

# 博士論文

## 画像・映像理解のための ノイズ特性推定とその応用

Estimating Noise Characteristics and its Applications  
for Image/Video Understanding



東京大学大学院  
情報理工学系研究科  
電子情報学専攻

48-077404 小林 理 弘

指導教員 佐藤 洋一 准教授

平成22年6月



# 概要

従来，信号処理の分野において，ノイズは信号を阻害するものであり，除去すべきものであるという認識が広く持たれてきた．コンピュータビジョンの分野においても，画像や映像のノイズ除去は黎明期から積極的に研究されてきたトピックであり，今日においてもなお重要な課題の1つとして扱われている．

一方，ノイズをただ排除するのではなく，ノイズの特性に着目し，積極的に利用することを目指す研究が近年注目されてきている．微小信号とノイズの振る舞いに関する研究は天文学の分野で興り，その後物理・化学・生物分野においても同様の研究が報告されている．

このような背景を踏まえ，本研究では画像や映像を対象としたノイズ特性の推定とその応用を目指す．先に述べたように，コンピュータビジョンの分野においても画像や映像に含まれるノイズの解析は伝統的な研究分野であるが，その多くはノイズ除去に関する研究であった．しかし，画像や映像の取得の背景には物理的・光学的な特性が隠れている以上，信号処理の過程で含まれるノイズの特性には信号に関する有益な情報が含まれていると考えられる．本研究では

- 物理的ノイズ
- 計測誤差

なる起源を異にする2種類のノイズを対象とし，各ノイズに関してその特性を推定し，応用することを目指す．いずれのノイズも信号や推定値の精度を悪化させる原因ともなるが，その振る舞いを正しくモデル化し，特性を推定することにより，取得した画像や映像の理解に役立てることができることを示す．

まず1つ目のノイズとして，物理的なノイズに着目した応用例を提案する．具体的には固定カメラで撮影された映像を対象とし，ノイズレベル関数と呼ばれる画素値の平均と分散を結ぶ関係式を改ざんの手がかりとして，映像内の改ざんを検出する手法を提案する．提案手法は与えられた映像からノイズ特性を求め，各領域のノイズレベル関数を推定する．ノイズレベル関数はカメラや撮影時のパラメータに依

存するため、ノイズレベル関数の一貫性を評価することによって他の映像から貼り付けられた領域を検出することができる。

はじめに静止シーンの仮定の下で提案手法の基本的な有効性を検証する。静止シーンでは各画素の時間方向の輝度変動はノイズのみに起因するので、時間方向に平均と分散を計算し、各画素のノイズ特性を求めることによりノイズレベル関数を推定することができる。改ざん箇所のノイズ特性はノイズレベル関数の推定において外れ値として振舞うため、外れ値を考慮した推定が必要である。そこで、本研究ではノイズ特性の確率的なモデルを仮定し、最尤推定の枠組みを用いた推定手法を提案する。改ざんの有無を潜在変数とする混合分布モデルを導入し、改ざんの事後確率とノイズレベル関数の推定をEMアルゴリズムを用いて同時に行うことにより、確率的に改ざんの確率を推定する。

続いて移動物体への対応を目指し、提案手法を拡張する。移動物体の場合、実空間において対応する点を追跡しながら分散を計算する必要があり、単純な時間方向の計算によってノイズ特性を求めることができない。しかし、前景の移動物体を画素精度で追跡することは技術的に困難であるので、領域的な追跡によってノイズ特性を求めるアプローチを取る。固定カメラを仮定すると、背景のノイズレベル関数は静止シーンと同様に求めることができる。空間的な分散には陰影・ノイズ・テクスチャの成分が含まれるが、陰影の成分は輝度変動を平面近似することによって除去し、ノイズの成分は背景のノイズレベル関数から推定することができる。よって、残ったテクスチャの分散の特性を検証することにより、ノイズ特性の一貫性を検証することができる。

次に2つ目のノイズに対する応用例として、計測誤差に起因するノイズの特性と伝播に着目することによって推定精度を向上させたアルゴリズムを提案をする。コンピュータビジョンのアルゴリズムの多くは用いる値の一部を既知と仮定し、実際の実験では予め校正により値を求めておく。しかし、現実的には校正にも誤差が混入し、アルゴリズムの推定値にも影響を与えるが、一般に校正誤差の影響はアルゴリズムには組み込まれていない。本研究では照度差ステレオからの形状復元に着目し、照明方向の校正誤差を考慮した形状復元の手法を提案する。照度差ステレオの多くは照明の方向を予め別の手法によって校正しておくが、光源方向の推定にも誤差が生じるため、その影響は推定された形状の法線や復元形状にも伝播する。

ノイズを考慮した形状復元の手法はこれまでに研究されてきているが、これらの手法は勾配に対して正規ノイズが混入した確率モデルを仮定し、最尤推定の枠組みで3次元形状を復元する。しかし、勾配に対するノイズモデルは経験的な仮定であ

り，また勾配は遮蔽輪郭線において発散するため，有限精度の計算機には不適である．一方，本研究では光源方向の校正誤差にもとづく確率モデルを導入し，校正誤差が物体形状の推定に与える影響を検証する．単位ベクトルである光源方向や法線の揺らぎを球面上の正規分布であるフィッシャー分布によってモデル化する．フィッシャーノイズモデルにもとづく最尤推定を用いた形状復元の手法を提案し，校正誤差を考慮することによって形状復元の精度を向上させることを目指す．

以上のように，本研究では異なる特性を持つノイズを対象としながらも，そのノイズ特性を推定することにより，画像・映像の理解や計算精度の向上に役立てることができることを示す．長年に渡ってノイズ除去の手法が研究されている一方で，ノイズ除去には物理的・技術的限界がある以上，本研究で提案するノイズの積極的な活用は，情報処理の新たな方針を与えるという点で重要であると言える．

# 目次

<b>第1章</b>	<b>序論</b>	<b>1</b>
1.1	本研究の背景	1
1.2	本研究の目的	2
1.2.1	応用手法1：物理ノイズの特性にもとづく改ざん検出	3
1.2.2	応用手法2：計測誤差のモデル化による推定精度向上	5
1.3	本論文の構成	6
<b>第2章</b>	<b>ノイズ特性の一貫性にもとづく映像の改ざん検出</b>	<b>7</b>
2.1	研究背景	7
2.2	提案手法の概要	12
2.3	関連研究	16
2.3.1	画像の特性を手がかりとしたカメラの特定手法	17
2.3.2	画像を対象とした改ざん検出手法	19
2.3.3	映像を対象とした改ざん検出手法	20
2.3.4	ノイズ特性の有効利用	21
<b>第3章</b>	<b>静止シーンの改ざん検出</b>	<b>23</b>
3.1	はじめに	23
3.2	CCDカメラのノイズ特性	26
3.2.1	ノイズレベル関数	26
3.2.2	ノイズレベル関数の性質	27
3.3	改ざん検出の流れ	31
3.3.1	基底関数の獲得	31
3.3.2	ノイズ特性の尤度	33
3.3.3	ノイズレベル関数の推定と改ざん検出	36
3.4	評価実験	39
3.4.1	実験室環境による定量的評価	39

3.4.2	実映像による定性的評価 . . . . .	45
3.4.3	映像圧縮の影響 . . . . .	45
3.5	まとめ . . . . .	50
<b>第4章</b>	<b>動的シーンの改ざん検出</b>	<b>53</b>
4.1	はじめに . . . . .	53
4.1.1	画素単位の対応点追跡手法の検討 . . . . .	54
4.1.2	領域的な追跡手法の検討 . . . . .	55
4.1.3	空間領域における分散を手がかりとした改ざん検出 . . . . .	56
4.2	改ざん検出の流れ . . . . .	58
4.2.1	テクスチャ分散と輝度の関係式 . . . . .	58
4.2.2	ノイズの分散成分の推定 . . . . .	59
4.2.3	改ざん検出の指標 . . . . .	62
4.3	評価実験 . . . . .	65
4.3.1	人工映像を用いた検証 . . . . .	65
4.3.2	実映像を用いた検証 . . . . .	68
4.4	まとめ . . . . .	73
<b>第5章</b>	<b>計測誤差を考慮したアルゴリズムの提案</b>	<b>75</b>
5.1	研究背景 . . . . .	75
5.2	提案手法の概要 . . . . .	79
<b>第6章</b>	<b>光源方向の校正誤差を考慮した照度差ステレオからの形状復元</b>	<b>83</b>
6.1	はじめに . . . . .	83
6.2	関連研究 . . . . .	85
6.2.1	照度差ステレオの拡張と誤差解析 . . . . .	85
6.2.2	勾配からの3次元形状復元 . . . . .	86
6.3	照度差ステレオにおける誤差の解析 . . . . .	89
6.3.1	照度差ステレオ . . . . .	89
6.3.2	照度差ステレオの校正誤差 . . . . .	90
6.3.3	校正誤差の法線への影響 . . . . .	94
6.4	フィッシャーノイズモデルを用いた形状復元 . . . . .	96
6.5	実験結果 . . . . .	97
6.5.1	人工画像を用いた検証 . . . . .	97

6.5.2	実画像を用いた検証 . . . . .	105
6.6	まとめ . . . . .	108
<b>第7章</b>	<b>結論</b>	<b>111</b>
7.1	本研究のまとめ . . . . .	111
7.1.1	ノイズ特性の一貫性にもとづく映像の改ざん検出について . . .	111
7.1.2	光源方向の校正誤差を考慮した形状復元について . . . . .	114
7.2	今後の課題 . . . . .	116
7.2.1	改ざん検出について . . . . .	116
7.2.2	形状復元について . . . . .	118
7.3	研究全体の総括 . . . . .	119
	謝辞	121
	参考文献	123
	発表文献	139

# 目次

2.1	改ざんされた画像の例	8
2.2	電子透かしの仕組み	11
2.3	CCD カメラの信号処理過程とノイズ源	13
2.4	ノイズレベル関数の模式図	15
3.1	カメラ応答関数とノイズレベル関数の例	29
3.2	実データにおけるノイズ特性	30
3.3	CCD カメラノイズの生成フロー	31
3.4	主成分分析によって得られた基底関数	34
3.5	改ざん映像におけるノイズ特性生成を表すベイジアンネットワーク	37
3.6	撮影された映像	41
3.7	改ざん領域	42
3.8	CRF(a) と CRF(b) を用いて改ざんした映像の推定結果	43
3.9	CRF(a) と CRF(c) を用いて改ざんした映像の推定結果	43
3.10	図 3.8 と図 3.9 の推定結果の定量的評価	44
3.11	より現実的なシーンを用いた改ざん映像	46
3.12	図 3.11(a) に示した映像の推定結果	47
3.13	図 3.11(b) に示した映像の推定結果	47
3.14	様々なコーデックで圧縮した映像のノイズ特性	49
4.1	オブティカルフローを用いた対応点追跡における問題点	55
4.2	線形カメラ応答関数におけるノイズレベル関数の基底関数	61
4.3	前景テクスチャ	65
4.4	背景	65
4.5	人工映像の背景のノイズレベル関数	66
4.6	輝度とテクスチャ分散の関係	67
4.7	失敗例 1：カメラ応答関数が線形に近い場合	69
4.8	失敗例 2：輝度の分布が偏っている場合	69

4.9	カメラ応答関数と正規化 RMSE . . . . .	70
4.10	実映像の背景のノイズレベル関数 . . . . .	71
4.11	撮影した実映像の指定フレーム . . . . .	71
4.12	実映像のテクスチャ分散 . . . . .	72
5.1	誤差の種類 . . . . .	76
5.2	照度差ステレオのノイズ伝播 . . . . .	79
5.3	勾配に対するノイズ $\Delta p$ が法線方向に与える影響 . . . . .	80
6.1	鏡面球を用いた光源方向の校正 . . . . .	84
6.2	ノイズによって回転する単位ベクトル . . . . .	91
6.3	合成した鏡面球の一例 . . . . .	92
6.4	光源方向の校正誤差のヒストグラムとフィッシャー分布の比較 . . . . .	93
6.5	推定された法線方向の誤差角ヒストグラムとフィッシャー分布の比較 . . . . .	95
6.6	照らされている光源の数と角度誤差の関係 . . . . .	98
6.7	シーン Sphere に対する形状復元結果の比較 . . . . .	100
6.8	シーン Vase に対する形状復元結果の比較 . . . . .	101
6.9	シーン Penny に対する形状復元結果の比較 . . . . .	102
6.10	校正誤差の大きさと形状復元精度の比較 . . . . .	103
6.11	光源の数と形状復元精度の比較 . . . . .	104
6.12	石膏球に対する形状復元結果の比較 . . . . .	106
6.13	石膏像に対する形状復元結果の比較 . . . . .	107

# 表目次

1.1	ノイズの種類による特性の違い . . . . .	2
2.1	特定範囲による特徴 . . . . .	17
3.1	様々なコーデックにおけるビットレートと検出精度 . . . . .	48
6.1	人工画像を用いた形状復元(図 6.7–図 6.9)の性能比較 . . . . .	99
6.2	実画像を用いた形状復元(図 6.12)の性能比較 . . . . .	105

# 第1章 序論

## 1.1 本研究の背景

従来、信号処理の分野において、ノイズは信号を阻害するものであり、除去すべきものであるという認識が広く持たれてきた。このため、電気・電子・機械・情報などの様々な工学分野において、ノイズの発生や外乱による情報の劣化の抑制、信号の冗長化による対ノイズ頑強性の向上、ノイズが混入した情報からの元情報の推定などといったノイズ除去の技術は、古くから研究されてきた伝統的な分野となっている。

一方、ノイズをただ排除するのではなく、ノイズの特性に着目し、積極的に利用することを目指す研究が近年注目されてきている。ノイズが微小信号に与える影響に関する研究は天文学の分野で興った [9]。ノイズに埋れた微小な周期信号と非線形応答が組み合わせられることによって、微小信号が増幅される現象は確率共鳴と呼ばれている。確率共鳴現象は物理・化学・生物分野においても同様の研究が報告されている [19, 31, 41, 111]。特に生物の信号処理や運動制御において観測された確率共鳴現象は、生体内というノイズが激しい環境において高精度かつ柔軟な信号処理を実現しているという点で工学的にも重要な示唆を含んでおり、生物の制御機構を模した温度制御 [121] や力制御 [120] が報告されている。

コンピュータビジョンの分野においても、画像や映像のノイズ除去は黎明期から積極的に研究されてきたトピックであり、今日においてもなお重要な課題の1つとして扱われている。しかし、上で述べた確率共鳴現象に対応する研究はまだ報告されていない。その理由の1つとして、画像や映像はデータの取得時に量子化されるため、微小な信号が量子化誤差に埋没してしまうことが考えられる。デジタルデータを対象とする信号処理では、確率共鳴現象が対象とする系とは異なるノイズの振る舞いがあると考えられる。

しかしながら、デジタル信号処理においてもノイズの振る舞いを解析することには重要な意味がある。代表的なイメージングデバイスの1つである CCD 素子のノイズモデルに関する考察は古くからされており [13, 44, 50, 77, 104]、応答関数などの力

表 1.1: ノイズの種類による特性の違い

	物理ノイズ	計測誤差
ノイズの発生要因	物理現象	系統誤差・偶然誤差
ノイズの一例	熱励起ノイズ	計測器の振動による値の乱れ
誤差分布の仮定方法	現象を解析してモデル化	経験則
代表的な確率分布	正規分布・ポアソン分布など	主に正規分布
ノイズ除去の手法	デバイスの改良・低温化など	複数回の測定など

メラ特性の推定に利用されてきた。また、取得された画像や映像に対してもノイズ特性は有効な情報として役立つことができる。Liuらは画像中のノイズ強度の上限を推定することにより、ノイズ除去フィルタのパラメータ設定や、エッジ検出の精度向上に適用する手法を提案した [69]。Martinらはノイズモデルが動的輪郭線の推定手法であるレベルセット法の推定精度に与える影響について考察した [76]。また、ノイズの統計的な性質を手がかりとして、撮影に用いられたカメラの特定 [71] や画像や映像に施された改ざんを検出する手法 [47, 72] が報告されている。

## 1.2 本研究の目的

1.1節で述べた背景を踏まえ、本研究では画像や映像を対象としたノイズ特性の推定とその応用を目指す。先に述べたように、コンピュータビジョンの分野においても画像や映像に含まれるノイズの解析は伝統的な研究分野であるが、その多くはノイズ除去に関する研究であった。しかし、画像や映像の取得の背景には物理的・光学的な特性が隠れている以上、信号処理の過程で含まれるノイズの特性には信号に関する有益な情報が含まれていると考えられる。

本研究では、起源の異なる2種類のノイズに着目する。1つ目は物理ノイズで、物理的な揺らぎを背景に持つノイズである。2つ目は計測誤差で、観測値の計測の不備や計測装置の有限な精度など、人為的な要因による誤差である。2種類のノイズの特性を表 1.1 にまとめる。

物理ノイズは量子的な揺らぎや熱励起などによって発生する観測量の揺らぎであり、その発生メカニズムは物理的な現象として考察される。画像処理において代表

的な物理ノイズは、電気的な信号増幅(カメラのパラメータでは一般にISOと呼ばれる設定値)を大きくした場合に顕著になる、電子回路内のノイズである。近年では電子デバイスの改良や撮像素子を冷却することにより[123]ノイズの低減化が進んでいるが、一方で省電力化などの理由により信号の強度も小さくなっているため、信号処理におけるノイズの影響は依然として大きい。

一方、計測誤差は人為的・または偶然によって起こる計測の揺らぎである。誤差論によれば、誤差には一定の条件下で一定の誤差が生じる系統誤差と、計測ごとに異なる誤差が発生する偶然誤差に分けることができる。前者は誤差が生じる原因が分かれば補正することができるが、後者は単一の測定では回避することのできない誤差である。例えば、測定の際に測定器が微妙に動揺し、測定値を揺らがせることが考えられるが、この動揺は予測できないため、偶然誤差に分類される。偶然誤差は同一条件において複数回計測を行い、その平均値を用いることによって誤差を減少させることができる。これは、ある正解値に対して誤差が対称に分布するという仮定にもとづく経験的な手法であるが、多くの場合において妥当な結果を示すため、日常的に用いられている手法である。しかし、観測対象によっては必ずしも経験的な手法が最適とは限らないことをここで強調しておく。

以上のように、2種類のノイズは異なる起源や特性を持ち、ノイズ除去の手法も異なるアプローチを取っていることが分かる。しかし、実際にはこれら複数のノイズが複雑に絡み合って観測されるため、日常生活においてこれらの違いを意識することは少ない。一方で、どちらのノイズもランダムな観測値を取るので予測することは難しいが、その発生メカニズムや統計的な性質は有効に用いることができることは前節で述べた背景からも推察される。本研究の目的は、これらの起源の異なるノイズを系統立てて扱い、有効利用する枠組みを提案することにある。続いて、本研究において各ノイズの特性を有効に用いた応用手法について詳しく述べる。

### 1.2.1 応用手法1：物理ノイズの特性にもとづく改ざん検出

物理ノイズの特性に着目した応用例として、本研究では映像中のノイズ特性の一貫性を手がかりとした改ざん検出手法を提案する。画像や映像のデジタル加工技術が進歩し、誰でも簡単に画像・映像の編集や加工が可能になってきている現在、悪意を持って施された改ざんが社会に与える影響は大きい。特に、裁判の証拠としてデジタルデータが用いられる際には、その真正性を保証する手法が必要不可欠である。

本研究では、映像に混入したノイズの一貫性を検証することにより、他の映像か

ら貼り付けられた領域を検出する手法を提案する．カメラで取得された映像には様々なノイズが混入するが，我々はフォトンショットノイズと呼ばれるノイズの性質に着目する．フォトンショットノイズは光子の量子的な揺らぎに起因するノイズであり，ポアソン分布に従うことが知られている．ポアソン分布は平均と分散の間に関連性があり，この関連性を Liu らはノイズレベル関数として定義した [69]．ノイズレベル関数はカメラの種類や撮影時のパラメータに依存するため，与えられた映像内でノイズレベル関数の一貫性を検証することにより，他の映像から貼り付けられた領域を検出することができる．

提案手法はまず静止シーンを仮定して提案手法の基本的な有効性を検証する．カメラと被写体が固定された静止シーンでは，各画素で時間方向の揺らぎはノイズに限定される．よって，各画素で時間方向の平均と分散を計算することにより，ノイズレベル関数を推定することができる．ただし，映像に改ざんが含まれる場合，改ざん箇所のノイズ特性は外れ値として振る舞うため，外れ値を考慮した推定を行う必要がある．

市販されているカメラのカメラ応答関数は一般にノンパラメトリックに定義されており，理論的な解析が難しい．そこで，ノイズレベル関数の基底関数表現を導入し，ノイズレベル関数を次元削減して推定する．カメラ応答関数のデータベースから様々なノイズを合成し，主成分分析を行うことによってノイズレベル関数の基底関数を取得する．そして，与えられた映像から各基底の結合係数を推定することにより，パラメトリックにノイズレベル関数を推定する．また，改ざん検出を確率モデルにもとづいて定式化し，推論的な改ざん検出手法を提案する．ノイズの特性にもとづく確率モデルを導入することにより，各画素の改ざんの事後確率を推定することができる．ここで，各画素の改ざんの確率とノイズレベル関数は同時に推定する必要があるため，EM アルゴリズムを用いた反復的な手法を採用し，推定を行う．

続いて，移動する改ざん箇所への対応を目指す．静止シーンでは各画素の時間的な平均と分散を計算することによってノイズ特性を求めることができたが，移動物体の場合は空間的に対応する点を追跡しながら平均と分散を計算する必要がある．一方，背景のノイズレベル関数は静止シーンと同様に推定できるため，移動物体のノイズレベル関数と比較することにより改ざんを検出する．対応点の探索にはオプティカルフローを用いた追跡手法などがあるが，画素精度で対応点探索は技術的に困難である．そこで，本研究では領域的な追跡によって移動物体のノイズ特性を推定するアプローチを取る．空間的に広がりのある領域の分散は，時間的な分散とは異なり陰影・テクスチャ・ノイズの影響が混在するため，ノイズ特性を評価するた

めには分散の成分を分離する必要がある．このうち陰影は区分的に平面を当てはめることによって除去し，背景のノイズレベル関数を用いてテクスチャとノイズの成分を分離する．改ざんの手がかりとしては，領域内の輝度とテクスチャ分散の間にある関係性を利用し，改ざんの指標を定式化する．

提案手法の評価には，ノイズモデルを模して合成した映像の他，実験室環境や屋内外で撮影された映像を用いる．提案手法は見た目には検出できない改ざんをノイズ特性を手がかりとすることで検出できることを示す．

## 1.2.2 応用手法2：計測誤差のモデル化による推定精度向上

本論文の後半では，計測誤差を対象とし，その伝播に着目してアルゴリズムの推定精度を向上させる枠組みを提案する．コンピュータビジョンのアルゴリズムの多くは，計算に用いる値の一部を既知であると仮定し，実際の実験においては他の手法を用いて校正し，推定を行う．しかし，現実的には校正にも誤差が含まれ，この誤差が推定にも影響を及ぼすはずであるが，多くのアルゴリズムは校正の誤差を考慮していない．校正の誤差を減らすために校正環境や実験装置を工夫することも重要であるが，高精度な計測装置の導入コストは大きく，また完全に除去できない偶発的な誤差は防ぐことができない．よって，校正における誤差を補償し，推定精度を高めることは重要であるといえる．このため，本研究では校正における誤差のモデルを導入し，校正誤差を考慮したアルゴリズム拡張の枠組みを提案する．

具体的な事例として，本研究では照度差ステレオの校正誤差について言及する．照度差ステレオはカメラと対象物体を固定し，光源の方向を変えながら複数枚の画像を撮影することにより，対象物体の形状（法線）を推定する手法である．一般の照度差ステレオでは光源の方向は予め校正によって求めておく．光源方向の校正には鏡面球を用いることが多い．先に述べたように，校正には誤差が混入し，法線や形状の推定にも影響を及ぼす．そこで，光源方向の校正誤差をモデル化し，誤差の振る舞いを推定することにより，形状復元の精度を向上させることを目指す．

誤差を考慮した形状復元の手法はこれまでも研究されてきている．対象物体の3次元形状は，照度差ステレオによって推定された法線ベクトルから勾配を計算し，積分することによって復元される．既存手法は勾配に対して正規分布するノイズが混入すると仮定し，最尤推定の枠組みで3次元形状を推定する．しかし，勾配に対する正規ノイズの仮定は経験的な手法であり，実験において混入するノイズの特性にもとづくものではない．

一方，本研究では光源方向の校正誤差を起源とする法線ベクトルのノイズモデルを導入し，最尤推定によって形状を復元する．光源方向や物体の法線ベクトルは，3次元の単位ベクトルで表すことができる．法線ベクトルは校正誤差によって真の方向からずれて観測されるが，単位ベクトルの条件により  $x, y, z$  の各成分は独立に変動せず，単位球の表面を回転するようにノイズが混入する．そこで，本研究では3次元の単位ベクトルに対するノイズモデルとして，フィッシャー分布にもとづく回転的なノイズモデルを仮定する．

まず，鏡面球による校正を模して人工画像を合成し，校正誤差のヒストグラムから光源方向の校正誤差の分布を検証する．また，同様に既知の形状を用いた照度差ステレオのシミュレーションにおいて光源方向の校正誤差を付加し，法線ベクトルの推定誤差分布を検証する．数値シミュレーションによる検証から光源方向の校正誤差と照度差ステレオによる法線ベクトルの推定誤差がフィッシャー分布することを確認した後，フィッシャーノイズモデルにもとづく形状復元の手法を提案する．既知の形状を用いた数値シミュレーションと実画像を用いた実験を通じて提案手法の有効性を定量的・定性的に評価する．

### 1.3 本論文の構成

以下，本論文は次の6つの章から構成される．第2章ではノイズ特性の一貫性にもとづく映像の改ざん検出に関して，研究の背景と先行研究を述べる．第3章では静止シーンの改ざん映像を仮定し，ノイズレベル関数の一貫性を手がかりとした改ざん検出の枠組みを提案・検証する．第4章において移動物体への拡張について触れる．一方，第5章ではコンピュータビジョンにおける計測誤差の取り扱いとアルゴリズムの推定誤差について概要を述べ，具体例として本研究が提案する校正誤差を考慮した照度差ステレオからの3次元形状復元について第6章で述べる．まず光源方向の校正誤差および照度差ステレオによる法線ベクトルの推定誤差の誤差分布を検証し，フィッシャー分布にもとづく推定誤差モデルについて述べる．最後に第7章で本研究の結論を述べる．

## 第2章 ノイズ特性の一貫性にもとづく 映像の改ざん検出

### 2.1 研究背景

近年のデジタル画像・映像編集技術の発達により，個人でも簡単に画像・映像の編集が行えるようになってきている．かつては画像・映像の編集は専門家が専用機器を用いて行うことが多かったが，現在ではパーソナルコンピュータ用のアプリケーションソフトが開発され，個人でも編集が手軽に行えるようになってきている．また，データサイズの大きい映像を扱うことができる，高性能なコンピュータが手軽に手に入れることができるようになったことも大きな要因の1つである．加えて，インターネット上で大量の画像や映像が共有されるようになり，自分好みの画像・映像ソースが手軽に手に入れられるようになった他，デジタル(ビデオ)カメラの普及により，自分で編集材料を撮影することも一般的となっている．

しかし，画像・映像編集の平易化は，一方で改ざんの危険性を高めている．画像や映像が悪意を持って改ざんされていた場合，その影響は非常に大きい．改ざん画像の例として，Los Angeles Times に掲載された画像を図 2.1 に示す．図 2.1(a) は図 2.1(b) と図 2.1(c) の一部をつなぎ合わせて作られた画像である．改ざん画像は兵士が市民に銃口を突きつけているセンセーショナルな場面を捉えたように見えるが，実際に撮影された写真とは意味が全く異なっている．特に被写体が政治的に重要な意味をもつ画像では，改ざんの影響は非常に大きい．

改ざんの脅威は政治的な場面のみならず，我々の身近にも存在する．例えば，デジタルカメラで撮影された画像や映像が裁判の証拠として採用される場合，改ざんの有無は証拠の信用性を左右する重要な要素である．デジタルカメラが手軽に外に持ち出せるようになり，個人で撮影した画像や映像が裁判における決定的な証拠として扱われることが容易に想定される．したがって，提出されたデジタルデータの真正性評価が重要視されてきている．アメリカ合衆国では連邦民事訴訟規則第 34 条が 2006 年に改定され，電子的に保存されたデータの取り扱いについて以下のように



(a) 改ざんされた画像



(b) 元画像 1



(c) 元画像 2

図 2.1: 改ざんされた画像の例 (文献 [29] より抜粋)

規定している。

*Rule 34. Producing Documents, Electronically Stored Information, and Tangible Things, or Entering onto Land, for Inspection and Other Purposes*

第34条 検査などを目的とした文書，電子的に保存された情報および有形物の提出または私有地への進入

*(a) In General.*

*A party may serve on any other party a request within the scope of Rule 26(b):*

(a) 総則

原告(被告)は，被告(原告)に対し，規則26条(b)項の範囲内において次に定める要求を行うことができる。

*(1) to produce and permit the requesting party or its representative to inspect, copy, test, or sample the following items in the responding party's possession, custody, or control:*

(1) 要求側またはその代表者に対し，回答側の所有，保管または管理下にある以下の項目を提出し，それらの検査，複製，調査，標本の抽出の許可を与えること。

*(A) any designated documents or electronically stored information – including writings, drawings, graphs, charts, photographs, sound recordings, images, and other data or data compilations – stored in any medium from which information can be obtained either directly or, if necessary, after translation by the responding party into a reasonably usable form; or*

(A) (文字記録，絵画，図表，グラフ，写真，音声記録，画像およびその他のデータやデータの編纂物を含む) 指定されたすべての書類や電子的に保存された情報。これらの情報は直接取得可能，または必要に応じて回答側が適当な形式に変換したものに記録されていること。

(後略)

また，American Law Reports ではデジタルデータの証拠物に対する包括的な根拠を立証するための方法の一覧がまとめられている．証拠データの信頼性を保証するためには，提出者は以下の事項について立証する必要があるとされている．

- *Reliability of the computer equipment*  
コンピュータ装置の信頼性
- *Manner in which the basic data was initially entered*  
基本データが最初に記録された方法
- *Measures taken to insure the accuracy of the data as entered*  
データが記録された時の精度を保証する尺度
- *Method of storing the data and the precautions taken to prevent its loss*  
データの保管方法や紛失予防策
- *Reliability of the computer programs used to process the data*  
データ処理に用いたコンピュータプログラムの信頼性
- *Measures taken to verify the accuracy of the program*  
プログラムの精度を証明するための尺度

上の2つの実例に示されるように，裁判における証拠として提出されたデジタルデータに対しては，その真正性を検証することの重要性が認められており，当事者同士でデータを検証することを認め，検証の指標が提案されている．

デジタルデータの真正性を保証する手法としては，古くから電子透かしが多く用いられてきた [21, 57, 66]．電子透かしは著作権情報などをデータに埋め込む手法であり，違法コピーなどへの対応策として用いられている．一般的な電子透かしの付加方法を図 2.2 に示す．ある画像が与えられたとき，画像情報と電子透かしから計算される情報を元の画像に付加することにより，新しい画像を得ることができる．また，このとき公開鍵または秘密鍵を同時に付加することもある．画像を検証する際には，付加とは逆に電子透かしの抽出を行う．画像の加工などにより電子透かしが

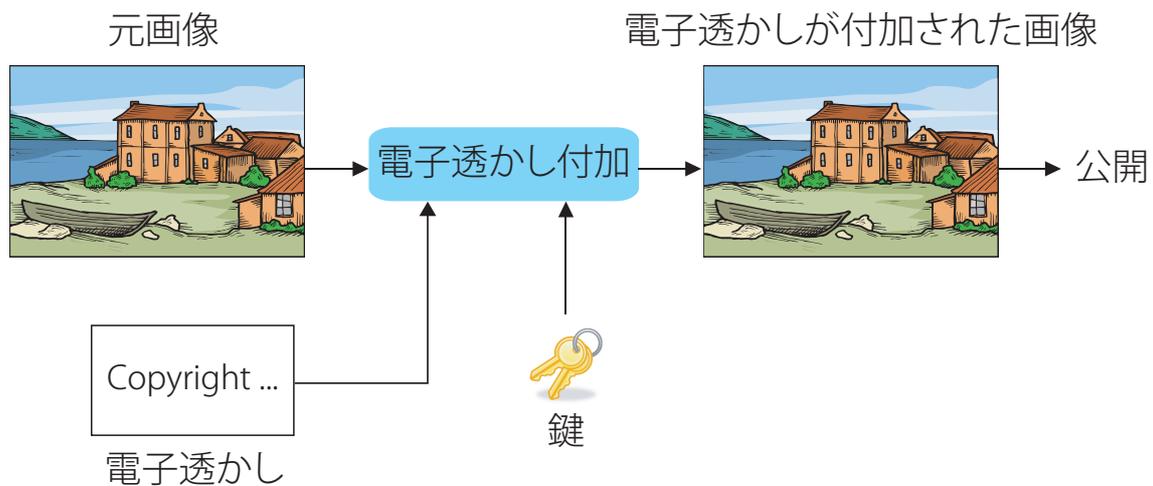


図 2.2: 電子透かしの仕組み

変化していることが考えられるため，抽出された電子透かしと対象となる電子透かしの類似度を計算することにより，画像の真正性を評価する．

電子透かしは可視性のものと不可視性のものがあり，用途に応じて使い分けられる．可視性のものは主に著作権情報を明示的に示すことが目的であり，ロゴやトレードマークが画像に埋め込まれることが多い．情報の埋め込み処理が速く，電子透かしが埋め込まれていることが直接確認できる点で有効であるが，一方で耐攻撃性が低く，また画像の見えを変化させてしまう欠点も持つ．不可視性の電子透かしは商用目的の画像に対して埋め込まれることが多い．画像の品質を劣化させずに情報を埋め込める点が特長であるが，多くの情報を埋め込もうとすると画像の劣化は避けられない．

電子透かしは現在最も一般的な著作権保護技術の1つであるが，その限界として，電子透かしはデータが公開される前に埋め込まれる必要があり，電子透かしを付加する前に公開されたデータに関しては真正性を評価できない点が挙げられる．これは，電子透かしの埋め込まれていない元画像が手に入れば，任意の電子透かしを埋め込んだ画像が作成できるためである．よって，商用目的ではなく個人的に撮影された画像や映像の真正性が争点となった場合，そのようなデータには電子透かしを埋め込むことは稀であるので，電子透かしによる真正性評価が有効な場面は少ないと言える．また，“電子透かしがデータから簡単に取り除かれたり，別の電子透かし

が付加されることはできない”という仮定も明確ではない。情報セキュリティの分野では暗号化技術において攻撃に対する頑強性を評価することが多いが、電子透かしにおいてはこの検証は未だ不十分と言える。例えば、文献 [22] では音楽データを対象とした電子透かしの対攻撃耐性について考察している。

電子透かしの限界である、事前処理の必要性という問題に対処するため、与えられたデータのみを使って画像や映像内の改ざんを検出する手法が提案されてきている [105]。改ざんを検出するための大きな枠組みとしては、データ内の相関を計算し、不自然に相関が高い箇所を検出するもの [38, 47, 108] と、データ内で一貫すべき特徴量を手がかりとするもの [48, 51, 52, 68, 72, 80, 86, 109, 117] に分類される。

画像を対象とした特徴量を手がかりとした手法としては、カメラ応答関数 [48, 68]、光源分布 [51]、収差 [52]、CCD の欠陥パターン [72]、高次の統計量 [80]、画像のリサンプリングが生じさせる画素間の相関 [86]、画像のぼけ [109]、JPEG の量子化テーブル [117] などを用いた手法が提案されてきている。また、映像を対象とする手法では、デインターレーシングパラメータ [107] やノイズ強度 [47] の一貫性を検証する手法が提案されている。

先行研究に共通する問題点として、画像や映像に含まれるノイズに対する頑強性が挙げられる。ノイズは画像や映像から計算される特徴量の推定精度を悪化させるため、改ざんの検出精度もノイズに大きく左右されることになる。特に改ざん検出では改ざん領域と背景領域の微妙な差を区別しなければならないため、ノイズの影響はより深刻になる。個人的に撮影された画像や映像では、商用目的のものとは異なり、画質が悪くノイズが大きいことも想定されるため、対ノイズ耐性は改ざん検出において重要な指標であると言える。

## 2.2 提案手法の概要

一方で、本研究では映像に混入したノイズの特性の一貫性を検証することにより、映像中の改ざん領域を検出する手法を提案する。現在普及しているデジタルカメラでは、撮像素子として CCD (Charge Coupled Device) または CMOS (Complementary

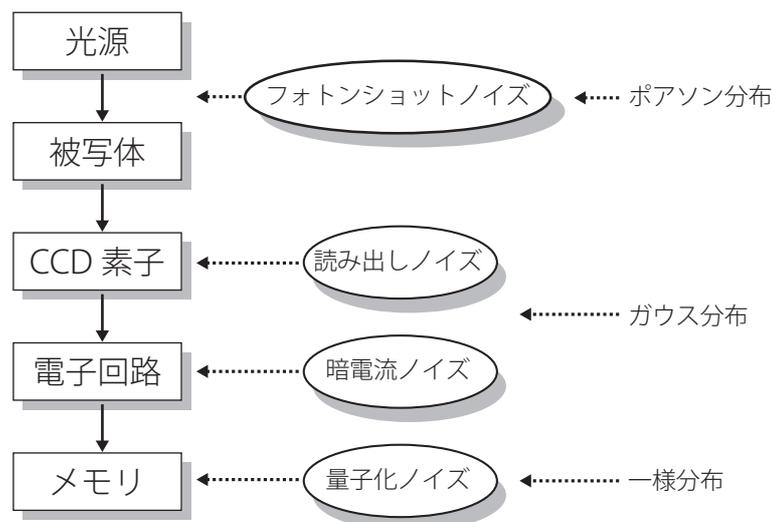


図 2.3: CCD カメラの信号処理過程とノイズ源

Metal Oxide Semiconductor) が用いられている．本研究では CCD 素子を対象としてノイズ特性の解析を行う．CCD 素子は入射した光子を電気信号に変換し，様々な信号処理を行った後，最終的にビット列にしてメモリに格納する．図 2.3 に CCD カメラにおける信号処理の過程と各過程で信号に混入するノイズ源のブロック図を示す．CCD カメラのノイズモデルは古くから研究されており [44, 96, 104]，信号処理の過程においては，フォトンショットノイズ，読み出しノイズ，暗電流ノイズ，量子化ノイズなどのノイズが信号に混入する．複数のノイズ源の中から，我々はフォトンショットノイズと呼ばれるノイズの特性に着目する．これは，フォトンショットノイズが次に述べるような特徴を持つためである．

1. 著しく暗い環境を除き，支配的なノイズ源である．
2. 画素値の平均値と分散の間に相関がある．

1 つ目の理由は，一般的な映像に対して提案手法が適用できる利点となる．また，2 つ目の理由は，改ざんの手がかりとなるノイズ特性を導出するための重要な特徴である．

フォトンショットノイズは，光子の量子的な性質に由来するノイズである．CCD 素子 [106] は露光時間の間に入射した光子の数に応じて電荷を発生する素子である．光

子の入射数はその量子的な性質により時間的に揺らぎを持ち、その分布はポアソン分布で表される。ポアソン分布は平均と分散の間に相関があるため、フォトンショットノイズも平均と分散の間に関連性を持つことになる。この平均と分散の関連性はカメラの特性や撮影時のパラメータに固有の特性であるため、フレーム内での一貫性を検証することによって改ざんを検出することができる。

画像の平均と分散の関連性については、Liuらがノイズレベル関数 (noise level function: NLF) [69] として定義した。Liuらはカメラ応答関数のデータベースから大量のカメラ応答関数を取得し、人工的なノイズを合成してノイズレベル関数を生成した。そして、主成分分析 [83] を用いてノイズレベル関数の基底関数を計算し、一般のノイズレベル関数を基底関数の線形和で表現した。Liuらは1枚の画像からノイズレベル関数を推定することにより、画像中のノイズの上限値を推定した。また、ノイズの上限値を推定することにより、ノイズ除去フィルタのパラメータを適応的に決定したり、エッジ検出の精度向上への利用方法を提案した。

一方で、我々はノイズレベル関数を改ざんの手がかりとして用いることを目指す。先に述べたように、映像に含まれるフォトンショットノイズは、平均値と分散の間に関連性がある。この関連性はまさにLiuらが提案したノイズレベル関数そのものである。したがって、与えられた映像からノイズレベル関数を推定し、その一貫性を検証することにより、改ざん箇所の検出を目指す。図2.4に改ざん映像のノイズ特性とノイズレベル関数の模式図を示す。1つの点が1つの画素のノイズ特性に対応する。また、青線が元映像のノイズレベル関数、赤線が改ざん領域のノイズレベル関数を表す。我々は与えられた映像からノイズ特性の点を計算し、続いてノイズレベル関数を推定、そしてその一貫性を評価し、改ざん箇所を特定するというアプローチを提案する。

本研究では、初めに特定の条件下で撮影された映像に対する改ざん検出手法を提案し、段階的に条件を緩和し対象映像を一般化する。まず、第3章において、被写体とカメラが固定された静止シーンを仮定し、ノイズ特性を手がかりとした改ざん検出手法の基本的な有効性を検証する。静止シーンにおいては、各画素の時間的な

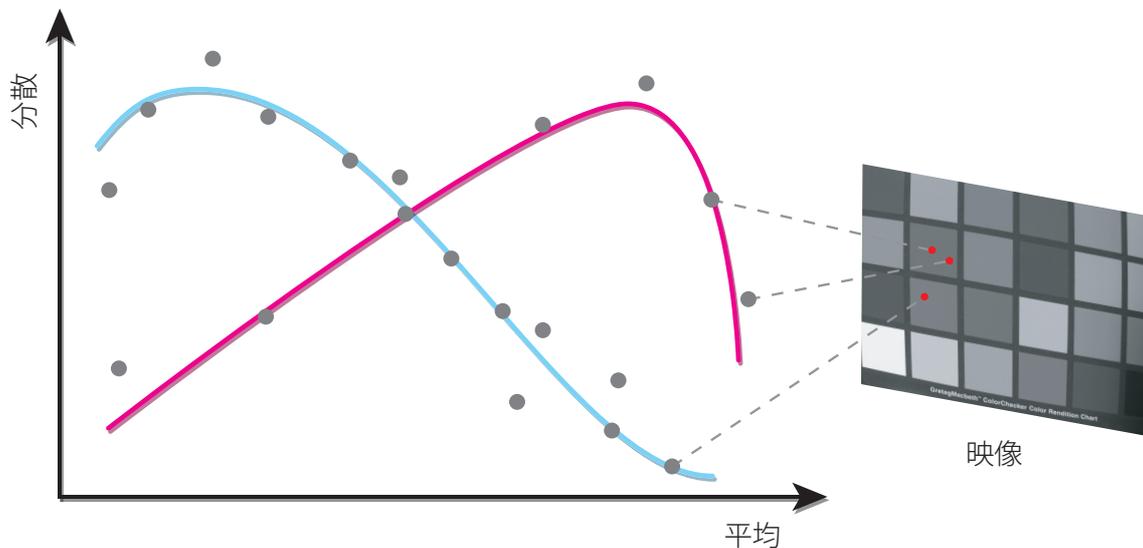


図 2.4: ノイズレベル関数の模式図

揺らぎはノイズのみに起因するため、テクスチャや陰影による輝度変化の影響を考慮することなくノイズ強度を求めることができる。したがって、各画素の時間的な平均と分散を計算し、計算されたノイズ特性の点からノイズレベル関数を推定することにより、その一貫性を検証する。しかし、一般にカメラ応答関数はノンパラメトリックに定義されているため、ノイズレベル関数もノンパラメトリックに表現され、解析的に推定することが難しい。そこで、Liuらの手法を踏襲し、ノイズレベル関数を基底関数を用いた線形結合で表現することにより、ノイズレベル関数の推定を行う。

また、各画素の改ざんの判定に確率的な推論の枠組みを導入する。有限のフレーム数から計算される分散は真の分散からずれて観測され、この揺らぎはカイ二乗分布 [82] に従う。そこで、計算されたノイズ特性の尤度をカイ二乗分布を用いて評価し、改ざんの尤度を求める。また、改ざんによるノイズ特性の違いは、潜在変数を用いて混合分布モデルとして表すことができる。各画素がどのノイズ特性に応じて振舞うかを表す事後確率と、各ノイズレベル関数のパラメータは同時に推定する必要があるため、我々はEMアルゴリズム [23] を用いてこれらの値を同時に推定する手法を提案する。

続いて第4章において改ざんされた移動物体を検出する手法を提案・検証する。移動物体の場合は静止シーンのように時間方向の平均と分散を計算することによりノイズ特性を求めることができない。画素単位に対応点追跡は技術的に困難であるため、本研究では領域的な追跡を用いて時間的な平均と分散の関係を求める。ただし、空間的に広がりのある領域の分散には、陰影・テクスチャ・ノイズの成分が混在するため、これらの分散成分を適切に分離する必要がある。陰影は先行研究で用いられた手法に則り、領域の輝度変化を平面近似することによって除去する。また、背景が静止していると仮定すると、背景のノイズレベル関数が推定できるため、注目領域のノイズ成分をノイズレベル関数から推定することができる。

分散の成分が分離できたら、改ざんの手がかりとして輝度とテクスチャ分散の関係性を利用する。線形のカメラ応答の仮定の下で、テクスチャの分散は輝度の2乗の関数となる。よって、異なる注目領域の明るさが時間的に変化すると仮定すると、輝度とテクスチャ分散の関係が2次曲線で記述できるか評価することができる。テクスチャ分散が2次曲線から乖離したときは、ノイズ成分を分離するために用いた背景のノイズレベル関数が、注目領域のノイズ特性と合致していないと結論づけることができる。したがって、2次曲線に対する適合度を評価することにより改ざんの判定をすることができる。

## 2.3 関連研究

本節では画像と映像を対象とし、データの真正性を評価する手法を紹介する。データの真正性評価は、大きく撮影されたカメラの特定とデータ内の改ざん検出に分類される。また、本研究が手がかりとして用いるノイズの特性を有効に利用した先行研究についても言及する。

表 2.1: 特定範囲による特徴

特定範囲	メーカーや型番	撮影したカメラ
同一製品で撮影した画像の区別	×	
事前情報の利用		×

### 2.3.1 画像の特性を手がかりとしたカメラの特定手法

本研究が提案する改ざん検出とはアプローチが異なるが、撮影に用いたカメラを画像から特定することも画像の真正性評価には重要である。カメラの特定は、その特定範囲によって次の2種類に区別することができる。

- カメラのメーカーや型番の特定
- 撮影した固有のカメラの特定

それぞれの特定範囲の特徴を表 2.1 に示す。

カメラのメーカーや型番まで特定する手法では、製品に固有の情報を手がかりとして特定を行う。このアプローチの特徴として、事前に製品ごとの特徴量をデータとして保持しておくことで、与えられた画像だけからメーカーや型番を特定することができる点がある。一方、共通する仕様の製品間ではカメラの識別ができないという限界がある。

CCD 素子が1組のカラーカメラでは、カラーフィルタを用いて空間解像度の低い画像を取得し、中間の画素を補間によって復元する(デモザイキング)。デモザイキングにおいて用いる補間アルゴリズムのパラメータはカメラモデルに固有であるため、Bayramらはモデルごとのパラメータを取得しておき、与えられた画像から一致するカメラモデルを推定する手法を提案した [7, 60]。Huangらもカラーフィルタからのデモザイキング処理に着目し、主成分分析を用いた画像内の相関の除去とニューラルネットを用いた識別モデルにより、カメラモデルの特定と画像内での異なる特性を持つ領域の検出を行う手法を提案した [49]。また、Gallagherらは人工画像ではデモザイキングが発生しないことを手がかりとし、デモザイキングの処理痕を検出す

ることによって写実的な人工画像と実画像を区別する手法を提案した [39] .

また, Mullikin らはカメラの信号処理における様々な特徴量を手がかりとしたカメラモデルの特定手法を提案した [79] . この手法は, カメラ応答関数の線形性, SN 比, 感度, 暗電流ノイズ, 空間周波数特性といった多角的な情報を用いることにより, 6 種類のカメラモデルを識別することができる .

Farid は JPEG 画像で用いられる量子化テーブルを手がかりとしてカメラの特定を行う手法を提案した [28] . JPEG 画像では, 元画像の周波数成分を決められた量子化テーブルの値で除算することによりデータ圧縮を行う . 量子化テーブルはメーカーやカメラモデルに固有であるので, 画像からテーブルを推定することにより, カメラモデルの特定を行うことができる . 204 種類のカメラの量子化テーブルを比較したところ, 62 種類のカメラは固有のテーブルを持ち, 残りのカメラは 2 種類から 28 種類の共通のテーブルを持つカメラが存在するという結果を得た . また, テーブルを共有するカメラの種類の期待値は 1.43 台であった .

一方, カメラ固有の情報を用いるアプローチでは, 同じ製品であっても区別することができる点が特徴である . しかし, カメラを特定するためには, 検証するカメラ実機, もしくはそのカメラで撮影された画像が必要になり, 事前情報を用いることはできない . カメラ固有の情報としては, CCD 画素の製造過程における欠陥を手がかりとするものが提案されている .

Lukáš らは固定パターンノイズと呼ばれる CCD 素子の特性の空間的な揺らぎを手がかりとしてカメラを特定する手法を提案した [71] . CCD 素子は製造過程における条件の揺らぎにより, 暗電流ノイズの大きさに空間的な揺らぎが生じる . この揺らぎは空間的にはランダムであるが時間的には一定であるため, 固定パターンノイズと呼ばれる . 固定パターンノイズは同じカメラモデルであっても異なるパターンを形成するため, カメラの指紋と見なすことができる . Lukáš らは与えられた画像から固定パターンノイズを推定し, 参照パターンと比較することによってカメラの特定を行った .

Bloy は Lukáš らの手法をより実用的に拡張し、実験環境で撮影されていない大量の参照画像から参照パターンを作成する手法を提案した [11]。また、カメラの特定がされていない画像セットに対して、クラスタリングにより同一のカメラから撮影された画像を選択し、参照パターンを作成する手法も同時に提案した。

### 2.3.2 画像を対象とした改ざん検出手法

画像を対象とした改ざんの検出手法は大きく2種類に分類することができる。即ち、(1) 同一画像内の複製と(2) 他の画像からの貼り付けを対象とした手法である。

Fridrich らは同一画像内の複製を対象とし、不自然に高い相関を持つ領域を検出することで改ざん箇所を特定する手法を提案した [38]。

以降の研究は他の映像から貼り付けられた領域を検出する手法を提案したものである。Lin らはカメラ応答関数の一貫性を手がかりとして改ざん検出を試みた [68]。被写体の輪郭線上において、カメラに入射する光は前景と背景の色情報の線形和で表わされる。カメラ応答関数はこの線形性を歪ませて画素値に変換するため、観測された画素値の線形性を復元するような逆関数を推定することにより、カメラ応答関数を求めることができる。カメラ応答関数の一貫性を検証することにより、他の画像から貼り付けられた領域を検出することができる。また、Hsu らはカメラ応答関数から計算される不変量を手がかりとして改ざん検出を行う手法を提案した [48]。

Johnson らは被写体の輝度と法線ベクトルから光源分布を推定し、その一貫性を検証することによって他の画像から貼り付けられた物体を検出する手法を提案した [51, 53]。また、同じく Johnson らは収差(光の波長によって屈折角が異なるために色ごとに焦点位置がずれる現象)に着目し、ずれのパラメータの画像内の一貫性を評価することによって改ざん箇所を検出する手法を提案した [52]。

Lukáš らはカメラ特定で用いた固定パターンノイズを改ざん検出にも応用した [72]。参照パターンと比較する点はカメラ特定と同じだが、画像内で参照パターンと異なるノイズパターンを持つ領域を特定することで、他のカメラから貼りつけられた領域を検出する。

あるデータ列(音声・画像など)に対して、閾値処理などの非線形変換を施すと、bicoherence と呼ばれる高次の統計量が不自然に高い値を示すことが先行研究により示されている [27]。Hsu らは与えられた画像に対して bicoherence を求めることにより、異なるソースの画像が貼り付けられてできた改ざん箇所を検出する手法を提案した [48]。

Popescu らは画像のリサンプリングにおける相関を検証することにより改ざん検出を行う手法を提案した [86]。複数の画像を合成して改ざん画像を作成する場合、元の画像のサイズを変更することが多い。画像サイズの変更にはリサンプリングが伴い、その過程で補間を行うため、リサンプリングされた画像では画素間に相関が生じる。この相関は画像内で一定であるはずなので、一貫性を検証することにより異なる処理でリサンプリングされた領域を検出することができる。

Wang らは画像の焦点ぼけを手がかりとする改ざん検出手法を提案した [109]。画像の焦点が合っていない場合、被写体からカメラに入射した光は CCD 素子上で 1 点に収束せず、円状に分布して観測される。レンズにおける屈折の式から、焦点ぼけの半径と被写体までの距離には関係があるため、画像内でこの関係が破綻している箇所を改ざんとして検出することができる。

Ye らは JPEG による画像圧縮に使われる量子化テーブルの一貫性を利用した検出手法を提案した [117]。もし改ざんが施されて複数の量子化テーブルにより圧縮された場合、blocking artifact measure と呼ばれる改ざんの指標が大きく計算される。この指標が不自然に高いブロックを改ざんとして検出することができる。

### 2.3.3 映像を対象とした改ざん検出手法

画像を対象とした改ざん検出手法に比べ、映像を対象とした手法はまだ少ない。映像を対象とした場合も、画像と同様に同一映像内の複製と他の映像からの貼り付けによって作られた改ざんを検出する手法に分けることができる。

Wang らは映像内の相関を計算することにより、同一映像内からの複製箇所を検出する手法を提案した [108]。また、ある時刻のフレーム全体が異なる時刻に複製され

た場合でも、フレーム間の相関を計算することによって複製を検出することができる。この手法は同一映像内の複製を対象としているため、他の映像から貼り付けられた領域は検出することができない。

Wangらは他の映像から貼り付けられた領域に対しても、デインターレーシングに用いられるパラメータの一貫性を検証することによる改ざん検出の手法も提案している [107]。デインターレーシングとは、縦方向の解像度が半分しかないインターレース方式で記録された映像を、補間などを用いてフル解像度の映像に変換する技術のことである。補間に用いられるパラメータは変換手法に固有であるため、映像内での一貫性を手がかりとして改ざん箇所を検出することができる。これら手法は他の映像から貼り付けられた領域を検出することができるが、映像の記録方式に依存する手法であるため、その適用範囲は限定される。

映像内のノイズを手がかりとした改ざん検出はHsuらによって提案されている [47]。フレームをブロックに分割し、ノイズ除去によってノイズ成分のみを抽出すると、同一映像内から複製されたブロックはノイズ成分が不自然に高い相関を持つ。一方、異なるテクスチャを持つ他の映像から貼り付けられた領域との間には不自然に低い相関が計算によって求まる。ブロック間の相関によって同一映像内の複製と異なる映像からの貼り付けを検出することができるが、ノイズの特性を求めるためのノイズ除去に検出の性能が大きく依存する。特に元映像と改ざん領域のノイズ強度が大きく異なる場合には正しくノイズを除去できないため、検出性能が低下するという問題がある。

### 2.3.4 ノイズ特性の有効利用

カメラのノイズ特性についてはデジタルカメラの黎明期から数多く研究されてきたが、その研究の多くは撮影された画像や映像からのノイズ除去を目的としたものである。一方、近年になり、画像や映像からノイズをただ除去するのではなく、ノイズの特性を積極的に用いることによってデータの理解などに役立てることを目指した研究が発表されてきている。本研究がノイズレベル関数の定義について引用し

たLiuらの報告 [69] では、1枚の画像からノイズレベル関数を推定することにより、ノイズ除去やエッジ検出に対して適用した事例が紹介されている。

Matsushitaらはノイズ特性を用いてカメラ応答関数を推定する手法を提案した [77]。輝度に混入するノイズはゼロを中心とした対称の分布を示すが、非線形のカメラ応答関数で変換されることにより、観測される画素値のノイズの分布形状は対称から歪む。そこで、観測された画素値の非対称なノイズ分布から逆変換された分布の対称性を評価することにより、カメラ応答関数の逆関数を推定することができる。同様に、Takamatsuらは輝度とノイズ強度の関係性 [102] や、probabilistic intensity similarity と呼ばれるノイズを考慮した画素値の類似度にもとづくカメラ応答関数の推定手法 [101] も提案しており、ノイズ特性を用いてカメラの特性を推定する手法は数多く提案されている。

## 第3章 静止シーンの改ざん検出

### 3.1 はじめに

本章では、ノイズ特性の一貫性を手がかりとした映像内の改ざん箇所検出手法の基本的な有効性の確認を目的するため、静止シーンを対象とした手法の提案と評価を行う。

被写体とカメラが固定された静止シーンの仮定は、監視カメラなどで撮影された映像に固定された物体が貼りつけられてできた改ざん映像などを想定している。静止シーンはノイズを手がかりとした改ざん検出の基本となる対象シーンであり、本研究の提案する、改ざん検出の手がかりとしてのノイズ特性の有効性を検証するための基礎となる。また、第4章では動的なシーンに対する改ざん検出を行うが、その際にも背景内の改ざんの有無を検証するために本手法は重要な役割を果たす。

静止シーンの特徴として、ノイズによる輝度値の揺らぎを簡単に計算できる点が挙げられる。静止シーンでは各画素が観測する被写体の位置が時間的に変化しないため、画素値の時間的な変化は全てノイズによるものと結論づけることができる。したがって、空間的な分散の計算時に問題となる、陰影・テクスチャ・ノイズの分離問題を考慮する必要がなく、各画素について独立に時間方向の分散を計算することにより、ノイズ強度を簡単に求めることができる。

本研究では、各画素で求められた時間的な平均と分散の関係性をノイズレベル関数として記述する。ノイズレベル関数はカメラ応答関数や撮影時のパラメータに依存するため、カメラ撮影時の固有情報と見なすことができる。我々は与えられた映像からこのノイズレベル関数を推定し、その一貫性を検証することによって改ざん箇所を検出することを目指す。

第2章で述べたように、提案手法が着目するフォトンショットノイズは、CCD素子に入射する光子数の時間的な揺らぎに起因するノイズである。光子はカメラ内で電子に変換され、様々な信号処理を経て画素値として観測されるため、フォトンショットノイズの強度と輝度値の関係性を直接観測することはできない。また、信号処理において未知なパラメータが存在するため、解析的にノイズレベル関数の推定することも難しい。したがって、与えられた映像のみからノイズレベル関数を推定することは大変困難である。そこで我々はLiuらの手法 [69] を踏襲し、事前知識としてノイズレベル関数の基底関数を用いる。基底関数の線形結合としてノイズレベル関数を表現することにより、推定されるパラメータの個数を減少させることができる。

また、提案手法は改ざん検出を確率的な枠組みで取り扱い、改ざんの事後確率を推定する。ノイズは確率的な現象であるため、計算されたノイズ特性も揺らぎを持っている。また、ノイズ特性の分布形状は入力映像に大きく依存するため、適応的なパラメータ調整が必要である。したがって改ざんの有無の判定境界を決定論的に取り扱うことは難しいため、確率モデルを用いた推論によって改ざんの事後確率を推定する。計算されるノイズ特性の揺らぎを適切に取り扱うため、我々は統計学の知見にもとづいて改ざんの尤度を定義する。具体的には、有限の標本数に起因する分散の揺らぎがカイ二乗分布 [82] に従うことにもとづき、ノイズレベル関数から示唆される分散と実際に計算された分散の差異をカイ二乗分布を用いて評価する。

ベイズ的な確率の取り扱いを用いることにより、改ざんは潜在変数を用いて混合確率分布モデルとして表現することができる。すなわち、各画素が従う(元となる映像の)ノイズ特性を潜在変数で表すと、各クラスに属する確率が改ざん・非改ざんの確率に対応する。すると、各映像ソースのノイズレベル関数が分かれば各画素が属するクラスの尤度を計算することができる。一方で、ノイズ特性の点が得られたとき、各映像ソースのノイズレベル関数を推定するためには、各点がどちらのノイズレベル関数に従うのかが分かっている必要がある。つまり、各画素が属する映像ソースの確率と各ノイズレベル関数のパラメータは同時に推定する必要がある。そこで我々はEMアルゴリズム [23] を用いて、これら2種類の値を同時に推定する手

法を提案する。

先行研究と比較した提案手法の特長は以下のとおりである。

1. 画素単位の検出が行える
2. 異なる映像ソースから貼りつけられた領域を検出できる
3. 対象映像の記録フォーマットを限定しない
4. 混入するノイズ強度の一貫性を仮定しない

1. は各画素において独立にノイズ特性を計算し、改ざんの確率を推定できることを意味する。これは、先行研究 [47] がブロック単位で改ざんを検出するため、ブロックの取り方に精度が依存し、また改ざん検出の空間分解能が小さいという限界に対する提案手法の利点である。2. は相関を手がかりとする先行研究 [108] に対する利点である。提案手法は映像内の特徴量の一貫性を手がかりとするため、同一ソース内の複製を対象とする先行研究では検出できない、異なる映像ソースから貼りつけられた領域を検出することができる。3. は記録フォーマットの特性を手がかりとする先行研究 [107] に対する利点である。与えられた映像ソースのフォーマットに依存しないという点で、提案手法はより多くの種類の映像に対して適用可能である。最後に、4. は提案手法と同じくノイズの特性を手がかりとする先行研究 [47] に対して提案手法が持つ利点である。先行研究では、背景映像と改ざん領域でノイズ強度が大きく異なっているとき、ノイズ除去が正確に行えず、検出精度が低下するという限界があった。提案手法はノイズ強度の違いを手がかりとするため、先行研究では検出できない改ざんを検出することができる。

以下、3.2 節では CCD カメラのノイズ特性とノイズレベル関数の導入について述べ、次いで 3.3 節で改ざん検出の枠組みについて説明する。そして 3.4 節で実験結果について述べ、最後に 3.5 節で本研究のまとめを行う。

## 3.2 CCDカメラのノイズ特性

提案手法では、映像に含まれるノイズを改ざんの手がかりとして用いる。本節ではまず CCD カメラにおける信号処理について 3.2.1 節で触れ、先行研究において提案されたノイズレベル関数について言及する。また、3.2.2 節ではノイズレベル関数を CCD カメラのノイズモデルから記述し、輝度の平均と分散の関係を導出する。

### 3.2.1 ノイズレベル関数

先に述べたように、CCD に入射する光子の数は量子的な揺らぎのためにノイズとして振る舞い、その分布はポアソン分布で表すことができる。ポアソン分布はその平均と分散に関係性があり、また考える乱数の規模が光子数のように十分大きいとき、正規分布に近似することができる。実際には光子数の平均と分散の関係を直接観測することは不可能であるが、その代わりに観測された画素値の平均と分散を計算することは可能である。提案手法では、映像内における画素値の平均と分散の関係を改ざんの指標として用いる。

ここで、Liu らが画像内のノイズ特性を推定し、エッジ検出などに用いたノイズレベル関数 (NLF) の定式 [69] について触れる。はじめに、本研究では以下の2つの仮定をする。まず1つ目に、映像に含まれるノイズはゼロ平均であり、真の画素値は平均化によって得ることができるものとする。次に、対象とする映像は静止シーンであると仮定する。すると、ある画素の時間的な平均と分散を計算することにより、ノイズ特性を求めることができる。

以上の仮定にもとづき、ノイズレベル関数を定式化する。観測された画素値  $O$  は、ノイズによってランダムに振舞うとする。この画素の時間的な平均を  $\mu$  とすると、ノイズレベル関数  $\tau(\mu)$  は次式によって定義される。

$$\tau(\mu) \equiv \text{Var}[O] = \text{E}[(O - \mu)^2] \quad (3.2.1)$$

文献 [69] 内 (2) 式とは異なり、ノイズレベル関数を標準偏差ではなく分散として定義していることに注意されたい。次節で示すように、ノイズレベル関数は量子効率

やカメラ応答関数 (camera response function: CRF) といったカメラ固有のパラメータの他, 露光時間やゲインといった撮影条件などに依存する。したがって, ノイズレベル関数は撮影条件を記録した固有情報として振る舞うため, その一貫性を検証することによって改ざん箇所を検出することができる。

### 3.2.2 ノイズレベル関数の性質

本節では CCD カメラ内の信号処理過程を考察することにより, 観測される画素値の平均や分散の振る舞いについて言及する。文献 [69] に記述されるように, カメラで撮影されメモリに格納される信号  $O$  は次式

$$O = f(a\{\eta(P + N_{PS}) + N_{DC} + N_R\}) + N_Q \quad (3.2.2)$$

で表される。ここで  $f(\cdot)$  はカメラ応答関数,  $a$  はゲイン,  $\eta$  は CCD 素子の量子効率,  $P$  は露光時間内に CCD 素子に入射した光子数である。また, 信号に混入するノイズ源として, フォトンショットノイズ  $N_{PS}$ , 暗電流ノイズ  $N_{DC}$ , 読み出しノイズ  $N_R$ , 量子化ノイズ  $N_Q$  の4種類を仮定する。ここで (3.2.2) 式において  $f$  に関して  $a\eta P$  の周りで1次のテイラー展開を施すと, 次式

$$O \simeq f(a\eta P) + af'(a\eta P)(\eta N_{PS} + N_{DC} + N_R) + N_Q \quad (3.2.3)$$

を得る。(3.2.3) 式より,  $O$  の平均は

$$\begin{aligned} \mu &= E[O] \\ &\simeq f(a\eta P) + af'(a\eta P) \cdot (\eta E[N_{PS}] + E[N_{DC}] + E[N_R]) + E[N_Q] \\ &= f(a\eta P) \end{aligned} \quad (3.2.4)$$

で計算される。ここで  $E[N_{PS}] = E[N_{DC}] = E[N_R] = E[N_Q] = 0$  と仮定した。(3.2.4) 式は平均を計算することによりノイズを除去できることを示唆している。一方,  $O$  の分散は

$$\sigma^2 \simeq a^2 f'^2(a\eta P)(\eta^2 P + \sigma_{DC}^2 + \sigma_R^2) + \sigma_Q^2 \quad (3.2.5)$$

で得られる．ここで  $\sigma_{\text{DC}}^2$ ,  $\sigma_{\text{R}}^2$ ,  $\sigma_{\text{Q}}^2$  はそれぞれ暗電流ノイズ，読み出しノイズ，量子化ノイズの分散である．(3.2.5) 式の第1項はフォトンショットノイズの分散に対応し，この強度はフォトンショットノイズがポアソン分布に従うことから光子数  $P$  に依存する．

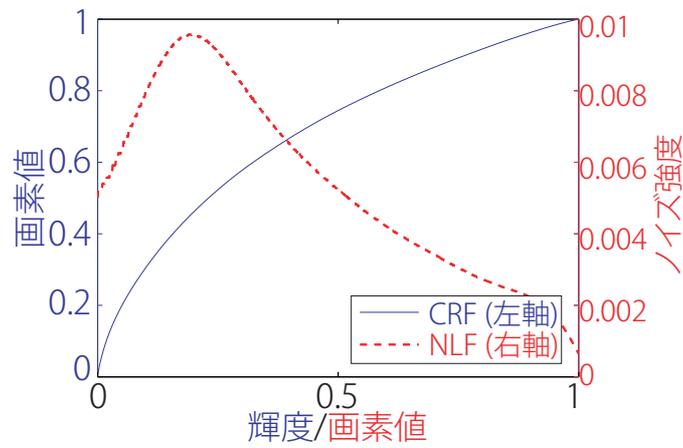
ここで(3.2.4) 式を(3.2.5) 式に代入し， $P$  を消去すると，

$$\sigma^2 = a^2 f'^2(f^{-1}(\mu)) \cdot \left( \frac{\eta}{a} f^{-1}(\mu) + \sigma_{\text{DC}}^2 + \sigma_{\text{R}}^2 \right) + \sigma_{\text{Q}}^2 \quad (3.2.6)$$

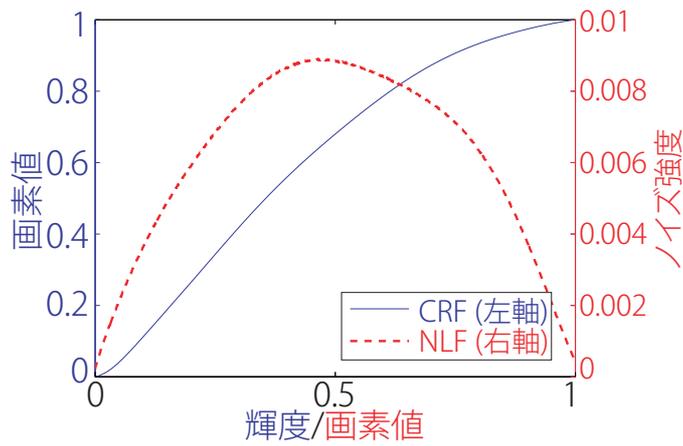
を得る．(3.2.6) 式は，画素値の平均  $\mu$  と分散  $\sigma^2$  を関連付ける式であるので，ノイズレベル関数の定義式である(3.2.1) 式を信号処理の側面から導いた式であると言える．また，(3.2.6) 式から，ノイズレベル関数は特にカメラ応答関数の微分  $f'$  に依存することが分かる．

ここでカメラ応答関数とノイズレベル関数の関係性を何例かの図を用いて示す．図3.1に3種類のカメラ応答関数と，合成ノイズによって得られた対応するノイズレベル関数を示す．青実線がデータベースから選ばれたカメラ応答関数で，赤破線が対応するカメラ応答関数から計算されたノイズレベル関数である．ノイズレベル関数の形状は，カメラ応答関数に大きく影響されることが分かる．また，(3.2.6) 式で示されたように，カメラ応答関数の微分が大きい箇所でノイズレベル関数の値が大きくなっていることが分かる．

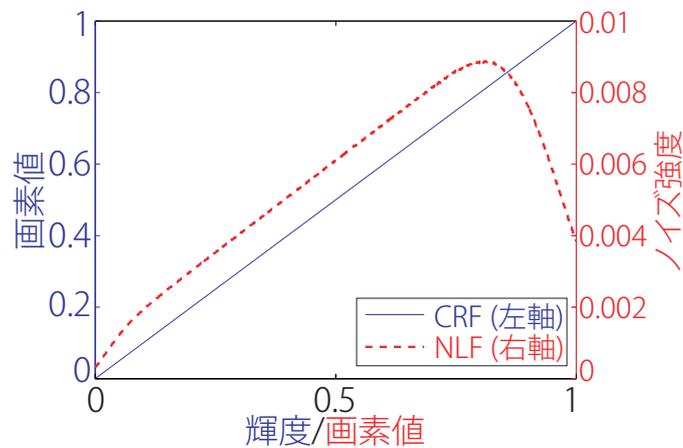
併せて実際にカメラによって撮影された映像から計算されたノイズ特性を図3.2に示す．各映像に用いられたカメラ応答関数は図3.1のものと対応し，以後 CRF(a)，CRF(b)，CRF(c) のラベルで各カメラ応答関数を扱うことにする．実データではノイズ強度が大きくなるほど描画された点が広く分散する．これは有限の標本数で分散を計算している影響であり，この分散の揺らぎに関しては3.3.2節で詳しく議論する．



(a) CRF No.27

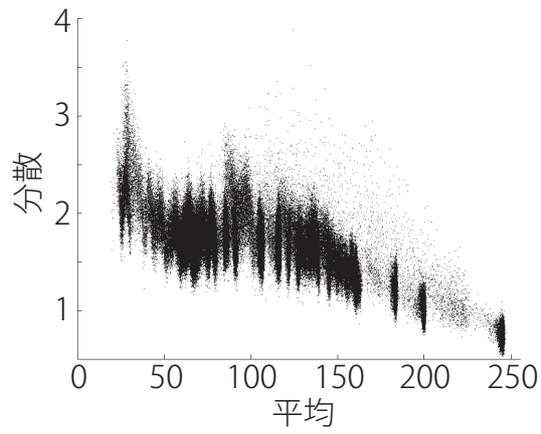


(b) CRF No.53

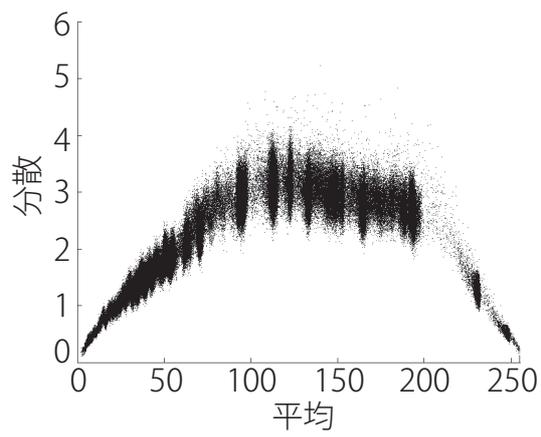


(c) CRF No.164

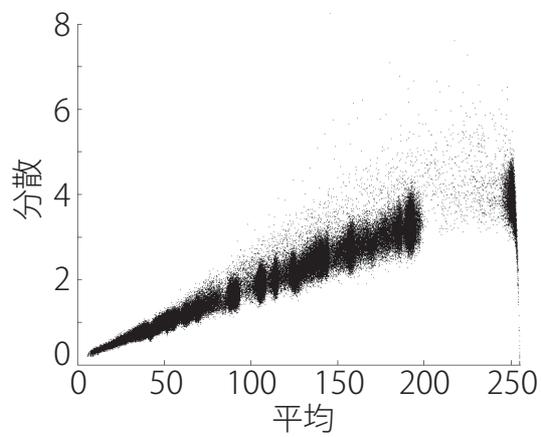
図 3.1: カメラ応答関数とノイズレベル関数の例．青実線と赤破線がカメラ応答関数とノイズレベル関数にそれぞれ対応する．



(a) CRF No.27



(b) CRF No.53



(c) CRF No.164

図 3.2: 実データにおけるノイズ特性

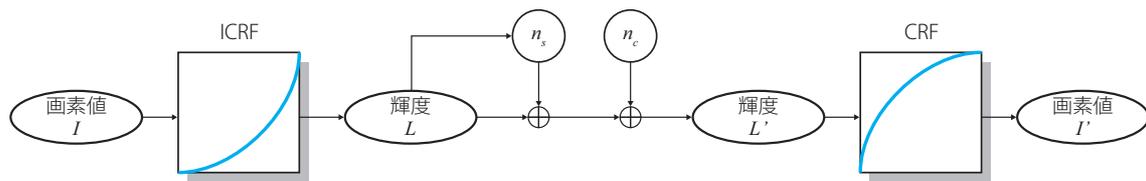


図 3.3: CCD カメラノイズの生成フロー

### 3.3 改ざん検出の流れ

前節では、ノイズレベル関数の定義を導入し、カメラの撮影条件に応じてノイズレベル関数の形状が変化することを示した。本節では、与えられた映像からノイズレベル関数を推定し、その一貫性を手がかりとして改ざん箇所を検出する手法の流れを説明する。

まず 3.3.1 節においてノイズレベル関数をパラメトリックに表現するために必要な基底関数の取得方法について述べる。次に、3.3.2 節において各画素の改ざんの尤度をノイズレベル関数と統計的な知見にもとづいて定義する。そして確率モデルにもとづく改ざん検出の手法を 3.3.3 節で述べる。

#### 3.3.1 基底関数の獲得

(3.2.6) 式ではカメラ応答関数とノイズレベル関数の関連性を示した。しかし、(3.2.6) 式にはカメラ応答関数の他にも未知のパラメータが多くあり、与えられた映像からノイズレベル関数を解析的に求めることは困難である。また、一般に用いられるカメラ応答関数はノンパラメトリックに表現されているため、ノイズレベル関数を解析的に推定することは困難である。そこで、我々は Liu らの用いた手法 [69] に倣い、合成されたノイズ特性に対して主成分分析を行い、ノイズレベル関数の基底関数を求めることにする。基底関数をノイズレベル関数の事前知識として用いることにより、解空間の次元を削減し、推定すべきパラメータの数を減らすことができる。

我々は公開されているデータベース<sup>1</sup> から 201 種類のカメラ応答関数を用意し、ノ

<sup>1</sup><http://www.cs.columbia.edu/CAVE>

イズを合成した．データベースではカメラ応答関数が 1024 個の点集合でノンパラメトリックに定義されているため，ノイズレベル関数も 1024 個の点集合で表現することができる．図 3.3 にノイズ生成のフローを示す．まずノイズを含まない画素値  $I$  を用意し，カメラ応答関数の逆関数  $f^{-1}$  によって輝度領域に逆写像し，ノイズを含まない輝度  $L$  を

$$L = f^{-1}(I) \quad (3.3.1)$$

と得る．次に，輝度  $L$  に対して輝度に依存するノイズ  $n_s$  と輝度と独立に混入するノイズ  $n_c$  を付加し，ノイズを含む輝度  $L'$  を

$$L' = L + n_s + n_c \quad (3.3.2)$$

と計算する．ここで各ノイズの分散は，定数パラメータ  $\sigma_s^2$  と  $\sigma_c^2$  を用いて  $\text{Var}[n_s] = \sigma_s^2 L$ ， $\text{Var}[n_c] = \sigma_c^2$  のように与える．最後にノイズレベル関数  $f$  によって再び画素値の領域に写像し，ノイズを含んだ画素値  $I'$  を

$$I' = f(L') \quad (3.3.3)$$

と得る．このようにして輝度を変化させながら 100 個のノイズを含んだ標本を合成し，平均と分散を計算した．同様の計算を 10 回繰り返し，平均と分散の平均を求めて滑らかなノイズ特性を取得した．こうして得られた 201 種類のノイズ特性データに対して主成分分析を行うことにより，平均関数  $\tau_0(\mu)$  と 6 つの基底関数  $\hat{\tau}_l(\mu)$  を求めた．各関数は 1024 次元のベクトルで表されるため，各関数に 8 次元の多項式を当てはめることにより次元削減を行った．図 3.4 に得られた平均関数と基底関数を示す．

基底関数を用いると，ノイズレベル関数は

$$\tau(\mu) = \tau_0(\mu) + \sum_{l=1}^6 \alpha_l \hat{\tau}_l(\mu) \quad (3.3.4)$$

のように基底関数の線形結合の形で表すことができる．すると，最小二乗の枠組みで結合係数  $\alpha_l$  を求めることにより，与えられた映像のノイズレベル関数を推定する

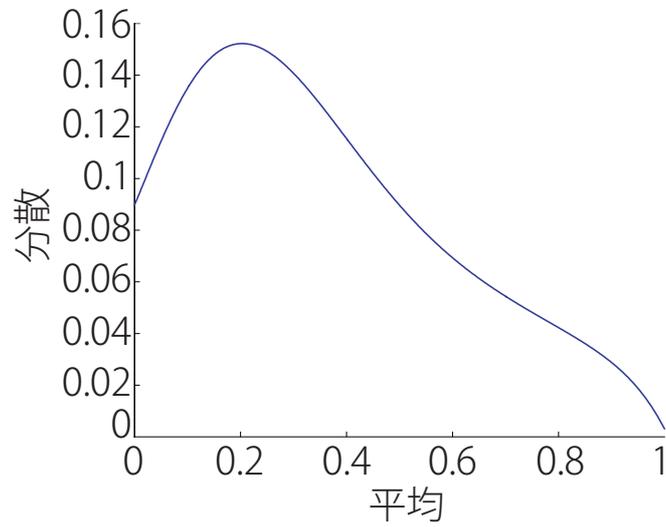
ことができる．ノイズレベル関数を用いると，映像内のある画素の平均  $\mu$  が与えられたとき，その画素の分散  $\sigma^2$  を  $\sigma^2 = \tau(\mu)$  と見積もることができる．ノイズレベル関数から見積られた分散と実際のデータから計算された分散を比較することにより，次節で述べるように各画素の改ざんの尤度を評価する．

### 3.3.2 ノイズ特性の尤度

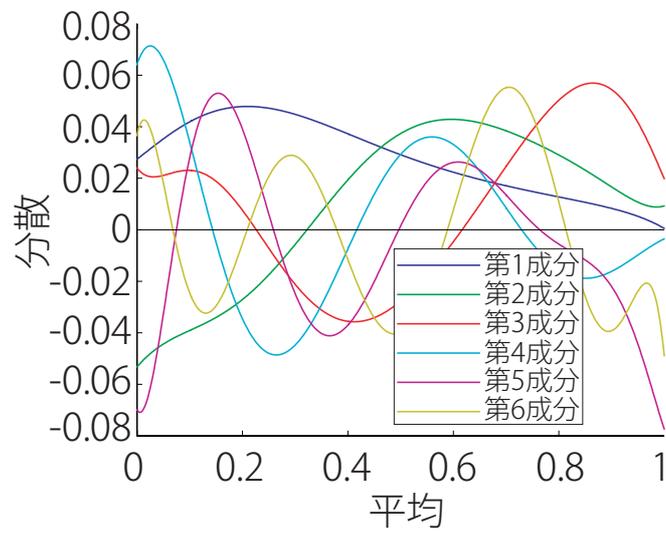
前節ではノイズレベル関数の事前情報として基底関数を取得し，推定すべきノイズレベル関数のパラメータについて述べた．本節では，推定されたノイズレベル関数から，各画素の改ざんの尤度を定式化する．図 3.2 で観測されるように，改ざんが施されていない映像においても計算された分散は揺らぎをもつ．したがって，各画素の分散がノイズレベル関数からずれて観測された場合，そのずれがノイズの揺らぎによるものなのか，改ざんによるものなのかを区別しなければならない．そこで，我々は統計的な立場から分散のずれの評価を行い，各画素が改ざんされた尤度を定式化する．

#### 統計学における用語の整理

我々は，与えられたデータに対してあるノイズ分布を仮定し，分布の特性を示すパラメータを推定することによってノイズ特性を評価する．統計学では，確率変数を生成する母集団と，データから観測される標本集団は明確に区別される [122]．本研究の場合，母集団はノイズが付加されて観測されうる全ての観測値の集合であり，標本集団は与えられた映像から実際に計算された観測値の集合である．母集団が従う確率分布を母集団分布と呼び，母集団分布を特徴付ける値を母数(パラメータ)と呼ぶ．本研究が手がかりとするフォトンショットノイズが従う分布はポアソン分布であるが，光子の数が十分多いことを踏まえ，正規分布  $\mathcal{N}(\mu, \sigma^2)$  に近似する．正規分布の場合，パラメータは母平均  $\mu$  と母分散  $\sigma^2$  の2種類ある．ノイズレベル関数を推定するためには，各輝度におけるノイズ分布のパラメータを推定する必要がある．しかし，我々が観測できるのは母集団から抽出された標本集団だけであるので，母集団の特性を完全に推定することはできない．



(a) 平均関数



(b) 基底関数

図 3.4: 主成分分析によって得られた基底関数

母集団から生成された標本  $\mathbf{x} = \{x_1, \dots, x_N\}$  が与えられたとき，母集団の分布に関する指標を求めることができ，この指標を統計量と呼ぶ．例えば，

$$m = \frac{1}{N} \sum_{n=1}^N x_n, \quad (3.3.5)$$

$$s^2 = \frac{1}{N-1} \sum_{n=1}^N (x_n - m)^2 \quad (3.3.6)$$

は代表的な統計量であり， $m$  と  $s^2$  はそれぞれ標本平均・標本分散と呼ばれる．データ数  $N$  が無限個あるとき，大数の法則から標本平均と標本分散はそれぞれ母平均と母分散に一致する．しかし，有限個数の標本集団から計算された統計量は，真のパラメータからずれて観測される．この揺らぎが図 3.2 の点の分布の原因であり，特にデータの数が分散の大きさに対して少ないとき，有限精度の影響は大きくなる．

### カイ二乗分布

$X_i$  を母平均  $\mu_i$ ，母分散  $\sigma_i^2$  の正規分布に従う  $N$  個の独立な確率変数とすると，カイ二乗値と呼ばれる次の統計量

$$Q = \sum_{i=1}^N \left( \frac{x_i - \mu_i}{\sigma_i} \right)^2 \quad (3.3.7)$$

はカイ二乗分布 ( $\chi^2$  分布) [82] に従う (すなわち  $Q \sim \chi_k^2$ ) . (3.3.7) 式は母平均と母分散で正規化された確率変数の二乗和であり，標本分散  $s^2$  を用いると

$$Q = \frac{(N-1)s^2}{\sigma^2} \quad (3.3.8)$$

と定義される．カイ二乗分布の確率密度関数は，

$$f(x; k) = \begin{cases} 0 & (x < 0) \\ \frac{(1/2)^{k/2}}{\Gamma(k/2)} x^{k/2-1} e^{-x/2} & (x \geq 0) \end{cases} \quad (3.3.9)$$

という形をとる．ここで  $\Gamma(\cdot)$  はガンマ関数である．カイ二乗分布はパラメータを 1 つ持ち， $X_i$  の自由度に等しい正の整数  $k = N - 1$  である．

<sup>1</sup>ここでは標本分散を不偏分散で定義した．標本分散を  $\frac{1}{N} \sum_{n=1}^N (x_n - m)^2$  で定義することもあるが，この場合は標本分散の期待値が母分散に一致しないため注意が必要である．

カイ二乗分布を用いると、あるカイ二乗値  $Q$  が観測される確率を表現することができる。すなわち、母分散が  $\sigma^2$  であるときに標本分散が  $s^2$  と得られる条件付確率は、カイ二乗分布を用いて

$$P(s^2|\sigma^2) = \chi_k^2 \left( \frac{(N-1)s^2}{\sigma^2} \right) \quad (3.3.10)$$

で計算される。

### ノイズレベル関数と改ざん尤度

(3.3.10) 式に表したように、カイ二乗分布はある母分散  $\sigma^2$  を仮定した時に標本分散  $s^2$  が観測される確率を与える。では、改ざん検出において、カイ二乗値の定義式 (3.3.8) 式で用いる母分散  $\sigma^2$  はどのようにして与えられるだろうか。ノイズレベル関数  $\tau(\cdot)$  が与えられたとき、分散は (3.2.1) 式によって推定される。したがって、ノイズ特性  $(m, s^2)$  をもつ画素が観測される確率は、

$$\begin{aligned} P(m, s^2|\sigma^2 = \tau(\mu)) &\simeq P(s^2|\sigma^2 = \tau(m)) \\ &= \chi_k^2 \left( \frac{(N-1)s^2}{\tau(m)} \right) \end{aligned} \quad (3.3.11)$$

と表すことができる。ただしここで  $\mu \simeq m$  と仮定した。(3.3.11) 式で計算される確率を評価することにより、ノイズ特性点  $(m, s^2)$  のノイズレベル関数からのずれを観測の揺らぎで説明できるかどうかを評価することができる。すなわち、(3.3.11) 式は  $(m, s^2)$  を観測した時のノイズレベル関数を  $\tau$  の尤度を表している。次節ではこの尤度を最大化するノイズレベル関数  $\tau(\cdot)$  を推定する。

### 3.3.3 ノイズレベル関数の推定と改ざん検出

3.1 節で述べたように、本研究では静止シーンの映像を対象とする。すると、各画素の時間的な平均  $m$  と分散  $s^2$  がそれぞれ (3.3.5) 式と (3.3.6) 式を用いて得られる。ノイズレベル関数を用いた尤度評価に先立ち、基底関数の定義に合わせるため、平均値は画素の最大値である 255、分散は観測値の最大値で正規化しておく。

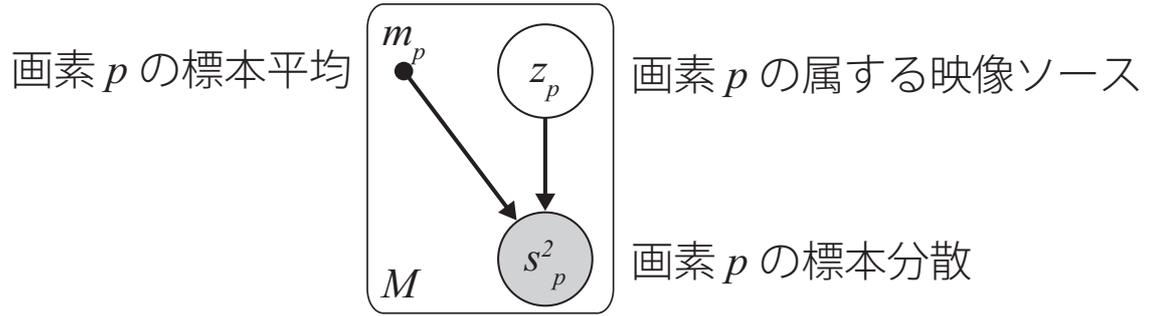


図 3.5: 改ざん映像におけるノイズ特性生成を表すベイジアンネットワーク

映像から画素数  $M$  だけのノイズ特性  $(m, s^2)$  が得られたら、改ざんの事後確率を評価する。ここで、与えられた映像の一部に他の映像から貼り付けられた改ざん領域がある場合、複数のノイズレベル関数が混在することに注意する。すなわち、各画素がどちらのノイズレベル関数に従う領域に属するのかを評価する必要がある。そこで画素  $p$  が属する映像ソースを表す潜在変数  $z_p$  を導入し、改ざん検出の問題を複数の確率分布（ノイズレベル関数）からなる混合確率分布モデルとして取り扱うことにする。ただし、本研究では含まれているノイズレベル関数の種類は2種類で既知とする。すなわち、 $z_p = \{1, 2\}$  である。図 3.5 に改ざん映像においてノイズ特性が観測される過程を表すベイジアンネットワークを示す<sup>2</sup>。ノイズ特性  $(m_p, s_p^2)$  が観測されたときに  $i$  番目のソースに属する事後確率は、ベイズの定理より

$$P(z_p = i | s_p^2; m_p) = \frac{P(s_p^2 | z_p = i; m_p) P(z_p = i)}{\sum_{j=1}^2 P(s_p^2 | z_p = j; m_p) P(z_p = j)} \quad (3.3.12)$$

で表される。(3.3.11) 式を用いると、尤度  $P(s_p^2 | z_p = i; m_p)$  は

$$\begin{aligned} P(s_p^2 | z_p = i; m_p) &= P(s_p^2 | \sigma^2 = \tau_i(m_p)) \\ &= \chi_k^2 \left( \frac{(N-1)s_p^2}{\tau_i(m_p)} \right) \end{aligned} \quad (3.3.13)$$

と書ける。

さて、ここで尤度の評価式 (3.3.13) 式を計算するためには、各映像ソースにおけるノイズレベル関数  $\tau(\cdot)$  が分かっている必要がある。そこで、与えられたノイズ

<sup>2</sup>ここで  $m_p$  は平均であるので本来ならば  $s_p^2$  同様確率変数として取り扱うべきであるが、本研究では (3.3.11) 式の導出と同様にパラメータとして取り扱っている。

特性の点群  $(m_p, s_p^2)$  ( $p = 1, \dots, M$ ) からノイズレベル関数を推定する．ただし異なる映像ソースに属する画素のノイズ特性点を含めて推定を行うと，それらの点は外れ値として振舞うため，正確な推定が行えない．したがって，各画素が属する映像ソースの確率を用いて重み付けされた曲線当てはめを行う必要がある．すると，各画素が属する映像ソースの確率と，各映像ソースのノイズレベル関数は同時に推定しなければならないため，我々はEMアルゴリズム [23] を用いてこれらの値を同時に推定する．

Eステップでは，まずある画素がソース1(ノイズレベル関数が  $\tau_1$ ) に属する尤度を(3.3.13)式で計算する．ソース2(同じく  $\tau_2$ ) に属する尤度も同様に計算する．次に，各ソースに属する事前確率を次式

$$P(z_p = i) = \frac{\sum_{p=1}^M P(s_p^2 | z_p = i; m_p)}{\sum_{j=1}^2 \sum_{q=1}^M P(s_q^2 | z_q = j; m_q)} \quad (3.3.14)$$

を用いて更新する．

Mステップでは，(3.3.4)式で示す  $i$  番目のソースのノイズレベル関数のパラメータ  $\{\alpha_1^{(i)}, \dots, \alpha_6^{(i)}\}$  を推定する．パラメータは次の誤差関数を最小化することによって求める．

$$E(\{\alpha_1^{(i)}, \dots, \alpha_6^{(i)}\}) = \sum_p w_p (s_p^2 - \tau_i(m_p))^2 + \lambda \int_0^1 |\tau_i''(\mu)| d\mu + \int_0^1 \xi(\tau_i(\mu)) d\mu \quad (3.3.15)$$

ここで  $w_p = P(z_p = i | s_p^2; m_p)$  は(3.3.12)式で得られる事後確率で，重み付きの最小二乗解を求めるためのデータ項である．第2項は平滑化項で， $\lambda$ によって平滑度が調節される．第3項はペナルティ項で，次のペナルティ関数によって表現される．

$$\xi(a) = \begin{cases} A & \text{if } a < 0, a > 1 \\ 0 & \text{otherwise} \end{cases} \quad (3.3.16)$$

ここで  $A$  は十分大きな値 ( $A \sim 10^5$ ) である．ペナルティ項はノイズレベル関数が0から1の間の値を取るように制限をかける働きをする．

EステップとMステップを収束するまで繰り返し、各画素が各ソースに属する事後確率と各ソースのノイズレベル関数のパラメータを同時に推定する。初期条件が収束解に影響しないことを確認の後、実験ではノイズレベル関数の形状の初期条件として、1番目のノイズレベル関数に対しては  $\{\alpha_1^{(1)}, \dots, \alpha_6^{(1)}\} = \{1, 0, 0, 0, 0, 0\}$ 、2番目のノイズレベル関数に対しては  $\{\alpha_1^{(2)}, \dots, \alpha_6^{(2)}\} = \{0, 1, 0, 0, 0, 0\}$  と設定した。また、事前確率の初期値はそれぞれのソースに対して0.9と0.1を設定した。

最後に、静止シーン映像に対する改ざん検出の計算フローを Algorithm 1 に示す。

## 3.4 評価実験

提案手法の有効性を確認するため、本研究では次の2種類の検証実験を行った。1つ目は実験室環境で撮影された映像を対象とした実験で、手法の定量的な評価を行った。2つ目はより現実的なシーンを撮影した映像を対象とした実験で、屋内外を撮影した映像を用いて定性的な評価を行った。

実験映像の撮影には、PointGrey Research 社製の IEEE 1394 カメラである Flea を使用し、30fps で撮影された VGA サイズ (640 × 480 画素) のグレースケール映像を用いた。撮影された映像は可逆圧縮である huffyuv コーデックを用いて圧縮し、1Mbps 以上のビットレートで保存した。

### 3.4.1 実験室環境による定量的評価

まずはじめに実験室環境において撮影された映像に対して提案手法を適用し、手法の定量的評価を行った。3.3 節で紹介したデータベースから3種類のカメラ応答関数を選び、Macbeth Color Checker Board を撮影した。カメラ応答関数が変わるとそれに伴って映像の明るさが変化するので、露光時間とゲインを調節して平均の明るさが近くなるように設定した。図 3.2 に示したノイズ特性は実験に用いた映像のものである。また、図 3.6 に撮影された映像の一例を示す。

---

**Algorithm 1** 静止シーン映像に対する改ざん検出

---

入力映像の各画素について，平均  $m_p$  と分散  $s_p^2$  を計算

[初期化]

 $\{\alpha_1^{(1)}, \dots, \alpha_6^{(1)}\} \leftarrow \{1, 0, 0, 0, 0, 0\}$ ,  $\{\alpha_1^{(2)}, \dots, \alpha_6^{(2)}\} \leftarrow \{0, 1, 0, 0, 0, 0\}$  $P(z_p = 1) \leftarrow 0.9$ ,  $P(z_p = 2) \leftarrow 0.1$ **for**  $iter = 1$  to  $MaxIter$  **do**

[Eステップ]

**for**  $NLF = 1$  to  $2$  **do**事後確率  $P(z_p = NLF | s_p^2, m_p)$  の計算  $\dots$  (3.3.12) 式**end for****for**  $NLF = 1$  to  $2$  **do**事前確率  $P(z_p = NLF)$  の更新  $\dots$  (3.3.14) 式**end for**

[Mステップ]

**for**  $NLF = 1$  to  $2$  **do**ノイズレベル関数のパラメータ  $\{\alpha_1^{(NLF)}, \dots, \alpha_6^{(NLF)}\}$  を更新  $\dots$  (3.3.15) 式**end for**

[収束判定]

**if** パラメータの変化  $< \epsilon$  **then**    **return****end if****end for**

---

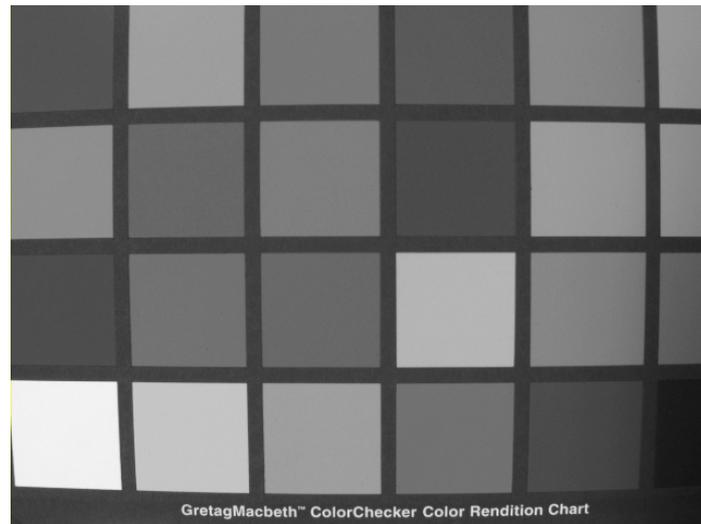


図 3.6: 撮影された映像

撮影された3種類の映像を用いて、改ざん映像を作成した。まず3種類の中から、背景(ソース1)となる映像と貼り付けられる映像(ソース2)の組み合わせを決め、16箇所の $100 \times 100$ 画素の領域に改ざんを施した。各改ざん領域は図3.7で示すように(a)から(p)でラベル付けした。改ざん領域では全フレームにわたってソース1の画素値をソース2の画素値で書き換えた。ソース1とソース2の組み合わせとしては、[ソース1:CRF(a), ソース2:CRF(b)]の組み合わせと[ソース1:CRF(a), ソース2:CRF(c)]の組み合わせを用いた。

3.3節で述べたように映像のノイズ特性を計算し、EMアルゴリズムを用いて2種類のノイズレベル関数のパラメータ $\{\alpha_1^{(i)}, \dots, \alpha_6^{(i)}\}$  ( $i = 1, 2$ )と、各画素が属するソースの事後確率 $P(z_p = i | s_p^2, m_p)$  ( $i = 1, 2$ )を推定した。提案手法は何か“真正であるか”についての指標は定義していないため、改ざん箇所は領域全体と比較して小さいと仮定し、映像内で主要な映像ソースを背景領域(ソース1)と判定する。ここで、主要とは事前確率 $P(z_p = i)$ によって評価する。

図3.8と図3.9に推定結果を示す。図3.8(a)は計算されたノイズ特性とノイズレベル関数の推定結果である。青線はソース1のノイズレベル関数、赤線はソース2のノイズレベル関数の推定結果を示す。各点は計算された1画素のノイズ特性を示し、色が各映像ソースに属する確率を示す。右のカラーバーに示すように、青いほどソー

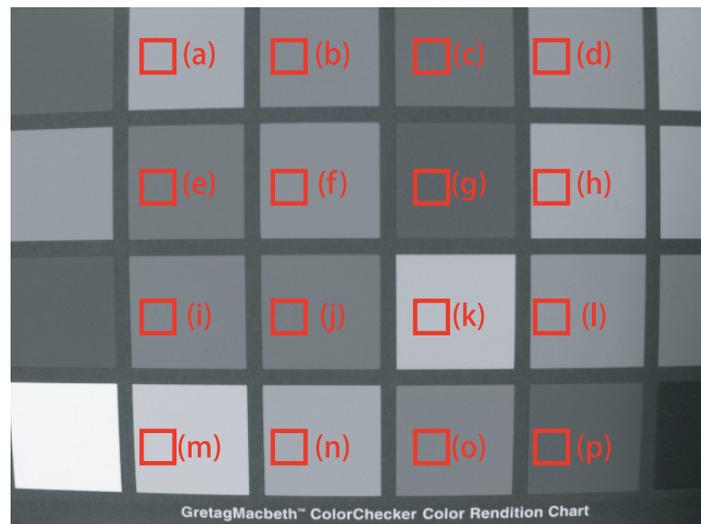


図 3.7: 改ざん領域

ス1に含まれる確率が高く(改ざんの確率が低い), 赤いほどソース2に含まれる確率が高い(改ざんの確率が高い). また, 図 3.8(b) は図 3.8(a) の各点を入力映像の画素に対応付けた結果で, 色は図 3.8(a) で示した確率に対応する.

図 3.8(a) ではソース1とソース2のノイズ特性がきれいに分離しているため, 2本のノイズレベル関数は正確に推定できている. このため, 図 3.8(b) の検出結果では改ざんが施された領域が正確に検出できている. 一方, 図 3.9(a) では各領域のノイズ特性がきれいに分離していない. しかし, 改ざんの検出結果図 3.9(b) を見ると図 3.8(b) ほどではないが改ざん箇所が正確に推定できていることが分かる.

また, 図 3.8 と図 3.9 に結果を示した映像を用いて, 提案手法の定量的評価を行った. 定量的評価は, 図 3.7 で示した 16 箇所の各改ざん領域と, 背景領域内における事後確率  $P(z_p = i | s_p^2, m_p)$  の平均で評価した. 各改ざん映像から計算した平均事後確率を図 3.10 に示す. 図 3.10(a) で示す CRF(a) と CRF(b) の組み合わせでは, すべての改ざん領域において正しい推定が得られている. 一方, 図 3.10(b) で示す CRF(a) と CRF(c) の組み合わせでは, いくつかの改ざん領域において誤った推定が得られていることが分かる. これは, 図 3.9(a) で示したように, 各ソースのノイズ特性がきれいに分離していないことが原因であると考えられる.

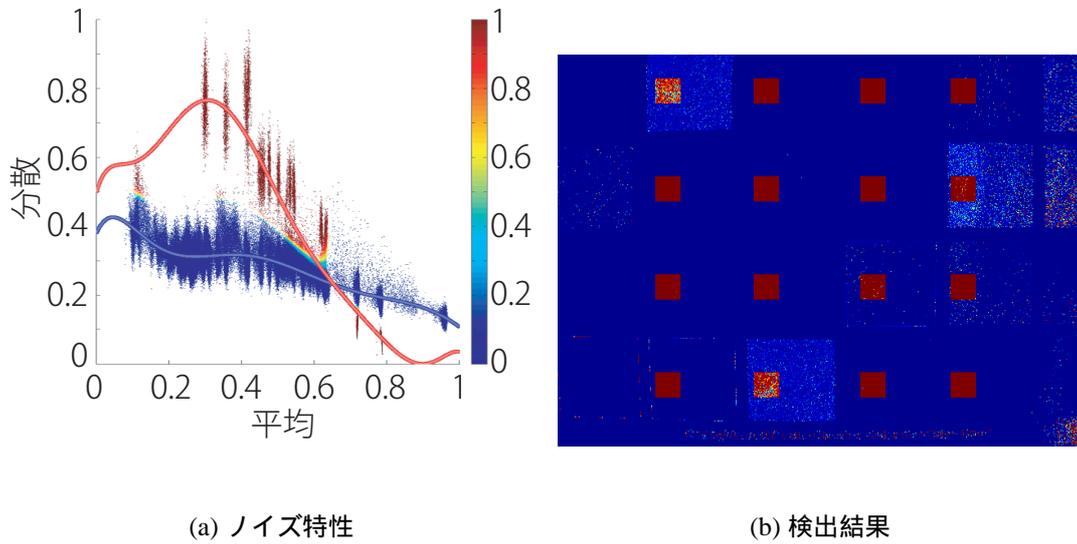


図 3.8: CRF(a) と CRF(b) を用いて改ざんした映像の推定結果 .

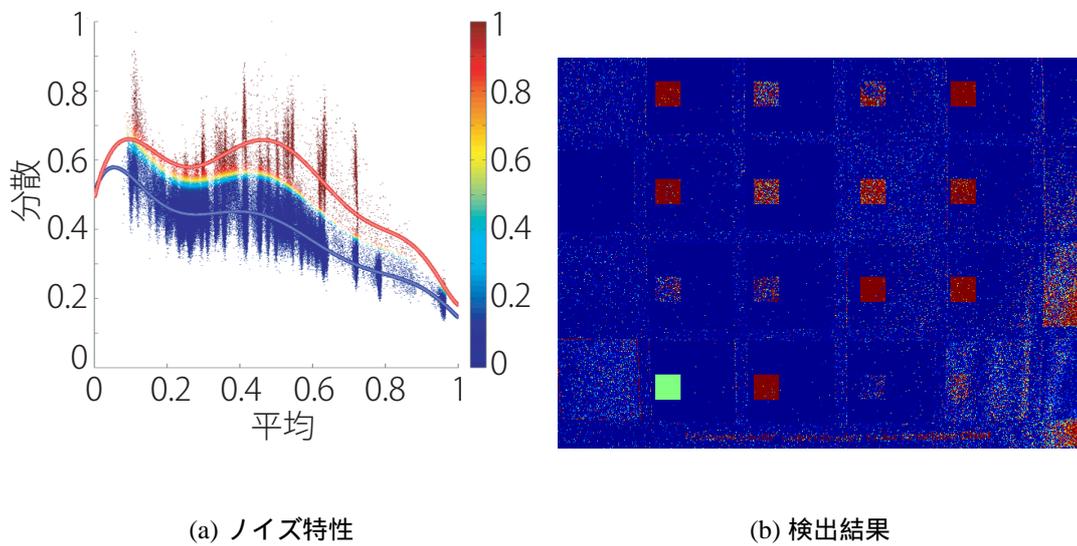
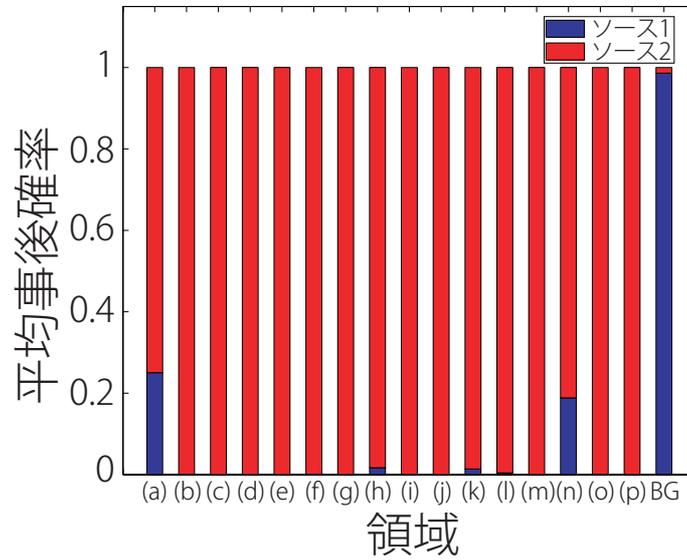
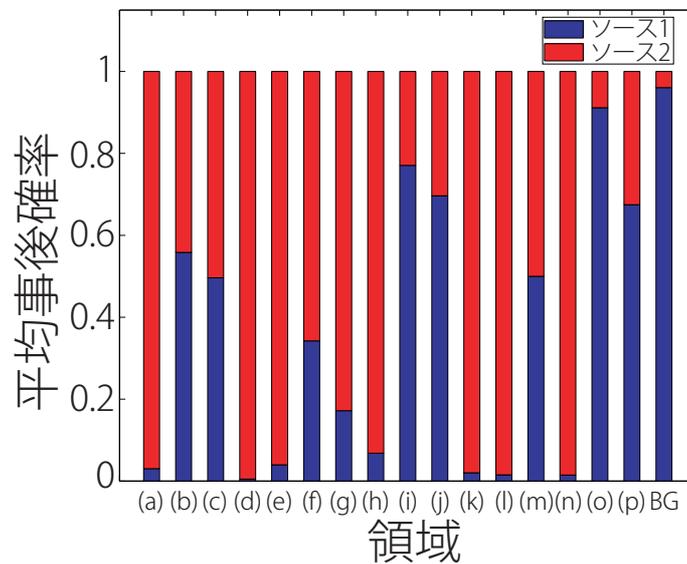


図 3.9: CRF(a) と CRF(c) を用いて改ざんした映像の推定結果 .



(a) CRF(a) と CRF(b)



(b) CRF(a) と CRF(c)

図 3.10: 図 3.8 と図 3.9 の推定結果の定量的評価．各改ざん領域と背景領域におけるソース1, ソース2に属する確率の平均を示す．

### 3.4.2 実映像による定性的評価

前節では実験室環境において提案手法の定量的評価を行った。本節ではより現実的な映像を対象として提案手法を適用し、評価を行う。図 3.11(a) は屋内で撮影したシーンを用いた改ざん映像で、本棚の映像中に紙箱の映像が貼り付けられている。一方、図 3.11(b) は屋外で撮影したシーンを用いた改ざん映像で、建物の映像に自転車の映像が貼り付けられている。改ざん作成には Adobe Premier を使用し、対象物体の切り出し・サイズ変更・明るさやコントラストの調整などを用いて、見た目が自然になるように加工を行った。現実的なシーンでは風などの影響によってシーンに動きが混入するため、静止シーンの仮定に反する画素が存在する。このような画素に対して時間的な分散を計算すると、対応画素の移動によって大きな時間分散を持つ。このような画素から計算されるノイズ特性の誤差の影響を排除するため、分散が閾値よりも大きな画素を改ざん検出の対象から除外し、判定不能として改ざんの事後確率を 0.5 とした。

提案手法を適用した結果を図 3.12 と図 3.13 に示す。図 3.12(b) に示すように、貼り付けられた箱の部分が正確に検出できていることが分かる。一方、図 3.13(b) では建物のガラスや壁の領域で誤った検出が分散しているが、これは図 3.13(a) で示す各領域のノイズ特性が重なった領域に相当することが原因である。しかし、改ざんされた自転車の領域では改ざんの事後確率が背景に比べて十分に高く、改ざん領域の検出が成功していることが分かる。

### 3.4.3 映像圧縮の影響

最後に映像の圧縮が改ざん検出に与える影響について考察する。3.4.1 節で使用した CRF(a) と CRF(b) の組み合わせの改ざん映像を様々なコーデックを用いて圧縮し、改ざん検出の精度にどのような影響があるか検証を行った。

比較対象として、前節までに用いた可逆圧縮コーデックである huffyuv の他、次の 3 種類のコーデックを用いた。(1)MPEG-2:1995 年に開発されたコーデックで、主にデジタル放送や記録メディアの符号化・復号化方式として利用されている。(2)Cinepak:1992



(a) 屋内シーン



(b) 屋外シーン

図 3.11: より現実的なシーンを用いた改ざん映像 .

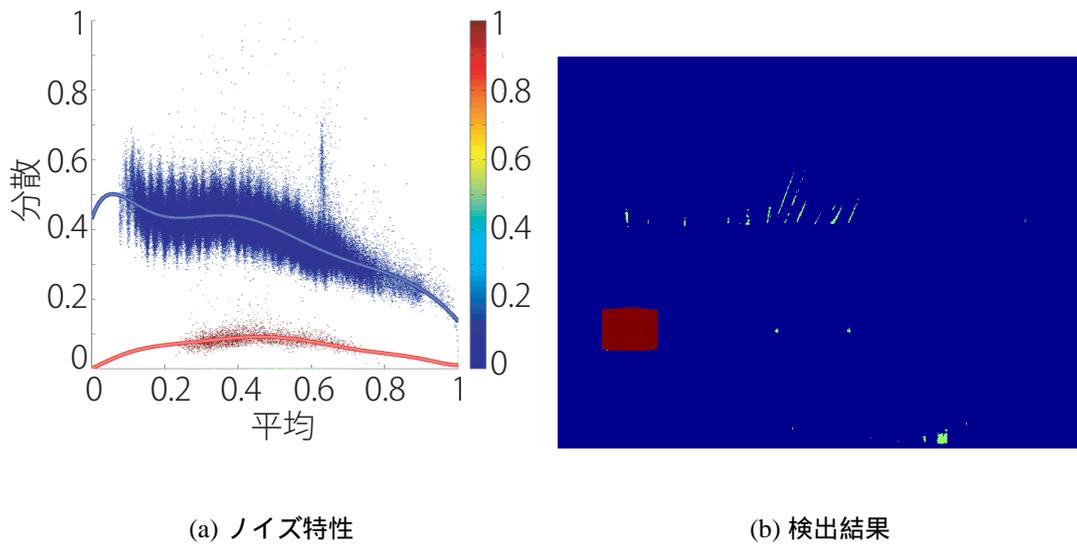


図 3.12: 図 3.11(a) に示した映像の推定結果 .

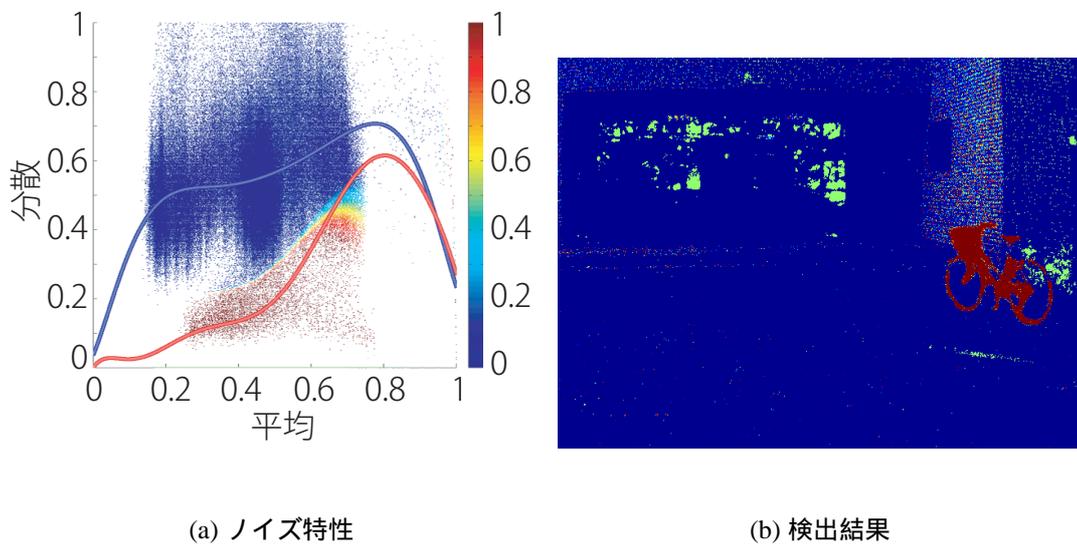


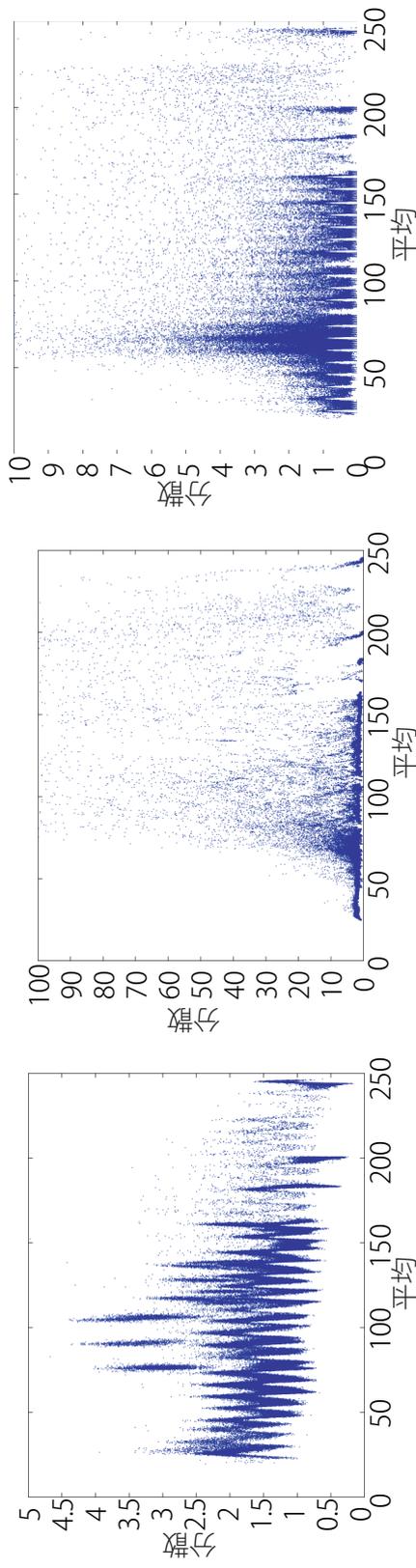
図 3.13: 図 3.11(b) に示した映像の推定結果 .

表 3.1: 様々なコーデックにおけるビットレートと検出精度

コーデック	ビットレート [kbps]	陽性率	陰性率
huffyuv	221184	0.9702	0.9860
MPEG-2	100000	0.4641	0.5503
Cinepak	16770	0.0620	0.9107
H.264	153	0.3938	0.5391

年に開発されたコーデック。初期の QuickTime や AVI で主に用いられた。(3)H.264 : 2003 年に開発されたコーデックで、従来のコーデックに比べて高い圧縮効率を実現している。各コーデックにおいては固定ビットレートを選択し、その他のパラメータは品質が最高となるように設定した。

改ざん検出の精度は、陽性率と陰性率の 2 つの指標を用いて評価した。陽性率は改ざん領域において改ざんされた映像ソースに判定される事後確率の平均として求めた。一方、陰性率は背景領域において背景の映像ソースに判定される事後確率の平均として求めた。表 3.1 に各コーデックで圧縮した映像のビットレートと改ざん検出精度を示す。表より提案手法の精度はコーデックに大きく依存することが分かる。これは、圧縮によって映像の品質が劣化し、新たなノイズが加わったことや、ノイズが除去されることによって提案手法が手がかりとする情報が失われてしまったことが原因と考えられる。図 3.14 に比較に用いた映像のノイズ特性を示す。MPEG-2 では比較的元映像のノイズ特性が保存されているため、表 3.1 に示すように陽性率・陰性率共に比較的高い値が得られている。Cinepak では品質劣化が激しく、分散が非常に大きくなっている。このことが陽性率の悪化を招いたと考えられる。H.264 もわずかに元のノイズ特性を残してはいるが、2 つのノイズレベル関数を分離することはできず、改ざん検出の性能は低くなっている。



(a) MPEG-2

(b) Cinepak

(c) H.264

図 3.14: 様々なコーデックで圧縮した映像のノイズ特性

### 3.5 まとめ

本研究では静止シーンを撮影した映像を対象とし、映像に含まれるノイズの特性の一貫性を検証することによって他の映像から貼り付けられた領域を検出する手法を提案した。提案手法はフォトンショットノイズに由来する輝度値の時間的な平均と分散の関係性に着目し、先行研究に倣ってノイズ特性をノイズレベル関数として定義した。ノイズレベル関数はカメラの種類や撮影条件に固有の形状を示すため、映像内で一貫していない箇所を求めることにより、他の映像ソースから貼りつけられた領域を検出することができる。

我々は人工データと実映像を用いて、撮影条件によってノイズ特性が異なることを示した。そして、ノイズレベル関数を推定するための事前情報として、合成ノイズを用いてノイズレベル関数の基底関数を取得し、ノイズレベル関数を基底関数の線形結合として表現する手法を導入した。また、改ざんを潜在変数を用いた混合確率分布モデルとしてモデル化し、確率的な推論の枠組みで改ざんの確率を推定する枠組みを提案した。各画素が各映像ソースに属する事後確率と各映像ソースのノイズレベル関数のパラメータは同時に推定する必要があるが、提案手法ではEMアルゴリズムを用いて反復的に計算することにより、これらを同時に推定した。計算されたノイズ特性に対して、提案手法はカイ二乗分布を用いて尤度を定式化することにより、有限個のデータを用いて計算された分散の揺らぎを適切に取り扱い、改ざんの尤度を評価することを目指した。

提案手法の有効性を示すため、実験室環境で撮影された映像を用いた定量的な評価と、より現実的なシーンを撮影した映像を用いた定性的な評価を行った。提案手法は2つのノイズレベル関数が混在する改ざん映像に対しても2種類のノイズレベル関数を分離することができ、各画素のノイズ特性が各ノイズレベル関数に従う確率も正確に推定することができた。2つの映像ソースのノイズ特性が重なる場合においても、提案手法は改ざんの確率を比較的正確に計算することができるため、改ざん箇所の候補を特定することにも有効に用いることができた。

提案手法の特長として、ノイズ強度の違いを手がかりとして用いるため、ノイズ

に対して頑健に改ざんを検出できる点が挙げられる。また、画素単位でノイズ特性を計算し、改ざんの確率を求めることができるため、先行研究として密な改ざん検出を行うことができる。

今後の課題としては、まず静止シーンの仮定を取り除き、動的なシーンに手法を拡張することが考えられる。これは第4章にて動的なシーンへの拡張を提案する。また提案手法は圧縮によって仮定にはないノイズが混入したり手がかりとなるノイズが除去された映像に対しては検出精度が低下することが確認された。圧縮による付加ノイズは、フォトンショットノイズとは異なり平均と分散の間に関係性はないと考えられる。よって、フォトンショットノイズに起因するノイズ強度を抽出することによって、圧縮による影響を取り除くような工夫が必要になる。また、提案手法では各画素の改ざんの確率を独立に推定していたが、改ざんはある領域に固まって施されることが多いため、隣接画素の情報を取り入れることによって改ざん検出の精度が向上することが期待される。

## 第4章 動的シーンの改ざん検出

### 4.1 はじめに

第3章では、対象映像として静止シーンを撮影した映像を仮定し、ノイズ特性を手がかりとして改ざんを検出する手法を提案した。しかし静止シーンの仮定では移動する物体が含まれる映像に対して手法を適用できないため、手法の実用範囲が限られるという問題があった。そこで、本章では静止シーンの仮定を緩和し、他の映像から貼りつけられた移動物体を検出する手法を提案する。本章でも改ざんの手がかりとしてノイズ特性の一貫性を検証するが、動的なシーンへ対応するため、静止シーンとは異なるアプローチによりノイズ特性を利用する。

本研究では、対象となる映像に対する仮定として、背景のみが静止したシーンを仮定する。これは監視カメラのように固定されたカメラで撮影された映像を想定し、前景の中から改ざんされた物体を検出することを目指す。監視カメラで撮影された映像は裁判の証拠として採用される可能性が高く、静止シーンと比較して手法の適用範囲が大きく広がると考えられる。

#### 提案手法の概要

次節以降で述べるように、画素単位や領域単位の追跡による追跡では本研究が必要とするノイズ特性を求めることは難しい。そこで、提案手法は時間的に対応する領域を数フレームに渡って手で指定することにより、領域的な分散を求める。領域的な分散にはノイズによる成分の他に陰影とテクスチャの成分が含まれるため、提案手法はこれらの成分を分離し、改ざん検出を行う。

改ざん検出の手がかりとして、領域内の輝度とテクスチャ分散の関係を利用する。線形のカメラ応答関数を仮定すると、テクスチャの分散は輝度の2乗の関数

となる．よって，テクスチャの分散が2乗の関数として表現できない場合は，ノイズとテクスチャの分離が正しく行えていないことを意味する．

ユーザの利便性を考慮し，指定する領域の数は数フレームに抑える必要がある．人工映像と実映像を用いた検証により，提案手法はわずかな標本からも撮影条件の違いによるテクスチャ分散の振る舞いを区別できることを示す．

#### 4.1.1 画素単位の対応点追跡手法の検討

静止シーンと比較した時の技術的な差異として，静止シーンの場合とは異なり時間的な平均と分散を計算することによって簡単にノイズ特性を求めることができないという点が挙げられる．静止シーンの場合は被写体が静止していたため，時間的な画素値の揺らぎは全てノイズに起因すると仮定することができた．しかし，被写体が移動する場合，ある画素に入射する光の時系列はシーン内の異なる点から入射した光が観測されたものである．よって，シーンの変動とノイズによる画素値の変動が混在して観測されることになる．

シーンの変動を補償する解決策として，シーン内で対応する点を追跡しながら画素値の統計量を計算するアプローチが考えられる．時間的な対応点追跡の手法としては，オプティカルフローが多く用いられている．オプティカルフローはコンピュータビジョンにおける基本問題の1つであり，古くから研究されている分野である [46,70]．最新の研究では，ピラミッド構造を用いて最初に粗いフロー推定を行ない，解像度を段階的に高めることで高精度なフロー推定を行う手法 [67,116] や，複数のフローを同時に扱うためのモデルを導入する手法 [93,115] などが提案されている．

アルゴリズムの向上により，画素単位で高精度な動きの検出が実現できるようになってきている一方で，オプティカルフローを用いた対応点追跡を本研究に導入するにあたり，図 4.1 に示すような問題が考えられる．1つ目は図 4.1(a) に示す追跡の頑強性と一様性の相反である．局所的な動き推定では，空間的に輝度が一樣(変化が少ない)方向に対して動きを一位に定めることができない，窓問題 (aperture problem) が存在する．そのため，対応点追跡を行う際にはコーナーやテクスチャなど不定性

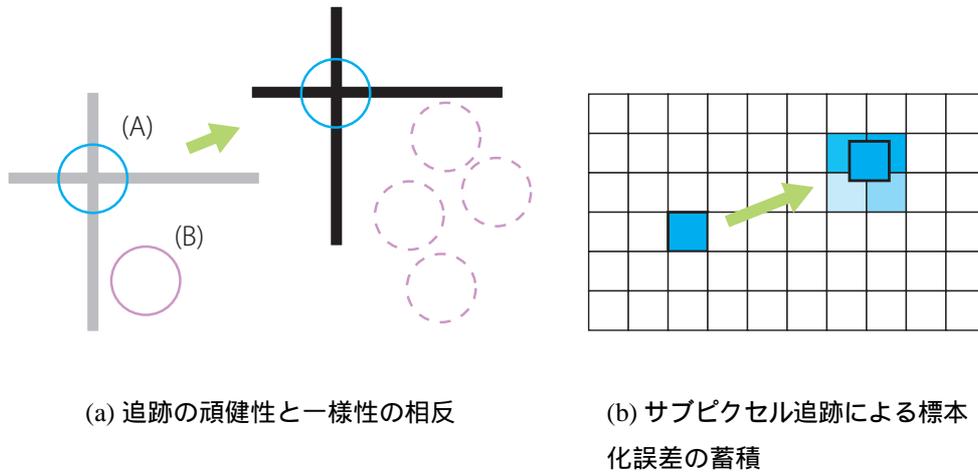


図 4.1: オプティカルフローを用いた対応点追跡における問題点

を生じない点(図中(A))を追跡することにより、窓問題の影響を少なくすることが多い[94]。一方で、追跡に適した点は輝度変化が大きな点であるので、追跡点の周囲の領域で空間的な分散を計算すると、テクスチャの成分を多く含むことになる。よって、本研究が手がかりとするノイズ特性を求めるために適した、一様な反射率を持つ領域(図中(B))は追跡を頑強に行うことができず、相反する性質を持つことが分かる。

また、2つ目の問題点として、サブピクセル追跡による標準化誤差の蓄積が挙げられる。オプティカルフローはある画素が次の時刻に移動した位置をサブピクセル精度で追跡することができる。しかし、図4.1(b)に示すように空間標準化によって追跡点の輝度情報は周囲の画素に分散して観測されるため、追跡を行ううちに誤差が蓄積し、正確なノイズ特性を求めることができない。

#### 4.1.2 領域的な追跡手法の検討

以上のような対応点追跡の問題点に対処するため、本研究では画素単位ではなく、有限領域における空間的な分散からノイズ特性を推定するアプローチを提案する。領域的な追跡に求められる特性としては以下のものが考えられる。

1. 各時刻において、異なるテクスチャ領域を包含する領域分割を行わないこと

## 2. 実空間上で対応する領域を追跡できること

そこで、領域追跡の先行研究が上の特性を満たすかどうか検討する。

オプティカルフローのような画素単位の動き推定ではなく、領域的な追跡を行う手法としては、人物や車両を対象とする研究が多い。一般に人物追跡や車両追跡の研究において、追跡結果は対象を囲う矩形や楕円領域(+重心位置)によって示される [2, 12, 20, 55, 59, 63, 78, 114]。これらの手法では、検出された領域からの対象物体の抽出は領域分割の問題となる。また、追跡の段階から輪郭線追跡を行う手法も提案されている [4, 14, 78]。しかし、これらの手法が行っていることは対象物体単位の追跡であり、各領域には異なるテクスチャが包含される。よって、これらの手法は特性 1. の条件を満たさないことが分かる。

また、画像を局所的に色が類似する領域に分割する様々な手法が提案されている [3, 73, 81, 92]。これらの手法は1枚の静止画に対する領域分割の手法であり、映像を対象としたものではない。また、各時刻に独立にこれらの手法を用いても、時間的な連続性を保証しないため、領域の時間的な変化や移動を追跡することはできない。よって、これらの手法は特性 2. の条件を満たさないことが分かる。

以上の考察により、領域的な追跡の先行研究を導入することによりノイズ特性を推定することは難しいと考えられる。

### 4.1.3 空間領域における分散を手がかりとした改ざん検出

#### 手動による計算領域指定

前節までに、先行研究における画素単位や領域的な追跡では、本研究が必要とするノイズ特性は求められないことを述べた。そこで、本研究では追跡をベースとして計算領域を自動的に指定するアプローチではなく、手動で計算領域を指定するアプローチを取ることにする。手動で計算領域を指定するためには、ユーザの利便性を考えると

- 空間的な指定量 (指定する領域の個数)

- 時間的な指定量 (指定するフレーム数)

を極力減らさなければならない。空間的な指定量に関しては、改ざんが物体単位で施されるため、ノイズ特性を求めるのに最適な箇所を数カ所指定すれば良いので、ユーザへの負担はそれほど大きくない。一方、時間的な指定量に関しては、推定を行うために必要な情報量を確保する必要から多くなることが考えられるため、アルゴリズムを工夫する必要がある。多くても数フレームから十数フレーム程度が現実的な限界であると言える。

#### 分散成分の分離問題

空間的に広がりのある領域から計算された分散には、陰影・テクスチャ・ノイズの分散が同時に混入し、分離できないという問題がある。この問題に対し、本研究ではまず陰影の変動を平面近似することによって除去し、またノイズ成分は背景領域から推定したノイズレベル関数から推定することで除去する。

分散の各成分を分離したら、映像内での特徴量の一貫性を検証することにより改ざんの検出を行う。ここで対象映像に対して線形のカメラ応答関数を仮定し、改ざんの検出には背理法を用いる。線形のカメラ応答関数を仮定すると、テクスチャの分散は輝度の2乗の関数となる性質がある。したがって、物体の明るさが時刻によって異なるとき、もし輝度とテクスチャの分散の関係が2乗の関係として記述できる場合は、その物体は背景領域と同じ撮影パラメータを用いて撮影されたと判断する。逆に輝度とテクスチャの分散の関係が2乗の関係として記述できない場合は、ノイズの分散が異なる、すなわち、背景のノイズレベル関数とは異なる特性を持つノイズが物体の輝度値に混入したと結論づけられるため、改ざんを検出することができる。

改ざんを検出するために必要な時間的な指定量は、輝度とテクスチャの分散が2次曲線上に分布するかどうかを判断できるだけの標本数によって決まり、データへの依存性が大きい。そこで、本研究では人工映像と実映像を用いた検証を通じて提案手法の実用性と限界を示す。

以下、4.2節では空間的な分散の成分分離と改ざんの指標について述べ、4.3節で評価実験により提案手法の有効性を検証する。最後に、4.4節において本研究の結論

を述べる。

## 4.2 改ざん検出の流れ

本手法では、移動物体を対象とした改ざん検出の実現のため、空間的な分散を計算することによりノイズ特性の一貫性を検証する。空間的な分散は陰影・テクスチャ・ノイズの分散を含むため、ノイズ特性を推定するためには分散成分の分離が必要である。そこで、本手法ではまず陰影による輝度変動を平面近似することによって除去し、ノイズの成分は背景領域から推定したノイズレベル関数を用いて除去することにより、分散に含まれる各成分の分離を行う。線形のカメラ応答関数の仮定のもとではテクスチャの分散と輝度が2乗の関係式で記述できるため、推定された2次曲線からの解離度を評価することによって改ざんを検出する。

### 4.2.1 テクスチャ分散と輝度の関係式

本節ではテクスチャの分散と輝度との関係式を導出する。本研究では、線形のカメラ応答関数で撮影された映像を仮定する。一般のカメラ応答関数で撮影された映像の場合、与えられた映像からカメラ応答関数を推定する先行研究の手法 [77, 101, 102, 112] を用いて輝度と画素値の関係を線形化できるので、この仮定はカメラ応答関数の推定精度の範囲内で一般性を損なわない。

提案手法では、ユーザは移動物体の時間的に対応する領域を手動で指定する。指定された領域は局所的なテクスチャを持つが、広域的には一様な特性であるとする。ユーザによって指定された領域の画素値の集合を  $x$  とすると、この領域の平均値と分散は次式

$$\mu_x \equiv E[x] \quad (4.2.1)$$

$$\sigma_x^2 \equiv \text{Var}[x] \quad (4.2.2)$$

として定義される。また、別の時刻において  $x$  に対応する領域の新たな画素値の集合を  $y$  とする。ノイズの影響を無視すると、光源との位置関係によって対象物体の輝

度が変化するとき，微小領域の輝度変動は線形と見なすことができる．本研究では線形のカメラ応答関数を仮定することから，新たな画素値は定数  $a$  を用いて  $y = ax$  と表すことができる．すると， $y$  に関する平均値と分散は，

$$\begin{aligned}\mu_y &\equiv E[y] \\ &= E[ax] = aE[x] \\ &= a\mu_x\end{aligned}\tag{4.2.3}$$

$$\begin{aligned}\sigma_y^2 &\equiv \text{Var}[y] \\ &= \text{Var}[ax] = a^2\text{Var}[x] \\ &= a^2\sigma_x^2\end{aligned}\tag{4.2.4}$$

のように  $x$  の平均と分散を用いて表すことができる．ここで (4.2.3) 式と (4.2.4) 式から，

$$\frac{\sigma_y^2}{\mu_y^2} = \frac{\sigma_x^2}{\mu_x^2} = \text{const.} \equiv \gamma\tag{4.2.5}$$

を得る．(4.2.5) 式は分散と輝度値の間にある定数  $\gamma$  で特徴付けられる関係があることを表し， $\gamma$  は領域内のテクスチャの強度に依存する不変量である．(4.2.5) 式から，テクスチャの分散  $\sigma_{\text{texture}}^2$  は平均輝度  $\mu$  を用いて

$$\sigma_{\text{texture}}^2 = \gamma\mu^2\tag{4.2.6}$$

と表すことができる．提案手法は対象物体が (4.2.6) 式を満たすかどうかを評価することにより改ざんの有無を判定する．

## 4.2.2 ノイズの分散成分の推定

先に述べたように，空間的な分散は陰影・テクスチャ・ノイズの成分を共に含むため，(4.2.6) 式による改ざんの評価を行うためにはこれらの成分を分離する必要がある．本節では領域に含まれる分散の成分を推定する手法について述べる．本研究では静止した背景を仮定するため，背景のノイズレベル関数は静止シーンと同様に

時間的な平均と分散を計算することにより求めることができる。したがって、背景から推定したノイズレベル関数と計算領域の平均を用いて、ノイズの分散成分を求めることができる。

本研究では線形のカメラ応答関数を仮定しているため、(3.2.6)式からノイズレベル関数も線形となることが予測される。しかし、カメラのダイナミックレンジの限界のため、観測される輝度はある範囲内に制限される(8ビットデータでは0~255)。したがって、輝度値の上限付近では飽和によって急速に分散が減少する。飽和による分散の低下を含めたノイズレベル関数の解析は困難であるため、静止シーンの場合と同様に、合成ノイズから基底関数を求めることにより、ノイズレベル関数を線形結合の形で表現することにする。

まず、一般のカメラ応答関数の場合と同様に輝度データを用意し、輝度に依存するノイズ  $n_s$  と輝度に依存しないノイズ  $n_c$  を付加する。ノイズが付加された輝度をカメラ応答関数によって画素値の領域に写像し<sup>1</sup>、ノイズを含む画素値の標本を取得する。次に、大量の標本から平均と分散を計算することにより、ノイズレベル関数を求める。ノイズの強度を変えながらノイズレベル関数を求め、最後に主成分分析によって基底関数を求める。各基底関数は9次の多項式によって近似する。合成ノイズから取得した基底関数を図4.2に示す。

基底関数を用いるとノイズレベル関数は(3.3.4)式と同様に線形結合の形で表現することができる。ここで、本研究では分散のスケール不定性を取り除くため、基底関数の結合係数  $\{\alpha_1, \dots, \alpha_6\}$  の他にスケール変数  $\alpha_0$  を加え、ノイズレベル関数の表現を

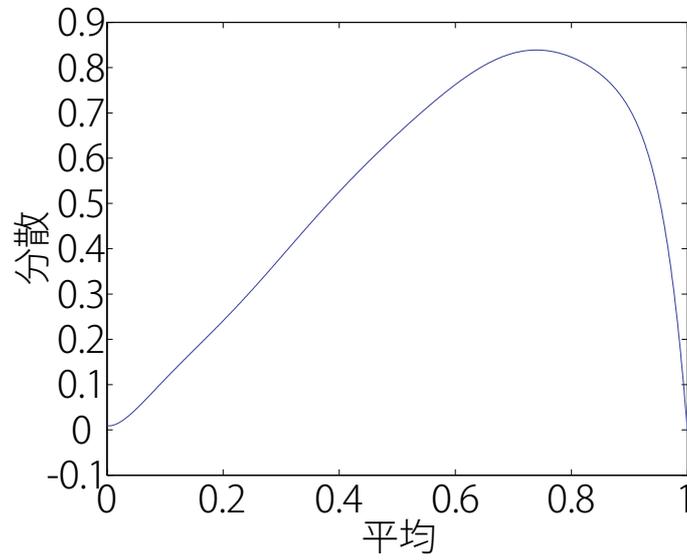
$$\tau(\mu) = \alpha_0 \left\{ \tau_0(\mu) + \sum_{l=1}^6 \alpha_l \hat{\tau}_l(\mu) \right\} \quad (4.2.7)$$

と拡張する。ここで

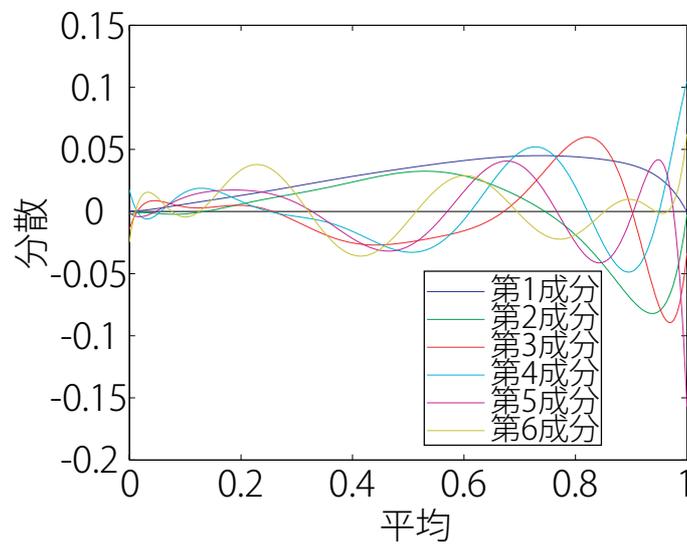
$$\begin{aligned} \beta_0 &\equiv \alpha_0 \\ \beta_l &\equiv \alpha_0 \alpha_l \quad (l = 1, \dots, 6) \end{aligned} \quad (4.2.8)$$

$$\hat{\tau}_0(\mu) \equiv \tau_0(\mu)$$

<sup>1</sup>本研究では線形のカメラ応答関数を用いるため、実装ではカメラ応答関数による写像は必要ない。



(a) 平均関数



(b) 基底関数

図 4.2: 線形カメラ応答関数におけるノイズレベル関数の基底関数

と再定義すると, (4.2.7) 式は

$$\tau(\mu) = \sum_{l=0}^6 \beta_l \hat{\tau}_l(\mu) \quad (4.2.9)$$

と書き直すことができ, スケールを考慮した関数の表現が行える.

背景領域の各画素について時間的な平均  $\mu_n$  と分散  $\sigma_n^2$  を求めた後, ノイズレベル関数のパラメータを次の最小化問題として解く.

$$\{\bar{\beta}_0, \dots, \bar{\beta}_6\} = \operatorname{argmin}_{\{\beta_0, \dots, \beta_6\}} \sum_n \left\| \sigma_n^2 - \sum_{l=0}^6 \beta_l \hat{\tau}_l(\mu_n) \right\|^2 \quad (4.2.10)$$

ただし, ノイズレベル関数の特性から以下の拘束式のもとで最小化を行う.

$$\sum_{l=0}^6 \beta_l \hat{\tau}_l(\mu) \geq 0 \quad \text{for } \mu = [0, 1] \quad (4.2.11)$$

### 4.2.3 改ざん検出の指標

本節では指定された計算領域からテクスチャとノイズの分散を計算し, 改ざんの判定を行う流れについて述べる. 空間的な分散には, 領域内の陰影, テクスチャおよびノイズによる分散の成分が含まれる.

$$\sigma_{\text{total}}^2 = \sigma_{\text{shading}}^2 + \sigma_{\text{texture}}^2 + \sigma_{\text{noise}}^2 \quad (4.2.12)$$

本研究では領域内の各輝度変動成分を分離し, ノイズ特性の一貫性を検証する.

まず, これらの成分から陰影による分散の影響を除去する. 本研究では Liu らが用いた手法 [69] にならって陰影による領域内の輝度変動を除去する. 計算領域を  $\Omega$  とし,  $\mathbf{v} = [x, y]^T$  を座標,  $I(\mathbf{v})$  を座標  $\mathbf{v}$  における輝度とする. 領域内の陰影が平面で近似できると仮定すると, 次式

$$\mathbf{a}^* = \operatorname{argmin}_{\mathbf{a}} \sum_{\mathbf{v} \in \Omega} \left\| I(\mathbf{v}) - \mathbf{a}^T \begin{bmatrix} \mathbf{v} \\ 1 \end{bmatrix} \right\|^2 \quad (4.2.13)$$

の最小化問題を解くことにより平面の当てはめができる. ここで  $\mathbf{a} = [a, b, c]^T$  は平面を表すパラメータである. 領域内の陰影を平面  $\mathbf{a}$  で推定できたので, 輝度から陰

影成分を次式

$$I'(\mathbf{v}) = I(\mathbf{v}) - [a, b, 0]^T \begin{bmatrix} \mathbf{v} \\ 1 \end{bmatrix} \quad (4.2.14)$$

によって除去する．ただし，平均輝度を保存するために  $c = 0$  とした．

続いて，陰影成分を除去した輝度  $I'(\mathbf{v})$  の平均  $\mu$  と分散  $\sigma^2$  を計算する． $I'(\mathbf{v})$  の分散にはテクスチャとノイズの成分が含まれるが，このうちノイズの分散をノイズレベル関数 ((4.2.9) 式) を用いて

$$\sigma_{\text{noise}}^2 = \tau(\mu) \quad (4.2.15)$$

と推定し， $\sigma^2$  から差し引くことによりテクスチャの分散  $\sigma_{\text{texture}}^2$  を得る．

以上の計算を異なる時刻で指定された計算領域に対して繰り返し，領域内の平均輝度  $\mathcal{M} \equiv \{\mu_1, \dots, \mu_T\}$  とテクスチャ分散  $\mathcal{S} \equiv \{\sigma_1^2, \dots, \sigma_T^2\}$  の組を得る．平均輝度とテクスチャ分散の間には (4.2.6) 式で示す関係があるので，次の二乗誤差

$$E(\gamma) = \sum_{t=1}^T (\sigma_t^2 - \gamma \mu_t^2)^2 \quad (4.2.16)$$

を最小にするパラメータ  $\gamma$  を求める．(4.2.16) 式を  $\gamma$  について微分することにより，最小二乗解を

$$\hat{\gamma} = \max \left( \frac{\sum_{t=1}^T \mu_t^2 \sigma_t^2}{\sum_{t=1}^T \mu_t^4}, 0 \right) \quad (4.2.17)$$

と得る．ここで，異なるノイズレベル関数から推定されたノイズの分散を差し引いた場合，テクスチャ分散が負になる場合を考慮し， $\gamma \geq 0$  となるように調整する．

最後に，推定された2次曲線と映像から計算されたテクスチャ分散の間の RMSE を評価することにより，改ざんの指標を求める．

$$\text{RMSE} = \frac{\sqrt{\sum_{t=1}^T (\sigma_t^2 - \hat{\gamma} \mu_t^2)^2 / T}}{|\sigma_{\text{max}}^2|} \quad (4.2.18)$$

ここでテクスチャの強度によるスケールの違いを吸収するために，観測された最大のテクスチャ分散  $\sigma_{\text{max}}^2$  で RMSE を正規化する．次節では，この改ざんの有無によって RMSE が大きく異なることを確かめる．

最後に，動的シーン映像に対する改ざん検出の計算フローを Algorithm 2 に示す．

---

**Algorithm 2** 動的シーンに対する改ざん検出

---

[背景]

背景映像からノイズレベル関数  $\tau(\cdot)$  を推定  $\cdots$  (4.2.10) 式

[前景]

**for**  $obj = 1$  to  $M$  **do**    **for**  $t = 1$  to  $T$  **do**        前景領域の陰影成分除去  $\cdots$  (4.2.14) 式        前景領域の平均  $\mu_t^{(obj)}$  と分散  $\sigma_{total\ t}^2^{(obj)}$  を計算        分散のノイズ成分  $\sigma_{noise\ t}^2^{(obj)} = \tau(\mu_t^{(obj)})$  を計算  $\cdots$  (4.2.15) 式        分散のテクスチャ成分  $\sigma_t^2^{(obj)} = \sigma_{total\ t}^2^{(obj)} - \sigma_{noise\ t}^2^{(obj)}$  を計算    **end for**        テクスチャ分散の分布パラメータ  $\hat{\gamma}^{(obj)}$  を計算  $\cdots$  (4.2.17) 式        理論曲線からの RMSE を計算  $\cdots$  (4.2.18) 式**end for**

---

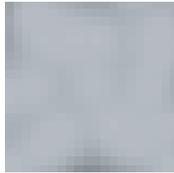


図 4.3: 前景テクスチャ



図 4.4: 背景

### 4.3 評価実験

本節では提案手法の有効性を実験により検証する。

まず、撮影パラメータによって正規化された RMSE が変化することを人工映像を用いて確認する。提案手法の基本的な有効性を確認した後、実映像を用いてユーザの利便性を含めた提案手法の実用性を検証する。

#### 4.3.1 人工映像を用いた検証

実映像を用いた実験では、提案手法が手がかりとするフォトンショットノイズ以外の外乱によって、提案手法の基本的な有効性が確認できない可能性があるため、はじめに人工映像を用いて検証を行う。

検証に用いた映像の前景と背景をそれぞれ図 4.3 と図 4.4 に示す。背景は全ての輝度におけるノイズ強度を得るするために  $640 \times 480$  画素のグラデーションを採用し、輝度依存するノイズを付加した。また、前景は一様な  $20 \times 20$  画素の領域にテクスチャを付加した画像を採用し、領域全体の輝度を縮尺  $a$  によって定数倍してノイズを付加した後、カメラ応答関数によって画素値空間に写像した。

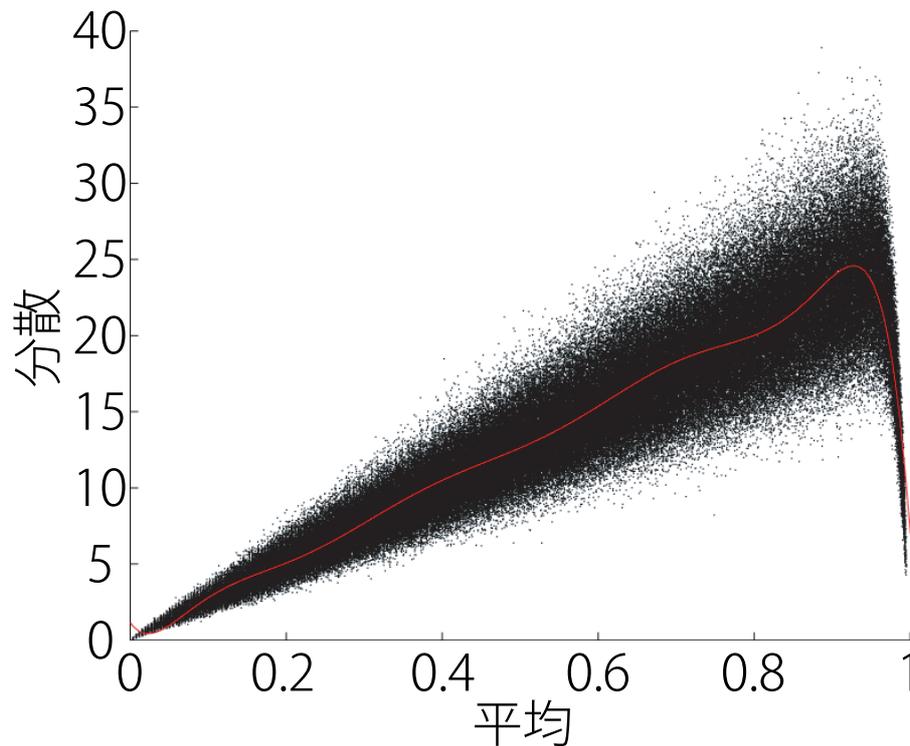
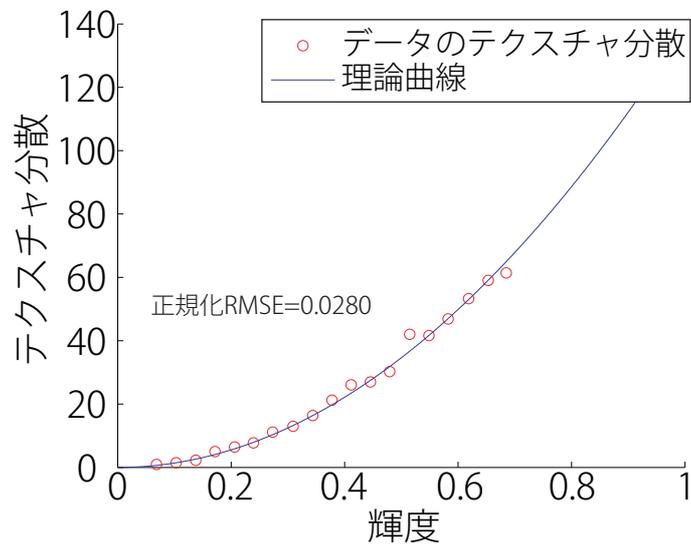


図 4.5: 人工映像の背景のノイズレベル関数

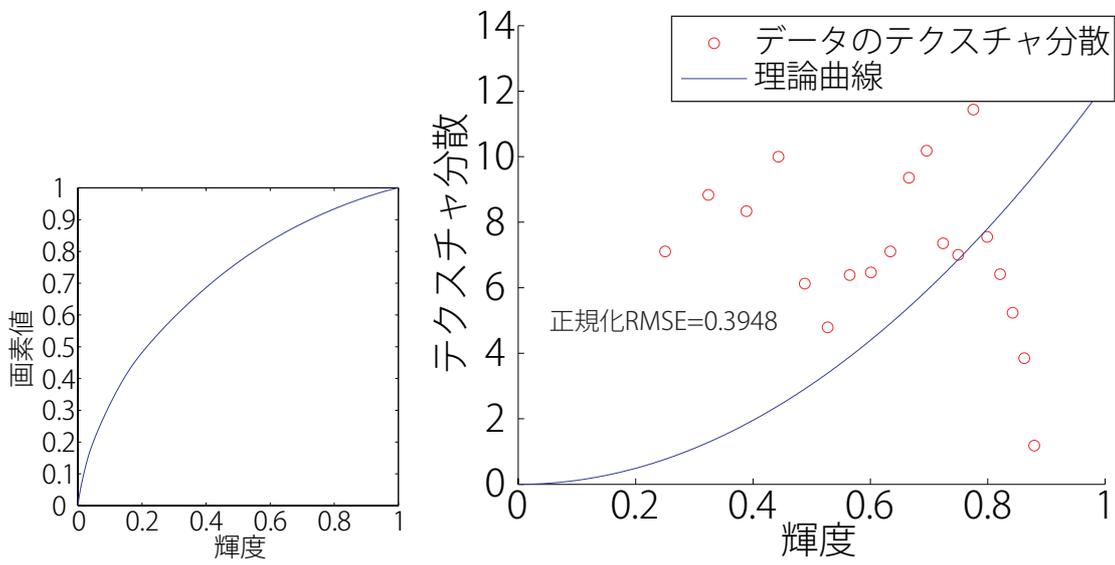
まず, 100 フレームの背景映像から求めたノイズ特性を図 4.5 に示す. 横軸は画素値の取りうる最大値 (=255) で正規化した平均, 縦軸は分散を示す. 各点は 1 画素のノイズ特性を示し, 赤線は基底関数を用いて推定したノイズレベル関数である. 線形カメラ応答関数におけるノイズレベル関数の基底関数を用いることにより, 画素値の最大値付近における分散の減少を正確に捉えていることが分かる.

前景については, 輝度値の縮尺  $a$  を 0.1 から 1 まで 0.05 ずつ変化させながら領域内の平均とテクスチャ分散を計算した. 図 4.6 に線形カメラ応答関数を用いた場合と非線形カメラ応答関数 (CRF No.10) を用いた場合のテクスチャ分散の変化を示す. 図から, 線形のカメラ応答関数の場合はテクスチャの分散が輝度の 2 乗で精度良く表現することができるが, 非線形の場合には異なる分布となることが分かる. また, 正規化 RMSE は線形カメラ応答関数の場合は 0.0280, 非線形カメラ応答関数場合は 0.3948 であった.

提案手法の限界として, 次に示すケースが確認された. まず, 提案手法はカメラ



(a) 線形カメラ応答関数のテクスチャ分散



(b) CRF No.10

(c) 非線形カメラ応答関数のテクスチャ分散

図 4.6: 輝度とテクスチャ分散の関係

応答関数の形状の違いによって改ざんの有無を判定するため、線形に近いカメラ応答関数の場合は線形のカメラ応答関数の場合との区別が難しい。例えば、図 4.7 に示すように、線形に近いカメラ応答関数の場合、正規化 RMSE は 0.0345 と線形カメラ応答関数の場合に近い。これは、静止シーンに対する手法も含めた本研究の限界である。また、提案手法は少数のテクスチャ分散から改ざんの有無を判定するため、輝度の分布によっては非線形なカメラ応答関数であっても RMSE が小さくなることもある。例えば図 4.8 に示す非線形なカメラ応答関数の場合では、輝度が暗い領域に偏って分布しており、この範囲では分布が 2 次曲線で表現できているため、RMSE の値が 0.0124 と小さくなっている。

様々なカメラ応答関数で撮影された映像に対する提案手法の有効性を示すため、第 3 章で用いたカメラ応答関数のデータベースから 201 種類のカメラ応答関数を取得し、同様に正規化 RMSE を計算した。図 4.9 にカメラ応答関数ごとの RMSE を示す。横軸がカメラ応答関数の番号で、縦軸が正規化 RMSE を表す。また、赤線が参考として示した線形カメラ応答関数の場合の正規化 RMSE である。図 4.7 や図 4.8 で示したように、非線形のカメラ応答関数であっても正規化 RMSE が小さくなるものが含まれてはいるが、ほとんどのカメラ応答関数において線形の場合と比較してかなり大きな値を示し、分離が可能であることが分かる。201 種類のカメラ応答関数のうち、線形のカメラ応答関数の正規化 RMSE 以下の値を示したものが約 4% にあたる 8 種類あり、カメラ応答関数全体の正規化 RMSE の平均は 0.3836 であった。これは線形カメラ応答関数の正規化 RMSE である 0.0280 の約 13.7 倍に相当する。

### 4.3.2 実映像を用いた検証

人工映像を用いた実験では、カメラ応答関数によってテクスチャ分散と理論曲線の正規化 RMSE が大きく異なることを示した。本節では、実映像を用いた実験によって提案手法の有効性を検証する。

実験用の映像の撮影には、静止シーンと同様に PointGrey Research 社製の IEEE1394 カメラである Flea を使用し、15fps で撮影された VGA サイズ (640 × 480 画素) のグ

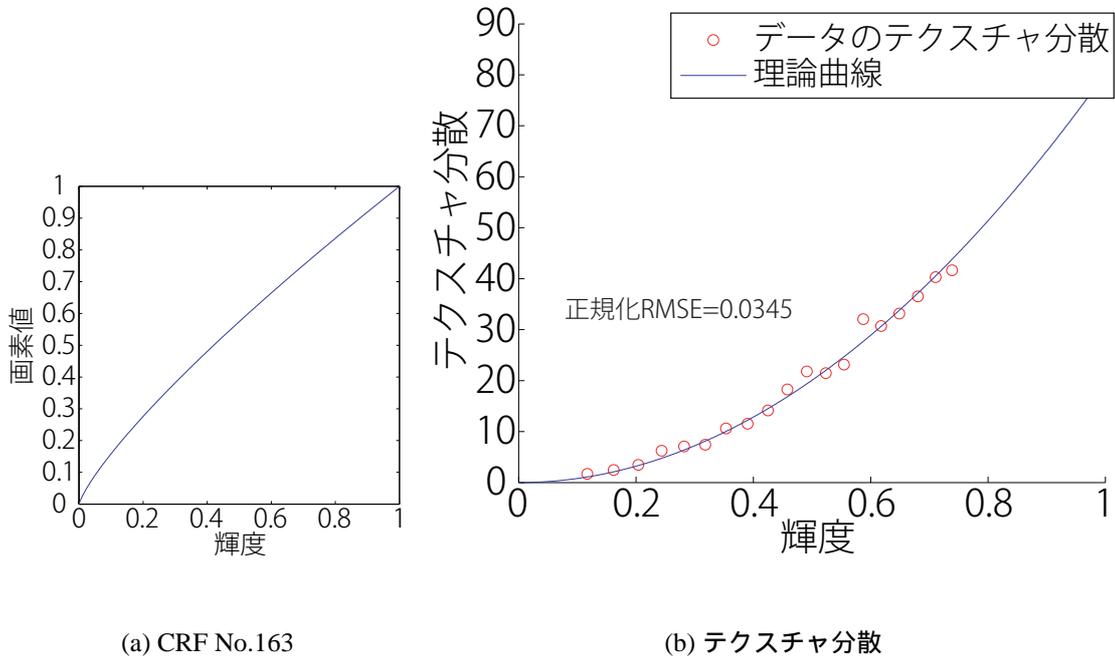


図 4.7: 失敗例 1 : カメラ応答関数が線形に近い場合

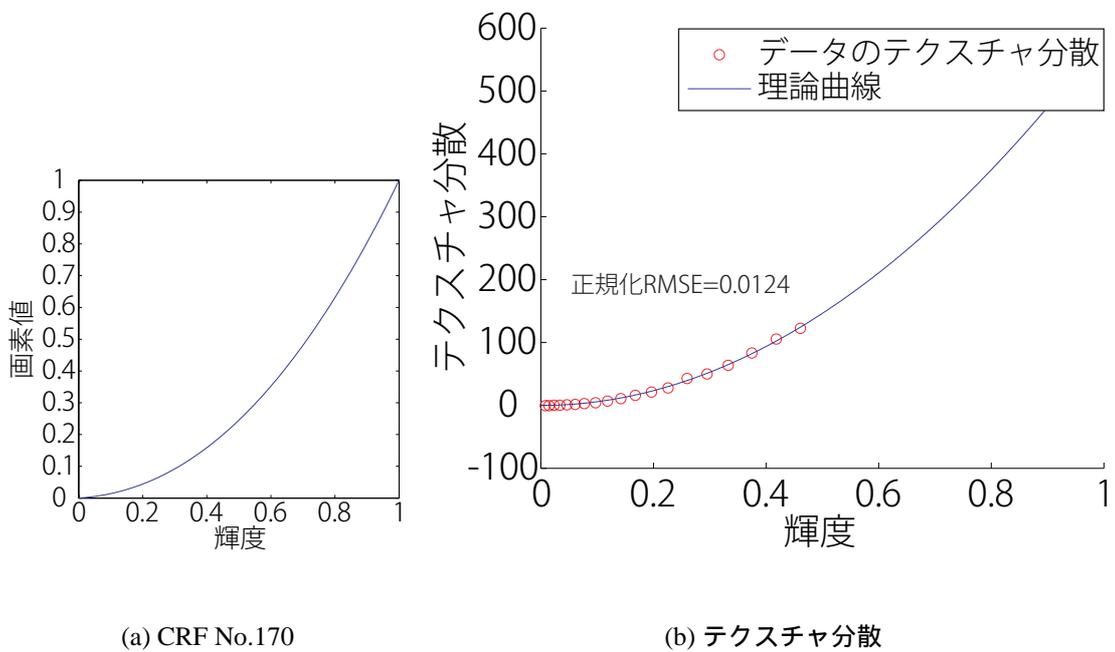


図 4.8: 失敗例 2 : 輝度の分布が偏っている場合

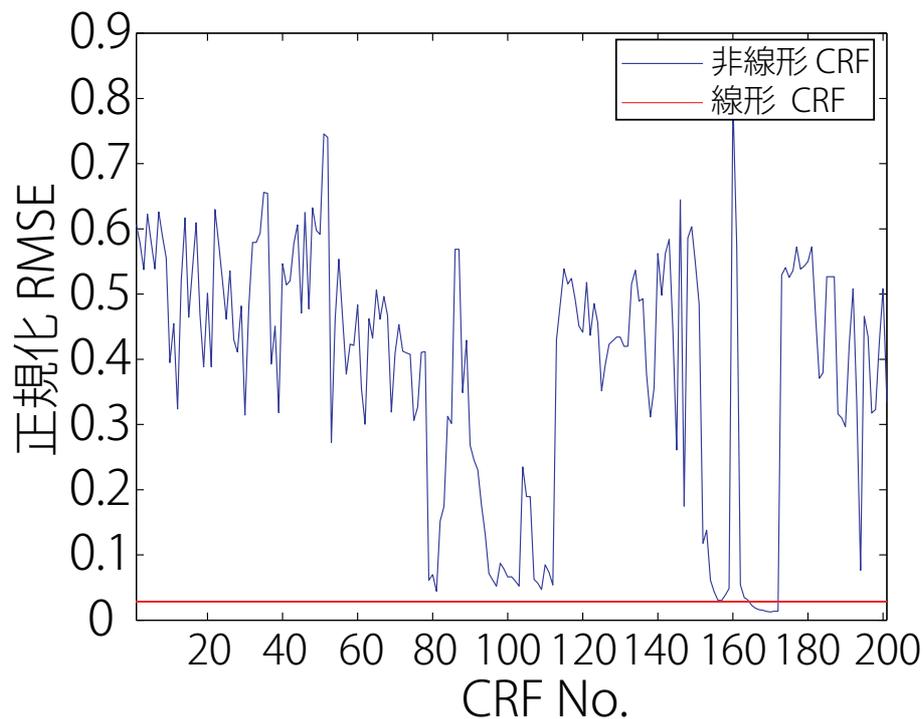


図 4.9: カメラ応答関数と正規化 RMSE

レースケール映像を撮影した．映像は huffyuv コーデックを用いて可逆圧縮した．映像の背景には様々な輝度が観測されるように Macbeth Color Checker Board を被写体とし，300 フレームの映像からノイズレベル関数を推定した．背景映像のノイズ特性と推定されたノイズレベル関数を図 4.10 に示す．

一方，前景には Color Checker Board 上に物体を置き，移動させながら撮影した．線形と非線形のカメラ応答関数を用いて移動物体を撮影し，フレーム合成を行った．この際，明るさやコントラストを調整して見た目が自然な合成を行った．図 4.11 に検証に用いた改ざん映像を示す．画像内上側の物体が改ざんされた被写体で，下側の物体が改ざんされていない物体である．また，四角で示した領域が分散の計算のために指定した領域である．計算領域は  $15 \times 20$  画素の大きさを固定し，位置は対応する箇所を手動で指定した．

図 4.11 に示した 5 つのフレーム内の計算領域からテクスチャ分散の分布を求めた結果を図 4.12 に示す．図より，背景と同じく線形のカメラ応答関数で撮影された下側の物体ではテクスチャ分散が 2 次曲線上に分布し，正規化 RMSE も小さい値を示

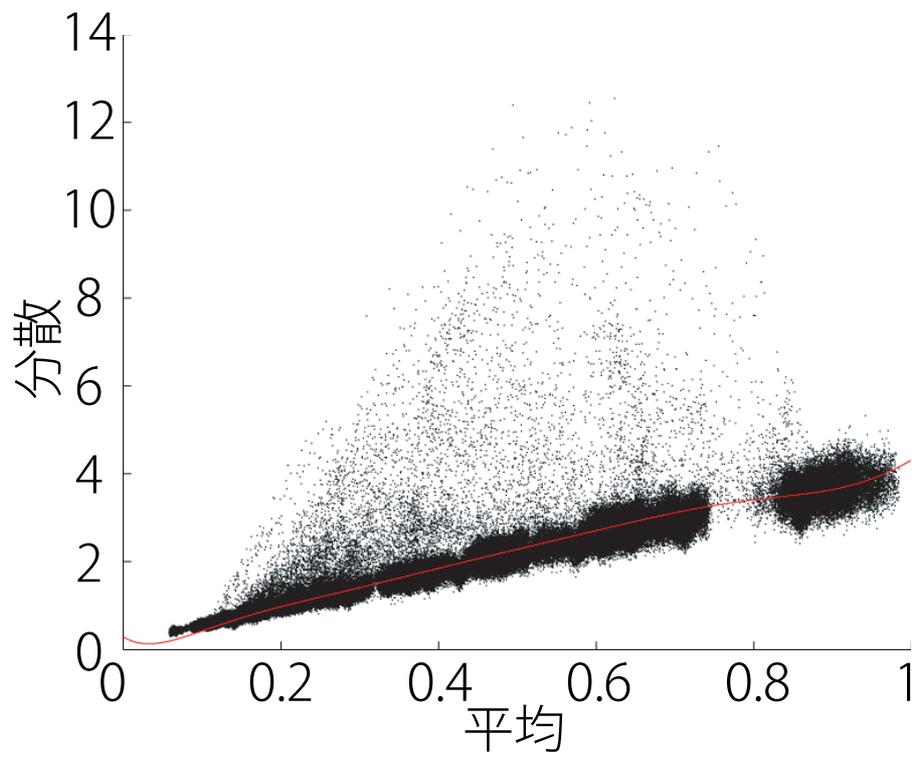


図 4.10: 実映像の背景のノイズレベル関数

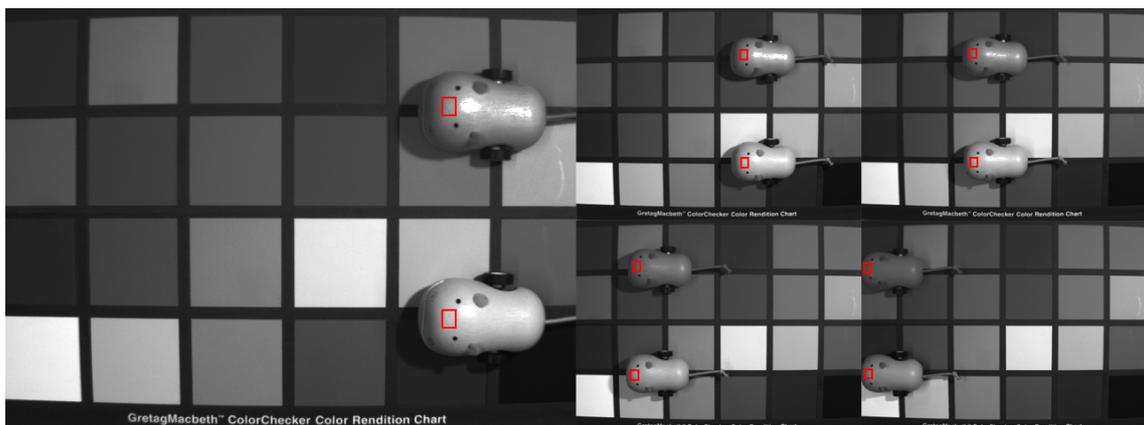
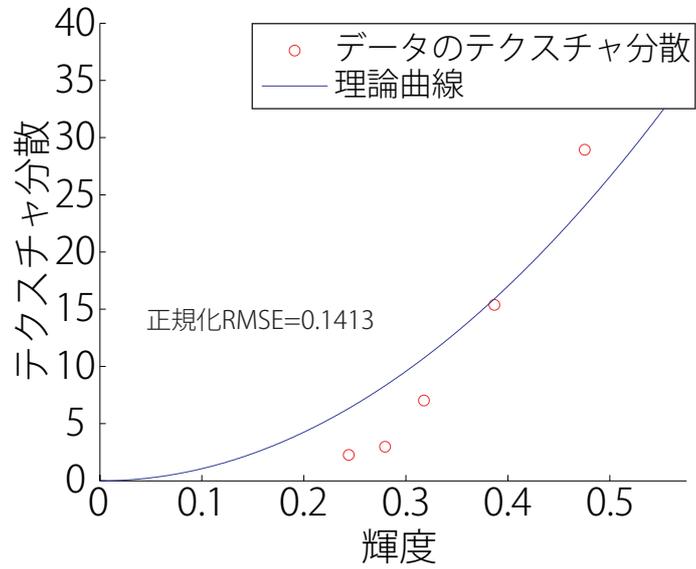
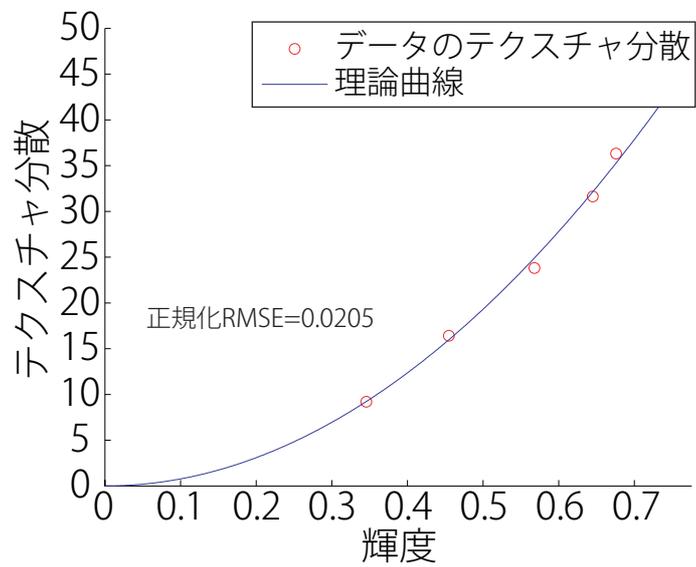


図 4.11: 撮影した実映像の指定フレーム．四角が指定した分散の計算領域



(a) 改ざん箇所のテクスチャ分散



(b) 非改ざん箇所のテクスチャ分散

図 4.12: 実映像のテクスチャ分散

しているが、非線形のカメラ応答関数で撮影された上側の物体ではテクスチャ分散が2次曲線から外れ、正規化 RMSE も大きくなっていることが分かる。本実験では  $15 \times 20$  画素の小さな領域かつ5フレーム分の少ない測定でもカメラ応答関数によるテクスチャ分散の振る舞いの違いを観測することができた。

## 4.4 まとめ

本研究では第3章で述べた静止シーンに対するノイズ特性を手がかりとした改ざん検出手法を動的なシーンに対応させるための拡張を行った。提案手法は移動物体のノイズ特性を推定するために、静止シーンに対する手法のように画素単位で平均と分散を計算せず、領域的な平均と分散を計算することによりノイズ特性を求めた。空間的な分散にはノイズの成分の他に陰影とテクスチャの成分が含まれるため、これらの影響を分離することによって改ざんの手がかりを取得した。改ざんの手がかりとしては、テクスチャ分散と輝度の間に2乗の関係があることを利用した。計算された分散からノイズによる成分を差し引き、輝度とテクスチャ分散の間の関係を検証することにより、前景と背景で撮影条件が異なる箇所を特定することができる。

我々は人工データと実映像を用いて、提案手法の有効性を検証した。様々なカメラ応答関数を用いて取得された映像のテクスチャ分散の関係を検証し、多くのカメラ応答関数において、テクスチャ分散と輝度の間の関係は2乗から大きく外れることを示した。実映像を用いた検証では、ユーザの利便性を含む設定条件の実用性を検証した。カメラを用いて撮影された改ざん映像から、各物体について微小領域を5フレーム指定することにより、改ざんの有無を区別することができた。

今後の課題としては、ユーザの利便性をさらに向上させるため、前景の抽出と領域的な追跡を実装し、改ざんの検出を全自動で行うことが考えられる。前景の抽出は背景差分の既存技術を用いて精度良く行うことができると考えられるが、領域的な追跡は現在発表されておらず、既存技術をそのまま適用できない。また、時間的な連続性を保持した領域追跡は難しい課題であり、実現には時間がかかると予想される。また、提案手法は改ざんの指標を正規化 RMSE によって評価するが、正規化

RMSE は改ざんのモデルにもとづいた確率的な指標ではないため，改ざんの度合いを絶対的に評価する指標の導入が望まれる．

# 第5章 計測誤差を考慮したアルゴリズムの提案

## 5.1 研究背景

コンピュータビジョンを含む多くの計算機科学の研究では、提案するアルゴリズム内の値の一部を既知と仮定することが多い。既知と仮定したパラメータは評価実験の際には予め校正によって値を取得しておく。例えば、ステレオビジョンでは、複数のカメラの相対的な位置関係を表す外部パラメータを校正によって予め取得しておく。カメラの外部パラメータや内部パラメータを校正する手法は数多く提案されてきている [45, 98, 103, 119]。

校正によって得られたパラメータを用いた未知変数の推定において注意しなければならない点として、校正自体にも誤差が生じることが挙げられる。図 5.1 に誤差の種類をまとめる。校正には巻尺や光度計のように求める量を直接計測する場合と、上で述べたカメラの外部パラメータのように間接的に測定する場合があるが、それぞれの計測方法における誤差について次にまとめる。

直接計測では、計測の際に生じる計測誤差が校正の精度を下げる要因となる。計測誤差は、誤差が生じる原因によって系統誤差と偶然誤差に分けられる。系統誤差は、測定者の癖や測定器の性能、測定条件などの原因による、計測値の真値からのずれである。系統誤差は決定的な要因によって生じる誤差であるため、誤差が生じる原因を正しく特定し除去することによって誤差を減少させることができる。偶然誤差は、測定の度にばらつく誤差である。例えば、秤を用いた質量の計測では、秤を置いた台の微妙な振動により測定の度に微妙に異なる値が観測される。偶然誤差は系統誤差のようにその原因を取り除くことはできないが、真の値に対して対称に

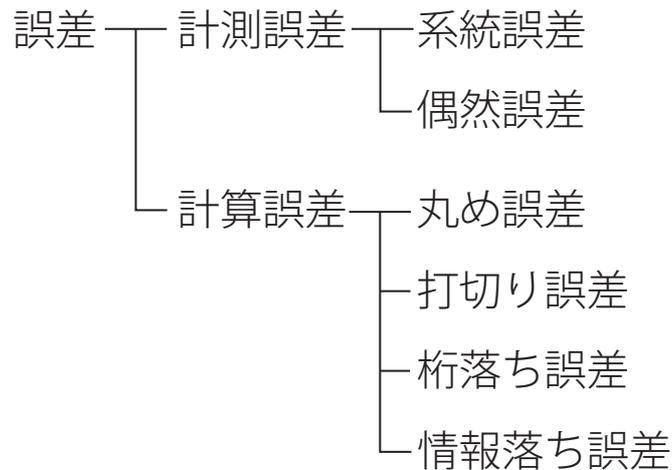


図 5.1: 誤差の種類

分布すると仮定し，複数回の測定を行いその平均値を用いることによって影響を抑えられることが多い。

計算機を用いた間接計測では，計測誤差に加えて計算誤差が影響を与える場合がある。計算誤差は計測された値から求めたい値を計算によって求める際に生じる誤差で，丸め誤差，打切り誤差，桁落ち誤差，情報落ち誤差に分けられる。丸め誤差は，有限の計算精度に起因する誤差である。計算機を用いた計算では，実数は浮動小数点型のデータとして扱われる。しかし，有限のビット列では実数を無限精度で正確に表現することはできないため，実数が近似されることによって生じる誤差である。打切り誤差は，無限級数や収束などの反復計算を無限ループさせないため，計算を途中で打ち切ることによって生じる誤差である。桁落ち誤差は，近い値を持つ浮動小数点数同士の減算などにより，有効桁数が小さくなることによって生じる誤差である。情報落ち誤差は，絶対値の非常に大きな数と小さな数の加減算を行ったとき，絶対値の小さい数が無視されてしまうために生じる誤差である。

誤差を小さくするためには様々な手法がある。計算誤差はアルゴリズム内で計算の順序を変えるなどの工夫によって影響を小さくする(増幅させない)ことができる場合がある。また，精度の高い計測装置を用いる，複数回の計測を行う，計算に用いるデータのビット数を増やすなどの処置により，誤差を小さくすることができる。しかし，これらの手法は経済的・時間的・計算空間的なコスト増加を招き，実用の

際には限界が生じる。また、計測精度は環境ノイズレベル以上には上げられないため、精度の限界問題は常に付きまとうことになる。

以上のような問題を解決するために、校正誤差を考慮したアルゴリズムが重要となってくる。多くのアルゴリズムでは、測定値に対してガウスノイズが混入したと仮定し、最尤推定によって真の値を推定する。正規分布は平均と分散が与えられた実数値確率分布における最大エントロピー分布である [10]。また、中心極限定理により、無作為抽出された標本の分布は抽出数が大きくなるにつれて正規分布に近づくことが知られている。よって、観測値が実数値であること以外に事前情報が何も無い場合、観測値の従う分布として正規分布が最も妥当であると考えられる。これが、一般に計測誤差に対して正規分布を仮定する理由である。

観測値が一定値の場合、正規分布を仮定したときの最尤値は標本平均に一致する [10]。これは偶然誤差を減少させるために複数回計測した値の平均を求めることの妥当性を裏付ける結果となっている。また、ある系の入出力データのサンプルからこの系の写像関数を回帰推定する場合、正規分布を仮定した確率モデルの最尤解は最小二乗法によって求めた解と同値である。このように、経験的に用いられてきた手法が、実はある誤差モデルの仮定の下での最尤解を与えるものであったという事実は興味深い。

計測誤差を考慮することによる精度向上は、理論体系としての誤差論だけではなく、計算機科学以外の応用分野でも研究されている。観測誤差を考慮した推定手法としては、カルマンフィルタ [54] が用いられることが多い。カルマンフィルタは、ある「状態」が時間的に遷移し、さらにその状態から何らかの写像の形で「観測量」が生成されると仮定する。そして、現時点での観測量と直前の時刻での状態推定値から、現時点での状態推定値と直後の時刻の状態予測値を求める。カルマンフィルタに類似する概念として、隠れマルコフモデル [25] がある。カルマンフィルタは状態変数が連続値であるが、隠れマルコフモデルでは離散値を取る。また、カルマンフィルタはガウスノイズが混入すると仮定する一方、隠れマルコフモデルでは任意の確率分布で状態の変化を記述することが可能な点が異なる。

実数値確率変数が従う分布は、事前情報が他に無ければ正規分布の仮定が最も妥当であることは既に述べた。しかし、カルマンフィルタや多くの誤差論では暗黙的に正規分布の仮定が用いられており、その妥当性には議論の余地があると言える。例えば、多くの測定機器では測定可能範囲の上下限が存在するため、この境界付近で正規分布する物理量は測定によって飽和を生じ、観測値が従う分布は切断正規分布となる。また、整数値確率変数が従う分布は、実数と近似できるオーダーになれば二項分布やポアソン分布などの整数値確率変数に対して定義された確率分布で表現する方が適当である。より適切な誤差モデルと分布を用いてアルゴリズムを再構成することにより、測定機器や測定方法のコストを増やすことなく、アルゴリズムの推定精度を向上させることができると考えられる。

本研究では、観測誤差が生じる背景にある母集団が従う確率モデルを考慮することにより、推定精度を向上させたアルゴリズムを提案する。具体的にはコンピュータビジョンの基本問題の1つである照度差ステレオからの3次元形状復元に着目する。照度差ステレオは、異なる光源環境下で撮影された画像から、被写体の法線を推定する手法である。照度差ステレオの手法の多くは、光源方向が既知であると仮定し、実際の実験では予め校正によって光源方向を取得しておく。しかし、校正にも誤差が含まれるため、照度差ステレオから推定された法線には誤差が混入し、さらに復元された3次元形状にも伝播することになる。

法線から計算された2次元勾配から、3次元形状を復元する手法は数多く提案されている。これらの手法は勾配に対してガウスノイズが混入したと仮定し、最尤推定から導出されたポアソン方程式 [1, 42, 95] や基底関数 [37, 40, 56, 62] を用いて適切な3次元形状の復元を試みている。しかし、勾配に対するガウスノイズの仮定は経験的な誤差モデルであり、その妥当性は明らかになっていない。したがって、光源方向の校正誤差が形状復元に与える影響を正しくモデル化することにより、形状復元の精度を向上させることができると考えられる。

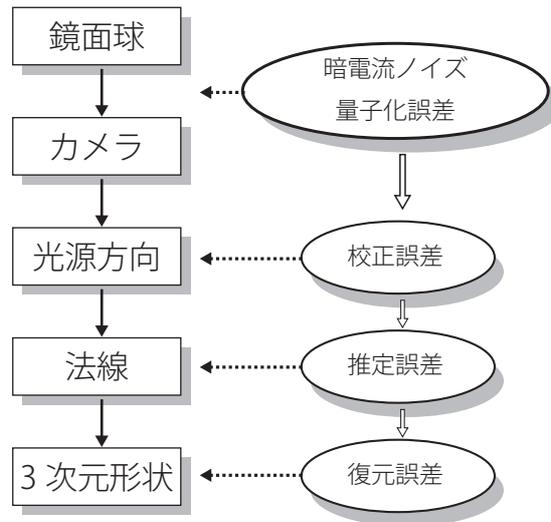


図 5.2: 照度差ステレオのノイズ伝播

## 5.2 提案手法の概要

計測誤差を考慮したアルゴリズムの精度向上の対象として、本研究では照度差ステレオにおける校正誤差に着目する。照度差ステレオ [113] は、様々な方向に配置された光源環境下で撮影された複数の画像から、物体の法線と反射率を推定する手法である。校正済みの照度差ステレオでは、光源の方向と強度は鏡面球や拡散球を用いて校正される。光源方向は、鏡面球の反射光のピークを検出することによって求めることができる。光源方向が正確に校正されれば正確な法線が推定されることが期待できるが、一般に画像ノイズなどの計測誤差や量子化誤差などの計算誤差が生じるために正確に光源方向を校正することは難しい。そのため、光源の校正に誤差が生じ、それが法線推定に伝播することは避けられないものと言える。

図 5.2 に、校正誤差が 3 次元形状の復元誤差まで伝播する過程を示す。鏡面球による光源方向の校正では、非常に暗い光源環境で画像を撮影するため、暗電流ノイズの影響が大きい。また、信号の標本化や画素値の量子化による誤差も校正に影響を与える。校正のために撮影された画像に混入する誤差は、光源方向の校正に誤差を与える。光源方向の校正誤差は照度差ステレオによって推定される物体の法線に伝播し、最終的に復元された 3 次元形状に影響を与えることになる。

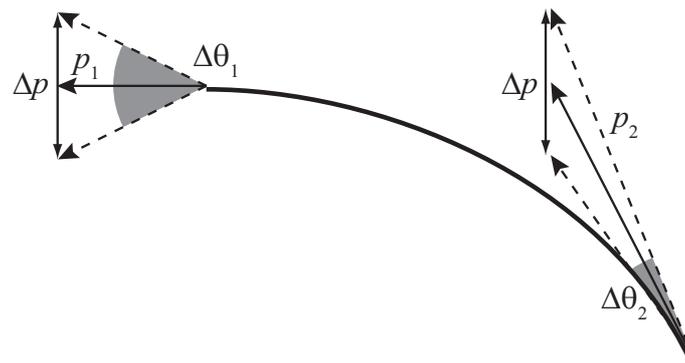


図 5.3: 勾配に対するノイズ  $\Delta p$  が法線方向に与える影響．水平な箇所における誤差角  $\Delta\theta_1$  は、勾配が大きい箇所における誤差角  $\Delta\theta_2$  よりも大きい．

照度差ステレオで推定される情報は物体の局所的な面方向成分であるため、物体の 3次元形状を復元するためには積分が必要になる．形状復元の手法は数多く提案されてきているが、ノイズの影響は法線から 3次元形状を復元する際に顕著になるため、復元誤差を減らす形状復元の手法が重要となる．Fraile と Hancock は最小全域木で構築した積分経路上で線積分する手法を提案した [36]．また、図 5.3 に示すように勾配に対してガウスノイズ  $\Delta p$  が混入するノイズモデルを仮定し、最尤推定の枠組みで形状を復元する手法が提案されている [1, 95]．他にも、ノイズの影響により可積分性を満たさない勾配場が与えられたとき、復元形状を様々な可積分な基底関数の線形結合として表現することにより形状復元を行う手法が提案されている [37, 40, 56, 62]．

しかし、既存手法の問題点として、ガウスノイズモデルが実際に正しいモデルであるかどうかは明らかになっていない点が挙げられる．また、図 5.3 に示すように、2つの勾配  $p_1, p_2$  が同じ程度の不確定性  $\Delta p$  を持っていると仮定する．このとき、水平な箇所における法線の揺らぎ  $\Delta\theta_1$  は、勾配が大きい箇所における揺らぎ  $\Delta\theta_2$  よりも大きいことが分かる．逆に、法線が同じ角度だけ不確定性を持っていたとき、この不確定性が勾配に与える影響は、勾配の大きい箇所の方が大きい．つまり、法線の不確定性と勾配の不確定性の関係性は、位置によって異なることになる．しかし、既存の手法ではこの関係性を考慮していない．

一方本研究では、ランバートな反射特性をもつ物体を対象として、照度差ステレ

オにおける光源方向の校正誤差の影響について考察し、この誤差を考慮に入れた形状復元の手法を提案する。はじめに、光源方向の校正誤差と照度差ステレオによって推定された法線方向の誤差の振る舞いを検証する。ここで光源方向や法線は3次元単位ベクトルで表されることに着目する。実数値確率変数に対する最大エントロピー分布は正規分布であったが、3次元単位ベクトルに対する最大エントロピー分布はフィッシャー分布であることが知られている [89]。そこで、まず光源方向の校正誤差と法線の推定誤差がフィッシャー分布に従うことを示す。法線の推定誤差の分布がフィッシャー分布で表せることを示した後、フィッシャー分布にもとづくノイズモデルを導入し、最尤推定を用いて形状復元を行う。校正誤差を正しく扱うことにより、提案手法が従来手法に比べてより正確に形状を復元できることを示す。

実験では、人工データを用いた定量的な評価に加え、実画像を用いた定量的評価と定性的評価によって提案手法の有効性を検証する。

## 第6章 光源方向の校正誤差を考慮した 照度差ステレオからの形状復元

### 6.1 はじめに

本章では、光源方向の校正誤差を考慮した照度差ステレオからの3次元形状復元手法の提案と評価を行う。

一般の照度差ステレオでは、画像を撮影する際に用いた光源の方向を予め校正によって求めておき、アルゴリズム内では既知の値として取り扱う。しかし、光源方向の校正には誤差が混入するため、物体の法線の推定値にも誤差が伝播し、最終的な出力である復元形状にも影響を与えることになる。鏡面球を用いた光源方向の校正の手法を図6.1に示す。鏡面球は光源の光を鏡面反射するため、光源方向を $\mathbf{l}$ 、反射位置での法線を $\mathbf{n}$ 、視線方向を $\mathbf{v}$ とすると、

$$\mathbf{v} = 2(\mathbf{n}^T \mathbf{l})\mathbf{n} - \mathbf{l} \quad (6.1.1)$$

の関係がある。ここで、視線方向を $\mathbf{v} = [0, 0, 1]^T$ とし、また反射位置での法線は鏡面球の形状が既知であることから求めることができるため、画像上での反射位置が分かれば光源方向を推定することができる。画像上での反射位置は広がりを持って観測されるが、重心位置を計算することでピーク位置をサブピクセル精度で求めることができる。

鏡面球を用いた校正は、非常に暗い環境で画像を撮影するため暗電流ノイズの影響が大きく、また画像が取得される際に標本化・量子化されることによる誤差が生じる。また、光源方向の校正は間接計測となるため、計算誤差が入る可能性がある。以上のような原因により、光源方向の校正値には誤差が混入することになる。

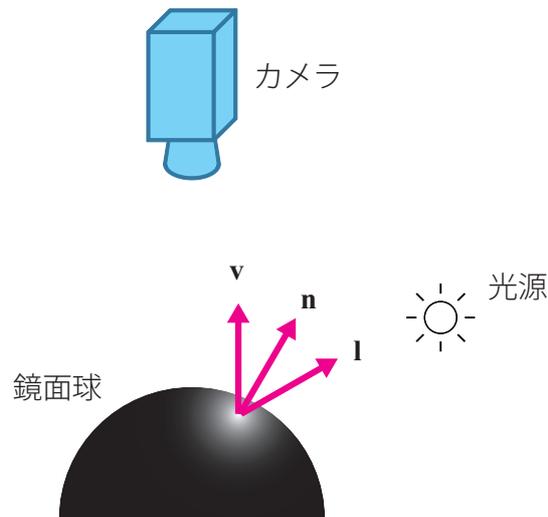


図 6.1: 鏡面球を用いた光源方向の校正

光源方向と物体の法線は、3次元の単位ベクトルで表現される。したがって、ベクトルの各成分は独立に揺らぐのではなく、ベクトルの長さを1に保つように拘束されることになり、その自由度は2である。このとき、単位ベクトルは単位球(ガウス球)の表面を回転するように揺らぐ。すると、3次元ベクトルの誤差モデルで一般に用いられる3次元正規分布を単位ベクトルの誤差モデルとして用いることはできない。 $n$ 次元正規分布に類似した特性を持つ、 $n$ 次元単位ベクトルが従う分布として、フォン・ミーゼス・フィッシャー分布 [75] がある。フォン・ミーゼス・フィッシャー分布は、 $(n-1)$ 次元超球面上に定義された確率分布である。特に、フォン・ミーゼス・フィッシャー分布の中で3次元単位ベクトルに対する分布をフィッシャー分布 [35] と言う。フィッシャー分布は3次元単位ベクトルに対する最大エントロピー分布であり、確率変数に関する事前情報が他にない場合に仮定する確率分布として最も妥当な分布である。

本研究では、ランバートな反射特性をもつ物体を対象とし、光源方向と法線の誤差がフィッシャー分布するという仮定の下で確率モデルを構築する。まず光源方向の校正において生じる誤差により、光源方向の校正値がフィッシャー分布することを確かめる。次いで、フィッシャー分布する光源方向の校正値を用いて照度差ステレオを行い、推定された法線がフィッシャー分布することを確かめる。法線の誤差

分布がフィッシャー分布で表現できることを確かめた後、フィッシャー分布にもとづく確率モデルを構築する。そして、最尤推定の枠組みにより3次元形状の復元を行う手法を提案する。

本研究では、勾配に対するガウスノイズを仮定した形状復元の先行研究とは異なり、照度差ステレオにおける誤差モデルにもとづく形状復元を提案する。既存手法に対する提案手法の有効性を、人工画像と実画像を用いて定量的・定性的に評価する。

本章の構成を以下に示す。まず、6.2節において照度差ステレオの拡張と誤差解析および3次元形状の復元に関する先行研究について述べる。6.3節で照度差ステレオの校正誤差のノイズ特性について記述し、ノイズモデルにもとづいた形状復元の手法を6.4節で提案する。6.5節では様々な従来手法と形状復元の精度を比較し、最後に6.6節で結論を述べる。

## 6.2 関連研究

本節では照度差ステレオの誤差解析と3次元形状の復元手法に関する先行研究を紹介する。照度差ステレオは古くから研究されているコンピュータビジョンの基本問題の1つであるが、研究の中心は光源方向の校正を必要としない手法の提案やより多くの物体の反射モデルへの拡張であり、手法内の誤差の解析についてはあまり研究されていない。一方、2次元勾配場からの3次元形状の復元手法は、照度差ステレオに限らず計算機科学の重要な分野の1つであり、様々な手法が提案されている。

### 6.2.1 照度差ステレオの拡張と誤差解析

Woodham [113] によって照度差ステレオが提唱された後、照度差ステレオは主に2つの軸で発展してきている。1つは光源方向の校正を必要としない照度差ステレオの提案である。物体の法線と光源方向が共に未知の場合、Generalized Bas-Relief (GBR) ambiguity と呼ばれる不定性が生じる [8]。GBR 不定性を解決するために、特異値分解 [26, 43, 118]、透視投影 [64]、球面調和関数 [5, 6, 87]、相互反射 [16] などを

利用する手法が提案された。もう1つの軸は、様々な反射特性をもつ物体に対して照度差ステレオを適用するための反射モデルの拡張である。初期の照度差ステレオは拡散反射成分のみを持つ単純な反射モデルを仮定したが、双方向散乱面反射率分布関数 (Bidirectional Reflectance Distribution Function: BRDF) [18] や二色性反射モデル [74, 88, 90] などに拡張した手法が提案されている。

一方、本研究の主張である誤差やノイズを考慮した照度差ステレオの改良に関する先行研究はまだ発展途中である。例えば、光源の配置に着目し、画像にノイズが混入したときにその影響を最小化する光源配置に関して、3光源の場合と一般の光源の場合で様々な考察がされている [24, 91, 97, 100]。これらの研究で行われた理論的・実験的考察では、仰角を一定にし円周上に均等に光源を配置した場合が誤差を最小化するという直感的な結論が得られている。

また、Kuparinen らは近似的に平面である物体を対象とし、画像ノイズとぼけを考慮した最適な形状復元の手法を提案した [65]。平面を仮定することでぼけの影響を削減し、さらにウィナーフィルタ [110] によってノイズとぼけを除去することにより、形状復元の精度が向上することを示した。

先行研究の提案する手法では、画像に混入するノイズやぼけを考慮することによって、照度差ステレオから復元された3次元形状の精度が向上してきている。しかし、画像のノイズやぼけの除去は画像解析の一分野として確立されている研究対象であり、既存手法を用いて予めノイズやぼけを取り除いた後に照度差ステレオを行うという手法も考えられる。一方、照度差ステレオに固有のノイズである校正誤差を対象とした精度向上の手法はまだ提案されていない。したがって、画像の品質向上に加えて本研究が提案する校正誤差を考慮した手法を用いることにより、従来の復元精度をさらにもう一段階向上させることが可能であると考えられる。

### 6.2.2 勾配からの3次元形状復元

本節では、2次元勾配から3次元形状を復元する先行研究を紹介する。照度差ステレオからの出力は法線であるので、法線から勾配を計算し、積分を行うことで3次

元形状を得ることができる。また、これらの手法は照度差ステレオからの3次元形状復元だけではなく、影除去 [33] や High Dynamic Range 画像の作成 [30] の他、勾配画像から強度画像を復元する際 [84, 99] にも応用できる技術である。ここで、推定された勾配場には誤差が含まれているため、この誤差が復元した形状に与える影響を最小化する必要がある。多くの既存手法では勾配場にガウスノイズが混入したと仮定し、入力勾配場にできるだけ近い勾配場をもつ形状を最尤推定によって求める。

形状を復元する最も単純なアプローチとして、勾配の値をある経路に沿って足し合わせる方法がある。Fraile と Hancock は、画像をグラフ構造として捉え、エッジの両側の法線の差にもとづいてエッジの重みを定義した [36]。そして、エッジの重みの和を最小とする最小全域木を積分の経路として用い、勾配を足し合わせることで形状を復元した。この手法は曲率の大きいエッジが経路に採用されるのを避け、曲がった形状を平面に近似して高さを求める際の誤差を最小化することができる。しかし、各ノードの高さを拘束する式が1つしかないため、ノイズや外れ値に対して敏感であるという問題点を抱える。

一方、ノイズを考慮した形状復元手法において、典型的なノイズモデルは勾配に対するガウスノイズモデルである。特に、ノイズの大きさが空間的に一様で、かつ勾配の  $x, y$  成分に等しい大きさのノイズが混入するような等方的なガウスノイズモデルを仮定すると、最尤解はポアソン方程式を解くことによって求まることが知られている。ポアソン方程式の数値解法は、変数の反復的な更新による古典的な手法 [42] の他に、固有値問題に落とし込むことで直接求める手法 [95] が提案されている。

ガウスノイズの仮定は非一様・非等方的なノイズに拡張することができる。Agrawal らは形状復元の既存の手法が拡張されたポアソン方程式の特別な場合として記述できることを主張した [1]。また、拡張されたポアソン方程式の実装の一例として、 $\alpha$ -surface と呼ばれる2値の重み付けを用いた形状復元の手法を提案した。 $\alpha$ -surface では RANSAC [34] と同様に与えられた勾配場の中から反復的にインライアと判定さ

れた値を抽出して形状復元を行う。RANSACでは形状復元を行う値をランダムに選択するが、 $\alpha$ -surfaceでは勾配の誤差がある許容値 $\alpha$ 以下である値を繰り返し加えることにより、決定論的にインライアを選択する点異なる。

Agrawalらは同論文[1]で、以下に示すような画像のノイズ除去などに用いられるエネルギー最小化手法が形状復元に応用できることを示した。

- M-estimator: 連続的な重み付けを用いたロバストなエネルギー最小化の手法の一つ。Huber, Cauchy, Tuckey関数などを用いて反復的に重みを更新する。
- Regularization: 画像ノイズの除去における不良設定問題を解くために正則化項を付け加えたエネルギー最小化手法[17]。エネルギー関数が非線形となるため、half-quadraticなエネルギー最小化を行う。
- Anisotropic Diffusion: 同じく画像ノイズの除去のために提案された手法[85]。エッジを保存するように、ポアソン方程式がアフィン変換によって変形される。

奥行きに不連続な箇所がない形状は可積分性を満たす(勾配場の回転が至る所で0)。しかし、照度差ステレオから得られる勾配場は一般に可積分性を満たさない。勾配場に可積分性を課す手法の一つとして、復元される形状を可積分性を満たす基底関数の線形結合として表現するアプローチがある。FrankotとChellappaはフーリエ関数を基底関数として使い、復元形状の勾配場と入力勾配場が近くなるような結合係数を推定した[37]。また、他の基底関数として余弦関数[40]、デルタ関数[56]、Shapeletと呼ばれる直交条件を満たさない基底関数[62]などを用いた手法が提案されている。

既存手法は照度差ステレオから得られる勾配に対してガウスノイズが混入する仮定のもとで形状を復元する。しかし、5.1節で述べたようにガウスノイズの仮定が正確かどうかはまだ明らかになっていない。一方、本手法は照度差ステレオにおける光源方向の推定に誤差が発生するという仮定の下で、法線ベクトルに対する誤差の分布を検証する。

また、形状復元のもう一つの問題である奥行きに不連続な箇所がある場合に対処

するために, Karacali と Snyder は部分可積分性の概念を導入した [56]. 不連続境界が既知であると仮定したとき, その両側で個別に結合係数を推定することにより, 不連続境界を正しく復元することができる. ただし, 本研究では連続な形状を対象としているため, 不連続境界の扱いについては言及しない.

### 6.3 照度差ステレオにおける誤差の解析

本節では, 照度差ステレオにおける光源方向の校正誤差と法線ベクトルの推定誤差がフィッシャー分布に従うことを示す. 6.3.1 節では照度差ステレオの基本的な原理について述べる. 6.3.2 節でフィッシャー分布の性質について述べ, 光源方向の校正誤差がフィッシャー分布に従うことを示す. また, 照度差ステレオから推定された法線がフィッシャー分布することを 6.3.3 節で示す.

#### 6.3.1 照度差ステレオ

$N$  個の光源下で観測されたランバートな表面特性をもつ物体を仮定すると, 観測される輝度  $\mathbf{i} = [i_1, \dots, i_N]^T$  は

$$\mathbf{i} = L\mathbf{b} \quad (6.3.1)$$

で表される. ここで  $L = [\mathbf{l}_1, \dots, \mathbf{l}_N]^T$  は光源方向を示す行列で,  $i$  番目の行ベクトルが  $i$  番目の光源の方向を表す. ここで, 全ての光源の強度は正規化されている, すなわち  $\|\mathbf{l}_i\| = 1$  ( $i = 1, \dots, N$ ) と仮定する. また,  $\mathbf{b} = \rho\mathbf{n}$  はスケールを考慮した法線ベクトルで, 反射率  $\rho$  と法線ベクトル  $\mathbf{n} = [n_x, n_y, n_z]^T$  の積で表される. もし  $N \geq 3$  枚の画像が与えられたとすると, (6.3.1) 式は解を持ち,  $\mathbf{b} = L^\dagger\mathbf{i}$  として得られる [113]. ここで  $L^\dagger = (L^T L)^{-1} L^T$  は  $L$  の擬似逆行列である.  $\mathbf{n}$  は定義より単位ベクトルであるので,  $\rho$  と  $\mathbf{n}$  は

$$\rho = \|\mathbf{b}\|, \quad \mathbf{n} = \mathbf{b}/\rho \quad (6.3.2)$$

で求めることができる.

光源方向行列  $L$  を求めるためには、予め校正が必要である。本研究では光源方向の校正に生じる誤差とその法線への影響に着目する。光源方向ベクトル  $l$  と法線ベクトル  $n$  は単位ベクトルであるので、 $x, y, z$  成分はノイズによって各々独立に動くのではなく、単位球面上を回転する。この回転誤差については次節で詳しく述べる。

### 6.3.2 照度差ステレオの校正誤差

光源方向を記述するための方法は様々であるが、は直交座標系のベクトルを用いると、3次元の単位ベクトルとして記述できる。先に述べたように、単位ベクトルの各成分はノイズによって各々独立に変化するのではなく、ベクトルの長さを1に保つように単位球面上を回転する。したがって、3次元正規分布を用いて単位ベクトルの誤差モデルを表現することは不適である。そこで我々は3次元単位球面上で定義された正規分布と類似した特性をもつフィッシャー分布 [35] を用いて単位ベクトルの揺らぎを記述する。

フィッシャー分布は単位ベクトルが真の方向からずれる角度誤差の確率分布を表す分布である。図 6.2 で示すような光源方向  $l = [l_x, l_y, l_z]^T$  の校正誤差が以下の確率密度分布の式

$$f_{\text{cart}}(\mathbf{l}; \hat{\mathbf{l}}, \kappa) = \frac{\kappa}{4\pi \sinh \kappa} \exp[\kappa \mathbf{l}^T \hat{\mathbf{l}}] \quad (6.3.3)$$

で表されるとき、ベクトル  $l$  はフィッシャー分布に従う。ここで  $\hat{\mathbf{l}}$  は真の光源方向で、 $\kappa$  は精度パラメータである。フィッシャー分布するノイズによって揺らぐ  $K$  個の単位ベクトルの標本  $\{l^{(1)}, \dots, l^{(K)}\}$  が与えられたとき、精度パラメータ  $\kappa$  は次式

$$\kappa = \frac{K-1}{K-R} \quad (6.3.4)$$

によって推定することができる [61]。ここで、 $R$  は標本の合成ベクトルの長さで、

$$R^2 = \left( \sum_{i=1}^K l_x^{(i)} \right)^2 + \left( \sum_{i=1}^K l_y^{(i)} \right)^2 + \left( \sum_{i=1}^K l_z^{(i)} \right)^2 \quad (6.3.5)$$

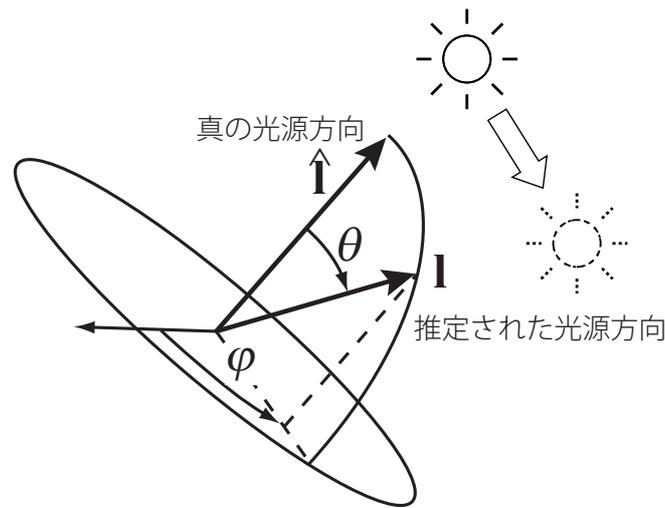


図 6.2: ノイズによって回転する単位ベクトル . 推定された光源方向  $\mathbf{l}$  は , フィッシャーノイズによって真の光源方向  $\hat{\mathbf{l}}$  から回転して観測される .  $\mathbf{l}$  は  $\hat{\mathbf{l}}$  を軸として対称に分布するため , 誤差の大きさは方位角  $\varphi$  に依存しない .

で計算される .  $\varphi$  と  $\theta$  をそれぞれ真の方向から測った方位角と極角とすると , (6.3.3) 式は極座標系を用いて

$$f_{\text{sph}}(\varphi, \theta; \kappa) = \frac{\kappa}{4\pi \sinh \kappa} \exp[\kappa \cos \theta] \quad (6.3.6)$$

と書き直せる . (6.3.6) 式は方位角  $\varphi$  に依存せず , 真の方向  $\hat{\mathbf{l}}$  の周りで対称に分布していることに注意する . 極角が  $\theta$  から  $\theta + d\theta$  の範囲で観測される確率から , 極角  $\theta$  に関する確率密度関数  $f_{\text{polar}}(\theta; \kappa)$  が

$$\begin{aligned} \int_{\partial S} f_{\text{sph}}(\varphi, \theta; \kappa) dS &= \int_{\varphi=0}^{2\pi} \frac{\kappa}{4\pi \sinh \kappa} \exp[\kappa \cos \theta] d\varphi d\theta \sin \theta \\ &= \frac{\kappa}{2 \sinh \kappa} \exp[\kappa \cos \theta] \sin \theta d\theta \\ \therefore f_{\text{polar}}(\theta; \kappa) &\equiv \frac{\kappa}{2 \sinh \kappa} \exp[\kappa \cos \theta] \sin \theta \end{aligned} \quad (6.3.7)$$

と得られる . ここで  $dS = d\varphi d\theta \sin \theta$  は極座標系における面積要素で ,  $\partial S$  は  $[\theta, \theta + d\theta]$  の範囲の微小な帯の領域を表す . 我々は以後の角度誤差ヒストグラムに対して  $f_{\text{polar}}(\theta; \kappa)$  を当てはめる .

6.1 節で述べたように , 照度差ステレオでは一般に光源方向が既知であると仮定し , 鏡面球を用いて校正を行うことが多い . 校正では , 鏡面反射の中心を求めるこ

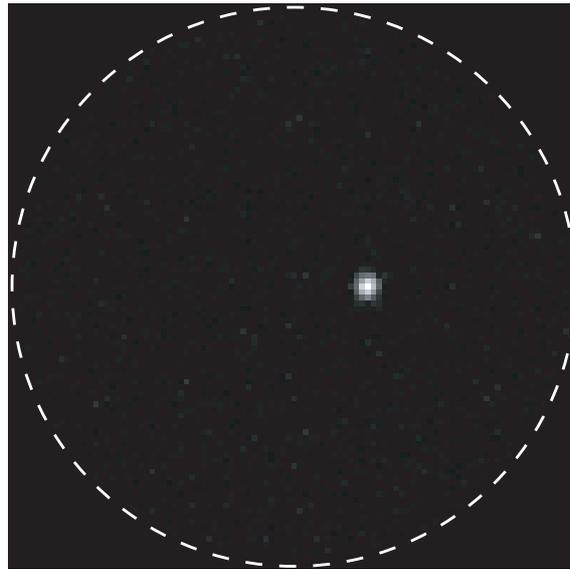


図 6.3: 合成した鏡面球の一例．白い破線が鏡面球の輪郭を示す．

とにより，光源方向を知ることができる．

角度誤差の分布を解析的に導出することはできないため，本論文では数値的に角度誤差の分布を得ることとする．はじめに誤差分布の直感的な説明を行う．校正において，画像に混入するノイズのため，計算された鏡面反射の中心は実際を中心位置からわずかにずれる．そしてこのずれは光源方向の推定を真の方向から回転させる．光源方向の回転量は画像ノイズの増加に伴って増加する．画像のノイズをガウスノイズと仮定すると，推定方向の分布は真の方向を中心とした釣鐘型の形状となることが予想される．我々は確率変数に関するこれ以上の事前情報を持っていないため，3次元単位ベクトルに対する最大エントロピー分布であるフィッシャー分布を用いて誤差をモデル化することが最も妥当であると判断する．

我々は光源方向の推定誤差分布を求めるため，図 6.3 に示すように実環境での校正を模して平行光源下の鏡面球の画像を合成した．カメラ内の信号処理の過程で画像に混入するノイズとして暗電流ノイズと量子化誤差を仮定し，観測された画素値に対してガウスノイズを付加した．合成した画像に対して上で述べた光源方向の校正手法を適用した．1,000回の試行によって得られた真の光源方向と推定された光源方向との誤差角のヒストグラムを図 6.4 に示す．赤線は(6.3.7)式で示した極角に対

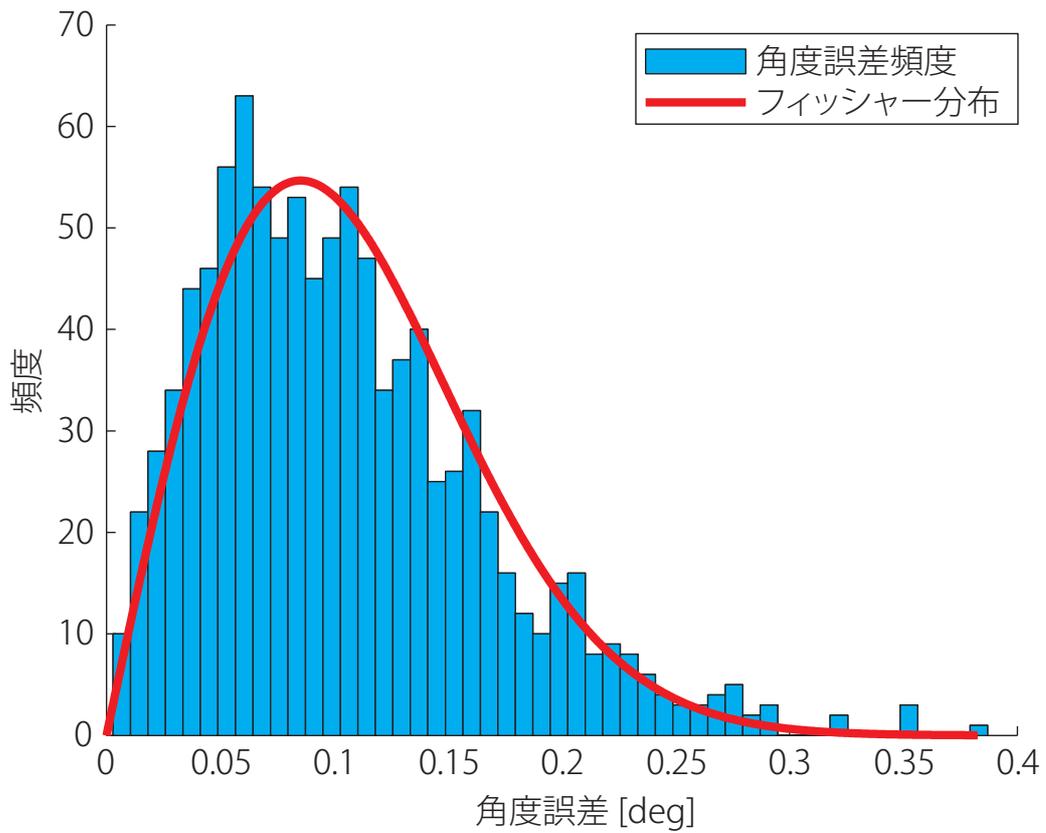


図 6.4: 光源方向の校正誤差のヒストグラムとフィッシャー分布の比較 .

する確率密度曲線  $f_{\text{polar}}(\theta)$  である。精度パラメータ  $\kappa$  は、(6.3.4) 式によって推定した。図より、画像ノイズを含んだ鏡面球による光源方向の校正では、フィッシャー分布に従う誤差角の確率分布が得られることが分かった。

### 6.3.3 校正誤差の法線への影響

前節において光源方向の校正誤差がフィッシャー分布することを示した。続いて、この校正誤差が照度差ステレオによって推定される法線に与える影響について考察する。

前節で述べたようなフィッシャー分布の特徴から、推定された光源方向は真の光源方向の周りに分布する。また、画像ノイズと光源方向の関係と同様に、光源方向の校正誤差が大きいほど、法線の推定誤差も大きくなることが予測される。したがって、法線の推定も真の方向を中心とした釣鐘型、すなわちフィッシャー分布で表すことができると考えられる。

我々はある光源配置の下で (6.3.1) 式を用いてシーンを合成した。合成された画像に対し、校正誤差を模して標準偏差が5度のフィッシャーノイズを付加した光源方向行列を用いて法線を推定した。

図 6.5 に法線の誤差角ヒストグラムと当てはめた確率密度曲線を示す。また、上に示した画像平面から見た極座標に、シーン合成に用いた光源の方向を示す。ヒストグラムは対応する法線の方向についての誤差角を示したものである。代表的な法線の方向として、(a) 視線方向、(b) ある光源方向の近く、(c) 中程度の仰角の3種類についてヒストグラムを示した。どの法線方向においても、角度誤差の分布がフィッシャー分布で表現できることが分かる。

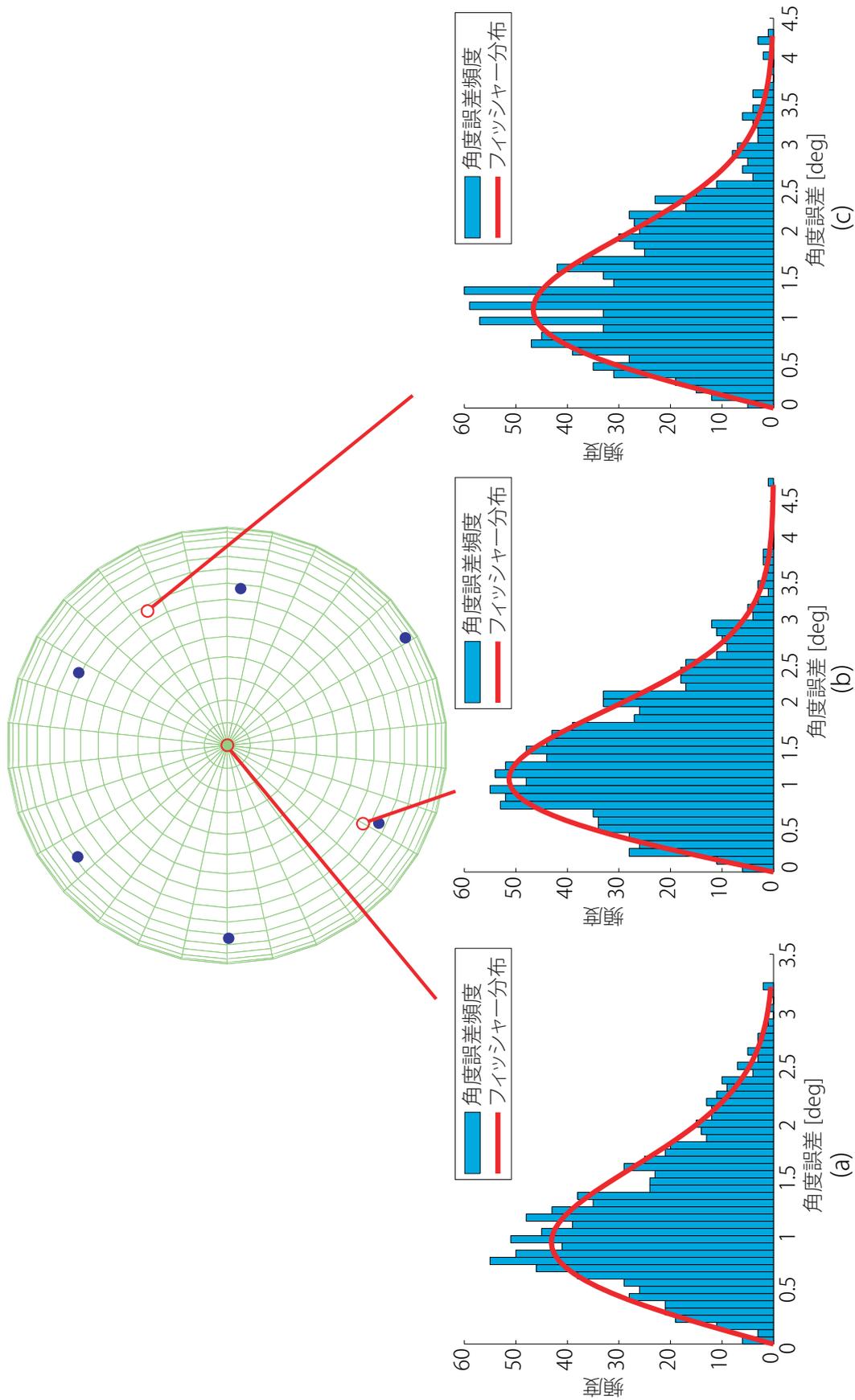


図 6.5: 推定された法線方向の誤差角ヒストグラムとフィッシャー分布の比較．上の極座標内の点は用いられた光源の方向を示す．下のヒストグラムは対応する位置の法線の方向における推定誤差を表す．

## 6.4 フィッシャーノイズモデルを用いた形状復元

前節では照度差ステレオにおける光源方向の校正誤差が法線の推定値に影響を与え、その誤差角がフィッシャー分布に従うことを示した。本節では、法線がフィッシャーノイズで揺らぐときの復元形状の尤度を定式化し、最尤推定によって形状復元を行う。

画素数が  $M$  である 2 次元平面について考える。  $N$  を法線ベクトルを並べた  $M \times 3$  の行列とし、  $Z$  を奥行きを縦に並べたベクトルとする。法線に対するフィッシャー分布ノイズの仮定の下で、法線  $N$  が観測されたときの形状  $Z$  の尤度は

$$\begin{aligned} P(N|Z) &= \prod_{i=1}^M P(\mathbf{N}_i | \bar{\mathbf{N}}_i) \\ &= \prod_{i=1}^M f_{\text{cart}}(\bar{\mathbf{N}}_i; \mathbf{N}_i, \kappa) \end{aligned} \quad (6.4.1)$$

と表せる。ここで  $\bar{\mathbf{N}}$  は形状  $Z$  から計算される法線で、

$$\bar{\mathbf{N}} = \frac{1}{\sqrt{Z_x^2 + Z_y^2 + 1}} \begin{bmatrix} -Z_x \\ -Z_y \\ 1 \end{bmatrix} \quad \left( Z_x = \frac{\partial Z}{\partial x}, Z_y = \frac{\partial Z}{\partial y} \right) \quad (6.4.2)$$

で計算される。(6.4.1) 式を最大化する形状  $Z^*$  は、

$$\begin{aligned} Z^* &= \operatorname{argmax}_Z P(N|Z) \\ &= \operatorname{argmin}_Z (-\log P(N|Z)) \\ &= \operatorname{argmin}_Z \left( -\sum_{i=1}^M \mathbf{N}_i^T \bar{\mathbf{N}}_i \right) \end{aligned} \quad (6.4.3)$$

と求まる。ただしここでは空間的に一様な精度パラメータ  $\kappa$  を仮定し、誤差関数から除外した。(6.4.2) 式より、(6.4.3) 式は  $Z$  に関して非線形であることが分かる。そこで我々は trust region 法 [15] にもとづく MATLAB の非線形最適化関数 *fminunc* を用いて誤差関数を非線形最小化することにより最適な形状  $Z^*$  を求めた。また、非線形最適化問題は多くの局所解を持つ。そこで、最適解から離れた局所解を避けるため、経験的に既存手法を用いて復元された形状を最適化の初期値とし、最適化を実行した。

## 6.5 実験結果

本節では光源方向に校正誤差が含まれる照度差ステレオに対する提案手法の有効性を様々な先行研究と比較することによって示す．照度差ステレオの対象として3種類の人工画像(1)Sphere, (2)Vase, (3)Pennyと2種類の実画像(4)Sphere, (5)Statueに対して形状復元の手法を適用した結果を比較した．

比較対象としては以下の復元形状手法を用いた．最小全域木を用いた線積分 [36], ポアソン方程式の直接解法 [95], Frankot と Chellappa のアルゴリズム [37], Shapelet を用いたアルゴリズム [62], また, 文献 [1] 内で用いられたエネルギー最小化手法として,  $\alpha$ -surface, M-estimator, Huber 関数を用いた Regularization, (Anisotropic) Diffusion の4種類である．Shapelet に関しては, 最小の半径が1でサイズが1.5倍ずつ増加する8つの基底関数を用いた．また, Regularization においては誤差  $s$  に対して  $\phi(s) = \sqrt{1 + s^2}$  なる誤差関数とバランスパラメータ  $\lambda = 10$  を用いた．

### 6.5.1 人工画像を用いた検証

最初に人工画像を用いたシーンを用いて形状復元の性能の概要を示す．6個の光源を配置し, それぞれ標準偏差が3度のフィッシャーノイズを付加し, 照度差ステレオによって法線を推定した．はじめに光源方向の校正誤差による法線ベクトルの推定誤差を示す．図6.6に図6.7で示すシーン Sphere の合成画像の一例と, 各画素を照らす光源の数, および真の法線からの仰角の絶対誤差分布を示す．遮蔽輪郭線の付近では, 照らされている光源の数が少ないため, 法線の角度誤差が大きくなっていることが分かる．また, 5.2節で述べたように, 勾配の大きい遮蔽輪郭線では法線の角度誤差が勾配の小さい領域に比べて大きくなる．したがって, 遮蔽輪郭線付近の勾配の誤差は極めて大きくなり, 勾配にもとづく既存の形状復元手法に大きな影響を与えることになる．

続いて図6.7から図6.9に各シーンの復元結果を示す．左上に正解形状, 左下に正解の光源方向と校正によって推定された光源方向, 右に各手法による復元結果が示

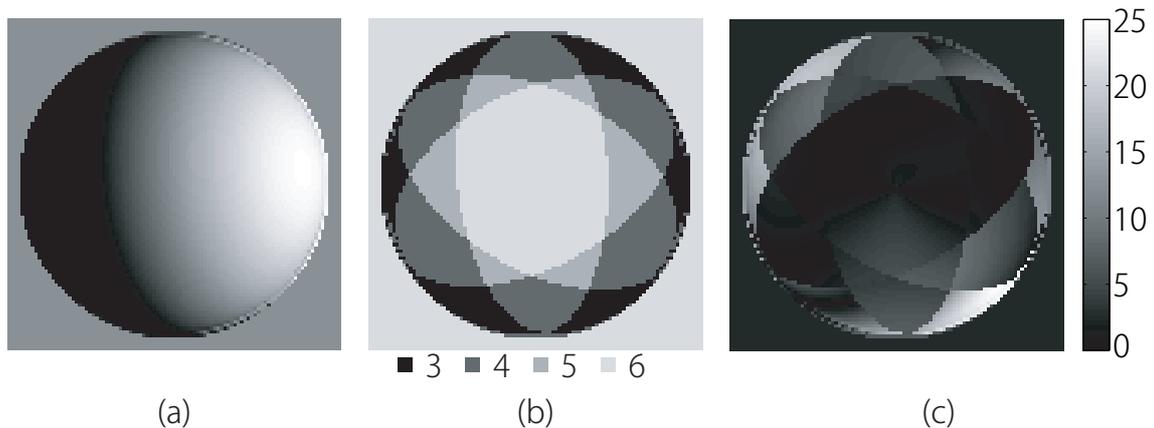


図 6.6: 照らされている光源の数と角度誤差の関係．(a) 合成されたシーン Sphere の例．(b) 画素を照らしている光源の数．(c) 法線ベクトルの仰角の絶対値誤差．

されている．各シーンの合成画像は，計算時間削減のために  $25 \times 25$  画素にダウンサンプリングした．表示のために復元形状が大きく発散している場合は一部を切り取っている．

ポアソン方程式，FC アルゴリズム，Shapelet および Regularization からの復元形状は，全体的に歪んだ結果となっている．線積分， $\alpha$ -surface，M-estimator および Diffusion の結果では全体的な形状の歪みはないが，遮蔽輪郭線付近で尖った形状が復元されている．以上のように，勾配に対するガウスノイズを仮定した既存の手法では，特に遮蔽輪郭線付近において誤差の影響を大きく受けている．一方，提案手法は法線に対する角度誤差を考慮した形状復元を行っているため，全体的な歪みもなくディテールも含めた正確な形状を復元できている．各手法から復元した形状と正解形状との間の RMSE(平均平方 2 乗誤差) の一覧を表 6.1 に示す．RMSE は各正解形状の高さの最大値で正規化した値である．各形状に対して提案手法が最も誤差の小さい復元結果を示した．

図 6.7～図 6.9 および表 6.1 では，光源方向の校正誤差が比較的大きい場合についての例を示した．次に，校正誤差の大きさと光源の数を変えて形状復元の精度を比較し，提案手法が有効な範囲を検証した．各実験条件について 100 回試行し，RMSE の平均を求めた．図 6.10 に 6 つの光源を用いた場合において，校正誤差の大きさ

表 6.1: 人工画像を用いた形状復元 (図 6.7–図 6.9) の性能比較

	線積分	ポアソン方程式	FC	Shapelet	$\alpha$ -Surface
Sphere	1.461	0.671	0.469	0.393	0.188
Vase	2.479	1.736	1.181	1.235	0.098
Penny	20.115	1.627	1.226	0.583	0.223

	M-estimator	Regularization	Diffusion	提案手法
Sphere	0.205	0.604	0.178	<b>0.048</b>
Vase	0.226	1.590	0.287	<b>0.058</b>
Penny	0.526	1.579	0.927	<b>0.115</b>

と復元形状の RMSE の比較を示す。  $\alpha$ -surface, M-estimator, 提案手法は校正誤差に対して頑強に復元が行えていることが分かる。提案手法は計算範囲内で常に最小の RMSE を記録した。また, 図 6.11 には各光源方向の推定結果に 3 度のフィッシャー分布を付加した場合において, 光源の数を変えながら復元形状の RMSE を比較した結果を示す。誤差の大きさを変えた場合と同様に, 提案手法は光源の数が少ない場合でも復元精度で従来手法を上回る結果が得られた。

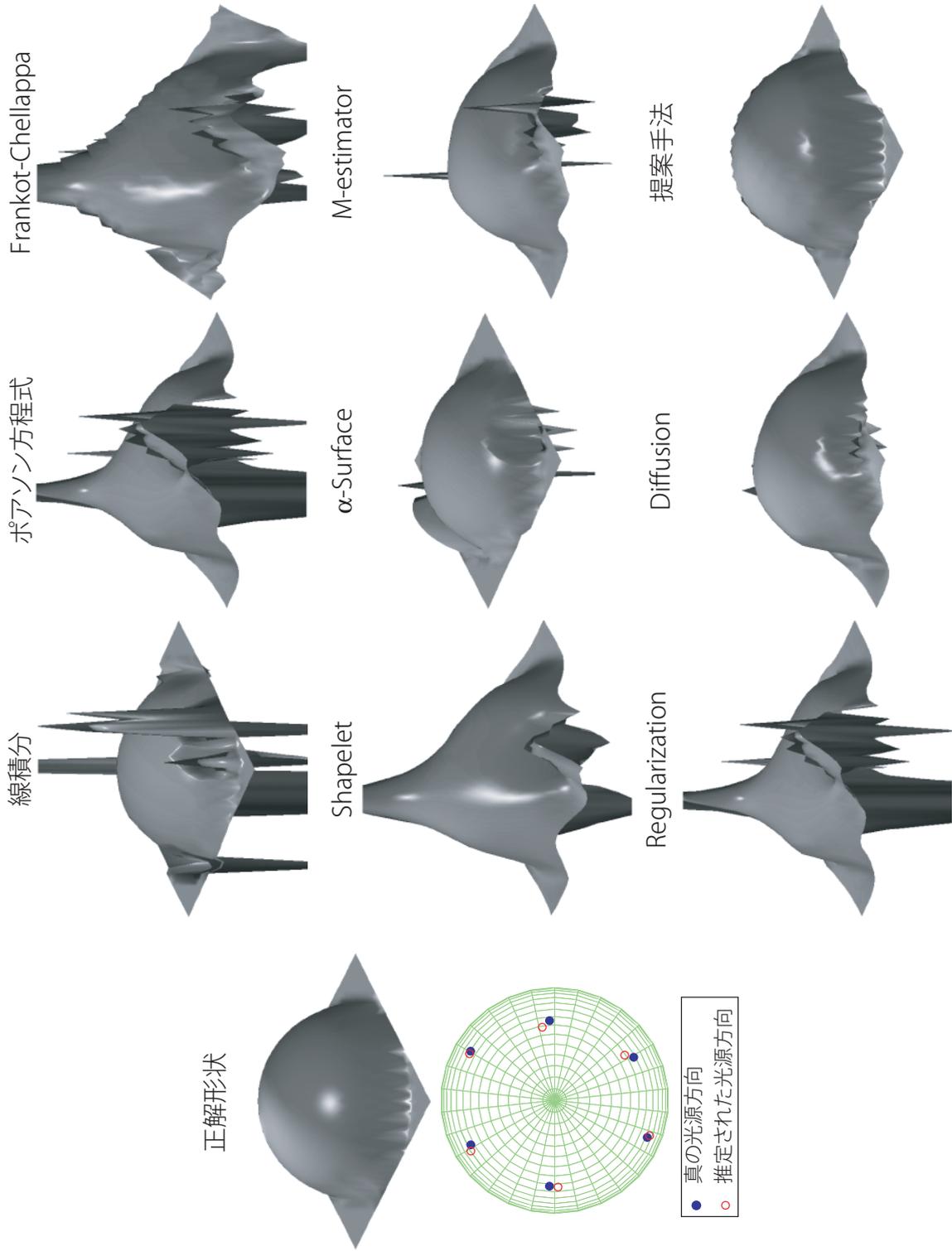


図 6.7: シーン Sphere に対する形状復元結果の比較

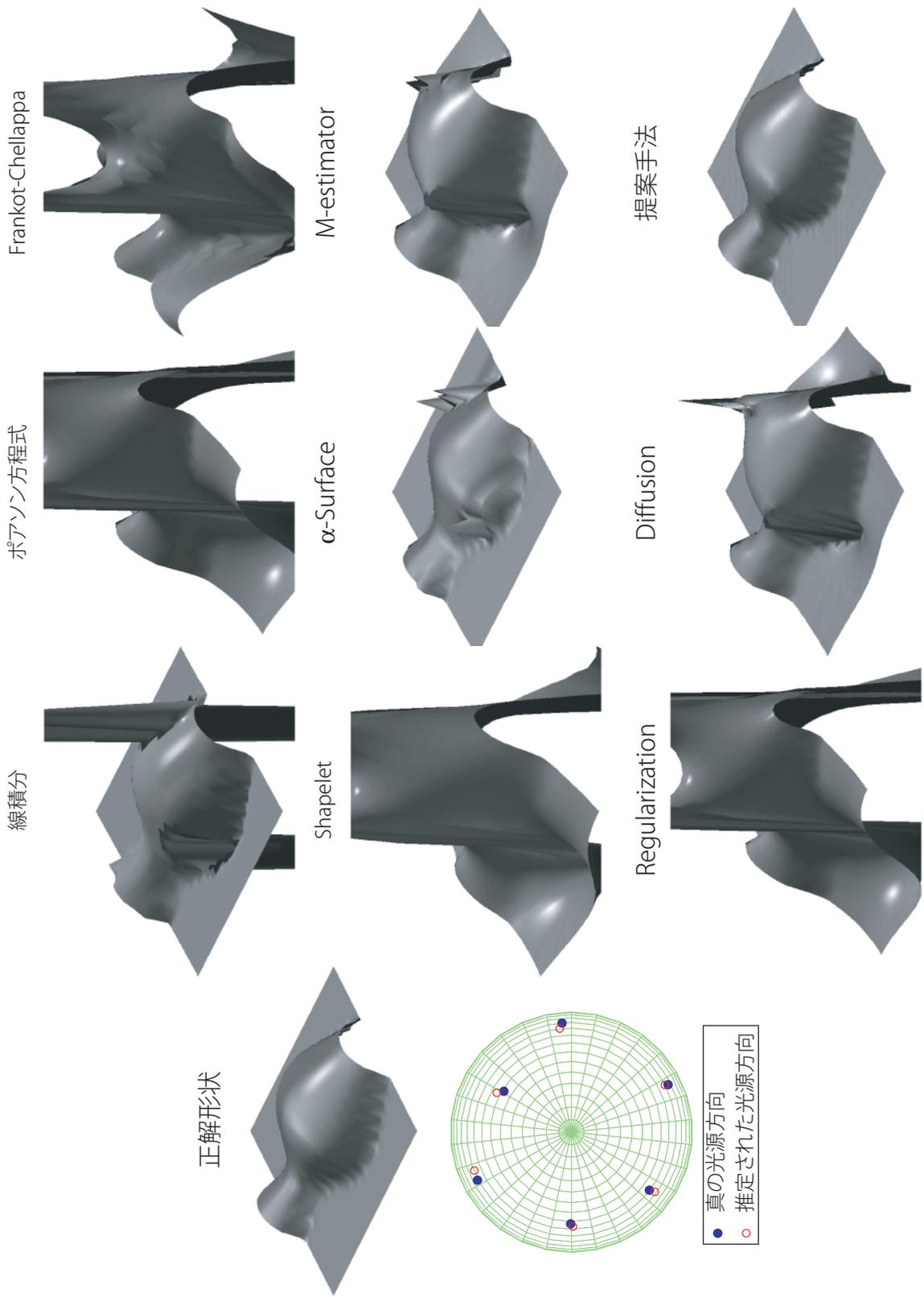


図 6.8: シーン Vase に対する形状復元結果の比較

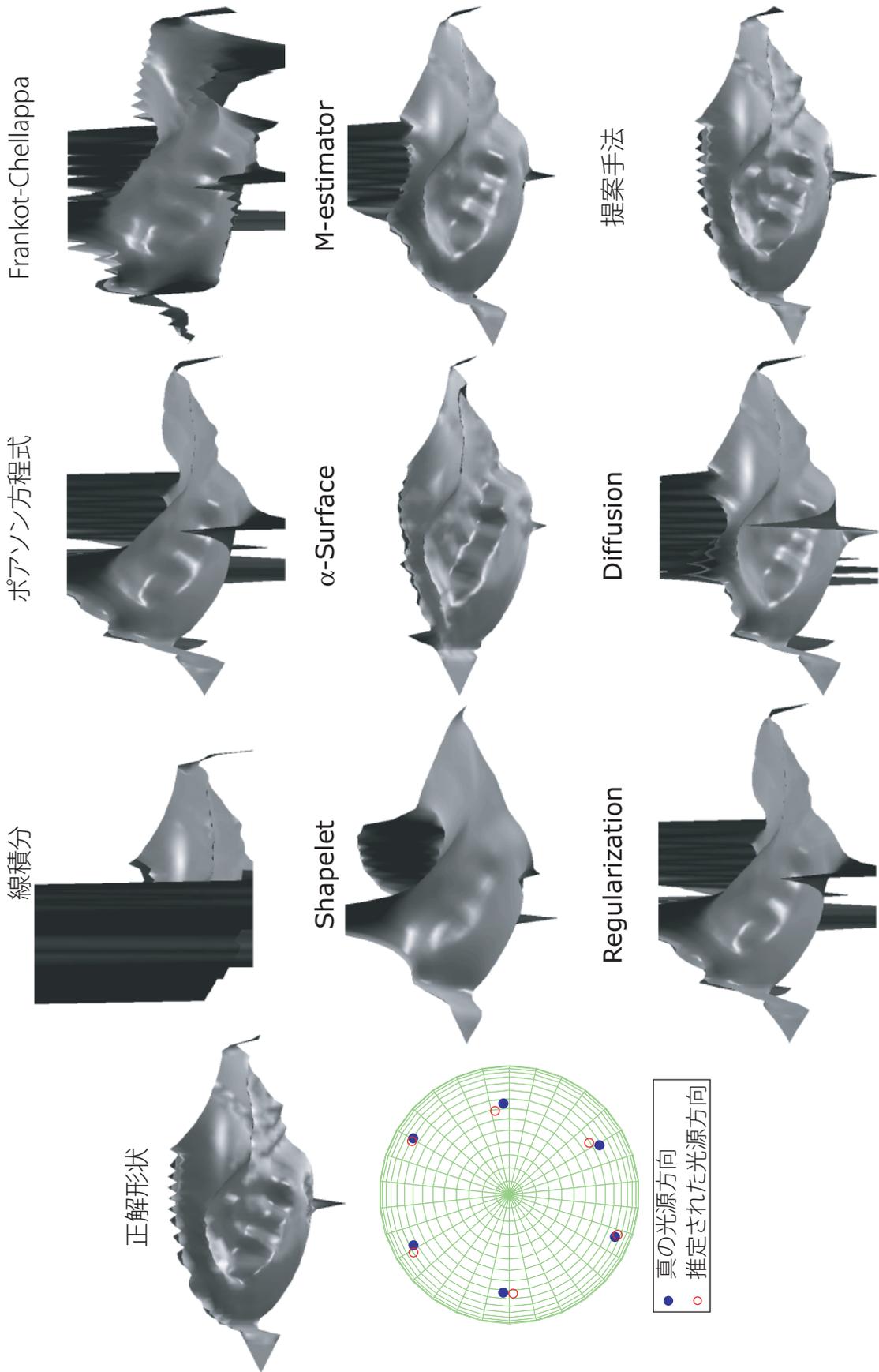


図 6.9: シーン Penny に対する形状復元結果の比較

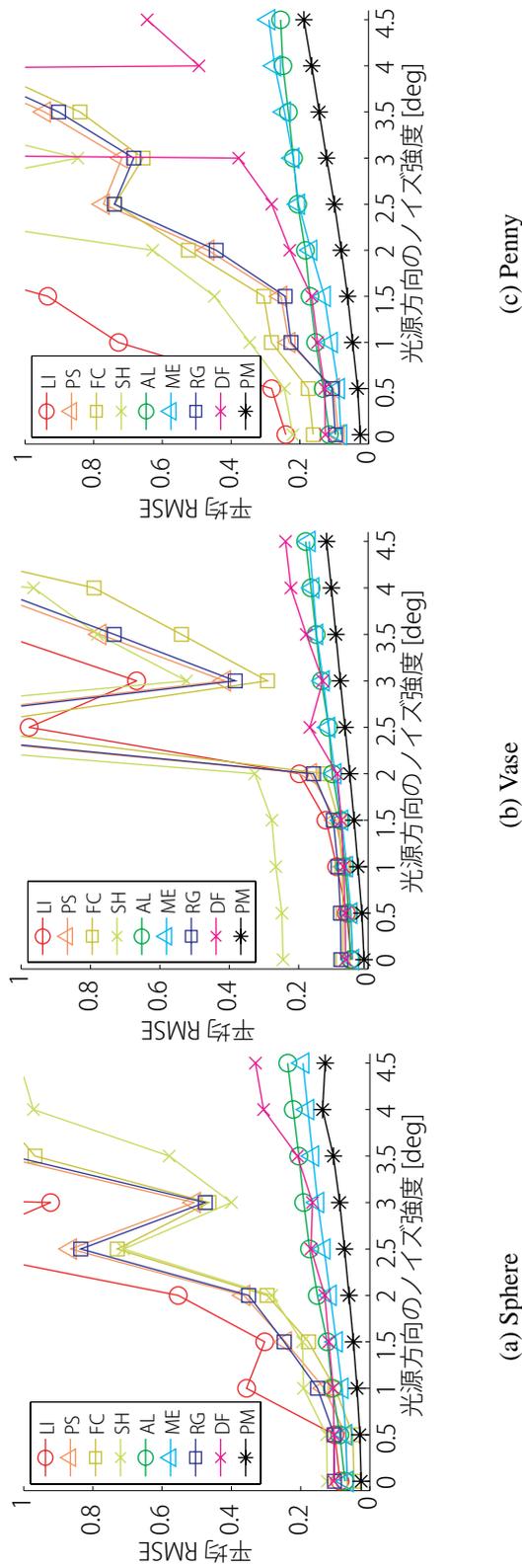


図 6.10: 校正誤差の大きさと形状復元精度の比較．凡例の略号はそれぞれ LI: 線積分, PS: ポアソン方程式の直接解法, FC: Frankot-Chellappa, SH: Shapelet, AL:  $\alpha$ -surface, ME: M-estimator, RG: Regularization, DF: Diffusion, PM: 提案手法を表す．

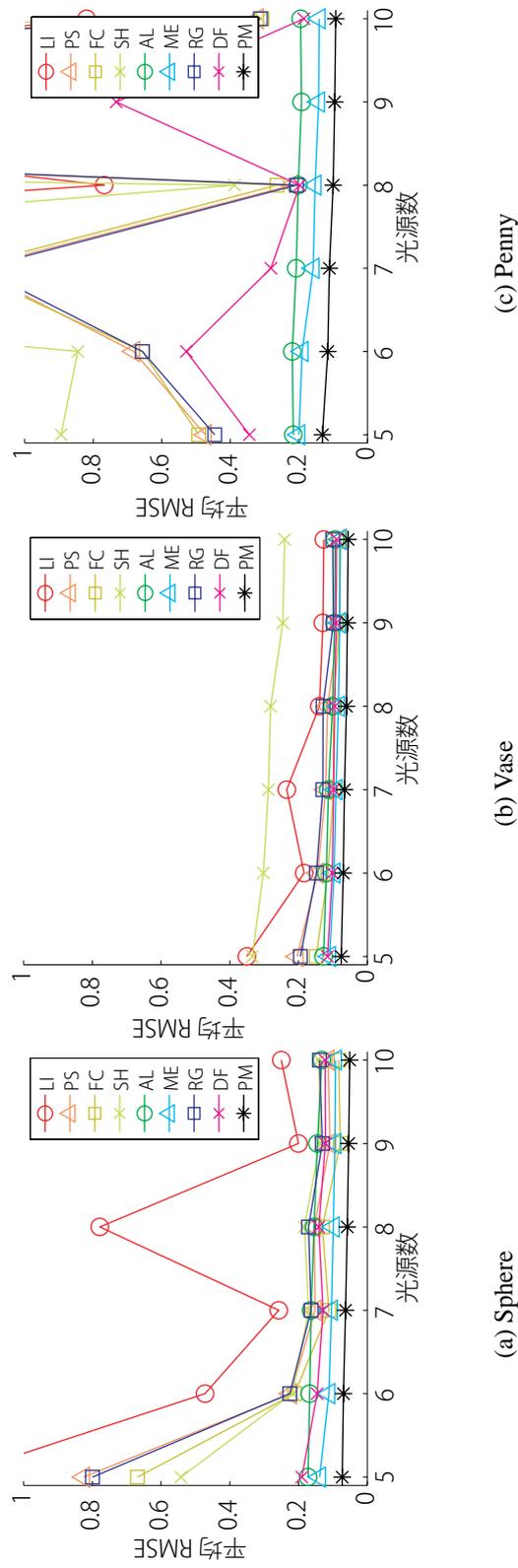


図 6.11: 光源の数と形状復元精度の比較.

表 6.2: 実画像を用いた形状復元 (図 6.12) の性能比較

	線積分	ポアソン方程式	FC	Shapelet	$\alpha$ -Surface
Sphere (real)	0.152	0.101	0.056	0.077	0.154
	M-estimator	Regularization	Diffusion	提案手法	
Sphere (real)	0.101	0.108	0.112	<b>0.055</b>	

### 6.5.2 実画像を用いた検証

続いて、実画像を用いた検証を行った結果を示す。9個の光源を使用し、被写体としてランバートな反射特性と仮定できる石膏製の球と胸像を用いた。各光源の方向は鏡面球を用いて校正した。図 6.12 と図 6.13 に撮影した画像と各手法で復元した形状を示す。

図 6.12 は石膏球に対する形状復元結果である。線積分と  $\alpha$ -surface は遮蔽輪郭線付近で尖った形状を復元した。また、ポアソン方程式、M-estimator、Regularization、Diffusion からは潰れた形状が得られた。FC アルゴリズムと Shapelet からは正確な形状が得られた。石膏球は形状が既知であるので、RMSE によって既存手法と提案手法の定量的な比較を行った。表 6.2 に比較結果を示す。人工画像を用いた評価と同じく、提案手法が最も復元誤差が小さいことが示された。

また、図 6.13 に石膏像を用いた復元結果を示す。石膏像の形状は計測していないので定性的な比較に留まるが、線積分や  $\alpha$ -surface は不連続な面が発生している。FC アルゴリズムと Shapelet からの復元形状は全体的に歪んだ形状となっている。ポアソン方程式と M-estimator は頭部のディテールが失われている。Regularization と Diffusion は潰れた形状を復元している。一方、提案手法は全体的な歪みもなく、ディテールまで復元していることが分かる。

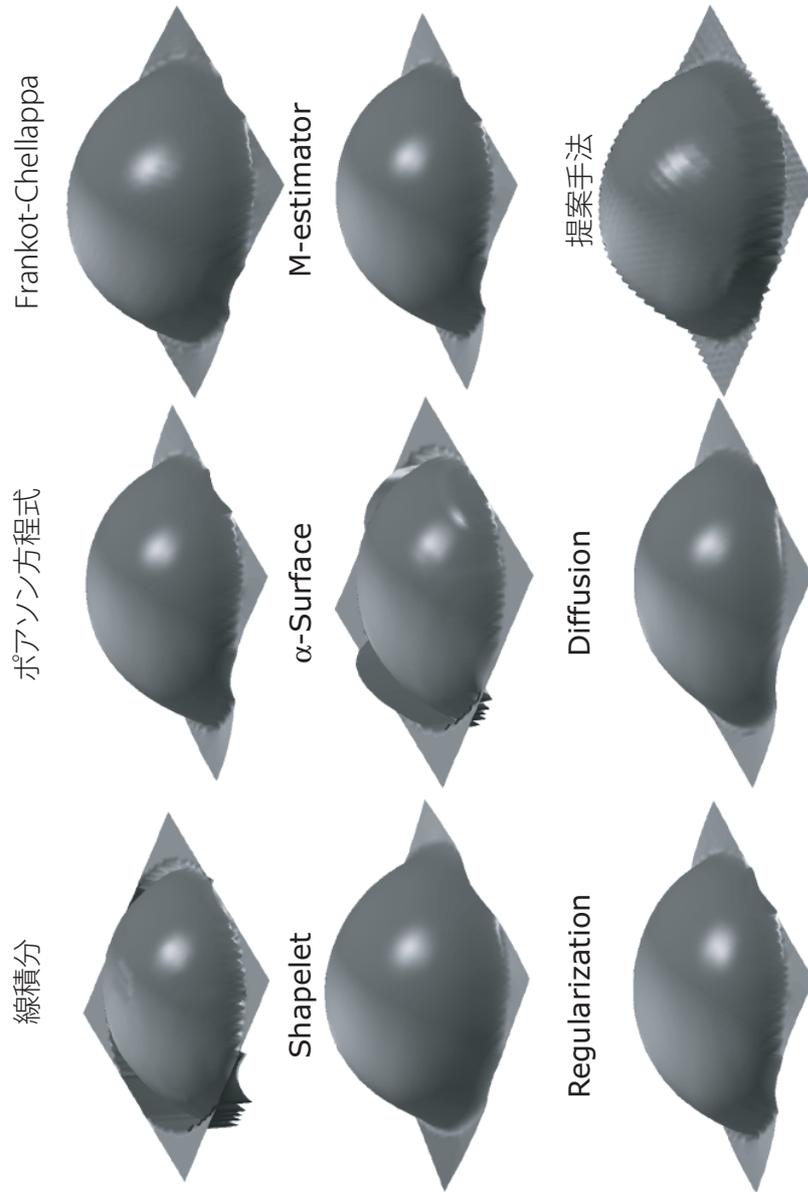
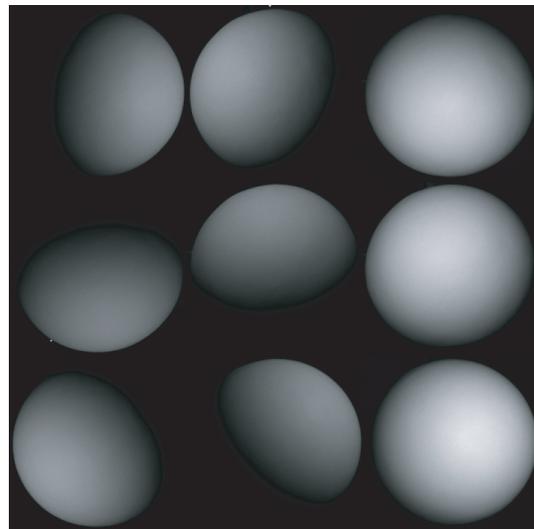


図 6.12: 石膏球に対する形状復元結果の比較



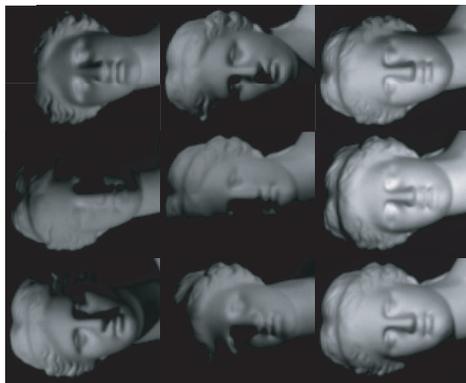
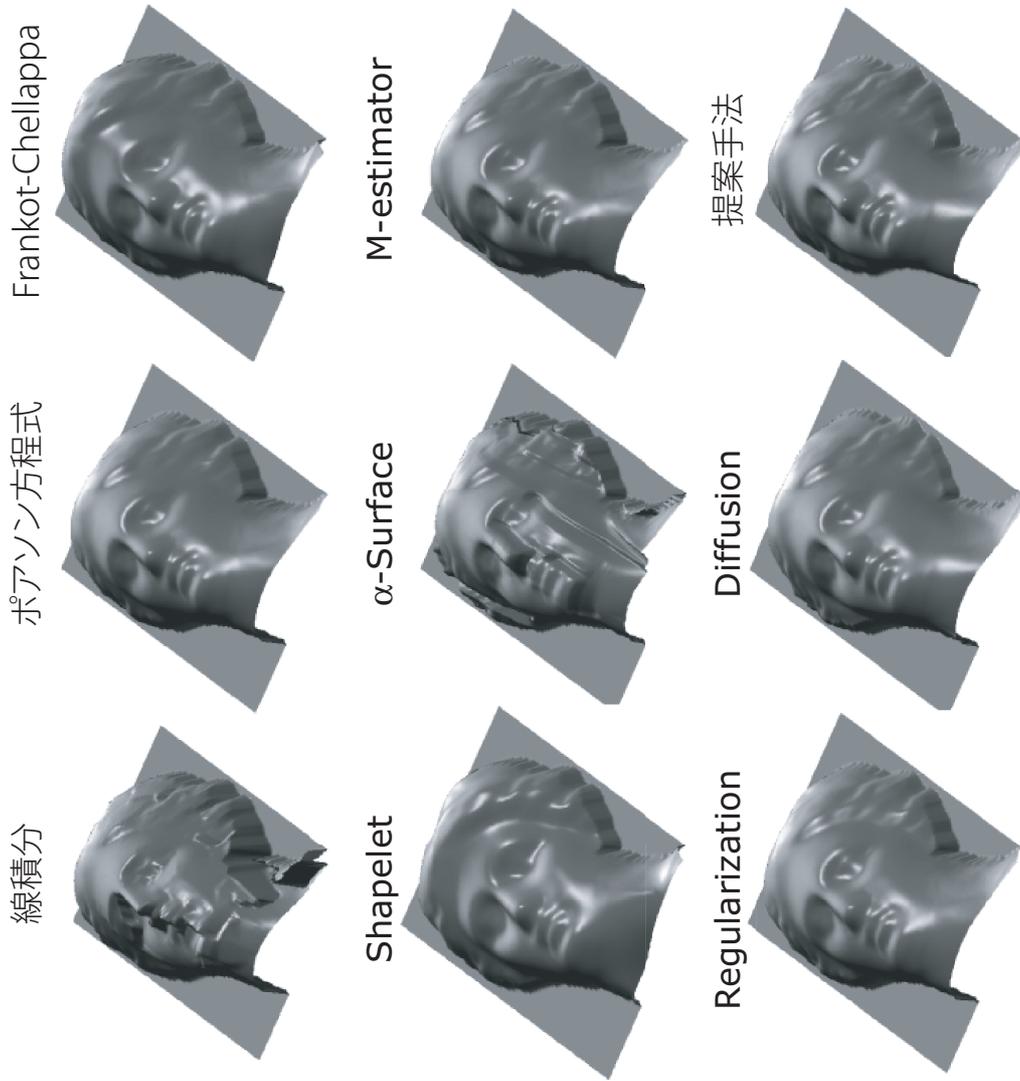


図 6.13: 石膏像に対する形状復元結果の比較

## 6.6 まとめ

本研究では照度差ステレオによって推定された法線に混入する誤差を考慮した形状復元手法を提案した。今回提案したフィッシャーノイズモデルは、光源方向の校正誤差が法線に与える影響のメカニズムにもとづくものである。我々は鏡面球を用いた校正によって生じる光源方向の校正誤差と、その校正誤差によって生じる法線の誤差がフィッシャー分布に従うことを示した。提案手法は大きな校正誤差や少ない光源数における照度差ステレオによって推定された法線からも、形状を正確に復元することができた。

一方、提案手法には次のような手法の限界がある。1つは、計算コストの大きさである。一例として、 $25 \times 25$  画素の人工データに対して、約 100 秒の計算時間を必要とした。また、 $90 \times 110$  画素の実画像に対する最適化には約 20 分要した。計算速度自体は本研究の本質的な問題ではないが、非線形の最適化では画素数に対して指数関数的に計算コストが増加するため、実用性を考えると計算速度を高める必要性が出てくる。この問題の解決法としては、検証実験において比較対象として用いた Regularization で用いられている最適化手法が挙げられる。Regularization の手法が最適化する関数は  $Z$  に関して非線形な関数となっているが、この手法では最適化関数を線形に近似し、反復計算を行うことにより非線形最適化している。

もう1つは非線形の最適化問題を扱うため、収束解が局所解に落ち込む可能性がある点が挙げられる。特に画像の解像度が増えると解空間が広大となるため、広域解から大きく離れた局所解に捕われてしまう可能性が高くなる。非線形最適化問題における局所解の回避は、機械学習における一般的な問題である。この問題を回避するためには、基底関数や多様体を用いた次元削減が考えられるが、形状に対する何らかの事前情報を用いると複雑な形状の復元精度が悪化することが懸念される。あるいは、ポアソン方程式の数値解法の1つであるマルチグリッド法 [32] のように、複数階層構造を用いた coarse-to-fine な最適化を行うことで、広域解に近い領域に範囲を絞り込むことができるかもしれない。

今後の理論的な課題として、照度差ステレオにおける誤差伝播を解析的に示すこ

とが挙げられる．本研究ではまず3次元単位ベクトルの角度誤差に対する確率分布を仮定し，数値シミュレーションを用いてその妥当性を評価した．しかし，解析的な結論はまだ導けていない．誤差伝播の振る舞いを解析的に示すことにより，提案手法をより一般的なノイズモデルに拡張することができると考えられる．例えば，ノイズ強度が空間的に変化する場合に，精度パラメータ  $k$  の空間的な分布をモデルに組み込むことで対応できると考えられる．

## 第7章 結論

### 7.1 本研究のまとめ

本論文では，物理ノイズと計測誤差という異なる起源をもつノイズに対して，それぞれのノイズ特性を解析することにより，与えられた画像・映像の理解に役立てる2種類の手法を提案した．異なる過程から観測されるノイズに対し，それぞれ適切にノイズの振る舞いをモデル化することにより，有効に活用できることを示すことができた．

以下，各種法について簡単にまとめる．

#### 7.1.1 ノイズ特性の一貫性にもとづく映像の改ざん検出について

物理ノイズの特性に着目した適用例では，与えられた映像内に他の映像を貼り付けることによって施された改ざんを，ノイズ特性の一貫性を手がかりとして検出する手法を提案した．静止画像に対する改ざん検出手法はこれまでに様々な手法が提案されてきたが，映像に対する研究はまだ始まったばかりである．映像に対する改ざん検出の既存手法ではノイズに対する脆弱性が問題点として挙げられるが，本研究ではノイズの特性を手がかりとして用いるため，ノイズに対して頑強な検出を実現した．本研究の主要な貢献を個別に述べる．

##### ノイズレベル関数による映像のノイズ特性の評価

本研究では改ざんの手がかりとして，ノイズ特性の一貫性を評価した．映像の改ざん検出に関する先行研究では，映像の記録形式に依存したり，ノイズに対して敏感であるなどの問題点があった．しかし，提案手法はノイズの特性にもとづいて改

ざんを検出するため、記録方式によらず、またノイズに対しても頑強な検出を実現した。

提案手法はノイズ特性をノイズレベル関数と呼ばれる画素値の平均と分散の関係を記述する関数で評価し、与えられた映像からノイズレベル関数を推定することにより一貫していない箇所を改ざんとして検出する。ノイズレベル関数はカメラの種類や撮影時のパラメータに依存するため、ノイズレベル関数が一貫していない領域は異なる条件下で撮影された映像から貼り付けられた箇所であることが推測される。

複数のカメラ応答関数を用いて映像を撮影し、撮影パラメータによってノイズレベル関数が増加することを示した。また、異なるパラメータを用いて撮影された映像から作られた改ざん映像では、異なるノイズ特性が混在して観測されることも示した。

#### ノイズレベル関数の事前情報の利用

ノイズレベル関数の形状は、カメラ応答関数に大きく依存する。一般のカメラ応答関数はノンパラメトリックに定義されており、解析的にノイズレベル関数を推定することができない。そこで、我々はLiuらの先行研究に則り、カメラ応答関数のデータベースから大量のカメラ応答関数を取得し、様々なパラメータを用いて人工的にノイズを合成した。合成したノイズからノイズレベル関数のサンプルを取得し、主成分分析を用いて一般のカメラ応答関数に対するノイズレベル関数の基底関数を取得した。ノイズレベル関数を基底関数を用いた線形結合の形で表現することにより、ノイズレベル関数をパラメトリックに推定することができ、ノイズレベル関数の推定における次元削減を実現した。

#### 確率的な推論を用いた改ざん検出

ノイズレベル関数が推定された後、確率モデルにもとづく推論的な改ざん検出手法を導入した。確率モデルにもとづく推論的な枠組みを用いることにより、経験的に判定境界を定めることなくデータに対して適応的で柔軟に対応出来る検出手法を実現した。提案手法は、有限の計算精度に起因する分散の揺らぎを補償するため、統計的な知見にもとづいて改ざんの確率を評価した。標本から計算された分散はその

生成モデルにある真の分散からずれて観測される．特に，標本数がノイズ強度に対して少ない場合，この揺らぎは深刻になる．分散の揺らぎはカイ二乗分布で表現することができ，我々はノイズレベル関数によって予測される真の分散と観測された分散のずれをカイ二乗分布で評価し，このずれが計算精度と改ざんのどちらに起因するものか確率的に評価した．

また，映像に改ざんが施されている場合は異なるノイズ特性が混在するため，ノイズレベル関数を推定するためには各画素のノイズ特性がどのノイズレベル関数に従うのかが分からなければ正しくノイズレベル関数を推定することができない．一方，各画素の改ざんの確率はノイズレベル関数が分かっていなければ評価することができない．即ち，各画素の改ざんの確率とノイズレベル関数の推定は同時に行わなければならない，所謂“Chicken-and-egg problem”となっている．そこで，我々はEMアルゴリズムを用いてこれらの値を反復的に推定することにより，同時推定を実現した．

#### 移動物体への対応

静止シーンにおけるノイズ特性にもとづく改ざん検出手法を提案した後，移動物体に対応するための手法の拡張を行った．静止シンの場合とは異なり，移動物体の場合は時間的に平均と分散を計算することでノイズ特性を求めることができない．画素単位の対応点追跡は技術的に困難であるので，空間的に広がりのある領域を追跡することによってノイズ特性を求めるアプローチを取った．このとき，空間的な分散にはノイズの成分の他に陰影とテクスチャの成分が含まれるため，事前処理として陰影の成分を除去し，さらに背景から推定したノイズレベル関数を用いてノイズの成分とテクスチャの成分を分離した．

改ざん検出の手がかりとしては，線形のカメラ応答の仮定の下でテクスチャの分散が輝度の2乗の関数となることを利用した．輝度の異なる少数の標本からテクスチャの分散を推定し，輝度とテクスチャの分散の関係が2次曲線に適合する度合いを評価した．テクスチャの分散が2次曲線で記述できないときには，背理法により背景と異なるノイズレベル関数にもとづくノイズが混入したと結論づけることがで

きるため、異なるカメラ応答で撮影された移動物体を判定することができた。

#### 改ざん検出手法のまとめ

最後に、本研究の成果を以下にまとめる。

- ノイズレベル関数を用いて与えられた映像のノイズ特性を記述し、撮影条件に対するノイズレベル関数の依存性を確認した。
- ノイズレベル関数の一貫性を検証することにより、画素単位で改ざんの有無を推定することができた。
- 有限精度の計算による分散の揺らぎを考慮し、統計的な知見にもとづいて改ざんの尤度を定義することにより、適応的な推定を実現した。
- 改ざんの推定を混合分布モデルにもとづいてモデル化し、EM アルゴリズムを用いて反復的な推定を行うことにより、各画素の改ざんの確率とノイズレベル関数の同時推定を実現した。
- 移動物体に対応するため、領域的な追跡のアプローチを提案し、テクスチャ分散と輝度との関係を手がかりとして改ざんを判定する手法を提案した。

#### 7.1.2 光源方向の校正誤差を考慮した形状復元について

計測誤差のノイズ特性に着目した適用例としては、光源方向の校正誤差を考慮した照度差ステレオからの形状復元手法を提案した。形状復元の先行研究では、勾配に対して正規分布するノイズが混入すると仮定し、最尤推定の枠組みで形状を積分している。しかし、勾配に対する正規ノイズの仮定は経験的であり、実際に照度差ステレオにおいて混入するノイズの特性にもとづくものではない。本研究では、照度差ステレオから推定された法線に対して混入するノイズの特性にもとづいて誤差モデルを構築し、形状の復元精度を向上させることを目指した。本研究の主要な貢献を個別に述べる。

### 単位ベクトルに対する誤差モデルの提案

照度差ステレオでは校正によって求めた光源方向と撮影された画像から物体の法線を推定する．光源方向と法線は共に3次元ベクトルであるが，その大きさが1に正規化された単位ベクトルで表現される．単位ベクトルの揺らぎは3次元正規分布では表現できないため，我々は単位球面上の正規分布であるフィッシャー分布を用いて単位ベクトルの揺らぎを評価する手法を提案した．

### 光源方向と法線の推定誤差の検証

光源方向の校正には一般に鏡面球を用いる．鏡面球は画像上の位置から法線方向が計算できるため，鏡面球上に当たる光源のピーク位置を測定することで光源方向を推定することができる．我々は鏡面球を用いた校正を模して数値シミュレーションを行い，画像ノイズの校正に対する影響を検証した．シミュレーションによって計算された正解方向からの角度誤差のヒストグラムに対して，フィッシャー分布は精度良くフィッティングできることが示された．

同様に，照度差ステレオを模した数値シミュレーションによって，光源方向の校正値にフィッシャー分布するノイズが混入した場合における，法線ベクトルの推定誤差を評価した．シミュレーションの結果，法線ベクトルの推定誤差もフィッシャー分布によって適切に表現できることが分かった．

### フィッシャーノイズモデルによる形状復元

数値シミュレーションを用いた検証から，法線ベクトルに対するフィッシャーノイズモデルの有効性が示された．そこで，フィッシャーノイズモデルにもとづき，最尤推定の枠組みで形状を復元する手法を構築・実装した．既知の形状を用いた人工画像による定量的評価に加え，実画像を用いた定量的・定性的評価によって提案手法を評価した．そして，提案手法は従来手法に比べ校正誤差の影響に対して頑強に形状を復元することができ，かつ形状のディテールを精度良く復元できることが示された．

## 形状復元手法のまとめ

最後に、本研究の成果を以下にまとめる。

- 校正誤差を考慮したアルゴリズムの枠組みを提案した。具体的には照度差ステレオにおける光源方向の校正誤差をモデル化し、誤差モデルにもとづく形状復元手法を提案した。
- 光源方向の校正誤差と法線ベクトルの推定誤差の分布を数値シミュレーションによって検証し、フィッシャー分布によって表現できることを示した。
- フィッシャーノイズモデルにもとづく形状復元手法を構築し、人工画像と実画像を用いて提案手法が形状復元の精度を向上させることを示した。

## 7.2 今後の課題

本節では本論文で紹介した研究において今後取り組む課題について触れる。

### 7.2.1 改ざん検出について

#### 実用性向上に向けた手法の拡張と検証

本研究では静止シーンに対する有効性の評価を示すことができたが、移動物体への対応についてはまだ人工的な映像と実験室環境で撮影された実映像を用いた評価評価に留まっており、より現実的な映像を用いた検証が必要である。屋内外で撮影された映像では、背景に微小な動きが入るなど、提案手法の仮定が成り立たない可能性がある。また、陰影の影響を除去しきれないといった原因によりノイズ特性の計算に誤差が生じ、正確な改ざん検出が行えないといった問題が想定される。さらに、本研究では異なる時刻における対応領域を手で与えていたが、実用性を考えると領域的な追跡によって対応領域を自動的に探索することが望ましい。領域的な追跡はそれ自体が1つの研究対象として成り立つ複雑な問題であり、追跡技術の知見を導入することによって正確な領域追跡を行う必要がある。

また、本研究では監視カメラを想定した固定カメラを仮定したが、適用可能な映像の範囲を広げるためには、背景も含んだ全領域が時間的に変化するシーンに対しても手法を拡張する必要がある。背景が固定されたシーンでは時間方向に平均と分散を計算することによりノイズ特性を求めることができたが、動的なシーンでは全ての領域について時間的な追跡を行う必要がある。本文でも述べたように、画素単位の追跡は技術的に困難であり、領域的な追跡はノイズとテクスチャの分離問題が生じる。したがって、提案手法とは異なるアプローチでノイズ特性を推定する手法が必要となる可能性がある。

例えば、本研究ではグレースケールの映像を用いたが、フルカラーの映像を用いれば情報が増えるため、映像に固有の色情報の手がかりが得られる可能性がある。ノイズレベル関数は3チャンネルにもそのまま拡張できるので、各色成分の分散は色空間において長楕円体の軸の長さに相当する。各軸の長さは平均の色と関係があるため、グレースケールの場合と同様に映像内で一貫する特性が定義できれば改ざん検出の手がかりとして用いることができると期待される。

#### ノイズレベル関数の他の応用

ノイズレベル関数はカメラや撮影条件に固有な特性を示すため、改ざん検出以外にも応用することができると考えられる。

例えば、画像に付加される情報としてEXIFがあるが、EXIF情報はネットワークに公開される時点で破棄される場合がある。EXIF情報の中にはカメラモデルや電気的なゲイン (ISO 値) など、撮影条件の情報も含まれており、画質改善のために有効に用いることができる可能性がある。したがって、EXIF情報が破棄された画像からノイズ特性を手がかりとしてカメラモデルや撮影パラメータを推定できることが期待される。カメラモデルの特定は画像の真正性評価にも通じるため、改ざん検出と同様に画像の信頼性向上への応用にも用いることができる。

また、カメラのモデルが与えられている場合には、撮影条件を変えた場合にノイズの強度がどのように変化するか推定することができるため、撮影条件に適応したノイズ除去への応用が期待される。これまでの適応的なノイズ除去の手法は、与え

られた画像からノイズ除去フィルターのパラメータを適応的に変化させることによって高品質なノイズ除去を削減してきたが、カメラノイズや撮影条件にもとづく事前知識を導入することにより、ノイズ除去の性能を高めることができると考えられる。

## 7.2.2 形状復元について

### 誤差分布の解析的な導出

本研究では照度差ステレオにおける光源方向の校正誤差と法線ベクトルの推定誤差がフィッシャー分布に従うと仮定し、数値シミュレーションを用いて検証を行った。しかし、解析的な導出はまだ行っていない。提案手法の妥当性を高めるためにも、誤差分布の解析的な導出は重要であると言える。

しかし、法線ベクトルの推定誤差は複数の光源の校正誤差が複雑に関連して観測されるため、誤差伝播の解析は非常に困難である。複雑な事象の解析的な確率モデルは、変分推論として研究されている。変分推論では確率分布の厳密な形状を導出することが困難な場合に、関数の形状を制限したり基底関数で表現することによって、確率分布の近似的な解を求める。

解析的な確率分布を導出することにより、より一般的なノイズモデルに対する拡張が容易に行えることができると考えられる。例えば、各光源の空間的な配置によって、ノイズ強度が空間的に変化する様子を解析的に推定することができると期待される。

### 確率モデルの一般化

提案手法では最も簡単な法線ベクトルの確率モデルとして、精度パラメータ  $\kappa$  が空間的に一様な場合を仮定した。しかし、実際にはノイズ強度は空間的に一様ではなく、区間的な強度分布を示すはずである。例えば光源ごとにノイズ強度が異なったり、光源の配置によって法線の推定精度は変化すると考えられる。したがって、精度パラメータの空間分布の関係を考慮に入れることにより、より現実環境を反映した高精度な形状復元ができると考えられる。

同様に、フィッシャー分布の一般化も検討している。フィッシャー分布は誤差が真の方向を軸として対称に分布する等方的な分布関数であるが、非等方的な分布に一般化することが考えられる。フィッシャー分布を一般化した分布はケント分布 [58] と呼ばれ、2つの独立したパラメータをもつ単位球面上の確率分布を表す。

勾配に対する正規ノイズ仮定が一樣・等方的な仮定によりポアソン方程式を導き、後に非一樣・非等方的な場合に拡張されたことにより復元精度が向上した。同様に、法線ベクトルに対するフィッシャーノイズの仮定も非一樣・非等方的な場合に拡張することによって復元精度が向上すると考えられる。

また、本研究では光源の方向のみに校正誤差が生じる場合の誤差モデルを考慮したが、同様の枠組みで光源強度の校正誤差や、画像ノイズが校正に与える影響についても誤差モデルを導入することによって精度を向上できると考えられる。

### 7.3 研究全体の包括

最後に本論文の総括として、本研究が提案したノイズ特性の有効活用の貢献について述べる。

本研究では、画像・映像からノイズの特性を推定し、ノイズを除去するのではなく画像・映像の理解に応用する枠組みを提案した。生成過程の異なるノイズに対して我々はノイズモデルを導入し、どちらのノイズ源も統計的な特性を元に有効活用できることを示した。

ノイズの特性を考慮することによる恩恵は大きく2つ考えられる。1つ目はノイズが確率的な現象であるため、画像や映像に対する固有の信号として取り扱える点である。改ざん検出はその検出アルゴリズムが分かっているならば、検出されないような改ざんを施すことは理論的には可能である。しかし、ノイズ特性にもとづく改ざん検出では、対称映像からノイズを完全に排除した上で特性の一貫性を保ちながらノイズを付加する必要がある。技術的には非常に困難である。したがって、ノイズを画像や映像に固有の指紋のように取り扱うことで、真正性評価に役立てることが期待される。

2つ目の恩恵は、既存の推定アルゴリズムの精度限界の向上である。序論で述べたように、ノイズの有効利用は特に生物の信号処理において頻繁に行われていると考えられており、ノイズの非常に大きな環境においても頑強に信号処理を行うことができる特長がある。一方、本研究では画像処理においてもノイズの特性を正しく把握することにより、ノイズの大きい環境下で撮影された画像や映像からもアルゴリズムの精度を向上させることができることを示した。今後も技術の進歩によりノイズの少ない良質な信号が取得できると考えられるが、現実的にはコストの問題でノイズ削減には限界がある。よって、ある程度のノイズ環境を前提とし、アルゴリズムを最適化することによって精度向上を図る提案手法の枠組みは、工学的にも意義のあるものと言える。

本研究を通じて、我々はノイズの扱いに対する新たな枠組を提案してきた。本論文ではノイズ活用の適用例の一例を提案するのみであったが、この分野がより充実し、体系化されることによって、画像処理の広範囲に渡るノイズ活用の枠組みが構築されることが期待される。

## 謝辞

本研究を遂行するにあたり、常日頃より熱心なご指導を頂きました佐藤洋一准教授に心から御礼申し上げます。博士課程から異なる分野に携わり、右も左も分からない状態からこの度晴れて修了することができたのも、佐藤准教授の細やかなご助言の賜であると感謝しております。

池内克史教授および相澤清晴教授には、アドバイザー教員として貴重なご意見を頂戴し、研究の意義を改めて考える機会を頂きましたこと、ここに感謝申し上げます。

マイクロソフトリサーチアジアにインターンした際には、松下康之主任研究員に様々なサポートを頂き、帰国後も再度にわたりご助言を頂きましたこと感謝しております。また、岡部孝弘助教には日頃よりの的確かつ鋭いご指摘を頂き、研究の完成度を高めるためにご尽力下さいましたこと御礼申し上げます。国立情報学研究所の佐藤いまり准教授には研究以外でも何かとお気遣い頂き、快適な研究室生活を送ることができましたこと感謝申し上げます。

研究室では、島野美保子研究員には研究の助言の他にも苦しい時期に労って頂き、お気遣い頂きました。菅野裕介氏と杉村大輔氏には公私共に熱く議論を交わして頂き、同級生として充実した生活を共に過ごすことができました。佐藤研究室の卒業生である木谷クリス真実氏、熊野史郎氏、劉玉宇氏および松井壮介氏には、様々な有益なご助言を頂きました。在籍中にお世話になった佐藤研究室のメンバー皆様に心より感謝いたします。

最後に、いつも暖かく見守り支えて下さった母と兄、そして研究者となる夢を与えて下さった亡父に心より感謝いたします。

平成 22 年 6 月 15 日

小林 理弘

## 参考文献

- [1] A. Agrawal, R. Raskar, and R. Chellappa. What is the range of surface reconstructions from a gradient field? *Lecture Notes in Computer Science*, Vol. 3951, pp. 578–591, 2006.
- [2] M. Andriluka, S. Roth, and B. Schiele. Monocular 3D pose estimation and tracking by detection. In *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2010.
- [3] J. Angulo and J. Serra. Modelling and segmentation of colour images in polar representations. *Image and Vision Computing*, Vol. 25, No. 4, pp. 475–495, 2007.
- [4] H. Bao and Z. Zhang. Motion objects segmentation using a new level set based method. *Advances in Multimedia Information Processing-PCM 2004*, pp. 312–318, 2005.
- [5] R. Basri, D. Jacobs, and I. Kemelmacher. Photometric stereo with general, unknown lighting. *International Journal of Computer Vision*, Vol. 72, No. 3, pp. 239–257, 2007.
- [6] R. Basri and D.W. Jacobs. Lambertian reflectance and linear subspaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 218–233, 2003.
- [7] S. Bayram, H. Sencar, N. Memon, and I. Avcibas. Source camera identification based on CFA interpolation. In *Proc. of IEEE International Conference on Image Processing*, Vol. 3, pp. III–69–72, 2005.

- [8] P.N. Belhumeur, D.J. Kriegman, and A.L. Yuille. The bas-relief ambiguity. *International Journal of Computer Vision*, Vol. 35, No. 1, pp. 33–44, 1999.
- [9] R. Benzi, G. Parisi, A. Sutera, and A. Vulpiani. Stochastic resonance in climatic change. *Tellus*, Vol. 34, pp. 10–16, 1982.
- [10] C.M. Bishop. *Pattern Recognition and Machine Learning*. Springer New York, 2006.
- [11] G.J. Bloy. Blind camera fingerprinting and image clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 30, No. 3, pp. 532–534, 2008.
- [12] R. Bodor, B. Jackson, and N. Papanikolopoulos. Vision-based human tracking and activity recognition. In *Proc. of Mediterranean Conference on Control and Automation*, Vol. 1, 2003.
- [13] Robert A. Boie and Ingemar J. Cox. An analysis of camera noise. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 14, No. 6, pp. 671–674, 1992.
- [14] A. Cavallaro, O. Steiger, and T. Ebrahimi. Tracking video objects in cluttered background. *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 15, No. 4, pp. 575–584, 2005.
- [15] MR Celis, JE Dennis, and RA Tapia. A trust region strategy for nonlinear equality constrained optimization. In *Proc. of SIAM Conference on Numerical Optimization*, p. 71. Siam, 1985.
- [16] M.K. Chandraker, F. Kahl, and D.J. Kriegman. Reflections on the Generalized Bas-Relief Ambiguity. In *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, p. 795. IEEE Computer Society, 2005.

- [17] P. Charbonnier, L. Blanc-Feraud, G. Aubert, and M. Barlaud. Deterministic edge-preserving regularization in computed imaging. *IEEE Transactions on Image Processing*, Vol. 6, No. 2, pp. 298–311, 1997.
- [18] E. Coleman and R. Jain. Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry. *Computer Graphics and Image Processing*, Vol. 18, No. 4, pp. 309–328, 1982.
- [19] J.J. Collins, C.C. Chow, and T.T. Imhoff. Stochastic resonance without tuning. *Nature*, Vol. 376, No. 6537, pp. 236–238, 1995.
- [20] D. Comaniciu, V. Ramesh, and P. Meer. Kernel-based object tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 564–575, 2003.
- [21] I.J. Cox, M.L. Miller, and J.A. Bloom. *Digital watermarking*. Morgan Kaufmann Publishers, 2002.
- [22] S.A. Craver, M. Wu, B. Liu, A. Stubblefield, B. Swartzlander, D.S. Wallach, D. Dean, and E.W. Felten. Reading between the lines: Lessons from the SDMI challenge. In *Proc. of USENIX Security Symposium*, 2001.
- [23] A.P. Dempster, N.M. Laird, D.B. Rubin, et al. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 39, No. 1, pp. 1–38, 1977.
- [24] O. Drbohlav and M. Chantler. On optimal light configurations in photometric stereo. In *Proc. of IEEE International Conference on Computer Vision*, Vol. 2, pp. 1707–1712 Vol. 2, 2005.
- [25] R.J. Elliott, L. Aggoun, and J.B. Moore. *Hidden Markov models: estimation and control*. Springer, 1995.

- [26] J. Fan and L.B. Wolff. Surface curvature and shape reconstruction from unknown multiple illumination and integrability. *Computer Vision and Image Understanding*, Vol. 65, No. 2, pp. 347–359, 1997.
- [27] H. Farid. Detecting digital forgeries using bispectral analysis. Technical report, MIT, 1999.
- [28] H. Farid. Digital image ballistics from JPEG quantization: A followup study. Technical Report TR2008-638, Department of Computer Science, Dartmouth College, 2008.
- [29] H. Farid. *Deception: Methods, Motives, Contexts and Consequences*, chapter Digital Doctoring: can we trust photographs?, pp. 95–108. Stanford University Press, 2009.
- [30] Raanan Fattal, Dani Lischinski, and Michael Werman. Gradient domain high dynamic range compression. In *Proc. of International Conference and Exhibition on Computer Graphics and Interactive Techniques*, pp. 249–256, 2002.
- [31] S. Fauve and F. Heslot. Stochastic resonance in a bistable system. *Physics Letters A*, Vol. 97, No. 1-2, pp. 5–7, 1983.
- [32] RP Fedorenko. A relaxation method for solving elliptic difference equations. *USSR Computational Mathematics and Mathematical Physics*, Vol. 1, No. 4, pp. 1092–1096, 1962.
- [33] G.D. Finlayson, S.D. Hordley, and M.S. Drew. Removing shadows from images. *Lecture Notes in Computer Science*, pp. 823–836, 2002.
- [34] M.A. Fischler and R.C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, Vol. 24, No. 6, pp. 381–395, 1981.

- [35] R. Fisher. Dispersion on a sphere. *Proc. of the Royal Society of London. Series A*, pp. 295–305, 1953.
- [36] R. Fraile and E. R. Hancock. Combinatorial surface integration. In *Proc. of IEEE International Conference on Pattern Recognition*, Vol. 1, pp. 59–62, 2006.
- [37] R.T. Frankot and R. Chellappa. A method for enforcing integrability in shape from shading algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 10, No. 4, pp. 439–451, 1988.
- [38] Jessica Fridrich, D. Soukal, and J. Lukáš. Detection of copy-move forgery in digital images. In *Proc. of Digital Forensic Research Workshop*, 2003.
- [39] A.C. Gallagher and Tsuhan Chen. Image authentication by detecting traces of demosaicing. In *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pp. 1–8, 2008.
- [40] A.S. Georghiades, P.N. Belhumeur, and D.J. Kriegman. From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 6, pp. 643–660, 2001.
- [41] A. Guderian, G. Dechert, K.P. Zeyer, and FW Schneider. Stochastic resonance in chemistry. 1. the Belousov- Zhabotinsky reaction. *Journal of Physical Chemistry*, Vol. 100, No. 11, pp. 4437–4441, 1996.
- [42] J.W Harris and H. Stocker. *Handbook of Mathematics and Computational Science*. Springer, 2006.
- [43] H. Hayakawa. Photometric stereo under a light source with arbitrary motion. *Journal of the Optical Society of America A*, Vol. 11, No. 11, pp. 3079–3089, 1994.

- [44] G.E. Healey and R. Kondepudy. Radiometric CCD camera calibration and noise estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 16, No. 3, pp. 267–276, 1994.
- [45] J. Heikkila and O. Silven. A four-step camera calibration procedure with implicit image correction. In *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1106–1112. Citeseer, 1997.
- [46] B.K.P. Horn and B.G. Schunck. Determining optical flow. *Artificial intelligence*, Vol. 17, No. 1-3, pp. 185–203, 1981.
- [47] Chih-Chung Hsu, Tzu-Yi Hung, Chia-Wen Lin, and Chiou-Ting Hsu. Video forgery detection using correlation of noise residue. In *Proc. of IEEE Workshop on Multimedia Signal Processing*, pp. 170–174, 2008.
- [48] Y.-F. Hsu and S.-F. Chang. Detecting image splicing using geometry invariants and camera characteristics consistency. In *Proc. of IEEE International Conference on Multimedia and Expo*, pp. 549–552, 2006.
- [49] Yizhen Huang and Yangjing Long. Demosaicking recognition with applications in digital photo authentication based on a quadratic pixel correlation model. In *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1–8, 2008.
- [50] K. Irie, AE McKinnon, K. Unsworth, and IM Woodhead. A model for measurement of noise in CCD digital-video cameras. *Measurement Science and Technology*, Vol. 19, p. 045207, 2008.
- [51] M.K. Johnson and H. Farid. Exposing digital forgeries by detecting inconsistencies in lighting. In *Proc. of Workshop on Multimedia and Security*, pp. 1–10, 2005.
- [52] M.K. Johnson and H. Farid. Exposing digital forgeries through chromatic aberration. In *Proc. of International Multimedia Conference*, pp. 48–55, 2006.

- [53] M.K. Johnson and H. Farid. Exposing digital forgeries in complex lighting environments. *IEEE Transactions on Information Forensics and Security*, Vol. 2, No. 3, pp. 450–461, 2007.
- [54] Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME—Journal of Basic Engineering*, Vol. 82, No. Series D, pp. 35–45, 1960.
- [55] S. Kamijo, Y. Matsushita, K. Ikeuchi, and M. Sakauchi. Traffic monitoring and accident detection at intersections. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 1, No. 2, pp. 108–118, 2000.
- [56] B. Karacali and W.E. Snyder. Partial integrability in surface reconstruction from a given gradient field. In *Proc. of International Conference on Image Processing*, Vol. 2, pp. II–525–II–528 vol.2, 2002.
- [57] S. Katzenbeisser and F. Petitolas. Information hiding techniques for steganography and digital watermarking. *EDPACS*, Vol. 28, No. 6, pp. 1–2, 2000.
- [58] J.T. Kent. The Fisher-Bingham distribution on the sphere. *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 71–80, 1982.
- [59] S. Khan, O. Javed, Z. Rasheed, and M. Shah. Human tracking in multiple cameras. In *Proc. of IEEE International Conference on Computer Vision*, Vol. 1, pp. 331–336, 2001.
- [60] M. Kharrazi, H.T. Sencar, and N. Memon. Blind source camera identification. In *Proc. of IEEE International Conference on Image Processing*, Vol. 1, pp. 24–27, 2004.
- [61] M. Kono. Statistics of paleomagnetic inclination data. *Journal of Geophysical Research*, Vol. 85, No. B7, pp. 3878–3882, 1980.

- [62] P. Kovesi. Shapelets correlated with surface normals produce surfaces. In *Proc. of IEEE International Conference on Computer Vision*, Vol. 2, pp. 994–1001, 17–21, 2005.
- [63] L. Kratz and K. Nishino. Tracking with local spatio-temporal motion patterns in extremely crowded scenes. In *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2010.
- [64] D.J. Kriegman and P.N. Belhumeur. What shadows reveal about object structure. *Journal of the Optical Society of America A*, Vol. 18, No. 8, pp. 1804–1813, 2001.
- [65] T. Kuparinen and V. Kyrki. Optimal reconstruction of approximate planar surfaces using photometric stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 31, No. 12, pp. 2282–2289, 2009.
- [66] Sin-Joo Lee and Sung-Hwan Jung. A survey of watermarking techniques applied to multimedia. In *Proc. of IEEE International Symposium on Industrial Electronics*, Vol. 1, pp. 272–277, 2001.
- [67] C. Lei and Y. Yang. Optical flow estimation on coarse-to-fine region-trees using discrete optimization. In *Proc. of IEEE International Conference on Computer Vision*, 2009.
- [68] Zhouchen Lin, Rongrong Wang, Xiaoou Tang, and Heung-Yeung Shum. Detecting doctored images using camera response normality and consistency. In *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 1, pp. 1087–1092, 2005.
- [69] Ce Liu, W.T. Freeman, R. Szeliski, and Sing Bing Kang. Noise estimation from a single image. In *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 1, pp. 901–908, 2006.

- [70] B.D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Proc. of International joint conference on artificial intelligence*, Vol. 3, pp. 674–679. Citeseer, 1981.
- [71] J. Lukáš, J. Fridrich, and M. Goljan. Determining digital image origin using sensor imperfections. In *Proc. of Society of Photo-Optical Instrumentation Engineers Conference*, Vol. 5685, pp. 249–260, 2005.
- [72] J. Lukáš, J. Fridrich, and M. Goljan. Detecting digital image forgeries using sensor pattern noise. In *Proc. of Society of Photo-Optical Instrumentation Engineers Conference*, Vol. 6072, pp. 362–372, 2006.
- [73] J. Malik, S. Belongie, T. Leung, and J. Shi. Contour and texture analysis for image segmentation. *International Journal of Computer Vision*, Vol. 43, No. 1, pp. 7–27, 2001.
- [74] S.P. Mallick, T.E. Zickler, D.J. Kriegman, and P.N. Belhumeur. Beyond lambert: Reconstructing specular surfaces using color. In *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 619–626, 2005.
- [75] KV Mardia and P.E. Jupp. *Directional statistics*. John Wiley & Sons Inc, 2000.
- [76] P. Martin, R. Philippe, et al. Influence of the noise model on level set active contour segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 26, No. 6, pp. 799–803, 2004.
- [77] Y. Matsushita and S. Lin. Radiometric calibration from noise distributions. In *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1–8, 2007.
- [78] S.J. McKenna, S. Jabri, Z. Duric, A. Rosenfeld, and H. Wechsler. Tracking groups of people. *Computer Vision and Image Understanding*, Vol. 80, No. 1, pp. 42–56, 2000.

- [79] J.C. Mullikin, L.J. van Vliet, H. Netten, F.R. Boddeke, G. van der Feltz, and I.T. Young. Methods for CCD camera characterization. In *Proc. of SPIE Image Acquisition and Scientific Imaging Systems*, Vol. 2173, pp. 73–84, 1994.
- [80] Tian-Tsong Ng, Shih-Fu Chang, and Qibin Sun. Blind detection of photomontage using higher order statistics. In *Proc. of International Symposium on Circuits and Systems*, Vol. 5, pp. V–688–V–691, 2004.
- [81] I. Omer and M. Werman. Color lines: image specific color representation. In *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2004.
- [82] K. Pearson. X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophical Magazine Series 5*, Vol. 50, No. 302, pp. 157–175, 1900.
- [83] K. Pearson. LIII. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine Series 6*, Vol. 2, No. 11, pp. 559–572, 1901.
- [84] Patrick Pérez, Michel Gangnet, and Andrew Blake. Poisson image editing. In *Proc. of International Conference and Exhibition on Computer Graphics and Interactive Techniques*, pp. 313–318, 2003.
- [85] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 12, No. 7, pp. 629–639, 1990.
- [86] A.C. Popescu and H. Farid. Exposing digital forgeries by detecting traces of re-sampling. *IEEE Transactions on Signal Processing*, Vol. 53, No. 2, pp. 758–767, 2005.

- [87] R. Ramamoorthi and P. Hanrahan. On the relationship between radiance and irradiance: determining the illumination from images of a convex Lambertian object. *Journal of Optical Society of America A*, Vol. 18, No. 10, pp. 2448–2459, 2001.
- [88] Y. Sato and K. Ikeuchi. Temporal-color space analysis of reflection. *Journal of the Optical Society of America A*, Vol. 11, No. 11, pp. 2990–3002, 1994.
- [89] H. Schaeben. Parameterizations and probability distributions of orientations. *Texture, Stress, and Microstructure*, Vol. 13, No. 1, pp. 51–54, 1900.
- [90] K. Schl  
”uns. Photometric stereo for non-Lambertian surfaces using color information. In *Proc. of International Conference on Computer Analysis of Images and Patterns*, p. 451. Springer-Verlag, 1993.
- [91] K. Schl  
”uns. The irradiance error and its effect in photometric stereo. Technical report, Computer Science Department of The University of Auckland, 1997.
- [92] L. Shafarenko, H. Petrou, and J. Kittler. Histogram-based segmentation in a perceptually uniform color space. *IEEE Transactions on Image Processing*, Vol. 7, No. 9, pp. 1354–1358, 1998.
- [93] Xiaohui Shen and Ying Wu. Sparsity model for robust optical flow estimation at motion discontinuities. In *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2010.
- [94] J. Shi and C. Tomasi. Good features to track. In *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 593–593, 1994.
- [95] T. Simchony, R. Chellappa, and M. Shao. Direct analytical methods for solving poisson equations in computer vision problems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 12, No. 5, pp. 435–446, 1990.

- [96] D.L. Snyder, A.M. Hammoud, and R.L. White. Image recovery from data acquired with a charge-coupled-device camera. *Journal of the Optical Society of America A*, Vol. 10, No. 5, pp. 1014–1023, 1993.
- [97] AD Spence and MJ Chantler. Optimal illumination for three-image photometric stereo using sensitivity analysis. *IEE Proc. -Vision, Image, and Signal Processing*, Vol. 153, No. 2, p. 149, 2006.
- [98] P.F. Sturm and S.J. Maybank. On plane-based camera calibration: A general algorithm, singularities, applications. In *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 1, pp. 432–437, 1999.
- [99] J. Sun, J. Jia, C.K. Tang, and H.Y. Shum. Poisson matting. In *Proc. of International Conference and Exhibition on Computer Graphics and Interactive Techniques*, pp. 315–321, 2004.
- [100] J. Sun, M. Smith, L. Smith, and A. Farooq. Examining the uncertainty of the recovered surface normal in three light photometric stereo. *Image and Vision Computing*, Vol. 25, No. 7, pp. 1073–1079, 2007.
- [101] J. Takamatsu, Y. Matsushita, and K. Ikeuchi. Estimating camera response functions using probabilistic intensity similarity. In *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1–8, 23–28, 2008.
- [102] J. Takamatsu, Y. Matsushita, and K. Ikeuchi. Estimating radiometric response functions from image noise variance. In *Proc. of IEEE European Conference on Computer Vision*, pp. 623–637. Springer, 2008.
- [103] R. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal of Robotics and Automation*, Vol. 3, No. 4, pp. 323–344, 1987.

- [104] Y. Tsin, V. Ramesh, and T. Kanade. Statistical calibration of CCD imaging process. In *Proc. of IEEE International Conference on Computer Vision*, Vol. 1, pp. 480–487, 2001.
- [105] Tran Van Lanh, Kai-Sen Chong, Sabu Emmanuel, and Mohan S Kankanhalli. A survey on digital camera image forensic methods. In *Proc. of IEEE International Conference on Multimedia and Expo*, pp. 16–19, 2007.
- [106] G. E. Smith W. S. Boyle. Charge coupled semiconductor devices. In *Bell System Technical Journal*, No. 49 in 4, pp. 587–593, 1969.
- [107] W. Wang and H. Farid. Exposing digital forgeries in interlaced and deinterlaced video. *IEEE Transactions on Information Forensics and Security*, Vol. 2, No. 3, pp. 438–449, 2007.
- [108] W. Wang and H. Farid. Exposing digital forgeries in video by detecting duplication. In *Proc. of Workshop on Multimedia & security in International Multimedia Conference*, pp. 35–42. ACM Press New York, NY, USA, 2007.
- [109] X. Wang, B. Xuan, and S. Peng. Digital image forgery detection based on the consistency of defocus blur. In *Proc. of International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pp. 192–195. IEEE, 2008.
- [110] N. Wiener. Extrapolation interpolation, and smoothing of stationary time series: with engineering applications. *Electrical engineering*, 1964.
- [111] K. Wiesenfeld and F. Moss. Stochastic resonance and the benefits of noise: from ice ages to crayfish and SQUIDs. *Nature*, Vol. 373, No. 6509, pp. 33–36, 1995.
- [112] Bennett Wilburn, Hui Xu, and Yasuyuki Matsushita. Radiometric calibration using temporal irradiance mixtures. In *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1–7, June 2008.

- [113] R.J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, Vol. 19, No. 1, pp. 139–144, 1980.
- [114] Mingjun Wu, Xianrong Peng, Qiheng Zhang, and Rujin Zhao. Patches-based Markov random field model for multiple object tracking under occlusion. *Signal Processing*, Vol. 90, No. 5, pp. 1518 – 1529, 2010. Special Section on Statistical Signal & Array Processing.
- [115] L. Xu, J. Chen, and J. Jia. A segmentation based variational model for accurate optical flow estimation. In *Proc. of IEEE European Conference on Computer Vision*, p. 684. Springer-Verlag, 2008.
- [116] L. Xu, J. Jia, and Y. Matsushita. Motion detail preserving optical flow estimation. In *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2010.
- [117] Shuiming Ye, Qibin Sun, and Ee-Chien Chang. Detecting digital image forgeries by measuring inconsistencies of blocking artifact. In *Proc. of IEEE International Conference on Multimedia and Expo*, pp. 12–15, 2007.
- [118] A.L. Yuille, D. Snow, R. Epstein, and P.N. Belhumeur. Determining generative models of objects under varying illumination: Shape and albedo from multiple images using svd and integrability. *International Journal of Computer Vision*, Vol. 35, No. 3, pp. 203–222, 1999.
- [119] Z. ZHANG. Flexible camera calibration by viewing a plane from unknown orientations. In *Proc. of IEEE International Conference on Computer Vision*, pp. 666–673, 1999.
- [120] 小林理弘, 石渡信吾. アクチュエータの確率的制御. 応用物理学関係連合講演会, 2007.

- [121] 石渡信吾, 小泉一弥. 確率共鳴による微弱信号検出とその応用. 応用物理学関係連合講演会, 2006.
- [122] 中井検裕, 縄田和満, 松原望. 基礎統計学 統計学入門. 東京大学出版会, 1991.
- [123] 浜松ホトニクス. FFT-CCD エリアイメージセンサの特性と使い方. Technical report, 浜松ホトニクス, 2003.

## 発表文献

### [論文誌]

1. Michihiro Kobayashi, Takahiro Okabe and Yoichi Sato, "Video Forgery Detection Based on Inconsistency in Noise Level Functions", *IEEE Transactions on Information Forensics and Security*, (Accept with mandatory minor revisions)

### [国際会議]

2. Michihiro Kobayashi, Takahiro Okabe and Yoichi Sato, "Detecting Video Forgeries Based on Noise Characteristics", *Pacific-Rim Symposium on Image and Video Technology 2009 (PSIVT2009)*, pp. 306-317, 2009.

### [国内会議(査読あり)]

3. 小林理弘, 岡部孝弘, 佐藤洋一, "ノイズ特性にもとづく映像の改ざん検出", 画像の認識・理解シンポジウム (MIRU2008), IS4-24, 2008.
4. 小林理弘, 岡部孝弘, 佐藤洋一, "複数のノイズレベル関数の同時推定による映像改ざん検出", 画像の認識・理解シンポジウム (MIRU2010), OS13-1, 2010.
5. 小林理弘, 松下康之, 岡部孝弘, 佐藤洋一, "校正誤差を考慮した照度差ステレオからの形状復元", 画像の認識・理解シンポジウム (MIRU2010), IS2-54, 2010.