

## QUAN HỆ MỐI GIỮA ĐẶC TRƯNG ẢNH VÀ CHẤT LƯỢNG ẢNH NÉN CÓ TỔN HAO

Lê Quang Tuấn\*, Vũ Đình Thành\*, Nguyễn Kim Sách\*\*

(\*) Khoa Điện & Điện Tử, Trường Đại Học Bách Khoa – ĐHQG-HCM

(\*\*) Trung Tâm Nghiên cứu Ứng Dụng Khoa học Truyền Hình

(Bài nhận ngày 24 tháng 6 năm 2002)

**TÓM TẮT:** Hiện nay, chúng ta thường nén các loại ảnh có đặc trưng rất khác nhau như ảnh tự nhiên, ảnh tổng hợp từ máy tính, ảnh y học... Với cùng một bộ nén có méo, giá trị tỷ số tín hiệu trên nhiễu đỉnh (Peak Signal to Noise-PSNR) có thể thay đổi rất lớn, đến 20-30dB, tùy theo đặc trưng của các loại ảnh. Làm sao chúng ta có thể thiết lập quan hệ giữa các loại ảnh khác nhau và chất lượng nén ảnh, thể hiện bởi tỉ số PSNR? Trong bài báo này chúng tôi trình bày một số giá trị đo đặc trưng ảnh CI (characteristic of image), và thiết lập hàm quan hệ toán học xác định với PSNR.

### 1. GIỚI THIỆU

Nhiều loại ảnh khác nhau (ảnh tự nhiên, ảnh tổng hợp, ảnh y học, ảnh kết hợp của nhiều loại ...) có đặc tính miền không gian khác nhau. Chúng ta quan sát và thấy rằng nén nhiều loại ảnh khác nhau với cùng một bộ nén ( dùng biến đổi wavelets, hoặc biến đổi DCT, hoặc biến đổi phủ lấp trực chuẩn-LOT, hoặc biến đổi phủ lấp lưỡng trực chuẩn-LBT ) và với cùng tốc độ nén bit (bits per pixel – bpp), giá trị PSNR sẽ thay đổi trong một phạm vi rộng đến 30 dB. Giá trị PSNR thay đổi trong phạm vi rộng như vậy phụ thuộc vào đặc trưng vốn có tự nhiên của các loại ảnh. Vấn đề đặt ra là giá trị PSNR thay đổi phụ thuộc vào đặc trưng nào của ảnh. Khi biết quan hệ giữa PSNR và đặc trưng, chúng ta có thể dự kiến trước giá trị PSNR của một loại ảnh xác định mà không cần nén.

Một ảnh gồm nhiều vùng có đặc điểm khác nhau về độ sáng như : vùng trơn có mức độ chói khá đồng đều, vùng nhiều đường biên, kết cấu (textures) có độ chói thay đổi lớn giữa các điểm ảnh lân cận. Chúng ta gọi những vùng ảnh tích cực là những miền ảnh có ảnh hưởng nhiều đến kết quả nén ảnh. Đó chính là những miền ảnh có nhiều đường biên sắc nét, đường kết cấu và những miền ảnh trơn rộng. Ảnh có nhiều vùng tích cực sẽ cho PSNR thấp với cùng bộ nén và cùng bpp. Mặt khác, người ta thấy rằng trong nén ảnh, phần lớn méo và biến dạng đều phát sinh ra từ các vùng tích cực và mắt người dễ dàng nhận thấy được trên ảnh giải nén.

Để định lượng các miền tích cực trong một ảnh, người ta thường dùng những số đo thống kê như giá trị trung bình, giá trị median, dải động, lệch chuẩn... của độ chói một ảnh. Chúng tôi ký hiệu số đo thể hiện đặc trưng ảnh là CI. Số đo đặc trưng ảnh CI thể hiện đặc trưng tự nhiên của ảnh và đồng thời thể hiện độ phức tạp trong quá trình nén ảnh có tổn hao. Do đó, số đo đặc trưng ảnh CI có mối liên hệ chặt chẽ với PSNR.

### 2. SỐ ĐO ĐẶC TRƯNG ẢNH CI

#### 2.1. Số đo CI theo thống kê ảnh [1]

Tính chất thống kê của ảnh được đánh giá bằng các tham số : trị trung bình, median, dải động, độ lệch chuẩn hay phương sai... Số đo thống kê bậc 1 như giá trị trung bình và median

của ảnh, đặc trưng về độ chói trung bình của mức xám ( các điểm ảnh trong một vùng ảnh). Số đo thống kê bậc 2 như độ lệch chuẩn đặc trưng độ tương phản mức xám và có thể mô tả bằng tính "trơn" tương đối của ảnh.

Chúng tôi ký hiệu CI1 là số đo đặc trưng cho ảnh tương ứng với số đo thống kê bậc 1. Trong bài này, CI1 là median của ảnh. CI2 tương ứng với số đo thống kê bậc 2; ở đây CI2 là độ lệch chuẩn của ảnh.

## 2.2. Số đo CI theo gradient các điểm ảnh

Người ta thường dùng các thuật toán phân đoạn ảnh khác nhau như Sobel, Canny, Prewitt, Roberts, Marr-Hildreth [2] ... để trích các đặc trưng của ảnh. Đó chính là tin tức về biên ảnh. Thuật toán phân đoạn ảnh bằng gradient điểm ảnh được thực hiện theo chiều ngang, chiều dọc hoặc cả hai tại các điểm biên ảnh. Chúng tôi sử dụng tin tức về biên ảnh có được từ thuật toán Sobel, làm số đo CI3. Công thức tính CI3 là :

$$CI3 = 100 * \left( \frac{1}{H * V} \sum_{i=1}^{H*V} K(i) \right) \quad (1)$$

trong đó H là chiều dài và V là chiều cao của ảnh, tính theo đơn vị điểm ảnh (pixel). K(i) là giá trị 1 tại đường biên và 0 ở ngoài biên.

Trong bài báo [3], tác giả đã sử dụng gradient của mọi điểm ảnh để xác định đặc trưng ảnh. Chúng tôi ký hiệu CI4 là số đo đặc trưng cho ảnh tương ứng với gradient của tất cả điểm ảnh trên một ảnh. Công thức tính CI4 là :

$$CI4 = \frac{1}{(H * V)} \left[ \sum_{i=1}^{V-1} \sum_{j=1}^H |I(i, j) - I(i+1, j)| + \sum_{i=1}^V \sum_{j=1}^{H-1} |I(i, j) - I(i, j+1)| \right] \quad (2)$$

trong đó H là chiều dài và V là chiều cao của ảnh, tính theo đơn vị điểm ảnh (pixel). I(i,j) là giá trị mức xám của điểm ảnh tại vị trí (i,j) trên ảnh.

Tuy nhiên, tác giả bài báo [3] đã sử dụng số đo trên để điều khiển tốc độ nén bit của video, không nghiên cứu tương quan giữa CI4 với đặc trưng của các loại ảnh khác nhau. Đó là điểm khác nhau cơ bản so với bài này.

## 3. PHƯƠNG PHÁP ĐÁNH GIÁ QUAN HỆ ĐẶC TRƯNG ẢNH VÀ PSNR

Với tỉ số nén bit bpp (bit per pixel) không đổi và cùng một bộ nén ảnh (có méo), ảnh có nhiều miền tích cực sẽ cho kết quả PSNR thấp hơn. Chúng ta có 04 số đo CI (CI1, CI2, CI3, CI4) đánh giá miền tích cực trên ảnh. Vấn đề đặt ra là xác định được số đo CI nào có quan hệ với PSNR.

Chúng tôi đã sử dụng một bộ gồm 68 ảnh mức xám kích thước 512x512, bao gồm nhiều loại để thử nghiệm tìm quan hệ giữa 02 số đo CI và PSNR. Tập ảnh thử được lấy từ cơ sở dữ liệu ảnh mẫu [4].

Chuẩn nén ảnh phổ biến nhất là JPEG. JPEG sử dụng biến đổi DCT. DCT có khuyết điểm tạo ra méo "ô vuông" trên ảnh giải nén khi nén với tỉ số lớn (bpp < 0.5). Hiện nay, các biến đổi NLOT [5], LOT [6] và LBT [7], [8] được phát triển từ DCT nhằm khắc phục nhược điểm của DCT. Mặt khác, biến đổi wavelets được phát triển mạnh tạo nên những bộ nén ảnh chất lượng vượt bộ nén ảnh dùng DCT. Tuy nhiên, bộ nén wavelets có thuật toán phức tạp hơn nhiều so với bộ nén dùng LOT hay LBT, với cùng hiệu quả và chất lượng nén. Đối với biến đổi wavelets, chúng tôi sử dụng bộ nén ảnh waveletes hiệu quả nhất hiện nay là

SPIHT(Set Partitioning in Hierarchical Trees) của Amir Said & William A. Pearlman [9], [10]. Bộ nén ảnh wavelets SPIHT sử dụng wavelet Cohen-Daubechies-Feauveau (CDF) 9/7 hoặc wavelets S+P, với thuật toán Phân hoạch Cây Phân cấp(SPIHT) có kết hợp với mã hoá thống kê số học (arithmetic coding). Sử dụng thuật toán phức tạp SPIHT, cần phải có trong bộ nén ảnh wavelets, đã cho kết quả chất lượng nén vượt trội so với bộ nén ảnh DCT (hay LOT,LBT).

Chúng tôi dùng bộ nén với biến đổi DCT và biến đổi LOT, LBT của Malvar [6][7][8]. Bộ nén ảnh DCT, LOT,LBT của Malvar đơn giản gồm biến đổi DCT(hoặc LOT,LBT) kết hợp với bộ lượng tử vô hướng bình thường, không có mã hoá thống kê Huffman ( hoặc mã hoá số học ) và không sử dụng thuật toán phức tạp nào cả. Gần đây, một số bộ nén dùng biến đổi LOT, LBT có kết hợp thuật toán phức tạp hơn sẽ cho kết quả hiệu quả và chất lượng nén ảnh tương tự như bộ nén ảnh wavelet SPIHT. Trong quá trình nén có mẹo dùng biến đổi DCT(hoặc LOT,LBT), vấn đề tính PSNR chính xác tương ứng với bpp cố định cần được giải quyết, để có kết quả đánh giá đặc trưng ảnh không bị sai lệch. Với giá trị bước lượng tử cho trước, chúng ta không thể có được bpp cố định sau khi nén. Vì vậy giá trị PSNR tương ứng với bpp không đổi sẽ được xác định từ phương pháp nội suy. Giá trị bpp cố định được chọn trong thử nghiệm là 0,5.

#### 4. KẾT QUẢ THỬ NGHIỆM

- Hình 1a) b) c) d) trình bày một số ảnh mẫu điển hình từ ảnh “trơn” được tạo từ máy tính như ‘slope.bmp’ đến các ảnh tự nhiên có vùng ảnh kết cấu và biên tăng dần như ‘lena.bmp’, ‘barbara.bmp’ và ‘baboon.bmp’. Mỗi ảnh có trình bày các số đo CI1 đến CI4 để dễ dàng so sánh đặc điểm các loại ảnh khác nhau. Cần lưu ý các ảnh tự nhiên ‘lena.bmp’, ‘barbara.bmp’ thường được cộng đồng xử lý ảnh dùng làm ảnh chuẩn để thử nghiệm các phương pháp nén ảnh mới.

- Tập gồm 28 ảnh khác nhau được lựa chọn từ tập 68 ảnh nói trên dùng để tính CI1, CI2, CI3, CI4 và PSNR của 04 loại nén ảnh DCT, LOT, LBT, wavelets SPIHT với bpp cố định là 0,5. Sau đó, chúng tôi quan sát quan hệ PSNR với lần lượt theo giá trị tăng dần của CI1, CI2, CI3, CI4, và nhận thấy PSNR với CI4 có quan hệ xác định. Để đơn giản việc trình bày, chúng tôi chỉ đưa ra quan hệ PSNR-CI4 trong hình 2. Trục tung ( hình 2 ) trình bày theo thứ tự ảnh từ 1 đến 28 và được sắp xếp theo giá trị CI4 tăng dần như sau : ‘slope’, ‘polygons2’, ‘275005’, ‘lamps’, ‘275000’, ‘275008’, ‘milkdrop’, ‘beauty’, ‘snake’, ‘redhead’, ‘Zelda’, ‘031’, ‘boat3’, ‘home4’, ‘lena’, ‘monarch’, ‘desk1’, ‘peppers’, ‘balloon’, ‘city1’, ‘goldhill’, ‘stream’, ‘leopard’, ‘384009’, ‘barbara’, ‘test8g’, ‘baboon’ và ‘cougar’.

Từ hình 2, với cùng bộ nén ảnh và bpp = 0,5 cố định, chúng ta thấy rõ tùy theo đặc trưng ảnh, giá trị PSNR thay đổi rất lớn từ 15 dB đến 55 dB. Mỗi loại bộ nén sẽ cho đường đặc tuyến chênh lệch nhau vài dB, nhưng điều quan trọng là tất cả các đường đặc tuyến trong hình 2 đều biến đổi tăng hay giảm theo CI4 của các loại ảnh khác nhau. Với cách sắp thứ tự ảnh theo giá trị CI 4 tăng dần trong hình 2, chúng ta thấy rõ số đo CI4 có quan hệ xác định với PSNR : giá trị CI4 của ảnh tăng sẽ dẫn đến PSNR tương ứng giảm. Tiếp theo, chúng tôi quan sát quan hệ giữa CI4 và CI1, CI2, CI3 theo giá trị CI4 tăng dần của tập 68 ảnh.

- Hình 3a) và 3b) trình bày quan hệ các số đo CI1, CI2, CI3, CI4 của tập 68 ảnh mẫu .bmp 512x512, gồm nhiều loại khác nhau. Trục tung ( hình 3a) và 3b ) đánh số thứ tự 68 ảnh, được liệt kê từ 1 đến 68 theo giá trị CI4 tăng dần như sau : ‘squares’, ‘slope’, ‘polygons2’, ‘linespr’, ‘275005’, ‘tbirds’, ‘275000’, ‘275008’, ‘room7’, ‘275001’,

'milkdrop', 'montage', 'dna', 'beauty', '275002', 'snake', 'redhead', 'Michelangelo\_David', 'Zelda', '031', '384003', 'boat3', 'boat1', 'home3', 'home4', 'autol', 'jet', 'lena', 'monarch', 'building6', '384011', 'desk1', 'truck', 'coral', 'peppers', 'balloon', 'caltrain007', 'tank', 'city1', 'city2', 'test1', 'boat', 'san\_juans', '384006', 'goldhill', '384010-2', 'stream', 'solings', '354009', 'leopard', 'spring1887', '384009', '384007', 'globe', 'barbara', '5.2.10', 'test2', 'tm', 'test8g', 'test3', 'test8r', 'test8b', 'ccitt7', '4.2.03', 'baboon', 'cougar', 'hand', 'test4'. Kết hợp hình 2 và 3a) và 3b), chúng ta nhận thấy giữa những số đo thống kê đặc trưng ảnh CI1, CI2 và PSNR không có mối liên hệ xác định. Nói cách khác, với một tỷ số nén bpp cho trước và sử dụng cùng bộ nén có méo, không có số đo đặc trưng theo tính chất thống kê ảnh có khả năng thể hiện được đặc trưng biến đổi lớn của PSNR theo các loại ảnh khác nhau.

Kết hợp hình 2 và 3b), chúng ta nhận thấy số đo theo gradient tại những điểm ảnh biên trên ảnh CI3 cũng không có mối quan hệ rõ ràng với PSNR. Số đo theo gradient của toàn bộ điểm ảnh-CI4- đặc trưng tốt nhất mối quan hệ giữa các loại ảnh khác nhau với PSNR. Ví dụ ảnh "test4" có CI4 lớn nhất sẽ cho PSNR nhỏ nhất và ảnh "squares" có CI4 nhỏ nhất cho PSNR lớn nhất.

#### 4. PHƯƠNG TRÌNH CI4-PSNR

Vì CI4 thể hiện đặc trưng vùng tích cực của toàn bộ ảnh, chúng ta sẽ tìm mối quan hệ của CI4 đối với chất lượng mã hoá có méo PSNR. Hình 3 thể hiện một quan hệ nghịch đảo giữa số đo CI4 với số đo PSNR. Những ảnh có giá trị CI4 thấp (như ảnh "slope") sẽ cho PSNR tốt hơn những ảnh có CI4 lớn (như ảnh "baboon").

Chúng tôi thiết lập phương trình thể hiện quan hệ giữa CI4 và PSNR trên một tập gồm 28 ảnh khác nhau như đã liệt kê trên. Quan hệ giữa CI4 và PSNR được xây dựng theo phương pháp bình phương bé nhất. Từ tập các cặp giá trị CI4 và PSNR, chúng ta có thể xem đây là quan hệ phi tuyến, theo hàm logarit như sau:

$$y = A \ln(x) + B \quad (3)$$

Trong phương trình CI4-PSNR, CI4 sẽ là đối số x và PSNR tương ứng với y. A và B là hai hằng số cần phải được xác định

Bằng cách lập trình matlab theo phương pháp bình phương bé nhất, chúng tôi xác định được phương trình CI4-PSNR cho mỗi loại bộ nén ảnh như sau :

\* Trường hợp bộ nén ảnh dùng DCT ( xem hình 4 b) : A = 56.5197; B = -9.4044 .

\* Trường hợp bộ nén ảnh dùng LOT ( xem hình 4 c) : A = 56.1813; B = -9.0996 .

\* Trường hợp bộ nén ảnh dùng LBT ( xem hình 4 d) : A = 57.2127; B = -9.2993 .

\* Trường hợp bộ nén ảnh wavelets SPIHT ( xem hình 4 a) : A= 62.3773; B= -10.5850.

Từ những hình 4a)b)c)d), chúng ta nhận thấy so với các bộ nén đơn giản DCT, LOT, LBT, khi số đo CI4 tăng dần, hiệu quả của bộ nén ảnh wavelets SPIHT sẽ giảm. Cụ thể độ vượt trội PSNR từ 5 dB giảm xuống còn 1-2 dB, khi CI4 giảm từ 1 xuống 40. Những hình 4a)b)c)d) cho phép chúng ta chọn lựa loại bộ nén sử dụng hiệu quả nhất cho một ảnh đã biết CI4. Với ảnh có CI4 > 10, chúng ta có thể sử dụng bộ nén LBT để có được một chất lượng giải nén tốt và thuật toán nén không quá phức tạp. Trong hệ thống nén ảnh sử dụng phần cứng, thuật toán nén LBT đơn giản hơn sẽ cho giá thành ít hơn với cùng chất lượng.

## 5. KẾT LUẬN

Từ bài báo này đã chứng tỏ rằng các đặc trưng khác nhau của ảnh có ảnh hưởng trực tiếp lên chất lượng (giá trị PSNR) nén có méo. Nén có méo với biến đổi wavelets hay biến đổi DCT, LOT, LBT đều phụ thuộc vào đặc trưng ảnh.

Số đo CI4 theo gradient ảnh thể hiện được sự khác biệt nội bộ các miền không gian điểm ảnh của những loại ảnh khác nhau. CI4 có mối liên hệ chặt chẽ với giá trị PSNR khi nén ảnh có méo. CI4 đã giải thích tại sao giá trị PSNR thay đổi rất lớn khi dùng cùng một bộ nén có tổn hao.

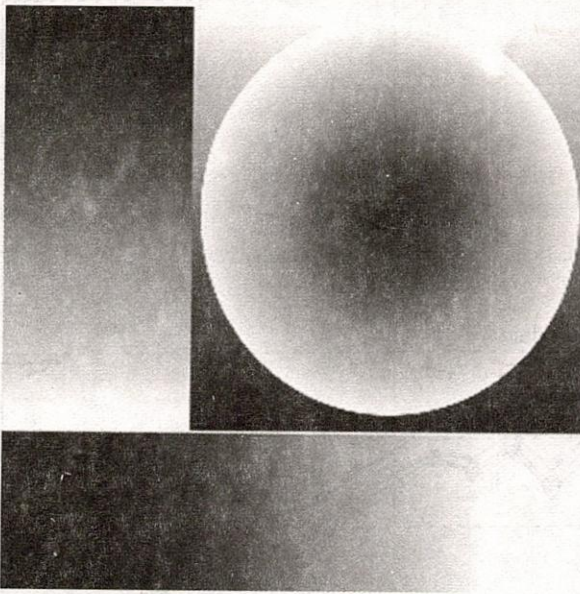
Phương trình CI4-PSNR xác định khả năng nén có tổn hao của một tập ảnh. Với một ảnh có giá trị CI4 được tính trước, chúng ta có thể tiên đoán được giá trị PSNR của ảnh từ phương trình CI4-PSNR, cho dù không thực hiện nén.

Phương trình CI4-PSNR có sai khác một ít ( từ 1dB đến 5 dB) khi dùng những loại bộ nén khác nhau ( wavelets, DCT , LOT hay LBT).

## THE NEW RELATIONSHIP BETWEEN IMAGE FEATURES AND LOSSY CODING PERFORMANCE

Le Quang Tuan, Vu Dinh Thanh, Nguyen Kim Sach

**ABSTRACT:** Nowadays, we need to compress and decompress many types of images such as natural, PC integrated, medical ones ... Depending on the characteristic of images, PSNR values can vary up to 30 dB with the same lossy coder. How to establish a new relationship between the features of different types of images and the lossy coding performance PSNR ? In this paper, we introduce some measures of image characteristic which are called CI (characteristic of image) and build a mathematical function of the relationship between CI and PSNR.



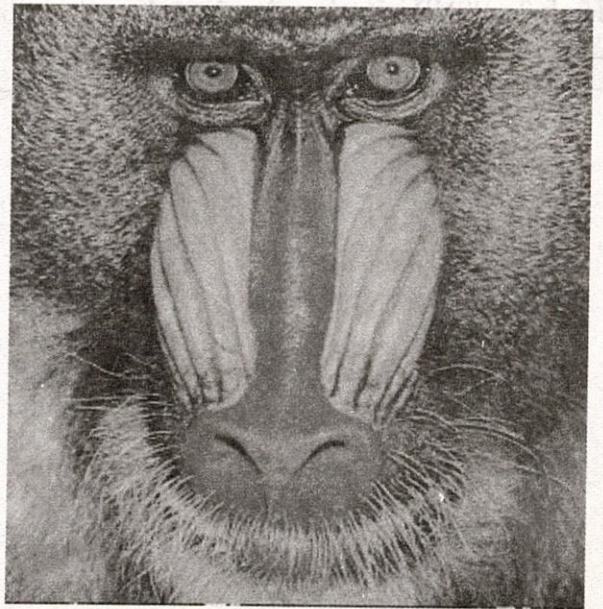
Hình 1a) Ảnh slope.bmp 512 x 512  
CI1 = 163.2500      CI2 = 9.7679  
CI3 = 0.6657      CI4 = 2.7186



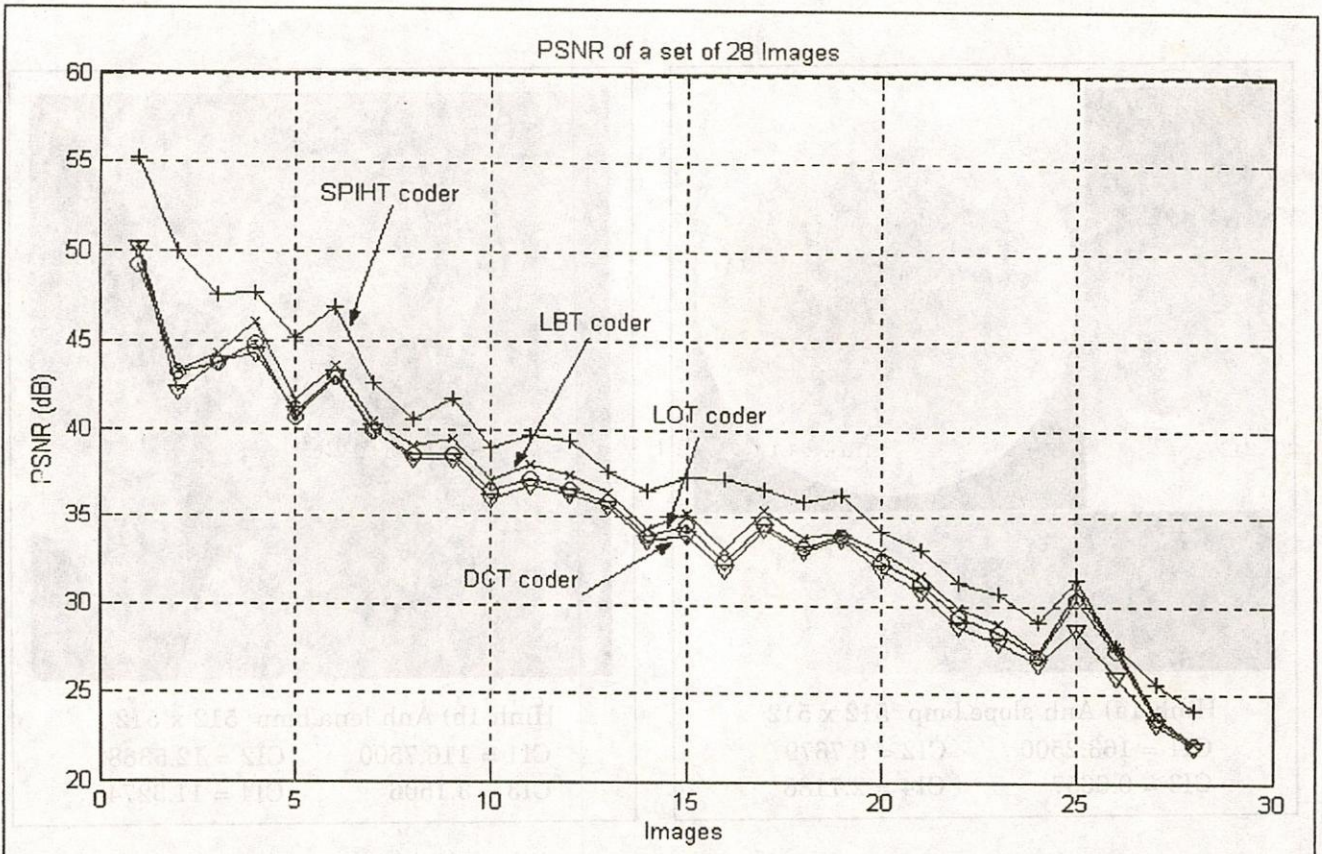
Hình 1b) Ảnh lena.bmp 512 x 512  
CI1 = 116.7500      CI2 = 12.5368  
CI3 = 3.1506      CI4 = 11.3274



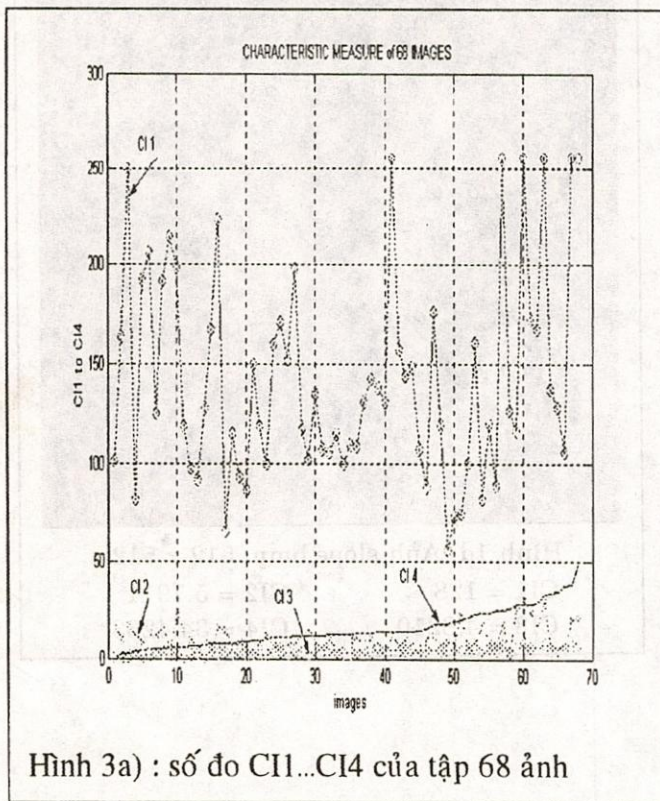
Hình 1c) Ảnh slope.bmp 512 x 512  
CI1 = 120      CI2 = 8.5242  
CI3 = 5.6007      CI4 = 23.5383



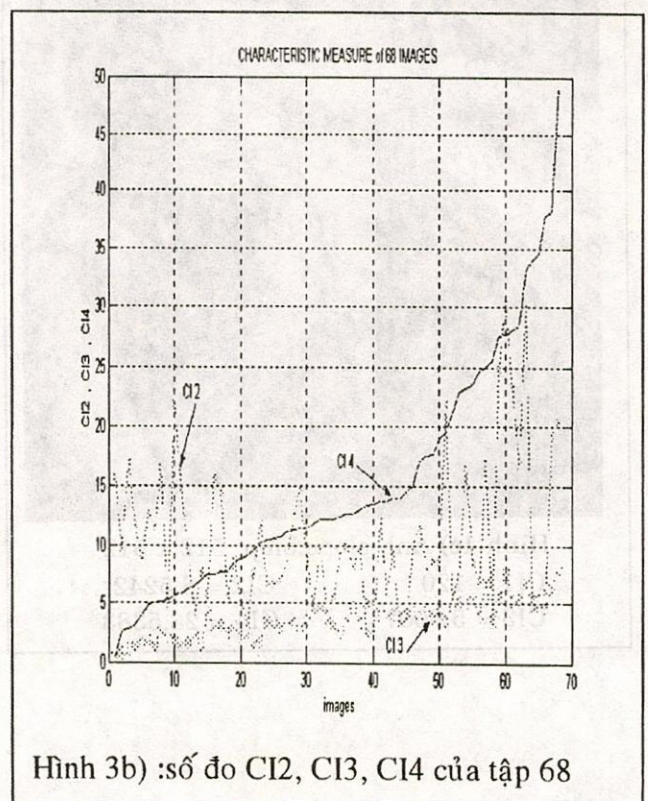
Hình 1d) Ảnh slope.bmp 512 x 512  
CI1 = 128      CI2 = 5.7911  
CI3 = 4.8210      CI4 = 34.4263



Hình 2 : PSNR của tập gồm 28 ảnh, với các bộ nén ảnh khác nhau và bpp cố định là 0,5.

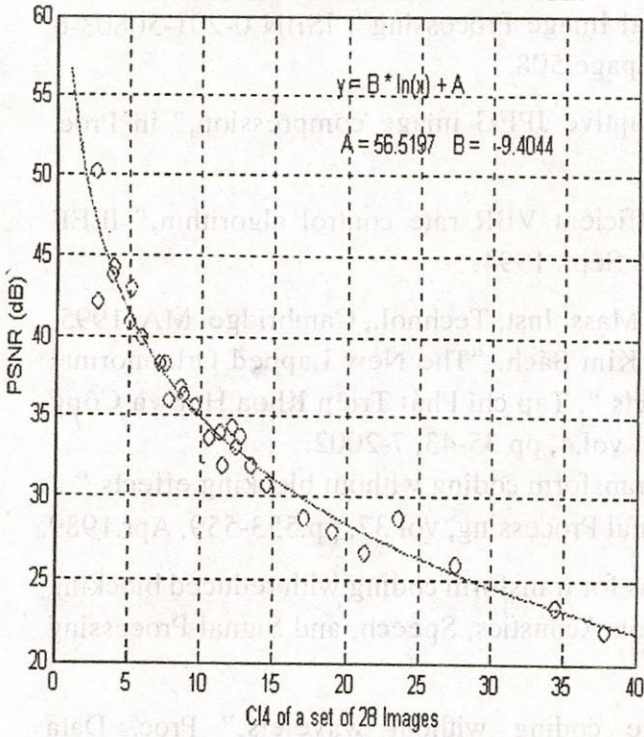


Hình 3a) : số đo CI1...CI4 của tập 68 ảnh



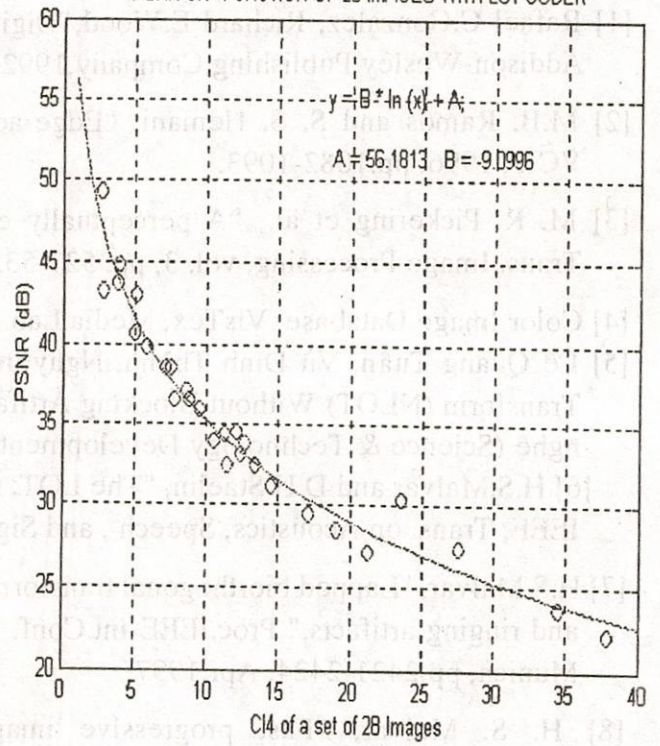
Hình 3b) : số đo CI2, CI3, CI4 của tập 68

"PSNR-CI4" FUNCTION OF 28 IMAGES WITH DCT CODER



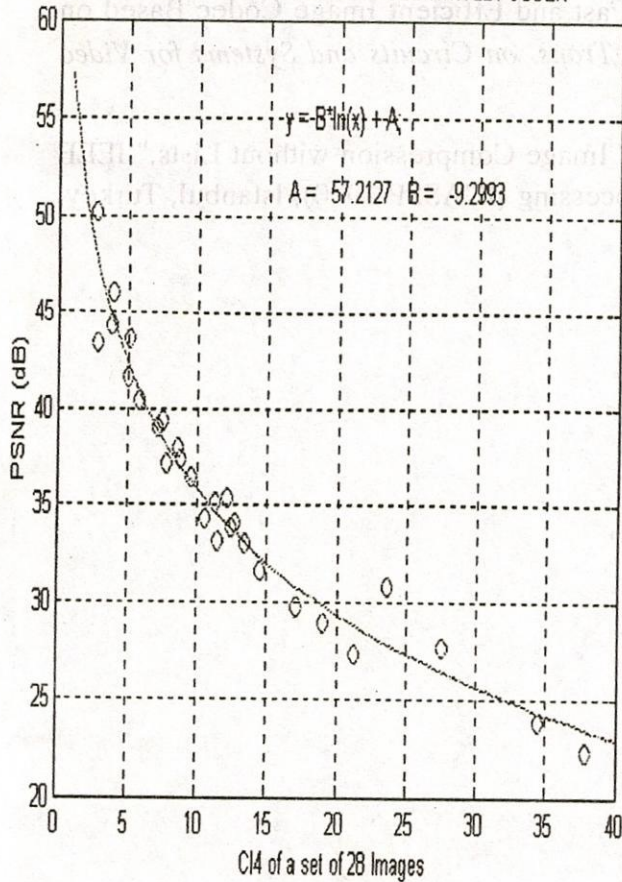
Hình 4a) Phương trình CI4-PSNR của bộ nén DCT

"PSNR-CI4" FUNCTION OF 28 IMAGES WITH LOT CODER



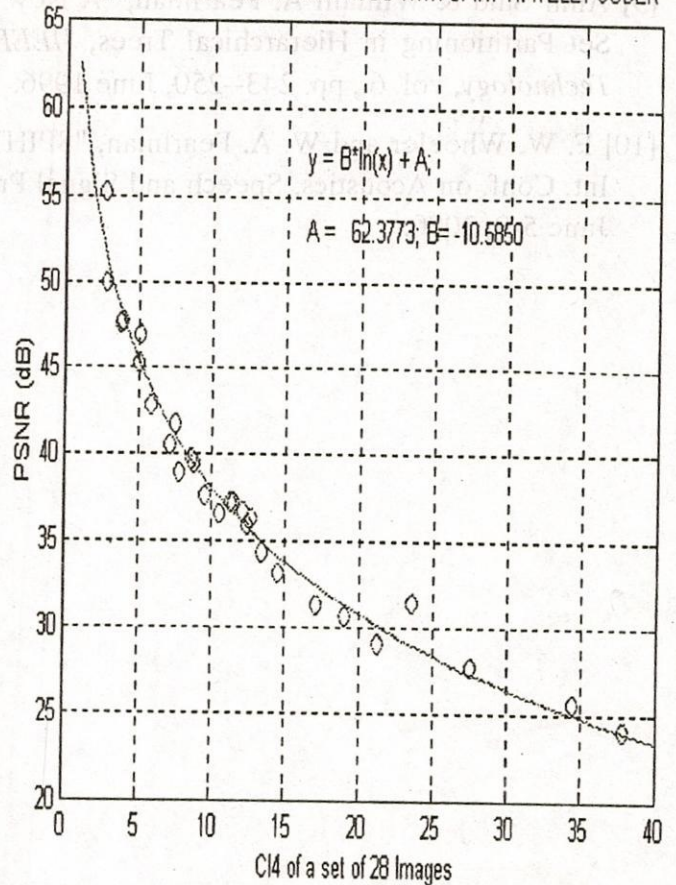
Hình 4b) Phương trình CI4-PSNR của bộ nén LOT

"PSNR-CI4" FUNCTION OF 28 IMAGES WITH LBT CODER



Hình 4c) Phương trình CI4-PSNR của bộ nén LBT

"PSNR-CI4" FUNCTION OF 28 IMAGES WITH SPIHT WAVELET CODER



Hình 4d) Phương trình CI4-PSNR của bộ nén SPIHT



## TÀI LIỆU THAM KHẢO

- [1] Rafael C.Gonzalez, Richard E.Wood,"Digital Image Processing", ISBN 0-201-50803-6 Addison-Wesley Publishing Company,1992, page 508.
- [2] M.B. Ramos and S. S. Hemani, "Edge-adaptive JPEG image compression," in Proc. VCIP, 1996, pp.1082-1093.
- [3] M. R. Pickering et al., "A perceptually efficient VBR rate control algorithm," IEEE Trans. Image Processing, vol. 3, pp. 527-532, Sept. 1994.
- [4] Color Image Database, VisTex, Media Lab., Mass. Inst. Technol., Cambridge, MA, 1995.
- [5] Lê Quang Tuấn, Vũ Đình Thành, Nguyễn Kim Sách, "The New Lapped Orthonormal Transform (NLOT) Without Blocking Artifacts ", Tạp chí Phát Triển Khoa Học và Công nghệ (Science & Technology Development), vol.4, pp 35-43, 7-2002.
- [6] H.S.Malvar and D.H.Staelin, "The LOT: transform coding without blocking effects," IEEE, Trans. on Acoustics, Speech , and Signal Processing, vol.37, pp.553-559, Apr.1989.
- [7] H.S.Malvar, "Lapped biorthogonal transforms for transform coding with reduced blocking and ringing artifacts," Proc.IEEE Int.Conf. on Acoustics, Speech, and Signal Processing Munich, pp.2421-2424, Apr.1997.
- [8] H. S. Malvar, "Fast progressive image coding without wavelets," Proc. Data Compression Conference, Snowbird, Utah, March 2000, pp. 243-252.
- [9] Amir Said & William A. Pearlman, "A New Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees, "IEEE Trans. on Circuits and Systems for Video Technology, vol. 6., pp. 243--250, June 1996.
- [10] F. W. Wheeler and W. A. Pearlman, "SPIHT Image Compression without Lists," IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP 2000), Istanbul, Turkey, June 5-9, 2000.