

PAPER NAME

lipi8.doc

WORD COUNT

2561 Words

CHARACTER COUNT

15830 Characters

PAGE COUNT

7 Pages

FILE SIZE

139.0KB

SUBMISSION DATE

Sep 25, 2022 8:28 PM GMT+7

REPORT DATE

Sep 25, 2022 8:28 PM GMT+7

● 15% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

- 15% Internet database
- 5% Publications database

● Excluded from Similarity Report

- Crossref database
- Submitted Works database
- Crossref Posted Content database
- Manually excluded text blocks

Pengkode Sinyal Suara Pada Laju 4 kbps : Menggunakan Model Sinusoida – Segmen Antar Puncak

Suhartono Tjondronegoro¹, Florentinus Budi Setiawan^{1,2}

¹Laboratorium Telekomunikasi radio dan Gelombang Mikro, STEI - ITB,
Jl. Ganesha 10 Bandung 40132, Indonesia

²Universitas Katolik Soegijapranata, Semarang, Indonesia
fbudisetiawan@yahoo.com

Abstract

Communications equipment needs to work at low bit rate, that can reduce the need of transmission bandwidth. The proposed speech coder generates coded signal at the rate of 4 kbps with low complexity. Thus, the transmission channel can be used for great number of communications connection. The low bit rate speech coder can be realized by using segmental sinusoidal model. By using this model, signal parameters are used to generate the synthetic speech signal.. In this paper, we proposed a sinusoidal model by using the peak of signal amplitude, called as segmental sinusoidal model. The proposed coder is combined with waveform interpolation model and codebook, thus it can work at low bit rate. The coder has MOS score of 3.8 (out of 5). It means that the perception quality is fairly good. The output data rate of encoder is 4 kbps or bellow, with complexity less than 10 MIPS(Million Instruction per Second).

Keywords : analysis, frequency, interpolation, peak, period, segmental, sinusoidal, synthesis

Abstrak

Perangkat komunikasi dituntut untuk dapat bekerja pada laju yang rendah untuk menurunkan penggunaan lebar pita transmisi. Pengkode sinyal usulan memiliki laju 4 kbps dengan kompleksitas rendah. Dengan demikian, kanal transmisi dapat dipakai untuk lebih banyak koneksi komunikasi. Pengkode sinyal suara dengan laju rendah dan kompleksitas rendah dapat direalisasi dengan model sinusoida secara segmental. Dengan menggunakan model ini, parameter sinyal berupa informasi periode dan informasi puncak dipakai untuk membangkitkan sinyal sintesis. Pada tulisan ini diusulkan sebuah model sinusoida menggunakan puncak amplituda sinyal, yang disebut sebagai model sinusoidal segmental. Pengkode tersebut dikombinasi dengan model interpolasi gelombang dan penggunaan bukukode untuk mendapatkan laju rendah. Pengkode memiliki skor MOS (Mean Opnion Score) sebesar 3,8 (dari 5) yang berarti bahwa kualitas sinyal cukup baik. Laju data keluaran pengkode adalah 4 kbps dengan kompleksitas kurang dari 10 juta instruksi per detik (MIPS,Million Instruction per Second) .

Keywords : analisis, frekuensi, interpolasi, periode, puncak, segmental, sinusoida, sintesis

1. Pendahuluan

Perangkat sistem komunikasi pada saat ini dituntut untuk dapat bekerja pada laju data dan kompleksitas rendah. Hal tersebut harus dipenuhi, mengingat jumlah kanal yang terbatas, namun terdapat banyak permintaan untuk melakukan komunikasi.

Pengkode usulan bekerja pada laju rendah, yaitu 4 kbps dengan kompleksitas yang cukup rendah dibandingkan dengan pengkode sejenis. Model sinusoida secara segmental dipakai sebagai dasar untuk membangun sistem pengkode sinyal suara pada laju rendah. Model sinusoida diterapkan dengan

berdasarkan asumsi bahwa sinyal suara memiliki karakteristik yang hampir periodik. Model tersebut mampu menjaga kualitas persepsi, khususnya dalam menjaga periodisitas sinyal sintesis yang dibangkitkan pada dekoder. Sinyal suara dapat dinyatakan sebagai kombinasi sinyal sinusoida dengan variasi amplituda, fasa dan frekuensi. Untuk model kuantisasi berdasarkan puncak ke puncak, puncak positif dan negatif yang berurutan dideteksi terlebih dahulu. Selanjutnya jarak waktu diantara puncak dikuantisasi. Dalam tulisan ini dijelaskan model baru usulan yang dibagi dalam segmen antar puncak berdasarkan model sinusoida.

Bagian sinyal diantara puncak positif dan negatif yang berurutan atau sebaliknya diestimasi dengan membangkitkan setengah periode sinyal sinusoida. Selisih level antar puncak yang berurutan diasumsikan sebagai dua kali amplituda cosinus.

Sinyal suara dapat dimodelkan dalam bentuk sinyal sinusoida [1][2] untuk sepanjang frame dengan lebar 15 ms sampai dengan 30 ms. Komponen sinyal sinusoida diambil dari parameter sinyal sinusoida [3] untuk dikirimkan ke penerima. Jumlah komponen sinusoida berkisar antara 40 sampai dengan 60 untuk membangkitkan sinyal sintesis. Model sinyal usulan, adalah pemodelan berbasis model sinusoidal secara segmental [4]. Sinyal diambil pada segmen yang lebarnya bervariasi tergantung di jarak waktu diantara puncak yang berurutan. Puncak mengandung arti sebagai nilai maksimum dan nilai minimum yang berfluktuasi sepanjang frame. Satu segmen mengandung arti bagian sinyal diantara puncak maksimum dan minimum yang berurutan atau bagian sinyal diantara puncak minimum dan maksimum yang berurutan. Satu segmen sinyal diantara puncak maksimum menuju puncak minimum berikutnya dapat dimodelkan sebagai setengah periode sinyal cosinus dari $\omega=0$ sampai dengan $\omega=\pi$. Sedangkan segmen sinyal diantara puncak minimum menuju ke puncak maksimum dapat dimodelkan sebagai setengah periode sinyal cosinus dari $\omega=\pi$ sampai dengan $\omega=2\pi$. Komponen sinyal sebanyak k dengan daya yang terbesar digunakan untuk mensintesis sinyal sepanjang frame. Sinyal rekonstruksi akan sama dengan sinyal asli jika terdapat sejumlah tak terhingga komponen sinyal k . Semakin banyak jumlah k , maka sinyal rekonstruksi akan semakin akurat.

Tulisan ini disusun dalam lima bagian. Setelah bagian pendahuluan ini adalah bagian kedua yang menjelaskan mengenai model sinusoida. Bagian berikutnya adalah perancangan enkoder dan dekoder berdasarkan model sinusoida. Bagian keempat

adalah mengenai hasil percobaan beserta dengan pembahasannya. Sedangkan yang terakhir adalah kesimpulan dari tulisan ini.

2. Model Sinusoida untuk Sinyal Suara

Model sinyal sinusoida untuk sinyal suara dapat dipakai untuk pengkodean sinyal suara seperti pada *Sinusoidal Transform Coding* (STC)[5][6][7]. Proses pada *sinusoidal transform coding* adalah mengambil beberapa komponen sinusoida yang memiliki amplituda terbesar. Jumlah komponen sinusoida yang dibutuhkan adalah berkisar antara 40 sampai dengan 60.

Dengan menggunakan model sinusoida secara segmental, dua sinyal berupa komponen DC dan frekuensi fundamental digunakan untuk melakukan estimasi sinyal. Satu segmen berarti bagian sinyal dari puncak maksimum menuju ke puncak minimum, atau sebaliknya. Jarak waktu diantara puncak maksimum ke- i disebut sebagai informasi periode dan dilambangkan sebagai $p_d(i)$. Puncak maksimum dan puncak minimum disebut sebagai informasi puncak dan dilambangkan sebagai $p_k(i)$. Informasi puncak diperoleh dengan mendeteksi puncak maksimum dan puncak minimum sepanjang frame. Informasi periode diperoleh dengan menghitung jarak waktu diantara puncak yang berurutan

Tahap pertama untuk memperoleh parameter sinyal adalah menetapkan satu frame dengan lebar 30 ms. Tahap berikutnya adalah menandai puncak sinyal, baik untuk puncak maksimum dan puncak minimum. Jarak waktu antara puncak ke- i menuju puncak ke- $(i+1)$ adalah sama dengan setengah periode sinyal estimasi, $p_d(i)$. Proses diteruskan untuk puncak berikutnya, sehingga deretan informasi periode dan informasi puncak didapatkan.

Sinyal rekonstruksi berdasarkan informasi periode dan puncak untuk segmen ke- i dari puncak maksimum menuju ke puncak minimum dapat dinyatakan sebagai :

$$s_{pv}(i,n) = \frac{p_k(i) + p_k(i+1)}{2} + \frac{p_k(i) - p_k(i+1)}{2} \cos\left(\frac{\pi \cdot (n - n_k(i))}{p_d(i)}\right) \quad (1)$$

Jika $p_k(0) > p_k(1)$, $i=0,2,4...(I-2)$ jika I genap
 $i=0,2,4...(I-1)$ jika I ganjil
 Jika $p_k(0) < p_k(1)$, $i=1,3,5...(I-1)$ jika I genap
 $i=1,3,5...(I-2)$ jika I ganjil

Dimana $n_k(i)$ adalah lokasi puncak $p_k(i)$. Koefisien komponen sinusoida untuk sinyal estimasi dari puncak minimum $p_k(i)$ menuju ke puncak maksimum berikutnya $p_k(i+1)$ dapat dinyatakan sebagai :

$$s_{vp}(i,n) = \frac{p_k(i) + p_k(i+1)}{2} + \frac{p_k(i+1) - p_k(i)}{2} \cos\left(\frac{\pi \cdot (n - n_k(i))}{p_d(i)}\right) \quad (2)$$

Jika $\cos(A) = -\cos(A+\pi)$, maka persamaan tersebut dapat dinyatakan sebagai :

$$s_{vp}(i,n) = \frac{p_k(i) + p_k(i+1)}{2} + \frac{p_k(i) - p_k(i+1)}{2} \cos\left(\frac{\pi \cdot (n - n_k(i))}{p_d(i)} + \pi\right) \quad (3)$$

Jika $p_k(0) > p_k(1)$, $i=1,3,5...(I-1)$ jika I genap
 $i=1,3,5...(I-2)$ jika I ganjil
 Jika $p_k(0) < p_k(1)$, $i=0,2,4...(I-2)$ jika I genap
 $i=0,2,4...(I-1)$ jika I ganjil

Sinyal estimasi sepanjang frame adalah deretan sinyal s_{pv} and s_{vp} untuk $i=0$ sampai dengan $i=I-1$. Berdasarkan penjelasan di atas, untuk $p_k(0) > p_k(1)$, sinyal rekonstruksi menggunakan model sinusoidal secara segmental dapat ditulis sebagai :

$$s_r(i,n) = \frac{p_k(i) + p_k(i+1)}{2} + \frac{p_k(i) - p_k(i+1)}{2} \cos\left(\frac{\pi \cdot (n - n_k(i))}{p_d(i)} + i\pi\right) \quad (4)$$

untuk $i = 0,1,2 \dots (I-1)$

Jika $p_k(0) < p_k(1)$, sinyal rekonstruksi dapat dinyatakan sebagai :

$$s_r(i,n) = \frac{p_k(i) + p_k(i+1)}{2} + \frac{p_k(i) - p_k(i+1)}{2} \cos\left(\frac{\pi \cdot (n - n_k(i))}{p_d(i)} + (i+1)\pi\right) \quad (5)$$

untuk $i = 0,1,2 \dots (I-1)$

3. Pengkode Sinyal Suara pada laju 4 kbps

Enkoder sinyal suara pada laju kurang dari atau sama dengan 4 kbps dirancang dalam beberapa blok dan algoritma. Bagian depan

dari enkoder terdiri atas detektor sinyal dan detektor pitch. Blok berikutnya adalah proses klasifikasi sinyal bergetar dan sinyal tidak bergetar, pengkode dengan model sinusoida secara segmental, dan pengkode formant. Mode operasi enkoder dibedakan dalam dua jenis proses untuk mendapatkan kualitas persepsi yang cukup tinggi [8][9]. Mode operasi terdiri atas : mode operasi pada keadaan tidak ada sinyal dan mode operasi jika terdeteksi ada sinyal. Mode operasi pada keadaan tidak ada sinyal adalah proses mengirimkan informasi ke dekoder bahwa dalam frame tersebut tidak ada sinyal. Mode operasi pada keadaan ada sinyal dibagi dalam dua jenis, yaitu proses untuk sinyal bergetar dan sinyal tidak bergetar. Masukan adalah sinyal suara yang dikuantisasi dalam format PCM linier 16 bit pada frekuensi 8 kHz.

Bagian awal dari pengkode adalah buffer sinyal untuk sinyal selebar 30 ms. Proses dilakukan dengan menerapkan jendela rektanguler. Selanjutnya sinyal masukan diklasifikasikan sebagai sinyal hening atau kondisi ada sinyal. Jika dinyatakan ada sinyal, maka proses berikutnya adalah deteksi periode pitch. Berdasarkan informasi periode pitch, sinyal dapat diklasifikasikan sebagai sinyal bergetar atau sinyal tidak bergetar. Jika periode pitch kurang dari 160 cuplikan, maka sinyal dinyatakan sebagai sinyal bergetar. Jika periode pitch 160 cuplikan atau lebih disebut sebagai sinyal tidak bergetar. Proses selanjutnya tergantung dari jenis sinyal. Untuk sinyal bergetar, sinyal karakteristik [10][11] diambil berdasarkan lebar pitch yang terdeteksi. Sinyal karakteristik dikuantisasi dengan model sinusoida secara segmental. Informasi formant untuk tiap frame dikirimkan ke dekoder untuk mengatur frekuensi tengah postfilter. Blok berikutnya adalah penelusuran indeks bukode perioda, puncak dan formant. Semua parameter sinyal terkode dikirimkan ke dekoder dalam laju kurang dari atau sama dengan 4 kbps.

Pitch adalah parameter yang penting dalam proses encoding. Berdasarkan nilai perioda pitch sinyal dapat diidentifikasi sebagai sinyal bergetar ataupun sinyal tidak bergetar.

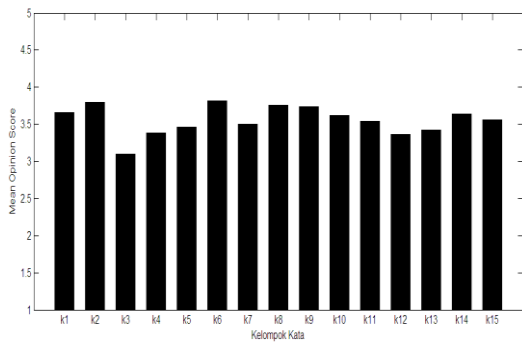
Periode pitch sinyal bergetar bervariasi antara 2.5 ms sampai dengan 20 ms. Periode pitch dideteksi dengan menggunakan proses otokorelasi. Pada tahap pertama, sinyal dideteksi nilai maksimumnya. Berdasarkan nilai maksimum tersebut dapat ditentukan nilai ambang untuk proses pemotongan tengah. Nilai ambang adalah sebesar setengah dari nilai sinyal maksimum yang ada dalam frame. Selanjutnya sinyal dalam frame dikurangi dengan nilai ambang tersebut. Jika nilai sinyal hasil pemotongan kurang dari nol, maka sinyal tersebut diberi nilai nol. Hal ini dimaksudkan untuk menurunkan kompleksitas proses perhitungan otokorelasi. Proses otokorelasi menghasilkan dua jenis pola sinyal, yaitu pola puncak-lembah-puncak dan pola puncak-lembah.

Sinyal bergetar dan sinyal tidak bergetar diklasifikasikan berdasarkan hasil proses deteksi periode pitch. Hasil proses otokorelasi berupa periode pitch dipakai sebagai acuan untuk menentukan jenis sinyal adalah sinyal bergetar atau sinyal tidak bergetar. Jika pola yang dihasilkan berbentuk puncak-lembah-puncak dan jika jarak diantara puncak lebih dari 2.5 ms dan kurang 20 ms dinyatakan sebagai sinyal bergetar. Jika pola yang dihasilkan adalah puncak-lembah, dan jarak antara puncak lebih dari 20 ms dinyatakan sebagai sinyal tidak bergetar.

Sinyal bergetar diambil selebar satu periode pitch, yang disebut sebagai sinyal

karakteristik, untuk diproses dengan kuantisasi sinusoda secara segmental. Sinyal karakteristik dipakai untuk membangkitkan sinyal bergetar pada periode berikutnya, sampai frame selebar 30 ms terpenuhi oleh sinyal [10][11]. Panjang sinyal karakteristik adalah selebar satu periode pitch. Tahap selanjutnya adalah kuantisasi sinusoida secara segmental. Untuk sinyal tidak bergetar, dilakukan proses desimasi untuk mendapatkan jumlah informasi periode dan puncak yang lebih kecil dari 40. Sinyal estimasi dibangkitkan berdasarkan informasi periode dan puncak.

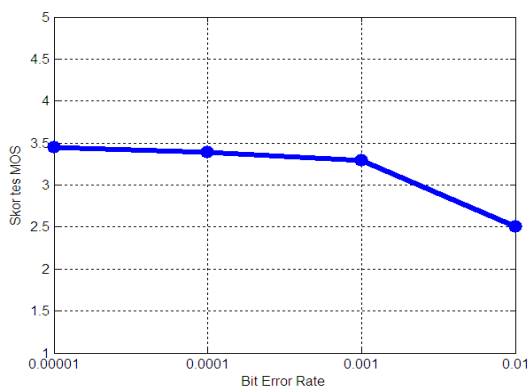
Ukuran informasi puncak dikurangi dengan menggunakan sepuluh buah bukode. Bukode dibangkitkan dengan algoritma k-means berdasarkan vektorkode training sehingga diperoleh bukode puncak. Jumlah indeks bukode bervariasi antara 6 sampai dengan 10 untuk mendapatkan hasil proses pengkodean yang optimum. Ukuran informasi periode juga dikurangi dengan menggunakan bukode. Indeks bukode untuk informasi puncak bervariasi antara 6 sampai dengan 10, seperti pada bukode puncak. Tingkat akurasi periode harus dijaga untuk mendapatkan persepsi penerima yang baik. Informasi formant sangat berguna pada proses filtering sinyal bergetar yang membangkitkan dekoder.



Gambar 3. Hasil tes MOS

Pengkode sinyal suara disimulasikan pada komputer dengan menggunakan program *coder 4 kbps* yang dibangun dengan menggunakan bahasa pemrograman C++. Hasil tes MOS (mean opinion score) untuk 15 buah deretan kata dalam bahasa Indonesia adalah 3,8 (dari 5). Tes dilakukan pada 46 orang dengan variasi jenis kelamin, latar belakang dan umur. Kompleksitas pengkode kurang dari 10 MIPS, dibanding jenis pengkode yang lain yang memiliki kompleksitas antara 0.01 MIPS sampai dengan 90 MIPS. Jumlah ruang memori yang dibutuhkan adalah kurang dari 16 kB.

Jika sinyal terkode dicampur dengan derau dengan tingkat kesalahan bit (*bit error rate*, BER) 10^{-5} , 10^{-4} , 10^{-3} dan 10^{-2} maka terjadi degradasi hasil tes MOS seiring dengan meningkatnya BER pada sisi dekoder.



Gambar 4. Nilai tes MOS pada BER tertentu
Tabel 1. Kinerja pengkode sinyal suara [12]

NO	NAMA PENGKODE	Laju (kbps)	MOS
1	INMARSAT	4,15	3,2
2	CELP (U.S. FS 1016)	4,8	3,2
3	QCELP (US CDMA)	1 – 8	3,4

4	VSELP (Seluler Jepang)	6,8	3,3
5	G.723.1 (H.323 dan H.324)	6,3	3,98
6	GSM Half Rate	5 – 6	3,4

Dibandingkan dengan beberapa pengkode sinyal suara pada laju di sekitar 4 kbps, pengkode usulan menghasilkan kualitas yang lebih baik.

Tabel 2. Kompleksitas pengkode suara[12]

NO	NAMA PENGKODE	Laju (kbps)	MIPS
1	PCM (G.711)	64	0,01
2	ADPCM (G.726)	16-40	~2
3	LD-CELP (G.728)	16	~19
4	RPE-LTP (GSM)	13	6
5	VSELP (IS-54)	8	13,5
6	CELP (DoD1016)	4,8	16
7	LPC-10e (DoD1015)	2,4	7
8	CS-ACELP (G.729)	8	17
9	CS-ACELP (G.723)	5,3/6,3	14,6/16
10	MELP	2,4	20

5. Kesimpulan

Sinyal suara dapat dikodekan dalam laju 4 kbps dan dapat direkonstruksi dengan kualitas yang cukup baik berdasarkan model sinusoida secara segmental. Hasil tes MOS adalah 3,8 untuk 15 deretan kata dalam bahasa Indonesia. Kompleksitas pengkode cukup rendah, yaitu kurang dari 10 MIPS. Hasil tersebut lebih baik daripada pengkode sinyal suara yang lain pada laju lebih besar atau sama dengan 4 kbps.

Daftar Pustaka

- [1] T.F. Quatieri, and R.J. McAulay, "Speech Transformation Based on a Sinusoidal Representation", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-34, 1986, pp. 1449-1464.
- [2] R.J. McAulay, and T.F. Quatieri, "Speech Analysis/Synthesis Based on a Sinusoidal Representation", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-34, 1986, pp. 744-754.
- [3] M. Lagrange, S. Marchand, and J.B. Rault, "Sinusoidal Parameter Extraction and Component Selection in A Non Stationary Model", *Proceedings of the 5th International Conference on Digital Audio Effects*, 2002, pp. 59-64.

- [4] FB Setiawan, S. Tjondronegoro, "Sinusoidal Model of the Speech Signal", Yogyakarta : Proceedings of National Seminar UTU, 2005.
- [5] S. Ahmadi, dan A.S. Spanias, "A New Phase Model for Sinusoidal Transform Coding of Speech", *IEEE Transactions on Speech and Audio Processing*, 6, 1998, pp. 495-501.
- [6] T. Abe, and M. Honda, "Sinusoidal Model Based On Instantaneous Frequency Attractor", *IEEE Transaction on Speech, Audio and Language Processing*, 14, 2006, pp. 1292-1300.
- [7] J. Epps, and W.H. Holmes, Speech Enhancement Using STC-based Bandwidth Extension, *Proceedings of 5th International Conference on Spoken Language Processing*, 1998, pp. 519-522.
- [8] B.S. Atal, V. Cuperman, and A. Gersho, *Advances in Speech Coding*, Massachusetts: Kluwer Academic Publishers, 1991.
- [9] A.M. Kondoz, : *Digital Speech : Coding for Low Bit Rate Communications Systems*, West Sussex, England : John Wiley & Sons Ltd, 1995.
- [10] O. Gottesman, and A. Gersho, "Enhanced Waveform Interpolative Coding at Low Bit-Rate", *IEEE Transactions on Speech and Audio Processing*, 9, 2001, pp. 1-13.
- [11] U. Sinervo , *Waveform Interpolation Speech Coding at 2.4-4.0 kb/s*, Master of Science Thesis, Tampere University of Technology, Finland, 2000.
- [12] N. Peleg, : *Introduction to Speech Coding*, The Signal and Image Processing Laboratory, Technion Electrical Engineering Department, Israel, 2005. <http://cs.haifa.ac.il/~nimrod/Compression/Speech/s3coding2004.pdf>.

● 15% Overall Similarity

Top sources found in the following databases:

- 15% Internet database
- 5% Publications database

TOP SOURCES

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	sintak.unika.ac.id	Internet	4%
2	citee2015.jteti.ft.ugm.ac.id	Internet	2%
3	hakenaudio.com	Internet	<1%
4	citeseerx.ist.psu.edu	Internet	<1%
5	mafiadoc.com	Internet	<1%
6	journal.ui.ac.id	Internet	<1%
7	premieraward.org	Internet	<1%
8	epubs.surrey.ac.uk	Internet	<1%
9	edocs.tu-berlin.de	Internet	<1%

10	pastel.archives-ouvertes.fr Internet	<1%
11	eet.unsw.edu.au Internet	<1%
12	biblio.ugent.be Internet	<1%
13	math.zju.edu.cn Internet	<1%
14	repository.lib.cuhk.edu.hk Internet	<1%
15	dspace.cc.tut.fi Internet	<1%
16	zeidler.ucsd.edu Internet	<1%
17	biblioteket.ehl.lu.se Internet	<1%

● Excluded from Similarity Report

- Crossref database
- Submitted Works database
- Crossref Posted Content database
- Manually excluded text blocks

EXCLUDED TEXT BLOCKS

Pengkode Sinyal Suara Pada Laju 4 kbps :Menggunakan Model Sinusoida - Segme...

sintak.unika.ac.id

Indonesiafbudisetiawan@yahoo.comAbstractCommunications equipment needs t...

citee2015.jteti.ft.ugm.ac.id

puncak maksimum menuju puncak minimumberikutnya dapat

sintak.unika.ac.id