

BAB II

TEORI COMPUTER VISION

2.1. Computer Vision

Computer vision adalah teknologi yang berkembang dengan sangat pesat dan telah banyak digunakan dalam kehidupan sehari-hari manusia mulai dari hal yang sederhana, sampai yang tidak biasa seperti UAV (pesawat tidak berawak). *Computer vision* adalah teknologi yang membuat komputer mampu melihat. Teori *computer vision* berasal dari teori kecerdasan buatan, yaitu mengambil informasi dan data dari citra yang diambil dengan menggunakan suatu algoritma. Data citra itu sendiri dapat berupa berbagai bentuk, tidak hanya berasal dari video kamera atau citra fotografi mentah, tetapi juga dari sensor panas atau pemindai medis. Tujuan dari bidang studi *computer vision* adalah memprogram komputer agar “memahami” sebuah gambaran atau ciri dari suatu citra. Secara lebih spesifik, tujuan dari penelitian *computer vision* termasuk pendeteksian dari suatu objek tertentu dari suatu citra, evaluasi dari hasil pengambilan citra, melacak objek dari suatu urutan citra, memetakan sebuah gambaran dalam citra menjadi model tiga dimensi, dan lain-lain. Tujuan-tujuan ini dapat dicapai dengan metode *pattern recognition*, pembelajaran statistik, *projective geometry*, pengolahan citra, dan teori-teori lain dalam *computer vision*. Tujuan lain dari *computer vision* adalah menghasilkan suatu keputusan dari objek yang nyata secara fisik dan suatu gambaran pengindraan dari suatu citra. Beberapa ahli berpendapat, tujuan dasar dari *computer vision* adalah membangun sebuah deskripsi dari sebuah citra, karena untuk menghasilkan suatu keputusan dari objek nyata, adalah suatu hal yang penting untuk membangun suatu deskripsi dari citra yang ditangkap. Studi *computer vision* termasuk didalamnya pengindraan dari objek, menerjemahkan informasi atau data yang dimasukkan ke dalam komputer, dan algoritma yang digunakan untuk memproses informasi citra (Shapiro dan Stockman, 2000:13). Sedangkan menurut definisi lain (Gonzalez dan Woods, 2001:2), tujuan utama dari studi *computer vision* adalah berusaha menyamai atau melebihi kemampuan

penglihatan manusia, termasuk didalamnya pembelajaran dan kemampuan untuk menarik kesimpulan serta mengambil aksi berdasarkan masukan visual.

Contoh dari penerapan *computer vision* adalah pencitraan medis yang telah banyak digunakan di rumah sakit-rumah sakit yang bernilai jutaan US dollar per tahun. Contoh lain dari *computer vision* digunakan di sistem-sistem berikut:

- Pengendalian proses
- Pendeteksian event/kejadian
- Pengorganisasian informasi
- Pemodelan objek atau lingkungan sekitar
- Interaksi
- Sistem informasi geografis

Penelitian tentang *computer vision* telah dimulai pada tahun 1970, dimana pada saat itu ada suatu penelitian yang bertujuan untuk meniru kecerdasan manusia dan memberikan robot kecerdasan untuk beberapa tujuan (Szeliski, 2009:11). Tetapi terkadang manusia tertipu dengan persepsinya tentang kemampuannya untuk melihat sesuatu, jadi orang sering berpikir adalah hal yang mudah untuk membuat sebuah aplikasi *computer vision* yang mampu menjelaskan apa yang dilihatnya (Szeliski, 2009). Setelah melalui banyak penelitian, akhirnya para ahli menemukan masalah-masalah yang ada dari studi *computer vision*. Walaupun studi *computer vision* telah dimulai sejak tahun 1970, bidang *computer vision* bisa disebut masih belum matang dan masih terlalu luas. Karena pada saat pengembangan awalnya, belum ada formula standar untuk memecahkan problem dalam *computer vision* dan metode bagaimana masalah tersebut diselesaikan.

Masalah yang membuat penerapan *computer vision* menjadi sulit adalah keterbatasan dari perangkat keras yang dipakai untuk mengambil citra dari objek di lapangan. Seringkali objek yang ditangkap terganggu oleh *noise* dari lingkungan sekitar objek. Untuk menyingkirkan *noise* tersebut, para ahli menggunakan algoritma tertentu yang dapat mengekstrak objek yang diinginkan, atau hanya meminimalisir *noise*.

Manusia dapat tertipu oleh kemampuan penglihatannya sendiri, tanpa mempertimbangkan proses yang sangat rumit dibaliknya. Otak manusia membagi sinyal

penglihatan menjadi banyak kanal yang menyalurkan informasi-informasi yang berbeda dan bermacam-macam. Otak manusia mempunyai sistem yang dapat mengidentifikasi bagian penting dari suatu citra untuk diperiksa dengan tanpa terganggu oleh citra objek lain. Hal ini sama dengan ketika manusia ingin fokus terhadap sesuatu. Bagaimanapun juga, dalam sistem *computer vision*, komputer menerima suatu *grid* yang berisi angka-angka dari kamera atau harddisk. Pada sebagian besarnya, tidak ada pengenalan objek yang bersifat *built-in*, tidak ada kendali otomatis atas fokus dan lobang lensa, tidak ada asosiasi silang dari pengalaman selama bertahun-tahun, dan ini yang membedakan dengan sistem penglihatan manusia. Dapat dikatakan bahwa sistem ini masih sangat naif. Dari angka-angka yang dikirimkan oleh perangkat keras perekam citra ke dalam komputer, baru dapat digunakan teknik-teknik dari *computer vision* untuk diolah menjadi suatu data atau informasi yang berguna (Bradski dan Kaehler, 2008:2-3).

Permasalahan lain dari *computer vision* adalah pose dari objek yang ditangkap, yang umumnya sangat sulit untuk dipecahkan. Disini belum ada suatu cara unik untuk membangun kembali sinyal tiga dimensi yang ditangkap, dan belum ada solusi secara definitif untuk memecahkan problem pose ini (Bradski dan Kaehler, 2008). Selain itu, data yang didapat seringkali menjadi rusak karena *noise* dan distorsi. Kerusakan ini muncul karena variasi seperti cuaca, pencahayaan, refleksi, dan pergerakan. Ada juga masalah lain seperti ketidaksempurnaan lensa dan pengaturan mekanisnya, dan *noise* elektrikal dari sensor. Halangan-halangan ini merupakan tantangan besar dari pengembangan *computer vision*.

Seperti yang dijelaskan Szeliski, ada 3 tingkatan deskripsi dari sistem pemrosesan informasi visual :

- Teori komputasi : Apakah tujuan dari tugas komputasi dan batasan-batasan apa yang telah diketahui atau dapat dibawa pada pemecahan masalah.
- Representasi dan algoritma : Bagaimana representasi dan algoritma dapat dipetakan ke dalam perangkat keras aktual seperti sistem penglihatan biologikal atau kepingan silikon? Sebaliknya, bagaimana batasan pada perangkat lunak dapat digunakan untuk memilih representasi dan algoritma? Dengan meningkatnya penggunaan chip grafis dan banyak arsitektur inti

untuk digunakan pada *computer vision*, pertanyaan ini menjadi relevan pada pengembangan *computer vision*.

Computer vision sendiri adalah suatu bidang yang sangat luas dan mempunyai banyak aspek, dan teknik-teknik yang dapat digunakan bermacam-macam, tergantung permasalahan dan tujuan pembuatan aplikasi *computer vision*.

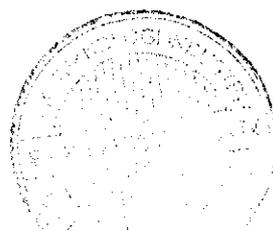
Aspek yang termasuk dalam studi *computer vision* adalah :

- a. *Image Processing*
- b. *Image Formation*
- c. *Image Segmentation*
- d. *Motion Detection and Tracking*
- e. *Feature Detection and Matching*
- f. *Projective Geometry*
- g. *Image Stitching and blending*
- h. *Optical Flow and Tracking*
- i. *Object Recognition*
- j. *Stereo Matching*
- k. *Multi-view Stereo and 3D Modelling*
- l. *Image-based Rendering*
- m. Dan lain-lain

Sesuai relevansi pada penelitian ini, yang akan dijelaskan hanya mencakup pengolahan citra (*image processing*) dan pengenalan objek.

2.2. Pengolahan Citra

Pengolahan Citra atau *image processing* adalah suatu ilmu untuk memanipulasi suatu citra (Crane, 1997:1). Ilmu ini mencakup bidang yang sangat luas dari teknik yang telah ada banyak dipakai pada aplikasi-aplikasi saat ini. Teknik-teknik ini beragam, mulai dari mengubah suatu citra, menyoroti suatu bagian tertentu dari citra, membuat citra baru dari bagian suatu citra lain, memperbaiki citra yang telah turun kualitasnya selama pengambilannya atau setelahnya. Studi pada pengolahan citra sendiri bermula dari dua area aplikasi : perbaikan dari informasi



yang berasal dari citra untuk kepentingan penafsiran manusia, dan pemrosesan data citra untuk penyimpanan, transmisi, dan representasi yang ditujukan untuk penglihatan mesin (Gonzalez dan Woods, 2001).

Tahap pertama dari sebagian besar aplikasi *computer vision* sendiri adalah penggunaan teknik-teknik pengolahan citra untuk melakukan pra-pemrosesan citra dan mengkonversinya menjadi bentuk yang memungkinkan untuk analisa lebih jauh lagi. Salah satu contoh dari operasi pengolahan citra adalah koreksi pencahayaan dan penyeimbangan warna, reduksi *noise*, peningkatan ketajaman, atau melakukan rotasi pada citra. Ketika beberapa menganggap bahwa dalam studi pengolahan citra tidak berhubungan dengan studi *computer vision*, justru sebagian aplikasi *computer vision* seperti *computational photography* atau bahkan pengenalan objek memerlukan desain yang cermat dari penggunaan teknologi pengolahan citra untuk mendapatkan hasil yang maksimal (Szeliski, 2009).

Pengolahan citra sering disamakan dengan komputer grafis, karena komputer grafis dan pengolahan citra adalah dua teknologi yang saling berkaitan. Akan tetapi, walaupun disini banyak kesamaan konsep antara pengolahan citra dan komputer grafis, keduanya adalah dua studi yang berbeda. Komputer grafis adalah penciptaan citra yang murni buatan, atau bisa disebut sintesis. Sedangkan pengolahan citra adalah manipulasi dari citra yang sudah ditangkap atau direkam sebelumnya. Komputer grafis bekerja dengan objek 2 dimensi dan 3 dimensi. Pengolahan citra biasanya bekerja dengan gambar 2 dimensi, walaupun tidak terbatas pada itu saja. Aplikasi pada industri diagnosa medis melakukan pengolahan citra pada data 3 dimensi. Dengan munculnya beberapa teknologi seperti "*volume visualization*" dan "*morphing*", garis pembatas antara komputer grafis dan pengolahan citra menjadi kabur. Sebuah citra didefinisikan sebagai fungsi dua-dimensi, yaitu $f(x,y)$, dimana x dan y adalah koordinat spasial, dan amplitudo dari f atas pasangan koordinat (x,y) apa saja disebut sebagai intensitas atau *gray level* (tingkatan abu-abu) dari citra pada poin tersebut. Ketika x,y , dan nilai amplitudo dari f secara keseluruhan adalah bilangan terbatas, kuantitas diskrit, maka gambar tersebut disebut gambar digital. Perlu dicatat bahwa citra digital terdiri dari sejumlah elemen terbatas, dimana setiap elemennya mempunyai lokasi dan nilai tertentu. Elemen-elemen ini merujuk pada

elemen citra, elemen gambar, dan juga *pixel*. Disini ada empat klasifikasi dari algoritma pengolahan citra : proses poin, proses area, proses geometrik, dan proses frame. Suatu proses poin memodifikasi sebuah nilai *pixel* berdasar pada posisi atau nilai awal dari *pixel* tersebut. *Pixel* adalah istilah yang banyak dipakai untuk menotasikan elemen dari citra digital. Suatu proses area memodifikasi nilai dari *pixel* berdasar dari nilai aslinya dan juga nilai dari *pixel* sekitarnya. Proses geometrik mengubah posisi atau susunan dari *pixel*. Suatu proses frame menciptakan nilai *pixel* berdasar dari operasi pada 2 atau lebih citra. Seringkali, efek yang diinginkan membutuhkan aplikasi dari beberapa proses yang berbeda. Sebuah alat yang cukup berguna untuk menganalisa dan kadang juga mempercepat operasi berbasis area atau proses area adalah algoritma Fourier Transform. Operasi berbasis area dapat mengolah data menjadi piramid citra (*image pyramids*) dan gelombang (*wavelets*), yang mana merupakan hal yang sangat penting untuk menganalisa citra dengan resolusi atau skala yang berbeda-beda dan untuk mempercepat beberapa operasi. Kelas penting lain dari operasi global image processing adalah transformasi geometrik seperti rotasi. Optimisasi global adalah salah satu pendekatan dari pengolahan citra, dimana operasi ini melibatkan estimasi optimal menggunakan model Markov Random Field dari Bayesian.

Penglihatan adalah bagian paling maju dari indera manusia, jadi tidak mengherankan jika citra memainkan peranan paling penting dalam pemahaman manusia. Bagaimanapun juga, tidak seperti manusia yang terbatas pada spektrum pita visual elektromagnetik (EM), mesin-mesin pencitraan mencakup sebagian besar spektrum elektromagnetik, mulai dari sinar gamma sampai gelombang radio. Mesin-mesin ini mampu beroperasi pada citra yang diciptakan oleh sumber dimana manusia tidak biasa berhubungan dengan citra-citra seperti ini. Dalam hal ini termasuk didalamnya adalah suara ultrasonik, mikroskop elektron, dan citra yang diciptakan komputer. Dengan demikian, pengolahan citra mencakup bidang aplikasi yang sangat beragam. Contoh aplikasi dari pengolahan citra :

a. Sains dan ruang angkasa

Teknik rekonstruksi dan filtering membuang "*noise*" dan mengembalikan bagian yang hilang dari citra satelit dan pesawat luar angkasa.

b. Film-film

Pada dunia perfilman, teknik pengolahan citra digunakan untuk mengubah suatu citra menjadi citra lain, membuang objek tertentu dari frame, dan membuat frame baru dengan mengkomposisikan bagian-bagian dari frame yang lain.

c. Pengkomposisian Citra

Pengkomposisian citra atau image composition juga digunakan pada film "Forrest Gump". Pemain utamanya telah ditambahkan pada segmen pada beberapa film dokumentasi sejarah, yang didalamnya banyak memunculkan presiden Amerika.

d. Industri medis

Industri medis telah lama menjadi pengguna dari pengolahan citra. Disini ada beragam teknologi pencitraan yang digunakan termasuk didalamnya adalah X-ray dan ultrasonik. Tomografi berbantu komputer (CT atau CAT), tidak banyak dipakai sampai tahun 1980-an. Dokter-dokter menggunakan pemindai CT untuk mendiagnosa bagian dari tulang seperti tulang belakang, tengkorak, dan tulang panggul. Magnetic Resonance Imaging (MRI) menghasilkan gambaran-gambaran dari jaringan lembut seperti syaraf tulang belakang dan juga jantung. Positron Emission Tomography (PET) menunjukkan pengukuran dari proses fisiologi dan kimia tubuh. Magnetic Source Imaging (MSI) mengamati sinyal listrik di dalam otak ketika berpikir dan melakukan fungsi motorik. Semua hasil pencitraan ini ketika telah disimpan di komputer dapat ditingkatkan dan dimanipulasi, sehingga dokter bisa berfokus pada area-area yang penting.

Teknik-teknik pengolahan citra juga banyak digunakan di bidang-bidang lain seperti fotografer, agen periklanan, dan penerbit. Dalam bidang pemetaan juga banyak digunakan teknologi pengolahan citra.

2.3. Pengenalan Pola dan Pengenalan Objek

Pengenalan objek merupakan salah satu permasalahan tersulit dalam studi *computer vision*. Sampai sekarang, belum ada algoritma serbaguna yang memungkinkan pembelajaran otomatis dari objek 3 dimensi yang berubah-ubah dan juga pengenalannya dan lokalisasi dari sebuah *scene* yang kompleks. Pendekatan untuk menangani tugas-tugas *computer vision* tingkat tinggi pada dasarnya

didominasi oleh metode pengenalan objek berdasarkan model (Hornegger, et. al., 2000).

Pada masa lalu, banyak aplikasi telah menghasilkan banyak tipe dari representasi model objek yang memungkinkan implementasi dari sistem pengenalan objek yang sangat bagus. Banyak algoritma pengolahan citra dan *computer vision* yang digolongkan atas solusi yang berlaku secara ad-hoc. Kebanyakan teknik ini menerapkan ide-ide intuitif yang secara spesifik diperuntukan aplikasi tertentu dan mengabaikan eksplorasi dari definisi model matematika yang akurat. Disini tidak ada teori formalisasi tunggal yang menyediakan *framework* untuk analisis analitikal dari keseluruhan sistem yang telah didesain. Untuk sebagian besar, hanya tinggal studi empirikal yang dapat membenarkan pemakaian pola representasi tertentu (Hornegger, et. al., 2000).

Dari semua tugas visual yang komputer lakukan, menganalisa suatu *scene* dari citra dan mengenali objek-objek tertentu dalam citra tersebut adalah hal yang masih menantang. Ketika komputer mencapai keakurasian pada saat membuat model 3 dimensi dari suatu bentuk sebuah adegan tertentu yang diambil dari sudut pandang yang berbeda-beda, komputer tidak dapat menyebutkan nama dari semua objek dan binatang yang muncul pada suatu gambar yang berusia 2 tahun, misalnya. Disini bahkan tidak ada suatu persetujuan diantara peneliti tentang bagaimana tahapan pelaksanaan ini dapat tercapai.

Pengenalan objek merupakan teknik yang sangat sulit, hal ini dikarenakan dunia nyata tersusun dari kumpulan objek yang campur aduk, dimana yang satu bercampur dengan yang lainnya dan tampil dengan berbagai pose. Selanjutnya, keragaman intrinsik dalam suatu kelas, seperti anjing misalnya, berdasar pada artikulasi kompleks yang tidak rigid dan variasi ekstrim dari bentuk dan rupa diantara ras yang berbeda-beda, membuatnya sangat sulit untuk dilakukan pencocokan secara mendalam dari basis data objek.

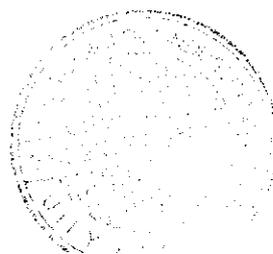
Permasalahan pengenalan dapat dipecah menjadi beberapa tahap. Contohnya, ketika sudah ditemukan apa yang akan dicari, maka permasalahan yang ada adalah deteksi objek, dimana hal ini melibatkan pemindaian cepat sebuah citra untuk menebak dimana kecocokan dapat muncul (Szeliski, 2009). Jika telah ada suatu objek

spesifik yang rigid, maka disini dapat dicoba dilakukan pengenalan, atau biasa disebut *instance recognition*, dimana disini dapat dicari poin dari ciri karakteristik dan kemudian dilakukan verifikasi bahwa pose dari citra yang ada telah cocok secara geometris.

Hal yang paling menantang dari pengenalan objek adalah pengenalan objek secara keseluruhan, dimana hal ini melibatkan pengenalan *instance* dari objek yang bervariasi secara ekstrim, seperti binatang misalnya (Szeliski, 2009). Teknik mungkin yang bergantung secara murni kepada kemunculan objek (biasa dikenal sebagai model algoritma '*bag of words*'), posisi relatif objek-objek tersebut (model '*part-based*'), atau bisa juga melibatkan segmentasi dari suatu citra menjadi suatu bagian-bagian yang mempunyai arti secara semantik. Di seluruh *instance* yang ada, pengenalan sangat bergantung pada konteks objek-objek yang mengelilinginya dan juga elemen-elemen dari adegan citra tersebut (Szeliski, 2009).

Sebagian besar permasalahan computer vision telah cocok pada pola standar dari permasalahan pengenalan objek seperti klasifikasi dan regresi. Sebuah citra digital f secara matematis dikenali sebagai suatu matrix dari nilai intensitas diskrit $f = [f_m, n] \ 1 \leq m \leq M, \ 1 \leq n \leq N$. Untuk pengolahan lebih lanjut, sebagian pekerjaan computer vision tingkat tinggi memerlukan pelabelan atau segmentasi dari citra-citra yang diolah (Hornegger, et. al., 2000). Berdasarkan hasil label-label dan segmentasi ini, pengenalan dan estimasi pose sebagai parameter objek harus diselesaikan. Penyelesaian dari pengenalan objek dan permasalahan estimasi pose dapat menggunakan *framework* probabilistik.

Ada 2 metode untuk melakukan pengenalan, yaitu deteksi suatu objek dengan menggunakan pencocokan pola dan pendeteksian bentuk menggunakan teknik transformasi Hough (Ritter dan Wilson, 1996). Salah satu metode yang paling pokok dari pendeteksian suatu objek adalah dengan pencocokan pola yang menggunakan acuan. Pada pencocokan acuan, replika dari suatu objek yang akan dicari dibandingkan dengan seluruh objek di dalam citra. Jika pencocokan pola antara acuan dan objek dalam citra cukup mendekati syaratnya, seperti melampaui ambang



yang telah diberikan, maka objeknya dapat digolongkan sama seperti objek pada acuan.

Transformasi Hough menyediakan metode yang serbaguna untuk mendeteksi bentuk yang dapat digambarkan sebagai suatu persamaan parametrik tertutup atau didalam bentuk tabular. Contoh dari bentuk yang dapat dijadikan sebagai parameter adalah garis, lingkaran, dan elips. Bentuk yang gagal dimasukkan ke dalam persamaan parametrik dapat dideteksi dengan menggunakan versi umum dari transformasi Hough yang menggunakan teknik dengan menggunakan tabel (Ritter dan Wilson, 1996).

Salah satu penerapan pengenalan objek adalah pada pengenalan wajah, yang menggunakan model *PCA(Principal Component Analysis)* dan pendekatan *Bayesian* untuk melakukan pengenalan dan klasifikasi.

1) Face Recognition

Dari semua tugas pengenalan objek komputer lakukan, pengenalan wajah adalah salah satu hal dimana tugas ini umumnya mempunyai kesuksesan tinggi untuk dikerjakan. Ketika komputer tidak dapat mengenali wajah seseorang dari ribuan wajah yang tertangkap video kamera (walaupun manusia juga belum bisa mengetahui wajah seseorang jika mereka tidak familiar dengannya), kemampuan komputer untuk membedakan seseorang dari sedikit objek, telah banyak digunakan pada aplikasi foto seperti Picasa dan iPhoto. Pengenalan wajah juga banyak digunakan pada beragam aplikasi, termasuk didalamnya Interaksi Manusia dan komputer (*HCI*), pengenalan identitas, login desktop, kontrol parental, dan juga monitoring pasien pada industri medis (Szeliski, 2009).

Suatu sistem pengenalan wajah mengidentifikasi wajah di dalam suatu citra dan video secara otomatis menggunakan komputer. Sistem ini mencakup beragam aplikasi, seperti aplikasi otentikasi biometrik, dan sistem kamera pengawas, interaksi manusia dan komputer, dan manajemen multimedia. Sistem pengenalan wajah umumnya terdiri atas empat bagian, yaitu : pendeteksian wajah, pensejajaran wajah, ekstraksi fitur pada wajah, dan klasifikasi wajah. Pendeteksian wajah menyediakan informasi tentang lokasi dan ukuran pada setiap wajah yang terdeteksi. Pada kasus video, wajah yang ditemukan memungkinkan untuk dapat

dilacak. Pada pensejajaran wajah, komponen pada wajah, seperti mata, hidung, dan mulut serta garis tepi pada wajah telah diketahui lokasinya, dan dari situ masukan citra wajah dinormalisasi secara geometris dan fotometris. Pada ekstraksi fitur, fitur yang dikira cukup berguna untuk membedakan orang satu dengan yang lainnya dipisahkan dari citra wajah yang telah mengalami normalisasi. Pada klasifikasi wajah, fitur vektor yang telah diekstrak dari masukan citra wajah dicocokkan dengan wajah yang telah didaftarkan sebelumnya ke dalam basis data, yang kemudian memberikan keluaran identitas dari wajah ketika data yang ada cocok dengan database (Li dan Lu, 2003).

Sistem pengenalan wajah saat ini baru bekerja secara maksimal ketika masukannya adalah wajah yang sepenuhnya menghadap ke depan dalam pencahayaan yang relatif sama, walaupun database yang ada sudah mengandung data pose dan pencahayaan yang sangat beragam juga (Szeliski, 2009).

Permasalahan ini dapat diselesaikan dengan menggunakan pengenalan pola dan pembelajaran mesin. Disini ada dua inti permasalahan: pertama, yaitu fitur-fitur apa sajakah yang digunakan untuk merepresentasikan suatu wajah, dan kedua yaitu bagaimana mengklasifikasikan suatu citra wajah baru berdasar pada representasi yang telah dipilih. Suatu sistem pengenalan wajah yang cukup memadai seharusnya dapat mengatasi permasalahan yang sangat bervariasi dari citra wajah pada sudut pandang, pencahayaan, ekspresi wajah, dan lain-lain.

Pendekatan dengan menggunakan fitur geometris adalah cara yang berdasar pada framework computer vision yang tradisional. Fitur wajah seperti mata, hidung, mulut, dan dagu kemudian dideteksi. Properti dan relasi seperti area, jarak, dan sudut diantara fitur wajah digunakan sebagai deskripsi wajah untuk keperluan pengenalan. Dengan menggunakan pendekatan ini, Kanade membangun sistem pengenalan wajah yang pertama di dunia. Keunggulan dari sistem ini adalah pada sisi ekonomisnya dan efisiensi pada penerimaan data reduksi dan minimnya sensitifitas pada pencahayaan dan sudut pandang. Kelemahan sistem ini sejauh ini, pendeteksian fitur wajah dan teknik pengukuran yang dikembangkan pada saat itu belum dapat diandalkan, dan informasi geometrik yang ada kurang mencukupi untuk pengenalan wajah (Li dan Lu, 2003).

Beberapa pendekatan yang cukup menarik pada teknik pengenalan wajah melibatkan penemuan lokasi dari beberapa fitur citra seperti mata, hidung, dan mulut, dan kemudian mengukur jarak diantara lokasi fitur-fitur ini. Pendekatan-pendekatan lain mengandalkan perbandingan tingkatan abu-abu dari citra yang diproyeksikan kepada ruang dimensional yang lebih rendah disebut *eigenfaces*, dan secara terpadu memodelkan bentuk dan kemunculan variasi menggunakan '*active appearance models*'.

2) Object Detection

Selama bertahun-tahun, telah dikembangkan beragam algoritma pendeteksian wajah yang mempunyai performa cukup cepat. Menurut referensi dari Yang et al., teknik deteksi wajah dapat diklasifikasikan sebagai teknik berdasar-ciri, berdasar-*template* (acuan), atau berdasar-rupa/tampilan. Teknik berdasar-ciri mencoba untuk menemukan lokasi dari ciri citra yang dapat dibedakan seperti mata, hidung, dan mulut, dan kemudian memverifikasi apakah ciri-ciri ini tersusun dengan benar secara geometris. Termasuk di dalam teknik-teknik ini adalah pendekatan awal pada teknik pengenalan wajah, sama seperti pendekatan lainnya yang berdasarkan *eigenspace modular*, *local filter jets*, dan *support vector machines*, dan *boosting*. Pendekatan pada teknik berdasar-acuan, seperti *Active Appearance Models*, dapat menangani beragam pose dan ekspresi yang bervariasi, tetapi biasanya memerlukan inisialisasi yang baik di dekat wajah sebenarnya, maka dari itu teknik ini tidak cocok untuk diterapkan pada aplikasi yang membutuhkan pendeteksian wajah secara cepat.

Teknik berdasar-rupa/tampilan menggunakan pendekatan melalui pemindaian potongan-potongan kecil berbentuk kotak yang saling menutupi dari citra yang akan diperiksa, dimana hasilnya kemudian akan diperbaiki menggunakan teknik *cascade* dari algoritma-algoritma deteksi yang lebih mahal, tetapi lebih selektif dan akurat. Kebanyakan dari pendekatan berbasis-rupa/tampilan sekarang lebih mengandalkan pada klasifikasi pelatihan menggunakan suatu set dari citra wajah terlabel dan label yang bukan wajah.

Sung and Poggio dan Rowley et al., menunjukkan dua pendekatan paling awal dari sistem deteksi wajah berdasar-tampilan/rupa, dan memperkenalkan beberapa inovasi yang kini banyak dipakai pada pekerjaan-pekerjaan selanjutnya

oleh peneliti lain. Untuk memulai, kedua sistem mengumpulkan suatu set dari potongan wajah yang telah terlabeli dan juga set potongan citra yang berisi objek bukan wajah. Dari citra wajah yang dikumpulkan ini kemudian dilakukan operasi tambahan seperti pembalikan/pencerminan, rotasi, scaling, dan translasi citra dengan sedikit perubahan untuk membuat sistem deteksi wajah yang ada tidak terlalu sensitif dengan efek-efek tersebut.

Setelah beberapa kali citra untuk pembelajaran didapatkan, dapat dilakukan beberapa pra-pemrosesan, seperti membuang rerata gradien (dari fungsi linear) dari citra untuk dapat mengimbangi efek-efek yang berbeda-beda, dan menggunakan persamaan histogram untuk mengimbangi kontras kamera yang bervariasi.

Dari bidang *object-detection* ini, menghasilkan suatu studi khusus pada pendeteksian wajah manusia atau *face detection*. Disini *face detection* berbeda dengan *face recognition*, karena hanya mendeteksi dimana letak wajah manusia pada citra masukan. Selama bertahun-tahun, telah dikembangkan banyak metode dan algoritma pendeteksian wajah manusia. Menurut Yang et al., teknik pendeteksian wajah dapat diklasifikasikan sebagai *feature-based*, *template-based*, atau *appearance-based*.

- Teknik *feature-based* mencoba mencari lokasi dari fitur citra yang sudah jelas seperti mata, hidung, dan mulut; kemudian melakukan verifikasi apakah fitur-fitur wajah ini sudah benar secara tatanan geometris. Teknik ini merupakan teknik yang muncul pada awal studi *face-detection*, juga pendekatan yang berbasis *eigenspaces*, *local filter jets*, *support vector machines*, dan *boosting*.
- Teknik *template-based* menggunakan pendekatan seperti *active appearance models (AAMs)*, dapat menangani bermacam-macam pose dan variasi ekspresi wajah, tetapi biasanya memerlukan inisialisasi yang tepat pada wajah, oleh karena itu tidak cocok untuk sistem yang membutuhkan pendeteksian yang cepat.
- Sedangkan teknik *appearance-based* memindai potongan-potongan persegi panjang diatas citra masukan untuk mencari kandidat region citra yang merupakan wajah, dan kemudian hasil pencariannya bisa disaring menggunakan suatu blok yang disebut *cascade* keputusan dan berisi algoritma deteksi (Viola

and Jones, 2004). Sebagian besar pendekatan *appearance-based* saat ini bergantung pada pengklasifikasi yang sudah dilatih menggunakan objek wajah dan bukan-wajah.

