

Resum

L'anàlisi i reconeixement d'imatges de documents històrics ha guanyat interès durant els darrers anys. La digitalització massiva juntament amb la interpretació dels documents digitalitzats permeten la preservació, l'accés i la indexació del llegat artístic, cultural i tècnic. L'anàlisi dels documents manuscrits és una subàrea d'interès excepcional. El principal interès consisteix no només en la transcripció del document a un format estàndard, sinó també en la identificació de l'autor d'un document davant d'un conjunt d'escriptors (l'anomenada identificació de l'escriptor).

La identificació de l'escriptor en documents manuscrits de text és una àrea activa d'estudi, i a la literatura és prolífica en contribucions significatives. No obstant, la identificació de l'escriptor en documents gràfics és encara un repte. El principal objectiu d'aquesta tesi és la identificació de l'escriptor de partitures musicals antigues, com a exemple de document gràfics. En referència a les partitures antigues, molts arxius històrics contenen un enorme volum de partitures musicals sense informació del seu compositor, i la recerca en aquest camp podria ser beneficiosa pels musicòlegs.

El marc de treball per a la identificació de l'escriptor proposat en aquesta tesi combina tres diferents aproximacions, corresponents a les principals contribucions científiques. La primera es basa en mètodes de reconeixement de símbols. Per a aquesta tasca, s'han proposat dos nous mètodes de reconeixement de símbols per fer front a les distorsions típiques dels símbols dibuixats a mà. El primer mètode està basat en l'alineament de seqüències (Dynamic Time Warping - DTW), on els símbols es descriuen emprant seqüències de vectors, i la proximitat entre símbols es mesura a partir de l'algoritme DTW. El segon mètode és el Model de Forma Difusa (Blurred Shape Model - BSM), que descriu els símbols emprant una funció de densitat de probabilitat que codifica la probabilitat de les densitats de les regions de la imatge.

La segona aproximació per la identificació de l'escriptor preprocés la imatge per obtenir línies de símbols musicals, i extreu informació de la inclinació i gruix de l'escriptura, les regions connexes, contorns i fractals. Finalment, la tercera aproximació extreu informació global, generant textures musicals a partir de les partitures, i extraient característiques de textura (filtres de Gabor i matrius de co-incidència).

Els bons resultats obtinguts demostren la idoneïtat de l'arquitectura proposada. Fins a on arriba el nostre coneixement, aquest treball és la primera contribució en la identificació de l'escriptor a partir d'imatges que contenen llenguatges gràfics.

Abstract

The analysis and recognition of historical document images has attracted growing interest in the last years. Mass digitization and document image understanding allows the preservation, access and indexation of this artistic, cultural and technical heritage. The analysis of handwritten documents is an outstanding subfield. The main interest is not only the transcription of the document to a standard format, but also, the identification of the author of a document from a set of writers (namely writer identification).

Writer identification in handwritten text documents is an active area of study, and the literature is prolific in noteworthy contributions. However, the identification of the writer of graphical documents is still a challenge. The main objective of this thesis is the identification of the writer in old music scores, as an example of graphic documents. Concerning old music scores, many historical archives contain a huge number of sheets of musical compositions without information about the composer, and the research on this field could be helpful for musicologists.

The writer identification framework proposed in this thesis combines three different writer identification approaches, which are the main scientific contributions. The first one is based on symbol recognition methods. For this purpose, two novel symbol recognition methods are proposed for coping with the typical distortions in hand-drawn symbols. The first one is a Dynamic Time Warping (DTW) based method, in which symbols are described by vector sequences, and a variation of the DTW-distance is used for computing the matching distance. The second one is called the Blurred Shape Model (BSM), in which a symbol is described by a probability density function that encodes the probability of pixel densities of image regions. The second writer identification approach preprocesses the music score for obtaining music lines, and extracts information about the slant, width of the writing, connected components, contours and fractals. Then, a k-NN classifier is used to categorize the document image. Finally, the third approach extracts global information about the writing, by generating texture images from the music scores and extracting textural features (Gabor features and co-occurrence matrices).

The high identification rates obtained in the experimental results demonstrate the suitability of the proposed ensemble architecture for the identification of the writer in music scores. To the best of our knowledge, this work is the first contribution on writer identification from images containing graphical languages.

Keywords: *Writer Identification, Graphics Recognition, Symbol Recognition, Optical Music Recognition.*