

自己相似性を用いた適応的な時間分解能選択に基づく映像符号化

島野美保子[†] チョンジーン^{††} 佐藤いまり^{††}

[†] 科学技術振興機構 さきがけ, 東京大学 生産技術研究所

^{††} 国立情報学研究所

E-mail: [†]miho@iis.u-tokyo.ac.jp, ^{††}gene.cheung@ieee.org, ^{††}imarik@nii.ac.jp

あらまし 映像に適した露光時間は、撮影シーンの照明環境に大きく影響される。一般に、光量の少ない環境下では露光不足を防ぐため、連続露光(露光時間 = 1 フレーム時間)の低フレームレート映像で撮影される。本論文では、このような映像を扱ったネットワークストリーミングにおいて、撮影映像よりも高フレームレートの映像を、デコーダで効率よく提供する手法を提案する。本手法は、連続露光映像の特性である自己相似性を用いた高時間分解能化を符号化技術に融合することで、再生する高フレームレート映像を高効率に生成する。これを可能にするため、エンコーダにおける最適なフレーム列の表現を、符号量と予測誤差の関係に基づいて、各フレームの時間分解能と量子化パラメタを適応的に選択することによって求める。更に、時間分解能と量子化パラメタの最適化において、コスト単調性に基づく高速アルゴリズムを実現する。実映像を用いた実験により、提案手法が非適応的手法に比較して PSNR で最大 1.3dB 上回ることを確認した。

キーワード 映像圧縮, 時間分解能, 超解像, 自己相似性

1. はじめに

映像に適した露光時間は、注目シーンの照明条件に大きく影響される。光量の少ない照明条件下では、1 フレーム時間よりも短い露光時間で撮影するのではなく(図 1(a)), 露光不足を防ぐため長い露光時間で撮影することが不可欠である(図 1(b))。通常、1 フレーム時間は露光時間よりも長い必要があり、長い露光時間による撮影では、動きぶれを含むような低フレームレートの映像に成らざるを得ない。

物理的あるいは経済的な制約のため、しばしばこのような光量不足の撮影シーンを見かける。例えば、撮影対象がカメラから遠く離れているため照明を当てることができない場合がある。また、夜間の監視カメラによるサーベイランスや観察のための広域にわたる長時間の照明は、経済的に高コストとなり現実的でない、あるいは夜間の動物の生息を妨害する可能性が生じる。従って、長い露光時間で撮影された低フレームレート映像、特に図 1(b) に示すような 1 フレーム時間を露光時間と等しくとり、フレームレートを最大にした連続露光映像を扱う必要がある。

連続露光映像では、カメラや被写体の動きによる動きぶれがしばしば見られる。最近、これらの影響を低減しつつフレームレートを上げるという画質改善を行うため、連続露光映像の特性である自己相似性を利用した映像の高時間分解能化手法が提案された[1]。この手法は、異なる露光時間の映像間では、動きぶれの見えパターンが相似になることを利用し、高時間分解能の見えパターンを復元する。これにより、1 フレームを時間方向に分解し

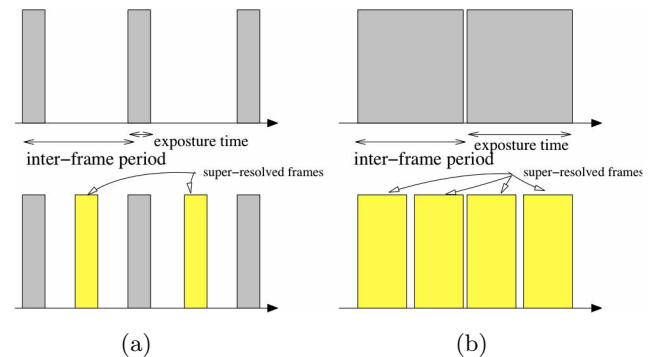


図 1 高時間分解能化 (a) 短露光時間映像(上)、フレーム補間手法で生成された高時間分解能映像(下)、(b) 連続露光映像(上)、高時間分解能化で生成された高時間分解能映像(下)。

て露光時間を短くし、フレームレートを増加する高時間分解能化を実現する。この高時間分解能化によって復元された高フレームレート映像は、オプティカルフロー[2]を用いてフレーム補間された従来手法に比べ、PSNR について最大 6dB 上回る改善がみられた。

昨今、ユーザーがサーバーからビデオストリーミングを受け取り、再生するというネットワークストリーミングが、著しく広がりつつある。このネットワークストリーミングを担う映像符号化技術の発展は不可欠であるが、以下の課題が存在する。それは、上述のような連続露光映像について、より時間分解能の高い映像をデコーダで生成するために、エンコーダにおける最も効率的な表現を考案することである。解決策の 1 つとしては、圧縮前にエンコーダで高時間分解能化を行う方法が挙げ

られる．しかし，この方法は撮影された映像よりも高フレームレートで圧縮を行わなければならない，多くの符号量を要する．別の選択肢としては，エンコーダで撮影映像をそのまま符号化し，これらの圧縮されたフレーム列に対してデコーダで高時間分解能化を施し，高フレームレート映像を復元する方法がある．しかしながら，生成される高フレームレート映像の画質は，圧縮時の量子化ノイズによって損なわれてしまう．また，デコーダにおける複雑性の制約（高時間分解能化に利用する近傍フレームの領域の制限）を厳しく設定することによっても，画質が低下する．

本論文では，撮影された連続露光映像よりも高時間分解能の映像を提供することを目的とし，符号量（レート）と予測誤差（歪み）のトレードオフを最適化するために，適切な時間分解能と量子化パラメタ（QP）を選択して符号化するという適応的な圧縮方法を提案する．提案手法の基本的な考え方を具体例を用いて以下に示す．

デコーダで高時間分解能化による生成が容易なフレームについては，符号量を節約するため，低い時間分解能で，且つ粗いQPによって符号化する．一方，圧縮後に生成することが難しいフレームについては，画質を保つため，高い時間分解能で，且つ精細なQPによって符号化するというものである．これらの異なる時間分解能と異なるQPのフレームの選択方法を，無閉路有向グラフ（DAG）の最短経路問題として解く．更に，予測器の時間分解能とQPにおけるコスト単調性の仮定に基づき，高速な探索アルゴリズムを提案する．適応的なフレーム選択により圧縮する提案手法が，同じ符号量で比較した非適応的手法に対し，PSNRで最大1.3dB性能が上回る結果を確認した．

以下の章は次の通り構成される．まず2.章において関連研究について要約し，3.章では自己相似性を利用した高時間分解能化技術の概要を示す．4.から6.章において，レート歪み特性の最適化問題を定式化し，最短経路探索アルゴリズムを説明する．7.章では評価実験により提案手法の性能を確認し，8.章に結論をまとめる．

2. 関連研究

映像の高時間分解能化 [1] と目的が類似している手法として，図 1(a) のように時間方向のフレーム補間を行う，低ビットレート映像符号化のための可変フレームレートに変換するフレーム補間手法が提案されている [3], [4]．これらの手法は，一般に，オプティカルフロー [2] による動き推定や動き補償に基づいているが，低フレームレートの連続露光映像では動きぶれを含むため，フレーム画像間における密な動きベクトルの対応がとれるとは限らない．また，図 1(b) のような連続露光映像では，低時間分解能の1フレームを複数の高時間分解能フレームに分解できないため，高時間分解能映像を生成できない．これらの従来手法と比較して，連続露光映像の特性である

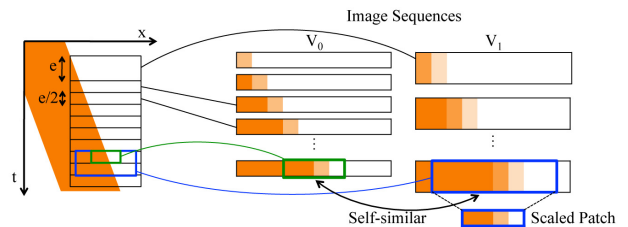


図 2 自己相似性．等速直線運動する物体を撮影した時間分解能の異なる 1 次元映像

自己相似性を利用した高時間分解能化手法は顕著に優れた復元性能を示している [1]．

提案する映像符号化のための最適な時間分解能およびQPのフレーム選択手法は，符号量と予測誤差の最適化による符号量配分問題における新しい試みである．関連手法としては，動き補償による映像符号化における符号量配分手法が最初に提案された [5]．また，最適なQP選択問題として，符号量制約下における予測誤差を最小化するため，映像内のフレームを間引く手法^{注1}も提案されている [6]．最近では，多視点画像のテクスチャと奥行き情報のサブセットの選択手法として，総符号量の制約下でデコーダにおける生成誤差を最小化するような，最適なQPでの差分符号化を実現した [7]．本手法は，以上の従来手法と2点において大きく異なる．第1は，符号量と予測誤差を最適化するため，異なる時間分解能およびQPの選択を行うという，従来検討されなかった映像符号化の問題設定である．第2には，予測器の時間分解能とQPの単調性に基づいた高速経路探索を実現した点である．

3. 自己相似性を利用した高時間分解能化

自然画像には，空間的に自己相似的なテクスチャパターンが様々なスケールで多数存在している．同様に，自然映像では時間方向にも，この自己相似性を観察することが可能である．ここで言う自己相似性とは，異なる時間分解能映像間で，各露光時間中の動きぶれの見えパターンが相似になることを指す．最近，この連続露光映像の特性である自己相似性を利用した映像の高時間分解能化手法が提案されている [1]．

まず，自己相似性を観察できる簡潔な例として，等速直線運動をする物体を連続露光で撮影した1次元映像を考える．図 2 に，露光時間 (=1 フレーム時間) $e/2, e$ を持つ2つの映像 V_0, V_1 を示す． V_1 のフレーム内の物体は， V_0 よりも露光時間が2倍のため2倍の距離を移動している．この際， V_1 で観察する微小領域のサイズを V_0

(注1): フレームの間引きによってもフレームレートが変換されることになるが，ここでの時間分解能とは，フレーム間引きとは異なり，低時間分解能フレームはその2倍の高時間分解能の2フレーム平均で表現される．本手法は，これらの複数の時間分解能のうち最適な時間分解能フレームを選択する点に着目する．

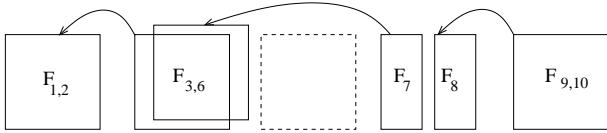


図 3 連続露光映像の符号化例．MR フレーム $F_{(1,2)}$ および $F_{(9,10)}$ はそのまま符号化される，LR フレーム $F_{(3,6)}$ は $F_{(3,4)}$ と $F_{(5,6)}$ のフレーム平均であり，SR フレーム F_7 と F_8 は MR フレーム $F_{(7,8)}$ から生成される．

の微小領域の 2 倍とするならば，図 2 のように， V_0 と V_1 の微小領域の見えは相似となる．この異なる時間分解能映像間の相似な関係は，2 次元にも拡張することが可能である．また，この相似な関係は，空間方向や時間方向のスケールを適応的に調整することにより，様々な運動やテクスチャの微小領域に拡張することが可能である．

これらの自己相似性を活用することにより，最大事後確率 (MAP) 推定の枠組みで 1 つの映像のみからの高時間分解能化を実現する．高時間分解能映像のフレーム平均が低時間分解能映像のフレームを成すという時間分解能映像間の映像復元条件を，高時間分解能フレームから低時間分解能フレームを構成する尤度として採用する．また，映像内の，ある動きぶれパターンに対応する相似なパターンを同一映像内から探索し，これを適応的にスケール変換することにより，動きぶれを半減させたパターンが生成できる．このような自己相似な事例を事前確率として利用する．これらの事前確率と映像復元条件を考慮した確率モデルによって，撮影映像の低時間分解能フレームを複数フレームに分解した高時間分解能映像を生成することができる．

4. レート歪み特性の最適化

本手法では，連続露光のフレームにおける適切な時間分解能の選択問題を，組み合わせの最適化として表す．初めに，選択可能なフレームの種類について概要を説明する．撮影された連続露光の N 枚の中時間分解能 (MR) フレームを $\mathcal{F}^M = \{F_{(1,2)}, F_{(3,4)}, \dots, F_{(2N-1,2N)}\}$ と表し， $2N$ 枚の高時間分解能化した SR (super-resolved) フレームを $\mathcal{F}^S = \{F_1, F_2, \dots, F_{2N}\}$ と表す．更に，撮影映像の連続する 2 フレームの平均である低時間分解能 (LR) フレームを $\mathcal{F}^L = \{F_{(1,4)}, F_{(3,6)}, \dots, F_{(2N-3,2N)}\}$ と表す．ここで， $F_{(i-3,i)}$ は MR フレーム $F_{(i-3,i-2)}$ と $F_{(i-1,i)}$ のフレーム平均である．

高時間分解能化によって，各 MR フレーム $F_{(i-1,i)} \in \mathcal{F}^M$ から，対応する 2 枚の SR フレーム F_{i-1} と F_i を生成できる．また，MR フレームの $F_{(i-1,i)}$ と $F_{(i-3,i-2)}$ の平均をとることで対応する 1 枚の LR フレーム $F_{(i-3,i)}$ を生成できる．逆に，LR フレームから，高時間分解能化によって MR フレームや SR フレームを生成できる．

本手法の目的は，ある総符号量で符号化された \mathcal{F}^S ， \mathcal{F}^M ，および \mathcal{F}^L から選択されたフレーム列を用いて，最

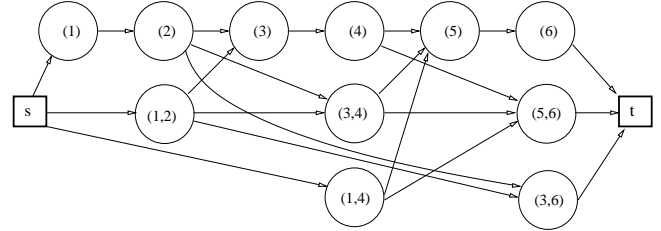


図 4 SR の 6 フレームと，開始 s ，終了 t ノードによる DAG 例．

高画質の SR フレームをデコーダで生成することである．SR フレームをエンコーダで符号化するならば，復号化後の画質は高いが，符号量が多い．一方，LR フレームをエンコーダで符号化するならば，符号量は少ないが，デコーダで生成される SR フレームの画質は低いと考えられる．このようなフレーム選択の組み合わせの例を図 3 に示す．

4.1 フレーム選択による無閉路有向グラフ

符号化するフレームの時間分解能の選択を行うため，以下のように無閉路有向グラフ (DAG) を作成する．図 4 に示すように，縦 3 列にわたり，SR フレーム \mathcal{F}^S ，MR フレーム \mathcal{F}^M ，および LR フレーム \mathcal{F}^L の 3 種類について，各フレーム F_x に対応するノード x を横方向に並べる．例えば，偶数の i に対し，縦方向は， F_i ， $F_{(i-1,i)}$ ，および $F_{(i-3,i)}$ に対応するノード i ， $(i-1,i)$ ，および $(i-3,i)$ が並べられる．これらのフレームは F_i に対応する異なる時間分解能のフレームである．更に，DAG の左右の両端には開始と終了ノード s ， t を設ける．

F_x を予測器として，ある時間分解能のフレーム F_y が符号化される関係を，ノード x から y へのエッジ $e_{x \rightarrow y}$ として表す．ノード i や (j,i) ， $\forall j$ は， $i+1$ や $(i+1,k)$ ， $\forall k$ の予測器となることが可能である．更に，開始ノード s からは，候補である最初のフレーム (I フレーム) F_1 ， $F_{(1,2)}$ ，および $F_{(1,4)}$ のノードへのエッジがある．終了ノード t は，最後のフレーム候補である F_{2N} ， $F_{(2N-1,2N)}$ ，および $F_{(2N-3,2N)}$ のノードからのエッジがある．

符号化されたフレームには 1 つの予測器であるフレームが存在し，またそのフレームにも 1 つの予測器が存在してというように，順に予測器が存在することになる．そのため，最後のフレームから I フレームである最初のフレームまで経路をたどることができる．この DAG における s から t への経路 p によって，映像符号化のためのフレーム選択が表される．従って，符号量と予測誤差を表すコストを各エッジに割り当てることが可能ならば，フレーム選択問題をコスト最小化問題として解くことができる．以下に，符号量と予測誤差のモデルについて説明する．

4.2 符号量モデル

経路長 L_p (s と t を除く) の経路 $p = \{s, p_1, p_2, \dots, p_{L_p}, t\}$

が与えられた場合に，経路 \mathbf{p} で符号化された映像の符号量 $R(\mathbf{p})$ は，選択フレームの総和として，

$$R(\mathbf{p}) = \sum_{l=1}^{L_{\mathbf{p}}} r^c(p_l, p_{l-1}) \quad (1)$$

と与えられる．ここで， $p_0 = s$ であり，1 フレーム目の符号量 $r^c(p_1, p_0)$ は，最初のフレーム選択 F_{p_1} で決定される．また， l 番目のフレームの符号量 $r^c(p_l, p_{l-1})$ は， l 番目のフレーム F_{p_l} の選択とその予測器 $F_{p_{l-1}}$ によって決まる．このようなマルコフモデルが有効であることは，最近報告されている [7]．

4.3 予測誤差モデル

エンコーダで選択されたフレームの経路 \mathbf{p} を用い，デコーダで高時間分解能化を行った映像の予測誤差 $D(\mathbf{p})$ は， $R(\mathbf{p})$ と同様に選択フレームの総和として，

$$D(\mathbf{p}) = \sum_{l=1}^{L_{\mathbf{p}}} d(p_l, \mathbf{p}) \quad (2)$$

と表される．ここで， $d(p_l, \mathbf{p})$ は，符号化されたフレーム F_{p_l} に対応する予測誤差である．符号化されたフレーム F_{p_l} がエンコーダで生成された SR フレームである場合，予測誤差 $d(p_l, \mathbf{p})$ は，単純に予測器 $F_{p_{l-1}}$ の選択にのみ依存する予測誤差 $d^c(p_l, p_{l-1})$ である．

一方， F_{p_l} が MR フレームあるいは LR フレームならば， F_{p_l} が対応する全ての SR フレームについて，デコーダで生成される際の生成誤差を考慮する必要がある． F_{p_l} が，MR フレーム $F_{(i-1,i)}$ ，あるいは LR フレーム $F_{(i-3,i)}$ である場合，この F_{p_l} はどの SR フレームの平均に相当するかというフレームの範囲を $(a(p_l), b(p_l))$ とおく．このとき，符号化されたフレーム F_{p_l} の予測誤差 $d(p_l, \mathbf{p})$ は，

$$d(p_l, \mathbf{p}) = \begin{cases} d^c(p_l, p_{l-1}) & \text{if } F_{p_l} \in \mathcal{F}^s \\ \sum_{k=a(p_l)}^{b(p_l)} d^s(k, \mathbf{p}) & \text{o.w.} \end{cases} \quad (3)$$

と記述できる．ここで， $d^s(k, \mathbf{p})$ は，選択されたフレームの経路 \mathbf{p} について，デコーダで高時間分解能化された SR フレーム F_k の生成誤差である．この $d^s(k, \mathbf{p})$ が \mathbf{p} に依存する理由は，SR フレーム F_k の生成に，符号化された映像の全フレームを使用して自己相似性を抽出し高時間分解能化を行うためである．従って，符号化された SR フレーム F_{p_l} の予測誤差 $d^c(p_l, p_{l-1})$ が局所的に p_l と p_{l-1} のみに依存するように， $d^s(k, \mathbf{p})$ を単純化することはできない．

4.4 DAG の最短経路

選択されたフレーム列 \mathbf{p} のための符号量 $R(\mathbf{p})$ と予測誤差 $D(\mathbf{p})$ のコストを用いて，最適な経路選択問題をラグランジュ最適化問題として次式のように定式化する．

$$\min_{\mathbf{p} \in \mathcal{P}} \Theta(\mathbf{p}) = D(\mathbf{p}) + \lambda R(\mathbf{p}) \quad (4)$$

ここで， \mathcal{P} は s から t への可能な経路集合で， λ はラグランジュ乗数である．式 (4) は，符号化されたフレーム F_{p_l} のラグランジュ関数のコスト $\theta(p_l, \mathbf{p})$ の総和として記述できる．

$$\Theta(\mathbf{p}) = \sum_{l=1}^{L_{\mathbf{p}}} \theta(p_l, \mathbf{p})$$

$$\theta(p_l, \mathbf{p}) = d(p_l, \mathbf{p}) + \lambda r^c(p_l, p_{l-1}) \quad (5)$$

各ラグランジュ関数のコスト $\theta(p_l, \mathbf{p})$ を経路 \mathbf{p} のエッジ e_{p_{l-1}, p_l} に割り当てる．これによって，DAG における経路 \mathbf{p} のコストが各エッジのコスト $c(e_{p_{l-1}, p_l})$ の総和で表され，最適なフレーム選択に相当する，DAG における最短コストの経路が決まる．次章で，最短経路の探索手法について説明する．

5. 高速探索アルゴリズム

式 (5) の最短経路を求めるためには，2 つの問題点がある．1 つは，エッジ e_{p_{l-1}, p_l} のラグランジュ関数のコスト $\theta(p_l, \mathbf{p})$ が，両端のエッジ p_{l-1} と p_l のみでなく，全経路 \mathbf{p} に依存する点である．2 つ目は，あるエッジのコスト $c(e_{p_{l-1}, p_l})$ が p_{l-1} と p_l のみの関数としても，多数のノードとエッジにおける最短経路を求めるには多くの計算量が必要である．以下に，これら 2 つの問題点について議論する．

5.1 反復計算による最短経路探索

1 つ目の問題は，各ステップにおいて固定したエッジコストで DAG の最短経路を算出するという反復計算により解決される．即ち，エッジ $e_{x \rightarrow y}$ の各コスト $c(e_{x \rightarrow y})$ はノード x と y により，他のノードとは独立に決定される．そして，エッジコスト更新による総コストの変化量が閾値 δ 以下になった時点で反復計算を終了する．詳細は，次の最短経路探索の反復計算手順に示す．

最短経路探索の反復計算手順

- 1: 1 つ前の経路 \mathbf{p}' を，全て SR ノードである経路に設定する．
- 2: $c(e_{x \rightarrow y})$ に $\theta(y, \mathbf{p}^i)$ を代入する．ここで，経路 \mathbf{p}^i は x と y を含み，その他の全てのノードは SR ノードとする．
- 3: 固定されたエッジコストから最短経路 \mathbf{p} を求める．このときの総コストを $\Theta'(\mathbf{p})$ とする．
- 4: $c(e_{p_{l-1}, p_l}) \leftarrow \max\{\theta(p_l, \mathbf{p}), c(e_{p_{l-1}, p_l})\}$.
- 5: if $|\Theta'(\mathbf{p}) - \Theta(\mathbf{p})| > \delta$ then
- 6: $\mathbf{p}' \leftarrow \mathbf{p}$. Goto ステップ 3 .
- 7: end if

この反復計算手順の根拠は以下のとおり説明できる．ステップ 2 では，各エッジ $e_{x \rightarrow y}$ に最小の $\theta(y, \mathbf{p}^i)$ を代入する．このときの経路 \mathbf{p}^i は，予測誤差を最小に抑えられるように x と y 以外の他の全てのノードを SR ノー

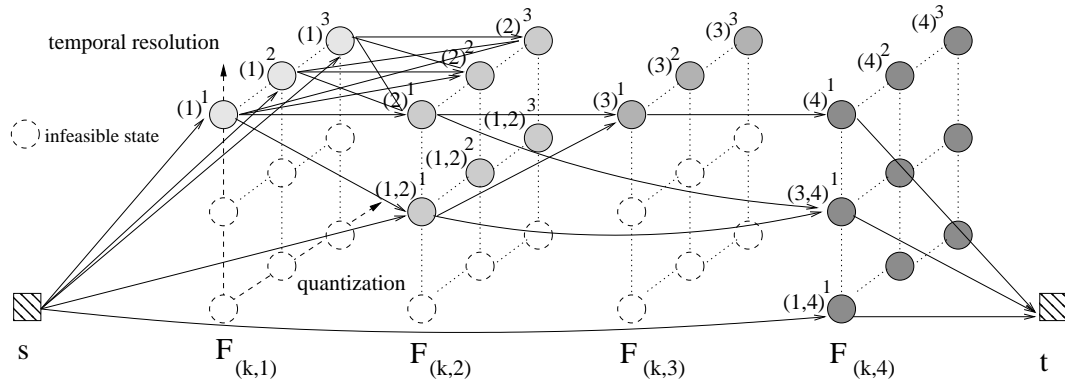


図5 4平面と、開始 s 、終了 t ノードによる3次元 DAG 例。

ドとする．これにより，次のステップで，全て SR ノードである初期経路とは異なる経路を探索することが可能である．以降，探索された最短経路 p の各エッジコストは p で新しく選択されたノードのみ増加することから，この手順を繰り返すことで収束する．更に，収束した最短経路のコストはステップ5により真の最短経路コストの上限を表す．

5.2 予測器の時間分解能における単調性

エッジコストが一定の場合，予測器の時間分解能のコスト単調性を仮定することで最短経路探索の高速化が可能である．一般に， $F_{(i-1,i)}$ のような高時間分解能フレームは， $F_{(i-3,i)}$ のような低時間分解能フレームよりも未来フレームを精度良く予測できると考えられる．従って，ある高時間分解能フレームを選択した際の局所的なコストが，低時間分解能よりも低いならば，この低時間分解能フレームは経路全体としても最適解には含まれない．より厳密には， $\psi(x)$ を s からノード x までの最短部分経路とおき，次のように記述できる．

[Lemma 1] 同一 SR フレームに相当する (ノードの縦列に並ぶ) 高時間分解能フレームを F_x ，低時間分解能フレームを F_y とする場合に， $\psi(x) < \psi(y)$ であれば， F_y は経路全体における最適解には含まれない．

[Proof of Lemma 1] 背理法で証明する．ノード $p_i^o = y$ を含む最短経路 p^o が存在すると仮定する．これと異なる新しい経路 $p^* = \{\psi(x), p_{i+1}^o, \dots, p_{L_{p^o}}^o, t\}$ を作成し， $\psi(x) < \psi(y)$ と仮定する．ここで，同一 SR フレームに相当する高時間分解能フレーム F_x ，低時間分解能フレーム F_y では予測器の時間分解能の単調性により， $c(e_{x \rightarrow p_{i+1}^o}) \leq c(e_{y \rightarrow p_{i+1}^o})$ である．従って，経路 p^* のコストは，最短経路 p^o のコストよりも少ないことになり，矛盾する．

Lemma 1 を，次のように最短経路探索の高速化に活用する．DAG におけるあるノードまでの最短部分経路は，左から右方向へ順に計算される．このとき，ノードの各縦列において， F_y よりも高時間分解能の F_x に対し

$\psi(x) < \psi(y)$ であれば，ノード y は最適な経路ではないため，DAG から削除することができる．

6. 量子化パラメタとデコーダにおける複雑性の導入

本章では，前述の時間分解能の選択のみではなく，量子化パラメタ (QP) を考慮した圧縮手法への拡張と，デコーダにおける複雑性を考慮したレート歪み特性の最適化について議論する．

6.1 量子化パラメタ選択による拡張

前章までは，各経路で固定された QP を用いてのレート歪み特性の最適化を考えてきた．今度は，符号化するフレームの時間分解能と QP の両方を自由に選択できるように，以下のように3次元の無閉路有向グラフ (DAG) を作成する．図5に示すように，時間分解能の異なる SR フレーム F_i ，MR フレーム $F_{(i-1,i)}$ と，LR フレーム $F_{(i-3,i)}$ の3つのセットで，ノード i の平面を作成する．即ち，平面 i の各ノード $x = f(x)^{q(x)}$ において，3つの縦列 $F_{f(x)}$ は，各々 F_i ， $F_{(i-1,i)}$ ，および $F_{(i-3,i)}$ に相当する．また，奥行き方向は，QP 値 $q(x)$ で符号化されたフレーム $F_{f(x)}$ を表す．

これらの平面には，実行不可能なノードも含まれる．例えば，平面1の縦2,3列目は，フレーム $F_{0,1}$ および $F_{-2,1}$ が存在しないため実行不可である．また，2次元の DAG と同様，左右の両端には開始と終了ノード s, t を設ける．QP 値 $q(x)$ で符号化された $F_{f(x)}$ を予測器として，ある時間分解能のフレーム $F_{f(y)}$ が QP 値 $q(y)$ で符号化される関係を，ノード x から y へのエッジ $e_{x \rightarrow y}$ として表す．この3次元 DAG における s から t への経路 p が，映像符号化のために選択されたフレームの時間分解能と QP に相当する．

高速な探索アルゴリズムとして，反復計算による最短経路探索は5.章と同様に計算できる．また，予測器の単調性についても同様に，時間分解能と QP の両方を考慮した単調性の仮定に基づき，高速な探索アルゴリズムを実現できる．

[Lemma 2] $f(x) = f(y)$ であるノード x および y の QP 値が $q(x) < q(y)$ である場合に, $\psi(x) < \psi(y)$ であれば, ノード y は経路全体における最適解には含まれない.

QP を考慮した Lemma 2 は, Lemma 1 と同様に証明できる. $q(x) = q(y)$ である Lemma 1 と, Lemma 2 を組み合わせて用い, 次のように最短経路探索の高速化を行う. まず, DAG における最短部分経路は, 平面 i のノードから i を増加する方向のエッジを計算して更新される. このとき, 各々の平面 i において, 平面 i の 2 つのノード x および y が 2 つの条件を満たすかを確認する: i) $\psi(x) < \psi(y)$, ii) $F_{f(x)}$ が $F_{f(y)}$ よりも, 同一あるいはより高時間分解能であり, 且つ QP 値が $q(x) \leq q(y)$. もしこれらを満たすならば, ノード y は最適な経路ではないため, DAG から削除することができる.

6.2 デコーダにおける複雑性考慮

新たに, デコーダにおける複雑性の制約 (高時間分解能化に利用する近傍フレームの領域の制限) を考慮したレート歪み特性の最適化を考える. 符号化されたフレーム F_{p_i} の予測誤差 $d(p_i, p)$ は式 (3) で表される. ここで, $d^s(k, p)$ は, 選択された経路 p のフレームを用いて, デコーダで高時間分解能化された SR フレーム F_k の生成誤差である. そこで, 複雑性を考慮するため, デコーダで自己相似性を抽出し高時間分解能化を行う際には, SR フレーム F_k の近傍の符号化された $2W$ フレーム (MR の W フレームに相当) を使用すると仮定する. これは, フレーム数 W が多くなるほど, 生成される SR フレームの画質は向上するが, デコーダにおいて高時間分解能化を行う際の複雑性が増すことを意味する. この複雑性による影響についても評価実験にて確認する.

7. 評価実験

実映像を用いて, 提案手法の評価実験を行った. フレーム画像サイズ 384×216 で 150 フレームの 2 種類の連続露光の交差点映像 (intrsct), および頭を揺らすくまのぬいぐるみの映像 (bear) を, MR フレーム \mathcal{F}^M として使用する. SR フレーム \mathcal{F}^S は, エンコーダで MR の各 150 フレームから, 3. 章の高時間分解能化によって生成される. また, LR フレーム \mathcal{F}^L は, 隣り合う MR フレームのフレーム平均として求められる.

符号量 r^c と予測誤差 d を計算するため, $\mathcal{F}^S, \mathcal{F}^M$, および \mathcal{F}^L の異なるフレームの組み合わせで, あるいは異なる量子化パラメタで, 標準的な H.264 により符号化する. MR フレーム \mathcal{F}^M と LR フレーム \mathcal{F}^L からの予測には, GOP(Group Of Picture) のサイズ, 例えば撮影映像の $N(= 30)$ フレーム列が, デコーダで SR フレームへの高時間分解能化に利用される. ここで, 予測誤差は, 高時間分解能化された符号化前の SR フレームと, 復号化後に再生される SR フレームとの誤差で表す. 以下の



(a)



(b)

図 6 生成された SR フレーム例 (a) intrsct と (b) bear 共に QP=35.

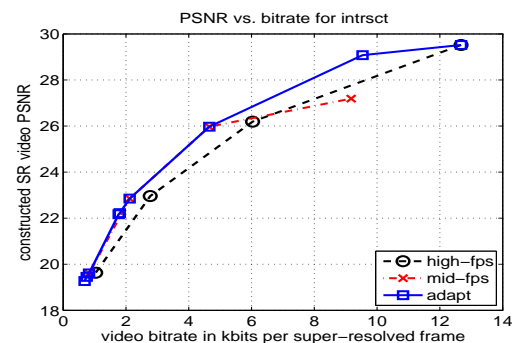


図 7 各圧縮手法によるレート歪み特性. x 軸は生成された SR フレームの平均符号量, y 軸は映像画質 (PSNR).

実験では, 異なる符号量 (レート) と予測誤差 (歪み) の関係を求めるため, ラグランジュの未定乗数 λ を変化させ, 最短経路探索により符号化に最適なフレームの選択を行う.

まず 7.1 節に, 最適な時間分解能フレーム選択によるレート歪み特性結果を示し, 7.2 節に最適な時間分解能および量子化パラメタ選択によるレート歪み特性結果を示す. また, 図 6 に, 符号化後の高時間分解能化により生成された SR フレームの例を載せる.

7.1 時間分解能選択によるレート歪み特性の最適化

初めに, 提案手法を用いた最適な時間分解能フレーム選択によるレート歪み特性を求めた. この際, 選択された全フレームにおいて, 量子化パラメタ (QP) は一定とする. 即ち, 各符号量について最適な固定 QP 値を用いて, 全フレームを符号化する. レート歪み特性について, 提案手法 (adapt) と, 単純に符号化される 2 つのケース, 全フレームについて符号化前に高時間分解能化を行う

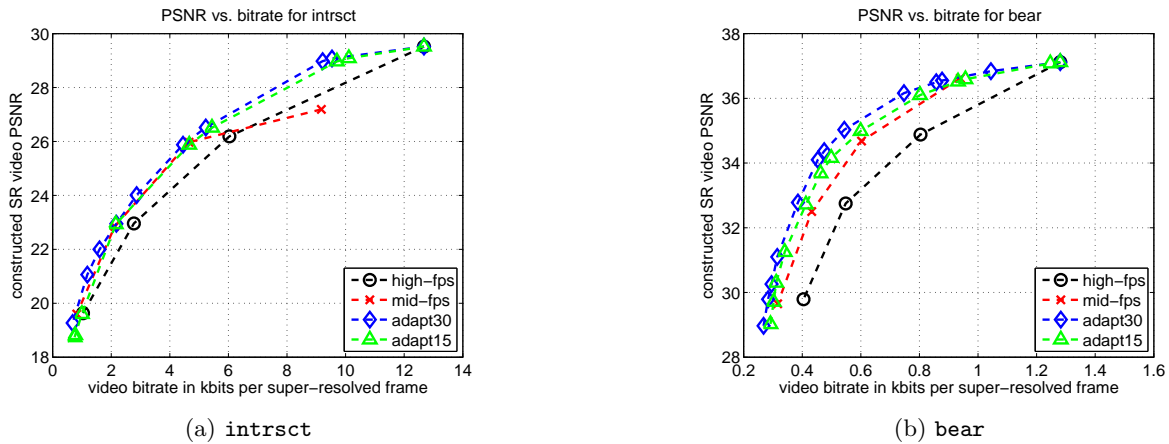


図 8 提案手法と各圧縮手法によるレート歪み特性. x 軸は生成された SR フレームの平均符号量, y 軸は映像画質 (PSNR).

場合 (high-fps) と, 撮影映像をそのまま符号化し, デコーダで高時間分解能化を行う場合 (mid-fps) との比較を行った.

図 7 に映像 intrsct のレート歪み特性を示す. 符号量の多い箇所では high-fps の方が mid-fps よりも PSNR が高く, 少ない符号量においては mid-fps の方が high-fps よりも高い. これは, high-fps の方が符号量を多く必要とするが, 高時間分解能化によって生成されるフレーム画像の画質を良好なまま伝送できることから, 直観的に理解できる. 次に, adapt は, 符号量の多い箇所では, high-fps より 1.1dB, また, mid-fps より 1.8dB 程も上回っている. これらの結果から, 異なる時間分解能のフレームを適応的に選択する提案手法の有効性が確認された.

7.2 時間分解能および量子化パラメタ選択によるレート歪み特性の最適化

次に, 2 種の映像について, 提案手法で求めた最適なフレームの時間分解能および量子化パラメタ (QP) の選択によるレート歪み特性を, 図 8 に示す. 前節と同様に, 提案手法と, 単純な 2 つの符号化手法: i) high-fps (全フレームを高時間分解能化した後に符号化), および ii) mid-fps (MR をそのまま符号化した後, 高時間分解能化) との比較を行った. high-fps と mid-fps については, レート歪み曲線が得られるよう固定 QP が調節される. デコーダの複雑性を考慮し, 高時間分解能化における自己相似性を探す領域を MR の 30 フレーム ($W = 30$) とした提案手法によるレート歪み特性を adapt30 に示す.

全ての符号量において, adapt30 は, high-fps と mid-fps のどちらよりも高い PSNR が得られている. 具体的には, PSNR 値は最大, 映像 intrsct で 1.2dB, 映像 bear では 1.3dB 程上回っている. 特に, 映像 intrsct では符号量の多い箇所, 映像 bear では符号量が中程度の箇所で, PSNR の向上が見られる. これらは, mid-fps

における MR フレームのみ, あるいは high-fps における SR フレームのみでは高性能が得られなかった部分について, 適応的な時間分解能および QP の選択が有効に働いたためと考えられる. また, 図 8(a) の全符号量において図 7 の adapt よりも上回る PSNR が得られていることから, 時間分解能のみでなく, QP を適応的に選択することの有効性も示している.

更に, デコーダにおける複雑性を考慮し, デコーダにおける制約を更に厳しく設定した実験も行った. デコーダにおける高時間分解能化時の自己相似性を探す領域を $W = 15$ に削減した場合の結果を, adapt15 に示す. レート歪み曲線から, adapt15 の性能は adapt30 よりも少し低下していることが確認できる. しかし, adapt15 の性能は high-fps および mid-fps の両方より高いことから, デコーダにおける複雑性の制約をより厳しく設定した場合にも, 提案手法は効果的であることが確認できる.

8. おわりに

本論文では, 復号後に高時間分解能映像を得るため, フレームの時間分解能と量子化パラメタ (QP) を適応的に選択することにより, レート歪み特性を最適化する手法を提案した. 提案手法は, 連続露光映像の特性である自己相似性を利用した高時間分解能化の技術を融合し, 異なる時間分解能フレームおよび QP の最適な組み合わせを求めることにより, 高い圧縮効率を実現する. 更に, 時間分解能と QP のコスト単調性に基づき高速な探索を可能にする. 実映像を用いた評価実験により, 時間分解能や QP を適応的に選択しない手法に比べて, 提案手法が PSNR で最大 1.3dB 上回ることを確認した. 今後は, フレーム画像内のブロック単位でレート歪み特性を最適化する方法について検討していく.

文 献

- [1] M. Shimano, T. Okabe, I. Sato and Y. Sato: "Video

- temporal super-resolution based on self-similarity”, The Tenth Asian Conference on Computer Vision, Queenstown, New Zealand (2010).
- [2] T. Brox, A. Bruhn, N. Papenberg and J. Weickert: “High accuracy optical flow estimation based on a theory for warping”, Proc. ECCV 2004 (LNCS3024), pp. 25–36 (2004).
 - [3] T. Chen: “Adaptive temporal interpolation using bidirectional motion estimation and compensation”, IEEE International Conference on Image Processing, Rochester, NY (2002).
 - [4] J. Zhai, K. Yu, J. Li and S. Li: “A low complexity motion compensated frame interpolation method”, IEEE International Symposium on Circuits and Ssystems, Kobe, Japan (2005).
 - [5] K. Ramchandran, A. Ortega and M. Vetterli: “Bit allocation for dependent quantization with applications to multiresolution and MPEG video coders”, IEEE Transactions on Image Processing, Vol. 3, no.5 (1994).
 - [6] S. Liu and C. J. Kuo: “Joint temporal-spatial bit allocation for video coding with dependency”, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 15, no.1, pp. 15–26 (2005).
 - [7] G. Cheung and V. Velisavljević: “Efficient bit allocation for multiview image coding & view synthesis”, IEEE International Conference on Image Processing, Hong Kong (2010).