# Recognizing arbitrarily connected and superimposed handwritten numerals in intangible writing interfaces 

Cheng-Chin Chiang*, Ren-Hong Wang, Bo-Ruei Chen<br>Department of Computer Science and Information Engineering, National Dong Hwa University, Shoufeng 974, Taiwan

## ARTICLE INFO

## Article history:

Received 20 December 2015
Received in revised form

## 25 June 2016

Accepted 8 July 2016
Available online 10 July 2016

## Keywords:

Intangible writing interface
Finger gesture recognition
Dynamic time warping
Graph path finding
Key numeral spotting


#### Abstract

An intangible writing interface is a new input mechanism that allows users to write in the air with hands or fingers. Since all written characters are connected in one writing trajectory, each writing trajectory is referred to as a one-stroke finger gesture. Because of no visual aid to indicate the effective writing area in the air, users often write conservatively the texts in a restrictive area to avoid invalid motion sensing. Consequently, each handwritten character may not only arbitrarily connect to others, but also highly superimpose over others. Aiming at addressing the challenging issue to segment and recognize the arbitrarily connected and superimposed characters in one-stroke finger gestures, this paper proposes a solution comprising (1) a simple but effective method to spot any numeral of interest, called a key numeral, from arbitrary one-stroke finger gestures and (2) a novel method to simultaneously accomplish the segmentation and recognition of all individual numerals in one-stroke finger gestures by a graphbased path finding algorithm. The evaluated performance shows that the proposed solution achieves satisfactory recognition accuracy even though no sophisticated matching features and context-dependent information are employed.


© 2016 Elsevier Ltd. All rights reserved.

## 1. Introduction

When using digital writing devices such as pen tablets and touch panels, users may have concerns about (1) the burden of carrying the externally bundled writing devices, (2) the worries of how to prevent devices from damage or loss, and (3) the sanitary issues of using publicly accessible writing utilities. An intangible writing interface (IWI) can relieve the users of the above concerns. Through an IWI, users can freely write in the air with hands or fingers, thereby avoiding both the load of carrying external devices and the indirect spread of diseases though device contacts.

A feasible and convenient IWI should allow a user to write multiple characters continuously in a writing trajectory. All characters are thus connected in a writing trajectory. Such a trajectory, which is called a one-stroke finger gesture, would incur great difficulty in segmenting and recognizing the connected characters in it. During the midair writing, a writer often writes the texts conservatively within a restrictive area to avoid invalid motion sensing because of no visual aid to clearly indicate the effective writing area. Consequently, the characters not only connect together in the gesture, but also distribute arbitrarily over the restrictive area. A further complication is that a character can highly overlap with others. Fig. 1 illustrates some examples of the

[^0]acquired one-stroke finger gestures. Apparently, the large overlapping and irregular layouts of characters complicate the segmentation and recognition of the individual characters in the gestures.

The goal of this paper is to propose a novel solution to tackle the problem of segmenting and recognizing handwritten numerals arbitrarily connected and superimposed in one-stroke finger gestures. We target at connected numeral gestures because the numeral set is not large and numeral gestures need no learning before using for common users. More importantly, numeral gestures are language-independent. Hence, numeral gestures are more suitable for popularizing the developed applications than other specific alphabetic or symbolic gestures. Besides the applications of phone number dialing and numeral password inputting, an intangible numeral writing interface also can remotely control multiple number-identified home appliances or select multiple numbered menu items through just a single finger gesture. The major technical contributions of our solution include (1) a simple but effective method to spot the occurrence of any designated numeral within an arbitrary one-stroke finger trajectory and (2) a graph-based path finding algorithm to simultaneously address both the segmentation and the recognition of numerals in onestroke finger gestures. Unlike many other related studies concerning the recognition of connected alphabetic characters, the proposed solution does not employ any context-dependent information because numeral strings reveal no lexical context


Fig. 1. Examples of connected handwritten numerals 456 in one-stroke finger gestures acquired from intangible writing interfaces.
dependencies. In addition, the segmentation and recognition problems can be addressed satisfactorily without extracting sophisticated features from the handwritten numerals. The higher structural complexity of some larger alphabetic gesture sets may often cause confusion that requires more delicate design of the recognition method.

## 2. Related Work

There have been extensive studies on handwritten character recognition. Two general approaches are offline recognition and online recognition [1-3]. Methods of the offline recognition approach scan each character as a 2-D image without keeping the order of writing strokes. Therefore, image processing and analysis become the key for the recognition. In contrast, the online recognition methods sequentially sample a series of positional data along the writing trajectory and thus gain additionally the temporal order of writing strokes. Thanks to the temporal order of strokes, the online recognition approach generally achieves better performance than the offline recognition approach [1] and therefore becomes a more popular design option for most writing-enabled devices [4-6]. In either approach, exploring good character features and designing good classifiers are two major critical issues. In the exploration of character features, there are spatial features and temporal features. Typical features include chain codes [7], stroke curvatures [8], DCT coefficients [9], wavelet packet coefficients [10], Fourier descriptors [11], trajectory and velocity modeling [12], and other structural features [13]. The common classifier designs include support vector machines (SVM) [14], neural networks [15], dynamic time warping (DTW) [16,17], and hidden Markov models (HMM) [18,19]. Among these designs, the DTW and the HMM prevail over others because of their better handling of temporal data. The HMM is a powerful statistical model for classifying temporal patterns. However, it requires a sophisticated learning process which involves an intricate training algorithm and requires a large number of training samples. Compared with the HMM, the DTW is much simpler and requires no complicated learning process, though its recognition may not be as robust as that of the HMM.

As for the segmentation of connected handwritten characters,
researchers have proposed some segmentation techniques. Some methods locate the cutting points between connected characters by inferring heuristic rules from character structures and text layouts [20-23]. Other methods employ lexical context dependency of characters to discern ambiguous cutting points or confusing numerals [24-26]. Unfortunately, the heuristic rules are often strongly dependent on character structure and text layouts and usually vulnerable to many exceptional cases. The rules are not applicable to the cases of highly superimposed and irregularly situated handwritten numerals. Yin et al. [27] proposed a lexicondriven recognition method for recognizing gestures of character strings. Besides the use of lexical dependency, their method also introduces four geometric models to characterize the geometric properties of individual redundant strokes (i.e., connecting strokes between characters), individual characters, characters connected by redundant strokes, and characters immediately connected with characters, respectively. Each geometric model uses a different set of features for the geometric characterization and demands a statistical training process. The design complexity of the lexicondriven recognition method is rather high. Vikram et al. [28] also presented an accelerated method to recognize connected character gestures in the air. They predefine a fixed database of words, with each word being a gesture of a character string. Given an input gesture of character string, the recognition is actually a fast search of the best word that best fits the input gesture. The fast search is guided by the searching of each individual character gesture contained in the word. Hence, this recognition still relies upon the lexicons defined in the word database. Furthermore, the searching speed would highly depend on the size of the word database. Besides, the number of permitted gestures is restricted by the fixed word database. For the recognition of connected numeral gestures, since a numeral string reveals no character-level or word-level lexical dependency, no context-dependent information can be employed to assist the segmentation or recognition of individual numerals. In recent years some researchers have developed a few IWIs [29-36]. These developed IWIs share one feature in common, i.e., requiring users to write one character at a time, which makes them less feasible for practical applications.

The rest of the paper is organized as follows. Section 3 presents the technical details of the proposed method, including the extraction of features, a method for spotting a specific numeral within a one-stroke finger gesture, and a method for segmenting and recognizing each individual numeral in one-stroke finger gestures. In Section 4, we conduct some experiments to evaluate and compare the performance of the proposed method. Finally, Section 5 ends this paper with some concluding remarks.

## 3. Proposed method

### 3.1. Overview

Given a one-stroke finger gesture, our method approximates the trajectory with a series of polylines which serve as the composing elements of the numerals in the trajectory. Each polyline is referred to as a substroke from which we extract two different simple features, the stroke orientation and the vertex orientation. Our procedure to segment and recognize individual numerals in single-stroke finger gestures comprises two phases, key numeral spotting (KNS) phase and recognition by concatenation (RBC) phase. The KNS phase aims to locate all possible occurrences of every numeral (i.e., 0 through 9) in the gesture. Each located occurrence is referred to as a key numeral. To spot key numerals in the gesture, an endpoint (including the starting point and stopping point) detection method is proposed based on dynamic time warping (DTW) algorithm [37]. With the detected endpoints of all spotted

# https://daneshyari.com/en/article/533075 

Download Persian Version:

## https://daneshyari.com/article/533075

## Daneshyari.com


[^0]:    * Corresponding author.

