ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

Recognition and Modeling of Characters, Words and Connecting Motions

Deepa. D, R. Dharmalingam

M.E. Student, Dept. of Applied Electronics, Anna University, Maharaja Institute of Technology, Coimbatore, India

Assistant Professor & Head, Dept. of ECE, Maharaja Institute of Technology, Coimbatore, India

ABSTRACT: We address air-writing recognition 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.Using different tracking technologies, the motion can be tracked explicitly with the position and orientation or implicitly with the acceleration and angular speed. In this work, we address the problem of motion gesture recognition for commandand control applications.Recognition of characters or words is accomplished based on six-degree-of freedom hand motion data. **.** We investigate the relative effectiveness of various feature dimensions of optical and inertial tracking signals and report the attainable recognition performance correspondingly.We also subjectively and objectively evaluate the effectiveness of air-writing and compare it with text input using a virtual keyboard.

KEYWORDS: Air-writing, handwriting recognition, usability study, 6-DOF motion, recognizing

I. INTRODUCTION

Detection of recognizing on the fly activities in a continuous stream of motion data without delimitation of the intent of movements is a fundamental problem different from the recognition problem. Itdeals with the ultimate problem of detecting and recognizing writing rendered by finger motion in the air, without any demarcation of the starting or ending points of the strokes, letters, and the letter sequences.Air-writing is more complicated than gesture recognition because of the interdependence among the involved “gestures.”

Localization of motion rendering 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.Recognition of characters 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.

Study of isolated air-writing is essential to provide the technological foundation for subsequent challenges.we adapt the six-degree-of-freedom (DOF) motion gesture recognition work for character recognition and incorporate context information to achieve recognition beyond the letter or character level. This extension includes the introduction of a modeling hierarchy, consisting of letter-based motion trajectory models and ligature models, to cope with the unique contiguous writing style of air-writing. Second, we conduct a usability study to demonstrate that air-writing is a preferred text input method on a motion-based user interface. We evaluate the text input performance of the proposed system and the use of a virtual keyboard with both subjective and objective metrics. we explain the feature extraction and normalization procedure and present the techniques for modeling motion characters and motion words.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

II.LITERATURE SURVEY

The computer mouse is one of the most common input devices on computers. However, there are many variations to the standard computer mouse. Computer mice can have a single button or many buttons along with a scroll wheel. Wireless mice are available. Mice also may have trackballs or use optical light to sense movement. However, while computer mice are because of several advantages, there are still a few disadvantages to computer mice as well.Computer mice need an unobstructed and flat surface to effectively monitor and manage user movements. However, flat surfaces may not always be available, especially as computer users become more mobile with laptop computers, and this may cause limited use of the mouse.

The proposed system converts air into mouse thus we can avoid the mentioned problems in existing system. In front of computer monitor an array of sensors arranged to detect the touch on thin layer of air. The output of sensors will be enough to get the location of touch and the position values will be send to the computer system and controlling mouse point through the MATLAB application.

Recognition and Modeling of charactersdeals with the recognition of isolated characters and words rendered in a single connected stroke fashion. Writing in the air is expected to have increased uncertainty compared to writing on a hard surface. It further presents the results of overlapped air-writing recognition with a usability study.We opt for a hybrid tracking system that is simple to control and convenient to use. Due to different physical meanings and dimensions of the motion data, it is less meaningful toapplythemethods mentioned above toourdataset for direct comparison.

III. RELATED WORK

Hidden Markov models (HMMs) are widely used for online handwriting recognition [4], [5]. In [6], 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. Many sign language recognition systems use HMMs with various sensing technologies, such as data gloves and visionbased techniques. Different motion sensing and tracking technologies impose various behavioral load on the user. The implementation of a gesture-control interface contains two key components: motion tracking and gesture recognition. We have to capture the motion before performing gesture recognition. Motion capture devices are essentially the input device for a motion-based user interface. Vision-based techniques provide more natural and unencumbered interaction. An ideal vision-based system poses no requirement on the user to wear markers, gloves, or long sleeves. It is difficult for a human to recognize the overlapped handwriting. we track air-writing with 6-DOF motion data (translation and rotation), which is also different from the conventional 2- D spatial trajectory of pen-based writing.

IV. PROPOSED ALGORITHM

A. UNIQUE WRITING STYLE:

a motion word is formed by connecting motion characters with ligature motions in-between. When there is no haptic or visual feedback, the ordinary left-to-right writing style is is difficult to maintain without overlap or shape distortion.we simply connect the pen-up and pen-down strokes to form unistroke letters altogether. , we ask the user to write every character of a word in a layer by-layer manner, overlapping all letters of the word in the same envisioned virtual box, a writing style we term “overlapped air writing,” which supersedes the usual connected writing style and appears to be more suitable for air-writing.

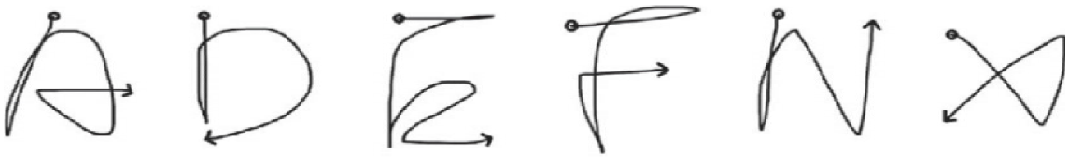
In Recognition ofmotion word is formed by connecting motion characters with ligature motions in-between. When there is no haptic or visual feedback, the ordinary left-to-right writing style is difficult to maintain without overlap or shape distortion. In air-writing, both allographs and different stroke orders can result in different spatiotemporal patterns for the same letter, which require separate models for classification. We have to ensure collection of sufficient data for modeling air-writing while keeping the recording process manageable in time

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

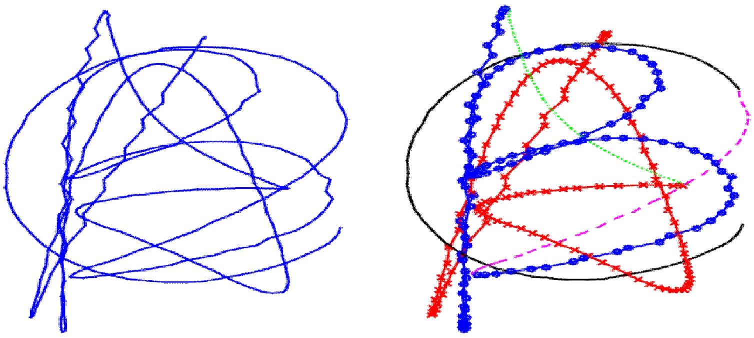
(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016



(a) A (b) D (c) E (d) F (e) N (f) X

Fig. 1. Illustrations of the unistroke writing of isolated letters. (a) A. (b) D. (c) E. (d) F. (e) N. (f)



(a) ABC (b) segmented ABC

Fig.2.(a) shows the projected trajectory of an air-writing of the letters ABC

It is difficult for a human to recognize the overlapped handwriting. To better illustrate the contrast and challenge, we manually segment the motion word and show the results in Fig. 1(b). The dash lines are ligature motions that connect characters A, B, and C. we track Recognizing on the Fly with 6-DOF motion data (translation and rotation). Figs. 1 and 2 demonstrate several key differences between air-writing (beyond single characters) and conventional hand-writing. The first major difference is the lack of penup/pen-down movesFigs. 1 and 2 demonstrate several key differences between air-writing (beyond single characters) and conventional hand-writing. The first major difference is the lack of penup/pen-down movesandhaptic feedbackandhapticfeedback.

B. *DESCRIPTION OF THE PROPOSED ALGORITHM:*

Aim of the proposed algorithm is touse a hybrid framework for 6-DOF motion tracking: the Worldviz PPT-X4 for optical tracking of the position of the infrared tracker and the Wii Remote Plus (Wiimote) for the inertial measurements of the acceleration and angular speed The system tracks a specially designed handheld device and provides both explicit (position and orientation) and implicit (acceleration and angular speed) 6-DOF data sampled at 60 Hz.Each isolated motion character (A to Z) was recorded ten times by every subject. For motion words, we select 40 words from common television channels, e.g., `ABC`, `CNN`, `FOX`, and common digital/Internet services, such as `TV`, `MUSIC`, and `GOOGLE`. The shortest word has two characters, and the longest one is `DISCOVERY`.

Step 1:*Feature Processing*

From the 6-DOF motion data, we derive five features (observations): position P and velocity V from optical tracking, orientation O, acceleration A, and angular speed W from inertial tracking. Let $P^o = [p_x, p_y, p_z]^\top$ denote the positions, and $V^o = [\Delta p_x, \Delta p_y, \Delta p_z]^\top$the rate of change in position. The orientation is represented in quaternion, $O^o = [q_w, q_x, q_y, q_z]^\top$..The implicit 6-DOF motion data form two features. Let$A^o = [a_x, a_y, a_z]^\top$denote the device-wise accelerations and denote the angular speeds in yaw, pitch, and roll, respectively. The superscript 0 indicates the original data from the tracking system before processing. The notations above represent the time sequences of a recorded air-writing in the corresponding coordinates, e.g. where N denotes the number of samples captured in a writing pattern.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

TABLE I
DURATIONS (IN NUMBER OF SAMPLES) OF MOTION CHARACTERS BY 22 SUBJECTS

	avg	std		avg	std		avg	std
A	159.5	37.4	J	60.6	12.0	S	92.7	17.8
B	156.7	37.1	K	136.8	26.4	T	88.6	16.4
C	77.3	19.8	L	64.8	15.0	U	73.5	14.1
D	118.7	24.9	M	146.4	29.4	V	67.2	11.5
E	190.8	48.6	N	115.7	21.1	W	110.4	19.3
F	132.6	27.4	O	85.1	17.4	X	91.3	16.1
G	149.7	35.1	P	107.6	20.3	Y	105.7	20.5
H	137.5	29.9	Q	119.4	26.4	Z	94.1	18.6
I	42.8	10.3	R	134.9	24.5			

The sample rate is 60 H

Step 2: *Air-Writing Modeling*

We use HMM models for motion characters can be readily concatenated to form a motion word with additional connecting ligature motions. We manually clusters characters according to the position of the starting and ending points.Gesture recognition typically involves a limited vocabulary set. It is relatively easy to collect sufficient data of each gesture and straightforward to model each gesture directly from its own recordings. However, the vocabulary of air-writing can easily be thousands of words, and it is difficult to collect enough data for every word in the vocabulary. The data sufficiency problem prevents a designer from directly using whole “word” models.We define the ligature as the motion from the ending point of the preceding character to the starting point of the following character.

TABLE II CLUSTERS FOR START AND END POINTS OF CHARACTERS

	Start point		End point
S1	BDEFHKLMNPRTUVWXYZ	E1	BDSX
S2	AIJOQ	E2	ITY
S3	CGS	E3	CEGHKLMQRZ
		E4	JP
		E5	AF
		E6	O
		E7	NUVW

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

We heuristically cluster the start and the end points of uppercase letters into three and seven groups, respectively, in Table II. The clustering reduces the number of ligature models to 21 (7 × 3). We label every ligature given the preceding and the following characters and form the model for a motion word. We further considered clustering ligatures with a data-driven decision tree based on likelihood and found that to be more effective.The decision tree attempts to find those contexts which make the largest difference in the calculated likelihood to distinguish clusters. Each question splits the current pool of training dataintotwosetsandincreasesthelikelihoodwiththeuseoftwo sets rather than one. We branch the decision tree by selecting the question that maximizes the increase of likelihood and repeat the process until the likelihood increase achieved by any question at any node is less than a prescribed threshold.

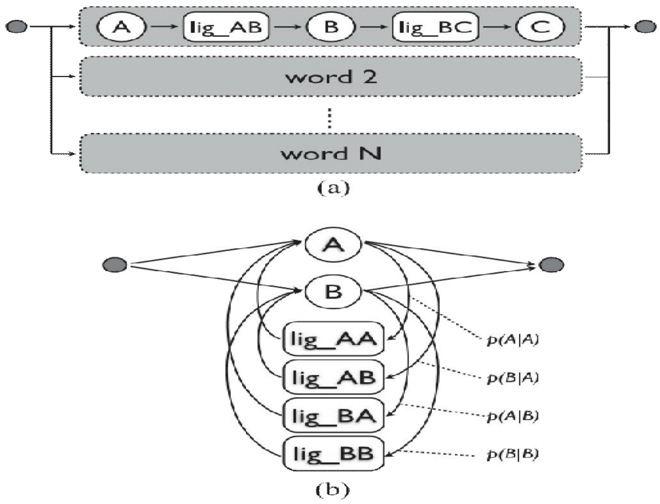


Fig. 3. Decoding word networks. (a) Word-based. (b) Letter-based (simplified).

In a motion word, we do not have or need the character-level segmentation. With the composite word HMM, character and ligature segmentation (alignment) can be simultaneously accomplished during recognition. Because the motion trajectories of characters and ligatures usually blend together, the ground truth of segmentation may be ambiguous.
We manually segmented all the motion words in the 40-word vocabulary recorded by subject M1. The manual segmentation is used for initial estimates of ligature models, which is proven to work better than ones that are initialized with zero means and global variances.

Step 3: *MOTION CHARACTER AND MOTION WORD RECOGNITION EVALUATION*
We perform motion word recognition with two approaches: word-based and letter-based recognition. In word-based recognition, we synthesize the HMM for every word in the vocabulary. For word-based word recognition, we use the refined HMMs of character and 21 hard clustered ligatures to build the decoding word network.Letter-based recognition decodes on a letter basis. Fig. 3(b) illustrates a simplified example of a letter-based decoding network that is built with letters A and B and the corresponding ligature models. The letter-based decoding network allows arbitrary decoded letter sequences and can handle OOV words. Another advantage is that letter-based word recognition allows progressive decoding while the user is writing, unlike the wordbased recognition that requires the user to complete a word.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

P	3.72	(3.60)
V	6.12	(2.88)
O	3.81	(5.05)
A	7.97	(7.38)
W	7.92	(3.34)
PV	1.61	(2.16)
AWO	1.84	(2.37)
$PVOAW$	1.05	(1.23)
$\hat{P}$	3.88	(3.55)
$\hat{V}$	6.15	(2.69)
$\hat{P}\hat{V}$	1.61	(2.06)
$\hat{P}\hat{V}OAW$	1.05	(1.33)

TABLE I11 :CER OF MOTION CHARACTER RECOGNITION

V.SIMULATION RESULTS

The average writing/typing time and total traverse distance for words of different length in Table 1. Because airwriting is recognized on a word basis, we report the average number of attempts to correctly input a word. Longer words tend to have higher recognition accuracy and hence need fewer attempts. The average writing time of a two-letter word is 3.9 (= 5.4/1.38) s. For virtual keyboard, we report the average number of extra keystrokes.

TABLE 1V
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	

The refined HMMs of character and decision-treeclustered ligatures to build the letter-based decoding word network. To further improve the recognition performance, we utilize the statistics of letter sequences of the vocabulary. We estimate the bigram language model for the 40-word and 1k-word vocabulary separately.Each ligature model depends on its previous and next characters; therefore, we can easily embed the conditional probabilities of the bigram language model into the transition arcs from characters to ligatures

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

TABLE V
RESULTS OF MOTION WORD RECOGNITION ON THE 1*k*-WORD VOCABULARY WITH SUBJECT "M1"

features	word-based	letter-based (backoff)	
	WER (%)	WER (%)	CER (%)
PV	0.80	1.90	0.66
*P^V^*	0.80	2.80	0.97
AW O	1.60	7.00	2.59
PV OAW	0.90	4.10	1.42
P^V^ OAW	0.90	5.00	1.73

TABLE V1
RESULTS OF MOTION WORD RECOGNITION ON THE 40-WORD VOCABULARY WITH 22 SUBJECTS

features	word-based		letter-based(backoff)			
	WER (%)		WER (%)		CER (%)	
	average	std	average	std	average	std
PV	0.045	(0.144)	10.59	(6.63)	3.48	(2.67)
*P^V^*	0.023	(0.104)	9.20	(5.36)	2.86	(1.82)
AW O	0.0	(0.0)	14.93	(11.11)	5.70	(4.86)
PV OAW	0.0	(0.0)	11.57	(8.23)	4.15	(3.48)
P^V^ OAW	0.0	(0.0)	10.61	(7.31)	3.65	(2.82)

The average WER and CER of leave-one-out cross validation with the bigram language model in Table V. The CER includes insertion, substitution, and deletion errors. The pure inertial *AWO* has the highest error rates. Words-per-minute (WPM) is a common performance metric for text input efficiency. WPM is computed based on correctly inputwordunits,whereonewordunitisfiveletters(keystrokes). The WPM of air-writing and virtual keyboard are 5.43 and 8.42 , respectively.

TABLE VII
AVERAGE WER OF DIFFERENT DESIGNS OF LETTER-BASED MOTION WORD RECOGNITION ON THE 40-WORD VOCABULARY WITH 22 SUBJECTS

features	decision-tree	decision-tree + bigram	decision-tree + bigram + 2-best	decision-tree + bigram (w/o backoff)
PV	17.27	10.59	4.73	2.73
AW O	23.09	14.93	8.16	2.75
PV OAW	15.16	11.57	5.77	2.18

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

V1.CONCLUSION AND FUTURE WORK

The simulation results showed that the proposed algorithm performs better with the virtual keyboard. Recognition and modeling of characters is suitable for infrequent and short text input on a motion based user interface. Although virtual keyboard is faster than air-writing, a virtual keyboard requires a display and precise pointing. We use hard clustering and decision tree for concatenating character and ligature models. Hard clustering is proven to be sufficient for word-based word recognition. Decision tree improves the performance of letter-based word recognition. The word-based word recognition achieves relatively low WER but is not able to recognize OOV words. The word-based recognizer is suitable for applications that have a limited vocabulary 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. The results suggest that air-writing is suitable for short and infrequent text input on a motion-based user interface.

REFERENCES

M. Chen, G. AlRegib, and B.-H. Juang, “Air-writing recognition—Part II: Detection and recognition of writing activity in continuous stream of motion data,” *IEEE Trans. Human-Mach. Syst*, to be published.

(2015). Leap motion. [Online]. Available: http://www.leapmotion.com

M. Chen, G. AlRegib, and B.-H. Juang, “Feature processing and modeling for6dmotiongesturerecognition,”*IEEETrans.Multimedia*,vol.15,no.3, pp. 561–571, Apr. 2013.

J. Makhoul, T. Starner, R. Schwartz, and G. Chou, “On-line cursive handwriting recognition using hidden markov models and statistical grammars,” in Proc. Workshop Human Lang. Technol., 1994, pp. 432–436. [5] J. Hu, M. Brown, and W. Turin, “HMM based online handwriting recognition,” IEEE Trans. Pattern Anal. Mach.Intell., vol. 18, no. 10, pp. 1039–1045, Oct. 1996.

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

[7] J. Hu, M. Brown, and W. Turin, “HMM based online handwriting recognition,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 10, pp. 1039–1045, Oct. 1996.

[8] J. Mantyj¨ arvi,¨ J. Kela, P. Korpipa¨a,¨ and S. Kallio, “Enabling fast and effortless customisation in accelerometer based gesture interaction,” in *Proc. 3rd Int. Conf. Mobile Ubiquitous Multimedia*, 2004, pp. 25–31.

[9] Z. Zafrulla, H. Brashear, T. Starner, H. Hamilton, and P. Presti, “American sign language recognition with the kinect,” in *Proc. 13th Int. Conf. Multimodal Interfaces*, 2011, pp. 279–286.

[10] N.PugeaultandR.Bowden,“Spellingitout:Real-timeASLfingerspelling recognition,” in *Proc. IEEE Int. Conf. Comput. Vision Workshops*, 2011, pp. 1114–1119. C. Amma, M. Georgi, and T. Schultz, “Airwriting: Hands-free mobile text input by spotting and continuous recognition of 3D-space handwriting with inertial sensors,” in *Proc. 16th Int. Symp. Wearable Comput.*, 2012, pp. 52–59.

[11] L.Jin,D.Yang,L.-X.Zhen,andJ.-C.Huang,“Anovelvisionbasedfingerwriting character recognition system,” in *Proc. 18th Int. Conf. Pattern Recog.*, 2006, vol. 1, pp. 1104–1107

[12] A.Schick,D.Morlock,C.Amma,T.Schultz,andR.Stiefelhagen,“Visionbased handwriting recognition for unrestricted text input in mid-air,” in *Proc. 14th ACM Int. Conf. Multimodal Interaction*, 2012, pp. 217–220.

[13] X. 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, pp. 85–93, Oct.–Dec. 2013.