Hand Gesture Recognition Using Ultrasonic Waves

Mohammed H. AlSharif, Mohamed Saad, and Tareq Y. Al-Naffouri
King Abdullah University of Science and Technology (KAUST), Saudi Arabia

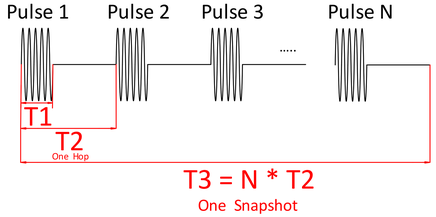
1. Motivation

- With nowadays spread usage of electronic devices and the advances in human-computer interface technology, gesture control is becoming an active research area.
- This technology has applications in consumer electronics, medical care, advertisement and many other applications.
- An attractive passive hand gesture recognition approach is based on using Ultrasonic waves.
- This approach insures user privacy, insensitive to illumination changes, and has lower power consumption and computational complexity compared to approaches based on using cameras.
- This work presents a hand gesture recognition system that can detect five different hand gestures with high accuracy using a single ultrasonic transmitter and a single ultrasonic receiver which makes the system suitable for devices like laptops and mobile phones.

2. System Parameters and Signal Design

- The system aims to detect five types of gestures, shown below

- The range of movement for the gesturing hand is 10 to 50 cm in depth and 40 cm horizontally with a speed not exceeding 1 m/s.
- The transmitted signal is a train of pulses designed to satisfy the requirements of the system



- The transmitter and receiver are separated by a distance of 1.1 cm, so in addition to the reflected signals, a self-interference signal is received directly from the transmitter.
- T_1, T_2, and T_3 are designed to reduce the overlap between reflections from the hand and the self interference signal (T_1 parameter), reduce the effect of multi-path reverberations (T_2 parameter) and improve SNR (T_3 parameter).
- Odd indexed pulses are up-Chirps and even indexed pulses are down-Chirps which provides orthogonality between consecutive pulses.

3. Signal Processing Techniques

- Processing the reflected signals aim is to estimate the range and RSS.
- The reflected signal from hand can be modeled as:

$$y[n] = \alpha x[n-d] + w[n] \tag{1}$$

Where α is attenuation factor, $x[n]$ is the transmitted signal, d is TOF from the transmitter to the hand then back to receiver, and $w[n]$ is AWGN.

- Cross-correlation is applied between the transmitted and received signal to estimate TOF (d) and RSS (α).
- De-cluttering is used to remove the self-interference and other unwanted reflections from static and slowly moving objects where the peaks associated with these objects will always appear at the same delay (TOF) over all cross-correlation frames.
- Subtracting the previous cross-correlation frames from the current one will remove them, and if we denote the cross correlation vector v_i where i refers to the index of the processed block, and let the cluttering factor be c, then the output of the de-clutter is given by:

$$\mathbf{v}_{\hat{i}} = \mathbf{v}_i - (c^i \mathbf{v}_1 + c^{i-1}(1-c)\mathbf{v}_2 + \ldots + c^0(1-c)\mathbf{v}_{i-1}) \tag{2}$$

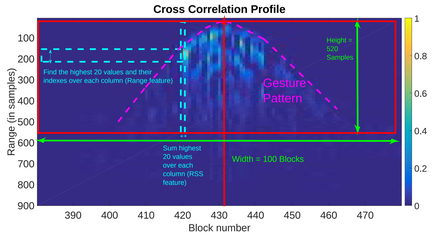
Where $c \in [0, 1]$.

- Cross correlating and de-cluttering produce Motion Frames in which peaks positions relates to the current range of gesturing hand and their amplitude indicates RSS.

4. Gesture Detection and Classification

- Support Vector Machine (SVM) classifier was chosen to detect and classify a predefined set of five gestures.
- Motion Frames obtained from consecutive blocks over time form a Motion Profile.
- On average a gesture occupies 2 seconds in time, generating a Motion Profile of 100 frames.
- Each gesture type has its unique Motion Profile.
- Two main features sets were engineered from the Motion Profile (100 frames length) to be used for detection and classification. The first set is RSS Vector, formed by summing the values of the highest 20 peaks over the profile length in each frame. The second set is Range Matrix, consisting of positions and values of the highest 20 peaks in each frame.

Cross Correlation Profile



Detection Classification Hierarchy

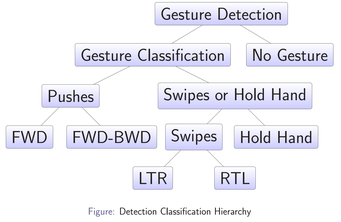


Figure: Detection Classification Hierarchy

Detection and classification is done in a hierarchy way where the first two levels in the hierarchy are classified using RSS vector, and the last two levels are classified using Range Matrix.

5. Experimental Evaluation

- A customized board with a mic and MEMS ultrasonic transmitter is connected to a PC through a sound-card.
- During experiments, data is recorded and saved to be processed off-line using MATLAB.
- A set of around 120 repetitions of each type of the five gestures was recorded.
- Test was done using cross validation where 8% of the data was taken as a test set and the rest was taken as a training set.
- The following confusion matrix shows the performance of the system

Gesture	Right-Left	Left-Right	Hold Hand	Fwd-Bwd	Fwd
Right-Left	**0.9423**	0.0385	0	0	0
Left-Right	0.0088	**0.9381**	0.0885	0	0
Hold Hand	0.0309	0.2165	**0.8247**	0	0.0103
Fwd-Bwd	0.0085	0.0169	0.1356	**0.8136**	0.1017
Fwd	0	0.0545	0.0545	0.0273	**0.9182**

6. Conclusions

- A single ultrasonic transmitter and a single receiver can be used to detect and classify five different types of hand gestures with high accuracy.
- The design of the transmitted signal uses LFM utilizing a small bandwidth.
- Least squares SVM classifier with a polynomial kernel function was used to detect and classify gestures.
- System evaluation shows an average accuracy of 88.7%.
- Since the proposed system is using a single ultrasonic transmitter and a single receiver, it can be suitable for devices like laptops and mobile phones.

References

[1] Rautaray, Siddharth S., and Anupam Agrawal. *Vision based hand gesture recognition for human computer interaction: a survey.*, Artificial Intelligence Review 43.1 (2015): 1-54.

[2] Gupta, S., Morris, D., Patel, S. and Tan, D., *Soundwave: using the doppler effect to sense gestures.*, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, 2012.

[3] Kalgaonkar, Kaustubh, and Bhiksha Raj, *One-handed gesture recognition using ultrasonic Doppler sonar.*, Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on. IEEE, 2009.

Emails: mohammed.alsharif, mohamed.saadeldin,tareq.alnaffouri@kaust.edu.sa