

カラー映像通信のための領域ベース多重記述符号化の検討

会沢 隆広[†] 堀内 隆彦[‡] 富永 昌治[‡]

[†]千葉大学大学院自然科学研究科 〒263-8522 千葉県千葉市稲毛区弥生町 1-33

[‡]千葉大学大学院融合科学研究科 〒263-8522 千葉県千葉市稲毛区弥生町 1-33

E-mail: [†] aizawa_t@graduate.chiba-u.jp, [‡] {horiuchi, shoji}@faculty.chiba-u.jp

あらまし 近年、マルチホップに代表される手軽な通信形態が盛んに議論されている。これらの通信ではエラーの頻度が高いため、品質の高い再現映像を得るためには工夫が必要となる。領域ベース多重記述符号化は、それらの問題を解決する手法の一つとして提案されているが、モノクロ映像にしか適用できなかった。本研究では、カラー映像通信を対象とした多重記述符号化方式を提案し、シミュレーションに基づいた検証を行ったので報告する。

キーワード 多重記述符号化, カラー画像通信, 画像符号化, 画質改善, ウェーブレット

Domain-Based Multiple Description Coding for Color Video Communication

Takahiro AIZAWA[†] Takahiko HORIUCHI[‡] and Shoji TOMINAGA[‡]

[†] Graduate School of Science and Technology, Chiba University 1-33 Yayoi-cho, Inage-ku, Chiba-shi, Chiba, 263-8522 Japan

[‡] Graduate School of Advanced Integration Science, Chiba University 1-33 Yayoi-cho, Inage-ku, Chiba-shi, Chiba, 263-8522 Japan

E-mail: [†] aizawa_t@graduate.chiba-u.jp, [‡] {horiuchi, shoji}@faculty.chiba-u.jp

Abstract In recent years, simple and convenient communication systems such as multihop network have been discussed. Since errors frequently occur in the network, a specific technique is required to obtain high quality image and video. Domain-Based Multiple Description Coding was proposed by Bajic et al. as a solution method. However, this method can be applied to only monochrome images. In this study, we propose a Multiple Description Coding method for the purpose of transmission of color images and videos. The performance of our method is verified based on transmission simulation.

Keyword Multiple Description Coding, Color Image Transmission, Image Coding, Image Quality Improvement, Wavelet

1. はじめに

様々な情報端末の普及に伴い、ネットワークを介して情報伝達をすることが多くなった。あらゆるところにネットワークが遍在するような「ユビキタス環境」も実現されつつある。ユビキタス環境における通信形態として、アドホック通信、センサネットワーク、IPv6などが盛んに議論されている。どこでも通信が出来る環境は便利である一方、従来の通信環境と比較して、無線通信を多用するため、障害物や端末の移動に伴う経路切断や電波の干渉などによって通信の信頼性が低下するという問題がある。本研究では、このような厳しい環境下において、高品質な画像通信を行うことを目指している。従来の画像通信に関する技術は、例えば文献[1]で詳細に報告されているが、ネットワーク上でのパケット損失による画質劣化については、十分な議論がなされていない。

パケット損失が多発するネットワークにおいて、効率的な画像通信の実現を目指した概念として多重記述

符号化 (Multiple Description Coding: MDC) がある[2].

この概念は、

- (1) 画像をあらかじめ複数かつ平等な信号 (description と呼ぶ) に分割し、その description を相手に送信する。
- (2) 受信側ではそれら description を合成する。もし、一部の description が失われてもある程度品質の劣化した画像を受信することが出来る。

の2段階の処理によって、高品質な画像通信を実現するものである。Bajic らは文献[3]において、エラーに対しても高品質な画像通信の実現を目的としたウェーブレットベースの Dispersive Packetization (以下 DP) を提案している。さらに、文献[4]において、DP に対して改良を加え、領域ベースの多重記述符号化に拡張している。本稿では、この拡張された手法を DP と称す。DP は、ネットワークでのエラー特性を考慮しながら、パケットを生成する。DP と同様にウェーブレットを用いている PZW[5]と比較して、パケット損失に対し

て、受信側での画質劣化が目立たないような工夫がなされている。しかしながら、DPに基づいたMDCは、従来モノクロ画像に対してのみ議論されていた。本稿では、DPを拡張することによって、カラー画像やカラー映像を扱える方法を提案する。

提案手法は、DPを各カラー信号に適用し、エラーが頻発する環境を通過させた後、受信側で各カラー信号に対して適切なエラー隠ぺい処理を施すことによって、復元画像を得る。本稿では、エラーによる色の変化に対する検討を行い、カラー映像通信の改善法について述べる。

2章では、Bajicらによって提案されたDPについて簡単に述べる。3章ではカラー画像を扱えるDPを提案する。4章ではシミュレーションによる実験を行い、提案手法の有効性を検証する。

2. Dispersive Packetization による画像通信

2.1. 画像通信システムの概要

DPの本質は画像を複数のパケットに分割するための手法である[3]。DPを用いた画像通信システムの流れを図1に示す。ネットワーク上ではパケットの衝突や通信障害などによる種々の要因によって、パケットの消失が発生することが想定される。消失した画像情報は、エラー隠ぺいによって補間され、品質の劣化を抑える工夫が施される。次節以降で、具体的な各要素技術を説明する。

2.2. Dispersive Packetization

紙面の都合上、本節ではDPの概念のみを説明する。詳細は文献[3-4]を参照されたい。図2にDPの概念図を示す。はじめに、図2に示すように画像を多重解像度解析によりウェーブレット展開する。図では、レベル2のウェーブレット展開例を示している。文献[3-4]では、ウェーブレットのカーネルとして、7/9 Antoniniを使用している。次に各サブバンドからZerotree走査により信号を抽出する。この処理は、画像空間の特定局所領域を周波数成分ごとに抽出したことを意味している。PZW[5]では、特定の局所領域(図2下の灰色部分)の情報を1つのパケットに収納していた。これによって、画像の乱れを局所的に抑えることが可能であった。しかしながら、通信でエラーが生じた場合に、局所領域の情報が完全に失われてしまうため、エラー隠ぺいが困難な問題点があった。DPでは、複数のパケットに分散させて局所領域の情報を収納する[3]。したがって、PZWと比較してエラーが拡散する弱点はあるものの、一箇所の情報が完全に失われることが少なくなり、エラー隠ぺいが有効に働く長所がある。

さて、パケット間で局所領域の情報を分散させる際

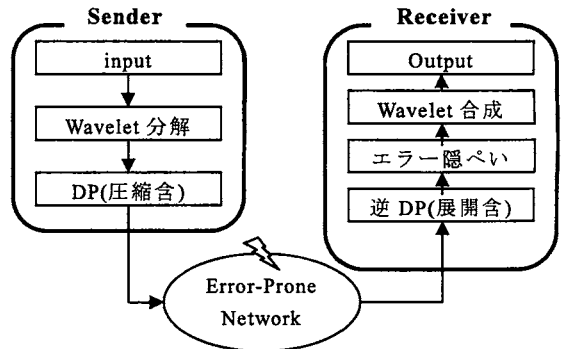
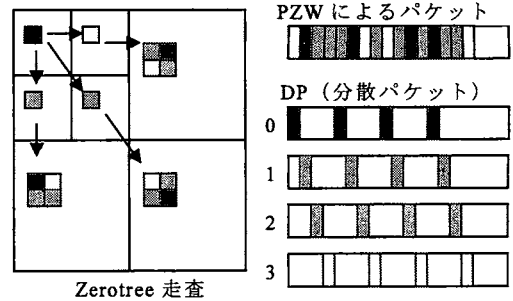


図1 DPを用いた画像通信システム



0	1	2	3	1	2	3	0	0	1	2	3	0	1	2	3
2	3	0	1	3	0	1	2	2	3	0	1	2	3	0	1
0	1	2	3	1	2	3	0	0	1	2	3	0	1	2	3
2	3	0	1	3	0	1	2	2	3	0	1	2	3	0	1
2	3	0	1	3	0	1	2	0	1	2	3	0	1	2	3
0	1	2	3	1	2	3	0	2	3	0	1	2	3	0	1
2	3	0	1	3	0	1	2	0	1	2	3	0	1	2	3
0	1	2	3	1	2	3	0	2	3	0	1	2	3	0	1
1	2	3	0	1	2	3	0	2	3	0	1	2	3	0	1
3	0	1	2	3	0	1	2	0	1	2	3	0	1	2	3
1	2	3	0	1	2	3	0	2	3	0	1	2	3	0	1
3	0	1	2	3	0	1	2	0	1	2	3	0	1	2	3
1	2	3	0	1	2	3	0	2	3	0	1	2	3	0	1
3	0	1	2	3	0	1	2	0	1	2	3	0	1	2	3

図2 DPの概念

に重要となるのは、次の2点である。(1)空間的に近隣にある画像情報は、できるだけ異なるパケットに収納する(パケットが消失しても、近隣のパケットの情報があれば補完しやすいため)、(2)空間的な、ある1箇所の画像情報(すなわち図2下の灰色部分)に対する情報は、周波数領域では出来るだけ異なるパケットに収納する(パケットが消失しても、一箇所の情報が集中して消失することがなくなる)。文献[4]ではこれら2点について検討されており、(1)については最適化問題として解決を図っている。また、(2)については、剰余シフト分割関数 $p^{(k)}: \mathbb{Z}^2 \rightarrow \{0,1,\dots,P-1\}$ を定義している。

$$p^{(k)}(\mathbf{x}) \equiv (p^{(0)}(\mathbf{x}) + k) \bmod P, k = 0,1,\dots,N-1 \quad (1)$$

ここで、 k はサブバンドの番号を表しており、 N はサブバンド数、 P は生成するパケット数である。式(1)は、低周波のサブバンド成分をどのパケットに収納するかを決定した後に、他のサブバンド信号に対するパケット分割の方法を与えている。図 2 下に、 $P=4, N=7$ のときの分割例を示す。図の中に示されている多くの数字が p であり、収納されるパケットの番号を意味している。空間的な情報に基いて分割されたパケットを、多重記述符号化における description とみなすため、一般に領域ベースの多重記述符号化と呼ばれている。

2.3. DP におけるエラー隠ぺい

パケットの消失が発生した場合、失われた情報が存在していた画像中の領域に対して、何らかの値を埋めなければいけない。その作業がエラー隠ぺい(error concealment)である。エラー隠ぺいを行うタイミングは、ウェーブレット合成を行う直前であり、サブバンドそれぞれに対して行う。

エラー隠ぺいのもっとも単純なものは、エラーで失われた領域をゼロとする方法であり、文献[3-4]ではこの方法を No Concealment(NC)と呼んでいる。また、周辺の画素情報から 2 次元的に線形補間する Bilinear Concealment(BC)が提案されている。他にも、種々の補間方法が考えられるが、文献[3]によれば、どの手法を用いてもある程度改善するが、BC と比較して劇的な改善が見られないため、演算量を考慮すると BC によるエラー隠ぺいが有用であると結論づけている。

3. カラー画像に対する Dispersive Packetization の提案

3.1. 色空間に対する検討

DP は、信号を復元することは考慮されているが、色情報の扱いに関しては全く考慮されていない。デジタルデバイスで扱われる色信号は、通常 RGB の 3 次元の信号から成り立っているため、RGB の各信号に対して独立に DP を行う方法が考えられる。しかし、各色信号を独立に復元した場合、カラーバランスが崩れることが容易に想定される。そこで、輝度-色差成分によって表現されている色空間に対して DP を行うことを提案する。本稿では、NTSC で用いられている YCrCb 色空間と、CIELAB 色空間での DP を考える。注意することは、使用する色空間によらず、エラーが存在しない場合に得られる再生画像は、量子化誤差を除いて原画像とほぼ等価になることである。したがって、エラー部分の見えの違いを比較することが重要となる。

3.2. カラー動画画像のエラー隠ぺい—Trilinear 法

文献[3]では、エラー隠ぺい方法として、動き補償を

考慮した 3 次元サブバンドビデオ符号器に対して検討されている。本稿では文献[4]で行われている、intra-frame のみからなるビデオに対してカラー化の検討を行う。intra-frame のみからなるビデオは時間的な相関を圧縮に用いていないため、一般にビットレートの点では不利であるが、inter-frame が存在しないため、1 枚のフレームを受信したらすぐに再生を開始できる特長がある。またエラーが後のフレームに伝播しない点が長所となる。特にこの長所は、エラーが多い環境下では画質改善の点で有効に働くと考えられる。

動画像に対するエラー隠ぺいは、静止画像と違って、空間的な相関に加えて、時間的なフレーム間相関を利用することができる。そこで、空間情報による 2 次元的な Bilinear Concealment に加えて、時間的にも線形補間を行う Trilinear Concealment(TC)を行うことによって、より適切な補間を行う。この時、空間的な情報と時間的な情報の比率が問題となる。この比率 $\alpha \in [0,1]$ は次式を満たすものとする。

$$TC(I(\mathbf{x},t)) = \alpha BC(I(\mathbf{x},t)) + (1-\alpha)I(\mathbf{x},t-1), (t > 1) \quad (2)$$

ここで、 \mathbf{x} は画像の座標、 $I(\cdot)$ は画素値、 $BC(\cdot)$ は線形補間、 $TC(\cdot)$ は時間軸方向の補間を加えた結果、 t はフレーム番号 ($t=0,1,\dots,n-1$) を意味する。 $t=0$ の時は、通常の BC を行う。

最適な α は、動画像のコンテンツに依存する。前のフレームとの間の動き量が大きい場合には、時間情報をあまり取り入れない方がよいと予想される。すなわち、動き量と α の間には強い相関があると考えられる。4 章では、シミュレーションによって、 α の値と再生画像との関係について考察する。

4. シミュレーションによる実験結果

4.1. 実験環境

実験に用いたカラー静止画像例として、Lenna 及び Mandrill を図 3 に示す。画像サイズはいずれも 512×512 である。また、カラー動画画像は図 4 の foreman[6]を使用した。本実験では、通信エラーのシミュレーションとして、ランダムエラーやバーストエラーを発生させることが可能な 2 状態マルコフモデルを使用した(図 5)。分割パケット数は 16、ウェーブレット展開の多重解像度のレベルは 4 とした。提案手法は、Visual Studio.NET を用いてインプリメントし、実験に用いた計算機の仕様は、Windows XP、CPU Pentium4 3.2GHz(2 CPU)、メモリ 1GB である。

送信画像と再現画像を客観的に評価するために、本稿では CIELAB 色差 ΔE_{ab} と SSIM[7]を使用した。SSIM は、Z.Wang らによって提案された視覚特性を考慮しながら画像コンテンツの構造的類似度を評価する指標であり、従来評価に使用されている PSNR よりも、実際

の見えとの相関が高いことが知られている。SSIM の値は 0~1 の値を取り、1 に近いほど 2 枚の画像は類似していることを意味する。動画像を対象とした SSIM には、Video Quality Studio ver. 0.3.5.79[8]を使用した。

4.2. 色空間の違いによる考察

はじめに、色空間の違いによる DP の結果への影響を考察する。Lenna 画像に対して、RGB, YCrCb, CIELAB の各色空間を用いて提案手法の DP を施した結果を図 6 に示す。図 5 に示したシミュレーションのためのエラーレートは $p_{GB}=0.125, p_{BG}=1$ とそれぞれ設定し、バーストエラーは用いていない。ここでは、色の乱れによる再現誤差を検証するために、評価尺度として ΔE_{ab} を用いた。

モノクロの紙面では表現が困難であるが、評価値に示されるように、RGB 色空間において DP を行うと、エラーによって色相が大きく変化することが確認された。一方、YCrCb と CIELAB 色空間における結果を比較すると、ほとんど差が見られなかった。この結果は、DP の色空間としては RGB よりも、YCrCb や CIELAB などの輝度-色差成分の色空間が適していることを意味している。演算量の観点から眺めると、YCrCb は単純な線形演算で求められるため、検証した 3 つの色空間では、YCrCb を用いるのが最適であると結論づけら

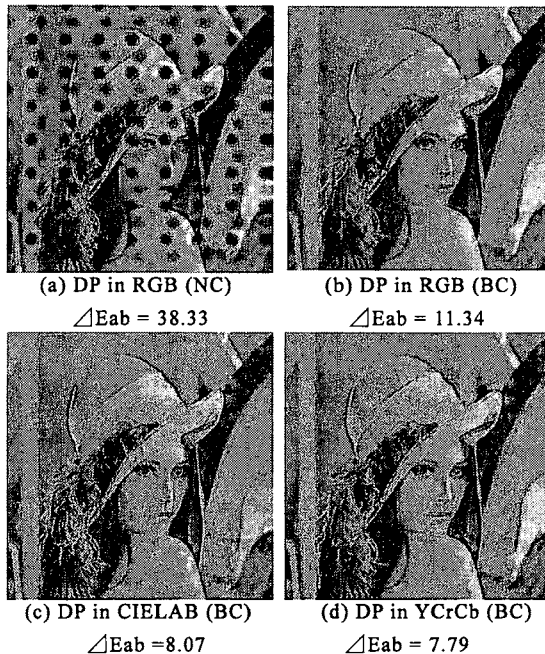


図 6 色空間による DP の違い

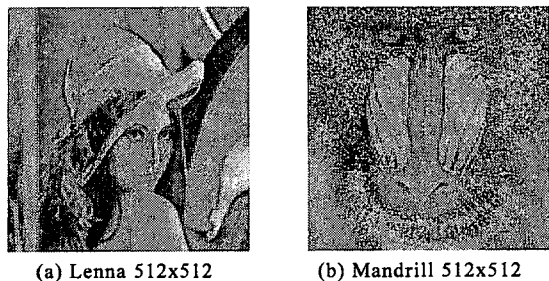


図 3 実験画像例 (カラー静止画)



図 4 実験画像 foreman (カラー動画)

Size	176 x 144 (QCIF)
Format	YUV 4:2:0
Frame rate	30fps
圧縮率	32 : 1
Bit rate	15kB/s

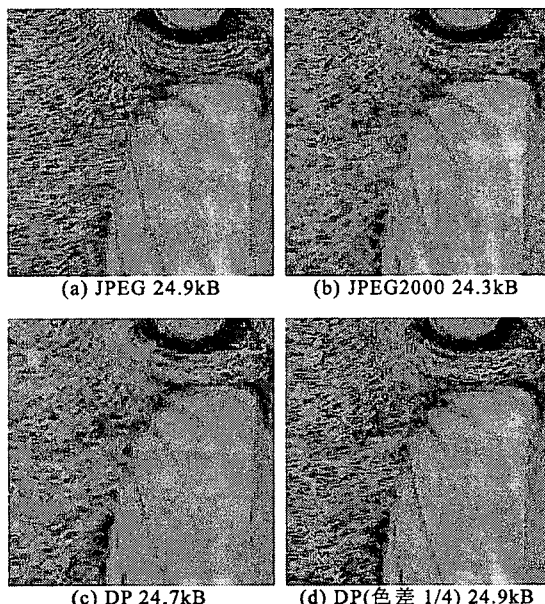


図 7 各チャンネルの重みによる DP の違い

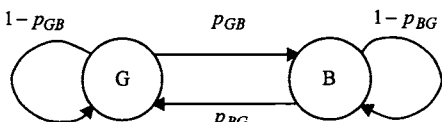


図 5 2 状態マルコフモデル

れる。この傾向は、Mandrill でも同様であった。

4.3. 他の符号化手法との考察

次に、通信における符号化の観点から、提案手法と JPEG 及び JPEG2000 との比較検証を行う。JPEG 圧縮には Photoshop CS を使用し、JPEG2000 には fnord 製の Photoshop プラグインを使用した。高周波成分を多く

含む Mandrill に対する実験結果を図 7 に示す。図 7 は、それぞれのデータ量をほぼ等しくした場合の再現画像である。図 7(a)~(c) を比較すると、提案手法の (c) は他手法と比較してボケによる劣化が目立つ結果となった。原因を考察した結果、提案手法では各色チャンネルを等価に処理していることに起因していることがわかった。一般に、人間の視覚特性は輝度に対して高い分解能を持っているため、JPEG では色差成分に対してはあらかじめ画像サイズを縦横それぞれ半分にするることによって、効率的な圧縮を実現している。そこで、提案手法においても、輝度チャンネルと色差チャンネルの解像度を変化させてみた結果が図 7(d) である。(a) および (b) と (d) を比較すると、画質に目立った差が見られない結果となった。したがって、提案手法の DP はエラーに対する隠ぺいのしやすさを備えつつ、JPEG や JPEG2000 と同等の符号化性能を有していることがわかる。

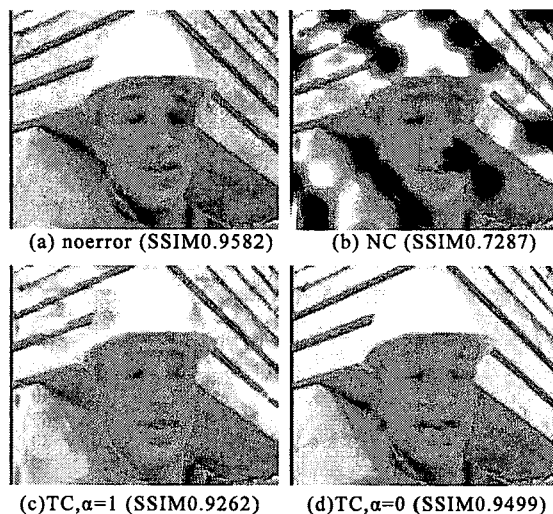


図 8 画像復元結果

4.4. カラー動画像に対する考察

次に、カラー動画像に対する考察を行う。実験に用いた foreman は、図 4 で示したように YUV 形式であり、これは RGB 信号から単純な線形演算によって求めることができる。また 4:2:0 形式であるので、図 7(d) と同様に、色差成分の信号はあらかじめ 1/4 まで間引かれている。

はじめに、エラーレートを $p_{GB} = 0.10, p_{BG} = 1$ と設定した場合に、式(2)における $\alpha = 0$ と $\alpha = 1$ (すなわち、エラーで失われた部分を完全に前のフレームで補間する場合と、時間的な情報を用いず BC のみを行う場合) に対する画像復元の結果を図 8 に示す。また全フレームにおける SSIM の評価結果を図 9 に示す。実験の結果、空間的に線形補間演算をしてエラー部分の画素を補間するより、エラーが起きた画素は前のフレームの情報でそのまま復元するだけで、大幅な画質の向上が見られることが分かる。しかし、フレーム番号 300 付近ではほとんど差が見られない。これは、フレーム番号 300 付近のシーンはカメラがパンしているため、動き量が多く、前のフレームの情報のみで復元すると、残像が強く残ってしまうためと考えられる。

そこで、適切な α の値の検討を行った。実験動画像として、背景を固定して半径 60px の円がランダムに飛び回る画像を生成した。図 10 に例を示す。背景画像は、解像度が 352×288 (CIF) の tempete[6] である。円の動き量 v をフレーム間の円の中心座標のピクセル移動量で定義する。すなわち、フレーム間の水平成分に対する円の中心座標のピクセル移動量を v_x 、垂直成分のピクセル移動量を v_y とすると、 v は式(3)で定義される。

$$v = \sqrt{v_x^2 + v_y^2} \quad (3)$$

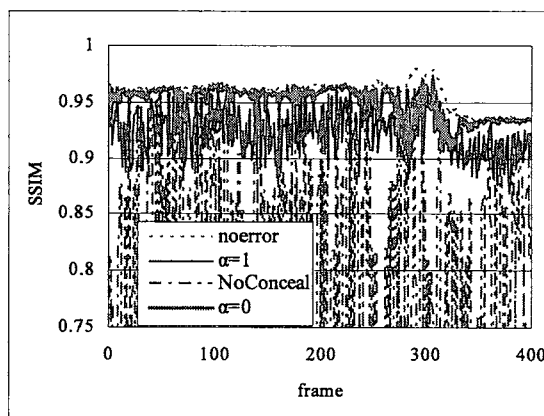


図 9 foreman 全フレーム結果



図 10 実験使用画像 (円が跳ね回る)

移動量 v が小さい時に $\alpha = 1$ として処理すると、図 11(a) の空の部分などで目立つようにエラー隠ぺいが不十分となり、逆に移動量 v が大きい時に $\alpha = 0$ とすると図 11(b) に示すようにエラー残像が発生する。円の各速度に対する α と評価値 SSIM との関係を図 12 に示す。評価値は輝度に対する 100 フレームの平均値である。グラフに示されるように、適切な α はフレーム間の画素の変化量に依存すると考えられる。円の半径を r 、と



(a) $v=10.2, \alpha=1$ (b) $v=82.46, \alpha=0$
 図 11 移動量 v と α による画質変化

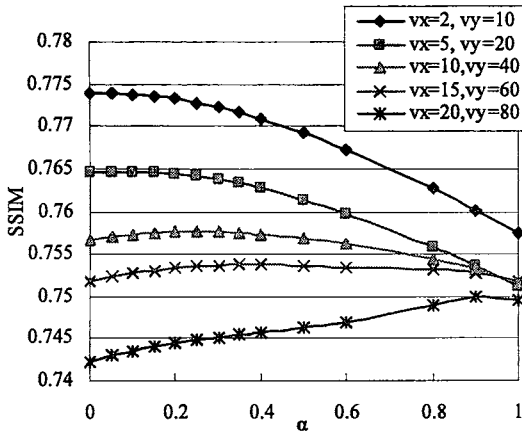


図 12 移動量 v と α による画質評価値

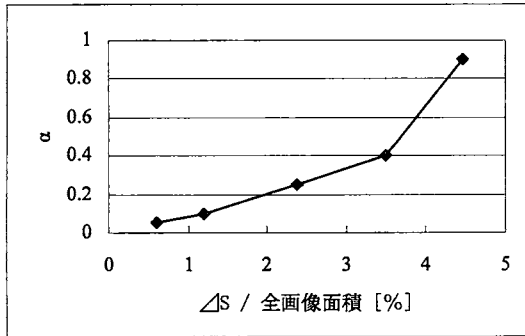


図 13 動き率と最適な α の関係

すると、1 フレームの間に画像の変化した面積 ΔS は、次式で与えられる。

$$\Delta S = \begin{cases} r^2(4\sin^{-1} \frac{v}{2r} + 2\sin(2\sin^{-1} \frac{v}{2r})), & v < 2r \\ \pi r^2, & v \geq 2r \end{cases} \quad (4)$$

この ΔS に関して、1 フレームあたりの面積 (352×288) に対する割合を横軸に、縦軸に SSIM が最大となる α の値をとったグラフを図 13 に示す。動き量と評価値を最大にする α との間には強い相関があることが分かる。したがって、前のフレームとの差分をとり、動き量を検出することが出来れば、動的に α を変更することによって、より適切な Trilinear が実現できると

考えられる。激しい動きをあまり含まない一般の動画像の場合は α をほぼ 0 としても問題がない。

5. 結論

本論文では、領域ベース多重記述符号化に用いられる Dispersive Packetization をカラー動画像に適用する方法を提案し、種々の性質を検証した。使用する色空間は輝度-色差空間が適切であり、演算量を考慮すると YCrCb が最適であった。また、カラー動画像については、画像の動きが非常に高速でない限り、前フレームの情報を利用した時間的なエラー隠ぺいを用いる手法が、空間的なエラー隠ぺいよりも大幅に画質が向上することが分かった。また、高速な動きに対する動画像エラー隠ぺい方法についても検討した。

静止画像に対して、より色の再現性を高めることや、動画像に対して、動き補償を取り入れたカラー動画像のエラー隠ぺい法の検討、及び時間軸方向の視覚特性を考慮した画像評価法の検討を行うことが今後の課題である。またアルゴリズム高速化を進める予定である。

文 献

- [1] 中島康之, “ブロードバンド・モバイルネットワークでの動画伝送・変換技術,” 電子情報通信学会誌, Vol.88, No.8, pp.666-671, Aug. 2005
- [2] V. Goyal, “Multiple description coding: Compression meets the network,” IEEE Signal Processing Magazine, Vol. 18, No. 5, pp. 74-93, Sept. 2001.
- [3] I. V. Bajic, J. W. Woods, and A. M. Chaudry, “Robust transmission of packet video through dispersive packetization and error concealment,” Proc. Packet Video Workshop (PV’2000), Cagliari, Sardinia, Italy, May 2000.
- [4] I. V. Bajic and J. W. Woods, “Domain-Based Multiple Description Coding of Images and Video”, IEEE Trans. on Image Processing, vol.12, no.11, pp.1211-1225, Oct 2003.
- [5] J. K. Rogers and P. C. Cosman, “Robust Wavelet Zerotree Image Compression with Fixed-Length Packetization,” in Proc. DCC’98, pp. 418-427, Snowbird, UT, 1998.
- [6] Video Traces Research Group, YUV Video Sequences, <http://trace.eas.asu.edu/>
- [7] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, “Image Quality Assessment: From Error Visibility to Structural Similarity,” IEEE Trans. on Image Processing, vol. 13, no. 4, pp.600-612, Apr. 2004.
- [8] VideoQualityStudio, <http://www.visumalchemia.com/vqstudio/>