BAB 2
LANDASAN TEORI

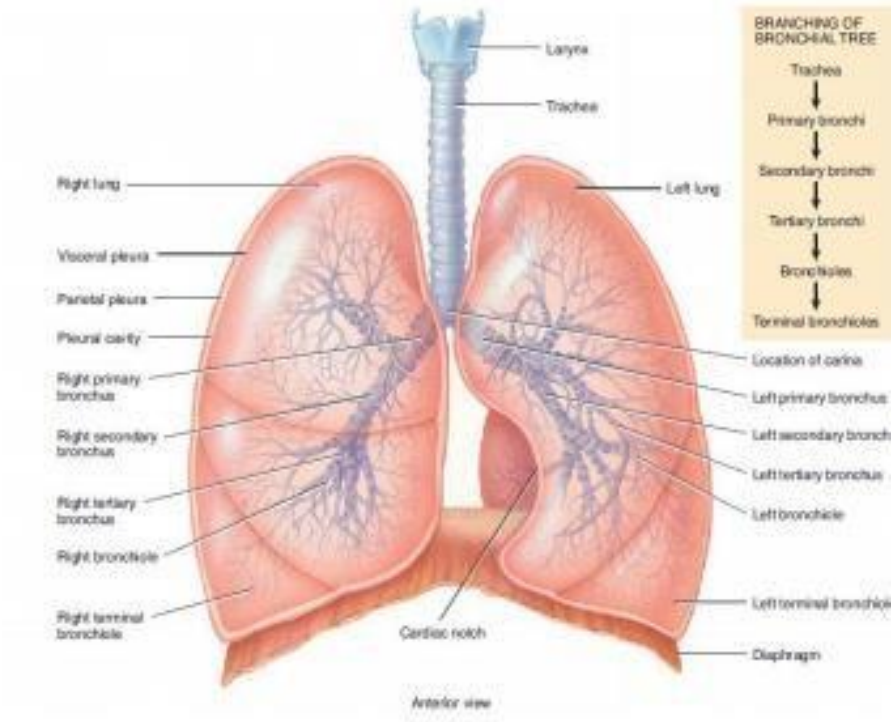
2.1 Paru - Paru

2.1.1 Anatomi Paru

Paru-paru terletak pada rongga dada, berbentuk kerucut yang ujungnya berada di atas tulang iga pertama dan dasarnya berada pada diafragma. Paru terbagi menjadi dua yaitu, paru kanan dan paru kiri. Paru-paru kanan mempunyai tiga lobus sedangkan paruparu kiri mempunyai dua lobus. Kelima lobus tersebut dapat terlihat dengan jelas. Setiap paru-paru terbagi lagi menjadi beberapa subbagian menjadi sekitar sepuluh unit terkecil yang disebut bronchopulmonary segments. Paru-paru kanan dan kiri dipisahkan oleh ruang yang disebut mediastinum.

Paru-paru dibungkus oleh selaput tipis yaitu pleura. Pleura terbagi menjadi pleura viseralis dan pleura pariental. Pleura viseralis yaitu selaput yang langsung membungkus paru, sedangkan pleura parietal yaitu selaput yang menempel pada rongga dada. Diantara 11 kedua pleura terdapat rongga yang disebut kavum pleura.

Paru manusia terbentuk setelah embrio mempunyai panjang 3 mm. Pembentukan paru di mulai dari sebuah Groove yang berasal dari Foregut. Pada Groove terbentuk dua kantung yang dilapisi oleh suatu jaringan yang disebut Primary Lung Bud. Bagian proksimal foregut membagi diri menjadi 2 yaitu esophagus dan trakea. Pada perkembangan selanjutnya trakea akan bergabung dengan primary lung bud. Primary lung bud merupakan cikal bakal bronchi dan cabang-cabangnya. Bronchial-tree terbentuk setelah embrio berumur 16 minggu, sedangkan alveoli baru berkembang setelah bayi lahir dan jumlahnya terus meningkat hingga anak berumur 8 tahun. Alveoli bertambah besar sesuai dengan perkembangan dinding toraks. Jadi, pertumbuhan dan perkembangan paru berjalan terus menerus tanpa terputus sampai pertumbuhan somatic berhenti.



Gambar 2. 1 Anatomi Paru

Sitem pernafasan dapat dibagi ke dalam sitem pernafasan bagian atas dan pernafasan bagian bawah.

1. Pernafasan bagian atas meliputi, hidung, rongga hidung, sinus paranasal, dan faring.
2. Pernafasan bagian bawah meliputi, laring, trakea, *bronchus*, bronchioles dan alveolus paru Pergerakan dari dalam ke luar paru terdiri dari dua proses, yaitu inspirasi dan ekspirasi. Inspirasi adalah pergerakan dari atmosfer ke dalam paru, sedangkan ekspirasi adalah pergerakan dari dalam paru ke atmosfer.

Agar proses ventilasi dapat berjalan lancar dibutuhkan fungsi yang baik pada otot pernafasan dan elastisitas jaringan paru. Otot-otot pernafasan dibagi menjadi dua yaitu,

1. Otot inspirasi yang terdiri atas, otot interkostalis eksterna, sternokleidomastoideus, skalenus dan diafragma.
2. Otot-otot ekspirasi adalah rektus abdominis dan interkostalis internus [16].

2.2 Pengolahan Citra (*Image Processing*)

Untuk melakukan pengenalan suatu objek berdasarkan pola yang dimiliki objek tersebut maka dibutuhkan teknik untuk melakukannya, teknik tersebut adalah teknik pengolahan citra. Pengolahan citra merupakan proses pengolahan dan analisis citra yang banyak melibatkan persepsi visual. Proses ini mempunyai ciri data masukan dan informasi keluaran yang berbentuk citra. Istilah pengolahan citra digital secara umum didefinisikan sebagai pemrosesan citra dua dimensi dengan computer. Dalam definisi yang lebih luas, pengolahan citra digital juga mencakup semua data dua dimensi. Citra digital adalah barisan bilangan nyata maupun kompleks yang diwakili oleh bit bit tertentu.

Citra merupakan istilah lain dari gambar yang merupakan komponen multimedia yang memegang peranan sangat penting sebagai bentuk informasi visual. Citra mempunyai karakteristik yang tidak dimiliki oleh data teks, yaitu kaya akan informasi. Citra digital adalah citra hasil digitalisasi citra kontinu (analaog). Tujuan dibuatnya citra digital adalah agar citra tersebut dapat diolah menggunakan computer atau piranti digital dan memperbaiki kualitas citra agar mudah diinterpertasi oleh manusia atau mesin (*computer*). Teknik teknik pengolahan citra mentransformasikan citra menjadi citra lain yang mempunyai kualitas lebih baik.

2.3 Langkah-langkah Penting dalam Pengolahan Citra (*image processing*)

Dalam pengolahan citra terdapat beberapa langkah penting yang harus dilakukan. Secara umum, langkah-langkah dalam pengolahan citra dapat di jabarkan menjadi beberapa langkah yaitu [3] :

1. Akuisisi Citra

Akuisisi citra adalah tahap awal untuk mendapatkan citra digital. Tujuan akuisisi citra adalah untuk menentukan data yang diperlukan dan memilih metode perekaman citra digital. Tahap ini dimulai dari objek yang akan diambil gambarnya, persiapan alat-alat, sampai pada pencitraan. Pencitraan adalah kegiatan transformasi dari citra tampak (foto, lukisan, gambar, patung, pemandangaan dan lain-lain) menjadi citra digital. Beberapa alat yang dapat digunakan untuk pencitraan adalah :

a. Video kamera

b. Kamera digital

c. Kamera konvesional dan converter analog to digital

d. *Scanner*

e. Photo sinar-x atau sinar infra merah

2. Pre Processing

Tahapan ini diperlukan untuk menjamin kelancaran pada proses berikutnya. Hal-hal penting yang dilakukan pada tingkatan ini diantaranya adalah:

a. Peningkatan kualitas citra (kontras, *brightness*, dan lain-lain)

b. Menghilangkan *noise*

c. Perbaikan citra (*image restoration*)

d. Transformasi (*image transformation*)

e. Menentukan bagian citra yang akan diobservasi

3. Segmentasi

   Tahapan ini bertujuan untuk mempartisi citra menjadi bagian-bagian pokok yang mengandung informasi penting. Misalnya, memisahkan objek dari latar belakang.

4. Representasi dan Deskripsi

   Dalam hal ini representasi merupakan suatu proses untuk merepresentasikan suatu wilayah sebagai suatu daftar titik-titik koordinat dalam kurva yang tertutup, dengan deskripsi luasan atau perimeternya. Setelah suatu wilayah dapat direpresentasikan, proses selanjutnya adalah melakukan deskripsi citra dengan cara seleksi ciri dan ekstraksi ciri (*Feature Extraction and selection*). Seleksi ciri 9 bertujuan untuk memilih informasi kuantitatif dari ciri yang ada, yang dapat membedakan kelas-kelas objek secara baik, sedangkan ekstraksi ciri bertujuan untuk mengukur besaran kuantitatif ciri setiap piksel, misalnya rata-rata, standar deviasi, koefisien variasi, *signal to nois ratio* (SNR), dan lain-lain.

5. Pengenalan dan Interpretasi

   Tahap pengenalan bertujuan untuk memberi label pada suatu objek yang informasinya disediakan oleh descriptor, sedangkan tahap interpretasi bertujuan untuk memberi arti atau makna kepada kelompok objek-objek yang dikenali.

6. Basis Pengetahuan

   Tahap pengenalan bertujuan untuk memberi label pada suatu objek yang informasinya disediakan oleh descriptor, sedangkan tahap

interpretasi bertujuan untuk memberi arti atau makna kepada kelompok objek-objek yang dikenali.

2.4 Elemen Citra

Dalam pengolahan citra objek yang digunakan adalah sebuah citra dari sebuah objek tertentu yang mana citra tersebut mengandung sejumlah elemen dasar. Elemen dasar tersebut di manipulasi dalam pengolahan citra, elemen tersebut adalah [3]:

1. Warna

   Warna adalah persepsi yang dirasakan oleh sistem visual manusia terhadap panjang gelombang cahaya yang dipantulkan oleh objek. Setiap warna mempunyai panjang gelombang. Warna yang diterima oleh mata merupakan hasil kombinasi cahaya dengan panjang gelombang berbeda. Kombinasi warna yang memberikan rentang warna yang paling lebar adalah *red* (R), *green* (G), *blue* (B).

2. Kecerahan (*Brightness*)

   Kecerahan disebut juga intensitas cahaya. Kecerahan pada sebuah piksel (titik) didalam citra bukanlah intensitas yang rell, tetapi sebenarnya adalah intensitas ratarata dari suatu area yang melingkupinya.

3. Kontras (*Contrast*)

   Kontras menyatakan sebaran terang dan gelap di dalam sebuah gambar. Citra dengan kontras rendah dicirikan oleh sebagian besar komposisi citranya adalah terang atau sebagian besar gelap. Pada citra dengan kontras yang baik, komposisi gelap dan terang tersebar secara merata.

4. Kontur (*Contour*)

Kontur adalah keadaan yang ditimbulkan oleh perubahan intensitas pada piksel yang bertetangga. Karena adanya perubahan intensitas, mata manusia dapat mendeteksi tepi objek didalam citra.

5. Bentuk (*Shape*)

   Bentuk adalah properti intrinsik dari objek tiga dimensi, dengan pengertian bahwa shape merupakan properti intrinsik utama untuk sistem visual manusia. Pada umumnya citra yang dibentuk oleh mata merupakan citra dwimatra (dua dimensi), sedangkan objek yang dilihat umumnya berbentuk trimatra (tiga dimensi). Informasi bentuk objek dapat diekstraksi dari citra pada permulaan prapengolahan dan segmentasi citra.

6. Tekstur (*Texture*)

   Tekstur diartikan sebagai distribusi spasial dari derajat keabuan di dalam sekumpulan piksel-piksel yang bertetangga. Jadi tekstur tidak dapat didefinisikan untuk sebuah piksel. Sistem visual manusia menerima informasi citra sebagai suatu kesatuan. Resolusi citra yang diamati ditentukan oleh skala dimana tekstur tersebut dipersepsi.

7. Waktu dan Pergerakan

   Respon suatu sistem visual tidak hanya berlaku pada faktor ruang, tetapi juga pada faktor waktu. Sebagai contoh, bila citra diam ditampilkan secara cepat, akan berkesan melihat citra yang bergerak.

8. Deteksi dan Pengenalan

   Dalam mendeteksi dan mengenali suatu citra, ternyata tidak hanya sistem visual manusia saja yang bekerja, tetapi juga ikut melibatkan ingatan dan daya pikir manusia.

2.5 Citra Analog dan Citra Digital

Secara umum terdapat 2 jenis citra yaiu citra analog dan citra digital. Citra analog adalah citra yang bersifat kontinu, seperti gambar pada monitor televisi, foto sinar-X, foto yang tercetak dikertas foto, lukisan, pemandangan, hasil CT scan, gambar-gambar yang terekam pada pita kaset, dan lain-lain sebagainya.

Citra analog tidak dapat direpresentasikan dalam komputer sehingga tidak dapat diproses dikomputer secara langsung. Oleh sebab itu, agar citra ini dapat diproses dikomputer, proses konversi analog ke digital harus dilakukan terlebih dahulu. Citra analog dihasilkan dari alat-alat analog, video kamera analog, kamera foto analog, Web Cam, CT scan, sensor ultrasound pada system USG, dan lainlain.

Citra Digital adalah citra yang dapat diolah oleh komputer dan citra digital yaitu gambar pada bidang dua dimensi. Dalam tinjauan matematis, citra merupakan fungsi kontinu dari intensitas cahaya pada bidang dua dimensi. Ketika sumber cahaya menerangi objek, objek memantulkan kembali sebagian cahaya tersebut. Pantulan ini ditangkap oleh alat-alat pengindera optik, misalnya mata manusia, kamera, scanner dan sebagainya. Bayangan objek tersebut akan terekam sesuai intensitas pantulan cahaya. Ketika alat optik yang merekam pantulan cahaya itu merupakan mesin digital, misalnya kamera digital, maka citra yang dihasilkan merupakan citra digital. Pada citra digital, kontinuitas intensitas cahaya dikuantisasi sesuai resolusi alat perekam [6].

2.6 Pola

Pola adalah entitas yang terdefinisi dan dapat diidentifikasi melalui ciri-cirinya (*features*). Ciri-ciri tersebut digunakan untuk membedakan suatu pola dengan pola lainnya. Ciri yang bagus adalah ciri yang memiliki daya pembeda yang tinggi, sehingga pengelompokan pola berdasarkan ciri yang dimiliki dapat dilakukan dengan keakuratan yang tinggi. Sebagai contoh,

Tabel 2. 1 Ciri Ciri Pola

Pola	Ciri
huruf	tinggi, tebal, titik sudut, lengkungan garis, dll
suara	amplitudo, frekuensi, nada, intonasi, warna, dll
tanda tangan	panjang, kerumitan, tekanan, dll
sidik jari	lengkungan, jumlah garis, dll

Ciri pada suatu pola diperoleh dari hasil pengukuran terhadap objek uji. Khusus pada pola yang terdapat di dalam citra, ciri-ciri yang dapat diperoleh berasal dari informasi wilayah/bentuk, keliling, luas, pusat massa, dst.

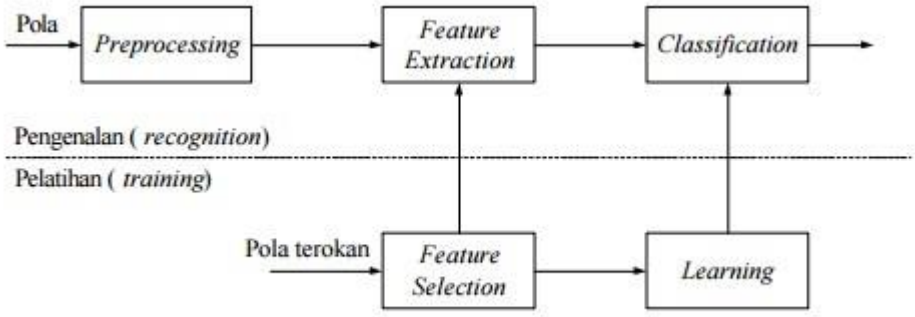
2.6.1 Sistem Pengenalan Pola

Pengenalan pola bertujuan menentukan kelompok atau kategori pola berdasarkan ciri-ciri yang dimiliki oleh pola tersebut. Dengan kata lain, pengenalan pola membedakan suatu objek dengan objek lain. Terdapat dua pendekatan yang dilakukan dalam pengenalan pola: pendekatan secara statistik dan pendekatan secara sintaktik atau struktural.

Pengenalan Pola secara Statistik Pendekatan ini menggunakan teori-teori ilmu peluang dan statistik. Ciri-ciri yang dimiliki oleh suatu pola ditentukan distribusi statistiknya. Pola yang berbeda memiliki distribusi yang berbeda pula. Dengan menggunakan teori keputusan di dalam statistik, menggunakan distribusi ciri untuk mengklasifikasikan pola. Contoh teori keputusan: Misalkan ada N pola yang dikenali, yaitu w1, w2, …, wN dan fungsi peluang atau kerapatan dari ciri-ciri pada pola diketahui. Jika x merupakan hasil pengukuran ciri-ciri, maka dapat dihitung.

$$p(\vec{x}|w_i) \qquad , i = 1, 2, \ldots, N \qquad \ldots (2.1)$$

Sistem pengenalan pola dengan pendekatan statistik ditunjukkkan oleh diagram pada Gambar 2.2.



Gambar 2. 2 Sistem Pengenalan Pola Dengan Pendekatan Statistic

Ada dua fase dalam sistem pengenalan pola:

(i) fase pelatihan dan

(ii) fase pengenalan. Pada fase pelatihan, beberapa contoh citra dipelajari untuk menentukan ciri yang akan digunakan dalam proses pengenalan serta prosedur klasifikasinya. Pada fase pengenalan, citra diambil cirinya kemudian ditentukan kelas kelompoknya.

2.6.1.1 *Preprocessing*

Proses awal yang dilakukan untuk memperbaiki kualitas citra (edge enhancement) dengan menggunakan teknik-teknik pengolahan citra yang sudah diejelaskan pada babbab sebelum ini.

2.6.1.2 *Feature Extraction*

Proses mengambil ciri-ciri yang terdapat pada objek di dalam citra. Pada proses ini objek di dalam citra mungkin perlu dideteksi seluruh tepinya, lalu menghitung propertiproperti objek yang berkaitan sebagai ciri. Beberapa proses ekstraksi ciri mungkin perlu mengubah citra masukan sebagai citra biner, melakukan penipisan pola, dan sebagainya.

2.6.1.3 *Classification*

Proses mengelompokkan objek ke dalam kelas yang sesuai.

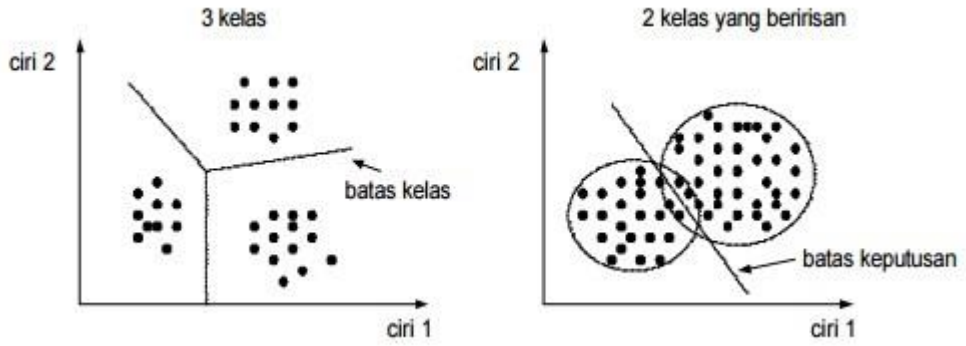
2.6.1.4 *Feature Selection*

Proses memilih ciri pada suatu objek agar diperoleh ciri yang optimum, yaitu ciri yang dapat digunakan untuk membedakan suatu objek dengan objek lainnya.

2.6.1.5 *Learning*

Proses belajar membuat aturan klasifikasi sehingga jumlah kelas yang tumpang tindih dibuat sekecil mungkin.

Kumpulan ciri dari suatu pola dinyatakan sebagai vektor ciri dalam ruang multi dimensi. Jadi, setiap pola dinyatakan sebagai sebuah titik dalam ruang. Ruang dibagi menjadi sejumlah sub-ruang. Tiap sub-ruang dibentuk berdasarkan pola-pola yang sudah dikenali kategori dan ciri-cirinya (melalui fase pelatihan). Lihat Gambar 2.5.



Gambar 2. 3 Contoh Pembagian Kelas Poin

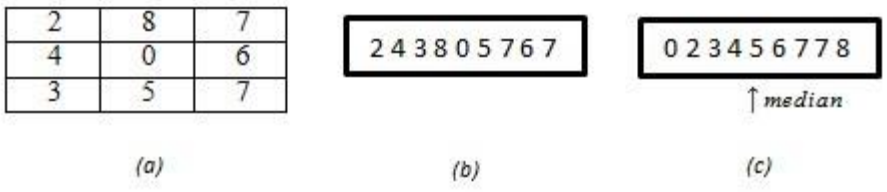
2.7. Median Filter

Konsep dasarnya adalah den gan menemukan nilai pixel yang memiliki nilai intensitas dari suatu pixel yang berbeda dengan nilai pixel yang ada didaerah sekitarnya, dan menggantinya dengan nilai yang lebih

cocok. Cara yang paling sederhana dalam mencapainya adalah dengan melakukan pencegahan atau pembatasan nilai pixel, sehingga suatu pixel tidak memiliki nilai intensitas yang diluar nilai yang ada di sekitarnya.

Untuk itu perlu mengetahui nilai intensitas pada suatu kelompok pixel. Pada pengerjaannya pada suatu daerah pixel seharusnya bagian yang merup akan nilai tertinggi dan terendah, dan nilai yang sebanding pada kedua bagian akhir distribusi dihilangkan. Sehingga hasilnya meninggalkan nilai median. Dari sana didapatkan median filter, dimana didapat seluruh nilai distribusi intensitas, dan dihasilkan gambar baru yang sesuai dengan nilai-nilai median yang ada.

Berbeda dengan *Gaussian filter* yang menghaluskan keseluruhan gambar, pada *median filter* terlihat bahwa proses penghalusannya terjadi pada daerah tepi gambar. Sehingga meski terjadi penghalusan gambar, *median filter* lebih kearah "melembutkan" gambar yang ada. dapat pula mengurangi komputasi dengan mencari nilai extrem yang perlu untuk diubah, atau dengan membatasi nilai median yang digunakan.



Gambar 2. 4 Mencari Median Dengan Ukuran 3x3

Terlihat diatas bahwa pengurutan nilai intensitas *pixel* wajib untuk digunakan. Bila menggunakan *buble sort* maka untuk setiap daerah *n* x *n*, maka diperlukan fungsi sebanyak O (*n*4). Maka untuk *median filter* yang memiliki daerah kecil, antara sebesar 3x3 atau 4x4, maka *buble sort* atau penyortiran lain baik untuk digunakan. Namun tidak terlalu baik bila daerah yang dipakai lebih dari atau sama dengan 5x5, atau nilai inensitasnya lebih terbatas.

Median filter umumnya menggunakan kernel dengan ukuran 3x3.Namun dapat pula menggunakan ukuran yang lebih besar. Selain itu sesuai dengan perkembanganya maka bentuk yang dipakai juga dapat bemacam-macam, seperti salib, dan garis (vertikal dan horisontal), yang terpusat pada titik tengahnya. Hal ini dimaksudkan agar proses yang dihasilkan menjadi lebih cepat, terutama karena jumlah *pixel* yang dihitung menjadi lebih sedikit. *median filter* cukup dikenal baik atas kemampuanya untuk menghilangkan *salt and paper noise*. Selain itu *median filter* akan meningkatkan kualitas gambar, sehingga memperjelas daerah tepi (*edge*) pada gambar

Hal ini terjadi karena pada daerah yang terletak pada bagian tepi suatu gambar, *filter* akan memproses data dan umumnya akan mendap atkan nilai yang sesuai dengan daerah yang memiliki nilai intensitas yang lebih besar (didalam atau diluar batas). Sehingga *filter* secara tidak langsung menentukan terdapat pada bagian mana *pixel* itu berada. Hal ini tentu saja membuat daerah tepi menjadi sedikit melebar, namun perlu diingat bahwa *pixel* melebarkan daerah tepi dari kedua belah sisi, sehingga hal ini menyebabkan daerah tepinya bisa lebih terlihat.

2.8. *Thresholding*

Salah satu metode yang sering digunakan dalam pengolahan citra digital atau image processing adalah thresholding citra. Thresholding citra adalah suatu metode yang digunakan untuk memisahkan antara obyek dan backgroundnya. Thresholding merupakan teknik yang sederhana dan efektif untuk segmentasi citra. Proses thresholding sering disebut dengan proses binerisasi. Pada beberapa aplikasi pengolahan citra, terlebih dahulu dilakukan threshold terhadap citra gray level untuk dapat menjadi citra biner (citra yang memiliki nilai level keabuan 0 atau 255). Sebuah citra hasil proses thresholding dapat disajikan dalam histogram citra untuk mengetahui penyebaran nilai-nilai intensitas piksel pada suatu citra/bagian tertentu dalam citra sehingga untuk citra bimodal, histogram dapat

dipartisi dengan baik (segmentasi objek dengan background) dan dapat ditentukan nilai threshold-nya.

Distribusi gray level tiap kelas mempunyai PDF(*probabilistic density function*) yang diasumsikan sebagai GMM. Dengan menggunakan algoritma *Hybrid optimal estimation* ini yang merupakan gabungan dari PSO dan EM algoritma. Algoritma ini digunakan untuk menyelesaikan funsi dari estimasi parameter dari distribusi gaussian. Algoritma ini digunakan untuk mendapatkan kurva GMM yang sesuai sehingga membuat proses thresholding citra menjadi lebih efektif.

Pada prinsipnya adalah melakukan pengubahan nilai intensitas warna piksel menjadi 2 nilai yaitu 0 atau warna hitam dan 1 atau warna putih, berarti nilai pixel dikelompkan menjadi dua kelompok seperti ditunjukan pada rumus berikut:

$$f(x,y) = \begin{cases} 0, & x < T \\ 1, & x \geq T \end{cases} \quad \ldots (2.2)$$

T = level threshold

Menentukan hasil threshold dilakukan pembandingan, jika nilai citra lebih dari T maka akan diubah menjadi warna putih dengan nilai 1. Jika nilai kurang dari T maka akan diubah menjadi warna hitam dengan nilai 0. Citra hasil threshold ini digunakan lebih lanjut untuk proses pengenalan.

2.9. Metode Watershed

Tujuan utama dari algoritma segmentasi berdasarkan konsep watershed adalah mencari garis watershed. Ide dasarnya sangat sederhana, andaikan sebuah lubang dibuat di setiap bagian minimumnya dan seluruh topografi memenuhi dari bawah dengan membiarkan air sampai ke lubang pada tingkat yang seragam, ketika air naik ke kolam penangkapan maka air akan bergabung. Oleh karena itu sebuah bendungan (dam) dibangun untuk mencegah penggabungan. Banjir akhirnya mencapai tahap di mana hanya

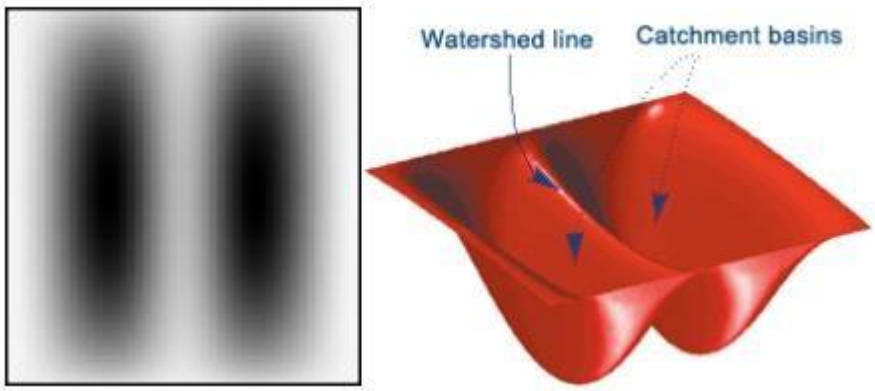
bagian atas bendungan yang bisa dilihat di atas garis air. Batas bendungan ini berhubungan dengan pembagian garis dalam watershed [12].

Salah satu aplikasi dasar dari segmentasi watershed adalah ekstraksi dari objek yang memiliki nilai dekat dan seragam terhadap background. Bagian citra yang mempunyai sifat variasi kecil di tingkat keabuan mempunyai nilai gradien yang kecil. Akan tetapi, pada kenyataannya sering melihat segmentasi watershed diaplikasikan ke gradien dari sebuah citra pada citra itu sendiri. Pada perumusan ini, regional minima dari kolam penangkapan berhubungan dengan nilai kecil dari gradien yang berhubungan ke objek yang diamati [13].

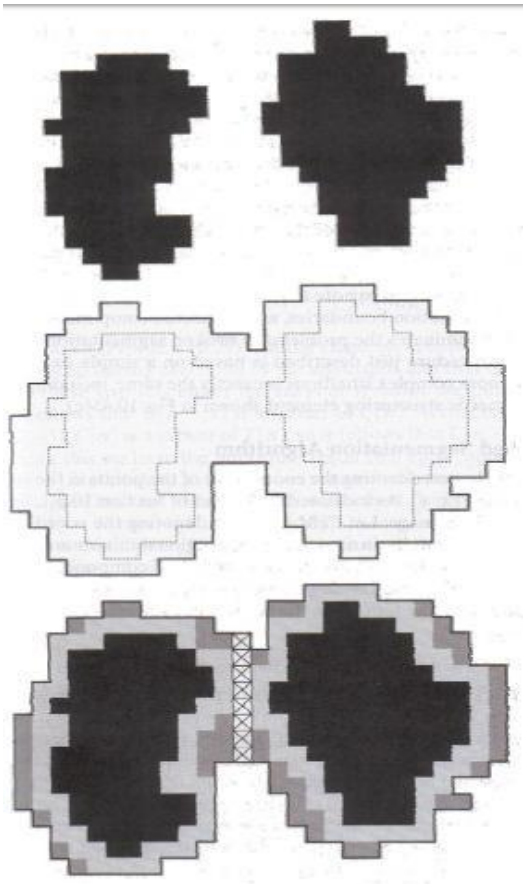
Watershed adalah salah satu pendekatan untuk segmentasi. Konsep dari *Watershed* adalah menggambarkan sebuah gambar dalam bentuk 3 dimensi dengan menganggap tingkat warna abu-abu adalah sebagai ketinggiannya dan dianggap bahwa semakin ke arah warna putih semakin tinggi. Jadi lebih cocok kalau dikatakan bahwa tingkat warna abu-abu adalah sebagai tingkat kedalamannya. Prinsip dari *Watershed* adalah mencari garis *watershed* (batas air) yaitu garis dimana titiktitiknya merupakan titik tertinggi dari penggambaran sebuah gambar ke dalam bentuk 3 dimensi.

2.9.1. Pembentukan Dam

Dam atau *watershed line* adalah hal yang paling penting dalam *watershed*. Pembuatan dam didasarkan pada gambar *binary*, yang merupakan anggota Z^2 . Cara termudah untuk membuat dam adalah dengan menggunakan *morphological dilation*.



Gambar 2. 5 Konsep Dasar *Watershed*



Gambar 2. 6 Pembuatan Dam

Pada Gambar 2.6 disebutkan ada 2 daerah (daerah dengan nilai minimum) yang akan di segmentasi pada kondisi pengisian air n-1 (gambar paling atas). Kemudian gambar selanjutnya menunjukan kondisi

pada saat pengisian air n. Pada saat ini 2 daerah tersebut menjadi 1 daerah. Pada kondisi seperti ini, dam harus dibuat untuk mencegah 2 daerah tersebut menyatu. Gambar paling bawah menunjukan proses pengisian air secara perlahan dengan melakukan *morphological dilation* menggunakan matriks 3x3. Pada pengisian atau *morphological dilation* pertama masih belum terjadi penyatuan 2 daerah (ditunjukan pada Gambar 2.7 paling bawah dengan warna abu-abu muda). Kemudian pada morphological dilation selanjutnya (warna abu-abu tua) terjadi penyatuan 2 daerah, maka titik di mana 2 daerah tersebut menyatu dibangun sebuah dam, yang ditandai dengan kotak yang berisi tanda silang. Proses ini dilakukan berulang hingga proses pengisian air n = maks+1.

2.9.2. Algoritma Watershed

Buat $M_1, M_2, M_3, \ldots, M_R$ menjadi kumpulan koordinat titik dalam daerah dengan nilai minimum dari sebuah gambar *g(x,y).* Gambar di sini pada umumnya sudah merupakan gambar yang telah diproses dengan *pre-processing* terlebih dahulu. Kemudian buat $C(M_i)$ menjadi kumpulan koordinat pada daerah pengisian yang memiliki hubungan dengan daerah minimum M_i (dianggap daerah pengisian dan daerah minimum membentuk komponen yang saling tersambung). Notasi *min* dan *max* digunakan untuk menandai nilai minimum dan nilai maksimum dari *g(x,y).* Kemudian buat *T[n]* menjadi kumpulan koordinat *(s,t)* di mana *g(s,t)* < *n*, sehingga dapat didefinisikan:

$$T[n]=\{(s,t)|f(x,y)<n\} \quad \ldots (2.3)$$

Secara geometri, *T[n]* adalah kumpulan koordinat dari *g(x,y)* yang berada di bawah daerah *g(x,y) = n*. Topografi tersebut kemudian diisi dengan penambahan integer mulai dari *n = min* hingga *n = max*. Pada setiap penambahan *n*, algoritma perlu mengetahui jumlah titik yang berada di bawah "kedalaman" *n*. Pada umumnya, daerah yang berada di bawah

g(x,y) = n diberi warna hitam atau nilai 0 dan yang berada di atasnya diberi warna putih atau nilai 1.

Kemudian buat $C_n(M_i)$ menjadi kumpulan koordinat titik pada daerah pengisian yang berhubungan dengan daerah minimum M_i yang diisi pada tahap *n*. $C_n(M_i)$ dapat dilihat sebagai gambar *binary* bila menggunakan persamaan:

$$C_n(M_i) = C(M_i) \cap T[n] \quad \ldots (2.4)$$

Dengan kata lain $C_n(M_i) = 1$ pada lokasi *(x,y)* jika *(x,y)* $\in C(M_i)$ dan *(x,y)* $\in$ *T[n]*, selain itu $C_n(M_i) = 0$. Dari sini dapat dikatakan bahwa $C[n] = \bigcup_{i=1}^{R} C_n(M_i)$ dan

$$C[\max+1] = \bigcup_{i=1}^{R} C(M_i)$$

C[n-1] adalah *subset* dari *C[n]* dan *C[n]* adalah *subset* dari *T[n]* maka *C[n-1]* adalah subset dari *T[n]*. Dari sini didapatkan bahwa tiap komponen yang terhubung dari *C[n-1]* memiliki 1 komponen yang terhubung dari *T[n]*. Jika *Q* adalah kumpulan komponen yang terhubung dalam *T[n]*, maka untuk tiap komponen yang terhubung $q \in Q[n]$, maka ada 3 kemungkinan:

1. $q \cap C[n-1]$ adalah kosong
2. $q \cap C[n-1]$ mengandung 1 komponen yang terhubung dari *C[n-1]*
3. $q \cap C[n-1]$ mengandung lebih dari 1 komponen yang terhubung dari *C[n-1]*

Jika kondisi c terjadi maka pengisian akan menyebabkan 2 daerah menjadi 1, maka titik pada *C[n-1]* menjadi milik dam atau *watershed line.*

2.10. Parameter Nilai Citra

2.10.1. Area

Untuk memudahkan mendeteksi dan pemahaman mengenai ciri-ciri pada suatu object. Dimisalkan saja pada sebuah citra yang diperiksa terdapat satu object dan langsung diterapkan dengan cara lebih sederhana. Notasi O dalam persamaanpersamaan berikut adalah titik atau piksel milik object dalam citra biner, sedangkan indeks *x* dan *y* adalah lokasi titik tersebut dalam matriks berukuran *m x n.*

$$A = \sum_{x=1}^{m} \sum_{y=1}^{n} O(x, y) \quad \ldots (2.5)$$

Nilai A didapatkan dengan cara membaca piksel piksel milik objectdan menambah nilai A dengan 1 setiap kali ditemukan piksel milik object. Dalam hal ini A yang merupakan luas proyeksi bayangan obyek di dalam citra disebut juga momen ordo ke nol dari bayangan object tersebut. Dalam matriks berikut diketahui area yang terdapat pada matriks tersebut.

0	0	0	1	1	1	
0	0	1	1	1	1	
0	1	1	1	1	1	
0	1	1	1	1	1	
0	1	1	1	1	1	
0	1	1	1	1	1	→ A = 27.

Gambar 2. 7 Contoh Nilai Area

Nilai Area yang telah didapatkan langsung disimpan kedalam *dataset* aplikasi.

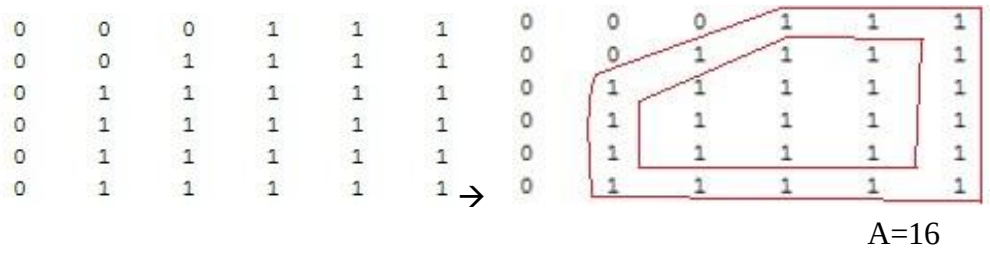
2.10.2. Perimeter

Perimeter adalah Bagian Terluar dari suatu object yang bersebelahan dengan Piksel atau puksel-puksel dari latar belakang. Oleh karena itu perimeter mempunyai beberapa definisi yang berbeda namun memiliki maksud yang sama. Ketiga definisi tersebut adalah :

1. Jumlah panjang garis yang memisahkan sepasang piksel p dan q dimana p € S dan q € Š

2. Jumlah langkah yang diambil dalam menemukan batas daerah
3. Jumlah piksel batas daerah

Jadi nilai perimeter suatu object dapat dicari dengan menghitung banyaknya piksel yang merupakan piksel-piksel yang berada pada perbatasan dari object tersebut, seperti terlihat pada matriks berikut



Gambar 2. 8 Contoh Nilai Perimeter

Nilai perimeter pada matriks tersebut adalah 16, dengan menjumlahkan nilai pada perbatasan pada object tersebut.

2.10.3. Compactness

Kombinasi beberapa sifat geometri object dapat digunakan untuk menganalisis bentuk object tersebut melalui perhitungan factor bentuk tak berdimensi. Kekompakan (*Compactness*) suatu object dapat diukur melalui analisis factor bentuk tak berdimensi dengan ketidaksamaan *isoperimetric* yang diwakili persamaan berikut

$$C = \frac{4 .\pi .A}{P^2} \quad \ldots (2.6)$$

Dimana C, A dan P masing-masing adalah *compactness*, Area dan Perimeter. Dengan menganalisis factor bentuk kekompakan, object-object dengan tepi yang rata akan memperlihatkan nilai yang berbeda-beda dengan tepi-tepi object yang bergerigi [15]. Jika dihubungkan dengan matriks yang Gambar 2.7 dan 2.8 yang diperoleh adalah :

$$C = \frac{4 .\pi .A}{P^2}$$

$$C = \frac{4 .3.14 .27}{16^2}$$

$$C = \frac{399.12}{256} = 1.55$$

Setelah nilai nilai dari Area, Perimeter dan *Compactness* diperolah maka langsung bisa dihitung data yang didapat dan disimpan sebagai data latih apabila proses *k-means clustering* telah diimplementasikan.

2.11 Metode k-means Clustering

Data Clustering merupakan salah satu metode Data Mining yang bersifat tanpa arahan (unsupervised). Ada dua jenis data clustering yang sering dipergunakan dalam proses pengelompokan data yaitu hierarchical (hirarki) data clustering dan nonhierarchical (non hirarki) data clustering. K-Means merupakan salah satu metode data clustering non hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih cluster/kelompok. Metode ini mempartisi data ke dalam cluster/kelompok sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu cluster yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam kelompok yang lain. Adapun tujuan dari data clustering ini adalah untuk meminimalisasikan objective function yang diset dalam proses clustering, yang pada umumnya berusaha meminimalisasikan variasi di dalam suatu cluster dan memaksimalisasikan variasi antar cluster.

Data clustering menggunakan metode K-Means ini secara umum dilakukan dengan algoritma dasar sebagai berikut [8] :

1. Tentukan Jumlah Cluster
2. Alokasikan data ke dalam cluster secara random
3. Hitung *Centroid*/rata-rata dari data yang ada masing-masing cluster
4. Alokasikan masing-masing data ke *centroid*/rata-rata terdekat
5. Kembali ke step 3, apabila masih ada data yang berpindah cluster atau apabila perubahan nilai *centroid*, ada yang di atas nilai *threshold* yang

ditentukan atau apabila perubahan nilai pada *Objective Function* yang digunakan diatas nilai *Threshold* yang ditentukan

Dalam tulisan ini beberapa hal terkait dengan metode K-Means ini berusaha untuk dijelaskan, termasuk di antaranya beberapa pengembangan yang telah dilakukan terhadap K-Means, beberapa permasalahan yang harus diperhitungkan dalam menggunakan metode K-Means dalam pengelompokan data, ulasan mengenai keberadaan K-Means di antara metode pengklasifikasian dengan arahan (*supervised*) dan tanpa arahan (unsupervised), ulasan singkat mengenai metode K-Means untuk dataset yang mempunyai bentuk khusus dan mixture modelling, serta algoritma dari metode-metode pengelompokan yang masih digolongkan sebagai pengembangan metode K-Means.

2.11.1. Perkembangan Penerapan K-Means

Beberapa alternatif penerapan K-Means dengan beberapa pengembangan teoriteori penghitungan terkait telah diusulkan. Hal ini termasuk pemilihan:

1. *Distance space* untuk menghitung jarak di antara suatu data dan centroid

2. Metode pengalokasian data kembali ke dalam setiap cluster

3. Objective function yang digunakan

2.11.2. *Distance Space* Untuk Menghitung Jarak Antara Data dan *Centroid*

Beberapa distance space telah diimplementasikan dalam menghitung jarak (distance) antara data dan centroid termasuk di antaranya L1 (Manhattan/City Block) distance space, L2 (Euclidean) distance space[10], dan Lp (Minkowski) distance space [9]. Jarak antara dua titik x1 dan x2 pada Manhattan/City Block distance space dihitung dengan menggunakan rumus sebagai berikut [8]:

$$D_{L_1}(x_2, x_1) = \|x_2 - x_1\|_1 = \sum_{j=1}^{p} |x_{2j} - x_{1j}| \quad \ldots 2.7$$

dimana:

p : Dimensi data

| . | : Nilai absolut

Sedangkan untuk L2 (*Euclidean*) distance space, jarak antara dua titik dihitung menggunakan rumus sebagai berikut[9]:

$$D_{L_2}(x_2, x_1) = \|x_2 - x_1\|_2 = \sqrt{\sum_{j=1}^{p} (x_{2j} - x_{1j})^2} \quad \ldots 2.8$$

dimana:

p : Dimensi data

Lp (*Minkowski*) distance space yang merupakan generalisasi dari beberapa distance space yang ada seperti L1 (*Manhattan/City Block*) dan L2 (*Euclidean*), juga telah diimplementasikan [9]. Tetapi secara umum *distance space* yang sering digunakan adalah *Manhattan* dan *Euclidean*. *Euclidean* sering digunakan karena penghitungan jarak dalam distance space ini merupakan jarak terpendek yang bisa didapatkan antara dua titik yang diperhitungkan, sedangkan *Manhattan* sering digunakan karena kemampuannya dalam mendeteksi keadaan khusus seperti keberadaaan *outliers* dengan lebih baik.

2.12. Perancangan Sistem

Perancangan sistem adalah sebagian gambaran, perencanaan dan pembuatan sketsa atau pengaturan dari beberapa elemen yang terpisah kedalam satu kesatuan yang utuh dan berfungsi. Untuk mencapai tujuan pada analisa sistem harus dapat mencapai sasaran-sasaran sebagai berikut :

• Perencanaan sistem harus berguna, mudah dipahami dan nantinya mudah digunakan.

• Rancangan sistem harus dapat mendukung tujuan utama perusahaan sesuai dengan yang telah didefinisikan pada tahap perancangan sistem yang dilanjutkan pada analisa sistem.

• Perancangan sistem harus efisien dan efektif untuk dapat mendukung pengolahan transaksi.

2.12.1. Konsep Dasar Data Flow Diagram (DFD)

DFD merupakan diagram yang menggunakan notasi-notasi atau simbol-simbol untuk menggambarkan arus dari data sistem yang telah ada atau sistem baru yang akan dikembangkan secara logika tanpa mempertimbangkan lingkungan fisik dimana data tersebut mengalir atau data tersebut akan disimpan.

2.12.2 Levelisasi DFD

Levelisasi pada DFD adalah penggambaran DFD dengan membagi DFD berdasarkan tingkatan dari tingkatan yang paling tinggi sampai dengan tingkat paling bawah yaitu :

• Diagram Konteks (diagram level 0) Diagram ini dibuat untuk menggambarkan sumber serta tujuan data yang akan diproses atau dengan kata lain diagram tersebut untuk menggambarkan sistem secara umum atau global dari keseluruhan sistem yang ada.

• Diagram Zero (diagram level 1) Diagram ini dibuat untuk menggambarkan tahapan proses lebih terperinci dari diagram kontek.

• Diagram Primitif / Detail (diagram level 2) Diagram ini dibuat untuk menggambarkan arus data secara lebih mendetail lagi dari tahapan proses yang ada di dalam diagram nol.

2.12.3. Komponen DFD

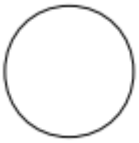
Ada 4 simbol yang digunakan dalam DFD yaitu :

• Terminator / Kesatuan Luar Terminator mewakili entitas eksternal yang berkomunikasi dengan sistem yang sedang dikembangkan. Komponen terminator digambarkan sebagai berikut :



Gambar 2. 9 Simbol Terminator

• Proses (Process) Suatu kegiatan sistem yang mentransformasikan dari masukkan ke keluaran. Komponen Proses digambarkan sebagai berikut :



Gambar 2. 10 simbol Proses

• Simpanan Data (Data Store) Simpanan data merupakan simpanan dari data yang dapat berupa :

- suatu file database

- suatu arsip atau catatan manual

- suatu table acuan manual Komponen Data Store digambarkan sebagai berikut :



Gambar 2. 11 Simbol Data Store

• Arus Data (Data Flow) Arus data ini mengalir diatara proses, simpanan data (Data store) dan kesatuan luar, serta menunjukkan arus dari data

yang berupa masukkan 14 dari sistem atau hasil dari proses sistem.

Komponen Data Flow digambarkan sebagai berikut :



Gambar 2. 12 Simbol Data Flow

2.13. Matlab

Matlab merupakan sebuah singkatan dari Matrix Laboratory, yang pertama kali dikenalkan oleh University of New Mexico dan University of Stanford pada tahun 1970. software ini pertama kali memang digunakan untuk keperluan analisis numerik, aljabar linier dan teori tentang matriks. Saat ini, kemampuan dan fitur yang dimiliki oleh Matlab sudah jauh lebih lengkap dengan ditambahkannya toolboxtoolbox yang sangat luar biasa. Beberapa manfaat yang didapatkan dari Matlab antara lain:

 Perhitungan Matematika Komputasi numerik

 Simulasi dan pemodelan

 Visualisasi dan analisis data

 Pembuatan grafik untuk keperluan sains dan teknik

 Pengembangan aplikasi, misalnya dengan memanfaatkan GUI.

Matlab dapat dipadang sebagai sebuah kalkulator dengan fitur yang lengkap. Pernah menggunakan kalkulator dengan degan fasilitas minimal, misalnya hanya terdapat fasilitas penambahan, pengurangan perkalian dan pembagian. Kalkulator yang lebih lengkap lagi adalah kalkulator scientific dimana fasilitas yang diberikan tidak hanya yang disebutkan di atas, melainkan sudah ada fungsi-fungsi trigonometri, bilangan kompleks, akar kuadrat dan logaritma. Nah, Matlab mirip dengan kalkulator tersebut, tetapi dengan fitur-fitur yang lengkap diantaranya dapat digunakan untuk

memprogram, aplikasi berbasis GUI dan lengkap dengan toolbox yang dapat dimanfaatkan untuk memecahkan masalah sains dan teknik.

Matlab memberikan kemudahan bagi para pengguna untuk menemukan bantuan sehubungan dengan semua fasilitas yang diberikan oleh Matlab. Misalnya, bantuan tentang bagaimana memulai Matlab pertama kali, trik pemrograman, membuat grafik 2 dan 3 dimensi, menggunakan tool akuisisi data, pengolahan sinyal, penyelesaian persamaan diferensial parsial.