## Short Abstract:

The problem of writer identification refers to the task of deciding on the authorship of a piece of handwritten document by comparing it against a set of samples saved in a database. Technology has aided in the design of hand held devices like tablets that can capture a person's handwriting. The main constituent of these devices is an electronic pen/stylus that helps in obtaining the dynamic information in handwriting through its tip. Online writer identification systems make use of such textual content for deriving the identity of the writer.

The present thesis focuses on proposing novel descriptors for a text-independent online writer identification system. In the first part of our work, we represent the spatio-temporal trace from a document with descriptors derived from a codebook. The derived descriptors take into account, the scores of each of the attributes in a feature vector, that are computed with regards of the proximity to their corresponding values in the assigned codevector of the codebook.

The codebook used for formulating the codebook descriptors in our first proposal is generated by applying the k-means algorithm. This however comes with a limitation that each feature vector gets assigned to only one prototype. In order to address this, we propose descriptors constructed from a set of dictionary atoms obtained in a sparse representation framework to characterize the segmented sub-strokes of the handwritten trace. These descriptors attempt to capture the similarity of the attributes of each feature vector with the corresponding value in the subset of dictionary atoms which possess a non-zero sparse coding coefficient.

In the last part of the thesis, we attempt to explore additional information from the sparse coding coefficients that can be useful for writer identification. The additional information is related to quantifying in an average sense, the degree of importance of each of the dictionary atoms with regards to the dynamic characteristics of the enrolled writers. For the computation of the same, we propose two separate strategies that are primarily based on analysis of histograms. Added to these contributions, we also propose an identification framework that makes use of saliency values that are adapted to each writer enrolled in the system.

The efficacy of all proposed descriptors in the thesis are demonstrated on the IAM and IBM-UB1 databases. The results obtained are found to be promising when compared to prior works.

Abstract-2012_146102009