

BAB II

STUDI PUSTAKA

2.1 Tinjauan Pustaka

Tinjauan pustaka dilakukan sebagai acuan dan kajian terhadap penelitian ini. Dengan dilakukannya tinjauan pustaka terhadap penelitian-penelitian terdahulu yang sesuai dengan topik pembahasan maka akan diperoleh metode terbaik guna mendukung proses penelitian ini, berkaitan dengan penerapan transformasi *Curvelet* untuk mendeteksi kecacatan permukaan manggis.

(Fitrada, 2010), dalam penelitian “Aplikasi *Image Processing* untuk Menentukan Tingkat Mutu Buah Naga (*Hylocereus undatus*) Secara *non-Destructive*”, tujuan dari penelitian tersebut, yaitu untuk mempelajari parameter mutu buah naga menggunakan teknologi pengolahan citra digital, mencari korelasi antara data hasil pengolahan citra digital dengan data pengukuran terhadap buah secara langsung dan menentukan parameter mutu dari hasil pengolahan citra digital yang dapat digunakan untuk menentukan kualitas buah serta melakukan validasi kualitas buah menggunakan data yang didapat dari hasil pengolahan citra. Parameter yang dijadikan penentu kualitas dari buah naga, yaitu nilai rata-rata dan nilai standar deviasi yang didapat dari hasil pengolahan citra digital buah naga. Hasil penelitian menyimpulkan bahwa teknologi pengolahan citra digital dapat melakukan validasi kualitas buah naga dengan tingkat akurasi mencapai 82% untuk mutu buah A, 50% untuk mutu buah naga B dan 78% untuk mutu buah naga C.

(Harjoko & Wiraharja, 2014) dalam penelitiannya “Pemrosesan Citra Digital untuk Klasifikasi Mutu Buah Pisang Menggunakan Jaringan Saraf Tiruan”, tujuan dari penelitian tersebut yaitu membuat sebuah system yang dapat mengklasifikasikan mutu buah pisang menggunakan pengolahan citra digital dan jaringan saraf tiruan. Dalam penelitiannya penulis mengambil data berupa *energy*, *homogeneity*, dan *contrast* yang didapatkan dari hasil ekstraksi citra digital dari perhitungan luas kerusakan yang ada pada buah pisang. Data tersebut digunakan sebagai parameter pemutusan buah. Dalam penelitian penulis menyebutkan konfigurasi terbaik model jaringan *backpropagation* untuk sistem klasifikasi mutu pisang adalah dengan laju pembelajaran sebesar 0,3 dan jumlah *neuron* pada lapisan tersembunyi sebanyak 10 *neuron*. Dengan konfigurasi tersebut, sistem mampu mengklasifikasikan mutu dengan tingkat akurasi sebesar 94 % dari 100 data yang diuji.

(Sandra, 2007), dalam penelitiannya “Pengembangan Pemutuan Buah Manggis untuk Ekspor Secara *Non Destruktif* dengan Jaringan Syaraf Tiruan”, tujuan dari penelitian tersebut yaitu untuk membangun sistem pemutuan buah manggis secara *non destruktif* dengan jaringan syaraf tiruan. Parameter input yang digunakan untuk pemutuan buah manggis bagian dalam adalah hasil dari pengolahan citra digital dengan perangkat penunjang yaitu *video capture* Matrix Meteor dan untuk pemutuan manggis bagian dalam yaitu dengan gelombang ultrasonik dengan frekuensi 50kHz. Dari penelitian tersebut menunjukkan hasil sortasi menggunakan metode pengolahan citra digital memiliki tingkat akurasi sebesar 94,05% untuk manggis kelas super, 95,7% untuk manggis kelas I dan 100% untuk manggis yang tidak layak ekspor. Namun dalam

penelitian tersebut penulis menggunakan citra RGB tanpa merubah ke dalam citra *greyscale* karena akan berpengaruh terhadap akurasi dan efektivitas waktu pemrosesan data.

(Marshalina, 2012), dalam penelitiannya “Klasifikasi Buah Mangga Berdasarkan Bentuk dan Warna dengan Metode *Curvelet*”, tujuan penelitian ini adalah untuk mengembangkan suatu sistem berbasis *software* pengolahan citra digital yang dapat mengidentifikasi jenis mangga dengan mendeteksi citra mangga. Citra tersebut diekstraksi menggunakan *Curvelet* untuk menentukan ciri apakah yang dapat digunakan sebagai parameter acuan nantinya. Hasil dari penelitian tersebut adalah berupa nilai dari setiap skala koefisien yang bervariasi yang kemudian dianalisis untuk menentukan ciri apakah yang dapat digunakan untuk proses klasifikasi

(Khoje1, Bodhe, & Adsul3, 2013), dalam penelitiannya “*Automated Skin Defect Identification System for Fruit Grading Based on Discrete Curvelet Transform*” tujuan dari penelitian ini untuk mengembangkan metodologi untuk menentukan kualitas permukaan buah pada resolusi citra yang bervariasi dari yang paling rendah sampai yang paling tinggi menggunakan metode multi skala transformasi *curvelet*. Studi ini menyimpulkan bahwa berdasarkan transformasi *curvelet* memberikan wawasan menjanjikan untuk memperkirakan kerusakan kulit buah dengan resolusi citra yang bervariasi.

(Ranjit K N, Chethan H K, & Naveena C, 2016) dalam penelitiannya “*Identification and Classification of Fruit Diseases*” tujuan dari penelitian ini adalah untuk mengidentifikasi dan mengklasifikasikan buah terhadap penyakit yang

menyebabkan pengurangan terhadap hasil panen yang diperoleh dan menyebabkan memburuknya varietas hasil buah dari budidaya. Metode yang digunakan pada penelitian tersebut, yaitu menggunakan sistem kecerdasan buatan berbasis pengolahan citra, yaitu menggunakan algoritma *K-means* dan *C-Meansclustering*. Hasil dari penelitian tersebut menunjukkan bahwa mengklasifikasikan penyakit pada buah menggunakan segmentasi *K-means* memiliki akurasi yang relatif lebih tinggi dibanding menggunakan algoritma *C-Meansclustering*.

(Raharjo, 2016), dalam penelitiannya “Analisa Klasifikasi Pohon Mangga Berdasarkan Tekstur Daun”. Metode yang digunakan penulis dalam menentukan klasifikasi jenis pohon mangga berdasarkan tekstur daun yaitu dengan *Linear Discriminant Analysis (LDA)*. Hasil penelitian menunjukkan tingkat akurasi yang diperoleh dari *Linear Discriminant Analysis (LDA)* mencapai 94.17% dengan fitur *GLCM*.

(Novitasari, 2012), Dalam penelitian “Analisis Identifikasi Serviks Normal dan Abnormal Berdasarkan Filter Gabor dan Ekstraksi Ciri Tekstur Statistik”. Tujuan dilakukan penelitian tersebut adalah untuk mengklasifikasi serviks normal dan tidak normal. Dari serangkaian tahapan yang dilakukan, didapatkan hasil akhir berupa citra yang dikelompokkan berdasarkan warna tekstur dan angka yang kemudian dianalisis. Proses analisis tekstur citra serviks untuk menentukan jenis serviks tersebut apakah normal atau abnormal. Nilai ekstraksi ciri pada penelitian tersebut terdiri dari *energy*, *entropy*, *smoothness*, dan standar deviasi. Dari keempat nilai yang dianalisis, nilai *smoothness* dan standar deviasi memiliki hasil yang berbanding terbalik dengan teori.

Kesimpulan dalam penelitian tersebut adalah akurasi program dalam mengidentifikasi serviks normal dan abnormal dengan tingkat akurasi tertinggi sebesar 92.85%.

(Gozali, Galuh, Burhanuddin, & Niam, 2014), dalam penelitian” Aplikasi Kematangan Tomat Berdasarkan Warna dengan Metode *Linear Discriminant Analysis* (LDA)”. Tujuan diadakannya penelitian tersebut adalah untuk mengidentifikasi kematangan tomat dilakukan dengan melakukan ekstraksi warna dan dengan menggunakan algoritma LDA (*Linear Discriminant Analysis*), tomat dikelompokkan dalam dua kelas yang telah ditentukan yaitu kelas matang dan kelas mentah. Kesimpulan dari penelitian tersebut menyebutkan bahwa dengan metode LDA kita dapat menentukan buah tomat dalam kelas matang dan kelas mentah dan metode LDA sangat akurat dalam menentukan kelas matang atau kelas mentah untuk buah tomat.

(Firdaus, Prnama, & WIsesty, 2016), dalam penelitian “Klasifikasi Kendaraan di Jalan Tol dengan Menerapkan Metode *Local Binary Pattern* dan *Linear Discriminant Analysis*”. Tujuan dari penelitian tersebut adalah untuk mengklasifikasikan jenis kendaraan dengan sistem kecerdasan buatan dengan metode yang digunakan dalam penelitian tersebut adalah *Gaussian Mixture Model* sebagai metode untuk memisahkan *background* dan *foreground*, *Local Binary Pattern* sebagai metode ekstraksi ciri, dan *Linear Discriminant Analysis* (LDA) sebagai metode klasifikasi. Penelitian menyimpulkan bahwa LDA sebagai metode klasifikasi dapat memberikan performansi 82.33% pada kondisi optimum yaitu: ukuran citra 60x60 piksel, LBP radius 2, dan pemotongan citra 25 region.

2.2 Landasan Teori

2.2.1 Manggis

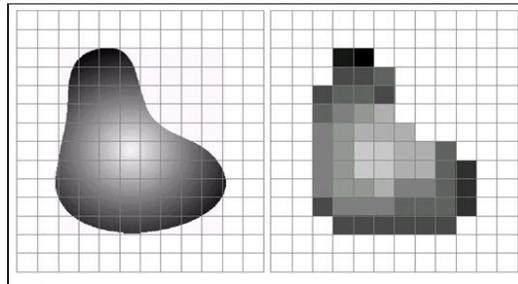
Manggis (*Garcinia mangostana* L.) merupakan tanaman buah yang berasal dari hutan tropis di kawasan Asia tenggara, yaitu kawasan hutan Malaysia dan Indonesia. Dari Asia Tenggara, tanaman ini menyebar ke daerah Amerika Tengah dan daerah tropis lainnya. Manggis adalah tanaman jangka panjang, manggis mampu berbuah ketika telah berumur 5 tahun (Kementrian Pertanian Republik Indonesia, 2012). Manggis menjadi buah komoditi ekspor Indonesia karena mendatangkan devisa tertinggi bagi Indonesia di antara buah-buahan lain. Selain tingginya permintaan dari manca negara, harga manggis yang relative tinggi, buah manggis dinobatkan sebagai komoditi ekspor unggulan (Reset Unggulan Strategis Nasional Buah, 2000).

Untuk kebutuhan ekspor, buah manggis dikelompokkan menjadi kelas Super, kelas I dan kelas II dengan standar kualitas FAO, ASEAN ataupun negara-negara pengimpor. Penentuan kelas manggis ditentukan dengan mengamati kualitas permukaan atau bagian luar manggis dan luasan kecacatan.

2.2.2 Pengertian Citra Digital

Citra adalah suatu representasi sebuah objek yang dihasilkan dari kombinasi sebuah titik, garis, bidang dan warna. Citra terbagi dua yaitu ada citra yang bersifat analog dan ada citra yang bersifat digital. Citra analog adalah citra-citra yang terbentuk dari sinyal kontinyu. Nilai intensitas cahaya pada citra analog memiliki range antara 0 s. d \sim . Alat akuisisi citra analog antara lain mata manusia dan kamera analog.

Sedangkan pada citra digital adalah citra yang terbentuk dari sinyal diskrit. Nilai intensitas cahaya pada citra digital bergantung pada ke dalam bit yang menyusunnya. Alat akuisisi citra digital antara lain yaitu kamera digital, kamera smartphone, mikroskop digital, scanner dan lain sebagainya.



Gambar 2. 1 Citra analog (kiri) dan citra digital (kanan)

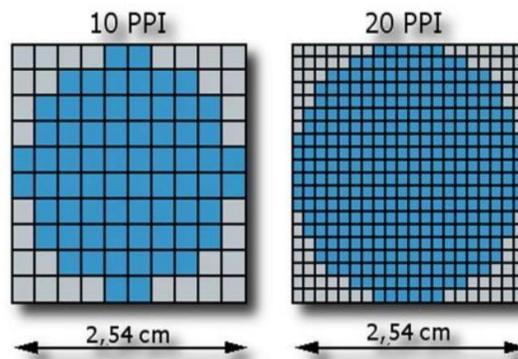
Citra digital direpresentasikan ke dalam sebuah matriks dimana indeks baris dan kolomnya menyatakan suatu titik pada citra tersebut dan elemen menyatakan tingkat keabuan pada titik tersebut. Citra $f(x, y)$ diproses oleh komputer dalam bentuk *array* $n \times m$.

$$f(x, y) \approx \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0,M-1) \\ f(1,0) & f(1,1) & \dots & f(1,M-1) \\ \vdots & \vdots & \vdots & \vdots \\ f(N-1,0) & f(N-1,1) & \dots & f(N-1,M-1) \end{bmatrix}$$

Gambar 2. 2 Citra digital dalam bentuk persamaan

Setiap elemen dari *array* disebut sebagai piksel dari sebuah citra dengan ukuran tertentu dan menunjukkan nilai intensitas keabuan piksel pada daerah yang tersebut.

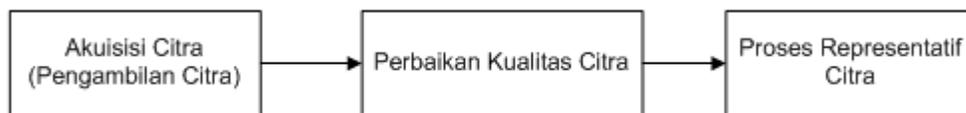
Piksel merupakan bagian terkecil dari sebuah citra digital berupa persegi-persegi bangun dasar yang berukuran kecil yang dihitung per inch setiap piksel hanya memiliki satu warna dan menempati sebuah tempat spesifik pada sebuah citra. Piksel sendiri berasal dari akronim bahasa Inggris *Picture Element* yang disingkat menjadi *Pixel*. Pada skala resolusi tertinggi, mesin cetak gambar berwarna dapat menghasilkan hasil cetak yang memiliki lebih dari 2.500 titik per inci dengan pilihan 16 juta warna lebih untuk setiap inci, dalam istilah komputer berarti gambar seluas satu inci persegi yang bisa ditampilkan pada tingkat resolusi tersebut sepadan dengan 150 juta bit informasi.



Gambar 2.3 Ilustrasi piksel dengan kerapatan 10 PPI(kanan) dan 20 PPI(kiri)

2.2.3 Pengertian Pengolahan Citra Digital

Pengolahan citra digital adalah sebuah manipulasi element-element matriks yang dihasilkan dari representasi sebuah gambar ke dalam bentuk matriks dengan bantuan computer dan software pengolahan citra digital. Pengolahan citra digital dapat dikelompokkan dalam dua jenis kegiatan, yaitu memperbaiki kualitas suatu gambar, sehingga dapat lebih mudah diinterpretasikan oleh mata manusia dan mengolah informasi yang terdapat pada suatu gambar untuk keperluan pengenalan objek secara otomatis.



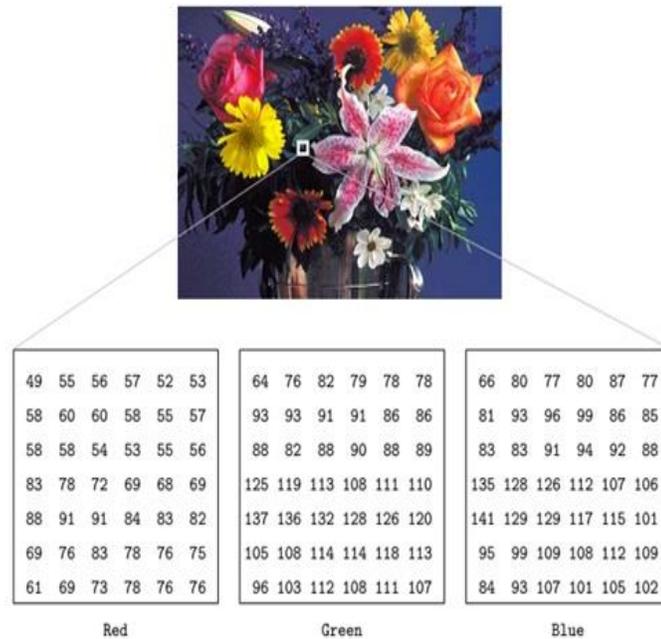
Gambar 2. 4 Proses Pengolahan citra untuk memperbaiki kualitas citra

Teknik-teknik pengolahan citra meliputi penajaman citra, penonjolan fitur tertentu dari suatu citra, kompresi citra dan perbaikan citra yang kabur (Basuki, 2005). Dalam pengambilan citra hanya citra digital yang dapat diproses oleh komputer digital, data citra yang dimasukkan berupa nilai-nilai angka yang menunjukkan nilai intensitas cahaya atau tingkat keabuan setiap piksel (Basuki, 2005).

2.2.4 Citra RGB

Citra RGB adalah sebuah gambar yang memiliki warna dasar, yaitu merah, hijau dan biru yang ditambahkan dengan berbagai cara untuk menghasilkan bermacam-macam warna.

Jika diproses ke dalam sebuah matriks citra RGB memiliki tiga matriks, yaitu matriks Red (warna merah), matriks Blue (warna biru) dan matriks Green (warna hijau) untuk merepresentasikan intensitas setiap warna pada daerah tertentu pada gambar sehingga menghasilkan beragam macam warna.



Gambar 2.5 Ilustrasi Citra RGB

2.2.5 Citra *Greyscale*

Citra grayscale merupakan citra digital yang hanya memiliki satu nilai kanal pada setiap pikselnya, dengan kata lain bagian *red* = *green* = *blue*. Nilai tersebut digunakan untuk menunjukkan tingkat intensitas. Warna yang dimiliki adalah warna dari hitam, keabuan dan putih. Tingkatan keabuan di sini merupakan warna abu dengan berbagai tingkatan dari hitam hingga mendekati putih (Putra, 2010). Pada citra grayscale warna hitam pada bagian ini memiliki nilai terendah, dan nilai warna putih

menjadi nilai terkuat, atau nilai warna putih lebih tinggi dari nilai warna hitam. Citra *grayscale* berbeda dengan citra biner atau citra “hitam putih”. Apabila citra hitam putih hanya mengenal dua warna yaitu “hitam” dan warna “putih” sedangkan *grayscale* mempunyai variasi yang banyak, karena ada nilai-nilai diantara nilai minimum (biasanya = 0) dan nilai maksimum. Banyaknya kemungkinan nilai minimum dan nilai maksimumnya bergantung pada jumlah bit yang digunakan.

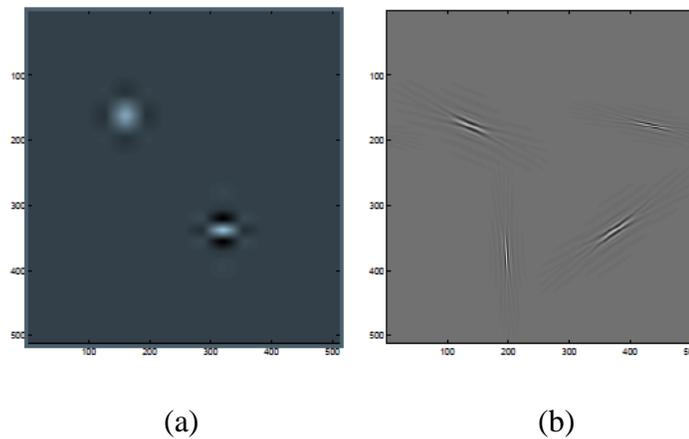
2.2.6 Transformasi *Curvelet*

Transformasi *curvelet* merupakan teknologi yang dapat memisahkan *noise* dari sinyal ke dalam dimensi frekuensi, *dip*, *azimuth* dan lokasi. Karena kelebihan ini transformasi *curvelet* dikenal dengan transformasi multi dimensi yang tidak dimiliki oleh teknologi transformasi lain semisal *Wavelet*. *Curvelet* banyak dimanfaatkan untuk keperluan diantaranya *denoising image*.



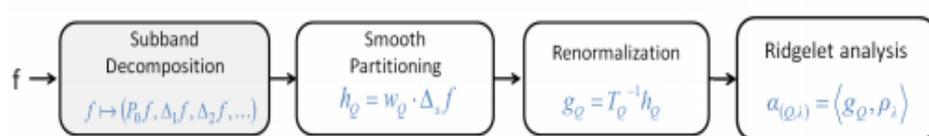
Gambar 2. 6 Aplikasi *Curvelet* pada *Image Processing* sebelum (kiri) dan setelah *denoising* (kanan)

Transformasi *curvelet* juga disebut transformasi multi skala terarah yang mampu penampilan *sparsa nonadaptive* secara optimal dari sebuah obyek yang memiliki banyak tepian. Transformasi *curvelet* dikembangkan sebagai salah satu solusi terhadap kelemahan yang dimiliki *wavelet*. Transformasi *Curvelet* ditemukan dalam riset untuk tujuan mengatasi keterbatasan transformasi *wavelet* pada tahun 1999 oleh Candes dan Donoho.

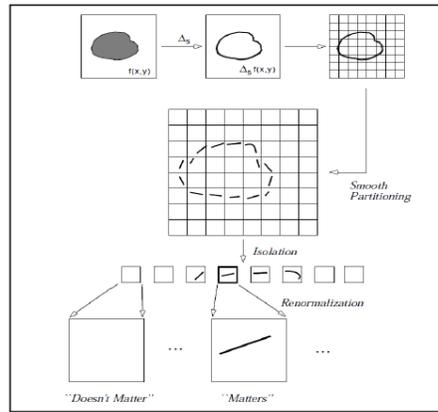


Gambar 2. 7 Hasil *Wavelet* (kiri) dan *Curvelet* (kanan)

Transformasi *Curvelet* dapat diuraikan (dekomposisi) dalam empat langkah: *Subband Decomposition*, *Smooth Partitioning*, *Renormalization*, *Ridgelet Analysis*. Dengan membalik urutan (*inversi*) secara matematis dapat diperoleh kembali (*rekonstruksi*) sinyal awalnya, disebut dengan transformasi *inversi*.



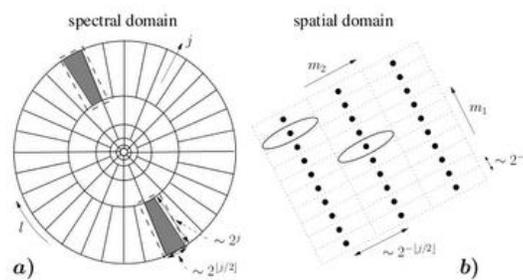
Gambar 2. 8 Tahapan dekomposisi *curvelet* I



Gambar 2. 9 Tahapan dekomposisi *curvelet* II

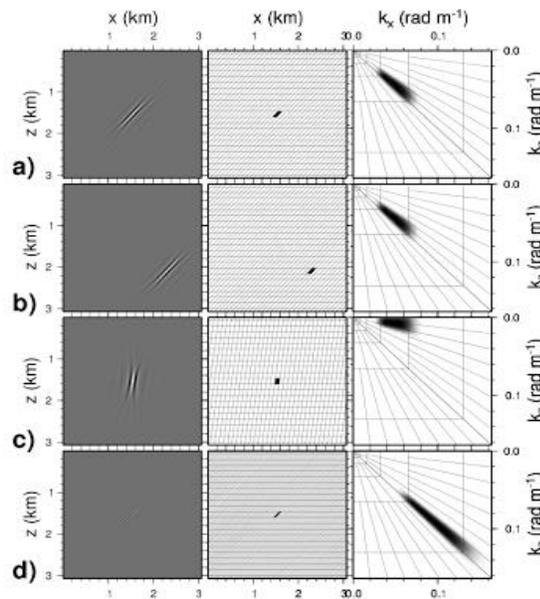
Tahapan melakukan *denoising* dengan transformasi *curvelet* adalah sebagai berikut:

1. melakukan FFT 2D pada data seismik.
2. melakukan *windowing* dari hasil pertama dengan Polar Grid dalam domain-domain frekuensi dengan demikian diperoleh *Curvelet Coefficien*.
3. melakukan Penskalaan (mendenoise) dari *Coefficien curvelet* tersebut, melakukan inversi untuk memperoleh kembali data seismik yang telah di-*denoise*.



Gambar 2. 10 Polar grid Curvelet dalam domain frekuensi (a) dan domain *spatial* (b) (Courtesy Douma and de Hoop)

Gambar di bawah ini adalah *curvelet* dalam domain *spatial* (kolom pertama) dan amplitude *spectra* pada polar grid (kolom ketiga) untuk masing-masing *curvelet*. Perhatikan bahwa *curvelet* untuk masing-masing lokasi polar grid memiliki ukuran(*size*), arah(*azimuth*), lokasi dan kemiringan tertentu.



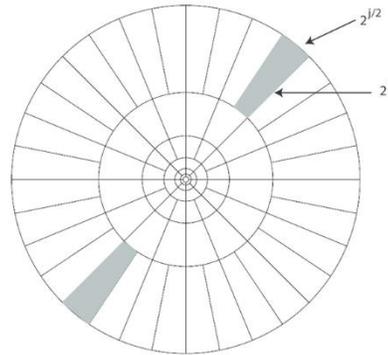
Gambar 2. 11 Curvelet dan Amplitude Spectra pada Polar Grid (Courtesy Douma and de Hoop)

Transformasi *curvelet* memiliki dua bentuk antara lain *Continuous Curvelet Transforms* dan *Discrete Curvelet Transform*.

1. *Continuous Curvelet Transforms*

Pada *Continuous Curvelet Transform* melakukan transformasi dengan cara membagi citra pada domain frekuensi sepanjang sudut radial melingkar menggunakan bagian-bagian jendela. *Continuous Curvelet Transform* bisa

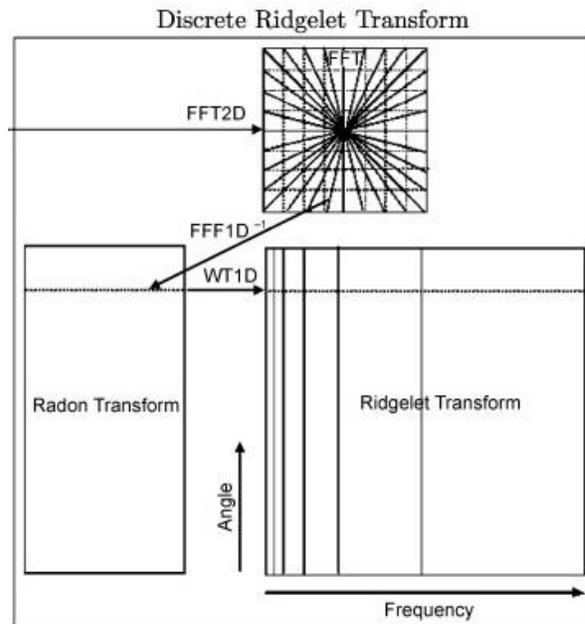
disebut curvelet versi awal. *Continuous Curvelet Transform* Ditunjukkan pada gambar 3. 13



Gambar 2. 12 Ilustrasi *Continuous Curvelet Transform*

2. *Discrete Curvelet Transform*

Discrete Curvelet Transform adalah pengembangan dari versi sebelumnya, yaitu *Continuous Curvelet Transform*. Pada *Discrete Curvelet Transform* melakukan transformasi dengan cara membagi gambar menggunakan kotak tengah yang ada pada gambar. Transformasi curvelet akan menghasilkan informasi yang berlebih (*redundant information*) sehingga dapat merepresentasikan sinyal yang berada pada tepian kurva gambar sehingga transformasi curvelet didesain ulang kemudian dan diperkenalkan sebagai *Fast Discrete Curvelet Transform (FDCT)*.



Gambar 2. 13 Ilustrasi *discrete curvelet transform*

Transformasi *curvelet discrete* dapat direpresentasikan dengan persamaan berikut:

$$C^D(j, l, k) = \sum_{0 \leq t_1, t_2 < n} f[t_1, t_2] \phi_{j,k,l}^D[t_1, t_2] \dots \dots \dots (2. 1)$$

Dengan memanfaatkan teknologi transformasi *curvelet* diharapkan hasil ekstraksi ciri akan memiliki tingkat akurasi yang baik karena telah dipisahkan sinyal dalam dimensi frekuensi dari noise yang ada pada citra yang akan diproses pada tahapan selanjutnya.

2.2.7 Ekstraksi Ciri

Ekstraksi ciri merupakan suatu proses pengambilan ciri unik dari suatu objek ke dalam sebuah nilai tertentu. Nilai yang didapatkan akan dianalisis dan digunakan sebagai bahan analisis. Untuk mendapatkan nilai dari suatu ciri dilakukan dengan cara menghitung jumlah titik atau piksel yang ditemui dalam setiap pengecekan, dimana pengecekan dilakukan dalam berbagai arah pada koordinat Cartesian dari citra digital yang dianalisis, yaitu vertikal, horizontal, diagonal kiri, dan diagonal kanan. Ciri citra dihitung menggunakan rumus sebagai berikut:

1. Rata-rata

Nilai rata-rata (*mean*) didapatkan dari hasil perhitungan dengan cara menjumlahkan nilai setiap elemen matriks mulai dari elemen matriks ke-1 hingga elemen matriks ke-N kemudian dibagi dengan banyak elemen matriks yang ada, Persamaan untuk menghitung nilai Rata-Rata adalah sebagai berikut:

$$\mu = \frac{1}{N} \sum_{i=1}^N X_i \dots\dots\dots (2. 2)$$

2. Standar Deviasi

Standar deviasi menunjukkan nilai rata-rata kontras dari intensitas citra, (Marshallina, 2012). Nilai standar deviasi (*standard deviation*) didapatkan dari hasil perhitungan dengan cara mengurangkan nilai setiap elemen matriks ke-1 hingga elemen matriks ke-N dengan nilai rata-rata setiap elemen matriks. Hasil pengurangan kemudian dikuadratkan dan dijumlahkan pada setiap elemen ke-1 sampai ke-N dan kemudian diakar kuadratkan untuk mendapatkan nilai standar

deviasi. Persamaan untuk menghitung nilai Standar Deviasi adalah sebagai berikut:

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2} \dots\dots\dots (2. 3)$$

3. *Entropy*

Nilai *entropy* menunjukkan ketidakteraturan distribusi derajat keabuan suatu citra. Nilai *entropy* akan terlihat besar untuk citra dengan transisi derajat keabuan merata dan bernilai kecil jika struktur citra tidak teratur, (Marshalina, 2012). Persamaan untuk menghitung nilai *entropy* adalah sebagai berikut:

$$H = - \sum_{i=1}^G p(d_i) \cdot \log_2 p(d_i) \dots\dots\dots (2. 4)$$

4. *Energy*

Energy mengacu pada homogenitas dari tekstur, jika nilai energi tinggi maka jumlah area yang homogen besar, dan sebaliknya jika nilai energi rendah maka jumlah area yang homogen sedikit atau kecil, (Marshalina, 2012). Nilai *energy* juga dapat dikatakan *inverse* dari nilai *entropy*. , dimana semakin tinggi nilai *entropy* maka nilai *energy* akan semakin rendah. Hal ini dikarenakan, nilai *energy* menunjukkan keteraturan distribusi derajat keabuan suatu citra. Persamaan untuk menghitung nilai *energy* adalah sebagai berikut:

$$E_k = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N |X_k(i, j)| \dots\dots\dots (2. 5)$$

Nilai hasil ekstraksi ciri yang didapat, bisa digunakan sebagai parameter acuan dalam proses klasifikasi data.

2.2.8 *Linear Discriminant Analysis (LDA)*

Analisis diskriminan adalah salah satu teknik statistik yang bisa digunakan untuk mengklasifikasi sebuah objek ke dalam satu atau lebih kelompok berdasarkan ciri tertentu. Analisis diskriminan bermanfaat pada situasi dimana sampel total dapat dikelompokkan menjadi group-group berdasarkan karakteristik variabel yang didapat dari sebuah perhitungan matematis. Tujuan utama dari analisis diskriminan adalah untuk mengklasifikasi sampel uji ke dalam kelompok tertentu berdasarkan ciri yang telah diambil. Metode analisis diskriminan yang sering digunakan sebagai sistem pembelajaran mesin adalah *Linear Discriminant Analysis (LDA)*.

Linear Discriminant Analysis (LDA) adalah sebuah metode yang digunakan dalam statistik pembelajaran mesin untuk mengetahui kombinasi *linear* dari ciri yang menjadi pemisahan dua atau lebih kelompok sampel. Kombinasi yang dihasilkan dapat digunakan sebagai pengklasifikasian *linier*, atau, lebih umum lagi, untuk pengurangan dimensi sebelum klasifikasi selanjutnya.

Banyak sekali tujuan dari metode *Linear Discriminant Analysis (LDA)* antara lain sebagai berikut:

1. Mengetahui apakah ada perbedaan yang jelas antar kelompok.
2. Melakukan klasifikasi terhadap objek dan untuk mengetahui apakah suatu objek termasuk pada grup 1 atau grup 2 atau lainnya.
3. Menentukan tingkat akurasi dari sebuah sistem dalam mengidentifikasi objek berdasarkan ciri tertentu.

Implementasi *Linear Discriminant Analysis (LDA)* ke dalam sebuah persamaan adalah sebagai berikut:

$$Y = b_0 + b_1X_1 + b_2X_2 + \dots + b_nX_n \dots\dots\dots (2. 6)$$

Keterangan:

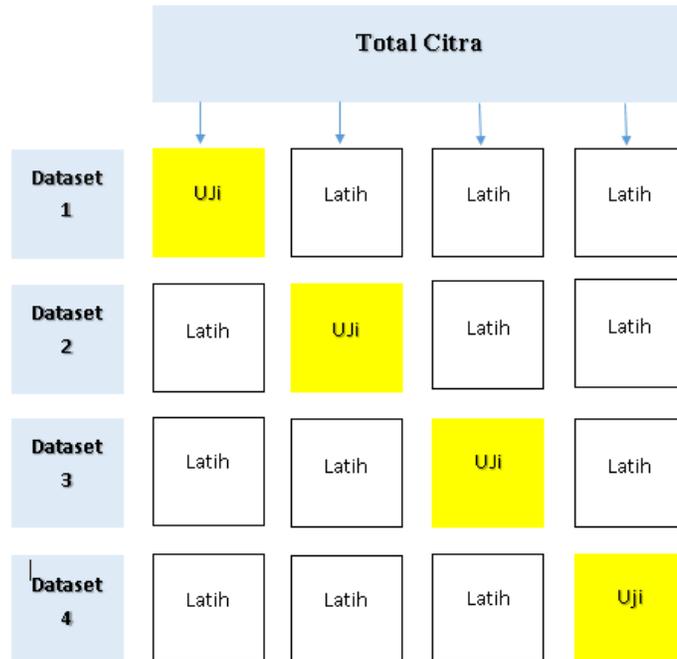
Y = nilai (skor) diskriminan dan merupakan variabel terikat.

X_n = variabel (atribut) ke-n dan merupakan variabel bebas

b_n = koefisien diskriminan/bobot dari variabel (atribut) ke-n

2.2.9 K-Fold Cross Validation

Cross Validation adalah salah satu metode yang digunakan untuk memvalidasi keakuratan sebuah model yang dibangun berdasarkan dataset tertentu. Pembuatan model biasanya bertujuan untuk melakukan klasifikasi terhadap suatu data baru yang belum pernah muncul di dalam dataset. Data yang digunakan dalam proses pembangunan data model disebut data *training*, sedangkan data yang akan digunakan untuk validasi model disebut dengan data *testing*. Salah satu metode *cross-validation* yang populer adalah *K-Fold Cross Validation*. Dalam teknik ini dataset dibagi menjadi sejumlah K-buah partisi secara acak. Kemudian dilakukan sejumlah K-kali eksperimen, dimana masing-masing eksperimen menggunakan data partisi ke-K sebagai data testing dan memanfaatkan sisa partisi lainnya sebagai data training (Muafiq, 2016).



Gambar 2. 4 Gambaran metode *K-fold Cross Validation*