Recognition of Online Handwritten Music Symbols

Jorge Calvo-Zaragoza, Jose Oncina, and Jose M. Iñesta

Department of Software and Computing Systems, University of Alicante
Carretera San Vicente del Raspeig s/n, 03690 Alicante, Spain
{jcalvo,oncina,inesta}@dlsi.ua.es

Abstract. An effective way of digitizing a new musical composition is to use an e-pen and tablet application in which the user's pen strokes are recognized online and the digital score is created with the sole effort of the composition itself. This work aims to be a starting point for research on the recognition of online handwritten music notation. To this end, different alternatives within the two modalities of recognition resulting from this data are presented: online recognition, which uses the strokes marked by a pen, and offline recognition, which uses the image generated after drawing the symbol. A comparative experiment with common machine learning algorithms over a dataset of 3800 samples and 32 different music symbols is presented. Results show that samples of the actual user are needed if good classification rates are pursued. Moreover, algorithms using the online data, on average, achieve better classification results than the others.

Keywords: Pen-based recognition, Optical Music Recognition

1 Introduction

There may be several reasons for exporting a music score to a digital format such as easier storage, distribution and reproduction. Many composers transcribe their daily compositions to take advantage of these benefits. To this end, conventional music score editors or OMR (Optical Music Recognition) systems can be used; however, it is more profitable to digitize a score while the composer is writing it. In general less attention has been devoted to building this kind of system, but below we mention some research that has already been carried out.

The first work for pen-based recognition was the *Presto* system [1], which received as input short gestures that were generally mnemonic of the music symbols. These gestures were processed and translated to the actual musical symbols. The main drawback of this approach is that it forces the writer to be adapted to the gesture alphabet. In [2] a system is proposed for pen-based recognition of primitives (lines, circles, arcs, etc.). After the recognition, it groups these primitives to reconstruct the musical symbols. Kian et al. [3] proposed the use of Hidden Markov Models (HMM) for recognition of a little set (8 types) of the most common musical symbols. George [4] used the image generated by

In MML 2013: International Workshop on Machine Learning and Music, ECML/PKDD, Prague, 2013.

the digital pen to learn an Artificial Neural Network (ANN) that would directly recognize musical symbols.

These studies have shown that the complete recognition of musical symbols written in the natural form of music is feasible. Nevertheless, none of them performed comparative experiments to assess which kinds of algorithms are more promising for this type of data. Furthermore, some of them only dealt with a very small set of the possible symbols of the music notation. The present work aims to show basic results for the task of recognizing online handwritten musical symbols. Experimentation with well-known machine learning algorithms over a large dataset is shown, so that the results can serve as a basis for future developments and comparisons.

2 Recognition of Online Handwritten Musical Symbols

The recognition task of online handwritten musical notation is focused on recognizing the musical symbols while the score is being written. Although this field is strongly related to offline OMR, it presents some relevant differences: the staff lines -one of the most important issues- does not interfere in the recognition since they are handled by the online system, the segmentation can be somehow performed intrinsically and the information about how the symbols are drawn is available.

From the strokes generated by an e-pen, an ordered set of points which indicate the path followed can be extracted. Each symbol can be drawn by one or more strokes. However, this is not the only kind of information that can be used; an image of the shape itself is also available for recognition (as it would be done in offline recognition). This dual nature of the data leads us to explore both ways to classify the symbols. Classification techniques for each of these modalities are presented in the following subsections.

2.1 Online Techniques

The recognition of the online modality uses the strokes generated by the pen. These strokes provide information about how the shape has been created segment by segment. Depending on the type of musical symbol and the pace of the writer, a greater or lower number of points will be generated. Since each sample has a different feature dimension, we will restrict ourselves to the use of nonparametric classification by using the Nearest Neighbor (NN) [5] technique.

The performance of this rule is strongly related to the dissimilarity measure $d(x, x')$ utilized. Two alternatives for the recognition of musical symbols are applied: edit distance with Freeman Chain Code [6] and Dynamic Time Warping (DTW) [7].

2.2 Offline Techniques

After drawing the symbol, an image can also be obtained by interpolating the points extracted from the strokes. The advantage of this representation is that it is robust against different speeds or different orders when writing the symbol.

The images obtained are resized to 20×20 binary images, as is commonly done in OMR classification systems [8]. On these 400-dimensional data, several algorithms can be used. Since this paper aims to establish a baseline, Nearest Neighbor (NN), Support Vector Machines (SVM) and Artificial Neural Networks (ANN) have been used.

3 Experimentation

Recognition experiments are shown in this section. The goal of this section is not to achieve low error rates, but to draw some conclusions about the recognition of this kind of data. The dataset used contains 3800 samples spread over 32 different musical symbols. These samples were drawn by 25 musicians using a *Samsung Galaxy Note 10.1* tablet and a stylus *S-Pen*. To cover more scenarios, some of them were experienced in handwritten composition while others have little composition experience. Each of them were asked to draw each class of the dataset four times in his/her own style.

Two experiments are conducted by means of cross validation. In the first, the samples of a particular musician are isolated, while the samples of the other remain as the training set (user-independent 25-fold cross-validation). In the latter, the samples of each writer are divided into four sets and each one is used separately (user-dependent 100-fold cross-validation). Figure 1 shows the mean error rate and the standard deviation of these experiments for each algorithm considered. The results of each fold are measured by a common $0-1$ loss function -the number of misclassified symbols divided by the number of total samples-.

As reported in the results of the user-independent experiment, algorithms are not robust when detecting samples from unseen users. All the algorithms provide high error rates, with a high variability in their results. However, online algorithms, on average, achieve better results than the others. In the user-dependent scenario, it is noticeable that algorithms using the online nature of the data have the best performance.

4 Conclusions

Nowadays, digitizing a new composition is a mandatory step for musicians. Much research has been devoted to the development of friendly music score editors but there is still no comfortable solution to this issue. The emergence of tablet computer devices allows musicians to approach the problem by using an electronic pen over a digital score. The recognition of online handwritten music notation is the task of recognizing the symbols generated in this situation. This paper

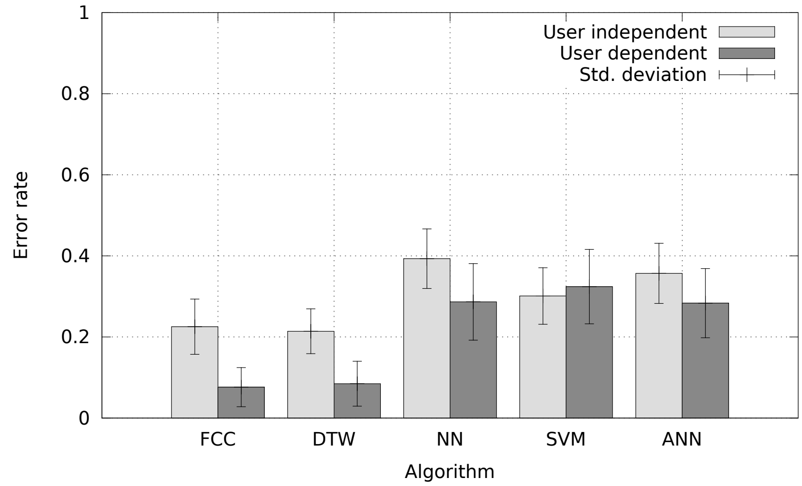

Fig. 1. Results of the two classification experiments.

shows the first basic results obtained by a comparative experiment with well-known machine learning algorithms on a dataset of 3800 samples and 32 different symbols. Two main conclusions were drawn from the experiments: it is difficult to achieve good classification rates in an user-independent scenario and, among the algorithms considered, those using the online modality obtain better results.

References

1. Anstice, J., Bell, T., Cockburn, A., Setchell, M.: The design of a pen-based musical input system. In: Computer-Human Interaction, 1996. Proceedings., Sixth Australian Conference on. (1996) 260–267
2. Miyao, H., Maruyama, M.: An online handwritten music score recognition system. In: Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on. Volume 1. (2004) 461–464
3. Lee, K.C., Phon-Amnuaisuk, S., Ting, C.Y.: Handwritten music notation recognition using hmm – a non-gestural approach. In: Information Retrieval Knowledge Management, (CAMP), 2010 International Conference on. (2010) 255–259
4. George, S.E.: Online pen-based recognition of music notation with artificial neural networks. Comput. Music J. **27**(2) (June 2003) 70–79
5. Aha, D.W., Kibler, D., Albert, M.K.: Instance-based learning algorithms. Mach. Learn. **6**(1) (January 1991) 37–66
6. Freeman, H.: On the encoding of arbitrary geometric configurations. Electronic Computers, IRE Transactions on **EC-10**(2) (1961) 260–268
7. Sakoe, H., Chiba, S.: Readings in speech recognition. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (1990) 159–165
8. Rebelo, A., Capela, G., Cardoso, J.S.: Optical recognition of music symbols: A comparative study. Int. J. Doc. Anal. Recognit. **13**(1) (March 2010) 19–31