Ultrasonic array doppler sensing for human movement classification

Ultrasonic Array Doppler Sensing for Human Movement Classification

Ruud J. G. van Sloun, Sriram Srinivasan, *Senior Member, IEEE*, Ashish Pandharipande, *Senior Member, IEEE*, and Piet C. W. Sommen

***Abstract*— Classification of human movements is an important problem in healthcare and well-being applications. An ultrasonic array Doppler sensing method is proposed for classifying movements from a given set. The proposed method uses velocity and angular information derived from Doppler frequencies and direction-of-arrival (DoA) by processing the signals at the receiver sensor array. Doppler frequency estimation is done by obtaining an initial estimate based on the Fourier transform in conjunction with a predictive tracker. A Root-MUSIC algorithm is used at the estimated Doppler frequencies to obtain DoA corresponding to the dominating moving object. Using speed, direction, and angle as features, a Bayesian classifier is employed to distinguish between a set of movements. The performance of the proposed method is evaluated using an analytical model of arm movements and also using experimental data sets. The proposed ultrasonic Doppler array sensor and processing methods provide a new, compact solution to human arm movement classification.**

***Index Terms*— Ultrasonic array, Doppler and DoA processing, movement classification.**

I. Introduction

HUMAN movement classification is of interest in a number of healthcare and personal well-being applications. In the healthcare domain, monitoring of a patient's recuperation from a Cerebrovascular Accident (CVA) or stroke has been shown to increase the effectiveness of rehabilitative interventions [1], [2]. The study in [3] considered the use of kinematic data - i.e. dynamics of arm movements (e.g. shoulder flexion and horizontal adduction), to characterize motor deficits in CVA patients. Currently, intensive physical and mental rehabilitation to allow patients to re-participate in society is performed by trained therapists in specialized rehabilitation centers. An upcoming approach that aims to resolve the future shortage of rehabilitation spaces and therapists is tele-rehabilitation [4]. With a growing shift from acute to chronic illnesses, rising healthcare costs due to population ageing [5], and the preference of people to live independently, such approaches are especially becoming important. Gaining insights into human movements is an important aspect that determines the efficiency of these remote monitoring solutions [6], [7]. In this paper, we consider an ultrasonic array Doppler sensing solution for classification of movements from a given set.

In [8] and [9], a wearable triaxial accelerometer was used as sensor for monitoring physical activities. In [10], subject-specific electromyography pattern classification techniques were studied to identify their functional tasks and differentiate these from the muscle activation patterns of stroke survivors. A sensor platform solution with a wearable sensor was presented in [11] for elderly health monitoring applications. A detailed review of ambulatory monitors for different clinical applications was done in [12]. It was reported that while wearable ambulatory monitoring sensor solutions are attractive, minimally obtrusive sensors are preferred in a number of home healthcare applications. In [13], methods for markerless pose recovery and human movement classification were presented using cameras from 3-D reconstructed volume data. However, privacy issues relating to the use of vision sensors remain a concern [14]. A radar sensor was used to estimate human motion features using a Boulic model in [15]. Ultrasound as a sensor modality for activity monitoring and classification has been shown to be effective, while being unobtrusive and addressing privacy concerns [16]–[18]. In [19], multiple ultrasonic sensors operating on time-of-flight principle were used to analyze human interactions for potential psychological applications. An acoustic Doppler sonar consisting of a single transmitter and distributed receivers was presented in [17] for single arm gestures. In these works, the receiver setup does not permit the use of beamforming, and as such angular information cannot be extracted.

The configuration considered in this paper is one where the sensor is placed in front of the human. Such a configuration is typical in scenarios where a patient in tele-rehabilitation needs to perform specific movements. The proposed sensor consists of an ultrasonic transmitter and a co-located linear receiver array, thereby allowing extraction of directional information from a correlation analysis of the received signals. Such an

Manuscript received February 5, 2014; revised March 24, 2014; accepted March 27, 2014. Date of publication April 9, 2014; date of current version July 1, 2014. The associate editor coordinating the review of this paper and approving it for publication was Dr. Stefan J. Rupitsch. The work of S. Srinivasan was done while with Philips Research.

R. J. G. van Sloun is with Philips Research, Eindhoven 5656 AE, The Netherlands, and also with the Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven 5612 AZ, The Netherlands (e-mail: ruud.van.sloun@philips.com).

S. Srinivasan was with Philips Research, Eindhoven 5656 AE, The Netherlands. He is now with Microsoft Corporation, Redmond, WA 98052 USA (e-mail: sriram.srinivasan@ieee.org).

A. Pandharipande is with Philips Research, Eindhoven 5656 AE, The Netherlands (e-mail: ashish.p@philips.com).

P. C. W. Sommen is with the Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven 5612 AZ, The Netherlands (e-mail: p.c.w.sommen@tue.nl).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSEN.2014.2315895

ultrasonic Doppler array sensing solution permits extraction of speed, direction and angular information which serve as rich features in movement classification. Furthermore, the array based approach yields a compact solution, in comparison to the distributed systems considered in previous works, and allows enhancement of the signal-to-noise ratio (SNR) for Doppler estimation.

The signals reflected from the human and the environment are processed at the receiver array. We first obtain an initial estimate of the Doppler frequency by processing the short time Fourier transform (STFT) of the signal spatially averaged over the array elements. This Doppler frequency estimate is then tracked over a time segment to obtain an improved estimate. The Doppler frequency in turn provides speed and direction of the human movement. A Root-MUSIC algorithm is then used to obtain angular information. The obtained speed, direction and angle are used as features to describe a movement class. A naive Bayes estimator is employed to classify human movements. For evaluating the proposed techniques, we constrain our attention to single-arm movements. Since rehabilitative interventions often focus on motor control of individual limbs and up to 85% of stroke patients initially show a motor deficit in the arm [20], we focus on arm movements. The performance of the Doppler estimation technique is evaluated using Doppler frequency profiles generated with an analytical model for human arm movement. Finally, the performance of the movement classifier is evaluated by experimental data using a four-element 1D array. The proposed ultrasonic Doppler array sensor and processing methods provide a new, compact solution to the arm movement classification problem, with its effectiveness evaluated with an analytical model and in an experimental setting.

II. System Description

A. Ultrasonic Doppler Sensor

Consider an ultrasonic array consisting of a single transmitter, with center frequency f_t, and co-located linear receiver array with P sensor elements. This choice is determined by the constraint of designing a sensor solution based on commercially available components; an array can be constructed at the receiver side at an operating frequency of $f_t = 40$ kHz, but commercially available transmitters at this frequency have larger component size and a half-wavelength separation is not possible to realize, thereby necessitating a single transmitter. The receiver has a narrow band-pass frequency response with center frequency f_t. The distance between two consecutive sensor elements is $\lambda_t/2$, where $\lambda_t = v_s/f_t$ and v_s is the speed of sound in air.

B. Signal Model

A wave traveling with velocity v_s, when reflecting off an object that is moving at a constant velocity v in the direction of the wave, undergoes a frequency shift due to Doppler effect. The resulting reflected frequency $\hat{f}$ is given by

$$\hat{f} = \frac{v_s + v}{v_s - v} f_t \approx (1 + \frac{2v}{v_s}) f_t, \tag{1}$$

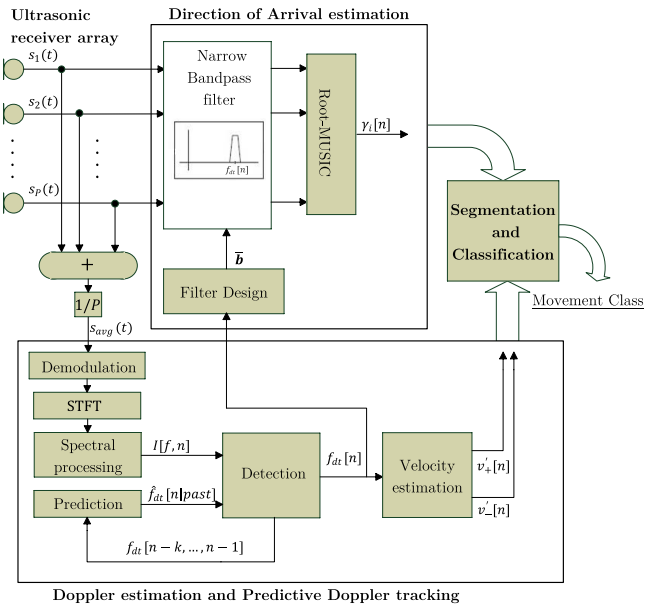

Fig. 1. System level overview of the receiver processing described in Section III.

where the above approximation holds since $v_s \gg v$. The received signal $r(t)$ has a phase shift $\alpha(t)$ given by

$$\alpha(t) = 2\pi \int_{\tau=0}^{t} \hat{f}(\tau) d\tau, \tag{2}$$

and using the approximation in (1), it follows that

$$\alpha(t) = 2\pi f_t [t + \frac{2}{v_s} \int_{\tau=0}^{t} v(\tau) d\tau], \tag{3}$$

from which we arrive at

$$r(t) = \sin(2\pi f_t [t + \frac{2}{v_s} \int_{\tau=0}^{t} v(\tau) d\tau]). \tag{4}$$

Now, the received signal at the p-th sensor element in the array as a result of N reflecting moving objects, where object i has velocity $\overline{v}_i(t)$ and $v_s \gg |\overline{v}_i(t)|$. Following (4), we then have the expression of the received signal at sensor element p as follows

$$s_p(t) = \sum_{i=0}^{N-1} a_{i,p}(t) \sin(2\pi f_t [t + \frac{2}{v_s} \int_{\tau=0}^{t} v'_{i,p}(\tau) d\tau] + \beta_{i,p}(t)) + \Psi(t) + w_p(t). \tag{5}$$

Here, $a_{i,p}(t)$ and $\beta_{i,p}(t)$ are respectively the amplitude and phase of the reflected wave from object i at time t, and $\Psi(t)$ is the contribution due to the reflections from stationary objects. The term $w_p \sim \mathcal{N}(0, \sigma_n^2)$ models the noise using a normal distribution having mean zero and standard deviation σ_n. The Doppler velocity $v'_{i,p}(t)$ is the component of $\overline{v}_i(t)$ in the direction of sensor element p.

III. Receiver Processing

The processing at the receiver array is divided into three modules that are illustrated in Fig. 1 and described in the following sections.

A. Doppler Estimation and Predictive Doppler Tracking

We first consider exploiting the received signals at multiple sensor elements to yield an improved Doppler velocity estimate. For this, consider the average of the P received signals

$$s_{avg}(t) = \frac{1}{P}\sum_{p=0}^{P-1} s_p(t) \tag{6}$$

$$= \sum_{i=0}^{N-1} a_i(t)\sin(2\pi f_t[t + \frac{2}{v_s}\int_{\tau=0}^{t} v_i'(\tau)d\tau] + \beta_i(t)) + \Psi(t) + u(t), \tag{7}$$

where $u(t) = \frac{1}{P}\sum_{p=0}^{P-1} w_p(t)$ and (7) follows under the assumption that $a_{i,p}(t)$, $v'_{i,p}(t)$ and $\beta_{i,p}(t)$ are independent of p. Since Doppler estimation is performed in the time-frequency domain, where the temporal resolution is limited by the desired spectral information, relative time delays due the array spacing are much smaller than the temporal resolution. Furthermore, the radius of propagation is much larger than the size of the array. Therefore, in this domain, the assumption in equation (7) is justified. Recalling that the sum of two independent normally distributed random variables $X \sim \mathcal{N}(\mu_x, \sigma_x^2)$ and $Y \sim \mathcal{N}(\mu_y, \sigma_y^2)$ is given by

$$X + Y = Z \sim \mathcal{N}(\mu_x + \mu_y, \sigma_x^2 + \sigma_y^2), \tag{8}$$

we have that $u \sim \mathcal{N}(0, \sigma_n^2/P)$. The averaging operation keeps the signal power unchanged but reduces the variance of the noise by a factor P, thus improving the signal-to-noise ratio (SNR).

By inspection of equation (7), we observe that $s_{avg}(t)$ is in fact a frequency modulated (FM) version of the transmitted signal $s_t(t) = \sin(2\pi f_t t)$. The FM message of $s_{avg}(t)$ is hence directly related to the object's velocity. We therefore frequency demodulate $s_{avg}(t)$ by applying the following steps. First $s_{avg}(t)$ is differentiated: $y(t) = \frac{d}{dt}s_{avg}(t)$. Then, $y(t)$ is heterodyned to f_{hd} to reduce the required bandwidth but not lose directionality: $y_{hd}(t) = y(t)\sin(2\pi(f_t - f_{hd})t)$. Finally, a low pass filter is applied: $s_{hd}(t) = LPF(y_{hd}(t))$. Although the processing up till here is implemented in the digital domain, it is presented in analog form for the sake of clarity.

The STFT is then calculated by segmenting the data into frames of length T [sec], corresponding to $W = f_s T$ samples, that overlap by 50%. Here, f_s is the sample frequency. This allows the spectrogram to have sufficient temporal resolution to describe the movements, while obtaining enough spectral information. To reduce spectral leakage, a Hamming window is chosen. The short time Fourier transform of $s_{hd}(t)$ in frame n is denoted by $STFT[f, n]$. The discrete form of the power spectral density for every frame is then given by:

$$S[f, n] = \frac{|STFT[f, n]|^2}{W}. \tag{9}$$

We compress the power spectral density by a logarithm:

$$S_{\text{dB}}[f, n] = 10\log_{10} S[f, n] = 20\log_{10}\frac{|STFT[f, n]|}{W}. \tag{10}$$

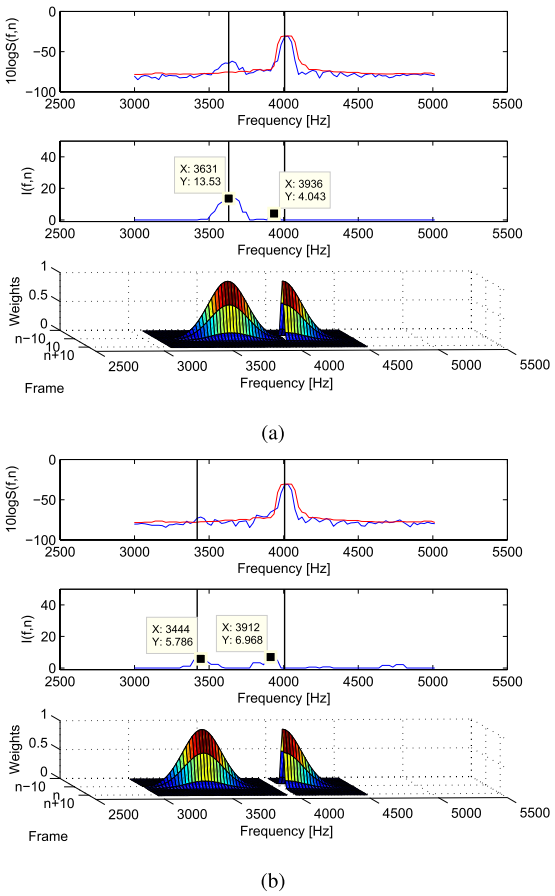

Fig. 2. Doppler tracking for two nearby time frames. In both Fig. 2(a) and (b), the top plot shows the power spectrum in a certain frame (blue) and the upper envelope of the static transmitter spectrum (red), the center plot shows the processed spectrum $I[f, n]$ and the bottom plot shows the weights. In the upper two plots, the tracked Doppler shifts are indicated by a vertical line.

Then, an upper envelope of the static spectrum is determined by applying a rank order filter [21] to the log power spectrum of the initial frames, when there is no movement yet. The result of this operation is shown in the top plot of Fig. 2(a) and Fig. 2(b). Here f_{hd} is 4000 Hz. Because the captured energy from the static reflections is typically high compared to its dynamic counterpart, for each frame the transmitter envelope is subtracted from the log power spectrum, clipped and subsequently processed using another rank order filter:

$$I[f, n] = \zeta_{B,q}(\max(0, S_{\text{dB}}[f, n] - \tilde{S}_{\text{dB}})) \tag{11}$$

where $\tilde{S}_{\text{dB}}$ is the upper envelope of the static spectrum and $\zeta_{B,q}$ denotes the rank order filtering operation having rank q and window (structuring element) B. Here, $q = \lceil 0.9 \times \text{cardinal}(B)\rceil$. The resulting spectrum $I[f, n]$ is shown in the center plot of Fig. 2(a) and Fig. 2(b). A naive Doppler shift estimate in frame n is then one that maximizes $I[f, n]$ over f

$$f_{dn}[n] = \max_f I[f, n]. \tag{12}$$

The naive approach assumes that the desired object always causes the Doppler shift having the highest power.

In real-world body movement applications, this is not necessarily the case. As the reflecting surface of the desired object (e.g. a hand) changes over time, its power at the Doppler frequency also varies. During that movement, another object with a large reflecting surface (e.g. the torso) may also be producing Doppler shifts. These movements are typically slow and hence cause small Doppler shifts, i.e. frequencies close to f_{hd}. Due to surface inconsistency, the power of these small Doppler shifts can sometimes be higher than the power of the desired object however, which causes the naive approach to fail. This effect is exemplified in Fig. 2. By inspecting the center plot of Fig. 2(a), one can easily identify the dominating Doppler shift located at 3631 Hz. Here, maximization of $I[f, n]$ would yield that frequency. However, when analyzing the spectral information obtained in a nearby later frame, shown in Fig. 2(b), the problem becomes evident. Here, maximization of $I[f, n]$ yields a Doppler frequency at 3912 Hz, which would imply that its Doppler shift changed abruptly w.r.t the previous frame. This situation is very unlikely to occur, as we know that human limb velocities and hence their Doppler shifts do not change abruptly during a movement. One would expect the current Doppler frequency to be close to the previous Doppler frequency. Taking into account that $I[f, n]$ shows a prominent local maximum in this region, leads us to reject the frequency obtained using the naive approach. In order to also enable Doppler detection in these situations, we consider a method which uses physical predictability of movements. For the sake of clarity, the method is divided into three steps which are repeated for each frame n. First, a probability density function, reflecting our belief that a trustworthy Doppler frequency is found at a certain frequency $f_{min} \leq f < f_{max}$ and frame $n \leq m < n+M$, is derived based on the Doppler frequencies found in the past. Then the PDF is transformed into a weighting function. This weighting function is applied to the processed spectrum $I[f, m]$ after which the result is maximized over f and m. Finally, the algorithm is designed such that it tracks negative f_{dt-} and positive f_{dt+} Doppler frequencies independently, using two distinct weighting functions and maximizers. This allows velocity estimation of objects moving towards and away from the sensor array at the same time. For tracking of f_{dt+}, f_{min} is set to f_{hd}. For f_{dt-}, f_{max} is set to f_{hd}. The tracker is described in detail in the following.

1) AR Prediction and Probability Modeling: The algorithm tracks the Doppler frequency $f_{dt}[n]$ of each frame by exploiting information about Doppler shifts in the previous frames. Since body movement is continuous and its velocity changes slowly w.r.t. the frame size, the Doppler frequency in the next frame will be close to the shift in the current frame. To incorporate this idea, we assign probabilities to the possible frequencies, thereby reflecting our state of belief. This probability density function (PDF) is modeled as a Gaussian, with standard deviation σ_f and mean value $\hat{f}_{dt}[n|past]$, the predicted Doppler frequency based on the past Doppler shifts. The prediction is performed by employing an autoregressive model of order k:

$$\hat{f}_{dt}[n|past] = f_{dt}[n-1] + \hat{\Delta f}[n|n-k, \ldots, n-1] \quad (13)$$

with

$$\hat{\Delta f}[n|n-k, \ldots, n-1] = \sum_{i=1}^{k} \Delta f[n-i]\chi_i + \epsilon_n \quad (14)$$

where $\Delta f[n-i] = f_{dt}[n-i] - f_{dt}[n-1-i]$ is the Doppler shift change between frames $n-i$ and $n-1-i$, its initial condition $f_{dt}[0] = f_{hd}$, ϵ_n is white noise and $\chi_1, \ldots, \chi_k$ are the model parameters, found using the Yule-Walker equations [22, Chapter 3].

Patients suffering from movement impairments may produce rather inconsistent Doppler shifts. To further increase robustness, we extend the search algorithm to the time domain by modeling the PDF as a truncated bivariate Gaussian, $\mathcal{N}_{trunc}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where

$$\boldsymbol{\mu} = [n, \hat{f}_{dt}[n|past]] \quad (15)$$

is the mean vector and

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_t^2 & 0 \\ 0 & \sigma_f^2 \end{bmatrix} \quad (16)$$

is the covariance matrix. Here, σ_t and σ_f denote the time (in frames) and frequency standard deviations respectively. Truncation assigns a probability of zero to finding the Doppler frequency in time frames $m < n$. This approach allows the Doppler tracker to find the nearest frame at which $I[f, m]$ reveals a trustworthy Doppler shift.

2) Weighting Function: The Gaussian PDF is reflected on the data by the frame-frequency weighting function $W_n[f, m]$, given by

$$W_n[f, m] = \begin{cases} Z_n[f, m], & \text{for } f_{min} \leq f < f_{max} \\ & \text{and } n \leq m < n+M \\ 0, & \text{else} \end{cases} \quad (17)\ (18)$$

where

$$Z_n[f, m] = e^{-\left(\frac{(m-n)^2}{2\sigma_t^2} + \frac{(f-\hat{f}_{dt}[n|past])^2}{2\sigma_f^2}\right)}. \quad (19)$$

The weighting function obtained at two nearby frames n is shown in the bottom plot of Fig. 2(a) and Fig. 2(b).

3) Detection: The Doppler frequency in frame m is then obtained by evaluating equation (20):

$$[f^\star[n], m^\star[n]] = \arg\max_{f,m}(W_n[f, m]I[f, m]). \quad (20)$$

Applying $W_n[f, m]$ incorporates the past information and knowledge of the system into the detection phase, addressing the issues discussed earlier. After finding the Doppler frequency at frame m, the Doppler shift at frame n is found by simple linear interpolation

$$f_{dt}[n] = \frac{f^\star[n] - f_{dt}[n-1]}{m^\star[n] - n + 1}. \quad (21)$$

Note that in Fig. 2(b), using the proposed method the correct Doppler frequency (at 3444 Hz) is detected although $I[f, n]$ is not a global maximum here.

From the tracked Doppler frequencies, the normal velocities v' can be calculated using (1) as

$$v'[n] = \frac{v_s}{2f_t}(f_{dt}[n] - f_{hd}), \quad (22)$$

TABLE I
ALGORITHM PARAMETERS

Parameter	Description	Value
v_s	Speed of sound	340.29 m/s
f_t	Transmitter frequency	40000 Hz
f_{hd}	Heterodyne frequency	4000 Hz
f_s	Sample frequency	96000 Hz
f_{min}	Minimum search frequency	3000 Hz
f_{max}	Maximum search frequency	5000 Hz
T	Frame size	32 ms
B	Rank order window	3 frames
k	AR model order	5
M	Search range time	10 frames
σ_f	Standard deviation frequency	200 Hz
σ_t	Standard deviation time	1 frame

where $f_{dt}[n] - f_{hd}$ is the Doppler shift (given by $\hat{f} - f_t$ in (1)) of the moving object.

B. Direction of Arrival Estimation

Since the radius of propagation is much larger than the size of the array, the far field assumption is valid. In that case, propagation can be described by plane waves. If the angle of incidence is not equal to zero, each array element will receive a slightly delayed version of the signal. Under the assumption that the signal is narrow banded, this delay corresponds to a phase shift, leading to the following array response vector:

$$a(\gamma) = [1, e^{j2\pi \frac{\lambda_t f}{2v_s} \sin\gamma}, \ldots, e^{j2\pi \frac{\lambda_t f}{2v_s}[P-1]\sin\gamma}]^T \quad (23)$$

where γ is the angle of incidence and f is the incoming wave's frequency.

As described in (5), each sensor measures the reflecting ultrasonic signal from both static (e.g. the room walls) and moving objects, collected in $\boldsymbol{s}(t) = [s_1(t), \ldots, s_P(t)]$. As we are just interested in the latter, directly applying direction of arrival estimation to the measured signals is unsuitable, as it would not be able to distinguish between the reflections caused by moving and static objects. In this section we describe a DoA estimation algorithm, which is able to estimate the angular position of a moving object with respect to the array.

Our method exploits the fact that moving objects cause a Doppler shift, which distinguishes them from static objects. For each frame n, a single sided bandpass FIR filter $\overline{\boldsymbol{b}}$ is applied to all $W \times P$ samples of $\boldsymbol{s}(t)$ belonging to n, such that we obtain P narrow-band signals with center frequency equal to the Doppler frequency $f_{dt}[n]$ at that frame. This enables the DoA estimation algorithm Root-MUSIC, described below, to focus only on the moving object. Denote the resulting narrow band-pass filtered signal vector of frame n by $\boldsymbol{h}_n(t)$.

Root-MUSIC [23], a variant of spectral MUSIC [24], replaces the costly iterative search of the latter by a fast polynomial rooting procedure, while having a higher spectral resolution. We first define the array correlation matrix $\boldsymbol{R}_h[n]$ by:

$$\boldsymbol{R}_h[n] = \mathbb{E}[\boldsymbol{h}_n(t)\boldsymbol{h}_n^H(t)]. \quad (24)$$

The algorithm relies on the property of $\boldsymbol{R}_h[n]$ that its eigenspace can be portioned in two orthogonal subspaces, the signal with added noise subspace and the noise only subspace. In our case, there is only one signal and the signal with added noise subspace thus corresponds to the highest eigenvalue. We then compute $C(z)$ by:

$$C(z) = U_n U_n^H \quad (25)$$

where U_n is the noise subspace. The Root-MUSIC algorithm identifies the root of $C(z)$ closest to the unit circle. The angular direction of arrival in degrees is then estimated using

$$\hat{\gamma}[n] = \arcsin(-\frac{v_s}{\pi \lambda_t f_{dt}[n]} \mathrm{Im}[\log(z)]). \quad (26)$$

For each frame, a new DoA is estimated using the above described procedure, thereby tracking the moving object.

TABLE II
CLASS DESCRIPTIONS: RIGHT SHOULDER

Class	Description
1	Horizontal abduction: left to center
2	Horizontal adduction: right to center
3	Extension: top to center
4	Horizontal adduction: center to left
5	Horizontal abduction: center to right
6	Flexion: center to top

C. Segmentation and Classification

Segmentation is based on the velocity feature as described in equation (22), and relies on the assumption that arm activities are separable in multiple movements when its velocity component in the direction of the ultrasound array is zero. The beginning and ending of a segment are defined as the times at which the normal velocity rises above or falls below 10% of the peak normal velocity, respectively. After segmentation, activities are classified based on the average DoA $\tilde{D}[l]$, positive velocity $\tilde{v}_+[l]$ and negative velocity $\tilde{v}_-[l]$ features in each segment l.

Denote the extracted dataset $\mathcal{D}$ by

$$\mathcal{D} = \{(\boldsymbol{g}_1, t_1), \ldots, (\boldsymbol{g}_L, t_L)\} \quad (27)$$

where the inputs $\boldsymbol{g}_l = [\tilde{D}[l], \tilde{v}_+[l], \tilde{v}_-[l]] \in \mathbb{R}^3$ are the features and t_l indicates the class in segment l. We use a 1-of-K coding scheme [25, Ch. 4], such that

$$t_{lk} = \begin{cases} 1, & \text{if } t_l \text{ in class } \mathcal{C}_k \quad (28) \\ 0, & \text{else} \quad (29) \end{cases}$$

is the binary class selection variable, which groups data of the same class. We assume Gaussian class-conditional distributions with a constant covariance matrix,

$$p(\boldsymbol{g}|\mathcal{C}_k, \boldsymbol{\Omega}) = \mathcal{N}(\boldsymbol{g}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_c), \quad (30)$$

and a multinomial prior on the classes, $p(\mathcal{C}_k)$. Since prior class probability differences may not be present, we consider $p(\mathcal{C}_k) = 1/K$, where K denotes the total number of classes.

For building the classifier, we use Maximum Likelihood to estimate the model parameters $\boldsymbol{\Omega} = \{\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_c\}$ from the data $\mathcal{D}$:

$$\hat{\boldsymbol{\Omega}} = \arg\max_{\boldsymbol{\Omega}} \log p(\mathcal{D}|\boldsymbol{\Omega}) \quad (31)$$

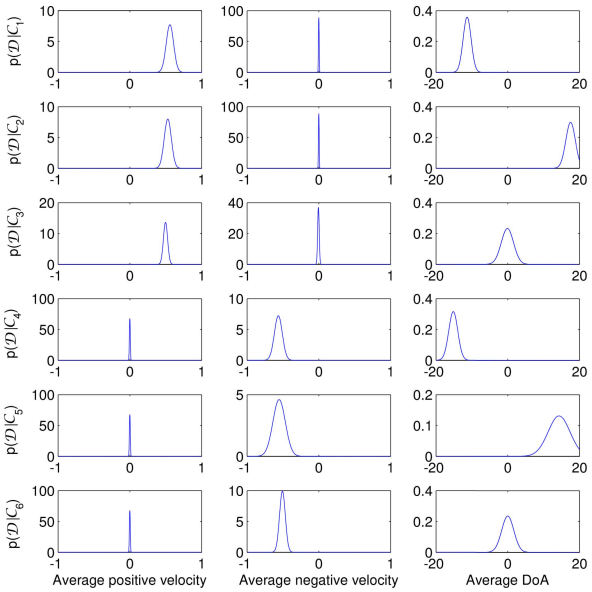

Fig. 3. Estimated probability distributions, obtained by training a naive Bayes model with Gaussian $p(\boldsymbol{g}|\mathcal{C}_k, \boldsymbol{\Omega})$.

where

$$\log p(\mathcal{D}|\boldsymbol{\Omega}) = \sum_{l,k} t_{lk} \log \mathcal{N}(\boldsymbol{g}_l|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_c) \tag{32}$$

We assume the features $\tilde{D}[l]$, $\tilde{v}_+[l]$ and $\tilde{v}_-[l]$ to be mutually independent given $\mathcal{C}_k$, and thus use naive Bayes [26] to estimate the probability distributions. The distributions $p(\mathcal{D}|\mathcal{C}_k)$ obtained for a measured dataset are shown in Fig. 3.

After training, we use the estimated model parameters to find the posterior class probability $p(\mathcal{C}_k|\boldsymbol{g}_{new}, \boldsymbol{\Omega})$. The class of a new segment is then found as

$$k^\star = \arg\max_k p(\mathcal{C}_k|\boldsymbol{g}_{new}, \boldsymbol{\Omega}). \tag{33}$$

D. Reference Method

As reference for Doppler estimation, we use a method for characterizing the frequency response to quantify velocity from the Doppler shifts by considering the 5th and 95th Percentiles as described in [15]. The percentiles are determined by first evaluating

$$P(f,n) = \frac{\sum_{f=f_{min}}^{f} S(f,n)}{\sum_{f=f_{min}}^{f_{max}} S(f,n)} \tag{34}$$

where f_{min} and f_{max} limit the evaluation to the frequency region of interest. The frequencies f for which $P(f,n)$ is equal to 5% and 95%, correspond to the negative and positive Doppler shifts respectively.

IV. Movement Model and Numerical Results

A. Movement Model

In this Section, we consider a simple model for human arm movement to generate Doppler frequency profiles, that serve as ground-truth. The generated profiles will be used to compare

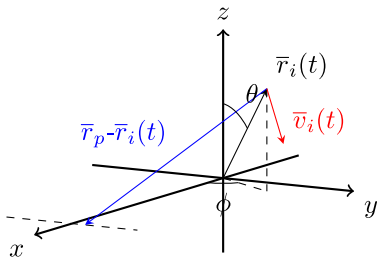

Fig. 4. The velocity vector of point $\overline{r}_i(t)$ is shown in red and the vector $\overline{r}' - \overline{r}$ in the direction of the p-th ultrasonic receiver element is shown in blue.

the proposed Doppler estimation method and the reference method [15]. Arm movements are modeled by a feedback control loop consisting of a plant and a Proportional-Integral-Derivative (PID) controller. The plant dynamics, derived by analyzing a lumped model of the human arm are given by

$$F_\theta(t) - b\dot{\theta}(t) = \kappa\ddot{\theta}(t) \tag{35}$$

$$\ddot{\theta}(t) = -\frac{b}{\kappa}\dot{\theta}(t) + \frac{1}{\kappa}F_\theta(t) \tag{36}$$

and

$$F_\phi(t) - b\dot{\phi}(t) = \kappa\ddot{\phi}(t) \tag{37}$$

$$\ddot{\phi}(t) = -\frac{b}{\kappa}\dot{\phi}(t) + \frac{1}{\kappa}F_\phi(t), \tag{38}$$

where b and κ are the damping and mass of the arm respectively and the inputs $F_\theta(t)$ and $F_\phi(t)$ represent the applied force due to muscle contraction in the $\hat{\theta}$ and $\hat{\phi}$ direction respectively. We incorporate noise in the motor commands as signal-dependent, with standard deviation increasing linearly with the magnitude of the motor command signal (control signal). This noise model has also been used in [27] and [28] and is consistent with empirical findings [29]. The complete plant dynamics in state space representation $\boldsymbol{\Sigma}_H$ can then be written as:

$$\boldsymbol{\Sigma}_H : \begin{cases} \dot{\boldsymbol{x}}(t) = \mathbf{A}\boldsymbol{x}(t) + \boldsymbol{B}(\boldsymbol{I} + \sigma\boldsymbol{\epsilon})\boldsymbol{u}(t) & (39) \\ \boldsymbol{y}(t) = \mathbf{C}\boldsymbol{x}(t) & (40) \end{cases}$$

where the plant's state vector $\boldsymbol{x}(t) = [\theta(t), \dot{\theta}(t), \phi(t), \dot{\phi}(t)]^T$, the input $\boldsymbol{F}(t) = [F_\theta(t), F_\phi]^T$ and the output $\boldsymbol{y}(t) = [\omega_\theta(t), \omega_\phi(t)]^T$. The matrix $\boldsymbol{I}$ represents an identity matrix, $\boldsymbol{\epsilon}$ is zero-mean Gaussian white noise with identity covariance matrix, and σ is the standard deviation. The above state space representation may be considered to be a generalization in two dimensions of the one-dimensional arm movement model considered in [27]. The system matrix is given by

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & -b/\kappa & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -b/\kappa \end{bmatrix}, \tag{41}$$

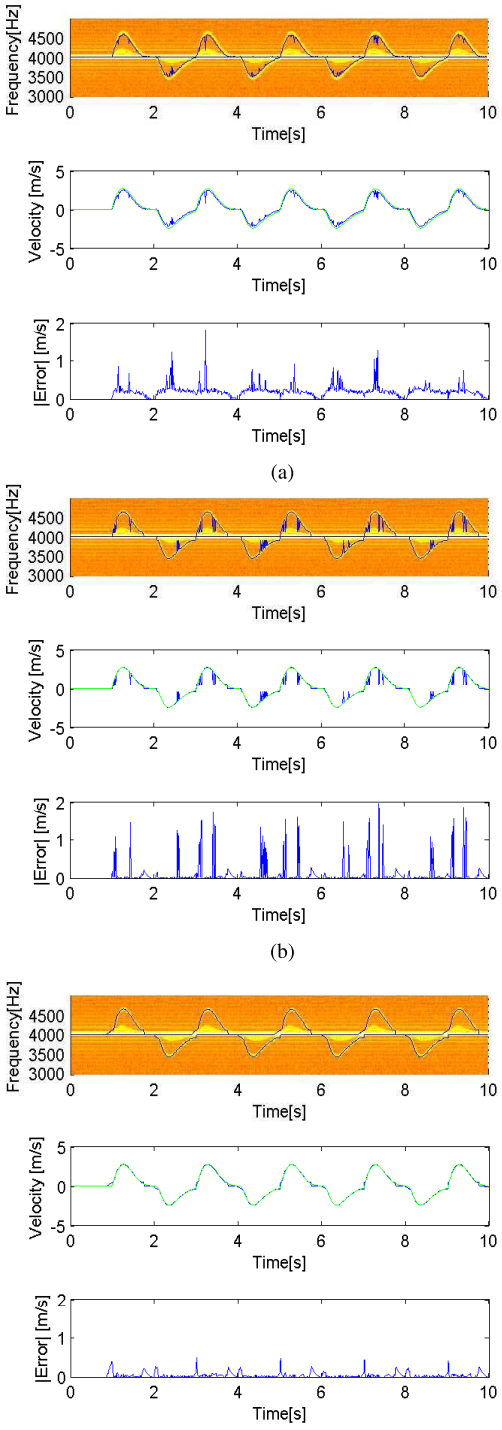

Fig. 5. Comparison of methods for Doppler tracking and velocity estimation, applied to a noiseless dataset generated using the movement model. The top plot shows the spectrogram with the detected Doppler shift, the center plot shows the estimated velocity (blue) compared to the ground truth (red) and the bottom plot shows the absolute error. (a) Reference method [15], equation (34). (b) Naive method, equation (12). (c) Proposed method, equation (21).

the input matrix

$$\boldsymbol{B} = \begin{bmatrix} 0 & 0 \\ 1/\kappa & 0 \\ 0 & 0 \\ 0 & 1/\kappa \end{bmatrix}, \tag{42}$$

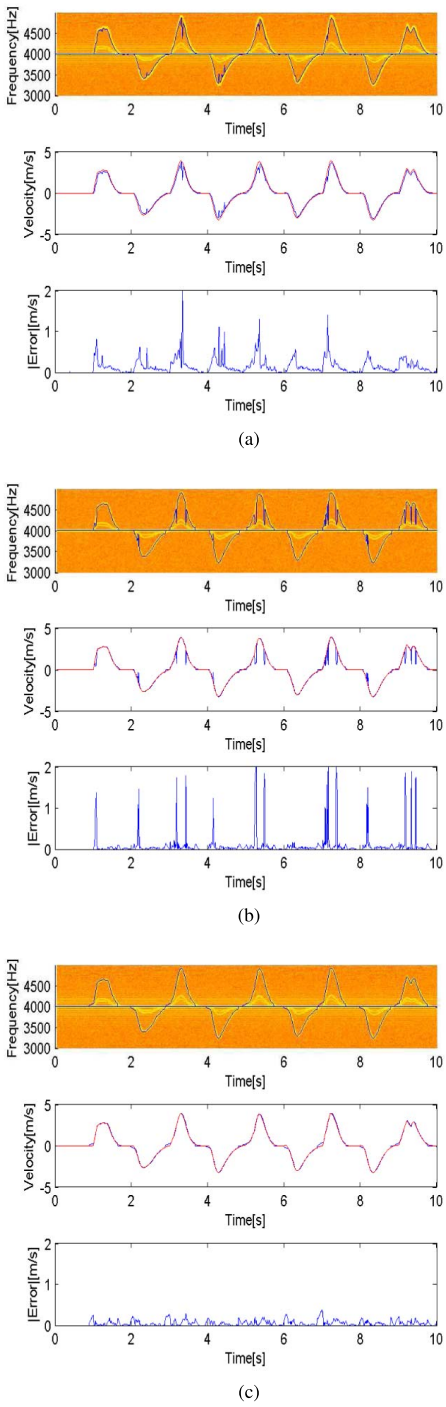

Fig. 6. Comparison of methods for Doppler tracking and velocity estimation, applied to a noisy dataset with $\sigma = 0.3$ generated using the movement model. The top plot shows the spectrogram with the detected Doppler shift, the center plot shows the estimated velocity (blue) compared to the ground truth (red) and the bottom plot shows the absolute error. (a) Reference method [15], equation (34). (b) Naive method, equation (12). (c) Proposed method, equation (21).

and the output matrix

$$\boldsymbol{C} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{43}$$

Although $\boldsymbol{\Sigma}_H$ is a simple approximation of the real, much more complex arm dynamics (see [28] for more sophisticated models), its behavior in the control loop suffices for simulation purposes. The PID controller's input is the position error vector

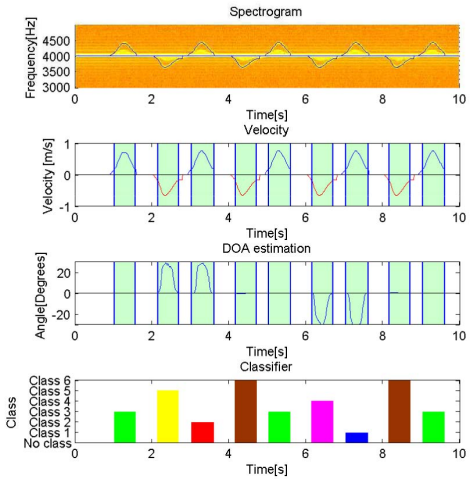

Fig. 7. Results for a dataset generated using the movement model. The top plot shows the spectrogram and the tracked Doppler shifts, the center plots show the velocities and DoA and the bottom plot shows the estimated class, as described in Table II. The detected segments are indicated by bars.

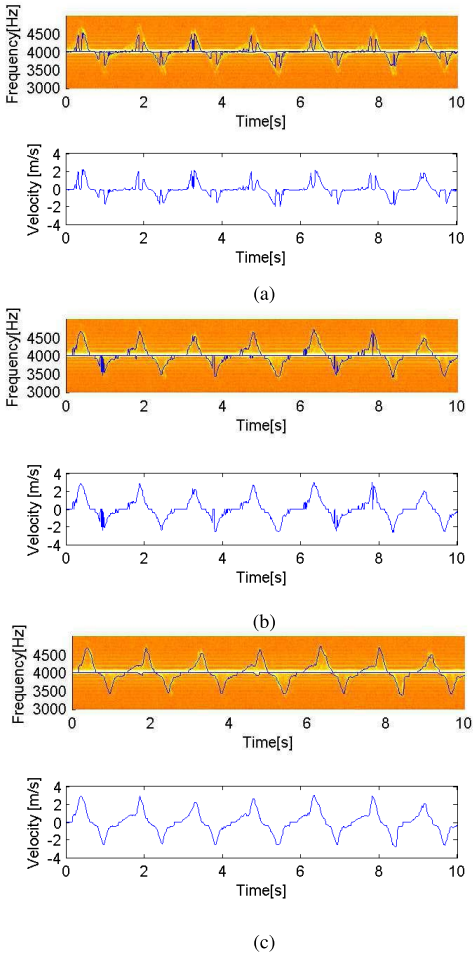

Fig. 8. Comparison of methods for Doppler tracking and velocity estimation, applied to an experimental dataset. The top plots show the spectrogram with the detected Doppler shift and the bottom plots show the estimated velocity. (a) Reference method [15], equation (34). (b) Naive method, equation (12). (c) Proposed method, equation (21).

$\boldsymbol{e}(t) = [\theta(t) - \theta_{des}(t), \phi(t) - \phi_{des}(t)]^T$, where $\theta(t)$ and $\phi(t)$ are found by integration of $\boldsymbol{y}(t)$ and $\theta_{des}(t)$ and $\phi_{des}(t)$ denote the desired position.

Finally, the received sensor array signals are determined by applying equation (5), where the velocity vector $\overline{v}_i(t)$ is denoted by:

$$\overline{v}_i(t) = \dot{r}_i(t)\hat{r} + r_i\omega_\phi(t)\hat{\phi} + r_i\omega_\theta(t)\sin\phi(t)\hat{\theta} \quad (44)$$

$$= r_i\omega_\phi(t)\hat{\phi} + r_i\omega_\theta(t)\sin\phi(t)\hat{\theta}, \quad (45)$$

and the normal vector $\hat{n}_{i,p}(t)$ is given by:

$$\hat{n}_{i,p}(t) = \frac{\overline{r}_p - \overline{r}_i(t)}{|\overline{r}_p - \overline{r}_i(t)|}. \quad (46)$$

Here, $\overline{r}_p$ and $\overline{r}_i(t) = [r_i, \phi(t), \theta(t)]$ are the vectors indicating the receiver's and object's position respectively, as illustrated in Fig. 4.

B. Numerical Results

The model allows us to validate Doppler tracking and velocity sensing since the ground truth velocity is known. We describe the moving arm by three reflecting surfaces at spherical positions $[r_1, \theta(t), \phi(t)]$, $[r_2, \theta(t), \phi(t)]$, and $[r_3, \theta(t), \phi(t)]$. Static objects and sensor noise, with power levels comparable to real measurements, are also added. The mass κ and damping b are set to 3 kg and 20 $[\frac{Ns}{m}]$ respectively. The PID parameters K_p, K_i and K_d are set to 50, 1 and 1 respectively. A comparison between Doppler tracking and velocity estimation methods is shown in Fig. 5 and Fig. 6 for the noiseless case and with $\sigma = 0.3$ respectively. The results are shown when applying:

(a) The reference method described in equation (34) (section III-D).
(b) Naive Doppler frequency estimate, given in equation (12).
(c) Doppler frequency estimate with tracking, given in equation (20).

By inspection of the bottom plots in Fig. 5(a)-(b), one can see that the naive method's error distribution is more spiky compared to the reference method's. Employing a tracker can yield drastic improvement in this particular scenario, as is verified in Fig. 5. Similar observations hold for the plots in Fig. 6.

Fig. 7 shows a section of the complete numerical results for a dataset simulated using the combined movement and signal model. By combining the acquired velocity and DoA features, all classes described in Table II are detected correctly.

V. Measurement Results

Several experimental datasets were acquired, using an ultrasound array consisting of $P = 4$ receivers. The algorithm parameters are given in Table I. For each set, the user stood in front of the horizontally oriented array, moved his arm in a specific pattern and the ultrasonic reflections were captured. Table II shows the six classes of movement and their description. All movements were performed with the user's right arm. A comparison of the reference and proposed method for Doppler and velocity sensing is shown in snapshot results in Fig. 8(a), Fig. 8(b) and Fig. 8(c). Note that the reference method fails to detect the largest velocities due to the surface inconsistency effects described in Section III-A. The proposed

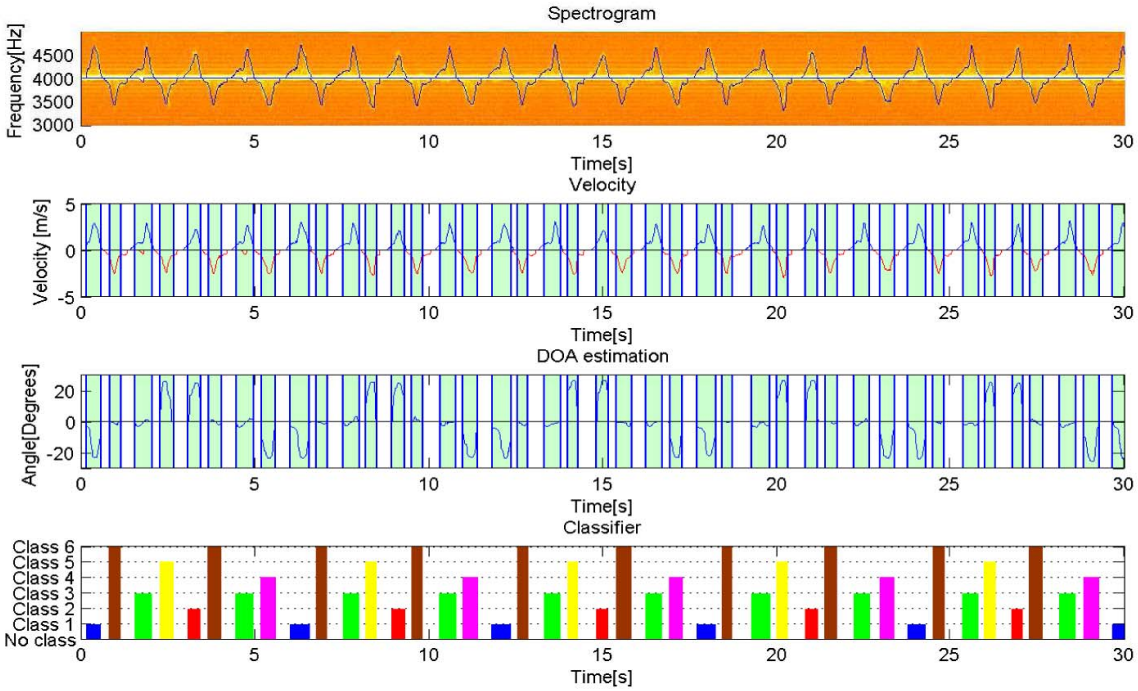

Fig. 9. Results for a section of the experimentally acquired dataset. The top plot shows the spectrogram and the tracked Doppler shifts, the center plots show the velocities and DoA and the bottom plot shows the estimated class, as described in Table II. The detected segments are indicated by bars.

TABLE III
CONFUSION MATRIX SHOWING CLASSIFICATION RESULTS

Class index	(Est.) 1	(Est.) 2	(Est.) 3	(Est.) 4	(Est.) 5	(Est.) 6
(Actual) 1	201	1	7	0	1	0
(Actual) 2	0	210	4	0	1	0
(Actual) 3	3	0	212	0	1	0
(Actual) 4	0	0	0	205	3	6
(Actual) 5	0	0	0	0	209	7
(Actual) 6	0	0	0	0	1	203

Doppler estimate on the other hand is able to better track Doppler, based on the performed arm movement.

The results of the proposed classification algorithm applied to one of the recorded datasets are visualized in Fig. 9. In this case, the following pattern was repeated using the right arm: $\{1, 6, 3, 5, 2, 6, 3, 4\}$. As can be seen from Fig. 9, all six classes are clearly distinguishable using the velocity and DoA features.

To analyze the classification performance, we show the confusion matrix to evaluate correct and incorrect classifications using the entire experimental dataset. We note that the experiments were done by a healthy user emulating tremors and other distortions when making arm movements under different conditions: at normal speed, movements with tremors, and very slow arm movements with tremors.

The overall misclassification probability was 2.7%. We note that the classification errors occur largely due to degraded quality of DoA estimation when arm movements are very slow and ridden with tremors.

VI. CONCLUSION AND DISCUSSION

A compact ultrasonic array sensor, consisting of a continuous wave signal transmitter and a co-located receiver array, and associated signal processing methods were proposed for human movement classification. The proposed method used predictive Doppler tracking and DoA estimation to classify a set of arm movements. By testing the method on both simulated and experimental datasets, we showed that the designed algorithm is able to extract velocity and angular information, allowing accurate classification from a set of arm movements.

Further testing of the proposed method in a clinical context is required to evaluate its use for healthcare monitoring applications. A complete monitoring solution in a tele-rehabilitation context would need movement type classification as well as movement quality estimation. The latter could be obtained by further analyzing the information provided from the features. Also, as suggested in [3] and [30], other than range of motion and peak velocity of the joint movement, movement time and velocity smoothness could be exploited to characterize motor quality and are topics of future investigation.

ACKNOWLEDGEMENT

The authors thank D. Caicedo for the useful comments and discussion on early drafts of the manuscript.

REFERENCES

[1] R. Boian *et al.*, "Virtual reality-based post-stroke hand rehabilitation," *Studies Health Technol. Inf.*, vol. 85, pp. 64–70, Jan. 2002.
[2] L. Eriksson, B. Lindström, and L. Ekenberg, "Patients' experiences of telerehabilitation at home after shoulder joint replacement," *J. Telemed. Telecare*, vol. 17, no. 1, pp. 25–30, 2011.
[3] M. Cirstea, A. Mitnitski, A. Feldman, and M. Levin, "Interjoint coordination dynamics during reaching in stroke," *Experim. Brain Res.*, vol. 151, no. 3, pp. 289–300, 2003.
[4] A. Garcia-Molina *et al.*, "Clinical program of cognitive telerehabilitation for traumatic brain injury," in *Proc. eChallenges*, 2010, pp. 1–10.
[5] J. M. Wiener and J. Tilly, "Population ageing in the United States of America: Implications for public programmes," *Int. J. Epidemiol.*, vol. 31, no. 4, pp. 776–781, 2002.

[6] C. Scanaill, S. Carew, P. Barralon, N. Noury, D. Lyons, and G. Lyons, "A review of approaches to mobility telemonitoring of the elderly in their living environment," *Ann. Biomed. Eng.*, vol. 34, no. 4, pp. 547–563, Apr. 2006.
[7] A. Godfrey, R. Conway, D. Meagher, and G. ÓLaighin, "Direct measurement of human movement by accelerometry," *Med. Eng. Phys.*, vol. 30, no. 10, pp. 1364–1386, 2008.
[8] D. Naranjo-Hernandez, L. Roa, J. Reina-Tosina, and M. Estudillo-Valderrama, "SoM: A smart sensor for human activity monitoring and assisted healthy ageing," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 11, pp. 3177–3184, Nov. 2012.
[9] D. Karantonis, M. Narayanan, M. Mathie, N. Lovell, and B. Celler, "Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring," *IEEE Trans. Inf. Technol. Biomed.*, vol. 10, no. 1, pp. 156–167, Jan. 2006.
[10] S. W. Lee, K. Wilson, B. Lock, and D. Kamper, "Subject-specific myoelectric pattern classification of functional hand movements for stroke survivors," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 19, no. 5, pp. 558–566, Oct. 2011.
[11] S. Abbate, M. Avvenuti, and J. Light, "MIMS: A minimally invasive monitoring sensor platform," *IEEE Sensors J.*, vol. 12, no. 3, pp. 677–684, Mar. 2012.
[12] T. Shany, S. J. Redmond, M. R. Narayanan, and N. H. Lovell, "Sensors-based wearable systems for monitoring of human movement and falls," *IEEE Sensors J.*, vol. 12, no. 3, pp. 658–670, Mar. 2012.
[13] J. Gu, X. Ding, S. Wang, and Y. Wu, "Action and gait recognition from recovered 3-D human joints," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 40, no. 4, pp. 1021–1033, Aug. 2010.
[14] K. E. Caine, "Visual sensing devices in home-care systems," in *Proc. 1st ACM Workshop Security Privacy Med. Home-Care Syst.*, 2009, pp. 61–62.
[15] P. van Dorp and F. C. A. Groen, "Feature-based human motion parameter estimation with radar," *IET Radar, Sonar Navigat.*, vol. 2, no. 2, pp. 135–145, Apr. 2008.
[16] B. Raj, K. Kalgaonkar, C. Harrison, and P. Dietz, "Ultrasonic Doppler sensing in HCI," *IEEE Pervas. Comput.*, vol. 11, no. 2, pp. 24–29, Feb. 2012.
[17] K. Kalgaonkar and B. Raj, "One-handed gesture recognition using ultrasonic Doppler sonar," in *Proc. IEEE ICASSP*, Apr. 2009, pp. 1889–1892.
[18] Y. Tsutsui, Y. Sakata, T. Tanaka, S. Kaneko, and M. Feng, "Human joint movement recognition by using ultrasound echo based on test feature classifier," in *Proc. IEEE Sensors*, Oct. 2007, pp. 1205–1208.
[19] R. Cheng, W. Heinzelman, M. Sturge-Apple, and Z. Ignjatovic, "A motion-tracking ultrasonic sensor array for behavioral monitoring," *IEEE Sensors J.*, vol. 12, no. 3, pp. 707–712, Mar. 2012.
[20] V. M. Pomeroy, C. A. Clark, J. S. G. Miller, J.-C. Baron, H. S. Markus, and R. C. Tallis, "The potential for utilizing the 'mirror neurone system' to enhance recovery of the severely affected upper limb early after stroke: A review and hypothesis," *Neurorehabil. Neural Repair*, vol. 19, no. 1, pp. 4–13, 2005.
[21] P. Soille, "On morphological operators based on rank filters," *Pattern Recognit.*, vol. 35, no. 2, pp. 527–535, 2002.
[22] P. Stoica and R. L. Moses, *Introduction to Spectral Analysis*, vol. 1. Upper Saddle River, NJ, USA: Prentice-Hall, 1997.
[23] A. Barabell, "Improving the resolution performance of eigenstructure-based direction-finding algorithms," in *Proc. IEEE ICASSP*, vol. 8. Apr. 1983, pp. 336–339.
[24] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, Mar. 1986.
[25] C. M. Bishop and N. M. Nasrabadi, *Pattern Recognition and Machine Learning*, vol. 1. New York, NY, USA: Springer-Verlag, 2006.
[26] D. Lowd and P. Domingos, "Naive Bayes models for probability estimation," in *Proc. 22nd Int. Conf. Mach. Learn.*, 2005, pp. 529–536.
[27] D. M. Wolpert, Z. Ghahramani, and M. I. Jordan, "An internal model for sensorimotor integration," *Science*, vol. 269, no. 5332, pp. 1880–1882, 1995.
[28] W. Li, "Optimal control for biological movement systems," Ph.D. dissertation, Dept. Comput. Sci., Univ. California, San Diego, CA, USA, 2006.
[29] D. E. Meyer, R. A. Abrams, S. Kornblum, C. E. Wright, and J. E. K. Smith, "Optimality in human motor performance: Ideal control of rapid aimed movements," *Psychol. Rev.*, vol. 95, no. 3, pp. 340–370, 1988.
[30] O. Rohrer *et al.*, "Movement smoothness changes during stroke recovery," *J. Neurosci.*, vol. 22, no. 18, pp. 8297–8304, 2002.

Ruud J. G. van Sloun was born in Roermond, The Netherlands, in 1990. He received the B.Sc. degree in electrical engineering from the Eindhoven University of Technology, Eindhoven, The Netherlands, in 2012, where he is currently pursuing the M.Sc. degree with the Biomedical Diagnostics Laboratory, Signal Processing Group.

During the summer period of 2013, he did an internship at Philips Research Eindhoven, where he was involved in arm movement classification using ultrasound. In 2013, he rejoined Philips Research for his master's thesis, specializing in the field of nonlinear ultrasound imaging. His current research interests include clinical applications of signal and image processing, sensor analytics, medical ultrasound, mathematics, and physics.

Sriram Srinivasan (M'06–SM'12) received the Ph.D. degree from the School of Electrical Engineering, KTH Royal Institute of Technology, Stockholm, Sweden, in 2005, and the master's degrees in computer science and mathematics from the Sri Sathya Sai Institute of Higher Learning, Prasanthinilayam, India, in 2001 and 1999, respectively. From 2006 to 2013, he was with Philips Research Laboratories, Eindhoven, The Netherlands, as a Senior Scientist, where he was involved in acoustical signal processing algorithms. He is currently a Senior Program Manager with Microsoft Corporation, Redmond, WA, USA, where he is involved in audio signal processing for real-time communication applications.

Ashish Pandharipande (SM'08) received the B.E. degree in electronics and communications engineering from Osmania University, Hyderabad, India, the M.S. degrees in electrical and computer engineering and mathematics, and the Ph.D. degree in electrical and computer engineering from the University of Iowa, Iowa City, in 1998, 2000, 2001, and 2002, respectively. In 2003, he was a Post-Doctoral Researcher with the University of Florida. From 2004 to 2006, he was with the Samsung Advanced Institute of Technology as a Senior Researcher. He has held visiting positions with AT&T Laboratories, Department of Electrical Communication Engineering, Indian Institute of Science, India, and Nanyang Technological University, Singapore.

Since 2006, he has been a Senior Scientist with Philips Research, Eindhoven, The Netherlands. Dr. Pandharipande is currently an Associate Editor of the IEEE SENSORS JOURNAL and the IEEE TRANSACTIONS ON SIGNAL PROCESSING, an Editor of the *EURASIP Journal on Wireless Communications and Networking*, and a member of the International Advisory Board, Lighting Research and Technology. His research interests are in the areas of cognitive wireless networks, sensor signal processing, control theory, and the application of these to lighting systems and energy management systems.

Piet C. W. Sommen received the Ingenieur degree in electrical engineering from the Delft University of Technology and the Ph.D. degree from the Eindhoven University of Technology in 1981 and 1992, respectively. From 1981 to 1989, he was with Philips Research Laboratories, Eindhoven, and has been with the Faculty of Electrical Engineering, Eindhoven University of Technology, since 1989, where he is currently an Associate Professor.

He is involved in different internal and external courses, all dealing with different basic and advanced signal processing topics. His main field of research is in adaptive array signal processing and new sampling techniques.

Dr. Sommen has been a Research Dean of the Faculty Board (1993–1995), a ProRISC Board Member (1993–2002), the Vice President of the IEEE Benelux Signal Processing Chapter (1998–2002), an Officer of the Administrative Board of the European Association for Signal Processing (European Association for Signal Processing) (1998–2003), a EURASIP Newsletter Editor (1998–2003), an Editor of the *Journal of Applied Signal Processing* (2002–2004), an Editor of a Special Issue on Signal Processing for Acoustic Communications (2003), a reviewer of the MEDEA+ project (2002–2004), and a Co-Chair of the International Workshop on Acoustic Echo and Noise Control (2005).