**Tugas Kelompok 2**

**Muhammad Haris :1810131210009**

**Rapiyah Hawannur :1810131120017**

**FACE RECOGNITION**

**ABSTRACT**

Face recognition presents a challenging problem in the field of image analysis and  
computer vision, and as such has received a great deal of attention over the last few years because of its many applications in various domains. Face recognition techniques can be broadly divided into three categories based on the face data acquisition methodology: methods that operate on intensity images; those that deal with video sequences; and those that require other sensory data such as 3D information or infra-red imagery. In this paper, an overview of some of the well-known methods in each of these categories is provided and some of the benefits and drawbacks of the schemes mentioned therein are examined. Furthermore, a discussion outlining the incentive for using face recognition, the applications of this technology, and some of the difficulties plaguing current systems with regard to this task has also been provided. This paper also mentions some of the most recent algorithms developed for this purpose and attempts to give an idea of the state of the art of face recognition technology.

Pengenalan wajah menghadirkan masalah yang menantang di bidang analisis gambar dan visi komputer, dan dengan demikian telah menerima banyak perhatian selama beberapa tahun terakhir karena: banyak aplikasinya di berbagai domain. Teknik pengenalan wajah secara luas dapat dibagi menjadi: tiga kategori berdasarkan metodologi akuisisi data wajah: metode yang beroperasi pada intensitas gambar-gambar; mereka yang berhubungan dengan urutan video; dan yang membutuhkan data sensorik lainnya seperti 3D informasi atau citra infra merah. Dalam makalah ini, ikhtisar dari beberapa metode terkenal di masing-masing kategori ini disediakan dan beberapa manfaat dan kerugian dari skema yang disebutkan didalamnya diperiksa. Selanjutnya, diskusi yang menguraikan insentif untuk menggunakan pengenalan wajah, aplikasi teknologi ini, dan beberapa kesulitan yang mengganggu sistem saat ini sehubungan dengan: tugas ini juga telah disediakan. Makalah ini juga menyebutkan beberapa algoritma terbaru dikembangkan untuk tujuan ini dan mencoba untuk memberikan gambaran tentang keadaan seni pengenalan wajah teknologiPengenalan wajah menghadirkan masalah yang menantang di bidang analisis gambar dan visi komputer, dan dengan demikian telah menerima banyak perhatian selama beberapa tahun terakhir karena: banyak aplikasinya di berbagai domain. Teknik pengenalan wajah secara luas dapat dibagi menjadi: tiga kategori berdasarkan metodologi akuisisi data wajah: metode yang beroperasi pada intensitas gambar-gambar; mereka yang berhubungan dengan urutan video; dan yang membutuhkan data sensorik lainnya seperti 3D informasi atau citra infra merah. Dalam makalah ini, ikhtisar dari beberapa metode terkenal di masing-masing kategori ini disediakan dan beberapa manfaat dan kerugian dari skema yang disebutkan didalamnya diperiksa. Selanjutnya, diskusi yang menguraikan insentif untuk menggunakan pengenalan wajah, aplikasi teknologi ini, dan beberapa kesulitan yang mengganggu sistem saat ini sehubungan dengan: tugas ini juga telah disediakan. Makalah ini juga menyebutkan beberapa algoritma terbaru dikembangkan untuk tujuan ini dan mencoba untuk memberikan gambaran tentang keadaan seni pengenalan wajah teknologi.

**INTRODUCTION**

Face recognition is a task that humans perform routinely and effortlessly in their daily lives.  
Wide availability of powerful and low-cost desktop and embedded computing systems has created an enormous interest in automatic processing of digital images and videos in a number of applications, including biometric authentication, surveillance, human-computer interaction, and multimedia management. Research and development in automatic face recognition follows naturally.

As one of the most successful applications of image analysis and understanding, face  
recognition has recently received significant attention, especially during the past  
few years. This is evidenced by the emergence of face recognition conferences such as the International Conference on Audioand Video-Based Authentication (AVBPA)  
since 1997 and the International Conference on Automatic Face and Gesture Recognition (AFGR) since 1995, systematic empirical evaluations of face recognition techniques (FRT), including the FERET [Phillips et al. 1998b, 2000; Rizvi et al. 1998], FRVT 2000 [Blackburn et al. 2001], FRVT 2002 [Phillips et al. 2003], and XM2VTS [Messer et al. 1999] protocols, and many commercially available systems (Table II). There are at least two reasons for this trend; the first is the wide range of commercial and law enforcement applications and the second is the availability of feasible technologies after 30 years of research. In addition, the problem of machine recognition of human faces continues to attract researchers from disciplines such as image processing, pattern recognition, neural networks, computer vision, computer graphics, and psychology.

The strong need for user-friendly systems that can secure our assets and protect our privacy without losing our identity in a sea of numbers is obvious. At present, one needs a PIN to get cash from  
an ATM, a password for a computer, a dozen others to access the internet, and so on. Although very reliable methods of biometric personal identification exist, for example, fingerprint analysis and retinal or iris scans, these methods rely on the cooperation of the participants, whereas  
a personal identification system based on analysis of frontal or profile images of the  
face is often effective without the participant’s cooperation or knowledge. Some of the advantages/disadvantages of different biometrics are described in Phillips et al.  
[1998]. Table I lists some of the applications of face recognition.

**Table I.** Typical Applications of Face Recognition

|  |  |
| --- | --- |
| Areas | Specific applications |
| Entertainment | Video game, virtual reality, training programs |
|  | Human-robot-interaction, human-computer-interaction |
|  | Drivers’ licenses, entitlement programs |
| Smart cards | Immigration, national ID, passports, voter registration |
|  | Welfare fraud |
| Information security | TV Parental control, personal device logon, desktop logon |
|  | Application security, database security, file encryption |
|  | Intranet security, internet access, medical records |
|  | Secure trading terminals |
| Law enforcement  and surveillance | Advanced video surveillance, CCTV control |
|  | Portal control, postevent analysis |
|  | Shoplifting, suspect tracking and investigation |

Research in face recognition is motivated not only by the fundamental challenges this recognition problem poses but also by numerous practical applications where human identification  
is needed. Face recognition, as one of the primary biometric technologies, became more and  
more important owing to rapid advances in technologies such as digital cameras, the Internet  
and mobile devices, and increased demands on security. Face recognition has several advantages over other biometric technologies: It is natural, nonintrusive, and easy to use. Among the  
six biometric attributes considered by Hietmeyer [12], facial features scored the highest compatibility in a Machine Readable Travel Documents (MRTD) [18] system based on a number of  
evaluation factors, such as enrollment, renewal, machine requirements, and public perception,  
shown in Figure 1.1.

A face recognition system is expected to identify faces present in images and videos automatically. It can operate in either or both of two modes: (1) face verification (or authentication),  
and (2) face identification (or recognition). Face verification involves a one-to-one match that  
compares a query face image against a template face image whose identity is being claimed.  
Face identification involves one-to-many matches that compares a query face image against all  
the template images in the database to determine the identity of the query face. Another face  
recognition scenario involves a watch-list check, where a query face is matched to a list of suspects (one-to-few matches). The performance of face recognition systems has improved significantly since the first automatic face recognition system was developed by Kanade [14]. Furthermore, face detection, facial feature extraction, and recognition can now be performed in “realtime” for images captured under favorable (i.e., constrained) situations. Although progress in face recognition has been encouraging, the task has also turned out to be a difficult endeavor, especially for unconstrained tasks where viewpoint, illumination, expression, occlusion, accessories, and so on vary considerably. In the following sections, we give a brief review on technical advances and analyze technical challenges.

Therefore, with this research in the form of an experiment, software engineering whose external form is an application with data with data from research in the form of image samples captured from a connected webcam or photo with a computer. Human face images are taken differently with each getting the same variation treatment, namely the slope of the face image position, face distance, and light intensity. Based on research journals conducted by (Muliawan et al., 2015) shows that facial recognition processing in the attendance system with using the Eigenface method on OpenCV can be done by inputting data user and facial data along with the password of each user. "Method Face Recognition for Personnel Identification Based on Facial Image” From the journal research conducted by (Kurniawan, 2014) discusses the design of online face recognition system for Semarang State University employees with Python and Opencv can be realized.

**PENDAHULUAN**

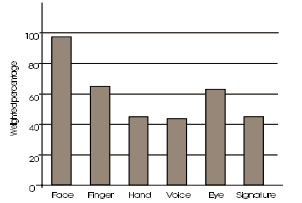
Pengenalan wajah adalah tugas yang dilakukan manusia secara rutin dan mudah dalam kehidupan sehari-hari. Ketersediaan luas desktop yang kuat dan berbiaya rendah dan sistem komputasi tertanam telah menciptakan minat yang sangat besar dalam pemrosesan otomatis gambar dan video digital di sejumlah aplikasi, termasuk otentikasi biometrik, pengawasan, interaksi manusia-komputer, dan manajemen multimedia. Penelitian dan pengembangan dalam pengenalan wajah otomatis mengikuti tentu saja.

Sebagai salah satu aplikasi paling sukses analisis dan pemahaman gambar, wajah pengakuan baru-baru ini mendapat perhatian yang signifikan, terutama selama masa lalu beberapa tahun. Hal ini dibuktikan dengan munculnya konferensi pengenalan wajah seperti: sebagai Konferensi Internasional tentang Otentikasi Berbasis Audio dan Video (AVBPA) sejak 1997 dan Konferensi Internasional tentang Wajah dan Gestur Otomatis Pengakuan (AFGR) sejak tahun 1995, evaluasi empiris sistematis teknik pengenalan wajah (FRT), termasuk: FERET [Phillips dkk. 1998b, 2000; Rizvi dkk. 1998], FRVT 2000 [Blackburn dkk. 2001], FRVT 2002 [Phillips dkk. 2003], dan XM2VTS [Messer dkk. 1999] protokol, dan banyak tersedia secara komersial sistem (Tabel II). Setidaknya ada dua alasan tren ini; yang pertama lebar jangkauan komersial dan penegakan hokum aplikasi dan yang kedua adalah ketersediaan teknologi yang layak setelah 30 tahun penelitian. Selain itu, masalah pengenalan wajah manusia oleh mesin terus menarik peneliti dari disiplin ilmu seperti pengolahan citra, pola, pengenalan, jaringan saraf, komputer visi, grafik komputer, dan psikologi.

Kebutuhan yang kuat akan sistem yang mudah digunakan yang dapat mengamankan aset kami dan melindungi privasi kami tanpa kehilangan identitas kami dalam lautan angka sudah jelas. Pada sekarang, seseorang membutuhkan PIN untuk mendapatkan uang tunai dari ATM, kata sandi untuk komputer, a selusin orang lain untuk mengakses internet, dan segera. Meskipun metode yang sangat andal untuk identifikasi pribadi biometrik ada, untuk contoh, analisis sidik jari dan retinal atau pemindaian iris, metode ini bergantung pada kerjasama para peserta, sedangkan sistem identifikasi pribadi berdasarkan analisis gambar frontal atau profil dari wajah seringkali efektif tanpa kerja sama atau sepengetahuan peserta. Beberapa kelebihan/kekurangan yang berbeda biometrik dijelaskan dalam Phillips et al. [1998]. Tabel I mencantumkan beberapa aplikasi pengenalan wajah.

|  |  |
| --- | --- |
| Areas | Specific applications |
| Entertainment | Video game, virtual reality, training programs |
|  | Human-robot-interaction, human-computer-interaction |
|  | Drivers’ licenses, entitlement programs |
| Smart cards | Immigration, national ID, passports, voter registration |
|  | Welfare fraud |
| Information security | TV Parental control, personal device logon, desktop logon |
|  | Application security, database security, file encryption |
|  | Intranet security, internet access, medical records |
|  | Secure trading terminals |
| Law enforcement  and surveillance | Advanced video surveillance, CCTV control |
|  | Portal control, postevent analysis |
|  | Shoplifting, suspect tracking and investigation |

Penelitian dalam pengenalan wajah dimotivasi tidak hanya oleh tantangan mendasar yang ditimbulkan oleh masalah pengenalan ini, tetapi juga oleh berbagai aplikasi praktis di mana identifikasi manusia dibutuhkan. Pengenalan wajah, sebagai salah satu teknologi biometrik utama, menjadi lebih dan lebih penting karena kemajuan pesat dalam teknologi seperti kamera digital, Internet dan perangkat seluler, dan peningkatan tuntutan keamanan. Pengenalan wajah memiliki beberapa keunggulan dibandingkan teknologi biometrik lainnya: Alami, tidak mengganggu, dan mudah digunakan. Diantara enam atribut biometrik yang dipertimbangkan oleh Hietmeyer [12], fitur wajah mencetak kompatibilitas tertinggi dalam sistem Machine Readable Travel Documents (MRTD) [18] berdasarkan sejumlah faktor evaluasi, seperti pendaftaran, pembaruan, persyaratan mesin, dan persepsi publik, ditunjukkan pada Gambar 1.1.

Gambar 1.1. Skenario penggunaan sistem MRTD biometrik untuk kontrol paspor (kiri), dan perbandingan berbagai fitur biometrik berdasarkan kompatibilitas MRTD (kanan, dari Hietmeyer [12] dengan izin).

Sistem pengenalan wajah diharapkan dapat mengidentifikasi wajah yang ada pada gambar dan video secara otomatis. Itu dapat beroperasi di salah satu atau kedua dari dua mode: (1) verifikasi wajah (atau otentikasi), dan (2) identifikasi wajah (atau pengenalan). Verifikasi wajah melibatkan pertandingan satu lawan satu yang membandingkan gambar wajah kueri dengan gambar wajah template yang identitasnya diklaim. Identifikasi wajah melibatkan kecocokan satu-ke-banyak yang membandingkan gambar wajah kueri dengan semua gambar template dalam database untuk menentukan identitas wajah kueri. Wajah lain skenario pengenalan melibatkan pemeriksaan daftar pantauan, di mana wajah kueri dicocokkan dengan daftar tersangka (satu-ke-beberapa pertandingan). Kinerja sistem pengenalan wajah telah meningkat secara signifikan sejak sistem pengenalan wajah otomatis pertama dikembangkan oleh Kanade [14]. Selanjutnya, deteksi wajah, ekstraksi fitur wajah, dan pengenalan sekarang dapat dilakukan dalam "waktu nyata" untuk gambar yang diambil dalam situasi yang menguntungkan (yaitu, dibatasi). Meskipun kemajuan dalam pengenalan wajah telah menggembirakan, tugas juga ternyata menjadi usaha yang sulit, terutama untuk tugas-tugas yang tidak dibatasi di mana sudut pandang, penerangan, ekspresi, oklusi, aksesori, dan sebagainya sangat bervariasi. Di bagian berikut, kami memberikan tinjauan singkat tentang kemajuan teknis dan menganalisis tantangan teknis.

Oleh karena itu, dengan adanya penelitian ini yang berbentuk eksperimen  
rekayasa perangkat lunak yang luarnya berupa aplikasi dengan data dengan data dari  
penelitian berupa sampel citra yang di capture dari webcam atau foto yang terhubung  
dengan komputer. Citra wajah manusia yang diambil berbeda-beda dengan masingmasing mendapatkan perlakuan variasi yang sama yaitu kemiringan posisi citra wajah,  
jarak wajah, dan intesitas cahaya. Berdasarkan jurnal penelitian yang dilakukan oleh (Muliawan et al., 2015) menunjukan bahwa pemrosesan pengenalan wajah pada system absensi dengan menggunakan metode Eigenface pada OpenCV dapat dilakukan dengan cara input data pengguna dan data wajah beserta password dari masing-masing pengguna. “Metod Face Recognition untuk Identifikasi Personil Berdasar Citra Wajah” Dari jurnal  
penelitian yang dilakukan oleh (Kurniawan, 2014) membahas tentang rancang bangun  
system face recognition online pegawai Universitas Negeri Semarang dengan Python  
dan Opencv dapat terealisasi. Oleh karena itu, dengan adanya penelitian ini yang berbentuk eksperimen rekayasa perangkat lunak yang luarnya berupa aplikasi dengan data dengan data dari penelitian berupa sampel citra yang di capture dari webcam atau foto yang terhubung  
dengan komputer. Citra wajah manusia yang diambil berbeda-beda dengan masingmasing mendapatkan perlakuan variasi yang sama yaitu kemiringan posisi citra wajah,  
jarak wajah, dan intesitas cahaya.

Berdasarkan jurnal penelitian yang dilakukan oleh (Muliawan et al., 2015) menunjukan bahwa pemrosesan pengenalan wajah pada system absensi dengan menggunakan metode Eigenface pada OpenCV dapat dilakukan dengan cara input data pengguna dan data wajah beserta password dari masing-masing pengguna. “Metode Face Recognition untuk Identifikasi Personil Berdasar Citra Wajah” Dari jurnal penelitian yang dilakukan oleh (Kurniawan, 2014) membahas tentang rancang bangun system face recognition online pegawai Universitas Negeri Semarang dengan Python dan Opencv dapat terealisasi.

**Face Recognition Processing**

Face recognition is a visual pattern recognition problem. There, a face as a three-dimensional  
object subject to varying illumination, pose, expression and so on is to be identified based on its two-dimensional image (three-dimensional images e.g., obtained from laser may also be used).  
A face recognition system generally consists of four modules as depicted in Figure 1.2: detection, alignment, feature extraction, and matching, where localization and normalization (face detection and alignment) are processing steps before face recognition (facial feature extraction and matching) is performed.

Face detection segments the face areas from the background. In the case of video, the detected faces may need to be tracked using a face tracking component. Face alignment is aimed at achieving more accurate localization and at normalizing faces thereby whereas face detection  
provides coarse estimates of the location and scale of each detected face. Facial components,  
such as eyes, nose, and mouth and facial outline, are located; based on the location points, the  
input face image is normalized with respect to geometrical properties, such as size and pose,  
using geometrical transforms or morphing. The face is usually further normalized with respect to photometrical properties such illumination and gray scale.

After a face is normalized geometrically and photometrically, feature extraction is performed to provide effective information that is useful for distinguishing between faces of different persons and stable with respect to the geometrical and photometrical variations. For face matching, the extracted feature vector of the input face is matched against those of enrolled.

faces in the database; it outputs the identity of the face when a match is found with sufficient  
confidence or indicates an unknown face otherwise. Face recognition results depend highly on features that are extracted to represent the face pattern and classification methods used to distinguish between faces whereas face localization and normalization are the basis for extracting effective features. These problems may be analyzed from the viewpoint of face subspaces or manifolds, as follows.

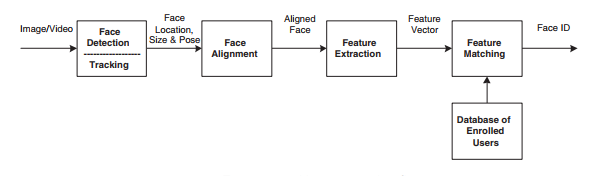
**Pemrosesan Pengenalan Wajah**

Pengenalan wajah adalah masalah pengenalan pola visual. Di sana, wajah sebagai tiga dimensi objek yang tunduk pada berbagai iluminasi, pose, ekspresi, dan sebagainya harus diidentifikasi berdasarkan objeknya gambar dua dimensi (gambar tiga dimensi misalnya, diperoleh dari laser juga dapat digunakan). Sebuah sistem pengenalan wajah umumnya terdiri dari empat modul seperti yang digambarkan pada Gambar 1.2: deteksi, penyelarasan, ekstraksi fitur, dan pencocokan, di mana lokalisasi dan normalisasi (wajah deteksi dan penyelarasan) adalah langkah-langkah pemrosesan sebelum pengenalan wajah (ekstraksi fitur wajah dan pencocokan) dilakukan.

Deteksi wajah membagi area wajah dari latar belakang. Dalam kasus video, wajah yang terdeteksi mungkin perlu dilacak menggunakan komponen pelacakan wajah. Penjajaran wajah ditujukan untuk mencapai lokalisasi yang lebih akurat dan menormalkan wajah dengan demikian sedangkan deteksi wajah memberikan perkiraan kasar lokasi dan skala setiap wajah yang terdeteksi. komponen wajah, seperti mata, hidung, dan mulut dan garis wajah, berada; berdasarkan titik lokasi, citra wajah masukan dinormalisasi sehubungan dengan sifat geometris, seperti ukuran dan pose, menggunakan transformasi geometris atau morphing. Wajah biasanya lebih dinormalisasi dengan hormat untuk sifat fotometrik seperti iluminasi dan skala abu-abu.

Setelah wajah dinormalisasi secara geometris dan fotometrik, ekstraksi fitur dilakukan untuk memberikan informasi yang efektif yang berguna untuk membedakan antara wajah orang yang berbeda dan stabil terhadap variasi geometris dan fotometrik. Untuk wajah pencocokan, vektor fitur yang diekstraksi dari wajah input dicocokkan dengan yang terdaftar.

wajah dalam database; itu menampilkan identitas wajah ketika kecocokan ditemukan dengan cukup percaya diri atau menunjukkan wajah yang tidak dikenal sebaliknya. Hasil pengenalan wajah sangat bergantung pada fitur yang diekstraksi untuk mewakili wajah pola dan metode klasifikasi yang digunakan untuk membedakan antara wajah sedangkan lokalisasi wajah dan normalisasi adalah dasar untuk mengekstraksi fitur yang efektif. Masalah-masalah ini dapat dianalisis dari sudut pandang subruang atau manifold wajah, sebagai berikut.



Gambar 1.2. Alur pemrosesan pengenalan wajah.

.

**Applications**

Face recognition is used for two primary tasks:

1. Verification (one-to-one matching): When presented with a face image of an unknown individual along with a claim of identity, ascertaining whether the individual is who he/she claims to be.

2. Identification (one-to-many matching): Given an image of an unknown individual, determining that person’s identity by comparing (possibly after encoding) that image with a database of (possibly encoded) images of known individuals.

There are numerous application areas in which face recognition can be exploited for these two purposes, a few of which are outlined below.

• Security (access control to buildings, airports/seaports, ATM machines and border checkpoints [2, 3]; computer/ network security [4]; email authentication on multimedia workstations).

• Surveillance (a large number of CCTVs can be monitored to look for known criminals, drug offenders, etc. and authorities can be notified when one is located; for example, this procedure was used at the Super Bowl 2001 game at Tampa, Florida [5]; in another instance, according to a CNN report, two cameras linked to state and national databases of sex offenders, missing children and alleged abductors have been installed recently at Royal Palm Middle School in Phoenix, Arizona [6]).

• General identity verification (electoral registration, banking, electronic commerce, identifying newborns, national IDs, passports, drivers’ licenses, employee IDs).

• Criminal justice systems (mug-shot/booking systems, post-event analysis, forensics).

• Image database investigations (searching image databases of licensed drivers, benefit recipients, missing children, immigrants and police bookings).

• “Smart Card” applications (in lieu of maintaining a database of facial images, the face-print can be stored in a smart card, bar code or magnetic stripe, authentication of which is performed by matching the live image and the stored template) [7].

• Multi-media environments with adaptive humancomputer interfaces (part of ubiquitous or contextaware systems, behavior monitoring at childcare or old people’s centers, recognizing a customer and assessing his needs) [8, 9].

• Video indexing (labeling faces in video) [10, 11].

• Witness face reconstruction [12]. In addition to these applications, the underlying techniques in the current face recognition technology have also been modified and used for related applications such as gender classification [13-15], expression recognition [16, 17] and facial feature recognition and tracking [18]; each of these has its utility in various domains: for instance, expression recognition can be utilized in the field of medicine for intensive care monitoring [19] while facial feature recognition and detection can be exploited for tracking a vehicle driver’s eyes and thus monitoring his fatigue [20], as well as for stress detection [21]. Face recognition is also being used in conjunction with other biometrics such as speech, iris, fingerprint, ear and gait recognition in order to enhance the recognition performance of these methods [8, 22-34].

**Aplikasi**

Pengenalan wajah digunakan untuk dua tugas utama:

1. Verifikasi (pencocokan satu-ke-satu): Saat disajikan dengan gambar wajah orang yang tidak dikenal bersama dengan klaim identitas, memastikan apakah individu adalah siapa yang dia klaim.

2. Identifikasi (pencocokan satu-ke-banyak): Diberikan sebuah gambar dari individu yang tidak dikenal, menentukan identitas dengan membandingkan (mungkin setelah penyandian) yang gambar dengan database gambar (mungkin dikodekan) dari individu yang dikenal.

Ada banyak area aplikasi di mana wajah pengakuan dapat dimanfaatkan untuk dua tujuan ini, beberapa di antaranya diuraikan di bawah ini.

• Keamanan (kontrol akses ke gedung, bandara/pelabuhan, Mesin ATM dan pos pemeriksaan perbatasan [2, 3]; komputer/keamanan jaringan [4]; otentikasi email pada multimedia stasiun kerja).

• Pengawasan (sejumlah besar CCTV dapat dipantau untuk mencari penjahat yang diketahui, pelanggar narkoba, dll. dan pihak berwenang dapat diberitahu ketika seseorang berada; misalnya, prosedur ini digunakan di Super Pertandingan Bowl 2001 di Tampa, Florida [5]; di tempat lain Misalnya, menurut laporan CNN, dua kamera terkait dengan database negara bagian dan nasional pelaku kejahatan seksual, anak-anak yang hilang dan tersangka penculik telah dipasang baru-baru ini di Royal Palm Middle School di Phoenix, Arizona [6]).

• Verifikasi identitas umum (pendaftaran pemilu, perbankan, perdagangan elektronik, mengidentifikasi bayi baru lahir, ID nasional, paspor, SIM, ID karyawan).

• Sistem peradilan pidana (sistem mug-shot/booking, analisis pasca-peristiwa, forensik).

• Investigasi database gambar (mencari database gambar pengemudi berlisensi, penerima manfaat, anak hilang, imigran dan pemesanan polisi).

• Aplikasi “Kartu Pintar” (sebagai pengganti pemeliharaan database gambar wajah, cetakan wajah dapat disimpan dalam kartu pintar, kode batang, atau strip magnetik, yang otentikasinya dilakukan dengan mencocokkan siaran langsung gambar dan template yang disimpan) [7].

• Lingkungan multi-media dengan antarmuka komputer manusia yang adaptif (bagian dari sistem yang ada di mana-mana atau kontekstual, pemantauan perilaku di penitipan anak atau pusat orang tua, mengenali pelanggan dan menilai kebutuhannya) [8, 9].

• Pengindeksan video (melabeli wajah dalam video) [10, 11].

• Rekonstruksi wajah saksi [12]. Selain aplikasi ini, teknik yang mendasarinya dalam teknologi pengenalan wajah saat ini juga telah dimodifikasi dan digunakan untuk aplikasi terkait seperti gender klasifikasi [13-15], pengenalan ekspresi [16, 17] dan pengenalan dan pelacakan fitur wajah [18]; masing-masing memiliki kegunaannya di berbagai domain: misalnya, ekspresi pengakuan dapat dimanfaatkan dalam bidang kedokteran untuk pemantauan perawatan intensif [19] sementara fitur wajah pengenalan dan deteksi dapat dimanfaatkan untuk melacak mata pengemudi kendaraan dan dengan demikian memantau kelelahannya [20], serta untuk deteksi stres [21]. Pengenalan wajah juga digunakan bersama dengan biometrik lain seperti ucapan, iris, sidik jari, telinga dan pengenalan gaya berjalan untuk meningkatkan pengenalan kinerja metode ini [8, 22-34].

**Face Recognition Techniques**

The method for acquiring face images depends upon the underlying application. For instance, surveillance applications may best be served by capturing face images by means of a video camera while image database investigations may require static intensity images taken by a standard camera. Some other applications, such as access to top security domains, may even necessitate the forgoing of the nonintrusive quality of face recognition by requiring the user to stand in front of a 3D scanner or an infra-red sensor. Therefore, depending on the face data acquisition methodology, face recognition techniques can be broadly divided into three categories: methods that operate on intensity images, those that deal with video sequences, and those that require other sensory data such as 3D information or infra-red imagery. The following discussion sheds some light on the methods in each category and attempts to give an idea of some of the benefits and drawbacks of the schemes mentioned therein in general (for detailed surveys, please see [44, 45]).

**Face Recognition from Intensity Images**

Face recognition methods for intensity images fall into two main categories: feature-based and holistic [46-48]. An overview of some of the well-known methods in these categories is given below.

**Teknik Pengenalan Wajah**

Metode untuk gambar wajah tergantung pada aplikasi yang diperolehnya. Misalnya, penerapan pengawasan mungkin paling baik dilayani dengan menangkap gambar wajah melalui video investigasi database gambar mungkin memerlukan gambar intensitas tinggi yang diambil oleh kamera standar. Beberapa aplikasi lain, seperti akses ke domain keamanan teratas, mungkin juga menerapkan kualitas pengenalan wajah yang tidak mengganggu dengan mengharuskan pengguna berdiri di depan 3D atau sensor infra merah. Oleh karena itu, tergantung pada akuisisi data wajah, pengenalan wajah dapat dilakukan secara luas dibagi tiga kategori: metode yang beroperasi pada gambar intensitas, metode yang berhubungan dengan urutan video, dan metode yang memerlukan data sensorik lain seperti informasi 3D atau infra-merah . perumpamaan. Diskusi berikut menjelaskan beberapa metode dalam setiap kategori dan memberikan gambaran tentang beberapa keuntungan dan kerugian dari skema yang disebutkan di dalamnya secara umum (untuk survei rinci, silakan lihat [44, 45]).

**Pengenalan Wajah dari Gambar Intensitas**

Metode pengenalan wajah untuk gambar intensitas terbagi dalam dua kategori utama: berbasis fitur dan holistik [46-48]. Ikhtisar dari beberapa metode terkenal dalam kategori ini diberikan di bawah ini.

**1.Featured-based**

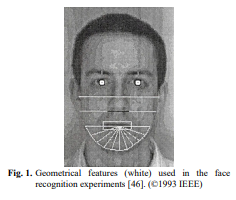
Feature-based approaches first process the input image to identify and extract (and measure) distinctive facial features such as the eyes, mouth, nose, etc., as well as other fiducial marks, and then compute the geometric relationships among those facial points, thus reducing the input facial image to a vector of geometric features. Standard statistical pattern recognition techniques are then employed to match faces using these measurements.

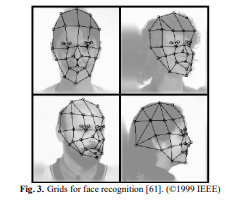
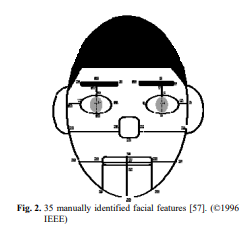
Early work carried out on automated face recognition was mostly based on these techniques. One of the earliest such attempts was by Kanade [49], who employed simple image processing methods to extract a vector of 16 facial parameters - which were ratios of distances, areas and angles (to compensate for the varying size of the pictures) - and used a simple Euclidean distance measure for matching to achieve a peak performance of 75% on a database of 20 different people using 2 images per person (one for reference and one for testing).

**Berbasis unggulan**

Pendekatan berbasis fitur pertama-tama memproses gambar input untuk mengidentifikasi dan mengekstrak (dan mengukur) fitur wajah yang khas seperti mata, mulut, hidung, dll., serta tanda fiducial lainnya, dan kemudian menghitung hubungan geometris di antara titik-titik wajah tersebut, sehingga mengurangi citra wajah masukan menjadi vektor fitur geometris. Teknik pengenalan pola statistik standar kemudian digunakan untuk mencocokkan wajah menggunakan pengukuran ini.

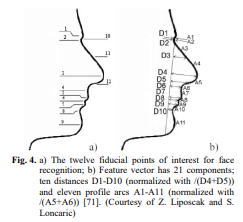
Pekerjaan awal yang dilakukan pada pengenalan wajah otomatis sebagian besar didasarkan pada teknik ini. Salah satu upaya paling awal adalah oleh Kanade [49], yang menggunakan metode pemrosesan gambar sederhana untuk mengekstraksi vektor dari 16 parameter wajah - yang merupakan rasio jarak, area, dan sudut (untuk mengimbangi ukuran gambar yang bervariasi) - dan menggunakan ukuran jarak Euclidean sederhana untuk pencocokan untuk mencapai kinerja puncak 75% pada database 20 orang yang berbeda menggunakan 2 gambar per orang (satu untuk referensi dan satu untuk pengujian).



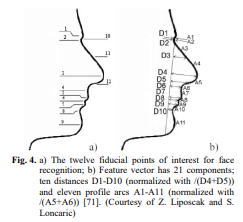


Another well-known feature-based approach is the elastic bunch graph matching method proposed by Wiskott et al. [59] . This technique is based on Dynamic Link Structures [60]. A graph for an individual face is generated as follows: a set of fiducial points on the face are chosen. Each fiducial point is a node of a full connected graph, and is labeled with the Gabor filters’ responses applied to a window around the fiducial point. Each arch is labeled with the distance between the correspondent fiducial points. A representative set of such graphs is combined into a stack-like structure, called a face bunch graph. Once the system has a face bunch graph, graphs for new face images can then be generated automatically by Elastic Bunch Graph Matching. Recognition of a new face image is performed by comparing its image graph to those of all the known face images and picking the one with the highest similarity value. Using this architecture, the recognition rate can reach 98% for the first rank and 99% for the first 10 ranks using a gallery of 250 individuals. The system has been enhanced to allow it to deal with different poses (Fig. 3) [61] but the recognition performance on faces of the same orientation remains the same. Though this method was among the best performing ones in the most recent FERET evaluation [62, 63], it does suffer from the serious drawback of requiring the graph placement for the first 70 faces to be done manually before the elastic graph matching becomes adequately dependable [64]. Campadelli and Lanzarotti [65] have recently experimented with this technique, where they have eliminated the need to do the graph placement manually by using parametric models, based on the deformable templates proposed in [50], to automatically locate fiducial points. They claim to have obtained the same performances as the elastic bunch graph employed in [59]. Other recent variations of this approach replace the Gabor features by a graph matching strategy [66] and HOGs (Histograms of Oriented Gradients [67].

Considerable effort has also been devoted to recognizing faces from their profiles [68-72] since, in this case, feature extraction becomes a somewhat simpler one-dimensional problem [57, 71]. Kaufman and Breeding [70] reported a recognition rate of 90% using face profiles; however, they used a database of only 10 individuals. Harmon et al. [68] obtained recognition accuracies of 96% on a database of 112 individuals, using a 17-dimensional feature vector to describe face profiles and utilizing a Euclidean distance measure for matching. More recently, Liposcak and Loncaric [71] reported a 90% accuracy rate on a database of 30 individuals, using subspace filtering to derive a 21- dimensional feature vector to describe the face profiles and employing a Euclidean distance measure to match them



Pendekatan berbasis fitur lain yang terkenal adalah metode pencocokan grafik tandan elastis yang diusulkan oleh Wiskott et al. [59] . Teknik ini didasarkan pada Struktur Tautan Dinamis [60]. Grafik untuk wajah individu dihasilkan sebagai berikut: satu set titik fiducial pada wajah dipilih. Setiap titik fidusia adalah simpul dari grafik yang terhubung penuh, dan diberi label dengan tanggapan filter Gabor yang diterapkan ke jendela di sekitar titik fidusia. Setiap lengkungan diberi label dengan jarak antara titik fidusia koresponden. Kumpulan perwakilan dari grafik tersebut digabungkan menjadi struktur seperti tumpukan, yang disebut grafik tandan wajah. Setelah sistem memiliki grafik tandan wajah, grafik untuk gambar wajah baru kemudian dapat dibuat secara otomatis oleh Pencocokan Grafik Tandan Elastis. Pengenalan citra wajah baru dilakukan dengan membandingkan grafik citranya dengan semua citra wajah yang diketahui dan memilih salah satu dengan nilai kemiripan tertinggi. Dengan menggunakan arsitektur ini, tingkat pengenalan dapat mencapai 98% untuk peringkat pertama dan 99% untuk 10 peringkat pertama menggunakan galeri 250 individu. Sistem telah ditingkatkan untuk memungkinkannya menangani pose yang berbeda (Gbr. 3) [61] tetapi kinerja pengenalan pada wajah dengan orientasi yang sama tetap sama. Meskipun metode ini adalah salah satu yang berkinerja terbaik dalam evaluasi FERET terbaru [62, 63], metode ini memiliki kelemahan serius yang mengharuskan penempatan grafik untuk 70 wajah pertama dilakukan secara manual sebelum pencocokan grafik elastis menjadi cukup dapat diandalkan. [64]. Campadelli dan Lanzarotti [65] baru-baru ini bereksperimen dengan teknik ini, di mana mereka telah menghilangkan kebutuhan untuk melakukan penempatan grafik secara manual dengan menggunakan model parametrik, berdasarkan template yang dapat dideformasi yang diusulkan pada [50], untuk secara otomatis menemukan titik fiducial. Mereka mengklaim telah memperoleh kinerja yang sama dengan grafik tandan elastis yang digunakan di [59]. Variasi terbaru lainnya dari pendekatan ini menggantikan fitur Gabor dengan strategi pencocokan grafik [66] dan HOGs (Histograms of Oriented Gradients [67].



Upaya yang cukup besar juga telah dikhususkan untuk mengenali wajah dari profil mereka [68-72] karena, dalam hal ini, ekstraksi fitur menjadi masalah satu dimensi yang agak sederhana [57, 71]. Kaufman dan Pemuliaan [70] melaporkan tingkat pengenalan 90% menggunakan profil wajah; namun, mereka menggunakan database hanya 10 individu. Harmon dkk. [68] memperoleh akurasi pengenalan 96% pada database 112 individu, menggunakan vektor fitur 17-dimensi untuk menggambarkan profil wajah dan memanfaatkan ukuran jarak Euclidean untuk pencocokan. Baru-baru ini, Liposcak dan Loncaric [71] melaporkan tingkat akurasi 90% pada database 30 individu, menggunakan penyaringan subruang untuk memperoleh vektor fitur 21-dimensi untuk menggambarkan profil wajah dan menggunakan ukuran jarak Euclidean untuk mencocokkannya.

**2.Holistic**

Holistic approaches attempt to identify faces using global representations, i.e., descriptions based on the entire image rather than on local features of the face. These schemes can be subdivided into two groups: statistical and AI approaches. An overview of some of the methods in these categories follows.

**2.Holistik**

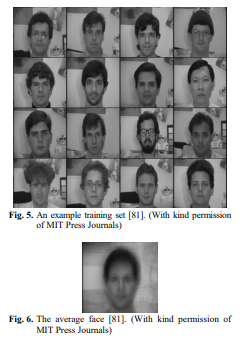
Pendekatan holistik berusaha mengidentifikasi wajah menggunakan representasi global, yaitu deskripsi berdasarkan keseluruhan gambar daripada fitur lokal wajah. Skema ini dapat dibagi menjadi dua kelompok: pendekatan statistik dan AI. Ikhtisar dari beberapa metode dalam kategori ini berikut.

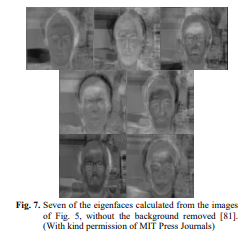
**Statistical**

In the simplest version of the holistic approaches, the image is represented as a 2D array of intensity values and recognition is performed by direct correlation comparisons between the input face and all the other faces in the database. Though this approach has been shown to work [75] under limited circumstances (i.e., equal illumination, scale, pose, etc.), it is computationally very expensive and suffers from the usual shortcomings of straightforward correlation-based approaches, such as sensitivity to face orientation, size, variable lighting conditions, background clutter, and noise [76]. The major hindrance to the directmatching methods’ recognition performance is that they attempt to perform classification in a space of very high dimensionality [76]. To counter this curse of dimensionality, several other schemes have been proposed that employ statistical dimensionality reduction methods to obtain and retain the most meaningful feature dimensions before performing recognition. A few of these are mentioned below.

Sirovich and Kirby [77] were the first to utilize Principal Components Analysis (PCA) [78, 79] to economically represent face images. They demonstrated that any particular face can be efficiently represented along the eigenpictures coordinate space, and that any face can be approximately reconstructed by using just a small collection of eigenpictures and the corresponding projections (‘coefficients’) along each eigenpicture.

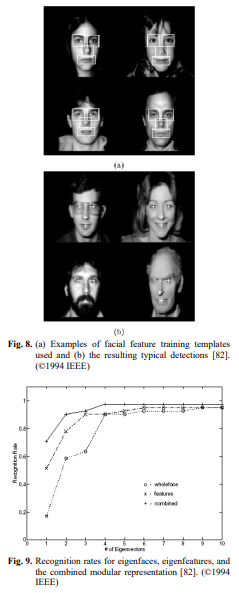
Turk and Pentland [80, 81] realized, based on Sirovich and Kirby’s findings, that projections along eigenpictures could be used as classification features to recognize faces. They employed this reasoning to develop a face recognition system that builds eigenfaces, which correspond to the eigenvectors associated with the dominant eigenvalues of the known face (patterns) covariance matrix, and then recognizes particular faces by comparing their projections along the eigenfaces to those of the face images of the known individuals. The eigenfaces define a feature space that drastically reduces the dimensionality of the original space, and face identification is carried out in this reduced space. An example training set, the average face and the top seven eigenfaces derived from the training images are shown in (Figs. 5, 6 and 7), respectively. The method was tested using a database of 2,500 images of 16 people under all combinations of 3 head orientations, 3 head sizes or scales, and 3 lighting conditions and various resolutions. Recognition rates of 96%, 85% and 64% were reported for lighting, orientation and scale variation. Though the method appears to be fairly robust to lighting variations, its performance degrades with scale changes.



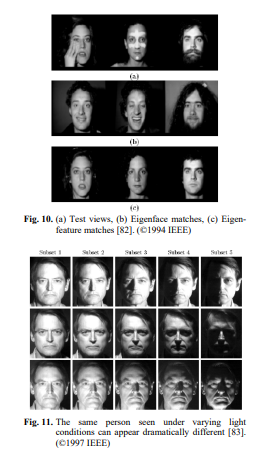


The capabilities of Turk and Pentland’s system have been extended in several ways in [82] and tested on a database of 7,562 images of approximately 3,000 people. A “multiple observer” method has been suggested to deal with large changes in pose: Given N individuals under M different views, one can either do recognition and pose estimation in a universal eigenspace calculated from the combination of NM images (parametric approach) or, alternatively, one can build a set of M separate eigenspaces, one for each of the N views (the view-based approach). The view-based approach is reported to have yielded better results than the parametric one. A modular “eigenfeatures” approach has also been proposed to deal with localized variations in the facial appearance where a low-resolution description of the whole face is augmented by additional higher resolution details in terms of the salient facial features (Fig. 8).

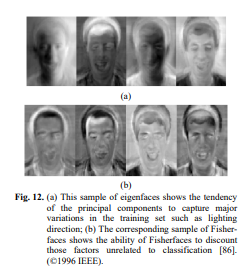
This system is reported to have produced slightly better results than the basic eigenfaces approach (Figs. 9, 10). Though no implementation has been reported, it has however been suggested in [81] that variation in scale be dealt with by employing multi-scale eigenfaces or by rescaling the input image to multiple sizes and using the scale that results in the smallest distance measure to the face space. PCA appears to work well when a single image of each individual is available, but when multiple images per person are present, then Belhumeur et al. [83] argue that by choosing the projection which maximizes total scatter, PCA retains unwanted variations due to lighting and facial expression. As stated by Moses et al. [84], “the variations between the images of the same face due to illumination and lighting direction are almost always larger than image variations due to a change in face identity” (see Fig. 11 for an example of this.) Therefore,



they propose using Fisher’s Linear Discriminant Analysis [85], which maximizes the ratio of the between-class scatter and the within-class scatter and is thus purportedly better for classification than PCA.



Conducting various tests on 330 images of 5 people (66 of each), they report that their method, called Fisherfaces, which uses subspace projection prior to LDA projection (to prevent the within-class scatter matrix from becoming degenerate), is better at simultaneously handling variations in lighting and expression. Swets and Weng [86] previously reported similar results when employing the same procedure not only for faces but also for general objects (90% accuracy on a database of 1316+298 images from 504 classes) (Fig. 12 shows some examples of eigenfaces and Fisherfaces and how Fisherfaces capture discriminatory information better than eigenfaces). It should be noted, however, that some recent work [87] shows that when the training data set is small, PCA can outperform LDA and also that PCA is less sensitive to different training sets.



The standard eigenfaces and the Fisherfaces approaches assume the existence of an optimal projection that projects the face images to distinct non-overlapping regions in the reduced subspace where each of these regions corresponds to a unique subject. However, in reality, that assumption may not necessarily be true since images of different people may frequently map to the same region in the face space and, thus, the regions corresponding to different individuals may not always be disjoint. Moghaddam et al. [88] propose an alternative approach which utilizes difference images, where a difference image for two face images is defined as the signed arithmetic difference in the intensity values of the corresponding pixels in those images. Two classes of difference images are defined: intrapersonal, which consists of difference images originating from two images of the same person, and extrapersonal, which consists of difference images derived from two images of different people. It is assumed that both these classes originate from discrete Gaussian distributions within the space of all possible difference images. Then, given the difference image between two images I1 and I2, the probability that the difference image belongs to the intrapersonal class is given by Bayes Rule as follows:



where d(I1, I2) = the difference image between two images I1 and I2 ΩI = the intrapersonal class ΩE = the extrapersonal class The formulation of the face recognition problem in this manner converts it from an m-ary classification problem (where m is the number of individuals in the database of known face images) into a binary classification problem which can be solved using the maximum a posteriori (MAP) rule – the two images are declared to belong to the same individual if P(ΩI|d(I1, I2)) > P(ΩE|d(I1, I2)) or, equivalently, if P(ΩI|d(I1, I2)) > ½. For a computationally more expedient approach, Moghaddam and Pentland [89] also suggest ignoring the extrapersonal class information and calculating the similarity based only on the intrapersonal class information. In the resulting maximum likelihood (ML) classifier, the similarity score is given only by P(d(I1, I2)| ΩI).

Numerous variations on and extensions to the standard eigenfaces and the Fisherfaces approaches have been suggested since their introduction. Some recent advances in PCA-based algorithms include multi-linear subspace analysis [90], symmetrical PCA [91], two-dimensional PCA [92, 93], eigenbands [94], adaptively weighted subpattern PCA [95], weighted modular PCA [96], Kernel PCA [97, 98] and diagonal PCA [99]. Examples of recent LDA-based algorithms include Direct LDA [100, 101], Direct-weighted LDA [102], Nullspace LDA [103, 104], Dual-space LDA [105], Pair-wise LDA [106], Regularized Discriminant Analysis [107], Generalized Singular Value Decomposition [108, 109], Direct Fractional-Step LDA [110], Boosting LDA [111], Discriminant Local Feature Analysis [112], Kernel PCA/LDA [113, 114], Kernel Scatter-Difference-based Discriminant Analysis [115], 2DLDA [116, 117], Fourier-LDA [118], Gabor-LDA [119], Block LDA [120], Enhanced FLD [121], Component-based Cascade LDA [122], and incremental LDA [123], to name but a few. All these methods purportedly obtain better recognition results than the baseline techniques.

One main drawback of the PCA and LDA methods is that these techniques effectively see only the Euclidean structure and fail to discover the underlying structure if the face images lie on a non-linear submanifold in the image space. Since it has been shown that face images possibly reside on a nonlinear submanifold [124-130] (especially if there is a perceivable variation in viewpoint, illumination or facial expression), some nonlinear techniques have consequently been proposed to discover the nonlinear structure of the manifold, e.g., Isometric Feature Mapping (ISOMAP) [130], Locally Linear Embedding (LLE) [126, 131], Laplacian Eigenmap [132], Locality Preserving Projection (LPP) [133], Embedded Manifold [134], Nearest Manifold Approach [135], Discriminant Manifold Learning [136] and Laplacianfaces [137].

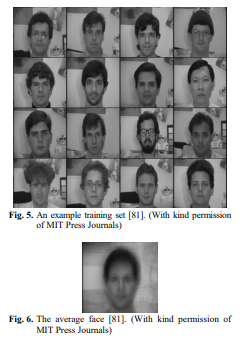
The eigenvectors found by PCA depend only on pairwise relationships between the pixels in the image database. However, other methods exist that can find basis vectors that depend on higher-order relationships among the pixels, and it seems reasonable to expect that utilizing such techniques would yield even better recognition results. Independent component analysis (ICA) [138], a generalization of PCA, is one such method that has been employed for the face recognition task. ICA aims to find an independent, rather than an uncorrelated, image decomposition and representation. Bartlett et al. [139] performed ICA on images in the FERET database under two different architectures: one treated the images as random variables and the pixels as outcomes; conversely, the second treated the pixels as the random variables and the images as outcomes. Both ICA representations outperformed PCA representations for recognizing faces across days and changes in expression. A classifier that combined both ICA representations gave the best performance. Others have also experimented with ICA [140-146] and have reported that this technique, and variations of it, appear to perform better then PCA under most circumstances.

Other subspace methods have also been exploited for the face recognition task: Foon et al. [147] have integrated various wavelet transforms and non-negative matrix factorizations [148] and claim to have obtained better verification rates as compared to the basic eigenfaces approach. In [149], an intra-class subspace is constructed, and the classification is based on the nearest weighted distance between the query face and each intra-class subspace. Experimental results are presented to demonstrate that this method performs better than some other nearest feature techniques.

A study and comparison of four subspace representations for face recognition, i.e., PCA, ICA, Fisher Discriminant Analysis (FDA), and probabilistic eigenfaces and their ‘kernalized’ versions (if available), is presented in [150]. A comprehensive review of recent advances in subspace analysis for face recognition can be found in [151].

**Statistik**

Dalam versi paling sederhana dari pendekatan holistik, gambar direpresentasikan sebagai array 2D dari nilai intensitas dan pengenalan dilakukan dengan perbandingan korelasi langsung antara wajah input dan semua wajah lain dalam database. Meskipun pendekatan ini telah terbukti bekerja [75] dalam keadaan terbatas (yaitu, pencahayaan yang sama, skala, pose, dll.), secara komputasi sangat mahal dan menderita kekurangan yang biasa dari pendekatan berbasis korelasi langsung, seperti sensitivitas terhadap orientasi wajah, ukuran, kondisi pencahayaan variabel, kekacauan latar belakang, dan kebisingan [76]. Hambatan utama untuk kinerja pengenalan metode pencocokan langsung adalah bahwa mereka mencoba untuk melakukan klasifikasi dalam ruang dimensi yang sangat tinggi [76]. Untuk melawan kutukan dimensi ini, beberapa skema lain telah diusulkan yang menggunakan metode pengurangan dimensi statistik untuk mendapatkan dan mempertahankan dimensi fitur yang paling berarti sebelum melakukan pengenalan. Beberapa di antaranya disebutkan di bawah ini.

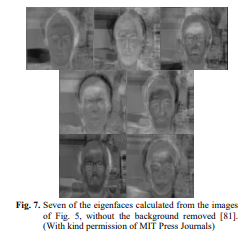


Sirovich dan Kirby [77] adalah yang pertama menggunakan Analisis Komponen Utama (PCA) [78, 79] untuk mewakili gambar wajah secara ekonomis. Mereka mendemonstrasikan bahwa setiap wajah tertentu dapat direpresentasikan secara efisien di sepanjang ruang koordinat eigenpictures, dan bahwa setiap wajah dapat direkonstruksi secara kira-kira hanya dengan menggunakan kumpulan kecil eigenpictures dan proyeksi yang sesuai ('koefisien') di sepanjang setiap eigenpicture.

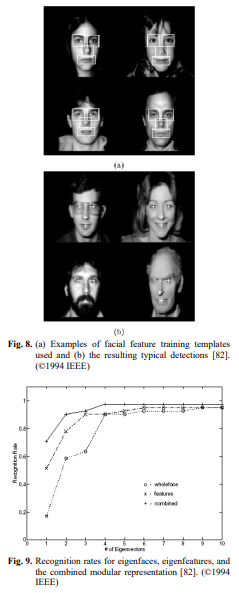
Turk dan Pentland [80, 81] menyadari, berdasarkan temuan Sirovich dan Kirby, bahwa proyeksi sepanjang gambar eigen dapat digunakan sebagai fitur klasifikasi untuk mengenali wajah. Mereka menggunakan alasan ini untuk mengembangkan sistem pengenalan wajah yang membangun eigenface, yang sesuai dengan vektor eigen yang terkait dengan nilai eigen dominan dari matriks kovarians wajah (pola) yang diketahui, dan kemudian mengenali wajah tertentu dengan membandingkan proyeksi mereka di sepanjang eigenface dengan proyeksi gambar wajah dari individu yang dikenal. Eigenfaces mendefinisikan ruang fitur yang secara drastis mengurangi dimensi ruang asli, dan identifikasi wajah dilakukan dalam ruang yang diperkecil ini. Contoh set pelatihan, wajah rata-rata dan tujuh eigenface teratas yang diturunkan dari gambar pelatihan ditunjukkan masing-masing pada (Gbr. 5, 6 dan 7). Metode ini diuji menggunakan database 2.500 gambar 16 orang di bawah semua kombinasi 3 orientasi kepala, 3 ukuran atau skala kepala, dan 3 kondisi pencahayaan dan berbagai resolusi. Tingkat pengenalan 96%, 85% dan 64% dilaporkan untuk pencahayaan, orientasi dan variasi skala. Meskipun metode ini tampaknya cukup kuat untuk variasi pencahayaan, kinerjanya menurun dengan perubahan skala.

Kemampuan sistem Turk dan Pentland telah diperluas dalam beberapa cara di [82] dan diuji pada database 7.562 gambar dari sekitar 3.000 orang. Metode "pengamat banyak" telah disarankan untuk menangani perubahan besar dalam pose: Mengingat N individu di bawah M pandangan berbeda, seseorang dapat melakukan pengenalan dan estimasi pose dalam ruang eigen universal yang dihitung dari kombinasi gambar NM (pendekatan parametrik) atau, sebagai alternatif, seseorang dapat membangun satu set M ruang eigen terpisah, satu untuk masing-masing N tampilan (pendekatan berbasis tampilan). Pendekatan berbasis tampilan dilaporkan telah memberikan hasil yang lebih baik daripada pendekatan parametrik. Pendekatan "fitur eigen" modular juga telah diusulkan untuk menangani variasi lokal dalam penampilan wajah di mana deskripsi resolusi rendah dari seluruh wajah ditambah dengan detail resolusi tambahan yang lebih tinggi dalam hal fitur wajah yang menonjol (Gbr. 8)

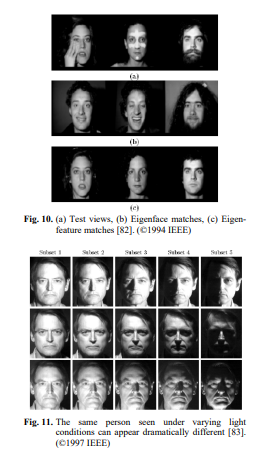
Sistem ini dilaporkan telah menghasilkan hasil yang sedikit lebih baik daripada pendekatan eigenfaces dasar (Gbr. 9, 10). Meskipun tidak ada implementasi yang dilaporkan, namun telah disarankan dalam [81] bahwa variasi dalam skala ditangani dengan menggunakan eigenfaces multi-skala atau dengan mengubah skala gambar input ke beberapa ukuran dan menggunakan skala yang menghasilkan ukuran jarak terkecil ke ruang wajah. PCA tampaknya bekerja dengan baik ketika satu gambar dari setiap individu tersedia, tetapi ketika ada banyak gambar per orang, maka Belhumeur et al. [83] berpendapat bahwa dengan memilih proyeksi yang memaksimalkan hamburan total, PCA mempertahankan variasi yang tidak diinginkan karena pencahayaan dan ekspresi wajah. Seperti yang diungkapkan oleh Musa et al. [84],



"variasi antara gambar wajah yang sama karena iluminasi dan arah pencahayaan hampir selalu lebih besar daripada variasi gambar karena perubahan identitas wajah" (lihat Gambar 11 untuk contoh ini.) Oleh karena itu, mereka mengusulkan menggunakan Analisis Diskriminan Linear Fisher [85],



yang memaksimalkan rasio pencar antara kelas dan pencar dalam kelas dan dengan demikian konon lebih baik untuk klasifikasi daripada PCA.



Melakukan berbagai tes pada 330 gambar dari 5 orang (masing-masing 66), mereka melaporkan bahwa metode mereka, yang disebut Fisherfaces, yang menggunakan proyeksi subruang sebelum proyeksi LDA (untuk mencegah matriks pencar dalam kelas menjadi merosot), lebih baik secara bersamaan menangani variasi dalam pencahayaan dan ekspresi. Swets dan Weng [86] sebelumnya melaporkan hasil yang sama ketika menggunakan prosedur yang sama tidak hanya untuk wajah tetapi juga untuk objek umum (90% akurasi pada database 1316+298 gambar dari 504 kelas) (Gbr. 12 menunjukkan beberapa contoh eigenface dan Fisherfaces dan bagaimana Fisherfaces menangkap informasi diskriminatif lebih baik daripada eigenfaces). Perlu dicatat, bagaimanapun, bahwa beberapa pekerjaan baru-baru ini [87] menunjukkan bahwa ketika set data pelatihan kecil, PCA dapat mengungguli LDA dan juga bahwa PCA kurang sensitif terhadap set pelatihan yang berbeda.

Eigenfaces standar dan pendekatan Fisherfaces mengasumsikan adanya proyeksi optimal yang memproyeksikan gambar wajah ke wilayah non-tumpang tindih yang berbeda di subruang yang dikurangi di mana masing-masing wilayah ini sesuai dengan subjek yang unik. Namun, pada kenyataannya, asumsi tersebut belum tentu benar karena citra orang yang berbeda mungkin sering dipetakan ke wilayah yang sama dalam ruang muka dan, dengan demikian, wilayah yang sesuai dengan individu yang berbeda mungkin tidak selalu terputus-putus. Moghaddam dkk. [88] mengusulkan pendekatan alternatif yang memanfaatkan gambar perbedaan, di mana gambar perbedaan untuk dua gambar wajah didefinisikan sebagai perbedaan aritmatika yang ditandatangani dalam nilai intensitas piksel yang sesuai dalam gambar tersebut. Dua kelas perbedaan gambar didefinisikan: intrapersonal, yang terdiri dari gambar perbedaan yang berasal dari dua gambar orang yang sama, dan ekstrapersonal, yang terdiri dari gambar perbedaan yang berasal dari dua gambar orang yang berbeda. Diasumsikan bahwa kedua kelas ini berasal dari distribusi Gaussian diskrit dalam ruang semua gambar perbedaan yang mungkin. Kemudian, mengingat perbedaan citra antara dua citra I1 dan I2, peluang bahwa citra selisih termasuk dalam kelas intrapersonal diberikan oleh Aturan Bayes sebagai berikut:



di mana d(I1, I2) = perbedaan citra antara dua citra I1 dan I2 I = kelas intrapersonal E = kelas ekstrapersonal Rumusan masalah pengenalan wajah dengan cara ini mengubahnya dari masalah klasifikasi m-ary (di mana m adalah jumlah individu dalam database citra wajah yang diketahui) ke dalam masalah klasifikasi biner yang dapat diselesaikan dengan menggunakan aturan maximum a posteriori (MAP) – kedua citra dinyatakan sebagai individu yang sama jika P(ΩI|d(I1 , I2)) > P(ΩE|d(I1, I2)) atau, ekuivalen, jika P(ΩI|d(I1, I2)) > . Untuk pendekatan komputasi yang lebih bijaksana, Moghaddam dan Pentland [89] juga menyarankan untuk mengabaikan informasi kelas ekstrapersonal dan menghitung kesamaan hanya berdasarkan informasi kelas intrapersonal. Dalam pengklasifikasi kemungkinan maksimum (ML) yang dihasilkan, skor kesamaan hanya diberikan oleh P(d(I1, I2)| I).

Banyak variasi dan ekstensi ke standar eigenfaces dan pendekatan Fisherfaces telah disarankan sejak diperkenalkan. Beberapa kemajuan terbaru dalam algoritma berbasis PCA termasuk analisis subruang multi-linear [90], PCA simetris [91], PCA dua dimensi [92, 93], pita eigen [94], PCA subpattern tertimbang adaptif [95], PCA modular tertimbang [96], PCA Kernel [97, 98] dan PCA diagonal [99]. Contoh algoritme berbasis LDA terbaru termasuk LDA Langsung [100, 101], LDA berbobot langsung [102], LDA Nullspace [103, 104], LDA ruang ganda [105], LDA berpasangan [106], Diskriminan Teratur Analisis [107], Dekomposisi Nilai Singular Umum [108, 109], LDA Langkah Pecahan Langsung [110], Meningkatkan LDA [111], Analisis Fitur Lokal Diskriminan [112], Kernel PCA/LDA [113, 114], Kernel Scatter -Analisis Diskriminan Berbasis Perbedaan [115], 2DLDA [116, 117], Fourier-LDA [118], Gabor-LDA [119], Blok LDA [120], Enhanced FLD [121], LDA Kaskade Berbasis Komponen [122 ], dan LDA inkremental [123], untuk menyebutkan beberapa. Semua metode ini konon memperoleh hasil pengenalan yang lebih baik daripada teknik dasar.

Salah satu kelemahan utama dari metode PCA dan LDA adalah bahwa teknik ini secara efektif hanya melihat struktur Euclidean dan gagal menemukan struktur yang mendasarinya jika gambar wajah terletak pada submanifold non-linear dalam ruang gambar. Karena telah ditunjukkan bahwa gambar wajah mungkin berada pada submanifold nonlinier [124-130] (terutama jika ada variasi yang terlihat dalam sudut pandang, iluminasi atau ekspresi wajah), beberapa teknik nonlinier telah diusulkan untuk menemukan struktur nonlinier dari manifold, misalnya, Pemetaan Fitur Isometrik (ISOMAP) [130], Penyematan Linier Lokal (LLE) [126, 131], Laplacian Eigenmap [132], Proyeksi Pelestarian Lokalitas (LPP) [133], Manifold Tertanam [134], Manifold Terdekat Pendekatan [135], Pembelajaran Berjenis Diskriminan [136] dan Laplacianfaces [137].

Vektor eigen yang ditemukan oleh PCA hanya bergantung pada hubungan berpasangan antara piksel dalam database gambar. Namun, ada metode lain yang dapat menemukan vektor basis yang bergantung pada hubungan orde tinggi di antara piksel, dan tampaknya masuk akal untuk berharap bahwa penggunaan teknik tersebut akan menghasilkan hasil pengenalan yang lebih baik. Analisis komponen independen (ICA) [138], generalisasi PCA, adalah salah satu metode yang telah digunakan untuk tugas pengenalan wajah. ICA bertujuan untuk menemukan dekomposisi dan representasi gambar yang independen, bukan tidak berkorelasi. Bartlett dkk. [139] melakukan ICA pada gambar dalam database FERET di bawah dua arsitektur yang berbeda: satu memperlakukan gambar sebagai variabel acak dan piksel sebagai hasil; sebaliknya, yang kedua memperlakukan piksel sebagai variabel acak dan gambar sebagai hasil. Kedua representasi ICA mengungguli representasi PCA untuk mengenali wajah sepanjang hari dan perubahan ekspresi. Pengklasifikasi yang menggabungkan kedua representasi ICA memberikan kinerja terbaik. Orang lain juga telah bereksperimen dengan ICA [140-146] dan telah melaporkan bahwa teknik ini, dan variasinya, tampaknya berkinerja lebih baik daripada PCA dalam sebagian besar keadaan.

Metode subruang lainnya juga telah dieksploitasi untuk tugas pengenalan wajah: Foon et al. [147] telah mengintegrasikan berbagai transformasi wavelet dan faktorisasi matriks non-negatif [148] dan mengklaim telah memperoleh tingkat verifikasi yang lebih baik dibandingkan dengan pendekatan eigenfaces dasar. Dalam [149], subruang intra-kelas dibangun, dan klasifikasi didasarkan pada jarak tertimbang terdekat antara wajah kueri dan setiap subruang intra-kelas. Hasil eksperimen disajikan untuk menunjukkan bahwa metode ini berkinerja lebih baik daripada beberapa teknik fitur terdekat lainnya.

Sebuah studi dan perbandingan empat representasi subruang untuk pengenalan wajah, yaitu, PCA, ICA, Fisher Discriminant Analysis (FDA), dan eigenfaces probabilistik dan versi 'kernalized' mereka (jika tersedia), disajikan dalam [150]. Sebuah tinjauan komprehensif kemajuan terbaru dalam analisis subruang untuk pengenalan wajah dapat ditemukan di [151].

**AI**

AI approaches utilize tools such as neural networks and machine learning techniques to recognize faces. Some examples of methods belonging to this category are given below.

In [152], 50 principal components were extracted and an auto-associative neural network was used to reduce those components to five dimensions. A standard multi-layer perceptron was exploited to classify the resulting representation. Though favorable results were received, the database used for training and testing was quite simple: the pictures were manually aligned, there was no lighting variation, tilting, or rotation, and there were only 20 people in the database.

Weng et al. [153] made use of an hierarchical neural network which was grown automatically and not trained on the traditional gradient descent method. They reported good results on a database of 10 subjects.

Lawrence et al [58] reported a 96.2% recognition rate on the ORL database (a database of 400 images of 40 individuals) using a hybrid neural network solution which combines local image sampling, a self-organizing map [154, 155] neural network (which provides dimensionality reduction and invariance to small changes in the image sample), and a convolutional neural network (which provides partial invariance to translation, rotation, scale and deformation). The eigenfaces method [80, 81] produced 89.5% recognition accuracy on the same data. Replacing the self-organizing map by the Karhunen-Loeve transform and the convolutional network by a multi-layer perceptron resulted in a recognition rate of 94.7% and 60% respectively (Fig. 13).

Eleyan and Demirel [156] used principal components analysis to obtain feature projection vectors from face images which were then classified using feed forward neural networks. Some tests on the ORL database using various numbers of training and testing images showed that the performance of this system was better than the eigenfaces [80, 81] one in which a nearest neighbor classifier was used for classification.

Li and Yin [157] introduced a system in which a face image is first decomposed with a wavelet transform to three levels. The Fisherfaces method [83] is then applied to each of the three low-frequency sub-images. Then, the individual classifiers are fused using the RBF neural network. The resulting system was tested on images of 40 subjects from the FERET database and was shown to outperform the individual classifiers and the direct Fisherfaces method.

Melin et al. [158] divided the face into three regions (the eyes, the mouth, and the nose) and assigned each region to a module of the neural network. A fuzzy Sugeno integral was then used to combine the outputs of the three modules to make the final face recognition decision. They tested it on a small database of 20 people and reported that the modular network yielded better results than a monolithic one.

Recently, Zhang et al. [159] proposed an approach in which a similarity function is learned describing the level of confidence that two images belong to the same person, similar to [88]. The facial features are selected by obtaining Local Binary Pattern (LBP) [160] histograms of the subregions of the face image and the Chi-square distances between the corresponding LBP histograms are chosen as the discriminative features. The AdaBoost learning algorithm, introduced by Freund and Schapire [161], is then applied to select the most efficient LBP features as well as to obtain the similarity function in the form of a linear combination of LBP feature-based weak learners. Experimental results on the FERET frontal image sets have shown that this method yields a better recognition rate of 97.9 % by utilizing fewer features than a previous similar approach proposed by Ahonen et al. [162].

Some researchers have also used the one-against-one approach [163] for decomposing the multi-class face recognition problem into a number of binary classification problems. In this method, one classifier is trained for each pair of classes, ignoring all the remaining ones. The outputs of all the binary classifiers are then combined to construct the global result. For binary classifiers with probabilistic outputs, pair-wise coupling (PWC) [164] can be used to couple these outputs into a set of posterior probabilities. Then, the test example is assigned to the class with the maximum posterior probability. One main disadvantage of PWC is that when a test example does not belong to either of the classes related to a binary classifier, then the output of that classifier is meaningless and can damage the global result. In [165], a new algorithm called PWC-CC (where CC stands for correcting classifier) is presented to solve this problem: for each binary classifier separating class ci from class cj, a new classifier separating the two classes from all other classes is trained. Even though PWC-CC performs better than PWC, it has its own drawbacks. In [166], a novel PWC-CC (NPWC-CC) method is proposed for the face recognition problem and the results of tests on the ORL database are presented to support the claim that it outperforms PWC-CC. In [167], the optimal PWC (O-PWC) approach is introduced and is shown to have better recognition rates that the PWC method. Feature extraction is done by using principal components analysis in [166] and by wavelet transform in [167]. In both [166] and [167], Support Vector Machines (SVMs) were employed as binary classifiers and the SVM outputs were mapped to probabilities by using the method suggested by Platt [168].

It should be noted that Support Vector Machine (SVM) is considered to be one of the most effective algorithms for pattern classification problems [169]. In general, it works as follows for binary problems [170]: First, the training examples are mapped to a high-dimensional feature space H. Then, the optimal hyperplane in H is sought to separate examples of different classes as much as possible, while maximizing the distance from either class to the hyperplane. SVM has been employed for face recognition by several other researchers and has been shown to yield good results [169, 171-175].

Hidden Markov models [176] have also been employed for the face recognition task. Samaria and Harter [177] used a one-dimensional HMM to obtain a peak recognition accuracy of 87% on the ORL database. They later upgraded the one-dimensional HMM to a pseudo two-dimensional HMM [178] and achieved a best recognition performance of 95% on the same database using half the images for training and the other half for testing. Nefian and Hayes III [179] reported a best recognition rate of 98% on the same training and testing sets using embedded HMM [180] face models, and they also claimed that their system was much faster than that of Samaria [178] and invariant to the scale of the face images.

Some other AI approaches utilized for the face recognition task include evolutionary pursuit [181, 182] and techniques [183, 184] based on boosting [161, 185]. These schemes have reportedly yielded promising results for various difficult face recognition scenarios.

**AI**

Pendekatan AI menggunakan alat seperti jaringan saraf dan teknik pembelajaran mesin untuk mengenali wajah. Beberapa contoh metode yang termasuk dalam kategori ini diberikan di bawah ini.

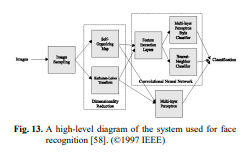
Dalam [152], 50 komponen utama diekstraksi dan jaringan saraf auto-asosiatif digunakan untuk mengurangi komponen tersebut menjadi lima dimensi. Sebuah perceptron multi-layer standar dieksploitasi untuk mengklasifikasikan representasi yang dihasilkan. Meskipun hasil yang menguntungkan diterima, database yang digunakan untuk pelatihan dan pengujian cukup sederhana: gambar disejajarkan secara manual, tidak ada variasi pencahayaan, kemiringan, atau rotasi, dan hanya ada 20 orang dalam database.

Weng dkk. [153] memanfaatkan jaringan saraf hierarkis yang tumbuh secara otomatis dan tidak dilatih pada metode penurunan gradien tradisional. Mereka melaporkan hasil yang baik pada database 10 mata pelajaran.

Lawrence et al [58] melaporkan tingkat pengenalan 96,2% pada database ORL (database 400 gambar dari 40 individu) menggunakan solusi jaringan saraf hibrida yang menggabungkan pengambilan sampel gambar lokal, peta yang mengatur sendiri jaringan saraf [154, 155] (yang memberikan pengurangan dimensi dan invarian terhadap perubahan kecil pada sampel gambar), dan jaringan saraf convolutional (yang menyediakan invarian parsial untuk translasi, rotasi, skala, dan deformasi). Metode eigenfaces [80, 81] menghasilkan akurasi pengenalan 89,5% pada data yang sama. Mengganti peta yang mengatur diri sendiri dengan transformasi Karhunen-Loeve dan jaringan konvolusi dengan perceptron multi-layer menghasilkan tingkat pengenalan masing-masing 94,7% dan 60% (Gbr. 13).

Eleyan dan Demirel [156] menggunakan analisis komponen utama untuk mendapatkan vektor proyeksi fitur dari citra wajah yang kemudian diklasifikasikan menggunakan jaringan saraf maju umpan. Beberapa pengujian pada database ORL menggunakan berbagai jumlah gambar pelatihan dan pengujian menunjukkan bahwa kinerja sistem ini lebih baik daripada eigenfaces [80, 81] di mana pengklasifikasi tetangga terdekat digunakan untuk klasifikasi.

Li dan Yin [157] memperkenalkan sistem di mana citra wajah pertama kali didekomposisi dengan transformasi wavelet menjadi tiga tingkat. Metode Fisherfaces [83] kemudian diterapkan pada masing-masing dari tiga sub-gambar frekuensi rendah. Kemudian, pengklasifikasi individu digabungkan menggunakan jaringan saraf RBF. Sistem yang dihasilkan diuji pada gambar 40 subjek dari database FERET dan terbukti mengungguli pengklasifikasi individu dan metode Fisherfaces langsung.



Melin dkk. [158] membagi wajah menjadi tiga wilayah (mata, mulut, dan hidung) dan menetapkan setiap wilayah ke modul jaringan saraf. Integral fuzzy Sugeno kemudian digunakan untuk menggabungkan output dari tiga modul untuk membuat keputusan pengenalan wajah akhir. Mereka mengujinya pada database kecil yang terdiri dari 20 orang dan melaporkan bahwa jaringan modular memberikan hasil yang lebih baik daripada jaringan monolitik.

Baru-baru ini, Zhang dkk. [159] mengusulkan pendekatan di mana fungsi kesamaan dipelajari menggambarkan tingkat kepercayaan bahwa dua gambar milik orang yang sama, mirip dengan [88]. Fitur wajah dipilih dengan memperoleh histogram Pola Biner Lokal (LBP) [160] dari subkawasan gambar wajah dan jarak Chi-kuadrat antara histogram LBP yang sesuai dipilih sebagai fitur diskriminatif. Algoritma pembelajaran AdaBoost, yang diperkenalkan oleh Freund dan Schapire [161], kemudian diterapkan untuk memilih fitur LBP yang paling efisien serta untuk mendapatkan fungsi kesamaan berupa kombinasi linier dari pembelajar lemah berbasis fitur LBP. Hasil eksperimen pada set gambar frontal FERET telah menunjukkan bahwa metode ini menghasilkan tingkat pengenalan yang lebih baik sebesar 97,9% dengan memanfaatkan lebih sedikit fitur daripada pendekatan serupa sebelumnya yang diusulkan oleh Ahonen et al. [162].

Beberapa peneliti juga menggunakan pendekatan satu lawan satu [163] untuk menguraikan masalah pengenalan wajah multi-kelas menjadi sejumlah masalah klasifikasi biner. Dalam metode ini, satu classifier dilatih untuk setiap pasangan kelas, mengabaikan semua yang tersisa. Keluaran dari semua pengklasifikasi biner kemudian digabungkan untuk membangun hasil global. Untuk pengklasifikasi biner dengan keluaran probabilistik, pair-wise coupling (PWC) [164] dapat digunakan untuk memasangkan keluaran ini ke dalam serangkaian probabilitas posterior. Kemudian, contoh tes ditugaskan ke kelas dengan probabilitas posterior maksimum. Salah satu kelemahan utama PWC adalah ketika contoh pengujian tidak termasuk dalam salah satu kelas yang terkait dengan pengklasifikasi biner, maka keluaran dari pengklasifikasi tersebut tidak berarti dan dapat merusak hasil global. Dalam [165], algoritma baru yang disebut PWC-CC (di mana CC adalah singkatan dari correcting classifier) ​​disajikan untuk memecahkan masalah ini: untuk setiap classifier biner yang memisahkan kelas ci dari kelas cj, classifier baru yang memisahkan dua kelas dari semua kelas lainnya adalah terlatih. Meskipun PWC-CC berkinerja lebih baik daripada PWC, ia memiliki kekurangannya sendiri. Dalam [166], metode PWC-CC (NPWC-CC) baru diusulkan untuk masalah pengenalan wajah dan hasil tes pada database ORL disajikan untuk mendukung klaim bahwa itu mengungguli PWC-CC. Dalam [167], pendekatan PWC (O-PWC) yang optimal diperkenalkan dan terbukti memiliki tingkat pengenalan yang lebih baik daripada metode PWC. Ekstraksi ciri dilakukan dengan menggunakan analisis komponen utama pada [166] dan dengan transformasi wavelet pada [167]. Dalam [166] dan [167], Support Vector Machines (SVMs) digunakan sebagai pengklasifikasi biner dan output SVM dipetakan ke probabilitas dengan menggunakan metode yang disarankan oleh Platt [168].

Perlu dicatat bahwa Support Vector Machine (SVM) dianggap sebagai salah satu algoritma yang paling efektif untuk masalah klasifikasi pola [169]. Secara umum, cara kerjanya sebagai berikut untuk masalah biner [170]: Pertama, contoh pelatihan dipetakan ke ruang fitur berdimensi tinggi H. Kemudian, hyperplane optimal di H dicari untuk memisahkan contoh kelas yang berbeda sebanyak mungkin, sambil memaksimalkan jarak dari kedua kelas ke hyperplane. SVM telah digunakan untuk pengenalan wajah oleh beberapa peneliti lain dan telah terbukti memberikan hasil yang baik [169, 171-175].

Model Hidden Markov [176] juga telah digunakan untuk tugas pengenalan wajah. Samaria dan Harter [177] menggunakan HMM satu dimensi untuk mendapatkan akurasi pengenalan puncak 87% pada database ORL. Mereka kemudian meningkatkan HMM satu dimensi menjadi HMM dua dimensi semu [178] dan mencapai kinerja pengenalan terbaik 95% pada database yang sama menggunakan setengah gambar untuk pelatihan dan setengah lainnya untuk pengujian. Nefian dan Hayes III [179] melaporkan tingkat pengenalan terbaik 98% pada set pelatihan dan pengujian yang sama menggunakan model wajah HMM [180] tertanam, dan mereka juga mengklaim bahwa sistem mereka jauh lebih cepat daripada Samaria [178] dan invarian dengan skala gambar wajah.

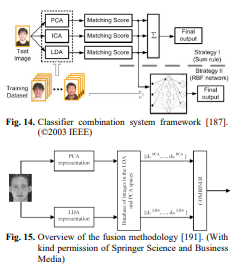
Beberapa pendekatan AI lain yang digunakan untuk tugas pengenalan wajah termasuk pengejaran evolusioner [181, 182] dan teknik [183, 184] berdasarkan peningkatan [161, 185]. Skema ini dilaporkan telah memberikan hasil yang menjanjikan untuk berbagai skenario pengenalan wajah yang sulit.

**Multiple Classifier Systems**

Since the performance of any classifier is more sensitive to some factors and relatively invariant to others, a recent trend has been to combine individual classifiers in order to integrate their complementary information and thereby create a system that is more robust than any individual classifier to variables that complicate the recognition task. Such systems have been termed as multiple classifier systems (MCSs) [186] and are a very active research area at present. Examples of such approaches proposed for face recognition include the following: Lu et al. [187] fused the results of PCA, ICA and LDA using the sum rule and RBF network-based [188] integration strategy (Fig. 14); Marcialis and Roli [189-191] combined the results of the PCA and LDA algorithms (Fig. 15); Achermann and Bunke [192] utilized simple fusion rules (majority voting, rank sum, Baye’s combination rule) to integrate the weighted outcomes of three classifiers based on frontal and profile views of faces; Tolba and Abu-Rezq [193] employed a simple combination rule for fusing the decisions of RBF and LVQ networks; Wan et al. [194] used a SVM and HMM hybrid model; Kwak and Pedrycz [195] divided the face into three regions, applied the Fisherfaces method to the regions as well as to the whole face and then integrated the classification results using the Choquet fuzzy integral [196]; Haddadnia et. al. [197] used PCA, the Pseudo Zernike Moment Invariant (PZMI) [198, 199] and the Zernike Moment Invariant (ZMI) to extract feature vectors in parallel, which were then classified simultaneously by separate RBF neural networks and the outputs of these networks were then combined by a majority rule to determine the final identity of the individual in the input image.

**Sistem Pengklasifikasi Ganda**

Karena kinerja setiap pengklasifikasi lebih sensitif terhadap beberapa faktor dan relatif invarian terhadap yang lain, tren baru-baru ini adalah menggabungkan pengklasifikasi individu untuk mengintegrasikan informasi pelengkap mereka dan dengan demikian menciptakan sistem yang lebih kuat daripada pengklasifikasi individu mana pun untuk variabel yang mempersulit tugas pengenalan. Sistem tersebut telah disebut sebagai sistem pengklasifikasi ganda (MCS) [186] dan merupakan area penelitian yang sangat aktif saat ini. Contoh pendekatan yang diusulkan untuk pengenalan wajah adalah sebagai berikut: Lu et al. [187] menggabungkan hasil PCA, ICA dan LDA menggunakan aturan penjumlahan dan strategi integrasi [188] berbasis jaringan RBF (Gbr. 14); Marcialis dan Roli [189-191] menggabungkan hasil dari algoritma PCA dan LDA (Gbr. 15); Achermann dan Bunke [192] menggunakan aturan fusi sederhana (pemungutan suara mayoritas, jumlah peringkat, aturan kombinasi Baye) untuk mengintegrasikan hasil pembobotan dari tiga pengklasifikasi berdasarkan tampilan frontal dan profil wajah; Tolba dan Abu-Rezq [193]



menggunakan aturan kombinasi sederhana untuk menggabungkan keputusan jaringan RBF dan LVQ; Wan dkk. [194] menggunakan model hibrida SVM dan HMM; Kwak dan Pedrycz [195] membagi wajah menjadi tiga region, menerapkan metode Fisherfaces pada region maupun seluruh wajah kemudian mengintegrasikan hasil klasifikasi menggunakan integral fuzzy Choquet [196]; Haddadnia et. Al. [197] menggunakan PCA, Pseudo Zernike Moment Invariant (PZMI) [198, 199] dan Zernike Moment Invariant (ZMI) untuk mengekstrak vektor fitur secara paralel, yang kemudian diklasifikasikan secara bersamaan oleh jaringan saraf RBF terpisah dan output dari jaringan ini kemudian digabungkan dengan aturan mayoritas untuk menentukan identitas akhir individu dalam gambar masukan.

**Advantages and Disadvantages**

The main advantage of the holistic approaches is that they do not destroy any of the information in the images by concentrating on only limited regions or points of interest [37]. However, as mentioned above, this same property is their greatest drawback, too, since most of these approaches start out with the basic assumption that all the pixels in the image are equally important [74]. Consequently, these techniques are not only computationally expensive but require a high degree of correlation between the test and training images, and do not perform effectively under large variations in pose, scale and illumination, etc. [200]. Nevertheless, as mentioned in the above review, several of these algorithms have been modified and/or enhanced to compensate for such variations, and dimensionality reduction techniques have been exploited (note that even though such techniques increase generalization capabilities, the downside is that they may potentially cause the loss of discriminative information [151]), as a result of which these approaches appear to produce better recognition results than the feature-based ones in general. In the latest comprehensive FERET evaluation [62, 63], the probabilistic eigenface [88], the Fisherface [83] and the EBGM [59] methods were ranked as the best three techniques for face recognition (Even though the EBGM method is feature-based in general, its success depends on its application of holistic neural network methods at the feature level).

**Keuntungan dan Kerugian**

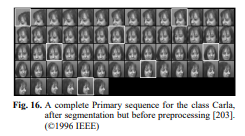
Keuntungan utama dari pendekatan holistik adalah bahwa mereka tidak menghancurkan informasi apa pun dalam gambar dengan berkonsentrasi hanya pada wilayah atau tempat tertentu yang terbatas [37]. Namun, seperti disebutkan di atas, properti yang sama ini juga merupakan kelemahan terbesarnya, karena sebagian besar pendekatan ini dimulai dengan asumsi dasar bahwa semua piksel dalam gambar sama pentingnya [74]. Akibatnya, teknik ini tidak hanya mahal secara komputasi tetapi memerlukan tingkat korelasi yang tinggi antara gambar uji dan pelatihan, dan tidak bekerja secara efektif di bawah variasi besar dalam pose, skala dan pencahayaan, dll. [200]. Namun demikian, seperti yang disebutkan dalam tinjauan di atas, beberapa dari algoritma ini telah dimodifikasi dan/atau ditingkatkan untuk mengimbangi variasi tersebut, dan teknik pengurangan dimensi telah dieksploitasi (perhatikan bahwa meskipun teknik tersebut meningkatkan kemampuan generalisasi, kelemahannya adalah bahwa mereka mungkin berpotensi menyebabkan hilangnya informasi diskriminatif [151]), akibatnya pendekatan ini tampaknya menghasilkan hasil pengenalan yang lebih baik daripada yang berbasis fitur pada umumnya. Dalam evaluasi FERET komprehensif terbaru [62, 63], metode eigenface probabilistik [88], Fisherface [83] dan EBGM [59] diperingkatkan sebagai tiga teknik terbaik untuk pengenalan wajah (Meskipun metode EBGM adalah fitur- berdasarkan secara umum, keberhasilannya tergantung pada penerapan metode jaringan saraf holistik di tingkat fitur).

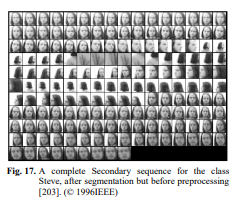
**Face Recognition from Video Sequences**

Since one of the major applications of face recognition is surveillance for security purposes, which involves realtime recognition of faces from an image sequence captured by a video camera, a significant amount of research has been directed towards this area in recent years.

A video-based face recognition system typically consists of three modules: one for detecting the face; a second one for tracking it; and a third one for recognizing it [201]. Most of these systems choose a few good frames and then apply one of the recognition techniques for intensity images to those frames in order to identify the individual [202]. A few of these approaches are briefly described below.

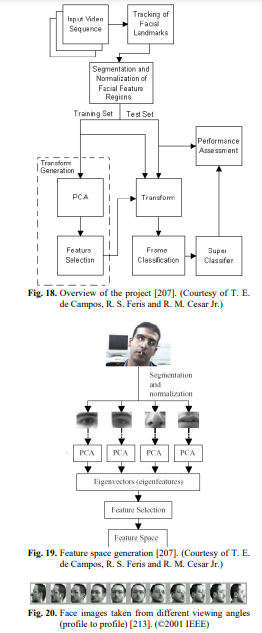
Howell and Buxton [203] employed a two-layer RBF network [204, 205] for learning/training and used Difference of Gaussian (DoG) filtering and Gabor wavelet analysis for the feature representation, while the scheme from [206] was utilized for face detection and tracking. Training and testing were done using two types of image sequences: 8 primary sequences taken in a relatively constrained environment, and a secondary sequence recorded in a much more unconstrained atmosphere (Figs. 16, 17). The image sequences consisted of 62 to 94 frames. The use of Gabor wavelet analysis for feature representation, as opposed to DoG filtering, seemed to yield better recognition results. The recognition accuracies reported varied quite a bit, ranging from 99%, using 278 images for training and 276 for testing, to 67%, using 16 training and 538 testing images.

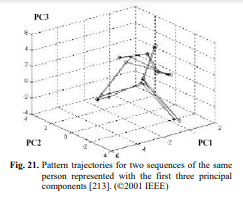




de Campos et al. [207] propose a recognition system which uses skin color modeling [208] to detect the face, then utilizes GWN [209] to detect prominent facial landmarks (i.e., the eyes, nose, and mouth) and to track those features. For each individual frame, eigenfeatures [82] are then extracted and a feature selection algorithm [210] is applied over the combination of all the eigenfeatures, and the best ones are selected to form the feature space. A couple of classifiers described in [211] are then applied to identify the individual in the frame and, finally, a superclassifier based on a voting scheme [212] performs the final classification for the entire video sequence (Figs. 18, 19). Good recognition results (97.7% accuracy) have been reported using 174 images of the eyes of 29 people (6 images per person).

Biuk and Loncaric [213] take image sequences in which the pose of the person’s face changes from -90 degrees to 90 degrees (Fig. 20). The sequences are projected into eigenspace to give a prototype trajectory for each known individual. During the recognition phase, an unknown face trajectory is compared with the prototype trajectories to determine the identity of the individual (Fig. 21). They tested the system on a database of 28 individuals (11 frames per person) and found that the system yielded excellent recognition results when all frames and 4 or more eigenfaces were used, but that the performance decreased when either parameter was decreased. They, however, propose a matching method which associates a score to each trajectory point and then makes the final match based on the maximum score: They claim that this enhancement enables their system to achieve better performance when a smaller number (< 4) of eigenfaces is used.





Recently, some approaches [214, 215] have employed a video-to-video paradigm in which information from a sequence of frames from a video segment is combined and associated with one individual. This notion involves a temporal analysis of the video sequence and a condensation of the tracking and recognition problems. Such schemes are still a matter of ongoing research since the reported experiments were performed without any real variations in orientation and facial expressions [216].

It is worth mentioning that several schemes have incorporated information from other modalities in order to recognize facial images acquired from video clips. For instance, [217] makes use of stereo information and reports a recognition accuracy of 90%, while [8] exploits both audio and video cues as well as 3D information about the head to achieve a 100% accuracy rate for 26 subjects. (For more information about recognition based on audio and video, see the Proceedings of the AVBPA Conferences [218]).

A detailed survey of recent schemes for face recognition from video sequences is provided in [219].

**Pengenalan Wajah dari Urutan Video**

Karena salah satu aplikasi utama pengenalan wajah adalah pengawasan untuk tujuan keamanan, yang melibatkan pengenalan wajah secara realtime dari urutan gambar yang diambil oleh kamera video, sejumlah besar penelitian telah diarahkan ke area ini dalam beberapa tahun terakhir.

Sistem pengenalan wajah berbasis video biasanya terdiri dari tiga modul: satu untuk mendeteksi wajah; yang kedua untuk melacaknya; dan yang ketiga untuk mengenalinya [201]. Sebagian besar sistem ini memilih beberapa frame yang baik dan kemudian menerapkan salah satu teknik pengenalan citra intensitas ke frame tersebut untuk mengidentifikasi individu [202]. Beberapa dari pendekatan ini dijelaskan secara singkat di bawah ini.

Howell dan Buxton [203] menggunakan jaringan RBF dua lapis [204, 205] untuk pembelajaran/pelatihan dan menggunakan pemfilteran Difference of Gaussian (DoG) dan analisis wavelet Gabor untuk representasi fitur, sedangkan skema dari [206] digunakan untuk deteksi dan pelacakan wajah. Pelatihan dan pengujian dilakukan dengan menggunakan dua jenis sekuens gambar: 8 sekuens utama yang diambil dalam lingkungan yang relatif terbatas, dan sekuens sekunder yang direkam dalam suasana yang jauh lebih bebas (Gbr. 16, 17). Urutan gambar terdiri dari 62 hingga 94 frame. Penggunaan analisis wavelet Gabor untuk representasi fitur, yang bertentangan dengan pemfilteran DoG, tampaknya menghasilkan hasil pengenalan yang lebih baik. Akurasi pengenalan yang dilaporkan cukup bervariasi, mulai dari 99%, menggunakan 278 gambar untuk pelatihan dan 276 untuk pengujian, hingga 67%, menggunakan 16 gambar pelatihan dan 538 gambar pengujian.

de Campos dkk. [207] mengusulkan sistem pengenalan yang menggunakan pemodelan warna kulit [208] untuk mendeteksi wajah, kemudian menggunakan GWN [209] untuk mendeteksi landmark wajah yang menonjol (yaitu, mata, hidung, dan mulut) dan untuk melacak fitur tersebut. Untuk setiap frame individu, fitur eigen [82] kemudian diekstraksi dan algoritma pemilihan fitur [210] diterapkan pada kombinasi semua fitur eigen, dan yang terbaik dipilih untuk membentuk ruang fitur. Beberapa pengklasifikasi yang dijelaskan dalam [211] kemudian diterapkan untuk mengidentifikasi individu dalam bingkai dan, akhirnya, pengklasifikasi super berdasarkan skema pemungutan suara [212] melakukan klasifikasi akhir untuk seluruh urutan video (Gbr. 18, 19). Hasil pengenalan yang baik (akurasi 97,7%) telah dilaporkan menggunakan 174 gambar mata dari 29 orang (6 gambar per orang).

Biuk dan Loncaric [213] mengambil urutan gambar di mana pose wajah seseorang berubah dari -90 derajat menjadi 90 derajat (Gbr. 20). Urutan diproyeksikan ke eigenspace untuk memberikan lintasan prototipe untuk setiap individu yang diketahui. Selama fase pengenalan, lintasan wajah yang tidak diketahui dibandingkan dengan lintasan prototipe untuk menentukan identitas individu (Gbr. 21). Mereka menguji sistem pada database 28 individu (11 frame per orang) dan menemukan bahwa sistem menghasilkan hasil pengenalan yang sangat baik ketika semua frame dan 4 atau lebih eigenfaces digunakan, tetapi kinerja menurun ketika salah satu parameter diturunkan. Namun, mereka mengusulkan metode pencocokan yang mengaitkan skor ke setiap titik lintasan dan kemudian membuat pertandingan final berdasarkan skor maksimum: Mereka mengklaim bahwa peningkatan ini memungkinkan sistem mereka mencapai kinerja yang lebih baik ketika jumlah eigenface yang lebih kecil (<4) digunakan.

Baru-baru ini, beberapa pendekatan [214, 215] telah menggunakan paradigma video-to-video di mana informasi dari urutan frame dari segmen video digabungkan dan dikaitkan dengan satu individu. Gagasan ini melibatkan analisis temporal dari urutan video dan kondensasi masalah pelacakan dan pengenalan. Skema seperti itu masih menjadi bahan penelitian yang sedang berlangsung karena eksperimen yang dilaporkan dilakukan tanpa variasi nyata dalam orientasi dan ekspresi wajah [216].

Perlu disebutkan bahwa beberapa skema telah memasukkan informasi dari modalitas lain untuk mengenali gambar wajah yang diperoleh dari klip video. Misalnya, [217] memanfaatkan informasi stereo dan melaporkan akurasi pengenalan 90%, sementara [8] memanfaatkan isyarat audio dan video serta informasi 3D tentang kepala untuk mencapai tingkat akurasi 100% untuk 26 subjek. (Untuk informasi lebih lanjut tentang pengenalan berdasarkan audio dan video, lihat Prosiding Konferensi AVBPA [218]).

Sebuah survei rinci skema terbaru untuk pengenalan wajah dari urutan video disediakan di [219].

**Advantages and Disadvantages**

Dynamic face recognition schemes appear to be at a disadvantage relative to their static counterparts in general, since they are usually hampered by one or more of the following: low quality images (though image quality may be enhanced by exploiting super-resolution techniques [220-223]); cluttered backgrounds (which complicate face detection [224]); the presence of more than one face in the picture; and a large amount of data to process [71]. Furthermore, the face image may be much smaller than the size required by most systems employed by the recognition modules [202].

However, dynamic schemes do have the following advantages over static techniques: the enormous abundance of data empowers the system to choose the frame with the best possible image and discard less satisfactory ones [203]. Video provides temporal continuity [203], so classification information from several frames can be combined to improve recognition performance. Moreover, video allows the tracking of face images such that variations in facial expressions and poses can be compensated for, resulting in improved recognition [225]. Dynamic schemes also have an edge over static ones when it comes to detecting the face in a scene, since these schemes can use motion to segment a moving person’s face [71].

**Keuntungan dan Kerugian**

Skema pengenalan wajah dinamis tampaknya kurang menguntungkan dibandingkan dengan skema statis pada umumnya, karena biasanya terhambat oleh satu atau lebih hal berikut: gambar berkualitas rendah (meskipun kualitas gambar dapat ditingkatkan dengan mengeksploitasi teknik resolusi super [220- 223]); latar belakang yang berantakan (yang mempersulit deteksi wajah [224]); kehadiran lebih dari satu wajah dalam gambar; dan sejumlah besar data untuk diproses [71]. Selanjutnya, citra wajah mungkin jauh lebih kecil dari ukuran yang dibutuhkan oleh sebagian besar sistem yang digunakan oleh modul pengenalan [202].

Namun, skema dinamis memiliki keuntungan sebagai berikut dibandingkan teknik statis: data yang sangat melimpah memungkinkan sistem untuk memilih bingkai dengan gambar terbaik dan membuang yang kurang memuaskan [203]. Video menyediakan kontinuitas temporal [203], sehingga informasi klasifikasi dari beberapa frame dapat digabungkan untuk meningkatkan kinerja pengenalan. Selain itu, video memungkinkan pelacakan gambar wajah sehingga variasi ekspresi wajah dan pose dapat dikompensasi, menghasilkan pengenalan yang lebih baik [225]. Skema dinamis juga memiliki keunggulan dibandingkan skema statis dalam mendeteksi wajah dalam sebuah adegan, karena skema ini dapat menggunakan gerakan untuk mensegmentasi wajah orang yang bergerak [71].

**Face Recognition from Other Sensory Inputs**

Though the bulk of the research on face recognition has been focused on identifying individuals from 2D intensity images, in recent years some attention has nevertheless been directed towards exploiting other sensing modalities, such as 3D or range data and infra-red imagery, for this purpose.

**Pengenalan Wajah dari Input Sensorik Lainnya**

Meskipun sebagian besar penelitian tentang pengenalan wajah telah difokuskan untuk mengidentifikasi individu dari gambar intensitas 2D, dalam beberapa tahun terakhir beberapa perhatian telah diarahkan untuk mengeksploitasi modalitas penginderaan lainnya, seperti data 3D atau rentang dan citra infra merah, untuk tujuan ini.

**3D Model-based**

The main argument in favor of using 3D information for face recognition appears to be that it allows us to exploit features based on the shape and the curvature of the face (such as the shape of the forehead, jaw line, and cheeks) without being plagued by the variances caused by lighting, orientation and background clutter that affect 2D systems [37, 226, 227]. Another argument for the use of depth data is that “at our current state of technology, it is the most straightforward way to input or record complex shape information for machine analysis” [228]. The obvious drawbacks of such approaches are their complexity and computational cost [202].

The following techniques are currently being used to obtain 3D information [226]:

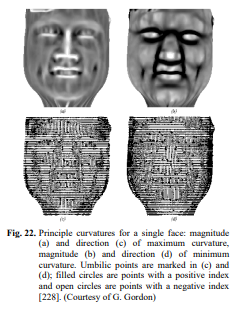
•Scanning systems: Laser face scanners produced by companies like Cyberware Inc. [229] and 3D Scanners Ltd. [230] seem to be producing highly accurate results; however, the cost of these commercial scanning services is obviously substantial.

•Structured light systems: These systems make use of the principles of stereo vision to obtain the range data. Their main advantage is that the only equipment they require is cameras and some kind of projection system. The primary drawback with such systems is that they can experience difficulty in resolving the shape of the pattern in the camera image.

•Stereo vision systems: These are systems that attempt to extract 3D information from two or more 2D images of the same object taken from different angles. They are limited to objects which will “generate a sufficient number of image features to allow for conclusive stereo matching. In the case of trying to establish the shape of a reasonably smooth object, such as the human face, these systems would be unable to generate an accurate surface shape. (Smooth surfaces can be 'roughed up' by the projection of a textured pattern onto the face)” [226].

•Reverse rendering/shape from shading: These techniques endeavor to construct the shape of an object using knowledge about illumination and the physical properties of the object.

Until recently, there appear to have been very few papers that describe attempts to recognize faces based mainly on range data or 3D data about the subjects’ faces. However, lately, there has been a revival of interest in this area and several new schemes of this sort have been proposed in the past few years. One of the earliest of such approaches is described in [228], where the principle curvatures of the face surface are calculated from range data (Fig. 22), after which this data - supplemented by a priori information about the structure of the face - is used to locate the various facial features (i.e., the nose eyes, forehead, neck, chin, etc.). The faces are then normalized to a standard position and re-interpolated onto a regular cylindrical grid. The volume of space between two normalized surfaces is used as a similarity measure. The system was tested using the face images of 8 people (3 images per person). Features were detected adequately for all faces. Recognition rates of 97% and 100% were reported for individual features and the whole face respectively.



Another approach described in [200] uses profiles (which have external contours consisting of rather rigid parts) instead of frontal images, captures 3D data by triangulation, and then does a 3D comparison of the profile data (Figs. 23, 24). This method requires a good deal of user cooperation and restrictions on the background, etc. in order for it to work.

[37] uses 3D data to normalize the results obtained from the face detection algorithm to a form more appropriate for the recognition engine, i.e., in this case, the 3D data is just being used to supplement rather than supplant existing face detection and recognition algorithms.

Some examples of more recent face recognition approaches based on 3D data include the following: Castellani et al. [231] approximate the range images of faces obtained by stereoscopic analysis using Multi-level B-Splines [232], and SVMs are then used to classify the resulting approximation coefficients. Some other techniques [227, 233, 234] first project the 3D face data onto a 2D intensity image, whereupon the projected 2D images are processed as standard intensity images. Yet other methods have been proposed for 3D face recognition based on local features [235], local and global geometric cues [236], profiles [237-240], and the rank-based decision fusion of various shape-based classifiers [241].

Several approaches have also been proposed that integrate 2D texture and 3D shape information. Such methods make use of the PCA of intensity images [242- 244], facial profile intensities [245], Iterative Closest Point (ICP [246]) [247, 248], Gabor wavelets [249], and Local Feature Analysis [250], etc. For instance, Wang et al. [249] extract 3D shape templates from range images and texture templates from grayscale images of faces, apply PCA separately to both kinds of templates to reduce them to lower-dimensional vectors, then concatenate the shape and texture vectors and, finally, apply SVMs to the resulting vectors for classification. In general, experiments with such systems indicate that combining shape and texture information reduces the misclassification rates of the face recognizer.

Comprehensive recent surveys of literature on 3D face recognition can be found in [251] and [252].

**Berbasis Model 3D**

Argumen utama yang mendukung penggunaan informasi 3D untuk pengenalan wajah tampaknya adalah bahwa hal itu memungkinkan kita untuk mengeksploitasi fitur berdasarkan bentuk dan kelengkungan wajah (seperti bentuk dahi, garis rahang, dan pipi) tanpa terganggu oleh varians yang disebabkan oleh pencahayaan, orientasi dan kekacauan latar belakang yang mempengaruhi sistem 2D [37, 226, 227]. Argumen lain untuk penggunaan data kedalaman adalah bahwa "pada kondisi teknologi kita saat ini, ini adalah cara paling mudah untuk memasukkan atau merekam informasi bentuk kompleks untuk analisis mesin" [228]. Kelemahan yang jelas dari pendekatan tersebut adalah kompleksitas dan biaya komputasi [202].

Teknik berikut saat ini digunakan untuk mendapatkan informasi 3D [226]:

• Sistem pemindaian: Pemindai wajah laser yang diproduksi oleh perusahaan seperti Cyberware Inc. [229] dan 3D Scanners Ltd. [230] tampaknya memberikan hasil yang sangat akurat; namun, biaya layanan pemindaian komersial ini jelas besar.

• Sistem cahaya terstruktur: Sistem ini menggunakan prinsip penglihatan stereo untuk mendapatkan data jangkauan. Keuntungan utama mereka adalah bahwa satu-satunya peralatan yang mereka butuhkan adalah kamera dan semacam sistem proyeksi. Kelemahan utama dengan sistem tersebut adalah bahwa mereka dapat mengalami kesulitan dalam menyelesaikan bentuk pola pada gambar kamera.

• Sistem penglihatan stereo: Ini adalah sistem yang mencoba mengekstrak informasi 3D dari dua atau lebih gambar 2D dari objek yang sama yang diambil dari sudut yang berbeda. Mereka terbatas pada objek yang akan “menghasilkan sejumlah fitur gambar yang cukup untuk memungkinkan pencocokan stereo yang meyakinkan. Dalam kasus mencoba menetapkan bentuk objek yang cukup halus, seperti wajah manusia, sistem ini tidak akan mampu menghasilkan bentuk permukaan yang akurat. (Permukaan halus dapat 'dikasar' dengan proyeksi pola bertekstur ke wajah)” [226].

• Rendering/bentuk terbalik dari shading: Teknik ini berusaha untuk membangun bentuk objek menggunakan pengetahuan tentang iluminasi dan sifat fisik objek.

Sampai saat ini, tampaknya hanya ada sedikit makalah yang menjelaskan upaya untuk mengenali wajah terutama berdasarkan data jangkauan atau data 3D tentang wajah subjek. Namun, akhir-akhir ini, ada kebangkitan minat di bidang ini dan beberapa skema baru semacam ini telah diusulkan dalam beberapa tahun terakhir. Salah satu yang paling awal dari pendekatan tersebut dijelaskan dalam [228], di mana kelengkungan prinsip permukaan wajah dihitung dari data rentang (Gbr. 22), setelah itu data ini - dilengkapi dengan informasi apriori tentang struktur wajah - digunakan untuk menemukan berbagai fitur wajah (yaitu, mata hidung, dahi, leher, dagu, dll.). Wajah-wajah tersebut kemudian dinormalisasi ke posisi standar dan diinterpolasi kembali ke dalam kotak silinder biasa. Volume ruang antara dua permukaan yang dinormalisasi digunakan sebagai ukuran kesamaan. Sistem diuji menggunakan citra wajah 8 orang (3 citra per orang). Fitur terdeteksi secara memadai untuk semua wajah. Tingkat pengenalan masing-masing 97% dan 100% dilaporkan untuk fitur individual dan seluruh wajah.

Pendekatan lain yang dijelaskan dalam [200] menggunakan profil (yang memiliki kontur eksternal yang terdiri dari bagian yang agak kaku) daripada gambar frontal, menangkap data 3D dengan triangulasi, dan kemudian melakukan perbandingan 3D dari data profil (Gbr. 23, 24). Metode ini membutuhkan banyak kerja sama pengguna dan pembatasan pada latar belakang, dll. agar dapat berfungsi.

[37] menggunakan data 3D untuk menormalkan hasil yang diperoleh dari algoritma deteksi wajah ke bentuk yang lebih sesuai untuk mesin pengenalan, yaitu, dalam hal ini, data 3D hanya digunakan untuk melengkapi daripada menggantikan algoritma deteksi dan pengenalan wajah yang ada. .

Beberapa contoh pendekatan pengenalan wajah yang lebih baru berdasarkan data 3D adalah sebagai berikut: Castellani et al. [231] memperkirakan rentang gambar wajah yang diperoleh dengan analisis stereoskopik menggunakan Multi-level B-Splines [232], dan SVM kemudian digunakan untuk mengklasifikasikan koefisien aproksimasi yang dihasilkan. Beberapa teknik lain [227, 233, 234] pertama-tama memproyeksikan data wajah 3D ke gambar intensitas 2D, di mana gambar 2D yang diproyeksikan diproses sebagai gambar intensitas standar. Namun metode lain telah diusulkan untuk pengenalan wajah 3D berdasarkan fitur lokal [235], isyarat geometrik lokal dan global [236], profil [237-240], dan fusi keputusan berbasis peringkat dari berbagai pengklasifikasi berbasis bentuk [241] .

Beberapa pendekatan juga telah diusulkan yang mengintegrasikan tekstur 2D dan informasi bentuk 3D. Metode tersebut menggunakan PCA dari citra intensitas [242-244], intensitas profil wajah [245], Titik Terdekat Iteratif (ICP [246]) [247, 248], wavelet Gabor [249], dan Analisis Fitur Lokal [250 ], dll. Misalnya, Wang et al. [249] mengekstrak templat bentuk 3D dari berbagai gambar dan templat tekstur dari gambar skala abu-abu wajah, menerapkan PCA secara terpisah ke kedua jenis templat untuk mereduksinya menjadi vektor berdimensi lebih rendah, lalu menggabungkan vektor bentuk dan tekstur dan, akhirnya, menerapkan SVM ke vektor yang dihasilkan untuk klasifikasi. Secara umum, eksperimen dengan sistem tersebut menunjukkan bahwa menggabungkan informasi bentuk dan tekstur mengurangi tingkat kesalahan klasifikasi dari pengenal wajah.

Survei literatur terbaru yang komprehensif tentang pengenalan wajah 3D dapat ditemukan di [251] dan [252].

**Infra-red**

Since thermal infra-red imagery of faces is relatively insensitive to variations in lighting [253], such images can hence be used as an option for detecting and recognizing faces. Furthermore, [254] argues that since infra-red facial images reveal the vein and tissue structure of the face which is unique to each individual (like a fingerprint), some of the face recognition techniques for the visible spectrum should therefore yield favorable results when applied to these images. However, there exist a multitude of factors that discourage the exploitation of such images for the face recognition task, among which figure the substantial cost of thermal sensors, the low resolution and high level of noise in the images, the lack of widely available data sets of infra-red images, the fact of infra-red radiation being opaque to glass (making it possible to occlude part of the face by wearing eyeglasses) [255] and, last but not least, the fact that infra-red images are sensitive to changes in ambient temperature, wind and metabolic processes in the subject [256] (Note that in [257], the use of blood perfusion data is suggested to alleviate the effect of ambient temperature).

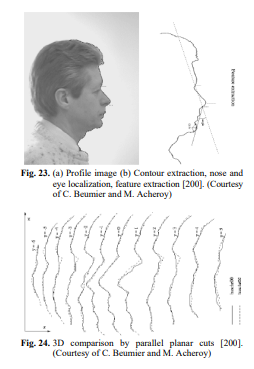
In [254], the eigenface technique [80, 81] was applied to a database of 288 hand-aligned low-resolution (160x120) images of 24 subjects taken from 3 viewpoints. The following recognition rates were reported: 96% for frontal views, 96% for 45 degrees views, and 100% for profile views.

Wilder et al. [258] compared the performance of three face recognition algorithms on a database of visible and infra-red images of 101 subjects and concluded that the recognition results for one modality were not significantly better than those for the other.

Socolinsky et al. [259] tested the eigenfaces [80, 81] and the ARENA [260] algorithms on a database of visible and infrared images of 91 distinct subjects (captured under various illumination conditions, with varying facial expressions, and with or without glasses using a sensor capable of imaging both modalities simultaneously [Figs. 25, 26 and 27]), and reported that the infra-red imagery significantly outperformed the visible one in all the classification experiments conducted under the various above-mentioned conditions. Selinger and Socolinsky [256] used the same database of 91 subjects and tested the performance of four face recognition algorithms (PCA, LDA, LFA and ICA) under the afore-mentioned conditions and reached the same conclusion, although they did concede that the apparent superiority of the infra-red approach may stem from the fact that their data did not contain sufficiently challenging situations (i.e., changes in temperature, wind, etc.) for the infra-red imagery, whereas it did so for the visible images.

Chen et al. [262] collected several datasets of images (both in infra-red and the visible spectrum) of 240 distinct subjects under various expressions and lighting conditions at different times (some of the images were taken in the same session while others were taken over a period of ten weeks [Fig. 28]). They studied the effect of temporal changes in facial appearance on the performance of the eigenfaces algorithm in both modalities and concluded that though the recognition accuracy was approximately the same for images taken in the same session, visible images outperformed the infra-red ones when there was a significant lapse in the time between which the training and test images were acquired. They attributed the lower performance of the infra-red imagery to variations in the thermal patterns of the same subject and the sensitivity of the infra-red imagery to the manual location of the eyes. They also found that the FACEIT [263] software performed better than eigenfaces in both modalities. However, the combination of the two classifiers using the sum rule [264] outperformed the individual classifiers as well as the FACEIT program. Latter experiments conducted by Chen et al. [261] on a larger set of data reconfirmed these results.

Other approaches have also been proposed recently which explore the fusion of face images captured under visible and infrared light spectrum to improve the performance of face recognition [255, 265-267].



A comprehensive review of recent advances in face recognition from infra-red imagery may be found in [268].

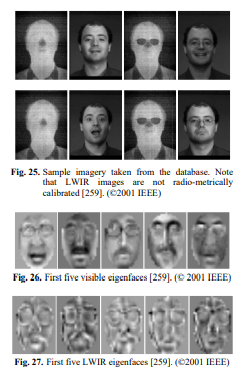
**Inframerah**

Karena citra termal inframerah wajah relatif tidak sensitif terhadap variasi pencahayaan [253], citra tersebut dapat digunakan sebagai opsi untuk mendeteksi dan mengenali wajah. Selanjutnya, [254] berpendapat bahwa karena citra wajah infra-merah mengungkapkan vena dan struktur jaringan wajah yang unik untuk setiap individu (seperti sidik jari), beberapa teknik pengenalan wajah untuk spektrum yang terlihat seharusnya memberikan hasil yang menguntungkan ketika diterapkan pada gambar-gambar ini. Namun, ada banyak faktor yang menghambat eksploitasi gambar tersebut untuk tugas pengenalan wajah, di antaranya adalah biaya sensor termal yang besar, resolusi rendah dan tingkat noise yang tinggi pada gambar, kurangnya kumpulan data yang tersedia secara luas. gambar infra-merah, fakta radiasi infra-merah yang buram ke kaca (memungkinkan untuk menutupi bagian dari wajah dengan memakai kacamata) [255] dan, terakhir namun tidak kalah pentingnya, fakta bahwa gambar infra-merah sensitif terhadap perubahan suhu lingkungan, angin dan proses metabolisme pada subjek [256] (Perhatikan bahwa dalam [257], penggunaan data perfusi darah disarankan untuk mengurangi efek suhu lingkungan).

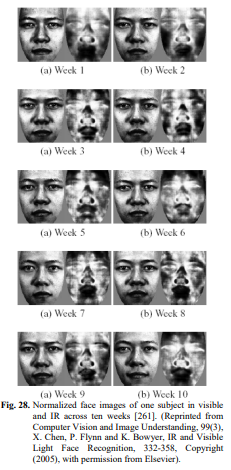
Dalam [254], teknik eigenface [80, 81] diterapkan pada database 288 gambar resolusi rendah (160x120) yang disejajarkan dengan tangan dari 24 subjek yang diambil dari 3 sudut pandang. Tingkat pengenalan berikut dilaporkan: 96% untuk tampilan depan, 96% untuk tampilan 45 derajat, dan 100% untuk tampilan profil.

Wilder dkk. [258] membandingkan kinerja tiga algoritma pengenalan wajah pada database gambar terlihat dan infra-merah dari 101 subjek dan menyimpulkan bahwa hasil pengenalan untuk satu modalitas tidak secara signifikan lebih baik daripada yang lain.

Sokolinsky dkk. [259] menguji algoritme eigenfaces [80, 81] dan ARENA [260] pada database gambar tampak dan inframerah dari 91 subjek berbeda (diambil dalam berbagai kondisi pencahayaan, dengan berbagai ekspresi wajah, dan dengan atau tanpa kacamata menggunakan sensor mampu mencitrakan kedua modalitas secara bersamaan [Gbr. 25, 26 dan 27]), dan melaporkan bahwa citra infra-merah secara signifikan mengungguli citra yang terlihat di semua eksperimen klasifikasi yang dilakukan di bawah berbagai kondisi yang disebutkan di atas. Selinger dan Socolinsky [256] menggunakan database yang sama dari 91 subjek dan menguji kinerja empat algoritma pengenalan wajah (PCA, LDA, LFA dan ICA) di bawah kondisi yang disebutkan di atas dan mencapai kesimpulan yang sama, meskipun mereka mengakui bahwa keunggulan pendekatan infra-merah mungkin berasal dari fakta bahwa data mereka tidak mengandung situasi yang cukup menantang (yaitu, perubahan suhu, angin, dll.) untuk citra infra-merah, sedangkan hal itu terjadi untuk gambar yang terlihat.

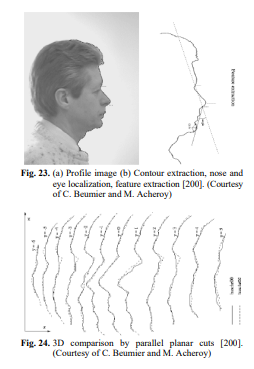


Chen dkk. [262] mengumpulkan beberapa kumpulan data gambar (baik dalam spektrum infra-merah dan tampak) dari 240 subjek berbeda di bawah berbagai ekspresi dan kondisi pencahayaan pada waktu yang berbeda (beberapa gambar diambil dalam sesi yang sama sementara yang lain diambil selama satu periode sepuluh minggu [Gbr. 28]). Mereka mempelajari efek perubahan temporal dalam penampilan wajah pada kinerja algoritma eigenfaces di kedua modalitas dan menyimpulkan bahwa meskipun akurasi pengenalan kira-kira sama untuk gambar yang diambil dalam sesi yang sama, gambar yang terlihat mengungguli yang infra-merah ketika ada selang waktu yang signifikan antara pelatihan dan gambar uji diperoleh. Mereka mengaitkan kinerja yang lebih rendah dari citra infra-merah dengan variasi dalam pola termal dari subjek yang sama dan sensitivitas citra infra-merah ke lokasi manual mata.



Mereka juga menemukan bahwa perangkat lunak FACEIT [263] berkinerja lebih baik daripada eigenfaces di kedua modalitas. Namun, kombinasi dari dua pengklasifikasi menggunakan aturan penjumlahan [264] mengungguli pengklasifikasi individu serta program FACEIT. Eksperimen terakhir dilakukan oleh Chen et al. [261] pada kumpulan data yang lebih besar menegaskan kembali hasil ini.

Pendekatan lain juga telah diusulkan baru-baru ini yang mengeksplorasi perpaduan gambar wajah yang diambil di bawah spektrum cahaya tampak dan inframerah untuk meningkatkan kinerja pengenalan wajah [255, 265-267].



Sebuah tinjauan komprehensif kemajuan terbaru dalam pengenalan wajah dari citra infra-merah dapat ditemukan di [268].

**Conclusions**

Face recognition is a challenging problem in the field of image analysis and computer vision that has received a great deal of attention over the last few years because of its many applications in various domains. Research has been conducted vigorously in this area for the past four decades or so, and though huge progress has been made, encouraging results have been obtained and current face recognition systems have reached a certain degree of maturity when operating under constrained conditions; however, they are far from achieving the ideal of being able to perform adequately in all the various situations that are commonly encountered by applications utilizing these techniques in practical life. The ultimate goal of researchers in this area is to enable computers to emulate the human vision system and, as has been aptly pointed out by Torres [225], “Strong and coordinated effort between the computer vision, signal processing, and psychophysics and neurosciences communities is needed” to attain this objective.

**Kesimpulan**

Pengenalan wajah adalah masalah yang menantang di bidang analisis citra dan visi komputer yang telah menerima banyak perhatian selama beberapa tahun terakhir karena banyak aplikasinya di berbagai domain. Penelitian telah dilakukan dengan giat di bidang ini selama empat dekade terakhir atau lebih, dan meskipun kemajuan besar telah dibuat, hasil yang menggembirakan telah diperoleh dan sistem pengenalan wajah saat ini telah mencapai tingkat kematangan tertentu ketika beroperasi dalam kondisi terbatas; namun, mereka jauh dari mencapai cita-cita untuk dapat tampil secara memadai di semua berbagai situasi yang biasa dihadapi oleh aplikasi yang menggunakan teknik ini dalam kehidupan praktis. Tujuan akhir para peneliti di bidang ini adalah untuk memungkinkan komputer meniru sistem penglihatan manusia dan, seperti yang telah ditunjukkan dengan tepat oleh Torres [225], “Upaya yang kuat dan terkoordinasi antara visi komputer, pemrosesan sinyal, dan komunitas psikofisika dan ilmu saraf diperlukan” untuk mencapai tujuan ini.

# Bibliography

Susim, T., & Darujati, C. (2021). PENGOLAHAN CITRA UNTUK PENGENALAN WAJAH (FACE. *Jurnal Syntax Admiration*.

Z. Li, S., & Jain, A. (2004). Handbook of. *ACADEMIA*.

Jafri, R., & R. Arabnia, H. (2009). A Survey of Face Recognition Techniques. *Journal of Information Processing Systems*.

man, a. (2003). Face recognition: A literature survey. *ACADEMIA*.

A. Selinger and D. A. Socolinsky, "AppearanceBased Facial Recognition Using Visible and Thermal  
Imagery: A Comparative Study," Equinox Corporation,  
Technical Report 02-01 February 2002.