

ABSTRACT

In this report we present a system tasked with identifying the writer's identity using on-line handwriting captured from a white board. We have to determine which person out of a group of individuals is the author of the text. We are using 7/8 of the data for extracting the training features and then are using the training features to train the Gaussian Mixture Models (GMMs). We are modeling the distribution of extracted features using Gaussian Mixture Density and will use the results obtained to predict the correct writer's identity.

INTRODUCTION

Identification using handwritten notes is based on the concept that every person writes uniquely and could be characterized based on the features present in their handwriting. Writer identification is done primarily on two types of data namely online and offline. In online identification, The data about the position and velocity of pen is sent through an IR sensor present at the tip, the text written is recorded through a triangular sensor/ receiver attached at one of the corners of the whiteboard. The output of the sensor are (x , y)-coordinates of the pen tip along with time stamps of each location.

WORK PLAN

1. Pre-processing of the online handwriting data
2. Extraction of meaningful features from the data acquired.
3. Modeling of distribution of feature vectors using GMMs.
4. Testing of the model's accuracy by giving it handwriting data input with unknown writer identity.

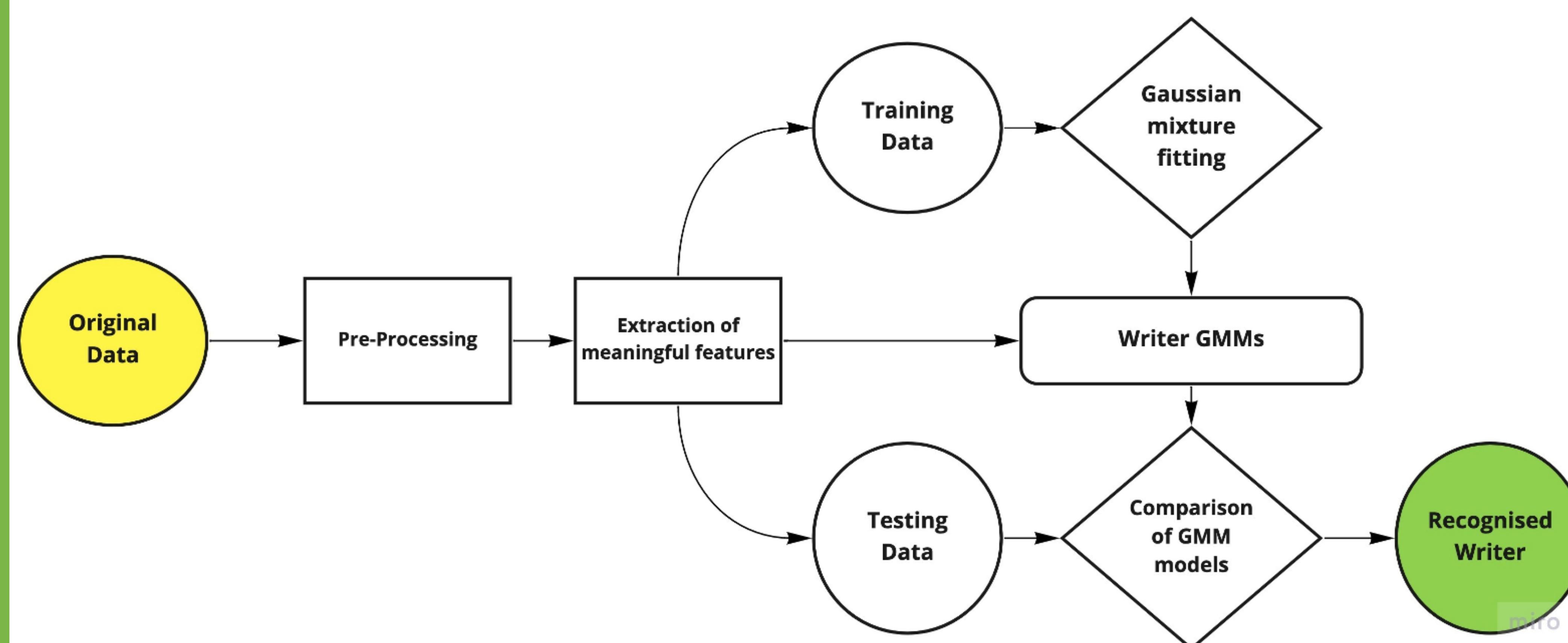
ONLINE WRITER IDENTIFICATION

Nishit Gaur (170108027)

Sameer Ghawana (170108034)

Supervised by: Dr. Suresh Sundaram
(Associate Proff. EEE Dept, IITG)

SYSTEM MODEL



Schematic representation of the model

REFERENCES

- [1] Schlapbach,A., Liwicki M., and Bunke H.(2007) "A Writer Identification System for On-line Whiteboard Data" Institute of Computer Science and Applied Mathematics, Universität Bern, Neubrückstrasse 10, CH-3012 Bern, Switzerland.
- [2] Namboodiri,A., Gupta,S.(2006) "Text Independent Writer Identification from Online Handwriting" Tenth International Workshop on Frontiers in Handwriting Recognition, Université de Rennes, La Baule (France).

RESULTS

We have trained models for various sets containing different counts of authors. Each model represents one author. We have conducted our experiment for two feature sets described below:

- **Feature Set 1:** In this feature set we have considered only point based properties.
- **Feature Set 2:** In this feature set we have considered both point based and stroke based properties.

Feature Set	No. of Authors	Accuracy
1	5	1.00
1	15	0.65
1	30	0.5
2	5	1.00
2	15	0.70
2	30	0.57

CONCLUSION

We have tested our model for three groups containing different number of authors and compiled our observations in the table above. We have observed that the accuracy values are better for feature set 1. This result was expected because in the feature set 2, we have considered only point based properties. While in the feature set 1, we have considered both point based and stroke based properties. Also, we have observed that as we increase the number of authors (classes), the accuracy results go down as expected.