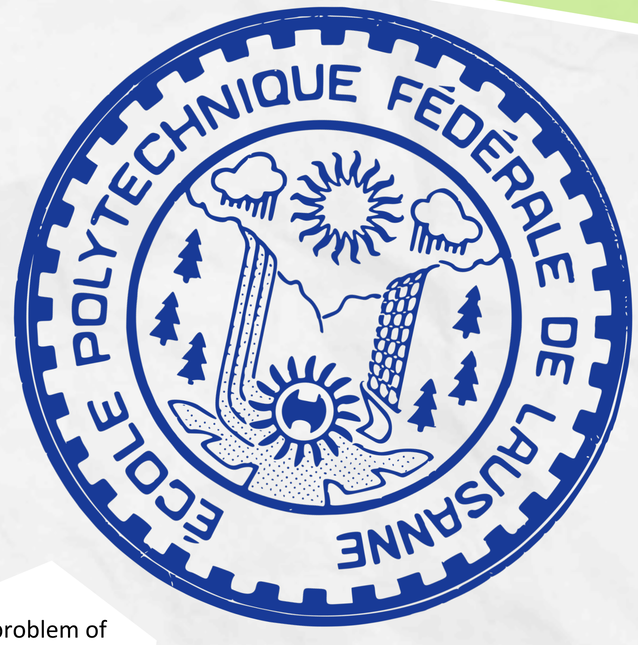
A SEMI-AUTOMATIC *Transcription* Tool for ancient manuscripts

Author:
Matthieu Simeoni
Supervisor:
Andrea Mazzei

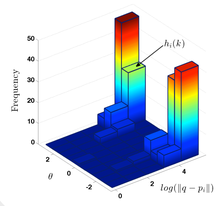
In this study, we investigate various techniques from the fields of shape analysis and image processing in order to construct a semi-automatic transcription tool for ancient manuscripts.

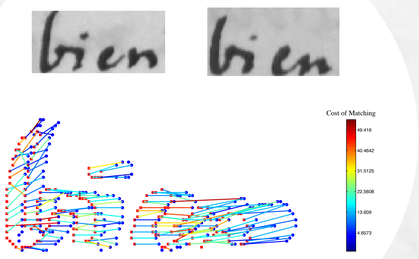


1 SHAPE MATCHING

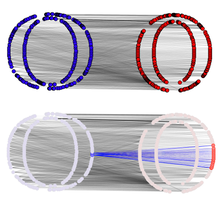
Shape matching is a **correspondence problem**: for each point on the first shape, we want to find the **"best" matching point** on the second shape. This matching can be efficiently performed by using **shape context.**

Shape context is a **rich local descriptor** of the shape.





The problem of shape matching is a problem of bipartite graph matching.

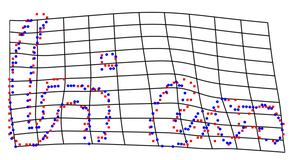


Potential matching points on the second shape for a point on the first shape.

2 SIMILARITY GRAPH

Using the shape matching procedure, we introduce different **shape distances** that we combine optimally in a single **similarity measure**.
This allows us to represent the shapes/words on a **similarity graph**.

We compute the amount of deformation necessary to align the two shapes.



Difference between the images before and after alignment

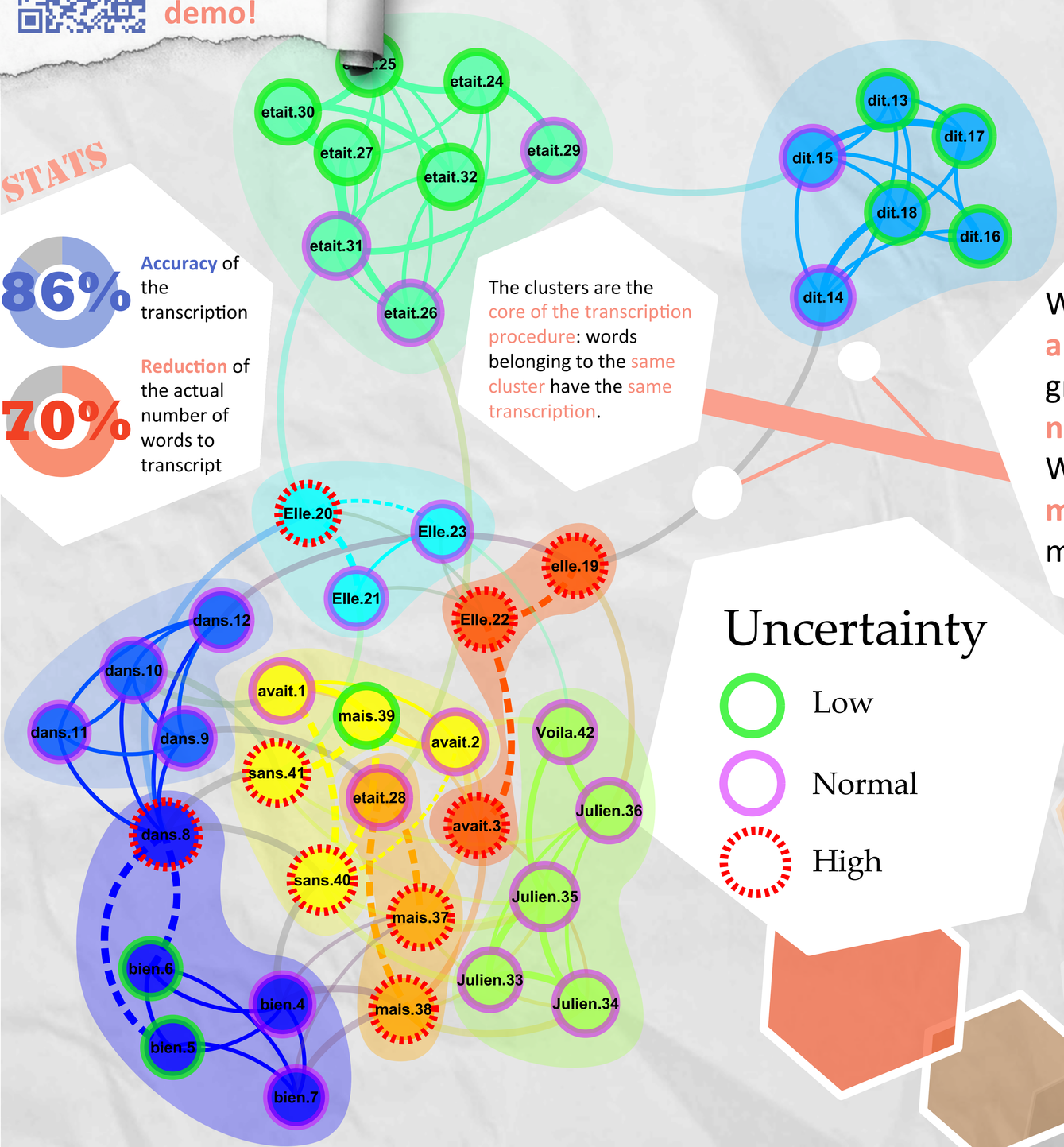
Check the online demo!

STATS

86% **Accuracy** of the transcription

70% **Reduction** of the actual number of words to transcript

The clusters are the core of the transcription procedure: words belonging to the same cluster have the same transcription.

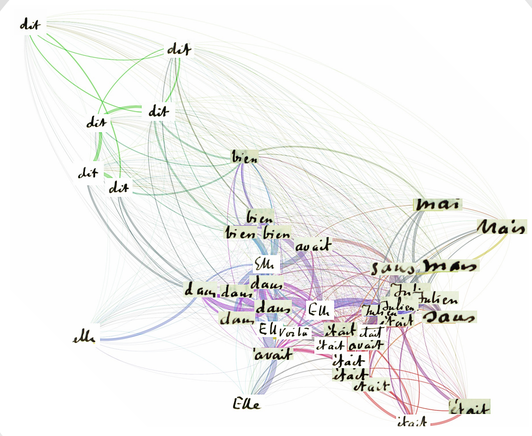


We perform a **clustering analysis** on the similarity graph using **the k-nearest neighbors** methodology.
We create an **uncertainty measure** to prevent from misclassification.

3 CLUSTERING

Uncertainty

- Low
- Normal
- High



Fully connected similarity graph of 42 handwritten words.

REFERENCES

[1] S. Belongie,J. malik, J. Puzicha, Shape Matching an Object Recognition using Shape Contexts, IEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, April 2002.