VPad: Virtual Writing Tablet for Laptops Leveraging Acoustic Signals

Presenter: Xiangyu Xu

Li Lu[1], Jian Liu[2], Jiadi Yu[1], Yingying Chen[2], Yanmin Zhu[1], Xiangyu Xu[1], Minglu Li[1]
Shanghai Jiao Tong University[1]
Rutgers, The State University of New Jersey[2]

Emerging Need for Touch Capability in Laptops

- More widely deployed touch screens
  - 97% smart devices are equipped with touch screens
  - The trend spreads into laptops
  - Most traditional laptops are incapable with touch screens
- Growing applications requiring touch screens
  - E.g., WRITEit, Drawboard PDF, etc.
  - Requiring writing & drawing function, which cannot be implemented o small touchpads on laptops
- Special situations requiring touch capability
  - Difficulty on inputting through keyboard
  - E.g.,
    - input on vibrated vehicles
    - user-friendly input for disabled persons

Related Works

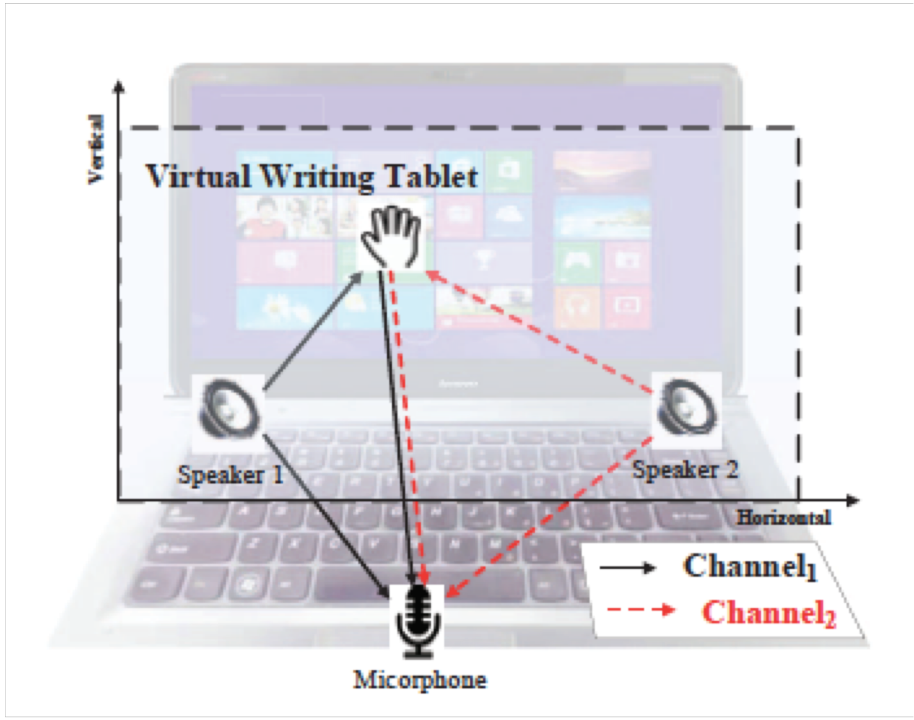
- Industrial products
  - E.g., Microsoft Kinect, Nintendo Wii
  - Vision-based solution: easy to be interfered by ambient factors, especially lights
- Research works
  - AAMouse & CAT: accurate tracking, but require additional devices
  - LLAP & FingerIO & Strata: wearable-oriented solution, but cannot adopted in laptops due to different audio components (i.e., only one microphone exists in laptops)
- Our work
  - Take one step forward to develop a device-free virtual writing tablet leveraging existing audio devices on traditional laptops without additional hardware

Outline

- Feasibility Study
- System Design
- Evaluation
- Conclusion

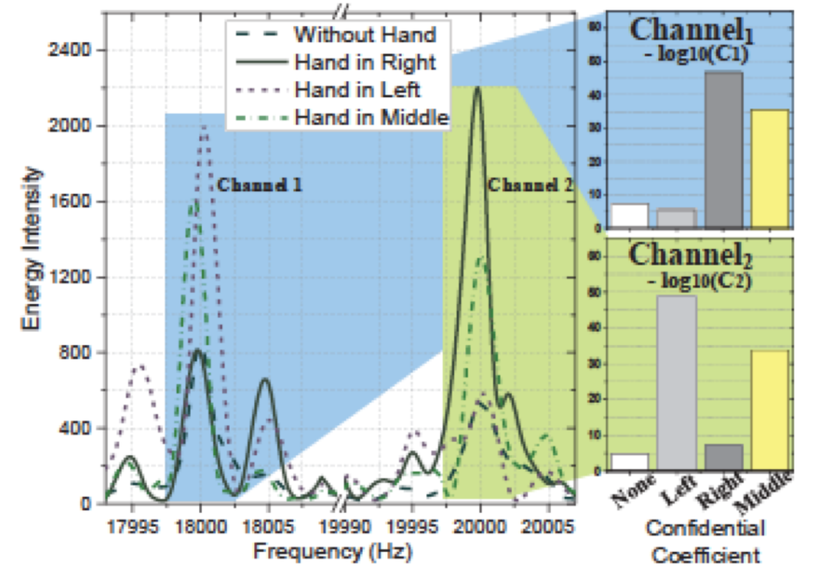
Movement Tracking

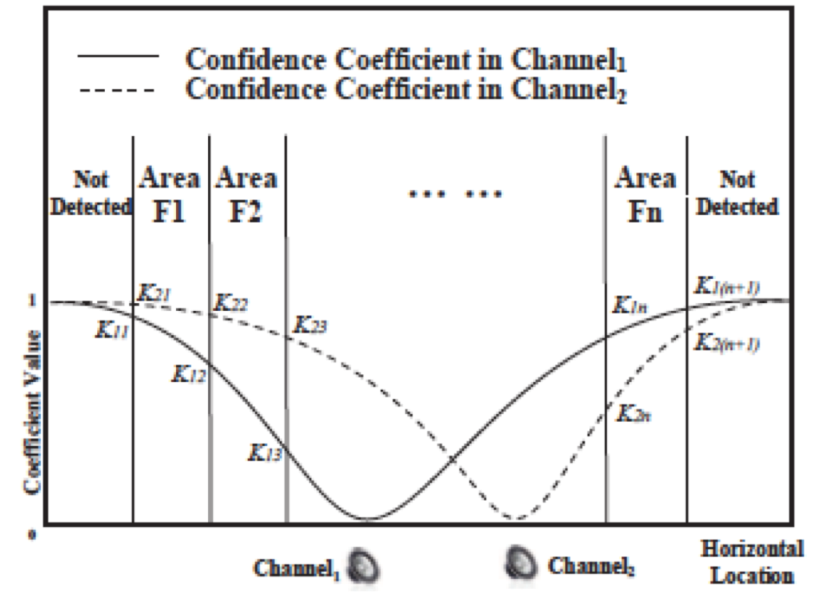
- Acoustic-based tracking
  - Tracking horizontal movement based on energy feature
  - Tracking vertical movement based on Doppler effect
- Basic audio devices in a laptop
  - Two speakers – 2 TX
    - Transmit acoustic signals under different frequencies
    - Serve as two channels to avoid interference
  - One microphone – 1RX
    - Receive acoustic signals at the same time



Horizontal Movement Tracking

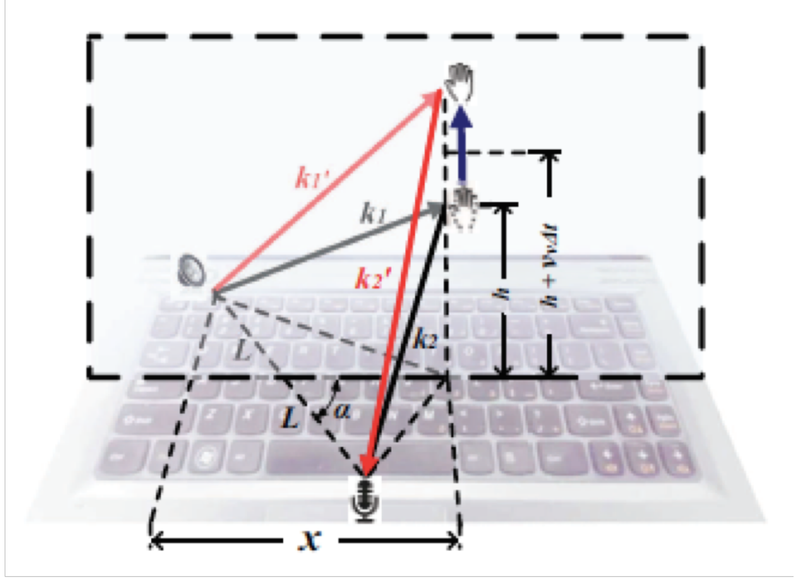
- Energy feature in each channel
  - Confidential coefficient $c = \int_{-\infty}^{t=-|t_0|} P(n,t) + \int_{t=|t_0|}^{\infty} P(n,t)$, where P(n, t) is the possibility distribution
  - When a hand is in the channel, the value of relative confidential coefficient becomes larger
- Energy feature
  - $<c_1, c_2>$ is unique under different position of a hand
  - Further divide the horizontal space into areas
- Track horizontal movement
  - During a tie period of Δt, a hand moves from area a to b
    - Area a and b is determined through matching $< c_1, c_2 >$
  - Moving velocity $v_h = \frac{(b-a)l_u}{\Delta t}$, where l_u is the length of an area

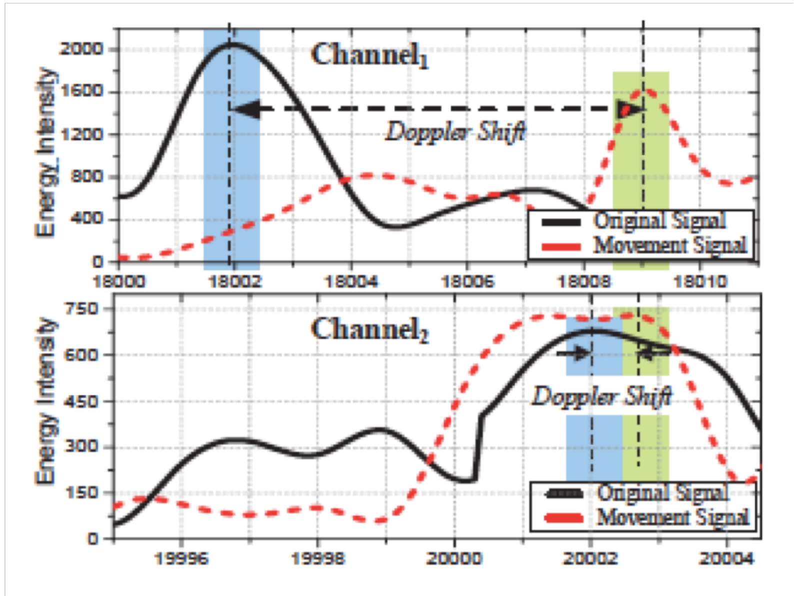




Vertical Movement Tracking

- Failure of energy feature in tracking vertical movement
  - Along with connect line between microphone & speaker
  - Insignificant energy feature changes in different vertical areas
- Doppler effect
  - Basic principle: $\Delta f = v f_0 / v_0$, where f_0 and v_0 are the frequency and speed of signal
- Tracking vertical movement
  - A hand movement may induce two kinds of Doppler shifts, i.e., horizontal & vertical
  - From received signal, obtain the whole Doppler shift Δf
  - By energy feature, obtain the horizontal Doppler shifts Δf_h
  - Based on $\Delta f = \sqrt{\Delta f_h^2 + \Delta f_v^2}$, obtain the vertical Doppler shift Δf_v
  - By Doppler effect, the vertical velocity v_v is obtained



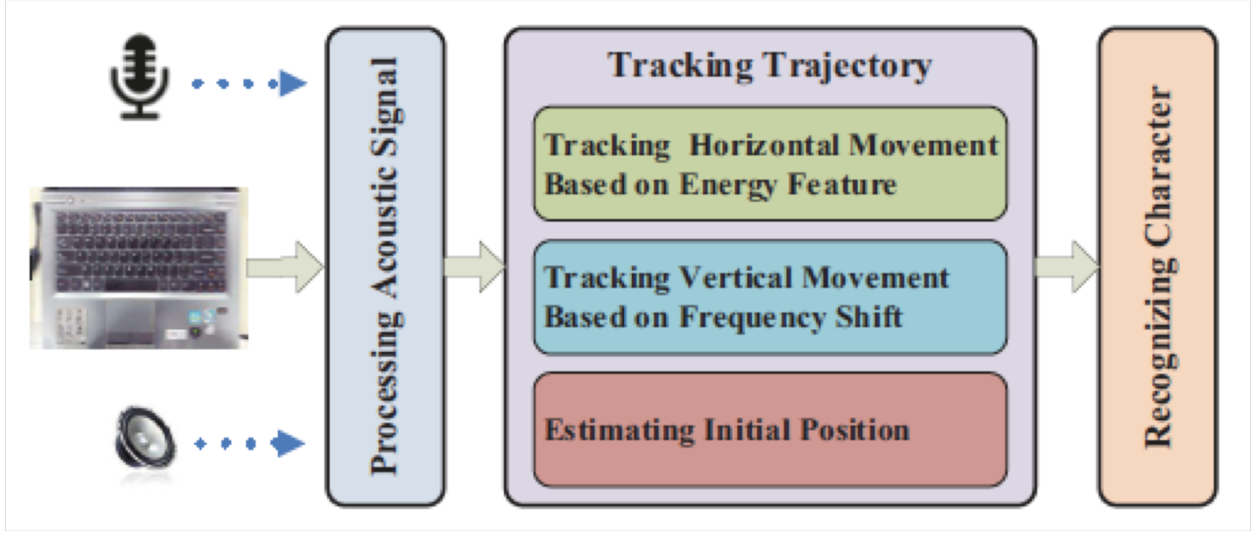


Outline

- Feasibility Study
- System Design
- Evaluation
- Conclusion

Overview

- Audio Device Setup
  - 18kHz & 20kHz acoustic signals transmitted by two speakers respectively
  - 44.1kHz sampling rate of microphone
- Three Processes:
  - Processing Acoustic Signal
  - Tracking Trajectory
  - Recognizing Character

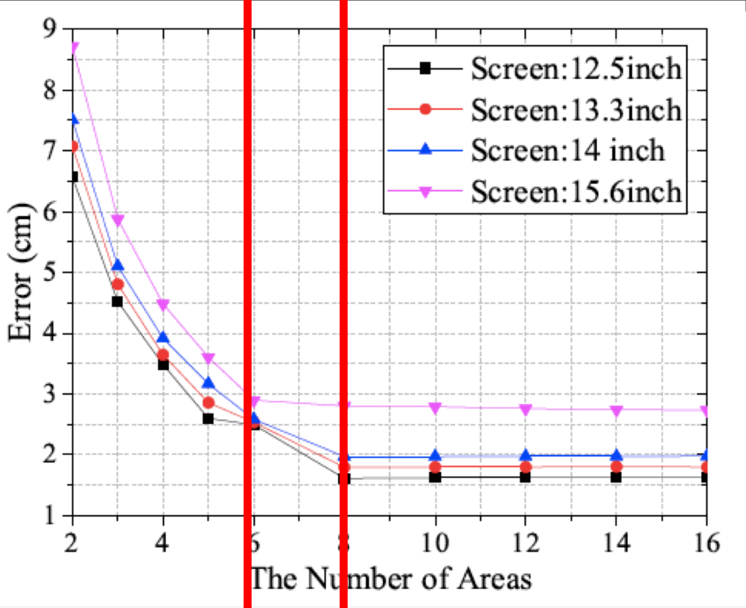


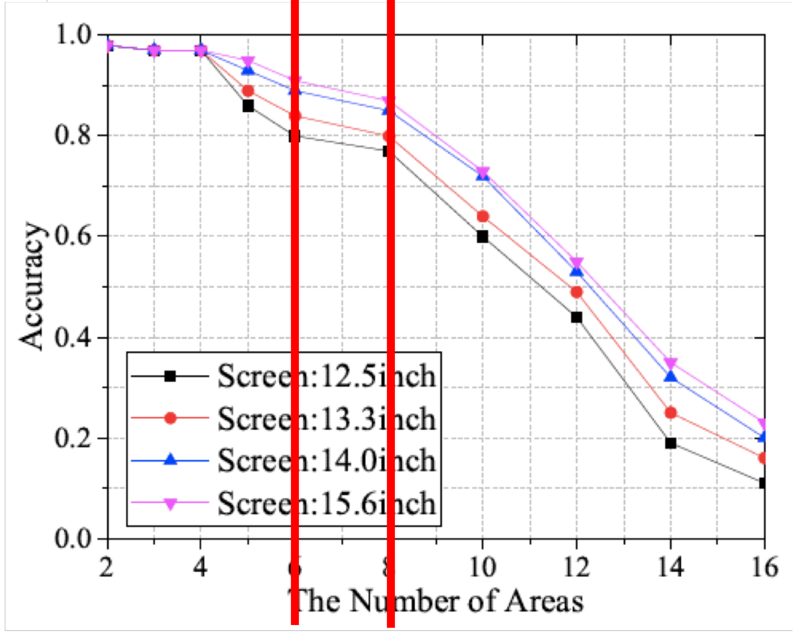
Processing Acoustic Signal

- Time-domain → Frequency Domain (FFT)
  - Goal: achieve accurate tracking → 1Hz frequency resolution
  - Problem: far from real-time tracking → 0.9s time interval
- Sliding-window overlap Fourier transformation
  - a sliding window whose length is 0.9s with step of 0.1s
  - 1Hz frequency resolution & 0.1s time resolution
  - Achieve both accurate & real-time tracking
- Some other problems
  - Fense Effect: require sampling points to be 2^n
    - Add redundant zero elements
  - Frequency Leakage Distortion:
    - Nonrectangular window (Hamming window in our implementation)

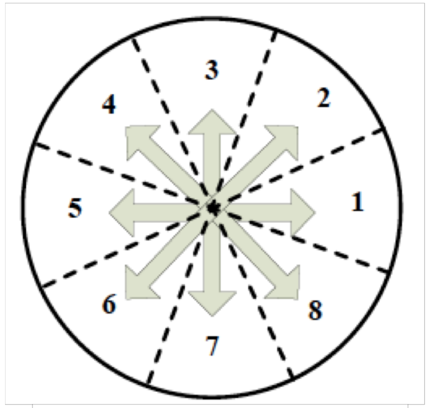
Tracking Trajectory

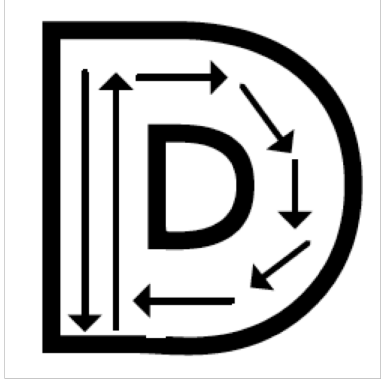
- Tracking Horizontal Movements
  - Energy feature-based approach
  - Key problem: how many horizontal areas should we divide?
  - By handling the tradeoff between errors and accuracy, 6-8 areas are appropriate
- Tracking Vertical Movements
  - Doppler effect-based approach
- Estimating Initial Position
  - Horizontal initial position is determined by energy feature
  - Vertical initial position is determined by TDoA
- Tracking trajectory
  - Combine with the three sub-processes

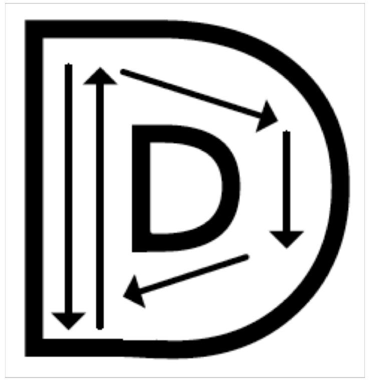




Recognizing Character







- Stroke direction sequence
  - Each character writing can be divided into a sequence of strokes
  - Each stroke belongs to a specific direction
  - The directions are categorized into 8 different ones in our design
- Potential direction sequences
  - Take writing D as example
    - Both [7,3,1,8,7,5,4] and [7,3,8,7,5] are valid for writing D
    - Construct a database containing all potential sequences
- Similarity comparing
  - Weighted Minimum Edit Distance (WMED)
    - If a stroke direction n_0 is substituted by another one n_1, the weight is

$$w = \begin{cases} |n_0 - n_1| & 1 \leq |n_0 - n_1| \leq 4 \\ 8 - |n_0 - n_1| & 5 \leq |n_0 - n_1| \leq 7 \end{cases}$$

    - Finally, choose k nearest ones as candidates

Outline

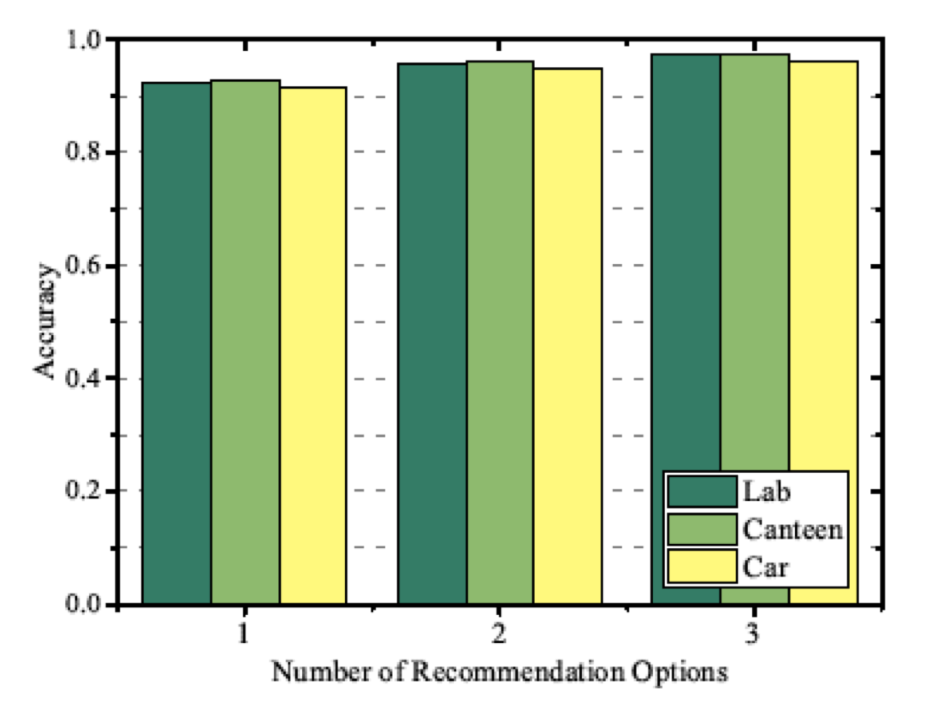
- Feasibility Study
- System Design
- Evaluation
- Conclusion

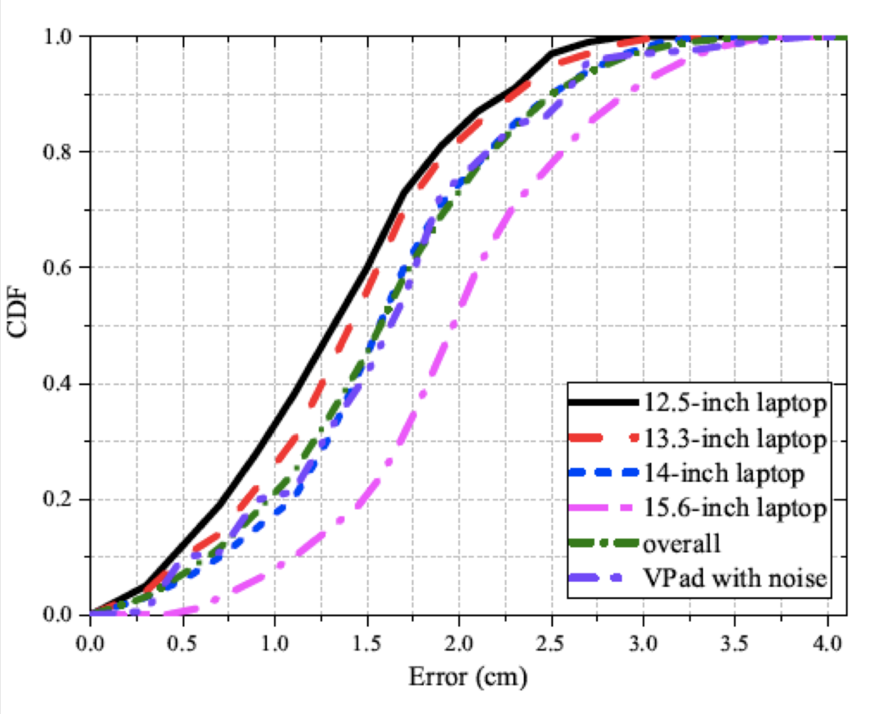
Experiment Setup

- 20 volunteers in 3 real environments respectively
  - Volunteers: 10 males and 10 females, whose ages range in [20,65]
  - Environments: lab (static and quiet),
    canteen (static but noisy),
    moving car (dynamic and noisy).
- 70 characters:
  - 26 capital letters (‘A’-’Z’)
  - 26 lower letters (‘a’-’z’)
  - 10 numbers (’0’-’9’)
  - 8 special characters (**Δ**,**Γ**,**Ω**,**Π**,**Σ**,Λ,V,∠)

Overall Performance

- Approach 95% accuracy in character recognition
- Achieve less than 1.6cm error in trajectory tracking



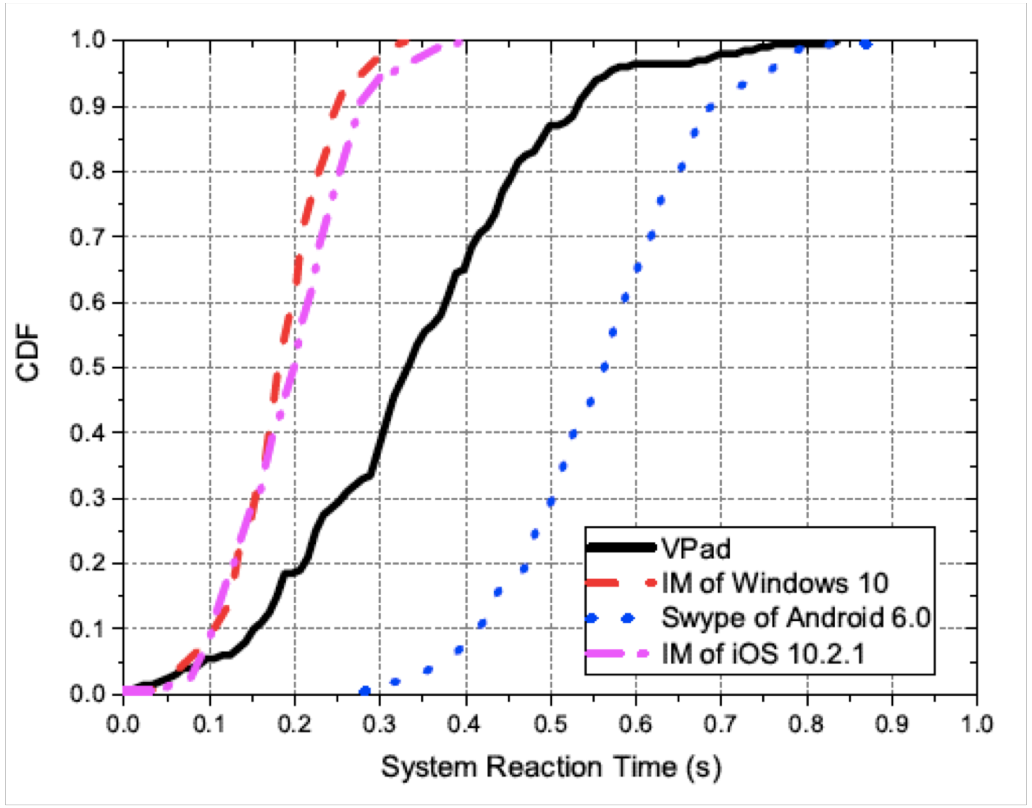


System Reaction Time

- Compare with other three common input methods (IMs) in touch screens
  - Default IM of Windows 10
  - Swype of Android 6.0
  - Default IM of iOS 10.2.1
- Reaction Time:
  - VPad: 0.34s
  - Windows 10: 0.19s
  - Swype: 0.56s
  - iOS: 0.21s

For third-party IMs, VPad responses more in real time

Lower time, due to better integration in OS



Outline

- Feasibility Study
- System Design
- Evaluation
- Conclusion

Conclusion

- Feasibility:
  - Tracking horizontal movements using energy features of acoustic signals
  - Tracking vertical movements using Doppler shifts of acoustic signals
- Contribution:
  - Propose VPad to enable traditional laptops with writing capability
  - Utilize both frequency shift and energy feature of acoustic signals to track hand movements
  - Employ a stroke direction sequence mode to recognize exact characters
- Evaluation: evaluate performances of VPad in three real environments
  - Approach 95% accuracy in character recognition
  - Achieve less than 1.6cm error in trajectory tracking

Thank you!

Q & A