# ON ERRORS IN SUBBAND CODING OF IMAGES

Mikołaj BLEJA

Technical University of Poznań, ul. Piotrowo 3 A, 60-965 Poznań,
P O L A N D

## RÉSUMÉ

Les erreurs causés par le filtrage et par le codage des signaux dans un système de codage d'images en sousbandes utilisant des filtres bidimensionnels à réssponse impulsionnelle inverses sont étudiée. Les problèmes pratiques d'application des systèmes du codage en sousbandes sont discutés.

## ABSTRACT

In the paper the errors due to filtering and coding of the signals in subband coder have been considered. Particularly reversive systems with recursive filters have been explored. Practical aspects of implementing of subband coder have been discussed.

## Introduction

Digital image compression is concerned with minimization of the number of bits (or other information carrying units) used to represent an image. The aim of compression is to minimize the memory for storage or/and the bandwidth for transmission. One of the promising methods for compression of digital image data is subband coding (SBC) [1-8,10], which have been studied extensively since its invention. A block diagram of SBC system is shown in Fig.1. In SBC system the two-dimensional (2-D) band of the image is splitted up into N two-dimensional (2-D) bands, which are then decimated. After decimation (downsampling) subbands are encoded using techniques matched to the statistics of given subband. In the decoder, after upsampling, the subbands are filtered and then summed up in order to reconstruct the original data.

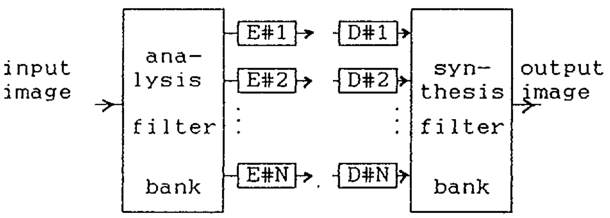


Fig.1. Scheme of subband coding
E#1-E#N -channel encoders
D#1-D#N -channel decoders

In the paper errors caused by coding and filtering in a particular subband coder (reversive systems with recursive filters) have been considered. Implementation aspects of SBC has been also discussed.

## Filter banks

Let us consider the basic four-subband SBC system with recursive filters. The scheme of partitioning of the frequency band is shown in Fig.2a. After splitting and downsampling the input image by factor two in each direction we get

$$U_{\mu_1\mu_2}(e^{j2\underline{\omega}})=\frac{1}{4}\sum_{\nu_1=0}^{1}\sum_{\nu_2=0}^{1} X(e^{j(\underline{\omega}+\underline{p})})H(e^{j(\underline{\omega}+\underline{s}+\underline{p})}) \qquad (1)$$

where $\underline{p}=(\nu_1\pi,\nu_2\pi)$ filter shift index, $\nu_1=0$ or 1, $\nu_2=0$ or 1, $\underline{s}=(\mu_1\pi,\mu_2\pi)$ is a subband shift index, where $\mu_1=0$ or 1, $\mu_2=0$ or 1 and $\underline{\omega}=(\omega_1,\omega_2)$, $X(\cdot)$ is the input of the system, $H(\cdot)$ represents transfer functions of the filters and $U(\cdot)$ subbands. After decoding, upsampling and filtering we get

$$Y(e^{j\underline{\omega}})=4\sum_{\mu_1=0}^{1}\sum_{\mu_2=0}^{1} \hat{U}_{\mu_1\mu_2}(e^{j2\underline{\omega}})\; H(e^{-(j\underline{\omega}+\underline{s})}) \qquad (2)$$

where $Y(\cdot)$ is the reconstructed image and $\hat{U}_{\mu_1\mu_2}(\cdot)$ - quantized subband.

Example of the filter characteristic set which realizes scheme from Fig.2a is shown on Fig.2b. Using recursive filters [3-5] we can design a system where the reconstructed signal $Y(\cdot)$ is free of magnitude and phase distortions in absence of quantization in channels. It is an advantage over SBC systems with FIR QMFs, where the design error is non-zero. Detailed considerations on filter banks for subband coding are given in [3,5].

## Error considerations

The problem of error analysis of subband coder have been studied by several authors (e.g. [10,11]). At first we will discuss the errors due to quantization in channels. In [11] it have been shown that for pdf-optimized quantizers the quantization process may be modeled with non-unity gain and additive noise (*gain-plus-additive-noise* quantizer model). This have been applied in [10] where the Lloyd-Max quantizer model have been incorporated into subband coder based on FIR QMFs. In [10] four error components have been calculated and extracted: filter bank design error, signal error, aliasing error and random error. For reversive systems based on pdf-optimized quantizers we achieve similar results, except that the filter bank design error is equal to zero. Nevertheless in most practical systems the quantizers in channels are not pdf-optimized ones (very often are matched to the properties of human visual system). It causes difficulties in analytical evaluation of quantization errors.

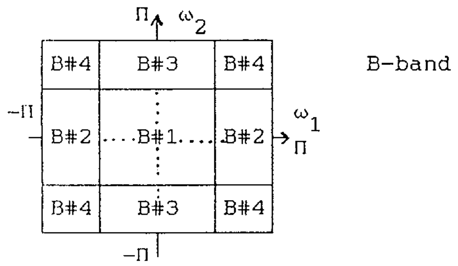


Fig.2a Four-subband frequency band partitioning scheme

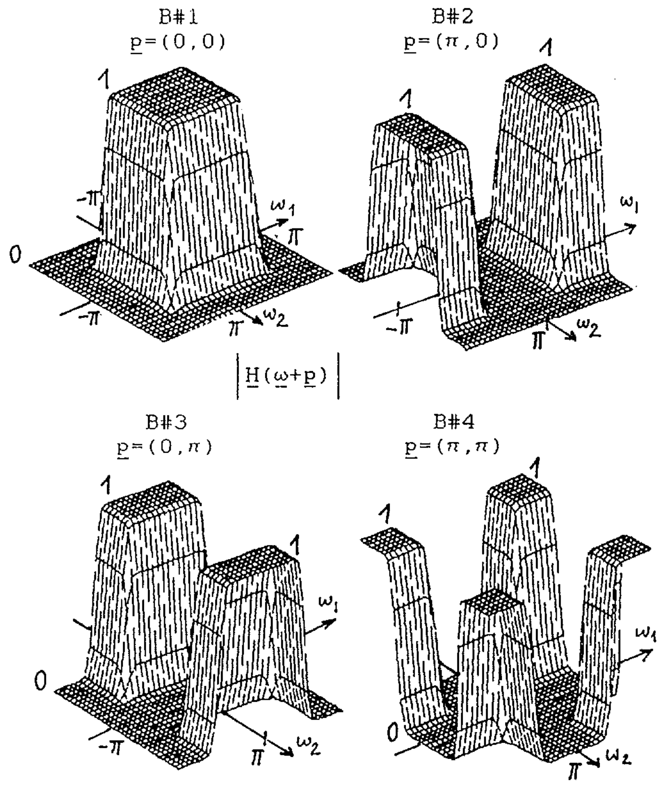


Fig.2b Amplitude responses of the filters in the four-subband SBC system

An example of non-pdf-optimized quantization is masking technique described in [5]. Now error analysis for this method will be presented. Let us consider SBC system, where the lowest subband, (B#1 in Fig.2a), is error free (8-bit PCM) coded and other subbands are coded by use of mask (window) with smoothed edges $w_{\mu_1\mu_2}(\underline{n})$, which is calculated from the subband data. The technique of high-frequency source quantization (subbands B#2,B#3,B#4) is the masking technique introduced above. Formally the quantized output signal for subband indexed $\mu_1\mu_2$ is equal to

$$\hat{u}_{\mu_1\mu_2}(\underline{n}) = u_{\mu_1\mu_2}(\underline{n})\, w_{\mu_1\mu_2}(\underline{n}) \qquad (3)$$

We calculate error for each subband

$$e^2_{\mu_1\mu_2}(\underline{n}) = u^2_{\mu_1\mu_2}(\underline{n})\,(1 - w_{\mu_1\mu_2}(\underline{n}))^2 \qquad (4)$$

assuming, that the error is a zero-mean signal we can calculate the error variance for each subband

$$\sigma^2_{e_{\mu_1\mu_2}} = \sigma^2_{u_{\mu_1\mu_2}} + \sigma^2_{\hat{u}_{\mu_1\mu_2}} - 2\sum_{\underline{n}} u^2_{\mu_1\mu_2}(\underline{n})\, w_{\mu_1\mu_2}(\underline{n}) \qquad (5)$$

All values of the mask function are so that:

$$0 \le w_{\mu_1\mu_2}(\underline{n}) \le 1 \qquad (5)$$

From this we get error variances for subband $\mu_1\mu_2$

$$\sigma^2_{e_{\mu_1\mu_2}} \le \sigma^2_{u_{\mu_1\mu_2}} - \sigma^2_{\hat{u}_{\mu_1\mu_2}} \qquad (6)$$

Quantization in given subband will generate error contributions at the receiver output within that frequency band. It is expected, that error contributions in different subbands are uncorrelated. In the reconstructed image the corresponding error variances will simply add.

$$\sigma^2_e \le \sigma^2_x - \sigma^2_{\hat{y}} \qquad (7)$$

We see that quality of reconstructed image is related to the difference between input and output image variance. The result may be used to set mask function parameters.
Using modulation property of 2-D signals we can easy compute aliasing error.

Now errors due to system implementation will be briefly discussed. In reversive SBC systems the image in decoder is processed in opposite direction to that in the encoder. The overall shift of the image after coding is equal to zero. Unfortunately the shifts of subbands are non-zero. To prevent shift errors (filter response truncation) some additional elements should be added at the end of each row and each column of the original image [3].

Other errors are transients errors due to filtering on the image boundaries. For transient effects suppressing some additional pixels should be added at the beginning of each row and column.

The number of pixels (rows, lines) to be added differs from image to image. For an average image 512x512 pixels about 120 lines must be additionally processed to avoid the above mentioned errors. Let us consider an image $x(\underline{n})$ N by M, where N- number of rows, M-number of lines. To reduce the number of samples which have to be added to the image following procedure can be applied (for simplification 1-D case will be considered):

step one: mirroring the image (line) with respect to its boundary

$$u(-k) := u(k), \qquad k=1,\dots,K \qquad (8a)$$
$$u(N+k) := u(n-k), \qquad k=1,\dots,K \qquad (8b)$$

where K- the number of pixels added N- number of pixels in a line

step two: smoothing of the mirrored boundaries with half of raised cosine window.

$$u(-k) := 0.5u(-k)\left(\cos\left(\frac{2\Pi k}{T}\right)+1\right) \qquad (9a)$$

for k=1,...,T-1 and

$$u(-k) := 0 \qquad (9b)$$

for k=T,...,K

$$u(N+k) := 0.5u(N+k)\left(\cos\left(\frac{2\Pi k}{T}\right)+1\right) \qquad (10a)$$

for k=1,...,T-1 and

$$u(N+k):=0 \qquad (10b)$$

for k=T,...,K

where T≤K, and T is the width of the window. In experiments values: T=10, K=10 and T=10, K=15 have been used. The number of lines and rows added may be chosen for each image individually. If the border region of the image is bright (high pixel values), the number of added lines and/or rows must be larger. For some pictures to reduce transients errors negative of the image can be processed.

Next group of errors is related to coefficients truncation and the type of arithmetic used in the system (fixed point, floating point). Type of arithmetic is very often related to the equipment used (digital signal processor, computer). In most cases high-speed processing is required. High-speed and minimum computational error are usually contradictory requirements. The number of multiplications and additions as well as the sensitivities of the filters must be small. To satisfy above conditions application of wave digital filters [9] in polyphase realization have been proposed in [3]. It have been shown, that substantial reduction in number of arithmetic operations and sensitivity can be obtained.

**Experiments**

The experiments have been performed for monochrome images (512x512 and 256x256, 8 bit/pixel). Four- to nineteen-subband coding schemes have been examined for standard images.

At first systems without quantization in channels have been explored.

Examined SBC-system architectures have been summarized in Table.1.

Table 1.

| # | filter bank | filter coefficients ([bit]) | image data [bit] |
|---|---|---|---|
| 1 | FLP,W,D | FLP | 8 ,16 |
| 2 | FLP,W,D | FXP (8,13,16) | 8 ,16 |
| 3 | FXP,W,D | FXP (8,13,16) | 8 ,13,16 |

FLP-floating-point
FXP-fixed-point
D-direct filter structure
W-wave filter structure

Floating point arithmetic was implemented using C-language **double** type. Fixed point arithmetic corresponds to that of TMS320xx fixed-point digital signal processor arithmetic (32-bit accumulator, 16-bit data, two-complement, saturated arithmetic). Direct structure and wave structure of IIR filters have been implemented. Several other experiments have been performed.

As an example comparison of three systems of configuration number 3 (Table.1) is given. The systems split the frequency band into 19 subbands (Fig.3). The filter banks consists of 5-th order 2-D recursive filters in the direct structure. The filters were designed as equivalents to the FIR QMFs of the type 24C [4,6]. No quantization have been performed. The values expressed in decibels (Table.2) are ratios of signal variances to error variances for each subband (after splitting and downsampling).

$$SNR_i = 10\log\frac{\sigma^2_{u_i}}{\sigma^2_{e_i}} \quad [dB] \qquad (11)$$

where $\sigma^2_{u_i}$ is the variance of the signal in subband $u_i(\underline{n})$ and $\sigma^2_{e_i}$ is the variance of the error $e(\underline{n}) = u_i(\underline{n}) - u^f_i(\underline{n})$ where $u^f_i(\underline{n})$ is the signal of i-th subband computed in fixed-point system.

Table.2

| | | | | |
|---|---|---|---|---|
| subband# | 19 | 18 | 17 | 16 |
| 8bit | 32.41 | 27.87 | 29.51 | 24.35 |
| 13bit | 42.41 | 38.73 | 34.52 | 35.81 |
| 16bit | 42.42 | 38.72 | 34.49 | 35.82 |
| [dB] | 12 | 13 | 14 | 15 |
| | 29.89 | 26.21 | 22.08 | 25.60 |
| | 30.36 | 26.70 | 22.08 | 26.81 |
| | 30.29 | 27.70 | 22.09 | 26.90 |
| | 11 | 10 | 9 | 8 |
| | 16.53 | 29.02 | 30.17 | 17.74 |
| | 48.17 | 35.22 | 36.69 | 41.42 |
| | 48.20 | 35.31 | 36.79 | 41.45 |

| | | | | |
|---|---|---|---|---|
| 3 | 4 | 5 | 6 | 7 |
| 32.13 | 30.34 | 32.76 | 29.14 | 31.89 |
| 49.80 | 46.88 | 45.39 | 29.68 | 39.93 |
| 50.47 | 46.89 | 45.40 | 29.59 | 39.91 |
| 1 | 2 | | | |
| 29.21 | 33.16 | | | |
| 50.01 | 47.75 | | | |
| 54.12 | 47.92 | | | |

From Table.2. we can state out, that the system with 13-bit coefficients gives similar results to that of 16-bit coefficients.

The error in higher bands is larger than for lower bands (e.g. subbands 1,2,3,4 versus 14,15,16,17). The signals in highest bands (8,15,16,17,18) have lower dynamic range and for most images can be set to zero (not transmitted).

Some results related to Table 2. are shown in Fig.4. for LENA image. There is no visible difference between reconstructions 4b,4c. Reconstructed image on Fig.4d is distorted due to coefficients quantization and computational errors (all data was 8-bit saturated). The square error for image 4d. is shown in Fig.4e. Reconstruction errors for LENA, INSTITUT and CAMERA-MAN image are summarized in Table 3.

The total reconstruction error increases with the number of bands.

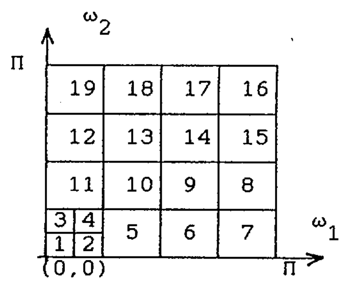


Fig.3. 19-subband frequency band partitioning scheme

From experiments we see, that the errors due to arithmetic and coefficients quantization are rather small. From this we can state out, that the task of filtering may be successfully performed using fixed-point digital signal processor. For TMS320xx processors register (16-bit coefficients) and immediate (13-bit coefficients) addressing will gives similar results from the error point of view.

In this paper only particular example of error analysis have been presented. Some practical advises of implementing a subband coder have been given. The research on this fiel will be continued.

Fig.4a Lena-original image

Fig.4b Lena-reconstruction, 16-bit filter coefficients, fixed-point arithmetic, 19-subbands, 16-bit data

Fig.4c Lena-reconstruction, 13-bit filter coefficients, fixed-point arithmetic, 19-subbands, 16-bit data

Fig.4d Lena-reconstruction, 8-bit filter coefficients, fixed-point arithmetic, 19-subbands, 8-bit data.

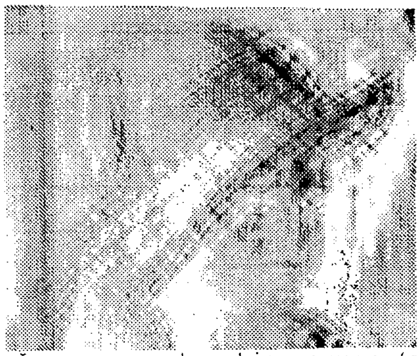

Fig.4e Lena-reconstruction error (square), 8-bit filter coefficients, fixed-point arithmetic, 19-subbands

Table.3.

| filter coefficients in [bit] | reconstruction MSE (mean square error) Cameraman | Lena | Institut |
|---|---|---|---|
| 16 | 1.50 | 1.39 | 1.692 |
| 13 | 2.01 | 1.63 | 1.90 |
| 8 | 96.71 | 48.80 | 37.59 |

(19-subbands)

### References


[1] VETERLI(M.). Multi-dimensional sub-band coding: some theory and algorithms, *Signal Processing* (1984), **6**, No. 2, pp.97-112.

[2] WOODS(J.), O'NEIL(S.). Subband coding of images. *IEEE Transactions* (1986), **ASSP-34**, No. 5, pp. 1279-1288.

[3] DOMAŃSKI(M.). Efficient Wave Filter Banks for Subband Coding of Images. *Archiv fuer Elektronik und Uebertragungstechnik* (1991) vol. 45, No.3, pp.160-167.

[4] BLEJA(M.), DOMAŃSKI(M.). Subband coding of monochrome images using nonseparable recursive filters, *Signal Processing V: Theories and Applications*, ed. L.Torres et al., Elsevier Science Publ.(North Holland), Amsterdam 1990, pp.841-844.

[5] BLEJA(M.), DOMAŃSKI(M.). Subband coding of monochrome images. Annales des Telecommunications (1990) vol.45, nr 9-10 1990.

[6] BLEJA(M.), DOMAŃSKI(M.). Results in subband coding of monochrome images. *Int. Symposium on Signals, Systems, and Electronics*, Erlangen 1989, pp. 314-317.

[7] KRONANDER(T.) A new approach to recursive mirror filters with a special application in subband coding of images. *IEEE Transactions* (1988), **ASSP-36**, pp. 1496-1500.

[8] RAMSTAD(T.) IIR filter bank for subband coding of images. *IEEE Int. Symposium on Signals and Systems* (1988), pp. 827-830.

[9] FETTWEIS(A.). Wave digital filters: theory and practice. *Proc. of the IEEE* (1986), **74**, pp. 270-327.

[10] WESTERLINK(P.H.), BIEMOND(J.), BOEKEE(D.E.). Quantization Error Analysis of image subband filter banks. *ISCAS* 1988, pp. 819-822.

[11]JAYANT(N.S), NOLL(P.). Digital Coding of Waveforms, Prentice Hall, 1984.