

## 1-6 DCT領域分割による解像度変換とその評価

An Evaluation of DCT Scaling Decoder

岩橋 政宏  
Masahiro IWAHASHIジー・シー・テクノロジー(株)  
GC Technology Corporation

**Abstract-** This paper proposes objective measures to examine scaled-down picture's quality. The measures enable us to theoretically compare the performance of transform domain decimation, such as Discrete Cosine Transform (DCT) scaling decoder, with that of time domain decimation with low pass filter or subband methods. The measures were derived from analysing frequency characteristics of the system which consists of the scaling decoder and an associated encoder. Applying them to comparisons, it was confirmed that increasing the number of DCT points reduces distortion on output signal's waveform. Expressing the theoretically derived distortion in Signal to Noise Ratio (SNR) we evaluated degree of transformation distortion of digital image data.

## 1. はじめに

デジタル画像の通信や蓄積を目的としたデータ圧縮方法が近年盛んに研究、開発されている。また、解像度やアスペクト比が相互に異なるアプリケーション間での互換性も、更なる検討課題となっている。そこでこれら双方を同時に実現する一手段として、離散コサイン変換(DCT) 係数の領域分割による解像度変換方法[1][2] (以下、DSD= DCT Scaling Decoder と呼ぶ) が、最近注目を集めている。本稿ではこの解像度変換に伴う画質劣化の程度を評価する新しい尺度を提案し、システムの理論的な評価や他の変換方法との比較検討を行う。

DSDに対する理論的な考察は、離散フーリエ変換との対応関係を求めることで、K.N. Ngan により試みられた[3]。また、J.M. Adamt らは、2:1変換を例として、変換に伴う画質劣化はほとんど見られないことを報告した[4]。しかし双方とも、出力信号に重量された変換誤差の評価は主観的判断に依存している。一方、A.N. Akansu らはフィルタバンクにおけるエイリアス成分を計算し、バンクに対する理論的な評価尺度を提案した[5]。しかし彼らが対象としたフィルタバンクは、帯域合成時に分割された帯域を全て加算するので、解像度変換機能は持っていない。

そこで本報告では、DSDとこれに関連するエンコーダから構成される、システム全体の周波数特性を解析することで、本変換方法を理論的に評価する尺度を提案する。そのために、DCT処理をマルチレート信号処理システムでの畳み込み演算(フィルタバンク処理)に帰着させ、周波数領域にてシステムの伝達関数とエイリアス成分を求めた。また理想特性とシステム特性との誤差エネルギーを求め、これにパーセバルの定理を適用することで、変換に伴う誤差量を信号対雑音比(SNR)で表現した。本稿の実験例では、これらの評価関数を利用してDSDシステムの特性を評価した。また2:1変換を例として、ローパスフィルタ処理後にダウンサンプルする方法やサブバンド分割による方法など、各種の変換方法間での比較を同じ評価基準(SNR)のもとで実施したので報告する。

## 2. 準備

この節では評価尺度を導出する為の準備として、システムの入出力関係を導く。これはDCT処理をフィルタバンク処理に置き換えることで周波数領域で計算する。これと理想特性との誤差を評価尺度として次節で提案する。

## 2.1 DCTとフィルタバンクの関係

DSDは、DCTされたN個の係数のうち、低域側のM個の係数をM点逆DCTすることで、画像の解像度をN:Mに変換する(N>M)。この処理はFIR (Finite Impulse Response) フィルタによる畳み込み演算(フィルタバンク処理)に帰着される[6]。即ち、N点DCTを行うことは、Nタップの分割フィルタ  $H_i(z)$  でN帯域に分割した後、N:1にダウンサンプルすることに等しく、その後、低域側のM点の係数をM点逆DCTすることは、各帯域信号をそれぞれ、1:MにアップサンプルしてからMタップの合成フィルタ  $G_i(z)$  で処理して全帯域を加算することに等しい(図1参照)。このときDCT係数と各フィルタ係数との関係は、

$$\text{分割フィルタ: } h_i(j) = e^{-j\omega_i(N-j)}, \quad i = 0, 1, \dots, N-1 \quad (1)$$

$$\text{合成フィルタ: } g_i(j) = e^{j\omega_i j}, \quad i = 0, 1, \dots, M-1 \quad (2)$$

$$\text{DCT係数: } e_{i,j} = \sqrt{2/N} d(i) \cos[i(2j+1)\pi/2N] \quad (3)$$

$$e'_{i,j} = \alpha \sqrt{2/M} d(i) \cos[i(2j+1)\pi/2M] \quad (4)$$

但し、 $d(0)=1/\sqrt{2}$ 、 $i \neq 0$ として  $d(i)=1$ 、 $\alpha$ は利得調整パラメータ、として与えられる。またフィルタの伝達関数は、係数を  $x(n)$ 、タップ数を  $L$ と置くと、

$$X(z) = \sum_{n=0}^{L-1} x(n) z^{-n} \quad (5)$$

で与えられ、直線位相特性を持つ。以上の結果から、各フィルタが図2、図3に示す様な帯域分割特性を持つことが分かる。

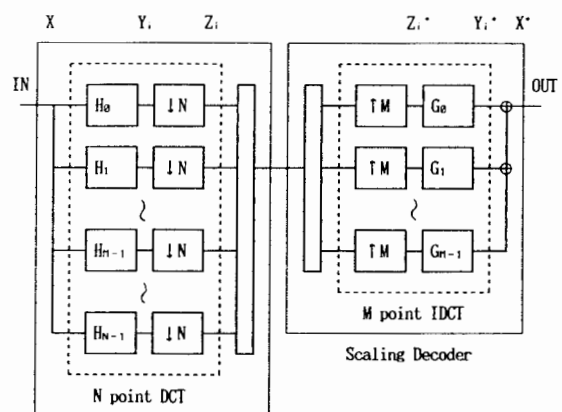


Fig.1 A structure of encoding and decoding system.  $H_i$  and  $G_i$  are analysis and synthesis filters in the  $i$ -th channel, respectively.

2.2 変換誤差の要因

このようなマルチレート信号処理では、各帯域信号が完全に帯域制限されていることが理想的であるが、図2、3から分かるように、有限のタップ長では帯域間干渉がある。従って  $N \neq M$  の条件下では、ダウンサンプルの結果として発生したエイリアスが帯域合成時に消滅する保証はない。また、システムの周波数振幅特性が全域通過特性になるとも限らない。一方、変換による波形歪みを除去する為には、システムがエイリアスのない全域通過特性を持つことが望ましい。こうした理想特性からのずれが変換による画質劣化の要因であると考えられる。但し、 $N=M$  の場合にはエイリアスが完全にキャンセルし、なおかつシステムが全域通過特性を持つので画質は劣化しない。

2.3 システムの入出力関係

そこでこれらの変換誤差を解析する為の準備として、システム内部での信号処理を  $z$  変換領域で記述する。但し、各フィルタは直線位相特性を持つので、以下、簡単のため振幅特性のみを示す。図1の各ステージにおける第  $i$  帯域の信号  $y_i(n)$ ,  $z_i(n)$ ,  $z_i'(n)$ ,  $y_i'(n)$  は、それらの  $z$  変換  $Y_i(z)$ ,  $Z_i(z)$ ,  $Z_i'(z)$ ,  $Y_i'(z)$  により以下の式で関係付けられる。但し、ここで  $z$  変換は、式(5)において、 $L$  を信号の長さ、 $x(n)$  を離散時間信号と解釈したものを意味する。

- a) 分割フィルタによる入力信号の帯域分割  

$$Y_i(z) = H_i(z) X(z), \quad i = 0, 1, \dots, N-1, \quad (6)$$
- b)  $N : 1$  ダウンサンプリング  

$$Z_i(z) = 1/N \sum_{k=0}^{N-1} Y_i(z^{1/N} e^{2\pi k j / N}) e^{2\pi k p j / N} \quad (7)$$

但し  $p = (N-1)/2$  (decimation timing)
- c)  $1 : M$  アップサンプリング  

$$Z_i'(z) = Z_i(z^M), \quad i = 0, 1, \dots, M-1, \quad (8)$$
- d) 合成フィルタによる不用帯域除去  

$$Y_i'(z) = G_i(z) Z_i'(z), \quad i = 0, 1, \dots, M-1, \quad (9)$$
- e) 各帯域の加算  

$$X^*(z) = \sum_{i=0}^{M-1} Y_i'(z) \quad (10)$$

以上をまとめると、

$$X^*(z) = 1/N \sum_{i=0}^{N-1} G_i(z) \sum_{k=0}^{N-1} H_i(v) X(v) e^{2\pi k p j / N} \quad (11)$$

但し  $v = z^{M/N} e^{2\pi k j / N}$

となる。こうして、入力信号  $X(z)$  と出力信号  $X^*(z)$  の関係が、分割フィルタ  $H_i(z)$  と合成フィルタ  $G_i(z)$  をパラメータとして  $z$  変換領域に於て関係付けられた。

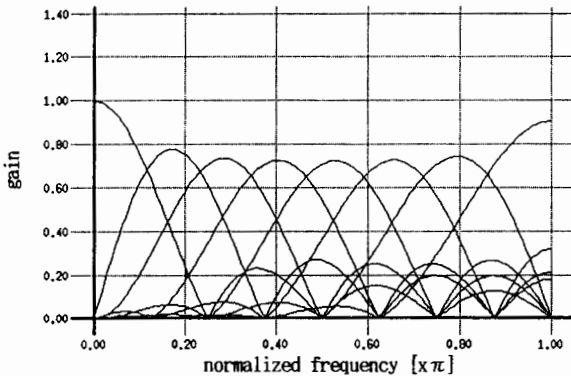


Fig.2 Frequency amplitude characteristics of DCT analysis filters.  $N=8$ .

3. 提案する評価尺度

以上の結果を用いることで、システムの周波数特性が容易に導出できる。実際、上記式(11)を、 $k=0$  の項とその他の項に分離してから、 $z=e^{j\omega}$  を代入すると、

$$X^*(e^{j\omega}) = [ A(e^{j\omega}) X(e^{j\omega M/N}) + B(e^{j\omega}) ] \quad (12)$$

但し、

$$A(e^{j\omega}) = 1/N \sum_{i=0}^{N-1} G_i(e^{j\omega}) H_i(e^{j\omega M/N}) \quad (12.a)$$

$$B(e^{j\omega}) = 1/N \sum_{i=0}^{N-1} G_i(e^{j\omega}) \sum_{k=1}^{N-1} H_i(u) X(u) e^{2\pi k p j / N} \quad (12.b)$$

但し  $u = e^{j(\omega M + 2\pi k j / N)}$

を得る。ここで、 $A(e^{j\omega})$  はシステム全体の伝達関数、 $B(e^{j\omega})$  は帯域間の干渉成分であるエイリアスを表す。こうして得られた関数が先に述べた変換誤差の要因を表している。例えば、 $N=M$  の場合には  $A(e^{j\omega})=1$ ,  $B(e^{j\omega})=0$  となり、システム全体がエイリアスのない完全なオールパス特性を持つことが確認できる。一方、 $N \neq M$  の場合にはこうした理想特性とは若干異なってくる。この誤差を調べることでシステムの評価が可能となる。

更に、例えば符号化時の量子化のような他の要因で発生する誤差との比較を行う為には、これらの変換誤差を、画質評価の指標として広く使われている SNR で表す。そこで理想的な変換の結果である  $X(e^{j\omega M/N})$  と、上記式(12)との誤差エネルギー密度を次式、

$$E_{power} = N/(M\pi) \int_0^{2\pi} [X(e^{j\omega M/N}) - X^*(e^{j\omega M/N})]^2 d\omega \quad (13)$$

$$ESNR = -10 \log_{10} E_{power} \quad (14)$$

で定義する。但し、積分の範囲は  $0 < \omega < \pi M/N$  とする。尚、ここでは誤差エネルギーを周波数領域で計算したが、パーセバルの定理によれば、これは空間領域で計算した値と等しくなることが保証されている[7]。以上、本稿では式(12.a)、(12.b)、(13)、(14)で定義された評価関数を評価尺度として利用することを提案する。

4. 実験結果と考察

本節では、前節で定義した評価関数を応用して各種方式を比較検討する。

4.1 DCT点数と画質劣化量の関係

2 : 1 変換を例として、このときの  $A(e^{j\omega})$ ,  $B(e^{j\omega})$  をそれぞれ図6a、図6bに示した。但し入力信号のスペクトルは  $X(e^{j\omega})=1$  と仮定した。図からは DCT 点数を上げるほど  $A(e^{j\omega})$  がオールパス特性に、 $B(e^{j\omega})$  が 0 に近づき、特に高域の画質劣化が抑えられることが確認できる。またこのことは平均誤差エネルギーを示す図5aや図5bからも確認出来る。同図では入力信号に自然画像を用いている。図5aでは縦、256ラインの一次元信号に256点高速フーリエ変換を適用してスペクトルを計算し、1ラインあたりの平均 ESNR を求めた。このとき、ESNR の値に  $\pm 9$  dB 程度の偏差があったことに注意する。また SNR が DCT 点数の増加に対して一様増加していないのは、その値が入力スペクトルと各フィルタ特性とのマッチングに依存するためと考えられる。

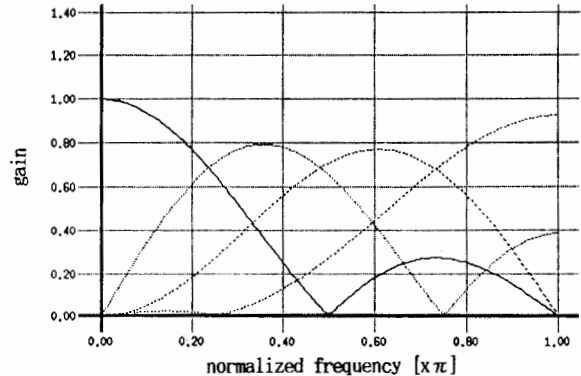


Fig.3 Frequency amplitude characteristics of DCT synthesis filters.  $M=4$ .

4.2 他の解像度変換方法との比較

図7にはローパスフィルタ処理後にダウンサンプルする方法や、サブバンドで2分割してから低域信号のみを再生する方法の例を示した。図6との比較から、4タップのSSKF (Symmetric Short Kernel Filter) [8] を用いたサブバンド分割による変換よりも4点DCTと2点逆DCTとの組合せによるDSDの方が良い特性を示すことが分かる。

4.3 変換比と誤差量の関係

図4には8:M (8 point DCT + M point IDCT) の変換を自然画像に対して行った場合の平均誤差エネルギーを示した。変換比率を100%に近づけるほど誤差が少なくなることが確認できる。

4.4 符号化誤差と変換誤差の割合

表1には符号化誤差量と解像度変換による誤差量とを別々に示した。表中の coding error はフルサイズの復号画像に重畳する誤差を表す。値はフレーム数で平均しており、符号化誤差のおおよそ値を示す。もう一方の error は1フレームの解像度変換誤差を表す。表からは、横方向に2:1に縮小した場合の変換誤差  $E_{power}$  は、伝送レートが4Mbpsでの符号化誤差の半分程度であることが分かる。横方向の縮小に比べて、縦方向の変換誤差が大きいのは、入力画像が飛び越し走査された動画像であることによる。

Table.1 Comparison between coding error and transformation error.

	coding error		2:1 transformation error		
	4Mbps	9Mbps	horizontal	vertical	total
$E_{SNR}$ [dB]	29	33	42	29	27
$E_{power}$ [ $\times 10^{-3}$ ]	1.25	0.50	0.63	1.25	1.88

sequence: flower garden, coding: Simulation Model 3 (N=12,M=3)

4.5 主観評価

8点DCTにより縦、横ともに2:1に縮小した画像では、その画質は鮮明であり、これ以上、何等かのフィルタ処理を施す必要はないという印象を受けた。ただ、若干、エイリアスと思われるチラつきが見られた。符号化誤差に比べれば概して目立ちにくい。またブロックノイズのような非ランダムな画質劣化は見られなかった。

8:6変換や8:7変換の場合、または、縦、横ともに2:1に縮小する際に3点DCTを使った場合には、変換誤差はほとんど見えなかった。

5. まとめ

DCT係数の領域分割による解像度変換方法について、変換に伴う画質劣化量を理論的に導出し、システムの評価尺度を提案した。この導出はDCT処理をフィルタバンク処理に帰着させ、システムの入出力関係を周波数領域で算出することで実施した。

本評価法の提案で、システムの理論的な評価や比較が可能となる。また、DSDと、サブバンド分割法やローパスフィルタ法とを、同一のフィルタバンク処理として扱ったため、両者の統一的な比較も行える。更にパーセバルの定理を用いて誤差量をSNRで表示したので、発生要因の異なる誤差間での比較が容易にできる。

実験例では提案した評価基準を利用して各種方法を比較検討し、4節に記した結論を得た。また評価尺度の妥当性を検証する為、実際に処理画像を見て主観評価した結果も併記した。

今後はこれらの評価法を用いてDSDの改善方法を検討する。

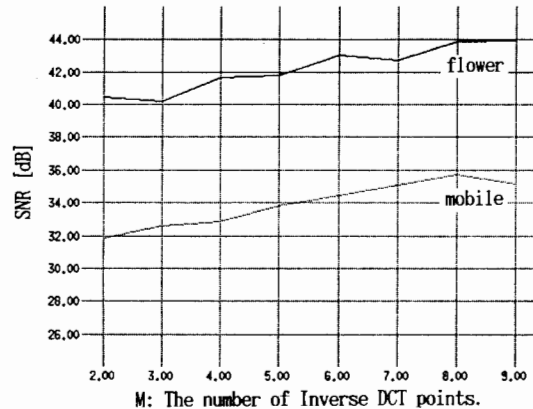


Fig.5a Averaged transformation error for images. (2:1, horizontal)

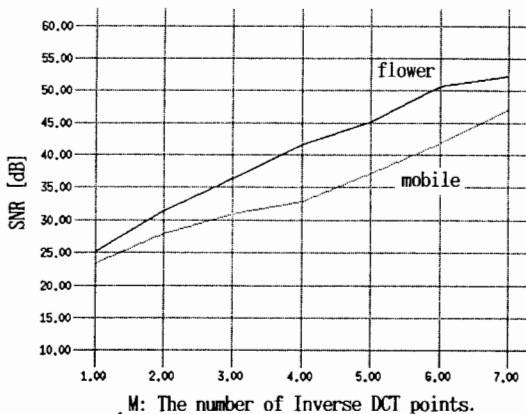


Fig.4 Averaged transformation error for images. (8:M, horizontal)

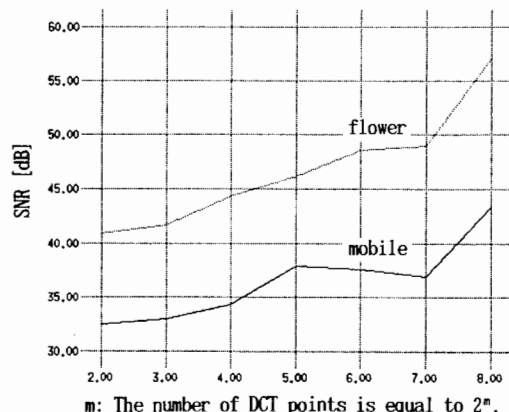


Fig.5b Transformation error of a spectrum of images. (2:1)

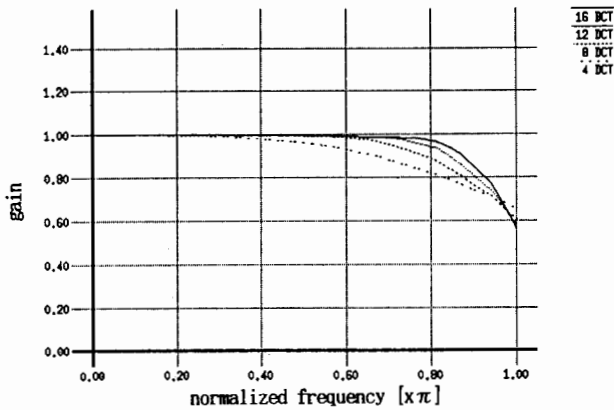


Fig. 6a Frequency characteristics of  $A(e^{j\omega})$  of DCT scaling decoder.

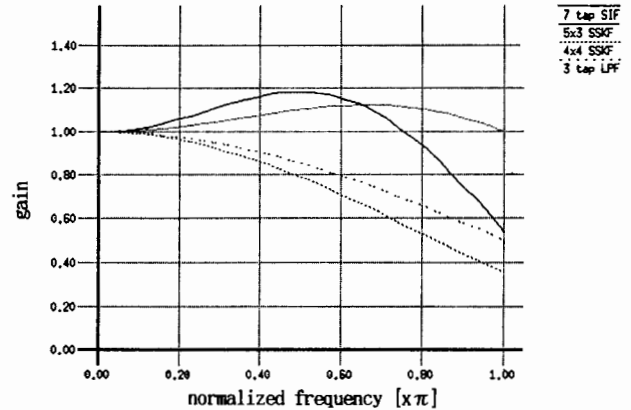


Fig. 7a Frequency characteristics of  $A(e^{j\omega})$  of some other methods.

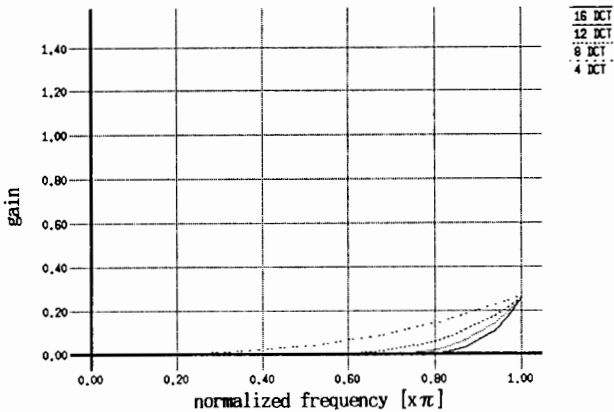


Fig. 6b Frequency characteristics of  $B(e^{j\omega})$  of DCT scaling decoder.  $X(e^{j\omega})$  is assumed to be 1 for all  $\omega$ .

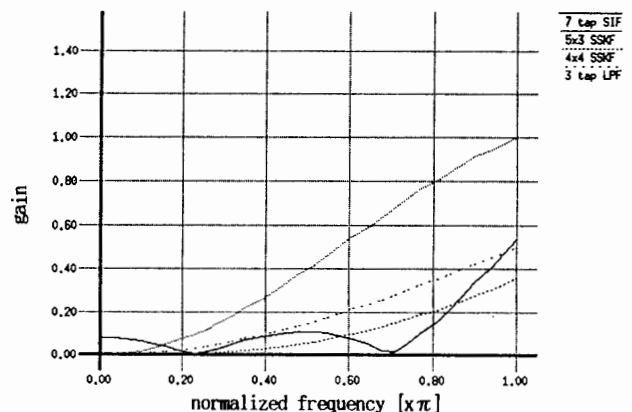


Fig. 7b Frequency characteristics of  $B(e^{j\omega})$  of some other methods.  $X(e^{j\omega})$  is assumed to be 1 for all  $\omega$ .

4x4 SSKF:  $(1+3z^{-1}+3z^{-2}+z^{-3})/8$ , 5x3 SSKF:  $(-1+2z^{-1}+6z^{-2}+2z^{-3}-z^{-4})/8$   
 LPF:  $(1+2z^{-1}+z^{-2})/4$ , SIF:  $(-29+88z^{-2}+138z^{-3}+88z^{-4}-29z^{-6})/256$

6. 参考文献

- [1] P. Delogne, O. Poncin, B. van Gaille, L. Vandendorpe, "Compatible Coding of Digital Interlaced HDTV," Proc. ISCAS' 92, pp.196-199, (May 1992)
- [2] Y. Yashima, H. Sakai and K. Sawada, "HDTV / Standard TV compatible coding based on DCT," Proc. ISCAS' 90, pp.1895-1898 (May 1990)
- [3] K. N. Ngan, "Experiments on Two-Dimensional Decimation in Time and Orthogonal Transform Domain," Signal Process., vol.11, pp. 249-263, 1986
- [4] J. M. Adant, P. Delogne, E. Lasker, B. Marq, L. Stroonants and L. Vandendorpe, "Block Operation in Digital Signal Processing with Application to A Measure of Aliasing Energy in Multiresolution Signal Decomposition," ISCAS' 92, pp.621-624, (May 1992)
- [5] J. W. Woods, "Subband Image Coding," Kluwer Academic Publishers (1991)
- [6] Rabiner and Gold "Theory and application of Digital Signal Processing", Prentice-Hall, INC., Englewood Cliffs, New Jersey (1975)
- [7] D. LeGall and A. Tabatabai, "Subband Coding of Digital Images using Symmetric Short Kernel Filters and Arithmetic Coding Techniques," Proc. ICASSP' 87, vol. 3, pp.1378-1381

ジー・シー・テクノロジー(株)  
アルゴリズム研究部



東京都 港区 南青山 7-1-5 3F南青山 6F 〒107  
 Telephone: 03-3498-7141, Facsimile: 03-3498-5224  
 Videophone: 03-5485-7551, e-mail: iwa@gctech.co.jp