

**PERBANDINGAN KOMPRESI TEKS DENGAN
MENGUNAKAN ARITHMETIC CODING DAN LZ77**

Skripsi



oleh
ANDI SUWITO
22043575

PROGRAM STUDI TEKNIK INFORMATIKA FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS KRISTEN DUTA WACANA
2013

**PERBANDINGAN KOMPRESI TEKS DENGAN
MENGUNAKAN ARITHMETIC CODING DAN LZ77**

Skripsi



Diajukan kepada Program Studi Teknik Informatika Fakultas Teknologi Informasi
Universitas Kristen Duta Wacana
Sebagai Salah Satu Syarat dalam Memperoleh Gelar
Sarjana Komputer

Disusun oleh

ANDI SUWITO
22043575

PROGRAM STUDI TEKNIK INFORMATIKA FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS KRISTEN DUTA WACANA
2013

PERNYATAAN KEASLIAN SKRIPSI

Saya menyatakan dengan sesungguhnya bahwa skripsi dengan judul:

PERBANDINGAN KOMPRESI TEKS DENGAN MENGGUNAKAN ARITHMETIC CODING DAN LZ77

yang saya kerjakan untuk melengkapi sebagian persyaratan menjadi Sarjana Komputer pada pendidikan Sarjana Program Studi Teknik Informatika Fakultas Teknologi Informasi Universitas Kristen Duta Wacana, bukan merupakan tiruan atau duplikasi dari skripsi keserjanaan di lingkungan Universitas Kristen Duta Wacana maupun di Perguruan Tinggi atau instansi manapun, kecuali bagian yang sumber informasinya dicantumkan sebagaimana mestinya.

Jika dikemudian hari didapati bahwa hasil skripsi ini adalah hasil plagiasi atau tiruan dari skripsi lain, saya bersedia dikenai sanksi yakni pencabutan gelar keserjanaan saya.

Yogyakarta, 11 September 2013



ANDI SUWITO
22043575

HALAMAN PERSETUJUAN

Judul Skripsi : PERBANDINGAN KOMPRESI TEKS DENGAN
MENGUNAKAN ARITHMETIC CODING DAN
LZ77

Nama Mahasiswa : ANDI SUWITO

N I M : 22043575

Matakuliah : Skripsi (Tugas Akhir)

Kode : TIW276

Semester : Genap

Tahun Akademik : 2012/2013

© UKDW

Telah diperiksa dan disetujui di
Yogyakarta,
Pada tanggal 11 September 2013

Dosen Pembimbing I



Dra. Widi Hapsari, M.T.

Dosen Pembimbing II



Erick Kurniawan, S.Kom., M.Kom.

HALAMAN PENGESAHAN

PERBANDINGAN KOMPRESI TEKS DENGAN MENGGUNAKAN ARITHMETIC CODING DAN LZ77

Oleh: ANDI SUWITO / 22043575

Dipertahankan di depan Dewan Penguji Skripsi
Program Studi Teknik Informatika Fakultas Teknologi Informasi
Universitas Kristen Duta Wacana - Yogyakarta
Dan dinyatakan diterima untuk memenuhi salah satu syarat memperoleh gelar
Sarjana Komputer
pada tanggal 14 Agustus 2013

Yogyakarta, 11 September 2013
Mengesahkan,

Dewan Penguji:

1. Dra. Widi Hapsari, M.T.
2. Erick Kurniawan, S.Kom., M.Kom.
3. Antonius Rachmat C., SKom., M.Cs
4. Budi Susanto, SKom., M.T.



Handwritten signatures of the examiners on a set of horizontal lines.



Dekan

(Drs. Wimmie Handiwidjojo, MIT.)

Ketua Program Studi

(Nugroho Agus Haryono, M.Si)

UCAPAN TERIMA KASIH

Suatu kebahagiaan yang besar bagi penulis dapat menyelesaikan Tugas Akhir atau Skripsi dengan judul “Kompresi Teks Dengan Menggunakan Arithmetic Coding Dan LZ77” ini dengan sebaik-baiknya.

Penulisan laporan Tugas Akhir ini merupakan salah satu syarat atau yang harus penulis penuhi untuk memperoleh gelar Sarjana Komputer. Selain itu juga bertujuan untuk mengasah kemampuan mahasiswa agar dapat menghasilkan sebuah karya tulis yang dapat dipertanggung-jawabkan secara ilmiah, sehingga dapat bermanfaat kedepannya bagi siapapun yang merujuk karya tulis ini sebagai referensi untuk penelitian ilmiah selanjutnya.

Dalam menyelesaikan Tugas Akhir ini banyak pihak-pihak yang baik secara langsung maupun tidak langsung sudah membantu dan mendorong semangat penulis untuk segera menuntaskan tugas dan kewajiban ini. Dan pada kesempatan ini, dengan segala rasa hormat dan kerendahan hati, penulis ingin mengucapkan Terima Kasih Banyak kepada semua pihak yang ada di bawah ini :

1. **Guru Buddha**, yang ajarannya tentang Kehidupan dan Pikiran selalu menguatkan penulis dalam melewati dan menjalani berbagai situasi sulit saat penggarapan Tugas Akhir ini.
2. **Restyandito, S.Kom., MSIS**. Selaku Dosen Pembimbing I yang terdahulu, yang sudah memberikan arahan, bimbingan, serta dorongan semangat yang penuh dengan kesabaran dan keramahan dalam menyusun Bab Satu hingga Bab Tiga.
3. **Dra. Widi Hapsari, M.T.** Selaku Dosen Pembimbing I yang menggantikan Dosen Pembimbing I yang terdahulu, yang juga sudah dengan sabar, ramah dan penuh ketelitian dalam membimbing penyusunan Bab Empat dan Bab Lima.
4. **Erick Kurniawan, S.Kom., M.Kom.** Selaku Dosen Pembimbing II yang selalu memandu penulis dalam penyusunan laporan dari awal Bab Satu hingga Bab Lima, serta selalu memberikan masukan-masukan dan perbaikan-perbaikan laporan.
5. **Kedua Orang Tua**, yang selalu memberikan dorongan moril yang sangat besar serta mendukung penulis apapun kondisi dan kendalanya.

6. **Adik-adik dan keluarga tercinta**, yang juga selalu memberikan dukungan yang penuh semangat lewat candaan dan tantangan.
7. **Peter Rave**, yang membantu memberikan pengetahuan dan cara pandang baru tentang hidup, sehingga membuat penulis menjadi lebih bersemangat dalam menjalani hari-hari, sehingga penulisan Tugas Akhir pun semakin lancar.
8. **Ko GGn dan Akiun**, yang selalu menemani penulis bersantai atau *refreshing* di tempat-tempat yang spektakuler untuk menghilangkan penat saat mengerjakan Tugas Akhir.
9. **Ko Modo**, yang banyak memberikan banyak dorongan mental, semangat, saran dan referensi-referensi tentang pemrograman, sehingga kegiatan penulisan Tugas Akhir ini juga terasa lebih ringan dan bersemangat.
10. **Segenap Staf dan Dosen Teknik Informatika**, yang sudah memberikan banyak sekali ilmu pengetahuan dan informasi lainnya selama masa perkuliahan.
11. Teman-teman dan pihak-pihak lain yang tidak dapat penulis sebutkan satu per satu yang membantu dalam memberikan dorongan dan semangat.

Semoga jasa kebajikan dan bantuan yang penulis terima dari mereka semua membuahkan kebahagiaan yang berkelimpahan bagi mereka dan keluarga di kehidupan ini dan di kehidupan yang akan datang.

Penulis mengharapkan kritik dan saran yang membangun untuk Tugas Akhir ini dengan kesadaran bahwa karya tulisan ilmiah ini masih jauh dari sempurna sehingga harapan untuk kedepannya adalah agar tulisan ini dapat lebih disempurnakan lagi supaya dapat menjadi acuan yang bermanfaat.

Semoga Semua Makhluh Hidup Berbahagia

Yogyakarta, Agustus 2013

Penulis

INTISARI

PERBANDINGAN KOMPRESI TEKS DENGAN MENGUNAKAN ARITHMETIC CODING DAN LZ77

Kompresi data merupakan sebuah teknik dalam ilmu teknologi informasi yang bertujuan untuk menyajikan data dalam bentuk yang lebih sederhana atau *compact*. Tujuan yang melandasi penelitian tugas akhir ini adalah mengimplementasikan dan menganalisa algoritma Arithmetic Coding dan LZ77 dalam mengkompresi data teks.

Pada ujicoba kompresi data akan membandingkan dua buah metode kompresi data yaitu dengan algoritma Arithmetic Coding dan algoritma LZ77. Ujicoba akan berupa pengkompresian dengan sampel data teks biasa, dimana akan diteliti mengenai jenis kata apa yang lebih baik hasilnya apabila dikompres dengan kedua buah metode di atas.

Dari hasil ujicoba pengimplementasian kedua buah metode didapati bahwa jenis teks yang dapat dikompresi adalah jenis teks yang memiliki karakter dengan redundansi huruf yang kemunculannya banyak atau dengan kata lain teks yang memiliki banyak karakter berulang.

Keywords : Arithmetic Coding, LZ77, Search buffer, Look-ahead buffer, Token, Tuple, Range, Rasio, Hexadecimal.

DAFTAR ISI

HALAMAN JUDUL	i
SAMPUL DALAM	ii
PERNYATAAN KEASLIAN SKRIPSI	iii
HALAMAN PERSETUJUAN	iv
HALAMAN PENGESAHAN	v
UCAPAN TERIMA KASIH	vi
INTISARI	viii
DAFTAR ISI	ix
DAFTAR GAMBAR DAN DAFTAR TABEL	xi
BAB 1 PENDAHULUAN	1
1.1. Latar Belakang Masalah	1
1.2. Perumusan Masalah	1
1.3. Batasan Masalah	1
1.4. Tujuan Penelitian	2
1.5. Metodologi	2
1.6. Sistematika Penulisan	2
BAB 2 TINJAUAN PUSTAKA	
2.1. Tinjauan pustaka	4
2.2. Pengertian Kompresi dan Dekompresi Data	5
2.3. Teknik Kompresi	6
2.3.1. Lossless Compression	6
2.3.2. Lossy Compression	6
2.4. Algoritma Arithmetic Coding	7
2.5. Algoritma LZ77	9
2.6. Perhitungan Rasio	10
BAB 3 PERANCANGAN SISTEM	11
3.1. Pemilihan Bahasa Pemrograman	11
3.2. Rancangan Kerja Sistem	12
3.2.1. Proses Kompresi Arithmetic Coding	12
3.2.2. Proses Kompresi LZ77	12

3.2.3. Proses Dekompresi Arithmetic Coding	15
3.2.4. Proses Dekompresi LZ77	15
3.3. Perancangan Antarmuka	18
3.3.1. Perancangan Tampilan Utama	18
3.3.2. Perancangan Tampilan Form Pengerjaan	18
BAB 4 IMPLEMENTASI DAN ANALISA SISTEM	20
4.1. Implementasi Antar Muka	20
4.1.1. Form Utama	20
4.1.2. Form Kompresi dan Dekompresi	21
4.2. Analisa	22
4.2.1. Analisa Kompresi Arithmetic Coding	23
4.2.2. Analisa Kompresi LZ77	26
BAB 5 KESIMPULAN DAN SARAN	43
5.1. Kesimpulan	43
5.2. Saran	43
DAFTAR PUSTAKA	44
LAMPIRAN A : Listing Program	

©UKDW

DAFTAR GAMBAR DAN TABEL

Daftar Gambar

Gambar	Keterangan	Halaman
2.1	Ilustrasi Kompresi dan Dekompresi data	6
2.2	Proses pengerjaan Arithmetic Coding	8
2.3	Proses encoding LZ77	10
3.1	Flowchart Kompresi Arithmetic Coding	13
3.2	Flowchart Kompresi LZ77	14
3.3	Flowchart Dekompresi Arithmetic Coding	16
3.4	Flowchart Dekompresi LZ77	17
3.5	Tampilan Form Utama	18
3.6	Tampilan Form Kompresi dan Dekompresi	19
4.1	Form Utama (Main Form)	20
4.2	Form Kompresi dan Dekompresi	21
4.3	Grafik hasil kompresi kalimat biasa Arithmetic Coding	24
4.4	Grafik hasil kompresi kalimat berulang Arithmetic Coding	26
4.5	Grafik hasil kompresi kalimat biasa LZ77 dengan <i>look-ahead buffer</i> 10	28
4.6	Grafik hasil kompresi kalimat biasa LZ77 dengan <i>look-ahead buffer</i> 30	29
4.7	Grafik hasil kompresi kalimat biasa LZ77 dengan <i>look-ahead buffer</i> 50	31
4.8	Grafik hasil kompresi kalimat biasa LZ77 dengan <i>look-ahead buffer</i> 70	32
4.9	Grafik hasil kompresi kalimat biasa LZ77 dengan <i>look-ahead buffer</i> 100	34
4.10	Grafik hasil kompresi kalimat berulang LZ77 dengan <i>look-ahead buffer</i> 10	36
4.11	Grafik hasil kompresi kalimat berulang LZ77 dengan <i>look-ahead buffer</i> 30	37
4.12	Grafik hasil kompresi kalimat berulang LZ77 dengan <i>look-ahead buffer</i> 50	39
4.13	Grafik hasil kompresi kalimat berulang LZ77 dengan <i>look-ahead buffer</i> 70	40
4.14	Grafik hasil kompresi kalimat berulang LZ77 dengan <i>look-ahead buffer</i> 100	42

Daftar Tabel

Tabel	Keterangan	Halaman
4.1	Kompresi dan Rasio kalimat biasa Arithmetic Coding	23
4.2	Hasil kompresi dan Rasio kalimat berulang Arithmetic Coding	25
4.3	Hasil kompresi dan Rasio kalimat biasa LZ77 dengan <i>look-ahead buffer</i> 10	26
4.4	Hasil kompresi dan Rasio kalimat biasa LZ77 dengan <i>look-ahead buffer</i> 30	28
4.5	Hasil kompresi dan Rasio kalimat biasa LZ77 dengan <i>look-ahead buffer</i> 50	29
4.6	Hasil kompresi dan Rasio kalimat biasa LZ77 dengan <i>look-ahead buffer</i> 70	31
4.7	Hasil kompresi dan Rasio kalimat biasa LZ77 dengan <i>look-ahead buffer</i> 100	33
4.8	Hasil kompresi dan Rasio kalimat berulang LZ77 dengan <i>look-ahead buffer</i> 10	35
4.9	Hasil kompresi dan Rasio kalimat berulang LZ77 dengan <i>look-ahead buffer</i> 30	36
4.10	Hasil kompresi dan Rasio kalimat berulang LZ77 dengan <i>look-ahead buffer</i> 50	37
4.11	Hasil kompresi dan Rasio kalimat berulang LZ77 dengan <i>look-ahead buffer</i> 70	39
4.12	Hasil kompresi dan Rasio kalimat berulang LZ77 dengan <i>look-ahead buffer</i> 100	41

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Banyak sekali teknik pengkompresian data yang diterapkan pada sebuah software kompresi data saat ini seperti WinZip, WinRAR, WinAce, 7-Zip dan lain-lain menggunakan lebih dari sebuah teknik atau algoritma di dalamnya untuk mengkompresi data-data yang ada.

Di dalam Tugas Akhir ini akan diteliti mengenai teknik pengkompresian dan pendekompresian data teks dengan menggunakan dua buah metode kompresi data untuk mengkompresi data teks. Dua buah metode kompresi yang akan dipakai adalah algoritma Arithmetic Coding dan LZ77 (Lempel Ziv 1977).

1.2. Perumusan Masalah

Adapun masalah-masalah yang ditemui dalam pengkompresian file teks dengan dua buah algoritma di atas adalah :

1. Pengaruh dari kemunculan karakter sebuah teks untuk menghasilkan kompresi yang lebih kecil dari file aslinya.

1.3. Batasan Masalah

Adapun batasan masalah yang ada dalam penulisan tugas akhir ini adalah sebagai berikut:

1. Menggunakan algoritma Arithmetic Coding dan LZ77 pada data teks biasa (*.txt).
2. Akan melakukan operasi Compressing dan Decompressing pada file teks.
3. Batasan nilai range Arithmetic Coding dari 0 (nol) sampai 1 (satu).
4. Mengubah file input LZ77 (untuk proses kompresi) dari ASCII ke heksadesimal.

1.4. Tujuan Penelitian

Tujuan yang melandasi penelitian tugas akhir ini adalah mengimplementasikan dan menganalisa algoritma Arithmetic Coding dan LZ77 dalam mengkompresi data teks.

1.5. Metodologi

Tugas Akhir ini akan dikerjakan dengan menggunakan metodologi studi pustaka sebagai sumber pembelajaran dan acuan teori-teori mengenai algoritma Arithmetic Coding dan LZ77 untuk mempelajari masalah dan langkah-langkah kompresi dan dekompresi data. Juga akan melakukan studi internet sebagai tempat bertanya seperti di forum-forum dan mailing list dan sumber acuan yang lebih luas untuk mencari materi yang diperlukan mengenai algoritma Arithmetic Coding dan LZ77.

Di dalam Tugas Akhir ini akan dilakukan percobaan dengan mengkompresi dan mendekompresi data teks yang berisi karakter dengan jumlah sedikit dan banyak dalam bentuk sebuah teks berisi dengan jumlah karakter yang pendek dan sebuah jumlah karakter yang lebih panjang.

1.6 Sistematika Penulisan

Agar laporan tugas akhir ini dapat dibaca dengan susunan yang baik dan tertata rapi, penulis menguraikan isi laporan dalam sistematika penulisan berikut ini. Sistematika penulisan laporan tugas akhir ini terdiri dari 5 (lima) bab, yaitu:

Bab 1 : PENDAHULUAN

Bab ini berisi mengenai latar belakang masalah yang akan dibahas, perumusan masalah, tujuan penelitian, batasan masalah, metodologi dan sistematika penulisan.

Bab 2 : TINJAUAN PUSTAKA

Bab ini berisi tentang teori-teori yang mendukung program yang dibuat terutama teori mengenai metode Arithmetic Coding dan LZ77 beserta dengan cara penyelesaiannya.

Bab 3 : PERANCANGAN SISTEM

Bab ini membahas mengenai perancangan program pengkompresian dan pendekompresian. Perancangan meliputi rancangan program yang dibuat secara terperinci dengan membahas mengenai input dan output.

Bab 4 : IMPLEMENTASI DAN ANALISIS SISTEM

Bab ini berisi mengenai implementasi program yang meliputi percobaan program dengan studi kasus dan analisis program.

Bab 5 : KESIMPULAN DAN SARAN

Bab ini berisi mengenai kesimpulan dari hasil pembuatan program dan analisis yang didapat serta saran-saran untuk pengembangan program lebih lanjut.

©UKDW

BAB 5

KESIMPULAN DAN SARAN

5.1. Kesimpulan

Berdasarkan hasil implementasi dan analisis sistem dengan menggunakan metode Arithmetic Coding dan LZ77 dalam pengkompresian file teks, maka diperoleh kesimpulan sebagai berikut :

1. Arithmetic Coding efektif mengkompres data teks dengan probabilitas kemunculan yang besar (kemunculan hurufnya banyak).
2. LZ77 tidak cocok diterapkan pada file dengan teks berupa kalimat dengan bahasa yang terstruktur atau dengan huruf yang munculnya secara acak (tidak berpola).
3. Semakin besar *search buffer* dan *look-ahead buffer* pada algoritma LZ77 maka keberhasilan kompresi yang didapat akan semakin baik (file hasil menjadi lebih kecil).

5.2. Saran

Beberapa saran yang dapat dirumuskan :

1. Algoritma Arithmetic Coding sebaiknya menggunakan batas nilai integer agar range yang bisa dikerjakan semakin banyak.
2. Untuk metode LZ77, Inputan huruf terlebih dahulu ke dalam bentuk basis yang lebih kecil, misalnya heksadesimal lalu di kompresi menggunakan LZ77.

DAFTAR PUSTAKA

- Alam, M. A. J. (2001). *Belajar sendiri borland delphi 6.0*. Jakarta: PT. Elex Media Komputindo. Kelompok Gramedia.
- Bib, P. S., Bukhari, F., & Solikha, N. (2006). Arithmetic coding for text compression. *4*(1), 1-10.
- Floyd, T. L. (2003). *Digital fundamentals with vhdl*. New Jersey: Prentice Hall. Pearson Education, inc.
- Gozali, F., & , M. (2004). Analisis perbandingan kompresi data dengan teknik arithmetic coding dan run length encoding. *4*(1), 37-52.
- Kandaga, T. (2006). Analisis penerapan kompresi dan dekompresi data dengan menggunakan metode statistik dan kamus. *Jurnal Informatika*, *2*(2), 81-91.
- Martha, B. B., & Venkataramani, Y. (2011). An efficient text compression for massive volume of data. *International Journal of Computer Applications*, *21*(5), 6-9.
- Martina, I. (2000). *36 jam belajar komputer delphi 5.0*. Jakarta: PT. Elex Media Komputindo. Kelompok Gramedia.
- Nelson, M., & Gailly, J. (1996). *The data compression book*. (2nd ed.). New York: M&T Books.
- Paul G, H., & Vitter, J. (1992). Analysis of arithmetic coding for data compression. *information processing & management* . *28*(6), 749-763.
- Platoš, J., & Dvorský, J. (2011). Word-based text compression. *Department of Computer Science VŠB – Technical University of Ostrava, Czech Republic*, *5*(1), 97–105.
- Raharjo, B. (2008). *Teknik pemrograman pascal*. (2 ed.). Bandung: Penerbit Informatika.

Salomon, D. (2007). *Data compression – the complete reference*. (4 ed.). London: Springer-Verlag London Limited.

Santoso, P. (2001). Studi kompresi data dengan metode arithmetic coding. *Jurnal Teknik Elektro*, 1(1), 14-18.

Sayood, K. (2006). *Introduction to data compression*. (3 ed.). San Francisco: Morgan Kaufmann Publishers.

©UKDW