

# Sammenfatning

Denne Ph.D. afhandling består af to dele, som hver især beskriver de to linjer af forskning jeg har udført. Afhandlingens titel *Statistics of Natural Image Geometry* afspejler min forskning angående statistikken af geometri i naturlige billeder, som udgør hovedvægten af denne afhandling (Part I). Den anden del af afhandlingen (Part II) omhandler analyse af videosekvenser med henblik på detektion og beregning af bevægelse.

## Statistik af naturlige billeder

Der er blevet argumenteret for, af denne og andre forfattere (Mumford, 1996; Srivastava et al., 2003), at kendskab til statistikken af naturlige billeder kan være nyttig ved udvikling af datamatsyn og billedanalysealgoritmer. Betragter man datamatsyn og billedanalyseproblemer som statistisk inferensproblemer kan man med fordel benytte den såkaldte Bayesianske inferens. Her er det nødvendigt at tildele en sandsynlighedsfordeling til a priori viden eller modelparametre. Der er i denne afhandling blevet argumenteret for at sådanne sandsynlighedsfordelinger kan konstrueres ud fra empirisk viden om statistikken af naturlige billeder. Fokus har specielt været på den statistiske variation af lokal geometri i billeder ud fra betragtningen, at lokal geometri udgør lav-niveau primitiver/features i mange datamatsynsopgaver.

I afhandlingen er den statistiske variation i høj-kontrast  $3 \times 3$  mikrobilleder af naturlige billeder blevet studeret. Der blev anvendt billeder med to forskellige modaliteter — billeder optaget med et konventionelt optisk kamera, samt billeder optaget med en 3D-laserskanner. Bemærk at begge billedtyper indeholder information om den betragtede scenes geometri. Det blev analyseret hvorledes disse mikrobilleder fordeler sig i tilstandsrummet med hensyn til en model af kanter. Den benyttede model er en udglattet trinkant, som danner en to-dimensional manifold i tilstandsrummet. Empiriske studier viser at naturlige optiske mikrobilleder fordeler sig tæt omkring kantmanifolden i tilstandsrummet. Fordelingen af optiske mikrobilleder med hensyn til afstand  $\theta$  til kantmanifolden er proportionel med  $\theta^{-2.5}$ . 3D-laser mikrobilleder

har en markant anderledes fordeling. Denne type mikrobilleder ligger tæt fordelt omkring punkter i tilstandsrummet, der kan betegnes som binære mikrobilleder. Nogle af disse binære mikrobilleder ligger på kantmanifolden, mens andre ligger spredt i tilstandsrummet. Fordelingen af 3D-laser mikrobilleder omkring binære mikrobilleder følger, ligesom optiske mikrobilleder, en potenslov. Forskellen i de to billedtypers fordeling af lokal geometri skyldes forskellene i de to modaliteters subpixel-egenskaber: Det optiske kamera kan sammenlignes med et udglattende filter og 3D-laserskanneren med et morfologisk filter som gør den lokale struktur skarpere.

Studiet af høj-kontrast  $3 \times 3$  pixel mikrobilleder er blevet udvidet med et studium af den statistiske variation af naturlige optiske billeder afbildet ind i det såkaldte jet-rum. Akserne i jet-rummet består af de partielle billedafledte, og i dette studium benyttes afledte op til tredje orden, hvilket definerer 3-jettet. Billedafledte defineredes i dette studium ved hjælp af lineær skalarumsteori. I denne teori introduceres en skala hvorpå billedafledte måles. Et punkt i jet-rummet beskriver den lokale geometri i et punkt i billedet. Igen fokuseres der på analyse af kanter og en analytisk kantmodel introduceres. Modellen består af skalarummet af en trinkant og den udspænder en to-dimensional  $C^\infty$  manifold i jet-rummet. Dette studium viser at den lokale geometri af naturlige billeder er tæt fordelt omkring kantmanifolden. Fordelingen af lokal geometri med hensyn til afstand  $\theta$  til kantmanifolden er proportional med  $\theta^{-0.7}$ , hvilket er i overensstemmelse med forsøgene med høj-kontrast  $3 \times 3$  pixel optiske mikrobilleder. Ydermere er denne fordeling skalainvariant.

Disse to studier antyder at de såkaldte features, dvs. kanter, blobs, hjørner etc., danner klynger eller manifolde i tilstandsrummet uafhængigt af billedrepræsentation, dvs. pixel-intensitetsværdier eller filtersvar. Dimensionen af en feature-manifold afhænger ikke af billedrepræsentationen, men bestemmes udelukkende af kompleksiteten af featuren eller antallet af frihedsgrader den pågældende feature besidder, f.eks. er dimensionen af kant manifolden 2 og dimensionen af manifolden for en isotrop blob 1. Det er rimeligt at antage at feature-manifolder er indlejret i hinanden og er del af et hierarki af manifolde, f.eks. er det oplagt at kantmanifolden er en del af manifolden for hjørner. Man kan forestille sig muligheden for at lære features ved at lære modeller for de tilhørende manifolde i tilstandsrummet.

Udover disse studier af lokal geometri, er der blevet arbejdet med en matematisk model for billeder. Denne model kaldes *den Brownske billedmodel* og er en stokastisk funktion på planet. Den Brownske model har en række attraktive egenskaber, blandt andet er den skalainvariant og har en Gaussisk sandsynlighedsfordeling. Vigtigst er dog at denne model fanger anden ordens statistikken af naturlige billeder. Empiriske studier viser nemlig at natur-

lige billeder har en skalainvariant kovariansstruktur. Det er i denne afhandling blevet vist empirisk at den Brownske billedmodel modellerer kovariansstrukturen af naturlige billeder afbildet ind i jet-rummet. Ydermere er den Brownske billedmodel og dens generalisering — *den fraktale Brownske billedmodel* — blevet anvendt som grundlag for en udledning af en metode til skalanormalisering af billedafledte i skalarum. Denne skalanormalisering er en udvidelse af en metode foreslået af Lindeberg (1998b) i forbindelse med en metode til automatisk skalaselektion i feature-detektorer.

## Bevægelsesanalyse

I afhandlingens anden del udvikles en algoritme til analyse af bevægelse i videosekvenser. I en bevægelsesanalyse ønsker man at beregne et vektorfelt som beskriver størrelse og retning af ændringen af intensitetsværdier i den analyserede videosekvens. Dette vektorfelt kaldes for *det optiske flow-felt*. Bevægelsesanalyse er vigtigt for løsningen af en række problematikker inden for datamatysyn. Fokus har i denne afhandling været på udviklingen af en metode til at beregne hastighedsvektorfelter med henblik på analyse af væskebevægelse. Her er det vigtigt ikke at lave forsimplende antagelser såsom at den analyserede scene kun indeholder stive legemer og at bevægelsen kan beskrives ved en affin transformation. Disse antagelser bryder generisk sammen ved analyse af væskebevægelse.

I den udviklede metode antages det, at det såkaldte normal-flow er givet på forhånd. Det optiske flow modelleres lokalt med en polynomisk model. Denne models parametre estimeres ved at tilpasse modellens projektion ind på den lokale iso-intensitetskurvenormal til det underliggende normal-flow. Modellens tangentielle komponent udledes ved en integration over et nabo-område. I metoden benyttes lineær skalarumsteori til at beregne afledte af billeder og normal-flow, samt at definere integrationsområdet. Modelparametrene estimeres lokalt ved at finde løsningen til et lineært ligningssystem. I metoden indgår også en teknik til automatisk selektion af de involverede skalaparametre.

I afhandlingen sammenlignedes den her beskrevne algoritme til beregning af optisk flow med andre kendte metoder. Denne sammenligning bestod af en kvalitativ vurdering af de involverede metoders evne til at analysere billedsekvenser med et kendt bevægelsesfelt. Sammenligningen viser at den her foreslåede metode producerer hastighedsfelter som er lige så gode, og i visse tilfælde bedre, end de bedste af de andre metoder. Det blev også sandsynliggjort at den her foreslåede metode kan beregne bevægelsesfelter af virkelige videosekvenser af væskebevægelse, som er en god approksimation af det faktiske bevægelsesfelt.