

誌上討論

河村, 江口, 重村氏の論文「ハフマンコード表の圧縮と
その応用」に対する意見

森 井 昌 克†

標題の論文で提案しているハフマン符号の符号語表の圧縮手法が過去に提案されている方法に比較して, 効果的でないことを指摘する。

**Comments on “Compression of Huffman Code Table
and Applications” by T. Kawamura, Y. Eguchi and
T. Shigemura**

MASAKATU MORII†

This letter points out that the compression technique of Huffman code table, which is presented by T. Kawamura and Y. Eguchi and T. Shigemura, is not efficient. Among others, we note that the efficient method has been proposed.

1. ま え が き

河村, 江口, 重村氏らの論文¹⁾ではハフマン符号の符号表(HC表)を小さな記憶量で表現する方法を開発し, 従来方法に比べて約1/4に圧縮できたと主張している。しかしながら, コンパクト木を表し, その総数を求めるために提案されているプロパーワード(proper word; PW)²⁾を利用することによって, 高々 $n-2$ ビットでハフマン符号木, すなわちHC表を表現し得ることが, すでに著者らによって提案されている³⁾。ここで n はハフマン符号化すべき文字(通報)の総数である。以下では通報 n を有するハフマン符号をサイズ n のハフマン符号と呼ぶ。

2. 本 論

森井, 今村は平面ハフマン符号(PHC)なる概念を提案し, どのようなハフマン符号に対しても, 同じ符号長(HC表のパターンの長さ)の組を有するPHCが必ず唯一存在することを証明した⁴⁾。さらに彼らはサイズ n で異なる符号長の組を有する平面ハフマン

符号の総数について考察し, EvenとLempleらが提案しているPWの総数を求める問題に等価であることを示した³⁾。すなわち, サイズ n のハフマン符号は $n-2$ ビットで表されるPWと符号長の組において, 1対1に対応することが示された。

以上の結果からサイズ n のハフマン符号語の各パターンの長さを記憶するのに要する記憶量は $n-2$ ビットであることが容易に理解できる。事実, 文献5)において, PWをHC表として, 効率的な符号化復号化を実現した方法が提案されている。

標題の文献1)ではHC表をハフマン符号語の各パターンの長さだけでなく, 「存在判定情報」を含めた記憶量で議論しているが, 本質となっているサイズ n のハフマン符号語の各パターンの長さの情報をさらにハフマン符号化した記憶量は, $n-2$ ビットに比較して, 極めて大きな値となる。したがって文献1)でのHC表の圧縮法は効果的ではない。

3. む す び

サイズ n のハフマン符号において各符号語パターンの長さの組を表すのに必要な記憶量が $n-2$ ビットであることを指摘し, 文献1)の主張が必ずしも正しくないことを示した。3氏の今後の検討を期待したい。

† 愛媛大学工学部情報工学科
Department of Computer Science, Faculty of
Engineering, Ehime University

参 考 文 献

- 1) 河村知行, 江口賢和, 重村哲至: ハフマンコード表の圧縮とその応用, 情報処理学会論文誌, Vol. 35, No. 2, pp. 267-271 (1994).
- 2) Even, S. and Lemple, A.: Generation and Enumeration of All Solutions of the Characteristic Sum Condition, *Inf. Control*, Vol. 21, pp. 476-482 (1972).
- 3) 森井昌克, 今村恭己: 符号表の組の異なるすべての2元ハルマン符号の生成について, 信学技報 IT 82-42, pp. 1-4 (1983).
- 4) 森井昌克, 今村恭己: 平面ハフマン符号の定義とその性質, 信学技報, AL 82-54, pp. 89-94 (1982).
- 5) Imamura, K. and Morii, M.: An Encoder and Decoder of Huffman-Shannon-Fano Code, *Proceedings of The VII-th Symp. on Information Theory and Its Application*, Vol. 1, pp. 61-65 (1984).

回 答

河 村 知 行[†] 江 口 賢 和[†] 重 村 哲 至[†]

Reply

TOMOYUKI KAWAMURA,[†] YOSHIKAZU EGUCHI[†] and TETSUJI SHIGEMURA[†]

ご指摘のようにサイズ n のハフマン符号のハフマン符号木は $n-2$ ビットで表現できます。しかし, HC表を用いて圧縮データを復元するためには, この $n-$

2 ビットのほかにハフマン符号と元の文字コードの対応表が必要です。この表の大きさは単純に表現すれば $n \log_2 n$ ビットです。階乗進法を用いても $\log_2 n!$ ビット程度が必要です。本論文 1) の CHT は, この意味において両方の情報を含んでいます。

[†] 徳山工業高等専門学校情報電子工学科
Department of Information and Electronics Engineering, Tokuyama College of Technology