

Kajian Penelitian Pemrosesan Bunyi dan Aplikasinya pada Teknologi Informasi (Study of Sound Processing and Application on Information Technology)

Ranny¹, Iping Supriana Suwardi², Tati Latifah Erawati Rajab³, Dessi Puji Lestari⁴

^{1,2,3,4}*Sekolah Teknik Elektro dan Informatika Institut Teknologi Bandung*

¹ranny.fransiska@gmail.com

²iping@informatika.org

Abstrak-- Hasil dari penelitian banyak digunakan dan dikembangkan pada aplikasi yang telah banyak dimanfaatkan pada kehidupan sehari-hari. Proses identifikasi bunyi menjadi salah satu penelitian yang banyak dilakukan. Identifikasi bunyi yang dilakukan oleh manusia berbeda satu sama lain. Misal pada suara detak jantung, pada pendengar umum, suara detak jantung tidak memiliki informasi apa pun terkait kesehatan, tapi jika suara detak jantung diperdengarkan pada ahli medik atau dokter, maka informasi yang diperoleh akan berbeda, dokter dapat mengidentifikasi suara detak jantung dikaitkan dengan kondisi kesehatan jantung. Selain dalam bidang medis, bunyi juga dimanfaatkan pada aplikasi berbasis bunyi dan suara pada Smart Homes. Namun, sebelum mengkaji tentang aplikasi pada Smart Homes dan aplikasi lain maka akan dibahas beberapa teori dasar tentang bunyi dan suara, seperti: teori suara dan bunyi, *noise* pada data suara, serta ekstraksi ciri suara bunyi yang secara spesifik akan menjelaskan tentang *Mel Frequency Cepstrum Coefficients* (MFCC). Berdasarkan hasil kajian dapat dibuat kerangka kerja aplikasi yang dibuat. Kerangka kerja yang disusun merupakan kerangka kerja yang umum dilakukan pada aplikasi dan penelitian tentang penggunaan data suara dan bunyi. Selain itu kajian ini akan menjabarkan tentang lingkup penelitian bunyi dan suara yang telah banyak dilakukan. Melalui penjabaran tentang lingkup penelitian didapatkan peluang penelitian yang dapat dilakukan pada data bunyi dan suara serta tantangannya.

Kata kunci - bunyi, suara, Smart Homes, pengenalan suara dan bunyi, ekstraksi ciri suara dan bunyi, MFCC.

Abstract-- *The results of research are widely used and developed in applications that have been used in human life. The process of sound identification becomes one of the many researches conducted in Computer Science research. The human activity of sound identification is different from one another. For example, in the heartbeat sound, in the*

general listener, the heartbeat does not have any information related to health, but if the heartbeat sound is played by a medical expert or doctor, the information obtained will be different, the doctor can identify the heartbeat associated with the heart health condition. In addition to the medical field, sound is also used in sound and voice based applications on Smart Home's. However, before reviewing the application on Smart Home's and other sound applications it will be discussed some basic theories about sound and voice, such as sound and voice theory, noise on sound data, and sound feature extraction that will specifically explain about Mel Frequency Cepstrum Coefficients (MFCC). Based on the results of the study can be made a framework of the application. Through the description of the scope of research we can get the research opportunities using the sound and voice data as well as challenges.

Keywords - sound data, voice data, Smart Home's, voice and sound recognition, sound and voice feature extraction, MFCC.

I. PENDAHULUAN

Penelitian tentang bunyi dan suara telah lama diteliti dan dikembangkan. Berdasarkan catatan yang ada, penelitian tentang bunyi diawali dengan keinginan manusia untuk meniru ucapan manusia. Berdasarkan pada keinginan tersebut, maka muncul penelitian yang mengembangkan teknologi untuk merekam dan memutar kembali bunyi. Pada buku yang berjudul "Voice Recognition" oleh Richard L Klevans and Robert D Rodman, disebutkan bahwa sejarah awal pengenalan suara adalah usaha perekaman bunyi [1].

Sistem pengenalan suara manusia dan bunyi juga merupakan bagian dari *Artificial Intelligence* karena memiliki tujuan untuk meniru kemampuan manusia dalam mendengar dan mengenali suara atau bunyi

tertentu. Kemampuan manusia dalam mengenali bunyi suara juga terbatas, terlebih jika informasi suara yang ditangkap sangat minim dan banyak gangguan atau *noise*. Secara umum tujuan dari penelitian pemrosesan bunyi adalah proses pengenalan, klasifikasi atau pengelompokan dimana manusia pun masih belum tentu memiliki kemampuan itu. Misal pada pemanfaatan bunyi pada bahan makanan di rumah tangga, secara harafiah manusia tidak memiliki kemampuan untuk membedakan bunyi tertentu yang dihasilkan pada makanan atau kegiatan yang berhubungan dengan makanan, misal memasak atau makan. Seperti pada penelitian oleh Kojima, Takamichi, dkk. yang mengenali sayuran dan buah berdasarkan bunyi potongnya [12]. Pada bidang medis, misal identifikasi suara paru-paru seperti pada penelitian oleh D. Wang, B. Fang, dkk [23] dan bunyi jantung seperti pada penelitian T. Gokhale [30]. Dalam hal ini kemampuan identifikasi suara atau bunyi terdapat pada seseorang yang memiliki informasi dan kemampuan dalam kasus yang lebih spesifik yaitu bidang kedokteran dan kesehatan. Pada bidang perkebunan, terdapat penelitian yang memanfaatkan bunyi gesek antara buah durian dengan suatu alat untuk mengetahui tingkat kematangan buah durian [31]. Pada bidang kesehatan balita terdapat juga penelitian yang menggunakan suara tangis bayi untuk mengidentifikasi tangis bayi [32].

Hingga saat ini berbagai program dan aplikasi yang menggunakan pemrosesan bunyi dan suara telah banyak dikembangkan, program dengan tujuan interaksi manusia dengan mesin menjadi populer untuk dikembangkan [33], [34]. Hingga berlanjut perkembangannya pada aplikasi *Smart Homes*. Program pendukung pada *Smart Homes* juga menjadi penelitian yang sedang dikembangkan saat ini [35], [2], [3].

Penelitian di bidang Teknologi Informasi akan berujung pada suatu aplikasi yang dapat digunakan dan bermanfaat bagi kehidupan manusia. Namun, pembuatan aplikasi selalu didahului oleh beberapa penelitian mendasar terlebih dahulu yang hasilnya digunakan dalam pembuatan aplikasi yang diinginkan. Pada aplikasi berbasis bunyi dan suara juga diperlukan penelitian mendasar untuk mendukung dibuatnya suatu aplikasi berbasis suara manusia dan bunyi. Terdapat topik penelitian mendasar yang umum dilakukan, antara lain pengembangan metode ekstraksi ciri suara dan bunyi [13]-[16] [20], metode klasifikasi dan *clustering* bunyi dan suara [4], [5], [6], [7], [8], [35], metode penghilangan *noise* dan gangguan pada data bunyi dan suara [35], [9], serta penelitian tentang metode pengenalan suara itu sendiri [35], [10], [11], [13]. Juga

terdapat topik tentang metode pengambilan data suara, terkait dengan alat dan teknik yang digunakan untuk merekam data suara atau bunyi yang akan digunakan sebagai data eksperimen pada penelitian. Juga terdapat penelitian yang membahas tentang penggunaan *tools* dan *software* pendukung berbasis bunyi dan suara [10], [14], [13]. Selain penelitian mendasar yang juga dibutuhkan, pengetahuan dan informasi tentang sinyal dan pengolahannya dibutuhkan mendukung penelitian yang akan dilakukan, seperti pada penelitian [15], [16].

Pada penelitian yang dikaji, akan dibahas dua aplikasi yang telah dipublikasikan pada seminar dan jurnal ilmiah. Kedua penelitian yang dikaji menggunakan bunyi atau suara sebagai data dan media penelitian serta aplikasinya. Pembahasan akan dijabarkan pada sub bab Teknologi Berbasis Suara Manusia dan Bunyi. Namun, sebelum masuk pada pembahasan penelitian, akan dibahas secara singkat tentang suara dan bunyi sebagai dasar informasi pada kedua penelitian yang akan dibahas. Selain itu juga akan dibahas tentang metode ekstraksi ciri suara dan bunyi yang diambil dari beberapa sumber penelitian. *Noise* pada bunyi dan suara juga dibahas karena pembahasan tentang *noise* tidak bisa lepas dari penelitian bunyi dan suara.

Kedua penelitian akan menjadi contoh aplikasi atau teknologi berbasis bunyi dan suara yang dimanfaatkan pada kehidupan sehari-hari. Masing-masing penelitian akan dilihat metode yang digunakan, teknik pengumpulan data dan pengolahannya serta hasil eksperimen yang dilakukan.

II. METODOLOGI PENELITIAN

A. Suara dan Bunyi

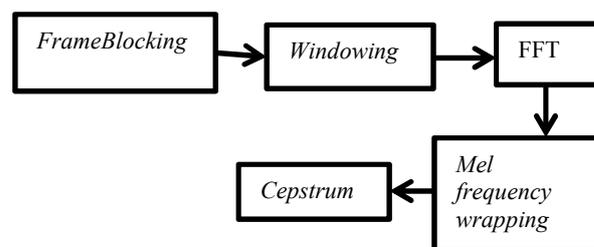
Pada lingkungan nyata, bunyi terdiri dari beberapa jenis, yaitu bunyi atau suara manusia, suara hewan, bunyi yang dihasilkan dari suatu kejadian atau benda mati, misal suara pintu diketuk, bunyi benda jatuh, bunyi klakson kendaraan, dan lain sebagainya. Bunyi merupakan gelombang yang dihasilkan akibat benda bergetar (sumber bunyi) yang merambat pada zat tertentu (zat peramban) [15], [16]. Gelombang bunyi yang dihasilkan dari sumber suara dapat berbeda satu dengan yang lain, bergantung dari sumber bunyi, media peramban bunyi, kecepatan bunyi, dan waktu [15], [16]. Pada penelitian di bidang Teknik Informatika, gelombang analog yang dihasilkan oleh sumber bunyi ini diubah menjadi sinyal digital melalui proses perekaman. Salah satu format file yang umum digunakan adalah *.wav*, isi dari file *wav* terdiri dari beberapa bagian yaitu "RIFF" chunk descriptor yang

berisi menyatakan bahwa file tersebut adalah file berformat .wav [17]. Bagian “RIFF” ini berisi dua sub-chunk yaitu “fmt” sub-chunk yang berisi informasi data, seperti *sample rate*, *byte rate*, *num channels*, etc. [17]. Sub chunk kedua adalah “data” yang berisi data mentah suara beserta ukurannya [17]. Bentuk data digital ini bergantung salah satunya pada sampling rate yaitu berapa banyak bit yang tersimpan pada satuan waktu tertentu (second atau mili second), yang biasanya berada pada satuan Hz (Nilai yang digunakan adalah 8000 Hz, 16000 Hz, 24000 Hz atau 48000 Hz, semakin tinggi frekuensinya semakin rinci data yang didapat [17].

1) *Noise pada Data Suara*. Noise pada sistem pengenalan suara dan bunyi menjadi penghambat pada peningkatan akurasi sistem. Namun, keberadaan noise tidak dapat dihindari, mengingat lingkungan pengambilan data harus senatural mungkin atau sesuai dengan kondisi nyata saat sistem akan diimplementasikan. Sistem yang diterapkan pada suatu lingkungan yang tetap, misal: kantor atau tempat tinggal akan mempermudah sistem memprediksi noise yang akan terjadi. Tapi jika sistem digunakan pada berbagai lingkungan yang tidak dapat diprediksikan, maka akan membuat noise menjadi sulit dihilangkan, misal: di stasiun, di dalam kendaraan umum, di jalan dan tempat lainnya yang memiliki banyak sumber suara dan bunyi.

Terdapat beberapa metode untuk mengurangi atau menghilangkan noise. Pada penelitian yang dilakukan oleh Valin, dkk penghilangan noise pada sistem dilakukan dengan menggunakan Echo Cancellation System [9]. Algoritma diimplementasikan pada SPEEX library dibawah lisensi GPL [9]. Metode yang digunakan adalah *Minimum Mean Square Estimator Short-Time Amplitude Spectrum Estimator* (MMSE-STSA).

2) *Ekstraksi Ciri Suara*. Data suara manusia terdiri dari kumpulan tinggi rendahnya suatu frekuensi. Perbedaan dari tinggi dan rendahnya frekuensi sinyal ini menjadi ciri bunyi yang ada. Namun, ini tidak cukup untuk menjadi ciri bagi suatu bunyi atau suara. Penangkapan ciri suara yang umum digunakan adalah metode MFCC. Metode MFCC menjadi populer karena mampu meniru kemampuan manusia dalam menangkap ciri suara atau bunyi. Penggunaan MFCC dapat dilihat pada beberapa penelitian, seperti pada [18], [19], [20], [21].



Gambar 1. Tahapan MFCC

Pada **Gambar 1** menunjukkan tahapan yang dilakukan pada MFCC [18]. Setiap tahapan memiliki tujuan dan pemrosesan yang berbeda. Metode MFCC merupakan metode ekstraksi ciri suara yang bertujuan untuk menangkap ciri suara berdasarkan sinyal-sinyal diskritnya [18]. Sinyal diskrit yang berbasis waktu akan diubah menjadi berbasis frekuensi yang lebih mudah diteliti. Metode MFCC dikembangkan berdasarkan studi psychophysical yang menyebutkan bahwa suara manusia tidak bersifat linear, pada MFCC hal ini menjadi dasar untuk melakukan filtering dengan skala mel. Skala mel bersifat linear pada frekuensi suara bernilai lebih kecil dari 1000 Hz sedangkan di atas 1000 Hz bersifat logaritmik [18]. Langkah selanjutnya adalah mengubah kembali spektrum log mel menjadi spektrum waktu menggunakan metode discrete cosine transform (DCT) dan hasilnya disebut sebagai mel frequency cepstrum coefficients. Hasil dari MFCC merupakan ciri suara yang menjadi input pada proses selanjutnya.

Tujuan dari frame blocking adalah melakukan segmentasi dari sinyal suara [18]. Segmentasi dilakukan karena kecepatan dari pengucapan kata tiap individu berbeda, dengan adanya segmentasi frame panjang data akan sama. Proses frame blocking akan dilakukan secara overlapping agar tingkat kontinuitas datanya tetap terjaga.

Tahapan selanjutnya adalah windowing. Tujuan dari tahap windowing adalah untuk menghaluskan hasil pemotongan dari tahapan frame blocking. Proses windowing ini untuk menyetarakan data pada tiap *frame* [18]. Metode yang umum digunakan adalah *Hamming window* [18]. Data suara yang diambil berada pada domain waktu, hal ini mengakibatkan kesulitan pada proses perhitungan dan analisa. Proses MFCC pada tahap *Fast Fourier Transform* (FFT), mengubah domain waktu menjadi domain frekuensi sehingga akan lebih mudah untuk dianalisis [18].

Setelah data diubah menjadi domain frekuensi maka tahap selanjutnya adalah *mel frequency wrapping*. Tahap ini menyesuaikan frekuensi yang didapat dengan frekuensi suara manusia, sehingga akan menghasilkan ciri suara yang bersesuaian dengan cara manusia menangkap ciri suara [18]. Tahap terakhir dari ekstraksi

ciri suara manusia adalah cepstrum. Data frekuensi diubah kembali menjadi data berdomain waktu. Hasil dari tahap cepstrum menjadi koefisien MFCC merupakan frekuensi dominan yang didapat dan digunakan sebagai hasil ekstraksi ciri suara [18].

Ukuran data akhir ekstraksi ciri suara adalah sebesar $n \times m$, dimana n adalah jumlah dari data yang diinginkan (di-input pada tahap cepstrum), sedangkan m adalah panjang data yang didapat saat proses perubahan data analog menjadi data digital dan hasil pada tahap frame blocking. Hasil dari MFCC ini menjadi input pada proses selanjutnya yaitu proses pelatihan dan pengenalan.

Terdapat beberapa penelitian yang membahas dan menjelaskan secara rinci penggunaan tools untuk melakukan ekstraksi ciri suara dan bunyi. Tools yang digunakan juga berbeda-beda pada setiap penelitian. Matlab dan Python merupakan dua bahasa pemrograman yang menyediakan library untuk melakukan pemrosesan data bunyi. Pada penelitian oleh J. Glover, V. Lassarini dan J. Timoney dilakukan pemrosesan bunyi dan suara menggunakan tools yang ada pada Python [22]. Pada bahasa pemrograman Python terdapat beberapa library yang menyediakan adanya pengolahan data suara yaitu SciPy. Pada library SciPy telah menyediakan pembacaan data file .wav, hingga tidak perlu dilakukan pengkodean manual untuk membaca file .wav. Keluaran yang dihasilkan berupa array satu dimensi, isi datanya adalah amplitudonya, sedangkan dimensinya adalah waktu (sampel). Pada library SciPy ini juga dapat menghasilkan nilai Fast Fourier Transform (FFT). Pada penelitian yang dilakukan oleh J. Glover, V. Lassarini dan J. Timoney, menggunakan library yang bersifat open source, yaitu Simpl. Simpl dapat digunakan pada bahasa C/C++ dan Python. Fungsi dari library Simpl adalah untuk melakukan pemodelan sinusoidal, yang dapat digunakan untuk menganalisa sinyal bunyi [22].

III. HASIL DAN PEMBAHASAN

A. Pemanfaatan bunyi pada makanan di rumah tangga (Cogknife)

Penelitian yang dilakukan pada penelitian dengan judul "CogKnife: Food recognition from their cutting sounds" adalah membangun sebuah prototype untuk mengenali bahan makanan dari bunyi potongnya yang diberi nama *CogKnife (Cognitive Knife)* [12]. Latar belakang penelitian yang dilakukan adalah karena *monitoring* dan perekaman data makanan pada rumah tangga menjadi penting untuk diteliti sebab terdapat beberapa permasalahan yang dapat ditimbulkan, antara

lain limbah makanan atau penyakit yang disebabkan dengan gaya hidup tertentu [12].

Penelitian ini menyebutkan bahwa *monitoring* konsumsi makanan dan aplikasinya merupakan bagian dari human *computer interaction*. Metode yang digunakan ada tiga yaitu kNN, SVM dan CNN, ketiga metode tersebut dibandingkan dan diujikan dengan metode *cross validation* [12].

Pada penelitian ini juga dijabarkan beberapa penelitian terkait untuk mendukung latar belakang penelitian yang dilakukan dan aplikasi yang dibuat dengan menggunakan data bunyi serta suara, antara lain:

Cooking support, terdapat beberapa aplikasi dan penelitian yang bertujuan untuk membuat sistem berhubungan dengan memasak, antara lain: *Cooking support system* yang diberi nama Panavi. Sistem ini mencatat suhu atau temperatur dan gerakan dari kualifikasi penggorengan dengan *thermocouple* dan *acceleration sensors* [12]. *Cooking activities for navigation*, dengan memasang beberapa sensor pada peralatan masak: kamera *infrared*, kamera digital, dan *microphone*, oleh Yamakawa, dkk, [24]. *Household food management*, sistem yang dibuat bertujuan untuk memantau dan mengatur pola makan dan makanan pada rumah tangga. Penelitian yang dilakukan antara lain: Pada Ganglbauer, E., dkk, 2012, disebutkan bahwa memonitor dan mencatat proses makan adalah hal yang penting bagi kesehatan [7]. Chi, P. Y, dkk, 2008, mengestimasi kalori dari makanan dengan menggunakan kamera yang dipasang pada meja dapur dan kompor [6]. Amift, dkk, 2005, penelitian yang dilakukan adalah mengklasifikasikan jenis makanan berdasarkan bunyi kunyahannya [4]. HAPILABS, Inc, 2016, penelitian bertujuan untuk mengukur kecepatan makan dengan cara mengukur jumlah suapan per menit [25]. Kadomura, dkk, 2014, penelitian yang dilakukan membuat *Sensing Fork* untuk mendeteksi dan memodifikasi kebiasaan cara makan [8].

Berdasarkan penelitian yang telah dilakukan di atas, sebagian besar penelitian membutuhkan alat atau sensor yang mahal karena perlu dirancang khusus sesuai dengan kebutuhan penelitian, sedangkan pada penelitian yang dilakukan hanya membutuhkan alat dan proses pemasangan alat yang sederhana dan memakan biaya yang sedikit.

Eksperimen yang dilakukan pada penelitian ini adalah dengan memasang sebuah *microphone* sederhana yang ditempelkan pada pisau yang akan digunakan untuk memotong makanan. Bunyi potong direkam per satu gerakan memotong, kemudian diambil data *spectrogram*-nya untuk data latih klasifikasi.

Digunakan enam buah sayuran dan buah: apel, pisang, wortel, kubis, daun bawang dan paprika.

Pengujian sistem dilakukan dengan *Cross Validation* dengan hasil akurasi: 95% menggunakan SVM, 83% dengan k-NN, 89% menggunakan CNN [12].

B. Pemanfaatan bunyi pada smart home.

Penelitian dengan judul “Complete Sound and Speech Recognition System for Health Smart Homes: Application to the Recognition of Activities of Daily Living,” membahas tentang pembuatan sistem Smart Homes, yang melatarbelakangi pembuatan sistem adalah peningkatan jumlah penduduk yang tidak bisa hidup secara mandiri sehingga membutuhkan bantuan untuk melakukan kegiatan dasar manusia, seperti makan, minum, mandi dan kegiatan dasar lainnya [35].

Sistem yang dibangun bertujuan untuk memberikan informasi yang membantu para penghuni rumah yang memiliki keterbatasan dalam melakukan kegiatan dasar di rumah. Informasi yang diberikan juga dijadikan peringatan bagi pengawas penghuni yang memiliki keterbatasan [35]

Paper ini merujuk pada beberapa penelitian pendahulu yang juga telah mengembangkan sistem untuk Smart Homes. Sebagian besar penelitian yang dirujuk menggunakan sensor pada sistem yang dibangun. Sensor yang digunakan adalah sensor suara atau bunyi menggunakan alat rekam *microphone* yang dipasang pada beberapa titik di rumah.

Penelitian ini mengangkat topik *sound and speech analysis* pada aplikasi *Smart Homes* sebagai *state of the art* dari penelitian yang dilakukan. *Sound* atau bunyi atau suara bisa berupa suara ucapan manusia, musik, atau bunyi dari kegiatan sehari-hari (seperti mencuci, mandi, berjalan, dan lain sebagainya).

Pada kajian penelitian ini juga dibahas hal-hal terkait yang dengan sistem yang dibuat, antara lain: *sound recognition*, *speech recognition*, aplikasi berbasis pengenalan suara manusia dan bunyi pada aplikasi untuk orang berkebutuhan khusus, dan aplikasi berbasis pengenalan suara manusia dan bunyi pada Smart Homes. Pembahasan ini menjadi merujuk pada penelitian yang telah dilakukan sebelumnya dari beberapa artikel ilmiah lainnya dan menjadi landasan penelitian dan sistem yang dibuat [35].

Penelitian tentang *sound recognition* telah banyak dikembangkan menggunakan metode *machine learning*, seperti NN, LVQ, dll. Metode ekstraksi ciri suara juga telah banyak dikembangkan pada beberapa penelitian terdahulu. Penerapan pengenalan suara juga telah banyak diimplementasikan pada perangkat Smart Home, misal pada penelitian oleh [26], untuk deteksi

kualitas dan volume air. Umumnya juga digunakan untuk mengukur suatu kejadian misal kejadian seseorang jatuh di rumah. Sebuah alat *accelerometer* dan sebuah *microphone* diletakan pada lantai. Gabungan antara bunyi dan getaran di lantai dideteksi. Penelitian tentang jumlah *microphone*, metode ekstraksi ciri, dan jenis ruangan pada rumah menjadi area atau topik penelitian yang dibahas, seperti pada [2] dan [3] yang menggunakan HMM dan MFCC, pada penelitian ini digunakan ruangan kamar mandi sebagai lingkungan eksperimennya.

Proses komunikasi manusia dilakukan dengan sangat mudah, namun dibalik proses komunikasi itu sebenarnya terdapat berbagai *variable* pada sinyal suara manusia. Berbagai hal mempengaruhi sinyal suara manusia tersebut, seperti konteks, dialek, lingkungan komunikasi terjadi. Jika terdapat suatu teks yang sama diucapkan berulang dan diucapkan pula oleh pembicara yang sama, maka sinyal akustik yang dihasilkan belum tentu sama. Metode yang dikembangkan harus mampu menangani berbagai variabel yang mempengaruhi terbentuknya sinyal akustik pada suara manusia [35].

Pada penelitian diberikan dua contoh aplikasi untuk menjelaskan hambatan pada pembuatan aplikasi *speech recognition*. Kedua aplikasi telah banyak dibangun dan diimplementasikan pada dunia nyata.

Contoh aplikasi yang pertama adalah aplikasi yang digunakan untuk mengubah suara menjadi teks pada komputer. Pada aplikasi ini dibutuhkan jenis data kata yang banyak dengan tingkat *error* yang relatif rendah, karena biasanya aplikasi menggunakan *microphone* yang bagus, dengan lingkungan yang lebih terkontrol (karena biasanya *microphone* yang digunakan terhubung dengan headset yang artinya dekat dengan sumber suara dan stabil jaraknya dengan sumber suara), hingga variabel *lingkungan* tidak terlalu berpengaruh [20]. Aplikasi juga membutuhkan tahapan belajar mengenali suara penggunaannya, karena input kata yang diberikan relatif panjang kalimat per kalimat.

Aplikasi kedua adalah aplikasi yang digunakan untuk memberikan perintah sederhana untuk mesin atau alat untuk menjalankan tugasnya. Perintah yang diberikan relatif menggunakan kata-kata yang singkat, hingga tidak diperlukan tahapan sistem belajar mengenali suara penggunaannya. Aplikasi ini membutuhkan data kata yang sedikit, namun memiliki tingkat *error* yang tinggi. Aplikasi digunakan pada lingkungan yang sulit dikontrol kondisinya, misal aplikasi harus mampu berjalan dengan baik pada berbagai kondisi, misal di tengah jalan, di kereta atau kendaraan umum, atau ada gangguan suara lain seperti suara tangis bayi yang menjadi *noise* bagi sistem. Selain

itu, *microphone* yang digunakan adalah *microphone* pada perangkat *mobile* yang umumnya memiliki *low pass filter*, sehingga sinyal yang dihasilkan tidak sebaik perangkat *microphone* pada komputer [35]. Berdasarkan kedua contoh di atas, penelitian dan aplikasi menerapkan metode yang berbeda berdasarkan tujuan dan kondisi aplikasi pengenalan suara yang dibuat.

C. Aplikasi berbasis pengenalan suara manusia dan bunyi pada aplikasi untuk orang berkebutuhan khusus.

Pada penelitian juga diberikan beberapa contoh aplikasi berbasis *speech recognition* yang telah berhasil dibuat dan dikembangkan pada penelitian sebelumnya yaitu aplikasi untuk orang dengan berkebutuhan khusus dan orang tua yang tidak mahir menggunakan komputer. Aplikasi yang dibuat menggunakan kata-kata singkat dan perintah sederhana. Beberapa aplikasi itu adalah yang dikembangkan oleh France Telecom Research and Development yang membuat aplikasi untuk orang berkebutuhan khusus atau orang jompo untuk melakukan panggilan darurat. Pada penelitian oleh Fezari & Bousbia-Salah, 2007, membuat aplikasi untuk mengontrol kursi roda dengan menggunakan suara manusia [27]. Pada penelitian oleh Kumiko et al., 2004, mengembangkan aplikasi bagi pengguna yang tidak bisa menggunakan keyboard pada komputer (misal pengguna yang sudah tua dan tidak mengenal teknologi dengan baik), maka fungsi kerja keyboard digantikan dengan memberikan kontrol berupa suara pada *keyboard* [28]. Pada penelitian oleh Anderson et al., 1999, juga mengembangkan sistem berbasis pengenalan suara bagi pengguna yang sudah tua untuk memperoleh informasi pada database dokumen [5].

D. Aplikasi berbasis pengenalan suara manusia dan bunyi pada Smart Homes

Pada penelitian ini membagi dua aplikasi Smart Homes berbasis pengenalan suara manusia dan bunyi berdasarkan tujuannya, yaitu *Home Automation dan Recognition of Distress in Smart Home* [35].

Pada penelitian oleh Wang, J.-F. & Lin, C.-B., 2008, penelitian yang dilakukan berbasis *sound classification* untuk menjalankan perintah yang diberikan dengan suara manusia, perangkat penangkap suara (*microphone*) dipasang pada langit-langit ruangan. Metode klasifikasi yang digunakan adalah *Support Vector Machine* dan metode *MFCC* digunakan untuk ekstraksi ciri suaranya [29].

Aplikasi kedua yang dibuat adalah untuk mendeteksi *distress* (kesulitan) pada Smart Homes. Hambatan ditekni melalui suara atau bunyi yang umumnya

terdengar saat kejadian *distress* terjadi, seperti bunyi gelas pecah, teriakan, atau benda jatuh. Penelitian yang mendasari aplikasi ini adalah penelitian yang dilakukan oleh [10]. Penelitian ini menghasilkan database suara yang dihasilkan dari kehidupan sehari-hari pada sebuah ruangan. Proses perekaman suara menggunakan dua *microphone* yang dipasang di dapur. Berdasarkan data yang dikumpulkan penelitian ini membedakan beberapa bunyi, seperti bunyi dering telepon, sendok jatuh, cangkir jatuh dan bunyi lainnya menggunakan metode *Gaussian Mixture Model*. Terdapat penelitian lain yang juga mengklasifikasikan bunyi yang ada pada perkantoran. Bunyi yang direkam berasal dari kegiatan sehari-hari di sebuah kantor, algoritma yang digunakan adalah *unsupervised*, yaitu dengan menggabungkan keseluruhan jenis data bunyi lalu dilakukan proses *clustering* [11]. Hasil dari pengelompokan bunyi dijadikan acuan untuk melakukan klasifikasi. Terdapat juga penelitian yang berfokus pada penggunaan alat rekam yang umum dan murah, yaitu *microphone* dan *soundcard*, data bunyi yang dikumpulkan juga berasal dari kegiatan *distress* di tempat tinggal [14].

Pada penelitian lain oleh Takahashi, dkk., mengembangkan sebuah robot yang mampu berkomunikasi dengan orang lansia untuk mengurangi rasa kesepian mereka. Robot yang dibuat menggunakan *voice synthesis* dan *speech recognition* [13]. Kesulitan dan hambatan pengembangan sistem pengenalan bunyi dan suara berasal dari kondisi perekaman. Umumnya, alat rekam atau *microphone* dipasang berjauhan dari *speaker* dan banyak dipasang di langit-langit ruangan. Kondisi lingkungan perekaman yang memiliki banyak *noise* dan gema akan mempersulit proses pengenalan. Hal ini menjadikan sistem memiliki kesulitan dalam memproses suara dan bunyi. Proses pengenalan dan sinyal suara diproses dan diadaptasi berdasarkan permasalahan yang ada, misal *noise* atau gema. Selain itu, pada pengenalan suara atau kata pada manusia membutuhkan model ucapan untuk identifikasi bunyi *phenome*, dan model bahasa untuk mengenali kata berdasarkan situasi tertentu (misal: memberikan perintah atau dalam percakapan). Sistem juga disesuaikan dengan kondisi pengguna, misal suara yang dihasilkan oleh para lansia memiliki karakteristik yang berbeda dengan non-lansia [33]. Pada penelitian oleh Baba, dkk., membangun model akustik khusus berdasarkan karakteristik pengguna, dalam hal ini para lansia [34]. Penelitian yang dilakukan adalah mengenali bunyi dan suara pada kehidupan sehari-hari di ruangan apartemen atau flat dengan input suara manusia berbahasa Perancis (*France*) untuk mengenali kondisi gawat atau bahaya [35]. Sistem yang dibangun

diberi nama AUDITHIS dan RAPHAEL. AUDITHIS merupakan sistem yang digunakan untuk mengenali bunyi, sedangkan RAPHAEL merupakan sistem untuk mengenali suara manusia. Segmentasi bunyi dan suara dilakukan dengan *Gaussian Mixture Models* (GMM) dengan ekstraksi ciri dengan *Linear-Frequency Cepstral Coefficients* (LFCC) [20]. Proses klasifikasi bunyi diuji dengan menggunakan GMM dan HMM, hasilnya HMM lebih unggul pada kondisi bebas noise sedangkan untuk bunyi dengan noise tinggi GMM lebih unggul, hingga GMM yang digunakan pada sistem yang dibuat. Sistem RAPHAEL dibangun secara independen, menggunakan tiga tahapan yaitu: tahap pengenalan fonem, tahap pengenalan kata dengan membentuk graph phoneme, tahap pengenalan kalimat dengan membentuk graph kata.

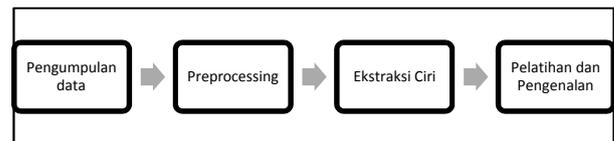
Percobaan yang dilakukan dengan membentuk corpus data dari 10 subjek orang dan beberapa bunyi dari kegiatan sehari-hari di beberapa ruangan seperti kamar mandi, dapur, ruang tamu, dan kamar tidur. Algoritma learning yang diujikan ada beberapa yaitu *Decision Tree* (C4.5), *Decision Table Majority* (DTM), *Naïve Bayesian Network* (NBayes) dan *Support Vector Machine* (SVM) dengan beberapa jenis set data, yaitu *Whole set*, *No sound*, *PIR only*, *Sound only*, GF, GW. *Whole set* merupakan keseluruhan data, *No sound* adalah data suara manusia saja tanpa bunyi lainnya, *PIR only* adalah suara manusia dan langkah kaki saja, *Sound only* adalah tanpa suara manusia hanya bunyi saja, GF adalah *Global Filtering* yaitu pemilihan atribut dengan *Global Filtering* dan pemilihan atribut dengan *Global Wrapping* adalah GW [35]. Berdasarkan hasil eksperimen menunjukkan metode NBayes memiliki akurasi paling baik menggunakan *Whole set*. Sedangkan untuk *No Sound*, *PIR only*, *Sound only*, dan GW metode SVM memiliki akurasi yang lebih tinggi dibanding metode lainnya. Pada data GF metode NBayes memiliki akurasi paling tinggi. Hingga dapat disimpulkan bahwa metode SVM dapat lebih menangani berbagai jenis data dari pada metode lainnya. Sedangkan data yang memiliki rata-rata akurasi tertinggi adalah GW yaitu 83.87 persen [35].

Berdasarkan penelitian dan survei dari beberapa penelitian yang telah dibahas pada sub bab sebelumnya, dapat dilihat bahwa terdapat beberapa permasalahan yang diteliti. Hal yang diteliti menjadi bagian-bagian kecil yang menjadi dasar aplikasi berbasis suara atau bunyi. Pada bab ini akan dilihat rangkuman secara umum hal-hal yang dapat menjadi topik penelitian. Pada bagian awal akan dijelaskan dengan model dan tahapan-tahapan penelitian yang umum dilakukan yang

kemudian dijadikan landasan untuk menentukan lingkup penelitian yang dapat dikembangkan.

Pada bagian berisi penjelasan ilmiah dari hasil penerapan metode penelitian yang telah ditetapkan pada sub bab 3.

1) *Kerangka kerja pada penelitian bunyi dan suara*. Secara umum terlepas dari tujuan dan sasaran pengguna aplikasi yang dibuat, dapat disimpulkan bahwa terdapat kerangka kerja yang digunakan. Pada **Gambar 2** menunjukkan kerangka kerja yang digunakan.



Gambar 2. Kerangka kerja

Langkah awal yang dikerjakan adalah pengumpulan data. Data yang dikumpulkan untuk menjadi sumber informasi bagi sistem yang dibuat. Data yang dikumpulkan bergantung pada tujuan penelitian dan aplikasi yang dibuat. Data dapat berupa bunyi atau suara yang disimpan ke dalam format digital untuk dapat diolah. Data bunyi atau suara yang dikumpulkan dapat berupa data bunyi yang direkam pada lingkungan atau kondisi nyata, dapat pula direkam pada lingkungan yang terkontrol untuk meminimalkan variabel dan noise yang tak terkontrol.

Langkah selanjutnya adalah *Preprocessing*, pada langkah ini bertujuan untuk mempersiapkan data yang akan digunakan pada sistem. Proses yang dilakukan antara lain: memotong-motong bunyi berdasarkan kriteria tertentu, misal: per kata, per kalimat atau per satuan waktu. Penguatan sinyal juga dapat dilakukan pada proses ini. Pengurangan atau pelemahan *noise* juga menjadi tahapan yang dilakukan guna meningkatkan kualitas data.

Ekstraksi ciri menjadi langkah berikutnya. Namun, pada beberapa penelitian ada yang menempatkan ekstraksi ciri pada tahapan *preprocessing* dan ada juga yang meletakkannya pada tahapan pelatihan, ini bergantung pada tujuan dan fokus penelitian yang dilakukan. Ekstraksi ciri suara atau bunyi dilakukan dengan tujuan untuk menentukan ciri dari bunyi yang didapat dengan melakukan pengolahan sinyal. Pengolahan sinyal secara *discrete* adalah metode yang umum digunakan. Melalui ekstraksi ciri diharapkan dapat ditentukan pembeda antar data bunyi yang didapat. Pada pengenalan kata digunakan ciri bunyi kata misalnya bunyi fonem, kecepatan pengucapan, dan intonasi.

Proses pelatihan dilakukan dengan memberikan label pada data ekstraksi ciri yang didapat. Proses pelatihan dapat juga dilakukan untuk membentuk model ciri bunyi yang akan digunakan untuk pengenalan. Proses pengenalan bunyi dilakukan dengan memanfaatkan basis data bunyi yang telah dibentuk pada proses pelatihan. Proses pengenalan menjadi tujuan umum dari suatu sistem pada level aplikasi.

2) *Lingkup penelitian*. Berdasarkan kerangka kerja yang telah dijabarkan, dapat ditentukan lingkup penelitian yang dapat dilakukan, antara lain: teknik pengumpulan data, pengolahan noise, teknik atau model ekstraksi ciri bunyi dan suara, teknik pelatihan dan pengenalan suara atau bunyi, penerapan penelitian pemrosesan bunyi. Teknik pengumpulan data dapat menjadi bagian penelitian yang dilakukan. Bagian ini menjadi penting untuk diteliti karena berkaitan dengan tujuan dan batasan penelitian yang akan dilakukan. Alat dan lingkungan pengambilan data juga menjadi faktor yang diperhatikan pada penelitian ini. Jika aplikasi yang dibuat digunakan pada lingkungan yang terkontrol akan berbeda teknik pengumpulan datanya pada lingkungan yang tidak terkontrol, baik dari kecepatan perekaman, alat yang digunakan, setting alat rekam dan lain sebagainya. Pada lingkup ini, berbagai penggunaan dan pemanfaatan perangkat lunak serta metode perekaman bunyi juga dapat dibahas dan diteliti.

Pengolahan *noise*, seperti yang telah dibahas sebelumnya noise menjadi kendala pada penelitian bunyi dan suara, terutama pada lingkungan yang tidak terkontrol. Penelitian yang dilakukan bertujuan mengembangkan metode untuk mengurangi *noise*. Terdapat pendekatan yang dapat digunakan yaitu dengan menguatkan sinyal yang bukan *noise*, ada juga pendekatan dengan mencari ciri bunyi noise, kemudian menghilangkan bunyi *noise* berdasarkan ciri yang didapat, atau dengan kata lain memisahkan antara bunyi sesungguhnya dengan bunyi *noise*.

Teknik atau model ekstraksi ciri bunyi dan suara menjadi salah satu lingkup penelitian yang dilakukan. Hal ini menjadi tantangan juga pada penelitian tentang bunyi dan suara. Bunyi dan suara tentu memiliki ciri yang berbeda berdasarkan sumber bunyi dan waktu tertentu. Ciri yang didapatkan harus dapat mewakili label atau kelompok tertentu, baik yang sudah ditentukan atau pun belum. Teknik yang dibangun juga harus menentukan ciri yang dapat membedakan suara yang satu dengan yang lain. Terdapat beberapa metode yang umum digunakan yaitu *Fast Fourier Transform* dan *Mel Frequency Cepstrum Coefficients*, atau metode ekstraksi ciri bunyi yang lain.

Teknik pelatihan dan pengenalan suara atau bunyi memiliki beberapa jenis pendekatan antara lain, klasifikasi dan clustering dengan menggunakan metode pada *Machine Learning* atau dengan pendekatan lain yang digunakan untuk melakukan pengenalan objek, misal dengan Regresi, *Hidden Markov Model*, *Support Vector Machine*, *k Nearest Neighbour*, *Neural Network* dan metode *Machine Learning* lainnya.

Penerapan penelitian pemrosesan bunyi dapat dilakukan pada beberapa bidang, antara lain untuk keamanan dan kontrol device, pada bidang perkebunan untuk deteksi kematangan buah seperti pada penelitian oleh Rong Phoophuangpairoj, yang melakukan penelitian tentang identifikasi kematangan buah durian berdasarkan bunyi gesek durian dengan alat tertentu [31]. Pada bidang otomotif untuk mendeteksi kerusakan mesin berdasarkan bunyi mesinnya seperti pada penelitian. Bidang kesehatan seperti pada penelitian [22] [30] yang menggunakan data bunyi jantung dan paru-paru untuk deteksi penyakit.

IV. KESIMPULAN

Setelah mengkaji penelitian tentang bunyi dapat disimpulkan bahwa masih banyak terdapat permasalahan yang dapat diangkat untuk menjadi penelitian baru, baik dari segi aplikasi maupun teori dan metode untuk mendukung pembuatan aplikasi. Pemanfaatan bunyi sebagai sumber data dan aplikasi juga masih sangat luas. Topik penelitian juga dapat dikembangkan berdasarkan subjek implementasi penelitian, seperti pada bidang otomotif, Smart Homes, dan kesehatan. Penelitian dengan bunyi dan suara masih dapat digali dengan memanfaatkan penelitian lain yang mungkin tidak berhubungan langsung dengan bunyi, seperti metode klasifikasi dan klastering atau metode lainnya. Diharapkan melalui kajian yang telah dibahas ini dapat memperluas wawasan dan membuka lingkup penelitian lainnya.

DAFTAR PUSTAKA

- [1] R. L. Klevans, R.D. Rodman. R. L. Klevans, R.D. Rodman 1997a. *Voice Recognition*. Artech House, Michigan University.
- [2] Mihail Popescu et al. Mihail Popescu et al. 2008b. "An Acoustic Fall Detector System That Uses Sound Height Information to Reduce the False Alarm Rate." In *30th Annual International IEEE EMBS Conference*, Vancouver, British Columbia, Canada, 4628–31.
- [3] Jianfeng Chen et al. Jianfeng Chen et al. 2005c. "Bathroom Activity Monitoring Based on Sound BT - Pervasive Computing." In eds. Hans -W. Gellersen, Roy Want, and Albrecht Schmidt. Berlin, Heidelberg: Springer Berlin Heidelberg, 47–61.

- [4] Oliver Amft, Mathias Stäger, Paul Lukowicz, Gerhard Tröster Oliver Amft, Mathias Stäger, Paul Lukowicz, Gerhard Tröster 2005d. "Analysis of Chewing Sounds for Dietary Monitoring." : 56–72.
- [5] S Anderson et al. S Anderson et al. 1999e. "Recognition of Elderly Speech and Voice-Driven Document Retrieval." In *1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. ICASSP99 (Cat. No. 99CH36258)*, , 145–48 vol.1.
- [6] Pei-Yu (Peggy) Chi, Jen-Hao Chen, Hao-Hua Chu, Jin-Ling Lo Pei-Yu (Peggy) Chi, Jen-Hao Chen, Hao-Hua Chu, Jin-Ling Lo 2008f. "Enabling Calorie-Aware Cooking in a Smart Kitchen." In *Proceedings of the 3rd International Conference on Persuasive Technology, PERSUASIVE '08*, Berlin, Heidelberg: Springer-Verlag, 116–27.
- [7] Eva Ganglbauer, Geraldine Fitzpatrick, Georg Molzer Eva Ganglbauer, Geraldine Fitzpatrick, Georg Molzer 2012g. "Creating Visibility: Understanding the Design Space for Food Waste." In *Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia, MUM '12*, New York, NY, USA: ACM, 1:1--1:10.
- [8] Azusa Kadomura et al. Azusa Kadomura et al. 2014h. "Persuasive Technology to Improve Eating Behavior Using a Sensor-Embedded Fork." In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp '14*, New York, NY, USA: ACM, 319–29.
- [9] J M Valin J M Valin 2007i. "On Adjusting the Learning Rate in Frequency Domain Echo Cancellation With Double-Talk." *IEEE Transactions on Audio, Speech, and Language Processing* 15(3): 1030–34.
- [10] Maunder D., Ambikairajah E., Epps J. Maunder D., Ambikairajah E., Epps J. 2008j. "Dual-Microphone Sounds of Daily Life Classification for Telemonitoring in a Noisy Environment." *Conference proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference 2008*(1): 4636–39.
- [11] A Harma, M F McKinney, J Skowronek A Harma, M F McKinney, J Skowronek 2005k. "Automatic Surveillance of the Acoustic Activity in Our Living Environment." In *2005 IEEE International Conference on Multimedia and Expo*, , 4 pp.
- [12] T Kojima et al. T Kojima et al. 2016l. "CogKnife: Food Recognition from Their Cutting Sounds." In *2016 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, , 1–6.
- [13] Shin-ya Takahashi, Tsuyoshi Morimoto, Sakashi Maeda, Naoyuki Tsuruta Shin-ya Takahashi, Tsuyoshi Morimoto, Sakashi Maeda, Naoyuki Tsuruta 2003m. "Dialogue Experiment for Elderly People in Home Health Care System." In *International Conference on Text, Speech and Dialogue*, eds. Václav Matoušek and Pavel Mautner. Berlin, Heidelberg: Springer Berlin Heidelberg, 418–23.
- [14] Dan Istrate, Michel Vacher, Jean-François Serignat Dan Istrate, Michel Vacher, Jean-François Serignat 2008n. "Embedded Implementation of Distress Situation Identification through Sound Analysis." *Journal on Information Technology in Healthcare (JITH)* 6(3): 204–11.
- [15] Richard E. Berg, David G. Stork, Brian Holmes Richard E. Berg, David G. Stork, Brian Holmes 1982o. "The Physics of Sound." *American Journal of Physics* 50(10): 953–54.
- [16] H.J. Pain H.J. Pain 2005p. *THE PHYSICS OF VIBRATIONS*. 6th ed. John Wiley & Sons Ltd.
- [17] "WAVE PCM Soundfile Format."
- [18] Todor Dimitrov Ganchev Todor Dimitrov Ganchev 2005r. Wire Communications Laboratory Department of Computer and Electrical Engineering University of Patras Greece "Speaker Recognition."
- [19] Rashidul Hasan, Mustafa Jamil, Golam Rabbani, Saifur Rahman Rashidul Hasan, Mustafa Jamil, Golam Rabbani, Saifur Rahman 2004s. "Speaker Identification Using Mel Frequency Cepstral Coefficients." *Proceedings of the 3rd International Conference on Electrical & Computer Engineering (ICECE 2004)* (December): 28–30.
- [20] Lindsalwa Muda, Mumtaj Begam, I. Elamvazuthi Lindsalwa Muda, Mumtaj Begam, I. Elamvazuthi 2010t. "Voice Recognition Algorithms Using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques." 2(3): 138–43.
- [21] Mandar Gilke, Pramod Kachare, Rohit Kothalikar, Varun Pius Rodrigues Mandar Gilke, Pramod Kachare, Rohit Kothalikar, Varun Pius Rodrigues 2012u. "MFCC-Based Vocal Emotion Recognition Using ANN." *2012 International Conference on Electronics Engineering and Informatics (ICEEI 2012)* 49(Iceei): 150–54.
- [22] John C. Glover, Victor Lazzarini, Joseph Timoney John C. Glover, Victor Lazzarini, Joseph Timoney 2011v. "Python for Audio Signal Processing."
- [23] Dongfang Wang, Shuangwei Wang Dongfang Wang, Shuangwei Wang 2016w. "The Research Progress about the Intelligent Recognition of Lung Sounds Normal Lung." In *2016 2nd IEEE International Conference on Computer and Communications The, IEEE*, 769–72.
- [24] Y. Yamakawa, T. Shoji, K. Kakusho, M. Minoh Y. Yamakawa, T. Shoji, K. Kakusho, M. Minoh 2005x. *Automatic Cooking Archiving with Spoken Dialogue with Assistant Agent*.
- [25] Inc HAPILABS Inc HAPILABS "Hapifork."
- [26] Alejandro Ibarz et al. Alejandro Ibarz et al. 2008z. "Design and Evaluation of a Sound Based Water Flow Measurement System BT - Smart Sensing and Context." In eds. Daniel Roggen et al. Berlin, Heidelberg: Springer Berlin Heidelberg, 41–54.

- [27] M. Fezari, Bousbia-SalahM. Fezari, Bousbia-Salah 2007aa. "Speech and Sensor in Guiding an Electric Wheelchair." *M. Aut. Conrol Comp. Sci.*
- [28] O. Kumiko et al.O. Kumiko et al. 2004ab. *Input Support for Elderly People Using Speech Recognition.*
- [29] J Wang, H Lee, J Wang, C LinJ Wang, H Lee, J Wang, C Lin 2008ac. "Robust Environmental Sound Recognition for Home Automation." *IEEE Transactions on Automation Science and Engineering* 5(1): 25–31.
- [30] T. GokhaleT. Gokhale 2017ad. "Machine Learning Based Identification of Pathological Heart Sounds." *Computing in Cardiology* 43: 553–56.
- [31] Rong PhoophuangpairrojRong Phoophuangpairroj 2014ae. "Durian Ripeness Striking Sound Recognition Using N-Gram Models with N-Best Lists and Majority Voting." In *Recent Advances in Information and Communication Technology*, eds. Sirapat Boonkrong, Herwig Unger, and Phayung Meesad. Cham: Springer International Publishing, 167–76.
- [32] Medhanita Dewi Renanti, Agus Buono, Wisnu Ananta KusumaMedhanita Dewi Renanti, Agus Buono, Wisnu Ananta Kusuma 2013af. "Infant Cries Identification by Using Codebook as Feature Matching, and MFCC as Feature Extraction." *Journal of Theoretical and Applied Information Technology* 56(3): 437–42.
- [33] J G Wilpon, C N JacobsenJ G Wilpon, C N Jacobsen 1996ag. "A Study of Speech Recognition for Children and the Elderly." In *1996 IEEE International Conference on Acoustics, Speech, and Signal Processing Conference Proceedings*, , 349–52 vol. 1.
- [34] A. Baba et al.A. Baba et al. 2004ah. "Acoustic Models of the Elderly for Large-Vocabulary Continuous Speech Recognition." *Electronics and Communications in Japan* 87: 49–57.
- [35] Michel Vacher et al.Michel Vacher et al. 2010ai. "Complete Sound and Speech Recognition System for Health Smart Homes: Application to the Recognition of Activities of Daily Living." In *New Developments in Biomedical Engineering*, , 645–73.