Research Plan:

Automated information extraction or information retrieval from document images
is an interesting area of research.
As document image databases in many different forms are growing everyday,
there is need of information summarizer which can report the key informations
from a large database of document images.

For example, extraction of emphasized texts (italic) from document images can play a big role in
meaningful information extraction or summarization of text.
It is seen that, a good number of publication houses use emphasized (italicized) words
for specifying the keywords, author's affiliations, etc. in the first page of their journal papers.
A system can be designed for quick information extraction to summarize the
content of the paper just by localization of the target emphasized words.
Very often old historical document pages are annotated by human beings to mark importance
by some means.
These annotations can be in the form of underlines, tick marks, cross marks, circles, etc.
By localization of the annotation objects, meaningful informations can be extracted
from this type of old historical image data bases.
So my research plan aims at developing algorithms for easy and robust information retrieval
from document images involving digital geometry, image processing techniques, etc.