Engineering Applications of Artificial Intelligence 59 (2017) 1–14

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

A robot calligraphy system: From simple to complex writing by human gestures

Fei Chao[a,*], Yuxuan Huang[a], Xin Zhang[a], Changjing Shang[b], Longzhi Yang[c], Changle Zhou[a], Huosheng Hu[a,d], Chih-Min Lin[a,e]

[a] Cognitive Science Department, Fujian Provincial Key Laboratory of Brain-like Computing, School of Informatics, Xiamen University, 361005, China
[b] Department of Computer Science, Aberystwyth University, UK
[c] Computer Science and Digital Technologies Department, Northumbria University, UK
[d] School of Computer Science and Electronic Engineering, University of Essex, UK
[e] Department of Electrical Engineering, Yuan Ze University, Taiwan

ARTICLE INFO

MSC:
00-01
99-00

Keywords:
Robotic writing
Robotic calligraphy
Human–robot interaction
Human gesture recognition
Classifier ensemble

ABSTRACT

Robotic writing is a very challenging task and involves complicated kinematic control algorithms and image processing work. This paper, alternatively, proposes a robot calligraphy system that firstly applies human arm gestures to establish a font database of Chinese character elementary strokes and English letters, then uses the created database and human gestures to write Chinese characters and English words. A three-dimensional motion sensing input device is deployed to capture the human arm trajectories, which are used to build the font database and to train a classifier ensemble. 26 types of human gesture are used for writing English letters, and 5 types of gesture are used to generate 5 elementary strokes for writing Chinese characters. By using the font database, the robot calligraphy system acquires a basic writing ability to write simple strokes and letters. Then, the robot can develop to write complex Chinese characters and English words by following human body movements. The classifier ensemble, which is used to identify each gesture, is implemented through using feature selection techniques and the harmony search algorithm, thereby achieving better classification performance. The experimental evaluations are carried out to demonstrate the feasibility and performance of the proposed method. By following the motion trajectories of the human right arm, the end-effector of the robot can successfully write the English words or Chinese characters that correspond to the arm trajectories.

1. Introduction

Handwriting is a highly demanding task involving both dynamics and kinematics, and is therefore normally regarded as a typically human motion (Potkonjak, 2011; Zeng et al., 2016). The kernel technology of robotic writing is to combine a number of basic actions (for instance, letters of the alphabet or Chinese character strokes) in order to generate complex functions (such as full sentences). A secondary goal would be to optimize the visual quality of the robotic output, with the help of human expert selection. Moreover the process should be easily adaptable, that is human users require robots to be able to quickly learn to handle new characters, or new motions. Such kernel technology can be exported to many other domains, particularly where there is a high demand for robotic imitation of repeated human movements. Notable examples would include medical rehabilitation training, where robotic movements could take advantage of scaling and repeated movements to assist patient to rehabilitate from small to large movements, as well as customized painting in automotive design finishes, or industrial welding of non-linear specialized shapes. With this in mind, the following work on robotic writing should be viewed as the test bed for a much wider range of practical applications.

The process of robotic handwriting requires the robot to obtain trajectory information, whether the strokes of Chinese characters, or the shape of English letters. A number of recent approaches have applied direct programming methods to embed a font database within the robot's control systems, which require complicated mathematical calculations and image processing work (Man et al., 2010; Ma et al., 2010). However, the imitation of human actions is considered as an effective learning method to transfer skills and knowledge from human beings to robots (Argall et al., 2009; Chersi, 2012; Liu et al., 2010; Bandera et al., 2012; Lu et al., 2014).

A wide range of applications using 3D human activity recognition has been introduced in recent years (Holte et al., 2012). It is very useful for robots to acquire new skills without the need of complex program-

* Corresponding author.
E-mail address: fchao@xmu.edu.cn (F. Chao).

http://dx.doi.org/10.1016/j.engappai.2016.12.006
Received 6 January 2016; Received in revised form 11 August 2016; Accepted 5 December 2016

ming and implementations (Goodrich and Schultz, 2007). Human users also prefer a convenient and natural way to directly control robots to copy characters and letters (Matsui and Katsura, 2013; Acosta-Calderon and Hu, 2005; Chao et al., 2014a). In particular, applying pen-tablets to obtain trajectory information (Kulvicius et al., 2012) is also a fast and even more natural way to write (Billard et al., 2008; Dillmann, 2004; Dillmann et al., 2010). However, using human gestures will provide more fine information towards controlling robots to write characters, e.g. wrist and elbow positioning data can be used to support robotic posture control. Beyond robotic writing, the ability to learn general human gestures should allow the proposed method to be adapted to handle dynamic robotic manipulations, such as the capturing and grasping of physical objects. Further, human gesture information can be effectually represented by medical EEG or EMG signals. Thus, this research grounds for EEG or EMG controlled robotic writing, a practical version of mind control.

Strokes of both Chinese characters and English letters can be represented by human arm trajectories. English consists of 26 different letters, which can be represented by 26 classes of human arm gestures. All Chinese characters are constructed by strokes. Yao et al. (2004) used 28 strokes to construct all of the Chinese characters. Compared with writing English letters by hand to writing Chinese characters by hand, robots not only need to know how to write the strokes of a character, but also need to consider the layout of each character's stroke. In addition, if human gestures are applied to represent strokes and letters, the gesture recognition problem must still be considered. In particular, gesture information is presented by a large group of three-dimensional points; thus, a recognition mechanism is required to identify each gesture precisely. The advantage of human gestures is to build a robotic action database through imitations. A robot develops a more complex action by assembling simple actions from the database. Thus the more actions there are included in the database, the better the resulting operational ability the robot can achieve.

This paper proposes a novel approach to robotic handwriting based on our initial preliminary work (Chao et al., 2014a,c). In the previous work, all the strokes had to be pre-programmed by human engineers, and only the writing actions were controlled through direct human gestures. The previous work had the obvious limitation that significant additional human programming was required to simply add a new font to the robot's database.

The work reported in Chao et al. (2014a) successfully supported free writing without repeated training or complex programming. However, because the robotic arm simply followed the demonstrator's movements, it was very difficult to improve the writing quality of the strokes. In this paper, the trajectories of human hand movements have two uses: (1) the trajectories can be recognized by a robot's classification methods and (2) the trajectories are the font shapes of the characters themselves. A number of different methods to classify gestural expressions have been reported in the literature (Tsoumakas et al., 2008; Zhou et al., 2012; Ramakrishna et al., 2013; Ramanathan et al., 2014; Ohn-Bar and Trivedi, 2014; Gong et al., 2014), including “human gesture corpora based methods” (Ruffieux et al., 2015), “dynamic Bayesian networks” (Suk et al., 2010), “Gaussian mixture modeling” (Vrigkas et al., 2014), “3D extremity movement observation” (Tran and Trivedi, 2012), and “hidden Markov models” (Mitra and Acharya, 2007; Richarz and Fink, 2011). However, this research is inspired from Schumacher et al.'s (2012) work, i.e. the problem is addressed by classifying trajectory segments comprising a fixed number of sampling points of human gestures.

Generally speaking, any conventional classifier could be used to recognize human hand gestures. A classifier ensemble can improve the performance of a single classifier system. However, an ensemble with too many classifiers may demand a large computational time. Classifier ensemble reduction (CER) aims to reduce the redundancy in a pre-constructed classifier ensemble, so as to form a much reduced subset of classifiers that can still deliver the same classification results (Tsoumakas et al., 2008; Diao et al., 2014, Yao et al., 2014). It is an intermediate step between ensemble construction and decision aggregation. Efficiency is one of the obvious gains from CER.

Removing redundant ensemble members may also lead to improved diversity within the group, and further increase the prediction efficiency of the ensemble (Diao and Shen, 2012). Existing literature approaches include techniques that employ clustering (Giacinto and Roli, 2001) to discover groups of models that share similar predictions, and subsequently prune each cluster separately. Other approaches use “reinforcement learning” (Partalas et al., 2009) and “multi-label learning” (Markatopoulou et al., 2010) to remove redundancy. In this paper, a new approach for CER that builds upon the ideas from existing feature selection techniques (Diao et al., 2014; Yao et al., 2016) is applied to classify a human demonstrator's gestures, so as to achieve a higher recognition rate for robotic writing.

In this work, a three-dimensional vision sensor, “Kinect”, is deployed to detect human right hand gestures. Kinect devices are widely applied in many robotic systems (Zaraki et al., 2014; Cazzato et al., 2014). The human hand's trajectories must be consistent with the character's trajectories. The robot system captures the human gestures and controls the robotic arm to write the designated trajectories. In particular, the captured trajectories are then converted to an array of hand trajectory data. A novel reduced classifier ensemble for recognition is used to improve the gesture recognition accuracy. The classifier ensembles are known to usually improve recognition performance in a wide range of pattern recognition tasks (Diao et al., 2014). A robot with a five DOFs arm receives the captured stroke trajectories, and kinematic algorithms are used to convert the stroke trajectories to the arm's joint values; then, the robot completes the writing task. This approach reduces the complexity of creating robotic writing, thereby enabling robots to exhibit higher flexibility. Additionally, the robotic writing guided by human gestures can introduce a natural and convenient way to control robots to execute complicated tasks. The main contributions of this paper are summarized as follows:

- The method for automated generation of robot's font database of Chinese character strokes and English letters from human arm's gestures (Sections 2 and 3.2).
- The method for empowering the robot's Chinese character and English word writing ability through the exploitation of the learned font database and given human gestures (Section 3).

The remainder of this paper is organized as follows. Section 2 describes the proposed framework and methodology used for robotic hand writing. In Section 3, human hand gestures and robot arm control are introduced. Section 4 presents the experimental results and discusses their implications. Finally, a brief conclusion and potential future work are given in Section 5.

2. The proposed approach

2.1. Robotic handwriting framework

Fig. 1 describes the framework of the robotic handwriting. First of all, a human demonstrator stands in front of the robotic system. The human uses one arm to perform predefined poses. The Gesture Sampling module is implemented by a Kinect device, which only captures the skeleton information of the human's poses. The skeleton information is sent to the Trajectory Capture module, in which the captured gestures are presented in 2-dimensional point arrays of the human's right arm trajectories. Then, the remaining approaches are divided into two phases: (1) the training phase, which includes classifier learning, classifier ensemble reducing, and obtaining trajectory information; and (2) the control phase, which uses the reduced classifier ensemble to identify the human gestures, and invokes the obtained trajectory information to write the identified strokes and