

## BAB II

### TINJAUAN PUSTAKA DAN LANDASAN TEORI

#### 2.1 Tinjauan Pustaka

Untuk mendukung penulisan ini digunakan beberapa landasan teori yang relevan serta berkaitan dengan pokok bahasan sebagai berikut :

- a. Penelitian yang dilakukan oleh Navneet Dalal and Bill Triggs, France : "*Histograms of Oriented Gradients for Human Detection*" menjelaskan bahwa metode HOG memungkinkan untuk mendeteksi manusia dengan 6 langkah dasar sebagai proses seperti, normalisasi warna, perhitungan gradien, orientasi bin dan sel, normalisasi kontras dengan penumpukan blok spasial, pengumpulan data *histogram*, dan klasifikasi menggunakan SVM.
- b. Penelitian yang kedua dilakukan oleh Yanwei Pang, Kun Zhang, Yan Yuan, Kongqiao Wang tentang "*Distributed Object Detection With Linear SVM*", komputasi rendah kompleksitas dan generalisasi tinggi dua gol penting untuk video deteksi objek. Kompleksitas komputasi rendah disini berarti tidak hanya kecepatan yang cepat tetapi juga konsumsi energi lebih sedikit. Sliding window metode objek deteksi dengan *Linear Support Vector Machine* (SVM) adalah framework untuk objek deteksi yang sudah umum.
- c. Penelitian yang ketiga dilakukan oleh Shunli Zhang, Xin Yu, Yao Sui, Sicong Zhao, and Li Zhang tentang "*Object Tracking With Multi-View Support Vector Machines*", metode *tracking* baru melalui *multi-view framework learning multi-view* dengan menggunakan *Support Vector Machine* (SVM), metode *tracking multi-view* SVM dibangun berdasarkan beberapa tampilan fitur dan strategi kombinasi baru. Untuk mewujudkan representasi yang komprehensif, kami memilih tiga jenis fitur, yaitu, nilai abu-abu skala, *histogram* gradien berorientasi (HOG), dan pola biner lokal (LBP), untuk melatih SVM sesuai.

## 2.2 Landasan Teori

### 2.2.1 Pengertian Computer Vision

*Computer Vision* suatu bidang yang bertujuan untuk membuat suatu keputusan yang berguna mengenai obyek fisik nyata dan keadaan berdasarkan atas sebuah citra. *Computer Vision* merupakan kombinasi antara pengolahan citra dan pengenalan pola. Hasil keluaran dari proses *Computer Vision* merupakan *image understanding*. Pengembangan bidang ini dilakukan dengan mengadaptasi kemampuan dari penglihatan manusia dalam mengambil informasi. Dalam disiplin ilmu, *Computer Vision* berkaitan dengan teori pada AI yang mengekstraksi informasi dari citra. Data citra dapat diperoleh dengan beberapa bentuk seperti urutan video, sudut pandang dari beberapa kamera, atau data multi-dimensional dari *scanner* medis. Beberapa area permasalahan dalam *Computer Vision* yaitu : *Recognition, Motion*, dan restorasi citra tahun 1989, Fungsi-fungsi yang umumnya dilakukan dalam banyak system *Computer Vision* adalah sebagai berikut:

- a. *Image acquisition*, berhubungan dengan pengambilan citra yang digunakan misalnya citra untuk pengukuran suhu yang umumnya diambil menggunakan kamera yang dilengkapi dengan sensor suhu.
- b. *Pre-processing*, yaitu tahapan yang dilakukan untuk mengekstraksi beberapa informasi yang dilakukan dengan melakukan pengurangan *noise*, peningkatan kontras, dan untuk citra berwarna dapat juga dilakukan *gray scaling*.
- c. *Feature extraction* adalah pengambilan fitur-fitur bermakna dari citra dan umumnya fitur-fitur tersebut dapat berbentuk tepi/*edge*, titik-titik yang terlokalisasi, bentuk atau pergerakan.
- d. *Detection / segmentation* yang merupakan penentuan wilayah dari titik-titik atau wilayah yang menjadi pokok perhatian.
- e. *High level processing* adalah fungsi dimana data diperiksa apakah data diterima untuk *verification* dan data dikenali untuk *recognition*.

- f. *Decision making* adalah fungsi yang menentukan apakah suatu data yang ada itu sesuai atau tidak dengan menilai seberapa besar kecocokan dengan sesuatu yang diinginkan. (Shapiro, 2001)

### 2.2.3 Pengertian Pengolahan Citra (*Image Processing*)

Pemrosesan citra atau *image processing* adalah sebuah teknik dimana data dari sebuah citra didigitisasi dan beberapa operasi matematika diaplikasikan ke dalam data, biasanya dengan komputer digital, untuk membuat sebuah citra yang lebih berguna atau menyenangkan pada mata manusia, atau untuk merepresentasikan beberapa dari tugas interpretasi dan pengenalan yang biasanya dilakukan oleh manusia. Lebih luasnya lagi, *image processing* adalah bentuk dari *information processing* dimana kedua input dan output berupa citra, seperti foto atau *frame* dari video. Kebanyakan teknik *image processing* memperlakukan citra sebagai signal dua dimensi dan mengaplikasikan teknik *signal processing* yang standard kedalamnya.

#### 1. Definisi Citra

Sebuah citra adalah kumpulan piksel-piksel yang disusun dalam larik dua dimensi. Piksel adalah sampel dari pemandangan yang mengandung intensitas citra yang dinyatakan dalam bilangan bulat. Indeks baris dan kolom (x,y) dari sebuah piksel dinyatakan dalam bilangan bulat. Piksel (0,0) terletak pada sudut kiri atas pada citra, indeks x bergerak ke kanan dan indeks y bergerak ke bawah. Konvensi ini dipakai merujuk pada cara penulisan larik yang digunakan dalam pemrograman komputer. (a) koordinat pada grafik matematika (b) koordinat pada citra Pengertian citra berdasarkan sifatnya dapat dikategorikan menjadi dua , yaitu citra analog dan citra digital.

- a. Citra Analog merupakan kesan visual yang dihasilkan oleh sistem saraf penglihatan manusia akibat stimulasi cahaya. Cahaya merupakan gelombang elektromagnetik yang berada didalam jangkauan 350nm sampai 780nm dari spektrum elektromagnetik dan diekspresikan sebagai distribusi energi  $L(\lambda)$ . Cahaya yang diterima oleh sebuah objek disebut iluminasi. Dimana  $p(\lambda)$

adalah tingkat reflektansi objek,  $L(\lambda)$  adalah distribusi energi cahaya dan  $I(\lambda)$  adalah iluminasi yang dihasilkan. Iluminasi yang dihasilkan akan ditangkap oleh sensor visual mata yang disebut batang dan kerucut. Sensor batang berjumlah sekitar seratus juta dan bersifat peka terhadap nilai iluminasi yang rendah (gelap), sedangkan sensor kerucut berjumlah sekitar 6,5 juta, peka terhadap nilai iluminasi yang tinggi (terang) dan warna.

- b. Citra digital dihasilkan melalui proses digitalisasi, yang terdiri atas proses sampling dan *quantizing*. Proses sampling adalah proses penangkapan gelombang cahaya oleh serangkaian sensor yang peka terhadap gelombang elektromagnetik cahaya tertentu, memiliki ukuran tertentu, dan disusun atas baris dan kolom, ukuran dan bentuk sensor, serta banyaknya sensor secara horizontal dan vertikal menentukan resolusi citra yang dihasilkan. Sensor-sensor ini akan menghasilkan tegangan- tegangan listrik yang sebanding dengan besar intensitas cahaya yang diterima, yang kemudian akan dikirimkan ke *quantizer* untuk ditentukan nilai-nilainya. Nilai-nilai inilah yang disebut dengan piksel. (Rafael C. Gonzalez, 2008)

### 2.2.6 Definisi *Histogram Of Oriented Gradient*

*Histogram Of Oriented Gradients* ini digunakan untuk mengekstraksi fitur pada obyek gambar dengan menggunakan obyek manusia. Berdasarkan langkahnya, proses awal pada metode HOG adalah mengkonversi citra RGB (*Red, Green, Blue*) menjadi *grayscale*, yang kemudian dilanjutkan dengan menghitung nilai gradien setiap piksel. Setelah mendapatkan nilai gradien, maka proses selanjutnya yaitu menentukan jumlah bin orientasi yang akan digunakan dalam pembuatan *histogram*. Proses ini disebut *spatial orientation binning*. Namun sebelumnya pada proses *gradient compute* gambar pelatihan dibagi menjadi beberapa cell dan dikelompokkan menjadi ukuran lebih besar yang dinamakan block. Sedangkan pada proses normalisasi blok digunakan perhitungan geometri R-HOG. Proses ini dilakukan karena terdapat blok yang saling tumpang tindih. Berbeda dengan proses pembuatan *histogram* citra yang menggunakan nilai – nilai intensitas piksel dari suatu citra atau bagian tertentu dari citra untuk pembuatan

*histogramnya*, berikut adalah tahapan dalam *Histogram Of Oriented Gradient* :  
(Navneet Dalal and Bill Triggs, 2003)

**a) Konversi Citra atau Normalisasi Warna**

Citra *true colour* adalah representasi citra berwarna yang memiliki tiga komponen utama yaitu merah, hijau dan biru (RGB). Masing-masing komponen pada citra *true colour* mempunyai 256 kemungkinan nilai. Citra *grayscale* memiliki 28 (256) kemungkinan nilai pada pikselnya. Nilai tersebut dimulai dari nol untuk warna hitam dan 255 untuk warna putih. Konversi Citra *true colour* ke Grayscale mengubah nilai piksel yang semula mempunyai 3 nilai yaitu *Red, Green, Blue* menjadi satu nilai yaitu keabuan. Berikut persamaan yang digunakan untuk mendapatkan nilai keabuan :

$$L = 0.144 * R + 0.587 * G + 0.299 * B$$

Keterangan :

L	: nilai keabuan pada piksel
0.144	: bobot untuk elemen warna merah (wR)
0.587	: bobot untuk elemen warna biru (wB)
0.299	: bobot untuk elemen warna hijau (wG)
R	: nilai intensitas elemen warna merah
B	: nilai intensitas elemen warna biru
G	: nilai intensitas elemen warna hijau

NTSC (*National Television System Committee*) mendefinisikan bobot untuk konversi citra *true colour* ke *grayscale* sebagai berikut : wR = 0.299, wB = 0.587, wG = 0.114.

Data masukan berupa citra *true colour* dan data keluaran berupa citra *Grayscale*.

**b) Gradient Compute**

Setelah proses konversi citra yaitu mengubah gambar dalam bentuk grayscale langkah selanjutnya adalah menghitung nilai gradien setiap piksel dalam gambar.

**c) Spatial Orientation Binning**

Untuk membuat sebuah *histogram* dibutuhkan nilai gradien dan nilai tersebut didapat dari nilai tiap piksel dalam sebuah gambar. Gambar kemudian akan dibagi menjadi *cells* dengan ukuran yang telah ditentukan. Jadi tiap *cells* dalam gambar akan dibuat histogramnya untuk mengetahui nilai dalam tiap *cells* karena tiap *cell* mempunyai nilai yang berbeda. Dalam pembuatan histogramnya diperlukan adanya bin untuk mengetahui nilai gradiennya. *Bin* akan ditentukan sendiri oleh pengguna. Dalam penelitian sebelumnya bin yang digunakan adalah 4 *bin orientation*

**d) Normalization Block**

Karena nilai gradien mempunyai nilai yang berbeda oleh karena itu diperlukan pengelompokkan *cells* menjadi lebih besar atau yang disebut dengan blok. Blok biasanya tumpang tindih karena setiap *cells* kontribusi nilai lebih dari sekali. Dalam normalisasi blok ini terdapat dua geometri blok utama yaitu blok persegi panjang R-HOG dan melingkar C-HOG akan tetapi dalam penelitian ini yang akan digunakan adalah geometri R-HOG. Hasil akhir dalam normalisasi block ini yaitu fitur. Dalam proses ini, blok yang tumpang tindih diselesaikan dengan R-HOG. Sedangkan dalam blok terdiri dari  $2 \times 2$  *cells*, dalam *detector windows* terdapat  $7 \times 15$  R-HOG dan menggunakan 4 bin orientasi sehingga diperoleh 1680 vektor dalam 1 *detector windows*. Jumlah vektor ini di dapat dari  $2 \times 2 \times 7 \times 15 \times 4$  dan vektor ini yang disebut sebagai fitur. Gambar 2.7 adalah contoh bentuk geometri R-HOG.

**e) Detector Windows**

*Detector windows* merupakan *windows* atau jendela berukuran  $64 \times 128$  yang digunakan untuk jendela pendeteksian. Jendela pendeteksian ini terdiri dari  $8 \times 8$  piksel dalam tiap *cells*.