

# B A B I

## PENDAHULUAN

### I LATAR BELAKANG

Pengembangan yang terbaru system remote sensing lebih canggih memungkinkan pengukuran radiasi dalam lebih banyak lagi interval spektral dibanding system-sistem sebelumnya. Salah satu contoh teknologi ini adalah AVIRIS sistem, yang mengumpulkan data image dalam 220 bands. Teknologi ini disebut juga dengan teknologi hyperspectral. Ketika banyaknya dimensi data dari resolusi spektral tinggi meningkat, diharapkan kemampuan untuk mendeteksi objek dengan lebih terperinci juga meningkat. Sehingga, dengan semakin banyaknya feature, biaya dan kompleksitas feature extractor dan classifier, diharapkan ketelitian akan klasifikasi akan meningkat juga.

Study tentang karakteristik dari ruang berdimensi tinggi dan implikasinya terhadap analisis data hyperspectral melahirkan sejumlah karakteristik yang tidak biasa dari data berdimensi tinggi. Salah satu contoh adalah kenyataan bahwa ruang berdimensi tinggi sebenarnya sebagian besarnya adalah kosong dan diasumsikan bahwa banyak sekali data yang redundant [JONES87].

Di dalam riset-riset yang telah ada, data dengan spectrum yang tinggi akan direpresentasikan sebagai distribusi data multivariat, sehingga memungkinkan untuk menggunakan teori-teori statistik dalam proses komputasinya. Atas dasar penyajian ini, analisa statistik multivariat akan digunakan untuk menghasilkan hasil kuantitatif

Data multivariat biasanya berlokasi pada sub ruang berdimensi rendah. Akibatnya memungkinkan untuk mengurangi dimensi tanpa kehilangan informasi yang signifikan dan pemisahan antara kelas-kelas. Namun untuk melakukannya, harus mempunyai alat untuk menemukan sub ruang yang tepat. Konsekuensi lain dari karakteristik ini adalah bahwa local neighborhood bisa dipastikan kosong. Sehingga karakteristik *high-dimensional space* dapat menimbulkan masalah pada algoritma feature-extraction seperti Principal Component Transform (PCT/PCA) [JIM99].

PCA efektif untuk mengompres informasi dalam himpunan data multivariate dengan menghitung proyeksi orthogonal yang memaksimalkan sejumlah varians data. Informasi kemudian dipresentasikan dalam bentuk komponen image, yang merupakan proyeksi image dalam eigenvektor. Komponen image berkorespondensi dengan nilai eigen terbesar dengan anggapan bahwa sebagian besar informasi gambar tetap terjaga. Yang tidak menguntungkan adalah isi informasi dari image hyperspectral pada beberapa hasil tidak selalu tepat dengan proyeksinya [AGUST00]

Oleh karena itu diperkenalkan suatu teknik untuk *exploratory analysis of multivariate data* oleh **Friedman and Tuckey**. Idennya adalah untuk memproyeksikan himpunan data berdimensi tinggi menjadi ruang data berdimensi rendah dengan tetap menyimpan informasi yang diinginkan yaitu dengan menggunakan Projection Pursuit.

Meskipun C. M. Bachmann dan T.F. Donato menyatakan bahwa algoritma PCA merupakan simple subset dari general class algoritma Projection Pursuit [BACH00], pada kenyatannya banyak praktisi menggunakan metode-metode

tersebut secara terpisah . Sehingga banyak peneliti juga mengamati perbedaan-perbedaan dari keduanya. Hal inilah yang menjadi topik dalam penelitian ini.

## II PERMASALAHAN

Principal Component Transform (PCT) pada PCA adalah metoda yang efektif untuk mengkompres informasi himpunan data multivariate yaitu dengan menghitung proyeksi orthogonal yang memaksimalkan sejumlah varians data [AGUST00]. PCT juga seringkali digunakan para peneliti dalam menganalisis data hyperspectral dengan hasil cukup memuaskan. PCT mengumpulkan informasi pada sumbu tertentu, dan informasi dianggap sebagai varian. Sehingga varian yang tinggi dinilai sebagai informasi yang tinggi. Kelemahannya adalah jika informasi yang tinggi tersebut merupakan varian noise, maka hasil proyeksi tidak sesuai lagi dengan apa yang ingin dicapai.

Bahkan untuk objek yang relative kecil dibandingkan dengan latar belakangnya, dengan metode PCT akan memberikan kontribusi yang kecil terhadap keseluruhan variannya . Sehingga PCT kerap kali gagal mengenali objek tersebut.

Projection Pursuit dipropose untuk menyelesaikan permasalahan ini. Dengan adanya sphering yang menghasilkan matriks transformasi dan translasi invarian pada tahap preprosesing, menjadikan data-data yang dianggap kosong maupun redundant untuk tidak diikutsertakan dalam transformasi sehingga hanya data-data yang dianggap “interesting” yang akan ditransformasikan. Sehingga dalam kasus target detection image hyperspectral, Projection Pursuits memberikan hasil yang lebih baik [AGUST00][CHIANG01]

Projection Pursuit adalah suatu proses reduksi data dengan mengoptimalkan Projection Index. Berbagai macam Index proyeksi terdapat dalam PP, sehingga memudahkan pengguna untuk memilih PI sesuai dengan kebutuhan. Salah satu PI yang penting yang digunakan dalam Projection Pursuit adalah yang digunakan Friedman yang berbasis pada mean-squared difference antara projection score distribution dan Gaussian distribution. Pada kasus lain Huber dan Jones and Sibson, menggunakan ukuran information-theoretic yaitu entropy, yang dihitung pada projection score distribution.

Sedangkan Iffarraguerri dan Chang menggunakan information divergence yaitu salah satu information-theoretic criterion yang dapat digunakan untuk mengukur distance antara dua distribusi peluang. Jimenez dan Landgrebe menggunakan Bhattacharyya distance sebagai PI. Chiang and Chang menggunakan skewness dan kurtosis sebagai basis PI.

Dalam penelitian ini yang menjadi permasalahan adalah bagaimana mereduksi data data dimana objek yang akan dikenali jauh lebih kecil dari dari latar belakangnya atau biasa disebut sebagai anomaly atau outliers. Dari beberapa PI yang diperkenalkan oleh **Friedman and Tuckey** dalam *exploratory analysis of multivariate data*, PI skewness dan Kurtosis dianggap cocok untuk kasus ini, karena skewness dan kurtosis digunakan dalam statistika untuk mengenali objek pencila yang biasa juga dikategorikan sebagai outlier atau anomaly dalam citra.

Paper yang menjadi rujukan utama untuk menjadi bahan penelitian ini adalah paper S-S. Chiang dan C-I . Chang dengan judul “*Unsupervised Target Detection in Hyperspectral Images Using Projection Pursuit*” yang membahas Projection Pursuit dengan menggunakan Projection Index Skewness dan Kurtosis.

Akan tetapi paper asli tidak mendeskripsikan secara lengkap tahapan-tahapan metoda yang dilalui , asumsi-asumsi yang dipergunakan maupun hasil yang diperoleh. Sehingga penulis membutuhkan literatur-literatur tambahan yang tidak terlalu mudah didapat sebagai pendukung maupun rekomendasi-rekomendasi pakar statistik yang ada disekitar penulis untuk dapat memperjelas pemahaman teori-teori dan asumsi-asumsi yang diambil.

### **III RUANG LINGKUP PERMASALAHAN**

Banyak riset telah dilakukan untuk meneliti perbedaan antara Principal Component dengan Projection Pursuit , dengan konklusi Projection Pursuit dianggap lebih baik dari Principal Component dalam mereduksi data [AGUST00][CHIANG01][BACH00]. Penelitian ini dilakukan untuk mengetahui alasan mengapa Projection Pursuit lebih baik dari Principal Component dalam ekstraksi data. Untuk itu permasalahan yang akan dijawab dalam penelitian ini adalah:

1. Bagaimana proses reduksi data yang terjadi dalam Projection Pursuit ?
2. Bagaimana perbandingannya dengan proses reduksi data pada Principal Component
3. Sejauh mana preprocessing dalam kedua metode dapat mempengaruhi proporsi kumulatif pemilihan Band?
4. Apakah sajakah keunggulan maupun kelemahan Projection Pursuit jika dibandingkan dengan Principal Component Transform?

5. Bagaimana perbandingan citra hasil reduksi dari PCT dan PP , kedua metoda baik PCT maupun PP sama-sama menggunakan perangkat lunak MATHLAB untuk pemilihan ciri.

#### **IV TUJUAN PENELITIAN**

Tujuan penelitian adalah mengkaji latar belakang teoritis tahapan-tahapan proses Projection Pursuit dalam mengimplementasikan mekanisme ekstraksi data dengan menggunakan Projection Index Skewness dan Kurtosis yang memanfaatkan translasi invarian dan matriks sphering pada preprosessingnya, maupun Pricipal Component yang memanfaatkan matriks vector eigen dan proporsi kumulatif kontribusi nilai eigen untuk mereduksi data citra asli.

Dengan membandingkan kedua metode reduksi data tersebut menggunakan perangkat lunak MATHLAB, dapat diketahui sejauh mana proses sphering pada preprosessing mempengaruhi keduanya. Sehingga dapat ditentukan kombinasi yang terbaik antara preprosessing dengan metode optimasi yang lebih baik..

Kedua metoda diujikan pada citra hasil penelitian terdahulu [citra CASI], sehingga hasil keduanya dapat dibandingkan secara fisis, baik dilihat dari proporsi kontribusi kumulatif datanya maupun citra hasil reduksinya.

#### **V BATASAN MASALAH**

Secara metode, perbedaan yang mendasar dari PCT dan PP adalah pada tahap preprosessing dan optimasi pemilihan band. Penelitian ini dibatasi pada perbandingan tahap preprosessingnya, yaitu dengan mencoba menukar tahap

preprocessing pada kedua metode dan melihat hasil ekstraksinya, juga menghitung proporsi proporsi kumulatifnya.

Perangkat lunak yang digunakan adalah MATHLAB, dengan alasan bahwa perangkat lunak ini telah teruji secara internasional dengan banyaknya penelitian-penelitian yang menggunakan MATHLAB dalam mengolah data, juga dengan banyak tersedianya key untuk pengolahan statistik yang digunakan, baik dalam PCT maupun PP.

## **VI METODOLOGI PENELITIAN**

Metodologi yang digunakan pada penelitian ini adalah sebagai berikut :

1. Melakukan studi kepustakaan yang menjadi konsep dasar dari penelitian

ini , yaitu mengenai :

- Citra Hyperspectral
- Skewness
- Kurtosis
- Matrik Korelasi
- Reduksi Data
- PCA/ PCT
- PP
- Algoritma Genetika

2. Merancang dan mengimplementasikan perangkat lunak untuk :

- Algoritma PCT
- Algoritma PP
- Algoritma Genetika : PI Skewness dan Kurtosis

3. Melakukan ujicoba reduksi data menggunakan semua algoritma
4. Melakukan uji coba reduksi data terhadap algoritma PP dan PCT dengan pertukaran preprocessingnya.
5. Melakuksn uji coba reduksi data berdasarkan korelasinya
6. Menganalisa hasil ujicoba, kemudian mengambil kesimpulan dari hasil analisa tersebut.

## **VII SISTEMATIKA PENULISAN**

Sistematika penulisan dalam tesis ini disusun berdasarkan metodologi penelitian yang telah dilakukan, yang diuraikan sebagai berikut :

1. Bab 1 berisi pendahuluan yang menjelaskan latar belakang, permasalahan, ruang lingkup permasalahan, tujuan penelitian, batasan masalah , metodologi penelitian, serta sistematika penulisan.
2. Bab 2 menjelaskan landasan teori yang diperlukan untuk memahami penelitian ini , yaitu berupa studi literatur yang telah dilakukan baik dari jurnal, buku, maupun publikasi lain, yang didapat dari perpustakaan-perpustakaan, internet maupun rekomendasi pakar.
3. Bab 3 menjelaskan metoda-metoda yang akan dipergunakan dalam penelitian
4. Bab 4 menjelaskan tahapan semua eksperimen dan evaluasi hasil uji coba terhadap PCA maupun Projection Pursuit dengan optimasi Projection Indeks menggunakan Genetic Algorithm .Uji coba reduksi data berdasarkan matrik korelasi dituangkan dalam bab ini.

5. Bab 5 yaitu penutup yang merupakan kesimpulan dari hasil yang diperoleh dari uji coba yang telah dilakukan dalam penelitian ini.
6. Daftar pustaka berisi berbagai bacaan referensi yang dikutip dalam penelitian maupun penulisan laporan ini.
7. Lampiran, berisi hal-hal yang menunjang kesimpulan laporan .

