

左右画像の相関による多眼3次元画像の効率的符号化

4F-2

浜村 倫行 八木原 英之 相田 仁 齊藤 忠夫

東京大学工学部

1 はじめに

近年、臨場感通信会議システムへの応用等で、3次元立体テレビへの期待が高まりつつある。しかし、3次元画像を伝送しようとする、眼数が増えるに従い莫大な帯域を消費してしまう。そこで本発表では、左右画像間の相関を積極的に利用した多眼3次元動画の効率的な符号化方式を提案し、その性能評価を行う。

2 Multiview Profile

現在ステレオ画像の符号化としては、MPEG2のMultiview Profile [1] がある。これは左眼は通常のMPEG2で符号化し、右眼は通常の動き補償だけでなく、左眼からの視差補償も行うものである。

しかし、左眼に対する右眼の圧縮率は数十%のオーダーであり、それほど大きな圧縮は達成できていない。また、多眼画像には対応していない。多眼画像では、眼数を増加させることによって、より現実に即したスムーズな立体画像を得ることができるが、これまでの符号化法だと莫大な帯域を消費してしまう。したがって、より高効率な符号化が必要とされる。

そこで、各眼画像を忠実に再現するのではなく、1眼の画像を送り、残りの画像はその画像をもとに立体に見える画像を作る、という考え方をとる。これにより、高い圧縮率を得られる符号化方式を提案する。

3 線形補間

ステレオ画像の立体視において、片眼画像の鮮鋭度を落しても最終的に知覚する立体画像の鮮鋭度はそれほど変わらないという報告がなされている [2]。これを利用して、ステレオ画像の伝送の際、片眼の解像度を落とすことで圧縮する方法が考えられている [3]。そこで以下の方法でシミュレーションを行った。

1. 左眼画像はそのままMPEG2を使って符号化する。

Coding of multi-view motion pictures by disparity compensation

Tomoyuki Hamamura, Hideyuki Yakiwara, Hitoshi Aida, Tadao Saito

Faculty of Engineering, The University of Tokyo

7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan

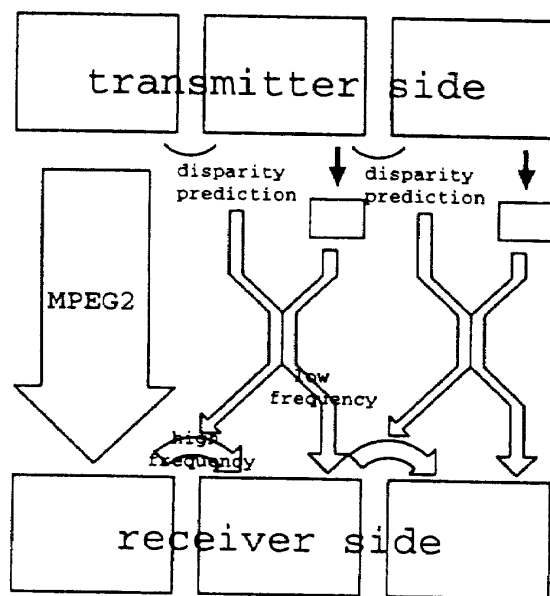


図1: DCT係数の補間

2. 右眼画像は 8×8 ブロック毎に平均値をとり、縦横 $1/8$ 解像度の画像をつくる。
3. その画像をMPEG2を使って符号化する。
4. 右眼画像は復号時に線形補間を行う。

線形補間を行うのはブロック歪みによるちらつきを抑えるためである。この方法で圧縮、復号した右眼画像は図1である。線形補間のため、全体的にはぼけた画像になっている。これは高周波成分がカットされたためと考えることができる。

4 DCT係数の補間

ステレオ画像では一般的に左右画像間の相関が非常に高い。そこで、左眼画像を参照して右眼画像の高周波成分を補間する方法を提案する。方法は以下の通りである。

1. 左眼画像はそのままMPEG2を使って符号化する。
2. 右眼画像を 8×8 ブロックに分割しそれぞれにDCT演算をする。4つのブロックをまとめてマクロブロックとする。



図 2: 原画像



図 3: 線形補間画像



図 4: DCT 係数補間画像

3. 左眼画像の探索領域の中で 16×16 のウィンドウを作り、そのウィンドウを4分割し、DCT演算を行う。
4. 右眼画像のマクロブロックと左眼画像のウィンドウで、直流成分を除いた交流成分の2乗誤差を求める。
5. 4.で求めた2乗誤差が一番小さいウィンドウを右眼画像におけるマクロブロックの対応点として、その相対位置を視差ベクトルとする。各マクロブロック毎に視差ベクトルを求める。
6. 右眼画像のブロック平均をとって縦横 $1/8$ に圧縮し、MPEG2で符号化する。そこに視差ベクトル情報を付加して伝送する。

以上が符号化である。左眼画像の中で最も高周波成分の似ているブロックを探し、その位置を視差ベクトルとして送ることで、右眼画像の高周波成分情報を補間している。

このようにして符号化したデータを復号するには以下のような手順を用いる。

1. 左眼画像と、縦横 $1/8$ にした右眼画像を復号する。
2. 送られてきた視差ベクトルを使って、右眼画像の各マクロブロック毎に該当する、左眼画像のウィンドウを特定し、DCT演算を行う。
3. DCT演算後に、その直流成分をマクロブロックの平均値に入れ替え、逆DCT演算を行う。

左眼画像における探索ではMPEGで動き補償する際に用いられている半画素精度の探索を行った。

この方法で圧縮して復号した右眼画像が図4である。明らかに画質が向上している。原画像と比べても遜色のない品質にまで良くなっていることから、線形補間とは違い、1眼から複数眼の画像を補間することもできる。

3種類のステレオ動画像（150フレーム）でのSNRを表1に示す。どの動画像においても約8dB程度上がっており、数字の上でも大きな画質向上が見られる。

	No.1	No.2	No.3
線形補間	27.42	20.20	25.07
DCT係数補間	35.42	28.65	32.93

表 1: 線形補間とDCT係数補間による復号画像のSNR [dB]

5 圧縮率

大きさ 640×480 の画像をNTSC程度の品質で伝送すると、単眼で6Mbps位であるから、各眼独立にMPEG2で圧縮するサイマルキャスト方式では(眼数) \times 6Mbps程度となる。一方、DCT係数補間によれば、1眼は6Mbps程度で送る必要があるが、残りの画像は1眼増える毎に縦横 $1/8$ の画像と視差ベクトル合わせて600kbps程度で符号化でき、全体で $6M + 600k \times (\text{眼数} - 1)$ bps程度の情報となる。眼数が増えるほど大幅な圧縮が可能である。

6 まとめ

多眼3次元画像を符号化する方式として、視差方向の相関を利用して、圧縮した画像に高周波成分を補間する方法を考案した。この方法により、画質の劣化を抑えつつ大幅な圧縮率を実現した。

参考文献

- [1] ISO 13818-2 AMD3 Multiview Profile, 1996
- [2] T.Mitsuhashi. "Subjective image position in stereoscopic TV systems - Considerations on comfortable stereoscopic images -". In Proc. SPIE, No.2179, pp.259-266,1994.
- [3] Michael G.Perkins. "Data Compression of Stereopairs". In IEEE Trans. Commun., Vol.40, pp. 684-696, April 1992.