

BAB II

TINJAUAN PUSTAKA

2.1 *Computer Vision*

Computer vision didefinisikan sebagai salah satu cabang ilmu pengetahuan yang mempelajari bagaimana komputer dapat mengenali obyek yang diamati atau diobeservasi. Arti dari *computer vision* adalah ilmu dan teknologi mesin yang melihat, dimana mesin mampu mengekstrak informasi dari gambar yang diperlukan untuk menyelesaikan tugas tertentu.

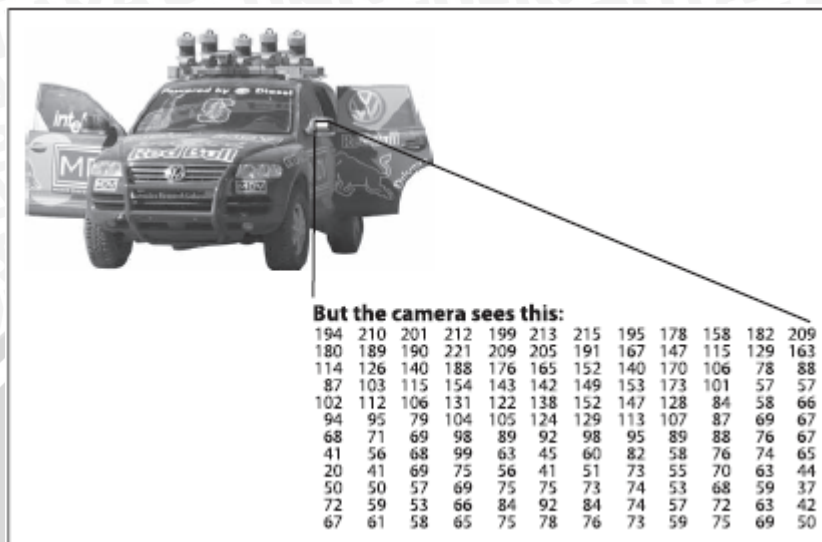
Computer vision adalah kombinasi antara :

- Pengolahan Citra (*Image Processing*), yaitu bidang yang berhubungan dengan proses transformasi citra/gambar (*image*). Proses ini bertujuan untuk mendapatkan kualitas citra yang lebih baik.
- Pengenalan Pola (*Pattern Recognition*), bidang ini berhubungan dengan proses identifikasi obyek pada citra atau interpretasi citra. Proses ini bertujuan untuk mengekstrak informasi/pesan yang disampaikan oleh gambar/citra.

Sebagai manusia, kita dapat melihat struktur benda-benda tiga dimensi yang ada di sekitar kita dengan mudah. Coba kita bayangkan tentang visual sebuah kacamata. Kita akan dengan mudah dapat mengenali seperti apa bentuk dan warnanya, serta dengan mudah kita dapat membedakan mana bagian frame dan mana lensa dari visual kacamata yang sedang kita bayangkan.

Karena manusia adalah makhluk visual, maka kita dapat terbodohi dengan mudah dan menganggap bahwa *computer vision* adalah bahasan yang mudah. Tidak pernah terpikir oleh kita betapa sulitnya untuk menemukan sebuah objek dalam sebuah gambar. Intuisi awal kita bisa sangat menyesatkan. Otak manusia membagi sinyal penglihatan menjadi beberapa kanal yang mengalirkan berbagai macam informasi ke dalam otak. Otak kita mempunyai sistem yang dapat mengidentifikasi, kemudian memeriksa mana bagian penting dari sebuah gambar sembari menurunkan kinerja pemeriksaan pada area yang lain (Bradski dan Kaehler, 2008).

Sedangkan di dalam *machine vision*, sebuah komputer menerima deretan angka dari sebuah kamera ataupun dari media penyimpanan.



Gambar 2.1 Bagi Komputer Bagian Kaca Mobil Hanya Deretan Angka

Sumber : Bradski dan Kaehler (2008)

2.2 Pengolahan Citra Digital

Pengolahan citra digital adalah sebuah disiplin ilmu yang mempelajari hal-hal yang berkaitan dengan perbaikan kualitas gambar (peningkatan kontras, transformasi warna, restorasi citra), transformasi gambar (rotasi, translasi, skala, transformasi geometrik), melakukan pemilihan citra ciri (*feature images*) yang optimal untuk tujuan analisis, melakukan proses penarikan informasi atau deskripsi objek atau pengenalan objek yang terkandung pada citra, melakukan kompresi atau reduksi data untuk tujuan penyimpanan data, transmisi data, dan waktu proses data. *Input* dari pengolahan citra adalah citra, sedangkan *input*-nya adalah citra hasil pengolahan (T. Sutoyo dkk, 2009).

2.3 Citra

Citra adalah gambar pada bidang dua dimensi. Ditinjau dari sudut pandang matematis, citra merupakan fungsi *continue* dari intensitas cahaya pada bidang dua dimensi. Sumber cahaya menerangi obyek, obyek memantulkan kembali sebagian dari berkas cahaya tersebut. Pantulan cahaya ini ditangkap oleh alat-alat

optik, misalnya mata pada manusia, kamera, *scanner*, dan sebagainya, sehingga bayangan obyek yang disebut citra terekam.

2.3.1 Citra Analog

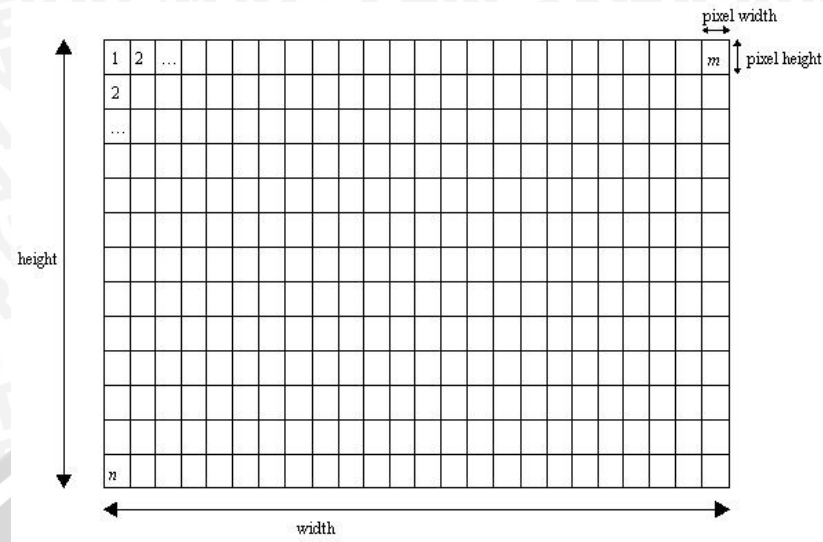
Citra analog adalah citra yang bersifat kontinu, seperti gambar pada monitor televisi, foto sinar-X, foto yang tercetak di kertas foto, lukisan, pemandangan alam, hasil CT scan, gambar-gambar yang terekam pada pita kaset, dan lain sebagainya. Citra analog tidak dapat dipresentasikan dalam computer sehingga tidak bisa diproses di komputer secara langsung. Oleh sebab itu, agar citra ini dapat diproses di computer, proses konversi analog ke digital harus dilakukan terlebih dahulu. Citra analog dihasilkan dari alat-alat analog, seperti video kamera analog, kamera foto analog, WebCam, CT scan, sensor rontgen untuk foto thorax, sensor gelombang pendek pada sistem radar, sensor ultrasound pada sistem USG, dan lain-lain (Sutoyo dkk, 2009).

2.3.2 Citra Digital

Citra digital dapat didefinisikan sebagai fungsi dua variabel, $f(x,y)$, dimana x dan y adalah koordinat spasial dan nilai $f(x,y)$ adalah intensitas citra pada koordinat tersebut. Teknologi dasar untuk menciptakan dan menampilkan warna pada citra digital berdasarkan pada penelitian bahwa sebuah warna merupakan kombinasi dari tiga warna dasar, yaitu merah, hijau, dan biru (*Red, Green, Blue – RGB*).

2.4 Pixel dan Resolusi Citra

Pixel merupakan satuan komponen terkecil yang menentukan ukuran dari suatu gambar (Gonzales dan Woods, 2002 seperti dikutip Astriawanti, 2012).



Gambar 2.2 Resolusi Citra

Sumber : <http://telestream.force.com/kb/>

Resolusi citra menyatakan ukuran bit dari suatu citra dalam satuan *pixel image*. Untuk memudahkan, kita asumsikan dengan gambar segi empat dengan panjang m (dalam kolom) dan tinggi n (dalam baris) dari matrix I , seperti pada gambar 2.3.

Banyaknya *pixel* persatuan panjang dikenal dengan satuan *dot per inch* (dpi). Semakin banyak titik per incinya, kualitas citra semakin baik dan detail, namun semakin banyak pula memori (bit) yang diperlukan untuk menyimpan citra tersebut.

2.5 Jenis-Jenis Citra

Informasi penting dari suatu citra adalah informasi mengenai intensitas cahaya yang tertangkap, ini erat hubungannya dengan warna. Suatu citra digital dapat dikelompokkan berdasarkan format intensitas penyusunnya, yaitu citra berwarna, citra skala keabuan (*grayscale*), dan citra biner (monokrom).

2.5.1 Citra Berwarna

Citra berwarna merupakan kombinasi dari tiga warna dasar yaitu merah, hijau, dan biru. Citra ini biasanya disebut citra RGB. Kombinasi dari tiga warna RGB menghasilkan warna yang khas untuk *pixel* yang bersangkutan. Setiap

komponen warna memiliki intensitas sendiri dengan nilai minimum 0 dan nilai maksimal 255 dan memiliki panjang 8 bit, hal ini menyebabkan setiap *pixel* pada citra RGB membutuhkan 3 *byte* di dalam media penyimpanan. Pada citra RGB mempunyai range warna $(2^8) (2^8) (2^8) = 2^{24}$, sehingga RGB mempunyai 16.777.216 kombinasi warna. Dalam windows citra berwarna dikenal dengan istilah *true color*.

2.5.2 Citra Grayscale

Citra *grayscale* merupakan citra digital dimana unsur RGB (*Red Green Blue*) hanya digunakan untuk menentukan tingkat intensitas pada suatu citra. Pada citra *grayscale* warna yang dimiliki adalah hitam, keabuan, putih. Tingkat keabuan merupakan warna abu dengan berbagai tingkatan dari hitam hingga mendekati warna putih. Pada citra *grayscale* 8 bit, masing-masing komponen memiliki nilai minimum 0 dan nilai maksimal 255 dan memiliki panjang 8 bit.

2.5.3 Citra Biner

Citra biner disebut juga citra monokrom yang hanya memiliki dua nilai yaitu 1 atau 0 yang jika diimplementasikan ke dalam 8-bit nilai yang digunakan adalah 0 dan 255. Biasanya citra biner ini digunakan pada pengolahan citra seperti segmentasi.

2.6 Noise

Noise adalah suatu bentuk kerusakan pada *image signal* yang disebabkan oleh gangguan eksternal. Gangguan pada citra umumnya berupa variasi intensitas suatu *pixel* yang tidak berkorelasi dengan *pixel-pixel* tetangganya (Yuwono, 2010 seperti dikutip Astriawanti, 2012).

2.7 Operasi Morfologikal

Morphologi adalah teknik pengolahan citra digital dengan menggunakan bentuk (*shape*) sebagai pedoman dalam pengolahan. Nilai dari setiap *pixel* dalam citra digital hasil diperoleh melalui proses perbandingan antara *pixel* yang bersesuaian pada citra digital masukan dengan *pixel* tetangganya.

Operasi morfologi bergantung pada urutan kemunculan dari *pixel*, tidak memperhatikan nilai numerik dari *pixel* sehingga teknik morfologi sesuai apabila digunakan untuk melakukan pengolahan *binary image* dan *grayscale image*.

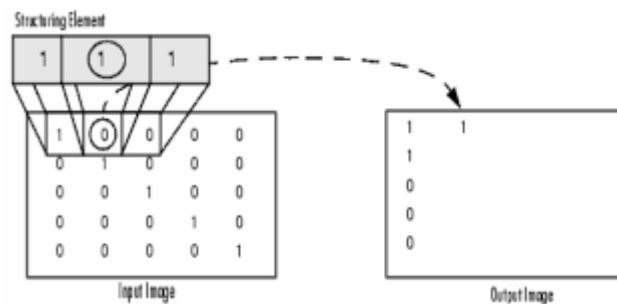
2.7.1 Dilasi

Dilasi adalah operasi morfologi yang akan menambahkan *pixel* pada batas antar objek dalam suatu citra digital. Operasi ini menggunakan aturan sebagai berikut:

“Untuk gambar *grayscale* maka nilai hasil operasi (*input pixel*) adalah nilai maksimal yang diperoleh dari himpunan *pixel* tetangganya. Dalam *binary image*, jika ada *pixel* tetangga yang bernilai 1 maka *input pixel* akan diset menjadi 1”.

Secara matematis, dilasi A oleh B dapat dituliskan sebagai berikut.

$$A \oplus B = \{z \mid (\hat{B})_z \cap A \neq \emptyset\} \quad \text{atau} \quad A \oplus B = \{z \mid [(\hat{B})_z \cap A] \subseteq A\} \quad (2-1)$$



Gambar 2.3 Proses Dilasi pada Citra Biner

Sumber : <http://rai46.blogspot.com/>

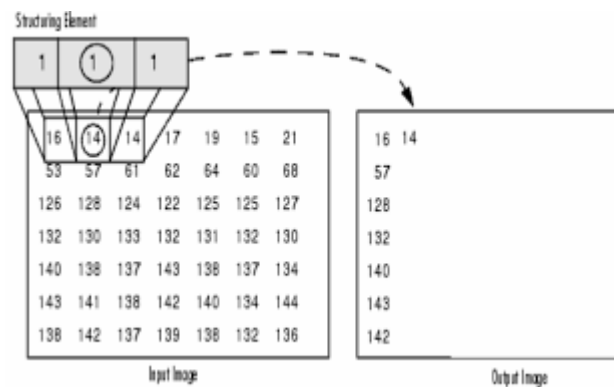
2.7.2 Erosi

Erosi adalah operasi morfologi yang akan mengurangi *pixel* pada batas antar objek dalam suatu citra digital. Operasi ini menggunakan aturan sebagai berikut:

“Untuk gambar *grayscale* maka nilai hasil operasi (*input pixel*) adalah nilai minimal yang diperoleh dari himpunan *pixel* tetangganya. Dalam *binary image*, jika ada *pixel* tetangga yang bernilai 0 maka *input pixel* akan diset menjadi 0”.

Secara matematis, erosi A oleh B dapat dituliskan sebagai berikut.

$$A \ominus B = \{z \mid (B)_z \subseteq A\} \quad (2-2)$$



Gambar 2.4 Proses Erosi pada Citra *Grayscale*

Sumber : <http://rai46.blogspot.com/>

2.7.3 Closing

Closing adalah kebalikan dari opening, yaitu proses dilatasi yang dilanjutkan dengan erosi. Proses ini berguna untuk menghilangkan *noise* berupa lubang-lubang kecil pada obyek.



Gambar 2.5 Contoh citra dengan *noise* (kiri) dan setelah dilakukan proses closing (kanan)

Sumber : <http://opencv-python-tutroals.readthedocs.org/>

2.8 Transformasi Intensitas Citra

Peningkatan kualitas citra dapat dilakukan melalui transformasi intensitas citra, yaitu besar intensitas setiap *pixel* pada citra diubah, tetapi posisi *pixel* tetap. Tranformasi ini dilakukan melalui sebuah fungsi yang disebut fungsi transformasi skala keabuan atau *Gray-Scale Transformation Function* (Castleman, 1996 seperti dikutip Sutoyo, 2009). Fungsi ini memetakan fungsi *input* $f_i(x,y)$ (yang bertindak sebagai citra *input*) menjadi fungsi *input* $f_o(x,y)$ (yang bertindak sebagai citra *input*). Untuk citra warna, fungsi GST diterapkan untuk ketiga elemen warna yang

ada (RGB). Penerapan ketiga elemen warna tadi tidak harus sama. Misalnya, elemen warna merah (*Red*) dikenai fungsi kontras, elemen warna hijau (*Green*) dikenai fungsi *brightness*, dan elemen warna biru (*Blue*) dikenai fungsi negasi. Menurut tingkat komputasi pengolahan citra digital, transformasi intensitas citra termasuk **operasi titik**, yaitu hasil proses suatu titik (*pixel*) tidak tergantung pada titik-titik tetangganya.

2.8.1 Transformasi Citra Warna menjadi Citra *Grayscale*

Citra warna bisa diubah menjadi citra *grayscale* dengan cara menghitung masing-masing elemen warna *Red*, *Green*, dan *Blue* yang dikalikan oleh koefisien tertentu. Secara matematis penghitungannya adalah sebagai berikut.

$$f_o(x, y) = 0,299 f_i^R(x, y) + 0,587 f_i^G(x, y) + 0,114 f_i^B(x, y) \quad (2-3)$$

Perbedaan koefisien untuk masing-masing elemen warna disebabkan karena tingkat kecerahan murni untuk masing-masing elemen, di mana elemen hijau adalah elemen paling cerah sehingga koefisiennya yang paling besar.

2.8.2 Operasi Ambang Batas (*Thresholding*)

Pada operasi ini nilai *pixel* yang memenuhi syarat ambang batas dipetakan ke suatu nilai yang dikehendaki. Dalam hal ini syarat ambang batas dan nilai yang dikehendaki disesuaikan dengan kebutuhan. Operasi *thresholding* yang mempunyai ketentuan berikut : nilai intensitas *input* $f_o(x, y) = 0$, bila nilai intensitas *input* $f_i(x, y) \leq 0$, nilai $f_o(x, y) = T_1$ bila $T_1 < f_i(x, y) \leq T_2$, nilai $f_o(x, y) = T_2$ bila $T_2 < f_i(x, y) \leq T_3$, ..., nilai $f_o(x, y) = T_{n-1}$ bila $T_{n-1} < f_i(x, y) \leq T_n$ secara matematis dapat ditulis sebagai berikut.

$$f_o(x, y) = \begin{cases} T_1, & f_i(x, y) \leq T_1 \\ T_2, & T_1 < f_i(x, y) \leq T_2 \\ T_3, & T_2 < f_i(x, y) \leq T_3 \\ \vdots & \\ T_n, & T_{n-1} < f_i(x, y) \leq T_n \end{cases} \quad (2-4)$$

Dalam hal ini $T_1, T_2, T_3, \dots, T_n$ adalah nilai-nilai yang dikehendaki, sedangkan $f_i(x, y) \leq T_1, T_1 < f_i(x, y) \leq T_2, T_2 < f_i(x, y) \leq T_3, \dots, T_{n-1} < f_i(x, y) \leq T_n$

adalah ambang batas yang diisyaratkan. Nilai-nilai ambang $T_1, T_2, T_3, \dots, T_n$ dan syarat ambang batasnya ditentukan sesuai kebutuhan.

Salah satu contoh operasi ambang batas yang cukup sering digunakan adalah operasi ambang tunggal. Operasi ambang tunggal adalah pemetaan citra *grayscale* menjadi citra biner (hanya mempunyai 2 warna saja, yaitu hitam dan putih), yang menggunakan fungsi transformasi sebagai berikut.

$$f_o(x, y) = \begin{cases} 0, & f_1(x, y) < 128 \\ 255, & f_1(x, y) \geq 128 \end{cases} \quad (2-5)$$

Hasilnya, *pixel-pixel* yang nilai intensitasnya di bawah 128 diubah menjadi hitam (nilai intensitasnya = 0), sedangkan *pixel-pixel* yang nilai intensitasnya di atas 128 diubah menjadi putih (nilai intensitas 255).

2.9 Properti Geometri

Setelah proses segmentasi objek selesai dilakukan, maka proses berikutnya adalah menganalisis objek untuk mengenali objek tersebut. Analisis objek didasarkan pada ciri khas (*feature*) geometri pada objek tersebut. Diasumsikan di dalam citra biner hanya terdapat 1 buah objek, dan berikut ini beberapa properti yang akan digunakan pada skripsi ini seperti yang diuraikan oleh Munir (2006).

- Luas atau Ukuran Objek (A)

$$A = \sum_{i=1}^n \sum_{j=1}^m f(i, j) \quad (2-6)$$

di mana (i, j) adalah koordinat *pixel* yang diperiksa dan $f(i, j) = 1$ jika (i, j) adalah *pixel* objek.

- Pusat massa

$$\bar{x} = \frac{\sum_{i=1}^n \sum_{j=1}^m j \cdot f(i, j)}{A}$$

$$\bar{y} = \frac{\sum_{i=1}^n \sum_{j=1}^m i \cdot f(i, j)}{A}$$

(2-7)

di mana (i, j) adalah koordinat *pixel* yang diperiksa dan $f(i, j) = 1$ jika (i, j) adalah *pixel* objek.

- Jarak Euclidean

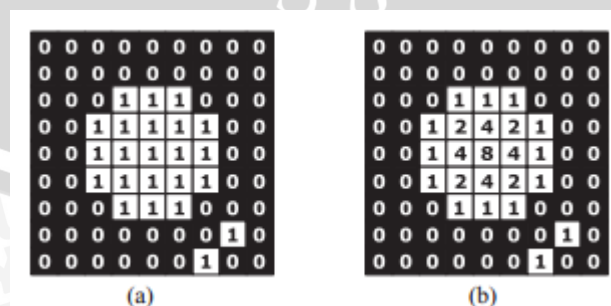
$$d_{\text{Euclidean}} = \sqrt{(i_1 - i_2)^2 + (j_1 - j_2)^2}$$

(2-8)

di mana (i₁, j₁) adalah koordinat dari *pixel* pertama/awal dan (i₂, j₂) adalah koordinat dari *pixel* kedua/tujuan.

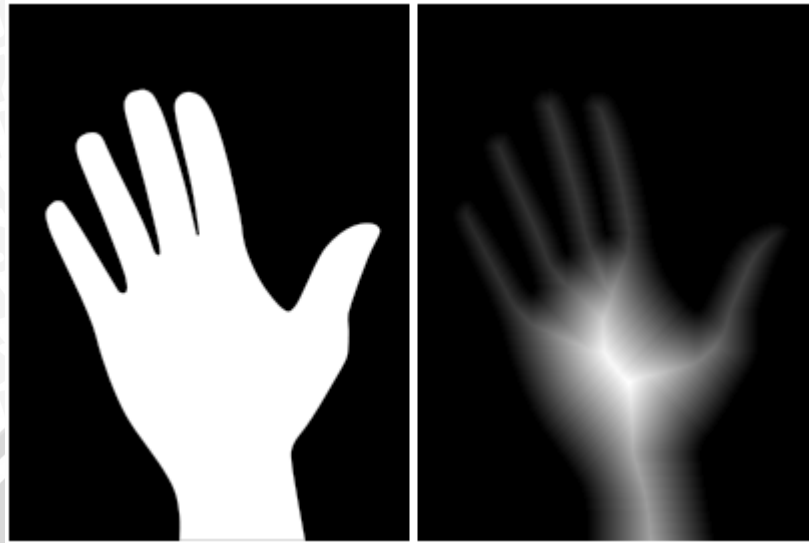
2.10 Euclidean Distance Transform (EDT)

Konsep utama dari EDT adalah menghitung jarak tiap *pixel* latar depan dengan *pixel* latar belakang yang terdekat. Pada skripsi ini, EDT digunakan untuk menentukan posisi tengah pada telapak tangan, dimana pada citra tangan hanya pada tengah telapak tanganlah *pixel-pixel* yang memiliki nilai EDT tinggi, karena di situlah luasan terbesar pada tangan.



Gambar 2.6 (a) Citra biner (b) Pengimplementasian EDT pada citra biner dengan nilai yang belum diakarkan

Sumber : Fabbri dkk. (2008)



Gambar 2.7 (a) Citra tangan dalam bentuk citra biner (b) EDT dari a
Sumber : <http://opencv-code.com/>

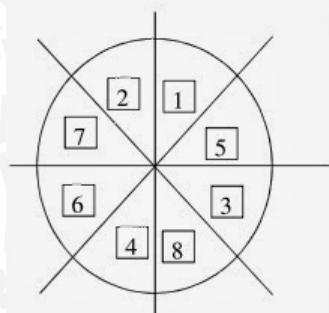
Berikut ini adalah algoritma dari *Euclidean Distance Transform* :

1. Mengelompokkan *pixel* hitam dalam satu *array*.
2. Mengelompokkan *pixel* putih dalam satu *array*.
3. Menghitung jarak Euclidean dari setiap *pixel* putih, mulai dari *array* pertama (dilambangkan *array* a) terhadap *pixel* hitam mulai dari *array* pertama sampai terakhir.
 - a. Mencari nilai minimal dari jarak-jarak tersebut.
 - b. Mengubah nilai *pixel* putih yang sedang diperiksa menjadi nilai minimal tersebut.
 - c. Ulangi proses 3 dengan urutan selanjutnya (a+1).
4. Proses berhenti jika *array* putih yang diperiksa sudah mencapai *index* terakhir.

2.11 *Midpoint Circle Algorithm* (MCA)

Lingkaran merupakan objek yang simetris sehingga karakteristik ini dapat dimanfaatkan untuk mengurangi pekerjaan pada saat menggambar lingkaran. Lingkaran dibagi menjadi 8 oktan, misalkan kita menyusun algoritma untuk

menggambarkan lingkaran di oktan pertama, maka koordinat untuk 7 oktan selanjutnya dapat ditentukan.



Oktan	x	y
1	x	y
2	-x	y
3	x	-y
4	-x	-y
5	y	x
6	-y	x
7	y	-x
8	-y	-x

Gambar 2.8 Pembagian Lingkaran Menjadi 8 Oktan

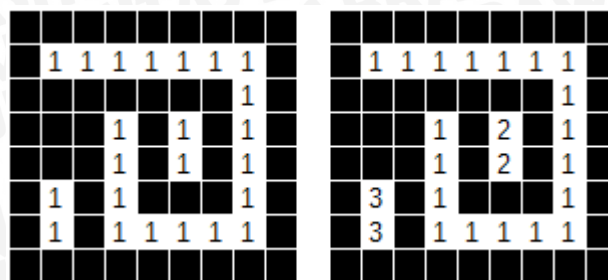
Sumber : <http://collegerlearn.blogspot.com/>

Langkah-langkah mengerjakan algoritma *midpoint circle* :

1. Input jari-jari (r) dan koordinat pusat lingkaran (x_c, y_c), kemudian tentukan koordinat untuk titik awal (x_0, y_0) = (x, r).
2. Hitung nilai awal untuk parameter keputusan $P_0 = 1 - r$.
3. Untuk setiap x_k , mulai dari $k = 0$, periksa/cek kondisi berikut :
 - a. Jika nilai $p_k < 0$ maka, titik selanjutnya pada lingkaran dgn pusat ($0,0$) adalah $x_{k+1} = x + 1$, nilai y tetap; dan $p_{k+1} = p_k + 2x + 1$
 - b. Selain itu berarti, $p_k > 0$ maka, titik selanjutnya adalah $x_{k+1} = x + 1$, $y_{k+1} = y - 1$; dan $p_{k+1} = p_k + 2x - 2y + 1$
4. Ulangi langkah 3 sampai 4, hentikan perhitungan ketika $x > \text{atau} = y$

2.12 Pelabelan Komponen Terkoneksi

Pelabelan komponen terkoneksi adalah sebuah algoritma yang berguna untuk mendeteksi region atau area yang terdapat pada citra biner. Algoritma ini bekerja dengan 2 kali meninjau seluruh data pada citra 2 dimensi. Tinjauan pertama bertujuan untuk memberikan label sementara pada setiap *pixel*nya dan dicatat. Tinjauan kedua bertujuan untuk mengganti label sementara tadi dengan label terkecil dari kelas yang sama.



Gambar 2.9 Contoh citra biner (kiri) dan citra setelah pengimplementasian pelabelan komponen terkoneksi (kanan)

Sumber : <http://www.codeproject.com/>

Algoritma ini mengecek konektivitas setiap *pixel* tetangga dari *pixel* yang sedang diproses. *Pixel* tetangga ini meliputi *pixel* pada diagonal kiri, diagonal kanan, horisontal dan vertikal dari *pixel* yang sedang diproses, atau biasa disebut 8-konektivitas. Alur dari pemberian label pada tinjauan pertama adalah sebagai berikut.

1. Tinjau tiap *pixel* dari citra.
2. Apabila *pixel* yang sedang ditinjau bukan *pixel* latar belakang maka dilakukan pengecekan tetangga dari *pixel* tersebut.
3. Apakah *pixel* tetangga sudah mempunyai label, jika iya maka beri label pada *pixel* yang sedang ditinjau sesuai dengan label utama dari *pixel* tetangga, jika tidak, maka beri label baru pada *pixel* yang sedang ditinjau.
4. Jika terdapat label berbeda pada *pixel* tetangga maka berikan label utama pada *pixel* tetangga tersebut. Label utama didapat dengan mengambil nilai label terkecil.

Setelah itu dilakukan tinjauan kedua yang berguna untuk mengubah semua label sementara ke label utama yang telah diberikan (Stefano dan Bulgarelli, 1999 seperti dikutip Relung, 2014).

2.13 *Moment Invariant*

Moment invariant dikenalkan oleh Hu pada tahun 1962 yang merupakan fungsi nonlinier yang invarian terhadap rotasi, translasi, dan skala serta dideferensialkan dalam momen geometri foto (Gonzales dan Woods, 2002 seperti dikutip Rizanti, 2013).

Momen dapat menggambarkan suatu objek dalam hal area, posisi, orientasi, dan parameter terdefinisi lainnya. Persamaan dasar dari momen suatu objek didefinisikan sebagai berikut

$$m_{ij} = \sum_x \sum_y x^i y^j a_{xy} \quad (2-9)$$

di mana orde dari momen adalah (i+j), x dan y menyatakan koordinat titik, sedangkan a_{xy} menyatakan intensitas titik. Momen tingkat ke-0 dan ke-1 (*zero and first-order moments*) didefinisikan sebagai berikut

$$\begin{aligned} m_{00} &= \sum_x \sum_y a_{xy} \\ m_{10} &= \sum_x \sum_y x.a_{xy} \\ m_{01} &= \sum_x \sum_y y.a_{xy} \end{aligned} \quad (2-10)$$

Pada citra biner yang mana a_{xy} akan bernilai 0 atau 1, momen tingkat ke-0 (m_{00}) adalah sama dengan area dari objek. Pusat dari area atau massa (*centroid*) adalah parameter yang baik untuk menyatakan lokasi dari objek. Pusat area dari objek didefinisikan sebagai berikut

$$x' = \frac{m_{10}}{m_{00}} \text{ dan } y' = \frac{m_{01}}{m_{00}} \quad (2-11)$$

dengan (x' , y') merupakan pusat koordinat dari objek. Nilai x' dan y' yang merupakan koordinat *centroid* inilah yang akan diekstraksi pada skripsi ini, yang terdiri dari *centroid* pada telapak tangan dan *centroid* pada tiap-tiap jari yang terangkat.