

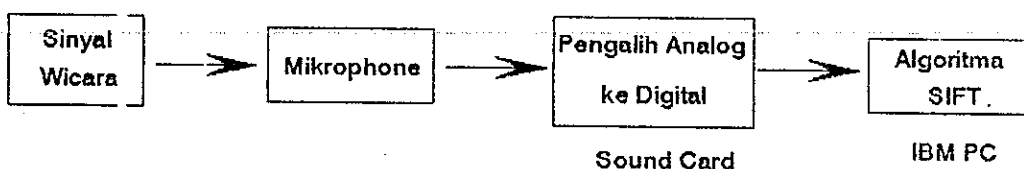
BAB III

METODE PENELITIAN

3.1. PENGANTAR

Untuk merealisasikan sistem algoritma SIFT, disamping software diperlukan pula perancangan sistem pengambilan data digital sinyal wicara. Sistem pengambilan data digital ini direalisasikan dengan hardware dan software.

Secara garis besar, konfigurasi sistem algoritma SIFT dapat dilihat pada gambar 3.1. Pada gambar tersebut, sinyal wicara dari Mikrophone dikondisikan lebih dahulu agar sesuai dengan kondisi yang diminta oleh masukan ADC yaitu melalui Sound Card. Sound Card ini dihubungkan dengan Komputer melalui slot I/O dan cara kerjanya dikontrol dengan Software dari Komputer IBM PC. Data digital hasil akusisi ini disimpan pada suatu file dalam Disket dan siap diolah lebih lanjut. Pengolahan data sinyal wicara tersebut diolah dengan Komputer secara software.



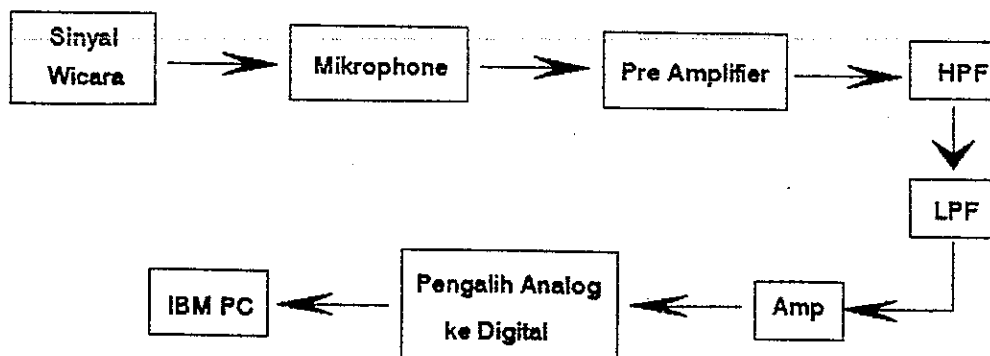
Gambar 3.1.

3.2. PERANCANGAN PERANGKAT KERAS

Perangkat keras yang diperlukan adalah perangkat keras untuk akuisisi data, yaitu untuk mengambil data sinyal wicara dari Mikrophone dan mengubah sinyal analog menjadi data diskrit. Realisasinya diwujudkan dengan bantuan Sound Card yang dihubungkan dengan Komputer IBM PC melalui slot I/O.

Bagan kotak sistem pengambilan data digital wicara dapat dilihat pada gambar 3.2. Sinyal masukan diumpankan di sebuah Mikrophone audio. Dalam pemakaiannya Mikrophone harus diletakkan dekat mulut Pembicara agar diperoleh sinyal dengan cukup kuat dan noise lingkungannya ikut terambil.

Sinyal wicara yang sudah tertangkap masih harus dikondisikan lebih dahulu. Pertama harus diperkuat dahulu karena Amplitudo sinyal dari Mikrophone masih sangat lemah, kemudian spektrum frekuensi sinyal wicara dibatasi dengan menggunakan LPF. (Low Pass Filter) agar tidak terjadi efek aliasing dengan frekuensi pencuplik.

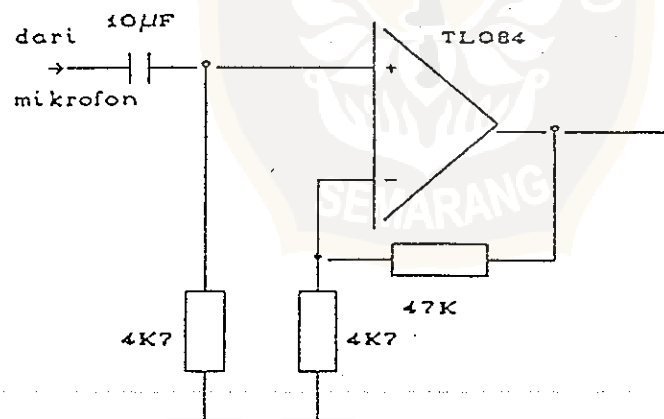


3.2.1. Penguatan Awal

Sinyal dari Mikrophone masih sangat lemah sehingga harus diperkuat lebih dahulu sebelum diumpankan ke LPF. Faktor penguat dapat diatur sampai 10 kali (20db) dengan konfigurasi Inverting Amplifier.

Mikrophone harus diletakkan didekat mulut Pembicara agar diperoleh S / N yang cukup besar yaitu arus sinyal wicara jauh lebih tinggi dibanding dengan aras noise lingkungan. Hal ini penting karena pengambilan data sinyal wicara dilakukan langsung dalam ruang kerja.

Untai Pre Amplifier tersebut dapat dilihat pada gambar 3.3.

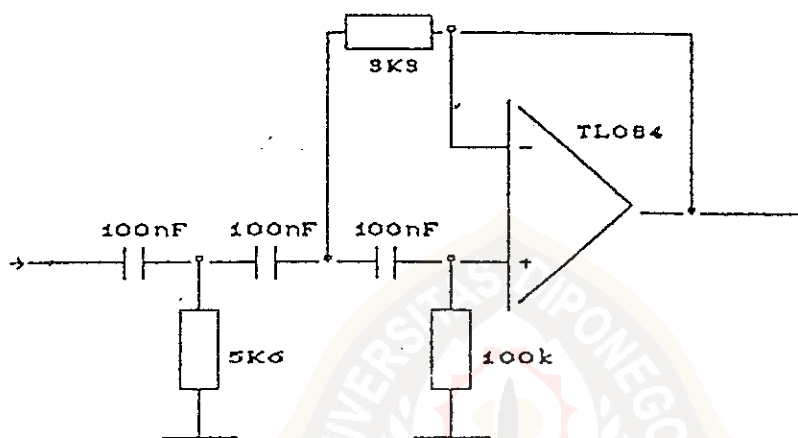


gambar 3.3

3.2.2. HPF (High Pass Filter)

Filter ini digunakan untuk menghambat suara yang bukan ucapan agar tidak ikut masuk, seperti dengung 50 Hz

yang ditimbulkan oleh jala-jala. Filter direalisasikan dengan sebuah op-amp agar diperoleh filter berorde-3 (18 db per oktaf) dengan frekuensi cut-off sekitar 300 Hz dan penguatan sama dengan satu. Untai filter ini dapat dilihat pada gambar dibawah ini (gambar 3.4.).

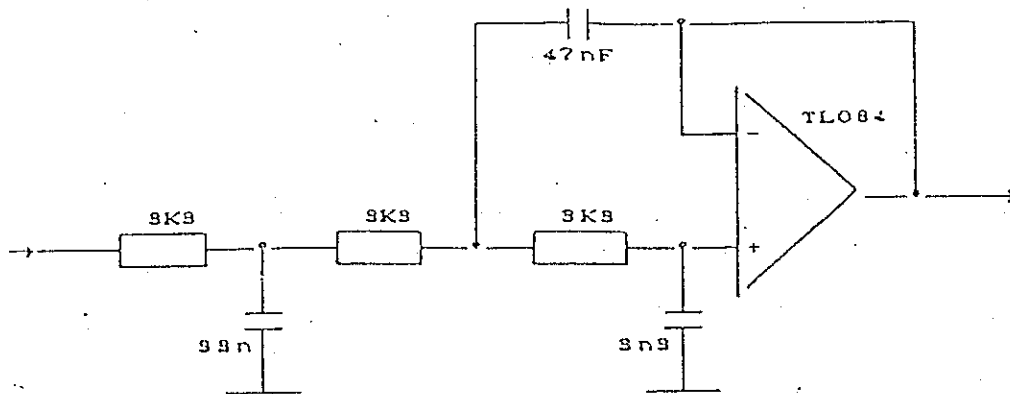


Gambar 3.4.

3.2.3. LPF (Low Pass Filter)

Sinyal wicara yang diambil harus dibatasi spektrum frekuensinya agar tidak terjadi efek aliasing dengan frekuensi pencuplik. Frekuensi pencuplik minimal 2 kali frekuensi sinyal wicara yang paling tinggi. Frekuensi manusia umumnya dapat dibatasi pada frekuensi 3400 Hz tanpa mengurangi kejelasan ucapan. Dengan demikian frekuensi cut off berorde-3 ini dibuat berada pada sekitar 3000 Hz dengan penguat sebesar satu kali.

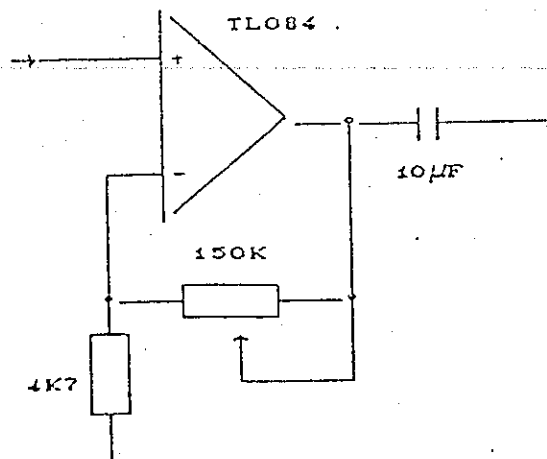
Untai filter aktif tersebut dapat dilihat pada gambar dibawah ini (gambar 3.5.)



gambar 3.5

3.2.4. Penguatan Akhir

Setelah sinyal dikondisikan, dan siap untuk dibaca maka tahap terakhir adalah memperkuat keseluruhan sinyal agar amplitudonya sesuai dengan jangkauan tegangan pada rangkaian sampling, yaitu sebesar 5 Vpp. Untuk itu penguatan dapat diatur sesuai dengan kebutuhan pada saat pengambilan yaitu berkisar antara 1 - 30 kali (0-30 db). Penguat dibentuk dengan konfigurasi Non_Inverting Amplifier yang dapat dilihat pada gambar dibawah ini :



3.2.3. PENGALIH ANALOG KE DIGITAL

Untuk mengubah sinyal analog ke digital digunakan bentuk Sound Card, yaitu *high performace data aquistition card*. Sound Card merupakan analog/digital I/O Card sehingga mampu bertindak sebagai ADC dengan resolusi 16 bit.

Pada TA ini penulis menggunakan Sound Card sebagai pengalih analog ke digital dengan spesifikasi :

- resolusi 16 bit
- frekuensi pencuplik antara 5 - 44,1 kHz

Sound Card mempunyai register yang memiliki alamat relatif terhadap alamat dasarnya. Register-register tersebut sangat berperan dalam pekerjaan yang dilakukan oleh Sound Card. Karena kontak terhadap Sound Card dilakukan oleh perangkat lunak melalui I/O port-nya, maka alamat dasar dari I/O port-nya ditentukan lebih dahulu agar nantinya dapat diakses. Bardasar alamat di Komputer IBM PC, dipilih alamat \$220 sebagai alamat dasar, karena biasanya port ini digunakan untuk suara(SCG) adalah port-port DSP (*Digital Signal Prossesor*). Nomor-nomor yang dapat digunakan adalah :

- \$226 untuk mereset DSP
- \$22A untuk membaca data di DSP
- \$22B untuk menulis perintah/data di DSP
- \$22C untuk membaca status tulis (7 bit)
- \$22E untuk membaca status data (7 bit)

3.3. PERANCANGAN PERANGKAT LUNAK

3.3.1. PERANCANGAN PERANGKAT LUNAK UNTUK PENGAMBILAN DATA DIGITAL SINYAL WICARA.

Program pengambilan data digital sinyal wicara dirancang dengan memanfaatkan driver routine yang sudah ada pada Sound card.

Setelah inisialisasi dan saluran masukan ditentukan maka langkah berikutnya menentukan frekuensi pencuplik, Frekuensi pencuplik yang diinginkan sebesar 8 KHz. Maka proses pengambilan data dapat dimulai. Setelah selesai proses pengambilan data, program meminta nama file tempat sampel disimpan. Data sampel wicara tersebut diolah oleh program SIFT (lihat diagram Alir).

3.3.2. PERANCANGAN PERANGKAT LUNAK ALGORITMA SIFT

3.3.2.1. PENDAHULUAN

Program ini bertujuan untuk mengekstrasi parameter sinyal wicara dari suatu data sinyal wicara yang disimpan pada suatu file dalam disket. Parameter sinyal wicaranya adalah periode Pitch dan frekuensi dasar serta pernyataan voiced / unvoiced, yang diperoleh dengan penganalisaan tiap frame data.

Pada awal program dibutuhkan pembacaan data input yang akan diproses serta mengambil isi file itu untuk mengolah data lebih lanjut.

Mula-mula data sinyal harus dilewatkan ke LPF untuk mencegah timbulnya aliasing saat didesimasi. Dalam hal ini LPF direalisasikan dengan procedure LPF. Setelah ini

data keluaran LPF siap di desimasi. Karena proses analisis wicara dilakukan untuk tiap frame data maka harus dibuat frame lebih dahulu, artinya data sinyal wicara dibagi dalam segmen-segmen. Banyaknya frame tergantung dari panjangnya sinyal wicara. Tiap frame diinginkan sebanyak 240 sampel. Berarti untuk sinyal wicara selama 1 sekon ada sebanyak 33 frame data. Frame 0 terdiri atas sampel ke-0 sampai ke-239, frame 1 terdiri atas sampel ke 240 sampai ke 479 dan seterusnya sampai frame ke 32. Selanjutnya proses dilakukan untuk tiap-tiap frame. Pertama diambil sampel ke-0 sampai sampel ke 239 sebagai frame 0, kemudian untuk frame ini masing-masing sampel dikurangi dengan sampel sebelumnya dan dikalikan dengan fungsi Hamming Window. Selanjutnya dihitung koefisien prediksinya dengan Procedure Hitung Parameter LPC. Kemudian data dilewatkan ke Inverse Filter dan dikalikan dengan fungsi Hamming Window, setelah itu di autokorelasikan untuk mengestimasi periode pitchnya dengan mencari puncak pada frame tersebut agar didapat hasil yang lebih akurat maka dilakukan Interpolasi disekitar puncak tersebut sehingga diperoleh periode pitch yang lebih teliti. Kemudian dari periode pitch tersebut dilakukan klasifikasi Voiced / Unvoiced. Proses berulang untuk frame berikutnya sampai frame berakhir.

3.3.2.2. Procedure LPF

Pada procedure ini masukan berupa deretan data sinyal wicara $S(n)$ dan keluarannya berupa deretan data $X(n)$

setelah dilewatkan ke LPF. Karena LPF yng digunakan merupakan Filter digital, maka realisasinya menggunakan persamaan matematis yaitu persamaan filter digital seperti yang telah disebutkan pada bab sebelumnya, persamaannya adalah :

$$U(n) = A_1 S(n) + A_2 U(n-1)$$

$$X(n) = A_3 U(n) + A_4 X(n-1) + A_5 X(n-2)$$

3.3.2.3. Prosedur Desimasi

Desimasi adalah penurunan kecepatan cuplikan sebesar m faktor m . Pada prosedur ini $m = 4$. Jadi kecepatan cupliknya diturunkan sebesar $1/4$ kalinya. Jika semula kecepataannya cuplikannya sebesar 8kHz, maka setelah di desimasikan menjadi 2 kHz. Jumlah datanya menjadi $1/4$ kalinya jumlah data semula. Desimasi direalisasikan dengan mengambil 1 sampel dari tiap m sampel. Pada program ini 1 sampel diambil dari tiap 4 sampel. Dengan jumlah data yang lebih sedikit dan akan mempercepat proses perhitungan, disamping itu juga akan menghemat memori penyimpanan data selama program berjalan. Pada prosedur ini masukan berupa deret data $X(n)$ yang merupakan data sinyal wicara setelah di filter. Sedang keluarannya berupa deret baru $W(n)$ yang merupakan deretan hasil desimasi.

3.3.2.4. Prosedur Buat Frame

Sampel data diambil mulai $n = 0$ sampai $n = (\text{jumlah sampel} - 1)$. Bila jumlah sampel dalam satu frame diinginkan sebanyak 240 sampel, maka dilakukan pengambilan

sampel dari data $n = 0$ sampai $n = 239$. Data yang disimpan dalam file tersebut merupakan segmen data sehingga memiliki besaran positif dan negatif. Pada procedure ini masukannya adalah deretan data $w(n)$ dan keluarannya adalah deretan data $P_j(n)$ untuk tiap frame

3.3.2.5. Procedure Hitung Parameter LPC

Yaitu untuk menghitung koefisien prediksi. Prosedurnya sebagai berikut :

1. Menghitung autokorelasi sinyal wicara $P_j(n)$ sesuai dengan persamaan berikut (yang sudah dibagi dengan jumlah sampel dalam satu frame)

$$\hat{r}_i = \frac{1}{N} \sum_{n=i}^{n-1} P_j(n) P_j(n-1)$$

dengan

\hat{r}_i = sampel autokorelasi ke- i

N = jumlah sampel dalam satu frame

$P_j(n)$ = sampel saat ke- n

$P_j(n-1)$ = data sampel data saat ke- i sebelum n

karena ada p orde maka i bergerak dari 0 sampai p .

Untuk $P = 4$ maka i bergerak dari 0 sampai 4. Jadi

dihitung nilai autokorelasi sampai r_4

2. Untuk $n = 0$ maka error ke-0, $E_0 = r_0$

3. Untuk $i = 1$ sampai P

- a. Hitung koefisien prediksi $a(n)$

untuk tiap-tiap tahapan

- b. hitung error untuk tahap berikutnya

3.3.2.6. Procedure Invers Filter

Masukan procedure in adalah masukan data $P_j(n)$ dan

koefisien prediksi a_1 yang sudah dihitung dalam procedure Hitung Parameter LPC. Keluaran procedure adalah deretan data $y(n)$ yang merupakan keluaran invers filter. Invers filter adalah filter digital dengan persamaan filternya seperti ditunjukkan oleh persamaan 4.12. Dengan demikian procedure ini akan menghitung keluaran inver filter untuk tiap frame data dengan persamaan :

$$Y(n) = P_j(n) + \sum_{i=1}^n a_i P_j(n-1)$$

$$n = 0, 1, 2, 3, \dots, N-1$$

dengan

$Y(n)$ = sampel data keluaran invers filter saat ke n

$P_j(n)$ = sampel data masukan invers filter

a_i = koefisien prediksi

P = orde koefisien prediksi

N = jumlah sampel dalam frame

Kemudian keluaran invers filter ini dikalikan dengan fungsi Hamming Window.

3.3.2.7. Procedure Autokorelasi

Procedure ini menghitung nilai autokorelasi dari deretan data $Y(n)$ yaitu keluaran invers filter, dengan persamaan :

$$r_n(n) = \sum_{j=0}^{N-1-n} Y(j) Y(j+n)$$

dengan

$r_n(n)$ = sampel autokorelasi ke $-n$

$Y(j)$ = sampel data keluaran invers filter ke- j

N = jumlah sampel dalam satu frame

Setelah sampel di autokorelasi $r_n(n)$ diperoleh harga tertentu, nilainya harus dinormalisasi dahulu terhadap r_0 yang merupakan nilai autokorelasi terbesar. Setelah dinormalisasi, dicari nilai puncak yaitu nilai autokorelasi tertinggi. Letak puncak inilah yang mengandung informasi periode pitchnya

3.3.2.8 Prosedur Interpolasi

Pada awal prosedur didefinisikan τ_0 = puncak. Nilai puncak yang mengandung informasi periode pitch ini akan diinterpolasi agar didapat periode pitch yang lebih akurat, caranya dengan menghitung nilai disekitar puncak yaitu : $\tau_{1/4}$, $i = 0, \pm 1, \pm 2$ Mula-mula nilai $\tau_{1/4}$ dan $\tau_{-1/4}$ dibandingkan.

Jika $\tau_{1/4} > \tau_{-1/4}$ maka nilai puncak dicari diantara ketiga nilai, yaitu τ_0 , $\tau_{1/4}$ dan $\tau_{1/2}$.

Sebaliknya jika nilai $\tau_{1/4} < \tau_{-1/4}$ maka dibandingkan nilai τ_0 , $\tau_{-1/4}$ dan $\tau_{-1/2}$. Setelah diperoleh nilai maka letak puncak tersebut merupakan informasi periode pitch yang sesungguhnya.

3.3.2.9. Prosedur Klasifikasi Voiced / Unvoiced

Klasifikasi *Voiced/Unvoiced* untuk tiap frame data didasarkan atas besarnya nilai puncak yang didapat dari prosedur interpolasi. Mula-mula nilai puncak pada frame k dibagi dengan faktor D seperti yang ditunjukkan oleh persamaan (4.60a dan 4.60b) sebagai penerapan nilai ambang variabel. Kemudian nilai puncak yang baru, dibandingkan

dengan dengan nilai ambang pertama. Pada tugas akhir ini nilai ambang pertama sebesar 0,35. Jika lebih besar, maka frame k voiced, sebaliknya bila puncak lebih kecil dari nilai ambang, maka perlu dilihat dahulu apakah frame (k-1) unvoiced . . Jika frame (k-1) unvoiced untuk frame k juga Unvoiced. Sebaliknya jika frame (k-1) voiced maka nilai ambang diturunkan menjadi 0,3. Jika puncak masih kecil dari nilai ambang ini maka frame adalah unvoiced.

Pada procedure ini juga dilakukan proses koreksi terhadap pernyataan yang salah, dimana suatu frame yang seharusnya Voiced dinyatakan Unvoiced dan sebaliknya yaitu pada kasus frame unvoiced diapit oleh frame Voiced, padahal berdasar kondisi stasioner hal ini tidak mungkin seperti dijelaskan pada bab sebelumnya.

