

3段階動画像動き検出法の開発

Three-step Motion Vector Estimation Algorithm

情報工学専攻

高瀬 裕介

Yusuke Takase

あらまし 動きベクトル検出処理 (Motion Vector Estimation : ME) のためのブロックマッチング (BM) 処理量を大幅に削減する「高速動画像動きベクトル検出アルゴリズム」を開発した。探索領域 (Search Area : SA) を水平方向と垂直方向に探索し、次に検出された最小差分絶対値和 (D_s) をしきい値と比較し, SA を変更して BM を行う高速動画像動きベクトル検出アルゴリズムを開発した。本アルゴリズムをテスト画像 “Susie” と “Whale-Show” (MP@ML, 4 Mbps, 30 fps) に適用し, 画質, BM 処理量を検討した。SA サイズ $[-p_1, p_1]$ が $[-64$ 画素, 64 画素] のとき全探索法の画質を維持しつつ, 平均 ME 速度 (平均 F_s) を全探索法 (Full Search : FS) では前者が 102.6 倍, 後者が 101.38 倍, 3階層サブサンプリング法 (3-SS); 高画質 (HQ)[2] では前者が 8.17 倍, 後者が 8.83 倍, 3-SS 法, 高速 (Faster)[2] では前者が 4.31 倍, 4.92 倍に高速化できた。

キーワード MPEG2, 動きベクトル検出, 探索領域サイズ, サブサンプリング

1. はじめに

地上デジタル放送が始まり, HDTV (high definition television, MP@HL, 1920 画素 \times 1080 ライン) ならびに HDTV 向けとして Blue-ray Disc のような次世代光ディスクが急速に普及すると予想される。これに伴い, データ量の極めて多い HDTV の符号化, 特に動きベクトル検出処理 (Motion Vector Estimation : ME) の高速化技術の開発が急務である。同時に, 符号化 LSI の低消費電力化が要求され, ME 演算量の削減が求められる。代表的な ME アルゴリズムの一つである, 全探索法 (Full Search : FS) は, 高精度に動きベクトルを検出できるが, 探索領域 (Search Area : SA) 内の全マクロブロック (Macro Block : M-Blk) に対して差分絶対値和 (D_s) を求めるため, ブロックマッチング数 (M_b) が極めて多く, HDTV 等の应用到に不適當である。一方, M_b を大幅に削減する方法としてサブサンプリング (Sub-Sampling : SS) 法 [1] があるが, SA 内の M-Blk を縦方向あるいは横方向に間引いてブロックマッチング (BM) を行うため, 高精度に動きベクトル (Motion Vector : MV) を検出できない。以上の問題を解決するために, 高速でかつ高精度な動きベクトルを検出するアルゴリズムを開発した。

本アルゴリズムを動きの遅い動画像 “Susie” と動きの速い動画像 “Whale-Show” (MPEG2, MP@ML) に適用した結果, 画質を維持しつつ (平均最小 D_s を FS の約 7% 増以内), ME 速度を SA サイズ $[-p_1, p_1]$ が $[-64$ 画素, 64 画素] のとき, FS の約 100 倍以上に高速化できた。

以下, 第 2 章でアルゴリズムについて述べ, 第 3 章で画像解析結果を述べた後, 第 4 章でまとめとする。

2. 3段階動画像動き検出法

本アルゴリズムは3段階で構成される。画像の動き (速さ, 量) は M-Blk 毎に変化する。本アルゴリズムの第 1 段階ではこの特性を利用して, 水平方向と垂直方向について探索し, 次に検出された最小差分絶対値和 (D_s) をしきい値と比較を行う。本アルゴリズムの第 2 段階ではこれを利用して, 決定された SA サイズを用いて, M-Blk を縦, 横共に 4:1 で間引いた探索を行う。本アルゴリズムの第 3 段階では, SA サイズを狭めて探索を行い BM を終了する。

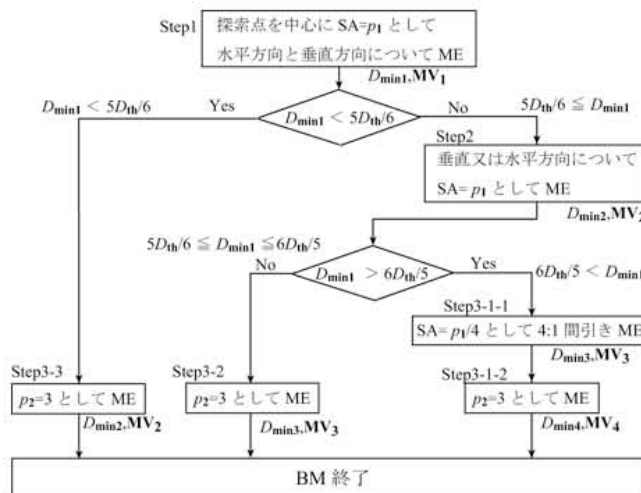


図 2.1 3段階動画像動き検出法のフローチャート

図 2.1 に本アルゴリズムの手順を示す。前フレームの平均最小 D_s をしきい値 D_{th} とする。まず, 図 2.2 に示すように探索の中心点から水平方向と垂直方向について SA を $[-p_1, p_1]$ として, 4:1 間引きで BM を行う。得られた差分絶対値和 (D_s) の最小値を D_{min1} , 対応する MV を MV_1 とする (Step 1)。

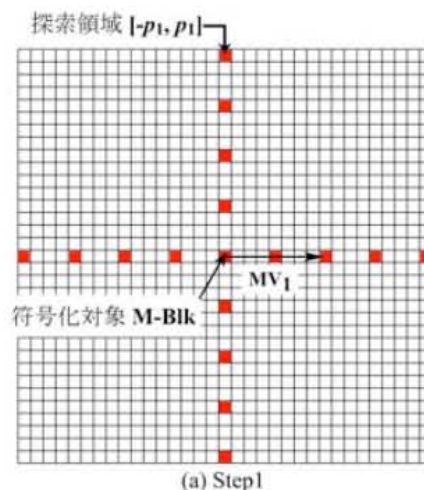


図 2.2 3段階動画像動き検出法の BM 手順 (Step 1)

次に, D_{min1} としきい値 (D_{th} =前フレームの平均最小 D_s) と比較する。 $D_{min1} < 5D_{th}/6$ の時, 第 3 段階へ移る (Step 3-3)。 $D_{min1} \geq 5D_{th}/6$ の時, 図 2.3 に示すように D_{min1}, MV_1 を得た M-Blk を中心として, 垂直方向 (図 2.3(b)) または水平方向 (図 2.3(c)) について SA を $[-p_1, p_1]$ として, 4:1 間引き BM を行う。得られた D_s の最小値を D_{min2} , 対応する MV を MV_2 とする (Step 2)。

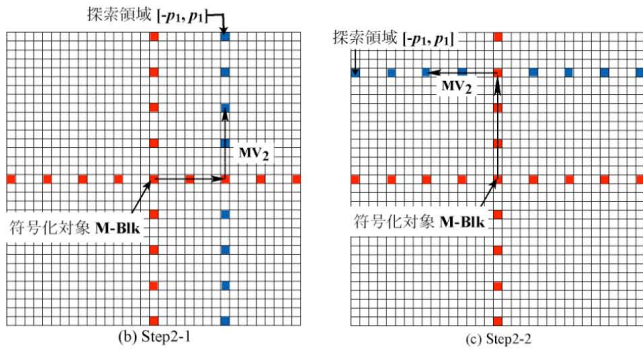


図 2.3 3段階動画像動き検出法の BM 手順 (Step2)

第3段階では、 $D_{min1} > 6D_{th}/5$ の時、 D_{min2} , MV_2 を得た M-Blk を中心として、SA を $[-p_1/4, p_1/4]$ とした 4:1 間引き BM を行う。得られた D_a の最小値を D_{min3} , 対応する MV を MV_3 とする (Step3-1-1). D_{min3} , MV_3 を得た M-Blk を中心として、SA を $[-3, 3]$ とした BM を行う (図 2.4(d)), 得られた D_a の最小値を D_{min4} , 対応する MV を MV_4 とし、BM を終了する (図 2.4(e)) (Step3-1-2).

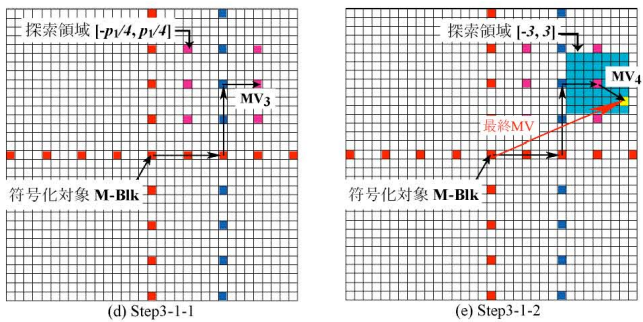


図 2.4 提案アルゴリズムの BM 手順 (Step3-1)

$6D_{th}/5 \geq D_{min1} \geq 5D_{th}/6$ の時、 D_{min2} , MV_2 を得た M-Blk を中心として、SA を $[-3, 3]$ とした BM を行う。得られた D_a の最小値を D_{min3} , 対応する MV を MV_3 とし、BM を終了する (Step3-2)。 (図 2.5(f)) $D_{min1} < 5D_{th}/6$ の時、 D_{min1} , MV_1 を得た M-Blk を中心として、SA を $[-3, 3]$ とした BM を行う。得られた D_a の最小値を D_{min2} , 対応する MV を MV_2 とし、BM を終了する (Step3-3)。 (図 2.5(g))

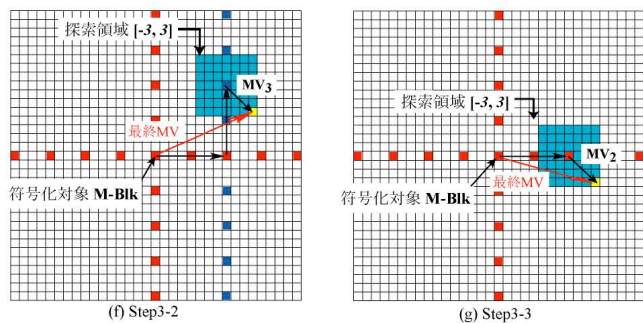


図 2.5 提案アルゴリズムの BM 手順 ((f)Step3-2, (g)Step3-3)

3. 画像解析

用いるテスト画像は、“Susie”と呼ばれる動きの遅い動画像と、“Whale-Show”と呼ばれる動きの速い動画像である。フレームサイズが MP@ML (720 画素×480 ライン) で、フレーム数は“Susie”が 150, “Whale-Show”が 450 である。以下では P ピクチャに対し

て MPEG-2 に準拠したソフトウェアエンコーダで処理速度と画質の解析を行う。画素データは符号無し 8b/画素, データレートは 4 Mbps, フレームレートは 30 fps, SA (p_1) は 16, 32, 48, 64, 80, 96, 112 である。なお、半画素精度は併用しない。図 3.1 に“Susie”の原画像を、図 3.2 に“Whale-Show”の原画像を示す。



図 3.1 テスト画像の原画像 ((a) “Susie”, (b) “Whale-Show”)

3.1 “Susie”の解析結果

テスト画像“Susie”を用いて、SA (p_1) が 64 のときの解析結果を表 3.1 に示す。提案アルゴリズムの平均最小 D_a は 657 で、FS の平均最小 D_a (= 623) の 5.46 % 増である。平均ピーク S/N 比 (平均 R_{sn}) は 40.896dB で、FS の平均 R_{sn} (= 40.9114 dB) の 0.0159dB 減となった。また、3-ss HQ, 3-SS Faster の平均 R_{sn} (= 40.8865 dB, 40.872 dB) とは、それぞれ 0.0090 dB 増 (向上), 0.0235 dB 増 (向上) になった。提案アルゴリズムは 3-ss HQ, 3-SS Faster を用いた場合より画質が向上している。以上の結果より、画質の劣化はなく、FS 並の画質が得られている事がわかる。

平均 N_b (= 140) は、FS の平均 N_b (= 14,423), 3-ss HQ の平均 N_b (= 1149), 3-SS Faster の平均 N_b (= 605) と比較して、それぞれ 0.97%, 12.2%, 23.1% に減少し、提案アルゴリズムを用いた場合の平均 F_s は FS, 3-ss HQ, 3-SS Faster の場合と比較して、それぞれ 102.57 倍, 8.21 倍, 4.32 倍に高速化され、3-ss よりも高速で動きベクトルを検出することがわかる。

表 3.1 符号化特性 (Susie, MP@ML, 4 Mbps, 30 fps, $p_1=64$)

アルゴリズム	平均 BM 数 (FS 比 [%])	平均 F_s [倍]	平均 D_a (FS 比 [%])	平均 R_{sn} [dB] (FS 比 [%])
FS	14423.5215 (100.0)	1.0000	623.0921 (+0.000%)	40.9114 (± 0.0000)
3-ss HQ*	1149.0964 (7.97)	12.55	636.1588 (+2.097%)	40.8865 (-0.0249)
3-SS Faster**	605.6766 (4.20)	23.82	645.9318 (+3.666%)	40.872 (-0.0394)
提案アルゴリズム	140.6187 (0.975)	102.6	657.1002 (+5.458%)	40.8955 (-0.0159)

*) 3-ss HQ は、 $m=0.25, n=1$ **) 3-SS Faster は、 $m=1, n=2$ 括弧内は FS との比較

図 3.2 に平均 D_a , 平均ピーク S/N 比 (平均 R_{sn}) と FS の SA サイズ $[-p_1, p_1]$ の関係を示す。SA サイズに関わらず、本アルゴリズムの平均最小 D_a は FS の平均最小 D_a の +7% 以内、本アルゴリズムの平均 R_{sn} は FS の平均 R_{sn} の約 -0.03 dB 以内で、FS と比べて遜色ない画質を維持している。 $p_1 = 16$ の時、本アルゴリズムの平均最小 D_a (= 765.9) は、FS の平均最小 D_a (= 749.9) の 3.55% 増 (劣化), 3-ss HQ, 3-SS Faster の平均最小 D_a (= 764.3, 792.0) の、それぞれ 0.21% 増 (劣化), 3.39% 減 (向上) である。本アルゴリズムは 3-ss よりも画質が向上している。同様に、 $p_1 = 16$ の時、本アルゴリズムの平均 R_{sn} (= 40.9534 dB) は、FS の平均 R_{sn} (= 40.9499 dB) の、0.0035 dB 増, 3-ss HQ, 3-SS Faster の平均 R_{sn} (= 40.9 dB, 40.8728 dB) の、それぞれ 0.0534 dB 増, 0.0806 dB 増 (向上) である。こちらでも本アルゴリズムは 3-SS より画質が向上している。

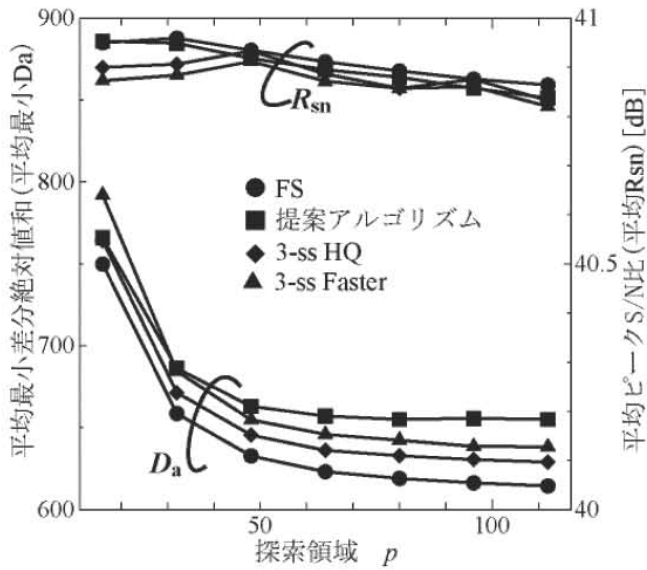


図 3.2 テスト画像“Susie”の平均 D_a , 平均 R_{sn} と探索領域サイズ p の関係 (MP@ML, MPEG-2, P 位 4:4:4, 40 枚, 4 Mbps, 30 fps, $p_1=16,32,48,64,80,96,112$).

図 3.3 に平均 M_b (実線), 平均 ME 速度 (平均 F_s) {(平均 $F_s = FS$ の平均 M_b) / (各アルゴリズムの平均 M_b)} (破線) と FS の SA サイズ $[p_1, p_1]$ の関係を示す。本アルゴリズムの平均 M_b は, FS のそれより小さく, $p_1 \geq 64$ で約 1/102 以下である。つまり, 本アルゴリズムの平均 F_s は FS の 102 倍以上速い。 $p_1 = 112$ の時, 本アルゴリズムの平均 M_b ($= 229$) は, FS の平均 M_b ($= 40019$), 3-ss HQ の平均 M_b ($= 3186$), 3-SS Faster の平均 M_b ($= 1655$) のそれぞれ 0.57%, 7.18%, 13.83% に減少し, 本アルゴリズムの平均 F_s は FS, 3-ss HQ, 3-SS Faster のそれぞれ 174.75 倍, 13.91 倍, 7.22 倍に高速化され, 3-SS よりも高速化されることがわかる。

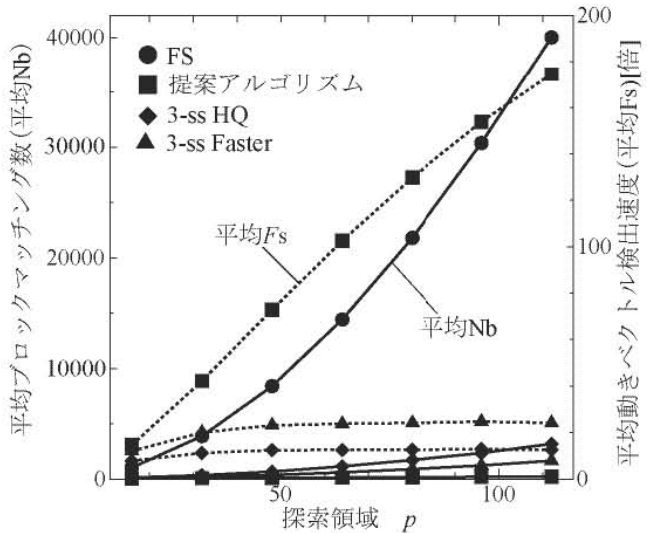


図 3.3 テスト画像“Susie”の平均 M_b , 平均 F_s と探索領域サイズ p の関係 (MP@ML, MPEG-2, P 位 4:4:4, 40 枚, 4 Mbps, 30 fps, $p_1=16,32,48,64,80,96,112$).

図 3.4 (a), (b) に FS, 本アルゴリズム の復号画像を示す。画質の優劣は見られない。



図 3.4 “Susie” の復号画像 (4Mbps, 30fps, $p_1 = 64$). (a)FS. (b) 本アルゴリズム

3.2 “Whale-Show”の解析結果

テスト画像“Whale-Show”を用いて, SA (p_1)が 64 のときの解析結果を表 3.2 に示す。提案アルゴリズムの平均最小 D_a は 3137 で, FS の平均最小 D_a ($= 2915$) の 7.63 % 増である。平均ピーク S/N 比 (平均 R_{sn}) は 27.417 dB で, FS の平均 R_{sn} ($= 27.5131$ dB) の 0.0159dB 減となった。また, 3-ss HQ, 3-SS Faster の平均 R_{sn} ($= 27.4034$ dB, 27.4004 dB) の, それぞれ 0.0143 dB 増 (向上), 0.0173 dB 増 (向上) になった。提案アルゴリズムは 3-ss HQ, 3-SS Faster を用いた場合より画質が向上している。以上の結果より, 画質の劣化はなく, FS 並の画質が得られている事がわかる。

平均 M_b ($= 140$) は, FS の平均 M_b ($= 14,423$), 3-ss HQ の平均 M_b ($= 1149$), 3-SS Faster の平均 M_b ($= 605$) と比較して, それぞれ 0.97%, 12.2%, 23.1% に減少し, 提案アルゴリズムを用いた場合の平均 F_s は FS, 3-ss HQ, 3-SS Faster の場合と比較して, それぞれ 102.57 倍, 8.21 倍, 4.32 倍に高速化され, 3-ss よりも高速で動きベクトルを検出することがわかる。

表 3.2 符号化特性(Whale-Show, MP@ML, 4 Mbps, 30 fps, $p_1=64$)

アルゴリズム	平均 BM 数 (FS 比[%])	平均 F_s [倍]	平均 D_a (FS 比[%])	平均 R_{sn} [dB] (FS 比[dB])
FS	14423.5068 (100.0000)	1.0000	2915.2017 (+0.000%)	27.5131 (± 0.0000)
3-SS HQ*	1254.2477 (8.6958)	11.49	3027.4578 (+3.8507%)	27.4034 (-0.1097)
3-SS Faster**	698.5719 (4.8432)	20.64	3135.3291 (+7.5510%)	27.4004 (-0.1127)
提案アルゴリズム	142.2763 (0.9864)	101.3	3137.7161 (+7.6328%)	27.4177 (-0.0954)

* 3-ss HQ は, $m=0.25, n=1$ ** 3-SS Faster は, $m=1, n=2$ 括弧内は FS との比較

図 3.5 に平均 D_a , 平均ピーク S/N 比 (平均 R_{sn}) と FS の SA サイズ $[p_1, p_1]$ の関係を示す。SA サイズに関わらず, 本アルゴリズムの平均最小 D_a は FS の平均最小 D_a の +9 % 以内, 本アルゴリズムの平均 R_{sn} は FS の平均 R_{sn} の約 -0.17 dB 以内で, FS と比べて遜色ない画質を維持している。 $p_1 = 16$ の時, 本アルゴリズムの平均最小 D_a ($= 3798.5$) は, FS の平均最小 D_a ($= 3679.5$) の 3.23% 増 (劣化), 3-ss HQ, 3-SS Faster の平均最小 D_a ($= 3781.1, 4071.1$) の, それぞれ 0.46% 増 (劣化), 6.69% 減 (向上) である。本アルゴリズムは 3-ss よりも画質が向上している。同様に, $p_1 = 16$ の時, 本アルゴリズムの平均 R_{sn} ($= 27.3215$ dB) は, FS の平均 R_{sn} ($= 27.1142$ dB) の, 0.2073 dB 増, 3-ss HQ, 3-SS Faster の平均 R_{sn} ($= 27.0095$ dB, 26.7884 dB) の, それぞれ 0.312 dB 増, 0.5331 dB 増 (向上) である。こちらも本アルゴリズムは 3-SS より画質が向上している。

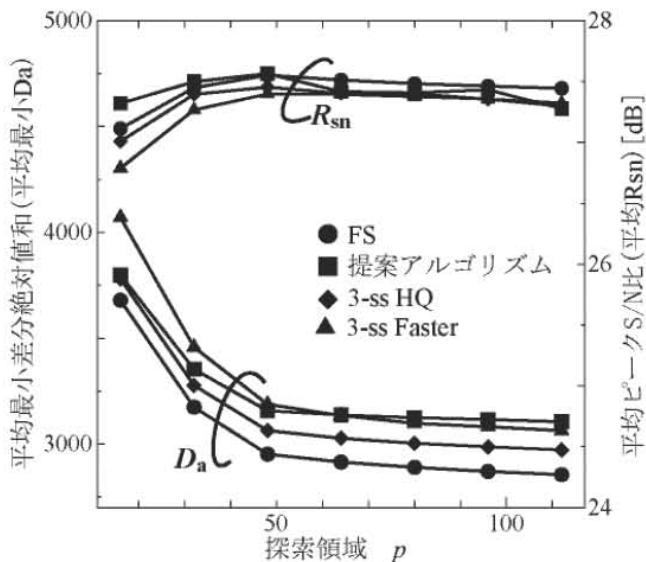


図 3.5 テスト画像“Whale-Show”の平均 D_a , 平均 R_{sn} と探索領域サイズ p の関係 (MP@ML, MPEG-2, Pt' 44:40 枚, 4 Mbps, 30 fps, $p_1=16,32,48,64,80,96,112$).

図 3.6 に平均 M_b (実線), 平均 ME 速度 (平均 F_s) (破線) と FS の SA サイズ p の関係を示す。本アルゴリズムの平均 M_b は, FS のそれより小さく, $p_1 \geq 64$ で約 1/101 以下である。つまり, 本アルゴリズムの平均 F_s は FS の 101 倍以上速い。 $p_1 = 112$ の時, 本アルゴリズムの平均 M_b ($= 233$) は, FS の平均 M_b ($= 40019$), 3-ss HQ の平均 M_b ($= 3477$), 3-SS Faster の平均 M_b ($= 1942$) のそれぞれ 0.58%, 6.70%, 11.99% に減少し, 本アルゴリズムの平均 F_s は FS, 3-ss HQ, 3-SS Faster のそれぞれ 171.75 倍, 14.92 倍, 8.33 倍に高速化され, 3-SS よりも高速化されることがわかる。

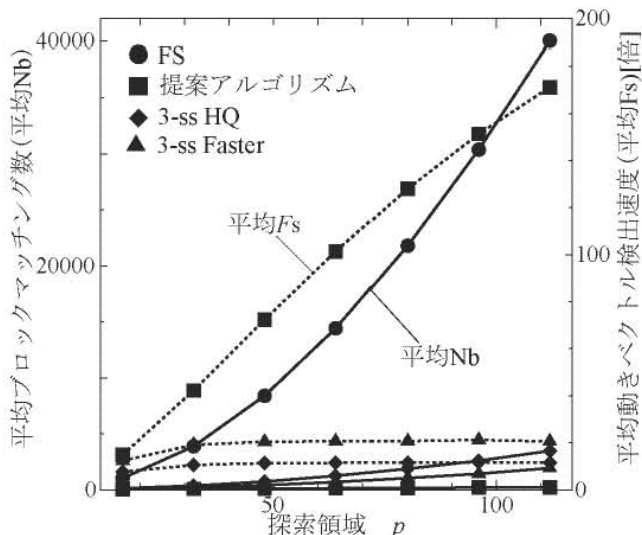


図 3.6 テスト画像“Whale-Show”の平均 M_b , 平均 F_s と探索領域サイズ p の関係 (MP@ML, MPEG-2, Pt' 44:120 枚, 4 Mbps, 30 fps, $p_1=16,32,48,64,80,96,112$).

図 3.7(a), (b) に FS, 本アルゴリズム の復号画像を示す。画質の優劣は見られない。

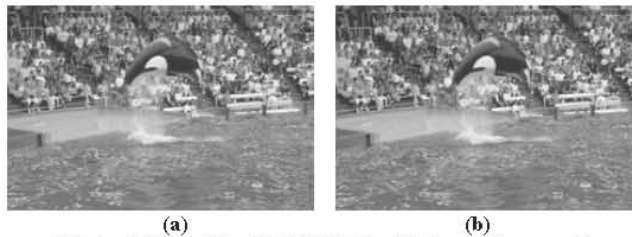


図 3.7 “Whale-Show”の復号画像 (4Mbps, 30fps, $p_1 = 64$). (a)FS, (b)本アルゴリズム

4. おわりに

探索領域(SA)を水平方向と垂直方向について探索し, 次に検出された最小差分絶対値和(D_a)をしきい値と比較し, SA を変更してブロックマッチング (BM) を行う 3 段階動画像動き検出法を開発した。

動きの遅いテスト画像“Susie”(MP@ML, 4Mbps, 30fps, $p_1 = 64$) を用いた場合, 平均最小 D_a は 657.1 となり, 全探索法(FS)の平均最小 D_a ($= 624.8$) の 5.45%増 (劣化) となった。3 階層サブサンプリング法(3-ss HQ, 3-SS Faster)の平均最小 D_a ($= 636.1, 645.9$) に対しては, それぞれ 3.29% 増, 1.73%増 (劣化) となった。平均ピーク SN 比 (平均 R_{sn}) は 40.8955 dB であり, FS の平均 R_{sn} ($= 40.9114$ dB) の 0.0159 dB 減少 (劣化) となった。3-ss HQ, 3-SS Faster の平均 R_{sn} ($= 40.8865$ dB, 40.872 dB) に対しては, それぞれ 0.0090 dB 増 (向上), 0.0235 dB 増 (向上) となった。本アルゴリズムは画質を維持し (平均最小 D_a を FS の 5.45%増, 平均 R_{sn} を FS の 0.0159 dB 減), 3-SS 法よりも大幅に向上した。

また, 平均ブロックマッチング数(平均 M_b)は 140 となり, FS の平均 M_b ($= 14,423$) の 0.97% となった。3-ss HQ, 3-SS Faster の平均 M_b ($= 1149, 605$) に対しては, それぞれ 12.2%, 23.1%の減少となった。平均動きベクトル検出速度 (平均 F_s) は FS の 102.57 倍に高速化できた。また, 3-ss HQ, 3-SS Faster の平均 F_s のそれぞれ 8.21 倍, 4.32 倍であり, 3-ss よりも速い結果となった。

動きの速いテスト画像“Whale-Show”(MP@ML, 4Mbps, 30fps, $p_1 = 64$) を用いた場合, 平均最小 D_a は 3137.7 となり, 全探索法(FS)の平均最小 D_a ($= 2915.2$) の 7.63%増 (劣化) となった。3 階層サブサンプリング法(3-ss HQ, 3-SS Faster)の平均最小 D_a ($= 3027.5, 3135.3$) に対しては, それぞれ 3.63% 増, 1.11%増 (劣化) となった。平均 R_{sn} は 27.4177 dB であり, FS の平均 R_{sn} ($= 27.5131$ dB) の 0.0954 dB 減 (劣化) となった。3-ss HQ, 3-SS Faster の平均 R_{sn} ($= 27.4034$ dB, 27.4004 dB) に対しては, それぞれ 0.0143 dB 増 (向上), 0.0173 dB 増 (向上) となった。本アルゴリズムは画質を維持し (平均最小 D_a を FS の 7.63%増, 平均 R_{sn} を FS の 0.0954 dB 減), 3-SS 法よりも大幅に向上した。

また, 平均ブロックマッチング数(平均 M_b)は 142 となり, FS の平均 M_b ($= 14,423$) の 0.98% となった。3-ss HQ, 3-SS Faster の平均 M_b ($= 1254, 698$) に対しては, それぞれ 11.3%, 20.3%の減少となった。平均動きベクトル検出速度 (平均 F_s) は FS の 101.38 倍に高速化できた。また, 3-ss HQ, 3-SS Faster の平均 F_s のそれぞれ 8.83 倍, 4.92 倍であり, 3-ss よりも速い結果となった。

謝辞

研究にあたり, 中央大学の榎本忠儀先生には御指導を頂き, 大変お世話になりました。また, 卒業された先輩方, 中央大学榎本研究室の皆様には多くの助言, 御協力を頂きました。ここに心より御礼を申し上げます。

参考文献

- [1] 榎本, 「画像 LSI システム設計技術」, コロナ社, pp. 77-91, 2003 年 9 月。
- [2] 小熊, 木谷, 榎本, 「探索領域を適応設定する 3 階層サブサンプリング動きベクトル検出法」, 信学技報, ICD2004-115, pp. 43-48, 2004 年 10 月 21 日。