama), ketiga haruslah bersifat permanen (karakteristik personal yang tidak berubah-ubah) dan terakhir dapat dihimpun (collectable), karakteristik ini mudah disajikan oleh sensor dan mudah dikuantisasikan dan dikuantifikasikan.

Gambar 1 Mekanisme Sistem Biometrik

Selain beberapa hal yang harus diperhatikan dari mekanisme ini adalah masalah kinerja (dalam mekanisme ini akurasi sistem, kecepatan, kehandalan) perlu mempertimbangkan adanya resource, faktor-faktor operasional dan pengembangan, dsb. Hal ini akan berpotensi sebagai kendala teknis. Selain itu adalah akseptabilitas (daya terima pengguna) akan mendorong keyakinan user terhadap akurasi dan kecepatan. Serta aspek circumvention yaitu aspek kemudahan sistem yang tidak bergantung alat, mekanisme operasional, dsb.

KESIMPULAN

Mekanisme identifikasi biometrik ini mampu mengidentifikasi individu-individu berdasarkan perbedaan lingkup karakteristik behaviour/psikologi (biometric identifier) yang diyakini bersifat unik. Selain itu identifier biometrik dianggap lebih reliable dibandingkan berdasarkan pemasukan token dan pengenalan knowledge. Identifikasi bioritmik sekarang terus berkembang mencakup kemampuan sensor dan infrastruktur identifikasi serta merupakan alternatif pilihan dalam pengembangan identifikasi personal.

DAFTAR PUSTAKA

- [1]. Anil K Jain, et. al., *Biometric Identification, Communications of The ACM*, Vol 43, No 2, pp. 91-99, 2000.
- [2]. Lin Hong, et. al., Integrating Faces and Fingerprints for Personal Identification, *IEEE Transactions On Pattern Analysis and Machine Intelligence*, Vol 20, No 12, pp. 1295-1307, 1998.
- [3]. Souheil Ben-Yacoub, et. al., *Fusion of face and Speech Data for Person Identity Verification, IEEE Transactions On Neural Network*, Vol 10, No 5, pp. 1065-1074, 1999.

UJI KETAHANAN WATERMARKING PADA DOMAIN WAVELET dan Frekuensi Terhadap Serangan *MOTION BLUR* DAN KOMPRESI JPEG

Hendry H. Wahid, Aris Sugiharto, Beta Noranita
Jurusan Matematika FMIPA UNDIP
Jl. Prof. Soedarto, Kampus UNDIP Tembalang, Semarang

Abstrak: *Watermarking* merupakan salah satu metode yang dikembangkan guna melindungi citra digital dari upaya penggandaan secara ilegal. Beberapa metode watermarking telah dilakukan pada domain wavelet dan frekuensi. Pada penelitian ini dikaji tentang ketahanan watermarking pada kedua domain dari serangan (*attack*) berupa pengkaburan (*blurring*) dan kompresi JPEG. Sebagai tolak ukur ketahanannya digunakan NC (*Normalized Cross correlation*) dengan membandingkan watermark terekstrak dengan watermark asli.

Kata Kunci : watermarking, kompresi JPEG, *motion blur*, NC.

PENDAHULUAN

Data digital pada era sekarang ini mengalami perkembangan yang sangat pesat. Banyak data digital dipertukarkan untuk berbagai kepentingan. Mulai dari kepentingan yang positif hingga kepentingan yang negatif. Salah satunya adalah adanya penggandaan secara ilegal seperti pembajakan CD, konflik kepemilikan citra digital dan sebagainya. Hal inilah yang mengakibatkan data digital menjadi salah satu pusat perhatian karena kemudahan data ini untuk digandakan tanpa takut atau khawatir akan adanya penurunan kualitas [2]. Sehingga banyak upaya atau metode yang dikembangkan guna melindungi data digital dari upaya penggandaan di atas.

Watermarking hadir sebagai salah satu alternatif untuk melindungi data digital dari usaha orang-orang yang tidak bertanggung jawab yang dengan seenaknya tanpa memperhatikan hak atas kekayaan intelektual (HAKI) dengan melakukan upaya manipulasi dan penggandaan tanpa ijin. Akan tetapi watermarking dalam kenyataannya juga sangat sering mengalami berbagai serangan. Serangan ini dapat berupa serangan alamiah seperti pemrosesan citra pada umumnya seperti proses rotasi, translasi, maupun cropping serta serangan yang benar-benar hanya bertujuan untuk menghilangkan watermark. Diantaranya serangan berupa pemrosesan citra digital seperti pengkaburan citra (*blurring*) dengan *Motion blur* dan kompresi JPEG. Pada penelitian ini diinginkan untuk mengetahui seberapa jauh efek serangan ini terhadap keutuhan watermark.

TINJAUAN PUSTAKA

Watermarking

Watermarking merupakan sebuah metode yang relatif baru yang dimanfaatkan untuk melindungi data digital dari upaya penggandaan atau manipulasi lainnya secara ilegal. Watermarking atau tanda air berbeda dengan tanda air pada uang kertas. Tanda air pada uang kertas masih dapat dilihat dengan mata telanjang (pada posisi tertentu), tetapi watermarking pada data digital disini tidak akan dirasakan kehadirannya oleh manusia tanpa menggunakan alat bantu mesin pengolah digital seperti komputer dan sejenisnya. Jadi watermarking dapat diartikan sebagai suatu teknik menyembunyikan data atau informasi "rahasia" ke dalam suatu data lain untuk "ditumpang", tetapi orang lain tidak menyadari akan kehadiran adanya data tambahan pada hostnya. Sehingga seolah - olah tidak ada perbedaan antara data host sebelum dan sesudah proses watermarking [4].

Beberapa aplikasi watermarking yang sering digunakan adalah :

a. *Owner Identification* (tanda pengenalan kepemilikan)

Pada aplikasi ini pemilik data dapat menanamkan informasi hak cipta pada data host, sehingga usaha untuk menghilangkan informasi hak cipta akan berdampak menurunkan kualitas data host.

b. *Proof of ownership* (Bukti kepemilikan)

Selain digunakan sebagai tanda pengenalan pemilikan, watermarking juga dapat digunakan sebagai bukti kepemilikan. Pembuktian ini diperlukan bilamana terjadi perselisihan hak kepemilikan atas data digital.

c. *Authentication* (Keaslian)

Watermarking dapat juga digunakan sebagai teknik untuk membuktikan keaslian suatu data digital. Hal ini disebabkan, watermark akan selalu melekat pada data host. Sehingga jika data host mengalami perubahan baik di cropping atau diubah ke dalam format lainnya maka watermarknya akan selalu bersama dengan data host.

d. *Fingerprinting*

Fingerprinting digunakan untuk menelusuri penggunaan ilegal terhadap data host. Pemilik data host dapat menanamkan watermark berbeda ke data host yang akan didistribusikan ke pelanggan yang berbeda. Dengan cara ini maka penggunaan ke pihak ketiga akan dapat dideteksi, karena adanya watermark yang berbeda untuk pelanggan yang berbeda.

e. *Medical safety*

Pada aplikasi ini, watermark yang berupa data pasien (nama, tanggal) dapat ditanamkan ke data host (medical image) sehingga dapat meminimalisir adanya kesalahan data.

f. *Broadcast Monitoring*

Pada aplikasi ini watermark ditanamkan ke dalam tiap video maupun suara sebelum ditayangkan oleh stasiun televisi atau radio. Untuk itu diperlukan stasiun pengamat otomatis yang akan menerima tayangan tersebut sehingga akan dapat mengekstrak informasi watermark yang dibawa dan sekaligus mencatat informasi tayangan yang muncul.

Motion Blur

Pengkaburan citra sering digunakan dalam proses pelembutan citra yang bertujuan untuk menekan gangguan (*noise*) pada citra. Gangguan pada citra umumnya berupa variasi intensitas suatu *pixel* yang tidak berkorelasi dengan *pixel-pixel* tetangganya. Secara kasat mata, gangguan mudah dilihat oleh mata, karena tampak berbeda dengan *pixel* tetangganya.

Pengkaburan citra didapatkan dengan mengkonvolusikan citra dengan sebuah penapis (filter). Penapis ini disebut juga penapis lolos-rendah (low-pass filter), karena menekan komponen yang berfrekuensi tinggi dan meloloskan yang berfrekuensi rendah.

$$g(x,y) = f(x,y) * h(x,y)$$

$h(x,y)$: fungsi penapis

Aturan untuk mendapatkan penapis lolos-rendah, adalah :

- 1) Semua koefisien penapis harus positif.**
- 2) Jumlah semua koefisien harus sama dengan 1**

Pengkaburan citra dengan *motion blur* merupakan efek yang disebabkan karena perpindahan objek atau bisa juga karena pergerakan kamera pada saat pengambilan gambar [6].

Fungsi penapis *motion blur* dihasilkan dengan menggunakan fungsi yang sudah ada pada software Matlab7.1 dengan menentukan parameter radius pengkaburan dan sudut yang dibentuk.

Kompresi JPEG

Data digital terutama citra memiliki ukuran file yang cukup besar. Hal ini mengakibatkan adanya beberapa permasalahan yang sering terjadi pada pemrosesan citra. Dengan ukuran file yang cukup besar memberi dampak pada ruang penyimpanan dan waktu transfer data. Untuk itu diperlukan upaya kompromi dengan menggunakan kompresi.

Sebenarnya kompresi merupakan upaya dilematis. Disatu sisi menguntungkan karena berkurangnya ukuran file tetapi disisi lain merugikan karena menurunnya kualitas citra.

Kompresi dibedakan menjadi dua jenis, yakni *lossless* dan *lossy*. Pada kompresi *Lossless* diperuntukkan ketika terdapat suatu persyaratan bahwa informasi asli tetap utuh. Pesan asli direkonstruksi kembali seperti aslinya. Contoh tipe kompresi ini adalah citra GIF dan BMP. Sedangkan kompresi *Lossy* juga menyimpan tempat, tetapi integritas citra asli tidak terjaga. Contoh metode ini terdapat pada citra JPG dan hasil kompresi sangat baik [8].

JPEG adalah salah satu standar kompresi citra yang dikembangkan oleh *Joint Photographic Expert Group*, yang didesain untuk kompresi citra *full-colour* atau *gray-scale*. JPEG menggunakan metode *lossy* ada informasi yang hilang, tetapi masih dapat ditolerir oleh persepsi mata. JPEG memanfaatkan keterbatasan mata manusia dalam melihat warna. Mata manusia tidak sensitif terhadap perubahan warna yang kecil, dibandingkan dengan perubahan kecerahan (*brightness*) yang kecil. Tetapi, jika dianalisis menggunakan komputer, akan terlihat kualitas citranya. Kompresi ini sangat cocok sekali diaplikasikan pada foto-foto digital (*Digital Images*).

Untuk mengetahui kualitas citra hasil kompresi, maka dikenal besaran PSNR (*peak signal to noise ratio*). PSNR memiliki satuan *decibel* (dB) yang dihitung untuk mengukur perbedaan antara citra semula dengan citra hasil kompresi dengan rumus [8] :

dimana :

$f'(i,j)$: *pixel* citra kompresi

$f(i,j)$: *pixel* citra semula.

b : 255 (nilai maksimum derajat keabuan)

M, N : Lebar dan Tinggi citra

Dari persamaan diketahui, PSNR berbanding terbalik dengan rms. Nilai rms yang rendah yang menyiratkan bahwa citra hasil kompresi tidak jauh berbeda dengan citra semula, akan menghasilkan PSNR yang tinggi. Semakin tinggi nilai PSNR-nya, maka kualitas citra akan semakin bagus, yang artinya bahwa proses kompresi tidak menurunkan kualitas citra, sebaliknya jika nilai PSNR-nya semakin kecil, maka proses kompresi menyebabkan penurunan kualitas citra.

PEMBAHASAN

Metode

Pada penelitian ini digunakan adalah citra digital yang sudah ditanami *watermark* pada domain wavelet dan frekuensi. Metode yang digunakan dapat digambarkan sebagai berikut :

Pada tahap awal, citra asli dan citra watermark dengan sistem watermarking [2] dilakukan proses penanaman sehingga diperoleh citra terwatermark. Selanjutnya citra terwatermark dikenakan serangan kompresi JPEG dengan menggunakan parameter quality tertentu dan *Motion Blur* dengan parameter radius dan sudut tertentu. Kemudian citra terwatermark yang telah terkena serangan dengan menggunakan sistem watermarking diekstrak untuk memperoleh citra watermark terekstrak. Citra watermark terekstrak inilah yang akan diuji kemiripannya dengan citra watermark asli. Pengujian dilakukan dengan menggunakan rumus *Normalized Cross Correlation* (NC) [3]

..... (2)

Hasil

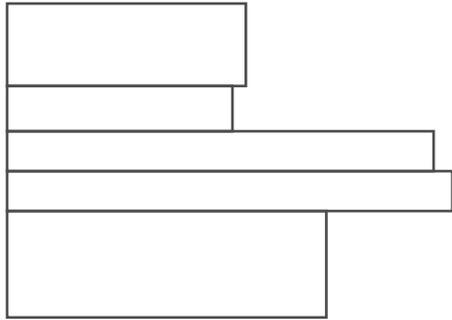
Citra yang disimulasikan adalah citra mandrilRGB dengan ukuran 256 x 256 berjenis RGB 24 bit, yang telah ditanami *watermark* pada domain wavelet Daubechies 4 dengan α 0.1, 0.5, 1.0, 1.5, 2.0 dan untuk domain frekuensi pada nilai K 10, 50, 100, 150, 200. dan citra watermarknya merupakan citra biner dengan ukuran 20 x 50.

(a) (b)
Gambar 3. (a) Citra mandrilRGB.bmp 256 x 256 yang telah disisipi watermark.
(b) Citra watermark1.bmp 20x50.

Data - data di atas kemudian disimulasikan dalam simulasi pengujian [7] dan pengekstrakan [10, 11] dengan menggunakan program aplikasi Graphical User Interface (GUI) Matlab 7.1 yang diperlihatkan pada gambar 4.

(a) (b)
(c)
Gambar 4. (a) Pengujian watermark. (b) Pengekstrakan watermark pada domain wavelet, (c) Pengekstrakan watermark pada domain frekuensi

Mula-mula citra mandrilRGB yang telah ditanami *watermark* dilakukan ekstraksi terlebih dahulu untuk mendapatkan nilai NC sebelum dilakukan serangan untuk masing-masing domain pada konstanta α dan K yang telah ditentukan. Pada Gambar 5.a diberikan serangan yang berupa *motion blur* pada nilai radius 5 dan sudut 30 atau kompresi JPEG pada nilai kualitas 20. Kemudian citra hasil serangan tersebut diekstrak kembali sesuai dengan domain penanamannya sehingga diperoleh *watermark* terekstrak hasil serangan. dengan nilai NC-nya. Secara lengkap dari beberapa percobaan yang dilakukan diperlihatkan pada beberapa tabel berikut:



ALPHA	NC	Watermark	K
A 10	B 10	Ekstrak	D 5
B 20	C 10	C 20	A 10
C 10	D 20	D 10	B 10
D 10	A 5	A 5	C 10
		B 20	

Kemudian dari tabel tersebut kita dapatkan matrik berdasarkan permintaan mesin oleh lahan, adalah sebagai berikut :

- 1 2 3 4 > pada lahan semangka (lahan 1)
- 2 3 4 1 > pada lahan jagung (lahan 2)
- 3 4 1 2 > pada lahan kacang (lahan 3)
- 4 1 2 3 > pada lahan melon (lahan 4)

sedangkan matrik waktu proses lahan oleh setiap mesin diperoleh sebagai berikut : (satuan waktu)

10	20	10	10
10	10	20	5
20	10	5	20
5	10	10	10

Dalam JSP ALGEN, sebuah solusi direpresentasikan sebagai sebuah individu. Contoh sebuah solusi atau individu adalah sebagai berikut :

1. 4 3 3 3 4 4 3 1 4 2 2 2 2 1 1

individu diatas merepresentasikan sebuah solusi atau jadwal sebagai berikut :

- o 1 yang pertama mengacu pada lahan 1 mesin yang pertama yaitu mesin no 1
- o 4 yang pertama mengacu pada lahan 4 mesin yang pertama yaitu mesin no 4
- o 3 yang pertama mengacu pada lahan 3 mesin yang pertama yaitu mesin no 3
- o 3 yang kedua mengacu pada lahan 3 mesin yang kedua yaitu mesin no 4
- o 3 yang ketiga mengacu pada lahan 3 mesin yang ketiga yaitu mesin no 1
- o 4 yang kedua mengacu pada lahan 4 mesin yang kedua yaitu mesin no 1
- o 4 yang ketiga mengacu pada lahan 4 mesin yang ketiga yaitu mesin no 2
- o Dan seterusnya

Gambar 4. Waktu tempuh

Setelah semua pekerjaan terselesaikan maka waktu yang ditempuh oleh jadwal tersebut adalah 145 satuan waktu.(terlihat pada Gambar 4)

Kemudian kedua matrik tersebut kami masukan ke dalam aplikasi JSP untuk mendapatkan jadwal pengerjaan lahan oleh masing-masing mesin dengan waktu yang optimal. Parameter Algen yang digunakan didefinisikan sebagai berikut :

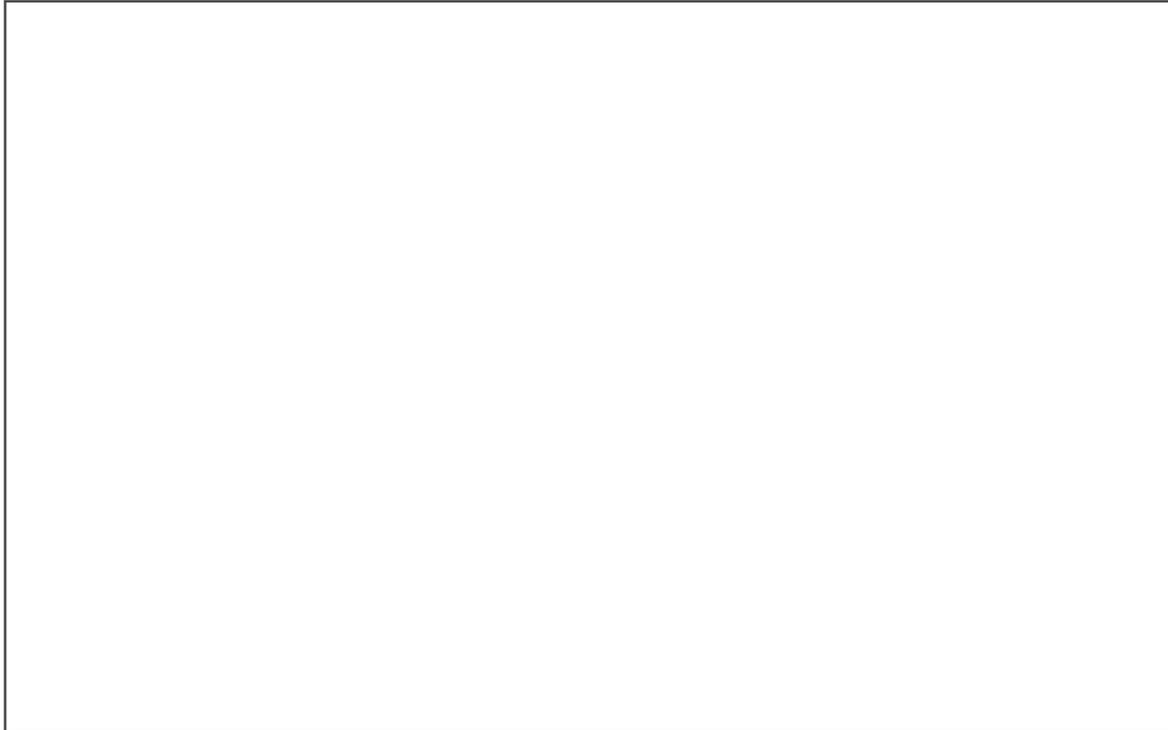
Laju Crossover : 0.8

Laju Mutasi : 0.3

Maksimum Generasi : 500

Eror : 3

Otput yang di peroleh dari aplikasi tersebut adalah sebagai berikut :



Gambar 5. Hasil Output

Diperoleh susunan jadwal sebagai berikut :

Hasil Jadwal Terpendek dari Generasi yang Terpilih

2 3 1 1 1 3 4 3 3 2 1 4 4 4 2 2 75
2 3 2 4 1 4 1 3 2 4 3 2 1 1 3 4 60

Hasil Akhir Jadwal Terpendek dari 13 Generasi

Jadwal Optimal / Terpendek : 2 3 2 4 1 4 1 3 2 4 3 2 1 1 3 4
Dengan Waktu : 60 satuan waktu
Pada Generasi ke : 3

Jadwal optimal / terpendek dapat diartikan sebagai berikut : (urutan jadwalnya)

- o 2 yang pertama mengacu pada lahan 2 mesin yang pertama yaitu mesin no 2
- o 3 yang pertama mengacu pada lahan 3 mesin yang pertama yaitu mesin no 3
- o 2 yang kedua mengacu pada lahan 2 mesin yang kedua yaitu mesin no 3
- o 4 yang pertama mengacu pada lahan 4 mesin yang pertama yaitu mesin no 4
- o 1 yang pertama mengacu pada lahan 1 mesin yang pertama yaitu mesin no 1
- o 4 yang kedua mengacu pada lahan 4 mesin yang kedua yaitu mesin no 1
- o 1 yang kedua mengacu pada lahan 1 mesin yang kedua yaitu mesin no 2
- o Dan seterusnya.

Gambar 6. Waktu tempuh jadwal optimal

Sehingga setelah semua pekerjaan terselesaikan maka waktu yang ditempuh oleh jadwal tersebut adalah 60 satuan waktu (terlihat pada Gambar 6).

KESIMPULAN DAN SARAN

Kesimpulan yang dapat diambil dari penelitian ini yaitu, algoritma genetika dapat digunakan dalam pencarian jadwal dalam permasalahan JSP dan dapat memberikan beberapa variasi solusi dalam permasalahan antrian mesin pada pengolahan lahan pertanian. Digunakan program penyelesaian JSP dengan algoritma genetika menggunakan software Delphi 6. Solusi yang didapatkan dalam penyelesaian JSP dengan algoritma genetika belum tentu merupakan hasil yang paling optimal. Hal ini dikarenakan algoritma genetika menggunakan bilangan random yang berperan dalam pencarian sehingga dengan nilai parameter yang sama dapat menghasilkan solusi yang berbeda pada waktu yang berbeda. Pada kasus-kasus nyata diharapkan dengan simulasi ini dapat memberikan gambaran umum tentang permasalahan yang terjadi dalam dunia pertanian. Untuk kelanjutannya sebaiknya diadakan studi lebih lanjut.

DAFTAR PUSTAKA

- 1] Davis, Lawrence, *Handbook of Genetic Algorithms*, Von Nostrand Reinhold, 1991.
- 2] Garen, J., *Multiobjective Job-Shop Scheduling With Genetic Algorithms Using a New Representation and Standard Uniform Crossover*, Department of Economics University of Osnabruck, Germany
- 3] Goldberg, David E., *Genetic Algorithm in Search, Optimization and Machine Learning*, Addison Wesley Publishing, 1989.
- 4] Michalewicz, Zbigniew, *Genetic Algorithms + Data Structure = Evolution Program*, Springer, 1992.
- 5] Oey, Kasin, Scott J. Manson, *Sceduling Batch Processing Machines in Complex Job Shops*, Department of Industrial Engineering University of Arkansas USA, 2001.
- 6] Ponnambalam, S.G, Aravindan, P., Sreenivasa Rao, P., *Comparative Evaluation of Genetic Algorithm for Job-Shop Sceduling*, Taylor & Francis Ltd, 2001.
- 7] Taillard, E, *Benchmarks Basic Scheduling Problems*, ORWP89/21 Dec, 1989.
- 8] Tsujimura, et. al., *Effects of Symbiotic Evolution in Genetic Algorithms for Job-Shop Scheduling*, Department of Industrial and Information Systems Engineering Ashikaga Institute of Technology, 2001.

WATERMARKING PADA DOMAIN FREKUENSI UNTUK Memberikan Identitas (Watermark) pada Citra Digital

Zaki Rakhmatulloh, Aris Sugiharto, Eko Adi Sarwoko
Jurusan Matematika FMIPA UNDIP
Jl. Prof. Soedarto, Kampus UNDIP Tembalang, Semarang

Abstrak: *Watermarking* merupakan salah satu solusi dalam memecahkan penggandaan ilegal produk digital. Pada penelitian ini *watermarking* citra digital ditransformasikan menggunakan *discrete cosine transform (DCT)*. Pada proses penanaman *watermark*, citra ditransformasikan menggunakan *DCT* menjadi domain frekuensi yang menghasilkan tiga area yaitu *Low Frequency (FL)*, *Medium Frequency (FM)*, dan *High Frequency (FH)*. Bit-bit *watermark* ditanam pada area *FM* dengan menggunakan nilai Koefisien Selisih (K). Kualitas citra ter-*watermark* diukur dengan *Peak Signal of Noise Ratio (PSNR)*. Semakin besar nilai K diperoleh nilai *PSNR* yang semakin kecil.

Kata Kunci: *watermarking, watermark, DCT, PSNR*

PENDAHULUAN

Selama ini penggandaan atas produk digital, seperti citra digital dilakukan begitu bebas dan leluasa secara ilegal. Hasil penggandaan tersebut memiliki kualitas yang sama dengan produk digital aslinya. Namun, pemegang hak cipta produk digital tidak mendapatkan royalti dari usaha penggandaan diatas, akibatnya pemegang hak cipta produk digital dirugikan atas usaha ilegal di atas.

Hampir semua produk digital yang tersebar di internet jarang mencantumkan informasi pemiliknya. Salah satu produk digital yang tersebar di internet adalah citra digital. Seseorang yang telah mendapatkan citra digital dapat mengklaim bahwa citra digital tersebut adalah hasil karyanya. Karena jika tidak ada bukti kepemilikan citra digital sebelumnya, maka setiap orang dapat mengklaim citra digital tertentu sebagai miliknya.

Berbagai upaya dilakukan untuk melindungi citra digital dari upaya penggandaan secara ilegal. Salah satunya adalah dengan *watermarking*, pada penggunaan *watermarking* akan disisipkan sebuah *watermark* sebagai identitas dari pemilik citra digital yang sah. Pemberian *watermark* dilakukan tanpa merubah citra digital secara langsung sehingga keberadaannya tidak merusak citra digital yang dilindungi. Jika seseorang membuka citra digital yang telah disisipi *watermark*, maka orang tersebut tidak akan menyadari bahwa di dalam citra tersebut telah terkandung label kepemilikan pembuatnya. Sebuah label *watermark* akan selalu terbawa kemana saja citra digital tersebut berada termasuk hasil penggandaannya. Jika dikemudian hari ada orang lain yang mengklaim bahwa citra digital tersebut adalah miliknya, maka pemegang hak cipta tersebut dapat membantahnya dengan cara mengekstrak kembali *watermark* dari citra digital yang disengketakan. Jika *watermark*-nya sama dengan yang dimiliki oleh pemegang hak cipta, maka orang tersebut merupakan pemegang hak cipta citra digital yang sebenarnya.

TINJAUAN PUSTAKA

Watermarking

Watermarking merupakan suatu bentuk dari *Steganography*. *Steganography* adalah ilmu yang mempelajari bagaimana menyembunyikan suatu data pada data yang lain [6]. Sehingga seseorang tidak menyadari kehadiran adanya data tambahan pada data tersebut. Jadi seolah-olah tidak ada perbedaan antara data sebelum dan sesudah proses *watermarking*. Disamping itu data yang ter-*watermark* harus tahan (*robust*) terhadap serangan-serangan baik

secara sengaja maupun tidak sengaja untuk menghilangkan data *watermark* yang terdapat didalamnya.

Watermarking memanfaatkan kekurangan-kekurangan sistem indera manusia seperti mata dan telinga. Sehingga *watermarking* juga dapat didefinisikan sebagai suatu cara penyembunyian data atau informasi tertentu ke dalam suatu data digital lainnya, tetapi tidak diketahui kehadirannya oleh indera penglihatan atau indera pendengaran manusia dan mampu menghadapi proses-proses pengolahan signal digital sampai pada tahap tertentu [2].

Terdapat dua proses dalam *watermarking*, yaitu proses penyisipan dan proses pengekstrakan. Proses penyisipan adalah proses menyisipkan *watermark* ke dalam citra digital yang akan disisipi. Untuk menyisipkan suatu *watermark* ke dalam citra digital, diperlukan bilangan selisih (K). K merupakan bilangan yang menjadikan *pixel-pixel* yang telah ditukar antara dua blok yang telah ditentukan sebagai area penanaman *watermark* memiliki selisih tertentu.

Gambar 1. Proses Penyisipan *Watermark*

Sedangkan proses pengekstrakan adalah proses mengekstrak kembali *watermark* yang telah tertanam pada citra terwatermark.

Untuk proses pengekstrakan dibutuhkan *watermark* asli dari citra ter-*watermark* sebagai pembanding ukuran dalam membentuk kembali *pixel-pixel watermark* yang telah ditanamkan dalam citra ter-*watermark*.

Gambar 2. Proses Pengekstrakan *Watermark*

Sifat dan Manfaat Watermarking

Untuk mendapatkan suatu digital watermarking yang baik, maka teknik yang digunakan memenuhi sifat-sifat di bawah ini [2] :

1. Tidak tampak (*Invisible*) untuk data digital seperti citra oleh pihak lain dengan menggunakan panca indera terutama indera penglihatan.
2. Tidak mudah dihapus atau diubah secara langsung oleh pihak yang tidak bertanggung jawab, dan tidak mudah terhapus atau berubah dengan adanya proses pengolahan sinyal digital.
3. Tidak menghambat proses penduplikasian tetapi penyebaran data digital tersebut tetap dapat dikendalikan dan diketahui.

Ada beberapa manfaat yang dapat dicapai dari penggunaan watermarking, sebagai suatu penyembunyian data pada data digital lain [1], yaitu:

1. *Tamper-proofing*

Watermarking digunakan sebagai alat indikator yang menunjukkan data digital asli telah mengalami perubahan dari aslinya (mengecek integritas data).

2. *Feature location*

Watermarking sebagai alat identifikasi isi dari data digital pada lokasi-lokasi tertentu, misalnya penamaan suatu objek tertentu dari beberapa objek yang ada pada suatu citra digital.

3. *Annotation/caption*

Watermark hanya digunakan sebagai keterangan tentang data digital itu sendiri.

4. *Copyright-Labeling*

Watermarking digunakan sebagai metoda untuk menyembunyikan label hak cipta pada data digital atau sebagai bukti autentik kepemilikan atas dokumen digital tersebut.

Discrete Cosine Transform (DCT)

Transformasi pada *watermarking* digunakan untuk menyederhanakan penyelesaian dan untuk mengetahui suatu informasi tertentu yang tidak tersedia sebelumnya.. *Discrete cosine transform (DCT)* memecah citra digital menjadi blok-blok kecil dengan ukuran yang tetap kemudian dikonversikan dari domain spatial menjadi domain frekuensi. *Discrete cosine transform (DCT)* merekonstruksi matrik citra ke dalam tiga area frekuensi yaitu *Low Frequency (FL)*, *Medium Frequency (FM)*, dan *High Frequency (FH)* [3].

Gambar 3. *Discrete cosine transform (DCT)*

Peak Signal to Noise Ratio (PSNR)

Peak Signal to Noise Ratio (PSNR) merupakan nilai (rasio) yang menunjukkan tingkat toleransi *noise* tertentu terhadap banyaknya *noise* pada suatu sinyal citra. *Noise* adalah kerusakan sinyal pada bagian tertentu dalam sebuah citra sehingga mengurangi kualitas sinyal tersebut. Dengan kata lain *PSNR* merupakan suatu nilai yang menunjukkan kualitas suatu sinyal citra. [7]

Untuk menentukan nilai *PSNR* digunakan rumus :



Sedangkan *MSE (Mean Square Error)* adalah kesalahan kuadrat rata-rata sinyal-sinyal piksel citra hasil pemrosesan sinyal terhadap sinyal citra asli. Rumus untuk menghitung *MSE* pada citra digital adalah sebagai berikut [4]:

$$\frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N (I_{ij} - O_{ij})^2$$

(untuk Citra Grayscale)

$$\frac{1}{3 \times M \times N} \sum_{i=1}^M \sum_{j=1}^N (I_{ij} - O_{ij})^2$$

(untuk Citra RGB)

dimana :

M : Baris matriks citra hasil pemrosesan.

N : Kolom matriks citra hasil

i : Baris matriks citra hasil pemrosesan.

$$j$$

I_{ij} : Piksel citra hasil pemrosesan.

$$O_{ij}$$

: Piksel citra asli.

i : index matriks (*Red = 1, Green = 2, dan Blue = 2*).

PEMBAHASAN

Metode

Pada penelitian ini semua bahan yang digunakan adalah citra digital yang mudah diperoleh di berbagai media. Metode yang digunakan dapat digambarkan sebagai berikut :

Pertama, citra original disisipi citra *watermark* menghasilkan citra terwatermark. Kualitas citra terwatermark ini kemudian diuji dengan parameter yang digunakan adalah *Peak Signal to Noise Ratio (PSNR)*.

Selanjutnya citra terwatermark yang telah diketahui nilai *PSNR*, diekstrak menghasilkan citra *watermark* terekstrak. *Watermark* terekstrak ini hanya digunakan sebagai pembanding kemiripan secara visual dengan citra *watermark* asli.

Metode ini dilakukan berulang-ulang dengan konstanta transformasi yang berbeda-beda.

Citra digital dengan nilai *PSNR* tertentu dapat dikategorikan ke dalam 5 kategori sebagai berikut [7]:

Tabel 1. Nilai *PSNR*

PSNR (dB)	Picture Quality
60	Excellent, no noise apparent
50	Good, a small amount of noise but picture quality good
40	Reasonable, fine grain or snow in the picture, some fine detail lost
30	Poor picture with a great deal of noise
20	Unusable

Gambar 4. Uji Kualitas Citra Terwatermark

Hasil

Citra yang disimulasikan adalah citra *Peppers.bmp*, berukuran 256 x 256 piksel, dengan jenis citra *RGB* 24 bit. Sedangkan citra *watermark*-nya adalah citra *zaki_3232.bmp* berjenis *grayscale* 8 bit, berukuran 32 x 32 piksel.

(a) (b)
Gambar 5. (a) *Peppers.bmp*, (b) *zaki_3232.bmp*

Kedua citra di atas kemudian disimulasikan dalam sistem *watermarking* dengan menggunakan *Graphical User Interface (GUI) MATLAB 7.1* [5] yang diperlihatkan pada gambar 6.



(a)



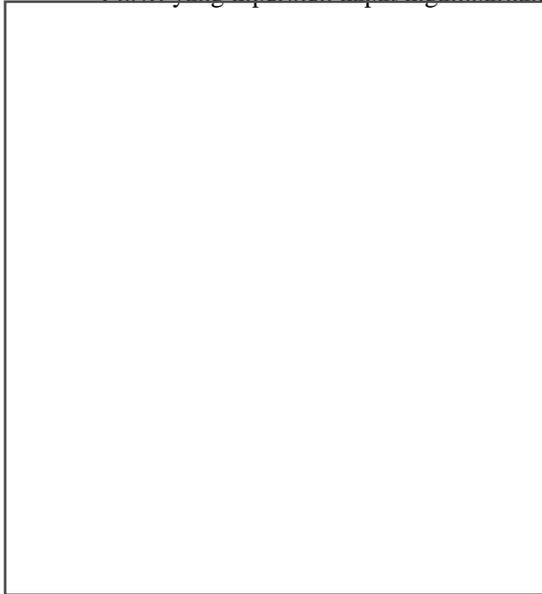
(b)

Gambar 6. (a) Penyisipan/ Penanaman Watermark, (b) Pengekstrakan *Watermark*

Gambar 6 (a) adalah simulasi proses penyisipan citra *watermark* ke dalam citra asli. *PSNR* citra ter-*watermark* (citra asli yang telah tersisipi *watermark*) dapat terlihat pada panel *Catatan Proses* pada tampilan simulasi. Sedangkan gambar 6 (b) adalah simulasi proses pengekstrakan kembali *watermark* yang telah tersisipi pada citra ter-*watermark*.

Pengujian dengan menggunakan Koefisien selisih (K) 1, 50, 100, 150, 200 diperoleh hasil *PSNR* dan *watermark* terekstrak sebagaimana pada Tabel 2.

PSNR yang diperoleh dapat digambarkan dalam bentuk grafik sebagai berikut :



Gambar 7 . Grafik *PSNR DCT*

Tabel 2. Nilai *PSNR* dan *Watermark* ekstrak

Koefisien Selisih (K)	DCT PSNR*	Watermark Terekstrak
1	46.2751	
50	38.3492	
100	33.204	
150	29.9508	
200	27.5787	

*PSNR** : *PSNR* untuk citra ter-*watermark*

Grafik pada gambar 7 menyatakan hubungan antara koefisien selisih (K) dengan *Peak Signal to Noise Ratio (PSNR)*. Angka-angka pada absis X menunjukkan skala K, sedangkan pada ordinat Y menunjukkan skala *PSNR*.

Dari grafik terlihat garis bergerak menurun dari kiri ke kanan. Hal ini berarti bahwa semakin besar nilai K yang digunakan dalam proses penyisipan *watermark*, maka akan berpengaruh pada penurunan nilai *PSNR*. Dengan kata lain semakin besar nilai K yang digunakan, maka semakin menurun kualitas citra ter-*watermark*.

Sedangkan pada proses ekstraksi, secara *visual* dapat terlihat bahwa nilai K berpengaruh pada

watermark ekstrak. Semakin besar nilai K yang digunakan, maka akan semakin mirip *watermark* yang terekstrak dengan *watermark* asli.

PENUTUP

Dari hasil yang diperoleh pada penelitian ini maka dapat disimpulkan bahwa semakin besar koefisien selisih (K) yang digunakan dalam proses transformasinya, berakibat pada semakin menurunnya kualitas citra *ter-watermark*, tetapi secara visual citra *watermark* ekstrak semakin mirip dengan citra *watermark* asli..

DAFTAR PUSTAKA

- [1]. Bender, W. Gruhl, D. Morimoto, N. Lu, A., Techniques for Data Hiding, *IBM System Journal*, Vol.35, 1996.
- [2]. H. Supangkat, dkk, *Paper : Watermarking sebagai Teknik Penyembunyian Label Hak Cipta pada Data Digital*. Institut Teknologi Bandung. 2000.
- [3]. Langelaar, G. Setyawan, I. Lagendijk, R.L., Watermarking Digital Image and Video Data, in *IEEE Signal Processing Magazine*, Vol. 17, pp. 20-43, 2000.
- [4]. Li Tan, Choo, *Still Image Compression using Wavelet Transform*, School of Information Technology and Electrical Engineering, The University of Queensland. Queensland. 2001.
- [5]. Littlefield, Bruce and Duane Hanselman, *MATLAB Bahasa Komputasi Teknis*, Andi and Pearson Education Asia Pte, Ltd., Yogyakarta. 2000.
- [6]. Schneiner, B., *Applied Cryptography: Protocols, algorithm, and Source Code in C*, New York: Wiley, 1994.
- [7]. www.cctv-information.co.uk/constant2/sn_ratio.html
- [8]. www.mathworks.com

Key

Penanaman
Watermark

Citra Original

Citra Terwatermark

Transformasi

Wavelet

Invers Transformasi Wavelet

Watermark

Key

Pengekstrakan
Watermark

Citra Terwatermark

Transformasi
Wavelet

Ubah Vektor Watermark menjadi Citra

Watermark Terekstrak

Watermark

Citra Digital

DWT

LL

LH

HL

HH

PSNR(dB)	Picture Quality
60	Excellent, no noise apparent
50	Good, a small amount of noise but picture quality good
40	Reasonable, fine grain or snow in the picture, some fine detail lost
30	Poor picture with a great deal of noise
20	
	Unusable

Sistem Watermarking

Citra Terwatermark

Citra Original

Citra Watermark

Uji Kualitas

Citra Watermark Terekstrak



Citra Asli

Blok range 8x8

Blok domain 16x16

Penskalaan

Uji Kemiripan

Transformasi affine

Parameter PIFS

Citra De-kompresi

Blok domain 16x16

Parameter PIFS

Blok range 8x8



Proses Penggolongan

Database Template

template2

Proses
Identifikasi

Sensor Biometrik

Extractor ciri

wajah

Extractor ciri

Sensor Biometrik

wajah

Alat Pencocokan ciri

Data terwatermark

Watermarking

Data Original

Watermark

Key

Gambar 1. Sistem Watermarking

Uji Kemiripan

Citra Watermark terekstrak

Kompresi JPEG dan Motion Blur

Citra Terwatermark

Sistem Watermarking

Citra Watermark



Gambar 2. Metode kompresi JPEG dan Uji kemiripan



Y
g
Start
Generate initial population P_t
Evaluate population P_t
While stopping criteria not satisfied Repeat
 $P_t = P_{t+1}$
End
Select elements from P_t to copy into P_{t+1}
Crossover elements of P_t and put into P_{t+1}
Mutation elements of P_t and put into P_{t+1}
Evaluate new population P_{t+1}
yes
no
Start
Input file data
atau membuat data

Proses Algoritma Genetika

output file data

End

Menentukan besarnya parameter Algoritma Genetika

4

3

1

2

3

4

2

1

4

3

2

1

1

3

4

2

m1

m2

m3

m4

Satuan waktu

5

10

20

30

35

55

45

65

75

95

105

115

125

135

140

145

4

3

1

2

3

4

2

1

4

3

2

1

1

3

4

2

m1

m2

m3

m4

Satuan waktu

5

10

20

30

35

FM

FH

FL

DCT

Citra Digital

Citra Terwatermark

Citra Original

Sistim Watermarking

Citra Watermark Terekstrak

Uji Kualitas

Citra Watermark

1

2

3

a

b

c