

UJI VALIDASI SUARA BERBASIS PENGENALAN SUARA (*VOICE RECOGNITION*) MENGUNAKAN *EASY VR 3.0*

Anjar Imario¹, Dodi Wisaksono Sudiharto², Endro Ariyanto³,

¹²³Program Studi Teknik Informatika, Fakultas Informatika, Universitas Telkom
Jl. Telekomunikasi No. 01, Terusan Buah Batu, Sukapura, Dayeuhkolot, Bandung, Jawa Barat
40257

Email: ¹jarimariojr@gmail.com, ²wsdodi@gmail.com, ³endroariyanto@gmail.com

Abstrak

Perintah suara dapat digunakan untuk melakukan pengoperasian sistem *home automation*. Terdapat dua metode untuk membangun sistem *home automation* berbasis perintah suara (*voice command*) yaitu dengan menggunakan pengenalan wicara (*speech recognition*) dan pengenalan suara (*voice recognition*). *Speech recognition* akan mengubah sinyal analog suara menjadi data digital yang akan dicocokkan data dengan pola tertentu yang disimpan di dalam basis data. Sehingga hasil yang didapatkan berupa teks yang sesuai dengan pola ucapan yang diberikan. Sedangkan *voice recognition* akan mengenali suatu suara dengan membandingkan pola karakteristiknya dengan sinyal suara yang menjadi referensi atau acuan yang sudah disimpan sebelumnya. Jadi dengan kata lain *speech recognition* dapat mengerti kata apa yang dikatakan oleh seseorang dan *voice recognition* dapat mengidentifikasi seseorang melalui suaranya. Dari definisi tersebut dapat terlihat bahwa faktor keamanan sistem dengan *speech recognition* masih rentan karena dapat dioperasikan oleh siapa saja yang mengetahui perintah suaranya. Studi ini merupakan penelitian awal untuk membangun sistem *home automation* berbasis *voice recognition*. Pada sistem ini digunakan *Easy VR 3.0* yang memiliki fitur *user-defined speaker dependent* yang dapat menyimpan suara sebagai referensi untuk dibandingkan dengan perintah suara berbasis *voice recognition*. Pengujian performansi dilakukan dengan memasukkan suara dari 30 user dan dibandingkan dengan 1 suara yang menjadi acuan.

Kata Kunci : *home automation, Easy VR 3.0, user-defined speaker dependent, voice recognition*

1. PENDAHULUAN

Perintah suara merupakan salah satu media pengoperasian sistem *home automation* yang banyak diminati. Pengoperasian sistem yang sangat mudah dan tidak membutuhkan banyak tenaga merupakan alasan utama perintah suara cocok digunakan untuk sistem *home automation*. Baik *speech* maupun *voice recognition* dapat digambarkan sebagai suatu proses di mana mesin atau program menerima dan menafsirkan dikte serta memahami dan menjalankan perintah yang diucapkan. Namun terdapat perbedaan antara *voice recognition* dengan *speech recognition*, di mana teknologi *voice recognition* dapat mengerti kata-kata seperti yang diartikulasikan sehingga memenuhi unsur biometrik (Findbiometrics, n.d.).

Studi ini merupakan penelitian awal di mana akan dibangun sebuah sistem yang dapat memenuhi unsur biometrik seperti intonasi suara untuk validasi perintah suara dari user yang spesifik. Sehingga unsur yang diteliti mencakup dua hal, yaitu pengenalan perintah suara dan pengenalan intonasi suara (Myers, 2004). Sehingga tidak selalu orang yang mengetahui perintah suara dapat diterima perintahnya oleh sistem bila suaranya memiliki intonasi yang berbeda dengan suara yang menjadi referensi dan tersimpan di dalam sistem (Jain, Ross, & Prabhakar, 2004).

Latar Belakang Masalah

Sistem berbasis perintah suara umumnya dapat diakses oleh semua orang. Maka dari itu diperlukan sistem yang memenuhi unsur *privacy*. Sistem yang menjamin unsur *privacy*, biasanya harus dapat memenuhi unsur biometrik atau *behaviour* dari penggunanya (Findbiometrics, n.d.; Prabhakar, Pankanti, & Jain, 2003)

Memproses pengenalan *user* berdasarkan unsur biometrik atau *behaviour* biasanya memerlukan perangkat komputasi yang besar. Umumnya hal ini membutuhkan komputer, seperti PC (*Personal Computer*) yang berperan khusus untuk mengolah perintah suara (Way, AbilitNet, & Charity, 2009). Tentunya akan kurang praktis bila tujuannya adalah untuk mengontrol otomasi

perangkat yang ada di rumah, seperti halnya sistem *home automation*. Sehingga solusi untuk masalah ini adalah bagaimana hal ini bisa diwujudkan pada perangkat yang kecil dan dapat menyatu pada perangkat yang digunakan untuk keperluan sehari-hari. Artinya, perangkat tersebut harus berukuran kecil dan memiliki kemampuan komputasi setidaknya dapat mengenali sebuah pola (*pattern*) dari suara.

Pada penelitian ini diajukan suatu desain sistem otomasi berbasis perintah suara dengan menggunakan modul *voice recognition* yaitu *Easy VR 3.0*. Modul *voice recognition* ini memiliki fitur pengenalan suara berupa *user-defined speaker dependent* (RoboTech, EasyVR 3, n.d.). Fitur ini dapat mengenali perintah suara *user* yang spesifik sesuai dengan suara yang sudah disimpan sebagai referensi untuk dibandingkan dengan perintah suara yang masuk ke sistem.

TEORI DASAR

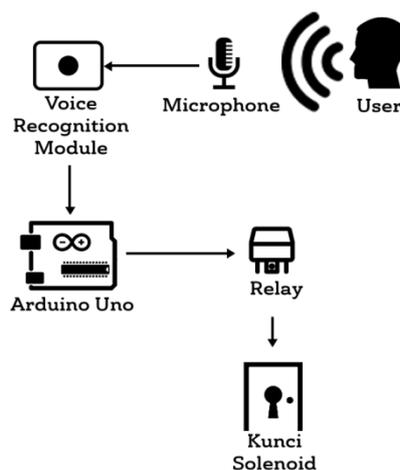
Voice recognition juga dikenal sebagai *speaker recognition* dirancang untuk mengenali siapa orang yang berbicara. Ini menggunakan karakter akustik suara yang berbeda antara individu (Findbiometrics, n.d.). Pola-pola akustik mencerminkan baik anatomi (seperti ukuran dan bentuk tenggorokan dan mulut) dan belajar pola perilaku (seperti nada suara dan gaya berbicara) (Frischholz & Diekmann, 2000; King, 2014). Sebelum dapat mengenali suara pembicara, metode ini membutuhkan beberapa pelatihan di mana sistem akan mempelajari suara, aksen dan nada pembicara. Hal ini umumnya dilakukan dengan merekam serangkaian kata-kata atau perintah tekstual oleh user yang berbicara melalui *built-in* atau mikrofon eksternal (RoboTech, EasyVR 3, n.d.).

Sistem ini menggunakan *Easy VR 3.0* sebagai modul pengenal suaranya. *Easy VR 3.0* memiliki fitur *user-defined speaker dependent* yang dapat menyimpan hingga 32 perintah suara dari user yang dapat digunakan dengan bahasa apapun (RoboTech, EasyVR 3, n.d.). *Easy VR 3.0* juga memiliki *shield interface 1.0* yang sesuai dengan beberapa jenis mikrokontroler dari *Arduino* seperti *Arduino Duemilanove*, *Uno*, *Mega*, *Leonardo*, dan *Due* (RoboTech, EasyVR 3, n.d.).

Easy VR 3.0 dapat berkomunikasi langsung dengan komputer menggunakan “bridge mode” pada *Arduino* dan dikonfigurasi dengan aplikasi *EasyVR Commander* (RoboTech, EasyVR 3 User Manual Release 1.0.15, n.d.). Terdapat juga beberapa aplikasi pendukung untuk membangun tabel suara (*sound table*) seperti *Sensory QuickSynthesis4* (RoboTech, EasyVR 3, n.d.).

2. DESAIN DAN IMPLEMENTASI.

Penelitian ini dimaksudkan untuk membangun sistem berbasis pengenalan suara sebagai fitur akses sistem. Sistem yang akan dianalisa menggunakan fitur *Speaker-Dependent* dari *Easy VR 3.0* untuk menyimpan perintah suara user. Berikut merupakan diagram blok dari sistem:



Gambar 1. Diagram blok sistem.

Pertama-tama sistem perlu diinisialisasi dengan cara melakukan pelatihan (*training*) pada perintah suara user (*Speaker-Dependent*). *Training* perintah suara ini bertujuan untuk menyimpan suara user secara spesifik. Sehingga diharapkan hanya user tersebut yang dapat mengakses sistem.

Training perintah suara dilakukan dengan menggunakan aplikasi *EasyVR Commander* (Hassan, Rejab, & Sapar, 2015).

Pada *EasyVR* memiliki tiga konfigurasi sensitifitas (*strictness control*) terhadap pengenalan suara, yaitu *easy*, *normal*, dan *hardest*. Pada penelitian ini dipilih konfigurasi *normal* dengan alasan merupakan tingkat konfigurasi yang dianggap umum atau rata-rata. Sehingga dalam hal ini sistem yang dibuat tidak dimaksudkan menggunakan tingkat sensitifitas yang terlalu tinggi.

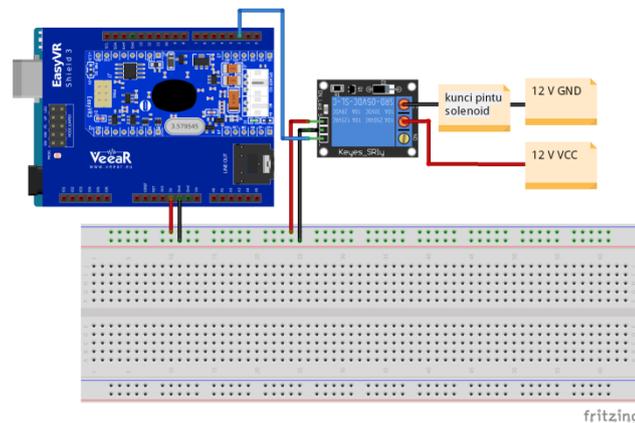
Dalam pengoperasiannya, user akan memberikan perintah suara melalui mikrofon. Kemudian akan dilakukan verifikasi terhadap suara tersebut apakah sesuai dengan data *training* yang disimpan sebelumnya atau tidak. Jika perintah suara valid, maka sistem akan merespon sesuai dengan perintah yang diberikan.

Komponen Perangkat Keras dan Perangkat Lunak yang Digunakan

Sistem ini menggunakan perangkat keras sebagai berikut (Hassan, Mat Rejab, & Sapar, 2015):

- *Arduino Uno R3* sebagai pengontrol aktuator.
- *Easy VR Shield 3.0* sebagai perangkat yang memverifikasi perintah suara.
- Aktuator (kunci pintu *Solenoid*).
- *Relay*, saklar untuk aktuator.

Sedangkan perangkat lunak yang digunakan adalah *Arduino IDE* dan *EasyVR Commander* (Hassan, Mat Rejab, & Sapar, 2015).



Gambar 2. Diagram *wiring* sistem

3. HASIL DAN PEMBAHASAN

Pengujian berupa validasi perintah suara yang menjadi referensi dengan perintah suara dari tiga puluh orang yang berbeda. Dari hasil pengujian tersebut dapat dilihat hasil performansi dari sistem.

Pengujian dilakukan dengan menggunakan dua perintah suara, yaitu “buka kunci” dan “kunci pintu”. Jarak user dari mikrofon diatur maksimal satu meter dan kondisi lingkungan dengan kebisingan suara dibawah 55 dB (kondisi perumahan dan pemukiman) (Indonesia, 1996). Sistem diinisialisasi dengan satu user valid yang menyimpan dua perintah suara tersebut. Adapun jenis kelamin dari suara yang menjadi acuan adalah berjenis kelamin laki-laki.

Pada pengujian akan diambil sebanyak 3 kali inputan untuk masing-masing kedua perintah tersebut. Sehingga untuk masing-masing *user* akan diambil data *voice command*-nya sebanyak 6 inputan. Jumlah *user* yang mengikuti proses pengujian ini adalah 30 orang.

Nilai terbaik dari pengujian adalah nol, yang artinya sistem mengenali bahwa *user* yang bersangkutan bukanlah yang berhak melakukan akses terhadap sistem. Sedangkan bila terdapat nilai selain itu, berarti masih terdapat celah yang menyebabkan *user* yang tidak berhak, dikenali oleh sistem sebagai *user* yang berhak melakukan akses pada sistem, dan ini merupakan nilai *error* dari implementasi sistem.

Hasil dari pengujian dapat dilihat pada tabel di bawah:

Tabel 1. Hasil pengujian validasi suara.

Nama	Jenis Kelamin	Buka_Kunci	Kunci_Pintu	Error (%)
Prima	L	0	0	0
Afrizal	L	2	2	67
Gita	P	0	0	0
Budi	L	0	1	17
Tyas	P	0	0	0
Lya	P	0	0	0
Ridho	L	0	0	0
Augi	P	0	0	0
Ayu	P	0	0	0
Arga	L	1	0	17
Ibnu	L	1	1	34
Hasbi	L	0	1	17
Zaki	L	1	0	17
Dodo	L	0	0	0
Sigit	L	0	1	17
Bayu	L	0	0	0
Lovel	P	0	0	0
Fasya	P	0	0	0
Hilman	L	0	1	17
Gilang	L	0	0	0
Cipaek	L	0	0	0
Fadli	L	0	0	0
Rian	P	0	0	0
Bintang	L	1	1	34
Keri	L	1	2	50
Adit	L	0	0	0
Wisnu	L	1	1	34
Amanda	P	0	0	0
Irul	L	0	1	17
Dian	P	0	0	0
Rata-rata error				11,3

Dari tabel hasil pengujian di atas terlihat bahwa probabilitas keberhasilan validasi perintah suara (tanpa memperhatikan masing-masing gender dari *user* yang melakukan akses) dari sistem yang diajukan adalah sebesar 88.7%.

Bila memperhatikan gender, terlihat dari data pada tabel 1, bahwa gender perempuan memiliki tingkat keberhasilan yang jauh lebih tinggi ketimbang yang laki-laki. Hal ini berkaitan dengan data suara yang menjadi acuan, yaitu suara laki-laki. Sehingga, manakala pengujiannya adalah suara perempuan, sistem akan mudah mengenalinya sebagai suara dari *user* yang tidak berhak melakukan akses ke sistem. Sebaliknya, sistem masih memiliki kesulitan dalam membedakan suara yang diuji, manakala suara tersebut merupakan suara dari *user* yang memiliki jenis kelamin laki-laki.

Selain terkait dengan suara acuan, kegagalan dari pengujian juga disebabkan karena adanya intonasi suara yang mirip dengan suara yang menjadi acuan meski ia berasal dari *user* yang berbeda.

Selanjutnya untuk mengetahui waktu respon dari sistem tepat setelah menerima masukan sampai dengan respon dari sistem. Hasilnya dapat dilihat di bawah:

Tabel 2. Hasil pengujian respon sistem.

Nama	Respon “buka kunci” (detik)			Respon “kunci pintu” (detik)		
	Ke-1	Ke-2	Ke-3	Ke-1	Ke-2	Ke-3
Prima	1,3	1,3	1,4	1,3	1,3	1,3
Afrizal	1,3	1,4	1,4	1,4	1,4	1,3
Gita	1,2	1,3	1,3	1,4	1,4	1,4
Budi	1,5	1,3	1,4	1,3	1,3	1,3
Tyas	1,3	1,4	1,3	1,3	1,2	1,5
Lya	1,3	1,2	1,3	1,3	1,3	1,2
Ridho	1,3	1,3	1,3	1,2	1,3	1,4
Augi	1,3	1,2	1,3	1,5	1,3	1,3
Ayu	1,3	1,3	1,3	1,3	1,4	1,2
Arga	1,3	1,3	1,3	1,3	1,4	1,4
Ibnu	1,4	1,4	1,4	1,3	1,3	1,3
Hasbi	1,3	1,3	1,3	1,4	1,3	1,4
Zaki	1,3	1,4	1,4	1,3	1,4	1,3
Dodo	1,4	1,2	1,2	1,3	1,2	1,3
Sigit	1,3	1,3	1,4	1,3	1,3	1,3
Bayu	1,3	1,4	1,4	1,4	1,4	1,3
Lovel	1,4	1,2	1,3	1,5	1,3	1,3
Fasya	1,4	1,3	1,2	1,3	1,3	1,2
Hilman	1,4	1,3	1,3	1,3	1,3	1,2
Gilang	1,2	1,3	1,5	1,2	1,3	1,3
Cipaek	1,3	1,3	1,4	1,3	1,3	1,3
Fadli	1,3	1,4	1,4	1,4	1,4	1,3
Rian	1,4	1,2	1,3	1,3	1,3	1,3
Bintang	1,3	1,5	1,2	1,3	1,4	1,3
Keri	1,3	1,3	1,4	1,3	1,4	1,3
Adit s	1,3	1,4	1,3	1,3	1,3	1,3
Wisnu	1,4	1,3	1,3	1,3	1,3	1,3
Amanda	1,2	1,3	1,3	1,3	1,3	1,3
Irul	1,3	1,3	1,2	1,3	1,3	1,4
Dian	1,2	1,3	1,5	1,3	1,4	1,3
Rata-rata waktu respon				1,32 detik		

Dapat dilihat dari tabel di atas bahwa waktu respon dari sistem tidak dipengaruhi oleh perintah suara yang diberikan oleh *user*. Waktu respon lebih dipengaruhi oleh performa dari mikrokontroler yang digunakan.

4. KESIMPULAN

Pengaturan *home automation* menggunakan perintah suara (*voice command*) merupakan pilihan yang populer. Namun terdapat permasalahan terkait akses terhadap perangkat rumah yang memiliki *privillage* yang tinggi, contohnya seperti pintu utama rumah. Penggunaan *voice recognition* dapat menjadi pilihan karena ia dapat mengidentifikasi seseorang melalui karakter suaranya.

Perangkat *Easy VR 3.0* dapat menjadi alternatif dalam membangun sistem berbasis *voice recognition*, dengan beberapa kelebihan seperti ukurannya yang kecil sehingga dapat menyatu pada perangkat yang digunakan untuk keperluan sehari-hari, memiliki fitur *user-defined speaker dependent*, sehingga untuk implementasinya dapat meminimalisir kompleksitas instalasi, dan konfigurasi fungsi perangkatnya yang relatif sederhana.

Dari percobaan terhadap tiga puluh *user*, dengan konfigurasi fungsi pengenalan suara dari *Easy VR 3.0* pada tingkat yang umum, masih bisa dikatakan bahwa keberhasilan implementasi dari sistem relatif tinggi, yaitu sebesar 88,7%. Sehingga dari beberapa hal yang bisa disimpulkan di atas, penggunaan perangkat ini sebagai bagian dari pengembangan *Home Automation System* adalah sangat layak

Saran

Pada pengujian penelitian ini masih memiliki beberapa kekurangan, seperti pengujian yang belum memperhatikan inputan suara berdasarkan gender, suku, serta tingkat sensitifitas pengenalan suara dari *Easy VR 3.0*. Pada penelitian selanjutnya setidaknya akan dilakukan pengambilan data dengan memperhatikan beberapa hal yang tadi disebutkan.

DAFTAR PUSTAKA

- Findbiometrics. (n.d.). Voice and Speech Recognition. (Findbiometrics) Diakses: 13 Mei 2016, dari <http://findbiometrics.com/solutions/voice-speech-recognition/>
- Frischholz, R. W., & Dieckmann, U. (2000). BiOLD: a multimodal biometric identification system. *Computer*, 33(2), 64-68.
- Hassan, N. F., Rejab, M. R. M., & Sapar, N. H. (2015). Implementation of Speech Recognition Home Control System Using Arduino. *ARPN Journal of Engineering and Applied Sciences*, 10(23), 17492-17498.
- Indonesia, P. (1996). Keputusan Menteri Lingkungan Hidup Nomor : KEP-48/MENLH/11/1996 tentang Baku Tingkat Kebisingan. Jakarta: Menteri Negara Lingkungan Hidup.
- Prabhakar, S., Pankanti, S., & Jain, A. K. (2003). Biometric recognition: Security and privacy concerns. *IEEE security & privacy*, 99(2), 33-42.
- Myers, L. (2003). Myers, L. (2004). An exploration of voice biometrics. *GSEC Practical Assignment version, 1*.
- King, R. O. (2014). Speech and Voice Recognition. Biometrics Research Group. Diakses: 14 Mei 2017, dari www.biometricupdate.com/wp-content/uploads/2014/05/Voice-Biometrics.pdf
- Prabhakar, S., Pankanti, S., & Jain, A. K. (2003). Biometric recognition: Security and privacy concerns. *IEEE security & privacy*, 99(2), 33-42.
- RoboTech. (n.d.). EasyVR 3. (RoboTech srl) Diakses: 13 Mei 2017, dari <http://www.veear.eu/products/easyvr3/>
- RoboTech. (n.d.). EasyVR 3 User Manual Release 1.0.15. RoboTech srl. Diakses: 14 Mei 2017, dari www.veear.eu/files/EasyVR-User-Manual.pdf
- Way, M. W., AbilitNet, & Charity, U. C. (2009). *Voice-Recognition Software - Advanced Features and Concepts* [Fact sheet]. Diakses: 14 Mei 2017, dari www.bbc.co.uk/accessibility/guides/factsheets/factsheet_VR_advanced.pdf