

Pembuatan Aplikasi Pendeteksi Nada Menggunakan IIR Comb Filter dan Peak2 Onset Detection

Klemens¹, Liliana², Gregorius Satia Budhi³

Program Studi Teknik Informatika, Fakultas Teknologi Industri, Universitas Kristen Petra

Jl Siwalankerto 121 – 131, Surabaya 60326

Telp. (031) - 2983455, Fax. (031) - 8417658

E-mail: aequalb@gmail.com, lilian@peter.petra.ac.id, greg@peter.petra.ac.id

ABSTRACT: Transkripsi musik adalah sebuah bidang riset yang aktif dalam *computer science*. Pemrosesan menggunakan filter sudah umum dipakai untuk memperbaiki kualitas audio maupun untuk menambahkan efek suara. Aplikasi *pitch detection* pada umumnya melakukan pemrosesan tambahan untuk menghilangkan nada dominan yang ditemukan. Menggunakan *IIR comb filtering* bukan sebagai algoritma penunjang namun sebagai algoritma untuk menentukan nada belum dieksplorasi dengan baik. Skripsi ini bertujuan untuk menguji performanya terhadap instrumen gitar dan piano. Metode yang dipilih adalah berdasarkan *IIR Comb Filter* untuk *pitch detection* dan *Peak2* untuk *onset detection*. Filter ini mengakibatkan hilangnya komponen frekuensi nada dominan tersebut, dan hasil filter yang sesuai dengan nada dasarnya akan memiliki nilai lebih kecil daripada hasil filter lainnya. *Peak2 onset detection* bekerja dengan dua *peak detector*. *Peak detector* yang pertama digunakan untuk mendeteksi perubahan yang cepat pada nilai amplitudo lokal. Untuk memperoleh *onset*, *peak detector* diulangi sekali lagi terhadap hasil *peak detector* pertama. Hasil pengujian menunjukkan hasil yang kurang untuk pengolahan suara instrumen gitar dan piano, sekitar 55% dari semua *onset* yang benar berhasil dideteksi oleh aplikasi. Sedangkan untuk pendeteksi nada, nada monofonik dideteksi dengan benar 81.25% dari 48 nada yang diujikan dan untuk nada polifonik 64.29% dari 84 nada yang diujikan.

Kata Kunci : *Pitch Estimation, Digital Signal Processing, Onset Detection, Transcription.*

ABSTRACT: Transcription of music is an active research problem in computer science and has many applications in *digital music*. The use of time-domain filter to improve sound quality and adding sound effects is commonplace, but the use of filter for pitch detection is rarely explored. This work thesis aims to create a pitch detection system based on digital filter and test it's performance in guitar and piano signals. IIR Comb Filter and Peak2 onset detection is chosen as the main algorithm for this application. IIR Comb Filter have notches, zero values corresponding to a certain frequency and all it's harmonic frequencies. If make IIR Comb Filter correspond to a pitch frequency, all its harmonic frequency will have zero values. Filtering this kind of filter with a pitched signal will result in elemination of a pitch frequency and all it's harmonics, and the filter corresponding to the input frequency will have a minimum output compared to the other filter's result. Peak2 onset detection works by taking the result of two peak detectors. The first peak detector detects sudden changes in amplitude in time-domain, the second peak detector takes the first peak detector's values as an input, producing a smother result. Test showed that the result

varied between instruments. Onset detection for guitar and piano performs rather poorly, about 55% of all the correct onsets were detected. For pitch estimation, the accuracy for the pitch detected is 81.25% for all monophonic signals and 64.29% for all polyphonic signals.

Keyword : *Pitch Estimation, Digital Signal Processing, Onset Detection, Transcription.*

1. Latar Belakang

Musik terkait erat dengan persepsi manusia. Musik dari jaman dahulu kala telah digunakan untuk menyampaikan informasi, memberikan kesenangan bagi orang lain, dan menyimpan informasi dari jaman yang lampau. Sekarang, dengan adanya internet, musik telah menjadi sebuah kebutuhan dan industri musik juga berkembang dengan pesat. Keberanian musisi baru untuk mengeksplorasi gaya seni yang baru menimbulkan kesulitan dalam pengelompokan dan pencarian musik. Kebutuhan akan pencarian informasi dengan lebih praktis menciptakan bidang khusus yaitu MIR. Riset [1] oleh LA Smith dkk. menciptakan sebuah metode yang menggunakan transkripsi lagu untuk mencari informasi musiknya. Transkripsi digunakan untuk menemukan nada dan ritme pada yang digunakan untuk menemukan lagu atau bagian lagu dari database yang ada.

Pitch, atau nada adalah sebuah faktor subjektif yang melibatkan persepsi seseorang terhadap frekuensi (atau frekuensi – frekuensi) yang menyusun sebuah dalam musik. Nada adalah salah satu identitas dari lagu, bersama dengan ritme. Dengan adanya ritme sebuah lagu dapat dibedakan dengan *noise*, karena *noise* tidak memiliki informasi *musical event* seperti halnya nada dan ketukan.

Riset tentang *pitch* estimation telah sangat banyak dilakukan. Riset [2] oleh Anssi Klapuri memunjukkan perbandingan antar metode ini untuk keperluan transkripsi musik. Penggunaan metode autocorrelation dan AMDF sebagai salah satu metode pencarian pitch telah diteliti dalam [3] dan [4]. Kedua metode ini mendasarkan pencarian nada berdasarkan perulangan sinyal pada time-domain. Pendekatan yang menggunakan analisa spektrum seperti pada [5] yang menggunakan kecocokan model nada piano untuk mencari nada. Ada pula yang meneliti perulangan dalam spektrometri dan menggunakannya untuk menganalisa nada. Vincent et al. [6] menggunakan harmonic spectral decomposition untuk menanalisa banyak nada. Seperti metode analisa spectrum, penggunaan Comb Filter untuk pendeteksi nada memanfaatkan perulangan yang terjadi akibat adanya frekuensi harmonik yang terjadi pada setiap instrument. Metode ini telah diteliti oleh beberapa sumber seperti [7] dan [8].

Onset adalah saat dimana sebuah (atau beberapa) nada mulai dimainkan, ditandai dengan perubahan dari keadaan tanpa nada menjadi “bernada”. Bernada disini dimaksudkan dengan nada yang terdengar kontras dan berbeda dengan suara latar, yang biasa disebut dengan nada dominan. Onset – onset ini menandakan mulainya sebuah *musical event*, dimana dalam sebuah musical event terdapat nada. Fenomena perubahan ini dapat dideteksi, dan pendeteksian ini berpengaruh pada *content delivery*, *compression*, *indexing*, dan *retrieval*. Segmentasi yang akurat membuka peluang bagi analisa suara digital yang lebih tidak rentan terhadap perubahan gaya bermusik dan juga menjadikan MIR lebih kuat terhadap *noise*.

Ada banyak cara untuk mendeteksi onset. Hal inilah yang diplotitasi oleh Kapil dalam karyanya algoritma Peak2 [9]. Peak2 cukup baik dalam mendeteksi nada dari dua instrumen yang digunakan, yaitu piano dan gitar. Implementasi aplikasi ini ditulis dalam C.

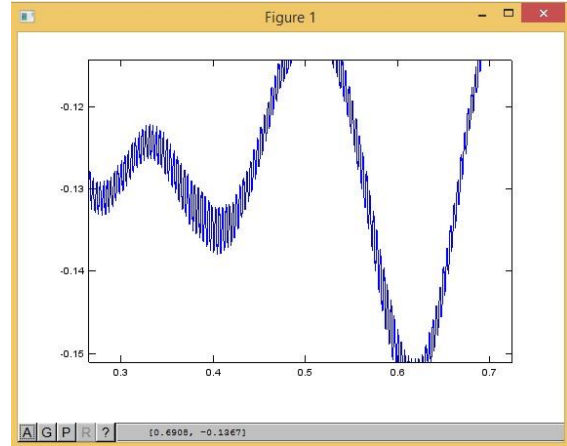
2. Landasan Teori

Metode deteksi *onset* yang adalah *time-domain based*, dengan menganalisa perubahan amplitudo yang ada. Hal ini diharapkan cukup untuk mendapatkan hasil yang cukup baik untuk diterapkan pada *pitch detection*. Metode yang digunakan adalah *peak2* untuk *onset detection* dan *IIR comb filter* untuk *pitch detection*. Pemilihan metode onset didasarkan pada performa pada [9] yang telah ditemukan baik untuk data gitar, sedangkan untuk data piano performanya masih belum teruji.

Pitch detection berdasarkan *IIR comb filter* digunakan karena semua instrumen memiliki frekuensi dasar dan frekuensi harmonik yang merupakan kelipatan dari frekuensi dasar tersebut. Perbedaan instrumen membawa perbedaan pada kekuatan volume frekuensi harmonik, yang mengakibatkan *inharmonicities*, dimana sebuah instrumen menghasilkan frekuensi harmonik lebih kuat daripada frekuensi dasarnya. Hal ini dapat diatasi dengan melakukan penghilangan pada nada dasar dan semua frekuensi harmoniknya.

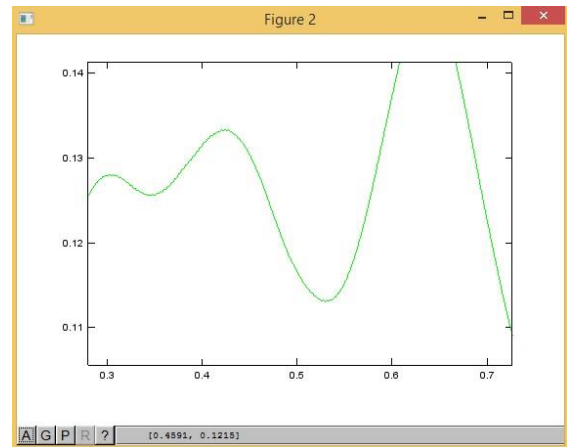
2.1 Algoritma Peak2

Pembuatan algoritma *peak2* didasari oleh bentuk sinyal gitar dan piano yang dapat dideskripsikan dengan *attack time* yang cepat dan *release time* yang lama. *Peak2* berdasarkan metode pada *time-domain* dan mencari perubahan pada amplitudo sinyal input untuk mendeteksi *onset*. Dalam metode ini, digunakan metode *exponential moving average* untuk mencari *peak* pada data. Metode ini dipilih karena kecepatan pemrosesannya yang relatif cepat dan cukup akurat untuk mendeteksi sinyal suara yang memiliki *slope* yang cukup tajam. Berikut ini adalah contoh dari penggunaan *peak2*.



Gambar 1. Gambar hasil peak detection pertama

Pada Gambar 1, diperlihatkan bahwa hasil dari *peak detection* pertama yang memiliki struktur yang berlembah lembah. *Peak detection* yang dideskripsikan pada metode *Peak2* menggunakan 2 kali *peak detection*. Pada hasil *peak detection* yang pertama, hasil yang diperoleh masih belum dapat digunakan sebagai *onset detection* karena masih banyak *local peak* yang perlu dihilangkan. Untuk memperhalus hasil yang diperoleh, *peak detection* yang kedua dilakukan atas data yang diperoleh pada *peak detection* yang pertama. Hasil dari *peak detection* kedua dapat dilihat pada Gambar 2.



Gambar 2. Gambar hasil peak detection kedua

Karena hasil yang lebih terstruktur dan hasil yang didapatkan dianalisa untuk diambil *peak*-nya, dan diidentifikasi waktu dimana *peak* itu berada. Waktu yang diperoleh sedikit bergeser ke arah depan dalam waktu (sedikit “terlambat”) dibandingkan onset asli yang diperoleh.

2.2 Comb Filter

Comb filter adalah filter yang memiliki lembah yang berulang. Bentuk yang demikian diciptakan atas penjumlahan sebuah sinyal dengan sinyal itu sendiri. Comb Filter yang paling standar adalah FIR Comb Filter, yang memiliki *transfer function* [10]

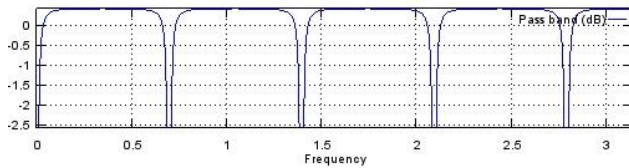
$$H(z) = 1 + \alpha z^{-k} \quad (1)$$

Comb filter memiliki nilai nol pada frekuensi yang dispesifikasi oleh ω , seperti yang terlihat pada rumus (1). Filter ini bersifat menghilangkan frekuensi ω sehingga untuk nilai ω yang merupakan frekuensi dari sebuah nada dasar, nada dasar tersebut akan dihilangkan. Selain itu, comb filter juga mempunyai nilai nol pada semua frekuensi kelipatannya yaitu $\omega, 2\omega, 3\omega \dots$ sehingga semua frekuensi harmoniknya akan juga terpengaruh dan menghilang.

Menurut penelitian pada [7], tingkat kesuksesan filter pada pendeteksian nada dasar mencapai 99.6% pada octave 3 dan 83.6% pada octave 3 sampai 5 pada violin dan clarinet. Penelitian ini diperbaiki oleh penelitian pada [10], yang menggunakan modifikasi dari FIR comb filter seperti yang dijelaskan pada [11]. Dengan menambahkan zero untuk setiap dalam rumus (1) kita memperoleh

$$H(z) = \frac{1-z^{-k}}{1-p^k z^{-k}} \quad (2)$$

Frequency response yang diperoleh dari filter pada (2) memiliki bentuk seperti yang dijelaskan pada Gambar 2. Dapat dilihat bahwa adanya perulangan pada *notch* yang ada sehingga nada dan harmoniknya dapat dihilangkan.

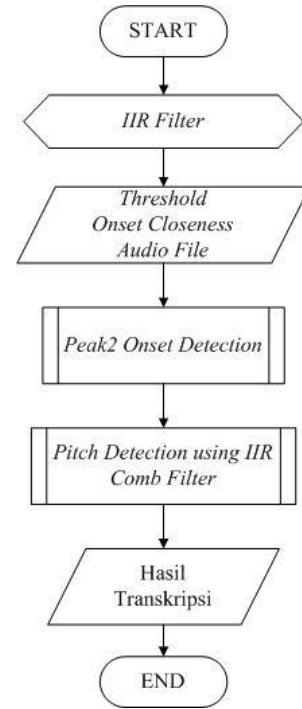


Gambar 3. Gambar Bentuk dari IIR Comb Filter

Paper [6] menuliskan bahwa penggunaan IIR comb filter adalah untuk menghilangkan nada dengan lebih akurat karena bentuk khusus dari IIR comb filter. Bentuk lembah yang lebih singkat ini memberikan sebuah performa pemfilteran yang lebih baik.

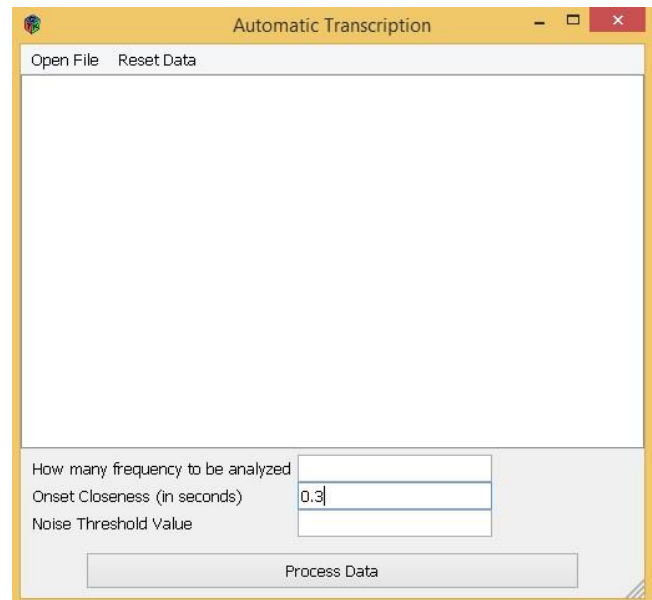
2.3 Desain sistem

Sistem yang digunakan membuat data onset dari algoritma peak2 dan kemudian menggunakan data hasil dari onset itu untuk menentukan nada. Pertama – tama program ini membuat filter yang diperlukan. Kemudian *onset detection* dijalankan atas sinyal suara. Hasil yang didapatkan oleh *onset detection* akan digunakan sebagai penanda waktu untuk *pitch detection*, sehingga bagian *pitch detection* dapat mengetahui waktu yang tepat untuk menjalankan pemfilteran data. Gambar 3. menunjukkan diagram alur garis besar jalannya program

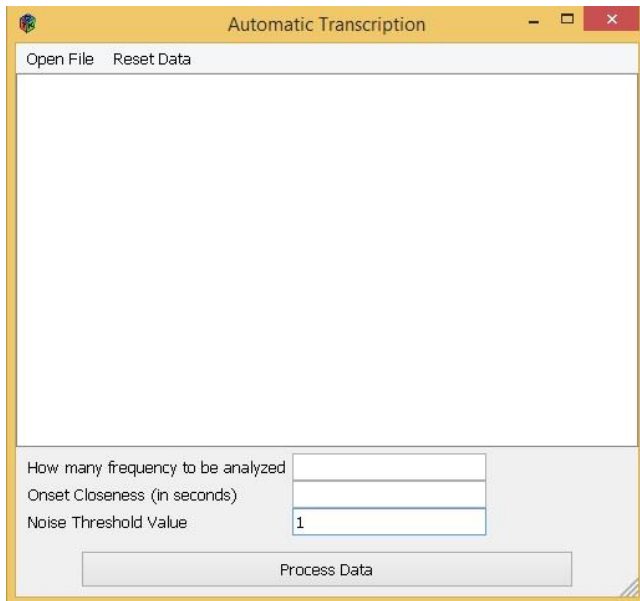


Gambar 4. Diagram Alur Garis Besar Program

Dengan menggunakan data filter yang dibuat sebelum program dimulai, sistem ini menjalankan semua filter untuk masing – masing frekuensi yang digunakan. Dari setiap hasil pemfilteran, hasil yang memiliki amplitudo minimal akan diambil dan kemudian menjadi input untuk iterasi kedua.



Gambar 5. Proses pengaturan *threshold* dan *onset closeness*



Gambar 6. Proses pengaturan *threshold* dan *onset closeness*



Gambar 7. Hasil transkripsi

3. IMPLEMENTASI DAN HASIL

3.1 Pengujian onset terhadap rekaman yang digunakan

Metode Peak2 menggunakan variabel untuk thresholding pada hasil moving average dan *onset closeness* untuk variabel yang menyatakan panjang *release time*. Nilai *threshold* yang digunakan adalah hasil perkalian dari hasil rata – rata dari *exponential moving average* dengan sebuah variabel. Nilai dari variabel ini dites dengan menggunakan suara suara yang sudah disiapkan. Berikut ini adalah nilai hasil dari pengujian pada nilai variabel, dibulatkan.

Variabel *onset closeness* adalah variabel yang menentukan panjang minimum yang dapat dimainkan oleh sebuah instrumen. Hal ini digunakan oleh sebab *pitch detection* memiliki panjang minimum untuk analisisnya. Jika variabel *onset closeness* ini dibuat terlalu dekat, maka selain *pitch detection* menjadi kurang akurat, jumlah onset yang tidak seharusnya ada juga bertambah. Apabila variabel *onset closeness* menjadi terlalu besar, maka banyak *onset* yang akan dilewatkan sebagai *onset* yang salah (*false onset*) dan kemungkinan *onset* yang benar juga akan terlewatkan. Variabel *threshold* akan menentukan tingkat sensitifitas algoritma terhadap *noise* yang ada. Semakin rendah nilai variabel ini semakin sensitif algoritma terhadap *noise* yang ada, dan akan ada banyak *noise* yang dianggap sebagai *onset*. Sebaliknya, nilai *threshold* yang terlalu tinggi akan menghilangkan banyak *onset* yang seharusnya dapat dideteksi.

Data monofonik yang digunakan memiliki jeda antar nada yaitu 0,5 – 1 detik, sedangkan data polifonik yang digunakan memiliki jeda antar nada 1- 1,5 detik. Jeda rata – rata pada data yang digunakan adalah sekitar 0.6 detik. Data monofonik yang digunakan dalam percobaan adalah data yang terbaik yang diperoleh dari rekaman menggunakan rekaman rumah, dengan sampling rate yang berbeda, yaitu 16000Hz untuk gitar dan 44100Hz untuk piano. Hal ini disebabkan karena perbedaan sampling rate berpengaruh pada pendeteksian nada.

Tabel 1. Hasil dari pengujian terhadap nilai variabel

Pengujian (Threshold / Onset Closeness)	Kesalahan Deteksi	Kesalahan (%)	Ketepatan
1x / 0.15	44 / 68	65%	35%
1x / 0.3	21 / 68	31%	69%
1x / 1	24 / 68	35%	65%
1.5x / 0.15	23 / 68	34%	66%
1.5x / 0.3	16 / 68	24%	76%
1.5x / 1	20 / 68	29%	71%
2x / 0.15	27 / 68	40%	60%
2x / 0.3	28 / 68	51%	49%
2x / 1	34 / 68	50%	50%

Didapatkan pada tabel 1 bahwa nilai variabel *threshold* 1.5 dan *onset closeness* sebesar 0.3 adalah nilai yang baik. Tabel ini juga menunjukkan bahwa *onset closeness* juga berpengaruh pada kesalahan deteksi, dengan nilai *onset closeness* yang semakin jauh dari nilai rata – rata rekaman, kesalahan akan semakin meningkat. Kesalahan yang terutama dari *onset closeness* adalah terlalu kecil, yang mengakibatkan aplikasi mendeteksi terlalu banyak nada. Kesalahan yang ada pada Tabel 1 memperlihatkan kesalahan ini pada tiga baris awal yang memiliki rata – rata nilai ketepatan yang lebih rendah dari tiga baris berikutnya.

Sedangkan nilai *threshold* berpengaruh pada nilai nilai yang terlalu besar, sehingga banya onset yang tidak terdeteksi seperti pada Tabel 1 pada tiga hasil terakhir, yang rata-rata ketepatannya lebih rendah dari tiga baris sebelumnya. Kesalahan yang terutama terjadi ketika *onset* yang benar dianggap sebagai *noise* karena tidak memiliki kenaikan amplitudo (volume) yang cukup.

3.2 Pengujian pitch pada suara monofonik sederhana

Pada pengujian ini digunakan nilai yang maksimal untuk pengujian masing - masing didapatkan bahwa tingkat keberhasilan penguian ini mencapai 81.33%. Pengujian ini memiliki tingkat ketepatan yang cukup tinggi karena pada instrumen gitar maupun piano, frekuensi harmonik yang ada cukup sesuai dengan spesifikasi filter. Adanya *error* yang berupa kesalahan analisa cukup banyak, tetapi yang lebih penting adalah metode ini lebih sering mengalami kesalahan oktaf, dimana sebuah nada terdeteksi satu oktaf dibawah nada sesungguhnya, seperti yang terlihat pada Tabel 2. Rata – rata ketepatan untuk gitar adalah sebesar 79,2%, sedangkan pada piano didapatkan rata – rata pengujian yang lebih baik, yaitu 83,3%. Kesalahan pada gitar adalah ketidak akuratan algoritma dalam mendata nada yang ada, dikarenakan pada gitar senar yang dipetik dapat menyebabkan senar lain ikut bergetar, sedangkan pada intrumen piano tidak menyebabkan hal seperti demikian. Kesalahan pendeteksian pada piano disebabkan oleh kelemahan dari IIR comb filter dimana IIR comb filter dengan spesifikasi frekuensi yang lebih rendah daripada frekuensi dasar

juga dapat menghapus frekuensi dasar serta frekuensi harmoniknya.

Tabel 2. Hasil dari pengujian terhadap suara monofonik

Nama Instrumen, Tangga Nada	Deteksi tepat	Ketepatan	Salah oktaf	Tidak tepat
Gitar, C	7	87,5%	1	0
Gitar, B	8	100%	0	0
Gitar, E	4	50%	1	3
Piano, C	8	100%	0	0
Piano, B	5	62,5%	3	0
Piano, E	7	87,5%	1	0

3.3 Pengujian pitch terhadap suara polifonik sederhana

Pada pengujian ini didapatkan bahwa tingkat keberhasilan pengujian ini mencapai 58,5%. Pada hasil yang didapatkan bervariasi, tetapi dapat terlihat bahwa piano memiliki ketepatan deteksi yang jauh rendah dari pada gitar. Hal ini dikarenakan piano tidak memiliki kekuatan volume pada frekuensi harmoniknya. Gitar memiliki kekuatan volume pada frekuensi harmoniknya, karena struktur instrumen yang secara fisik. Hal ini menyebabkan ketika terpengaruh oleh penghilangan IIR comb filter, ada bagian - bagian tertentu dari sinyal yang ikut menghilang. Tabel 3 menunjukkan hasil dari deteksi nada pada gitar dan piano untuk input nada polifonik. Rata - rata yang didapatkan dari pendeteksian nada polifonik pada gitar adalah 65,5% sedangkan pada piano rata - rata yang didapatkan adalah sebesar 61,1%.

Tabel 3. Hasil dari pengujian terhadap suara polifonik

Nama Chord	Deteksi tepat	Ketepatan	Salah oktaf	Tidak tepat
Gitar, CMaj, C7, Cm	12	80%	2	1
Gitar, BMaj, B7, Bm	5	33,3%	3	7
Gitar, EMaj, E7, Em	15	83,3%	0	3
Piano, CMaj, C7, Cm	5	41,7%	3	2
Piano, BMaj, B7, Bm	8	66,7%	3	2
Piano, EMaj, E7, Em	9	75%	3	1

4. Kesimpulan

Dari pengujian yang dilakukan dapat disimpulkan bahwa:

- Pencarian *onset* bergantung pada nilai *threshold* dan nilai *onset closeness* yang digunakan. Karena hal ini, algoritma *onset detection* kurang menunjukkan hasil yang memuaskan karena kurangnya fleksibilitas yang ditunjukkan oleh penggunaan nilai *threshold* dan *onset closeness*.
- Pencarian *pitch* bergantung pada jenis instrumen yang digunakan. Pada instrumen piano yang tidak memiliki banyak frekuensi harmonik, *pitch detection* pada nada

tunggal menghasilkan nada yang lebih tepat karena sedikit frekuensi yang mengganggu. Sedangkan pada gitar, *pitch detection* pada nada tunggal menghasilkan nada banyak nada yang salah karena adanya suara harmonik lain.

- Pencarian *pitch* pada suara polifonik pada jenis instrumen yang digunakan. Pada instrumen piano yang memiliki sedikit frekuensi harmonik, *pitch detection* tidak dapat menebak data yang dihasilkan setelah filter menghilangkan nada dan semua harmoniknya. Pada gitar, filter yang menghilangkan suara harmonik menyisakan sebagian dari suara yang ada, sehingga pendeteksian nada - nada berikutnya mendapatkan input yang baik.

5. REFERENCES

- [1] Smith, Lloyd A., Rodger J. McNab, and Ian H. Witten. *Music Information Retrieval Using Audio Input*. (1997) Proceedings of the AAAI: Intelligent Integration and Use of Test, Image, Video and Audio Corpora : 12-16
- [2] Klapuri, Anssi, Tuomas Virtanen, Jan-Markus Holm (2005). *A perceptually motivated multiple-f0 estimation method*. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics.
- [3] Rabiner, L. (1977). *On the use of autocorrelation analysis for pitch detection*. Acoustics, Speech and Signal Processing, IEEE Transactions on, 25(1), 24-33.
- [4] Muhammad, G. (2008, December). *Noise robust pitch detection based on extended AMDF*. In Signal Processing and Information Technology, 2008. ISSPIT 2008. IEEE International Symposium on (pp. 133-138). IEEE.
- [5] Emiya, V., Badeau, R., & David, B. (2010). *Multipitch estimation of piano sounds using a new probabilistic spectral smoothness principle*. Audio, Speech, and Language Processing, IEEE Transactions on, 18(6), 1643-1654.
- [6] Vincent, E., Bertin, N., & Badeau, R. (2010). *Adaptive harmonic spectral decomposition for multiple pitch estimation*. Audio, Speech, and Language Processing, IEEE Transactions on, 18(3), 528-537.
- [7] Tadokoro, Y. O. S. H. I. A. K. I., T. A. K. E. O. Morita, and Michiru Yamaguchi. (2003) *Pitch detection of musical sounds noticing minimum output of parallel connected comb filters*. TENCON 2003. Conference on Convergent Technologies for Asia-Pacific Region. Vol. 1.
- [8] Miwa, T., Y. Tadokoro, and T. Saito, *The Problems of Transcription using Comb Filters for Musical Instrument Sounds and Their Solutions*. 2000, TECHNICAL REPORT OF IEICE
- [9] Krishnamurthy, Kapil. (n.d.) *Generation of Control Signals using Pitch and Onset Detection for an Unprocessed Guitar Signal*. Final paper at Center for Computer Research in Music and Acoustics Stanford University.
- [10] Gainza, M., Lawlor, R., Coyle, E.: *Multi pitch estimation by using IIR comb filters*. 47th. International Symposium focused on Multimedia Systems and Applications (ELMAR), Zadar, 2005.
- [11] Pei, S. C., & Tseng, C. C. (1998). A comb filter design using fractional-sample delay. *Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on*, 45(5), 649-65