



TUGAS AKHIR - TF 141581

ESTIMASI F₀ BERBASIS IFAS

**DIPTA NUSBANGGA HAIKAL AHMADA
NRP 2311040000083**

**Dosen Pembimbing
Dr. Eng. Dhany Arifianto, S.T, M.Eng**

**JURUSAN TEKNIK FISIKA
Fakultas Teknologi Industri
Institut Teknologi Sepuluh Nopember
Surabaya 2018**



FINAL PROJECT - TF 141581

IFAS-BASED F_0 ESTIMATION

**DIPTA NUSBANGGA HAIKAL AHMADA
NRP 2311040000083**

**Advisor
Dr. Eng. Dhany Arifianto, S.T, M.Eng**

**DEPARTMENT OF ENGINEERING PHYSICS
Faculty of Industrial Technology
Sepuluh Nopember of Technology Institute
Surabaya 2018**

PERNYATAAN BEBAS PLAGIASI

Saya yang bertanda tangan di bawah ini:

Nama : Dipta Nusbandga Haikal Ahmada

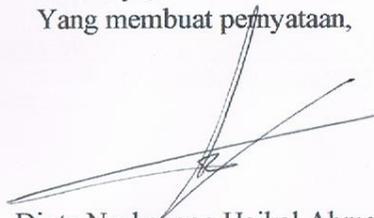
NRP : 2311040000083

Jurusan : Teknik Fisika FTI-ITS

dengan ini menyatakan bahwa tugas akhir saya berjudul "Estimasi F0 Berbasis IFAS" adalah bebas plagiasi. Apabila pernyataan ini tidak benar, maka saya bersedia menerima sanksi sesuai ketentuan yang berlaku.

Demikian surat pernyataan ini saya buat dengan sebenar-benarnya.

Surabaya, 17 Januari 2018
Yang membuat pernyataan,



Dipta Nusbandga Haikal Ahmada
2311040000083

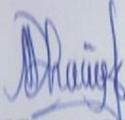
LEMBAR PENGESAHAN
ESTIMASI F0 BERBASIS IFAS
TUGAS AKHIR

Oleh:
Dipta Nusbanga Haikal Ahmada
NRP. 2311040000083

Surabaya, 17 Januari 2018

Mengetahui/ Menyetujui

Pembimbing



Hany Arifianto, ST., M.Eng
IPN. 197310071998021001



Kepala Departemen
Teknik Fisika FTI-ITS

Agus Muhammad Hatta, ST., MSI, Ph.D.
NIPN. 197809022003121002

LEMBARAN PENGESAHAN

ESTIMASI F0 BERBASIS IFAS

TUGAS AKHIR

Diajukan Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Teknik

pada

Bidang Studi Rekayasa Akustik dan Fisika Bangunan
Program Studi S-1 Jurusan Teknik Fisika
Fakultas Teknologi Industri
Institut Teknologi Sepuluh Nopember

Oleh:

DIPTA NUSBANGGA HAIKAL AHMADA

NRP. 2311040000083

Disetujui oleh Tim Penguji Tugas Akhir :

1. Ir. Dhany Arifianto, ST.,M.Eng  (Pembimbing)
2. Ir. Wiratno Argo Asmoro, M, Sc. (Ketua Penguji)
3. Andi Rahmadiansah, ST., MT. (Penguji I)

ESTIMASI F_0 BEBAS IFAS

Nama : Dipta Nusbandha Haikal Ahmada
NRP : 2311040000083
Jurusan : Teknik Fisika FTI - ITS
Pembimbing : Dr. Dhany Arifianto ST, M.Eng

Abstrak

Pada penelitian tugas akhir ini, dilakukan eksperimental implementasi metode estimasi F_0 berbasis *instantaneous frequency amplitude spectrum* (IFAS) pada sinyal suara (*speech*). F_0 merupakan frekuensi yang berfungsi mengkarakterisasi struktur harmonik pada sinyal suara. Dalam implementasinya, F_0 merupakan frekuensi yang memaksimalkan fungsi evaluasi IFAS. Metode estimasi F_0 berbasis IFAS merupakan metode estimasi F_0 yang cukup handal dan akurat jika dibandingkan metode lainnya, seperti: *cepstrum* dan *autocorrelation*. Di samping itu, dalam implementasi estimasi F_0 berbasis IFAS, didefinisikan suatu besaran yang mendeskripsikan keberadaan struktur harmonik dalam sinyal suara secara kuantitatif, yang disebut sebagai *harmonicity measure*. Dalam implementasinya, Harmonicity measure digunakan selektor yang memilih optimal interval frekuensi yang padanya fungsi evaluasi diterapkan, sehingga, error estimasi F_0 dapat direduksi secara substansial.

Pada implementasi estimasi F_0 berbasis IFAS, meskipun *harmonicity measure* memberikan hasil yang optimal, namun keberadaan representasi struktur harmonik masih belum jelas. Hal ini ditunjukkan pada hasil representasi struktur harmonik *instantaneous frequency* (IF), dimana, pada transisi frekuensi harmonik satu ke frekuensi harmonik berikutnya menunjukkan adanya

diskontinuitas. Selain memberikan struktur harmonik yang masih belum jelas, diskontinuitas pada representasi struktur harmonik, membatasi interval frekuensi evaluasi. Oleh karena itu, pada penelitian tugas akhir ini, diusulkan formulasi IF berbasis pada Gaussian window dengan formulasi IF yang dapat mereduksi diskontinuitas pada representasi struktur harmonik IF. Hasil dari IF Gaussian Window diperoleh representasi struktur harmonik yang lebih jelas. Pengujian dari hasil estimasi F_0 berbasis IFAS Gaussian window pada database suara yang terdiri dari 2 pria dan 2 wanita dengan 20 pengucapan, menunjukkan hasil *fine pitch error* pria sebesar 6.2934% dan wanita sebesar 4.9833% dibandingkan dengan metode STRAIGHT.

Kata Kunci: F_0 , Gaussian Window, IF, IFAS

IFAS-based F_0 ESTIMATION

Name : Dipta Nusbanga Haikal Ahmada
NRP : 2311040000083
Departement : Teknik Fisika FTI - ITS
Supervisor : Dr. Dhany Arifianto ST, M.Eng

Abstract

In this final project, experimental implementation of F_0 method based on instantaneous frequency estimation amplitude spectrum (IFAS) on the speech signal. F_0 is a frequency functioning to characterize the harmonic structure of the sound signal. In its implementation, F_0 is frequency that can maximizes IFAS evaluation function. The IFAS F_0 estimation method is a sufficient F_0 estimation method reliable and accurate when compared to other methods, such as: cepstrum and autocorrelation. In addition, in the implementation of an IFAS-based F_0 estimate, a value is defined to describe the existence of harmonic structure in the speech signal quantitatively, which is called the harmonicity measure. In its implementation, Harmonicity measure is used as selector that selects the optimal frequency interval on it evaluation function is applied, so, the estimation error F_0 can be reduced substantially.

In the implementation of an IFAS-based F_0 estimate, although the harmonicity measure gives optimal results, the existence of a harmonic structure representation remains unclear. This matter shown in the result of the representation of the instantaneous frequency (IF) harmonic structure, where, at the harmonic frequency transition one to the next harmonic frequency indicates the presence discontinuity. In addition to give the harmonic structure is still unclear, discontinuities on harmonic structure representation,

limiting the evaluation frequency interval. Therefore, on research of this final task, proposed IF formulation bebasis on Gaussian window with IF formulation that can reduce discontinuity on the representation of IF harmonic structure. Result of IF Gaussian Window obtained the representation of harmonic structure which is clearer. Testing of IFD F_0 bebasis estimation results Gaussian window on a sound database consisting of 2 men and 2 women with 20 pronunciations, showing fine pitch error for male is 6.2934% and for women is 4.9833% compared with the STRAIGHT method.

Keywords: F_0 , *Gaussian Window, IF, IFAS*

KATA PENGANTAR

Alhamdulillah, segala puji dan syukur penulis panjatkan kepada Allah SWT atas limpahan rahmat, hidayat, serta petunjuk-Nya sehingga penulis dapat menyelesaikan tugas akhir yang berjudul “ Estimasi F_0 Berbasis IFAS ”. Penulis telah banyak mendapatkan bantuan dari berbagai pihak dalam menyelesaikan tugas akhir ini. Atas berbagai bantuan itu penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada

1. Bapak dan Ibu yang telah memberikan dukungan, bimbingan, serta doa yang terus mengalir. membutuhkan.
2. Dr. Dhany Arifianto, S.T., M.Eng yang telah sabar dalam motivasi, bimbingan, dan ilmunya dalam menyelesaikan tugas akhir ini, terima kasih atas segala perhatian dan bantuannya.
3. Bapak Agus Muhammad Hatta, ST., M.Si, Ph.D selaku Ketua Jurusan Teknik Fisika ITS dan Dr. Ir. Ali Musyafa', M.Sc selaku dosen wali penulis yang telah sabarmemberikan petunjuk, bimbingan, serta ilmu yang sangat bermanfaat.
4. Bapak dan Ibu dosen Teknik Fisika yang telah memberikan ilmu selama kuliah serta Bapak/Ibu Karyawan Teknik Fisika ITS yang sudah banyak membantu penulis selama kuliah.
5. Semua anggota Laboratorium Rekayasa Akustik dan Fisika Bangunan Teknik Fisika ITS yang telah banyak membantu dalam proses pengambilan data maupun pengolahan data serta telah memberikan dukungan demi kelancaran tugas akhir.
6. Teman-teman seperjuangan TA-wan/wati yang sama-sama mengambil tugas akhir bidang minat Akustik dan Fisika Bangunan, terima kasih telah memberikan semangat dan

mendampingi penulis dalam suka dan duka saat mengerjakan.

7. Semua pihak yang telah mendukung dan memberikan doa yang tidak bisa penulis sebutkan satu per satu.

Penulis menyadari bahwa penulisan laporan tugas akhir ini tidaklah sempurna. Oleh karena itu, sangat diharapkan kritik dan saran yang membangun dari semua pihak sehingga mencapai sesuatu yang lebih baik lagi. Penulis juga berharap semoga laporan ini dapat menambah wawasan yang bermanfaat bagi pembaca.

Surabaya, 24 January 2018

Penulis

DAFTAR ISI

BAB I

PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Permasalahan	3
1.3 Tujuan	3
1.4 Batasan Masalah	4
1.5 Sistematika Laporan	4

BAB II

TINJAUAN PUSTAKA	5
2.1 Konsep Dasar Frekuensi	5
2.1.1 Total Energi dan Frekuensi Rata-rata	8
2.1.2 Instantaneous Frequency (IF)	10
2.2 Sinyal Analitik	13
2.3 Bandpass Filter Bank	15
2.4 Formulasi IF	18
2.5 Amplitudo Spektrum IF (IFAS)	22
2.6 Estimasi F_0	22
2.7 Harmonicity Measure	23

BAB III

METODOLOGI	27
3.1 Estimasi F_0 berbasis IFAS	27
3.2 Klasifikasi <i>Voiced-Unvoiced</i>	29

BAB IV

HASIL DAN PEMBAHASAN	33
4.1 Estimasi F_0 berbasis IFAS	33
4.2 Klasifikasi <i>Voiced-Unvoiced</i>	40

4.3 Perbandingan IFAS dengan STRAIGHT 42

BAB V

KESIMPULAN DAN SARAN 45

4.1 Kesimpulan 45

4.2 Saran 45

DAFTAR PUSTAKA 47

DAFTAR GAMBAR

2.1	Analogi putaran roda terhadap gerak periodik [7]	5
2.2	Transformasi spektrum sinyal real ke sinyal analitik[7]	14
2.3	Bandpass filter bank[7]	16
2.4	Ilustrasi IF[7]	19
2.5	Ilustrasi IFAS[1][2]	22
2.6	Representasi hasil implementasi fungsi evaluasi F_0	24
2.7	Representasi hasil implementasi <i>harmonicity measure</i>	25
2.8	Ilustrasi distribusi harga <i>harmonicity measure</i> [1]	26
3.9	<i>Flowchart</i> estimasi F_0	28
4.10	Representasi hasil implementasi IF Blackman, dan IF Gaussian untuk input sinyal tonal	35
4.11	Representasi hasil implementasi STFT, IF Blackman, dan IF Gaussian untuk input suara vokal	36
4.12	Representasi hasil implementasi STFT, IF Blackman, dan IF Gaussian untuk input suara campuran	37
4.13	Representasi hasil implementasi estimasi F_0 berbasis IFAS .	39
4.14	Representasi hasil implementasi penentuan harga ambang untuk suara vokal	40
4.15	Representasi hasil implementasi penentuan harga ambang untuk suara campuran	41
4.16	Perbandingan F_0 STRAIGHT dengan F_0 IFAS Gaussian window pada naraswara wanita 'fala'.	42
4.17	Perbandingan F_0 STRAIGHT dengan F_0 IFAS Gaussian window pada naraswara wanita 'fena'.	43
4.18	Perbandingan F_0 STRAIGHT dengan F_0 IFAS Gaussian window pada naraswara pria 'mmht'.	44
4.19	Perbandingan F_0 STRAIGHT dengan F_0 IFAS Gaussian window pada naraswara pria 'mjra'.	44

BAB I

PENDAHULUAN

1.1 Latar Belakang

Sinyal suara merupakan sinyal akustik yang diproduksi atas kerja sama antara paru-paru, *epiglottis*, *vocalcords*, *vocaltract*. Untuk menghasilkan sinyal suara, organ paru-paru menekan udara keluar melewati *epiglottis*, *vocalcords* bergetar dan meng*interrupt* aliran udara yang keluar dari paru-paru dengan gerakan membuka-menutup menghasikan gelombang tekanan semi-periodik. *Vocaltract* terdiri dari *pharynx*, *oral cavity*, dan *nasal cavity*, berfungsi *shaping* gelombang tekanan semi-periodik tersebut menjadi sinyal suara yang nantinya ditransmisikan melalui media udara[6]. Sinyal suara terbagi menjadi dua tipikal, yakni, tipikal *voiced* dan tipikal *unvoiced*. Tipikal *voice* merupakan tipikal sinyal suara berupa sinyal periodik yang mana pada tipikal suara ini memiliki struktur harmonik pada spektrum sinyal. Sedangkan tipikal *unvoiced* merupakan tipikal suara yang menyerupai sinyal derau dan tidak memiliki struktur harmonik tertentu.

Sebagaimana pada karakteristik sinyal periodik umumnya dimana karakterisasi sinyal terletak pada frekuensi dasar (F_0) dan frekuensi harmonik ($F_1, F_2, \dots, F_n = n \cdot F_0$). Fitur-fitur suara, seperti intonasi, stress, dan gender sangat dipengaruhi oleh F_0 . Fitur-fitur ini disebut sebagai fitur suprasegmental yang diasosiasikan sebagai melodi, irama, serta penegasan pada makna ucapan, yang merupakan bagian dari aspek inteligensi sinyal suara. Oleh karena itu, F_0 sangat berguna untuk identifikasi serta sintese sinyal suara. Namun, karena adanya proses stokastik pada sinyal suara serta perubahan struktur harmonik pada sinyal yang begitu cepat, menyebabkan penentuan F_0 menjadi sangat sulit.

Adalah metode estimasi F_0 berbasis pada amplitudo spektrum IF (*Instantaneous Frequency*), yang diturunkan dari STFT (*Short-Time Fourier Transform*) sebagai model bandpass filter, merupakan metode estimasi F_0 yang cukup handal baik dalam keadaan normal maupun keadaan dimana sinyal derau terintegrasi dengan sinyal suara. Berdasarkan laporan sebelumnya[1][2], bahwa, metode ini jauh lebih handal jika dibandingkan dengan metode klasik seperti cepstrum dan autocorrelation. Pada estimasi F_0 berbasis amplitudo spektrum IF, F_0 merupakan element dari himpunan frekuensi tertentu yang memaksimalkan output fungsi evaluasi estimasi F_0 . Walaupun lebih unggul dari metode klasik, metode estimasi F_0 ini masih menunjukkan adanya error sebagai akibat dari struktur harmonik sinyal yang kurang begitu jelas. Untuk mengatasi hal ini, diterapkan metode *smoothing* berbasis DP (*Dynamic Programming*)[3]. Adapun alternatif lain agar error estimasi F_0 dapat direduksi secara substansial adalah dengan menerapkan konsep *harmonicity measure* pada estimasi F_0 [1]. Manfaat *harmonicity measure* ini adalah sebagai selektor band IF mana yang memberikan representasi struktur harmonik yang maksimal. Semakin besar harga *harmonicity measure* semakin berpeluang pula eksistensi struktur harmonik pada sinyal suara.

Adapun masalah lainnya pada metode estimasi F_0 berbasis amplitudo spektrum IF ini, yakni, walaupun *harmonicity measure* memberikan hasil perhitungan yang optimal, pada struktur harmonik yang direpresentasikan oleh kurva IF menunjukkan adanya kejanggalan dimana ketika perubahan frekuensi harmonik ke frekuensi harmonik berikutnya pada kurva IF mengalami diskontinuitas. Sehingga, saat proses evaluasi F_0 dimana IF sebagai input fungsi evaluasi harus dilakukan seleksi terlebih dahulu. Oleh

karena itu, pada penelitian kali ini mengusulkan formulasi IF Gaussian window agar nantinya diskontinuitas pada kurva IF dapat direduksi sehingga diperoleh representasi struktur harmonik yang lebih baik.

1.2 Permasalahan

Dari latar belakang yang telah dipaparkan di atas, maka, dapat dapat diambil permasalahan untuk diangkat sebagai permasalahan pada penelitian, yaitu:

- Bagaimana implementasi estimasi F_0 berbasis IFAS dengan perhitungan IF menggunakan formulasi IF Gaussian window agar dapat mereduksi diskontinuitas sehingga didapatkan representasi struktur harmonik yang lebih jelas.
- Bagaimana klasifikasi *voiced* (sinyal suara yang kaya dengan struktur harmonik) dan *unvoiced* (sinyal suara yang menyerupai sinyal derau) pada sinyal suara input berdasarkan perolehan implementasi estimasi F_0 pada permasalahan pertama diatas.

1.3 Tujuan

Tujuan pada penelitian tugas akhir ini, yaitu:

- Mengimplementasi estimasi F_0 berbasis IFAS dengan perhitungan IF menggunakan formulasi IF Gaussian window agar dapat mereduksi diskontinuitas sehingga didapatkan representasi struktur harmonik yang lebih jelas.
- Mengklasifikasikan *voiced* dan *unvoiced* pada sinyal suara input berdasarkan perolehan implementasi estimasi F_0 pada tujuan pertama.

1.4 Batasan Masalah

Adapun batasan pada penelitian yang dilakukan adalah basis data suara yang digunakan berasal dari satu orang pembicara dengan naraswara wanita dan laki-laki kemudian hasil perhitungan estimasi nantinya akan dievaluasi dengan membandingkan hasil estimasi F_0 berbasis IFAS dengan metode STRAIGHT.

1.5 Sistematika Laporan

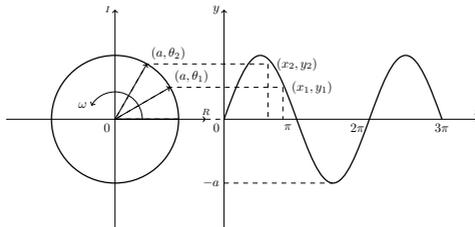
Penulisan laporan tugas akhir ini dibagi menjadi lima bab. Adapun susunan dari laporan ini adalah yang pertama Bab 1 sebagai pendahuluan dimana pada bab ini membahas hal-hal yang melatarbelakangi penelitian pada tugas akhir ini, kemudian hal-hal yang menjadi topik permasalahan penelitian hingga kepada tujuan dan batasan masalah dalam penelitian. Bab 2 berisi materi-materi yang menjadi dasar teori IFAS pada penelitian tugas akhir ini. Selanjutnya adalah Bab 3, yang merupakan bab yang menguraikan mengenai metodologi penelitian bagaimana implementasi estimasi F_0 berbasis IFAS, kondisi eksperimental, hingga pada bagaimana cara mengevaluasi data. Kemudian pada bab 4 akan menjelaskan mengenai hasil dan analisa data. Terakhir bab 5 adalah kesimpulan dari hasil pembahasan permasalahan yang ditetapkan pada penelitian tugas akhir ini.

BAB II TINJAUAN PUSTAKA

2.1 Konsep Dasar Frekuensi

Frekuensi didefinisikan sebagai ukuran jumlah putaran tiap peristiwa dalam interval waktu yang diberikan. Untuk menghitung frekuensi, seseorang perlu menetapkan interval waktu, kemudian menghitung jumlah peristiwa yang terjadi dan membaginya dengan interval waktu yang diberikan sebelumnya. Hasil perhitungan frekuensi dinyatakan dalam satuan *Hertz* (Hz). Frekuensi 1 Hz menyatakan satu peristiwa yang terjadi dalam satu detik. Secara alternatif seseorang bisa mengukur waktu antara dua buah peristiwa (menyebutnya sebagai periode, T), kemudian frekuensi, F , diperoleh sebagai hasil resipokal dari periode, $F = 1/T$. Analogi pada gerak sebuah roda, frekuensi dapat bertanda positif atau negatif yang mengindikasikan arah putaran roda (*clockwise* atau *counter clockwise*). Gerak sebuah roda dapat dimodelkan sebagai fasor. Fasor merupakan representasi matematis pergerakan suatu objek sebagai fungsi sinusoidal kompleks,

$$x(t) = a(t) e^{j\phi(t)}, \quad (2.1)$$



Gambar 2.1: Analogi putaran roda terhadap gerak periodik [7]

dimana $x(t)$ merupakan posisi objek, $a(t)$ adalah amplitudo pada t detik, dan $\phi(t)$ adalah fase sudut. Jika $\phi(t)$ adalah linier terhadap waktu dan $a(t)$ adalah konstan, maka $x(t)$ merepresentasikan gerak melingkar sederhana. Konsep frekuensi dapat digunakan untuk menjelaskan perubahan siklis posisi yang ditunjukkan oleh $x(t)$ dimana $\phi(t)$, misal pada gerak melingkar sederhana, ekuivalen dengan $2\pi Ft$.

Sebagaimana gambaran sinyal sebagai gerak melingkar, perubahan $x(t)$ yang dinyatakan dengan persamaan (2.1) dengan $\phi(t)$ yang linier dan $a(t)$ konstan, merupakan contoh model sinyal monokromatik atau CW (*Continuous Wave*). Dalam konteks sistem linier, model sinyal sederhana ini merupakan fungsi eigen dari sistem. Jika input pada sistem merupakan sinyal sinusoidal dengan frekuensi tertentu (misal $2\pi F$), maka output sistem juga merupakan sinyal sinusoidal dengan frekuensi yang sama namun amplitud atau fase sudut berbeda dengan sinyal input. Sistem linier tidak merubah karakteristik sinyal input jika sinyal input tersebut merupakan sinyal sinusoidal. Adapun pada eksistensi sinyal $x(t)$ di alam bebas dimana kompleksitas bentuk sinyal muncul akibat adanya fenomena *time-varying* ($a(t)$ dan $\phi(t)$ bervariasi terhadap waktu) dan sinyal *additive-noise* yang menyebabkan pemodelan sinyal sulit ditentukan secara deterministik demi mendeskripsikan perilakunya. Tidak cukup sampai disini, untuk pemodelan sinyal $x(t)$, sinyal dipandang sebagai kombinasi linier sinyal-sinyal basis (dikenal sebagai konsep dekomposisi sinyal). Dalam konteks sistem linier, akan lebih bermanfaat jika sinyal $x(t)$ didekomposisi menjadi sinyal sinusoidal. Konsep dekomposisi sinyal ke dalam basis sinusoidal disebut sebagai konsep dekomposisi Fourier. Pada konsep dekomposisi Fourier, sembarang sinyal periodik dapat direpresentasikan kembali sebagai

superposisi sinyal sinusoidal kompleks harmonik yang diekspresikan sebagai

$$x(t) = \sum_{k=0}^K \frac{a_k}{\sqrt{T}} e^{j\omega_k t}, \quad (2.2)$$

dimana K merupakan jumlah komponen basis sinusoidal, a_k adalah spektrum yang bersesuaian dengan komponen basis sinusoidal ke- k , dan ω_k adalah frekuensi harmonik ke- k ($\omega_k = k\omega_0$, $k \in \mathbb{Z}$, $\omega_0 = 2\pi/T$). Pada persamaan (2.2), komponen-komponen basis sinusoidal adalah saling independen dengan komponen basis sinusoidal lainnya. Misal, diberikan dua basis sinusoidal dengan frekuensi harmonik ω_1 dan ω_2 . Kedua basis sinusoidal ini akan saling independen jika dan hanya jika korelasi kedua basis ini memberikan harga yang minimum,

$$\int_{-\frac{\pi}{\omega_0}}^{\frac{\pi}{\omega_0}} e^{j\omega_k t} e^{-j\omega_l t} dt = \begin{cases} \frac{2\pi}{\omega_0}, & \omega_k = \omega_l \\ 0, & \omega_k \neq \omega_l \end{cases}. \quad (2.3)$$

Persamaan (2.2) tidak hanya memberikan representasi struktur basis sinyal $x(t)$ namun juga memberikan pengetahuan mengenai konsep separasi sumber sinyal. Pada berbagai aplikasi pemrosesan sinyal, khususnya aplikasi yang melibatkan konsep dekomposisi Fourier, proses separasi dilakukan dengan mengestimasi harga a_k . Dengan mengkombinasikan persamaan (2.2) dengan (2.3), estimasi harga a_k diekspresikan sebagai

$$a_k = \frac{1}{\sqrt{T}} \int_{-\frac{\pi}{\omega_0}}^{\frac{\pi}{\omega_0}} x(t) e^{-j\omega_k t} dt. \quad (2.4)$$

Adapun pada kasus sinyal $x(t)$ merupakan sinyal non-periodik yang terbatas pada interval waktu tertentu dimana agar padanya

persamaan (2.2) dan persamaan (2.4) berlaku, maka sinyal non-periodik $x(t)$ dipandang sebagai sinyal periodik dengan harga T mendekati limit tak hingga. Jika kuantitas k adalah kontinyu, maka persamaan (2.2) dan persamaan (2.4) diekspresikan kembali sebagai

$$x(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} a(\omega) e^{j\omega t} d\omega, \quad (2.5)$$

$$a(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x(t) e^{-j\omega t} dt. \quad (2.6)$$

Persamaan (2.5) disebut sebagai invers transformasi Fourier dan persamaan (2.6) disebut sebagai transformasi Fourier. Dengan transformasi Fourier, kita bisa memperoleh representasi spektrum yang dapat memberikan informasi mengenai distribusi rapat energi basis sinyal sinusoidal, serta informasi mengenai frekuensi pusat dimana seluruh distribusi energi terlokalisir padanya. Adapun sinyal $x(t)$ pada persamaan (2.2) hingga persamaan (2.6) adalah konsisten terhadap komponen-komponen basis sinusoidal penyusunnya. Dengan kata lain, tidak menutup kemungkinan perubahan basis sinusoidal penyusun terjadi yang jika hal ini terbukti maka dekomposisi Fourier menjadi ambigu.

2.1.1 Total Energi dan Frekuensi Rata-rata

Dalam pemrosesan sinyal, perdeskripsian suatu sinyal, misal $x(t)$, dapat diperoleh dengan melakukan perhitungan total energi dan frekuensi rata-rata padanya. Motivasi dasar perhitungan energi pada sinyal $x(t)$ adalah mengetahui seberapa besar energi yang dibutuhkan untuk membangkitkan sinyal yang serupa. Total energi sinyal secara matematis didefinisikan sebagai

$$E = \int_{-\infty}^{\infty} |x(t)|^2 dt = \int_{-\infty}^{\infty} x(t) \overline{x(t)} dt. \quad (2.7)$$

$\overline{x(t)}$ merupakan konjugat dari $x(t)$, $|x(t)|^2$ adalah energi persatuan waktu (energi sesaat). Tidak hanya pada domain t , perhitungan energi total dapat diperoleh dari domain ω . Jika $X(\omega)$ merupakan spektrum sinyal $x(t)$, maka perhitungan total energi dapat dinyatakan sebagai

$$E = \int_{-\infty}^{\infty} |X(\omega)|^2 d\omega = \int_{-\infty}^{\infty} X(\omega)\overline{X(\omega)} d\omega. \quad (2.8)$$

$|X(\omega)|^2$ adalah rapat energi spektrum. Persamaan (2.8) merupakan cara lain mengekspresikan persamaan (2.7) dengan frekuensi sebagai domain analisis. Kesetaraan perhitungan total energi pada domain waktu dengan domain frekuensi dapat dibuktikan sebagai berikut.

$$\begin{aligned} E &= \int_{-\infty}^{\infty} x(t)\overline{x(t)} dt = \frac{1}{2\pi} \iiint_{-\infty}^{\infty} X(\omega)\overline{X(u)} e^{j(\omega-u)t} dt d\omega du \\ &= \iint_{-\infty}^{\infty} X(\omega)\overline{X(u)}\delta(\omega-u) d\omega du = \int_{-\infty}^{\infty} X(\omega)\overline{X(\omega)} d\omega. \end{aligned}$$

Sedangkan motivasi dasar perhitungan frekuensi rata-rata yakni mengadaptasi pada konsep statistik dalam mengukur pemusatan data. Jika $x(t)$ merupakan sinyal monokromatik, maka perhitungan frekuensi rata-rata menunjukkan frekuensi diaman distribusi energi spektrum terlokalisir padanya. Jika $|X(\omega)|^2$ dipandang sebagai densitas rapat energi spektrum, maka formulasi frekuensi rata-rata diekspresikan sebagai

$$\langle \omega \rangle = \int_{-\infty}^{\infty} \omega |X(\omega)|^2 d\omega. \quad (2.9)$$

Kesetaraan analisis pada domain t juga berlaku pada perhitungan frekuensi rata-rata. Adapun kesetaraan perhitungan frekuensi rata-rata dengan memandang $|x(t)|^2$ sebagai densitas rapat energi domain waktu,

$$\begin{aligned} \langle \omega \rangle &= \int_{-\infty}^{\infty} \omega |X(\omega)|^2 d\omega = \frac{1}{2\pi} \iiint_{-\infty}^{\infty} \omega x(\tau) \overline{x(t)} e^{j(t-\tau)} d\omega d\tau dt \\ &= \iint_{-\infty}^{\infty} x(\tau) \overline{x(t)} \frac{1}{j} \frac{\partial}{\partial t} \delta(t - \tau) d\tau dt = \int_{-\infty}^{\infty} \overline{x(t)} \frac{1}{j} \frac{\partial}{\partial t} x(t) dt. \end{aligned}$$

2.1.2 Instantaneous Frequency (IF)

Pada penjelasan sebelumnya bahwa munculnya kompleksitas bentuk pada suatu sinyal merupakan akibat dari fenomena *time-varying* pada sinyal tersebut. Fenomena *time-varying* dapat diekspresikan sebagai fenomena modulasi amplitudo-frekuensi (AM-FM). Mengingat sinyal periodik sembarang dapat direpresentasikan kembali sebagai kombinasi linier basis sinyal sinusoidal, model sinyal periodik, $x(t)$, dengan tanpa mengabaikan fenomena *time-varying* dinyatakan sebagai

$$x(t) = \sum_{k=0}^K a_k(t) \cos \left(\underbrace{\phi_{0,k} + 2\pi \int_0^t F_k(t) dt}_{\phi_k(t)} \right). \quad (2.10)$$

Persamaan (2.10) disebut sebagai model sinyal multi-komponen modulasi AM-FM. Model ini muncul di berbagai aplikasi pemrosesan sinyal seperti radar, pencitraan medis, dan pemrosesan sinyal audio. Fokus pada pemrosesan sinyal audio, model sinyal

AM-FM dapat mendeskripsikan karakteristik perubahan konten frekuensi sumber sinyal suara pada interval waktu tertentu dan juga mampu mendeskripsikan karakteristik resonansi *vocal-tract* pada proses produksi sinyal suara.

Sebagaimana pada persamaan (2.10) terdapat dua parameter penting, $A_k(t)$ (spektrum sesaat) dan $\phi(t)$ (fase sudut sesaat). Untuk menentukan dua parameter ini, terlebih dahulu dilakukan proses separasi sumber komponen sinyal AM-FM. Akan tetapi, agar proses separasi sumber ini tidak memberikan hasil yang ambigu akibat fenomena *time-varying* maka perlu mengasumsikan pada perubahan spektrum, $A_k(t)$, tidaklah terlalu signifikan terhadap waktu, yang artinya bahwa sinyal $x(t)$ adalah cukup konsisten terhadap komponen-komponen penyusunnya pada interval waktu analisis. IF, yakni frekuensi sesaat, dapat diperoleh dari turunan pertama parameter $\phi_k(t)$ terhadap waktu. Pada konteks gerak melingkar dimana sudut fase sesaat mendeskripsikan perubahan sudut sesaat dan IF mendeskripsikan laju melingkar sesaat yang diekspersikan sebagai

$$\text{IF} \triangleq F_k(t) = \frac{1}{2\pi} \frac{\partial}{\partial t} \phi_k(t). \quad (2.11)$$

IF juga dapat diperoleh dari konsep perhitungan frekuensi rata-rata yang dinyatakan pada persamaan (2.9). Misal, jika $x_k(t)$ adalah sinyal AM-FM yang diekstrak dari sinyal utama pada frekuensi pusat spektrum lokal ω_k dengan lebar spektrum ΔB , maka spektrum sinyal $x_k(t)$ adalah

$$X_k(\omega) = \text{rect} \left(\frac{\omega - \omega_k}{\Delta B} \right) X(\omega). \quad (2.12)$$

Agar separasi sumber diharapkan memisahkan frekuensi sumber spesifik, maka ΔB harus cukup sempit agar perubahan spektrum

lokal tidak terlalu signifikan dalam retang ΔB , sehingga, diperoleh

$$\begin{aligned} x_k(t) &= \int_{-\infty}^{\infty} \text{rect}\left(\frac{\omega - \omega_k}{\Delta B}\right) X(\omega_k) e^{j\omega_k t} d\omega \\ &\approx X(\omega_k) e^{j\omega_k t} \Delta B \frac{\sin(\pi \Delta B t)}{\pi \Delta B t}. \end{aligned} \quad (2.13)$$

Adapun harga ΔB dipilih mendekati harga *infinitesimal*, $\Delta B \rightarrow \delta B$. Hal ini menjadikan sinyal AM-FM sebagai sinyal monokromatik (sinyal dengan spektrum yang bersesuaian dengan garis tunggal spektral tegak (*narrowband*)),

$$x_k(t) = X(\omega_k) \delta B e^{j\omega_k t} = a_k e^{j\phi_k(t)} \quad (2.14)$$

Untuk total energi pada sinyal $x_k(t)$, perhitungan dievaluasi pada interval waktu sesaat sebab kita ingin memperhitungkan kuantitas sesaat pada sinyal,

$$E_{\omega_k, t_0} = \int_{t_0}^{t_0 + \delta t} x_k(t) \overline{x_k(t)} dt \approx x_k(t_0) \overline{X(\omega_k)} e^{-j\omega_k t_0} \delta B \delta t, \quad (2.15)$$

$$\xi(\omega_k, t_0) = \frac{E_{\omega_k, t_0}}{\delta B \delta t} = x(t_0) \overline{X(\omega_k)} e^{-j\omega_k t_0}. \quad (2.16)$$

$\xi(\omega_0, t_0)$ adalah densitas rapat energi spektrum sesaat yang dikenal sebagai rapat energi Rihaczek. Dengan mensubstitusikan rapat energi Rihaczek dan $x_k(t)$ sebagai $x(t)$ dalam perhitungan frekuensi rata-rata maka diperoleh

$$\begin{aligned} \langle \omega \rangle &= \int_{-\infty}^{\infty} \omega \xi(\omega, t) d\omega = \int_{-\infty}^{\infty} \omega x(t) \overline{X(\omega)} e^{-j\omega t} d\omega \\ &= jx(t) \frac{\partial}{\partial t} \overline{x(t)} = |x(t)|^2 \left[\frac{\partial \phi(t)}{\partial t} - \frac{1}{ja(t)} \frac{\partial a(t)}{\partial t} \right] \end{aligned} \quad (2.17)$$

Karena $|x(t)|^2$ merupakan besaran densitas, maka harga $|x(t)|^2$ adalah sama dengan satu, dan karena $a(t)$ tidaklah terlalu signifikan terhadap perubahan waktu mengingat bahwa sinyal utama adalah konsisten terhadap sinyal-sinyal penyusun, maka diperoleh definisi IF sebagai

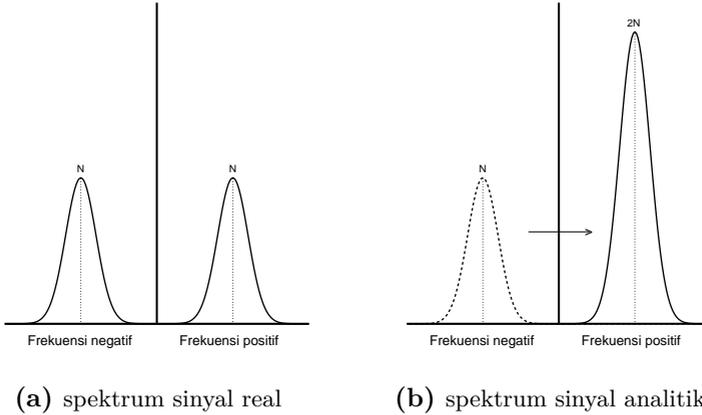
$$\text{IF} \triangleq \frac{1}{2\pi} \frac{\partial \phi(t)}{\partial t} = \frac{\langle \omega \rangle}{2\pi}. \quad (2.18)$$

2.2 Sinyal Analitik

Tipikal sinyal suara atau audio di alam bebas adalah sebagian besar merupakan sinyal real. Saat merekam suara dengan melewati kepada perangkat microphone kemudian menyimpannya pada harddisk atau penyimpanan lain, pada dasarnya kita melakukan pemrosesan terhadap sinyal real. Karakteristik spektrum sinyal real adalah simetris terhadap sumbu garis tegak yang memotong frekuensi pusat nol sehingga pada spektrum sinyal real terdapat komponen sumber frekuensi negatif dan positif dengan proporsi yang sama. Hal ini, dapat menimbulkan masalah ketika melakukan perhitungan IF dengan menggunakan persamaan (2.9). Dengan karakteristik spektrum yang simetris, perhitungan IF akan selalu menghasilkan harga nol. Gabor mengusulkan sebuah representasi sinyal yang disebut sebagai sinyal analitik, yakni mencari komponen *quadrature* (imaginer) aditif yang diperoleh dengan memproses sinyal real menggunakan transformasi Hilbert. Jika $z(t)$ merupakan sinyal analitik untuk sinyal real $x(t)$, akan hubungan matematis sinyal analitik dinyatakan sebagai

$$z(t) = a(t) e^{j\phi(t)} = x(t) + j\mathcal{H}[x(t)], \quad (2.19)$$

$$\mathcal{H}[x(t)] = Pr \int_{-\infty}^{\infty} \frac{x(\tau)}{(t - \tau)\pi} d\tau.$$



Gambar 2.2: Transformasi spektrum sinyal real ke sinyal analitik[7]

Dalam perspektif domain spektrum bahwa penambahan komponen *quadrature* pada sinyal real adalah ekuivalen dengan proses eliminasi komponen frekuensi sumber negatif. Seperti pada tujuan awal representasi analitik, proses eliminasi frekuensi sumber negatif menjamin perhitungan IF tidak akan selalu menghasilkan angka nol. Uraian sinyal analitik dalam perspektif spektrum dinyatakan sebagai berikut.

$$\begin{aligned}
 z(t) &= \frac{2}{\sqrt{2\pi}} \int_0^{\infty} X(\omega) e^{j\omega t} d\omega = \frac{1}{\pi} \int_0^{\infty} \int_{-\infty}^{\infty} x(\tau) e^{-j\omega(\tau-t)} d\tau d\omega \\
 &= \frac{1}{\pi} \int_{-\infty}^{\infty} x(\tau) \left[\frac{1}{j(\tau-t)} + \pi\delta(\tau-t) \right] d\tau \\
 &= x(t) + j \int_{-\infty}^{\infty} \frac{x(\tau)}{\pi(t-\tau)} d\tau.
 \end{aligned}$$

Adapun persamaan $\mathcal{H}[x(t)]$ yakni pada perhitungan pecahan pada bagian imajiner akan bermasalah ketika harga τ pada penyebut adalah sama dengan t . Perhitungan pecahan ini dilakukan dengan pendekatan limit yang dikenal sebagai pendekatan asas Cauchy. Pendekatan harga asas Cauchy dinotasikan sebagai Pr dan representasi pendekatan limit persamaan $\mathcal{H}[x(t)]$ adalah

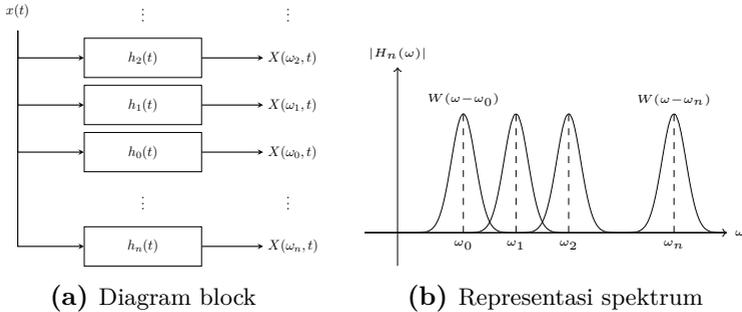
$$Pr \int_{-\infty}^{\infty} \frac{x(\tau)}{\pi(t - x\tau)} d\tau \triangleq \lim_{\epsilon \rightarrow 0} \left[\int_{t-1/\epsilon}^{t-\epsilon} + \int_{t+\epsilon}^{t+1/\epsilon} \right] \frac{x(\tau)}{\pi(t - \tau)} d\tau$$

Dengan demikian ekspresi sinyal analitik sebagai penjumlahan sinyal real dengan sinyal *quadrature* merupakan implikasi dari proses eliminasi spektrum frekuensi negatif. Jika $Z(\omega)$ merupakan spektrum sinyal analitik dan $X(\omega)$ merupakan spektrum sinyal real, maka secara matematis relasi kedua spektrum dinyatakan sebagai

$$Z(\omega) = \begin{cases} 2X(\omega), & \omega > 0 \\ X(0), & \omega = 0 \\ 0, & \omega < 0 \end{cases} \quad (2.20)$$

2.3 Bandpass Filter Bank

Separasi sumber merupakan metode utama dalam pemrosesan sinyal terutama jika pemrosesan dilakukan pada sinyal yang dimodelkan sebagai superposisi komponen-komponen basis sinusoidal. Umumnya, untuk memisahkan atau menseleksi komponen sumber spesifik, sinyal dilewatkan pada suatu bandpass filter. Peran bandpass filter ialah menseleksi spektrum yang terlokalisir pada frekuensi pusat dan lebar interval tertentu dan mereduksi spektrum diluar interval ini. Sehingga untuk memisahkan seluruh komponen-komponen sumber penyusun, sinyal dilewatkan



Gambar 2.3: Bandpass filter bank[7]

pada sekumpulan bandpass filter yang disusun secara paralel dimana masing-masing bandpass filter bekerja secara independen. Sekumpulan filter ini disebut sebagai bandpass filter bank.

Secara teoritis, proses separasi sumber dapat dilakukan dengan menerapkan transformasi Fourier pada suatu sinyal input. Transformasi Fourier dapat dipandang sebagai bandpass filter bank ideal. Adapun batasan pada transformasi Fourier, sinyal input haruslah konsisten terhadap komponen-komponen penyusun pada interval waktu analisis (stasioner). Transformasi Fourier waktu-pendek (*Short-Time Fourier Transform, STFT*) merupakan modifikasi dari transformasi Fourier yang secara khusus digunakan untuk sinyal non-stasioner. Pada STFT, sinyal non-stasioner dipartisi menjadi segmen-segmen waktu pendek yang padanya segmen merupakan sinyal stasioner (perubahan pola amplitudo-amplitudo spektrum sinyal pada segmen tidaklah terlalu signifikan). Proses partisi ini dilakukan dengan cara mengalikan sinyal dengan suatu fungsi pembatas (*tapper*) pada sinyal non-stasioner yang disebut sebagai fungsi window. Fungsi win-

dow merupakan fungsi real, simetris terhadap titik pusat yang membagi segmen waktu-pendek tunggal sama panjang, dan tidak berharga nol pada interval $[-T, T]$. Blackman, Gaussian, dan Hamming merupakan contoh fungsi window yang umum digunakan.

Fokus pada sinyal suara, jika $x(t)$ mendeskripsikan perubahan pola sinyal suara total, dan $X(\omega, t)$ merupakan spektrum sinyal suara waktu-pendek dengan t sebagai waktu pusat segmen tunggal spesifik, maka STFT dinyatakan sebagai

$$X(\omega, t) = \int_{-\infty}^{\infty} x(\tau) w(\tau - t) e^{-j\omega\tau} d\tau. \quad (2.21)$$

STFT dapat dipandang sebagai bandpass filter bank ideal. Misal, jika dipilih komponen bandpass filter ke- n dengan spektrum yang terlokalisir pada frekuensi $\omega_n = n\Delta B$, ΔB adalah resolusi dari bandpass filter bank, output komponen bandpass filter ke- n adalah

$$\begin{aligned} e^{-j\omega_n t} \int_{-\infty}^{\infty} x(\tau) w(\tau - t) e^{-j\omega_n(\tau - t)} d\tau &= X(\omega_n, t) \\ e^{-j\omega_n t} \left[x(t) * w(-t) e^{j\omega_n t} \right] &= X(\omega_n, t) \quad (2.22) \\ e^{-j\omega_n t} H(\omega_n, t) &= X(\omega_n, t), \end{aligned}$$

dimana $*$ merupakan operator konvolusi, $H(\omega_n, t)$ merupakan output bandpass filter ke- n dengan impulse respon filter adalah $w(-t) e^{j\omega_n t}$ yang causal dengan selang integral efektifnya adalah $[-T, T]$. $H(\omega_n, t)$ merupakan sinyal analitik, karena jika fokus pada spektrum positif maka bandpass filter akan mereduksi spektrum frekuensi negatif. Adapun pada $H(\omega_n, t)$, dimana lebar spektrum $H(\omega_n, t)$ ditentukan lebar spektrum fungsi winow $w(t)$,

sehingga agar memperoleh IF pada sinyal suara maka lebar spektrum fungsi window $w(t)$ haruslah cukup sempit agar solusi IF sebagai turunan pertama fase sudut $H(\omega_n, t)$. Disamping itu, jika spektrum fungsi window $w(t)$ cukup sempit, maka kecil sekali kemungkinan terjadi aliasing saat proses separasi sumber disekitar frekuensi Nyquist.

2.4 Formulasi IF

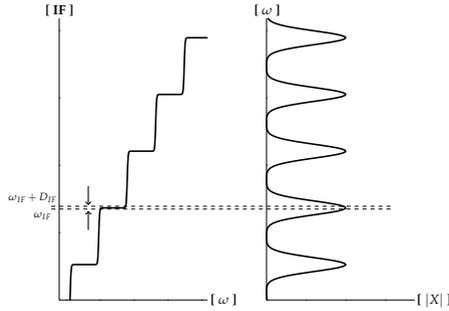
Sebagaimana pada penjelasan sebelum mengenai IF, bahwa IF dapat diperoleh dari turunan pertama sudut fase sinyal analitik. Sehingga untuk memproses sinyal real agar memperoleh IF maka sinyal real harus dikonversi terlebih dahulu ke bentuk sinyal analitik. Tentunya proses konversi sinyal ke bentuk analitik tidak perlu dilakukan jika IF dihitung dengan formulasi IF yang diturunkan dari STFT sebagai bandpass filter ideal. Misal, $x(t)$ merupakan sinyal real non-stasioner dalam domain waktu, transformasi Fourier waktu-pendek diekspresikan sebagai

$$X(\omega, t) = \int_{-\infty}^{\infty} x(\tau) w(\tau - t) e^{-j\omega t} d\tau, \quad (2.23)$$

$$X(\omega, t) = e^{-j\omega t} H(\omega, t),$$

dimana $w(t)$ adalah fungsi window dan $H(\omega, t)$ adalah output bandpass filter berfrekuensi pusat ω . Adapun maksud dari variabel t adalah STFT diaplikasikan pada selang waktu pendek sinyal non-stasioner dimana selang waktu-pendek ini berpusat pada t . IF diperoleh dari turunan pertama $\angle H(\omega, t)$ terhadap t . Sedangkan $\angle H(\omega, t)$ adalah bagian imajiner dari $\log_e H(\omega, t)$. Jika IF dinotasikan sebagai $\lambda(\omega, t)$, maka formulasi IF adalah

$$\lambda(\omega, t) = \frac{\partial}{\partial t} \angle H(\omega, t) = \text{Im} \frac{1}{H(\omega, t)} \frac{\partial H(\omega, t)}{\partial t} \quad (2.24)$$



Gambar 2.4: Ilustrasi IF[7]

Jika mensubstitusikan $H(\omega, t)$ sebagai output bandpass filter bank STFT yang dinyatakan dengan persamaan (2.23), maka formulasi IF dinyatakan sebagai

$$\lambda(\omega, t) = \omega + \text{Im} \frac{1}{X(\omega, t)} \frac{\partial X(\omega, t)}{\partial t} \quad (2.25)$$

$$\frac{\partial X(\omega, t)}{\partial t} = - \int_{-\infty}^{\infty} x(\tau) \frac{\partial w}{\partial t}(\tau - t) e^{-j\omega\tau} d\tau.$$

Alternatif lain dalam memrepresentasikan persamaan IF, yakni dengan menerapkan kondisi dimana output bandpass filter, $H(\omega, t)$, memenuhi syarat kekontinyuan domain kompleks. Jika $H(\omega, t)$ memenuhi syarat kekontinyuan domain kompleks, maka berlaku

$$\text{Im} \frac{1}{H(\omega, t)} \frac{\partial H(\omega, t)}{\partial t} = \text{Re} \frac{1}{H(\omega, t)} \frac{\partial H(\omega, t)}{\partial \omega}. \quad (2.26)$$

Agar terpenuhinya syarat kekontinyuan pada $H(\omega, t)$, maka fungsi window $w(t)$ pada impulse respon bandpass filter adalah fungsi

window Gaussian yang diekspresikan sebagai $w(t) = e^{-t^2/T^2}$, dimana T merupakan parameter yang mengontrol lebar fungsi window. Syarat kekontinyuan ini merupakan implikasi dari prinsip Cauchy-Riemann dalam teori bilangan kompleks.

Untuk menyatakan persamaan IF secara eksplisit dengan mempertimbangkan syarat kekontinyuan domain kompleks, terlebih dahulu memodelkan sinyal $x(t)$ waktu-pendek sebagai superposisi harmonik fungsi basis sinusoidal kompleks, $x(t) = \sum_n a_n e^{j\omega_n t}$. Pada model superposisi harmonik ini, IF tidak lain merupakan kelipatan bulat frekuensi dasar ω_0 , $\lambda = \{\dots, n\omega_0, (n-1)\omega_0, (n-2)\omega_0, \dots\}$. Karena $H(\omega, t)$ merupakan output bandpass filter ideal, maka frekuensi respon $H(\omega, t)$ akan bersesuaian dengan komponen-komponen spektrum fungsi basis sinusoidal, yakni, $X_n(\omega) = \delta(\omega - n\omega_0)$. Oleh karena itu, jika berfokus pada komponen tunggal sinyal $x(t)$ berfrekuensi ω_0 dan $w(t)$ adalah Gaussian window, maka diperoleh formulasi IF, yaitu

$$\operatorname{Re} \frac{1}{W(\omega - \omega_0, t)} \frac{\partial W(\omega - \omega_0, t)}{\partial \omega} = -T^2(\omega - \omega_0), \quad (2.27)$$

$$\lambda(\omega, t) = \omega_0,$$

dimana $W(\omega)$ merupakan frekuensi response fungsi window Gaussian. Adapun IF komponen sinyal lain, misal $n\omega_0$, maka kita perlu menggeser satu persatu frekuensi pusat $W(\omega)$, yang artinya kita memilih komponen bandpass filter lain dengan $n\omega_0$ sebagai frekuensi pusat. Namun, dalam perosedur perhitungan IF yang dinyatakan pada persamaan (27) kita tidak perlu menggeser frekuensi pusat $W(\omega)$, tetapi cukup mentransformasikan $-jx(t) \cdot w(t)$ sebagai substituen untuk $\partial W/\partial \omega$, lalu mentransformasikan $x(t) \cdot w(t)$ sebagai substituen untuk $W(\omega)$ dan memposisikan ω_0 secara eksplisit pada persamaan (27) untuk memperoleh IF.

Jika kembali kepada penjelasan sebelumnya pada topik band-pass filter bank, bahwa respon impulse komponen tunggal band-pass filter bank adalah causal dimana selang integral efektif respon impulse filter adalah $-T < t < T$. Maksud dari filter causal ini ialah jika diterapkan operasi konvolusi pada sinyal input dengan filter causal, maka pada sinyal output tidak mengalami distorsi fase sudut, yang diekspresikan sebagai

$$\frac{\partial \lambda(\omega, t)}{\partial \omega} = 0. \quad (2.28)$$

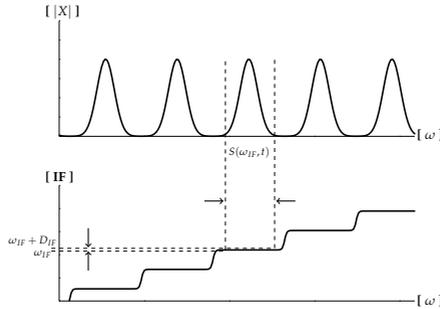
Kemudian pada persamaan (27) dimana parameter T^2 kita ganti dengan konstanta K . Lalu menerapkan persamaan (28) agar mendapatkan konstanta K yang sesuai. Sehingga diperoleh formulasi IF yaitu

$$K = \max \left[T^2, \left[\operatorname{Re} \frac{1}{W(\omega - \omega_0, t)} \frac{\partial W(\omega - \omega_0, t)}{\partial \omega} \right]^2 - \operatorname{Re} \frac{1}{W(\omega - \omega_0, t)} \frac{\partial^2 W(\omega - \omega_0, t)}{\partial \omega^2} \right], \quad (2.29)$$

$$\operatorname{Re} \frac{1}{W(\omega - \omega_0, t)} \frac{\partial W(\omega - \omega_0, t)}{\partial \omega} = -K (\omega - \omega_0), \quad (2.30)$$

$$\lambda(\omega, t) = \omega_0.$$

Operator max pada persamaan (29) adalah sebagai konsistensi pada persamaan (27). Hal ini dikarenakan pada argumen kedua pada operator max akan bernilai nol akibat turunan pertama dan kedua spektrum $W(\omega)$ bernilai nol sehingga menjadikan hasil perhitungan IF menjadi tak terdefinisi.



Gambar 2.5: Ilustrasi IFAS[1][2]

2.5 Amplitudo Spektrum IF (IFAS)

Berdasarkan pembahasan sebelumnya mengenai bandpass filter bank, bahwa magnitudo spektrum output bandpass filter ekuivalen dengan magnitudo spektrum STFT, $|H(\omega, t)| = |X(\omega, t)|$. Kemudian jika kita mengambil integral $|H(\omega, t)|$ pada waktu lokal konstan t dengan integral efektif ω , yakni, $\Omega_0 = \{\omega | \lambda_0 \leq \lambda(\omega, t) \leq \lambda_0 + \Delta\lambda\}$, maka amplitudo spektrum IF (IFAS) diekspresikan sebagai

$$S(\lambda_0, t) = \lim_{\Delta\lambda \rightarrow 0} \frac{1}{\Delta\lambda} \int_{\Omega_0} |H(\omega, t)| d\omega. \quad (2.31)$$

Untuk kemudahan penotasi IFAS pada pembahasan selanjutnya amplitudo spektrum IF pada waktu lokal konstan t dinotasikan sebagai $S(\lambda_0)$.

2.6 Estimasi F_0

Misalkan $S(\lambda)$ merupakan amplitudo spektrum IF sinyal suara input. F_0 merupakan frekuensi, misal F , yang memaksimalkan fungsi evaluasi. Adapun fungsi evaluasi ini mentransformasikan

$S(\lambda)$ ke bentuk besaran yang mendeskripsikan kemungkinan frekuensi F sebagai kandidat kuat F_0 . Secara matematis fungsi evaluasi F_0 dinyatakan sebagai

$$\eta(F) = 10^{-\beta/F} \int_{\lambda_0}^{\lambda_1} S(\lambda) \Lambda(\lambda, F) d\lambda, \quad (2.32)$$

dimana β merupakan konstanta real dan

$$\Lambda(\lambda, F) = \begin{cases} 0, & \lambda/F < \pi \\ \frac{1}{2} (\cos(\lambda/F) + 1), & \lambda/F \leq \pi. \end{cases} \quad (2.33)$$

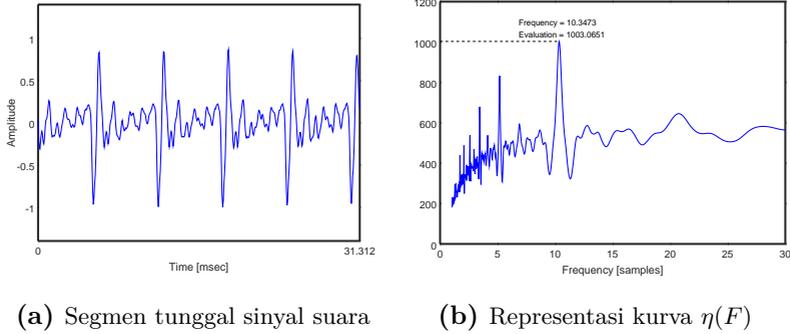
Jika pada interval $[\lambda_0, \lambda_1]$ terdapat struktur harmonik dengan F_0 sebagai frekuensi dasar, $\eta(F)$ akan menunjukkan maksimum lokal pada $F = F_0/n$, $n = 1, 2, 3, \dots$. Sedangkan $10^{-\beta/F}$ berlaku sebagai pembobot agar F tidak terjebak pada maksimum lokal F_0/n , $n = 2, 3, \dots$, dengan kata lain, prioritas F_0 adalah F yang tertinggi bersesuaian dengan maksimum lokal $\eta(F)$. Adapun pada persamaan (32), karena $S(\lambda)$ ekuivalen dengan amplitudo spektrum STFT, maka evaluasi F_0 dapat langsung diterapkan pada representasi spektrum STFT tanpa harus menghitung terlebih dahulu $S(\lambda)$. Karenanya, persamaan (32) dapat ditulis kembali sebagai

$$\eta(F) = 10^{-\beta/F} \int_{\Omega} |X(\lambda)| \Lambda(\lambda, F) d\lambda, \quad (2.34)$$

dimana $\Omega = \{\omega | \lambda_0 \leq \lambda(\omega) \leq \lambda_1\}$ dan asumsi pada integral persamaan (32) dan (34) memenuhi kriteria konvergensi Lebesgue.

2.7 Harmonicity Measure

Estimasi F_0 akan memberikan hasil yang berarti jika pada spektrum sinyal terdapat struktur harmonik, adapun sebaliknya F_0 akan memberikan hasil error. Oleh karena itu, sebelum



Gambar 2.6: Representasi hasil implementasi fungsi evaluasi F_0

melakukan estimasi, kita perlu mendefinisikan terlebih dahulu eksistensi struktur harmonik pada sinyal secara kuantitatif yang disebut sebagai *harmonicity measure*. Dengan *harmonicity measure* tidak hanya memberikan pendeskripsian secara kuantitatif mengenai struktur harmonik, tetapi mampu digunakan sebagai selektor interval frekuensi yang kaya dengan struktur harmonik sehingga error estimasi F_0 dapat direduksi secara substansial. Secara matematis, *harmonicity measure* didefinisikan sebagai

$$\epsilon_{\lambda_0, \lambda_1}(F) = \frac{1}{m(\Omega)} \int_{\Omega} \mathcal{C}(\lambda(\omega), F) d\omega, \quad (2.35)$$

dimana $\Omega = \{\omega | \lambda_0 \leq \lambda(\omega) \leq \lambda_1\}$ dan

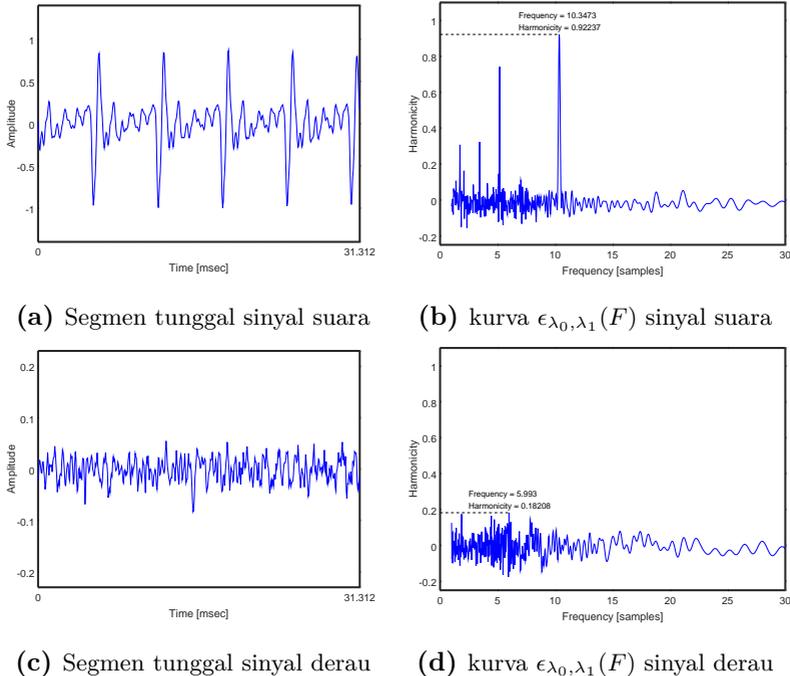
$$\mathcal{C}(\lambda, F) = \begin{cases} 0, & \lambda/F < \pi/2 \\ \cos(\lambda/F), & \lambda/F \leq \pi/2. \end{cases} \quad (2.36)$$

$m(\Omega)$ adalah *Lebesgue's measure* interval Ω atau panjang interval Ω atau jika interval Ω merupakan himpunan diskrit terhingga

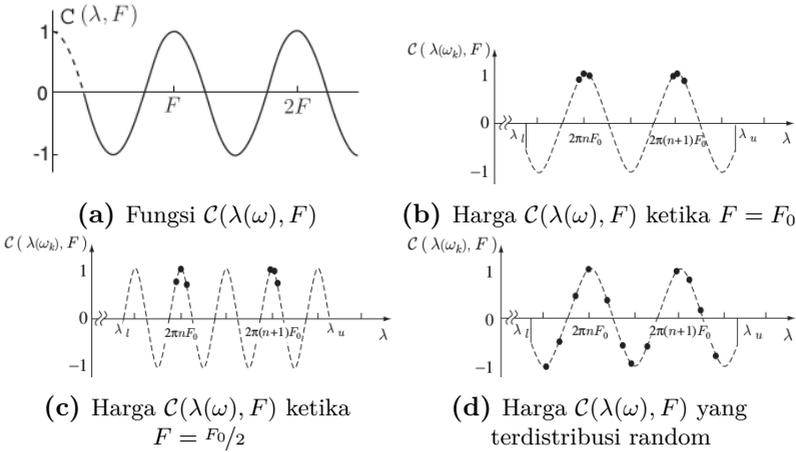
maka *Lebesgue's measure* dari Ω adalah jumlah elemen yang termasuk dalam himpunan. Oleh karena itu, agar error estimasi F_0 dapat direduksi secara substansial, maka, evaluasi F_0 diterapkan pada interval domain IF yang memaksimalkan harga *harmonicity measure*,

$$P_{\lambda_0, \lambda_1} = \max_P \epsilon_{\lambda_0, \lambda_1}(F). \quad (2.37)$$

Harga P_{λ_0, λ_1} berkisar antara -1 hingga 1. Jika mendapati struktur harmonik pada spektrum sinyal, maka P_{λ_0, λ_1} akan bernilai



Gambar 2.7: Representasi hasil implementasi *harmonicity measure*



Gambar 2.8: Ilustrasi distribusi harga *harmonicity measure*[1]

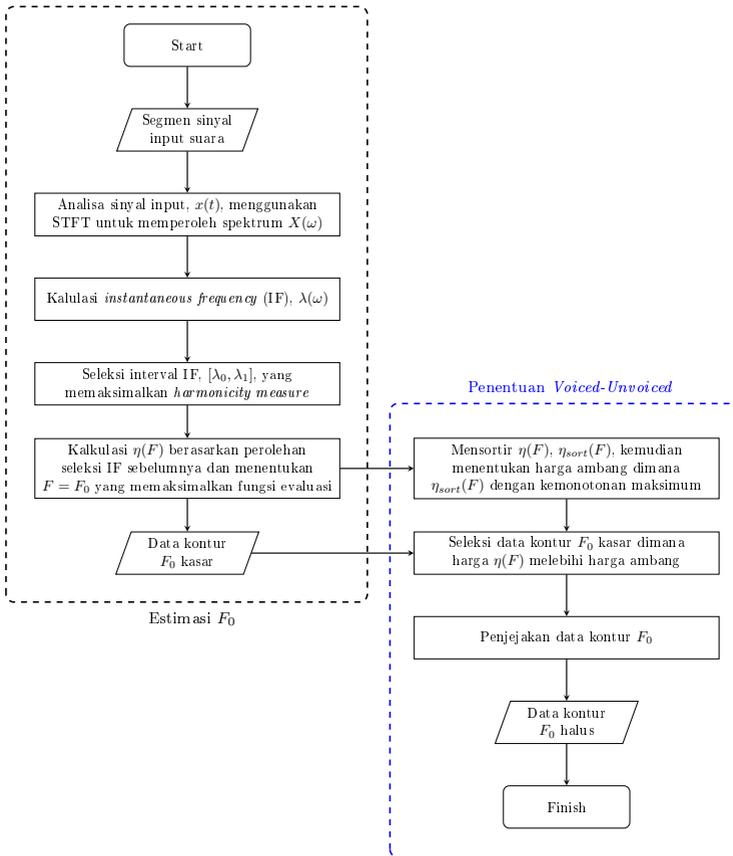
mendekati 1 pada $F = F_0/n$, $n = 1, 2, 3, \dots$. Sebaliknya jika mendapati struktur harmonik yang kurang jelas pada spektrum sinyal maka harga P_{λ_0, λ_1} berada disekitar angka 0 akibat IF yang terdistribusi secara random.

BAB III METODOLOGI

3.1 Estimasi F_0 berbasis IFAS

Berdasarkan uraian sebelumnya bahwa estimasi F_0 berbasis IFAS Gaussian window merupakan metode yang diimplementasikan pada penelitian tugas ini. Pada implementasi estimasi F_0 , sinyal suara disampling pada frekuensi 16 kHz dan dipartisi menjadi beberapa segmen dengan panjang segmen tunggal adalah 1001 sampel (62.562 msec) dan pergeseran segmen ke segmen berikutnya adalah 16 sampel (1 msec). Kemudian pada implementasi, perhitungan spektrum segmen sinyal suara menggunakan transformasi Fourier diskrit (DFT) dengan 1024 partisi frekuensi sampling dengan harapan akurasi frekuensi berada pada level 0.1% dari jarak pucak maksimum lokal dengan *zero-crossing* pertama. Selanjutnya, menghitung IF pada segmen sinyal dengan menggunakan formulasi IF Gaussian-window yang dinyatakan pada persamaan (29) dan (30) atau dengan Formulasi IF Blackman-window yang dinyatakan pada persamaan (24) dan (25). Selanjutnya, menghitung *harmonicity measure* (persamaan (35) dan (36) berdasarkan perolehan IF agar diperoleh pendeskripsian secara kuantitatif eksistensi struktur harmonik pada segmen tunggal sinyal suara yang mana semakin besar harga maksimum yang diperoleh dari perhitungan *harmonicity measure* semakin jelas pula eksistensi struktur harmonik.

Dan tahap terakhir adalah mengevaluasi struktur harmonik yang direpresentasikan oleh IF dan amplitudo spektrum segmen sinyal suara dengan fungsi evaluasi yang dinyatakan dengan persamaan (32) atau (34). F_0 merupakan elemen frekuensi dari himpunan frekuensi yang memaksimalkan fungsi evaluasi. Him-



Gambar 3.9: *Flowchart* estimasi F_0

puan frekuensi ini berkisar dari 40 Hz hingga 400 Hz dengan resolusi 0.36 Hz. Pada fungsi evaluasi kita mensetting harga β sebesar 12 Hz, ω_0 sebesar 31.25 Hz (2 sampel frekuensi), dan ω_1 sebesar 5 kHz (320 sampel frekuensi).

Frekuensi sampling	16000 Hz
Inisialisasi L	1001 sampel (62.562 msec)
Fungsi window	Blackman dan Gaussian
Pergeseran window	32 sampel (2 msec)
Variasi α -Gaussian window	1-8, resolusi 0.25
Interval estimasi frekuensi	40-400 Hz, resolusi 0.36
Parameter β	12 Hz

Table 3.1: Tabel pengkondisian parameter-parameter estimasi F_0

Adapun pada formulasi IF Gaussian-window dimana terdapat parameter T yang harus ditentukan yang mana penentunya berdasarkan parameter α pada fungsi Gaussian. Hubungan antara parameter T dengan parameter α adalah $T \triangleq (L-1)/(2 \cdot \alpha)$ dengan L adalah panjang integral efektif fungsi Gaussian. Rentang parameter α yang digunakan berkisar 1 hingga 8 dengan resolusi 0.25. Harga α yang dipilih adalah yang memaksimalkan harga *harmonicity measure*. Sedangkan pada formulasi IF Blackman-window dimana pemilihan IF agar memaksimalkan harga *harmonicity measure* dengan memvariasikan separuh panjang integral efektif fungsi Blackman, $(L-1)/2$, sebesar 200, 225, 250, 300, 400, dan 500. Penentuan variasi panjang efektif integral fungsi Blackman ditentukan berdasarkan pada penelitian yang dilakukan pada [1][2]. Kemudian dari keseluruhan pengkondisian parameter-parameter pada estimasi F_0 berbasis IFAS dapat dirangkum dan dinyatakan sebagai tabel 1.

3.2 Klasifikasi *Voiced-Unvoiced*

Implementasi penentuan *voiced-unvoiced* pada sinyal suara dilakukan berdasarkan pada fungsi evaluasi F_0 IFAS, $\eta(F)$. *Voiced*

merupakan tipikal sinyal suara yang kaya dengan struktur harmonik sehingga pada tipikal suara ini akan memiliki *harmonicity measure* yang cukup signifikan jika dibandingkan dengan tipikal suara *unvoiced*. Tidak hanya pada harga *harmonicity measure*, tetapi, pada tipikal *voiced* memiliki level energi yang cukup tinggi jika dibandingkan dengan tipikal *unvoice*. Adapun strategi penentuan tipikal *voiced* atau *unvoiced* terbagi menjadi dua tahap, yaitu, penentuan harga ambang *voiced-unvoiced* dan penjejakan pada kontur F_0 .

Pada tahap penentuan harga ambang *voiced-unvoiced*, fungsi evaluasi F_0 IFAS pada tiap segmen sinyal suara dihitung terlebih dahulu. Setelah perhitungan fungsi evaluasi F_0 IFAS didapatkan, menghitung harga rata-rata fungsi evaluasi F_0 IFAS pada tiap segmen kemudian mensortir hitungan rata-rata fungsi evaluasi F_0 IFAS tiap segmen sehingga didapatkan kurva hitungan rata-rata monoton naik, $\eta_{sorted}(F)$. Untuk Harga ambang, dipilih pada segmen dimana kurva $\eta_{sorted}(F)$ memiliki harga maksimum pada turunan kedua. Akan tetapi, pada proses pencarian harga ambang pada kurva $\eta_{sorted}(F)$ yang mana harga ambang dibatasi tidak melebihi harga median $\eta_{sorted}(F)$ atau harga persentil yang ditetapkan. Dengan diperolehnya harga ambang, maka segmen sinyal suara dengan hitungan rata-rata $\eta(F)$ kurang dari harga ambang tergolong sebagai tipikal *unvoiced*. Kemudian pada tahap kedua dilakukan penjejakan pada kontur F_0 . Tujuan penjejakan kontur F_0 adalah memastikan kekontinyuan pada kontur dimana prosentasi deviasi F_0 dengan F_0 tetangga tidak melebihi prosentasi yang ditetapkan. Jika prosentase deviasi F_0 melebihi prosentasi yang ditetapkan, maka dilakukan eliminasi (*disetting* ke harga nol). Adapun jika prosentasi deviasi melebihi 10% maka tergolong sebagai gross error, sehingga jika yang diharapkan kontur

F_0 yang halus maka penentuan toleransi prosentasi deviasi haruslah kurang dari 10%. Setelah memastikan kekontinyuan pada kontur F_0 diperoleh kontur F_0 sepotong-sepotong. Pada kontur F_0 sepotong-sepotong, terdapat kontur F_0 yang bukan tergolong sebagai tipikal *voice* pada sinyal suara yang mana panjang dari kontur F_0 ini adalah kurang 13 segmen. Jika panjang kontur F_0 kurang dari 13, maka dilakukan eliminasi. Harga 13 segmen diperoleh berdasarkan *trial-and-error*. Setelah itu, penentuan tipikal *voice* pada sinyal suara adalah segmen-segmen sinyal suara yang padanya kontur F_0 berharga tidak nol.

"Halaman ini sengaja dikosongkan"

BAB IV HASIL DAN PEMBAHASAN

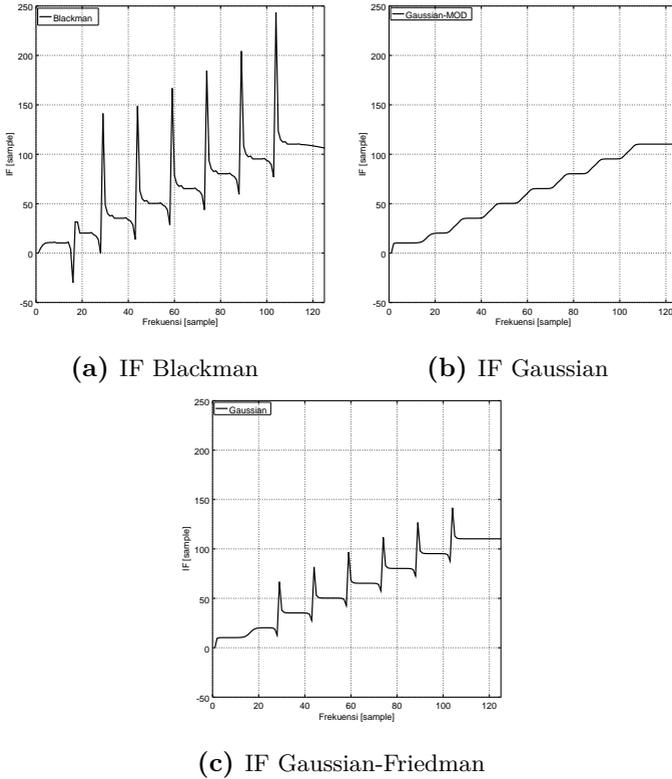
4.1 Estimasi F_0 berbasis IFAS

Pada tahap awal estimasi F_0 berbasis IFAS adalah melakukan analisa pada sinyal suara agar diperoleh representasi spektrum. Metode analisis yang digunakan adalah metode transformasi Fourier. Metode transformasi Fourier merupakan metode yang memandang sembarang sinyal input terutama sinyal input periodik sebagai hasil dari kombinasi linier fungsi-fungsi basis sinusoidal harmonik. Analogi pada konsep dasar kimia dimana sembarang material kompleks tidak lain merupakan tersusun atas atom-atom dasar yang tertulis pada tabel unsur periodik, atau analogi pada bangunan-bangunan rumah dimana tidak lain tersusun atas batu bata yang berbentuk sederhana. Sembarang sinyal input berperan sebagai material kompleks atau bangunan rumah sedangkan fungsi-fungsi basis sinusoidal harmonik berperan sebagai atom-atom dasar penyusun material kompleks atau sebagai batu bata sederhana yang menyusun berbagai bentuk arsitektur bangunan rumah.

Pada transformasi Fourier, sinyal-sinyal fungsi basis sinusoidal harmonik dikorelasikan pada sinyal input dengan melakukan operasi *sum-of-product* sehingga dihasilkan koefisien-koefisien yang disebut sebagai spektral. Kumpulan-kumpulan spektral inilah yang membentuk representasi spektrum. Semakin besar harga spektral yang dihasilkan, semakin besar pula kemungkinan fungsi basis sinusoidal sebagai penyusun dasar sinyal input. Pada representasi spektrum, disamping memberikan informasi mengenai eksistensi sinyal fungsi basis sinusoidal dalam sinyal input tetapi juga memberikan informasi mengenai distribusi energi yang terdapat sinyal input terhadap sinyal-sinyal penyusunnya. Selain itu,

dengan representasi spektrum kita dapat mengamati pola struktur harmonik sinyal input. Dengan teramatinya struktur harmonik sinyal input, kita dapat menentukan frekuensi dasar (F_0) sebagai karakteristik pada sinyal input.

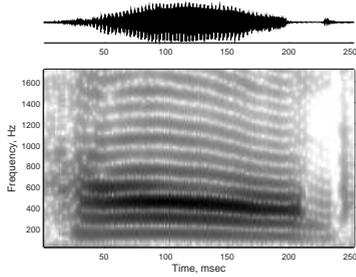
Pada sinyal suara yang mana merupakan sinyal non-stasioner yang memiliki properti akustik (distribusi *formant*, F_0 , dll) pada interval waktu yang pendek, Transformasi Fourier tidak dapat diterapkan langsung sedemikian rupa. Sebab, sinyal non-stasioner merupakan sinyal dengan representasi spektrum yang berubah seiring dengan perubahan waktu. Jika penerapan transformasi Fourier secara langsung dilakukan, maka pada proses korelasi yakni operasi *sum-of-product* akan memberikan hasil ambigu yang mana sinyal input tidak konsisten terhadap fungsi-fungsi basis sinusoidal penyusun. Oleh karena itu, sebelum melakukan penerapan transformasi Fourier pada sinyal suara, pada sinyal suara dilakukan proses partisi. Proses partisi pada sinyal suara adalah mengalikan sinyal suara dengan kumpulan fungsi-fungsi window yang disusun secara berdampingan yang mana fungsi window bisa saling tumpang tindih dengan fungsi window yang berdekatan. Transformasi Fourier diterapkan pada tiap-tiap partisi secara terpisah. Kombinasi transformasi Fourier dengan proses partisi disebut sebagai transformasi Fourier waktu-pendek (*Short-Time Fourier Transform*, STFT). Kemudian, hasil STFT dari tiap partisi sinyal disusun sedemikian rupa sehingga membentuk matriks yang mana tiap-tiap kolom matriks merupakan hasil STFT dari tiap-tiap partisi. Jika matriks hasil STFT ini diplot dengan garis ordinat sebagai domain frekuensi dan garis absis sebagai domain waktu dan elemen-elemen matriks direpresentasikan sebagai gradasi warna, maka diperoleh representasi spektrum yang disebut sebagai spektrogram. Dengan spektrogram, kita dapat



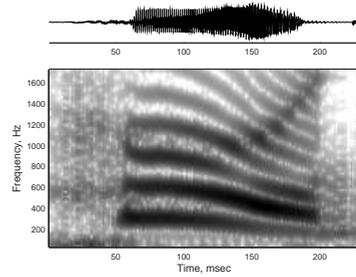
Gambar 4.10: Representasi hasil implementasi IF Blackman, dan IF Gaussian untuk input sinyal tonal

mengamati perubahan distribusi energi dan struktur harmonik tiap waktu lokal pada sinyal suara. Hasil spektrogram yang diperoleh dari data input suara dapat dilihat pada gambar 4.11 dan gambar 4.12

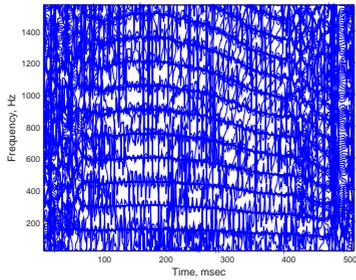
Tahap selanjutnya pada estimasi F_0 berbasis IFAS adalah



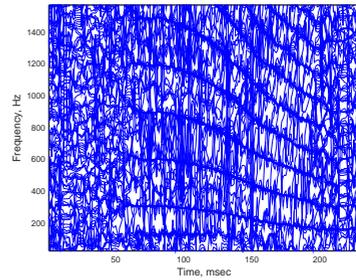
(a) Spektrogram vokal pria



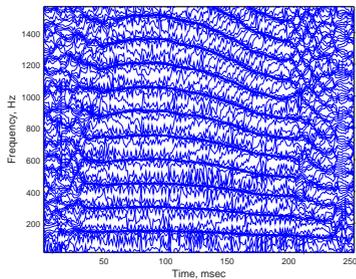
(b) Spektrogram vokal wanita



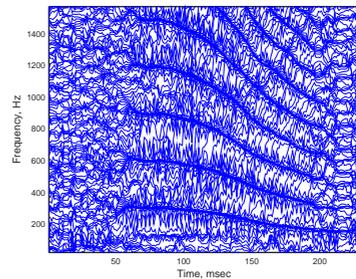
(c) IF Blackman vokal pria



(d) IF Blackman vokal wanita

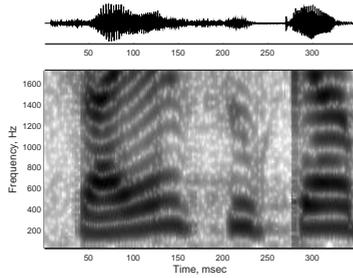


(e) IF Gaussian vokal pria

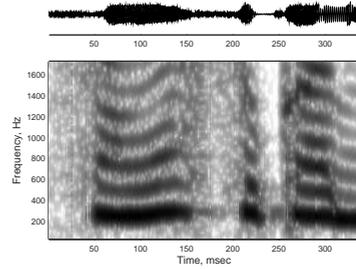


(f) IF Gaussian vokal wanita

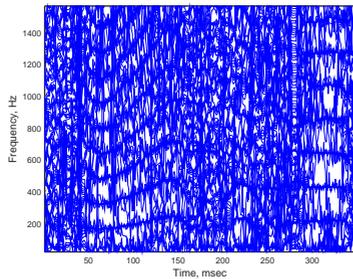
Gambar 4.11: Representasi hasil implementasi STFT, IF Blackman, dan IF Gaussian untuk input suara vokal



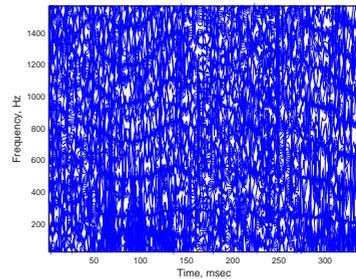
(a) Spektrogram campuran pria



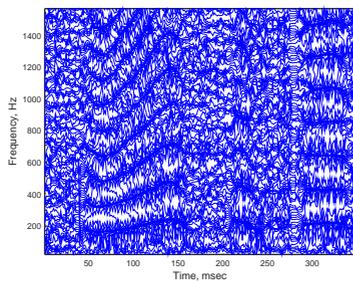
(b) Spektrogram campuran wanita



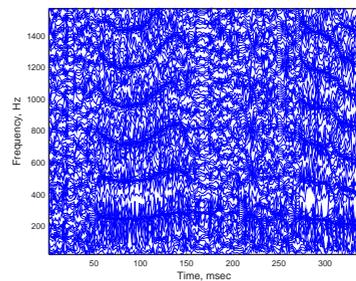
(c) IF Blackman campuran pria



(d) IF Blackman campuran wanita



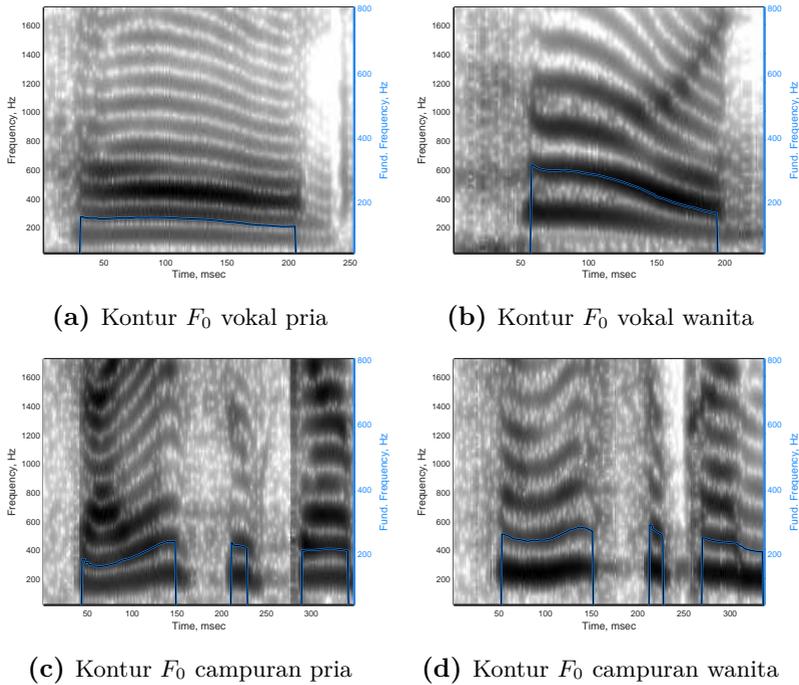
(e) IF Gaussian campuran pria



(f) IF Gaussian campuran wanita

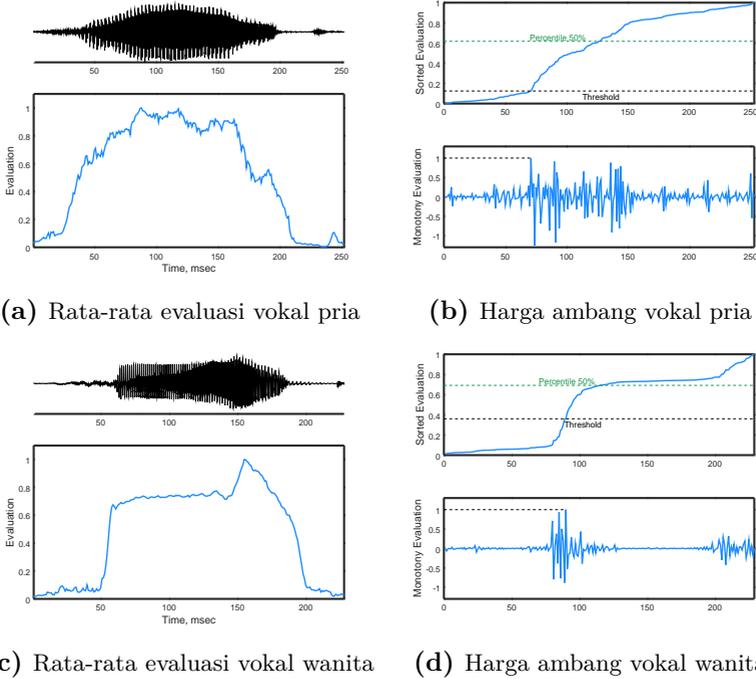
Gambar 4.12: Representasi hasil implementasi STFT, IF Blackman, dan IF Gaussian untuk input suara campuran

kalkulasi *Instantaneous Frequency* (IF) pada sinyal suara. Prosedur kalkulasi IF pada sinyal suara adalah sama dengan STFT yang mana kalkulasi diterapkan pada tiap partisi secara terpisah. Penjelasan detail mengenai IF dapat dilihat pada bab 2, namun, pada dasarnya adalah IF merupakan deskriptor baik untuk sinyal dengan representasi spektrum yang berubah terhadap waktu. Pada penelitian ugass akhir ini, mengusulkan formulasi IF dengan gaussian sebagai fungsi window yang mana pada penelitian sebelumnya menggunakan formulasi IF dengan Blackman sebagai fungsi window. Usulan formulasi IF pada penelitian tugas akhir ini merupakan modifikasi dari formulasi IF Gaussian window yang diusulkan oleh Friedman, dimana, modifikasi diterapkan pada koefisien T^2 yang secara detail dapat dilihat pada bab 2. Efek dari modifikasi koefisien T^2 dimana diskontinuitas pada kurva IF muncul ketika terjadi transisi dari frekuensi harmonik ke frekuensi harmonik berikutnya dapat direduksi secara substansial. Dampak munculnya diskontinuitas ini adalah menjadikan estimasi F_0 berbasis IFAS terbatas pada estimasi frekuensi tertentu yana mana estimasi F_0 berbasis IFAS pada penelitian sebelumnya yang dilakukan Arifianto terbatas pada frekuensi kF_s/N dengan F_s adalah frekuensi sampling, $k \in \mathbb{Z}$, dan N adalah ukuran partisi frekuensi transformasi Fourier. Dengan usulan formulasi IF Gaussian window dengan koefisien termodifikasi, estimasi F_0 berbasis IFAS tidak terbatas pada frekuensi tertentu sehingga resolusi frekuensi pada estimasi F_0 dapat ditingkatkan. Adapun tahap seleksi dimana perhitungan IF dilakukan secara berulang dengan mevariasikan lebar fungsi window sehingga diperoleh kurva IF dengan harga *harmonicity measure* yang tinggi mendekati harga maksimum normal. Hasil IF Blackman window dan IF Gaussian window dapat dilihat pada gambar 4.11 dan gambar 4.12



Gambar 4.13: Representasi hasil implementasi estimasi F_0 berbasis IFAS

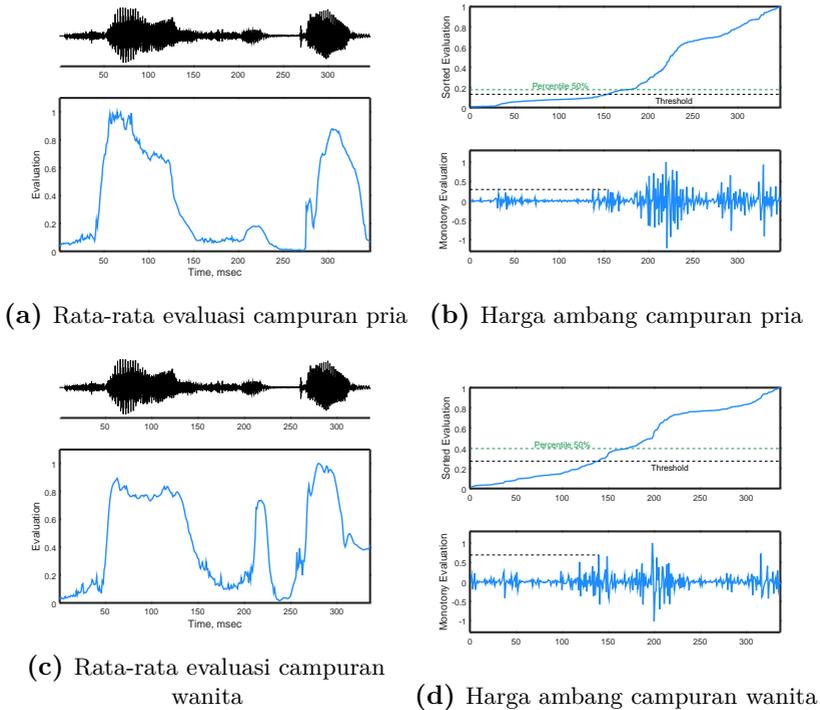
Tahap terakhir pada estimasi F_0 berbasis IFAS adalah proses evaluasi frekuensi. Secara detail mengenai evaluasi frekuensi pada estimasi F_0 berbasis IFAS dapat dilihat pada bab 2, sedangkan penentuan parameter-parameter pada proses evaluasi frekuensi dapat dilihat pada bab 3. Hasil dari estimasi F_0 berbasis IFAS dapat dilihat pada gambar 4.13.



Gambar 4.14: Representasi hasil implementasi penentuan harga ambang untuk suara vokal

4.2 Klasifikasi *Voiced-Unvoiced*

implementasi penentuan tipikal *voice-unvoice* pada data suara adalah dengan menggunakan metode *thresholding* dimana penentuan harga ambang berdasarkan harga kemonotonan kurva fungsi evaluasi rata-rata yang telah disortir terlebih dahulu. Harga ambang yang digunakan merupakan harga kemonotonan kurva fungsi evaluasi yang memiliki harga turunan kedua yang domi-

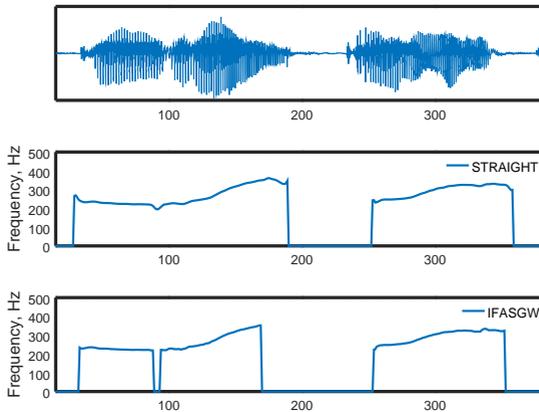


Gambar 4.15: Representasi hasil implementasi penentuan harga ambang untuk suara campuran

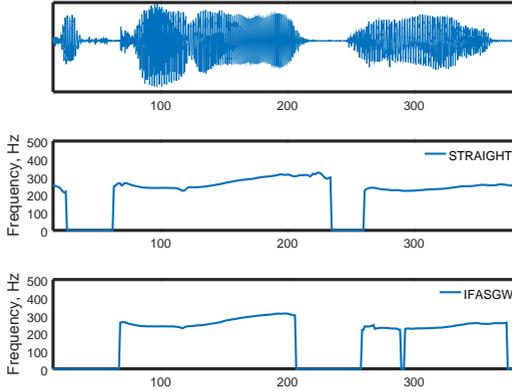
nan. Hasil penentuan harga ambang dapat dilihat pada gambar 4.14 untuk suara vokal dan gambar 4.15 untuk suara campuran. Kemudian untuk penjejakan kontur F_0 yang diperoleh dimana toleransi deviasi F_0 dengan F_0 tetangga *disetting* pada harga kurang dari 10% agar diperoleh kontur F_0 halus.

4.3 Perbandingan IFAS dengan STRAIGHT

STRAIGHT merupakan *software* berbasis pada MATLAB untuk analisis sinyal suara yang dikembangkan oleh profesor Hideki Kawahara, Wakayama University, Japan. Analisis sinyal suara STRAIGHT, mulai modifikasi hingga pada sintese sinyal suara, berdasarkan pada konsep VOCODER. Di dalam STRAIGHT, terdapat fungsi library untuk mengekstrak F_0 . Adapun pada penelitian tugas akhir yang dilakukan, membandingkan hasil estimasi F_0 STRAIGHT dengan hasil estimasi F_0 berbasis IFAS Gaussian window (klasifikasi tipikal *voiced-unvoiced* menggunakan metode harga ambang kemonotonan maksimum). Contoh hasil perbandingan STRAIGHT dengan IFAS Gaussian window dapat dilihat pada gambar berikut.



Gambar 4.16: Perbandingan F_0 STRAIGHT dengan F_0 IFAS Gaussian window pada naraswara wanita 'fala'.

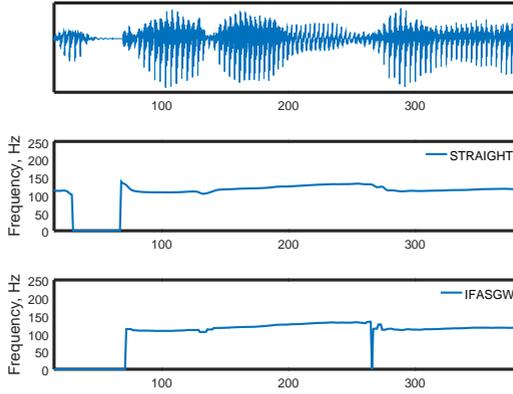


Gambar 4.17: Perbandingan F_0 STRAIGHT dengan F_0 IFAS Gaussian window pada naraswara wanita 'fena'.

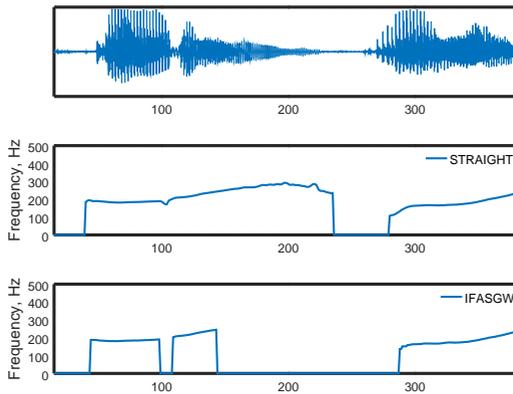
Perhitungan error estimasi F_0 Gaussian window didasarkan pada perhitungan *fine-pitch* error terhadap hasil estimasi F_0 STRAIGHT, yang diekspersikan sebagai,

$$Error = \frac{1}{K} \sum_{k=0}^{k=K} \frac{f - f_{ref}}{f_{ref}} \times 100\%,$$

dimana K adalah jumlah frame dengan tipikal *voiced*. Hasil perhitungan error pada pengujian sinyal suara dengan naraswara 2 pria dan 2 wanita yang masing-masing memiliki 20 pengucapan, diperoleh 4.9833% untuk naraswara wanita dan 6.2934% untuk naraswara pria.



Gambar 4.18: Perbandingan F_0 STRAIGHT dengan F_0 IFAS Gaussian window pada naraswara pria 'mmht'.



Gambar 4.19: Perbandingan F_0 STRAIGHT dengan F_0 IFAS Gaussian window pada naraswara pria 'mjra'.

BAB V KESIMPULAN DAN SARAN

4.1 Kesimpulan

Berdasarkan penelitian tugas akhir yang telah dilakukan, dapat disimpulkan bahwa representasi kurva IF yang diperoleh dari formulasi IF Gaussian window dengan koefisien T^2 yang termodifikasi dimana diskontinuitas yang terjadi pada saat transisi frekuensi harmonik ke frekuensi harmonik berikutnya dapat direduksi secara substansial. Dengan formulasi IF Gaussian window koefisien termodifikasi, estimasi F_0 berbasis IFAS tidak lagi terbatas pada frekuensi tertentu sehingga resolusi estimasi frekuensi dapat ditingkatkan. Dan Pengujian dari hasil estimasi F_0 berbasis IFAS Gaussian window pada database suara yang terdiri dari 2 pria dan 2 wanita dengan 20 pengucapan, menunjukkan hasil error *emphfine pitch* pria sebesar 6.2934% dan wanita sebesar 4.9833% dibandingkan dengan metode STRAIGHT.

4.2 Saran

Saran yang dapat disampaikan pada penelitian tugas akhir ini adalah pada klasifikasi *voiced-unvoiced* masih belum memberikan hasil yang memuaskan, dimana, terlihat pada sebagian hasil estimasi F_0 terdapat sinyal dengan pola periodik masih tergolong sebagai *unvoiced*. Adapun saran metode klasifikasi *voiced-unvoiced* yang digunakan adalah metode yang diusulkan pada jurnal [1] yang mana sebagai rujukan utama pada penelitian tugas akhir ini.

"Halaman ini sengaja dikosongkan"

DAFTAR PUSTAKA

- [1] T. Tanaka, T. Kobayasi, D. Arifianto, T. Masuko, “Fundamental frequency estimasi based on instantaneuous frequency spectrum”, Proc. ICASSP, vol-I, pp.329-332, Orlando, USA, May 2002.
- [2] D. Arifianto, and T. Kobayasi, “Performance Evaluation of IFAS-based fundamental frequency estimator in noisy enviromtment,” 4259 Nagatsuta, Midory-ku, Yokohama, Japan, 256-8502. 2002.
- [3] T. Abe, T. Kobayasi, and S. Imai, “Robust pitch estimation with harmonics enhancement in noisy enviromtments based on instantaneuous frequency,” in proc. ICSLP-96, Oct, 1996, pp.1277-1280.
- [4] T. Abe, T. Kobayasi, and S. Imai, “Harmonics estimation based on instantaneuous frequency and its application to pitch determination of speech,” IEICE Trans. On Information and systems, vol. E78-D, no.9, pp. 1188-1194, 1995.
- [5] D. Friedman, “Instantaneuous frequency distribution vs. time: an interpretation of the phase structure of speech”, ICASSP 1985, pp. 1121-1124.
- [6] M. Honda, NTT CS Laboratories, Speech synthesis technology based on speech production mechanism, How to observe and mimic speech production by human, Journal of the Acoustical Society of Japan, Vol. 55, No. 11, pp. 777-782, 1999 (in Japanese).

- [7] B. Boashash, "Estimating and Interpreting The instantaneous Frequency of a Signal-Part 1: Fundamentals," proc. Of the IEEE, vol. 80. no. 4, April, 1992.
- [8] Hideki Kawahara, Ikuyo Masuda-Katsuse, and Alain de Cheveigne. Restruturing speech representations using pitch-adaptive time-frequency smoothing and an instantaneous-frequency-based F0 extraction. *Speech Communication*, Vol. 27, No. 3-4, pp. 187-207, 1999.

BIODATA PENULIS



Dipta Nusbangga Haikal Ahmada lahir di Banyuwangi pada tanggal 18 Agustus 1992 sebagai anak pertama dari tiga bersaudara. Penulis sudah menunjukkan prestasi negatif sejak menempuh pendidikan dasar di SD Al-Irsyad Al-Islamiyyah karena salah satu dari murid-murid yang bermasalah mulai dari kelas satu hingga kelas enam. Kebermasalahan penulis terus berlanjut hingga bangku sekolah menengah pertama, yaitu di

SMP 1 Glagah Banyuwangi. Penulis merupakan salah satu dari kelompok orang kikuk, bermimpi ingin menjadi orang keren namun tidak terwujud, tidak berprestasi, hingga pada tahap sebagai objek bulan-bulanan oleh teman-teman lain hanya karena tidak keren. Penulis merasa putus asa dan tidak ada gairah hidup karena keadaan ini berlangsung hingga pada saat kelulusan jenjang pendidikan menengah. Penulis terus-menerus merasa bimbang dan ingin sekali mengetahui jalan keluar apa yang harus diambil, hingga akhirnya penulis menemukan sebuah buku. Buku ini berceritakan tentang nabi Sulaiman a.s., yang mana, buku ini mengisahkan bagaimana nabi sulaiman mendapatkan mukjizat dari Allah SWT. Buku ini bercerita, sebelum Allah memberikan mukjizat pada nabi sulaiman, Allah memberi kepada nabi Sulaiman lima pilihan. Pilihan pertama adalah harta, yang kedua adalah tahta atau kedudukan, yang ketiga adalah kegagahan dunia, yang keempat adalah pasukan atau kekuatan, dan yang

terakhir adalah ilmu. Nabi Sulaiman terdiam sejenak, hingga pada akhir pilihan nabi Sulaiman jatuh pada pilihan terakhir yakni ilmu. Karena ilmu yang dipilih maka Allah memberikan semua pilihan kepada nabi Sulaiman. Penulis merasa terinspirasi dengan cerita ini, dan pada jenjang pendidikan menengah keatas, yaitu MAN 1 Banyuwangi, penulis berusaha keras dan rajin dalam menuntut ilmu. Dan pada akhirnya penulis menjadi salah satu murid berprestasi hingga pernah menjuarai olimpiade fisika tingkat regional sebagai *runner up* ke dua. Setelah lulus jenjang pendidikan atas di MAN 1 Banyuwangi, Penulis melanjutkan pendidikan di ITS jurusan Teknik Fisika hingga saat ini namun dengan waktu tempuh perkuliahan yang sedikit molor. Namun, hal itu semata-mata untuk mencari ilmu dan keridhoan Allah SWT.