論文 97-22-7-14

Multidimensional Uniform Cubic Lattice Vector Quantization for Wavelet Transform Coding

Jae Sik Hwang*, Yong Jin Lee**, Hyun Wook Park** *Regular Members*

웨이브렛변환 영상 부호화를 위한 다차원 큐빅 격자 구조 벡터 양자화

正會員 황 재 식*, 이 용 진**, 박 현 욱**

※본 논문은 1995년 한국과학재단 95-0100-23 과제의 일환으로 연구되었음

ABSTRACT

Several image coding algorithms have been developed for the telecommunication and multimedia systems with high image quality and high compression ratio. In order to achieve low entropy and distortion, the system should pay great cost of computation time and memory. In this paper, the uniform cubic lattice is chosen for Lattice Vector Quantization (LVQ) because of its generic simplicity. As a transform coding, the Discrete Wavelet Transform (DWT) is applied to the images because of its multiresolution property. The proposed algorithm is basically composed of the biorthogonal DWT and the uniform cubic LVQ. The multiresolution property of the DWT is actively used to optimize the entropy and the distortion on the basis of the distortion-rate function. The vector codebooks are also designed to be optimal at each subimage which is analyzed by the biorthogonal DWT. For compression efficiency, the vector codebook has different dimension depending on the variance of subimage. The simulation results show that the performance of the proposed coding method is superior to the others in terms of the computing complexity and the PSNR in the range of entropy below 0.25 bpp.

요 약

날로 발전하는 무선통신 시스템 및 멀티미디어 통신 시스템에서 영상 데이터의 고화질과 고압축이 중요한 목표가 되고 있으며 이를 실현하기 위한 알고리즘의 연구가 현재 활발히 진행되고 있다. 그러나 이러한 시스템을 구축하는 데 있어서 많은 계산량과 메모리가 요구되고 있다. 본 논문에서는 간단한 유니폼 큐빅 격자를 벡터 양자화의 코드북으로 선택하였고 다해상도 표현이 가능한 웨이브렛 변환을 영상에 적용하였다. 여기서 제안한 시스템은 기

*삼성전자 멀티미디어연구소 선임연구원
** 한국과학기술원 정보 및 통신공학과
論文番號:96307-0930
接受日字:1996年 9月 30日

본적으로 쌍직교 웨이브렛 변환과 유니폼 격자벡터 양자화로 구성되어 있는 데 엔트로피와 왜곡 문제의 해결에 웨이브렛의 다해상도 특성을 유용하게 사용하였다. 격자구조의 코드북의 성능은 격자의 형태 및 벡터의 차원에 의존한다. 따라서 쌍직교 웨이브렛 변환으로 분해된 각각의 부영상들의 분산 값에 의해 최적으로 할당된 비트 율에 따라 코드북의 차수를 결정하였다. 모의 실험결과 제안한 방법이 기존의 방법보다 계산량이 줄었고 0.25bpp이하에서의 엔트로피와 왜곡에 있어서 좋은 성능을 보여 주고 있다.

Ⅰ. Introduction

The demand for data transmission at low bit rate and high accuracy requirements has been increased day by day, according to the development of mobile and multimedia communications. Therefore, the research on the development of data compression for digital communication system has been actively going on. The new image compression algorithms based on the fractal theory, the model based coding theory, the object oriented coding theory, and wavelet transform theory have been developed. Specially, the wavelet transform has recently become quite popular and its main property is multiresolution, or multiscale view of signal analysis [1]. The concept of multiresolution analysis was introduced by Mallat who viewed the orthonormal bases of wavelets as a vehicle for multiresolution analysis [1][2]. Daubechies achieved the construction of orthonormal wavelet bases by combination of Mallat's ideas and restrictions on filters [3]. On the other hand, the pyramidal lattice vector quantization (PLVQ) has good performance in terms of MSE (Mean Square Error) and complexity. Of course, the non-structured VQ based on LBG (Lide, Buzo and Gray) algorithm [4] was proven to take a good performance in MSE but this algorithm is computationally expensive and hard to implement in real time image processing. The PLVQ overcomes this shortcome because it does not need the training that yields sometimes local optimum VQ codebook.

This paper proposes a new compression scheme using wavelet transform and LVQ which has more powerful performance in computation time and distortion than other compression algorithms [5][6][7][17] below 0.25 bpp. The uniform cubic lattice (Z^n) is used as the basic lattice for LVQ because it can be implemented with simple architecture. Section II deals with DWT (Discrete Wavelet Transform) using biorthogonal wavelet basis. In section III, pyramid VQ and the uniform cubic lattice are briefly reviewed, and encoding/decoding algorithms are described.

Section IV presents simulation results and compares with other algorithms in terms of entropy, distortion and complexity.

Ⅱ. Wavelet Transform

Transform and subband coding of image are widely used in image compression. The DCT (Discrete Cosine Transform) provides good energy compaction, but it suffers from blocking artifacts that become more pronounced at higher compression ratio. The DWT, which is closely related to subband coding, has a mutiresolution property and is more efficient in describing the abrupt changes of edges. The quantization of the transformed coefficients can be realized in consideration of the human visual system because the scale of multiresolution basis function varies in logarithmic scale [8].

1. DWT (Discrete Wavelet Transform)

By the paraunitary QMF (Quadrature Mirror Filter) bank, the DWT can be realized easily [9]. A basic structure for biorthogonal wavelet transform is shown in Fig. 1 [10][11]. The biorthogonal basis is used in this paper and there are two pairs of filters: One pair is $h[n]$ and $g[n]$, which are typically lowpass and highpass filters for signal analysis, respectively. The other pair is $h^*[n]$ and $g^*[n]$, which are dual basis of $h[n]$ and $g[n]$, respectively used for signal synthesis. The wavelet coefficients obtained from a two-channel filter bank with filters $h[n]$ and $g[n]$, followed by down-sampling by 2, are given as follows,

$$x_h[n]=\sum_k h[2n-k]x[k] \quad (1)$$

$$x_g[n]=\sum_k g[2n-k]x[k] \quad (2)$$

where $x_h[n]$ is input signal, $x_h[n]$ and $x_g[n]$ are the channel signals depicted in Fig. 1.

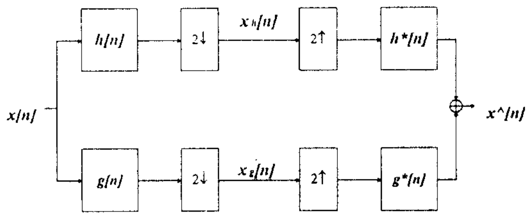


Fig. 1 Basic structure for biorthogonal wavelet transform

If the relationship between analysis and synthesis filters are given by

$$g^*[n]=(-1)^n h[-n+1] \quad (3)$$

$$g^*[n]=(-1)^n h^*[-n+1] \quad (4)$$

$$\sum_n h[n]h^*[n+2k]=\delta[k] \quad (5)$$

the perfect reconstruction can be realized as

$$\hat{x}[n]=\sum_n (h^*[n-2k]x_h[k]+g^*[n-2k]x_g[k])=x[n] \quad (6)$$

In our simulation, Daubechies97 filters (analysis filter with 9 taps and synthesis filter with 7 taps) are adopted [12]. For 2-D DWT, 1-D filters are applied to 2-D image. At first, apply lowpass filter $h[n]$ to the original image along the horizontal direction and then fill the left half plane of the transformed image with the filtered outputs, and apply highpass filter $h[n]$ to the original image along the horizontal direction and fill the right half plane of the transformed image with the filtered outputs. So, the new transformed image is created, which consists of the lowpass filtered outputs on the left half plane and the highpass filtered ones on the right half plane. After this processing in the horizontal direction, apply lowpass filter $g[n]$ to the new created transformed image along the vertical direction and fill the upper half plane with the filtered outputs, and apply highpass filter $g[n]$ along the vertical direction and fill the bottom half plane with the filtered outputs. Therefore, four kinds of subband images are created, denoted by HH, HL, LH, and LL, as shown in Fig. 2. The same procedure above is repeated four times on the subband image LL in our simulation.

Ⅲ. Pyramidal Vector Quantization and Uniform Cubic Lattice

1. Pyramidal Vector Quantization

In the paper written by T.R.Fisher[13], when the transformed coefficients have an independent, identically distributed multivariate Gaussian distribution, the surfaces of equal probability for multidimensional vector are ordinary spheres when the components of the vector have no correlation with each other. Whereas in the case of Laplacian sources, those are pyramids. The PVQ uses codewords corresponding to points in the cubic lattice that also lies on a particular pyramid.

2. Uniform Cubic Lattice

The uniform cubic lattice (Z^n) is used as the basic lattice for PVQ which is written as

$$Z^n=\{Y=(y_1, y_2, \ldots, y_n | y_i \in Z)\} \quad (7)$$

where Z is the integer space. Let us define n-D hyperpyramidal surface, $S(n, m)$, as follows

$$S(n, m)=\left\{X: \sum_{i=1}^{n} |x_i|=m\right\} \quad (8)$$

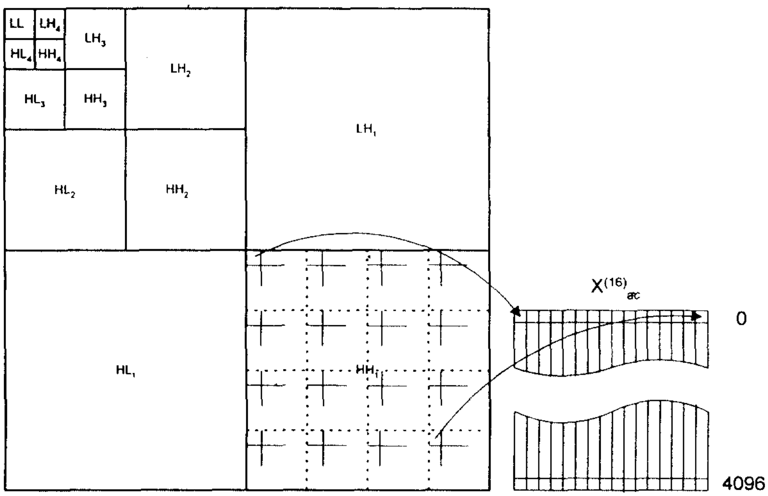


Fig. 2 Subband images and vector assembling scheme

where n is the vector dimension and m is the radius of the hyperpyramidal surface. Let $N(n, m)$ be the number of lattice points on $S(n, m)$, then the recursive formula proven by Fischer [13] is used to calculate $N(n, m)$ as follows,

$$N(n, m) = N(n-1, m) + N(n-1, m-1) + N(n, m-1), \quad \text{for } n \geq 2,\ m \geq 2 \tag{9}$$

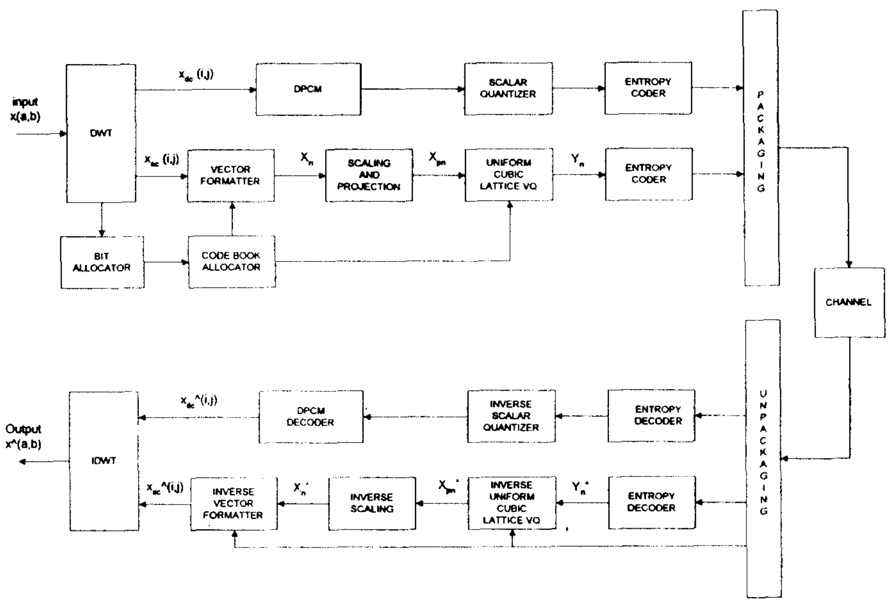


Fig. 3 Overall block diagram of the proposed encoding and decoding system

Ⅳ. Proposed Alogrithm

Since the source image coefficients can be well transfomed to the coefficients in the several subimages with the Laplacian distribution by the DWT, n-D vector of coefficients in the subimage is mapped into n-D space with pyramidal distribution. The n-D vector is quantized into unique lattice points on n-D hyperpyramid surface, where the dimension is decided by the allocated bitrate of the subimage.

The overall block diagram of the proposed encoding and decoding system is shown in Fig. 3. The encoding procedure is described as follows;

1. Encoding Algorithm

Step 1: The original image is decomposed into subband images through the multi-level wavelet transform.

Step 2: The DC coefficients, which are lowest subband image, are so important that the lossless compression method is adopted. The DPCM (Differential Pulse Code Modulation) method is applied to DC coefficients, whereas PLVQ (Pyramidal Lattice Vector Quantization) is applied to AC coefficients of the other subband images. The AC coefficients, $X_{a,c}(i, j)$, are assembled to vectors X_n which have the given dimension n. The pixels in real image have high correlation with neighborhood pixels. The coefficients in transformed image have still some correlation with the neighborhood pixels although the transform makes the coefficients decorrelated. For this reason, the vector formatter assembles the dimensional vectors like Fig. 2 so that the components of vectors have less correlation with each other.

Step 3: Determine the optimal scale factor, α, with a given compression ratio and image quality. The compression ratio can be adjusted by changing the scale factor α. Every vector in subband images of LH, HL, and HH in Fig. 2 is scaled with the factor of α as follows,

$$X_{sn} = \alpha X_n \tag{10}$$

Step 4: If the scaled vectors X_{sn} are outside of the maximum pyramidal surface, that is, the value $\|X_{sn}\| = \sum_{i=1}^{n} |x_{si}|$ is greater than the maximum radius M, then those vectors are projected on the pyramidal surface $S(n, M)$ by projection theorem [14]. In this case, the maximum radius M should be determined for optimum compression performance. Therefore, all the scaled and projected vectors X_{pn} are on or inside the hyperpyramidal surface $S(n, M)$.

Step 5: Every component x_{pi} of vector X_{pn} are rounded to y_i so that y_i has only integer value. These quantized vectors are denoted by $Y_n = (\{y_1, y_2, ..., y_n) : y_i \in Z\}$.

Step 6: For the quantized vectors $Y_n = \left\{(y_1, y_2, ..., y_n) : \sum_{i=1}^{n} |y_i| = m\right\}$, there are $N(n, m)$ lattice points, which can be calculated from the equation (9). One to one mapping is possible between the vector Y_n and an integer in the range $[0, ..., N(n, m)-1]$. Therefore, the unique indexing can be performed [13].

Step 7: Those index values are entropy coded and transmitted through the communication channel.

2. Decoding Algorithm

Decoding procedure is given as the following steps:

Step 1: The received bit streams through the communication channel are entropy decoded to the vector index values.

Step 2: The decoded index values are converted to the corresponding vectors

Step 3: Rescale the vectors using rescale factor $\frac{1}{\alpha}$.

Step 4: The rescaled vectors are deassembled to the corresponding subband images. The DC coefficients are decoded by DPCM. The reconstructed image is obtained by inverse DWT of the deassembled subband images.

3. Vector Dimension and Bitrate

The performance of Z^n LVQ is affected by the dimension of the vector. Therefore, the dimension is decided by the bit allocation algorithm. The optimal bit allocation of each subband images is determined by the variance of subband images in order to approach the target bitrate. The relationship between the variance and optimal bit allocation is given by [15]

$$r_i(\theta) = \max\left[0.0, \frac{1}{2}\log_2\left(\frac{\sigma_i}{\theta}\right)\right] \tag{11}$$

where σ_i is a variance of i_{th} subband image and θ is a variance to be adjusted until the following equation is satisfied.

$$R_T = r_{dc} + \sum_{i=1}^{3L} r_i(\theta) \tag{12}$$

where L is the multiresolution level, R_T is the total target bitrate and r_{dc} is the bitrate at the DC subband image. The variance of the DWT coefficients increases for lower-frequency subband images, which means that the coefficients at the lower frequency is more important. Therefore, the subband images which have higher variance are assigned higher bitrate than those have lower variance under the distortion restriction. The bitrate is increasing as the codebooks dimension is decreasing. The bitrate range corresponding to the dimension n of Z^n are summarized at Table 1.

Table 1. The bitrate range for the dimension n of Z^n

Lattice	Z^{16}	Z^8	Z^4	Z^2
Bitrate	0 bpp~ 0.4 bpp	0.4 bpp~ 0.7 bpp	0.7 bpp~ 1.4 bpp	above 1.4 bpp

Ⅳ. Simulation Results

The Lena image, whose resolution is 512×512, was used for comparison study of PSNR (Peak Signal to Noise Ratio) and entropy. The proposed algorithm

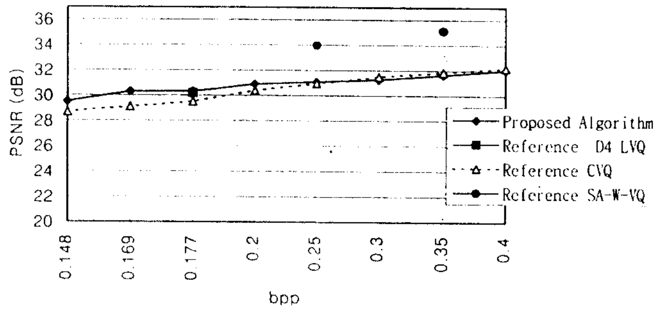


Fig. 4 PSNR of the proposed and the referenced algorithms

was compared with the CVQ (Classified VQ) [6], D^4 LVQ and SA-W-VQ [17]. The PSNR of the proposed algorithm is better than the D^4 LVQ and superior to the referenced CVQ as shown in Fig. 4. The SA-W-VQ is better than the proposed algorithm at the entropy above 0.25 bpp. However, the proposed algorithm is optimized at the entropy below 0.25 bpp for the real time video processing at the very low bit-rate communication network. By these reasons, the computation complexity is importantly considered. To measure the computation complexity, the number of operations per vector (o.p.v) is counted. The number of o.p.v is computed for one vector of dimension as follows.

i) The analysis of the LBG complexity

nL differences +nL squares +(n−1)L sums = (3n−1)L (o.p.v).

where L is codebook size, and n is the number of vector components.

ii) The analysis of D^n LVQ complexity

n roundoffs + n − 1 sums + n differences (to find δ) + n absolutes (to find absolute value of δ) + n − 1 differences (to find maximum value of the absolute of δ) + 1 round off (to find the integer of the component having absolute value of δ) + 2 modulations + 1 test (to find the even value) = 4n + 1 (o.p.v).

iii) The analysis of SA-W-VQ complexity

Because the SA-W-VQ uses the lattice Λ_n, 4n + 1 + 32

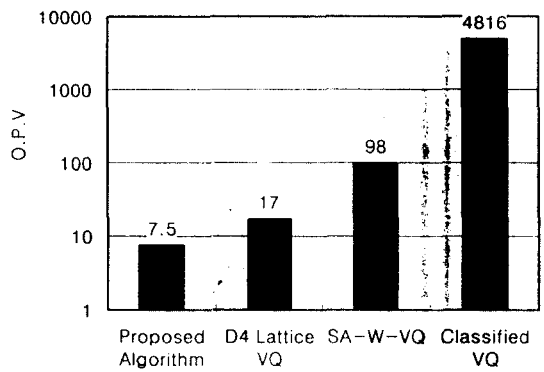


Fig. 5 The complexity results by the calculation of O.P.V (log scale) (O.P.V: No.of Operation Per Vector)

adds +1 multiply = 4n +34(o.p.v)
iv) The analysis of Z^n LVQ complexity
n roundoffs = n (o.p.v).

Fig. 5 shows the computation complexity by the measured o.p.v. This complexity measure is just the operation number of one vector, whereas the extra computation for the LBGs training operation is not included in this o.p.v. In the CVQ, its vector dimension n is 4 and codebook size L is 256. In the D^4 LVQ, its vector dimension n is 4 and 16 in case of the SA-W-VQ. On the other hand, the average vector dimension n of the proposed algorithm is 7.5 for Z^n LVQ.

V. Conclusions

In this paper, a new image coding algorithm was proposed, which is based on the theory of the wavelet transform and the lattice vector quantization. The DWT-based coding has better performance than the block DCT-based coding in terms of the blocking effects and the image quality. Also the VQ has good compression performance by Shannons information theory. The LVQ is the most efficient algorithm for implementation among several VQ algorithms. The proposed algorithm employed the uniform cubic lattice VQ to reduce the computation complexity. Finally the proposed algorithm improved the compression performance, computation complexity and storage cost.

References

1. S. G. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," IEEE Trans. on PAMI Vol. 2 No 7, July 1989.
2. G. Kaiser, A Friendly Guide to Wavelets, printed by Birkhauser, pp144-175, 1994.
3. I. Daubechies, "The Wavelet Transform, Time-Frequency Localization and Signal Analysis," IEEE Trains on IT, Vol 36, No 5, Sep 1990.
4. Y. Linde, A. Buzo, R. Gray, "An Algorithm for Vector Quantizer Design," IEEE, Trans on comm, Jan 1980.
5. M. Barlaud, P. Sole, T. Gaidon, M. Antonini, "Pyramidal Lattice Vector Quantization for Multiscale Image Coding," IEEE Trans on IP, Apr 1992.
6. Y. Huh, J. J. Hwang, C. K. Choi, R. Queiroz, "Classified Wavelet Transform Coding of Images using VQ," ICIP-94, Nov 1994.
7. Y. Huh, J. J. Hwang, K. R. Rao, "Classified Wavelet Transform Coding of Images using Two Channel conjugate VQ," ICIP-94, Nov 1994.
8. M. Antonini, M. Barlaud, P. Mathieu, I. Daubachies, "Image Coding Using Wavelet Transform," IEEE Trans on IP, Vol 1 No2, Apr 1992.
9. P. P. Vaidyanathan, Multirate Systems and Filter Banks, printed by Prentice Hall, pp457-539, 1993.
10. M. Vetterli, "Wavelets and Filter Banks: Theory and Design," IEEE Trans on SP, Sep 1992.
11. P. P. Vaidyanathan, "Theory and Design of M channel Maximally Decimated QMF with Arbitrary M having PR property," IEEE Trans on Acoustics, Apr 1987.
12. M. Barlaud, P. Sole, T. Gaidon, M. Antonini, P. Mathieu, "Pyramidal Lattice Vector Quantization for Multiscale Image Coding," IEEE Trans on IP, Vol 3, No 4, July 1994.

13. T. R. Fischer, "A Pyramid Vector Quantizer," IEEE Trans on IT, Vol. IT-32, pp 568-583, uly 1986.

14. A. Gersho, R. M. Gray, Vector Quantization and Signal Compression, Kluwer Academic Publishers, 1992.

15. M. A. Sid-Ahmed, Image Processing, printed by Mc GrawHill, pp 471-479, 1995

16. A. K. Jain, Fundamentals of Digital Image Processing, printed by Prentice Hall International Inc, pp150 -180, 1995.

17. E. da Silva, D. Sampson, "A Successive Approximation Vector Quantizer for Wavelet Transform Image Coding," IEEE Trans on IP, Vol 5, No 2, Feb 1996.

황 재 식(Jae-Sik Hwang) 정회원

1987년 2월:연세대학교 전자공학과(공학사)

1996년 2월:한국과학기술원 정보 및 통신공학과(공학석사)

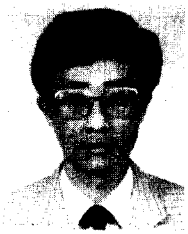
1987년 1월~현재:삼성전자 멀티미디어연구소 선임연구원

이 용 진(Yong-Jin Lee) 정회원

1990년 2월:한국과학기술원 과학기술대학 전기 및 전자공학과(공학사)

1995년 8월:한국과학기술원 정보 및 통신공학과(공학석사)

1995년 9월~현재:한국과학기술원 정보 및 통신공학과 박사과정



박 현 욱(Hyun-Wook Park) 정회원

1981년 2월:서울대학교 전기공학과(공학사)

1983년 2월:한국과학기술원 전기 및 전자공학과(공학석사)

1988년 2월:한국과학기술원 전기 및 전자공학과(공학박사)

1989년 7월~1992년 3월:Univ. of washington

1992년 5월~1993년 7월:삼성전자 정보컴퓨터연구소 수석연구원

1993년 8월~현재:한국과학기술원 정보 및 통신공학과 부교수

※주관심분야:디지털 신호처리, 영상처리, Image computing system, 의료영상