

**PEMBELAJARAN ANAK BERKEBUTUHAN KHUSUS
BERDASARKAN MODEL PENGENALAN SUARA
MENGUNAKAN MATLAB DAN MIKROKONTROLER
ATMEGA16**

Beni Utomo¹, Turahyo², Bagus Priyo Tomo³
^{1,2,3} STITEK Bontang

¹mir4kel@gmail.com, ²turahyoahyo@ymail.com, ³mistersains@gmail.com

Abstrak

Penelitian diawali dengan merekam suara huruf vokal A, I, U, E dan O menggunakan mikrophone yang terhubung dengan MATLAB. Setiap lafal yang terekam selanjutnya diekstrak dan dianalisa untuk mengetahui karakteristik dari sinyal suara yang diperoleh dengan menggunakan formant. Formant merupakan frekuensi alami atau resonansi yang dihasilkan dari jejak vokal ketika seseorang berbicara. Metode yang digunakan untuk mengestimasi formant menggunakan analisis Cepstral. Analisis Cepstral dilakukan dengan cara menghitung spektrum dari Transformasi Fourier, dengan menyatakan invers Transformasi Fourier dari bentuk logaritmik spektrum energinya, dan dengan hanya mengambil koefisien orde rendah (8 sampai 16) dari inversnya. Hasil dari proses tersebut adalah sinyal cepstrum. Periode puncak dari sinyal tersebut diestimasi dan log magnitude yang *smooth* diperoleh dari cepstrum. Formants diestimasi dari spektral tersebut menggunakan syarat-syarat pada kisaran frekuensi formant dan ketinggian relatif dari puncak spektral pada frekuensi formant.

Selanjutnya berdasarkan hasil tersebut klasifikasi untuk mengenal huruf A, I, U, E, dan O berupa sinyal dilakukan dan bisa dibuat suatu basisdata. Mikrokontroler dipakai untuk menampilkan suatu sinyal yang masuk setelah sebelumnya dicocokkan dengan data yang telah dibuat melalui dua alat, yaitu melalui layar LCD dan melalui lampu LED..

Kata kunci: formant, fourier, cepstral, mikrokontroler, LED, LCD

A. PENDAHULUAN

Penelitian ini akan lebih diarahkan pada masalah komunikasi bicara anak berkebutuhan khusus terutama pada anak yang berusia diatas 2 tahun dan dibawah 5 tahun. Pada usia tersebut baru bisa diketahui apakah seorang anak memang mempunyai gangguan wicara atau hanya mengalami penundaan bicara yang biasa terjadi pada anak-anak. Untuk anak-anak yang mempunyai gangguan wicara, biasanya terapi wicara dilakukan oleh pengajar atau terapis dengan perbandingan 1:1 yaitu 1 terapis untuk 1 anak. Kegiatan terapi salah satunya dilakukan dengan pengenalan huruf abjad. Secara manual terapis meminta anak untuk mengucapkan suatu huruf tertentu dan anak akan mengucapkan atau menirukan lafal tersebut. Pada bagian inilah penelitian ini akan dilaksanakan dan sebagai batasan masalah huruf yang akan digunakan adalah huruf vokal yaitu a,i,e,o, dan u.

Pada saat anak melafalkan suatu huruf, maka kriteria benar tidaknya atau bisa tidaknya anak melafal akan sangat tergantung dari pendengaran terapisnya. Hal tersebut bisa menimbulkan beberapa interpretasi dari apa yang telah dilafalkan anak. Kondisi tempat anak berkebutuhan khusus biasanya tidak setenang tempat belajar anak biasa. Dengan bantuan teknologi yang akan

dibuat melalui penelitian ini, cara memberikan penilaian terhadap lafal anak dalam mengucapkan huruf atau kata akan dianalisis menggunakan komputer. Dengan demikian akan diketahui dengan lebih jelas, pada bagian mana anak mempunyai masalah pada pelafalannya. Dengan data tersebut, terapis atau pengajar tentunya mempunyai informasi yang lebih dari cukup untuk menentukan tindakan selanjutnya dalam memberikan pembelajaran pada anak, karena kondisi anak satu dengan lainnya bisa sangat berbeda.

Tujuan utama penelitian ini adalah menghasilkan suatu alat yang bisa mengukur pelafalan huruf abjad dari anak-anak yang mempunyai keterbatasan dalam berbicara. Harapannya, dengan adanya alat ini, para pendidik dan pengajar anak berkebutuhan khusus mempunyai alat untuk mengenali pada bagian mana anak kesulitan untuk melafalkan huruf, dengan demikian para pendidik dan pengajar bisa memberikan terapi wicara dengan lebih tepat .

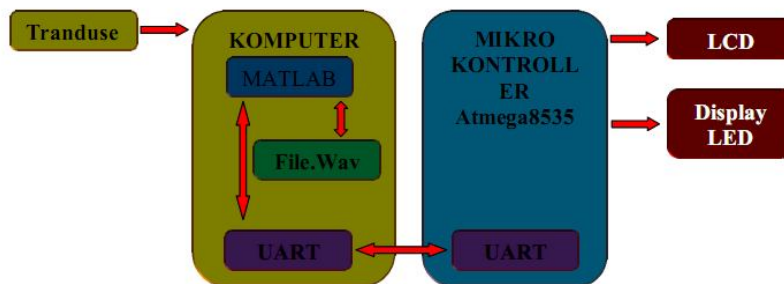
B. METODE PENELITIAN

Penelitian ini dilakukan melalui dua tahapan yang bisa dilaksanakan secara terpisah namun saling mendukung. Tahapan pertama merupakan tahapan analisis suara menggunakan MATLAB, sedangkan tahap kedua adalah tahapan menyusun peralatan yang bisa dipakai anak untuk bermain dan belajar menggunakan mikrokontroler ATMEGA16 atau ATMEGA8535.

Desain Sistem Secara Umum

Sistem yang akan dibuat terdiri dari dua bagian besar, yaitu komputer yang didalamnya sudah berisi program MATLAB serta mikrokontroler seri ATMEGA16 atau ATMEGA8535. Sistem bekerja dengan cara sebagai berikut. Komputer terhubung dengan mikrophone dan mikrokontroler. Mikrophone berfungsi sebagai alat untuk memasukkan suara menjadi sinyal yang selanjutnya oleh MATLAB sinyal suara tersebut disimpan dalam file berekstensi .wav. file .wav tersebut selanjutnya diekstrak oleh MATLAB untuk mengetahui kekhasan dari sinyal tersebut. proses dilanjutkan dengan melakukan sintesis dari sinyal tersebut. kegiatan tersebut dilakukan berulang-ulang hingga terbentuk basis data sinyal suara.

Mikrokontroler berperan dalam menampilkan menampilkan hasil dari sinyal tersebut melalui LCD atau dot matrix. Data yang telah dibuat tadi selanjutnya ditanamkan pada mikrokontroler menggunakan downloader. Setelah program berhasil ditanam maka mikrokontroler bisa tidak dihubungkan dengan komputer.



Gambar 1 Desain Sistem

Analisis Suara

Suara yang direkam menggunakan mikrophone akan dianalisa menggunakan MATLAB. Analisa sinyal tersebut meliputi proses ekstraksi, identifikasi, dan sintesis. Proses ekstraksi dilakukan menggunakan formant. Formant merupakan frekuensi alamiah atau resonansi yang dihasilkan ketika seseorang berbicara. Ada beberapa metode untuk mengekstrak formant dari suatu sinyal. diantaranya adalah metode cepstrum dan LP (Linear Predictive). Berikut adalah sinyal lafal huruf

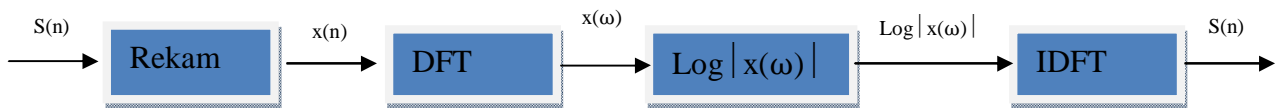
vokal yang direkam menggunakan MATLAB dengan sumbuX menyatakan domain waktu dalam detik dan sumbuY menyatakan amplitudo.

File yang tersimpan di MATLAB tersebut selanjutnya diproses menggunakan transformasi Fourier. Transformasi Fourier merupakan suatu transformasi yang dikenakan pada sinyal yang hanya terjadi satu kali dan tentu saja tidak bersifat periodik. Untuk sinyal yang periodik analisisnya menggunakan deret Fourier sedangkan untuk sinyal yang tidak periodik dianalisis menggunakan integral Fourier. Transformasi Fourier mempunyai formula (1) sedangkan inversnya (2):

$$\hat{f}(\omega) = \mathfrak{F}(f) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)e^{-i\omega x} dx \quad \dots\dots\dots(1)$$

$$f(x) = \mathfrak{F}^{-1}(f) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(\omega)e^{i\omega x} d\omega \quad \dots\dots\dots(2)$$

Secara umum proses pengolahan sinyal hingga mendapatkan formant ditunjukkan oleh blok diagram berikut ini.



Gambar 1 Blok Diagram menghitung Spektrum Sinyal (Cepstrum)

Suatu cepstrum merupakan hasil dari proses invers transformasi fourier (IDFT) dari suatu bentuk logaritmik untuk estimasi spektrum suatu sinyal. ada cepstrum kompleks, cepstrum real, cepstrum tenaga (power), dan cepstrum phase. Power cepstrum banyak ditemukan dalam aplikasi analisis bicara manusia.

Metode formant ini dilakukan dengan cara membagi sinyal menjadi blok-blok waktu. Selanjutnya dengan menggunakan blok-blok tersebut disusun model AR untuk setiap blok dan pilih 3 frekuensi positif pertama untuk pole sebagai frekuensi formant.

Transformasi Fourier Diskret/ DFT(Discrete Fourier Transform)

Sinyal yang direkam merupakan sinyal waktu kontinu sehingga sebelum diproses sinyal perlu diubah dulu menjadi sinyal diskret. Proses perubahan tersebut bisa dilakukan dengan cara mengambil titi-titik sampel atau *sample-point* dari sinyal kontinu dengan interval waktu tertentu. Berikut adalah bentuk sinyal waktu diskret.

Untuk fungsi $f(x)$, misalkan panjang sinyal adalah 2π , untuk sejumlah N pengukuran, dari $f(x)$ diambil interval $0 \leq x \leq 2\pi$ sehingga bisa diperoleh titik-titik sampel:

$$x_0 = 0, x_1 = \frac{2\pi}{N}, x_2 = \frac{2\pi.2}{N}, x_3 = \frac{2\pi.3}{N}, \dots\dots\dots$$

Atau secara umum : $x_k = \frac{2.\pi.k}{N}$ untuk $k = 0,1,2,\dots, N - 1$.

Dengan kata lain interval $0 \leq x \leq 2\pi$ dibagi menjadi N subinterval sama besar.

Pandang suatu polinomial trigonometri bentuk kompleks berikut:

$$q(x) = \sum_{n=0}^{N-1} c_n \cdot e^{i \cdot n \cdot x_k}$$

Dengan $q(x)$ merupakan nilai interpolasi dari setiap fungsi $f(x)$ pada setiap titik x_k yaitu:

$$q(x_k) = f(x_k) = f_k$$

Diperoleh:

$$f_k = f(x_k) = q(x_k) = \sum_{n=0}^{N-1} c_n e^{i \cdot n \cdot x_k} \text{ untuk nilai } k = 0, 1, 2, \dots, N-1.$$

Selanjutnya dicari nilai-nilai dari $c_0, c_1, c_2, \dots, c_{N-1}$.

Sehingga formula transformasi DFT (3) dan IDFT (4) adalah:

$$\hat{f}(\omega) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) e^{\frac{-2\pi i \omega x}{N}} \dots\dots\dots(3)$$

$$f(x) = \frac{1}{N} \sum_{\omega=0}^{N-1} \hat{f}(\omega) e^{\frac{+2\pi i \omega x}{N}} \dots\dots\dots(4)$$

Transformasi $\log|x(\omega)|$

Setiap sinyal suara yang direkam $s(n)$ merupakan sinyal yang terdiri dari dua komponen, komponen pertama adalah sinyal dari sumber suara $e(n)$ dan komponen kedua adalah sinyal yang berasal dari komponen sistem $h(n)$. Kedua komponen ini membentuk sistem sinyal suara. Pada pemrosesan sinyal, kedua komponen tersebut harus dipisahkan. Tujuan dari analisis spektral adalah untuk memisahkan kedua sinyal tersebut. sinyal hasil resultan $s(n)$ bisa dipandang sebagai konvolusi dari $e(n)$ dan $h(n)$ dan bisa dinyatakan dalam domain waktu (5) dan domain frekuensi (6)

$$s(n) = e(n) * h(n) \dots\dots\dots(5)$$

$$S(\omega) = E(\omega) \bullet H(\omega) \dots\dots\dots(6)$$

Magnitudo dari spektrum barisan $S(\omega)$ bisa dinyatakan dalam:

$$|S(\omega)| = |E(\omega)| \bullet |H(\omega)| \dots\dots\dots(7)$$

Jika kedua persamaan tersebut diambil logaritmanya diperoleh:

$$\log|S(\omega)| = \log|E(\omega)| + \log|H(\omega)| \dots\dots\dots(8)$$

Tampak bahwa spektrum barisan akhirnya dinyatakan dalam kombinasi linear dari sumber dan komponen sistem dalam domain frekuensi. Hal tersebut sesuai dengan kenyataan bahwa sinyal suara memang dipengaruhi atau terbentuk dari sumber suara dan komponen sistem. Langkah terakhir untuk mendapatkan cepstrum atau spektral adalah dengan mengambil invers dari (8), dan diperoleh:

$$c(n) = IDTF(\log|S(\omega)|) = IDTF(\log|E(\omega)| + \log|H(\omega)|) \quad \dots\dots(9)$$

Sehingga hasil akhir dari proses adalah sinyal dalam domain waktu.

Setelah cepstrum $c(n)$ diperoleh, proses selanjutnya adalah pengambilan nilai-nilai tertentu dari cepstrum dengan kriteria tertentu. Proses ini dinamakan *liftering*. Proses *liftering* sama dengan proses *filtering* pada domain frekuensi hanya sekarang dilakukan pada domain waktu. Ada dua jenis *liftering*, yaitu *low-time liftering* dan *high-time liftering*. Yang akan digunakan sekarang adalah *low-time liftering* karena berguna serupa dengan mengekstrak lintasan karakteristik vokal pada domain frekuensi. *Low-time liftering* dinyatakan dengan:

$$\omega[n] = \begin{cases} 1 & ; \quad 0 \leq n \leq L_c \\ 0 & ; \quad L_c \leq n \leq \frac{N}{2} \end{cases} \quad \dots\dots\dots(10)$$

dengan L_c merupakan panjang dari cut off dan $\frac{N}{2}$ adalah setengah panjang cepstrum.

Karakteristik lintasan vokal bisa diperoleh dengan mengkalikan cepstrum $c(n)$ dengan *low-time liftering* $\omega_e[n]$ yaitu:

$$c_e[n] = \omega_e[n].c(n) \quad \dots\dots\dots(11)$$

Transformasi Fourier diskret dikenakan pada persamaan tersebut dan mengakibatkan sinyal berubah menjadi logaritma spektrum magnitudo:

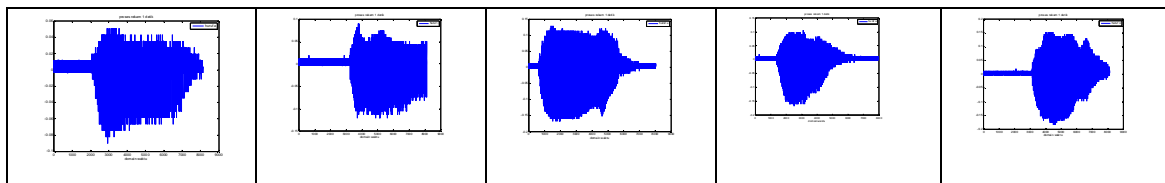
$$\log(H|\omega) = DFT[c_e[n]] \quad \dots\dots\dots(12)$$

Parameter lintasan penting seperti lokasi formant dan bandwidth bisa diperoleh dari lintasan spektrum vokal tersebut. Lokasi formant bisa diperkirakan dengan mengambil puncak-puncak dari spektrum tersebut.

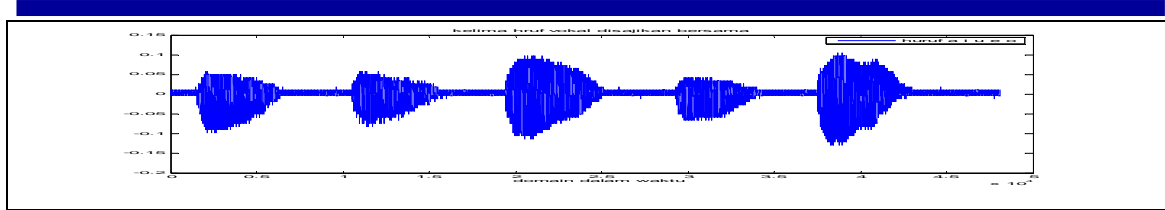
C. HASIL PENELITIAN DAN PEMBAHASAN

Sinyal Hasil Rekaman

Berikut adalah contoh hasil rekam sinyal suara lafal huruf vokal.



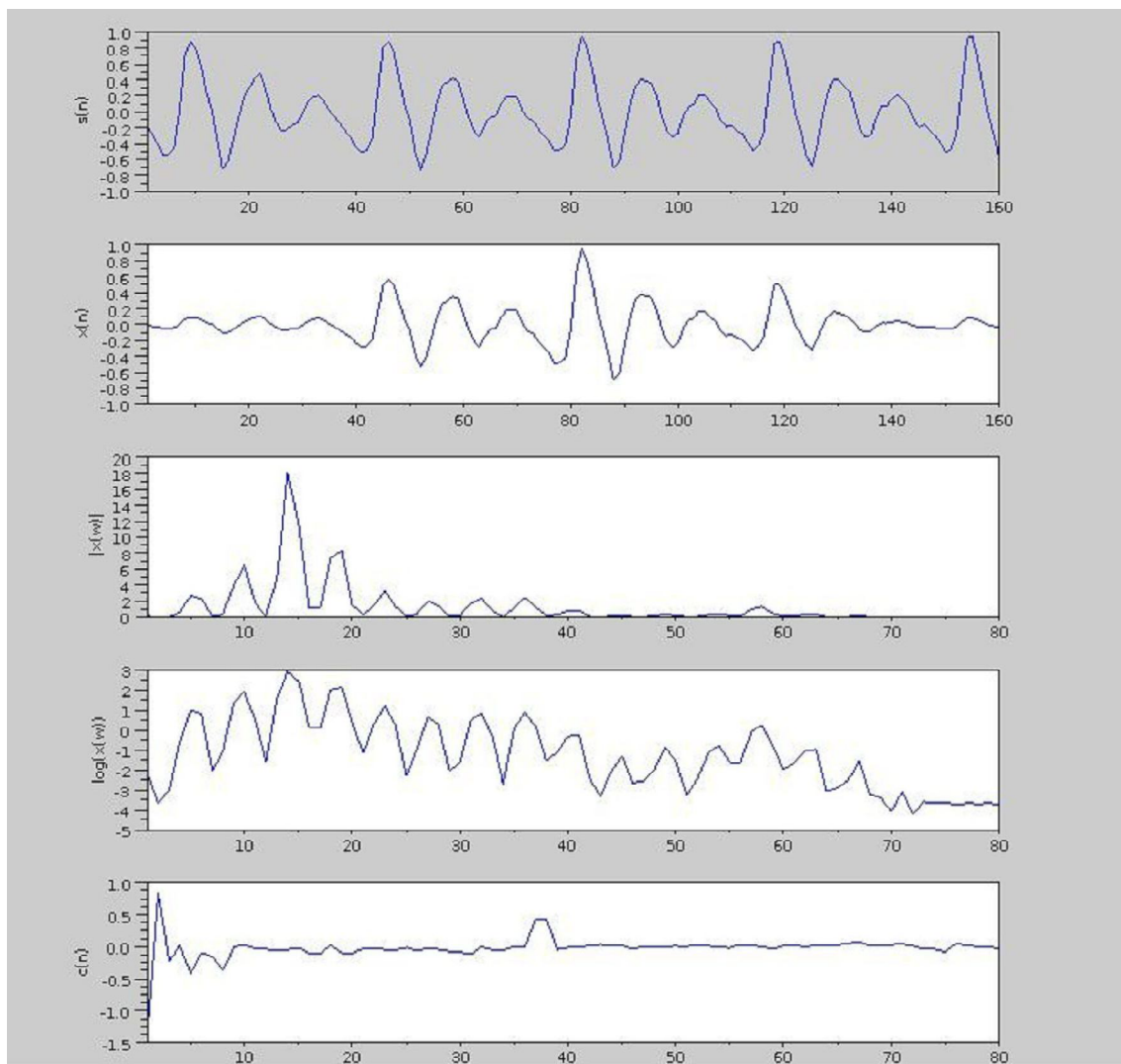
Gambar 2 sinyal lafal a,i,u,e, dan o



Gambar 3 Sinyal a,i,u,e, dan o dalam satu frame

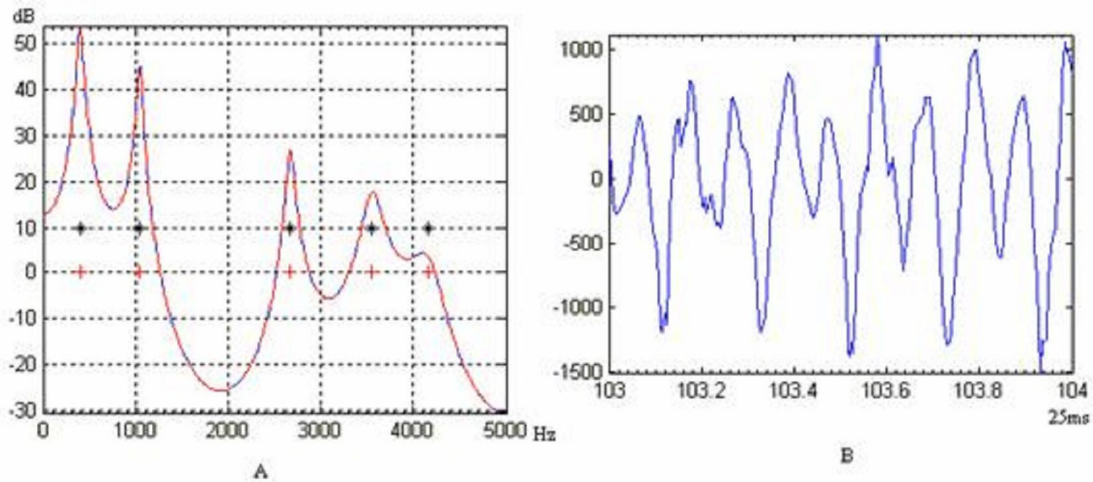
Hasil Analisis Sistem

Sinyal hasil rekam tersebut selanjutnya diproses dan hasilnya adalah cepstrum. Cepstrum dianalisa lagi menggunakan transformasi Fourier diskret dan hasil akhirnya adalah formant. Formant inilah yang menjadi karakteristik dari suatu sinyal.



Gambar 4 Segment Sinyal dan Cepstrumnya

Dari kegiatan-kegiatan tersebut, pada akhirnya diperoleh hasil akhir dari analisa sinyal suara, yaitu formant untuk setiap sinyal yang dianalisis. Berikut adalah gambaran lokasi formant yang dimaksud.



Gambar 5 Model Sinyal Linear Predictive. A) Model LP Respon Sistem, tanda + dan * merupakan lokasi formant. B) Sinyal dalam Domain Waktu, terlihat seakan-akan ada sinyal periodik.

Desain Perangkat Keras

Perangkat keras yang digunakan adalah mikrokontroler seri ATMEGA16 atau 8535. Desain sederhana dari sistem terdiri dari 1 buah mikrokontroler yang dipakai untuk memproses data, 1 Berikut adalah sistem perangkat keras sederhana yang bisa menampilkan karakter huruf vokal a,i,u,e, dan o.



Gambar 6 Desain Miktokontroler

D. KESIMPULAN DAN SARAN

1. Ciri atau karakteristik dari suatu abjad yang dinyatakan suara bisa dinyatakan menggunakan formant meskipun masih perlu penyesuaian.
2. Pengendalian mikrokontroler terhadap display LED dan LCD dapat dilakukan dengan baik sesuai dengan program yang dibuat menggunakan AVRCodeVision dengan bahasa C.
3. Mikrokontroler ATMEGA16 siap untuk dilakukan penggabungan kendali dengan komputer yang menggunakan MATLAB sebagai perangkat lunaknya.

E. DAFTAR PUSTAKA

Adriana S. Ginanjar.2010. *Panduan Praktis Mendidik Anak Autis: Menjadi Orang Tua Istimewa*, Dian Rakyat, Jakarta

-
- Baihaqi MIF. dan M. Sugiarmun. 2011. *Memahami dan Membantu Anak ADHD*, Refika Aditama, Bandung.
- Gopi,ES. 2007. *Algorithm Collection for Digital Signal Processing Applications Using MATLAB*. Springers. Netherland.
- Huzaemah. 2009. *Kenali Autisme Sejak Dini*. Penerbit Gramedia, Jakarta
- Karris, Steven T.2005 . *Electronic Devices and Amplifier Circuits*. Orchard Publications. Fremonts, California.
- Karris, Steven T.2005 . *Signals and Systems, 3rd Editions*. Orchard Publications. Fremonts, California.
- Kreuzig, Erwin. 2006. *Advanced Engineering Mathematics*. John Wiley and Sons, New York
- Permanarian Somad dan Tati H, .1995. *Bina Wicara dan Persepsi Bunyi*, Jakarta, Dirjen Dikti Depdikbud.
- Semmlow.,J.L. 2004 .*Biosignal And Biomedical Image Processing MATLAB-Based Applications*. Marcell Dekker Inc. USA
- Subkhan, 2011, *Sistem Pengaman Pintu Bioskop Dengan Menggunakan Kartu dan Sistem Suara Berbasis Mikrokontroler*, Universitas Gunadarma.
- Sukarman. 2006. *Komunikasi Perangkat Keras Menggunakan Perangkat Lunak MATLAB, Seminar Nasional II Sumber Daya Manusia Teknologi Nuklir*, Sekolah Tinggi Teknologi Nuklir Batan, Yogyakarta, 21-22 Desember 2006.
- Suparno. 2007. *Bahan Ajar Cetak: Pendidikan Anak Berkebutuhan Khusus*. Direktorat Jenderal Pendidikan Tinggi: Departemen Pendidikan Nasional.
- Vitria, Rikki .2008 . *Komunikasi Data Serial Multipoint Menggunakan Teknik RS485 Half Duplex*, *Jurnal Ilmiah Poli Rekayasa* ,Volume 3, Nomor 2, Maret 2008, Halaman 67-73.
- Fauzia, Y. 2010. *Apa Dan Bagaimana Autisme*. Lembaga Penerbit Fakultas Ekonomi Universitas Indonesia, Jakarta