Gland Segmentation of Hyperspectral Images for Cancer detection in Colon

Vinay Chandragiri (120101018, Class of 2016, chandragiri@iitg.ac.in)

Undergraduate Thesis, Department of Computer Science and Engineering, Indian Institute of Technology Guwahati, Assam, India

Abstract

Hyperspectral Imaging is the new modality in medical applications which is being used in Remote sensing. The main idea behind segmentation is to identify cancerous cells among the tissues. I am trying to address the problem of classifying cells by segmentation for cancer detection in colon tissue. The dimensionality problem has been tackled by Band Selection based on Independent Component Analysis. Future work involves comparision of unsupervised and supervised segmentation.

Keywords: *Hyperspectral Imaging, ICA, Clustering, Colon tissue*

Background

Hyperspectral Imaging is an emerging technology in the field of *Medical Imaging*. Instead of RGB values for each pixel, every pixel in the image contains a continous electromagnetic spectrum and is very useful in characterising the objects with immense accuracy and precision. HSI is capable of capturing both spectral information as well as spatial information in one shot.

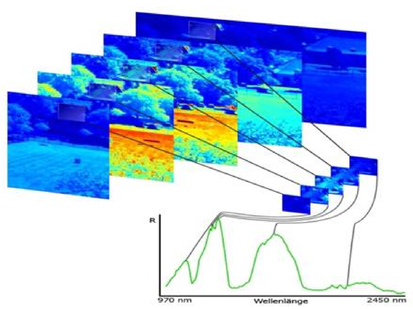

Fig. 1: Overview: HSI

Objectives

1. *Feature Selection* or *Feature Reduction* of HSI for dimensionality reduction.
2. Segmentation of glands and as well as classification of cells based on the properties of reflectance of different materials involved in the tissue.

Curse of Dimensionality

In a single HSI *(say of dimension A* x *B* x *C)*, where A x B represents number of pixels and C represents the number of spectral. Considering the medical data used, A and B were in order of 500 - 600 and C in the order 220 - 240. Each pixel in a image is a C-dimensional vector and there are more than 2,00,00 pixels in one image.

Features

Features are attributes from raw data such that there value make an instance. Different statistical methods such as *Clustering*, *Mutual Information*, *KL divergence* were explored in feature selection. *Independent Component Analysis*(ICA based feature selection was used because HSI are **highly correlated.**

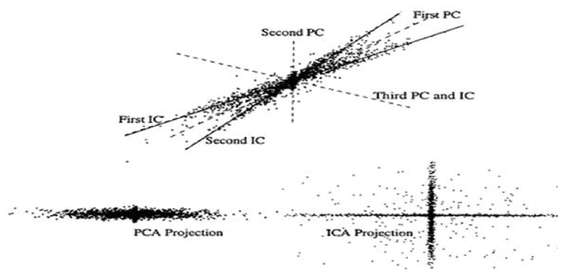

Fig. 2: Why ICA and Why not PCA ?

Band Selection based on ICA

The most crucial part of my work in Phase-1.

1. The absolute weight coefficients corresponding to each row are sorted and a band sequence is obtained.
2. The **Weight** matrix obtained from **ICA** determines how a particular band contributes to each material in the hyperspectral image.
3. This means that the bands with higher absolute weight coefficients will have more spectral information than the other bands.

Results

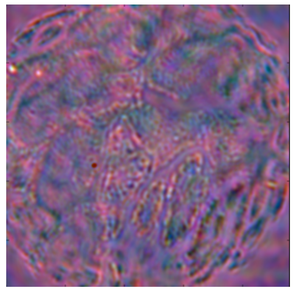

Fig. 3: Original Image

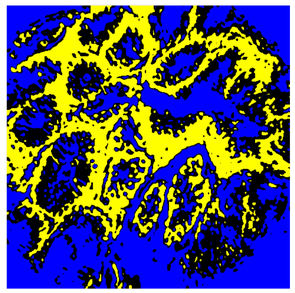

Fig. 4: Clustered Bands 195 & 198 from all 226

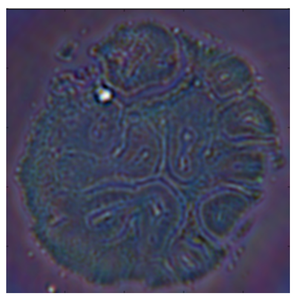

Fig. 5: Original Image

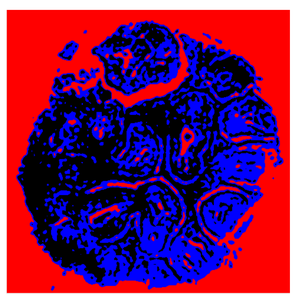

Fig. 6: Clustered Bands 192 & 194 from all 226

ICA Algorithm

Given X, assuming it as a linear combination of independent sources S and the mixing matrix A are to be determined such that X = AS. The independent sources are determined from the equation S = (inv)A * X i.e S = WX where W is termed as the Unmixing matrix.

1. *Input:* X of (N x M) represents N dimensional Sample. We can also fix the number of components we want to*(say C)*.
2. *Output:* W of (N x N), the Unmixing matrix.

Pseudocode:

```
for p in 1 to C:
    w_p ← Random vector of length N
    while w_p changes
        w_p ← (1/M) X g(w_p^T X) − (1/M) g'(w_p^T X) 1 w_p
        w_p ← w_p − Σ_{j=1}^{p−1} w_p^T w_j w_j
        w_p ← w_p / ||w_p||
Output: W = [w_1 ; ... ; w_C]
Output: S = WX
```

References

- A Comparative Analysis of Dimension Reduction Algorithms on Hyperspectral Data.*Kate Burgers, Yohannes Fessehatsion, Jia Yin Seo, Sheida Rahmani*
- Independent Component Analysis: Algorithms and Applications *Aapo Hyvrinen & Erkki Oja*

Conclusions

The results obtained are good as we can see a **structure emerging** but nothing quantitative can be inferred as it is unsupervised. This technique which I have implemented is new to medical imaging.