

BAB II

TINJAUAN PUSTAKA DAN LANDASAN TEORI

2.1. TINJAUAN PUSTAKA

Beberapa penelitian yang telah dijadikan rujukan pada penulisan tugas akhir ini didapatkan dari Nasrullah Ratu B.S.L dan Deby Faisol Akbar (2014), Andrian Wijayana (2015), Muhammad Aulia Firmansyah dan Prof.Dr.Ir.Iping Supriana (2015), serta Muhammad Roynaldi Prabowo (2017) yang telah melakukan penelitian terkait *object tracking* dapat diperlihatkan pada Tabel 2.1.

Tabel 2.1. Penelitian sebelumnya terkait *object tracking*

Peneliti	Tahun	Metode	Hasil
Nasrullah Ratu B.S.L dan Deby Faisol Akbar	2014	1. <i>Background subtraction</i> 2. Kalman Filter	1. 81,27% 2. 85,2%
Andrian Wijayana	2015	ASIFT dan Mean Shift	30%
Muhammad Aulia Firmansyah dan Prof.Dr.Ir.Iping Supriana	2015	1. Multiple Object Tracking 2. Estimasi posisi 3. Deteksi objek	1. 81,72% 2. 87,95% 3. 76,38%
Muhammad Roynaldi Prabowo	2017	Gaussian Mixture 1. Recall 2. Precision	1. 71% 2. 72%

Nasrullah Ratu B.S.L dan Deby Faisol Akbar (2014) melakukan penelitian tentang *Object Tracking Berbasis Background Substraction dan Kalman Filter*, penelitian proses pendeteksian objek bergerak terdapat berbagai macam metode yang digunakan. Beberapa diantaranya yang sering digunakan adalah *background subtraction* dan *Kalman Filter*. Kedua metode ini memiliki fungsi yang berbeda terhadap pengolahan citra. Metode yang digunakan dalam penelitian ini adalah pemrograman dengan bahasa Matlab, sedangkan video yang digunakan adalah video yang dibuat sendiri melalui perangkat perekam video. Penelitian ini berusaha untuk memadukan kedua metode ini untuk membantu dalam *object tracking* pada video. Hasil yang diperoleh adalah dengan metode *substraction* objek berhasil dilakukan *tracking* yaitu sebesar 81,27%, dan dengan penambahan metode *Kalman filter*, estimasi pergerakan objek dapat dilakukan hingga akurasi 85,2%.

Andrian Wijayana. Universitas Telkom (2015) melakukan penelitian tentang penggunaan metode ASIFT dan *Mean Shift* dalam analisis serta implementasi *object tracking*. Penelitian ini bermula dari semakin tingginya kebutuhan manusia akan sistem keamanan berbasis *tracking* yang bekerja secara otomatis. Dalam sistem *object tracking*, ekstraksi ciri menjadi salah satu tugas utama dalam melacak sebuah objek, dalam hal ini ciri yang digunakan diharuskan mampu tahan terhadap berbagai kondisi karena objek selalu bergerak bebas dalam video. Hal ini berarti ekstraksi ciri yang tahan terhadap berbagai kondisi sangat dibutuhkan dalam sistem *object tracking*.

Penelitian ini menggunakan metode *Affine Scale Invariant feature Transform* (ASIFT) yang merupakan metode ekstraksi ciri yang *fully invariant* serta metode *Mean Shift* dimana keduanya bekerja dengan membuat sistem yang dapat melacak objek pada berbagai kondisi objek dimana menggunakan inputan gambar dan video. Inputan ini kemudian diproses mulai dari *oreprocessing* hingga proses *tracking*. Hasil dari penelitian ini menunjukkan bahwa kondisi lingkungan objek terhadap banyaknya kemungkinan objek

terdeteksi dengan tepat. Kondisi lingkungan terkontrol cenderung memiliki hasil yang lebih baik dari lingkungan yang tidak terkontrol. Selain itu, nilai *threshold* dan radius yang digunakan juga sangat mempengaruhi hasil *matching* objek. Berdasarkan penelitian, *threshold* dengan nilai 0,9 dan radius 10% memiliki kecenderungan dapat mendeteksi objek dengan tepat. Hasil penelitian juga menunjukkan bahwa metode ASIFT dan *Mean Shift* dapat mengatasi permasalahan perubahan *point of view* yang terjadi pada objek dengan akurasi 30%.

Muhammad Aulia Firmansyah dan Prof.Dr.Ir.Iping Supriana (2015) melakukan penelitian tentang sistem keamanan rumah yang ditunjang oleh metode *multiple object tracking* dan estimasi posisi. Penelitian ini membahas tentang sistem keamanan rumah yang sering digunakan oleh masyarakat adalah CCTV. Walaupun begitu, masih terdapat kekurangan pada sistem tersebut, seperti penggunaannya yang tidak praktis. Untuk mengetahui kejadian pada rekaman CCTV, rekaman harus diperiksa dari awal hingga akhir. Hal ini akan menjadi masalah apabila rekaman tersebut berdurasi sangat lama. Oleh sebab itu, diajukan sistem yang dapat secara otomatis melacak kejadian yang ada pada rekaman tersebut.

Sistem yang diajukan pada penelitian ini adalah sistem yang melakukan proses *multiple object tracking* dan estimasi posisi. Proses *multiple object tracking* adalah suatu proses melacak seluruh objek bergerak yang terekam dalam video. Pada sistem ini, proses *multiple object tracking* terbagi menjadi beberapa tahap. Tahap-tahap tersebut ialah tahap *proprocessing*, tahap *detection*, dan tahap *tracking*. Proses estimasi posisi adalah suatu proses melakukan estimasi terhadap posisi objek pada koordinat nyata berdasarkan koordinat objek pada video. Prototipe sistem dibuat dan kemudian diuji pada beberapa rekaman CCTV. Parameter pengujiannya adalah akurasi sistem dalam melacak suatu objek dan rata-rata waktu yang dibutuhkan untuk melacak satu buah objek. Dari pengujian tersebut, didapatkan rata-rata akurasi deteksi

sebesar 76,38%, akurasi tracking sebesar 81,72%, dan akurasi estimasi posisi sebesar 87,95% dengan rata-rata waktu pelacakan per objek.

Muhammad Roynaldi Prabowo (2017) melakukan sebuah penelitian yang memberikan output berupa sistem pendeteksi objek pada video di dalam air. Video tersebut diambil dengan menggunakan sebuah *action camera*. Penelitian ini merupakan langkah awal dalam pengembangan teknologi penangkapan ikan, dan berbagai sistem pengawasan bawah air lainnya. Proses pendeteksian objek dalam penelitian ini dilakukan dengan cara mengurangkan nilai-nilai piksel pada *frame* saat ini dengan nilai-nilai piksel pada model background sebelumnya. Pada pendeteksian objek didalam air, background selalu berubah-ubah karena pengaruh intensitas cahaya dan gelombang arus bawah air. Maka dari itu diperlukan sebuah metode pembuat latar belakang yang dapat beradaptasi, yaitu metode *Gaussians Mixture Model*.

Pada metode ini nilai sebuah piksel latar belakang ditentukan oleh nilai intensitas piksel-piksel sebelumnya, dan apabila terjadi perubahan intensitas piksel yang drastis dan mengakibatkan nilai piksel di luar nilai ambang, maka piksel ini akan diidentifikasi sebagai objek, kemudian piksel latar belakang yang baru akan diubah berdasarkan nilai piksel ini. Setelah model background didapatkan, maka akan dibandingkan dengan *frame* selanjutnya, proses perbandingan ini dilakukan pada piksel-piksel di setiap *frame* hingga *frame* terakhir pada sebuah citra video. Video diambil di tiga kondisi berbeda yaitu pagi, siang, dan malam hari. Tingkat keberhasilan metode ini dievaluasi berdasarkan nilai *recall* dan *precision*, dimana diperoleh nilai rata-rata dari ketiga video sebesar 71% untuk nilai *recall*-nya, dan 72% untuk nilai *precision*-nya.

2.2. LANDASAN TEORI

2.2.1. Citra

Citra merupakan suatu representasi (gambar), kemiripan, atau imitasi dari suatu objek. Citra sebagai keluaran suatu sistem perekaman data bersifat optic berupa foto, bersifat analog berupa sinyal-sinyal video seperti gambar pada monitor televisi, atau bersifat digital yang dapat langsung disimpan pada suatu media penyimpanan.

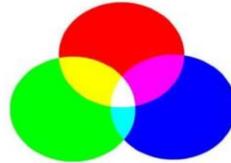
Secara harfiah, gambar pada bidang dwimarta (dua dimensi) disebut citra (*image*). Jika dilihat dari sudut pandang matematis, fungsi *continue* dari suatu intensitas cahaya pada bidang dwimarta merupakan sebuah citra. Sumber cahaya yang menerangi objek, kemudian dipantulkan kembali sebagai berkas cahaya tersebut. Pantulan cahaya tersebut kemudian ditangkap oleh alat-alat optik seperti mata manusia, kamera, pemindai (*scanner*), dan sebagainya. Sehingga dapat diketahui bahwa hakikat bayangan objek merupakan sebuah citra yang terekam. (Permadi dan Murinto, 2015)

Citra pada beberapa standar umumnya merupakan tampilan suatu titik yang berada pada ruang tiga dimensi. Ada beberapa standar yang digunakan pada citra antara lain *Red, Green, Blue* (RGB), *Hue, Saturation, Value* (HSV), dan *Hue, Luminosity, Saturation* (HLS). Standar yang akan digunakan pada penelitian ini adalah *Red, Green, Blue* (RGB) dimana memiliki tiga indeks warna sesuai namanya, yaitu merah, hijau, dan biru. Ketiga indeks warna tersebut memiliki kriteria indeks masing-masing yaitu antara 0-255 atau 0-256 bit. Dari citra berwarna yang memiliki tiga ruang dimensi dapat disederhanakan menjadi satu dimensi *grayscale*. Sebuah citra keabuan yang memiliki nilai antara 0-255 disebut dengan *grayscale*. Nilai tersebut menunjukkan tingkat derajat keabuan atau kecerahan dari citra (0=hitam/gelap dan 255=putih/terang). Untuk mengkonversi citra RGB

menjadi *Grayscale* dapat dilakukan dengan perhitungan mencari nilai rata-rata antara ketiga indeks dalam citra RGB. Dari hasil konversi citra yang telah dilakukan ke dalam RGB maka nilai dalam indeks *Grayscale* dapat diasumsikan mewakili nilai dan informasi yang terkandung dalam citra RGB.

a. Citra Warna

Citra warna dapat direpresentasikan ke dalam beberapa kanal (*channel*) seperti yang ditunjukkan Gambar 2.1. yang dapat menyatakan komponen-komponen warna penyusunnya.

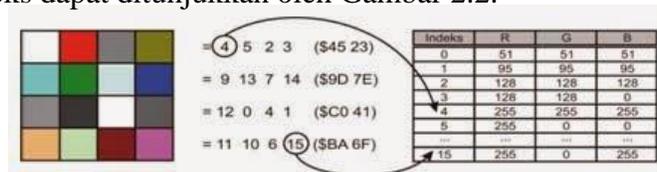


Gambar 2.1 Citra warna (true color)

Banyaknya kanel (*channel*) yang digunakan tergantung pada model warna yang akan digunakan pada citra tersebut. Kemudian intensitas suatu titik pada sebuah citra berwarna adalah suatu kombinasi dari tiga intensitas yaitu suatu derajat keabuan merah ($f_{merah}(x,y)$), hijau ($f_{hijau}(x,y)$), dan biru ($f_{biru}(x,y)$). (Wiryadinata, dkk. 2016)

b. Citra Warna Berindeks

Pada format citra warna indeks merupakan suatu format yang memberikana informasi setiap titik yang merupakan indeks dari suatu tabel yang berisi informasi warna yang tersedia, yang disebut palet warna (*color map*). Adapun untuk *detail* gambar warna citra berindeks dapat ditunjukkan oleh Gambar 2.2.

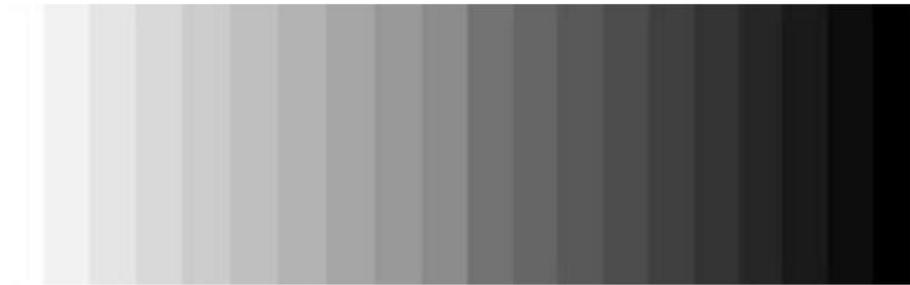


Gambar 2.2 Citra warna berindeks

Jumlah bit yang dibutuhkan untuk mendapatkan informasi setiap titik bergantung pada jumlah warna yang tersedia dalam palet warna.

c. Citra *Grayscale*

Grayscale merupakan warna-warna pixel yang memiliki rentang gradiasi warna hitam dan putih. Citra skala keabuan dapat ditunjukkan oleh Gambar 2.3.

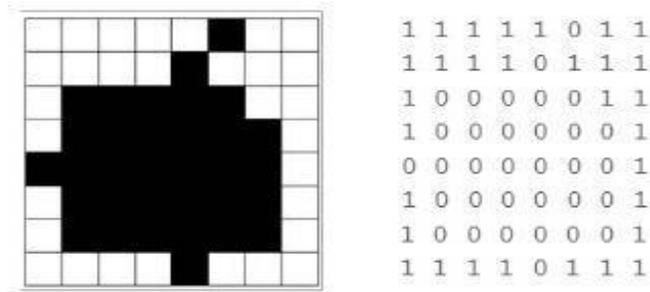


Gambar 2.3 Citra skala keabuan (*grayscale*)

Derajat keabuan merupakan format pada citra ini, karena terdapat warna abu-abu diantara warna minimum (hitam) dan warna maksimum (putih). (Wiryadinata, dkk. 2016)

d. Citra Biner

Suatu citra biner hanya memiliki dua nilai keabuan yaitu 0 dan 1. Maka 1 bit sudah cukup untuk merepresentasikan nilai pixel yang ditunjukkan pada Gambar 2.4.



Gambar 2.4 Citra biner dan representasinya dalam data digital

Citra biner didapat dari proses pemisahan piksel berdasarkan derajat keabuan yang dimilikinya. Pada proses peminoran dilakukan dengan cara membulatkan ke atas ataupun ke bawah untuk setiap nilai suatu keabuan dari piksel yang berada di atas atau bawah harga ambang. *Thresholding* adalah suatu metode untuk menentukan besarnya nilai ambang. Apabila nilai yang terdeteksi masih di dalam nilai batas ambang maka diumpangkan nilainya 1 (putih). Namun apabila nilai di luar ambang batas maka diumpangkan nilai tersebut 0 (hitam) (Wiryadinata, dkk. 2016). Maka hasil citra setelah di *thresholding* tersebut akan hitam putih karena terdapat batas ambang seperti yang ditunjukkan pada Gambar 2.5.



Gambar 2.5 Hasil *Thresholding*

2.2.2. Pengolahan Citra

Pengolahan citra digital (Digital Image Processing) merupakan ilmu yang mempelajari tentang teknik-teknik mengolah citra. Sebuah citra yang dimaksud adalah objek pada video yang diam (*image*) ataupun objek pada video yang bergerak yang berasal dari webcam/kamera. Yang dimaksud dengan digital adalah pengolahan citra atau gambar yang dilakukan secara digital menggunakan bantuan computer. Pada dasarnya sebuah sistem pengolahan citra terdiri dari objek yang akan diproses yang akan digunakan

untuk merepresentasikan objek ke dalam bentuk citra digital, hasil dari pemrosesan tersebut kemudian menghasilkan suatu citra digital baru untuk dianalisis sehingga proses analisis citra yang dilakukan mendapatkan hasil yang diinginkan.

Pengolahan citra digital memiliki beberapa tujuan antara lain adalah memperbaiki kualitas gambar dilihat dari aspek *radiometric* (peningkatan kontras, transformasi warna, restorasi citra) serta dari aspek *geometric* (rotasi, translasi, skala, transformasi *geometric*), melakukan proses penarikan informasi dan deskripsi objek atau pengenalan objek yang terkandung pada citra, melakukan kompresi atau reduksi data untuk tujuan penyimpanan data, transmisi data, dan waktu proses data.

2.2.3. Image Stitching

Image Stitching merupakan salah satu teknik pengolahan citra digital yaitu *perspective warp* untuk menemukan suatu bagian pada citra yang memiliki kemiripan dengan citra *template* (acuan). *Image stitching* ini merupakan salah satu ide yang digunakan untuk menjelaskan bagaimana otak kita mengenali kembali bentuk-bentuk atau pola-pola. *Template* dalam konteks rekognisi pola menunjuk pada konstruk internal yang jika cocok (*match*) dengan stimulus penginderaan mengantar pada rekognisi suatu objek, dimana pengenalan pola dapat terjadi kesesuaian antara stimulus indera dengan bentuk mental internal. Gagasan ini mendukung bahwa sejumlah besar *template* telah tercipta melalui pengalaman hidup kita. (Efraim,dkk. 2015)

2.2.4. Motion Detection

Untuk mendeteksi objek bergerak, kita memilih background gambar dari gambar sequence pertama. Kemudian masing-masing gambar dalam proses dikurangi oleh gambar latar belakang. Secara teoritis, daerah nol dalam

hasil pengurangan adalah lokasi dari gerak objek. Bagaimanapun juga, ada beberapa hal yang lain yang muncul dalam percobaan ini yang dapat menjadikan perubahan background secara keseluruhan pada percobaan ini. Beberapa hal tersebut antara lain adalah: *noises*.

Perubahan cahaya dan gerakan bayangan benda. Beberapa hal ini dapat menyebabkan video tracking menjadi terlihat mengalami gangguan oleh karena itu diperlukan sebuah video dengan background tetap dan memiliki kontras warna yang cukup baik. Metode yang kami usung terdiri dari dua tahap utama yaitu: deteksi objek dan pelacakan objek. (Nasrulloh dan Deby, 2014)

2.2.5. OpenCV (Open Computer Vision)

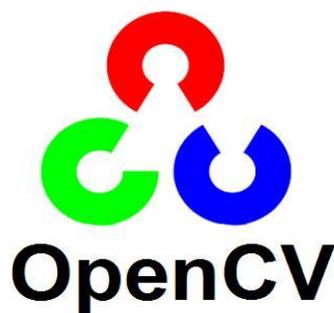
OpenCV (Open Computer Vision) merupakan sebuah open source library dari fungsi pemrograman yang terutama fungsi ini ditujukan untuk keperluan pengolahan citra secara realtime, yang dimana openCV ini dikembangkan oleh sebuah pusat penelitian Intel Rusia di Nizhny Novgorod, dan sekarang sudah didukung oleh Willow Garage dan Iteseez. OpenCV sendiri sudah memiliki lebih dari 500 algoritma yang sudah dioptimalkan untuk keperluan pengolahan citra dan video.

OpenCV dibuat untuk menyelesaikan kebanyakan masalah computer vision yang sudah ditentukan. Untuk mencapai tingkat optimasi yang tinggi diperlukan ketegasan yang merupakan aspek dari OpenCV. Hal tersebut dimaksudkan untuk aplikasi real-time dan dirancang untuk sebuah proses eksekusi yang sangat cepat. OpenCV sendiri memiliki beberapa modul yang dapat digunakan untuk proses langsung sesuai fungsionalitasnya yang dipaparkan pada table dibawah ini.

Intel Open Source Computer Vision Library merupakan kepanjangan dari OpenCV yang sekurang-kurangnya terdiri dari tiga ratus fungsi-fungsi C, bahkan bias lebih. OpenCV dapat dioperasikan pada

computer berbasis windows ataupun linux. Ada beberapa contoh aplikasi dari OpenCV adalah pada Human Computer Interaction(interaksi manusia computer) antara lain Object Identification (identifikasi objek), Segmentation (segmentasi) dan Recognition (pengenalan), Face Recognition (pengenalan wajah), Gesture Recognition (pengenalan gerak isyarat), Motion Tracking (penjajakan gerakan), Ego Motion (gerakan ego) dan Motion Understanding (pemahaman gerakan), Structure From Motion (gerakan dari struktur), Mobile Robotics (robot-robot yang bergerak)

OpenCV menggunakan sebuah tipe face detector yang disebut Haarcascade Classifier. Face detector akan bekerja jika ada sebuah image yang berasal dari file gambar atau live video, face detector tersebut akan menguji tiap lokasi image dan akan mengklasifikasi file image atau live video tersebut sebagai wajah atau bukan wajah. Dalam mengklasifikasi wajah tersebut digunakan sebuah pemisalan skala yang tetap, contohnya 50×50 pixel. Jika sebuah wajah pada image lebih besar ataupun lebih kecil dari ukuran pixel yang sudah ditentukan, maka classifier akan terus menerus jalan beberapa kali untuk mencari wajah pada gambar tersebut. Logo dari OpenCV dapat ditunjukkan pada Gambar 2.6.



Gambar 2.6 logo OpenCV

Dalam melakukan tugasnya classifier menggunakan data yang sudah disimpan pada file XML untuk memutuskan bagaimana cara mengklasifikasi tiap lokasi image. OpenCV akan menggunakan 4 data XML untuk mendeteksi wajah depan, dan satu untuk wajah profile.

2.2.6. Python

Python merupakan salah satu bahasa pemrograman yang bersifat interpreter, interaktif, object-oriented dan dapat beroperasi hampir di semua platform, seperti keluarga UNIX, Mac, Windows dan lainnya. Python termasuk dalam salah satu bahasa pemrograman yang mudah untuk dipelajari karena sintaksnya yang jelas dan elegan, yang dikombinasikan dengan menggunakan module-module siap pakai dan struktur data yang tingkat tinggi dan efisien. Bahasa pemrograman python sendiri diciptakan di Belanda pada tahun 1990 oleh Guido Van Rossum dan namanya diambil dari acara televisi kesukaan. Guido yaitu Monty Python's Flying Circus. Kemudian python digunakan sebagai bahasa pemrograman yang dipakai secara luas dalam industry dan Pendidikan karena sederhana, ringkas, sintaks intuitif dan memiliki pustaka yang luas. Logo Python dapat ditunjukkan pada Gambar 2.7.



Gambar 2.7 Logo Python

Python juga merupakan pemrograman yang berorientasi pada objek (OOP). Data yang terdapat dalam python adalah sebuah objek yang telah dibuat dari kelas (*class*). Pada pemrograman yang berorientasi pada objek merupakan suatu alat ampuh yang mampu untuk mengembangkan perangkat lunak yang dapat digunakan kembali.

2.2.7. Kamera Webcam

Web camera (Webcam) merupakan suatu sebutan bagi sebuah camera real-time yang gambarnya dapat diakses atau dilihat melalu *World Wide Web*, program *instant messaging*, atau aplikasi *video call*. Webcam merupakan sebuah kamera video digital kecil yang dapat dihubungkan kekomputer (PC) melalui port USB ataupun port COM, ada juga yang sudah tidak memakai USB dan sudah tetanam langsung di laptop. Kamera webcam yang digunakan dalam penelitian ini adalah kamera webcam Logitech C270 seperti ditunjukkan pada Gambar 2.8.



Gambar 2.8 kamera webcam C270

Webcam memiliki beberapa tipe seperti Slim1320 (True 1.3 Mega Pixels High Performance Webcam), Slim 2020AF (Mega Pixel Webcam Auto Focus), Eye 312 (Simplify Instant Video and Chat), Eye 110 (Instant Video Messenger Webcam) serta i-Look 1321 (Advance 1.3 Mega Pixel Camera), dll. Sekarang hampir semua kamera digital dan HP bisa dijadikan sebagai kamera web (webcam). (Andre, 2016)