



Universitetet
i Stavanger

DET TEKNISK-NATURVITENSKAPELIGE FAKULTET

MASTEROPPGAVE

Studieprogram/spesialisering: Informasjonsteknologi, signal og bildebehandling	Vårsemesteret, 2010 Åpen
Forfatter: Inge Rasch Sandvik (signatur forfatter)
Fagansvarlig: Ivar Austvoll	
Tittel på masteroppgaven: En litteraturstudie i ansiktsgjenkjenning	
Studiepoeng: 30	
Emneord: Ansiktsgjenkjenning, P.C.A., eigenfaces, symmetri, MatLab	Sidetall: 48 + vedlegg/annet: 8 Stavanger, 15. juni, 201

Innholdsfortegnelse

Sammendrag	4
Kapittel 1 - Introduksjon.....	5
Kapittel 2 - Ansiktsgjenkjenning	7
2.1 Hva menes med ansiktsgjenkjenning?	7
2.2 Hvordan foregår ansiktsgjenkjenning?.....	8
2.3 En oversikt over fagfeltet	10
2.4 Forskning i andre fagfelt som er relevante for ansiktsgjenkjenning	17
Er ansiktsgjenkjenning en dedikert prosess?.....	17
Er oppfatningen av ansikter basert på analyse av kjennetegn?.....	17
Rangering av viktighet for ansiktskjennetegn.....	18
Karikaturer	19
Særpreg i ansikter	19
Rollen romlig frekvensanalyse har i ansiktsgjenkjenning	19
Er gjenkjenning uavhengig av perspektiv?.....	19
Effekt av lysforandring	19
Bevegelse og ansiktsgjenkjenning	20
Ansiktsuttrykk	20
Påvirkning på fagfeltet	20
2.5 Bildedatabaser	21
The AR Face Database.....	22
The Facial Recognition Technology (FERET) Database	22
Face Recognition Data, University of Essex	22
Databasene Yale og Yale B	22
Andre databaser.....	23
Kapittel 3 - Utplukking av ansiktets egenskaper	24
3.1 "Face space"	24
3.2 Prinsipalkomponentanalyse, P.C.A.....	26
3.3 Teorien bak P.C.A.....	27
3.4 Ansiktssymmetri	32
Kapittel 4 – Eksperimenter og resultater	34
4.1 Implementering	34

4.2 Eksperiment 1	36
4.3 Eksperiment 2	41
Kapittel 5 - Konklusjon	44
Kapittel 6 - Referanser.....	46
Kapittel 7 Appendiks	49
7.1 ansiktsgjenkjenning.m	50
7.2 lag_treningsdatabase.m	51
7.4 egenbilder.m.....	52
7.5 projeksjon.m	54
7.6 lag_testdatabase.m	54
7.7 gjenkjenning.m	55

Sammendrag

Innledningsvis legges grunnlaget for masteroppgaven fram. Dette beskriver noen bruksområder for ansiktsgjenkjenning, og aktualiserer behovet for denne typen systemer. Videre konkretiseres problemstillingen, og er formulert på følgende vis: "Denne masteroppgaven har til hensikt å være en litteraturstudie av fagfeltet ansiktsgjenkjenning, hvor en tar utgangspunkt i artikkelen "Symmetry, probability, and recognition in face space" av Lawrence Sirovich og Marsha Meytils, [1], fra desember 2008. Den vil også forsøke å gjenskape resultatene fra denne artikkelen".

Hovedresultatene en har kommet fram til i litteraturstudiet er en nærmere forklaring av begrepet ansiktsgjenkjenning, og de forskjellige delene dette består av. En har også presentert tilstanden for fagfeltet. Det har blitt brukt mye tid på prinsipalkomponentanalyse, og implementasjon av denne metoden i MatLab ref. Implementeringen har dannet grunnlaget for å forsøke å gjenskape resultatene fra [1].

Av hovedinntrykk, sitter en igjen med følgende:

Ansiktsgjenkjenning er et fagfelt som har opplevd mye framgang de siste 20 årene. Spesielt ansiktsgjenkjenning i kontrollerte omgivelser begynner å nærme seg perfektjonert, og en beveger seg mot at "state of the art"-algoritmene er eid og beskyttet av kommersielle selskaper, og er ikke tilgjengelig.

Forskning innenfor fagfeltet har nå begynt å ta for seg de vanskelige problemstillingene, som kan gjøre ansiktsgjenkjenning mer anvendelig for praktiske formål hvor en har ukontrollerte forhold. Her er det snakk om gjenkjenning fra video, gjerne med lav pikseloppløsning, slik som mange overvåkingskameraer. En har også tatt fatt i gjenkjenning hvor det høy variasjon i belysning og hvilket perspektiv en ser bildene fra.

Forsøket på å gjenskape resultatene fra [1] har vært nokså vellykket. Fra eksperiment 1 klarer en å få feilrater som er i nærheten av de som presenteres i artikkelen. Resultatene fra eksperiment 2 støtter derimot ikke hypotesen om sammenhengen mellom sidebelysning og odde egenverdier fra [1].

Kapittel 1 - Introduksjon

25. august 2009 meldte CBS at ansiktsgjenkjenning hadde stoppet et identitetstyveri. Historien som fortelles er at en mann i USA oppsøker "Bureau of Motor Vehicles" for å skaffe seg et ID-kort, som kan utgi han for å være en annen person. Han har alle de riktige papirene, men da han skal ta bilde til det nye ID-kortet, blir han identifisert av ansiktsgjenkjenningssystemet. Ansiktet hans blir gjenkjent som 10 andre personer i databasen, og mannen blir politianmeldt og siktet for identitetstyveri.

Å bruke ansiktsgjenkjenning som et ledd i bekjemping av kriminalitet, har vært et populært mål blant politikere i tiden etter 11. september 2001, spesielt i USA og Storbritannia. Det har blitt foretatt mange prøveprosjekter på blant annet automatisk identifikasjon av kriminelle på flyplasser og lignende offentlige steder, rundt om i verden. Så langt det er forfatteren kjent, har det ikke blitt gjort noen forsøk på å benytte ansiktsgjenkjenning i større omfang i Norge, verken fra privat eller offentlig sektor. En har derimot benyttet en annen type biometrisk gjenkjenning (engelsk: biometric recognition), nemlig fingeravtrykklesing på flyplasser, for å bekrefte at passasjerene som går om bord i fly, er de samme som har sjekket inn.

Ansiktsgjenkjenning har fått mye kritikk for å være upresist. På en flyplass i Tampa, i USA, fortelles det at ansiktsgjenkjenningssystemet kun identifiserte flyplassansatte i 53 % av tilfellene, i 2009. En har altså langt i fra perfektionert ansiktsgjenkjenning, men som gjenkjenningssystem har det fortsatt en del fordeler over andre former for biometrisk gjenkjenning, slik som for eksempel irislesing eller fingeravtrykklesing. Hovedargumentet for å bruke ansiktsgjenkjenning, er at systemet slipper direkte kontakt med brukerne. Altså, en slipper at personene som skal gjenkjennes må legge fingeren på en fingeravtrykksleser, eller må stikke øynene inn i en irisleser. Det eneste en må gjøre mot et ansiktsgjenkjenningssystem, er å smile til kameraet. Eller helst bør en nok egentlig la være å smile, da gjenkjenningssystemene oftest virker best på nøytrale ansikter. Av andre fordeler kan en dra fram kostnadene forbundet med ansiktsgjenkjenning. Selv kameraer med relativt høy oppløsning, og god optikk er blitt relativt billig de siste årene.

En ser altså fra denne innledende teksten, at ansiktsgjenkjenning fortsatt er et fagfelt det er interesse for å videreutvikle, og forsøke å perfektionere. Et viktig argument som brukes for å fortsette forskning, er at vi fortsatt vet for lite om hvordan menneskenes hjerner fungerer. Legger en dette faktum sammen med at mennesker er svært dyktige på ansiktsgjenkjenning, er det fortsatt mange som er av den oppfatning at det er mulig å gjøre gjenkjenningssystemer bedre.

Denne masteroppgaven har derfor til hensikt å være en litteraturstudie av fagfeltet ansiktsgjenkjenning, hvor en tar utgangspunkt i artikkelen "Symmetry, probability, and recognition in face space" av Lawrence Sirovich og Marsha Meytils, [1], fra desember 2008. Den vil også forsøke å gjenskape resultatene fra denne artikkelen. Innholdet i artikkelen vil bli tatt opp grundigere i teorikapitlene, men som en innledning kan det sies at den bygger videre på en del av Lawrence Sirovichs tidligere forskning, og tar for seg oppgaven med å finne sannsynlighetsfordelingen for egenvektorene, som i fagmiljøet ofte kalles egenansikter (engelsk: eigenfaces), til et sett med en stor samling med bilder av ansikter. Disse egenvektorene danner grunnlaget for gjenkjenning av ansikter fra settet. Konklusjonen som dras i artikkelen bygger på resultatet fra [2], som sier at dersom en utvider settet, til å inkludere bildene speilet om midten, vil en få egenvektorer som fordeler seg som like og odde. Dersom en ser bilder av hvordan de odde og like egenvektorene blir seende ut, se Figur 3.5, kan en se hvordan Sirovich og Meytils videre konkluderer med at de odde egenvektorene må komme fra støy, i form av sidebelysning. Det oppdages videre at de fleste av disse odde egenvektorene opptrer med relativt lav sannsynlig. Konklusjonen som dras, noe forenklet, er dermed at dersom en kun baserer gjenkjenning på de like egenvektorene, vil en fjerne store deler av støy fra sidebelysning, og kunne oppnå bedre gjenkjenning.

Valget av masteroppgave kommer fra forfatterens ståsted om at resultatene, fra denne artikkelen, kan være viktige funn innen ansiktsgjenkjenning, og de ser ut til å være bidragsgivende til at en kommer litt nærmere et mer nøyaktig estimat for hvor mye informasjon menneskets hjerne benytter til gjenkjenning av ansikter. Arbeidet i oppgaven vil bestå av to deler: en litteraturstudie og en eksperimentdel. Litteraturstudien vil ta for seg en del artikler som er relevante for [1], og en del artikler som tar for seg andre temaer innen ansiktsgjenkjenning. I litteraturdelen fokuseres det først og fremst på de artiklene som er relevante for [1]. Eksperimentdelen vil prøve å gjenskape og verifisere resultatene fra [1], ved hjelp av matematikkverktøyet MatLab. En vil her også adressere noen uklarheter rundt resultatene fra artikkelen, slik som at det ikke foreligger noen resultatet for hvordan gjenkjenning uten symmetriutvidingen, som anses for å være den mest vanlige måten å gjøre gjenkjenning på i dag, måler seg mot de nye resultatene.

Kapittel 2 - Ansiktsgjenkjenning

2.1 Hva menes med ansiktsgjenkjenning?

Ordet ansiktsgjenkjenning blir ofte brukt om flere forskjellige oppgaver. Dette skyldes kanskje mest at det engelske språket har hatt mye innflytelse på ordet, og som kjent er det ikke alltid slik at alle oversetter tekster, ord og uttrykk likt. I denne masteroppgaven har det blitt valgt å bruke ordet på samme måte som det engelske "face recognition" stort sett brukes, og dermed valgt at det omfatter gjenkjenning av ansikter. Dette er ikke det samme som ansiktsdeteksjon, som i denne oppgaven tilsvarer det engelske "face detection".

Ansiktsgjenkjenning refererer altså til den prosessen som har et antall identifiserte og uidentifiserte ansikter, og som finner ut hvilket av de identifiserte ansiktene de uidentifiserte ansiktene ligner mest på. I noen tilfeller setter en også en slags terskel i gjenkjenningen, slik at en også kan avgjøre at et uidentifisert ansikt ikke tilhører et av de identifiserte individene. Altså, ansiktsgjenkjenning betegner det at en kan knytte et ansikt til et individ.

Ofte blir ansiktsgjenkjenning forvekslet med ansiktsdeteksjon. Dårlig oversatte manualer til intelligente kameraer må ta en del av skylden for dette, da en del av disse betegner ansiktsdeteksjon som ansiktsgjenkjenning. Ansiktsdeteksjon refererer til prosessen hvor en ønsker å oppdage et ansikt i et bilde. Denne prosessen kan være til stor hjelp i ansiktsgjenkjenning, da en kan finne ansiktet i et bilde. Siden en vet "hva som er ansikt" kan en fjerne mye av informasjonen som ikke er ansikt, før en prøver å sammenligne informasjonen med den en har for de identifiserte individene. I kameraer brukes funksjonen til å avgjøre hva som skal stilles skarpt, siden en som oftest ønsker å få menneskene i bildet til å være i fokus. I denne masteroppgaven har fokuset vært på ansiktsgjenkjenning, og derfor kommer ikke ansiktsdeteksjon til å bli diskutert.

2.2 Hvordan foregår ansiktsgjenkjenning?

I utgangspunktet deler en vanligvis ansiktsgjenkjenning inn i tre deler. Dette gjøres for eksempel i [3]. Del 1 består vanligvis av en eller annen form for ansiktsdeteksjon. Dette kan for eksempel gjøres av en algoritme, som finner et ansikt i et bilde, og tar vekk uvesentlig informasjon. En annen mulighet er at et menneske visuelt detekterer ansiktet, tar vekk uvesentlig informasjon for hånd. Uvesentlig informasjon refererer her til bildebakgrunn, og eventuelle andre objekter som ikke er tilknyttet ansiktet. Det er også vanlig å knytte en slags normalisering av bildene til denne delen. Da skalerer, roterer og translaterer en ansikter slik at de kan ses på som sammenlignbare. Sirovich og Meytils skalerer ansiktene, i bildene de benytter i [1], slik at avstanden mellom øynene er den samme for alle bildene, og roterer de slik at en kan trekke en tenkt vertikal linje mellom øynene, slik som i Figur 2.1. Bildene kan nå være på en form slik som i Figur 2.2.



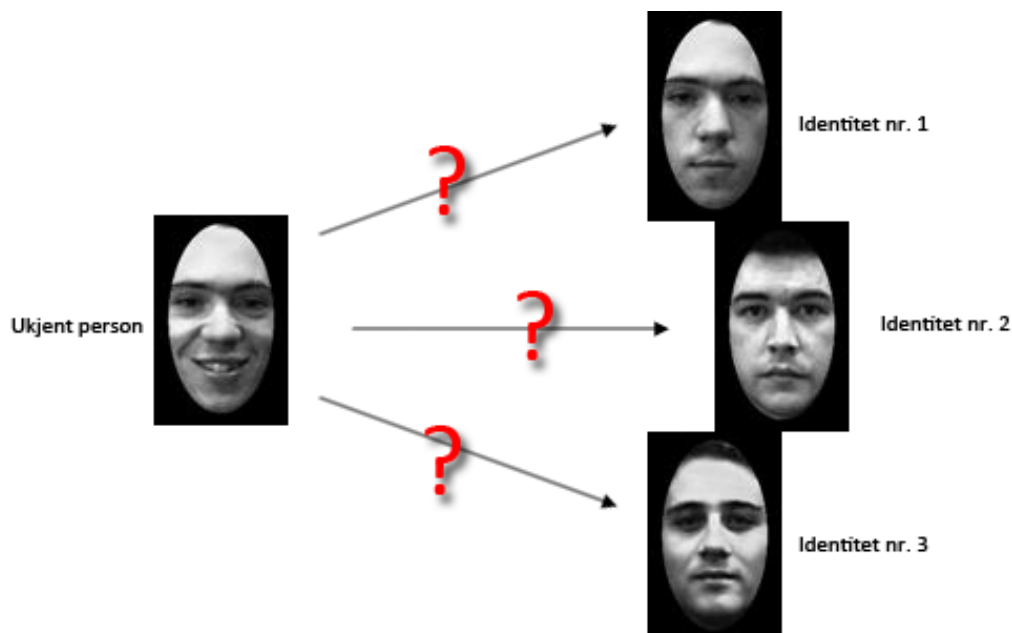
Figur 2.1: Et ansikt er vanligvis ganske symmetrisk om en tenkt symmetrilinje mellom øynene. Figuren er laget med et bilde fra "Face Recognition Data, University of Essex", [4].



Figur 2.2: Prinsipiell illustrasjon av hvordan et ansikt detekteres og normaliseres. Figuren er laget med bilder fra "Face Recognition Data, University of Essex", [4].

Del 2 av ansiktsgjenkjenningen inneholder mye større variasjon i metodene som tas i bruk, men et hovedtrekk er at en vanligvis forsøker å komme fram til kun den informasjonen som er nødvendig for gjenkjenning, for hvert bilde. I kapittel 2.3 blir det sett litt nærmere på de mest dominerende metodene for dette steget. I dette arbeidet er det prinsipalkomponentanalyse (engelsk: Principal Component Analysis), ofte forkortet til P.C.A..

Til slutt, i del 3, gjør en den faktiske gjenkjenningen. Også her varierer metodene som benyttes relativt mye. Det overordnede målet i dette steget er at en sammenligner egenskapene, fra del 2, fra et bilde med kjent identitetstilhørighet med et bilde med ukjent identitetstilhørighet. Avhengig av hva gjenkjenningssystemet skal oppnå, dras en konklusjon på problemstillingen. Dette kan være at egenskapene er like nok til at det antas at bildene tilhører samme identitet. En annen mulighet kan være at en har et sett med ansikter, og bestemmer hvilket ansikt det ukjente ansiktet ligner mest på. Hva gjenkjenningen skal oppnå, er forsøkt illustrert i Figur 2.3.

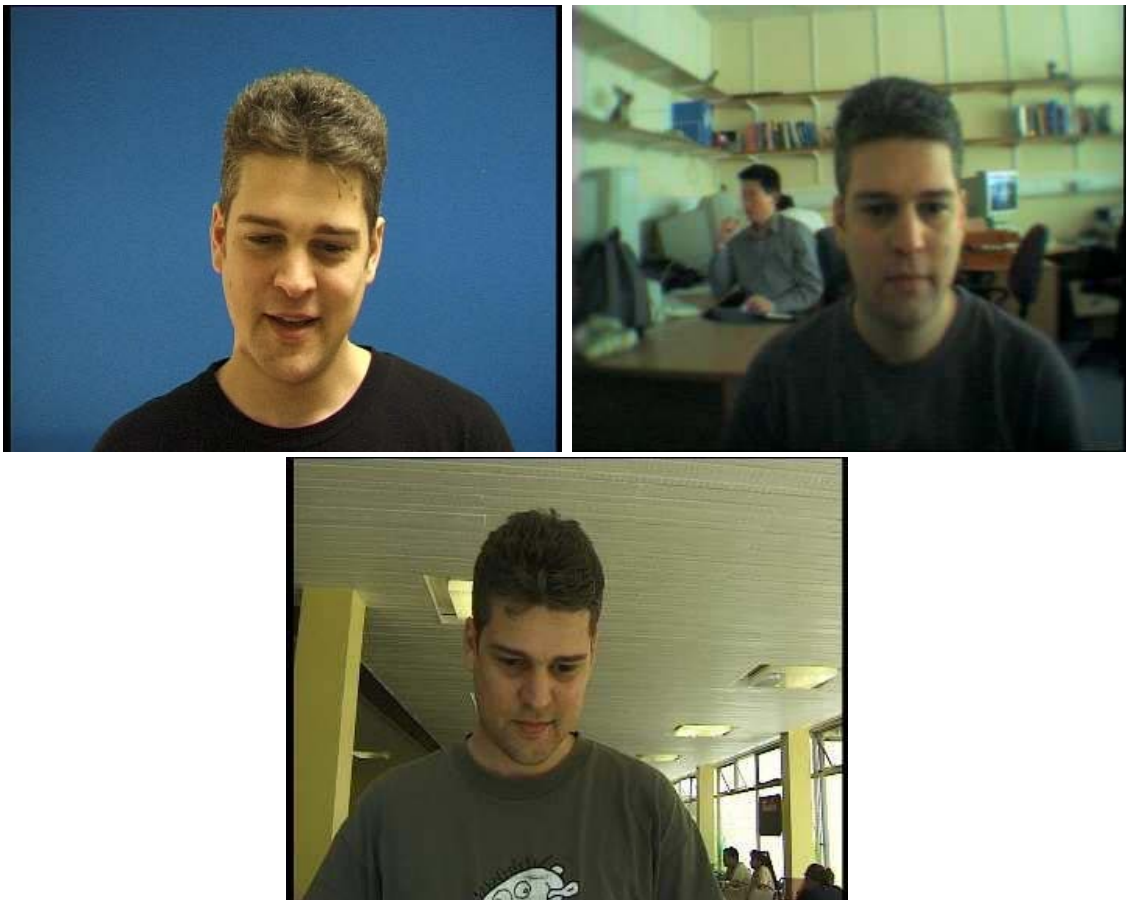


Figur 2.3: Ansiktsgjenkjenning handler om å avgjøre om en person, i et bilde, er den samme som personen i et annet bilde. Figuren er laget med bilder fra "Face Recognition Data, University of Essex", [4].

2.3 En oversikt over fagfeltet

Automatisk gjenkjenning av ansikter er fortsatt et utfordrende problem, til tross for at det har fått mye oppmerksomhet de siste førti årene. Selv til dags dato er det ingen metode som har utmerket seg som robust for alle anvendelser av ansiktsgjenkjenning. Luis Torres, [3], forsøkte i 2004 å svare på spørsmålet: "Er det håp for ansiktsgjenkjenning?". Han konkluderer at en ikke vil ha en god løsning på ansiktsgjenkjenningsproblemet på noen år enda, for mer komplekse situasjoner. For enklere, klart avgrensede situasjoner, hevder han at det fortsatt er håp. I årene etter 2004 har en sett en del framgang i denne typen problemer. Historien i starten på innledningen er bevis på dette.

Torres forsøker også å sette fokus på de forskjellige vanskelighetsgradene i problemer innen ansiktsgjenkjenning, i [3]. I problemer, som anses som enkle, har en liten variasjon i forholdene rundt hvordan bildene som skal sammenlignes fotograferes, og i hvordan ansiktet i bildet er posisjonert i forhold til kameraet. Det er også liten variasjon i ansiktsuttrykk mellom bildene. Når en snakker om forholdene rundt fotograferingen, menes det her forhold som belysning, avstand fra kamera til personen i bildet, fokus, antall piksel i det digitaliserte bildet og så videre. Eksempler på enkle problemstillinger kan ses i Figur 2.4. Denne typen problemer anses som mer eller mindre løst, men videre forskning er nyttig for å redusere feilrater ytterligere.



Figur 2.4: Eksempel på enkle problemstillinger innen ansiktsgjenkjenning. Bildene er hentet fra BANCA bildedatabasen, se [5].

Straks problemene går utenfor forholdene som karakteriserer de som enkle, øker vanskelighetsgraden betraktelig, og feilratene øker betraktelig. Blant vanskelige problemstillinger kan en trekke fram de viktigste. Disse er: stor variasjon i ansiktsuttrykk mellom bildene, stor variasjon i belysning mellom bildene, store tidsintervaller mellom fotografering av bildene, dårlig pikseloppløsning på bildene, alle former for tildekte ansikter, skjegg, barter og briller. Eksempler på vanskelige problemstillinger kan ses i Figur 2.5.

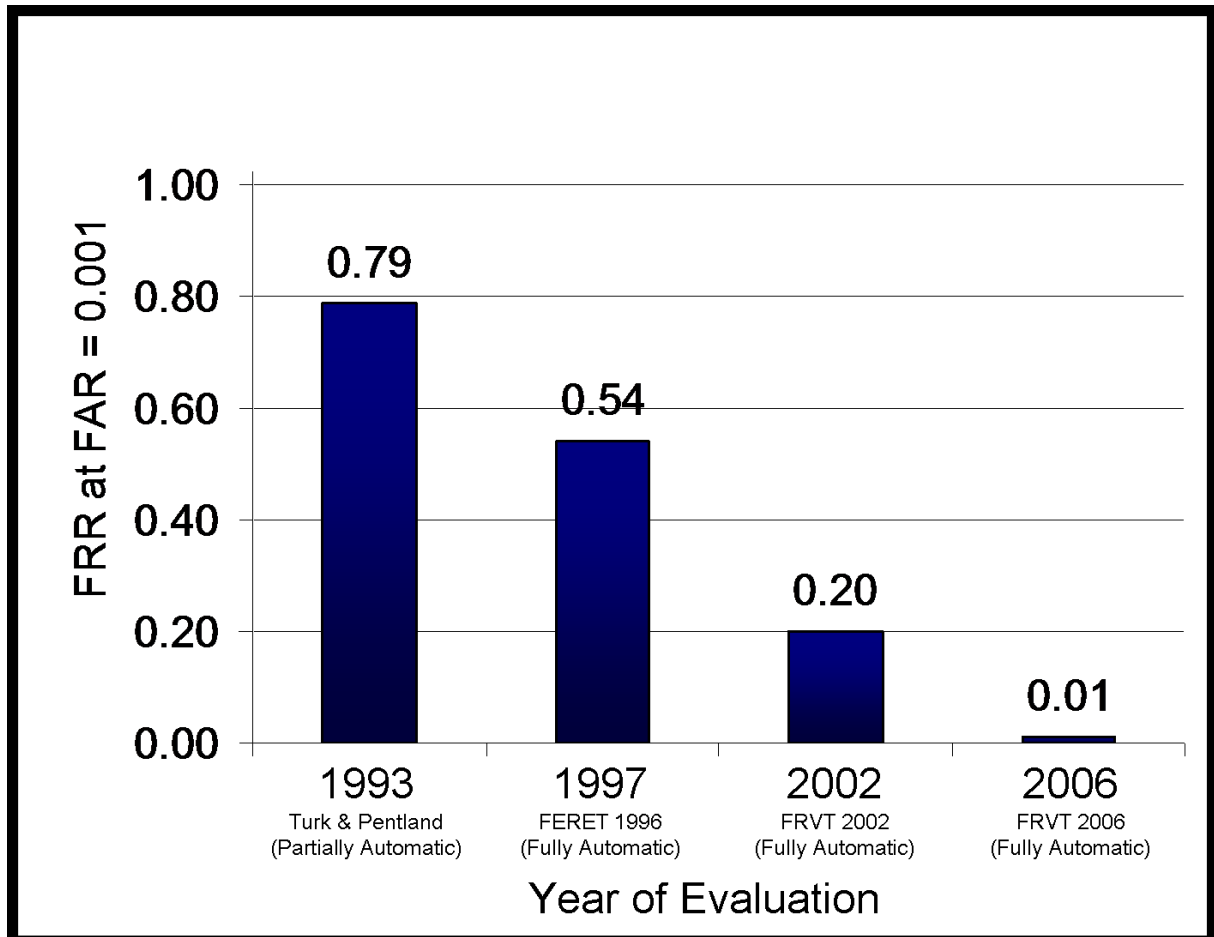


Figur 2.5: Eksempler på bilder som bidrar til vanskelige problemstillinger innen ansiktsgjenkjenning. Her er det vist bilder med lav pikseloppløsning, flere ansikter, skjegg og briller. Figuren er hentet fra [3].

Et problem, av et annet slag, som gjør det vanskelig å avgjøre hvilke metoder som er best, er den store variasjonen i hvordan en måler feilrater og lignende i metodene. Forskjellige bildedatabaser brukes på forskjellige problemstillinger, og alle databasene har styrker og svakheter. Bildedatabasene blir nærmere omtalt i kapittel 0. På slutten av 1990-tallet vokste FERET-testprotokollen fram, som et svar på dette problemet. Denne testprotokollen har blitt en standard for sammenligning av metoder innen ansiktsgjenkjenning, men bildene i FERET-databasen dekker ikke alle de problemstillingene ansiktsgjenkjenning står ovenfor. FERET-testprotokollen ble laget av P. Phillips, H. Moon, S. Rizvi og P. Rauss i [6], [7], [8] og [9]. Arbeidet begynte som et statsstøttet prosjekt, med hensikt å evaluere prototyper av ansiktsgjenkjenningssystemer, midt på 1990-tallet. Evalueringene hadde den hensikten at de skulle gi U.S.A.s regjering og byråer for lovshåndheving informasjon, som kunne være til hjelp med å avgjøre hvor og hvordan ansiktsgjenkjenning kan brukes.

Dette arbeidet har gått over til en rekke testprosjekter, som har blitt utført i 2000, 2002 og 2006, av "National Institute of Standards and Technology". Disse testprosjektene går under navnet "Face Recognition Vendor Tests" (FRVT), [10]. De har fortsatt arbeidet med å være opplysende for U.S.A.s statlige organer, samtidig som de forsøker å identifisere i hvilke retninger ansiktsgjenkjenning behøver mer forskning.

Selve testene som gjennomføres, gjøres upartisk, uten innflytelse fra personene som har utviklet algoritmene. I Figur 2.6 kan en se hvordan progresjonen for fagfeltet har vært på enkle problemer siden begynnelsen på 1990-tallet, fram til den siste FRVT, utført i 2006. En kan se at framgangen har vært veldig stor.

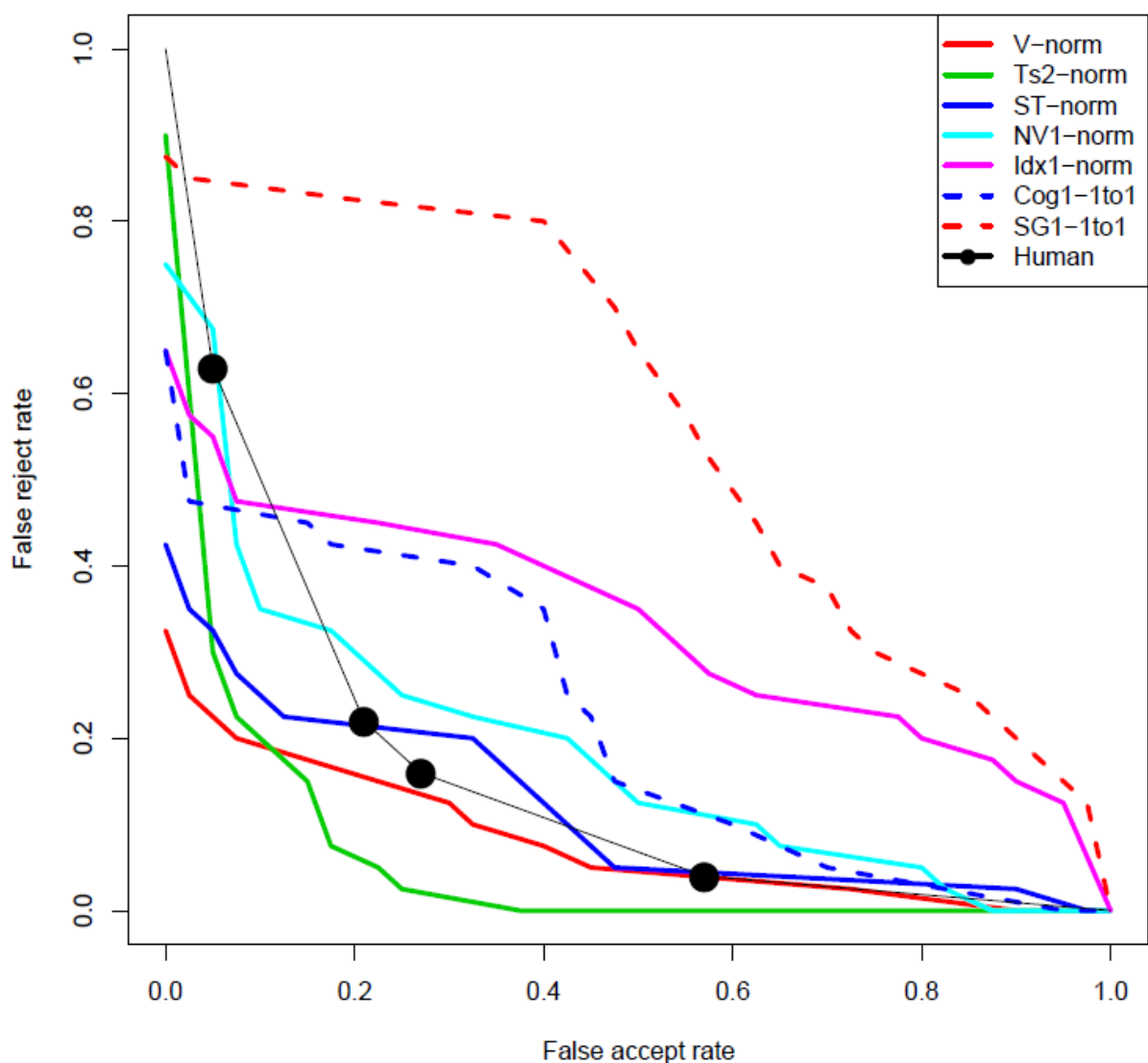


Figur 2.6: Reduksjonen i feilrate for de beste algoritmene, dokumentert av FERET, FRVT 2002 og FRVT 2006 evalueringene. Figuren er hentet fra [10]. Forkortelsene står for: "False Reject Rate" (FRR) og "False Accept Rate (FAR)".

Konklusjonen på FRVT fra 2006 sier at det er tre hovedkomponenter som har bidratt til forbedringer av resultatene, siden FRVT 2002. Disse er: A: selve gjenkjenningsteknologien, B: høyere oppløsning på bildene, og C: helhetlig forbedret kvalitet som følge av mer konsekvent belysning. Ut i fra forholdene, er det blitt estimert hvor stor forbedring i gjenkjenning som kommer som følge av gjenkjenningsteknologien. Denne ble estimert til å være mellom en faktor på 4 og 6, avhengig av hvilken algoritme det er snakk om.

I FRVT 2006 ble det også gjort tester hvor en sammenligner algoritmer med menneskers egenskaper til å kjenne igjen ukjente ansikter. Med ukjente ansikter er det her snakk om fremmede personer menneskene ikke kjenner. Altså ikke venner, familie og lignende. Resultatet av testene ble at mennesker har en feilrate på ca. 5% for å kjenne igjen fremmede mennesker. Dermed vil algoritmene som klarer å oppnå lavere feilrate enn 5% utklassere et menneske. Av algoritmene som ble testet, var det 6 av 7 algoritmer som var bedre eller like gode som mennesker.

Dessverre er ikke disse algoritmene publisert, og er kommersielt beskyttet av firmaene som har utviklet de. Disse firmaene er: Viisage, Tsinghua, SAIT, Neven Vision, Identix, Cognitec og Sagem. En sammenligning av disse algoritmene, mot mennesker, er gitt i Figur 2.7.



Figur 2.7: Figur 8 i [10] viser hvordan de forskjellige algoritmene presterer når det testes med varierende belysning. Perfekt gjenkjenning er hjørnet nederst til venstre, hvor F.R.R.=F.A.R.=0.

I FRVT 2006 er det ikke lagt fram noen forslag til nye forskningsområder, så her presenteres listen fra FRVT 2002:

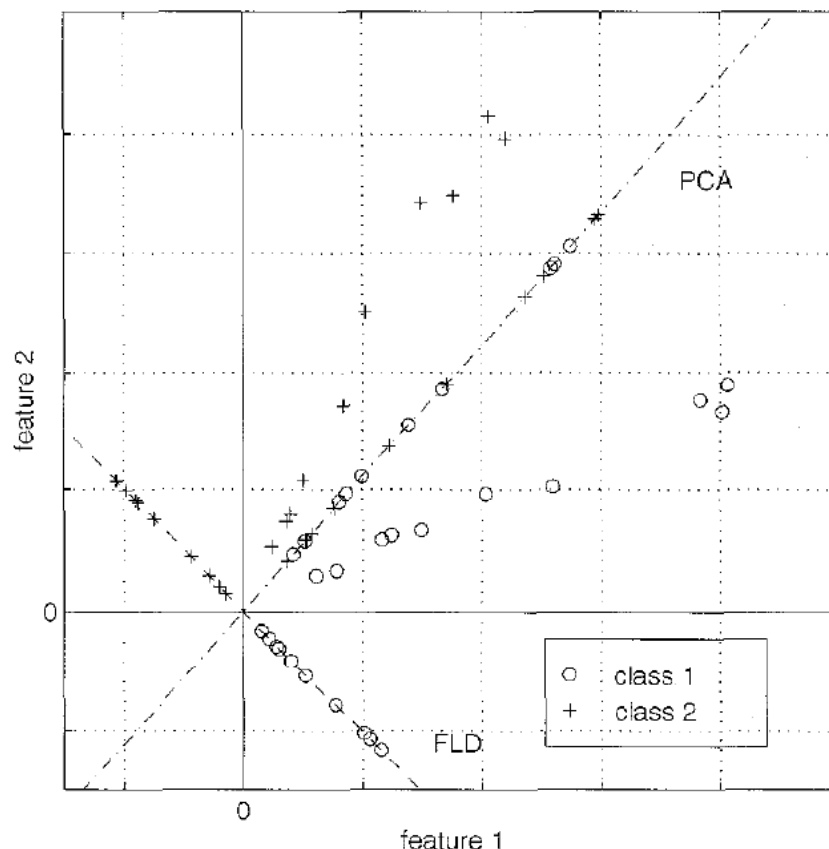
- Gjenkjenning fra utendørs ansiktsbilder.
- Gjenkjenning fra ikke-frontale ansiktsbilder.
- Gjenkjenning ved lave "False accept/alarm"-rater.
- Forstå hvorfor menn er enklere å gjenkjenne enn kvinner.
- Større forståelse for hvilke innspill demografiske faktorer har på utførelse.
- Utvikle bedre statistiske metoder for å forstå utførelse.
- Utvikle bedre modeller for prediksjon av identifisering i store bildedatabaser.
- Forstå effekten av algoritme- og systemtrening på kovariant utførelse.
- Integrasjon av "morphable models" med ansiktsgjenkjenningsutførelse.
- Forstå hvorfor videosekvensene i FRVT 2002 ikke forbedret prestasjoner.

Fram til 2006 ser en at framgangen innen ansiktsgjenkjenning har vært stor. Etter hvert som en nå begynner å se på problemer med frontale ansiktsbilder som mer eller mindre løst, har forskningen gått over til å sette fokuset mer mot vanskeligere problemstillinger. Framgangen her ser ut til å ha vært relativt liten. På grunn av dette regnes fortsatt [11], av W. Zhao, R. Chellappa, A. Rosenfeld og P.J. Phillips, som en god oversikt over fagfeltet ansiktsgjenkjenning.

Siden det ble bestemt at masteroppgaven først og fremst skulle dreie seg om gjenkjenning fra vanlige bilder, har temaene rundt tredimensjonale bilder, video og bilder tatt fra infrarødt lys blitt utelatt. For vanlige bilder begynte en først å kjenne igjen ansikter med metoder basert på geometrisk analyse av ansiktene. Denne formen for gjenkjenning gav dårlige resultater, til tross for at den ble svært mye brukt før, i følge [3].

Det var først da Sirovich og Kirby introduserte statistiske metoder, i [12], at ansiktsgjenkjenning begynte å se lovende ut. Metoden ble utvidet til å være brukbar til klassifisering, av Turk og Pentland i [13], og går nå under navnet P.C.A. Denne metoden danner basisen for mange av metodene som brukes innen ansiktsgjenkjenning i dag. Siden mye av arbeidet i denne oppgaven har vært basert rundt P.C.A., er denne beskrevet i detalj i kapittel 3.2, men for oversikten sin del, kan det nevnes at denne baserer seg på prosjektering av bildene, ned i et underrom med lavere dimensjon.

En annen metode som har fungert godt til ansiktsgjenkjenning er Fishers lineære diskriminant analyse (engelsk: Fishers Linear Discriminant Analysis), ofte forkortet til F.L.D. eller L.D.A.. Denne metoden ble introdusert i [14], av P. Belhumeur, J Hespanha og D. Kriegman, i 1997. F.L.D. metoden er en klassifiseringsmetode, som kan brukes sammen med P.C.A.. Den er basert på klassifiseringsprinsippet om å forsøke å maksimere forholdet mellom determinanten til spredningsmatrisen for spredning innad i klassene, og determinanten til spredningsmatrisen for spredning mellom klassene. For nærmere forklaringer rundt klassifisering, og spredningsmatriser anbefales læreboken [15] av Richard Duda, Peter Hart og David Stork. En klasse blir i dette klassifiseringstilfellet en person. Dette gjør at en får klassespesifikk projeksjon av bildene. For å illustrere fordelene med klassespesifikk projeksjon, kan en se på et toklasseproblem hvor projeksjonene fra hver klasse ligger i et lineært underrom. Resultatet av et slikt problem, for P.C.A. og F.L.D. kan se ut noe slikt som i Figur 2.8. Fra et klassifiseringsståsted ønsker en å finne de optimale grensene for å avgjøre hvilken klasse nye projeksjoner skal klassifiseres som. Fra Figur 2.8 ser en at F.L.D. separerer klassene så mye som mulig, mens P.C.A. smører de sammen. Selv om P.C.A. oppnår større spredning, sprer F.L.D. de projiserte verdiene på en slik måte at klassifiseringen forenkles.



Figur 2.8: En sammenligning av P.C.A. og F.L.D. for et toklasseproblem, hvor dataene ligger i et lineært underrom. En kan se at F.L.D. separerer klassene, mens P.C.A. faktisk smører de sammen. Figuren er hentet fra [14].

I [14] gjøres det en sammenligning av F.L.D. mot tre andre metoder: korrelasjonsbasert gjenkjenning, en variant av "lineært underrom"-metoden fra [16] og P.C.A.. Den korrelasjonsbaserte gjenkjenningen er basert på en "nærmeste nabo"-klassifiserer, en standard klassifiseringsalgoritme fra fagfeltet mønstergjenkjenning/klassifisering, se [15]. Sammenligningen konkluderer med at F.L.D. er en god del bedre enn de andre metodene. Nyere kilder, slik som [3], hevder derimot at ingen ansiktsgjenkjenningsalgoritme har noen klar fordel over de andre, i alle problemstillinger.

Den senere tiden er det P.C.A. og F.L.D. som har dominert ansiktsgjenkjenning, sammen med en del metoder som baserer seg på tredimensjonale bilder. Av andre metoder har mange klassifiseringsalgoritmer blitt foreslått. Nevrale nett, se [15], har blitt brukt til både ansiktsgjenkjenning, se [17], og en del problemstillinger som ligner på ansiktsgjenkjenning, slik som avgjøring av kjønn, se [18], og klassifisering av ansiktsuttrykk. "Hidden Markov model", H.M.M., en variant av nevrale nett har også blitt brukt, men ikke fått stor utbredelse. "Support Vector Machine", S.V.M., se [19], har blitt brukt under klassifiseringen av ansikter, men er heller ikke så mye i bruk lenger.

2.4 Forskning i andre fagfelt som er relevante for ansiktsgjenkjenning

Siden gjennombrudd i forskning innen fagfeltene ansiktsgjenkjenning, psykologi og nevrologi ofte påvirker hverandre, tas det her med en oversikt over gjennombrudd fra psykologi og nevrologi, som kan ha betydning for ansiktsgjenkjenning. Denne teksten er hentet fra [11], og gjengitt på norsk.

”Generelt bruker menneskenes gjenkjenningssystem et stort spektrum med stimuli, fra mange, om ikke alle, sansene. Disse brukes enten individuelt, eller sammen, for å lagre eller hente fram bilder av ansikter. I mange tilfeller brukes det også kontekstbasert kunnskap. Omgivelsene spiller for eksempel en viktig rolle når det kommer til å gjenkjenne ansikter, i forhold til hvor de ”hører hjemme”. Dessverre er det håpløst å forsøke å utvikle et system som etterligner alle ferdighetene til mennesker, med dagen teknologi, men et datasystem har andre fordeler over en hjerne. Mennesker har for eksempel en begrensning på hvor mange personer de kan huske. En viktig fordel med datasystemer er derfor muligheten til å håndtere store datasett med ansiktsbilder. Siden de fleste systemer består av intensitetsdata fra todimensjonale bilder, begrenses litteraturen her til det som omhandler menneskets visuelle oppfattning.

Mange studier innen psykologi og nevrologi har direkte relevans for ingeniører, som designer algoritmer eller systemer for gjenkjenning av ansikter. En fullstendig oversikt over relevante studier innen psykologi og nevrologi er utenfor omfanget til denne oppgaven. Her blir det kun presentert funn som potensielt kan være viktige for utviklingen av ansiktsgjenkjenningssystemer. For detaljer henvises lesere til de refererte kildene.

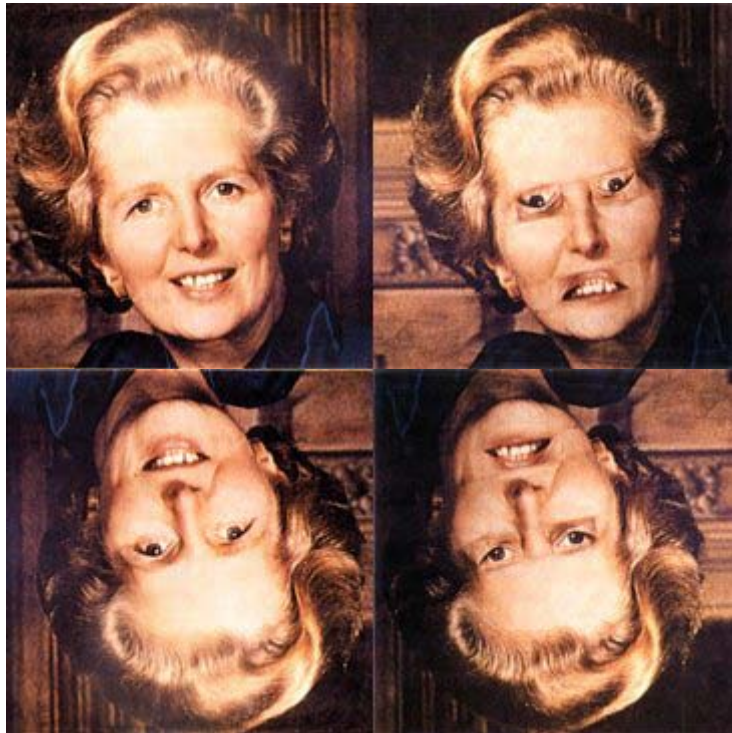
Er ansiktsgjenkjenning en dedikert prosess?

Funn som støtter teorien om at ansiktsgjenkjenning er en dedikert prosess kommer tre kilder, i følge [20]. A: Mennesker har lettere for å huske ansikter enn andre objekter, dersom de presenteres i normal orientering (med håret øverst, og haken nederst). B: Prospagnosiapasienter mangler egenskapen til å gjenkjenne personer fra ansiktet, men har vanligvis ingen andre agnosia. De er altså i stand til å kjenne igjen personer etter stemmer, hårfarge, klær og lignende. Det kan bemerkes at prospagnosiapasienter klarer å oppfatte om et objekt er et ansikt eller ikke. Så til tross for at de er i stand til å oppfatte øyne, nese osv., klarer de ikke å bruke denne informasjonen til identifikasjon. C: Det argumenteres for at nyfødte instinktivt tiltrekkes av ansikter. De ser ut til å foretrekke å følge med på ansiktslignende objekter, foran objekter som ikke ligner ansikter. Det strides om dette videre bekrefter at ansiktsgjenkjenning er en dedikert prosess, slik som i [21].

Er oppfattningen av ansikter basert på analyse av kjennetegn?

Både helhetlig informasjon, og informasjon hentet fra kjennetegn er viktige for oppfatning og gjenkjenning av ansikter, se [22]. Resultater fra studier foreslår at helhetlig informasjon brukes til grovere gjenkjenning, samtidig som informasjon hentet fra kjennetegn brukes til finere gjenkjenning. Dersom det finnes fremtredende kjennetegn, er det mulig helhetlig gjenkjenning ikke brukes i det hele tatt. For eksempel, tenderer mennesker til å

fokusere på fremtredende egenskaper som store øye, rar nese og lignende, i studier som fokuserer på å huske ansikter. Et sterkt bevis for denne teorien kan ses i [23], basert på Thatcher-illusjonen fra [24]. I denne illusjonen, inverteres munn og øyne i et ansikt. Dersom dette gjøres for et bilde i normal orientering, oppfattes resultatet som grotesk. Gjøres det derimot på et bilde som er snudd på hodet, oppfattes ansiktet som ganske normalt, og inverteringen er mye mindre merkbar, se Figur 2.9.



Figur 2.9 Tacher illusjonen ble oppkalt etter Storbritannias tidligere "Prime Minister", Margaret Thatcher, fordi dette bildet, av henne, ble brukt for å demonstrere den. Bildet kan finnes mange steder på Internett, men opphavet er uvisst. Det originale bildet er i gråtoner, og viser bare to av bildene ovenfor. Illusjonen viser hvordan en ved å rotere øyne og munn, får et grotesk bilde (øverst til høyre). Dersom også ansiktet roteres, oppfattes bildet som ganske normalt (nede til høyre).

Rangering av viktighet for ansiktskjennetegn

Hår, ansiktets fasong, øyne og munn (ikke nødvendigvis i denne rekkefølgen) har blitt ansett som å være viktige for oppfatning og gjenkjenning av ansikter. Flere studier har vist at nesen spiller en veldig liten rolle i denne sammenhengen. Det har også kommet fram at den øvre delen av ansiktet er viktigere for gjenkjenning, enn den nedre delen. Attributter som vakkerhet og tiltrekning spiller også inn på gjenkjenningsraten. Vakre personer har best gjenkjenningsrate, deretter de minst attraktive personene. De gjennomsnittlig vakre personene har dårligst gjenkjenningsrate.

Karikaturer

Perkins definerer i [25] formelt en karikatur som framheving av et eller annet mål som varierer fra person til person. En nese varierer fra person til person, og kan derfor være brukbar som et symbol når en lager en karikatur av en person. Antall ører derimot, er ikke nyttig informasjon i en karikatur. Karikaturer inneholder ikke like informasjon som et fotografi, men de klarer allikevel å få fram nok kjennetegn til at personen kan gjenkjennes.

Særpreg i ansikter

Studier viser at ansikter med særpreg holder seg bedre i hukommelsen, enn vanlige ansikter. Dersom en skal ta en avgjørelse på om et objekt er et ansikt eller ikke, tar det lengre tid å avgjøre at et utypisk ansikt er et ansikt, enn for vanlige ansikter.

Rollen romlig frekvensanalyse har i ansiktsgjenkjenning

Tidlige studier, som [26] og [27], konkluderte med at informasjon i de lave romlige frekvensbåndene har en dominant rolle i ansiktsgjenkjenning. Senere studier, som [28], viste at de romlige frekvensbåndene har ulike roller i de forskjellige gjenkjenningsoppgavene. For eksempel spiller de lave frekvensbåndene en stor rolle i å avgjøre hvilket kjønn en person har. Identifisering av personen krever derimot de høyfrekvente frekvensene. De lavfrekvente komponentene bidrar altså til den globale beskrivelsen av et ansikt, mens de høyfrekvente komponentene bidrar til de finere detaljene som kreves for identifisering.

Er gjenkjenning uavhengig av perspektiv?

Mye arbeid innefor objektgjenkjenning, slik som [29] og [30], har blitt gjort under hypotesen framsatt av Marr, i [31], om at forskjellige perspektiv på objekter kan analyseres slik at de kan uttrykkes mer eller mindre perspektivuavhengig. I senere tid har det blitt drøftet om objektgjenkjenning virkelig er perspektivuavhengig. For ansiktsgjenkjenning ser det ut som at hukommelsen er veldig perspektivavhengig. Hill, med flere, viser i [32] at generalisering fra et perspektiv til et annet generelt er dårlig.

Effekt av lysforandring

Det har lenge vært uformelt kjent at det er vanskelig å kjenne igjen personer fra fotografiske negativer. I [33] har det blitt utført en del eksperimenter for å utforske om vanskelighetene knyttet til gjenkjenning med negativer, og bilder snudd på hodet, kommer fra at en reverserer den tilsynelatende retningen på belysningen. Denne studien demonstrerte at bunnbelyste bilder faktisk gjør det vanskeligere å kjenne igjen personer. I [34] demonstreres det hvor viktig toppbelyste ansikter er for identifisering.

Bevegelse og ansiktsgjenkjenning

En studie utført i [35] viser at kjente ansikter er lettere å gjenkjenne, dersom de vises i bevegelse i en sekvens, i stedet for i enkeltbilder. Denne observasjonen har, i [22], blitt utvidet til å vise at bevegelse gjør gjenkjenning enklere, under mange ulike former for forvring av forholdene, slik som for fotografiske negativer, vending av bildene eller terskelsetting (sort hvitt framstilling av bildene). Studiet viser at det også ser ut til at fordelene kommer uavhengig av om informasjonen det sammenlignes med kommer som statiske bilder eller dynamiske sekvenser.

Ansiktsuttrykk

Basert på nevrologiske studier, slik som [36], ser det ut som analyse av ansiktsuttrykk foregår i parallell med gjenkjenning. En del prosopagnotiske pasienter, som har vansker for å gjenkjenne personer fra ansikter, er i stand til å kjenne igjen ansiktsuttrykk. Pasienter med "organic brain syndrome" er dårlige på å kjenne igjen ansiktsuttrykk, men er flinke på ansiktsgjenkjenning. For vanlige mennesker er det også mulig å se parallelle egenskaper for analyse av ansiktsuttrykk, og ansiktsgjenkjenning. Det påstås at mennesker er i stand til å separere ansiktsgjenkjenning og fokusert synsprosessering. For eksempel er det mulig å lete etter personer med en tykk bart, uten å forsøke å identifisere personene en ser på."

Påvirkning på fagfeltet

En kan finne mange eksempler på hvordan oppdagelser i psykologi og nevrologi har påvirket ansiktsgjenkjenning. I [12] er det for eksempel tatt utgangspunkt i at ansiktsgjenkjenning ikke er avhengig av den nedre delen av ansiktet. Bildene er beskåret, slik at kun øyne og nese tatt med. Her så resultatene lovende ut, for P.C.A.. Det ble observert at de beskårede bildene hadde lavere feilrate, når færre egenbilder brukes. Senere er det blitt antatt at bilder med mindre variasjon, har behov for færre egenbilder for å uttrykkes med lavere feilrate. Siden de beskårede bildene har mindre variasjon, kan en ikke konkludere noe fra disse resultatene. Sirovich hevder senere at ved å bruke hele ansikter, oppserverer en lavere feilrater.

Et annet eksempel kan være arbeidet fra [37]. Denne artikkelen tar utgangspunkt i hypotesen, som F. Krouse setter fram i [38]. Denne sier at mennesker har lavere feilrate for gjenkjenning av ansikter, dersom de ses 3/4 vridd fra fronten på ansiktet. I [37] kommer Valentin, Abdi og Edelman fram til at selv om at dette perspektivet gir lavere feilrate, behøver en ikke å lagre dette bildet for å oppnå effektene av bedre feilrate. En vinner like mye på å inkludere frontal og side perspektivbilder.

2.5 Bildedatabaser

”På grunn av dets varierende og komplekse tredimensjonale struktur, blir et utseende til et ansikt påvirket av et stort antall faktorer. Disse omfatter blant annet identitet, ansiktsuttrykk, belysning, alder og hår. Utvikling av algoritmer til ansiktsgjenkjenning, som er robuste ovenfor variasjoner i disse faktorene, krever bildedatabaser med stor nok størrelse, og nøyaktig kontroll av faktorene. Det trengs også standard bildedatabaser slik at det er mulig å sammenligne algoritmer. Arbeidet med å utvikle en bildedatabase, av høy kvalitet, er et ressurskrevende arbeid, men tilgjengelighet av bildedatabaser er viktig, dersom fagfeltet skal kunne ha progresjon.”

- Ralph Gross, [39], gjengitt på norsk.

I Figur 2.10 er det gitt en oversikt over noen av de viktigste bildedatabasene, som er tilgjengelig for alle som arbeider med ansiktsgjenkjenning. I dette kapitlet blir i hovedsak de databasene som ble vurdert til eksperimentene i Kapittel 4 presentert. Utover dette, tas det også med noen databaser for å vise bredden i hva som er tilgjengelig. Dersom det ønskes mer informasjon om databaser, henvises lesere til [39] eller [40], som begge inneholder store oversikter over bildedatabaser.

Database	No. of subjects	Pose	Illumination	Facial Expressions	Time
AR	116	1	4	4	2
BANCA	208	1	++	1	12
CAS-PEAL	66 – 1040	21	9 – 15	6	2
CMU Hyper	54	1	4	1	1 – 5
CMU PIE	68	13	43	3	1
Equinox IR	91	1	3	3	1
FERET	1199	9 – 20	2	2	2
Harvard RL	10	1	77 – 84	1	1
KFDB	1000	7	16	5	1
MIT	15	3	3	1	1
MPI	200	3	3	1	1
ND HID	300+	1	3	2	10/13
NIST MID	1573	2	1	++	1
ORL	10	1	++	++	++
UMIST	20	++	1	++	1
U. Texas	284	++	1	++	1
U. Oulu	125	1	16	1	1
XM2VTS	295	++	1	++	4
Yale	15	1	3	6	1
Yale B	10	9	64	1	1

Figur 2.10: En oversikt over alle databasene, som blir omtalt i [39]. Figuren er hentet fra samme artikkel. I tilfeller hvor det ikke er målt antall forekomster, eller dette ikke er tatt hensyn til under utviklingen av databasen, er disse merket med ”++”.

Da bildedatabaser skulle vurderes til eksperimentene i Kapittel 4, ble det satt opp en del kriterier for hvilke egenskaper databasen burde ha. Da det ble valgt å se vekk fra ansiktsdeteksjon, var det mest interessant med en database med bilder, som allerede var normalisert. En slik database var det ikke mulig å oppspore, så primæregenskapen en så etter ble derfor at databasen skulle inneholde omtrent like mange individer som det som ble brukt av Sirovich og Meytils, i [1]. Det ble også bestemt å forsøke å inkludere kun kaukasoidske (hvite) menn, da dette ble gjort i samme artikkel.

The AR Face Database

“The AR Face Database” ble utarbeidet ved Computer Vision Center i Barcelona i 1998, [41]. Databasen inneholder bilder av 116 individer, hvorav 63 er menn, og 53 er kvinner. Personene ble fotografert to ganger, med intervaller på to uker. Faktorer som belysning og kamerainstillinger ble nøye kontrollert under fotografering, slik at forholdene skulle være like for alle bildene.

Denne databasen er den Sirovich og Meytils bruker i [1]. Det framgår ikke hvor alle de 94 bildene i deres bildedatabase kommer fra, da AR-databasen kun inneholder 63 menn.

The Facial Recognition Technology (FERET) Database

FERET-databasen ble laget som en del av FERET-testene, som ble sponset av U.S.A., se [6]. Feilrater og kjøretider for nye ansiktsgjenkjenningsskemaer, utført med denne databasen, blir ofte brukt som et standardmål for sammenligning mot andre algoritmer. Databasen inneholder 1199 individer, men i distribusjonen er ikke disse sortert etter kjønn, rase eller lignende. På grunn av at normalisering av databasen måtte skje for hånd, ble det for tidkrevende å bruke denne bildedatabasen.

Face Recognition Data, University of Essex

Denne bildedatabasen ble satt sammen av Dr. Libor Spacek, og består av tidligere studenter fra Universitetet i Essex. De fleste individene er derfor i aldersgruppen 18-20 år. Totalt er det 395 individer i databasen, fordelt i tre kategorier: menn, kvinner og ansatte, for forskjellige vanskelighetsgrader. Noen av individene har briller og/eller bart/skjegg. Databasen er tilgjengelig fra [4].

Databasene Yale og Yale B

Databasene fra universitetet i Yale, se [42] og [43], består av relativt få individer, i mange forskjellige situasjoner. Siden sidebelysning var spesielt interessant for problemstillingene i eksperimentdelen, ble også denne bildedatabasen brukt. Den første databasen består av 15 individer, og den andre, “Yale B”, består av 10 individer. I den første databasen er fokuset lagt mot forskjellige ansiktstrykk, briller, skjegg og lignende. “Yale B” er mer rettet mot forskjellig belysning, og forskjellige vinkler mellom kamera, og fronten på ansiktene.

Andre databaser

Av andre databaser har en de som tar for seg bestemte etniske folkeslag, slik som "CAS-PEAL Face Database", se [44], som tar for seg kinesiske personer. SCface-databasen inneholder bilder tatt i ukontrollerte omgivelser innendørs, fra 5 forskjellige overvåkningskameraer. Denne databasen er rettet mot vanskelige, mer praktiske, problemstillinger i ansiktsgjenkjenning, se [45].

Kapittel 3 - Utplukking av ansiktets egenskaper

3.1 "Face space"

I forrige kapittel ble det presentert hvordan ansiktsgjenkjenning ofte blir foretatt i praksis. Det ble nevnt at en ofte deler gjenkjenningsprosessen inn i tre deler: ansiktsdeteksjon, utplukking av den informasjonen som er viktig for gjenkjenning og ansiktsidentifisering (engelsk: "face detection", "feature extraction" og "face identification"). Dette kapitlet tar for seg den andre delen i gjenkjenningsprosessen, nemlig utplukkingen av den informasjonen som er viktig for gjenkjenning.

Så hvorfor ønsker en egentlig å begrense hvilken informasjon en skal se på i ansiktsgjenkjenningssammenheng? Som kjent fra for eksempel [46] av Rafael Gonzalez og Richard Woods, blir et fotografi typisk digitalisert, som et sett med pikselintensitet i et todimensjonalt gitter. Noen velger også i en mer matematisk sammenheng å se på dette som en matrise bestående av lysintensitet for hver piksel. Siden en ofte ønsker å ha tett romlig sampling av fargeintensitet i scenen en tar bilde av, inneholder bilder ofte svært mye informasjon. Dersom en ønsker å oppbevare store mengder identiteter å gjenkjenne ansiktene mot, kan en fort få problemer med lagringsplass, og treg algoritmeutførelse dersom en bruker all bildeinformasjonen. En av de store utfordringene innen ansiktsgjenkjenning er derfor knyttet til representasjon av ansiktene som inngår i systemet.

Lawrence Sirovich har sammen med andre forsket mye på å finne et estimat for hvor mye informasjon som behøves for å representere et menneskeansikt. Et øvre tak for denne informasjonen blir satt til å være pikseldimensjonen, for bilder hvor det fortsatt er mulig for et menneske å kjenne igjen et ansikt. I artikkelen [2] antas det at et menneske klarer å kjenne igjen identiteten til et ansikt som fyller et bilde på 128x128 piksler, noe som er i nærheten av $O(10^4)$ dersom en ønsker å benytte "Big O"-notasjon. Som et dimensjonsestimat for settet med alle mulige ansikter, er dette veldig høyt, og det antas at den faktiske dimensjonen er en mye lavere enn dette. I løpet av de forskningsartiklene Sirovich, har utgitt i samarbeid med andre, slik som [12], [2], [47], [48] og [1], kommer estimatet til slutt ned mot at en kan uttrykke alle menneskeansikter med litt mindre enn 70 dimensjoner. Dette påstås å være i samsvar med psykologiske studier, som [49] og [50]. Dimensjonene en snakker om her er altså det vektorrommet som uttrykkes av så mange basisvektorer som det er dimensjoner. Når det er snakk om ansikter, kaller en gjerne dette vektorrommet for ansiktsrommet (engelsk: face space). Dersom en fortsetter denne tankemåten, blir det etter hvert en naturlig problemstilling å forsøke å finne andre basiser som er bedre egnet til å uttrykke ansiktsrommet, enn lysintensiteten i et bilde.

Denne problemstillingen har vært opphav til mye forskning på fagfeltet ansiktsgjenkjenning, og har ført fram til flere metoder som blir mye brukt i dag. Av disse to, tar vi, som nevnt tidligere, for oss P.C.A.. Selve uttrykket "face space" har sitt opphav i artikkelen [13], av M. Turk og A. Pentland.

3.2 Prinsipalkomponentanalyse, P.C.A.

Som kapittel 3.1 fortalte, er en av de store utfordringene innen ansiktsgjenkjenning knyttet til representasjon av ansiktene som inngår i systemet. En ønsker altså med bakgrunn i argumentene derfra å forsøke å representere hver enkelt person med kun den informasjonen, som er nødvendig for gjenkjenning. Det er denne utfordringen som har dannet grunnlaget for prinsipalkomponentrepresentasjon av ansikter.

Dersom en ser for seg at et bilde av et ansikt består av en rekke egenskaper, og hvor hver av disse egenskapene uttrykker en dimensjon i ansiktsrommet, er det naturlig å anta at en del av disse egenskapene er irrelevant for gjenkjenning. For eksempel vet vi at belysning påvirker et bilde, men at denne belysningen ikke påvirker identiteten til personen som skal gjenkjennes. En vet også at unøyaktigheter og støy fra kameralinsen også påvirker bildet, men heller ikke disse påvirker identiteten til personen som skal gjenkjennes. Slik er det mulig å liste opp egenskaper som er uinteressante for selve gjenkjenningen. Ideelt sett ønsker en derfor å representere ansiktene, i databasen en gjenkjenner mot, kun med den informasjonen, altså den dimensjonen som er nødvendig for korrekt gjenkjenning. Dessverre vet vi fortsatt lite om hvordan mennesker gjenkjenner hverandre, hvilke egenskaper som vektlegges, eller hvor stor dimensjon ansiktsrommet har. I 1987 foreslo Kirby og Sirovich, i [12], å bruke "Karhunen-Loève"-transformen til å finne en ny basis for bilder. Denne basisen består av et sett med egenvektorer fra kovariansmatrisen til bildesettet, vektet med sine respektive egenverdier. I statistikkitteratur kaller en ofte denne basistransformasjonen for prinsipalkomponentanalyse. Videre ble det satt fram en hypotese om at en kunne avgjøre identitet ved hjelp av euklidsk avstand mellom projeksjoner av ansikter med kjent identitet, og projeksjonene av nye ansikter. Denne hypotesen ble framsatt av Turk og Pentland, i [13]. Det strides om hvorvidt dette kan regnes som den første skikkelige ansiktsgjenkjenningsmetoden, men det er i alle fall sikkert at den var mye bedre enn andre metoder som eksisterte på slutten av 1980-tallet.

3.3 Teorien bak P.C.A.

I dette avsnittet kommer masteroppgaven til å følge den notasjonen Sirovich, Kirby og Meytils benytter i [12], [2] og [1]. I [12] viser Kirby og Sirovich hvordan en kan foreta P.C.A. på et sett med bilder. Artikkelen [1] viser ikke en gjennomgang av metoden, men refererer til [2] og [12]. Forskjellen fra [1] til [12] er at [1] gjør bruk av ansiktssymmetrien, som blir introdusert i [2]. For enkelhets skyld, vises først framgangsmåten fra [12].

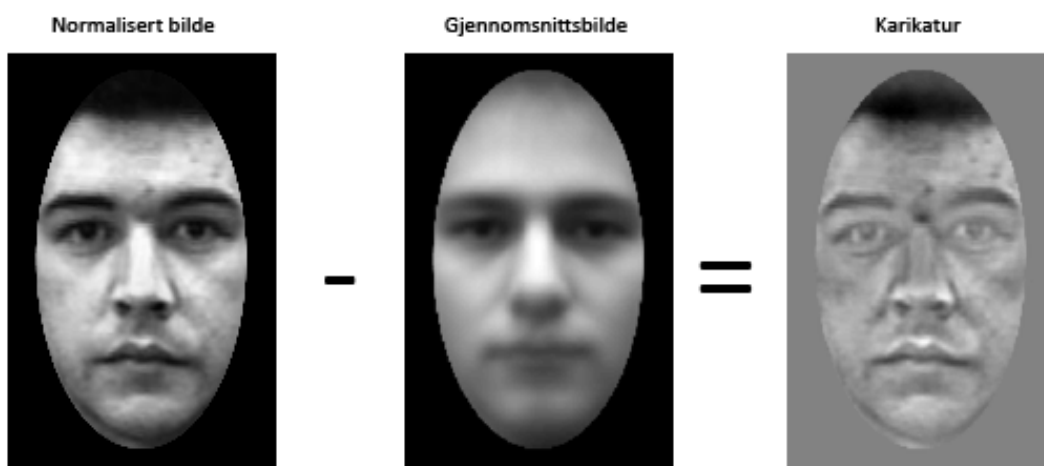
Utgangspunktet for metoden er at en har et sett med ansikter, som har blitt fotografert som et sett med bilder $\{\tilde{\varphi}_n(\mathbf{x})\}$, hvor \mathbf{x} er en skalar funksjon av posisjon, $\mathbf{x} = (x,y)$, og n indikerer hvilket bilde det er snakk om. Hvert bilde $\tilde{\varphi}_n(\mathbf{x})$ representerer da en matrise bestående av gråverdiintensitetene til hver piksel i bildet. I teksten i dette kapittelet er det underforstått at når det er snakk om bilder, menes det stort sett bilder som kun viser et ansikt, slik som de normaliserte bildene i Figur 4.2. Disse bildene ønsker en å representere på en mer effektiv måte, enn ved bruk av lysintensitet forbundet med hver piksel. P.C.A. er den lineære transformen som gir en ny basis med størst mulig varians. Altså, den basisen som gir mest spredning mellom de projiserte bildene.

Før en starter med selve prinsipalkomponentanalysen, kan det altså være hensiktsmessig å utføre steg som ansiktsdeteksjon, normalisering av bildene, fjerning av bakgrunn og lignende steg fra det som omtales som del 1 av ansiktsgjenkjenning i kapittel 2.2, se Figur 2.2. I framgangsmåten fra [2], ville en her også introdusert noen av stegene som inngår i bruk av ansiktssymmetri, og disse vil det bli sett nærmere på i kapittel 0.

For at tilnærmingen til den nye basisen skal minimere minste kvadraters feil, må en trekke gjennomsnittsbildet fra hvert enkelt bilde. Gjennomsnittsbilde finner en slik som i likning 3.1, hvor M er her det totale antallet bilder i settet.

$$\bar{\varphi} = \sum_{n=0}^M \tilde{\varphi}_n(\mathbf{x}) \quad (3.1)$$

Etter gjennomsnittsbildet er beregnet, kan altså dette trekkes fra hvert bilde i settet. Slik det antydes i likning 3.2. I mange tilfeller kaller en bildene, med gjennomsnittsbildet subtrahert, for karikaturer (engelsk: caricatures), da de kan minne om karikaturtegninger. Et eksempel for hvordan karikaturene blir seende ut, kan ses i Figur 3.1.



Figur 3.1: Når en trekker gjennomsnittsbildet fra et bilde, kalles resultatet en karikatur. Figuren er laget med bilder fra "Face Recognition Data, University of Essex", [4], og resultater fra MatLab implementeringen av P.C.A..

$$\{\tilde{\Phi}_n(\mathbf{x})\} = \{\tilde{\varphi}_n(\mathbf{x}) - \bar{\varphi}\} \quad (3.2)$$

Det neste steget mot å finne en ny basis som bildene skal uttrykkes med, er å gjøre alle bildene om til kolonnevektorer, og sette disse som kolonnene i en matrise, \tilde{A} , som i likning 3.3.

$$\tilde{A} = [\vec{\Phi}_1 \quad \vec{\Phi}_2 \quad \dots \quad \vec{\Phi}_n] \quad (3.3)$$

De nye basisvektorene består nå i følge Karhunen-Loève transformen av egenvektorene til kovariansmatrisen til dette settet. Kovariansmatrisen beregnes med likning 3.4.

$$\tilde{C} = \tilde{A} * \tilde{A}^t \quad (3.4)$$

Fra lineær algebra, vet en at egenvektorene, og egenverdiene til kovariansmatrisen, \tilde{C} , kan finnes fra ligningen, vist i likning 3.5. Se for eksempel definisjonen på egenverdier og egenvektorer fra [51]. Egenvektorene har her blitt gitt navn \vec{u}_1 til \vec{u}_n .

$$\tilde{C} * \vec{u}_n = \lambda_n * \vec{u}_n \quad (3.5)$$

Da størrelsen på matrisen \tilde{C} ofte blir veldig stor, kan det være hensiktsmessig å bruke et triks fra [13]. En vet at en matrise med størrelse $P \times M$ har det minste av $P-1$ og $M-1$ egenverdier, som er ulik null. Siden antallet bilder som brukes, M , som oftest er mindre enn antall piksel, P , i bildene, er $P-1$ mindre enn $M-1$. På grunn av dette er det hensiktsmessig å søke å finne egenvektorene til en tilsvarende matrise, med størrelse $P \times P$, $\tilde{A}^t * \tilde{A}$, slik som i likning 3.6.

$$\tilde{A}^t * \tilde{A} * \vec{u}_n = \lambda_n * \vec{u}_n \quad (3.6)$$

Dersom en multipliserer begge sidene av denne ligningen med \tilde{A}^t , ser en fra likning 3.7 at egenvektorene $\tilde{A} * \vec{u}_n$ vil være egenvektorene for matrisen \tilde{C} , og egenverdiene vil være de samme.

$$\tilde{A} * \tilde{A}^t * \tilde{A} * \vec{u}_n = \lambda_n * \tilde{A} * \vec{u}_n \quad (3.7)$$

Fra teorem 9, kapittel 7, i [51], vet en at dersom en har en orthonormal basis bestående av egenvektorene til matrisen $\tilde{A}^t * \tilde{A}$, sortert etter synkende egenverdi, $\lambda_1 \geq \dots \geq \lambda_n$, så er settet med egenvektorer $\{\tilde{A} * \vec{u}_n\}$ en orthogonal basis for kolonnene i \tilde{A} . Fra likning 3.3 vet en at kolonnene i \tilde{A} er bildene av ansiktene. Derfor er egenvektorene $\{\tilde{A} * \vec{u}_n\}$ for substituttmatrisen $\tilde{A}^t * \tilde{A}$ den nye basisen fra Karhunen-Loève transformen. Siden egenverdiene kommer i synkende rekkefølge etter størrelse, vil vektorrommet, utspendt av den nye basisen, ha mest informasjon i retning av basisvektorene med lavest indeks. På en litt mer uformell måte kan en si at den første egenvektoren er den viktigste, den andre inneholder litt mindre informasjon, og de siste egenvektorene inneholder mer eller mindre bare støy dersom bildesettet er stort nok. På grunn av hvordan disse egenvektorene blir seende ut, når de omformes fra kolonnevektorer til bilder med samme størrelse som de originale bildene, kalles de ofte for egenbilder (engelsk: eigenpictures), egenansikter (engelsk: eigenfaces) eller spøkelsesbilder (engelsk: ghost images), se Figur 3.2. I denne figuren ser en hvordan de egenverdiene med lavest indeks begynner å minne om støy. Her er egenbildene vist fra 1 til 76, fra Eksperiment 1, kapittel 4.2.



Figur 3.2: En kan representere et bildesett med nye basisvektorer. Et slikt sett med basisvektorer kan bli seende ut slik som bildene over. Med P.C.A. blir bildene seende mer og mer ut som støy, jo høyere indeks av basisvektorene en ser på. Figuren er laget med bilder fra "Face Recognition Data, University of Essex", [4], og resultater fra MatLab implementeringen av P.C.A..

Nå som en har funnet en ny basis for ansiktene representert av bildesettet, $\{\tilde{\varphi}_n(\mathbf{x})\}$, er det neste steget i P.C.A. å projisere alle bildene over i denne basisen. Siden en nå kun har like mange basisvektorer som en har bilder i settet, ser en hovedfordelen av denne nye representasjonen. I stedet for at hvert bilde må representeres med en verdi per piksel i bildet, behøver en bare å representere hvert bilde med en verdi for hver ny basisvektor, se Figur 3.3. Etter en har projisert bildene av kjente og ukjente personer inn den nye basisen, blir det et klassifiseringsproblem å avgjøre hvilken identitet et ukjent bilde hører til. En tar da steget inn i tredje del av gjenkjenningsfase, nemlig den delen som omhandler den faktiske gjenkjenningen.



Figur 3.3: Ved å foreta en projeksjon ned i den nye basisen, kan bilder representeres som et sett med vektinger for hver basisvektor. Et bilde kan ses på som å bestå av "litt av egenbilde 1, litt mindre av egenbilde 2 osv". Figuren er laget med bilder fra "Face Recognition Data, University of Essex", [4], og resultater fra MatLab implementeringen av P.C.A..

3.4 Ansiktssymmetri

I artikkelen, [2], av Kirby og Sirovich, introduseres det en videreutviklet variant av P.C.A., som også tar høyde for at menneskeansikter har høy grad av symmetri om en tenkt symmetriakse, som går vertikalt gjennom midten av ansiktet. Denne utvidede varianten av P.C.A. blir også senere brukt i artikkelen, [1], av Sirovich og Meytils. I eksperimentdelen skal vi forsøke å gjenskape resultatene fra denne artikkelen.

Denne utvidelsen baserer seg på symmetribetraktningen om den tenkte symmetriaksen, som går vertikalt gjennom midten av ansikter, se **fig**. Ideen med å utnytte den naturlige symmetrien i signaler har vært brukt innenfor både talegjenkjenning, [52], og turbulente strømnings [53], og har vist seg å forbedre metoder hvor det er større grad av symmetri. Det er derfor naturlig for Kirby og Sirovich å forsøke å utnytte symmetrien i ansikter til å forbedre P.C.A., i artikkelen [2].

Selve implementeringen av denne ekstra symmetribetraktningen kan gjøres på flere forskjellige måter. Ideen er at en utvider det originale bildesettet, av ansikter, med de samme bildene speilet om symmetriaksen. Dersom en gjør dette, vil egenvektorene i den nye basisen fordele seg som like og odde. For å spare lagringsplass kan en faktisk oppnå dette uten å doble mengden lagringsplass bildesettet, $\{\tilde{\varphi}_n(\mathbf{x})\}$, behøver. Dersom en fortsatt lar $\mathbf{x} = (x,y)$, kan en se på det nye settet som $\{\tilde{\varphi}_n(x,y)\} \cup \{\tilde{\varphi}_n(-x,y)\}$. I [2] vises det hvordan en matematisk kommer fram til at kovariansmatrisen \tilde{C} , for det utvidede bildesettet kan ses på som en sum av en like og en odde kovariansmatrise, \tilde{C}_{like} og \tilde{C}_{odde} . Det vises også at basisene for disse to kovariansmatrisene sammen utspenner det samme vektorrommet som den originale kovariansmatrisen, \tilde{C} , for det utvidede bildesettet. En definerer videre to nye matriser $\tilde{\alpha}$ og $\tilde{\beta}$, med kolonnevektorer slik som i likning 3.8 og likning 3.9.

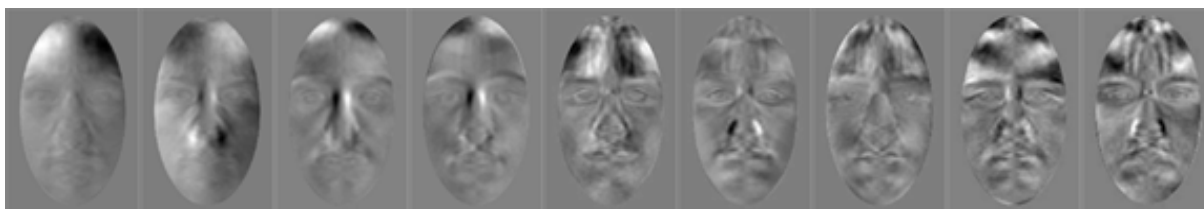
$$\vec{\alpha}_n = \vec{\Phi}_n + \vec{\Phi}_{n, \text{speilet}} \quad (3.8)$$

$$\vec{\beta}_n = \vec{\Phi}_n - \vec{\Phi}_{n, \text{speilet}} \quad (3.9)$$

En kan da dele P.C.A. algoritmen opp i to dekkoblede problemer. Dersom en substituerer matrisene $\tilde{\alpha}$ og $\tilde{\beta}$ med \tilde{A} , i likning 3.7, og finner to sett med egenvektorer og egenverdier. Disse egenvektorene fordeler seg som like og odde, slik som det er vist i Figur 3.4 og Figur 3.5. En kjenner igjen like egenvektorer med at de er symmetriske om symmetriaksen, og odde egenvektorer med at de er inverse om symmetriaksen. Ønsker en å holde kontroll på hvilke egenvektorer som er like og odde, bør en velge å implementere symmetribetraktningen som disse to dekkoblede problemene.



Figur 3.4: De like egenbildene skiller seg fra vanlige egenbilder ved at de er speilede om symmetriaksen. Figuren er laget med bilder fra "Face Recognition Data, University of Essex", [4], og resultater fra MatLab-implementeringen av P.C.A..



Figur 3.5 De odde egenbildene skiller seg fra vanlige egenbilder ved at inverse om symmetriaksen. Figuren er laget med bilder fra "Face Recognition Data, University of Essex", [4], og resultater fra MatLab-implementeringen av P.C.A..

Nå som en har fått egenansiktene fordelt som like og odde, kan en gjøre en del antagelser om hvilke av disse egenansiktene som faktisk er viktige eller i det hele tatt nødvendige for gjenkjenningen. Hypotesen for denne problemstillingen, er det som danner grunnlaget for artikkelen [1].

Sirovich og Meytils antar i **1** at de odde egenbildene representerer støy i for av sidebelysning, og at disse er unødvendige for gjenkjenning. Sidebelysning kan derfor fjernes uten at det går på bekostning av gjenkjenning.

Som en tilleggsantagelse antas det, i **1**, at pannen i stor grad fungerer som et speil når en bruker blits. Dette påvirker i stor grad, sammen med hår, egenbildene med indeks 3 og 5. Disse tilsvarer egenbilder nummer 2 og 4 i likebasisen. Dersom en fjerner disse, er håpet at en forbedrer gjenkjenningsratene.

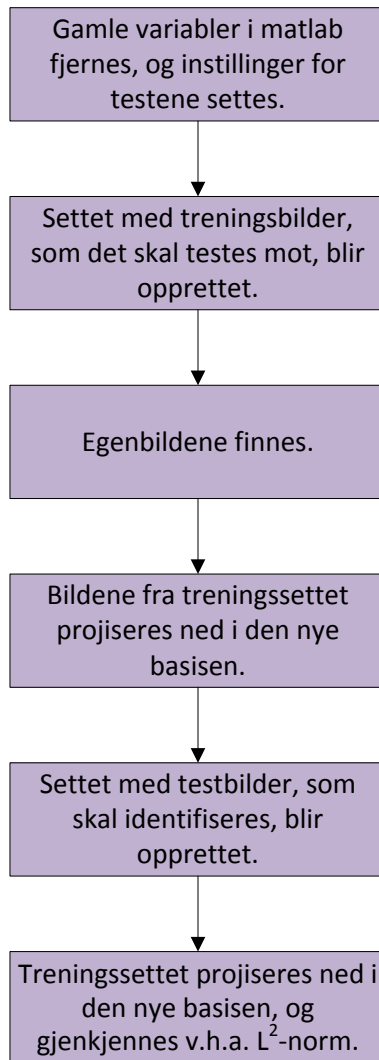
Kapittel 4 – Eksperimenter og resultater

4.1 Implementering

Eksperimentdelen i denne masteroppgaven har tatt for seg P.C.A., hvor det overordnede målet har vært å forsøke å gjenskape de resultatene Sirovich og Meytils kommer fram til i artikkelen [1]. Eksperimentene har blitt utført i Matworks MatLab, versjon 2009a, et matematikkverktøy basert på høynivåprogrammering, [54]. Implementeringen har blitt realisert som et gjenkjenningssystem som utfører gjenkjenning både med og uten bruk av ansiktssymmetri. Det er også mulig å bestemme hvilke egenbilder som skal benyttes under gjenkjenningen hvor ansiktssymmetri brukes. Gjenkjenningssystemet følger framgangsmåten fra kapittel 3.3 og kapittel 0, og er organisert i en struktur som kan ses i Figur 4.1. Hovedfilen, ansiktsgjenkjenning.m, kaller funksjoner som er definerte i de resterende filene. MatLab-kildekoden for gjenkjenningssystemet er vedlagt i Kapittel 7 Appendiks, og er kommentert slik at det fremgår hva som blir gjort. Selve koden som er blitt skrevet, er basert på framgangsmåten for "eigenfaces", slik den er beskrevet av Turk og Pentland, i [13].

Utvidelsen til å ta med de speilede bildene kan matematisk sett gjøres uten å utvide behovet for lagringsplass. I implementeringen har det ikke blitt gjort noe forsøk på å få til dette. På grunn av dette blir det også opprettet en ekstra speilet treningsdatabase. Derfor dobler faktisk denne implementeringen behovet for lagringsplass.

Programflyt



Figur 4.1: Programflyt for P.C.A.-algoritmen.

4.2 Eksperiment 1

Hovedeksperimentet har gått ut på å gjenskape resultatet fra [1]. Det har blitt tatt utgangspunkt i ansiktsdatabasen til Dr. Libor Spacek, fra Universitetet i Essex, [4]. Denne databasen inneholder 359 individer, med 20 bilder av hvert individ. Herfra har vi benyttet to bilder fra 76 individer i den delen av databasen som kalles "Faces94". For å prøve å oppnå sammenlignbare resultater med [1], har disse individene blitt plukket ut. De er hvite menn som ikke bruker briller. De 20 bildene per individ, er i denne delen av databasen tatt samtidig som individene snakker. Derfor er det noe variasjon i ansiktspositur, fra bilde til bilde. Det ble tidligere nevnt at det plukkes ut to bilder for hvert av disse individene. Disse har vært et bilde av et nøytralt ansikt, og et tilfeldig valgt av de andre bildene fra samme individ.

Normalisering av bildene, slik det blir omtalt i kapittel 2.2, er gjort manuelt i Adobe Photoshop, [55]. Her har bildene blitt redusert til gråtoner, rotert og skalert, slik at øynene overlapper to kryss. Deretter har det blitt lagt en oval maske over bildet, slik at bakgrunnen fjernes. Treningssettet for dette eksperimentet, som består av bildene med nøytrale ansiktsuttrykk, blir dermed seende ut slik som i Figur 4.2. Testsettet blir seende ut slik som i Figur 4.3. Det består av bildene med ansikter med tilfeldig positur.

Selve eksperimentet ble utført slik som i ref, hvor en forsøker å bruke en ny basis bestående av alle egenbildene, kun de like egenbildene og kun de like egenbildene, hvor det 2. og det 4. egenbildet fjernes.

Eksperiment 1		
Basis for projeksjon	Antall feilklassifiseringer	Feilrate
Standard P.C.A. basis	6/76	0,08
Like og odde egenbilder	31/76	0,42
Kun like egenbilder	6/76	0,08
Like egenbilder, uten 2. og 4. egenbilde.	6/76	0,08

Tabell 4.1

Resultatet av dette eksperimentet, som kan ses i Tabell 4.1, ser ut til å være i samsvar med resultatene fra [1], se Tabell 4.2, uten å være like gode. Bruk av både like og odde egenbilder, gir relativt høy feilrate. Bruk av kun like egenbilder gir lavere feilrate. Bruk av kun like egenbilder, hvor egenbilde 2 og egenbilde 4 er fjernet, gir lavest feilrate.

Det er ikke umulig at resultatene hadde vært bedre, dersom det hadde vært færre bilder med ansikter i veldig annerledes positur, enn nøytralt ansikt. Sirovich og Meytils bemerker at den eneste feilklassiferingen i deres eksperiment kommer som følge av at personen har lukkede øyne i testbildet.

Resultatene fra [1]		
Basis for projeksjon	Antall feilklassifiseringer	Feilrate
Standard P.C.A. basis	Ingen data.	Ingen data.
Like og odde egenbilder	82/94	0,87
Kun like egenbilder	21/94	0,22
Like egenbilder, uten 2. og 4. egenbilde.	4/94	0,04

Tabell 4.2



Figur 4.2: En oversikt over treningsdatabasen til systemet. Alle bildene kommer fra "Face Recognition Data, University of Essex", [4], normalisert for hånd, med Adobe Photoshop.



Figur 4.3: En oversikt over testdatabasen til systemet. Alle bildene kommer fra "Face Recognition Data, University of Essex", [4], normalisert for hånd, med Adobe Photoshop.

Det er stor sjanse for at resultatet kunne vært annerledes dersom en hadde hatt en større database med personer, eller forsøkt å bruke flere treningsbilder per person. Innledningsvis, ble det samme eksperimentet gjort på to dårligere normaliserte bildesett. I dette tilfellet ble feilraten så høy at eksperimentet ble forkastet. Allikevel kan det være verd å nevne at denne framgangsmåten for ansiktsgjenkjenning ser ut til å være veldig avhengig av god normalisering av bildene, for at en skal få akseptable feilrater. Sirovich og Meytils legger selv fram at metoden ikke gjør noe forsøk på å være et godt gjenkjenningssystem for datasett med større forandringer i ansiktspositur, [2], og denne konklusjonen støttes av resultatene fra dette eksperimentet.

4.3 Eksperiment 2

For å utfylle eksperimentene fra artikkelen, [1], av Sirovich og Meytils, ble det i denne masteroppgaven også gjort en del eksperimenter rundt databaser med bilder med større og mindre grad av sidebelysning, både fra høyre og venstre side. Målet med eksperimentene var å forsøke å verifisere påstanden om at den delen av bildene som projiseres ned i de odde egenbildene kommer som et resultat av sidebelysnings. Sirovich og Meytils hevdet også at ved å fjerne disse egenbildene ville en bli kvitt unødvendig informasjon fra sidebelysning, og derfor oppnå bedre gjenkjenningsresultater. Artikkelen deres har ikke vist til noen resultater som faktisk støtter denne påstanden, og derfor ble det eksperimentert rundt dette.

For å teste denne påstanden ble de to ansiktsdatabasene fra Yale University benyttet. Disse går under navnene "The Yale Face Database", [42], og "The Yale Face Database B", [43]. Begge disse databasene inneholder bilder av personer med nøytrale ansiktspositurer, tatt med sidebelysning. Eksperimentene dekket høyrebelyste bilder, venstrebelyste bilder, ekstremt høyrebelyste bilder og ekstremt venstrebelyste bilder. De vanlig belyste bildene brukte en lyskilde ca. 50 grader rotert vekk fra midten av ansiktene til personene, mens de "ekstremt" sidebelyste bildene brukte en lyskilde rotert ca. 70 grader vekk fra midten av ansiktene. Ansiktene med normale ansiktspositurer kan ses i Figur 4.4. Venstrebelyste ansikter kan ses i Figur 4.5, og ekstremt venstrebelyste bilder kan ses i Figur 4.6.



Figur 4.4: En oversikt over treningsdatabasen til systemet. Bildene kommer fra databasene fra universitetet i Yale, se [42] og [43], normalisert for hånd, med Adobe Photoshop.



Figur 4.5: En oversikt over testdatabasen til systemet, i den venstrebelyste testen. Bildene kommer fra databasene fra universitetet i Yale, se [42] og [43], normalisert for hånd, med Adobe Photoshop.



Figur 4.6: En oversikt over testdatabasen til systemet, i den ekstremt venstrebelyste testen. Bildene kommer fra "Yale B database", se [43], og er normalisert for hånd, med Adobe Photoshop.

I tilfellet hvor de 20 høyre eller venstrebelyste bildene fra Yale databasene brukes som testbilder i gjenkjenningssystemet, får en resultater tilsvarende Tabell 4.3. En ser i dette tilfellet ingen merkbare fordeler med å benytte en egenbasis basert på et utvidet bildesett.

Ekspirement 2: Venstrebelyste testbilder

Basis for projeksjon	Antall feilklassifiseringer	Feilrate
Standard P.C.A. basis	17/20	0,85
Like og odde egenbilder	17/20	0,85
Kun like egenbilder	17/20	0,85
Like egenbilder, uten 2. og 4. egenbilde.	17/20	0,85

Tabell 4.3

Ekspirement 2: Ekstremt venstrebelyste testbilder

Basis for projeksjon	Antall feilklassifiseringer	Feilrate
Standard P.C.A. basis	7/9	0,77
Like og odde egenbilder	7/9	0,77
Kun like egenbilder	7/9	0,77
Like egenbilder, uten 2. og 4. egenbilde.	7/9	0,77

Tabell 4.4

Også i tilfellet der en benytter de "ekstremt" sidebelyste bildene som testbilder, ser en ingen fordel av å benytte en egenbasis basert på et utvidet bildesett. En kan derfor spekulere på om det er noe hold i påstanden om at å la være å projisere bilder ned i odde egenverdier faktisk tar vekk sidebelysningskomponentene i bilder. En lignende studie, [56], av Mili Shah, som implementerer en annen form for ansiktssymmetriutvidelse i P.C.A., kommer fram til resultater som er mer i samsvar med resultatene fra dette eksperimentet. Her sammenligner Shah S.P.S.V.D., "Symmetry-Preserving Singular Value Decomposition", som er den andre formen for symmetriutvidelse som ble nevnt tidligere, med vanlig P.C.A.. Resultatene av denne sammenligningen kan ses i Figur 4.7. S.P.S.V.D. er i denne figuren sammenlignet med S.V.D. (engelsk: Singular Value Decomposition). Mili Shah hevder selv at implementeringen av S.V.D. i denne artikkelen, [56], er matematisk ekvivalent med P.C.A.. Konklusjonen som dras er at S.P.S.V.D. er minst like god som vanlig S.V.D., men algoritmen behøver bare halvparten så lang kjøretid. Dersom vi sammenligner disse resultatene med de som denne masteroppgaven produserer, ser vi at en feilrate i størrelsesorden 70%-80% ikke er helt urimelig. Resultatene støtter også at det ikke behøver å være stor forskjell mellom vanlig P.C.A. og P.C.A. utvidet med speilede bilder.

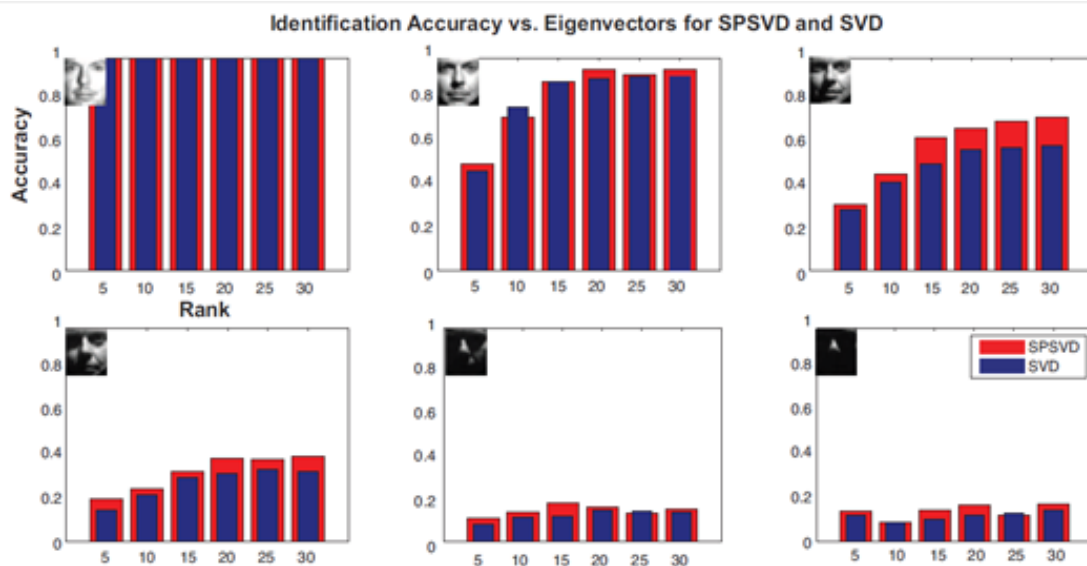


FIG. 4.3. Percentage of correct identification between the SPSVD (red) and the SVD (blue) with increased lighting variation and increased dimensionality of the face database.

Figur 4.7: Resultater fra et lignende eksperiment, utført i [56].

Konklusjonen fra dette eksperimentet er at det trolig er lite hold i påstanden om at i et bildesett, utvidet til å inneholde speilede bilder, er sidebelysningskomponentene representert i de odde egenbildene.

Kapittel 5 - Konklusjon

Innledningsvis ble det nevnt at målene med masteroppgaven skulle være et litteraturstudie, med utgangspunkt i artikkelen, [1], av Lawrence Sirovich og Masha Meytils. Det skulle også gjøres et forsøk på å gjenskape resultatene fra denne artikkelen i egne eksperimenter. Grunnlaget for denne konklusjonen vil derfor være å drøfte hvorvidt disse målene har blitt nådd.

For litteraturstudier er det ofte vanskelig å dra en konklusjon på hvorvidt en har lykkes eller ikke. Derfor presenteres kort de hovedinntrykkene forfatteren sitter igjen med, etter endt litteraturstudie.

Ansiktsgjenkjenning er et fagfelt som har opplevd mye framgang de siste 20 årene. Spesielt ansiktsgjenkjenning i kontrollerte omgivelser begynner å nærme seg perfektjonert, og en beveger seg mot at "state of the art"-algoritmene er eid og beskyttet av kommersielle selskaper, og er ikke tilgjengelig.

Forskning innenfor fagfeltet har nå begynt å ta for seg de vanskelige problemstillingene, som kan gjøre ansiktsgjenkjenning mer anvendelig for praktiske formål hvor en har ukontrollerte forhold. Her er det snakk om gjenkjenning fra video, gjerne med lav pikseloppløsning, slik som mange overvåkingskameraer. En har også tatt fatt i gjenkjenning hvor det høy variasjon i belysning og hvilket perspektiv en ser bildene fra.

Litteraturstudiet har også tatt for seg P.C.A. i ganske stor detalj. Dette har vært en matematisk krevende del av oppgaven, og en kunne kanskje ha redusert dybden av denne delen noe, til fordel for mer dybde om flere andre algoritmer. Dette ville naturlig nok ha gått ut over implementeringen av symmetridelen i eksperimentene, noe som var en av hovedproblemstillingene for oppgaven.

På grunn av lite dokumentasjon av hvordan denne symmetridelen ble implementert i [1], har implementeringen, som danner grunnlaget for eksperimentene, tatt mye tid. Forklaringen av framgangsmåten er spredt utover flere artikler, og stiller nokså høye krav til leserens matematikkunnskaper.

Forsøket på å gjenskape resultatene fra [1] har vært nokså vellykket. Fra Eksperiment 1 klarer en å få feilrater som er i nærheten av de som presenteres i artikkelen. Forskjellene her kommer sannsynligvis fra normaliseringen av bildene, som er gjort for hånd. En del av bildene som brukes i eksperimentet har nok også litt for store variasjoner i ansiktsuttrykkene mellom treningsbildene og testbildene.

I eksperiment 2 klarer en ikke å vise noen sammenheng mellom sidebelysning og de odde egenbildene.

I innledningen ble det også nevnt at valget av masteroppgave kommer fra forfatterens ståsted om at observasjonene i [1] kan være viktige funn innen ansiktsgjenkjenning. Som en avslutting på oppgaven kan det sies at forfatterens oppfatning har forandret seg noe. Observasjonene kan nok være nyttige, men det er stilles en del krav til systemet for å kunne dra nytte av de. Problemstillingen for videre arbeid rundt de, blir om det er mulig å dra nytte av de i praktiske anvendelser.

Kapittel 6 - Referanser

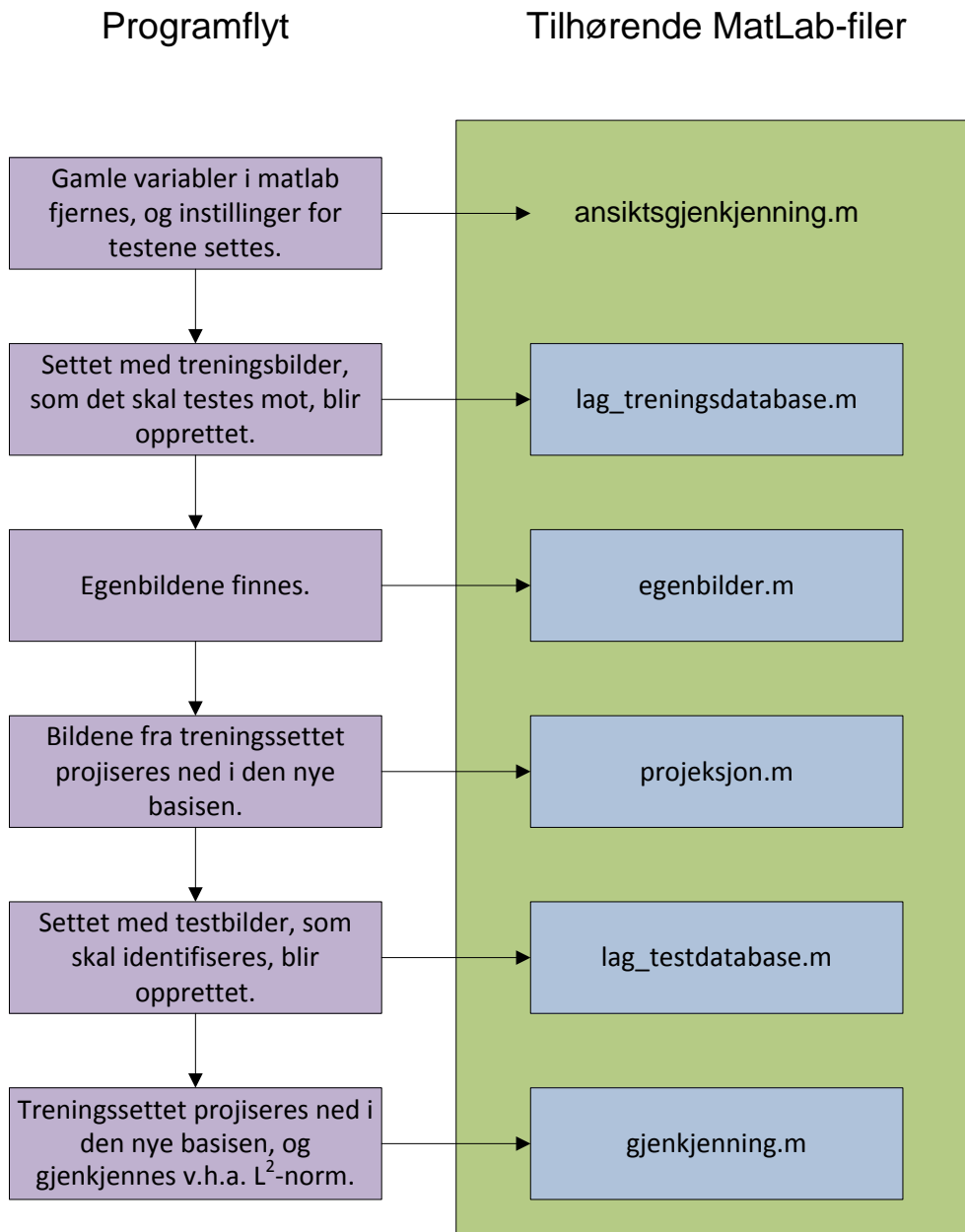
1. Sirovich, L. and M. Meytils, *Symmetry, probability, and recognition in face space*. PNAS - Proceedings of the National Academy of Sciences, 2009. **106**(17): s. 6895-6899.
2. Kirby, M. and L. Sirovich, *Application of the Karhunen-Loeve procedure for the characterization of human faces*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1990. **12**(1): s. 103-108.
3. Torres, L., *Is there any hope for face recognition?* WIAMIS, Proc. of the 5th International Workshop on Image Analysis for Multimedia Interactive Services, 2004.
4. Spacek, D.L. *Face Recognition Data*. 2008 [sitert 2010 13. juni]; Tilgjengelig fra: <http://cswww.essex.ac.uk/mv/allfaces/index.html>.
5. Bailly-Bailliere, E., et al., *The BANCA database and evaluation protocol*. Audio- and Video-Based Biometric Person Authentication (AVBPA), 2003: s. 625–638.
6. Phillips, P.J., P. Rauss, and S. Der, *FERET (Face Recognition Technology) Recognition Algorithm Development and Test Report*. Technical Report ARL-TR 995, 1996.
7. Phillips, P.J., et al., *The FERET Evaluation Methodology for Face- Recognition Algorithms*. Proceedings, IEEE Conference on Computer Vision and Pattern Recognition, 1997: s. 137-143.
8. Rizvi, S.A., P.J. Phillips, and H. Moon, *A Verification Protocol and Statistical Performance Analysis for Face Recognition Algorithms*. Proceedings, IEEE Conference on Computer Vision and Pattern Recognition, 1998: s. 833-838.
9. Phillips, P.J., et al., *The FERET Testing Protocol*. Face Recognition: From Theory to Applications, Berlin: Springer-Verlag, 1998: s. 244-261.
10. Phillips, P.J., et al. *FRVT 2006 and ICE 2006 Large-Scale Results*. 2006 [sitert 2010 13. juni]; Tilgjengelig fra: <http://www.frvt.org/FRVT2006/docs/FRVT2006andICE2006LargeScaleReport.pdf>.
11. Zhao, W., et al., *Face Recognition: A Literature Survey*. ACM Computing Surveys, 2003: s. 399-458.
12. Sirovich, L. and M. Kirby, *Low-dimensional Procedure for the Characterization of Human Faces*. Journal of the Optical Society of America A - Optics, Image Science and Vision, 1987. **4**(3): s. 519-524.
13. Turk, M. and A. Pentland, *Eigenfaces for Recognition*. Journal of Cognitive Neuroscience, 1991. **3**(1): s. 71-86.
14. Belhumeur, P.N., J.P. Hespanha, and D.J. Kriegman, *Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997. **19**(7): s. 711-720.
15. Duda, R.O., P.E. Hart, and D.G. Stork, *Pattern Classification*. Second ed. 2007: John Wiley & Sons, Inc.
16. Shashua, A., *Geometry and Photometry in 3D Visual Recognition*. PhD Thesis, MIT, 1994.
17. Kohonen, T., *Self-Organization and Associative Memory*. Berlin: Springer-Verlag, 1988.
18. Brunelli, R. and T. Poggio, *Hyper BF Networks for Gender Classification*. Proceedings, DARPA Image Understanding Workshop, 1992: s. 311-314.
19. Guo, G., S.Z. Li, and K. Chan, *Face Recognition by Support Vector Machines*. Proc. of the IEEE International Conference on Automatic Face and Gesture Recognition, 2000: s. 196-201.

20. Ellis, H.D., *Introduction Aspects of Face Processing: Ten Questions in Need of Answers*. Aspects of Face Processing, Dordrecht: Nijhoff, 1986: s. 3-13.
21. Biederman, I. and P. Kalocsai, *Neural and Psychophysical Analysis of Object and Face Recognition*. Face Recognition: From Theory to Applications, Berlin: Springer-Verlag, 1998: s. 3-25.
22. Bruce, V., P.J.B. Hancock, and A.M. Burton, *Human Face Perception and Identification*. Face Recognition: From Theory to Applications, Berlin: Springer-Verlag, 1998: s. 51-72.
23. Barlett, J.C. and J. Searcy, *Inversion and Configuration of Faces*. Cognitive Psychology, 1993. **25**: s. 281-316.
24. Tompson, P., *Margaret Thatcher - A New Illusion*. Perception, 1980. **9**(483-484).
25. Perkins, D., *A Definition of Caricature and Recognition*. Studies in the Anthropology of Visual Communication, 1975. **2**: s. 1-24.
26. Harmon, L.D., *The Recognition of Faces*. Scientific American, 1973. **229**(5): s. 71-82.
27. Ginsburg, A.P., *Visual Information Processing Based on Spatial Filters Constrained by Biological Data*. AMRL Technical Report, 1978: s. 78-129.
28. Sergent, J., *Microgenesis of Face Perception*. Aspects of Face Processing, Dordrecht: Nijhoff, 1986.
29. Tarr, M.J. and H.H. Bulthoff, *Is Human Object Better Described by Geon Structural Descriptions or by Multiple Views - Comment on Biederman and Gerhardstein (1993)*. Journal of Experimental Psychology: Human Perception and Performance, 1995. **21**: s. 71-86.
30. Biederman, I., *Recognition by Components: A Theory of Human Image Understanding*. Psychological Review, 1987. **94**: s. 115-157.
31. Marr, D., *Vision*. San Francisco: W. H. Freeman, 1982.
32. Hill, H., P.G. Schyns, and S. Akamatsu, *Information and Viewpoint Dependence in Face Recognition*. Cognition, 1997. **62**: s. 201-222.
33. Johnston, A., H. Hill, and N. Carman, *Recognizing Faces: Effects of Lighting Direction, Inversion and Brightness Reversal*. Cognition, 1992. **40**: s. 1-19.
34. Hill, H. and V. Bruce, *Effects of Lighting and Matching Facial Surfaces*. Journal of Experimental Psychology: Human Perception and Performance, 1996. **22**: s. 986-1004.
35. Knight, B. and A. Johnston, *The Role of Movement in Face Recognition*. Visual Cognition, 1997. **4**: s. 265-274.
36. Bruce, V., *Recognizing Faces*. London: Lawrence Erlbaum Associates, 1988.
37. Valentin, D., H. Abdi, and B. Edelman, *What represents a face: A computational approach for the integration of physiological data*. Perception, 1997. **26**: s. 1271-1288.
38. Krouse, F., *Effects of pose, pose change, and delay on face recognition performance*. Journal of Applied Psychology, 1981. **66**: s. 651-654.
39. Gross, R., *Face Databases*. Handbook of Face Recognition, Springer-Verlag, 2005.
40. Grgic, P.D.M. and P.D.K. Delac. *Face Recognition Homepage - Databases*. 2010 [sitert 2010 13. juni]; Tilgjengelig fra: <http://www.face-rec.org/databases/>.
41. Martinez, A.R. and R. Benavente, *The AR face database*. Technical Report 24, ComputerVision Center(CVC) Technical Report, Barcelona, 1998.
42. Georghiades, A.S., P.N. Belhumeur, and D.J. Kriegman. *Yale Face Database*. 1997 [sitert 2010 13. juni]; Tilgjengelig fra: <http://cvc.yale.edu/projects/yalefaces/yalefaces.html>.

43. Georghiades, A.S., P.N. Belhumeur, and D.J. Kriegman, *From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose*. IEEE Trans. Pattern Anal. Mach. Intelligence, 2001. **23**(6): s. 643-660.
44. Gao, W., et al., *CAS-PEAL large-scale Chinese face database and evaluation protocols*. Technical Report JDL-TR-04-FR-001, Joint Research & Development Laboratory, 2004.
45. Grgic, M., K.Delac, and S. Grgic, *SCface – surveillance cameras face database*. Multimedia Tools and Applications, Springer Netherlands, 2009.
46. Gonzalez, R.C. and R.E. Woods, *Digital Image Processing*. Third ed. 2008: Prentice-Hall, Inc.
47. Penev, P.S. and L. Sirovich, *The Global Dimensionality of Face Space*. Fourth IEEE International Conference on Automatic Face and Gesture Recognition, 2000. Proceedings, 2000: s. 264-270.
48. Meytils, M. and L. Sirovich, *On the Dimensionality of Face Space*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007. **29**(7): s. 1262-1267.
49. Freiwald, W. and D. Tsao. *Single-unit recording in fMRI-identified macaque face patches: Evidence for a second domain-specific module*. 2006 [sitert 2010 06. mai]; Tilgjengelig fra:
www.abstractsonline.com/viewer/viewAbstractPrintFriendly.asp?CKey={11A9C9EE-3DA4-4585-8C2E-8A19558C318A}&SKey={F8DC3D3D-E02F-46BE-870C-EBA5C9E47285}&MKey={D1974E76-28AF-4C1C-8AE8-4F73B56247A7}&AKey={3A7DC0B9-D787-44AA-BD08-FA7BB2FE9004}.
50. Tsao, D., et al., *A cortical region consisting entirely of face-selective cells*. Science, 2006. **311**: s. 670-674.
51. Lay, D.C., *Linear Algebra and its Applications*. Third ed. 2002: Pearson Education.
52. Watanabe, S., *Karhunen-Loève expansion and factor analysis theoretical remarks and applications*. Proc. 4th Prague Conf. Inform. Theory, 1965.
53. Sirovich, L., *Turbulence and the dynamics of coherent structures, Part II: Symmetries and transformations*. Quart. Appl. Math., 1987. **XLV**(3): s. 573-582.
54. MathWorks. *MatLab 2009a*. Tilgjengelig fra: <http://www.mathworks.com/>.
55. Adobe. *Photoshop CS4*. Tilgjengelig fra: <http://www.adobe.com/>.
56. Shah, M. *Symmetric Eigenfaces*. 2009; Tilgjengelig fra:
http://math.loyola.edu/reports/tr2009_01.pdf.

Kapittel 7 Appendiks

Appendiksen viser hvordan MatLab-programmet blir seende ut. Programflyten er vist sammen med sine tilhørende MatLab-filer i Figur 7.1.



Figur 7.1: Programflyt med tilhørende filstruktur.

7.1 ansiktsgjenkjenning.m

```
%% Rydd opp i MatLab før start.
clear all
close all
clc
%% Innstillinger.
% Biblioteksti til bildene til treningsdatabasen.
biblioteksti =
'f:\Masteroppgave\Algoritmer\Gjenkjenningssystem\Treningsdatabase';
% Bildetekst for bildene til treningsdatabasen.
% For bilde1, bilde2 etc. er bildetekst 'bilde'.
bildetekst = 'img';
% Start og stoppnummer for bildene til databasene.
% For eksempel fra bilde1 til bilde100.
bildenrstart = 1;
bildenrstop = 76;
% Filtype for bildene.
filtype = '.jpg';
% Biblioteksti til bildene til testdatabasen.
biblioteksti_test =
'f:\Masteroppgave\Algoritmer\Gjenkjenningssystem\Testdatabase';
% Bildetekst for bildene til testdatabasen.
bildetekst_test = 'test';
% Hvilken egenbasis som skal ligge til grunn for gjenkjenningen med
% ansiktssymmetri, og P.C.A..
% Muligheter: 'full', 'like' og 'like_reduisert_med_2_og_4'.
egenbasis = 'like_reduisert_med_2_og_4';

%% Oppretting av treningsdatabasen.
[phi, phi_speilet] =
lag_treningsdatabase(biblioteksti, bildetekst, bildenrstart, bildenrstop, filtype);

%% Utrekning av egenbilder for treningsdatabasen.
[egenbilder_sym, egenbilder, gjennomsnittsbilde, a] =
egenbilder(phi, phi_speilet, egenbasis);

%% Projeksjon av bildene fra treningsdatabasen inn i ansiktsrommet.
[projiserte_bilder_sym, projiserte_bilder] =
projeksjon(egenbilder_sym, egenbilder, a);

%% Oppretting av testdatabasen.
[testdatabase] =
lag_testdatabase(biblioteksti_test, bildetekst_test, bildenrstart, bildenrstop, filtype, gjennomsnittsbilde);

%% Gjenkjenning av bildene fra testdatabasen.
gjenkjenning(testdatabase, projiserte_bilder_sym, projiserte_bilder, egenbilder_sym, egenbilder);
```

7.2 lag_treningsdatabase.m

```
function [phi,phi_speilet] =
lag_treningsdatabase(biblioteksti,bildetekst,bildenrstart,bildenrstoppp,filtype)
% Funksjonsbeskrivelse
%

%% Oppretter databasen.
phi = [];
phi_speilet = [];
for i = bildenrstart:bildenrstoppp
    % Laster inn ett og ett bilde, for hver iterasjon av for-løkken.
    if i < 10
        nullpadding = '00';
    elseif i < 100
        nullpadding = '0';
    else
        nullpadding = '';
    end
    bildesti_og_navn =
strcat(biblioteksti,'\',bildetekst,nullpadding,int2str(i),filtype);
    bilde = double(rgb2gray(imread(bildesti_og_navn)));

    % Henter størrelsen på bildet.
    [radlengde kolonnelengde] = size(bilde);
    % Gjør bildet om til en kolonnevektor ved hjelp av reshape-funksjonen,
    % med så størrelsen til bildet som parametere.
    kolonnevektor = reshape(bilde',radlengde*kolonnelengde,1);
    % Legger så elementene inn i databasen phi_like.
    phi = [phi kolonnevektor];

    % Oppretter den speilede varianten av bildet, slik at den kan legges
    % til, eller trekkes fra bildeverdien når alfa og beta skal beregnes.
    speilet_bilde = fliplr(bilde);
    kolonnevektor = reshape(speilet_bilde',radlengde*kolonnelengde,1);
    % Legger så elementene inn i databasen phi_like.
    phi_speilet = [phi_speilet kolonnevektor];
end
```

7.4 egenbilder.m

```
function [egenbilder_sym,egenbilder,gjennomsnittsbilde,a] =
egenbilder(phi,phi_speilet,egenbasis)
%
%

%% Beregner gjennomsnittsbilde.
gjennomsnittsbilde = mean(phi,2);
gjennomsnittsbilde_speilet = mean(phi_speilet,2);
gjennomsnittsbilde_sym = mean([gjennomsnittsbilde
gjennomsnittsbilde_speilet],2);

%% Henter ut antall bilder.
bildeantall = size(phi,2);

%% Oppretter alfa og beta matrisene.
alfa = [];
beta = [];
a = [];
for i = 1:bildeantall
    % Finner først karikaturen for hvert bilde, for hver iterasjon av
    % for-løkken.
    karikatur = phi(:,i) - gjennomsnittsbilde;
    karikatur_sym = phi(:,i) - gjennomsnittsbilde_sym;
    % For hver iterasjon legges deretter det speilede bildet til
    % den korresponderende karikaturen. Resultatet er kolonnevektorene i
    % matrisa alfa.
    alfa = [alfa (karikatur_sym + phi_speilet(:,i) -
mean([gjennomsnittsbilde gjennomsnittsbilde_speilet],2))]];
    % For hver iterasjon trekkes deretter det speilede bildet fra
    % den korresponderende karikaturen. Resultatet er kolonnevektorene i
    % matrisa beta.
    beta = [beta (karikatur_sym - phi_speilet(:,i) -
mean([gjennomsnittsbilde gjennomsnittsbilde_speilet],2))]];
    % Når symmetri ikke benyttes samles karikaturene i matrisa a.
    a = [a karikatur];
end

a_sym = a;

for i=1:bildeantall
    karikatur_sym = phi_speilet(:,i) - gjennomsnittsbilde_sym;
    a_sym = [a_sym karikatur_sym];
end

%% Beregner egenbildene
% Finner substituttmatrisene, L_like og L_odde, for kovariansmatrisene i
% de to dekoplete PCA-problemstillingene.
L_like = alfa'*alfa;
L_odde = beta'*beta;
% Finner substituttmatrisen L, for a i tilfellet uten symmetri.
L = a'*a;

% Beregner deretter egenverdiene, og de tilhørende egenvektorene for
% substituttmatrisene.
[egenvektorer_L_like,egenverdier_L_like] = eig(L_like);
[egenvektorer_L_odde,egenverdier_L_odde] = eig(L_odde);
% Tilsvarende gjøres for tilfellet uten symmetri.
[egenvektorer_L,egenverdier_L] = eig(L);
```

```

% Videre sorteres egenvektorene etter synkende egenverdi.
[egenvektorer_L,egenverdier_L] = sortem2(egenvektorer_L,egenverdier_L);

% Til slutt beregnes de egentlige egenbildene ved å multiplisere
% karikaturene med egenvektorene fra substituttmatrisen.
egenbilder_like = alfa*egenvektorer_L_like;
egenbilder_odde = beta*egenvektorer_L_odde;

% Egenbildene for symmetritilfellet må også sorteres etter sorteres
% etter synkende egenverdi.
switch egenbasis
    case 'full'
    case 'like'
        % Fjerner odde egenbilder.
        egenvektorer_L_odde = [];
        egenverdier_L_odde = [];
        egenbilder_odde = [];
        beta = [];
    case 'like_redusert_med_2_og_4'
        % Fjerner odde egenbilder.
        egenvektorer_L_odde = [];
        egenverdier_L_odde = [];
        egenbilder_odde = [];
        beta = [];
        % Fjerner like egenbilder med indeks 4.
        egenvektorer_L_like(:,4) = [];
        egenverdier_L_like(:,4) = [];
        egenbilder_like(:,4) = [];
        alfa(:,4) = [];
        % Fjerner like egenbilder med indeks 2.
        egenvektorer_L_like(:,2) = [];
        egenverdier_L_like(:,2) = [];
        egenbilder_like(:,2) = [];
        alfa(:,2) = [];
    otherwise
        disp('Feil: Det er definert hvilken egenbasis som skal benyttes. ');
        disp('Benytter derfor full egenbasis');
end

[Ml Nl] = size(egenverdier_L_like);
[Mo No] = size(egenverdier_L_odde);
[egenbilder_sym,egenverdier_sym] = sortem2([egenbilder_like
egenbilder_odde],[egenverdier_L_like zeros(Ml,No); zeros(Mo,Nl)
egenverdier_L_odde]);

% Tilsvarende beregnes de egentlige egenbildene også for tilfellet uten
symmetri.
egenbilder = a*egenvektorer_L;

```

7.5 projeksjon.m

```
function [projiserte_bilder_sym,projiserte_bilder] =
projeksjon(egenbilder_sym,egenbilder,a)
%
%
%% Projeksjon av bildene fra treningsdatabasen inn i ansiktsrommet.
projiserte_bilder_sym= [];
antallBilder = size(a,2);
for i = 1:antallBilder
    projeksjon = egenbilder_sym'*a(:,i);
    projiserte_bilder_sym = [projiserte_bilder_sym projeksjon];
end

%% Tilsvarende gjøres for tilfellet uten symmetri.
projiserte_bilder= [];
antallBilder = size(a,2);
for i = 1:antallBilder
    projeksjon = egenbilder'*a(:,i);
    projiserte_bilder = [projiserte_bilder projeksjon];
end

end
```

7.6 lag_testdatabse.m

```
function [testdatabse] =
lag_testdatabse(biblioteksti,bildetekst,bildenrstart,bildenrstop,filtype,
gjenomsnittsbilde)
% Funksjonsbeskrivelse
%
%% Oppretter databasen.
testdatabse = [];
for i = bildenrstart:bildenrstop
    % Laster inn ett og ett bilde, for hver iterasjon av for-løkken.
    if i < 10
        nullpadding = '00';
    elseif i < 100
        nullpadding = '0';
    else
        nullpadding = '';
    end
    bildesti_og_navn =
strcat(biblioteksti,'\',bildetekst,nullpadding,int2str(i),filtype);
    bilde = double(rgb2gray(imread(bildesti_og_navn)));

    % Henter størrelsen på bildet.
    [radlengde kolonnelengde] = size(bilde);
    % Gjør bildet om til en kolonnevektor ved hjelp av reshape-funksjonen,
    % med så størrelsen til bildet som parametere.
    kolonnevektor = reshape(bilde',radlengde*kolonnelengde,1);
    % Trekker så gjenomsnittsbilde fra kolonnevektorene.
    kolonnevektor = kolonnevektor - gjenomsnittsbilde;
    % Legger så elementene inn i databasen phi_like.
    testdatabse = [testdatabse kolonnevektor];
end
```

7.7 gjenkjenning.m

```
function
gjenkjenning(testdatabase,projiserte_bilder_sym,projiserte_bilder,egenbilde
r_sym,egenbilder)
%
%

%% Finner først P.C.A. for bildene i testdatabasen.
projiserteTestbilder_sym= [];
antall_bilder_test = size(testdatabase,2);
for i = 1:antall_bilder_test
    projeksjon = egenbilder_sym'*testdatabase(:,i);
    projiserteTestbilder_sym = [projiserteTestbilder_sym projeksjon];
end

projiserteTestbilder= [];
antall_bilder_test = size(testdatabase,2);
for i = 1:antall_bilder_test
    projeksjon = egenbilder'*testdatabase(:,i);
    projiserteTestbilder = [projiserteTestbilder projeksjon];
end

%% Beregner videre Euklidisk avstand til alle projeksjonene av bildene fra
% treningsdatabasen.
feilklassifisering = 0;
antall_bilder_trening = size(egenbilder,2);
for i = 1:antall_bilder_test
    euklidisk_avstand_sym = [];
    for j = 1:antall_bilder_trening
        euk_avst = (norm(projiserteTestbilder_sym(:,i) -
projiserte_bilder_sym(:,j)))^2;
        euklidisk_avstand_sym = [euklidisk_avstand_sym euk_avst];
    end
    % Testbildet skal ha minst euklidisk avstand til sitt tilsvarende bilde
i
    % treningsdatabasen.
    [Minste_euklidiske_avstand, indeks] = min(euklidisk_avstand_sym);
    if i ~= indeks
        feilklassifisering = feilklassifisering + 1;
    end
    % Vis hvilke bilder som blir klassifisert som hva.
    %disp(['Klassifisering: ' int2str(i) ' som ' int2str(indeks)]);
end

disp(['Feilklassifiseringer med bruk av ansiktssymmetri: '
num2str(feilklassifisering)]);
disp(['Feilklassifiseringsrate: '
num2str(feilklassifisering/antall_bilder_test)]);
```

```

% Samme klassifisering utføres for tilfellet uten symmetri.
feilklassifisering_uten_symmetri = 0;
antall_bilder_trening = size(egenbilder,2);
for i = 1:antall_bilder_test
    euklidisk_avstand = [];
    for j = 1:antall_bilder_trening
        euk_avst = (norm(projiserteTestbilder(:,i) -
projiserte_bilder(:,j)))^2;
        euklidisk_avstand = [euklidisk_avstand euk_avst];
    end
    % Testbildet skal ha minst euklidisk avstand til sitt tilsvarende bilde
i
    % treningsdatabasen.
    [Minste_euklidiske_avstand, indeks] = min(euklidisk_avstand);
    if i ~= indeks
        feilklassifisering_uten_symmetri = feilklassifisering_uten_symmetri
+ 1;
    end
    % Vis hvilke bilder som blir klassifisert som hva.
    %disp(['Klassifisering: ' int2str(i) ' som ' int2str(indeks)]);
end

disp(['Feilklassifiseringer uten bruk av ansiktssymmetri: '
num2str(feilklassifisering_uten_symmetri)]);
disp(['Feilklassifiseringsrate: '
num2str(feilklassifisering_uten_symmetri/antall_bilder_test)]);
end

```