International Journal of Scientific Research in Science, Engineering and Technology
Print ISSN: 2395-1990 | Online ISSN : 2394-4099 (www.ijsrset.com)
doi : https://doi.org/10.32628/IJSRSET

AIR-Writing Word Recognition

Sushruthi Sheri*[1], Palle karthik Reddy*[2], Mr. Narsappa*[3]
[1]B.Tech Student,[2]B. Tech Student,*[3]Professor
ECM Department, JB Institute of Engineering & Technology, Hyderabad, Telangana, India

ARTICLEINFO

Article History:
Accepted: 01 April 2023
Published: 10 April 2023

Publication Issue
Volume 10, Issue 2
March-April-2023

Page Number
369-374

ABSTRACT

Air-writing refers to writing of linguistic characters or words in a free space by hand or finger movements. Airwriting differs from conventional handwriting; the latter contains the pen-up-pen-down motion, while the former lacks such a delimited sequence of writing events. We address air-writing recognition problems in a pair of companion papers. In Part 1, recognition of characters or words is accomplished based on 6 degrees-of-freedom hand motion data. We address air-writing on two levels: motion characters and motion words. Isolated air-writing characters can be recognized similar to motion gestures although with increased sophistication and variability. For motion word recognition in which letters are connected and superimposed in the same virtual box in space, we build statistical models for words by concatenating clustered ligature models and individual letter models. Hidden Markov model is used for airwriting modeling and recognition.

We show that motion data along dimensions beyond a 2D trajectory can be beneficially discriminative for air-writing recog-nition. We investigate the relative effectiveness of various feature dimensions of optical and inertial tracking signals, and report the attainable recognition performance correspondingly. The proposed system achieves a word error rate of 0.8% for word-based recognition and 1.9% for letter-based recognition. We also subjectively and objectively evaluate the effectiveness of airwriting and compare it to text input using a virtual keyboard. The words-per-minute of airwriting and virtual keyboard are 5.43 and 8.42, respectively.

Keywords : Handwriting Recognition, Air-Writing, 6-DOF Motion, Usability Study

I. INTRODUCTION

MOTION gestures provide a complimentary modality for general human-computer interaction. Motion gestures are meant to be simple so that a user can easily memorize and perform them. However, motion gestures themselves are not expressive enough to input text for motion-based control.

We define "air-writing" as writing letters or words with hand or finger movements in a free space. Air-writing is especially useful for user interfaces that do not allow the user to type on a keyboard or write on a trackpad/touchscreen, or for text input for smart system control, among many applications. Isolated letters written in the air involves a sequence of hand or finger movements. Although any snapshot of such move-ments can be considered a realization of a motion gesture, air-writing is more complicated than gesture recognition because of the interdependency among the involved "gestures". In conventional handwriting, a sequential discrete stroke structure is made. A stroke is an isolated writing trajectory between the pen-up/pen-down events. In contrast, air-writing is rendered on a virtual plane without visual or haptic feedback, and lacks the delimited sequence of writing events. Air-writing is also more complex for automatic recognition than cursive style writing on paper due to the lack of a concrete anchoring or reference position; the person who performs air-writing can only use an imaginary coordinate to guide the writing motion. The variability of motion data that represents a letter is thus considerably broader in air-writing than in paper writing.

From a user's perspective, air-writing can be realized in several ways. The first and the most essential is writing of individual isolated letters in an imaginary box in the space, one at a time. The second is the writing of multiple letters across the space from left to right in a style much like writing on a paper. Finally, one can also write several letters, stacked contiguously one over another in the same imaginary box. We call these, isolated, connected, and overlapped air-writing respectively.

The problem of air-writing recognition can be approached progressively. Isolated air-writing carries the assumption that the hand motion to render a letter has already been roughly localized in time and in space. Localization of motion render-ing may be accomplished by use of a tracker, which can be easily turned on or off, to signify the beginning and ending of a writing activity. The localization is only approximate and not fluctuation-free because most users cannot precisely synchronize the tracker control (on-off) and the true writing trajectory. This is similar to the notorious problem of end-pointing in spoken utterance recognition even with a push-to-talk control.

Between the approximate endpoints, the motion trajectory forms a letter that resembles a uni-stroke writing. Study of isolated air-writing is essential to provide the technological foundation for subsequent challenges. Beyond isolated let-ters, recognition of "word" poses two additional challenges: the contiguous writing of letters without segmentation, and the incorporation of sequential constraints between letters. The distinction between connected and overlapped air-writing mainly arises from system usability; the latter requires less limb movement. From the viewpoint of technology develop-ment, techniques for overlapped air-writing can be applied to connected-letter air-writing and we shall address overlapped air-writing with emphasis.

Demonstration of the masked output

The paper is organized as follows. In the next section, we discuss the related prior work. Section III describes the motion tracking system and data acquisition procedures for air-writing. In Section IV, we explain the feature extraction and normalization procedure, and present the techniques for modeling motion characters and motion words. The experiment setup and results are given in Section V. We present the usability study in Section VI. The discussion and conlusion are in Section VII.

II. LITERATURE AND SURVEY

Traditional handwriting styles include cursive or print let-ters. These writing styles vary with writers and are often mixed in actual handwriting. To make it easier for a machine to recognize and quicker for a user to write, letters are simplified into single-stroke styles. The Graffiti alphabet [4] best exemplifies the uni-stroke handwriting. Because the uni-stroke alphabet differs from conventional writing, a novice user needs to learn and practice the writing system to attain entry speed.

There are other text input modalities in addition to typing and writing.One alternative approach is a mixture of typing and writing. In Quickwriting [5], [6], a user swipes strokes on zones and sub-zones to input the associated characters. TwoStick [7] applies a similar concept to the two joysticks on a gamepad. Swype [8] allows a user to enter words on a soft keyboard by sliding from the first letter of a word to its last letter and uses a language model to guess the intended word. Similar to typing on a virtual keyboard, swiping strokes also requires the user's attention to the visual feedback while inputing text and is not eyes-free.

A review of automated handwriting recognition can be found in [9]. Hidden Markov models (HMMs) are widely used for online handwriting recognition [10], [11]. In [12], ligature models are proposed to address online recognition of cursive handwriting, in which successive letters are connected without explicit pen-up moves. Motion-based handwriting can also be considered in parallel to motion gestures or sign language. Motion gesture recognition has been studied with different types of motion tracking devices [3], [13]. Sign language is more sophisticated than motion gestures. Many sign language recognition systems use HMMs with various sensing technologies, such as data gloves and vision-based techniques [14], [15]. In [16], [17], air-writing recognition was achieved with inertial sensors attached to a glove. Jin et al. [18] proposed a vision-based approach for finger-writing character recognition. Schick et al. [19] also proposed a vision-based hand tracking system that recognizes handwriting in mid-air. In [20], finger writing in the air is tracked with a depth sensor. Different motion sensing and tracking technologies impose various behavioral load on the user. As an example, wearing data gloves is often considered by many users as an undesirable burden and may change the wearing user's motion behavior. In our earlier work [21], we opt for a hybrid tracking system (and the accompanying device) that is simple to control.

3. Recurrent Neural Network (or RNNs)

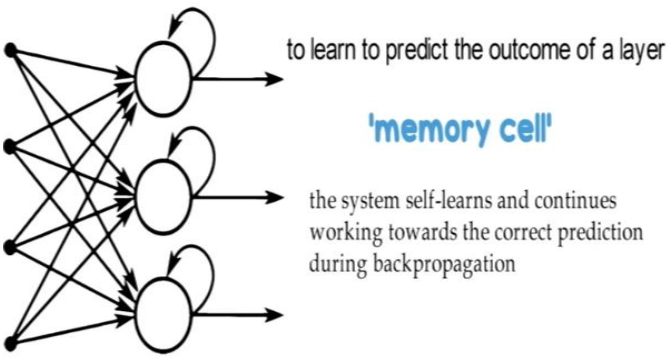

TABLE IX. USABILITY RESULT OF AIR-WRITING AND VIRTUAL KEYBOARD (SUBJECTIVE RATING FROM 1 TO 5)

Question	air-handwriting	virtual keyboard
1. Intuitiveness [5: most intuitive]	4.10	4.75
2. Arm fatigue level [5: no fatigue]	3.05	3.10
3. Vote for inputing a short word (2-3 letters)	16	4
4. Vote for inputing a long word (4+ letters)	11	9
5. Satisfaction of recognition performance [5: most satisfied]	4.25	-

Because relatively large and unconstrained control motions are involved. Our study indicates the speed for these alternative text input methods on a motion-based user interface.

The objective metrics show that air-writing is roughly 1.5 times slower and 3 times longer in motion footprint than the virtual keyboard. However, we get quite interesting results from the subjective evaluation as shown in Table IX. Air-writing is a variation of conventional writing, and virtual key-board follows the same metaphor of typing on a touchscreen. Both methods are intuitive to users and have neutral scores for the arm fatigue level. Motions in the air involve more muscles than keyboard or touch-based interaction and thus cause more fatigue. Even though the motion footprint of air-writing is three times larger, it does not directly reflects arm fatigue ratings. The arm fatigue level relates to the writing or typing style. For example, air-writing could cause less fatigue for a user who rests the elbow and writes with the upper arm and wrist than a user who holds the whole arm in the air. The layout of the virtual keyboard is fixed for all subjects. To cover all keys, it requires a larger range of movement, e.g., the distance between key Z and Backspace is about 60 cm (1200 pixels). Six subjects mention that the keyboard layout is too big. Reducing the size of the keyboard layout can reduce the motion footprint. However, smaller keys can be prone to "typing" errors and require more precise pointing motions. The majority of users choose air-writing for short text input (2-3 letters), and about half of users prefer air-writing for long text input (4+ letters).

Based on our study, air-writing may not be fast enough for general-purpose text input, but it is suitable for infrequent and short text input on a motion-based user interface, where conventional writing or typing is not available. Although virtual keyboard is faster than air-writing, a virtual keyboard requires a display and precise pointing. Typing on a virtual keyboard requires two foci of attention (FOA), i.e., the user needs to pay attention to the keyboard and then the input result. On the contrary, air-writing is a single-FOA task. The user does not necessarily need the visual feedback of writing and achieves "eyes-free" text input. Air-writing recognition does not require precise pointing and is applicable to a broader range of motion tracking systems.

There are other usability issues of air-writing from user feedback. The box-writing style appears to be easy to learn, but it needs some practice to write with the specified stroke order. In our current system, writing with different stroke orders can cause errors in recognition, especially for shorter words. Five users suggest to write without constraints on the stroke order, and four users would like to write without holding a button.

III. CONCLUSION

In this work, we attempt to recognize air-writing with a 6-DOF motion tracking system. The writing motion is tracked with the position and orientation in the global frame, and the acceleration and angular speed in the device-wise coordinates. The air-writing recording process is very time consuming. To make the recording process feasible, we place constraints on stroke orders and upper-case letters with limited vocabulary to refine the scope of air-writing data acquisition without losing too much generality. From these motion data, we derive five basic features for observations of HMMs and form the combination of pure optical, pure inertial, and complete 6-DOF features. Although the handwriting is defined purely by the planar shape, we show that motion information beyond the spatial trajectory is informative for air-writing recognition.

Air-writing is uni-stroke without pen-up/pen-down infor-mation. The writing style and motor control are different from ordinary pen-based writing due to lack of haptic and vision feedback. We separate air-writing in two levels: motion characters and motion words. Motion characters are handled similar to motion gestures, and each character is modeled with a HMM. A motion word can be modeled by concatenating character and ligature models. We present two approaches to model ligatures: hard clustering and decision tree. The former is proven to be sufficient for word-based word recognition. The latter provides better capability of ligature modeling, which improves the performance of letter-based word recognition. The word-based word recognition achieves relatively low WER but is not able to recognize out-of-vocabulary words. The word-based recognizer is suitable for applications that have a limited vocabulary and stringent requirement on the accuracy. On the other hand, letter-based word recognition has around 10% WER but can handle arbitrary letter sequences and progressive decoding. To substantially improve the letter-based recognition accuracy, the system can provide suggestions with n-best decoding and lets the user choose the right one.

A user study investigates input speed, motion footprint, physical strain, and subjective evaluation of two motion-based text input methods: air-writing and virtual keyboard. The results suggest that air-writing is suitable for short and infrequent text input on a motion-based user interface.

IV. REFERENCES

[1]. M. Chen, G. AlRegib, and B.-H. Juang, "Air-writing recognition, part 2: Detection and recognition of writing activity in continuous stream of motion data," IEEE Trans. HumanMach. Syst, in press.

[2]. "Leap motion," http://www.leapmotion.com, 2015.

[3]. M. Chen, G. AlRegib, and B.-H. Juang, "Feature processing and modeling for 6d motion gesture recognition," IEEE Trans. Multimedia, vol. 15, no. 3, 2013, pp. 561–571.

[4]. C. H. Blickenstorfer, "Graffiti: Wow!" Pen Computing Magazine, vol. 1, 1995, pp. 30–31.

[5]. K. Perlin, "Quikwriting: continuous stylus-based text entry," in Proc. of the 11th annual ACM symposium on User interface software and technology, 1998, pp. 215–216.

[6]. P. Isokoski and R. Raisamo, "Quikwriting as a multi-device text entry method," in Proc. of the third Nordic conference on Human-computer interaction, 2004, pp. 105–108.

[7]. T. Koltringer,¨ P. Isokoski, and T. Grechenig, "Twostick: writing with a game controller," in Proc. of Graphics Interface, 2007, pp. 103–110.

[8]. "Swype — type fast, swype faster," http://www.swype.com/, 2013.

[9]. R. Plamondon and S. Srihari, "Online and off-line handwriting recogni-tion: a comprehensive survey," IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 1, 2000, pp. 63 –84.

[10]. J. Makhoul, T. Starner, R. Schwartz, and G. Chou, "On-line cursive handwriting recognition using hidden markov models and statistical grammars," in Proc. of the workshop on Human Language Technology, 1994, pp. 432–436.

[11]. J. Hu, M. Brown, and W. Turin, "Hmm based online handwriting recognition," IEEE Trans. Pattern Anal. Mach. Intell., vol. 18, no. 10, 1996, pp. 1039–1045.

[12]. B. K. Sin and J. H. Kim, "Ligature modeling for online cursive script recognition," IEEE Trans. Pattern Anal. Mach. Intell., vol. 19, no. 6, 1997, pp. 623–633.

[13]. J. Mantyj¨arvi,¨ J. Kela, P. Korpipa¨a,¨ and S. Kallio, "Enabling fast and effortless

customisation in accelerometer based gesture interaction," in Proc. of the 3rd Int'l Conf. on Mobile and ubiquitous multimedia, 2004, pp. 25–31.

[14]. Z. Zafrulla, H. Brashear, T. Starner, H. Hamilton, and P. Presti, "American sign language recognition with the kinect," in Proc. of the 13th intl. conf. on multimodal interfaces, 2011, pp. 279–286.

[15]. N. Pugeault and R. Bowden, "Spelling it out: Real-time asl finger-spelling recognition," in Computer Vision Workshops (ICCV Work-shops), 2011 IEEE International Conference on. IEEE, 2011, pp. 1114–1119.

[16]. C. Amma, D. Gehrig, and T. Schultz, "Airwriting recognition using wearable motion sensors," in Proc. of the 1st Augmented Human Intl. Conf., 2010, pp. 10:1–10:8.

[17]. C. Amma, M. Georgi, and T. Schultz, "Airwriting: Hands-free mobile text input by spotting and continuous recognition of 3d-space handwrit-ing with inertial sensors," in Proc. of the 16th International Symposium on Wearable Computers, 2012, pp. 52–59.

[18]. L. Jin, D. Yang, L.-X. Zhen, and J.-C. Huang, "A novel vision based finger-writing character recognition system," in Pattern Recognition, 2006. ICPR 2006. 18th International Conference on, vol. 1, 2006, pp. 1104–1107.

[19]. A. Schick, D. Morlock, C. Amma, T. Schultz, and R. Stiefelhagen, "Vision-based handwriting recognition for unrestricted text input in mid-air," in Proc. of the 14th ACM intl. conf. on Multimodal interaction, 2012, pp. 217–220.

[20]. . Zhang, Z. Ye, L. Jin, Z. Feng, and S. Xu, "A new writing experience: Finger writing in the air using a kinect sensor," MultiMedia, IEEE, vol. 20, no. 4, 2013, pp. 85–93.

[21]. M. Chen, G. AlRegib, and B.-H. Juang, "6DMG: A new 6D motion gesture database," in Proc. of the third annual ACM conf. on Multimedia systems, 2012.

[22]. "Worldviz ppt-x," http://www.worldviz.com, 2015.

[23]. P. O. Kristensson, T. Nicholson, and A. Quigley, "Continuous recogni-tion of one-handed and two-handed gestures using 3d full-body motion tracking sensors," in Proc. of the ACM international conference on Intelligent User Interfaces, 2012, pp. 89–92.

[24]. T. Brants and A. Franz, "Web 1T 5-gram version 1," Linguistic Data Consortium, Philadelphia, 2006.

[25]. L. Rabiner and B.-H. Juang, Fundamentals of Speech Recognition. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993.

[26]. L. E. Baum, T. Petrie, G. Soules, and N. Weiss, "A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains," The Annals of Mathematical Statistics, vol. 41, no. 1, 1970, pp. 164–171.

[27]. S. Katz, "Estimation of probabilities from sparse data for the language model component of a speech recognizer," IEEE Trans. Acoustics, Speech and Signal Processing, vol. 35, no. 3, 1987, pp. 400–401.

[28]. C.-M. Karat, C. Halverson, D. Horn, and J. Karat, "Patterns of entry and correction in large vocabulary continuous speech recognition systems," in Proc. of the SIGCHI conference on Human Factors in Computing Systems, 1999, pp. 568–575.

Cite this article as :

Sushruthi Sheri, Palle Karthik Reddy, Mr. Narsappa, "AIR-Writing Word Recognition", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Online ISSN : 2394-4099, Print ISSN : 2395-1990, Volume 10 Issue 2, pp. 369-374, March-April 2023.

Journal URL : https://ijsrset.com/IJSRSET2310254