

OPTIMALISASI SINYAL MIKROFONE UNTUK KEPERLUAN PENGENALAN POLA SUARA

Oleh:
Arba'i Yusuf*

Abstrak

Tulisan ini menjelaskan bagaimana mengoptimalkan sinyal audio dari mikrofone yang akan dihubungkan dengan input audio komputer untuk keperluan speech recognition, dan voice recognition. Sinyal yang keluar dari mikrofone adalah sinyal analog yang mempunyai frekuensi 20Hz – 20KHz dengan amplitudo yang selalu berubah-ubah. Terkadang dalam aplikasi tertentu dibutuhkan sinyal yang mempunyai level audio dengan amplitudo tetap agar proses pengenalan tetap akurat. Dalam tulisan ini dijelaskan cara membuat level sinyal audio dari mikrofone selalu tetap, yaitu dengan menggunakan audio kompresor. Prinsipnya adalah jika sinyal yang keluar dari mikrofone kecil akan dikuatkan dan jika sinyal yang keluar dari mikrofone besar akan diturunkan. Rasio sinyal yang dihasilkan oleh kompresor ini adalah 1:1 hingga 10:1.

Kata Kunci: Sinyal Audio, Mikrofone, Kompresor, Speech/Voice Recognition

Abstract

This paper describes how to optimize audio signal from the microphone that connected with computer soundcard that use for speech recognition and voice recognition. Output signal from microphone is an analog signal, frequency 20 Hz – 20 KHz with an amplitude that is always changing. Sometimes it takes a signal in certain applications that have a fixed amplitude of the audio level so that the recognition process can be accurately. To stabilize the analog signals are used audio compressor. The basic principle is that if the analog signal is small will be amplified and if the analog signal is bigger will be decrease. The ratio signal from audio compressor is 1:1 to 10:1.

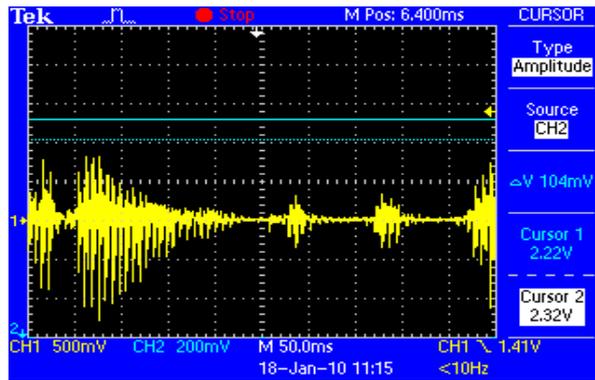
Key word: Audio signal, microphone, compressor, speech/voice recognition

1. PENDAHULUAN

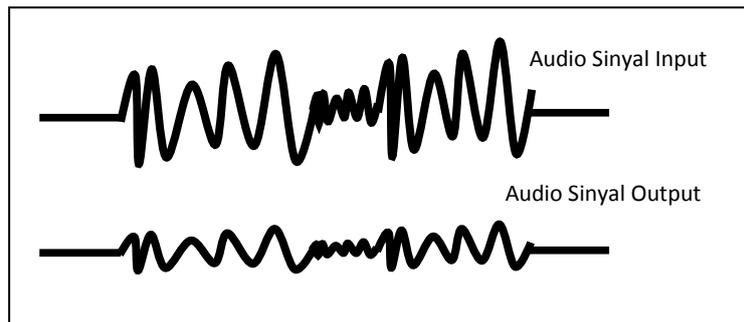
Suara adalah suatu fenomena fisik yang dihasilkan oleh getaran benda yang berupa sinyal analog dengan amplitudo yang berubah secara kontinyu terhadap waktu. Suara berhubungan erat dengan rasa “mendengar”. Suara/bunyi biasanya merambat melalui udara. Suara/bunyi tidak bisa merambat melalui ruang hampa. Selama bergetar, perbedaan tekanan terjadi di udara sekitarnya. Pola osilasi yang terjadi dinamakan sebagai gelombang. Gelombang mempunyai pola sama yang berulang pada interval tertentu, yang disebut sebagai periode. Contoh suara periodik: instrument musik, nyanyian burung, dll. Contoh suara nonperiodik: batuk, percikan ombak, dll. Manusia membuat suara dengan frekuensi 50Hz – 10KHz, sedangkan sinyal suara musik memiliki frekuensi: 20Hz – 20KHz. Sistem multimedia menggunakan suara yang berada dalam range pendengaran manusia dikenal sebagai “AUDIO”, dan gelombangnya sebagai “ACCOUSTIC SIGNALS”. Suara diluar range pendengaran manusia dapat dikatakan sebagai “NOISE” (getaran yang tidak teratur dan tidak berurutan dalam berbagai frekuensi, tidak dapat didengar manusia).

Gambar 1.1 memperlihatkan salah satu contoh sinyal suara manusia dengan pola tertentu yang keluar dari mikrofone. Sinyal tersebut akan dimasukkan ke dalam komputer untuk dilakukan pemrosesan sinyal secara digital. Salah satu contoh aplikasi dalam pemrosesan sinyal digital adalah *speech recognition, dan voice recognition*. Dalam pengolahan untuk aplikasi *speech recognition, dan voice recognition*, diperlukan sinyal audio yang mempunyai level amplitudo tetap. Jika level sinyal kecil akan dikuatkan dan jika level sinyal besar akan diturunkan. Dalam proses penguatan dan penurunan level sinyal tersebut dinamakan kompresi. Dalam proses kompresi, pola sinyal tidak boleh berubah harus tetap mempunyai pola tertentu sesuai dengan sinyal aslinya. Tujuan dari tulisan ini adalah membahas mengenai perancangan kompresi sinyal audio untuk inputan komputer. Gambar 1.2 menunjukkan contoh proses kompresi sinyal audio.

*Peneliti Pusat Teknologi Satelit - LAPAN



Gambar 1.1. Contoh Sinyal Suara Manusia

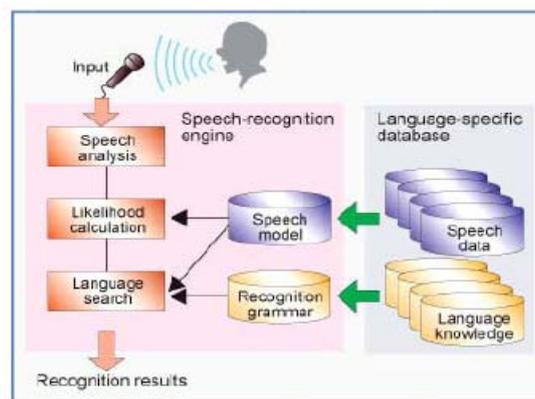


Gambar 1.2. Contoh Sinyal Kompresi

2. DASAR TEORI

Terdapat empat langkah utama dalam sistem pengenalan suara :

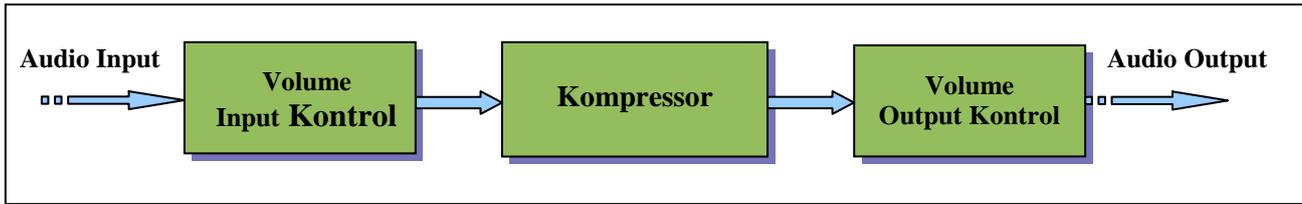
1. Penerimaan Data Input
2. Ekstrasi, yaitu penyimpanan data masukan pembuatan data base.
3. Perbandingan atau pencocokan, yaitu pencocokan data yang baru dengan data suara (pencocokan tata bahasa)
4. Validasi identitas



Gambar 2.1. Teori Pengenalan Pola Suara

2.1 Rancangan Hardware

Gambar 2.2 berikut menjelaskan blok diagram perancangan sistem kompresi sinyal audio.

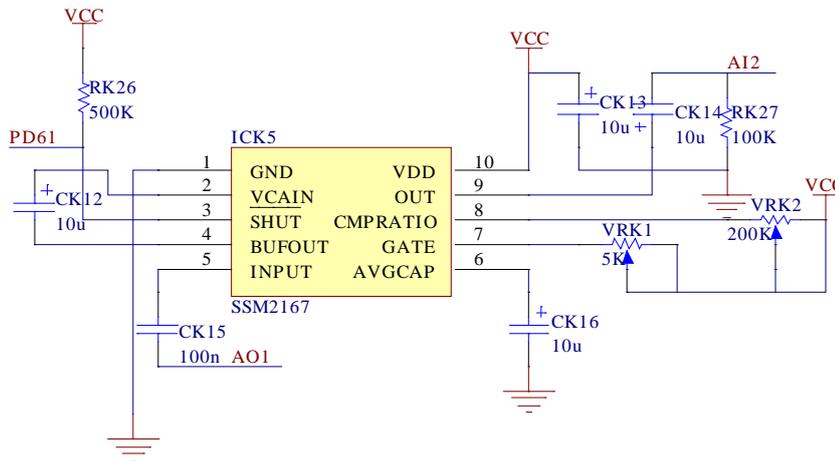


Gambar 2.2. Blok Diagram Kompresi Sinyal Audio

Bagian yang paling penting dalam sistem kompresi audio adalah blok kompresor yang berfungsi mengatur sinyal output audio pada level tertentu. Dalam perancangan ini digunakan IC kompresor bertipe SSM2167 dari *Analog Device*. Tegangan input kompresor maksimal adalah 1.25V dan tegangan output kompresor adalah 700 mV. Di bagian input dan output kompresor diberi rangkaian volume kontrol. Kontrol volume pada input kompresor digunakan untuk mengatur level audio supaya tidak melebihi 1.25 volt. Sedangkan kontrol volume pada output kompresor digunakan untuk mengatur level audio yang akan masuk ke komputer. Level audio pada komputer bervariasi ada yang 1 volt dan ada yang 2 volt maksimumnya.

2.2 Rangkaian Kompresor

Gambar 2.3 berikut menjelaskan rangkaian perancangan audio kompresor.



Gambar 2.3. Rangkaian Audio Kompresor

Gambar 2.3 memperlihatkan rangkaian audio kompresor yang menggunakan IC kompresor SSM2167. Pengaturan rasio kompresi dapat dilakukan dengan mengatur variabel resistor VRK2. Sedangkan pengaturan level noise digunakan variabel resistor VRK1. Tabel 2.1 dan Tabel 2.2 berikut menjelaskan pengaturan resistor untuk kompresi dan noise.

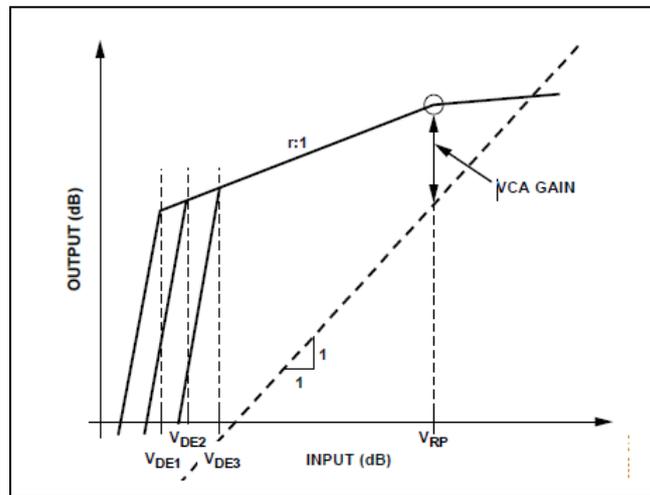
Tabel 2.1. Setting Rasio Kompresi

Ratio Kompresi	Nilai Resistor
Rasio 1:1	0 Ω (Hubung singkat ke V ⁺)
Rasio 2:1	15 K Ω
Rasio 3:1	35 K Ω
Rasio 5:1	75 K Ω
Rasio 10:1	175 K Ω

Tabel 2.2. Setting Noise Gate Threshold

Noise Gate (dBV)	Nilai Resistor
-40	0 Ω (Hubung singkat ke V ⁺)
-48	1 K Ω
-54	2 K Ω
-55	5 K Ω

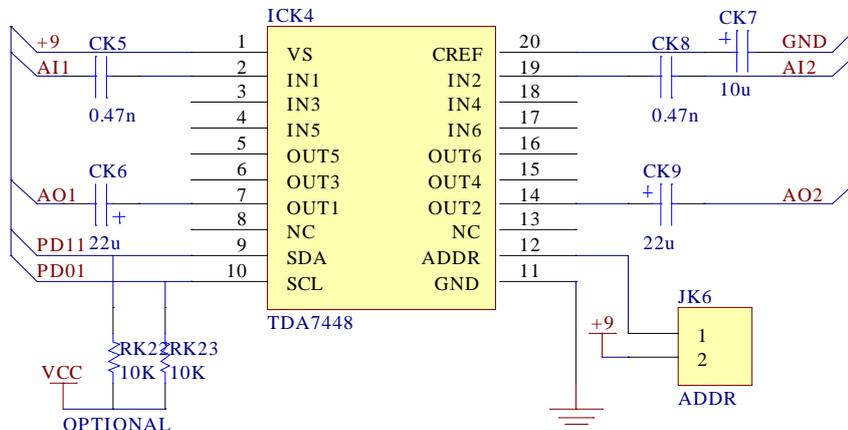
Tabel 2.1 menjelaskan pengaturan rasio kompresi, jika resistor diset sebesar 0 ohm menunjukkan rasio kompresi 1:1 maka sinyal audio tidak ada yang dikompres. Nilai resistor 15 K ohm menunjukkan rasio kompresi 2:1 begitu seterusnya. Sedangkan tabel 2.2 menjelaskan pengaturan noise. Nilai resistor 0 ohm berarti bahwa noise thresholdnya adalah -40 dBV, nilai resistor 1 K berarti bahwa noise threshold -48 dBV dan seterusnya. Perlu diingat bahwa semakin tinggi rasio kompresi, akan semakin tinggi pula noise yang dihasilkan. Gambar 2.4 berikut memperlihatkan perbandingan kompresi terhadap noise yang dihasilkan.



Gambar 2.4. Perbandingan Kompresi Terhadap Noise

2.3. Rangkaian Volume Kontrol

Gambar 2.5 berikut menjelaskan rangkaian perancangan volume kontrol untuk audio kompressor.

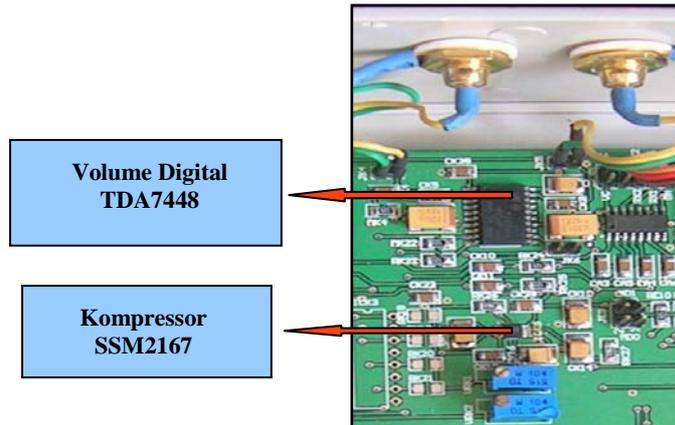


Gambar 2.5. Rangkaian Volume Kontrol

Gambar 2.5 adalah rangkaian volume kontrol, terdapat enam buah channel input dan enam buah channel output. Untuk perancangan audio kompresor ini dibutuhkan 2 channel saja yaitu untuk input kompresor dan output kompresor. Untuk input kompresor digunakan channel 1 yaitu pin nomor 2 dan outputnya pin nomor 7. Sedangkan untuk output kompresor digunakan channel 2 yaitu pin nomor 19 dan outputnya pin nomor 14. Untuk pengontrolan volume digunakan komunikasi I2C yang dapat dihubungkan dengan mikrokontroler. Bagaimana protokol untuk mengakses volume ini dapat dilihat dalam datasheet TDA7448.

2.4. Foto Alat Kompresor dan Volume Digital

Gambar 2.6 berikut adalah bagian dari sistem *speech recognition* dan *voice recognition* yang telah dibuat. Komponen utama adalah sebuah IC kompresor SSM2167 dan volume digital TDA7448.

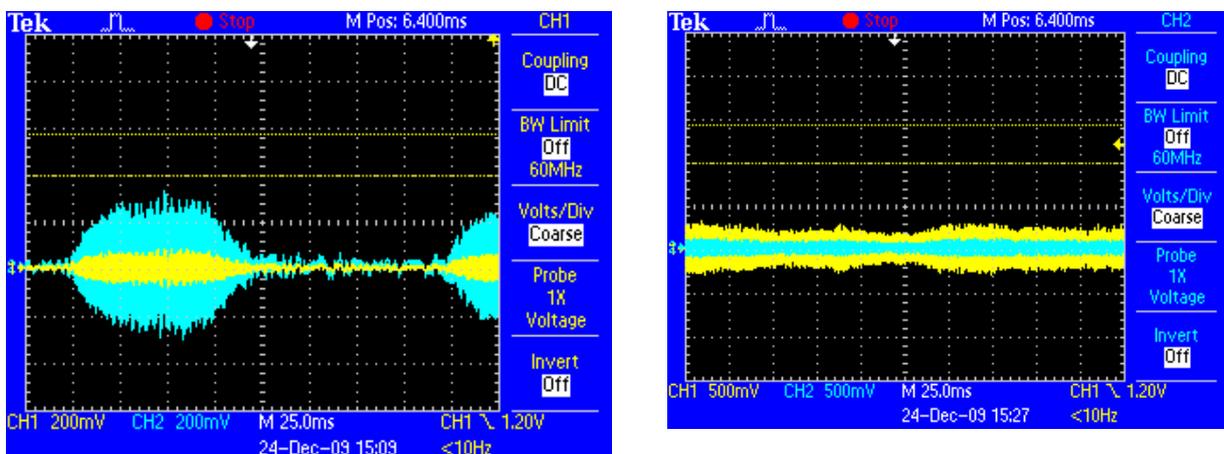


Gambar 2.6. Foto Alat Kompresor dan Volume Digital

3. HASIL PERCOBAAN

Percobaan optimalisasi sinyal mikrofone dilakukan menggunakan sinyal dengan frekuensi audio, sumber sinyal bisa dari sinyal generator, suara manusia atau dari suara radio. Pengetesan pertama dilakukan menggunakan suara manusia dengan jarak mikrofone dengan mulut orang yang berbicara dibuat dekat dan jauh sehingga menghasilkan amplitudo yang berubah-ubah. Kemudian percobaan kedua menggunakan sinyal suara musik dari sebuah radio kemudian dilihat apakah modul kompresor dapat bekerja dengan baik. Setting volume digital diset 0db.

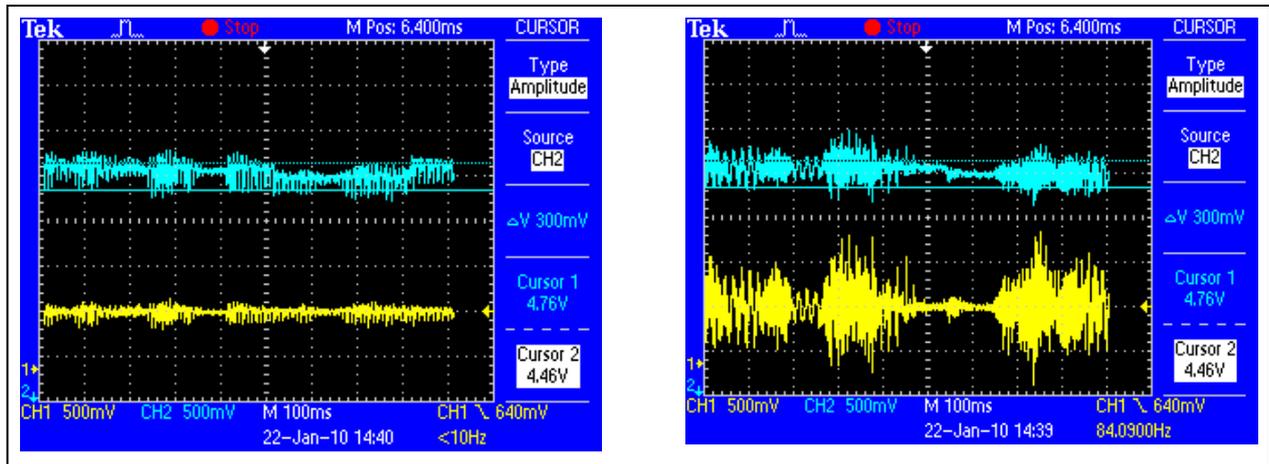
3.1. Percobaan Menggunakan Suara Manusia



Gambar 3.1. Sinyal Hasil Percobaan Menggunakan Suara Manusia

Gambar 3.1. adalah hasil percobaan sinyal dengan menggunakan suara manusia. Gambar tersebut terdiri dari dua warna, yaitu biru dan kuning, gambar berwarna kuning adalah sinyal input sedangkan gambar berwarna biru adalah sinyal output. Gambar disebelah kiri menunjukkan suara manusia jauh dari mikrofone dengan jarak kira-kira 50 cm, sedangkan gambar disebelah kanan menunjukkan suara manusiadekat dengan mikrofone dengan jarak kira-kira 5 – 10 cm. Hasil percobaan menunjukkan hasil bahwa sinyal yang lemah akan dikuatkan dan sinyal yang kuat akan diturunkan. Ini menunjukkan kompresor dapat bekerja dengan baik.

3.2. Percobaan Menggunakan Suara Music



Gambar 3.2. Sinyal Hasil Percobaan Menggunakan Suara Musik

Gambar 3.2. memperlihatkan hasil percobaan menggunakan suara musik dari radio. Gambar dengan warna kuning adalah sinyal input, gambar dengan warna biru adalah sinyal output. Gambar disebelah kiri menunjukkan suara dengan volume kecil dan gambar disebelah kanan menunjukkan suara dengan volume besar. Hasil percobaan menunjukkan output kompresor mempunyai level amplitudo sama dan pola juga sama. Ini menunjukkan kompresor dapat bekerja dengan baik.

4. KESIMPULAN DAN SARAN

Telah dibuat suatu prototipe hardware yang dapat mengoptimalisasi sinyal mikrofone untuk keperluan *speech recognition* dan *voice recognition*. Dari hasil percobaan menunjukkan hasil bahwa sinyal suara yang lemah akan dikuatkan dan sinya suara yang kuat akan diturunkan. Ini menunjukkan bahwa hasil kompresi dapat bekerja dengan baik. Pengesetan kompresi akan mempengaruhi noise output, semakin besar kompresinya maka semakin tinggi pula noise yang dihasilkan.

DAFTAR PUSTAKA

- Unknown, "*Low Voltage Microphone Preamplifier with Variable Compression and Noise Gating*", Datasheet SSM2167, Analog Device, 2009.
- Unknown, "*6 Channel Volume Controller*", Datasheet TDA7448, STMicroelectronics, 2004.
- Lawrence Labiner, "*Fundamentals of Speech Recognition*", Prentice-Hall, 1993.

DAFTAR RIWAYAT HIDUP PENULIS

DATA UMUM

Nama Lengkap : ARBA'I YUSUF
Tempat & Tgl. Lahir : KEDIRI, 10 APRIL 1976
Jenis Kelamin : LAKI-LAKI
Instansi Pekerjaan : PUSTEK ELEGAN LAPAN
NIP. / NIM. : 197604102006041027
Pangkat / Gol. Ruang : PENATA MUDA / IIIB
Jabatan Dalam Pekerjaan : PENELITI
Agama : ISLAM
Status Perkawinan : KAWIN

DATA PENDIDIKAN

SLTA : SMAN3 KEDIRI Tahun : 1995
STRATA 1 (S.1) : UNIVERSITAS BRAWIJAYA Tahun : 2001
STRATA 2 (S.2) : UNIVERSITAS INDONESIA Tahun : 2010

ALAMAT

Alamat Rumah : JL. PAKIS I/24 RT04 RW09 TAMAN YASMIN 2 BOGOR
HP. : 0818813840
Alamat Kantor / Instansi : JL. CAGAK SATELIT KM 0.4 RANCABUNGUR BOGOR
Telp. : .0251621667.
E-mail : arbaibrwlpn@yahoo.com

Hasil Diskusi

Pertanyaan :

1. Berapa jarak minimum dan maksimum mikrophone terhadap orang yang berbicara supaya sistem tersebut masih mampu memproses? (Bp. Dwiyanto, LAPAN)

Jawaban:

1. Jarak minimum adalah sedekat mungkin dan jarak maksimum adalah sekitar satu meter, lebih dari satu meter sistem kurang responsif.