

復号器を省略可能な動画像の符号化法*

3Q-7

大塚 雄三† 千種 康民†
東京工科大学‡

1 はじめに

近年、画像データを情報圧縮するための画像の符号化法に関する研究が数多くなされている。我々は、これまでに、多階調画像の情報圧縮法として、画像の2値化による符号化法に着目してきた。この方法による最大の利点は圧縮後のデータが人間の肉眼で認識可能なデータ構造をもつことにある。

これまでの我々の研究では、Hopfield 型のニューラルネットワーク (HNN) [1] を用いて、多階調の静止画像を2値化処理する符号化法とその復号化法を提案 [2][3] し、その有効性について示している。

本稿ではこの HNN による符号化法を動画像へ適用し、多階調の動画像を図1のように x-y-t の3次元空間と見なして、3次元2値化する。これにより符号化した2値の動画像のアニメーション表示は、人間の視覚特性によって、階調を持った画像のように認識させることが可能となる。この視覚特性を利用し、本稿では人間の視覚に復号器の役割をさせることで、復号器を事実上省略可能にする新しい動画像の符号化・復号化システムを提案する。

2 HNN による符号化と復号化

HNN による画像の符号化法において、画像中の1画素(ユニット)を一つのニューロンとし、注目ユニットの近傍ほど結合荷重が大きくなる。ユニットはランダムに抽出し、ネットワークのエネルギーが収束するまで以下の処理を繰返し行う。

$$X_i^{n+1} = \begin{cases} 1 & (\text{if } \sum_{j \neq i} W_{ij} X_j^n \geq U_i) \\ 0 & (\text{otherwise}) \end{cases}$$

ただし、 X_i^n は注目ユニット i の第 n 回目の出力、 W_{ij} は結合荷重、 U_i はしきい値とする。

3次元符号化は $n \rightarrow \infty$ によって実現される。つまり、 $\{X_i^\infty\}$ が最終的に得られる符号である。この収束したときの $\sum_{j \neq i} W_{ij} X_j^\infty$ をしきい値画像とすると、2値

画像、しきい値画像、HNN による復号処理過による

更新値の中間値をとることで、復号処理に原画像に近づくような方向性を持たせることができる。

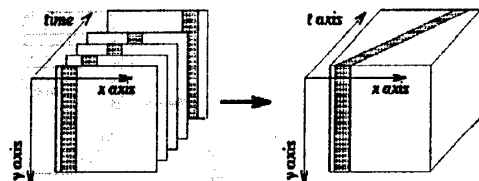


図1: 動画像の3次元処理

3 動画像への適用

動画像処理の場合、空間特性だけでなく、時間特性についても考慮する必要がある。本稿では、HNN による符号化法におけるユニットを x-y-t の3次元空間からランダムに抽出し、また、図2のように時間と空間が注目ユニットに接近しているほど結合荷重を大きくすることで、時間情報を含めた符号化処理を実現している。

この2値に収束した動画像は静止画における復号処理同様に、元の動画像に近づくように方向性を持たせて復号処理をできるが、この場合わざわざ復号化処理を行わなくても、符号化した動画像を連続表示することにより、みかけ上人間の視覚には階調が復元されているように認識される。すなわち、人間の視覚に復号処理の役割をさせることで、復号器を省略可能にする符号化法が実現できる。

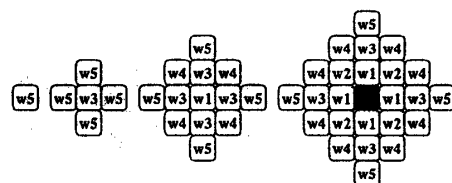


図2: 結合荷重

*Image Cording for Moving Pictures without Decoder

†Yuzo OHTSUKA, Yasutami CHIGUSA
(E-mail chigusa@cc.teu.ac.jp)

‡Tokyo Engineering University, Japan

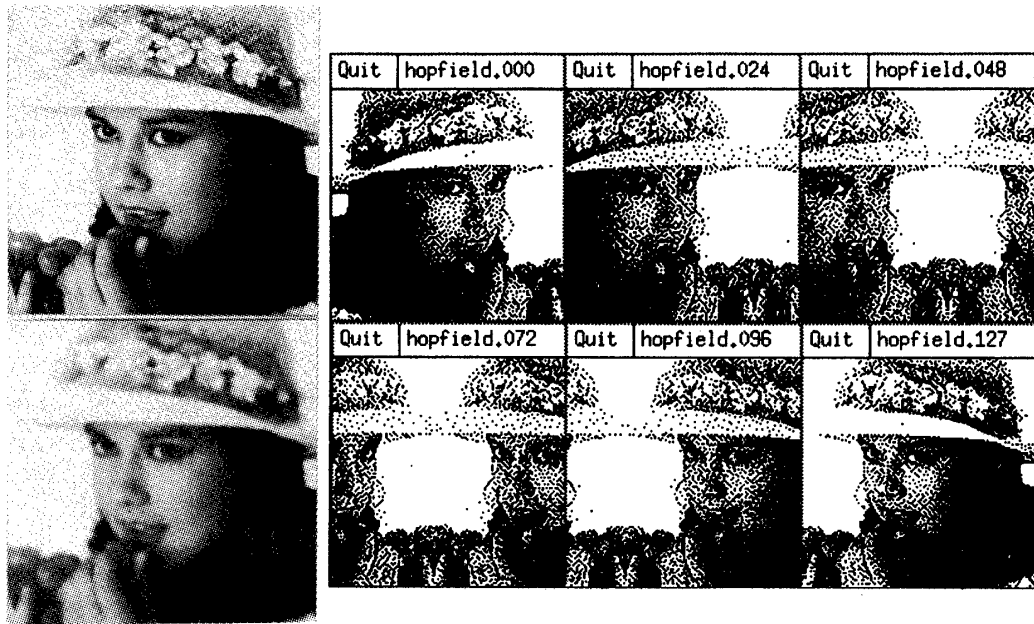


図3: シミュレーション例 (左上: 原画像, 左下: 運動視モデルによる画像, 右: 2値の動画像のサンプル)

4 シミュレーション結果

我々はこれまでにこの動画像が瞬間的に人間の目にどのように捉えられているのか van Santen and Sperling 運動視モデルを用いてシミュレーション^[4]している。ここでは、この運動視モデルによるシミュレーション画像を実際の60枚/秒で2値に符号化した動画像のアニメーション表示と比較した。運動視モデルによる画像では空間フィルタが強すぎて少しぼやけた画像となる。それに対し、実際のアニメーション表示では、より鮮明であるが、サンドストームのかかった画像になってしまう。このとき、サンドストームの影響を少なくする手段として、表示速度(60枚/秒)を増加する方法と視野角あたりの画素数を増加する方法の2つを試みたが、実際には視野角あたりの画素数を増加するほうがより効果的であった。

この運動視モデルによる画像と、実際のHNNによる符号化処理によって得られた2値の動画像を比較することは大変意義がある。これは、運動視モデルによる画像を2値に符号化した動画像の連続表示と同じように映すことで、視覚が画像をどのように処理しているのか具体的なモデルのパラメータを知ることが可能となり、人間の視覚処理に基づいて適切に結合荷重、表示速度などの値を設定することができるからである。

5 まとめ

x-y-tの3次元空間において、HNNによる符号化法は、動画像においても、2値でありながら、擬似的に中間調を表現できる時間特性を考慮した符号化法である。これによって得られた2値の動画像の連続表示は、

復号器を使用せずとも、錯視によって人間の視覚に、原画像に近い階調が再現されているように認識できることを示した。

今後の課題として、人間の視覚特性と運動視モデルについてさらに研究し、結合荷重、表示速度など、より人間の視覚処理に近づくように検討することで、サンドストームなどによるノイズの影響を少なくする必要がある。

参考文献

- [1] J.J Hopfield, "Neural network and physical-systems with emergent collective computational abilities", in Proc. Natl. Acad. Sci, USA, col. 79, pp.2554-2558, 1982.
- [2] Y. Chigusa, K. Suzuki, "An Image Reconstruction System By Neural Network with Median Filter", IEEE ISCAS '93 Chicago 1993.
- [3] 鈴木健介, 千種康民, "神経回路網にメディアンフィルタを使用した2値画から多階調画像の復元システム", 情報処理学会第46回全国大会, 4K-3, 1993.
- [4] 大塚雄三, 鈴木健介, 千種康民, "復号器の不要な動画像圧縮法", 情報処理学会第48回全国大会, 4U-4, 1994.