Engineering Applications of Artificial Intelligence 25 (2012) 392–403

Contents lists available at SciVerse ScienceDirect

Engineering Applications of Artificial Intelligence

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

Freely-drawn sketches interpretation using SVMs-chain modeling ☆

Kun Yang [a,b], Zhijun Li [a,*], Jingwei Ye [c]

[a] *Department of Automation, Shanghai Jiao Tong University, Donghuan Rd. 800, 200240, China*
[b] *Department of Mathematics, Shanghai Jiao Tong University, Donghuan Rd. 800, 200240, China*
[c] *School of Microelectronics, Shanghai Jiao Tong University, Donghuan Rd. 800, 200240, China*

ARTICLE INFO

Article history:
Received 27 October 2010
Received in revised form
2 October 2011
Accepted 3 October 2011
Available online 21 October 2011

Keywords:
Sketch recognition
Scale space filtering
Support Vector Machines
Hidden Markov Models

ABSTRACT

The growing popularity of tablet PCs and intelligent pen-centric computing has increased the importance of freehand sketch recognition algorithms. In this paper, the proposed method integrates the temporal, spatial and geometric constraint information to improve the recognition accuracy. To interpret the sketch as an incremental process, the paper investigates the use of the information fusion technique with Support Vector Machines (SVMs) chain for modeling and understanding the spatial and temporal information of sketch sequences. Online sketch recognition is achieved through the use of the SVMs-chain for systematically modeling the dynamic and stochastic behaviors of the sketch. To validate its efficiency, the experimental results in various domains and the comparison with traditional Hidden Markov Models have been presented.

1. Introduction

Currently, the pen-centric computing becomes popular. Microsofts Tablet PC has established a beachhead as a commercially available platform for pen-centric applications. Batteries, displays, and computing hardware continue to improve. There are increasing demands for the software that deals with complex data, such as images, 3D models, and video, using the electric pen input. Online recognition has two main advantages. First, it interprets and displays the pen input when entered for users convenience. For example, an engineer drawing of a mechanical system receives feedback to distinguish the different parts of the system, which sets the stage for users further input. Second, an online sketching system can utilize temporal information as well as geometric information of the drawing process.

Sketch recognitions are often characterized by the presences of spatial or geometric constraints, and temporal features. These aspects pose a number of technical challenges in developing sketch recognition systems. Numerous techniques have been proposed for sketch recognition within the past several years. However, some sketch approaches utilize the model or template based methods without regarding to the temporal information, such as Alvarado and Davis (2004) and Shilman (2002), which formulate the recognition problem on the basis of objects' implicit or explicit representations in terms of their structure (such as components, subcomponents, and their relationships). Therefore, these approaches may require high computational burden that makes them unsuitable for real-time recognition of realistic sketches, and require experts to specify domain knowledge or manually input structural descriptions of the objects. Other similar recognition algorithms were presented in Gennari et al. (2005), and Shilman and Viola (2004).

Recently, the Hidden Markov Models (HMMs) and SVMs have been explored on the recognition of the handwriting gesture, the degraded text and the symbols due to their robustness to cope with incomplete information and distortions. In Su et al. (2009), it proposes a segmentation-free strategy based on the first order HMMs to handle the off-line recognition of Chinese's handwriting. In Labush et al. (2008), the authors train SVMs on the extracted features to obtain classification performance in the digit recognition task. Moreover, there appears a few methods for recognizing sketch using HHMs by integrating the temporal information. In Anderson et al. (2004), a HMMs-based method was described that uses chain-code-like features to recognize the isolated symbols. In Sezgin and Davis (2005), a sketch recognition system utilizing HMMs was presented which does segmentation and recognition by combining outputs of HMMs.

Compared with HMMs, we find that (i) unlike HMMs, the proposed SVMs-chain can estimate some complex distribution pattern of the training data robustly with the stable computation cost; (ii) the SVMs-chain provides a simplistic probability model

☆ This work is supported by Mobile Computing Education research program of Microsoft Research Asia (No. FY08-RES-THEME-154), Natural Science Foundation of China under Grant (Nos. 60804003 and 61174045) and the project of International Cooperation of Ministry of Science and Technology of China (No. 2011DFA10950) and Shanghai Pujiang Program (No. 08PJ 1407000).

* Corresponding author, Tel./fax: +86 21 34204616.
E-mail addresses: yang.kun@acm.org (K. Yang), zjli@ieee.org (Z. Li), yejingwei@ic.sjtu.edu.cn (J. Ye).

0952-1976/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.engappai.2011.10.001

that directly deals with prior knowledge rather than presetting some parameters of the structures and the mixtures for HMMs; (iii) the accuracy of the inference by the SVMs-chain seems to be less sensitive than the HMMs to the loss of access to sets of observations.

Therefore, in this paper, a new statistical learning algorithm for the sketch recognition–SVMs-chain is proposed. It incorporates the strategies of graphic model methods, such as HMMs, with SVM's robustness in estimation of distribution pattern. It transforms the discriminant function of SVMs, into a pseudo probability density function, which can learn the distribution pattern of the training data in the feature space.

The main contributions of the paper are listed as follows:

(i) a new method for extracting the feature vertices combined with dynamic programming technique is proposed, and less time and space for computing are needed while presenting a good performance;
(ii) a modified primitive representation for human-handed sketch is utilized for feature abstraction;
(iii) the pseudo probability density function is presented for estimating the unknown probability distribution based on SVMs method; and
(iv) developing a sketch recognition based on our proposed SVMs-chain, which outperforms traditional HMMs approach in freehand sketch recognitions.

The paper is organized as follows. Existing algorithms for feature vertices extraction are briefly introduced and the adaptive window algorithm is presented in Section 2.1. Section 2.2 describes the way of encoding the free hand sketch. Section 3 presents the algorithm of probability distribution estimated by our proposed pseudo probability density function. Section 4 presents the evaluation results to show a significant advance in this proposed method compared with HMMs.

2. Sketch encoding

Compared with the traditional approaches by abstracting features from static sketches, the real-time data collection by digitalizing pen input can provide more information for facilitating the sketch recognition. The existing approaches for feature extraction can be rawly divided into two classes: one is the global-level feature extraction, which works when the complete sketch is available. The class of these methods is implemented in Kimura et al. (1993), Fonseca and Jorge (2000), Fonseca et al. (2002). The other is the stroke-level feature extraction, for example, the primitives representation in Sezgin and Davis (2006). Both classes have different emphasis and strengths. For example, the global-level features provide the global geometric information, such as bounding box, area of the convex hull, centroid. However, it is only efficient in the recognition of some simple sketches. In contrast, the stroke-level features provide more detailed information deriving from each stroke and their sequential order. This makes the use of temporal information feasible. Similar works are introduced in Williems et al. (2009), which shows that using the stroke-based features can improve the classification accuracy by 10% and 50% compared with the global feature representations.

2.1. Feature vertex detection

In the early stage of sketch recognition, a sequence of vertices along the freely-drawn line sampled in the temporal order is obtained for preprocessing and the structure of strokes can be recorded. A fast and stable algorithm to extract the structural information, represented with the feature vertices, from these sampled points of the original sketch is needed.

There are two contradict reasons, which make the feature vertex selection complicated: (i) the rawly sampled points are redundant for representing the structure of a sketch; (ii) the few points recorded are insufficient to depict a sketch. Therefore, an appropriate subset of these points should be saved as the feature vertices in order to make some balance between both.

2.1.1. Adaptive windows

The selection criteria for feature vertices take the turning angle $A_t(t)$ of each point along the strokes into account. It is a supplementary angle of the included angle defined by three successive points shown in Fig. 1. The feature vertices, which can effectively reserve the basic structure of sketch, are those points positioned at the corners of sketch. Hence, they are usually accompanied with the larger value of $A_t(t)$ than their neighboring points. The point with local maximal $A_t(t)$ can be selected as the candidates of feature vertices. In addition, the two endpoints of a stroke should be feature vertices, of which the $A_t(t)$ are manually set with large values.

The adaptive window approach is proposed to extract the local maximal $A_t(t)$ as shown in Fig. 2. The turning angle of a single point is compared with its neighborhoods within the limited window, $[t-w/2, t+w/2]$, where, t and w are the position of the

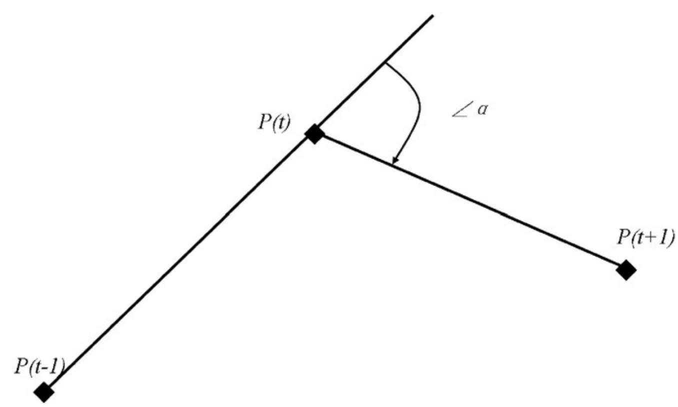

Fig. 1. The turning angle $\angle\alpha = A_t(t)$.

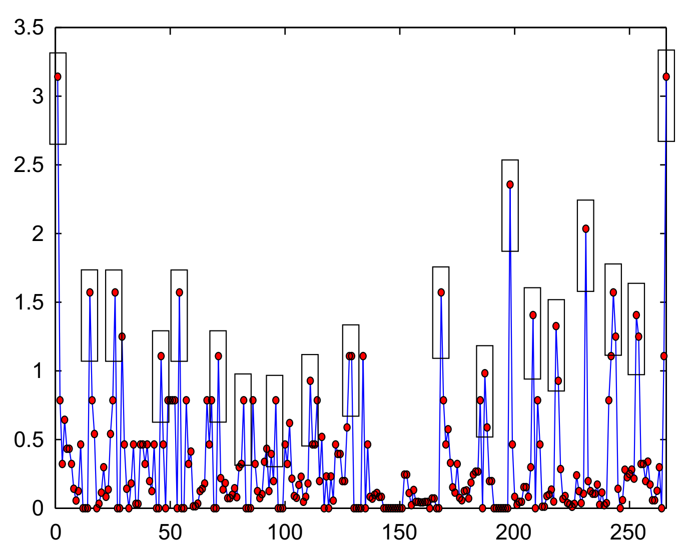

Fig. 2. Adaptive windows for picking the local maximums, window width $w=4$.