

画像のエッジとサンプリング位置の位相差に 着目した動きベクトル検出に関する研究

Research on Motion Estimation Considering
Phase Difference between an Edge Model and a
Sampling Position of Digital Images

2015年2月

早稲田大学大学院 基幹理工学研究科
情報理工学専攻 画像情報研究

木村 淳一

目次

第1章 序論

1.1	研究の背景	1
1.2	研究の目的	3
1.3	本論文の構成	4
1.4	本研究の位置づけ	6
1.4.1	研究全体の位置づけ	6
1.4.2	MV検出単位	9
1.4.3	処理に用いるフレームの時間的位置関係	9
1.4.4	MVの精度	9
1.4.5	並進以外の動きへの対応	10
1.5	従来の研究と問題点	11
1.5.1	参照画像生成・評価関数を改善する方法	11
1.5.2	MVの連続性を用いる方法	15
1.5.3	画面の大局的な情報を用いる方法	17
1.5.4	インプリメンテーション	21

第2章 複数フレームを用いた動きベクトル検出

2.1	はじめに	22
2.2	ブロックマッチングの誤差解析	22
2.3	複数のフレームを用いたMV検出方法の提案	28
2.3.1	提案法の評価関数	28
2.3.2	フィルタ係数の決定	32
2.3.3	時間方向適応処理	33
2.4	評価	34
2.4.1	シミュレーション実験条件	34
2.4.2	シミュレーション実験結果と考察	37
2.4.2.1	提案方法の特性	37
2.4.2.2	輝度差／BM誤差の改良による方式との比較	39
2.4.2.3	MV連続性を利用する方式との比較	41
2.4.2.4	インプリメンテーションに関する考察	43
2.4.2.5	平行移動以外の動きの画像に対する考察	44
2.5	むすび	47

第3章	maxDVSEを用いた動きベクトル検出	
3.1	はじめに	48
3.2	maxDVSEの算出	49
3.3	maxDVSEを用いたMV検出方法	50
3.3.1	アンカーブロックの判定	51
3.3.2	動きベクトルの確定	52
3.3.3	詳細アルゴリズム	52
3.4	評価	54
3.4.1	シミュレーション	54
3.4.2	シミュレーション結果と考察	55
3.4.2.1	BM誤差の分布	55
3.4.2.2	DVSEの特性	56
3.4.2.3	MV検出誤り率および処理時間の結果	57
3.4.2.4	ノイズ画像に対する結果	59
3.4.2.5	異なるブロックサイズ	60
3.4.2.6	並進ベクトル以外を含むシーケンスでの結果	61
3.4.2.7	アンカーベクトルの評価	63
3.4.2.8	誤りの分析	64
3.4.2.9	提案方法の画像適応性	66
3.4.2.10	第2章のMV検出方法との関係	66
3.5	むすび	67
第4章	画像モデルを用いた動き補償フレーム間予測効率の検討	
4.1	はじめに	68
4.2	モデル化	69
4.2.1	画像のモデル化	69
4.2.2	MCのモデル化	70
4.2.3	動き補償処理	71
4.2.3.1	画像モデルを用いた動き補償	71
4.2.3.2	単画素予測 (Simple prediction)	72
4.2.3.3	ハーフペル予測 (Half pel prediction)	73
4.2.3.4	平均値型予測 (Average Type Prediction)	74
4.3	動き補償誤差電力の定量的評価	76
4.3.1	予測効率の定義	76

4.3.2	予測効率の計算	77
4.4	実画像を用いたシミュレーション	81
4.4.1	シミュレーション	81
4.4.2	シミュレーション結果	81
4.5	ノイズ・変形に対する各予測方式の耐性評価	83
4.6	簡略手法による平均値型予測の予測効率の比較	85
4.7	むすび	86
第5章	結論	
5.1	まとめ	88
5.2	今後の課題	89
5.3	期待される効果	89

謝辞

参考文献

図一覽

図1-1	動きベクトルの応用例	1
図1-2	繰り返しパターンが多く含まれる画像の例	3
図1-3	繰り返しパターンにおける誤りMVの分布例	4
図1-4	本研究の位置付け	7
図1-5	2章の技術の適用範囲	8
図1-6	3章の技術の適用範囲	8
図1-7	MV検出	10
図1-8	小数精度画素	12
図1-9	小数精度MVの検出方法	13
図1-10	パラボラフィッティングによる小数精度MV検出	14
図1-11	マルチベースラインにおけるカメラ位置	15
図1-12	MV predictionに用いるMV	16
図1-13	3Dサーチの候補MV位置	17
図1-14	4分木階層画像	18
図1-15	Middleburyベンチマーク (2013/9/6)	21
図2-1	周期画像における動きベクトル	24
図2-2	画像エッジ部のサンプリング誤差	27
図2-3	動きベクトルの判定例とサンプルブロック	27
図2-4	$F_{N+i}(kX+i v_1X)$ の画素の位置	29
図2-5	フレーム間距離に比例した動きベクトルにより示される画素の位置	30
図2-6	提案方法のフィルタレスポンス	33
図2-7	テストシーケンスと正解ベクトル	35
図2-8	時間方向適用処理ありと適用処理なし	39
図2-9	提案方法とMBLの比較	40
図2-10	処理時間とMV検出誤り率	44
図2-11	提案方法とMDP_flow2の比較	45
図3-1	DVSEの最大値候補箇所	50
図3-2	ブロックマッチング誤差とmaxDVSEとの関係の例	51
図3-3	候補動きベクトルから動きベクトルの選択例	52
図3-4	提案方法の詳細アルゴリズム	53

図3-5	正解ベクトルのブロックマッチング誤差の分布	55
図3-6	候補MV選定閾値を変化させた時の誤り率の変化	56
図3-7	ノイズ画像に対する誤り率	58
図3-8	各方式のMV比較	59
図3-9	異なるブロックサイズでの誤り率比較	60
図3-10	他のシーケンスのシミュレーション結果	62
図3-11	提案方法とMDPの細部の比較	62
図3-12	アンカーブロックの位置	64
図4-1	エッジモデルおよびインパルスモデル	70
図4-2	予測画素の生成法	71
図4-3	MPEG-2における予測構造	75
図4-4	ε_0 の確率密度関数	76
図4-5	エッジモデルを用いた予測効率	78
図4-6	インパルスモデルを用いた予測効率	78
図4-7	単画素予測に対する予測効率比	79
図4-8	平均値型予測の予測距離比と予測効率の関係	80
図4-9	シミュレーション結果	83
図4-10	予測距離ごとの位相差の関係	85
図4-11	図4-10のモデルによる予測誤差	86

表一覽

表 2-1 提案方法とそのフィルタ係数	32
表 2-2 シミュレーション諸元	34
表 2-3 シミュレーション結果(1)	38
表 2-4 シミュレーション結果(2)	42
表 2-5 処理時間とMV検出誤り率	43
表 3-1 シミュレーション諸元	54
表 3-2 シミュレーション結果	58
表 3-3 処理時間	58
表 3-4 異なるブロックサイズでの結果	60
表 3-5 他のシーケンスのシミュレーション結果 (参考)	63
表 3-6 他のシーケンスのシミュレーション結果 (参考) ($\text{thc} = \sqrt{2}$) ...	63
表 3-7 アンカーベクトルの統計量	64
表 3-8 動きベクトル検出の誤り原因	65
表 3-9 第 2 章・第 3 章の方法の比較	66
表 4-1 シミュレーション条件	81

第1章 序論

1.1 研究の背景

動画像の処理において動きベクトル (Motion Vector, MV) 検出は重要な技術となっている。MPEG-2[1]や H.264[2]に代表される動画像符号化 (圧縮蓄積, あるいは圧縮伝送) を始め, 異なるフレームレートの機器間を高画質に接続あるいは通信するためのフレームレート変換[3] [4], フレーム内挿処理[5][6], 液晶テレビの画像の動きによるぼけを改善し高画質化するためのフレーム内挿[7] [8] [9], 複数のフレームを用いて画像の高解像度化を図る超解像処理[10][11], カメラの手振れ補正[6]など, 民生機器を含めて多くの分野で MV 検出が使われている。例えば, 図 1-1 に示すように液晶ディスプレイのようなホールド型の表示装置では, 画面が動くと, 物体の表示位置と人間の脳での予測位置の差の大きさにより生じるホールドぼけが生じる。この解消法として通常の 60Hz のフレームから MV を用いて生成した新たなフレームを内挿し, 120Hz の表示を行うことによりホールドぼけのない高画質を実現することができる[9]。

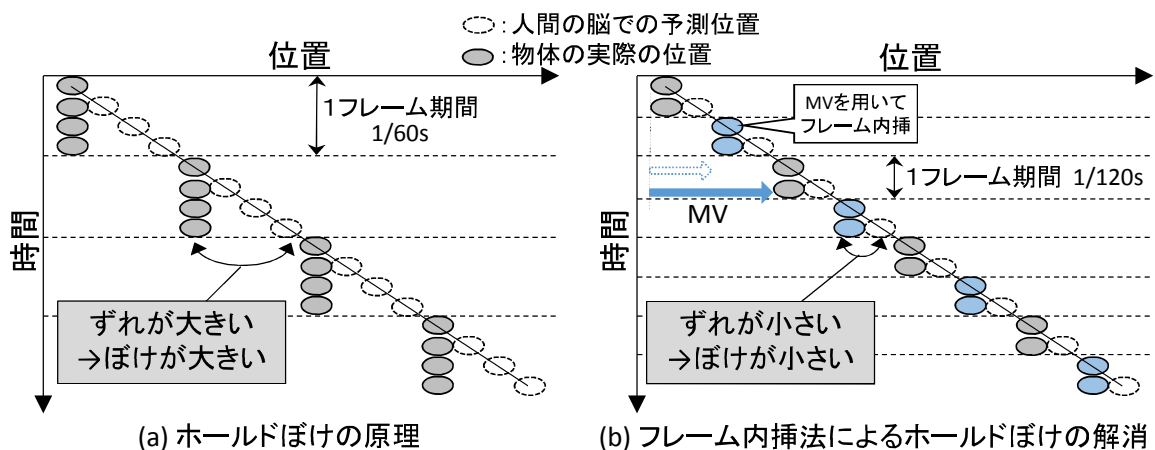


図 1-1 動きベクトルの応用例ー液晶ディスプレイにおけるホールドぼけの解消 ([9]に加筆)

Figure 1-1 An example of application using motion vector estimation

この他にも, 2つあるいはそれ以上の異なる位置から撮影した画像から距離情報を抽出するステレオビジョン処理[12]などにも MV 検出と同様の技術が使われている。このように, MV 検出は身近な様々な機器において利用されており, 今後ますます MV 検出を利用した機器やソフトウェアが普及することが想定される。

MV 検出の技術は、TV 電話や TV 会議に用いる動画像符号化の標準 H.261[13]の標準化から 30 年弱をかけて発展してきた。その牽引役はデジタル放送や、DVD、Blu-ray といった蓄積メディアなどの民生品である。しかし、動画像符号化の場合は、もし検出した MV が誤っていても、検出した MV により生成される予測画像と符号化する画像の間の差分画像の情報を符号化し補正する。このため、検出する MV は、必ずしも被写体やカメラの動きを正しく表していなくても問題が少なかった。もし、誤った MV を検出した場合、差分画像の情報量が増え、限られた情報量に抑えるために、該当画像あるいはその後の画像に割り当てられる符号量を削減し、その結果、画質が劣化するという、いわば間接的な劣化であった。1990 年代から 2000 年台初頭にかけては LSI の集積度などハードウェアの性能が十分でなく、あるいは、CPU の速度・処理性能なども不足していたこともあり、MV 検出の検討は、少ない処理量で、いかに高速に、いかに広範囲の MV を検出するかに注力されてきた[14][15][16][17]。

2000 年代中ごろより、動画像符号化もより高い圧縮率を求められ、また、ハードウェア・CPU の処理能力の向上に伴い、フレーム内挿、超解像、ステレオビジョンなどの新たなアプリケーションが民生機器に搭載されるようになってきた。こうしたアプリケーションにおいては、被写体あるいはカメラの動きを表す「正確な MV」を検出することが必要となる。たとえば、フレーム内挿において、正しい MV を検出できないと、誤った画像を挿入することになり、画面内の動きが不連続になる致命的な画質の劣化を招いてしまう。

さらに近年では、スマートフォンをはじめとするモバイル機器の急激な普及に伴い、ソフトウェアでの実装も重要度を増してきている。モバイル機器では消費電力を低く抑えるために、こうした MV 検出処理を少ない処理数、あるいは高い CPU 処理効率を実現する MV 検出アルゴリズムの必要性が高まっている。

こうした正確な MV を検出する研究は進んできているものの、まだ、その性能は実用上必ずしも十分でない。特に、図 1-2 に示すような同じ形の窓や、織物の模様のように、類似したパターンが規則的に並ぶ「繰返しパターン」画像では、本来対応すべきパターンとは別の誤ったパターンとの対応がとられ、正確な MV を検出できない「MV 検出誤り」が多く発生することが文献 [8]にも指摘されている。こうした MV 検出誤りは、特に、パターンの繰返し数が多く、また、繰返しパターンの領域が広い場合に、より顕著になる。

正確な MV 検出のために、これまで、小数精度の MV を用いる方法や、空間方向の MV の連続性（相関）を利用する方法、時間／空間双方の相関を利用する方法などが開発されてきたが、特に、繰返しパターンに対しては十分な性能が得られていない。また、近年では、ステレオビジョンの応用を目的とした 2 つの画像間の対応点検出において、画像の大局的な情報を用い、かつ、MV の連続性、被写体の形状、画素値の変化等の最適化を図る手法が開発され[18]、高い性能が得られている。この技術は MV 検出にも適用が可能であるが、処理量が多い点が問題である。このように、少ない処理量を実現することと、正確な MV を検出することが両立できないことが課

題である。

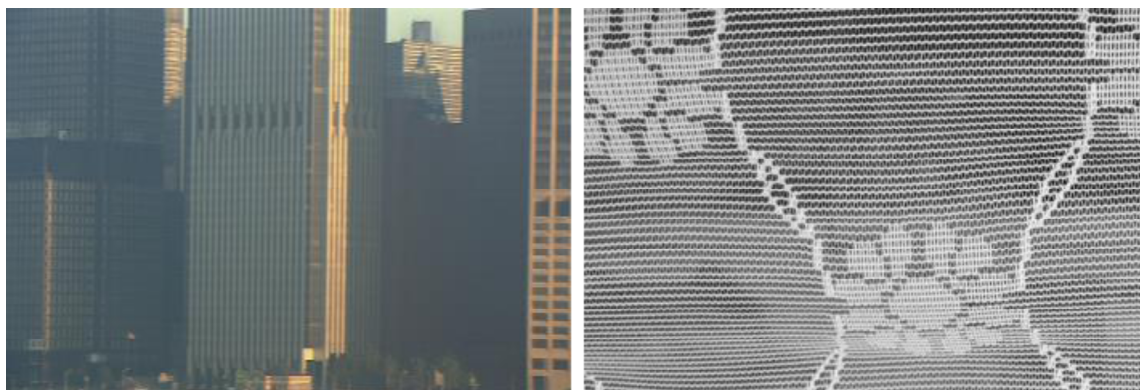


図 1-2 繰り返しパターンが多く含まれる画像の例
Figure 1-2 Examples of spatial periodic pattern

1.2 研究の目的

そこで本研究では、ブロックマッチングを用いて、少ない処理量で、被写体あるいはカメラの動きを表す正確な MV を検出することを第 1 の目的とする。対象画像は、画面全体あるいは画面内の大きな領域が並進している画像を対象とし、特に、MV 検出誤りが多く発生する「繰り返しパターン」を多く含む画像を主たる対象とする。並進以外の動きの場合も、局所的に見れば並進の組み合わせと考えることができるため、並進の動きを正確に検出できることが、他の動きの MV 検出においても重要な技術となる。

繰り返しパターンを多く含む画像の場合、検出誤りの MV の大きさ、方向の分布は図 1-3 のように離散的になる。すなわち、図のような繰り返しパターン（図では動きがわかりやすいように一部のパターンの輝度を変化させてある）が動いている場合、正解の MV（図右の原点から●印への MV）の代わりに、パターン 1 つ分となりの MV（例えば、原点から図右の●印の右側の○印への MV）が検出される場合のように、正解 MV の他に、繰り返しパターンのピッチで誤った MV が分布する。

本研究では、こうした離散的な分布の MV の中から正しい MV（図 1-3 右の●印の MV）を検出できるか否かを判断することにより、MV 検出精度の評価を行う。

本研究では、第 2 の目的として、MV を用いた応用の一つである動き補償フレーム間予測 (MC: Motion Compensation) 方式の評価方法に関して検討を行う。MC 方式の評価方法では、評価画像に依存しない理想的な MC の予測効率を求める手法を開発する。

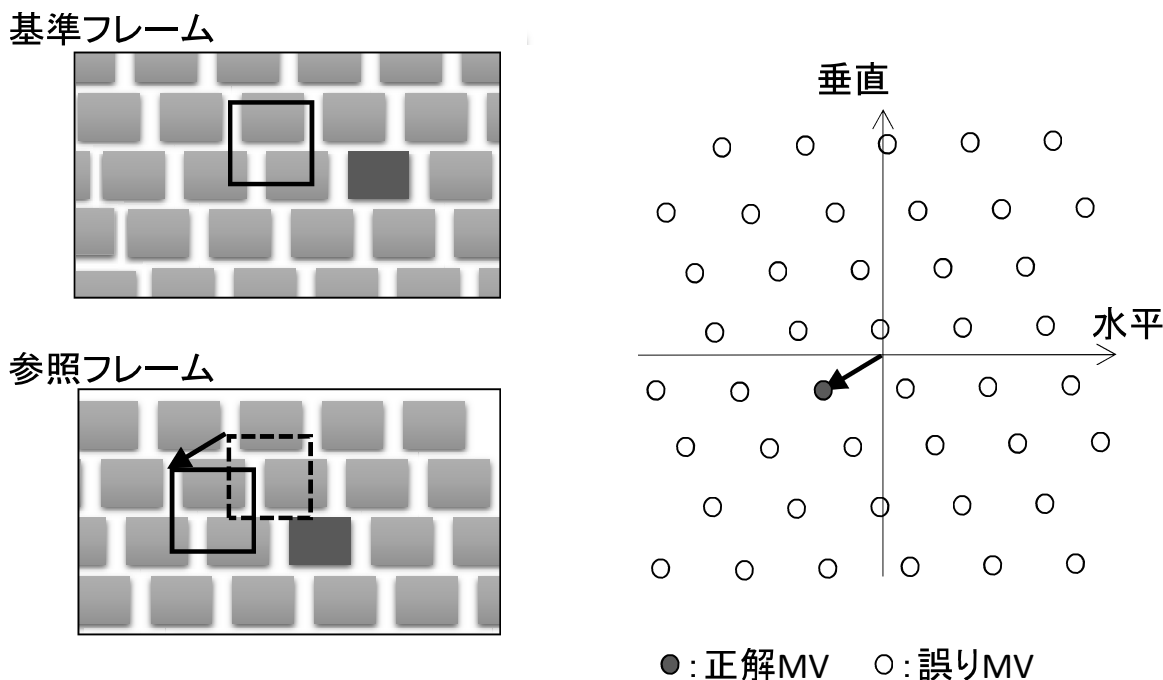


図 1-3 繰り返しパターンにおける誤り MV の分布例

Figure 1-3 An example of distribution of false MV in spatially periodic pattern

1.3 本論文の構成

本論文では、前節に示した目的を達成するために、画像のエッジ位置とサンプリング位置の位相差に着目する。まず、第2章の前半において、画像のエッジ位置とサンプリング位置の位相差に着目し、画像モデルを用いて正確な MV が検出できない原因を分析する。次に、この分析結果を基に、第2章の後半、第3章において正確な MV を検出する方法を2つのアプローチから提案する。また、第4章においては、画像のエッジ位置とサンプリング位置の位相差に着目した画像モデルを、MV 検出の応用である、動き補償フレーム間予測方法の予測方法の比較に適用する手法を提案する。

第2章[19][20]では、画像のエッジ位置とサンプリング位置の位相差に着目し、まず、画像モデルを用いて、正確な MV が検出できない原因の分析を行い、MV 検出誤りの原因の一つが MV の量子化誤差であることを示す。次に、分析結果を基に、MV の量子化誤差の影響を排除するために、複数のフレームを用いた MV 検出方法を提案する。第2.2節では、1次元の画像モデルを用い、正しい MV と誤った MV のそれぞれの場合のブロックマッチング誤差を算出し比較することにより、MV の量子化誤差成分により生じるブロックマッチング誤差 (DVSE : block

matching Distortion caused by motion Vector Sampling Error) が MV 検出誤りの原因であることを示す。第 2. 3 節では、複数のフレームを用いて時間方向にフィルタ処理を施すことにより、第 2. 2 節で示した MV の量子化誤差成分を相殺し、DVSE の影響を低減した MV 検出方法を提案する。基準フレームを中心に、時間軸の前方向および後方向の双方に、基準・参照フレームのフレーム間距離の逡倍に位置するフレーム、2～6 フレームを用いてフィルタ処理を行うことにより MV の量子化誤差成分を相殺する評価関数 Temporally-Predicted Difference (TPD) を提案する。最適なフレーム数、フィルタ係数を算出し、フィルタ処理により生成された新たな参照ブロックを用いてブロックマッチングを行うことにより正確な MV を検出できることを示す。さらに、提案方法の応用として、異なるフレーム数から生成した複数の参照画像を適用的に用いる方法を提案する。適応処理により、時間方向に不規則な変化がある動画像における MV 検出誤りを低減する。第 2. 4 節では、第 2. 3 節で提案した MV 検出方法の評価を行う。評価には繰り返しパターンを多く含む実画像のシーケンスを用いて、シミュレーションにより提案方法と既存の方法の MV の正解率の比較を行う。さらに、提案方法と既存の手法を組み合わせることにより、MV 検出誤りをさらに低減できることを示す。あわせて、提案方法の処理量の評価を含めた考察を行う。

第 3 章[21][22][23]では、第 2. 2 節の MV 検出誤りの分析結果を基に、MV の量子化誤差の影響が存在することを前提とした、第 2 章とは別のアプローチによる MV 検出方法を提案する。第 3. 2 節では、第 2. 2 節にて示した MV 検出誤りの原因である DVSE の最大値（近似値、maxDVSE）の算出方法を示す。maxDVSE は、MV が正確な MV であった場合でも取り得る最大のブロックマッチング誤差である。halfpel の MV 検出と同様の処理を行うことにより、基準画像のみからブロックごとに maxDVSE を算出できることを示す。第 3. 3 節では、maxDVSE が正確な MV が取り得る最大のブロックマッチング誤差であることを利用した MV 検出方法を提案する。ブロックごとに maxDVSE の値を算出し、次にフルサーチの MV の探索を実行する。このとき、ブロックマッチング誤差が maxDVSE 以下の全ての MV をそのブロックの MV 候補として記憶しておく。1 フレーム全てのブロックマッチング処理が完了した後に、記憶しておいた MV から、MV が一意的に定まるアンカーブロックを検出し、ブロック間の MV の連続性を用いて、アンカーブロックの MV と類似した MV を周囲のブロックから再帰的に探索することにより各ブロックの最終的な MV を確定する。第 3. 4 節では、第 3. 3 節で提案した MV 検出方法の評価を行う。繰り返しパターンを多く含む実画像シーケンスを用いたシミュレーションにより MV の正解率の評価を実施し、繰り返しパターン画像においても MV 検出誤りが少ないことを、他方式と比較して実証する。あわせて処理量が少ないことを示す。これに加え、ノイズ画像への耐性、ブロックサイズを変えた時の影響、提案方法の検出誤りの箇所の評価などを考察する。

第 4 章[24][25][26][27]では、画像モデルを用いて、MV を用いた応用の一つである動き補償フレーム間予測 (MC: Motion Compensation) 方式の評価を行う。1 次元の画像モデルと、単画素

予測，ハーフペル予測，平均値型予測の3つのMC方式のモデルを用いて，それぞれのMC方式の予測効率を定量的に算出する方法を提案する．第4.2節ではMCを単画素予測，ハーフペル予測，平均値型予測の3つの予測値生成法のタイプに分類し，これらMC方式の動作を数式にて表現をするモデル化を行う．あわせて，周波数帯域をパラメータとするエッジモデルおよびインパルスモデルの2つの画像モデルを定義し，それぞれの画像モデルとMC方式のモデルを用いて各MC方式の予測効率を算出する方法を提案する．提案方法では，画像モデルのエッジあるいはインパルスの位置とサンプリング位置の位相差を確率変数とし，予測誤差電力の期待値を定量的に算出し，この予測誤差電力の期待値の逆数を予測効率と定義する．特に，平均値型予測では平均値処理により誤差電力が半減するだけでなく，2つの画像の予測距離比に応じて位相差成分を相殺する効果が現れ，予測効率が高くなることを示す．第4.4節では実画像シーケンスを用いたシミュレーションにより，第4.3節で提案した予測効率の評価を行う．実画像に含まれるノイズや変形の影響により若干の差はあるものの，ほぼ計算値に近い結果が得られ，提案した方法の妥当性を示す．

1.4 本研究の位置づけ

1.4.1 研究全体の位置づけ

図1-4にMV検出の研究の中での本研究の位置付けを示す．図1-4では，MV検出の研究を「アプリケーション」，「MV検出方法」，「参照画像生成・評価関数」の大きく3つのレイヤに分け，例として該当する著名な方法を記載している．

「アプリケーション」レイヤは，実際の応用用途に対応したレイヤであり，図では「アプリケーション」レイヤの中を，さらに，video coding等の具体的な応用用途である最上位サブレイヤと，MVの評価の方法・実用上の制約，block matching等の実装，full pel precision（整数画素精度のMV検出）等のアプリケーションの基本的な仕様に相当する最下位のサブレイヤとに分けている．

「MV検出方法」レイヤは，MVをどうやって検出するかのアプローチの汎用的なレイヤである．その中には大きくfull search等の基本方式のサブレイヤ（下位）と，基本方式を組み合わせ，あるいは性能向上・処理量削減のための付加的な処理を加えた応用方式のサブレイヤ（上位）に分けている．

「参照画像生成・評価関数」レイヤは，MV検出の基本である2つ，あるいはそれ以上のフレームの，対応する小領域画像の類似性を定量的に評価する手法である．具体的には小領域画像の各対応する画素値の生成，画素値の差分，あるいは特徴量の差分を求める評価関数のレイヤである．「参照画像生成・評価関数」レイヤも，SSD（Sum of Squared Difference），SAD（Sum of Absolute Difference），NCC（Normalized Cross Correlation）等の単純な画素間の差分関数の

サブレイヤ（下位）と、参照画像を生成する、上位のサブレイヤに分けている。

図 1-4 において、本論文の 2 章～4 章の項目は、それぞれ太枠ハイライトにて図示した位置づけとなる。2 章の複数フレームを用いた MV 検出は、主に参照画像生成・評価関数のレイヤとなる（ただし、章の中では、MV 検出方式との組み合わせにより、MV 検出方式としてベンチマークを行っている）、3 章の maxDVSE を用いた MV 検出は、MV 検出方式のレイヤ、そして、4 章の画像モデルを用いた動き補償フレーム間予測効率の検討は、アプリケーションのレイヤとなる。

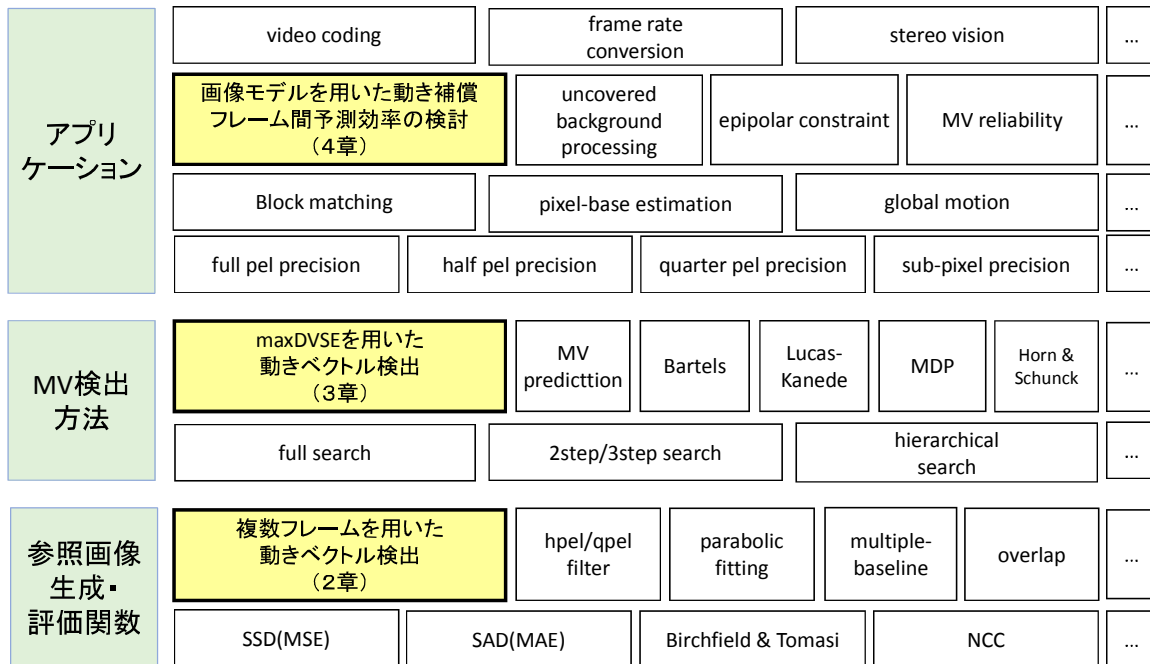


図 1-4 本研究の位置付け

Figure 1-4 The position of this work

図 1-5, 図 1-6 に 2 章, 3 章, それぞれの技術が適用できるレイヤを色付けしている（該当のレイヤよりも上位のレイヤは橙色, 下位のレイヤは青色）。特に濃い色（橙色および青色：図 1-5 の SSD, 図 1-6 の full search, SSD）は、各章で評価した方式の組み合わせである。2 章の複数フレームを用いた MV 検出は参照画像生成・評価関数のレイヤであるため、MV 検出方式のレイヤとの組み合わせが可能となる。ただし、例外として 3 章の方式へは適用はできない。2 章では DVSE の影響を排除しているのに対して、3 章では DVSE の影響が存在することを前提として処理をしており、同じ課題を別のアプローチから取り組んでおり排他的になっているためである。

アプリケーション	video coding	frame rate conversion	stereo vision		...		
	画像モデルを用いた動き補償 フレーム間予測効率の検討 (4章)	uncovered background processing	epipolar constraint	MV reliability	...		
	Block matching	pixel-base estimation	global motion		...		
	full pel precision	half pel precision	quarter pel precision	sub-pixel precision	...		
MV検出方法	maxDVSEを用いた動きベクトル検出 (3章)	MV prediction	Bartels	Lucas-Kanede	MDP	Horn & Schunck	...
	full search	2step/3step search		hierarchical search		...	
参照画像生成・評価関数	複数フレームを用いた動きベクトル検出 (2章)	hpel/qpel filter	parabolic fitting	multiple-baseline	overlap	...	
	SSD(MSE)	SAD(MAE)	Birchfield & Tomasi	NCC		...	

図 1-5 2章の技術の適用範囲

Figure 1-5 Applicable layer of technology described in section 2

アプリケーション	video coding	frame rate conversion	stereo vision		...		
	画像モデルを用いた動き補償 フレーム間予測効率の検討 (4章)	uncovered background processing	epipolar constraint	MV reliability	...		
	Block matching	pixel-base estimation	global motion		...		
	full pel precision	half pel precision	quarter pel precision	sub-pixel precision	...		
MV検出方法	maxDVSEを用いた動きベクトル検出 (3章)	MV prediction	Bartels	Lucas-Kanede	MDP	Horn & Schunck	...
	full search	2step/3step search		hierarchical search		...	
参照画像生成・評価関数	複数フレームを用いた動きベクトル検出 (2章)	hpel/qpel filter	parabolic fitting	multiple-baseline	overlap	...	
	SSD(MSE)	SAD(SAE)	Birchfield & Tomasi	NCC		...	

図 1-6 3章の技術の適用範囲

Figure 1-6 Applicable layer of technology described in section 3

1.4.2 MV 検出単位

本研究では小領域単位での類似度比較による MV 検出を対象としている。MV 検出は、小領域画像単位のもの、画素単位のものに大別できる。小領域画像単位の MV は通常矩形（例えば 8 画素×8 画素、以下ブロックと呼ぶ）の領域あたり 1 つの MV であり、画素単位の MV は文字通り 1 画素あたり 1 つの MV となる。小領域画像単位の MV 検出処理では、2 つのフレームの間における、小領域画像単位での画像の類似性の判定が基本となる。また、画素単位の MV 検出においても、検出対象とする 1 つの画素を中心として周囲の画素を含む領域単位で画像の類似性を判定するケースが多い。すなわち、本研究が対象とする、2 つのフレームの間での小領域画像単位での画像の類似性の判定が、MV 検出において基本の技術となる。

1.4.3 処理に用いるフレームの時間的位置関係

MV 検出の一般的な処理を図 1-7 に示す。基準画像内の小領域画像と類似した小領域画像を参照画像から検出し、その位置のずれを MV とする。基準画像と、参照画像の時間的な位置関係により図 1-7 に示したように、2 通りの場合が存在する。すなわち、図にて Type 1 と示したように、過去の参照画像から MV を検出する場合と、Type 2 と示したように未来のフレームを参照画像とする場合である。いずれのタイプを採用するかは、応用分野により決まる。たとえば、動画像の符号化では Type 1 が中心となり、テレビ受像機のフレーム内層処理では Type 2 が使われることが多い。MV 検出（整数精度）以降の処理は、いずれのタイプも共通の処理となるため、本研究はいずれのタイプの MV 検出にも適用が可能となる。検出される MV の符号が、同じシーケンスの動き、同じ MV の符号の定義（MV の各要素の正の値が上下、左右それぞれどちらの向きを示すか）の条件下であっても、2 つのタイプでは正負が逆転した MV が検出される点異なる。

1.4.4 MV の精度

本研究 2 章・3 章の MV の精度は整数精度とし、図 1-7 の太枠とした「MV 検出（整数精度）」の箇所を対象としている。

MV は応用によって必要な精度（1/4 画素精度等小数の細かさ）が異なる。小数精度の MV 検出は、整数精度の MV を検出した後に、検出した整数精度の MV の近傍を、あるいは整数精度の MV のいくつかを用いて、小数精度の MV を検出することが主流となっている。他の方法として、たとえば、1/2 画素精度（half pel, hpel）の MV 検出を行う場合、原画像のサイズを縦横 2 倍に補間した拡大画像を生成し、拡大画像を用いて MV 検出を行い、検出された MV を 1/2 にすることにより精度よく half pel MV を検出することができる。しかし、この方法は拡大処理が追加となる上、MV 探索の処理量が 4 倍になるため実用的には用いられることは少ない。このように、MV 検出では本研究にて対象としている整数精度の MV 検出が基本の処理となる。

MV 検出（整数精度）の前に，縮小画像を用いて MV 検出を行う階層探索は，その処理量の少なさからも実用上用いられることが多い．たとえば，原画を縦横それぞれ 1/2 に縮小し MV の検出を行い，検出された MV を 2 倍にスケール（いわば"2 画素精度"）した後，その MV の周辺あるいは，その MV を参考に原画像のサイズにて整数画素精度の MV 検出を行う．こうした処理においても，縮小画像を基準に見れば，通常の整数精度の MV 検出と同じ処理となる．

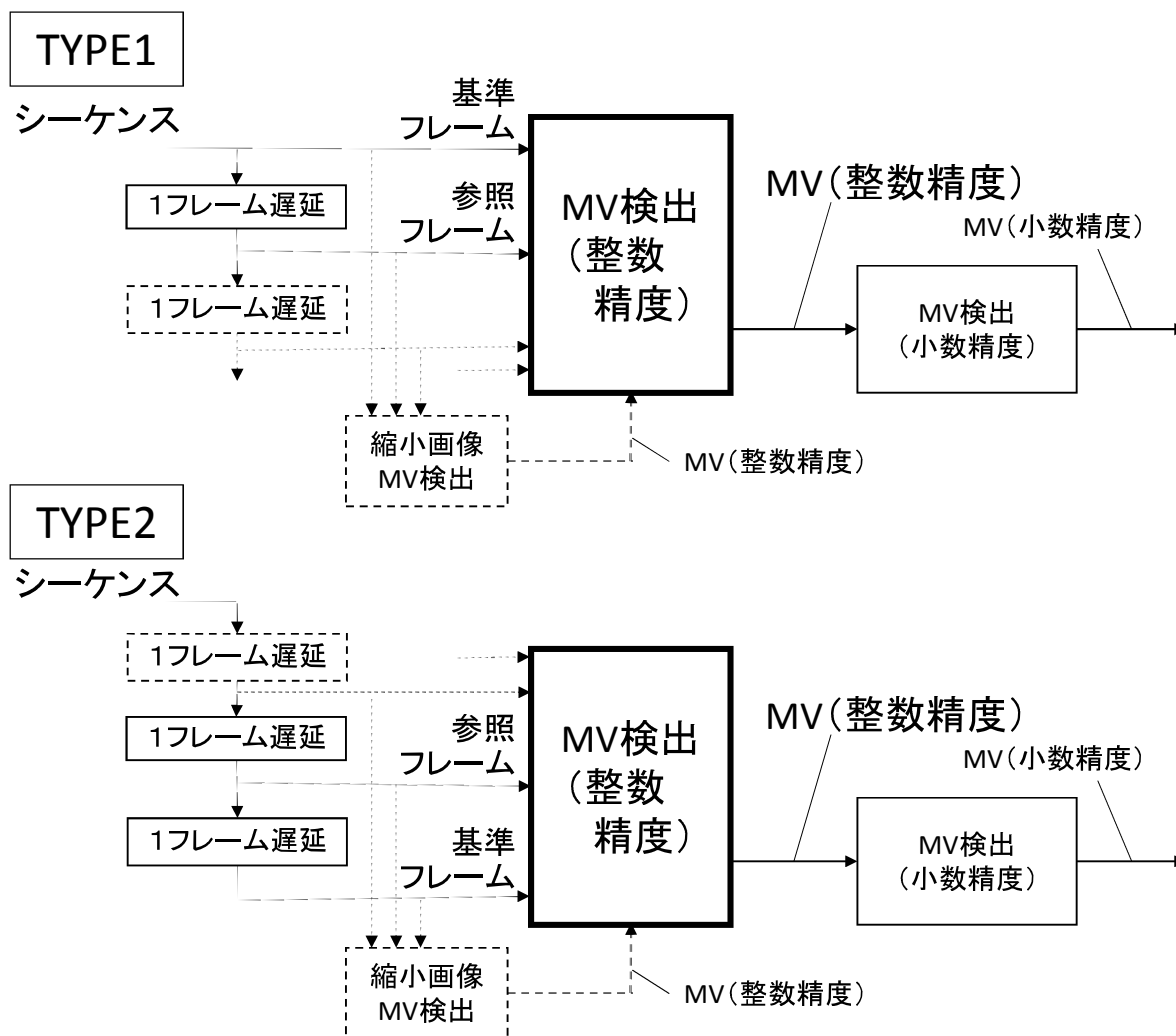


図 1-7 MV 検出

Figure 1-7 Motion estimation

1.4.5 並進以外の動きへの対応

MV 検出を検討する上では，フレーム間の回転，スケール，輝度変化への対処が問題となる．フレーム間の回転，スケール等が大きい応用分野では，画像の局所特徴量によるマッチング，例えば SIFT(Scale-Invariant Feature Transform)[28]などが用いられることもある．しかし，

テレビ放送や個人撮影のビデオを代表とする動画像の処理においては、画面全体は回転あるいはスケールリングしていても、その程度は少なく、局所的、すなわち 8 画素×8 画素程度のブロック等の小領域の近傍の動きで見ると、並進の動きとして扱うことができる。ただし、輝度変化が大きい場合は通常の MV 検出とは別の方法が必要となり、本研究では対象としていない。

1.5 従来の研究と問題点

繰返しパターンを含め、MV 検出誤りを低減し、カメラら被写体の正しい動きを検出する方法として、これまでさまざまな方法が検討されている。本節では、参照画像生成・画素差分の評価関数を改善する方法、MV の連続性を用いる方法、画面の大局的な情報を用いる方法に分けて、個々の方法とその問題点を示す。

1.5.1 参照画像生成・評価関数を改善する方法

MV 検出において、ブロックマッチング、画素単位の MV、いずれにおいても 2 つのフレームの類似位置の検出することが必要となる。類似位置の検出においては、基準画像と参照画像の画素との差分の評価を行うための、参照画像生成法の改善あるいは評価関数の改善により正確な MV を検出することができる。具体的には、参照画像生成の一つである小数精度の MV を用いる方法は、文献[8]に繰返しパターンにおける MV 検出誤りを軽減する方法として有効であるとの結果が示されている。小数精度の MV 検出の具体的な検出手法としては halfpel[1]、パラボラフィッティング[29][30]、Birchfield & Tomasi の手法[31]が知られている。小数精度の MV を用いることにより、大域的なデータを用いることができない場合においても MV 検出誤りは低減する。しかし、その低減量は後述する文献[18][32]に比べると少ない。以下、各方法の概要を記す。なお、以下の概要では、説明を簡単にするために一次元の信号を用いる。

halfpel [33][1], quarterpel [2]

halfpel(hpel) は MPEG-1[33], MPEG-2[1], H.264[2]にて、quarterpel(qpel)は H.264[2]にて採用されている、それぞれ 1/2 画素精度、1/4 画素精度の MV である。それぞれの規格では hpel あるいは qpel の MV がビットストリームに出現した場合の、その時の予測画像の生成法、すなわち MV 指定された画素位置の値の生成法を以下のように規定している。

MPEG-1, MPEG-2 の場合 (図 1-8(a) の X の位置)

$$F(i + 1/2) = (F(i) + F(i + 1) + 1) \gg 1 \quad (1.1)$$

ここで、" $a \gg n$ "は、 a を右へ n ビット右シフト (除算) の演算を示す。H.264 の場合 (図 1-8(b) の X の位置 $F(i + 1/2)$, Y,Z の位置 $F(i + 1/4)$, $F(i + 3/4)$) の値は以下となる。

$$F(i + 1/2) = \text{clip}((F(i - 2) - 5F(i - 1) + 20F(i) + 20F(i + 1) - 5F(i + 2) + F(i + 3) + 16) \gg 5)$$

(1.2)

$$F(i + 1/4) = (F(i) + F(i + 1/2) + 1) \gg 1$$

(1.3)

$$F(i + 3/4) = (F(i + 1) + F(i + 1/2) + 1) \gg 1$$

(1.4)

これらを図示すると図 1-8 となり，○印の整数画素位置のそれぞれの画素の値から，補間フィルタにより，□の位置の画素の値を求めている。

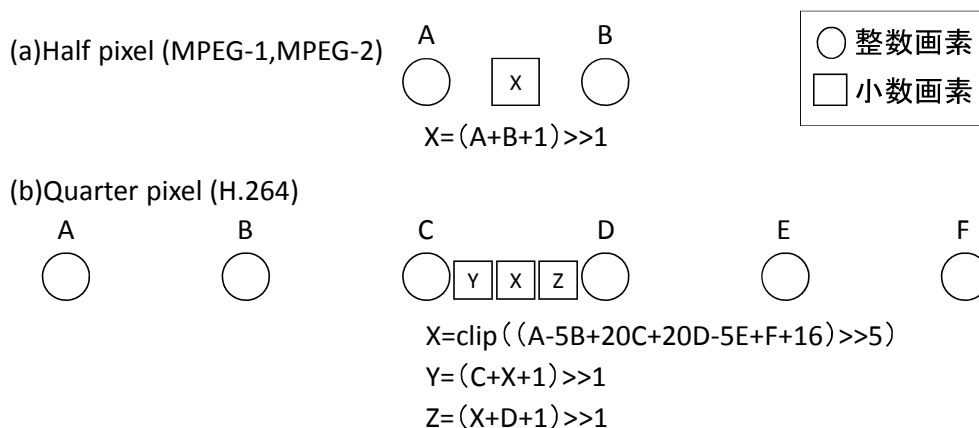
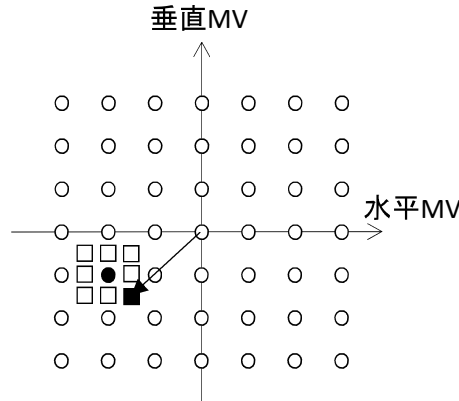


図 1-8 小数精度画素

Figure 1-8 sub-pixels

上記の規格では MV の検出法は規定していないが，図 1-9 のように整数精度の MV (図の●位置) を求めた後に，その周囲の小数精度の MV を検出する two-step search (図の□■の 8 点を探索し，例えば結果として■を検出) が一般的である[34].



- : 整数精度MV(検出) ■: 整数精度MV(検出)
- : 整数精度MV(候補) □: 整数精度MV(候補)

図 1-9 小数精度 MV の検出方法

Figure 1-9 half pel MV estimation

Birchfield & Tomasi の手法 [31]

画素の差を評価する評価関数として Birchfield & Tomasi [31]の手法が知られている. Birchfield & Tomasi はステレオビジョンにおける左右画像の対応を求める方法である. 左目画像の位置 j の画素 (画素値 $F_L(j)$) に対応する右目画像の画素位置 i (画素値 $F_R(i)$) を求めるとき,

$$F_R^+ = F_R\left(i + \frac{1}{2}\right) = \frac{1}{2}(F_R(i) + F_R(i + 1)) \quad (1.5)$$

$$F_R^- = F_R\left(i - \frac{1}{2}\right) = \frac{1}{2}(F_R(i) + F_R(i - 1)) \quad (1.6)$$

$$F_{min} = \min(F_R^-, F_R^+, F_R(i))$$

$$F_{max} = \max(F_R^-, F_R^+, F_R(i))$$

として, 両画素の差の評価値 $D(j, i, F_L, F_R)$ を以下のように定義する.

$$D(j, i, F_L, F_R) = \max\{0, F_L(j) - F_{max}, F_{min} - F_L(j)\} \quad (1.7)$$

画素の差が F_{max} と F_{min} の間の場合は, 差の評価値を 0 としている.

Parabolic fitting (PF) [29]

小数精度 MV の精度, すなわち小数成分の細かさをもつ方法としてパラボラフィッティング (Parabolic fitting) [29]が知られている. パラボラフィッティングは, 図 1-10 に示すように

画像信号値を放物線で近似することにより小数精度の MV を検出するものである。具体的には、図 1-10 のように 3 点から放物線の極小値より、小数成分 ε を以下による求める。

$$\varepsilon = \frac{F(i-1) - F(i+1)}{2F(i-1) - 4F(i) + 2F(i+1)} \quad (1.8)$$

ここで $F(i)$ は位置 i における画素の値である。

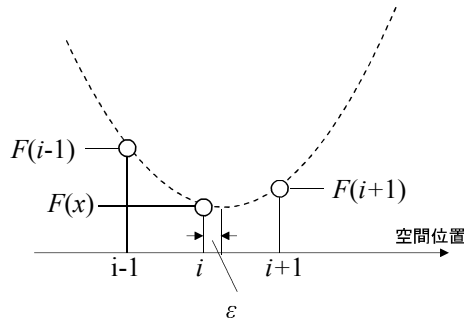


図 1-10 パラボラフィッティングによる小数精度 MV 検出 [29]

Figure 1-10 Sub-pixel MV estimation by parabolic fitting

パラボラフィッティングは、空間方向の MV の精度を高めるには有効な方法である。しかし、すなわち図 1-10 における放物線の最下点の輝度レベルの値の精度が十分に得られないため、正確な MV を検出するとの観点からは十分な精度が得られない。

Multiple-Baseline (MBL) [35]

マルチベースライン (Multiple-Baseline) は、ステレオビジョンの応用として、図 1-11 に示すように配置された複数のカメラを用いて被写体からカメラの列への距離 (法線) を求める手法である。 ζ を距離の逆数、 c を定数 (測距の応用ではカメラの焦点距離に相当)、 $F_i(\cdot)$ をカメラ i の画像信号、 Ψ をウィンドウ (ブロックマッチングではブロックに相当) としたとき、式(1.9)の評価関数を最小にする ζ を求める。

$$D(x, \zeta) = \sum_{i=1}^n \sum_{j \in \Psi} [F_0(x+j) - F_i(x + cB_i\zeta + j)]^2 \quad (1.9)$$

上式にて、カメラ P_0 とカメラ P_i の画面での被写体の画素位置の差、すなわち、カメラ P_0 の画像を基準画像、カメラ P_i の画像を参照画像としたときの MV を v_i とすると、以下となる。

$$v_i = cB_i\zeta$$

これを単一のカメラで撮影した動画像の MV 検出に適応する場合は、被写体が静止しており、カ

カメラが等速で動き、P0, P1, P2, ... の位置でそれぞれのフレームをキャプチャした場合に該当する。P0-P1間のMVを v_1 としたとき、 $i=2$ 以降のフレームのMVを

$$v_i = i v_1$$

とした上で、(1.9)式の評価関数を最小とすることに相当する。

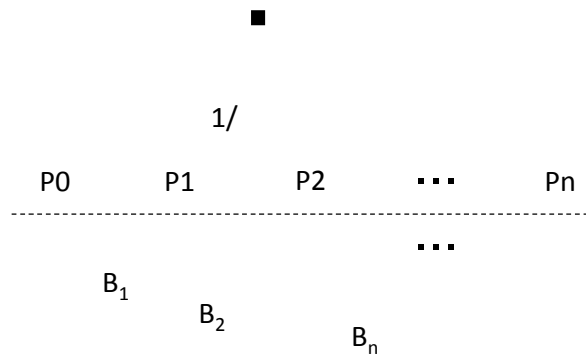


図 1-11 マルチベースラインにおけるカメラ位置

Figure 1-11 Camera position in multiple-baseline

マルチベースラインはステレオビジョンの応用においては安定した性能が得られるが、繰返しパターンに適用すると十分な精度が得られない。

1.5.2 MV の連続性を用いる方法

空間方向あるいは時間／空間方向のMVの連続性（相関）を用いることによりMV検出誤りを低減する手法が知られている[32][36][37][38]。しかし、繰返しパターンに対しては必ずしも十分な性能が得られない場合がある。また、実用上は、画素の差分とMVの連続性の重み付けをどのように設定するかが大きな課題である。さらに、MV検出の性能を高めるために画面全体での大局的なレベルでMVの連続性を最適化する場合、反復処理を行い、結果としてMV検出の処理時間（処理量）が増加する課題がある。

Horn and Schunck (HS) [36]

Horn と Schunck は動きの検出の手法の中で、輝度間の差分 D_b に隣接MV間の差分（距離） D_c の成分（連続性）を加えることによりMV検出精度を向上させている。具体的には、MVの連続性を考慮した(1.10)式の評価関数を提案している。ここで、 α は輝度差分とMV連続性の比率を設定するパラメータである。

$$D = \iint (\alpha^2 D_c^2 + D_b^2) dx dy \tag{1.10}$$

MV prediction (PMV) [37]

HS の MV 連続性の概念の実装形態の 1 つとして MV prediction がある。周囲の MV 検出済のブロックの MV から、これから MV 検出するブロック（検出ブロック）の MV を予測し、重み付けを行う方法である。最もシンプルな方法は検出ブロックの左隣の検出済の MV (図 1-12 の MV1) を予測に用いる方法である。文献[37] では、処理図 1-12 に示すように、MV を検出するブロック（図の太枠）の周囲の 3 つのブロックの MV (MV1, MV2, MV3) を用いて(1.11)式により予測 $MV(P_x, P_y)$ を求め、予測 MV と検出する MV との差の大きさによって評価に重み付けを行うことにより MV の連続性を MV 検出の評価に取り入れている。

$$\begin{aligned} P_x &= \text{median}(MV1_x, MV2_x, MV3_x) \\ P_y &= \text{median}(MV1_y, MV2_y, MV3_y) \end{aligned} \tag{1.11}$$

ここで、添字 x, y はそれぞれの MV の水平成分、垂直成分を示す。median(\cdot) は中央値をとる関数である。

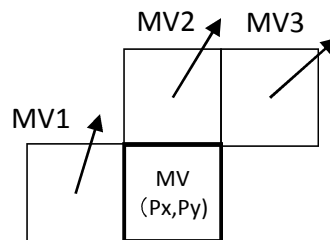


図 1-12 MV prediction に用いる MV

Figure 1-12 MVs used for MV prediction

HS では反復処理により、(1.10)式を最小化しているのに対して、PMV では検出済の MV のみを利用するため、反復処理は不要になる。簡便な手法であり、ある程度の効果が得られるため、動画像符号化の処理では採用されることが多い。

3D recursive search [38], Bartels (BT) [32]

Haan は文献[38]にて、前フレームの MV を予測に取り入れ、時間方向の連続性を考慮した 3D recursive search を提案している。Bartels は Haan と共に、これを改良した方式を提案している(BT) [32]。BT では、MV 候補として、図 1-13 に示すように空間方向候補 MV (2 つ、図にて検出ブロック C に対して、 S_1, S_2) と、時間方向候補（前フレームあるいは 1 回前の反復処理の結果）MV (図の T_1) を候補とし、候補 MV をアップデートしながら、

$$SAD(v) + \lambda \text{Pen}(v) \tag{1.12}$$

を最小化する．ここで， $SAD(\cdot)$ はブロックの画素差の絶対値和（Sum of Absolute Distance）， λ は定数， $Pen(\cdot)$ はMV連続性を表すペナルティ関数であり，BTでは P_I から P_V の5種類が提案されている．事前評価により最も性能が高かったペナルティ関数 P_{III} の場合，以下となる．

$$P_{III}(v) = \frac{1}{|N|} \sum_{x \in N} \rho(v, D(x)) \tag{1.13}$$

ここで， $|N|$ は検出ブロックを中心とした領域 N のブロック数， $D(x)$ は領域 N の各MV， $\rho(a, b)$ は2つのMVの距離（マンハッタン距離）である．

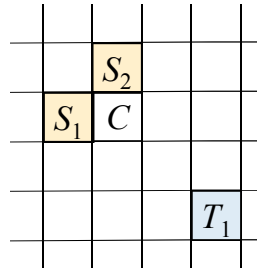


図 1-13 3D サーチの候補 MV 位置 [32]

Figure 1-13 Candidate vectors position in 3D recursive search

1.5.3 画面の大局的な情報を用いる方法

人間の場合は，画面全体を俯瞰することにより被写体やカメラの正確な動きを正しく検出することができる．本項では，この原理を適用した方法を「画面の大局的な情報を用いる方法」として示す．画面の大局的な情報を用いる方法としては，大きく「画面の一部の情報を用いる方法」と「画面全体の情報を用いる方法」に便宜的に分けている．前者はいわばボトムアップ的な方法であり，統合したブロックあるいは領域を用いてMVの検出を行う．後者は，当初より画面全体の情報を利用することを前提とした方法である．なお，画面の一部の情報を用いる方法においても，結果的に全画面の情報を用いることもあるため，画面全体の情報を用いる方法との境界は明確ではない．

(1) 画面の一部の情報を用いる方法

hierarchical search (HAS) [39]

hierarchical search では図 1-14 に示すような4分木階層画像（縮小画像）を生成し，上位の階層，例えば図では階層 $h+1$ の画像を用いてMV検出を実施し，次に階層 h の画像を用いてMV検出を行う．図に示すように，階層 $h+1$ の1つのブロックで検出されたMVを階層 h の同じ位置に存在する4つのブロックの初期MVとしてMV検出を行う．上位の階層（例えば階層 $h+1$ ）

では、ブロックサイズが画面サイズに対して相対的に大きくなるため、大局的な MV を検出することが容易になる。さらに、下位の階層の MV 探索にて上位の階層で検出した MV を継承することにより下位の階層においてもより安定した MV を得ることが可能となる。

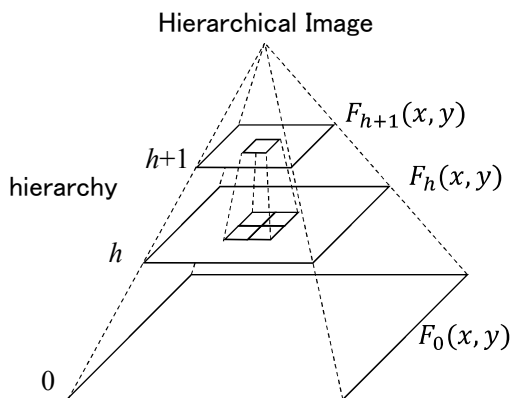


図 1-14 4分木階層画像 [39]

Figure 1-14 Quad tree hierarchical image

HAS は大局的な動きを考慮する上では有効な方法であるが、画面内が大きくいくつかの領域に分かれ、かつそれらの中の被写体が異なる動きをしている場合（例えば、ウィンドウ画像の合成や、シーンチェンジのエフェクト等）には、最上位のレイヤにおいて正しい MV が検出できず、これが下位に波及することがある。

Pyramidal implementation of the Lucas-Kanade (PLK) [40]

上述した HAS の応用例として知られている MV 検出方法が PLK[40] である。PLK はベースとなっている勾配法による Lucas-Kanade 探索[12]を階層的に実行することにより MV 検出の精度を向上させている。

平坦ブロック統合[41]

Yoshida[41]は「動きベクトルの信頼度」の考え方をを用いることにより、精度が低いと判定される推定ベクトルを隣接ブロックとの統合を通して信頼度を上げる方法を提案している。動きベクトルの信頼度は、検査する方向 θ に対して(1.14)式により定義される[42]。ここで C は定数、 f_x, f_y はそれぞれ画像 $F(x, y)$ のそれぞれ偏微分である。 $R(\theta)$ は平坦なブロックでは値が大きくなり、信頼性が低いとされる。直線状のエッジ部では θ がエッジと同じ方向の場合は $R(\theta)$ の値は大きく、エッジと垂直の方向では $R(\theta)$ の値は小さくなる。Yoshida の方法では、ブロックの $R(\theta)$ の最大値 ÷ 最小値が所定値以下でかつ、最大値が所定以上であるときは平坦なブロックであると判定する。平坦ブロックは周囲の平坦ブロックと統合され信頼度が向上させている。

$$R(\theta) = C \frac{\sin^2 \theta \sum_{\psi} f_x^2 + \cos^2 \theta \sum_{\psi} f_y^2 - 2 \sin \theta \cos \theta \sum_{\psi} f_x f_y}{(\sum_{\psi} f_x^2)(\sum_{\psi} f_y^2) - (\sum_{\psi} f_x f_y)^2} \quad (1.14)$$

繰返しパターン統合 [43]

坂本は文献[43]において、ステレオビジョンの応用のための2つの画像間の繰返しパターン部分の対応付けの方法を提案している。提案手法では、以下の4つの前提条件：

- ・繰返しパターン領域が3次元空間内でほぼ同一の平面上にある
- ・繰返しパターン要素の面積は3次元空間内でほぼ等しい
- ・繰返しパターンの要素の平均輝度・分散はほぼ等しい
- ・繰返しパターン自身の奥行きに対して、カメラからパターンまでの距離は十分長い

に対して、以下の5つの特徴量から構成される点が、ほぼ（超）平面上に分布することを利用して、連続した繰返しパターン部分の抽出を行っている。

5つの特徴量：

- ・要素の重心位置： (u, v)
- ・要素の面積： S
- ・要素内の画素の平均輝度： m
- ・要素内の画素の輝度分散： σ^2

こうした手法では、抽出を誤ると、その後の処理に影響が出るため、抽出の精度が課題となる。

(2) 画面全体の情報を用いる方法

Global motion vector (GMV) [44] [45]

GMVは画面全体のカメラモーションを表す、拡大縮小、回転、平行移動（水平、垂直）の4つのパラメータを算出する手法であり [45]、動画像符号化等に応用されている [44]。ブロックごとにMVを検出した後に画面全体のMVを平滑化し、その後に拡大縮小、回転、平行移動の各パラメータを順に検出している。文献[45]では、クラスタリングによりインライヤ（カメラモーションのみを反映したデータ）とアウトライヤ（カメラモーション以外の物体の動きなどのデータ）を分離している。GMVは単一のカメラモーションの検出には有効であるが、ウィンドウ画像の合成や、シーンチェンジのエフェクト等、複数の動きが合成された動画像に対しては適用ができない。

ステレオビジョン

ステレオビジョンにおける両眼2つの画像間の対応付け方法はMV検出に適用が可能である。その一般的な方法は以下となる。 [46]。

- (1) SIFT(Scale-Invariant Feature Transform)[28]等により画面全体の特徴点を抽出

(2)2つの画像間の特徴点の対応候補を作成

(3)これらの対応点候補から, RANSAC[47]等により, 両眼カメラの位置・姿勢により定まるエピポラ方程式を満たす対応点を探索

これらの応用として, 例えば, 金澤[48]は, エピポラ拘束の検証の前の段階で, 画素差の残差に関する確信値, 局所相関に関する確信値, 空間相関に関する確信値, 大域的整合性に関する確信値等を算出し, 確信値の高い対応から大域的な性質を推定した方法を提案している.

MDP-Flow2(MDP) [18]

ステレオビジョンの対応点検出の方法として, 文献[18]では, SIFT[23] 画面全体の大域的な対応点探索, 階層処理, グラフカットによる MV 連続性の最適化を用いることにより, その精度を高めている. 画面全体の大域的な動き検出を行うことにより, 繰返しパターンが画面全体に及ぶような場合においても MV 検出誤り率を低く抑えている. MDP 処理は以下の通りである.

1. Construct pyramids for both of the images and set the initial level $l=0$ and $u_i=0$ for all pixels.
2. Propagate u_i to level $l+1$.
3. Extended Flow Initialization.
 - 3.1 Detect and match SIFT features in level $l+1$.
 - 3.2 Perform patch matching in level $l+1$.
 - 3.3 Generate multiple flow vectors as candidates.
 - 3.4 Optimize flow using QPBO (Quadratic Pseudo-Boolean Optimization)
4. Continuous Flow Optimization.
 - 4.1 Compute the α map.
 - 4.2 Solve the TV/ l_1 energy function
5. Occlusion-aware Refinement
6. If $l \neq n-1$ where n is the total number of levels, $l=l+1$ and go to step 2

特に step 3 の部分が MV 検出に対応する部分であり, 複数候補の MV からグラフカット(QPBO, Quadratic Pseudo-Boolean Optimization)により最適な MV を算出している. MDP は図 1-15 に示すように, 2015 年 1 月時点で, Middlebury ベンチマーク [49]にて最も高い性能を示している. MDP の課題はその処理量の多さである.

Optical flow evaluation results Statistics: Average SD R2.5 R5.0 R10.0 A90 A95 A99
 Show images: below table above table in window Error type: endpoint angle interpolation normalized interpolation

Average interpolation error	avg. rank	Mequon (Hidden texture)				Schefflera (Hidden texture)				Urban (Synthetic)				Teddy (Stereo)				Backyard (High-speed camera)				Basketball (High-speed camera)			
		im0	GT	im1	untext	im0	GT	im1	untext	all	disc	im1	untext	all	disc	im1	untext	all	disc	im1	untext	all	disc	im1	untext
MDP-Flow2 [68]	10.4	2.89 ⁴	5.38 ⁴	1.19 ¹		3.47 ⁸	5.07 ⁸	1.26 ¹		3.66 ⁹	6.10 ²⁸	2.48 ⁵²		5.20 ²	7.48 ⁹	3.14 ⁶		10.2 ³	12.8 ⁵	3.61 ²⁰		6.13 ¹⁴	11.8 ¹²	2.31 ¹⁸	
DeepFlow [86]	16.9	2.98 ¹³	5.67 ¹⁸	1.22 ³¹		3.88 ³⁸	5.78 ³⁸	1.52 ⁴⁷		3.62 ⁴	5.93 ²⁴	1.34 ²		5.39 ¹⁹	7.20 ⁵	3.17 ¹⁰		11.0 ²⁷	13.9 ³⁰	3.63 ³⁰		5.91 ⁴	11.3 ³	2.29 ¹⁰	
CBF [12]	19.7	2.83 ¹	5.20 ¹	1.23 ⁴⁵		3.97 ⁴⁵	5.79 ³⁹	1.56 ⁴⁰		3.62 ⁴	5.47 ²	1.60 ¹³		5.21 ³	7.12 ²	3.29 ³⁸		10.1 ¹	12.6 ¹	3.62 ²⁵		5.97 ⁷	11.57 ⁷	2.31 ¹⁸	
SuperFlow [81]	21.6	2.94 ⁷	5.56 ¹²	1.24 ⁵⁴		3.99 ⁴⁷	5.78 ³⁸	1.67 ⁶⁵		4.06 ³¹	5.55 ⁴	1.83 ²⁸		5.52 ³⁰	7.07 ¹	3.20 ¹⁹		10.2 ³	12.7 ³	3.68 ³⁸		6.13 ¹⁴	11.8 ¹²	2.24 ¹	
Aniso. Huber-L1 [22]	22.8	2.95 ⁹	5.44 ⁶	1.24 ⁵⁴		4.42 ⁷⁰	6.27 ⁶⁸	1.67 ⁶⁵		3.79 ²⁰	5.70 ⁹	1.50 ⁷		5.31 ⁸	7.42 ⁸	3.24 ³¹		11.1 ³⁴	14.0 ⁴⁰	3.61 ²⁰		5.91 ⁴	11.4 ⁵	2.24 ¹	
NNF-Local [87]	23.4	2.92 ⁶	5.51 ¹⁰	1.19 ¹		3.30 ²	4.71 ²	1.26 ¹		3.65 ⁸	5.91 ²²	2.29 ⁴⁹		5.76 ⁵¹	8.70 ⁸⁸	3.55 ⁷⁸		10.3 ⁶	12.9 ⁸	3.60 ¹⁵		6.42 ³⁹	12.4 ³⁸	2.34 ²⁸	
CLG-TV [48]	23.5	2.94 ⁷	5.45 ⁷	1.25 ⁸³		4.26 ⁵⁹	6.17 ⁸⁸	1.60 ⁶⁵		3.68 ¹³	5.73 ¹¹	1.73 ²²		5.36 ¹³	7.41 ⁷	3.32 ⁵²		11.1 ³⁴	14.0 ⁴⁰	3.57 ⁵		5.88 ³	11.3 ³	2.26 ³	
NN-field [71]	23.5	2.98 ¹³	5.70 ¹⁹	1.20 ⁵		3.31 ³	4.73 ³	1.26 ¹		4.69 ⁸⁷	5.91 ²²	2.03 ³⁸		5.99 ⁶⁹	9.13 ⁸⁰	3.57 ⁷⁹		10.3 ⁶	12.8 ⁵	3.60 ¹⁵		6.24 ²²	12.0 ²⁰	2.31 ¹⁸	
IDDF-TV [63]	24.3	3.07 ²⁶	5.91 ²³	1.23 ⁴⁵		3.71 ²⁶	5.47 ²⁹	1.40 ²⁷		3.70 ¹⁶	6.27 ⁶⁸	1.58 ³¹		5.26 ⁵	7.60 ¹⁶	3.17 ⁴⁴		11.0 ²⁷	13.9 ³⁰	4.47 ⁶³		6.37 ²⁶	12.4 ³⁸	2.30 ¹⁸	

図 1-15 Middlebury ベンチマーク(2014/1/4)[49]

Figure 1-15 Middlebury benchmark as of Jan. 4 2015

1.5.4 インプリメンテーション

冒頭で示したような組み込み機器においては、ハードウェアあるいはデジタルシグナルプロセッサにより MV 検出をインプリメントすることが多い[9][50]. こうしたインプリメンテーションでは、画素データの並列処理、パイプライン処理により連続的にかつ計画的に処理をすることにより、低コストで高速な、効率のよい信号処理を実現している. こうした並列性、パイプライン性を活かせる効率のよい MV 検出方法の 1 つがブロックマッチング(BM)である. BM は、1 ブロックに 1 つの MV であるとの制約はあるものの、通常の動画像の場合、ブロック程度の少領域では、そこに含まれる画素は同じ MV になることが多いため、BM は有効な手法として用いられている.

一方、画面内の画素データをランダムにアクセスする大域的な処理や、直前の処理の内容によって次の処理が大きく変わるアルゴリズムを用いて高性能な MV 検出処理を実現しようとする、処理の並列性、パイプライン性が低下し、処理全体の効率が低下する問題がある.

インプリメンテーションの観点からは、「参照画像生成・画素差分の評価関数を改善する方法」が最もインプリメンテーションの効率がよく、「MV の連続性を用いる方法」、「画面の大局的な情報を用いる方法」の順にアクセスするデータの範囲が広く、かつランダムに近くなるためインプリメンテーションの効率が低下する. ここで、注意する点は、インプリメンテーションの効率と処理量あるいは処理時間は必ずしも一致しない. インプリメンテーションの効率が高い方法の方が、よりハードの設計が容易になりハード自体もコンパクトに、消費電力も少なくなる傾向にあると考える. また、ソフトウェアによるインプリメンテーションの場合は、キャッシュのヒット率や、パイプラインの処理の乱れの少なさなどの処理効率が高くなる方向になり、より低い周波数のプロセッサで処理を行うことができる可能性が高くなる. たとえば、先に示した MDP [18] の場合、ソフトウェアの処理時間は他の方法に対して 100 倍以上多くなっている (第 2 章, 第 3 章参照)

第2章

複数フレームを用いた動きベクトル検出

2.1 はじめに

正確なMVを検出するために、これまで、小数精度のMVを用いる方法や、空間方向のMVの連続性（相関）を利用する方法、時間／空間双方の相関を利用する方法などが開発されてきたが、特に、繰返しパターンに対しては十分な性能が得られていない。一方、ステレオビジョンの応用を目的とした2つの画像間の対応点検出において、画像の大局的な情報を用い、かつ、MVの連続性、被写体の形状、画素値の変化等の最適化を図る手法が開発され、高い性能が得られている。この技術はMV検出にも適用が可能であるが、処理量が多い点が課題である。このように、少ない処理量を実現することと、正確なMVを検出することが両立できないことが課題である。

本章では、ブロックマッチング（BM）をベースに、BM誤差の評価関数を改良することにより、少ない処理量にて正確なMVを検出する方法を提案する。具体的には、BM誤差の評価関数に複数フレームを用いることにより、繰返しパターン画像のMV検出誤りを低減させる方法を検討する。

提案する評価関数は、MV連続性など、より高度な方式と組み合わせることによりさらにMV検出誤り率を低減することが可能となる。実際に本章における提案手法は、文献[37]（PMV）や文献[32]（BT）など、MV連続性の評価に利用する方法と組み合わせることにより、MV検出誤りをさらに低減している。

本章の構成は以下である。第2.2節では、ブロックマッチング誤差を解析し、MV検出の誤りの原因がMVのサンプリングによる誤差に起因するBM誤差DVSE (block matching Distortion caused by motion Vector Sampling Error)であることを示す。第2.3節では、複数のフレームを用いることによりDVSEの影響を小さくし、MV検出誤りを低減する方法を提案する。第2.4節ではシミュレーションにより提案方法の妥当性の確認をした後、他方式とMV検出誤り率の比較を行い、提案方法の効果を示す。

2.2 ブロックマッチングの誤差解析

まず、MV検出誤りが発生するメカニズムを、従来手法である、MSE (mean square error)によるフルサーチ（FS）を用いて、繰返しパターン画像のMVを検出する場合を例にとり解析する。

あるシーケンスにおけるN番目のフレーム（基準画像）の、座標画像 (x, y) の画像信号を $F_N(x, y)$ としたとき、N番目のフレームの小領域 Ψ からM番目のフレームに対するMVは、式(2.1)のように、BM誤差 $D_\Psi(V_x, V_y)$ を最小とするベクトル (V_x, V_y) である。ここで、 Ω はMVの探索範囲である。MSEを評価関数とすると $D_\Psi(V_x, V_y)$ は式(2.2)のようになる。

$$(V_x, V_y) = \arg \min_{(V_x, V_y) \in \Omega} (D_\Psi(V_x, V_y)) \quad (2.1)$$

$$D_\Psi(V_x, V_y) = E_\Psi \left\{ [F_M(x + V_x, y + V_y) - F_N(x, y)]^2 \right\} \quad (2.2)$$

ここで、 $E_\Psi \left\{ [F_M(x + V_x, y + V_y) - F_N(x, y)]^2 \right\}$ は $(x, y) \in \Psi$ における信号 $p(x, y)$ の平均電力である。なお、小領域 Ψ は、例えば 8×8 画素のブロックなどに相当する。

以下、繰返しパターン画像全体が一方向に並進する例を考える。まず、画像信号をサンプリングされていない連続系のデータとして検討を行い、その後サンプリングされた離散系の検討を行うことにより、サンプリングがMV検出に影響を与えていることを示す。

説明を簡単にするため、繰返しパターンが単位時間（フレーム間隔）に水平方向に実数 V_0 動く場合（垂直方向の動きは0）を検討する。 x を実数として、画像信号を、 y 成分を省略して $F_N(x)$ と表記する。信号 $F_N(x)$ にてN番目とM番目（ $N \neq M$ ）の2つのフレーム間の関係を以下のように定義する。

$$\delta_T(x) = F_M(x + V_0) - F_N(x) \quad (2.3)$$

$\delta_T(x)$ は、N番目のフレームの画像信号と、動きベクトル V_0 （1次元で説明しているため V_0 はスカラーである）により補正されたM番目のフレームの画像信号の差分である。動画像符号化の場合にはMVによる予測誤差に相当する。前提条件ではフレーム間の時間変動はほとんど存在しないことから、 $\delta_T(x)$ は0に近い値となる。

一方、画像内においては、繰返しパターンの水平方向の繰返し周期（距離）を実数 L として、

$$\delta_L(x) = F_N(x + L) - F_N(x) \quad (2.4)$$

と定義する。 $\delta_L(x)$ は座標 x における、水平距離 L 離れた画像信号との差成分であり、画像に水平方向の周期性があることから、繰返し部分では $\delta_L(x)$ は0に近い値となる。

図2-1のように、正確な動き量 V_0 に対して、正しい動きベクトル $V_1 (=V_0)$ と、誤った動きベクトル $V_2 (=V_1+L)$ との2つのMVについて考える。

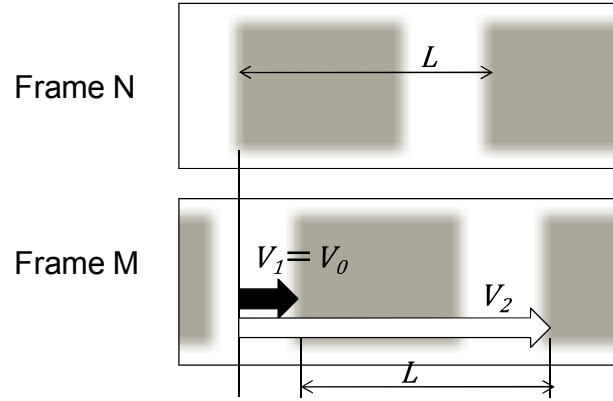


図2-1 周期画像における動きベクトル(黒矢印: 正確なベクトル, 白矢印: 誤ベクトル)
 Figure 2-1 Motion vectors of a periodic texture sequence
 (Black arrow: ture vector, white arrow: false vector)

小領域 Ψ における, 動きベクトル V_1 に対するBM誤差 $D_\Psi(V_1)$ は, 式(2.5)で表される.

$$D_\Psi(V_1) = E_\Psi\{[F_M(x + V_1) - F_N(x)]^2\} = E_\Psi\{\delta_T(x)^2\} \quad (2.5)$$

一方, 動きベクトル V_2 に対するBM誤差は, 式(2.3) , (2.4)を適用することにより, 次式となる.

$$\begin{aligned} D_\Psi(V_2) &= E_\Psi\{[F_M(x + V_2) - F_N(x)]^2\} \\ &= E_\Psi\{[\delta_T(x + L) + \delta_L(x)]^2\} \end{aligned} \quad (2.6)$$

式(2.5), (2.6)にて, 対象画像の時間方向の変化が少ないとの前提から, $\delta_T(x) \cong 0$ とすると,

$$D_\Psi(V_1) = E_\Psi\{\delta_T(x)^2\} \cong 0 \quad (2.5)'$$

$$D_\Psi(V_2) = E_\Psi\{[\delta_T(x + L) + \delta_L(x)]^2\} \cong E_\Psi\{[\delta_L(x)]^2\} \quad (2.6)'$$

となり, 画像内の繰返しパターンに少しでも変化がある, すなわち誤差 $|\delta_L(x)|$ が微小ながらも0以外の値をとり $|\delta_L(x)| \gg |\delta_T(x)| \cong 0$ であれば, $D_\Psi(V_1) < D_\Psi(V_2)$ となるため, BM誤差の小さい動きベクトル V_1 (正しいMV) が検出される.

上記の検討は, サンプリングされていない連続系での検討であり, MVの大きさを実数とした場合である. 実際には, MVの大きさはサンプリングされ離散値となる. 画素のサンプリング間隔を X , k を整数として, $x = kX$ と表すと, 式(2.3), (2.4)はそれぞれ式(2.7), (2.8)となる.

$$\delta_T(kX) = F_M(kX + V_0) - F_N(kX) \quad (2.7)$$

$$\delta_L(kX) = F_N(kX + L) - F_N(kX) \quad (2.8)$$

ここで V_0, L は実数である.

MVの水平成分 V_0, V_1, V_2 を, $-\frac{X}{2} \leq \varepsilon_i < \frac{X}{2}$ の範囲の実数 ε_i と整数 v_i を用いて,

$V_0 = V_1 = v_1X + \varepsilon_1$, $V_2 = v_2X + \varepsilon_2$ と定義する. なお, v_1, v_2 は動画像符号化処理におけるMV (単位は"画素"である) に相当する. このとき, 動きベクトル v_1 に対するBM誤差は

$$\begin{aligned} D_\Psi(v_1X) &= E_\Psi\{[F_M(kX + v_1X) - F_N(kX)]^2\} \\ &= E_\Psi\left\{[\Delta_{-\varepsilon_1}F_N(kX) + \delta_T(kX - \varepsilon_1)]^2\right\} \end{aligned} \quad (2.9)$$

となる. ここで

$$\Delta_a F_N(x) = F_N(x + a) - F_N(x) \quad (2.10)$$

である. なお, a の絶対値が小さいときには $\Delta_a F_N(x) \cong a \frac{d}{dx} F_N(x)$ であり, $\Delta_a F_N(x)$ は a と原画の輝度値の微係数に比例する.

式(2.9)と式(2.5)を比較すると, サンプリングされた系においてはBM誤差に, 元の画像の微係数に依存する成分が含まれている.

同様に, 動きベクトル v_2 に対するBM誤差は

$$\begin{aligned} D_\Psi(v_2X) &= E\{[F_M(kX + v_2X) - F_N(kX)]^2\} \\ &= E_\Psi\left\{[F_N(kX + v_2X - (v_1X + \varepsilon_1)) + \delta_T(kX + v_2X - V_0) - F_N(kX)]^2\right\} \\ &= E_\Psi\left\{[\Delta_{-\varepsilon_2}F_N(kX) + \delta_L(kX - \varepsilon_2) + \delta_T(kX + v_2X - V_0)]^2\right\} \end{aligned} \quad (2.11)$$

と表すことができる. 式(2.6)と式(2.11)の関係も, 式(2.5)と式(2.9)との関係同様に, サンプリングされた系においてはBM誤差に, 元の画像の微係数を含む成分が加わっている.

MV検出においては式(2.9)と式(2.11)のそれぞれのBM誤差が計算され, BM誤差が小さくなる

MVが選ばれる．式(2.9)と式(2.11)の大小を検討するために両式を比較する．

例えば，FSにおいてMV検出しやすい例として，時間方向の変化がなく，空間方向へはわずかな輝度変化（グラデーション等）がある例を考えると，以下となる．

$$\delta_T(x) = 0, \quad \delta_L(x) = \delta_0 \neq 0$$

ここで， δ_0 は0に近い定数である．これらの条件を式(2.9)，(2.11)に代入するとそれぞれ以下が得られる．

$$D_\Psi(v_1X) = E_\Psi \left\{ \left[\Delta_{-\varepsilon_1} F_N(kX) \right]^2 \right\} \quad (2.12)$$

$$D_\Psi(v_2X) = E_\Psi \left\{ \left[\Delta_{-\varepsilon_2} F_N(kX) + \delta_0 \right]^2 \right\} \quad (2.13)$$

例えば，小領域 Ψ が画像のエッジ部を含み， $F_N(x)$ が図2-2のような波形の場合には，

$-\frac{X}{2} \leq \varepsilon_1, \varepsilon_2 < \frac{X}{2}$ であるため， $\Delta_{-\varepsilon_1} F_N(kX), \Delta_{-\varepsilon_2} F_N(kX)$ の値は

$$F^- - F_N(kX) \leq \Delta_{-\varepsilon_i} F_N(kX) < F^+ - F_N(kX)$$

となる($i=1,2$)． $\Delta_{-\varepsilon_1} F_N(kX), \Delta_{-\varepsilon_2} F_N(kX)$ の値はエッジの振幅が大きく，傾きが急峻なほど，取り得る値の絶対値が大きくなる．この結果，式(2.13)では δ_0 の項が相対的に小さくなり， $\Delta_{-\varepsilon_1} F_N(kX)$ と $\Delta_{-\varepsilon_2} F_N(kX)$ の絶対値の大小，すなわち， $\varepsilon_1, \varepsilon_2$ の大小が，動きベクトル v_1, v_2 の選択判定に影響することになる．

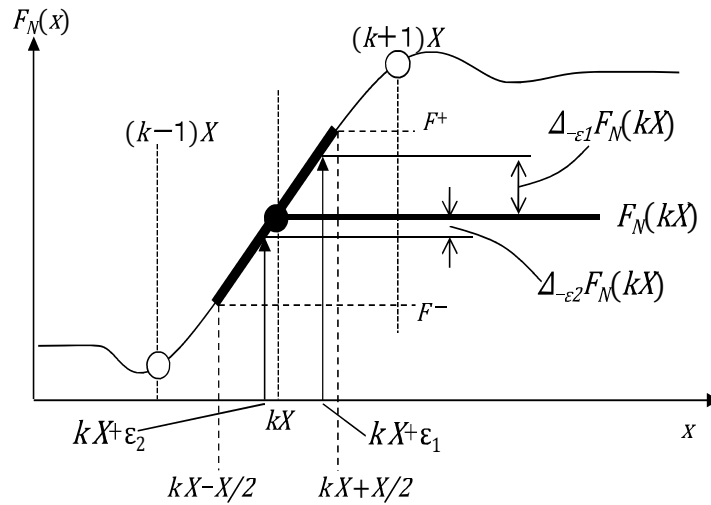


図2-2 画像エッジ部のサンプリング誤差
 Figure 2-2 Sampling error around an edge

図2-3に2つのMVの判定例を示す. 図の右のブロック (ブロックサイズ 8×8 , エッジ部では $F_N((k+1)X) - F_N(kX)$ の最大が64) が, $\delta_0=6$ の特性で水平方向に繰返し現れる場合に, v_1, v_2 のいずれが選ばれるかを, $\varepsilon_1, \varepsilon_2$ をパラメータにして示している. $\varepsilon_1, \varepsilon_2$ が, 図のグレーの部分の場合, 動きベクトル v_1 が選ばれ, (正解), 白の部分の場合, 動きベクトル v_2 が選ばれる (不正解). 図の例では, $\varepsilon_1, \varepsilon_2$ の分布が均一とすると, 白の部分の面積比, すなわち約1/3の確率で不正解の動きベクトルが選ばれることになる.

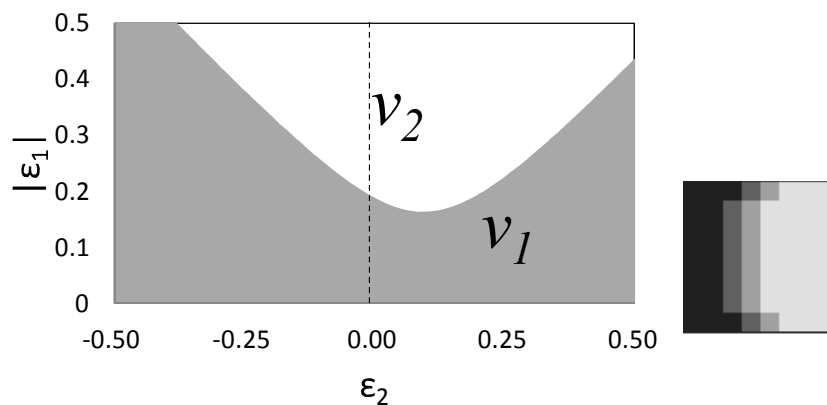


図2-3 動きベクトルの判定例(左)とサンプルブロック(右)
 Figure 2-3 An example of decision of motion vector (left) and sample image (right)

以上のように、離散系、すなわち、通常のデジタル画像においては、式(2.12)、(2.13)に示した、MVのサンプリングによる誤差に起因するBM誤差のために、MV検出誤りが発生する。このサンプリング誤差をDVSE (block matching Distortion caused by motion Vector Sampling Error)と呼ぶ。サンプリング誤差が ε のときの、小領域 Ψ に対するDVSEの値 $DVSE(\Psi, \varepsilon)$ は式(2.14)となる。

$$DVSE(\Psi, \varepsilon) = E_{\Psi}\{[\Delta_{-\varepsilon}F_N(kX)]^2\} \quad (2.14)$$

なお、上記のMVのサンプリング誤差 ε はMVの小数成分であるため正確な値を求めることはできないが、小数精度のMV検出を行うことにより概算値は求めることができる。しかし、後述の結果に示すように繰返しパターンの多い画像においては、小数精度のMV検出を行っても、十分なMV検出誤り低減ができていない。

上記のDVSEの影響により、画面内に類似したパターンが2か所存在すればMV検出誤りが発生する可能性がある。特に繰返しパターン画像においては、誤ったMVの候補が多数存在するため、正確なMVよりもBM誤差が小さいMVが出現する確率が高くなり、MV検出誤りが多数発生する。例えば、図2-3の例では、類似した個所が2か所であれば、正しいMVを検出できる確率は約2/3であった。類似した個所がNか所の場合は $(2/3)^N$ 、例えばN=10であれば、正しいMVを検出できる確率（正解率）は1.7%に低下することになる。

2.3 複数のフレームを用いたMV検出方法の提案

2章で示したように従来のMSEを用いたFSではDVSEによりMV検出誤りが発生する。本章では、複数のフレームを用いることによりDVSEの影響を小さくし、MV検出誤りを低減する方法を提案する。

2.3.1 提案法の評価関数

今、N番目のフレームの $x=kX$ の画素に注目し、整数*i*を用いてN+i番目のフレームに着目する。N+1番目のフレームのMVを v_1X とし、フレーム間隔に比例して逡倍された動きベクトル $i v_1X$ により示される画素、すなわち $F_{N+i}(kX + i v_1X)$ は図2-4の●となる。これらを、図2-4を、エッジ位置（例えば、エッジの中心）を基準に重ね合わせると図2-5となる。例えばフレームN+1の画素の位置がちょうど○の位置、すなわち $\varepsilon_1 = 0$ であればDVSEの影響なくMVを検出できることになる。しかし、実際には $\varepsilon_1 \neq 0$ となり、図中*d*で表されるDVSEの影響が生じてしまい、前章で示したように、正しいMVが得られなくなる。そこで、提案方法では、 $F_{N+i}(kX + i v_1X)$ ($i \neq 0$) を用いて式(2.15)のフィルタにより $F_{N+i}(kX)$ の位置○印の画素の予測値 $\hat{F}_N(kX)$ を生成することにより、DVSEの影響を低減することを考える。

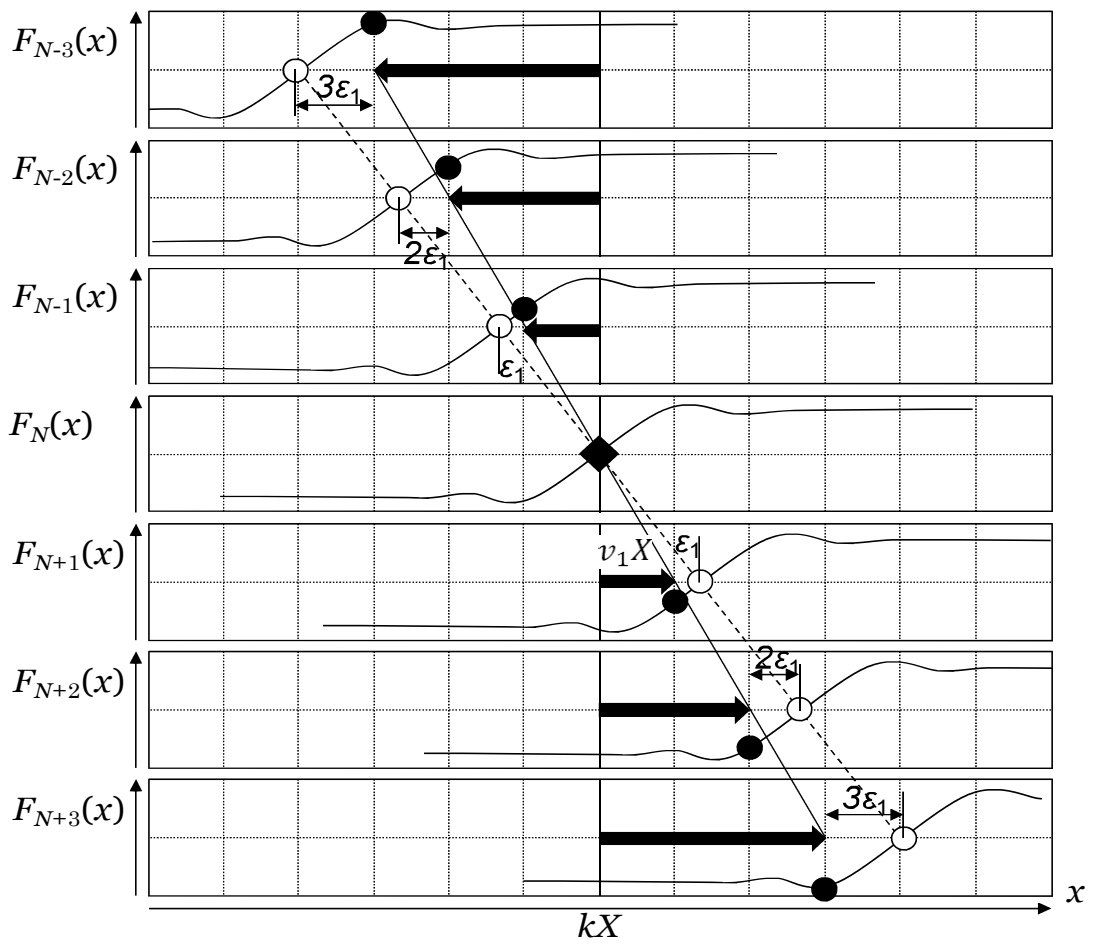


図2-4 $F_{N+i}(kX + i v_1 X)$ の画素の位置(●)
 Figure 2-4 Position of pixels $F_{N+i}(kX + i v_1 X)$

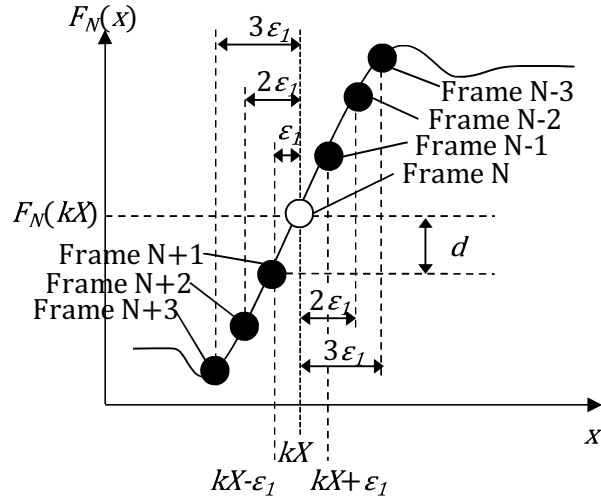


図2-5 フレーム間距離に比例した動きベクトルにより示される画素の位置

Figure 2-5. Pixels indicated by motion vectors which are proportional to frame distance.

$$\hat{F}_N(kX) = \sum_{i \in \mathbb{Z}, i \neq 0} (\alpha_i F_{N+i}(kX + i v_1 X)) \quad (2.15)$$

ここで、 α_i は以下の条件を満たす。

$$\sum_{i \in \mathbb{Z}, i \neq 0} \alpha_i = 1$$

式(2.15)により得られる予測値と基準画像の画素値 $F_N(kX)$ との差分電力を評価関数すると、BM誤差 $\hat{D}_\Psi(v_1 X)$ は以下となる。

$$\begin{aligned} \hat{D}_\Psi(v_1 X) &= E_\Psi \left\{ \left[\left(\sum_{i \in \mathbb{Z}, i \neq 0} (\alpha_i F_{N+i}(kX + i v_1 X)) \right) - F_N(kX) \right]^2 \right\} \\ &= E_\Psi \left\{ \left[\sum_{i \in \mathbb{Z}, i \neq 0} (\alpha_i (F_N(kX - i \varepsilon_1) - F_N(kX))) + DT \right]^2 \right\} \end{aligned}$$

$$= E_{\Psi} \left\{ \left[\sum_{i \in \mathbb{Z}, i \neq 0} (\alpha_i (\Delta_{-i\varepsilon_1} F_N(kX))) + DT \right]^2 \right\} \quad (2.16)$$

ここで、 DT は、時間軸方向の誤差の成分

$$DT = \sum_{i \in \mathbb{Z}, i \neq 0} (\alpha_i \delta_T^i(kX - i\varepsilon_1))$$

$\delta_T^i(\cdot)$ は式(2.3)を i 番目のフレームに拡張したもので式(2.17)となる。

$$\delta_T^i(x) = F_{N+i}(x + iV_0) - F_N(x) \quad (2.17)$$

式(2.16)にて、 $\sum_{i \in \mathbb{Z}, i \neq 0} (\alpha_i (\Delta_{-i\varepsilon_1} F_N(kX)))$ が、MV検出誤りを引き起こす成分である、

$$\begin{aligned} & \sum_{i \in \mathbb{Z}, i \neq 0} (\alpha_i (\Delta_{-i\varepsilon_1} F_N(kX))) \\ &= \sum_{i \in \mathbb{Z}, i \neq 0} (\alpha_i (F_N(kX - i\varepsilon_1) - F_N(kX))) \\ &= \sum_{i \in \mathbb{Z}, i \neq 0} (\alpha_i F_N(kX - i\varepsilon_1)) - \sum_{i \in \mathbb{Z}, i \neq 0} (\alpha_i F_N(kX)) \\ &= \sum_{i \in \mathbb{Z}, i \neq 0} (\alpha_i (F_N(kX - i\varepsilon_1))) - F_N(kX) \sum_{i \in \mathbb{Z}, i \neq 0} (\alpha_i) \\ &= \sum_{i \in \mathbb{Z}, i \neq 0} (\alpha_i (F_N(kX - i\varepsilon_1))) - F_N(kX) \end{aligned}$$

となるため、

$$\sum_{i \in \mathbb{Z}, i \neq 0} (\alpha_i F_N(kX - i\varepsilon_1)) \approx F_N(kX) \quad (2.18)$$

である適切なフィルタ係数 α_i を選ぶことができれば、式(2-16)のDVSEの原因となる項 $\sum_{i \in \mathbb{Z}, i \neq 0} (\alpha_i (\Delta_{-i\varepsilon_1} F_N(kX)))$ を0に近づけ、BM誤差におけるDVSE成分を低減することができる。

以上のように、1つのMV候補に対して、基準画像の前後のフレームを用いて式(2.15)により予測値 $\hat{F}_N(kX)$ を生成し、この予測値とのMSEが最小となるMV候補をこのブロックのMVとすることによりMV検出誤りを低減することが可能となる。式(2.15)にて予測した画素値を用いた評価関数をTemporally-Predicted Difference (TPD)と呼ぶ。また、2次元に拡張したTPDによる

評価関数を式(2.19)に示す.

$$\hat{D}_\psi(V_x, V_y) = E_\psi \left\{ \left[\left(\sum_{i \in \mathbb{Z}, i \neq 0} (\alpha_i F_{N+i}(x + iV_x, y + iV_y)) \right) - F_N(x, y) \right]^2 \right\} \quad (2.19)$$

2.3.2 フィルタ係数の決定

フィルタ係数 α_i の生成方法は、入力画像から最適化する方法と、周波数特性的に最適化する方法がある. 本研究では、フィルタ係数を固定値として実装を容易にすることを優先し、周波数特性が最適となるフィルタ係数 α_i を求めた. 図2-5および式(2.18)から、フィルタ係数 α_i は式(2.18)の左辺が補間フィルタとして動作するように設定すればよいことがわかる. ここで、フィルタ係数 α_i を求めるにあたり、2つの異なるアプローチにて検討を行う. 1つは i の絶対値が奇数の画素のみを用いたサンプリング間隔 $2\varepsilon_1$ の信号に対する補間フィルタとしての検討、そしてもう一つはサンプリング間隔 ε_1 の全画素を用いるフィルタとしての検討である. 前者は、一般的な1:2の補間フィルタであるため α_i は容易に求めることができる、本研究では実装の容易性も考慮して、表2-1に示す、タップ数2,4,6の3種類のフィルタを検討した. また、それぞれの評価関数をTPD-2T, TPD-4T, TPD-6Tと表記する.

一方、全画素を用いるフィルタでは、図2.5の●印のサンプリング間隔 ε_1 が、画素のサンプリング間隔 X の1/2以下であることから、画像のサンプリング周波数を $f_s (=1/X)$ としたとき、 $f_s/4$ 以下の周波数における透過特性が最適とする係数を求めればよい. 用いるフレームは $|i| \leq 3$ とし、フィルタ係数を $a_i/64$ (a_i は整数)、 $a_0 = -64$ ($\alpha_0 = -1.0$)としたとき、 $f_s/4$ 以下の周波数における利得が最小となる a_i を求めた. このフィルタ係数を用いたときの評価関数をTPD-S6T (Short 6 tap) と表記し、求めたフィルタ係数を表2-1の最下行に示す.

表2-1 提案方法とそのフィルタ係数

Table 2-1 Filter coefficients of proposed methods

method	filter coefficients				
	α_1, α_{-1}	α_2, α_{-2}	α_3, α_{-3}	α_4, α_{-4}	α_5, α_{-5}
TPD-2T	0.50				
TPD-4T	0.59	0	-0.09		
TPD-6T	0.63	0	-0.17	0	0.05
TPD-S6T	0.75	-0.3	0.05		

なお、表2-1の係数は全て固定値である. $\alpha_0 = -1.0$ とした時の各評価関数のフィルタレスポンスを図2-6に示す. 低周波域にて利得が0に近いほどよい特性となる. 図2-6から、TPD-

2T/4T/6Tの中ではTPD-6Tの特性が優れ、特に $f_s/4$ 以下の周波数の領域ではTPD-S6TはTPD-6Tと同等以上の特性が得られることがわかる。なお、従来の1フレームのみを用いた場合の評価関数をMSEと表記する。MSEを、これらフィルタ係数を用いて表現すると、 $\alpha_1=1$, $\alpha_i=0$ ($i \neq 1$)となる。

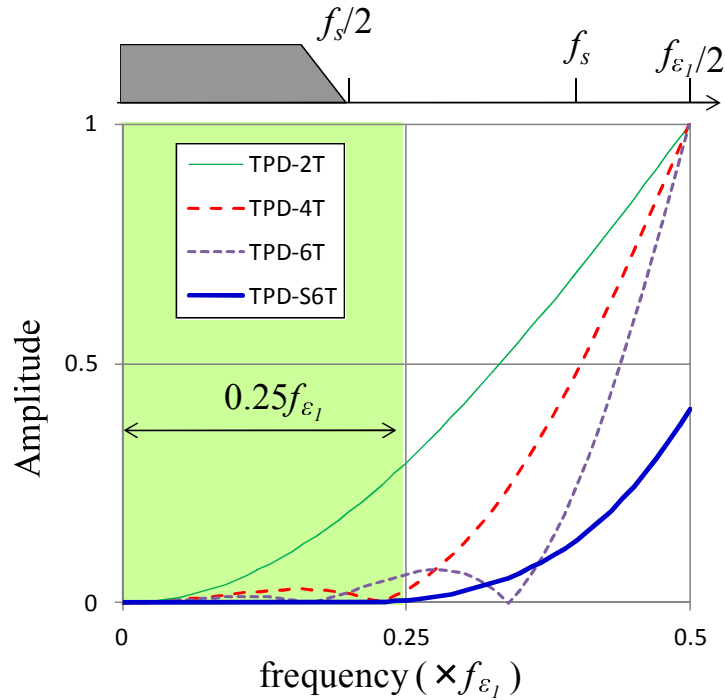


図2-6 提案方法のフィルタレスポンス

Figure 2-6 Filter response of proposed evaluation function

2.3.3 時間方向適応処理

前節にて示した提案方法は複数フレームを用いるため、動きが変動するシーケンスや、複数の動きが含まれるシーケンスにおいてMV検出誤りが増加することが懸念される。この対処法として式(2.20)の適応処理を行う。

$$\hat{D}_\psi(V_x, V_y) = \min \left(\hat{D}_\psi^L(V_x, V_y), w \cdot \hat{D}_\psi^S(V_x, V_y) \right) \quad (2.20)$$

ここで、 $\hat{D}_\psi^L(V_x, V_y)$ および $\hat{D}_\psi^S(V_x, V_y)$ はタップ数の異なる評価関数、 w は評価関数の重みである。本検討では、 $\hat{D}_\psi^L(V_x, V_y)$ はTPD-6T, TPD-S6T, TPD-4Tのいずれか、 $\hat{D}_\psi^S(V_x, V_y)$ はTPD-2T, MSEのいずれかとした。MVが時空間方向に一定であるときに、検出精度の低い $\hat{D}_\psi^S(V_x, V_y)$ が選ばれにくくするために重み $w (> 1)$ を設定している。 w の値は予備実験により2.0とした。なお、

組み合わせる2つの評価関数により、例えばTPD-6TとTPD-2Tの組み合わせであればTPD-6T/2Tのように表記する。

2.4 評価

2.4.1 シミュレーション実験条件

2.3章で提案したBMの評価関数の特性を検証するため、シミュレーションを行った。表2-2にシミュレーションの諸元を、図2-7にテストシーケンスを示す。

(a)(b)(d)(e)に用いたSkyscrapers, Harbor Scene, Church, Buildings along the Canalは文献[51]のシーケンスである。(c)は、高解像度のレースのカーテンの静止画より合成した動画シーケンスであり、6.4画素/フレームの一定速度で画面が動き、空間方向のエリアジングはない。(d)(e)は空間方向に複数の動きが混在するシーケンスである、(d)は動画像の画面内に他のシーケンスをウィンドウ表示（表示位置は画面内固定）している。(e)は2つのシーケンスの境界部が時間とともに左に移動する（ワイプ）。(f)は、繰返しパターンを含む布を水平方向にパニングしたシーケンスであり、手振れ（フレーム毎にMVが±1程度変動）を含む。(f)では、時間方向にMVの揺らぎのある場合を評価する。

表2-2 シミュレーション諸元

Table 2-2 Simulation conditions

sequences	352×240pel, 15fps, 15frames (a)Skyscrapers(skys),(b)Harbor Scene(habr), (c)lace, (d)Skyscrapers+Church(chur), (e)lace+Buildings along the Canal(canl), (f) cloth
block size	8×8pel
search area	H±15pel, V±7pel
accuracy	full pel (integer)
evaluated blocks	40x26blocks (320x208pel) in frame 6 through 11 excluding flat blocks and distortional blocks

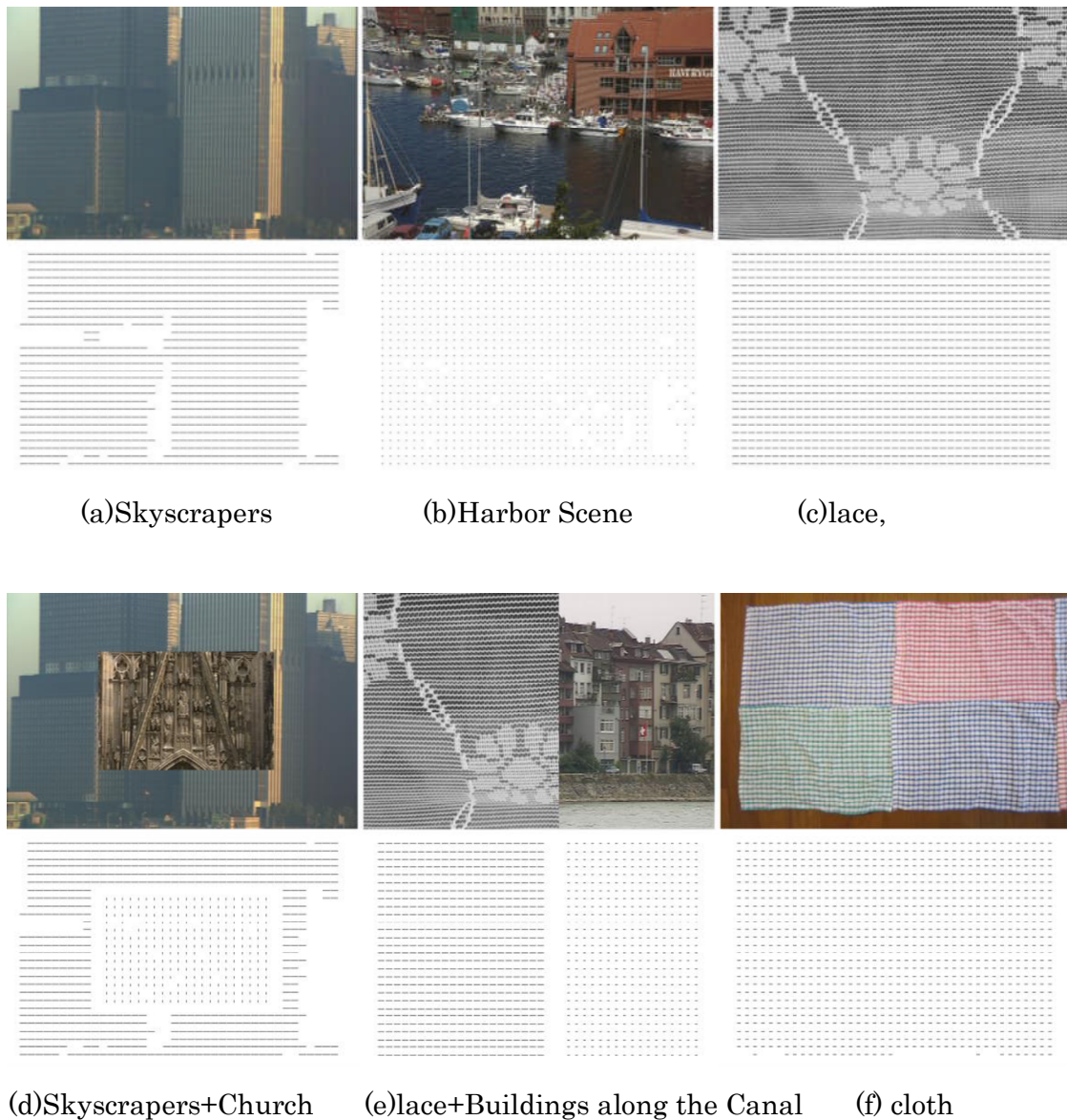


図2-7 テストシーケンスと正解ベクトル

Figure 2-7 Test sequences and grand truth (top: sequence , bottom: grand truth)

いずれのシーケンスもブロック単位で正解ベクトルを設定している。シミュレーションでは、正解ベクトルは整数精度のベクトルとし、(c)以外のシーケンスは、フルサーチにより検出した動きベクトルの中から、実際の画像の動きに合致したベクトルを手動で抽出して作成している。(c)はシーケンス作成時に設定した動きベクトルから正解ベクトルを設定している。

繰返しパターン画像においては、局所解MVは、例えば、図1-3のように、繰返しパターンの分布に対応して、離散的に分布する。本研究では、これら局所解MVの中から正しいMV(図1-3では●印)を選択できるか否かを、MV検出誤り率(100%−正解率)にて評価する。すなわち、

MV検出誤り率は $(1 - n_{\text{correct}}/n_{\text{eval}})$ にて定義する. ここで n_{correct} はMVが正解であったブロックの個数であり, 検出したMVと正解MVのユークリッド距離が th_c 以下のときを正解としている (本実験では指定のない限り $th_c=1$). n_{eval} は評価対象のブロック数であり, 以下のブロックを除いたブロックが評価対象である. なお, 以下にて"maxDVSE"は, DVSEの最大値の近似値であり, 正確なMVが取り得るBM誤差の最大値である (第3章参照).

評価対象外のブロック:

(1) 平坦なブロック

$$\text{maxDVSE} < th_f \times \text{number_of_pixels_per_block}, th_f = 2. \text{ (i.e., 128 for an } 8 \times 8 \text{ block),}$$

(2) 変形したブロック, 時間方向の変化の大きいブロック

$$\text{minimum distortion} > \text{maxDVSE}$$

(3) 画面の端のブロック

左右各2ブロック, 上下各1ブロック

(4) シーケンスの先頭5フレーム・末尾4フレームのブロック

なお, シミュレーションにおいては, いずれの方式も, (1)(2)(3)のブロックにおいても, MV検出処理は実行しており, MV検出誤り率の評価時にのみ計算対象から除外をしている.

提案方法のシミュレーションでは, 提案する評価関数を用いて探索エリアを整数画素単位でフルサーチにより探索する方法 (単に評価関数名で表記, 例: TPD-S6T) と, 提案する評価関数をBT[32]あるいはPMV[37]に適用する方法 (例えばTPD-S6T+BT, TPD-4T+PMVと表記) を実施する. 比較対象として, 以下の方式を採用する.

輝度差/BM誤差の改良による方式

輝度差/BM誤差の改良による方式 (MV連続性を利用しない方式) の比較対象は, MSE, half pel[33] (hpel), quarter pel[2][37] (qpel), パラボラフィッティング[29] [30] (PF), マルチベースライン[35] (MBL) とした. hpel, qpelは, それぞれの解像度の補間画像を生成したのちに全探索を行っている (探索回数はMSEのそれぞれ4倍, 16倍となる). また, PFではMSEにより得た全ての極小値に対してPFを適用し推定されるBM誤差が最小のMVを採用する. MBLに用いるフレームは, TPD-2T, TPD-S6Tと同じフレーム, すなわち基準画像の前後各1フレーム(MBL2)および前後各3フレーム(MBL6)とした.

MV連続性を利用する方式

MV連続性を利用する方式の比較対象は, PMV[37], BT[32], MDP-flow2[18]とした. BTは大域的なデータアクセスに制約を加えた方式である. 予備実験にて最も性能がよいsmoothness prior P_m を採用し, パラメータは $\lambda=400$ (提案方式と組み合わせ最高性能) と $\lambda=1000$ (BT単独での最高性能, 以下"BT1000") を使用している. MDP-flow2はMV検出誤りの低い方式であ

る[49]¹. smoothパラメータは6とし、画素単位で得られるMVをmedianフィルタによりブロック単位のMVに変換している(256個のMVの水平・垂直成分の各128番目を採用).

2.4.2 シミュレーション実験結果と考察

2.4.2.1 提案方法の特性

表2-3にシミュレーション実験結果を示す. 表2-3より, TPD-2T, 4T, 6T, S6Tは, いずれもMSEよりもMV検出誤りが大幅に低減していることがわかる. TPD-2T, 4T, 6Tの中では, 画面が一方向に動くシーケンス(a)(b)(c)ではタップ数が長い方がMV検出誤りが少ない. しかし, 動きが完全に一定であるシーケンス(c)にてTPD-6TのMV検出誤りの低減が著しい一方, 若干のMVの変動があるシーケンス(a)にてTPD-6TよりもTPD-4TのMV誤りがより低減されていることから, タップ数が長くなるほどMVの時間変動に影響されやすくなっていることがわかる. 画面内あるいは時間方向に動きが変化するシーケンス(d)(e)(f)でも, 基準画像よりフレーム間距離の長いフレームの影響により, TPD-6TのMV検出誤りが増加する傾向にある. TPD-S6Tは, 全シーケンスを通して, MV検出誤りが低く抑えられており, MV検出誤りが, シーケンス(a)~(c)では, MSEの1/5以下, シーケンス(d)~(f)では1/3~1/4となっている.

時間方向適応処理を採用した提案方法では, 画面内あるいは時間方向に動きが変化するシーケンス(d)~(f)でのMV検出誤りが低減され, 適応処理の効果が確認できる. TPD-S6T/MSEではTPD-S6Tに比べシーケンス(d)(e)ではMV検出誤りが約30%~60%低減されている. 基準画像の前後各3フレームにおいて遜倍したMVの指し示す先が異なる動きの領域であった場合などでは, TPD-S6Tによる誤差値では正しいMVが得られない. こうしたブロックでは, MSEが選ばれ, 正しいMVが検出される確率が高まる. 図2-8にシーケンス(e)におけるTPD-S6T/MSE(左)とTPD-S6T(右)のワイプの境界部付近でのMV検出例を示す. 図左の枠内のブロックが, MSEが選ばれたブロックである. これらのブロックでは, TPD-S6Tは誤ったMVを検出しているが, TPD-S6T/MSEは正しいMVを検出している.

シーケンス(f)では, 手振れの程度によりフレーム毎にMSEが選ばれる比率が大きく異なっている. TPD-S6T/MSEはMV検出誤り率のフレーム単位での変動が少なく安定している. 一方, 時間方向適応処理を行っても, シーケンス(a)~(c)でのMV検出誤りへの影響はわずかであり, 時間方向適応処理の副作用が小さいことが確認できた.

¹MDP-flow2は文献[49]のサイトにて,2015年1月29日時点で, Optical flow のInterpolation error (Average, interpolation) 評価のランキング1位である.

表2-3 シミュレーション結果(1)
Table 2-3 Simulation results(1)

Method		ME error rate (%)						
		(a)skys	(b)habr	(c)lace	(d)skys/ chur	(e)lace/ canl	(f)cloth	average
TPD	2T	14.32	2.67	52.01	9.01	36.88	13.33	21.37
	4T	11.09	1.91	19.40	8.64	20.99	15.17	12.87
	6T	14.00	1.55	1.80	11.68	13.94	19.11	10.35
	S6T	7.30	1.58	0.81	6.46	12.29	9.38	6.30
	2T/MSE	14.41	2.69	52.04	8.05	31.89	12.04	20.19
	4T/MSE	11.12	1.95	19.40	6.94	14.74	12.63	11.13
	6T/MSE	13.28	1.58	1.80	8.88	6.55	15.00	7.85
	S6T/MSE	7.35	1.64	0.81	4.46	4.59	7.62	4.41
	4T/2T	10.26	1.95	19.40	6.36	19.94	14.73	12.11
	6T/2T	10.92	1.64	1.80	6.86	11.73	17.38	8.39
	S6T/2T	6.99	1.62	0.81	4.77	9.19	8.77	5.36
FS(MSE)		44.23	8.31	70.70	24.37	43.29	39.90	38.47
hpel		35.60	8.79	51.69	16.90	32.44	27.13	28.76
qpel		20.16	8.35	44.30	9.17	28.38	3.29	18.94
PF		35.68	10.86	67.93	20.89	42.47	27.68	34.25
MBL2		35.50	5.07	61.39	20.29	42.27	35.58	33.35
MBL6		40.09	8.21	59.83	23.42	50.30	57.41	39.88

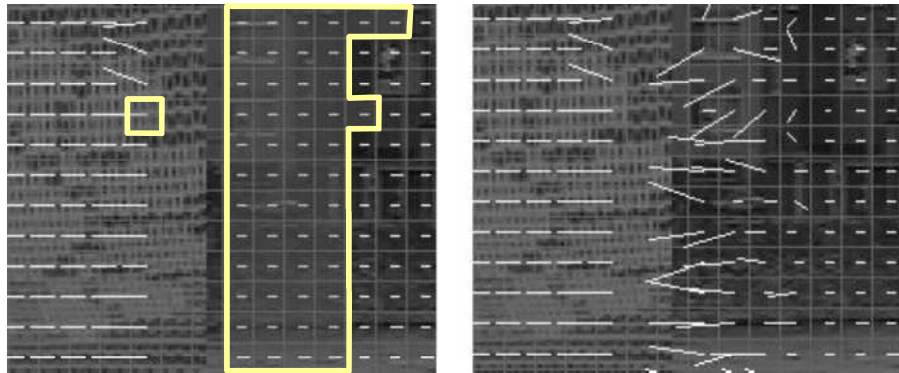


図2-8 時間方向適用処理あり(左)と適用処理なし(右)

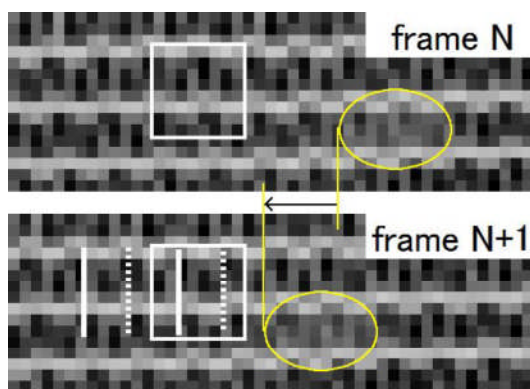
Figure 2-8 An example of temporal adaptive mode (left) and adaptive off (right) (short tap evaluate function is selected at framed blocks)

2.4.2.2 輝度差／BM誤差の改良による方式との比較

hpel, qpel, PF, MBLは概ね従来のFS(MSE)よりもMV検出誤りは低減している。しかし、提案方法との比較においては、提案方法TPD-4T, 6T, S6Tは、ほとんどの場合においてこれらの方式よりもMV検出誤りが低減されている。

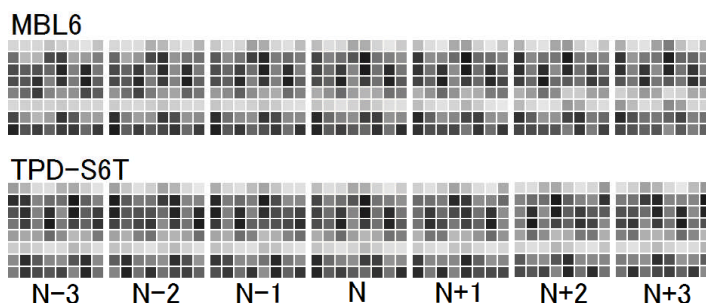
図2-9に提案方法TPD-S6TとMBL6の比較を、シーケンス(c)を例に示す。図2-9(a)上(フレームN)の中央左枠内のブロックがMVの検出対象のブロックである。図2-9(a)下(フレームN+1)を参照画像としてMVを求める例である。図2-9(a)下において、左右2本の点線で示したブロックがMBL6により検出されたMVが示すブロックの位置、左右2本の太線で示したブロックがTPD-S6Tによるブロック位置である。検出対象ブロックの位置は四角の位置であるため、それぞれのMVは、MBL6が検出したMVは水平方向へ2画素/フレーム、TPD-S6Tは6画素/フレームである。正解ベクトルは6.4画素/フレームである(図2-9(a)の楕円で示した、輝度が周囲よりも若干高い箇所動きからもわかる)ため、提案方法TPD-S6Tが検出したMVは正解、MBL6が検出したMVは誤りである。

MBL6, TPD-S6TそれぞれのMVを検出する時に用いた6フレームのブロックを図2-9(b)に示す(中央の"N"を除く6つのブロック)。表示されたブロックは、それぞれの方式により検出されたMV(フレームNからフレームN+1へのMV)をフレーム間距離に応じて逡倍したMVが指し示す位置のブロックである。中央("N")の列のブロック(MBL6, TPD-S6Tともに同じ画像)が図2-9(a)上(フレームN)の枠線内のブロックに。図2-9(b)の"N+1"のブロックが、図2-9(a)下(フレームN+1)に図示したブロック(上述した点線間あるいは太線間)にそれぞれ対応する。

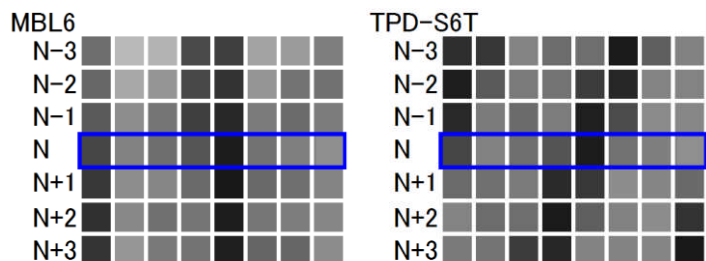


(a) template frame (N) and reference frame (N+1).

Two dotted lines: ref. block by MBL6's MV, two thick lines: ref block by TPD-S6T's MV



(b) blocks used for block matching



(c) block lines used for block matching (2nd lines of (b))

図2-9 提案方法とMBLの比較.

Figure 2-9 Comparison between proposed method (TPD-S6T) and MBL6

MVを検出する時に用いたブロックの2ライン目を、時間順に縦方向に並べたものを図2-9(c)に示す. 図2-9(c)見ると, MBL6は, どのフレームも同じ画素値となるようなブロックを選んでいる. これは, 式(2-11)あるいは(2-13)におけるMVの量子化誤差 ε_2 が,

$$-3\varepsilon_2 = -2\varepsilon_2 = -\varepsilon_2 = \varepsilon_2 = 2\varepsilon_2 = 3\varepsilon_2,$$

すなわち $|\varepsilon_2| \sim 0$ となるMVを見つけることに相当し, 結果として誤ったMVが選ばれている.

一方TPD-S6Tでは、図2-9(c)右の図のように、画像が各ブロックで少しずつ位置が変化（図にて斜め線に見える）、すなわち、各画素値が時間方向にゆるやかに変化している。TPD-S6Tでは、こうした時間方向の画素値の変化を利用してMVの量子化誤差、すなわちDVSEの影響を排除することにより、画面内のパターン間の微小な変化（例えば、式(2.12)(2.13)の δ_0 ）を検知しやすくしている。

このように、提案方法とMBLの処理は、一見類似しているが、処理の思想は大きく異なっており、MV検出との観点からは、表2-3の結果からも、提案方法が優れていることがわかる。

2.4.2.3 MV連続性を利用する方式との比較

表2-4に、提案方法とMV連続性を利用する方式（PMV, BT）の組み合わせと、MV連続性を利用する方式PMV, BT, MDP-flow2との比較を示す。

まず、提案方法内の比較では、MV連続性と組み合わせた方がMV検出誤りが低減され、pmvとの組み合わせよりもBTとの組み合わせの方が概ねMV検出誤りがより低減されている。時間方向適応処理との組み合わせでは、TPD-FSとの適応処理の方が、TPD-2Tとの適応処理よりもMV検出誤りが低減されている。

提案方法では、TPD-S6T/MSE+PMVが、PMV, TPD-S6T/MSEそれぞれの単独の方式よりもMV検出誤りが低減している。同様に、TPD-S6T/MSE+BTについても、BT, TPD-S6T/MSEそれぞれの単独の方式よりも概ねMV検出誤りが低減している。なお、BTにおけるMVの重みパラメータ λ は、BT単独の場合は $\lambda=1000$ の時の平均MV検出誤り率が最も低いが、提案手法と組み合わせた場合は $\lambda=400$ の時の平均MV検出誤り率が最も低くなっている。DVSEの影響を排除することによりBM誤差の絶対値が小さくなり、MVの連続性の重み付けが変わったと考えられる。以上のように、提案方法の評価関数を既存の方式と組み合わせることによりMV検出誤りがより低減できることが確認できる。

さらに、TPD-S6T/MSE+BTは、MDP-flow2と同等のMV検出誤り率となっており、6シーケンスのMV誤り率の単純平均では、TPD-S6T/MSE+BTのMV検出誤り率は、MDP-flow2の約1/2となっている。

表2-4 シミュレーション結果(2)
Table 2-4 Simulation results (2)

Method	ME error rate (%)						average
	(a)skys	(b)habr	(c)lace	(d)skys/ chur	(e)lace/ canl	(f)cloth	
TPD-S6T(参考)	7.30	1.58	0.81	6.46	12.29	9.38	6.30
TPD-2T+PMV	2.35	0.85	51.16	10.86	36.25	10.96	18.74
TPD-4T+PMV	2.57	0.61	17.65	10.67	19.87	13.02	10.73
TPD-6T+PMV	2.37	0.59	1.08	12.12	13.75	17.07	7.83
TPD-S6T+PMV	1.11	0.43	0.46	10.55	11.97	7.88	5.40
TPD-S6T/MSE(参考)	7.35	1.64	0.81	4.46	4.59	7.62	4.41
TPD-2T/MSE+PMV	2.35	0.89	51.20	9.03	31.21	10.12	17.47
TPD-4T/MSE+PMV	2.57	0.65	17.65	7.36	13.69	10.81	8.79
TPD-6T/MSE+PMV	1.61	0.65	1.08	5.66	6.64	13.36	4.83
TPD-S6T/MSE+PMV	1.24	0.43	0.46	6.90	4.42	6.44	3.32
TPD-2T +BT	0.00	0.00	23.50	4.00	13.98	0.00	6.91
TPD-4T +BT	0.00	0.00	0.76	4.11	5.89	0.73	1.92
TPD-6T+BT	16.91	1.60	1.70	12.67	14.94	21.12	11.49
TPD-S6T+BT	0.00	0.00	0.00	4.82	7.90	0.36	2.18
TPD-2T/MSE+BT	0.00	0.00	23.50	3.32	14.94	0.00	6.96
TPD-4T/MSE+BT	0.00	0.00	0.76	2.54	0.04	0.48	0.64
TPD-6T/MSE+BT	16.43	1.60	1.70	10.70	7.12	18.39	9.32
TPD-S6T/MSE+BT	0.00	0.00	0.00	2.72	0.18	0.20	0.52
TPD-4T +BT1000	0.00	0.00	0.00	9.36	8.81	0.00	3.03
TPD-S6T +BT1000	0.00	0.00	0.00	15.92	9.21	0.00	4.19
TPD-4T/MSE+BT1000	0.00	0.00	0.00	5.86	0.04	0.00	0.98
TPD-S6T/MSE+BT1000	0.00	0.00	0.00	11.30	0.05	0.00	1.89
PMV	21.51	5.93	69.80	13.80	42.45	38.11	31.93
BT ($\lambda = 400$)	0.00	0.00	42.05	3.60	21.58	0.00	11.21
BT1000 ($\lambda = 1000$)	0.00	0.00	16.89	5.63	14.65	0.00	6.20
MDP	0.00	1.53	1.87	0.00	2.94	0.07	1.07

2.4.2.4 インプリメンテーションに関する考察

提案方法は、BM回路の前段にFIRフィルタ処理を追加することで実現できる。フィルタの係数は固定であり、また、アクセスする画素も予め定まっているため回路の設計は比較的容易である。メモリ量やメモリバンド幅は、TPD-S6T/MSEの場合、FS(MSE)に比べ約6倍増加する。しかし、プロセスの微細化に伴い、メモリ量の制約は緩和させる方向にある。また、アクセスする領域が予め定まっているため、LSIの内部メモリにロードしてLSI内部に閉じた処理が可能となり、バンド幅増加の影響も少なくなる。

一方、処理の遅延に関しては、TPD-S6T/MSEの場合、最低3フレーム分の遅延（15frame/sのフレームレートの場合は200ms）が発生してしまう。従って、3フレーム分の遅延が許容されない低遅延アプリケーションに対しては、TPD-S6Tをベースとする方式は適用することができない。なお、TPD-2Tをベースとする方式は、従来（MSE）と同じ遅延量での処理が可能となる。

各処理の平均計算時間を表2-5、図2-10に示す。処理時間はいずれもCore2 Duo 2.27GHzによる実測値である。なお、MDP_flow2は画素単位のベクトルからブロック単位のベクトルへの変換時間を含んでいない。処理の内容や、MVの検出範囲などが異なるため単純な比較はできないが、TPD-S6T/MSE+BTの処理時間は、MDP_flow2に比べ大幅に少なくなっている。TPD-S6T/MSE+BTはBTに比べて処理時間は2倍程度かかっているが、MV検出誤り率は1/20以下になっている。

表2-5 処理時間とMV検出誤り率

Table 2-5 Processing time vs ME error rate

方式	処理時間 (msec/frm)	MV検出誤り率 (平均) (%)
TPD-S6T	543	6.30
TPD-S6T/MSE	657	4.41
TPD-S6T/MSE+BT	714	0.52
PMV	143	31.93
BT	314	11.21
BT1000	314	6.20
MDP-flow	114900	1.07

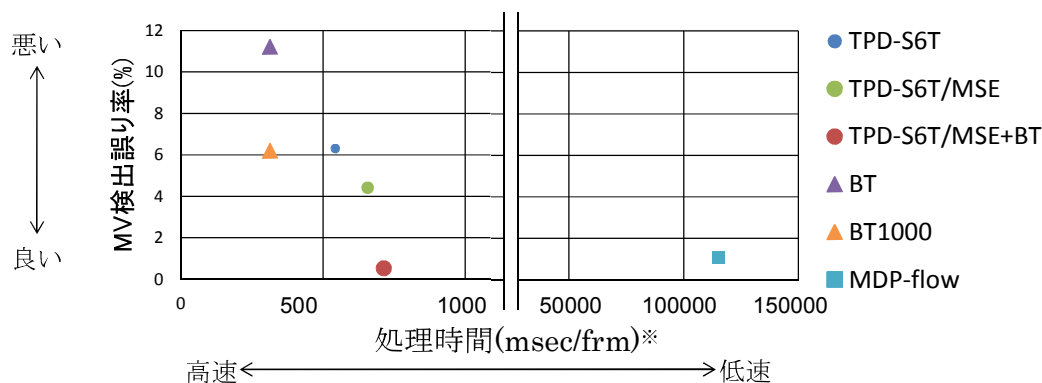


図2-10 処理時間とMV検出誤り率

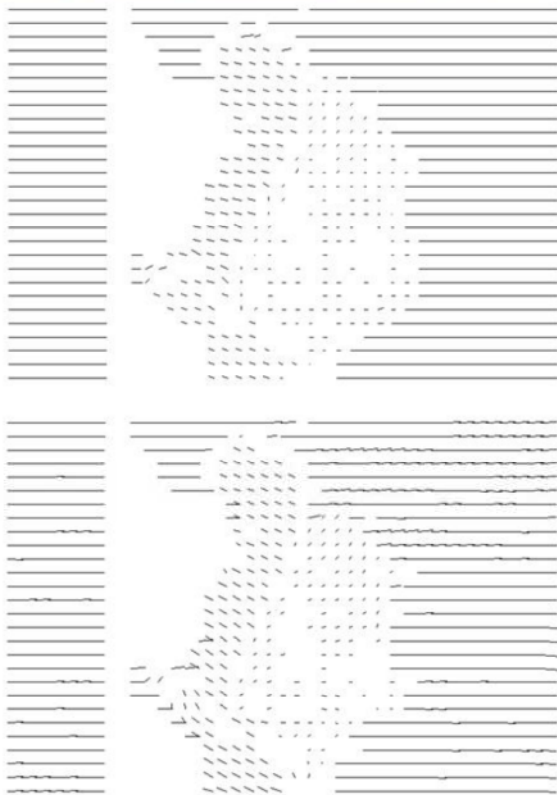
Figure 2-10 Processing time vs ME error rate

2.4.2.5 平行移動以外の動きの画像に対する考察

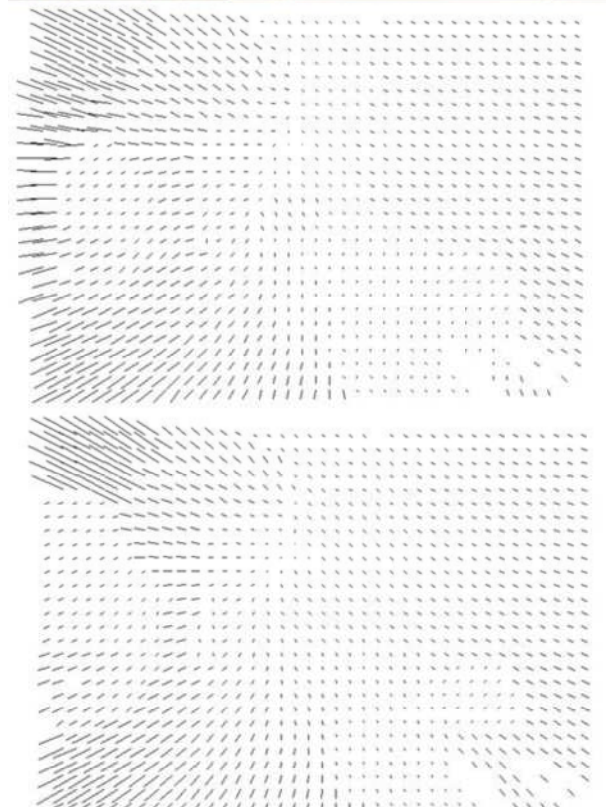
図2-11に、本方式を他の画像に適用したときのMV検出結果をMDP-flow2の結果とともに示す。本編での評価と同様に画面端ブロック、平坦なブロック、歪の大きいブロック（オクルージョン等）は除いて表示してある。

図2-11のシーケンス(g)(h)(i)は文献[51]のシーケンスである。正解ベクトルを正確に設定できないため本編の定量的な評価から除外をしている。(g)は回転、(h)はフォローパン、(i)は走行による背景の流れ、(j)は水平方向に移動するゾーンプレートのシーケンスである。

シーケンス(g)では、提案方法、MDP-flow2ともにほぼ正確なMVが検出できている。シーケンス(h)も、提案方法、MDP-flow2ともに概ね正しいMVが得られている。ただし、オクルージョン（人物右側）の周辺では、MDP-flow2がより正確なMVを検出している。一方、MDP-flow2は背景部のMVにわずかな揺らぎが見られる。シーケンス(i)でも、提案方法、MDP-flow2ともに概ね正しいMVが得られている。画面左端の柵の付近の動きは提案方法の方が正確である。シーケンス(j)では、提案方法はほぼ正しいMVを検出している（わずかに垂直方向にゆらぎがある）。一方、MDP-flow2は特に左右の水平高周波部にMV検出誤りが見られる。処理の冒頭で特徴点が見つけられなかったことが原因と思われる。



(g)



(h)

図2-11(1) 提案方法(TPD-S6T/FS+BT)とMDP_flow2の比較(上段より:原画, TPD-S6T/FS+BT, MDP-flow2

Figure 2-11(1) Comparison between proposed method (TPD-S6T/FS+BT) and MDP-flow2 (original, TPD-S6T/FS+BT,MDP-flow2 from top)

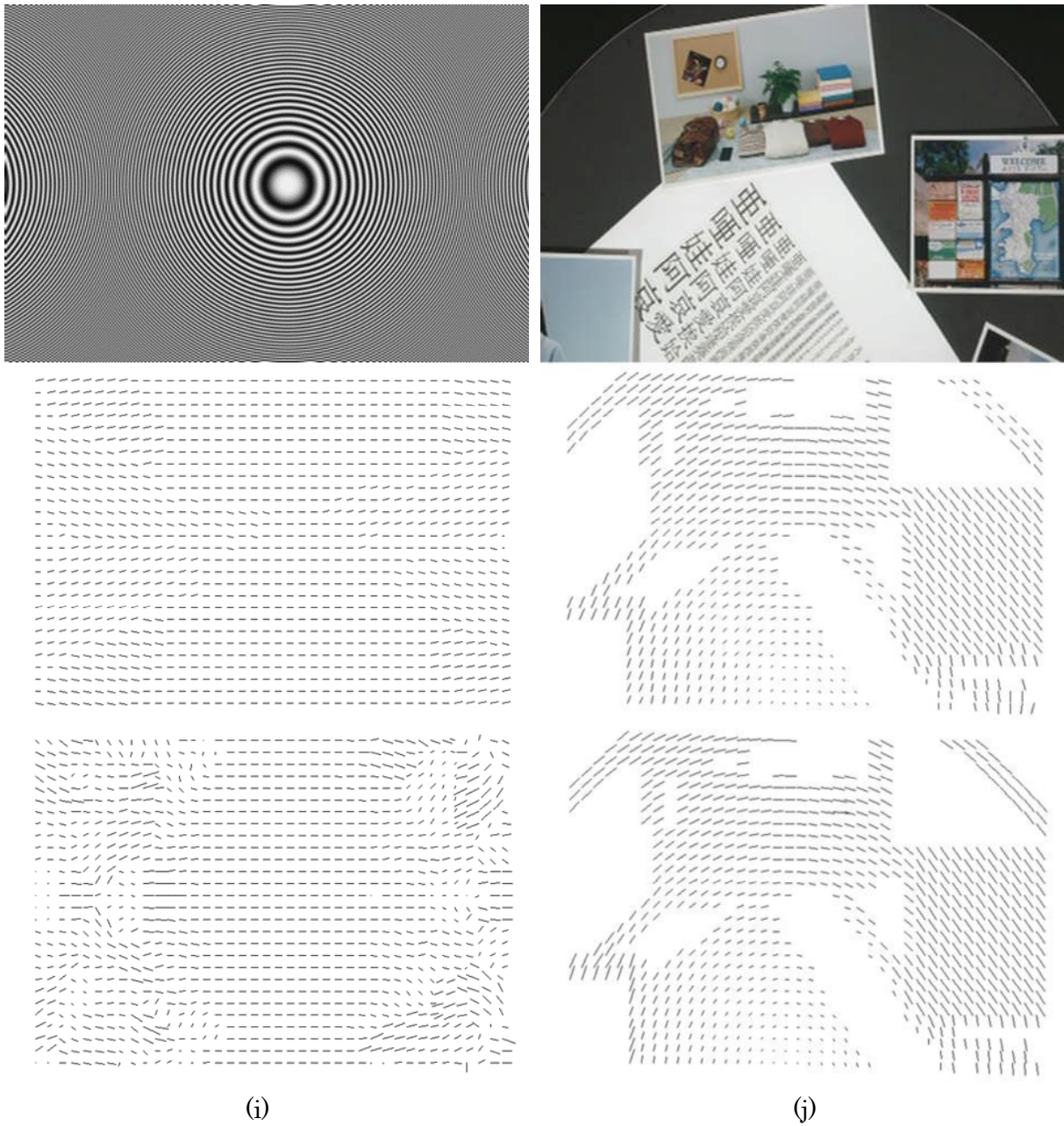


図2-11 (2) 提案方法(TPD-S6T/FS+BT)とMDP_flow2の比較(上段より:原画, TPD-S6T/FS+BT, MDP-flow2)

Figure 2-11(2) Comparison between proposed method (TPD-S6T/FS+BT) and MDP-flow2 (original, TPD-S6T/FS+BT,MDP-flow2 from top)

2.5 むすび

画面や被写体の動きのMVを正しく検出することを目的として、ブロックマッチングにおけるMV誤検出のメカニズムを解析し、その結果から、複数のフレームを用いた評価関数TPD (Temporally-Predicted Difference)を用いたMV検出方式を提案した。

まず、画像のエッジ位置とサンプリング位置の位相差に着目し、画像モデルを用いて、正確なMVが検出できない原因の分析を行い、MV検出誤りの原因がMVの量子化誤差であることを示した。1次元の画像モデルを用い、正しいMVと誤ったMVのそれぞれの場合のブロックマッチング誤差を算出し、MVの量子化誤差成分により生じるブロックマッチング誤差 (DVSE: block matching Distortion caused by motion Vector Sampling Error) がMV検出誤りの原因であることを示した。

次に、分析結果を基に、MVの量子化誤差の影響を排除するために、複数のフレームを用いたブロックマッチングの評価関数TPDおよびTPDを用いたMV検出方法を提案した。TPDは、基準画像を中心に、時間軸の前方向および後方向の双方に、基準・参照画像のフレーム間距離の逡倍に位置するフレーム、2～6フレームを用いてフィルタ処理を行うことにより、MV検出誤りの原因であるMVの量子化誤差成分を相殺する効果を持つ。さらに、提案方法を用いて、異なるフレーム数からフィルタ処理により生成した複数の参照画像を用いることにより、時間方向に不規則な変化があるシーケンスにも対応する適応処理方法を提案した。

提案した方法は、実画像を用いたシミュレーションにより効果を検証した。評価には繰り返しパターンを多く含む実画像のシーケンスを用い、予め算出した正解ベクトルを用いて、MV検出誤り率として評価を行った。TPDを用いたMV検出方法は従来の評価関数よりもMV検出誤り率を低減することができることを示した。また、TPDと既存の手法を組み合わせることにより、MV検出誤りをさらに低減でき、提案方法の1つであるTPD-S6T/MSEとBTを組み合わせることにより、最も性能の高いMV検出方式の1つであるMDP-flow2よりもMV検出誤りを低減できる。あわせて、提案方法のソフトウェアの処理時間より処理量の評価を行い、実用的な処理量であることを示した。

本研究では、平坦なブロックや、歪の大きいブロックを除いた、MVが一意に定まるブロックを対象として検討を行った。ブロックサイズを適応的に小さくすることなどにより、現在対象外としている、歪の大きいブロックのMV検出に提案方法が適用できるかを検討することが今後の課題である。また、フィルタ係数を入力画像から学習し最適化する方法も今後の課題である。

第3章

maxDVSEを用いた動きベクトル検出

3.1 はじめに

本章では、ブロックマッチング（BM）によるMV検出において、被写体あるいはカメラの動きを表す正確なMVを検出することを目的とする。特に、繰返しパターンを含む画像を処理したときの、MV検出誤りを低減することを目的とし、第2.1節のMV検出誤りの分析結果を基に、MVの量子化誤差の影響が存在することを前提とした、第2章とは別のアプローチによるMV検出方法を検討する。

第2.1節で示したように、MVのサンプリングによる誤差に起因するBM誤差DVSE (block matching Distortion caused by motion Vector Sampling Error)のために、MV検出誤りが発生する。本章ではDVSEの最大値の近似値（maxDVSE）を利用したMV検出方法を提案する。

第1章で示したように、空間方向MVの連続性（MVの差分）を用いることによりMV検出誤りを低減する手法が知られている。一方で、画素の差分とMVの差分の重み付けの設定や、画面全体での大局的なレベルでMVの連続性を最適化する反復処理の処理時間（処理量）増加が課題である。

maxDVSEを用いることにより、画素の差分とMVの連続性の重み付けの設定を不要とするとともに、MVの連続性の最適化における反復処理を行うことなしに、十分なMV検出性能を実現する。

なお、本章では、繰返しパターンを多く含む画像が並進している画像を主に対象とする。また、第2章と同様に、MVが一意に定まらないブロック、すなわち、平坦なブロックや、変形・オクルージョン・回転・複数の動き混在等が存在するブロックに対するMV検出は対象としない。こうした平坦な、あるいは変形したブロックでは、MVは、周囲の「MVが一意に定まるブロック」のMVを正しく検出できれば、それらを利用してMVを検出することができる[52]。したがって、本検討の意義は高いと考える。空間方向の精度の細かさ（サブピクセル精度のMV検出）を求めることも、第2章と同様に、本章の対象ではない。

本章の構成は以下である。第3.2節では、MV検出誤りの解析結果から、誤りの原因となるDVSEを間接的に表す指標であるmaxDVSEを算出する方法を示す。第3.3節では、正しいMVのBM誤差がmaxDVSE以下であることを利用して、フルサーチ中にBM誤差がそれぞれのブロックのmaxDVSE以下であるMVを候補ベクトルとして記録しておき、後に各ブロックのベクトルを確定してゆく方法を提案する。第3.4節ではシミュレーションにより提案方法の妥当性の確認をした後、他方式とMV検出誤り率の比較を行い、提案方法の効果を示す。

3.2 maxDVSEの算出

MV検出誤りのメカニズムは第2.1節で述べたように、式(2.12)、(2.13)に示した、MVのサンプリング誤差に起因するBM誤差(DVSE) $E_{\Psi}\{[\Delta_{-\varepsilon}F_N(kX)]^2\}$ である。時間方向に画像の変化がない場合においても、DVSEは発生し、MV検出誤りが生じる。特に繰返しパターン画像においては、誤ったMVの候補が多数存在するため、正確なMVのBM誤差よりもBM誤差が小さいMVが出現する確率が高くなり、MV検出誤りが多数発生する。

DVSEの値はMVを検出する小領域 Ψ ごと、かつBM誤差を計算するMVごとに変わる。DVSEを計算する方法の1つはサブピクセル精度のMV検出である[1][29][30][31]。しかし、サブピクセル精度のMV検出では、結果として十分なMV検出誤りを低減することができない(第3.4節参照)。本章では、DVSEの代わりに、DVSEの最大値の近似値(maxDVSE)に着目する。maxDVSEの計算の方法を示し、maxDVSEを用いた、正確なMVを検出するための、MV検出手法を提案する。

maxDVSE

正確なMVのBM誤差にDVSEが含まれ、その取り得る最大値がmaxDVSEである。正確なMVのDVSEを含まないBM誤差(すなわち、時間方向の変化)が少ないと考えると、「正確なMVのBM誤差はmaxDVSE以下である」と考えられる。

DVSEの最大値は、図2-2の例では、1つのエッジにおいて、 ε が+1/2または-1/2のとき、 $F_N(kX) - F^-$ または $F^+ - F_N(kX)$ が最大になるため、このいずれかのときDVSEが最大になる。これは、DVSEの最大値、正確には最大値の近似値が原画像のみから算出できることを示している。なお、上述の ε の区間において、 $F_N(kX)$ に極大値または極小値が存在する場合(例:図2-2の $(k+1/2)X$ から $(k+3/2)X$ の区間)には必ずしも ε が+1/2または-1/2において最大値にはならない。しかし、極大値または極小値において $F_N(kX)$ の微係数が0になることから、 ε が+1/2または-1/2のときの $F_N(\cdot)$ の値が実際の最大値(極大値あるいは極小値)をよく近似できる。2次元の画像に適用すると、図3-1に示すようになる。○●印を画像サンプリング位置としたとき、着目点(●)に対して、A~Hの8点の箇所が、DVSEが最大となる点の候補である。この原理を用いて、DVSEのとり得る最大値の近似値maxDVSEを以下に定義する。

$$\text{maxDVSE} = \max_{(\mu, \nu) \in \Phi} \left\{ \sum_{(x, y) \in \Psi} [\hat{F}_{N, (\mu, \nu)}(x, y) - F_N(x, y)]^2 \right\} \quad (3.1)$$

ここで、 Ψ はMV検出単位の小領域、 Φ は動きサンプリング誤差が最大になる候補位置(図3-1のA~H)であり、サンプリング間隔を1としたとき、

$$\Phi = \left\{ \begin{array}{l} (-0.5, -0.5), (0, -0.5), (0.5, -0.5), \\ \quad (-0.5, 0), (0.5, 0), \\ (-0.5, 0.5), (0, 0.5), (0.5, 0.5) \end{array} \right\}$$

となる。

$\hat{F}_{N,(\mu,\nu)}(x,y)$ はN番目のフレームにおいて、 $F_N(x,y)$ を (μ,ν) だけ、サンプリング点をシフトした位置の信号である。例えば $(\mu,\nu) = (0.5, 0)$ の時は、図3-1中央●の画素のサンプリング位置を、Eの位置にずらした画像信号（補間画像）となる。

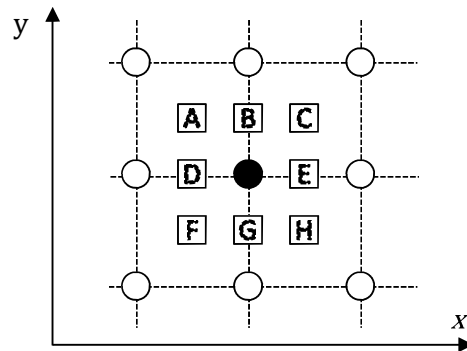


図3-1 DVSEの最大値候補箇所(A~H)

Figure 3-1 Candidates of position which makes DVSE maximum

3.3 maxDVSEを用いたMV検出方法

MV検出の問題点は、正確な（正しい）MVでも、そのBM誤差が最大でmaxDVSEになる点である。BM誤差が最小のMVを選択しても正確なMVとは限らない。特に繰返しパターンを多く含む画像の場合、多数のMV候補が存在するため、BM誤差が正確なMVの場合よりも小さくなるMVが存在する確率が高くなり、さらにMV検出誤り率が高くなる。

空間方向のMVの連続性を用いることによりMV検出誤りを低減することができる。しかし、画素の差分とMVの差分の重み付けの設定や、画面全体での大局的なレベルでMVの連続性を最適化する反復処理の処理時間（処理量）増加が課題である。MVの差分の重み付けの設定の問題は、「画素の差分」と「MVの差分」をそれぞれ連続的に扱っているために発生している。「正確なMVのBM誤差はmaxDVSE以下である」ことに着目すると、「画素の差分」（BM誤差）を、BM誤差がmaxDVSE以下であるか否かの離散的な二値情報として扱うことができる。これは、「画素の差分」（BM誤差）がmaxDVSE以下であればBM誤差の大小で優劣をつけないことを表す。「画素の差分」を二値情報とすることにより、MVの連続性を最適化する問題は、「画素の差分」がmaxDVSE以下であるMVのみに対して、「MVの差分」を考えればよいことになる。

提案方法は以下の3つのステップにより実行される。以下、小領域 Ψ を「ブロック」と置き

換えて説明をする。

- STEP1 画像内の全ブロックに対してブロックマッチング（フルサーチ）を実行し，BM誤差が各ブロックの maxDVSE 以下の全てのMVを，ブロックごとに候補MVとして記録する。
- STEP2 MVを一意に特定できるブロックを探索し，このブロックのMVを基準に，隣接ブロック間のMVの連続性を利用して，候補MVから1つのMVを選定し，そのブロックのMVとして確定する。
- STEP3 STEP2にて確定されなかったブロックのMVを確定する。

以下，MVを一意に特定できるブロック（アンカーブロック）の検出法を3.3.1に，MV確定処理の概要を3.3.2に，これらのインプリメンテーションを3.3.3に示す。なお，MVを一意に特定できるブロックをアンカーブロック，そのMVをアンカーベクトルと定義する。

3.3.1 アンカーブロックの判定

画像内の繰返しパターン部においては図3-2(a)のようにBM誤差が maxDVSE 以下となるMVがいくつか存在する。しかし，繰返しパターンの端部や，パターンの中の周期・形状が不規則となる，特異点の部分においては，MVが一意に定まり，図3-2(b)に示すように，BM誤差が maxDVSE 以下となるMVの数が1となる。このようにBM誤差が maxDVSE 以下となるMVの数が1となるとき，このMVをアンカーベクトルとし，対応するブロックをアンカーブロックと判定する。

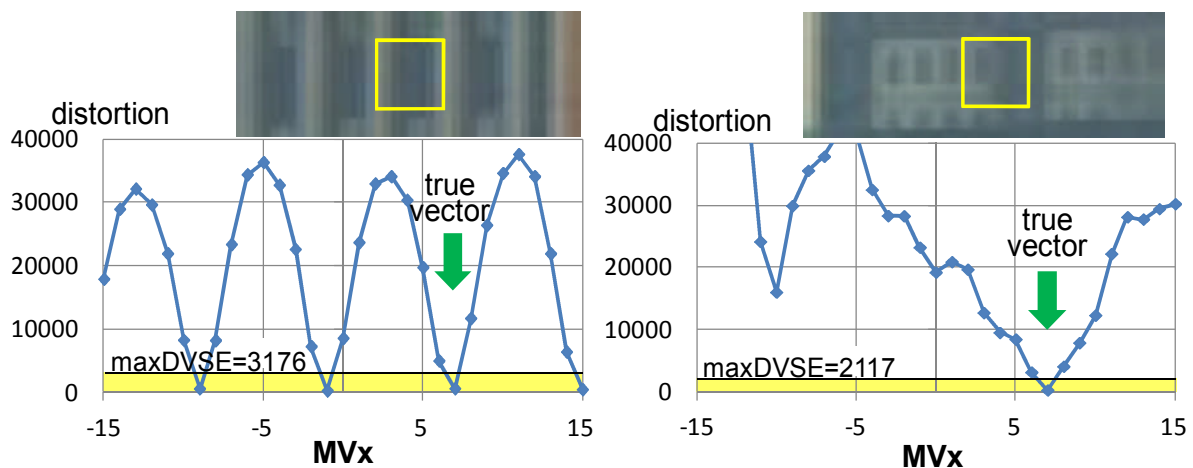


図3-2 ブロックマッチング誤差と maxDVSE との関係の例 ($MV_y=0$)

Figure 3-2 Examples of block matching distortion and maxDVSE

3.3.2 動きベクトルの確定

STEP1において、ブロックごとに、BM誤差がmaxDVSE以下である候補MVを記録した後、STEP2では、隣接するMVの連続性を利用して、候補MVの中からブロックの最終的なMVを確定する。図3-3に、MVの確定処理の動作例を示す。図3-3は画像内のブロックを左上から順にスキャンし、アンカーブロックとしてブロックpが検出された例である。この時、アンカーベクトル(3,0)をブロックpのMVとして確定した後、ブロックpの4近傍のブロックのうち、MVが未確定のブロック (q,r,t) について、候補MVを検索する。候補MVの中に、アンカーベクトルに近いMVがあれば、それらのMVをそれぞれのブロックのMVとして確定する（ブロックq,r, 図中○印）。アンカーベクトルに近いMVがない場合は、そのブロックのMVは未確定のままとする。新たに確定したブロックに対して、ブロックpの場合と同様に4近傍のブロックを検索して順次MVを確定していく。例えばブロックqに隣接するブロックuでは、ブロックqのMV(3,1)に近い△印を示したMVが検出され確定される。

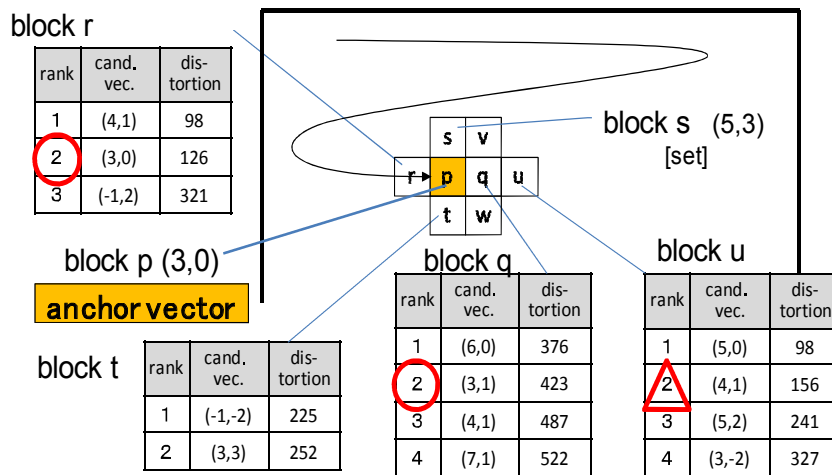


図3-3 候補動きベクトルから動きベクトルの選択例

Figure 3-3 An example of selecting motion vector from candidate vectors

3.3.3 詳細アルゴリズム

以上のアルゴリズムの詳細を示す。MV検出処理は、MVを検出する基準画像と、参照画像、それぞれの輝度信号を用い、画像（フレーム）単位で処理が完結する。

STEP1ではフレーム内のすべてのブロックに対して、(3.1)式を用いてmaxDVSEを計算する（図3-4 2行目、以下図番号省略）。その後、各ブロックにフルサーチを実行する。フルサーチにおいて、BM誤差がそのブロックのmaxDVSEよりも小さい場合は、そのMVを候補として記録しておく（5行目）。各ブロックにおいてBM誤差が最小となるMVも同時に記録しておく（7行目）。

STEP2では、まず、ブロック毎に、そのブロックのMVが確定したか否かを示す"*nd_flag*"をセットする (9行目, セット=未確定). 次に、フレーム内のブロックをスキャンし、3. 3. 2項で示したアンカーブロックを検索する (11行目). アンカーブロックが見つかったら、そのブロックのMV (アンカーベクトル) を引数としてprocedure "*set_MV_to_block*"を実行する (13行目).

```

STEP1:
1  for each block in a frame (blk_i) {
2      calculate maxDVSE ;
3      for each MV in search area(MV_j) {
4          distortion = calculate block distortion for MV_j ;
5          add MV_j to MV_candidate_list[blk_i] if distortion ≤ maxDVSE ;
6      }
7      store minimum distortion MV to min_MV[blk_i] ;
8  }

STEP2:
9  set nd_flag /*not determined' flag*/ to MV status of each block ;
10 for each block in a frame (blk_i) {
11     if blk_i is 'anchor block' {
12         MV = get MV from MV_candidate_list[blk_i] ;
13         set_MV_to_block(MV, block_i) ;
14     }
15 }

STEP3:
16 for each block in a frame (blk_i){
17     if MV status of blk_i is 'not detected' {
18         MV = min_MV[blk_i] ;
19         store MV as the motion vector of blk_i ;
20     }
21 }

-----
22 procedure set_MV_to_block( MV_p /* AMV*/, blk_p ){
23     clear nd_flag to MV status of blk_p ;
24     store MV_p as the motion vector of blk_p ;
25     for each 4-neighboring block(blk_j){
26         for each MV in MV_candidate_list[blk_j] (MV_k){
27             set_MV_to_block(MV_k, blk_j)
28             if |MV_k - MV_p| ≤ 1 && MV status of blk_j is 'not detected' ;
29         }
30 }

```

図3-4 提案方法の詳細アルゴリズム

Figure. 3-4 Detail algorithm of proposed method

procedure "*set_MV_to_block*"では、まず、引数のMVをそのブロックのMVとして確定し (23,24行目), 次にそのブロックの4近傍のブロックに類似したMVがないか検索する (26行目). 近傍のブロックに記録されているMVの中に、類似したMVが存在し、かつ、そのブロックが未確定のブロックであった場合、procedure "*set_MV_to_block*"を再帰的に呼び出すことにより、その類似したMVを、その近傍ブロックのMVとして確定する (27行目). なお、MV間の

ユークリッド距離が1以下の場合を「類似」とした。ここで着目すべき点は、近傍ブロックのMVの確定の判定基準に、MVが「一致」ではなく「類似」を用いている点である。近傍ブロックのMVが「類似」していることを許容することにより、画像内においてMVが空間方向に徐々に変化している場合も正しいMVを検出することができる

3.4 評価

3.4.1 シミュレーション

第3.2節で提案したMV検出方法を検証するためのシミュレーションを行った。諸元を表3-1にテストシーケンスを図3-4に示す。テストシーケンスは、第2章と同じ図2-7の6シーケンスを用いた。

表3-1 シミュレーション諸元
Table 3-1 Simulation conditions

test sequence	352×240pel(SIF), 15fps, 15frames, 6sequences	
block matching	block size	8×8 pel (fixed size)
	search area	H±15pel, V±7pel
	BM criterion	mean square error
	vector precision	full pel (integer)
maxDVSE interpolation	4tap filter(-0.125, 0.625, 0.625, -0.125) interpolation	
evaluation	percentage of blocks which have correct vectors	
evaluated blocks	maxDVSE≥128 and #of candidate vectors≥1 (excluding edge blocks)	

繰り返しパターン画像においては、局所解MVは繰り返しパターンの分布に対応して、離散的に分布するため、第2章における評価と同様に、これら局所解MVの中から正しいMVを選択できるか否かを、MV検出誤り率（100%－正解率），すなわち， $(1 - n_{\text{correct}}/n_{\text{eval}})$ に評価する。ここで n_{correct} はMVが正解であったブロックの個数であり，検出したMVと正解MVのユークリッド距離が th_c 以下のときを正解としている（本実験では指定のない限り $th_c=1$ ）。 n_{eval} は評価対象のブロック数であり，以下のブロックを除いたブロックが評価対象である（第2章の評価に比べ(3)(4)が異なっている）：

- (1) 平坦なブロック

$$\text{maxDVSE} < th_r \times \text{number_of_pixels_per_block}, th_r = 2. (\text{i.e., } 128 \text{ for an } 8 \times 8 \text{ block}),$$

(2) 変形したブロック，時間方向の変化の大きいブロック

minimum distortion > maxDVSE

(3) 画面の端のブロック

左右各1ブロック，上下各1ブロック

(4) シーケンスの最終フレームのブロック

いずれの方式も，(1)(2)(3)のブロックにおいて，MV検出処理は実行しており，MV検出誤り率の評価時にのみ計算対象から除外をしている。

提案方法の比較対象として，フルサーチとして，通常の整数画素精度のフルサーチ("MSE")，quarter pel ("QPEL") [2]，Birchfield & Tomasi ("B&T") [31]，パラボラフィッティング ("PF") [29]。また，隣接MVの連続性を用いる方式として，JM PMV ("PMV") [37]，階層Lucas-Kanade ("PLK") [40])，Bartels P_{III} ("BT") [32]，MDP-flow2 ("MDP") [18]を採用した。QPEL，B&T では4倍あるいは2倍（画素数は16倍あるいは4倍）にアップサンプリングした画像を用いてフルサーチを実行している。アップサンプリングした画像を用いた方が，整数精度のMVを検出した後にそのMVの周囲を小数精度で探索するtwo-stepサーチに比べて検出精度が高い。PFでは，BM誤差が極小値となるすべてのMVを記録した後，これらのMVに対してパラボラフィッティングを適用しMVを求めている。MDP，PLKでは画素単位でMVを求めた後，ブロックを構成する画素（64画素）の各MVを64タップのmedian filterによりブロックMVに変換した。PMV, PLK, BT, MDPでは，smoothness weightingなどのパラメータは予備実験により定めた。各方式とも，全テストシーケンスにて同じパラメータを用いて，全シーケンスの平均のMV検出誤り率が最も低減されるパラメータを採用している。なお，BTでは，smoothness prior P_{III}を採用し，パラメータは $\lambda=1000$ としている。

3.4.2 シミュレーション結果と考察

3.4.2.1 BM誤差の分布

図3-5に，“Skyscrapers,” “Harbor scene,” “Buildings along the canal,” “Church”の先頭8フレームの正解MVのBM誤差の分布を示す。図ではブロックのBM誤差をmaxDVSEで正規化している。正規化されたBM誤差が1.0より小さければ，正解MVが候補MVの中に含まれることとなる。図3-5では98.9%の正解MVのBM誤差がmaxDVSE以下となっている。一方，図3-5からは，BM誤差の値がmaxDVSEに近い大きな値を持つ正解MVが存在することが確認できる。

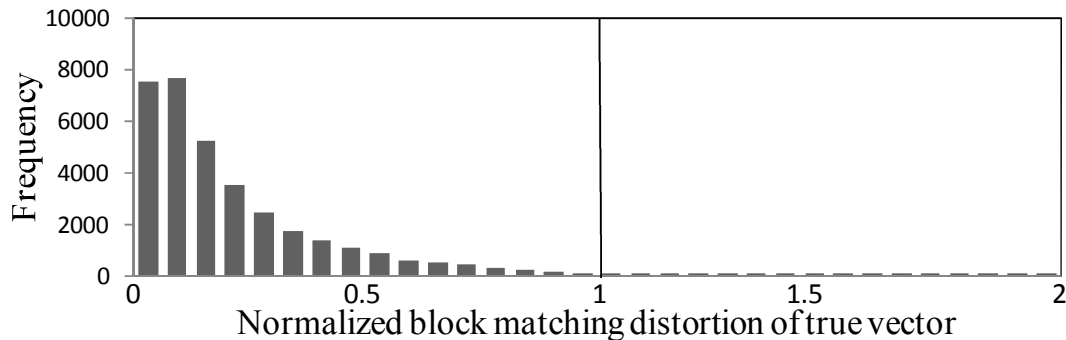


図3-5 正解ベクトルのブロックマッチング誤差の分布(maxDVSEで正規化)

Figure 3-5 Histogram of block matching distortion of true vector (normalized by maxDVSE)

3.4.2.2 DVSEの特性

正解ベクトルのBM誤差がmaxDVSE以下になっていることを確認する別の実験結果を図3-6に示す。図3-6はMV候補を記録する時にmaxDVSEを β 倍した閾値を使用した実験結果である。すなわち、図3-4の5行目の判定を以下のように変更している。

add MV_j to MV_list[blk_i] if distortion $\leq \beta \times \text{maxDVSE}$

$\beta=1$ の時の提案方法であり、図3-6では、シーケンスごとに $\beta=1$ の時のMV検出誤り率との差分をプロットしている(図3-6の縦軸が0より大きいと、 $\beta=1$ の時のよりもMV検出誤りが多くなっていることを表す)。 β が1より小さくなると正確なMVが各ブロックのMV候補に入らない確率が増え、逆に β が1より大きくなるとMV候補の中に明らかに正解でないMVが増加する。こうしたMV候補数の変化はMVの選定にも影響するが、それ以外にアンカーベクトルの判定にも大きく影響する。MV候補が少なくなると本来アンカーブロックでないブロックの候補MVの数が1となり、誤ってアンカーブロックと判断されてしまうケースが出てくる。この場合、アンカーベクトルが正確なMVであるとは限らない。一方、MV候補が多くなると、本来はアンカーブロックであるブロックにノイズのMVが加わることによりアンカーブロックと判断されないケースが発生する。

図3-6を見ると、 β が0から1に近づくに従いMV検出誤り率は低減する。 β が1を超えるとMV検出誤り率が同等あるいは増加する。一部、MV検出誤り率がわずかに減少する($\beta=1$ のときよりも誤りが少なくなる)シーケンスがあるが、このシーケンス、画面全体が一定方向に動いているシーケンスであるため、アンカーベクトルの数が減っても全体のMV検出誤り率に影響が少なかったと考えられる。全体的には、 $\beta=1$ の時の最適であり、提案方法の原理が妥当であることがわかる。

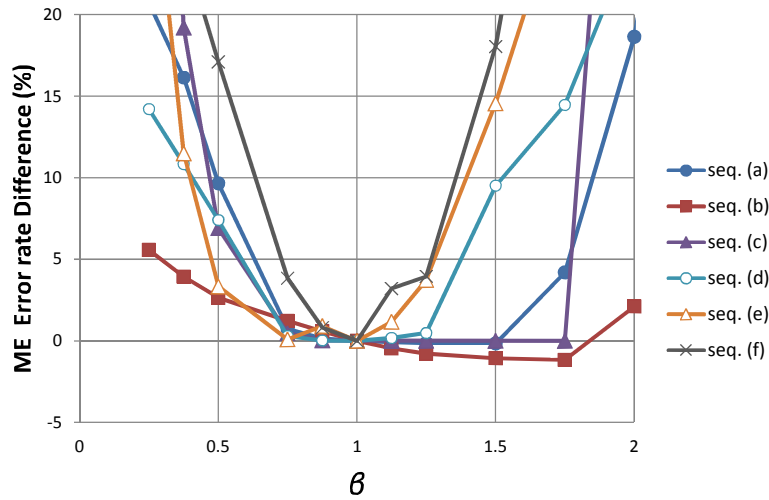


図3-6 候補MV選定閾値を変化させた時の誤り率の変化

Figure 3-6 ME error rate versus candidate MV threshold

3.4.2.3 MV検出誤り率および処理時間の結果

表3-2にシミュレーション結果および、提案方法と他の方式との比較、図3-7に検出したMVの比較を示す。表3-2の最下行以外はMV検出誤り率であり0%に近い（値が少ない）方が、性能が高いことを示す。表3-3は平均の処理時間であり、左カラムが1フレームのMV検出にかかる秒数(msec)、中央カラムが提案方法の処理時間（処理量）を1とした時の相対的な処理時間である。処理時間の算出はCore2 Duo 2.27GHzにて行った。なお、MDP、PLKでは画素単位のMVの算出までの処理時間を示しており、画素単位のMVからブロック単位のMVへの変換時間は含まれていない。

提案方法は平均のMV検出誤り率が最も低減されている。MDPが僅差で2番目となっており、いくつかのシーケンスではMDPのMV検出誤り率が下回っている。しかし、処理時間を見るとMDPは提案方法の1000倍以上の時間がかかっている。FSの誤り率は高い。FSの中でも最もMV検出誤り率が低いQPELでも提案方法よりも15%以上高い誤り率となっている。PFは空間方向に精度の高いMV（サブピクセル精度のMV）を検出できることで知られているが、BM誤差方向の精度が不十分なため、繰返しパターンへの対応を正確に判別することができていない。BTは画面が一方向にのみ動く(a)(b)(f)において、最も低いMV検出誤り率になっている。しかし、複数の動きから構成されるシーケンスではMV検出誤り率が増加している。

表3-2 シミュレーション結果

Table 3-2 Results of simulation

Method		ME error rate (%)						
		(a)skys	(b)habr	(c)lace	(d)skys/ chur	(e)lace/ canl	(f)cloth	average
proposed		0.63	1.27	0.00	1.18	2.65	0.96	1.11
FS	MSE	42.23	10.11	69.35	24.73	43.57	43.87	38.96
	QPEL	19.38	9.28	40.79	10.41	27.02	5.00	18.65
	B&T	32.26	9.59	30.17	16.88	22.05	23.76	22.45
	PF	35.27	12.16	66.12	22.34	42.35	33.29	32.25
PMV		18.02	3.38	63.34	22.12	39.14	28.96	29.16
PLK		4.26	1.67	13.89	8.92	4.50	4.47	7.13
BT ($\lambda=1000$)		0.00	0.00	17.14	4.50	12.39	0.02	5.67
MDP		0.02	2.32	1.19	0.00	2.98	0.58	1.18

表3-3 処理時間

Table 3-3 Processing time

Method		processing time*		average ME error rate (%)
		ms/frm	relative	
proposed		143	---	1.11
FS	MSE	139	1.0	38.96
	QPEL	1385	9.7	18.65
	B&T	336	2.4	22.45
	PF	308	2.2	32.25
PMV		143	1.0	29.16
PLK		3400	23.8	7.13
BT		314	2.2	5.67
MDP		144900	1014.3	1.18

* Core2 Duo 2.27GHz. Conversion time from pixel MV to block MV is not included in MDP and PLK

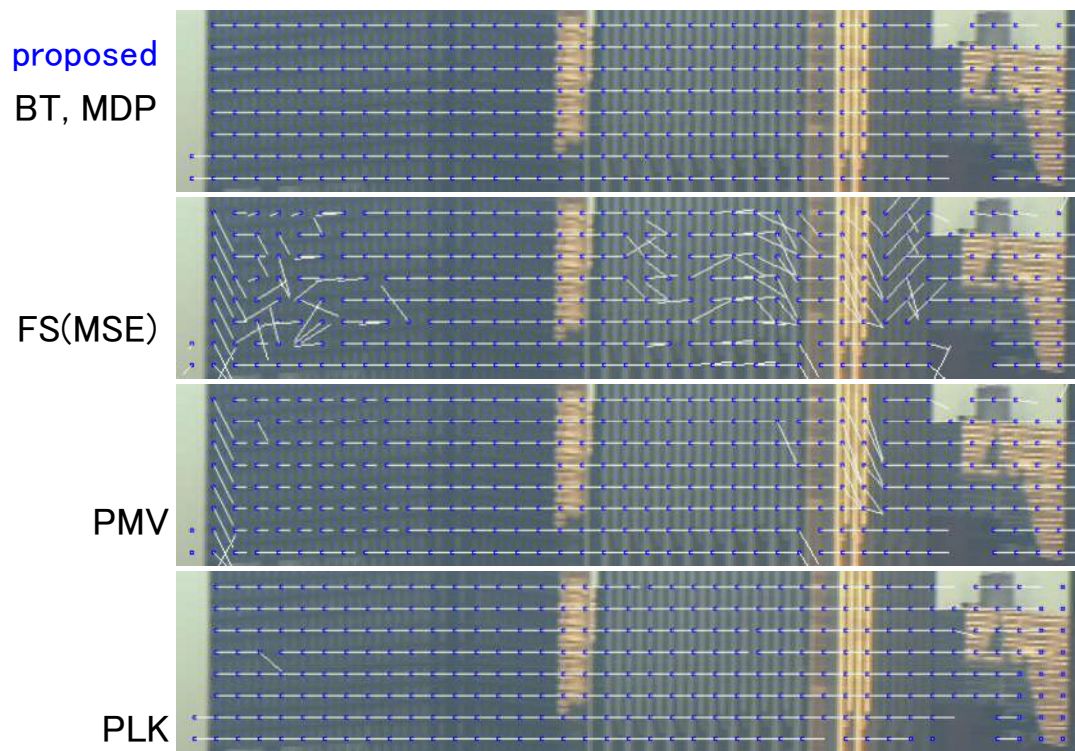


図3-7 各方式のMV比較

Figure 3-7 Comparison of MVs

図3-7では検出したMVを白線で、ブロックの中心位置を青点で示している。なお、提案方式、BT、MDPはMV検出誤りがないため同じ共通のMV画像としている。FS(MSE)は繰り返しパターンの類似性によりMVが大きく乱れている様子がわかる。PMVはFSに比べれば改善されているものの、一旦MVが誤ると周囲に誤りが波及する傾向にある。PLKのMV検出誤りは画素単位のMVからの変換に用いるmedian filterのノイズ除去効果もあり、MV検出誤りは比較的少なく、かつその発生位置は単発的である。

平均の処理時間は、提案方法がMSE、PMVとほぼ同等の少なさである。この結果から、図3.4のSTEP2およびSTEP3にかかる計算時間は無視できる程度に少ないことがわかる。

3.4.2.4 ノイズ画像に対する結果

図3-8にノイズ画像に対する耐性の実験結果を示す。テストシーケンスに乱数でガウスノイズを加え、その時のMV検出誤り率（正解MVはノイズを加える前のもの）を測定した。ノイズが35dB以上と少ない場合は提案方法、MDPともにほぼ同等の性能となる。しかし、30dBになると、MDPではMV検出誤り率の増加が0.4%程度であるのに対して、提案方法のMV検出誤り率はノイズがない場合に比べ2.7%増加する。なお、MDPでは画素単位のMVからブロック単位のMVへの変換にmedian filterを使っているため、median filterのノイズ除去（誤MV排除）の

効果がMDPのノイズ耐性に寄与している。ノイズ画像に対する耐性は今後の課題である。

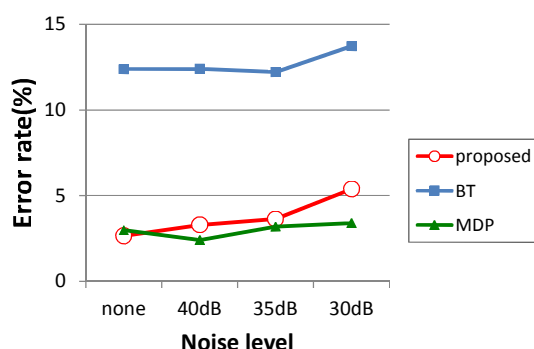


図3-8 ノイズ画像に対するMV検出誤り率(シーケンス(f))

Figure 3-8 MV error rate with noisy sequences

3.4.2.5 異なるブロックサイズ

表3-4と図3-9に、 16×16 画素、 4×4 画素のブロックに提案方法を適用した結果を示す。表3-2に示した 8×8 画素ブロック以外のブロックサイズにおいても、概ねMV検出誤り率が低減されている。しかし、表3-4のシーケンスを個別に見ると、シーケンス(e)において 4×4 画素ブロックのMV検出誤り率が大幅に増加していることがわかる。シーケンス(e)では、いくつかのフレームにおいてアンカーブロックが検出できず、その結果、図3.4のSTEP3によりBM誤差が最小となるMV (FSと同じMV) が選ばれたことが原因である。参考までに、3.4.2.2にて用いた β を0.75とした場合の結果を表3-4の右端のカラムに示す。表からもわかるように、 $\beta=0.75$ と1より小さくし、アンカーベクトルの数を増やすことにより、4%近いMV検出誤り率が得ることができる。しかし、 β を1よりも小さくすると、アンカーベクトルが正しいか否かは保証されず、また、正しいMVが候補に含まれない場合も発生する。このため、他のシーケンスではMV検出誤り率が $\beta=1$ の場合に比べやや増加している。 β の値の変更は本質的な解決ではなく、これ以外の方法にてアンカーブロックを安定的に求めることも今後の課題である。

表3-4 異なるブロックサイズでの結果

Table 3-4 Results for different block sizes

Method	ME error rate (%)						average
	(a)skys	(b)habr	(c)lace	(d)skys/ chur	(e)lace/ canl	(f)cloth	
16x16	0.94	0.11	0.00	3.55	0.57	0.46	0.94
8x8	0.63	1.27	0.00	1.18	2.65	0.96	1.11
4x4	1.04	2.4	4.66	1.47	21.82	2.15	5.59
4x4($\beta=0.75$)	1.67	5.71	1.58	1.91	4.18	4.56	3.27

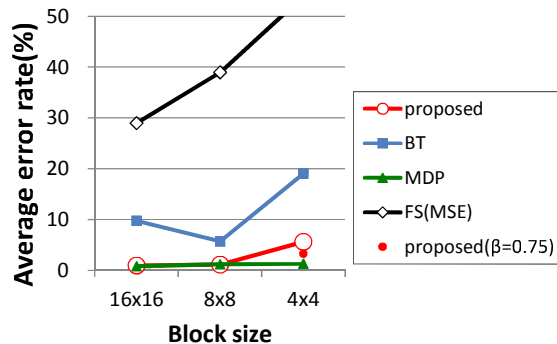


図3-9 異なるブロックサイズでの誤り率比較

Figure 3-9 Results for different block sizes

3.4.2.6 並進ベクトル以外を含むシーケンスでの結果

並進ベクトル以外を含むシーケンスは提案方法の検討対象外である。しかし、MVが空間方向に徐々に変化している場合（カメラのズームや被写体の回転など）も、局所的に見れば並進ベクトルと考えることができる。本提案方法も最大1画素/ブロックの割合でMVが変化することを許容しているため、こうした並進ベクトル以外を含むシーケンスへ適用することができる。本項では"Rotating disk"（回転）、"Driving"（ズームと同等）、"Mobile & calendar"（複数の動きの異なるオブジェクト、CIFフォーマット）を用いた実験結果を示す。ここで、注意すべき点は、シーケンスの性格上、正解MVが必ずしも正しく設定できていない場合がある点である。その意味で、特に、表3-5に示す定量的な評価はあくまでも参考値として見ていただきたい。図3-10に検出したMVを示す。図3-10による主観評価では、提案方法はMSEよりも優れ、BTと同等あるいはわずかに提案方法が優れている。

図3-11は提案方法とMDPの細部の比較である。特にMDPの検出結果が誤りであった個所を示している。上段が基準画像（Template image）、下段が検出した中央部のブロックのMVにより動き補償した画像（Compensated image）であり、上下の画像の位置が一致していれば正解、ずれていれば誤りである。図3-11の例では、MDPが数画素程度ずれていることがわかる。

各シーケンスの3フレームを定量的に評価した結果を表3-5に示す。提案方法のMV検出誤り率はMDPよりも6.8%、BTよりも0.8%、それぞれ高く（悪く）なっている。提案方法のMV検出誤り箇所を調べると、斜め方向のMVでの誤りが多く見られた。本評価では、正解の判定の閾値 $th_c=1$ としているが、斜め方向のMVではわずかな違いが誤りと判断される。例えば、正解が(2,-2)であったとき、検出されたMVが(3,-3)ならば、たとえ実際の動きが(2.5,-2.5)であっても、誤りと判断されてしまう。参考までに $th_c=\sqrt{2}$ としたシミュレーション結果を表3-6に示す。表3-6では提案方法のMV検出誤り率はBTよりも低減され、主観評価に近い結果が得られている。これらの結果から提案方法はMDPにはやや劣るものの、BTに比べればMV検出誤り時のMVの距離差が少なく、安定したMVを検出している。

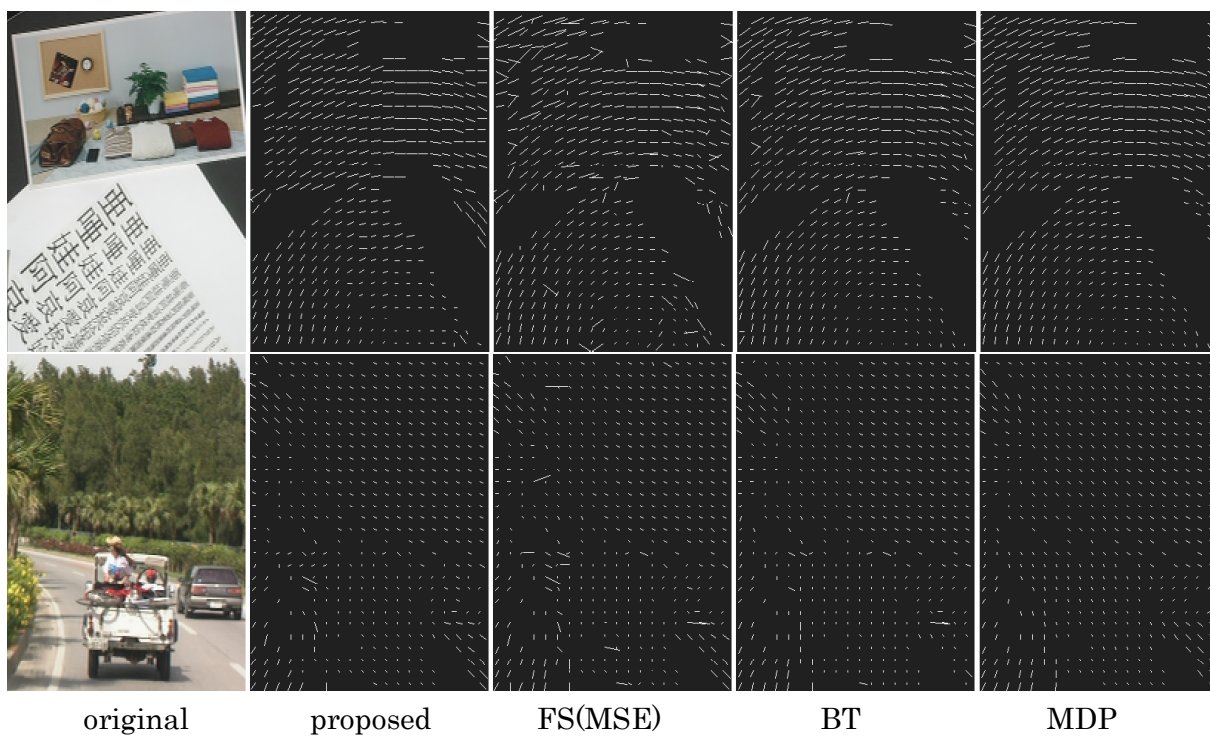
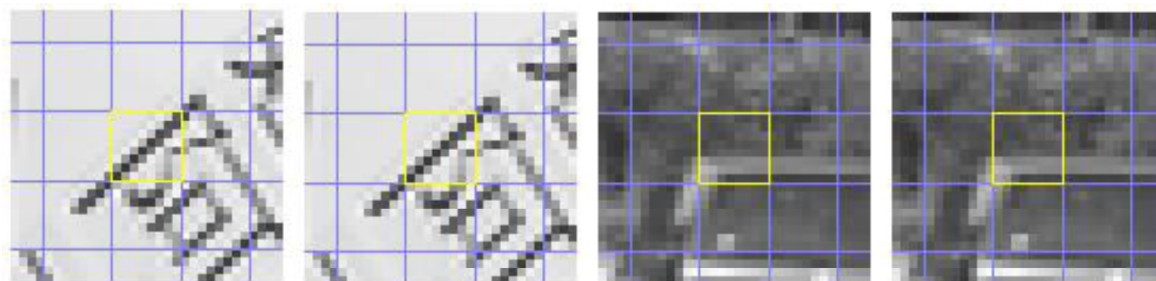


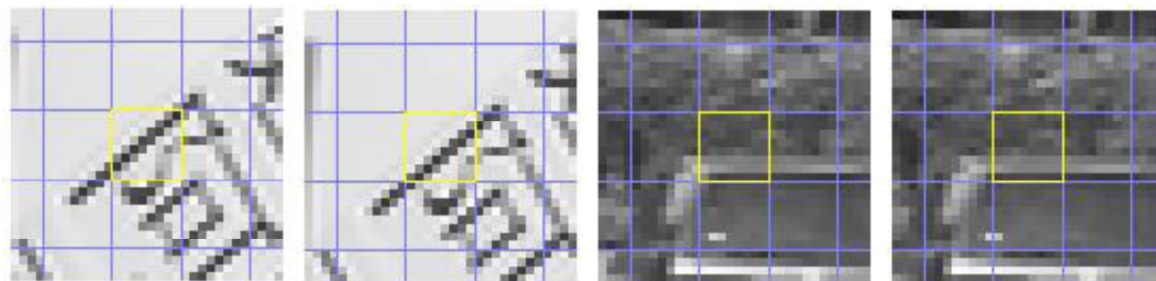
図3-10 他のシーケンスのシミュレーション結果

Figure 3-10 Simulation results for other sequences

Template image



Compensated image



proposed

MDP

proposed

MDP

図3-11 提案方法とMDPの細部の比較

Figure 3-11 Comparison of details for proposed method and MDP

表3-5 他のシーケンスのシミュレーション結果(参考)

Table 3-5 Results for other sequences

Sequences	ME error rate (%)			
	proposed	FS (MSE)	BT	MDP
Rotating disk	21.42	27.77	15.84	12.62
Driving	13.87	10.57	18.72	6.63
Mobile & calendar	7.10	14.29	5.55	2.77
average (%)	14.13	17.54	13.37	7.34

表3-6 他のシーケンスのシミュレーション結果(参考) ($th_c = \sqrt{2}$)

Table 3-6 Results for other sequences ($th_c = \sqrt{2}$)

Sequences	ME error rate (%)			
	proposed	FS(MSE)	BT	MDP
Rotating disk	11.79	21.44	8.22	8.71
Driving	7.06	6.55	13.42	3.43
Mobile & calendar	1.37	9.26	2.35	1.57
average (%)	6.76	12.42	8.00	4.57

3.4.2.7 アンカーベクトルの評価

表3-7にアンカーベクトルの平均個数, アンカーベクトルの検出誤り率, シーケンス全体のMV検出誤り率(値は表3-2と同じ)を, 図3-12に, アンカーブロックの位置を○印として示す. アンカーベクトルの検出誤り率は, すべてのアンカーベクトル数に対する正解でないアンカーベクトル数の比率であり, 0%であることが望ましい.

提案方法では, もし, アンカーベクトルが誤っていても, 周囲のブロックの候補MVの中に, 誤ったアンカーベクトルに類似したMVがなければ, 誤りはアンカーブロックにのみとどまり, 周囲のブロックに伝播することはない. 実際に, 表3-7では半数のシーケンスにおいて, シーケンス全体のMV検出誤り率がアンカーベクトルの誤り率を下回っている.

表3-7 アンカーベクトルの統計量

Table 3-7 Statics of anchor vectors

Sequences	# of anchor vec/frm	ME error rate of anchor vectors (%)	ME error rate (total) (%)
(a) Skyscrapers	19	0.00	0.63
(b) Harbor scene	111	3.52	1.27
(c) Lace	3	0.00	0.00
(d) (a)+Church	91	0.06	1.18
(e) Lace + canal	59	57.80	2.65
(f) Cloth	2	16.94	0.96

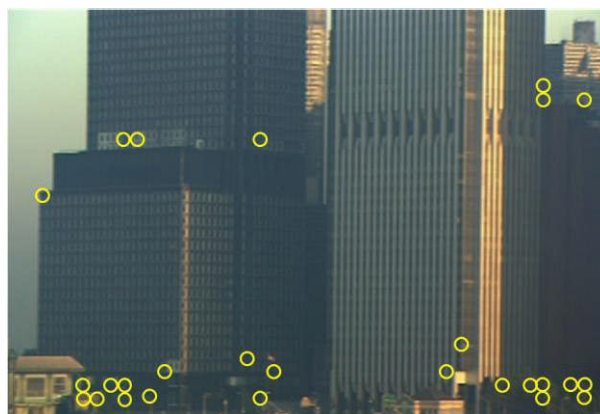


図3-12 アンカーブロックの位置(○印)

Figure 3-12 Anchor blocks (circled)

3.4.2.8 誤りの分析

提案方法について、誤った動きベクトルの解析を行った。まず、誤りの原因は以下である。

- (1) 歪等の理由で正解ベクトルが候補ベクトルに選ばれない (候補漏れ)
- (2) 該当ブロックが含まれる領域に初期アンカーベクトルが存在しない (アンカーベクトル不在)
- (3) 誤った初期アンカーベクトルにより、誤ったベクトルが選択される (誤アンカーベクトル伝播)
- (4) 隣接する異なる動きをする領域のベクトルが選択される (ベクトル漏込)

シーケンス(a)(b)(d)に関して、誤り個所の頻度を解析した結果を表3-8に示す。ここで、(3)のアンカーベクトルの誤りの影響を示す指標、および(4)の隣接する領域の影響を測る指標として、以下を定義する。

誤伝搬率：誤ったアンカーベクトル1つあたりの、アンカーベクトルの誤りに起因するMV
 検出誤り数（アンカーベクトル自体の誤りは除く：誤り伝搬数）。例えば、アン
 カーベクトルが誤り、それに連続する5つのブロックのMVが誤れば、誤り伝搬数
 を4となる。他に9個（全部で10個）のアンカーベクトルの誤りがあり、それら
 はいずれも誤りが伝搬していない（伝搬誤り数0）とすると、このフレームの誤伝
 搬率は0.4となる（ $= (4+0 \times 9) \div 10$ ）

漏込率：2つの、動きが異なる領域の境界ブロック一辺（8画素）の数に対する漏れ込み誤
 りベクトル数の比率。例えば、 3×3 の9個のブロックの領域があった場合、境
 界の辺数は12（1辺2画素 $\times 4 \div 8$ 画素）となり、もし、9ブロックのMVう
 ち1つのMVが誤って周囲の（別の動きの）領域のMVとして検出された場合は漏
 込率 $= 1 \div 12 = 8.3\%$ となる。

表3-8 動きベクトル検出の誤り原因

Table 3-8 Analysis of the causes of motion vector detection error

sequences	誤り原因(%)				誤伝 播率	漏込 率(%)	備考
	(1)候補 漏れ	(2) 不在	(3) 伝搬	(4) 漏込			
(a) Skyscrapers	38	62	0	—	0	—	(2)は15フレーム中、2 フレームのみに集中
(b) Harbor scene	99	0	1	—	0.04	—	(1)の3/4が画面中の海 面の波部分.
(d) (a)+Church	68	30	0	2	0	0.3	

全体として、一番頻度の高い原因は、(1)の正解ベクトルが候補ベクトルにない「候補漏れ」
 であった。これは、図3-4における、横軸が1.0より右側、BM誤差がmaxDVSEを超えている正
 解ベクトルのブロック（全体の1.1%）に該当する。

(2)の初期アンカーベクトルが存在しない場合は、2番目になるが、(2)に関しては、シーケン
 ス(a)のうち2フレームにのみ集中的に生じている。これらのフレームではパニングにより新た
 に画面に現れてくる被写体の部分にて「アンカーベクトル不在」の領域が存在している。

(3)の初期アンカーベクトルが誤った場合の影響は非常に小さい。初期アンカーベクトルが
 誤っていても、前項で述べたように隣接するブロックの候補ベクトルに、誤ったアンカーベク
 トルと同じベクトルが存在しなければ、アンカーベクトルの誤りは伝播しない。実際に、誤っ
 た初期アンカーベクトル1つあたりの、初期アンカーベクトルの誤りに起因する動きベクトル
 誤りの数（誤伝播率）はシーケンス(b)では0.04個、すなわち20個の誤ったアンカーベクトルが

あった場合でも、それらのアンカーベクトルの周囲のMVへ誤りが伝搬するのは1か所のみである（伝播して誤るベクトル数を1とした場合）で、残りの19個のアンカーベクトルの誤りは伝播せずに、アンカーベクトルの箇所ブロックのみが独立して誤ることを示している。

(4)の異なる動きの領域が存在する場合は、異なる動きのブロックが接しており、かつ一方のブロックに、他方の動きベクトルと同じベクトルが存在する場合の誤りは少ない。2つの領域の境界8画素（ブロックの1辺）あたりの漏れ込み誤り数（漏込率）は、該当するシーケンス(d)においても0.3%である。

3.4.2.9 提案方法の画像適応性

提案方法の大きな特徴の1つに入力画像への自動適応性を挙げることができる。たとえば、BT、MDPなどは、MVの連続性の重みづけパラメータの選定が重要であり、本評価でも、多くのパラメータ候補に対して事前実験を行い、最適なパラメータを定めた後にシミュレーションを行った。一方、提案方法は、MV検出処理の中でキーとなるパラメータmaxDVSEが入力画像自身から生成している。このため、基本的にパラメータの調整を行うことなしにシミュレーションを行っている。

3.4.2.10 第2章のMV検出方法との関係

第2章のMV検出方法と第3章のMV検出方法は正確なMVを検出する同じ目的で検討されている。その結果、繰返しパターンの多い画像では、いずれも従来の手法に対して、MV検出誤りを低減することでできている。表3-9に、第2章、第3章両方法の特長をまとめる。評価対象のブロック数の違いはあるものの、MV検出誤り率では、第2章の方法（TPD-S6T/MSE+BT）が優れ、処理量では第3章の方法が優っている。

表3-9 第2章・第3章の方法の比較

Table 3-9 Comparison of the proposed ME methods (chap. 2 and chap. 3)

方式	特長
第2章 (TPD-S6T/MSE+BT)	<ul style="list-style-type: none"> ・MV検出誤り率が低い（高性能） ・性能が安定している ・他の方式との組み合わせが容易
第3章 (maxDVSEを用いた方法)	<ul style="list-style-type: none"> ・処理量（処理時間）が短い（約1/5） ・処理遅延が少ない ・必要なフレームメモリ数が少ない

3.5 むすび

本章では、ブロックマッチングにおいて、被写体あるいはカメラの動きを表す「正確なMV」を検出することを目的としたMV検出方法を提案した。提案方法は、繰り返しパターンを含む画像あるいは画像の一部が一定方向に動いているシーケンスを対象としている。

まず、MV検出誤りの原因である、MVの量子化誤差成分により生じるブロックマッチング誤差（DVSE：block matching Distortion caused by motion Vector Sampling Error）の最大値 $\max DVSE$ を定量化する手法を示した。

次に、正確なMVのブロックマッチング（BM）誤差が $\max DVSE$ 以下である点に着目したMV検出方法を提案した。ブロックごとにフルサーチのMV探索を実行し、このとき、BM誤差が $\max DVSE$ 以下のMVをそのブロックのMV候補として全て記憶しておき、1フレーム全てのブロックマッチング処理が完了した後に、記憶しておいたMVから各ブロックのMVを確定させてゆく。各ブロックのMV確定においては、ブロック間のMVの連続性を用いる。1フレーム全体から、MVが一意的に定まるアンカーブロックを求める。アンカーブロックの周辺の各ブロックについて、各ブロックのMV候補の中にアンカーブロックのMV（アンカーMV）と類似しているMVを探索し、類似MVが存在すればそのブロックのMVとして確定する。これを再帰的に繰り返し各ブロックのMVを確定する。

提案した方法は、実画像を用いたシミュレーションにより効果を検証した。評価には繰り返しパターンを多く含む実画像のシーケンスを用い、予め算出した正解ベクトルを用いて、MV検出誤り率として評価を行った。提案方法はフルサーチと同程度の処理量にて、大部分のシーケンスにおいてMV検出誤り率3%以下（97%以上の正解率）が得られ、提案方法が正確なMVを安定して検出できることが確認できた。他の方式と比べても、よりも低いMV検出誤り率を実現することを確認した。合わせて、ノイズ画像への耐性、ブロックサイズを変えた時の影響、提案方法の検出誤りの箇所の評価などを分析考察した。

ノイズ画像への耐性の検討、アンカーベクトルの安定的な検出方法、可変ブロックサイズへの対応は今後の課題である。

第4章

画像モデルを用いた動き補償フレーム間予測効率の検討

4.1 はじめに

デジタル動画像符号化技術の標準化ISO/IEC13818-2 | Recommendation ITU-T H.262 (MPEG-2video) [1]では、それまでの標準であるITU-T H.261[13]やISO/IEC11172-2 (MPEG-1 video) [33]と同様に「動き補償フレーム間予測符号化 (Motion Compensation, MC)」と「離散コサイン変換 (DCT)」を基本とした符号化技術を採用しているが、特にMCを改良することにより、高圧縮・高画質な符号化を実現している。

こうした、MPEG-2 videoの標準を決める過程においては、数多くのMC方式が提案され、予測の効率等が比較検討されてきた。これらの比較は、H.261のRM (Reference Model)を用いた方法[53]に代表される参照モデルを利用した方法が採用された。参照モデルを利用する方法では、実形態に近い符号化システム (例えば、MPEG-2 videoではTM (Test Model) [34]) を構築し、検討するMCをその符号化システムに組み込み、動画像データを一定符号量の下で符号化し、画質やS/N等を測定する実験を行い比較評価する方法である。参照モデルを利用する方法は動画像標準化の代表的な検討方法となっており、MPEG-1のSM (Simulation Model) [54] や、その後の標準化MPEG-4のVM (Verification Model) [55]、H.264/MPEG4-AVCのJM (Joint Model) [37]、H.265/HEVCのHM (HEVC Test Model) [56]などで採用されている。

こうした動画像処理方式の検討において最大の問題点の一つは、評価画像の選択である。画像は1つのフレームの内容自体が、さまざまなエッジや、平坦面、緩やかな輝度の変化、テクスチャなど千差万別のパーツの組み合わせから構成されており、さらに動画像ではこれらがさまざまな方向に、さまざまな領域単位で移動し、輝度や色の変化、形の変形、ノイズや特殊効果などが発生し付加され、これらの組み合わせバリエーションは、そのオーダさえも計数できないほど膨大になる。これらのすべての組み合わせを評価することはとうてい不可能であり、動画像符号化の標準化においても限られた数の評価画像を用いて評価を行わざるをえなかった。例えば、1991年に開催されたMPEG-2の方式プロポーザル評価[57]に用いられたシーケンスは4種類であった。

また、映像機器の製品化等において、ある定まった処理方式 (例えば動画像符号化) を実装した機器を商品化する場合、方式や実装に不具合がないかを調べるために、多くの実験者が、数か月かけて、さまざまな動画像をその機器で再生あるいは撮影を行い、画質の不具合がないことを確認している。しかし、数か月かけての確認でも動画像の被写体と動きの組み合わせの総数に比

べればその比率は零に限りなく近いとあると考えられる。こうした、数か月かけての評価は、1つの、すでに定まった処理方式に限定した場合には可能あるが、その方式の初期の検討の時点で、膨大な映像コンテンツを処理して評価比較を行うことは、方式の試行錯誤やパラメータの調整等を繰り返し実行する必要があるため、実質上不可能である。

膨大な画像データを使用することなしに方式の評価を行う評価方法として画像をモデル化して検討を行う方法である。しかし、動画は様々なバリエーションを持つため、モデルの構築が課題である。従来は、ARモデルを用いた解析法[58]が紹介されている。

本章では、画像モデルを用いて、MCの符号化効率の評価を行うことを目的とする。MCの予測誤差が主に画像内の物体の輪郭（エッジ）周辺に多く発生することに注目し、モデル化したエッジを用いて、サンプリング点とエッジとの位相成分のずれに起因する誤差電力を確率的手法により解析する方法を検討する[25][26][27]。特に、これを一般的なモデルに拡張してMCの予測効率を解析する方法を提案する。同時に提案方法を用いた解析結果を示す。

本章の構成は以下である。第4.2節では、画像モデルとして、周波数帯域をパラメータとするエッジモデルおよびインパルスモデルの2つのモデルを、MCモデルとして単画素予測、ハーフペル予測、平均値型予測の3つの予測値生成法のモデルを提案する。第4.3節にて、それぞれの画像モデルとMC方式のモデルを用いて各MC方式の予測効率を算出する方法を提案する。提案方法では、画像モデルのエッジあるいはインパルスの位置とサンプリング位置の位相差を確率変数とし、予測誤差電力の期待値を定量的に算出し、この予測誤差電力の期待値の逆数を予測効率と定義する。第4.4節では実画像シーケンスを用いたシミュレーションにより、第4.3節で提案した予測効率の評価を行う。第4.5節、第4.6節ではそれぞれノイズ耐性、簡易的な予測効率計算手法に関して述べる。

4.2 モデル化

4.2.1 画像のモデル化

画像モデルの設定にあたり、以下に示す4項目を前提条件とする。

- ・画像入力時のカメラの特性による影響（例えば、撮像管の蓄積効果等）はない。
- ・画像内の物体は変形あるいは回転せずに、平行に移動する。
- ・画面端の影響やアンカバードバックグラウンド（移動物体通過後の背景の露出）による影響はない。
- ・画像モデルは1次元モデルとする。

以上の条件は、例えば、カメラをゆっくりパンニングしながら静止した物体を撮影するような状況に対応する。

MCの予測誤差が主に画像内の物体のエッジ周辺に多く発生するため、画像モデルとして、ステップ関数を周波数 B にて帯域制限した式(4.1)に示すエッジモデル $S(x, B)$ を第一の画像モデル

とする.

$$S(x, B) = 2B \int_{-\infty}^x \frac{\sin(2\pi By)}{2\pi By} dy \quad (4.1)$$

ここで, x は画像位置 (1次元) である. また, 2つのエッジが組み合わせられた特殊な場合として, インパルス周波数 B にて帯域制限した式(4.2) に示すインパルスモデル $P(x, B)$ を第二の画像モデルとする.

$$P(x, B) = \frac{\sin(2\pi By)}{2\pi By} \quad (4.2)$$

エッジモデル, インパルスモデルの波形を図4-1に示す.

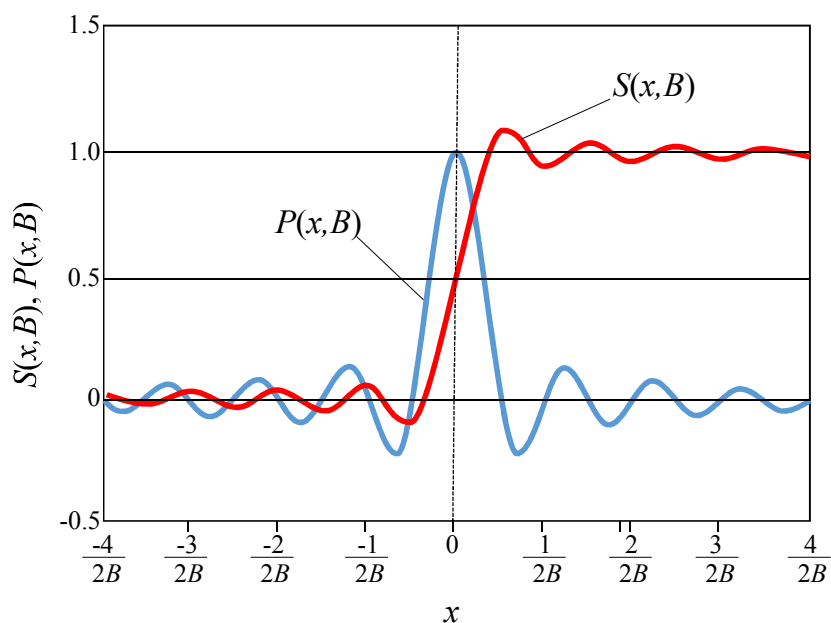


図4-1 エッジモデルおよびインパルスモデル

Figure 4-1 Waveform of edge model an impulse model

4.2.2 MCのモデル化

本節ではMCの効果を検討するためのモデル化の検討を行う. MCの効果を検討するに先立ち, MCの効果を

- ・予測画素値生成法に依存する部分
- ・予測画素選択構造に依存する部分

に分けて考える。前者は、例えば1つの予測画素を生成するための参照画像の画素の数、空間補間／時間補間の選択、また、時間補間する場合はその時間（予測距離）の比など、画素単位の操作に該当する。一方、後者は、MC単位の大きさ・形状、あるいは動きベクトルの探索法、範囲など、画素よりも大きい単位での操作に該当する。

提案する方法は前者の予測効率を検討するものである。したがって、後者の予測画素選択構造に関しては選択が理想的に行われ、選択の誤りあるいは選択時の制限に起因する予測効率の低下はないものと仮定している。

具体的には、以下の3種の予測の予測画素値生成法を解析する（図4-2）

- ・単画素予測（1つの画素を用いて1つの画素を予測）
- ・ハーフペル予測（空間補間した画素を用いて予測）
- ・平均値型予測（時間軸の異なる2画像を平均値化した画素を用いて予測）

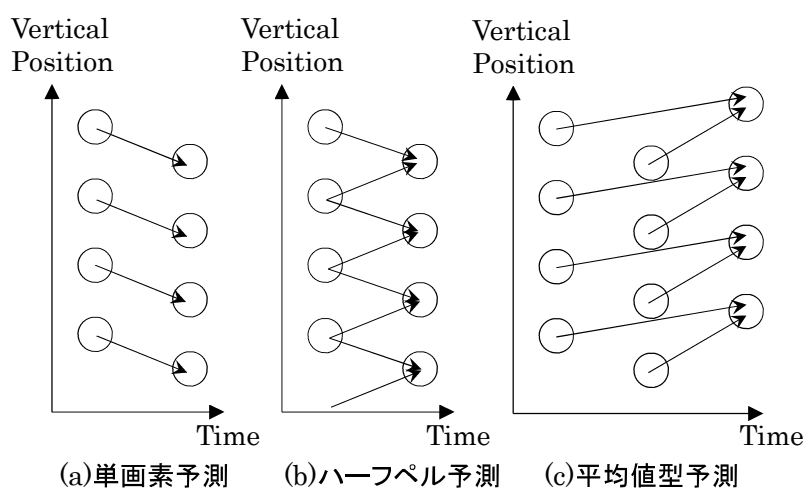


図4-2 予測画素の生成法

Fig.4-2 Methods of calculating predictive pixels

4.2.3 動き補償処理

4.2.3.1 画像モデルを用いた動き補償

X を1次元の空間軸としたとき、連続な関数である画像モデル $m(X)$ を用いて、次のようにサンプリングされた画像モデル $M(i, \varepsilon)$ を定義する

$$M(i, \varepsilon) = m(iX - \varepsilon) \tag{4.3}$$

ここで、 i は整数、 X はサンプリング間隔、 ε は $\varepsilon \in [-\frac{X}{2}, \frac{X}{2})$ で定義されるサンプリング点と画像基準位置との位相成分である。式 (4.3) の $m(x)$ として $S(x, B)$ あるいは $P(x, B)$ を適用することにより動き補償動作を記述する。

4.2.3.2 単画素予測(Simple prediction)

参照画像の1つの画素を予測画素とする予測を単画素予測（以下smp1予測）と呼ぶ。ある時刻に

$$M_{ref}(i) = M(i, \varepsilon_1) = m(iX - \varepsilon_1) \quad (4.4)$$

で表される画像が、次の時刻に距離 V_0 だけ左に移動した場合、新しい画像 $M_{new}(i)$ は

$$M_{new}(i) = m(iX - \varepsilon_1 + V_0) \quad (4.5)$$

と表すことができる。ここで、 ε_1 は $\varepsilon_1 \in [-\frac{X}{2}, \frac{X}{2})$ で定義される位相成分である。いま、新たな画像 $M_{new}(i)$ を $M_{ref}(i)$ を用いてMC予測する。ここで、 $[a]$ を a を越えない最大の整数。また、以下に定義する表記法

$$\langle a \rangle_b = b \left[\frac{a}{b} + 0.5 \right] - a \quad (4.6)$$

を用いて

$$v = \left[\frac{V_0 - \varepsilon_1}{X} + 0.5 \right] \quad (4.7)$$

$$\varepsilon_0 = vX - (V_0 - \varepsilon_1) = \langle \varepsilon_1 - V_0 \rangle_x \quad (4.8)$$

を定義する。(4.7), (4.8) 式を用いると、(4.5) 式は

$$M_{new}(i) = m((i + v)X - \varepsilon_0) = M(i + v, \varepsilon_0) \quad (4.9)$$

と表すことができる。ここで、 v はMC予測に用いるおおまかな動きベクトル、 ε_0 は移動後の画

像の位相成分である。一方、動き補償後の予測画像 $M_{prd}(i)$ は、動きベクトル v と、その補正項 \hat{v} を用いて

$$M_{prd}(i) = M_{ref}(i + v + \hat{v}) \quad (4.10)$$

と表すことができる。補正項 \hat{v} は $\varepsilon_0, \varepsilon_1$ の位相差を補償するための動きベクトルで $\hat{v} = \{-1, 0, 1\}$ の範囲にて、 $M_{new}(i) - M_{prd}(i + v + \hat{v})$ の予測誤差電力を最小にする値である。 \hat{v} の値は、 $(\varepsilon_1 - \varepsilon_0)/X$ を四捨五入した値で近似できる。

(4.9), (4.10)両式より、この予測における予測誤差電力 D_{smpl} は

$$\begin{aligned} D_{smpl} &= \sum_{i=-\infty}^{\infty} \left(M_{new}(i) - M_{prd}(i + v + \hat{v}) \right)^2 \\ &= \sum_{i=-\infty}^{\infty} \left(m((i + v)X - \varepsilon_0) - m((i + v + \hat{v})X - \varepsilon_1) \right)^2 \\ &= \sum_{i=-\infty}^{\infty} \left(m(iX - \varepsilon_0) - m((i + \hat{v})X - \varepsilon_1) \right)^2 \end{aligned} \quad (4.11)$$

となる。 \hat{v} は位相成分 $\varepsilon_0, \varepsilon_1$ によって決まる値であるため、 D_{smpl} は $\varepsilon_0, \varepsilon_1$ をパラメータとする関数として表される。また(4.11)式より、 D_{smpl} は v には依存しないため、以降、 $v=0$ とすることにより、(4.11)式は次のように記述できる。

$$D_{smpl}(\varepsilon_0, \varepsilon_1) = \sum_{i=-\infty}^{\infty} \left(M(i, \varepsilon_0) - M(i + \hat{v}, \varepsilon_1) \right)^2 \quad (4.12)$$

4.2.3.3 ハーフペル予測(Half pel prediction)

ハーフペル予測（以下hpel予測）では、 ε_2 を位相成分に持つ画像モデルを用いて次の関数を定義する。

$$H_{refs}(i, frc, \varepsilon_2) = \sum_{j=1}^{N \cdot S - 1} C(j) \hat{M} \left(i + \frac{j - j_0 + frc}{S}, \varepsilon_2 \right) \quad (4.13)$$

ここで、

$$\hat{M}(x, \varepsilon) = \begin{cases} M(x, \varepsilon) & \text{if } x \text{ is integer} \\ 0 & \text{otherwise} \end{cases} \quad (4.14)$$

$C(j)$ は補間フィルタ係数 ($j = 0$ が中心), また, S はハーフペル補間のアップサンプリングレート, N はフィルタタップ数, j_0 はフィルタタップの位相補正項である. frc は小数成分を表す整数値で $0 \leq frc < S$ であり, frc/S が実際の小数成分になる.

MPEG-2のhpel予測では $U = 2, N = 2, C(-1) = C(1) = 0.5, C(0) = 1.0, j_0 = 2$ で, すなわち

$$H_{ref2}(i, frc, \varepsilon_2) \Big|_{N=2} = \begin{cases} M(i, \varepsilon_2) & frc = 0 \\ \frac{M(i, \varepsilon_2) + M(i+1, \varepsilon_2)}{2} & frc = 1 \end{cases} \quad (4.15)$$

また, 例えば4タップのhpel予測としては, $U = 2, N = 4, C(-3) = C(3) = -0.125, C(-2) = C(2) = 0, C(-1) = C(1) = 0.625, C(0) = 1.0, j_0 = 4$ などがある.

式(4-15)の関数を用いると, hpel予測の予測誤差電 D_{hpel} は

$$D_{hpel}(\varepsilon_0, \varepsilon_2) = \sum_{i=-\infty}^{\infty} \left(M(i, \varepsilon_0) - H_{ref2}(i + \hat{v}_2, \hat{v}_{2frc}, \varepsilon_2) \right)^2 \quad (4.16)$$

と表すことができる. ここで \hat{v}_2 は動きベクトル補正項 $\{-1, 0, 1\}$, \hat{v}_{2frc} は動きベクトルの小数成分 $\{0, 1\}$ である. すなわち, 動きベクトルは $\hat{v}_2 + \hat{v}_{2frc}/2$ となる. \hat{v}_2 および \hat{v}_{2frc} は, smpl予測時と同様に予測誤差電力を最小とする値であり, 近似的には, $2(\varepsilon_2 - \varepsilon_0)/X$ を四捨五入した値が $2 \cdot \hat{v}_2 + \hat{v}_{2frc}$ となる.

4.2.3.4 平均値型予測 (Average Type Prediction)

平均値型予測 (以下ave予測) では時間軸の異なる2つの画像を用いて対応する画素ごとに平均値処理を行い, 予測画像を生成する. 具体例としては, デュアルプライム予測, デュアルフィールド予測 (以下, デュアルフィールド予測をデュアルプライム予測も含む総称として用いる: DF予測) や, 時間的に前後の画像から予測するインターポレイテッド予測 (以下IP予測) がある. これらは同一の記法にて定義できる.

一般に予測に用いる2つの画像はそれぞれの参照画面内から独立に, 予測誤差電力の最小のものが求められる. したがって, ave予測の予測誤差電力 D_{ave} は, ε_3 を位相成分に持つ第2の予測画像を用いて

$$D_{ave}(\varepsilon_0, \varepsilon_1, \varepsilon_3) = \sum_{i=-\infty}^{\infty} \left(M(i, \varepsilon_0) - \frac{M(i + \hat{v}_1, \varepsilon_1) + M(i + \hat{v}_3, \varepsilon_3)}{2} \right)^2 \quad (4.17)$$

被予測画像から第1の予測画像への距離と, 第2の予測画像への距離の比を μ ($|\mu| \geq 1$) とすると, ε_3 は式(4.8)式を利用して

$$\varepsilon_3 = \langle \varepsilon_0 + \mu V_0 \rangle_X = \langle \varepsilon_0 + \mu(\varepsilon_1 - \varepsilon_0) \rangle_X \quad (4.18)$$

と表される. 式 (4.18) 式は ε_3 が独立変数でないことを示す. このことから, \hat{v}_3 も独立変数でなくなり, これらをそれぞれ $\varepsilon_3(\varepsilon_0, \varepsilon_1, \mu)$, $\hat{v}_3(\varepsilon_0, \varepsilon_1, \mu)$ と表すことにより, 式 (4.17) 式は以下のようなになる.

$$D_{ave}(\varepsilon_0, \varepsilon_1) = \sum_{i=-\infty}^{\infty} \left(M(i, \varepsilon_0) - \frac{M(i + v_1, \varepsilon_1) + M(i + \hat{v}_3(\varepsilon_0, \varepsilon_1, \mu), \varphi_3(\varepsilon_0, \varepsilon_1, \mu))}{2} \right)^2 \quad (4.19)$$

μ の例としては, $\mu=2$ の時がデュアルプライム予測 (予測距離比が1:2の2つの画像を用いたDF予測に相当), $\mu=1.5$ の時が予測距離比2:3の画像を用いたDF予測, $\mu=-1$ の時が前後の等距離の画像から予測したIP予測 (等IP予測, MPEGにおけるM=2のBピクチャ等. 図4-3参照), $\mu=-2$ の時が前後の一方が2倍の距離の画像から予測したIP予測 (M=3のBピクチャ等) に相当する.

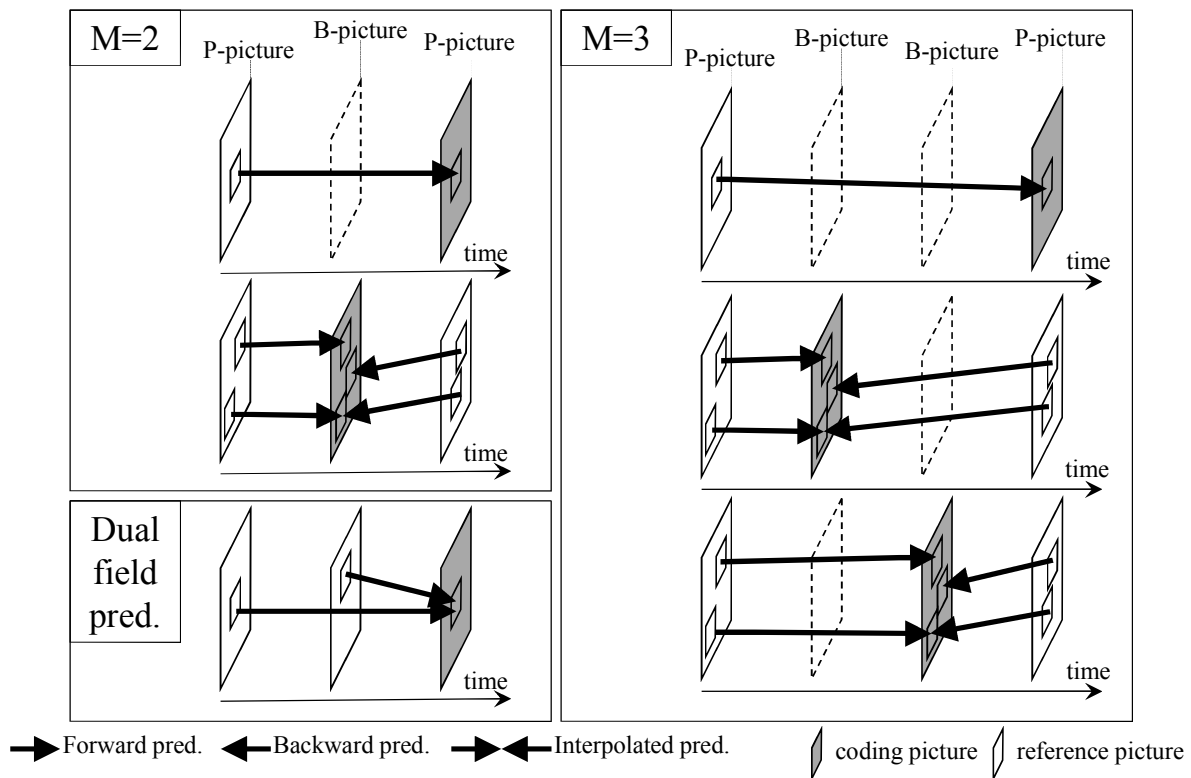


図4-3 MPEG-2における予測構造

Figure 4-3 Interframe prediction methods in MPEG-2

4.3 動き補償誤差電力の定量的評価

4.3.1 予測効率の定義

MCの予測誤差電力は式(4.12), (4.16), (4.19)のように原画像とサンプリング位置との位相成分の関数となる. 一般にエッジあるいはインパルスの位置とサンプリング点は相関がないため, 位相成分 $\varepsilon_0, \varepsilon_1$ 等は図4-4に示すように, $[-X/2, X/2]$ で一様分布のランダム変数と考えることができる. そこで, 動き補償の予測効率を予測誤差電力の位相成分($\varepsilon_0, \varepsilon_1$)に対する期待値の逆数と定義する. smpl 予測では, 予測効率 E_{smpl} は

$$E_{\text{smpl}} = \frac{1}{E\{D_{\text{smpl}}\}} = \frac{1}{\iint_{-\infty}^{\infty} D_{\text{smpl}}(\varepsilon, \Psi) f_{01}(\varepsilon, \Psi) d\varepsilon d\Psi} \quad (4.20)$$

となる. $f_{01}(\varepsilon, \Psi)$ は $\varepsilon_0 = \varepsilon, \varepsilon_1 = \Psi$ の結合確率密度関数であり, ε_0 と ε_1 は互いに独立であるため

$$f_{01}(\varepsilon, \Psi) = f_0(\varepsilon)f_1(\Psi) = \begin{cases} \frac{1}{X^2} & \varepsilon, \Psi \in [-X/2, X/2] \\ 0 & \text{otherwise} \end{cases} \quad (4.21)$$

となる. ここで $f_0(\varepsilon), f_1(\Psi)$ はそれぞれ $\varepsilon_0, \varepsilon_1$ の確率密度関数である. この結果, smpl 予測の予測効率は以下の式で表すことができる.

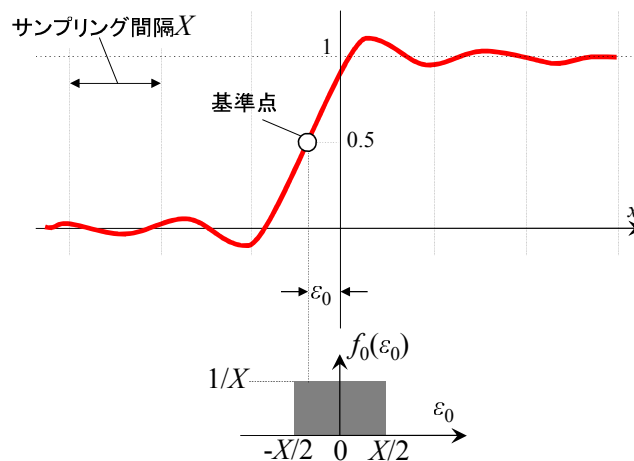


図4-4 ε_0 の確率密度関数

Figure 4-4 Density function of ε_0

$$E_{\text{smpl}} = \frac{X^2}{\int_{-\frac{X}{2}}^{\frac{X}{2}} \int_{-\frac{X}{2}}^{\frac{X}{2}} D_{\text{smpl}}(\varepsilon, \Psi) d\varepsilon d\Psi} \quad (4.22)$$

同様に、hpel予測、ave予測の予測効率は

$$E_{\text{hpel}} = \frac{1}{E\{D_{\text{hpel}}\}} = \frac{X^2}{\iint_{-\infty}^{\infty} D_{\text{hpel}}(\varphi, \Psi) d\varphi d\Psi} \quad (4.23)$$

$$E = \frac{X^2}{\iint_{-\infty}^{\infty} D_{\text{ave}}(\varphi, \Psi) d\varphi d\Psi} \quad (4.24)$$

となる。

4.3.2 予測効率の計算

式(4.22)(4.23)(4.24)にエッジモデル、インパルスモデル代入し、帯域制限周波数 B をパラメータとして数値的に計算した結果を図4-5、図4-6に示す。また、smpl予測に対する各予測の予測効率比を図4-7に示す。なお、図4-5、図4-6は比較を容易にするために、予測効率の逆数（予測誤差電力の期待値）を縦軸にとってある。動きベクトルおよびハーフペルの動きベクトルの補正項計算には、予測誤差電力を最小とするものを用いた。図中、IP予測のパラメータ M はMPEGにおける非Bピクチャの周期である（図4-3参照）。

MPEGで採用されている各予測方式は、smpl予測に比べると予測効率が1.5倍以上になっている。いずれの方式も周波数が高くなるにつれ相対的に効率が悪くなるが、その傾向はDF予測が最も弱い。

周波数が低い領域において、hpel予測は予測誤差の原因となる「位相のずれ」の範囲が、smpl予測に比べ $1/2$ の範囲になるため、誤差電力の期待値が約 $1/4$ になる。また、DF予測では2つの予測画像の誤差の振幅が、平均値処理によりそれぞれ $1/2$ になるため、誤差電力の期待値が約 $1/2$ になっている。 $M=2$ のIP予測の場合は、DF予測と同様の効果の他に、後に説明する位相差相殺の効果により予測誤差比が高くなっている。

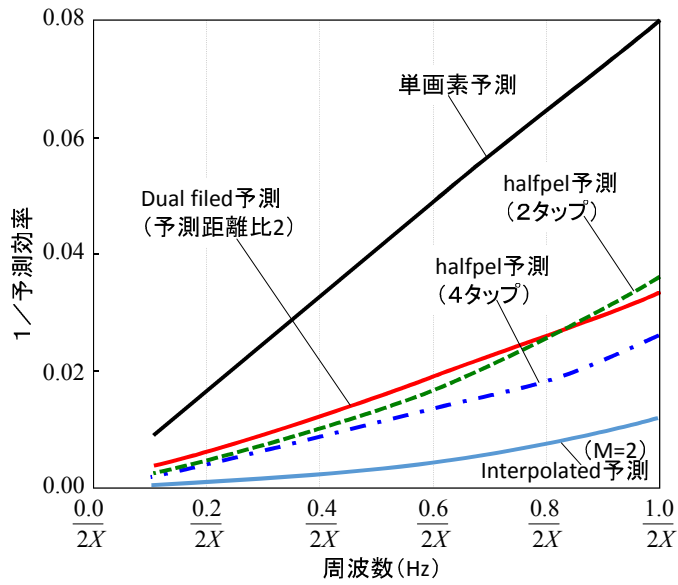


図4-5 エッジモデルを用いた予測効率

Fig.4-5 Prediction efficiency with edge model

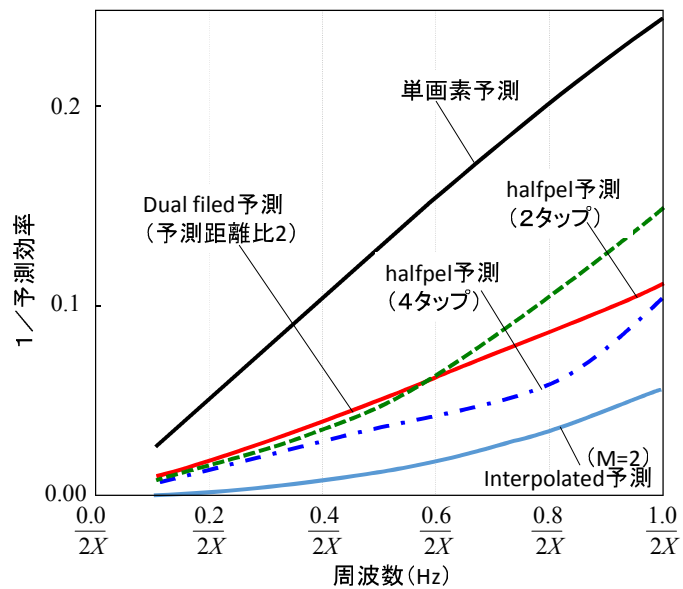


図4-6 インパルスモデルを用いた予測効率

Fig.4-6 Prediction efficiency with impulse model

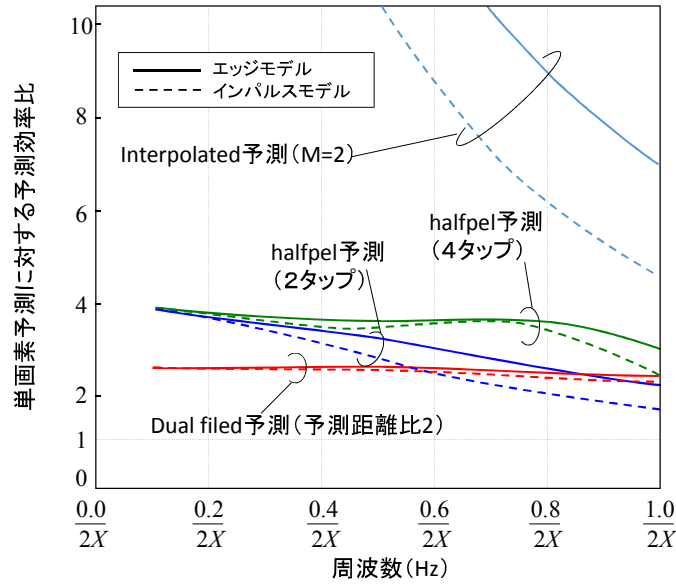


図4-7 単画素予測に対する予測効率比

Fig.4-7 Relative prediction efficiency to simple prediction

図4-8は各種のave予測の予測効率の比較である。エッジモデルを用い、 B を $\frac{0.8}{2X}$ に固定した上で、ave予測の2つの画像の予測距離比 μ を変化させた時の予測効率の逆数を示している。図では μ の絶対値が1未満の時の予測効率は、 $1/\mu$ に対応する値を表示

している。 $\mu=1$ の時がsmpl予測に相当する。予測距離比が整数比の時に極大と極小を繰り返しながら、smpl予測の誤差電力の $1/2$ に収束する(4.6節参照)。予測効率が最大になる点は、IP予測では等距離(距離比-1)、DF予測では距離比2のときである。

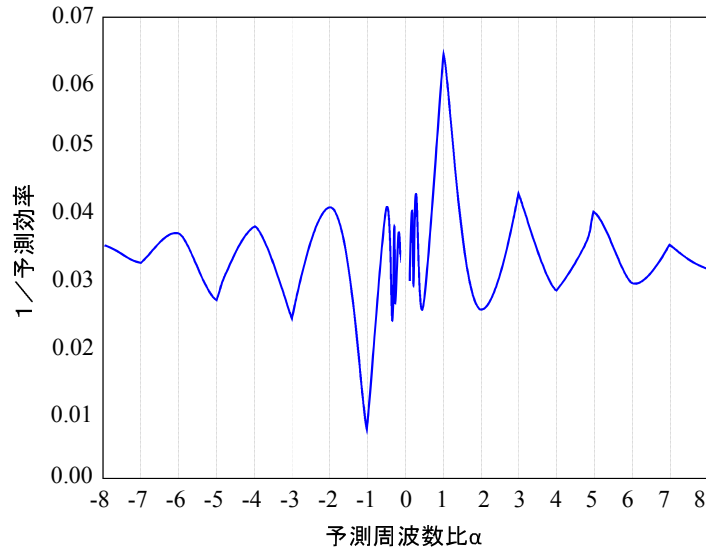


図4-8 平均値型予測の予測距離比と予測効率の関係

Fig.4-8 Prediction distance ratio of average type prediction vs. prediction efficiency

等IP予測では，被予測画像の位相に対する，2つの予測画像の位相の位相差が絶対値が同じで符号が反転しているため相殺し，他の場合に比べ予測効率は突出して高くなる．以下，予測誤差が最も大きい「信号値変化の急峻部」（図4-1においてエッジモデルでは $x = 0$ 付近の区間，インパルスモデルでは $x = \pm \frac{1}{8B}$ 付近の区間）の画素に着目することにより，この「位相差相殺」のメカニズムについて説明する．

予測距離比 μ が -1 の時には，式(4.18)より

$$\varepsilon_3 = \langle 2\varepsilon_0 - \varepsilon_1 \rangle_x \quad (4.25)$$

エッジモデルを用いて，被予測画素値を， $x_0 = \varepsilon_0$ を用いて $S(-x_0, B)$ ，第1の予測画素値を $S(-x_1, B)$ ，第2の予測画素値として， $x_3 = 2x_0 - x_1$ を用いて $S(-x_3, B)$ とする． x_3 に対応する画素は式(4.25)を満たす．ここで「信号値変化の急峻部」を直線で近似すると，エッジモデルでは

$$S(x, B) = gx + 0.5 \quad (4.26)$$

と表せる．式(4.26)を用いて予測画素は

$$\frac{S(-x_3, B) + S(-x_1, B)}{2} = \frac{-g(x_3 + x_1)}{2} + 0.5 = -gx_0 + 0.5 = S(-x_0, B) \quad (4.27)$$

となり，被予測画像に一致する．実際には(4.26)式の近似が成り立たない部分において若干の予測誤差が生じるが，他の予測方法に比べ予測効率は飛躍的に向上する．インパルスモデルでは直線部分がエッジモデルに比べ少ないため，位相の相殺による効果が少なくなり，符号化効率比はやや低くなる．DF予測および，等IP予測以外のIP予測（不等IP予測）の予測効率に関しては4.4.2を参照のこと．

4.4 実画像を用いたシミュレーション

4.4.1 シミュレーション

提案した画像モデルに比較的近い実画像を用いて，シミュレーションにて予測効率を測定した．シミュレーションの諸元を表1に示す．使用した画像は，非常に高い周波数成分を含むエッジおよびインパルスが多く，画面全体がゆっくりパンニングしている部分である．予測画像には原画を用い，被予測画像よりも広い範囲からの予測を許している．

表4-1 シミュレーション条件
Simulation specifications

シーケンス	“Mobile&Calendar” のパンニング部分（水平352 画素，垂直96画素，top field），36 field. 被予測画像は30 field
予測距離	forward : 1~4 frames backward : 1~2 frames (-1, -2と表示)
予測方法	単画素予測（整数ペル，フィールド予測） ハーフペル予測（2タップ，単画素予測適応） ハーフペル予測（4タップ，単画素予測適応） デュアルフィールド予測（整数ペル） インターポレイテッド予測（整数ペル）
ブロック サイズ	16×16，輝度信号
探索範囲	整数ペル：H：±15，V：±15，フルサーチ ハーフペル：整数ペル探索後にH：±0.5，V：±0.5 評価値：誤差電力

4.4.2 シミュレーション結果

図4-9にシミュレーション結果として，smpl予測に対する予測効率比を計算値と共に示す．hpel予測の効率比は，同じ画像を用いてsmpl予測した時の予測誤差電力に対する比，DP予測，IP予測の効率比は，予測に用いた2つの画像によるsmpl予測誤差電力の平均値に対する比を用いた．全体的には計算値とほぼ近い値あるいは，やや悪い値になっている．前提条件にも記したように，

本方法は位相成分による予測誤差を対象としており、ノイズ、変形を含まないことを前提としている。これに対し、実測値は位相成分による予測誤差の他に、ノイズや変形等（以下、総称してノイズ成分と呼ぶ）による予測誤差を含んでいる。以下、計算値と実測値との誤差について考察を行う。

2タップhpel予測では、計算値と実測値が比較的近い値となっている。しかし、予測距離が遠くなるに従い、予測誤差全体に占めるノイズ成分による予測誤差の比率が高くなる。ノイズ成分による予測誤差は、2タップhpel予測に対しても、smpl予測に対しても、ほぼ同程度になると予想されるため、実測値の予測効率比は1に近づき、計算値と実測値との差が比較的大きくなる。

DF予測あるいは不等IP予測以外の場合も同様にノイズの影響が考えられる。しかし、4.5に後述するように、ノイズの影響はhpel予測に比べて弱く、計算値と実測値が近い値となっている。

一方、等IP予測と4タップhpel予測では、位相成分による予測誤差は大幅に減少しているため、ノイズ成分による予測誤差の比率が相対的に高くなり、計算値と実測値との差が大きくなっていると考えられる。

さらに、等IP予測と4タップhpel予測では、以下に説明するように、方式的にノイズ成分による影響を受けやすいため、計算値との差が大きくなっている。等IP予測による効果は、主に平均値処理による誤差電力半減の効果と、4.2節で説明した位相差相殺による効果に分けられる。これら2つの寄与度は、位相差相殺によるものが大きい（平均値処理により誤差電力は約 $1/2$ になり、位相差相殺によりさらに約 $1/3 \sim 1/4$ 程度になっている）。

位相差を相殺するためには、被予測画像および2つの予測画像の対応する画素がそれぞれ正確な位置（時間位置および空間位置）になければならない。しかし、実画像ではサンプル位置やエッジの動く速度に若干のジッタが含まれていると考えられる。実験に用いたような高い周波数成分を含む画像では、サンプリングの点のわずかなずれでもサンプル値は大きく変化し、位相差相殺による予測誤差削減ができなくなる。この結果、予測効率は不等IP予測やDP予測の予測効率に近づく。これらの現象は図4-8において μ が -1 をはずれると急激に予測効率が低下している点にも現れている。

4タップhpel予測では、次節に示したように、他の予測に比較してノイズ成分による予測誤差電力の増加の度合いが大きい。また、位相成分による予測誤差自体も少ないため、ノイズ成分による予測誤差の比率が相対的に高くなり、比較的急に予測効率比が1付近に近づく傾向がある。

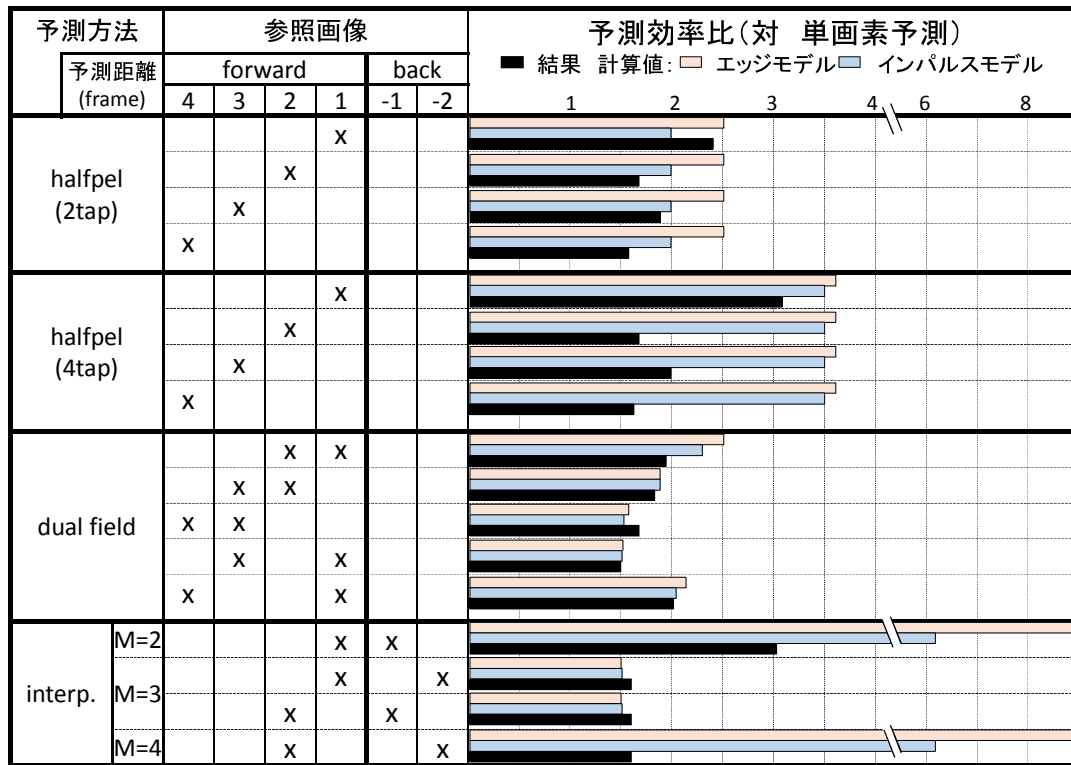


図4-9 シミュレーション結果
Fig.4-9 Results of the simulation

4.5 ノイズ・変形に対する各予測方式の耐性評価

予測に用いる参照画像にノイズあるいは変形が生じたときの予測誤差電力への影響を算出する。被予測画像を $A(i)$ 、参照画像の画素列を $B(i)$ 、参照画像の各画素へのノイズを $n_B(i)$ 、動きベクトルを0とすると、単画素予測での予測誤差電力の増分 dD_{smpl} は

$$\begin{aligned}
 dD_{smpl} &= \sum_i (B(i) + n_B(i) - A(i))^2 - \sum_i (B(i) - A(i))^2 \\
 &= \sum_i n_B(i)^2 - 2 \sum_i n_B(i)(B(i) - A(i))
 \end{aligned}
 \tag{4.28}$$

ノイズの生じている位置と予測誤差の大きい位置との間に相関がなく、式(4.28)の第2項目を省略すると

$$dD_{smpl} = \sum_i^{\infty} n_B(i)^2 \quad (4.29)$$

同様に2タップのハーフペル予測の予測誤差電力の増分は、動きベクトルの端数が0.5の場合は

$$dD_{hpel2} = \sum_i^{\infty} (C(-1)n_B(i) + C(1)n_B(i+1))^2 \quad (4.30)$$

ここで、 $C(-1)$ 、 $C(1)$ は4.2.3.3で定義したフィルタ係数である。式(4.30)を展開し、ノイズの自己相関を0と仮定すると、式(4.30)は以下のようになる。

$$dD_{hpel2} = (C(-1)^2 + C(1)^2) \sum_i^{\infty} n_B(i)^2 = \frac{1}{2} \sum_i^{\infty} n_B(i)^2 \quad (4.31)$$

また、動きベクトルの端数が0.0の場合は単画素予測と同じになるため、2タップのハーフペル予測の予測誤差電力の増分は確率的に、以下のように予測誤差電力の増分は単画素予測の75% (3/4)になる。

$$d\bar{D}_{hpel2} = \frac{1}{2} (dD_{smpl} + dD_{hpel2}) = \frac{3}{4} \sum_i^{\infty} n_B(i)^2 \quad (4.32)$$

4タップのハーフペル予測の場合は、ベクトルの端数が0.5の時、以下となる。

$$dD_{hpel4} = (C(-3)^2 + C(-1)^2 + C(1)^2 + C(3)^2) \sum_i^{\infty} n_B(i)^2 = \frac{52}{64} \sum_i^{\infty} n_B(i)^2 \quad (4.33)$$

ベクトルの端数が0.0の場合も考慮に入れると

$$d\bar{D}_{hpel4} = \frac{1}{2} (dD_{smpl} + dD_{hpel4}) = \frac{58}{64} \sum_i^{\infty} n_B(i)^2 \quad (4.34)$$

となり、予測誤差電力の増分は単画素予測の90%になり、2タップの場合よりも悪くなる。

平均値型予測の場合は、第二の参照画像のノイズを $n_D(i)$ とし、その電力が $n_B(i)$ の電力と同じでかつ、それぞれの間に相関がないと仮定すると

$$dD_{dp} = \sum_i \left(\frac{n_B(i) + n_D(i)}{2} \right)^2 = \frac{1}{2} \sum_i n_B(i)^2$$

(4.35)

となり，最も予測誤差電力の増分は少なくなる．

4.6 簡略手法による平均値型予測の予測効率の比較

平均値型予測では，予測距離比によって位相差相殺の効果が異なり，予測効率が異なる．本節では，位相差相殺の効果（特に図4-8の特性）を容易に理解できるように，簡単な手法を用いて視覚的に説明する．

本簡略手法では，予測誤差が最も大きい「信号値変化の急峻部」の画素にのみ着目し，その画素の予測誤差を算出する．なお，該当画素の周辺では信号値は式(4.26)で近似できるものとする．

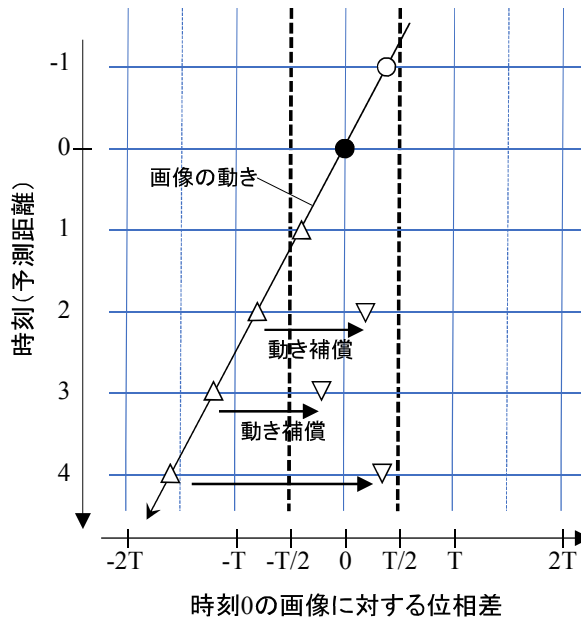


図4-10 予測距離ごとの位相差の関係

Fig.4-10 Dependence of differential phase in each prediction distance

例えば，画像の動きが図4-10の中の斜めの線の方角とし，被予測画像の位相（図4-10●印）を基準位置とすると，各時刻の予測画像の位相と，被予測画像の位相との位相差は図4-10の○や△印の値となる．第1の予測画像を時刻-1の画像とし，第2の予測画像を適当な時刻（>0）から選択する．もし，選択した位相差が $[-T/2, T/2]$ を越えた場合は，動きベクトルによって $[-T/2, T/2]$ に入るように動き補償する（▽印）．このとき，(4.27)式にも示されているように，第2の予測画像の位相差（位相差2）と第1の予測画像位相差（位相差1）の和が0になるときにIP予測の予測誤差

が0になる。

位相差1を $-T/2$ から $T/2$ まで変化させたときの位相差2の負数を図4-11 (1) ~ (3) に示す。図中の点線は位相差1に相当し、図の陰をつけた部分が予測誤差に相当する。点線と直線の交点では予測誤差が0となる。 $T=1$ として誤差電力を計算すると、(1) 0.00, (2) 0.29, (3) 0.11となり、以下、 μ の絶対値が増加するに従い、 μ が奇数では極大、偶数では極小を繰り返しながら0.166...に収束する。

DF予測の場合は、時刻1の画像を第1の予測画像とし、位相差2は図4-11 (4) ~ (6) のようになる。(4) はsmpl予測に相当し、誤差電力0.33。以下、(5) 0.13, (6) 0.22となり、IP予測とは逆に μ が奇数の時に予測誤差が偶数の時よりも多くなる。

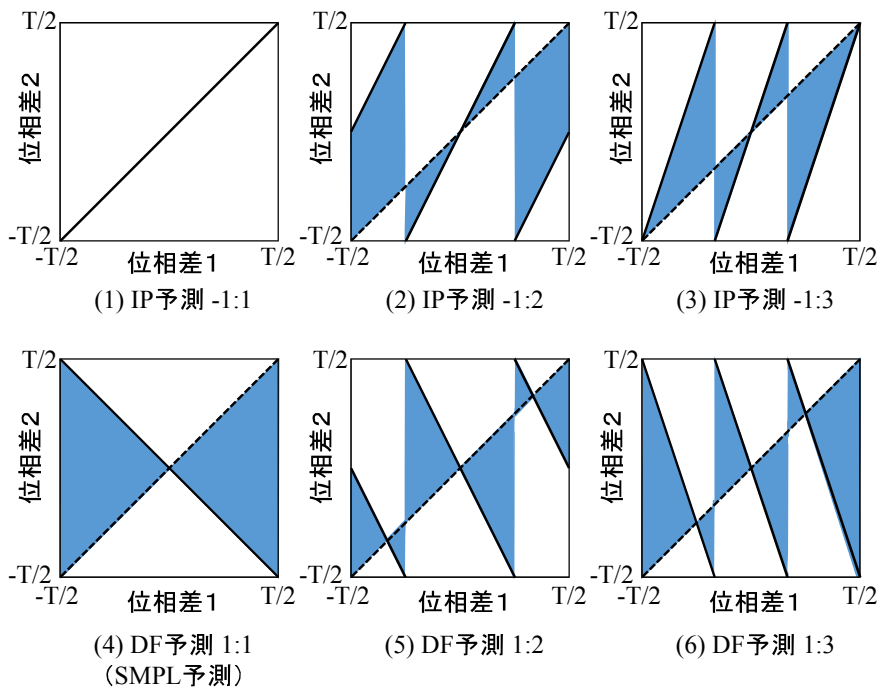


図4-11 図4-10のモデルによる予測誤差

Fig.4-11 Simple examples of prediction error using the model illustrated in Fig. 4-10

4.7 むすび

動き補償フレーム間予測の符号化効率の評価方法として、サンプリング点と画像位置との位相成分のずれに着目した画像モデルを用いた評価方法を提案した。

まず、画像モデルとMC方式モデルを提案した。画像モデルは、MCの予測誤差が主に画像内の物体の輪郭（エッジ）周辺に多く発生することに注目して、ステップ関数（エッジ）およびインパルス関数をベースに、帯域制限の周波数と、サンプリング点との位相差をパラメータとした関

数とした。MC方式モデルとして、単画素予測、ハーフペル予測、平均値型予測の3つの予測値生成方式をモデルとして定式化した。

次に、各画像モデルを各MC方式モデルに適用し、画像モデルのサンプリング点との位相差を確率変数と考え、予測誤差電力の期待値を計算することにより、各MC方式の予測効率（誤差電力期待値の逆数）を算出した。特に、平均値型予測では平均値処理により誤差電力が半減するだけでなく、2つの画像の予測距離比に応じて位相成分を相殺する効果が現れ、予測効率が高くなることを示した。この現象は前後等距離から予測するIP予測において最も顕著になる。

評価結果は、実画像を用いた予測効率の実測結果と比較した。実画像に含まれるノイズや変形の影響により若干の誤差はあるものの、ほぼ計算値に近い結果が得られ、提案した方式の妥当性が示された。

本章にて提案した方法より、画像モデルを用いて方式検討を行うことが可能であることが示された。画像モデルを用いて方式検討を行うことが実現することにより、以下の効果が期待できる。

- 「時間短縮」：方式検討の検討期間を短縮できる。多数の評価画像を用いて処理を行う必要がないため、短時間で方式の基礎検討を実施することができる。
 - 「方式選択早期化」：方式検討の初期の段階で多数の方式候補の中からの絞り込みが容易にできる。評価画像を用いた方法の場合、画像のいずれの特性が評価結果に影響しているかが不明確であるため、方式の取舍選択の判断を誤る場合がある。画像モデルを用いることにより、その画像モデルで代表される特性に対する方式の適性が明確に判断できるため、特性の悪い方式等を早い段階で排除することができる。
 - 「特性把握」：選択した方式の特性（得失）を容易に把握することができ、早い時点から対策等を検討することができる。例えば、MCにて高周波成分を含む画像が不得意な点などを初期段階で把握することができるため、早い時点からその対策（例えば、フィルタをかける、他の方式との適用処理等）を検討することが可能である。
 - 「定量評価」：評価画像の特性によらない方式の定量的な評価を行うことができる。
- これらの効果が、実際の方式検討において得られることを検証することが今後の課題である。

第5章

結論

5.1 まとめ

本論文では、画像のエッジ位置とサンプリング位置の位相差に着目し、画像モデルを用いることにより、動画像のMV検出、およびMC方式の予測効率の検討を行った。

第1章では、研究の背景と目的を述べるとともに、本研究の位置づけを明確にした。あわせて、これまでの関連研究の概要と課題を説明した。

第2章、第3章では、主に繰り返しパターンを多く含む動画像を中心に、被写体あるいはカメラの動きを表す「正確なMV」を検出する方式を提案した。1次元のエッジのモデルを用いて、MV検出誤りの原因が、MVの量子化誤差成分により生じるブロックマッチング誤差（DVSE：block matching Distortion caused by motion Vector Sampling Error）であることを明らかにし、これを基に、第2章では、MVの量子化誤差の影響を低減するMV検出方法を、第3章では、DVSEの最大値を利用したMV検出方法を提案した。

第2章では、MVを検出する2つのフレームの前後2～6フレームを利用して、時間方向のフィルタ処理を行って参照画像を生成する評価関数Temporally-Predicted Difference (TPD)を提案した。TPDを用いることにより、MVの量子化誤差の影響が相殺され、正確なMVを求めることができる。また、利用するフレーム数（タップ数）の異なる2つのTPDを適応的に選択する方法により、時間方向に不規則な変化がある映像に対しても提案方法が効果的に適用できる。さらに、MVの空間方向の連続性を利用する方法と組み合わせる方法を提案し、実画像を用いたシミュレーションにより効果を検証した。予め算出した正解ベクトルを用いて、MV検出誤り率を指標として評価を行い、TPDを用いたMV検出方法が他の方法よりも優れていることを示した。あわせて、ソフトウェアの処理時間より処理量の評価を行い、実用的な処理量であることを示した。

第3章では、まず、MV検出誤りの原因であるDVSEの最大値 $\max DVSE$ を算出する手法を示した。次に、正確なMVのブロックマッチング（BM）誤差が $\max DVSE$ 以下である点に着目し、ブロック単位にBM誤差が $\max DVSE$ 以下のMVをそのブロックのMV候補として全て記憶しておき、1フレーム全てのBM処理が完了した後に、記憶しておいたMVから各ブロックのMVを確定させてゆく。各ブロックのMV確定においては、MVが一意的に定まるアンカーブロックのMV（アンカーMV）とのMVの連続性を用いる。提案した方法は、実画像を用いたシミュレーションにより、MV検出誤り率を算出して評価を行った。提案方法はフルサーチと同程度の処理量にて、他の方式よりもMV検出誤り率を低減することを示した。あわせて、ノイズ画像への耐性、ブロックサイズを変えた時の影響、提案方法の検出誤りの箇所の評価などを分析し考察した。

第4章では、画像モデルとMC方式モデルを提案し、これらを用いて各MC方式の予測効率を

求める方法を提案した。画像モデルは、MCの予測誤差が主に画像内の物体の輪郭（エッジ）周辺に多く発生することに注目して、ステップ関数（エッジ）およびインパルス関数をベースにした関数とし、帯域制限の周波数と、サンプリング点との位相差をパラメータとした。MC方式モデルは、単画素予測、ハーフペル予測、平均値型予測の3つの予測値生成方式をモデルとして定式化した。次に、各画像モデルを各MC方式モデルに適用し、各MC方式の予測効率（誤差電力期待値の逆数）を計算する方法を提案した。画像モデルのサンプリング点との位相差を確率変数と考え、各MC方式の予測誤差電力の期待値を計算する。算出された予測効率は、実画像を用いた予測効率の実測結果と比較実験を実施した。実画像に含まれるノイズや変形の影響により若干の誤差はあるものの、ほぼ計算値に近い結果が得られ、提案した方式の妥当性が示された。

5.2 今後の課題

本提案では、平坦なブロックや、歪の大きいブロックを除いた、MVが一意に定まるブロックを対象としている。回転やズームの動きを含む画像における性能は一部確認しているが、回転部や、変形を伴う動きに対する性能の確認および、これらに対応する既存の方式との組み合わせの検討は今後の課題である。また、小さなサイズのブロック、ノイズを含む画像において高いMV検出性能を安定的に得るための検討も今後の課題である。また、各章の個別の課題としては以下がある。

第2章の時間方向のフィルタに関しては、フィルタ係数を入力画像から学習し最適化する方法もある。提案方法との比較が今後の課題である。第3章においては、小さなサイズのブロックの場合を含み、アンカーブロックを安定的に求めることが今後の課題である。また、可変ブロックサイズへの対応は今後の課題である。第4章においては、方式検討の「時間短縮」、「方式選択早期化」、「特性把握」、「定量評価」の効果が、実際の方式検討において得られることを検証することが今後の課題である。

また、画像のエッジとサンプリング位置の位相差の関係に着目することにより、他の画像処理方式における新たな方式を創出し、これを用いた新たなアプリケーションを検討していくことが今後の大きな課題である。

5.3 期待される効果

第2章、第3章の成果により、正確なMV検出を必要とするアプリケーションにおいて、より少ない処理量にて、高画質な処理を実現することが期待できる。処理量が少なくなることにより、低電力の処理が可能となり、モバイル機器を含めて、より多くの機器にMV検出処理を搭載することが可能となる。

第4章にて提案した方法により、画像モデルを用いて方式検討を行うことが可能であることが示された。画像モデルを用いて方式検討を行うことが実現することにより、方式検討の「時間短

縮」,「方式選択早期化」,「特性把握」,「定量評価」が容易となる. この結果, 方式検討から最終的な製品の評価に至るまでの工数が低減され, 短期間に効率的に優れた性能の方式を検討することが可能となる.

以上のように, 本論文にて提案した方法により, 将来, より多くの機器に, より短期の検討期間にて, **MV**検出を用いた動画像アプリケーションを実装することが可能となり, こうしたアプリケーション, サービス, コンテンツがより一層普及することが期待できる.

また, これまで, **MV**検出, あるいは**MC**処理において, 画像のエッジとサンプリング位置の位相差の関係が着目されることは少なく, 具体的な検討もあまり行われていなかった. 今後, 画像のエッジとサンプリング位置の位相差の関係に着目することにより, 他の画像処理方式や, アプリケーションにおいて新たな方式が創出されることが期待される.

謝辞

本論文を纏めるにあたり、ご指導、ご助言、ご支援を賜りました全ての方々に深く御礼を申し上げます。

懇切丁寧な御指導、御助言を賜りました早稲田大学 甲藤二郎教授に心から感謝の意を表します。また、広島大学 児玉明准教授には、論文投稿・学会発表および学位論文審査において多くの御指導、御助言を賜りました。心より御礼申し上げます。

学位論文審査において、貴重なご指導とご助言を頂きました早稲田大学 渡辺裕教授、同学亀山渉教授に心より感謝申し上げます。

ゼミや日常の場において議論をさせていただくとともに、研究環境構築を含め研究生活のさまざまな点においてお世話になりました早稲田大学小松尚久研究室卒業生の大木哲史氏、市野政嗣氏、披田野清良氏、鶴丸和宏氏、星健太郎氏に深く感謝の意を表します。甲藤二郎研究室の社会人博士課程卒業生の坂本静生氏、松尾康孝氏、泉川晴紀氏には毎月行われた社会人博士課程ゼミにおきまして、様々な意見をいただき、大変有益な議論をさせていただきました。深く感謝いたします。また、早稲田大学小松尚久研究室卒業生、同学甲藤二郎研究室卒業生・現役生の皆様に深く感謝いたします。

早稲田大学 富永英義名誉教授には学部、修士課程において、元（株）日立製作所中央研究所吹抜敬彦技師長、元（株）日立製作所中央研究所 滝沢正明氏には（株）日立製作所中央研究所入社時より、それぞれ、研究の基礎をご指導いただきました。これらの指導があったからこそ、本論文をまとめることができたと考えております。心より御礼申し上げます。

最後に、2011年にご逝去された元早稲田大学 小松尚久教授には、博士課程入学時からご指導、ご助言を賜り、また、本研究の全体の方向性をご指導いただきました。感謝の念に堪えません。完成した本論文をご覧いただけないことが、ただただ残念でなりません。謹んでご冥福をお祈りいたします。

2015 年 2 月

参考文献

- [1] ISO/IEC, "Information technology generic coding of moving pictures and associated audio (MPEG-2)," ISO/IEC 13818-2 (1995)
- [2] ISO/IEC, "Information technology - coding of audio-visual objects- Part 10: Advanced video coding (H.264)," ISO/IEC 14496-10 (2004)
- [3] 川田亮一, 浜田高宏, 松本修一, “動き補正テレビ方式変換の改善”, 映情学誌, vol.50, no.9, pp.1577-1586 (Sep. 1997)
- [4] 田中豊, 大村俊郎, 岡田清孝ほか, “動き補正型フレーム数変換法を用いたHDTV-PAL 方式変換装置の構成と画質評価”, 信学論 (D), vol.J70-D, no.8, pp.1535-1545 (Aug. 1987)
- [5] 藤原慎矢, 田口亮, “ブロックマッチング動き補償に基づくフレーム間補間法”, 信学論 (A), vol.J90-A, no.5, pp.477-483 (May 2007)
- [6] 平賀督基, “組込み向けイメージングテクノロジーのライセンスビジネスで成長し続けるモルフォ〜手ブレ補正技術から被写体自動認識技術まで〜”, 映情学誌. Vol.64, No. 1, pp.76-80 (2010年1月)
- [7] 吉田育弘, 藤根俊之, 山本健一郎ほか, “大型LCD-TV用倍速フレームレート変換技術とその画質改善効果の検証”, 映情学誌, vol.62, no.5, pp.778-787 (May 2008)
- [8] C.N. Cordes and G.de Haan, “Key requirements for high quality picture-rate conversion,” Digest of the SID'09, vol. XV, Book II, pp.850-853 (May 2009)
- [9] 山内日美生, “液晶テレビの動画表示性能を改善する FrameBooster”, 東芝レビュー vol.64, no.6, pp.15-18 (Jun. 2009)
- [10] O. A. Omer and T. Tanaka, “Extraction of high-resolution frame from low-resolution video sequence using region-based motion estimation,” IEICE Trans. Fundamentals, Vol.E93-A, no.4, pp.742-751 (Apr. 2010)
- [11] S. C. Park, M. K. Park, and M. G. kang, “Super-resolution image reconstruction: a technical overview,” IEEE Signal Processing Magazine, vol. 20, no. 3, pp. 21-36 (May 2003)
- [12] B.D. Lucas and T. Kanade, “An Iterative Image Registration Technique with an Application to Stereo Vision,” Proceedings of the international joint conference on artificial intelligence, pp.674-679 (1981)
- [13] ITU-T, “Recommendation H261 Video Codec for Audiovisual Services at $p \times 64$ kbit/s” (1990)
- [14] L.M. Po and W.C. Ma, “A novel four-step search algorithm for fast block motion estimation,” IEEE

- Trans. on CSVT, Vol. 6, No. 3, pp.313-317 (Jun. 1996)
- [15] A.M. Tourapis, O.C. Au, and M.L. Liou, "Fast Motion Estimation using Circular Zonal Search," ISCAS-2001, Vol. 5, pp.183-186 (May 2001)
- [16] C. Zhu, X. Lin, and L.P. Chau, "Hexagon-based search pattern for fast block motion estimation", IEEE Trans. CSVT, Vol. 12, No.5, pp.349-355 (May 2002)
- [17] Z. Chen, J. Xu, Y. He, and J. Zheng, "Fast integer-pel and fractional-pel motion estimation for H.264/AVC," Visual Communication and Image Representation, Vol. 17, No.2, pp.264-290 (Oct. 2005)
- [18] L. Xu, J. Jia, and Y. Matsushita, "Motion detail preserving optical flow estimation," IEEE Trans. on PAMI, vol.34, no.9, pp.1744-1757 (Sep. 2012)
- [19] 木村淳一, 小松 尚久, 甲藤二郎, 児玉明, "複数フレームを用いたブロックマッチングによる繰り返しパターンにロバストな動きベクトル検出", 信学論 (D), Vol. J97-D, No.9, pp.1361-1373 (2014年9月)
- [20] 木村淳一, 小松尚久, 甲藤二郎, 児玉明, "複数フレームを用いたブロックマッチング動きベクトル検出精度向上手法に関する検討", 2013画電年大, R3-2 (Jun. 2013)
- [21] Jun-ichi Kimura, Naohisa Komatsu, Jiro Katto, and Mei Kodama, "A Study on Motion Estimation using Block Matching for Spatially Periodic Patterns," ITE Trans. on MTA, Vol. 3, No. 1, pp.58-66 (Jan. 2015)
- [22] 木村淳一, 小松 尚久, 甲藤二郎, 児玉明, "繰り返しパターンを持つ画像におけるブロックマッチング動きベクトル検出の誤り解析と検出精度向上手法に関する検討", 2014画電年大, R1-2 (Jun. 2014)
- [23] J. Kimura, and N. Komatsu, "Accurate Motion Estimation for Image of Spatial Periodic Pattern," 28th Picture Coding Symposium (PCS 2010), P2-24, pp.250-253 (Dec. 2010)
- [24] 木村淳一, 松本孝浩, "画像モデルを用いた動き補償フレーム間予測効率の検討", テレビジョン学会誌 Vol.50, No.7, pp.965~973 (1996年7月)
- [25] 木村淳一, 松本孝浩, "ランダム変数を用いた動画像符号化予測効率の検討", PCSJ93, 4. 13 (Oct. 1993)
- [26] 木村淳一, 松本孝浩, "動き補償予測効率の検討", テレビ学技報, Vol.18, No.46, pp.7-12 (Aug. 1994)
- [27] J. Kimura and T. Matsumoto, "A Consideration on the Prediction Efficiency of Motion Compensation using an Edge Model," Int. Conference on Communication Technology (ICCT'94), 57.06 (June 1994)
- [28] D.G. Lowe, "Distinctive image features from scale-invariant keypoints," Int. J. of Computer Vision (IJCV), vol.60, no2, pp.91-110 (Jan. 2004)

- [29] V.N. Dvorchenko, "Bounds on (deterministic) correlation function with applications to registration," IEEE Trans. on PAMI, Vol.5, No.2, pp.206-213 (Mar. 1983)
- [30] 清水雅夫, 奥富正敏, "領域ベースマッチングのための2次元同時サブピクセル推定法," 信学論(D-II), vol.J87-D-II, no.2, pp.554-564 (Feb. 2004)
- [31] S. Birchfield and C. Tomasi, "A pixel dissimilarity measure that is insensitive to image sampling," IEEE Trans. on PAMI, 20(4), pp.401-406 (Apr. 1998)
- [32] C. Bartels and G.de Haan, "Smoothness constraints in recursive search motion estimation for picture rate conversion," IEEE Trans. on CSVT, vol.20, no.10, pp.1310-1319 (Oct. 2010)
- [33] "ISO/IEC 11172-2 Information Technology - Coding of Pictures and Associated Audio for Digital Storage Media at up to about 1.5Mbit/s" (Nov. 1992)
- [34] "Test Model 5", ISO/IEC JTCl/SC29/WG11/N0400 (Apr. 1993)
- [35] M. Okutomi and T. Kanade, "A Multiple-Baseline Stereo," IEEE Trans. on PAMI, vol.15, no.4, pp.353-363, Apr. 1993.
- [36] B. Horn, and B. Schunck, "Determining Optical Flow," Artificial Intelligence, vol.17, pp.185-203 (Aug. 1981)
- [37] ISO/IEC, "14496-10 AVC Reference Software (JM18.6) ," <http://iphome.hhi.de/suehring/tml/> (2014)
- [38] G.de Haan, P. W.A.C. Biezen, H. Huijgen, and O. Ojo, "A True-Motion Estimation with 3-D Recursive Search Block Matching," IEEE trans on CSVT, vol. 3, no. 5, pp.368-379 (May 1993)
- [39] 富永英義, 小松尚久, 宮下壮史, 花村剛, "階層画素情報を用いた動画像における動き量検出方式", 信学論(D-II), Vol. J72-D2, No.3, pp.395-403 (1989年3月)
- [40] J. Bouguet, "Pyramidal implementation of the Lucas-Kanade feature tracker: description of the algorithm," Technical report, OpenCV Document (2000)
- [41] T. Yoshida, H. Katoh, and Y. Sakai, "Block matching motion estimation using block integration based on reliability metric," 1997 Int. Conference on Image Processing (ICIP'97), vol.2, pp.152-155 (Oct. 1997)
- [42] 吉田俊之, 宮本敦司, 酒井義則, "動画像の動きベクトルに対する信頼度関数とその応用", 信学論(D-II), Vol.J80-D2, No.5, pp.1192-1201, (1997年5月)
- [43] 坂本大樹, 金澤靖, "パターン特徴量空間を用いた繰り返しパターンを含む画像間の対応付け", 信学技報 PRMU2009-52, pp.77-82 (Jun. 2009)
- [44] 秦泉寺久美, 石橋聡, 小林直樹, "カメラモーション抽出によるスプライト自動生成", 信学論(D-II), Vol. J82-

- D2, No.6, pp.1018-1030 (1999年6月)
- [45] 秦泉寺久美, 渡辺裕, 小林直樹, “スプライト生成のためのグローバルモーション算出法と符号化への適用”, 信学論(D-II), Vol. J83-D2, No.2, pp.535-544 (2000年2月)
- [46] 菅谷保之, 金谷健一, “画像の三次元理解のための最適化計算[III]—基礎行列の計算—”, 信学誌, Vol. 92, No. 6, pp.463-468 (2009年6月)
- [47] M.A. Fischler and R.C. Bolles, “Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography,” *Comm. ACM*, Vol.24, No.6, pp.381–395 (Jun. 1981)
- [48] 金澤靖、金谷健一, “大域的な整合性を保証するロバストな画像の対応づけ”, 情報処理学会論文誌コンピュータビジョンとイメージメディア, Vol. 44, No. SIG17 (CVIM8), pp.70-77 (2003年12月)
- [49] S. Baker, D. Scharstein, J. P. Lewis, S. Roth, M. Black, and R. Szeliski, “A database and evaluation methodology for optical flow,” *Proc. 13th Int. Conf. Computer Vision*, pp. 1–8 (Oct. 2007) (<http://vision.middlebury.edu/flow/>)
- [50] 中嶋満雄, “大型FPDテレビ向け映像信号処理技術,” 日立評論, vol.89, no.10, pp.778-779 (Oct. 2007)
- [51] 映像情報メディア学会, “テレビジョン・システム評価用テストチャート「DVD版システム評価用標準動画像シリーズ」(SIF)”
- [52] 井口和久, 曾根原源, 野尻裕司, “方式変換におけるアンカバー処理の検討,” 信学技報, vol.95, no.504, IE95-132, pp.81-88 (1996年2月)
- [53] S. Okubo, “Reference Model Methodology—A Tool for the Collaborative Creation of Video Coding Standards,” *Proceedings of the IEEE*, Vol. 83, No. 2, pp. 139–150 (Feb. 1995)
- [54] MPEG simulation model editorial group, “MPEG Video Simulation Model Three(SM3) ,” Doc. ISO/IEC JTC1/SC29/WG11 N0010 (Jul. 1990)
- [55] MPEG Video Group, “MPEG-4 video verification model 2.0,” Doc. ISO/IEC JTC1/SC29/WG11 N1260 (Mar. 1996)
- [56] “HEVC Test Model (H+M) Documentation,” <http://hevc.hhi.fraunhofer.de/HM-doc/> (2011)
- [57] T. Hidaka, and K. Ozawa, "ISO/IEC JTC1 SC29/WG11; Report on MPEG-2 subjective assessment at Kurihama," *Signal Processing: Image Communication*, Vol. 5, No. 1–2, pp.127-157 (Feb 1993)
- [58] 甲藤二郎, “オーバーラップ動き補償の特性解析”, 信学技報, IE94–8, DSP94–8 (1994年4月)

早稲田大学 博士（工学） 学位申請 研究業績書

(2015年 1月 現在)

種類別	題名、発表・発行掲載誌名、発表・発行年月、連名者（申請者含む）
○論文	A study on motion estimation using block matching for spatially periodic patterns ITE Trans. on MTA, Vol. 3, No. 1, pp.58-66 (Jan. 2015) <u>Jun-ichi Kimura</u> , Naohisa Komatsu, Jiro Katto, Mei Kodama
○論文	複数フレームを用いたブロックマッチングによる繰り返しパターンにロバストな動きベクトル検出 電子情報通信学会論文誌 D, Vol. J97-D, No. 9, pp. 1361-1373 (2014年9月) <u>木村淳一</u> , 小松 尚久, 甲藤二郎, 児玉明
○論文	画像モデルを用いた動き補償フレーム間予測効率の検討 テレビジョン学会誌, Vol. 50, No. 7, pp. 965~973 (1996年7月) <u>木村淳一</u> , 松本孝浩
講演	繰り返しパターンを持つ画像におけるブロックマッチング動きベクトル検出の誤り解析と検出精度向上手法に関する検討 2014画像電子学会年次大会, R1-2 (2014年6月) <u>木村淳一</u> , 小松 尚久, 甲藤二郎, 児玉明
講演	複数フレームを用いたブロックマッチング動きベクトル検出精度向上手法に関する検討 2013画像電子学会年次大会, R3-2 (2013年6月) <u>木村淳一</u> , 小松 尚久, 甲藤二郎, 児玉明
○国際 会議	Accurate motion estimation for image of spatial periodic pattern 28th Picture Coding Symposium (PCS 2010), P2-24, pp.250-253 (Dec. 2010) <u>Jun-ichi Kimura</u> , Naohisa Komatsu, Mei Kodama
講演	The Requirements for HDD for Consumer Electronics Products 情報ストレージ研究推進機構 (SRC) 第19回技術報告会 (2005年6月) <u>木村淳一</u>
講演	パネルディスクカッション 携帯情報端末 (IMT-2000等) 用 MPEG4-LSI 技術 電子情報通信学会 技術研究報告 ICD2000-59, pp. 97-101 (2000年8月) 榎本忠儀, 吉本雅彦, <u>木村淳一</u> , 他
国際 会議	Perceived quality and bandwidth characterization of layered MPEG-2 video encoding Proc. SPIE 3845, Multimedia Systems and Applications II, 3845-37, pp.308-319 (Nov. 1999) <u>Jun-ichi Kimura</u> , Fouad A. Tobagi, Jose-miguel Pulido, Peder J. Emstad
講演	領域分割符号化における画質制御法の検討 1996年画像符号化シンポジウム (PCSJ96), P-2. 5, pp. 21-22 (1996年10月) <u>木村淳一</u>

早稲田大学 博士（工学） 学位申請 研究業績書

種類別	題名、発表・発行掲載誌名、発表・発行年月、連名者（申請者含む）
講演	領域分割符号化における領域境界の表現法に関する一検討 電子情報通信学会総合大会 情報・システム(2), D-219, p. 7 (1996年3月) <u>木村淳一</u>
国際 会議	A study on contour representation method in gray level block truncated coding 1996 Picture Coding Symposium (PCS96) (Mar. 1996) <u>Jun-ichi Kimura</u>
講演	ブロック内領域分割による画像符号化の検討 1995年画像符号化シンポジウム(PCSJ95), 3-7, pp. 39-40 (1995年10月) <u>木村淳一</u>
講演	Dual-prime予測の動きベクトル検出法に関する検討 電子情報通信学会総合大会 情報・システム(2), D-325, p. 51 (1995年3月) <u>木村淳一</u> , 小林孝之
講演	動き補償予測効率の検討 テレビジョン学会技術報告, Vol. 18, No. 46, pp. 7-12 (1994年8月) <u>木村淳一</u> , 松本孝浩
国際 会議	A Consideration on the prediction efficiency of motion compensation using an edge model 1994 Int. Conf. on Communication Technology (ICCT'94), 57.06 (Jun. 1994) <u>Jun-ichi Kimura</u> , Takahiro Matsumoto
講演	ランダム変数を用いた動画画像符号化予測効率の検討 1993年画像符号化シンポジウム(PCSJ93), 4. 13, pp. 99-100 (1993年10月) <u>木村淳一</u> , 松本孝浩
講演	低遅延画像符号化方式に関する検討 1992年画像符号化シンポジウム(PCSJ92), 2. 2, pp. 49-52 (1992年10月) <u>木村淳一</u> , 松本孝浩, 中橋知子
講演	私はTV電話を家庭に普及させたい—ユーザ・メーカ・両側から見たTV電話の将来— 電子情報通信学会 第6回通信サービスシンポジウム (1992年4月) <u>木村淳一</u>
講演	DCT係数を用いた動きベクトル探索方式の特性評価 電子情報通信学会 92年 春季全国大会, D-297 (1992年3月) <u>木村淳一</u> , 滝沢正明
講演	ハイブリッドアーキテクチャを用いた小型画像CODEC テレビジョン学会技術報告, Vol. 15, No. 49, pp. 55-62 (1991年9月) <u>木村 淳一</u> , 滝沢 正明, 志村 隆則, 大熊 克己, 兼岡 則幸

早稲田大学 博士（工学） 学位申請 研究業績書

種類別	題名、発表・発行掲載誌名、発表・発行年月、連名者（申請者含む）
講演	DCT係数を評価関数に用いた動きベクトル探索方式の検討 電子情報通信学会 91年秋期全国大会, D-117 (1991年9月) <u>木村淳一</u> , 滝沢正明
講演	卓上オールインワン動画像TV電話の開発 —小型画像CODECの開発— 電子情報通信学会 91年 春季全国大会, D-420 (1991年3月) <u>木村淳一</u> , 志村隆則, 滝沢正明, 井澤裕司, 渡辺浩巳, 他
国際 会議	Image compression using truncated block DCT 1990 Picture Coding Symposium(PCS90), 13.1 (Mar. 1990) <u>Jun-ichi Kimura</u>
講演	ブロック歪除去に関する検討 1989年画像符号化シンポジウム(PSCJ89), 4.3, pp. 49-50 (1989年10月) <u>木村淳一</u> 井澤裕司 柴田巧一 滝沢正明
講演	中国語漢字パターンデータベースの作成 1986年情報学シンポジウム (1986年1月) <u>木村淳一</u> , 富永英義
	講演 他 23件 (連名)
特許	Image encoding and decoding apparatus US patent 5,701,160 <u>Junichi Kimura</u> , Taizo Kinoshita
	特許 他 主筆 40件以上, 連名 100件以上
標準化 寄書	Fujitsu-Hitachi proposal for MPEG2 ISO-IEC/JTC1/SC29/WG11 MPEG 1991-150 (Nov. 1991) Tadashi Saitoh, <u>Jun-ichi Kimura</u> , Kiichi Matsuda 他
標準化 寄書	Scene change handling without picture skipping in low delay mode CCITT SG XV Working Party XV/1 Experts Group for ATM Video Coding AVC-328 (Sep. 1992) Japan (<u>Jun-ichi Kimura</u>)
標準化	Requirements document iVDR(Information Versatile Disk for Removable usage) Consortium Service working group (2005) iVDR Consortium Service working group (<u>Jun-ichi Kimura</u> : Director of iVDR Consortium, Chairperson of Service working group)
	標準化寄書 他多数